GEOWEB APPLICATION FOR SPECIES ASSOCIATION WITH VEGETATION TYPE AND CLIMATIC REGIMES IN WESTERN HIMALAYA OF INDIAN LANDSCAPE

APARNA KULKARNI
March, 2016

ITC SUPERVISOR
prof. dr. A. A. Voinov

IIRS SUPERVISORS
Dr. Arijit Roy
Dr. Harish Karnatak

GEOWEB APPLICATION FOR SPECIES ASSOCIATION WITH VEGETATION TYPE AND CLIMATIC REGIMES IN WESTERN HIMALAYA OF INDIAN LANDSCAPE

APARNA KULKARNI
Enschede, The Netherlands, March, 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Geoinformatics

ITC SUPERVISOR
prof. dr. A. A. Voinov

IIRS SUPERVISORS
Dr. Arijit Roy
Dr. Harish Karnatak

THESIS ASSESSMENT BOARD:

Chair	: prof. dr. M. J. Kraak
ITC Professor	: prof. dr. M. J. Kraak
External examiner	: Dr. S. P. S. Kushwaha
ITC Supervisor	: prof. dr. A. A. Voinov
IIRS Supervisor	: Dr. Arijit Roy
IIRS supervisor	: Dr. Harish Karnatak

OBSERVERS:

ITC OBSERVER	: dr. V. A. Tolpekin
ITC OBSERVER	: dr. N. A. S. Hamm
IIRS OBSERVER	: DR. S. K. Srivastav
IIRS OBSERVER	: DR. Sameer Saran

Dedicated to Appa, Amma and my teacher Prof. I. V. Rao

ABSTRACT

The natural ecosystem is diminishing day by day due to human interventions, rapid climate change, etc. This has resulted in the domination of invasive species over species of high economic, cultural and medicinal important. The study of individual species becomes complex and therefore, species groups with similar traits and response to environment known as plant functional types are studied. The objective of this research was to identify actual and potential plant functional types and change in these types for 1 km x 1km grid with the help of an algorithm integrated in a GeoWeb application. The algorithm followed the matrix approach consisting of traits, community and ecological variables matrices. The traits were grouped using k-means clustering. The dissimilarities between communities were calculated using Bray-Curtis dissimilarity measure and dissimilarities of environmental variables with Euclidean distance. The correlation between communities and environmental variables was calculated using Pearson correlation coefficient. The GeoWeb application was developed using Geodjango python web framework with the algorithm running at the backend .This type of WebGIS application which is first of its kind, helps to study the effects climate change on vegetation and take necessary measures to protect the natural assets of the country.

Keywords: Plant functional types, algorithm, clustering, traits, correlation, Geodjango

ACKNOWLEDGEMENTS

Firstly, I would like to thank my parents who have been supportive, encouraging and cooperative throughout my life. I express my sincere gratitude to my IIRS supervisors Dr. Harish Karnatak and Dr. Arijit Roy for showing faith in me and providing immense support and motivation. I would like to thank my ITC supervisor prof. dr. A. A. Voinov for his valuable remarks and suggestions during research work. I am thankful to all IIRS and ITC faculty for their guidance throughout the course.

My sincere thanks to Dr. S.K. Srivastav for making sure of our progress at every phase of research and providing valuable suggestions. I also thank Dr. Sameer Saran, Course coordinator for being supportive at all times.

I would also like to thank dr. Nicholas Hamm and dr. V. A. Tolpekin for their kindness and support throughout the course.

Special thanks to tea shop (BCMC) for providing amazing and refreshing tea which helped me to keep my stress at bay and work on my research with constant energy and hopes. Worth a special mention is Rajan, who has been encouraging and being all time support emotionally at all difficult times. Heartily thanks to Sreeja who was with me at all times.

I also thank Raj, Gokul P, Gokul, Sanjay, Kavitha, Deepti, Priyanka, Pooja, Dhruval for their constant help. I thank my wonderful Msc friends Solanki, Sayan, Amit, Nitin, Maral, Ipshit, Grace, Melkom, Honore, John, Erica for the memorable days at ITC and cherished Europe trip. I thank Swati, Vineet, Anchit, Rohit, Prithvi, Ram Sir, Sushil, Arsh, Arnab, Nikita, Shweta, Shilpa, Vidya, Roopa for spreading happiness all the time. I specially thank the CMA guys of IIRS for their constant help in software related problems.

Last but not the least I thank one and all who were behind the successful completion of my research.

Aparna Kulkarni

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

This chapter deals with the brief background of the research, motivation for the research, objectives and the formulated research questions.

1.1. Background

India is one of the mega-biodiversity nations having huge diversity in flora and fauna which includes plants which are economically and medicinally important, many of which are endemic, endangered species, etc. India also harbours four biodiversity hotspots i.e. The Himalaya, Indo-Burma, The Western Ghats and Sri Lanka, Sundaland including the Andaman and Nicobar Islands (Roy et al., 2012). This biodiversity richness is diminishing due to various reasons like increasing human population, industries, global climate change, deforestation, land use and land cover change, invasive species and many other, which in turn affects the ecosystem stability.

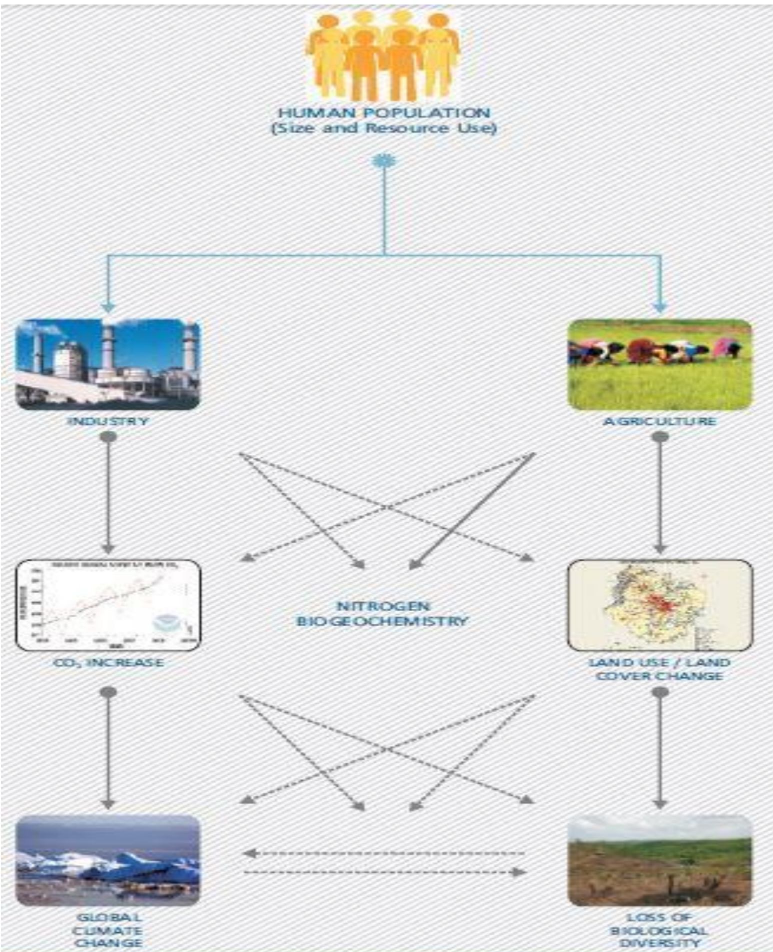


Figure 1-1 Variables leading to the loss of biological diversity (Roy et al., 2012)

The natural ecosystem is affected by various factors such as human interaction, climate change, land use change, soil, topography, etc. Due to this, several non-native species are present in some regions instead of native species. These non-native species when introduced intentionally or unintentionally outside their natural habitats suppress the native species. For example in Gangetic plain region of India, the deodar forest which is the potential vegetation has been replaced by thorn forest. To study this vegetation change, studying the response of individual species is a complex task and hence, the species must be grouped (Díaz & Cabido, 1997). This group can be of species based on their similar impact on environment, similar responses to disturbances (Condit, Hubbell, & Foster, 1996) or user-defined group of species with similar traits and identical response to ecosystem. This group of species is termed as Plant Functional Types (PFT) (Körner & Jeltsch, 2008). The criteria for PFT system has been emphasized on their structural and functional properties (Box, 1995). The structural properties may be the herbs, shrubs, trees, etc. and the functional properties may be varying time of growth, photosynthetic processes, etc. Some examples of PFTs are Needle leaf evergreen trees, Tall shrubs, Broad leaf deciduous trees, etc.

Identification of these types are important because, in an ecosystem, PFTs play different roles regarding energy processes and also their abundance estimation. This estimation is highly relevant to the assessment of ecosystem function (Díaz & Cabido, 1997) as well as to predict the result of global change on ecosystem processes and vegetation by using PFTs to describe patterns in plant communities (Valério DePatta Pillar & Sosinski, 2003).

1.2. Motivation and Problem Statement

The natural environment is changing rapidly due to changing climate, land use, human interaction, etc. Due to this, the species composition and biodiversity are affected. Many species are vanishing from regions dominated by land use, anthropogenic factors and human impact. The plant species belonging to a particular ecosystem evolves in balance with other animal, plant and insect species within their proximity. When they are away from their normal range, they become competitive over the native flora in the absence of predators. Thus, the native or potential species get replaced by non-native species. To study this vegetation change, studying individual species change is not recommended. Hence, the plants are grouped based on different factors such as plant traits, ecological variables, soil factors, etc., and these groups are called plant functional types. These functional types helps to understand the community response to environmental change.

.

Many remote sensing techniques were adopted to identify plant functional types. For example MODIS data produced PFT map of coarser resolution for entire globe and the accuracy was less (Sun & Liang, 2008).

Also, the community structure and environmental variables were not considered. With the help of an algorithm, all these conditions can be taken into account for identifying plant functional types.

All previous methods for identifying PFTs was done on desktop level. Desktop based Geographic Information System (GIS) limits the number of users. It is confined to predefined software and operating system. Therefore, with the development of web technologies, GIS has been shifted from desktop level to web. Web based GIS application is platform independent; multiuser environment allows to integrate user defined datasets. Hence, with the help of an algorithm for identifying PFTs integrated in a WebGIS application which is the first attempt can assist in studying the change in PFTs as well as take necessary measures for the conservation of threatened species.

1.3. Research Objectives

The main objective is to develop a GeoWeb application for assessment of Plant Functional Types (PFTs) and the species associations in these PFTs with respect to vegetation type and climatic regimes.
The following are the sub-objectives through which the main objective can be achieved:

- ✓ To identify Actual and Potential PFTs
- ✓ To identify the regions where these PFTs are likely to change.
- ✓ To design and develop a GeoWeb application

1.4. Research Questions

- ✓ How the PFTs are to be identified based on environmental factors and vegetation type?
- ✓ How is uncertainty of PFT associated with the definition of PFT and what ranges of environmental factors may correspond to same PFT?
- ✓ What functionalities should the application offer to the users?

1.5. Innovation Aimed At

To develop a new algorithm based on environmental parameters to identify PFT for 1 kmx1 km grid for Western Himalayan region in the Indian landscape as a GeoWeb application.

1.6. Thesis Structure

The thesis accounts for the work done so far for this particular research work in five chapters. The *First* chapter gives a brief summary about this research, the objectives to be accomplished and research questions formulated from the research objectives. The *Second* chapter describes about the concepts of various terms used and previous work that has been done related to this research work. The *Third* chapter describes the study area chosen for this research and the different dataset used in the research. The fourth chapter describes with the methodology and the methods adopted for this research. The *fifth* chapter is on the results obtained and the discussions on it. *Sixth* chapter is about the conclusions and the recommendations of the research.

2. LITERATURE REVIEW

2.1. Plant Traits

Research has been carried out to answer the questions relating to the organisms distribution across various habitats or their influence on ecosystem functioning. Previous studies used Shannon diversity index, species richness for describing functional diversity and composition of community (Schwoertzig, Poulin, Hardion, & Trémolières, 2016) out of which trait-based approach is opted for describing community structure (Laporte & Garnier, 2012). A trait is defined as "Any morphological, physiological or phenological feature measurable at the individual level, from the cell to the whole-organism level" (Laporte & Garnier, 2012). Traits are properties of organisms that can be measured and which influences the performance of organisms (Salmon et al., 2014). Examples of plant traits are specific leaf area, leaf lifespan, specific root length, plant life form, plant height, wood Density, bark thickness, etc. (Laporte & Garnier, 2012).

The biodiversity dynamics and the distribution of species which depends upon certain mechanisms are identified by plant traits. It helps to understand how the communities have been shaped and to understand the species response to disturbances such as land use change, fragmentation, agricultural practices (Salmon et al., 2014). Plant traits provide deep understanding of the plant community assembly (Kooyman, Cornwell, & Westoby, 2010). In developing plant community, only the species adapted to biotic and abiotic conditions of the region successfully establish themselves. Apart from abiotic and biotic filters, dispersal filter play major role in selecting species subgroup from regional pool and the filtering results in the trait distribution among species that co-occur in community. Thus, traits provide information about species assembly (Bochet & García-Fayos, 2015). Trait knowledge provides insight about properties of vegetation change along different geographic gradients (Westoby & Wright, 2006). Also, they have capability for predicting ecosystem properties. These properties relate to functional attributes of species i.e. the traits (Pla, Casanoves, & Di Rienzo, 2012). According to V. D. P. D. Pillar (1999), the traits are those which are used to define plant types optimize association of environmental factors and vegetation. The selection of traits follows two approaches i.e. through experience in the past and through computer algorithms for selecting optimal subset of traits.

The traits are divided into quantitative and qualitative variables. Quantitative traits include continuous variables like leaf area and discrete traits like leaflets, and Qualitative traits include nominal and ordinal traits i.e. the traits must be recognized by its presence or absence. For example, if the species is evergreen or not (Pla et al., 2012).

A significant work of comparison between trait convergence and divergence patterns has been done by Valério D. Pillar, Duarte, Sosinski, & Joner (2009) in community gradients. Species in a particular community which have similar ecological requirements leads to trait convergence and association of species in a community may have different traits and leads to trait divergence. The two terms to describe convergence and divergence patterns adopted are trait convergence assembly pattern (TCAP) and trait divergence assembly pattern (TDAP). The study of TCAP and TDAP relation to ecological gradients helps to study community patterns and predict ecosystem function and community structure (Valério D. Pillar et al., 2009).

2.2. Plant Functional Types

On a global scale, the study of vegetation response to changing environment cannot depend on plant species because their distribution across globe is limited. The vegetation must be described with the help of plant functional types instead of species (V. D. P. D. Pillar, 1999) to facitate the study of the predictions of vegetation response. Pillar & Sosinski (2003) defines plant functional types as group of species irrespective of their evolutionary relationships that consist of similar traits and similar association to given set of variables. These variables can be plants response factors like disturbance regime, soil conditions or plant effects in the ecosystem like litter accumulation, biomass production. A functional group is also formed by group of species with similar functional characteristics co-existing in a community, and these groups can be defined using traits generated through cluster analysis. For example, species in a community with similar nitrogen content and photosynthetic process related to woody density is termed as plant functional type (Pla et al., 2012). Other examples of plant functional type are Needle leaf evergreen tree, Needle leaf deciduous tree, Broadleaf evergreen tree, Broadleaf deciduous tree, etc. (Bonan, Levis, Kergoat, & Oleson, 2002).

Important traits for determining plant functional types are plant size, plant permeance, plant architecture, leaf form, leaf structure, metabolism and reproduction. Plant size is measured about other plants. Plant permanence is measured particularly for woody versus non-woody structure and root longevity. Plant architecture is defined by branching pattern and growth form. Leaf form is expressed as leaf shape and size. Leaf structure is captured by shade tolerance and leaf area. Metabolism is captured by respiration and photosynthesis, and reproduction includes dispersal strategy and phenology.

For plant functional type's classification, identification of factors that limit plant type distribution is important, and the factors are:

- ✓ A natural spotlight on capacity in connection to ecological conditions, for example, climatic conditions, frequently as showed in structural elements;
- ✓ A physiological spotlight on plants internal function, especially at metabolism level.
- ✓ A geophysical spotlight on how atmosphere is affected by plant functions and affecting larger- energy and water balances (Box, 1995).

To evaluate the influence of changing climatic conditions, plant functional types assessment is widely adopted for determining species composition (Müller, Overbeck, Pfadenhauer, & Pillar, 2007). Anderson & Hoffman (2011) defines plant functional types as "acting and reacting units which respond to changes in environmental conditions". Plant functional types are important in studying disturbances where the set of characteristics may respond similarly which in turn helps to understand why and how the ecosystem changes. The sorting of plant species based on environmental gradients can be easily understood with the help of functional types (Schwoertzig et al., 2016). Further, plant functional types helps to compare complex and taxonomically distinct systems (Anderson & Hoffman, 2011).

Several research has been performed on plant functional types. Hodgson, Wilson, Hunt, Grime, & Thompson (1999) worked on allocating -Competitiveness (C) -Stress tolerance (S) -Ruderality (R) functional type to an unknown species using predictor variables. The position of individual species is indicated by predictor variables about each axis of CSR system. Competitiveness is described by low disturbance and low stress, Stress tolerance by low disturbance and high stress and Ruderality by low stress and high disturbance. Fourth environmental possibility i.e. high disturbance and high stress does not support vegetation at all. The steps for CSR allocation included definition of 'gold standard' for CSR type, predictor variables selection, predictor regression building and CSR type allocation to unknown herbaceous species. This method has extensive potential for interpreting ecosystem properties. The CSR approach was tested in Britain and can be applied for typical taxa of southern parts of Europe.

Similar work has been done by Souza, Forgiarini, Longhi, & Oliveira (2014) on subtropical tree species classification by ecological groups detection from traits in southern Brazil. The ecological groups were identified based on the hypothesis that along gradients where there are variations in traits, and there exist discontinuities of species. The traits used for species groups were leaf size, height, wood density, crown depth, stem slenderness and growth rates. To identify trait based species clusters, both non-hierarchical and hierarchical methods were used. Hierarchical clustering was carried out using Euclidean distances along with Unweighted Pair Group Method with Arithmetic Mean (UPGMA) technique and to identify non-hierarchical k-means groups, Simple Structure Index criterion was employed. The ecological group's distribution identified by cluster analysis was assessed by using Principal Component Analysis (PCA). The authors concluded that k-means groups were more effective than hierarchical groups regarding height gradients and resource capture.

Another significant research was carried out by Valério DePatta Pillar & Sosinski (2003) on finding optimal plant types in Southern Brazil by searching traits numerically with the help of recursive algorithm. The algorithm uses three matrices: trait matrix in which the populations are described by traits, community matrix where in the different communities are described by populations and environment variables matrix in which the community sites are described by environmental factors. The algorithm used UPGMA clustering for defining Plant Functional Types (PFT). UPGMA algorithm condenses the similarity matrix step wise. At each step, it identifies the pair with highest similarity in the matrix, and the values that corresponds to these groups or objects is averaged with other groups or objects (Legendre & Birks, 2012). The identified PFTs showed that their behaviour in communities is highly connected with the environmental factors. This method was tested for assessing the outcome of grazing levels and N-fertilizer on communities of natural grassland. The results revealed that optimized trait subset and corresponding PFTs are highly associated with environmental factors than non-optimal species or trait data. Also, the PFTs defined by polythetic approach gave better results than monothetic approach.

2.3. Clustering

Ecologists are interested in combining similar entities into groups to find complex relationships of ecological systems for achieving dimensionality reduction so that ordination technique can be used or to discard redundant data (McGarigal, Stafford, & Cushman, 2000). Suppose a group of species show similar distribution across an environmental gradient, and then they can be grouped together to form a cluster. Clustering is one of the techniques of data mining most useful for identifying distributions and patterns in the data. It is a method of grouping similar data points such that the data points in a given cluster are similar than data points of different cluster (Halkidi, Batistakis, & Vazirgiannis, 2001). Depending on the criteria used, clustering may result in different data partition and thus pre-processing is essential before using clustering method. The steps of clustering process is as follows

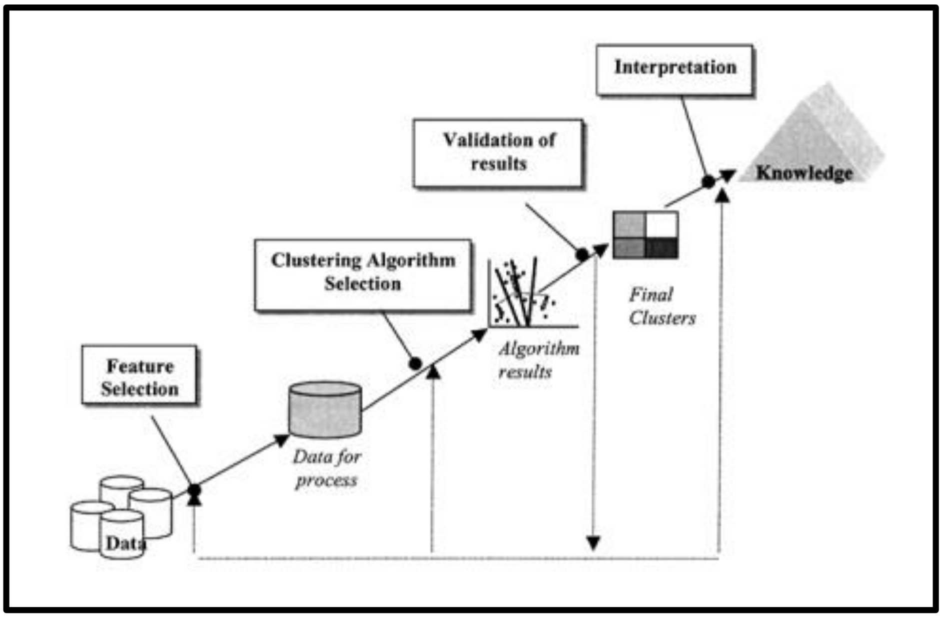
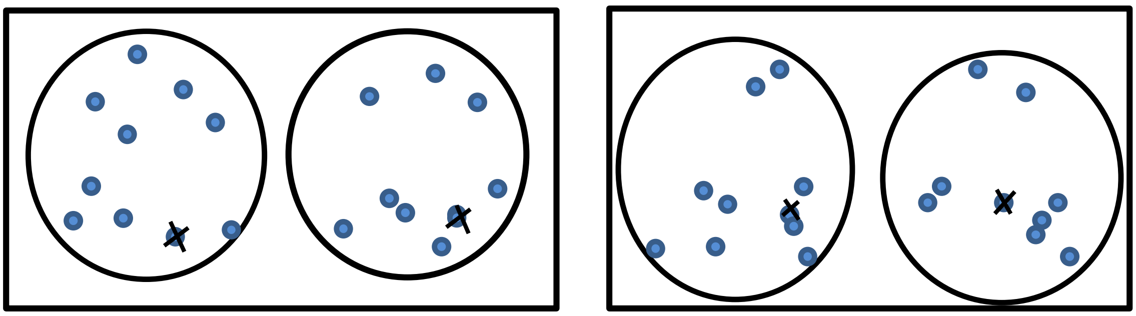


Figure 2-1 clustering process steps (Halkidi et al., 2001)

McGarigal et al., (2000) divides clustering method into hierarchical and partitioning methods. The hierarchical method partitions the instances in bottom up or top down approach. The hierarchical method results in dendrogram representing change of groupings at similarity level and clustered data is obtained at a desired level of similarity by cutting the dendrogram. Partitioning method, on the other hand, relocates instances from initial partitioning by moving the instances among different clusters. In this case, the number of clusters must be predefined. Error Minimization Algorithms belong to partitioning method. In this method, the distance is measured between each instance and its representative value and reduces error criterion. The widely used error criterion is Sum of Squared Error (SSE) which uses Euclidean distance as distance measure. K-means algorithm employs squared error criterion. K-means can be applied to different types of data. It is robust, simple and efficient (Wu, 2012).

In a multidimensional data set, K-means algorithm finds K number of clusters given by user. Each cluster consists of a central point which is called centroid. It finds optimum solution by following an iterative approach. The algorithm first selects set of centroids from data points randomly. Each data point is assigned to nearest centroid to form the initial clusters. Along each data dimension, new centroid is calculated by taking the mean value and results in new centroids. The data points are assigned to new centroids following the same procedure until centroids don't change (Thilina Hewaragamunige, 2014). Figure shows K-means clustering example for K=2. Each cluster center is denoted by 'x'.



(a) (b)

Figure 2-2 Example of k-means clustering (a)-Initial centroids (b)-Re-assigned centroids.

2.4. WebGIS

In desktop based or standalone applications, which is also known as single user environment requires high computation capabilities as well as predefined operating system and software at user end. With the advent of web technology, one of which is the client-server architecture allows geospatial data to be accessed in a shared environment (Karnatak, Saran, Bhatia, & Roy, 2007). Also, it offers many desktop GIS applications on web. WebGIS provides many advantages when compared to traditional desktop GIS. Multiple users can access the WebGIS application simultaneously. It is easy to use compared to desktop GIS as the end user is not expected to know GIS for accessing the WebGIS application.

WebGIS has grown rapidly from visualisation where in GIS data is displayed as maps, querying GIS data to providing analytical functions such as distance, area measurement, geocoding, etc. as well as simulations and virtual 3D on web. Future trend of WebGIS is towards setting up clustering solution to analyse big data for decision making. Also, geoprocessing requirement in web is a challenge.

Another advancement in the web environment is the concept of web service which provides high data interoperability and information access for distributed environment. It works on Extensible Markup Language (XML) which is data definition language and is based on Service-Oriented Architecture (SOA). These web services are very useful in GIS domain. The GIS components provide different methods to analyse, manipulate, visualize geo spatial data. In these methods, the interoperability is necessary integrate and share spatial data seamlessly in various GIS applications. Various organizations develop their storage structures, data models, and hence, problem of interoperability may arise. Therefore, Open Geospatial Consortium (OGC) has promoted open standards for GIS services.

There are six major categories provided by OGC services which include catalogue services, processing services, encoding, data services, portrayal services and other services like GeoRSS, Geospatial objects, etc. These categories consists of various OGC standards such as Simple Features Implementation Specifications, Styled Layer Descriptor (SLD) Implementation Specification, Web Feature Service (WFS), Web Coverage Service (WCS), Web Map Service (WMS), etc. (Percivall, 2008). These standards adapt to changing requirements, help to introduce new technologies seamlessly, provide security, authentication and implement N-tired architectures.

2.5. Python

Python is a powerful open source high level general purpose programming language developed under OSI ("Python," n.d.). It is dynamically typed and interpreted language. It is progressively used by people spreading over from customary bioinformatics to climate modellers and consists of extensive libraries for scientific programming ("why Python?," n.d.). Python modules like numpy, sklearn, scipy are mainly used for numeric and scientific analysis and matplotlib for data plotting.

2.5.1. Numpy

Numpy is a basic package in python licensed under BSD license used for scientific programming especially for matrix manipulations (Nathan Lemoine, 2013). It contains N-dimensional array object, robust functions as well as ability to define arbitrary data types that helps to integrate with different databases ("NumPy," n.d.).

2.5.2. Scipy

Scipy module in Python is mainly used for scientific computing ("SciPy," n.d.). It consist if scipy stack with set of core packages numpy for numerical computation. Scipy library which consists of toolboxes that area domain specific and numerical algorithms including statistics, optimization and many others, matplotlib which is used for plotting data points in 2D as well as 3D environment, pandas which provides data structures that can be used easily, IPython for interactive interface.

2.5.3. Scikit-learn

Scikit-learn is a machine learning package in python. It consists of efficient tools for data analysis and data mining. It supports different methods like classification, regression, clustering, dimensionality reduction, model selection, feature extraction and normalization, etc. ("scikit-learn," n.d.)

3. STUDY AREA AND MATERIALS USED

This section gives the details of study area and the reason for selecting the particular study area. It also discusses the materials used for this research

3.1. Study Area

The study area chosen for this research is Western Himalayas of India. It comprises of Himachal Pradesh, Jammu and Kashmir, and Uttarakhand. The total study area size is approximately 3, 00,000 square kilometres. It is home for wide variety of medicinal, flowering plants and other species of economic value. The complex Himalayan terrain delivers favourable climatic conditions which helps species to grow. But this species pool is facing threat due to increasing human intervention, land use change, climate change and increased anthropogenic activities.

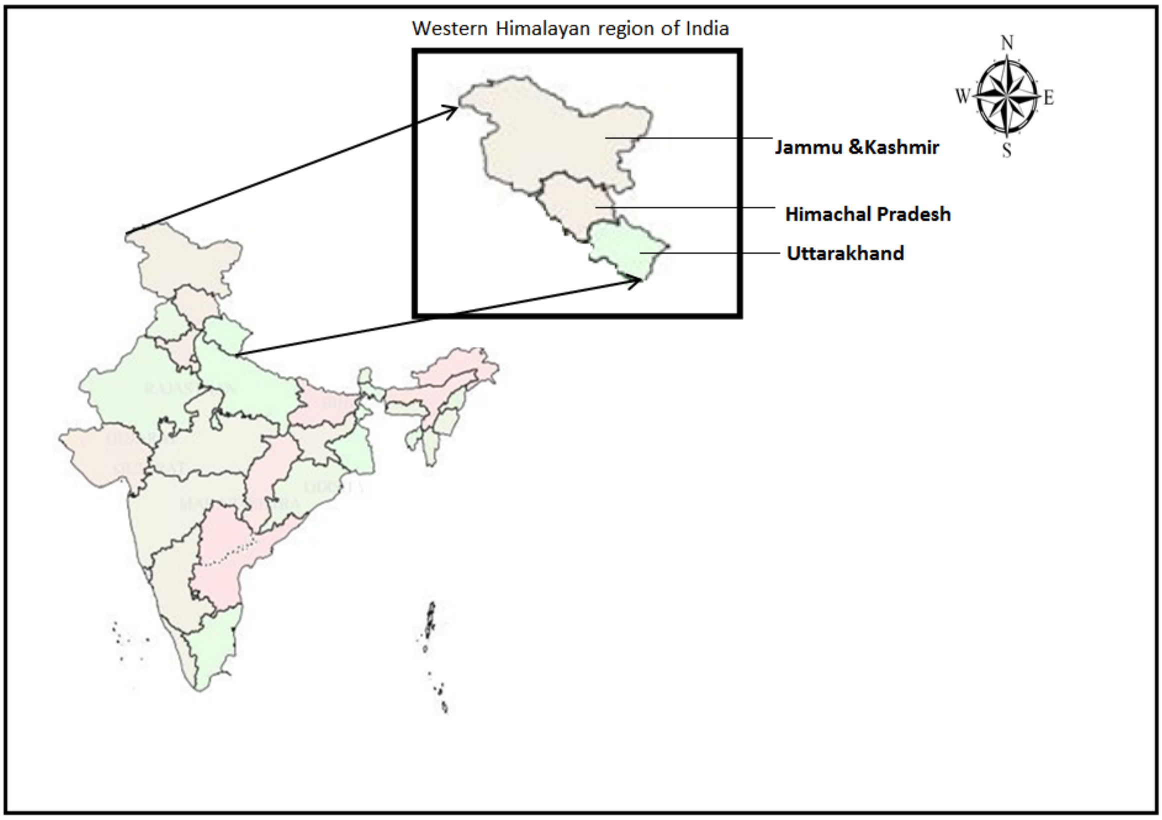


Figure 3-1 Geographical Location of Study Area (Source: http://bhuvan.nrsc.gov.in/bhuvan_links.php)

Figure 3-2 Google Earth view of the study area

3.2. Materials Used

In this study, vegetation type map of national level project 'Biodiversity characterization at Landscape Level' is used. The vegetation type for Western Himalayas consists of 13 types. For the chosen study area, the field sample plots are 1,688 in number and consists of 1,732 species. Along with these dataset, climate data which is 1km gridded monthly minimum, maximum temperatures and precipitation, soil and elevation data is also used. The data source of each dataset is described in the table below.

Table 3-1 Dataset used

DATA	SOURCE
Vegetation type map	**Biodiversity Characterization at Landscape Level (National level project of India)**
Field sample plots	**Biodiversity Characterization at Landscape Level (National level project of India)**
Climate data	**World Clim ("WorldClim - Global Climate Data," n.d.)**
Soil data	**National Bureau of Soil Survey**
Elevation data	**Shuttle Radar Topography Mission (SRTM)**

4. RESEARCH METHODOLOGY AND IMPLEMENTATION

This chapter deals with the overall methodology adopted for this research and the methods used to obtain the results and the implementation.

4.1. Methodology

The main objective of this research is to develop a GeoWeb application for assessment of Plant Functional Types (PFT) and the species associations in these PFTs with respect to vegetation type and climatic regimes. To achieve this objective the following methodology was adopted.

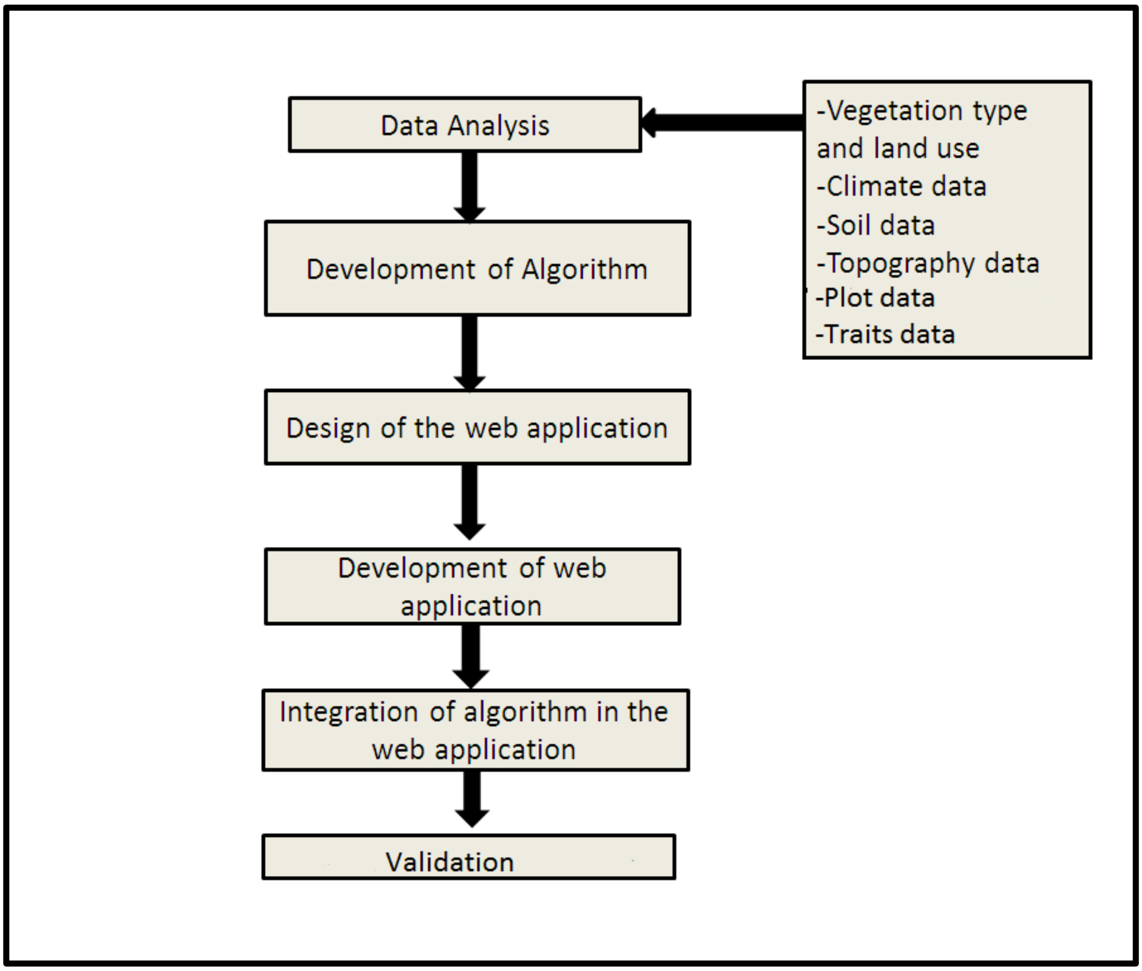


Figure 4-1 Overall Methodology

Explanation of the adopted methodology is as follows, For the adopted methodology, the actual and potential vegetation type map, climate data, soil data, topography data, plot data and trait data were analysed. These dataset were used as input to the algorithm. The algorithm was developed in python using various statistical and machine learning methods. A GeoWeb application was designed and the developed algorithm was integrated in the web application. The web application was developed using open source tools such as Geodjango python web framework, PostgreSQL object relational database, Leaflet JavaScript library and with the algorithm running at the backend, the results were displayed on map and finally the results were validated.

4.2. Implementation

This section explains the implementation of the methodology step by step

4.2.1. Data Analysis

The obtained data from various sources were first analysed. Below is the description about the datasets used for this research.

Actual Vegetation type map and Potential Vegetation type map

DOS-DBT project generated the actual vegetation type map for entire India. The map was clipped to fit the study area. It consisted of 31 classes. These classes were recoded according to Champion and Seth's (1968) classification and the resulting dataset consisted of 12 classes. For potential vegetation type the dataset was modelled at IIRS for entire India using Champion and Seth's (1968) classification. According to this classification, forest ecosystems consisting of unique phenology and physiognomy are driven by climatic conditions and are classified as type groups. Hierarchical approach was carried out for classification (Bandhu, 2011). The dataset was clipped to fit Western Himalayas. It consisted of 11 vegetation classes.

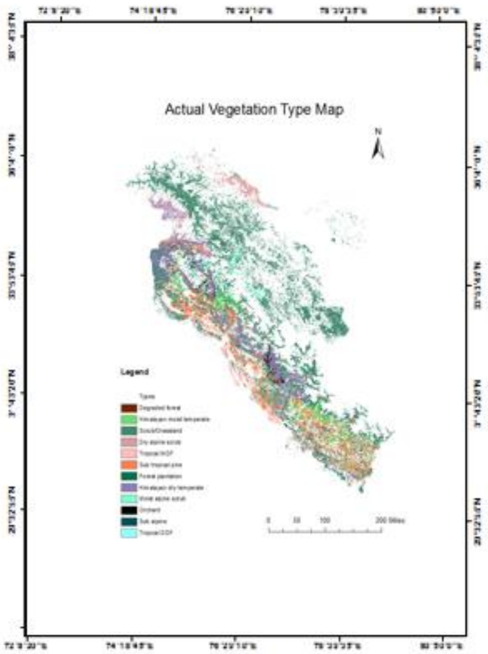


Figure 4-3 Actual Vegetation Type Map

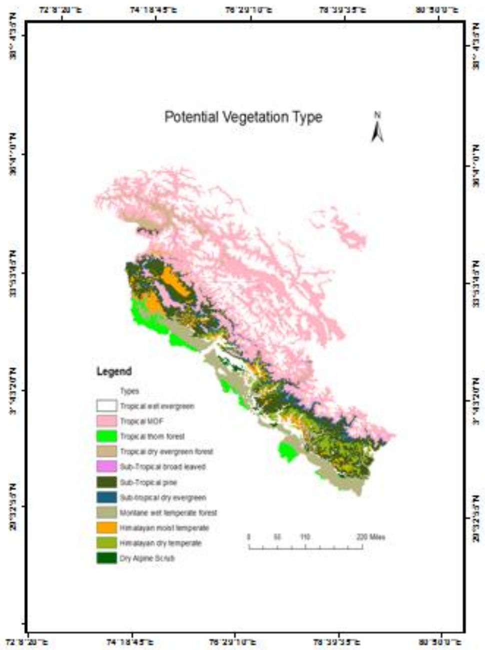


Figure 4-2 Potential Vegetation Type Map

Climate data

Climatic conditions are one of the important factors for species association. Temperature and Precipitation data is required to understand how the ecosystem will respond to variability in climatic conditions (Young, Watts, Taylor, & Post, 2016). The climate data which is 1km gridded data consisting of minimum temperature, maximum temperature, precipitation was downloaded from World Clim.

Soil data

Soil plays very important role in plants growth. The soil data was obtained from National Bureau of Soil Survey for entire India. It was clipped to fit the study area. The soil type was taken as one of the environment variable.

Topography data

Altitude is yet another important factor for species composition. Elevation data was obtained from SRTM.

Plot data

The study area consisted of 1633 plots. The plots were 0.04 ha i.e. 20 m x 20 m to 0.1 ha i.e. 31.62 m x 31.62 m (Roy et al., 2012).

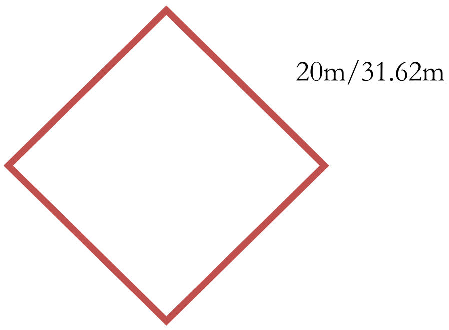


Figure 4-4 Sample plot

Species Data for identifying actual and potential PFTs

For identifying actual PFTs, the algorithm was implemented on four plots (1-Him-54, 1-Him-506, 1-Him-490 and 1-Him-515) in Himachal Pradesh due to time constraint. However it can be implemented to remaining plots. Similarly, to identify potential PFTs, the algorithm was implemented on the dominant species of Himachal Pradesh from Champion and Seth classification so that comparison can be made to identify the change in plant functional types. The actual and potential species which were taken into account is shown below.

Table 4-1 Actual Species

Species
Oxalis corniculata L.
Trifolium repens L.
Poa annua L.
Fragaria nubicola Lindley ex Lacatia
Pinus roxburghii
Cedrus deodara

Potential Species

Species
Betula
Rhododendron arboreum
Pinus roxburghii
Cedrus deodara
Lyonia ovalifolia
Ilex dipyrena
Myrica
Acer
Corylus
Ulmus
Abies pindrow
Cornus
Litsea umbrosa
Castanopsis
Pinus wallichiana
Arundinaria falcate

Traits data

From literature four important traits were selected for analysis. The traits for selected plant species were obtained from TRY-DB as well as literature. Trait codes were applied to all traits according to Pillar & Sosinski (2003).

	PlantHeightInterval text	PlantHeightID [PK] integer
1	<7.5	1
2	7.5-14	2
3	14-22.5	3
4	22.5-30	4
5	30-40	5
6	>40	6

i) Plant height (cm)

	LeafLengthInterval text	LeafLengthID [PK] integer
1	null	0
2	<4.3	1
3	4.3-8.6	2
4	8.7-13	3
5	13.1-17.3	4
6	17.3-21	5
7	>21	6

ii) Leaf length (cm)

	LeafAreaInterval text	LeafAreaID [PK] integer
1	null	0
2	<11.2	1
3	11.2-22.3	2
4	22.4-36.48	3
5	36.49-38.85	4
6	38.86-67.2	5
7	>67.2	6

iii) Leaf area (cm^2)

	PlantVegetativePhenologyTrait text	PlantVegetativePhenologyID [PK] integer
1	spring	1
2	summer	2
3	spring and summer	3

iv) Plant vegetative phenology

Figure 4-5 Traits

Next section of this chapter explains about the development of the algorithm to find the actual and potential plant functional types using the above mentioned datasets

4.2.2. Development of Algorithm

Defining the plant types is major problem in PFT identification. To define plant types, selection of suitable traits is very important. Numerical analysis helps in searching the relevant traits (Valério DePatta Pillar & Sosinski, 2003). In this study, an algorithm to identify actual and potential PFTs was developed and implemented using python programming language. The input matrices required for algorithm were developed in PostgreSQL relational database while the processing was performed in python. The algorithm flow to identify actual and potential plant functional types is shown below.

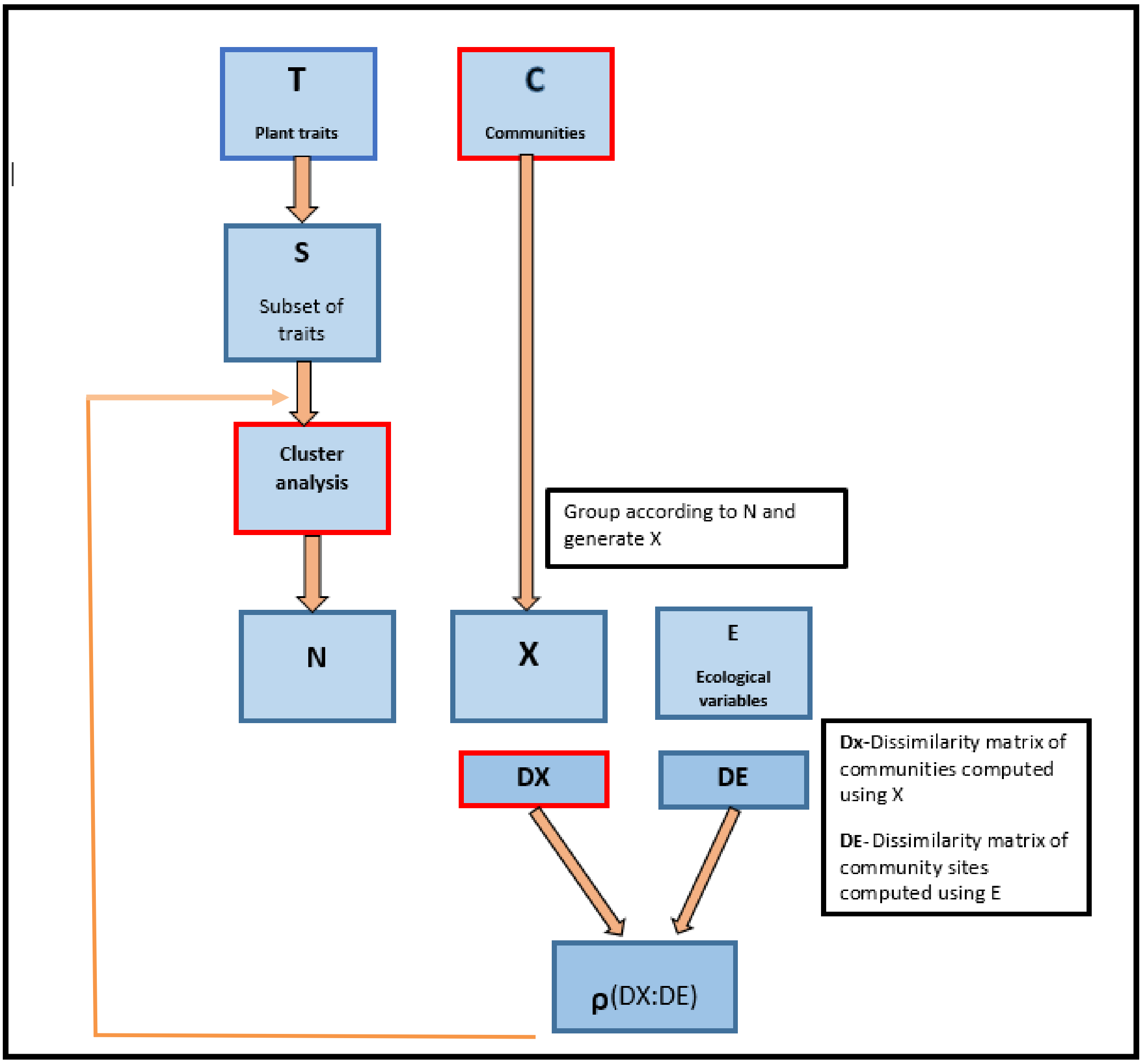


Figure 4-6 Algorithm

How does algorithm work?

Below is the step wise process of the algorithm
Require Input trait matrix, community matrix and ecological variables matrix: T, C, E

Step1: Takes trait matrix as input
T <- input Trait Matrix

Step2: Computes the subset of traits
S<- subset(T)

Step3: Performs K-means clustering on each subset of the traits to generate N clusters
N<- for i=1 to 15, i++{
Kmeans(S[i])
}

Step4: Takes community matrix as input and groups it according to N to generate X (PFTs) matrix
X<- for each N
group(C according to N)

Step5: Compute Bray-Curtis dissimilarity matrix (Dx) using X
Dx<- Bray-Curtis (for each X)

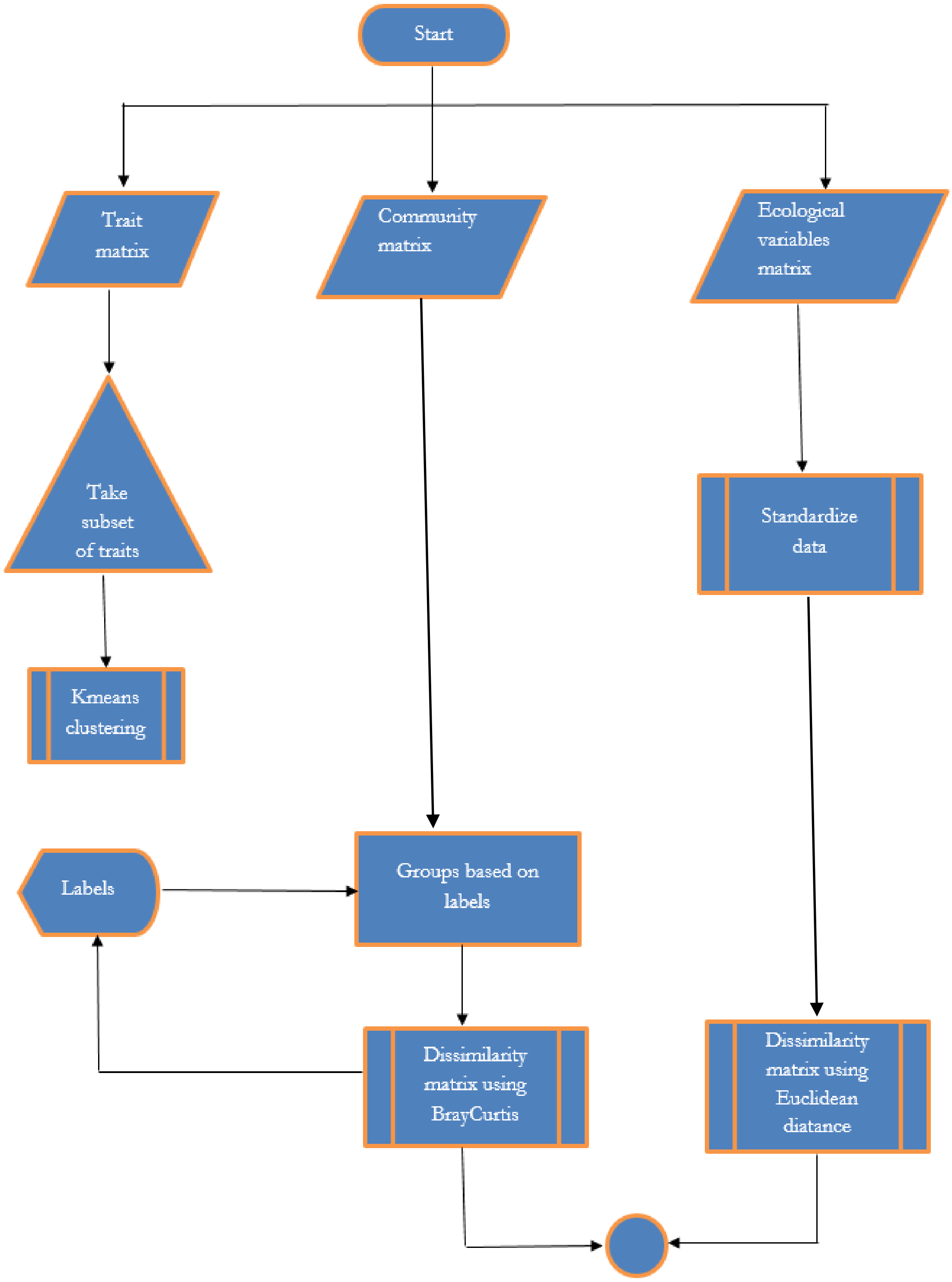
Step6: Takes Ecological variables (E) as input matrix and computes dissimilarity matrix (D_E) using Euclidean distance.
D_E<- Euclidean distance (E)

Step7: Calculates Pearson correlation coefficient $(Dx; D_E)$.
Co<-Correlation($Dx; D_E$)

Step8: Repeats Step3 to Step7

Step9: Finds the maximum correlation coefficient value and corresponding traits indicating PFTs maximally associated with environmental variables.
max(Co), S, X
PFT<-X

Flow chart of algorithm



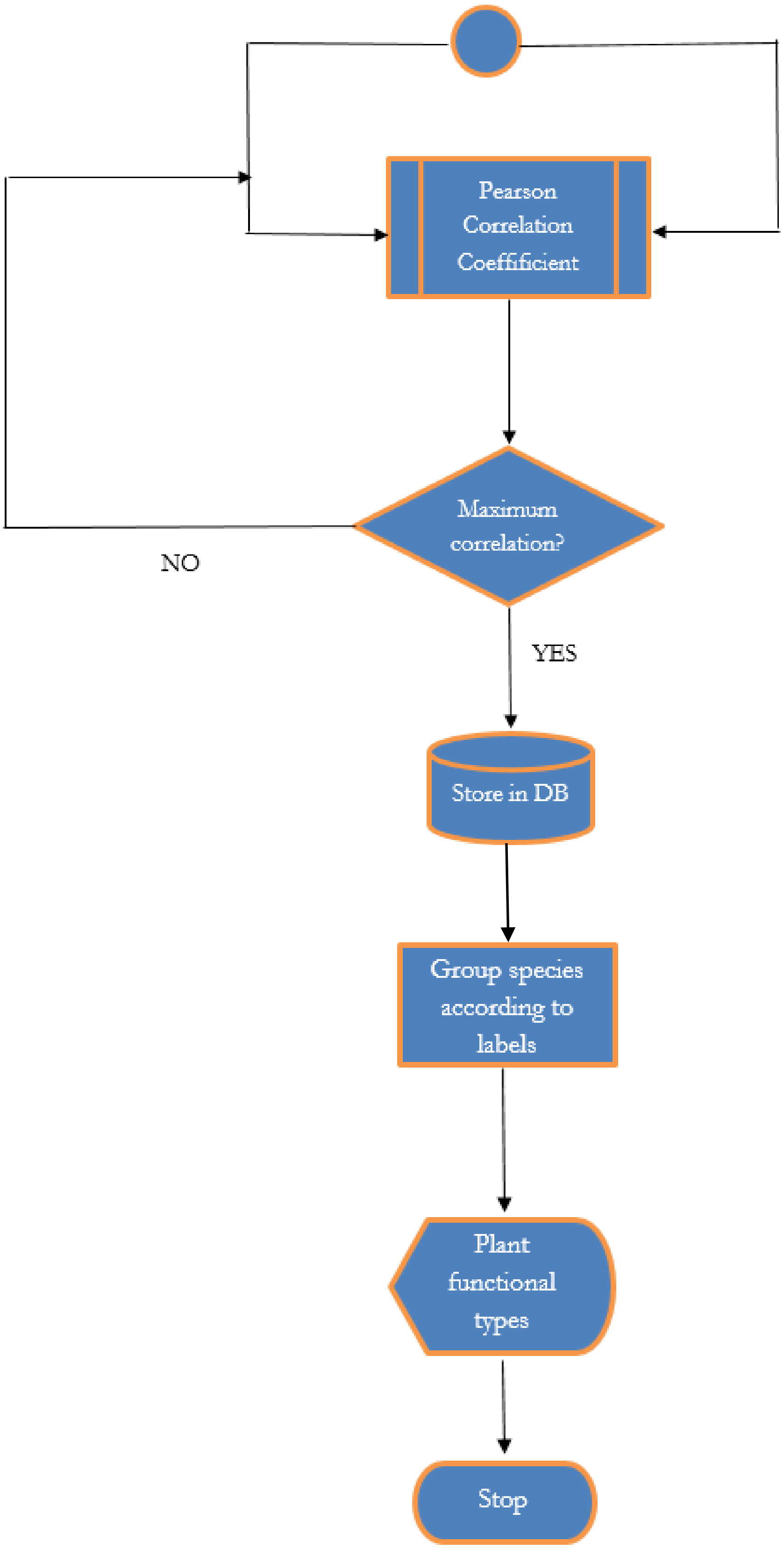


Figure 4-7 Flowchart of Algorithm

4.2.2.1. Inputs to the Algorithm

i) Traits matrix (T)

Species were described by traits in this matrix. Trait matrix was defined as

T=Species x Traits

$$\begin{array}{l} \\ \text{Sp1} \\ \text{Sp2} \\ . \\ . \end{array}\begin{array}{c} \text{Trait1, Trait2 ,Trait3, Trait4} \\ \left[\begin{array}{c} \quad \\ \quad \\ \quad \\ \quad \end{array}\right]_{nxn} \end{array}$$

Figure 4-8 Trait Matrix

The above trait matrix was created for actual and potential species which is shown below.

Actual species

The actual species are those currently existing in the plot 1-Him-54. The dominant species of the plot is shown below.

	species text	speciesid bigint	plantheghtid integer	leaflengthid integer	leafareaid integer	plantvegphenologyid integer
1	Oxalis corniculata L.	1	5	1	1	2
2	Pinus roxburghii Sargent	2	6	6	0	2
3	Trifolium repens L.	3	2	1	1	2
4	Poa annua L.	4	5	2	1	1
5	Fragaria nubicola Lindley ex Lacatia	5	4	1	0	2
6	Cedrus deodara (Roxb. ex Lambert.) G.Don.	6	6	1	0	1

Figure 4-9 Trait Matrix for Actual species

Potential species

The potential species are the species which were supposed to be present in the selected plot but have changed over the course of time due to human impact, anthropogenic factors, climate change etc. The potential species were taken for Himalayan moist temperate class of Himachal Pradesh from Champion and Seth classification.

	species character varying	speciesID [PK] integer	Plantheight integer	Leaf length integer	Leaf area integer	PlantVegetativePhenology integer
1	Cedrus deodara	1	6	1	0	1
2	Abies pindrow	2	6	1	0	0
3	Pinus wallichiana	3	6	5	0	2
4	Arundinaria falcata	4	6	2	1	0
5	Rhododendron arboreum	5	6	3	3	1
6	Litsea umbrosa	6	6	2	3	1
7	Lyonia ovalifolia	7	6	4	0	2
8	Cornus	8	6	5	0	3
9	Ilex dipyrena	9	6	2	0	2
10	Myrica	10	6	2	2	2
11	Castanopsis	11	6	3	6	2
12	Acer	13	6	2	6	1
13	Betula	14	6	2	3	1
14	Corylus	15	6	3	5	1
15	Ulmus	16	6	2	2	0
16	Pinus roxburghii	17	6	6	1	2
*						

Figure 4-10 Trait Matrix for Potential species

ii) Community matrix (C)

The community matrix was constructed by using actual vegetation type map. The actual species were considered as rows and vegetation type as column values. 12 classes for vegetation type were obtained for the study area. Community matrix describes the species composition in each vegetation type. It is defined as

C= Species x Communities

$$\begin{array}{ll} & \text{vegType1, vegType2, vegType3, vegType4.....} \\ \begin{array}{l} \text{sp1} \\ \text{sp2.} \\ . \\ . \end{array} & \left[\begin{array}{c} \\ \\ \\ \\ \end{array} \right]_{\text{nxn}} \end{array}$$

Figure 4-11 Community Matrix

Community matrix for actual species

speciesID [PK] integer	Tropical DDF integer	Himalayan moist temperate integer	Forest plantation integer	Orchards integer	Scrub/grassland integer	Sub tropical pine integer	Degraded forest integer	Himalayan dry temperate integer	Moist alpine scrub integer	Tropical MDF integer	Dry alpine scrub integer	Sub alpine forest integer
1	0	281	8	0	297	582	144	757	308	170	16	0
2	0	431	68	18	864	736	168	1996	310	344	21	0
3	0	11	91	0	2218	125	80	105	886	14	100	0
4	0	490	0	7	528	1124	340	1423	236	83	135	0
5	0	846	1	22	224	256	11	739	20	3	53	0
6	0	143	5	0	780	1135	14	69	0	100	0	0

Figure 4-12 Community matrix for Actual species

Community matrix for potential species

The community matrix for potential species was prepared by using potential natural vegetation type map. Here, potential species were taken as rows and vegetation type as columns. 11 classes were obtained for the study area. It is defined as

C= Species x Communities

Tropical Thorn forest integer	Himalayan moist temperate integer	Tropical dry evergreen forest integer	Tropical broadleaved integer	Sub-tropical dry evergreen integer	Tropical pine integer	Montane wet temperate forest integer	Himalayan dry temperate integer	Tropical MDF integer	Dry alpine s integer
0	70	0	0	0	0	0	70	0	0
0	2	0	0	0	0	0	0	0	0
0	2	0	0	0	0	0	15	0	0
0	2	0	0	0	0	0	0	0	0
0	2	0	0	0	10	70	0	0	0
0	2	0	0	0	0	0	0	0	0
0	2	0	0	0	10	0	0	0	0
0	2	0	0	0	0	0	0	0	0
0	2	0	15	0	0	0	0	0	0
0	2	0	15	0	10	0	0	0	0
0	2	0	0	0	0	0	0	0	0
0	0	0	0	0	0	15	15	0	0
0	2	0	70	0	0	15	0	0	0
0	2	0	0	0	0	0	0	0	0
0	2	0	0	0	0	0	0	0	0
0	0	0	0	0	70	0	0	0	0

Figure 4-13 Community Matrix for Potential Species

iii) Ecological variables matrix (E)

It describes the communities with environmental variables such as minimum temperature, maximum temperature, precipitation, altitude and soil type. It is defines as

E = Community sites x ecological variables

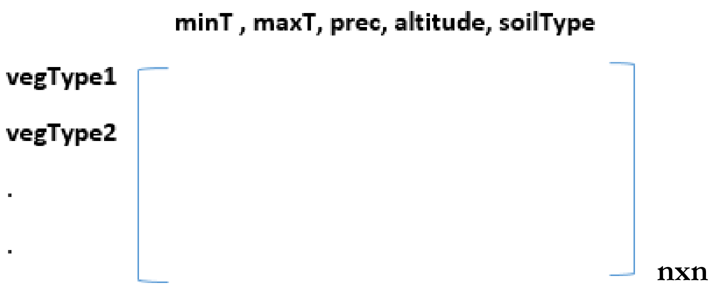


Figure 4-14 Ecological Variables Matrix

	veg_type character varying	CommunityID [PK] integer	mintemp numeric(6,2)	maxtemp numeric(6,2)	prec numeric(6,2)	altitude numeric(6,2)	soil_temp numeric(6,2)	soil_type integer
1	Tropical DDF	1	7.00	38.00	1398.00	284.00	4.00	118
2	Himalayan moist temperate	2	1.00	24.00	1394.00	2010.00	35.00	28
3	Forest plantation	3	9.00	38.00	1500.00	214.50	51.50	118
4	Orchards	4	7.00	38.00	1349.00	264.00	100.00	48
5	Scrub/grassland	5	-1.50	26.70	1373.00	1909.00	29.70	29
6	Sub tropical pine	6	1.00	28.00	1637.00	1402.00	35.00	20
7	Degraded forest	7	10.00	41.00	818.00	414.00	3.00	20
8	Himalayan dry temperate	8	-1.00	24.00	1141.00	1951.50	3.00	12
9	Moist alpine scrub	9	-21.00	10.00	119.00	4463.00	100.00	48
10	Tropical MDF	10	9.00	38.00	1626.00	526.00	3.00	1652
11	Dry alpine scrub	12	-9.00	15.00	932.00	3600.00	100.00	20
12	Sub alpine forest	13	1.00	23.00	1529.00	2427.00	3.00	20

Figure 4-15 Ecological Variables Matrix (1)

4.2.2.2. K-means cluster analysis

The next step in the development of the algorithm is the cluster analysis. Here it was performed using K-means clustering. Below is the description about the clustering method.

A cluster is defined as set of data points which are associated close to each other. Each cluster has its centroid. K-means algorithm initially chooses set of centroids randomly. Then, each data point is assigned to the centroid which is closest to form set of clusters. For each cluster, new centroid is calculated by considering mean value along each dimension of data that belong to the cluster and new centroids are formed. The data points are assigned to new centroids following the same procedure as before. This continues till centroids don't change (Thilina Hewaragamunige, 2014).

Consider a data set $P = \{x1, x2, x3, x4, .. xn\}$ that has to be clustered. K-means can be defined as an objective function which depends on distance of data points to the centroid of the cluster. It is given as,

$$\min_{\{m_k\},1\leq k\leq K} \sum_{k=1}^{K} \sum_{x\in C_k} \pi_x \operatorname{dist}(x, m_k)$$ Equation 4-1

In the above equation,

- ✓ π_x is the weight belonging to x and x belong to C_k,
- ✓ the total number of data objects given to the cluster C_K is given as m_k
- ✓ K is the total number of clusters set by the user
- ✓ "dist" measures the distance between the object and the centroid m_k, $1 \leq k \leq K$

K-means cluster analysis was applied on each subset of traits. The question that arises when using K-means is value of K. Based on trial and error method, K value was assigned to 3. The algorithm generated 3 centroids for each trait subset. The labels for each centroid were named as 0, 1 and 2 respectively.

4.2.2.3. Dissimilarity matrix of communities

A community is defined as a set of species present in the same habitat. Comparing communities help to understand the relationship between variables that define species and communities (Pavoine, Dufour, & Chessel, 2004). The community matrix describes the species composition as presence and absence of species.

Dissimilarity is defined as the relationship between different community sits. It explains how the species composition is differed between sites (De'ath, 1999). There are several dissimilarity measures like Kendall, Manhattan, Gower metric, Euclidean, Bray-Curtis, Chord, Chi-squared etc. Out of all the dissimilarity measures, Bray-Curtis, Manhattan and Kulczynski measures are related to ecological distance monotonically and proportional to smaller ecological distance (Faith, Minchin, & Belbin, 1987). Bray-Curtis dissimilarity measure is expressed as,

$$\sum_i (X_{ij} - X_{ik}) / \sum_i (X_{ij} + X_{ik})$$ Equation 4-2

i & j – two different community sites

dij - similarity between i & j

xi & xj – set of n attributes which defines i & j

The community matrix was grouped according to species grouped in traits matrix to generate X (PFTs). For each X matrix, Bray-Curtis dissimilarity measure (Dx) was calculated. The ecological variables matrix was standardized and then distance matrix (D_E) was calculated for the same X matrix using Euclidean distance. It is given by,

$$[\sum_i (X_{ij} - X_{ik})^2]^{\frac{1}{2}}$$ Equation 4-3

Pearson Correlation Coefficient

According to Zou, Tuncali, & Silverman (2003), the linear or non-linear i.e. polynomial, logistic, exponential relationship between any two continuous variables can be analyzed with the help of correlation. Pearson correlation coefficient measures linear relationship between two variables. The correlation coefficient value ranges from -1 to +1 where in -1 represents negative correlation, 0 as no correlation and +1 positive correlation. If Xi and Yi are two bivariate data where $i = 1,2,3 \dots n$, then Pearson correlation coefficient, r can be computed by using the below formula

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}}$$ Equation 4-4

$\overline{X}$, $\overline{Y}$ are the sample means of Xi and Yi (Zou et al., 2003).

The Pearson correlation coefficient was calculated between dissimilarity matrix of communities computed using X and dissimilarity matrix of communities computes using ecological variables.

For the above algorithm development following python modules were used.

4.2.3. Design and Development of GeoWeb Application

Development in the web and also the internet has provided a wide range of opportunities for the geoscientists for analysing and visualizing the GIS data on web. For the development of this Geoweb application, following tools were used.

Table 4-2 GeoWeb development tools

Database Server	PostgreSQL
Spatial Database	Postgis
Web Framework	Geodjango in Python
Web server	Django Development Server
GIS Server	Geoserver
JavaScript Library	Leaflet
User Interface	HTML5, CSS

Below is the detailed explanation about each of the tools used.

PostgreSQL

For this research, PostgreSQL object relational database is used as it supports postgis. It is an open source database under PostgreSQL License. It can be run in Windows, Linux, UNIX, Mac operating systems. The architecture has earned reputation for its data integrity, correctness and reliability. It is completely ACID compliant and supports joins, stored procedures, triggers, views, foreign keys including different data types such as NUMERIC, CHAR, BOOLEAN, INTEGER, INTERVAL, DATE, TIMESTAMP and VARCHAR. Ability to store images, video which are large binary objects is yet another functionality of PostgreSQL. It consists of programming interfaces for Python, Perl, .Net, Java etc. ("PostgreSQL," n.d.)

Postgis Spatial Database

It is an extender of PostgreSQL object-relational database that allows to perform spatial queries. It supports spatial objects such as geography, geometry and raster. It makes PostgreSQL rich in features, fast and robust by adding operators, indexes and functions to spatial types. It includes major functionalities like

- Supports creation and conversion of geometry, buffer, generalization, convex hull, reprojection etc.
- Importing and rendering of different vector data formats such as GML, KML, GeoJSON, Shapefile and raster data formats like JPEG, GeoTIFF, PNG using SQL ("Boundless," n.d.)
- High speed spatial querying with the help of R tree over Generalized Search Tree (GiST) spatial index.
- It supports third party applications like geoserver, mapserver tec.

Geodjango Python Web Framework

The algorithm was developed in python and to integrate a python code in web application, a web framework in python is required. One such framework is Geodjnago. It is a geographic web framework which extends Django python web framework. Django is an MVC (Model, View and controller) pattern where in models describes the data structure/database schema, views controls what user sees and controller describes the framework and URL parsing.

Geodjango makes building of webgis applicaltions as simple as possible. Some of the important features of Geodjango are

- Support for Open Geospatial Consortium (OGC) geometries like *Point, MultiPoint, LineString, Polygon, MultiPolygon, MultiLineString, Geometry Collection* in Django model fields.
- Django's Object Relational Mapper (ORM) helps to query and manipulate spatial data.
- It is loosely coupled and consists of python interfaces for various operations of geometry operations and data formats
- Provides functionality to edit geometry attributes inside admin ("Geodjango Tutorial," n.d.)

Leaflet

Leaflet is a javascript library for mobile and desktop web mapping applications developed by Vladimir Agafonkin along with other developers. It consists of numerous plugins which supports GIS formats for web and simple to use (Sanchez, n.d.). It provides responsive web application. As the GeoWeb application was developed as a desktop web mapping application, it can further be extended as a mobile application using leaflet.

HTML5 and CSS

HTML5 is the new version of standard HTML which consists of new attributes, elements and behaviors along with several functions like Semantics, Connectivity, Multimedia, 2D/3D graphics, Performance, Device access etc.
CSS is a stylesheet language for beautifying the HTML pages. It describes how the HTML document must be presented to user.

GeoServer

Geoserver is a software server which is open source and allows users to edit and share GIS data. It is written in Java. It is designed for interoperability can take multiple formats of GIS data and publish as a service using protocols compliant to OGC standards such as Web Map Service (WMS), Web Feature Service(WFS), Web Coverage Service (WCS) ("GeoServer," n.d.). It has the ability to connect OpenLayers, Bing Maps and Google Maps that are web based maps with the existing information.

4.2.3.1. Architecture

Below diagram shows the architecture of the GeoWeb application.

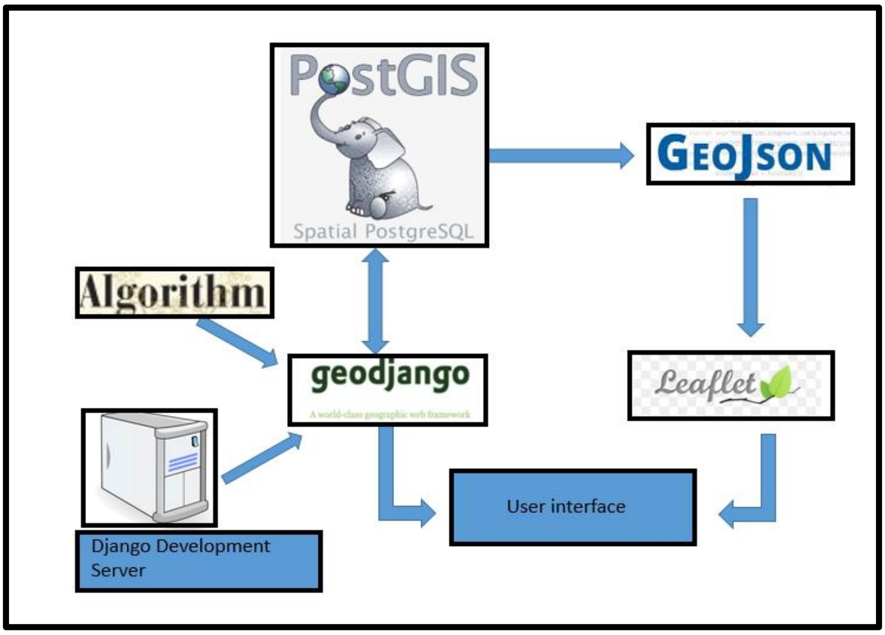


Figure 4-16 Web Architecture

GeoDjango Installation and Connection with Postgis

The Geodjango python web framework comes along with Django. It runs in Django development server. Django was installed through *pip* in command line

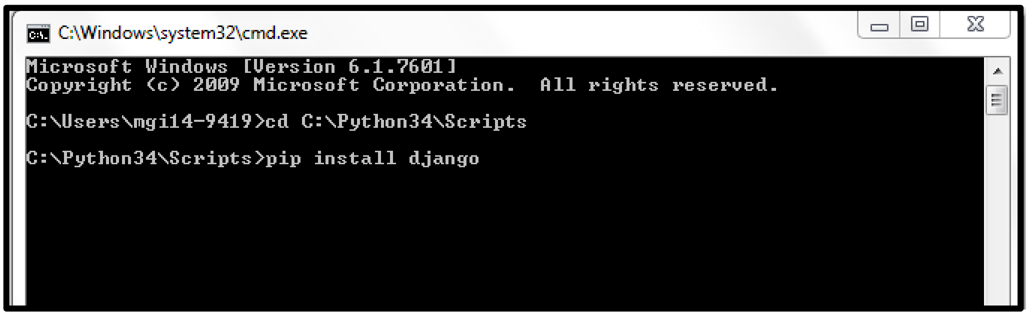


Figure 4-17 Geodjango installation

Once Django gets installed, it creates a file called django-admin.py. A project must be created and it can have any number of applications. To create a project and application the following command was used,

```
django-admin.py startproject geodjango
cd geodjango
python manage.py startapp PFT
```

The generated folder structure was,

```
geodjango
   |----geodjango
   |       |----settings.py
   |       |----urls.py
  -PFT     |
          -views.py
          -templates
```

To connect Postgis to Geodjango, database settings was configured in settings.py file of the project as below,

```
DATABASES = {
    'default': {
        'ENGINE': 'django.contrib.gis.db.backends.postgis',
        'NAME': 'geodjango',
        'USER':'postgres',
        'PASSWORD':'postgres',
        'HOST':'localhost',
        'PORT':'5433',
    }
}
```

Figure 4-18 Database connection

Model, View and Template (MVT)

A **model** is a single data source which contains fields of data and behavior of the data stored. Usually, a single database is mapped to each model. A model is a subclass of *django.db.models.Model* and database field is represented by each attribute of the model. Models for actual PFTs, potential PFTs, file upload were created.

To map the created models to database the following commands were used

```
python manage.py makemigrations
python manage.py migrate
```

Example of a model of actual PFTs and the corresponding table in database is shown below,

```
class Existing_Data_APFT(models.Model):
    #spid = models.AutoField(primary_key=True)
    species=models.CharField(max_length=255)
    x=models.FloatField()
    y=models.FloatField()
    geom = models.PointField(srid=4326)

    rastvalues=models.IntegerField()
    listvalues=models.IntegerField()
    objects=models.GeoManager()


    def __str__(self):
        return "%s %s %s %s %s" % (self.species, self.geom.x, self.geom.y, self.rastvalues, self.listvalues)
```

id [PK] serial	species character varying(255)	x double precision	y double precision	geom geometry(Point,4326)	listvalues integer	rastvalues integer
1	Oxalis corniculata L.	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0	0
2	Pinus roxburghii	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	1	0
3	Trifolium repens L.	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0	0
4	Poa annua L.	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	2	0
5	Fragaria nubicola Lindley ex Lacatia	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0	0
6	Cedrus deodara	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0	0

Figure 4-19 Django Model

A **view** is a function that takes HTTP request and returns HTTP response like web page contents, a 404 error, XML document, image etc. It contains the logic required to return the response. The request object is always the first argument in a view function ("Writing Views," n.d.). The logic for getting data from database as geojson data in order to display on map was written for actual PFTs and potential PFTs. The logic for user login, registration, user uploaded data and change in PFTs were written as view functions

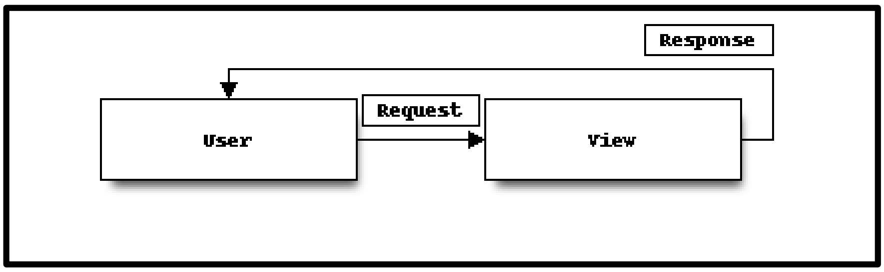


Figure 4-20 Django View ("Writing Views," n.d.)

A **template** separates the application logic from presentation layer. It is an engine to define the layer presented to user. A template's layout can be used any number of times. Another functionality supported by template is template extending where different web pages can use same HTML parts ("Template Extending," n.d.). A template folder consists of HTML and CSS files.

When user sends request, the corresponding view function takes the request, and gives response to the user by rendering through the corresponding template file. An example of a view rendering template file is shown below.

```
def index(request):
    return render_to_response("index.html")
```

Figure 4-21 Django view rendering template.

4.2.4. Integration of Algorithm in Web Application

The developed algorithm to identify actual and potential PFTs must be integrated in the web application in order to view the results on web. Geodjango framework makes integration of python code very simpler. Any python code can be called inside a view function with the help of Popen() function.

```
p = Popen(["python", "labels_to_db.py"], cwd=script_dir, stdout=PIPE, stderr=PIPE)
```

Figure 4-22 Integration of algorithm in web

The working of GeoWeb application is summarized below,

i) User request is taken
ii) Checks the URL name and searches the name in urls.py
iii) Once the URL name is found, searches the corresponding view function in views.py
iv) Executes the view function and renders the corresponding template.

4.2.5. VALIDATION

Validation is one of the important step in any research. In this research the validation was performed in the following way.

The algorithm was run on four plots to identify the actual PFTs and potential PFTs. However validation must be done on potential PFTs since the actual PFTs are ground data which were obtained from 'Biodiversity Characterization at Landscape Level' project. To identify the change in potential and actual PFTs, 75% of the plots consisting of dominant species were taken that belong to Himalayan moist temperate class and 16 potential species of Himalayan moist temperate class from Champion and Seth classification. This result was validated with the remaining 25% of the plots that belong to Himalayan moist temperate class.

5. RESULTS AND DISCUSSION

This chapter is about the results obtained using above adopted methodology and methods, and detail discussions on it

5.1. Results

5.1.1. Algorithm Results for Actual Plant Functional Types

The below results are shown for one plot of Himachal Pradesh of western Himalaya. The species in this plot consisted of Oxalis corniculata, Pinus roxburghii, Trifolium repens, Poa annua, Fragaria nubicola and Cedrus deodara.

5.1.1.1. K-means clustering

The K-means clustering was applied for each trait subset. The total number of trait subsets obtained were 15. The plot shows the grouping of species based on traits. Each cluster in the plot is denoted by specific colour. The clusters were plotted using matplotlib library in python. The scatter plot for one trait subset obtained is shown below. The result for all trait subset is shown is appendix.

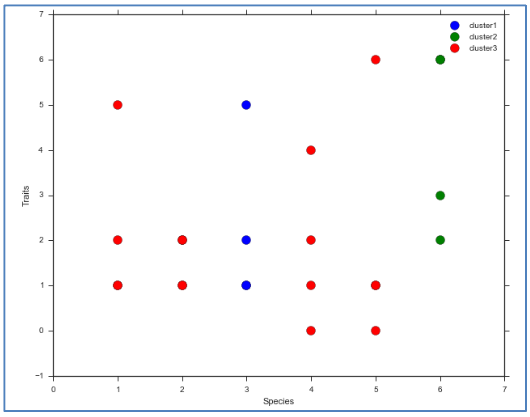


Figure 5-1 – Scatterplot of species and trait subset - Plant height, Leaf length, Leaf area, Plant phenology

5.1.1.2. Community Dissimilarity matrix

The community matrix was grouped according to the clusters of trait subset and 15 resultant community matrices were obtained. In case of actual species there were total of 12 communities. The dissimilarity matrix for communities were calculated using 'Bray-Curtis' dissimilarity measure. The result for one community matrix is shown below. The result for all community matrix is shown in appendix.

The communities consisted of the following vegetation classes

Table 5-1: List of communities for actual PFT

Communities for Actual PFT
Tropical dry deciduous forest
Himalayan moist tempearte
Forest plantation
Orchard
Scrub/grassland
Sub tropical pine
Degraded forest
Himalayan dry tempearte
Moist alpine scrub
Tropical MDF
Dry alpine scrub
Sub alpine forest

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8792321									
4	1.0000000	0.9606701	0.5859031								
5	1.0000000	0.3378428	0.9383464	0.9803386							
6	1.0000000	0.3074350	0.9341579	0.9789720	0.3097614						
7	1.0000000	0.4209554	0.7275591	0.9061685	0.6643207	0.6449264					
8	1.0000000	0.4150369	0.9482423	0.9835516	0.3664129	0.1233397	0.7116572				
9	1.0000000	0.5332765	0.8404795	0.9473171	0.4630800	0.5736084	0.3579049	0.6088387			
10	1.0000000	0.5431356	0.6432990	0.8730905	0.7443874	0.7288655	0.1583949	0.7819127	0.4292875		
11	1.0000000	0.7081218	0.4533965	0.7898833	0.8440414	0.8339949	0.4091201	0.8680436	0.6254072	0.3508353	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(o)

Figure 5-2: Community dissimilarity matrix for clusters of one trait subset-Plant height, Leaf length, Leaf area, Plant phenology

5.1.1.3. Ecological variables Dissimilarity matrix

The 12 communities were compared based on environmental variables such as minimum temperature, maximum temperature, altitude, precipitation and soil type using 'Euclidean Distance'. The obtained result is shown below

	1	2	3	4	5	6	7	8	9	10	11
2	1.9041448										
3	0.3119941	1.9891029									
4	2.0692109	0.3287696	2.1204194								
5	1.9037930	0.1884854	1.9918689	0.4417695							
6	0.7696283	1.5586445	0.8133481	1.6316467	1.5965278						
7	2.5542711	1.4276569	2.6337213	1.6222538	1.3083737	2.5965528					
8	0.5717481	2.1378706	0.6266615	2.3410011	2.1067792	1.2914427	2.4777711				
9	0.6378446	2.2743567	0.4433358	2.3656931	2.2928791	0.9220713	2.9964272	0.9877272			
10	2.1223095	3.9116424	2.1198850	4.1181596	3.8858238	2.8375187	4.0627746	1.7997278	2.1188004		
11	4.0184527	3.3475022	4.0393129	3.3142088	3.5149575	3.8035018	3.9810250	4.2288860	4.1490728	5.3848898	
12	4.3890700	5.9414590	4.4444503	6.2052961	5.8917870	5.1325042	5.7112499	3.9710615	4.5215297	2.4397809	7.1383039

Figure 5-3: Dissimilarity matrix of Communities based on Environmental variables

5.1.1.4. Pearson Correlation Coefficient

The correlation coefficient between community dissimilarity matrix and environmental factors dissimilarity matrix was calculated using Pearson correlation coefficient. The result is shown below

Table 5-2: Correlation between community and environmental dissimilarity matrix and corresponding trait subset

Trait Subsets	Correlation
Plant height	0.03
Leaf length	0.05
Leaf area	0.06
Plant Phenology	0.07
Plant height, leaf length	0.07
Plant height, Leaf area	0.11
Plant height, Plant phenology	0.11
Leaf Length, Leaf area	0.03
Leaf length, Plant phenology	0.05
Leaf area, Plant phenology	0.06
Plant height, Leaf length, Leaf area	0.06
Plant height, Leaf length, Plant phenology	0.11
Plant height, Leaf area, Plant phenology	0.06
Leaf length, Leaf area, Plant phenology	0.06
All traits	0.12

The maximum correlation coefficient was obtained was 0.12 and the corresponding trait subset were Plant height. Leaf length, Leaf area and Plant phenology. The corresponding labels of the species and trait clusters were mapped to the database table and accessed through Django view and template to display the resulting plant functional types on web.

```
[0, 2, 0, 0, 0, 1]
```

id [PK] serial	species character varying(255)	x double precision	y double precision	geom geometry(Point,4326)	listvalues integer
1	Oxalis corniculata L.	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0
2	Pinus roxburghii	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	2
3	Trifolium repens L.	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0
4	Poa annua L.	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0
5	Fragaria nubicola Lindley ex Lacatia	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	0
6	Cedrus deodara	76.320083	32.233167	0101000020E61000005B40683D7C145340D3C08F6AD81D4040	1

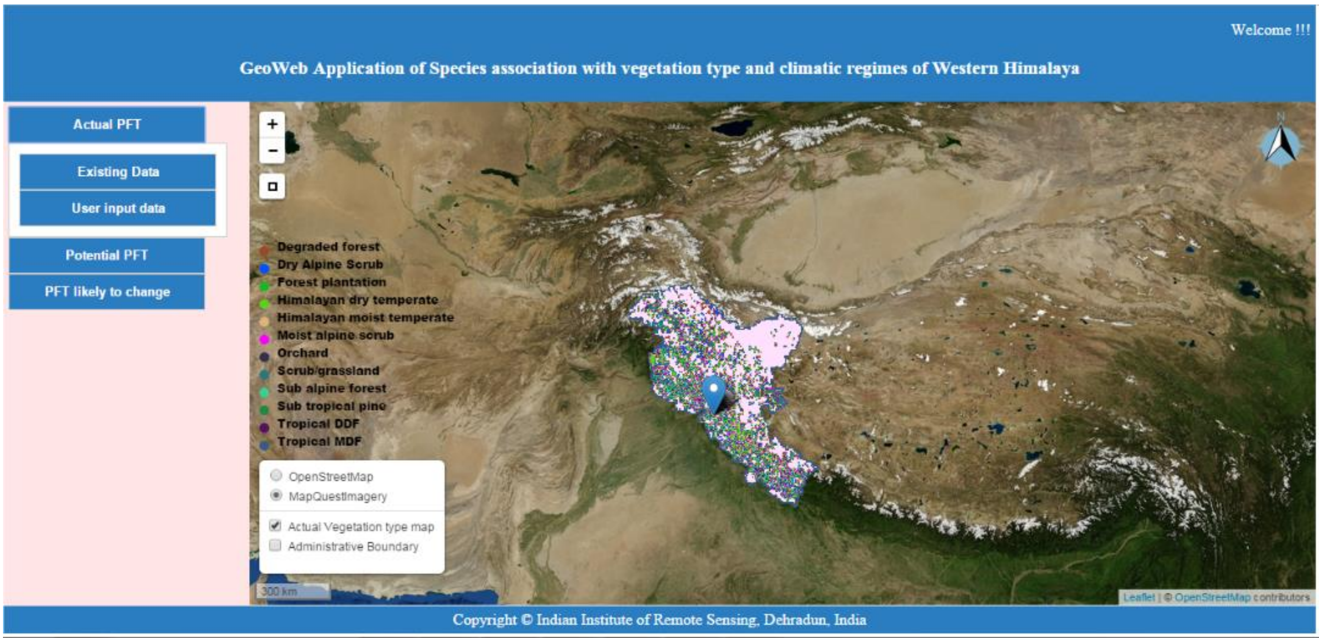


Figure 5-4 One plot overlaid on actual vegetation type map

The same methodology was implemented for three more plots of Himachal Pradesh. The plant functional types was shown on a pie chart as a pop up for each plot. The colours of pie chart indicate the different plant functional types of the corresponding plot. The different PFTs of one plot is shown below.

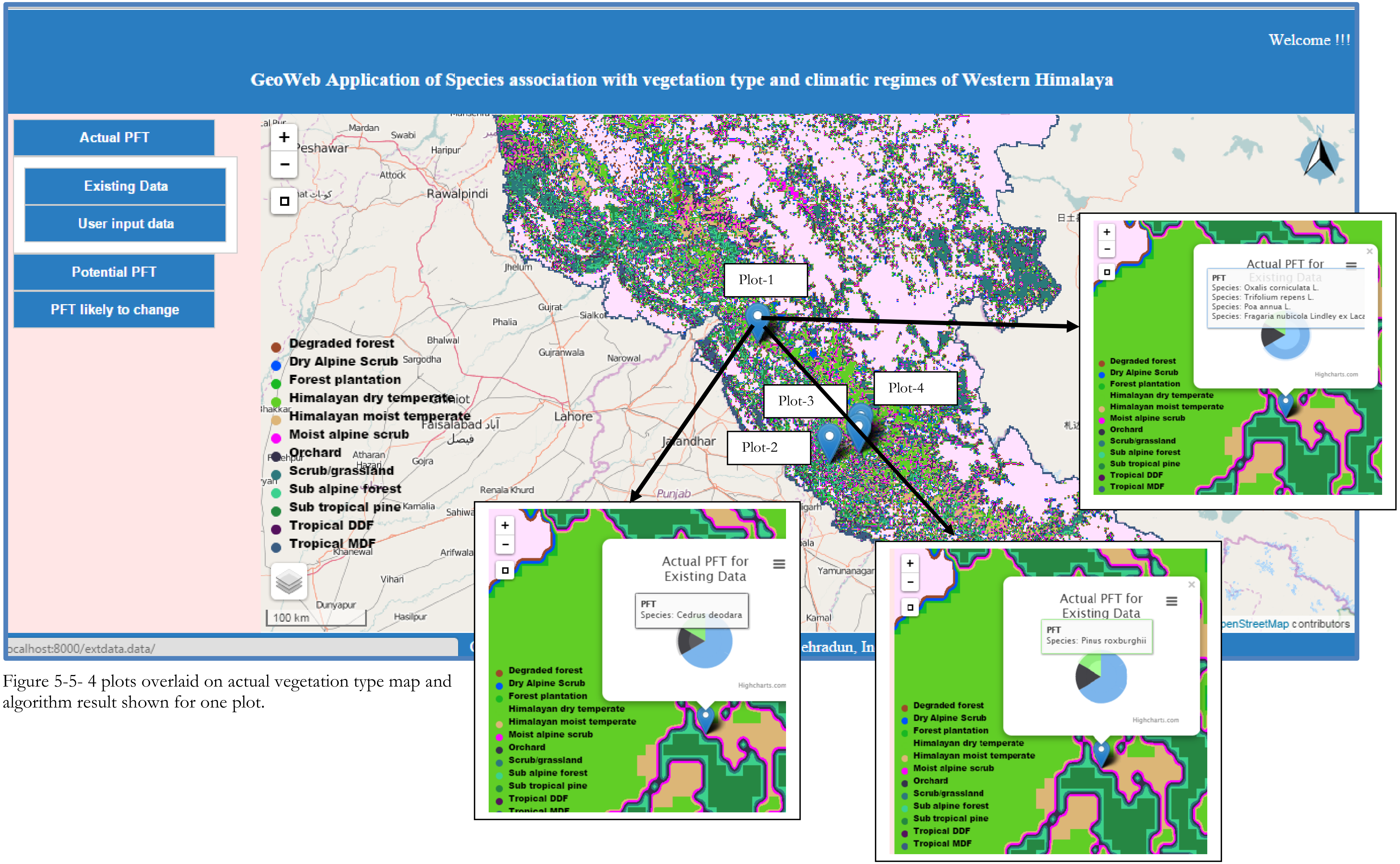


Figure 5-5- 4 plots overlaid on actual vegetation type map and algorithm result shown for one plot.

The results for all four plots is shown below

Table 5-3 – Actual PFTs for four different plots

Plot number	PFT 1	PFT 2	PFT 3
1-Him-54	**Cedrus deodara**	**Pinus roxburghii**	**Fragaria nubicola Lindley ex Lacatia, Poa annua L, Trifolium repens L., Oxalis corniculata**
1-Him-506	**Cedrus deodara**	**Pinus wallichiana**	**Oxalis corniculata, Pinus roxburghii**
1-Him-515	**Pinus wallichiana**	**Cedrus deodara, Oxalis corniculata, Pinus roxburghii, Taraxacum officinale**	
1-Him-490	**Cedrus deodara**	**Trifolium repens L.,**	**Oxalis corniculata, Fragaria nubicola Lindley ex Lacatia**

5.1.2. Algorithm results for Potential Plant Functional Types

The same procedure was followed as above for identifying the potential plant functional types. The below results are shown for Himachal Pradesh of western Himalaya. The dominant species of Himachal Pradesh were taken from Champion & Seth classification. The potential species consisted of Cedrus deodara, Abies pindrow, Pinus wallichiana, Arundinaria falcate, Rhododendron arboretum, Litsea umbrosa, Lyonia ovalifolia, Cornus, Ilex dipyrena, Myrica, Castanopsis, Acer, Betula, Corylus, Ulmus and Pinus roxburghii. The results for each method is shown below.

5.1.2.1. K-means clustering

The K-means was performed similarly as performed for actual species. The result is shown for one trait subset. The result for all trait subset is shown in appendix.

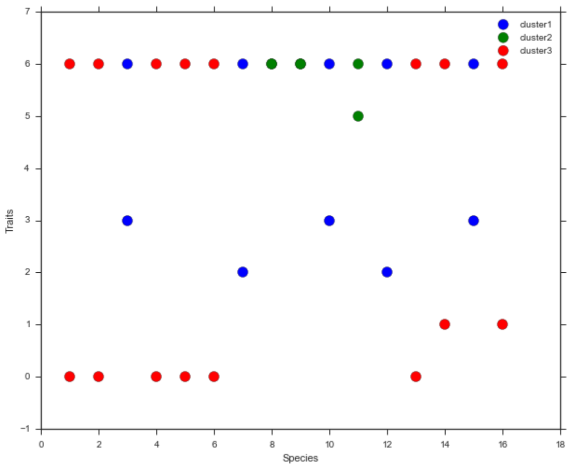


Figure 5-6: Scatterplot of species and trait subset -Plant height, Leaf area.

5.1.2.2. Community Dissimilarity matrix

In case of potential species, there were total of 11 communities. The dissimilarity matrix for communities were calculated using 'Bray-Curtis' dissimilarity measure. The result is shown for one community matrix and rest is shown in appendix. The communities consisted of the following:

Table 5-4: List of communities for potential PFT

Communities for Potential PFT
Tropical thorn forest
Himalayan moist temperate
Sub-tropical dry evergreen
Sub-tropical broadleaved
Sub-tropical dry evergreen
Sub-tropical pine
Montane wet temperate forest
Himalayan dry temperate
Tropical moist deciduous forest
Dry alpine scrub
Tropical wet evergreen

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	0.58333333	1.00000000	1.00000000	0.62500000						
6	1.00000000	1.00000000	1.00000000	0.86666667	0.89090909					
7	0.06666667	1.00000000	1.00000000	1.00000000	0.62264151	1.00000000				
8	0.64705882	1.00000000	1.00000000	1.00000000	0.26829268	1.00000000	0.68421053			
9	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000		
10	0.17647059	1.00000000	1.00000000	0.88679245	0.58730159	0.83695652	0.11111111	0.73913043	1.00000000	
11	0.47826087	1.00000000	1.00000000	0.76470588	0.26760563	0.76744186	0.52000000	0.50000000	1.00000000	0.37931034

Figure 5-7: Community dissimilarity matrix for clusters of one trait subset. (f)-Plant height, Leaf area

5.1.2.3. Ecological variables Dissimilarity matrix

The 11 communities were compared based on environmental variables such as minimum temperature, maximum temperature, altitude, precipitation and soil type using 'Euclidean Distance'. The obtained result is shown below.

	1	2	3	4	5	6	7	8	9	10
2	4.2031603									
3	3.4237296	2.7085605								
4	3.2803805	2.5121159	0.6501869							
5	4.0815965	0.4387268	2.7175699	2.4576830						
6	3.7541597	1.1675577	2.1511862	2.1361763	1.1446900					
7	4.0048399	3.5638414	2.5061916	2.0223047	3.5136564	3.7679824				
8	4.3390270	0.9762774	3.1124133	2.7270156	0.8757634	1.9609978	3.2643948			
9	3.4847066	2.6500755	0.5400076	1.0866460	2.6812230	1.9553228	2.9524005	3.1997338		
10	3.3102337	2.6577839	0.5775238	1.0873430	2.6809404	1.9440527	2.9574434	3.2035992	0.1811995	
11	6.1450069	2.9662098	5.4881208	5.1020135	2.8990609	3.9658209	5.2887734	2.3983010	5.5213261	5.5242071

Figure 5-8 - Dissimilarity matrix of Communities based on Environmental variables

5.1.2.4. Pearson Correlation Coefficient

The correlation coefficient between community dissimilarity matrix and environmental factors dissimilarity matrix was calculated using Pearson correlation coefficient. The result is shown below

Table 5-5: Correlation between community and environmental dissimilarity matrix and corresponding trait subset

Trait Subsets	Correlation
Plant height	**-0.07**
Leaf length	**-0.05**
Leaf area	**-0.12**
Plant Phenology	**-0.06**
Plant height, leaf length	**-0.06**
Plant height, Leaf area	**-0.13**
Plant height, Plant phenology	**-0.06**
Leaf Length, Leaf area	**-0.11**
Leaf length, Plant phenology	**-0.12**
Leaf area, Plant phenology	**-0.09**
Plant height, Leaf length, Leaf area	**-0.07**
Plant height, Leaf length, Plant phenology	**-0.12**
Plant height, Leaf area, Plant phenology	**-0.09**
Leaf length, Leaf area, Plant phenology	**-0.07**
All traits	**-0.07**

The maximum negative correlation i.e. -0.13 and the corresponding trait subset were Plant height and Leaf area. The corresponding labels of the trait clusters were mapped to the database table and accessed through Django view and template to display the resulting plant functional types on web.

```
[0, 0, 1, 0, 0, 0, 1, 2, 2, 1, 2, 1, 0, 1, 1, 0]
```

id [PK] serial	species character varying(255)	x double precision	y double precision	listvalues integer	rastvalues integer	geom geometry(Point,4326)
1	Cedrus deodara	76.320083	32.233167	0		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
2	Abies pindrow	76.320083	32.233167	0		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
3	Pinus wallichiana	76.320083	32.233167	1		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
4	Arundinaria falcata	76.320083	32.233167	0		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
5	Rhododendron arboreum	76.320083	32.233167	0		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
6	Litsea umbrosa	76.320083	32.233167	0		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
7	Lyonia ovalifolia	76.320083	32.233167	1		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
8	Cornus	76.320083	32.233167	2		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
9	Ilex dipyrena	76.320083	32.233167	2		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
10	Myrica	76.320083	32.233167	1		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
11	Castanopsis	76.320083	32.233167	2		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
12	Acer	76.320083	32.233167	1		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
13	Betula	76.320083	32.233167	0		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
14	Corylus	76.320083	32.233167	1		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
15	Ulmus	76.320083	32.233167	1		0101000020E61000005B40683D7C145340D3C08F6AD81D4040
16	Pinus roxburghii	76.320083	32.233167	0		0101000020E61000005B40683D7C145340D3C08F6AD81D4040

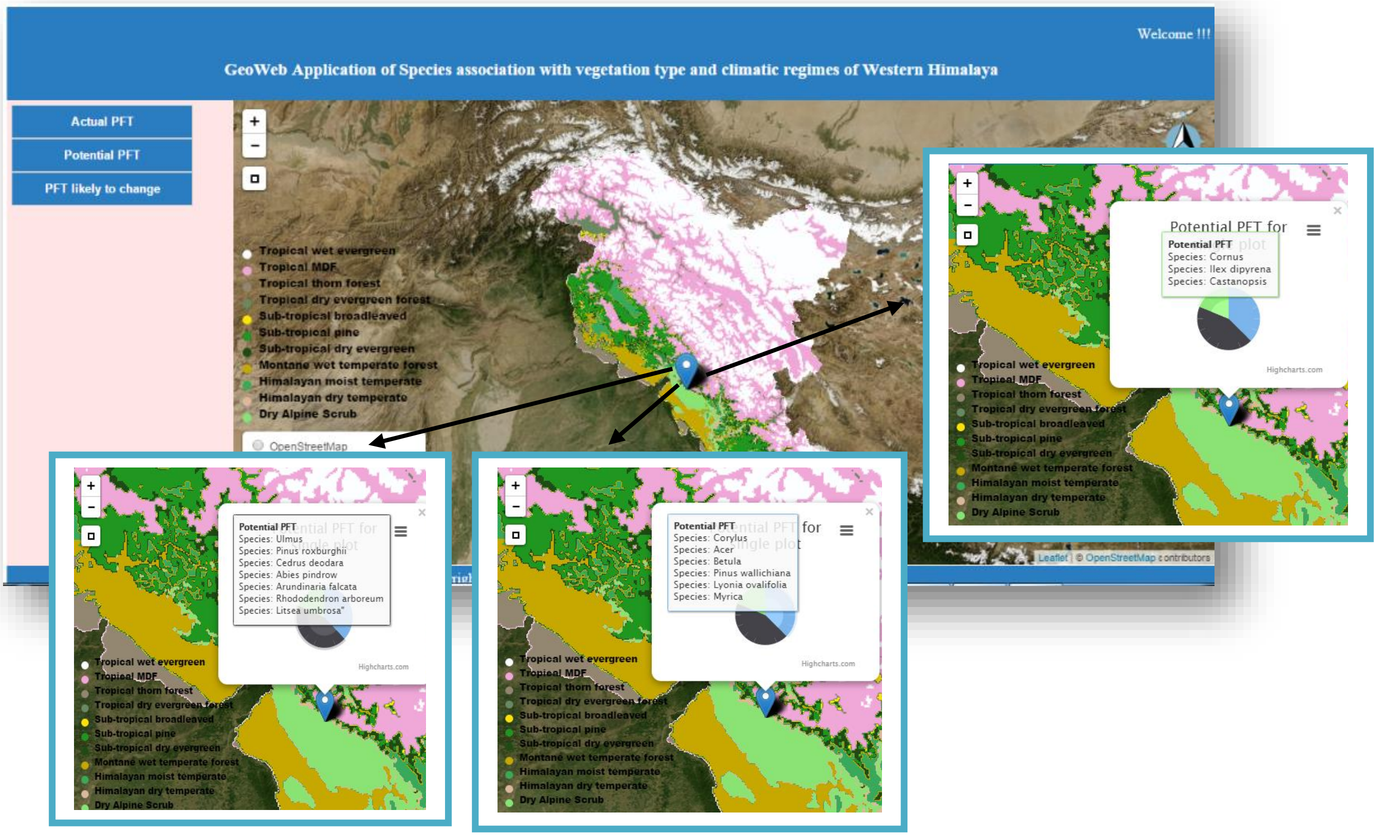


Figure 5-9- Marker indicating the potential PFTs of Himachal Pradesh overlaid on potential natural vegetataion type map.

Table 5-6: Potential plant functional types across Himachal Pradesh

PFT 1	PFT 2	PFT 3
Cornus	Corylus	Ulmus
Ilex dipyrena	Acer	Pinus roxburghii
Castanopsis	Betula	Cedrus deodara
	Pinus wallichiana	Abies pindrow
	Lyonia ovalifolia	Arundinaria falcata
	Myrica	Rhododendron arboreum
		Litsea umbrosa

5.1.3. Change in Plant Functional Types

The actual and potential plant functional types were compared for one plot of Himachal Pradesh and plant functional types likely to change was displayed as column graph using highcharts. The result showed 88% for PFTs likely to change and 12% for PFTs to remain unchanged. The result is shown below

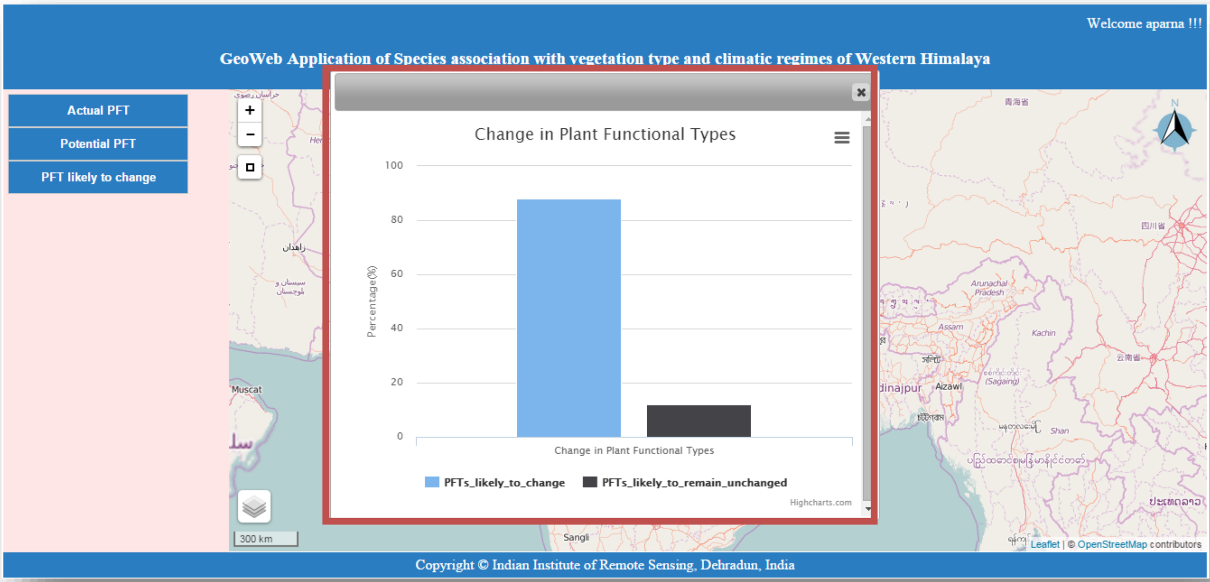


Figure 5-10: Change in plant functional types from potential to actual.

5.1.4. Functionalities of GeoWeb application

The various user functionalities provided in the GeoWeb application is shown below.

✓ Login Page

The registered users are allowed to access the login page. The new users are requested to register before accessing the web application.

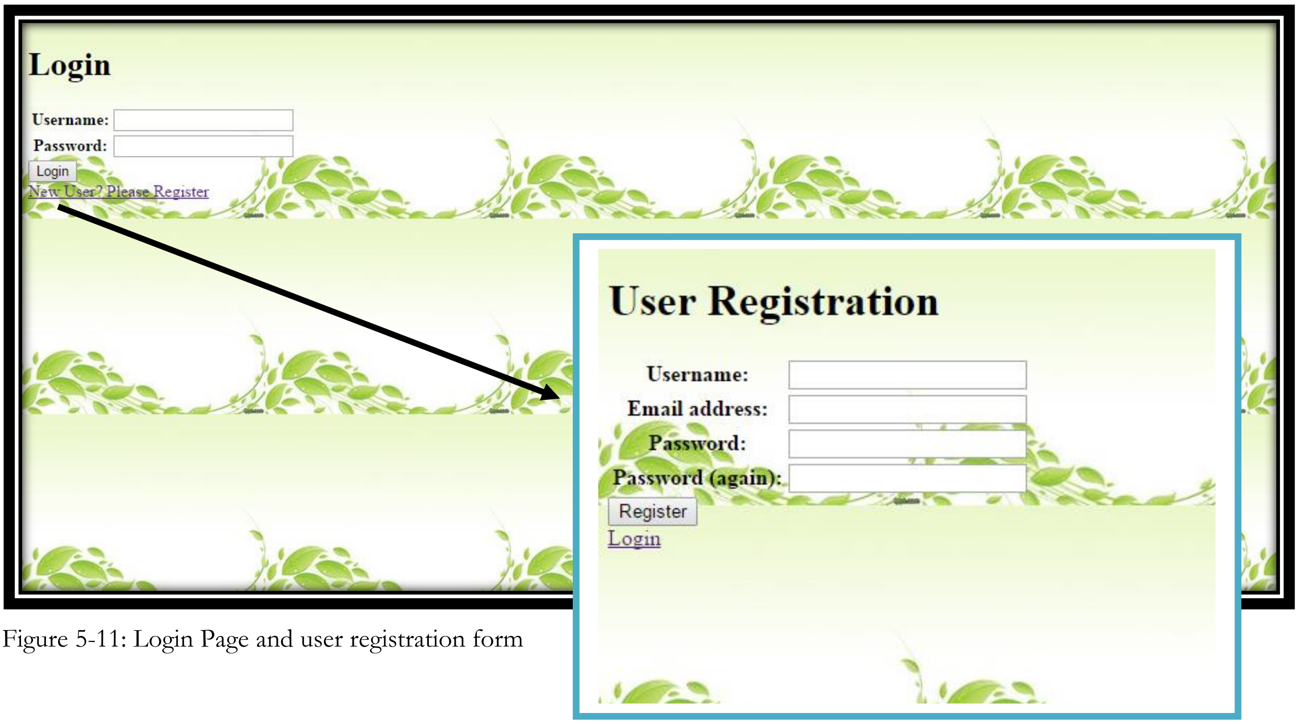


Figure 5-11: Login Page and user registration form

- ✓ User data upload

  The users can upload their data such as species and traits data in csv format with the specified fields. The web application validates the fields, stores the data in database, processes the data and provides desired result.

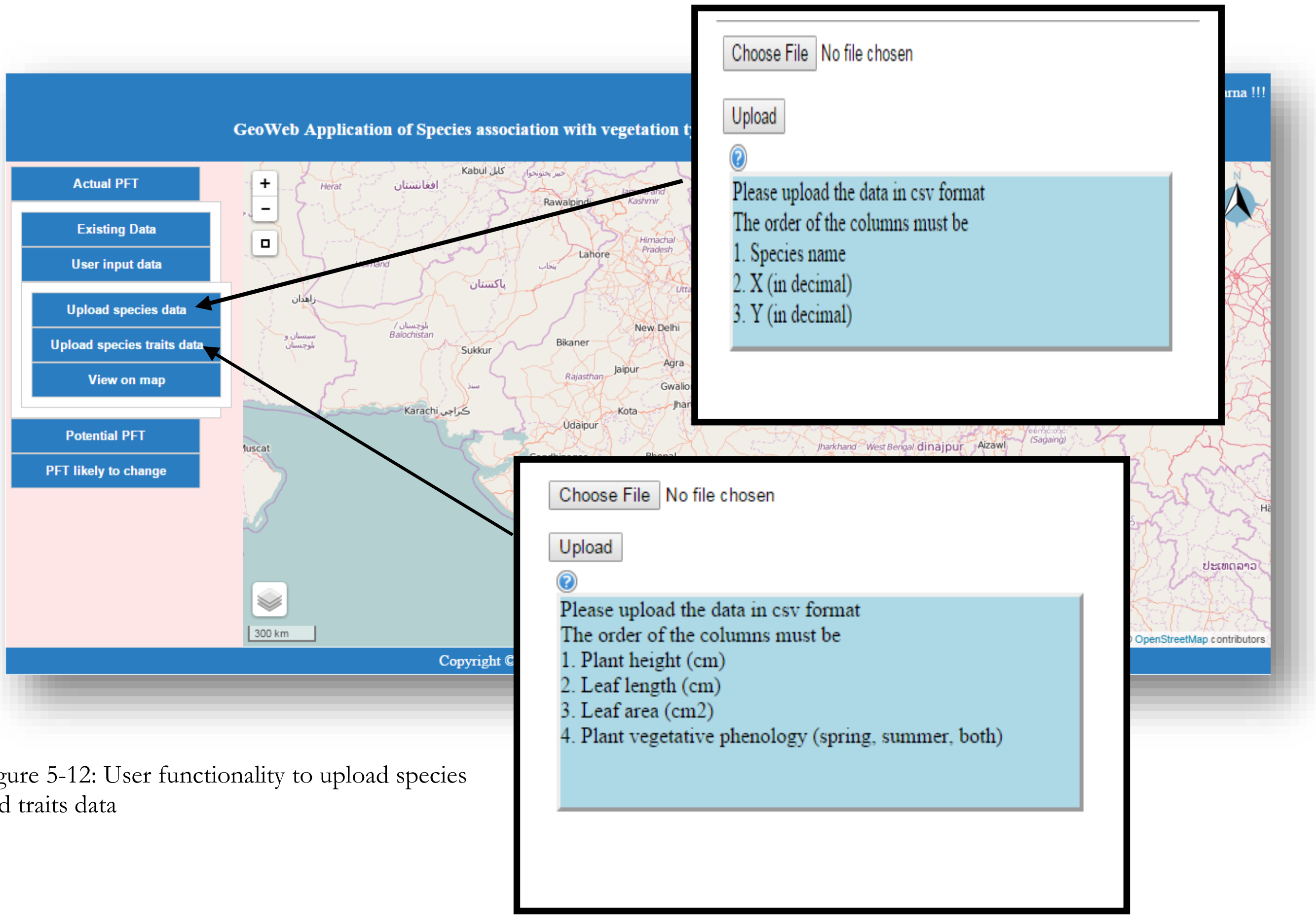


Figure 5-12: User functionality to upload species and traits data

- ✓ User interface

  The various UI functionalities like changing the base map, overlaying actual and potential vegetation type map and administrative boundaries were provided to users.

Figure 5-13: Base map with satellite view (Source: http://www.mapquest.com/)

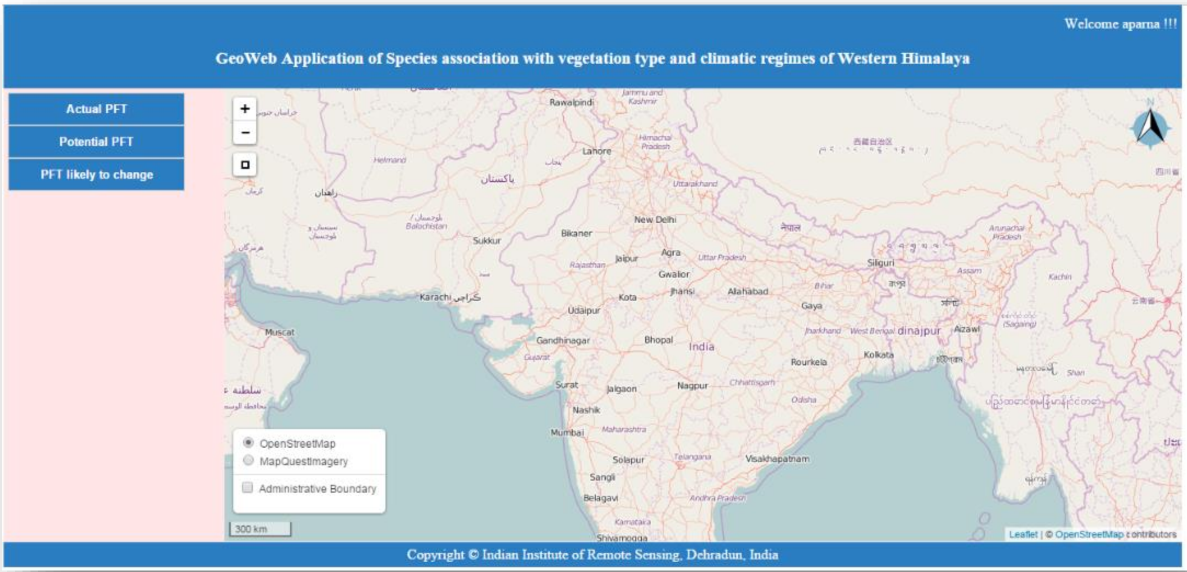


Figure 5-14: Base map. (Source: http://www.openstreetmap.org/)

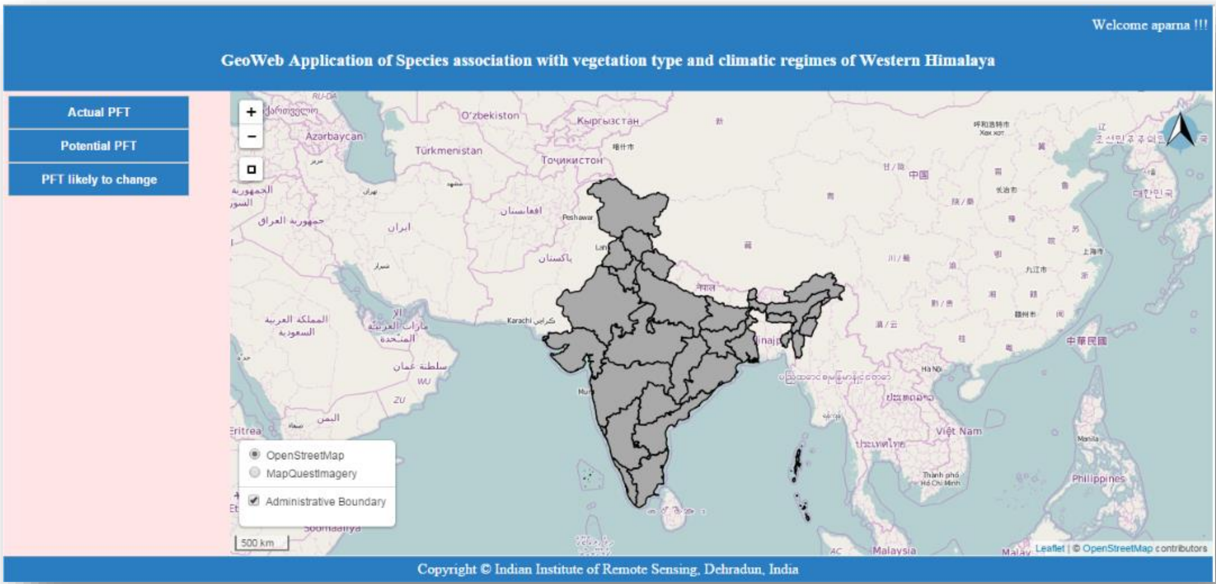


Figure 5-15: Base map with administrative boundaries of India. (Source: Survey of India)

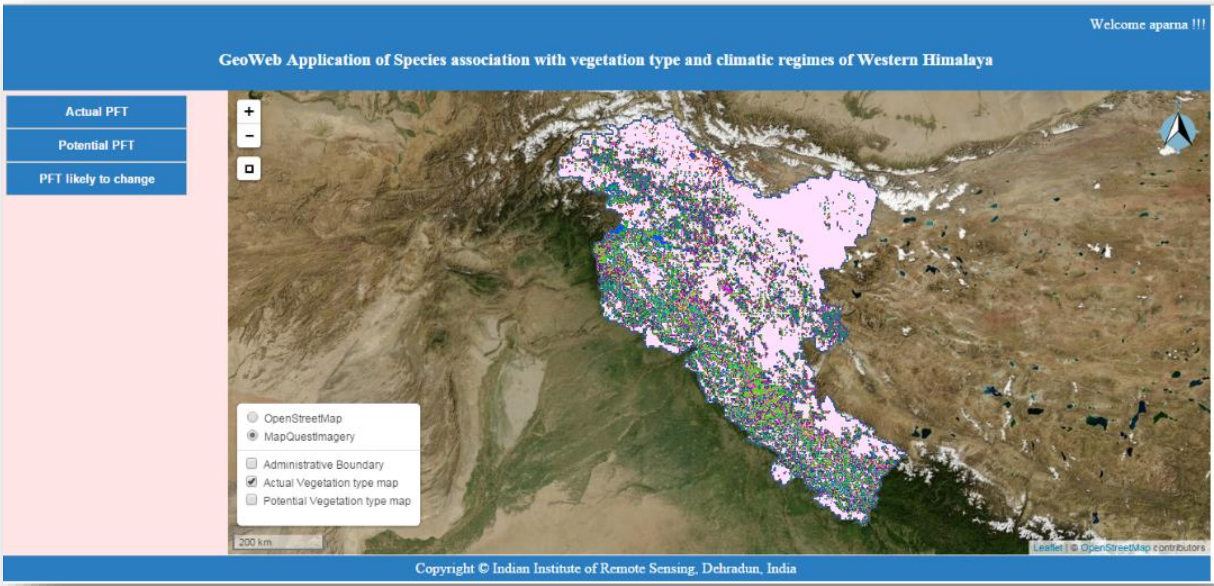


Figure 5-16: Base map overlaid with actual vegetation type map. (Source: DOS-DBT project)

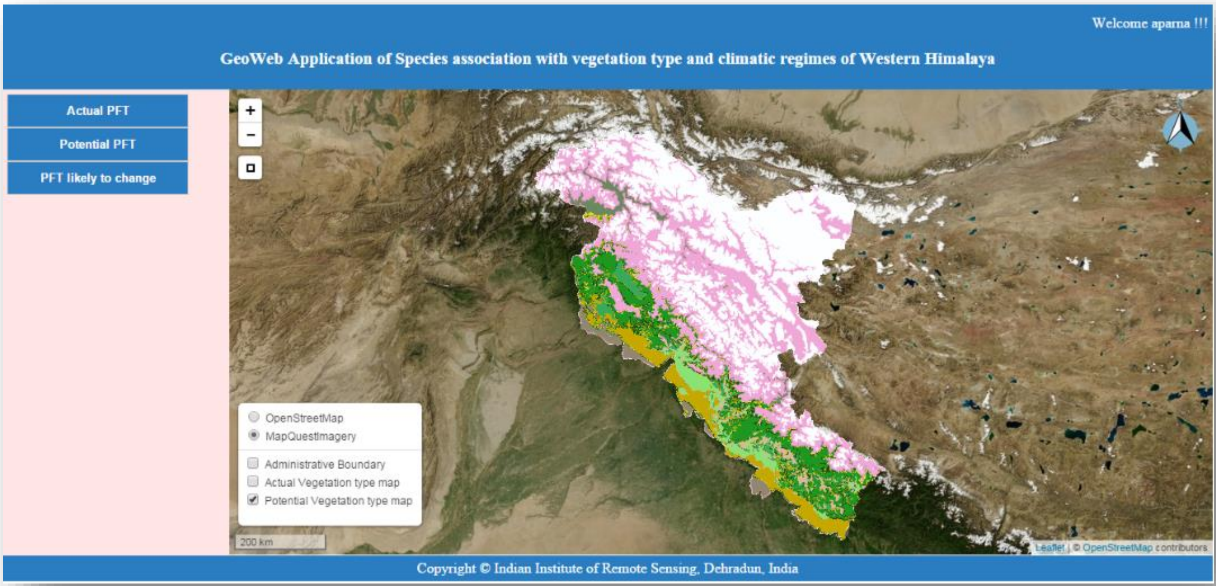


Figure 5-17: Base map overlaid with potential natural vegetation type map. (Source: IIRS)

5.1.5. Validation

The validation was performed on potential species since actual species are ground truth data. The dominant species for 75% of the plots of Himalayan moist temperate vegetation class were compared with the potential species of Himalayan moist temperate vegetation class. Then 25% of the plots consisting of dominant species were compared with potential species. The results are shown below.

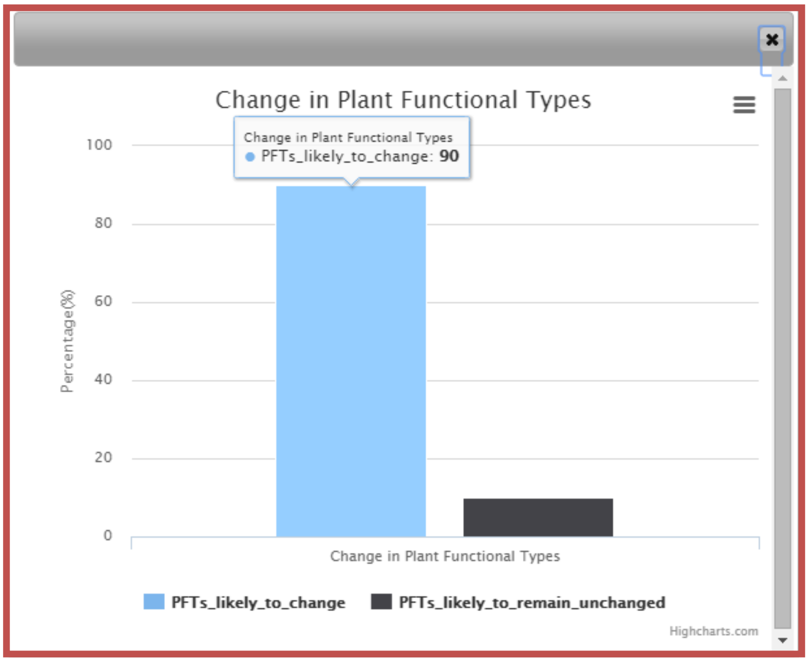


Figure 5-18: Comparison of 75% of dominant actual species with potential species

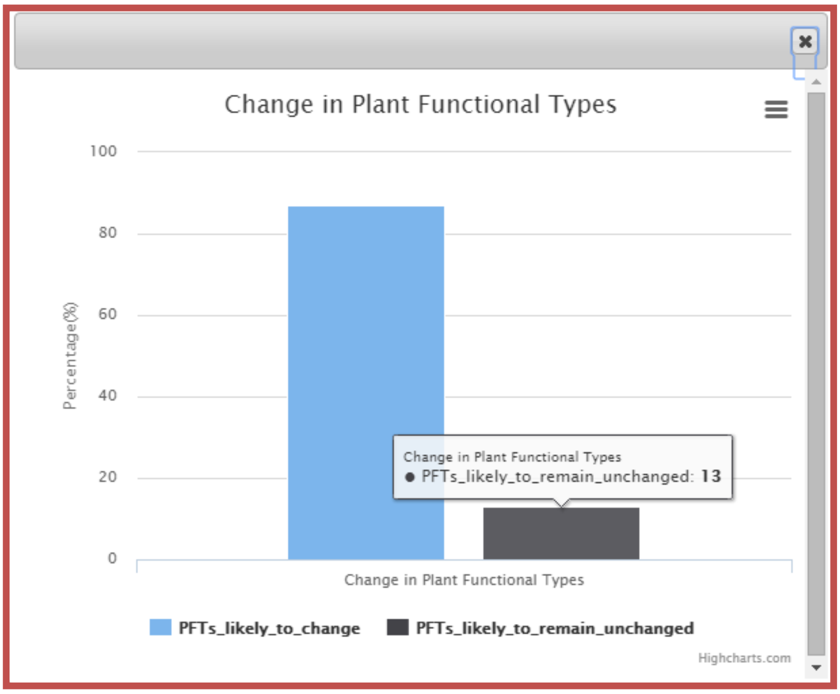


Figure 5-19: Comparison of 25% of dominant actual species with potential species

5.2. Discussion

This section discusses the results obtained for actual and potential plant functional types and the plant functional types likely to change. The objective of this research was to develop an algorithm for identification of plant functional types for 1 km x 1 km grid as a GeoWeb application.

For identification of actual and potential plant functional types, the algorithm performed K-means clustering for each trait subset followed by computing dissimilarity matrix for communities using 'Bray Curtis' and dissimilarity matrix for ecological variables using 'Euclidean' distance and then the correlation between communities and ecological factors were computed.

Clustering is one of the data analysis method which groups the items that are similar to each other. When species are grouped as a cluster, it represents the geographical region where it has established itself. This indicates that these regions have favourable climatic conditions and other environmental factors which help the species community to sustain themselves (Watts & Worner, 2009). In this research, K-means clustering was carried out to group the similar traits.

According to Watts & Worner (2009), K-means is more efficient when compared to self-organizing map (SOM) in computation. In this study, K-means clustering was applied on each subset of traits matrix (Figure 5-20 (a)-(o) and Figure 5-21 (a)-(o)) so that the correlation with ecological factors can be maximized. The K-means clustering helped to identify the discrete species groups and thus interpretation was easy. The K value in K-means was assigned to 3. Each trait subset resulted in three clusters along with the labels 0 1 and 2 that was mapped to species of trait matrix for grouping the species based on communities and environmental variables. The species clusters obtained for each plot were varying. For example, in plot 1-Him-54 and 1-Him-506 Cedrus deodara was not associated with any other species but in plot 1-Him-515 it was associated with Oxalis corniculata, Pinus roxburghii and Taraxacum officinale (Table 5-2). This indicates that Cedrus deodara had established itself in a particular environmental gradient. But due to disturbances, it was found to exist in other geographical region due to presence of favorable environmental and climatic conditions.

The communities were compared for species composition using 'Bray-Curtis' dissimilarity measure. This dissimilarity measure checks for similarity between two communities for species. The value of dissimilarity ranged from 0 to 1 in the obtained result (Figure 5-22 (a)-(o) and Figure 5-7 (a)-(o)). The value 0 indicated that the two communities are similar for all the species and value 1 indicated the two communities are dissimilar for all the species. For actual PFTs, out of 66 community pairs, 31% of community pairs did not share any species and 69% of the community pairs shared few species (Figure 5-23 (o)). For potential PFTs, out of 45 community pairs, 71% of community pairs did not share any species and 29% of community pairs shared few species (Figure 5-24 (f)).

The communities were then compared using 'Euclidean distance' dissimilarity matrix based on ecological factors. The data in ecological variables was standardized to mean of 0 and standard deviation of 1. For actual PFTs, 12 communities were compared for similarity in environmental factors like minimum temperature, maximum temperature, precipitation, altitude and soil type and for potential PFTs, 11 communities were compared. The distance values showed the contributions of each of the ecological factors between the two communities.

Further, correlation between environmental dissimilarity matrix and community dissimilarity was calculated using Pearson correlation coefficient. For actual PFTs, the correlation values ranged from 0.03 to 0.12 (Table 5-7). The results showed positive correlation between communities and environmental factors and maximum correlation was obtained for all traits i.e. plant height, leaf length, leaf area and plant phenology. This showed that the environmental conditions are suitable for these traits for the species to survive in the given spatial region. The reason for low correlation values might be due to few number of species which comprised of 31% of similar species between two communities and traits which influences the behaviour of species in a community. For potential PFT, the correlation values ranged from -0.05 to -0.13 (Table 5-8) which showed negative correlation between communities and environmental variables and maximum correlation i.e. -0.13 was obtained for plant height and leaf area.

The reason for finding both actual and potential PFTs was to compare how the potential PFTs have been changed over the course of time. For example, in plot 1-Him-54, the potential species like Cornus, Ilex dipyrena, Castanopsis, Corylus, Acer, Betula, Pinus wallichiana, Lyonia ovalifolia, Myrica, Ulmus, Abies pindrow, Arundinaria falcata, Rhododendron arboreum, Litsea umbrosa have been changed to Fragaria nubicola Lindley, Poa annua, Trifolium repens L. and Oxalis corniculata. The graph (Figure 5-25) showed more than 80% change in PFTs which indicated that there was drastic change in traits distribution over the course of time. This change may be due to increasing human population, change in environmental conditions, anthropogenic factors, deforestation, forest fire etc. According to Blundo, Malizia, & González-Espinosa (2015), the trait distribution is mainly influenced by climatic and topographic changes. Through change in PFTs i.e. change in trait and environment relationship, the effects of climate change as well as structure of the community can be studied. It also helps to predict the functional types which will be able to survive in spite of changing environment (Blundo et al., 2015).

The obtained potential PFTs was validated by dividing the total plots into 75% and 25% plots consisting of dominant species. When 75% of the plots with dominant species was compared with potential species, 90% change in PFTs were obtained and with 25% of the plots, 87% change in PFTs were obtained (Figure 5-18, 5-19). Thus, for almost all the plots more than 80% of potential species are dominated by other species. This helps to study the climate change as well as take necessary measures to protect the endangered species.

All the above discussed results were visualized on a GeoWeb application for four plots. The web application was developed using Django python web framework, postgis spatial database and leaflet javascript library.

Django python web framework provides more functionality when compared to older Common Gateway Interface (CGI) libraries. It consists of model, view and template pattern as discussed in section 4.2.3.1. The application logic can be separated from user interface through a template system. Database can be defined and accessed through python based API, creation of SQL statements can be automated. HTML forms can be generated and data can be retrieved from them through an API. It also includes logic for features like cookie access, user logins and session management (Burch, 2010). An important feature of Django is that it consists of a geographic extension called 'GeoDjango' for development of WebGIS application. It supports OGC geometries and web services as well as built in API's for Geospatial Data Abstraction Library (GDAL) and Geometry Engine- Open Source (GEOS) libraries.

Leaflet javascript library requires less coding when compared to OpenLayers due to wide variety of plugins for web mapping applications. The size of minified builds is 126 KB when compared to OpenLayers which 465 KB (Limoine, 2015). It provides responsive web application that can be extended to mobile applications as well.

The developed GeoWeb application provided functionality for the users to upload their data in csv format with specified set of fields. When the data is uploaded the application validates the fields and the data is stored in PostgreSQL database server. The algorithm takes this data as input, processes the data at server side and provides desired results to the client. The application also provided MapQuestImagery and OpenStreetMap as Basemap along with functionality to overlay actual and potential vegetation type map. To limit the server load, user session was taken care of by using django.contrib.sessions app in Django. The GeoWeb application is planned to be deployed publically on a web server. This kind of WebGIS application is the first attempt to study the community structure by analysing the plant functional types for 1 km x 1km grid for western Himalaya of Indian region.

6. CONCLUSION AND RECOMMENDATION

6.1. Conclusion

Earlier, many remote sensing techniques were adopted for mapping plant functional types. The PFTs were extracted from land cover data which were of coarser resolution. For community land model, a global PFT map has been produced by MODIS team. However, the MODIS PFT dataset demands validation and the accuracy of the dataset appears to be less (Sun & Liang, 2008). Also, the important community specific environmental conditions which determine PFTs are less observed by remote sensing techniques. Therefore a new approach was required to identify PFTs which can take into account all the factors like plant traits, community structure and environmental conditions. All these factors can be analysed numerically (Valério DePatta Pillar & Sosinski, 2003). Therefore, in this research the developed algorithm followed numerical approach for identifying actual and potential PFTs.

The numerical approach in the algorithm with the trait, community and environmental variables matrix helped to identify the plant functional types that were maximally associated with environmental variables. Also, in this research the combination of different traits were considered to find the traits combination that was highly correlated with ecological variables which in turn helped to identify the actual and potential plant functional types. The developed algorithm identified actual and potential plant functional types for 1 km x 1 km grid which is of a finer resolution compared to PFTs identified through remote sensing techniques for coarser resolution. Also, the algorithm was integrated with the GeoWeb application which helped to visualize the results as well as analyse the user input data.

The end result of this research showed the change in PFTs from potential to actual which indicated the species which are highly important due to economic importance, ethical, cultural and scientific reasons are diminishing. Thus the developed WebGIS application can contribute for taking necessary measures for conservation of biodiversity which in turn may contribute to the provision of ecological services and functioning of healthy natural ecosystem (Brandful Cobbinah, Black, & Thwaites, 2015).

6.1.1. Answers to the research questions

1. *How the PFTs are to be identified based on environmental factors and vegetation type?*

The PFTs were identified using numerical approach. In this approach, three matrices were used i.e. trait matrix which consisted of species and their corresponding traits, community matrix with communities as vegetation type and species composition in each vegetation type, environmental variables matrix with communities and environmental factors such as minimum temperature, maximum temperature, altitude, precipitation and soil type. The species were grouped based on traits using K-means clustering. To find how these species are associated with environmental variables, correlation between community and environmental variables dissimilarity matrix was calculated. The corresponding cluster of species were identified as plant functional types associated with environmental factors and vegetation type.

2. ***How is uncertainty of PFT associated with the definition of PFT and what ranges of environmental factors may correspond to same PFT?***

The developed algorithm followed the numerical approach where in, clustering was performed on each trait subset i.e. combinations of traits like plant height and leaf length, plant height and leaf area, leaf area and plant phenology etc. Thus each trait subset resulted in set of clusters. The communities were grouped according to these clusters and correlation between each of the community groups and ecological variables was calculated to find which trait subset is maximally associated with environmental variables. Thus the trait subset which gave maximum correlation value are the plant functional types that correspond to given environmental variables. Thus the uncertainty of PFT associated with definition of PFT was taken care.

3. What functionalities should the application offer to the users?

The Developed GeoWeb application offered various functionalities to the users such as

- ✓ User login and registration.
- ✓ User data upload for species and traits data
- ✓ Geo-visualization functions such as overlaying administrative boundaries, actual and potential vegetation type map, satellite and street view for basemap.

6.2. Recommendation

In this research, the end product was the developed algorithm for identification of plant functional types integrated with the GeoWeb application for multi user environment. There are certain observations in improving the performance of the algorithm as well as the web application which could be taken into account for the future work. They are,

- ✓ The correlation value between the communities and environmental variables could be improved by considering more number of species and their traits.

- ✓ The 'Bray-Curtis' dissimilarity measure was applied for communities which consisted of dominant species. Similarly, other dissimilarity measures could be tried if dominant as well as non-dominant species are considered.

- ✓ The GeoWeb application developed was confined to desktop application. It could be extended for mobile application.

7. LIST OF REFERENCES

References

Anderson, P. M. L., & Hoffman, M. T. (2011). Grazing response in the vegetation communities of the Kamiesberg, South Africa: Adopting a plant functional type approach. *Journal of Arid Environments*, *75*(3), 255–264. http://doi.org/10.1016/j.jaridenv.2010.10.012

Bandhu, K. (2011). *Principles and Practice of Silviculture*. Khanna Bandhu.

Blundo, C., Malizia, L. R., & González-Espinosa, M. (2015). Distribution of functional traits in subtropical trees across environmental and forest use gradients. *Acta Oecologica*, *69*, 96–104. http://doi.org/10.1016/j.actao.2015.09.008

Bochet, E., & García-Fayos, P. (2015). Identifying plant traits: A key aspect for species selection in restoration of eroded roadsides in semiarid environments. *Ecological Engineering*, *83*, 444–451. http://doi.org/10.1016/j.ecoleng.2015.06.019

Bonan, G. B., Levis, S., Kergoat, L., & Oleson, K. W. (2002). Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Global Biogeochemical Cycles*, *16*(2), 5–1–5–23. http://doi.org/10.1029/2000GB001360

Boundless. (n.d.). Retrieved from http://boundlessgeo.com/products/opengeo-suite/postgis/

Box, E. O. (1995). Factors determining distributions of tree species and plant functional types. *Vegetatio*, *121*(1-2), 101–116. http://doi.org/10.1007/BF00044676

Brandful Cobbinah, P., Black, R., & Thwaites, R. (2015). Biodiversity conservation and Livelihoods in rural Ghana: Impacts and coping Strategies. *Environmental Development*, *15*, 79–93. http://doi.org/10.1016/j.envdev.2015.04.006

Burch, C. (2010). Django, a web framework using Python: tutorial presentation. *Journal of Computing Sciences in Colleges*, *25*(5), 154–155. Retrieved from http://dl.acm.org/citation.cfm?id=1747137.1747166

Condit, R., Hubbell, S. P., & Foster, R. B. (1996). Assessing the response of plant functional types to climatic change in tropical forests. *Journal of Vegetation Science*, *7*(3), 405–416. http://doi.org/10.2307/3236284

De'ath, G. (1999). Extended dissimilarity: a method of robust estimation of ecological distances from high beta diversity data. *Plant Ecology*, *144*(2), 191–199. http://doi.org/10.1023/A:1009763730207

Díaz, S., & Cabido, M. (1997). Plant Functional Types and Ecosystem Function in Relation to Global Change. *Journal of Vegetation Science*, *8*(4), 463–474. http://doi.org/10.2307/3237198

Faith, D. P., Minchin, P. R., & Belbin, L. (1987). Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*, *69*(1-3), 57–68. http://doi.org/10.1007/BF00038687

Geodjango Tutorial. (n.d.). Retrieved from http://www.igg.go.ug/static/files/publications/gasasira-verification-report___.pdf

GeoServer. (n.d.). Retrieved from http://docs.geoserver.org/2.6.5/user/introduction/overview.html

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On Clustering Validation Techniques. *Journal of*

Intelligent Information Systems, *17*(2/3), 107–145. http://doi.org/10.1023/A:1012801612483

Hodgson, J. G., Wilson, P. J., Hunt, R., Grime, J. P., & Thompson, K. (1999). Allocating C-S-R Plant Functional Types: A Soft Approach to a Hard Problem. *Oikos*, *85*(2), 282. http://doi.org/10.2307/3546494

Karnatak, H. C., Saran, S., Bhatia, K., & Roy, P. S. (2007). Multicriteria Spatial Decision Analysis in Web GIS Environment. *GeoInformatica*, *11*(4), 407–429. http://doi.org/10.1007/s10707-006-0014-8

Kooyman, R., Cornwell, W., & Westoby, M. (2010). Plant functional traits in Australian subtropical rain forest: partitioning within-community from cross-landscape variation. *Journal of Ecology*, *98*(3), 517–525. http://doi.org/10.1111/j.1365-2745.2010.01642.x

Körner, K., & Jeltsch, F. (2008). Detecting general plant functional type responses in fragmented landscapes using spatially-explicit simulations. *Ecological Modelling*, *210*(3), 287–300. http://doi.org/10.1016/j.ecolmodel.2007.08.002

Laporte, M.-A., & Garnier, E. (2012). ThesauForm—Traits: A web based collaborative tool to develop a thesaurus for plant functional diversity research. *Ecological Informatics*, *11*, 34–44. http://doi.org/10.1016/j.ecoinf.2012.04.004

Legendre, P., & Birks, H. J. B. (2012). Clustering and Partitioning (pp. 167–200). http://doi.org/10.1007/978-94-007-2745-8_7

Limoine, E. (2015). A response to Iván's "Leaflet vs OpenLayers" talk. Retrieved from http://erilem.net/a-response-to-ivans-leaflet-vs-openlayers-talk.html

McGarigal, K., Stafford, S., & Cushman, S. (2000). Cluster Analysis. In *Multivariate Statistics for Wildlife and Ecology Research* (pp. 81–128). New York, NY: Springer New York. http://doi.org/10.1007/978-1-4612-1288-1_3

Müller, S. C., Overbeck, G. E., Pfadenhauer, J., & Pillar, V. D. (2007). Plant Functional Types of Woody Species Related to Fire Disturbance in Forest–Grassland Ecotones. *Plant Ecology*, *189*(1), 1–14. http://doi.org/10.1007/s11258-006-9162-z

Nathan Lemoine. (2013). Climate Change Ecology. Retrieved from https://climateecology.wordpress.com/2013/04/23/python-compliments-rs-shortcomings/

NumPy. (n.d.). Retrieved December 6, 2015, from http://www.numpy.org/

Pavoine, S., Dufour, A.-B., & Chessel, D. (2004). From dissimilarities among species to dissimilarities among communities: a double principal coordinate analysis. *Journal of Theoretical Biology*, *228*(4), 523–537. http://doi.org/10.1016/j.jtbi.2004.02.014

Percivall, G. (2008). OGC's Open Standards for Geospatial Interoperability. In *Encyclopedia of GIS* (pp. 800–805). Boston, MA: Springer US. http://doi.org/10.1007/978-0-387-35973-1_904

Pillar, V. D., Duarte, L. da S., Sosinski, E. E., & Joner, F. (2009). Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *Journal of Vegetation Science*, *20*(2), 334–348. http://doi.org/10.1111/j.1654-1103.2009.05666.x

Pillar, V. D. P. D. (1999). On the identification of optimal plant functional types. *Journal of Vegetation Science*, *10*(5), 631–640. http://doi.org/10.2307/3237078

Pillar, V. D., & Sosinski, E. E. (2003). An improved method for searching plant functional types by numerical analysis. *Journal of Vegetation Science*, *14*(3), 323. http://doi.org/10.1658/1100-9233(2003)014[0323:AIMFSP]2.0.CO;2

Pla, L., Casanoves, F., & Di Rienzo, J. (2012). *Quantifying Functional Biodiversity*. Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-007-2648-2

PostgreSQL. (n.d.). Retrieved from http://www.postgresql.org/about/

Python. (n.d.). Retrieved December 6, 2015, from https://www.python.org/about/

Roy, P. S., Kushwaha, S. P. S., Murthy, M. S. R., Roy, A., Kushwaha, D., Reddy, C. S., … Porwal, M. C. (2012). *Biodiversity Characterization At Landscape Level*. Indian Institute of Remote Sensing. Retrieved from http://bis.iirs.gov.in/

Salmon, S., Ponge, J. F., Gachet, S., Deharveng, L., Lefebvre, N., & Delabrosse, F. (2014). Linking species, traits and habitat characteristics of Collembola at European scale. *Soil Biology and Biochemistry*, *75*, 73–85. http://doi.org/10.1016/j.soilbio.2014.04.002

Sanchez, I. (n.d.). Leaflet. Retrieved from https://github.com/Leaflet/Leaflet

Schwoertzig, E., Poulin, N., Hardion, L., & Trémolières, M. (2016). Plant ecological traits highlight the effects of landscape on riparian plant communities along an urban–rural gradient. *Ecological Indicators*, *61*, 568–576. http://doi.org/10.1016/j.ecolind.2015.10.008

scikit-learn. (n.d.). Retrieved from http://scikit-learn.org/stable/

SciPy. (n.d.). Retrieved from http://www.scipy.org/about.html

Souza, A. F., Forgiarini, C., Longhi, S. J., & Oliveira, J. M. (2014). Detecting ecological groups from traits: a classification of subtropical tree species based on ecological strategies. *Brazilian Journal of Botany*, *37*(4), 441–452. http://doi.org/10.1007/s40415-014-0084-z

Sun, W., & Liang, S. (2008). Methodologies for Mapping Plant Functional Types. In *Advances in Land Remote Sensing* (pp. 369–393). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-1-4020-6450-0_14

Template Extending. (n.d.). Retrieved from http://tutorial.djangogirls.org/en/template_extending/index.html

Thilina Hewaragamunige. (2014). Implementation and Use of The K-Means Algorithm. Retrieved from http://www.cs.colostate.edu/~anderson/cs545/index.html/lib/exe/fetch.php?media=assignments:solutions1:one.pdf

Watts, M. J., & Worner, S. P. (2009). Estimating the risk of insect species invasion: Kohonen self-organising maps versus k-means clustering. *Ecological Modelling*, *220*(6), 821–829. http://doi.org/10.1016/j.ecolmodel.2008.12.016

Westoby, M., & Wright, I. J. (2006). Land-plant ecology on the basis of functional traits. *Trends in Ecology & Evolution*, *21*(5), 261–268. http://doi.org/10.1016/j.tree.2006.02.004

why Python? (n.d.). Retrieved from http://www.programmingforbiologists.org/about/why-python/

WorldClim - Global Climate Data. (n.d.). Retrieved August 24, 2015, from http://www.worldclim.org/current

Writing Views. (n.d.). Retrieved from http://www.effectivedjango.com/tutorial/views.html

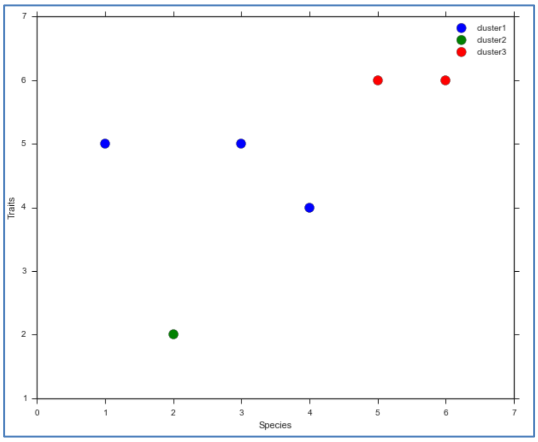
Wu, J. (2012). Cluster Analysis and K-means Clustering: An Introduction (pp. 1–16). http://doi.org/10.1007/978-3-642-29807-3_1

Young, A. B., Watts, D. A., Taylor, A. H., & Post, E. (2016). Species and site differences influence climate-shrub growth responses in West Greenland. *Dendrochronologia*, *37*, 69–78. http://doi.org/10.1016/j.dendro.2015.12.007

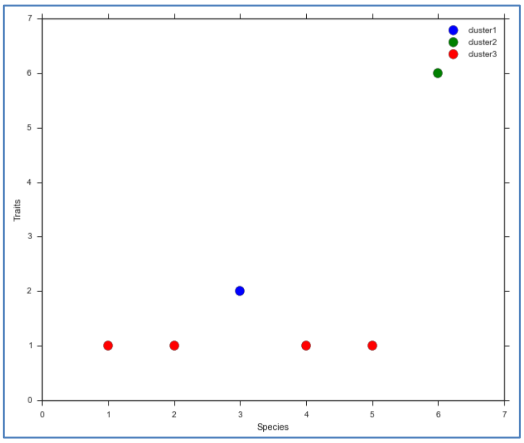
Zou, K. H., Tuncali, K., & Silverman, S. G. (2003). Correlation and Simple Linear Regression1. *Radiology*, *227*(3), 617–628. http://doi.org/10.1148/radiol.2273011499

8. APPENDIX

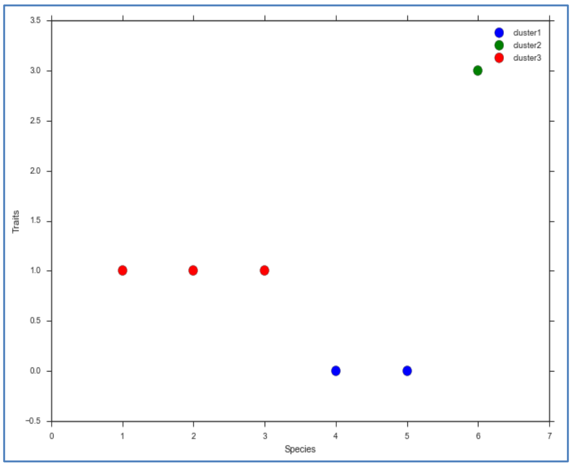
Kmeans clustering for all the trait subsets –Actual PFT



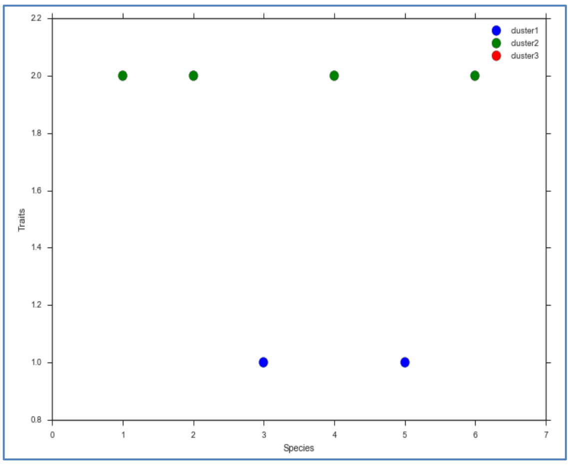
(a)



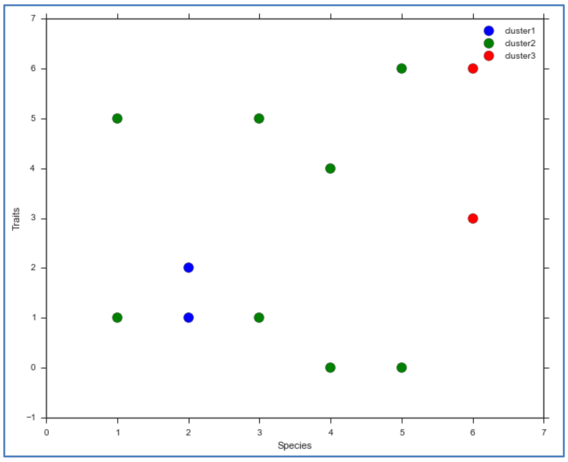
(b)



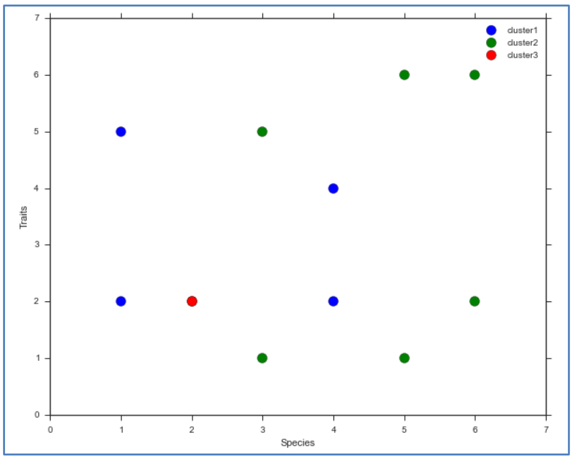
(c)



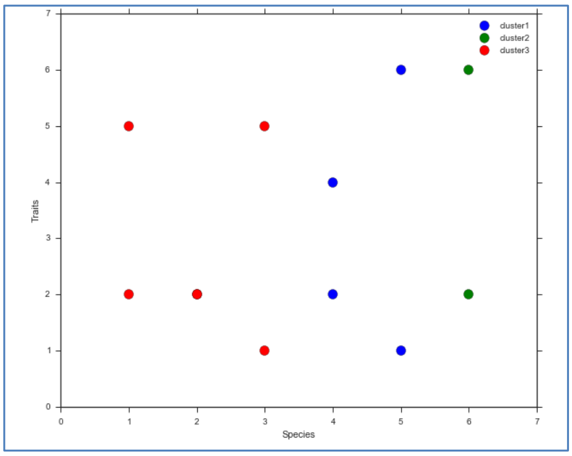
(d)


cluster1
cluster2
cluster3
Traits
Species

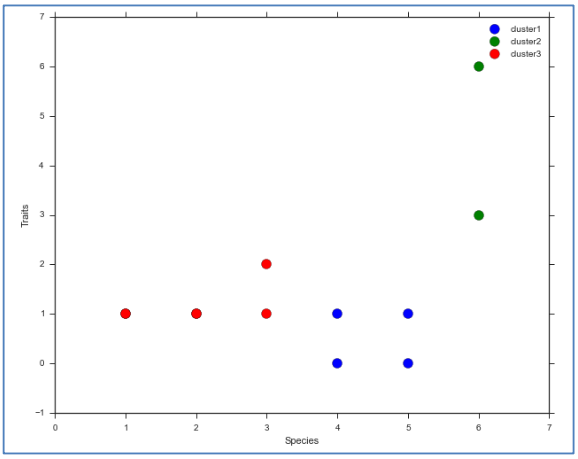
(e)


cluster1
cluster2
cluster3
Traits
Species

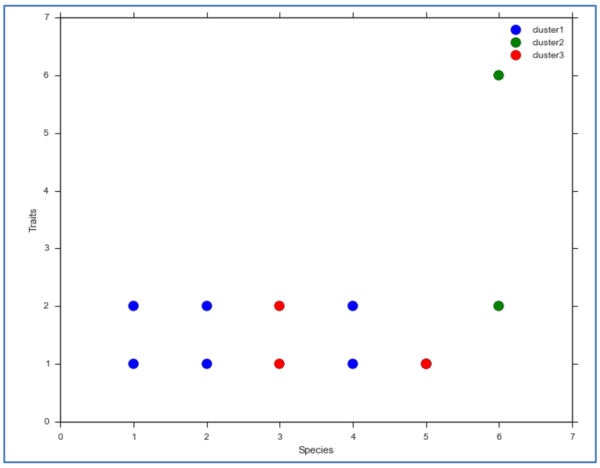
(f)


cluster1
cluster2
cluster3
Traits
Species

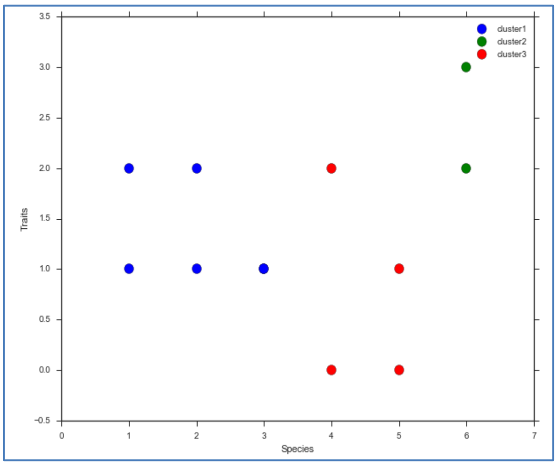
(g)


cluster1
cluster2
cluster3
Traits
Species

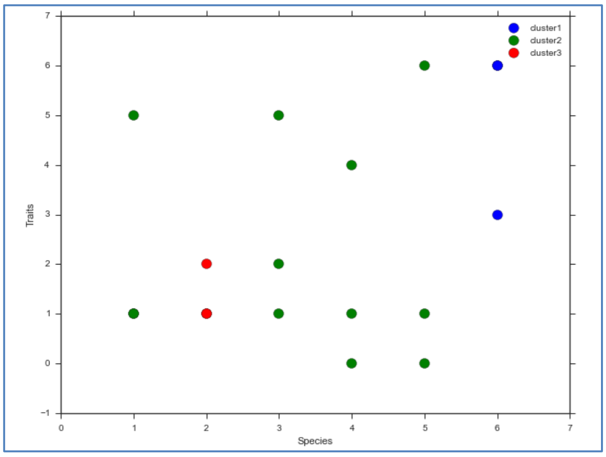
(h)


cluster1
cluster2
cluster3
Traits
Species

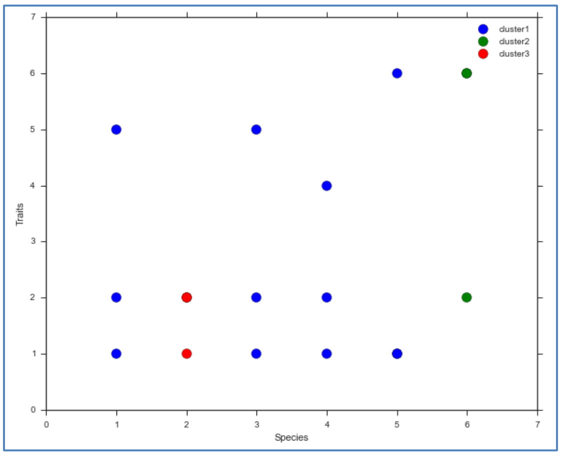
(i)


cluster1
cluster2
cluster3
Traits
Species

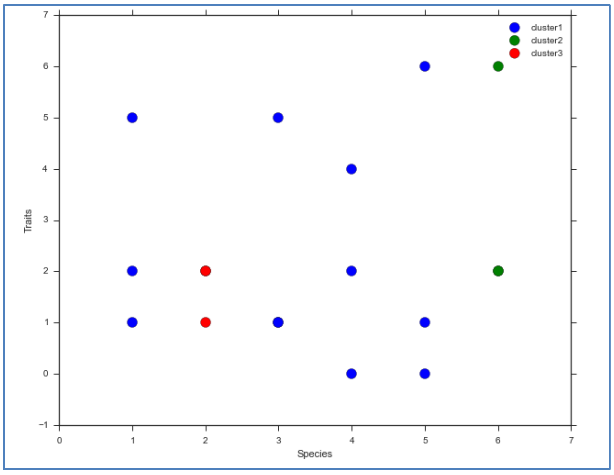
(j)


cluster1
cluster2
cluster3
Traits
Species

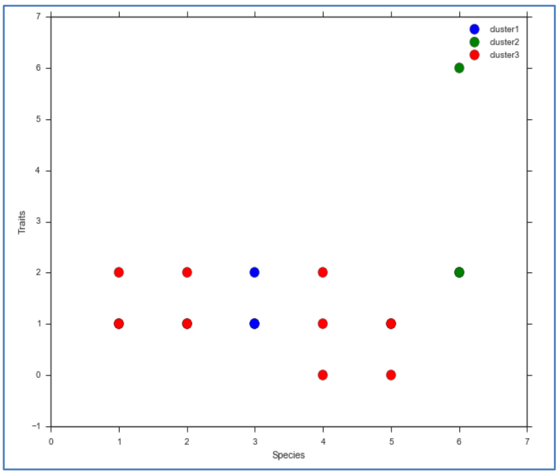
(k)


cluster1
cluster2
cluster3
Traits
Species

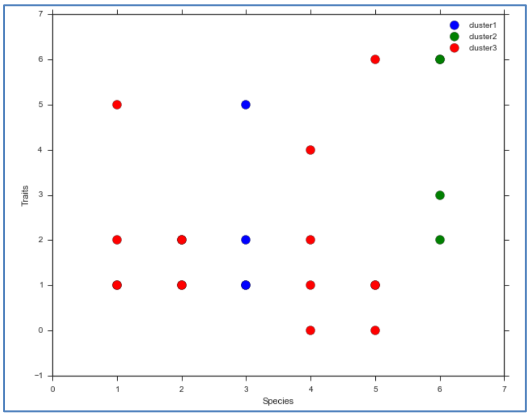
(l)


cluster1
cluster2
cluster3
Traits
Species

(m)


cluster1
cluster2
cluster3
Traits
Species

(n)



(o)

Figure 8-1: Scatterplot of species and each trait subset (a)- Plant height (b)- Leaf length (c)-Leaf area (d)-Plant phenology (e)-Plant height, Leaf length (f)-Plant height, Leaf area (g)-Plant height, Plant phenology (h)-Leaf length, Leaf area (i)- Leaf length, Plant phenology (j)-Leaf area, Plant phenology (k)- Plant height, Leaf length, Leaf area (l)-Plant height, Leaf length, Plant phenology (m)- Plant height, Leaf area, Plant phenology (n)-Leaf length, Leaf area, Plant phenology (o)- Plant height, Leaf length, Leaf area, Plant phenology

Community Dissimilarity matrix - Actual PFT

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8543158									
4	1.0000000	0.9582036	0.5727273								
5	1.0000000	0.4339941	0.9319434	0.9810407							
6	1.0000000	0.3775974	0.9162430	0.9765293	0.2517758						
7	1.0000000	0.4883407	0.6279570	0.8830846	0.7328864	0.6788971					
8	1.0000000	0.3959676	0.9342455	0.9816978	0.3116000	0.3219852	0.7410195				
9	1.0000000	0.4250379	0.8210036	0.9479801	0.4723430	0.3844001	0.3984903	0.4860564			
10	1.0000000	0.5102881	0.6099211	0.8764783	0.7461333	0.6943493	0.4099252	0.7539204	0.4365400		
11	1.0000000	0.7427780	0.3052209	0.7473118	0.8758594	0.8482372	0.3992606	0.8799409	0.6882494	0.6400385	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(a)

	1	2	3	4	5	6	7	8	9	10	11
2	1.00000000										
3	1.00000000	0.85431579									
4	1.00000000	0.95820365	0.76363636								
5	1.00000000	0.55574301	0.93194335	0.98104074							
6	1.00000000	0.47662338	0.91624304	0.97652934	0.22020521						
7	1.00000000	0.48834066	0.62795699	0.91044776	0.73288638	0.67889714					
8	1.00000000	0.42531889	0.93424553	0.98169782	0.49060000	0.32640654	0.74101950				
9	1.00000000	0.48683322	0.62915327	0.88847584	0.73196967	0.67782959	0.05998682	0.74012652			
10	1.00000000	0.51028807	0.60992108	0.92641261	0.74613333	0.69434932	0.09177430	0.75392039	0.15332429		
11	1.00000000	0.74277800	0.30522088	0.74731183	0.87585943	0.84823722	0.47689464	0.87994089	0.46175115	0.47064485	
12	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000

(b)

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8792321									
4	1.0000000	0.9606701	0.5859031								
5	1.0000000	0.3378428	0.9383464	0.9803386							
6	1.0000000	0.3074350	0.9341579	0.9789720	0.3097614						
7	1.0000000	0.4209554	0.7275591	0.9061685	0.6643207	0.6449264					
8	1.0000000	0.4150369	0.9482423	0.9835516	0.3664129	0.1233397	0.7116572				
9	1.0000000	0.5332765	0.8404795	0.9473171	0.4630800	0.5736084	0.3579049	0.6088387			
10	1.0000000	0.5431356	0.6432990	0.8730905	0.7443874	0.7288655	0.1583949	0.7819127	0.4292875		
11	1.0000000	0.7081218	0.4533965	0.7898833	0.8440414	0.8339949	0.4091201	0.8680436	0.6254072	0.3508353	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(c)

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8792321									
4	1.0000000	0.9606701	0.5859031								
5	1.0000000	0.3378428	0.9383464	0.9803386							
6	1.0000000	0.3074350	0.9341579	0.9789720	0.2361943						
7	1.0000000	0.4209554	0.7275591	0.9061685	0.6643207	0.6449264					
8	1.0000000	0.4150369	0.9482423	0.9835516	0.3992135	0.1971710	0.7116572				
9	1.0000000	0.1484642	0.8404795	0.9473171	0.4630800	0.4359989	0.3579049	0.5307945			
10	1.0000000	0.5431356	0.6432990	0.8730905	0.7443874	0.7288655	0.1985216	0.7819127	0.4292875		
11	1.0000000	0.7081218	0.4533965	0.7898833	0.8440414	0.8339949	0.4091201	0.8680436	0.6254072	0.3508353	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(d)

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8543158									
4	1.0000000	0.9582036	0.5727273								
5	1.0000000	0.3808520	0.9319434	0.9810407							
6	1.0000000	0.3418831	0.9162430	0.9765293	0.2930432						
7	1.0000000	0.4883407	0.6279570	0.8830846	0.7328864	0.6788971					
8	1.0000000	0.4162666	0.9342455	0.9816978	0.3876000	0.3606720	0.7410195				
9	1.0000000	0.1499243	0.8261769	0.9479801	0.4723430	0.4721931	0.4096146	0.4860564			
10	1.0000000	0.5102881	0.6099211	0.8764783	0.7461333	0.6943493	0.1461591	0.7646045	0.5036378		
11	1.0000000	0.7427780	0.3253012	0.7473118	0.8758594	0.8482372	0.3992606	0.8799409	0.6882494	0.3743985	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(e)

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8543158									
4	1.0000000	0.9582036	0.5727273								
5	1.0000000	0.3808520	0.9319434	0.9810407							
6	1.0000000	0.2850649	0.9162430	0.9765293	0.2301274						
7	1.0000000	0.4883407	0.6279570	0.8830846	0.7328864	0.6788971					
8	1.0000000	0.3959676	0.9342455	0.9816978	0.2614000	0.1250138	0.7410195				
9	1.0000000	0.4386673	0.8210036	0.9479801	0.4756408	0.3844001	0.3984903	0.4860564			
10	1.0000000	0.5102881	0.6099211	0.8764783	0.7461333	0.6943493	0.4099252	0.7539204	0.5173808		
11	1.0000000	0.7427780	0.3052209	0.7473118	0.8758594	0.8482372	0.3992606	0.8799409	0.6882494	0.6400385	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(f)

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8792321									
4	1.0000000	0.9606701	0.5859031								
5	1.0000000	0.3378428	0.9383464	0.9803386							
6	1.0000000	0.3074350	0.9341579	0.9789720	0.3097614						
7	1.0000000	0.4209554	0.7275591	0.9061685	0.6643207	0.6449264					
8	1.0000000	0.4150369	0.9482423	0.9835516	0.3664129	0.1233397	0.7116572				
9	1.0000000	0.5332765	0.8404795	0.9473171	0.4630800	0.5736084	0.3579049	0.6088387			
10	1.0000000	0.5431356	0.6432990	0.8730905	0.7443874	0.7288655	0.1583949	0.7819127	0.4292875		
11	1.0000000	0.7081218	0.4533965	0.7898833	0.8440414	0.8339949	0.4091201	0.8680436	0.6254072	0.3508353	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(g)

	1	2	3	4	5	6	7	8	9	10	11
2	1.00000000										
3	1.00000000	0.91602745									
4	1.00000000	0.95840708	0.69774920								
5	1.00000000	0.52622565	0.92858109	0.98690078							
6	1.00000000	0.29701398	0.87853692	0.97723971	0.47502676						
7	1.00000000	0.54163934	0.56039964	0.89366516	0.78985689	0.65975610					
8	1.00000000	0.40245714	0.90326127	0.98206449	0.52884849	0.12838202	0.72243409				
9	1.00000000	0.63119984	0.81855670	0.96509469	0.45861893	0.66592361	0.51937984	0.72346939			
10	1.00000000	0.50901054	0.77822581	0.87870968	0.81468754	0.69736022	0.09904153	0.75413712	0.56846473		
11	1.00000000	0.81273692	0.23367199	0.80084746	0.88747683	0.81144632	0.38985737	0.84872753	0.72321719	0.56114484	
12	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000

(h)

	1	2	3	4	5	6	7	8	9	10	11
2	1.00000000										
3	1.00000000	0.91602745									
4	1.00000000	0.95840708	0.69774920								
5	1.00000000	0.52622565	0.92858109	0.98690078							
6	1.00000000	0.29701398	0.87853692	0.97723971	0.47502676						
7	1.00000000	0.54163934	0.56039964	0.89366516	0.78985689	0.65975610					
8	1.00000000	0.40245714	0.90326127	0.98206449	0.52884849	0.12838202	0.72243409				
9	1.00000000	0.63119984	0.81855670	0.96509469	0.45861893	0.66592361	0.51937984	0.72346939			
10	1.00000000	0.50901054	0.77822581	0.87870968	0.81468754	0.69736022	0.09904153	0.75413712	0.56846473		
11	1.00000000	0.81273692	0.23367199	0.80084746	0.88747683	0.81144632	0.38985737	0.84872753	0.72321719	0.56114484	
12	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000

(i)

	1	2	3	4	5	6	7	8	9	10	11
2	1.00000000										
3	1.00000000	0.85431579									
4	1.00000000	0.95820365	0.76363636								
5	1.00000000	0.55574301	0.93194335	0.98104074							
6	1.00000000	0.47662338	0.91624304	0.97652934	0.17893787						
7	1.00000000	0.48834066	0.62795699	0.91044776	0.73288638	0.67889714					
8	1.00000000	0.42531889	0.93424553	0.98169782	0.17720000	0.12501382	0.74101950				
9	1.00000000	0.30540131	0.82100362	0.95019369	0.47564083	0.38440014	0.39849027	0.48605636			
10	1.00000000	0.51028807	0.60992108	0.92641261	0.74613333	0.69434932	0.06458192	0.75392039	0.42279709		
11	1.00000000	0.74277800	0.30522088	0.74731183	0.87585943	0.84823722	0.47689464	0.87994089	0.71990408	0.47064485	
12	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000

(j)

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8792321									
4	1.0000000	0.9606701	0.5859031								
5	1.0000000	0.3378428	0.9383464	0.9803386							
6	1.0000000	0.3074350	0.9341579	0.9789720	0.3097614						
7	1.0000000	0.4209554	0.7275591	0.9061685	0.6643207	0.6449264					
8	1.0000000	0.4150369	0.9482423	0.9835516	0.3664129	0.1233397	0.7116572				
9	1.0000000	0.5332765	0.8404795	0.9473171	0.4630800	0.5736084	0.3579049	0.6088387			
10	1.0000000	0.5431356	0.6432990	0.8730905	0.7443874	0.7288655	0.1583949	0.7819127	0.4292875		
11	1.0000000	0.7081218	0.4533965	0.7898833	0.8440414	0.8339949	0.4091201	0.8680436	0.6254072	0.3508353	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(k)

	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8792321									
4	1.0000000	0.9606701	0.5859031								
5	1.0000000	0.3378428	0.9383464	0.9803386							
6	1.0000000	0.3074350	0.9341579	0.9789720	0.3097614						
7	1.0000000	0.4209554	0.7275591	0.9061685	0.6643207	0.6449264					
8	1.0000000	0.4150369	0.9482423	0.9835516	0.3664129	0.1233397	0.7116572				
9	1.0000000	0.5332765	0.8404795	0.9473171	0.4630800	0.5736084	0.3579049	0.6088387			
10	1.0000000	0.5431356	0.6432990	0.8730905	0.7443874	0.7288655	0.1583949	0.7819127	0.4292875		
11	1.0000000	0.7081218	0.4533965	0.7898833	0.8440414	0.8339949	0.4091201	0.8680436	0.6254072	0.3508353	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

(l)

	1	2	3	4	5	6	7	8	9	10	11
2	1.00000000										
3	1.00000000	0.91602745									
4	1.00000000	0.95840708	0.69774920								
5	1.00000000	0.52622565	0.92858109	0.98690078							
6	1.00000000	0.29701398	0.87853692	0.97723971	0.47502676						
7	1.00000000	0.54163934	0.56039964	0.89366516	0.78985689	0.65975610					
8	1.00000000	0.40245714	0.90326127	0.98206449	0.52884849	0.12838202	0.72243409				
9	1.00000000	0.63119984	0.81855670	0.96509469	0.45861893	0.66592361	0.51937984	0.72346939			
10	1.00000000	0.50901054	0.77822581	0.87870968	0.81468754	0.69736022	0.09904153	0.75413712	0.56846473		
11	1.00000000	0.81273692	0.23367199	0.80084746	0.88747683	0.81144632	0.38985737	0.84872753	0.72321719	0.56114484	
12	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000

(m)

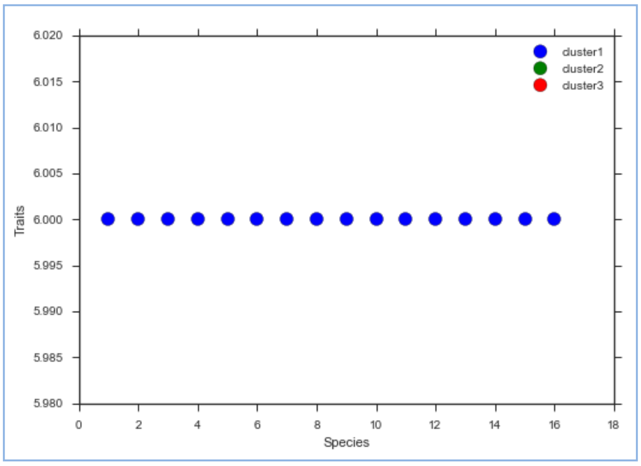
	1	2	3	4	5	6	7	8	9	10	11
2	1.00000000										
3	1.00000000	0.85431579									
4	1.00000000	0.95820365	0.76363636								
5	1.00000000	0.55574301	0.93194335	0.98104074							
6	1.00000000	0.47662338	0.91624304	0.97652934	0.17893787						
7	1.00000000	0.48834066	0.62795699	0.91044776	0.73288638	0.67889714					
8	1.00000000	0.42531889	0.93424553	0.98169782	0.17720000	0.12501382	0.74101950				
9	1.00000000	0.30540131	0.82100362	0.95019369	0.47564083	0.38440014	0.39849027	0.48605636			
10	1.00000000	0.51028807	0.60992108	0.92641261	0.74613333	0.69434932	0.06458192	0.75392039	0.42279709		
11	1.00000000	0.74277800	0.30522088	0.74731183	0.87585943	0.84823722	0.47689464	0.87994089	0.71990408	0.47064485	
12	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000

(n)

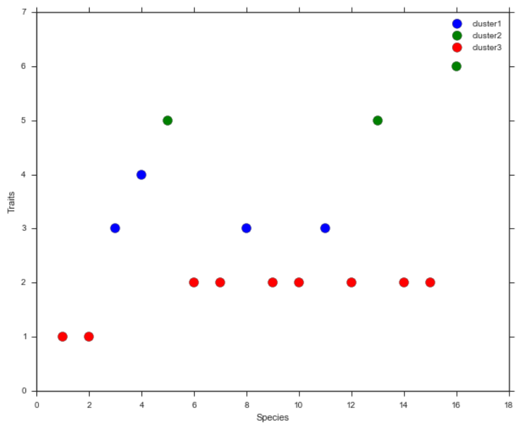
	1	2	3	4	5	6	7	8	9	10	11
2	1.0000000										
3	1.0000000	0.8792321									
4	1.0000000	0.9606701	0.5859031								
5	1.0000000	0.3378428	0.9383464	0.9803386							
6	1.0000000	0.3074350	0.9341579	0.9789720	0.3097614						
7	1.0000000	0.4209554	0.7275591	0.9061685	0.6643207	0.6449264					
8	1.0000000	0.4150369	0.9482423	0.9835516	0.3664129	0.1233397	0.7116572				
9	1.0000000	0.5332765	0.8404795	0.9473171	0.4630800	0.5736084	0.3579049	0.6088387			
10	1.0000000	0.5431356	0.6432990	0.8730905	0.7443874	0.7288655	0.1583949	0.7819127	0.4292875		
11	1.0000000	0.7081218	0.4533965	0.7898833	0.8440414	0.8339949	0.4091201	0.8680436	0.6254072	0.3508353	
12	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000

Figure 8-2: Community dissimilarity matrix for clusters of each trait subset. (a)- Plant height, (b)- Leaf length, (c)-Leaf area, (d)-Plant phenology, (e)-Plant height, Leaf length, (f)-Plant height, Leaf area, (g)-Plant height, Plant phenology, (h)-Leaf length, Leaf area, (i)- Leaf length, Plant phenology, (j)-Leaf area, Plant phenology, (k)- Plant height, Leaf length, Leaf area, (l)-Plant height, Leaf length, Plant phenology, (m)- Plant height, Leaf area, Plant phenology, (n)-Leaf length, Leaf area, Plant phenology, (o)- Plant height, Leaf length, Leaf area, Plant phenology

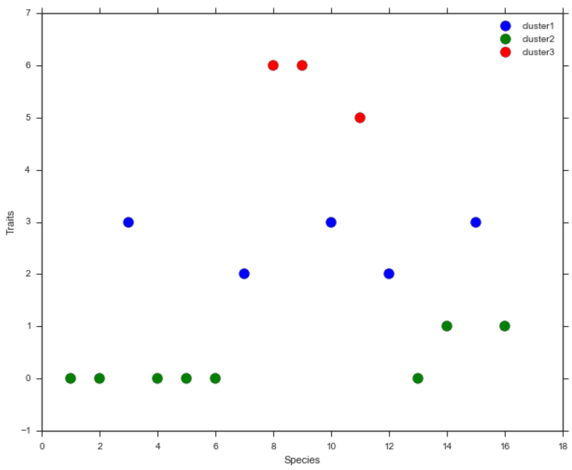
K-means clustering – Potential PFT



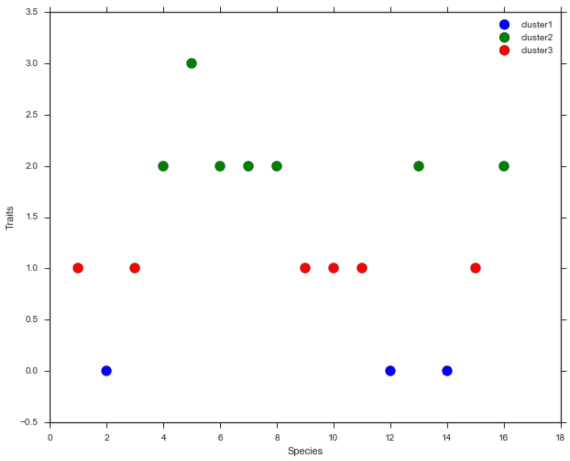
(a)



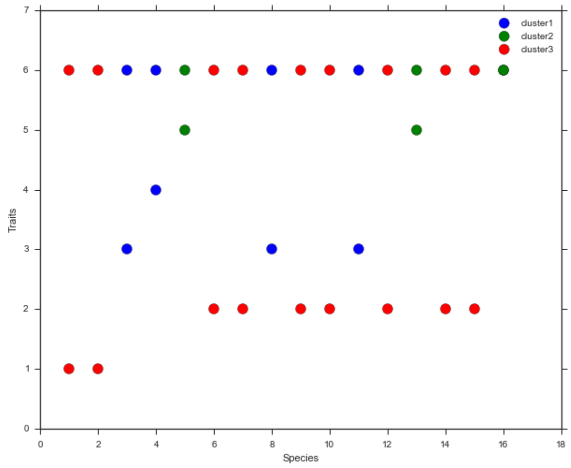
(b)



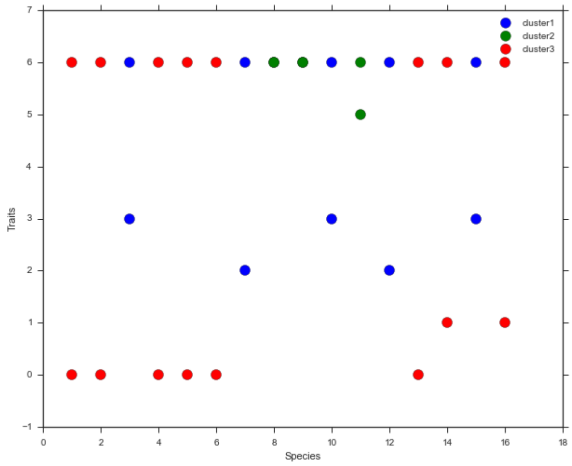
(c)



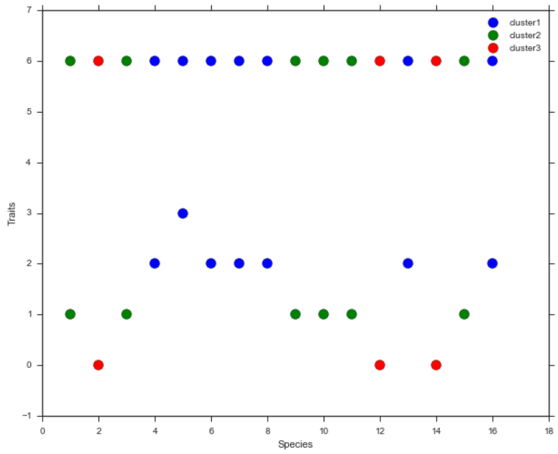
(d)


cluster1
cluster2
cluster3
Traits
Species

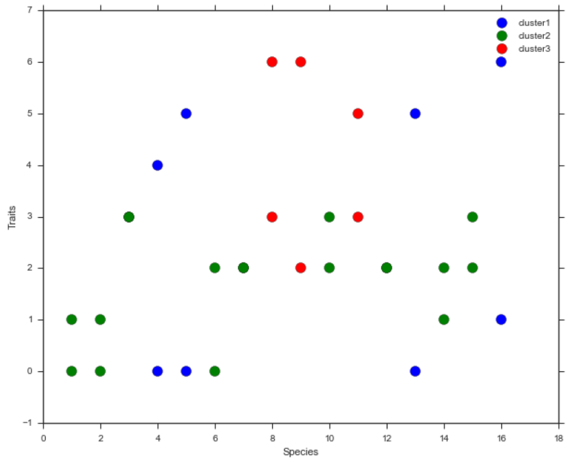
(e)


cluster1
cluster2
cluster3
Traits
Species

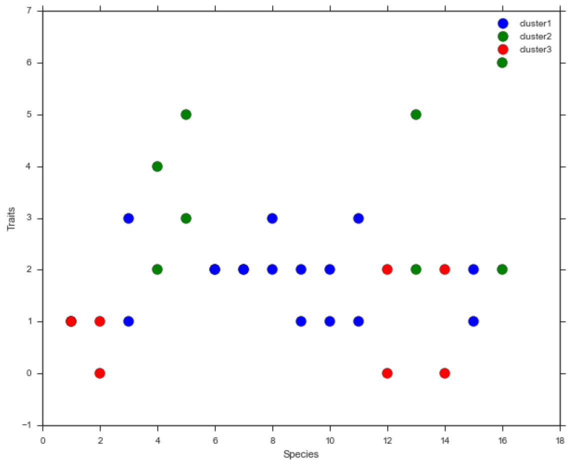
(f)


cluster1
cluster2
cluster3
Traits
Species

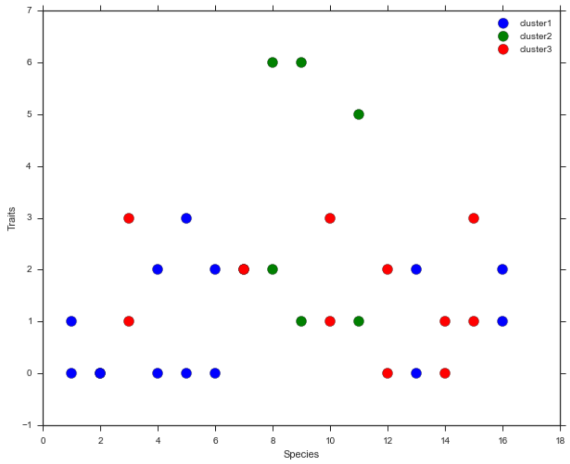
(g)


cluster1
cluster2
cluster3
Traits
Species

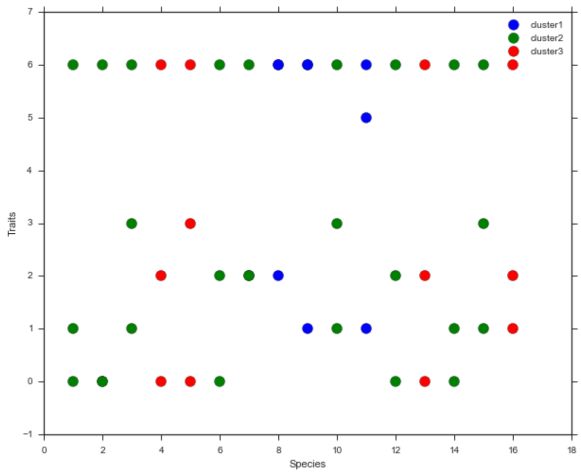
(h)


cluster1
cluster2
cluster3
Traits
Species

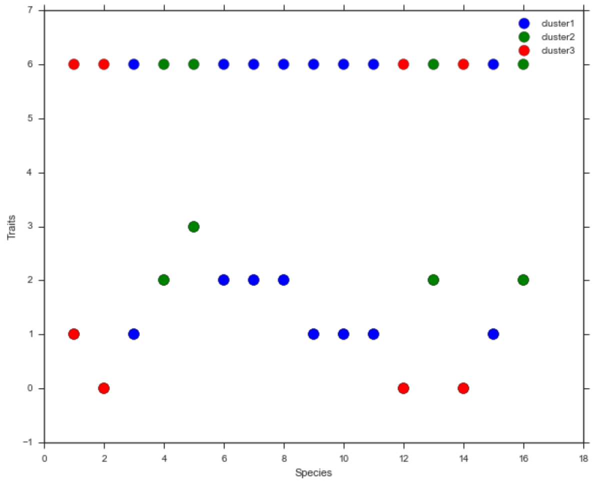
(i)


cluster1
cluster2
cluster3
Traits
Species

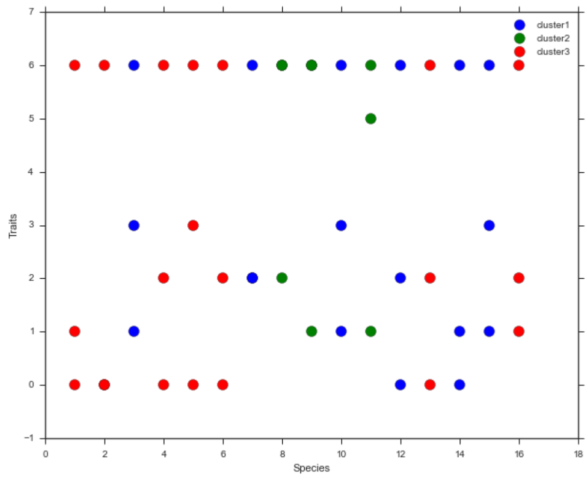
(j)



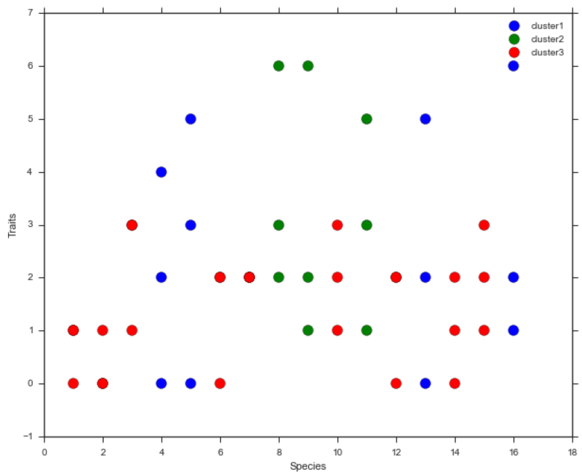
(k)



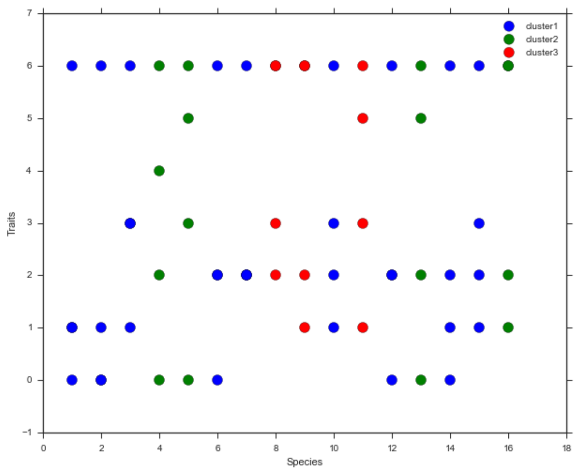
(l)



(m)



(n)



(o)

Figure 8-3: (a)- Plant height, (b)- Leaf length, (c)-Leaf area, (d)-Plant phenology, (e)-Plant height, Leaf length, (f)-Plant height, Leaf area, (g)-Plant height, Plant phenology, (h)-Leaf length, Leaf area, (i)- Leaf length, Plant phenology, (j)-Leaf area, Plant phenology, (k)- Plant height, Leaf length, Leaf area, (l)-Plant height, Leaf length, Plant phenology, (m)-

Plant height, Leaf area, Plant phenology, (n)-Leaf length, Leaf area, Plant phenology, (o)- Plant height, Leaf length, Leaf area, Plant phenology

Community dissimilarity- Potential PFT

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	0.00000000	1.00000000								
4	1.00000000	0.02040816	1.00000000							
5	0.00000000	1.00000000	0.00000000	1.00000000						
6	1.00000000	0.02040816	1.00000000	0.00000000	1.00000000					
7	1.00000000	0.02040816	1.00000000	0.00000000	1.00000000	0.00000000				
8	1.00000000	0.02040816	1.00000000	0.00000000	1.00000000	0.00000000	0.00000000			
9	0.00000000	1.00000000	0.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000		
10	0.00000000	1.00000000	0.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	0.00000000	
11	0.00000000	1.00000000	0.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	0.00000000	0.00000000

(a)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	1.00000000	1.00000000	1.00000000	0.75000000						
6	0.69230769	1.00000000	1.00000000	0.96078431	0.75438596					
7	0.45454545	1.00000000	1.00000000	1.00000000	1.00000000	0.77777778				
8	0.33333333	1.00000000	1.00000000	1.00000000	1.00000000	0.73913043	0.68421053			
9	0.47826087	1.00000000	1.00000000	1.00000000	1.00000000	0.78378378	0.03030303	0.70000000		
10	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	
11	0.73913043	1.00000000	1.00000000	0.95402299	0.71717172	0.08108108	0.81818182	0.70000000	0.82352941	1.00000000

(b)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	0.00000000								
4	1.00000000	1.00000000	1.00000000							
5	0.84126984	1.00000000	1.00000000	0.89090909						
6	0.80198020	1.00000000	1.00000000	0.45454545	0.73333333					
7	0.80645161	1.00000000	1.00000000	1.00000000	0.77011494	0.76744186				
8	0.70000000	1.00000000	1.00000000	1.00000000	0.74789916	0.35483871	0.64705882			
9	0.82352941	1.00000000	1.00000000	1.00000000	0.78835979	0.80198020	0.09677419	0.70000000		
10	1.00000000	1.00000000	1.00000000	0.42857143	0.74789916	0.61290323	1.00000000	1.00000000	1.00000000	
11	0.83783784	1.00000000	1.00000000	0.88679245	0.11764706	0.72413793	0.64705882	0.73913043	0.67567568	0.73913043

(c)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	1.00000000	1.00000000	1.00000000	0.75000000						
6	0.69230769	1.00000000	1.00000000	0.96078431	0.75438596					
7	0.45454545	1.00000000	1.00000000	1.00000000	1.00000000	0.77777778				
8	0.33333333	1.00000000	1.00000000	1.00000000	1.00000000	0.73913043	0.68421053			
9	0.47826087	1.00000000	1.00000000	1.00000000	1.00000000	0.78378378	0.03030303	0.70000000		
10	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000	
11	0.73913043	1.00000000	1.00000000	0.95402299	0.71717172	0.08108108	0.81818182	0.70000000	0.82352941	1.00000000

(d)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	1.00000000	1.00000000	1.00000000	0.00000000						
6	0.83870968	1.00000000	1.00000000	0.88000000	0.88000000					
7	0.50000000	1.00000000	1.00000000	1.00000000	1.00000000	0.89130435				
8	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000			
9	0.47826087	1.00000000	1.00000000	1.00000000	1.00000000	0.88826816	0.02857143	1.00000000		
10	0.33333333	1.00000000	1.00000000	1.00000000	1.00000000	0.81651376	0.71428571	1.00000000	0.70000000	
11	0.73913043	1.00000000	1.00000000	0.86813187	0.86813187	0.10614525	0.82857143	1.00000000	0.82352941	0.70000000

(e)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	0.58333333	1.00000000	1.00000000	0.62500000						
6	1.00000000	1.00000000	1.00000000	0.86666667	0.89090909					
7	0.06666667	1.00000000	1.00000000	1.00000000	0.62264151	1.00000000				
8	0.64705882	1.00000000	1.00000000	1.00000000	0.26829268	1.00000000	0.68421053			
9	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000		
10	0.17647059	1.00000000	1.00000000	0.88679245	0.58730159	0.83695652	0.11111111	0.73913043	1.00000000	
11	0.47826087	1.00000000	1.00000000	0.76470588	0.26760563	0.76744186	0.52000000	0.50000000	1.00000000	0.37931034

(f)

	1	2	3	4	5	6	7	8	9	10
2	0.70000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	1.00000000	1.00000000	1.00000000	0.78181818						
6	1.00000000	1.00000000	1.00000000	0.87755102	0.71739130					
7	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000				
8	1.00000000	1.00000000	1.00000000	1.00000000	0.78142077	0.06432749	1.00000000			
9	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	0.00000000	1.00000000		
10	1.00000000	1.00000000	1.00000000	0.75257732	0.07103825	0.75438596	1.00000000	0.82352941	1.00000000	
11	1.00000000	1.00000000	1.00000000	0.42857143	0.53125000	0.89655172	1.00000000	1.00000000	1.00000000	0.47826087

(g)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	0.00000000								
4	1.00000000	1.00000000	1.00000000							
5	0.90338164	1.00000000	1.00000000	0.87755102						
6	0.85915493	1.00000000	1.00000000	0.88461538	0.81052632					
7	1.00000000	0.00000000	0.00000000	1.00000000	1.00000000	1.00000000				
8	0.85000000	1.00000000	1.00000000	1.00000000	0.88700565	0.07103825	1.00000000			
9	0.76923077	1.00000000	1.00000000	1.00000000	0.81308411	0.73451327	1.00000000	0.70000000		
10	1.00000000	1.00000000	1.00000000	0.84210526	0.13580247	0.90476190	1.00000000	1.00000000	1.00000000	
11	1.00000000	1.00000000	1.00000000	0.42857143	0.71962617	0.85840708	1.00000000	1.00000000	1.00000000	0.64705882

(h)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	0.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	1.00000000	1.00000000	1.00000000	0.61904762						
6	1.00000000	1.00000000	1.00000000	0.71428571	0.67857143					
7	1.00000000	1.00000000	1.00000000	1.00000000	0.59183673	0.87012987				
8	0.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000			
9	1.00000000	1.00000000	1.00000000	1.00000000	0.64601770	0.88165680	0.09677419	1.00000000		
10	1.00000000	1.00000000	1.00000000	0.52542373	0.36986301	0.37984496	0.73913043	1.00000000	0.76923077	
11	1.00000000	1.00000000	1.00000000	0.71717172	0.59292035	0.05325444	0.80645161	1.00000000	0.82352941	0.30769231

(i)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	0.31034483	1.00000000	1.00000000	0.91489362						
6	0.69072165	1.00000000	1.00000000	0.85185185	0.70833333					
7	0.14285714	1.00000000	1.00000000	1.00000000	0.41176471	0.60784314				
8	0.70000000	1.00000000	1.00000000	1.00000000	0.79797980	0.16167665	0.61904762			
9	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000		
10	0.70000000	1.00000000	1.00000000	0.15151515	0.71717172	0.67664671	0.71428571	0.82352941	1.00000000	
11	0.70000000	1.00000000	1.00000000	0.81818182	0.71717172	0.01796407	0.61904762	0.17647059	1.00000000	0.64705882

(j)

	1	2	3	4	5	6	7	8	9	10
2	1.0000000									
3	1.0000000	0.0000000								
4	1.0000000	1.0000000	1.0000000							
5	1.0000000	1.0000000	1.0000000	0.9090909						
6	0.7979798	1.0000000	1.0000000	0.7500000	0.8571429					
7	0.4400000	1.0000000	1.0000000	1.0000000	1.0000000	0.6296296				
8	1.0000000	0.0000000	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000			
9	0.8378378	1.0000000	1.0000000	1.0000000	1.0000000	0.3023256	0.6842105	1.0000000		
10	0.4814815	1.0000000	1.0000000	0.9230769	0.8369565	0.5625000	0.1111111	1.0000000	0.7391304	
11	1.0000000	1.0000000	1.0000000	0.5789474	0.6969697	0.6279070	1.0000000	1.0000000	1.0000000	0.7391304

(k)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	0.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	1.00000000	1.00000000	1.00000000	0.61904762						
6	1.00000000	1.00000000	1.00000000	0.71428571	0.67857143					
7	1.00000000	1.00000000	1.00000000	1.00000000	0.59183673	0.87012987				
8	0.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000			
9	1.00000000	1.00000000	1.00000000	1.00000000	0.64601770	0.88165680	0.09677419	1.00000000		
10	1.00000000	1.00000000	1.00000000	0.52542373	0.36986301	0.37984496	0.73913043	1.00000000	0.76923077	
11	1.00000000	1.00000000	1.00000000	0.71717172	0.59292035	0.05325444	0.80645161	1.00000000	0.82352941	0.30769231

(l)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	1.00000000								
4	1.00000000	1.00000000	1.00000000							
5	0.58333333	1.00000000	1.00000000	0.64705882						
6	1.00000000	1.00000000	1.00000000	0.82222222	0.88888889					
7	0.06666667	1.00000000	1.00000000	1.00000000	0.62264151	1.00000000				
8	0.64705882	1.00000000	1.00000000	1.00000000	0.26829268	1.00000000	0.68421053			
9	1.00000000	1.00000000	0.00000000	1.00000000	1.00000000	1.00000000	1.00000000	1.00000000		
10	0.17647059	1.00000000	1.00000000	0.85185185	0.58730159	0.83516484	0.11111111	0.73913043	1.00000000	
11	0.47826087	1.00000000	1.00000000	0.69811321	0.26760563	0.76377953	0.52000000	0.50000000	1.00000000	0.37931034

(m)

	1	2	3	4	5	6	7	8	9	10
2	1.0000000									
3	1.0000000	1.0000000								
4	1.0000000	1.0000000	1.0000000							
5	0.3103448	1.0000000	1.0000000	0.6000000						
6	0.8245614	1.0000000	1.0000000	0.9259259	0.8352941					
7	0.1428571	1.0000000	1.0000000	1.0000000	0.4117647	0.7727273				
8	0.7000000	1.0000000	1.0000000	1.0000000	0.7979798	0.4190871	0.6190476			
9	1.0000000	1.0000000	0.0000000	1.0000000	1.0000000	1.0000000	1.0000000	1.0000000		
10	0.7000000	1.0000000	1.0000000	0.8681319	0.7171717	0.2946058	0.7142857	0.8235294	1.0000000	
11	1.0000000	1.0000000	1.0000000	0.4285714	0.7241379	0.8245614	1.0000000	1.0000000	1.0000000	0.7000000

(n)

	1	2	3	4	5	6	7	8	9	10
2	1.00000000									
3	1.00000000	0.00000000								
4	1.00000000	1.00000000	1.00000000							
5	0.30232558	1.00000000	1.00000000	0.85454545						
6	0.35483871	1.00000000	1.00000000	0.87755102	0.27272727					
7	0.64705882	1.00000000	1.00000000	1.00000000	0.59183673	0.76744186				
8	0.00000000	1.00000000	1.00000000	1.00000000	0.30232558	0.35483871	0.64705882			
9	0.70000000	1.00000000	1.00000000	1.00000000	0.64601770	0.80198020	0.09677419	0.70000000		
10	1.00000000	1.00000000	1.00000000	0.69072165	0.62790698	0.61290323	1.00000000	1.00000000	1.00000000	
11	0.83783784	1.00000000	1.00000000	0.44444444	0.71717172	0.82795699	0.41666667	0.83783784	0.33333333	0.83783784

(o)

Figure 8-4: shows community dissimilarity matrix for clusters of each trait subset. (a)- Plant height, (b)- Leaf length, (c)-Leaf area, (d)-Plant phenology, (e)-Plant height, Leaf length, (f)-Plant height, Leaf area, (g)-Plant height, Plant phenology, (h)-Leaf length, Leaf area, (i)- Leaf length, Plant phenology, (j)-Leaf area, Plant phenology, (k)- Plant height, Leaf length, Leaf area, (l)-Plant height, Leaf length, Plant phenology, (m)- Plant height, Leaf area, Plant phenology, (n)-Leaf length, Leaf area, Plant phenology, (o)- Plant height, Leaf length, Leaf area, Plant phenology