Received: 17 February 2017 | Revised: 21 April 2017 | Accepted: 18 May 2017

DOI: 10.1111/geb.12610

CORRESPONDENCE

WILEY Global Ecology and Biogeography A Journal of Macroecology

Using the KDE method to model ecological niches: A response to Blonder et al. (2017)

Dear Editor:

Recently, we noted (Qiao, Escobar, Saupe, Ji, & Soberón, 2016) that multivariate kernel density estimation (KDE) may not outperform other methods for estimating hypervolume geometries and, moreover, under certain circumstances the algorithm would not detect 'holes' in environmental space, as Blonder, Lamanna, Violle, and Enquist (2014) had proposed. In our original note (Qiao et al., 2016), we explained that KDE (a) is sensitive to both sample size and environmental dimensionality, (b) may overestimate niche volumes in low dimensions and constrict niche volume estimates in high dimensions, and (c) is useful only to the extent that the realized niche is sought and not the fundamental niche. Here, we also note that bandwidth is a crucial parameter for KDE, and its selection needs to be evaluated rigorously and all assumptions stated clearly.

In their response to our comments, Blonder et al. (2017) indicated that (a) KDE output depends in useful ways on dataset size and bias, (b) other species distribution modelling methods make equally stringent but different assumptions about dataset bias, (c) we made an incorrect data transformation in our original experiments that may result in unfair comparisons, and (d) hypervolume methods are more general than KDE and have other benefits for niche modelling. We address these points below, which we divide into two main categories: methodological concerns and theoretical concerns.

METHODOLOGICAL CONCERNS

Blonder et al. (2017) criticized our transformation of units during the modelling process. We had log-transformed the data when constructing KDE hypervolumes following the log-transformed data framework in the 'hypervolume' R package demo code (see https://cran.r-project.org/web/packages/hypervolume/index.html), developed by Blonder et al. (2014). However, we agree with with Blonder et al. (2017) that this transformation is not necessary if hypervolumes are subsequently delineated and plotted in untransformed space. Consequently, we re-ran our analyses using untransformed units and two bandwidth configurations to explore their influence on model results (Supplementary Information Figure S1): the default estimated using the 'estimate_bandwidth()' function in the 'hypervolume' package from Blonder et al. (2014), and the value obtained from this function when divided by two to obtain a model with high fit with the data.

Our results remained consistent using the default bandwidth from estimate_bandwidth(). That is to say, our previous conclusions (Qiao et al., 2016) hold regardless of whether we use transformed or untransformed data; KDE overestimated niche volumes in low environmental dimensions, underestimated niche volumes in high dimensions, was unable to detect holes in environmental space with low sample size (Supporting Information Figures S2c and S3c) and was plagued by decreased sensitivity (Supporting Information Figure S4). The other evaluation metrics, including specificity, hypervolumes and the Jaccard similarity index (Supporting Information Figures S5–S7), show patterns similar to those reported in our previous analyses (Qiao et al., 2016).

Our re-analysis using a smaller bandwidth detected holes (Supporting Information Figures S2.3c and S3.3c), but at the cost of increased type II error (Supporting Information Figures S1.3a, S2.3a and S3.3a). This is an important intervention in the model parameterization that should not be overlooked; a pragmatic a posteriori parameterization was necessary for KDE to reconstruct the 'holey' niche of the virtual species effectively. Of course, when dealing with data from real species with unknown niche shapes, bandwidth selection would be more complex. The bandwidth is a crucial parameter for KDE, which we noted in our original manuscript, and bandwidth selection deserves further research. We recommend that researchers using KDE explain their assumptions during bandwidth selection, explore a series of bandwidth configurations and present the results of these models for more informed conclusions.

THEORETICAL CONCERNS

The remaining points made by Blonder et al. (2017) are primarily conceptual in nature. Blonder et al. (2014, 2017) argue that fundamental niches can have holes and complex shapes in higher dimensions. Although we argue that this is still far from certain, what is most relevant for discussions herein is that niches in high dimensions may be highly clustered in the central regions of environmental space (Drake, 2015), such that any holes in niche 'hyperspace' are, once again, difficult to identify and determine. Blonder et al. (2017) quote several references in the literature to suggest that fundamental niches may have complex forms. However, our own perusal of these references indicates that these estimates (obtained either from first-principle models or experimental data) have either simple convex shapes, or the data presented include only a few points, making the estimation of 'complex shapes' a doubtful exercise at best. This point, however, may be moot. KDE can be a useful method to fit shapes in multivariate spaces; interpretation about the meaning of the shapes may be best left to the researchers.

Global Ecology and Biogeography A Journal of Macroecology WILEY

We note that an overfitted KDE may be no more informative than using the original species occurrence records to identify the occupied environments, and a complex model with high fit (e.g., narrow KDE bandwidth) would be redundant. Comparison of KDE with other, more physiologically realistic methods (e.g., range bagging; Drake, 2015) is warranted. We agree that KDE is a promising method and should be included in the toolbox of ecological niche modellers. KDE has both pros and cons that may not be present in other algorithms, making it complementary and not opposed to other methods. Indeed, our original cautionary note (Qiao et al., 2016) was inspired by our interest in preventing the adoption of single algorithms as 'silver bullets' for characterizing fundamental and/or realized niches for any given species. The algorithm of choice depends on the nature of the research question, as Blonder et al. (2017) also note, and on the nature of the research data.

ACKNOWLEDGMENTS

The authors thank Pedro Peres-Neto and Ben Blonder for a productive discussion about applications of hypervolume methods in modern ecology.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

Huijie Qiao[1], Luis E. Escobar[2], Erin E. Saupe[3,4], Liqiang Ji[1], Jorge Soberón[5]

[1]*Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing, China*

[2]*Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, St Paul, Minnesota*

[3]*Department of Geology & Geophysics, Yale University, New Haven, Connecticut*

[4]*Department of Earth Sciences, South Parks Road, Oxford, United Kingdom*

[5]*Biodiversity Institute and Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas*

Correspondence

Luis E. Escobar, Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, 135 Skok Hall, 2003 Upper Buford Circle, St. Paul, MN 55108, USA.

Email: lescobar@umn.edu

Editor: Pedro Peres-Neto

REFERENCES

Blonder, B., Lamanna, C., Violle, C., & Enquist, B. J. (2014). The *n*-dimensional hypervolume. *Global Ecology and Biogeography*, *23*, 595–609. doi:10.1111/geb.12146

Blonder, B., Lamanna, C., Violle, C., Enquist, B. J. (2017). Using *n*-dimensional hypervolumes for species distribution modeling: A response to Qiao et al. (2016). *Global Ecology and Biogeography*, *26*, 1071–1075.

Drake, J. M. (2015). Range bagging: A new method for ecological niche modelling from presence-only data. *Journal of The Royal Society Interface*, *12*, 20150086.

Qiao, H., Escobar, L. E., Saupe, E. E., Ji, L., & Soberón, J. (2016). A cautionary note on the use of hypervolume kernel density estimators in ecological niche modelling. *Global Ecology and Biogeography*, *26*, 1066–1070.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.