GPM-API: A Python Interface to Access the Global Precipitation Measurement Mission Satellites Open Data Archive

Gionata Ghiggi[1], Alexis Berne[1]
[1]Environmental Remote Sensing Laboratory, LTE, EPFL, Lausanne, Switzerland

<Objectives>

- **Unified access** to all GPM products **(26 passive microwave sensors** and **2 radars)**
- Make the **GPM archive** easier **exploitable,** also for **educational purposes**
- **Simplify** code tasks associated to data **download, reading, manipulation, data extraction, analysis**, **visualization and archiving** of GPM data
- **Increase** scientist **productivity** and foster research reproducibility
- Accelerate **model prototyping** and **development**
- Promote the **improvement** of precipitation retrieval **algorithms**
- **Widen** the **GPM users' community** to students and people with limited background

<Streamlined Workflow>

DOWNLOAD

```python
#### Download and Search GPM Data
gpm_api.download(
    product="2A-DPR",
    product_type="RS",
    version=7,
    start_time="2020-07-05 02:00:00", end_time="2020-07-08 06:00:00",
)
```

READING

```python
#### Open Archive in a CF-compliant xarray.Dataset
ds = gpm_api.open_dataset(
    product="2A-DPR",
    product_type="RS",
    version=7,
    start_time="2020-07-05 02:00:00", end_time="2020-07-08 06:00:00",
)
# Read only required data into RAM memory
da = ds["precipRateNearSurface"].compute()
```

MANIPULATION

```python
#### Slice n-dimensional (N-D) Arrays using Named Dimensions
# – Index based selection
ds_subset = ds.isel({
            "range": slice(0, 10),
            "along_track": slice(0, 100),
            "cross_track": slice(18, 28),
})
# – Value based selection
ds_subset = ds.sel(time=slice("2020-07-06 03:00", "2020-07-07 03:00"))

#### Straighforward Geospatial Manipulations
# – Retrieve list of overpass slices
list_isel_dict = ds.gpm_api.get_crop_slices_by_country("Switzerland")
list_isel_dict = ds.gpm_api.get_crop_slices_by_extent([6, 11, 45, 48])
# – Select a single overpass
isel_dict = list_isel_dict[0]
ds_crop = ds.isel(isel_dict)
```

DATA EXTRACTION

```python
#### N-D Object Labelling
ds = ds.ximage.label(
    variable="precipRateNearSurface",
    min_value_threshold=1,
    min_area_threshold=5,
    label_name="label_precip_area",
)
gpm_api.plot_labels(ds["label_precip_area"])

#### N-D Patch Extraction
ds_patch_gen = ds.ximage.label_patches(
    label_name="label_precip_area",
    patch_size={"cross_track": -1,"along_track": 100},
    centered_on="max",
    variable="precipRateNearSurface",
)
gpm_api.plot_patches(ds_patch_gen, variable="precipRateNearSurface")
```

ANALYSIS

```python
#### Community-based Retrievals
ds["EchoTop40dBZ"] = ds.gpm_api.retrieve("EchoTop40dBZ")
ds["MESH"] = ds.gpm_api.retrieve("MESH")

#### Powerful Visualizations
ds["EchoTop40dBZ"].gpm_api.plot_map()
ds["zFactorFinal"].isel(along_track=30).gpm_api.plot_transect()
```

ARCHIVING

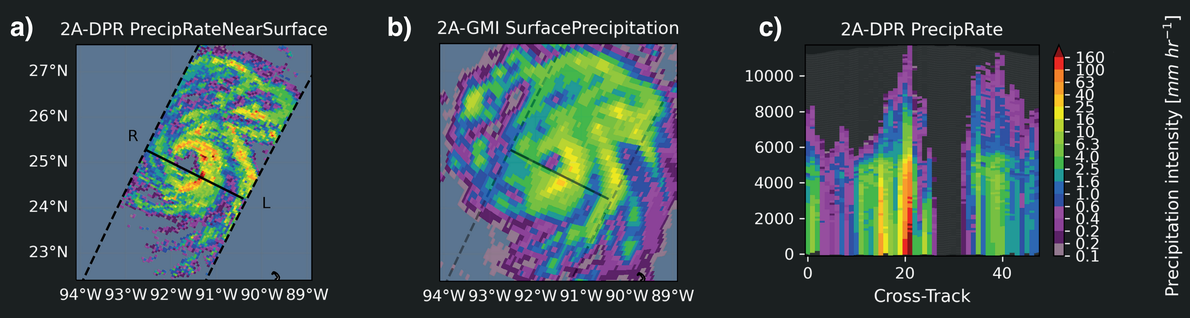
```python
#### Optimized Storage to Zarr
# – File size reduction up to 5x
# – Massive speed up in data reading (5x * n_threads)
ds.gpm_api.to_zarr("my_gpm_archive.zarr")

#### Geographic Archive Binning to Apache Parquet Dataset
ds.gpm_api.write_bucket(bucket_dir="my_gpm_bucket.parquet", lonbin=5, latbin=5)
```


<Software Stack>

xarray · dask · Zarr · Parquet · Apache Arrow

<Applications>

Local Analyses



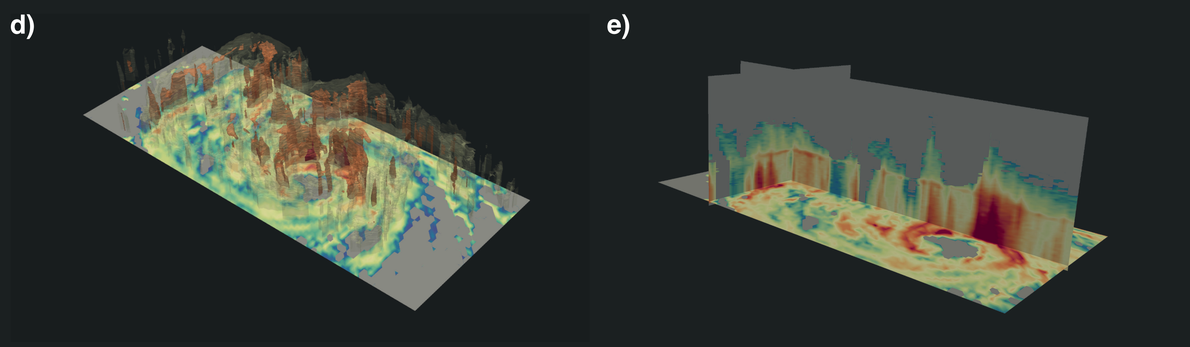


Figure 1. GPM DPR and GMI overpass during Hurricane Zeta on October 28, 2020, at approximately 08:25 UTC. (a) and (b) present a comparative analysis of surface precipitation estimates by GPM DPR and GPM GMI. The cross-track transect orientation for the vertical profile shown in (c) is indicated by the (L)eft and (R)ight labels at the ends of the transect line in (a). (d) and (e) showcase advanced 3D visualizations of DPR Ku-Band corrected reflectivities highlighting the spatial structure of the hurricane. In (d), GPM DPR isosurfaces are delineated at 30, 40, and 50 dBZ.

Global Analyses

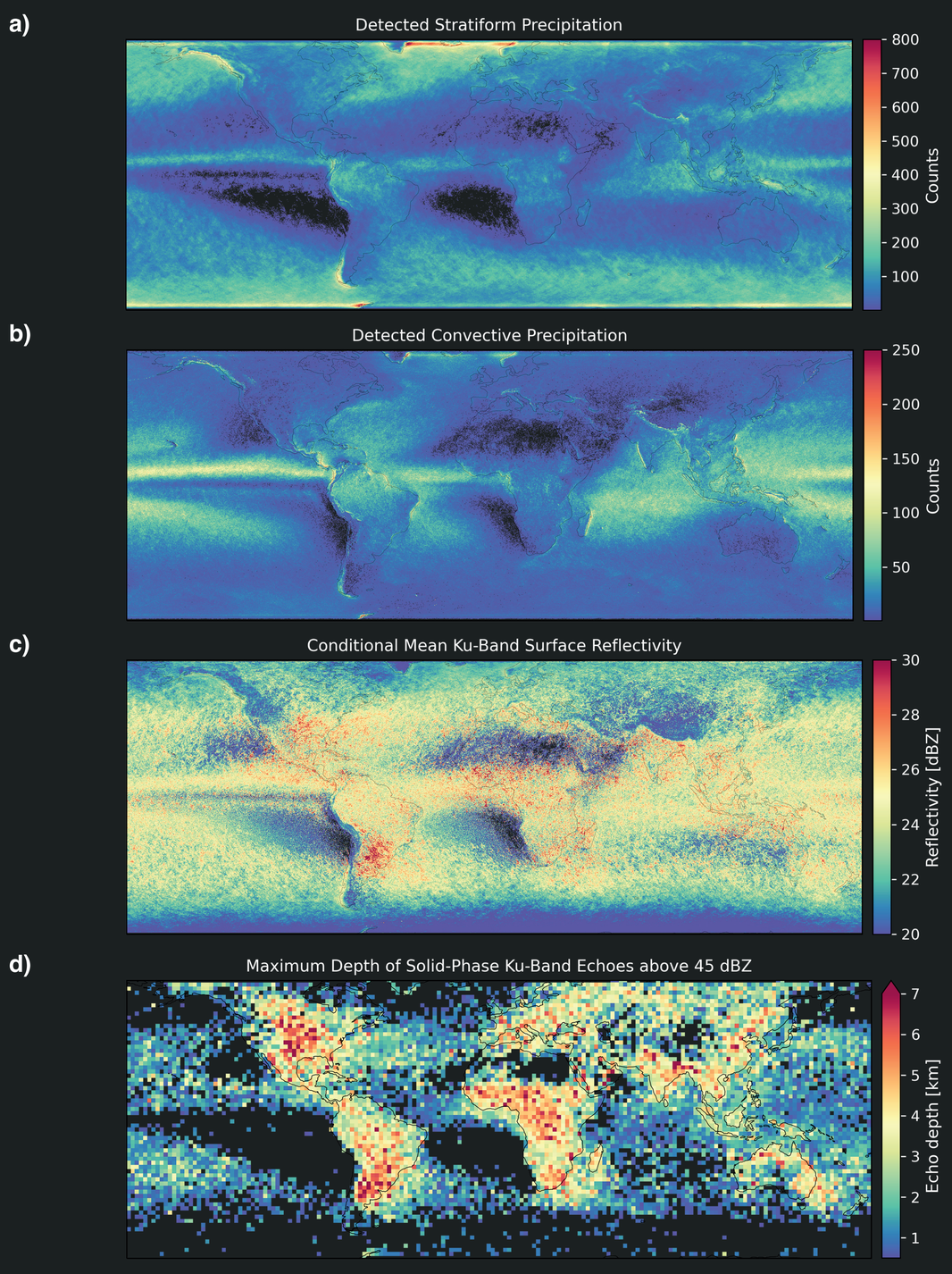


Figure 2: Spatial aggregation of GPM DPR measurements (2019-2023). (a) and (b) depict the count of convective (a) and stratiform (b) precipitation profiles within a grid of 0.1° x 0.1°. (c) reveals the conditional mean Ku-band surface reflectivity (at 0.1° x 0.1°) when precipitation is detected. (d) showcases the maximum depth of Ku-band echoes below 0°C and above 45 dBZ, with a spatial binning of 2° x 2°.

GET INVOLVED.
Get started by installing the software with "pip install gpm-api".
Contribute with your algorithms/retrievals, expertise and ideas.
Open a discussion on the GPM-API GitHub repository
(at https://github.com/ghiggi/gpm_api)