Indian Institute of Astrophysics

Summer Project 2019

Evaluation of Planetary Boundary Layer (PBL) height using multi-reanalysis and CALIOP observations over Hanle and Merak sites

Sachin P

Guide – Dr. Shantikumar Singh Nigombam

ABSTRACT

Planetary Boundary Layer (PBL) height is defined as the turbulent mixing layer near the earth's surface (0-3km). PBL is strongly influenced on Astronomical seeing (day and night) due to variations of the optical refractive index. The present work describes about the characteristics of PBL heights from various reanalysis datasets (e.g., Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2), ECMWF Re-Analysis (ERA)-Interim (ERA-Interim) and ERA-5 (version 5) for the period of 1979–2018 with spatial resolution of 0.25x0.25 degree, and temporal resolution of 3 hours for ERA-Interim, and hourly for MERRA-2 and ERA-5. These data are stored in Hierarchical *Data* Format (HDF) and Network Common *Data* Format (NetCDF). Such data formates are extracted for Hanle and Merak sites using Python. Since PBL height are related to night-time seeing and day-time r0 values, the present work is examining the influence of seeing and r0 values on PBL heights for Hanle and Merak sites, respectively and the study is under progress. The present work was also initially planned to study the PBL heights using the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) Lidar data from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO), but due to some technical issues, the present work has dropped.

CERTIFICATE

This is to certify that **Mr. Sachin P** has completed the project work titled - **Evaluation of Planetary Boundary Layer (PBL) height using multi-reanalysis and CALIOP observations over Hanle and Merak sites**, under my guidance and supervision during the period June 3, 2019 to July 5, 2019 at the Indian Institute of Astrophysics, Bangalore, India.

Guide
Dr. NS Shantikumar
Indian Institute of Astrophysics
Bangalore, India

ACKNOWLEDGEMENTS

I would like to acknowledge Dr. NS Shantikumar for guiding be throughout the project and helping me understand the theory behind PBLH and other atmospheric parameters.

I would also like to thank Dr. Hwan-Jin Song, Seoul National University, who provided NetCDF4 files for PBLH over Hanle and Merak along with example codes in IDL to extract them without which this project would have been really hard.

I also want to thank Dr. Sridharan, Dr. B Ravindara and K Prabu of the Indian Institute of Astrophysics for providing with Seeing and cloud cover data from their previous work and institute archives.

Table of Contents

INTRODUCTION

Planetary Boundary Layer

Boundary layer is said to be the fluid that is near to a surface and is directly influenced by it. There is a lot of viscosity involved. Planetary boundary layer is a layer where we see a lot of turbulence. It has a very profound influence in physics of atmosphere and strongly influences cloud processes, ocean surface fluxes and so on. The height of the boundary layer is a great measure of PBL and can be related to cloud cover, and to turbulence profiles that relate to astronomical seeing. [1]

Knowing planetary boundary layer height over a certain location can therefore help us in astronomical site surveys. Here the PBLH is extracted for Hanle (32.75 N,79.00 E) and Merak (33.5 N,78.5 E) where Indian Astronomical Observatory(IAO) of the Indian Institute of Astrophysics(IIA) is set up. It's effect on astronomical seeing, and cloud cover can be understood, along with other parameters like wind gust and Temperature.

The PBLH is calculated as reanalysis data and are available from various different sources including ERA-Interim, ERA5 from European Centre for Medium-range Weather Forecasts (ECMWF), and from Modern-era Retrospective Analysis for Research & Applications (MERRA) from Global Modelling and Assimilation Office (GMAO) collaborated with NASA's Earth Science Mission.

Datasets and file structure

Various organisations use variational analysis and data assimilation systems along with forecast models to reprocess archived observations and combine various parameters to predict variables that cannot be directly observed. Reanalysis data provide a multivariate, and spatially complete record of the global atmospheric circulation and successive generations of reanalyses makes the data better and more accurate. [3]

These data are available in the form of NetCDF4 and HDF files. This project included the extraction of these files and read required data from it and store it in formats with simpler usability or in a database.

To download these data, basic parameters of each dataset including their spatial and temporal resolutions are required. A brief comparison between the reanalysis datasets (ERA-Interim, ERA5 and MERRA) are produced in the form of a table below. [3, 4]

ERA-Interim	ERA5	MERRA
• IFS 4D-Var • Spatial resolution – 0.75x0.75 • Temporal Res – 3 hourly for surface and 6 hourly for upper atmospheric • Sequential 12 hour cycles, each using prior cycle's forecast and combining it with available observations.	• Upgraded IFS 4D-VAR • Spatial resolution – 0.28125x0.28125 • Temporal res – hourly for surface and 3 hourly for some upper atmospheric • Similar to ERA-I but also combines it with error estimates of EDA(ensemble of data assimilation)	• Upgraded IFS 4D-VAR • Spatial resolution – 0.28125x0.28125 • Temporal res – hourly for surface and 3 hourly for some upper atmospheric • Similar to ERA-I but also combines it with error estimates of EDA(ensemble of data assimilation)

As much as the data gets better each generation and the models and the observations used are improved, reanalysis data is still a model to predict unavailable variables at various time, and location. Accuracy of reanalysis data will be lesser than satellite observed data.

CALIPSO

One such satellite data from where PBLH can be extracted is the Cloud Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) that captures the attenuated backscatter, cloud emissivity and particle size along with images.
The CALIPSO Payload consists of [11]:

- **Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP)** comprising of one 1064nm receiver of backscatter intensity and two channels measuring orthogonally polarized signal from 532nm backscatter.
- **Wide Field Camera (WFC)**, a fixed nadir-viewing imager covering 630-670nm spectral region.
- **Imaging Infrared Radiometer (IIR)**, providing measurements at three channels in the thermal infrared window region at 8.7 mm, 10.5 mm, and 12.0 mm, which optimise CALIOP/IIR retrievals.

The spatial resolution of CALIPSO is 1/3km, 1 km and 5 km. The satellites of the A-train are in a 705km sun synchronous polar orbit, giving a 16-day repeat cycle over a location [10]

PBLH can be derived from backscatter data using methods described in [9] based on the theory that maximum backscattering is observed on top of PBL due to temperature inversion as described in [8]. A more recent method as described by [8] is described in Fig 1.1 below.

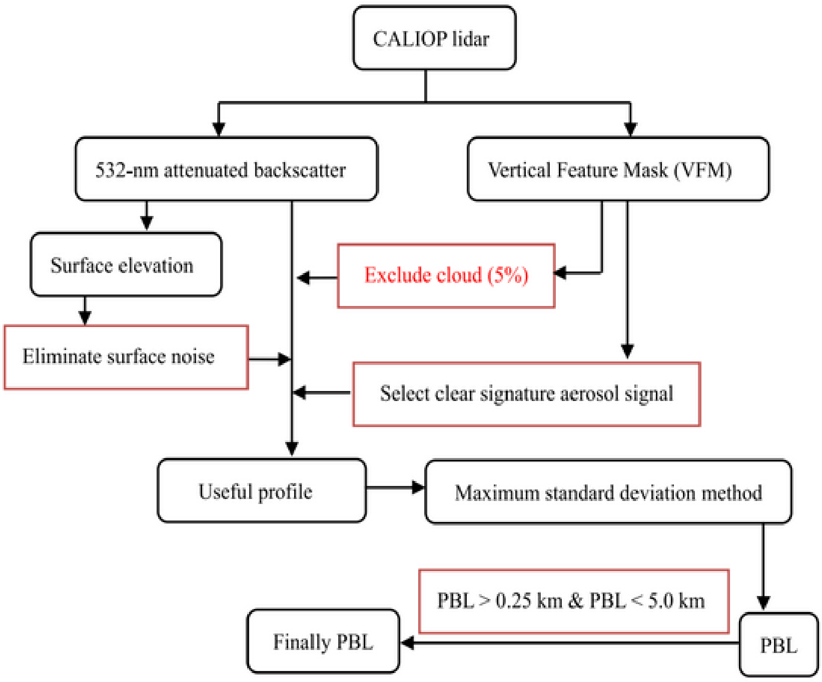

Fig 1.1: Describes the method followed in [7] to derive PBLH using backscatter data from CALIPSO.

A simpler adaptation of this method to derive PBLH from satellite observation data of CALIPSO was a part of this project but was eventually dropped due to technical issues pertaining to acquiring CALIPSO data within alotted time.

Once the file is downloaded using API/directly having understood a bit about different datasets, it is required to know to extract data from these file formats, and for that an attempt was made to understand the basic file structure of these file formats.

Hierarchy data format or HDF is a library and multi-object file format for the transfer of graphical and numerical data. One can access the structure and multi-type data without outside information. [6].

Network Common Data Format or NetCDF also stores array oriented scientific data. NetCDF is an abstraction that supports a view of data as a collection of self-describing, portable objects that can be accessed through a simple interface. Array values may be accessed directly, without knowing details of how the data are stored. Auxiliary information about the data, such as what units are used, may be stored with the data. [5].

Both of these formats allow flexible reading of data. Direct access instead of sequential access, and are widely used by the scientific community to store and read large data easily.

The structure of file structure HDF and NetCDF broadly include Groups and datasets. Each of these have headers that include information about the data in form of objects called dimensions, attributes and variables, and then the array of variable data itself as seen in Fig 1.2. The file structure of HDF and NetCDF became very similar with newer versions and NetCDF4 and HDF5 are currently interoperable.

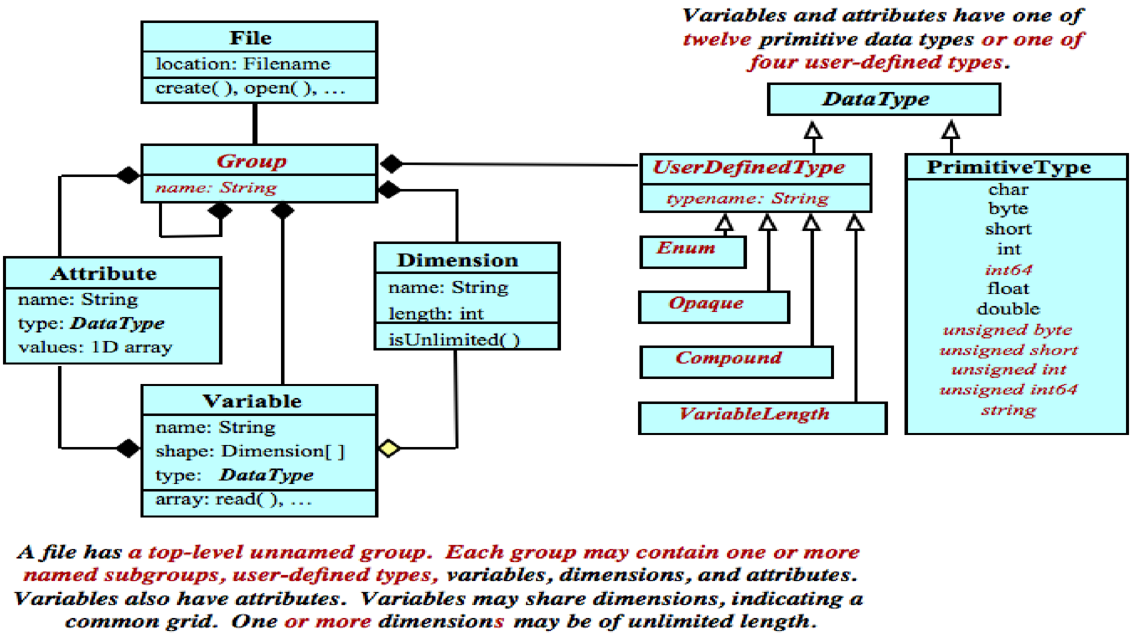

Fig 1.2: Describes file structure of HDF File [6]

These file structures are program and system independent and do not require users to worry about compatibility. They are accessed and manipulated through programming models that are part of these files and libraries in various computer languages that can access these programming models to give user the ability to manipulate the abstract data stored in the files. A simple flowchart of this is given in Fig 1.3.

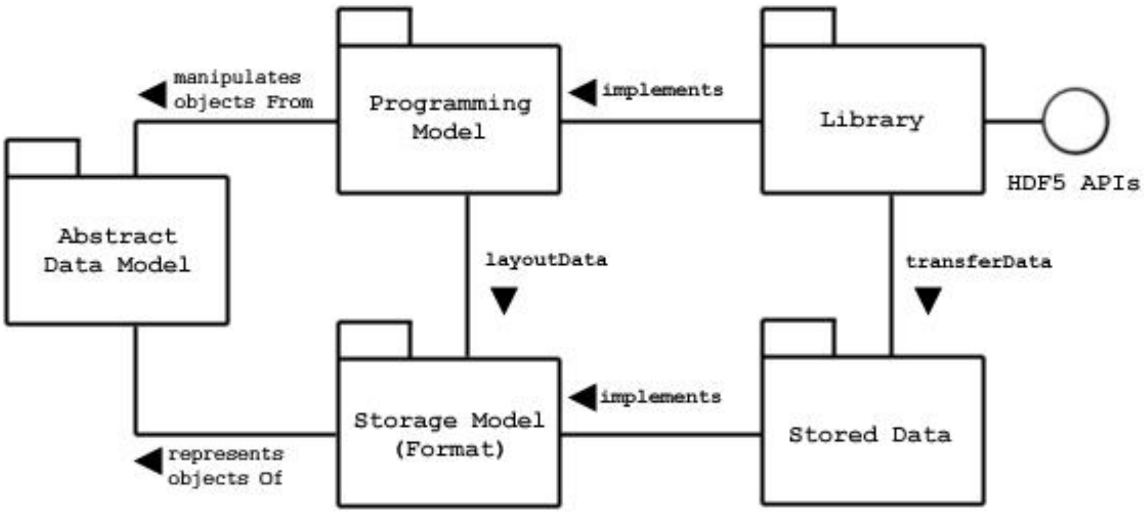

Fig 1.3: Describes how HDF files are accessed [6]

Seeing data and FITS Images

Once PBLH data can be extracted it becomes necessary to obtain seeing data for the sites, and analyse PBLH and seeing. Astronomical seeing is apparent blurring of the stars due to turbulence in the atmosphere.

Seeing is characterised by FWHM of the seeing disk of the telescope or by the Fried's parameter (r0) that tells us the size of the lump of turbulence seen while Imaging [11, 12].

Seeing data of an image from a telescope is stored along with the image files using the Flexible Image Transport System (FITS), file format. FITS file format allows storage of image as an array as well as other Header Data Units (HDU) that contain information about the image in form of dictionaries and tables referred to as cards in the HDU's of FITS files. [13]

DATA EXTRACTION

The reanalysis data from ERA Interim, ERA5 and MERRA was downloaded from the respective organisations online archive as NetCDF4 files. Python programming language was used to extract data from these files and store them in .csv or .txt format. These formats allow simpler navigation and manipulation of the data.

Python has a module to handle NetCDF4 files, called NetCDF which has various functions and attributes to open .nc files and read their content and extract the data out of them. The Dataset() method allows reading of .nc files, which can later be combined with methods such as Dimensions(), Variables(), and Attributes() to access required header information and data from the NetCDF4 files. The Varaibles() method returns the data stored under a header in form of a masked numpy array, which can be stored in a database. [14]

The x-array package also supports reading .nc files once the netCDF4 module is installed along with it. X-array has opendataset() method that allows one to open a .nc file and then the to_dataframe() method can be used to convert it into a Pandas dataframe. For those already proficient with Pandas dataframe, this method can be more convenient. [15]

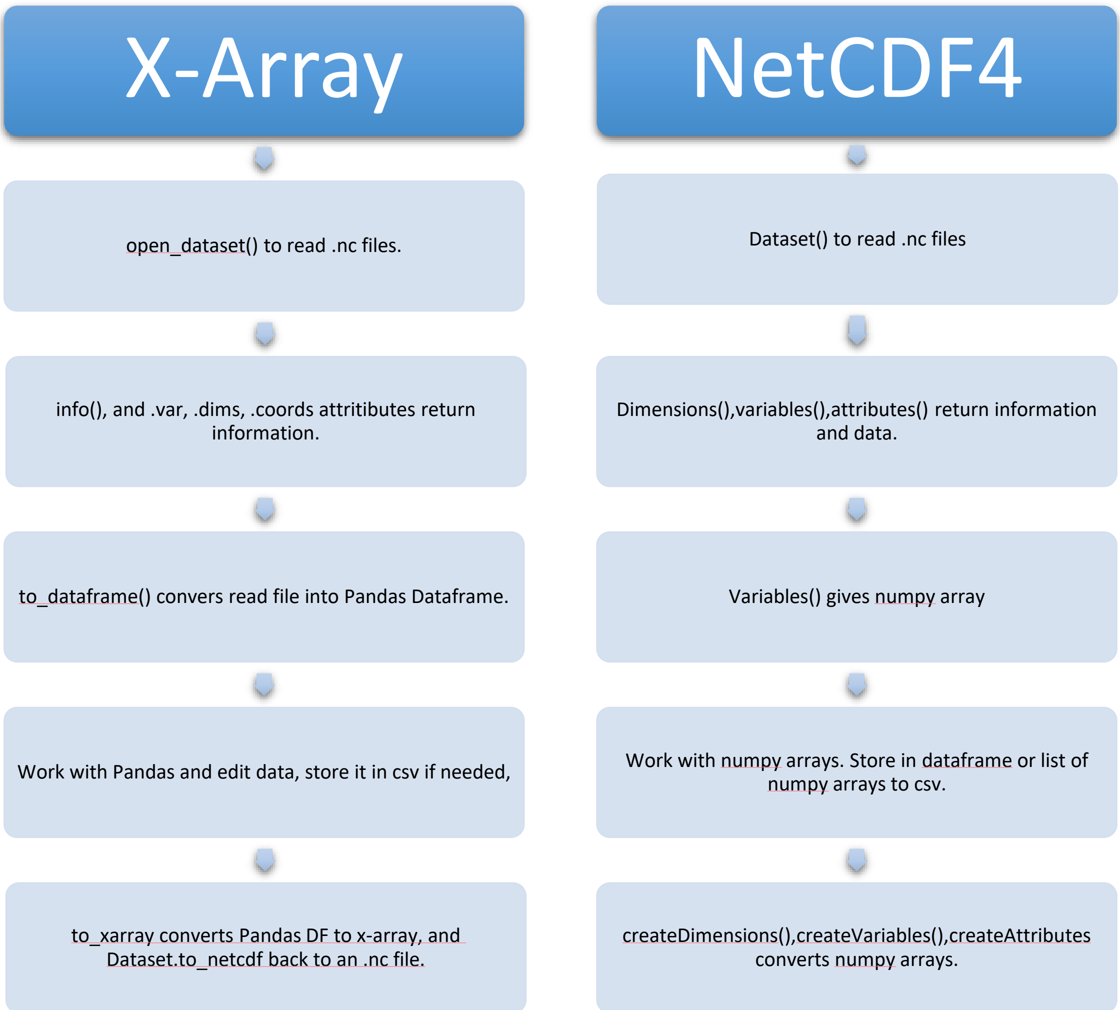

Fig 2.1 describes in brief the basic methods to work with .nc files in python. More information can be obtained from [14, 15]

The x-array methods were extensively used and data was manipulated with Pandas to get hourly, monthly medians, along with segregation into daytime and nighttime PBLH Values.

Seeing Data can similarly be extracted using astropy package that comes with FITS support. Methods such as header(), and data attribute allows one to access data value of a given header. Once a FITS file is available fits.open() reads the FITS file and the info() method can be used to understand the dimensions and HDU list of the FITS file. [13]

FITS file come as individual files with individual observation times. One needs to make sure to download sufficient FITS files at different date & time, and extract observation time and FWHM value (or r0 value) and store them to obtain a comprehensive seeing data. Some data were extracted, and others were provided for the purpose of the Project by the institute from the archives.

The PBLH data and seeing data were the time-matched, and corresponding medians and means were taken to analyse them.

ANALYSIS AND DISCUSSION

In this section a number of plots are presented along with a brief discussion about them. Some of the plots obtained during this project are compared with earlier plots from previous papers.

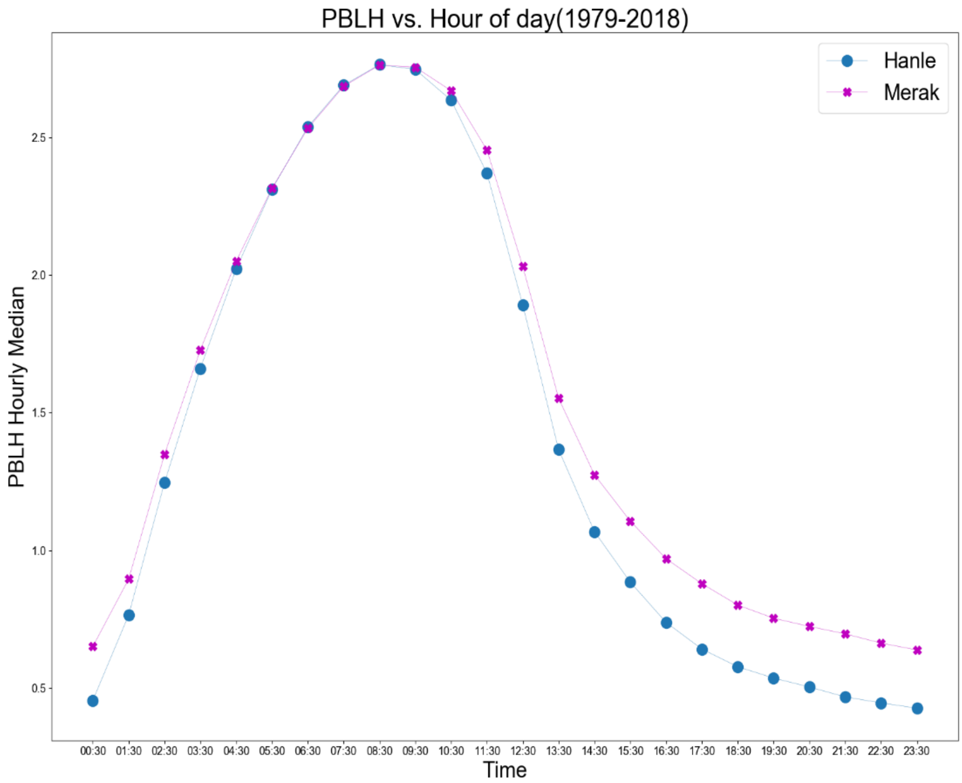

Fig 3.1 PBLH hourly

Fig 3.1 has a hourly median plot of PBLH (in km) of Hanle and Merak obtained from MERRA for the year 1980-2018. We can observe the increase in height of the PBLH during daytime due to heat that causes hot air to rise up. The higher the height of PBLH the higher the turbulence profiles, there is a sharp rise at sunrise and a slow decline at sunset as can be seen. Further discussion is provided in [1, 2].

Fig 3.2 is a similar Fig on the monthly median against PBLH (in km) for Hanle and Merak from MERRA data from 1980-2018. The monthly trend shows a peak during summer months and a sharp dip mid-august due to Monsoon, and then a slow rise and again a sharp dip at winter.

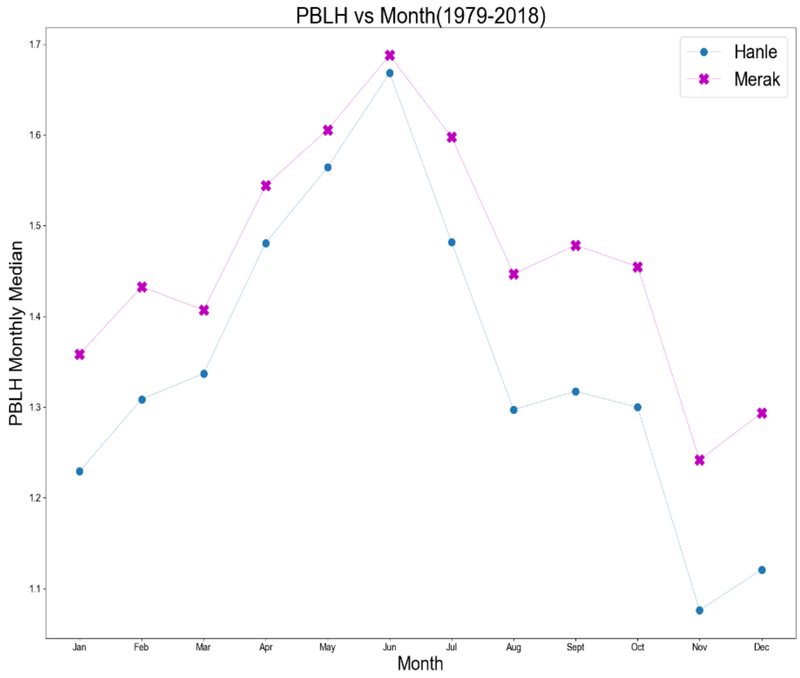

Fig 3.2 PBLH Monthly

Fig 3.3 shows us the correlation between ERA5 and MERRA data. A recent study [16] has compared different reanalysis data with in-situ observations and found the MERRA data to have a higher correlation than ERA5 and ERA-Interim. This could be because of better assimilation techniques and methods used in MERRA as described briefly in [2].

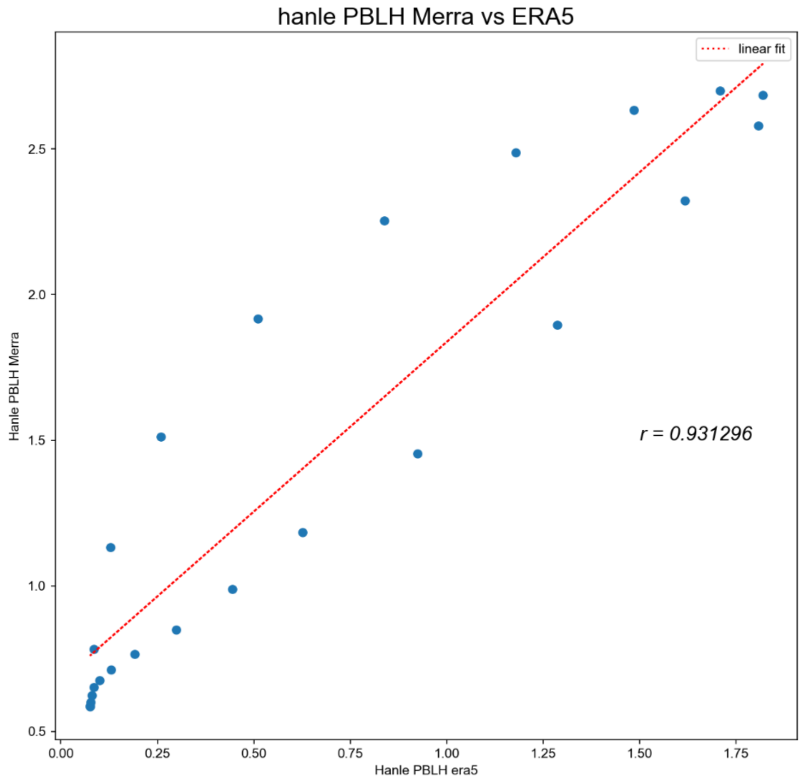

Fig 3.3 PBLH ERA5 MERRA Correlation

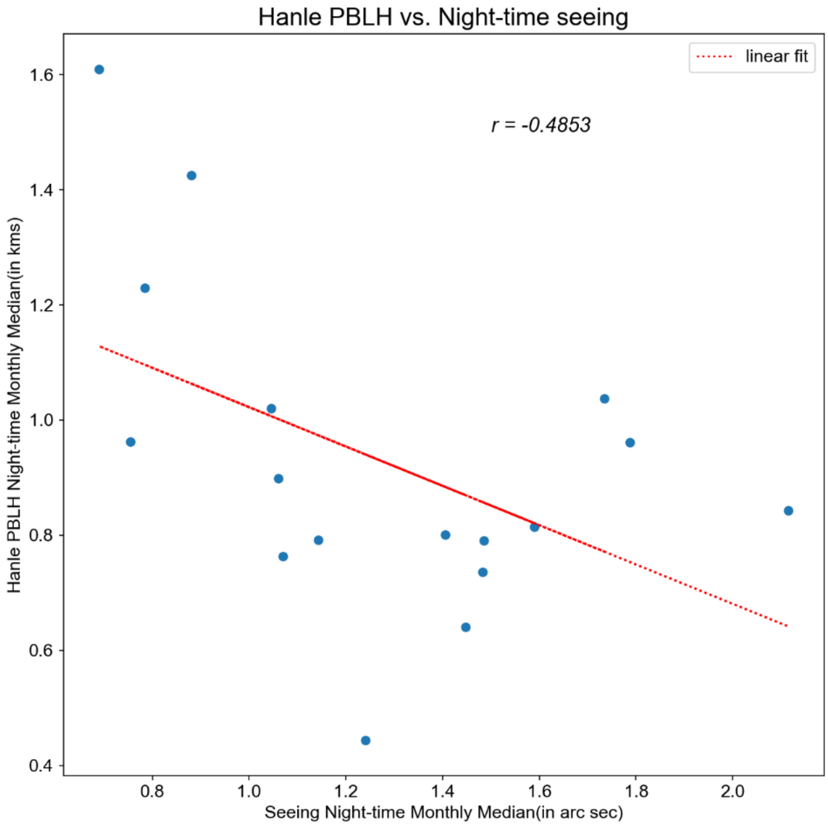

Fig 3.4 Hanle PBLH Seeing Correlation

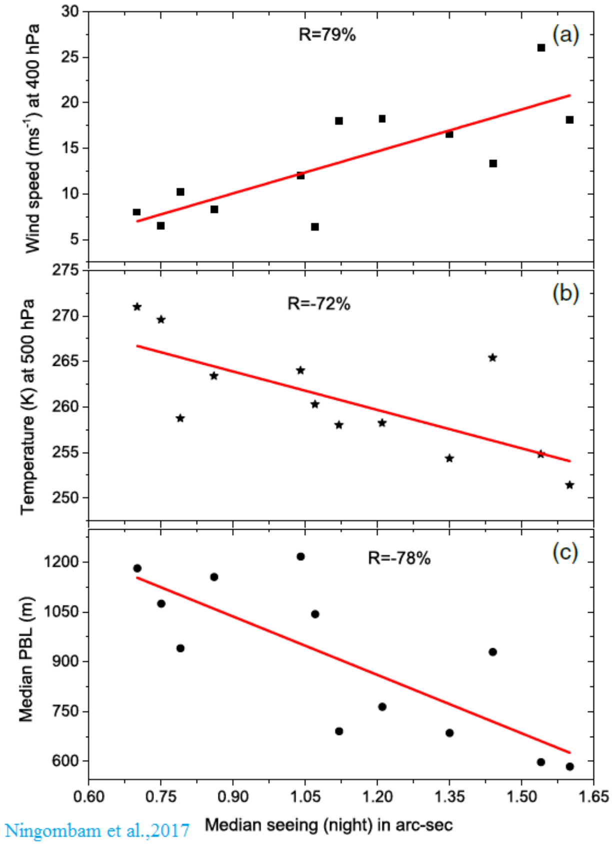

Fig 3.5 [2] – Seeing correlation with other parameters

Fig 3.4 draws a plot between monthly median PBLH of Hanle(in km) vs. monthly median night time seeing (FWHM in arc-sec) for the year 2011-2013. Data was time-matched and then the medians were taken. This has been compared with a more prolific plot from [2] that draws out the correlation between seeing and PBL, Temperature and Wind gust in Fig 3.5. Though the correlation obtained in 3.4 is lesser than 3.5, with the data for two years only the trend is clearly observed. This shows us that PBLH and Seeing are negatively correlated.

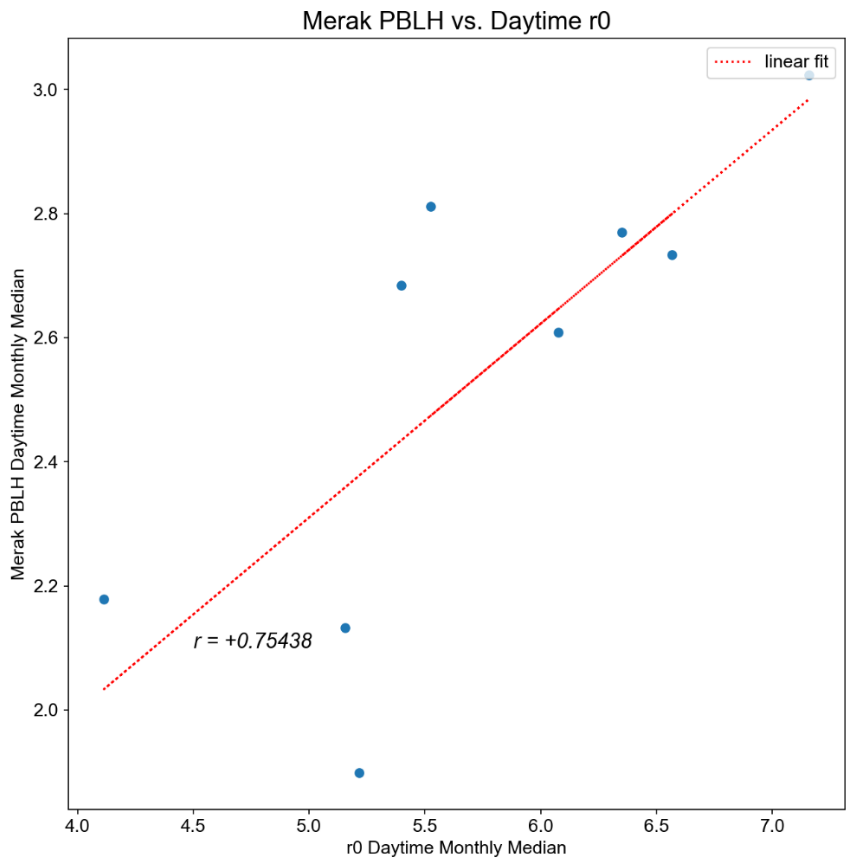

Fig 3.6 Merak PBLH r0 Correlation

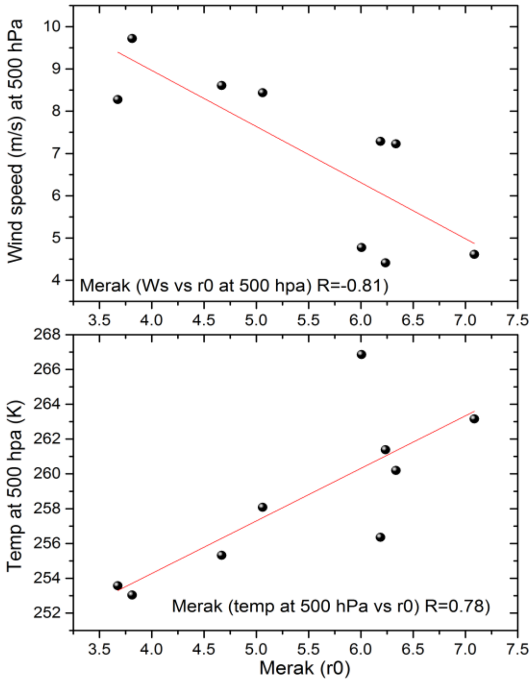

Fig 3.7 [2] – Merak r0 correlation with other parameters

Fig 3.6 has the monthly PBLH median values of Merak plotted against monthly r0 values for Merak from 2008 October to 2009 June. Seeing (FWHM) and r0 are inversely related as $s=k\ /r0$ [11]. Hence we see a direct correlation between PBLH and r0 values, and the correlation is significant. This can be compared to correlation of Temperature and Wind gust with r0 of Merak, and we get results as expected.

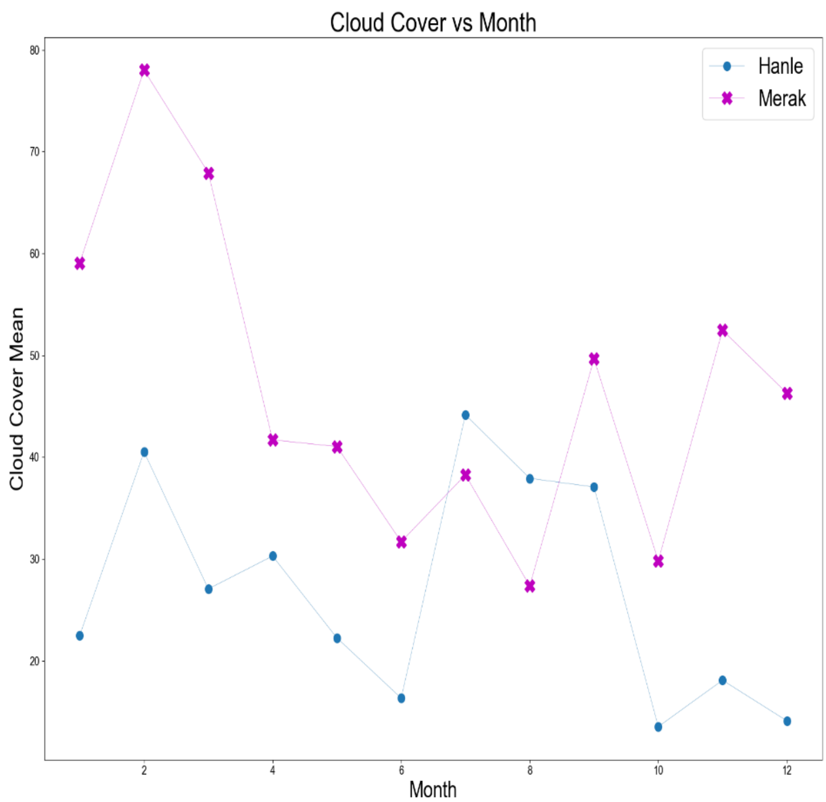

Fig 3.8 Cloud Cover Monthly

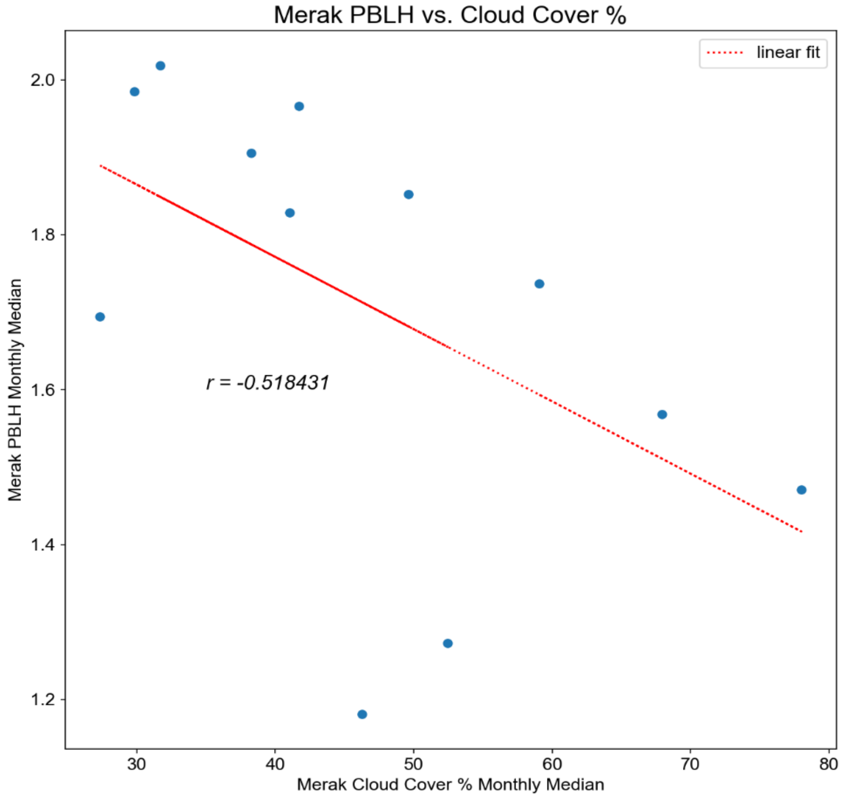

Fig 3.9 Merak PBLH Cloud Cover correlation

Fig 3.8 describes the monthly cloud cover percentage over Hanle and Merak during 2018, obtained from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) – Aqua and Terra. This can be compared with Fig 3.2 PBLH Monthly.

Fig 3.9 plots the monthly PBLH of Merak in 2018, with cloud cover % and draws a negative correlation showing us that higher the cloud cover, lower the height of PBL, and higher the blurring caused due to astronomical seeing also. Hence a low PBLH not only implies a higher seeing, but also gives a higher probability of cloud cover being there, making astronomical observations harder.

RESULT

The project aimed at majorly extracting PBLH data from reanalysis dataset, and evaluating it with other parameters. Once extraction was done, the analysis showed us the relation between PBLH and other atmospheric parameters that are influenced by it like Atmospheric Seeing and Cloud Cover.

A negative correlation of PBLH and Seeing shows us that higher the PBLH values lesser the Seeing and better the site for astronomical observations. A similar relation applies to PBLH and cloud cover as well. Having seen the values of PBLH, Seeing and cloud cover over Hanle and Merak for certain times of the year, it can be inferred the Hanle is a slightly better candidate for astronomical observations as compared to Merak though the difference is not significant.

The derivation of PBLH from CALIOP observations could not be completed due to time factor and technical issues. The groundwork of extracting data has been laid, and it can be explored in the future.

REFERENCES

1. von Engeln, A., & Teixeira, J. (2013). A Planetary Boundary Layer Height Climatology Derived from ECMWF Reanalysis Data. Journal Of Climate, 26(17), 6575-6590. doi: 10.1175/jcli-d-12-00385.1
2. Ningombam, S., Kathiravan, S., Parihar, P., L. Larson, E., Mohanan, S., & Angchuk, D. et al. (2017). Astronomical site survey report on dust measurement, wind profile, optical turbulence, and their correlation with seeing over IAO-Hanle. Experimental Astronomy, 43(2), 145-165. doi: 10.1007/s10686-017-9525-6
3. Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P. , Poli, P. , Kobayashi, S. , Andrae, U. , Balmaseda, M. A., Balsamo, G. , Bauer, P. , Bechtold, P. , Beljaars, A. C., van de Berg, L. , Bidlot, J. , Bormann, N. , Delsol, C. , Dragani, R. , Fuentes, M. , Geer, A. J., Haimberger, L. , Healy, S. B., Hersbach, H. , Hólm, E. V., Isaksen, L. , Kållberg, P. , Köhler, M. , Matricardi, M. , McNally, A. P., Monge-Sanz, B. M., Morcrette, J. , Park, B. , Peubey, C. , de Rosnay, P. , Tavolato, C. , Thépaut, J. and Vitart, F. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q.J.R. Meteorol. Soc., 137: 553-597. doi:10.1002/qj.828
4. (Gelaro, R., McCarty, W., Suárez, M., Todling, R., Molod, A., & Takacs, L. et al. (2017). The Modern-Era Retrospective Analysis for esearch and Applications, Version 2 (MERRA-2). *Journal Of Climate*, *30*(14), 5419-5454. doi: 10.1175/jcli-d-16-0758.1
5. https://www.unidata.ucar.edu/software/netcdf/docs/netcdf_introduction.html
6. https://support.hdfgroup.org/products/hdf4/HDF-FAQ.html#1
7. Liu, J., Huang, J., Chen, B., Zhou, T., Yan, H., & Jin, H. et al. (2015). Comparisons of PBL heights derived from CALIPSO and ECMWF reanalysis data over China. Journal Of Quantitative Spectroscopy And Radiative Transfer, 153, 102-112. doi: 10.1016/j.jqsrt.2014.10.011
8. Melfi, S., Spinhirne, J., Chou, S., & Palm, S. (1985). Lidar Observations of Vertically Organized Convection in the Planetary Boundary Layer over the Ocean. Journal Of Climate And Applied Meteorology, 24(8), 806-821. doi: 10.1175/1520-0450(1985)024<0806:loovoc>2.0.co;2
9. Jordan, J., Smith, D., Dawson, S., & Dawson, A. (2010). Holocene relative sea-level changes in Harris, Outer Hebrides, Scotland, UK. Journal Of Quaternary Science, 25(2), 115-134. doi: 10.1002/jqs.1281
10. https://www-calipso.larc.nasa.gov/resources/project_documentation.php
11. http://www.eso.org/gen-fac/pubs/astclim/papers/lz-thesis/node11.html
12. Martinez, P., Kolb, J., Sarazin, M., & Tokovinin, A. (2010). On the difference between seeing and image quality: when the turbulence outer scale enters the game. *The Messenger*, *141*, 5-8
13. https://docs.astropy.org/en/stable/io/fits/
14. https://unidata.github.io/netcdf4-python/netCDF4/index.html
15. http://xarray.pydata.org/en/stable/
16. Comparative Evaluation of the Third-Generation Reanalysis Data for Wind Resource Assessment of the Southwestern Offshore in South Korea. (2018). Atmosphere, 9(2), 73. doi: 10.3390/atmos9020073