Pacific Northwest NATIONAL LABORATORY

PNNL-32567

GCAM Base Year Update

January 2022

Steven J Smith
Maridee Weber
Tanner Stelmach
Haewon McJeon

U.S. DEPARTMENT OF ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

**Available to DOE and DOE contractors from
the Office of Scientific and Technical
Information,
P.O. Box 62, Oak Ridge, TN 37831-0062**
www.osti.gov
**ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov**

**Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov**

PNNL-32567

GCAM Base Year Update

January 2022

Steven J Smith
Maridee Weber
Tanner Stelmach
Haewon McJeon

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Acknowledgement

This research was supported by the Earth and Biological Sciences Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Report

Project Summary

In FY21 we worked to update the Global Change Analysis Model (GCAM) to match historical energy consumption and technology cost trends and near-term economic growth expectations using available historical data out to 2019. We have done this in a way that builds the capacity to more easily update the model in the future to make GCAM a robust and readily updatable tool for decarbonization analysis.

Introduction and Background

The Global Change Analysis Model (GCAM) has long been used to conduct analysis of future human-Earth system pathways, including pathways that meet cumulative emissions, forcing, or temperature targets. There is an increasing need for analysis that bridges near-term and longer-term timescales. In order to do this effectively, GCAM projections need to reflect historical trends and near-term expectations.

The current version of GCAM has some limitations in terms of its near-term projections. The last historical year in GCAM (termed the last GCAM base year below) is currently 2015, which is now over 6 years in the past. Trends over the first five years of the GCAM projection can diverge significantly from actual historical trends, which means that near-term projections are starting out at an ahistorical starting point. This is a practical problem (e.g., the model is "wrong" for these years) and this also decreases the accuracy and utility of near-term scenario results. In practice, when doing applied analysis, GCAM modelers have spent a substantial amount of time adjusting the near-term behavior of GCAM so that it better matches history, which takes time away from substantive analysis.

In addition, costs for some technologies, particularly solar, wind, and batteries, have been dropping rapidly and these trends are not necessarily captured in the model. Although recent updates have improved these assumptions, there are still likely some technologies for which historical cost decreases have outpaced model assumptions.

In the FY21 portion of this project we will work to update GCAM to match historical energy consumption and technology cost trends and near-term economic growth expectations. The overall goal is to make GCAM a robust and readily updatable tool for decarbonization analysis.

Approach

This work will be performed by updating code and data within the *gcamdata* package (Bond-Lamberty et al. 2019), which generates the input data files need to run the GCAM model. *gcamdata* takes in a wide variety of data sources, processes those into common spatial regions where needed (generally countries) and then aggregates and processes the data to create the set of XML (eXtensible Markup Language) input files needed by GCAM. The current version of *gcamdata* is setup to produce a GCAM input dataset with a last historical base year of 2015. We have updated the code and necessary datasets to a model base year of 2019. We have done this in a way that allows more user flexibility in selection of the GCAM model base year, as described further below.

The majority of the work will be updating and improving *gcamdata* R code. While the code is currently written to be flexible, there are a number of places where the processing either implicitly or explicitly assumes a specific relationship between input data files and the model base year. For example, implicitly assuming that historical input data is available up to the model base year. Our approach will be to generalize the code along the following principles:

- Where it is necessary for historical input data to be present out to the last model base-year, we will add explicit checks for this so that the code will error in an informative manner if the necessary data is not present.
- Where the code is currently assuming historical data is present, but the data is being implicitly copied forward now in data files, this will be made explicit in code, which reduces the effort required to update the base-year in the future.
- Where data is incomplete, but can be readily filled in or extrapolated, the code will do this and report out the number of regions impacted. (For example, in one current test, population data was not available for 2 small countries for the test base-year. In this case the code will automatically copy existing data forward to complete the historical dataset and report out as a diagnostic message that data for two countries was extrapolated.).

In general, we will update the data system so that it is more flexible in setting the model base-year while using all available historical data. While in the past *gcamdata* was configured so that the last GCAM base-year was identical with the last historical data year for key data (such as gross domestic product (GDP), population, technology costs, and energy consumption), we now will allow the last GCAM base-year to be any year up to and including the last year historical data is available (figure 1).

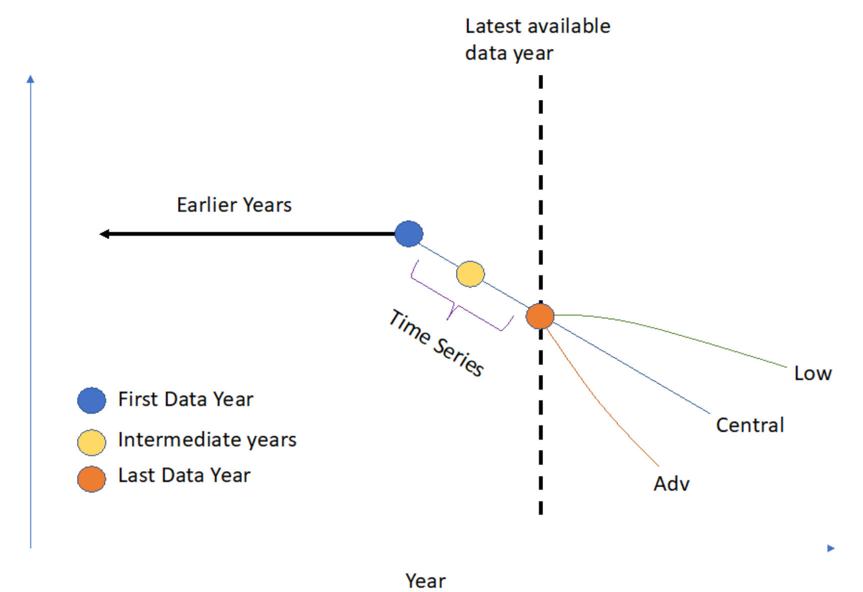

Figure 1 – New approach to blending of historical data time series with scenario projections in GCAM. Historical time series are used up until the latest year these are available (even if after the last GCAM base-year) where these then can diverge depending on the future scenarios. Where time series are available throughout the historical time period these will be used.

In some cases, such as technology costs, these data are not readily available before a specific historical year so values are kept constant before this point (as illustrated in the Figure). Note this historical assumption has no practical impact on GCAM output results since the model is populated with historical calibration values for historical years.

For the new implementation we have developed for this project we will have a number of historical data series available up to 2019, for example, as shown in Figure 1. These would all follow historical trajectories until 2019 and then diverge from there depending on the future scenario. However, the last GCAM model base year can be set in the system to be any year up to and including 2019.

For new analysis, we would use the 2019 base year, as this provides the most up to date starting point for GCAM.

However, this flexible capability also allows us to re-generate GCAM input datasets for model setups using an older base year, for example 2015 in the current GCAM 5.4 release version. This allows analysis pipelines that are set-up using current versions of GCAM to be easily updated to use more recent historical information, but without the need to re-develop all the input files that are set up for a different base year (which is very time consuming and often not practical for a project already underway).

Data Needs

The key dataset for the update is the latest version of the IEA global energy balances (https://www.iea.org/data-and-statistics/data-product/world-energy-balances). This has been purchased using separate LDRD funds under a PNNL-wide license so that this dataset can be used anywhere within PNNL. This has been processed into the input format expected by *gcamdata* and incorporated into the system.

Another key input dataset are updated pathways for future technology costs from the NREL 2021 Annual Technology Baseline for 2019-2050 (https://atb.nrel.gov). Code to incorporate this data has also been developed, although not yet fully implemented in *gcamdata*.

Updated economic growth projections from the International Monetary Fund (IMF) from the April 2021 outlook has been used to provide economic activity (e.g., Gross Domestic Product – GDP) pathways for the near-term future (IMF 2021). As in the current data system, economic growth following the SSP pathways will be used to provide multiple scenarios after the last year of the IMF near-term projections.

We did not work on updating historical agriculture and land-use data from FAO (and other sources). The system is configured to extrapolate the data available in the system, which means that this is the last priority for updating. While we ultimately will want to update these data as well, these values change less rapidly than energy-system characteristics, so are a second priority to update.

Note that GCAM already is calibrated using non-CO2 greenhouse gas and air pollutant emissions from the latest release (O'Rourke et al. 2021) from the Community Emissions Data System (CEDS; Hoesly et al. 2018). This version of the CEDS dataset extends to 2019, so no updates are required.

Results

We have updated the GCAM data system to generate input an input dataset that contains updated historical time series for energy production and consumption, economic growth rates, and electric system technology costs. *gcamdata* now produces a complete set of XML GCAM data inputs, although additional work is needed to debug the input data and correct base-year calibration mis-matches that are still present in this version.

We have set the first future model period to 2023 as an example of the flexibility of the code. At present, it is necessary to set the first future period to a (assumed) post-COVID year where the energy system is back to something closer to a nominal state. The COVID-related economic declines and sub-sequent rebound are accounted for by use of the most recent near-term IMF economic projections.

Impacts/Benefits

The new model set-up will allow analysis using the most recent historical information and near-term projections. This prepares GCAM for analysis of decarbonization and other system pathways over near- and long-term time horizons.

References

Bond-Lamberty B., K.R. Dorheim, Y. Cui, R.L. Horowitz, A.C. Snyder, K.V. Calvin, L. Feng et al. (2019). The gcamdata package: preparation synthesis and tracking of input data for the GCAM integrated human-earth systems model." *Journal of Open Research Software*, 7: 6. DOI: https://doi.org/10.5334/jors.232.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J.-I., Li, M., Liu, L., Lu, Z., Moura, M. C. P., O'Rourke, P. R., and Zhang, Q. (2018) Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), *Geosci. Model Dev.,* 11**,** 369-408. doi: 10.5194/gmd-11-369-2018.

IMF, World Economic Outlook- Managing Divergent Recoveries, 2021 (International Monetary Fund. Washington, DC, April 2021).

O'Rourke, P. R, Smith, S. J., Mott, A., Ahsan, H., McDuffie, E. E., Crippa, M., Klimont, S., McDonald, B., Z., Wang, Nicholson, M. B, Feng, L., and Hoesly, R. M. (2021, February 05). Community Emissions Data System (Version v_2021-04-21). Zenodo. http://doi.org/10.5281/zenodo. 4737769.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov