SPATIAL AND TEMPORAL COMPONENTS OF ENVIRONMENTAL PREDICTORS IN SPECIES DISTRIBUTION MODELING:

A CASE STUDY USING BELL'S VIREO

(*VIREO BELLII)*

By

CASSONDRA M. WALKER

Bachelor of Science in Wildlife Ecology
Oklahoma State University
Stillwater, Oklahoma
2009

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2018

SPATIAL AND TEMPORAL COMPONENTS OF ENVIRONMENTAL PREDICTORS IN SPECIES DISTRIBUTION MODELING:

A CASE STUDY USING BELL'S VIREO (VIREO BELLII)

Dissertation Approved:

Monica Papeș

Dissertation Adviser

Amy Frazier

Adam Mathews

Mark Gregory

Jennifer Grindstaff

Tim O'Connell

ACKNOWLEDGEMENTS

I am grateful for the generous support of the Welder Wildlife Foundation in funding this project. I would also like to thank the South Central Arc Users Group, the Payne County Audubon Society, both of which provided scholarships that made this research possible.

Further, I would like to mention how appreciative I am to the citizen science community for the huge benefit they provide by the sheer volume of data collection effort.

Name: CASSONDRA WALKER

Date of Degree: DECEMBER, 2018

Title of Study: SPATIAL AND TEMPORAL COMPONENTS OF ENVIRONMENTAL PREDICTORS IN SPECIES DISTRIBUTION MODELING: A CASE STUDY USING BELL’S VIREO

Major Field: INTEGRATIVE BIOLOGY

Abstract:

The largest threats to biodiversity are global climate change and habitat loss, both of which are global concerns due to decreases in species’ populations. Understanding species’ responses to both threats is needed and a common practice used is species distribution modeling (SDM). SDM is a predictive modeling technique, which incorporates environmental conditions associated with species presence locations to derive species-environment relationships that are used to predict geographic locations of species across space and time. As the nature of SDM is both spatial and temporal, the scale of data used affects model performance and predictions. Specifically, the grain and extent of predictor variables influences model performance and hence, model interpretation. I set out to address spatial and temporal scale concerns in SDM using Bell’s Vireo (Vireo bellii), a Neotropical migratory songbird, as a case study. Bell’s Vireo is a species of concern that has shown declining trends across its range, where it inhabits threatened landscapes such as riparian and shrubland-grassland ecotones. Here I describe the use of Bell’s Vireo presence locations to address the role of extent, effects of resampling and grain size, as well as the temporal aspects of environmental predictors in SDM.

First, I compared model performance and potential distributions across three study area extents under eight variable selection techniques and five species’ occurrence data compilations. Overall, I was able to show interactions among these model components, specifically that data quality influenced model performance more than study extent size and variable subset. Second, to investigate the effects of grain size manipulation on SDM, I compared twelve grain sizes resampled using three upscaling techniques. My results showed that model performance in terms AUC was influenced by resampling method, but not grain size, whereas the model performance metric, omission error, was not influenced by resampling technique or grain size, whereas prediction of potentially suitable area was influenced by both resampling and grain size. Last, when investigating temporal effects on SDM performance, I found that more temporally explicit variables, such as seasonal variables, did not necessarily improve model performance although it did increase proportions of suitable area compared to annual variables.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

CHAPTER II

CHAPTER III

CHAPTER IV

LIST OF FIGURES

Figure Page

CHAPTER II

CHAPTER III

CHAPTER IV

CHAPTER I

DISSERTATION OVERVIEW

Decreasing species' populations are a global concern (Ceballos et al. 2017) and the largest threats to biodiversity are global climate change (Rosenzweig et al. 2008, Spooner et al. 2018) and habitat loss (Newbold et al. 2015). Cataloging biodiversity to understand species' responses to both threats is needed for future conservation action (Anderson 2018). Developments in statistical models have provided researchers with tools to evaluate geographic distributions of species, specifically species distribution modeling (SDM). SDM is a predictive modeling technique, which incorporates environmental conditions associated with species presence locations to derive species-environment relationships that are used to predict geographic locations of species across space and time (Elith and Leathwick 2009, Franklin 2010, Peterson et al. 2011). This capability has allowed scientists to model species' predicted responses to land cover/land use change (Zhang et al. 2012), future climate scenarios (Araujo et al. 2004), invasive species (Jiménez-Valverde et al. 2011) and conservation actions (Guisan et al. 2013, Villero et al. 2017).

The applications of SDM to ecological research have increased with some modifications of this modeling technique (Lobo et al. 2010, Brotons 2014, Yackulic and Ginsberg 2016), although the modeling approach is the same. Guisan and Zimmerman (2000) suggest four steps in the SDM framework: conceptual model formulation, statistical model formulation, model calibration, and model evaluation. The conceptual model must consider the ecological basis on which the model will be built (i.e. the affiliation of the species to environmental predictors, geographic extent, and other modeling parameters) (Franklin 2010). Ultimately, the foundation for SDM lies in niche theory (see Peterson et al. 2011 for full discussion) and decisions in the conceptual model should be based on this theory. The statistical model must be an appropriate algorithm for the data associated with the conceptual model (Guisan and Zimmermann 2000). Many approaches exist in SDM, but statistical methods such as generalized linear models (GLMs), or generalized additives models (GAMs), and machine learning (decision trees, random forests, artificial neural networks, and maximum entropy) are the most widely used (Franklin 2010). Model calibration requires the adjustment of the model for parameter estimation and model evaluation uses measures of agreement to assess the model's performance (Franklin 2010). Each step in the SDM framework should be considered carefully as is suggested by a large body of literature (e.g. Heikkinen et al. 2006, Anderson and Gonzalez 2011, Acevedo et al. 2012, Bean et al. 2012, Miller 2012, Aguirre-Gutierrez et al. 2013, Braunisch et al. 2013, Boria et al. 2014, Bucklin et al. 2015, Boria and Blois 2018, Connor et al. 2018).

SDM requires presence data, which are becoming more accessible with the digitization of museum records (Newbold 2010, Anderson 2012), large online repositories (e.g. GBIF Telenius 2011), and citizen-science initiatives (e.g. eBird Sullivan et al. 2009). One of the SDM methods, the Maxent algorithm (Phillips et al. 2006), is based on the maximum entropy principle which estimates a species' geographic distribution under maximum entropy (i.e. a uniform distribution) based on constraints (derived from the presence data), thus creating a model which considers all known aspects of the distribution without additional assumptions (Phillips and Dudik 2008). Maxent uses "features" to constrain the distribution to what is known about the species' use of the environment since responses to environmental conditions can be non-linear (Elith et al. 2011). The distribution upon which constraints are placed is dependent on the random sampling of the background (i.e. area studied). The distribution of environmental conditions captured by the background data is contingent on the geographic region used (i.e. the extent of the study) (VanDerWal et al. 2009), and the area defined by each sample (i.e. grid size) (Mertes and Jetz 2018). Thus, the way the environmental variables are defined (i.e. the scale) across the region of interest has direct implications on model formation.

Scale refers to the spatial and temporal properties of an environmental predictor. Spatially, the total area under consideration is called the "extent", and the area attributed to each measurement or value is the "grain"; Wiens (1989) describes these as the upper and lower spatial bounds since we cannot generalize further than the extent of the study and cannot know data attributes below the grain size. This is particularly important in SDM as these delimiting characteristics of environmental predictors will determine the variation expressed in model formation (i.e. distribution defined by the background samples), and changes in grain and extent will change the variation (between-grain variation and within-grain variation) of environmental conditions (Wiens 1989, Levin 1992). Changes in variation depend on how extent and grain are varied (Anderson 2018): under a constant grain size, larger extent sizes increase between-grain variation as they are likely to include rare conditions, whereas at one extent size larger grains include more spatial heterogeneity (increased within-grain variation) that is no longer detectable, reducing between-grain variation (Wiens 1989). Variation changes are not always predictable as they are related to the spatial structure of the regions' heterogeneity (Palmer 1988, Mertes and Jetz 2018). Further, the variation of environmental predictors is related to the temporal range and frequency of observations (Wiens 1989, Schneider 1994). We know environmental conditions change over different temporal periods (e.g. hourly, daily, seasonally, yearly, etc.), therefore, when conditions are measured and for how long, will affect the perceived variation (Schneider 1994).

While it is important to consider the variation of environmental predictors in terms of model development, this only addresses half of the concern surrounding scale in SDM. The other half of scale's importance is finding the appropriate scale for predictor variables at which species respond (McGill 2010). It has been proposed several times that there exists a series of scales at which species will perceive and respond to environmental conditions (Wiens 1989, Urban 2005, McGill 2010, Mertes and Jetz 2018) and SDM should be conducted at these scales to reflect the ecological processes occurring (Yackulic and Ginsberg 2016). Modeling must occur across a gradient (spatial and temporal) to find scale thresholds where species' responses change (Wiens 1989, Wheatley and Johnson 2009, McGarigal et al. 2016).

I set out to address spatial and temporal scale concerns in SDM using Bell's Vireo (*Vireo bellii*), a Neotropical migratory songbird, as a case study. Bell's Vireo is a species of concern that has shown declining trends across its range (Sauer et al. 2011), where it inhabits threatened landscapes such as riparian and shrubland-grassland ecotones (Noss et al. 1995, Sleeter et al. 2013). As little data exist about the ecological and habitat requirements of this species, the need for further research is clear. Here I describe the use of Bell's Vireo presence locations to address the role of extent, effects of resampling and grain size, as well as the temporal aspects of environmental predictors in SDM.

As mentioned earlier, changes in extent size influence the observed variation in environmental conditions from which background samples are selected, and is an important consideration in SDM conceptual model development (Guisan and Zimmermann 2000, Barve et al. 2011, Acevedo et al. 2012). The extent itself can produce an 'overfit' or overly complex model, when background sampling includes large amounts of heterogeneity, as the model begins to match the presence locations so closely it underestimates predictions (Phillips and Dudik 2008, Barve et al. 2011). On the other hand, low variation in background sampling may create 'underfit' or overly simple models that do not discriminate well and overestimate predictions (Barve et al. 2011). As Maxent model creation is associated with the relationship of presence conditions to background conditions, the quality of occurrence records and the selection of environmental predictors are critical components that influence model output independently (Wisz et al. 2008, Synes and

Osborne 2011, Boria et al. 2014, Bradie and Leung 2017). However, extent, presence locations, and environmental variables are vital components of model creation and fitting and it is likely they may interact in ways that ultimately influence model performance. Since a large portion of Bell's Vireo range encompasses the central United States, it is comprised of ecoregions of varying levels organization (Omernik 1987, McMahon et al. 2001). Therefore, it was possible to delineate non-arbitrary, ecologically relevant extents within a hierarchical framework that demonstrated decreased variation with decreased spatial extent. In addition, Bell's Vireo is a loud and frequent vocalizer, allowing for easy detection and ensuring adequate presence records for modeling. Lastly, habitat selection in Bell's Vireo is poorly understood outside of a preference for nest placement in dense, low vegetation (Parody and Parker 2002), allowing for experimentation in variable subsets. By manipulating spatial extent, presence locations, and environmental variable subsets, I was able to investigate the potential interactions among these components. Specifically, I compared model performance and distributional predictions across three study area extents under eight variable selection techniques and five occurrence data complications.

To investigate scaling in ecological research, multiple scales (extent and grain) should be utilized (Wheatley and Johnson 2009) and while this can be a fairly straightforward process for changing extents, this is not the case for increasing grain size. Multiple techniques have been developed for resampling or upscaling gridded data, particularly in the remote sensing field (Atkinson 1988, Hay et al. 1997, Jensen 2005) and these methodologies use the surrounding grains or cells in different ways to recalculate values for the larger environmental grain. Model performance in response to grain size manipulations is inconclusive at best, where some studies show increased performance (Guisan et al. 2007a, Guisan et al. 2007b, Revermann et al. 2012, Suarez-Seoane et al. 2014), no change in performance (Guisan et al. 2007a), or decreased performance (Guisan et al. 2007a, Guisan et al. 2007b, Seo et al. 2009, Gottschalk et al. 2011, Hanberry 2013, Song et al. 2013). These conflicting results may be due to insufficient grain size comparisons, as several studies compared three or less (Guisan et al. 2007a, Guisan et al. 2007b, Revermann et al. 2012, Suarez-Seoane et al. 2014). Finding appropriate response grains (Mertes and Jetz 2018) of species will require the use of numerous grain sizes. Thus, it is surprising that the effects of resampling have yet to be investigated. While Bell's Vireo has been reported to choose dense vegetation (Parody and Parker 2002) for nest placement, the scale at which environmental conditions shape suitable habitat for Bell's Vireo is unknown, proving an opportunity to test for response scales. Additionally, dense vegetation associated with Bell's Vireo presences is expected to occur irregularly across the spatial extent as shrubby vegetation tends to be dynamic (Myster 2012), providing an opportunity to test the homogenizing effects of resampling. My expectation was that different resampling techniques used to increase grain sizes of environmental variables would likely create altogether different model predictors. By accounting for variation change in the geographical background, I investigated the effects of three resampling methodologies on four environmental variables across 12 grain sizes.

In addition to differential responses of species to spatial scaling of environmental conditions, species are also influenced by the temporal characteristics of environmental predictors (Wiens 1989). Environmental conditions can be measured frequently or infrequently and over different temporal ranges, each providing a distinct snapshot of the environment at a given time period. Since many environmental conditions change over time, the variation may be misrepresented. Seasonality plays a large role in ecological phenomena (Levin 1992) for many taxa (Jacobi and Cary 1996, García 2008, Milakovic et al. 2012, McClure et al. 2013, Fynn et al. 2014, Varner et al. 2014, Johnson et al. 2016), and incorporating seasonality into models can improve SDM performance (Smeraldo et al. 2018). Migrating organisms such as birds show distinct phenology (i.e. seasonally variable ecological requirements) for breeding and wintering seasons (Engler et al. 2014). When categorizing habitat use between these two seasons, researchers have used the terms "niche-followers" and "niche-switchers" (Nakazawa et al. 2004, Engler et al. 2017). The former refers to species utilizing the same environmental conditions between the two seasons and the latter to species displaying distinct differences in between breeding and wintering seasons (Engler et al. 2017). Bell's Vireo exhibits two distinct population clusters, one occurring in the southwest portion of the United States and the other in the central portion of the United States. Each population cluster in comprised of two subspecies, *V.b. arizonae*, and *V.b. pusillus* in the west and *V.b. bellii* and *V.b. medius* in the east. The western population is closely associated with cottonwood-willow riparian areas and the eastern population with shrubby grasslands, although both populations utilize denser areas than the surrounding vegetation (Parody and Parker 2002). A recent genetic analysis from Klicka et al. (2016) indicates that the west-east population divide, occurring near the Arizona/New Mexico border, actually delineates two separate species, each encompassing the two sub-species. Thus, Bell's Vireo provides an ideal opportunity to investigate the influence of seasonal versus annual climatic variables in structuring the distribution of a species, as both populations utilize the same wintering grounds but select different breeding habitats in distinct geographical locations. Capitalizing on the unique breeding habitat associations, my objectives were to explore the use of seasonal and annual environmental predictors in SDM, investigate if spatial extent influenced variable contribution of seasonal and annual variable, and lastly to test the performance of models built with seasonal and annual variables to predict distributions in the future under for an ensemble global circulation model and two greenhouse gas emission scenarios.

REFERENCES

Acevedo, P., A. Jimenez-Valverde, J. M. Lobo, and R. Real. 2012. Delimiting the geographical background in species distribution modelling. Journal of Biogeography 39:1383-1390.

Aguirre-Gutierrez, J., L. G. Carvalheiro, C. Polce, E. E. van Loon, N. Raes, M. Reemer, and J. C. Biesmeijer. 2013. Fit-for-purpose: Species distribution model performance depends on evaluation criteria - Dutch hoverflies as a case study. PLOS One 8:11.

Anderson, C. B. 2018. Biodiversity monitoring, earth observations and the ecology of scale. Ecology Letters 21:1572-1585.

Anderson, R. P. 2012. Harnessing the world's biodiversity data: Promise and peril in ecological niche modeling of species distributions. Pages 66-80. Blavatnik awards for young scientists 2011.

Anderson, R. P., and I. Gonzalez. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecological Modelling 222:2796-2811.

Araujo, M. B., M. Cabeza, W. Thuiller, L. Hannah, and P. H. Williams. 2004. Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. Global Change Biology 10:1618-1626.

Atkinson, P. 1988. The interrelationship between resampling method and information extraction technique. Pages 521-527 in International Geoscience and Remote Sensing Symposium, 'Remote Sensing: Moving Toward the 21st Century'.

Barve, N., V. Barve, A. Jimenez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberon, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling 222:1810-1819.

Bean, W. T., R. Stafford, and J. S. Brashares. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography 35:250-258.

Boria, R. A., and J. L. Blois. 2018. The effect of large sample sizes on ecological niche models: Analysis using a North American rodent, *Peromyscus maniculatus*. Ecological Modelling 386:83-88.

Boria, R. A., L. E. Olson, S. M. Goodman, and R. P. Anderson. 2014. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. Ecological Modelling 275:73-77.

Bradie, J., and B. Leung. 2017. A quantitative synthesis of the importance of variables used in Maxent species distribution models. Journal of Biogeography 44:1344-1361.

Braunisch, V., J. Coppes, R. Arlettaz, R. Suchant, H. Schmid, and K. Bollmann. 2013. Selecting from correlated climate variables: A major source of uncertainty for predicting species distributions under climate change. Ecography 36:971-983.

Brotons, L. 2014. Species distribution models and impact factor growth in environmental journals: Methodological fashion or the attraction of global change science. PLOS One 9:5.

Bucklin, D. N., M. Basille, A. M. Benscoter, L. A. Brandt, F. J. Mazzotti, S. S. Romanach, C. Speroterra, and J. I. Watling. 2015. Comparing species distribution models constructed with different subsets of environmental predictors. Diversity and Distributions 21:23-35.

Ceballos, G., P. R. Ehrlich, and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proceedings of the National Academy of Sciences 114:E6089.

Connor, T., V. Hull, A. Viña, A. Shortridge, Y. Tang, J. Zhang, F. Wang, and J. Liu. 2018. Effects of grain size and niche breadth on species distribution modeling. Ecography 41:1270-1282.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: Ecological explanation and prediction across space and time. Pages 677-697. Annual review of ecology evolution and systematics. Annual Reviews, Palo Alto.

Elith, J., S. J. Phillips, T. Hastie, M. Dudik, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of Maxent for ecologists. Diversity and Distributions 17:43-57.

Engler, J. O., D. Rödder, D. Stiels, and M. I. Förschler. 2014. Suitable, reachable but not colonized: Seasonal niche duality in an endemic mountainous songbird. Journal of Ornithology 155:657-669.

Engler, J. O., D. Stiels, K. Schidelko, D. Strubbe, P. Quillfeldt, and M. Brambilla. 2017. Avian SDMs: Current state, challenges, and opportunities. Journal of Avian Biology 48:1483-1504.

Franklin, J. 2010. Mapping species distributions: Spatial inference and prediction. Cambridge University Press, New York.

Fynn, R. W. S., M. Chase, and A. Röder. 2014. Functional habitat heterogeneity and large herbivore seasonal habitat selection in northern Botswana. South African Journal of Wildlife Research 44:1-15.

García, A. 2008. The use of habitat and time by lizards in a tropical deciduous forest in western Mexico. Studies on Neotropical Fauna and Environment 43:107-115.

Gottschalk, T. K., B. Aue, S. Hotes, and K. Ekschmitt. 2011. Influence of grain size on species–habitat models. Ecological Modelling 222:3403-3412.

Guisan, A., C. H. Graham, J. Elith, F. Huettmann, and N. S. Distri. 2007a. Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions 13:332-340.

Guisan, A., R. Tingley, J. B. Baumgartner, I. Naujokaitis-Lewis, P. R. Sutcliffe, A. I. T. Tulloch, T. J. Regan, L. Brotons, E. McDonald-Madden, C. Mantyka-Pringle, T. G. Martin, J. R. Rhodes, R. Maggini, S. A. Setterfield, J. Elith, M. W. Schwartz, B. A. Wintle, O. Broennimann, M. Austin, S. Ferrier, M. R. Kearney, H. P. Possingham, and Y. M. Buckley. 2013. Predicting species distributions for conservation decisions. Ecology Letters 16:1424-1435.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135:147-186.

Guisan, A., N. E. Zimmermann, J. Elith, C. H. Graham, S. Phillips, and A. T. Peterson. 2007b. What matters for predicting the occurrences of trees: Techniques, data, or species' characteristics? Ecological Monographs 77:615-630.

Hanberry, B. B. 2013. Finer grain size increases effects of error and changes influence of environmental predictors on species distribution models. Ecological Informatics 15:8-13.

Hay, G. J., K. O. Niernann, and D. G. Goodenough. 1997. Spatial thresholds, image-objects, and upscaling: A multiscale evaluation. Remote Sensing of Environment 62:1-19.

Heikkinen, R. K., M. Luoto, and R. Virkkala. 2006. Does seasonal fine-tuning of climatic variables improve the performance of bioclimatic envelope models for migratory birds? Diversity and Distributions 12:502-510.

Jacobi, G. Z., and S. J. Cary. 1996. Winter stoneflies (*Plecoptera*) in seasonal habitats in New Mexico, USA. Journal of the North American Benthological Society 15:690-699.

Jensen, J. R. 2005. Introductory digital image processing. 3rd edition. Pearson Prentice Hall, Upper Saddle River, NJ.

Jiménez-Valverde, A., A. T. Peterson, J. Soberón, J. M. Overton, P. Aragón, and J. M. Lobo. 2011. Use of niche models in invasive species risk assessments. Biological Invasions 13:2785-2797.

Johnson, J. H., R. Abbett, M. A. Chalupnicki, and F. Verdoliva. 2016. Seasonal habitat use of brook trout and juvenile steelhead in a Lake Ontario tributary. Journal of Freshwater Ecology 31:239-249.

Klicka, L. B., B. E. Kus, P. O. Title, and K. J. Burns. 2016. Conservation genomics reveals multiple evolutionary units within Bell's Vireo (*Vireo bellii*). Conservation Genetics 17:455-471.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.

Lobo, J. M., A. Jimenez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. Ecography 33:103-114.

McClure, C. J. W., B. W. Rolek, and G. E. Hill. 2013. Seasonal use of habitat by shrub-breeding birds in a southeastern national forest. The Wilson Journal of Ornithology 125:731-743.

McGarigal, K., H. Y. Wan, K. A. Zeller, B. C. Timm, and S. A. Cushman. 2016. Multi-scale habitat selection modeling: A review and outlook. Landscape Ecology 31:1161-1175.

McGill, B. J. 2010. Matters of scale. Science 328:575-576.

McMahon, G., S. M. Gregonis, S. W. Waltman, J. M. Omernik, T. D. Thorson, J. A. Freeouf, A. H. Rorick, and J. E. Keys. 2001. Developing a spatial framework of common ecological regions for the conterminous United States. Environmental Management 28:293-316.

Mertes, K., and W. Jetz. 2018. Disentangling scale dependencies in species environmental niches and distributions. Ecography 41:1604-1615.

Milakovic, B., K. L. Parker, D. D. Gustine, R. J. Lay, A. B. D. Walker, and M. P. Gillingham. 2012. Seasonal habitat use and selection by grizzly bears in northern British Columbia. The Journal of Wildlife Management 76:170-180.

Miller, J. A. 2012. Species distribution models: Spatial autocorrelation and non-stationarity. Progress in Physical Geography 36:681-692.

Myster, R. W., editor. 2012. Ecotones between forest and grassland. Springer, New York.

Nakazawa, Y., A. T. Peterson, E. Martínez-Meyer, A. G. Navarro-Sigüenza, and F. Moore. 2004. Seasonal niches of Nearctic-Neotropical migratory birds: Implications for the evolution of migration. The Auk 121:610-618.

Newbold, T. 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. Progress in Physical Geography: Earth and Environment 34:3-22.

Newbold, T., L. N. Hudson, S. L. L. Hill, S. Contu, I. Lysenko, R. A. Senior, L. Börger, D. J. Bennett, A. Choimes, B. Collen, J. Day, A. De Palma, S. Díaz, S. Echeverria-Londoño, M. J. Edgar, A. Feldman, M. Garon, M. L. K. Harrison, T. Alhusseini, D. J. Ingram, Y. Itescu, J. Kattge, V. Kemp, L. Kirkpatrick, M. Kleyer, D. L. P. Correia, C. D. Martin, S. Meiri, M. Novosolov, Y. Pan, H. R. P. Phillips, D. W. Purves, A. Robinson, J. Simpson, S. L. Tuck, E. Weiher, H. J. White, R. M. Ewers, G. M. Mace, J. P. W. Scharlemann, and A. Purvis. 2015. Global effects of land use on local terrestrial biodiversity. Nature 520:45.

Noss, R. F., E. T. LaRoe, and M. J. Scott. 1995. Endangered ecosystems of the United States: A preliminary assessment of loss and degradation. United States Department of the Interior, National Biological Service, Washington D.C.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118-125.

Palmer, M. W. 1988. Fractal geometry: A tool for describing spatial patterns of plant communities. Vegetation 75:91-102.

Parody, J. M., and T. H. Parker. 2002. Biogeographic variation in nest placement: A case study with conservation implications. Diversity and Distributions 8:11-19.

Peterson, A. T., J. Soberon, R. G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araujo. 2011. Ecological niches and geographic distributions. Princeton University Press, Princeton.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Phillips, S. J., and M. Dudik. 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31:161-175.

Revermann, R., H. Schmid, N. Zbinden, R. Spaar, and B. Schroder. 2012. Habitat at the mountain tops: How long can rock ptarmigan (*Lagopus muta helvetica*) survive rapids climate change in the Alps? A multi-scale approach. Journal of Ornithology 153:891-905.

Rosenzweig, C., D. Karoly, M. Vicarelli, P. Neofotis, Q. Wu, G. Casassa, A. Menzel, T. L. Root, N. Estrella, B. Seguin, P. Tryjanowski, C. Liu, S. Rawlins, and A. Imeson. 2008. Attributing physical and biological impacts to anthropogenic climate change. Nature 453:353.

Sauer, J. R., J. E. Hines, J. E. Fallon, K. L. Pardieck, D. J. Ziolkowski, and W. A. Link. 2011. The North American breeding bird survey, results and analysis 1966 - 2010. Version 12.07.2011. USGS Patuxent Wildlife Research Center Laurel, MD.

Schneider, D. C. 1994. Quantitative ecology: Spatial and temporal scaling. Academic Press Inc., San Diego, California.

Seo, C., J. H. Thorne, L. Hannah, and W. Thuiller. 2009. Scale effects in species distribution models: Implications for conservation planning under climate change. Biology Letters 5:39-43.

Sleeter, B. M., T. L. Sohl, T. R. Loveland, R. F. Auch, W. Acevedo, M. A. Drummond, K. L. Sayler, and S. V. Stehman. 2013. Land-cover change in the conterminous United States from 1973 to 2000. Global Environmental Change 23:733-748.

Smeraldo, S., M. Di Febbraro, L. Bosso, C. Flaquer, D. Guixé, F. Lisón, A. Meschede, J. Juste, J. Prüger, X. Puig-Montserrat, and D. Russo. 2018. Ignoring seasonal changes in the ecological niche of non-migratory species may lead to biases in potential distribution models: Lessons from bats. Biodiversity and Conservation 27:2425-2441.

Song, W., E. Kim, D. Lee, M. Lee, and S.-W. Jeon. 2013. The sensitivity of species distribution modeling to scale differences. Ecological Modelling 248:113-118.

Spooner, F. E. B., R. G. Pearson, and R. Freeman. 2018. Rapid warming is associated with population decline among terrestrial birds and mammals globally. Global Change Biology 24:4521-4531.

Suarez-Seoane, S., E. Virgos, O. Terroba, X. Pardavila, and J. M. Barea-Azcon. 2014. Scaling of species distribution models across spatial resolutions and extents along a biogeographic gradient. The case of the Iberian mole *Talpa occidentalis*. Ecography 37:279-292.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. Ebird: A citizen-based bird observation network in the biological sciences. Biological Conservation 142:2282-2292.

Synes, N. W., and P. E. Osborne. 2011. Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. Global Ecology and Biogeography 20:904-914.

Telenius, A. 2011. Biodiversity information goes public: GBIF at your service. Nordic Journal of Botany 29:378-381.

Urban, D. L. 2005. Modeling ecological processes across scales. Ecology 86:1996-2006.

VanDerWal, J., L. P. Shoo, C. Graham, and S. E. William. 2009. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? Ecological Modelling 220:589-594.

Varner, D. M., G. R. Hepp, and R. R. Bielefeld. 2014. Movements and seasonal use of habitats by rural and urban female mottled ducks in southeast Florida. The Journal of Wildlife Management 78:840-847.

Villero, D., M. Pla, D. Camps, J. Ruiz-Olmo, and L. Brotons. 2017. Integrating species distribution modelling into decision-making to inform conservation actions. Biodiversity and Conservation 26:251-271.

Wheatley, M., and C. Johnson. 2009. Factors limiting our understanding of ecological scale. Ecological Complexity 6:150-159.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.

Wisz, M. S., R. J. Hijmans, J. Li, A. T. Peterson, C. H. Graham, A. Guisan. 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions 14:763-773.

Yackulic, C. B., and J. R. Ginsberg. 2016. The scaling of geographic ranges: Implications for species distribution models. Landscape Ecology 31:1195-1208.

Zhang, M.-G., Z.-K. Zhou, W.-Y. Chen, J. W. F. Slik, C. H. Cannon, and N. Raes. 2012. Using species distribution modeling to improve conservation and land use planning of Yunnan, China. Biological Conservation 153:257-264.

CHAPTER II

GEOGRAPHIC EXTENT OF ENVIRONMENTAL PREDICTORS, COMBINED WITH SPECIES' PRESENCE RECORDS, INFLUENCES MODEL PERFORMANCE AND ESTIMATED POTENTIAL DISTRIBUTIONS

Abstract

Correlative methods under the ecological niche modeling or species distribution modeling category estimate species' potential distributions from occurrence data and environmental variables and have been used frequently to study biogeography, ecology, and conservation of species. However, choices of occurrence data and environmental variables interact with study area extent to create complexity in species distribution modeling (SDM) evaluations and predictions. To assess the effects of these interactions, I compared model performance and potential distributions across three study area extents under eight variable selection techniques and five species' occurrence data compilations (including spatially clustered data). I used Bell's Vireo (*Vireo bellii*), an IUCN near threatened, migratory songbird, as a case study to highlight the complex effects of extent size, occurrence data, and variable selection on potential distributions obtained with Maxent, a maximum entropy modeling approach. Overall I found that medium extent models had higher performance scores (AUC) but highly clustered occurrence data resulted in lower performance models. At all three extents, occurrence data quality influenced model predictions but was most noticeable at the large extent. I showed that the degree to which

spatially biased occurrence data determine model performance and potential distribution estimates varies with the extent used to train models. Thus, the choice of study area extent should not be made independently from the selection of occurrence locations. Additionally, the selection of variables within models directly influences model output and should be considered carefully with regard to ecological relevance. SDM is commonly used to address management issues such as estimating distributions of rare species, planning species' reintroductions, and designing reserve networks based on maps of species richness and endemism. My study adds to the growing awareness of user choices of data in SDM and advocates careful deliberation in both model construction and interpretation.

Introduction

Species distribution modeling (SDM), also referred to as ecological niche modeling (ENM), is a wide-spread technique that seeks to characterize the conditions suitable for a species' survival based on available species' occurrences and environmental information associated with these occurrences (Guisan and Zimmermann 2000; Araujo and Peterson 2012). In simplest terms, the SDM process has three elements: an ecological model, which is the theory tested; a data model, which encompasses the data used; and a statistical model, or choice of modeling algorithm (Austin 2002). Numerous studies have investigated how these components affect SDM output and interpretation (Liu et al. 2005; Guisan et al. 2007; Moudry and Simova 2012; Guillera-Arroita et al. 2015). Users of SDM must weigh the assumptions of the chosen modeling algorithm with the data available, particularly for presence-background algorithms such as Maxent (Elith et al. 2011; Yackulic et al. 2013), since the effects of violating model assumptions are unclear (Bell and Schlaepfer 2016). Generally, model performance is influenced by user's decisions regarding: occurrence data (Graham et al. 2008; Lozier et al. 2009), environmental variable selection (Johnson and Gillingham 2005), extent size (Barve et al. 2011), and modeling algorithm (Austin 2002; Austin 2007; Aguirre-Gutierrez et al. 2013). Species distribution and ecological niche theory should provide the foundation for user choice (Franklin 2010; Peterson et al. 2011). For example, the selection of occurrence data ultimately determine the sampling of a species' range. Further, choices in environmental predictors, such as which ones to include and the resolution of the variables, can shape the perceived relationships between the species and environment. Addressing implications of user choice on modeling methodologies is an evolving topic and here, rather than treating them separately, I address the combined influences of occurrence data, variable selection, and study area extent on model performance and associated potential distribution estimates.

Occurrence data are often spatially biased as a result of sampling techniques or lack thereof (Dennis and Thomas 2000; Reddy and Davalos 2003; Elith et al. 2006; Schulman et al. 2007). This often translates into bias in sampling environmental conditions and thus, if the modeling protocol does not account for sampling bias, the occurrence records may not provide an accurate model of conditions associated with the distribution of the species (Yackulic et al. 2013). Phillips et al. (2009) suggested introducing bias into the selection of background data that mimics the bias in occurrence data. Often, quality occurrence data means working with small datasets, which can decrease the predictive ability of the modeling algorithm (Stockwell and Peterson 2002). However, small datasets can be effective at uncovering environmental conditions associated with species' distribution if the small number of occurrence locations captures the environmental conditions across the entire range of the species (Hernandez et al. 2006; Pearson et al. 2007).

SDM uses occurrence locations to extract environmental conditions that may shape species' distributions, so the choice of environmental variables must be considered carefully. A wide range of techniques has been used in SDM to select environmental variables (Austin and Van Niel 2011; Synes and Osborne 2011), yet no consensus has been reached. Some studies advocate the use of ecologically meaningful variables (Peterson et al. 2011; Mod et al. 2016) or suggest using variables that may be linked to life history traits (Heikkinen et al. 2006; Tanner et al. 2017), whereas others take a more statistical approach and recommend removing (Dormann et al. 2013; Merow et al. 2013) or minimizing collinearity (Palaoro et al. 2013) using AIC (Muscarella et al. 2014; Warren et al. 2014) or stepwise variable selection approaches (Luoto et al. 2005; Zeng et al. 2016). Synes and Osborne (2011) showed strong positive relationships between number of variables and model performance and negative relationships between the performance measure and number of pixels predicted present. Although selection of variables has been a subject of disagreement, there is support for the notion that the environmental variables should not only inform the models about the ecological processes shaping species' distributions but also allow for transferability of models between regions of interest (Anderson and Raza 2010; Bradie and Leung 2017).

The ecological insight gained from the combination of occurrence data and environmental variables depends on the study area extent since it defines the range of environmental conditions sampled in both occurrence and background locations (Phillips et al. 2006; Elith et al. 2011). Underestimation or model overfit may result from a study area that does not capture well the environmental variation shaping species' distributions (Jimenez-Valverde et al. 2011; Peterson et al. 2011; Sanchez-Fernandez et al. 2011). Adversely, a study area containing too much noise in the sampled environmental conditions will prevent the model from uncovering the ecological processes at work, producing an underfit model which overestimates (Lobo et al. 2010) and can artificially inflate model performance (Jimenez-Valverde et al. 2008). Barve et al. (2011) state that robustness and validity of models are highly dependent upon extent of study region and discuss three ways that extent can influence SDM: 1) impact model training (VanDerWal et al. 2009; Anderson and Raza 2010), 2) impact model testing (Jimenez-Valverde et al. 2008; Lobo et al. 2008), and 3) impact comparisons between models. Both Barve et al. (2011) and Acevedo et al. (2012) offer suggestions on choosing extents for SDM, such as selecting biotic regions, model-based reconstructions, dispersal potential, or trend surface analyses.

It is likely that choices of occurrence data and variable subsets interact with the extent of the study area to create complexity in SDM evaluations and predictions. To our knowledge, the three factors (occurrence dataset, variable selection, and extent of study area) have not been analyzed in conjunction. Thus, our objective was to investigate potential impacts of extent size through the experimental manipulation of occurrence data and variable selection approaches. Specifically, I compared model performance and distributional predictions across three study area extents under eight variable selection techniques and five occurrence data compilations. I selected an IUCN near threatened, migratory songbird (*Vireo bellii*, Bell's Vireo) as a case study to highlight the complex effects of extent size, occurrence data, and variable selection on models and distributional estimates obtained with Maxent, a commonly used SDM algorithm. I selected Bell's Vireo for this study for several reasons: 1) it is a declining species of concern (IUCN 2013); 2) very little published research is available about its environmental requirements; and 3) the high vocalization rates make the species easily detectable in the field, thus minimizing identification and detection errors in the presence dataset.

Bell's Vireo *(Vireo bellii*) is a migratory songbird often found in grassland and shrubland ecotones and other low shrubby vegetation throughout the central United States (Lebbin et al. 2010). Bell's Vireo taxonomy currently recognizes four subspecies (Ridgway and Friedmann 1919): *V.b. bellii*, *V.b. medius*, *V.b. arizonae*, and *V.b. pusillus*. A recent phylogenetic analysis by Klicka et al. (2016), found that the Bell's Vireo species is likely two species, delineated by an east/west divergence near the New Mexico and Arizona border. The two populations are found in unique habitats, with the western population associated with willow-dominated riparian vegetation (Franzreb 1987) and the eastern population showing more generalist habitat choices associated with lower vegetation. Further, Parody and Parker (2002) found variation in net site selection across the range of Bell's Vireo. Thus, I excluded the western population from this analysis.

Although the Great Plains, which encompasses a large portion of the eastern population of Bell's Vireo, has not undergone significant land cover reduction in the past, the overall productivity and structure have dramatically changed (Noss et al. 1995; Sleeter et al. 2013). This is a likely factor in the decline of Bell's Vireo, but demographic information specific to the central United States is scant (Budnik et al. 2000). Most Bell's Vireo literature encompasses their mention in avian assemblages (Remsen et al. 1996; Thogmartin et al. 2009) or anecdotal reports on nesting and behavior (Pitelka and Koestner 1942; Mumford 1952; Dunkin and Guthery 2010), or investigations related to Brown-headed Cowbird (*Molothrus ater*) parasitism (Parker 1999; Budnik et al. 2001). There is a clear need to better understand environmental determinants of Bell's Vireo distribution across the eastern portion of the species range. Thus, I use a single, data limited species as a case study to call attention to possible interactions of extent, variable selection, and occurrence data in SDM. The implications of these interactions need to be considered when SDM and resulting potential distribution estimates are used in conservation and management of species.

Methods

Study Area

I used three study area extents (small, medium, large) in the modeling process, which comprised the distributional range of Bell's Vireo across the Great Plains of the United States. I implemented a nested design of the study extents, where the smallest extent was completely contained by the medium and large extents (Figure 1). To delineate the largest spatial extent I selected ecoregions, as defined by US Environmental Protection Agency (Omernik 1987; McMahon et al. 2001), that encompassed all Bell's Vireo locations reported by citizen scientists in the eBird database (eBird.org) for all years (eBird 2015), specifically, the Great Plains and western portions of the Eastern Temperate Forests (EPA Level I ecoregions). I used a single EPA

Level II ecoregion, the South Central Semiarid Prairies, which was contained within the Great Plains, to define the medium spatial extent. Lastly, the smallest spatial extent was comprised of the Central Great Plains (EPA Level III ecoregion). I use ecoregion boundaries because areas within the boundaries share similar biotic and abiotic capabilities, with ecoregions at lower levels (i.e. II, III, IV) having more homogeneity (McMahon et al. 2001). This approach has been used in other SDM investigations and provides both a straightforward and an effective method for choosing extents (Soberón 2010; Barve et al. 2011).

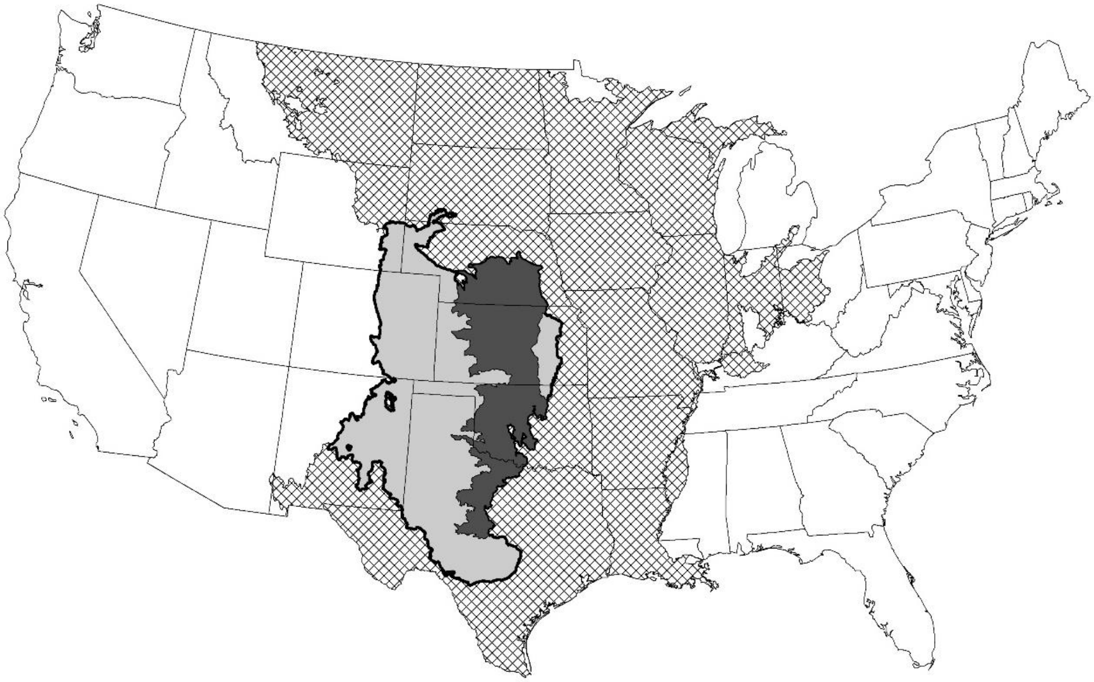


Figure 1. Nested study area extents used in the study. The study area encompasses the range of Bell's Vireo (*Vireo bellii*). Three extents were defined based on United States Environmental Protection Agency (EPA) classification of ecoregions. The smallest extent (Central Great Plains, EPA Level III ecoregion) covered ~ 275,000 km2 (colored dark grey), the medium extent (South Central Semiarid Prairies, EPA Level II ecoregion) was three times larger with an area of 863,000 km2 (shown in light grey), and the largest extent (Great Plains and western portions of the Eastern Temperate Forests, EPA Level I ecoregions) covered an area of 3,520,000 km2 (crosshatched).

Data Acquisition and Processing

I acquired three types of environmental data layers, climatic, vegetative, and topographic, for use in the modeling process for the Bell's Vireo breeding season (May – August), during 2004 – 2014. Four climatic variables were downloaded from PRISM (PRISM 2004) at a spatial resolution of 4 km: Mean Temperature, Maximum Temperature, Minimum Temperature, and Precipitation. Two vegetation variables, the Normalized Difference Vegetation Index (NDVI) and the Leaf Area Index (LAI) from NASA's MODIS sensor, with a spatial resolution of 1 km, were downloaded using the USGS's Land Processes Distributed Active Archive Center (LP DACC) (NASA LP DAAC 2004-2014). Lastly, six topography variables with a spatial resolution of 1 km were obtained from the USGS National Hydrography Dataset (USGS 2014): elevation, aspect, slope, flow accumulation, flow direction, and wetness index.

For both climatic and vegetation variables I calculated the mean, minimum, and maximum values across May-August for ten years (2004-2014) and created new variables representing the mean, minimum, and maximum for temperature, precipitation, NDVI, and LAI. In total, the modeling experiments were based on 18 variables: six climate (minimum, mean, and max of temperature and precipitation), six vegetation (minimum, mean, and max of NDVI and LAI), and six topography (elevation, aspect, slope, flow accumulation, flow direction, and wetness index). All environmental variables were re-projected to WGS84 and resampled to 4 km, using nearest neighbor, then clipped to the three study area extents.

Bell's Vireo locations were obtained from eBird (eBird 2015), a citizen science program developed by the Cornell Lab of Ornithology (Sullivan et al. 2014), that allows individuals to submit bird observations to an online database for which researchers can request access. Individual eBird checklist locations are associated either with a stationary point of observation or a distance traveled for which all bird sightings are reported. I restricted Bell's Vireos locations to those that were stationary observations or traveling observations that did not exceed 1 km for my sampling timeframe (2004 – 2014). By limiting the distance on the traveling checklist locations, I increased the likelihood that each observation location could be associated with a 4 km pixel, representing the spatial resolution of our environmental variables.

Model Development

To fully investigate the combined influence of study area extent, variable choice, and occurrence data bias on model calibration performance and prediction, I ran several experiments where I manipulated both the occurrence data and the variable selection approaches for the three extents. I produced four subsets of the original presence dataset to simulate two common occurrence data biases, specifically, small sample sizes and spatially biased locations. To simulate smaller sample sizes, I reduced the number of occurrence locations using a random selection of both 50% and 25% of the original data at each of the three extents. To create our spatially biased locations (clustered data), I split the data latitudinally such that 50% the occurrences were in the northern half of the extent and the remaining 50% in the southern portion (Figure 2), keeping the records from the northern half. Using this technique allowed me to simulate a restricted dataset, which did not sample the environment across the accessible range of the study species. Thus, our original occurrence dataset and the four manipulated datasets were: 1) all presences after removing duplicate and erroneous records (100P); 2) a random selection of 50% of the original presences (50R); 3) a random selection of 50% from the 50R data, representing 25% of the original data (25R); 4) a selection of the northern 50% of the 100P data (50C); and 5) a random selection of 50% from the 50C data, representing 25% of the original data (25C). These subsets of occurrence data simulate datasets using all data available (100P), smaller datasets (50R and 25R) and spatially biased sampling occurrence data (50C and 25C). Occurrence locations for species of interest are often obtained from museum collections and citizen science programs such as eBird. However, these are often opportunistically collected, containing bias and often violating the assumptions of presence-background modeling (Phillips et al. 2009; Yackulic et al. 2013), such as fully saturated suitable habitat and constant detection probability (Wiens et al. 2009). Our goal was to highlight the modeling implications of spatial bias at differing extents. Since I specifically constructed spatial bias in 50C and 25C subsets, I assumed they contained more spatial bias than our 100P occurrence subset.

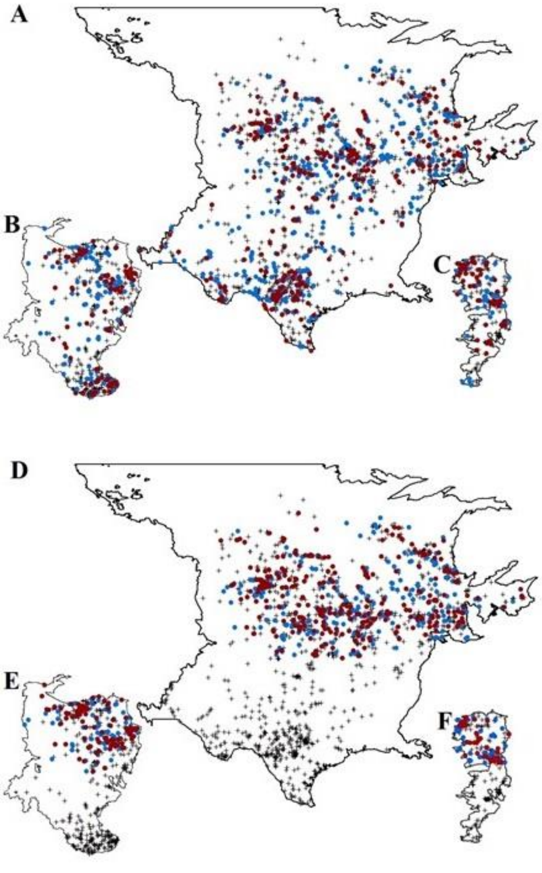


Figure 2. Map of random (A – C) and clustered (D – F) occurrence data for each extent, large (A and D), medium (B and E) and small (C and F). Black crosses indicate testing occurrences, blue circles represent 50% data subsets for random (50R) and clustered (50C) training occurrences where red circles show 25% data subsets for training data that was random (25R) or clustered (25C).

Each occurrence subset was used with eight variable selection approaches, for each of the three extents. I selected variables for inclusion into the models following eight criteria: 1) all 18 variables (FULL), ignoring collinearity and ecological relevance of the variables; 2) a technique to address collinearity (NOCOL), where a subset of variables with Pearson's $|r| < 0.7$ was used (Dormann et al. 2013; Dormann et al. 2008); 3) a climate only approach using only temperature and precipitation variables (CLM); 4) a vegetation only approach using only NDVI and LAI variables (VEG); 5) a topographic only approach using elevation, aspect, slope, flow accumulation, flow direction, and wetness index (TOPO); 6) maximum climate and vegetation values (MAX) across the sampling timeframe (2004-2014) and also elevation and slope; 7) mean climate and vegetation values (MEAN) and also elevation and slope; and finally 8) the minimum climate and vegetation values (MIN) and the same two topographic.

To generate the models I used Maximum Entropy (Maxent), a presence-background algorithm that defines environmental constraints for species based on available presence data, contrasted to randomly selected samples from the background environment of the study area (Phillips et al. 2006). Higher probabilities of suitability are given to locations within the study region, which have environmental conditions more similar to the environmental conditions of known presences (Phillips et al. 2006). I randomly divided our presence locations from each of the five occurrence subsets into a training and testing dataset, each containing 50% of the available data. I also allowed 50% of the background points to be sampled during the modeling process at each of the three extents (small, medium, large). To convert Maxent's continuous suitability predictions to spatial binary suitable/unsuitable predictions, I applied a 10% minimum training presence threshold. This threshold leaves out 10% of the training points, and is commonly used for datasets where the error is unknown (but should still be accounted for) because this threshold is less sensitive to outliers or incorrectly located occurrences (Peterson 2006; Peterson et al. 2011).

As extent choice directly affects the sampled background, I quantified the variation of each variable at each extent. I used two first order variation metrics, range and coefficient of variation to describe the sampled variation at each extent. Range is the maximum value minus the minimum value and the coefficient of variation is mean divided by the standard deviation. I compared the change range and CV to changes in model performance metrics and model predictions

Model Performance Metrics

I compared and evaluated models using two performance criteria, the area under the curve of receiver operating characteristic plot (ROC AUC) and testing omission error. The ROC AUC is a common method of model comparison where a plot is constructed with sensitivity (fraction of presences predicted present) and 1-specificity (fraction of absences predicted absent), modified for SDM to proportion of area predicted present (Phillips et al. 2006; Phillips et al. 2009). The area under this curve with a value of 1 indicates perfect discrimination of presence and absence, whereas models with AUC > 0.7 are considered reliable (Swets 1988; Fielding and Bell 1997). Some authors have shown that AUC values may overstate the performance of models (Austin 2007; Yackulic et al. 2013); such as over-fit models (Lobo et al. 2008; Jimenez-Valverde 2012) and others showed that the AUC estimates are sensitive to many model parameters (Peterson et al. 2008; Hanczar et al. 2010). In addition to evaluating individual model AUC values, I also calculated the average AUC of models at each extent, occurrence treatment, and variable selection. Since AUC incorporates the proportion of the extent predicted present, it is influenced by extent size and there is general agreement that comparing AUC values between different extents is an erroneous comparison due to the differences in extent background (Lobo et al. 2008; Barve et al. 2011; Acevedo et al. 2012). Here, I compare AUC across extents to highlight that extent, occurrences, and environmental variables combined influence model performance metrics such as AUC. Further, Lobo et al (2008) states that multiple methods of model evaluation should be used to assess and compare models since any given single indicator can have some level of bias associated with it. Despite this, most modeling studies continue to use AUC as the single measure of model evaluation. Thus, here I also used omission error, which is the percentage of known locations not classified as presences by the model. Typically, lower omission error indicates higher model performance (Mouton et al. 2010). Since I used the 10% minimum training presence threshold, I selected 20% as the testing OE threshold for “acceptable” models as this allowed for maximum 10% additional error.

Model Predictions of Potential Distributions

Besides evaluating the Maxent models, I also compared geographic predictions of models using three criteria, the proportion of study area predicted suitable, the map kappa statistic, and variable contribution. Additionally, since the number of variables for the models differed, I investigated any correlations between the evaluation criteria and the number of variables. To compare Maxent thresholded predictions, I used the proportion of the study area that was predicted suitable for each model. I used this output to evaluate models because models that are ‘over-fit’ are generally unable to predict suitable areas outside the training dataset and usually have lower proportions of predicted suitable area, whereas models that do not discriminate well between presence and background tend to have higher proportions of predicted suitable area (Peterson et al. 2011). For additional model prediction comparison I used the map kappa statistic, which assesses the accuracy of a prediction compared to observed phenomena (Pontius 2000), with the implementation proposed by Hagen (2002) that takes into account both the quantity and location of prediction pixels. In this study, I compared predictions from each unique occurrence subset (50R, 25R, 50C, 25C) and variable subset (FULL, NOCOL, CLIM, VEG, TOPO, MAX, MEAN, MIN) to the prediction obtained using 100P occurrence data for each variable selection approach at each extent. This allowed us to quantify prediction similarity between models.

Contribution of Variables to Model Accuracy Gain

Since varying the extent of the study area could affect variable contribution to model accuracy gain, I compared variable importance ranking across models. Maxent ranks environmental variables by calculating the contribution of each variable to overall model accuracy gain as proportion of contributions of all variables (Phillips et al. 2006). Lastly, I checked for correlations between the number of variables and all evaluation metrics (AUC, omission error, kappa, predicted area, and percent contribution of the highest performing variable).

Results

Variation of Environmental Variables by Extent

Of the 18 variables used in models, 15 showed some decrease in range (three showed no change) from the large to medium extents and seven showed decreases at the smallest extent, with the remaining 11 showing no change from the medium extent (Table 1). Elevational variables showed both the smallest and largest amount of proportional change in variable range values between the large and medium extents. Comparing large to medium extent variable ranges, vegetation variables showed the least proportional change and temperature showed the most change. Using coefficient of variations as a variation metric, 12 variables showed decreased variability at the medium extent whereas six showed increases. All 18 variables showed decreased variation, lower coefficient of variation values, when compared to the medium extent. I found less variability at the smallest extent, and in general, more variation occurred at the largest extent, with some exceptions.

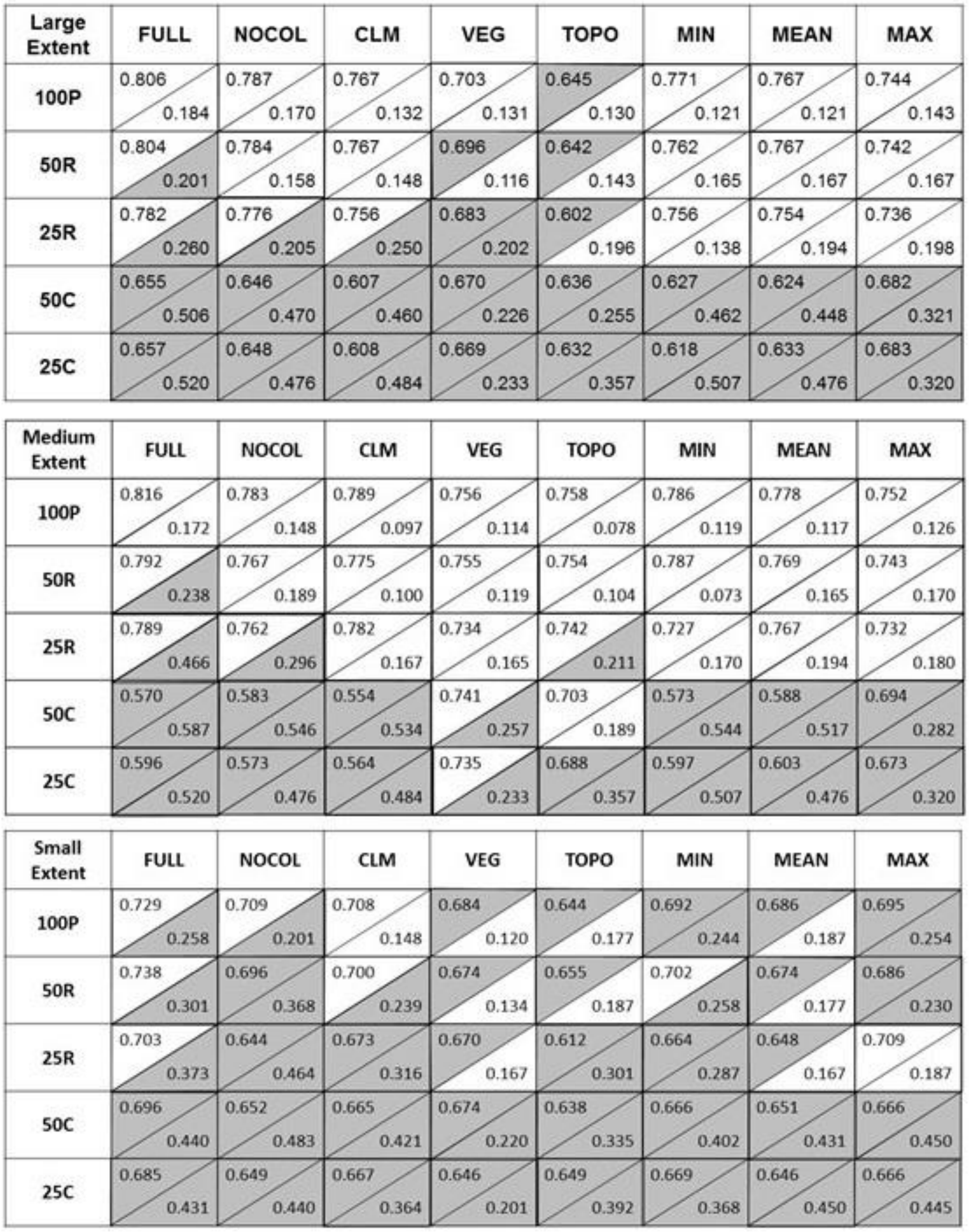
Table 1. Calculated range and coefficient of variation (CV) for the 18 variables at each of the three extents (Broad, Medium, and Small). Bold numbers show variables that have increased variation (measured by either range or CV) from the next largest extent; up arrows also indicate this trend. Dashes denote variables that have the same variation (range or CV) from the next largest extent. Numbers with no annotation correspond to decreasing variation.

	Broad Extent		Medium Extent				Small Extent			
	Range	CV	Range		CV		Range		CV	
max temp	20.8	0.09	13.9		0.06		8.6		0.05	
max prcp	742.7	0.31	578		**0.32**	↑	377.5		0.22	
max ndvi	1.299	0.20	0.914		0.18		0.914	–	0.12	
max lai	25.4	1.02	25.2		**1.40**	↑	25		0.92	
mean temp	21.4	0.17	13.1		0.11		7.8		0.09	
mean prcp	180.5	0.29	110.8		0.28		84.2		0.16	
mean ndvi	1.188	0.31	0.903		0.27		0.903	–	0.18	
mean lai	25.37	1.82	25.22		2.52		**25.31**	↑	1.73	
min temp	26	0.57	17.4		0.41		10.8		0.26	
min prcp	59.3	0.87	40.9		**1.15**	↑	40.9	–	0.85	
min ndvi	1.01	0.52	0.705		0.33		0.703		0.25	
min lai	25.39	4.93	25.39	–	**5.11**	↑	25.39	–	4.55	
aspect	35990	0.65	35976		**0.65**	↑	35968		**0.66**	↑
slope	2184	1.35	981		1.01		383		0.72	
flow	3196055	27.22	195821		11.76		195821	–	9.72	
flow dir	127	1.42	127	–	**1.45**	↑	127	–	1.41	
topo	2205	0.36	1822		0.35		1754		0.34	
elev	3248	0.77	2322		0.49		780		0.29	

Model Performance Metrics

Although overall model performance measured by AUC was relatively low across all extents (AUC range 0.554 – 0.853), I found that model performance measured by AUC varied with extent size (Figure 3). Medium extent models showed the most range in AUC values (0.554 – 0.816), followed by the large extent (0.602 – 0.806) then the smallest extent (0.612 – 0.738). Further, models at the medium extent showed the highest AUC values, but only slightly. Mean AUC for all medium extent models was 0.710, followed by large extent models with a mean of 0.702, and finally the small extent (mean = 0.674). Generally, models with AUC values greater than 0.700 are considered adequate models (Swets 1988; Fielding and Bell 1997). Medium extent models with non-clustered occurrence data (100P, 50R and 25R) had AUCs greater than 0.700.

Clustering occurrence data (50C and 25C) reduced model performance at the medium extent. Models from the smallest extent had the lowest performance scores, with a majority of models based on non-clustered and clustered data exhibiting AUCs less than 0.700. Large extent models showed performance between medium and small extents, with some models' AUC values greater than the 0.700 threshold and some falling below. At all three extents, clustered datasets produced models that performed worse than models with all presences (100P) and randomly selected presences (50R and 25R). The most variation in model performance between non-clustered and clustered data occurred at the medium extent and the least variation at the small extent. Additionally, models based on 100P and 50R showed the most similarity in AUC for the eight sets of variables. I found that models using FULL variable set consistently produced higher AUC values across large, medium, and small extents. For non-clustered data, models based on TOPO variable set exhibited the lowest AUC values at the large and small extent, but MAX models appeared to perform worse than the others at the medium extent. The AUC values were highly influenced by extent size, data clustering, and variable selection techniques (Figure 3). Mean AUC values for models averaged across occurrence data types regardless of extent and variables selection showed that 100P and 50R models performed better (mean AUC 0.739 and 0.734, respectively), with 25R showing moderate performance (0.716). The clustered datasets produced models with poorer performance (mean AUCs of 0.644 for 50C and 0.643 for 25C).



Large Extent	FULL	NOCOL	CLM	VEG	TOPO	MIN	MEAN	MAX
100P	0.806 0.184	0.787 0.170	0.767 0.132	0.703 0.131	0.645 0.130	0.771 0.121	0.767 0.121	0.744 0.143
50R	0.804 0.201	0.784 0.158	0.767 0.148	0.696 0.116	0.642 0.143	0.762 0.165	0.767 0.167	0.742 0.167
25R	0.782 0.260	0.776 0.205	0.756 0.250	0.683 0.202	0.602 0.196	0.756 0.138	0.754 0.194	0.736 0.198
50C	0.655 0.506	0.646 0.470	0.607 0.460	0.670 0.226	0.636 0.255	0.627 0.462	0.624 0.448	0.682 0.321
25C	0.657 0.520	0.648 0.476	0.608 0.484	0.669 0.233	0.632 0.357	0.618 0.507	0.633 0.476	0.683 0.320

Medium Extent	FULL	NOCOL	CLM	VEG	TOPO	MIN	MEAN	MAX
100P	0.816 0.172	0.783 0.148	0.789 0.097	0.756 0.114	0.758 0.078	0.786 0.119	0.778 0.117	0.752 0.126
50R	0.792 0.238	0.767 0.189	0.775 0.100	0.755 0.119	0.754 0.104	0.787 0.073	0.769 0.165	0.743 0.170
25R	0.789 0.466	0.762 0.296	0.782 0.167	0.734 0.165	0.742 0.211	0.727 0.170	0.767 0.194	0.732 0.180
50C	0.570 0.587	0.583 0.546	0.554 0.534	0.741 0.257	0.703 0.189	0.573 0.544	0.588 0.517	0.694 0.282
25C	0.596 0.520	0.573 0.476	0.564 0.484	0.735 0.233	0.688 0.357	0.597 0.507	0.603 0.476	0.673 0.320

Small Extent	FULL	NOCOL	CLM	VEG	TOPO	MIN	MEAN	MAX
100P	0.729 0.258	0.709 0.201	0.708 0.148	0.684 0.120	0.644 0.177	0.692 0.244	0.686 0.187	0.695 0.254
50R	0.738 0.301	0.696 0.368	0.700 0.239	0.674 0.134	0.655 0.187	0.702 0.258	0.674 0.177	0.686 0.230
25R	0.703 0.373	0.644 0.464	0.673 0.316	0.670 0.167	0.612 0.301	0.664 0.287	0.648 0.167	0.709 0.187
50C	0.696 0.440	0.652 0.483	0.665 0.421	0.674 0.220	0.638 0.335	0.666 0.402	0.651 0.431	0.666 0.450
25C	0.685 0.431	0.649 0.440	0.667 0.364	0.646 0.201	0.649 0.392	0.669 0.368	0.646 0.450	0.666 0.445

Figure 3. Performance metrics for each model and occurrence dataset combination across the three extents. For each cell, upper values represent area under the curve (AUC) and lower omission error values. Metrics in white represent acceptable performance (AUC ≥ 0.70 and omission error ≤ 0.10), whereas grey values indicate poor performance. The matrix of values is designed to provide a framework to help users better visualize model performance issues associated with variable subsets, occurrence data quality, and extent size. For example, boxes completely white indicate higher performing models (higher AUC and lower omission error).

Extent size did not affect model performance as evaluated by omission error to the same degree as AUC, indicating that perhaps other factors such as occurrence data and model variables are more influential than extent size (Figure 3). Since I used the 10% training omission error threshold for converting Maxent continuous probability of suitability predictions into binary suitable/not suitable maps, testing omission error rates greater than 10% indicate error in the model, and I considered acceptable models with less than 20% omission error. I found the highest variation in omission error for medium extent models, showing both the highest and lowest omission error rates compared to the models at the other two extents. Models at the large extent showed consistently lower omission error. Models based on non-clustered presence data overwhelmingly showed lower omission error rates than clustered data. In particular, the clustered occurrences at the large extent showed nearly identical trends and the testing omission error ranged from 10% to 40% higher than identical models of non-clustered data. The testing omission error of FULL and NOCOL models was consistently high for all extents and occurrence data types, although these models did not necessarily show the single highest omission error. Of all models constructed, the medium extent FULL model based on clustered data exhibited the two largest omission error values, 67% (25C) and 58.7% (50C). Omission error was most affected by data clustering and variable selection technique, rather than extent size (Figure 3).

Model Predictions of Potential Distributions

The percentage of predicted suitable area varied across extents, data occurrence, and variable selection techniques (Figure 4 A, B, and C). Medium extent models showed less variation in the amount of predicted suitable area, but they did not necessarily show the lowest values when compared to the large or small extent models. The three extents showed similar ranges of predicted suitable area, although the predicted suitable area of medium extent models was slightly lower than the values for large and small extent models. Non-clustered data models predicted higher percentages of predicted suitable area compared to clustered data models; the difference is striking at the large and medium extents (Figure 5). Across model types, VEG, TOPO, and MAX models in most cases appeared to predict higher percentages of predicted suitable area at all three extents, whereas FULL and NOCOL models generally predicted less area as suitable for the three extents.

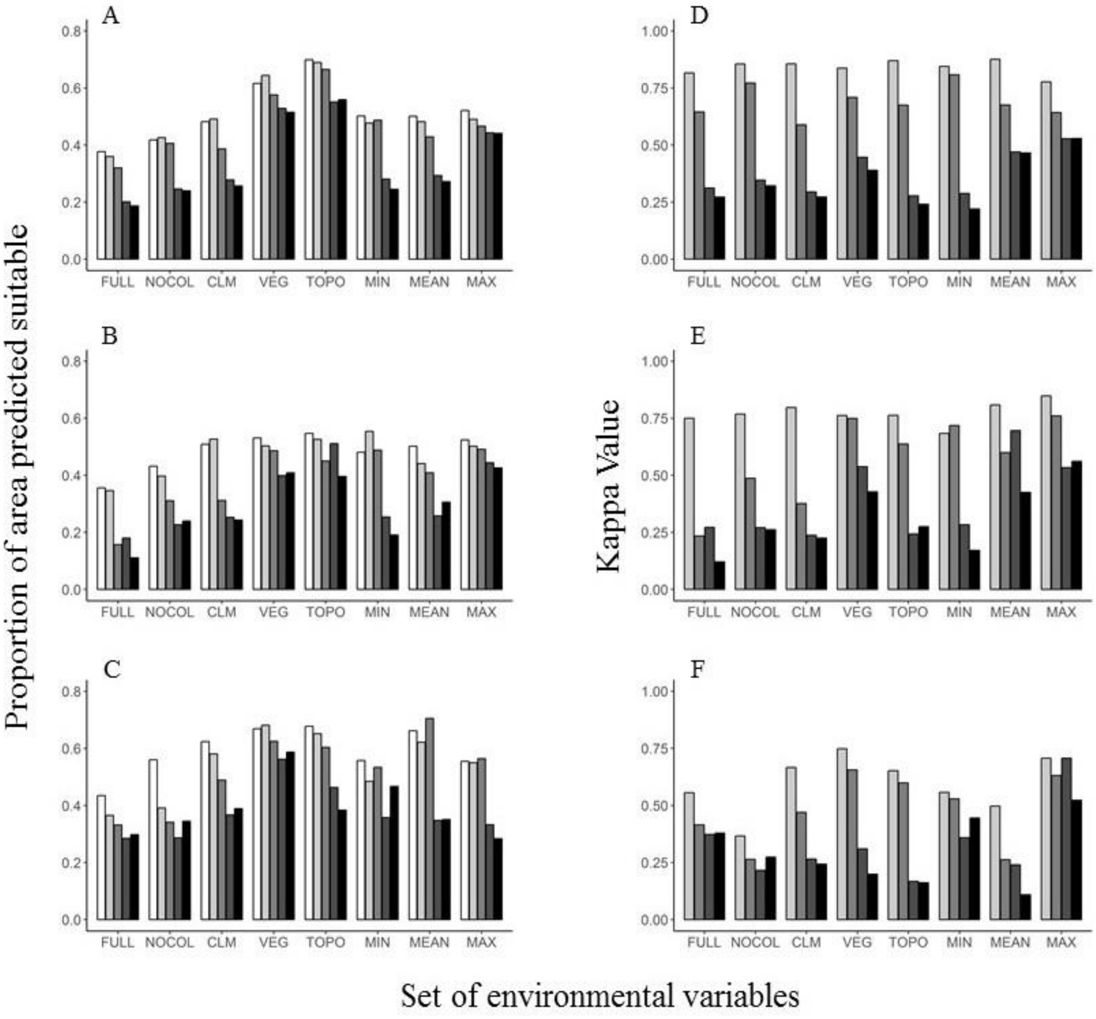


Figure 4. Evaluation metrics (predicted area and the kappa) for large (A & D) medium (B & E) and small (C & F) extents. The five occurrence datasets, 100P (white), 50R (light grey), 25R (medium grey), 50C (dark grey) and 25C (black) are shown for the predicted area metric, while only non-100P (non-white) datasets are shown for the kappa metric as the kappa measures the similarity in predictions between each of the four datasets to the 100P data set.

The map kappa statistic showed that occurrence data quality highly influences similarity of model predictions to the models obtained with all occurrence data (100P). Clustered occurrences showed lower kappa values than randomly selected occurrence data (Fig. 4 D, E, and F). This effect is apparent at all three extents, but is most noticeable at the large extent. In fact, at the large extent, 50R models show high similarity in predictions, 80% or more, and 25R models, although more variable in map kappa statistic, show 60% and higher similarity. Differences across variable selection models were highly dependent on extent; for example, MEAN models showed higher prediction similarity at the large extent, but very little similarity at the smallest extent.

While both performance and prediction metrics provide insightful information about overall model quality, the combination of these together can guide a user's confidence in the model. High AUC values generally indicate models with high discrimination, whereas models unable to predict conditions associated with presence data will exhibit high AUC and low predicted suitable area and high omission error. Our models did not appear to show these characteristics. However, models with high predicted suitable area and very low omission error indicate models that are overly broad (much of the study extent is predicted as suitable in order to correctly predict testing occurrences), although they may have acceptable AUC performance. At the large and medium extent, TOPO models demonstrate this, as they have moderate AUC values ranging from 0.602 – 0.758, omission error ranging from 10 – 35%, and predicted suitable area 39 – 68%. In some cases, these models predict nearly 70% of the extent as suitable.

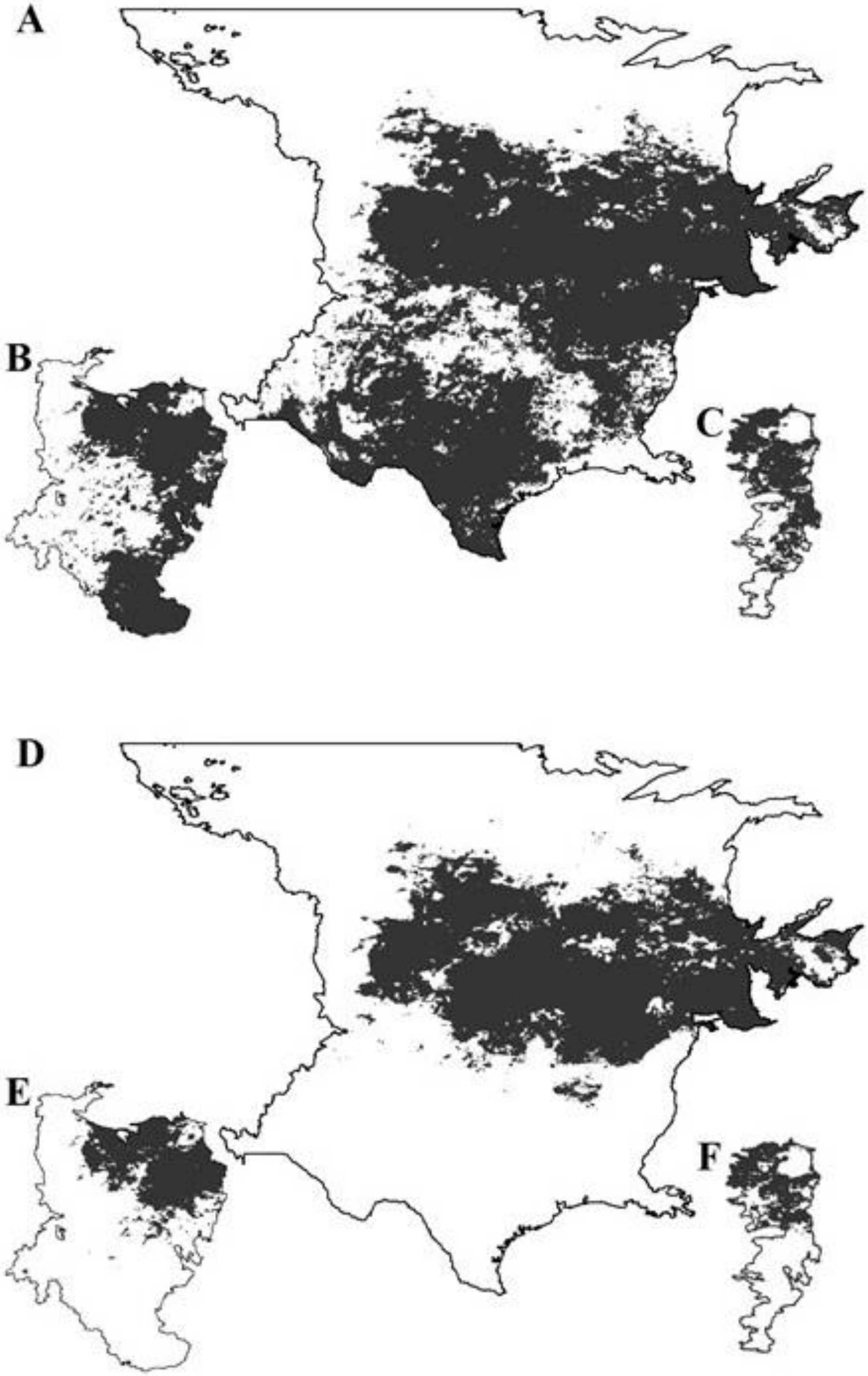


Figure 5. Potential suitability for Bell’s Vireo predicted by CLM models for 50R (A – C) and 50 C (D – F) occurrence data sets for the three study area extents.

Contribution of Variables to Model Accuracy Gain

Overall, highest variable contribution did not differ greatly across the three extents (see Table S.2. in the Supplementary Material). Variation in contribution ranged from 16.9% to 89.5% and was more apparent within extents based on variable selection techniques and not based on clustering of occurrence data. The TOPO model generally always produced the variable with the highest contribution; in fact, elevation was the top variable in every TOPO model for all extents and data types. Conversely, the top variable for the VEG models was not consistent and regularly produced low contributing variables for both non-clustered and clustered data at all three extents.

At the largest extent, a temperature variable was the top contributing variable for every model (except VEG and TOPO, which did not include any temperature variable). Min, mean, and max temperature showed the most decrease in variation with extent size reduction and were not a top contributing variable at the medium extent but did regularly show up as a high contributing variable at the smallest extent. Precipitation and both NDVI and LAI contributed most to models at the medium extent (except TOPO), and showed less decrease in variation across extents.

When selecting variables for the eight model types, the number of variables ranged from 6 to 18, with most models (CLM, VEG, TOPO, MIN, MEAN, and MAX) having six variables, the FULL model comprising of 18, and NOCOL having between 11 and 13 variables, depending on extent of study area. Correlations between the number of variables and the evaluation metrics (AUC, omission error, predicted suitable area, kappa values, and first variable contribution) showed several strong relationships (see Appendix 3). In particular, predicted suitable area and first variable contribution exhibited strong negative correlations with the number of variables in the model. In contrast, omission error and in some instances map kappa statistic showed a positive relationship; AUC showed both positive and negative associations.

Discussion

Our results revealed that model performance can indeed depend on the combination of study area extent, occurrence data, and environmental variable choices. This is cause for concern as SDM is used to address many important topics such as: climate change (Hijmans and Graham 2006; Forester et al. 2013), invasive species (Jones et al. 2010; Mainali et al. 2015), biodiversity assessments (Rodriguez-Castaneda et al. 2012), and conservation management (Johnson and Gillingham 2005; Mizsei et al. 2016). These applications may use a variety of extents, ranging from local to continental or global, and the conclusions derived from the evaluation and assessment of models should be considered carefully. With the wide use of SDM, there exists a need to outline best modeling practice (Jimenez-Valverde et al. 2008; Jarnevich et al. 2015) and standardize model calibration and evaluation (Lobo et al. 2008). Several papers discuss obstacles in the modeling process associated with occurrence data (Lozier et al. 2009), environmental variable selection (Johnson and Gillingham 2005), and extent size (Barve et al. 2011). I integrated these three main factors affecting model performance into a single study, of a widely used algorithm (Maxent) using a data deficient species of conservation concern.

Influence of Choice of Data on Model Performance

Biased occurrence data, specifically data that are clustered, do not adequately sample the environment within the whole distribution of the species (Phillips et al. 2009), thus they produced inaccurate models. Models calibrated with clustered occurrence data exhibited lower AUC values than models using other occurrence datasets. This was especially true when models contained variables that were not ecologically relevant to *V. bellii*. However, extent size adds another dimension of complexity in that at the large and small extents clustered models performed worst but the same models were high performing at the medium extent. The varying performance of models using different occurrence datasets across extents may be attributed to lower variation in environmental conditions extents. This low environmental variation may be mitigating the effects of clustered data and allowing all occurrence datasets to sample the simpler environment adequately. Alternatively, certain extents may be appropriate for sampling environmental conditions that operate at particular spatial scales shaping distributions (Luoto et al. 2007). Additionally, I showed that the number of variables influenced AUC as the full models generally had higher AUC values, although the correlation between the number of variables and AUC did not show a clear relationship. Evidently, not all models are useful, such as our topography model, yet if I use the AUC >0.7 rule of thumb, I classify these models as useful, giving a false sense that these variables explain the spatial distribution of species.

Reporting the omission error associated with models is not as common practice in the literature as is reporting AUC. Yet, this metric is a meaningful way to assess the discriminatory ability of the model. Not surprisingly, clustered occurrence data consistently produced larger omission error rates, most likely due to the unequal sampling of environmental conditions to train the model. Additionally, omission error showed a positive relationship with the number of variables used in the model, with weaker relationships occurring for clustered data. Most of the variation in omission error comes from the spatial bias associated with clustered data, which occurs at all extents, although, like AUC, at the smaller extent, clustering does not have a strong effect.

Model Predictions of Potential Distributions

Thresholded model predictions (presence-absence maps) are often used to find new species locations (Alfaro-Saiz et al. 2015; Mizsei et al. 2016), identify areas of conservation importance (Johnson and Gillingham 2005; Micchi de Barros Ferraz et al. 2012), investigate biological invasions (Jones et al. 2010; Mainali et al. 2015), and assess climate change impacts on biodiversity (Hijmans and Graham 2006; Forester et al. 2013). Indeed, maps of potential distributions are perhaps the most valuable aspect of SDM. In my study, when comparing the amount of the study area that was predicted as suitable, models at the medium extent produced the narrowest predictions and exhibited moderate amounts of omission error. Thus, these models do an acceptable job of discriminating areas of appropriate environmental conditions whereas models at the large and small extents showed relatively larger predicted areas, with overall higher omission error for the small extent models and lower for the large extent models, possibly hinting at 'overfit' and 'underfit' models, respectively. Spatially biased occurrence locations always produced smaller predicted areas at all three extents, most likely due to reduced sampling of environmental conditions, particularly at the largest extent. Additionally, using more variables decreased suitable area predictions, although at the small extent the models may have not incorporated enough environmental variability, making the number of variables less important in constraining the model. However, VanDerWal et al. (2009) found that the number of variables contributing to the model decreased with extent size. Vale et al. (2014) recommend caution in choosing study area extent because smaller extents may not predict marginal or suboptimal habitat used by the species at the periphery of their geographic range, however determining the appropriate extent is often not possible due to a lack of ecological knowledge about the species (Anderson and Raza 2010; Barve et al. 2011).

When comparing model predictions, I found that spatially biased models (with clustered occurrence data) almost always showed the least similarity with full occurrence models, most noticeably at the largest extent. Both spatial bias in occurrence data and the variables used in the model can produce wildly different predictions, an outcome further exacerbated by the extent of study area. This may be due to a change in the top contributing variables for models at different extents. Variable importance is strongly linked to extent size (VanDerWal et al. 2009; Jarnevich et al. 2015), for example climatic variables generally contribute more at large spatial extents and biotic variables appear to constrain distributions more at smaller extents (Luoto et al. 2007). Bradie and Leung (2017) analyzed 2,040 published Maxent models that used 400 unique environmental variables and found that temperature was one of the most frequently used variable and mean temperature contributed more to models than did minimum or maximum temperature; however, the study area extent was not taken into consideration. I found that mean temperature always outperformed minimum and maximum temperature at the largest extent, but did not contribute more to either the medium or small extent models. Rather, minimum or maximum temperature or other variables generally outperformed mean temperature at these two extents, which agrees with the notion that ecological processes operate at different spatial scales (McGill 2010).

Confidence in models and potential distribution estimates may be misplaced when model performance drives decisions or parameters are selected in absence of ecological knowledge. The effects of occurrence data (Johnson and Gillingham 2005; Graham et al. 2008; Lozier et al. 2009; Anderson 2012), environmental variable selection (Johnson and Gillingham 2005), and extent size (Barve et al. 2011) on model performance have been well documented, but not concomitantly. I used Bell's Vireo as a case study to investigate the interacting influences of these data choices for a presence-background modeling algorithm (Maxent). I showed that the degree to which spatially biased occurrence data shape model performance (AUC and omission error) and model predictions (area predicted suitable and map kappa statistic) varies across the three extents. My study adds to the growing awareness of user choice in SDM and advocates for careful deliberation in both model construction and interpretation.

Considerations for Bell's Vireo

My study describes environmental conditions constraining distributions of Bell's Vireo on relatively short time frames (decadal). Across its range, Bell's Vireo shows noticeable variation in habitat utilization (Parody and Parker 2002) and thus the environmental conditions shaping the species distribution are expected to change across different extents. At the largest extent, I found temperature to contribute the most to shaping the distribution, usually twice as much as the second most important variable. The difference in temperature between presence and background locations was likely more distinct at the large extent where all three temperature variables showed the most variation. Precipitation was generally the second most important variable for clustered models as well as contributing the most at the medium extent. In the Bell's Vireo ecological niche model constructed by Klicka et al. (2016), four variables representing temperature and precipitation (mean annual temperature, mean temperature of the warmest quarter, mean annual precipitation, and precipitation of the warmest quarter) were used to describe the current niche. Although the authors do not report the contribution of the variables, the model predicted well the current Bell's Vireo range including populations not represented in model training (Klicka et al. 2016).

Interestingly, vegetation variables performed well at the medium extent, but only for non-clustered models. This suggests that, at a regional scale, vegetation density and cover best discriminate between presence and background, but only when occurrence data are not spatially biased (i.e. a random selection of occurrence across the study extent). This is likely due to findings presented by Parody and Parker (2002) describing that nest placement of Bell's Vireo occurred in the densest vegetation within habitat patches, suggesting the strong influence of vegetation physiognomy at more local scales (Rotenberry 1985). Despite the clear role of vegetation at smaller extents, my results showed that at the smallest extent temperature variables still best described Bell's Vireo occurrences. Likely, the environmental conditions that interact to shape the distribution of Bell's Vireos at this smaller extent operate at a smaller spatial grain size than the 4 km used in this study.

Recommendations and Future Directions

First, I recommend a study extent design that considers the spatial and environmental properties of the occurrence data in association with the background selection. Secondly, clustered occurrences should be removed, as models based on smaller datasets, non-clustered, performed better. Lastly, the selection of environmental variables directly influences model output and should reflect the expected variability across the study extent, in addition to ecological relevance to species studied.

Spatial extent is only one part of spatial scale; its complement, grain size, also influences the environmental variability sampled, and thus the estimates of species distributions obtained with SDM (Saura 2002; Barve et al. 2011; Suarez-Seoane et al. 2014). I did not investigate the interaction of occurrence data and variable selection with spatial resolution (grain size). Since extent size did directly affect model performance and predictions, I expect that the spatial resolution of the environmental variables is yet another aspect of data input that must be considered. However, choice of grain size is restricted by availability of environmental datasets, thus its variability is more limited than that of study area, which is defined by the investigator.

REFERENCES

Acevedo, P., A. Jimenez-Valverde, J. M. Lobo, and R. Real. 2012. Delimiting the geographical background in species distribution modelling. Journal of Biogeography 39:1383-1390.

Aguirre-Gutierrez, J., L. G. Carvalheiro, C. Polce, E. E. van Loon, N. Raes, M. Reemer, and J. C. Biesmeijer. 2013. Fit-for-purpose: species distribution model performance depends on evaluation criteria - Dutch hoverflies as a case study. PLOS ONE 8: e63708.

Alfaro-Saiz, E., M. E. Garcia-Gonzalez, S. del Rio, A. Penas, A. Rodriguez, and R. Alonso-Redondo. 2015. Incorporating bioclimatic and biogeographic data in the construction of species distribution models in order to prioritize searches for new populations of threatened flora. Plant Biosystems 149:827-837.

Anderson, R. P. 2012. Harnessing the world's biodiversity data: promise and peril in ecological niche modeling of species distributions. In Blavatnik Awards for Young Scientists 2011, vol. 1260 Annals of the New York Academy of Sciences 126066-80.

Anderson, R. P., and A. Raza. 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (Genus *Nephelomys*) in Venezuela. Journal of Biogeography 37:1378-1393.

Araujo, M. B., and A. T. Peterson. 2012. Uses and misuses of bioclimatic envelope modeling. Ecology 93:1527-1539.

Austin, M. P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecological Modelling 157:101-118.

Austin, M. P. 2007. Species distribution models and ecological theory: a critical assessment and some possible new approaches. Ecological Modelling 200:1-19.

Austin, M. P., and K. P. Van Niel. 2011. Improving species distribution models for climate change studies: variable selection and scale. Journal of Biogeography 38:1-8.

Barve, N., V. Barve, A. Jimenez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberon, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling 222:1810-1819.

Bell, D. M., and D. R. Schlaepfer. 2016. On the dangers of model complexity without ecological justification in species distribution modeling. Ecological Modelling 330:50-59.

Bradie, J., and B. Leung. 2017. A quantitative synthesis of the importance of variables used in Maxent species distribution models. Journal of Biogeography 44:1344-1361.

Budnik, J. M., D. E. Burhans, M. R. Ryan, and F. R. Thompson. 2001. Nest desertion and apparent nest protection behavior by Bell's Vireos in response to Cowbird parasitism. The Condor 103:639-643.

Budnik, J. M., M. R. Ryan, and F. R. Thompson. 2000. Demography of Bell's Vireos in Missouri grassland-shrub habitats. Auk 117:925-935.

Dennis, R. L. H., and C. D. Thomas. 2000. Bias in butterfly distribution maps: the influence of hot spots and recorder's home range. Journal of Insect Conservation 4:73-77.

Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carre, J. R. Garcia Marquez, B. Gruber, B. Lafourcade, P. J. Leitao, T. Muenkemueller, C. McClean, P. E. Osborne, B. Reineking, B. Schroeder, A. K. Skidmore, D. Zurell, S. Lautenbach. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27-46.

Dormann, C. F., O. Purschke, J. R. G. Marquez, S. Lautenbach, and B. Schroeder. 2008. Components of uncertainty in species distribution analysis: a case study of the Great Grey Shrike. Ecology 89:3371-3386.

Dunkin, S. W., and F. S. Guthery. 2010. Bird nesting in Chickasaw plum related to age of plum in Oklahoma. American Midland Naturalist 164:151-156.

eBird. 2015. Basic Dataset Version: Ebd_Relnov-2015. Cornell Lab of Ornithology, Ithaca, New York.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, et al. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151.

Elith, J., S. J. Phillips, T. Hastie, M. Dudik, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of Maxent for ecologists. Diversity and Distributions 17:43-57.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38-49.

Forester, B. R., E. G. DeChaine, and A. G. Bunn. 2013. Integrating ensemble species distribution modelling and statistical phylogeography to inform projections of climate change impacts on species distributions. Diversity and Distributions 19:1480-1495.

Franklin, J. 2010. Mapping species distributions: spatial inference and prediction. Cambridge University Press, New York.

Franzreb, K. 1987. Endangered status and strategies for conservation of the Least Bell's Vireo. Western Birds 18:43 - 49.

Graham, C. H., J. Elith, R. J. Hijmans, A. Guisan, A. T. Peterson, B. A. Loiselle. 2008. The influence of spatial errors in species occurrence data used in distribution models. Journal of Applied Ecology 45:239-247.

Guillera-Arroita, G., J. J. Lahoz-Monfort, J. Elith, A. Gordon, H. Kujala, P. E. Lentini, M. A. McCarthy, R. Tingley, and B. A. Wintle. 2015. Is my species distribution model fit for purpose? Matching data and models to applications. Global Ecology and Biogeography 24:276-292.

Guisan, A., C. H. Graham, J. Elith, F. Huettmann, and N. S. Distri. 2007. Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions 13:332-340.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135:147-186.

Hagen, A. 2002. Multi-method assessment of map similarity. 5th AGILE Conference on Geographic Information Science.

Hanczar, B., J. P. Hua, C. Sima, J. Weinstein, M. Bittner, and E. R. Dougherty. 2010. Small-sample precision of ROC-related estimates. Bioinformatics 26:822-830.

Heikkinen, R. K., M. Luoto, and R. Virkkala. 2006. Does seasonal fine-tuning of climatic variables improve the performance of bioclimatic envelope models for migratory birds? Diversity and Distributions 12:502-510.

Hernandez, P. A., C. H. Graham, L. L. Master, and D. L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling Methods. Ecography 29:773-785.

Hijmans, R. J., and C. H. Graham. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. Global Change Biology 12:2272-2281.

IUCN. 2013. International Union for Conservation of Nature Red List of Threatened Species Version 2013.2.

Jarnevich, C. S., T. J. Stohlgren, S. Kumar, J. T. Morisette, and T. R. Holcombe. 2015. Caveats for correlative species distribution modeling. Ecological Informatics 29:6-15.

Jimenez-Valverde, A. 2012. Insights into the Area Under the Receiver Operating Characteristic Curve (AUC) as a discrimination measure in species distribution modelling. Global Ecology and Biogeography 21:498-507.

Jimenez-Valverde, A., N. Barve, A. Lira-Noriega, S. P. Maher, Y. Nakazawa, M. Papes, J. Soberon, J. Sukumaran, and A. T. Peterson. 2011. Dominant climate influences on North American bird distributions. Global Ecology and Biogeography 20:114-118.

Jimenez-Valverde, A., J. M. Lobo, and J. Hortal. 2008. Not as good as they seem: the importance of concepts in species distribution modelling. Diversity and Distributions 14:885-890.

Johnson, C. J., and M. P. Gillingham. 2005. An evaluation of mapped species distribution models used for conservation planning. Environmental Conservation 32:117-128.

Jones, C. C., S. A. Acker, and C. B. Halpern. 2010. Combining local- and large-scale models to predict the distributions of invasive plant species. Ecological Applications 20:311-326.

Klicka, L. B., B. E. Kus, P. O. Title, and K. J. Burns. 2016. Conservation genomics reveals multiple evolutionary units within Bell's Vireo (*Vireo bellii*). Conservation Genetics 17:455-471.

Lebbin, D. A., M. J. Parr, and G. H. Fenwick. 2010. The American Bird Conservatory guide to bird conservation. University of Chicago Press, Chicago.

Liu, C. R., P. M. Berry, T. P. Dawson, and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385-393.

Lobo, J. M., A. Jimenez-Valverde, and J. Hortal. 2010. The uncertain nature of absences and their importance in species distribution modelling. Ecography 33:103-114.

Lobo, J. M., A. Jimenez-Valverde, and R. Real. 2008. AUC: a misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17:145-151.

Lozier, J. D., P. Aniello, and M. J. Hickerson. 2009. Predicting the distribution of Sasquatch in Western North America: anything goes with ecological niche modelling. Journal of Biogeography 36:1623-1627.

Luoto, M., J. Poyry, R. K. Heikkinen, and K. Saarinen. 2005. Uncertainty of bioclimate envelope models based on the geographical distribution of species. Global Ecology and Biogeography 14:575-584.

Luoto, M., R. Virkkala, and R. K. Heikkinen. 2007. The role of land cover in bioclimatic models depends on spatial resolution. Global Ecology and Biogeography 16:34-42.

Mainali, K. P., D. L. Warren, K. Dhileepan, A. McConnachie, L. Strathie, G. Hassan, D. Karki, B. B. Shrestha, and C. Parmesan. 2015. Projecting future expansion of invasive species: comparing and improving methodologies for species distribution modeling. Global Change Biology 21:4464-4480.

McGill, B. J. 2010. Matters of scale. Science 328:575-576.

McMahon, G., S. M. Gregonis, S. W. Waltman, J. M. Omernik, T. D. Thorson, J. A. Freeouf, A. H. Rorick, and J. E. Keys. 2001. Developing a spatial framework of common ecological regions for the conterminous United States. Environmental Management 28:293-316.

Merow, C., M. J. Smith, and J. A. Silander, Jr. 2013. A practical guide to Maxent for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36:1058-1069.

Micchi de Barros Ferraz, K. M. P., S. F. de Barros Ferraz, R. C. de Paula, B. Beisiegel, and C. Breitenmoser. 2012. Species distribution modeling for conservation purposes. Natureza & Conservacao 10:214-220.

Mizsei, E., B. Uveges, B. Vagi, M. Szabolcs, S. Lengyel, W. P. Pfliegler, Z. T. Nagy, and J. P. Toth. 2016. Species distribution modelling leads to the discovery of new populations of one of the least known European snakes, *Vipera ursinii graeca*, in Albania. Amphibia-Reptilia 37:55-68.

Mod, H. K., D. Scherrer, M. Luoto, and A. Guisan. 2016. What we use is not what we know: environmental predictors in plant distribution models. Journal of Vegetation Science 27:1308-1322.

Moudry, V., and P. Simova. 2012. Influence of positional accuracy, sample size and scale on modelling species distributions: a review. International Journal of Geographical Information Science 26:2083-2095.

Mouton, A. M., B. De Baets, and P. L. M. Goethals. 2010. Ecological relevance of performance criteria for species distribution models. Ecological Modelling 221:1995-2002.

Mumford, R. E. 1952. Bell's Vireo in Indiana. The Wilson Bulletin 64:224-233.

Muscarella, R., P. J. Galante, M. Soley-Guardia, R. A. Boria, J. M. Kass, M. Uriarte, and R. P. Anderson. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods in Ecology and Evolution 5:1198-1205.

Noss, R. F., E. T. LaRoe, and M. J. Scott. 1995. Endangered ecosystems of the United States: a preliminary assessment of loss and degradation. United States Department of the Interior, National Biological Service, Washington D.C.

Omernik, J. M. 1987. Ecoregions of the conterminous United-States. Annals of the Association of American Geographers 77:118-125.

Palaoro, A. V., M. M. Dalosto, G. C. Costa, and S. Santos. 2013. Niche conservatism and the potential for the crayfish *Procambarus clarkii* to invade South America. Freshwater Biology 58:1379-1391.

Parker, T. H. 1999. Responses of Bell's Vireos to brood parasitism by the Brown-Headed Cowbird in Kansas. The Wilson Bulletin 111:499-504.

Parody, J. M., and T. H. Parker. 2002. Biogeographic variation in nest placement: a case study with conservation implications. Diversity and Distributions 8:11-19.

Pearson, R. G., C. J. Raxworthy, M. Nakamura, and A. T. Peterson. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeography 34:102-117.

Peterson, A. T. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3:59-72.

Peterson, A. T., M. Papes, and J. Soberon. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecological Modelling 213:63-72.

Peterson, A. T., J. Soberon, R. G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araujo. 2011. Ecological niches and geographic distributions. Princeton University Press, Princeton.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecological Applications 19:181-197.

Pitelka, F. A., and E. J. Koestner. 1942. Breeding behavior of Bell's Vireo in Illinois. The Wilson Bulletin 54:97-106.

Pontius, R. G. 2000. Quantification error versus location error in comparison of categorical maps. Photogrammetric Engineering and Remote Sensing 66:1011-1016.

PRISM. 2004. Prism Climate Group. in Oregon State University http://prism.oregonstate.edu, accessed 13 March 2014.

Reddy, S., and L. M. Davalos. 2003. Geographical sampling bias and its implications for conservation priorities in Africa. Journal of Biogeography 30:1719-1727.

Remsen, J. V., Jr., S. W. Cardiff, and D. L. Dittmann. 1996. Timing of migration and status of Vireos (*Vireonidae*) in Louisiana. Journal of Field Ornithology 67:119-140.

Ridgway, R., and H. Friedmann. 1919. The Birds of North and Middle America: a descriptive catalogue of the higher groups, genera, species, and subspecies of birds known to occur in North America, from the Arctic lands to the Isthmus of Panama, the West Indies and other Islands of the Caribbean Sea, and the Galapagos Archipelago. Smithsonian Institution Press, Washington D.C.

Rodriguez-Castaneda, G., A. R. Hof, R. Jansson, and L. E. Harding. 2012. Predicting the fate of biodiversity using species' distribution models: enhancing model comparability and repeatability. PLOS ONE 7: e44402.

Rotenberry, J. T. 1985. The role of habitat in avian community composition-physiognomy or floristics. Oecologia 67:213-217.

Sanchez-Fernandez, D., J. M. Lobo, and O. L. Hernandez-Manrique. 2011. Species distribution models that do not incorporate global data misrepresent potential distributions: a case study using Iberian Diving Beetles. Diversity and Distributions 17:163-171.

Saura, S. 2002. Effects of minimum mapping unit on land cover data spatial configuration and composition. International Journal of Remote Sensing 23:4853-4880.

Schulman, L., T. Toivonen, and K. Ruokolainen. 2007. Analyzing botanical collecting effort in Amazonia and correcting for it in species range estimation. Journal of Biogeography 34:1388-1399.

Sleeter, B. M., T. L. Sohl, T. R. Loveland, R. F. Auch, W. Acevedo, M. A. Drummond, K. L. Sayler, and S. V. Stehman. 2013. Land-cover change in the conterminous United States from 1973 to 2000. Global Environmental Change 23:733-748.

Soberón, J. M. 2010. Niche and area of distribution modeling: a population ecology perspective. Ecography 33:159-167.

Stockwell, D. R. B., and A. T. Peterson. 2002. Effects of sample size on accuracy of species distribution models. Ecological Modelling 148:1-13.

Suarez-Seoane, S., E. Virgos, O. Terroba, X. Pardavila, and J. M. Barea-Azcon. 2014. Scaling of species distribution models across spatial resolutions and extents along a biogeographic gradient. The case of the Iberian Mole *Talpa occidentalis*. Ecography 37:279-292.

Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich, A. Farnsworth, D. Fink, et al. 2014. The eBird enterprise: an integrated approach to development and application of citizen science. Biological Conservation 169:31-40.

Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. Science 240:1285-1293.

Synes, N. W., and P. E. Osborne. 2011. Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. Global Ecology and Biogeography 20:904-914.

Tanner, E. P., M. Papes, R. D. Elmore, S. D. Fuhlendorf, and C. A. Davis. 2017. Incorporating abundance information and guiding variable selection for climate-based ensemble forecasting of species' distributional shifts. PLOS ONE 12: e0184316.

Thogmartin, W. E., M. Gallagher, N. Young, J. J. Rohweder, F. Durbian, and M. G. Knutson. 2009. Avian assemblages in the Lower Missouri floodplain. Wetlands 29:552-562.

USGS. 2014. U.S. Geological Survey Earth Resources Observation and Science Center, Hydro 1k Elevation Derivative Database. in NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center. https://lpdaac.usgs.gov Sioux Falls, South Dakota accessed 13 March 2014.

Vale, C. G., P. Tarroso, and J. C. Brito. 2014. Predicting species distribution at range margins: testing the effects of study area extent, resolution and threshold selection in the Sahara-Sahel Transition Zone. Diversity and Distributions 20:20-33.

VanDerWal, J., L. P. Shoo, C. Graham, and S. E. William. 2009. Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? Ecological Modelling 220:589-594.

Warren, D. L., A. N. Wright, S. N. Seifert, and H. B. Shaffer. 2014. Incorporating model complexity and spatial sampling bias into ecological niche models of climate change risks faced by 90 California vertebrate species of concern. Diversity and Distributions 20:334-343.

Wiens, J. A., D. Stralberg, D. Jongsomjit, C. A. Howell, and M. A. Snyder. 2009. Niches, models, and climate change: assessing the assumptions and uncertainties. Proceedings of the National Academies of Science 106:19729-19736.

Yackulic, C. B., R. Chandler, E. F. Zipkin, J. A. Royle, J. D. Nichols, E. H. C. Grant, and S. Veran. 2013. Presence-only modelling using Maxent: when can we trust the inferences? Methods in Ecology and Evolution 4:236-243.

Zeng, Y. W., L. B. Wei, and D. C. J. Yeo. 2016. Novel methods to select environmental variables in Maxent: a case study using invasive crayfish. Ecological Modelling 341:5-1

CHAPTER III

GRAIN SIZE DOES NOT AFFECT OVERALL HETEROGENEITY OF ENVIRONMENTAL PREDICTORS, BUT MODIFIES MODEL PERFORMANCE AND ESTIMATED POTENTIAL DISTRIBUTIONS OF SPECIES

Abstract

Ecological pattern and processes are shaped by spatial and temporal scale components to which species respond making multi-scale investigations needed to find the scales at which species respond to environmental predictors. In species distribution modeling (SDM) the spatial scale of environmental predictors has been shown to influence both model performance and predicted potential distributions. Resampling environmental predictors is common in the SDM modeling framework, but the influence of different resampling techniques on environmental predictors used in SDM and the overall effect on model performance and potential distributions is unknown. Similarly, increases in grain size of the environmental predictors has led to an inconclusive understanding of its influence on model performance. To assess these interactions, I compared model performance and potential distributions across twelve grain sizes, which were resampled using three different methodologies, nearest neighbor, bilinear interpolation, and aggregation. I used Bell's Vireo (*Vireo bellii*) to test the impacts of resampling and grain size on SDM performance and predictions. Overall, I found that resampling techniques did not meaningfully affect the overall heterogeneity of the environmental predictors, and interestingly did not differ

between the techniques. Further, my results showed that model performance in terms AUC was influenced by resampling method, but not grain size, whereas the model performance metric, omission error, was not influenced by resampling technique or grain size. However, model prediction of potentially suitable area was influenced by both resampling and grain size. I suggest that the spatial structure of environmental heterogeneity in the study extent exerts a large influence on variation changes due to resampling technique as well as grain size. Thus, modelers should consider the interactions between the study area extent, resampling method, and grain size as well as the expected scale of species response for SDM investigations.

Introduction

Species-habitat associations are comprised of complex ecological relationships that vary along spatial and temporal scales, as well as across species and landscapes (Wiens 1989). Marceau (1999) defines scales as "a continuum through which entities, patterns, and processes can be observed." The importance of scale in ecological research has long been documented (Levin 1992, Chave 2013) and the discussion now encompasses scaling in species habitat modeling and species distribution modeling (SDM) (Guisan et al. 2007a, Barve et al. 2011, Martin and Fahrig 2012, Song et al. 2013, Suarez-Seoane et al. 2014). Species respond to ecological processes operating at multiple spatial scales that shape observed species' distributions (McGill 2010). Scale is as an important consideration in SDM as other model specifications such as choice of occurrence data (Wisz et al. 2008, Boria and Blois 2018), environmental variables (Braunisch et al. 2013, Bradie and Leung 2017), study region (Yackulic and Ginsberg 2016), algorithm (Elith et al. 2006), and model parameters (Anderson and Gonzalez 2011, Elith et al. 2011). Further, scale can interact with these model characteristics to influence model performance (Bean et al. 2012, Vale et al. 2014, Connor et al. 2018).

Scale is comprised of extent and grain, each exerting their own influence on SDM performance and predictions. Extent is the total area under consideration whereas grain refers to the spatial size of the observation (Wiens 1989). Across a landscape, between grain and within grain variation is constrained by extent and grain size, respectively (Levin 1992, Anderson 2018). Additionally, observed variation across an extent is also highly linked to the system and will influence how ecological phenomena are measured under different extents and grain sizes as well as time (Wiens 1989). This creates a challenge for SDM, as individuals using a landscape are often responding to multiple habitat variables at multiple spatial and temporal scales, especially for heterogeneous areas (With et al. 1997). SDM measures the suitability of geographic landscapes to species and is susceptible to scale effects (Song et al. 2013). To estimate the distribution of a species, SDM uses known presence locations and the associated environmental conditions (Phillips et al. 2006, Elith et al. 2011), thus the geographic extent, which comprises the possible range in environmental conditions used in the model, should be carefully considered (Barbet-Massin et al. 2010). For presence-background algorithms such as Maxent, background points are selected from the training extent to delineate a suitability signal by contrasting presence and background conditions (Elith et al. 2011). An extent that is overly broad may disrupt the model's ability to characterize this signal due to increased heterogeneity contained in the background, influencing both model performance (Lobo et al. 2008) and predicted suitability (Bean et al. 2014). Smaller extents may underrepresent environmental conditions in the training background creating biased models that do not fully represent species' potential distributions (Barbet-Massin et al. 2010, Sanchez-Fernandez et al. 2011). Decreasing extent size is generally thought to reduce variation in the background conditions sampled (Song et al. 2013), however this is dependent on the landscape and environmental variables considered (first chapter, this dissertation). The variation in environmental covariate conditions shapes the distribution of background samples and the fitted distribution of presence locations (Elith et al. 2011), thus affecting comparisons, as they depend on model predictions. Variation of background samples is also affected by grain size since increasing grain homogenizes the landscape, inhibiting model discrimination between presence and background environmental conditions (Song et al. 2013). The limited research of SDM-grain relationships shows mixed model performance, measured by AUC, in relation to coarsening grain sizes (Seoane et al. 2004, Guisan et al. 2007a, Seo et al. 2009, Gottschalk et al. 2011, Bellamy et al. 2013, Suarez-Seoane et al. 2014, Scales et al. 2017, Connor et al. 2018). More specifically, when investigating multiple taxa and study regions, Guisan et al. (2007a) concluded that model performance, decreased, increased, or did not change with larger grains, likely due to interacting influences. However, at coarser resolutions, Seo et al. (2009) found that, in combination with decreasing AUC, predicted suitable area for trees increased under future climate conditions, likely overestimating potential distributions. Coarsening grain sizes decreased rare land cover types and resulted in mixed model performance for 13 bird species, further suggesting confounding interactions between species and landscapes (Gottschalk et al. 2011). Song et al. (2013) compared the influence of resampling cell values to represent a neighborhood statistic (keeping original grain size) to aggregating cells into larger cells representing the same sized neighborhood and found that the first technique slightly increased AUC whereas the second approach decreased model performance measured by AUC. They concluded that the information surrounding a data cell is more important for predicting species' distributions than the actual data cell value (Song et al. 2013). More recently, an analysis using virtual species to test SDM behavior regarding grain size changes revealed that using environmental variables at scales to which the species responds produces the best performing models (Connor et al. 2018). The scales at which species are most influenced by environmental conditions vary with species, geographic extent, and environmental predictors (McGill 2010) and disentangling the effects of grain can be difficult, as true species-environment relationships are not known.

The observed reaction of species to particular scales has been named multiple times, most notably, "domains of scale" by Wiens (1989), "characteristic scale" (Urban 2005), "intrinsic scale" (Wu et al. 2006) and specifically for grain, "response grain" (Mertes and Jetz 2018). To identify these few but important scales (McGarigal et al. 2016), a continuous scale spectrum must be employed (Wiens 1989, Wheatley and Johnson 2009). This allows for the detection of breaks or thresholds within the domain of scale for the observed ecological process that indicate a change in the scale relationship (Wiens 1989, Levin 1992, Marceau 1999). Since no single measurement scale is likely to capture the entirety of an ecological process, multi-scale ecological data, is needed.

Remote sensing data provides environmental data on multiple spatial and temporal scales (Kerr and Ostrovsky 2003, Anderson 2018) and incorporating remotely sensed data into SDM may lead to increased model performance (Leyequien et al. 2007, Buermann et al. 2008). Although for plant-based SDM, remotely sensed data have introduced bias, especially for plant distribution models as these variables may actually map species (Bradley et al. 2012). Since species interact with their environment at multiple spatial and temporal scales, environmental parameters derived from remote sensing may better capture variation in the environment, which fashions the observed species' distributions across space. Many environmental predictor data sets, particularly climate data, are interpolations of long-term data sets (e.g. Worldclim (Hijmans et al. 2005) and PRISM (Daly et al. 2000)); whereas remote sensing directly measures many ecological processes occurring on the surface of the Earth (He et al. 2015). Currently in orbit there are countless Earth observing systems differing in the temporal, spatial, and spectral specifications that are used to measure variation on the Earth's surface, thus providing opportunities to investigate species-scale relationships that were not possible in the past (Turner et al. 2003, Anderson 2018).

Despite increased data availability with the expansion of remote sensing technology, many scale related obstacles still exist, most notably, the modifiable areal unit problem (MAUP) (Openshaw 1984). The MAUP phenomena is simply the fact that there is a large amount of ways to divide a study area extent into different grains (both sizes and shapes). Each unique defined grain will produce an observed relationship only applicable at that "scale" and changes with aggregation. Thus, much work is still needed to identify meaningful scales influencing species-environmental relationships (Marceau 1999).

Environmental predictors that match the scale most influential to species provide the most accurate model predictions, is the foremost objective of SDM. Thus, finding the appropriate extent and grain size for the most relevant environmental predictors is the optimal ecological framework for conservation professionals, decision makers, and managers. For presence-background SDM, specifically Maxent, the sampled background for environmental predictors is the basis for comparison to environmental conditions associated with presences (Elith et al. 2011).

To fully establish ideal scales for species, and identify critical thresholds, multiple spatial grains must be investigated. While many studies have only compared a few grain sizes (Guisan et al. 2007a, Guisan et al. 2007b, Revermann et al. 2012, Bean et al. 2014, Georgian et al. 2014, Suarez-Seoane et al. 2014), it is now more common to see multiple grain size comparisons (Seo et al. 2009, Bellamy et al. 2013, Liang et al. 2013, Pradervand et al. 2013, Song et al. 2013, Connor et al. 2018, Mertes and Jetz 2018). However, despite the increasing interest, a consensus has not been reached on the full extent of grain size effects on SDM. Additionally, since altering grain size is expected to modify the background-presence comparison, as changes in measurement scales are expected to change variation (Wiens 1989, Levin 1992), quantifying the change in a variable at each grain size is needed. To my knowledge, no studies have quantified variation change with grain manipulation, although Connor et al. (2018) reported the standard deviation of the original grain size for the two landscapes used and Mertes and Jetz (2018) grouped variables by spatial structure (fine, intermediate, and coarse). Further, Song (2013) calculated mean elevation change for increasing grain size and Gottschalk et al. (2011) calculated percent land cover change and land use configuration for each grain size. Measuring variation change is crucial since upscaling methodologies may produce disparate environmental predictors.

Traditional methods for increasing grain size or upscaling have been developed in the remote sensing field to resample spectral imagery (Atkinson 1988, Hay et al. 1997, Jensen 2005). Resampling is the process by which new cell/pixels values are assigned to locations that are not the original cell/pixel location (Atkinson 1988). Multiple resampling techniques exist, each of which uses a different methodology to calculate the new cell value, which results in different outputs (Atkinson 1988). The expectation exists that different resampling techniques used to increase grain sizes of environmental variables may create different altogether different model predictors.

The relationship between upscaling technique and grain size manipulation on SDM performance has not yet been evaluated. Thus, my objective was to investigate the choice of resampling methodology on grain size manipulation studies for SDM, by accounting for variation change in the geographical background. Specifically I tested the effects of three resampling techniques on four environmental predictors at 12 grain sizes. I expect that classic methods of evaluating model performance will be affected by resampling technique as well as grain size.

Methods

Study Species

Bell’s Vireo (*Vireo bellii*) is a migratory songbird found throughout central United States in the Great Plains, as well as southern portions of the US including parts of south Texas, New Mexico, Arizona, California, and portions of Northern Mexico. The eastern population (delineated between New Mexico and Arizona; Figure 1) constitutes two subspecies, *V.b bellii* and *V.b. medius*, and is associated with low, dense, shrubby vegetation and is often found in grassland shrub mottes and shrubland ecotones (Lebbin et al. 2010). Although closely linked to denser shrub environments, Bell's Vireo shows considerable variation in habitat choice and nest site selection in these areas (Parody and Parker 2002). I excluded the western population of Bell's Vireo, which comprises the subspecies *V. b. pusillus* and *V.b. arizonae*, as it is riparian dependent and closely linked with willow-dominated vegetation stands (Franzreb 1987). I selected Bell's Vireo for this study for several reasons: 1) it is a declining species of concern (IUCN 2013); 2) very little published research is available about its environmental requirements; and 3) the high vocalization rates make the species easily detectable in the field, thus minimizing identification and detection errors in the presence dataset.

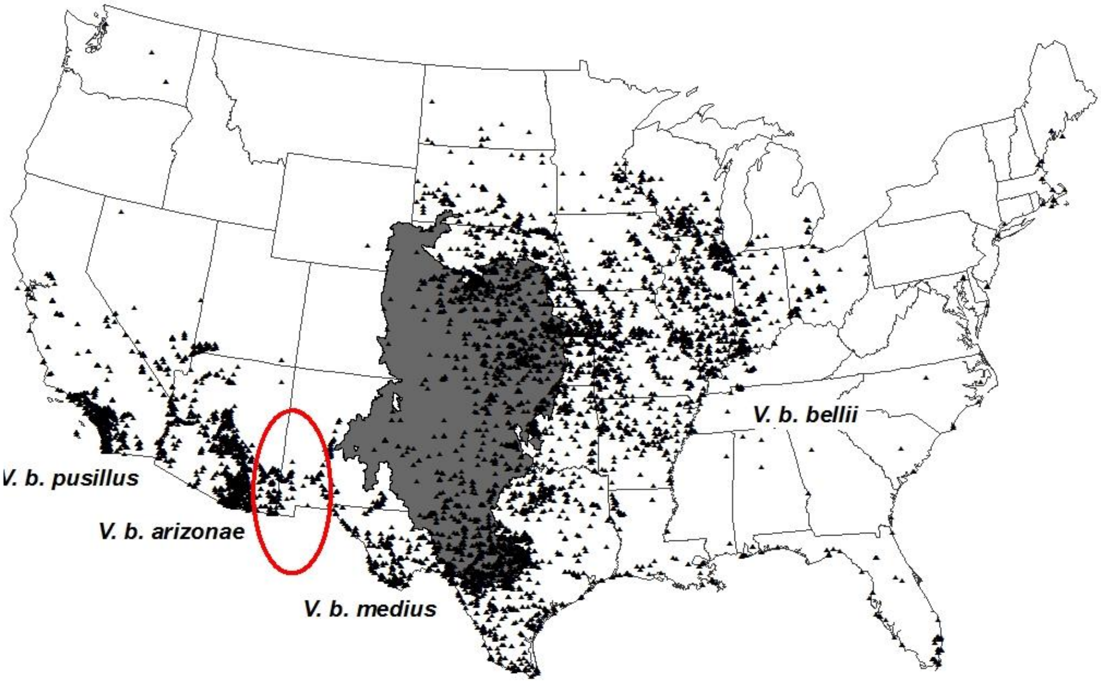


Figure 1 Bell's Vireo (*Vireo bellii*) occurrence records, black triangles, downloaded from eBird for May through August during 2004 – 2014. Not all records from Ebird were identified to the subspecies level, however for all four subspecies; at least some records did include this information. Red ellipse indicates the area delineating the eastern and western populations (modified from Klicka et al 2016). The study extent (South Central Semiarid Prairies) is shown in grey.

Study Area

The distribution of Bell’s Vireo encompasses a large portion of the central United States, and I selected the south-central portion of the Great Plains for my research. This area is defined as the South Central Semiarid Prairies, an EPA Level II ecoregion (Omernik 1987, McMahon et al. 2001). The South Central Semiarid Prairies provided a unique opportunity because was a moderately large area (~ 863,000 km2) in which both subspecies of the Bell’s Vireo eastern population occurred. Additionally, its location suggested environmental variation that would possibly explain the observed differential use of the area by Bell’s Vireo occurrence data. I defined my study extent using ecoregion boundaries because areas within the boundaries share similar biotic and abiotic capabilities (McMahon et al. 2001) and this approach is found in other SDM investigations because it is a straight-forward and an effective method for choosing extents (Soberón 2010, Barve et al. 2011).

Occurrence Data

I obtained Bell’s Vireo locations from eBird (Sullivan et al. 2009), a citizen science platform, which allows individuals to submit bird observations to an online database from which researchers can request access. Locations of birds in eBird have several collections types: traveling count, stationary count, historical, incidental, and area observation. I restricted Bell’s Vireos locations to those that were stationary observations or incidental observations for my sampling timeframe (2004 – 2014).

Environmental Variables

I selected four variables that were expected to influence the distribution of Bell’s Vireo across the South Central Semiarid Prairies extent. The four variables, aspect, wetness, normalized difference vegetation index (NDVI), and land surface temperature (LST), each contribute to shaping the variation of environmental conditions across the South Central Semiarid Prairies. In a region with little elevational change such as the Great Plains, topographic characteristics such as aspect affects microclimate conditions creating local scale variation across the landscape (Bennie et al. 2008). As the South Central Semiarid Prairies is located on the western edge of the Great Plains, it is more susceptible to rain shadow effects from the Rocky Mountains, making precipitation intermittent and relatively low (Borchert 1950). I included the wetness index as a proxy for fine scale precipitation as I expected areas with more consistent levels of precipitation to exhibit higher soil wetness. NDVI is a measure of vegetation health/biomass (Rouse et al. 1974) that is used in SDM and habitat modeling (Kerr and Ostrovsky 2003, Turner et al. 2003, Rose et al. 2015). As Bell's Vireo is highly associated with dense vegetation patches, such as shrub mottes that can differ from surrounding vegetation types, such as grassland, I included NDVI as a variable that may show a distinct signal at occurrence locations when compared to the background. Lastly, I included LST since temperature variables generally perform well in SDM and are likely to influence species distributions (Bradie and Leung 2017). I specifically selected LST over other temperature variables such as ambient temperature, because I expected local conditions to influence LST variation across the landscape, playing a role in creating suitable conditions for Bell's Vireo.

Remote Sensing Data

I calculated the four environmental variables from remote sensing data, specifically SRTM (Shuttle Radar Topographic Mission) and Landsat 7 ETM+. I used Landsat 7 data due to its high spatial resolution (30 m) and mission dates covering the temporal range of my occurrence data (Young et al. 2017). I downloaded 66 Landsat scenes (images) for each of the 11 years of interest (2004 – 2014), selecting scenes with ≤ 20% cloud cover. I selected scenes from the Bell's Vireo breeding season with a temporal priority of: late June, early July, early June, late July, late May, early August, early May, late August. As Landsat has a temporal resolution of 14 days, a scene is collected twice for each month, once during the first half (early) and once for the last half (late).

If a particular scene for one year was not available with ≤ 20% over the breeding season, I did not download a scene, thus for any given year there were coverage gaps. I chose both Level 1 and Level 2 products; Level 1 products provide data for each of the 8 bands as digital numbers, whereas Level 2 products provides preprocessed surface reflectance (for bands 1 - 5) and top of atmosphere brightness data (for band 6). In 2003, the Landsat 7 Scan Line Corrector failed, resulting in data voids of ~22% of the images. Thus, all 2,772 images (924 scenes for each data product) were corrected using a gap fill algorithm in the program ENVI version 5.3 (Exelis Visual Information Solutions, Boulder, Colorado) for bands 1, 2, 3, 4, 5, 6, and 7, before further processing. The gap fill algorithm could not resolve some scenes and so these areas also exhibited data gaps. Corrected scenes were mosaicked into yearly raster surfaces after which I used the cloud mask provided with each scene to remove any clouds before variable calculations.

Aspect

I downloaded 30 m SRTM digital elevation model (DEM) (NASA 2002) from the Land Processes Distributed Active Archive Center (LPDAAC) data portal AppEEARS (https://lpdaacsvc.cr.usgs.gov/appeears). The DEM was converted to topographic aspect using ArcMap v 10.4 (ESRI 2016), by fitting a plane to a 3x3 moving window for each elevation cell value. The rate of change in both the x and y directions are calculated and converted to compass direction (measured 0 – 360, with due north at 0 and 360) of the plane face, which is the value entered for each cell (Burrough and McDonell 1998).

Wetness Index

The tasseled-cap transformation is an orthogonal transformation that compresses spectral data into new axes (Kauth and Thomas 1976) representing four indices: soil brightness, vegetation greenness, yellow stuff/wetness, and non-such or noise. Crist (1985) proposed a tasseled cap transformation that utilized reflectance values, needed for multi-scene transformation. I calculated at-satellite reflectance using the corrected Landsat 7 level 1 products (Huang et al. 2001). Following Huang et al. (2002) I calculated yearly soil wetness index for the South Central Semiarid Prairies using the following formula:

$$wetness = 0.2626Band1 + 0.2141Band2 + 0.0926Band3 + 0.0656Band4 - 0.7629Band5 - 0.5388Band7$$

Normalized Difference Vegetation Index

Due to vegetation's ability to reflect light from 0.7 to 1.1 µm but absorb light 0.4 – 0.7 µm, a ratio between these two areas of the electromagnetic spectrum provides a measure of the density of 'greenness' or vegetation (Rouse et al. 1974). To derive NDVI for the study extent, I used the corrected yearly-mosaicked Level 2 surface reflectance images using the following formula:

$$NDVI = \frac{Band4 - Band3}{Band4 + Band3}$$

Land Surface Temperature

Yearly LST for South Central Semiarid Prairies was estimated using the methodology outlined in Walwender et al. (2014), applied to multiple scenes. Calculating LST requires multiple steps to produce the inputs needed. Walawender et al. (2014) defines a single channel LST algorithm (from (Jiménez-Muñoz and Sobrino 2003, Sobrino et al. 2004, Jimenez-Munoz et al. 2009) as:

$$LST = \gamma\left[\frac{1}{\varepsilon}(\psi_1 L_s + \psi_2) + \psi_3\right] + \delta$$

Here, γ and δ are Planck's function dependent parameters, and $\psi1$, $\psi2$, $\psi3$ are atmospheric functions, and ε is land surface emissivity. Using the Landsat level 1 product, I computed at-satellite reflectance (Ls) and used the level 2 product, at sensor brightness temperature, as Ts to find values for γ and δ.

$$\gamma = \left\{ \frac{c_2 L_s}{T_s^2} \left[\frac{\lambda^4 L_s}{c_1} + \frac{1}{\lambda} \right] \right\}^{-1}$$

$$\delta = -\gamma * L_s + T_s$$

Here, Planck's radiation constants, c1 and c2, are 1.19104 108 W µm4 sr-1 and 1.43877 104 µm K, respectively. The effective wavelength of Landsat 7 ETM+ band 6, λ, is 11.27 µm.

The three atmospheric functions, ψ1, ψ2, ψ3, were calculated by:

$$\psi 1 = \frac{1}{\tau} \qquad \psi 2 = -\mathrm{L}\downarrow - \frac{L\uparrow}{\tau} \qquad \psi 3 = \mathrm{L}\downarrow$$

Here, τ, L↑, and L↓, are atmospheric transmissivity, up-welling atmospheric radiance, and down-welling atmospheric radiance, respectively. These parameters were acquired from the web-based interface Atmospheric Correction Parameter Calculator (http://atmcorr.gsfc.nasa.gov) (Barsi et al. 2003, Barsi et al. 2005). This atmospheric calculator uses the date, time and location of a point to model the atmosphere and produce τ, L↑, and L↓ at a 1x1 degree resolution from which individual locations are interpolated. I created yearly 1x1 degree grid for the South Central Semiarid Prairies and for the centroid of each cell and recorded the date and time of image acquisition as well as the spatial location. Using this information, I queried the Atmospheric Correction Parameter Calculator to obtain the three atmospheric parameters for each location for each year, which were then used to calculate ψ1, ψ2, ψ3.

Lastly, to determine surface emissivity (ε) for the LST equation, I used the relationship between NDVI and emissivity (Van De Griend and Owe 1993, Valor and Caselles 1996), following the equation:

$$\varepsilon = \varepsilon \mathrm{v}_\lambda * P_v + \varepsilon s_\lambda * (1 - P_v) + C_\lambda$$

where εvλ, the assumed emissivity of full vegetation, is 0.99, εvλ, the assumed emissivity of soil, 0.96, Pv is the proportion of vegetation, and Cλis surface roughness correction. Pv (Carlson and Ripley 1997) was calculated from:

$$P_v = \left(\frac{NDVI - NDVI_s}{NDVI_v - NDVI_s}\right)^2$$

Here I used the NDVI variable previously calculated, and used the global values of NDVIs = 0.2 (NDVI of soil cells) and NDVIv = 0.5 (NDVI of vegetation cells) proposed by Soberon and Raissouni (2000). For the surface roughness correction, I used the equation:

$$C_\lambda = (1 - \varepsilon s_\lambda) * \varepsilon \mathrm{v}_\lambda * F' * (1 - P_v)$$

I used the value of 0.55 for F´, a geometrical factor (between 0 and 1), which Sobrino et al. (1990) calculated as the mean value for a rough and heterogeneous surface.

I averaged yearly variables across the time-period of the study (2004 – 2014) to produce one variable for each environmental parameter. The resolution of Landsat 7 ETM+ Level 2 products was slightly larger than 30m (~ 32.6m), so I resampled the environmental parameters to 32.6m using the nearest neighbor technique and established this as my base resolution. For each resampling technique (see below), I started at the smallest grain size and resampled to the calculated grain size.

Because I used remote sensing imagery to calculate variables, my environmental parameter final outputs had gaps of missing data from areas of persistent cloud cover or areas where no images with < 20% clouds were available for the study time period (11 of the 924 scenes), with the exception of aspect, which had complete coverage.

Resampling Technique

I resampled the environmental parameters from my base resolution of ~33m to 11 additional resolutions in multiples of 33. Specifically, these resolutions were: 99m, 165m, 231m, 333m, 429m, 561m, 660m, 759m, 891m, 990m, and 1089m. I used three resampling techniques: nearest neighbor, bilinear interpolation, and aggregation, see Figure 2. Resampling was done in ArcMap (ESRI 2016), where each new grain size was resampled from the base resolution of 33m.

To quantify the variation of each environmental parameter at each grain size resampled under each technique, I calculated the coefficient of variation (CV). I used ArcMap (ESRI 2016) to calculate basic statistics on each variable (mean, standard deviation, minimum, and maximum) and divided the standard deviation by the mean to standardize the variation as a proportion. This allowed us to compare variation between variables, grain sizes and resampling techniques.

Nearest Neighbor

In nearest neighbor (NN) resampling, inverse logic is used to assign a value from the nearest input cell to the output cell (Jensen 2005). The nearest neighbor resampling will upscale to any grain size from the original resolution and has two advantages. First, no pixel values are altered and there will be no missing data values in the output (Konecy 2003). Second, it is an efficient method as it requires the least amount of computation (Jensen 2005). Data loss occurs with this technique, but as the output data comes from the nearest neighbor, the loss is not necessarily due to homogenization.

Bilinear Interpolation

The Bilinear Interpolation resampling technique (BIL), assigns output cell values based on a weighted distance of the four closest input cells, where closer cells are weighted more (Jensen 2005), and it allows for the output cell to be any size (Figure 2). As the output cell value is dependent upon the surrounding input cells, data loss or homogenization is inevitable, although the magnitude of data loss is a product of the spatial structure of an environmental parameter across the landscape.

Aggregation

Aggregation (AGG) is a resampling technique that aggregates or clusters a specified number of cells into a new output cell, thus the output cell size must be a multiple of the input cell size. This means that the extent under consideration can be modified by either truncating the output for areas where the number of cells for an output raster does not exist or expanding it by using the available input cells to calculate the output cell. As with the previous technique, data homogenization occurs, but is largely driven by landscape characteristics (Hunsaker et al. 2001). To aggregate to larger grain sizes, I used the mean method, where the average of the input cells was used as the value for the output cell.

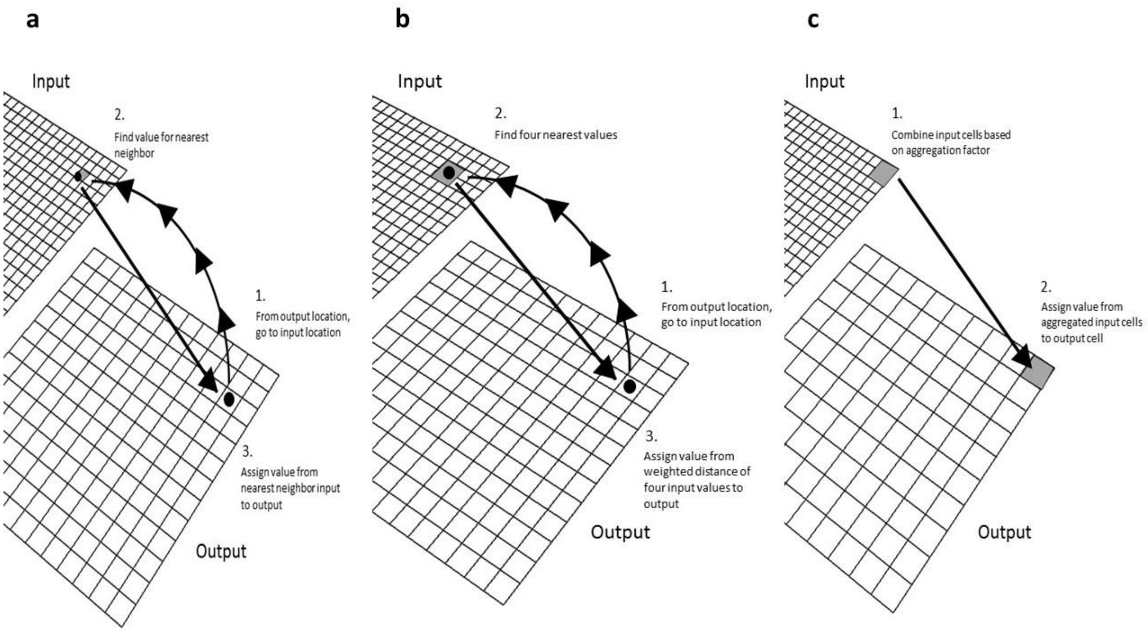


Figure 2 Illustration of resampling techniques used to upscale environmental parameters, modified from Figure 7-8 in Jensen (2005). (a) Nearest Neighbor, (b) Bilinear Interpolation, and (c) Aggregation.

Species Distribution Modeling

I used the Maximum Entropy (Maxent) algorithm to generate models at each grain size for Bell's Vireo using the four environmental parameters: aspect, LST, NVI, and wetness. Maxent is a presence-background algorithm that estimates species distributions by comparing the variation of environmental parameters of presence locations to that of points randomly selected from the background (Elith et al. 2011). Regions which have environmental conditions more similar to occurrence locations will be given higher suitability probabilities (Phillips et al. 2006).

I ran an initial model for each grain size to find spatially unique locations that fell within a data cell for all environmental variables. I then selected the occurrence locations that were common for all models to use for model training and testing. This ensured that the same number of occurrences were used for each model. Additionally, I calculated the number of background points representing ~1% of my study region for each grain size so that the background was proportionately sampled across models. I used the crossvalidate feature of Maxent to create five models each using a unique 20% of the occurrences used as training and the remaining 80% used as testing. I averaged the five models to produce one overall prediction. I selected a logistic, continuous prediction, which I thresholded with the 10% minimum training presence. This threshold leaves out 10% of the training points having the lowest suitability. The advantage to this threshold is that it can be used for datasets where the error is unknown (but should still be accounted for) as this threshold is less sensitive to outliers or incorrectly located occurrences (Peterson 2006, Peterson et al. 2011). Additionally, this threshold produces more restrictive predictions (Radosavljevic and Anderson 2014). I used Maxent’s default settings for regularization parameter, maximum iterations, convergence threshold, and prevalence as these have performed well (Phillips and Dudik 2008).

Model Evaluation

Model Performance

I used two model performance metrics to evaluate models and compare across grain sizes and resampling technique. The first was the area under the cure of receiver operating characteristic plot (ROC AUC) and the second, testing omission error. In SDM, the traditional ROC AUC plot (sensitivity, or fraction of presences predicted present vs. 1-specificity, or fraction of absences predicted absent) is modified to sensitivity vs. proportion of area predicted present (Phillips et al. 2006, Phillips et al. 2009). Models with perfect discrimination will have an area under this curve (AUC) with a value of 1, whereas models with AUC > 0.7 are considered acceptable (Swets 1988, Fielding and Bell 1997). Since the proportion of the area predicted present is highly influenced by the classification threshold as well as the size of the extent, AUC can be an unreliable criterion of model evaluation when used alone (Lobo et al. 2008, Jimenez-Valverde

2012). I also use omission error to evaluate models as a single indicator can be misleading about overall model performance (Lobo et al. 2008). Omission error is the percentage of known locations incorrectly classified as absences by the model. In general, lower omission rates indicate higher model performance (Mouton et al. 2010). I used another model to evaluate model performance, the proportion of study area predicted suitable. Maxent produces thresholded models based on the selected threshold, in this case, 10% training minimum training, to convert continuous suitability predictions to binary potential distributions. This metric is useful for evaluating models as poor performing models can have both low predicted suitable area (model unable to predict suitable areas outside the training occurrence locations) and high (models unable to discriminate between presence locations and background) (Peterson et al. 2011). I used the non-parametric Kruskal-Wallis test to compare average AUC, omission error, and predicted suitable area across resampling technique and grain size. I used the R (R Core Team) package DescTools (Signorell et al. 2018) to calculate Kruskal-Wallis comparisons.

Variable Contribution

I expected some correlation across the four variables as they are ecologically related and calculated from the same imagery (Landsat 7 ETM+) data. I also expected changes in correlations across variables with changing grain size, so I calculated correlational coefficients for all four variables at reach grain size for each resampling technique. I used SDMtoolbox 2.0 (Brown 2014, Brown et al. 2017) to calculate correlations for all variables at all grainsizes for all three resampling methods. Additionally, I compared variable contribution to model accuracy gain for all models. Maxent computes a percent contribution for each variable to the model.

Results

Environmental Variation

Across the three resampling techniques, I found that within variable variation, measured by CV, showed little to no change, with the exception of aspect (Figure 3). Aspect showed the highest variation of the four variables: under NN, aspect CV did not change and under BIL several fluctuations occurred but seemed to be noisy artifacts. However, when resampled using AGG, aspect steadily lost variation (CV decreased from 0.62 to 0.12). The remaining three variables showed little change in CV across the three resampling techniques, although the amount of variation across the environmental variables was quite large. The most variation occurred in NDVI (~0.40), followed by WET (~0.20), and lastly LST having the least amount of variation (~0.01) across all resolutions. NN and BIL showed the most similarity in CV across increasing grain size.

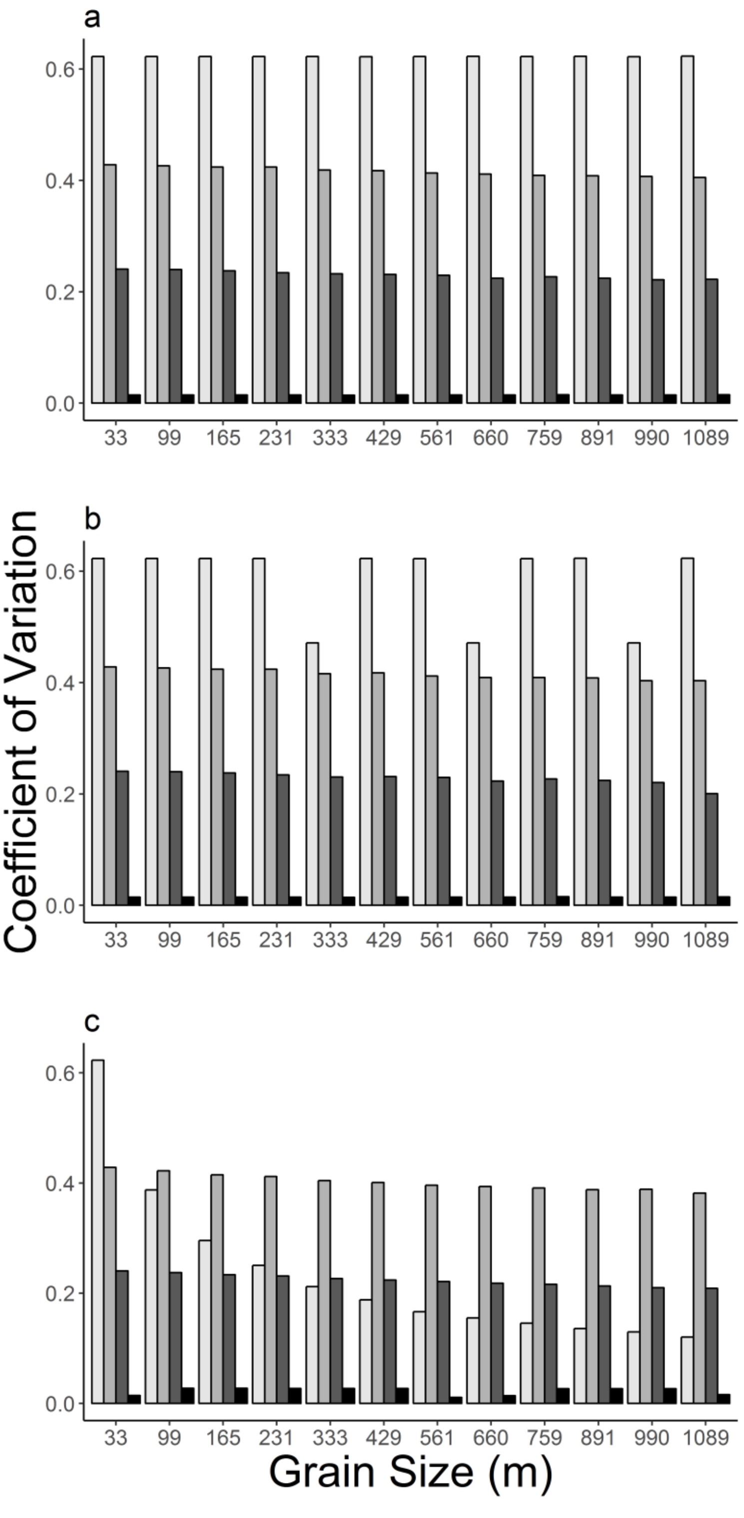


Figure 3 Coefficient of Variation (standard deviation/mean) for the four environmental parameters, aspect (light grey), NDVI (medium grey), wetness (dark grey), and land surface temperature (black) at each grain size for each resampling technique: (a) Nearest Neighbor, (b) Bilinear Interpolation, and (c) Aggregation.

Model Performance

Overall, model performance, measured by AUC, was moderate (0.747 – 0.842). I found that model performance did significantly differ for models based on resampling technique: NN (average AUC = 0.798), BIL (average AUC = 0.798), and AGG (average AUC = 0.805) (Table 1) for the Kruskal-Wallis comparison. Both NN and BIL produced models with remarkably similar performance (Figure 4), as the AUC generally decreased with grain size, but only slightly. AUC across the AGG resampled grain sizes showed less deviation in model performance (see Figure 4 (C)). However, when I compared model performance across grain sizes within each resampling method, I found no significant differences for based on resampling technique (Table 1). When model performance for each grain size was compared across the resampling techniques, no significant differences were found although 759m did come close to having significantly different model performance (p = 0.058) from the rest of the grain sizes.

Table 1 P-values for Kruskal-Wallis comparisons of average model AUC, OE, and PSA. Tech compared all models across resampling technique. Grain-NN, compares models across each grain size for the nearest neighbor resampling technique, Grain-BIL and Grain-AGG are for bilinear interpolation and aggregation respectively. Tech-30 compares all models at 30 m resolution across resampling technique, as such with models up to 990 m resolution. Numbers in bold are significant comparisons (< 0.05).

	p-value		
	AUC	OE	PSA
Tech	**0.012**	0.916	**0.000**
Grain-NN	0.363	0.999	**0.000**
Grain-BIL	0.323	0.999	**0.000**
Grain_AGG	0.719	1.000	**0.000**
Tech-33	1.000	1.000	1.000
Tech-99	0.993	0.970	**0.001**
Tech-165	0.934	0.992	**0.009**
Tech-231	0.691	0.865	**0.011**
Tech-333	0.645	0.910	**0.009**
Tech-429	0.112	0.820	0.112
Tech-561	0.293	0.990	0.332
Tech-660	0.746	0.853	0.250
Tech-759	0.058	0.992	0.321
Tech-891	0.182	0.970	0.426
Tech-990	0.275	0.679	0.925
Tech-1089	0.400	1.000	**0.035**

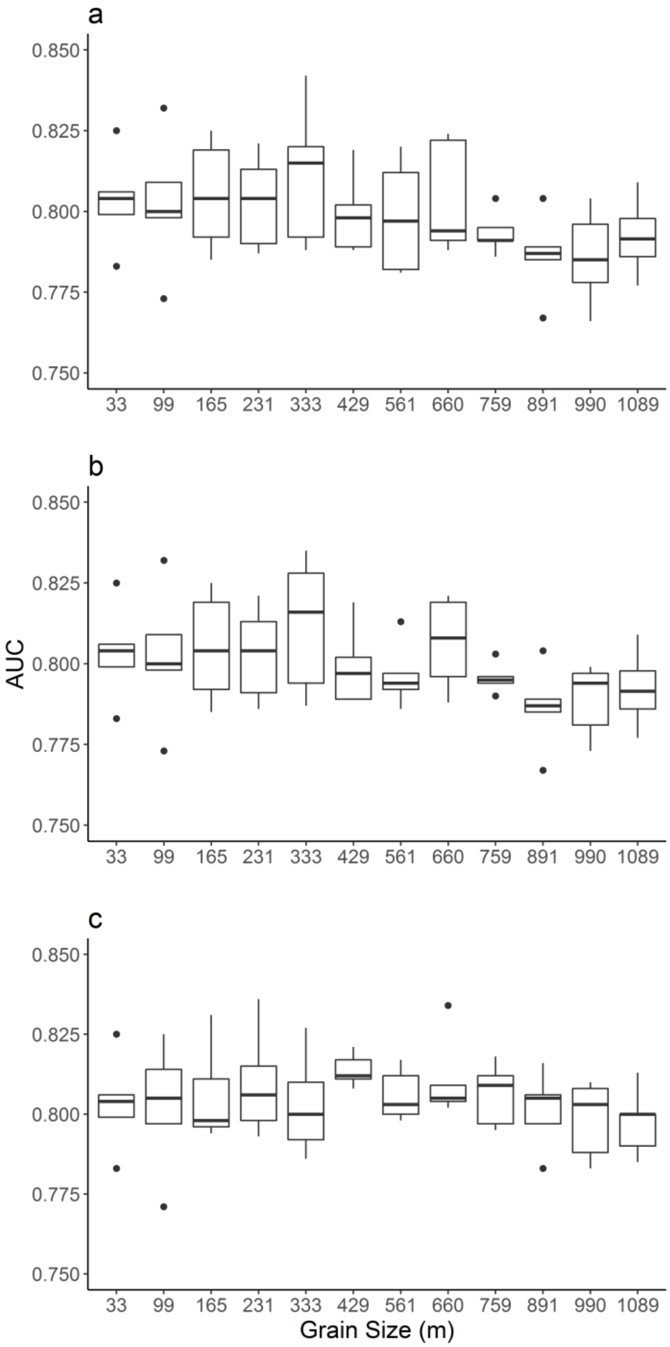


Figure 4 Boxplot showing AUC values for the five Maxent replicates for each grain size resampled with each technique (a) nearest neighbor, (b) bilinear interpolation, and (c) aggregation.

Overall, omission error was relatively low, ranging from 0.05 to 0.18 and did not significantly differ by resampling technique (Table 1). Omission error was so similar across resampling technique and grain size (Figure 5) that no Kruskal-Wallis comparison was significant (Table 1).

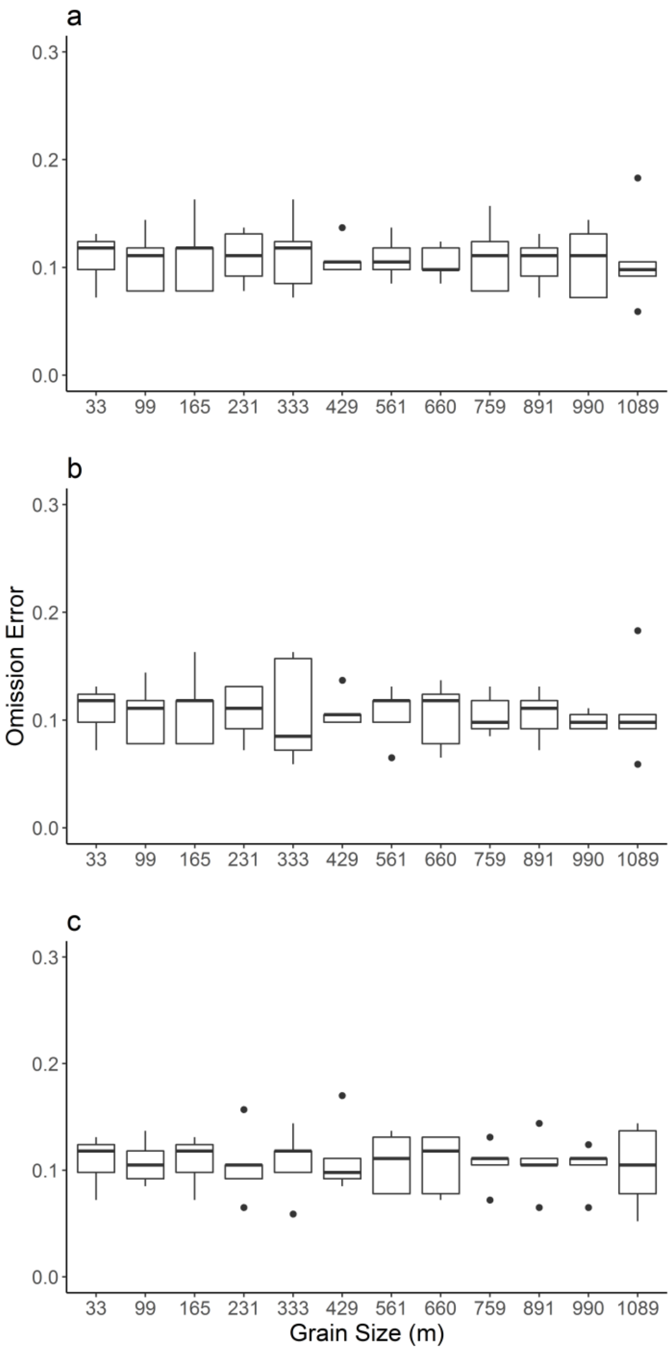


Figure 5 Boxplot of omission error for the five Maxent replicates at each grain size for (a) nearest neighbor, (b) bilinear interpolation, and (c) aggregation.

Using the 10% minimum training threshold, Maxent calculated the proportion of the study area that is predicted suitable (PSA) for each replicate. I found relatively little variation in PSA across models (0.440 – 0.558) with an increase in PSA when NN and BIL techniques were used and general having more PSA at each grain size when AGG was used (Figure 6). Variation in PSA across grain sizes was most evident with AGG. Kruskal-Wallis test showed that PSA across resampling techniques was strongly significant ($p < 0.0000$) as was PSA compared across grain size within each sampling technique (NN, BIL, and AGG; $p < 0.000$; Table 1). However, when each grain size was compared across resampling methods, Kruskal-Wallis tests showed significant differences ($p < 0.01$ and $p < 0.05$) occurred at smaller grain sizes (99m – 333m) than larger grain sizes (429m – 990m), with the exception of 1089m.

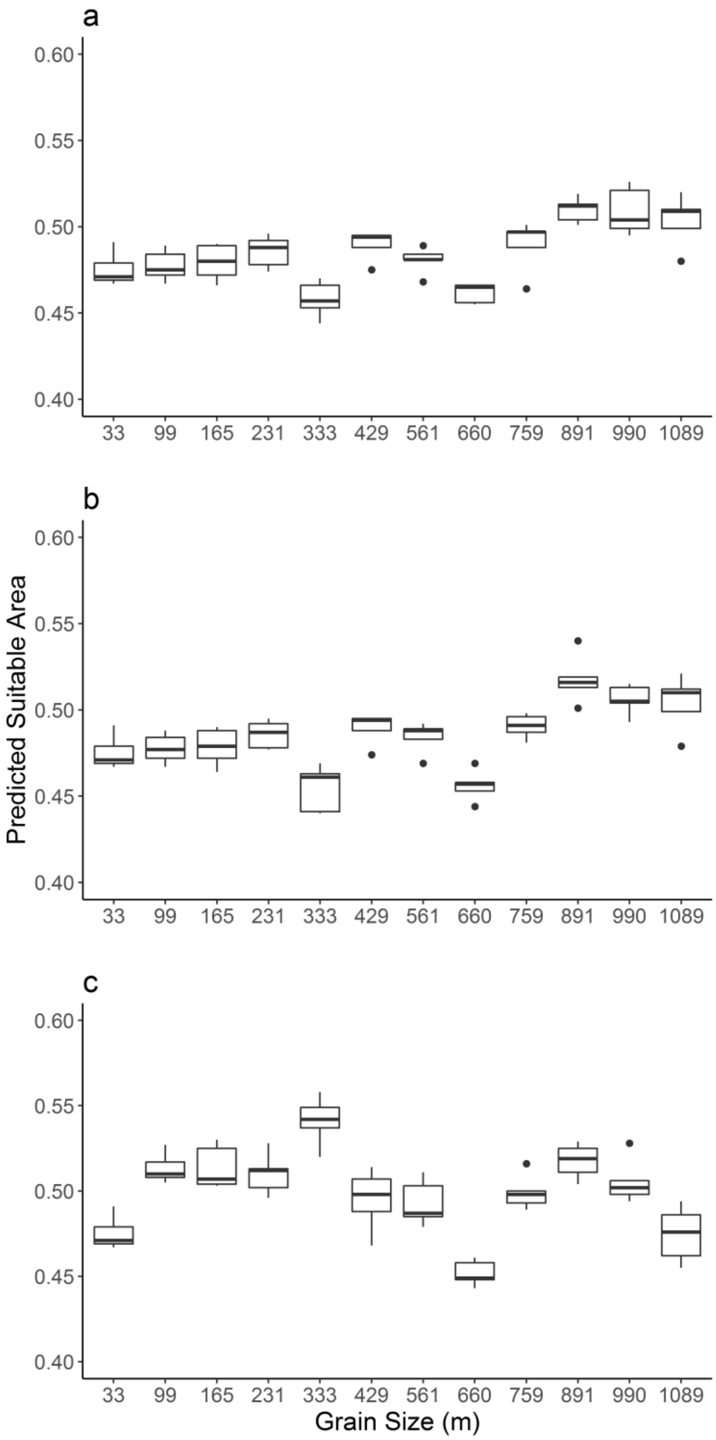


Figure 6 Boxplot of predicted area for the five Maxent replicates at each grain size for (a) nearest neighbor, (b) bilinear interpolation, and (c) aggregation.

Variable Contributions

After calculating correlation coefficients for all four variables at all 12 grain size for each of the three resampling techniques, I found that only one variable pair (WET and NDVI) showed a consistent strong relationship at all grain sizes (r = -0.659 to -0.747) for both AGG and BIL (Table 2). The correlations between LST and NDVI were intermediate (r ~ 0.700). Perhaps the most noticeable change in relationship across grain size was between LST and NDVI, which showed noticeable increases in correlation at 600m and 990m grain sizes from -0.385 to -0.725, when resampled using AGG. However, this relationship was stronger for grain sizes resampled with BIL, except for 600m and 990m. No clear trend in correlations between AGG and BIL resampling was evident, as some variable pair relationships (e.g. LST) were stronger in BIL and others in AGG, (aspect relationships).

Table 2 Correlational coefficients calculated between all environmental parameters at each grain size. Numbers in bold in bottom left of each matrix are r- values for the AGG resampling technique and numbers in grey in upper right of matrix are for BIL. NN and BIL showed similar r-values across the grain sizes so only one is shown here. Underlined numbers show r ≥ 0.700.

30 m

	ASP	LST	NDVI	WET
ASP		-0.011	0.018	-0.013
LST	**0.009**		-0.692	0.571
NDVI	**0.024**	**-0.384**		-0.716
WET	**-0.015**	**0.349**	**-0.717**	

90 m

	ASP	LST	NDVI	WET
ASP		-0.008	0.014	-0.012
LST	**0.009**		-0.690	0.574
NDVI	**0.024**	**-0.384**		-0.712
WET	**-0.015**	**0.349**	**-0.717**	

150 m

	ASP	LST	NDVI	WET
ASP		-0.006	0.011	-0.009
LST	**0.013**		-0.677	0.566
NDVI	**0.028**	**-0.385**		-0.700
WET	**-0.013**	**0.346**	**-0.727**	

210 m

	ASP	LST	NDVI	WET
ASP		-0.005	0.016	-0.007
LST	**0.019**		-0.676	0.584
NDVI	**0.046**	**-0.375**		-0.698
WET	**-0.045**	**0.353**	**-0.729**	

300 m

	ASP	LST	NDVI	WET
ASP		-0.009	0.024	-0.016
LST	**0.022**		-0.683	0.590
NDVI	**0.048**	**-0.384**		-0.698
WET	**-0.024**	**0.345**	**-0.747**	

390 m

	ASP	LST	NDVI	WET
ASP		-0.010	0.017	-0.014
LST	**0.024**		-0.682	0.584
NDVI	**0.047**	**-0.336**		-0.681
WET	**-0.023**	**0.317**	**-0.705**	

510 m

	ASP	LST	NDVI	WET
ASP		-0.004	0.017	-0.012
LST	**0.030**		-0.669	0.593
NDVI	**0.053**	**-0.344**		-0.685
WET	**-0.018**	**0.310**	**-0.729**	

600 m

	ASP	LST	NDVI	WET
ASP		0.001	0.018	-0.008
LST	**0.004**		-0.670	0.567
NDVI	**0.053**	**-0.703**		-0.669
WET	**0.001**	**0.557**	**-0.727**	

690 m

	ASP	LST	NDVI	WET
ASP		-0.008	0.014	0.002
LST	**0.039**		-0.671	0.564
NDVI	**0.053**	**-0.337**		-0.659
WET	**0.007**	**0.305**	**-0.714**	

810 m

	ASP	LST	NDVI	WET
ASP		0.002	0.010	-0.004
LST	**0.042**		-0.660	0.581
NDVI	**0.439**	**-0.340**		-0.676
WET	**-0.001**	**0.310**	**-0.746**	

900 m

	ASP	LST	NDVI	WET
ASP		-0.002	0.020	-0.012
LST	**0.442**		-0.672	0.596
NDVI	**0.050**	**-0.343**		-0.696
WET	**-0.011**	**0.312**	**-0.742**	

990 m

	ASP	LST	NDVI	WET
ASP		0.003	0.006	-0.003
LST	**0.005**		-0.667	0.566
NDVI	**0.043**	**-0.725**		-0.657
WET	**-0.011**	**0.582**	**-0.729**	

Variable contribution to model accuracy gain was strikingly similar between NN and BIL resampling techniques, with WET contributing the most to models at all grain sizes (range = 44.7% at 333m and 58.5% at 429m; Figure 7). LST was always the second most contributing variable at all grain sizes for NN and BIL resampling techniques (range 28.5% at 990m and 38.9% at 99m), followed by NDVI (range 7.1% at 99m and 21.2% at 759m) and lastly, ASP (range = 0.3% at 759m and 3.2% at 990m). Contribution of the four environmental parameters was much more varied for the AGG resampling technique, except for WET which was still generally the highest contributing variable (10 of 12 grain sizes), but showed more variation in its contribution (range = 69.5% at 99m and 19.2% at 1089m) than for either NN or BIL. For AGG grain sizes 660m and 1089m, WET was not the highest contributing variable, but rather LST and NDVI. The most change in variable contribution under AGG resampling technique occurred for NDVI, which increased its contribution to model accuracy gain with increasing grain size, where WET appeared to decrease contribution with increasing grain. Overall, variable importance was more volatile when cells were resampled with AGG than either NN or BIL.

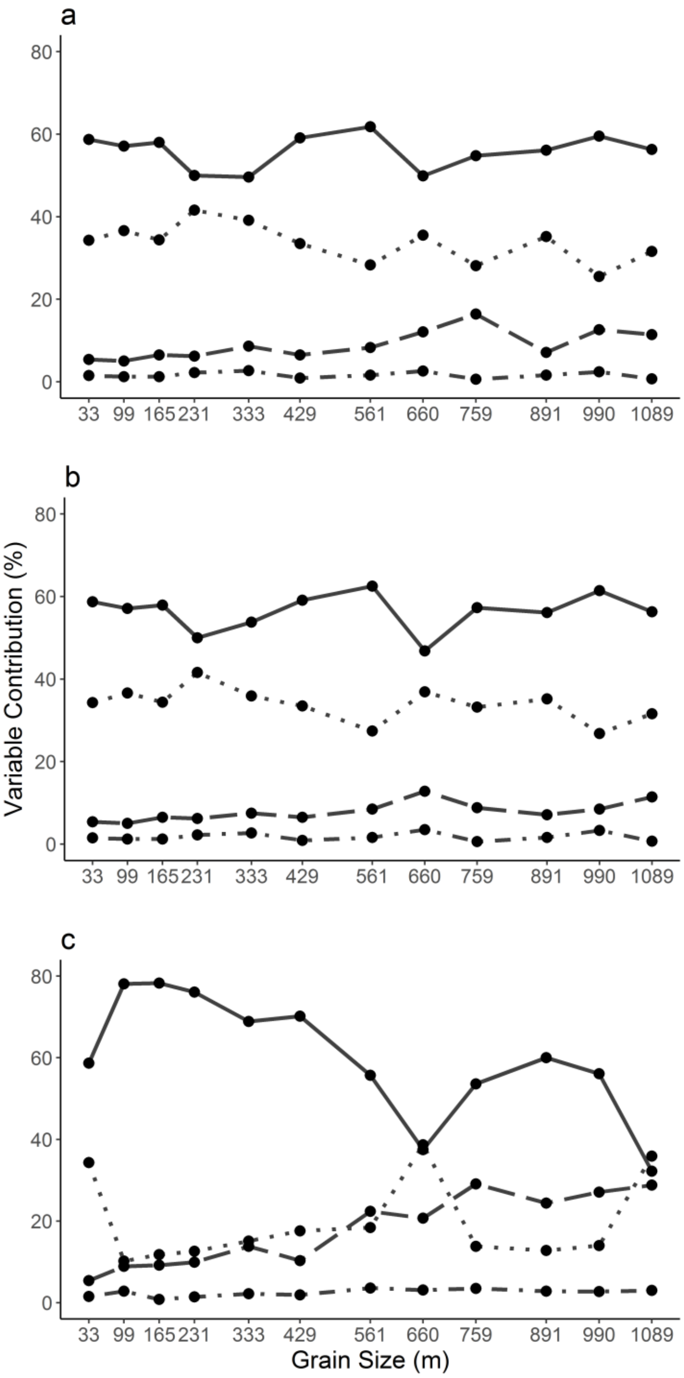


Figure 7 Average variable contribution for the five replicates for each grain size for (A) nearest neighbor, (B) bilinear interpolation, and (C) aggregation. Variables: wetness (solid line), land surface temperature (dotted line), NDVI (dashed line), aspect (dot-dashed line).

Although predicted suitable area differed slightly but significantly between resampling techniques, these differences did not change the spatial predictions (Figure 8). All models showed higher suitability for Bell’s Vireo in the northeast and southeast portions of the South Central Semiarid Prairies. For NN and BIL, model performance and variable contribution were nearly identical, thus leading to similar spatial predictions of suitability. More interestingly is the change of variable contribution in the AGG technique that still converges on a very similar suitability prediction as NN and BIL. The overall prediction of these models follow expected ecological relevancy with wetter conditions and denser vegetation increasing to the east.

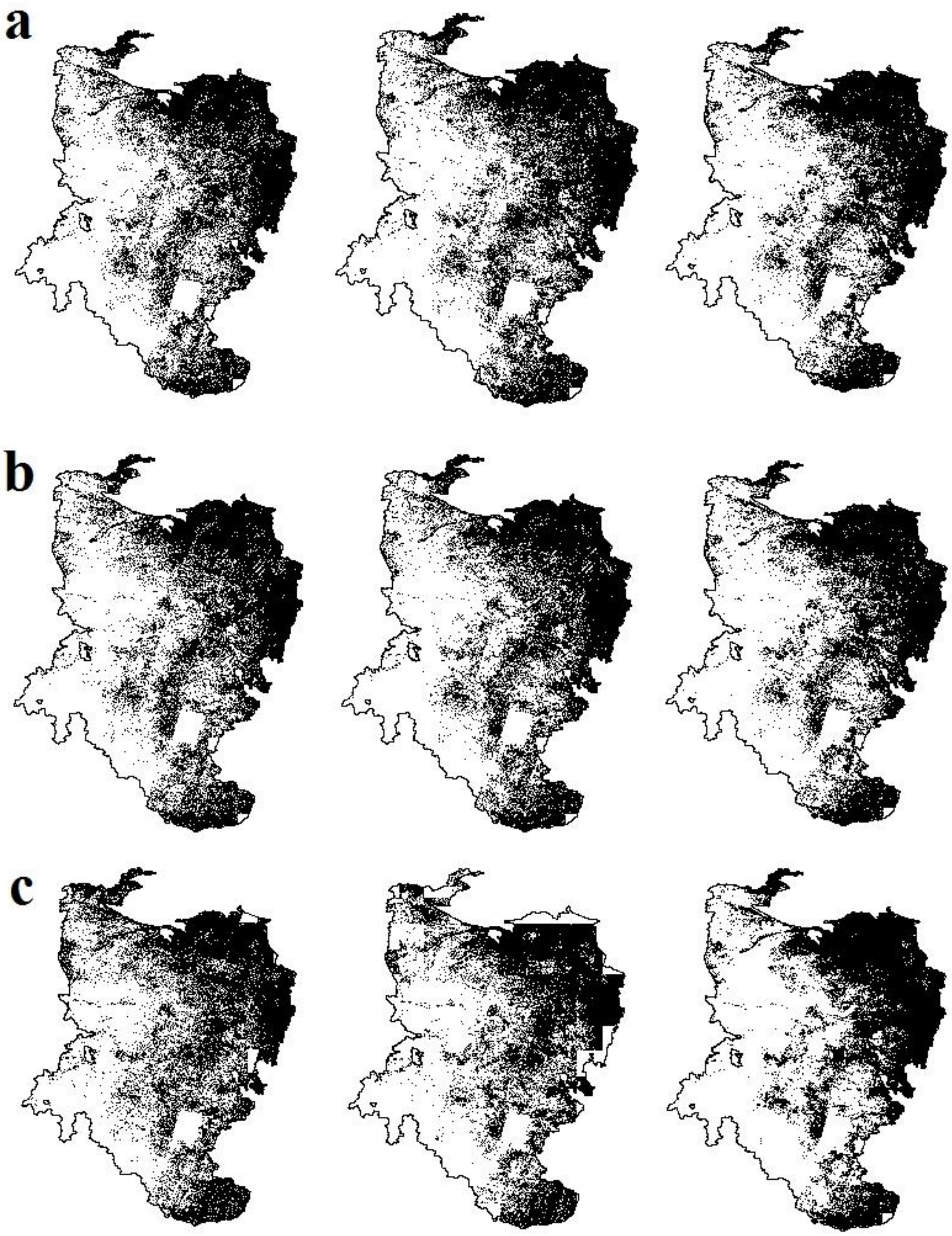


Figure 8. Thresholded (10% minimum training) Maxent predictions for Bell's Vireo using three resampling methods (a) Nearest Neighbor, (b) Bilinear Interpolation, (c) Aggregation for three grain sizes, 99 m (left), 429 m (middle) and 1089 m (right).

Discussion

A key challenge in species-environment investigations is uncovering the scale at which these relationships occur. In SDM, studies of changing grain size have varying conclusions about the effects on model performance (Seoane et al. 2004, Guisan et al. 2007a, Seo et al. 2009, Gottschalk et al. 2011, Bellamy et al. 2013, Song et al. 2013, Suarez-Seoane et al. 2014, Connor et al. 2018). Surprisingly, little quantification of environmental parameter variation has occurred in previous research considering that variation is expected to change across grain sizes, extents, and ecoregions (Wiens 1989). Additionally, no attention has been given to resampling technique, although Mertes and Jetz (2018) acknowledged that resampling changes the properties of the environmental predictors. Resampling is a common practice in SDM and little is known about how this may influence model output. Environmental variables used in SDM come in varying resolutions that are generally altered to an ecologically relevant grain or to match other predictors used in the model. Although information loss is expected with upscaling, the loss is not necessarily the same across sampling techniques (Atkinson 1988, Hay et al. 1997, Hunsaker et al. 2001). Nearest neighbor does not change values, bilinear interpolation creates new variables based on the location of a neighborhood of cells, and aggregation forms new values from summary statistics calculated on a moving window of cells. How the observed variation changes is also related to the amount of heterogeneity in the study extent and the spatial structure of the variable (Hunsaker et al. 2001).

By calculating the coefficient of variation (CV), I was able to show environmental variation changes in respect to resampling technique. Specifically, I found that the landscape or underlying ecoregion making up the background affected the variation as well as the environmental parameters used. For my study region, I chose variables with high (aspect), medium (NDVI and LST) and low (WET) variation across the region and found that variation was preserved with resampling, with the exception of aspect. This is likely due to aspect’s lack of spatial structure across the study region, whereas the remaining variables showed distinct structuring. Variation is expected to decrease with increasing grain size due to homogenization (Wiens 1989), especially with the AGG technique, but my results show that this may not be the case for some ecoregions. Both NN and BIL consider the proximity of neighboring input cells in relation to the output cell location, which should move with increasing grain. This indicates that the cells selected as the nearest neighboring should also change; creating new values for output cells at each grain size (Jensen 2005). Because both of these techniques had constant environmental variation (for three environmental parameters), this also indicates strong spatial structure structuring of these variables (i.e. a gradient or spatial autocorrelation). The adjacency of similar values would produce the observed preservation of environmental variation (for NDVI, LST, and WET) and in cases such as this, the spatial structure is likely more susceptible to variation changes with extent manipulation rather than grain (Anderson 2018). Xu et al. (2018) demonstrated that upscaled remote sensing images when compared to images collected at the larger resolution showed differing classification accuracies based on resampling techniques with varying data loss.

Generally, as the relationship of variation in background versus presence locations change, differences in model performance should occur. However, I found that even though variation within environmental parameters did not change across the resampling techniques, in some cases model performance did. Model performance measured by AUC was generally not impacted by the resampling method and showed a decreasing trend with increasing grain size. The more homogenizing effect of the AGG technique appeared to create models that did not show as much variation in AUC as did both NN and BIL. This is surprising as variable importance showed relatively large amounts of dissimilarity. Changes in model behavior such as variable importance usually suggest changes in model performance, and while AUC did not significantly change, the amount of predicted suitable area was significantly larger for the AGG resampling technique. Despite the difference in predicted suitable area across NN, BIL, and AGG, omission error did not differ overall. The significant differences in predicted suitable area at the smaller grain sizes (33m – 333m) may suggest that AGG created upscaled environmental parameters with less data gaps. As the technique simply averages a neighborhood of cells were ignored and the output would be one cell with the new averaged value. This decrease in missing data appears to not be the case for NN and BIL, which use inverse logic.

The overall subtle differences in model performance, while surprising are not unprecedented. Pradervand et al. (2013) modeled distributions for 239 plant species at six grain sizes in the Swiss Alps and found little difference in AUC. Additionally, thematic resolution was more important than spatial resolution for the distribution of several tree species in montane temperate forest in northeast China (Liang et al. 2013). The importance of scale, represented by grain, may differ between sessile and vagile organisms as studies using bats (Bellamy et al. 2013) and birds (Gottschalk et al. 2011) showed difference in model performance (both increases and decreases) as measured by AUC and were species specific. Guisan et al. (2007a), using only two grain sizes, showed the opposite that plant models were more affected by coarsening grain than birds. Both Connor et al. (2018) and Mertes and Jetz (2018) accounted for the spatial structure of study regions when investigating the influence of grain size on model accuracy. For seven virtual species, AUC decreased with increasing grain size in both heterogeneous and homogenous landscapes, although species were better modeled on heterogeneous landscapes when grain size was not considered. When the scale at which a species responds to the environment is known, larger grain sizes degrade model performance more than smaller grain sizes (relative to response grain) (Connor et al. 2018). Using 3,000 virtual species simulations, Mertes and Jetz (Mertes and Jetz 2018) found that the spatial structure of environmental variables was a large determinant in the performance of SDMs at different grain sizes. Environmental predictors exhibiting fine-scale heterogeneity performed poorly when upscaled, as many important features of the environment were lost in the aggregation process and they warn that overly coarse environmental data (in relation to species response grain) will produce spurious models that misconstrue the species-environment relationship (Mertes and Jetz 2018). Using virtual species to better understand dynamics between response grains and environments does provide insights into scaling considerations for SDM, however much work is needed to better match species response grains to environmental grains for real world applications.

My study exemplifies the difficulty in acquiring data that are both directly measured and complete, as cells with missing data may influence model performance. Remote sensing data are susceptible to clouds (Kechu and Wentao 1984, Scales et al. 2017), shadows (Ranson and Daughtry 1987), and other atmospheric effects (Morimoto et al. 1979), but on the other hand can better match the temporal resolution of occurrence data. Missing data values likely contributed to model performance overall predictions, due to upscaling. This is most evident using AGG, as certain grain sizes produced “blocky” data voids at the study area margins. Remote sensing data has been shown to improve model performance (Buermann et al. 2008, Deblauwe et al. 2016), although in some cases can introduce bias if used incorrectly (Bradley et al. 2012). However, the field of remote sensing can offer more than just data to the scale question in SDM. Woodcock and Strahler (1987) identified three factors to consider when determining the best classification scale for an image: desired output, method used to extract information, and spatial structure of the scene. These considerations are a useful guide in the development of models and echo many of the suggestions published in the SDM literature. Perhaps the biggest advantage of remote sensing data in the SDM framework is the real-time observation of the Earth’s surface allowing for more nuanced study extents and temporal investigations, as well as better capture the variation in environmental conditions (He et al. 2015).

For SDM utilizing the Maxent algorithm, an awareness of the variation in the background environment is fundamental as background conditions inform the model about prior expectations before the environmental conditions of presence data is accounted for (Merow et al. 2013). As resampling is expected to change variation (Wiens 1989) and spatial structure (Hunsaker et al. 2001) in environmental conditions, resampling methods that minimize distortion of these environmental predictor characteristics are ideal. However, the choice of resampling method is dependent upon the study extent, the variables under consideration, and the species of interest, as these act in conjunction to frame scale dependent responses of species. Since Maxent uses species locations to constrain the expected distribution, to produce a predicted species specific distribution as it conforms to what is known without the inclusion of unknown parameters (Phillips et al. 2006). Thus, accurate depictions of the study region and presence locations in terms of the environmental conditions is needed.

While this study provided a first look at how resampling may effect SDM, there are still many aspects that remain unknown; specifically, how the underlying spatial structure of environmental variables is altered. Users must recognize that decisions in the modeling process will potentially influence on model performance and overall usability in unexpected ways. When it comes to resampling, there is still much left to understand. Designing models that are specific to the species and study objectives is generally the best approach, based on these two factors, I suggest exploring resampling methods by comparing measures of background variation as these different methods are used. Additionally, multiple grain and extents should be investigated in conjunction with the biology of the species to find response scales for environmental predictors.

Conclusion

I investigated the effects of resampling environmental predictors on model performance and was able to show that for the study region, the effects were significant in one regard, predicted suitable area and not others, AUC and omission error. In addition, I tested the effects of grain size on SDM performance using a large study extent, short temporal range, and multiple remote sensing derived environmental variables. While I found results similar to previous studies, they suggest that many of these modeling parameters interact making the modeling process highly case-specific.

REFERENCES

Anderson, C. B. 2018. Biodiversity monitoring, earth observations and the ecology of scale. Ecology Letters 21:1572-1585.

Anderson, R. P., and I. Gonzalez. 2011. Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. Ecological Modelling 222:2796-2811.

Atkinson, P. 1988. The interrelationship between resampling method and information extraction technique. Pages 521-527 in International Geoscience and Remote Sensing Symposium, 'Remote Sensing: Moving Toward the 21st Century'.

Barbet-Massin, M., W. Thuiller, and F. Jiguet. 2010. How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? Ecography 33:878-886.

Barsi, J. A., J. L. Barker, and J. R. Schott. 2003. An atmospheric correction parameter calculator for a single thermal band earth-sensing instrument. Pages 3014-3016 vol.3015 in IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477).

Barsi, J. A., J. R. Schott, F. D. Palluconi, and S. J. Hook. 2005. Validation of a web-based atmospheric correction tool for single thermal band instruments. Page 7 in Optics and Photonics 2005. SPIE.

Barve, N., V. Barve, A. Jimenez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberon, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling 222:1810-1819.

Bean, W. T., L. R. Prugh, R. Stafford, H. S. Butterfield, M. Westphal, and J. S. Brashares. 2014. Species distribution models of an endangered rodent offer conflicting measures of habitat quality at multiple scales. Journal of Applied Ecology 51:1116-1125.

Bean, W. T., R. Stafford, and J. S. Brashares. 2012. The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. Ecography 35:250-258.

Bellamy, C., C. Scott, and J. Altringham. 2013. Multiscale, presence-only habitat suitability models: Fine-resolution maps for eight bat species. Journal of Applied Ecology 50:892-901.

Bennie, J., B. Huntley, A. Wiltshire, M. O. Hill, and R. Baxter. 2008. Slope, aspect and climate: Spatially explicit and implicit models of topographic microclimate in chalk grassland. Ecological Modelling 216:47-59.

Borchert, J. R. 1950. The climate of the central North American grassland. Annals of the Association of American Geographers 40:1-39.

Boria, R. A., and J. L. Blois. 2018. The effect of large sample sizes on ecological niche models: Analysis using a North American rodent, *Peromyscus maniculatus*. Ecological Modelling 386:83-88.

Bradie, J., and B. Leung. 2017. A quantitative synthesis of the importance of variables used in Maxent species distribution models. Journal of Biogeography 44:1344-1361.

Bradley, B. A., A. D. Olsson, O. Wang, B. G. Dickson, L. Pelech, S. E. Sesnie, and L. J. Zachmann. 2012. Species detection vs. Habitat suitability: Are we biasing habitat suitability models with remotely sensed data? Ecological Modelling 244:57-64.

Braunisch, V., J. Coppes, R. Arlettaz, R. Suchant, H. Schmid, and K. Bollmann. 2013. Selecting from correlated climate variables: A major source of uncertainty for predicting species distributions under climate change. Ecography 36:971-983.

Brown, J. L. 2014. SDMtoolbox: A python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. Methods in Ecology and Evolution 5:694-700.

Brown, J. L., J. R. Bennett, and C. M. French. 2017. SDMtoolbox 2.0: The next generation python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. PeerJ 5:e4095.

Buermann, W., S. Saatchi, T. B. Smith, B. R. Zutta, J. A. Chaves, B. Milá, and C. H. Graham. 2008. Predicting species distributions across the Amazonian and Andean regions using remote sensing data. Journal of Biogeography 35:1160-1176.

Burrough, P. A., and R. A. McDonell. 1998. Principles of geographic information systems. Oxford University Press, New York.

Carlson, T. N., and D. A. Ripley. 1997. On the relation between ndvi, fractional vegetation cover, and leaf area index. Remote Sensing of Environment 62:241-252.

Chave, J. 2013. The problem of pattern and scale in ecology: What have we learned in 20years? Ecology Letters 16:4-16.

Connor, T., V. Hull, A. Viña, A. Shortridge, Y. Tang, J. Zhang, F. Wang, and J. Liu. 2018. Effects of grain size and niche breadth on species distribution modeling. Ecography 41:1270-1282.

Crist, E. P. 1985. A tm tasseled cap equivalent transformation for reflectance factor data. Remote Sensing of Environment 17:301-306.

Daly, C., G. H. Taylor, W. P. Gibson, T. W. Parzybok, G. L. Johnson, and P. A. Pasteris. 2000. High-quality spatial climate data sets for the United States and beyond. Transactions of the ASAE 43:1957-1962.

Deblauwe, V., V. Droissart, R. Bose, B. Sonke, A. Blach-Overgaard, J. C. Svenning, J. J. Wieringa, B. R. Ramesh, T. Stevart, and T. L. P. Couvreur. 2016. Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. Global Ecology and Biogeography 25:443-454.

Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129-151.

Elith, J., S. J. Phillips, T. Hastie, M. Dudik, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of Maxent for ecologists. Diversity and Distributions 17:43-57.

Environmental Systems Research Institute (ESRI). 2016. ArcGIS Release 10.4. Redlands, CA.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38-49.

Franzreb, K. 1987. Endangered status and strategies for conservation of the Least Bell's Vireo. Western Birds 18:43 - 49.

Georgian, S. E., W. Shedd, and E. E. Cordes. 2014. High-resolution ecological niche modelling of the cold-water coral *Lophelia pertusa* in the Gulf of Mexico. Marine Ecology Progress Series 506:145-161.

Gottschalk, T. K., B. Aue, S. Hotes, and K. Ekschmitt. 2011. Influence of grain size on species–habitat models. Ecological Modelling 222:3403-3412.

Guisan, A., C. H. Graham, J. Elith, F. Huettmann, and N. S. Distri. 2007a. Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions 13:332-340.

Guisan, A., N. E. Zimmermann, J. Elith, C. H. Graham, S. Phillips, and A. T. Peterson. 2007b. What matters for predicting the occurrences of trees: Techniques, data, or species' characteristics? Ecological Monographs 77:615-630.

Hay, G. J., K. O. Niernann, and D. G. Goodenough. 1997. Spatial thresholds, image-objects, and upscaling: A multiscale evaluation. Remote Sensing of Environment 62:1-19.

He, K. S., B. A. Bradley, A. F. Cord, D. Rocchini, M.-N. Tuanmu, S. Schmidtlein, W. Turner, M. Wegmann, and N. Pettorelli. 2015. Will remote sensing shape the next generation of species distribution models? Remote Sensing in Ecology and Conservation 1:4-18.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965-1978.

Huang, C., B. Wylie, L. Yang, C. Homer, and G. Zylstra. 2002. Derivation of a tasseled cap transformation based on Landsat 7 at-satellite reflectance. International Journal of Remote Sensing 28:1741-1748.

Huang, C., L. Yang, C. Homer, B. K. Wylie, and J. E. Vogelmann. 2001. At-satellite reflectance: A first order normalization of Landsat 7 ETM+ images. Publications of the US Geological Survey paper 109.

Hunsaker, C. T., M. F. Goodchild, M. A. Fridel, and T. J. Case, editors. 2001. Spatial uncertainty in ecology. Springer-Verlag, New York, NY.

Jensen, J. R. 2005. Introductory digital image processing. 3rd edition. Pearson Prentice Hall, Upper Saddle River, NJ.

Jimenez-Munoz, J. C., J. Cristobal, J. A. Sobrino, G. Soria, M. Ninyerola, X. Pons, and X. Pons. 2009. Revision of the single-channel algorithm for land surface temperature retrieval from Landsat thermal-infrared data. IEEE Transactions on Geoscience and Remote Sensing 47:339-349.

Jiménez-Muñoz, J. C., and J. A. Sobrino. 2003. A generalized single-channel method for retrieving land surface temperature from remote sensing data. Journal of Geophysical Research: Atmospheres 108.

Jimenez-Valverde, A. 2012. Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. Global Ecology and Biogeography 21:498-507.

Kauth, R. J., and G. S. Thomas. 1976. The tasseled cap-a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. Laboratory for Applications of Remote Sensing Symposium 4B:41-51.

Kechu, Y., and W. Wentao. 1984. On removing cloud interference in remote-sensing images. Journal of Electronics 1:1-18.

Kerr, J. T., and M. Ostrovsky. 2003. From space to species: Ecological applications for remote sensing. Trends in Ecology & Evolution 18:299-305.

Konecy, G. 2003. Geoinformation: Remote sensing, photogrammetry and geographic information systems. Taylor and Francis, New York.

Lebbin, D. A., M. J. Parr, and G. H. Fenwick. 2010. The American bird conservatory guide to bird conservation. University of Chicago Press, Chicago.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.

Leyequien, E., J. Verrelst, M. Slot, G. Schaepman-Strub, I. M. A. Heitkönig, and A. Skidmore. 2007. Capturing the fugitive: Applying remote sensing to terrestrial animal distribution and diversity. International Journal of Applied Earth Observation and Geoinformation 9:1-20.

Liang, Y., H. S. He, J. S. Fraser, and Z. Wu. 2013. Thematic and spatial resolutions affect model-based predictions of tree species distribution. PLOS One 8:e67889.

Lobo, J. M., A. Jimenez-Valverde, and R. Real. 2008. AUC: A misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17:145-151.

Marceau, D. J. 1999. The scale issue in the social and natural sciences. Canadian Journal of Remote Sensing 25:347-356.

Martin, A. E., and L. Fahrig. 2012. Measuring and selecting scales of effect for landscape predictors in species-habitat models. Ecological Applications 22:2277-2292.

McGarigal, K., H. Y. Wan, K. A. Zeller, B. C. Timm, and S. A. Cushman. 2016. Multi-scale habitat selection modeling: A review and outlook. Landscape Ecology 31:1161-1175.

McGill, B. J. 2010. Matters of scale. Science 328:575-576.

McMahon, G., S. M. Gregonis, S. W. Waltman, J. M. Omernik, T. D. Thorson, J. A. Freeouf, A. H. Rorick, and J. E. Keys. 2001. Developing a spatial framework of common ecological regions for the conterminous United States. Environmental Management 28:293-316.

Merow, C., M. J. Smith, and J. A. Silander, Jr. 2013. A practical guide to Maxent for modeling species' distributions: What it does, and why inputs and settings matter. Ecography 36:1058-1069.

Mertes, K., and W. Jetz. 2018. Disentangling scale dependencies in species environmental niches and distributions. Ecography 41:1604-1615.

Morimoto, T., R. Kumar, and L. C. B. Molion. 1979. Effect of the atmosphere on the classification of Landsat data. Symposium on Machine Processing of Remotely Sensed Data 243.

Mouton, A. M., B. De Baets, and P. L. M. Goethals. 2010. Ecological relevance of performance criteria for species distribution models. Ecological Modelling 221:1995-2002.

National Aeronautics and Space Administration (NASA). 2002. National Imagery and Mapping Agency (NIMA), German Aerospace Center (DLR), and Italian Space Agency (ASI), Shuttle Radar Topography Mission (SRTM) Elevation Dataset: U.S. Geological Survey, Sioux Falls, SD.

Omernik, J. M. 1987. Ecoregions of the conterminous united-states. Annals of the Association of American Geographers 77:118-125.

Openshaw, S. 1984. Ecological fallacies and the analysis of areal census data. Environment and Planning 16:17-31.

Parody, J. M., and T. H. Parker. 2002. Biogeographic variation in nest placement: A case study with conservation implications. Diversity and Distributions 8:11-19.

Peterson, A. T. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3:59-72.

Peterson, A. T., J. Soberon, R. G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araujo. 2011. Ecological niches and geographic distributions. Princeton University Press, Princeton.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Phillips, S. J., and M. Dudik. 2008. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31:161-175.

Phillips, S. J., M. Dudik, J. Elith, C. H. Graham, A. Lehmann, J. Leathwick, and S. Ferrier. 2009. Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecological Applications 19:181-197.

Pradervand, J.-N., A. Dubuis, L. Pellissier, A. Guisan, and C. Randin. 2013. Very high resolution environmental predictors in species distribution models: Moving beyond topography? Progress in Physical Geography: Earth and Environment 38:79-96.

R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Radosavljevic, A., and R. P. Anderson. 2014. Making better Maxent models of species distributions: Complexity, overfitting and evaluation. Journal of Biogeography 41:629-643.

Ranson, K. J., and C. S. T. Daughtry. 1987. Scene shadow effects on multispectral response. IEEE Transactions on Geoscience and Remote Sensing GE-25:502-509.

Revermann, R., H. Schmid, N. Zbinden, R. Spaar, and B. Schroder. 2012. Habitat at the mountaintops: How long can rock ptarmigan (*Lagopus muta helvetica*) survive rapids climate change in the Alps? A multi-scale approach. Journal of Ornithology 153:891-905.

Rose, R. A., D. Byler, J. R. Eastman, E. Fleishman, G. Geller, S. Goetz, L. Guild, H. Hamilton, M. Hansen, R. Headley, J. Hewson, N. Horning, B. A. Kaplin, N. Laporte, A. Leidner, P. Leinagruber, J. Morisette, J. Musinsky, L. Pintea, A. Prados, V. C. Radeloff, M. Rowen, S. Saatchi, S. Schil, K. Tabor, W. Turner, A. Vodacek, J. Vogelnaann, M. Wegmann, and D. Wilkie. 2015. Ten ways remote sensing can contribute to conservation. Conservation Biology 29:350-359.

Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering. 1974. Monitoring vegetation systems in the Great Plains with ERTS. Proceedings of the 3rd Earth Resource Technology Satellite (ERTS) symposium 1:48-62.

Sanchez-Fernandez, D., J. M. Lobo, and O. L. Hernandez-Manrique. 2011. Species distribution models that do not incorporate global data misrepresent potential distributions: A case study using Iberian diving beetles. Diversity and Distributions 17:163-171.

Scales, K. L., E. L. Hazen, M. G. Jacox, C. A. Edwards, A. M. Boustany, M. J. Oliver, and S. J. Bograd. 2017. Scale of inference: On the sensitivity of habitat models for wide-ranging marine predators to the resolution of environmental data. Ecography 40:210-220.

Seo, C., J. H. Thorne, L. Hannah, and W. Thuiller. 2009. Scale effects in species distribution models: Implications for conservation planning under climate change. Biology Letters 5:39-43.

Seoane, J., J. Bustamante, and R. Díaz-Delgado. 2004. Are existing vegetation maps adequate to predict bird distributions? Ecological Modelling 175:137-149.

Andri Signorell et al. 2018. DescTools: Tools for descriptive statistics. R package version 0.99.25.

Soberón, J. M. 2010. Niche and area of distribution modeling: A population ecology perspective. Ecography 33:159-167.

Sobrino, J. A., V. Caselles, and F. Becker. 1990. Significance of the remotely sensed thermal infrared measurements obtained over a citrus orchard. ISPRS Journal of Photogrammetry and Remote Sensing 44:343-354.

Sobrino, J. A., J. C. Jiménez-Muñoz, and L. Paolini. 2004. Land surface temperature retrieval from Landsat TM 5. Remote Sensing of Environment 90:434-440.

Sobrino, J. A., and N. Raissouni. 2000. Toward remote sensing methods for land cover dynamic monitoring: Application to morocco. International Journal of Remote Sensing 21:353-366.

Song, W., E. Kim, D. Lee, M. Lee, and S.-W. Jeon. 2013. The sensitivity of species distribution modeling to scale differences. Ecological Modelling 248:113-118.

Suarez-Seoane, S., E. Virgos, O. Terroba, X. Pardavila, and J. M. Barea-Azcon. 2014. Scaling of species distribution models across spatial resolutions and extents along a biogeographic gradient. The case of the Iberian mole *Talpa occidentalis*. Ecography 37:279-292.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. Ebird: A citizen-based bird observation network in the biological sciences. Biological Conservation 142:2282-2292.

Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. Science 240:1285-1293.

Turner, W., S. Spector, N. Gardiner, M. Fladeland, E. Sterling, and M. Steininger. 2003. Remote sensing for biodiversity science and conservation. Trends in Ecology & Evolution 18:306-314.

Urban, D. L. 2005. Modeling ecological processes across scales. Ecology 86:1996-2006.

Vale, C. G., P. Tarroso, and J. C. Brito. 2014. Predicting species distribution at range margins: Testing the effects of study area extent, resolution and threshold selection in the Sahara-Sahel transition zone. Diversity and Distributions 20:20-33.

Valor, E., and V. Caselles. 1996. Mapping land surface emissivity from ndvi: Application to European, African, and South American areas. Remote Sensing of Environment 57:167-184.

Van De Griend, A. A., and M. Owe. 1993. On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. International Journal of Remote Sensing 14:1119-1131.

Walawender, J. P., M. Szymanowski, M. J. Hajto, and A. Bokwa. 2014. Land surface temperature patterns in the urban agglomeration of Krakow (Poland) derived from Landsat-7/ETM+ data. Pure and Applied Geophysics 171:913-940.

Wheatley, M., and C. Johnson. 2009. Factors limiting our understanding of ecological scale. Ecological Complexity 6:150-159.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.

Wisz, M. S., R. J. Hijmans, J. Li, A. T. Peterson, C. H. Graham, A. Guisan. 2008. Effects of sample size on the performance of species distribution models. Diversity and Distributions 14:763-773.

With, K. A., R. H. Gardner, and M. G. Turner. 1997. Landscape connectivity and population distributions in heterogeneous environments. Oikos 78:151-169.

Woodcock, C. E., and A. H. Strahler. 1987. The factor of scale in remote sensing. Remote Sensing of Environment 21:311-332.

Wu, J., K. B. Jones, H. Li, and O. L. Loucks, editors. 2006. Scaling and uncertainty analysis in ecology: Methods and applications. Springer, Netherlands.

Xu, K., Q. Tian, Y. Yang, J. Yue, and S. Tang. 2018. How up-scaling of remote-sensing images affects land-cover classification by comparison with multiscale satellite images. International Journal of Remote Sensing:1-27.

Yackulic, C. B., and J. R. Ginsberg. 2016. The scaling of geographic ranges: Implications for species distribution models. Landscape Ecology 31:1195-1208.

Young, N. E., R. S. Anderson, S. M. Chignell, A. G. Vorster, R. Lawrence, and P. H. Evangelista. 2017. A survival guide to Landsat preprocessing. Ecology 98:920-932.

CHAPTER IV

SEASONALITY OF ENVIRONMENTAL CONDITIONS CONTRIBUTES TO DIFFERENCES IN MODEL PERFORMANCE AND ACCURACY OF ESTIMATED POTENTIAL DISTRIBUTIONS OF SPECIES

Abstract

The interactions between species and their environment occur across spatial and temporal scales. These are not always independent influences as they often interact to shape species' responses to the environment. Temporally, species exhibit a wide range of interactions with their environment occurring across varying time frames. Since many species exhibit phenology, changes in ecological relationships across seasons, understanding seasonal influences on species distributions is paramount. Using seasonal variables in the species distribution modeling (SDM) framework is rare, and usually focuses on species level responses to seasons. However, populations within a single species can show variation in phonological responses to the environment. I use Bell's Vireo, a species comprised of two distinct populations (western and eastern), which exhibit distinct difference in breeding season habitat use, as a case study to investigate the inclusion of seasonally explicit environmental predictors in SDM. Specifically, I compared seasonal and annual climatic variables expected to shape available habitat internally and externally for the breeding ranges of both populations. Additionally, I projected these models

under two emission scenarios (4.5 and 8.5) for three future time periods (2020, 2050, and 2080). I found differences between the two populations in climatic conditions at the breeding range locations compared to wintering locations, suggesting that within the species, the populations are both utilizing similar but also different climatic conditions from wintering areas. Further, the inclusion of seasonal variables did not necessarily improve models for both populations. However, the use of seasonal variables did significantly show increased proportions of suitable area compared to annual variables for all models and future projections.

Introduction

The distribution of organisms across landscapes is the expression of natural selection and adaptions to environmental conditions (Morrison et al. 2006). However, fully understanding species-habitat relationships requires the consideration of scale (Wiens 1989, Levin 1992). Species have distinct responses to environmental conditions at different scales (Wiens 1989, Fisher et al. 2011, Thornton and Fletcher 2014). Spatial scale consists of two separate but linked components: extent (geographic region under analysis) and grain (spatial size of observation). Changes in extent and grain modify perceived ecological phenomena, by constraining or encompassing variation. Independently, increases in extent incorporate more spatial variation across the study extent, whereas increases in grain decrease variability across the study extent; although, the spatial structuring of ecological phenomena may determine the degree to which variation is altered. For example, highly structured environmental conditions have values located near similar values (e.g. temperature gradient) and changes in grain or extent may not necessarily change observed variation if grain values are resampled with similar surrounding values (Mertes and Jetz 2018).

In addition to differential responses of species to the spatial scale of environmental variables, species are also influenced by the temporal scale of environmental conditions (Wiens 1989).

Temporal scale has similar attributes of spatial scale, such as an extent (length of time under observation) and grain (the frequency of observations) to which species will uniquely respond (Schneider 1994, Morrison et al. 2006). The temporal variation of environmental variables can exhibit structuring, such as temporal autocorrelation, which may shape species' response (Moloney and Levin 1996, Metcalf and Koons 2007, Meyer et al. 2007). Interactions between spatial and temporal scales configure the ecological processes to which species react, making predictions about species habitat relationship difficult.

Species distribution modeling (SDM) is a widespread technique utilized in predicting species distributions to further our understanding of species-habitat relationships (Guisan and Thuiller 2005). In short, SDM uses geographical locations of species to extract environmental conditions associated with presences to predict new environmentally suitable areas across geographic extents. The perceived response of species to environmental conditions in SDM is highly dependent on the extent and grain of environmental explanatory variables (Guisan et al. 2007, Seo et al. 2009, Barve et al. 2011, Hanberry 2013, Song et al. 2013, Bean et al. 2014, Suarez-Seoane et al. 2014, Vale et al. 2014, Connor et al. 2018, Mertes and Jetz 2018). Although temporal scales are expected to influence SDM, less work has quantified this effect (Reside et al. 2010, Pickens and King 2014, Fernandez et al. 2017, Hereford et al. 2017, Bonthoux et al. 2018).

Most variables available for SDM are either interpolated surfaces (e.g. WorldClim (http://www.worldclim.org), PRISM (http://prism.oregonstate.edu)), or directly measured remote sensing surfaces (e.g. normalized difference vegetation index, land surface temperature), each of which exhibits tradeoffs in their use. For example, interpolated surfaces often have long temporal ranges but have data for all cells/pixels within the modeled geographic extent, whereas remote sensing variables can offer short temporal ranges but may exhibit missing data due to cloud obstruction or sensor malfunctions. Generally, the availability of environmental data strongly influences decisions on the scale of environmental predictors used in SDM.

The incorporation of temporally explicit predictors in SDM is an important consideration as environmental conditions can exhibit variation over multiple time-periods (e.g. hourly, daily, monthly). Temporal variation in environmental predictors can be more important for some organisms than others, such as sessile and vagile or resident and migratory species. Specifically, long temporal variables imply that the conditions are suitable long-term in an area where a species presence location occurs (Reside et al. 2010), which is not always the case. Pickens and King (2014) demonstrated the importance of multi-temporal variables in habitat models for systems, such as wetlands, which are temporally dynamic. For oceanic environments, Fernandez et al. (2017) acknowledged that low-frequency temporal data did not always reproduce the variation found at higher temporal frequency, resulting in misleading predictions. Further, temporal effects may have distinct implications for a species, as coarse predictions of suitability for the Mediterranean blue mussel were not influenced by temporal scale; whereas the reproductive dynamics were highly depend on the temporal scale of data (Montalto et al. 2014). Bonthoux et al. (2018) found that single date NDVI models performed better than multi-data models when the appropriate time period was used. Similarly, short-term weather models performed better than long-term climate models for many Australian birds, of which some were highly mobile and others more sedentary, indicating the influence of weather variation is important and likely to influence species under climate change (Reside et al. 2010).

Temporal variation is not limited to time ranges, but also the time of year, as seasonality plays a large role in ecological phenomena (Levin 1992). Seasonal habitat use has been shown for numerous taxa including insects (Jacobi and Cary 1996), fish (Johnson et al. 2016), reptiles (García 2008), birds (McClure et al. 2013, Varner et al. 2014), and mammals (Milakovic et al. 2012, Fynn et al. 2014). Further, Smeraldo et al. (2018) was able to show that models which consider a species phenology (i.e. seasonally variable ecological requirements) better estimate a species' niche or distribution. This is especially true for migratory species like many bird species, which exhibit differential habitat use for breeding and wintering seasons (Engler et al. 2014). Habitat use during the breeding season is thought to fall into one of two categories: “niche-followers” that utilize similar environmental conditions in wintering and breeding, and “niche-switchers” that utilize distinct environments between the two seasons (Nakazawa et al. 2004, Engler et al. 2017). Additionally, some species show high variation in breeding season environmental selection (Parody and Parker 2002), further complicating the temporal aspect of the species-habitat relationship. Since migratory bird abundance is proportionally related to the amount of suitable breeding habitat available (Sherry and Holmes 1993), the need to decipher spatio-temporal influences on migratory bird distributions is crucial.

I used Bell’s Vireo (*Vireo bellii*), a Neotropical migratory songbird, as my study system as the species exhibits two distinct population clusters (west and east) inhabiting environmentally distinct areas. The western population (hereafter BEVI-west) consists of two subspecies *V.b. arizonae*, and *V.b. pusillus*, which are associated with cottonwood-willow dominated riparian vegetation within southern California and Arizona (Franzreb 1987). The eastern population (BEVI-east) comprised of *V.b. bellii*, *V.b. medius* is affiliated with dene shrubby vegetation in a variety of landscapes such as grasslands and shrublands (Budnik et al. 2000) and demonstrates a more generalist approach to habit selection (early-successional dense grassland-shrubland areas) compared to BEVI-west (riparian areas). Habitat selection is expected to differ between the two populations as Parody and Parker (2002) were able to measure variation in the environmental features of nest sites across the two populations of Bell’s Vireo. A recent genetic analysis from Klicka et al. (2016) indicates that the west-east population divide, occurring near the Arizona/New Mexico border, actually delineates two separate species, each encompassing two sub-species. Thus, Bell’s Vireo provides an ideal opportunity to investigate the influence of seasonal versus annual climatic variables in structuring the distribution of a species (or very closely related two species), which selects habitat distinctive from background (e.g. riparian areas) in one geographic region and habitat exhibiting only subtle differences from the background (e.g. shrubby grasslands) in a different region.

I use climatic variables as proxies for habitat variables, since climatic conditions shape habitat at broad and fine scales (McGill 2010). My objectives were three-fold. First, I explored the use of seasonal and annual variables in SDM, to describe suitable habitat for the two Bell's Vireo population. Secondly, I investigated if study area extent changed the contribution of seasonal and annual variables. Lastly, I tested the performance of seasonal and annual variables to predict distributions into the future under two greenhouse gas emission scenarios.

Methods

Occurrence Data

I downloaded Bell's Vireo occurrence locations from several and to some degree overlapping sources, specifically I acquired 292,333 locations from eBird (eBird 2017) (Sullivan et al. 2009), 105,128 records from GBIF (GBIF 2018) (Telenius 2011), and 97,175 records from BISON (BISON 2018). I divided the records into two temporal data sets, pre-1990 and 2011-2018. For both temporal data sets, I identified BEVI-west occurrences from BEVI-east using two criteria: 1) location, by which I kept records that were located within the known breeding range of each population and 2) sub-species identification, for the area where overlap in the two populations' breeding ranges occurs; I only retained records that included a sub-species identification. I further filtered occurrences by the precision of the coordinates, only keeping records with at least two significant digits (~ 1 km) for both longitude and latitude. I removed records that were associated with atlas or gridded sampling (e.g. 10 min block sampling) and records associated with high locational uncertainty (e.g. Breeding Bird Survey, county centroids) or traveling checklists ≥ 0.25 km. Lastly, I removed duplicate records and occurrences located within the same data cell. Using the same filtering criteria, I complied a data set of historical wintering Bell's Vireo occurrences (pre-1990), BEVI-winter. After filtering occurrence locations for both Bell's Vireo populations, my four data sets were BEVI-west training which encompassed all pre-1990 occurrence locations (85), BEVI-west testing which included locations collected between 2011 and 2018 (1,006), BEVI-east training (73), and BEVI-east testing (1,968). I used the final occurrence data sets for both BEVI-west and BEVI-east to create convex polygons with 25 km buffers representing the breeding range training regions; the BEVI-west range encompassed 79,964 km2 and BEVI-east covered 2,032,780 km2 (Figure 1).

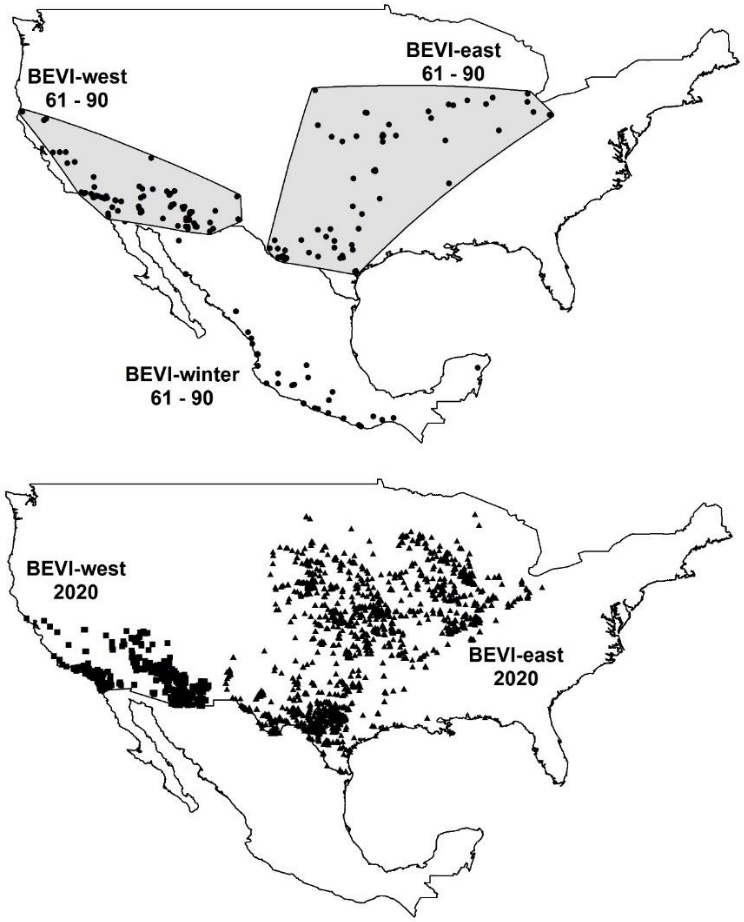


Figure 1 Top: Pre-1990 occurrence locations for BEVI-west, BEVI-east, and BEVI-winter. Breeding range polygons are shown in grey for both populations. Bottom: Independent testing data (2011 – 2018) BEVI-west (squares) and BEVI-east (triangles) used for 2020 model projections.

Environmental Data

I used six 1 km variables acquired from the ClimateNA dataset (Wang et al. 2016); three annual variables, annual heat moisture (AHM), mean annual temperature (MAT), mean annual precipitation (MAP) and three seasonal (breeding) variables, summer heat moisture (SHM), summer average temperature (TAV-S) and summer precipitation (PPT-S). I chose AHM and SHM since these variables represent the ratio between the warmest temperature and mean precipitation, annually and seasonally respectively. Higher values of either variable indicate hotter and drier conditions with lower values indicating cooler and wetter conditions. I also included temperature and precipitation, as these are highly variable annually and seasonally and were expected to differ between the breeding ranges of the two Bell's Vireo populations. I downloaded historical data (1961-1990) based on the CRU-TS 3.22 data set (Mitchell and Jones 2005) and future climate projections for three 30 year time periods 2011-2030 (hereafter 2020), 2041-2070 (2050), and 2071-2100 (2080) using an ensemble projection of 15 Atmosphere-Ocean General Circulation Models (AOGCM) (Knutti et al. 2013). General Circulation Models (GCMs) are designed to model the climate system by incorporating multiple properties of the atmosphere (physical, chemical, and biological), atmosphere-ocean interactions, ocean properties, and land/ice surface properties (Harris et al. 2014). It is important to acknowledge that these models do not predict future climate conditions; rather they describe climatic conditions under a suite of possible future climate forcing scenarios (Weaver et al. 2013). GCMs use future emission scenarios from four Representative Concentration Pathways (PCPs) (2.6, 4.5, 6.0, and 8.5) describing different policies ranging from carbon emission mitigation (2.6) to no mitigation policy (8.5) (Harris et al. 2014).

I clipped both the annual and seasonal variables (1961 – 1990) to two training extents, the breeding range for each population (BEVI-west and BEVI-east) and for the United States and Mexico combined (US-MEX). I extracted variable values from the 61-90 US-MEX environmental dataset to BEVI-west, BEVI-east, and BEVI-winter to compare the distribution of annual and seasonal variables. I created density plots for breeding and wintering ranges for each variable.

Further, I calculated basic statistics, minimum, maximum, mean, and standard deviation for each variables for each training region using ArcMap (ESRI 2016). I then calculated the coefficient of variation (standard deviation /mean) to describe the heterogeneity in the environmental predictors for the different training regions.

Species Distribution Modeling

I used the presence-background algorithm Maxent (v 3.4.1) to generate annual and seasonal models for the two populations of Bell's Vireo. Maxent defines constraints from the combination of environmental data associated with presence and background locations to fit a distribution that maximizes entropy across a geographic extent (Elith et al. 2011). Locations within the geographic extent that have conditions similar to those at species' presence locations are assigned higher probabilities of suitability. The choice of background is a crucial since it informs the shape of the probability distribution: selecting from too small a background can produce overfit models whereas too large a background can create oversimplified models (VanDerWal et al. 2009). As such, I used two methods to define the geographic extent from which background sampling would occur. In the first approach, I used the polygons depicting the breeding ranges for each population to delineate the extent of background sampling. Models created with this approach were then projected to US-MEX extent. For the second approach, I created models at the extent of US-MEX. I used both methods to create models with annual variables and seasonal variables from 1961-1990 for BEVI-west and BEVI-east. Historic occurrence data sets were separated into training and testing data using cross validation and three replicates. Since the extent size for background sampling differed, I constrained the number of background samples to 1% of the training extents. To convert Maxent's continuous suitability output to binary predictions, I used the 10% training presence threshold, which excludes 10% of the training data with the lowest suitability predictions to create a more restricted prediction (Radosavljevic and Anderson 2014). I took these "historic" models and projected them to the three future time spans: 2020, 2050, and 2080. For projected 2020 models, I used the 2011-2018 occurrence data sets as independent testing data.

Model Evaluation

I evaluated model performance using three criteria: the area under the curve of the receiver operating characteristic plot (ROC AUC), testing omission error, and predicted suitable area. Generally, ROC AUC is used to compare models by using a plot of sensitivity and 1-specificity. For SDM, this plot is modified to sensitivity and proportion of area predicted suitable (Phillips et al. 2006, Elith et al. 2011). AUC values > 0.7 are considered acceptable (Swets 1988, Fielding and Bell 1997), with higher AUC values indicating higher discriminatory power. Since AUC is sensitive to model parameters (Peterson et al. 2008, Hanczar et al. 2010), I considered model performance indicated by AUC in conjunction with omission error and predicted suitable area (Lobo et al. 2008). Omission error is the percentage of known locations not classified as presences by the model, with lower omission error rates indicating higher model performance (Mouton et al. 2010). Lastly, I also calculated the proportion of area predicted suitable to evaluate model performance, from thresholded binary suitability maps. Relatively lower proportions indicate model overfitting, where the model is unable to predict suitable areas outside the presence locations, whereas relatively large proportions of suitable area indicates poor model discrimination such that larger areas must be predicted to include the presence locations. I compared omission error across 61-90 models and for 2020 models. I used the omission error rates calculated by Maxent, for 61 – 90 models; however, I calculated the 2020 omission error rate. For each 2020 model, I took the three replicate thresholded predictions and created a composite prediction by adding them together and selecting the area where all three showed agreement. Using this composite prediction, I calculated the omission error rates using 2011-2018 test data.

I also evaluated variable contribution for seasonal and annual variables since Maxent calculates the percent contribution of each variable to model based on training gain. As I ran three replicates, Maxent calculated the average percent contribution for variables across the three replicates. I used this output to compare variables across training regions and temporal range. Lastly, I extracted values for all variables based the pre-1990 breeding occurrence locations of both populations (BEVI-west and BEVI-east) as well as pre-1990 wintering range as well as winter seasonal variables (temperature and precipitation). I created density plots in R using the ggplot package (Wickham 2016) to compare the distribution of environmental conditions encountered by Bell’s Vireos in their wintering range and breeding ranges.

Results

Environmental Data Variation across Ranges

After calculating the mean and standard deviation at all time periods (61-90, 2020, 2050, and 2080) and RCPs (4.5 and 8.5) for seasonal variables at the two occurrence dependent training ranges, BEVI-west and BEVI-east, as well as for the larger training range (US-MEX), I determined the coefficient of variation (CV), see Table 1. Annual variables exhibited more disparity in CV values than seasonal variables. For annual variables, AHM showed the most variation across all training ranges and time periods with most for US-MEX (61-90) and the least for BEVI-east. Generally, MAP showed higher variation than MAT, although the difference was dependent on training range as well as time period and RCP. Similar to annual variables, SHM showed the CV values as well as the most variation in CV, with PPT-S having the second highest CV values followed by TAV-S.

Table 1 Coefficients of variation for both annual variables (AHM, MAT, MAP) and seasonal variables (SHM, TAV-S, PPT-S) at each of the three training regions (BEVI-west, BEVI-east, and US-MEX 61-90) as well as for each projection (US-MEX 2020, 2050, 2080) and RCP (4.5 and 8.5). CV values over 1 indicate high variance and values lower than 1 indicate less variation.

	Breeding Range		US-MEX						
	BEVI-west	BEVI-east	61-90	2020 (4.5)	2020 (8.5)	2050 (4.5)	2050 (8.5)	2080 (4.5)	2080 (8.5)
AHM	0.808	0.467	1.036	1.020	1.004	0.995	1.003	1.003	0.993
MAT	0.310	0.251	0.495	0.427	0.421	0.391	0.366	0.372	0.319
MAP	0.843	0.313	0.608	0.605	0.605	0.603	0.607	0.605	0.602
SHM	0.957	0.286	1.719	1.669	1.651	1.623	1.603	1.641	1.554
TAV-S	0.200	0.087	0.197	0.178	0.177	0.169	0.160	0.162	0.144
PPT-S	0.865	0.213	0.705	0.690	0.697	0.679	0.685	0.678	0.648

Unsurprisingly, all US-MEX variables showed higher variation than either BEVI-west or BEVI-east. US-MEX at time period 61-90 showed the higher CV than any other period of time or RCP scenario. Interestingly, BEVI-east always showed the least amount of variation for both annual and seasonal variables, despite covering an area significantly larger than BEVI-west. Additionally, all three seasonal variables showed lower variation than annual variables for BEVI-east, whereas two seasonal variables, SHM and PPT-S, showed slightly higher variation than annual variables for BEVI-west.

Density plots (Figure 2) showed that annual and seasonal variables differed across the two breeding ranges (BEVI-west, green and BEVI-east, red) and the wintering range (BEVI-winter, blue). For the three annual variables, BEVI-winter generally showed a high density of “extreme” conditions, this is not surprising as temperature and precipitation are generally higher at lower latitudes. The BEVI-west population showed higher frequency use of locations with more moderate annual temperatures as well as drier conditions than BEVI-east. Seasonally, BEVI-west and BEVI-east populations used locations with similar temperature regimes, but relatively drier locations were used by BEVI-west. Summer seasonal temperature and precipitation in the wintering location for Bell’s Vireo were higher than the seasonal variables experienced by each population in its respective breeding range. However, when I compared the temperature and precipitation from breeding season and wintering season localities, I found differences between the two populations (Figure 3). Specifically, temperature was relatively similar across the three geographic locations, but BEVI-east exhibited a preference for locations with higher precipitation than found in its wintering locations.

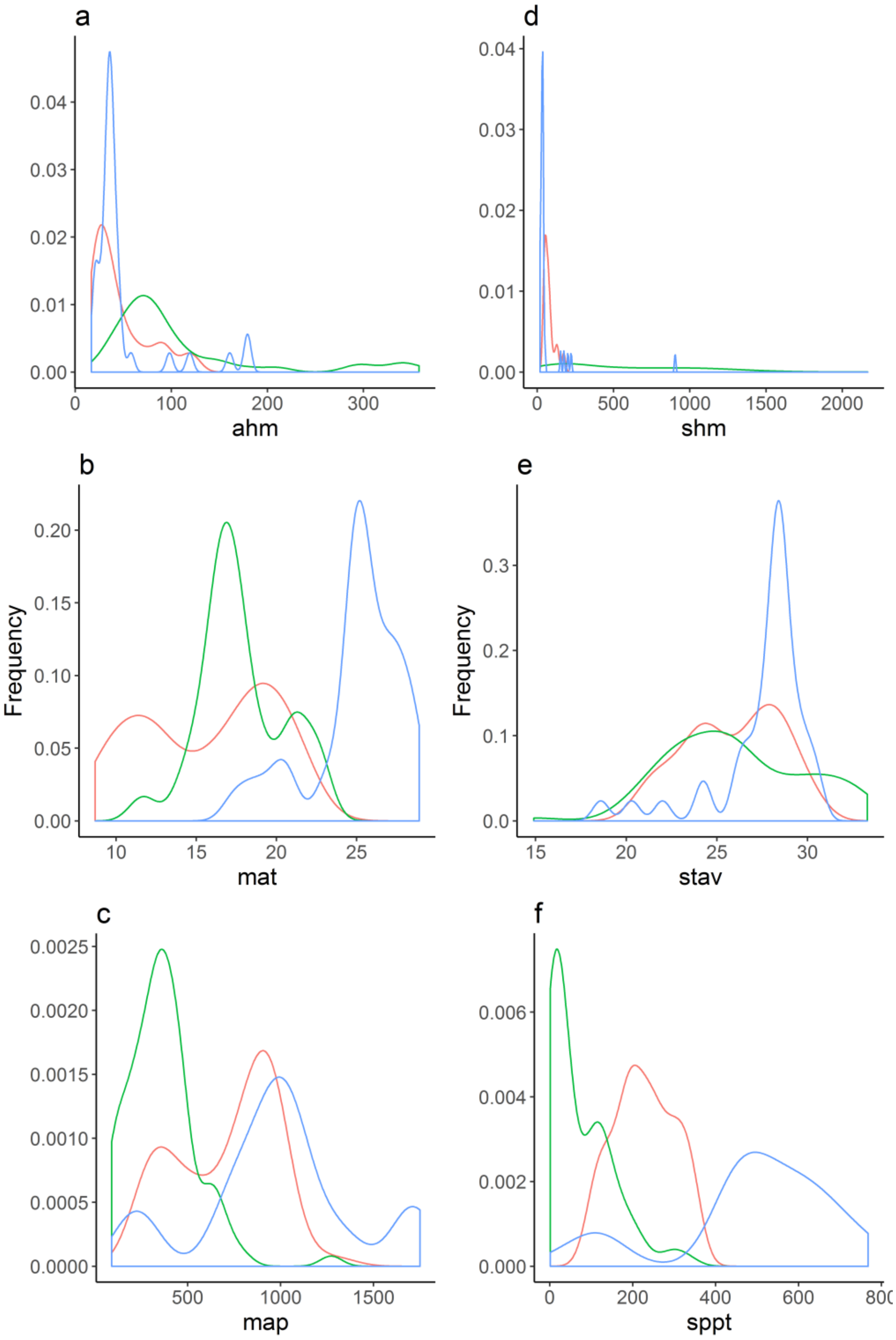


Figure 2. Density plots for annual variables (a – c) and seasonal variables (d – f) for BEVI-west (green), BEVI-east (red), and BEVI-winter (blue). Annual heat moisture (ahm), mean annual temperature (mat), mean annual precipitation (map), summer heat moisture (shm), summer average temperature (stav), and summer precipitation (sppt).

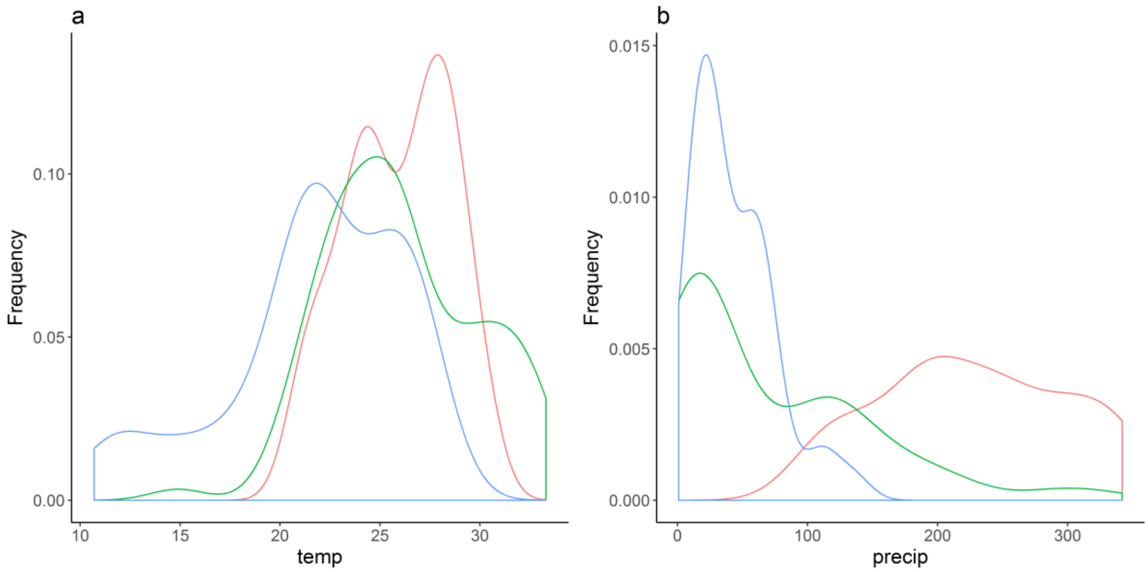


Figure 3. Density plots for (a) temperature and (b) precipitation for BEVI-west (green) and BEVI-east (red) breeding ranges (summer seasonal variables) and BEVI-winter (blue) locations (winter seasonal variables).

Model Performance and Variable Contribution to Models

All BEVI-east models showed lower AUC (range = 0.661 – 0.842), and higher omission error (0.123 – 0.221) compared to BEVI-west (AUC and omission error range, 0.713 – 0.928 and 0.117 – 0.165, respectively; see Table 2). Models trained on the breeding range for both BEVI-west and BEVI-east performed worse than models trained at US-MEX. The difference in AUC between these models was more pronounced for BEVI-east (ΔAUC = 0.154) compared to BEVI-west (ΔAUC =0.104). Seasonal models showed slightly lower AUC performance for both populations when compared to annual models, with the exception of BEVI-west trained at US-MEX. Models for the eastern population of Bell’s Vireo showed higher omission error than the western population did, although all omission error was ≤ 0.221.

Table 2. Model performance metrics, AUC, omission error (OE), and predicted suitable area (PSA), as well as variable contribution for annual and seasonal models for both populations of Bell's Vireo. All replicates are numbered (i.e. Rep 1) with the means for all three replicates underlined.

	Annual Models															
	BEVI-west breeding				BEVI-west US-MEX				BEVI-east breeding				BEVI-east US-MEX			
	Rep 1	Rep 2	Rep 3	Mean	Rep 1	Rep 2	Rep 3	Mean	Rep 1	Rep 2	Rep 3	Mean	Rep 1	Rep 2	Rep 3	Mean
AUC	0.799	0.768	0.831	0.799	0.899	0.921	0.890	0.903	0.734	0.695	0.628	0.686	0.844	0.782	0.901	0.842
OE	0.069	0.250	0.071	0.130	0.310	0.000	0.179	0.163	0.24	0.125	0.125	0.163	0.08	0.5	0.083	0.221
PSA	0.446	0.368	0.457	0.424	0.143	0.189	0.151	0.161	0.558	0.66	0.668	0.629	0.335	0.233	0.277	0.282
AHM	0.1	1.6	0.7	0.8	50.1	59.7	61.6	57.2	44.8	54.9	18.9	39.5	16.8	19.5	14.5	17
MAT	78.6	74.3	74.7	75.9	48.0	37.6	37.2	40.9	13.3	28.9	32.5	24.9	65.7	48.5	62.6	58.9
MAP	21.3	24.1	24.6	23.3	1.9	2.7	1.2	1.9	41.9	16.2	48.6	35.6	17.5	31.9	22.9	24.1
	Seasonal Models															
AUC	0.756	0.659	0.723	0.713	0.926	0.958	0.900	0.928	0.721	0.58	0.683	0.661	0.801	0.806	0.857	0.821
OE	0.103	0.250	0.143	0.165	0.138	0.000	0.214	0.117	0.08	0.333	0.083	0.165	0.16	0.208	0	0.123
PSA	0.481	0.444	0.523	0.483	0.109	0.315	0.104	0.176	0.694	0.692	0.748	0.711	0.38	0.367	0.402	0.383
SHM	8.1	8.5	7.9	8.2	64.9	77.2	79.6	73.9	69.6	65.2	78	70.9	20	13.2	16.4	16.6
TAV-S	47.3	54.1	50.7	50.7	13.0	15.9	17.3	15.4	15.9	20.4	5.3	13.9	59	66.7	59.6	61.7
PPT-S	44.6	37.4	41.4	41.1	22.1	6.9	3.1	10.7	14.4	14.4	16.7	15.2	21.1	20.1	24	21.7

In addition to lower AUC values and higher omission error (Table 2), BEVI-east models exhibited noticeably more predicted suitable area compared to BEVI-west models. The amount of predicted suitable area for breeding range trained models was larger than US-MEX trained models, in some cases over twice as large. This is was also the case for seasonal variables compared to annual variables, although the difference in the amount of area predicted suitable was only slightly higher for seasonal models.

Variable contribution was highly dependent on training range, and differed dramatically between populations (Table 2). For both east and west Bell's Vireo populations, the highest contributing variables for breeding range models were the lowest contributing variable for US-MEX models. For example, AHM contributed most to BEVI-east models and added little to model discrimination for BEVI-west, however the opposite is true for US-MEX models. I found that seasonal variables showed the same trend.

Model Projections

Omission error (calculated with 2011-2018 data) for models projected to 2020 was lower for BEVI-west models than BEVI-east models in all cases (Table 3). For BEVI-east, seasonal models showed lower omission error (1.73% - 24.14%) than annual models (33.38% - 39.43%). However, for BEVI-west, annual models exhibited better model prediction with lower omission error rates (3.88% - 5.77%) compared to seasonal models (9.94% - 16.00%). Additionally, models trained at the breeding range did a poorer job of predicting BEVI-east locations leading to higher omission error than models trained for US-MEX, whereas for BEVI-west US-MEX models showed higher omission error. For both BEVI-east and BEVI-west, 2020 models for RCP 4.5 showed slightly higher omission error than RCP 8.5 models.

Table 3. Omission error rates for 2020 projected (2011 – 2030) models using an independent testing dataset of 2011-2018 occurrences.

	Annual			
	BEVI-west		BEVI-east	
	Breeding	US_MEX	Breeding	US_MEX
2020 (RCP 4.5)	0.039	0.056	0.394	0.356
2020 (RCP 8.5)	0.038	0.057	0.348	0.333
	Seasonal			
2020 (RCP 4.5)	0.099	0.170	0.241	0.017
2020 (RCP 8.5)	0.091	0.160	0.223	0.010

For BEVI-west, 2020 models trained at the US-MEX extent produced more conservative predictions (Figure 3, top, column B and D), whereas for BEVI-east annual models (Figure 4, bottom, column A and B) were more conservative than seasonal models. Seasonal models for breeding range predicted the largest amount of climatic suitability for both populations, with BEVI-east covering a substantial portion of the US and Mexico. Most models predicted the most climate range expansion (in blue) at the northern portions of predictions, with the exception of BEVI-east models where increased suitability was predicted in the wintering range in central and southern Mexico. Interestingly, the breeding model with the largest predicted suitable area (Figure 3, bottom, column C) also exhibited the most range contraction under 2020 scenarios (in yellow).

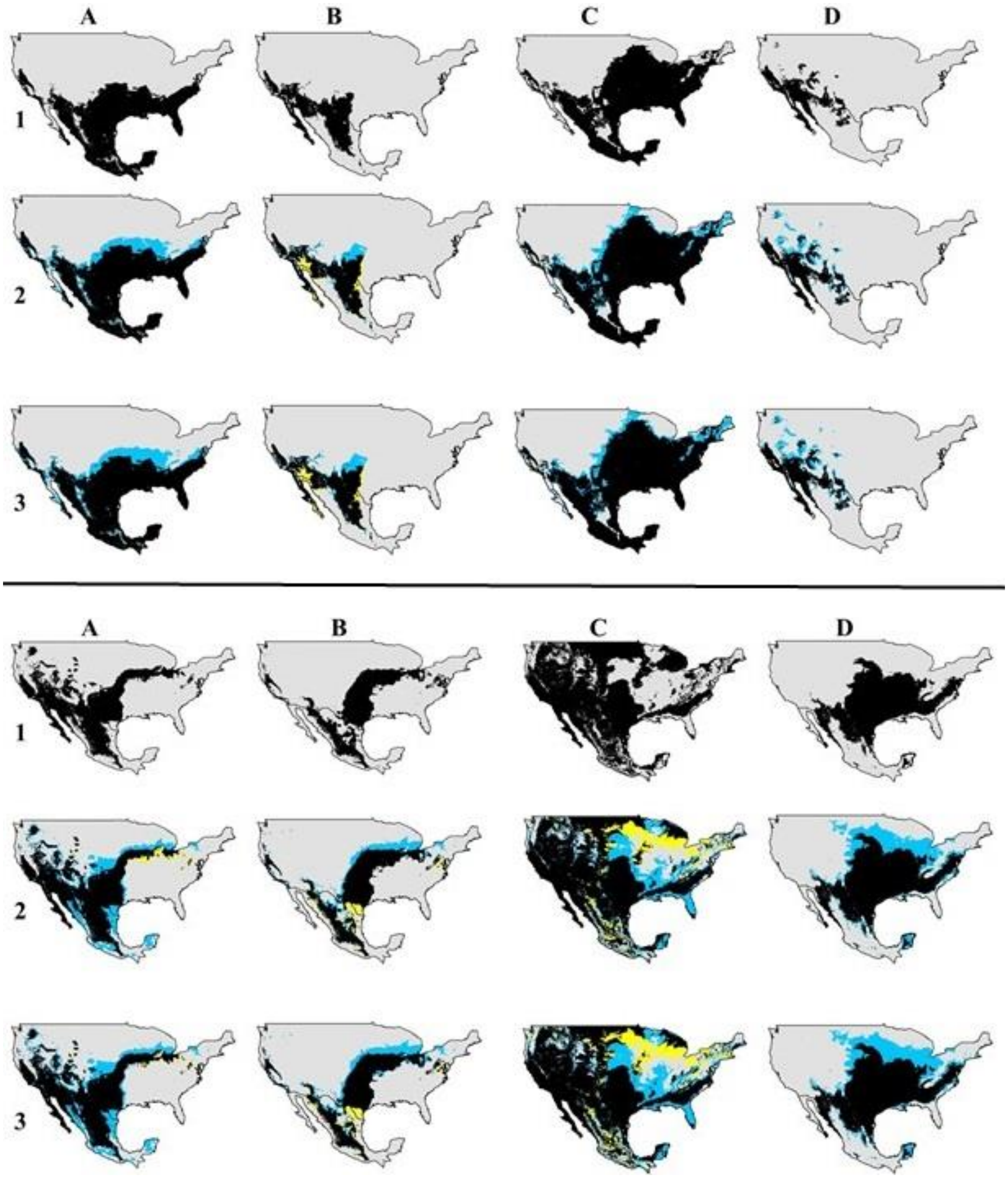


Figure 4. Model projections for 1961-1990 (row 1) and 2020 projections under RCP 4.5 (row 2) and RCP 8.5 (row 3) for both the western (top) and eastern (bottom) population of Bell’s Vireo. Annual models trained at the breeding range are show in column A and US-MEX trained models in column B. Column C and D are breeding range trained and US-MEX trained models, respectively, for seasonal models. Black indicates mutual predictions between 1961-1990 and 2020, blue indicates new areas predicted by 2020 (climatic range expansion) and yellow indicates areas predicted by 1961-1990 and not 2020 (climatic range contraction due to future changes in environmental conditions).

Overall, 2050 models showed the same trends as 2020 models with increased range expansion (see areas in blue for Figure 4). For BEVI-east, RCP 4.5 (Figure 5, bottom, row 2) showed more range contraction (in yellow) than RCP 8.5 (row 3), which appeared to correspond to higher range expansion (in blue) for RCP 8.5. The same is true for 2080 (Figure 6) predictions where all models show increasing by more northern range expansion. A troublesome trend for annual US-MEX models for both populations (Figure 6, column B) is loss of wintering range in favor of more northerly expansion. Seasonal breeding range models (Figure 6, column C) predict nearly all of the US and Mexico climatically suitable under both RCP 4.5 (row 2) and RCP 8.5 (row 3).

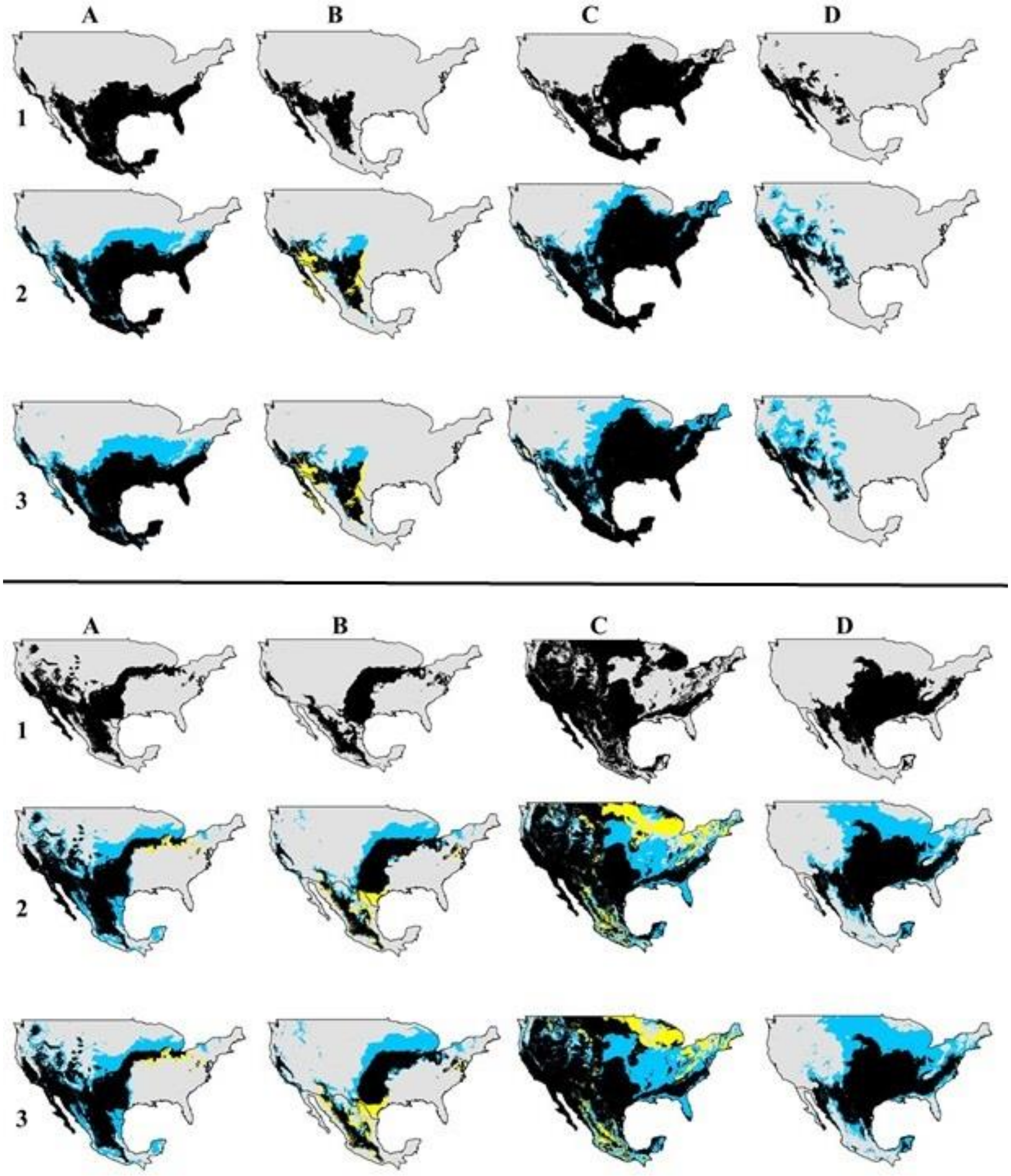


Figure 5. Model projections for 1961-1990 (row 1) and 2050 projections under RCP 4.5 (row 2) and RCP 8.5 (row 3) for both the western (top) and eastern (bottom) population of Bell’s Vireo. Annual models trained at the breeding range are show in column A and US-MEX trained models in column B. Column C and D are breeding trained and US-MEX trained models, respectively, for seasonal models. Black indicates mutual predictions between 1961-1990 and 2050, blue indicates new areas predicted by 2050 (climatic range expansion) and yellow indicates areas predicted by 1961-1990 and not 2050 (climatic range contraction due to future changes in environmental conditions).

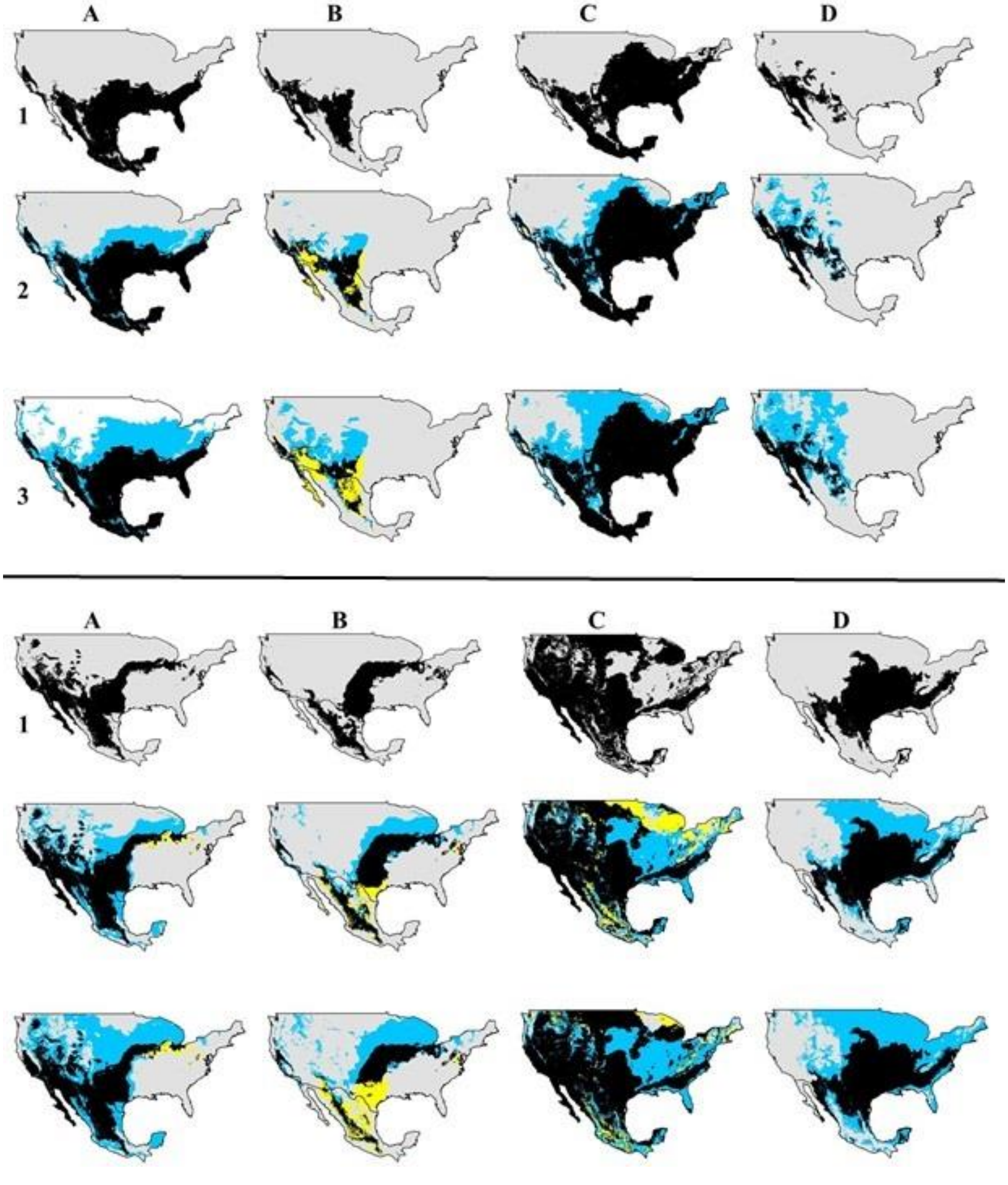


Figure 6. Model projections for 1961-1990 (row 1) and 2080 projections under RCP 4.5 (row 2) and RCP 8.5 (row 3) for both the western (top) and eastern (bottom) population of Bell's Vireo. Annual models trained at the breeding range are show in column A and US-MEX trained models in column B. Column C and D are breeding trained and US-MEX trained models, respectively, for seasonal models. Black indicates mutual predictions between 1961-1990 and 2080, blue indicates new areas predicted by 2080 (climatic range expansion) and yellow indicates areas predicted by 1961-1990 and not 2080 (climatic range contraction due to future changes in environmental conditions).

Discussion

I showed that model performance was influenced by temporal scale of the variables, the population, and the training area extent. For BEVI-west, the population exhibiting more distinct habitat areas, seasonal variables performed better when discriminating suitable habitat at the US-MEX extent. Here, the summer heat index was by far the highest contributing variable, which makes sense since the mean SHM value delineating BEVI-west habitat was nearly four times higher (482.32) than the mean for US-MEX (118.47). This shows that the BEVI-west population is utilizing areas that are hotter and drier than background conditions. However, for within range seasonal models, temperature and to a slightly lesser degree, precipitation were better at differentiating areas suitable to BEVI-west populations. The cottonwood-willow riparian areas used by the BEVI-west population likely do exhibit unique temperature and precipitation regimes from the background, whereas the heat moisture index is probably not as variable across the breeding range and thus contributes very little to model performance. Interestingly, annual variables produce a better model for BEVI-west breeding range, with temperature contributing a significant amount to model discrimination, followed distantly by precipitation. The better performance of the annual variables may hint at a yearly process shaping the area into varying levels of suitability. For example, riparian areas have longer wet periods, directly influencing the vegetation and forming microclimate and vertical structure variation.

For BEVI-east, which exhibits less defined habitat, seasonal variables did not perform better than annual variables at either training extent, although variable contribution was different. All three annual variables contributed similarly to delineating within range discrimination whereas seasonally, summer heat index was most important. Since BEVI-east individuals use a range of dense vegetation, which can be located in a wide array of landscapes such as grasslands, shrublands, forest edges, fence lines, and riparian areas, no one variable best discriminates between all of these areas annually. However, seasonally, variation in temperature and precipitation makes the summer heat moisture index a good discriminator. At the US-MEX extent, temperature then precipitation were the highest contributing variables, both seasonally and annually. As BEVI-east individuals occupy areas in the Central US, the selection of temperature and precipitation as most important annually and seasonally aligns with the ecology of the Great Plains.

When both seasonal and annual models were projected to 2020 for both RCP scenarios (4.5 and 8.5), the lowest omission error was exhibited by seasonal models under RCP 8.5 for both populations, however omission error was only slightly lower than under RCP 4.5. This should be interpreted carefully as seasonal models predicted significantly more suitable area than annual models, indicating that variables more specific to each population would yield more realistic future predictions, although other studies did not appear to have this issue (Nakazawa et al. 2004, Reside et al. 2010, Smeraldo et al. 2018). I did not focus on optimizing variables in this study, and so doing is expected to increase model performance and more importantly produce more realistic and useful models. This is notable considering that the two populations of this species likely respond to different variables at different spatial and temporal scales. Conservation actions taken at the species level would be ineffective without the regional context of each population. What constitutes high quality habitat differs between BEVI-west and BEVI-east, even though both populations exhibit a preference for the densest vegetation within selected areas (Parody and Parker 2002). Joos et al. (2014) showed that territory success is tied to habitat quality and the selection of territories is adaptive for BEVI-east and that nest-site placement is also adaptive for both BEVI-west and BEVI-east (Parody and Parker 2002). My models show that climatically, many areas of the US and Mexico are suitable, however the biotic interactions shaping the dense vegetation preferred by both species is lacking in my models. Since both populations show plasticity of within-site use, although more so in the case of BEVI-east, preserving the landscapes and ecological processes that produce the appropriate habitat is essential.

When I compared the climatic conditions between the two breeding ranges and the wintering range, I found high agreement for annual heat moisture, and high overlap for mean precipitation. I found that mean annual temperature exhibited a distinct difference between ranges, perhaps suggesting that both populations of Bell's Vireo are switching their temperature niche between wintering and breeding ranges. Looking at the summer average temperature, I noticed that during the breeding season, both populations have a wider range of temperatures. For precipitation, both annual and seasonal variables show high overlap, indicating that perhaps Bell's Vireos seek out breeding conditions with similar annual and seasonal precipitation values, which likely contribute to much of the habitat structure that the birds are associated with. I also found that breeding and wintering ranges showed similar annual heat moisture, but seasonally that appears to not be the case as both breeding ranges show more variation in summer heat moisture than what is available in their wintering range during the breeding period.

Nakazawa et al. (2004), used SDM to investigate the climatic niche of 21 migratory bird species by creating distributional predictions from seasonal models (breeding and wintering) that were projected either onto the same season (breeding projected to breeding) or different season (breeding projected to wintering). For Bells' Vireo, they found that wintering predictions based on breeding season models performed nearly identical to wintering predictions from wintering season models, with the same trend for breeding season (Nakazawa et al. 2004). Based on Nakazama et al. (Nakazawa et al. 2004), Bell's Vireo follows a climatic niche between wintering and breeding seasons. Although I did not create wintering range models, I did compare the variation in conditions associated with the wintering range and breeding ranges of both populations. The lower latitude of wintering locations allowed for higher annual and seasonal temperature and precipitation compared to breeding range conditions, suggesting that these climatic conditions likely would not support the breeding habitat used by either population. However, when compared the frequency of use at locations for both breeding ranges and wintering locations, I found that both populations appear to follow a temperature component of the climatic niche, but not a precipitation component for BEVI-east. These findings suggest a more complex relationship between seasonal environmental conditions, perhaps attributed to habitat features structured by microclimate (e.g. vegetation density).

I have shown that incorporating temporal variables in SDM influences both model performance and predictions. For migratory birds incorporating aspects of shifting ecological requirements across time into SDM, especially under future climate scenarios, will benefit conservation efforts, specifically to address conservation of important flyways and breeding/wintering habitats (Engler et al. 2017). This may be especially critical for migrating birds since recent research by Martin and Fahrig (2018) has shown that specialists generally travel farther than generalist bird species and climate change is expected to uncouple migration times to seasonal availability of habitat (Kellermann and van Riper 2015). However, even resident birds have shown seasonality changes as well (Reside et al. 2010). Thus, in addition to spatial scale considerations such as extent (Barve et al. 2011, Song et al. 2013, Vale et al. 2014) and grain (Hanberry 2013, Connor et al. 2018, Mertes and Jetz 2018), temporal scale should also be part of the decision making process (Engler et al. 2014, Williams et al. 2017). I acknowledge that incorporating temporal variables is not straightforward as species phenology is as unknown as species' responses to spatial scales. Besides selecting scales based on biological input (Wiens 1989, Levin 1992), Wheatley and Johnson (2009) suggest that continuous scales should be utilized to find the "domains of scale" (Wiens 1989) where species' responses change. When one must consider three "scales", extent, grain, and temporal, and contend with limited data (Scales et al. 2017) as is often the case, moving toward more spatially and temporally explicit SDM will be a slow and challenging process.

Conclusions

My work adds to the growing literature describing the inclusion of phenology-related environmental predictors in SDM (Nakazawa et al. 2004, Reside et al. 2010, Smeraldo et al. 2018). However, I show that the effects of seasonality are not consistent within a species, as populations are influenced by differences in geographic locations. I demonstrated that the seasonal ecology of populations within a species may be driven by both different temporal scales (i.e. season or annual) and different geographic extents (i.e. intra-breeding range and inter-breeding range). Intraspecific variation in distributional ranges must be accounted for especially when projecting to future climates, as the species-environment relationship must be accounted for when predicting future distributions.

REFERENCES

Barve, N., V. Barve, A. Jimenez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberon, and F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. Ecological Modelling 222:1810-1819.

Bean, W. T., L. R. Prugh, R. Stafford, H. S. Butterfield, M. Westphal, and J. S. Brashares. 2014. Species distribution models of an endangered rodent offer conflicting measures of habitat quality at multiple scales. Journal of Applied Ecology 51:1116-1125.

BISON. 2018. Vireo bellii. (Accessed through Biodiversity Information Serving Our Nation (BISON), https://bison.usgs.gov, 2018-10-01).

Bonthoux, S., S. Lefèvre, P.-A. Herrault, and D. Sheeren. 2018. Spatial and temporal dependency of ndvi satellite imagery in predicting bird diversity over France. Remote Sensing 10:1136.

Budnik, J. M., M. R. Ryan, and F. R. Thompson. 2000. Demography of Bell's Vireos in Missouri grassland-shrub habitats. Auk 117:925-935.

Connor, T., V. Hull, A. Viña, A. Shortridge, Y. Tang, J. Zhang, F. Wang, and J. Liu. 2018. Effects of grain size and niche breadth on species distribution modeling. Ecography 41:1270-1282.

eBird Basic Dataset. Version: EBD_relNov-2017. Cornell Lab of Ornithology, Ithaca, New York. Nov 2017.

Elith, J., S. J. Phillips, T. Hastie, M. Dudik, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of Maxent for ecologists. Diversity and Distributions 17:43-57.

Engler, J. O., D. Rödder, D. Stiels, and M. I. Förschler. 2014. Suitable, reachable but not colonized: Seasonal niche duality in an endemic mountainous songbird. Journal of Ornithology 155:657-669.

Engler, J. O., D. Stiels, K. Schidelko, D. Strubbe, P. Quillfeldt, and M. Brambilla. 2017. Avian SDMs: Current state, challenges, and opportunities. Journal of Avian Biology 48:1483-1504.

Environmental Systems Research Institute (ESRI). (2016). ArcGIS Release 10.4. Redlands, CA.

Fernandez, M., C. Yesson, A. Gannier, P. I. Miller, and J. M. N. Azevedo. 2017. The importance of temporal resolution for niche modelling in dynamic marine environments. Journal of Biogeography 44:2816-2827.

Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38-49.

Fisher, J. T., B. Anholt, and J. P. Volpe. 2011. Body mass explains characteristic scales of habitat selection in terrestrial mammals. Ecology and Evolution 1:517-528.

Franzreb, K. 1987. Endangered status and strategies for conservation of the Least Bell's Vireo. Western Birds 18:43 - 49.

Fynn, R. W. S., M. Chase, and A. Röder. 2014. Functional habitat heterogeneity and large herbivore seasonal habitat selection in northern Botswana. South African Journal of Wildlife Research 44:1-15.

García, A. 2008. The use of habitat and time by lizards in a tropical deciduous forest in western Mexico. Studies on Neotropical Fauna and Environment 43:107-115.

GBIF.org (2nd October 2018) GBIF Occurrence Download https://doi.org/10.15468/dl.eljky1.

Guisan, A., C. H. Graham, J. Elith, F. Huettmann, and N. S. Distri. 2007. Sensitivity of predictive species distribution models to change in grain size. Diversity and Distributions 13:332-340.

Guisan, A., and W. Thuiller. 2005. Predicting species distribution: Offering more than simple habitat models. Ecology Letters 8:993-1009.

Hanberry, B. B. 2013. Finer grain size increases effects of error and changes influence of environmental predictors on species distribution models. Ecological Informatics 15:8-13.

Hanczar, B., J. P. Hua, C. Sima, J. Weinstein, M. Bittner, and E. R. Dougherty. 2010. Small-sample precision of roc-related estimates. Bioinformatics 26:822-830.

Harris, R. M. B., M. R. Grose, G. Lee, N. L. Bindoff, L. L. Porfirio, and P. Fox-Hughes. 2014. Climate projections for ecologists. Wiley Interdisciplinary Reviews: Climate Change 5:621-637.

Hereford, J., J. Schmitt, and D. D. Ackerly. 2017. The seasonal climate niche predicts phenology and distribution of an ephemeral annual plant, *Mollugo verticillata*. Journal of Ecology 105:1323-1334.

Jacobi, G. Z., and S. J. Cary. 1996. Winter stoneflies (*Plecopter*a) in seasonal habitats in New Mexico, USA. Journal of the North American Benthological Society 15:690-699.

Johnson, J. H., R. Abbett, M. A. Chalupnicki, and F. Verdoliva. 2016. Seasonal habitat use of brook trout and juvenile steelhead in a Lake Ontario tributary. Journal of Freshwater Ecology 31:239-249.

Joos, C. J., F. R. Thompson, and J. Faaborg. 2014. The role of territory settlement, individual quality, and nesting initiation on productivity of Bell's Vireos (*Vireo bellii bellii*). Journal of Avian Biology 45:584-590.

Kellermann, J. L., and C. van Riper. 2015. Detecting mismatches of bird migration stopover and tree phenology in response to changing climate. Oecologia 178:1227-1238.

Klicka, L. B., B. E. Kus, P. O. Title, and K. J. Burns. 2016. Conservation genomics reveals multiple evolutionary units within Bell's Vireo (*Vireo bellii*). Conservation Genetics 17:455-471.

Knutti, R., D. Masson, and A. Gettelman. 2013. Climate model genealogy: Generation cmip5 and how we got there. Geophysical Research Letters 40:1194-1199.

Levin, S. A. 1992. The problem of pattern and scale in ecology. Ecology 73:1943-1967.

Lobo, J. M., A. Jimenez-Valverde, and R. Real. 2008. AUC: A misleading measure of the performance of predictive distribution models. Global Ecology and Biogeography 17:145-151.

Martin, A. E., and L. Fahrig. 2018. Habitat specialist birds disperse farther and are more migratory than habitat generalist birds. Ecology 99:2058-2066.

McClure, C. J. W., B. W. Rolek, and G. E. Hill. 2013. Seasonal use of habitat by shrub-breeding birds in a southeastern national forest. The Wilson Journal of Ornithology 125:731-743.

McGill, B. J. 2010. Matters of scale. Science 328:575-576.

Mertes, K., and W. Jetz. 2018. Disentangling scale dependencies in species environmental niches and distributions. Ecography 41:1604-1615.

Metcalf, C. J. E., and D. N. Koons. 2007. Environmental uncertainty, autocorrelation and the evolution of survival. Proceedings: Biological Sciences 274:2153-2160.

Meyer, K. M., K. Wiegand, D. Ward, and A. Moustakas. 2007. The rhythm of savanna patch dynamics. Journal of Ecology 95:1306-1315.

Milakovic, B., K. L. Parker, D. D. Gustine, R. J. Lay, A. B. D. Walker, and M. P. Gillingham. 2012. Seasonal habitat use and selection by grizzly bears in northern British Columbia. The Journal of Wildlife Management 76:170-180.

Mitchell, T. D., and P. D. Jones. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology 25:693-712.

Moloney, K. A., and S. A. Levin. 1996. The effects of disturbance architecture on landscape-level population dynamics. Ecology 77:375-394.

Montalto, V., G. Sarà, P. M. Ruti, A. Dell'Aquila, and B. Helmuth. 2014. Testing the effects of temporal data resolution on predictions of the effects of climate change on bivalves. Ecological Modelling 278:1-8.

Morrison, M. L., B. G. Marcot, and R. W. Mannan. 2006. Wildlife-habitat relationships: Concepts and applications. 3rd edition. Island Press, Washington D.C.

Mouton, A. M., B. De Baets, and P. L. M. Goethals. 2010. Ecological relevance of performance criteria for species distribution models. Ecological Modelling 221:1995-2002.

Nakazawa, Y., A. T. Peterson, E. Martínez-Meyer, A. G. Navarro-Sigüenza, and F. Moore. 2004. Seasonal niches of Nearctic-Neotropical migratory birds: Implications for the evolution of migration. The Auk 121:610-618.

Parody, J. M., and T. H. Parker. 2002. Biogeographic variation in nest placement: A case study with conservation implications. Diversity and Distributions 8:11-19.

Peterson, A. T., M. Papes, and J. Soberon. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. Ecological Modelling 213:63-72.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190:231-259.

Pickens, B. A., and S. L. King. 2014. Linking multi-temporal satellite imagery to coastal wetland dynamics and bird distribution. Ecological Modelling 285:1-12.

Radosavljevic, A., and R. P. Anderson. 2014. Making better Maxent models of species distributions: Complexity, overfitting and evaluation. Journal of Biogeography 41:629-643.

Reside, A. E., J. J. VanDerWal, A. S. Kutt, and G. C. Perkins. 2010. Weather, not climate, defines distributions of vagile bird species. PLOS One 5:e13569.

Scales, K. L., E. L. Hazen, M. G. Jacox, C. A. Edwards, A. M. Boustany, M. J. Oliver, and S. J. Bograd. 2017. Scale of inference: On the sensitivity of habitat models for wide-ranging marine predators to the resolution of environmental data. Ecography 40:210-220.

Schneider, D. C. 1994. Quantitative ecology: Spatial and temporal scaling. Academic Press Inc., San Diego, California.

Seo, C., J. H. Thorne, L. Hannah, and W. Thuiller. 2009. Scale effects in species distribution models: Implications for conservation planning under climate change. Biology Letters 5:39-43.

Sherry, T. W., and R. T. Holmes. 1993. Are populations of Neotropical migrant birds limited in summer or winter? Implications for management. Pages 47-57 in D. M. Finch and P. W. Stangel, editors. Status and management of Neotropical migratory birds. United States Forest Service, Fort Collins, Colorado.

Smeraldo, S., M. Di Febbraro, L. Bosso, C. Flaquer, D. Guixé, F. Lisón, A. Meschede, J. Juste, J. Prüger, X. Puig-Montserrat, and D. Russo. 2018. Ignoring seasonal changes in the ecological niche of non-migratory species may lead to biases in potential distribution models: Lessons from bats. Biodiversity and Conservation 27:2425-2441.

Song, W., E. Kim, D. Lee, M. Lee, and S.-W. Jeon. 2013. The sensitivity of species distribution modeling to scale differences. Ecological Modelling 248:113-118.

Suarez-Seoane, S., E. Virgos, O. Terroba, X. Pardavila, and J. M. Barea-Azcon. 2014. Scaling of species distribution models across spatial resolutions and extents along a biogeographic gradient. The case of the Iberian mole *Talpa occidentalis*. Ecography 37:279-292.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. Ebird: A citizen-based bird observation network in the biological sciences. Biological Conservation 142:2282-2292.

Swets, J. A. 1988. Measuring the accuracy of diagnostic systems. Science 240:1285-1293.

Telenius, A. 2011. Biodiversity information goes public: GBIF at your service. Nordic Journal of Botany 29:378-381.

Thornton, D. H., and R. J. Fletcher. 2014. Body size and spatial scales in avian response to landscapes: A meta-analysis. Ecography 37:454-463.

Vale, C. G., P. Tarroso, and J. C. Brito. 2014. Predicting species distribution at range margins: Testing the effects of study area extent, resolution and threshold selection in the Sahara-Sahel transition zone. Diversity and Distributions 20:20-33.

VanDerWal, J., L. P. Shoo, C. Graham, and S. E. William. 2009. Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? Ecological Modelling 220:589-594.

Varner, D. M., G. R. Hepp, and R. R. Bielefeld. 2014. Movements and seasonal use of habitats by rural and urban female mottled ducks in southeast Florida. The Journal of Wildlife Management 78:840-847.

Wang, T., A. Hamann, D. Spittlehouse, and C. Carroll. 2016. Locally downscaled and spatially customizable climate data for historical and future periods for North America. PLOS One 11:e0156720.

Weaver, C. P., R. J. Lempert, C. Brown, J. A. Hall, D. Revell, and D. Sarewitz. 2013. Improving the contribution of climate model information to decision making: The value and demands of robust decision frameworks. Wiley Interdisciplinary Reviews: Climate Change 4:39-60.

Wheatley, M., and C. Johnson. 2009. Factors limiting our understanding of ecological scale. Ecological Complexity 6:150-159.

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.

Williams, H. M., M. Willemoes, and K. Thorup. 2017. A temporally explicit species distribution model for a long distance avian migrant, the common cuckoo. Journal of Avian Biology 48:1624-1636.

CHAPTER V

CONCLUSIONS

The research presented here adds to the continuing conversation surrounding scale in species distribution modeling (SDM). In general as the spatial scale of an ecological processes becomes more coarse, the temporal range also increases, thus one must observe a process at the appropriate spatial and temporal scale while also considering the perspective of the species (Wiens 1989). Current SDM methodology, specifically the application of virtual species (where the species-environment relationship is known) has addressed this scaling consideration as models performed best when using variables scaled to the response of the species (Connor et al. 2018, Mertes and Jetz 2018). However, a functional modeling framework for SDM utilizing current knowledge on species does not exist as of yet. I have shown that by utilizing a multi-scale approach for a single species, insight into species-environment relationships is possible.

First, by using a spatially nested study design, I was able to show that the size of the study extent enhanced the negative effects of spatially biased presence locations. Further, variable contribution to model accuracy gain changed with extent size, as did variation in the environmental variables across the three extents. Changes in extent size alters the variation of environmental conditions across the study area and determine how the environment is quantified with background sampling. Further, spatially biased presence locations provide a skewed representation of the environment used by the species. Together, this serves as a cautionary tale that user decisions regarding presence data and extent of training do not always act independently within the modeling process, and instead may interact in unexpected ways.

Secondly, I highlighted the importance of data resampling techniques for studies manipulating the grain size of environmental predictors. Variation in environmental variables is generally expected to decrease with increasing grain size, although I argue that spatial structuring of the environment may have played a large role. In practice, the change in variation due to either resampling technique or grain size shapes the area from which background sampling occurs, thus providing the statistical distribution of environmental conditions to which presences are compared. Altering this relationship changes the model, its performance and geographical prediction.

Lastly, I addressed the temporal scale of environmental predictors in SDM, as ecological phenomena occur at intrinsic temporal scales (Wiens 1989, Mayor et al. 2009). Specifically, I illustrated that the importance of temporal variation in environmental conditions changed with extent as well as with population. By comparing seasonal versus annual environmental predictors for two populations of a migrating bird, I showed that conditions discriminating between suitable areas differed within the breeding range when compared to outside the breeding range. Both occurrence data and environmental predictors define the environment as a snapshot, capturing the variation in background conditions and presence locations at a particular spatial and temporal scale. The interaction between temporally mis-matched background and presence locations undermines the niche theory assumptions upon which SDM is based.

Processes occurring at multiple spatial and temporal scales structure the environment and organisms respond to the environment at multiple scales making scaling the core of nearly all ecological relationships (Mayor et al. 2009). The need to understand species' responses the environment across spatial and temporal scales is pressing as multiple threats to biodiversity exist (Ceballos et al. 2017, Spooner et al. 2018). To answer this call, the SDM framework may need to be altered to better recognize the importance of scale.

First, identifying the species' response grain (Mertes and Jetz 2018) of species is paramount, but this can only be achieved if a more exploratory workflow exists. Traditionally, environmental variables are subjected to pre-model evaluation techniques to establish "best performing" variables for inclusion. However, Mertes and Jetz (2018) demonstrated that the spatial structure of the variable and the response grain of species interact to provide a clearer picture of the species-environment relationship, but only at the right resolution. A workflow that addresses this phenomena was proposed for three uses of SDM: inferring niche relationships, evaluating scale dependence, and predicting occurrences (Mertes and Jetz 2018). However, this workflow does not address the importance of extent, or the temporal scale of variables. Spatial extent selection is highly tied to study aims. A suggested approach is to train the model with a restricted extent, then apply the model to the larger extent under consideration (Anderson and Raza 2010) to avoid overfitting. I suggest a modified approach where the training region is delineated based on heterogeneity. For example, an ecoregion is expected to have more homogeneity within the region compared to without. This creates an extent that is ecologically defined by the processes occurring within it, thus the species occurring within it should exhibit a relationship to the environmental conditions that can be captured if measured at the appropriate scale. The temporal scale of data is perhaps the most poorly considered aspect in SDM (Smeraldo et al. 2018), as most species show temporal responses (Miller-Rushing and Weltzin 2009). Increasing the incorporation of temporal scale in SDM requires a new approach to niche-theory addressing the temporal component. However, a priori knowledge of species ecology should allow users to assimilate appropriate data into SDM. SDM has only been used for a couple of decades, but has undergone rapid development. Further improvement of data, statistical models, and evaluation techniques will continue to move the field forward.

REFERENCES

Anderson, R. P., and A. Raza. 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: Preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. Journal of Biogeography 37:1378-1393.

Ceballos, G., P. R. Ehrlich, and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proceedings of the National Academy of Sciences 114:E6089.

Connor, T., V. Hull, A. Viña, A. Shortridge, Y. Tang, J. Zhang, F. Wang, and J. Liu. 2018. Effects of grain size and niche breadth on species distribution modeling. Ecography 41:1270-1282.

Mayor, S. J., D. C. Schneider, J. A. Schaefer, and S. P. Mahoney. 2009. Habitat selection at multiple scales. Écoscience 16:238-247.

Mertes, K., and W. Jetz. 2018. Disentangling scale dependencies in species environmental niches and distributions. Ecography 41:1604-1615.

Miller-Rushing, A. J., and J. Weltzin. 2009. Phenology as a tool to link ecology and sustainable decision making in a dynamic environment. New Phytologist 184:743-745.

Smeraldo, S., M. Di Febbraro, L. Bosso, C. Flaquer, D. Guixé, F. Lisón, A. Meschede, J. Juste, J. Prüger, X. Puig-Montserrat, and D. Russo. 2018. Ignoring seasonal changes in the ecological niche of non-migratory species may lead to biases in potential distribution models: Lessons from bats. Biodiversity and Conservation 27:2425-2441.

Spooner, F. E. B., R. G. Pearson, and R. Freeman. 2018. Rapid warming is associated with population decline among terrestrial birds and mammals globally. Global Change Biology 24:4521-4531.

Wiens, J. A. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.

APPENDICES

Appendix 1. List of variables used in Chapter 2 models

Name	Variable
mean temp	Mean Temperature
min temp	Minimum Temperature
max temp	Maximum Temperature
mean prcp	Mean Precipitation
min prcp	Minimum Precipitation
max prcp	Maximum Precipitation
mean ndvi	Mean Normalized Difference Vegetation Index
min ndvi	Minimum Normalized Difference Vegetation Index
max ndvi	Maximum Normalized Difference Vegetation Index
mean lai	Mean Leaf Area Index
min lai	Minimum Leaf Area Index
max lai	Maximum Leaf Area Index
elev	Elevation
aspect	Aspect
slope	Slope
flow dir	Flow Direction
flow	Flow Accumulation
topo	Wetness Index

Appendix 2. Percent contribution of the highest performing variable for each model at the three extents used in Chapter 2

Model	Extent	100P	50R	25R	50C	25C
FULL	Small	21.7	18.9	21.0	36.2	30.2
	Medium	25.7	30.4	16.9	30.4	23.8
	Large	30.5	28.8	30.5	27.3	28.9
NOCOL	Small	35.0	30.0	31.7	47.0	56.9
	Medium	31.0	30.5	29.3	40.3	26.7
	Large	53.9	53.7	53.4	57.4	50.7
CLM	Small	43.4	29.5	30.7	47.8	40.1
	Medium	34.5	30.2	27.7	34.7	43.2
	Large	62.0	66.0	60.0	44.2	46.3
VEG	Small	30.2	45.0	31.4	36.6	53.7
	Medium	32.6	52.3	45.5	34.1	46.4
	Large	38.0	40.5	40.8	36.1	32.1
TOPO	Small	70.5	69.5	77.9	75.5	77.8
	Medium	78.6	74.5	66.9	79.1	74.4
	Large	68.7	65.9	58.5	89.8	85.2
MIN	Small	47.2	50.9	38.5	68.2	40.4
	Medium	43.9	44.6	50.9	40.1	49.3
	Large	65.4	64.3	71.5	63.2	59.4
MEAN	Small	45.5	29.0	47.9	77.3	74.3
	Medium	31.8	37.2	35.0	52.8	49.7
	Large	66.2	67.3	58.4	55.0	54.4
MAX	Small	47.0	32.0	49.3	69.3	60.8
	Medium	55.7	48.0	53.3	47.0	43.1
	Large	49.0	51.3	41.1	39.5	40.0

Appendix 3. Pearson Correlation Coefficient for the number of variables and the five model evaluation metrics used in Chapter 2

Large Extent

	AUC	Omission Error	Kappa	Predicted Suitable Area	Variable Contribution
100P	0.566	0.936	-0.266	-0.694	-0.651
50R	0.569	0.777	0.618	-0.672	-0.721
25R	0.476	0.565	0.577	-0.619	-0.667
50C	0.193	0.535	0.611	-0.587	-0.415
25C	0.241	0.475	0.472	-0.548	-0.436

Medium Extent

	AUC	Omission Eerror	Kappa	Predicted Suitable Area	Variable Contribution
100P	0.756	0.840	-0.610	-0.945	-0.458
50R	0.265	0.828	-0.148	-0.876	-0.458
25R	0.546	0.909	0.581	-0.862	-0.634
50C	-0.358	0.474	0.522	-0.575	-0.425
25C	-0.324	0.544	0.381	-0.691	-0.699

Small Extent

	AUC	Omission Error	Kappa	Predicted suitable Area	Variable Contribution
100P	0.672	0.447	-0.288	-0.800	-0.651
50R	0.745	0.709	0.813	-0.835	-0.552
25R	0.297	0.691	0.696	-0.843	-0.540
50C	-0.065	0.483	0.651	-0.486	-0.486
25C	-0.065	0.460	0.450	-0.421	-0.676

VITA

Cassondra Meredith Walker

Candidate for the Degree of

Doctor of Philosophy

Thesis: SPATIAL AND TEMPORAL COMPONENTS OF ENVIRONMENTAL PREDICTORS IN SPECIES DISTRIBUTION MODELING: A CASE STUDY USING BELL'S VIREO (VIREO BELLII)

Major Field: Integrative Biology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Integrative Biology at Oklahoma State University, Stillwater, Oklahoma in December, 2018.

Completed the requirements for the Bachelor of Science in Wildlife Ecology at Oklahoma State University, Stillwater, Oklahoma in December, 2009.

Experience:

Graduate Research/Teaching Associate, Department of Integrative Biology, Oklahoma State University, Stillwater, OK. 2013-2018.

Professional Memberships:

Welder Wildlife Foundation
American Ornithologists Union