This README describes the different datasets used for the paper "**The 2019 Methane Budget And Uncertainties At 1 Degree Resolution And Each Country Through Bayesian Integration Of GOSAT Total Column Methane Data And A Priori Inventory Estimates**"

1) Prior and Posterior Emissions and their uncertainties by Sector at 1 degree resolution

The file TopDownEmissions_GOSAT_GEOS_CHEM_2019.nc contains the prior and posterior emissions and their uncertainties (squared) by sector at 1 degree resolution. Note that the uncertainties are only appropriate at the grid cell (i.e. 1 degree) resolution. As the uncertainties are correlated between grid cells, projecting the emissions to a coarser scale results in a too small uncertainty. To properly project to this coarser resolution you should use the prior or posterior covariances provided in the other links on this web site.

An alternative, ad-hoc approach for calculating uncertainties at coarser scales could be to 1) linearly sum the uncertainties for the selected region (this assumes the uncertainties are 100% correlated for the selected region, then 2) linearly sum the square of the uncertainties (the value provided by the corresponding matrix in this file) and take the square root. 3) Take the average between these values.

For the formal approach the user will need to use the prior and posterior covariances provided by the corresponding files available on this web site.

2) Prior covariances

Prior covariances are provided in a zip file by sector (e.g. CoalPrior.zip). Within each zip file are netcdf files related to 8 regions. The label might be for an entire continent (e.g. N. America) or for the country that contains the largest emissions within a region (e.g. India, Russia). The emission sector is given as the first part of the name (e.g. Coal). However two names were shortened or renamed to reflect other inventories: Live = Livestock and Farm = Rice

Each a priori covariance contains five variables. While each file is for a specific region

(e.g. N. America), the longitude and latitude arrays are for the whole globe.

Here are the lon/lt ranges for each region the ranges are lon1,lon2,lat1,lat2
N. America ,-175,-40,25,80
S. America: -130,-30,-65,25
Europe/N. Africa: -24,60,20,80
Africa: -22,60,-40,20
Russia: 60,179,50,90
India: 60,90,5,50
Asia 90,179,5,50
Australia/Indonesia 90,179,-45,5,

The variables contained in each file are:
1) ch4_emissions contains a vector of emissions with units of Tg CH4/ year

2) The lonlatindex is the index corresponding to ch4_emissions and it is where ch4_emissions are larger than zero

A value of 0 in the lonlatindex refers to the zeroth element in the longitude and latitude array or -179.5, -89.5. A value of 1 would be -178.5,-89.5, A value of 360 corresponds to -179.5,-88.5

3 and 4) The a priori covariances contain latitude and longitude matrices and are the same for all files and contain the lat/lon values for the whole globe and are generated for user convenience.

5) SA is the prior covariance corresponding to ch4_emissions. Units are Tg CH4/yr (squared) or Tg CH4/yr * Tg CH4/yr

3) Posterior Covariances

Posterior covariances have a slightly different naming e.g. convection global_gosat_2010_2018_regionemissiontype.nc. The variables in the posterior covariance have slightly different names but correspond to the variables in the prior covariances.

1) posterior_emission contains a vector of emissions with units of Tg CH4/ year (can be compared to ch4_emissions in the prior covariance)

2) The lonlatindex is the index corresponding to posterior_emission
A value of 0 in the lonlatindex refers to the zeroth element in the longitude and latitude array or -179.5, -89.5. A value of 1 would be -178.5,-89.5, A value of 360 corresponds to -179.5,-88.5

3) posterior_covariance is posterior prior covariance corresponding to posterior_emission. Units are Tg CH4/yr (squared) or Tg CH4/yr * Tg CH4/yr

4) CH4_emissions.nc: prior fluxes (not emissions!) for GOSAT 2019 Inversion (Qu et al. 2021)

5) Post_SF_emissions.nc: posterior scaling factors that when applied to prior fluxes yield posterior fluxes

6) cluster_map.pkl: mapping from flux state vector element index to lat/lon coordinate

7) S_poscov.pkl: posterior scaling factor covariance matrix that when applied to prior error covariance matrix yields posterior flux covariance matrix