

**Data Set Name:**

Greenhouse Gas Observing Network

**Abstract:**

Design an observing network to monitor emissions of a greenhouse gas (GHG) in California given time series of synthetic observations and tracers from weather model simulations.

**Source:**

This data was created under work funded by the National Institute of Standards and Technology and Laboratory Directed Research and Development projects at the Lawrence Livermore National Laboratory. The work was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344, and is released under UCRL number LLNL-MI-668913. Refer to our published paper or contact dducas at alum.mit.edu for dataset questions.

**Data Set Information:**

This data set contains time series of greenhouse gas (GHG) concentrations at 2921 grid cells in California created using simulations of the Weather Research and Forecast model with Chemistry (WRF-Chem). Each grid cell covers an area of 12 km by 12 km, and there is one data file per grid cell. Each file contains 16 time series of GHG concentrations. The data points in the time series are spaced 6 hours apart (4 samples per day) over the period May 10 – July 31, 2010. The first 15 rows are time series of GHG tracers released from 14 distinct spatial regions in California and one outside of California. The last row corresponds to the time series of “synthetic GHG observations” generated with EDGAR emissions of HFC-134a scaled by a factor 0.7 and with noise added.

Using this data, the goals are to (1) use inverse methods to determine the optimal values of the weights in the weighted sum of 15 tracers that best matches the synthetic observations, (2) and use optimization methods to determine the best locations to observe GHGs to constrain the inversion. We used a Bayesian method for (1) and genetic algorithms for (2).

Further details about the data and methods are given in the publication "Designing optimal greenhouse gas observing networks that consider performance and cost," Geoscientific Instrumentation Methods and Data Systems.

**Attribute Information:**

Each file in the data set is labeled ghg.gid.siteWXYZ.dat, where WXYZ is an integer location ID described in our manuscript.

At each location,

Rows 1-15: GHG concentrations of tracers emitted from regions 1-15

Row 16: GHG concentrations of synthetic observations

Columns 1-327: GHG concentrations every 6 hours from May 10 – July 31, 2010.

All GHG concentrations are in units of parts per trillion.

**Relevant Papers:**

Lucas, D. D., Yver Kwok, C., Cameron-Smith, P., Graven, H., Bergmann, D., Guilderson, T. P., Weiss, R., and Keeling, R.: "Designing optimal greenhouse gas observing networks that consider performance and cost," Geoscientific Instrumentation Methods and Data Systems, (2015).

[<http://www.geosci-instrum-method-data-syst-discuss.net/4/705/2014/gid-4-705-2014.html>]

**Citation Requests:**

Please cite our *final* paper in Geoscientific Instrumentation Methods and Data Systems.

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.