

This is a repository copy of *HRDS: A Python package for hierarchical raster datasets*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/146051/>

Version: Published Version

Article:

Hill, Jonathan orcid.org/0000-0003-1340-4373 (2019) HRDS: A Python package for hierarchical raster datasets. *Journal of open source software*. 1112. ISSN: 2475-9066

<https://doi.org/10.21105/joss.01112>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

HRDS: A Python package for hierarchical raster datasets

Jon Hill¹

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

DOI: [10.21105/joss.01112](https://doi.org/10.21105/joss.01112)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Submitted: 16 November 2018

Published: 13 May 2019

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC-BY](#)).

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 multiscale 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 Development Team, 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 (Fig. 1). 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. Linear buffers are used as this makes the fewest assumptions about the data. Alternatives could include other functions, such as sigmoidal, which may be included in future releases, depending on use cases. 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. The software assumes all rasters are using the same coordinate system and datum.

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

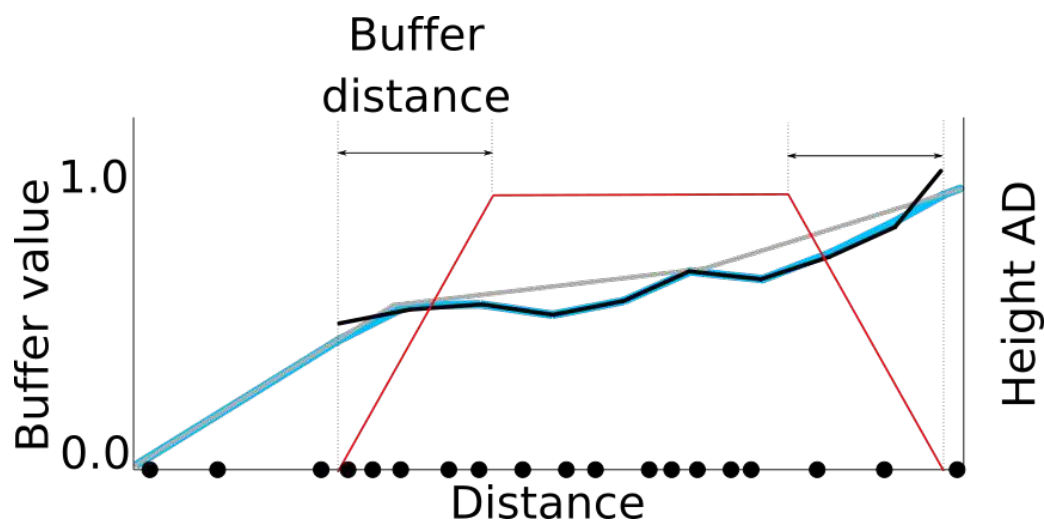


Figure 1: Conceptual image demonstrating the blending of two data sets (coarse resolution - grey, high resolution: black). The buffer generated (red line) goes from zero at the edge of the high resolution dataset to 1.0 over the user-prescribed distance. The resulting data can then be calculated at the required points (black dots on x-axis) as a function of the buffer and the two datasets (light blue thick line).

Acknowledgements

I acknowledge the University of York and World University Network for funding my sabbatical visit which enabled this work to be completed. Also thanks to Tristan Salles and Jody Webster for hosting me at the University of Sydney.

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 Development Team. (2018). *GDAL/OGR Geospatial Data Abstraction Library, Version 2.1.3.3*. Open Source Geospatial Foundation. Retrieved from <http://www.gdal.org>
- 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