Python Interface to RSIG (pyrsig)

Presented: Barron H. Henderson

Date: 2023-05-05

RSIG Team: Jim J. Szykman, Luke Valin, Todd Plessel, Matt Freeman

***Disclaimer**: The views expressed in this presentation are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.*

barronh.github.io/pyrsig/ https://barronh.github.io/pyrsig/

pyrsig

Search docs

TABLE OF CONTENTS

- pyrsig User's Guide
- pyrsig Example Gallery
- pyrsig package

pyrsig User's Guide View page source

pyrsig User's Guide

Python interface to RSIG Web API

RSIG server prepares the data!

The key value of *pyrsig* is to present RSIG data in pandas DataFrames and xarray Datasets. This makes it easy to do advanced analyses in a pythonic way. Example analyses are provided, but the sky is the limit.

Getting Started

The best way to get started is to install (see below) and then explore the examples gallery.

Installation

pyrsig is avalable through pypi.org, but is still in rapid development. You can get the latest release from pypi via the command below.

```
pip install pyrsig
```

Get data examples

Examples showing how to get data.

Get DataFrame for PurpleAir PM25

Get DataFrame for TropOMI NO2

Get DataFrame for AQS ozone

Get IOAPI formatted NetCDF TropOMI NO2

Get List of Possible Coverages

Get COARDS formatted NetCDF TropOMI NO2

Get data examples

Examples showing how to get data.

Get DataFrame for PurpleAir PM25

Get DataFrame for TropOMI NO2

Get DataFrame for AQS ozone

Get IOAPI formatted NetCDF TropOMI NO2

```
import pyrsig

rsigapi = pyrsig.RsigApi(bdate='2022-03-01')
df = rsigapi.to_dataframe('tropomi.offl.no2.nitrogendioxide_tropospheric_column')
print(df.shape, *df.columns)
# (303444, 4) Timestamp(UTC) LONGITUDE(deg) LATITUDE(deg) nitrogendioxide_tropospheric_column(molecules/
```

pyrsig

Search docs

TABLE OF CONTENTS

pyrsig User's Guide

pyrsig Example Gallery

Get data examples

Oversample examples

Timeseries examples

Get data examples

Oversample examples

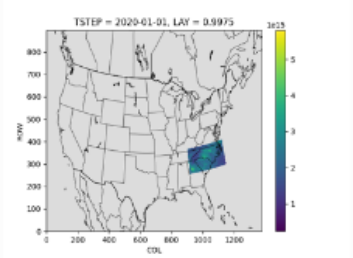
Oversample CONUS at 4k

Timeseries examples

pyrsig package

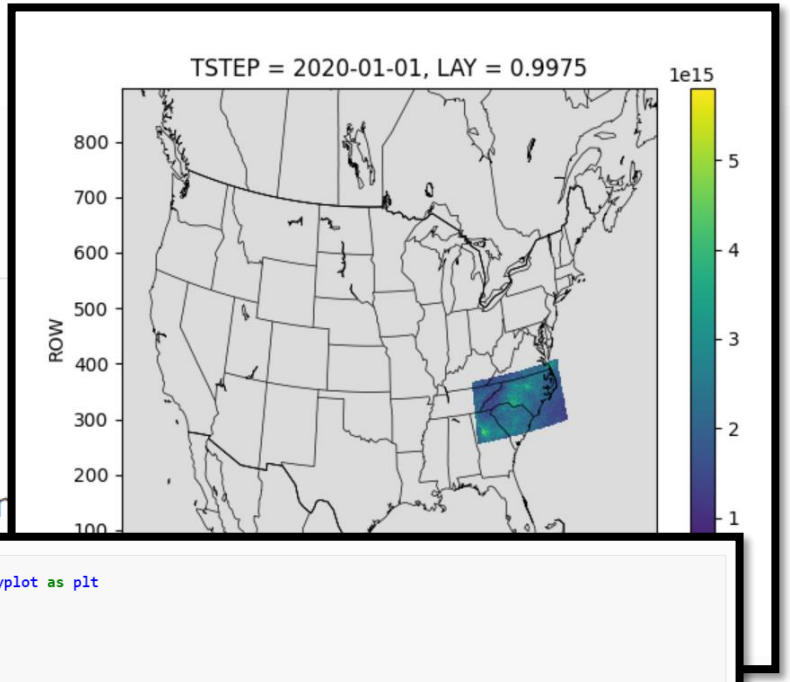
Oversample examples

Examples showing data oversampling of satellite data or



Oversample CONUS at 4k

Previous




```python
import matplotlib.pyplot as plt
import pyrsig
import pandas as pd
import xarray as xr
import pycno
import os


# Create a working directory
gdnam = '4US1'
bdate = '2021-01-01'
edate = '2021-01-15'
wdir = f'{gdnam}/{bdate[:4]}'

os.makedirs(gdnam, exist_ok=True)

rsigapi = pyrsig.RsigApi(
    bdate=bdate, bbox=(-85, 33, -75, 37),
    encoding={"zlib": True, "complevel": 1, "_FillValue": -9.999e36},
    workdir=wdir, grid_kw='4US1'
)

# Update to download daily averages instead of hourly
rsigapi.grid_kw['REGRID_AGGREGATE'] = 'daily'

# Loop over days
dss = []
for bdate in pd.date_range('2020-01-01', '2020-01-15'):
    print(bdate)
    try:
        ds = rsigapi.to_ioapi(key='tropomi.offl.no2.nitrogendioxide_tropospheric_column', bdate=bdate)
        dss.append(ds)
    except Exception as e:
```

pyrsig

Search docs

TABLE OF CONTENTS

pyrsig User's Guide

pyrsig Example Gallery

Get data examples

Oversample examples

Timeseries examples

Get data examples

Oversample examples

Timeseries examples

NYC VIIRS AOD vs TropOMI NO2

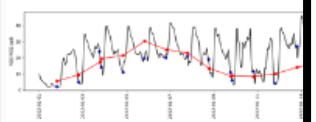
Phoenix AQS vs TropOMI

pyrsig package

Timeseries examp

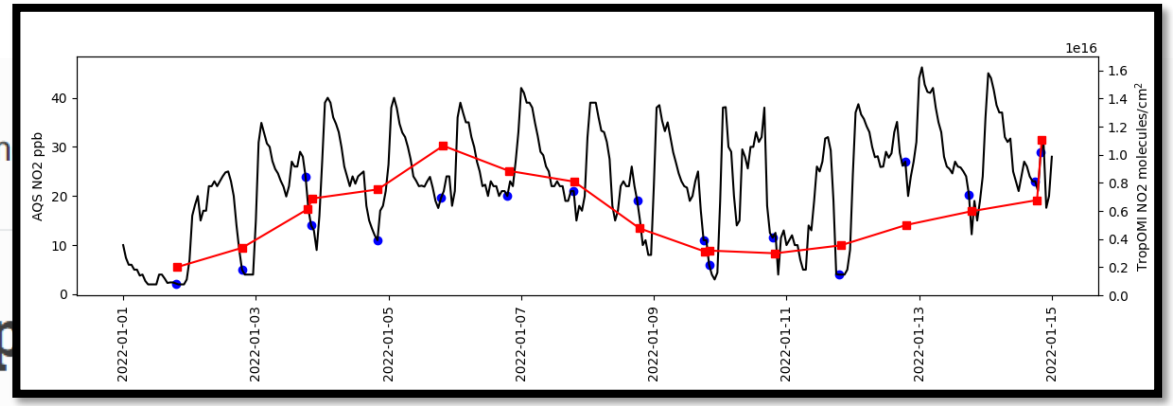
Examples showing timeseries analyses that il

NYC VIIRS AOD vs TropOMI NO2



Phoenix AQS vs TropOMI

Previous




```
import matplotlib.pyplot as plt
import pyrsig

# Create an RSIG api isntance
# Define a Time and Space Scope: here end of February around Phoenix
rsigapi = pyrsig.RsigApi(
    bdate='2022-01-01', edate='2022-01-15',
    bbox=(-112.3, 33.25, -111.85, 33.65)
)

# Get AQS NO2 with dates parsed and units removed from column names
aqsdf = rsigapi.to_dataframe('aqs.no2', parse_dates=True, unit_keys=False)

# Get TropOMI NO2
tomino2df = rsigapi.to_dataframe(
    'tropomi.offl.no2.nitrogendioxide_tropospheric_column',
    unit_keys=False, parse_dates=True
)


# Create spatial means for TropOMI and AQS
tomids = (
    tomino2df.groupby('time').median()['nitrogendioxide_tropospheric_column']
)
aqsds = aqsdf.groupby(['time']).median()['no2']

# Subset AQS to overpass times
oaqsds = aqsds.loc[aqsds.index.isin(tomids.index.floor('1h'))] # just overpass t

# Create axes with shared x
fig, ax = plt.subplots(figsize=(12, 4),
                       gridspec_kw=dict(bottom=0.25, left=0.05, right=0.95))
ax.tick_params(axis='x', labelrotation = 90)
tax = ax.twinx()
```

Run Example On Colab

General

- colab.research.google.com
  - Open a new notebook
  - Add “%pip install --user netcdf4 pyproj pycno pyrsig”
  - And run it
- barronh.github.io/pyrsig
  - Navigate to example
  - Copy code from example
  - Paste it in your colab notebook
  - Run it

Today prepackaged examples:

- gist.github.com/barronh/6faf2008a0cf0ac9e49543f31c300601
- gist.github.com/barronh/497e3418fb66ddce2d04a4c8d68e31a3

https://gaftp.epa.gov/Air/aqmg/bhenders/tutorials.html

Questions?

pyrsig: henderson.barron@epa.gov

RSIG: rsig@epa.gov