Data Intake Processing and Verification Report

Background Information

- **Original Dataset Name:** ODIAC Fossil Fuel Emission Dataset
- **GHG Center Dataset Title:** ODIAC Fossil Fuel CO_2 Emissions
- **Dataset Provider:** NASA/NIES
- **Date Obtained:** June 5, 2024
- **Location Obtained from:** http://doi.org/10.17595/20170411.001
- **Data Location in GHG Center:** odiac-ffco2-monthgrid-v2023
- **Data POC(s):** Dr. Tomohiro Oda
- **Dataset File Type(s):** GeoTIFF
- **Projection (if different from WGS84):** NA

Data Transfer Confirmation

An MD5 checksum is used to detect high-level errors within data transmissions
- Results for individual checksum file comparisons of pre-transfer and post-transfer for few files are shown below:

Filename	MD5 Original file	MD5 Downloaded file
odiac2023_1km_excl_intl_2201.tif.gz	6aaa315ee69d20b1bf4634d8c575cca3	6aaa315ee69d20b1bf4634d8c575cca3
odiac2023_1km_excl_intl_1004.tif.gz	f2b33f753c56c871a6d22c166c953a28	f2b33f753c56c871a6d22c166c953a28

- All files were transferred successfully
- Report any individual file issues: NA

Data Intake Process

- https://us-ghg-center.github.io/ghgc-docs/data_workflow/odiac-ffco2-monthgrid-v2023_Data_Flow.html

Overall Dataset Statistics:

- Data file reads confirmed: 23 years * 12 months, total of 276 files
- Mean, min, max across all files:

	Minimum	Maximum	Mean	Standard Deviation
Original Dataset	0.0	1635993.875	0.895	267.301

Cog transformed dataset	0.0	1635993.875	0.895	267.301

- File range (most cases will be all files)
- Bounding Box of all data
- Link to transformation record in Jupyter Notebook:
  https://github.com/US-GHG-Center/ghgc-docs/blob/main/cog_transformation_scripts/odiac-ffco2-monthgrid-v2023_Transformation_Code.py
- All values are in expected range (catches out of range values)

Specific, Random Checks / Visual Confirmation

- More detailed statistics for few files are shown below (randomly chosen)

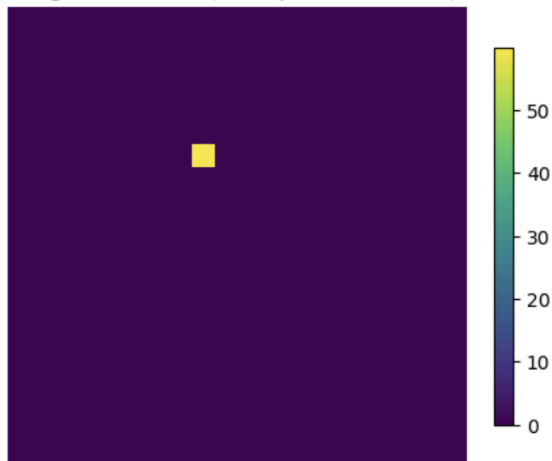   ○ Statistics were performed for the following:
        ■ February, 2000:

	Minimum	Maximum	Mean	Standard Deviation
Original Dataset	0.0	864731.375	0.585	193.043
Cog transformed dataset	0.0	864731.375	0.585	193.043

        ■ November, 2022:

	Minimum	Maximum	Mean	Standard Deviation
Original Dataset	0.0	1267992.625	0.843	248.995
Cog transformed dataset	0.0	1267992.625	0.843	248.995

- The data comparison for May, 2010 was performed at a specific location by subsetting using the indices [2430:2450, 1200:1220]

Original Dataset (at a specific location)	COG file (at a specific location)
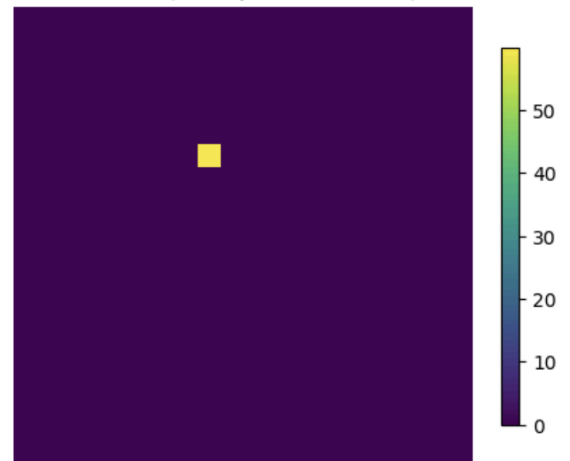
	Minimum	Maximum	Mean	Standard Deviation
Original Dataset	0.0	51.346	0.1283	2.564
Cog transformed dataset	0.0	51.346	0.1283	2.564

Summary

- We are confident that the transformation and display of data in GHG Center is correct
- These are no problems we have identified in the data.
- Link to Data Usage Notebook
- Link to US GHG Center Data Catalog overview page

Report Completed on:
MSFC POC for questions: Deborah Smith, Siddharth Chaudhary