HRDS: A Python package for hierarchical raster datasets

Jon Hill[1]

1 Department of Environment and Geography, University of York, UK

Summary

Multi-scale modelling of geophysical domains requires data, such as bathymetry or topography to set up the initial conditions. These data are typically in the form of GIS rasters and can be derived from a number of sources. A single data source is commonly used which has a fixed spatial resolution. However, in multi-scale models, e.g. (Martin-Short et al., 2015), the spatial scale of the model can vary by four or more orders of magnitude, e.g. from kilometre- to sub-metre-scale. In order to use the high resolution data set in the area of highest model resolution this limited-area highest resolution data must be blended with a wider area coarse resolution dataset. A choice therefore has to be made to either sacrifice some resolution or create a very large data file. For a wide region zooming into metre-scale processes this data file could be terabytes in size when re-sampled at the resolution of the highest resolution dataset. This problem is particularly acute when using GIS tools such as `qmesh` (A. Avdis, Candy, Hill, Kramer, & Piggott, 2018) to generate meshes from contours or other derivatives of the raster data.

`HRDS` is a solution to this problem as it allows a user to extract data from an arbitrary point from a stack of rasters, each of which can have different extents and resolutions. `HRDS` is based on the `GDAL` library (GDAL/OGR contributors, 2018) so can load all common raster formats. As these rasters are unlikely to agree on the topographic/bathymetric height where they overlap, a linear distance buffer is created to smoothly blend the two datasets together. The user can specify the distance over which this buffer acts. The user gives HRDS a base raster which is low resolution but covers the whole extent of the domain to be modelled and then a stack of other rasters in priority order (typically increasing in resolution) along with a corresponding list of buffer distances. `HRDS` then calculates the linear buffers and stores these as rasters. The user can then request data at an arbitrary point which is calculated via bilinear interpolation. There are limitations in this approach in that rasters cannot be partially overlapped, all but the base raster must be entirely contained within another raster. This may be resolved in future versions.

This software solves a particular problem when using multiscale numerical models, in that in using high resolution meshes, high resolution data is required, but for spatially limited regions. By blending a heirachy of data sources, `HRDS` overcomes this problem and enables multiscale numerical problems to use spatially appropriate data to be used with minimal effort.

Acknowledgements

References

Avdis, A., Candy, A. S., Hill, J., Kramer, S. C., & Piggott, M. D. (2018). Efficient unstructured mesh generation for marine renewable energy applications. *Renewable Energy*, *116*, 842–856. doi:10.1016/j.renene.2017.09.058

GDAL/OGR contributors. (2018). *GDAL/OGR Geospatial Data Abstraction software Library*. Open Source Geospatial Foundation.

Martin-Short, R., Hill, J., Kramer, S. C., Avdis, A., Allison, P. A., & Piggott, M. D. (2015). Tidal resource extraction in the Pentland Firth, UK: Potential impacts on flow regime and sediment transport in the Inner Sound of Stroma. *Renewable Energy*, *76*(0), 596–607. doi:10.1016/j.renene.2014.11.079