

AgriFoodPy: a package for modelling food systems

Juan P. Cordero¹, Kevin Donkers², Ian Harrison³, Sarah L. Bridle¹, Angelina Frankowska⁴, Michelle Cain⁴, Neil Ward⁵, Jez Frendenburgh⁵, Edward Pope², Alana Kluczkovski⁶, Ximena Schmidt⁷, Jacqueline Silva⁸, Christian Reynolds⁹, Katherine Denby⁶, Bob Doherty¹⁰, and Aled Jones¹¹

¹ Department of Environment and Geography Wentworth Way, University of York, Heslington, York, YO10 5NG, UK ² Met Office Hadley Centre, Met Office, Fitz Roy Road, Exeter, Devon, UK ³ School of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, UK ⁴ Centre for Environmental and Agricultural Informatics, School of Water, Energy and Environment, Cranfield University, Cranfield MK43 0AL, UK ⁵ School of Environmental Sciences, University of East Anglia, Norwich, UK ⁶ Centre for Novel Agricultural Products (CNAP), Department of Biology, University of York, York, YO10 5DD, UK ⁷ Equitable Development and Resilience Research Group, College of Engineering, Design and Physical Science, Brunel University London, London, UB8 3PH, UK ⁸ Global Academy of Agriculture and Food Systems, The University of Edinburgh. Charnock Bradley Building, Easter Bush Campus, EH25 9RG. ⁹ Centre for Food Policy, City, University of London, Northampton Square, London, EC1V 0HB, UK ¹⁰ School for Business and Society, University of York ¹¹ Global Sustainability Institute, Anglia Ruskin University, Cambridge CB1 1PT, UK ¶ Corresponding author

DOI: 10.xxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970

Published: unpublished

License

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

Summary

AgriFoodPy is an open-source Python package for dataset manipulation, interoperability, simulation, and modeling of agrifood systems.

By employing xarray (Hoyer & Hamman, 2017) as the primary data structure, AgriFoodPy provides access methods to manipulate tabular data by extending xarray functionality via accessors.

A separate repository, agrifoodpy_data, is also maintained in parallel to provide access to local and global agrifood datasets, including geospatial land use and classification data (Morton, 2022), food supply (FAO, 2022), life cycle assessment (Poore & Nemecek, 2018), and population data (Nations, 2022). The AgriFoodPy framework is region-agnostic and provides facilities to model and simulate processes and interventions regardless of their geographic origin. As an example, current functionality can be used to construct predictions on the effect of dietary change on greenhouse emissions, the carbon sequestration of transforming agricultural land use, or consequences of modifying farming practices on operational costs.

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 for data-driven policymaking. AgriFoodPy is developed and maintained by a diverse community of domain experts with a focus on software sustainability and interoperability.

Version 0.1 provides basic table manipulation methods to extend the coordinate dimensions of xarray Datasets and DataArrays, extract summary statistics and indicators, plotting methods to analyze and display data. It also includes a library of intervention models such as diet change, afforestation and agroecology and land carbon sequestration. We have also built an interface to external an atmospheric model (Leach et al., 2021). We plan to extend this library of interfaces to incorporate other open source tools, including socioeconomic, land use, environmental, health, and policy modelling packages.

44 Future releases will provide access to more models and community-contributed datasets
45 formatted using xarray. Additionally, AgriFoodPy will implement a pipeline manager to perform
46 end-to-end simulations of agrifood systems, which can be used to speed up the comparison of
47 multiple scenarios.

48 Statement of need

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

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

60 Existing datasets and analysis software usually rely on non-standardized data structures and
61 predominantly closed-source code. This hinders research and independent scrutiny of food
62 system intervention projections and the impact of policy on environmental, socio-economic,
63 and health indicators. Moreover, this forces researchers to routinely expend significant effort
64 replicating or re-developing existing code to reduce and analyze data. Additionally, the opacity
65 of some data sources and analysis choices makes it difficult to draw conclusions from equivalent
66 comparisons between different interventions and policy decisions.

67 Few open initiatives exist focused on analysis and modelling of agrifood and environmental
68 related data, e.g., The Environmental Data Science book (<https://edsbook.org/>). The
69 research community has developed open-source packages that address some individual aspects of
70 modelling agrifood systems, such as geospatial imaging (e.g., Geopandas ([Jordahl et al., 2020](#)),
71 Rasterio ([Gillies & others, 2013](#))), atmospheric and climate modeling (Fair ([Leach et al., 2021](#)))
72 in Python, and other open softwares in other languages, such as agriculture and farming
73 (APSIM ([Holzworth et al., 2014](#))) and life cycle assessment (OpenLCA, www.openlca.org).

74 AgriFoodPy provides a consistent standard for agrifood data distribution, while also allowing
75 external models and packages to coexist and interoperate, allowing a holistic approach to
76 agrifood modeling.

77 Plans for future use in research and communication include the FixOurFood agrifood calculator
78 (<https://fixourfood.streamlit.app/>), an interactive modelling tool to evaluate the effect of food
79 system transformations in the UK. There are also plans to publish a paper on global diets their
80 social and environmental impacts.

81 Acknowledgements

82 JPC, SLB, AF, MC, KD, BD, AK and FixOurFood's work was supported by the UKRI
83 Transforming Food Systems Strategic Priority Fund (grant number BB/V004581/1). NW &
84 JF are grateful for funding from the AFN Network+ (UKRI Agri-food for Net Zero Network+)
85 Grant Award EP/X011062/1 CR was funded through Transforming the UK Food System for
86 Healthy People and a Healthy Environment SPF Programme, Grant Award BB/V004719/1
87 Healthy soil, Healthy food, Healthy people (H3). This article has benefited from comments
88 and suggestions of the following people: Daniel Lewis, Joe Kennedy, Nada Saidi, Adam Amara.

References

- 89
- 90 FAO. (2022). *FAO. FAOSTAT. License: CC BY-NC-SA 3.0 IGO. Extracted from:*
 91 *https://www.fao.org/faostat/en/#data/FBS. Date of access: 22-06-2023.*
- 92 Gillies, S., & others. (2013--). *Rasterio: Geospatial raster i/o for Python programmers.*
 93 Mapbox. <https://github.com/rasterio/rasterio>
- 94 Holzworth, D. P., Huth, N. I., deVoil, P. G., Zurcher, E. J., Herrmann, N. I., McLean,
 95 G., Chenu, K., van Oosterom, E. J., Snow, V., Murphy, C., Moore, A. D., Brown, H.,
 96 Whish, J. P. M., Verrall, S., Fainges, J., Bell, L. W., Peake, A. S., Poulton, P. L.,
 97 Hochman, Z., ... Keating, B. A. (2014). APSIM - evolution towards a new generation
 98 of agricultural systems simulation. *Environmental Modelling & Software*, 62, 327–350.
 99 <https://doi.org/10.1016/j.envsoft.2014.07.009>
- 100 Hoyer, S., & Hamman, J. (2017). Xarray: N-D labeled arrays and datasets in Python. *Journal*
 101 *of Open Research Software*, 5(1). <https://doi.org/10.5334/jors.148>
- 102 Jordahl, K., Bossche, J. V. den, Fleischmann, M., Wasserman, J., McBride, J., Gerard,
 103 J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D.,
 104 Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur,
 105 A., ... Leblanc, F. (2020). *Geopandas/geopandas: v0.8.1 (Version v0.8.1).* Zenodo.
 106 <https://doi.org/10.5281/zenodo.3946761>
- 107 Leach, N. J., Jenkins, S., Nicholls, Z., Smith, C. J., Lynch, J., Cain, M., Walsh, T., Wu, B.,
 108 Tsutsui, J., & Allen, M. R. (2021). FaIRv2.0.0: A generalized impulse response model
 109 for climate uncertainty and future scenario exploration. *Geoscientific Model Development*,
 110 14(5), 3007–3036. <https://doi.org/10.5194/gmd-14-3007-2021>
- 111 Morton, C. G. ;O'Neil, R. D.;Marston. (2022). *Land cover map 2020 (1km summary*
 112 *rasters, GB and n. ireland).* NERC EDS Environmental Information Data Centre. <https://doi.org/10.5285/d6f8c045-521b-476e-b0d6-b3b97715c138>
- 114 Nations", "United. (2022). *United nations, department of economic and social affairs,*
 115 *population division (2022). World population prospects 2022: Summary of results. UN*
 116 *DESA/POP/2022/TR/NO. 3.*
- 117 Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers
 118 and consumers. *Science*, 360(6392), 987–992. <https://doi.org/10.1126/science.aag0216>