

Armadillo Mapper: A Case Study of an Online Application to Update Estimates of Species' Potential Distributions

Tropical Conservation Science
Volume 10: 1–5
© The Author(s) 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1940082917724133
journals.sagepub.com/home/trc



Xiao Feng¹, Mariela C. Castro², Evan Linde³, and Monica Papeş⁴

Abstract

Ecological niche modeling (ENM) uses species' known occurrences and environmental variables to estimate potential geographic distributions. The potential distribution of a species can be updated and refined after new occurrences are collected; therefore, we developed an online application (Armadillo Mapper [AM]) that automatically updates the ecological niche model for the hairy long-nosed armadillo (*Dasyurus pilosus*) with new observations and provides a revised potential distribution map. AM can be used by researchers and conservation practitioners without a deep knowledge of geographic information system (GIS) techniques and ENM. Here, we introduce this online application and demonstrate the use of our application with a new observation of *D. pilosus*. AM has the potential to be extended to other species and to bridge the gap between distributional knowledge synthesized in a computer lab and conservation efforts carried out in the field.

Keywords

Armadillo, *Cryptoproctus pilosus*, *Dasyurus pilosus*, ecological niche model, Maxent, *Shiny*, species distribution model

The hairy long-nosed armadillo (*Dasyurus pilosus*) is endemic to the cloud forest and subparamo habitats of Peru (Wetzel, 1985) and is unique among armadillo species because of the long hair concealing the carapace (Castro et al., 2015; Wetzel & Mondolfi, 1979). Due to the lack of knowledge on this species, Superina, Pagnutti, and Abba (2014) have suggested that scientists investigating armadillos give priority to *D. pilosus*. One of the first points to be elucidated before any new field surveys is where to search for this species. Aiming to answer this question, Feng, Castro, McBee, and Papeş (2017) compiled *D. pilosus* historical distributional data, assessed the relationship between *D. pilosus* records and climatic conditions (i.e., climatic niche) using ecological niche modeling (ENM), and provided a potential distribution map for the species.

Using ENM for predicting species' potential distributions has helped in discovering rare species in field surveys (Fois, Fenu, Cuenca Lombraña, Cogoni, & Bacchetta, 2015; Menon, Choudhury, Khan, & Peterson, 2010). Such application of ENM should be iterative, with newly collected distributional information (i.e., occurrences) incorporated in the modeling process to refine or update estimates of potential distribution.

The updating procedure is important for improving the accuracy of the predicted potential distribution because more information of the species' climatic niche will be included in the modeling process (Saupe et al., 2012).

Shortly after the potential distribution of *D. pilosus* was published (Feng et al., 2017), we were informed that the species was observed in Cajamarca Department, Peru, in 2011 (A. Tovar Narvaez and S. Sánchez Tello, personal communication, April 4, 2017). The significance of this observation is twofold.

¹Department of Integrative Biology, Oklahoma State University, Stillwater, OK, USA

²Laboratório de Paleontologia de Ribeirão Preto, FFCLRP, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brazil

³High Performance Computing Center, Oklahoma State University, Stillwater, OK, USA

⁴Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA

Received 12 June 2017; Revised 10 July 2017; Accepted 13 July 2017

Corresponding Author:

Xiao Feng, Department of Integrative Biology, Oklahoma State University, 501 LSW, Stillwater, OK 74078, USA.
Email: Xiao.Feng@okstate.edu



First, this provides support for the prediction of potential distribution in Cajamarca, where, to our knowledge, no other specimen has been collected. Second, in a broader sense, this illustrates the usefulness of ENM in investigating the distribution of rare or little studied species (Peterson et al., 2011). New species' occurrences can be used to update a niche model and refine the estimated potential distribution, thus better guiding future field surveys (Feng et al., 2017). This new observation inspired us to develop a mechanism to update the prediction of *D. pilosus* potential distribution for the benefit of researchers continuing to collect new records in the field.

Updating potential distributions requires knowledge of GIS techniques and ENM, but some conservation practitioners may not have skills or time to readily replicate the GIS analysis and ENM process. Moreover, ENM is continuously being developed (Feng & Papeş, 2017; Merow, Wilson, & Jetz, 2017; Muscarella et al., 2014; Phillips, Anderson, Dudík, Schapire, & Blair, 2017), thus the best practices of ENM keep changing. On the other hand, researchers have been trying to develop user-friendly ENM software that bridge modeling techniques and conservation applications; one of the notable developments is Wallace, an open-source platform that guides users through a benchmark ENM procedure (Kass, Vilela, Aeillo-Lammens, Muscarella, & Anderson, 2017). Different model parameters and settings influence model fit and predictions of potential distributions, but conservation practitioners are less likely to compare or optimize predictions, as they are probably more interested in a final product of practical use. To narrow down the knowledge gap of modeling best practices and to address the real-world need of revising potential distribution maps, we built an online application that automatically updates the potential distribution of *D. pilosus* based on new occurrences.

Implementation of Armadillo Mapper (AM)

We built Armadillo Mapper (AM), a web application that takes new occurrences from users as data input and generates an updated potential distribution map. We implemented AM using *Shiny*, a web application framework for R (Chang, Cheng, Allaire, Xie, & McPherson, 2016). AM mainly relies on two R scripts: one draws the user interface and the other implements the spatial analysis, model training, and distribution prediction on a server. The user interface can be displayed in either English or Spanish. The data analysis and map display mainly rely on several R packages: *dismo* (Hijmans, Phillips, Leathwick, & Elith, 2013), *ggplot2* (Wickham, 2009), *raster* (Hijmans et al., 2016), *sp* (Pebesma & Bivand, 2005), and *leaflet* (Cheng, Karambelkar, & Xie, 2017). AM is hosted on a *Shiny* server with the CentOS 7 operating system at Oklahoma

State University High Performance Computing Center (<http://armadillomapper.fengxiao.info>). The source code of AM is available at GitHub (<https://github.com/shandongfx/ArmadilloMapper>).

Users can upload new occurrences to AM in two ways: enter coordinates manually (up to five records) or upload all coordinates in a .csv file (batch input). AM updates the potential distribution based on the new occurrences and the original occurrence data set compiled in Feng et al. (2017). The modeling procedure is the same as that used in Feng et al. (2017), which includes the following steps: decreasing spatial bias of the occurrence data set, randomly selecting background samples from a two decimal degree spatial buffer around occurrences, using linear and quadratic features and the default regularization value in Maxent algorithm (Phillips & Dudík, 2008), selecting climatic variables based on collinearity and variable contribution to model accuracy, and transforming the model prediction into a potential distribution map based on the threshold of minimum training presence. AM displays the updated potential distribution in either a static figure or an interactive map. The static figure is easier to transfer to other documents, whereas the interactive map overlays the potential distribution on a base map (Cheng et al., 2017), which can be zoomed in or out by the user and should be more useful in practice. AM can also display the climatic niche in environmental space along two climatic variables (annual mean temperature and annual precipitation). This function may help to compare climatic conditions of new occurrences to those of previously collected specimens.

Users can also share their occurrence data by submitting them to AM. After reviewing the new occurrences, we will make them publicly available by adding them to the database, so that other users can view such information as well as apply the information as baseline for future updates.

A Case Study of AM

We demonstrate the use of AM with a recent observation of *D. pilosus* (A. Tovar Narvaez and S. Sánchez Tello, personal communication, April 4, 2017) at 78.4492°S and 6.64722°W (Bambamarca District, Hualgayoc Province, Cajamarca Department) as data input. After clicking the button "Update map," the updated map of the species' potential distribution, displayed on the right side of the screen (Figure 1), shows a broader area of suitability than the potential distribution estimated by Feng et al. (2017). By switching to the "interactive map" tab, the user can view the potential distribution plotted on an interactive map (Figure 2), whereas switching to the "climatic niche" tab provides a view of the niche in two-dimensional environmental space. We found that the climatic conditions of the new occurrence are within the boundary of the

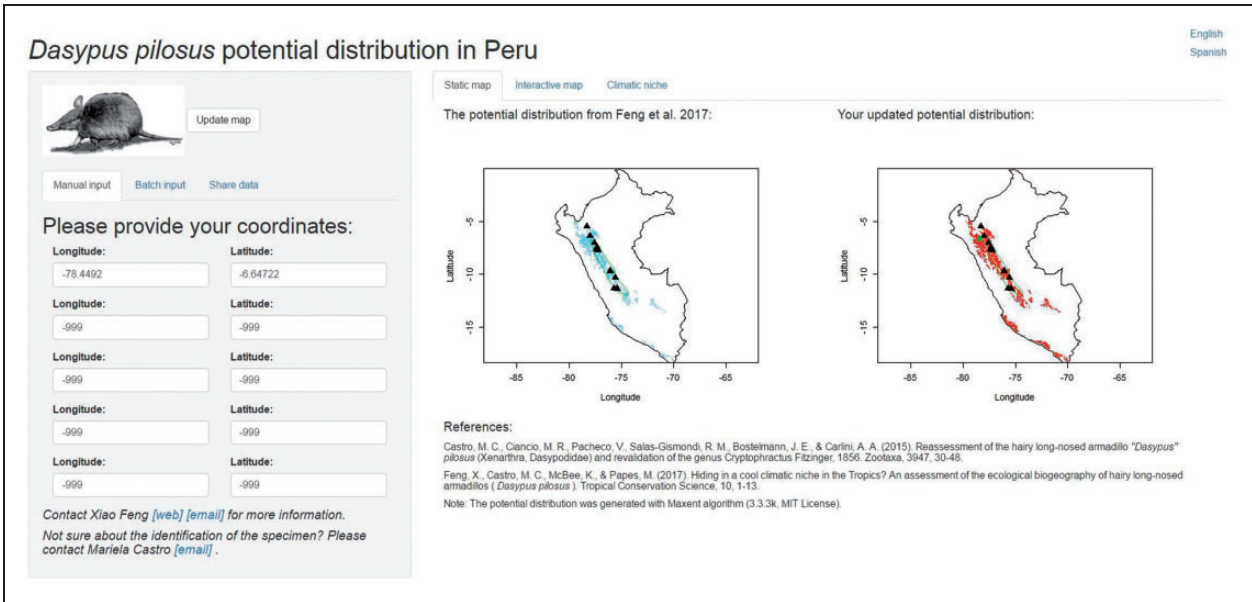


Figure 1. The default interface of Armadillo Mapper. The data input module is on the left side, where users can provide coordinates of new observations either manually or in batch mode. The right side shows the updated potential distribution (in red) compared to the potential distribution (in blue) predicted by Feng et al. (2017). The new occurrence is shown as a green triangle and the occurrences compiled in Feng et al. (2017) are shown as black triangles. The green polygon represents the International Union for Conservation of Nature range map.

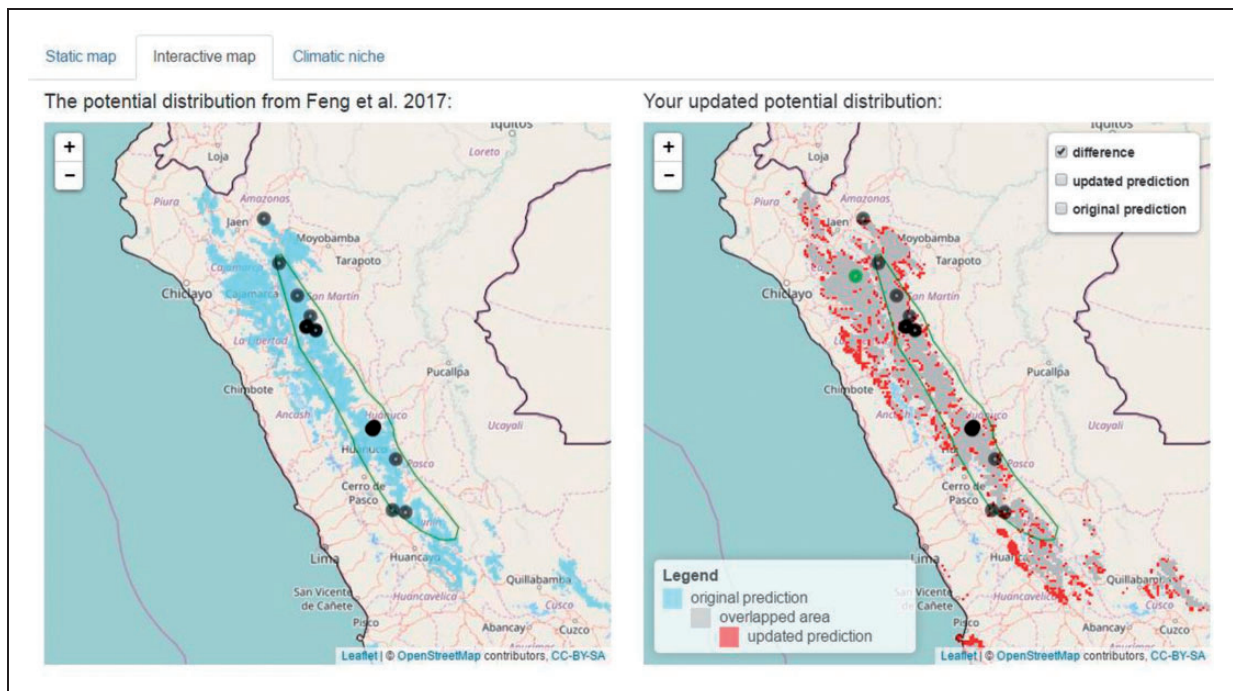


Figure 2. Implementation of the interactive map in Armadillo Mapper. The left side shows the original potential distribution and the right side shows the difference between the original and the updated potential distributions (blue if only predicted suitable by the original model, red if only predicted suitable by the updated model, or gray if predicted suitable by both models). The predictions are shown in an interactive mode, overlaid on a base map, in which users can zoom in and out. The new occurrence is shown as a green circle and the occurrences compiled in Feng et al. (2017) are shown as black circles. The green polygon represents the International Union for Conservation of Nature range map.

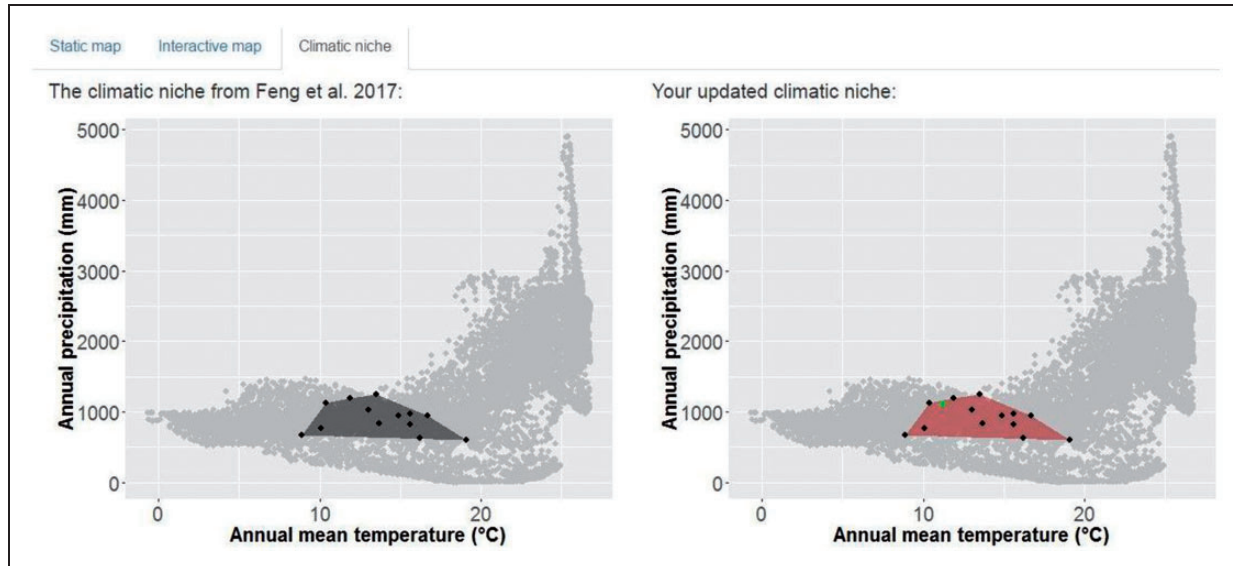


Figure 3. Illustration of climatic requirements for *Dasylops pilosus* in two-dimensional environmental space. The black points represent the climatic conditions associated with the geographic coordinates of specimens compiled in Feng et al. (2017) and the black polygon is the convex hull of the specimens, illustrating the possible boundary of the precipitation and temperature requirements of *D. pilosus*. The new occurrence is shown as a green point on the right graph and the red polygon is the updated convex hull of occurrences. The gray points represent climatic conditions associated with the pixels surrounding specimens in a two decimal degree spatial buffer.

conditions associated with previous *D. pilosus* specimens (Figure 3).

Discussion

AM provides a solution to directly and iteratively update the potential distribution map of *D. pilosus*, as new occurrences are collected. This web application does not require advanced knowledge of GIS techniques or ENM from its users and thus will be easily accessible by conservation practitioners focusing on applications of ENM.

The recent observation of *D. pilosus* lends support to the utility of the predicted potential distribution of Feng et al. (2017). Further, including newly collected occurrences will refine the prediction of potential distribution and thus better fulfill the objective of using ENM to facilitate surveys of rare species.

Compared to existing tools used in ENM, our AM is similar to Wallace (Kass et al., 2017). One notable difference is that our application is simplified in terms of user input (i.e., only new presence records) because the modeling parameters and procedures have been previously validated, so the potential distribution from our application can be easily used by other scholars who may not be experts in ENM. The framework of this application (user interface and interaction between the interface and the server) may be extended to other species, although such extension will require additional, species-specific, modeling settings.

Implications for Conservation

The various armadillo species are drawing increasing attention for research and conservation efforts. We expect that GIS and ENM will be indispensable investigative tools for synthesizing ecological and distributional knowledge. The web application we developed will shorten the time between synthesizing distributional knowledge on a computer and carrying out conservation efforts in the field. By providing an avenue to update and refine the potential distribution of *D. pilosus*, the AM web application will facilitate better targeted field surveys and relevant biogeographical and ecological studies.

Authors' Note

The computing for this project was performed at the High Performance Computing Center, Oklahoma State University.

Acknowledgments

We thank Antonio Tovar Narvaez and Segundo Sánchez Tello for sharing the photos and location information for the new occurrence of *D. pilosus* in Cajamarca Department. We also thank the anonymous reviewer for constructive comments on an earlier draft of this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research has been partially supported by FAPESP grant (Proc. 2014/23815-2) to MCC.

References

- Castro, M. C., Ciancio, M. R., Pacheco, V., Salas-Gismondi, R. M., Bostelmann, J. E., & Carlini, A. A. (2015). Reassessment of the hairy long-nosed armadillo “*Dasypus*” *pilosus* (Xenarthra, Dasypodidae) and revalidation of the genus *Cryptophractus* Fitzinger, 1856. *Zootaxa*, *3947*, 30–48.
- Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2016). *Shiny: Web application framework for R (Version 0.13.2)*. Retrieved from <https://CRAN.R-project.org/package=shiny>.
- Cheng, J., Karambelkar, B., & Xie, Y. (2017). *leaflet: Create interactive web maps with the javascript 'leaflet' library (Version 1.0.1)*. Retrieved from <http://CRAN.R-project.org/package=leaflet>
- Feng, X., Castro, M. C., McBee, K., & Papeş, M. (2017). Hiding in a cool climatic niche in the Tropics? An assessment of the ecological biogeography of hairy long-nosed armadillos (*Dasypus pilosus*). *Tropical Conservation Science*, *10*, 1–13.
- Feng, X., & Papeş, M. (2017). Physiological limits in an ecological niche modeling framework: A case study of water temperature and salinity constraints of freshwater bivalves invasive in USA. *Ecological Modelling*, *346*, 48–57.
- Fois, M., Fenu, G., Cuenca Lombraña, A., Cogoni, D., & Bacchetta, G. (2015). A practical method to speed up the discovery of unknown populations using species distribution models. *Journal for Nature Conservation*, *24*, 42–48.
- Hijmans, R. J., Eten, J. v., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J. A., . . . Shortridge, A. (2016). *raster: Geographic data analysis and modeling (Version 2.5-8)*. Retrieved from <http://cran.r-project.org/web/packages/raster/>.
- Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2013). *dismo: Species distribution modeling (Version 0.8-5)*. Retrieved from <http://CRAN.R-project.org/package=dismo>
- Kass, J. M., Vilela, B., Aeillo-Lammens, M. E., Muscarella, R., & Anderson, R. P. (2017). *Wallace: A modular platform for reproducible ecological modeling (Version 0.6.3)*. Retrieved from <https://wallaceEcoMod.github.io>
- Menon, S., Choudhury, B. I., Khan, M. L., & Peterson, A. T. (2010). Ecological niche modeling and local knowledge predict new populations of *Gymnocladus assamicus* a critically endangered tree species. *Endangered Species Research*, *11*, 175–181.
- Merow, C., Wilson, A. M., & Jetz, W. (2017). Integrating occurrence data and expert maps for improved species range predictions. *Global Ecology and Biogeography*, *26*, 243–258.
- Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. (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.
- Pebesma, E. J., & Bivand, R. S. (2005). Classes and methods for spatial data in R. *R News*, *5*, 9–13.
- Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., & Araújo, M. B. (2011). *Ecological niches and geographic distributions*. Princeton, NJ: Princeton University Press.
- Phillips, S. J., Anderson, R. P., Dudík, M., Schapire, R. E., & Blair, M. E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, *40*, 887–893.
- Phillips, S. J., & Dudík, M. (2008). Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, *31*, 161–175.
- Saupe, E. E., Barve, V., Myers, C. E., Soberón, J., Barve, N., Hensz, C. M., . . . Lira-Noriega, A. (2012). Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, *237–238*, 11–22.
- Superina, M., Pagnutti, N., & Abba, A. M. (2014). What do we know about armadillos? An analysis of four centuries of knowledge about a group of South American mammals, with emphasis on their conservation. *Mammal Review*, *44*, 69–80.
- Wetzel, R. M. (1985). Taxonomy and distribution of armadillos, Dasypodidae. In: G. G. Montgomery (ed.) *The evolution and ecology of Armadillos, Sloths, and Vermilinguas* (pp. 23–46). Washington, DC: Smithsonian Institution Press.
- Wetzel, R. M., & Mondolfi, E. (1979). The subgenera and species of long-nosed armadillos, genus *Dasypus* L. In: J. F. Eisenberg (ed.) *Vertebrate ecology in the Northern Neotropics* (pp. 43–63). Washington, DC: Smithsonian Institution Press.
- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer.