The Journal of Open Source Software

gdess: A framework for evaluating simulated atmospheric CO_2 in Earth System Models

Daniel E. Kaufman[1], **Sha Feng**[2], **Katherine V. Calvin**[1], **Bryce E. Harrop**[2], **and Susannah M. Burrows**[2]

1 Joint Global Change Research Institute, Pacific Northwest National Laboratory, College Park, MD, USA **2** Atmospheric Sciences and Global Change Division, Pacific Northwest National Laboratory, Richland, WA, USA

DOI: 10.21105/joss.03690

Software

- Review
- Repository
- Archive

Editor: Pending Editor

Submitted: 03 September 2021
Published: 12 September 2021

Summary

Atmospheric carbon dioxide (CO_2) plays a key role in the global carbon cycle and global warming. Climate-carbon feedbacks are often studied and estimated using Earth System Models (ESMs), which couple together multiple model components—including the atmosphere, ocean, terrestrial biosphere, and cryosphere—to jointly simulate mass and energy exchanges within and between these components. Despite tremendous advances, model intercomparisons and benchmarking are aspects of ESMs that warrant further improvement (Fer et al., 2021; Smith et al., 2014). Such benchmarking is critical because comparing the value of state variables in these simulations against observed values provides evidence for appropriately refining model components; moreover, researchers can learn much about Earth system dynamics in the process (Randall et al., 2019).

We introduce gdess (a.k.a., Greenhouse gas Diagnostics for Earth System Simulations), which parses observational datasets and ESM simulation output, combines them to be in a consistent structure, computes statistical metrics, and generates diagnostic visualizations. In its current incarnation, gdess facilitates evaluating a model's ability to reproduce observed temporal and spatial variations of atmospheric CO_2. The diagnostics implemented modularly in gdess support more rapid assessment and improvement of model-simulated global CO_2 sources and sinks associated with land and ocean ecosystem processes. We intend for this set of automated diagnostics to form an extensible, open source framework for future comparisons of simulated and observed concentrations of various greenhouse gases across Earth system models.

Statement of need

Thorough evaluation of simulated atmospheric CO_2 concentrations—by comparing against observations—requires multiple diagnostics, metrics, and visualizations. During the past decade, such evaluations have utilized certain common methods, such as aggregating in situ measurements into latitude bands and detrending of multidecadal time series to investigate seasonal cycles (Chevallier et al., 2019; Jing et al., 2018; Keppel-Aleks et al., 2013; Liptak et al., 2017; Ott et al., 2015; Weir et al., 2021). However, the construction of diagnostics used in these evaluations has not been automated in an open-source tool available to the broader atmospheric modeling community. Thus, each modeling or analysis team has had to decide on and code their own preferred set of diagnostics, resulting in redundancies and potential inconsistencies among efforts.

Several software packages have been developed to streamline the application of diagnostics for ESM benchmarking. These tools share related functionality with gdess, and some have directly inspired the gdess design and our development approach. For example, the ESM Evaluation Tool (ESMValTool; Eyring, Righi, et al. (2016); Eyring et al. (2020)) has been used to generate specific figures from the literature—including a comparison of column-averaged CO_2 values as performed by Gier et al. (2020)—but does not provide for tailored processing of varied CO_2 data sources. We adopted the term *recipe* from its use by ESMValTool. The International Land Model Benchmarking (ILAMB) System (Collier et al., 2018) excels at intercomparisons between multiple land models and has been used to benchmark inferred CO_2 concentrations against surface station measurements (Wu et al., 2020). In contrast to gdess, ILAMB provides the means to evaluate emulated results but not prognostic simulations for CO_2 (Keppel-Aleks, 2021).

Design and data sources

gdess is written in Python ("version 3"; Python Core Team, 2015; Van Rossum & Drake, 2009). A comprehensive readme file and docstrings throughout the open source codebase (https://github.com/E3SM-Project/gdess) provide documentation and guidance, and Continuous Integration tests facilitate further code development and maintenance. Data variables are represented and handled in memory using xarray, an open-source Python package for working with labeled multi-dimensional arrays (Hoyer & Hamman, 2017).

As shown in Figure 1, gdess is organized into modular components. A *Collection* class encapsulates source-specific attributes and methods for each data source (described below) and each Collection inherits common attributes from a parent *Multiset* class. Each diagnostic recipe, defined in a separate module file (e.g., `surface_trends.py`), instantiates and uses Collection objects to handle the loading and pre-processing of data. Additionally, visualization functions (e.g., time-series, annual cycles) are accessible from any instance of a Collection or Multiset so that data sources can be inspected individually—i.e., without the need to run one of the comparative diagnostic recipes.

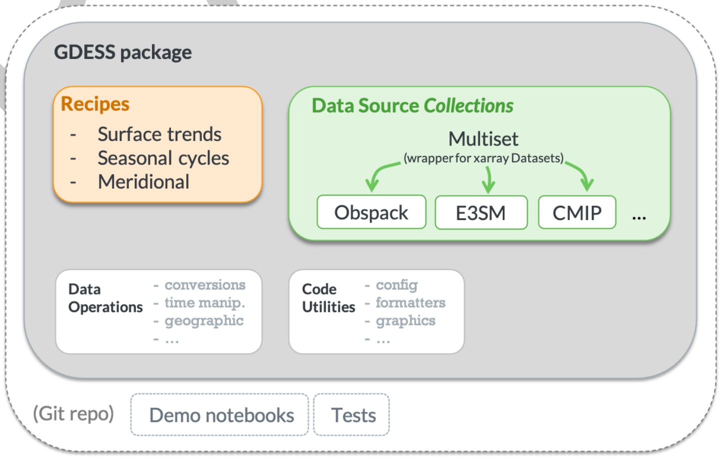

Figure 1: Schematic of the gdess code structure.

gdess can process data from three sources: Globalview+, CMIP, and E3SM. Data from surface observing stations must be retrieved from the NOAA Global Monitoring Laboratory (GML) Globalview+ version 6.0 Observation package (Obspack; Schuldt et al. (2020); Masarie et al. (2014)). In situ and flask measurements can be used from approximately 200 stations whose data in Obspack spans at least a 12 month period (Figure 2).

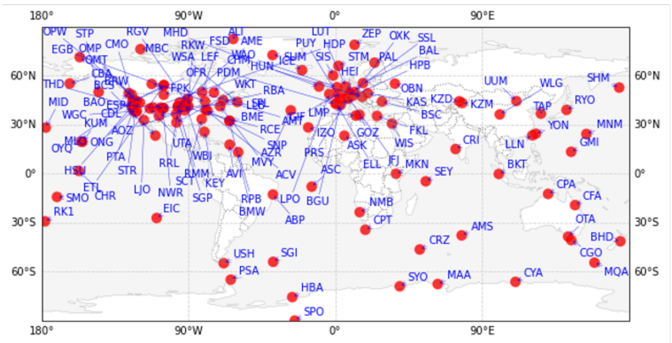

Figure 2: Global map showing surface observing station locations (red circles) and their three-letter site codes, as recorded in Obspack and used in gdess.

We distinguish between the model results from two different sources: (i) simulations by the Energy Exascale Earth System Model (E3SM), and (ii) other Earth system models participating in the latest, Version 6, Coupled Model Intercomparison Project (CMIP6). E3SM is a global modeling system composed of multiple coupled subcomponent models: atmosphere, ocean, land, ice (Burrows et al., 2020; Golaz et al., 2019). In this study, our focus is on evaluating CO_2 mole fractions in the atmospheric component, which is called the E3SM atmosphere model (EAM) and which has been described in detail by Rasch et al. (2019).

CMIP6 organizes the setup, experimental design, and intercomparisons of simulations performed using numerous global climate models. Data from CMIP6 are accessed either via locally stored files—downloaded directly from Earth System Grid Federation (ESGF) data nodes—or programmatically via the *intake-esm* package, which is a gdess dependency maintained as part of the *Pangeo* project. By default, comparisons in gdess use data from the 'esm-hist' experiment, which contains CO_2 emission-driven simulations that span the period of 1850 to 2014—i.e., an "all-forcing simulation of the recent past with atmospheric CO_2 concentration calculated" (Eyring, Bony, et al., 2016). We expect model output from any CMIP6 experiment could be used by specifying the appropriate data identifier or file location, although additional testing would be needed to confirm expected behavior.

Functionality

This section describes and provides example output from the three diagnostic recipes implemented in gdess. These recipes can be initiated either from a terminal or from within a running Python kernel. The command-line interface consists of the gdess command, followed by the type of recipe, and then options for each recipe—e.g., which observing station(s) to use for comparison. Within a Python kernel, options are specified via a dictionary object.

Multidecadal trend

Skillful simulation of the historical multidecadal trend in atmospheric CO_2 is a necessary condition for an ESM to be an effective tool for conducting climatological projections and analyses. The research questions one might address with this diagnostic recipe (see example output in Figure 3) include: What are the long-term biases in the model simulation? How does the simulated increase in CO_2 mixing ratios compare to surface measurements?

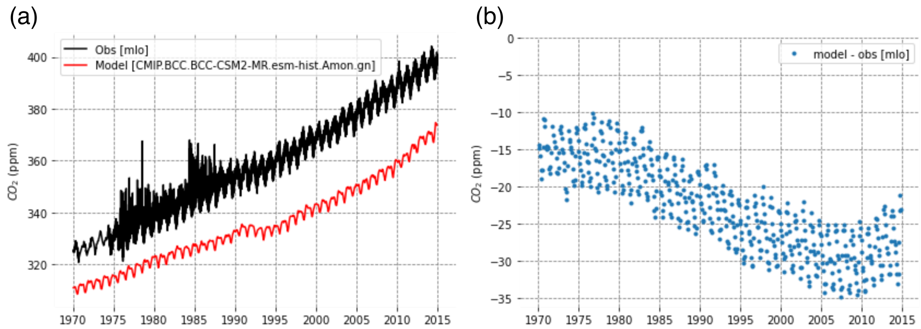

Figure 3: Example output of the `surface_trends` recipe, showing (a) individual time series and (b) differences between simulated and observed concentrations of surface-level atmospheric CO_2 at the Mauna Loa Observatory, Hawaii (MLO).

Seasonal cycle

Because of the substantial impact primary production and respiration have on CO_2 concentrations, evaluating the seasonal cycle at a given location can help disentangle the effects of biological from physical processes. The seasonal cycle can be quantified by "the projection of an atmospheric time series onto a suitably defined subset of orthogonal basis functions, the choice of which depends on the length of the series involved" (Straus, 1983). For computing the seasonal cycle, we detrend the time series by fitting a function composed of both polynomial and harmonic terms, following the procedure of Sweeney et al. (2015) and originally proposed by Thoning & Tans (1989). Example output of the `seasonal cycle` recipe is shown in Figure 4.

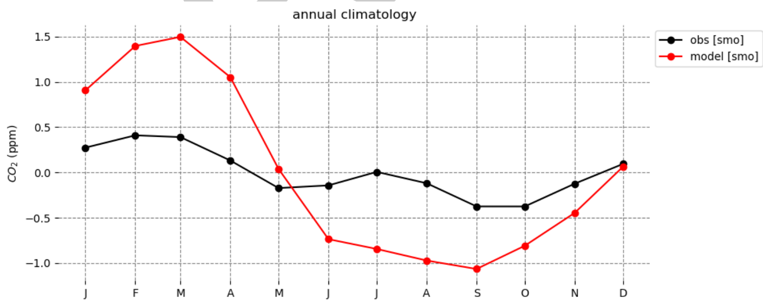

Figure 4: Example output of the `seasonal_cycle` recipe, comparing annual climatologies of surface atmospheric CO_2 concentrations at the American Samoa Observatory, Tutuila Island (SMO).

Meridional gradient

By comparing CO_2 concentrations across observing sites distributed globally, we can assess whether simulated transport and mixing is skillfully reproducing spatial gradients. For instance, the surface CO_2 flux signals at lower latitudes (30-45N) are moved to northern boreal latitudes and also to the south by large scale circulation. Spatial analysis can reveal evidence of southward movement toward (sub)tropical convection that becomes mixed with Hadley circulation or northward movement toward midlatitude synoptic weather patterns and the Ferrell circulation (Denning et al., 1999; Schuh et al., 2019; Stephens et al., 2007). Figure 5 shows example output of the `meridional` recipe.

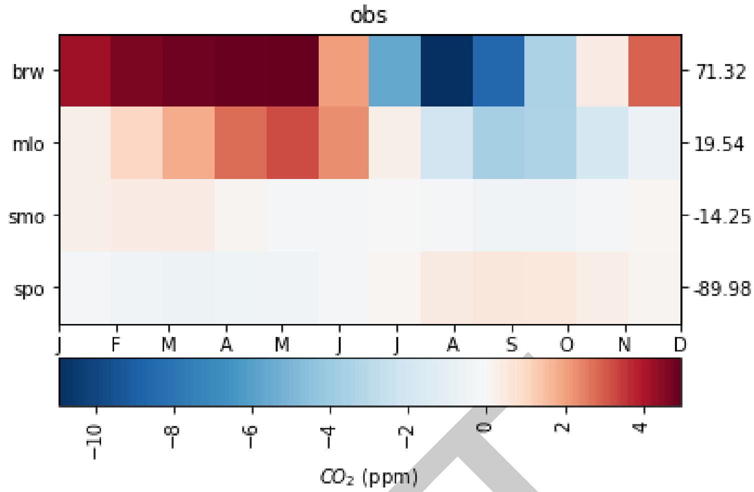

Figure 5: Example output of the `meridional` recipe, comparing the seasonal cycle across latitudes, at locations of user-specified surface stations.

Outlook

Currently, `gdess` is helping to assess simulations using the biogeochemistry configuration of E3SM, with the aim of exploring carbon-climate interactions. In addition to the three implemented recipes (multidecadal trends, seasonal cycles, and meridional gradients), current development includes two other methods—by which CO_2 was also evaluated by Keppel-Aleks et al. (2013)—vertical gradients and interannual variability. Future releases may evaluate vertical gradients using aircraft data from Globalview+ Obspack, include satellite data, and extend to data for other greenhouse gases, such as methane.

Acknowledgements

We thank Drs. Colm Sweeney and Kirk Thoning, at the NOAA Global Monitoring Laboratory, for providing code and support for implementing the curve fitting methods. A dataset file provided via the Obspack from the Mauna Loa surface observing station is included in the tests directory with permission from the data provider, Keeling et al. (2001). This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy (DOE), Office of Science, Office of Biological and Environmental Research. Data analysis described in this work relied on computational resources provided by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract DE-AC02-05CH11231. The Pacific Northwest National Laboratory (PNNL) is operated for DOE by Battelle Memorial Institute under Contract DE-AC05-76RLO1830.

References

Burrows, S. M., Maltrud, M., Yang, X., Zhu, Q., Jeffery, N., Shi, X., Ricciuto, D., Wang, S., Bisht, G., Tang, J., Wolfe, J., Harrop, B. E., Singh, B., Brent, L., Baldwin, S., Zhou, T., Cameron-Smith, P., Keen, N., Collier, N., … Leung, L. R. (2020). The DOE E3SM v1.1 Biogeochemistry Configuration: Description and Simulated Ecosystem-Climate

Responses to Historical Changes in Forcing. *J. Adv. Model. Earth Syst.*, *12*(9), 1–59. https://doi.org/10.1029/2019MS001766

Chevallier, F., Remaud, M., O'Dell, C. W., Baker, D., Peylin, P., & Cozic, A. (2019). Objective evaluation of surface- and satellite-driven carbon dioxide atmospheric inversions. *Atmos. Chem. Phys.*, *19*(22), 14233–14251. https://doi.org/10.5194/acp-19-14233-2019

Collier, N., Hoffman, F. M., Lawrence, D. M., Keppel-Aleks, G., Koven, C. D., Riley, W. J., Mu, M., & Randerson, J. T. (2018). The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation. *J. Adv. Model. Earth Syst.*, *10*(11), 2731–2754. https://doi.org/10.1029/2018MS001354

Denning, A. S., Holzer, M., Gurney, K. R., Heimann, M., Law, R. M., Rayner, P. J., Fung, I. Y., Fan, S.-M., Taguchi, S., Friedlingstein, P., Balkanski, Y., Taylor, J., Maiss, M., & Levin, I. (1999). Three-dimensional transport and concentration of SF6 A model intercomparison study (TransCom 2). *Tellus B: Chemical and Physical Meteorology*, *51*(2), 266–297. https://doi.org/10.3402/tellusb.v51i2.16286

Eyring, V., Bock, L., Lauer, A., Righi, M., Schlund, M., Andela, B., Arnone, E., Bellprat, O., Carvalhais, N., Cionni, I., Cortesi, N., Crezee, B., L. Davin, E., Davini, P., Debeire, K., De Mora, L., Deser, C., Docquier, D., Earnshaw, P., … Zimmermann, K. (2020). Earth System Model Evaluation Tool (ESMValTool) v2.0 - An extended set of large-scale diagnostics for quasi-operational and comprehensive evaluation of Earth system models in CMIP. *Geosci. Model Dev.*, *13*(7), 3383–3438. https://doi.org/10.5194/gmd-13-3383-2020

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K. D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., … Williams, K. D. (2016). ESMValTool (v1.0)-a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. *Geosci. Model Dev.*, *9*(5), 1747–1802. https://doi.org/10.5194/gmd-9-1747-2016

Fer, I., Gardella, A. K., Shiklomanov, A. N., Campbell, E. E., Cowdery, E. M., De Kauwe, M. G., Desai, A., Duveneck, M. J., Fisher, J. B., Haynes, K. D., Hoffman, F. M., Johnston, M. R., Kooper, R., LeBauer, D. S., Mantooth, J., Parton, W. J., Poulter, B., Quaife, T., Raiho, A., … Dietze, M. C. (2021). Beyond ecosystem modeling: A roadmap to community cyberinfrastructure for ecological data-model integration. *Global Change Biology*, *27*(1), 13–26. https://doi.org/10.1111/gcb.15409

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., … Zhu, Q. (2019). The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2089–2129. https://doi.org/10.1029/2018MS001603

Hoyer, S., & Hamman, J. (2017). xarray: N-D labeled arrays and datasets in Python. *Journal of Open Research Software*, *5*(1). https://doi.org/10.5334/jors.148

Jing, Y., Wang, T., Zhang, P., Chen, L., Xu, N., & Ma, Y. (2018). Global atmospheric CO2 concentrations simulated by GEOS-Chem: Comparison with GOSAT, carbon tracker and ground-based measurements. *Atmosphere (Basel).*, *9*(5). https://doi.org/10.3390/atmos9050175

Keeling, C. D., Piper, S. C., Bacastow, R. B., Wahlen, M., Whorf, T. P., Heimann, M., & Meijer, H. A. (2001). *Exchanges of atmospheric CO2 and 13CO2 with the terrestrial* *biosphere and oceans from 1978 to 2000. I. Global aspects* (S. D. Scripps Institution of Oceanography, Ed.; No. 01-06; p. 88). UC San Diego: Library – Scripps Digital. https://escholarship.org/uc/item/09v319r9

Keppel-Aleks, G. (2021). personal communication.

Keppel-Aleks, G., Randerson, J. T., Lindsay, K., Stephens, B. B., Keith Moore, J., Doney, S. C., Thornton, P. E., Mahowald, N. M., Hoffman, F. M., Sweeney, C., Tans, P. P., Wennberg, P. O., & Wofsy, S. C. (2013). Atmospheric carbon dioxide variability in the community earth system model: Evaluation and transient dynamics during the twentieth and twenty-first centuries. *J. Clim.*, *26*(13), 4447–4475. https://doi.org/10.1175/JCLI-D-12-00589.1

Liptak, J., Keppel-Aleks, G., & Lindsay, K. (2017). Drivers of multi-century trends in the atmospheric CO2 mean annual cycle in a prognostic ESM. *Biogeosciences*, *14*(6), 1383–1401. https://doi.org/10.5194/bg-14-1383-2017

Masarie, K. A., Peters, W., Jacobson, A. R., & Tans, P. P. (2014). ObsPack: A framework for the preparation, delivery, and attribution of atmospheric greenhouse gas measurements. *Earth Syst. Sci. Data*, *6*(2), 375–384. https://doi.org/10.5194/essd-6-375-2014

Ott, L. E., Pawson, S., Collatz, G. J., Gregg, W. W., Menemenlis, D., Brix, H., Rousseaux, C. S., Bowman, K. W., Liu, J., Eldering, A., Gunson, M. R., & Kawa, S. R. (2015). Assessing the magnitude of CO 2 flux uncertainty in atmospheric CO 2 records using products from NASA's Carbon Monitoring Flux Pilot Project. *J. Geophys. Res. Atmos.*, *120*(2), 734–765. https://doi.org/10.1002/2014JD022411

Python Core Team. (2015). *Python: A dynamic, open source programming language. Python Software Foundation*. https://www.python.org/

Randall, D. A., Bitz, C. M., Danabasoglu, G., Denning, A. S., Gent, P. R., Gettelman, A., Griffies, S. M., Lynch, P., Morrison, H., Pincus, R., & Thuburn, J. (2019). 100 Years of Earth System Model Development. *Meteorol. Monogr.*, *59*, 12.1–12.66. https://doi.org/10.1175/amsmonographs-d-18-0018.1

Rasch, P. J., Xie, S., Ma, P. L., Lin, W., Wang, H., Tang, Q., Burrows, S. M., Caldwell, P., Zhang, K., Easter, R. C., Cameron-Smith, P., Singh, B., Wan, H., Golaz, J. C., Harrop, B. E., Roesler, E., Bacmeister, J., Larson, V. E., Evans, K. J., … Yang, Y. (2019). An Overview of the Atmospheric Component of the Energy Exascale Earth System Model. *J. Adv. Model. Earth Syst.*, *11*(8), 2377–2411. https://doi.org/10.1029/2019MS001629

Schuh, A. E., Jacobson, A. R., Basu, S., Weir, B., Baker, D., Bowman, K., Chevallier, F., Crowell, S., Davis, K. J., Deng, F., Denning, S., Feng, L., Jones, D., Liu, J., & Palmer, P. I. (2019). Quantifying the Impact of Atmospheric Transport Uncertainty on CO2 Surface Flux Estimates. *Global Biogeochemical Cycles*, *33*(4), 484–500. https://doi.org/10.1029/2018GB006086

Schuldt, K. N., Mund, J., Luijkx, I. T., Jacobson, A. R., Cox, A., Vermeulen, A., Manning, A., Beyersdorf, A., Manning, A., Karion, A., Hensen, A., Arlyn Andrews, Frumau, A., Colomb, A., Scheeren, B., Law, B., Baier, B., Munger, B., Paplawsky, B., … Loh, Z. (2020). *Multi-laboratory compilation of atmospheric carbon dioxide data for the period 1957-2019; obspack_co2_1_GLOBALVIEWplus_v6.0_2020-09-11*. NOAA Earth System Research Laboratory, Global Monitoring Division. https://doi.org/10.25925/20200903

Smith, M. J., Palmer, P. I., Purves, D. W., Vanderwel, M. C., Lyutsarev, V., Calderhead, B., Joppa, L. N., Bishop, C. M., & Emmott, S. (2014). Changing How Earth System Modeling is Done to Provide More Useful Information for Decision Making, Science, and Society. *Bulletin of the American Meteorological Society*, *95*(9), 1453–1464. https://doi.org/10.1175/BAMS-D-13-00080.1

Stephens, B. B., Gurney, K. R., Tans, P. P., Sweeney, C., Peters, W., Bruhwiler, L., Ciais, P., Ramonet, M., Bousquet, P., Nakazawa, T., Aoki, S., Machida, T., Inoue, G., Vinnichenko, N., Lloyd, J., Jordan, A., Heimann, M., Shibistova, O., Langenfelds, R. L., … Denning, A. S. (2007). Weak Northern and Strong Tropical Land Carbon Uptake from Vertical Profiles of Atmospheric CO2. *Science*, *316*(5832), 1732–1735. https://doi.org/10.1126/science.1137004

Straus, D. M. (1983). On the Role of the Seasonal Cycle. *Journal of Atmospheric Sciences*, *40*(2), 303–313. https://doi.org/10.1175/1520-0469(1983)040%3C0303:OTROTS%3E2.0.CO;2

Sweeney, C., Karion, A., Wolter, S., Newberger, T., Guenther, D., Higgs, J. A., Andrews, A. E., Lang, P. M., Neff, D., Dlugokencky, E., Miller, J. B., Montzka, S. A., Miller, B. R., Masarie, K. A., Biraud, S. C., Novelli, P. C., Crotwell, M., Crotwell, A. M., Thoning, K., & Tans, P. P. (2015). Seasonal climatology of CO 2 across North America from aircraft measurements in the NOAA/ESRL Global Greenhouse Gas Reference Network. *J. Geophys. Res. Atmos.*, *120*(10), 5155–5190. https://doi.org/10.1002/2014JD022591

Thoning, K. W., & Tans, P. P. (1989). Atmospheric carbon dioxide at Mauna Loa Observatory. 2. Analysis of the NOAA GMCC data, 1974-1985. *J. Geophys. Res.*, *94*(D6), 8549–8565. https://doi.org/10.1029/JD094iD06p08549

Van Rossum, G., & Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace. ISBN: 1441412697

Weir, B., Ott, L. E., Collatz, G. J., Kawa, S. R., Poulter, B., Chatterjee, A., Oda, T., & Pawson, S. (2021). Bias-correcting carbon fluxes derived from land-surface satellite data for retrospective and near-real-time assimilation systems. *Atmospheric Chemistry and Physics*, *21*(12), 9609–9628. https://doi.org/10.5194/acp-21-9609-2021

Wu, G., Cai, X., Keenan, T. F., Li, S., Luo, X., Fisher, J. B., Cao, R., Li, F., Purdy, A. J., Zhao, W., Sun, X., & Hu, Z. (2020). Evaluating three evapotranspiration estimates from model of different complexity over China using the ILAMB benchmarking system. *Journal of Hydrology*, *590*, 125553. https://doi.org/10.1016/j.jhydrol.2020.125553