Hybrid Deep Learning and Numerical Weather Prediction Enabled Crop Predictions

Dissertation

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology

in

Data Science

by

Souvik Roy

(Registration No. 20-27-08)

Under the Guidance of

Dr. Bharath Ramkrishna
Assistant Professor
Department Of Applied Mathematics, DIAT (DU)
Pune
(Internal Supervisor)

Mr. Manmeet Singh
Scientist 'C'
IITM
Pune
(External Supervisor)

Department of Applied Mathematics

Defence Institute of Advanced Technology

(Deemed to be University)
Girinagar, Pune 411025

April 2022

Dedication

To each and everyone who supported and motivated me

Approval Sheet

Thesis Titled : Hybrid Deep Learning and Numerical Weather Prediction Enabled Crop Predictions

Undertaken By : **Souvik Roy** (Reg. No. 20-27-08)

is approved for the award of the degree of Master of Technology in Data Science programme.

(Internal Supervisor)

(External Supervisor if any)

(Member from Allied Department)

(External Subject Expert)

Date :

Place :DIAT, Pune

(Chairman)

Certificate of Course Work

This is to certify that **Souvik Roy** has successfully completed the necessary course work as required for award of M. Tech in Data Science. The details of the course work are as follows:

Sr. No.	Course No.	Course Name	Credits
1	AM 603D	Computer Oriented Optimization Methods	4
2	AM 604D	Statistical Computing for Data Science	4
3	AM 606D	Scientific Computing	4
4	AM 607D	Data Structures and Algorithms with C	4
5	CE 615A	Intelligent Algorithms	4
6	CE 696A	Artificial Intelligence and DSS	4
7	AM 623D	Machine Learning	4
8	AM 624D	Data Science: Tools and Techniques	4
9	CE 694	Big Data Analysis and Algorithms	4
10	AM 628D	Computational Number Theory and Cryptography	4
11	CE 631	Deep Learning	4
12	CE 632	Computer Vision	4
13	AM 609D	Data Science: Industrial Perspectives	4
14	AM 651D	M.Tech Dissertation - 1	14
15	AM 652D	M.Tech Dissertation - 2	14

DIAT Pune — Controller of Examinations

Dated :

Certificate

This is to certify that the dissertation titled "**Hybrid Deep Learning and Numerical Weather Prediction Enabled Crop Predictions**" submitted by **Souvik Roy** (Reg. No. 20-27-08) in partial fulfillment of the requirement for the award of degree of Master of Technology in Data Science.

Dr. Bharath Ramkrishna,
Internal Supervisor
Assistant Professor
Department of Applied Mathematics,
DIAT, Girinagar, Pune-411025
Date:
Place:DIAT,Pune

Mr. Manmeet Singh,
External Supervisor
Scientist 'C'
IITM, Pashan, Pune
Date:
Place:IITM,Pune

COUNTERSIGNED

Dr. SVSSNVG Krishna Murthy,
HOD,Department of Applied Mathematics,
Defence Institute of Advanced Technology (DIAT),
Girinagar, Pune

Declaration

This is to certify that the dissertation report titled "**Hybrid Deep Learning and Numerical Weather Prediction Enabled Crop Predictions**" comprises of original work and due citations have been made in the text to all other material used. The dissertation has not been submitted in any form for another degree or diploma at any other institute/university.

Date:
Place:

Souvik Roy,
Reg.No. 20-27-08,
Department of Applied Mathematics,
DIAT, Girinagar, Pune

Dr. Bharath Ramkrishna,
Internal Supervisor
Assistant Professor
Department of Applied Mathematics,
DIAT, Girinagar, Pune-411025
Date:
Place:DIAT,Pune

Mr. Manmeet Singh,
External Supervisor
Scientist 'C'
IITM, Pashan, Pune
Date:
Place:IITM,Pune

Declaration For Plagiarism Check

This is to certify that M Tech thesis entitled "**Hybrid Deep Learning and Numerical Weather Prediction Enabled Crop Predictions**" submitted by "**Souvik Roy** " bearing Registration No "**20-27-08** " under the supervision of "**Mr. Manmeet Singh**" , Scientist 'C', Indian Institute of Tropical Meteorology (IITM), and "**Dr. Ramkrishna Bharath**" in the Department of Applied Mathematics is the Original research work done by the candidate.

We have read the provision of "**DIAT (DU) Plagiarism Policy** " and it is certified that all the conditions prescribed in the aforesaid policy are compiled with in respect of above mentioned M Tech thesis.

The thesis has been checked for the Plagiarism and the report is submitted along with the thesis for further processing.

- ✓ Originality content (including the contents from his own publications): 95%
- ✓ Similarity Reproduction of the content from other sources: 5%

We are aware that any issue related to plagiarism in future will have to be addressed by the candidate and the Supervisor (s) concerned.

Name of the Candidate:
Signature:
Date:

Name of Supervisor (s):
1. 2.

Signature:

1. 2.

Date:

Acknowledgement

I would like to express my sincere gratitude to **Dr. Bharath Ramkrishna**, Assistant Professor, Department of Applied Mathematics, for motivating me all through this work. I would like to thank for his timely guidance and support.

I would like to express my sincere gratitude to **Mr. Manmeet Singh**, Scientist 'C', Indian Institute of Tropical Meteorology, Pune for providing me the opportunity to take up this project and more importantly, his precious guidance and motivation, which helped me to conquer the difficulties in research and development of this project.

I also like to express my sincere gratitude to **Dr. Rajib Chattopadhyay**, Scientist 'E', Indian Institute of Tropical Meteorology, Pune for his valuable support and encouragement.

Above all, I offer my sincere thanks and gratitude to the merciful God Almighty for the providence and grace, He has given throughout this attempt.

Souvik Roy
M.Tech in Data Science

Abstract

Agriculture has always been an important sector for every person who is involved directly or indirectly in this sector. The traditional methods are still widely used in agriculture but as the advancement of technology is increasing day by day so the effectiveness of products produced by agriculture such as crops, fruits, and many more products are getting enhanced. The help of modern technologies such as Artificial Intelligence, Remote Sensing, and various other technologies are helping the agriculture sector to boom faster which is a great sign for people involved in this sector.

The main goal of this project is to use one of the advancements which are using Deep Learning model with the help of UNET architecture which will be used to forecast the yield of vegetation or crops for which in this project is NDVI (Normalized Difference Vegetation Index) and by the help of Meteorological methods to convert the data of vegetation or crops into a suitable format to make it useful for the Deep Learning model to proper forecast the yields of vegetation.

In this project, seasonal forecasting of vegetation is done where monsoon months are crucial for farmers for predicting the vegetation of future monsoons that will help the farmers to put crops as per the prediction and also will help reduce the wastage of crops. There are several case studies like forest fires, droughts, etc., which are taken in the project to evaluate the deep learning model which will give an idea of whether the model is giving a forecast which is of correct significance when it comes to predicting vegetation in different parts of the world.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The vegetation always plays a vital role in the enhancement of ecosystems but in recent times due to heavy involvement of human interaction with nature such as exploitation of forests, illegal mining and many more have played a huge role in the declining of the vegetation and the related resources like food production, crop yielding, etc. Normalized Difference Vegetation Index (NDVI) is extensively used in most of the research works connected to vegetation. Several methods had been used in the last few years related to NDVI research works such as ANN (Artificial Neural Networks), MLP (Multiple Layer Perceptron), and many more deep learning architectures at the beginning phase of these research works but have some limitations which is the reason extensive studies are made to use other deep learning models for NDVI. The main purpose of the project is that it will use a newly developed concept or model named Cubed Sphere [2]. The data is for the whole world and the format of the NDVI will be converted into the Cubed Sphere format and then these files will be further applied to the Deep Learning model which is UNET architecture [1]. After the model is applied then evaluation metrics will be used to test this newly made concept of Cubed Sphere is working for NDVI data.

1.1 Normalized Difference Vegetation Index

Normalized Difference Vegetation Index (NDVI) is a vegetation index that shows the intensity of vegetation around the world. Mostly NDVI lies in the range -1 to 1 in which the negative values of NDVI closer to -1 represent water, cloud, or snow. The values closer to 0 which can range from -0.1 to 0.1 depicts areas of barren land, land containing rock, sand, etc. The values ranging from 0.2 to 0.4 mostly show shrubs and grasslands and at last, the values which are closer to 1 show temperate and tropical rainforests. Mostly NDVI for an area containing dense vegetation the canopies will tend to have values in the range closer to 0.3 to 0.8.

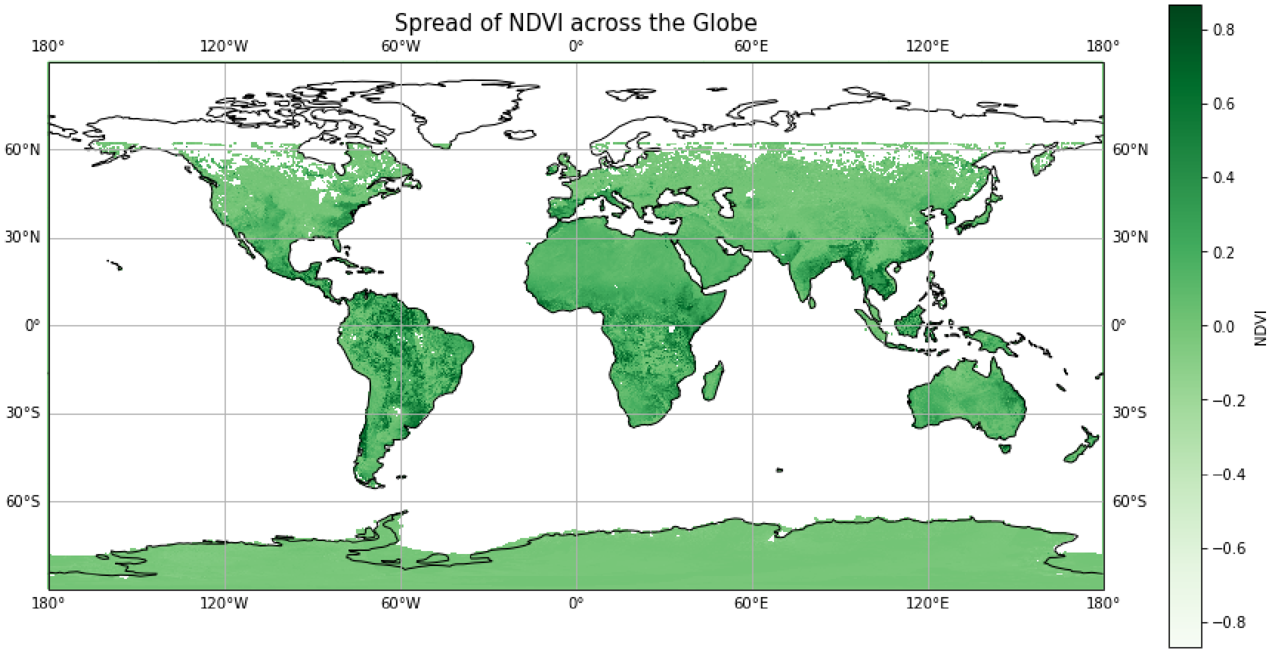


Figure 1.1: Plot of NDVI spread over the entire globe

NDVI generally uses the NIR (Near-infrared) and red channels in the formula. The vegetation which is healthily represented by chlorophyll reflects NIR more and the green light compared to the other wavelengths present. But it mostly absorbs more of the red and blue light from the atmosphere. This is the reason our eyes see green when seeing vegetation. The resolution of the NDVI data taken is 3600*7200.

S.No	Target Variables	Name(Units)	Min-Max Value
1	NDVI	Normalized Difference Vegetation Index(1)	-0.0999 - 0.9307

Table 1.1: Details of NDVI Variable

```
<xarray.Dataset>
Dimensions:    (latitude: 3600, longitude: 7200, ncrs: 1, nv: 2, time: 1)
Coordinates:
  * latitude   (latitude) float32 89.97 89.93 89.88 ... -89.88 -89.93 -89.97
  * longitude  (longitude) float32 -180.0 -179.9 -179.9 ... 179.9 179.9 180.0
  * time       (time) datetime64[ns] 2018-01-01
Dimensions without coordinates: ncrs, nv
Data variables:
    crs        (ncrs) int16 ...
    lat_bnds   (latitude, nv) float32 ...
    lon_bnds   (longitude, nv) float32 ...
    NDVI       (time, latitude, longitude) float32 ...
    TIMEOFDAY  (time, latitude, longitude) datetime64[ns] ...
    QA         (time, latitude, longitude) int16 ...
Attributes: (12/48)
    title:                              Normalized Difference Vegetation ...
    institution:                        NASA/GSFC/SED/ESD/HBSL/TIS/MODIS-...
    Conventions:                        CF-1.6, ACDD-1.3
    standard_name_vocabulary:           CF Standard Name Table (v25, 05 J...
    naming_authority:                   gov.noaa.ncei
    license:                            See the Use Agreement for this CD...
    ...                                 ...
    PercentValidDaytimeData:            32.81
    PercentValidDaytimeLand:            32.81
    PercentValidClearDaytimeLand:       3.80
    PercentValidDaytimeLandInCloudShadow:   0.96
    PercentValidClearDaytimeWater:      0.00
    PercentValidDaytimeWaterInCloudShadow:  0.00
```

Figure 1.2: Details of NETCDF file containing NDVI variable

The formula generates results between -1 and 1. If there are low values which mean low reflectance in the red channel and high values or the high reflectance in the Near-infrared channel, this will help us get a high value of NDVI and vice versa.

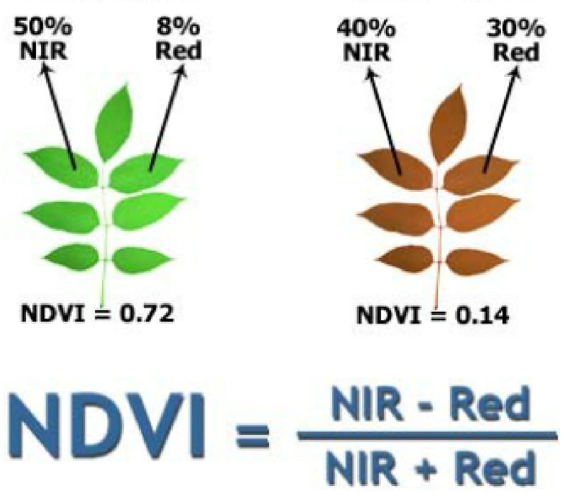


Figure 1.3: Formula for calculating NDVI. Note. The figure is taken from (https://ece.montana.edu/seniordesign/archive/SP15/OpticalWeedMapping/ndvi.html).

NDVI helps to see the intensity and also to measure whether vegetation is healthy or not. If there is a need to see the change in vegetation over time then the best way is to perform the atmospheric correction.

1.2 U-Net

The U-Net architecture was designed by Olaf Ronneberger et.al. [13], which is generally used for segmenting the images of BioMedical. It is a type of CNN(Convolutional Neural Network) that generally does tasks on the classification of images. The input is an image and the output is only one label, but in the case of biomedical, there is a requirement of distinguishing the probability of having a disease and also localizing the areas where there is a chance of abnormality. U-Net helps to solve this problem and it does localization by classifying each pixel so that the input and output share a similar size.

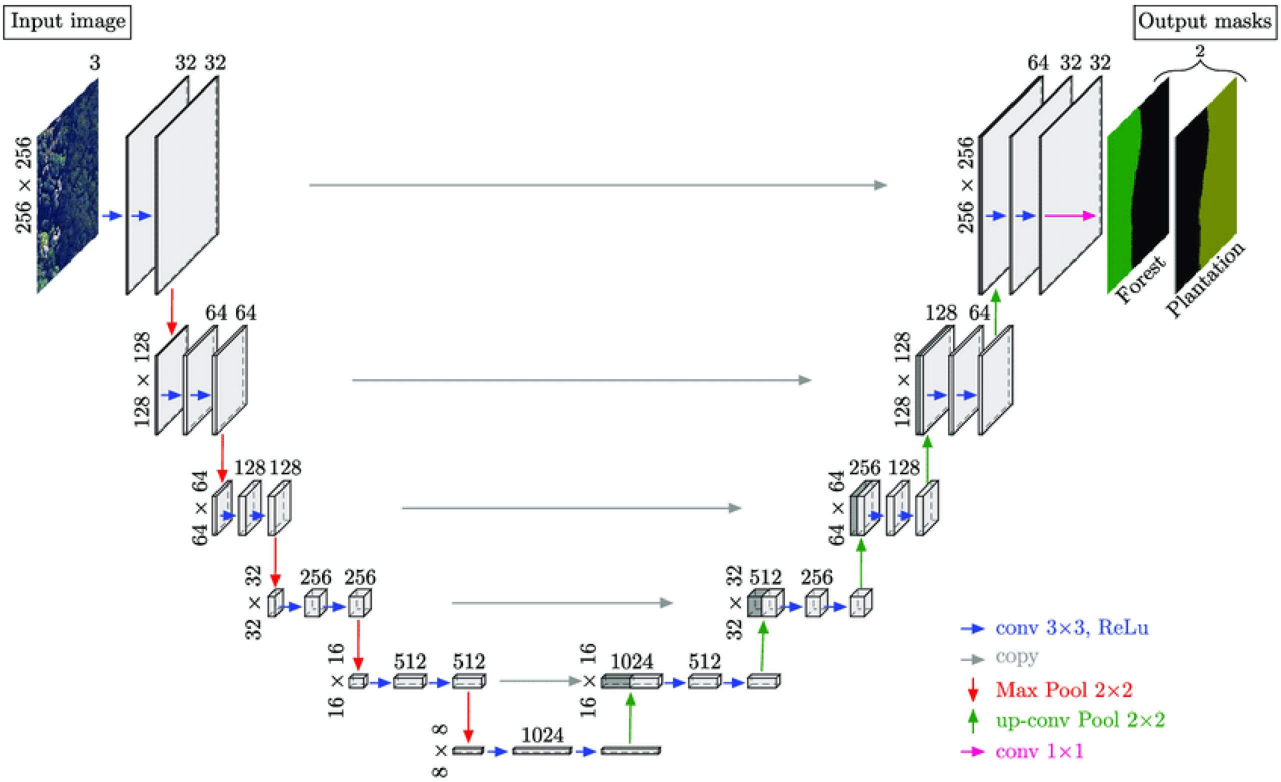


Figure 1.4: Diagram for describing U-Net architecture. Note. The diagram is taken from [13](Fig. 1).

U-Net architecture has a “U” shape. The architecture is symmetric and has two parts which are the left part and the right part. The left part is called the contracting path, which generally constitutes the Convolutional process whereas the right part is the expansive path, which generally constitutes transposed 2d Convolutional layers which can be depicted as an Upsampling technique.

The contracting path consists of:-

convlayer1 - convlayer2 - maxpooling - dropout(optional)

The process generally shows two Convolutional layers and the number of channels generally changes from 1 to 64 which will increase the depth of an image with the help of the Convolution process. The red arrow that is pointing down is the max-pooling layer which helps to halve down the size of an image in the implementation there is a requirement of using padding=”same” and also there is a chance of some padding issues. This process is repeated three times after the first process and then it reaches the bottommost which makes the fact that 2 Convolutional layers get built and there is no max-pooling yet which makes the image resized to 28x28x1024.

In the expansive path, upsizing of an image is done from its original size. The formula follows:

conv2dtranspose - concatenate - convlayer1 - convlayer2

Transposed convolution is a way to expand the image size which can be called upsampling. Some padding is done on the image which is followed by the convolutional operation. Then the image is upsized from 28x28x1024 to 56x56x512 with the corresponding image concatenated from the contracting path thus making an image of size 56x56x1024. This helps to get a more precise prediction. Two other convolution layers are added in lines 4 and 5 and repeated three times which makes us reach the uppermost part of the architecture.

The last step is reshaping the image as per the prediction requirement which is a Convolution layer with one filter of size 1x1 and there is no dense layer in the whole architecture.

Chapter 2

Literature Review

Manmeet Singh and Bipin Kumar et.al. [1], suggested that the understanding of the relationship of other variables with the formation of precipitation is very beneficial. The UNET architecture is used to learn the global data-driven models for precipitation and the results showed that the residual learning-based UNET architecture can explain the relationships for the precipitation variable and it can be helpful in the dynamical operational models which are compared with the results that further improve the forecasting of precipitation.

Jonathan A.Weyn and Dale R.Durran et.al. [2], proposed further improvements in CNN (Convolutional Neural Networks), created an offline volume conservative mapping to the newly formed concept called Cubed Sphere, and also minimized the loss function that was occurring in the various steps of the prediction sequence. The model predicts weather forecasts which are stable and does so for several weeks and also uses few input atmospheric state variables.

Liuqing Ji and Ke Fan et.al. [3] come up with an idea for the seasonal prediction of NDVI for the spring months which are from April to June over Eurasia. The approach is based on year-to-year increments and the prediction model used in the project is SVD (Singular Value

Decomposition). The predictors that are used in the prediction model are:- surface air temperature (SAT-CFS scheme), sea surface temperature (SSTP scheme), sea-ice cover (SICBS scheme), and winter North Atlantic Oscillation (NAO scheme). A statistical scheme is also included in the project included with the four predictors and to evaluate cross-validation technique is used with an approach of one year out and finally getting the hybrid scheme which is the prediction skill at the best.

Mathyam Prabhakar and K.A. Gopinath et.al. [4], demonstrated that assessment of crop damage in an area due to hailstorms is a tough task but provides relief to the farmers. The project shows the usage of satellite data being multispectral, identifying the areas of hailstreaks and the difference of NDVI from pre to post hailstorm events. With the change in NDVI, they classified different crops and calculated Kappa Coefficient for evaluation.

In the year 2012, J. V. Revadekar and Yogesh K. Tiwar et.al. [5], illustrated that NDVI data from 2000 to 2010 are taken to evaluate the variability being spatio-temporal over the Indian Region. The project shows that there is variability in NDVI which is linked with the factors like rainfall, temperature, etc. Over the Indian region the changes in rainfall, temperature, and also in different months like the monsoon months over the NDVI will help to do the seasonal prediction. The technique used here is the regression method.

Eric S. Kasischke and Nancy H. F. French et.al. [6], expressed the NDVI data from AVHRR which will be used to see the area contained by fire over Alaska. The technique is used to map the boundaries of fire in the early summer. It also depicts that AVHRR NDVI data sometimes does not capture the actual forest fire area over a region so it is suggested that the areal estimation can be improved.

Dyah R. Panuju and Bambang H. Trisasongko1n et.al. [7], indicated that forest fires are a problem in Indonesia the need for verifying the hot spots which is the need as an indication of forest fires. But these get restricted by the medium and high-resolution data. A time-series data is made for NDVI data to relate it with the forest fire accidents and to relate the changes associated with it.

M. E. Brown and D. J. Lary et.al. [8], showed that there is a struggle to create a vegetation dataset that is derived from various sensors which will be used for time series analysis.

For this shortcoming, ANN (Artificial Neural Network) is used in this project to map NDVI from AVHRR (Advanced Very High Resolution Radiometer) with MODIS (Moderate Resolution Imaging Spectroradiometer) to get the difference between these sensors. ANN is used to remove the atmospheric and the effects caused by the sensors on NDVI.

Habib Aziz Salim and Xiaoling Chen et.al. [9], illustrated the changes in the variability of NDVI annually which is taken for the period from 1982 to 1993. This project included a relationship between NDVI and rainfall in Sudan where NDVI values are taken for monitoring the drought.

The inferences that can be taken from the literature survey are that most of the test cases taken for evaluating the models are in particular areas or countries like Indonesia, Brazil, etc. Then most of the predictions or forecasting is done are either for a few days or months or a year depending on the case study of the particular country. The techniques used in the papers are common machine learning and deep learning techniques which works only for particular areas, specific case study, small data, etc. The data is mostly taken from either AVHRR or MODIS or both.

Chapter 3

Methodology

3.1 Description of Data Variables

There are input variables and the target variable which is used for the U-Net model. As the target variable is NDVI, there are four input variables which are Precipitation, Near-Surface Air Temperature, Monthly mean air temperature at 200hPa, and Monthly mean air temperature at 850hPa. The details of the input variables in tabular form are as follows:-

S.No	Input Variables	Name(Units)	Min-Max Value
1	pr	Precipitation(kg m-2 s-1)	0 - 0.002
2	tas	Near-Surface Air Temperature(K)	220 - 320
3	ta200	Monthly mean air temperature at 200hPa(K)	200 - 235
4	ta850	Monthly mean air temperature at 850hPa(K)	230 - 310

Table 3.1: Details of Input Variables

The graphs for representing the input variables are as follows:-

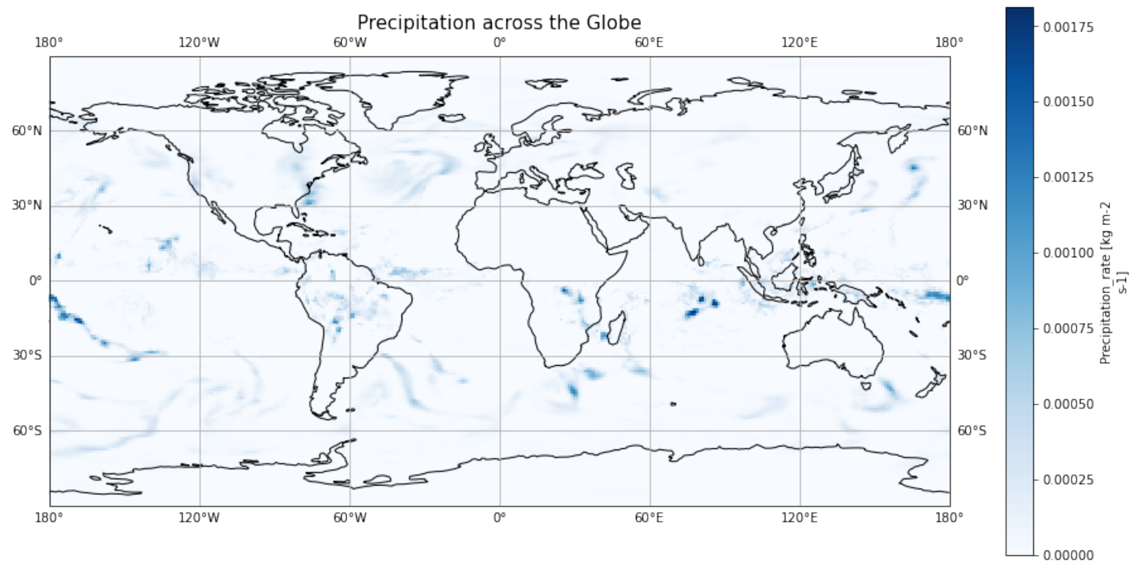


Figure 3.1: Plot for Precipitation Input Variable

The precipitation variable is a CFSv2 data in which the resolution is 190*384. The data is six hourly intervals starting from the years 1982 to 2010. The data shows the precipitation rate across the globe.

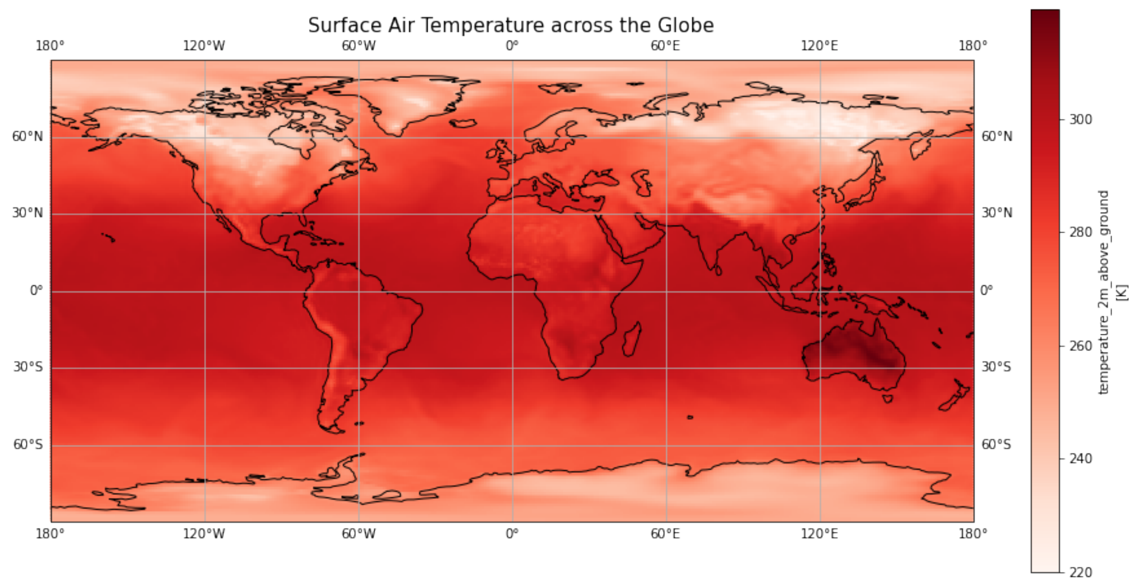


Figure 3.2: Plot for Surface Air Temperature Input Variable

The surface air temperature variable is a CFSv2 data in which the resolution is 190*384. The data is six hourly intervals starting from the years 1982 to 2010. The data gives an overview of the temperature levels for both land and ocean.

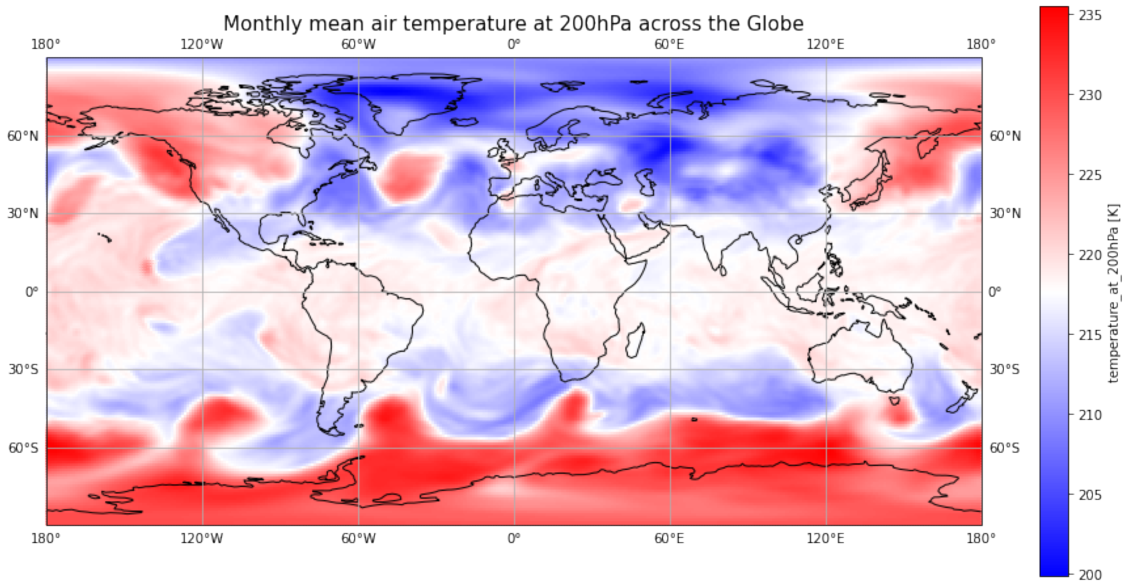


Figure 3.3: Plot for Monthly mean air temperature at 200hPa Input Variable

The air temperature variable is a CFSv2 data in which the resolution is 181*360. The data is six hourly intervals starting from the years 1982 to 2010. The data is measured at 200hPa showing variations in air temperature at this pressure across the globe.

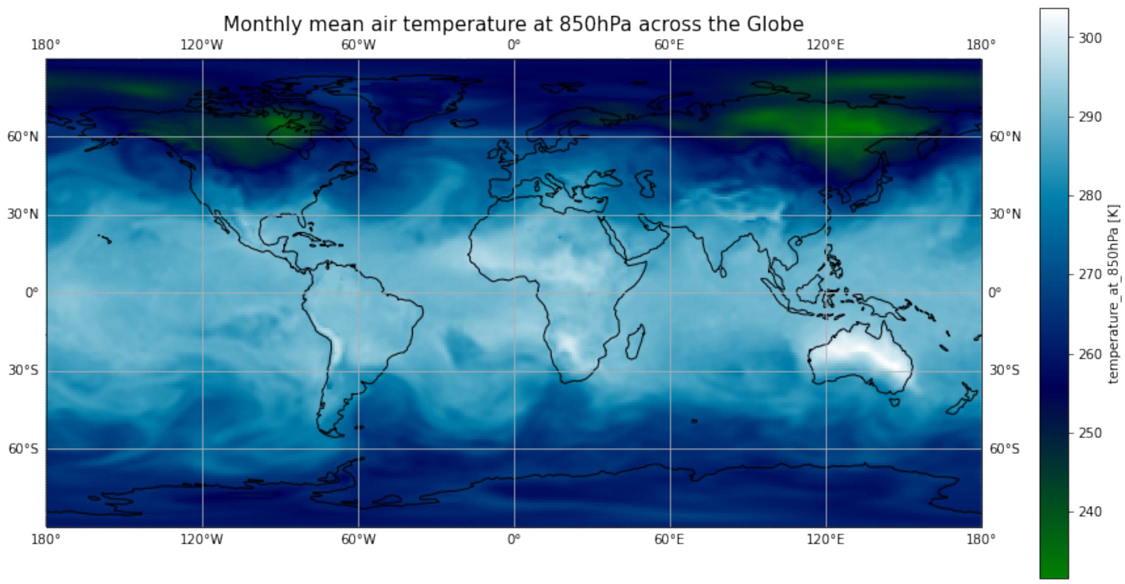


Figure 3.4: Plot for Monthly mean air temperature at 850hPa Input Variable

The air temperature variable is a CFSv2 data in which the resolution is 181*360. The data is six hourly intervals starting from the years 1982 to 2010. The data is measured at 850hPa showing variations in air temperature at this pressure across the globe.

The input variables are having data with a span of six hours which means the hours will be six, twelve, eighteen, etc in the data with the date which is from the years 1982 to 2010. The input variables are CFSV2 data variables.

3.2 Proposed Model

The standard representation of the model with a detailed explanation is mentioned below. The block diagram provides a clear explanation of each of the steps used for crop yield forecasting.

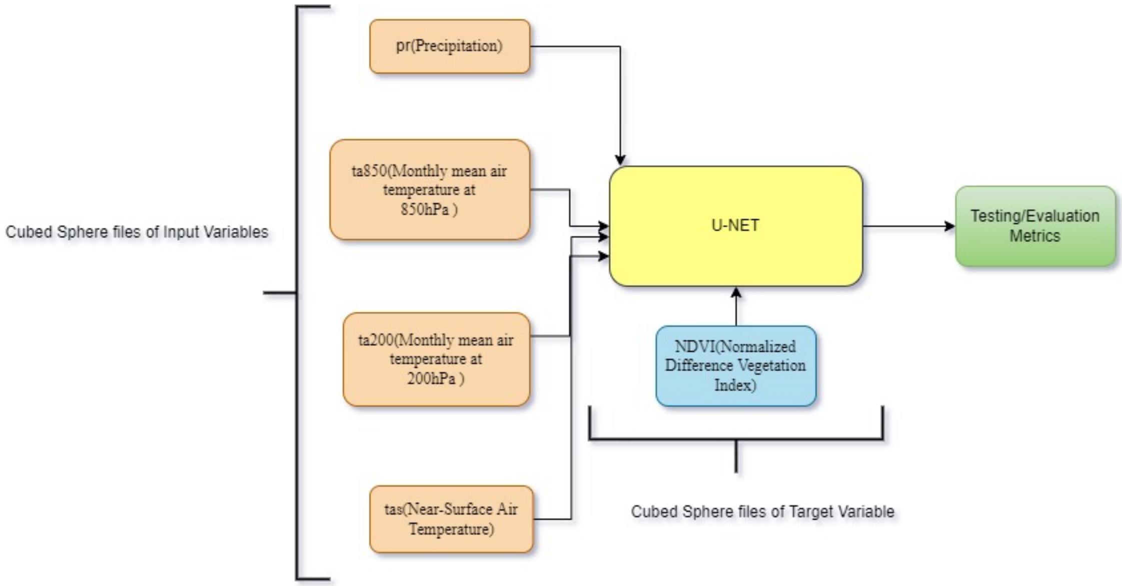


Figure 3.5: Methodology used for forecasting yield of crops

First, the input and target variables are preprocessed to make it in the form of a Cubed Sphere. After the Cubed Sphere is made then the next step is to code the U-Net model according to the Cubed Sphere method. Then the model is trained with the Cubed Sphere files of input and the target variable. Then the last step is to post-process the output files which generate from the training and then evaluation and results are generated with some case studies.

3.2.1 Proposed Algorithm

The algorithm is mostly based on making the Cubed Sphere with a schematic diagram given below:-

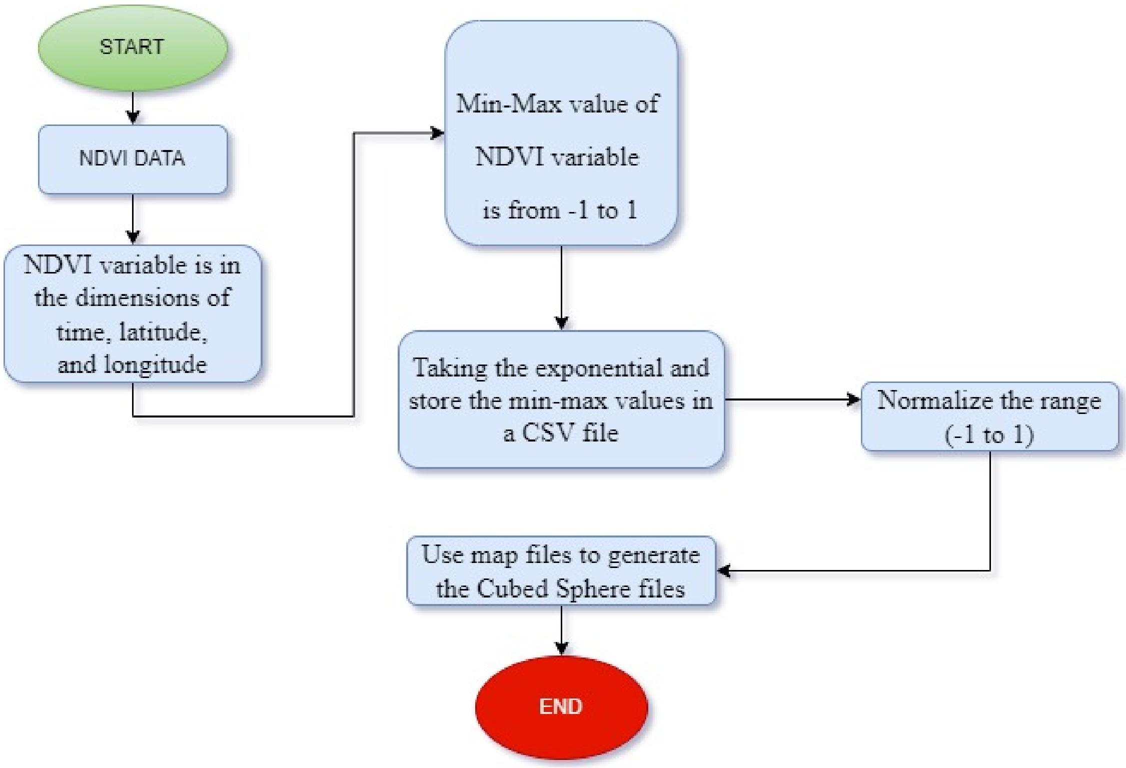


Figure 3.6: Systematic representation for generating Cubed Sphere

Before proceeding with the Cubed Sphere of the input variables, as the prediction is seasonal so to do that each input variable is made to form with an interval of 210 days and a mean is taken after that. This means that for every input variable each file will have 210 days time step with seven dates embedded in it. The next step is to use this method and create new files which will be used in the generation of the Cubed Sphere files. The files generated will be in the NETCDF format. Then the code is made regarding this and the U-Net model is made with it.

3.2.2 Algorithm flow

NDVI data is downloaded which is from the year 1981 to 2021 taken from [14]. The data downloaded is the daily data which means the NDVI data is available every day. There are two years 1981 and 2021 where there are missing data for several days so these two years have been discarded for better performance of the Deep Learning model. The next process is to recognize if a variable is in a tripolar grid or not so which in this case is no as the variables are in normal time latitude and longitude format which is required for further processing.

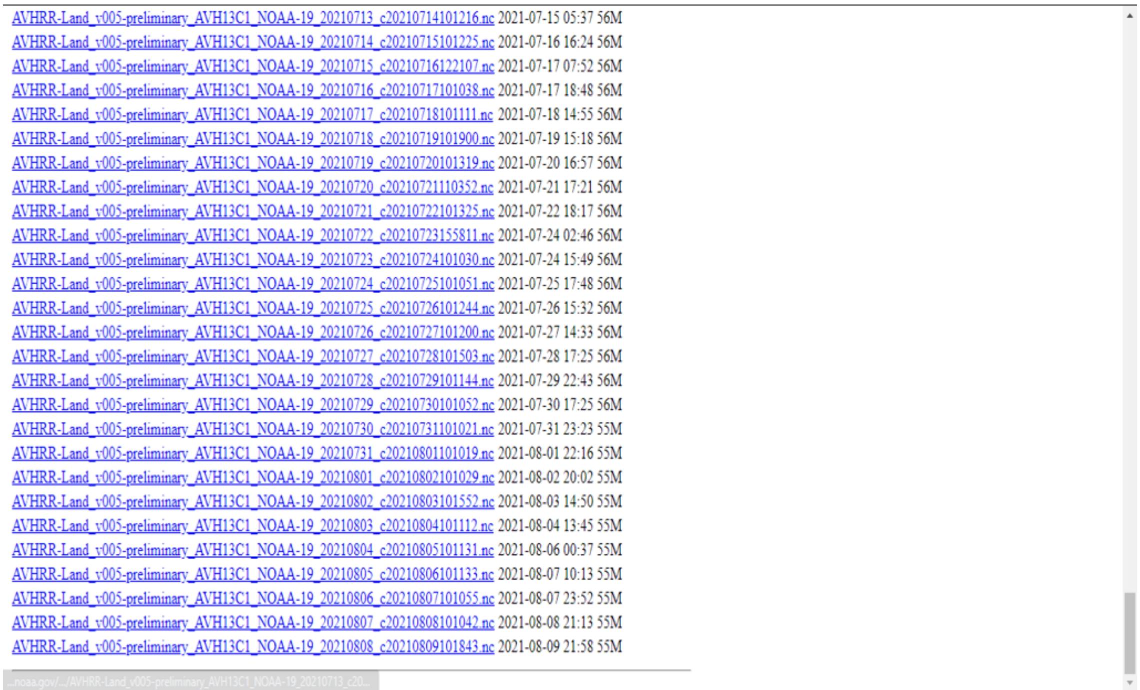
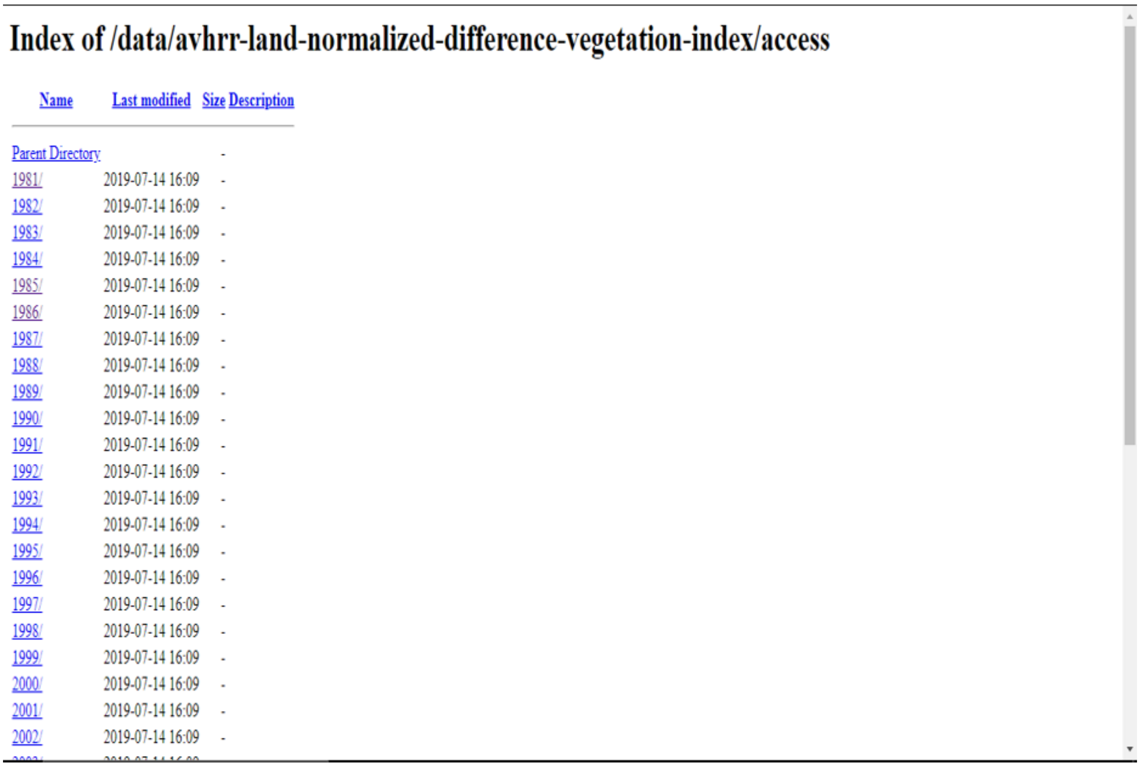

AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210713_c20210714101216.nc 2021-07-15 05:37 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210714_c20210715101225.nc 2021-07-16 16:24 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210715_c20210716122107.nc 2021-07-17 07:52 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210716_c20210717101038.nc 2021-07-17 18:48 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210717_c20210718101111.nc 2021-07-18 14:55 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210718_c20210719101900.nc 2021-07-19 15:18 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210719_c20210720101319.nc 2021-07-20 16:57 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210720_c20210721110352.nc 2021-07-21 17:21 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210721_c20210722101325.nc 2021-07-22 18:17 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210722_c20210723155811.nc 2021-07-24 02:46 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210723_c20210724101030.nc 2021-07-24 15:49 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210724_c20210725101051.nc 2021-07-25 17:48 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210725_c20210726101244.nc 2021-07-26 15:32 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210726_c20210727101200.nc 2021-07-27 14:33 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210727_c20210728101503.nc 2021-07-28 17:25 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210728_c20210729101144.nc 2021-07-29 22:43 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210729_c20210730101052.nc 2021-07-30 17:25 56M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210730_c20210731101021.nc 2021-07-31 23:23 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210731_c20210801101019.nc 2021-08-01 22:16 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210801_c20210802101029.nc 2021-08-02 20:02 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210802_c20210803101552.nc 2021-08-03 14:50 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210803_c20210804101112.nc 2021-08-04 13:45 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210804_c20210805101131.nc 2021-08-06 00:37 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210805_c20210806101133.nc 2021-08-07 10:13 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210806_c20210807101055.nc 2021-08-07 23:52 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210807_c20210808101042.nc 2021-08-08 21:13 55M
AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210808_c20210809101843.nc 2021-08-09 21:58 55M
...noaa.gov/.../AVHRR-Land_v005-preliminary_AVH13C1_NOAA-19_20210713_c20...

Figure 3.7: Daily Data for every year of AVHRR NDVI. Note. The data for AVHRR NDVI is taken from [14].



Index of /data/avhrr-land-normalized-difference-vegetation-index/access

Name	Last modified	Size	Description
Parent Directory		-	
1981/	2019-07-14 16:09	-	
1982/	2019-07-14 16:09	-	
1983/	2019-07-14 16:09	-	
1984/	2019-07-14 16:09	-	
1985/	2019-07-14 16:09	-	
1986/	2019-07-14 16:09	-	
1987/	2019-07-14 16:09	-	
1988/	2019-07-14 16:09	-	
1989/	2019-07-14 16:09	-	
1990/	2019-07-14 16:09	-	
1991/	2019-07-14 16:09	-	
1992/	2019-07-14 16:09	-	
1993/	2019-07-14 16:09	-	
1994/	2019-07-14 16:09	-	
1995/	2019-07-14 16:09	-	
1996/	2019-07-14 16:09	-	
1997/	2019-07-14 16:09	-	
1998/	2019-07-14 16:09	-	
1999/	2019-07-14 16:09	-	
2000/	2019-07-14 16:09	-	
2001/	2019-07-14 16:09	-	
2002/	2019-07-14 16:09	-	

Figure 3.8: AVHRR NDVI Data starting from 1982 to 2021

Since NDVI is a target variable and the input variables used for the NDVI variable are in normal time, latitude, and longitude format so the target variable also needs to match that format which is already available for NDVI. If there were tripolar grid format then CDO (Climate Data Operator) will be used to convert into general time, latitude, and longitude format and use it for further process. Then there is a need to merge the daily data into a single file for every year and finally need to mask the land as a batch of 10 years files area. Due to the mask of the land, it becomes essential to mask the land and oceanic variables needed for the exponential of the data array. The NDVI variable is a land variable.

The resulting data array's value range becomes so huge after taking exponential that there is a need to normalize to the range from 0 to 1 using min-max normalization and the maximum and minimum values are stored in a CSV file. After all these steps there is a reduction in the value range from -1 to 1. For NDVI the minimum and maximum values are already in the range -1 to 1 so there is no need for min-max normalization and can take the exponential directly. The last step is to use the Map file generation and Cubed Sphere generation that uses DLWP (Deep Learning Weather Prediction) which in the background uses TempestRemap [2].

The map file that is generated at first is stored in a location that can be reused to save computation time as it is quite expensive. After the map file is generated then the ApplyOfflineMap executable will generate a remapped temporary file that helps, in turn, to give results in the Cubed Sphere generation with CubeSphereRemap() module imported from DLWP.remap. The same process is done for the input variables to create Cubed Sphere files and the method is similar from the generation of CSV file and forwards to the other methods till the generation of Cubed Sphere files.

After this, the NDVI Cubed Sphere files are used in UNET architecture and evaluation metrics are used to check the performance of the model.

Chapter 4

Implementation

4.1 Data Preprocessing

The first step for converting the NDVI data into the Cubed Sphere data is merging the daily data into the yearly files which are shown in Fig. 4.1:-

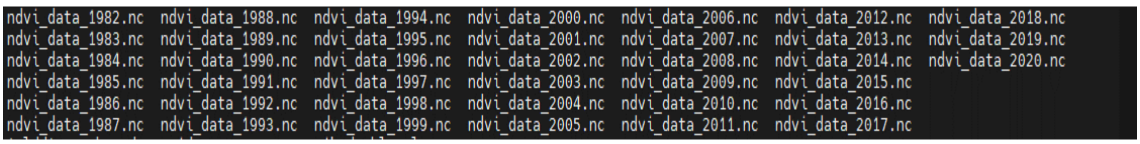


Figure 4.1: Generation of Yearly files

This step is mainly done because the issue of memory error comes into the picture which means that when we try to merge the daily data into yearly files with a resolution of 3600*7200, then memory error comes into the picture which is shown in Fig. 4.2:-

```
MemoryError: Unable to allocate 68.9 GiB for an array with shape (9253440000,) and data type timedelta64[ns]
```

Figure 4.2: Memory Error while merging the daily data into yearly files

So now the resolution is reduced to 720*1440 and then the second step is to convert these yearly files into a batch of 10-10 year files which is shown in Fig. 4.3 with details:-

```
<xarray.Dataset>
Dimensions:  (lat: 720, lon: 1440, time: 3651)
Coordinates:
  * time     (time) datetime64[ns] 2011-01-01 2011-01-02 ... 2020-12-31
  * lon      (lon) float32 0.125 0.375 0.625 0.875 ... 359.1 359.4 359.6 359.9
  * lat      (lat) float32 -89.88 -89.62 -89.38 -89.12 ... 89.38 89.62 89.88
Data variables:
    NDVI     (time, lat, lon) float64 ...
Attributes:
    CDI:          Climate Data Interface version 1.9.10 (https://mpimet.mpg.d...
    Conventions:  CF-1.6
    history:      Mon Sep 27 21:38:22 2021: cdo remapbil,/lus/dal/mtechstuden...
    CDO:          Climate Data Operators version 1.9.10 (https://mpimet.mpg.d...
```

Figure 4.3: Details of one of the 10-10 year batch files

Now the next step is to create a CSV file from the 10-10 year batch files where the CSV file will contain the minimum and maximum value of the NDVI variable which is shown in Fig. 4.4:-

```
var-lev,min,max
NDVI,0.9048374234292109,2.7177381362594346
```

Figure 4.4: Minimum and Maximum value of NDVI stored in CSV file

```
[0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.90483742342921
09, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.90483742342
92109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.90483742
34292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.9048374234292109, 0.90483
74234292109, 0.9048374234292109, 0.9048374234292109]

[2.7177381362594346, 2.7177381362594346, 2.7171947147628224, 2.7171947147628224, 2.7171947147628224, 2.7177381362594346, 2.71773813625943
46, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.71773813625
94346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.71773813
62594346, 2.7177381362594346, 2.7177381362594346, 2.7177381362594346, 2.7171947147625115, 2.7171947147625115, 2.7171947147625115, 2.71719
47147625115, 2.7171947147625115, 2.7177381362594346]

0.9048374234292109

2.7177381362594346

  var-lev  ...       max
0    NDVI  ...  2.717738
```

Figure 4.5: Extracting minimum and maximum value from every 10-10 year batch files

After this step, the generation of map files is created with the help of the concept used in [2]. The map files are done to create the Cubed Sphere files for the variable.

```
map_CS96_LL720x1440.nc  map_LL720x1440_CS96.nc
```

Figure 4.6: Map files generated for NDVI variable

Now the last step is to use these map files and the CSV file generated to create the Cubed Sphere files which are shown in Fig. 4.7:-

```
<xarray.Dataset>
Dimensions:  (face: 6, height: 96, time: 3651, width: 96)
Coordinates:
  * time     (time) datetime64[ns] 2011-01-01 2011-01-02 ... 2020-12-31
  * face     (face) int64 0 1 2 3 4 5
  * height   (height) int64 0 1 2 3 4 5 6 7 8 9 ... 87 88 89 90 91 92 93 94 95
  * width    (width) int64 0 1 2 3 4 5 6 7 8 9 ... 86 87 88 89 90 91 92 93 94 95
Data variables:
    lon      (face, height, width) float64 ...
    lat      (face, height, width) float64 ...
    NDVI     (time, face, height, width) float32 ...
```

Figure 4.7: Details of Cubed Sphere files for NDVI

```
remap_converted_grid_NDVI_1982_1991_normalized_CS96.nc  remap_converted_grid_NDVI_1997_2006_normalized_CS96.nc
remap_converted_grid_NDVI_1983_1992_normalized_CS96.nc  remap_converted_grid_NDVI_1998_2007_normalized_CS96.nc
remap_converted_grid_NDVI_1984_1993_normalized_CS96.nc  remap_converted_grid_NDVI_1999_2008_normalized_CS96.nc
remap_converted_grid_NDVI_1985_1994_normalized_CS96.nc  remap_converted_grid_NDVI_2000_2009_normalized_CS96.nc
remap_converted_grid_NDVI_1986_1995_normalized_CS96.nc  remap_converted_grid_NDVI_2001_2010_normalized_CS96.nc
remap_converted_grid_NDVI_1987_1996_normalized_CS96.nc  remap_converted_grid_NDVI_2002_2011_normalized_CS96.nc
remap_converted_grid_NDVI_1988_1997_normalized_CS96.nc  remap_converted_grid_NDVI_2003_2012_normalized_CS96.nc
remap_converted_grid_NDVI_1989_1998_normalized_CS96.nc  remap_converted_grid_NDVI_2004_2013_normalized_CS96.nc
remap_converted_grid_NDVI_1990_1999_normalized_CS96.nc  remap_converted_grid_NDVI_2005_2014_normalized_CS96.nc
remap_converted_grid_NDVI_1991_2000_normalized_CS96.nc  remap_converted_grid_NDVI_2006_2015_normalized_CS96.nc
remap_converted_grid_NDVI_1992_2001_normalized_CS96.nc  remap_converted_grid_NDVI_2007_2016_normalized_CS96.nc
remap_converted_grid_NDVI_1993_2002_normalized_CS96.nc  remap_converted_grid_NDVI_2008_2017_normalized_CS96.nc
remap_converted_grid_NDVI_1994_2003_normalized_CS96.nc  remap_converted_grid_NDVI_2009_2018_normalized_CS96.nc
remap_converted_grid_NDVI_1995_2004_normalized_CS96.nc  remap_converted_grid_NDVI_2010_2019_normalized_CS96.nc
remap_converted_grid_NDVI_1996_2005_normalized_CS96.nc  remap_converted_grid_NDVI_2011_2020_normalized_CS96.nc
```

Figure 4.8: Generated Cubed Sphere files for NDVI

So this is the last step for generating Cubed Sphere files for the NDVI variable. The same process is done for all the four input variables and at last, these files will be used in the U-Net model for training.

4.2 Training

The training process first consists of taking all the files one by one from the four input variables and the target variable. Then the next step is to make the NDVI variable in the form of 210 days of data and take the mean before training so that the dimensions of the Cubed Sphere files for all the input and target variables are the same. Then these files are used to start training with keeping an eye on training loss and validation loss. The training data is taken for years 1982-2000, validation data from 2001-2002 and test data from 2003-2010 for all the variables.

```
(4133, 6, 96, 96, 28) (4133, 6, 96, 96, 7) (584, 6, 96, 96, 28) (584, 6, 9
Epoch 1/1000000
517/517 [==============================] - 132s 205ms/step - loss: 0.0638

Epoch 00001: val_loss improved from inf to 0.06362, saving model to model-
Epoch 2/1000000
517/517 [==============================] - 101s 195ms/step - loss: 0.0561

Epoch 00002: val_loss improved from 0.06362 to 0.05809, saving model to mo
Epoch 3/1000000
517/517 [==============================] - 108s 209ms/step - loss: 0.0528

Epoch 00003: val_loss improved from 0.05809 to 0.05584, saving model to mo
Epoch 4/1000000
517/517 [==============================] - 98s 188ms/step - loss: 0.0510 -

Epoch 00004: val_loss improved from 0.05584 to 0.05428, saving model to mo
Epoch 5/1000000
517/517 [==============================] - 103s 197ms/step - loss: 0.0500
```

Figure 4.9: Training Process for NDVI by U-Net

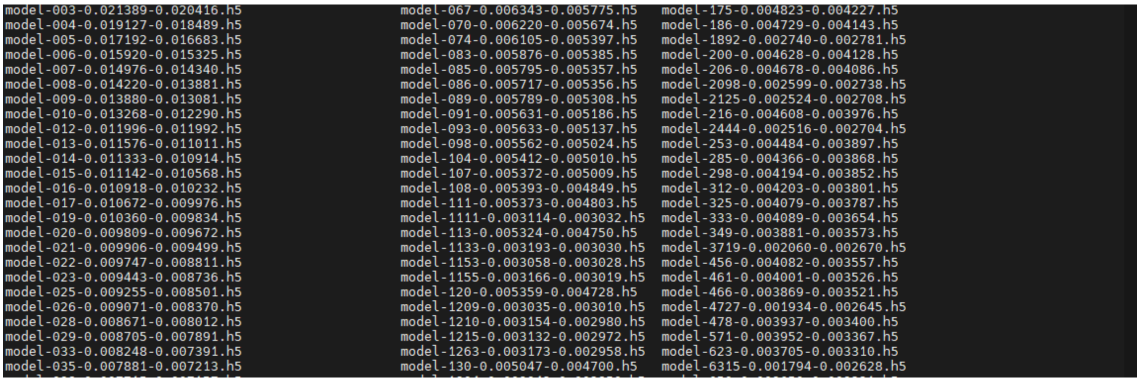


Figure 4.10: Generated h5 files while training

These h5 files are generated while training which will be used in the testing of the NDVI.

4.3 Testing

The testing is done with the help of h5 files generated while training. This is the post-processing step which in simple terms can be said as the steps used to create Cubed Sphere files which are recreated in reverse order. Hereafter post-processing, the normal NETCDF files will come which will thus in the case be used for evaluation.

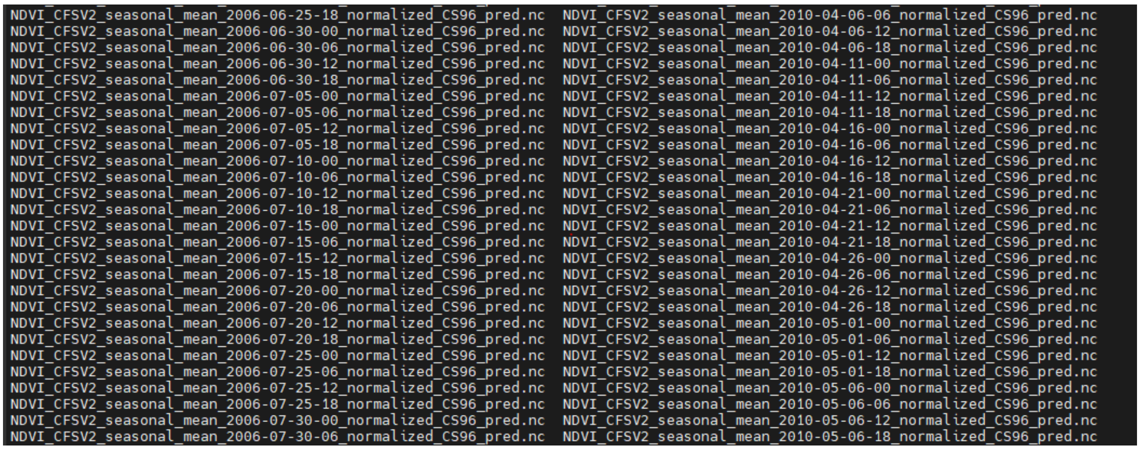


Figure 4.11: Cubed Sphere files generated while post processing method

```
NDVI_CFSV2_seasonal_mean_2010-05-01-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-01-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-06-00_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-06-06_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-06-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-06-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-11-00_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-11-06_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-11-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-11-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-16-00_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-16-06_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-16-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-16-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-21-00_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-21-06_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-21-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-21-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-26-00_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-26-06_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-26-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-26-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-31-00_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-31-06_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-31-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-05-31-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-06-05-00_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-06-05-06_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-06-05-12_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-06-05-18_normalized_CS96_pred_to_rll_inverse_norm.nc
NDVI_CFSV2_seasonal_mean_2010-06-10-00_normalized_CS96_pred_to_rll_inverse_norm.nc
```

Figure 4.12: Generated files for evaluation

With these files the post-processing steps are complete and the evaluation process starts.

Chapter 5

Results And Discussions

The first result is to plot the monthly leads with the parameters being AVHRR, DL, and Bias (DL-AVHRR) which is shown below for Jan, Feb, and March, and these type of plots is also done for April, May up to December.

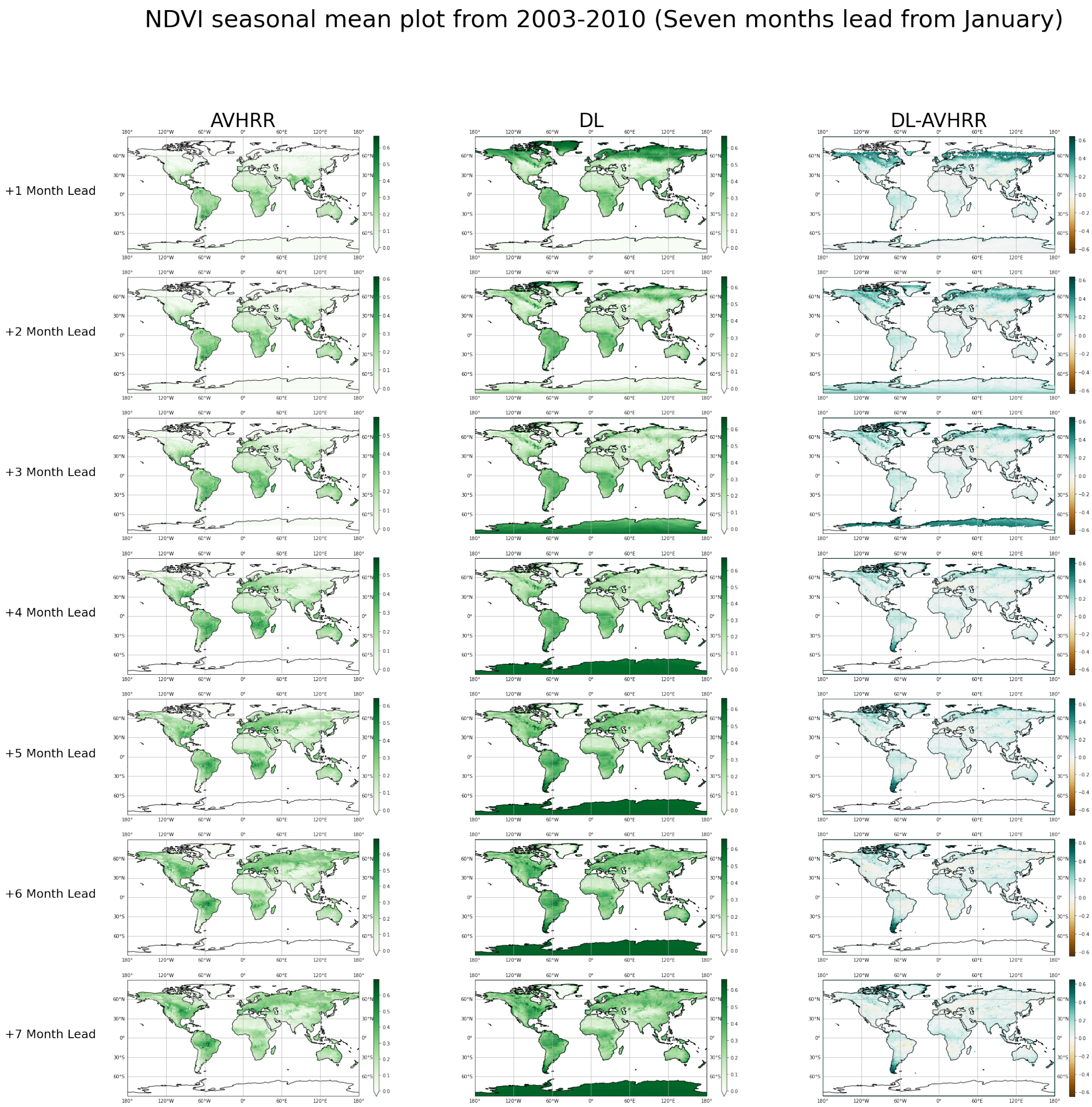


Figure 5.1: Plot for seven months lead from January

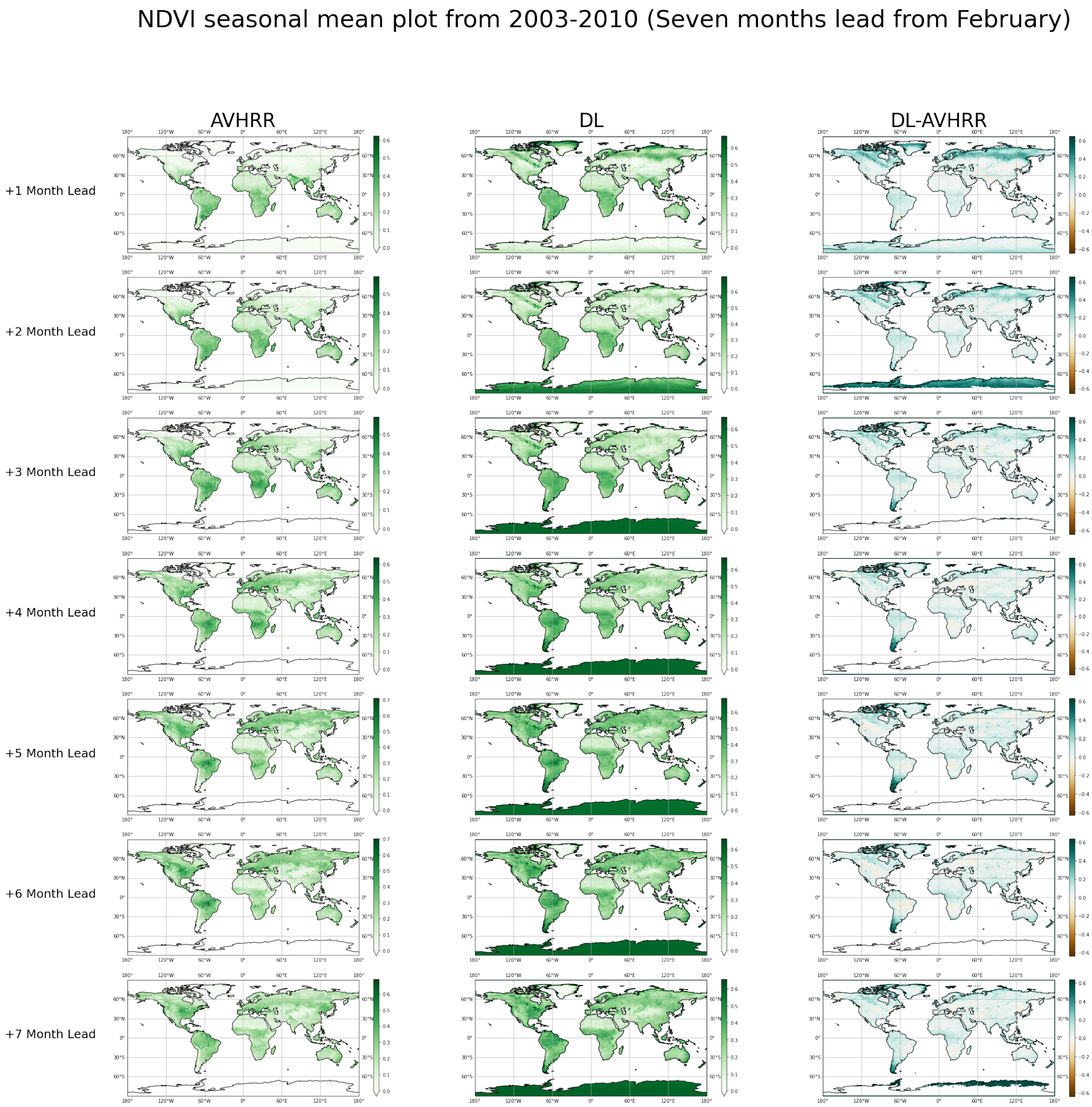


Figure 5.2: Plot for seven months lead from February

The monthly lead plots can be inferred by the fact that these results can be taken as a reference from [11] to predict wheat yield in the range of 30 days before harvesting. As this can be a better way to check for the results derived as the lead days is 210 days which is 7 times 30 days.

So, [11] is taken as a reference to validate the results and compare prediction with ground truth for crop yield forecasting. As this is about the intensity of vegetation so for this case, the case study of forest fire is taken into account to evaluate the model's performance.

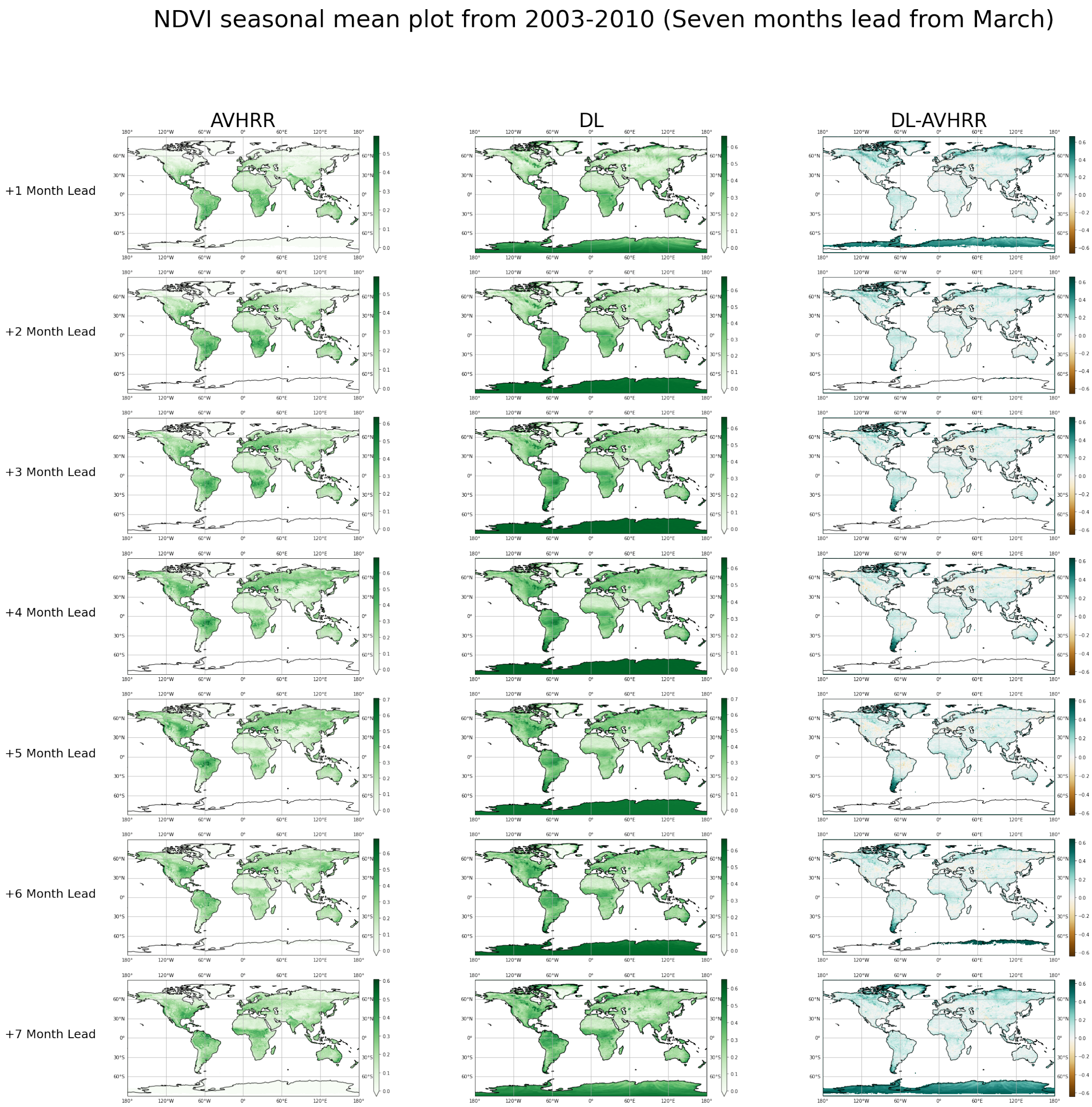


Figure 5.3: Plot for seven months lead from March

These seven months lead is let's say +1 month lead will be from the monthly mean of January if we start the lead from January, then +2 months will be January and February and so on up to +7 months. The mean is taken for the years in test cases starting from 2003 to

2010. This means that if we want January means then it will be taken for all the January from 2003-2010 and likewise be going on.

The bias is calculated which in this case is low for the NDVI variable and the bias is done by subtracting AVHRR which is the original NDVI from AVHRR from DL which is the output from the Deep Learning Model.

The next result is in context with the forest fires around the world. There are six forest fire cases taken from every continent which can help to analyze the different NDVI before and after the forest fires around the globe.

The six test cases are taken are as follows:-

1. 2005 Torres del Paine Fire:- The forest fire started in February 2005 and lasted for 10 days in the Torres del Paine National Park located in Chile, South America.
2. 2006 Table Mountain Fire:- The fire started in January 2006 and lasted for 2 days in the Table Mountain National Park located in South Africa.
3. 2008 California Wildfires:- The wildfires started in April 2008 and lasted for around 7 months which is November across California, North America.
4. Black Saturday Bushfires of 2009 (Victoria):- The forest fire started on 7th February in 2009 and went up to 14th March across Victoria State, Australia.
5. 2009 Southeast Asian Haze:- The haze started in June 2009 and lasted for around 4 months which is September during the monsoon season across Indonesia, Malaysia, and Singapore.
6. 2009 Mediterranean Wildfires:- The wildfires started in July 2009 and lasted for around 1 month across France, Greece, Italy, Spain, and Turkey.

The table is made to check the values originally in the AVHRR data and the predicted values from the Deep Learning model which is shown below:-

Case Studies on Forest Fires	NDVI Value in AVHRR	NDVI Value in DL
2005 Torres del Paine Fire	0.25	0.57
2006 Table Mountain Fire	0.15	0.59
2008 California Wildfires	0.23	0.36
Black Saturday Bushfires of 2009 (Victoria)	0.18	0.57
2009 Southeast Asian Haze	0.30	0.56
2009 Mediterranean Wildfires	0.26	0.34

Table 5.1: Case study of forest fires

The last result is the prediction of NDVI for India in the monsoon months from 2003 to 2010. The months taken are June, July, August, and September (JJAS). The mean is taken for these months for testing years and a bar plot is made which is shown below:-

The bar plot shows that the average NDVI value from AVHRR in June, July, August, and September which are the monsoon months is the average around 0.25 which can be seen in the paper [5]. There are several cases in which we see a dip in the NDVI value in monsoon can be due to floods, thunderstorms, etc for the whole of India.

In [5], there is a month-wise NDVI for the period 1981 to 2000 which showed that the NDVI values decrease for June around 0.25 and then increase from July to September up to 0.3, so in this case for the years 2003 to 2010, taking means of June, July, August, September, the average AVHRR NDVI is around 0.23-0.25 and the DL NDVI is around 0.33-0.35.

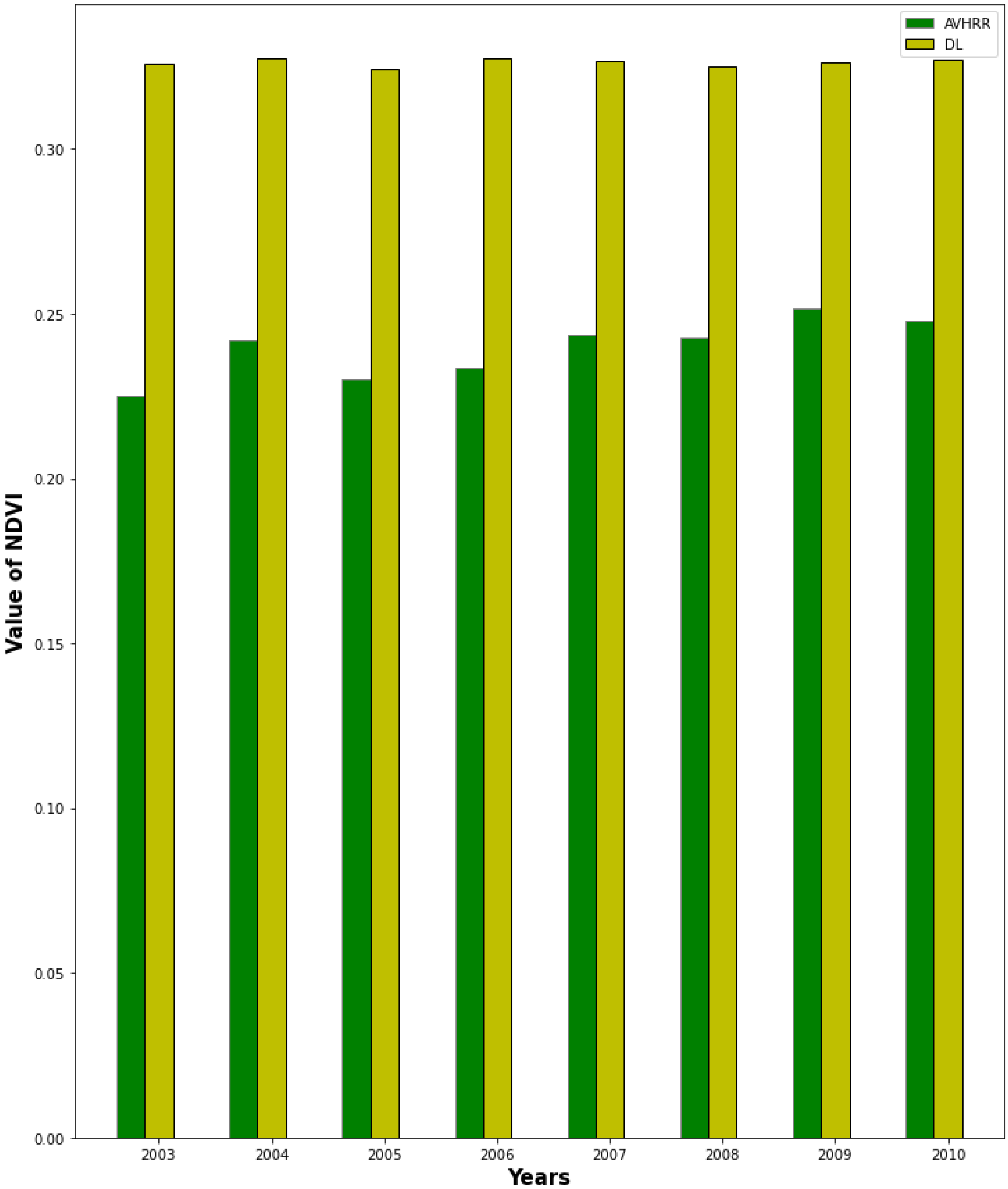


Figure 5.4: Bar plot for NDVI in India for monsoon months from 2003 to 2010

Chapter 6

Conclusion

The NDVI data which is taken from AVHRR with the range of years from 1982 to 2021 is analyzed to determine the seasonal forecasting of the NDVI variable using the U-Net model. The data analyzed is taken from the whole world, so the case studies can be taken from anywhere around the world and determine NDVI which in turn helps to understand the crop yield forecasting.

The seasonal forecasting of NDVI is taken based on 210 days lead which can be interpreted as 7 months lead. For example, the current month is January so after 7 months which is August then we can calculate NDVI in August. This will help us determine how much vegetation is gonna be available so farmers can harvest crops according to it.

The evaluation is done by taking several test cases including forest fires. This will help to understand the intensity of vegetation before and after the forest fires. Nowadays, we see forest fires cases growing over the years, so with the evaluation of seasonal forecasting of NDVI, there are great chances of getting an idea of evaluating the damage done by forest fires.

Most of the research work done on NDVI or in crop yield forecasting had focused on certain areas and not on the whole globe. This project will help you analyze any area around the globe with just the slicing of latitude and longitude. This will help people to understand one place and go with the research as per their area of interest around the world.

The seasonal forecast will help the farmer to get an idea about the intensity of vegetation in a certain area before the crop yielding season starts. The best use case will be targeting the monsoon season where the evaluation of NDVI will help them understand how much crop is needed to get a better yield and thus save the wastage of crops.

The damage done by floods, forest fires, etc, to vegetation can also be evaluated. This will help to evaluate the damage done by the forest fires and floods and can help to prepare for the next events of these cases.

Bibliography

[1] Manmeet Singh, Bipin Kumar, Suryachandra Rao, Sukhpal Singh Gill, Rajib Chattopadhyay, Ravi S Nanjundiah, Dev Niyogi. Deep learning for improved global precipitation in numerical weather prediction systems. arXiv.org, Physics, Atmospheric and Oceanic Physics (physics.ao-ph); Machine Learning (cs.LG), arXiv:2106.12045v2 [physics.ao-ph], 24 Aug 2021.

[2] Jonathan A.Weyn, Dale R.Durran, Rich Caruana. Improving Data-Driven Global Weather Prediction Using Deep Convolutional Neural Networks on a Cubed Sphere. arXiv:2003.11927v1 [physics.ao-ph], 15 Mar 2020.

[3] Liuqing Ji, Ke Fan . Climate Prediction of Satellite-Based Spring Eurasian Vegetation Index (NDVI) using Coupled Singular Value Decomposition (SVD) Patterns. Remote Sens. 2019, 11, 2123; doi:10.3390/rs11182123.

[4] Mathyam Prabhakar, K.A. Gopinath, A.G.K. Reddy, M. Thirupathi, Ch. Srinivasa Rao. Mapping hailstorm damaged crop area using multispectral satellite data. https://doi.org/10.1016/j.ejrs.2018.09.001

[5] J. V. Revadekar, Yogesh K. Tiwar, K. Ravi Kumar. Impact of climate variability on NDVI over the Indian region during 1981–2010. http://dx.doi.org/10.1080/01431161.2012.697642

[6] Eric S. Kasischke, Nancy H. F. French, Peter Harrell, Norman L. Christensen, Jr, Susan L. Ustin, Donald Barry. Monitoring of Wildfires in Boreal Forests Using Large Area AVHRR NDVI Composite Image Data. REMOTE SENS. ENVIRON. 45:61-71 (1993).

[7] Dyah R. Panuju, Bambang H. Trisasongko1n, Budi Susetyo, Mahmud A. Raimadoya, Brian G. Lees. Historical Fire Detection of Tropical Forest from NDVI Time-series Data: Case Study on Jambi, Indonesia. ITB J. Sci. Vol. 42 A No. L 2010, 49-66.

[8] M. E. BROWN, D. J. LARY, A. Vrieling, D. Stathaki, H. Mussa . Neural networks as a tool for constructing continuous NDVI time series from AVHRR and MODIS. http://dx.doi.org/10.1080/01431160802238435.

[9] Habib Aziz Salim, Xiaoling Chen, Jianya Gong. Analysis of Sudan Vegetation Dynamics Using NOAA AVHRR NDVI Data from 1982-1993. https://dx.doi.org/10.3923/ajes.2008.1.15.

[10] Xiao Xiang Zhu, Devis Tuia, Lichao Mou, Gui-Song Xia , Liangpei Zhang, Feng Xu, Friedrich Fraundorfer. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. CoRR, abs/1710.03959.

[11] Lopresti, M. F., Di Bella, C. M., Degioanni, A. J. Relationship between MODIS-NDVI data and wheat yield: A case study in Northern Buenos Aires province, Argentina. Information Processing in Agriculture, 2(2), 73–84. doi:10.1016/j.inpa.2015.06.001.

[12] Huang, J., Wang, X., Li, X., Tian, H., Pan, Z.,. Remotely Sensed Rice Yield Prediction Using MultiTemporal NDVI Data Derived from NOAA's-AVHRR. PLoS ONE, 8(8), e70816. doi:10.1371/journal.pone.0070816.

[13] Olaf Ronneberger, Philipp Fischer, Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. https://doi.org/10.48550/arXiv.1505.04597.

[14] https://www.ncei.noaa.gov/data/avhrr-land-normalized-difference-vegetation-index/access/