Argus: An interactive application to enable scientific discovery through multi-sector and multi-scale visual analytics

Zarrar Khan[1], **Sean X. Tang**[1], **Anna M. Warmka**[1], **Chris R. Vernon**[1], **Mengqi Zhao**[2], **and Thomas Wild**[1, 2]

1 Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA **2** Earth System Science Interdisciplinary Center (ESSIC), University of Maryland, College Park, MD, USA

Summary

`Argus` is an R Shiny App to interactively visualize data across scenarios, parameters, and regions. Argus is designed to ingest simple `.csv` data tables which can be loaded directly into the applciation or from a `url` if the data is hosted online. The platform comes preloaded with various shapefiles for countries, states, river basins and in the case of the United States, U.S. counties. The platform processes the input data into interactive maps, charts and tables to allow users to easily explore their data across regions, parameters, time periods and scenarios. The primary purpose of Argus is for users to be able to highlight key messages from their datasets and to this end the application allows users to subset their data by choosing relevant scenarios, regions and time periods, annotate each page with storylines and then save the state of the application via bookmarks to be shared as urls or `.rds` files which can be loaded back up in the application at a later time and place. The application is designed to be used as a tool by scientists to easily curate and share large datasets with their audiences, in an impactful and interacive way. While Argus can be used for any data that is spatial and temporal in nature the application has been developed specifically for large multi-sector global modeling outputs such as those from the open source Global Change Analysis Model (GCAM) model (Calvin et al., 2019).
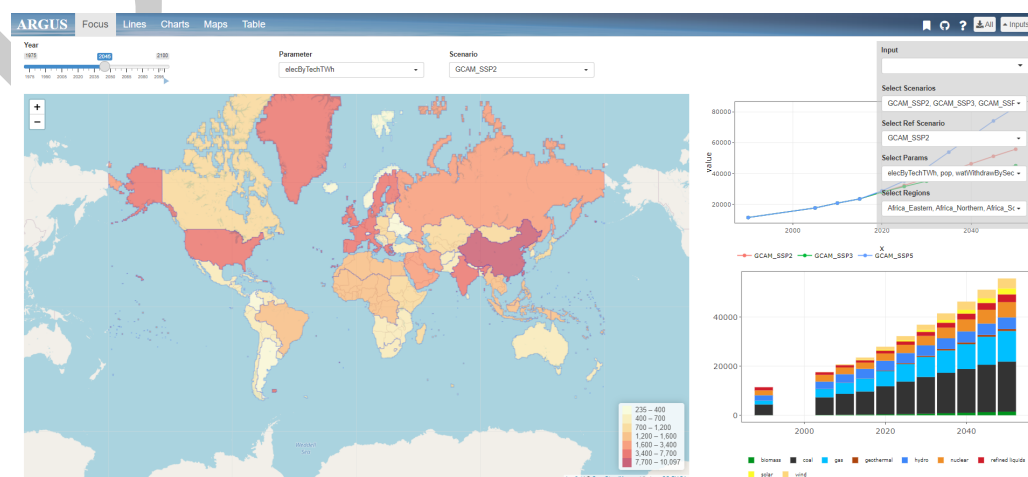
Figure 1: `Argus` landing page

26 `Argus` can be accessed via this [Home Page](). A detailed [User Guide]() is also available which
27 walks users through all the features of `Argus`.

Statement of need

29 Global multi-sector models have advanced significantly over the years and continue to push
30 the boundaries of spatial, sectoral and temproal resolution and detail. This push towards
31 more detail results in increasing complextiy as well as a larger number of both inputs and
32 outputs. Output databases from these models (e.g. GCAM (Calvin et al., 2019), MESSAGE
33 (Huppmann et al., 2019)) can have up to several GB of data capturing combinations of
34 thousands of sectors and sub-sectors; subregions ranging from countries to cities to river
35 basins; and time periods ranging from decades to hours. Climate models outputs such as
36 those from the Weather Research & Forecasting (WRF) model (Powers et al., 2017) or the
37 Coupled Model Intercomparison Project Phase (CMIP) series (Eyring et al., 2016) can be
38 even larger and range up to several Tera Bytes of data at globally gridded resolutions.

39 A key challenge that continues to become more critical with this increasing complexity of
40 models is the ability to synthesize the outputs and extract relevant trends and messages.
41 Visual and interactive communications are a particularly effective method of delivering key
42 messages from complex topics and large datasets, with audeinces having been shown to have
43 a higher retention of knowledge and comprehension of ideas when using such tools Janvrin et
44 al. (2014).

45 In response to this, interactive data visualization dashboards are starting to be used across
46 several fields of study to analyze large, complex data sets. Examples include the World Re-
47 sources Institute's (WRI's) WRI Aqueduct(WRI, 2021), the International Institute for Applied
48 Systems Analysis's (IIASA's) [Global Hotspots Explorer]() (IIASA, 2021), the Pacific North-
49 west National Lab's (PNNL's) [Hector UI]() (Evanoff et al., 2020), the Intergovernmental Panel
50 on Climate Change's (IPCC's) [WGI Interactive atlas]() for the IPCC Sixth Assessment Report
51 [] and the Model for the Assessment of Greenhouse Gas Induced Climate Change ([MAG-
52 ICC]())(Meinshausen et al., 2011). These interactive applciations have an advantage over static
53 applications because they allow users to choose and subset information that is most relevant
54 and useful for their own purposes (Janvrin et al., 2014).

55 `Argus` addresses these key issues by allowing users to interactively visualize and subset relevant
56 portions of their datasets as well as annotate the outputs in order to deliver a final product
57 that highlights key messages and storylines from complex and large datasets. Users are also
58 able to save the state of Argus at any point in time so they can return to it later on further
59 modification as well as for easy sharing with other users and audiences, who can then interact
60 with the data themselves. Additionally, while most of other applications come with limited
61 preloaded datasets (with good reason as they are meant to be viewers for particular data),
62 Argus is designed to be used with user datasets and is thus not restricted to any particular
63 model or type of analysis.

Functionality

65 An Argus user can choose to input their own spatial and temporal data via a URL or .csv
66 file. In addition `Argus` has also been customized to be used directly with the Global CHange
67 Analysis Model output databases (Calvin et al., 2019). Users can at any time change which
68 scenarios are selected, set the reference scenario, and select or deselect different parameters
69 and regions. Changing these selections will instantaneously change the data visualizations
70 throughout the application.

Argus includes several functions for data visualization, separated as different tabs within the application. Note that all functions are explained in full detail in the User Guide, which includes documentation for each individual tab.

- Focus: The focus tab gives an overview of the inputted data, and the user can select a year, parameter, and scenario they would like to visualize. Based on the selections, a world map, comparative line plot, and categorical bar chart will populate.
- Lines: The lines tab creates line plots of each parameter by scenario over time. Within this tab there is also a compare regions option, which will create several line graphs for each parameter, broken up by region.
- Charts: The charts tab shows bar charts for each parameter by scenario over time, breaking the parameters up categorically. Within this tab, there are also difference charts that compare each scenario to the specified reference scenario.
- Maps: The maps tab shows maps for each parameter by scenario, splitting up the map by the desired regions. In this tab, there is also an option to look at the regional differences between each scenario compared to the reference scenario.
- Table: The table tab displays all inputted data, within which the user can search for specific values using the search bar. Additionally, the table can be filtered in each column.
- Bookmarks: Argus also includes a function to create a bookmark to save any changes made and allow others to view the application exactly how the user left it. Detailed instructions on creating, sharing, and loading bookmarks can be found in the Bookmarks section of the User Guide.
- Preloaded Data: `Argus` comes with a number of preloaded datasets which will be updated regularly. These include curated data sets from official GCAM model releases as well other datasets of interests.
- Data Story: Each tab in `Argus` has a collapsable `Data Story` button on top which allows users to annotate the visuals seen on the particular tab in order to highlight particular aspects of what is being seen.

Within Argus, the user can also download all figures created based on their data, or they can choose to download figures from specific tabs.

Finally, the application is continually being developed to support the needs of its users, and some of the features described may change. Additionally, new features are likely to be included in future releases. Updated documentation of all features and their functionality will always be available in the User Guide.

Acknowledgements

References

Calvin, K., Patel, P., Clarke, L., Asrar, G., Bond-Lamberty, B., Cui, R. Y., Vittorio, A. D., Dorheim, K., Edmonds, J., Hartin, C., Hejazi, M., Horowitz, R., Iyer, G., Kyle, P., Kim, S.,

114 Link, R., McJeon, H., Smith, S. J., Snyder, A., … Wise, M. (2019). GCAM v5.1: Repre-
115 senting the linkages between energy, water, land, climate, and economic systems. *Geosci-*
116 *entific Model Development*, *12*(2), 677–698. https://doi.org/10.5194/gmd-12-677-2019

117 Evanoff, J. E., Vernon, C. R., Waldhoff, S. T., Snyder, A. C., & Hartin, C. A. (2020).
118 Hectorui: A web-based interactive scenario builder and visualization application for the
119 hector climate model. *The Journal of Open Source Software*, *5*(PNNL-SA-156861). https:
120 //doi.org/10.21105/joss.02782

121 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor,
122 K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6)
123 experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–
124 1958.

125 Huppmann, D., Gidden, M., Fricko, O., Kolp, P., Orthofer, C., Pimmer, M., Kushin, N.,
126 Vinca, A., Mastrucci, A., Riahi, K., & others. (2019). The MESSAGEix integrated as-
127 sessment model and the ix modeling platform (ixmp): An open framework for integrated
128 and cross-cutting analysis of energy, climate, the environment, and sustainable develop-
129 ment. *Environmental Modelling & Software*, *112*, 143–156.

130 IIASA. (2021). *Global Hotspots explorer*. International Institute for Applied Systems Analysis.
131 https://hotspots-explorer.org/

132 Janvrin, D. J., Rachke, R. L., & Dilla, W. N. (2014). Making sense of complex data using
133 interactive data visualization. *Journal of Accounting Education*, *32*, 31–48. https://doi.
134 org/10.1016/j.jaccedu.2014.09.003

135 McIntyre, D. R., & Wolff, F. G. (1998). An experiment with WWW interactive learning in
136 university education. *Computers & Education*, *31*, 255–264. https://doi.org/10.1016/
137 S0360-1315(98)00025-6

138 Meinshausen, M., Raper, S. C., & Wigley, T. M. (2011). Emulating coupled atmosphere-ocean
139 and carbon cycle models with a simpler model, MAGICC6–part 1: Model description and
140 calibration. *Atmospheric Chemistry and Physics*, *11*(4), 1417–1456. https://doi.org/10.
141 5194/acp-11-1417-2011

142 Otten, J. J., Cheng, K., & Drewnowski, A. (2015). Infographics and public policy: Using
143 data visualization to convey complex information. *Health Affairs*, *13*(11), 1901–1907.
144 https://doi.org/10.1377/hlthaff.2015.0642

145 Powers, J. G., Klemp, J. B., Skamarock, W. C., Davis, C. A., Dudhia, J., Gill, D. O., Coen, J.
146 L., Gochis, D. J., Ahmadov, R., Peckham, S. E., & others. (2017). The weather research
147 and forecasting model: Overview, system efforts, and future directions. *Bulletin of the*
148 *American Meteorological Society*, *98*(8), 1717–1737.

149 WRI. (2021). *WRI Aqueduct*. World Resources Institute. https://www.wri.org/aqueduct