

# Dataset Validation

## Background Information

- Original Dataset Name: ODIAC Fossil Fuel Emission Dataset
- GHGC Dataset Title: ODIAC Fossil Fuel CO<sub>2</sub> Emissions
- Dataset Provider: NASA/NIES
- Date Obtained: 06/01/2023
- Location Obtained from: [https://db.cger.nies.go.jp/dataset/ODIAC/DL\\_odiac2022.html](https://db.cger.nies.go.jp/dataset/ODIAC/DL_odiac2022.html)
- Data location in GHGC: odiaac-fossil-fuel-co2-emissions-2022
- Data POC(s): Dr. Tomohiro Oda, Dr. Lesley Ott
- Dataset File Type(s): GeoTIFF
- Projection (if different from WGS84): NA

## Checksum results (if used in data transfer):

An SHA-256 checksum is used to detect high-level errors within data transmissions

- Results from individual checksum file comparisons of pre-transfer and post-transfer
  - All files were transferred successfully

Filename	SHA 256 Original file	

- Report any individual file issues: NA

## Overall Statistics:

- Data file reads confirmed: 22 years \* 12 months, total of 264 files
- Mean, min, max across all files:
  - Original dataset:
  - COG tranformed dataset:
- Distribution of values across all data (by variable)
  - Original dataset
  - Transformed

- What file range most cases will be all files
- Bounding Box of all data
- Link to transformation record in Jupyter Notebook
  - Script to transform GeoTIFF to COG
- All values are in expected range (catches out of range values)

## Specific, Random Checks / Visual Confirmation

- Visual example and side by side comparison
- More detailed statistics for specific files (randomly chose)
  - Statistics was performed for the following files:
    - 2010-03
    - 2018-08
    - 2006-12
    - 2003-01
- Data comparison at a few specific locations

## Summary

- We are confident that the transformation and display of ODIAC 2022 data in GHGC is correct.