portalcasting: Supporting automated forecasting of rodent populations

Juniper L. Simonis[1,2], **Glenda M. Yenni**[1], **Ellen K. Bledsoe**[1,3], **Erica M. Christensen**[4], **Henry Senyondo**[1], **Shawn D. Taylor**[4], **Hao Ye**[1], **Ethan P. White**[1], and **S. K. Morgan Ernest**[1]

1 University of Florida 2 DAPPER Stats 3 University of Regina, Saskatchewan 4 Jornada Experimental Range

Summary

The portalcasting package provides a model development, deployment, and evaluation system for forecasting how ecological systems change through time, with a focus on a widely used long-term study of mammal population and community dynamics, the Portal Project (Brown, 1998; Brown & Davidson, 1977; Ernest, Yenni, Allington, Bledsoe, Christensen, Diaz, Goheen, et al., 2018). The software is designed to encourage the active engagement of the broad ecological modeling community in understanding the dynamics of this well-sampled empirical system and forecasting future changes. The initial infrastructure for continuous analysis and forecasting at the site provided a functional system for running existing models repeatedly (White et al., 2018). However, it did not facilitate the development and evaluation of new models. The portalcasting package combines existing R packages for modeling (e.g., JAGS), forecasting (e.g., forecast) and data management (e.g., portalr) into a continuous pipeline that produces new forecasts and allows for new model development.

The process of code development and exploration outside of a production pipeline is known as sandboxing and is key to scientific progress in a forecasting setting, as it substantially lowers the effort required for new users to contribute models. The portalcasting R package (R Core Team, 2020) simultaneously provides the engine for production forecasting for this system and a portable sandbox that allows users to set up a fully-functioning replica of the production system on their own computer for testing new models. The package also leverages a container based approach to enable fully reproducible models to be developed and tested for the production system. The associated Docker image is automatically rebuilt with each portalcasting release.

The portalcasting package automates the acquisition and updating of near real-time data; analysis of the data using models; and making, reporting, and evaluating ecological forecasts. A combination of a simple yet powerful API with advanced configuration options allow users to both quickly add new models that run using default configurations and also produce highly specific modifications to the underlying settings.

Statement of need

Accurate and up-to-date forecasts are needed in ecology for tasks ranging from endangered species management, disease epidemiology, and invasive species control. However, the field generally lacks many of the tools necessary for operationalizing ecological forecasts. As a result research on automated ecological forecasting requires years of development for each system of interest. Further, engaging the ecological modeling community in these efforts is challenging, especially as a novice. The portalcasting package helps address this gap by allowing other

researchers to quickly engage in the development of forecasting models for a widely known and used long-term ecological study with a long history of open data (Ernest, Yenni, Allington, Bledsoe, Christensen, Diaz, Geluso, et al., 2018; Ernest et al., 2009; Ernest, Yenni, Allington, Bledsoe, Christensen, Diaz, Goheen, et al., 2018; Morgan Ernest et al., 2016) and associated open source software (Christensen et al., 2019). The software will be used as a key component of a series of ecological forecasting competitions to engage the field broadly in how to make and improve ecological forecasts and the resulting models will be integrated into the production forecasting system. While portalcasting is designed for the Portal Project forecasting system, it also lays the groundwork for a generalized approach to ecological forecasting pipelines that work seamlessly in production and sandbox environments, thus facilitating broad scientific development in an ecological forecasting setting.

Acknowledgements

References

Brown, J. H. (1998). The desert granivory experiments at Portal. In W. J. Resetarits & J. Bernardo (Eds.), *Experimental ecology: Issues and perspectives* (pp. 71–95). Oxford University Press, Oxford, UK.

Brown, J. H., & Davidson, D. W. (1977). Competition between seed-eating rodents and ants in desert ecosystems. *Science*, *196*(4292), 880–882. https://doi.org/10.1126/science.196.4292.880

Christensen, E. M., Yenni, G. M., Ye, H., Simonis, J. L., Bledsoe, E. K., Diaz, R. M., Taylor, S. D., White, E. P., & Ernest, S. K. M. (2019). portalr: An R package for summarizing and using the Portal Project Data. *Journal of Open Source Software*, *4*(33), 1098. https://doi.org/10.21105/joss.01098

Ernest, S. K. M., Valone, T. J., & Brown, J. H. (2009). Long-term monitoring and experimental manipulation of a Chihuahuan desert ecosystem near Portal, Arizona, USA. *Ecology*, *90*(6), 1708–1708. https://doi.org/10.1890/08-1222.1

Ernest, S. K. M., Yenni, G. M., Allington, G., Bledsoe, E. K., Christensen, E. M., Diaz, R., Geluso, K., Goheen, J. R., Guo, Q., Heske, E., Kelt, D., Meiners, J. M., Munger, J., Restrepo, C., Samson, D. A., Schutzenhofer, M. R., Skupski, M., Supp, S. R., Thibault, K. M., … Valone, T. J. (2018). Weecology/PortalData 1.64.0. In *Zenodo.* https://doi.org/10.5281/zenodo.1475410

Ernest, S. K. M., Yenni, G. M., Allington, G., Bledsoe, E., Christensen, E., Diaz, R., Goheen, J., Guo, Q., Heske, E., Kelt, D., Meiners, J., Munger, J., Restrepo, C., Samson, D., Schutzenhofer, M., Skupski, M., Supp, S., Thibault, K., Taylor, S., … Valone, T. J. (2018). The Portal Project: A long-term study of a Chihuahuan desert ecosystem. *bioRxiv.* https://doi.org/10.1101/332783

Morgan Ernest, S. K., Yenni, G. M., Allington, G., Christensen, E. M., Geluso, K., Goheen, J. R., Schutzenhofer, M. R., Supp, S. R., Thibault, K. M., Brown, J. H., & Valone,

T. J. (2016). Long-term monitoring and experimental manipulation of a Chihuahuan desert ecosystem near Portal, Arizona (1977–2013). *Ecology*, *97*(4), 1082–1082. https://doi.org/10.1890/15-2115.1

R Core Team. (2020). *R: A language and environment for statistical computing*. https://www.R-project.org/

White, E. P., Yenni, G. M., Taylor, S. D., Christensen, E. M., Bledsoe, E. K., Simonis, J. L., & Ernest, S. K. M. (2018). Developing an automated iterative near-term forecasting system for an ecological study. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.13104