Response to reviewer comments on:

Omara, M., Gautam, R., et al. Developing a spatially explicit global oil and gas infrastructure database for characterizing methane emission sources at high resolution (https://essd.copernicus.org/preprints/essd-2022-452/)

Reviewer 2

The oil and gas infrastructure mapping is very important to monitoring and modeling the GHG emissions for limiting climate change. This work provides the most complete dataset so far that collects infrastructure information from online sources. The infrastructure type and geolocation from this work can be very useful to the GHG emission inventory developments and modeling from facility to regional and global scale with remote sensing images. Also, it can also be used as the ground-truth dataset for the oil and gas infrastructure identification with remote sensing images and machine learning approach. The methods are clearly described, and the paper also provides a detailed bottom-up emission inventory case study using this dataset.

We thank Reviewer 2 for these helpful feedback on our manuscript.

I'm very excited to download and look at the OGIM_v1 dataset (OGIM_v1.gpkg), however, I found the currently provided dataset should be further modified before it can be used in other research:

1. It would be better to provide the mapping from column names as well as the shortcuts used in the entries (especially for the "FAC_TYPE") to their detailed meanings, maybe provide a table in the supplementary or in the description of Zenodo?

We have included an appendix to the Main Text with full description of the OGIM data attributes.

2. Figure 6. I see oil and natural gas infrastructures in some countries, such as China and India, but I cannot find them in the OGIM_v1 dataset. There is no "China" or "Indian" in the "COUNTRY".

There is indeed oil and gas infrastructure datasets for China and India in the OGIM database. We note that these are countries with limited or no public data for oil and natural gas wells (Figure 6 (a)) and, as such, searching the OGIM database for wells information may return null values. However, also as shown in Figure 6, the OGIM database includes several other publicly available oil and gas infrastructure datasets for these countries, including LNG facilities, pipelines, refineries, and offshore platforms.

3. It will be interesting to see how the emissions of the natural gas compressor as well as the natural gas flaring changed from remote sensing images after the Ukraine war if the infrastructure information is provided by the dataset. So, if the Russian data

are not included in the current version, it should be stated in the paper unless it will be available soon.

We have included the following description in the Data Availability section of the manuscript:

- *"The current version of the publicly available OGIM database does not include compressor station locations for Russia (shown in the map on Figure 6). Future updates to the OGIM database may include these datasets when appropriate permissions to make them publicly accessible are obtained."*

4. I tried to extract the pipeline from the OGIM_v1 dataset, however, there is no such type either from "FAC_TYPE" or "geometry". The same issue also exists for the fields or basins.

The OGIM_v1 dataset does indeed include pipeline datasets and oil and gas fields and basins. The following are the layer names in the OGIM dataset:

```
['Oil_and_Natural_Gas_Wells',
 'Natural_Gas_Compressor_Stations',
 'Gathering_and_Processing',
 'Tank_Battery',
 'Offshore_Platforms',
 'LNG_Facilities',
 'Crude_Oil_Refineries',
 'Petroleum_Terminals',
 'Injection_Disposal_and_Underground_Storage',
 'Stations_Other',
 'Natural_Gas_Flaring_Detections',
 'Equipment_and_Components',
 'Oil_and_Natural_Gas_Production',
 'Oil_Natural_Gas_Pipelines',
 'Oil_and_Natural_Gas_Fields',
 'Oil_Natural_Gas_Basins',
 'OGIM_v1_Data_Catalog'
]
```

We note that the OGIM_v1 dataset was developed and tested using open-source software, including Python 3.7 and QGIS. We have not tested and do not guarantee that the dataset will be accessible in other proprietary software or GIS programs. We have added the following sentence in the "Data Availability" section:

- *"OGIM_v1 was developed and tested using open-access software (Python 3.7 and QGIS)."*

5. Line 360-365: "We quantitatively assess data quality in each country for which open oil and gas data for these facilities are available in the OGIM_v1 database". This will be another advantage of this dataset if all the entries are labeled with quality scores, and users with different research purposes can easily select them without any data cleaning processes. But I did not see the score from the dataset.

I uploaded a simple test code, maybe I missed them?

We have uploaded an update to the OGIM database (OGIM_v1.1) on Zenodo that includes the addition of VIIRS flaring data. Data quality scores, as described in the Main Text, are incorporated in the database attributes.