Data Intake Processing and Verification Report

Background Information

- **Original Dataset Name:** Replication Data for: Geostationary satellite observations of extreme and transient methane emissions from oil and gas infrastructure
- **GHG Center Dataset Title:** Geostationary Satellite Observations of Extreme and Transient Methane Emissions from Oil and Gas Infrastructure
- **Dataset Provider:** NOAA, Harvard University
- **Date Obtained:** November 2024
- **Location Obtained From:** https://doi.org/10.7910/DVN/EQWHCG
    - Plume masks (for visualization purposes) delivered directly to the US GHG Center
- **Data Location in GHG Center:** goes-ch4plume-v1
- **Data POC(s):** Dr. Shobha Kondragunta, Dr. Daniel Varon, Dr. Tailong He
- **Dataset File Type(s):** NetCDF
- **Projection (if different from WGS84):** GOES defined projection

Data Transfer Confirmation

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 shows all files were transferred successfully and no individual files had any transfer issues.

Data Intake Process

- https://us-ghg-center.github.io/ghgc-docs/data_workflow/goes-ch4plume-v1_Data_Flow.html

Dataset Statistics

- Statistics across selected plume locations:

Plume Locations	Minimum (mol CH_4/m^2)	Maximum (mol CH_4/m^2)	Mean (mol CH_4/m^2)	Standard Deviation
BV1-1	-0.4	0.43	0.02	0.04
BV1-2	-0.57	0.5	0.03	0.05

BV2-1	-0.24	0.41	0.03	0.04
BV2-2	-0.13	0.32	0.01	0.03
IN-2	-0.13	0.42	0.04	0.07
PB-1	-0.04	0.16	0.04	0.05

Note:

All the provided data were in NetCDF format, using a projection defined by GOES. Each file represented a gridded product with dimensions of X × Y pixels. Files were generated at 5-minute intervals, and each file included a corresponding plume mask in binary format. This plume mask was used to clip the data, isolating the relevant pixels, which were fewer than X × Y.

The clipped data were then reprojected to the WGS 84 coordinate system. Statistics and visualizations for each reprojected file were shared with the data provider and subsequently approved. It is important to note that due to the reprojection process, the data values were interpolated during regridding, and as a result, they may not exactly match the original values.

- Link to transformation record in Jupyter Notebook
- All values are in expected range

Summary

- We are confident that the transformation and display of data in the GHG Center is correct
- There were no problems 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: Jeanné le Roux, Siddharth Chaudhary