




1 interflow: A Python package to organize, calculate, and
2 visualize sectoral interdependency flow data

3 **Kendall L. Mongird¹, Konstantinos Oikonomou¹, Juliet S. Homer¹, and**
4 **Jennie S. Rice¹**

5 **¹** Pacific Northwest National Laboratory, Richland, WA, USA

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

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

Editor: 

Submitted: 30 March 2022

Published: unpublished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](https://creativecommons.org/licenses/by/4.0/)).

6 **Summary**

7 Many economic sectors rely on an uninterrupted “upstream” supply of a resource to conduct
8 their primary functions, leaving them vulnerable to adverse effects should that resource flow be
9 interrupted or compromised ([OECD, 2017](#); [U.S. EPA, 2010](#)). Well-known examples of these
10 relationships include water demand by the energy sector (e.g., thermoelectric cooling for nuclear
11 generation) ([Grubert & Sanders, 2018](#); [Webber, 2017](#)) and energy demand by the water sector
12 (e.g., electricity required to treat or move water in the public water sector) ([Congressional](#)
13 [Research Service, 2017](#)) though many others exist. Being able to calculate and document
14 these interdependencies and evaluate where the greatest cross-sectoral intensities and flows
15 exist can reveal opportunities to enhance the overall network. Despite the implications and
16 potential impacts, however, these interconnections and flows have been historically complex to
17 analyze and understand.

18 The interflow package provides a flexible tool to organize, calculate, and visualize (using
19 Sankey diagrams and other visualizations) sectoral interdependency flows for multiple subsectors
20 and resources (Figure 1). This tool can help decision-makers, researchers, and other audiences
21 more easily pull meaning from these interdependencies to reveal multi-faceted opportunities
22 and risks. interflow can help investigate questions such as (1) which sectors have high
23 cross-resource dependencies, (2) how does demand for a resource in various sectors compare
24 across regions, and (3) where the sectoral and regional opportunities are for enhanced efficiency,
25 security, and resiliency.

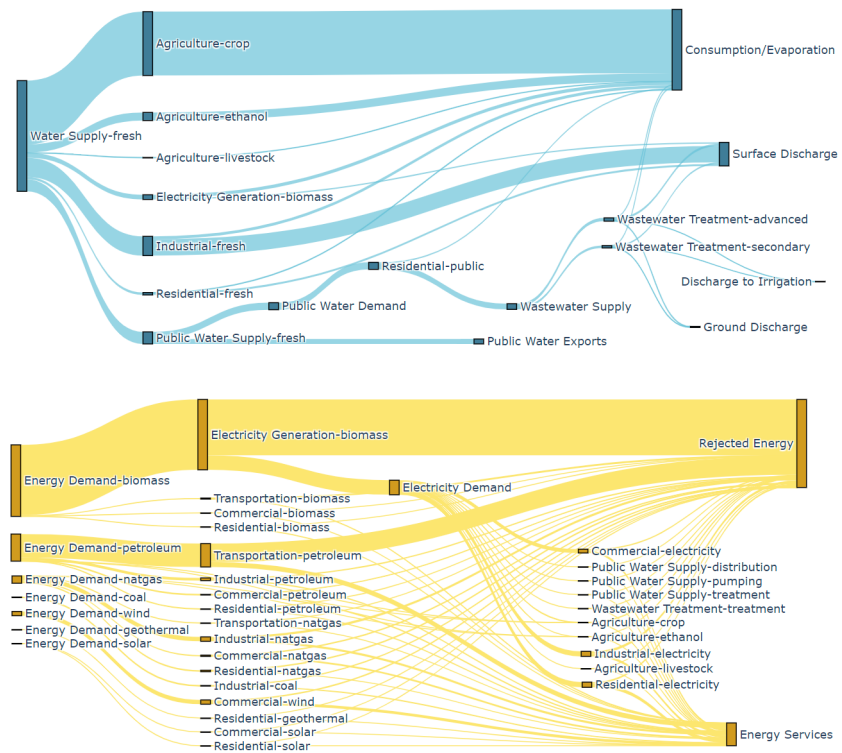


Figure 1: Example Sankey Diagram Demonstrating the Flow of a Resource Across Sectors.

Statement of Need

To the best of our knowledge, there is no open-source software option available that facilitates the calculation, organization, and visualization of multiscale, multisector interdependencies in one place. Software exists for visualizing pre-calculated sectoral flow values as Sankey diagrams such as the proprietary e!Sankey software (iPoint-systems gmbh, 2022) or the Python library matplotlib (Hunter, 2007), but these tools do not offer a way to calculate the demand of a cross-resource type and build out new sectoral connections based on the result. The interflow package aims to fill this gap and serve as a flexible and open-source option for conducting multi-resource sectoral interdependency data calculations, producing multisectoral datasets, and visualizing the results.

Sectoral interdependency analysis itself is not a new area of research (Curmi et al., 2013; Greenberg et al., 2017; Liu et al., 2016), but the publications in this area typically return only the end-product of their analysis (e.g., output values or diagrams). The detailed methodology or algorithms developed to conduct their calculations are often not provided in an accessible or transparent manner. Researchers looking to build upon or modify the existing assumptions and data are left to redevelop the calculation structure from scratch as a result. The interflow package provides a consistent and open-source calculation framework to enable reproducibility.

Design and Methodology

The interflow package iterates through a user-provided tabular input data to (1) collect known resource flow values (e.g., water demand) between provided sectors, (2) calculate new sector flow values for a secondary resource type from the initial flow (e.g., energy demand based on the water flow) using provided cross-resource intensity coefficients, and (3) build upstream and downstream sector connections to carry those calculated and collected flows.

49 Collected and calculated flows are then compiled in a way that can be used in a variety of
50 data visualization functions offered in the package.

51 The basic methodology described above is repeated for all regions provided by the user in
52 the input data. Though the interflow package comes equipped with extensive sample data
53 to evaluate U.S. county-level water and energy flows across various sectors, it can conduct
54 analysis for any region (e.g., country, province), any sector or group of sectors (e.g., electricity
55 generation, agriculture), and any group of resources (e.g., water, energy, food, carbon, land)
56 and is limited only by the input data that the user provides. More information on the input data
57 requirements can be found in the [generalizability documentation](#). All code used to generate
58 the sample water and energy input data for all US counties is additionally included as part of
59 the package.

60 A Pandas DataFrame ([McKinney & others, 2010](#)) containing collected and calculated flow
61 values between sectors for each region and resource type is returned as output from the
62 `calculate()` function. The DataFrame output can be directly used with other package
63 functions to generate a variety of visualizations which each utilize the Plotly Python package
64 ([Plotly Technologies Inc., 2015](#)) including (1) Sankey diagrams showing the network of flows
65 across sectors for a chosen region, (2) stacked bar charts of inflow and outflow values for
66 sectors in a region, and (3) a choropleth map to compare flow values across regions. The
67 visualizations can be used to compare dependencies across and between sectors at various levels
68 of sub-sector granularity. Given that interflow collects and calculates values for each region
69 provided in the input data, the output can additionally be used for region-wise comparisons of
70 flows and intensities. For more information on the key outputs see the [outputs documentation](#).

71 Acknowledgements

72 This research was supported by the U.S. Department of Energy, Water Power Technologies
73 Office, as part of research in Integrated Water Power Resilience.

74 References

- 75 Congressional Research Service. (2017). *Energy-water nexus: The water sector's energy use*.
76 [PDF]. <https://sgp.fas.org/crs/misc/R43200.pdf>
- 77 Curmi, E., Fenner, R., & Richards, K. et al. (2013). Visualising a stochastic model of
78 californian water resources using sankey diagrams. *Water Resources Management*, 27,
79 3035–3050. <https://doi.org/10.1007/s11269-013-0331-2>
- 80 Greenberg, H. R., Simon, A. J., Singer, S. L., & Shuster, E. P. (2017). *Development of energy-*
81 *water nexus state-level hybrid sankey diagrams for 2010*. [PDF]. https://flowcharts.llnl.gov/content/assets/docs/2010_United-States_EnergyWater.pdf
- 82
- 83 Grubert, E., & Sanders, K. T. (2018). Water Use in the United States Energy System: A
84 National Assessment and Unit Process Inventory of Water Consumption and Withdrawals.
85 *Environmental Science & Technology*, 52(11). <https://doi.org/10.1021/acs.est.8b00139>
- 86 Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science &*
87 *Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- 88 iPoint-systems gmbh. (2022). *e!Sankey - show the flow*. <https://www.ifu.com/e-sankey/>
- 89 Liu, Y., Hejazi, M., Kyle, P., Kim, S. H., Davies, E., Miralles, D. G., Teuling, A. J., He, Y.,
90 & Niyogi, D. (2016). Global and regional evaluation of energy for water. *Environmental*
91 *Science & Technology*, 50(17), 9736–9745. <https://doi.org/10.1021/acs.est.6b01065>
- 92 McKinney, W., & others. (2010). Data structures for statistical computing in python.
93 *Proceedings of the 9th Python in Science Conference*, 445, 51–56.

- 94 OECD. (2017). *The land-water-energy-nexus: Biophysical and economic conse-*
95 *quences.* [https://read.oecd-ilibrary.org/environment/the-land-water-energy-nexus_](https://read.oecd-ilibrary.org/environment/the-land-water-energy-nexus_9789264279360-en#page4)
96 [9789264279360-en#page4](https://read.oecd-ilibrary.org/environment/the-land-water-energy-nexus_9789264279360-en#page4)
- 97 Plotly Technologies Inc. (2015). *Collaborative data science.* Plotly Technologies Inc. [https://](https://plot.ly)
98 plot.ly
- 99 U.S. EPA. (2010). *Understanding water sector interdependencies.* [PDF]. [https://nepis.epa.](https://nepis.epa.gov/Exe/ZyPDF.cgi/60001QB5.PDF?Dockkey=60001QB5.PDF)
100 [gov/Exe/ZyPDF.cgi/60001QB5.PDF?Dockkey=60001QB5.PDF](https://nepis.epa.gov/Exe/ZyPDF.cgi/60001QB5.PDF?Dockkey=60001QB5.PDF)
- 101 Webber, M. E. (2017). *Chapter 2.2.6 - Water for Electricity Generation in the United States.*
102 *In Competition for water resources: Experiences and management approaches in the US*
103 *and europe.* Elsevier. ISBN: 9780128032374

DRAFT