5-1-2010

Estimating the Carbon Fluxes using the CASA Model in the Southern United States

Venkata Narendra Appala Rongali

ESTIMATING THE CARBON FLUXES USING THE CASA MODEL IN THE SOUTHERN UNITED STATES

By

Venkata Narendra Appala Rongali

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master in Science
in Electrical Engineering
in the Department of Electrical and Computer Engineering

Mississippi State, Mississippi

May 2010

ESTIMATING THE CARBON FLUXES USING THE CASA MODEL IN THE SOUTHERN UNITED STATES

By

Venkata Narendra Appala Rongali

Approved:

Nicolas H. Younan
Professor and Department Head
Dept. of Electrical and Computer Engg.
Major Advisor

Surya S. Durbha
Assistant Research Professor
Dept. of Electrical and Computer Engg.
Co-Major Advisor

Jenny Q. Du
Associate Professor
Dept. of Electrical and Computer Engg.
Committee Member

James E. Fowler
Professor
Dept. of Electrical Engineering
Graduate Program Director

Sarah A. Rajala
Dean of Bagley College of Engineering

Name: Venkata Narendra Appala Rongali

Date of Degree: May 1, 2010

Institution: Mississippi State University

Major Field: Electrical Engineering

Major Professor: Dr. Nicolas H. Younan

Title of Study: ESTIMATING THE CARBON FLUXES USING THE CASA MODEL IN THE SOUTHERN UNITED STATES

Pages in Study: 90

Candidate for Degree of Master of Science

A minute change in human body temperature can get one collapsed permanently. If this is the case with human body, one might wonder what would be the case with Earth. The result is the greenhouse. The main function of the greenhouse gases (GHG's) is to tap energy from the sun and prevent the heat energy escaping to space, thus sustaining life on earth. Because of increased human activity, industrialization, and deforestation, the concentration of the GHG's in the atmosphere has increased, resulting in a temperature rise. Considering the effects of temperature rise, caused by the GHG's, one should know the ways to minimize them. In order to do this, an estimation of the amount of GHG's is important. The CASA model is one such model that estimate the GHG's and also the amount of carbon in the atmosphere by estimating the Net Ecosystem Productivity (NEP) and Net Primary Productivity (NPP).

DEDICATION

I would like to dedicate this thesis to God, to my parents to whom I owe everything, and to my sisters, in-laws and friends, who have been my continual support.

ACKNOWLEDGEMENTS

I would like to thank my major advisor, Dr. Nicolas H. Younan, for his advice and support while working on this research. I would also like to thank my co-major advisor, Dr. Surya S. Durbha, for his supervision, trust, support and guidance from the early stages of this research. Finally, I would like to acknowledge my committee member, Dr. Jenny Q. Du, for her support throughout the research.

I gratefully acknowledge the financial support provided by the GEOSYSTEMS RESEARCH INSTITUTE (GRI) AT MISSISSIPPI STATE UNIVERSITY FOR THIS RESEARCH for this research. I would like to thank fellow graduate students and faculties of the Geosystems Research Institute and the Department of Electrical and Computer Engineering for their support and advice. I would like to thank the National Aeronautics and Space Administration - Rapid Prototyping Center (NASA - RPC) people for funding this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A minute change in human body temperature makes one sick. The change can be to such an extent that one could possibly die. If this is the case with the human body, one might wonder how a change like this might affect the Earth. The result of rise in temperature is the greenhouse effect. The main function of the greenhouse gases (GHG) is to tap the energy from the sun and prevent the heat energy from escaping back to space, thus sustaining life on earth. Because of increased human activity, industrialization, and deforestation, the concentration of the greenhouse gases in the atmosphere has increased, causing a rise in temperature. This rise in temperature is referred to as global warming.

In observing the temperatures of the past century, warmest temperatures have been recorded in the past two decades, with 2005 being the hottest year followed by years 1998, 2002, 2003 and 2004. Over the past thirty years, the Earth has warmed by 0.6° C or 1.08° F. Over the past one hundred years, it has warmed by 0.8° C or 1.44° F [1]. The temperature rise has adverse effects on the living chain. The consequences of the rise in temperature include sea levels rising, melting of the polar ice caps, global warming which further leads to drought and deforestation (by forest fires), decrease in the quantity of moisture, extinction of living beings and increased global warming. Keeping in mind the causes and effects of this rise in temperature; one should know the ways to minimize or decrease them. In order to do this, one should first know the methods of estimating the amount of GHG's and then apply relevant methods to decrease them. Several methods can be followed to estimate the amount of GHG's and other energy exchanges between the atmosphere and biosphere. The continuous field model, the Vegetation Ecosystem Modeling and Analysis Project (VEMAP) and the CASA model are a few of the models that estimate the amount of carbon in the atmosphere. 6.5 billion tons of CO_2 are released into the earth's atmosphere each year [2].

Background

Carbon Sequestration

Carbon sequestration is the storage of carbon through a biological process such as fossil formation, a chemical process, such as acid rain, or a physical process such as volcanic eruptions. The carbon that is usually stored is the carbon that is absorbed from the atmosphere. The carbon sequestration helps in reducing global warming by decreasing the amount of carbon that is present in the atmosphere and storing it in a solid form.

Carbon sequestration is a growing problem today because of the technology development and industrialization. It is the biggest problem that the environment faces. The carbon sequestration problem, if not addressed, may have an adverse effect in the future on the plants, the environment, humans, and the whole Earth. Global warming is a consequence of the carbon sequestration from burning fossil fuels.

The combined effort of many countries for the mitigation of the greenhouse gases came into existence via the Kyoto protocol [3].

Kyoto Protocol

Kyoto Protocol is an international agreement about climate change. The main objective behind the development of the Kyoto protocol is the stabilization of the climate. It sets goals for countries to reduce the greenhouse gases like methane, carbon, and chlorofluorocarbons (CFCs). The protocol was adopted in Kyoto, Japan on December 11, 1997 and has been effective since February 16, 2005.

The protocol aims to reduce the GHG's, helping countries reach their goal by means of marketing such as:

a) Emission trading, which is known as the 'Carbon Market' or 'Carbon Credits'
b) Clean development mechanism
c) Joint implementation

By emission trading it does not mean encouraging the GHG and trade them to other countries in order to reach their target. Emission trading means the redistribution of emitted carbon and other GHG.

Carbon Credit

The Kyoto protocol sets a price for carbon gas. According to the protocol, each country has an allowance of carbon that can be released into the atmosphere. This limitation in turn sets a cutoff point for the carbon released by industries. Industries which have not used their allowance can trade the carbon with industries that have exceeded the limit, thus balancing the overall cutoff. Industries that have emissions below the allowed quota are said to possess Carbon Credits. Industries can trade the carbon for a monetary value or for other purposes in the open market. Those industries which have exceeded the allowed quota can either employ superior machinery to limit the carbon released or trade the carbon with other industries.

The carbon credits provide flexibility for the industries. The carbon can be traded not only within the country, but it can also be a trading among nations thus leading to a collective balance and decrease in the amount of carbon that is released into the atmosphere.

Source and Sink

Source and Sink of a particular region are estimated by the net amount of carbon, that a particular region or country emits or absorbs. If a region emits carbon, then it is called a source of carbon, and if a region absorbs carbon, then it is a sink of carbon. Since the Net Ecosystem Productivity (NEP) and Net Primary Productivity (NPP) indicate the amount the carbon that is being absorbed, hence NEP or NPP also give an estimation of source or sink. A positive NEP indicates a sink while a negative NEP indicates a source [4].

Determining the source and sink plays an important role in decision making, designing new methods and technology, and controlling the amount of carbon released.

IPCC and Global Warming

The Intergovernmental Panel on Climate Change (IPCC) is an organization formed by the joint collaboration of the United Nations Environment Program and the World Meteorological Organization. The main objective of the IPCC is to review and assess the climate information and have an understanding of the climatic changes that are taking place [5].

There are three different working groups that function together to review and assess the climate. The main objective of the working groups is to assess the physical aspects of the climate, such as the temperature changes, the carbon cycle, the amount of gases present in the atmosphere, especially the gases which contribute for temperature rise, the amount of rainfall, and the ocean and sea level. Physical changes in the climate can vary from global warming to rains and forest fires. The working groups also review the satellite data and climate models; estimate the vulnerability of various systems to climate change, the effects of climate change and how to adapt to the consequences of the change. This change includes the alteration in water resources, ecosystem, food, forests, industry, and human development. Furthermore, these groups take measures to decrease the greenhouse gas emissions and their effects both short-term and long-term, taking into consideration the development and management of the way the wastes and gases are being released by various sources like industry, agriculture, forest, and human development.

NEP and NPP

Regions are distinguished as sources and some as sinks based on the parameters that give the net exchange of carbon that takes place in the atmosphere. Such parameters include the net ecosystem productivity (NEP), the net primary productivity (NPP), the gross primary productivity (GPP), and the net ecosystem exchange (NEE); with NEP and NPP being the direct outputs of the CASA model.

The net ecosystem productivity is the carbon absorbed and released by plants, i.e., the net carbon exchange between the ecosystem and the atmosphere. The net primary productivity is the net carbon absorption rate by living plants only. It is the plant photosynthesis rate minus the photorespiration rate.

The GPP and NPP determine how fast the carbon is taken up by the ecosystem while NEP tells how much carbon is stored by the ecosystem. By knowing NEP or NPP, the damage to the vegetation due to climate change can be estimated. If NEP is negative, then the region is a source of carbon. If NEP is positive, then the region is a sink for carbon.

Objectives

The main objective of this study is to estimate the carbon fluxes in the southern United States via the CASA model. Several inputs that are used to run the CASA model are the normalized difference vegetation index (NDVI), the precipitation data (PPT), soil data, air temperature, land mask data, vegetation type and solar radiation.

Various aspects that are involved in the implementation of the model are:

- Downloading the necessary data (satellite data), direct or indirect use of various other satellite data to form the raw input data for the CASA model.
- Converting the raw data into the necessary format (to NetCDF format) needed for the CASA model.
- Applying a projection transformation on the data.
- Converting the raw data into a formatted data involves several steps like HDF to GeoTIFF conversion and then to NetCDF conversion and using several softwares, like MODIS Tool, Arc Map, etc.
- Writing Linux and Perl scripts that make the process automatic and also run the model without the user intervention for all the states.
- Creating and formatting the CASA output data sets as input for the Weather Research Forecasting (WRF) model.
- Viewing and analyzing the outputs of the CASA model using visualizing tools like ncBrowse and ncView.

The outputs that are of special interest are the Net Ecosystem Productivity (NEP) and Net Primary Productivity (NPP), which help to characterize carbon source and sinks.

Thesis Organization

Chapter I provides an introduction to the terms used in the thesis, while Chapter II gives the literature overview, all the related work being done in the NASA - CASA project and about the weather research and forecasting (WRF) model. Chapter III describes how the research is carried about, the inputs and outputs, and all the scripts written. Chapter IV talks about the results and discusses the NEP and NPP based on the results and outputs of the CASA model. Chapter V tells about the contribution towards the project, what else could be done in getting the required outputs, and what more could be achieved. It also describes the data sets that can be used so that the CASA model is independent of the input data sets.

CHAPTER II

LITERATURE OVERVIEW

The NASA Carnegie-Ames-Stanford Approach, commonly known as the NASA-CASA, is used in the estimation of the carbon in the atmosphere. The CASA model is an approach for the estimation of the carbon that is being released and absorbed, hence allowing one to estimate whether a particular region is a source or a sink for carbon. This estimation is necessary which acts as a feedback to the contributors, who in turn can reduce the GHGs released into the atmosphere.

The research is focused on the multi-year or monthly data sets. The eco-system varies greatly when considered locally and globally. The research also focuses on the regional or global scale study of the GHG. The CASA model is mainly focused on the estimation and exchange of carbon within the eco-system, while the estimation of other gases, like nitrogen, is also possible.

CASA Model and Data Assimilation

Most of the energy exchange between all the living and non-living things within the atmosphere take place through CO_2 component. CO_2 measuring instruments provide additional data for research and forecasting and estimating the cause and consequences of GHG's. The NOAA - TIROS Operational Vertical Sounder (TOVS) [6], SCIAMACHY

[7] and AIRS [8] are few satellites which provide us with CO_2 concentrations by observing the spectral levels in the CO_2 absorption bands. These instruments led to developing a CO_2 specific instrument, the OCO (Orbiting Carbon Observatory) [9].

The OCO instrument also measures solar radiation after reflection at the Earth's surface by aiming at the sun glint over the oceans, which enhances the observed signal. The enhancement in the observed signal helps in sensitive CO_2 measurements. The measuring grid of OCO is small, thus increasing the probability of cloud free measurements [10].

Carbon from the atmosphere enters and leaves the land through plants and water. This process where carbon enters the land from atmosphere and back to land is called as carbon cycle as shown in Figure 1. NPP and NEP are the pathways by which carbon enters the biosphere from the atmosphere. A few of the models which estimate the NEP and NPP are the NASA-CASA model, the VEMAP model [11], the Continuous Field Model [12], the Global Production Efficiency model (GLO-PEM), the Century model, the Global Biome Model-Biogeochemical Cycle (BIOME-BGC) and the Terrestrial Ecosystem Model (TEM) [13]

The Carnegie Ames Stanford Approach is a simulation model that is used in analyzing the impacts of various greenhouse gases, especially carbon, on the ecosystem. The analysis can be done on a regional scale basis or on a global scale. The NASA – CASA has done significant research in the examination of the exchange on a global scale while a regional scale study has also gained significant momentum. CASA computes NPP as a function of absorbed photosynthetically active radiation (APAR), light use efficiency (LUE) ε^*, Temperature T , and moisture W i.e.,

$$NPP5 = 5f APAR\ PAR\ 5 \quad (2.1)$$

where *f*APAR is the fraction of APAR, solar radiation is converted to PAR by multiplying by 0.5 and *f*APAR x PAR equals APAR. ε^* was set to 0.405 as described by Potter et al., 1993 [14]

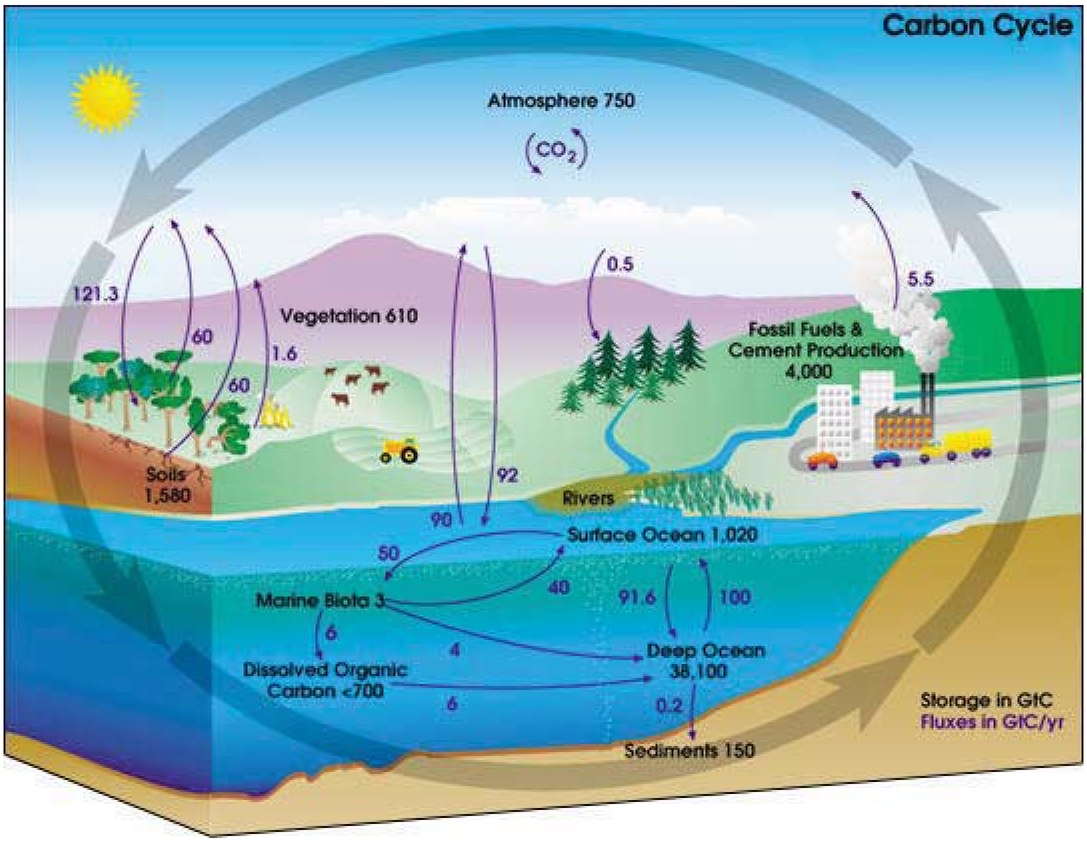

Figure 1 Carbon Cycle [15]

WRF Model

The Weather Research and Forecasting (WRF) [16] model is a collaborative work of the National Center for Atmospheric Research (NCAR), the National Centers for Environmental Prediction (NCEP), the Air Force Weather Agency (AFWA) and the National Oceanic and Atmospheric Administration (NOAA) designed to predict weather for forecasting and atmospheric research. A more specialized version is the hurricane WRF (HWRF), which is used to forecast and track cyclones. The WRF model operates at a resolution of 4 km to 12.5 km. The WRF model can also be used for regional climate modeling, chemistry and air quality research and prediction, cloud and storm modeling, and data assimilation. It can be made to work on parallel processors.

The WRF model uses a 3-D grid to represent the atmosphere, topographical land information and observational data to define initial conditions for the model [17].

Land and Biosphere Models

The physical representation of the ecosystem is what is called a model. The model describes what is happening in a place, specifically the exchange of energy between the atmosphere and biosphere. The models can be distinguished into atmosphere models and biosphere models.

The atmosphere models describe the exchange with respect to height above sea level. They estimate the effect of the various gases that are being released into the atmosphere. The estimation varies with altitude. The atmosphere models can be further classified into global scale atmospheric models, where the exchange is estimated on a global scale such as for a country or for the whole Earth, and regional scale atmospheric models, which concentrates on a particular region, such as the WRF model.

The biosphere models, unlike the atmospheric models, try to estimate the exchange from land to the atmosphere considering the biological phenomenon. The land biosphere models, such as the CASA model, estimate the exchanges between the atmosphere and land (soils), plants, water (molasses), vegetations, and organisms.

CHAPTER III

METHODOLOGY

Comprehensive Description of the CASA Model

Basic Implementation of the CASA

The CASA model takes NDVI, PPT, soil, land mask, air temperature, vegetation and solar radiation as inputs and gives out NEP, NPP, Leaf Area Index (LAI), soil carbon (SoilC) etc. as outputs. The inputs are of 8 km cell resolution. The list of outputs of the CASA can be found in figures 3 and 4. The CASA model can be run in either LINUX or Windows operating systems. It is run at command prompt as shown below

make; ./casa.out

The internal working of the CASA to give the required output (that is of interest) NPP can be understood by the following set of equations (1.1) through (1.11) [18]

$$NPP = \max_{t} \{ NPP(x,t), 0 \} \quad (1.1)$$

$$NPP(x,t) = IPAR(x,t)\,\varepsilon(x,t) \quad (1.2)$$

where:

IPAR = the intercepted photosynthetically active radiation,

= the light utilization efficiency at a grid cell ‘x’ in a month‘ t ‘,

NPP = the net primary productivity, with

$$IPAR(x,t) = SOL(x,t)\ FPAR(x,t)\ 0.5 \quad (1.3)$$

where:

SOL = the solar radiation incident on the grid cell x in the month t

FPAR = the fraction of photosynthetically active radiation

0.5 = accounts for the fact that approximately half of the incoming radiation is the PAR in 0.4 – 0.7 μm

$$FPAR(x,t) = min\left\{\left[\frac{SR(x,t)}{SR_{max}-SR_{min}} - \frac{SR_{min}}{SR_{max}-SR_{min}}\right], 0.95\right\} \quad (1.4)$$

$$SR(x,t) = \frac{1+NDVI(x,t)}{1-NDVI(x,t)} \quad (1.5)$$

where:

FPAR = the fraction of photosynthetically active radiation and is a linear function of AVHRR advanced very high resolution radiometer simple ratio (SR)

NDVI = the normalized difference vegetation index at a grid cell ‘x’ in the month‘t’

SR_{min} = SR for unvegetated lands and equals 1.08 for all grid cells

0.95 = the finite upper limit of the leaf area

$$\varepsilon(x,t) = T_{\varepsilon 1}(x,t)\, T_{\varepsilon 2}(x,t) W_{\varepsilon}(x,t)\, \varepsilon^* \quad (1.6)$$

where:

$T_{\varepsilon 1}$ and $T_{\varepsilon 2}$ = the effects of temperature stress

W = is the effects of water stress

ε^* = is the maximum possible efficiency which equals to 0.56

$$T_{\varepsilon 1}(x,t) = 0.8 + 0.02\, T_{opt}(x) - 0.005\, (T_{opt}(x))^2 \quad (1.7)$$

$$T_{\varepsilon 2}(x,t) = \frac{1.1814}{(1+e^{0.2\{T_{opt}(x)-10-T(x,t)\}})(1+e^{0.3\{T(x,t)-10-T_{opt}(x)\}})} \tag{1.8}$$

$$W_{\varepsilon}(x,t) = 0.5 + 0.5\frac{EET(x,t)}{PET(x,t)} \tag{1.9}$$

where:

$T_{opt}(x)$ = the temperature in the month when NDVI is max in the year

EET = the actual estimated evapotranspiration come from the soil moisture

PET = the monthly potential evapotranspiration come from the soil moisture

$$SOILM(x,t) = \begin{cases} SOILM(x,t-1) - [PET(x,t) - PPT(x,t)]RDR, \quad and\ PPT(x,t) < PET(x,t) \\ SOILM(x,t-1) + [PPT(x,t) - PET(x,t)], \quad and\ PPT(x,t) \geq PET(x,t) \end{cases} \tag{1.10}$$

where:

PPT = the average precipitation at a month t in the grid cell x

RDR = the relative drying rate scalar for potential water extraction as a function of soil moisture.

$$RDR = \frac{1+a}{1+a\,\Theta^{b}} \tag{1.11}$$

where:

'a' and 'b' are texture dependant empirical coefficients and

Θ is the volumetric moisture content (m/m)

The following are the inputs that are fed to the NASA – CASA [19] model.

Input Data Sets

Normal Difference Vegetation Index (NDVI)

Plants are the major source of carbon. By the process of respiration plants release carbon dioxide into the atmosphere. Usually plants release carbon during the night and take in carbon dioxide during the day. The amount of intake or release depends on the active area of the leaf that the plant uses. The NDVI gives the vegetation index of a plant, i.e., the effective area normalized to get an average value.

The healthiness of a plant or vegetation is indicated by NDVI. Healthy vegetation reflects very well in the near infrared while absorbing most of the visible light. Unhealthy vegetation reflects more of the visible light. Thus, one can determine the healthiness of the vegetation by measuring the reflectance of the near infrared radiation and the reflectance of the visible radiation. NDVI can be formulated as

NDVI = (NIR – VIS) / (NIR + VIS). NDVI is unit less since it is a ratio and ranges between -1 and +1, while a 0 indicates no vegetation and 1 indicate dense healthiest vegetation. Figure 2 shows how NDVI is calculated considering an example of spring and fall seasons.

The satellites can measure the *f*APAR that is absorbed by the vegetation. The data of interest for NDVI is downloaded from the Earth Observing System Data Gateway in HDF (Hierarchical Data Format) [20]. The HDF file format that is used in the management of large and complex data contains several other data sets. The HDF file, of interest, contains several other data sets like NDVI, Enhanced Vegetation Index (EVI), Vegetation Index (VI) quality, red reflectance, near infrared (NIR) reflectance, blue reflectance, middle infrared (MIR) reflectance, vie zenith angle, sun zenith angle, relative azimuth angle, and pixel reliability, of which NDVI is of special interest to us. All the data sets are of 1km monthly.

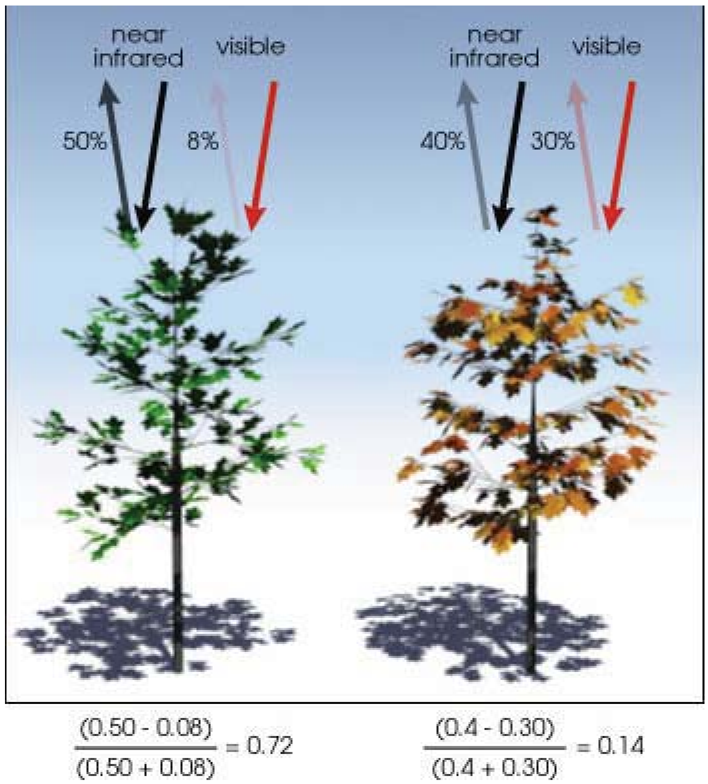

Figure 2   Calculation of NDVI [21]

The NDVI data can be extracted by using a tool MODIS TOOL. The MODIS tool can convert HDF to GeoTIFF that can be formatted in ArcMap. Each region consists of twelve data sets that represent twelve months. Each of the NDVI data set corresponding to each month is extracted using the state soil data as mask; thereby NDVI and soil data sets are of the same size and a respective pixel of NDVI correspond to the respective pixel of soil.

Precipitation (PPT)

Precipitation is the amount of water that reaches the ground. Precipitation directly affects the carbon storage. An increase in precipitation increases the carbon storage. The greater the precipitation is the greater is the plant growth; hence the greater is the amount of carbon that is being stored.

The process of storing the atmosphere carbon (usually greenhouse gases) into a solid form by various physical and biological processes is known as carbon sequestration. The carbon sequestration thus decreases the amount of greenhouse gases in the atmosphere by converting them into solid form thereby help in reducing global warming.

The following are steps involved in converting raw precipitation data into a formatted input for the CASA model. The data is downloaded from the PRISM website [22]. After downloading the gridded precipitation data from the website; the file is unzipped and added an extension '.txt' at the end of the unzipped file. The file then changes into ASCII formatted file. The downloaded file is made up of individual monthly precipitation data for the whole North America. The size of the data is 1405 x 621. There are 12 data sets that represent 12 months. The unit of precipitation is mm/month.

Air Temperature (AIRT)

Temperature is an important parameter for reactions to take place. A slight change in the operating temperature alters a process a lot. Sea level rising and ozone hole expansion cause the chemicals to breakdown and react with the atmospheric gases, altogether creating worse environment.

The temperature data fed as an input to the CASA model is not exactly the mean temperature data sets. It is the average of the maximum and minimum temperatures. The units of the temperature data is degrees Celsius. The temperature data sets are also downloaded from the PRISM website [22].

Soil

Soil forms a major form of carbon source. The amount of carbon that is stored in soil is twice the amount of carbon that the plants store. Different soil types release or take different amounts of carbon depending on its nature, such as coarse, fine, sand, and clay. Carbon is exchanged between the atmosphere and the soil through plants.

The soil data is downloaded from Geospatial Data gateway [23].

Solar Radiation (SOLRAD)

Solar radiation, without which there would not be any existence of living things on this Earth, is an important natural and renewable energy. Usually radiation consists of the ultraviolet and near infrared, i.e., the visible and near visible radiations whose wavelength is short and ranges from 0.3μm to 3.0μm hence called a short wave solar radiation.

The solar radiations after filtering the ultraviolet radiations from the Earth's atmosphere reach the surface of the earth in less magnitude. This decrease in the quality and quantity of the radiation is beneficial to the living organisms on Earth, thus leading to their growth. The solar radiation data is usually measured in watts $/m^2$. The solar radiation data is downloaded from the Earth System Research Laboratory website [24] and [25].

Land Mask

Landmask is an input to the CASA model, which tells the model about the existence of land and water. When logic 1 is present it means it is land. If logic 0 is present it means water. At points where the landmask file shows logic 0, the code runs faster since no calculation needs to done. Because of this, there is a minimal use of the available memory and consequently less computation.

Vegetation.

The vegetation input of the CASA indicates the type of plants growing in a particular area. For example the vegetation can be broadleaf, evergreen, deciduous, grass, wood, cultivated, mixed, barren soil or swamps. The type of vegetation decides the amount of carbon that is released or absorbed. If the vegetation is broadleaf, then at particular seasons there is a net carbon absorbed. During other seasons there is a net carbon release.

Each kind of vegetation has a different NDVI, which in turn regulates the temperature and rainfall. The vegetation decides whether a particular area is a source or sink of carbon.

Parameters Table

The parameters table is one of the main input file of the CASA model. The parameters table file specifies the initial parameter values for CASA variables and constants. The table specifies the size of the input files in rows and columns. All the input files have the same size except when mixing global data and regional data. The table also mentions the latitude and longitude of interest in case of non gridded data, whether the input data sets are yearlong data sets or monthly data sets.

Information about deforestation, carbon dioxide, and all inputs used are citied in the file. The table also describes how the model should give the output, like monthly or yearly or monthly average, vegetation information, year of data fed to the model, etc. The memory constraints, such as dynamic or fixed allocation of the memory, the output and input data format i.e., the byte order, the non available data or bad values, and whether the output should be in binary or NetCDF or non-gridded data format are specified in the parameters file.

Landcover Table

The landcover table, an input to the CASA model specifies the vegetation classes of the data. The table specifies the type of vegetation such as the broad leaf or needle leaf, evergreen or deciduous or mixed, grass or shrubs or moss or lichens, cultivated or non-cultivated, the wood age, the normalized difference vegetation index minimum and NDVI max values.

Soil Table

The soil table is an optional input file that the CASA model uses. The soil table specifies the soil texture, the soil type such as coarse, medium coarse, medium, fine medium, fine, ice and organic, sand, and clay etc. The type of soil depends on the kind of region or area that is of interest.

Other Inputs

Other input files are annual and perennial files, which indicate the perennial agriculture, deforestation files give what fraction of each pixel is deforested, latitude and longitude files that specify the boundary when a non gridded data is being used.

The CASA model includes options to model CO_2 effects, option to model deforestation and agriculture effects, dynamically allocate memory with no limitation on the size of the input data sets, and provide the user with a choice about the output format.

Output Data Sets

Once the CASA model is implemented, the output directories are created. Depending on the user choice of output data type, the output data is either a binary file or a NetCDF file, which is stored in their respective directories and sub directories. If monthly averages are to be outputted, the monthly values are saved as per month; if chosen yearly, the outputs are saved as yearly files.

Figure 3 gives the outputs that a CASA outputs once the model is finished running, while Figure 4 gives the CPOOLS outputs. Soil carbon (SOILC) output of the CASA model for the state of Mississippi is in good agreement with the ground truth.

AET	:	actual evapotranspiration (mm/month or mm/year)
APAR	:	absorbed photosynthetically active radiation (MJ/m^2/month)
CPOOLS	:	carbon pools (gC/m^2)
FAPAR	:	fractions of photosynthetically active radiation
LAI	:	leaf area indices
LEAFFR	:	leaf allocation fractions
NBP	:	net biome productivity (gC/m^2/month or gC/m^2/year)
NEP	:	net ecosystem productivity (gC/m^2/month or gC/m^2/year)
NPP	:	net primary productivity (gC/m^2/month or gC/m^2/year)
NPPMOIST	:	moisture down regulator for npp calculations
NPPTEMP	:	temperature down regulator for npp calculations
PET	:	potential evapotranspiration (mm/month or mm/year)
RESP	:	respiration (gC/m^2/month or gC/m^2/year)
ROOTFR	:	root allocation fraction
SOILC	:	total soil carbon or nitrogen (gC/m^2)
STEMFR	:	stem allocation fraction

Figure 3 CASA Outputs

CPOOL_01	:	leaf carbon pool (gC/m^2)
CPOOL_02	:	wood carbon pool (gC/m^2)
CPOOL_03	:	root carbon pool (gC/m^2)
CPOOL_04	:	surface metabolic carbon pool (gC/m^2)
CPOOL_05	:	surface structure carbon pool (gC/m^2)
CPOOL_06	:	soil metabolic carbon pool (gC/m^2)
CPOOL_07	:	soil structure carbon pool (gC/m^2)
CPOOL_08	:	coarse woody debris carbon pool (gC/m^2)
CPOOL_09	:	surface microbian carbon pool (gC/m^2)
CPOOL_10	:	soil microbian carbon pool (gC/m^2)
CPOOL_11	:	slow carbon pool (gC/m^2)
CPOOL_12	:	passive carbon pool (gC/m^2)
CPOOL_13	:	black carbon pool (gC/m^2)

Figure 4 CASA Model's CPOOL Outputs

Data Extraction

NDVI Data Extraction

The following steps are involved in converting a raw HDF NDVI input into a formatted input for running the CASA model.

Firstly, extract the NDVI data subset from the parent HDF file using the MODIS Re-projection Tool (MRT). Save the extracted raw NDVI data as a GeoTIFF file; this is a convenient file extension that can be easily processed in ArcMap. Figure 5 is a screen shot of the NDVI data extraction using the MRT tool in windows operating system. The MRT tool can be downloaded from the Land Processes Distributed Active Archive Center (LP DAAC) [26].

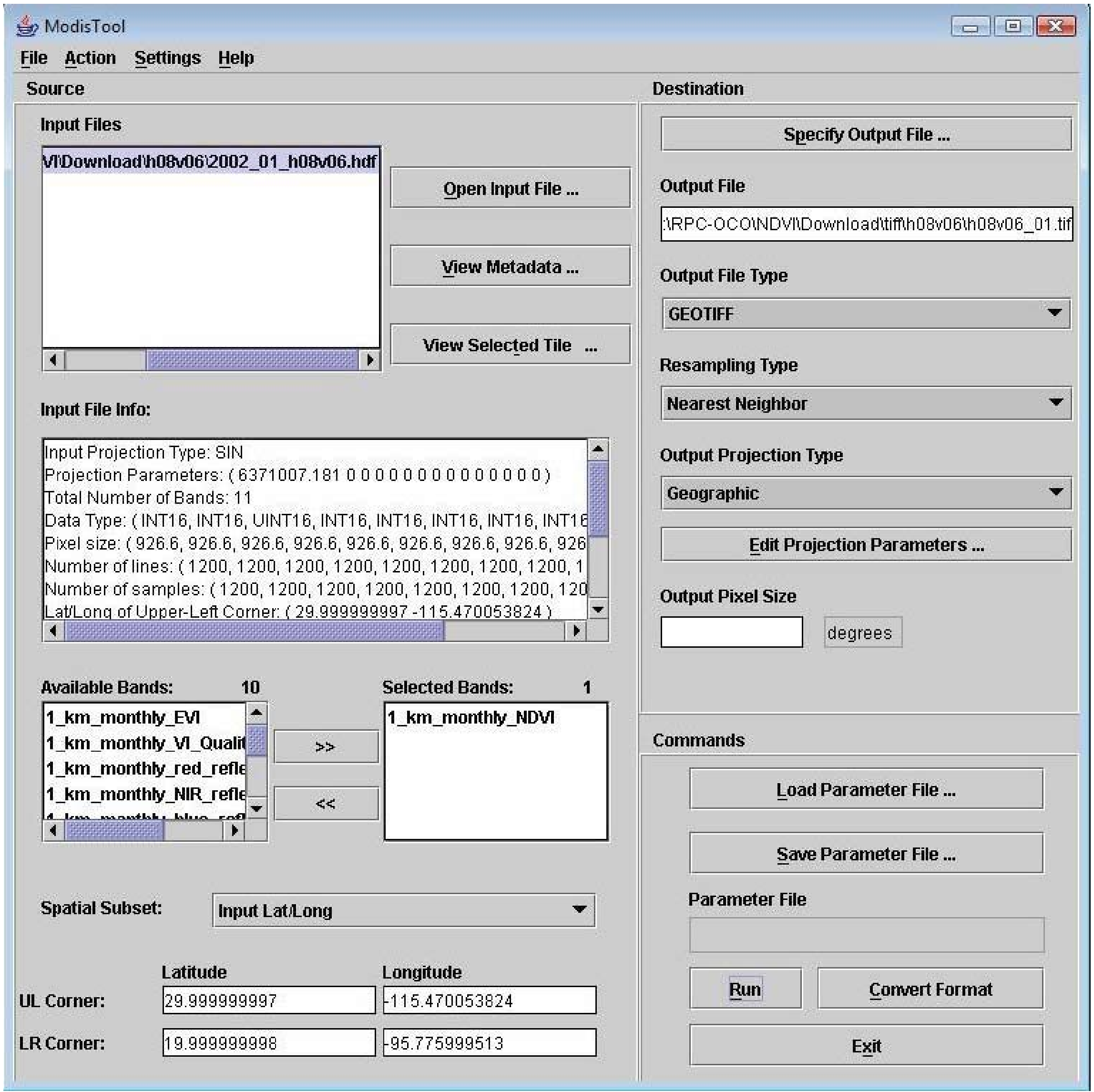

Figure 5 MODIS Re-projection Tool [26]

The downloaded data is in the shape of a tile. A decision is made as to which states falls into which tile. This decision can be done by looking at the latitudes and longitudes of the tiles and individual states. If in case a state falls in more than one tile, then all the tiles in which the state falls are mosaicked. Figure 6 gives the tile number, which help in determining the tiles in which the states fall. Further individual tiles and state latitudes and longitudes are given in the Tables 3 and 4 in appendix A. Choice of the tile, as to which states falls in what tile is shown in Table 5 in appendix A.

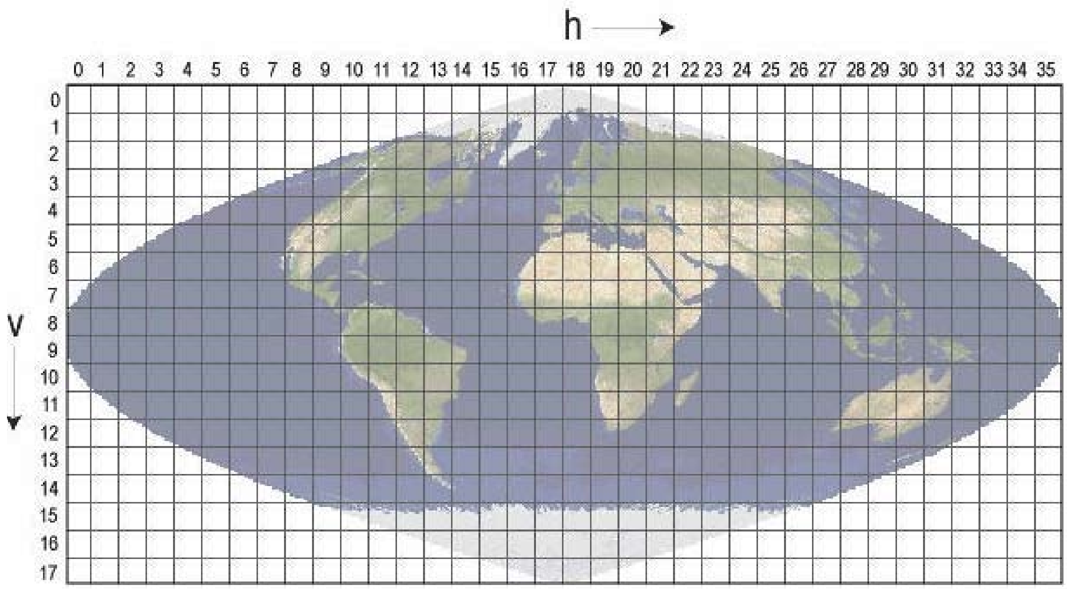

Figure 6 Sinusoidal Earth [27]

Next, the co-ordinate system of the NDVI data is changed from GCS_WGS_1984 (Geographic co-ordinate system, World Geodetic Survey 1984) to GCS_North_American_1983. The GCS_North_American_1983 is the co-ordinate system of the soil data, which is used as mask.

Figure 7 shows how it is done:

Run ArcMap→ Arc Toolbox → Data Management Tools → Projections and transformations → Raster→ Double click "Project Raster" → Give the path of "Input raster" and "Output raster" → "Output coordinate system" → select → Geographic coordinate system → North America → North American Datum 1983.prj → Add → OK.

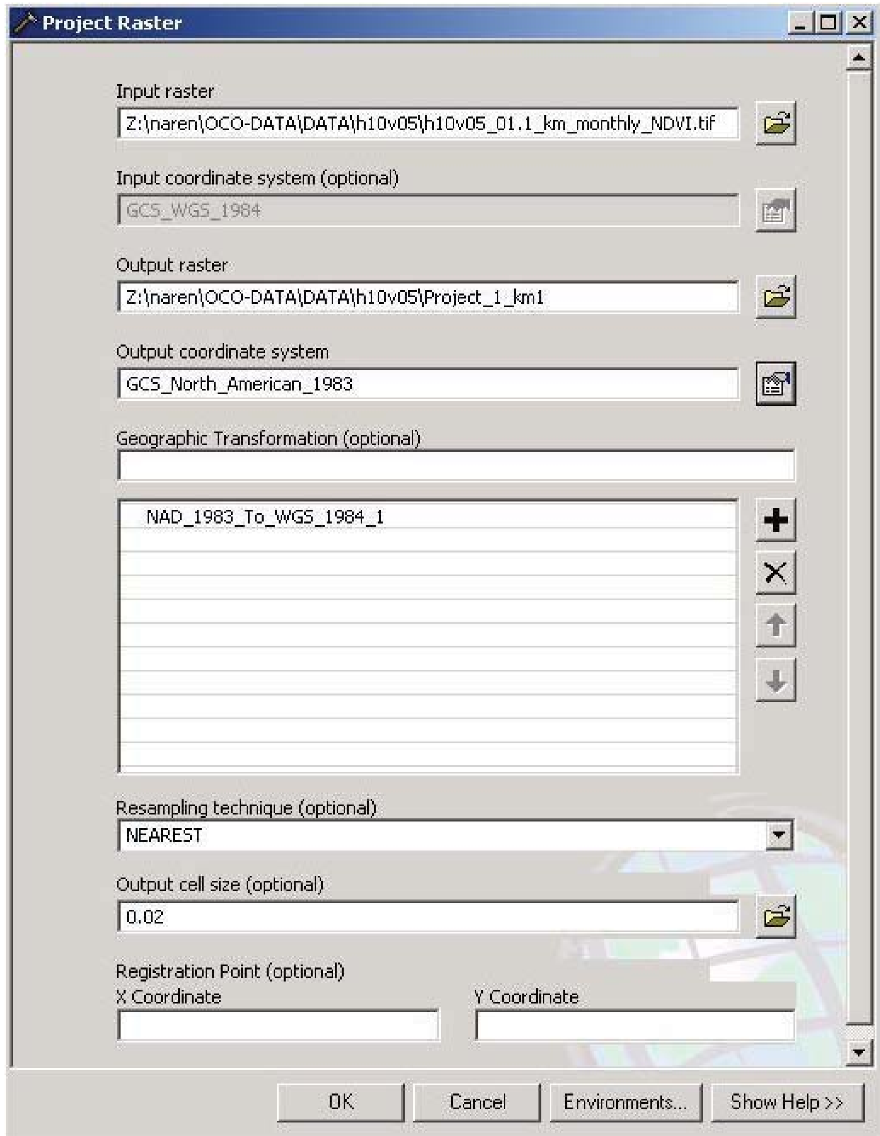

Figure 7 ArcMap Changing the Co-ordinate System [28]

The output of the ArcMap after changing the coordinate system is shown in Figure 8, where the tile h10v05 is displayed.

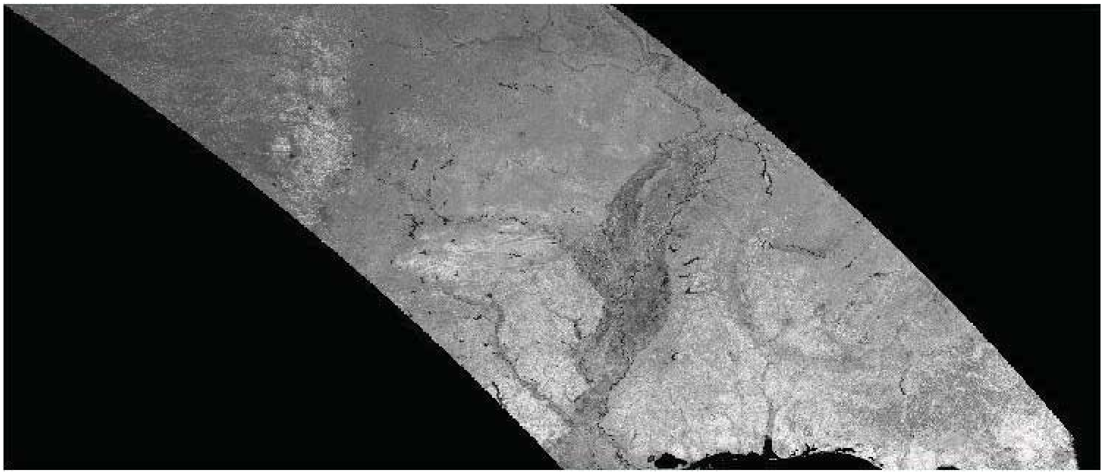

Figure 8 Output of Project Raster Tool of ArcMap, Tile h10v05

The original value of the raw NDVI data is in the range of -3000 to 10,000. The negative sign indicates that no data is present at the point. Positive values of NDVI are all that is needed, hence positive values needed to be extracted.

Figure 9 and 10 shows, how the process is carried out and the output ArcMap after the positive values are extracted (tile displayed is h10v05).

ArcMap→ ArcToolbox → Spatial Analyst Tools → Extraction → Extract by Attributes → Values > 0 →OK → OK

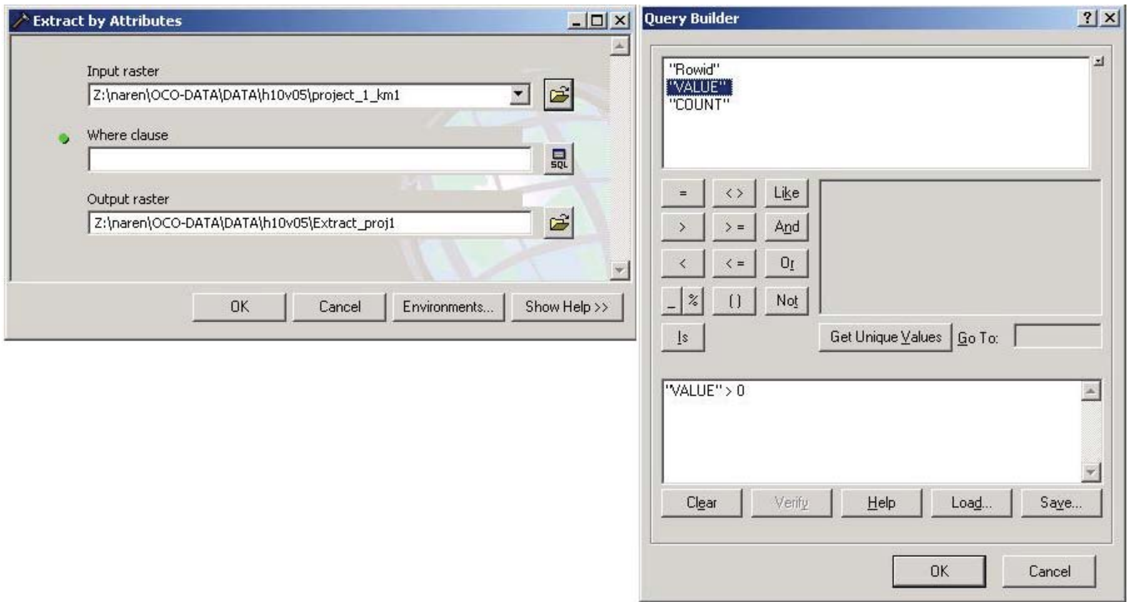

Figure 9 ArcMap Extraction of Positive Values

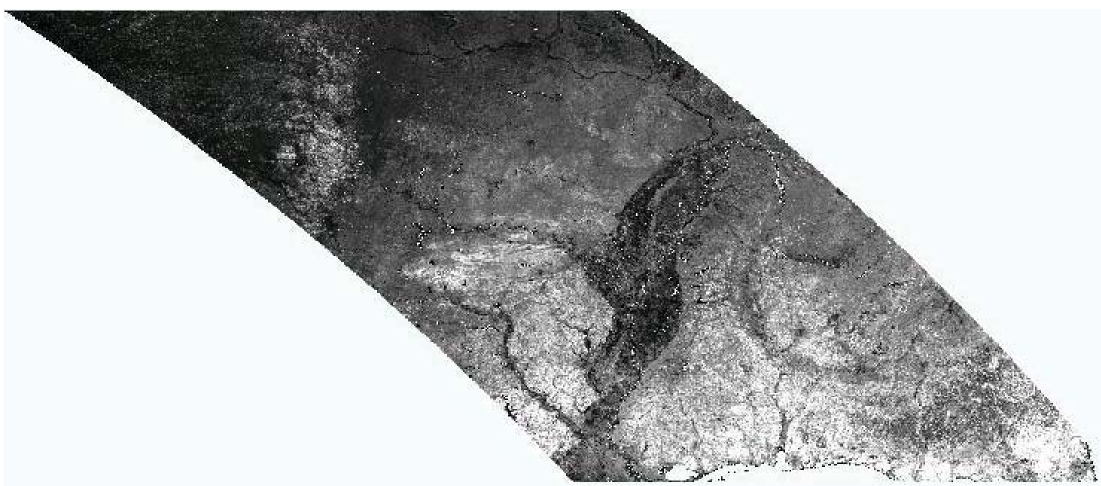

Figure 10 ArcMap Output after Positive Value Extraction

The extracted positive values of the NDVI are converted to floating point numbers as shown in Figure 11 below, while Figure 12 is the output after converting to floating point. ArcMap → Add the NDVI raster file with positive value → Spatial

Analyst → Raster Calculator → Float → select layers → OK. A floating layer will be created.

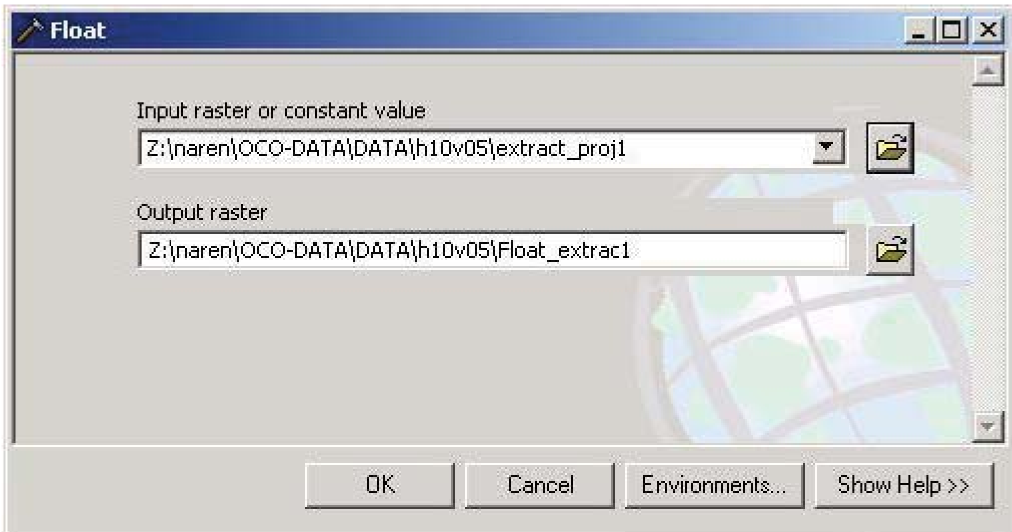

Figure 11 ArcMap Conversion to Float

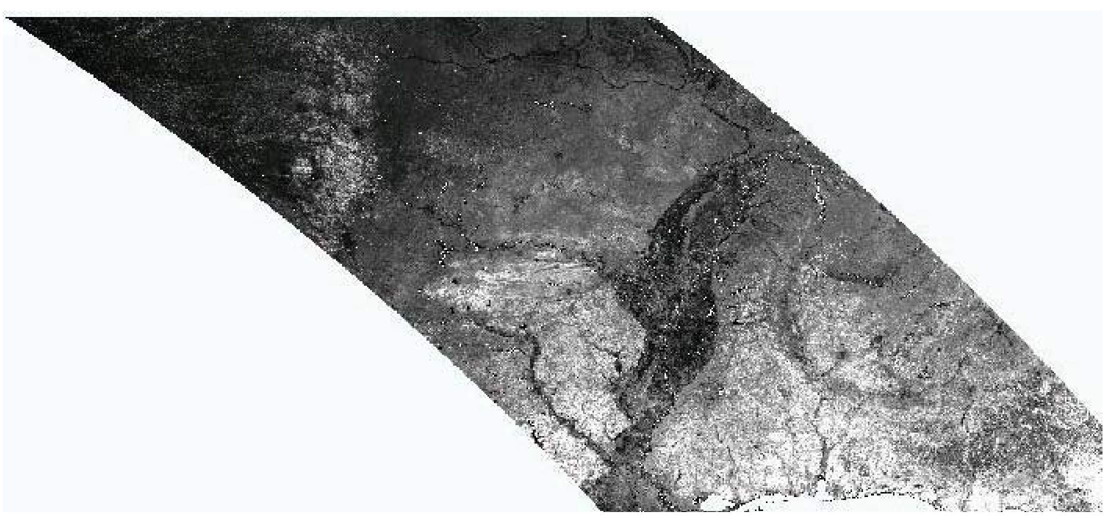

Figure 12 ArcMap Output after Converting to Float

The value range of NDVI data is changed to 0 - 1 corresponding to CASA model which is shown in Figure 13 and its respective output is shown in Figure 14. .

ArcMap →Spatial Analyst →Raster Calculator →"[Calculation] / 10000" → Evaluate. A layer named "Calculation2" will be created.

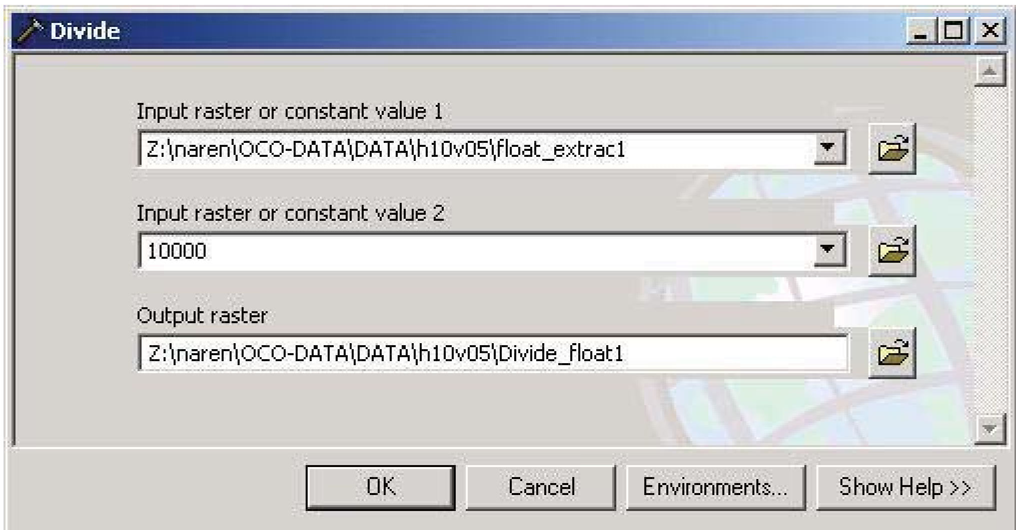

Figure 13 ArcMap Division by a Constant Value

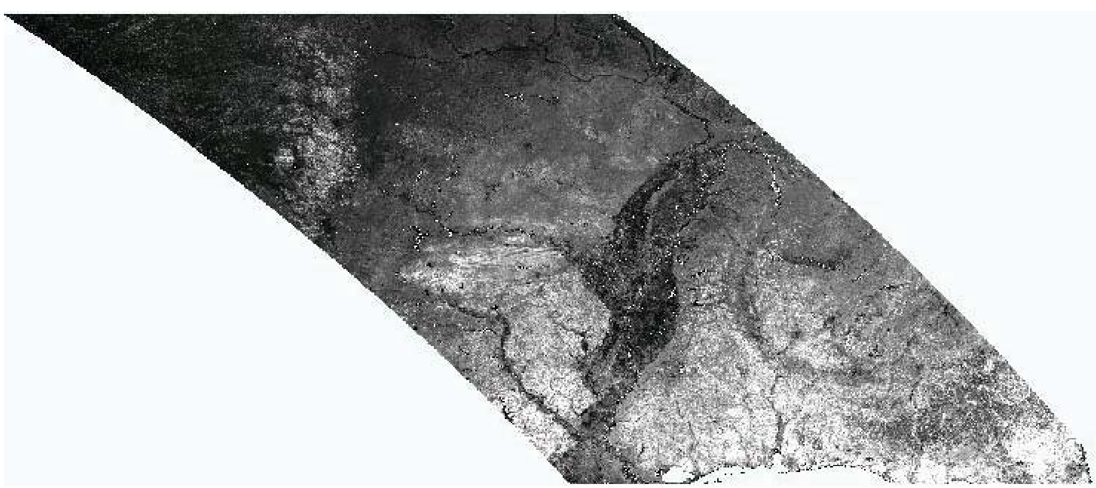

Figure 14 ArcMap Output after Division

Until here the data processed is the data in the form of tiles for example h08v06, which indicates it is a tile with horizontal number 8 and vertical tile number 6.

Extraction of individual state data can be done by applying the corresponding soil data as mask. As can be seen from Figure 15, the mask data is the gsmsoilmu_a_ms.shp, which is the soil shape file for the state of Mississippi. Figure 16 gives the output after extracting by the shape file. The outputs thereafter this stage can be viewed as individual states rather than tiles.

ArcMap → ArcToolbox → Spatial Analyst Tools → Extraction → Extract by Mask.

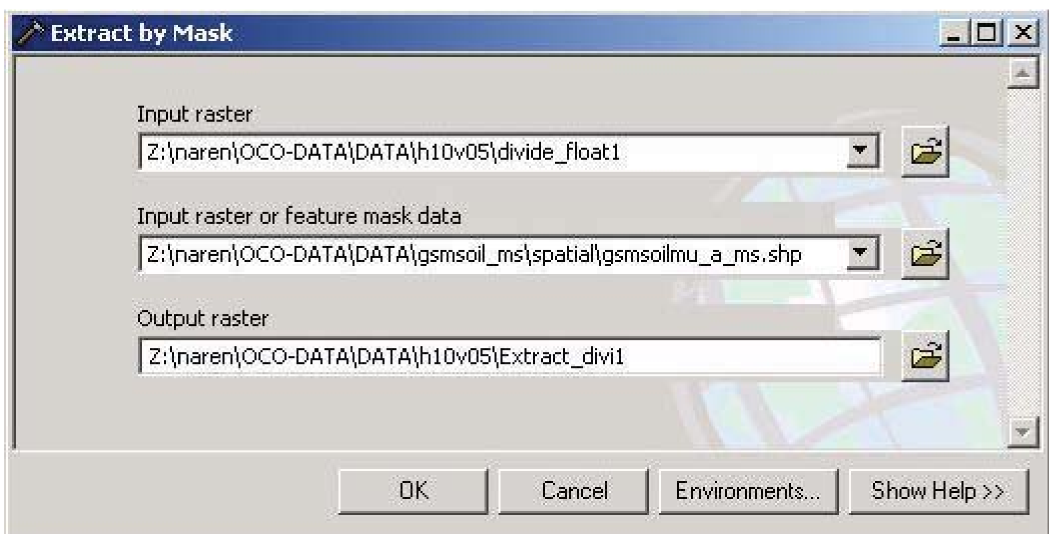

Figure 15 ArcMap Extract by MASK (Soil Data Shape File)

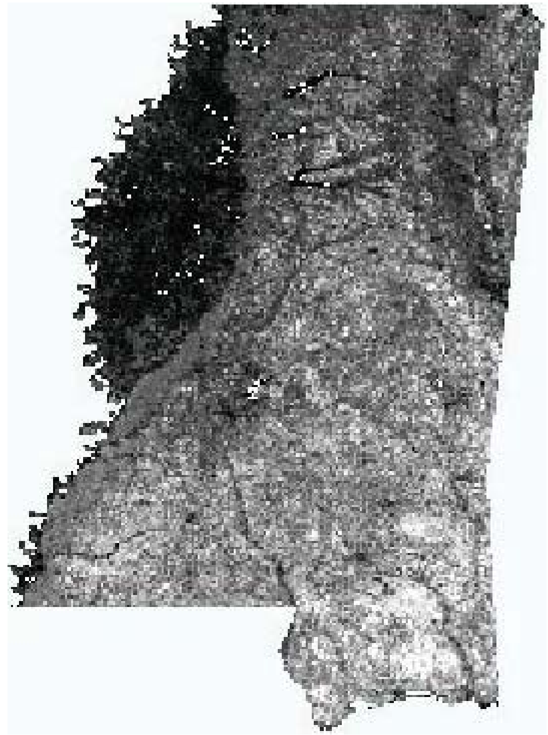

Figure 16 ArcMap Output after Extraction by Mask

The output at this stage is individual, state monthly data. The data has to be further formatted such that the NDVI data is of the same size as all other inputs of that state. The size of the individual state depends on the area of the state; the larger the state the bigger is the size of the data matrix.

The output raster image is further formatted for a cell size of 0.04. The cell size of 0.04 is the 8km resolution, as shown in Figure 17, which same as the resolution of the soil data.

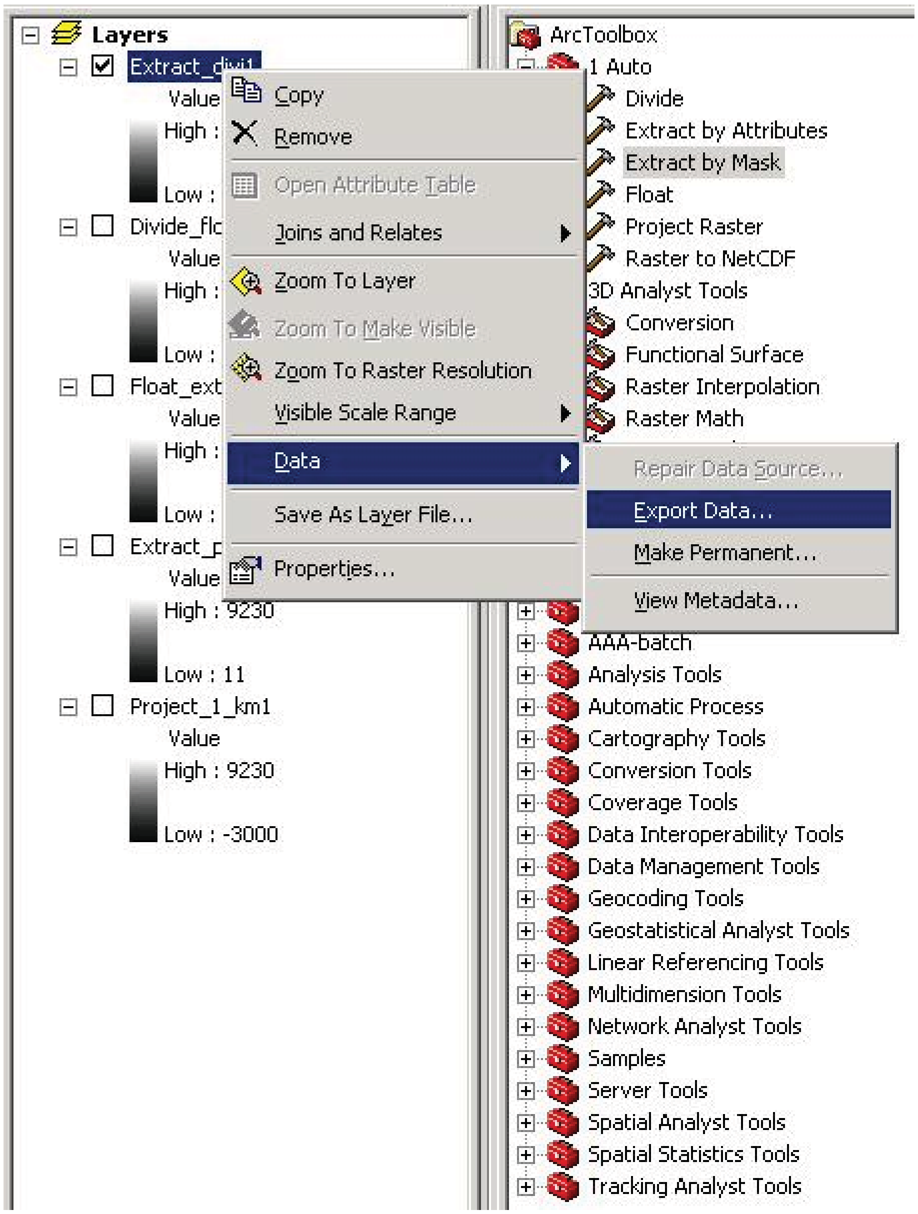

Figure 17 Different Layers after Processing

In ArcMap, as depicted in Figure 18, right click “NDVI_01” → Data→ Export data→ Square; cell size 0.04→ Give the location path and file name “ndvi_ms_01.tif”; Format “TIFF” → Save. Figure 19 is the final image file of the data. Further files are dealt only NetCDF format.

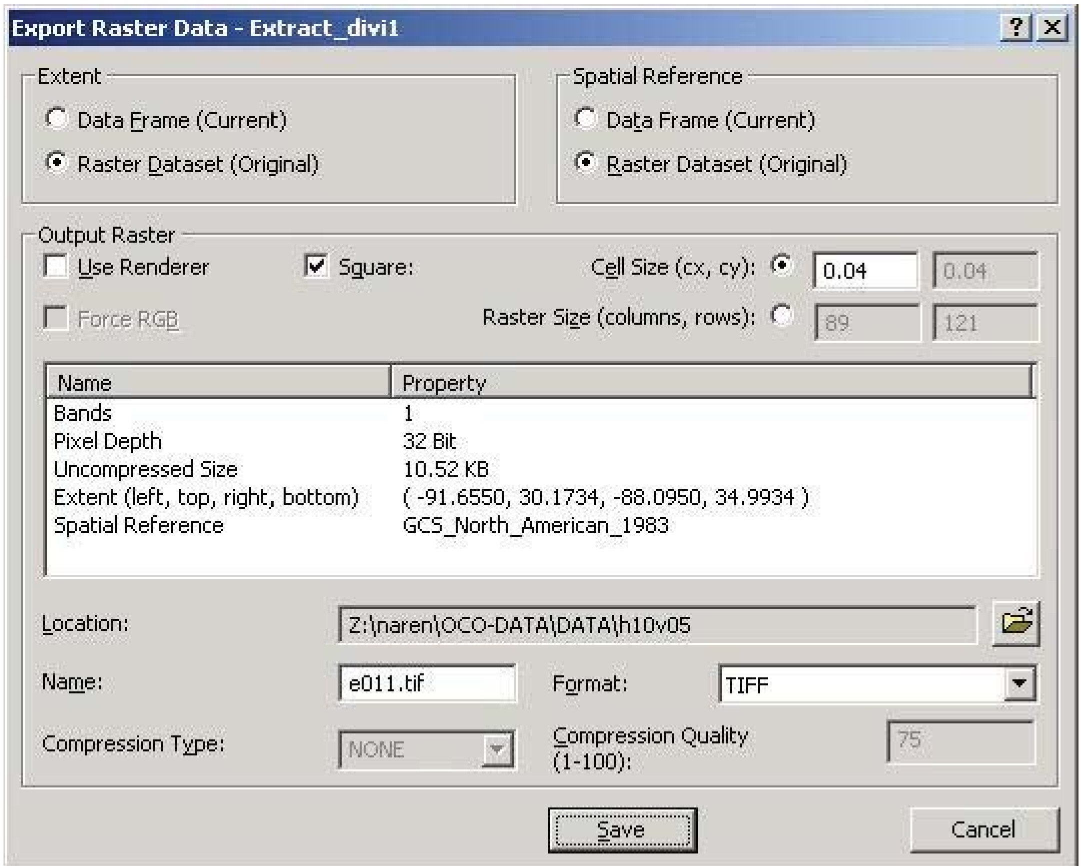

Figure 18 ArcMap Export to Square Cell Size 0.04

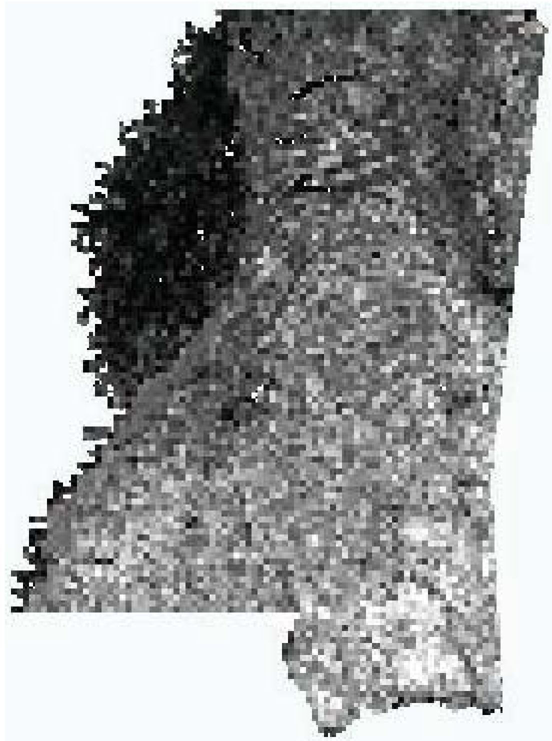

Figure 19 ArcMap Output after Export to Square Cell Size 0.04

The corresponding netcdf file for the month of January is produced. This can be observed from the Figure 20 below

In ArcMap, ArcToolbox → Multidimension Tools → Raster to NetCDF. The X dimension is chosen to be longitude, while the Y dimension is chosen to be latitude.

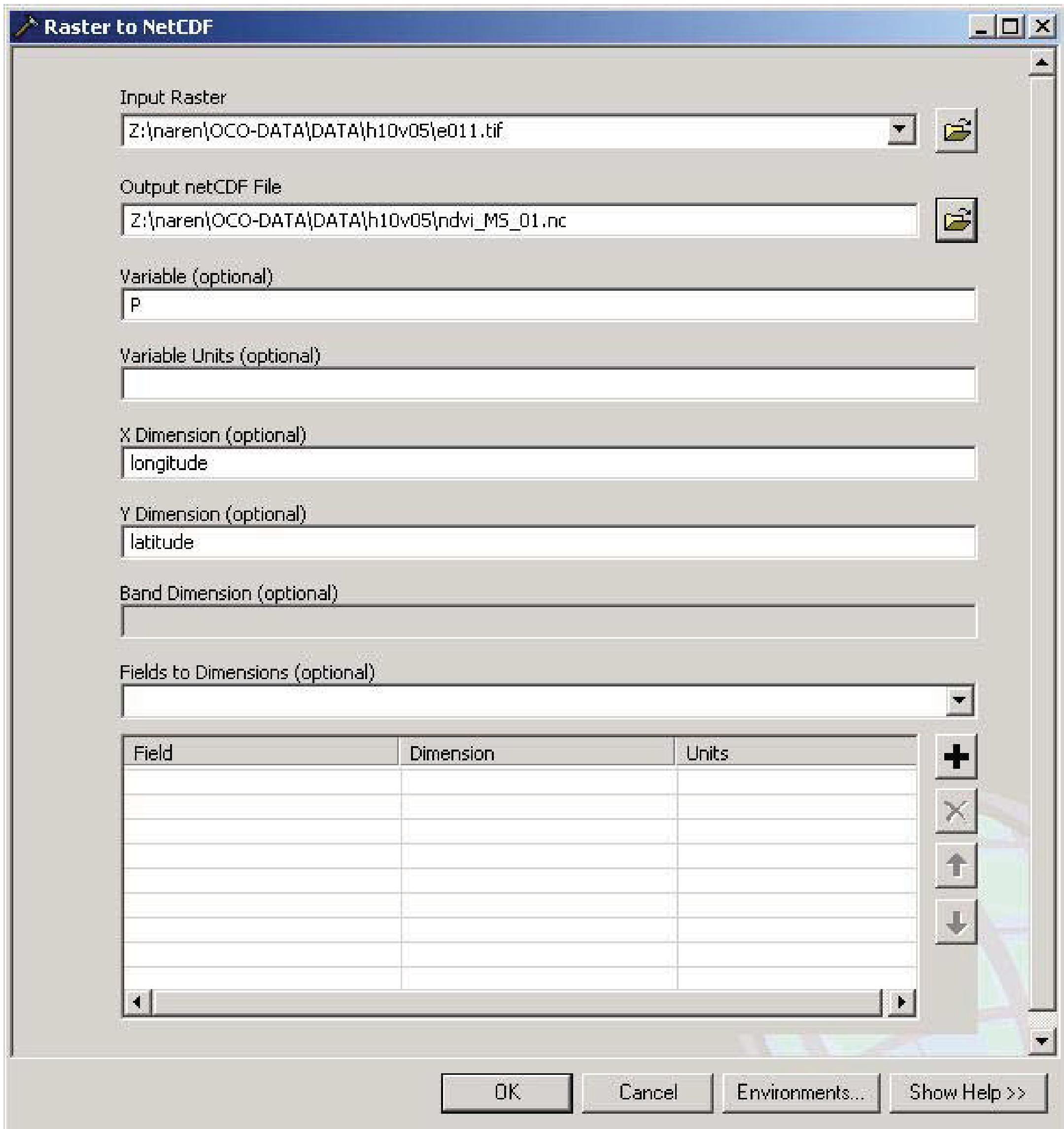

Figure 20 ArcMap Raster to NetCDF Conversion

Further bash scripts are written to concatenate the twelve months individually formatted data to form a single NetCDF file. This single NetCDF file is the end 'NDVI' input for the CASA model. The bash scripts used can be found in the appendix B.

Precipitation Data Extraction

While processing an ASCII or a text file make sure that the text file contains a proper header, like the one shown in Figure 21 below, and that the 'xllcorner' and 'yllcorner' are right.

```
ncols 1405
nrows 621
xllcorner -125.020833
yllcorner 24.062500
cellsize 0.04
NODATA_value -9999
```

Figure 21 Sample Header of Precipitation Data

The ASCII or text file is converted into a raster image by doing:

In ArcMap → AcrToolbox → Conversion Tools → To Raster → ASCII to Raster as in Figure 22.

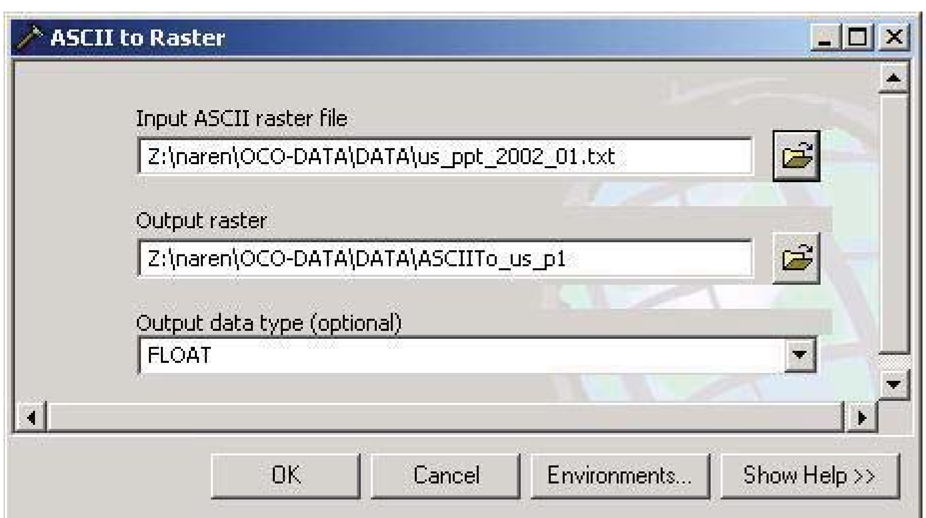

Figure 22 ArcMap ASCII to Raster Conversion

The output after the ASCII file is converted to a raster file is the image of North America as shown in Figure 23 below

Figure 23 ArcMap Output after ASCII to Raster Conversion

The co-ordinate system of the raster image is defined by running ArcMap→ Arc Toolbox → Data Management Tools → Projections and transformations → Raster→ Double click “Project Raster” → Give the path of “Input raster” and “Output raster” → “Output coordinate system” → select → Geographic Coordinate systems → North America → North American Datum 1983.prj → Add → OK. The output cell size is defined to be 0.04 so that the resolution is the same as the resolution of the soil data.

To make sure that the data values are between 0 – 1000 corresponding to the CASA model the data is divided by factor of 1000 as depicted in Figure 24. In ArcMap → Spatial Analyst → Raster Calculator → “ ppt _ms_01 / 1000” → Evaluate.

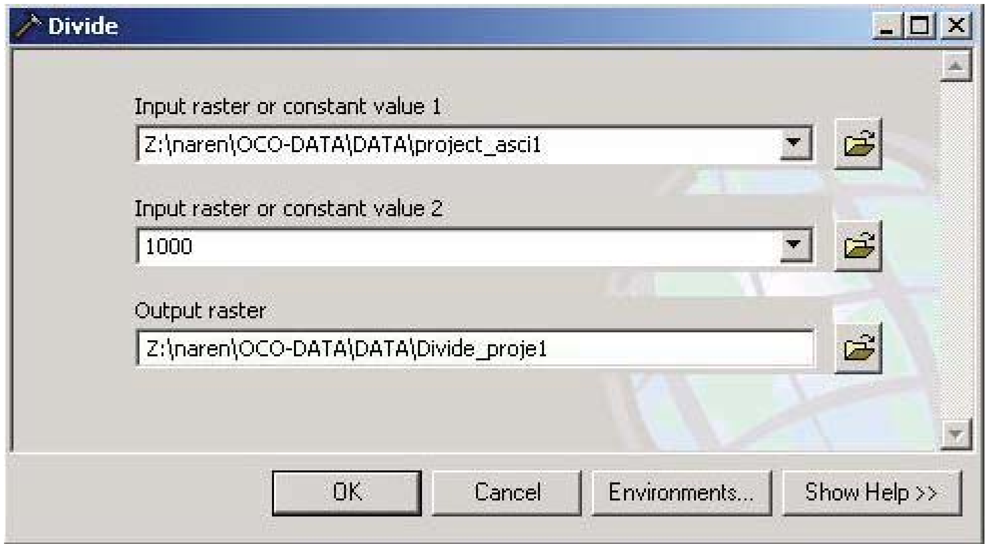

Figure 24 ArcMap Division by Constant Value

Now individual state monthly precipitation data is extracted by using the corresponding soil data as mask. ArcMap → ArcToolbox → Spatial Analyst Tools → Extraction → Extract by Mask.

Figures 25 and 26 display the process during the extraction with soil data as mask and the output of the extracted layer overlapped on the input file respectively.

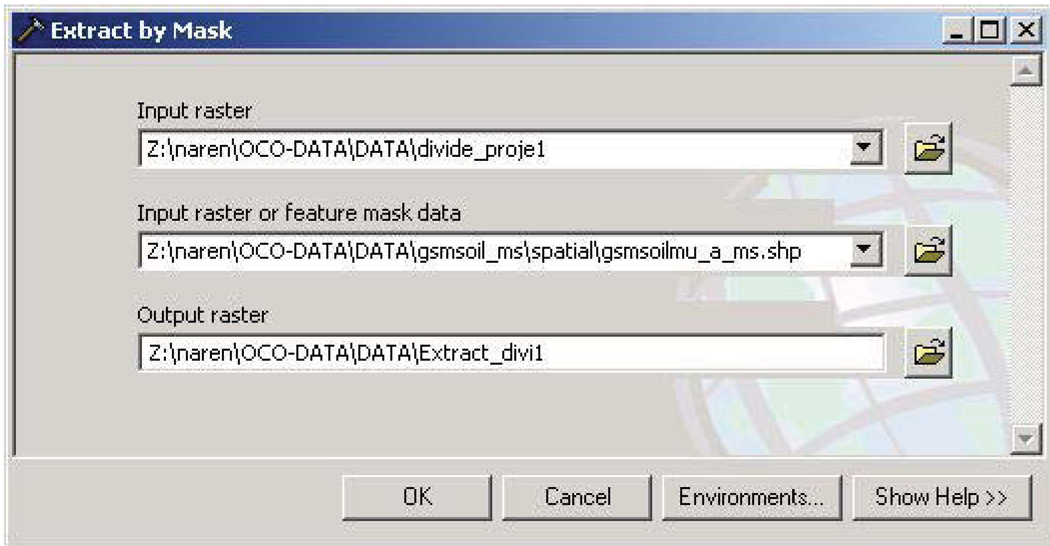

Figure 25 ArcMap Extraction by Mask

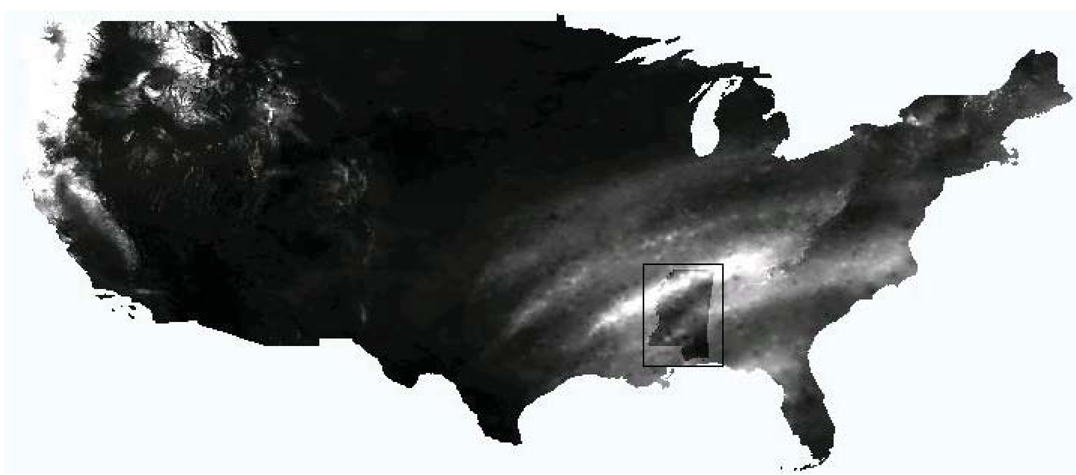

Figure 26 O/P after Extraction by Mask (Inside Extracted MS State)

The new extracted layer as in Fig 26 and the new image will be that of the mask applied. The output in (d) is further processed so that it has the same resolution, i.e., of the same matrix size as that of the rest of the input data sets like NDVI and SOIL.

In ArcMap, right click the new created layer → Data→ Export data→ Square; cell size 0.04→ Give the location path and file name “ppt_ms_01.tif”; Format “TIFF” → Save

The corresponding netcdf file for the month is produced by In ArcMap, ArcToolbox → Multidimension Tools → Raster to NetCDF. The X dimension is chosen for longitude while the Y dimension is chosen for latitude.

Further bash scripts concatenate the twelve months; individually formatted data to form a single netcdf file. This single netcdf file is the end ‘ PPT ’ input of the CASA model.

Soil Data Extraction

Soil data, which is an important input file is used as a mask file in the extraction of remaining data sets. The final output soil file is a shape file, which is further converted into NetCDF file to form as an input to the CASA model. Figure 27 through 31 shows the steps involved in downloading the correct soil data sets.

Figure 27   Soil Data Extraction Step 1

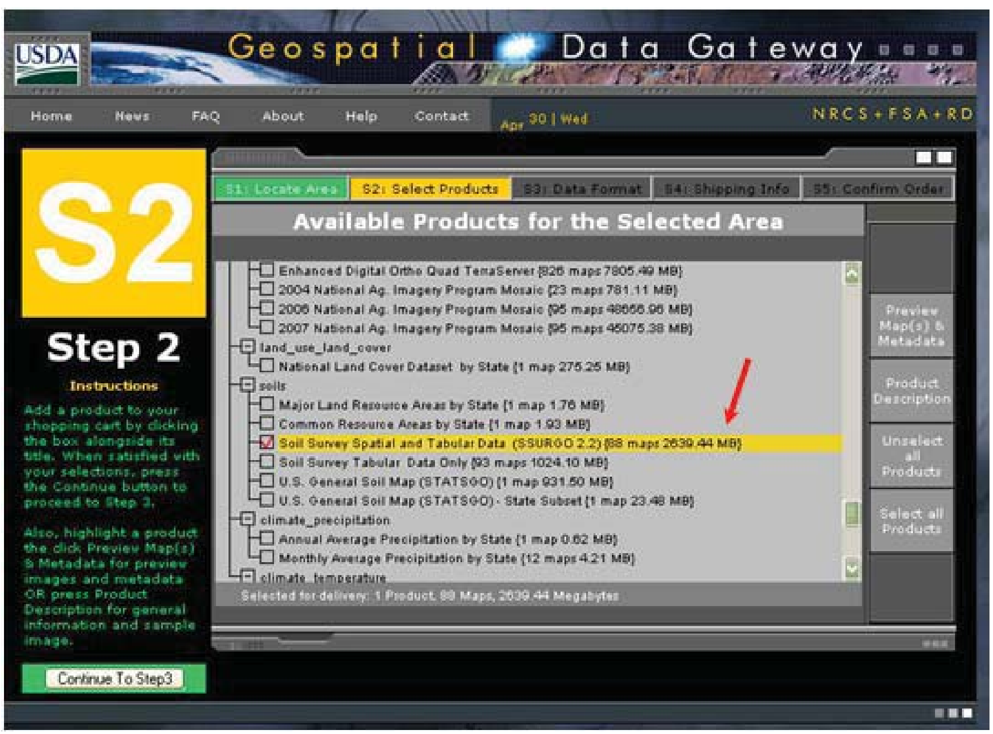

Figure 28   Soil Data Extraction Step 2

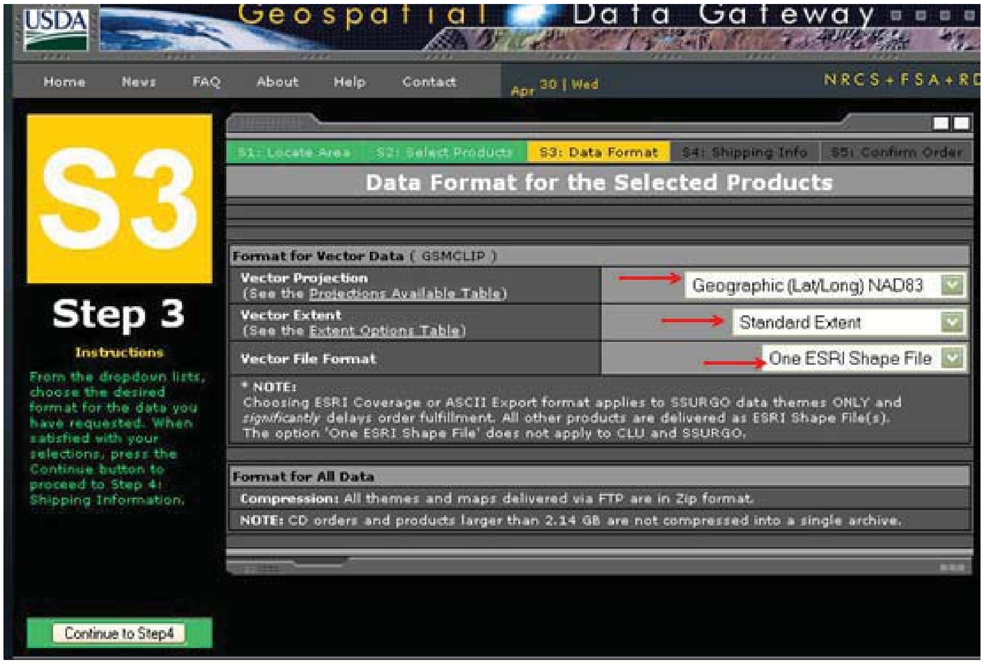

Figure 29 Soil Data Extraction Step 3

Figure 30 Soil Data Extraction Step 4

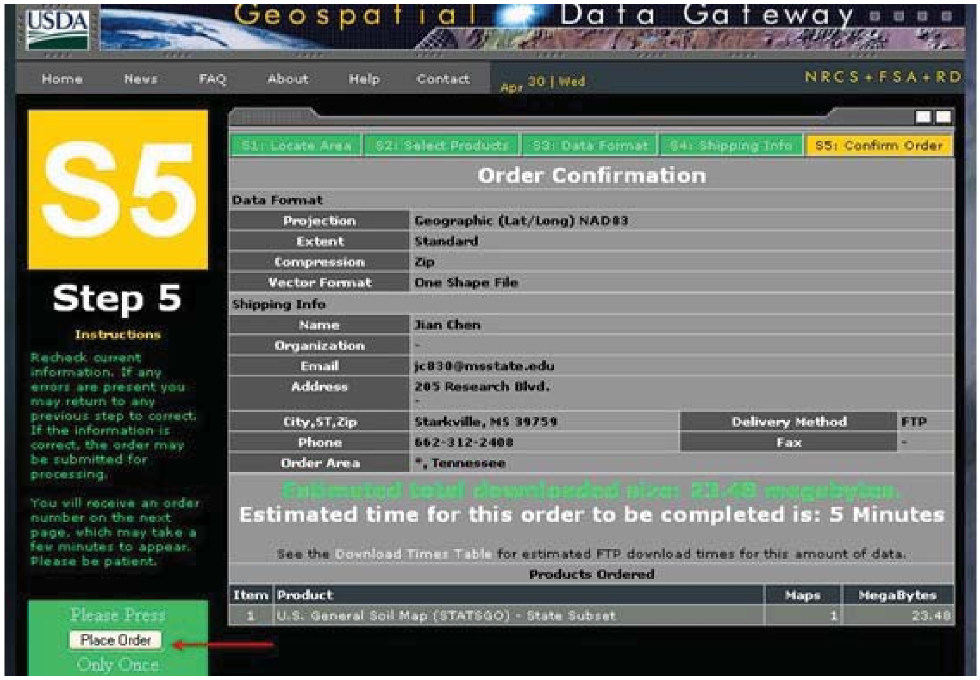

Figure 31 Soil Data Extraction Step 5

Downloaded the template database by going to Soil Data Mart and download the “Template database”. Download the general US template database in one of the three versions (Access 2002, 2000, 97) depending on the Microsoft Access version in the computer used. The downloaded file name is “soil_db_US_2000”.

Tabular data is imported into Access Template Database by downloading the soil data from your e-mail box and unzip it. The next step is to import tabular data into Access Template Database by following the instruction of “SSURGO Data Packaging and Use”.There is a folder for each county; and five items are included in this folder as follows: spatial, tabular, readme.txt, soil_metadata_ms001.txt, soil_metadata_ms001.xml.

The template database (soil_db_US_2000.mdb) is copied into this folder. Double click soil_db_US_2000.mdb; input the path of tabular data. Then, the importation will be finished automatically.

First, install ArcInfo workstation 9.2 and ArcDesktop 9.2 completely but without any service package. You need to uninstall Microsoft Internet Explorer 7 if you have it in the computer because ArcGIS is not compatible with IE 7. Second, download and install Soil Data Viewer 5.2 (SDV) from http://soildataviewer.nrcs.usda.gov/download52.aspx

Produced a thematic map as shown Run Arcmap→ Add data (e.g. in folder “spatial” named “soilmu_a_ms001.shp”); and SDV is run under Arcmap (Click gray area in Arcmap and check “soil data viewer tools” before run SDV under Arcmap) open the same maplayer (soilmu_a_ms001.shp) and the corresponding database (soildb_US_2000.mdb) in SDV → In “attribute folders”, select “Soil Physical Properties”, then select “Surface texture” → In “Report Options” “Select All” → Map. A thematic map named “Surface Texture” appeared in Arcmap

Produced raster file by following Arcmap, ArcToolbox—Convesion Tools—To rater—Feature to raster, as shown in Figure 32. When the conversion completed, a raster file named “Feature_Sd150” was created and appeared in Arcmap.

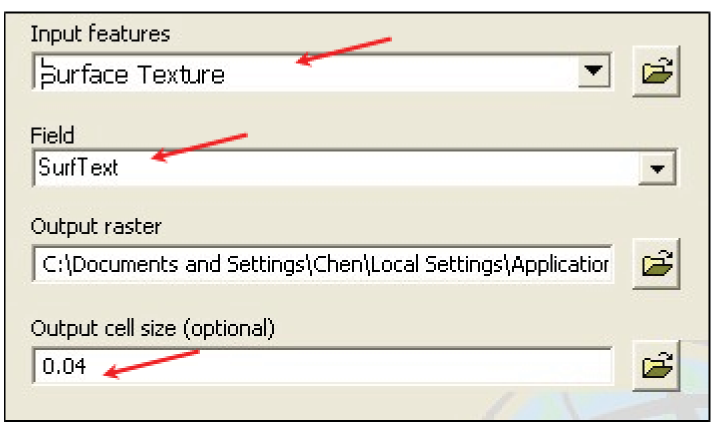

Figure 32 Producing a Raster File

The data is reclassifying by following steps in Arcmap, drag the created raster file ("Feature_Sd150") to the top layer and select it→ Spatial Analyst→ Reclassify Change new values to the value that corresponds to soil classes in the CASA model and then group the same values by pressing "Ctrl" key→ OK. Then a new raster file named "Reclass of Feature_Sd150" was created.

Figure 33 shows the process of reclassifying of the soil type along with their values while Table 1 is the whole soil classification.

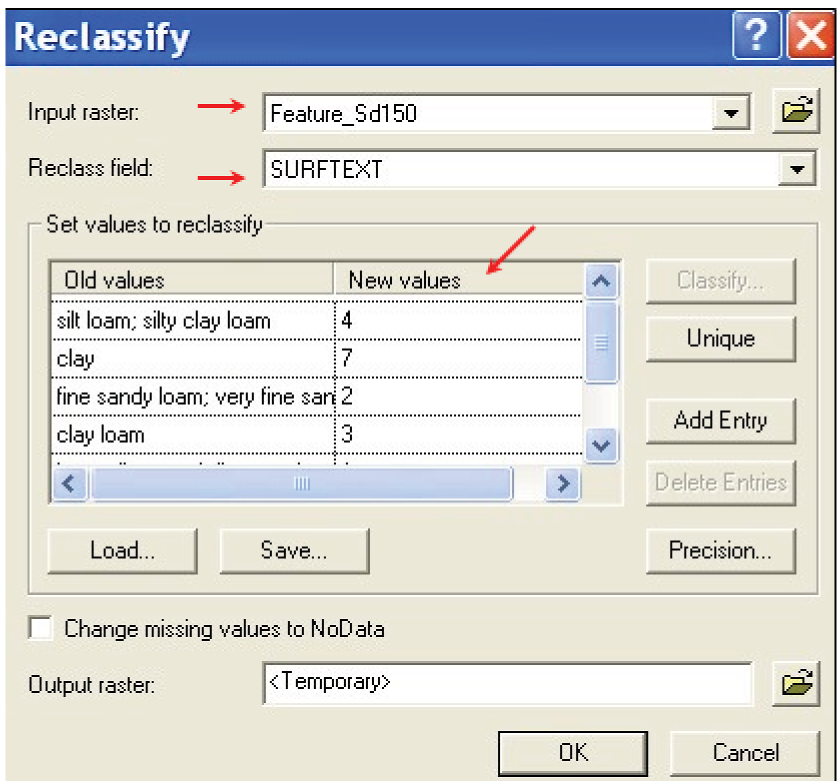

Figure 33 ArcMap Reclassifying of Soil Data

Table 1 Tabular Form of Reclassifying Soil Data

SA Class	CASA Soil name	Sand	Clay	Soil names
1	Coarse	0.8	0.1	Coarse sand; Sand; Fine sand; Very fine sand; Loamy coarse sand; Loamy sand; Loamy fine sand; Loamy very fine sand; Silt
2	Medium/coarse	0.5	0.25	Coarse sandy loam; Sandy loam; Fine sandy loam; Very fine sandy loam; Sandy clay loam; Loam
3	Medium	0.4	0.35	Clay loam; SANDY CLAY
4	Fine/medium	0.1	0.2	Silt loam; Silty clay loam
5	Fine	0.25	0.6	**Silty clay**
6	Ice	-999	-999	
7	Organic	0.2	0.7	Clay

The output raster is created by doing in ArcMap, right click “Reclass of ms001 - ms001” → Data→ Export data→ Square; cell size 0.04→ Give the location path and file name “soil_ms001.tif”; Format “TIFF” → Save as illustrated in Figure 34.

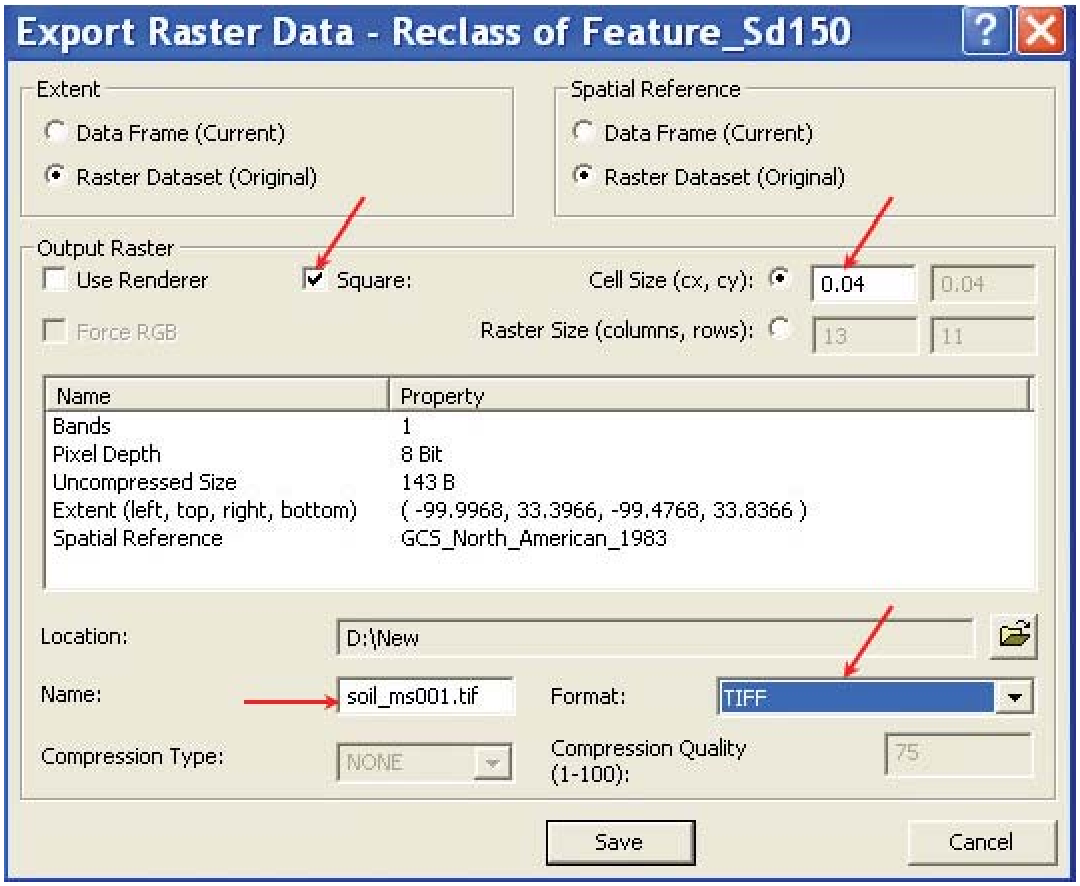

Figure 34 ArcMap Exporting of Soil Data

The state soil data is produced by repeating steps from page 45 to 48 starting by producing a thematic map, to get all of the county soil data in a state. In Arcmap, ArcToolbox → Data Management Tools → Rasters → Mosaic To New Raster. By this tool, shown in Figure 35, you can composite all of the counties together into one file.

Input all of county soil data; give the output location, raster dataset name and coordinate system and make cell size 0.04.

In the “Coordinate system for raster (optional)”—Select-- Geographic Coordinate Systems-- North America-- North American Datum 1983.prj—Add—OK.

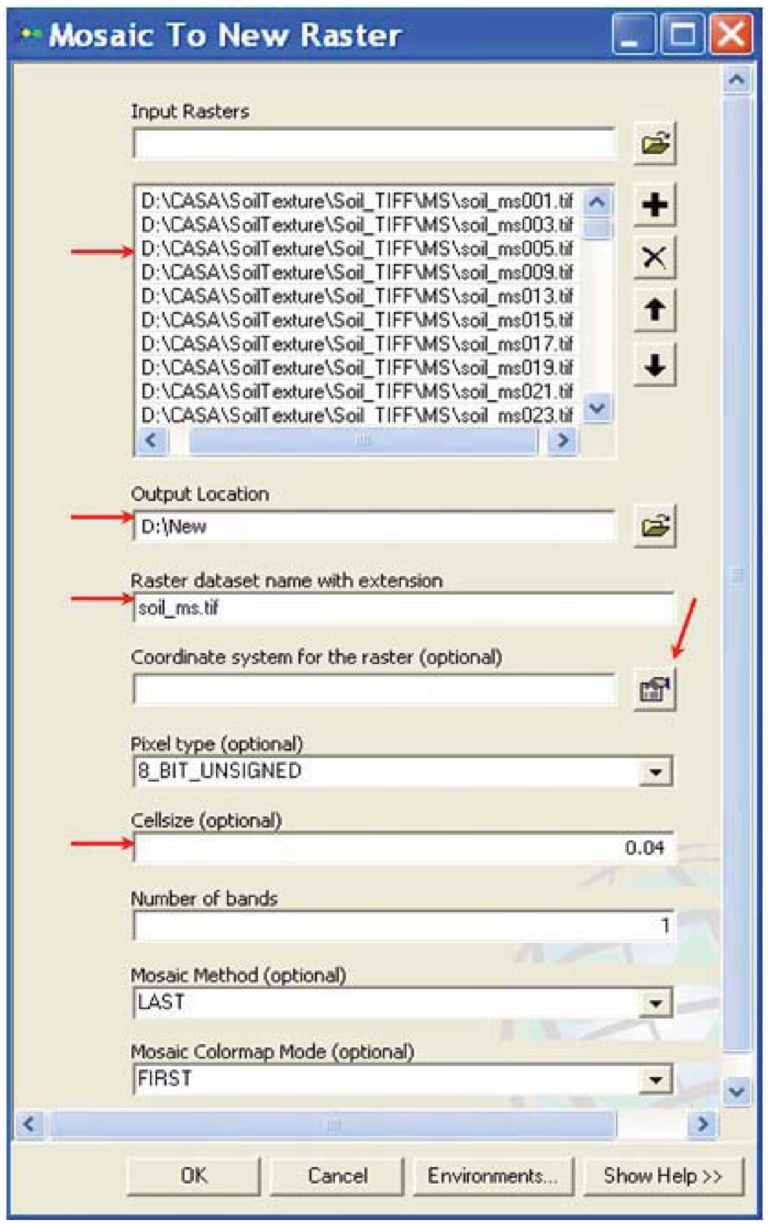

Figure 35 ArcMap Mosaicking

NetCDF file is produced by doing the following in ArcMap, ArcToolbox → Mutidimension Tools → Raster to NetCDF. When doing a regional scale study, the resolution of all the inputs must match. For county level study, county soil data needs to be formatted one by one. Then, all the formatted county data of a state are combined together to get a state data. This combined state data has high resolution. For a state level study, one state data can be formed directly. This state level data has relatively low resolution compared to the county level data.

Temperature Data Extraction

The min and the max temperatures data sets are downloaded from the Parameter-elevation Regressions on Independent Slopes Mode (PRISM) website. Bash scripts are written for a batch processing. The bash script takes the January or the 1st file from the Min folder and similar January or 1st file from the Max folder and averages the value, as in matrix operations and stores in another folder Average as a text file.

Further header is added to the average temperature obtained in step (a) such as the number of rows and columns, the latitude and longitude corners, cell size, and no data value. These header values are very important for further processing in ArcMap. The data size can be found by importing the text file into Matlab and the latitude and longitude corners can be found by plotting the data as image in Matlab.

Once the header is added to the text file, it is initially converted into a raster image using the process ASCII to Raster conversion. From here on, the process is repeated same as done for the precipitation data.

Solar Data Extraction

Solar radiation data is downloaded from the NLDAS. The data size is of 464 x 224 for the whole of North America. The downloaded data is in the form of a text file, and necessary header is added to the data file as explained in the Air Temperature section (b) on page 37. Steps on page 37 - page 41 are repeated for all the 12 files, which correspond to the 12 monthly files.

Scripts Written

Scripts help in running a process, in a much faster way and are an effective way of getting a job done. Bash and Perl scripts are the simplest scripts, where one line handlers can do the whole job. Running similar processes multiple numbers of times or running a process recursively, formatting can be done using the scripts.

Scripts are written to convert the HDF files into netcdf files, i.e., extract the required data set from the HDF file then format and convert them to netcdf file format. This is done for the twelve data sets which are later concatenated to form a single input file of CASA. Further scripts are also written to break the interested CASA outputs, the yearly data sets into monthly data sets, which are used as inputs to the Weather research Forecasting model coupled with Chemistry (WRF-Chem). The yearly data sets are converted into a text format in order to plot graphs to know the pattern of the outputs.

Softwares and Tools

Software helps in more interaction between the data and the user. It also helps in analyzing data sets in an effective way by displaying them as images. The various kinds of software that are used during the process, i.e., preparing/formatting the raw data sets for CASA and, viewing and analyzing the output data, are the MRT (MODIS re-projection tool), ArcMap and ArcInfo, Matlab, Mexnc, Snctools [29], NcView and ncBrowse, MS-Excel.

The MRT Win tool is used to convert the HDF files into GeoTIFF files. Since the HDF file has large amounts of data sets, the MRT Win tool helps in extracting a particular band of data that is of interest and converting it to GeoTIFF.

The ArcMap tool is used in formatting and converting the raw data as inputs for the CASA model. This tool is the major one used in the process of forming the inputs to the CASA. In ArcMap, the data is processed for a new geographic coordinate system, and then the needed data is extracted from the original data sets and converted to NetCDF format.

The Matlab together with Mexnc and SNC tools helps in viewing the data in Matlab. Mexnc is a mex file interface for netcdf files with that of Matlab. Mexnc is a mex file that sits in between the Matlab and NetCDF library files while the snc tool is a set of Matlab files.

NcView is a software tool that is used in viewing and analyzing the CASA inputs and outputs. The ncBrowse (in Windows OS) and ncView (in Linux OS) are the tools used. MS Excel is a powerful mathematical tool that is used in plotting the graphs. The graphs plotted are of NEP, NPP, and LAI state-wise and of southern USA. The graphs form a main part in the analysis of the sources and sinks.

Potter's Regional Scale Study

Dr. Christopher Potter is currently a NASA senior research scientist in the Ecosystem Science and Technology Branch at Ames Research Center. Dr. Potter and Dr. Steven A. Klooster, a senior scientist at the Ames Research Center (ARC) are the people behind the CASA project and the designing of the CASA model.

Dr. Potter's and Dr. Klooster's regional scale study gained a lot of importance among the environmentalists. So far, regional scale studies that are carried out and that are being carried out are of the Amazon Rainforest, Boreal Forest of Canada, California, Central California Coast, Southeastern US, Virginia State Forest, and Yellowstone.

The present research concentrates mainly on the southern US, which includes Alabama, Arkansas, Florida, Georgia, Mississippi State, North Carolina, South Carolina, Tennessee, Texas and Virginia State. The regional scale study helps a lot in regional assessment of the kind of land or vegetation and helps in decision making to reduce the GHG by making changes to the kind of vegetation etc.

Potter's regional scale study about North American carbon sink for the period 1982 through 1998 estimates the NEP for CO_2 has an annual sink of +0.2 Pg C per year and +0.3 Pg C per year with a nearly zero NEP flux during the cooler periods [30].

Potter and Klooster also estimated carbon emissions of 0.2 to 1.2 Pg C per year in the Brazilian Amazon region, which is mainly because of the daily fire activity in the region [30]. It can be concluded that the Brazilian Amazon region is a source of a carbon [31]. CASA estimated that 7.4 Pg of carbon is stored in dead wood pools, while surface soils stored 28.6 Pg of carbon in North America [32] and [33].

The regional scale studies of the NASA CASA project estimate the net primary production, predict the soil trace gas emissions and emissions from the soils. They also estimate the annual emission of $N - NH_3$ from soils, forest, and industries and assess the climate changes.

The Potter's CASA model is not confined to regional scale study only, but it is also implemented on a global basis. The model estimates a global NPP of 48 Pg C per year and concludes that over 70% of terrestrial NPP occurs between 30° N and 30° S latitude [14].

CHAPTER IV

RESULTS AND DISCUSSIONS

CO_2 fertilization is responsible for a majority of the small C sink, while climate change contributes a small amount. Birdsey et. al. suggest that land use change and intensive ecosystem management are responsible for large C sink compared to CO_2 fertilization and climate change [34, 35].

NcView Plots of NPP and NEP

The NcView plots of the Net primary Productivity (NPP) and Net Ecosystem productivity (NEP) on a regional scale help in analyzing the contribution of carbon by a particular region or states towards the global source and sink. Further, they also help in determining the significant addition of land cover NPP towards the global source and sink of carbon.

Figures 36 – 47 show the net primary productivity and net ecosystem productivity plots of Alabama, Arkansas, Georgia, Mississippi, South Carolina, and Texas. The NPP for the state of Alabama during the month of May was between 2 – 3 g C/m^2/month while it was between 5 – 7 g C/m^2/month towards south of Alabama. For the state of Arkansas, the NPP varied between -1 and 7 g C/m^2/month with most of the region having NPP values of 2 - 4 g C/m^2/month. From the NPP plots in Figures 36, 38, 40, 42 and 44, the

NPP varied between 1.5 - 3 g C/m^2/month for most of the region. For Texas, the NPP was between 1 - 4 g C/m^2/month and 7 - 10 g C/m^2/month as illustrated in Figure 46. The white patches as can be observed in Figure 46 and 47, NPP and NEP for the month of May for TEXAS respectively correspond to the missing values or bad data.

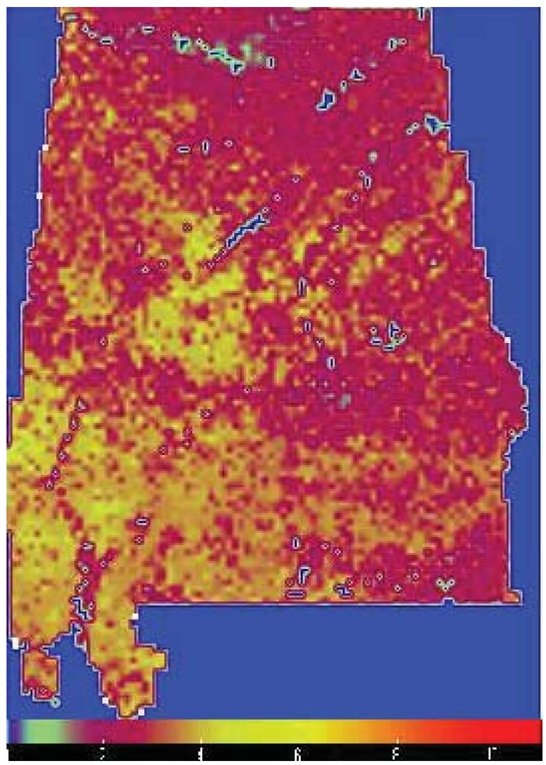

Figure 36   NPP for the Month of May for ALABAMA

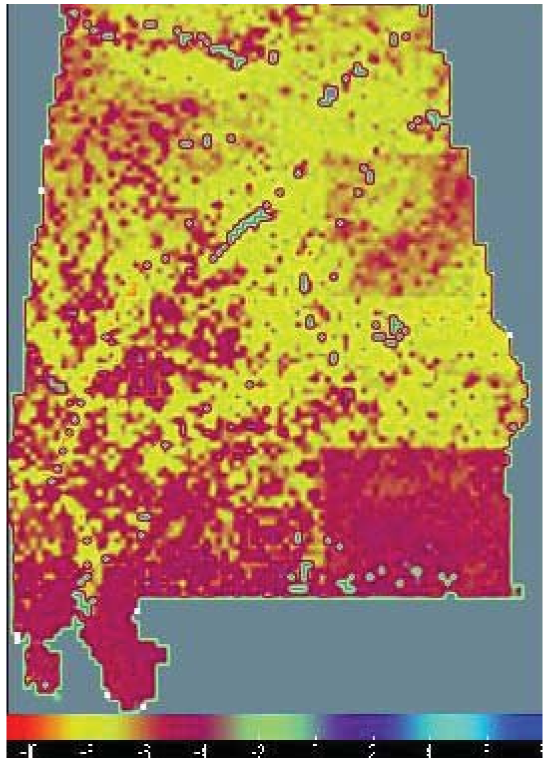

Figure 37   NEP for the Month of May for ALABAMA

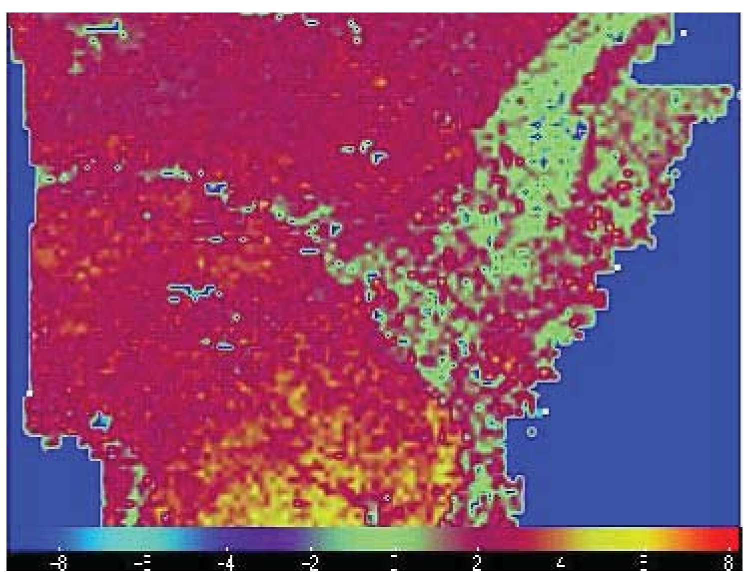

Figure 38   NPP for the Month of May for ARKANSAS

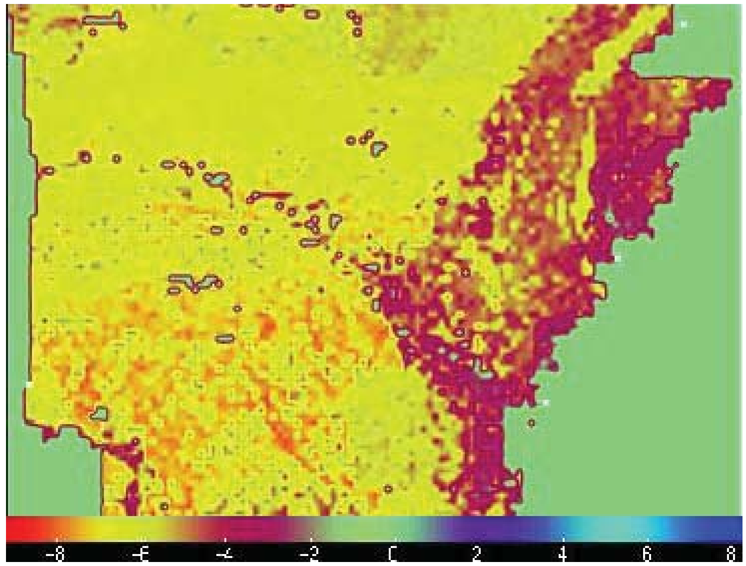

Figure 39 NEP for the Month of May for ARKANSAS

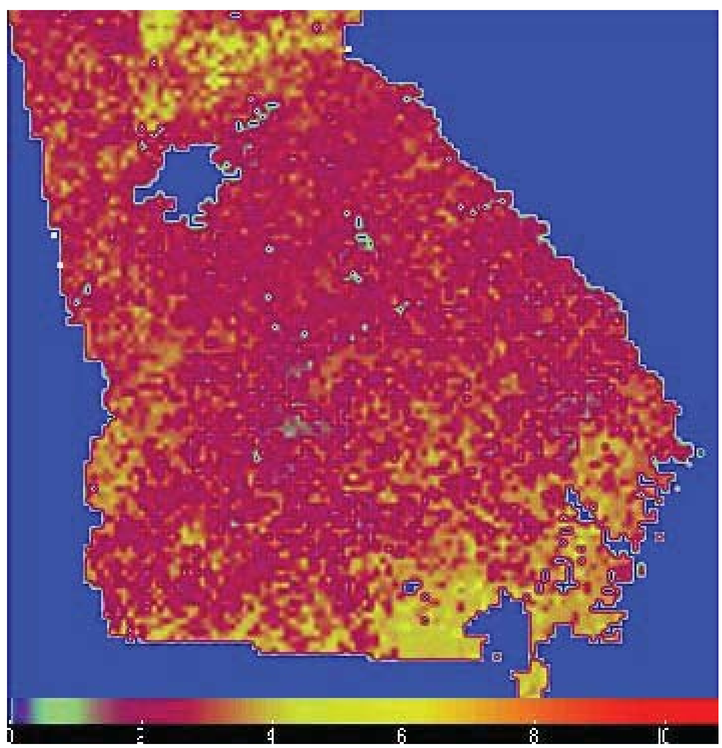

Figure 40 NPP for the Month of May for GEORGIA

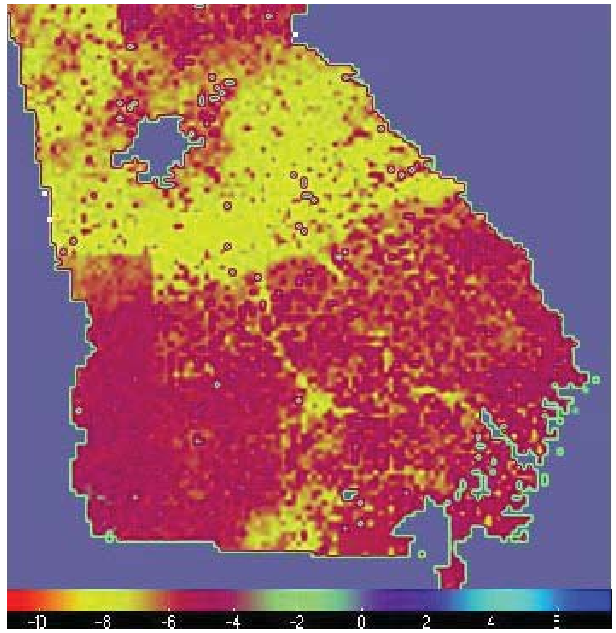

Figure 41   NEP for the Month of May for GEORGIA

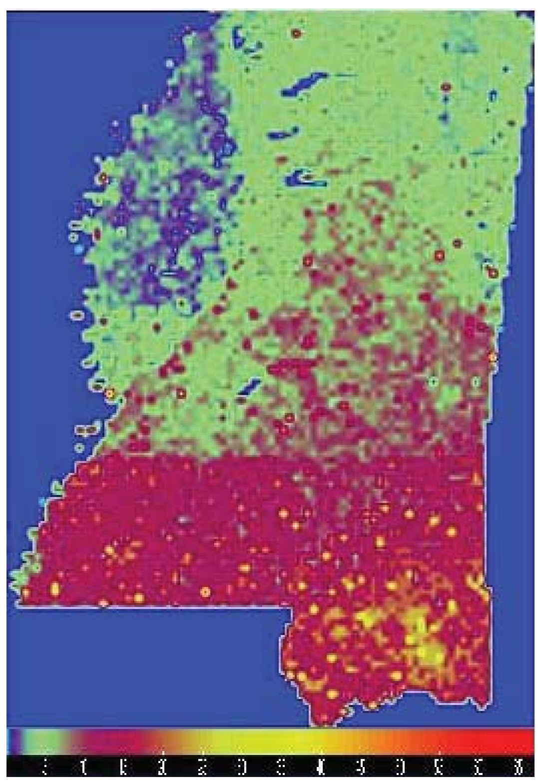

Figure 42   NPP for the Month of May for MISSISSIPPI

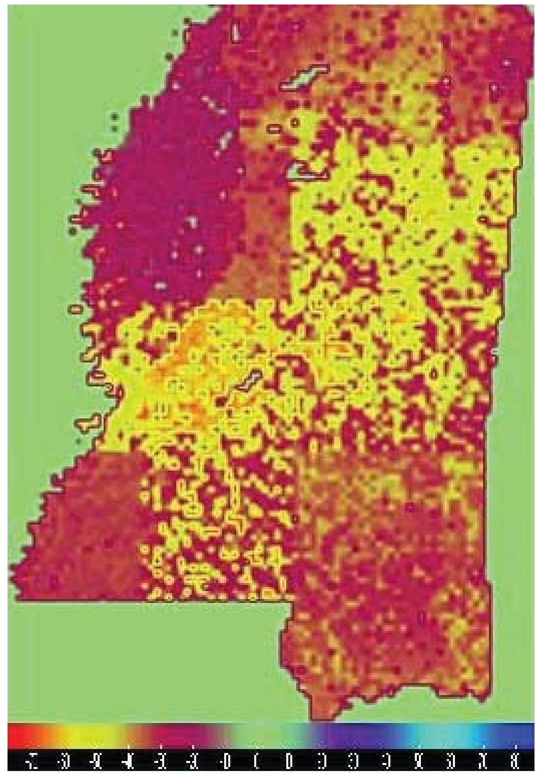

Figure 43   NEP for the Month of May for MISSISSIPPI

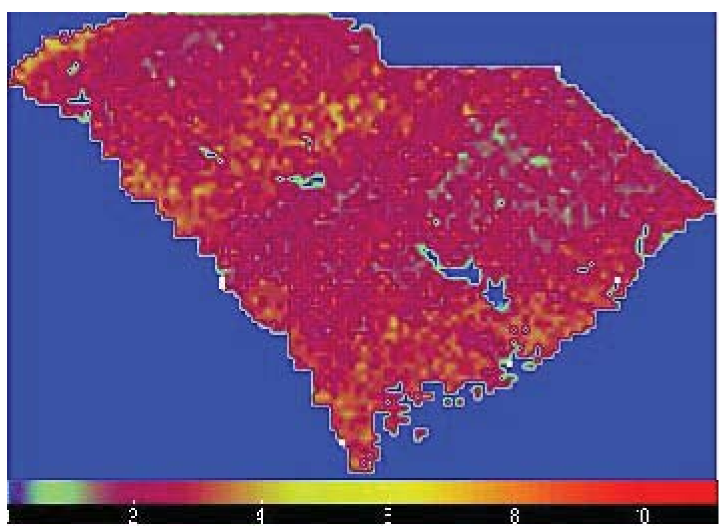

Figure 44   NPP for the Month of May for SOUTH CAROLINA

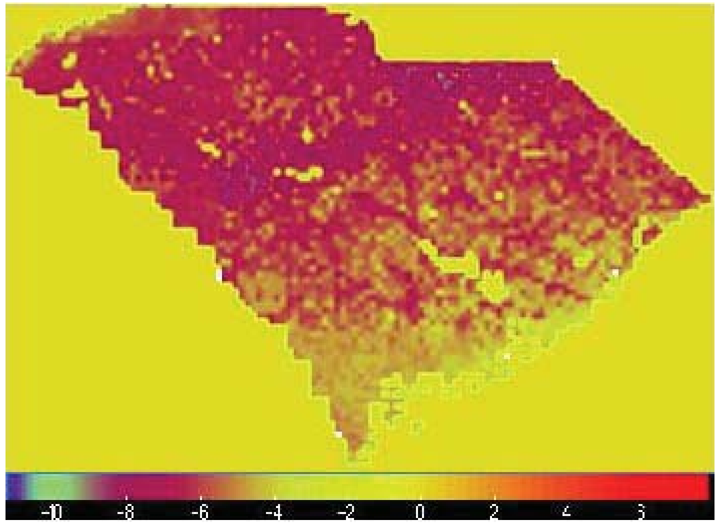

Figure 45   NEP for the Month of May for SOUTH CAROLINA

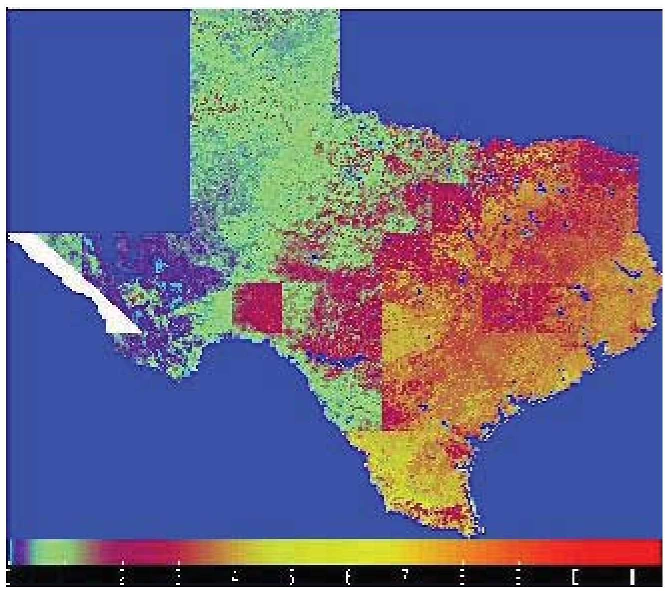

Figure 46   NPP for the Month of May for TEXAS (white, for missing values)

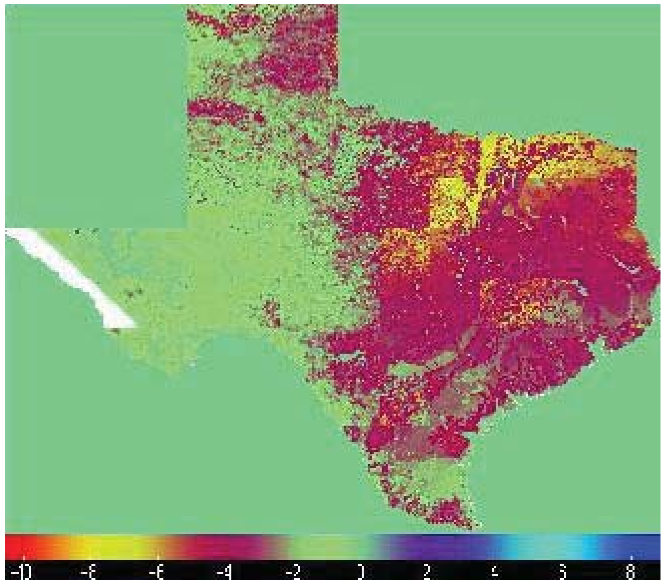

Figure 47 NEP for the Month of May for TEXAS (white, for missing values)

Forests, woodlands and crops of the southern United States have showed NPP increases summer, also suggesting an increase in the length of the growing season. Regions of large NPP increases occurred in the southeastern United Sates.

Observing Table 2 and Figure 48 the south central, southeastern, and eastern United States had maximum monthly NPP trends in June through October; the growing season appears to be lengthening here as well as at high latitudes. The southeastern United States had NPP increases in spring and fall that appear to be somewhat independent of climate driving factors. Hicke et al. [2002] showed that this region's large increases computed by CASA were matched by increases calculated using forest inventory data [12].

The NPP trend for the southern United States follows the regular pattern; with NPP being high during the summer and low during the winter as shown in Figure 48. The NPP was found to be ~165 g C /m^2/month during the month of January where it increases as the summer approaches. During the summer, June through August, the NPP is high as compared to other months varying between ~890 g C /m^2/month and ~1200 g C /m^2/month. The high NPP for September can be understood by the fact that the summer in the southern United States is likely from mid May through mid October; this is different as compared to the meteorological summer, which is May through August. During the winter season, December through March, the NPP is low varying between ~165 g C /m^2/month and ~400 g C /m^2/month.

The NPP for different states are in good agreement with the CASA - Carbon Query and Evaluation Support Tools (CQUEST). Average NPP for the state of Alabama was found to be ~560 g C /m^2/year while the CQUEST NPP varied between ~500 – ~800 g C /m^2/year. The Table 2 shows the average NPP calculated for different states and the CQUEST NPP range.

In southern US forests still represent the predominant land cover 63% and crops occupy about 28% [36]. As we go towards the east of southern US i.e. from Arkansas towards Virginia it can be observed from the Table 2 that there is a decrease in the average NPP. The states of Alabama, Mississippi, Georgia, and Florida have the highest average NPP, which is because of the large crop lands, deciduous and evergreen forests. It can be further observed that, the states with highest average NPP occur between latitudes 30° N and 40° N. The tile h10v05 with longitudes 104.4° W – 80.82° W and latitudes 30° N and 40° N is a major land cover source contributing towards high NPP in the southern US. The current total NPP for the southern US is estimated to be ~875 T g C per year which in good agreement with Nemani et. al. and Mickler et. al. [36, 37].

Table 2 CASA Estimated NPP and CQUEST Estimated NPP

State	Estimated NPP	CQUEST estimated
Texas	230.89	200 – 600
Arkansas	556.16	500 – 600
Mississippi	554.63	450 – 800
Alabama	560.73	500 – 800
Georgia	475.22	500 – 750
Florida	506.08	300 – 700
Tennessee	356.26	400 – 700
South Carolina	346.9	400 – 600
North Carolina	327.82	400 – 700
Virginia	218.89	400 – 650

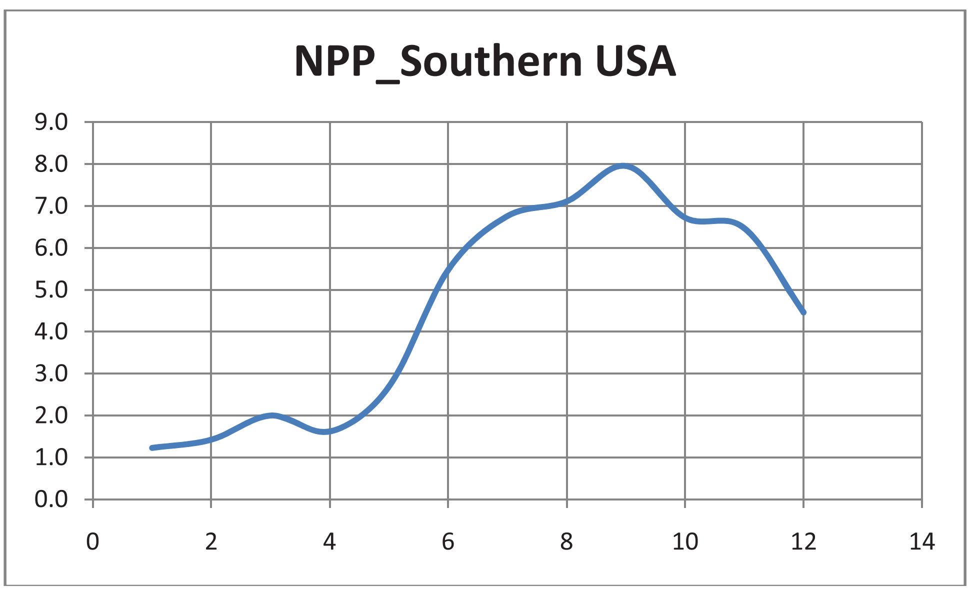

Figure 48 Average NPP Plot for Southern United States of America

NPP Plots at Random Location

Figures 51 through 59 show that NPP is high during the months of June through August and low from November to March for all the states. It can be inferred that NPP is estimated to be high during the summer season. During the winter months the NPP is estimated to be low or non-existent. This observation can be clearly explained by the plant photosynthesis process.

During the summer, daylight is long, and plants continue to prepare food through photosynthesis by using water absorbed from the roots and CO_2 from the air. Hence, plants have more time to convert the atmosphere carbon into solid carbon. Therefore NPP is high during the summer months during which the region acts as a sink for carbon. The day length decreases and nights increase during the months of November through March. The highest NPP for most of the states can be observed during the summer months especially during the month of July, which is usually the month with highest temperatures and precipitations.

At nights, plants prepare food by the process of photorespiration, where CO_2 is released into the atmosphere. Since nights are long, more CO_2 is released into the atmosphere. During these seasons, the region acts as less sink (because of shorter days and longer nights), which in a sense is referred to as, a source for carbon. This trend usually gives the bell shaped curve for NPP.

The NPP patterns usually follow a bell shaped curve. The NPP plots for different states vary in the shape of the curve. The curves differ in relative size and breadth of the summer peak. Some have a narrow peak, as shown in Figure 59 (the NPP plot for the state of Virginia) and some have broad peaks, like the states of FL, GA, and AL. From these plots one can infer that half of the annual NPP of the states occurs during the summer, i.e., June through August.

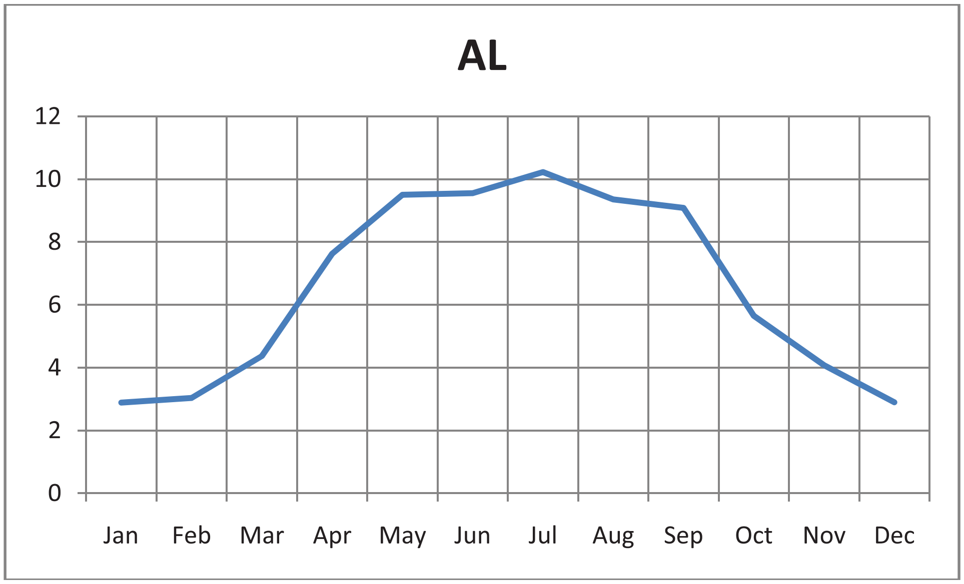

Figure 49 NPP Plot for the State of ALABAMA

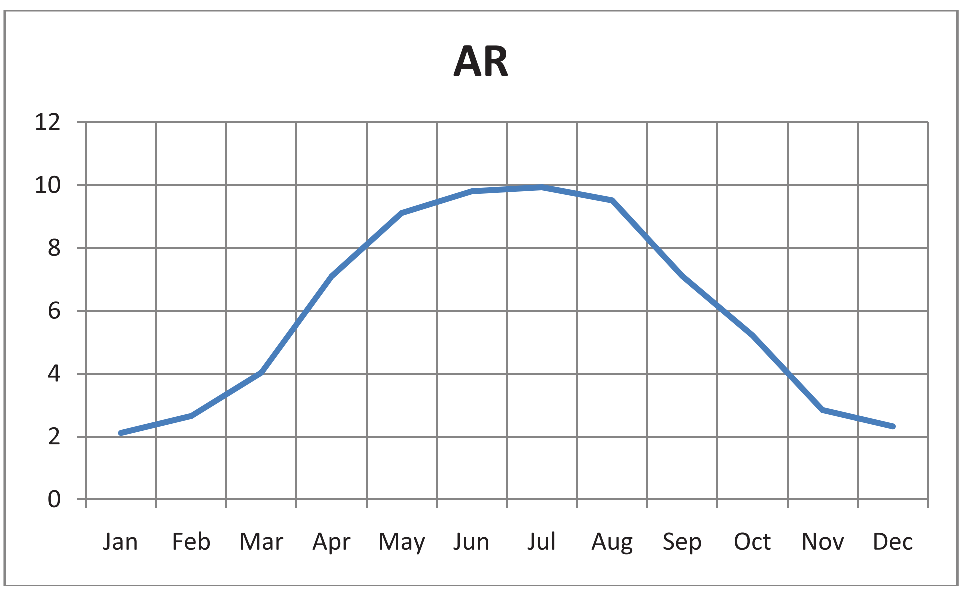

Figure 50 NPP Plot for the State of ARKANSAS

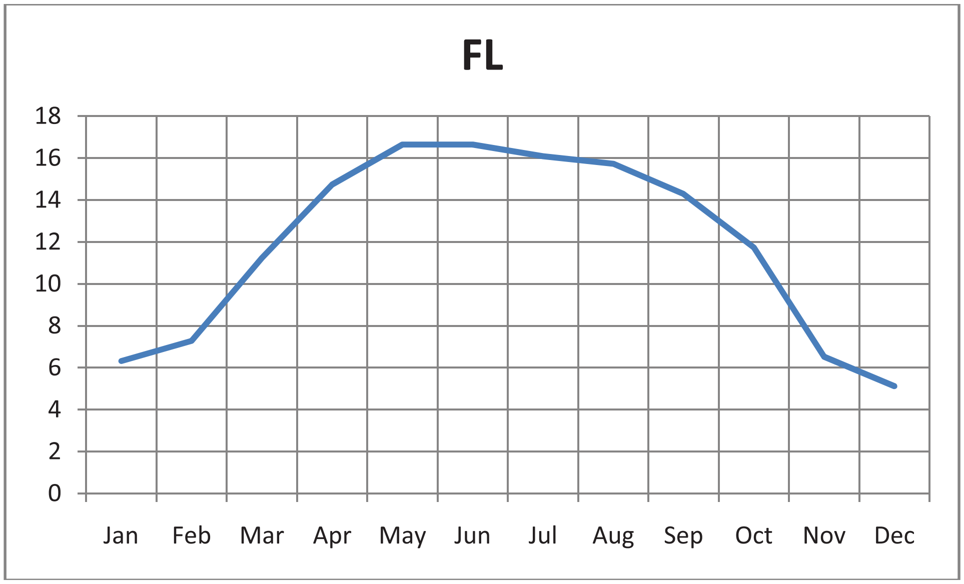

Figure 51 NPP Plot for the State of FLORIDA

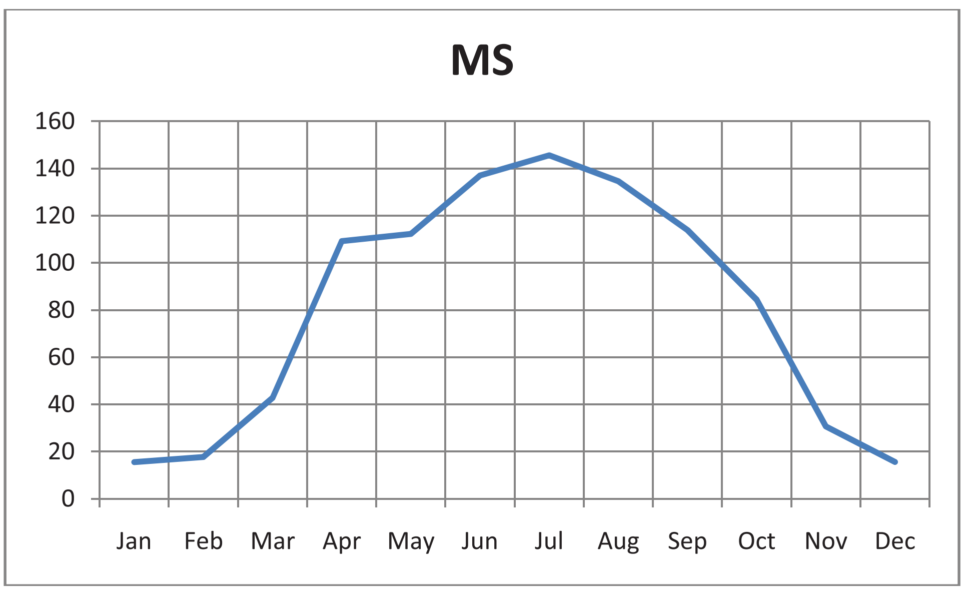

Figure 52 NPP Plot for the State of MISSISSIPPI

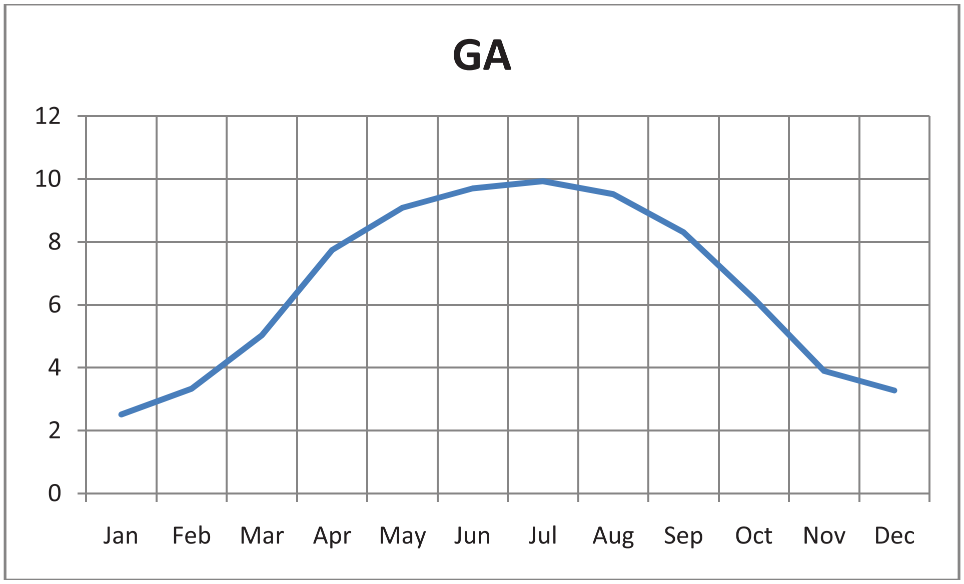

Figure 53 NPP Plot for the State of GEORGIA

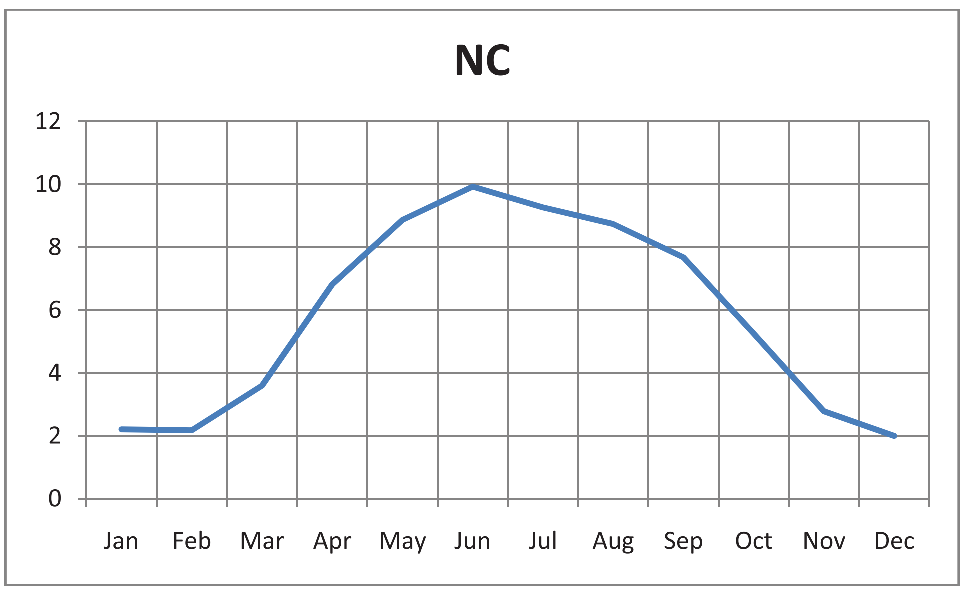

Figure 54 NPP Plot for the State of NORTH CAROLINA

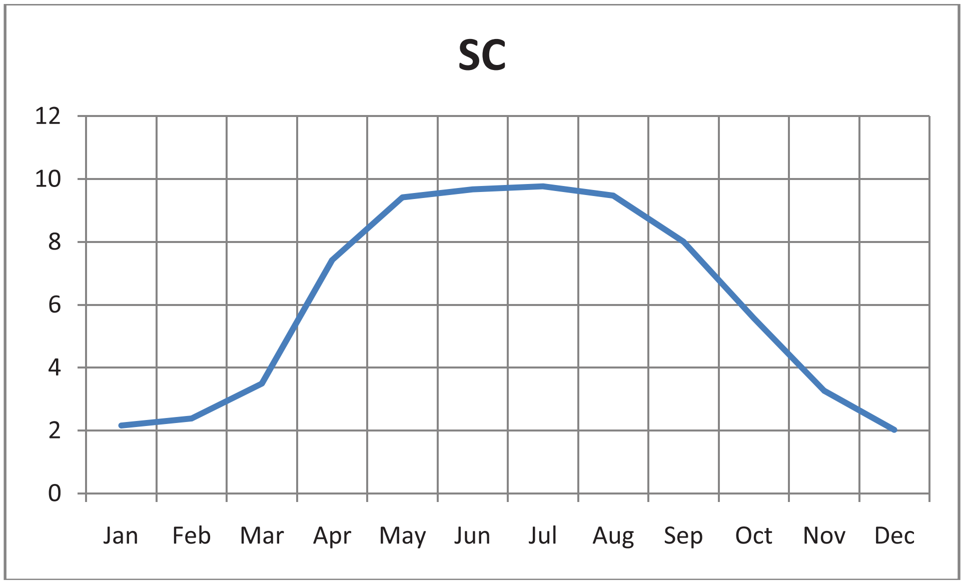

Figure 55 NPP Plot for the State of SOUTH CAROLINA

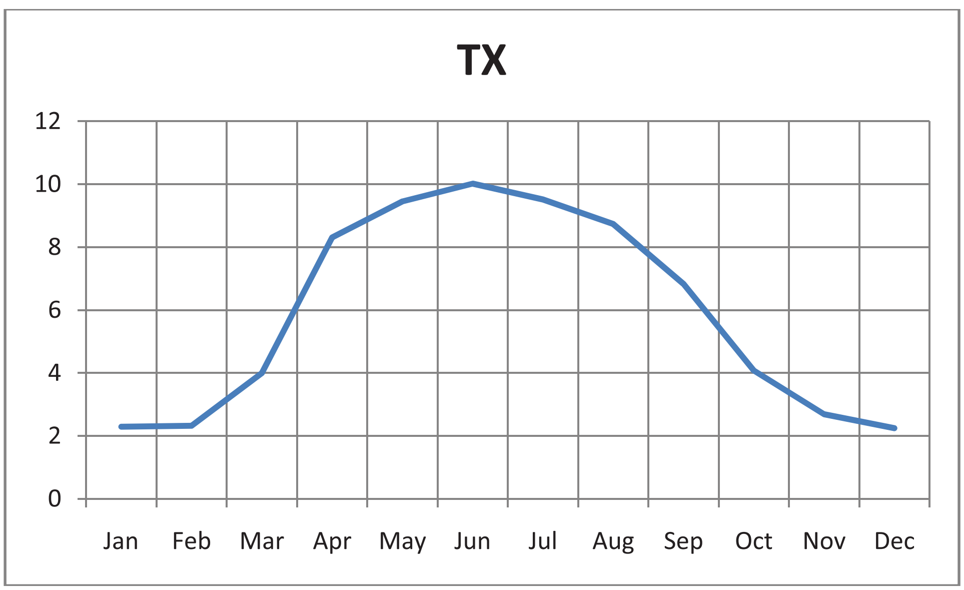

Figure 56 NPP Plot for the State of TEXAS

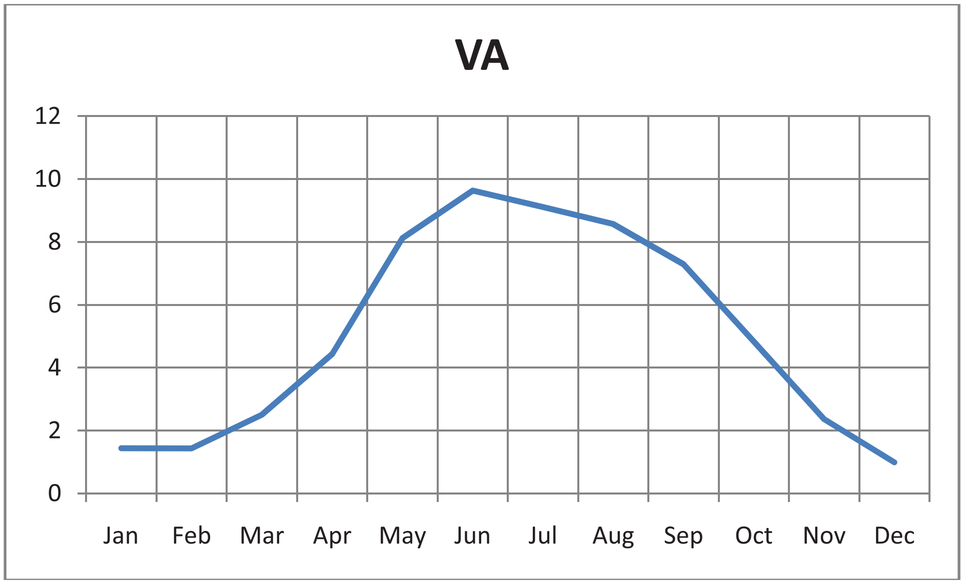

Figure 57 NPP Plot for the State of VIRGINIA

CHAPTER V

CONCLUSION

As more and more industrialization is occurring there is a necessity for more and more accurate estimations of the carbon that is being released into the atmosphere. This helps companies or countries to trade for carbon which can in a way reduce the amount of carbon being released and in another way provide a way of economy for the companies.

The CASA model approach is used to estimate NEP and NPP. The NPP trends calculated here agree with field based measurements for the State of Mississippi. Long term satellite data, such as NDVI, precipitation, soil, air temperature and solar radiation data, is used in the estimation of NEP and NPP. The Net Primary Productivity patterns for the southern United Sates are observed to identify regions of sources and sinks.

There is a need to know the amount of carbon being absorbed and released into the atmosphere. It is necessary to learn to reduce the amount of greenhouse gases and increase forestation. It is imperative to realize one's responsibility in reforming the eco-system and thus the environment.

Suggestions

The CASA model results agree with the trends of carbon throughout the year, i.e., sink during the fall and summer seasons and less sink or source during the spring and winter seasons. For more accurate and better results, the data sets should be more accurate.

Unlike the air temperature data that has been used here, the average of maximum and minimum temperatures; a mean temperature data would have been used. The average temperature and mean temperature are in no way the same. Hence, using the mean air temperature data sets would be more meaningful.

Land mask data is an important data set which identifies land and water. This data set in particular is good in terms of computational aspects; however, the data set should be more accurate to distinguish between water molasses, which are a major source for the sink of carbon.

NPP, one of the CASA model outputs, is mainly driven by NDVI, while temperature and precipitation have less variation on NPP and solar radiation data has greater NPP variation. Hence, data sets are to be chosen such that they do not have much variation on NPP. Most of the data used are the satellite measured data; hence the satellites should be capable to measure all that are responsible for carbon sink.

What More Can Be Done?

The CASA model predicts regions of sources and sinks. Some places might be more of a sink of carbon while others might be more of a source of carbon. If one can determine the reason as to why a particular region is a source or a sink, apart from determining the nature of the region or land, it would have been an important aspect for companies, government, and, most importantly farmers. There could be a way in which one can also determine the kind of crop or vegetation that is being carried out in a particular area or whether a particular area is industrialized area or not.

REFERENCES

[1] The National Aeronautics and Space Administration http://www.nasa.gov/

[2] The Department for Environment, Food and Rural Affairs, U.K. http://www.defra.gov.uk/

[3] United Nations Framework Convention on Climate Change www.unfccc.int

[4] Potter C, Klooster S, Genovese V, Myneni R: Satellite data helps predict terrestrial carbon sinks. *Eos – Transactions of the American Geophysical Union* 2003, 84:502-508

[5] Intergovernmental Panel on Climate Change http://www.ipcc.ch/

[6] Smith, W. L., Woolf, H. M., Hayden, C. M., Wark, D. Q., and McMillin, L. M.: The TIROS-N operational vertical sounder, Bull. Amer. Meteor. Soc., 60, 1177–1187, 1979.

[7] Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., No¨el,S., Rozanov, V. V., Chance, K. V., and Goede, A. P. H.: SCIAMACHY: Mission objectives and measurement modes, J. Atmos. Sci., 56, 127–150, 1999.

[8] Aumann, H. H. and Pagano, R. J.: Atmospheric infrared sounder on Earth observing system, Optical Eng., 33, 776–784, 1994

[9] Orbiting Carbon Observatory http://oco.jpl.nasa.gov/

[10] S. Houweling, F.-M. Breon, I. Aben, C. Rödenbeck, M. Gloor, M. Heimann, and P. Ciais. "Inverse modeling of CO_2 sources and sinks using satellite data: a synthetic inter-comparison of measurement techniques and their performance as a function of space and time"; Atmospheric Chemistry and Physics, 4. Pgs 523 – 538, 2004.

[11] Vegetation-Ecosystem Modeling and Analysis Project http://www.cgd.ucar.edu/vemap/

[12] Hicke, J. A., G. P. Asner, J. T. Randerson, C. J. Tucker, S. O. Los, R. Birdsey, J. C. Jenkins, C. B. Field, and E. A. Holland, 2002, Satellite-derived increases in net primary productivity across North America, 1982-1998. Geophysical Research Letters. 29(10): 1427, doi:10.1029/2001GL013578.

[13] Hicke, J. A., G. P. Asner, J. T. Randerson, S. Los, R. Birdsey, J. C. Jenkins, C. Tucker, C. Field, Trends in North American net primary productivity derived from satellite observations, 1982-1998, *Global Biogeochemical Cycles*, *16*, 10.1029/2001GB001550, 2002

[14] Potter CS, Randerson JT, Field CB, Matson PA, Vitousek PM, Mooney HA, Klooster SA: Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochem, Cycles* 1993, 7:811-841.

[15] http://earthobservatory.nasa.gov/Features/CarbonCycle/Images/carbon_cycle_diagram.jpg

[16] Janjic, Z., T. Black, M. Pyle, E. Rogers, H.-Y. Chuang, and G. DiMego, 2005: High resolution applications of the WRF NMM. Extended abstract, 21st Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction, American Meteorological Society, July 31 August 5, 2005, Washington, DC, 21 pp.

[17] Darren J. Kerbyson, Kevin J.Barker, Kei Davis; Analysis of the WRF Model on Large-Scale Systems; published in *Parallel Computing: Architectures, Algorithms and Applications* ,C. Bischof, M. B¨ucker, P. Gibbon, G.R. Joubert, T. Lippert, B. Mohr,F. Peters (Eds.), John von Neumann Institute for Computing, J¨ulich, NIC Series, Vol. 38, ISBN 978-3-9810843-4-4, pp. 89-98, 2007

[18] ftp://geo.arc.nasa.gov/pub/glemis/811-820.pdf

[19] NASA – CASA Approach http://geo.arc.nasa.gov/sge/casa/index.html

[20] NDVI Data Download http://edcimswww.cr.usgs.gov/pub/imswelcome/

[21] http://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_2.php

[22] Precipitation Data http://www.prism.oregonstate.edu/products/matrix.phtml

[23] Soil Data Download http://datagateway.nrcs.usda.gov/NextPage.aspx

[24] Earth System Research Laboratory (ESRL) http://www.cdc.noaa.gov/

[25] North America Land Data Assimilation System NLDAS http://ldas.gsfc.nasa.gov/

[26] The Land Processes Distributed Active Archive Center (LP DAAC) https://lpdaac.usgs.gov/lpdaac/tools/modis_reprojection_tool

[27] http://geospatialmethods.org/search/

[28] http://www.esri.com/software/arcgis/arcinfo/index.html

[29] Mexnc, SNC Tools and Matlab NetCDF Tool box http://mexcdf.sourceforge.net/index.html

[30] Christopher Potter, Vanessa Brooks Genovese, Steven Klooster, Matthew Bobo, Alicia Torregrosa," Biomass burning losses of carbon estimated from ecosystem modeling and satellite data analysis for the Brazilian Amazon region", Atmospheric Environment 35 (2001) 1773-1781.

[31] Christopher Potter, Vanessa Brooks Genovese, Steven Klooster, Alicia Torregrosa, "Biomass burning emissions of reactive gases estimated from satellite data analysis and ecosystem modeling for the Brazilian Amazon region", Journal of Geophysical Research, VOL. 107, NO. 0, 10.1029/2000JD000250, 2002

[32] Potter, C., S. Klooster, R. Nemani, V. Genovese, S. Hiatt, M. Fladeland, and P. Gross (2006), Estimating Carbon Budgets for U.S. Ecosystems, *Eos Trans. AGU*, 87(8), doi:10.1029/2006EO080001.

[33] Houghton, R.A., J.L. Hackler and K.T. Lawrence. 1999. The U.S. carbon budget: contributions from land-use change. Science 285:574-578.

[34] Birdsey, R. A., and L. S. Heath, Carbon changes in U.S. forests, in Productivity of America's Forests and Climate Change, edited by L. A. Joyce, pp. 56– 70, U.S. Dep. Agric. For. Serv., Rocky Mountain For.Range Exp. Stn., Fort Collins, Colo., 1995.

[35] Schimel, D., et al., Contribution of increasing CO2 and climate to carbon storage by ecosystems in the United States, Science, 287, 2004– 2006, 2000.

[36] C.Milesi, C.D. Elvidge, R.R. Nemani and S.W. Running , "Impact of urban sprawl on net primary productivity in the Southeastern United States", Numerical Terradynamic Simulation Group / School of Forestry University of Montana Missoula, MT 59812

[37] Robert A. Mickler, Todd S. Earnhardt, and Jennifer A. Moore, "Modeling and Spatially Distributing Forest Net Primary Production at the Regional Scale ", Mantech Environmental Technology, Inc., Raleigh, North Carolina; North Carolina State University, Raleigh, North Carolina; and U.S. Department of Agriculture Forest Service, Raleigh, North Carolina

APPENDIX A

TILE DISTRIBUTION

Tiles and State Distribution of Latitudes and Longitudes

Table 3 Individual Tiles Longitude and Latitude Distribution

TILE Number	Min Longitude	Max Longitude	Min Latitude	Max Latitude
h08v06	-115.47	-97.77	20	30
h09v05	-117.50	-92.40	30	40
h09v06	103.90	85.12	20	30
h10v04	124.58	91.38	40	50
h10v05	104.4	80.82	30	40
h10v06	92.38	74.50	20	30
h11v04	108.9	78.32	40	50
h11v05	91.38	69.27	30	40
h11v06	80.83	63.8	20	30
h12v04	93.35	65.26	40	50
h12v05	78.32	57.72	30	40

Table 4 Individual State Longitude and Latitude Distribution

State	Min Longitude (W)	Max Longitude (W)	Min Latitude (N)	Max Latitude (N)
AL	**84.88**	**88.47**	**30.18**	**35.0**
AR	**89.65**	**94.62**	**33.0**	**36.5**
FL	**80.03**	**87.64**	**24.50**	**31.0**
GA	**80.84**	**85.61**	**30.37**	**34.99**
MS	**88.1**	**91.65**	**30.2**	**35.0**
NC	**75.5**	**84.32**	**33.83**	**36.6**
SC	**78.53**	**83.4**	**32.03**	**35.22**
TN	**81.65**	**90.32**	**35.0**	**36.7**
TX	**93.52**	**106.65**	**25.83**	**36.5**
VA	**75.25**	**83.7**	**36.53**	**39.5**

Table 5   Decision of Tile into Which a State Falls

State	Tiles in which the state falls
AL	h10v05
AR	h10v05
FL	h10v06 ; h10v05; h11v05
GA	h10v05 and h11v05
MS	h10v05
NC	h11v05
SC	h11v05
TN	h10v05 and h11v05
TX	h08v06; h09v06; h09v05; h10v05
VA	h11v05

ArcMap Process Flow and NcView

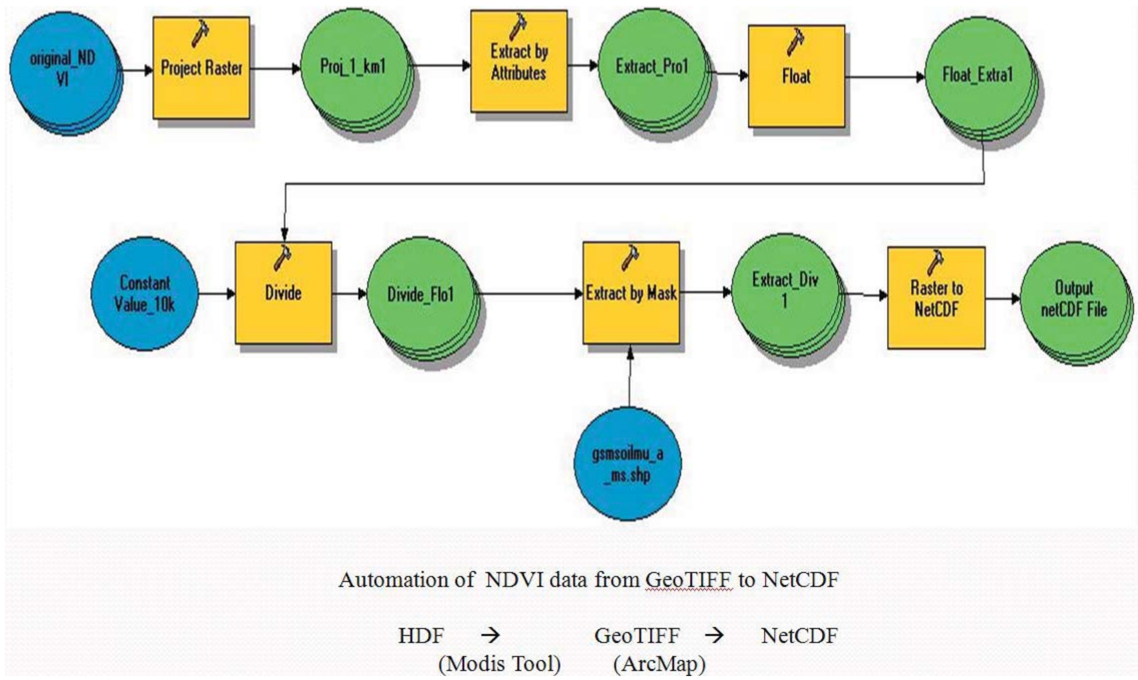

Figure 58 ArcMap NDVI Extraction Process Flow for One State

The above figure is a pictorial description of how the extraction of raw NDVI data, i.e., HDF format, is converted into ndvi.nc, the CASA input. Modis tool is used to extract the NDVI data from a bunch of data in HDF format and saved as images. The Arcmap tool uses this image and after a series of process converts it to netcdf NDVI monthly files. Bash scripts later concatenate the twelve monthly files to form a single NDVI input file. Mississippi state soil data is used as mask.

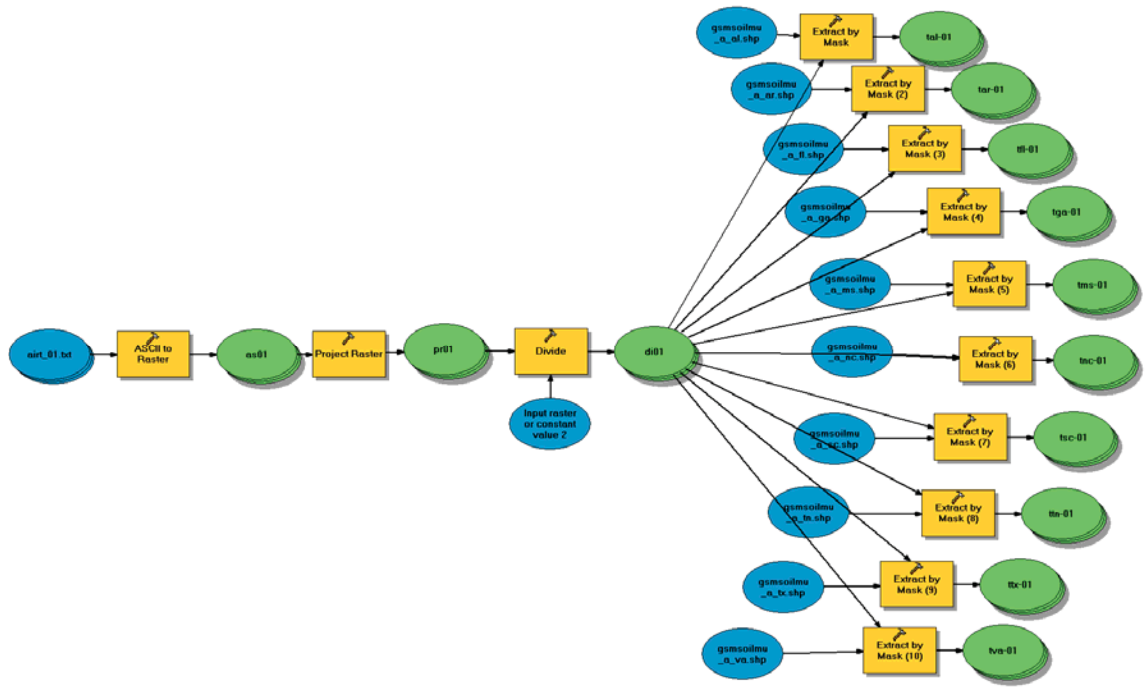

Figure 59 ArcMap Temperature Extraction Process Flow for Ten States

The above figure shows the sequence of process that takes place in converting a raw temperature data to form monthly files of all the states at one time. The individual data is extracted by applying individual mask, which is the soil data mask of respective states.

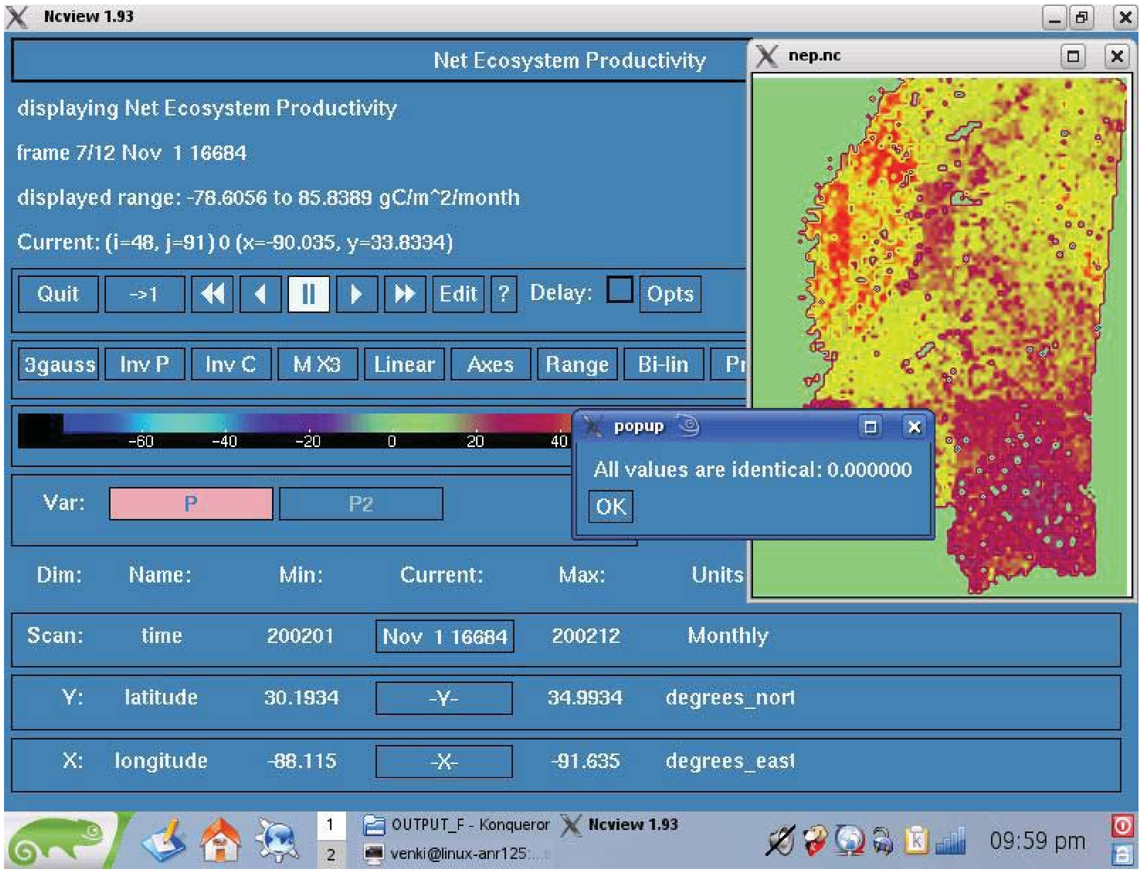

Figure 60   NEP Plot for MS State using NcView (point values for water body)

The figure is a screenshot of NcView software tool that is used in viewing the netcdf inputs and outputs. The screenshot is a NEP CASA output for the month of July where a random point is chosen. Since the mask for water is assumed to be zero, a click on the water body shows all values are identical to zero.

APPENDIX B

PERL AND BASH SCRIPTS

Script 1

Take each yearly output of the CASA and break them into monthly files of the same size as that of monthly inputs.

The following script is run as ./break.csh output.nc size

E.g. ./break.csh npp.nc 10800

```
#!/bin/bash

#$1 is the command line argument for the bash script
#which is the file name that one in interested in.

for file in $1
do
      #the variable P is dumped into the file a01
      ncdump -v P $1 > a01.cdl

      #the last line i.e '}' and the 1st 34 lines
      #i.e., the header of the netcdf file are removed
      cat a01.cdl | head -n-1 | tail -n+33 > a02.txt

      # the last characted i.e., ';' is removed
      sed -e 's/;//g' a02.txt > a03.txt
done
#creating respective directories
mkdir $1_monthly
```

```
#perl script is run with $2 being command line
#argument which is the num of elements in each month i.e.
row x column

perl BREAK_DATA.pl a03.txt $2

mv Set* $1_monthly
rm a0*

#end of program
echo "DONE with breaking the data into 12 months"
```

Script 2

BREAK_DATA.pl

Perl script to break the data into monthly files. The actual logic in breaking the data is in BREAK_DATA.pl

```
#Counter the number of elemenets seperated by a comma

#!/usr/local/bin/perl

#/Change test.txt to the file name desired
#ARGV[0] is interested file name i.e perl test.pl ARGV[0]

ARGV[1] ...
open(handle,$ARGV[0]) || die("could not open file");
@raw_data=<handle>;
close(handle);
$counter = 0;
$current_set = 0;
$set_size = $ARGV[1]; # give the number of data to be
reterieved (here it;s 120 x 90 = 10800)
@Array_with_all_data =();
```

```
foreach $Line_var (@raw_data)
{
      chomp($Line_var);
      $Line_var =~ s/^\s+//; # remove leading space
      $Line_var =~ s/\s+$//; # remove trailing space
      @individualdata = split(',',$Line_var); # Split each
line in comma
      push(@Array_with_all_data,@individualdata);
}

for($i=0;$i<@Array_with_all_data/$set_size;$i++)
{
      #Change the file name here
      open(handle,'>>Set_'.$i) || die ("Could not open
file");
      for($j=0;$j<$set_size;$j++)
      {
            if($j%7 == 0 )
            {
                  print handle
@Array_with_all_data[$i*$set_size+$j] .",".""\n";
            }
            else
            {
                  print handle
@Array_with_all_data[$i*$set_size+$j].",";
            }
      }
}
```

Script 3

ndvi_conv2nc.csh

The script converts raw input of NDVI from HDF to NetCDF format and then extracts NDVI from the output and finally concatenates all monthly files.

```
#!/bin/bash

#!!!!!!!!!!!!!START PROGRAM !!!!!!!!!!!!

#********* Start of HDF to NetCDF conversion***********
for file in h*
do
     cd $file
# Convert the available hdf files in to netcdf files
# the netcdf files names are the same as their
# corresponding hdf files

     filelist='ls *.hdf'
     for ff in *.hdf
     do
          ncl_convert2nc $ff
     done

#*********End of HDF to NetCDF conversion*******
#********* Start of NDVI extraction *********
# intialization such that output file has distinct names
     A=1
# operate on the available netcdf files

     filelist='ls *.nc'
     for ff in *.nc
```

```
      do
      # renames the variable 1_km_monthly_NDVI to a new
variable ' P '
            ncrename -v 1_km_monthly_NDVI,P $ff -O temp.nc

      # extracts the data from the netcdf file corresponding
to the variable
      # 'P'(the NDVI data)and saves as a netcdf file
            ncks -v P temp.nc out$A.nc
            A=$((A+1))
      done

#********** End of NDVI extraction **********

#********** Start of NDVI data concatenation ********

# concatenate all the 12 months data of the variable
'P'(NDVI data) and saves it
# as a netcdf file
      ncecat out[123456789].nc out1[0-2].nc $file.nc

#********** End of NDVI data concatenation *********

# clearing the temporary files for optimized use of hard
disk
      rm temp.nc
      rm out*.nc
      rm 2002*.nc

      cd ..
done
      #!!!!!!!!!!!!!! END PROGRAM!!!!!!!!!!!!!!
```