The Journal of Open Source Software

AgriFoodPy: a package for modelling food systems

Juan P. Cordero [1¶], **Sarah L. Bridle** [1], **Kevin Donkers**[2], **Ian Harrison**[3], **Angelina Frankowska**[4], **Michelle Cain**[4], **Katherine Denby**[1], **Bob Doherty**[1], **Alana Kluczkovski**[1], **Jez Frendenburgh**[5], **Neil Ward**[5], **Ximena Schmidt**[6], **Christian Reynolds**[7], **Jacqueline Silva**[8], **and Aled Jones**[9]

1 Department of Environment and Geography Wentworth Way, University of York, Heslington, York, YO10 5NG, UK **2** **3** **4** **5** Vice Chancellor's Office, University of East Anglia, Norwich, NR4 7TJ UK **6** **7** **8** ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

`AgriFoodPy` is an open-source Python package for interoperability of datasets, simulation and modelling of agrifood systems. By employing xarray (Hoyer & Hamman, 2017) as the primary data structure, AgriFoodPy provides access to a range of available agrifood data through manually formated arrays, including geospatial, time series and tabular data from different open sources. The framework is region agnostic and will provide facilities to model and simulate arbitrary data regardless of its geographic origin.

Open-source code and community development will allow a transparent view into analysis choices and data sources, which can help provide trustworthy evidence-based support to data-driven policy making. AgriFoodPy is developed and maintained by a diverse community of domain experts with a focus on software sustainability and interoperability.

Version 0.1 provides access global agrifood datasets including geospatial land use and classification data (Morton, 2022), food supply (FAO, 2022), life cycle assestment (Poore & Nemecek, 2018) and population data(Nations", 2022). Also included is a library of intervention models across a range of observables and indicators, which connects with pre-existing atmospheric, land use change, socio-economic and human health models.

Future releases will provide access to more models and community contributed datasets formatted using xarray, as well as implementing a pipeline manager to perform end-to-end simulations of agrifood systems and predictions of diferent scenarios.

Statement of need

Providing food for an ever-growing population while reducing the impact of human activity on the environment has become one of the main current global challenges. Local and and intergovernmental independent committees (https://www.theccc.org.uk/) (https://www.ipcc.ch/) have reported the importance of food production on climate change. The scenarios and projections in their reports also highlighted the need for precise and transparent modelling of the different aspects of the food system to help stakeholders understand the effects of consumption patterns and farming practices.

Coordinated efforts to achieve a sustainable food system must originate from effective policy-making based on evidence, careful choice of metrics and indicators to describe the state of the food system, and accurate estimates of how these metrics change under different scenarios.

Existing datasets and analisys software usually rely on non standardized data structures and predominantly closed-source code. This difficults research and independent scrutiny of food

system inverventions projections and the impact of policy on environmental, socio-economic and health indicators. Moreover, this forces researchers to routinely expend significant effort replicating or re-developing existing code to reduce and analyse data. Additionally, the opacity of some data sources and analysis choices make it difficult to draw conclusions from equivalent comparisons between different interventions and policy decisions.

Several open initiatives exist focused on analysis and modelling of agrifood and enviromental related data, Research community has also developed open-source python packages that adress some individual aspects of modelling agrifood systems, such as geospatial imaging (e.g. `Geopandas` (Jordahl et al., 2020), `Rasterio` (Gillies & others, 2013--)), atmospheric and climate modelling (`Fair` (Leach et al., 2021)). Other open softwares also exist in other languages, including agriculture and farming (`APSIM` (Holzworth et al., 2014)), life cycle assestment (`OpenLCA`, www.openlca.org).

AgriFoodPy will provide a consistent standard for agrifood data distribution, while also allowing external models and packages to coexist and interoperate to allow a holistic approach to agrifood modelling.

Acknowledgements

References

FAO. (2022). *FAO. FAOSTAT. License: CC BY-NC-SA 3.0 IGO. Extracted from: Https://www.fao.org/faostat/en/#data/FBS. Date of access: 22-06-2023*.

Gillies, S., & others. (2013--). *Rasterio: Geospatial raster i/o for Python programmers*. Mapbox. https://github.com/rasterio/rasterio

Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., Chenu, K., van Oosterom, E. J., Snow, V., Murphy, C., Moore, A. D., Brown, H., Whish, J. P. M., Verrall, S., Fainges, J., Bell, L. W., Peake, A. S., Poulton, P. L., Hochman, Z., … Keating, B. A. (2014). APSIM – evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software*, *62*, 327–350. https://doi.org/https://doi.org/10.1016/j.envsoft.2014.07.009

Hoyer, S., & Hamman, J. (2017). Xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, *5*(1). https://doi.org/10.5334/jors.148

Jordahl, K., Bossche, J. V. den, Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., … Leblanc, F. (2020). *Geopandas/geopandas: v0.8.1* (Version v0.8.1). Zenodo. https://doi.org/10.5281/zenodo.3946761

Leach, N. J., Jenkins, S., Nicholls, Z., Smith, C. J., Lynch, J., Cain, M., Walsh, T., Wu, B., Tsutsui, J., & Allen, M. R. (2021). FaIRv2.0.0: A generalized impulse response model for climate uncertainty and future scenario exploration. *Geoscientific Model Development*, *14*(5), 3007–3036. https://doi.org/10.5194/gmd-14-3007-2021

Morton, C. G. ;O'Neil., R. D.;Marston. (2022). *Land cover map 2020 (1km summary rasters, GB and n. ireland)*. NERC EDS Environmental Information Data Centre. https://doi.org/10.5285/d6f8c045-521b-476e-b0d6-b3b97715c138

Nations", "United. (2022). *United nations, department of economic and social affairs, population division (2022). World population prospects 2022: Summary of results. UN DESA/POP/2022/TR/NO. 3*.

Poore, J., & Nemecek, T. (2018). Reducing food’s environmental impacts through producers and consumers. *Science*, *360*(6392), 987–992. https://doi.org/10.1126/science.

87 aaq0216