An accessible GeoComputation

Shawn W. Laffan

School of Biological, Earth and Environmental Sciences, University of New South Wales, Australia
*Email: shawn.laffan@unsw.edu.au

Researchers in GeoComputation typically have access to comparatively high powered computational systems, large amounts of data storage, ample knowledge for where and how such data can be found and, perhaps most importantly, the ability to cobble it all together into a (semi-)coherent system of scripts and code libraries, possibly using a polyglot of languages and coding paradigms.

However, many researchers outside GeoComputation and similar fields have no such access to systems and data, and are either intimidated by, or disinterested in, tool development, even in a single language environment. This issue is magnified for those outside of research environments who need and want access to model results to get their work done, but do not have access to researchers or analysts who can generate them.

Making analytical tools accessible beyond the academy is an ongoing concern across computational research domains. Cloud based systems are important for such processes, but are usually not free, and do not readily address the issues of data discovery and access.

Biodiversity is a case in point. We are in the sixth mass extinction event in Earth's history, with ongoing threats due to processes such as climate changes and land use and land cover change. Understanding the spatial distribution of biodiversity in the past, present and future is essential for understanding how it all fits together and for the allocation of scarce conservation resources. It is unusual for conservation practitioners to have skills in modelling and access to computational resources.

The Biodiversity and Climate Change Virtual Laboratory (BCCVL) and EcoCloud projects have both been developed to enable such access. Both are major collaborative project between researchers across multiple universities, with primary funding from the Australian Research Data Commons (ARDC) and its antecedent organisations. They are open source projects that are freely accessible to anyone, and provide access to algorithms and data with a high performance computing back-end.

The BCCVL (https://bccvl.org.au) provides an easy to use interface for species distribution modelling applying up to 17 algorithms for multiple species simultaneously, with these results projected into the future under nine different potential emissions scenarios (Hallgren et al. 2015). Further analysis types support the identification of hotspots of endemic species, ensemble analyses to combine SDM results from different algorithms, as well as more specialised analyses such as for species traits and migratory species. All of this can be applied across terrestrial, marine and freshwater aquatic systems, with a range of visualisations and performance metrics being provided. Many essential data sets are provided, with support for users to upload their own.

The EcoCloud (https://ecocloud.org.au) is a broader project that provides access to virtual desktops (CoESRA) and interactive coding environments (RStudio and Jupyter) backed by high performance computing environments. Having such power available means that analytical workflows can be developed with realistically sized data sets, rather than scaled-down versions that can conceal crippling performance bottlenecks. Perhaps the part of EcoCloud with the broadest appeal is its Data Explorer. This provides a search interface across a wide range of environmental data for Australia (47,000 data sets at the time of writing), with code snippets making it a simple copy-and-paste exercise to load these data into a coding environment to integrate into analytical workflows.

Platforms like the BCCVL and EcoCloud remove some of the cognitive burden needed to implement and understand models and analytical tools, and obviate many of the computational limitations faced by researchers who will ultimately use the tools developed in GeoComputation and related fields.

References

Hallgren, W., Beaumont, L., Bowness, A., Chambers, L., Graham, E., Holewa, H., Laffan, S., Mackey, B., Nix, H., Vanderwal, J., Warren, R., & Weis, G. (2016) The Biodiversity and Climate Change Virtual Laboratory: Where Ecology meets Big Data. *Environmental Modelling and Software*, 76, 182-186.