Leveraging open-source tools and development best practices for IAM analysis in CMIP6 & IPCC SR1.5

International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Matthew J. Gidden and Daniel Huppmann; International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria

Introduction

Integrated Assessment Models (IAMs) have a long and important history in the modeling and analysis of critical drivers of social and environmental change. Their use is varied, but perhaps most iconically displayed in the many IPCC assessment and special reports. However, calls to increase the understandability, usability, and transparency of both models and results have been made in recent years. The IAM community is beginning to leverage open source software and collaborative development practices, as modeling teams have begun not only to release model codes and workflows under open source licenses (e.g., GCAM and MESSAGEix) but also to build community tools that help coordinate model intercomparison exercises and allow for standard visualization of results. We here present two efforts to develop and utilize novel toolboxes that bring collaborative and open tool development to the IAM community: the consistent harmonization of IAM results across multiple modeling teams for CMIP6 utilizing the *aneris* Python library (https://github.com/iiasa/aneris), and the analysis and visualization of scenarios for the IPCC 1.5C Special Report using the *pyam* Python library (https://github.com/iamconsortium/pyam). Both tools were developed as open-source software projects using best practices from the domain of scientific software development.

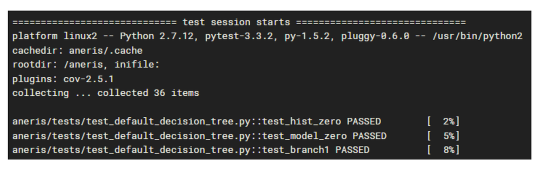
Best Practices

There are a number of best practices scientists can employ to make more reusable and robust tools. Many of these are outlined in lessons provided by the Software Carpentry foundation (ask about a workshop!) [1]. These include but are not limited to:

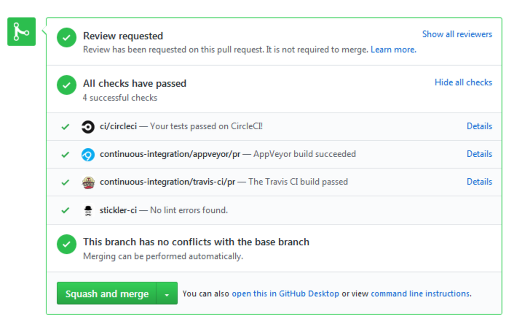
1. Use version control

GitHub

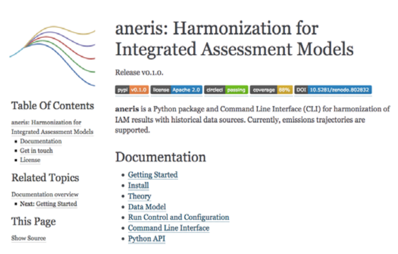
2. Test your code (and use them with every change!)



3. Leverage Continuous Integration (CI) for collaborative work



4. Write (automatic) documentation



5. Make your tool installable

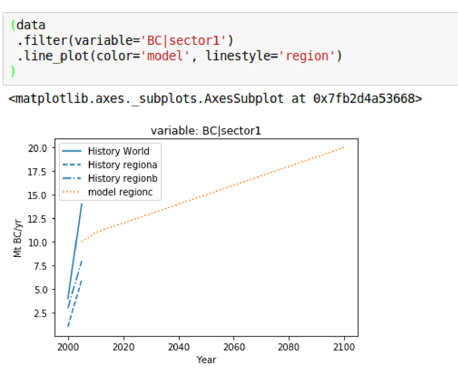
CONDA

Via Conda

```
conda install -c conda-forge pyam
```

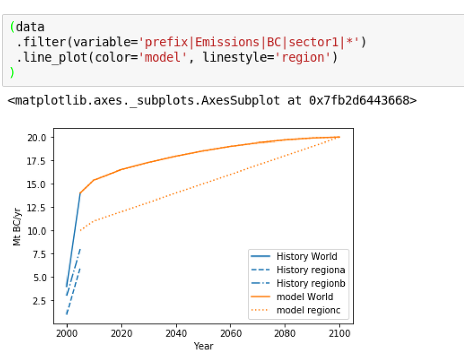
Harmonizing IAM Data for CMIP6

Aneris (http://software.ene.iiasa.ac.at/aneris/) allows for harmonization across regional ranges and variables. In this stylized example, we have historical data available at two micro-regions (regiona, regionb). Our model data includes a single a macro-region (regionc) which comprises both of the two micro-regions.



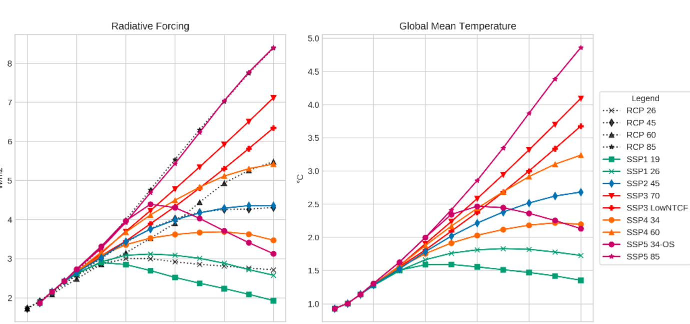
```
for scenario in driver.scenarios():
    driver.harmonize(scenario)
harmonized, metadata, diagnostics = driver.harmonized_results()
```

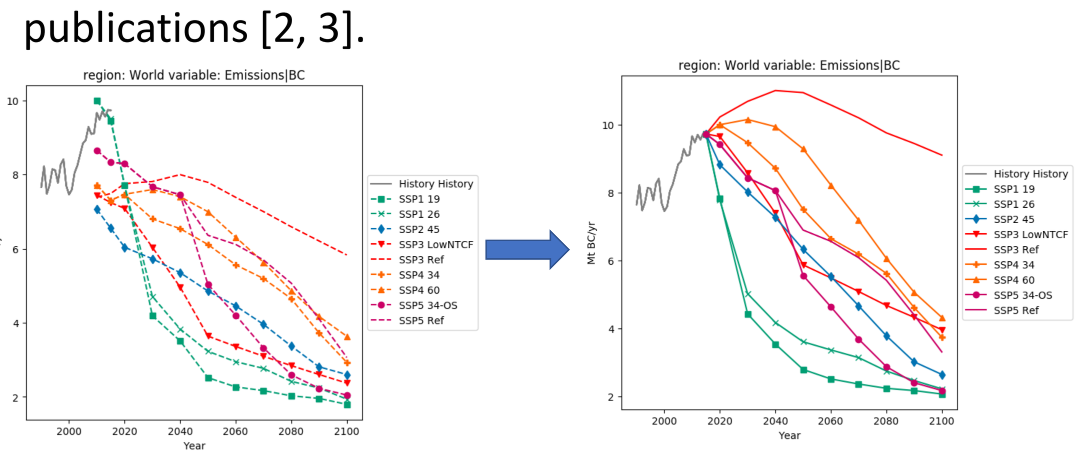
Given a historical dataset, model data, and regional mappings between the two, *aneris* will automatically harmonize the model data to historical data. It uses a decision-tree based approach to choose an appropriate harmonization method, which depends on the difference between model and historical values in the harmonization year and the resulting harmonized trajectory.

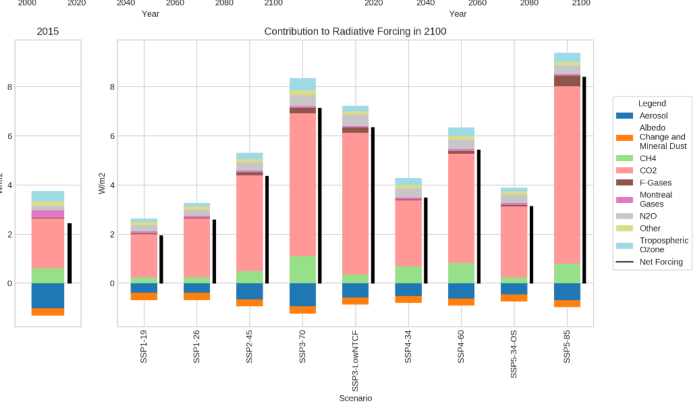


Usage in CMIP6

Aneris was used extensively by IAM teams within ScenarioMIP to quickly iterate through harmonization assessments, which resulted in two peer-reviewed publications [2, 3].







Contact: gidden@iiasa.at

Analyzing IAM Data in the IPCC SR1.5

Pyam (*http://software.ene.iiasa.ac.at/pyam/*) is a tool for easily analyzing and visualizing IAM results. Some of the key features include:

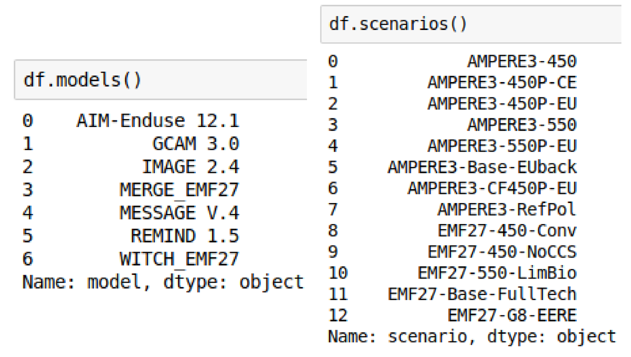
- Easily filter and manipulate data in the IAMC timeseries format
- An interface similar in feel and style to a *pandas.DataFrame*
- Advanced visualization and plotting functions
- Diagnostic checks for non-reported variables or timeseries values to validate scenario data
- Categorization of scenarios according to timeseries data or metadata for further analysis

A Quick Tour

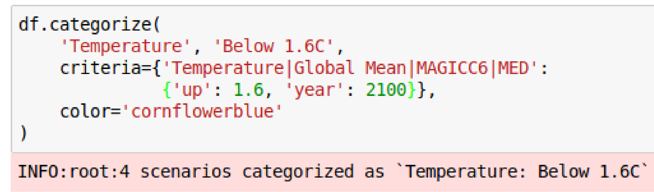
Pyam can read from a variety of file formats, including *csv*, *xls(x)*, and IIASA database snapshots.

```
df = pyam.IamDataFrame(data='tutorial_AR5_data.csv')
INFO:root:Reading `tutorial_AR5_data.csv`
```

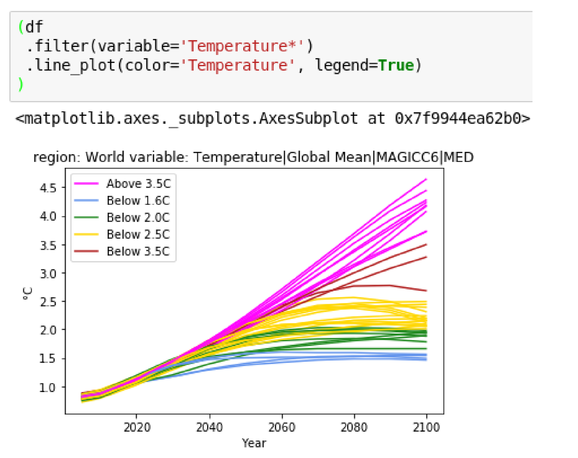
Summary information of the dataset is easy to query.



Data can be easily filtered and categorized, for example to classify different scenarios by their end of century temperature outcomes.

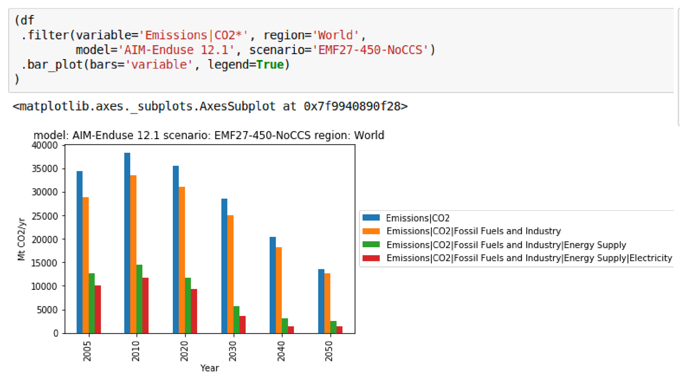


Plotting can then be applied directly to the categorized variable or any other variable of interest.

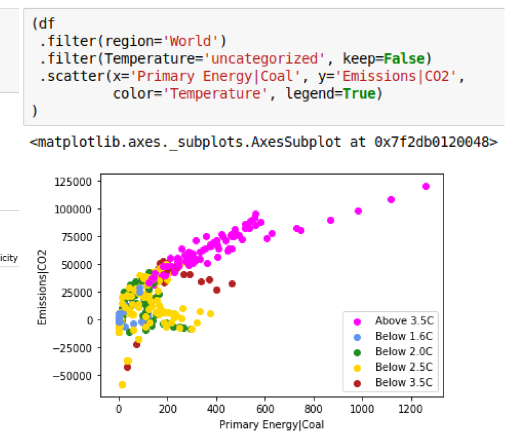


A wide variety of plotting features are supported in addition to *line_plot*(), including

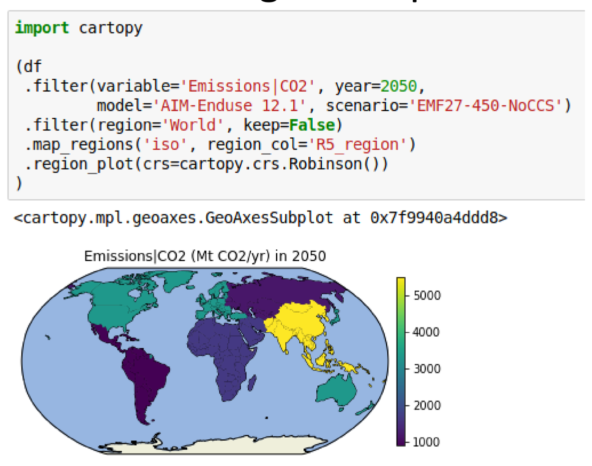
Bar Plots



Scatter Plots



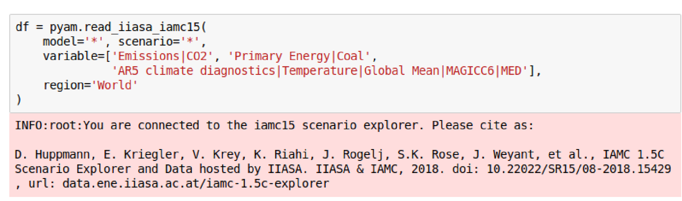
Plotting on Maps



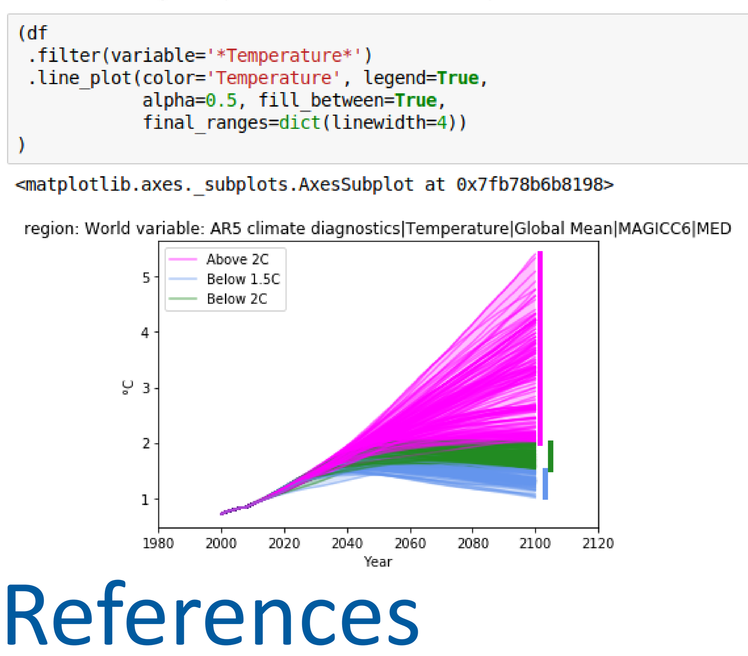
Analyzing IPCC SR1.5 Results

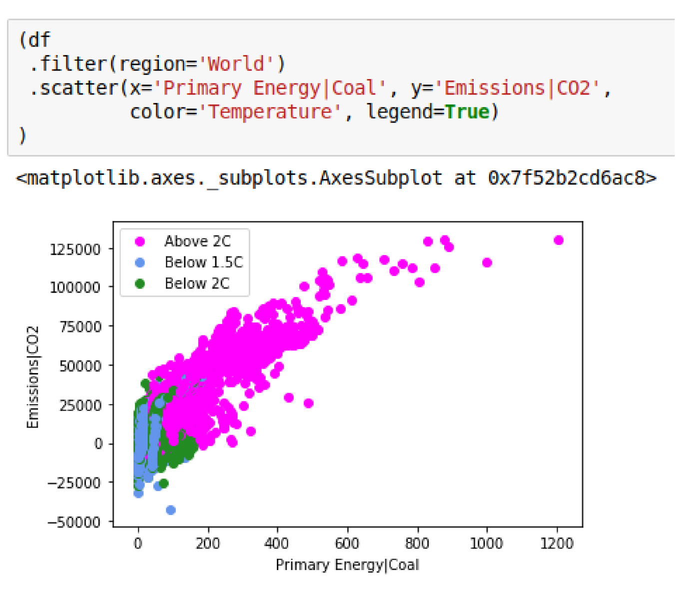
Pyam was used extensively to perform analysis and assessment of IAM scenarios by the IPCC in the SR1.5 [4].

You can directly connect to the new scenario explorer to categorize and analyze the data.



Then begin your own analysis!





References

[1] https://software-carpentry.org/lessons/
[2] Gidden, M. J. *et al.*., "A methodology and implementation of automated emissions harmonization for use in integrated assessment models," Environmental Modelling & Software, vol. 105, pp. 187 –200, 2018
[3] Gidden, M. J. *et al.*, "Global emissions pathways under different socioeconomic scenarios for use in CMIP6: A dataset of harmonized emissions trajectories through the end of the century," Geoscientific Model Development Discussions, 2018
[4] Daniel Huppmann et al., *Scenario analysis notebooks for the IPCC Special Report on Global Warming of 1.5°C*. 2018. doi: 10.22022/SR15/08-2018.15428 url: https://github.com/iiasa/ipcc_sr15_scenario_analysis