List of Useful Python Packages for Atmospheric, Ocean and Climate Sciences

wxguy

2022-05-25 21:10:00 +0530

Contents

Here is the list of Python modules, packages, and programs that I keep on referring to or visiting often. Most of the packages can be installed using `conda` , some using `pip` , and a few need to be installed using `setup.py` script. Descriptions of packages are taken from official sites or documentation. The list is exhaustive and hopes that some will benefit from my research while looking for the specific use case.

Integrated Development Environments or Text Editors

PyCharm Community Edition. It comes with a lot of goodies by default. Install it only if you have enough hardware resources.

Sublime Text. Install Sublime Text with Anaconda and Tabnine plugins to make it a truly remarkable IDE.

Spyder. Highly recommended if you are working on Scientific applications. This is also the recommended IDE by the scientific community.

ipython. It is not a direct scientific Python package. IPython provides a rich architecture for interactive computing.

Core Packages

numpy. The fundamental package for scientific computing with Python. Core package around in which other scientific modules are built.

matplotlib. A comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib makes easy things easy and hard things possible.

scipy. Fundamental algorithms for scientific computing in Python. It provides algorithms for optimization, integration, interpolation, eigenvalue problems, algebraic equations, differential equations, statistics, and many other classes of problems.

sympy. SymPy is a Python library for symbolic mathematics. It aims to become a full-featured computer algebra system (CAS) while keeping the code as simple as possible in order to be comprehensible and easily extensible. SymPy is written entirely in Python.

Data Input/ Output

NetCDF

netcdf4-python. A Python interface to the netCDF C library by unidata who is also the creator of netCDF itself. Most of the other easy-to-use NetCDF libraries are built on top of this library.

xarray. It is inspired by and borrows heavily from pandas, the popular data analysis package focused on labeled tabular data. It is particularly tailored to working with netCDF files, which were the source of xarray's data model, and integrates tightly with dask for parallel computing.

GRIB 1/2

pygrib. A high-level Python interface to ECCODES library for GRIB file IO operations. If you are working on GRIB2 files, you can make use of this library.

HDF5

h5py. Pythonic interface to the HDF5 binary data format. It lets you store huge amounts of numerical data, and easily manipulate that data from NumPy.

Satellites

satpy. It is a python library for reading, manipulating, and writing data from remote-sensing earth-observing satellite instruments. Satpy provides users with readers that convert geophysical parameters from various file formats to the common Xarray DataArray and Dataset classes for easier interoperability with other scientific Python libraries.

WxRadar

wradlib. The ωradlib project has been initiated in order to facilitate the use of weather radar data as well as to provide a common platform for research on new algorithms. It supports various DWR radar formats available in the market.

pycinrad. PyCINRAD is an open-source weather radar library that supports the reading of all mainstream radar formats in China and provides some practical algorithms and visualizations.

pycwr. The China Weather Radar Toolkit supports most of China's radar formats(WSR98D, CINRAD/SA/SB/CB, CINRAD/CC/CCJ, CINRAD/SC/CD).

Radiosondes

pyigra. The National Centres for Environmental Information provides a large global archive for vertical soundings also known as radiosondes. PyIGRA downloads and extracts data from the Integrated Global Radiosonde Archive data set and prepares them in a more handy data format.

Other Formats

pygdal. GDAL, the Geospatial Data Abstraction Library, is an access and translator library for raster and vector geospatial data formats, often known as the swiss knife of geospatial. pygdal is a Python interface to the GDAL library.

argos. Argos is a GUI for viewing and exploring scientific data, written in Python and Qt. It has a plug-in architecture that allows it to be extended to read new data formats. At the moment plug-ins are included to read HDF-5, NetCDF-4, WAV, Exdir, NumPy binary files, and various image formats, but a plug-in could be written for any data that can be expressed as a Numpy array.

siphon. Siphon is a collection of Python utilities for downloading data from remote data services. Much of Siphon's current functionality focuses on access to data hosted on a THREDDS Data Server, although access to data hosted by other web technologies is being quickly added.

mkmov. This utility is designed to make a movie from a NetCDF file or stitch together a series of .png files. The interface is by command line and everything is done in one line.

Data Science, Machine Learning, and Artificial Intelligence

pandas. A fast, powerful, flexible, and easy-to-use open-source data analysis and manipulation tool, built on top of the Python programming language. This is the base package for working with a statistical dataset in Python.

seaborn. It is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

theano. Theano is a Python library that allows you to define, optimize, and efficiently evaluate mathematical expressions involving multi-dimensional arrays.

tensorflow. TensorFlow is an open-source library for fast numerical computing. This module makes it easy for beginners and experts to create machine learning models.

keras. Keras is an API designed for human beings, not machines. Keras follows best practices for reducing cognitive load: it offers consistent & simple APIs, it minimizes the number of user actions required for common use cases, and it provides clear & actionable error messages.

pytorch. PyTorch has a range of tools and libraries that support computer vision, machine learning, and natural language processing. The library is open-source and is based on the Torch library.

scikit-learn. It is a powerful Python library that was originally generated to serve the purpose of data modeling and building machine learning algorithms. It has a simple, engaging, and consistent interface that is exceptionally user-friendly, making it easy to use and share data.

geopandas. It is an open-source project to make working with geospatial data in python easier. GeoPandas extends the datatypes used by pandas to allow spatial operations on geometric types.

pytables. PyTables is a package for managing hierarchical datasets and is designed to efficiently and easily cope with extremely large amounts of data.

Atmospheric Science

metpy. MetPy is a collection of tools in Python for reading, visualizing, and performing calculations with weather data.

wrf-python. A collection of diagnostic and interpolation routines for use with output from the Weather Research and Forecasting (WRF-ARW) Model.

wrf_install. automate the downloading & building process of the wrf model with all dependencies.

getgfs. Extracts weather forecast variables from the NOAA GFS forecast with no obscure, platform-specific, dependencies.

pyart. It is a Python module containing a collection of weather radar algorithms and utilities.

prob_meteogram. An intuitive probabilistc meteogram based on ensemble forecasts.

SkewTplus. SkewTplus provides useful tools for plotting and analyzing atmospheric sounding data. In particular, it provides useful classes to handle the awkward skew-x projection.

tropycal. Python package intended to simplify the process of retrieving and analyzing tropical cyclone data, both for past storms and in real-time, and is geared towards the research and operational meteorology sectors.

aoslib. aoslib is a Python library of standard atmospheric and oceanic sciences calculation routines. It exists mainly so we're all not writing our own routines to calculate potential temperature, isentropic potential vorticity, etc.

imdlib. IMDLIB is a python package to download and handle binary gridded data from the India Meteorological Department (IMD).

verif. Verif is a command-line tool that lets you verify the quality of weather forecasts for point locations. It can also compare forecasts from different forecasting systems (that have different models, post-processing methods, etc).

eofs. eofs is a Python package for EOF analysis of spatial-temporal data. Using EOFs (empirical orthogonal functions) is a common technique to decompose a signal varying in time and space into a form that is easier to interpret in terms of spatial and temporal variance.

windspharm. windspharm is a Python package for performing computations on global wind fields in spherical geometry. It provides a high-level interface for computations using spherical harmonics.

windrose. A windrose, also known as a polar rose plot, is a special diagram for representing the distribution of meteorological data, typically wind speeds by class and direction. This is a simple module for the matplotlib python library, which requires NumPy for internal computation.

tephigrams. Python package for plotting tephigrams.

pymeteo. Skew-T/Log-P plotting routine, meteorology, and interfacing with CM1. This package provides routines for plotting Skew-T/Log-P diagrams and working with CM1 output.

metar. metar is a python package for interpreting METAR and SPECI coded weather reports.

PythonMETAR. Yet another Python package for interpreting METAR and SPECI weather reports.

xcape. xcape provides an efficient workflow for the calculation of parameters commonly used in meteorology, allowing for post-processing of 4D gridded model or observed data using xarray for parameters including Convective Available Potential Energy (CAPE) and Storm Relative Helicity (SRH). This is made possible by leveraging wrapped Fortran routines and NumPy and allows for process parallelization via Dask.

act. Atmospheric Community Toolkit (ACT) is an open-source Python toolkit for working with atmospheric time-series datasets of varying dimensions. The toolkit is meant to have functions for every part of the scientific process; discovery, IO, quality control, corrections, retrievals, visualization, and analysis.

tcpyPI. tcpyPI, 'pyPI' for short, is a set of scripts and notebooks that compute and validate tropical cyclone (TC) potential intensity (PI) calculations in Python.

pyresample. Pyresample is a python package for resampling geospatial image data. It is the primary method for resampling in the SatPy library, but can also be used as a standalone library.

gridded. gridded is a single API for accessing/working with gridded model results on multiple grid types.

xESMF. xESMF is a Python package for regrinding. It is powerful, easy to use, and fast.

uxarray. Uxarray aims to provide xarray-styled functionality for unstructured grid datasets following ugrid conventions.

sharppy. SHARPpy is a collection of open-source sounding and hodograph analysis routines, a sounding plotting package, and an interactive, cross-platform application for analyzing real-time soundings all written in Python. It was developed to provide the atmospheric science community with a free and consistent source of sounding analysis routines.

pywrf. Python tools to launch and process WRF simulations. The code is a little old but still useful.

wrfplot. Python script to plot various WRF-ARW outputs. Command-line driven by passing options. The script is written in Python 2.x but gives you insight on processing wrfout files.

new_wrf_plotting_functions. Yet another collections of functions for plotting WRF data.

wrftools. Functions for post-processing WRF output data.

wrf-aurorun. Scripts to automate a WRF model run.

Data Visualisation

cartopy. This package is designed for geospatial data processing to produce maps and other geospatial data analyses. It makes use of the powerful PROJ, NumPy, and Shapely libraries and includes a programmatic interface built on top of Matplotlib for the creation of publication-quality maps.

colormaps. Colormaps is a library of a collection of colormaps or color palettes for Python. It's written in Python with matplotlib and NumPy as dependencies. You can use Colormaps to customize matplotlib plots.

cmocean. This package contains colormaps for commonly-used oceanographic variables. Most of the colormaps started from matplotlib colormaps, but have now been adjusted using the viscm tool to be perceptually uniform.

psyplot. psyplot is an open-source python project that mainly combines the plotting utilities of matplotlib and the data management of the xarray package. The main

purpose is to have a framework that allows a fast, attractive, flexible, easily applicable, easily reproducible, and especially an interactive visualization of your data.

bokeh. Bokeh is an interactive visualization library that targets modern web browsers for presentation. Bokeh can help anyone who would like to quickly and easily connect powerful PyData tools to interactive plots, dashboards, and data applications.

pyferret. Python interface to famous ferret visualisation tool for oceanographers.

[gradspy(]http://cola.gmu.edu/grads/gadoc/python.html). Official Python extension of Grads plotting software.

py3grads. Yet another Python 3 Interface to GrADS.

pygmt. PyGMT is a library for processing geospatial and geophysical data and making publication-quality maps and figures. It provides a Pythonic interface for the Generic Mapping Tools (GMT), a command-line program widely used in the Earth Sciences.

geoviews. GeoViews is a Python library that makes it easy to explore and visualize geographical, meteorological, and oceanographic datasets, such as those used in weather, climate, and remote sensing research.

geocat-viz. The GeoCAT-viz repo contains tools to help plot data, including convenience and plotting functions that are used to facilitate plotting geosciences data with Matplotlib, Cartopy, and possibly other Python ecosystems plotting packages.

pylustrator. Pylustrator is software to prepare your figures for publication in a reproducible way. This means you receive a figure representing your data alongside a generated code file that can exactly reproduce the figure as you put them in the publication, without the need to readjust things in external programs.

pyproj. Python interface to PROJ (cartographic projections and coordinate transformations library). Most of the Plotting packages which require geo-reference uses this package.

Climate and Forecast (CF) Conventions

cftime. Python library for decoding time units and variable values in a netCDF file conforming to the Climate and Forecasting (CF) netCDF conventions.

cf-plot. A set of Python routines for making the common contour, vector, and line plots that climate researchers use.

cf-python. An Earth Science data analysis library that is built on a complete implementation of the CF data model.

cfdm. This library implements the data model of the CF (Climate and Forecast) metadata conventions http://cfconventions.org and so should be able to represent and manipulate all existing and conceivable CF-compliant datasets.

cf-checker. The CF Checker is a utility that checks the contents of a NetCDF file and complies with the Climate and Forecasts (CF) Metadata Convention.

cf_xarray. cf_xarray mainly provides an accessor (DataArray.cf or Dataset.cf) that allows you to interpret Climate and Forecast metadata convention attributes present on xarray objects.

Climate Science

climetlab. CliMetLab is a Python package aiming at simplifying access to climate and meteorological datasets, allowing users to focus on science instead of technical issues such as data access and data formats.

aospy. aospy (pronounced A - O - S - py) is an open-source Python package for automating your computations that use gridded climate and weather data (namely data stored as netCDF files) and the management of the results of those computations.

xclim. xclim is a library of functions to compute climate indices from observations or model simulations. Its objective is to make it as simple as possible for users to compute indices from large climate datasets and for scientists to write new indices with very little boilerplate.

climate-indices. This project contains Python implementations of various climate index algorithms which provide a geographical and temporal picture of the severity of precipitation and temperature anomalies useful for climate monitoring and research.

Oceanographic Science

verde. Verde is a Python library for processing spatial data (bathymetry, geophysics surveys, etc) and interpolating it on regular grids (i.e., gridding).

oceanspy. It is a Python package to facilitate ocean model data analysis and visualization.

sea-py. Python Tools for Oceanographic Analysis. It is not a package by itself. It is a webpage with a collection of useful oceanography-related analysis tools and tutorials.

xmitgcm. It is a python package for reading MITgcm binary MDS files into xarray data structures. By storing data in dask arrays, xmitgcm enables parallel, out-of-core analysis of MITgcm output data.

octant. A Python library used for post-processing the output of PMCTRACK - vorticity-based cyclone tracking algorithm.

pyroms. Pyroms is a collection of tools to process input and output files from the Regional Ocean Modeling System, ROMS. It was originally started by Rob Hetland as a google code project, then he morphed it into octant, also at google code. Frederic Castruccio then created a fork and renamed it back to pyroms.

ctd. Tools to load hydrographic data as pandas DataFrame with some handy methods for data pre-processing and analysis. This module can load SeaBird CTD (CNV), Sippican XBT (EDF), and Falmouth CTD (ASCII) formats.

seawater. This is a Python library for calculating the properties of seawater. The package uses the formulas from Unesco's joint panel on oceanographic tables and standards, UNESCO 1981 and UNESCO 1983 (EOS-80).

gsw. This Python implementation of the Thermodynamic Equation of Seawater 2010 (TEOS-10) is based primarily on NumPy ufunc wrappers of the GSW-C implementation.

ttide_py. Direct conversion of T_Tide to Python.

pytides. Pytides is a small Python package for the analysis and prediction of tides. Pytides can be used to extrapolate the tidal behavior at a given location from its previous behaviour.

pywavelets. It is open-source wavelet transform software for Python. It combines a simple high-level interface with low-level C and Cython performance.

pywafo. It is WAFO toolbox Python routines for statistical analysis and simulation of random waves and random loads.

okean. ocean modeling and analysis tools. But unfortunately, there is no documentation provided on the site.

pymc. PyMC3 allows you to write down models using an intuitive syntax to describe a data-generating process.

pygridgen. A Python interface to Pavel Sakov's gridgen library.

mixsea. Ocean mixing parameterizations in python.

xgcm. xgcm is a python package for working with the datasets produced by numerical General Circulation Models (GCMs) and similar gridded datasets that are amenable to finite volume analysis.

Geo Spatial

qgis. QGIS is a free and open-source cross-platform desktop geographic information system application that supports viewing, editing, printing, and analysis of geospatial data. It also has a Python API to extend its ability.

pysal. Python spatial analysis library is an open-source cross-platform library for geospatial data science with an emphasis on geospatial vector data written in Python.

geopy. It is a Python client for several popular geocoding web services. geopy makes it easy for Python developers to locate the coordinates of addresses, cities, countries, and landmarks across the globe using third-party geocoders and other data sources.

Astrophysics, Seismology Etc.,

argopy. argopy is a python library dedicated to Argo data access, manipulation, and visualisation for standard users as well as Argo experts.

yt. yt is a tool for querying, analyzing, and visualizing objects or regions of interest to identify emergent properties in data available in a variety of real-world research data formats. Initially developed for use by professional astronomers, yt can be applied in a variety of domains including astrophysics, seismology, nuclear engineering, molecular dynamics, and oceanography.

sunpy. SunPy is a Python-based software library that provides tools for performing research using direct observations of the Sun and Heliosphere.

You can download this article from here for free.