<u>Template/Pro forma for Submission</u>

NMHS-Himalayan Institutional Project Grant

NMHS-FINAL TECHNICAL REPORT (FTR)

Demand-Driven Action Research and Demonstrations

NMHS Reference No.:	GBPNI/NMHS-2017-18/MG-18/554 dated 26.02.2018

Date of Submission:	3	0	0	3	2	0	2	2
	d	**d**	**m**	**m**	**y**	**y**	**y**	**y**

PROJECT TITLE (IN CAPITAL)

HYPERSPECTRAL IMAGING FOR SHARPER DEFINITIONS OF HIMALAYAN ECOSYSTEMS AND ITS HIGH-VALUE PLANT SPECIES UNDER CLIMATE UNCERTAINTIES

Project Duration: *from* (**<u>26.02.2018</u>**) *to* (**<u>30.11.2021</u>**).

***Submitted to*:**

Er. Kireet Kumar

Scientist 'G' and Nodal Officer, NMHS-PMU

National Mission on Himalayan Studies, GBP NIHE HQs

Ministry of Environment, Forest & Climate Change (MoEF&CC), New Delhi

E-mail: nmhspmu2016@gmail.com; kireet@gbpihed.nic.in; kodali.rk@gov.in

***Submitted by*:**

[*Dr Prashant K. Srivastava*]

[*Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh*]

[*Contact No.: 75719 27744*]

[*E-mail: prashant.just@gmail.com*]

NMHS-Final Technical Report (FTR) *template*

Demand-Driven Action Research Project

DSL: Date of Sanction Letter

2	8	0	2	**2**	**0**	**1**	**8**
d	d	m	m	y	y	y	y

DPC: Date of Project Completion

3	0	1	1	**2**	**0**	**2**	**1**
d	d	m	m	y	y	y	y

Part A: Project Summary Report

1. Project Description

i.	Project Reference No.	GBPNI/NMHS-2017-18/MG-18/554 dated 26.02.2018					
ii.	Type of Project	Small Grant		Medium Grant	√	Large Grant	
iii.	Project Title	**Hyperspectral Imaging for Sharper Definitions of Himalayan Ecosystems and Its High-Value Plant Species Under Climate Uncertainties**					
iv.	State under which Project is Sanctioned	**Uttar Pradesh**					
v.	Project Sites (the IHR States covered) (Maps to be attached)	Uttarakhand, Jammu & Kashmir					
vi.	The scale of Project Operation	Local		Regional		Pan-Himalayan	√
vii.	Total Budget/ Outlay of the Project	2,23,74,560					

viii.	Lead Agency	*Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh*
	Principal Investigator (PI)	Dr Prashant Kumar Srivastava Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh
	Co-Principal Investigator (Co-PI)	• Prof. Rajesh Kumar Mall, Professor, Coordinator DST Mahamana Centre for Excellence in Climate Change Research, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, UP • Prof. A S Raghubanshi, Professor, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, UP • K. Chandra Sekhar, Scientist F, G.B. Pant National Institute of Himalayan Environment & Sustainable Development, Almora, Uttarakhand • Dr Irfan Rashid, Assistant Professor, Department of Earth Sciences, University of Kashmir • Dr K V Ramesh, CSIR Fourth Paradigm Institute, NAL Belur Campus, Bangalore, India
ix.	Project Implementing Partners	1. Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, UP 2. G.B. Pant National Institute of Himalayan Environment & Sustainable Development, Almora, Uttarakhand 3. Department of Earth Sciences, University of Kashmir 4. CSIR Fourth Paradigm Institute, NAL Belur Campus, Bangalore, India
	Key Persons / Point of Contacts with Contact Details, Ph. No, E-mail	Dr Prashant Kumar Srivastava Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh prashant.iesd@bhu.ac.in, prashant.just@gmail.com +91-7571927744

2. Project Outcomes

2.1. Abstract

The Quantification of biophysical and biochemical parameters is very crucial for monitoring the high-value plants in the Himalayan region, their sustainable utilization as well as the conservation of related resources. In all, there were six objectives in the project. The first one was the **creation of the Spectral library and standardization for selected medicinal and threatened species**. The second one was **Spectral and image analysis for discrimination of selected economically important plant species** followed by **Model building for retrieval of Biochemical and Biophysical properties of selected species**. **Projection and scenario for distribution of selected species** was the fourth objective followed by the fifth objective of **fine scaling of space-time map of the high-value species in the Himalayas**. The final objective was the **management and conservation of species based on previous objectives**. To fulfil these objectives, extensive field samplings were conducted in Pindari, Nandadevi biosphere reserve, Jageshwer, Almora, Ranikhet, Gwaldham, and Munsiyari regions in Uttarakhand and, Dachigam National National Park, Gulmarg Wildlife sanctuary in Kashmir. A total of 400 plus species were sampled during the study and a spectral library of species was prepared.

The First objective resulted in classifying the spectral species through collected spectra for which PRISMA data were acquired. Classification techniques namely spectral angular mapper (SAM) were used to classify the species over the acquired Hyperspectral data.

The second objective resulted in the development of a spectral library of the identified species. For this **SHRIME** (Spectral Tool of Himalayan Rare, Invasive, Medicinal, and Economical Plant Species) was developed.

The third objective resulted in the retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species *Taxus wallichiana*. The development and validation of hyperspectral indices for anti-cancerous Taxol content estimation were conducted as part of this objective. Further, Taxol Content Variation was also explored in altitudinal variation in the Himalayan regions. The species distribution of Taxol content through empirical methods and radiative transfer models (physical methods) were conducted. Taxol Mapping was also explored through drone quasi-hyperspectral images. As part of this objective, a list of tentative compounds identified in *Rhododendron arboreum* using UPLC and HRMS was prepared. Sampling and Analysis of Viburnum grandiflorum and Parrotiosis jacquemontiana were also done for its biochemical and biophysical properties.

For the fourth objective, fine-grained species distributions were modelled using Sentinel-2, Sentinel-5P, Moderate Resolution Imaging Spectroradiometer (MODIS), ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) and Shuttle Radar Topography Mission (SRTM) along with other climatic variables over the sampling regions. Several machine learning algorithms were incorporated to establish the relation between physical and climatic variables to estimate the probability distribution of species. The phenology assessment was carried out for two decades to the effect of an elevational gradient, temperature, and precipitation on the start of the season (SOS) and end of the season (EOS) in major forest types of the Kumaon region of the western Himalaya using MODIS NDVI time series data (2001-2019). The study disclosed that due to winter warming and summer dryness, despite a warming trend in pre-season or springtime the onset of the vegetation growth cycle shows a delayed trend across the vegetation types. The potential distribution of Rhododendron arboreum, a medicinal plant species found within the foothills of the Himalayas was also evaluated using Conventional Species Distribution Models (SDM) namely BIOCLIM and Maxent, a Machine Learning variant as well as Convolutional Neural Network (CNN).

For the fifth objective, High-resolution weather information is generated for the study region for 2000-2016 and CMIP6 climate scenarios are organized. As part of this objective, Investigation into the possibility of applying functional variables retrieved from Moderate Resolution Imaging Spectroradiometer (MODIS) onboard sensor data to map the most realistic species distribution limits of two alpine treeline species, namely *Betula utilis D.Don* and *Rhododendron campanulatum* over the Himalayan biodiversity hotspot was conducted. Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. For each model, the congruence of predictions was assessed and made pairwise comparisons to assess the performance of the models. As part of this objective, bias-corrected, statistically downscale models drawn from the NASA, Earth Exchange Global Daily Downscaled Projections - Coupled Model Intercomparison Project Phase 5 (NEX-GDDP-CMIP5) were examined.

As part of objective 6, Traditional knowledge among communities residing in alpine regions of Western Himalaya was conducted. Gunji, Kuti, Napalchyu, Navi and Kutti of Byans Valley Pithoragarh villages were taken for this study. Questionnaire-based surveys gathered information on demographic profiles, biodiversity, and traditional knowledge where 75 informants of different age groups from five villages were selected. The Perceptions were recorded for 31 high-value medicinal and economical plants (belonging to 24 families) in which Maximum high-value species (23) were identified as herbs, followed by shrubs (5) and trees (3).

2.2. Objective-wise Major Achievements

S. No.	Objectives	Major achievements (in bullets points)
1.	Detect and identify the medicinal plants, rare, endangered, and other economically important plant species concerning pedological and climatic conditions using the handheld Hyperspectral Spectroradiometer, AVIRIS Next Generation (NG) data and field observations in the Himalayas.	• For classifying the spectral species through collected spectra and PRISMA data were acquired. • Classification techniques namely spectral angular mapper (SAM) were used to classify the species over the acquired Hyperspectral data. • Spectral and image analysis for discrimination of selected economically important plant species such as Berberis, Polygonum polystachyum, Rhododendron Barbatum, Rhododendron Arboreum, Taxus Wallichiana, Bamboo and Fern are over.

2.	Spectral library of the economically important plant species to use with the Hyperspectral satellite and airborne data for large-scale quantification.	• Field surveys were conducted over study areas to collect plant samples. • The collected samples (fresh twigs and flowers) of each species were utilized for the preparation of their spectral libraries where the spectra in form of reflectance were recorded and stored using a hyperspectral Spectroradiometer. • Development of a spectral library of the identified species. For this **SHRIME** (Spectral Tool of Himalayan Rare, Invasive, Medicinal, and Economical Plant Species) was developed. • This could be utilized to create a spectral library in terms of reflectance and first derivative for the Himalayan plant species.

3.	Development of forward and inverse models for the retrieval of biophysical and biochemical parameters from economically important plant species using the Hyperspectral data.	• Retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species *Taxus wallichiana*. • For the retrieval of biochemical parameters, various indices were generated and three different types of filters, Average Mean, Savitzky Golay, and Fast Fourier Transform (FFT) one at a time followed by feature selection on each denoised spectra were applied to smooth the spectra for better estimation of biochemical parameters. • The development and validation of hyperspectral indices for anti-cancerous Taxol content estimation were conducted as part of this objective. • Further, Taxol Content Variation was also explored in altitudinal variation in the Himalayan regions. • The species distribution of Taxol content through empirical methods and radiative transfer models (physical methods) were conducted. Taxol Mapping was also explored through drone quasi-hyperspectral images. • As part of this objective, a list of tentative compounds identified in *Rhododendron arboreum* using UPLC and HRMS was prepared. • The algorithm for retrieval of biochemical properties of Taxus Wallichiana, and Pinus are over and Rhododendron, Parrotiopsis jacquemontiana, and Viburnum grandiflorum are underway.

4.	To provide fine-scale space-time inventory of the high-value medicinal plants, rare, endangered, and other economically important plant species.	• Fine-grained species distributions were modelled using Sentinel-2, Sentinel-5P, Moderate Resolution Imaging Spectroradiometer (MODIS), ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) and Shuttle Radar Topography Mission (SRTM) along with other climatic variables over the sampling regions. • Several machine learning algorithms were incorporated to establish the relation between physical and climatic variables to estimate the probability distribution of species. • The phenology assessment was carried out for two decades to the effect of an elevational gradient, temperature, and precipitation on the start of the season (SOS) and end of the season (EOS) in major forest types of the Kumaon region of the western Himalaya using MODIS NDVI time series data (2001-2019). • The potential distribution of Rhododendron arboreum, a medicinal plant species found within the foothills of the Himalayas was also evaluated using Conventional Species Distribution Models (SDM) namely BIOCLIM and Maxent, a Machine Learning variant as well as Convolutional Neural Network (CNN).

5.	Projection of future climate change scenarios and its possible effect on the distribution of medicinal plants, rare, endangered and other economically important plant species.	• High-resolution weather information is generated for the study region for 2000-2016 and CMIP5 climate scenarios are provided. • As part of this objective, Investigation into the possibility of applying functional variables retrieved from Moderate Resolution Imaging Spectroradiometer (MODIS) onboard sensor data to map the most realistic species distribution limits of two alpine treeline species, namely Betula utilis D.Don and Rhododendron campanulatum over the Himalayan biodiversity hotspot was conducted. • Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. For each model, the congruence of predictions was assessed and made pairwise comparisons to assess the performance of the models. • As part of this objective, bias-corrected, statistically downscale models drawn from the NASA, Earth Exchange Global Daily Downscaled Projections - Coupled Model Intercomparison Project Phase 5 (NEX-GDDP-CMIP5) were examined.

6.	Development of knowledge-based planning of an appropriate management system to safeguard medicinal plants, rare, endangered and other economically important plant species from unplanned harvesting, Identification of new ecological niches in higher Himalayas under climate change to protect the species and improve the livelihood conditions of local and regional communities.	• Traditional knowledge among communities residing in alpine regions of Western Himalaya was conducted. Gunji, Kuti, Napalchyu, Navi and Kutti of Byans Valley Pithoragarh villages were taken for this study. • Questionnaire-based surveys gathered information on demographic profiles, biodiversity, and traditional knowledge where 75 informants of different age groups from five villages were selected. The Perceptions were recorded for 31 high-value medicinal and economical plants (belonging to 24 families) in which Maximum high-value species (23) were identified as herbs, followed by shrubs (5) and trees (3).

2.3. Outputs in terms of Quantifiable Deliverables*

S. No.	Quantifiable Deliverables*	Monitoring Indicators*	Quantified Output/ Outcome achieved	Deviations made, if any, & Reason thereof:
1.	Spectral library and standardization for selected medicinal and threatened species;	Periodic, regular updates on Spectral Library of the economically important plant species (Nos)	**SHRIME** (Spectral Tool of Himalayan Rare, Invasive, Medicinal, and Economical Plant Species) tool is live and running.	No deviation.
2.	Spectral and image analysis for discrimination of selected economically	No. of New Database/ Datasets generated on the identified	Apart from DB created in terms of the SHRIME tool, where the spectral signature of more than 250 species are captured, there are five	No deviation.

	important plant species;	dynamics, i.e., biochemical, and biophysical properties of selected species.	species whose dynamics in terms of biochemical and biophysical properties are captured. The species are *Betula utilis D.Don*, *Rhododendron campanulatum, Taxus wallichiana, Rhododendron arboretum and Pyracantha crenulata D.Don.*	
3.	Model for retrieval of Biochemical and Biophysical properties of selected species;	Fine-scale space-time distribution maps of economically important species in the target area	Fine-scale space-time distribution maps were generated for the species *Betula utilis D.Don*, *Rhododendron campanulatum, Taxus wallichiana, Rhododendron arboretum and Pyracantha crenulata D.Don.*	No deviation.
4.	Projection and scenario for distribution of selected species;	Assessment report/ manual on safeguarding the medicinal and threatened plant species.	The species distribution limits of two alpine treeline species, namely Betula utilis D.Don and Rhododendron campanulatum over the Himalayan biodiversity hotspot were conducted. Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs	No deviation.

			combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. This not only has helped in the assessment of the prevalent condition but would also help in the identification of the practices required for safeguarding the medicinal and threatened plant species.	
5.	Fine-scale space-time map of the high-value species in the Himalayas;	Policy/ Strategic frameworks/ drafts (No.) for assisting the regional decision-making in the identified area and dynamics.	Traditional knowledge among communities residing in alpine regions of Western Himalaya was conducted. Gunji, Kuti, Napalchyu, Navi and Kutti of Byans Valley Pithoragarh villages were taken for this study. 77% of respondents claimed that the medicinal plants are decreasing in their natural surroundings 92% of medicinal and economical importance plants harvest from the wild population for their household	No deviation.

			use and commercial trading 45% of respondents agreed with Govt. subsidy-based medicinal plant cultivation, while 55% refused to cite they would go for the cultivation of economical plant species than medicinal ones.	
6.	Plan for management and sustainable harvest of species based on above.	Other Publications and knowledge products (Nos.)	13 publications as a whole are published and 7 were presented at reputed conferences. Couple of publications are under review. Complete list is given in Appendix 2.	No deviation.

(*) As stated in the Sanction Letter issued by the NMHS-PMU.

2.4. Strategic Steps with respect to Outcomes (in bullets)

S. No.	Particulars	Number/ Brief Details	Remarks/ Attachment
1.	New Methodology developed	5	Details are given in Appendix 5
2.	New Models/ Process/ Strategydeveloped	14	Details are given in Appendix 5
3.	New Species identified	0	NA

S. No.	Particulars	Number/ Brief Details	Remarks/ Attachment
4.	New Database established	6	1. Developed SHRIME tool which provides a spectral library for Himalayan economic and medicinally important species. 2. Development of novel earth observation variables (Eco-physiological and phenological) for accurate niche modelling in the Himalayas. 3. Estimations of biophysical parameters such as Chlorophyll content, Leaf Area Index (LAI), Diameter at Breast Height (DBH), Leaf water content, Height, Soil moisture, Soil roughness 4. Biodiversity Mapping through PRISMA 5. Fine scale inventory and spatial database for biochemical and bipphysical paramaters were developed. 6. Robust DB of biochemical parameters of various medicinal plant species were developed. Details are provided in Appendix 5
5.	New Patent, if any	Under process	IPR Under process
	I. Filed (Indian/ International)		
	II. Granted (Indian/ International)		
	III. Technology Transfer (if any)		
6.	Others (if any)	NA	NA

3. Technological Intervention

S. No.	Type of Intervention	Brief Narration on the interventions	Unit Details (No. of villagers benefited / Area Developed)
1.	Development and deployment of indigenous technology	The random stratified sampling technique was used to collect data on tree species in the field. Sample plots were laid across the forests of Almora and Pithoragarh districts of Uttarakhand and Dachigam National Park (Jammu & Kashmir). Sample plots were selected across all forest types in the study areas. A total of 45 sample plots were laid based on a stratified random sampling approach. A square-shaped plot of 0.1 ha (31.6 × 31.6 m) was established, and data was collected between 2019 and 2021. Field sample plot geo-locations are recorded using the Global Positioning System (GPS). Tree species were identified by local taxonomists' expertise and regional floras. GBH (Girth at Breast Height) were measured at ~130 cm from the ground using 5 m self-retractable metric measuring tapes, and heights were recorded for all the trees inside the plot that had greater than 10 cm GBH using a Nikon forestry pro laser range finder. The collected data were evaluated for estimation of tree biomass across forests of study sites.	

		Conventional in situ estimations of biophysical parameters such as Chlorophyll content, Leaf Area Index (LAI), Diameter at Breast Height (DBH), Leaf water content, Height, Soil moisture, Soil roughness were carried for various forest species. Measured DBH and Height of different forest tree species was used to estimate above ground biomass of these species using species-specific allometric volume equations. Quadrat sampling method was adopted for species diversity measurement. Plots of 30 m × 30 m were considered for the quadrat sampling, which was divided into 9 subplots with 10 m × 10 m size. In the centre of each sampling subplot, the geographical latitude and longitude readings were collected using Garmin eTrex 10 GPS. Within each subplot, the total number of species with > 20 cm Girth at Breast Height (GBH) (1.3 m) was thoroughly counted, which was used in estimating different species diversity indices like Shannon-Weiner Index, Margalef Index, McIntosh index, Brillouin index and Simpson index etc. Plant samples were also collected for laboratory estimation of different plant biochemical estimation such as carotenoids, taxol, berberine, etc . Spectra of collected leaf samples were measured using ASD FieldSpec Spectroradiometer. Simultaneously drone data were also collected for the study area.	

2.	Diffusion of High-end Technology in the region	Details are given in Appendix 5.	
3.	Induction of New Technology in the region	Details are given in Appendix 5.	
4.	Publication of Technological / Process Manuals	Details are given in Appendix 2.	
	Others (if any)		

4. New Data Generated over the Baseline Data

S. No.	New Data Details	Status of Existing Baseline	Additionality and Utilisation New data
1	Developed SHRIME tool which provides a spectral library for Himalayan economic and medicinally important species.	As such, no benchmark data was available to date. This would be the first of its kind baseline	This would help the upcoming researchers in adding more to the existing literature.
2	Development of novel earth observation variables (Eco-physiological and phenological) for accurate niche modelling in the Himalayas.	There exists few ecological or phenological data in past literature but this perspective of taking Eco-physiological and phenological variables for any study is explored for the first time.	This would help the upcoming researchers in adding more to the existing literature.
3	High Resolution UAV Datasets (Thermal/Visible/IR)	As such, no benchmark data was available to date. This would be the first of its kind baseline	This would help the upcoming researchers in adding more to the existing literature.
4	Geotagged Data of Economical Species	Very few data available.	The current deliverables would help building a robust benchmark to study IHR at hyper local level.

5	Geotagged Data of Medicinal Species	Very few data available.	The current deliverables would help building a robust benchmark to study IHR at hyper local level.

5. Demonstrative Skill Development and Capacity Building/ Manpower Trained

S. No.	Type of Activities	Details with number	Activity Intended for	Participants/Trained			
				SC	**ST**	**Woman**	**Total**
1.	Workshops	**5**	1. 2nd National Workshop on "Techniques in Hyperspectral Data Analysis and Processing" (27 January - 31 January 2020)			**15**	**42**
			2. Online Workshop on "Species Distribution Modelling Using R" (15-16 March 2021)			**17**	**28**
			3. One-week Online Workshop on "QGIS, Image Processing and Species Distribution Modelling" (30 August - 04 September 2021).			**19**	**33**
			4. Drone Training			**2**	**6**
			5. Graduate Seminars			**43**	**63**

2.	On Field Training	**6**	Extensive field samplings were conducted in Pindari, Nandadevi biosphere reserve, Jageshwer, Almora, Ranikhet, Gwaldham, and Munsiyari regions in Uttarakhand and, Dachigam National National Park, Gulmarg Wildlife sanctuary in Kashmir.			**8**	**45**
3.	Skill Development	**1**	A Six Lecture Course on Data Assimilation (22-24 March 2022)			**3**	**15**
4.	Academic Supports	**5**	Training were conducted for various on field sampling using Drone, Spectroradiometer. The biochemical treatment of the collected samples were also demonstrated at University level. A complete overview of data generation to storage to data science concepts were also taught at the PG/PhD level students.			**5**	**20**

6. Linkages with Regional & National Priorities (SDGs, INDC, etc)/ Collaborations

S. No.	Linkages /collaborations	Details	No. of Publications/ Events Held	Beneficiaries

1.	Sustainable Development Goal (SDG)	1. The Species Distribution Models (SDMs) studies carried out in this grant contribute to SDG 15 and are crucial for biodiversity conservation in the Himalayas. The output data can be utilised for conservation, restoration activities and sustainable utilisation of resources. 2. Mitigating the effects of climate change on treeline species is one of the most important global ecological problems. Future projections of climatic species niche changes constitute the backbone for impact mitigation and adaptive management strategies in the Himalayas in response to short- and long-term changes. This attempt gives insight concerning SDG 13. 3. The plan for management and sustainable harvest of species provided in this project is contributing to SDG 12 (Sustainable Consumption and Production Patterns), more clearly to 12.2 4. Biophysical variables are crucial in understanding ecosystem functioning and processes, and thus have a key role in biodiversity conservation. Monitoring of biophysical parameters in the present grant contributes to SDG 15.4 and SDG 15.1.	1 Publication 1 Publication	Biodiversity conservation and management departments Climate change impact mitigation agencies Indigenous people

2.	Climate Change/INDC targets	India's commitment to achieving five nectar elements (Panchamrit) as their climate action plan. The outcomes of the current project would help propagate a healthy and sustainable way of living based on traditions and values of conservation and moderation, through the motto 'LIFE'– 'Lifestyle for Environment' as a key to combating climate change". The outcomes would also help the policymakers in better adaptation to climate change by enhancing investments in development programmes in the Himalayan region by identification of prone areas. The project has trained many of the resources from the technological perspective as well as from the out of box thinking. This would help in building capacities, creating the domestic framework and international architecture for quick diffusion of cutting-edge climate technology in India and joint collaborative R&D for such future technologies.	1 Publication (under communication)	
3.	International Commitments	Carbon maps derived from remote sensing are widely used by scientists and policymakers. The carbon stock assessment undertaken as part of this award contributes to the Paris Agreement's goals (2015). The biophysical variables and SDMs developed in this grant contribute to the GEOBON Essential Biodiversity Variables (EBVs).		

4.	Bilateral engagements	• G.B. Pant National Institute of Himalayan Environment & Sustainable Development, Almora, Uttarakhand • Department of Earth Sciences, University of Kashmir • CSIR Fourth Paradigm Institute, NAL Belur Campus, Bangalore, India • Center for Quantitative Economics and Data Science, Birla Institute of Technology, Mesra, Ranchi, Jharkhand, 835215, India • Conservation Ecology Lab, Department of Botany, Baba Ghulam Shah Badshah University Rajouri, Jammu and Kashmir, 185234, India • Dr. Harisingh Gour Central University, Sagar - 470003, MP • CSIR - Central Institute of Medicinal and Aromatic Plants, Lucknow, India • Centre for Oceans, Rivers, Atmosphere and Land Sciences (CORAL), Indian Institute of Technology, Kharagpur 721302 • Shiv Nadar University, Greater Noida	9	
5.	National Policies	As a part of India's National Action Plan on Climate Change, sustaining the Himalayan Ecosystems and adaptation efforts would be boosted by the outcome of this study.		

6.	Other collaborations	• Department of Geography, Harokopio University of Athens, El. Venizelou 70, Kallithea, 17671 Athens, Greece • Department of Forestry and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI 53706, US	4	

7. Project Stakeholders/ Beneficiaries and Impacts

S. No.	Stakeholders	Support Activities	Impacts
1.	Gram Panchayats	Provided information on medicinal plants.	Policy development for extraction of medicinal plants from the wild and their cultivation on farmlands.
2.	Govt Departments (Agriculture/ Forest)	India's National Action Plan on Climate Change.	Sustaining the Himalayan Ecosystems and adaptation efforts would be boosted by the outcome of this study.
3.	Villagers	Collection of medicinal plant species, their identification and biochemical treatment were carried out with villagers at Khati, Dwali and Furkiya.	Importance of species for maintaining the biodiversity amid climate change challenges has offered the much needed awareness drive amond the villagers
4.	SC Community	NA	NA
5.	ST Community	NA	NA
6.	Women Group	NA	NA
	Others (if any)		

8. Financial Summary (Cumulative) The consolidated UC-SE are attached as Annexure I.

Attached the consolidated and audited Utilization Certificate (UC) and Year wise Statement of Expenditure (SE) as **Annexure I.**

9. Major Equipment/ Peripherals Procured under the Project** (if any)

S. No.	Name of Equipment's	Cost (INR)	Utilisation of the Equipment after project
1.	FieldSpec 4 SpectroRadiometer	Rs. 49,83,020.00	The Spectroradiometer, its accessories and UAV would not only be utilized for data capturing across Himalayan Regions but also would be utilized in building the resource pool for making India self-reliant as far as Hyperspectral Remote Sensing is considered.
2.	FieldSpec 4 SpectroRadiometer Accessories	Rs. 13,66,496.00	
3.	Lookout VTOL™- X-Mapper Series Industrial Unmanned Aerial System	Rs. 17,49,875.00	

Details are provided in **Annexures III &IV.

10. Quantification of Overall Project Progress

S. No.	Parameters	Total (Numeric)	Remarks/ Attachments/ Soft copies of documents
1.	IHR States Covered	2 (Uttarakhand and Jammu & Kashmir)	*Appendix 4*
2.	Project Site/ Field Stations Developed	*Pindari Valley and Dachigam National Park*	
3.	New Methods/ Modeling Developed	*12*	*Appendix 1 and 5*
4.	No. of Trainings arranged	*5*	*Appendix 2*
5.	No of beneficiaries attended trainings	*166*	*Appendix 2*
6.	Scientific Manpower Developed (PhD/M.Sc./JRF/SRF/ RA):	*3 PhD* *4 JRF* *2 SRF* *4 Research*	

		Associate *8 Postgraduates*	
7.	SC stakeholders benefited		
8.	ST stakeholders benefited		
9.	Women Empowered	*94*	*Appendix 2*
10.	No of Workshops Arranged along with level of participation	*5*	*Appendix 2*
11.	On field Demonstration Models initiated	*NA*	
12.	Livelihood Options promoted	*NA*	
13.	Technical/ Training Manuals prepared	*NA*	
14.	Processing Units established	*NA*	
15.	No of Species Collected	469	
16.	New Species identified	--	
17.	New Database generated (Types):	*2 (tool & data product)*	*Appendix 3*
	Others (if any)		

11. Knowledge Products and Publications:

S. No.	Publication/ Knowledge Products	*Number*		*Total Impact Factor*	*Remarks/* Enclosures
		National	International		
1.	Journal Research Articles/ Special Issue:		9	37.764	*Appendix 1*
2.	Book Chapter(s)/ Books:		2 Books published, 1 book will be appeared in 2023. 3 Book Chapters, 7 conference presentations.	NA	Pandey, P.C., Srivastava, P.K., Balzter, H., Bhattacharya, B., Petropoulos, G.P. (Eds.), Hyperspectral Remote

S. No.	Publication/ Knowledge Products	*Number*		*Total Impact Factor*	*Remarks/* Enclosures
		National	International		
					Sensing: Theory and Applications, ISBN 978-0-08-102894-0. Elsevier George Petropoulos and Srivastava, P.K. (2021) GPS and GNSS technology in Geosciences, Elsevier, United States. ISBN: 9780128186176
3.	Technical Reports	1			In Process
4.	Training Manual (Skill Development/ Capacity Building)	1			In Process
5.	Papers presented in Conferences/Seminars		7	NA	*Appendix 1*

* Please append the list of KPs/ publications (with impact factor and further details) with due Acknowledgement to NMHS.

12. Recommendation on Utility of Project Findings, Replicability and Exit Strategy

Particulars	Recommendations

Utility of the Project Findings	The deliverables of the current project would act as a baseline for the identification and retrieval of Himalayan Medicinal, Economical, Rare, Invasive Plant Species. The SHRIME tool developed as a part of this project would be the first of its kind initiative that will boost the research in the Himalayan Ecosystem. The models developed could be scaled across the Himalayan region for offering sustainable solutions for Himalayan Ecosystems under climatic uncertainties.
Replicability of Project	The dataset would be utilized for researchers across the domains on request. The models published as part of various objectives are developed under open-source platforms and thus easily replicated by the researchers.
Exit Strategy	The Indian Himalayan Region (IHR) not only is complex in terms of geomorphology but also terms of reachability. This work is a genuine effort towards bridging the gaps and offering solutions in terms of sustainable Himalayan Ecosystems under climatic uncertainties. The offerings in terms of benchmarked SHRIME tool where the spectral library of various Himalayan Medicinal, Economical, Rare, Invasive Plant Species would be a giant leap to the researchers working in this domain. There is a great scope in enriching the potential of this tool in terms of informed decision-making to various challenges across IHR. Various models developed as part of deliverables would help offer customized and localized solutions to the natives as well as build a natural ecosystem amidst the growing concern of Climate Change.

(PROJECT PROPONENT/ COORDINATOR

(Signed and Stamped)

Dr. Prashant K. Srivastava
Principal Investigator
(P-07/683)
IESD, BHU
Varanasi-221005

Place: BHU, Varanasi

Date: 29/11/2022

!

!

PART B: PROJECT DETAILED REPORT

1 EXECUTIVE SUMMARY

The Indian Himalayan Region (IHR) is one of the world's unique mountain systems. The variations in topography, climate, and a vast altitude range of <100 m to >7000 m asl provide habitats for different life forms to grow and flourish. Information generation on diversity, distribution and extended availability of the high value (ecologically, medicinally, or economically) plant species is of prime importance. The Himalayas, which represent one of the global biodiversity hotspots, is rich in the diversity of such high-value plant species. They play an integral role in the structure and functioning of the forest ecosystem. Plant diversity is an integral element of the global ecosystems and provides stability to the ecosystem and maintains the ecological balance (Brugière & Scholte, 2013; Rands et al., 2010). It also provides ecosystem services and goods for the well-being and long-term survival of mankind. However, the continued growth of human populations and per capita consumption has resulted in unsustainable exploitation of biological diversity, exacerbated by climate change (IPCC, 2007; UNFCC, 2015) and other anthropogenic environmental impacts (Rands et al., 2010). Overexploitation of species, expansion of invasive alien species, climate change, forest degradation and destruction of unique habitats are the major factors of declining biodiversity (Butchart et al., 2010). During the last century, the planet has lost 50% of its wetlands, 40% of its forests and around 60% of global ecosystem services have been degraded in just 50 years (Patrick ten Brink, 2011). According to (Mora et al., 2011) second half of the last century has suffered from a high rate of plant extinction; a loss of 137 species per day. This pattern of species extinction is considered 1000–10,000 times faster which could be occurred naturally (Hilton-Taylor et al., 2000)). Because of the above, quantification of biophysical [Leaf Area Index (LAI)] and biochemical (Chlorophyll and carotenoid content) properties is very crucial for monitoring the high-value plants in the Himalayan region. Due to terrain complexity, it becomes very difficult for field-based sampling methods in mountain areas. Some places are very difficult to approach, and some are even unfeasible. In this context, Remote Sensing (RS) and GIS (Geographical Information System) tools and techniques play an important role in vegetation sampling and monitoring. Therefore, their detection and quantification of biophysical and biochemical factors are prime elements in vegetation monitoring, sustainable utilization, and resource conservation. Hyperspectral acquisition provides a spectral response in narrow and continuous spectral bands, with significant improvements when compared with broad bands in terms of spectral resolution. The spectral profile obtained through the radiometer can be used for the creation of a spectral digital library and subsequently, it can be used for the detection and monitoring of medicinal, rare, endangered, and economically important plant species. Hyperspectral remote sensing data can be further used for the estimation of phenolic contents as well as biophysical and biochemical properties in the plants through forward and inverse modelling approaches. Because of changes in the temperature regime, there is a shifting in the above-mentioned plants towards a higher altitude. Some species that are not adapted or acclimatized are on the verge of extinction. Identification and monitoring of these species and their biophysical and biochemical parameters using hyperspectral remote sensing can provide the conservation of vulnerable species. Further, it will help detect the economically important species and their distribution in a short time over the complex Himalayan terrain and can also predict their extinction or survival chances through the projected climate data over the Himalayas.

2 INTRODUCTION

2.1 Background of the Project

The project was proposed for the Quantification of biophysical and biochemical parameters which is very crucial for monitoring the high-value plants in the Himalayan region, their sustainable utilization as well as the conservation of related resources. In all, there were six objectives in the project. The first one was the creation of the Spectral library and standardization for selected medicinal and threatened species. The second one was Spectral and image analysis for discrimination of selected economically important plant species followed by Model building for retrieval of Biochemical and Biophysical properties of selected species. Projection and scenario for distribution of selected species was the fourth objective followed by the fifth objective of fine scaling of space-time map of the high-value species in the Himalayas. The final objective was the management and conservation of species based on previous objectives. To fulfil these objectives, extensive field samplings were conducted in Pindari, Nandadevi biosphere reserve, Jageshwer, Almora, Ranikhet, Gwaldham, and Munsiyari regions in Uttarakhand and, Dachigam National National Park, Gulmarg Wildlife sanctuary in Kashmir. The exhaustive species list in the Himalayan range is given in Table 1.

Table 1 Available species in the complete Himalayan Range

Name	**Height**	**Type**	**Location**	**Remarks (** • **Altitude** • **Blooming Season** • **Medicinal Value)**
Nepata laevigata	1-3 feet	Herb	Pangi, Himachal	2500 mts

Nepata leucophylla benth	2-3 feet	Herb	Kumaon, Himalayan, Mussoorie, dhanolti	
Nepata longibracteata benth	2-3 feet	Herb	Kazork spanganak, lang Sea-Kangri, thukje, pang, debring (hemis National Park), stakna	
Nyctanthes arbor trisitis L. (Parijat)	2-3 feet	Small tree and deciduous shrub	South India, Mumbai, Delhi, near Allahabad	
Ocimum basilicum L. (basil)	2-3 feet	Herb	All over India	
Organium vulgare L. (Himalayan Marjorum/ oregano)	0.6 ft- 2.6ft	Perennial Herb	Badhangari, Chamoli Aeradev, Almora Kamedi Devi, Bageshwar Patal Bhuvneshwar, Pithoragarh Liti, Bageshwar Shama, Bageshwar Badhangari, Chamoli Dronagiri, Almora Purara, Bageshwar	Alt. 1500 – 2800 m

			Shama, Bageshwar	
Vitex negundo L.	10-20 ft tall	Shrubs or small trees	Bulekha Uttaranchal, Dehradun	
Zanthoxylum armatum DC.		Small Plants	Pithoragarh, Lohaghat, dana	In evergreen forests, between 90-1000 m.
Thuja Orientalis L. / Platycladus Orientalis	More than 20 m/ 10 to 200 feet		Shimla	15 m (49ft) by 5 m (16ft)
Thymus linearis benth.		Shrub	Hemis national park, pang, sangtha, burwa, Nanda Devi National Park, Martuli Pancholi, Badrinath, jaspa, Zanskar, Balta (away from 20 km from Dachigam national Park)	
Thymus Serphyllum L.	40 cm	Perennial Shrub	Dachigam National Park, Chandanwadi (near Dachigam), Mussoorie (UK)	
Valeriana hardwickii Wall	40- 50 cm	Perennial Herb	Dachigam national park, palchan, Marh (Himachal), Nanda Devi	1500-4000 m

			national park and Dehradun	
• ***Valeriana jatamansi* Jones**	15-65 cm tall	Perennial Herb	Burwa (near Manali), Mussoorie, Jawarna (Himachal), forest Range near Mussoorie	
• **Rhododendron anthopogon**	2-3 ft	Alpine Shrub	Chitkul (3600 m) in Kinnaur district, Great Himalayan National Park (3500 m) in Kullu district, Tungnath, Uttarakhand at an altitude of approx. 3600 m	
• **Senecio Rufinervis DC**	3 ft tall	Woody Herb	In the Himalayas in Uttaranchal and Nepal.	1800-3000 m
• **Skimmia anquetilia Tayl.**	4 ft	Aromatic erect or creeping shrub	Harinagar, Nartola (District Nainital)	
• **Skimmia lauerola (DC)**	4 ft		Kedarnath, Agasthamani, Ransi Mountain, Shimla	4000 m

• ***Selinum nuifolium* Salisb.**	3–6 feet (1–1.83 m).	Herb with biennial tubers	Govind pashu vihar national park, Dachigam national park	
• ***Senecio Nudicaulis buch***	30-70 Cm	Perennial Herb	Jim Corbett, tanda Deshral, Zanskar (Hemis National Park)	
• **Solnum Xanthcarpum**	50-70 Cm	Herb	occurs throughout India, in dry situations as a weed along the roadsides and wastelands	
Satcys sericea wall.ex benth		-	-	-
Tanacetum longifolium wall.	erect hairy herbaceous plant		Kedarnath wildlife sanctuary,	
Tanacetum gracile Hook.	30–60 cm tall	Small plant	Near Hemis National Park, Ladakh Himalaya, Ganglas, Ladakh on the way to Nubra valley (altitude 3500 m a.s.l.)	July to September
Tanacetum nubigenum Wall	Stems of 18-33 Cm	Perennial Herb	Malari in the Chamoli district of Uttar Pradesh,	September-October

Perovskia abrontanoides Kar.	100 Cm	Perennial subshrub	Shey, phey (mountains) near Hemis National Park, Dargoo, Skurbhuchan, Dah, Damkhar	Dry stream beds on mountain
Persea duthiei (king) Kosterm/ Machilus duthiei	25 m	Large trees	Barmdeo (Bhabar), Kumaon	1500–3200 m.
Persea gamblei(king Ex Hook f.) Kosterm/ Machilus gambeli	up to 25 m tall;	Trees	Jaunsar-Bawar is a hilly region, 85 km from Mussoorie, Kalatope, Chamba, Dehradun, Rishikesh	2100 mts
Phoebe lanceolate (Nees) Nees (lauraceae)	8 m tall.	Trees	Dehradun, Lacheen, Siliguri, Nagaland	
Pinus roxburghii Sarg.	4-15(-20) m tall	Trees	Dehradun, Lohaghat, Pithoragarh, Almohra, Kausani, Dwarahat, Gwaldam, Salyana, Uttarkashi, Mussorie	
Piper betle L. (Piperaceae)		perennial creeper	Bihar, Bengal, Orissa,	

			Tamilnadu and Karnataka.	
Pleurospernum angelicoides (Wall ex. DC) Benth ex. C.B. Clarke	10-30 ft tall	Tree	Niti, Chamoli (Uttarakhand), North-Western Himalaya, India	
Cedrus deodara / Deodar	50 m high	Tree	Kashmir to Garhwal at altitude ranges from 1210 to 3050 m	
Abies himalayensis /Pinus Pindrow /Pinus Spectabilis /var Pindrow	50 m Tall	tree		
Justicia Adhoatoda Linn	1.2-6 m tall	perennial shrub	Dehradun, Haldwani,	
Agave Americana	25–30 ft (8–9 m) tall.			
Centella Asiatica Linn	perennial, creeper herb, attains height up to 15cm (6inches)	medicinal herb	Eastern Himalaya	Temperate and tropical swampy areas in many regions of the world. Threatened at IUCN
Mangifera indica	Large tree to 150 ft (45 m) growing wild in the eastern Himalayas,	tree	Jim Corbett, Rudraprayag, zirkapur (Chandigarh)	Elevation: 980 to 4,200 ft (300-1,300 m).

	Nepal, and the Andaman Islands.			
Holarrhena antidysenterica	height up to 13 m	A small deciduous tree. Leaves:		
Carissa spinarum	upto 4m.	spiny shrub	Tanda desharal, Gurugram	
Acorus calamus Linn.		semiaquatic, perennial, aromatic herb with creeping rhizomes		
Arnebia benthamii Wallich	Height of 60 Cm	Perennial Plant		
Arisaema tortuosum	Max 2m	Scrub and Alpine Meadows in the Himalaya. Perennial	Pahalgam Valley, Kashmir Himalayas	
Calotropis procera/ Madar	1-4 m tall.	shrub		
• **Artemisia annua**	30 to 100 cm tall	annual herb		
Artimisia Nilagirica	1-4 m height	Aromatic Shrub		
Arster flaccidus	3-30 cm	Perennial	Hemis national	

	tall	Herb	park, balthi, Kelong, Kishwar throughout Kashmir region	
Centipedia minima	1,200 metres	Prostrate glabrous herbs		
Cichorium intybus	1-4 feet tall	Bushy Perennial Herb	Munsayri (Nanda Devi national park), garbyang	
Erigeron asteroids roxb.	15-25(45) m tall 40-60(90) cm	evergreen tree	Shimla, Dehradun, Bharmour (near dhauladar range, Dhamashala)	IUCN endangered W. Himalaya at 1800-300 m on limestone substrates
Rhododendron arboretum	40 to 50 ft tree up to 14 m in height & 2.4 m in girth	evergreen shrub or small tree with a showy display of bright red flowers	Dehradun, Shimla, Hemis National Park, Nainital, Mukteshwar Near Almohra, Mussorie	elevations of 4500 to 10,500 ft
Emblica Officinalis				
Bahuinia variegata	20-40 ft tall	Orchid Tree	Haridwar, Dehradun	
Quercus leucotrichophora Cam.	25 m	Evergreen Tree	Garhwal Himalaya-Chandrabadani , Chaurangikhal,	

			Ghuttu, Khirsoo and Mussoorie	
Aesculus indica.	20-30 m tall	Chestnut is a tall deciduous spreading shady tree.	Dehradun, Haldwani, Ranikhet, Nanda Devi National Park, Nanital, Dungari, ghes, Dhoru Shadab (Srinagar)	3,000 m above sea level
Albizia lebbek	25 m high	Deciduous trees	Dehradun, haldawani, Kapurthala, Nakudar	
Ficus carica	15-30 ft	Small Tree	Kalpa,	
Ficus religiosa	25 m high	Deciduous trees	Haridwar Rishikesh, Nanital, Dehradun, Bageshwar, Rudrayag, Haldwani	
Myrica esculanta Buch Ham.	6 - 8 m in height	small tree	Haldwani, Shimla, patal Bhuvneshwar, Pithoragarh, Chamoli, Gwaldham, Shimla	
Rheum australe	1ft long	Branched Clusters	Nanada Devi national Park, Zanskar zumdo	

Prunus persica	up to 15 m tall	Trees,	Uttarkashi, Dehradun	
Murraya koenigii	3-5 m tall	deciduous aromatic shrub	Chandigarh, Haridwar, Dehradun, Tanda Deshral	
Sapindus mukrossi	25 m	Tall Tree	Dehradun, Haldawani, Pittoragarh	
Taxus baccata Linn.	30 m tall	under-story tree	Found in the temperate Himalayas at the elevation of 6000-11000 ft.	Flowering occurs from March to May and seeds ripen between August and November of the same year. Vulnerable
Cinnamomum glanduliferum (Wall.) Meisn. (Lauraceae)		Leaves false camphor tree or Nepal camphor tree		• 1,500–2,500 m • March to May • Leaves are used as a stimulant, carminative, and to treat coughs and colds
Cinnamomum Tamala (Buch. -Ham.) T. Nees and Nees (Lauraceae)		Leaves, root		• 1,000–2,000 m • Jan – March • tejpatta leaves to treat gastric problems
Juglans regia L. (Juglandaceae)		Bark, leaves		• June-July • walnut Leaf essential oil from Kashmir,

				antibacterial
Juniperus indica Bertol. (Cupressaceae)		Leaves, berries		• 2600-5200 m • leaves and berries are used to treat fevers, coughs, skin diseases
Murraya koenigii (L.) Spreng. (Rutaceae)		Leaves, bark, twigs		• 1500 m • April-may • sweet neem, Kadi Patta anthelmintic and in blood disorders
Eucalyptus globules Labill. Lyptus (Myrtaceae)		Leaf, bark and, oil		Bluegum
Melia azedarach L. Daikan (Meliaceae)		Leaves barks and fruits		• In the spring • Chinaberry tree Bark and fruit extract are used to kill parasitic roundworms. In Manipur, leaves and flowers are used as a poultice in nervous headache. Leaves, bark, and fruit are insect repellent. Seed-oil is used in rheumatism. Wood-extract is

				used in asthma.
Pistacia khinjuk Socks. Kakarsingi (Anacardiaceae)		Leaves, petiole, and branches		Antimicrobial
Pterocarpus marsupium Roxb Bijesar (Fabaceae)		Gum leaf and flower	.	• 1,200 m September-October. • Beeja patta the wood and bark of the tree are known for anti-inflammation property and their anti-diabetic activity
Shorea robusta Roxb. ex. Sal (Dipterocarpaceae)		Leaves and bark		Sal
Taxus baccata auct. (Non-L.) Thuner (Taxaceae)		Leaves and bark		anti-cancer agents
Pistacia integerrima Stews		The whole plant (leaf, bark, root, and galls)		its galls have been utilized for the treatment of cough, asthma, dysentery, liver disorders and for snake bite.
Flacourtia indica		root, Leaves, Fruits		Kancu Skin diseases, Poisonous biting, Jaundice
Ginkgo biloba		leaves		treatment for blood

				disorders and memory problems, enhancement of cardiovascular function and to improve eye health

The Pindari Basin is very rich in terms of forest resources and diversity. From the valley region to the highly elevated alpine meadows, locally known as Kharakor Bugyal, a rich diversity of plants is found. In the mid slopes, Chir(pine) is common, while in the upper reaches, temperate coniferous forests, mainly Banj(oak), Tilonj (Quercesdilitata) and Devadar(Cedrusdeodara), are found extensively. The diversity of the forest in Pindari as per altitude is given in Table 2.

Table 2 Forest Diversity Based on Altitude in Pindari

Belt/Altitude	**Geographical Area**	**Main Species**
Valley regions/below 1000 m	Along the valley of the Pindar River	Eucalyptus, Dendrocalamus sp
Middle altitude 1000 m to 1600 m	The slope of the various streams such as Kaver Gadhera, Ming Gadhera, Pranmati and Atagarh	Pine dominate (blue pine and chir forest), Kaver Gadhera, Ming Gadhera, and Pranmati Gadhera
Temperate zone 1600 m to 2000 m	Watershed regions, Love-Kush tope, Kanpur Garhi, Khankhrakhet, Shubhtal-Chhaltal, Kurur-Kwarad and Sol-Dungri	Deodar forests (Cedrusdeodara), oak forests (Quercus species), fir (Abiespindrow) & spruce (Picea smithiana), ringal (bamboo)

		forests (Dendrocalamus spp.)
Alpine Meadows/between 2600 m to the Snow line	Bedni Bugyal, Roopkund and Shail Samunder	Dominated by herbs Ringal (bamboo)
3000-4000m	Alpine Meadows	Osmunda-tussock grass, tussock grass, tussock grass-forb, tussock grass-Sedge, Sedge- forb, Sedge, early successional Rumaxnepalensis

2.2 Overview of the Major Issues to be Addressed

A total of 400 plus species were sampled during the study and a spectral library of 250+ species was prepared.

The First objective resulted in classifying the spectral species through collected spectra for which PRISMA data were acquired. Classification techniques namely spectral angular mapper (SAM) were used to classify the species over the acquired Hyperspectral data.

The second objective resulted in the development of a spectral library of the identified species. For this SHRIME (Spectral Tool of Himalayan Rare, Invasive, Medicinal, and Economical Plant Species) was developed.

The third objective resulted in the retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species Taxus wallichiana. The development and validation of hyperspectral indices for anti-cancerous Taxol content estimation were conducted as part of this objective. Further, Taxol Content Variation was also explored in altitudinal variation in the Himalayan regions. The species distribution of Taxol content through empirical methods and radiative transfer models (physical methods) were conducted. Taxol Mapping was also explored through drone quasi-hyperspectral images. As part of this objective, a list of tentative compounds identified in Rhododendron arboreum using UPLC and HRMS was prepared. Sampling and Analysis of Viburnum grandiflorum and Parrotiosis jacquemontiana were also done for its biochemical and biophysical properties.

For the fourth objective, fine-grained species distributions were modelled using Sentinel-2, Sentinel-5P, Moderate Resolution Imaging Spectroradiometer (MODIS), ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) and Shuttle Radar Topography Mission (SRTM) along with other climatic variables over the sampling regions. Several machine learning algorithms were incorporated to establish the relation between physical and climatic variables to estimate the probability distribution of species. The phenology assessment was carried out for two decades to the effect of an elevational gradient, temperature, and precipitation on the start of the season (SOS) and end of the season (EOS) in major forest types of the Kumaon region of the western Himalaya using MODIS NDVI time series data (2001-2019). The study disclosed that due to winter warming and summer dryness, despite a warming trend in pre-season or springtime the onset of the vegetation growth cycle shows a delayed trend across the vegetation types. The potential distribution of Rhododendron arboreum, a medicinal plant species found within the foothills of the Himalayas was also evaluated using Conventional Species Distribution Models (SDM) namely BIOCLIM and Maxent, a Machine Learning variant as well as Convolutional Neural Network (CNN).

For the fifth objective, High-resolution weather information is generated for the study region for 2000-2016 and CMIP6 climate scenarios are organized. As part of this objective, Investigation into the possibility of applying functional variables retrieved from Moderate Resolution Imaging Spectroradiometer (MODIS) onboard sensor data to map the most realistic species distribution limits of two alpine treeline species, namely Betula utilis D.Don and Rhododendron campanulatum over the Himalayan biodiversity hotspot was conducted. Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. For each model, the congruence of predictions was assessed and made pairwise comparisons to assess the performance of the models. As part of this objective, bias-corrected, statistically downscale models drawn from the NASA, Earth Exchange Global Daily

Downscaled Projections - Coupled Model Intercomparison Project Phase 5 (NEX-GDDP-CMIP5) were examined.

As part of objective 6, Traditional knowledge among communities residing in alpine regions of Western Himalaya was conducted. Gunji, Kuti, Napalchyu, Navi and Kutti of Byans Valley Pithoragarh villages were taken for this study. Questionnaire-based surveys gathered information on demographic profiles, biodiversity, and traditional knowledge where 75 informants of different age groups from five villages were selected. The Perceptions were recorded for 31 high-value medicinal and economical plants (belonging to 24 families) in which Maximum high-value species (23) were identified as herbs, followed by shrubs (5) and trees (3).

2.3 Baseline Data and Project Scope

Though a lot has been studied in the IHR, one of the core aspects that needs special attention is the plant species that are present in the Himalayan region. The presence of Medicinal, Economical, Rare, and Invasive plant species for a very long time has been of prime interest to many. A thorough study in terms of their presence amidst climate change is most crucial than ever. The degradation of the ecosystems along with the loss of these species needs to be explored at the local level. To bridge this gap, sadly none of the existing data could act as a baseline and thus the current project fills this gap and this makes the scope of this project an important one. The baseline developed in this project would help the policymakers in building a cumulative solution in the context of water, biodiversity, energy, livelihoods and associated disaster risk reduction.

2.4 Project Objectives and Target Deliverables (as per the NMHS Sanction Order)

Table 3 Objective-wise Deliverables

S. No.	Objectives	Major achievements (in bullets points)

1.	Detect and identify the medicinal plants, rare, endangered, and other economically important plant species concerning pedological and climatic conditions using the handheld Hyperspectral Spectroradiometer, AVIRIS Next Generation (NG) data and field observations in the Himalayas.	• For classifying the spectral species through collected spectra and PRISMA data were acquired. • Classification techniques namely spectral angular mapper (SAM) were used to classify the species over the acquired Hyperspectral data. • Spectral and image analysis for discrimination of selected economically important plant species such as Berberis, Polygonum polystachyum, Rhododendron Barbatum, Rhododendron Arboreum, Taxus Wallichiana, Bamboo and Fern is over.
2.	Spectral library of the economically important plant species to use with the Hyperspectral satellite and airborne data for large-scale quantification.	• Field surveys were conducted over study areas to collect plant samples. • The collected samples (fresh twigs and flowers) of each species were utilized for the preparation of their spectral libraries where the spectra in form of reflectance were recorded and stored using a hyperspectral Spectroradiometer. • Development of a spectral library of the identified species. For this **SHRIME** (Spectral Tool of Himalayan Rare, Invasive, Medicinal, and Economical Plant Species) was developed. • This could be utilized to create a spectral library in terms of reflectance and first derivative for the Himalayan plant species.

3.	Development of forward and inverse models for the retrieval of biophysical and biochemical parameters from economically important plant species using the Hyperspectral data.	• Retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species *Taxus wallichiana*. • For the retrieval of biochemical parameters, various indices were generated and three different types of filters, Average Mean, Savitzky Golay, and Fast Fourier Transform (FFT) one at a time followed by feature selection on each denoised spectra were applied to smooth the spectra for better estimation of biochemical parameters. • The development and validation of hyperspectral indices for anti-cancerous Taxol content estimation were conducted as part of this objective. • Further, Taxol Content Variation was also explored in altitudinal variation in the Himalayan regions. • The species distribution of Taxol content through empirical methods and radiative transfer models (physical methods) were conducted. Taxol Mapping was also explored through drone quasi-hyperspectral images. • As part of this objective, a list of tentative compounds identified in *Rhododendron arboreum* using UPLC and HRMS was prepared. • The algorithm for retrieval of biochemical properties of Taxus Wallichiana, and Pinus are over and Rhododendron, Parrotiopsis jacquemontiana, and Viburnum grandiflorum are underway.

4.	To provide fine-scale space-time inventory of the high-value medicinal plants, rare, endangered, and other economically important plant species.	• Fine-grained species distributions were modelled using Sentinel-2, Sentinel-5P, Moderate Resolution Imaging Spectroradiometer (MODIS), ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) and Shuttle Radar Topography Mission (SRTM) along with other climatic variables over the sampling regions. • Several machine learning algorithms were incorporated to establish the relation between physical and climatic variables to estimate the probability distribution of species. • The phenology assessment was carried out for two decades to the effect of an elevational gradient, temperature, and precipitation on the start of the season (SOS) and end of the season (EOS) in major forest types of the Kumaon region of the western Himalaya using MODIS NDVI time series data (2001-2019). • The potential distribution of Rhododendron arboreum, a medicinal plant species found within the foothills of the Himalayas was also evaluated using Conventional Species Distribution Models (SDM) namely BIOCLIM and Maxent, a Machine Learning variant as well as Convolutional Neural Network (CNN).

5.	Projection of future climate change scenarios and its possible effect on the distribution of medicinal plants, rare, endangered and other economically important plant species.	• High-resolution weather information is generated for the study region for 2000-2016 and CMIP6 climate scenarios are organized. • As part of this objective, Investigation into the possibility of applying functional variables retrieved from Moderate Resolution Imaging Spectroradiometer (MODIS) onboard sensor data to map the most realistic species distribution limits of two alpine treeline species, namely Betula utilis D.Don and Rhododendron campanulatum over the Himalayan biodiversity hotspot was conducted. • Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. For each model, the congruence of predictions was assessed and made pairwise comparisons to assess the performance of the models. • As part of this objective, bias-corrected, statistically downscale models drawn from the NASA, Earth Exchange Global Daily Downscaled Projections - Coupled Model Intercomparison Project Phase 5 (NEX-GDDP-CMIP5) were examined.

6.	Development of knowledge-based planning of an appropriate management system to safeguard medicinal plants, rare, endangered and other economically important plant species from unplanned harvesting, Identification of new ecological niches in higher Himalayas under climate change to protect the species and improve the livelihood conditions of local and regional communities.	• Traditional knowledge among communities residing in alpine regions of Western Himalaya was conducted. Gunji, Kuti, Napalchyu, Navi and Kutti of Byans Valley Pithoragarh villages were taken for this study. • Questionnaire-based surveys gathered information on demographic profiles, biodiversity, and traditional knowledge where 75 informants of different age groups from five villages were selected. The Perceptions were recorded for 31 high-value medicinal and economical plants (belonging to 24 families) in which Maximum high-value species (23) were identified as herbs, followed by shrubs (5) and trees (3).

3 METHODOLOGIES, STRATEGY AND APPROACH

3.1 Methodologies used for the study

The potential distribution of Rhododendron arboreum, a medicinal plant species found within the foothills of the Himalayas was evaluated using Conventional Species Distribution Models (SDM) namely BIOCLIM and Maxent, a Machine Learning variant as well as Convolutional Neural Network (CNN). The flowchart of Linear models, namely Maxent and BIOCLIM and Convolutional Neural Network (CNN) architecture is given in Figures 1,2, and 3 respectively.

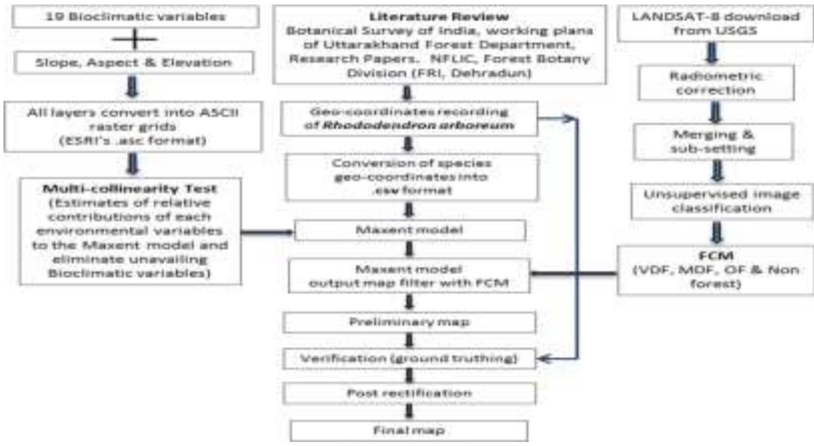


Figure 1 Maximum Entropy Model

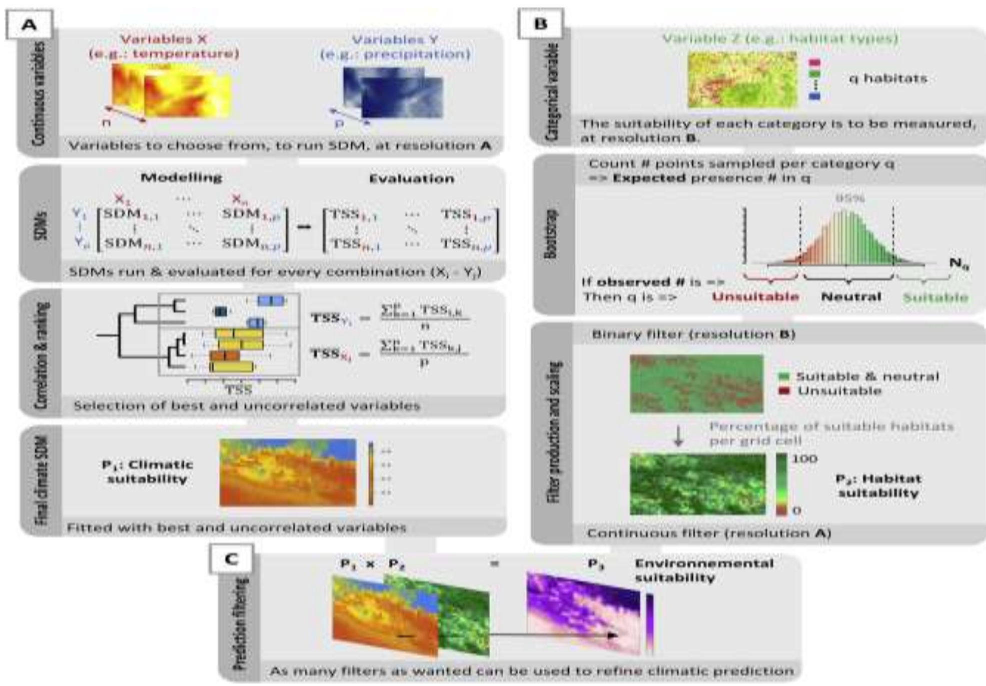


Figure 2 Bioclimatic Envelope Model

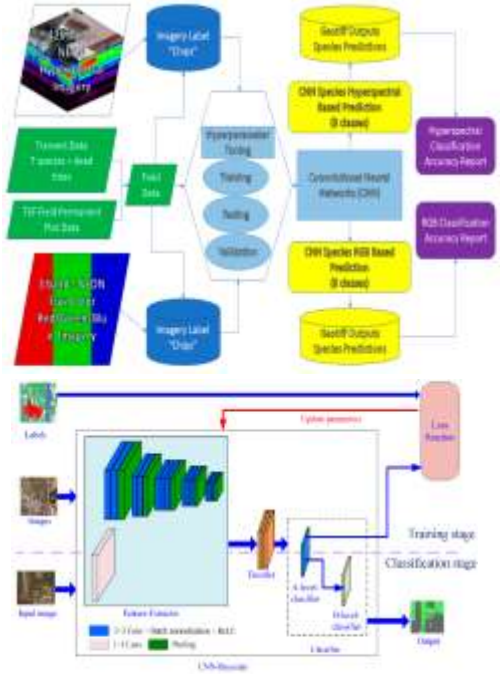


Figure 3 Convolutional Neural Network (CNN) architecture

Convolutional Neural Network (CNN) is a kind of deep learning-based model for processing multidimensional data that follows a grid pattern. The algorithm is developed in such a manner that the algorithm learns and adapts spatial hierarchies of features by itself from lower to higher levels of the pattern. Mathematically it is composed of three layers or building blocks: convolution, pooling and fully connected layers. Feature extraction is carried out using the first two layers and mapping of the extracted features to output is carried out by the third layer. The convolution layer has a crucial role in CNN that constitutes a heap of convolution which is a kind of linear operation mathematically. CNN is highly efficient in processing images which are rich in features at some or all places using small grid-like parameters known as the kernel. The kernel is a feature extractor that can be optimized and is applied at each of the feature points. The output of one-layer acts as an input to the next layer which directs the extracted features to grow hierarchically into a complex set of features. The procedure of optimizing kernels is known as training and it is done to minimize the difference between the output and ground labelled data using backpropagation and gradient descent algorithms. A convolution layer constitutes a combination of linear and nonlinear operations called convolution and activation respectively. Convolution is used for feature extraction in which a kernel is applied on an input tensor. A feature map is thus obtained through the product of kernel elements and tensor input. The procedure is then repeated on multiple kernels to obtain random feature maps which represent different feature extractors. The hyperparameters involved in convolution operations are the size and number of kernels. The size could be anything from 3x3 to 5x5 to 7x7 and the kernel could be chosen randomly. An activation operation is a nonlinear operation that takes the output of the convolution operation. The activation function used nowadays is rectified linear unit (ReLU). Previously used activation functions are sigmoid or hyperbolic tangent (tanh) functions. A pooling layer offers downsampling functionality that decreases the dimensionality of the feature maps to achieve translation invariance to alterations and biases incorporated and thus helps in reducing the no of learnable parameters. The details and complete work is given in (Anand et al., 2021).

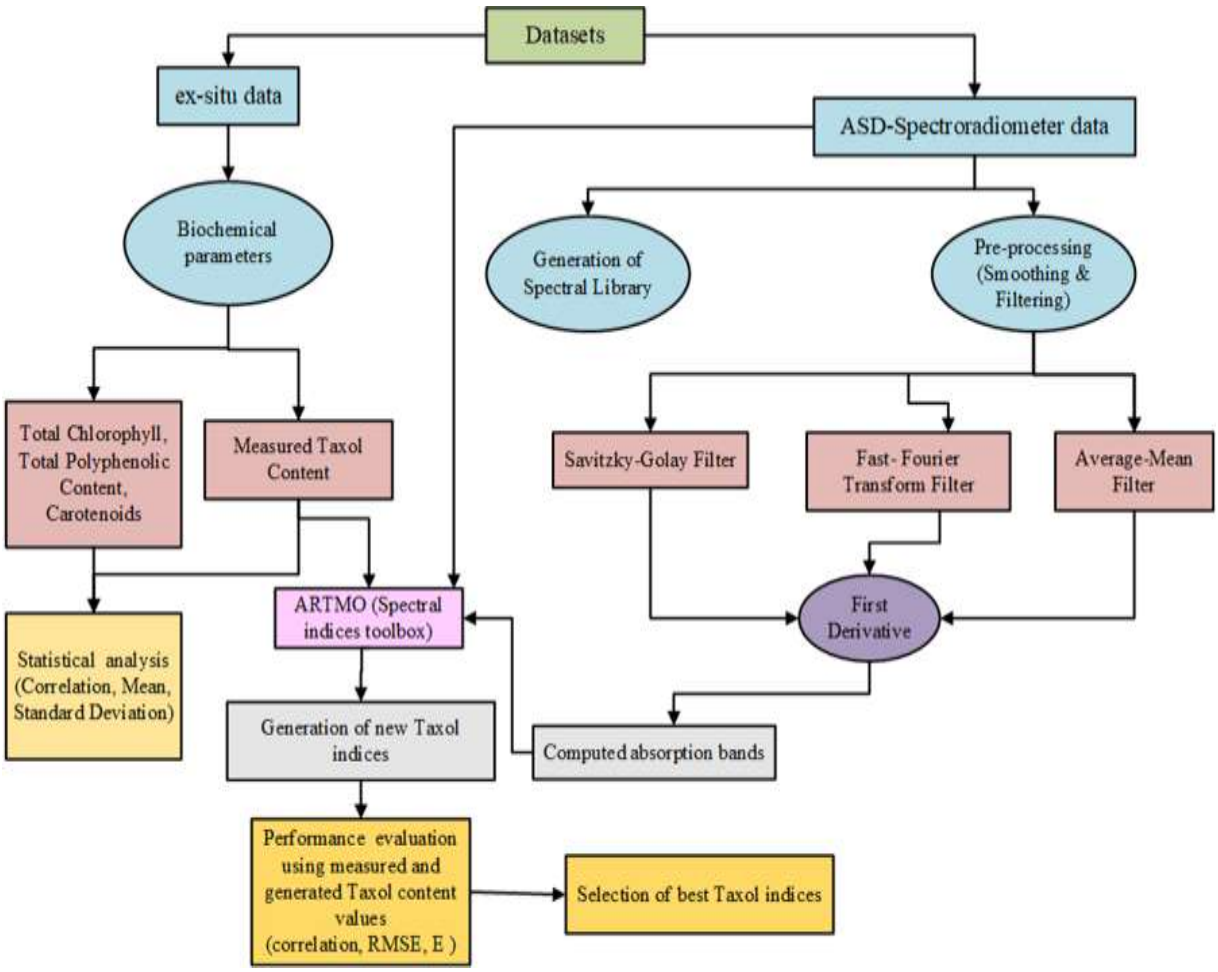


Figure 4 Development and Validation of hyperspectral indices for anti-cancerous Taxol content estimation in the Himalayan region

Retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species Taxus wallichiana is carried out as per Figure 4. The details of the work are given in (Gupta et al., 2022). Further, Taxol Content Variation was also explored in altitudinal variation in the Himalayan regions. The species distribution of Taxol content through empirical methods and radiative transfer models (physical methods) were conducted as given in (Gupta et al., 2021).

Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. For each model, the congruence of predictions was assessed and made pairwise comparisons to assess the performance of the models. The following data were utilized to develop NEOVs as specified in Table 4.

Table 4 Details of satellite data were used to develop novel EOVs

Sensor and Product	**Used variable**	**Ecological relevance**	**Original Resolution (m)**	**Rescaled at**	**Scaling factor**	**Resampling algorithm**	**Unit**	**Processing platform**	**Source**	**Citation**
MODIS Terra Daily EVI	EVI	Temporal EVI identifies the dynamics of the “greenness” of an area. It provides information about plant health and a proxy measure of species richness at a spatiotemporal scale	500	1000	0.0001	Nearest Neighbor		GEE ee.ImageCollection("MODIS/MOD09GA_EVI")	NASA LP DAAC at the USGS EROS Center	https://doi.org/10.5067/MODIS/MOD13A2.061

		(Chang et al., 2014).								
MOD16 A2.006: Terra Net Evapotranspiration 8-Day Global 500m	Total evapotranspiration (ET)	Evapotranspiration is an indirect measurement of plant available water that determines the presence of a plant community. For instance, focal species in this investigation are typically dependent on moisture availability (Mod et al., 2016).	500	1000	0.1	Nearest Neighbor	kg/m^2/8day	GEE ee.ImageCollection("MODIS/006/MOD16A2")	NASA LP DAAC at the USGS EROS Center	https://doi.org/10.5067/MODIS/MOD16A2.006 and https://modis-land.gsfc.nasa.gov/pdf/MOD16ATBD.pdf

MOD11 A1.006 Terra Land Surface Temperature and Emissivity Daily Global 1km	Daytime Land Surface Temperature (LST_Day_1km)	Temperature directly regulates plant growth and governs the seasonality of growth phases such as greening, browning, length of seasons, etc. Plants' physiological thresholds are governed by minimum and maximum temperatures in terms of heat and frost resistance. Focal species are	1000	NA	0.02	NA	Kelvin	GEE ee.ImageCollection("MODIS/061/MOD11A1")	NASA LP DAAC at the USGS EROS Center	https://doi.org/10.5067/MODIS/MOD11A1.006

		affected by direct growth limitation governed by low growing season temperatures in the tree and timberline ecotone. LST is used to quantify the surface energy and water nexus Mod et al. (2016); Bobrowski et al. (2018).								
MOD15A2H.006: Terra Leaf Area Index/FPAR 8-Day Global 500m	Leaf Area Index (LAI_500m)	Leaf area index (LAI) is an indicator of the seasonality of vegetation	500	1000	0.1	Nearest Neighbor	sq. meter/sq. meter	GEE ee.ImageCollection("MODIS/061/MOD15A2H")	NASA LPDAAC at the USGS EROS	https://doi.org/10.5067/MODIS/MOD15A2H.006

		cover, which may be used to quantify leaf phenology dynamics and is strongly tied to the carbon and water cycles. For the present study, dry season LAI is a crucial observation for focal species distribution. During this season, *B. utilis* holds very low LAI, while evergreen broad-leaved							Center	

		species like *R. campanulatum* hold a good amount of LAI (Naithani et al., 2013; Singh and Negi 2018).								
MOD17A2H.006: Terra Gross Primary Productivity 8-Day Global 500M 500m	Gross primary production (GPP)	GPP varies over functionally distinct vegetation groups such as evergreen and deciduous vegetation, and low-abrupt changes in GPP show low habitat heterogeneity. Our focal	500	1000	0.0001	Nearest Neighbor	kg*C/m^2	GEE ee.ImageCollection("MODIS/006/MOD17A2H")	NASA LP DAAC at the USGS EROS Center	https://doi.org/10.5067/MODIS/MOD17A2H.006

		species produce high productivity in the tree line ecotone. *B. utilis* and *R. campanulatum* exhibits conspicuous and inconspicuous seasonal variation in GPP respectively (Hu et al., 2018).								
DEM		Climatic conditions of Himalayan Forest ecosystems are governed by topographic factors. Slight changes in these				-	meters	-	Alaska Satellite Facility	

		factors (elevation, slope, aspect) cause large differences in microclimates to occur by altering temperature, irradiation, precipitation, humidity, the thickness of soil, etc., and this complex interaction plays a significant role in determining the species distributions of focal species (Bhattacharyya and Sanjappa, 2008;								

		Oke and Thompson, 2015).								
MODIS Terra Daily EVI	Phenometrics retrieved from EVI	Phenometrics represent the circannual biological events of vegetation (species). Phenometrics developed based on phenophases are mostly species-specific, and these phenophases are more prominent in deciduo	500	1000	0.0001	Nearest Neighbor		GEE and CropPhenology R Package ee.ImageCollection("MODIS/MOD09GA_EVI") and https://github.com/SofanitAraya/CropPhenology	NASA LP DAAC at the USGS EROS Center	https://doi.org/10.5067/MODIS/MOD13A2.061

		us and evergreen species. Therefore, these variables are significant in the detection of focal species (Araya et al., 2018).								

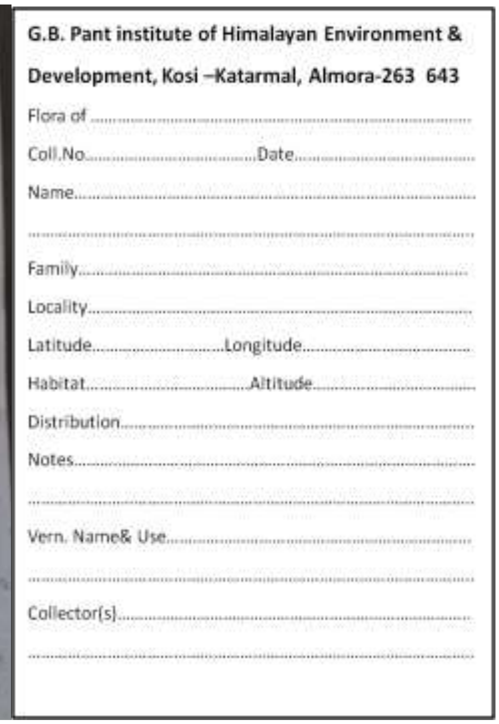
Table 5 List of finally accepted environmental variables for model building

B. utilis				***R. campanulatum***			
BCM	**PhenoM**	**BiophyM**	**HM**	**BCM**	**PhenoM**	**BiophyM**	**HM**
Cbio03	Offset Time	Mean ET of the warmest month	TIEVIBeforeMax	Cbio03	Offset Time	Min LST of Coldest Month	Offset Time
Cbio05	Max Value	Mean ET of warmest quarter	GreenUpSlope	Cbio06	Max Value	Max LST of Warmest Month	Max Value
Cbio07	Max Time	Mean EVI of warmest quarter	BrownDownSlope	Cbio07	Max Time	LST Seasonality (CV)	Max Time
Cbio08	TIEVIBeforeMax	Annual EVI range	BeforeMaxT	Cbio08	TIEVIBeforeMax	Isothermality LST	GreenUpSlope
Cbio15	GreenUp Slope	Mean LAI of the warmest month	Onset Value	Cbio15	GreenUp Slope	Annual Mean Diurnal Range LST	TIEVIAfterMax

Cbio 17	BrownDownSlope	Mean LAI of coldest quarter	Asymmetry	Cbio 17	TIEVIAfterMax	Mean LAI of warmest quarter	BrownDownSlope
Cbio 18	BeforeMaxT	Annual LAI range	Mean ET of the warmest month	Cbio 18	BrownDownSlope	Annual LAI range	BeforeMaxT
	Onset Value	Annual Mean Diurnal Range LST	Mean EVI of warmest quarter		BeforeMaxT	Mean EVI of warmest quarter	Asymmetry
	Asymmetry	Isothermality LST	Annual EVI range		Onset Value	Mean ET of coldest quarter	Onset Time
		LST Seasonality (CV)	Mean LAI of the warmest month		Asymmetry	Annual ET range	Annual Mean Diurnal Range LST
		Max LST of Warmest Month	Mean LAI of coldest quarter		Onset Time	Mean GPP of warmest quarter	Isothermality LST
		Mean GPP of warmest quarter	Annual LAI range			Mean GPP of coldest quarter	LST Seasonality (CV)
		Mean GPP of coldest quarter	Max LST of Warmest Month				Max LST of Warmest Month
			LST Seasonality (CV)				Min LST of Coldest Month
			Isothermality LST				Mean LAI of warmest quarter
			Annual Mean Diurnal Range LST				Annual LAI range
			Cbio 17				Mean ET of coldest quarter
			Cbio18				Annual ET range
			Cbio 15				Cbio 18
			Cbio 05				Cbio 17
			Cbio 08				Cbio 15
			Cbio 03				Cbio 03

3.2 Preparatory Actions and Agencies Involved

Field Sampling was conducted to collect the species. The samples were kept in a Ziploc bag for the next few hours. The samples collected were then crushed and stored in liquid nitrogen until the immediate analysis. Hence the values reported after analysis are represented in terms of fresh weight (FW). The full-range spectroradiometer (350-2500 nm) was used to capture the spectral signature of the leaves. The species were processed for Herbarium documentation. The blotter paper dried specimens were brought to the laboratory for the preparation of herbarium sheets. After poisoning with Mercuric Chloride ($HgCl_2$) the specimens were mounted on herbarium sheets with the help of glue and then stitched for more safety (Figure 5). Then finally the labels were pasted on each herbarium sheet (Figure 6). 2 specimens of every individual species were prepared. One set of specimens of each species will be submitted to the herbarium of Botanical Survey of India, Northern Regional Circle, Dehradun (BSD).



G.B. Pant institute of Himalayan Environment & Development, Kosi –Katarmal, Almora-263 643

Flora of ....

Coll.No.......... Date..........

Name..........

....

Family..........

Locality..........

Latitude.......... Longitude..........

Habitat.......... Altitude..........

Distribution..........

Notes..........

....

Vern. Name& Use..........

....

Collector(s)..........

....

Figure 5 Complete herbarium sheet *Figure 6 Standard herbarium label*

The following Teams were involved in the collection of data. The pictures of the sampling team are given in Figures 7 and 8.

- Dr Prashant Kumar Srivastava, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh

- K. Chandra Sekhar, Scientist F, G.B. Pant National Institute of Himalayan Environment & Sustainable Development, Almora, Uttarakhand
- Dr Irfan Rashid, Assistant Professor, Department of Earth Sciences, University of Kashmir

Figure 7 Field Sampling in Pindari

t

Figure 8 Field Sampling in Dachigam National Park

3.3 Details of Scientific data collected and Equipments Used

The properties of the Fieldspec4 spectroradiometer is given in Figure 9 and 10.

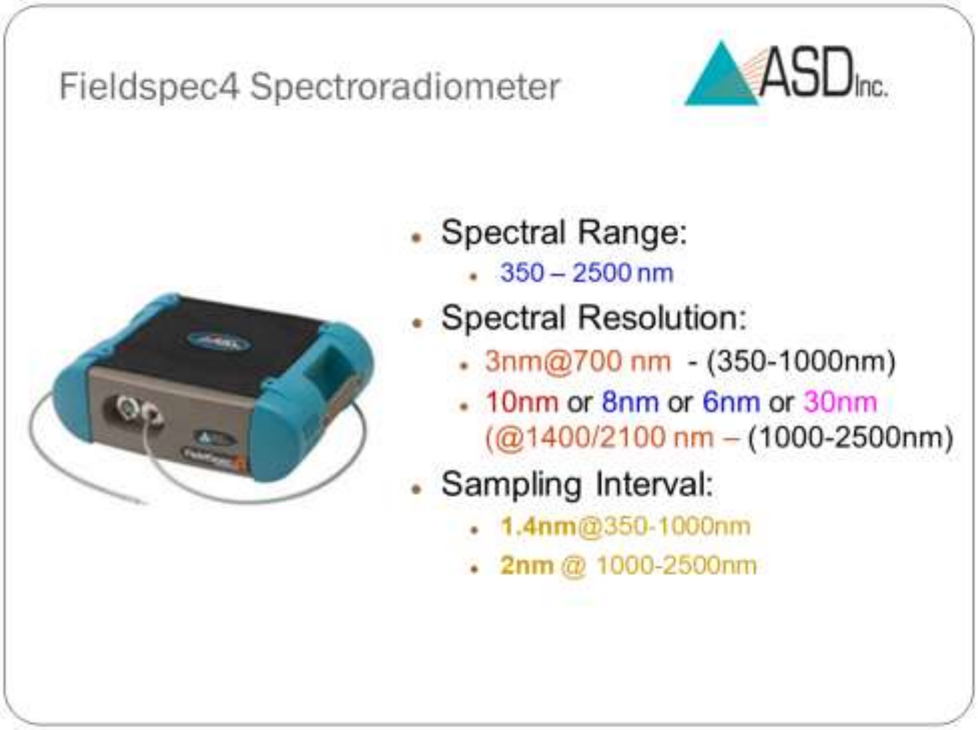


Figure 9 FieldSpec4 Spectroradiometer Configuration

Figure 10 FieldSpec4 Spectroradiometer Configuration Continued.

The detailed characteristics of the FieldSpec4 Spectroradiometer as per (Pandey et al., 2021) are given in Table 6.

Table 6 Characteristics of the Spectroradiometer Used

Characteristics	**ASD FieldSpec**
Wavelength	(350 - 2500 nm)
Spectral Resolution	3nm@700 nm - (350-1000nm) 10nm or 8nm or 6nm or 30nm (@1400/2100 nm – (1000-2500nm)
Diodes	Triple diode scanning design
FOV	Permanent Fiber optic cable with Wide 25˚ Field of View and Pistol grip for easy operation. Easy mountable limiting fore optic lenses for the smaller field of view.
Sampling Interval	1.4nm@350-1000nm 2nm @ 1000-2500nm
Time to Collect Spectrum	Collects the complete spectrum (350 – 2500nm) in 0.1 second.
Configuration	No need to configure the diodes as the entire unit is calibrated to function, not as three individual units. Here optimize features, and automatically adjusts integration time and scaling properties for each diode according to the light conditions to optimize the amount of radiation received, which results in a decreased signal-to-noise ratio increased accuracy and a seamless spectrum. Small glitches sometimes are observed at the stitching wavelengths.
Integration Time	Pre-Configured.
Dark Current	It is equipped with an internal shutter that closes off the light sources when taking dark scenes. This system, called drift lack is helpful in taking accurate dark scenes to eliminate electronic signalling.
Sequential Measurement	It has the feature to take the sequential measurements. No need for manual intervention. Saving spectra by spectrum numbers, no files to saves and the interval

	between saves can be specified by the user.
Data Management	Spectra stored is recognizable by FieldSpec ViewSpec Pro. This data needs to be converted into ASCII format using ViewSpec. It doesn't have the ability to open previously saved spectra files and needs the ViewSpec program to do this. It has a primitive directory structure which is time-consuming to navigate.
Attachment and Hardware	Excellent hardware design. Heavy but easy to use and self-contained.
Cost	It costs approximately ten times that of the Apogee/StellarNet model.
Software	The software (FieldSpec Pro RS3) is poorly designed. FieldSpec4 with wireless communication to the controller laptop provides flexibility to work on any platform
Resolution	
Operating Temperature	
Summary	To summarize, it has an excellent hardware design, three diode scanning design far exceeds than single diode APOGEE/Stellar/ Net Spec PAR/NRI. It's a pre-calibrated and ready to use in its own backpack. All in one design eliminates the clutter and Hustle of multiple units, external battery pack, and excess firing. Water stress could be measured which is indicated by several majority water bands above 1100 nm.

3.4 Primary Data Collected

Pindari glacier is found in the upper reaches of Kumaon Himalaya and situated at an elevation of 5200 m. to the southeast of Nanda Devi, Nanda Kot (28° 43' 55" to 30° 30' 12" N and 78° 44' 30" to 80° 45' E), between altitudes 1500-3000 m. Dachigam National Park is located in the Zabarwan Range of the western Himalayas (34.1372° N, 75.0378° E). The variation in altitude is vast, ranging from 5500 ft to 14000 ft above mean sea level. Due to this vast variation, the park is demarcated into an uneven region. With the help of available literature (Dwivedi et al., 2019; Samant et al., 2007) pertaining to the economic and medicinal value of plants, the high-value species were targeted in the study area. Munsiyari is one of the north-eastern districts of Kumaon Himalaya and is a part of Kailash Sacred Landscape (KSL) between 29° 50' 50'' to 30° 49' 41'' N latitudes and 79° 56' 43'' to 80° 29' 23'' E longitudes, Panchachuli 2 (6,904 m) is the highest peak in this area. The average temperature is 16°C and the average annual precipitation is 1500 mm. However, temperature and precipitation patterns extremely vary depending on season and elevation. Each climatic zone supports a variety of forests and supports a wide variety of native and endemic biodiversity habitats. The location maps of the study area are represented in Figures 11,12,13 and 14.

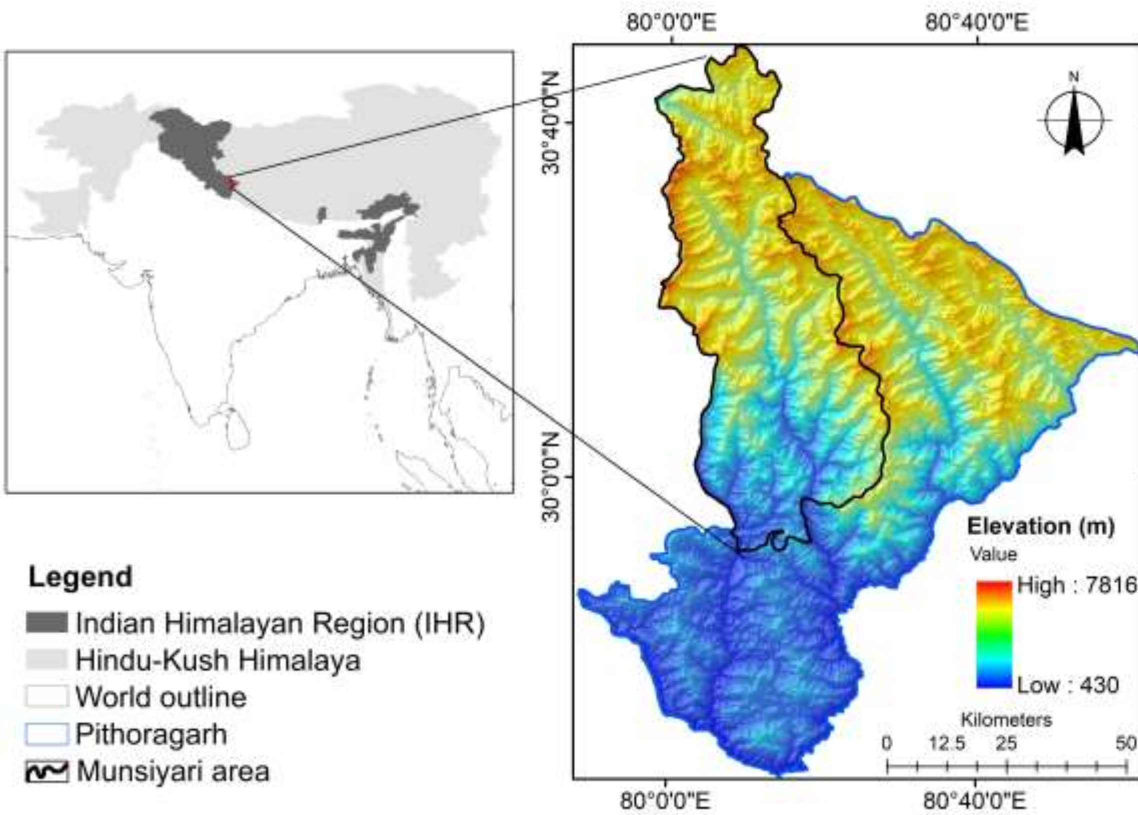


Figure 11 Location map of the study area: Munsiyari

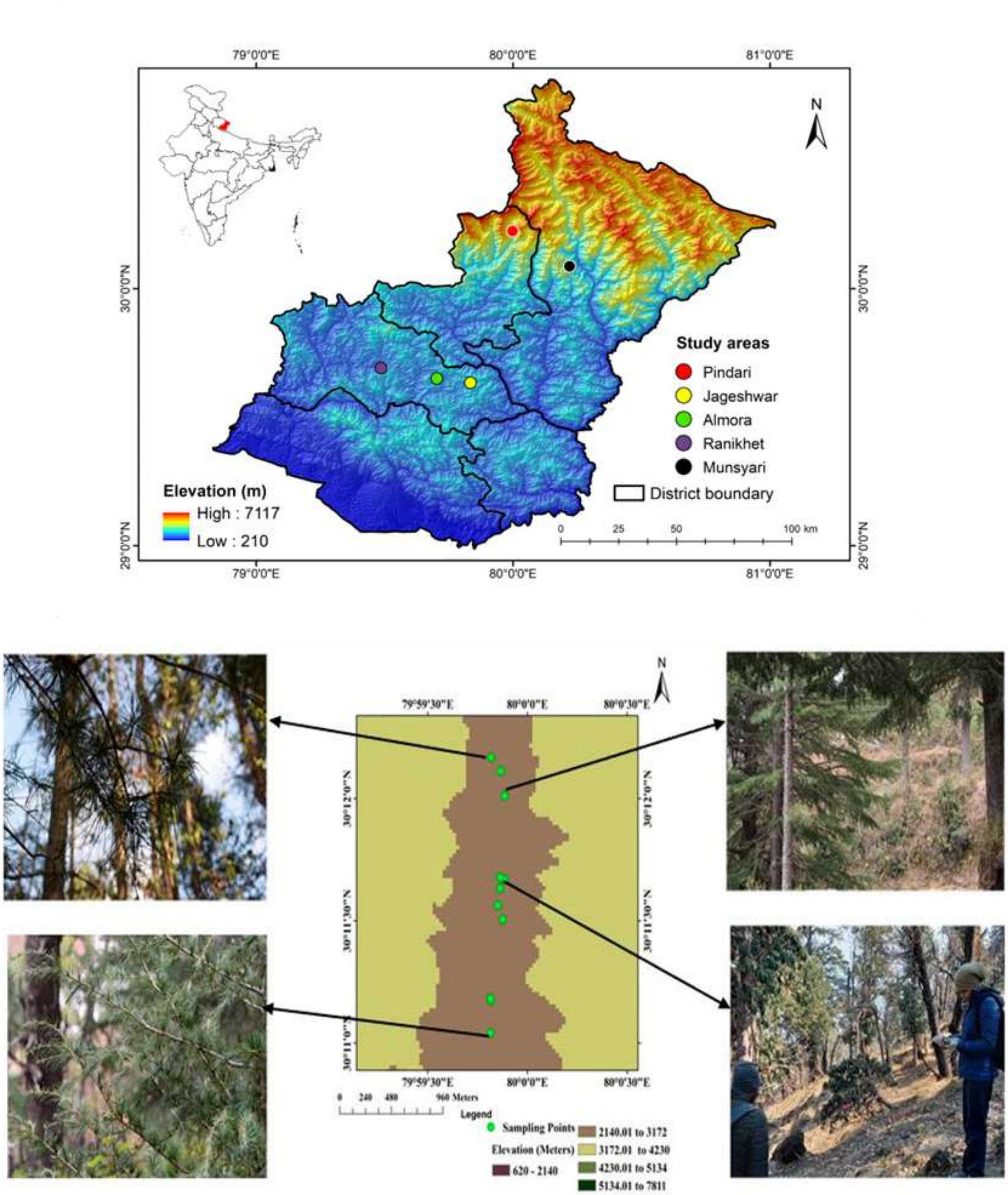


Figure 12 Location map of study area Pindari. Jageshwer, ALmora, Ranikhet

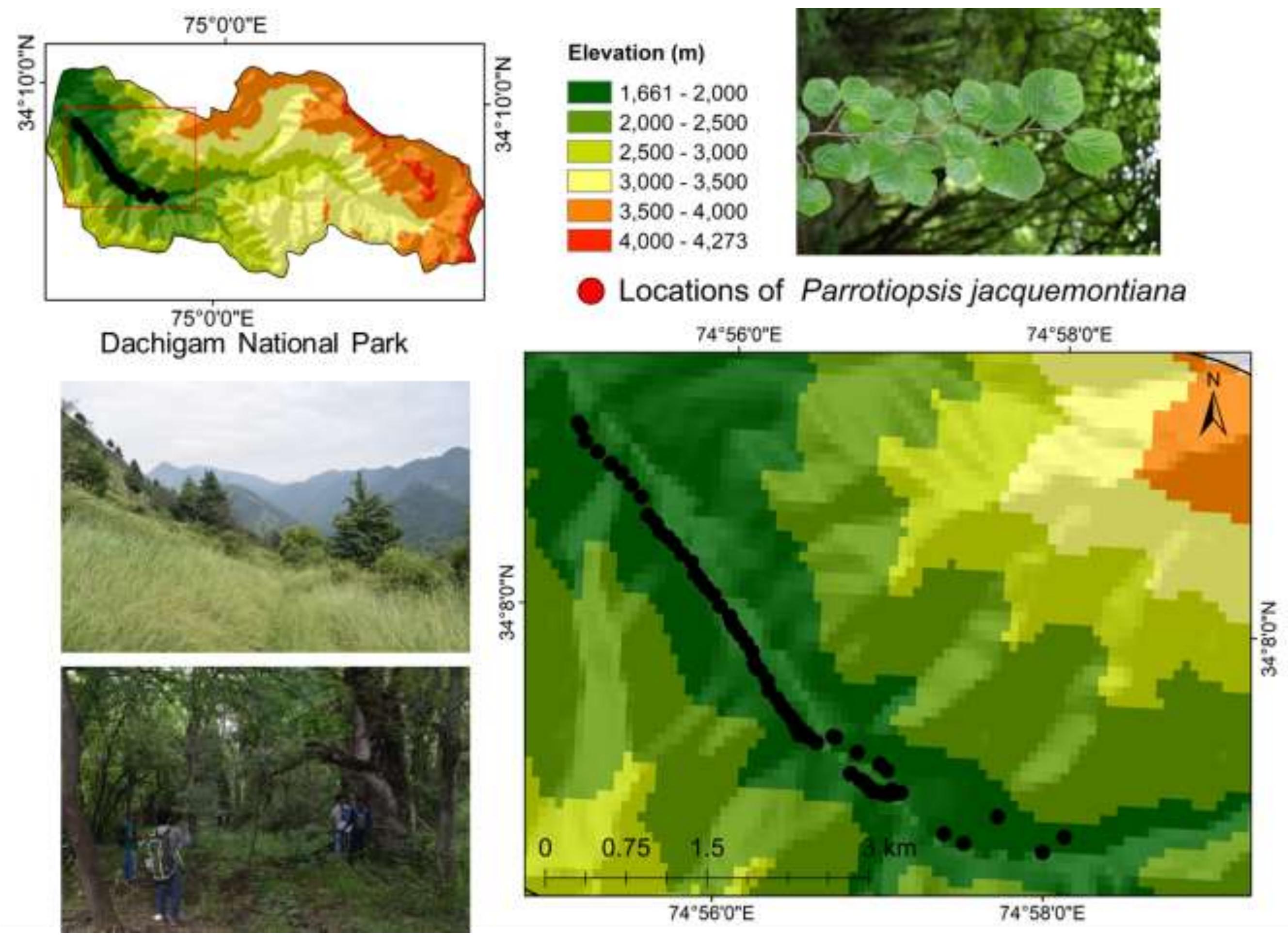


Figure 13 Location map of study area Dachigam National Park

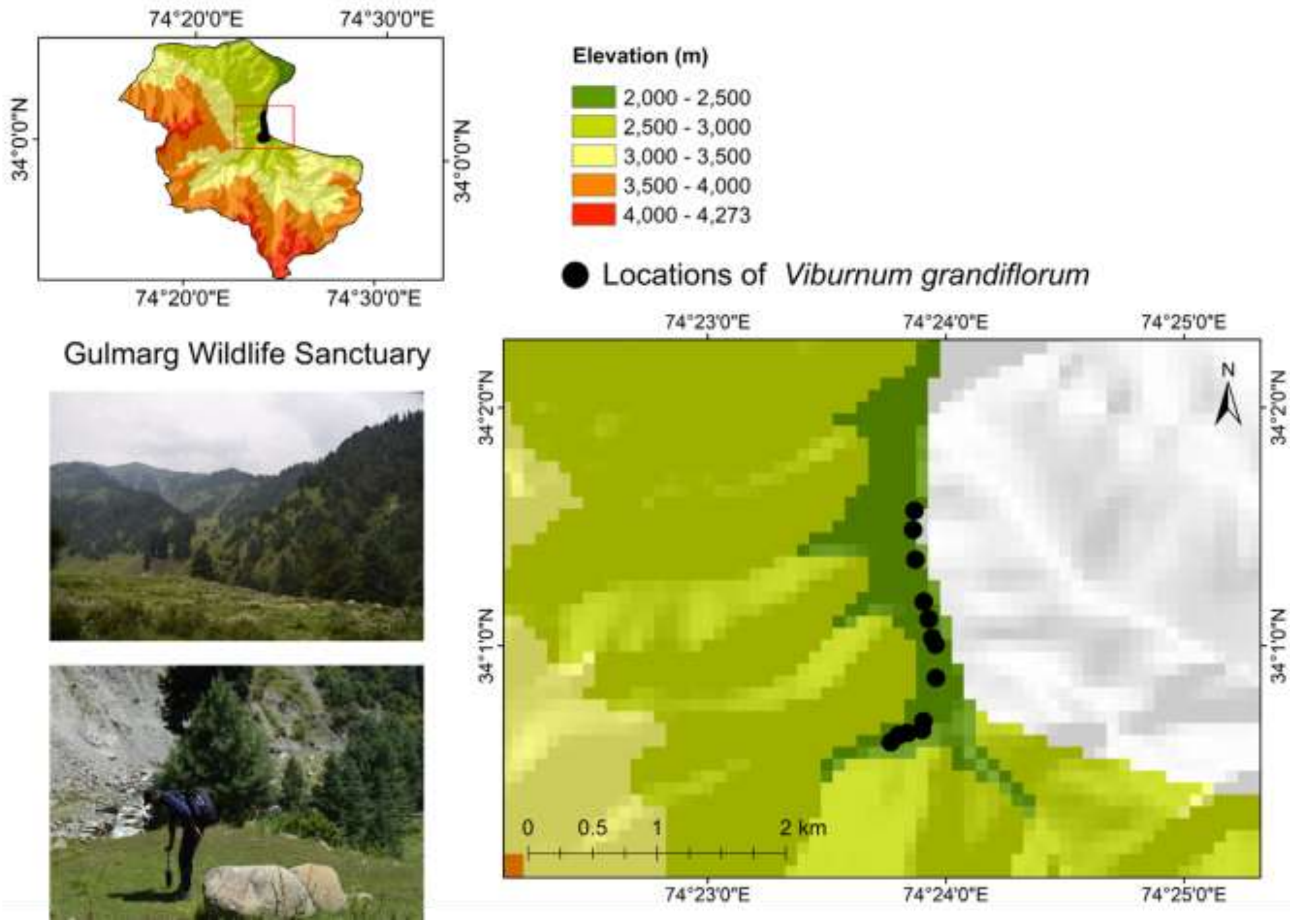


Figure 14 Location map of study area Gulmarg Wildlife Sanctuary

3.5 Details of Field Survey arranged

Field surveys were conducted in the Gwaldam area of Chamoli district and Munsiari area (Khaliya top, Kalamuni, Thamrikund, Jeoljibi, Askot, Baram, Lumti, Madkot) of Pithoragarh district in the state of Uttarakhand. The field sampling was conducted in End of September 2019 and at the End of February 2020 in Pindari Glacier and Almora, Kausani, Jageswer, and Ranikhet respectively. During the period April 2020 – March 2021 field expeditions were conducted in the Gwaldam area of Chamoli district and Munsiyari area (Khaliya top, Kalamuni, Thamrikund, Jeoljibi, Askot, Baram, Lumti, Madkot) of Pithoragarh district in the state of Uttarakhand.

3.6 Strategic Planning for each Activity

3.6.1 Methodology: Random sampling was done in the study sites to take an overview of the total plant diversity of the area. Then the high-value plant species were categorized and targeted for extensive study.

3.6.2 Phytochemical analyses: Phytochemical analyses of two important species viz. *H. Salicifolia* and *Ilex dipyrena* collected during a previous field visit to the Pindari area were conducted. The berries were collected for the estimation of different phytochemicals present in the species. Polyphenolic content (total phenolic content, total flavonoid content, total flavonols content, total tannin content, total proanthocyanidin content) and antioxidant potential were analyzed. *H. salicifolia* is an important hardy deciduous shrub, belonging to the family Elaeagnaceae and popularized as Seabuckthorn. Their yellowish-orange berries are consumed to meet food supplements by the rural populace of Himalaya. Seabuckthorn is an environmentally and financially vital species of the Himalayan region, known as 'Ladakh Gold', 'Golden Bush' or 'Gold Mine', in Himachal Pradesh, of Uttarakhand Ames in Chamoli, Ameel in Uttarkashi and Chook in Pithoragarh district. Besides this, the phytochemical analyses of fruits of *Ilex dipyrena* were also done (Figure 15). Fresh berries of *Hippophae salcifolia* & *Ilex* were randomly collected from thirty individuals. 20 g fresh berries extracted with 200 ml of ethanol (80%), methanol (80%) distilled water. Extractions were homogenized (Ultrasonicater Toshiba-India) for 5 minutes, placed in the water bath for 1 h (30°C) and orbital shaker for 14 hrs (22°C±1°C), respectively. Then Ultrasonicated, centrifuged and filtered before conducting the analyses.

Figure 15 Hippophae salcifolia and Ilex dipyrena

3.6.3 Results: The collected species and their family are given in Tables 7 and 8.

Floristic diversity: A total of 42 species (including 12 lichen species) represented by 25 families (6 lichen families) and 36 genera (11 lichen genera) were sampled during the study (Table 1, Graph 1). Out of the sampled 42 species, 17 were trees, 10 were shrubs, 3 were herbs and 12 were lichens.

Table 7 Plant species sampled from different regions of Bageshwar and Almora districts

Sl. No.	Plant taxa	Family	Almora	Bageshwar
1.	Acacia delbata	Mimosaceae	+	-
2.	Achyranthes aspera	Amaranthaceae	+	-
3.	Aconitum hetrophyllum	Ranunculaceae	-	+
4.	Adhatoda vasica	Acanthaceae	+	-
5.	Adiantum edgeworthii	Pteridaceae	-	+
6.	Aesculus indica	Sapindaceae	-	+
7.	Alnus nepalensis	Betulaceae	-	+
8.	Anemone rivularis	Ranunculaceae	-	+
9.	Artemisia maritima	Asteraceae	+	-
10.	Asparagus racemosus	Asparagaceae	+	-
11.	Aster falconeri	Asteraceae	-	+
12.	Bauhinia variegata	Fabaceae	+	-
13.	Begonia picta	Begoniaceae	-	+
14.	Berberis aristata	Berberidaceae	+	+
15.	Berberis asiatica	Berberidaceae	+	-
16.	Berberis jaeschkeana	Berberidaceae	-	+

17.	Berberis umbellate	Berberidaceae	-	+
18.	Bergenia ciliate	Saxifragaceae	+	-
19.	Bergenia ligulata	Saxifragaceae	+	-
20.	Betula utilis	Betulaceae	-	+
21.	Bidens pilosa	Asteraceae	+	-
22.	Bistorta affinis	Polygonaceae	-	+
23.	Boenninghausenia albiflora	Rutaceae	-	+
24.	Bombax ceiba	Malvaceae	+	-
25.	Callistemon	Myrtaceae	+	-
26.	Cannabis sativa	Cannabaceae	+	+
27.	Cedrus deodara	Pinaceae	+	-
28.	Chenopodium album	Amaranthaceae	+	-
29.	Cicerbita macrorhiza	Asteraceae	-	+
30.	Clematis sp.	Ranunculaceae	+	-
31.	Colebrookea oppositifolia	Lamiaceae	+	-
32.	Corylus jacquemontii	Betulaceae	-	+
33.	Cotoneaster microphyllum	Rosaceae	-	+
34.	Cupressus torulosa	Cupressaceae	+	-
35.	Cyananthus lobatus	Campanulaceae	-	+
36.	Cyathula tomentosa	Amaranthaceae	+	-
37.	Dactylorhiza hatagirea	Orchidaceae	-	+
38.	Datura metal	Papaveraceae	+	-
39.	Delphinium brunonianum	Ranunculaceae	-	+

40.	Desmodium sp.	Fabaceae	+	-
41.	Diploknema butyracea	Sapotaceae	+	-
42.	Dodonaea viscosa	Sapindaceae	+	-
43.	Dolomiaea macrocephala	Asteraceae	-	+
44.	Elsholtzia fruticosa	Lamiaceae	-	+
45.	Eucalyptus sp.	Myrtaceae	+	-
46.	Eupatorium	Asteraceae	+	-
47.	Euphorbia esula	Euphorbiaceae	+	-
48.	Ficus auriculata	Moraceae	+	-
49.	Ficus palmate	Moraceae	+	-
50.	Flemingia sp.	Fabaceae	+	-
51.	Fragaria nubicola	Rosaceae	-	+
52.	Galinsoga parviflora	Asteraceae	-	+
53.	Galium aparine	Rubiaceae	-	+
54.	Gaultheria trichophylla	Ericaceae	-	+
55.	Geranium himalayense	Geraniaceae	+	+
56.	Geranium wallichianum	Geraniaceae	-	+
57.	Gravillea robusta	Proteaceae	+	-
58.	Grewia optiva	Tiliaceae	+	-
59.	Hedera nepalensis	Araliaceae	-	+
60.	Hedychium spicatum	Zingiberaceae	-	+
61.	Hemiphragma heterophyllum	Scrophulariaceae	-	+
62.	Hippophae salicifolia	Elaeagnaceae	-	+

63.	Hypericum oblongifolium	Hypericaceae	+	-
64.	Hypericum uralum	Hypericaceae	-	+
65.	Impatiens sulcata	Balsaminaceae	-	+
66.	Indigofera indica	Fabaceae	+	-
67.	Inula cuspidate	Asteraceae	-	+
68.	Jacaranda mimosifolio	Bignoniaceae	+	-
69.	Jasminum humile	Oleaceae	+	-
70.	Juniperus communis	Cupressaceae	+	-
71.	Lactuca dissecta	Asteraceae	-	+
72.	Lecanthus peduncularis	Urticaceae	-	+
73.	Litsea sp.	Lauraceae	+	-
74.	Lycopodium japonicum	Lycopodiaceae	-	+
75.	Lyonia ovalifolia	Ericaceae	-	+
76.	Mahonia jaunsarensis	Berberidaceae	+	-
77.	Mahonia napalensis	Berberidaceae	+	-
78.	Mimosa bicolour	Fabaceae	+	-
79.	Morina longifolia	Caprifoliaceae	-	+
80.	Morus alba	Moraceae	+	-
81.	Murraya koenigii	Rutaceae	+	-
82.	Myosotis caespitosa	Boraginaceae	-	+
83.	Myrica esculenta	Myricaceae	+	-
84.	Nardostachys jatamansi	Valerianaceae	-	+
85.	Nasturtium rentals	Tropaeolaceae	+	-

86.	Origanum vulgare	Lamiaceae	-	+
87.	Oxalis corniculata	Oxalidaceae	+	-
88.	Oxyria digyna	Polygonaceae	-	+
89.	Parnassia nubicola	Saxifragaceae	-	+
90.	Pedicularis bicornuta	Orobanchaceae	-	+
91.	Pedicularis hoffmesisteri	Orobanchaceae	-	+
92.	Persicaria capitata	Polygonaceae	-	+
93.	Persicaria nepalensis	Polygonaceae	-	+
94.	Phlomis bracteosa	Lamiaceae	-	+
95.	Picrorhiza kurroa	Plantaginaceae	-	+
96.	Pinus roxhburghi	Pinaceae	+	-
97.	Pittosporum eriocarpum	Pittosporaceae	+	-
98.	Plectranthus barbatus	Lamiaceae	-	+
99.	Polygonatum cirrhifolium	Asparagaceae	-	+
100.	Polygonatum verticillatum	Asparagaceae	-	+
101.	Polygonum capitatum	Polygonaceae	+	-
102.	Polygonum polystachyum	Polygonaceae	-	+
103.	Potentilla cuneata	Rosaceae	-	+
104.	Prenanthes brunoniana	Asteraceae	-	+
105.	Primula denticulata	Primulaceae	+	-
106.	Prinsepia utilis	Rosaceae	+	-
107.	Prunella vulgaris	Lamiaceae	-	+
108.	Prunus cerasoides	Rosaceae	+	-

109.	Pyrus pashia	Rosaceae	-	+
110.	Quercus floribunda	Fagaceae	+	+
111.	Quercus glauca	Fagaceae	+	-
112.	Quercus leucotrichophora	Fagaceae	+	-
113.	Ranunculus sceleratus	Ranunculaceae	+	-
114.	Rheum austral	Polygonaceae	-	+
115.	Rhododendron arboreum	Ericaceae	+	+
116.	Rhododendron barbatum	Ericaceae	-	+
117.	Rhododendron campanulatum	Ericaceae	-	+
118.	Roscoea purpurea	Zingiberaceae	-	+
119.	Rosmarinus officinalis	Lamiaceae	+	-
120.	Rubia cordifolia	Rubiaceae	+	+
121.	Rubus ellipticus	Rosaceae	+	-
122.	Rubus paniculatus	Rosaceae	-	+
123.	Rumex hastatus	Polygonaceae	+	-
124.	Salix tetrasperma	Salicaceae	+	-
125.	Salvia lanata	Lamiaceae	+	-
126.	Saussurea taraxacifolia	Asteraceae	-	+
127.	Saxifraga moorcroftiana	saxifragaceae	-	+
128.	Saxifraga parnassifolia	saxifragaceae	-	+
129.	Sedum multicaule	Crassulaceae	-	+
130.	Selinum vaginatum	Apiaceae	-	+
131.	Senecio laetus	Asteraceae	-	+

132.	Sibbaldia parviflora	Rosaceae	-	+
133.	Sinopodophyllum hexandrum	Berberidaceae	-	+
134.	Smilex aspera	Smilacaceae	+	-
135.	Solanum torum	Solanaceae	+	-
136.	Solidago virgaurea	Asteraceae	-	+
137.	Sorbus ursine	Rosaceae	-	+
138.	Stachys melissifolia	Lamiaceae	-	+
139.	Stellaria media	Caryophyllaceae	+	-
140.	Strobilanthes atropurpureus	Acanthaceae	-	+
141.	Strobilanthes bracteata	Acanthaceae	+	-
142.	Swertia ciliate	Gentianaceae	-	+
143.	Swertia cordata	Gentianaceae	-	+
144.	Synotis alata	Asteraceae	-	+
145.	Taxus wallichiana	Taxaceae	-	+
146.	Tectona grandis	Lamiaceaea	+	-
147.	Thalictrum	Ranunculaceae	-	+
148.	Toona ciliata	Meliaceae	+	-
149.	Trachycarpus takil	Arecaceae	+	-
150.	Trifolium repens	Fabaceae	+	-
151.	Urtica diocia	Urticaceae	+	-
152.	Valeriana wallichii	Valerianaceae	+	-
153.	Verbascum thapsus	Scrophulariaceae	+	-
154.	Viburnum mullaha	Adoxaceae	-	+

155.	Viola sp.	Violaceae	+	-
156.	Woodfordia fruticosa	Lythraceae	+	-
157.	Yucca gloriosa	Asparagaceae	+	-
158.	Zanthoxylum armatum	Rutaceae	+	-
159.	Ziziphus mauritiana	Rhamnaceae	+	-

Table 8 Sampled species during April 2020-March 2021 from Gwaldam and Munsiari areas of Uttarakhand

S. No.	**Taxa**	**Family**	**Gwaldam**	**Munsiari**	**Habit**	**Uses**	**Part(s) used**
1.	*Acacia dealbata* A. Cunn.	Mimosaceae	+	+	Tree	Fuel	Wood
2.	*Adhatoda vasica* Nees	Acanthaceae	-	+	Shrub	Adhatoda is useful for curing coughs, colds and asthma.	
3.	*Aesculus indica* (Wall. ex Cambess.) Hook.	Sapindaceae	-	+	Tree	-	-
4.	*Alnus nepalensis* D. Don	Betulaceae	+	+	Tree	-	-
5.	*Berberis aristata* DC.	Berberidaceae	+	+	Shrub	It is claimed to possess antibacterial, anti-inflammatory, antipyretic and antiseptic properties, the herb is used as	Root

						a cholagogue, stomachic, laxative and diaphoretic.	
6.	*Berberis asiatica* Roxb.	Berberidaceae	+	+	Shrub	Roots are used in treating ulcers, urethral discharges, ophthalmia, jaundice, fevers etc. The bark and wood are crushed in Nepal then boiled in water, strained and the liquid evaporated until a viscous mass is obtained. This is antibacterial, laxative and tonic. It is taken internally to treat fevers and is used externally to treat conjuctivitis and other inflammations of the eyes. Tender leaf	Bark, root, wood

						buds are chewed and held against affected teeth for 15 minutes to treat dental caries. The fruit is cooling and laxative.	
7.	*Berberis umbellata* Wall. ex G. Don	Berberidaceae	-	+	Shrub	-	-
8.	*Bergenia ciliata* Sternb.	Saxifragaceae	+	+	Herb	The root of Chinese Bergenia is rich in tannins and is used as a tonic in the treatment of fevers, diarhoea and pulmonary affections. The root has a high reputation in indigenous systems of medicine for dissolving stones in the kidneys. The root is bruised and applied externally to	Root

						treat boils and ophthalmia.	
9.	*Betula utilis* D.Don	Betulaceae	-	+	Tree	Bark was used for writing in ancient India.	Bark
10	*Cedrus deodara* (Roxb. ex D. Don) G. Don	Pinaceae	+	-	Tree	Timber	Wood
11	*Colebrookea oppositifolia* Lodd.	Lamiaceae	+	+	Shrub	-	-
12	*Cotoneaster microphyllum*	Rosaceae	+	+	Shrub	-	-
13	*Cupressus torulosa* D. Don ex Lamb.	Cupressaceae	+	+	Tree	-	-
14	*Cyathula tomentosa* (Roth) Moq.	Amaranthaceae	+	+	Shrub	-	-
15	*Diploknema butyracea* (Roxb.) H.J. Lam	Sapotaceae	-	+	Tree	Food, fuel wood, timber, fodder and other products. It is also used to cure many diseases and body disorders by the community people such as	Whole plant

						in inflammation of tonsils, rheu-matism, itching, ulcers, hemorrhage, etc. The wax obtained from seeds is used in cosmetics, soaps.	
16	*Dolichousnea longissima* (Ach.) Articus	Parmeliaceae	-	+	Lichen	The Baiga tribe of Madhya Pradesh uses this lichen along with other ingredients for treating bone fracture; the Nitinaht Indians of Vancouver Island use the lichen for wound dressing as it has absorbent quality; In Turkey it is used in the treatment of gastric ulcer in local folk medicine.	Whole plant (Thallus)
17	*Hemiphragma heterophyllum*	Scrophulariaceae	-	+	Herb	-	-

	Wall.						
18	*Heterodermia diademata* (Taylor) D.D. Awasthi	Physciaceae	-	+	Lichen	In India Nepalese of Chaunje Basti, near Gangtok in Sikkim apply this lichen on cuts and wound as plaster to protect from water and infection.	Whole plant (Thallus)
19	*Hypotrachyna cirrhata* (Fr.) Divakar, A. Crespo, Sipman, Elix & Lumbsch	Physciaceae	-	+	Lichen	It is used in Ayurvedic and Unani medicine under the name 'Chharila' as carminative and aphrodisiac and considered useful in dyspepsia, spermatorrhoea, amenorrhoea, calculi, diseases of blood and heart, stomach disorders, enlarged spleen, bronchitis, bleeding piles, Scabies,	Whole plant (Thallus)

							Leprosy, excessive salivation, soreness of throat, tooth-ache and general pain; Gaddi tribe of Kangra valley, Himachal Pradesh use it as material for sacrificial fire (havan samgri); Bhaiga, Bhil, Bhilala, Gond, Korka, Muria of Madhya Pradesh and Lepcha ans Nepalese of Sakyong valley in Sikkim use it as a spice in meat and vegetables.	
20	*Indigofera indica* Lam.	Fabaceae	+	+	Herb	It is used for Glowing Skin, Sores, Ringworm, Blisters, Hair rejuvenation, natural hair darkening,	Leaves

						removing Worms in Teeth and Gums, Sores on Skin, Liver, Urinary problems and poison, Mouth ulcers or Canker sores, Kidney disease, Insects in Ear, Black Hair and Hair fall, Dog bite.	
21	*Lobaria retigera* (Bory) Trev.	Lobariaceae	-	+	Lichen	Skin diseases	Whole Plant (Thallus)
22	*Lyonia ovalifolia* (Wall.) Drude	Ericaceae	+	+	Tree	The young leaves and buds are toxic, but they are used externally as an infusion to treat skin diseases and external parasites.	Leaves
23	*Myrica esculenta* Buch.-Ham. ex D. Don	Myricaceae	+	+	Tree	Edible	Fruits
24	*Nephromopsis nephromoides*	Parmeliaceae	-	+	Lichen	-	-

	(Nyl.) Ahti & Randlane						
25	*Parmotrema reticulatum* (Taylor) M. Choisy	Parmeliaceae	-	+	Lichen	The tea is commonly prepared in late afternoon and left for one night before being drunk to relieve discomfort from kidney disorders or venereal disease.	Whole Plant (Thallus)
26	*Peltigera polydactylon* (Neck.) Hoffm.	Peltigeraceae	-	+	Lichen	In India Lepchas and Nepalese of Skyong Valley, north Sikkim apply the paste of *P. polydactylon* to cuts to stop bleeding and as an antiseptic, that cure wounds	Whole Plant (Thallus)
27	*Pinus roxburghii* Sarg.	Pinaceae	+	+	Tree	The turpentine obtained from the resin of all pine trees is antiseptic, diuretic,	Resin

						rubefacient and vermifuge. It is a valuable remedy used internally in the treatment of kidney and bladder complaints and is used both internally and as a rub and steam bath in the treatment of rheumatic affections. It is also very beneficial to the respiratory system and so is useful in treating diseases of the mucous membranes and respiratory complaints such as coughs, colds, influenza and TB. Externally it is a very beneficial treatment for a	

						variety of skin complaints, wounds, sores, burns, boils etc and is used in the form of liniment plasters, poultices, herbal steam baths and inhalers. The wood is diaphoretic and stimulant. It is useful in treating burning of the body, cough, fainting and ulcers.	
28	*Prinsepia utilis* Royle	Rosaceae	+	+	Shrub	-	-
29	*Pyracantha crenulata* (D.Don.)M.Roem.	Rosaceae	+	+	Shrub	Wild edible	Fruits
30	*Quercus leucotrichophora* A. Camus	Fagaceae	+	+	Tree	Fuel, fodder	Leaves, wood
31	*Quercus semecarpifolia* Sm.	Fagaceae	-	+	Tree	Fuel, fodder	
32	*Ramalina*	Ramalinaceae	-	+	Liche	Anti-bacterial	Whole

	conduplicans Vain.				n		plant (Thallus)
33	*Ramalina sinensis* Jatta	Ramalinaceae	-	+	Lichen	Anti-bacterial, Anti-fungal	Whole plant (Thallus)
34	*Rhizocarpon geographicum* (L.) DC.	Rhizocarpaceae	-	+	Lichen	-	-
35	*Rhododendron arboreum*	Ericaceae	+	+	Tree	Wild edible	Fruits
36	*Rhododendron campanulatum* D. Don	Ericaceae	-	+	Tree	-	-
37	*Rhododendron* sp.	Ericaceae	-	+	Tree	-	
38	*Rubus ellipticus* Sm.	Rosaceae	+	+	Shrub	Wild edible	Fruit
39	*Shorea robusta*	Dipterocarpaceae	-	+	Tree	Timber	Wood
40	*Sulcaria sulcata* (Lév.) Bystrek	Parmeliaceae	-	+	Lichen	Anti-viral, used for dizziness, kidney infections, weakness, heart palpitation, night sweating, edema,	Whole plant (Thallus)

						impetigo and sores, used for aching back and legs, traumatic bleeding, menstrual irregularities, uterus prolapses, epilepsy, paralysis, impotence, drunk decoction or applied to the affected area.	
41	*Taxus wallichiana* Zucc.	Taxaceae	-	+	Tree	Anti-cancerous	Bark
42	*Usnea splendens* Stirt.	Parmeliaceae	-	+	Lichen	-	-

3.7 Activity wise Time frame followed [using Gantt/ PERT Chart]

4 KEY FINDINGS AND RESULTS

4.1 Major Research Findings

Objective-wise findings are described in detail. **Objective 1: Discrimination of spectral species in parts of Nanda Devi Biosphere Reserve**

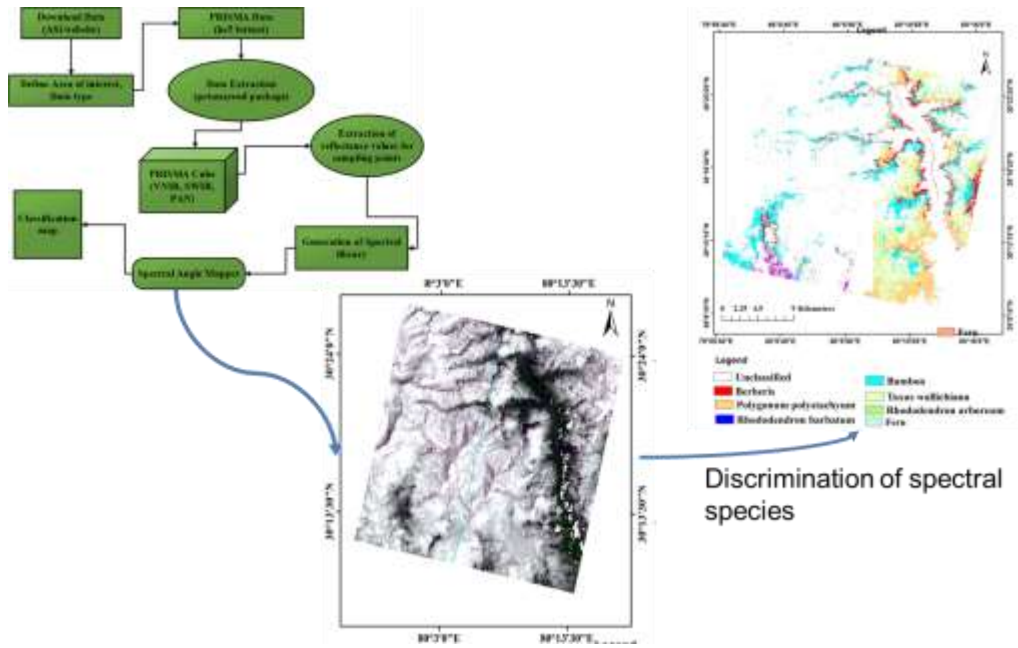


Figure 16 Discrimination of spectral species

Discrimination of species using PRISMA hyperspectral data has been completed for the parts of Nandadevi Biosphere Reserve. Spectral and image analysis for discrimination of selected economically important plant species such as Berberis, Polygonum polystachyum, Rhododendron Barbatum, Rhododendron Arboreum, Taxus Wallichiana, Bamboo and Fern is over.

Objective 2:Spectral library of the economically important plant species to use with the Hyperspectral satellite and airborne data for large-scale quantification

❖ Development of the SHRIME

Highlights

- India's first hyperspectral library for Himalayan plants
- Loaded with spectrum > 400 spectral species
- Offers a dynamic platform for basic analysis and understanding of biophysical and hyperspectral aspects of Himalayan plants
- Location, importance, conservation status and other remote sensing parameters including vegetation indices

Figure 17 SHRIME tool

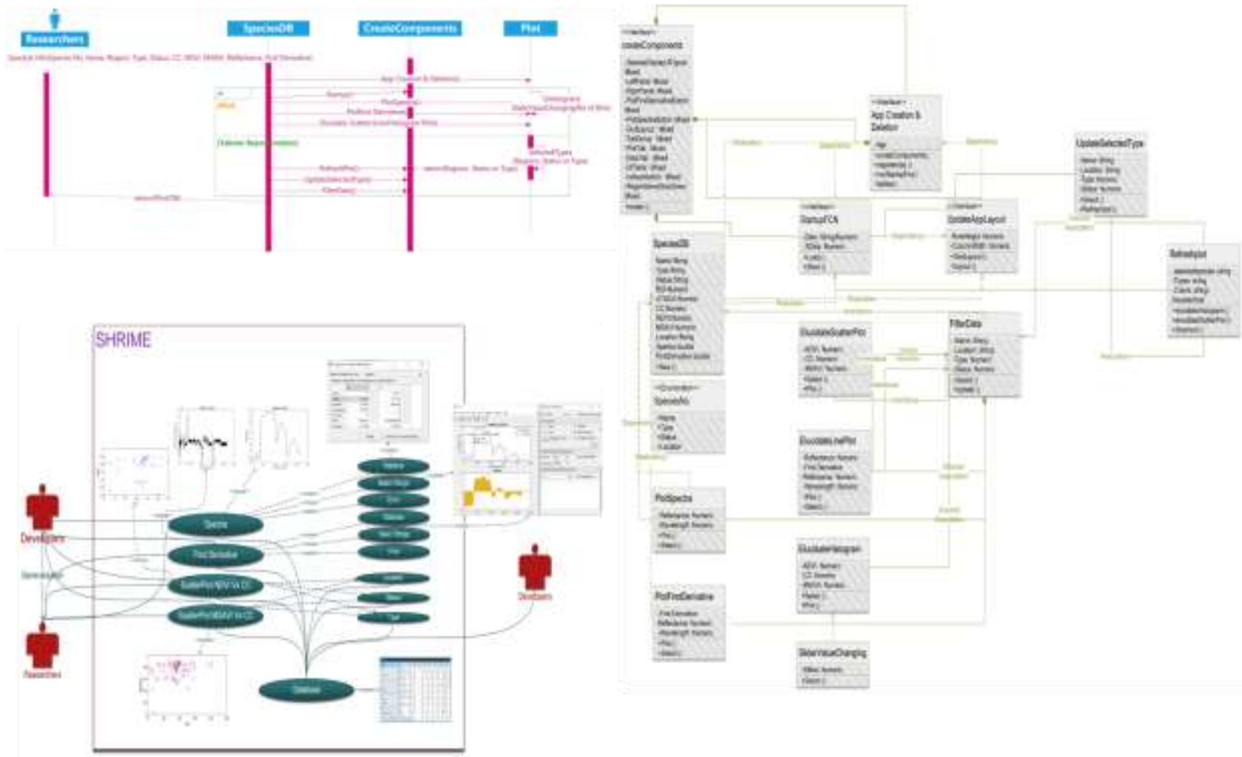


Figure 18 Class and UML Diagram of SHRIME.

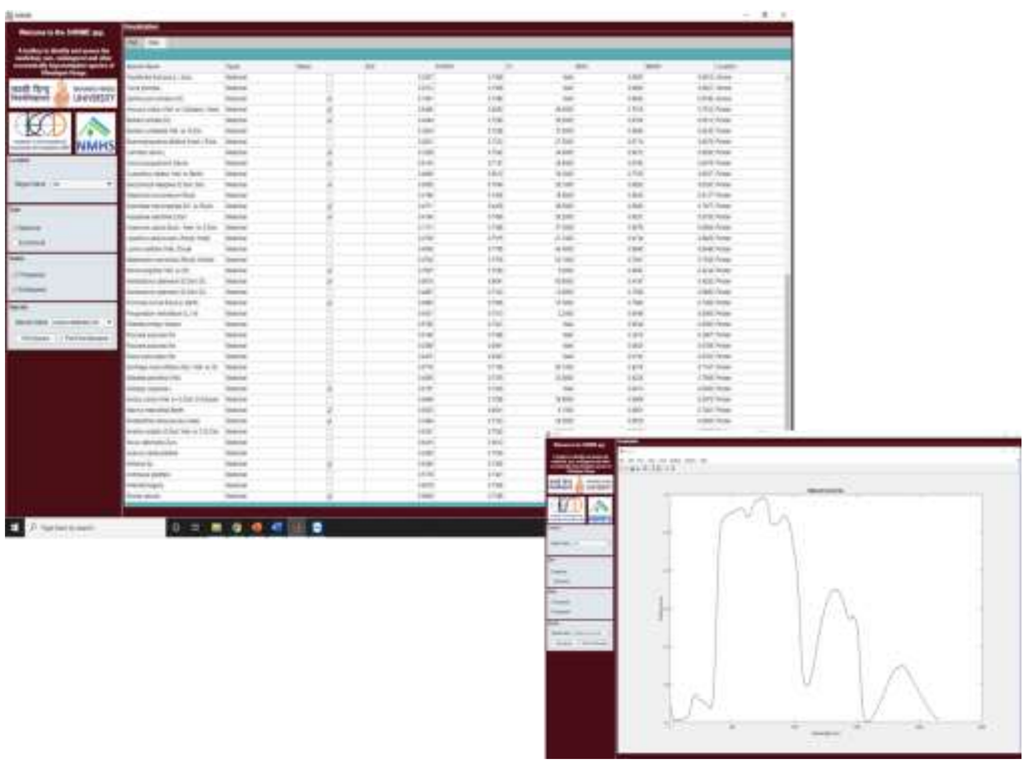


Figure 19 Snapshot of the SHRIME tool

The Spectral tool of Himalayan Rare, Invasive, Medicinal and Economical (SHRIME) Plant Species is a MATLAB tool (Manish Kumar Pandey, 2022) that could be utilized to create a spectral library in terms of reflectance and first derivative for the Himalayan plant species and a toolbox to be installed to extend its functionality further. This tool is developed for the Quantification of biophysical and biochemical parameters, which are crucial for monitoring the high-value plants in the Himalayan region, their sustainable utilization, and the conservation of related resources. The ultimate goal of this tool is in offering a dynamic platform for reflectance spectrum and derivative spectrum generation of various plant species, creation of the spectral library, application of basic fitting operations on the spectra generated, understanding the relationship between the chlorophyll content with NDVI/MSAVI and a database of the various Himalayan species in terms of location, type, status and other parameters.

Objective: 3 Development of forward and inverse models for the retrieval of biophysical & biochemical parameters from economically important plants using the Hyperspectral data

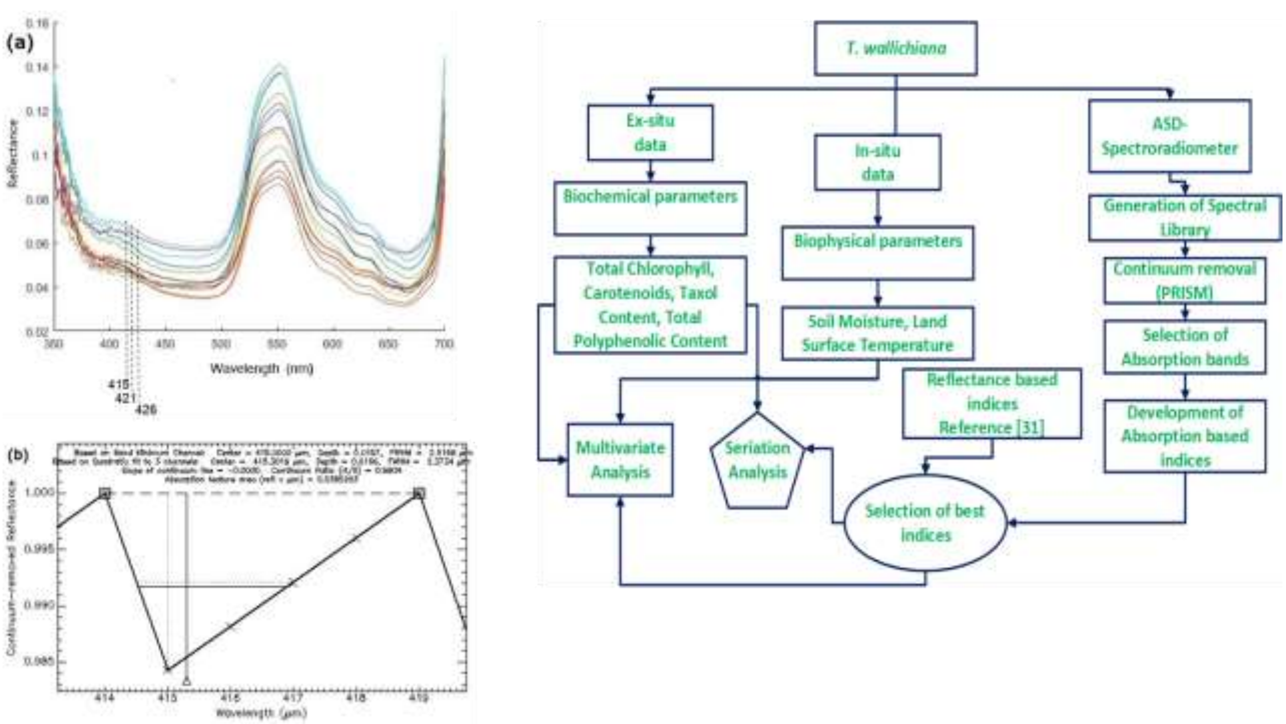


Figure 20 Flowchart for Understanding Influencing Factors for Taxol Content Variation in High Altitude Himalayan Region

Retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species Taxus wallichiana. For the retrieval of biochemical parameters, various indices were generated and three different types of filters, Average Mean, Savitzky Golay, and Fast Fourier Transform (FFT) one at a time followed by feature selection on each denoised spectra were applied to smooth the spectra for better estimation of biochemical parameters. The development and validation of hyperspectral indices for anti-cancerous Taxol content estimation were conducted as part of this objective, the details of which are available (Gupta et al., 2022). Further, Taxol Content Variation was also explored in altitudinal variation in the Himalayan regions, the details of which are available (Gupta et al., 2021). Taxol Mapping was also explored through drone quasi-hyperspectral images as depicted in Figure 25.

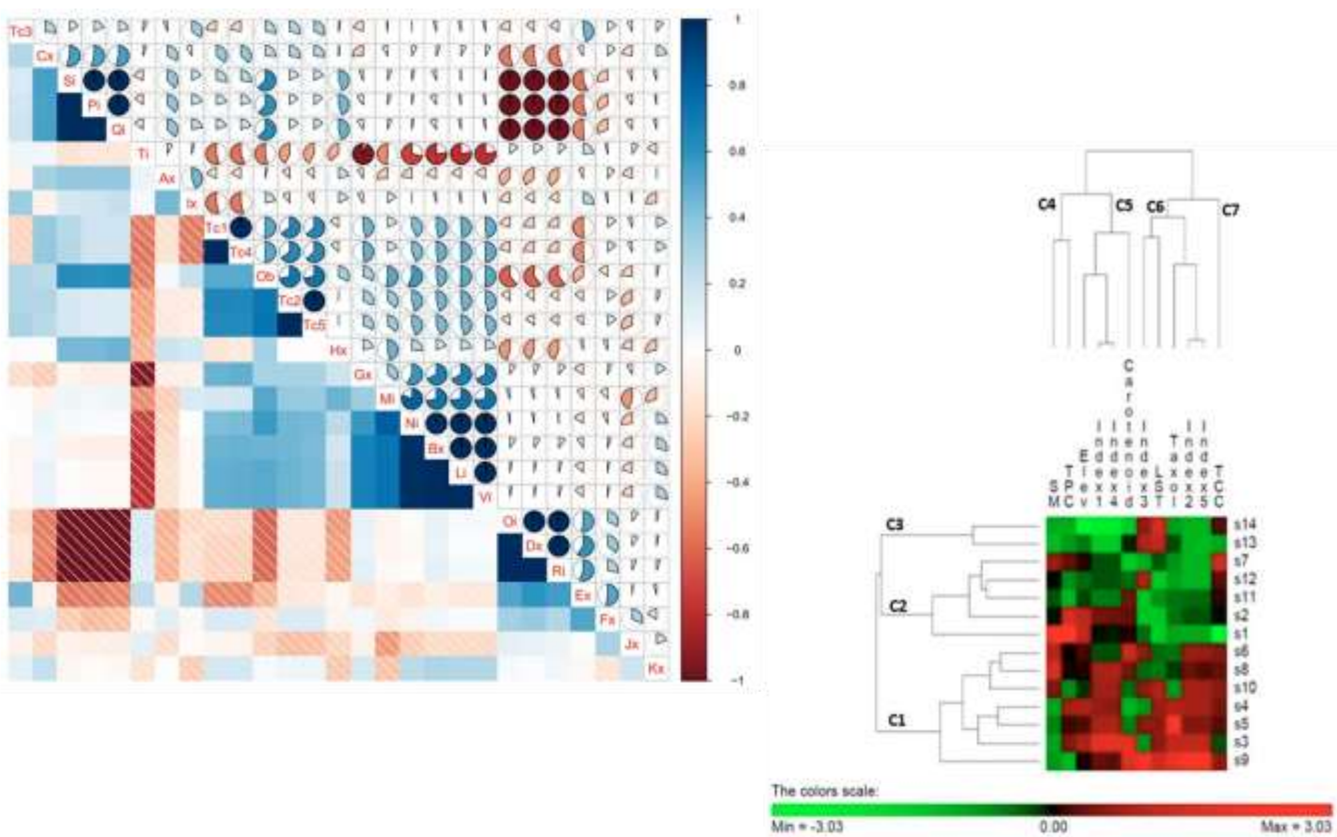


Figure 21 Pearson’s correlation between the selected wavelength represented as ‘x’ as a subscript, absorption indices represented with ‘i’ as a subscript, and measured TC represented as ‘Ob’.

Figure 21 illustrates the correlation among all the possible absorption band values along with the absorption and reflectance indices values that showed a significant correlation with the taxol content. The absorbance wavelengths are represented with ‘x’ as a subscript and the absorption indices values are represented with ‘i’, while the reflectance-based indices are represented with TC. Figure 21 shows that the measured taxol content (Ob) showed the nearest positive correlation with reflectance-based indices TC 2 ($r = 0.741$) and TC5 ($r = 0.742$). The Ob values also showed a positive correlation with absorption indices Ni ($r = 0.565$) and Mi ($r = 0.561$), while a significant negative correlation was observed between Ob and Ri ($r = 0.604$) and Oi ($r = -0.615$). The parameters Si, Pi, and Qi also showed a significant positive correlation, but the magnitude values were out of range. The seriation column matrix plot Cluster C4 contains TPC and SM, while cluster C5 contains carotenoids and elevation along with TC 1 (Index 1) and TC 4 (Index 4). This implies that the elevation, TPC, SM, and Carotenoids values show a relation with TC 1 and TC 4. This implies that elevation, due to the change in albedo, has a direct relationship with biochemical and edaphic properties. The C6 cluster contains TC3 (Index 3), LST, Taxol, TC 2 (Index 2), and TC 5 (Index 5). Carotenoids show a close relationship with both indices-generated values TC 4 and TC 3. Similarly, the Taxol content was found to be feeble with LST, but the correlation coefficient was insignificant to consider. The taxol content has a close correlation with TC 2 and TC 5 with good correlation values and does not relate to any other variable, even in the hierarchical sense. Hence, it can be said that TC 2 and TC 5 can only retrieve taxol, not any other common foliar pigment found in the visible region of electromagnetic spectra. The Total Chlorophyll Content (TCC) behaved as a runt. It consistently showed the same range values without being majorly affected by any other variable. The TCC is supposed to vary with the season, species, age of the plant, and forest type. The samples S6, S8, S10, S4, S5, S3, and S9 are grouped in one cluster, C1, while cluster C2 includes samples S7, S12, S11, and S2. Samples S14 and S15 clustered together in C3. As depicted in Figure 21, several samples from cluster C1 were characterized with similar behaviour. These samples were measured in-between altitudes of 2826 and 3003 m, which is also the dense region of the forest in the sampling area. Therefore, these samples must have a strong influence on the forest and its ecosystem. The samples clustered under C2 and C3 shared the same hierarchy, due to their presence either at the high or low altitude of the sampling elevation, which was marked by human intervention at the lower altitude or ecosystem change at the higher altitude.

Table 9 Selected Taxol indices

Sl No.	Reflectance based Taxol Indices
1	TC 1 = (R426-R421)/(R426+R421)
2	TC 2 = (R415-R421)/(R415+R421)
3	TC 3 = (R601-R608)/(R601-R608)
4	TC 4 = (R421/R426)
5	TC 5 = (R415/R421)

Table 10 Selected Regression model obtained with considerable r value for PRISMA data for taxol estimation

Sl No.	Regression Equation	Correlation Coefficient
1	TC = 2.6006 TC1 + 0.0219 r= 0.798	O.798
2	TC = -1.6035 TC2 - 0.0309 r=-0.798	-0.798
3	TC = -0.1783 TC3 - 0.0214 r= 0.384	0.384

The species distribution of Taxol content through empirical methods and radiative transfer models (physical methods) were conducted as depicted in Figures 22 and 23.

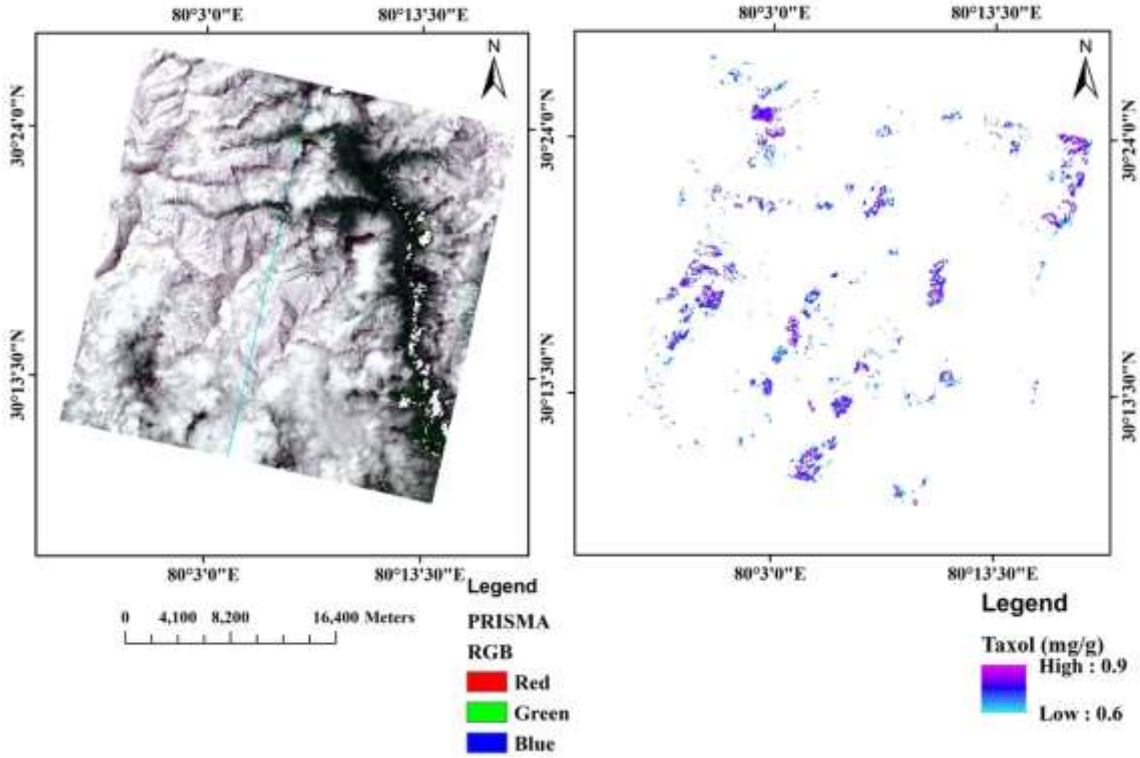
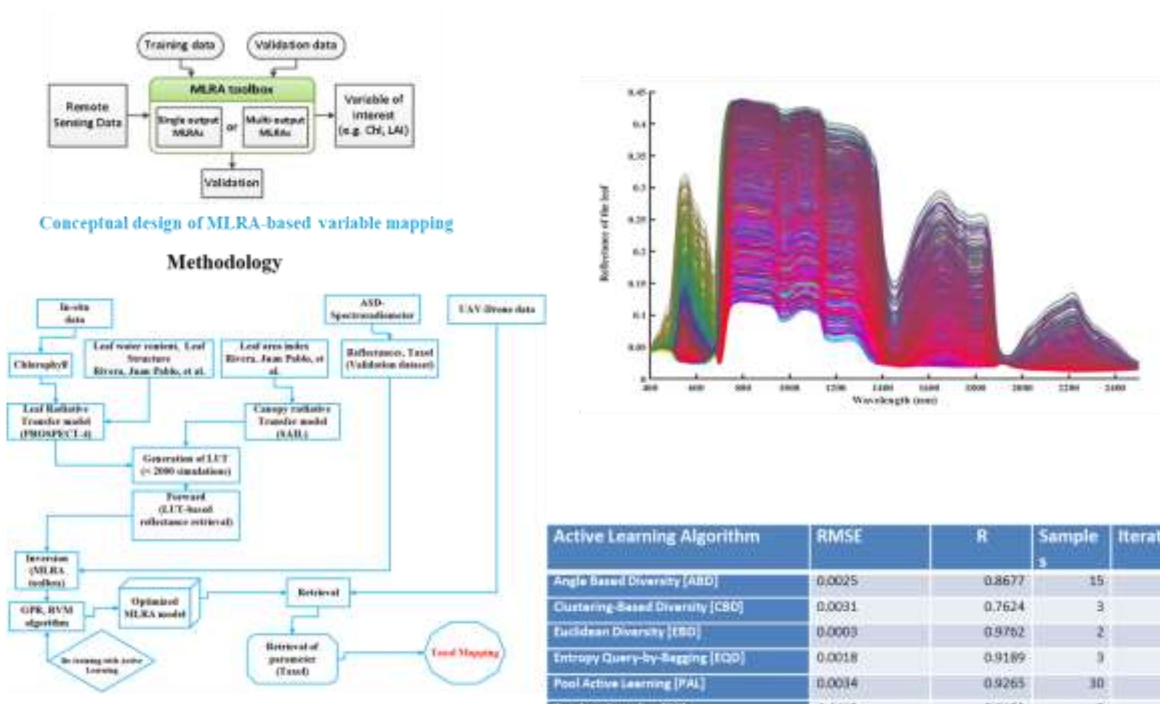


Figure 22 Modelling Spatial distribution of *Taxol* content (Empirical method)



Active Learning Algorithm	RMSE	R	Samples	Iteration
Angle Based Diversity [ABD]	0.0025	0.8677	15	3
Clustering-Based Diversity [CBD]	0.0031	0.7624	3	3
Euclidean Diversity [EBD]	0.0003	0.9762	2	2
Entropy Query-by-Bagging [EQD]	0.0018	0.9189	3	3
Pool Active Learning [PAL]	0.0034	0.9265	30	3
Random Sampling [RS]	4.4412	0.9152	3	3
Residual Active Learning [RSAL]	6.666	0.9003	15	3

Figure 23 Modelling Spatial distribution of Taxol content using radiative transfer model (Physical method)

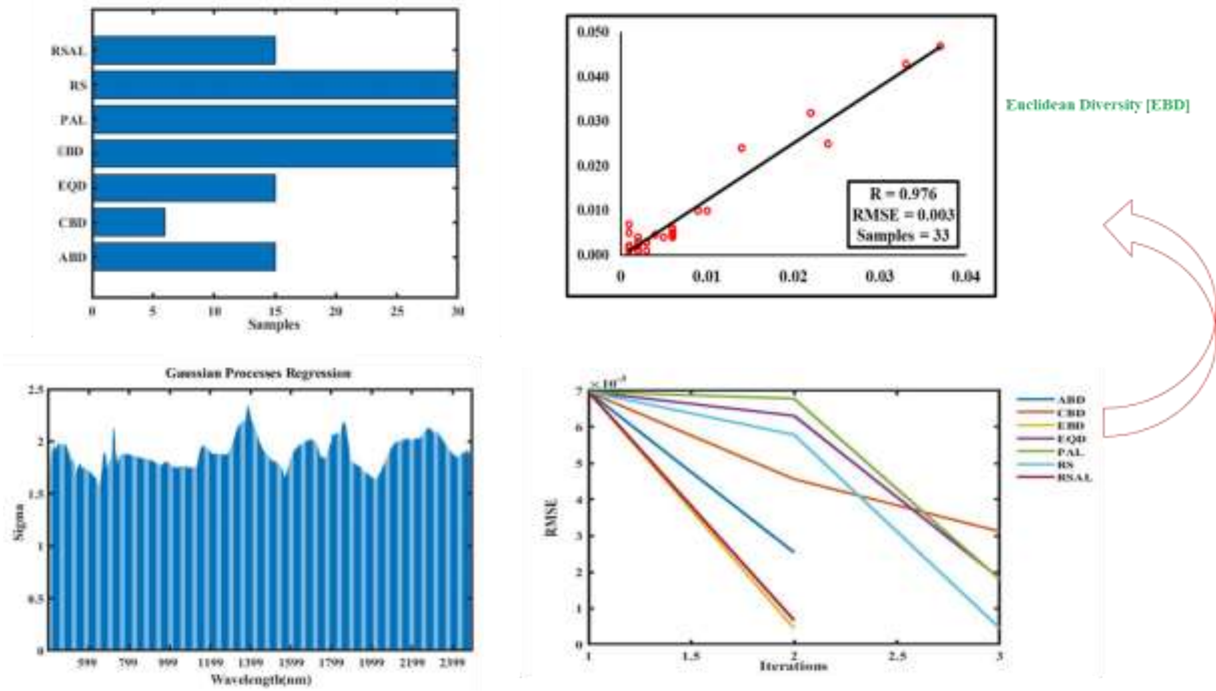


Figure 24 Best algorithm for Taxol retrieval

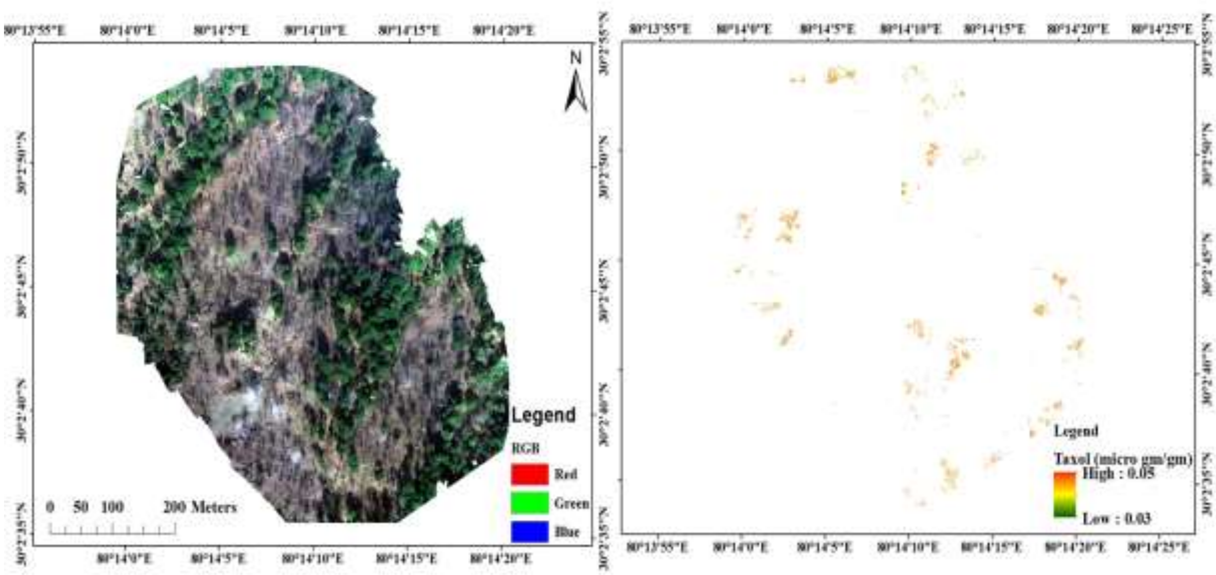


Figure 25 Taxol Mapping using drone quasi hyperspectral

Taxus wallichiana Zucc. (*T. wallichiana*) also known as Himalayan yew is found very promising for cancer treatment. The needles/leaves of the *T. wallichiana* are one of the valuable sources of taxoid. Retrieval of Biomedicinal property of *Taxus wallichiana* was retrieved using indices on spectroradiometric data. The smoothening and filtering technique was applied to spectroradiometer data. Average Mean filter, Savitzky Golay, and FFT filtering technique were paired with derivative analysis to further select the most optimal band for Taxol. The derivative analysis is a feature selection process that is efficient in capturing the subtle difference in the spectra required to locate any specific feature present in the spectra. It works best when paired with an optimal pre-processing technique.

Each spectrum of *T. wallichiana* after application of the filter was followed by the first derivative to select the absorption bands. These selected wavebands were specified in ARTMO in the SI generation toolbox. The derivative analysis presented a different range of wavelengths specified under Vis (370-700nm), Near SWIR (NSWIR-1350-1450nm), and Far SWIR (FSWIR- 1800-2500nm) regions. The Regression-based models on spectral indices are typically empirical formulae that enable the mapping of biophysical and biochemical parameters over a large area. The specified regression-based retrieval strategies with the SI toolbox were analysed and then the input of the wavelength selected using derivative analysis was made along with the raw spectral spectroradiometer data. The model input of spectral information was specified with measured Taxol content as an input variable.

Novel Taxol indices are generated by identifying various combinations from the spectroradiometer raw data between a spectral range of 350-2500 nm. The index for each selected combination wavelength was tested. ARTMO model spectral indices (SI) were specified with the preselected wavelength using feature selection (derivative analysis). Using statistical techniques, the Taxol content (TC) retrieval accuracies of newly developed Taxol models were investigated. The best-selected wavelength for band two-band combination yielded a statistically significant correlation for the Average Mean filter. Based on statistical performance two best models were selected. The two-band combination makes processing time much quicker. Various curve fitting was also tested with real observed and model-generated Taxol content data, where linear curve fitting came out to be the most promising. Model LTC-TC 5 (TC = 0.4154 * TC 5 + 0.0251 and LTC-TC 8 (TC = 0.4175 * TC 8 + 0.0248) showed a correlation of 0.719, 0.718. RMSEr value of both models is relatively equivalent i.e., 0.578 and 0.576 for models LTC-TC 5 and LTC-TC 8.

The model generation with ARTMO also gives the best results when the most appropriate bands after smoothening were given as input. The linear curve fitting performed best between modelled data values from ARTMO, and TC estimated with HPLC analysis. The wavelength used for the best two indices generation was the same. Hence it suggests that Taxol content can be allocated using reflectance values in indices format from the visible range of 415-421 +/- 5 nm window.

Pyracantha crenulata (D. Don) M. Roemer (Rosaceae), commonly called Himalayan firethorn, is an evergreen thorny shrub species found in open slopes that grows between 1,000 and 2,400 m above mean sea level (a.m.s.l.), from Himalaya to South-West China and Myanmar. *P. crenulata* leaves were sampled between elevation 1370-1890m at Ranikhet, Kausani, Jageshwar, and Bhimtal of the Garhwal region of Uttrakhand. The spectra of each sample were recorded using an ASD spectroradiometer. The samples were then stored in liquid nitrogen for further anthocyanin estimation using the destructive method.

P. crenulata is reported to have many traditional and ethnomedicinal uses. It is used in the treatment of hepatic, cardiac, stomach, and skin disease. Recently its fruit extract was found to have antiurolithogenic and diuretic activities and antihypertensive. Moreover, its fruits are considered a potent source of natural antioxidants. Since this herb is found on open slopes it's an estimation on a broader scale that becomes the next logical step for further uses. Usually, plant health is determined by the pigment content present in it. The conventional approach for its estimation involves extraction in organic solvents for subsequent spectrophotometric measures which is time-consuming and unsuitable at a larger scale. Hence the development of vegetation indices and the Statistical linear regression (SLR) model gives more specific indices suitable for the selected species. In this study we evaluated a) three of the already developed indices for anthocyanin at various wavelengths b) modified old indices with hyperspectral data c) a developed SLR model for the estimation of anthocyanin for *P. crenulata*.

With the help of the retrieval model more specified and modified anthocyanin indices were developed. The three most common already developed indices Normalized difference type index, modified anthocyanin content index (mACI), and modified Anthocyanin reflectance index (mARI) were tested for the anthocyanin estimation. mARI was found most appropriate with a correlation of 0.606 at wavelengths 1368nm and 816nm for *P. crenulata*. These results suggest with hyperspectral data the increased number of bands gives better and more accurate indices for retrieval of biochemical vegetation variables.

Table 11 List of tentative compounds identified *Rhododendron arboreum* using UPLC and HRMS

EXPERIMENTAL				THEORETICAL			TENTATIVE COMPOUND	
M-H	**Rt**	**fragments**	**λmax (nm)**	**M-H**	**fragments**	**λmax (nm)**	**formula**	**Name**
447.1143	6.058	300.0303, 151.0041	256, 348	447.1105	301.0353, 179.028, 151.0258	256, 307	$C_{21}H_{20}O_{11}$	Quercitrin
433.0939	5.718	384.9915, 300.0305, 151.0041	256, 352	433.0797	301, 261, 191	256, 356	$C_{20}H_{18}O_{11}$	Quercetin-3-D-xyloside
433.0892	5.417	387.1108, 300.0301, 151.0045	255, 354	433.0797	301, 261, 191	256, 356	$C_{20}H_{18}O_{11}$	Quercetin-3-D-xyloside
863.1954	2.283	711.1422, 573.1095 411.0757	279	863.3000	711.09, 695.01, 573.220	308,280	-	A-type Procyanidin trimer 2
463.0929	4.412	300.0301, 205.0519	255,354	463.0895	301.036, 151.003, 178.998	255,354	$C_{21}H_{20}O_{12}$	Quercetin 3-β-D-glucoside
497.1329	8.510	299.1026, 179.0363	250,267	497.1298	-	254	$C_{22}H_{25}O_{13}$	4-O-Glucopyranosyl-5-O-caffeoylshikimic acid

The current sample analysis focuses on Rhododendron arboreum. Rhododendron arboreum Sm. The world’s most famous rhododendron is known as the national flower of Nepal. The word Rhododendron is derived from the Greek word ‘rhodo’ means rose and dendron means tree. The flowers being sweet and sour taste are used in the preparation of squash, jams, jellies, and local brew. It is used in folklore medicine to treat many disorders and is often used as a substitute for another famous plant drug Rohitaka. Pharmaceutical companies and traditional practitioners of Nepal are using its bark, as a substitute for Rohitaka i.e., *Tecomella undulate (Sm.)* Seem. *R. arboreum* is reported to be effective as a diuretic, choleretic, chronic diarrhoea, anti-irritable bowel syndrome therapy, and astringent. The present work focuses on finding out the other phytoconstituents of the R. arboreum leaves and flowers, the extracts were analysed are being analysed using the UPLC technique. It is observed that qualitatively linoleic acid, flavone 40-OH, Benzenepropanoic acid, hexadecanoic acid (CAS) palmitic acid, and 9,12-Octadecadienoic acid are commonly present in leaves. The compounds identified in the present study belong to different categories e.g., palmitic acid and linoleic acid are fatty acids; Palmitic acid has excellent potential as a free scavenger of free radicals. 9,12-Octadecadien-1-ol possess healing potential against arthritis and inflammation. 9,12 octadecadienoic acid can be used as an anti-inflammatory, anti-microbial, and anti-arthritis agent. Linoleic acid has antimicrobial activities. The existence of these phytoconstituents in R. arboreum points towards its expectant candidature for remedial use. Hence it's an understudy to quantify one of the characteristic compounds.

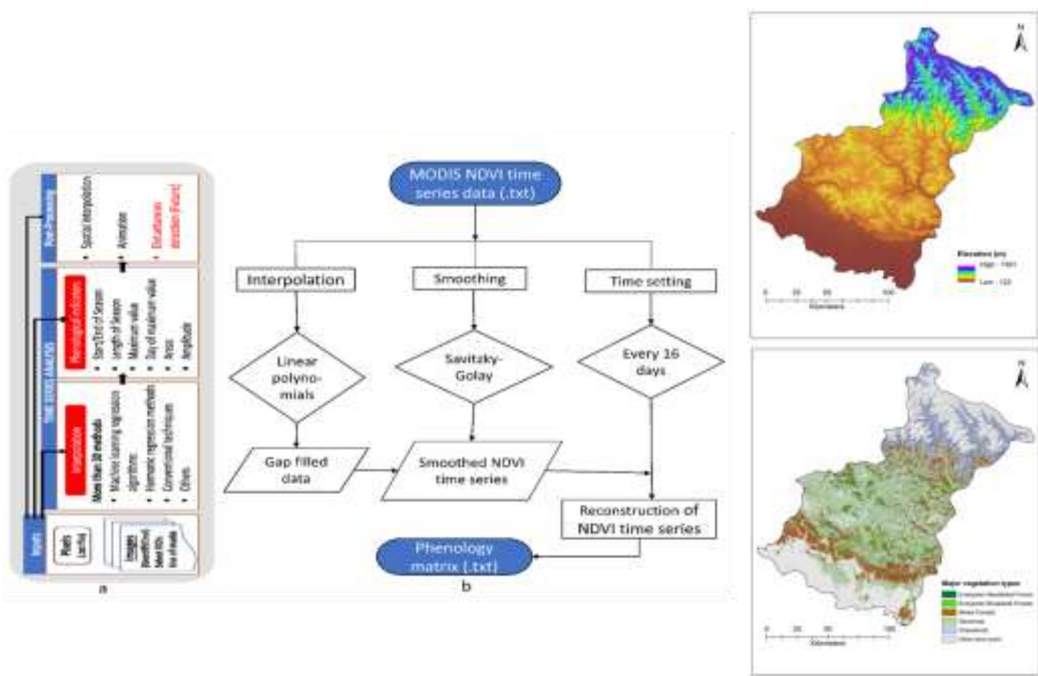


Figure 26 Impact of climate change on the phenology of major forest types in Kumaon Himalaya

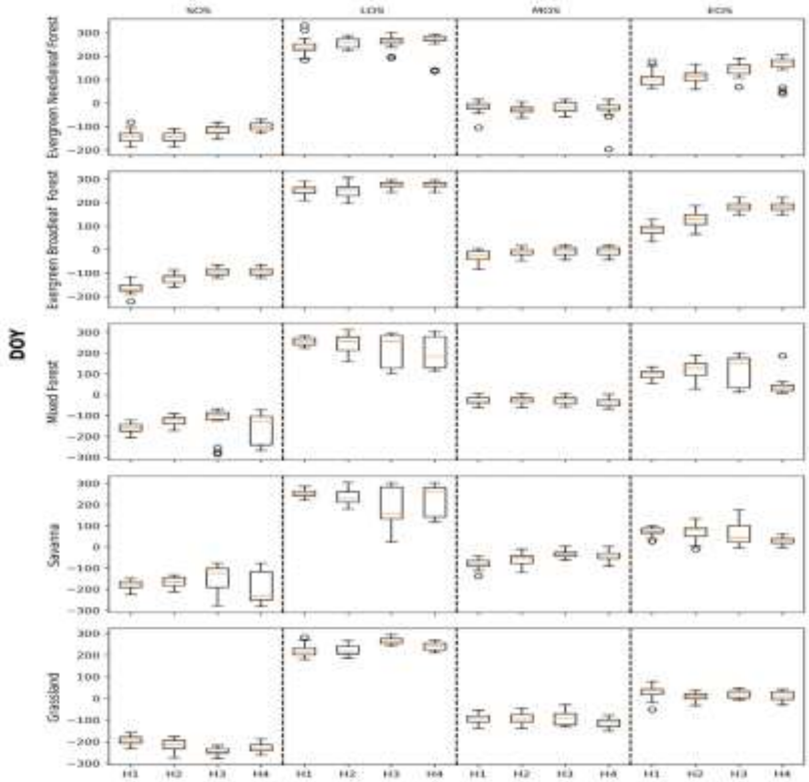


Figure 27 Phenometrics distribution along the elevation gradient across the vegetation types

The impact of climate change on the phenology of major forest types in Kumaon Himalaya is depicted in Figure 26. Figure 27 depicts the phenometric distribution along the elevation gradient across the vegetation types. The positive and negative values indicate starting in the current year and ending in the next year respectively. Figure 28 and 29 depict the Spatial distribution of field data values (SPAD & Soil Moisture) of *Parrotiopsis jacquemontiana* and Viburnum grandiflorum respectively.

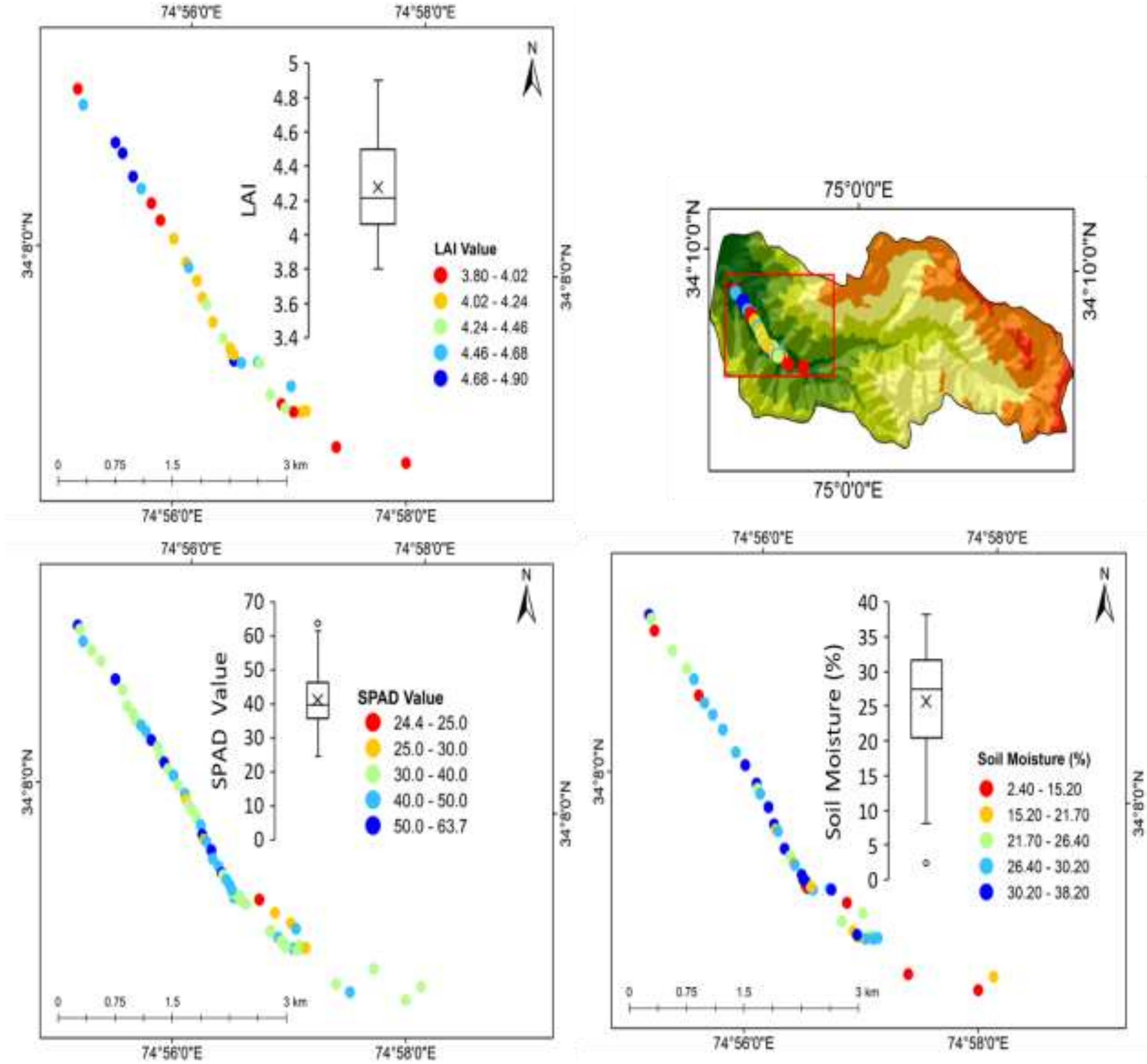


Figure 28 Spatial distribution of field data values (LAI, SPAD & Soil Moisture) of Parrotiopsis jacquemontiana

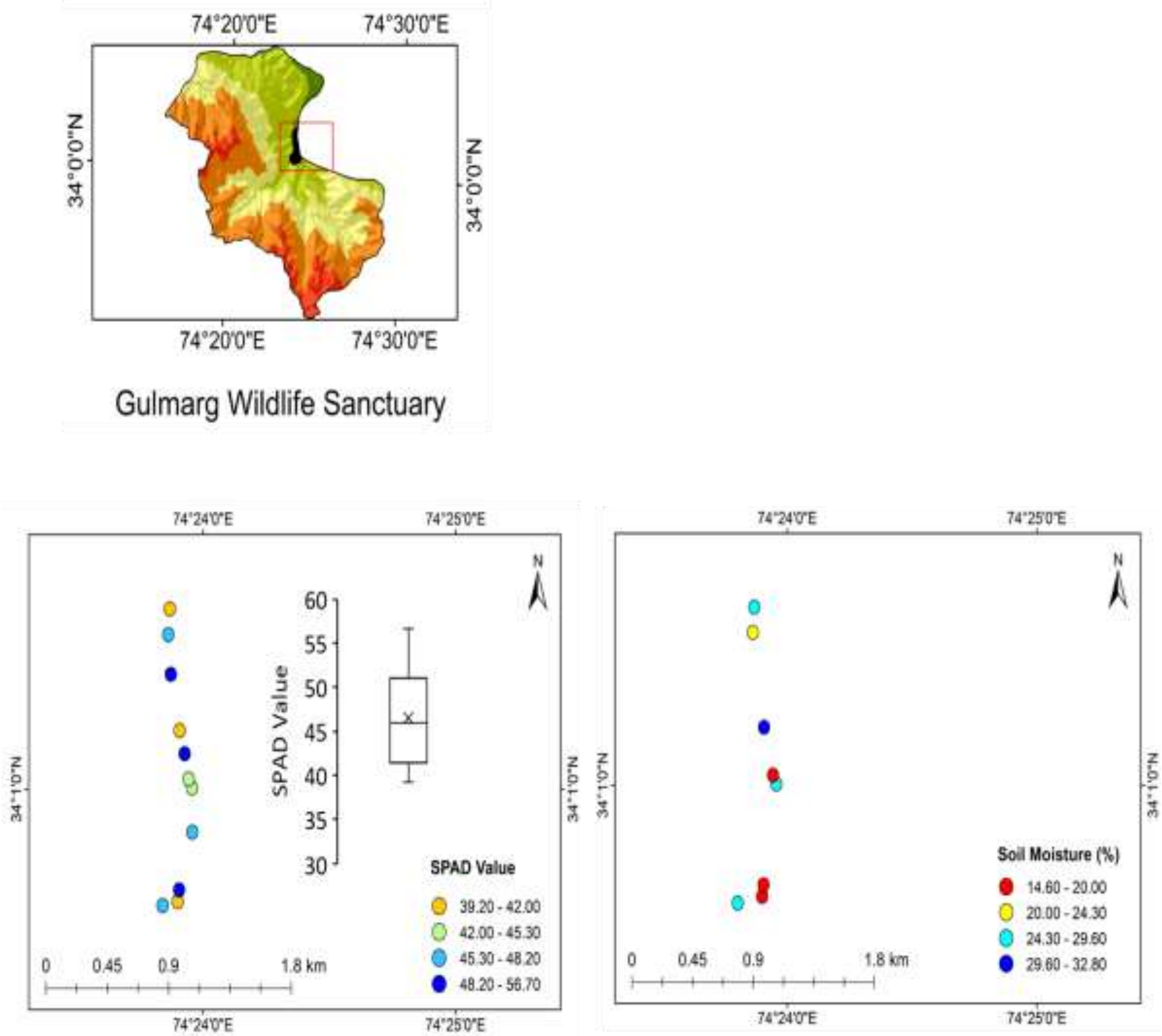


Figure 29 Spatial distribution of field data values (SPAD & Soil Moisture) of Viburnum grandiflorum

Investigated the effect of climate change on the start of the season (SOS) and end of the season (EOS) in major forest types (along the elevation gradient) of the Kumaon region. SOS indicates a delayed trend along the elevational gradient till mid-latitude and shows an advancing pattern thereafter. Winter warming, winter precipitation, and hot and dry spring played a crucial role in the delay of SOS. Winter warming and winter precipitation push EOS days further due to high winter temperatures and low winter precipitation. Due to winter warming and summer dryness, despite a warming trend in pre-season or springtime, the onset of the vegetation growth cycle shows a delayed trend across the vegetation types.

Objective 4: Fine-scale space-time map of the high-value species

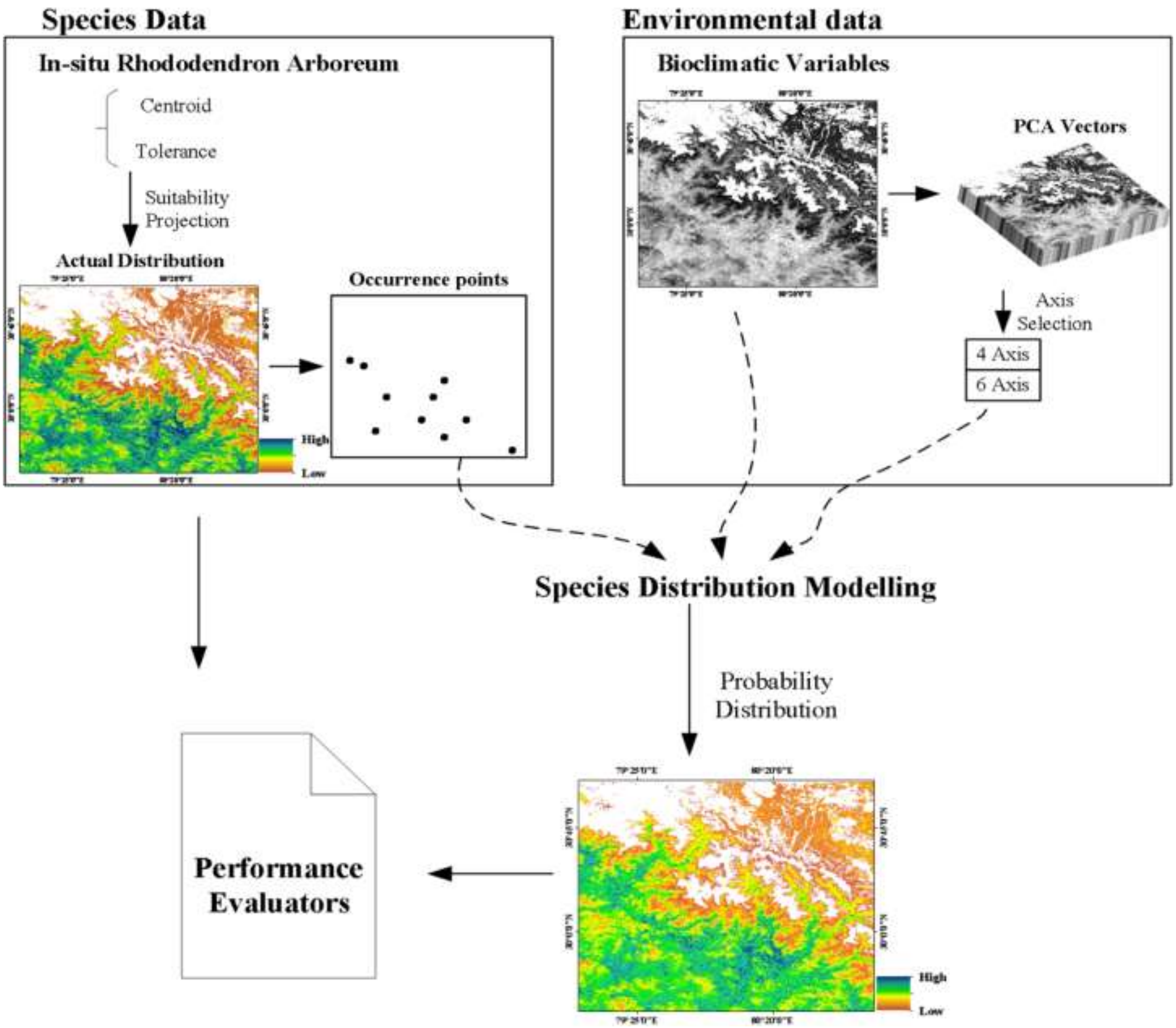


Figure 30 Flow diagram of the conventional approaches for species distribution modelling

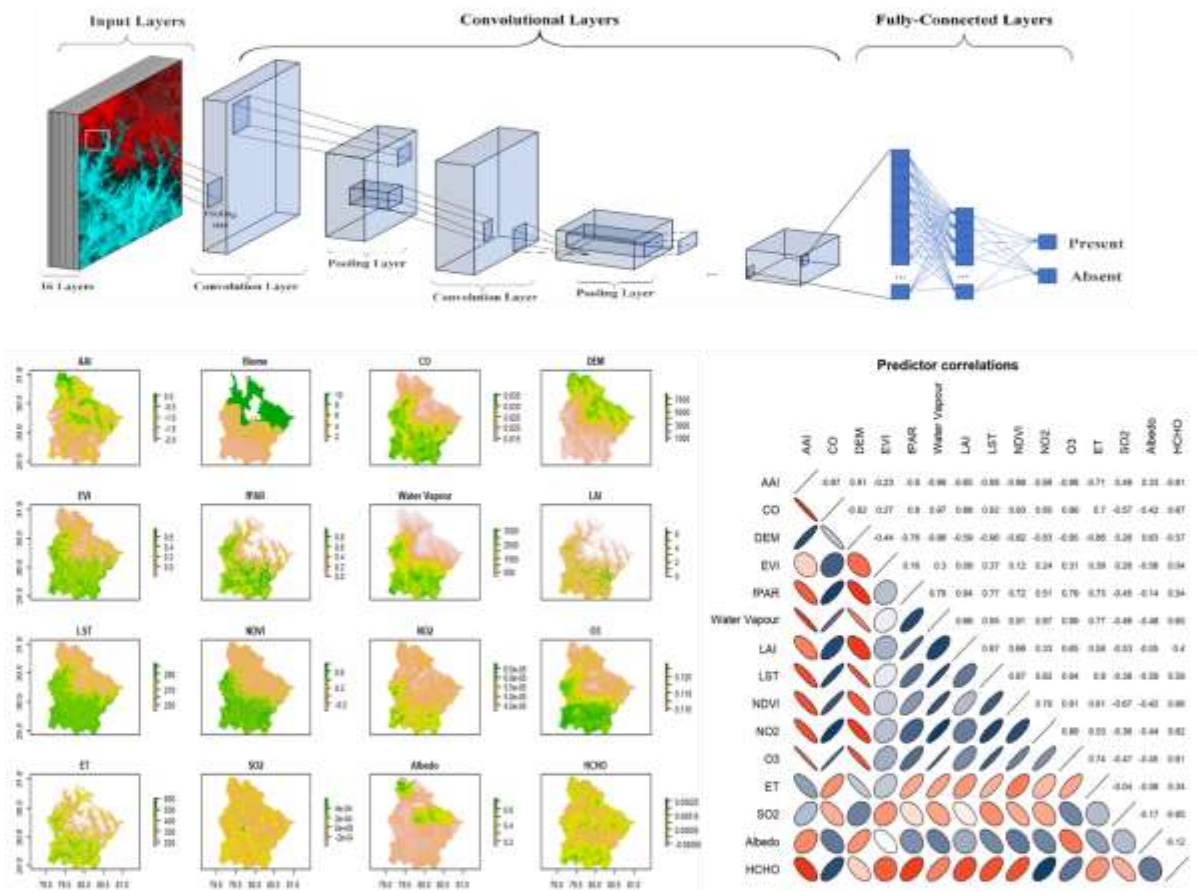


Figure 31 CNN architecture, Spatial Distribution of various input parameters and square correlation matrix between input parameters.

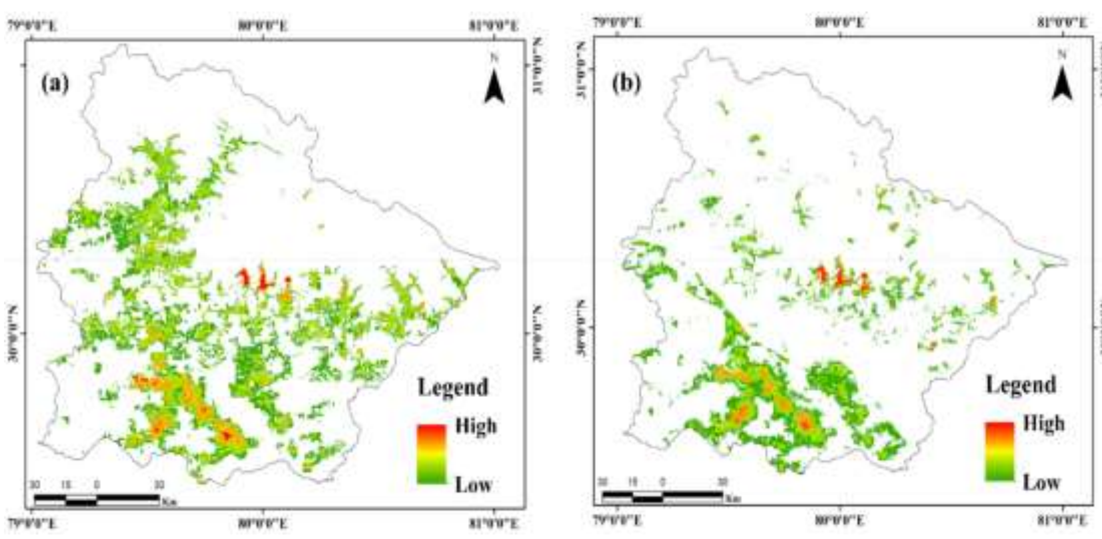
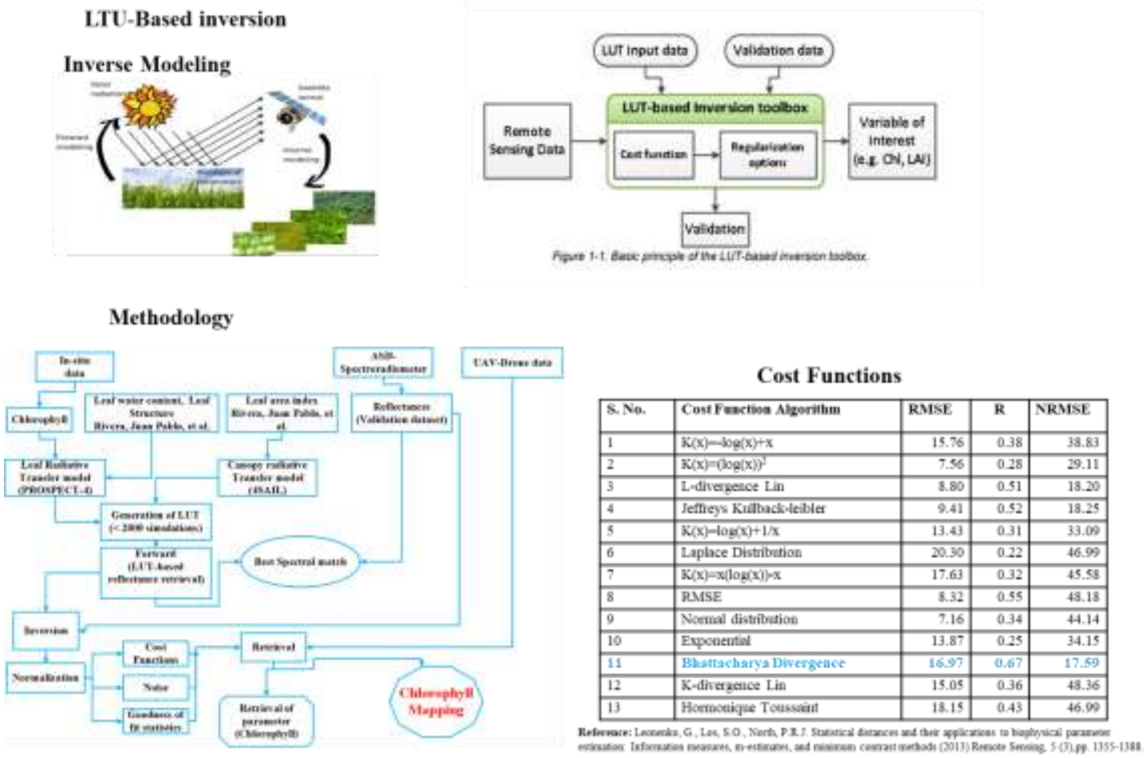


Figure 32 Probability distribution of Rhododendron arboreum using (a) BIOCLIM and (b) CNN Models.

As part of this objective, a novel approach towards establishing a CNN architecture and testing the performance of CNN in SDM and its comparison with other well-established SDMs namely, BIOCLIM and Maxent (Anand et al., 2022) was conducted. This study was conducted on the foothills of the Himalayas, where the altitudinal variation is very drastic and varies from 416 to 7801 m above mean sea level. This high-altitude Himalayan range constitutes a heterogeneous ecosystem and is home to many rare/endangered, medicinally, and economically important plant species. One of the major economically and medicinally important plant species, Rhododendron arboreum, was tracked and mapped in this study using different SDMs. Based on its occurrence and several ecological and bioclimatic satellite-based observations, the probability distribution of the Rhododendron arboreum was established. The CNN-based probability distribution model outperformed the presence-only-based BIOCLIM model with an AUC score of 0.917. The CNN-based prediction was also found to be more precise and accurate and with significantly less overestimation, whereas the AUC values of the BIOCLIM model were found to be 0.68 with a high overestimation. The superiority of CNN implies the role of nonlinear parameters in predicting the probability of species distribution. The scalability of the current solution on a global scale, the addition of some other important parameters, and an ensemble of all the available SDMs need to be explored in future work. An increase in the presence of the Rhododendron species is an indication of strong soil retention, which, in turn, is fruitful for other vegetation to grow and flourish. Apart from this, an increase in the green vegetation fraction and a decrease in shade fraction were found to be associated with a higher likelihood of Rhododendron. This increased likelihood of using the models would offer researchers an opportunity to understand the vegetation distribution and to contribute to the restoration of the ecology and biodiversity conservation in the protected areas so that a provision could be established for sustainable ecosystem services.



S. No.	Cost Function Algorithm	RMSE	R	NRMSE
1	K(x)=-log(x)+x	15.76	0.38	38.83
2	K(x)=(log(x))^2	7.56	0.28	29.11
3	L-divergence Lin	8.80	0.51	18.20
4	Jeffreys Kullback-leibler	9.41	0.52	18.25
5	K(x)=log(x)+1/x	13.43	0.31	33.09
6	Laplace Distribution	20.30	0.22	46.99
7	K(x)=x(log(x))-x	17.63	0.32	45.58
8	RMSE	8.32	0.55	48.18
9	Normal distribution	7.16	0.34	44.14
10	Exponential	13.87	0.25	34.15
11	Bhattacharya Divergence	16.97	0.67	17.59
12	K-divergence Lin	15.05	0.36	48.36
13	Hormonique Toussaint	18.15	0.43	46.99

Reference: Leonenko, G., Los, S.O., North, P.R.J. Statistical distances and their applications to biophysical parameter estimation: Information measures, m-estimates, and minimum contrast methods (2013) Remote Sensing, 5 (3), pp. 1355-1388.

Figure 33 Flowchart for Chlorophyll Retrieval and Mapping through UAV data

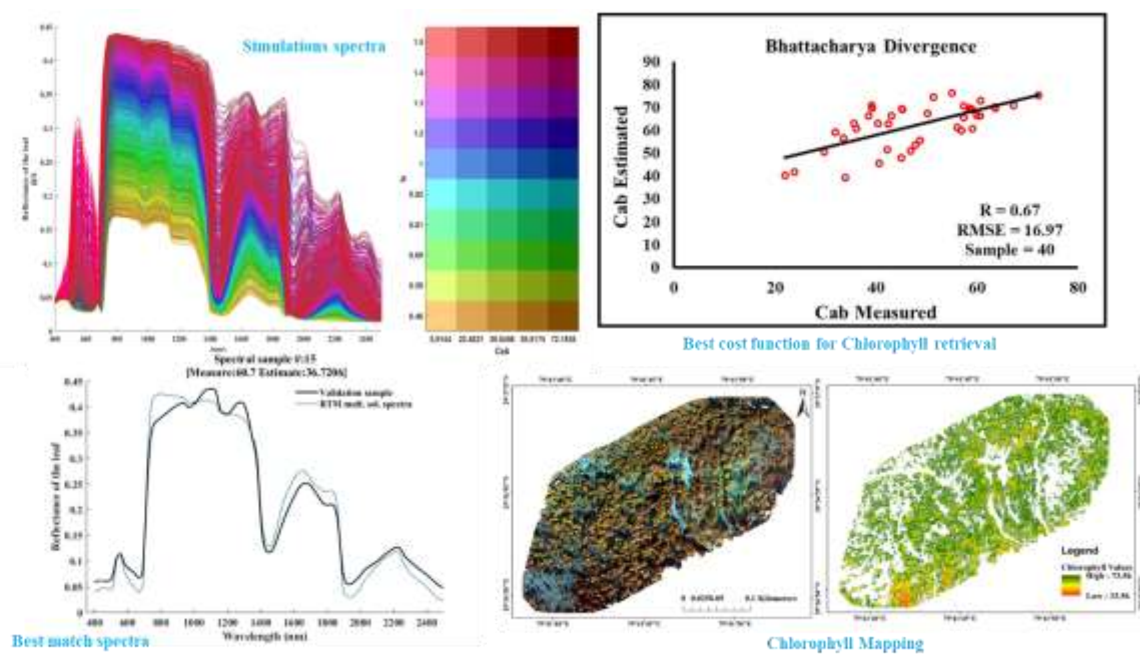


Figure 34 Chlorophyll Retrieval and Mapping through UAV data

Objective 5: Fine-scale space-time map of the high-value species

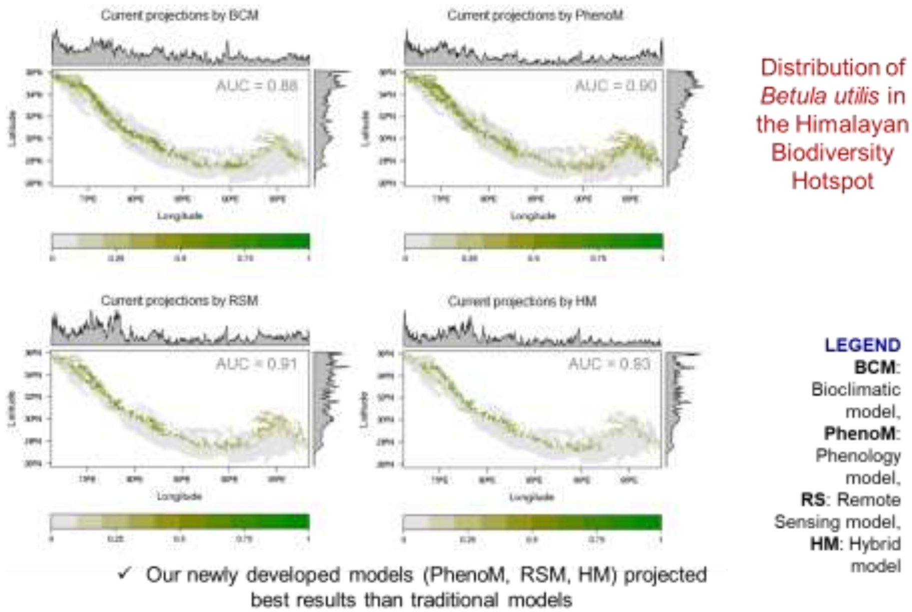


Figure 35 Distribution of Betula utils in Himalayan Biodiversity Hotspot

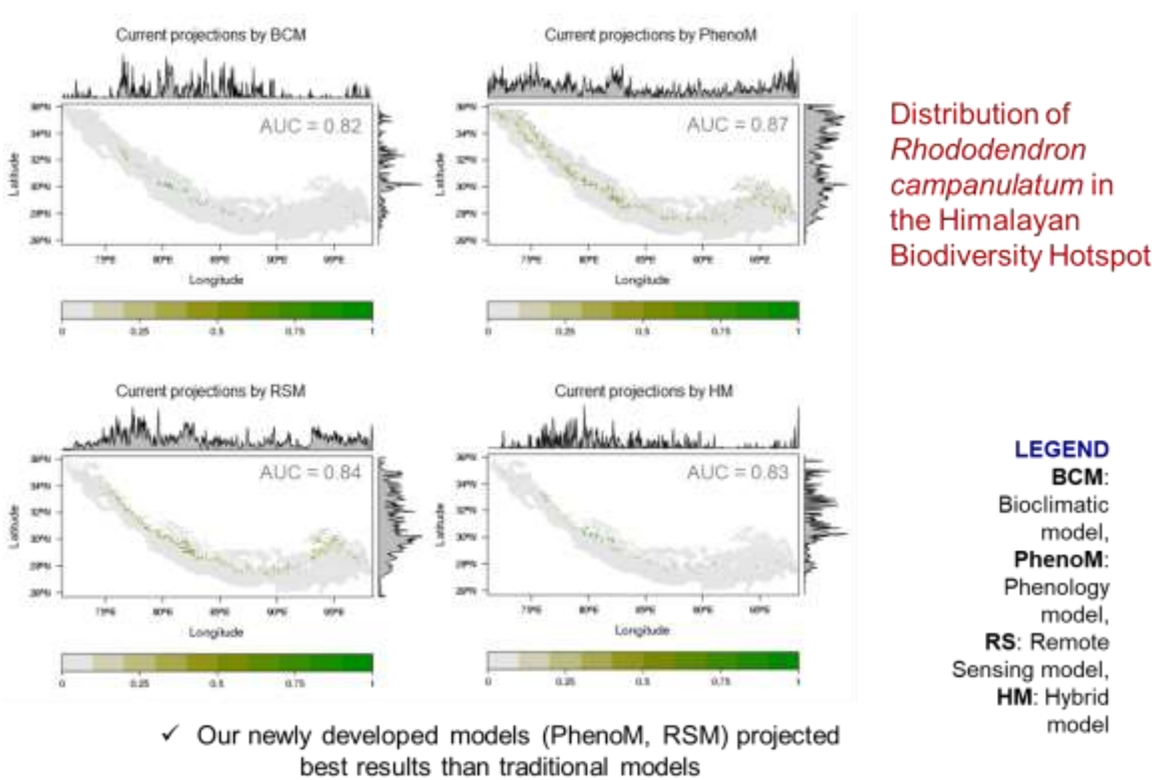


Figure 36 Distribution of Rhododendron campanulatum in Himalayan Biodiversity Hotspot

Table 12 Model evaluation statistics

Model	*Betula utilis*			*Rhododendron campanulatum*		
	AUC	TSS	Accuracy	AUC	TSS	Accuracy
BCM	0.9	0.62	0.75	0.85	0.67	0.95
PhenoM	0.91	0.68	0.77	0.87	0.62	0.89
RSM	0.91	0.68	0.88	0.84	0.64	0.89
HM	0.93	0.72	0.78	0.83	0.62	0.94

This investigation unfolded that our NEOVs (Satish, K.V., Dugesar, V., Pandey, M.K., Srivastava, P.K., Pharswan, D.S., & Wani, 2022; Satish, 2022) have the competency to model the reliable distribution of species over mountain ranges. Therefore, we strongly recommend integrating them into mountain species distribution modelling. We recommend researchers explore NEOVs with various variable combinations along with prior knowledge of species autecology for realistic modelling of Himalayan species. Projections of PhenoM, BiophyM, and HM showed close accordance with actual distributions and field experts' knowledge. This research finds that the integration of EO variables into SDMs undoubtedly improves SDMs. NEOVs developed in this study can be freely accessed from BCCVL and figshare. The moderate, high, and very high confidence classes and niche overlap maps could be prioritized for conservation applications and ecological investigations. We argue that building models with explicit species data and high-resolution spatiotemporal NEOVs could provide new opportunities for modelling the realized niches of species. Eventually, these maps can be validated with substantial ground data and spectral species maps developed from imaging spectroscopy. For the conservation and management of Himalayan Forest ecosystems, these precise understandings are necessary.

Objective 5: Projection of future climate change scenarios and its possible effect on the distribution of medicinal plants, rare, endangered and other economically important plants species.

In this study, we examine bias-corrected, statistically downscale models drawn from the NASA, Earth Exchange Global Daily Downscaled Projections - Coupled Model Intercomparison Project Phase 5 (NEX-GDDP-CMIP5)

CMIP5 Models	Institution, Country	Atmospheric Resolution	NEX-GDDP resolution
1-Geophysical Fluid Dynamics Laboratory Climate Model, version3 (GFDL-CM3)	National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, U.S.A	2.5º X 2º	0.25°X 0.25°
2-Institute of Numerical Mathematics Coupled Model, version 4.0 (INMCM-4)	Institute of Numerical Mathematics, Russia	2°X1.5º	0.25°X 0.25°
3-Max Plank Institute Earth System Model, low resolution (MPI-ESM-LR)	Max Plank Institute for Meteorology, Germany	1.875°X1.8653º	0.25°X 0.25°
4-Meteorological Research Institute Coupled Atmosphere–Ocean General Circulation Model, version 3 (MRI-CGCM3)	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, Japan	1.125°X1.1215º	0.25°X 0.25°
5-The second–generation Canadian Earth System model (CanESM2)	Canadian Centre for Climate Modelling and Analysis, Canada	2.8125ºX2.7906º	0.25°X 0.25°
6-Geophysical Fluid Dynamics Laboratory Earth System Model with Modular Ocean Model, version 4 (GFDL-ESM2M)	National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, U.S. A	2.5°X 2.0225º	0.25°X 0.25°

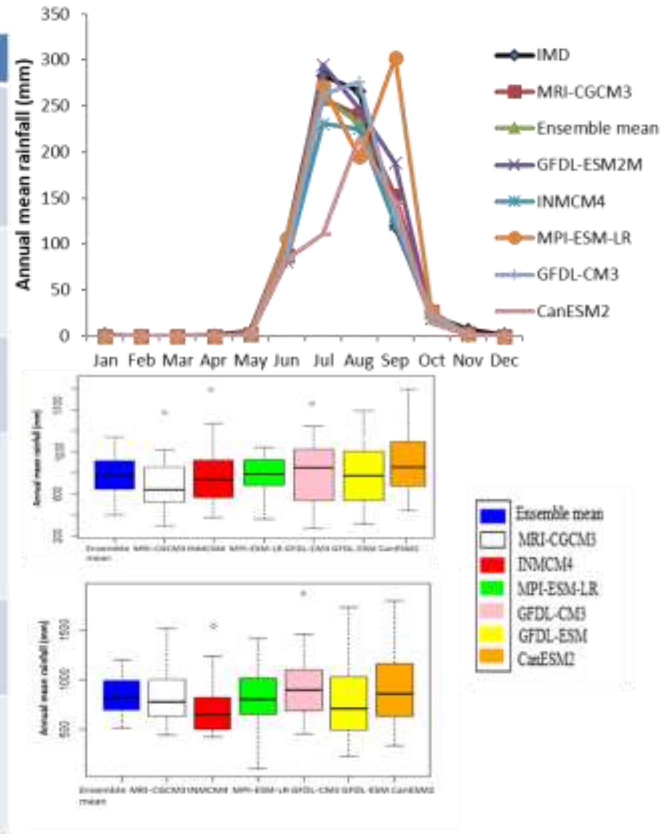


Figure 37 Bias corrected, statistically downscale models

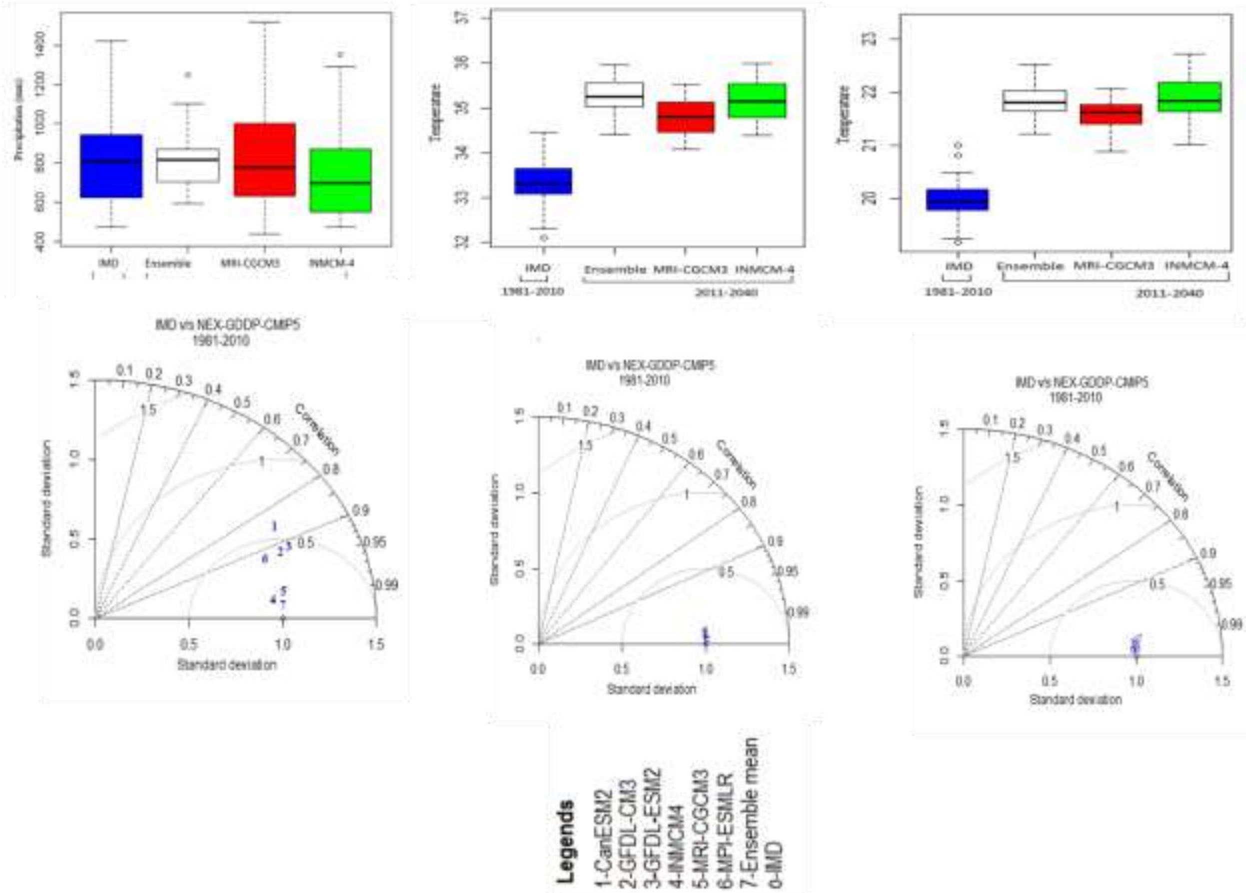


Figure 38 Model evaluation statistics

Objective 6: Traditional knowledge among communities residing in alpine regions of Western Himalaya

For this objective, Questionnaire-based surveys gathered information on demographic profile, biodiversity, and traditional knowledge in villages: Gunji, Kuti, Napalchyu, Navi and Kutti of Byans Valley Pithoragarh. 75 informants of different age groups from five villages were selected. One of the villages is shown in Figure 39. Perceptions were recorded for 31 high-value medicinal and economical plants (belonging to 24 families). Maximum high-value species (23) were herbs, followed by shrubs (5) and trees (3). *Aconitum heterophyllum, Allium stracheyi*, *Angelica glauca*, *Arnebia benthamii*, *Dactylorhiza hatagirea*, *Fagopyrum sculentum* , *Fritillaria roylei* , *Hippophae salicifolia, Paris polyphylla*, *Picrorhiza kurrooa, Podophyllum hexandrum , Polygonatum verticillatum* were high medicinal plant species. Some of these are represented in Figure 41.

Figure 39 One of the villages during the study

Yartsa Gunbu (*Ophiocordyceps sinensis*) the Caterpillar Fungus came out as the highest economically important species with its price of rupees 15,00000/-KG which is represented in Figure 40.

Figure 40 Ophiocordyceps sinensis

Figure 41 Medicinal Species

It was found that 77% of respondents claimed that the medicinal plants are decreasing, 17% claimed that there is no change, and 6% claimed an increasing trend. 8% of the respondents said they cultivate medicinal plants on their agricultural land, while the remaining 92% harvest from the wild population for their household use and commercial trading. When the community people were asked will they prefer the cultivation of medicinal plants in their fields if any financial assistance provided, 45% were agreed, while 55% refused and said that it is not beneficial for that, rather than cultivating medicinal plants, they preferred to go for *Ophiocordyceps sinensis collection* because it is highly economical value. Community people do not practice any particular exercise for biodiversity conservation, although a few respondents said that during recent years due to road construction a large proportion of the land has been degraded, which has led to biodiversity and habitat loss. The responses could be represented pictorially in Figure 42

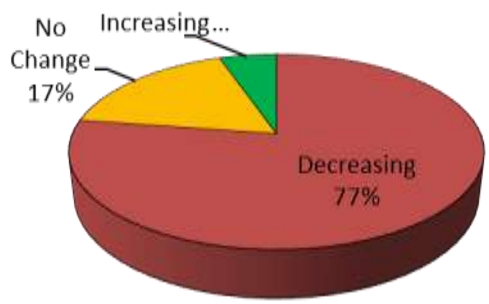


Figure 42 Hippophae salcifolia and Ilex dipyrena

Modelling Rhododendron arboreum distribution using CNN: In the process of understanding the impact of biogeoclimatic variables on particular species, we designed a CNN architecture to model the distribution of *Rhododendron arboreum*. We found that the CNNs are better at capturing the non-linear pattern using the combination of different satellite observations and they outperformed the conventional SDM approaches like BIOCLIM. Further, we found that an increase in the presence of the Rhododendron species is an indication of strong soil retention, which at higher altitudes can support other vegetation to grow and flourish. We also found that an increase in the green vegetation fraction and a decrease in shade fraction were associated with a higher likelihood of Rhododendron. After analyzing the climate variables, we found that there is a drastic change in climate patterns in the last couple of decades. The impacts of climate change are visible on the total carbon stock of the region which is gradually shifting to the higher latitudes.

Retrieval of biochemical and biophysical parameters: For estimating medicinal (biochemical) properties via a nondestructive method. Filtering, smoothing, and feature selection was applied to hyperspectral spectra. The average mean filter in combination with the first derivative on hyperspectral reflectance-based radiometer data produced the best for Taxol spectral indices, model development, and eventual Taxol content prediction. The best model showed a correlation of 0.719 with a relative root mean square error (RMSEr) value of 0.678 for taxol content prediction. Biophysical parameters including elevation, land surface temperature, and soil moisture along with other biochemical parameters like chlorophyll, carotenoids, and polyphenols were measured for the T. wallichiana ecosystem. By applying Principal Component analysis to the above-measured variables, the most sensitive parameters of the ecosystem defining the habitat of T. wallichiana were identified.

Other species like Pyracantha crenulata were also studied for their biochemical properties, i.e., anthocyanin. P. crenulata anthocyanin content was measured destructively and various developed indices for anthocyanin were tested to validate the most accurate indices to calculate anthocyanin in the case of P. crenulata.

Phenology assessment: This study disclosed that due to winter warming and summer dryness, despite a warming trend in pre-season or springtime, the onset of the vegetation growth cycle shows a delayed trend across the vegetation types of Kumaon Himalaya. The SOS show a delayed trend of 1.3, 2.9, and 1.75 days/100m in evergreen needle leaf forest (ENF), evergreen broadleaved forest (EBF) and mixed forest (MF) respectively and advancement of 4, 3.6, 1.4 days/100m in MF, Savanna and Grassland respectively. Similarly, the EOS show a delay of 1.8, 3.8, and 4.36 respectively in ENF, EBF and MF and show an advancement of 1.6, 1.5, 0.74 days/100m in MF, Savanna and Grassland respectively.

- ***Mapping potential distributions of timber and tree line species:*** EOVs-based models yielded realistic distributions compared to bioclimatic models. The study finds that there is a moderate variation in the niche space between *R. campanulatum* and *B. utilis*.
- ***Climate change projections of treeline species:*** This investigation unfolded that the optimal climatic niche of R. campanulatum likely to undergo substantial squeeze over the northern-end and the central Himalayas under emission scenarios. A few optimal climatic niches are likely to remain intact over Himachal and Uttarakhand and keep hopes for R. campanulatum conservation under climate change.

4.2 Key Results

- ***Modelling Rhododendron arboreum distribution using CNN:*** This research found CNN based species distribution modelling approach is outperforming other conventional methods like Maxent and BIOCLIM.

- ***Estimation and Generation of Carbon Stock Product:*** The dense patches of carbon stocks are shifting to higher altitudes, and it is directly proportional to the increase in temperature, precipitation and melting of snow cover.

- ***Retrieval of biochemical and biophysical parameters:*** The models were developed for the retrieval of biochemical (medicinal molecule) parameters of Taxus wallichiana -Novel Taxol indices are generated by identifying various combinations from the spectroradiometer data between a spectral range of 350-2500 nm. Taxol content can be allocated using reflectance values in indices format from the visible range of 415-421 +/- 5 nm window. PCA explains the most variance-causing factor among the crucial biophysical and biochemical parameters to bring out the most suitable habitat. PC 1 also shows moderate loadings for the measured taxol content (> 0.70) and elevation (>0.65), thus highlighting that taxol and elevation are the most variables among the measured variables. Results on *P. crenulata* disclosed that a modified anthocyanin content index (mACI) was found most appropriate with a correlation of 0.606 at wavelengths of 1368 nm and 816 nm for *P. crenulata.*
- ***Phenology assessment:*** This study investigated the effect of an elevational gradient, temperature and precipitation on the start of the season (SOS) and end of the season (EOS) in major forest types of the Kumaon region of the western Himalaya using MODIS NDVI time series data (2001-2019).

- ***Mapping potential distributions of timber and tree line species:*** Accurately mapped niches of B. utilis and R. campanulatum over Himalaya. This study developed novel earth observation variables (EOVs) for the enhancement of species distribution modelling. Models including EOVs were competitive for study species, and models without EOVs had considerably poor model performance and explanatory strength. Among the three machine learning algorithms tested (artificial neural networks, generalised boosting model, and maximum entropy), maximum entropy produced the most promising predictions for BCM, PhenoM, RSM, and HM.

- ***Impacts of climate change on treeline species:*** Results unfolded that the distribution of R. campanulatum is governed by annual temperature range, the minimum temperature of the coldest month and the precipitation of the warmest quarter. This study found that there is an upward and downward climatic niche shift in the future and currently optimal climate is expected to become unfavourable under emission scenarios and vice versa.

4.3 Conclusion of the study

Modelling Rhododendron arboreum distribution using CNN: This study provided a novel approach towards establishing a CNN architecture and testing the performance of CNN in SDM and its comparison with other well-established SDMs namely, BIOCLIM. The superiority of CNN implies the role of nonlinear parameters in predicting the probability of species distribution. The scalability of the current solution on a global scale, the addition of some other important parameters, and an ensemble of all of the available SDMs need to be explored in future work. This increased likelihood of using the models would offer researchers an opportunity to understand the vegetation distribution and to contribute to the restoration of the ecology and biodiversity conservation in the protected areas so that a provision could be established for sustainable ecosystem services.

Retrieval of biochemical and biophysical parameters: Overall, research indicates that hyperspectral data provides better and more accurate indices for determining the biochemical properties of medicinal plants. Absorption-based hyperspectral indices are useful in identifying the bands, but reflectance-based hyperspectral indices are more accurate in predicting taxol content non-destructively. The involvement of the Multivariate technique and the Seriation technique in the crucial parameters for any habitat is a combination of factors rather than a single factor.

Phenology assessment: The present study unfolded those disparities in phenology changes caused by climate change in major vegetation types of Himalayan Forest ecosystems. This study highlighted the necessity for rigorous research into plant phenology in understudied Himalayan forests. New understandings and knowledge produced in the study could assist climate change impact mitigation actions in the Kumaon Himalaya.

Mapping potential distributions of timber and tree line species: This investigation unfolded that our newly developed EOVs have the competency to model the reliable distribution of species over mountain ranges. Therefore, we strongly recommend integrating them into mountain species distribution modelling. We recommend researchers explore EOVs with various variable combinations along with prior knowledge of species autecology for realistic modelling of Himalayan species. Among the different confidence level classes of projections, moderate, high, and very high classes, and niche overlapping maps can be given priority for ecological and conservation studies.

We argue that model building with accurate unbiased species data and improved spatial resolutions of EOVs could open new windows to model species realized distribution maps, and such attempts are crucial for precise conservation and management of Himalayan Forest ecosystems.

Climate change projections of treeline species: The study proposed that climatic niche contraction grids be taken on a high-priority basis to minimize anthropogenic pressures on the habitat of *R. campanulatum*. The expansion grids are to be prioritized for ecological and climate change impact studies such as upward shift recruitments. Drastic shrinkage of very high and high confidence should be prioritized for threat mitigation, habitat rejuvenation, and restoration. Priority might be given to newly established suitable climatic spaces for conservation and reintroduction projects.

5 OVERALL ACHIEVEMENTS

5.1 Achievement on Project Objectives. Defining contribution of deliverables in overall Mission

S. No.	Quantifiable Deliverables*	Monitoring Indicators*	Quantified Output/ Outcome achieved
1.	Spectral library and standardization for selected medicinal and threatened species;	Periodic, regular updates on Spectral Library of the economically important plant species (Nos)	**SHRIME** (Spectral Tool of Himalayan Rare, Invasive, Medicinal, and Economical Plant Species) tool is live and running.
2.	Spectral and image analysis for discrimination of 4 selected economically important plant species;	No. of New Database/ Datasets generated on the identified dynamics, i.e., biochemical, and biophysical properties of selected species.	Apart from DB created in terms of the SHRIME tool, where the spectral signature of more than 250 species are captured, there are five species whose dynamics in terms of biochemical and biophysical properties are captured. The species are *Betula utilis*

			D.Don, *Rhododendron campanulatum, Taxus wallichiana, Rhododendron arboretum and Pyracantha crenulata D.Don.*
3.	Model for retrieval of Biochemical and Biophysical properties of selected species (04 No.);	Fine-scale space-time distribution maps of economically important species in the target area	Fine-scale space-time distribution maps were generated for the species *Betula utilis D.Don*, *Rhododendron campanulatum, Taxus wallichiana, Rhododendron arboretum and Pyracantha crenulata D.Don.*
4.	Projection and scenario for distribution of selected species (4 No.);	Assessment report/ manual on safeguarding the medicinal and threatened plant species.	The species distribution limits of two alpine treeline species, namely Betula utilis D.Don and Rhododendron campanulatum over the Himalayan biodiversity hotspot were conducted. Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM),

			and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. This not only has helped in the assessment of the prevalent condition but would also help in the identification of the practices required for safeguarding the medicinal and threatened plant species.
5.	Fine-scale space-time map of the high-value species in the Himalayas (04 No.);	Policy/ Strategic frameworks/ drafts (No.) for assisting the regional decision-making in the identified area and dynamics.	Traditional knowledge among communities residing in alpine regions of Western Himalaya was conducted. Gunji, Kuti, Napalchyu, Navi and Kutti of Byans Valley Pithoragarh villages were taken for this study. 77% of respondents claimed that the medicinal plants are decreasing in their natural surroundings 92% of medicinal and economical importance plants harvest from the wild population for their household use and commercial trading 45% of respondents agreed with Govt. subsidy-based medicinal plant cultivation,

			while 55% refused to cite they would go for the cultivation of economical plant species than medicinal ones.
6.	Plan for management and sustainable harvest of species (01 No.) based on above.	Other Publications and knowledge products (Nos.)	8 publications in International Journal are published. 3 More publications are in pipeline.

5.2 Establishing New Database/Appending new data over the Baseline Data

- Developed SHRIME tool which provides a spectral library for Himalayan economic and medicinally important species. As such, no benchmark data was available to date. This would be the first of its kind baseline. This would help the upcoming researchers in adding more to the existing literature.
- Development of novel earth observation variables (Eco-physiological and phenological) for accurate niche modelling in the Himalayas. There exists few ecological or phenological data in past literature but this perspective of taking Eco-physiological and phenological variables for any study is explored for the first time.

The details of these are specified in Appendix 3 for reference.

5.3 Generating Model Predictions for different variables

- DB was created in terms of the SHRIME tool, where the spectral signature of more than 250 species are captured, there are five species whose dynamics in terms of biochemical and biophysical properties are captured.
- Fine-scale space-time distribution maps were generated for the species *Betula utilis D.Don*, *Rhododendron campanulatum, Taxus wallichiana, Rhododendron arboretum and Pyracantha crenulata D.Don.*
- The species distribution limits of two alpine treeline species, namely Betula utilis D.Don and Rhododendron campanulatum over the Himalayan biodiversity hotspot were conducted. Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM),

biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. This not only has helped in the assessment of the prevalent condition but would also help in the identification of the practices required for safeguarding the medicinal and threatened plant species.

- Retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species *Taxus wallichiana*.
- For the retrieval of biochemical parameters, various indices were generated and three different types of filters, Average Mean, Savitzky Golay, and Fast Fourier Transform (FFT) one at a time followed by feature selection on each denoised spectra were applied to smooth the spectra for better estimation of biochemical parameters.
- The development and validation of hyperspectral indices for anti-cancerous Taxol content estimation were conducted as part of this objective.
- Further, Taxol Content Variation was also explored in altitudinal variation in the Himalayan regions.
- The species distribution of Taxol content through empirical methods and radiative transfer models (physical methods) were conducted. Taxol Mapping was also explored through drone quasi-hyperspectral images.
- As part of this objective, a list of tentative compounds identified in *Rhododendron arboreum* using UPLC and HRMS was prepared.
- Fine-grained species distributions were modelled using Sentinel-2, Sentinel-5P, Moderate Resolution Imaging Spectroradiometer (MODIS), ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) and Shuttle Radar Topography Mission (SRTM) along with other climatic variables over the sampling regions.
- Several machine learning algorithms were incorporated to establish the relation between physical and climatic variables to estimate the probability distribution of species.
- The phenology assessment was carried out for two decades to the effect of an elevational gradient, temperature, and precipitation on the start of the season (SOS) and end of the season (EOS) in major forest types of the Kumaon region of the western Himalaya using MODIS NDVI time series data (2001-2019).
- The potential distribution of *Rhododendron arboreum*, a medicinal plant species found within the foothills of the Himalayas was also evaluated using Conventional Species Distribution Models (SDM) namely BIOCLIM and Maxent, a Machine Learning variant as well as Convolutional Neural Network (CNN).

5.4 Technological Intervention

NA

5.5 On field Demonstration and Value-addition of Products

NA

5.6 Promoting Entrepreneurship in IHR

NA

5.7 Developing Green Skills in IHR

The project has trained many of the resources from the technological perspective as well as from the out of box thinking. This would help in building capacities, creating the domestic framework and international architecture for quick diffusion of cutting-edge climate technology in India and joint collaborative R&D for such future technologies.

5.8 Addressing Cross-cutting Issues

India’s commitment to achieving five nectar elements (Panchamrit) as their climate action plan. The outcomes of the current project would help propagate a healthy and sustainable way of living based on traditions and values of conservation and moderation, through the motto ‘LIFE’– ‘Lifestyle for Environment’ as a key to combating climate change".

The outcomes would also help the policymakers in better adaptation to climate change by enhancing investments in development programmes in the Himalayan region by identification of prone areas. Carbon maps derived from remote sensing are widely used by scientists and policymakers. The carbon stock assessment undertaken as part of this award contributes to the Paris Agreement's goals (2015). The biophysical variables and SDMs developed in this grant contribute to the GEOBON Essential Biodiversity Variables (EBVs).

6 PROJECT’S IMPACTS ON IHR

6.1 Socio-Economic Development

- Awareness of the Himalayan rich heritage in terms of biodiversity would help the natives in building sustainable solutions in terms of Farming Medicinal and Economical Plant species that not only would be a revenue generator.

- Formation of Self groups to locally create a revenue model in terms of catering for the huge supply of medicinal, and economical requirements.

- Creating the resource pool that can identify the plant species and their mapping for fulfilling the necessities and improving the socio-economic conditions.

6.2 Scientific Management of Natural Resources In IHR

- The project has trained many of the resources from the technological perspective as well as from the out of box thinking. This would help in building capacities, creating the domestic framework and international architecture for quick diffusion of cutting-edge climate technology in India and joint collaborative R&D for such future technologies.

6.3 Conservation of Biodiversity in IHR

- The Species Distribution Models (SDMs) studies carried out in this grant contribute to SDG 15 and are crucial for biodiversity conservation in the Himalayas. The output data can be utilised for conservation, restoration activities and sustainable utilisation of resources.
- Mitigating the effects of climate change on treeline species is one of the most important global ecological problems. Future projections of climatic species niche changes constitute the backbone for impact mitigation and adaptive management strategies in the Himalayas in response to short- and long-term changes. This attempt gives insight concerning SDG 13.
- The plan for management and sustainable harvest of species provided in this project is contributing to SDG 12 (Sustainable Consumption and Production Patterns), more clearly to 12.2
- Biophysical variables are crucial in understanding ecosystem functioning and processes, and thus have a key role in biodiversity conservation. Monitoring of biophysical parameters in the present grant contributes to SDG 15.4 and SDG 15.1.

6.4 Protection of Environment

- India’s commitment to achieving five nectar elements (Panchamrit) as their climate action plan. The outcomes of the current project would help propagate a healthy and sustainable way of living based on traditions and values of conservation and moderation, through the motto ‘LIFE’– ‘Lifestyle for Environment’ as a key to combating climate change".
- The outcomes would also help the policymakers in better adaptation to climate change by enhancing investments in development programmes in the Himalayan region by identification of prone areas.
- Increase the green cover by developing the agro-forestry system. This not only would provide a check on landslides but also help in growing organic content.

6.5 Developing Mountain Infrastructures

- The deliverables of the current work would be utilized in creating natural ecosystems in IHR that not only would be sustainable but also would engage the natives in developing the rich heritage of the Himalayas in terms of Medicinal and Economical Plant species.

6.6 Strengthening Networking in IHR

- The project has involved resources with skillset Ecology, Agriculture, Remote Sensing, Physics, Computer Science, Data Science and Earth Sciences. The deliverables of the current project have trained many of the resources from the technological perspective as well as from the out of box thinking. This would help in building capacities, creating the domestic framework and international architecture for quick diffusion of cutting-edge climate technology in India and joint collaborative R&D for such future technologies.

7 EXIT STRATEGY AND SUSTAINABILITY

7.1 How effectively the project findings could be utilized for the sustainable development of IHR

The deliverables of the current project would act as a baseline for the identification and retrieval of Himalayan Medicinal, Economical, Rare, Invasive Plant Species. The SHRIME tool developed as a part of this project would be the first of its kind initiative that will boost the research in the Himalayan Ecosystem. The models developed could be scaled across the Himalayan region for offering sustainable solutions for Himalayan Ecosystems under climatic uncertainties.

7.2 Efficient ways to replicate the outcomes of the project in other parts of IHR

The dataset would be utilized for researchers across the domains on request. The models published as part of various objectives are developed under open-source platforms and thus easily replicated by the researchers.

7.3 Identify other important areas not covered under this study needs further attention

Development of Climatic Network for better understanding of Information Loss or Gain between various parameters for informed decision at the micro level. The tools could be extended for many more functionalities in terms of Machine Learning capabilities.

7.4 Major recommendations for sustaining the outcome of the projects in future

- The Indian Himalayan Region (IHR) not only is complex in terms of geomorphology but also terms of reachability. This work is a genuine effort towards bridging the gaps and offering solutions in terms of sustainable Himalayan Ecosystems under climatic uncertainties. The offerings in terms of benchmarked SHRIME tool where the spectral library of various Himalayan Medicinal, Economical, Rare, Invasive Plant Species would be a giant leap to the researchers working in this domain. There is a great scope in enriching the potential of this tool in terms of informed decision-making to various challenges across IHR. Various models developed as part of deliverables would help offer customized and localized solutions to the natives as well as build a natural ecosystem amidst the growing concern of Climate Change.

- A mission-level effort is required to offer solutions from the amalgamation of private bodies, state, and central government offerings to solve the micro-level problems. The problems should be highlighted through socio-economic, climatic challenges and livelihoods. Short-term goals should be given for better results. A few of the offerings could be skill development, sustainable tourism, and cultivation in a phased manner according to shifting in the altitudes. Development of the Himalayan Agricultural Landscape is most important as far as medicinal and economical plant species are considered.

8 REFERENCES/BIBLIOGRAPHY

Anand, A., Pandey, M. K., Srivastava, P. K., Gupta, A., & Khan, M. L. (2021). Integrating multi-sensors data for species distribution mapping using deep learning and envelope models. *Remote Sensing*, *13*(16), 1–17. https://doi.org/10.3390/rs13163284

Anand, A., Srivastava, P. K., Pandey, P. C., Khan, M. L., & Behera, M. D. (2022). Assessing the niche of Rhododendron arboreum using entropy and machine learning algorithms: role of atmospheric, ecological, and hydrological variables. *Journal of Applied Remote Sensing*, *16*(04). https://doi.org/10.1117/1.JRS.16.042402

Brugière, D., & Scholte, P. (2013). Biodiversity gap analysis of the protected area system in poorly-documented Chad. *Journal for Nature Conservation*, *21*(5), 286–293. https://doi.org/10.1016/j.jnc.2013.02.004

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., … Watson, R. (2010). Global Biodiversity: Indicators of Recent Declines. *Science*, *328*(5982), 1164–1168. https://doi.org/10.1126/science.1187512

Dwivedi, T., Kanta, C., Singh, L. R., & Sharma, D. I. P. (2019). A list of some important medicinal plants with their medicinal uses from Himalayan State Uttarakhand, India. *Journal of Medicinal Plants Studies*, *7*, 106–116.

Gupta, A., Singh, P., Srivastava, P. K., Pandey, M. K., Anand, A., Chandra Sekar, K., & Shanker, K. (2022). Development of hyperspectral indices for anti-cancerous Taxol content estimation in the Himalayan region. *Geocarto International*, 1–17. https://doi.org/10.1080/10106049.2021.1983031

Gupta, A., Srivastava, P. K., Petropoulos, G. P., & Singh, P. (2021). Statistical Unfolding Approach to Understand Influencing Factors for Taxol Content Variation in High Altitude Himalayan Region. *Forests*, *12*(12), 1726. https://doi.org/10.3390/f12121726

Hilton-Taylor, C., Pollock, C. M., Mittermeier, R. A., & Brackett, D. A. (2000). *2000 IUCN red list of threatened species*.

IPCC. (2007). *Working Group II: Impacts, adaptation and vulnerability. Geneva: Intergovernmental Panel on Climate Change.*

Manish Kumar Pandey, P. K. S. (2022). SHRIME: Spectral tool of Himalayan Rare, Invasive, Medicinal and Economical Plant Species. *Communicated*.

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How Many Species Are There on Earth and in the Ocean? *PLoS Biology*, *9*(8), e1001127. https://doi.org/10.1371/journal.pbio.1001127

Pandey, P. C., Pandey, M. K., Gupta, A., Singh, P., & Srivastava, P. K. (2021). Spectroradiometry: Types, Data Collection, and Processing. In *Advances in Remote Sensing for Natural Resource Monitoring* (pp. 9–27). Wiley. https://doi.org/10.1002/9781119616016.ch2

Patrick ten Brink. (2011). *The economics of ecosystems and biodiversity in national and international policy making.*

Rands, M. R. W., Adams, W. M., Bennun, L., Butchart, S. H. M., Clements, A., Coomes, D., Entwistle, A., Hodge, I., Kapos, V., Scharlemann, J. P. W., Sutherland, W. J., & Vira, B. (2010). Biodiversity Conservation: Challenges Beyond 2010. *Science*, *329*, 1298–1303.

Samant, Pant, S., Singh, M., Lal, M., Singh, A., Sharma, A., & Bhandari, S. (2007). Medicinal plants in Himachal Pradesh, north western Himalaya, India. *International Journal of Biodiversity Science & Management*, *3*(4), 234–251. https://doi.org/10.1080/17451590709618177

Satish, K.V., Dugesar, V., Pandey, M.K., Srivastava, P.K., Pharswan, D.S., & Wani, Z. A. (2022). Seeing from space makes sense: Newly developed phenological and earth observation variables accurately map species distributions over Himalaya. *Communicated*.

Satish, K. V. et al. (2022). Ensemble machine learning models reveal good and bad news for Bell Rhododendron (Rhododendron campanulatum D. Don.) in Himalaya under climate change. *Communicated*.

UNFCC. (2015). *Paris agreement. In Report of the Conference of the Parties to the United Nations Framework Convention on Climate Change*.

9 ACKNOWLEDGEMENT

This work is funded by the National Mission on Himalayan Studies, G.B. Pant National Institute of Himalayan Environment (NIHE), Ministry of Environment, Forest & Climate Change (MoEF&CC), Government of India.

APPENDICES

Appendix 1 – Details of Technical Activities
Appendix 2 – Copies of Publications duly Acknowledging the Grant/ Fund Support of NMHS
Appendix 3 – List of Trainings/ Workshops/ Seminars with details of trained resources and dissemination material and Proceedings
Appendix 4 – List of New Products (utilizing the local produce like NTFPs, wild edibles, bamboo, etc.)
Appendix 5 – Copies of the Manual of Standard Operating Procedures (SOPs) developed
Appendix 6 – Details of Technology Developed/ Patents filled
Appendix 7 – Any other (specify)

Annexure I Consolidated and audited Utilization Certificate (UC) and Year wise Statement of Expenditure (SE)

Annexure-I

Consolidated and Audited
Utilization Certificate (UC) and Statement of Expenditure (SE)

For the Period: 1st April 2021-30th Nov 2021

1.	Title of the project/Scheme/Programme:	Hyperspectral imaging for sharper definitions of Himalayan ecosystems and its high value plant species under climate uncertainties.
2.	Name of the Principle Investigator & Organization:	Dr. Prashant K. Srivastava, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh
3.	NMHS-PMU, G.B. Pant National Institute of Himalayan Environment, Kosi-Katarmal, Almora, Uttarakhand Letter No. and Sanction Date of the Project:	GBPNI/NMHS-2017-18/MG-18/554 dated 26.02.2018
4.	Amount received from NMHS-PMU, G.B. Pant National Institute of Himalayan Environment, Kosi-Katarmal, Almora, Uttarakhand during the project period (Please give number and dates of Sanction Letter showing the amount paid):	**Rs. 23,50,840.00** GBPNI/NMHS-2017-18/MG-18/554/466/126/309/354
5.	Total amount that was available for expenditure (Including commitments) incurred during the project period:	₹2,65,371.00
6.	Actual expenditure (including commitments) incurred during the project period:	₹66,35,139.00
7.	Unspent Balance amount refunded, if any (Please give details of Cheque no. etc.):	**NA**
8.	Balance amount available at the end of the project:	₹-41,05,692.00
9.	Balance Amount:	₹-41,05,692.00
10.	Accrued bank Interest:	₹ 9,274.00

Certified that the expenditure of **₹25,29,447.00 (Twenty-Five Lakh Twenty-Nine Thousand Four Hundred Forty-Seven Only)** mentioned was actually incurred and **₹41,92,456.00** was Committed on the project/ scheme for the purpose it was sanctioned.

Date:

the Consolidated statement of Expenditure is issued on the certified SOE & UC received from the respective Institution of the copis of the project

(Signature of Principal Investigator) Dr. P.K. Srivastava Assistant Professor Institute of Environment & Sustainable Development Banaras Hindu University Varanasi-221005	(Signature of Registrar/ Finance Officer) 14/5 सहायक कुलसचिव (लेखा) Asstt. Registrar A/cs काशी हिन्दु विश्वविद्यालय Banaras Hindu University 14/5/22	(Signature of Head of the Institution) 19/5/22 उप कुलसचिव (विकास) Deputy Registrar (Development) काशी हिन्दु विश्वविद्यालय Banaras Hindu University वाराणसी / Varanasi-22100 15/5/22 18/5/22

OUR REF. No.

ACCEPTED AND COUNTERSIGNED

Date:

COMPETENT AUTHORITY
NATIONAL MISSION ON HIMALAYAN STUDIES (GBP NIHE)

NMHS 2020 Final Technical Report (FTR) – Project Grant 2 of 8

EXPENDITURE STATEMENT
NATIONAL MISSION ON HIMALAYAN STUDIES

Name of the Project/Fellowships: **Hyperspectral imaging for sharper definitions of Himalayan ecosystems and its high value plant species under climate uncertainties. (P07-683) up to 30th November 2021_ Consolidated**

Statement showing the expenditure of the period from 2021-2022
Sanction No. &Date:

1. Total outlay of the project/Fellowship : 2, 23, 74,560/-
2. Date of Start of the Project/fellowship : 26-02-2018
3. Duration : 3 years
4. Date of Completion :

a) Amount received during the financial year : **₹23,50,840.00**
b) Unspent amount carried forward from pervious Financial year : **₹2,65,371.00**
c) Total amount available for Expenditure (a+b) : **₹26,16,211.00**

S.No.	Budget head	Amount Carried forward	Amount received	Amount received+ amount carried forward	Actual Expenditure	Amount Balance Actual	Actual committed	Amount Balance committed
1	Salaries	₹-343220.00	₹1125840.00	₹782620	₹967383.00	₹-184763.00	₹1665997.00	₹-1850760.00
2	Permanent Equipment Purchased (Item-wise)	₹50691.00	0	₹50691	₹50082.00	₹609.00	₹0	₹609.00
3	Contingency	₹219743.00	₹175000.00	₹394743	₹212104.00	₹182639.00	₹582639.00	₹-400000.00
4	Consumables	₹320388.00	₹300000.00	₹620388	₹412188.00	₹208200.00	₹808200.00	₹-600000.00
5	Travel	₹17309.00	₹275000.00	₹292309	₹259590.00	₹32719.00	₹0	₹32719.00
6	Activities & other costs	₹460.00	₹375000.00	₹375460	₹528100.00	₹-152640.00	₹835620.00	₹-988260.00
7	Institutional charges	0	₹100000.00	₹100000	₹100000.00	₹0	₹300000.00	₹-300000.00
8	Accrued bank Interest	₹7018.00	₹7619.00	₹9274.00	₹0.00	₹9274.00	₹0.00	₹9274.00
9	**Total**	₹2,65,371.00	₹23,50,840.00	₹26,16,211.00	₹25,29,447.00	₹86764.00	₹41,92,456.00	₹-41,05,692.00
10	Amount allowed to be Carried forward to the next financial year				**₹-41,05,692.00**			

Cont.........2.

2.

Certified that the expenditure of **₹25,29,447.00 (Twenty-Five Lakh Twenty-Nine Thousand Four Hundred Forty-Seven Only)** mentioned was actually incurred and **₹41,92,456.00** was Committed on the project/ scheme for the purpose it was sanctioned.

Date:
Place:

The consolidated statement of Expenditure is
Based on the certified SOE and UC received
From the respective Institution of the Co-PIs of the project

Signature
Name:
PI
Dr. B.K. Srivastava
Assistant Professor
Institute of Environment &
Sustainable Development
Banaras Hindu University
Varanasi-221005

Signature सहायक कुलसचिव (लेखा)
Asstt. Registrar A/cs
Name: हिन्दू विश्वविद्यालय
Banaras Hindu University
Finance Officer
14/5
19/5
14/5/22

Signature 19/5/22
Name: उप कुलसचिव (विकास)
Deputy Registrar (Development)
Head of the Organization हिन्दू विश्वविद्यालय
Banaras Hindu University
वाराणसी / Varanasi-221005
19/5/22
18/5/22

OUR REF. No.

ACCEPTED AND COUNTERSIGNED

Date:

COMPETENT AUTHORITY
NATIONAL MISSION ON HIMALYAN STUDIES (GBPNIHESD)

NMHS 2020 Final Technical Report (FTR) – Project Grant 4 of 8

FORM GFR 12 – A
[(See Rule 238 (1)]

UTILIZATION CERTIFICATE FOR THE YEAR 2021-22 in respect
Of Non-Recurring Grants

1. Name of the Project/Fellowships: **Hyperspectral imaging for sharper definitions of Himalayan ecosystems and its high value plant species under climate uncertainties. (P07-683) up to 30th November 2021_ Consolidated**

2. Whether recurring grants: **non-recurring**

3. Name of the Grantee Org.: NATIONAL MISSION ON HIMALAYAN STUDIES

4. Grants position at the beginning of the financial year: 2021-2022
   (i) Cash in Hand/Bank: **₹ 50691 + 210.00 (Interest)**
   (ii) Unadjusted advances: NIL
   (iii) Total: **₹ 609.00 + 210.00 (Interest)**

5. Details of grants received, expenditure incurred and closing balances:

Unspent Balances of Grants received years	Interest Earned thereon	Interest deposited back to the Govt.	Grant received during the year			Total Available funds (1+2+3+4)	Expenditure (excluding commitments)	Closing Balances (5-6)
1	2	3	4			5	6	7
			Sanction No. (i)	Date (ii)	Amount (iii)			
₹ 50,691.00 + **₹ 210.00 (Interest)**	**₹ 0.00**	₹ 0-	- GBPNI/NMHS-2017-18/MG-18	20-JUNE-2018	**₹ 0.00**	**₹ 50,691.00**	**₹50,082.00**	**₹ 609.00** + **₹ 210.00 (Interest)**

Component wise utilization of grants:

Grant-in-aid– General	Grant-in-aid– Salary	Grant-in-aid–creation of capital assets	Total
₹ 31,16,980.00	-	**₹ 31,16,371.00**	**₹ 31,16,371.00**

Details of grants position as on 30-11-2021
   (i) Cash in Hand/Bank: - **₹ 609.00**
   (ii) Unadjusted Advances:
   (iii) Total: **₹ 609.00**

Cont...........2

2.

Certified that I have satisfied myself that the conditions on which grants were sanctioned have been duly fulfilled/are being fulfilled and that I have exercised following checks to see that the money has been actually utilized for the purpose for which it was sanctioned:

(i) The main accounts and other subsidiary accounts and registers (including assets registers) are maintained as prescribed in the relevant Act/Rules/Standing instructions (mention the Act/Rules) and have been duly audited by designated auditors. The figures depicted above tally with the audited figures mentioned in financial statements/accounts.

(ii) There exist internal controls for safeguarding public funds/assets, watching outcomes and achievements of physical targets against the financial inputs, ensuring quality in asset creation etc. & the periodic evaluation of internal controls is exercised to ensure their effectiveness.

(iii) To the best of our knowledge and belief, no transactions have been entered that are in violation of relevant Act/Rules/standing instructions and scheme guidelines.

(iv) The responsibilities among the key functionaries for execution of the scheme have been assigned in clear terms and are not general in nature.

(v) The benefits were extended to the intended beneficiaries and only such areas/districts were covered where the scheme was intended to operate.

(vi) The expenditure on various components of the scheme was in the proportions authorized as per the scheme guidelines and terms and conditions of the grants-in-aid.

(vii) It has been ensured that the physical and financial performance under has been according to the requirements, as prescribed in the guidelines issued by Govt. of India and the performance/targets achieved statement for the year to which the utilization of the fund resulted in outcomes given at Annexure – I duly enclosed.

(viii) The utilization of the fund resulted in outcomes given at Annexure – II duly enclosed (to be formulated by the Ministry/ Department concerned as per their requirements/ specifications.)

(ix) Details of various schemes executed by the agency through grants-in-aid received from the same Ministry or from other Ministries is enclosed at Annexure –II (to be formulated by the Ministry/Department concerned as per their requirements/ specifications).

Date:
Place:

The consolidated statement of Expenditure is Based on the certified SOE and UC received From the respective Institution of the Co-PIs of the project

Signature	Signature 14/5	Signature 19/5/22
Name:	Name: कुलसचिव (लेखा) Asstt. Registrar A/cs	Name: उप कुलसचिव (विकास) Deputy Registrar (Development)
PI	Finance Officer: वित्त अधिकारी Banaras Hindu University 14/5/22 14/5/22	Head of the Organization: काशी हिन्दू विश्वविद्यालय Banaras Hindu University वाराणसी / Varanasi-221005 8/5/22 18/5/22
Dr. P.K. Srivastava Assistant Professor Institute of Environment & Sustainable Development Banaras Hindu University Varanasi-221005		

NMHS 2020 Final Technical Report (FTR) – Project Grant 6 of 8

FORM GFR 12 – A

[(See Rule 238 (1)]

UTILIZATION CERTIFICATE FOR THE YEAR 2021-22 in respect Of Recurring Grants

1. Name of the Project/Fellowships: **Hyperspectral imaging for sharper definitions of Himalayan ecosystems and its high value plant species under climate uncertainties (P07-683) up to 30th November 2021 _Consolidated**
2. Whether recurring grants: Recurring
3. Name of the Grantee Org.: NATIONAL MISSION ON HIMALAYAN STUDIES
4. Grants position at the beginning of the financial year: 2021-2022
   (i) Cash in Hand/Bank: **₹ 2,14,680.00 + ₹ 6808.00 (Interest)**
   (ii) Unadjusted advances: NIL
   (iii) Total: **₹ -41,06,301.00 + ₹ 9,064.00.00 (Interest)**
5. Details of grants received, expenditure incurred and closing balances:

Unspent Balances of Grants received years	Interest Earned thereon	Interest deposited back to the Govt.	Grant received during the year			Total Available funds (1+2+3+4)	Actual Expenditure	Amount Balance Actual (5-6)	Committed Expenditure	Amount Balance Committed (7-8)
1	2	3	4			5	6	7	8	9
			Sanction No. (i)	Date (ii)	Amount (iii)					
₹ 2,14,680.00 + ₹ 6808 (Interest)	**2256.00**	**0**	**GBPNI/NMHS-2017-18/MG-18**	**20-JUNE-2018**	**₹23,50,840.00**	**₹ 25,65,520.00**	**₹ 25,29,447.00**	**₹86,155.00**	**₹ 41,92,456.00**	**₹ -41,06,301.00 + ₹ 9,064.00 (Interest)**

Component wise utilization of grants:

Grant-in-aid– General	Grant-in-aid– Salary	Grant-in-aid–creation of capital assets	Total

Details of grants position as on 30-11-2021

(i) Cash in Hand/Bank: **₹ -41,06,301.00.00**

(ii) Unadjusted Advances: NIL

(iii) Total: **₹ -41,06,301.00.00**

NMHS 2020 Final Technical Report (FTR) – Project Grant 7 of 8

Cont.........2.

Certified that I have satisfied myself that the conditions on which grants were sanctioned have been duly fulfilled/are being fulfilled and that I have exercised following checks to see that the money has been actually utilized for the purpose for which it was sanctioned:

(i) The main accounts and other subsidiary accounts and registers (including assets registers) are maintained as prescribed in the relevant Act/Rules/Standing instructions (mention the Act/Rules) and have been duly audited by designated auditors. The figures depicted above tally with the audited figures mentioned in financial statements/accounts.

(ii) There exist internal controls for safeguarding public funds/assets, watching outcomes and achievements of physical targets against the financial inputs, ensuring quality in asset creation etc. & the periodic evaluation of internal controls is exercised to ensure their effectiveness.

(iii) To the best of our knowledge and belief, no transactions have been entered that are in violation of relevant Act/Rules/standing instructions and scheme guidelines.

(iv) The responsibilities among the key functionaries for execution of the scheme have been assigned in clear terms and are not general in nature.

(v) The benefits were extended to the intended beneficiaries and only such areas/districts were covered where the scheme was intended to operate.

(vi) The expenditure on various components of the scheme was in the proportions authorized as per the scheme guidelines and terms and conditions of the grants-in-aid.

(vii) It has been ensured that the physical and financial performance under has been according to the requirements, as prescribed in the guidelines issued by Govt. of India and the performance/targets achieved statement for the year to which the utilization of the fund resulted in outcomes given at Annexure – I duly enclosed.

(viii) The utilization of the fund resulted in outcomes given at Annexure – II duly enclosed (to be formulated by the Ministry/ Department concerned as per their requirements/ specifications.)

(ix) Details of various schemes executed by the agency through grants-in-aid received from the same Ministry or from other Ministries is enclosed at Annexure –II (to be formulated by the Ministry/Department concerned as per their requirements/ specifications).

Date:
Place:

The consolidated statement of Expenditure is
Based on the certified SOE and UC received
From the respective Institution of the Co-PIs of the project

Signature	Signature 14/5	Signature 19/5
Name:	Name: Asstt. Registrar A/cs, Banaras Hindu University	Name: Deputy Registrar (Development), Banaras Hindu University, Varanasi-221005
PI	Finance Officer 14/5	Head of the Organization 15/5/22
Dr. P.K. Srivastava, Assistant Professor, Institute of Environment & Sustainable Development, Banaras Hindu University, Varanasi-221005	Final Technical Report (FTR) – Project Grant	8 of 8 18/5/22

Annexure III Consolidated Assets Certificate

Annexure-III

Consolidated Assets Certificate

Assets Acquired Wholly/ Substantially out of Government Grants

(Register to be maintained by Grantee Institution)

Name of the Sanctioning Authority: NMHS-PMU, GBNIHE

1. Sl. No.
2. Name of Grantee Institution: Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh
3. No. & Date of sanction order: GBPNI/NMHS-2017-18/MG-18/554 dated 26.02.2018
4. Amount of the Sanctioned Grant: Rs. 81,00,000.00
5. Brief Purpose of the Grant: Quantification of biophysical and biochemical parameters are very crucial for monitoring the high-value plants in the Himalayan region, its sustainable utilization as well as the conservation of related resources that constitutes of six objectives. First one is the creation of Spectral library and standardization for selected medicinal and threatened species. The second one is Spectral and image analysis for discrimination of selected economically important plant species followed by Model building for retrieval of Biochemical and Biophysical properties of selected species. Projection and scenario for distribution of selected species is the fourth objective followed by the fifth objective of fine scaling of space-time map of the high-value species in the Himalaya. The final objective is the management and conservation of species based on previous objectives.
6. Whether any condition regarding the right of ownership of Govt. in the property or other assets acquired out of the grant was incorporated in the grant-in-aid Sanction Order: NA
7. Particulars of assets actually credited NA or acquired **FieldSpec 4 SpectroRadiometer, Its accessories and Lookout VTOL™- X-Mapper Series Industrial Unmanned Aerial System**
8. Value of the assets as on **NA**
9. Purpose for which utilised at present Spectroradiometer is being utilized for capturing the spectra of the plant species for building the spectral library and UAV are utilized in conducting aerial surveillance, mapping, aerial photography, thermal imagery, and topographic monitoring of the Himalayan Forests.
10. Encumbered or not No
11. Reasons, if encumbered **NA**
12. Disposed of or not No
13. Reasons and authority, if any, for disposal **NA**
14. Amount realised on disposal **NA**

Any Other Remarks:

Prashant K. Srivastava
Principal Investigator
(P-07/683)
IESD, BHU
Varanasi-221005

(PROJECT INVESTIGATOR)

(Signed and Stamped)

(FINANCE OFFICER)

(Signed and Stamped)

Head
Department of Environment
& Sustainable Development
Banaras Hindu University

निदेशक
Director
पर्यावरण एवं धारणीय विकास संस्थान
Institute of Environment
& Sustainable Development
काशी हिन्दू विश्वविद्यालय
Banaras Hindu University

(HEAD OF THE INSTITUTION)

(Signed and Stamped)

उप कुलसचिव (विकास)
Deputy Registrar (Development)
काशी हिन्दू विश्वविद्यालय
Banaras Hindu University
वाराणसी / Varanasi-221005

NMHS 2020 Final Technical Report (FTR) – Project Grant 1 of 4

Annexure IV List of Inventory of Assets/ Equipment/ Peripherals

Annexure-IV

List or Inventory of Assets/ Equipment/ Peripherals

S. No.	Name of Equipment	Quantity	Sanctioned Cost	Actual Purchased Cost	Purchase Details
1	FieldSpec 4 SpectroRadiometer	1	81,00,000.00	80,99,391.00	Rs. 4983020.00
2	FieldSpec 4 SpectroRadiometer Accessories	1			Rs. 13,66,496.00
3.	Lookout VTOL™-X-Mapper Series Industrial Unmanned Aerial System	1			Rs. 17,49,875.00

Dr. Prashant K. Srivastava
Principal Investigator
(P-07/683)
IESD, BHU
Varanasi-221005

(PROJECT INVESTIGATOR)

(Signed and Stamped)

(FINANCE OFFICER)

(Signed and Stamped)

18/6/22

(HEAD OF THE INSTITUTION)

(Signed and Stamped)

4/6/2022

निदेशक
Director
पर्यावरण एवं धारणीय विकास संस्थान
Institute of Environment & Sustainable Development
काशी हिन्दू विश्वविद्यालय
Banaras Hindu University

Head
Department of Environment & Sustainable Development
Banaras Hindu University

उप कुलसचिव (विकास)
Deputy Registrar (Development)
काशी हिन्दू विश्वविद्यालय
Banaras Hindu University
वाराणसी / Varanasi-221005

20/6/22

NMHS 2020 Final Technical Report (FTR) – Project Grant 2 of 4

Annexure V Letter of Head of Institution/Department confirming Transfer of Equipment Purchased under the Project to the Institution/Department

Annexure-V

Letter of Head of Institution/Department confirming Transfer of Equipment Purchased under the Project to the Institution/Department

To,

The Convener, Mountain Division
Ministry of Environment, Forest & Climate Change (MoEF&CC)
Indira Paryavaran Bhawan
Jor Bagh, New Delhi-110003

Sub.: Transfer of Permanent Equipment purchased under Research Project titled "**Hyperspectral Imaging For Sharper Definitions Of Himalayan Ecosystems And Its High Value Plant Species Under Climate Uncertainties**" funded under the NMHS Scheme of MoEF&CC – reg.

Sir/ Madam,

This is hereby certified that the following permanent equipment purchased under the aforesaid project have been transferred to the Implementing Organization/ Nodal Institute after completion of the project:

1. FieldSpec 4 SpectroRadiometer
2. FieldSpec 4 SpectroRadiometer Accessories
3. Lookout VTOL™- X-Mapper Series Industrial Unmanned Aerial System

4/6/2022

Head
Department of Environment
& Sustainable Development
Banaras Hindu University

4/6/2022

Director
Institute of Environment
& Sustainable Development
Banaras Hindu University

Head of Implementing Organization: 4/6/22
Name of the Implementing Organization: IESD, BHU
Stamp/ Seal: Dr. Prashant K. Srivastava
Date: Principal Investigator
(P-07/683)
IESD, BHU
Varanasi-221005

Copy to:

1. The Nodal Officer, NMHS-PMU, National Mission on Himalayan Studies (NMHS), G.B. Pant National Institute of Himalayan Environment (NIHE), Kosi-Katarmal, Almora, Uttarakhand-263643

NMHS 2020 Final Technical Report (FTR) – Project Grant 3 of 4

Annexure-VI

Details, Declaration and Refund of Any Unspent Balance

Please provide the details of refund of any unspent balance and transfer the balance amount through RTGS (Real-Time Gross System) in favor of **NMHS GIA General** and declaration on the official letterhead duly signed by the Head of the Institution.

There is no balance to be returned, rather as per approval vide reference no GBPNI/NMHS-2017-18/MG-18/554/466/126/309/354/207 dated 30-11-2011, the total fund sanctioned by NMHS and available for utilization is Rs. 4935180.00.

Kindly note the further Bank A/c Details as follows:

Name of NMHS A/c:	NMHS GIA General
Bank Name & Branch:	Central Bank of India (CBI), Kosi Bazar, Almora, Uttarakhand 263643
IFSC Code:	CBIN0281528
Account No.:	3530505520 (Saving A/c)

3/6/22
Dr. Prashant K. Srivastava
Principal Investigator
(P-07/683)
IESD, BHU
Varanasi-221005

4/6/2022
Head
Department of Environment
& Sustainable Development
Banaras Hindu University

4/6/2022
निदेशक
Director
पर्यावरण एवं धारणीय
Institute of Environment
विकास संस्थान
& Sustainable Development
काशी हिन्दू विश्वविद्यालय
Banaras Hindu University

Appendix 1 Copies of Publications duly Acknowledging the Grant/ Fund Support of NMHS

Research Publications in Journals: -

1. Satish, K.V., Dugesar, V., Pandey, M.K., Srivastava, P.K., Pharswan, D.S., and Wani, Z.A (2022). **Development of forty-nine novel earth observation variables (NEOVs) for improving species distribution models in the Himalaya**. figshare. Dataset. https://doi.org/10.6084/m9.figshare.20049125
2. Manish Kumar Pandey, P. K. S. (2022). SHRIME: Spectral tool of Himalayan Rare, Invasive, Medicinal and Economical Plant Species. Communicated.
3. Satish, K. V., Srivastava, P. K., Behera, M.D., and Khan, M. L. (2023). **Ensemble machine learning models for Himalayan Bell Rhododendron distribution prediction under different climate change scenarios**. Environment, Development and Sustainability (Under review).
4. Dugesar V, Satish KV, Pandey MK, Srivastava PK, Petropoulos GP, **Performance Assessment of Sentinel-2 LAI products and data fusion for development of LAI datasets over high-altitude Himalayan forests: A Case Study**, Remote Sensing. (Under Review)
5. Satish, K.V., Vikas Dugesar, Manish K. Pandey, Prashant K. Srivastava, Dalbeer S. Pharswan, and Zishan Ahmad Wani. 2023. "**Seeing from Space Makes Sense: Novel Earth Observation Variables Accurately Map Species Distributions over Himalaya.**" Journal of Environmental Management 325 (January): 116428. https://doi.org/10.1016/j.jenvman.2022.116428.
6. Dugesar V, Satish KV, Pandey MK, Srivastava PK, Petropoulos GP, Anand A, Behera MD. **Impact of Environmental Gradients on Phenometrics of Major Forest Types of Kumaon Region of the Western Himalaya**. *Forests*. 2022; 13(12):1973. https://doi.org/10.3390/f13121973
7. Akash Anand, Prashant K. Srivastava, Prem Chandra Pandey, Mohammed L. Khan, and Mukund Dev Behera "**Assessing the niche of Rhododendron arboreum using entropy and machine learning algorithms: role of atmospheric, ecological, and hydrological variables,**" Journal of Applied Remote Sensing 16(4), 042402 (26 May 2022). https://doi.org/10.1117/1.JRS.16.042402
8. Ayushi Gupta, Prachi Singh, Prashant K. Srivastava, Manish K. Pandey, Akash Anand, K. Chandra Sekar & Karuna Shanker (2022) **Development of hyperspectral indices for anti-cancerous Taxol content estimation in the Himalayan region**, Geocarto International, DOI: 10.1080/10106049.2021.1983031
9. Gupta A, Srivastava PK, Petropoulos GP, Singh P. **Statistical Unfolding Approach to Understand Influencing Factors for Taxol Content Variation in High Altitude Himalayan Region.** *Forests*. 2021; 12(12):1726. https://doi.org/10.3390/f12121726
10. Anand A, Pandey MK, Srivastava PK, Gupta A, Khan ML. **Integrating Multi-Sensors Data for Species Distribution Mapping Using Deep Learning and Envelope Models**. Remote Sensing. 2021; 13(16):3284. https://doi.org/10.3390/rs13163284

Research Publications in Conferences: -

1. Gupta, A., Srivastava, P. K., and Shanker, K.: **Investigating the links between primary metabolites of medicinal species with leaf hyperspectral reflectance**, EGU General Assembly

2022, Vienna, Austria, 23–27 May 2022, EGU22-7726, https://doi.org/10.5194/egusphere-egu22-7726, 2022.
2. V. Dugesar, P. K. Srivastava and V. K. Kumra, "**Retrieval and Validation of Sentinel 2 LAI Product: A Comparison with Global Products Over High-Altitude Himalayan Forests**," IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, 2022, pp. 5648-5651, doi: 10.1109/IGARSS46834.2022.9884160.
3. V. Dugesar and P. K. Srivastav, "**Appraisal of Sentinel-2 Derived Vegetation Indices Using Uav Mounted With Visible-Ir Sensors**," *2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2022, pp. 1-4, doi: 10.1109/WHISPERS56178.2022.9955085.
4. V. Dugesar and P. K. Srivastav, "**Above-ground Biomass Estimation and Validation Using Multisource Remote Sensing Data, and Linear and Machine Learning Approaches Over High-altitudinal Himalayan Landscape**" (Submitted)

<u>Research Publications in Books: -</u>

1. Gupta, A., Srivastava, P.K., Satish, K., Chauhan, A. and Pandey, P.C. (2022). **Challenges and Future Possibilities Toward Himalayan Forest Monitoring. In Advances in Remote Sensing for Forest Monitoring** (eds P.C. Pandey and P. Arellano). https://doi.org/10.1002/9781119788157.ch14
2. Pandey, P.C., Pandey, M.K., Gupta, A., Singh, P. and Srivastava, P.K. (2021). **Spectroradiometry: Types, Data Collection, and Processing**. In Advances in Remote Sensing for Natural Resource Monitoring (eds P.C. Pandey and L.K. Sharma). https://doi.org/10.1002/9781119616016.ch2
3. Srivastava, Prashant K., Ramandeep Kaur M. Malhi, Prem Chandra Pandey, Akash Anand, Prachi Singh, Manish Kumar Pandey, and Ayushi Gupta. 2020. "**Revisiting Hyperspectral Remote Sensing: Origin, Processing, Applications and Way Forward**." In *Hyperspectral Remote Sensing*, 3–21. Elsevier. https://doi.org/10.1016/B978-0-08-102894-0.00001-2.

<u>Presentation:-</u>

1. An extended abstract was Presented entitled "**Classification of Medicinally Important Himalayan Plant Species using Hyperspectral Remote Sensing and Machine-Learning Algorithms**" in **ISRS-ISG National Symposium on "Remote sensing for environmental monitoring and climate change assessment: Opportunity and Challenges**" At: **Space Application Center**, Ahmedabad on 18-19 December 2020.
2. An extended abstract was Presented entitled " **Selection of Sensitive Bands for Anthocyanin Estimation for Pyracantha crenulata**" in **ISRS-ISG National Symposium on "Remote sensing for environmental monitoring and climate change assessment: Opportunity and Challenges**" At: **Space Application Center**, Ahmedabad on 18-19 December 2020.
3. A Poster was presented entitled "**Hyperspectral Sensors for Biochemical Parameter Detection in Medicinally Important Plant**" in **Indo U.S. Conference on Bioengineering & Regenerative Medicine (ICBR 2020)** organized at School of Biochemical Engineering, IIT BHU, Varanasi on 27-29 February 2020.

Estimation and Generation of Carbon Stock Product using Machine Learning Algorithms through Invocation of Cloud-Based Web Services

Akash Anand[1,*], Sumit K Chaudhary[1], Manish K. Pandey[1], Ayushi Gupta[1], Prachi Singh[1], Prashant K. Srivastava[2]

[1]Remote Sensing Laboratory, Institute of Environment & sustainable development, Banaras Hindu University, Varanasi-221005

[2]DST-Mahamana Centre for Excellence in Climate Change Research, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, India-221005

**Corresponding author e-mail: anand97aakash@gmail.com*

Abstract

A vast amount of carbon is stored in terrestrial ecosystem's benchmarking biophysical models, global carbon cycle and biogeochemical properties of regional as well as global ecology. Terrestrial carbon stocks are a critical climate change driver as well as a driving factor in forest conservation policies like Reducing Emissions from Deforestation and Forest Degradation (REDD+), and carbon offset protection plans. Remotely sensed carbon maps are widely used by scientists and policymakers, but the scope and coverage are confined to regional scale. As of now, there is no global product that provides carbon stock at a scale appropriate to modelling and decision-making applications with a frequent temporal data pool. There is a need to harmonize and integrate different remote sensing product to generate a comprehensive and temporally consistent carbon stock products, especially in the ridgy topographical regions like the Himalayas. The present study deals with an approach in which a comprehensive analysis is done to estimate total carbon stock from the year 2005-2018 using MODIS products namely, Gross Primary Productivity (GPP) and Net Primary Productivity (NPP). The model is trained on different non-linear machine learning algorithms and the best performing algorithm is adopted for generating final carbon stock product. The product is validated using the rigorous in-situ sampling which was done using non-destructive methods over the study area. The analysis is conducted over Garhwal division that lies in the lower foothills of the Himalayas where the regional topography, as well as the ecological diversity, is very heterogeneous, having regional biogeographical complexities along with extreme weather conditions, that makes regular in-situ monitoring of biophysical parameters quite challenging. The existing vulnerabilities of climate change are making an irrecoverable loss to the extreme regions that call for their regular monitoring. To achieve this, the decadal variation of several climatic parameters like Snow Cover (%), Land Surface Temperature (LST) and Precipitation are explored and a comparison has been made with the respective observed carbon stock. The output depicts a drastic variation in the regional climatic parameters that implies a direct impact on the regional ecosystem resulting in the growth of carbon stock in the higher altitudes. The dense carbon pools within the range are found to be shifting towards the higher altitudes which is indicative of the origin of rare species that were detected during the field sampling. The experiments were conducted and analysed on Google Earth Engine Platform and the final carbon stock product is generated as a user interface to be assessed by end-users and policymakers for better management and monitoring of terrestrial carbon stock.

Keyboards- Carbon Stock, Google Earth Engine, Himalayan region, Climatic Variables, Altitudinal Variation.

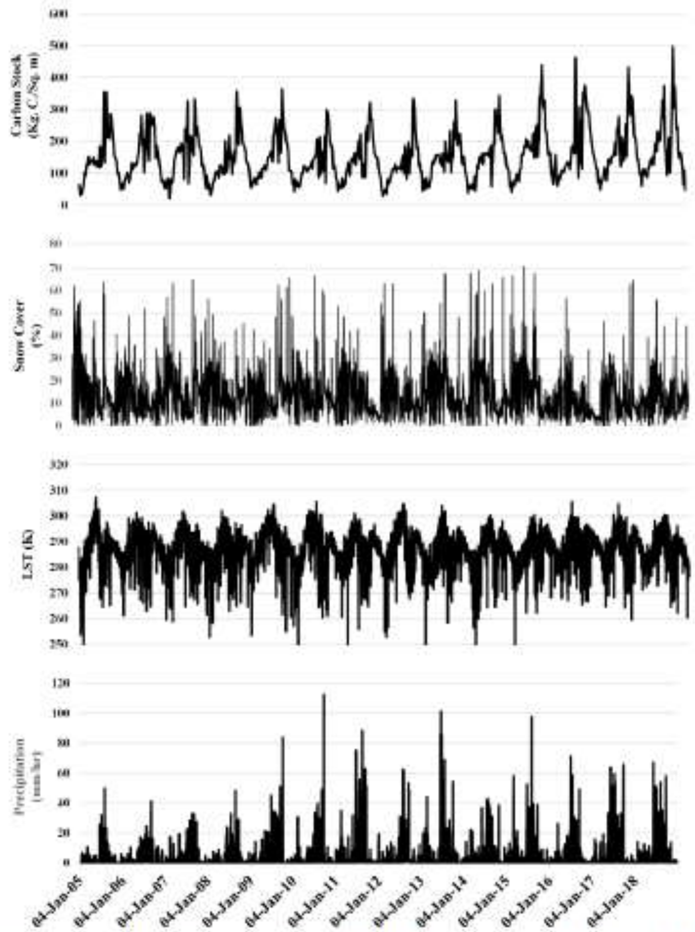
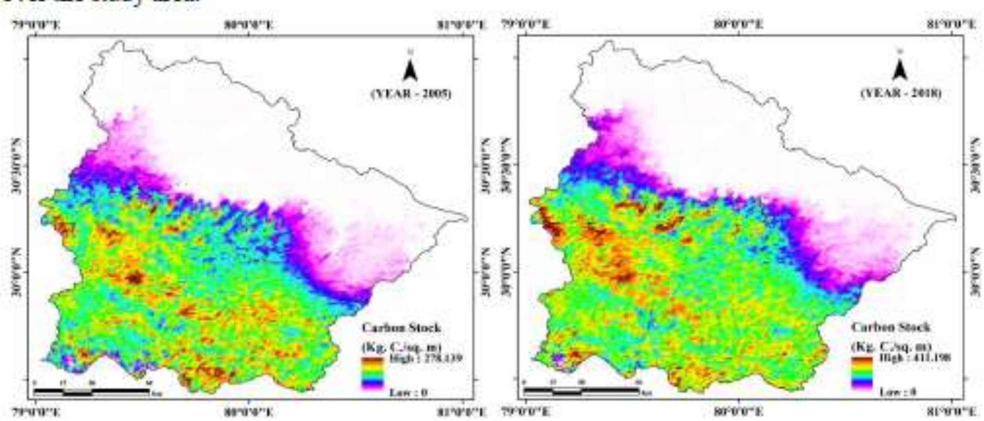


Figure 1. Temporal variation of (a) carbon stock (b) snow cover (c) LST and (d) precipitation over the study area.



Classification of Medicinally Important Himalayan Plant Species using Hyperspectral Remote Sensing and Machine-Learning Algorithms

Manish Kumar Pandey[1,*], Akash Anand[1], Ayushi Gupta[1], Prachi Singh[1], Prashant Kumar Srivastava[1,2] and A S Raghubanshi[1]

[1]*Remote Sensing Laboratory, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi 221005, India;*
[2]*DST Mahamana Center for Excellence in Climate Change Research, Banaras Hindu University, Varanasi 221005, India;*
**Corresponding author e-mail: pandey.manish@live.com*

Abstract

Quantification of biophysical and biochemical parameters are very crucial for monitoring the high-value plants in the Himalayan region, its sustainable utilization as well as the conservation of related resources. This would not only help in addressing the challenges of emerging climate change issues but also aid in understanding and prediction of several terrestrial ecosystem functions. The tough terrains of Himalayan ranges make the task of field sampling and hence lab-based retrieval of biophysical and biochemical parameters a challenging one. The use of remote sensing is very popular from several decades as of now and considered to be a cost-effective ad rapid solution for monitoring and retrieval of various vegetation parameters. The advancement of technologies in the current era has popularized the usage of hyperspectral sensors that generate data in a large number of contiguous bands spanning over a vast range of the electromagnetic spectrum. The advantage of hyperspectral sensors lies in its ability to detect even minute variations in the vegetation parameters. The high dimensional hyperspectral remote sensing data makes the task of information extraction a challenging one. This generated data possesses characteristics 3V's of Big data, i.e., Volume, Velocity and Variety. Volume represents the amount of data generated, Velocity represents the speed at which this data is generated and Variety represents the types of data generated, which in the current scenario is a combination of the structured text of reflectance value of various species, the images of various species, their satellite images and GPS coordinates. The aim of the current work is threefold. In the first step, optimal narrow wavelengths were detected which characterizes the chlorophyll content which is a biochemical vegetation parameter of two medicinally important Himalayan species, namely *Taxus wallichiana* Zucc. also known as Himalayan yew and *Berberis jaeschkeana*. These species were identified in Nanda Devi Bioreserve, Uttarakhand, India. The Leaf Spectroradiometry data of these species were measured using ASD FieldSpec® 4 Spectroradiometer. This data was used to identify best spectral wavelengths from the electromagnetic spectrum of 350 nm to 2500 nm that gave the highest correlation with chlorophyll content of these medicinally important plant species. In the second step, Chlorophyll content was computed corresponding to the collected spectra and optimal wavelength in blue, green, red and Near Infrared regions of electromagnetic radiation were identified and used for developing twelve hyperspectral indices. A relationship was established between the developed indices and chlorophyll content to evaluate the best hyperspectral index. In the third step, various machine learning algorithms were utilized to classify these two species bases on the value of spectra and chlorophyll content. The classifiers that were used for the experimentation were Bagging, Support Vector Machine (LibSVM) and K Nearest Neighbours (IBK). The classes were represented as 0 for *Berberis jaeschkeana* and 1 for *Taxus wallichiana* Zucc. The flowchart of the methodology is depicted in Figure 1.

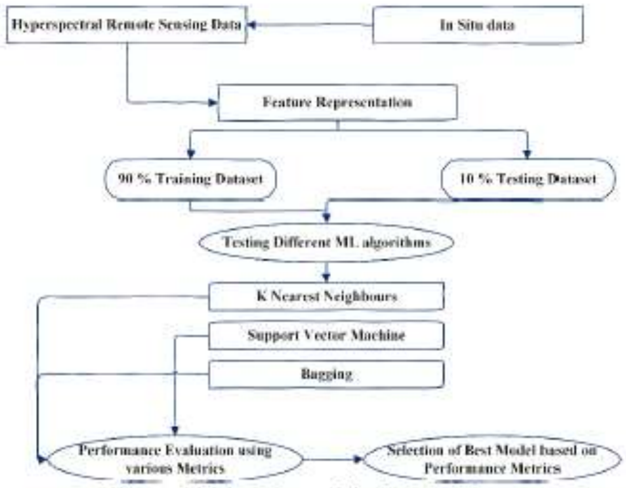


Figure 10 Flowchart of the Methodology

The output of the first two steps validates the consistency of the relationship with which hyperspectral indices showcase the variability with the biochemical parameters. The performance evaluation of the classifiers in the third case was made using various metrics like Overall Accuracy, Kappa Coefficient, Precision, Recall and F1 score. The classifier, K Nearest Neighbour was found to best performing than other classifiers with Overall Accuracy, Kappa Coefficient, Precision and Recall values of 93.5484 %, 0.8697, 93.5 and 93.5 respectively. Another measure of performance evaluation is the Receiver Operating Characteristics curve or ROC curve. Its application is in threshold independent assessment which illustrates the model performance at various discrimination thresholds. The ROC curve of the classifiers for the classification of both the curves are depicted in Figure 2.

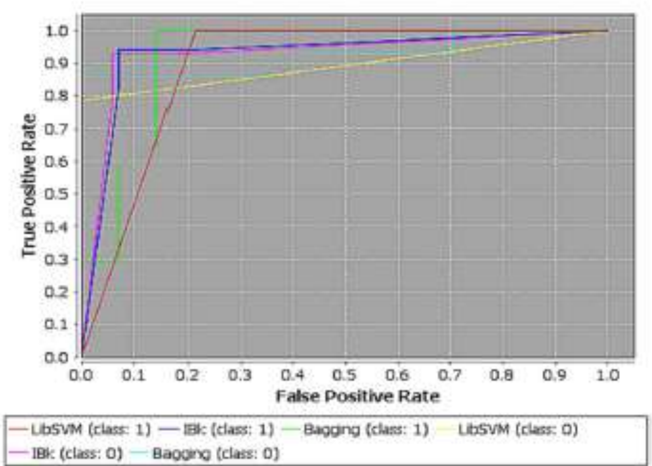


Figure 11 ROC for classifiers trained using 10-fold cross-validation.

Keywords - Big Data; Support Vector Machine; Bagging; K Nearest Neighbour; Hyperspectral Data; Vegetation Parameter.

Selection of Sensitive Bands for Anthocyanin Estimation for *Pyracantha crenulata*

Ayushi Gupta[1,*], Prachi Singh[1], Manish K. Pandey[1], Akash Anand[1], Prashant K. Srivastava[1,2] and K. S. Chandra Sekar[3]

[1]*Remote Sensing Laboratory, Institute of Environment & sustainable development, Banaras Hindu University, Varanasi-221005*

[2]*DST-Mahamana Centre for Excellence in Climate Change Research, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, India-221005*

[3]*Centre for Biodiversity Conservation and Management (CBCM), G.B. Pant National Institute of Himalayan Environment, Kosi-Katarmal, Almora, Uttarakhand, India- 263643*

**Corresponding author ayushi.gupta10@bhu.ac.in*

Abstract

Physically-based radiative transfer models (RTMs) help in the development of inversion models to accurately retrieve atmospheric and vegetation properties from remotely sensed data. However, advanced RTMs can be computationally troublesome, which makes them impractical in many real-time applications. To overcome this problem, a substitute RTMs is proposed through proxy meta-models. The Metamodels approximates the functioning of RTMs through statistical learning regression (SLR) methods that cater to many new applications because of their computational efficiency and outstanding accuracy. Anthocyanins are a group of polyphenolic pigments that are universally found in the plants. The anthocyanins not only play a crucial role in reproduction by attracting pollinators and seed dispersers but also acts protection shield against various abiotic and biotic stresses. Several studies revealed them to have properties that are useful in many ailments as well as immunity boosters. In the current study, the existing anthocyanin reflectance index was tested based on hyperspectral data (between 350- 2500nm) that was acquired using ASD spectroradiometer for a medicinally important species *Pyracantha crenulata*. The three most common indices, Normalized Difference Type Index, modified Anthocyanin Content Index (mACI), and modified Anthocyanin Reflectance Index (mARI) were evaluated for the estimation of anthocyanin to develop the SLR model. The mARI was found to be best performing with a correlation value of 0.606 at wavelengths 1368nm and 816nm. Thus, mARI can be utilized for the estimation of anthocyanin in the Himalayan region.

Pyracantha crenulata (D. Don) M. Roemer (Rosaceae), commonly called as Himalayan firethorn, is an evergreen thorny shrub species found in open slopes that grows between 1,000 and 2,400 m above mean sea level (a.m.s.l.), from Himalaya to South-West China and Myanmar. *P. crenulata* is reported to have many traditional and ethnomedicinal uses. It is used in the treatment of hepatic, cardiac, stomach, and skin disease (Khare 2004). Recently its fruit extract was found to have antiurolithogenic and diuretic activities (Bahuguna, Rawat et al. 2009) and antihypertensive (Negi, Singh et al. 2018) in nature. Moreover, its fruits are considered as a potent source of natural antioxidants (Bhatt, Rawat et al. 2017). Since this herb is found on open slopes, it's estimation on a broader scale becomes the next logical step for further uses. Usually, plant health is determined by its pigment content. The conventional approaches for its estimation generally involve extraction in organic solvents for subsequent spectrophotometric measures which is time-consuming and unsuitable at a larger scale. Hence, the development of vegetation indices and Statistical linear regression (SLR) model is found to be an effective solution that offers specific indices more suitable for the selected species. In this study, a threefold evaluation is conducted, viz. a) three of the already developed indices for anthocyanin at various wavelengths b) modified old indices with hyperspectral data c) developed SLR model for the estimation of anthocyanin for *P. crenulata*.

P. crenulata leaves samples were collected at an elevation that ranges between 1370- 1890m at Ranikhet, Kausani, Jageshwar, and Bhimtal of Garhwal region of Uttrakhand. The spectra of each sample were recorded using an ASD spectroradiometer. The samples were then stored in liquid nitrogen for the estimation of anthocyanin using the destructive method (Flores, Oosterhaven et al. 2005). One of the most utilized indices, normalized vegetation type index along with two other anthocyanin indices were tested using contiguous hyperspectral data between 350-2500 nm. The datasets were divided into two sets between calibration and validation set as depicted in Figure 1. Modified Anthocyanin Indices were developed using the retrieval model.

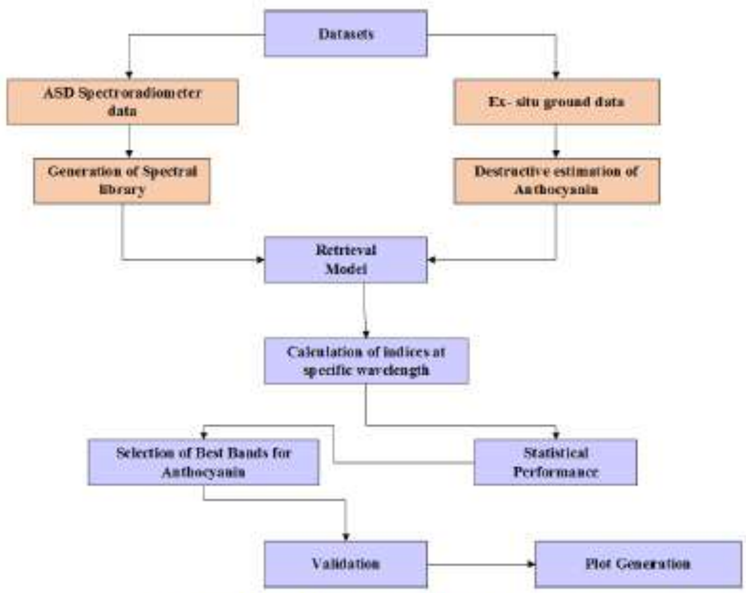


Figure 1. Flowchart of methodology opted

The three most common already developed indices Normalized difference type index, modified anthocyanin content index (mACI), and modified Anthocyanin reflectance index (mARI) were tested for the estimation of anthocyanin. The mARI was found to be most appropriate for the estimation of anthocyanin of *P. crenulate* based on the correlation value of 0.606 at wavelengths 1368nm and 816nm. The output of the current work suggests that a better and accurate estimation of indices for the retrieval of biochemical parameters would be better facilitated using the hyperspectral remote sensing data.

EGU General Assembly 2022

EGU22-7726
https://doi.org/10.5194/egusphere-egu22-7726
EGU General Assembly 2022

Investigating the links between primary metabolites of medicinal species with leaf hyperspectral reflectance

Ayushi Gupta[1], Prashant K Srivastava[1], and Karuna Shanker[2]
[1]Remote Sensing Laboratory, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, India
[2]CSIR - Central Institute of Medicinal and Aromatic Plants, Lucknow, India

Recent studies have shown that the turnover in tree species composition across edaphic and elevational gradients is strongly correlated with functional traits. However, our understanding of functional traits has been limited by the lack of detailed studies of foliar chemistry across habitats and the logistical & economic challenges associated with the analysis of plant functional traits at large geographical scales. Advances in remote sensing and spectroscopic approaches that measure spectrally detailed light reflectance and transmittance of plant foliage provides accurate predictions of several functional chemical traits. In this study, *Pyracantha crenulata* (D. Don) M. Roemer has been used, which is an evergreen thorny shrub species found in open slopes between 1,000 and 2,400 m above mean sea level. *P. crenulata* is used in the treatment of hepatic, cardiac, stomach, and skin disease. In this study the P. crenulata leaves samples spectra were recorded using an ASD spectroradiometer and following primary metabolites such as chlorophyll, anthocyanin, phenolic, and sterol were analyzed. The spectroradiometer data were preprocessed using filter and then reduced to a few sensitive bands by applying feature selection to the hyperspectral data. The band values were directly correlated with the measured values. The analysis indicates a significant correlation between *P. crenulata* primary metabolite in the Visible and Infrared region (VISIR). This result suggests that molecules that have important functional attributes could be identified by VISIR spectroscopy, which would save a lot of time and expense as compared to wet laboratory analysis.

remote sensing

MDPI

Article

Integrating Multi-Sensors Data for Species Distribution Mapping Using Deep Learning and Envelope Models

Akash Anand [1], Manish K. Pandey [1], Prashant K. Srivastava [1,*], Ayushi Gupta [1] and Mohammed Latif Khan [2]

[1] Remote Sensing Laboratory, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi 221005, Uttar Pradesh, India; anand97aakash@gmail.com (A.A.); pandey.manish@live.com (M.K.P.); ayushi.gupta10@bhu.ac.in (A.G.)

[2] Department of Botany, Dr. Harisingh Gour Central University, Sagar 470003, Madhya Pradesh, India; mlkhan@dhsgsu.edu.in

* Correspondence: prashant.iesd@bhu.ac.in

Abstract: The integration of ecological and atmospheric characteristics for biodiversity management is fundamental for long-term ecosystem conservation and drafting forest management strategies, especially in the current era of climate change. The explicit modelling of regional ecological responses and their impact on individual species is a significant prerequisite for any adaptation strategy. The present study focuses on predicting the regional distribution of *Rhododendron arboreum*, a medicinal plant species found in the Himalayan region. Advanced Species Distribution Models (SDM) based on the principle of predefined hypothesis, namely BIOCLIM, was used to model the potential distribution of *Rhododendron arboreum*. This hypothesis tends to vary with the change in locations, and thus, robust models are required to establish nonlinear complex relations between the input parameters. To address this nonlinear relation, a class of deep neural networks, Convolutional Neural Network (CNN) architecture is proposed, designed, and tested, which eventually gave much better accuracy than the BIOCLIM model. Both of the models were given 16 input parameters, including ecological and atmospheric variables, which were statistically resampled and were then utilized in establishing the linear and nonlinear relationship to better fit the occurrence scenarios of the species. The input parameters were mostly acquired from the recent satellite missions, including MODIS, Sentinel-2, Sentinel-5p, the Shuttle Radar Topography Mission (SRTM), and ECOSTRESS. The performance across all the thresholds was evaluated using the value of the Area Under Curve (AUC) evaluation metrics. The AUC value was found to be 0.917 with CNN, whereas it was 0.68 with BIOCLIM, respectively. The performance evaluation metrics indicate the superiority of CNN for species distribution over BIOCLIM.

Keywords: spatial distribution modelling; convolutional neural network; *Rhododendron arboreum*; biodiversity management; ecological responses

check for updates

Citation: Anand, A.; Pandey, M.K.; Srivastava, P.K.; Gupta, A.; Khan, M.L. Integrating Multi-Sensors Data for Species Distribution Mapping Using Deep Learning and Envelope Models. *Remote Sens.* **2021**, *13*, 3284. https://doi.org/10.3390/rs13163284

Academic Editor: Parth Sarathi Roy

Received: 6 June 2021
Accepted: 11 August 2021
Published: 19 August 2021

1. Introduction

The Himalayan ecosystem is experiencing a continuous temperature rise, and the impact of climate change can be seen very clearly in the Himalayas, which demonstrates the need to monitor the Himalayan ecosystem even more [1,2]. The Himalayas are home to several medicinally and economically important plant species, *Rhododendron* species with botanical name *Rhododendron arboreum Sm.* from the family Ericaceae is among one of them [3–5]. It is widely spread in Himalayas, South India, and Sri Lanka [4]. With tremendous biological significance, it can sustain itself in the fragile ecotone between the alpine and subalpine biomes. Despite being identified as a medicinally important plant species, the geographical distribution and geospatial modelling of *Rhododendron arboreum* have not been explored to its fullest and needs to be deciphered, which will further benefit the formulation of conservation strategies [6]. The literature review of past studies offered isolated information on the distribution pattern, frequency of the species, genetic diversity, and net productivity of *Rhododendron arboreum*, particularly in Himalayan regions. The studies primarily focus on the areas that are found over the Mussoorie hills of the Uttarakhand [7], Himachal Pradesh [8], and Garhwal division of Himalayas [9] and mainly showcased the threat of habitat fragmentation and frequency degradation of *Rhododendron arboreum* over the Himalayan region. Therefore, a holistic and cohesive research approach is needed to track the distribution of these species so as to generate baseline data for future research programmes, management practices, and conservation policies.

Throughout the centuries, researchers have observed and documented the linear relationship between species distribution and their local physical ecosystem and the role of climate and altitudinal variation in their occurrence, which can be found in the available scientific writings of the early 19th century [10–12]. Identifying the parameters for establishing the relationship between species and the environment is the core step for simulating the geographical distribution of any species [13,14]. Species Distribution Modelling (SDM) or Ecological Niche Modelling (ENM) is widely used in biogeography [15], macroecology [16], and biodiversity [17] research to model the geographical distribution of species. It is a statistical tool that performs habitat suitability analysis using high-dimensional digital data through regression or machine learning algorithms. Distribution modelling is achieved by establishing a linear or nonlinear relationship between regional climatic and ecological conditions. As each species is adapted to specific tolerance zones (also known as niches), SDM helps to identify the environmental constraints and simulates the n-dimensional input data to produce habitat suitability [18].

Generally, the SDM techniques take in geocoded input data of species distribution and establish a relationship with the regional environmental and climatic conditions to map its distribution throughout an area of interest [14,19,20]. One of the most commonly used SDM is BIOCLIM, which is also used in the present study. BIOCLIM is one of the earliest developed SDM algorithms, and it is mainly used by ecologists due to its easy-to-use algorithm, which is more accessible than other models. BIOCLIM was first introduced by reporting bioclimatic profiles and distribution maps of 73 species. After introducing BIOCLIM, several researchers applied the model in different bioclimatic conditions and received better results [21–24]. They found BIOCLIM to be quite consistent. The reason behind BIOCLIM's popularity is credited to its predefined assumptions and less complex training algorithm.

The theoretical aspect of SDM considers that a given species is likely to be found in a single privileged ecological niche under an ecosystem of unimodal distribution [25]. Still, an actual scenario is more complex and diverse than a hypothetical niche. To overcome this limitation, deep neural networks would be a good option, as their architecture favours high order multi-dimensional feature interactions without constraining their functional form [26]. Deep neural networks have shown significantly better results in image classification, and there are several cases where single-layered neural networks have been used for SDM [27,28]. Recently, a study by [29] has shown a better deep neural network prediction ability in SDM that has been found to be even better performing than the conventional ecological SDM models. A detailed discussion of convolution neural networks in handling and learning non-linear features are given in later sections.

In this paper, a study was conducted to map the distribution of *Rhododendron arboreum* using linear models, namely BIOCLIM, and Convolutional Neural Network (CNN) architecture has been proposed to a establish nonlinear relationship between input parameters. A total of 16 environmental and climatic parameters acquired from different satellites are used as input parameters, which influence the distribution of a particular species to a greater extent in the given study area.

2. Materials and Methods

2.1. Study Area

The current study was conducted within four districts of Uttarakhand, India, namely Chamoli, Almora, Bagheshwar, and Pithoragarh, which is situated at the foothills of the Himalayas. Geographically, the study area lies between 28°43′22.42″ to 31°27′22.06″N latitude and 77°34′20.28″ to 81°2′34.35″E longitude with an area of around 20,736.99 km^2. The study area is shown in Figure 1. Being situated in the Himalayan Mountain range, the variation in ground elevation is very sharp, varying from 416 to 7801 m from mean sea level for the present study area. Due to the variation in elevation and unique climatic settings, this region is immensely rich with thousands of different plant species and has a remarkable diversity in flora and fauna [30,31]. This region experiences evenly distributed rainfall throughout the year and has an average temperature of 23.4 °C [32]. According to terrestrial ecoregion classifications that have been previously performed, the ecoregions found in the present study area are tropical and subtropical moist broadleaf forest, coniferous forest, temperate broadleaf and mixed forest, temperate conifer forest, and montane grassland and shrublands, and approximately 20% of the area is covered with snow throughout the year.

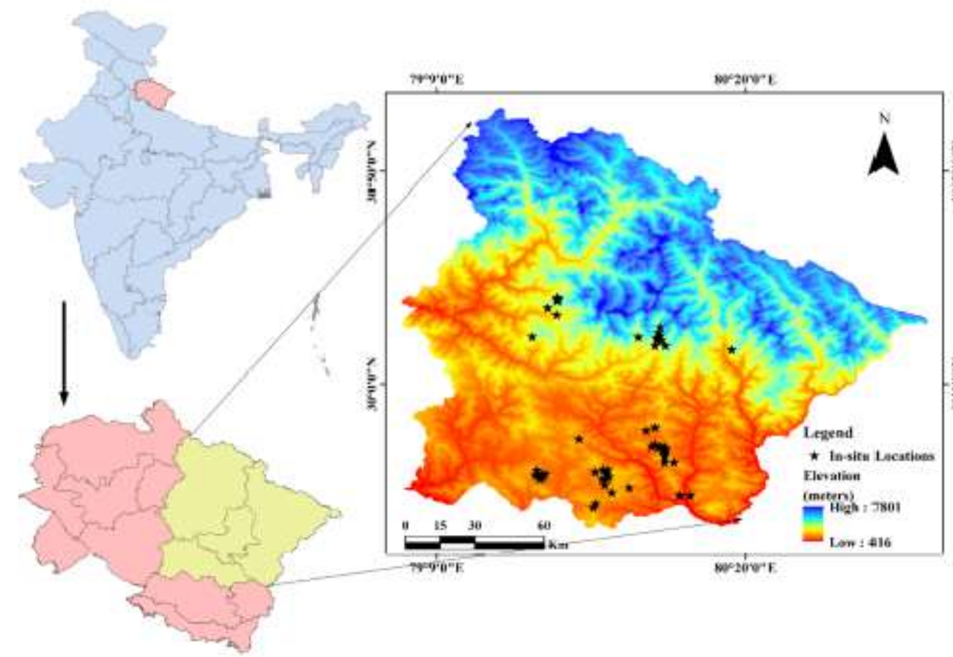


Figure 1. Location of the study area in the Western Himalayan region.

2.2. Target Species and Occurrence Data

Rhododendron species are found in the Himalayan range, which exhibits vast biological significance in the fragile ecotone of the alpine and subalpine zones [33]. Among the variety of Rhododendrons, *Rhododendron arboreum* is also a common species in the Western Himalayan region and can be found at an elevation range of 1200–4000 m above mean sea level. It shows some characteristics of invasive species and has a high medicinal value that makes the study of its distribution and the impact of climatic and ecological parameters on its growth very important [34]. Not only does this species carry medicinal value, but it is also very highly valued economically [35]. Medicinally, it is found to possess anti-cancer, immunomodulatory, anti-inflammatory, hepatoprotective, antidiabetic, antioxidant, antidiarrheal, adaptogens, antimicrobial, and antinociceptive properties, among others [35]. Economically, it has found its usage in squash, local brew, jellies, jams, and *sherbet* (*rhodojuice*). The juice from the leaves is used to encounter bed bugs bites. The wood from Rhododendron can be used to make tools used in agriculture such as '*Khukri*' handles, etc. [36]. The leaves are used for decoration purposes in houses as well as in temples [37]. The wood is also used for the preparation of charcoal and can be used as a fuel. Some studies reported that consuming the squash made from the flowers can serve as a treatment for mental retardation [38,39], and flowers along with the roots and bark were found to be effective in treating digestive, heart, and respiratory complications [40]. The leaves of the plant burnt with juniper leaves are used to cleanse the air [41]. Menstrual cramps and heartaches are treated with the juice and squash made out of these flowers [42]. The extracts of the plants have also been utilized in curing nasal bleeding [43], headache, fever, rheumatism, wounds, dysentery [44], cough, skin diseases, liver malfunction, piles, worms, and jaundice as well as for preliminary cancer treatment [45].

Since phonological responses are better observed during the flowering season, the ground data sampling was done in September 2019, March 2020, and March 2021 at different elevation ranges [46]. The amount of ground data available for certain species, and particularly for *Rhododendron arboreum*, are limited, which makes it crucial to undertake the possibility of bias based on the sample size. As per the study by [47], they found that the SDM models with a smaller size consistently performed poorly and suggested that for reliable accuracy, the sample size must be greater than 30. In purview of the studies conducted by several researchers and considering the local geography, a total 65 homogenous patches of *Rhododendron arboreum* were identified and geotagged within the study area using the handheld Garmin GPS (Global Positioning System) with a horizontal accuracy of 95% ± 9.3 m. According to Wisz et al. [47], a small sample size (>30) is generally useful in exploratory modelling, and considering that the present study is conducted at a regional scale for a single target species, 65 sample points are enough for regional distribution modelling. The observed multiple species occurrence within the pixels were removed by applying spatial rarefication, which gave a single occurrence point per pixel. The photographs of the *Rhododendron arboreum* captured during field sampling are given in Figure 2.

Figure 2. Field photographs of Rhododendron arboreum.

2.3. *Environmental Variables*

The environmental variables used in this study include different bioclimatic variables acquired from various active satellites. Recent developments in satellite sensors have enabled access to several ecological and climatic information derived from satellite observations at higher spatial and temporal resolutions. Current studies have used MODIS (https://modis.gsfc.nasa.gov/data accessed on 5 June 2021), Sentinel-2 (https://sentinel.esa.int/web/sentinel/missions/sentinel-2/data-products accessed on 5 June 2021), Sentinel-5P (https://sentinel.esa.int/web/sentinel/missions/sentinel-5/data-products accessed on 5 June 2021), ECOSTRESS (https://ecostress.jpl.nasa.gov/data accessed on 5 June 2021), and SRTM (https://www2.jpl.nasa.gov/srtm accessed on 5 June 2021) satellite observations as an input parameter for the SDM algorithms. The satellite products are acquired during the sampling period (September 2019, March 2020, and March 2021) and then averaged, so that it can be used as a single product for each parameter. The Leaf Area Index (LAI) [48] and a fraction of Photosynthetically Active Radiation (fPAR) were retrieved from the MODIS data product. The motive behind using both LAI and fPAR establishes their direct relationship with the surface photosynthesis, evapotranspiration, and net primary productivity of the plants that is further utilized to estimate the water cycle processes, terrestrial energy, biophysical, and biochemical properties of the regional vegetation. Sentinel 2 optical data were used to estimate NDVI and EVI values at a fine scale of 10 m, which helped in understanding the vegetation status throughout the study area and were also used to mask the non-vegetative lands. Although there are a number of vegetation indices, which can play a crucial part in identifying the species distribution including Modified Soil Adjusted Vegetation Index (MSAVI) as well as other soil and ground surface adjusted indices, the target specie Rhododendron Arboreum is found in the dense forest cover of Himalayas and is independent of any soil and surface distortion; therefore, only NDVI and EVI were considered for SDM. Sentinel-5P is one of the most recent satellite missions from the European Space Agency (ESA), which is a combined mission of the European Union. It can take atmospheric measurements with the high spatial and temporal resolution and is utilized to retrieve several atmospheric parameters. Presently, eight different Sentinel-5P parameters have been used to establish a relationship between the existence of target species and to model species distribution which includes, the Aerosol Absorption Index (AAI), CO density, water vapor column, columnar NO_2, columnar O_3 level, SO_2 density, surface albedo, and tropospheric Formaldehyde (HCHO) density [49]. Evapotranspiration (ET) and Land Surface Temperature (LST) are also included to study the response of the target species with regional land processes. It was retrieved from the ECOSTRESS satellite. Elevation is an important parameter when studying the species distribution. It has a huge role in species growth and distribution due to the changing conditions at varying altitudes; therefore, SRTM Digital Elevation Model (DEM) data are used to retrieve the elevation factor at a spatial resolution of 30 m. A significant factor that is also included is terrestrial ecoregions to acquire the regional biome information to understand the diversity of the Himalayan ranges that coexist with each other. The ecoregion the vector data provided by [50] was used, which was attributed into 14 classes, 5 of which are used in the current study area that is found in the foothills of the Himalayas, namely tropical and subtropical moist broadleaf forests, tropical and subtropical coniferous forests, temperate broadleaf and mixed forests, temperate coniferous forests, and montane grassland/shrubland. The input parameters and their source mission are listed in Table 1.

All of the input parameters were used based on their linear and non-linear correlation with the occurrence data. The parameters are the yearly average for year 2020, considering the cloud free pixels only. Additionally, a correlation matrix was plotted to interpret the relationship between each parameter.

Table 1. Input parameters.

Satellite/Vector Data	Parameter	Unit	Spatial Resolution	Description
MODIS	LAI	Unitless	500 m	Defined as the projected area of leaves per unit of ground surface area.
	fPAR	Unitless	500 m	The fraction of photosynthetically active radiation (400–700 nm) absorbed by an integrated plant canopy.
Sentinel-5P	Aerosol Absorption Index	Unitless	0.01 arc degree	Indicates the elevated absorbed aerosols in the atmosphere.
	Vertically integrated CO column density	mol/m^2	0.01 arc degree	CO is an important atmospheric trace gas and a major atmospheric pollutant. A major source of CO is biomass burning and the oxidation of hydrocarbons.
	Water vapour column	mol/m^2	0.01 arc degree	A major greenhouse gas that directly impacts plant growth as well as photosynthesis.
	The total vertical column of NO_2	mol/m^2	0.01 arc degree	A trace gas mostly found in the troposphere and stratosphere that can harm plant growth with an increase in its concentration
	The total atmospheric column of O_3	mol/m^2	0.01 arc degree	Acts as a shield for the biosphere from solar ultraviolet radiation. It is an important greenhouse gas, and its high concentration can be harmful to the vegetation.
	SO_2 vertical column density	mol/m^2	0.01 arc degree	Has a major impact on local and global climate change and is directly and indirectly related to plant growth and distribution.
	Surface Albedo	Unitless	0.01 arc degree	The flux per unit area received at the surface, and it shows low values in dense forest sue to its high absorption.
	Tropospheric HCHO column number density	mol/m^2	0.01 arc degree	An intermediate gas in most of the oxidation chains of non-methane organic compounds. The inter-annual variations of HCHO distribution result from the oxidation in organic hydrocarbons from vegetation, fires, industrial sources, and temperature changes.
Sentinel-2	NDVI	Unitless	10 m	A simple indicator to assess whether or not the observed target contains green vegetation.
	EVI	Unitless	10 m	An optimized vegetation index to enhance the vegetation signal by decoupling the canopy background signal and reduction in atmospheric noises.
ECOSTRESS	Evapotranspiration	W/m^2	70 m	The latent heat flux coming from the earth's surface in the form of evaporation and plant transpiration.
	Land Surface Temperature	Kelvin	70 m	The radiative skin temperature of the earth's surface derived from solar radiation.
SRTM	DEM	Meters	30 m	An array of equally spaced elevation values referenced horizontally by a geographical coordinate system.
Terrestrial Ecoregions	Biome	Vector data		The classification of different types of forest present worldwide. The biome classification used for the present study has 14 different types of forest classes.

2.4. Ecological Niche Modelling

Broadly, there are two methods to model the ecological niche: a mechanistic approach and a correlative approach [47]. The mechanistic approach deals with the physiological limiting mechanisms of species intolerance in ecological conditions. In this approach, the growth parameters are taken into consideration, such as soil pH, nitrogen content in the soil, incoming solar radiation, carbon dioxide intake by plants, etc. [51]. The correlative approach, also known as an empirical approach, uses the environmental variables that are reasonably expected to affect the growth of a particular species. The basis of the correlative approach is the interrelation between the observed parameters within the identified species location, which is used to establish a relationship between the parameters to model the species distribution for an entire area [52]. Having run the algorithm, a species distribution map can be generated using the established relationship. At this stage, the model's ability can also be tested using a set of species occurrence data that was not used in model development via suitable statistical parameters [53]. The SDM used in the present study only includes the presence of the BIOCLIM model and CNN based SDM. The representation of the conventional methodology is given in Figure 3.

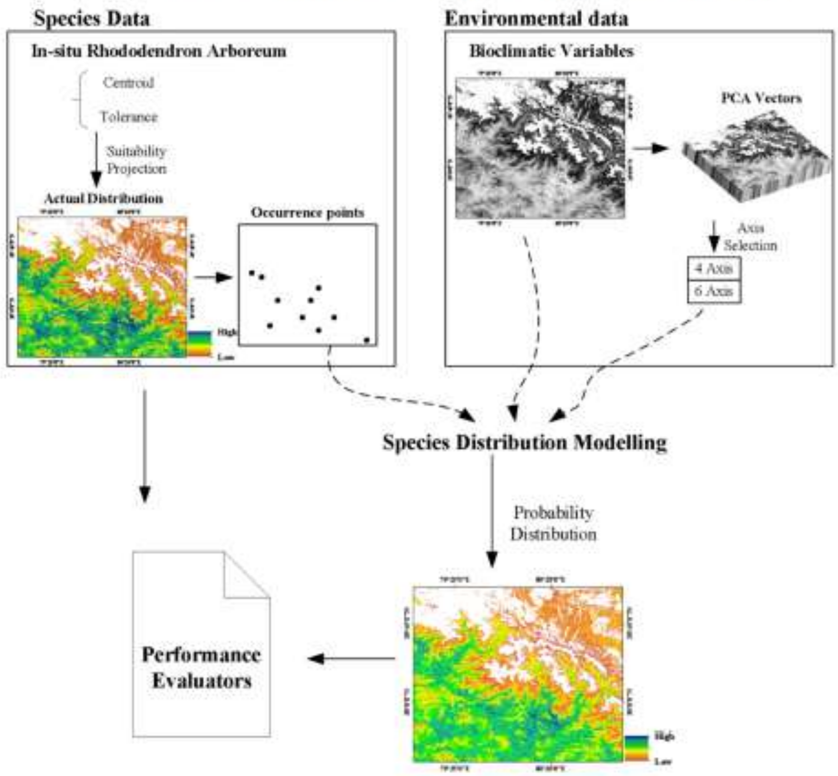


Figure 3. Flow diagram of the conventional approaches for species distribution modelling.

2.4.1. BIOCLIM

BIOCLIM is one of the first SDM algorithms to be introduced by [54]. The BIOCLIM is a widely used SDM due to its easy-to-use graphical user interface and wide application area. After the introduction of BIOCLIM, many researchers published their work using this algorithm, including the work of [55], in which they discussed the application of BIOCLIM in building ecographic regions and ways to improve the estimation of the ecological distance between patches in meta-population landscape dynamics. The study by [55] also pointed out some pros and cons, which include the error associated with the climatic parameters, defining the ranking of factors, and taxonomic uncertainty. Early BIOCLIM applications occurred between 1984–1991 in terms of ecology and conservational biology and were addressed by [56].

BIOCLIM is based on a bioclimatic envelope model, which is widely used to predict the potential species distribution, which does not account for any possible interrelation between variables. Being an intuitively simple model, it assigns equal weight to each variable and produces binary predictions [57,58]. To predict the probability species distribution, BIOCLIM compares the values of input variables of known locations to the values of unknown pixels. The closer the value of an unknown pixel to the available pixel, the more suitable the location is for a particular species to be found. BIOCLIM is simple and intuitive. It is susceptible to over prediction and as specified, does not account for the interactions between input variables [59].

2.4.2. CNN

A deep neural network is a multi-layered model that can learn complex nonlinear relationships between the input parameters. The current study is an attempt that has been made to use the Convolutional Neural Network (CNN) architecture for SDM. During the last two decades, there has been a huge increase in deep learning and advanced machine learning algorithms in a variety of research fields [60,61]. Deep learning conducts high-level data abstraction using a hierarchical architecture consisting of multiple interconnected layers with multiple artificial neurons. The neurons receive the input values and multiply them with the specific weight obtained through optimization. Thereafter, the weighted sum is transformed through the nonlinear activation function to further pass it to the neurons of the next layer. The CNN architecture is represented in Figure 4 and the pseudocode is given in Table S1. Through this procedure, the network will learn through the optimal set of weights between the neurons in the adjoining layers and will maximize the network performance, which would help the neurons focus on specific patterns in the data. In the final layer, the parameters are passed through the SoftMax function, which transforms them into probabilities that sum to 1, as shown in Equation (1).

$$\hat{p}_k = \sigma(s(x))_k = \frac{\exp(s_k(x))}{\sum_{j=1}^{K} \exp(s_j(x))} \quad (1)$$

where K is the total number of classes, $s(x)$ is a vector with the weight of each class for instance x, and $\sigma(s(x))_k$ is the calculated probability of x belonging to class k as per the assigned weight. Although there have been several works that have been conducted with SDM using shallow networks containing a single hidden layer, their performance is not as good as that of the multi-layer networks [62]. The authors of [63], used a multi-layered network for distribution modelling and achieved a better performance compared to the single-layered networks.

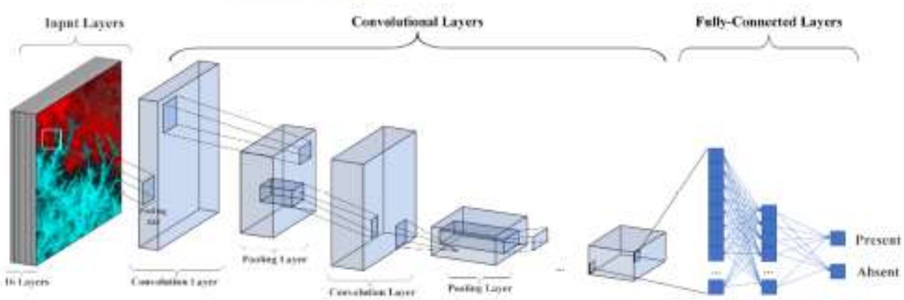


Figure 4. CNN architecture.

A Convolutional Neural Network (CNN) is a type of deep learning-based model for processing multidimensional data that follows a grid pattern [60]. The model is developed in such a way that the algorithm learns and adapts to the spatial hierarchies of features by itself from the lower to the higher levels of the pattern. Mathematically, it is composed of three layers or building blocks: convolution, pooling, and fully connected layers. Feature extraction is conducted using the first two layers and mapping the extracted features to the output is conducted by the third layer.

Convolution is used for feature extraction, in which a kernel is applied to an input tensor. A feature map is thus obtained through the product of kernel elements and tensor input. The procedure is then repeated on multiple kernels to obtain random feature maps that represent different feature extractors. The hyperparameters involved in convolution operations are the size and number of kernels. The size could be anything from 3×3 to 5×5 to 7×7, and the kernel could be chosen randomly.

A pooling layer offers downsampling functionality that decreases the dimensionality of the feature maps to achieve translation invariance to the alterations and the biases incorporated and thus helps in reducing the number of learnable parameters. There are two types of pooling operations, namely Max Pooling and Global Average Pooling [64]. The first one extracts speckles from the input feature maps and offers maximum values in each of the speckles and leaves the remaining values unattended. The second one downsamples a feature map with a size equaling product of height and width into an array of a one cross one by averaging the elements of each feature map by retaining the depth of the feature map. The advantage of Global Average Pooling lies in reducing the number of learnable parameters along with offering the CNN with variable sized input.

The features extracted by the convolution layers followed by downsampling by the pooling layers are mapped using a subset of fully connected layers to the final output of the network. The fully connected layer is executed with the ReLU function [65,66]. Mathematically, the Rectifier can be described as:

$$f(x) = x^{+} = max(0, x) \tag{2}$$

where x is the input to the neuron. A unit employing the Rectifier is known as the Rectifier linear unit (ReLU).

The performance evaluation of the model is conducted by tuning the learnable parameters, kernels, and weight by a loss function through the forward propagation followed by updating these parameter values through an optimization algorithm either by backpropagation or gradient descent.

2.5. Model Validation

All of the input parameters are resampled in a single grid size of 100 m and are converted into the same file format. Out of the in-situ occurrences of *Rhododendron arboreum* at ground locations, only 70% of the data were used in calibrating the model, whereas the remaining 30% of the data were used to test the model. In any type of modelling, performance evaluation is an essential task. In terms of validation of species probability distribution, the AUC (Area Under ROC (Receiver Operating Characteristics) Curve) is one of the most used performance evaluation metrics [67]. The primary application of the ROC curve is in the threshold independent assessment that characterizes the model performance at various discrimination thresholds. This application was found in raster-based studies focusing on predicting land use and land cover, species distribution modelling, risk assessment, and other probability mappings.

The AUC is generated by plotting the True Positive Rate (TPR) versus the False Positive Rate (FPR) at varied thresholds. The TPR is also known as sensitivity, probability of detection, or recall, and the FPR is also known as the probability of false alarm. Therefore, an accurate model will generate a ROC curve away from the 1:1 line, and a less accurate model will have a ROC curve towards the 1:1 line. The range of the AUC varies from 0 to 1. The closer the value is to 1, the better the prediction is. The plots can be described mathematically as:

$$\text{TPR or Sensitivity or Recall or Probability of Detection} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \quad (3)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100 \quad (4)$$

$$\text{FPR or Probability of false alarm} = 1 - \text{Specificity} \quad (5)$$

Here, TP stands for true positive, and FP is false positive, where specificity is also termed the true negative rate. The TPR provides the percentage of correctly predicted instances of species other than rhododendron, whereas specificity provides the percentage of correctly predicted instances of rhododendron distribution.

Thereafter, Cohen's kappa is also calculated to support the AUC value. Being one of the most popular performance evaluation indices, it is considered to be less complex and dependent on prevalence. The kappa value ranges from −1 to +1, where +1 indicates the perfect agreement. Other than kappa, the True Skill Statistic (TSS) is also incorporated, as it corrects the unimodel dependency of kappa. TSS is widely used in ecology, and it can be explained as

$$\text{TSS} = \text{Sensitivity} - \text{Specificity} - 1 \quad (6)$$

3. Results

The spatial species distribution is highly associated with regional environmental conditions, climatic variability, and land use [68,69]. The species distribution is simulated using the correlation models between the dependent as well as independent parameters. These models were generated through the presence-only data, presence/absence data, and pseudo-presence locations of the species. A total of 16 input parameters were taken from different satellite observations to model the potential distribution of Rhododendron arboreum confined to the current study area. The in situ species locations were recorded to be used as training and testing data and to retrieve the corresponding ecological and climatic satellite observations [70,71]. To understand the overall objective of the work, analysis was conducted on the distribution of the input parameters followed by the intercorrelation between them.

3.1. Assessing the Distribution of Input Parameters

As several input parameters were used from different satellite observations, a statistical downscaling was first performed to achieve a common spatial resolution of 100 m to be given as the model input. The statistically resampled images of different input parameters are shown in Figure 5. The yearly average was taken for each parameter to incorporate the overall variation throughout the year. AAI was found to range between −2.196 to 0.071, in which the higher values were distributed where the higher altitudes have an upper limit of 7771 m and a lower limit of 379 m from mean sea level. This drastic variation in elevation permits rare species to grow in an extraordinary ecosystem, and it is the main reason for the higher species heterogeneity in this region. The EVI and NDVI derived from the Sentinel-2 optical data varies from −0.19 to 0.77 and −0.28 to 0.83, respectively. The lower and lower-middle altitude locations tend to have higher NDVI/EVI values than the higher altitudes. As the LST has a linear relationship with altitude, a drastic variation in the upper and lower limits of LST can be found to be in the range of 242.2 to 306.1 Kelvin, respectively, which reflects the presence of glaciers on top of the Himalayan mountains. Due to the presence of dense vegetation at the lower altitudes, the values of water vapour, fPAR, LAI, and ET are also higher in the foothills and lower in the upper Himalayas. At the same time, atmospheric constituents like ozone, nitrogen dioxide, and carbon monoxide also show high values at the lower altitude, where the vegetation density and the presence of anthropogenic factors contributing to their concentration are relatively higher. However, the concentration of SO_2 and HCHO are very low and are evenly distributed throughout, which directly relates to industrial and transportation activities, which is very low in these areas. The SO_2 varied from -0.0002 to 0.0004 mol/m^2, and HCHO varied from -0.00005 to 0.00025 mol/m^2.

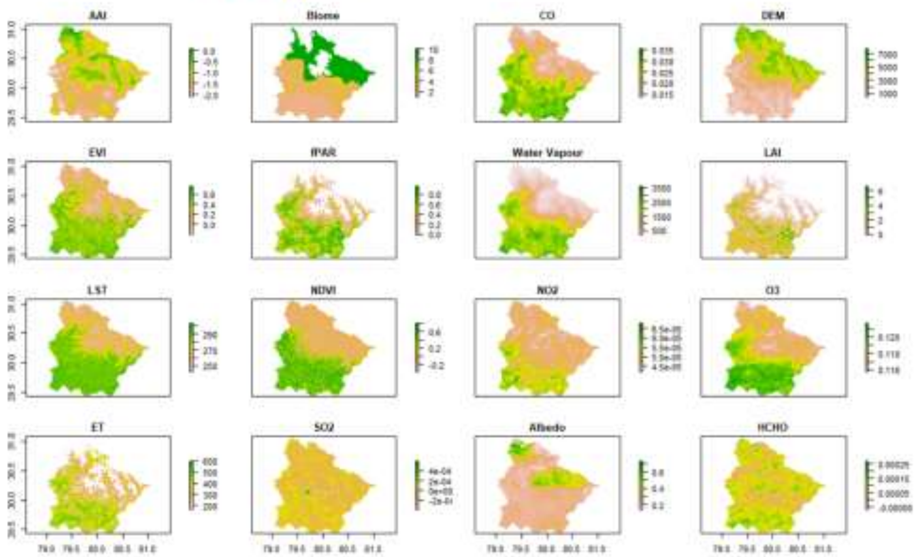


Figure 5. Spatial Distribution of various input parameters.

3.2. *Understanding Parameter Intercorrelation*

A correlation matrix plot was drawn as depicted in Figure 6 to understand the relationship between the input parameters. Total fifteen parameters except for the biome layer, which is in the vector form, were used in the correlation matrix. A highly linear or nonlinear relationship shows a relation/dependency or non-relation/non-dependency between the parameters. The representation can be explained in terms of the values varying from -1 to $+1$. The $-$ve value represents the negative relationship, and the +ve value describes the positive relationship. The depiction of the negative relationship is in orange, where the higher correlation value is visualized through the steeper circular shape, and vice versa for the positive relationship. No relationship is represented by the correlation value of zero that is represented by a perfect circular shape, and the colour becomes whitish. It can be observed that many parameters are related to each other. A highly linear relationship exists between Sentinel-5p based ozone and carbon monoxide as well as with ozone and water vapour with a correlation value of 0.99. It can also be observed that the linear relationship of DEM with NDVI, LST, and water vapour is very high, which shows the variation in the local geometry and the influence of regional ecological and climatic parameters. A lower correlation value is observed between the atmospheric parameters and the vegetation indices, especially for EVI, LAI, and ET.

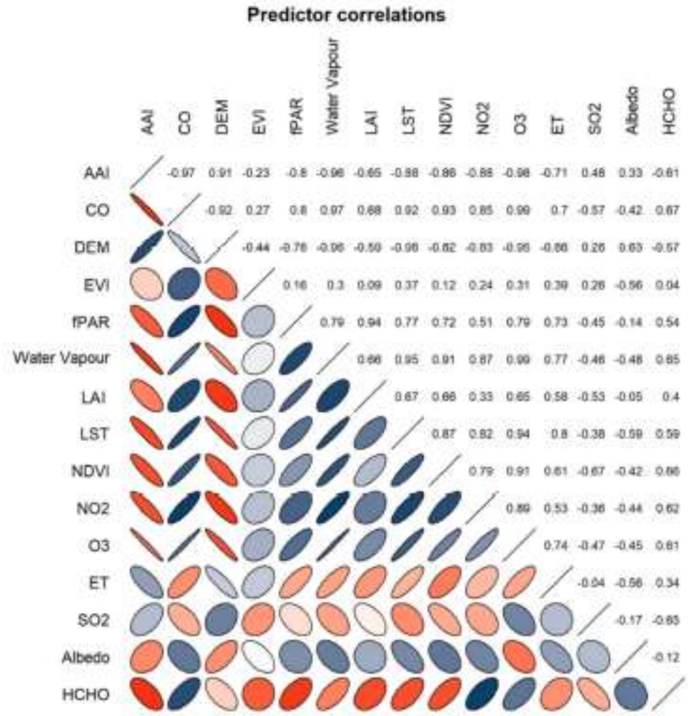


Figure 6. Square correlation matrix between input parameters.

3.3. Spatial Distribution of *Rhododendron arboreum*

To simulate the potential distribution of *Rhododendron arboreum*, linear and a nonlinear SDM were used in the current study using 16 a priori input parameters. The input parameters were taken from different satellite observations followed by data resampling to match their spatial resolution to achieve a standard resolution of 100 m. The probability distribution is classified into four classes, namely very low, low, high, and very high, according to their distribution. The presence, based only on the BIOCLIM model, predicts the probability distribution of species using a linear correlation, as shown in Figure 7a. Apart from a well-established presence only algorithm, a deep learning-based convolution neural network model was used to establish a nonlinear relationship between the input parameters to predict the probability of species distribution. A CNN based architecture was used to train the model according to the known locations and was fitted on different layers. The perfect combination of layers and activation functions was then used to predict the species distribution Figure 7b.

3.4. Model Validation and Comparison

The accuracy or performance of the probability distribution of the incorporated models was compared using the AUC, TSS, and kappa coefficient that characterize the performance of the models with an in situ validation dataset. Table 2 shows the statistical performance for BIOCLIM and CNN based the probability distribution of *Rhododendron arboreum*. A lower AUC value was obtained by the BIOCLIM models, which is 0.639, based on the in situ points reserved for the validation purpose. The AUC values for the CNN-based probability distribution was found to be 0.917, which is considered to be very good compared to the BIOCLIM that was given an AUC of 0.68. In addition to the AUC value, TSS and kappa, with values 0.652 and 0.94, respectively, also gave support to the applicability of CNN in comparison to conventional SDM's such as BIOCLIM. These values showcase the superiority of deep learning models for species probability distribution using the given set of ecological and bioclimatic parameters.

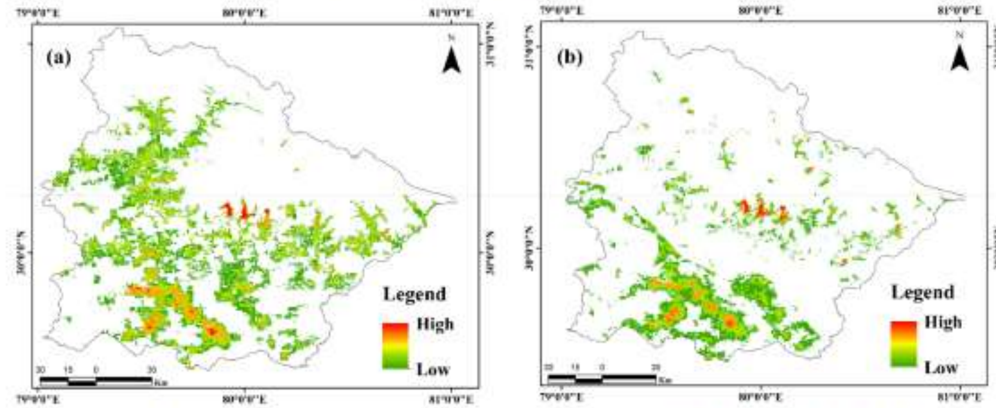


Figure 7. Probability distribution of Rhododendron arboreum using (a) BIOCLIM and (b) CNN Models.

Table 2. Statistical performance analysis of BIOCLIM and CNN.

	BIOCLIM	CNN
AUC	0.68	0.917
Kappa	0.76	0.94
TSS	0.44	0.652

4. Discussion

Understanding the dynamics and distribution of the forest ecosystem is a crucial step towards biodiversity conservation. The recent advancements in statistical machine learning models and the availability of reliable datasets help researchers build policies towards conservation and sustainable solutions to achieve this conservation. The SDMs came into existence in the mid-1980s with very limited ecological and climatic datasets, and from there, they have evolved with the regular integration of newer and more reliable datasets.

Recently, a great variety of SDMs have been used to model the distribution of species, in which the most popular are the ones that are based on the non-linear modelling approach followed by statistical and rule-based methods. Among machine learning models, Maxent is widely used due to its user-friendly interface and simple background algorithm. Maxent accuracy, as per [72], has provided robust predictions with an AUC of 0.75 and for BIOCLIM, an AUC of 0.65, whereas a similar result is achieved by [73] with an AUC of 0.73 and 0.66 for Maxent and BIOCLIM, respectively. Another well performing algorithm is the Boosted Regression Tree (BRT), which performs slightly better than Maxent. As per [74], they archived an overall AUC of 0.81 using BRT. Statistical and rule-based methods are among the conventional approaches that are not consistent with the changes in regional ecology.

The BIOCLIM model is one of the oldest yet most used SDMs due to its simple algorithm and easy parameterization [75]. It is based on a linear bioclimatic envelop model that assigns equal weight to each variable and offers a binary prediction. A pervious study indicated the use of BIOCLIM to predict species distribution, and similar studies were conducted in the past. The introduction of machine learning and its integration in SDM revolutionized the probability distribution modelling approach. The variety in modelling approaches and the increased number of datasets has made this model one of the most globally accepted SDMs. There is a growing concern for the establishment of the nonlinear relationship between the bioclimatic parameters through innovative approaches such as deep learning-based models. It has been observed that the BIOCLIM model is overestimates the species distribution and the higher probability of species occurrence at the higher altitude. Moreover, it has been observed that the distribution pattern of the predicted *Rhododendron arboreum* distribution using CNN architecture is quite different at some places than from conventional BIOCLIM models. The current work proposed deep learning-based CNN architecture for probability distribution modelling and proved to perform better than the traditional BIOCLIM model. There was an underestimation of species distribution observed in CNN than BIOCLIM. The distribution probability in CNN was precise, and it was found that the majority of *Rhododendron arboreum* is distributed in the southern part of the study area where the vegetation density is high. Some high probability patches at the higher altitudes are commonly predicted by both the models.

However, the scalability of the current outcome needs to be tested on a global scale. Apart from this, some limitations, namely uncertainties associated with the input data, the assigned weights, and some important biotic parameters, need to be handled for future work. An ensemble of all of these available methods needs to be explored in the future to establish the linear and nonlinear relationship between the dependent parameters to predict any one species out of the multiple species that are available in a location. Additionally, there is a need to perform sensitivity analysis to understand variable impact on the target variable, instead of forcing all of the variables into the model. This would reduce the algorithm complexity and computational demand.

In spite of achieving significant accuracy and popularity in the field of correlation modelling, there is still not a single algorithm that can be recommended. Deep neural networks are showing more promising results, but they are still to be tested in different ecological settings.

5. Conclusions

This study is a novel approach towards establishing a CNN architecture and testing the performance of CNN in SDM and its comparison with other well established SDMs namely, BIOCLIM. This study was conducted on the foothills of the Himalayas, where the altitudinal variation is very drastic and varies from 416 to 7801 m above mean sea level. This high-altitude Himalayan ranges constitutes a heterogeneous ecosystem and is home to many rare/endangered, medicinally, and economically important plant species. One of the major economically and medicinally important plant species, Rhododendron arboreum, was tracked and mapped in this study using different SDMs. Based on its occurrence and several ecological and bioclimatic satellite-based observations, the probability distribution of the Rhododendron arboreum was established. The CNN based probability distribution model outperformed the presence only based BIOCLIM model with an AUC score of 0.917. The CNN based prediction was also found to be more precise and accurate and with significantly less overestimation, whereas the AUC values of the BIOCLIM model were found to be 0.68 with a high overestimation. The superiority of CNN implies the role of nonlinear parameters in predicting the probability of species distribution. The scalability of the current solution on a global scale, the addition of some other important parameters, and an ensemble of all of the available SDMs need to be explored in future work. An increase in the presence of the Rhododendron species is an indication of strong soil retention, which, in turn, is fruitful for other vegetation to grow and flourish. Apart from this, an increase in the green vegetation fraction and a decrease in shade fraction was found to be associated with a higher likelihood of Rhododendron. This increased likelihood using the models would offer researchers an opportunity to understand the vegetation distribution and to contribute to the restoration of the ecology and biodiversity conservation in the protected areas so that a provision could be established for sustainable ecosystem services.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/rs13163284/s1. Table S1: Methodology of the convolution and full connected layer.

Author Contributions: Conceptualization, data curation, formal analysis: A.A., M.K.P., and P.K.S.; funding acquisition: P.K.S.; supervision: P.K.S. and M.L.K.; validation, visualization, writing—original draft: A.A., M.K.P., and P.K.S.; writing—review and editing: A.G., P.K.S., and M.L.K. All authors have read and agreed to the published version of the manuscript.

Funding: National Mission on Himalayan Studies, G.B. Pant National Institute of Himalayan Environment (NIHE), Almora, Uttarakhand, India.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the satellite datasets used in this study are available free of charge. Links are given as follows: MODIS (https://modis.gsfc.nasa.gov/data accessed on 5 June 2021), Sentinel-2 (https://sentinel.esa.int/web/sentinel/missions/sentinel-2/data-products accessed on 5 June 2021), Sentinel-5P (https://sentinel.esa.int/web/sentinel/missions/sentinel-5/data-products accessed on 5 June 2021), ECOSTRESS (https://ecostress.jpl.nasa.gov/data), and SRTM (https://www2.jpl.nasa.gov/srtm accessed on 5 June 2021).

Acknowledgments: The authors are thankful to the National Mission for Himalayan Studies (NMHS), G.B. Pant National Institute of Himalayan Environment (NIHE) for the necessary financial assistance and support throughout this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sabin, T.; Krishnan, R.; Vellore, R.; Priya, P.; Borgaonkar, H.; Singh, B.B.; Sagar, A. Climate change over the Himalayas. In *Assessment of Climate Change over the Indian Region*; Krishnan, R., Sanjay, J., Gnanaseelan, C., Mujumdar, M., Kulkarni, A., Chakraborty, S., Eds.; Springer: Singapore, 2020; pp. 207–222.
2. Kraaijenbrink, P.D.; Bierkens, M.; Lutz, A.; Immerzeel, W. Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. *Nature* **2017**, *549*, 257–260. [CrossRef]
3. Kala, C.P. Status and conservation of rare and endangered medicinal plants in the Indian trans-Himalaya. *Biol. Conserv.* **2000**, *93*, 371–379. [CrossRef]
4. Veera, S.N.; Panda, R.M.; Behera, M.D.; Goel, S.; Roy, P.S.; Barik, S.K. Prediction of upslope movement of Rhododendron arboreum in Western Himalaya. *Trop. Ecol.* **2019**, *60*, 518–524. [CrossRef]
5. Srivastava, P. Rhododendron arboreum: An overview. *J. Appl. Pharm. Sci.* **2012**, *2*, 158–162.
6. Bhandari, M.S.; Meena, R.K.; Shankhwar, R.; Shekhar, C.; Saxena, J.; Kant, R.; Pandey, V.V.; Barthwal, S.; Pandey, S.; Chandra, G. Prediction mapping through maxent modeling paves the way for the conservation of Rhododendron arboreum in Uttarakhand Himalayas. *J. Indian Soc. Remote Sens.* **2020**, *48*, 411–422. [CrossRef]
7. Jain, A.; Pandit, M.K.; Elahi, S.; Jain, A.; Bhaskar, A.; Kumar, V. Reproductive behaviour and genetic variability in geographically isolated populations of Rhododendron arboreum (Ericaceae). *Curr. Sci.* **2000**, *79*, 1377–1381.
8. Sharma, G. *Development and Charaterization of UGMS Markers for Genetic Diversity Analysis in Rhododendron Arboreum*; Guru Kashi University: Punjab, India, 2013.
9. Chauhan, D.; Lal, P.; Singh, D. Composition, population structure and regeneration of Rhododendron arboreum Sm. temperate broad-leaved evergreen forest in Garhwal Himalaya, Uttarakhand, India. *J. Earth Sci. Clim. Chang.* **2017**, *8*, 430. [CrossRef]
10. Humboldt, A.V.; Bonpland, A. *Ideen Zu Einer Geographie Der Pflanzen Nebst Einem Naturgemälde Der Tropenländer*; Cotta: Tübingen, Germany, 1807.
11. De Candolle, A. *Géographie Botanique Raisonnée Ou Exposition Des Faits Principaux Et Des Lois Concernant La Distribution Géographique Des Plantes De L'époque Actuelle*; V. Masson: Paris, France, 1855; Volume 2.
12. Udvardy, M.F. Notes on the ecological concepts of habitat, biotope and niche. *Ecology* **1959**, *40*, 725–728. [CrossRef]
13. Priti, H.; Aravind, N.; Shaanker, R.U.; Ravikanth, G. Modeling impacts of future climate on the distribution of Myristicaceae species in the Western Ghats, India. *Ecol. Eng.* **2016**, *89*, 14–23. [CrossRef]
14. Adhikari, D.; Barik, S.; Upadhaya, K. Habitat distribution modelling for reintroduction of Ilex khasiana Purk., a critically endangered tree species of northeastern India. *Ecol. Eng.* **2012**, *40*, 37–43. [CrossRef]
15. Franklin, J. Moving beyond static species distribution models in support of conservation biogeography. *Divers. Distrib.* **2010**, *16*, 321–330. [CrossRef]
16. Vasconcelos, T.S.; Rodríguez, M.Á.; Hawkins, B.A. Species distribution modelling as a macroecological tool: A case study using New World amphibians. *Ecography* **2012**, *35*, 539–548. [CrossRef]

17. Rodríguez, J.P.; Brotons, L.; Bustamante, J.; Seoane, J. The application of predictive modelling of species distribution to biodiversity conservation. *Divers. Distrib.* **2007**, *13*, 243–251. [CrossRef]
18. Lorena, A.C.; Jacintho, L.F.; Siqueira, M.F.; De Giovanni, R.; Lohmann, L.G.; De Carvalho, A.C.; Yamamoto, M. Comparing machine learning classifiers in potential distribution modelling. *Expert Syst. Appl.* **2011**, *38*, 5268–5275. [CrossRef]
19. Elith, J.; Leathwick, J.R. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. of Ecol. Evol. Syst.* **2009**, *40*, 677–697. [CrossRef]
20. Rana, S.K.; Rana, H.K.; Luo, D.; Sun, H. Estimating climate-induced 'Nowhere to go'range shifts of the Himalayan Incarvillea Juss. using multi-model median ensemble species distribution models. *Ecol. Indic.* **2021**, *121*, 107127. [CrossRef]
21. Beaumont, L.J.; Hughes, L.; Poulsen, M. Predicting species distributions: Use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions. *Ecol. Model.* **2005**, *186*, 251–270. [CrossRef]
22. Doran, B.; Olsen, P. Customizing BIOCLIM to investigate spatial and temporal variations in highly mobile species. In Proceedings of the 6th International Conference in GeoComputation, Brisbane, Australia, 24–26 September 2001.
23. Xu, Y.; Zhou, P.-Y.; Wang, Y.; Chen, Z.-X.; Ma, R.; Yu, S.-F. Assessment of risk of introduction of pine wood nematode, bursaphelenchus xylophilus in Yunnan Province using BIOCLIM ecological niche model. *J. Yunnan Agric. Univ.* **2008**, *23*, 746–753.
24. Bhatta, K.P.; Robson, B.A.; Suwal, M.K.; Vetaas, O.R. A pan-Himalayan test of predictions on plant species richness based on primary production and water-energy dynamics. *Front. Biogeogr.* **2021**, *13*, e49459.
25. Mamgain, A.; Uniyal, P.L. Species Distribution Modelling of Rhododendron arboreum Sm.–A Keystone Species, in India and Adjoining Region. *Int. J. Ecol. Environ. Sci.* **2018**, *44*, 261–286.
26. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT press: Cambridge, MA, USA, 2016.
27. Lek, S.; Delacoste, M.; Baran, P.; Dimopoulos, I.; Lauga, J.; Aulagnier, S. Application of neural networks to modelling nonlinear relationships in ecology. *Ecol. Model.* **1996**, *90*, 39–52. [CrossRef]
28. Thuiller, W. BIOMOD–optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Chang. Biol.* **2003**, *9*, 1353–1362. [CrossRef]
29. Zhang, J.; Li, S. A Review of Machine Learning Based Species' Distribution Modelling. In Proceedings of the 2017 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), Wuhan, China, 2–3 December 2017; pp. 199–206.
30. Kumar, K. *Water Management in Himalayan Ecosystem: A Study of Natural Springs of Almora*; Indus Publishing: New Delhi, India, 1996.
31. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]
32. Tewari, A.P. Recent changes in the position of the snout of the Pindari glacier (Kumaon Himalaya), Almora District, Uttar Pradesh, India. In *Proceedings of the Role of Snow and Ice in Hydrology, Banff Symposia, September 1972*; WMO-IAHS-Unesco: Geneva, Switzerland, 1973; pp. 1144–1149.
33. Singh, K.; Rai, L.; Gurung, B. Conservation of rhododendrons in Sikkim Himalaya: An overview. *World J. Agric. Sci.* **2009**, *5*, 284–296.
34. Secretariat, G. GBIF backbone taxonomy. *Checklist Dataset* **2017**, *10*. Available online: https://www.gbif.org/dataset/d7dddbf4-2cf0-4f39-9b2a-bb099caae36c (accessed on 5 June 2021).
35. Rawat, P.; Rai, N.; Kumar, N.; Bachheti, R. Review on Rhododendron arboreum—A magical tree. *Orient. Pharm. Exp. Med.* **2017**, *17*, 297–308. [CrossRef]
36. Paul, A.; Khan, M.L.; Arunachalam, A.; Arunachalam, K. Biodiversity and conservation of rhododendrons in Arunachal Pradesh in the Indo-Burma biodiversity hotspot. *Curr. Sci.* **2005**, *89*, 623–634.
37. Chauhan, N.S. *Medicinal and Aromatic Plants of Himachal Pradesh*; Indus Publishing: New Delhi, India, 1999.
38. Watts, J.S. *When a Billion Chinese Jump: How China Will Save Mankind—Or Destroy It*; Simon and Schuster: New York, NY, USA, 2010.
39. Singh, V.K.; Ali, Z.A. *Herbal Drugs of Himalaya*; Today & Tomorrow's Printers and Publishers: Delhi, India, 1998.
40. Singh, N.; Ram, J.; Tewari, A.; Yadav, R. Phenological events along the elevation gradient and effect of climate change on Rhododendron arboreum Sm. in Kumaun Himalaya. *Curr. Sci.* **2015**, *108*, 106–110.
41. Paul, A.; Khan, M.L.; Das, A.K. *Utilization of Rhododendrons by Monpas in Western Arunachal Pradesh, India*; Assam University: Silchar, India, 2010.
42. Negi, V.S.; Maikhuri, R.; Rawat, L.; Chandra, A. Bioprospecting of Rhododendron arboreum for livelihood enhancement in central Himalaya, India. *Environ. We Int. Jouranl Sci. Technol.* **2013**, *8*, 61–70.
43. Uniyal, S.K.; Singh, K.; Jamwal, P.; Lal, B. Traditional use of medicinal plants among the tribal communities of Chhota Bhangal, Western Himalaya. *J. Ethnobiol. Ethnomedi.* **2006**, *2*, 1–8. [CrossRef] [PubMed]
44. Sharma, P.; Samant, S. Diversity, distribution and indigenous uses of medicinal plants in Parbati Valley of Kullu district in Himachal Pradesh, Northwestern Himalaya. *Asian J. Adv. Basic Sci.* **2014**, *2*, 77–98.
45. Zhasa, N.; Hazarika, P.; Tripathi, Y. Indigenous knowledge on utilization of plant biodiversity for treatment and cure of diseases of human beings in Nagaland, India: A case study. *Int. Res. J. Biol. Sci.* **2015**, *4*, 89–106.
46. Kumar, P. Assessment of impact of climate change on Rhododendrons in Sikkim Himalayas using Maxent modelling: Limitations and challenges. *Biodivers. Conserv.* **2012**, *21*, 1251–1266. [CrossRef]

47. Wisz, M.S.; Hijmans, R.; Li, J.; Peterson, A.T.; Graham, C.; Guisan, A.; NCEAS Predicting Species Distributions Working Group. Effects of sample size on the performance of species distribution models. *Divers. Distrib.* **2008**, *14*, 763–773. [CrossRef]
48. Yang, W.; Tan, B.; Huang, D.; Rautiainen, M.; Shabanov, N.V.; Wang, Y.; Privette, J.L.; Huemmrich, K.F.; Fensholt, R.; Sandholt, I. MODIS leaf area index products: From validation to algorithm improvement. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1885–1898. [CrossRef]
49. Martin, R.; Parrish, D.; Ryerson, T.; Nicks, D., Jr.; Chance, K.; Kurosu, T.; Jacob, D.J.; Sturges, E.; Fried, A.; Wert, B. Evaluation of GOME satellite measurements of tropospheric NO_2 and HCHO using regional data from aircraft campaigns in the southeastern United States. *J. Geophys. Res. Atmos.* **2004**, *109*, 1–11. [CrossRef]
50. Olson, D.; Dinerstein, E.; Wikramanayake, E.; Burgess, N.; Powell, G.; Underwood, E.; d'Amico, J.; Itoua, I.; Strand, H.; Morrison, J. Terrestrial ecoregions of the world: A new map of life on earth. *BioScience* **2001**, *51*, 933–938. [CrossRef]
51. Kamei, J.; Pandey, H.; Barik, S. Tree species distribution and its impact on soil properties, and nitrogen and phosphorus mineralization in a humid subtropical forest ecosystem of northeastern India. *Can. J. For. Res.* **2009**, *39*, 36–47. [CrossRef]
52. Kala, C.P.; Mathur, V.B. Patterns of plant species distribution in the Trans-Himalayan region of Ladakh, India. *J. Veg. Sci.* **2002**, *13*, 751–754. [CrossRef]
53. Pearson, R.G. Species' distribution modeling for conservation educators and practitioners. *Synth. Am. Mus. Nat. Hist.* **2007**, *50*, 54–89.
54. Nix, H.A. A biogeographic analysis of Australian elapid snakes. *Atlas Elapid Snakes Aust.* **1986**, *7*, 4–15.
55. Haydon, D.T.; Pianka, E.R. Metapopulation theory, landscape models, and species diversity. *Ecoscience* **1999**, *6*, 316–328. [CrossRef]
56. Guisan, A.; Thuiller, W. Predicting species distribution: Offering more than simple habitat models. *Ecol. Lett.* **2005**, *8*, 993–1009. [CrossRef]
57. Parthasarathy, U.; Saji, K.; Jayarajan, K.; Parthasarathy, V. Biodiversity of Piper in South India–application of GIS and cluster analysis. *Curr. Sci.* **2006**, *91*, 652–658.
58. Rameshprabu, N.; Swamy, P. Prediction of environmental suitability for invasion of Mikania micrantha in India by species distribution modelling. *J. Environ. Biol.* **2015**, *36*, 565.
59. Booth, T.H.; Nix, H.A.; Busby, J.R.; Hutchinson, M.F. BIOCLIM: The first species distribution modelling package, its early applications and relevance to most current MAXENT studies. *Divers. Distrib.* **2014**, *20*, 1–9. [CrossRef]
60. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
61. Guo, Y.; Liu, Y.; Oerlemans, A.; Lao, S.; Wu, S.; Lew, M.S. Deep learning for visual understanding: A review. *Neurocomputing* **2016**, *187*, 27–48. [CrossRef]
62. Fukuda, S.; De Baets, B.; Waegeman, W.; Verwaeren, J.; Mouton, A.M. Habitat prediction and knowledge extraction for spawning European grayling (Thymallus thymallus L.) using a broad range of species distribution models. *Environ. Model. Softw.* **2013**, *47*, 1–6. [CrossRef]
63. Harris, D.J. Generating realistic assemblages with a joint species distribution model. *Methods Ecol. Evol.* **2015**, *6*, 465–473. [CrossRef]
64. Lin, M.; Chen, Q.; Yan, S. Network in network. *arXiv Prepr.* **2013**, arXiv:1312.4400.
65. Agarap, A.F. Deep learning using rectified linear units (relu). *arXiv Prepr.* **2018**, arXiv:1803.08375.
66. Schmidt-Hieber, J. Nonparametric regression using deep neural networks with ReLU activation function. *Ann. Stat.* **2020**, *48*, 1875–1897.
67. Hanley, J.A.; McNeil, B.J. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **1982**, *143*, 29–36. [CrossRef]
68. Pandey, P.C.; Anand, A.; Srivastava, P.K. Spatial distribution of mangrove forest species and biomass assessment using field inventory and earth observation hyperspectral data. *Biodivers. Conserv.* **2019**, *28*, 2143–2162. [CrossRef]
69. Anand, A.; Malhi, R.K.M.; Pandey, P.C.; Petropoulos, G.P.; Pavlides, A.; Sharma, J.K.; Srivastava, P.K. Use of Hyperion for Mangrove Forest Carbon Stock Assessment in Bhitarkanika Forest Reserve: A Contribution Towards Blue Carbon Initiative. *Remote Sens.* **2020**, *12*, 597. [CrossRef]
70. Malhi, R.K.M.; Anand, A.; Srivastava, P.K.; Kiran, G.S.; Petropoulos, G.P.; Chalkias, C. An Integrated Spatiotemporal Pattern Analysis Model to Assess and Predict the Degradation of Protected Forest Areas. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 530. [CrossRef]
71. Malhi, R.K.M.; Anand, A.; Mudaliar, A.N.; Pandey, P.C.; Srivastava, P.K.; Sandhya Kiran, G. Synergetic use of in situ and hyperspectral data for mapping species diversity and above ground biomass in Shoolpaneshwar Wildlife Sanctuary, Gujarat. *Trop. Ecol.* **2020**, *61*, 106–115. [CrossRef]
72. Elith, J.; Graham, C.H.; Anderson, R.P.; Dudík, M.; Ferrier, S.; Guisan, A.; Hijmans, R.J.; Huettmann, F.; Leathwick, J.R.; Lehmann, A. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **2006**, *29*, 129–151. [CrossRef]
73. Graham, C.H.; Elith, J.; Hijmans, R.J.; Guisan, A.; Townsend Peterson, A.; Loiselle, B.A.; Group, N.P.S.D.W. The influence of spatial errors in species occurrence data used in distribution models. *J. Appl. Ecol.* **2008**, *45*, 239–247. [CrossRef]
74. Hallman, T.A.; Robinson, W.D. Comparing multi-and single-scale species distribution and abundance models built with the boosted regression tree algorithm. *Landsc. Ecol.* **2020**, *35*, 1161–1174. [CrossRef]
75. Busby, J.R. BIOCLIM-a bioclimate analysis and prediction system. *Plant. Prot. Q.* **1991**, *61*, 8–9.

forests

MDPI

Article

Statistical Unfolding Approach to Understand Influencing Factors for Taxol Content Variation in High Altitude Himalayan Region

Ayushi Gupta [1], Prashant K. Srivastava [1,*], George P. Petropoulos [2] and Prachi Singh [1]

[1] Remote Sensing Laboratory, Institute of Environment and Sustainable Development, Banaras Hindu University, Uttar Pradesh 221005, India; ayushi.gupta10@bhu.ac.in (A.G.); prachisngh246@gmail.com (P.S.)
[2] Department of Geography, Harokopio University of Athens, El. Venizelou 70, Kallithea, 17671 Athens, Greece; gpetropoulos@hua.gr
* Correspondence: prashant.just@gmail.com

Abstract: Taxol drugs can be extracted from various species of the taxaceae family. It is an alkaloid (metabolic product) used for the treatment of various types of cancer. Since taxol is a metabolic product, multiple aspects such as edaphic, biochemical, topographic factors need to be assessed in determining the variation in Taxol Content (TC). In this study, both sensor-based hyperspectral reflectance data and absorption-based indices were tested together for the development of an advanced statistical unfolding approach to understand the influencing factors for TC in high altitude Himalayan region. Seriation analysis based on permutation matrix was applied with complete linkage and a multi-fragment heuristic scaling rule along with the common techniques such as Principal Component Analysis (PCA) and correlation to understand the relationship of TC with various factors. This study also tested the newly developed taxol indices to rule out the possibility of overlapping of TC determining bands with the foliar pigment's wavelengths in the visible region. The result implies that *T. wallichiana* with a high TC is found more in its natural habitat of deep forest, relating it indirectly to elevation in the case of the montane ecosystem. Taxol is the most varying parameter among the measured variables, followed by hyperspectral Taxol content (TC) indices such as TC 2, TC 5, and carotenoids, which suggests that the indices are well versed to capture variations in TC with elevation.

Keywords: taxol; sensor-based indices; biophysical variables; biochemical variables; hyperspectral; principal component analysis; seriation analysis

Citation: Gupta, A.; Srivastava, P.K.; Petropoulos, G.P.; Singh, P. Statistical Unfolding Approach to Understand Influencing Factors for Taxol Content Variation in High Altitude Himalayan Region. *Forests* **2021**, *12*, 1726. https://doi.org/10.3390/f12121726

Received: 8 October 2021
Accepted: 1 December 2021
Published: 7 December 2021

1. Introduction

Recent studies have shown that the turnover in tree species composition across edaphic and elevational gradients can be strongly correlated with the functional traits [1]. These factors affect plant growth via various means and can be used to characterize different ecosystems. The major determining components of vegetation include biochemical constituents that are central to their physiological form and function, along with water, chlorophyll, and accessory pigments, nitrogen, cellulose, starch, sugars, lignin, and protein. These are the mandatory parameters for describing the nutritional status of any tree of a particular ecosystem [2,3], while the secondary metabolites such as terpenes, sesquiterpenes, phytosterols, etc., are more useful to humans [4], which makes the plant economically valuable.

The majority of studies have been carried out to acknowledge and retrieve these determining variables and the effects on vegetation using various models and remote sensing techniques, but the relative effects of all these factors have not been addressed intricately with proper research findings [5]. The spatial and temporal variation of these properties offers great help in understanding and evaluating physiological conditions such as photosynthesis, evapotranspiration, secondary metabolites formation, and deriving plans for the conservation of the ecosystem [2,6]. Now, researchers are focusing on covariation trait studies to determine definite functional indicators for ecological and biodiversity conservation [7,8].

Taxus wallichiana Zucc. is a tree species that belongs to the family Taxaceae and is popularly known as the Himalayan yew, which is globally distributed in Europe, North America, North India, Pakistan, China, and Japan. In Asia, its variation and availability extend from Afghanistan through the great Himalayas to the Philippines, and it is widely distributed in countries such as Pakistan and India. Recently, it gained widespread attention on a global scale because its leaves and bark were found to be rich in taxol, which is a potential anti-cancerous drug [9]. Taxol is known to have first been isolated from the bark of *Taxus brevifolia*, and since then, taxol and related bioactive taxoids have been reportedly found in the various other species of the same genus Taxus. Due to the overexploitation of this group of species, it is currently endangered as per IUCN and on the verge of extinction. Moreover, several species are disappearing at an alarming rate mainly at higher altitudes due to over-harvesting, habitat destruction, and abrupt climate change [10,11]. Altitudinal variation influences the ecological factors and, thus, the ecosystem. Factors including soil nutrients, precipitation, and mean temperature directly or indirectly affect the secondary metabolite amount and biological activities of the plants [7].

The extraction and estimation of secondary metabolites such as TC are always expensive, time-consuming, and tedious. The non-destructive method of taxol estimation for conservation and planning becomes vital. A few researchers such as Kokaly et al. [1] have characterized the plant phenolics (another secondary metabolite) [12] to their hyperspectral signature at a 1660-nanometer wavelength. Phenolics are characterized using continuum removal, which is a technique used to isolate and analyze the features in reflectance spectra acquired using hyperspectral sensors. Hyperspectral remote sensing (HSR) is a new dimension of remote sensing with a higher number of band data in a continuous form that gives fine resolution to obtain detailed information on the object [13]. Many scientists have characterized different species in the same area using HSR [14]. Keystone species conservation is the next logical step that can be brought by HSR [15]. 'Curse of dimensionality' is the phrase used for high dimensional hyperspectral data. This problem can be remedied by indices used for the retrieval of a particular parameter. These indices are easy to use and require less time and a less sophisticated system to compute [16].

The canopy confounding variables such as foliar nitrogen, chlorophyll, cellulose, etc., are successfully estimated using vegetation indices when applied to remote sensing data. The spectral wavelengths region near 550 and 700 nm, as well as the red-edge region (680–780 nm), have been utilized for assessing chlorophyll by many researchers in hyperspectral remote sensing [17–19]. Wang et al. [20] estimated nitrogen accurately in cases of broadleaf, needle leaf, and mixed forests plots using Normalized Difference Nitrogen Index (NDNI) centered at 1510 nm. However, an indirect relationship occurs between nitrogen and chlorophyll that generates a correlation between Near InfraRed (NIR) reflectance (800–850 nm) and canopy foliar mass-based nitrogen concentration [21]. Similarly, reflectance in the visible wavelengths 400–700 nm is dominated by absorptions from foliar pigments [22]. Among the pigments, chlorophyll a and b have the strongest effect over absorption in the visible region, followed by carotenoids and anthocyanins [23]. Hence, more extensive research is required to characterize any metabolite apart from foliar pigments in the visible region.

In this study, an effort has been made to assess the efficiency of different hyperspectral indices developed to understand TC variations. It also included various statistical unfolding techniques such as covariation, correlation, and the extent of various edaphic, topographic, biochemical properties, and sensor-based indices values for understanding TC variations.

2. Materials and Methodology

2.1. Study Area

Pindari glacier, which is situated in the Central Himalaya of Almora District of Uttaranchal state, was used as a study area for this research. Pindari glacier spread around the length of 5 km within an elevation range of 2400–3000 m. The climate of this region is categorized into the following three seasons: winters, summers, and monsoon. The long cold winter range from October to March with temperatures reaching below freezing point. In contrast, the maximum temperature seen in summers is around 30 °C. Mostly cloudy conditions exist throughout the monsoon months (June to September) because of the disturbances of western regions. The average annual rainfall is 930 mm, which mainly occurs during July and August [24]. Broadly, the area is divided into two climatic zones that could be categorized as (i) Lower montane zone: elevation range of 1800–2400 m above mean sea level (amsl), and (ii) Upper montane zone: elevation range of 2400–3000 m amsl. There is more precipitation in the upper zone and is more in terms of snowfall than showers [25]. The vegetation of the Pindari region comprises Pinus, Acer, Juglans, Cupressus, Quercus, Taxus, Berberis, and Rhododendron, which can be found around the region of Phurkia and the Pindari Glacier, as shown in Figure 1. This region is mostly covered by dense forests with high availability of medicinally important species [26–28]. The survey mainly focused on the collection of medicinally important species, i.e., *Taxus wallichiana*.

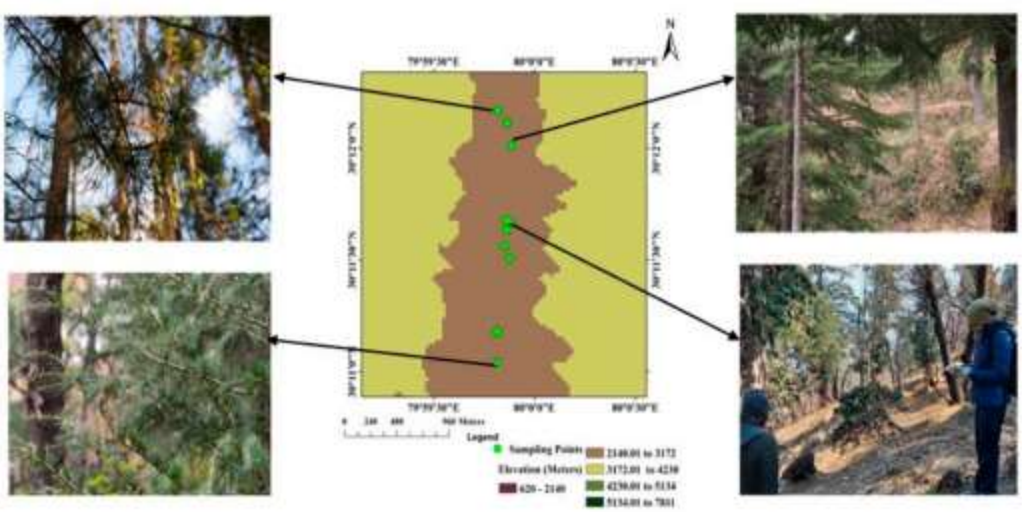


Figure 1. Sampling locations of T. wallichiana in Nanda Devi Biosphere Reserve.

2.2. Sample and Radiometer Data Collection

The sample was collected on 26 September 2019 and 29 September 2019 at different locations along with a Global Positioning System (GPS) coordinate that varied between altitudes of 3039 and 2292 m in the Nanda Devi Biosphere Reserve (NDBR) in the state of Uttarakhand, which is located in the Western Himalayan Highland Biogeographic Zone. The samples were kept in a Ziploc bag for the next few hours. The samples collected were then crushed and stored in liquid nitrogen until the immediate analysis. Hence, all the values reported after analysis are represented in terms of fresh weight (FW).

The full range (350–2500 nm) FieldSpec spectroradiometer developed by Analytical spectral devices (ASD) was used to capture the spectral reflectance of the leaves and pre-processed using ViewSpecPro software Version 6.2 by Malvern Panalytical, Malvern, United

Kingdom. ASD uses a fore-optic system to measure the spectral radiance and reflectance of any object. It distributes the signal via a fiber optic bundle to a fixed diffraction grating spectrometer. It uses three different types of detectors enabling a spectroradiometer to record the whole spectra from 350–2500 nm. Spectral reflectance is the part or a fraction of incident electromagnetic radiation that is reflected from any interface. The reflectance is plotted with a wavelength known as the reflectance spectrum or spectral reflectance curve, which is the end product of the device [29,30]. A Spectralon reference panel was used to optimize and adjust the sensitivity of the instrument.

2.3. Robustness of Indices

2.3.1. Reflectance Based Indices

A study was conducted on the Himalayan region for the development of the best Taxol indices [31]. Three different filtering techniques were applied, namely, Savitzky and Golay (S. Golay), Fast Fourier transformation, and Average Mean Filter, prior to feature selection. S. Golay uses simplified least-square-fit intricacy for smoothing. A mean filter takes the mean spectral value of nearest points within the considered window as the new value of the middle point of the window. The Fourier domain digital filter is a simple trapezoid characterized by four indices (N1, N2, N3, N4). Digital filtering is implemented simply by multiplying the Fourier domain signal by the appropriate filter function, that is, the signal points between N1 and N2 multiplied by y (value of the slope). The processed spectra after each filter application on *T. wallichiana* spectra were then applied with feature selection (first derivative). From the transformed spectra, the absorbance region at certain wavelengths was selected. The reflectance file of *T. wallichiana* spectra in text format, measured taxol content along with the wavelength selected were taken as inputs in the Automated Radiative Transfer Models Operator (ARTMO) model and the two band indices, suitable for TC estimation, were developed.

The indices developed using Average Mean smoothened wavelength, revealed a significant correlation with the measured taxol values. The five most appropriate taxol indices were selected which were developed by Average-Mean filtered wavelengths as listed in Table 1.

Table 1. Selected Taxol indices from Gupta et al. where R is reflectance band [31].

SI No.	Reflectance Based Taxol Indices
1	TC 1 = (R426 − R421)/(R426 + R421)
2	TC 2 = (R415 − R421)/(R415 + R421)
3	TC 3 = (R601 − R608)/(R601 + R608)
4	TC 4 = (R421/R426)
5	TC 5 = (R415/R421)

Since taxol detection using hyperspectral data is not a phenomenon that has been explored much, all the wavelengths suggested by Gupta et al. [31] obtained from three different filtering techniques were considered. This was performed to rule out any possibilities of missing out even small absorption peaks that could indicate a taxol presence on hyperspectral data.

2.3.2. Absorption Based Indices

Continuum removal is referred to as baseline normalization and has been commonly used in laboratory infrared spectroscopy. This technique is an estimate of the other absorptions present in the spectrum. In that sense, continuum removal is most often performed on absorption features.

Continuum removal was applied to the selected absorption features. Continuum removal normalizes reflectance spectra in order to allow for a comparison of individual absorption features from a common baseline [32]. The continuum is a convex hull fitted over the top of a spectrum to connect the local spectrum maxima.

$$Rc'_{(\lambda i)} = \frac{R_{(\lambda i)}}{R_{c(\lambda i)}} \quad (1)$$

Here, in Equation (1), the continuum-removed reflectance is $Rc'_{(\lambda i)}$, the reflectance value is $R_{(\lambda i)}$ for each waveband in the absorption pit, and the reflectance level of the continuum line (convex hull) is $R_{c\ (\lambda i)}$ at the corresponding wavelength. The first and last spectral data values are on the hull and, therefore, the first and last values of the continuum removed spectrum is equal to 1. This process enhances the absorption pits' output, whose values are between 0 and 1 [33]. Three variables were calculated from the continuum removed absorption features, viz. Continuum removed derivative reflectance (CRDR), band depths (BD), and band depth ratio (BDR). Collectively, this has been termed spectral feature analysis [34].

Processing Routines in IDL for Spectroscopy Measurements (PRISM) have the feature for automated spectral feature analysis [1]. Using this feature section option, the user can select the initial and final continuum endpoints on the spectrum to be analyzed. The PRISM software applies continuum removal to each spectrum separately and derives the spectral feature parameters (e.g., center, depth, width, area, etc.). PRISM performs continuum removal twice and gives the feature parameters as follows: (1) selection of start and endpoints; and (2) an automatically attuned set of continuum endpoints. PRISM searches for improved continuum endpoints on both sides of the absorption feature, by searching for nearby channels that have continuum-removed values higher than the initial endpoints. The new endpoint channels are referred to as the adjusted endpoints, as shown in Figure 2b. This software is added as an extension to ENVI 5.1 Aliso Viejo, CA, USA. This software can be downloaded from (https://pubs.usgs.gov/of/2011/1155/ accessed on 12 February 2021). This corresponding wavelength absorption feature area was used to develop absorption-based indices.

2.4. Soil Moisture and LST

The soil moisture (SM) and soil temperature are better known as Land Surface Temperature (LST). The in-situ measurements during sampling were carried out using Steven's HydraGo instrument. HydraGo is a rugged SM sensor that measures the dielectric spectrum of the soil based on the 'dielectric impedance' at a 50 MHz radiofrequency (https://stevenswater.com/products/hydrago-s/ accessed on 12 February 2021). The reflected signals measure the soil dielectric permittivities that correspond to the SM and bulk soil electrical conductivity (EC). The device communicates wirelessly with the HydraMon app using Bluetooth. The app displays soil moisture content, temperature, conductivity, and dielectric permittivity for immediate viewing. The date and time of each measurement were recorded along with this measurement and the GPS location was measured using a handheld Garmin GPS receiver.

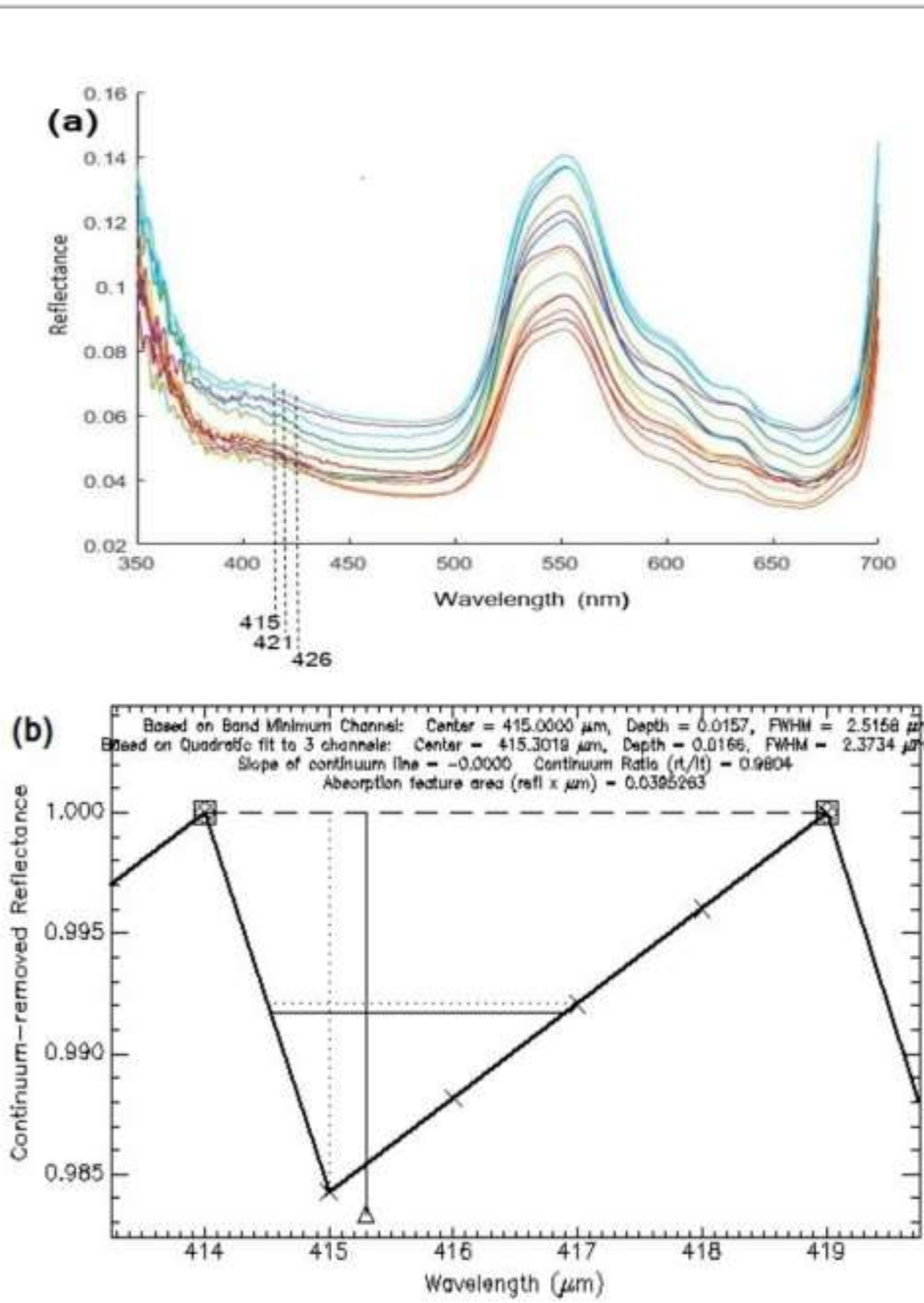


Figure 2. Spectral feature analysis (**a**) sectional view of the spectra highlighting small dip (**b**) dip (feature) center position with feature depth and width highlighted after application of continuum removal for a single sample spectrum.

2.5. Determination of Chlorophyll (TCC), Total Phenolic Content (TPC), and Taxol Content (TC)

For the estimated total chlorophyll content (TCC), a crushed leaf sample was homogenized with 2 mL of acetone (80%) and then centrifuged for 10,000× rpm for 15 min at 4 °C. An amount of 0.5 mL of supernatant was taken from the main solution and mixed with 4.5 mL of acetone. The solution mixture was analyzed for chl. a, chl. b, and carotenoids using a spectrophotometer (Thermo scientific UV–Vis Spectrophotometer). The TPC in various plant samples was estimated using the Folin–Ciocalteu (F–C) colorimetric method. The plant leaf samples were added with 2 mL of ice-cold 95% (*vol*/*vol*) methanol and then were homogenized. The samples were kept in the dark for 48 hr. The samples were then centrifuged again (13,000× rpm for 5 min). To the 150 µL supernatant of this plant extract, 900 µL of distilled water was added, followed by 225 µL of F–C reagent, was added to the solution and it was permitted to stand still for 5 min at room temperature. Then, 1.125 mL of 2% sodium carbonate was added and mixed thoroughly. Along with the samples, the blank was also prepared without the supernatant plant extract but with the other entire constituent. The prepared samples and the blank were set aside in the dark for 15 min at room temperature. The absorbance of the samples and the blank were noted using a spectrophotometer @750 nm. The TPC was calculated with a standard curve based on gallic acid. The TPC results were expressed in milligrams of gallic acid equivalent (GAE) per gram fresh weight (FW) (mg GAE/g FW) [35,36].

For TC, 1 g of crushed leaves was deflated with hexane using sonication. The hexane portions were discarded, and aliquots of methanol were concentrated using a rotary evaporator, extracted in chloroform, then dried under reduced pressure using a rotary evaporator, and then re-dispersed in methanol (1 mL). Taxoid standard paclitaxel (Sigma, St. Louis, MI, USA) as used as a standard in HPLC for quantification. The working solution of paclitaxel was prepared from standard methanol. The UV-DAD scanned acquisitions of Taxol were performed at 230 nm. The percentage of Taxol was calculated using Equation (2) [37].

$$\text{Taxol (\%) Content} = \frac{\text{Ar}_{\text{-sample}} \times \text{Conc}_{\text{-std.}}\left(\frac{\text{mg}}{\text{mL}}\right)}{\text{Ar}_{\text{-std.}} * 1000 \times \text{Conc}_{\text{-sample}}\left(\frac{\text{g}}{\text{mL}}\right)} \times 100 \quad (2)$$

where Ar_{std} and $\text{Ar}_{\text{sample}}$ are the areas under the peak associated with the standard or reference and sample taxoid, respectively, and $\text{Conc}_{\text{-sample}}$ and $\text{Conc}_{\text{-std}}$ are the concentrations of the sample and reference taxoid, respectively [37].

The standard methodology of collecting in-data and samples was followed as described earlier in Sections 2.2 and 2.4. The samples of the plants underwent multiple tests (Section 2.5) for obtaining ex-situ data. The ASD spectroradiometer was used as an input to extract absorption values from the spectra of the *T. wallichiana* via PRISM software [31,32]. The absorption values were then used to develop two-band absorption indices. The measured taxol content along with selected bands, reflectance-based indices values, absorption-based indices values were subjected to Pearson correlation in R studio software Boston, MA, USA. It brings out the most suitable indices for taxol estimation. The measured TC along with the selected hyperspectral indices and ex-situ data were subjected to Pearson's correlation to check their correlations. This step highlighted the correlation between the variable measured and the indices selected. Bartlett's sphericity test was applied to the correlation matrix variables to test the assumption that variances are equal across groups. The elevation for the samples collected between 3039 and 2292 m was divided into three groups based on elevation, and One-way Analysis of Variance (ANOVA) was applied to test the significance of elevation with each set of parameters. ANOVA was followed by Scheffe's test, which was applied to identify the significance of this difference among designated groups. PCA was done on SPSS Version 22 developed by IBM Armonk, New York, NY, USA. This further was applied to accomplish a significant reduction in the dimensionality of the original data set and bring the most varying variables to the foreground. Multivariate analysis was to highlighted the role of elevation and parameters for a suitable habitat for *T. wallichiana*. The flowchart depicting the methodology is shown in Figure 3.

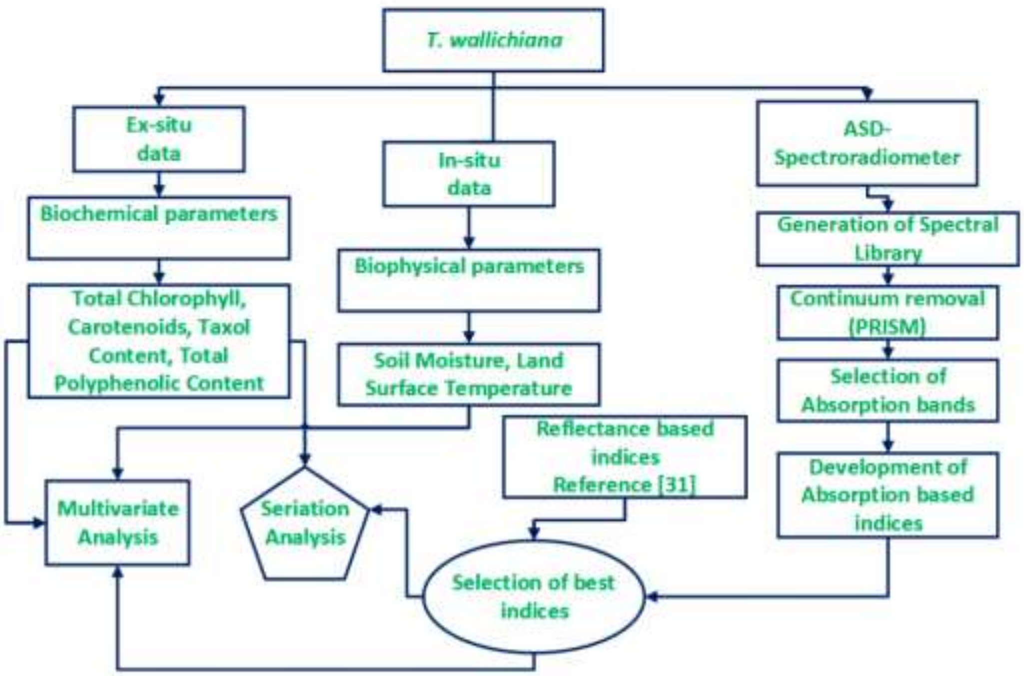


Figure 3. Flowchart for methodology opted for the current work.

3. Results and Discussion

3.1. Comparative Analysis between Indices, Selected Wavelengths, and Measured Taxol Content

The absorption indices were developed using two-band absorption values in different combinations. More than 84 absorption indices combinations were tested using the two bands. While these combinations were tried in the preliminary stage, it was observed that the indices utilizing the bands 415 and 670 nm outperformed any other indices developed utilizing other wavelengths. The indices with significant correlation are plotted in Figure 4, in what is known as a correlogram. A correlogram or Auto Correlation function is a visual way to show the serial correlation in data that change over time.

Figure 4 illustrates the correlation among all the possible absorption band values along with the absorption and reflectance indices values that showed a significant correlation with the taxol content. The absorbance wavelengths are represented with 'x' as a subscript and the absorption indices values are represented with 'i', while the reflectance-based indices are represented with TC. Figure 4 shows that the measured taxol content (Ob) showed the nearest positive correlation with reflectance-based indices TC 2 ($r = 0.741$) and TC5 ($r = 0.742$). The Ob values also showed a positive correlation with absorption indices Ni ($r = 0.565$) and Mi ($r = 0.561$), while a significant negative correlation was observed between Ob and Ri ($r = 0.604$) and Oi ($r = -0.615$). The parameters Si, Pi, and Qi also showed a significant positive correlation but the magnitude values were out of range; hence, they were discarded. In a general sense, the absorption-based indices showed a significant correlation, but the indices were more likely to capture the trend of the real values rather than quantifying near the measured taxol values. In contrast, the reflectance-based indices captured trends along with the magnitude of the real values. The reflectance-based indices (TC 2 and TC 5) showed the highest correlation to the measured taxol content (Ob). The wavelength absorption values that were found most closely to the Ob values were Bx and Dx, which are centered at 415 and 670 nm, respectively. The positive correlating factors to the Ob values were majorly allocated in the center of Figure 4, which implies that the difference between the modelled and observed values was found the least in terms of magnitude. The center region of the graph majorly consists of indices values, either reflectance-based or absorption-based. This highlights the fact that to exploit hyperspectral data, more techniques need to be explored to process the data. The correlation heat map as per [4] may indicate the metabolite signature on the spectra, but to make those data useful information, more techniques need to be implemented.

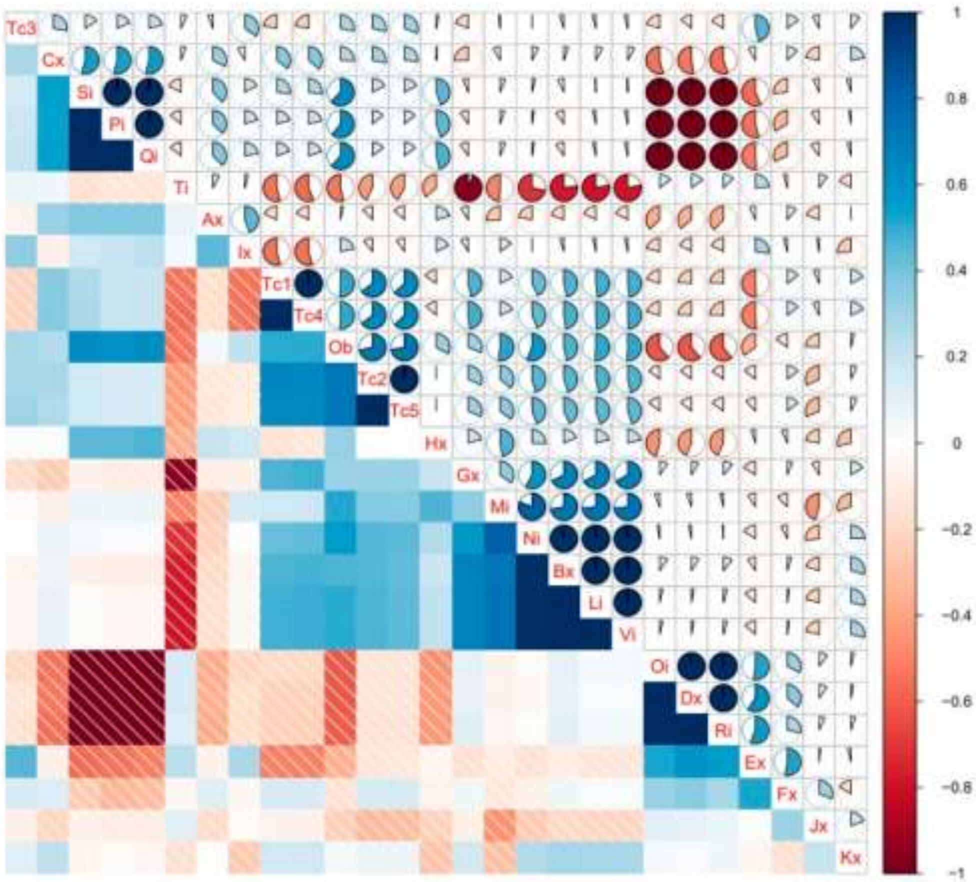


Figure 4. Pearson's correlation between the selected wavelength represented as 'x' as a subscript, absorption indices represented with 'i' as a subscript, and measured TC represented as 'Ob'.

The absorption indices that show a significant correlation and a magnitude value within the range, as compared to the measured TC, are listed in Table 2. Since the reflectance-based indices outperformed any other indices and wavelengths, they were further considered for the next set of statistical operations along with other important variables.

Table 2. Selected Taxol absorption-based indices where R is reflectance band.

Sl No.	Absorption Based Taxol Indices
1	Ni = (R415 − R670)/(R415 + R670)
2	Mi = (R415 − R2272)/(R415 + 2272)
3	Ri = R670/R1181
4	Oi = R670/R975

3.2. Descriptive Statistics

The land surface temperature (LST) showed a slightly higher temperature at the low altitude of around 15.8 (°C) at an elevation of 2992 m. With the varying altitude, the temperature also changed. The samplings were conducted in the rainy season with the average LST recorded as 14.171 ± 1.002 (°C). By following the soil temperature, the soil moisture values show a mean value of 43.950 ± 5.500 (%). The final values show that the TC varied between 0 and 0.037 mg/g FW, with an average of 0.011 mg/g FW ± 0.012. The values of total phenolic content (TPC) ranged from 72.656 to 94.676 mg GAE/g FW, with an average value of 79.973 ± 6.418. The correlation concerning elevation for TPC was found to be 0.672, which is significant. This shows that the TPC shows a clear positive change with elevation (as in Table 3). It clearly shows that medicinal plants also carry phenolic content in them, which indirectly shows the redox properties that are responsible for their antioxidant properties. The *p*-value was 0.006, which is also less than 0.05. This clearly shows that there is not much of a significant difference among the TPC content values, but it does increase with the elevation. The total polyphenol content (TPC) has a positive correlation with elevation, while it shows a negative correlation with the temperature. TPC is not limited by ecosystem boundaries but limited by human interventions. TPC values do increase with elevation. The high TPC concentration reported in *T. wallichiana* suggests that the medicinal plant contains high antioxidant activity, which makes it more beneficial.

Table 3. Pearson's correlation matrix among edaphic parameters, topographic parameters, biochemical- and indices-generated parameters of *Taxus wallichiana* needles.

Variables	Elevation	SM	Taxol Content	TCC	Carotenoids	TPC	LST	TC 1	TC 2	TC 3	TC 4	TC 5
Elevation	1.000											
SM	0.333	1.000										
Taxol content	0.277	−0.478	1.000									
TCC	−0.238	−0.340	0.186	1.000								
Carotenoids	0.435	0.065	0.333	−0.162	1.000							
TPC	0.658	0.516	−0.070	−0.438	0.001	1.000						
LST	−0.445	−0.654	0.478	0.312	−0.080	−0.533	1.000					
TC 1	0.789	0.192	0.495	0.087	0.456	0.380	−0.158	1.000				
TC 2	0.372	−0.186	0.715	0.310	0.508	−0.116	0.401	0.656	1.000			
TC 3	−0.393	−0.296	0.341	0.104	0.168	−0.543	0.536	−0.218	0.304	1.000		
TC 4	0.782	0.188	0.497	0.090	0.454	0.379	−0.153	1.000	0.658	−0.219	1.000	
TC 5	0.357	−0.191	0.715	0.313	0.504	−0.122	0.410	0.646	1.000	0.311	0.648	1.000

The total chlorophyll content (TCC) values vary from 2.013 to 4.194 mg/g. The average total chlorophyll concentration was found to be 3.541 ± 0.504. The correlation of total chlorophyll content with elevation came out to be insignificant. Similarly, the correlation between total chlorophyll and taxol was found to be insignificant. The values for carotenoids vary between 0.703 to 0.982 mg/g. The average carotenoid value came out to be 0.836 ± 0.087. This correlation between elevation and carotenoid was found to be insignificant (Table 3). The correlation between total chlorophyll and carotenoids was statistically insignificant. This inverse relationship between chlorophyll and carotenoid directs toward the inference that carotenoids increase in the senescence stage and chlorophyll is reported to be higher than carotenoids in the growing period in the case of many plants [38].

TPC shows a significant positive correlation with Soil Moisture (SM) of 0.516 and a significant negative correlation with LST. Soil Moisture shows an inverse relationship with LST, which is obvious as the temperature is the driving factor for water movement [39]. Here, the TPC relationship clearly indicates that TPC is affected by temperature [40].

The correlation matrix between the parameters of indices from TC 1 to TC 5 with elevation verifies the fact that the reflectance/albedo varies at different altitudes. This indicates that the indices developed for biochemical factors have the effect of altitude within them. Similarly, Mokarram et al. [41] have indicated that the vegetation growth is highest between the elevations of 1500 to 3000 m, with high values of the Normalized Difference Vegetation Index (NDVI), the Enhanced Vegetation Index (EVI), and the Difference Vegetation Index (DVI). It can be seen that TC 1 shows a clear relationship with TC 2 and TC 5 while TC 2 correlates with TC 4, but all these indices are probable indices for the same parameter, i.e., taxol. Hence, their mutual correlation is expected. Although, it is noticeable that TC 2 shows a perfect correlation with TC 5, and both TC 2 and TC 5 show the highest correlation with Taxol, which indicates that both the indices are useful in calculating Taxol content using hyperspectral reflectance data. TC 2 and TC 5 also show correlation with carotenoids, but the correlation cannot be considered significant. In the correlation matrix in Table 3, it can be seen that neither TC 2 nor TC 5 showed any significant correlation with any other foliar characteristic than taxol nor the taxol indices. This showed that the taxol indices developed using the wavelength of the visible region (415 and 421 nm) can be uniquely characterized for the same. Taxol did not show any significant correlation with elevation directly. Similar results were also obtained for taxol by Priyanka et al. After the application of ANOVA, the null hypothesis was accepted for soil moisture, total chlorophyll, near-surface temperature, taxol content, and carotenoids. The null hypothesis was rejected for TPC based on the Fcrit and p-value. The F values of 13.875 exceed the critical value of 3.88, which signifies that there is a difference among groups. Pairwise Scheffe's showed that there is no statistical significance among the various classes for TPC based on elevation.

3.3. Multivariate Analysis

Bartlett's sphericity test shows a calculated $\chi^2 = 243.932$, which is greater than the critical value $\chi^2 = 22.362$ ($p = 0.05$), thus the null hypothesis of equal variance among groups was rejected, indicating that PCA can accomplish a significant reduction in the dimensionality of the original data set [42]. To determine the principal component that explains the major attributes of *T. wallichiana*, PCA was applied, as shown in Table 4 and Figure 5.

PCA was performed on the combined (edaphic, topographic, and biochemical properties, and indices values) correlation matrix dataset in order to identify a condensed set of features that could capture and explain most of the variance in the data for *T. wallichiana*. The Scree plot and Table 4 highlight the factor loadings, eigenvalues, and variance described by each PC. According to the criteria set by [43], an eigenvalue greater than one was considered as a principal component. The factor loading of more than 0.650 was considered a contributing factor since the sample size was less than 100. PCA rendered three principal components with eigenvalues > 1, explaining almost 80.00% of the total variance of the data. The parameter PC 1, describing 39.17%, has strong positive factor loadings (>0.80) on TC 1, TC 2, TC 4, and TC 5. PC 1 also shows moderate loadings for the measured taxol content (>0.70) and elevation (>0.65), thus highlighting that taxol and elevation are the most varying variables among the measured parameters. In the case of *T. wallichiana*, the TC varies with age and seasons; hence, this variance is expected [44,45]. PC 2 explained 31.11% of the total variance and has moderate negative loading on LST. PC 2 also shows strong positive loading with TPC and SM. The loadings and scores of the first two PCs (PC1 and PC2) are plotted in Figure 5. The loadings plot (Figure 5) shows the distribution of all the parameters in the first (upper right) and fourth (lower right) quadrants. The factor loading lines joining the variables along with the length of the line passing through the origin in the plot of the factor loadings are indicative of the low and high contribution of the variables to the samples. The closeness of the lines of two variables signifies the strength of their mutual correlation, which was also adequately shown using a correlation matrix. The assemblage of TC, Carotenoids, elevation, TPC, and SM in the loadings plot suggests their significant mutual positive correlation.

Table 4. Loadings of experimental variables on the PCs for the combined data set of *Taxus wallichiana* (red indicates the most significant component).

Variables	PC 1	PC 2	PC 3
Elevation	0.684	0.630	−0.021
Soil Moisture (SM)	−0.036	0.743	0.179
Measured Taxol content (Taxol)	0.727	−0.417	0.009
Total Chlorophyll content (TCC)	0.147	−0.522	−0.697
Carotenoids	0.652	0.035	0.532
Total Polyphenolic Content (TPC)	0.179	0.828	0.036
Land Surface Temperature (LST)	0.122	−0.850	0.001
TC1	0.891	0.321	−0.205
TC2	0.902	−0.337	0.053
TC3	0.066	−0.700	0.512
TC4	0.891	0.317	−0.208
TC5	0.896	−0.347	0.055
Eigen Value	4.700	3.733	1.155
% Variance	39.17%	31.11%	9.63%
Cumulative % Variance	39.17%	70.28%	79.91%

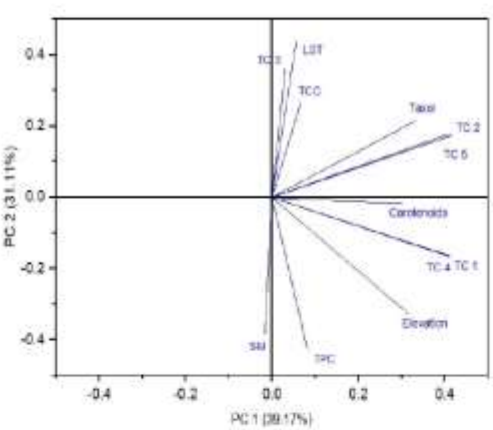


Figure 5. Loading plot of Principal Component Analysis for Taxus wallichiana.

3.4. Seriation Analysis

Generally, as the amount of medicinal compound is dependent upon numerous factors, it primarily involves elevation, temperature, and the ecosystem. Here, seriation was conducted to bring out the arrangement of various factors related to TC. The relationship becomes highly subjective but has some principal components that are commonly associated with biochemical parameters. The results of this study suggest that a landscape variable such as altitude is important in influencing other biochemical parameters as well as the secondary metabolites associated with the selected species. In order to analyze the data more clearly, the seriation plot and dendrogram were generated and are shown in Figure 6. The samples at every 50 m were pooled together as one sample. The objective functions during the iteration for the row and the column were obtained as 0.707 and 0.644, respectively, while the sum of all the pairwise distances in the neighboring rows (path length) was found to be 43.484 and the neighboring column (path length) was found to be 31.869. The Complete linkage rule was utilized for both the row and column, while the multi fragment heuristic (MF) scaling rule was used for tree seriation. The dissimilarity analysis used in seriation was based on the Euclidean distance measurement.

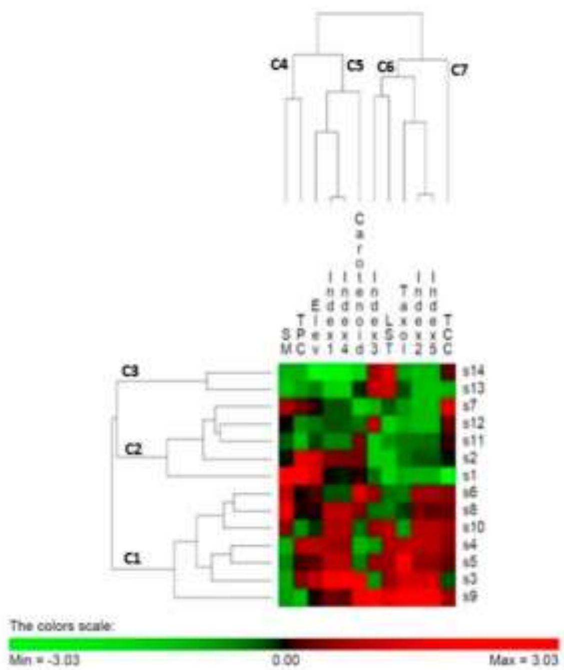


Figure 6. Seriation Analysis of the samples.

The seriation column matrix plot Cluster C4 contains TPC and SM, while cluster C5 contains carotenoids and elevation along with TC 1 (Index 1) and TC 4 (Index 4). This implies that the elevation, TPC, SM, and Carotenoids values show a relation with TC 1 and TC 4. This implies that elevation, due to the change in albedo, has a direct relationship with biochemical and edaphic properties. The C6 cluster contains TC3 (Index 3), LST, Taxol, TC 2 (Index 2), and TC 5 (Index 5). Carotenoids show a close relationship with both indices-generated values TC 4 and TC 3. Similarly, the Taxol content was found to be feeble with LST, but the correlation coefficient was insignificant to consider. The taxol content has a close correlation with TC 2 and TC 5 with good correlation values and does not relate to any other variable, even in the hierarchical sense. Hence, it can be said that TC 2 and TC 5 can only retrieve taxol, not any other common foliar pigment found in the visible region of electromagnetic spectra. The Total Chlorophyll Content (TCC) behaved as a runt.

It consistently showed the same range values without being majorly affected by any other variable. The TCC is supposed to vary with the season, species, age of the plant, and forest type [46].

The samples S6, S8, S10, S4, S5, S3, and S9 are grouped in one cluster, C1, while cluster C2 includes samples S7, S12, S11, and S2. Samples S14 and S15 clustered together in C3. In Figure 6, several samples from cluster C1 were characterized with similar behavior. These samples were measured in-between altitudes of 2826 and 3003 m, which is also the dense region of the forest in the sampling area. Therefore, these samples must have a strong influence on the forest and its ecosystem. The samples clustered under C2 and C3 shared the same hierarchy, due to their presence either at the high or low altitude of the sampling elevation, which was marked by human intervention at the lower altitude or ecosystem change at the higher altitude.

4. Conclusions

The reflectance-based indices are more useful in quantifying taxol content using hyperspectral data. More techniques such as indices and algorithms need to be applied for the exploitation of hyperspectral data so that these data can be converted into useful information, as absorption bands at particular wavelengths are not providing any sufficient information to make the HSR data more useful. In the case of multispectral satellite data, absorption-based indices may be used to quantify the taxol content.

The result of statistical analysis suggests that the density of the forest determines the range of parameters measured, which, in the case of the montane ecosystem, is indirectly determined by elevation. Therefore, elevation along with aspect and slope in many respects determines the microclimate, and thus, plays an important role in foliar and edaphic properties in the case of the montane ecosystem. Chlorophyll does not show any significant change in a species under the same forest canopy therefore, it might be used as a health indicator at the canopy scale but cannot be used as an indicator to decide the number of secondary metabolites in the same species.

The relationship in the case of taxol with elevation suggests that the taxol content does not vary with elevation but is affected by temperature. It is the most varying variable among the measured variables, followed by elevation and carotenoids. The frequency of the plant becoming less near the edge of the ecosystem (ecotone) and the amount of taxol content in *T. wallichiana* near these regions was also low. Beyond 3100 m, more of a grassland ecosystem exists in the NDBR. The samples showing similar behavior in terms of parameters were found between elevations of 2800 to 3000 m in the NDBR. This region is characterized by dense forests in the NDBR. The *T. wallichiana* plant shows low taxol content near the timberline at Phurkia and near the point of human intervention at Khati (the last habitable point in the valley). This makes our understanding of this highly medicinal plant more refined. *T. wallichiana* with a high taxol content is found more in its natural habitat in the absence of human intervention and ecosystem change. Taxol indices TC 2 and TC 5, which were developed using visible range wavelengths (415 and 421 nm) of the hyperspectral data, have been related to the taxol content and not related with other foliar variables, which might be attempted in the future. This study can be expanded to other regions for taxol estimation, but the availability of ground hyperspectral data is a challenge. The canopy chemistry and its relationship with remote sensing hyperspectral data is a challenge, as there are thousands of compounds in the same species. The implementation of more sophisticated techniques applied with HSR holds the key to future research in canopy chemistry.

Author Contributions: Conceptualization, P.K.S.; Formal analysis, A.G., P.K.S. and P.S.; Funding acquisition, P.K.S. and G.P.P.; Investigation, A.G., P.K.S. and P.S.; Methodology, A.G. and P.K.S.; Project administration, P.K.S.; Resources, P.K.S. and G.P.P.; Software, A.G. and P.S.; Supervision, P.K.S.; Validation, A.G.; Visualization, P.K.S.; Writing—original draft, A.G. and P.K.S.; Writing—review and editing, P.K.S. and G.P.P. All authors have read and agreed to the published version of the manuscript.

Funding: A.G. is funded under the University Grant Commission's Junior Research Fellowship program. This work is funded by the *National Mission on Himalayan Studies*, G.B. Pant National Institute of Himalayan Environment (NIHE), Ministry of Environment, Forest & Climate Change (MoEF & CC), Government of India.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the datasets are generated during the study.

Acknowledgments: The authors are thankful to the University Grant Commission for the necessary financial assistance. The authors also acknowledge the Central Institute of Medicinal and Aromatic Plants, Lucknow, India for providing the necessary laboratory support for the study. The authors also extend their sincere thanks to National Mission on Himalayan Studies, G.B. Pant National Institute of Himalayan Environment (NIHE) for the necessary funding.

Conflicts of Interest: There is no conflict of interest.

References

1. Kokaly, R.F.; Skidmore, A.K. Plant phenolics and absorption features in vegetation reflectance spectra near 1.66 μm. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *43*, 55–83. [CrossRef]
2. Lu, B.; He, Y.; Liu, H.H. Mapping vegetation biophysical and biochemical properties using unmanned aerial vehicles-acquired imagery. *Int. J. Remote Sens.* **2018**, *39*, 5265–5287. [CrossRef]
3. Peng, Y.; Zhang, M.; Xu, Z.; Yang, T.; Su, Y.; Zhou, T.; Wang, H.; Wang, Y.; Lin, Y. Estimation of leaf nutrition status in degraded vegetation based on field survey and hyperspectral data. *Sci. Rep.* **2020**, *10*, 1–12. [CrossRef] [PubMed]
4. Fine, P.V.; Salazar, D.; Martin, R.E.; Metz, M.R.; Misiewicz, T.M.; Asner, G.P. Exploring the links between secondary metabolites and leaf spectral reflectance in a diverse genus of Amazonian trees. *Ecosphere* **2021**, *12*, e03362. [CrossRef]
5. Hui, C. Carrying capacity, population equilibrium, and environment's maximal load. *Ecol. Model.* **2006**, *192*, 317–320. [CrossRef]
6. Asner, G.P.; Martin, R.E. Conservation. Spectranomics: Emerging science and conservation opportunities at the interface of biodiversity and remote sensing. *Glob. Ecol.* **2016**, *8*, 212–219. [CrossRef]
7. Jugran, A.K.; Bahukhandi, A.; Dhyani, P.; Bhatt, I.D.; Rawal, R.S.; Nandi, S.K. Impact of altitudes and habitats on valerenic acid, total phenolics, flavonoids, tannins, and antioxidant activity of Valeriana jatamansi. *Appl. Biochem.* **2016**, *179*, 911–926. [CrossRef]
8. Díaz, S.; Kattge, J.; Cornelissen, J.H.; Wright, I.J.; Lavorel, S.; Dray, S.; Reu, B.; Kleyer, M.; Wirth, C.; Prentice, I.C. The global spectrum of plant form and function. *Nature* **2016**, *529*, 167–171. [CrossRef] [PubMed]
9. Nisar, M.; Khan, I.; Simjee, S.U.; Gilani, A.H.; Perveen, H. Anticonvulsant, analgesic and antipyretic activities of Taxus wallichiana Zucc. *J. Ethnopharmacol.* **2008**, *116*, 490–494. [CrossRef] [PubMed]
10. Shi, Q.W.; Kiyota, H. New natural taxane diterpenoids from Taxus species since 1999. *Chem. Biodivers.* **2005**, *2*, 1597–1623. [CrossRef] [PubMed]
11. Kumari, N.; Srivastava, A.; Dumka, U.C. A long-term spatiotemporal analysis of vegetation greenness over the Himalayan Region using Google Earth Engine. *Climate* **2021**, *9*, 109. [CrossRef]
12. Lin, D.; Xiao, M.; Zhao, J.; Li, Z.; Xing, B.; Li, X.; Kong, M.; Li, L.; Zhang, Q.; Liu, Y. An overview of plant phenolic compounds and their importance in human nutrition and management of type 2 diabetes. *Molecules* **2016**, *21*, 1374. [CrossRef]
13. Hill, J.; Buddenbaum, H.; Townsend, P.A. Imaging spectroscopy of forest ecosystems: Perspectives for the use of space-borne hyperspectral earth observation systems. *Surv. Geophys.* **2019**, *40*, 553–588. [CrossRef]
14. Shang, X.; Chisholm, L.A. Classification of Australian native forest species using hyperspectral remote sensing and machine-learning classification algorithms. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *7*, 2481–2489. [CrossRef]
15. Anand, A.; Pandey, M.K.; Srivastava, P.K.; Gupta, A.; Khan, M.L. Integrating Multi-Sensors Data for Species Distribution Mapping Using Deep Learning and Envelope Models. *Remote Sens.* **2021**, *13*, 3284. [CrossRef]
16. Zarco-Tejada, P.J.; Morales, A.; Testi, L.; Villalobos, F. Spatio-temporal patterns of chlorophyll fluorescence and physiological and structural indices acquired from hyperspectral imagery as compared with carbon fluxes measured with eddy covariance. *Remote Sens. Environ.* **2013**, *133*, 102–115. [CrossRef]
17. Li, F.; Miao, Y.; Feng, G.; Yuan, F.; Yue, S.; Gao, X.; Liu, Y.; Liu, B.; Ustin, S.L.; Chen, X. Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crop. Res.* **2014**, *157*, 111–123. [CrossRef]
18. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [CrossRef]
19. Singh, P.; Srivastava, P.K.; Malhi, R.K.M.; Chaudhary, S.K.; Verrelst, J.; Bhattacharya, B.K.; Raghubanshi, A.S. Denoising AVIRIS-NG data for generation of new chlorophyll indices. *IEEE Sens. J.* **2020**, *21*, 6982–6989. [CrossRef]
20. Wang, Z.; Wang, T.; Darvishzadeh, R.; Skidmore, A.K.; Jones, S.; Suarez, L.; Woodgate, W.; Heiden, U.; Heurich, M.; Hearne, J. Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sens. Environ.* **2016**, *8*, 491. [CrossRef]

21. Ollinger, S.V.; Richardson, A.D.; Martin, M.E.; Hollinger, D.Y.; Frolking, S.E.; Reich, P.B.; Plourde, L.C.; Katul, G.G.; Munger, J.W.; Oren, R. Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 19336–19341. [CrossRef] [PubMed]
22. Fernandes, M.R.; Aguiar, F.C.; Silva, J.M.; Ferreira, M.T.; Pereira, J.M. Spectral discrimination of giant reed (Arundo donax L.): A seasonal study in riparian areas. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 80–90. [CrossRef]
23. Hennessy, A.; Clarke, K.; Lewis, M.J.R.S. Hyperspectral classification of plants: A review of waveband selection generalisability. *Remote Sens.* **2020**, *12*, 113. [CrossRef]
24. Pandey, S.K.; Singh, A.K.; Hasnain, S. Grain-size distribution, morphoscopy and elemental chemistry of suspended sediments of Pindari Glacier, Kumaon Himalaya, India. *Hydrol. Sci. J.* **2002**, *47*, 213–226. [CrossRef]
25. Saxena, K.; Maikhuri, R.; Rao, K.; Nautiyal, S. *Assessment Report: Nanda Devi Biosphere Reserve, Uttarakhand, India as a Baseline for Further Studies Related to the Implementation of Global Change in Mountain Regions (GLOCHAMORE) Research Strategy*; Assessment Report; UNESCO, New Delhi Office: New Delhi, India, 2010.
26. Joshi, S.; Upreti, D.K. Lichenometric studies in vicinity of Pindari Glacier in the Bageshwar district of Uttarakhand, India. *Curr. Sci.* **2010**, *99*, 231–235.
27. Singh, R.; Kumar, S.; Kumar, A. Climate change in Pindari region, Central Himalaya, India. In *Climate Change, Glacier Response, and Vegetation Dynamics in the Himalaya*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 117–135.
28. Joshi, S.; Upreti, D.; Das, P. Lichen diversity assessment in Pindari glacier valley of Uttarakhand, India. *Geophytology* **2011**, *41*, 25–41.
29. Mac Arthur, A.; MacLellan, C.J.; Malthus, T. The fields of view and directional response functions of two field spectroradiometers. *IEEE Trans. Geosci. Remote Sens. Environ.* **2012**, *50*, 3892–3907. [CrossRef]
30. Srivastava, P.K.; Malhi, R.K.M.; Pandey, P.C.; Anand, A.; Singh, P.; Pandey, M.K.; Gupta, A. Revisiting hyperspectral remote sensing: Origin, processing, applications and way forward. In *Hyperspectral Remote Sensing*; Elsevier: Amsterdam, The Netherlands, 2020; pp. 3–21.
31. Gupta, A.; Singh, P.; Srivastava, P.K.; Pandey, M.K.; Anand, A.; Chandra Sekar, K.; Shanker, K. Development of hyperspectral indices for anti-cancerous Taxol content estimation in the Himalayan region. *Geocarto Int.* **2021**, 1–14. [CrossRef]
32. Kokaly, R.F. Investigating a physical basis for spectroscopic estimates of leaf nitrogen concentration. *Remote Sens. Environ.* **2001**, *75*, 153–161. [CrossRef]
33. Schmidt, K.; Skidmore, A. Exploring spectral discrimination of grass species in African rangelands. *Int. J. Remote Sens.* **2001**, *22*, 3421–3434. [CrossRef]
34. Schmidt, K.; Skidmore, A. Spectral discrimination of vegetation types in a coastal wetland. *Remote Sens. Environ.* **2003**, *85*, 92–108. [CrossRef]
35. Kamboj, A.; Gupta, R.; Rana, A.; Kaur, R. Application and analysis of the Folin Ciocalteu method for the determination of the total phenolic content from extracts of Terminalia bellerica. *Eur. J. Biomed. Pharm. Sci.* **2015**, *2*, 201–215.
36. Gupta, A.; Lamba, P.; Gupta, D.; Verma, D. Medicinal Evaluation of different flowers from Asteraceae Family. *Bull. Environ. Sci. Res.* **2019**, *8*, 10–14.
37. Shanker, K.; Negi, A.S.; Chattopadhyay, S.K.; Sashidhara, K.; Kaur, T.; Gupta, M.; Agrawal, P.; Misra, A. Determination of paclitaxel, 10-DAB, and related taxoids in Himalayan Yew using reverse phase HPLC. *J. Herbs Spices Med. Plants* **2008**, *13*, 25–44. [CrossRef]
38. Croft, H.; Chen, J.; Wang, R.; Mo, G.; Luo, S.; Luo, X.; He, L.; Gonsamo, A.; Arabian, J.; Zhang, Y. The global distribution of leaf chlorophyll content. *Remote Sens. Environ.* **2020**, *236*, 111479. [CrossRef]
39. Cuo, L.; Zhang, Y.; Bohn, T.J.; Zhao, L.; Li, J.; Liu, Q.; Zhou, B. Frozen soil degradation and its effects on surface hydrology in the northern Tibetan Plateau. *J. Geophys. Res. Atmos.* **2015**, *120*, 8276–8298. [CrossRef]
40. Ghaderpour, E.; Ben Abbes, A.; Rhif, M.; Pagiatakis, S.D.; Farah, I.R. Non-stationary and unequally spaced NDVI time series analyses by the LSWAVE software. *Int. J. Remote Sens.* **2020**, *41*, 2374–2390. [CrossRef]
41. Mokarram, M.; Sathyamoorthy, D. Modeling the relationship between elevation, aspect and spatial distribution of vegetation in the Darab Mountain, Iran using remote sensing data. *Modeling Earth Syst. Environ.* **2015**, *1*, 1–6. [CrossRef]
42. Vorapongsathorn, T.; Taejaroenkul, S.; Viwatwongkasem, C. A comparison of type I error and power of Bartlett's test, Levene's test and Cochran's test under violation of assumptions. *Songklanakarin J. Sci. Technol.* **2004**, *26*, 537–547.
43. Singh, K.P.; Malik, A.; Sinha, S.; Singh, V.K.; Murthy, R.C. Estimation of source of heavy metal contamination in sediments of Gomti River (India) using principal component analysis. *Water Air Soil Pollut.* **2005**, *166*, 321–341. [CrossRef]
44. Nadeem, M.; Rikhari, H.; Kumar, A.; Palni, L.; Nandi, S. Taxol content in the bark of Himalayan Yew in relation to tree age and sex. *Phytochemistry* **2002**, *60*, 627–631. [CrossRef]
45. Yang, L.; Zheng, Z.-S.; Cheng, F.; Ruan, X.; Jiang, D.-A.; Pan, C.-D.; Wang, Q. Seasonal dynamics of metabolites in needles of Taxus wallichiana var. mairei. *Molecules* **2016**, *21*, 1403. [CrossRef] [PubMed]
46. Li, Y.; He, N.; Hou, J.; Xu, L.; Liu, C.; Zhang, J.; Wang, Q.; Zhang, X.; Wu, X. Factors influencing leaf chlorophyll content in natural forests at the biome scale. *Front. Ecol. Evol.* **2018**, *6*, 64. [CrossRef]

Development of hyperspectral indices for anti-cancerous Taxol content estimation in the Himalayan region

Ayushi Gupta[1], Prachi Singh[1], Prashant K. Srivastava[1,2*], Manish K. Pandey[1], Akash Anand[1], K. S. Chandra Sekar[3], and Karuna Shanker[4]

[1]Remote Sensing Laboratory, Institute of Environment and Sustainable Development, Banaras Hindu University, U.P

[2]DST-Mahamana Centre for Excellence in Climate Change Research, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, India

[3]G.B. Pant, National Institute of Himalayan Environment (NIHE), Kosi-Katarmal, Almora, Uttarakhand, India

[4]CSIR - Central Institute of Medicinal and Aromatic Plants, Lucknow, India

[*]Corresponding Author: prashant.iesd@bhu.ac.in

Abstract

Monitoring and management of rare and economically important species in the highly complex terrain are challenging and thus need advanced technological development. In this study, the hyperspectral radiometer data of *Taxus wallichiana* were acquired at highly complex terrain of the Pindari region of the Himalaya and processed by using several sophisticated algorithms to deduce Taxol content in the plants. The spectroradiometer data were denoised through three different types of smoothing filters such as Average Mean, Savitzky Golay, and Fast Fourier Transform (FFT) followed by feature selection for allocation of best bands for Taxol content estimation. The results showed that the Average Mean filter in combination with feature selection performed best for Taxol spectral indices generation, model development, and Taxol content prediction. The best model showed a correlation of 0.719 with a relative root mean square error (RMSEr) value of 0.678 for Taxol content prediction.

Keywords Taxol spectral indices; Taxol model development; hyperspectral data; smoothening and filtering; *Taxus wallichiana*

1. Introduction

According to WHO (World Health Organization) estimates, 80% of the population worldwide count on herbal medicines for some aspect or the other for their primary health care needs. Approximately 2/3rd of the plants accounted in the modern medical system found their health care origin in Asian countries. Apart from the rural population depending on indigenous systems of medicine (Ekor 2014) even modern medicine derives its inspiration from the indigenous medicinal system (Yuan *et al.* 2016). Many researchers have also emphasized that modern medicine should take the precious experience of natural products and traditional medicine (Pan *et al.* 2014, Yuan *et al.* 2016). In India, around 30% household uses traditional medicine (Srinivasan *et al.* 2017), which is plentiful in India.

Medicinal plants contain phytopigments and bioactive compounds which contribute to their physiological function and medicinal properties (Mohamed *et al.* 2010). The most important phytochemical components that are responsible for medicinal properties for any plant are alkaloids, tannins, flavonoids, and phenolic compounds (Geetha and Geetha 2014). One of the major plants i.e., *Taxus wallichiana* Zucc. (*T. wallichiana*) also known as Himalayan yew is found very promising for cancer treatment. The needles/leaves of the *T.wallichiana* is one of the valuable sources of taxoid (Appendino *et al.* 1992, Bala *et al.* 1999) Paclitaxel (trade name Taxol) is a tricyclic diterpenoid (alkaloid) and considered as an efficient anti-cancerous drug (Zhu *et al.* 2019). The extraction and refinement of Taxol are time-consuming, difficult, expensive, and tedious because of the low yields (van Rozendaal *et al.* 2000). The taxanes are isolated from the Taxus plant material by complex extraction procedures and analyzed using sophisticated HPLC–UV or LC-MS methods (Fu *et al.* 2009, Chakchak and Zineddine 2013, Sadeghi-Aliabadi *et al.* 2015). However, the population of these species has seen a large reduction due to its excessive demand and collection of this anti-tumor and anti-cancerous drug. A study conducted exhibited that these trees were spoiled due to bark-stripping practices. Moreover, these species are very slow-growing (Suffness 1995). Hence, management of this important resource at a larger scale becomes necessary which can be achieved using remote sensing.

The identification and differentiation among various medicinally important species using remote sensing are often limited, by the ability of spectral variance, which can discriminate the minute spectral differences among species (Clark *et al.* 2005). Hyperspectral remote sensing can serve as a suitable solution since it contains several (mainly between 64 and 256) contiguous fine resolution bands with a bandwidth of 1 to 10 nm, providing noteworthy levels of subtle feature differences to provide fine spectral variations among tree species (Peerbhay *et al.* 2013, Srivastava *et al.* 2020b). These fine spectral bands data are characteristically high-dimensional but also found to be highly correlated with the vegetation parameters tested (Landgrebe 2002). Recently plant phenolics were also characterized in vegetation reflectance at 1660 nm (Kokaly *et al.* 2015). Clearly, retrieval of such canopy information via remotely sensed data involves analytical means which are proficient in translating the spectral response data into practical information.

The high dimensionality of these datasets also cause many problems such as heavy computational processor requirements and high data storage cost. To process high-dimensional data efficiently, dimensionality reduction (DR) becomes essential. Band Selection (BS) is one of the techniques of DR that selects a subset of bands preserving their physical meaningful data with the benefit of keeping intact the relevant original information in the data (Srivastava *et al.* 2020a). The BS method derivative analysis is an amalgamation of a variance-based and shape-based (derivative) approach for feature identification (Tsai and Philpot 1998b) having better separability (larger Jefferies-Matsushita (JM) distances) to differentiate among groups. The apprehension of data dimensions and training data sizes in hyperspectral emphasizes the need to compress valuable information into the least number of bands (Ling *et al.* 2019, Singh *et al.* 2020a).

Uncertainties are often seen in spectral datasets caused due to atmospheric disturbances. Hence, for optimal band selection, smoothening (data pre-processing) becomes a part of the derivative analysis application (Torrecilla *et al.* 2009). Many studies have modified and reviewed these filtering and smoothening techniques to develop a set of cross-platform tools for the analysis of spectral data (Tsai *et al.* 2002). Comparatively few scholars have taken the derivative approach for the analysis of hyperspectral data used in remote sensing due to its limitation (Torrecilla *et al.* 2009). The regression-based models on spectral indices are characteristically empirical formulae aiding the plotting of several biochemical parameters consequential from remotely sensed data. Since it is empirical in nature, it remains undefined to up to what extent this selected regression model works well, till all the band combinations and curve-fitting functions are evaluated (Rivera *et al.* 2014b, Pandey *et al.* 2019).

According to the study of J. P. Rivera et. al (Rivera *et al.* 2014a) many hyperspectral indices have been tested for the retrieval of LAI and Chlorophyll using HyMap sensor data during the

SPARC-2003 campaign in Barrax, Spain. 12 chlorophyll spectral indices for chlorophyll inversion have been developed using ASD spectroradiometer data such as Vogelmann red edge index, Zarco-Tejada, Miller index (ZMI), modified normalized difference vegetation index (mNDVI), modified normalized difference index (mND) etc (Lin *et al.* 2012), (Zagajewski *et al.* 2018). The vast majority of these SIs and their association with desired parameters have been established through experimental work. According to the above studies development of indices can be a successful approach for the retrieval of the biochemical parameters using hyperspectral data. These studies are based on the parametric regression approach (Gupta *et al.* 2014, Verrelst *et al.* 2019). Hence expanding the knowledge of the Hyperspectral for sophisticated biochemical parameters estimation becomes the next logical step in this direction. In the purview of the above, the main aims of this study are 1) Estimation of alkaloid Taxol and reporting its concentration in the Pindari region of Himalaya. 2) Development of indices sensitive for Taxol content estimation through denoised hyperspectral data. 3) Development of robust assessment method to evaluate various indices, spectral band settings, and curve-fitting functions for retrieval of Taxol content in the Himalayan region.

2. Material and Methods

2.1. Study Area

The Nanda Devi biosphere reserve is in Chamoli, Pithoragarh, and Bageshwar districts of the state of Uttarakhand is located in Western Himalayas lying in Highland Biogeographic Zone (2a). The climatic year of the Nanda Devi biosphere reserve has been distinguished into three seasons mainly- summer (April- June), rainy season (June-September), and winter (October-March). The average annual rainfall is 930 mm, out of which 48% occurs in two months (July-August). The maximum temperature range varies between 11 to 24°C and the minimum temperature varies between 3 to 7°C. The present study site selected is rich in medicinal herbs as well as trees. *T. wallichiana* is one of the species which is prominent and highly medicinal in nature. The medicinal compound is one of the major reasons for the plant's declining population. The sampling was done in the rainy season as it is the most favorable season for plant growth. Broadly the area is divided into two climatic zones that could be categorized as (i) Lower montane zone: elevation range of 1800-2400 m above mean sea level (amsl), (ii) Upper montane zone: elevation range of 2400- 3000 m amsl. The precipitation is more in the upper zone is more in terms of snowfall than showers (Gaur, 1999). The samples of *Taxus wallichiana* were collected are as shown in figure 1.

2.2. Sample collection and analysis

Ground sample was collected in the Pindari region of Himalaya during the dates 26/09/2019 and 29/09/2019 at different locations with a varying altitude of 2292-3039 m in the Nanda Devi Biosphere Reserve (NDBR) as shown in figure 1 along with hyperspectral radiometer measurements. The leaves of the plant sample were collected from two to three locations from the same tree to make the leaf sample homogenous for each located tree of *T. wallichiana*. The plant samples collected over the NDBR were of the same phenological stage. The spectra of these leaves were recorded using a handheld ASD spectroradiometer. The properties of the selected plants are displayed in table 1. The samples were then were crushed in liquid nitrogen for further analysis. For Taxol extraction, 1 g of crushed leaves was deflated with hexane using sonication. The deflated samples were filtered and then percolated using 25 ml methanol, each time repeated thrice using sonication. The hexane portions were rejected and methanol aliquots were collected together and then concentrated using a rotary evaporator. The samples were extracted in distilled water (50 ml). For the chloroform partitioning, the extracted water sample was then successively extracted by the solvent extraction process five times with 50 ml of chloroform each time. The chloroform extracted sample was then pooled together (250 ml) for each sample and dried under reduced pressure using a rotary evaporator, then re-dissolved in methanol (1 ml) (Shanker *et al.* 2008).

The liquid chromatography was done at room temperature on a Symmetry® C18, (both 250 mm × 4.6 mm i.d, 5.0 µm particle size) with an ultraviolet-diode array detector (UV-DAD). Chromatographic surroundings were augmented by regulating the composition and potential of hydrogen (pH) of the mobile phase for replicable results. The chromatographic solvents used for isocratic runs were: (a) Methanol and (b) Water (0.05% Acetic Acid) (62:38, v/v). The flow rate for the mobile phase was 1.0 mL min^{-1}. The working solution of paclitaxel was prepared from standard using methanol. Insertions of samples were done using a sample injector of a 20 µL loop. The UV-DAD scanned acquisitions of Taxol at 230 nm. The percentage of Taxol was calculated using equation (1) (Shanker *et al.* 2008).

$$\text{Taxol Content (\%)} = \frac{Ar_{sample} * Conc_{std}\left(^{mg}/_{ml}\right)}{Ar_{std} * 1000 * Conc_{sample}\left(^{g}/_{ml}\right)} * 100 \quad (1)$$

where Ar_{std} and Ar_{sample} are the areas under peak associated with the standard or reference and sample taxoid, respectively, and $Conc_{sample}$ and $Conc_{std}$ are the concentration of sample and reference taxoid, respectively (Shanker *et al.* 2008).

2.3. Data pre-processing

Data pre-processing is a crucial step. It has been stated that a key issue of applying filters for pre-processing is to allow the smoothening techniques to match the scale of the spectral features of interest (Bruce *et al.* 2001).

2.3.1. Savitzky–Holay Smoothing

Savitzky and Golay uses simplified least square fit intricacy, smoothing and derivatives. The general equation of the simplified least square convolution can be represented as equation (2)

$$S^* = \frac{\sum_{i=-m}^{m} C_i S_{j+i}}{n} \quad (2)$$

where S is the original spectral information, S* is the resultant (smoothed) spectral information, C_i is the coefficient for the i^{th} spectral value of the filter (smoothing window), and n is the number of convoluting integers. The index j is the running index of the original ordinate data table. The smoothing array (filter size) consists of 2 m + 1 points, where m is the half-width of the smoothing window (Tsai and Philpot 1998a).

2.3.2. Mean filter Smoothening

A mean filter takes the mean spectral value of nearest points within the considered window and the new value of j is the midpoint of the chosen window as given in equation (3)

$$S_j = \frac{\sum s_i}{n} \quad (3)$$

where n is number of sampling points. If the user specifies an even number of points as the filter size, the mean is assigned to the new value of the nearest point right of the center (longer wavelength) (Tsai and Philpot 1998a).

2.3.3. Fast Fourier transform (FFT)

The Fast Fourier process of digital filtering has been used for many years to process chemical signals. The basic equation for digital filtering is the correlation equation (4):

$$c(\pm n\Delta x) = \sum_x a(x)b(x \pm n\Delta x)\ n = 0, 1, 2\,. \qquad (4)$$

where an (x) is the original signal, b(x) is the filter function, c(nΔx) is the filtered signal, and Δx is the sampling interval. Eq. (4) points to the filtered signal which is obtained by estimating the sum of the products of the signal and the filter function when the filter function is shifted across the whole signal waveform. In simple terms written as equation (5) (Singh *et al.* 2020b).

$$\begin{tikzcd} a(x) \arrow[d] & * & b(x) \arrow[d] & = & c(x) & & (5) \\ A(f) & * & B(f) & = & C(f) \arrow[u] & & (6) \end{tikzcd}$$

The digital filters were implemented using the Fourier transform route as illustrated by Equations (5) and (6). The output of the FFT subroutine consists of two series, X(J) and Y(J), which are the real and imaginary components of the transform. X(J) resembles to A(f) in Eq. (6) (Betty and Horlick 1976).

2.4. Feature Selection

Derivative Spectral Analysis (DSA) algorithms were implemented to facilitate the extraction of "useful" information from hyperspectral data. Absorption features in reflectance spectra are enhanced using derivative spectroscopy. A derivative of a set of consecutive values (a spectrum). Adding the derivatives as features in the identification process optimizes and minimizes the number of bands required to achieve acceptable results due to larger JM distances (Tsai *et al.* 2002).

In the process here, spectral derivatives were assessed using a finite approximation algorithm. For the first-order derivative of a spectrum, s(λ), the estimation is based on equation (7).

$$\frac{\partial s}{\partial \lambda} \approx \frac{s(\lambda_j) - s(\lambda_i)}{\Delta\lambda} \qquad (7)$$

where Δλ is the separation between adjacent bands, i.e., $\Delta\lambda = \lambda_j - \lambda_i$ and $\lambda_j > \lambda_i$

However, the procedures do not work at the ends of the spectrum therefore, the resultant spectrum is shorter than the original by the width of the filter. It is noteworthy to remember that no new information is created by using derivatives (Tsai *et al.* 2002). Qualitative information regarding pigment concentration has been obtained based on the wavelength position of absorption features in derivative spectra (Louchard *et al.* 2002)

2.5. Automated Radiative Transfer Models Operator

The Spectral Index (SI) assessment (Verrelst *et al.* 2011, Verrelst *et al.* 2013) through Automated Radiative Transfer Models Operator (ARTMO) has been implemented in this study. It is based on parametric and non-parametric regression along with physically-based inversion using a lookup table (LUT). Because of complicated processing steps, a combined approach of the Radiative

Transfer Model (RTM) and vegetation indices have been introduced in the ARTMO on the MATLAB platform. With the help of the Spectral Index (SI) assessment, new generic indices have been developed in this study. Afterward, a statistical regression model with various curve fitting was implemented, which allows the relation of satellite data between desired biochemical parameters by using (ex-situ) calibration data. First, for each SI, all spectral band combinations are correlated against the generated dataset. The results are generated with the dataset that has formerly been segregated into a calibration and validation set. The obtained SI models are evaluated with multiple linear goodness-of-fit measures like the r, and the relative root means squared error (RMSEr). By investigating all bands against each other in a correlation matrix, ARTMO helps to identify redundant bands and to overcome Hughes' phenomenon or "curse of dimensionality" (Rivera *et al.* 2014b). The workflow of the entire methodology is shown in **figure 2.**

3. Results and Discussion

3.1. Statistical analysis of biochemical properties

Here, chlorophyll content, carotenoid content, total polyphenolic content (TPC), and taxol content (TC) were statistically analysed. The outcomes of the biochemical analysis are shown in **figure 3**, which suggests that common biochemical properties such as total chlorophyll and carotenoids are present in *T. wallichiana* in a certain specified range. The chlorophyll content varied between 2.014 to 4.195 mg/g with an average of 3.541 +/- 0.501. Carotenoids varied between 0.704 to 0.983 mg/g with an average of 0.836 +/- 0.087 respectively. TPC and TC varied between 94.676 to 72.656 mg GAE/g of its fresh weight (FW) with an average of 79.901 ± 6.271 mg GAE/g of FW and 0 to 0.037 mg/g FW with an average of 0.011 +/- 0.012 mg/g FW. TPC showed a significant correlation coefficient (0.672) with altitude. The other biochemical properties did not show any statistically significant correlation with elevation. The frequency of the *T. wallichiana* becomes lesser after 3040 m as more of a grassland ecosystem naturally exists at Nanda Devi Biosphere Reserve. The highest Taxol content concerning elevation is recorded between 2850 to 3000 m in Nanda Devi Biosphere Reserve. The TPC values and their correlation with elevation are strong unlike taxol due to low temperatures at higher altitudes up to elevation 3100 m. TPC and TC relations show that medicinal plants also carry phenolic content in them that indirectly relates to redox properties which are responsible for their antioxidant effects (Heinig and Jennewein 2009, Baba and Malik 2015). *T. wallichiana* plant showed lowest Taxol content near timberline in the upper montane zone beyond which grassland ecosystem (3040 m) at Phurkia, which is the ecotone region and near lower montane zone at the point of human intervention Khati (last habitable point in the valley) (Rai *et al.* 2019). The lowest TC measured was 0.001 mg/g and 0 mg/g of FW at the upper montane zone and lower montane zone respectively.

3.2. Denoising and feature selection of hyperspectral data

The smoothening and filtering technique used here Average Mean filter, Savitzky Golay, and FFT paired with derivative analysis further to select the most optimal band for Taxol. The derivative analysis, a feature selection process is an efficient process in capturing the subtle difference in the spectra required to locate any specific feature present in the spectra. It works best when paired with an optimal pre-processing technique. The spectra depicted in figure 4(a) are the raw spectra of T. wallichiana while figures 4(c), (e), (f) are denoised spectra. The derivative of all the spectra is depicted in figure 4(b), (d), (f), (h). A high-resolution clear sectional view of individual spectra of *T. wallichiana* is represented in figure 5(a), (b) which showed a significant difference among the three methods of denoising.

In **figure 5(a)** it is clearly visible how each smoothening algorithm transformed in regards to the raw spectra of *T. wallichiana*. In the case of moving averages, a least-squares fit is made to a zero-order polynomial (i.e., a straight horizontal line or a constant value). Typically, these features are flattened by other (simpler) averaging points within the filter window (Tsai and Philpot 1998b). The primary factor controlling the degree of smoothing is the size (bandwidth) of the filter window used for convolution or averaging. It can be seen from table 2 from TC 4 - TC 8 that all the bands lies in-between 420-610 nm range, which is the initial reflectance wavelength of the spectra. It was found that the Average Mean filter may not have presented the nearest value to the raw spectra but maintained the peak of the spectra as shown in figure 5(a).

The Savitzky-Golay filtering technique makes use of frequency data or spectroscopic (peak) data. For frequency data, this smoothening method is more effective at conserving the high-frequency components of the signal while upholding the profile and height of waveform peaks (in their case, Gaussian-shaped spectral peaks) (Persson and Strang 2003). The Savitzky-Golay method was the least successful compared to moving average filter and FFT to de-noise the spectra with small disturbance as shown in **table 2.** The wavelength allocated for Taxol content after applying the Savitzky-Golay belongs to the SWIR region (TC 9 to TC 14) of the spectra. The signal after denoising in figure 5(a), (b) have lost the pattern, as well as its reflectance magnitude, was also changed.

In the case of the FFT filter as can be seen from **figures 5(a) and (b)** it is neither enhanced nor reduced the raw signal after application keeping the information in the signal intact but it did lose the patterns (peaks and dips). The wavelength selected using FFT filter showed a negative correlation of 0.320 at 960-970 nm with the SR indices generated values. A negative correlation was observed in the NIR region. Similarly, the derivative analysis of raw spectra shows a negative correlation of 0.370 near the 958-968 nm range. This implies terpenes have a negative correlation in the NIR region and FFT pre-processing is close to raw spectra. FFT pre-processing does modify the signal in such a way that the subtle difference is preserved in the spectra after feature selection. This filter proves to have better correlation results for the Taxol indices generation than the Savitzky Golay filter as shown in table 2 (TC 15 and TC 16).

Contrary to the above filter techniques used, Savitzky Golay changes the spectra magnitude and pattern creating more loss of information. Previously, many studies have stated that the mean filter algorithm is not as good as Savitzky Golay but in our case, retrieval of biochemical variables like Taxol from hyperspectral data was found most suitable. Usually, the first parameter which is retrieved from electromagnetic spectra using Hyperspectral remote sensing is Chlorophyll. This chlorophyll is majorly allocated in electromagnetic spectra after 480 nm (Yang *et al.* 2015), but the visible region expands between 370–700 nm. Hence the information between (370- 450) nm is discarded as noise. TC is majorly allocated in the visible region where noise is much. Moreover, TC is very less hence locating these small peaks for the same was most appropriately done using an Average Mean filter in combination with derivative analysis. The Average Mean filter removed the noise without compromising the ability to resolve fine spectral detail. FFT data provided a smooth spectrum preserving the magnitude of the signal but during absorbance band selection, the number of bands selected due to FFT transformation was very less (Betty and Horlick 1976) making it the second-best filter. These filtered spectra followed by feature selection led to the selection of wavelengths. The advantage and disadvantages of each filter technique were judged based on a statistical correlation between the indices generated values with real estimated values from the field data.

Each spectrum of *T. wallichiana* after application of filter was followed by first derivative to select the absorption bands. These selected wavebands were then included in the ARTMO SI generation toolbox. The derivative analysis presented a different range of wavelengths specified under VIs (370-700 nm), Near SWIR (NSWIR-1350-1450 nm), and Far SWIR (FSWIR- 1800-2500 nm) regions (Hennessy *et al.* 2020). Some of these wavelengths are represented in table 2. The derivative analysis in figure 5(b) suggests that the uneven raw spectra were much smoother after filter application.

3.3. Taxol model development

The regression models are based on spectral indices are typically empirical equations enabling the mapping of biophysical and biochemical parameters over a large area. The specified retrieval strategies within the SI toolbox were first analyzed and then the wavelength selected using the derivative analysis from the raw spectral spectroradiometer data are provided in the text file format.

Novel Taxol indices are generated by identifying various combinations from the spectroradiometer raw data between a spectral range of 350-2500 nm as shown in table 2. The index for each selected combination wavelength was tested. ARTMO model spectral indices (SI) were specified with the preselected wavelength using feature selection (derivative analysis). Using statistical techniques, the Taxol content (TC) retrieval accuracies of newly developed Taxol models were investigated. The best-selected wavelength for band two-band combination yielded a statistically significant correlation for the Average Mean filter. Based on statistical performance two best models were selected. Various curve fitting was also tested with real observed and model-generated Taxol content data, among which the linear curve fitting was found best as shown in **table 3**. The visual comparative representation of all the models is shown in **figure 6** using a Taylor plot. Model LTC-TC 5 and LTC-TC 8 showed a very high correlation of 0.719, 0.718. RMSEr values of both the models are found relatively equivalent i.e., 0.578 and 0.576 for model LTC-TC 5 and LTC-TC 8 respectively. After testing different indices at the selected wavelength, Average Mean denoised spectra found to be the best filter for indices generation, which in combination with the feature selection showed the best statistical performance among all other models. LTC-TC 5 and LTC-TC 8 are the top-performing models which are formed using the combination of wavelengths selected from the visible range values. These values provided ideal results for deriving TC from hyperspectral reflectance data. Similar high relation of this class of compounds i.e., Taxol is recently reported to have an association with the visible region between 400-500nm (Fine *et al.* 2021). Hence, the results obtained are consistent with Taxol indices generated using the visible region reflectance values.

The model generation with ARTMO also gives the best results when the most appropriate bands after smoothening were given as input. The linear curve fitting performed best between modelled data values from ARTMO and TC estimated with HPLC analysis. Hence, the overall results suggests that Taxol content can be quantified using the hyperspectral reflectance in the visible range of 415-421 +/- 5 nm.

4. Conclusion

The result of statistical analysis suggests that the elevation along with its ecosystem climatic conditions plays an important role in variation in phytoconstituents. The highest Taxol content concerning elevation is recorded between 2850 to 3000 m in Nanda Devi Biosphere Reserve. For retrieval of biomedicinal molecule Taxol from hyperspectral data, the average mean filter was found most suitable. TC found in leaves of *T. wallichiana* was very less in terms of quantity hence locating these small peaks corresponding to it in the reflectance curve was most appropriately done using an Average Mean filter in combination with derivative analysis for indices generation. The SI assessment through ARTMO provides a systematized approach in a streamlined way for the selection as well as the assessment of the most precise and sensitive SI formulations which can be used for parameter retrieval using hyperspectral datasets. ARTMO generated SI clearly shows that the TC can be quantified using spectral indices and the model developed using these indices shows the best results in terms of r and RMSEr. The linear curve fitting with modelled data values from ARTMO correlated best with measured TC. Empirical methods like the regression-based model are the best tool to monitor the health of the plant on a real-time basis as it takes less time to compute and is easy to use.

In the future, sampling at more locations in the Himalayas will be performed with the inclusion of seasonality to check the robustness of the model developed. The other region between 370-480 nm of the spectra also needs to be rigorously analysed as it holds more valuable information. This requires a network for the collection of ground samples which becomes quite difficult due to elevation and harsh weather in the region. The current study will reduce the time and tedious effort of researchers and will make the management of canopy-level information much simpler. Compared to conventional labour-intensive on-site measurements, the proposed method will deliver quick information about the Taxol content and thus can help in protecting and managing forest resources more realistically.

Funding

A.G.'s is funded under the University Grant Commission's Junior Research Fellowship program. This work is funded by the *National Mission on Himalayan Studies,* G.B. Pant National Institute of Himalayan Environment (NIHE), Ministry of Environment, Forest & Climate Change (MoEF&CC), Government of India.

Acknowledgment

The authors are thankful to the University Grant Commission and National Mission for Himalayan Studies for the necessary financial assistance and support throughout. The authors also acknowledge the Institute of Environment and Sustainable Development, Banaras Hindu University, and Central Institute of Medicinal and Aromatic Plants, Lucknow, India for providing the necessary laboratory support for the study. The authors also extend their sincere thanks to NMHS, G.B. Pant National Institute of Himalayan Environment (NIHE) for their constant support in this work. The authors also extend their gratitude to Dr. Jochem Verrelst, Imaging Processing Laboratory (IPL) at the University of València, Spain for providing the ARTMO tool and guidance for the presented work.

Conflict of interest

There is no conflict of interest

References

Appendino, G., Gariboldi, P., Gabetta, B., Pace, R., Bombardelli, E. & Viterbo, D.J.J.O.T.C.S., Perkin Transactions 1, 1992. 14β-hydroxy-10-deacetylbaccatin iii, a new taxane from himalayan yew (taxus wallichiana zucc.). (21), 2925-2929.

Baba, S.A. & Malik, S.A., 2015. Determination of total phenolic and flavonoid content, antimicrobial and antioxidant activity of a root extract of arisaema jacquemontii blume. *Journal of Taibah University for Science,* 9 (4), 449-454.

Bala, S., Uniyal, G., Chattopadhyay, S., Tripathi, V., Sashidhara, K., Kulshrestha, M., Sharma, R., Jain, S., Kukreja, A. & Kumar, S.J.J.O.C.A., 1999. Analysis of taxol and major taxoids in himalayan yew, taxus wallichiana. 858 (2), 239-244.

Betty, K.R. & Horlick, G.J.a.S., 1976. A simple and versatile fourier domain digital filter. 30 (1), 23-27.

Assessing the niche of *Rhododendron arboreum* using entropy and machine learning algorithms: role of atmospheric, ecological, and hydrological variables

Akash Anand,[a] **Prashant K. Srivastava**,[a,*] **Prem C. Pandey**,[b] **Mohammed L. Khan**,[c] **and Mukund D. Behera**[d]

[a]Banaras Hindu University, Institute of Environment and Sustainable Development, Remote Sensing Laboratory, Varanasi, Uttar Pradesh, India
[b]Shiv Nadar University, Greater Noida, Uttar Pradesh, India
[c]Dr. H. S. Gour University, Department of Botany, Sagar, Madhya Pradesh, India
[d]CORAL, IIT Kharagpur, IIT Kharagpur, West Bengal, West Bengal, India

Abstract. Species distribution models (SDMs) have been used extensively in the field of landscape ecology and conservation biology since its origin in the late 1980s. But there is still a void for a universal modeling approach for SDMs. With recent advancements in satellite data and machine learning algorithms, the prediction of species occurrence is more accurate and realistic. Presently, four machine learning and regression-based algorithms, namely, generalized linear model, maximum entropy, boosted regression tree, and random forest (RF) are used to model the geographical distribution of *Rhododendron arboreum*, which is economically and medicinally important species found in the fragile ecosystem of Himalayas. To establish complex relation between the occurrence data and regional climatic and land use parameters, several satellite products, namely, MODIS, Sentinel-5p, GPM, ECOSTRESS, and shuttle radar topography mission (SRTM), are acquired and used as predictor variables to the different SDM algorithms. The performance evaluation has been conducted using the area under curve (AUC), which showed the best result for Maxent (AUC = 0.871) and poor result was observed for RF (AUC = 0.755) among all. The overall prediction confirmed the distribution of *Rhododendron arboreum* in the mid to high altitudes of central and southern parts of the Garhwal Division. We provide crucial evidence that combining multisatellite products using machine learning algorithms can provide a much better understanding of species distribution that can eventually help the researchers and policymakers to take the necessary step toward its conservation. [DOI: 10.1117/1.JRS.16.042402]

Keywords: species distribution model; Maxent; boosted regression trees; *Rhododendron arboreum*; Himalayan ecology; species occurrence.

Paper 210393SS received Jun. 30, 2021; accepted for publication Jan. 7, 2022; published online May 26, 2022.

1 Introduction

Since the start of the century, humans started to recognize the importance of regional ecosystem in explaining the distribution of flora and fauna. As a result, the understanding of species-specific geographical presence has become an important aspect especially considering the global concerns of climate change, altitudinal range shift, species invasions, and depletion of endangered species.[1,2] Modeling the potential distribution of a plant species is typically achieved by one (or more) of the several modeling methods. They use exploration of diversity patterns exploration to investigate the distribution depending upon species identity and based on different input parameters, such as climatic data, land use data, soil data, presence/absence data, climatic condition, and its projection data for the generation of suitability maps. These models are sensitive to abundance patterns, altitudinal variations,[3] latitudinal variation ranges, and climate change scenarios. The technique used to model species geographical distribution is termed as species

*Address all correspondence to Prashant K. Srivastava, prashant.iesd@bhu.ac.in; prashant.just@gmail.com

distribution models (SDM), also known as ecological niche modeling, bioclimatic envelop modeling, or bioclimatic modeling. It provided solutions to some of the core issues in ecology, evolution, and its conservation. With the advancement in SDM algorithms, there is still a need to better understand the nonlinear interactions of species with local parameters as prediction based on extrapolation was found to be nonrobust, especially with the conventional approaches.[4,5]

In recent years, several studies have been carried out to solve computational problems and implemented neural network[6] and machine learning models in the study, which are valuable tools for modeling many phenomena in ecology, mathematics, medical, economics, physics, and engineering.[6–8] Some of their significant applications were introduced in the research works of Jamali et al.,[6] Radmanesh and Ebadi,[7] Fouladi et al.,[8] Rafieipour et al.,[9] Heydarpour et al.,[10] and Altaher et al.[11] Also, the authors in Refs. 6 and 8 given several remarkable studies on the theory, analysis, and recent historical development of the neural network and computational studies.

According to the niche theory,[12] a species can only be found in a region where the combination of local bioclimatic gradients allows the species to have positive population growth. This theory conceptualizes the regional species environment and its occurrence considering the absence of immigration. While further extending the theory, it can also be realized that the variation in species traits allows them to inhabit different niches or cohabit in a particular spatial extent. These interactions are ecologically complex and nonlinear; therefore, the role of machine learning is crucial in understanding their distributions.[13,14] But before the introduction of machine learning in SDM, several theories and models were proposed by ecologists and researchers widely used to predict the distribution of plants and animals. BIOCLIM[15] and DOMAIN[16] are among the earliest SDMs that received global acceptance due to their less complex algorithm and easy to use interface. For establishing the nonlinear relation between input parameters, several machine learning iterative algorithms are proposed that give much better accuracy than the linear models. Boosted regression trees (BRT)[17] and random forest (RF)[18] are among the widely accepted iterative models, especially for modeling species distribution.[19] On the other hand, the maximum entropy (Maxent) model[20] is based on envelop model, which takes the presence-only data as its input parameter. Another one is based on the conventional regression-based learning technique called generalized linear model (GLM),[21] which can consider multiple measurement levels of response values using different link functions. Among the above mentioned SDM's, Maxent is widely expected algorithm due to its robust and nonlinear modeling techniques.[22] Maxent models are able to satisfy all known variables without any subjective assumptions, which is not present in earlier SDM models (such as Bioclim/DOMAIN). Therefore, it is more robust than earlier SDM because of the following inherent merits that involve improved mathematical modeling, machine learning, and statistical tools with better predictive accuracy. These SDM models have efficient deterministic algorithms that can be benefit to predict species' optimal probability distribution at the study sites. They are less sensitive to the various environmental variables and changes occurring in them. They consider interactions between environmental variables and minimize overfitting problems.

Major drawback of SDMs are the availability of non-uniform and relatively lesser field observations as compared to the area of interest, therefore models are generally extrapolated beyond their sampling sites. These spatial and geographical-extrapolations based on limited species sampling often lead to spurious results. A major limitation of macroecological SDMs is the inability to predict species identity and thus mainly involved species richness, i.e., emergent ecosystem properties implemented for exploring macroecological phenomena. Even the probability distribution is not uniform in earlier SDMs, thus the stability is lower than expected. Species distribution results depend on the spatial resolution chosen for the extent mapping, and also temporal aspects play a significant role in species. Therefore, models having functioning of species ecological distribution at the relevant scale are needed.[23] References 24 and 25 suggested that the unavailability of data or insufficient data-based predictions using extrapolation are limiting to the true species distribution in the region, which was supported by the study conducted by Ref. 26. Therefore, while using SDM, one has to understand data quality, sufficient data, predictor variables (hydrometeorology), and reliability of the models for distribution output. There is some development that happened in past, but there is still a lack of modeling techniques for understanding the complex relationship between different regional input parameters. As per the fundamental assumptions of SDM, the target species is considered to be in equilibrium with the predictor variables, which is highly criticized in the past and still there is not a relevant alternative.[27] The recent developments in theoretical ecology, remote sensing techniques, and modeling algorithms have now enabled the ecologists to model near real-time distribution of the species, especially with the data coming at much finer spatial and temporal resolutions.[28–30] Also with the introduction of sensors such as Sentinel-5p and ECOSTRESS, it is now possible to assess the relation of greenhouse gases and evapotranspiration with species distribution, which was missing in the conventional studies. Earlier, ecologists preferred to use the data from Worldclim,[31] NCEP,[32] and ECMWF[33] as predictor variables that provide spatially interpolated atmospheric datasets at 0.1 deg to 2.5 deg of spatial resolution. The coarse resolution data were the major source of uncertainty and error, also they did not allow the model to predict the distribution of species at the regional scale. Particularly for the topographical conditions of the Himalayas, which varies drastically, it needed fine-scale satellite products.

The Himalayas being home to thousands of economically, medicinally, and rare flora and fauna is experiencing global climate change.[34,35] In their study, they reported that the overall warming in the Himalayas is consistently increasing for the past 100 years, and the rate is much higher than the global average of 0.74°C.[36,37] The temporal change in the distribution of species has been reported by several researchers and their impact on regional ecology.[38–40] One such species is *Rhododendron arboreum*, which comes from the Ericaceae family and dominantly found in the Himalayas, South India, Nepal, and Sri Lanka.[41,42] It is an economically and medicinally important species and sustains itself in the fragile ecotone of alpine and subalpine regions. The continuous change in regional climatic conditions is imperative to model the distribution of species so the biodiversity and conservation of the ecosystem can be maintained.

The main contribution of this work is to uncover the following:

- Impact of environmental variables on the distribution of *Rhododendron arboreum* at the study site.
- Linking variables, such as normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), evapotranspiration (ET), fraction of photosynthetically active radiation (fPAR), water vapor, leaf area index (LAI), land surface temperature (LST), precipitation, ozone, NOx, albedo, aerosol absorbing index (AAI), and digital elevation model (DEM) for understanding the distribution of species in Himalayan environment.
- Assessment of the variables on geographical distribution of *Rhododendron arboreum* through machine learning and entropy models.

Therefore, in purview of the above and considering uniformity in probability and stability of Maxent, this study is focused on establishing the relation between different bioclimatic and environmental parameters to model the distribution of *Rhododendron arboreum* within the study area. In Sec. 2, we provide an overview of the study site, specifications of target species, modeling algorithm employed in the study, and performance evaluation metrics. Section 3 explains the model result and discussion part, in which the species distribution maps along with the discussion are presented. Section 4 provides some conclusions and gives suggestions for future research.

2 Materials and Methods

2.1 *Study Area*

The Himalayas is one of the most complex and diverse ecoregions, and it offers rich biodiversity and has been home to thousands of floras and fauna. The complex topography allows some of the rare, medicinal, and economically important species to grow within this region. This study is conducted within the Garhwal Division of Uttarakhand state where the elevation ranges from 416 to 7801 m above mean sea level. As shown in Fig. 1, this region has several biomes, namely tropical evergreen and deciduous broadleaf forest, tropical and subtropical coniferous forest, temperate coniferous forest, temperate savanna, grassland, and shrubland as well as it has a significant number of glaciers as well. The region falls under subtemperate to temperate climate,

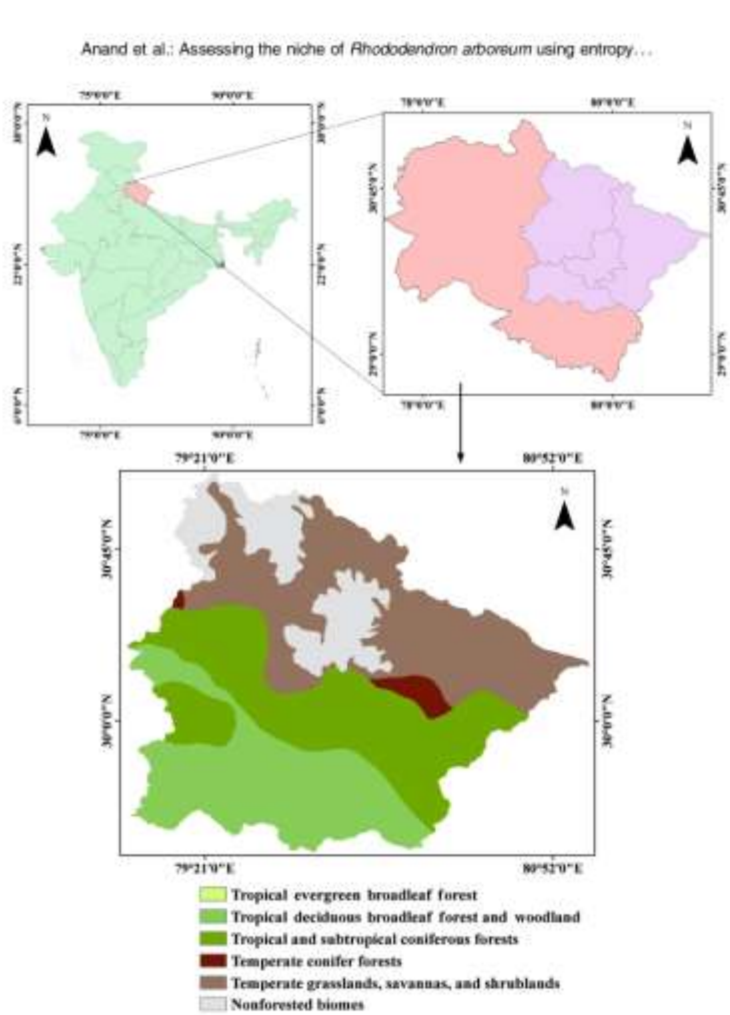


Fig. 1 Study area map showing the distribution of biomes.

and the hilly terrain with densely forested slopes receives significant rainfall from mid-June till September with several occasional rain events throughout the year, whereas ~20% of the area is covered with snow throughout the year. Due to its rich ecology and complex environmental conditions, it has always attracted the attention of researchers and scientists from around the world, especially with the recent trend in global warming, this region is showing early impacts of climate change that makes this region more important.

2.2 *Target Species and Occurrence Data*

The geographical distribution of *Rhododendron arboreum* is performed for the study area using *in situ* occurrence data and predictor variables. *Rhododendron arboreum* has a great biological significance and dominantly found in the Himalayas. The occurrence of this species was reported between 1200 and 4000 m above mean sea level.[42] *Rhododendron arboreum* is a high valued species both in terms of medicinal and economic importance, also it is reported by ecologists that it possesses some characteristics of invasive species.[43,44] Medicinally, *Rhododendron arboreum* is anticancerous, antioxidant, antidiabetic hepatoprotective, antimicrobial, diarrhoea, and antinociceptive.[44] Whereas a study in Ref. 45 reported that the squash made of *Rhododendron arboreum* is used for the treatment of intellectual disabilities. Economically, the species used in making jams, local brews, squash, and in different juices, also its wood is used as fuel, and the leaves are used for treating the bedbug bites. With such importance, mapping its distribution is very crucial for ecosystem conservation and making necessary strategies for its protection.[46] The *in situ* data collection was conducted within the study area during September 2019 and Match 2021, covering the complex topography of the region. A total of 70 homogeneous patches of *Rhododendron arboreum* were identified at different elevations. Among the collected occurrence data, two-third used for model development and one-third is used for validation purpose. To mark the occurrence of the species, handheld Garmin GPS is used with a horizontal accuracy of 95% ± 9.3 m.

2.3 *Predictor Variables*

For establishing complex relationship between the occurrence data and local ecosystem, several environmental predictor variables are acquired from different satellite sensors. The selection of predictor variables is performed on the basis of their direct impact on the vegetation. The environmental predictor variables include MODIS products, namely, NDVI, EVI, fPAR, and LAI. The sentinel-5p sensor provides information related to the greenhouse gases with a high spatial and temporal resolution that makes it very crucial in the regional ecological study considering climate change. The Sentinel-5p product used as predictors is AAI, water vapor, albedo, ozone (O_3), nitrogen dioxide (NO_2), carbon monoxide (CO), and sulfur dioxide (SO_2).[47] Apart from the climatic parameters provided by Sentinel-5p, LST and (ET) are acquired from ECOSTRESS sensor.[48] Precipitation data are obtained from GPM dataset. Topography plays a vital role in studying species distribution, especially in the topographically complex regions such as the Himalayas. Therefore, DEM provided by the SRTM[49] is used as an input parameter for the calculation of SDM. Also, the regional biome is taken into consideration as forest biomes directly influence the species distribution.[50]

2.4 *Species Distribution Modeling Algorithms*

The regional dynamics and distribution of species, especially their prediction under the influence of climate change, is a crucial issue for ecological biodiversity and conservation. The concept of SDM started in early 1980s, but before that, mapping of species distribution is only limited to *in situ* sampling. With the availability of satellite-based climatic data, several presence-only based model is designed and tested in the early 2000s, which includes BIOCLIM and DOMAIN. After that, several other models were introduced based on establishing the correlation between occurrence data and the climatic predictor variables. These models are easy to implement but did not provide significant accuracy and failed on the regional scale. As the interaction between species and respective ecological parameters is extremely complex, nonlinear models must establish their relationship. Therefore, some of the most popular and widely used machine learning algorithms are trained to model the geographical distribution of *Rhododendron arboreum* within the study area.

2.4.1 *Maxent*

Maxent is one of the widely used SDM that works on estimating the probability distribution through maximum entropy of input parameters. Reference 51 defined that entropy is the measure of the total number of choices involved in selecting a particular feature. The basic concept of Maxent was first introduced by Ref. 52, which stated that the best way to ensure the approximation is by testing the results with known positions, and the distribution must have maximum entropy, it is also known as the maximum-entropy principle. Maxent owes its success in species distribution because the entropy reaches a minimal value that is highest to that of a species among the probability distributions of all the species. This is achieved by predicting the occurrence of species through the distribution, which is mostly spread out or tends to have a uniform distribution. This, in turn, necessitates having the information of all the environmental variables of the locations. The Maxent algorithm uses many points taken during ground data sampling that are referred to as background points, and these background points define the current environmental variables.[53] These variables are used as constraints that limit the rule for the predicted distribution. In general, Maxent considers linear/quadratic/product/threshold/hinge/categorical features as constraints that define the rule to confine the expected distribution. These features have different implications for the constraints. This work has used a categorical feature called "biome," which is defined as types of regional land use. This constraint specifies the proportion of predicted occurrences in each category to be as close to the proportion of observed occurrences in each category.

In this work, two probability densities are calculated. These densities provide the relative likelihood of all environmental variables over the range of background points. The algorithm then calculates the ratio between these probability densities to find the relative ecological suitability for the occurrence of the rhododendron for the given point in the study area. In this manner, the Maxent chooses the distribution that has maximum similarity between the environmental characteristics of the given climate and the locations where the required species are supposed to be abundant. This is the raw output of the algorithm, which is logically transformed by considering the prevalence value. In this work, the value is taken as 0.5, which implies that the species is present in half of all the possible locations. The limitation of this algorithm lies in the fact that it provides environmental suitability rather than the predicted probability of occurrence.[41]

The mathematical measure of the uniformity of a conditional distribution $\hat{\pi}(x)$, which is provided by the conditional entropy, is given as

$$H(\hat{\pi}) = -\sum_{x \in X} \hat{\pi}(x) \ln \hat{\pi}(x). \tag{1}$$

The entropy is bounded from below by zero, the entropy of a model with no uncertainty at all, and from above by $\ln \hat{\pi}(x)$, the entropy of the uniform distribution over all possible $|\hat{\pi}(x)|$ of x. This acts as a base for presenting the principle of maximum entropy.

To select a model from a set S allowed probability distribution, choose the model $p^* \in S$ with maximum entropy $H(\hat{\pi})$:

$$p^* = \arg\max_{p \in S} H(\hat{\pi}). \tag{2}$$

It can be shown that p^* is always well-defined, that is, there is always a unique model p^* with maximum entropy in any constrained set S. The threshold chosen in the current method is the value for which the sum of sensitivity and specificity is highest, and the probability model prediction will be transformed to a binary score of the presence or absence of the species. The required solution is achieved by maximizing the gain function that is a penalized maximum likelihood function that is given as

$$\text{Gain} = \frac{1}{m}\sum_{i=1}^{M} z(x_j)\lambda - \log\sum_{i=1}^{N} Q(x_j)e^{z(x_j)\lambda} - \sum_{j=1}^{J} |\lambda_j| * \beta * \sqrt{s^2[z_j]/M}, \tag{3}$$

where the likelihood of the presence data are the sum of predicted values at presence locations is given by $\frac{1}{m}\sum_{i=1}^{M} z(x_j)\lambda$, the likelihood at all the background locations is the sum of the predicted values at background locations is given by $\sum_{i=1}^{N} Q(x_j)e^{z(x_j)\lambda}$ and overfitting penalty to be used in regularization is given as

$$\sum_{j=1}^{J} |\lambda_j| * \beta * \sqrt{\frac{s^2[z_j]}{M}}, \tag{4}$$

where β is the regularization coefficient, $s^2[z_j]$ represents the variance of feature j at presence locations, and $Q(x_j)$ is the prior distribution. A significant characteristic of Maxent is regularization that helps in reducing the overfitting of the model. This is achieved by setting confidence intervals across the constraints and excluding the features that are not significant. The regularization used is the least absolute shrinkage and selection operator.

2.4.2 *Random forest*

Random forest algorithm[54] is developed using the classification and regression tree approach, and it has shown significant performances in different application of remote sensing, including forestry,[55] ecology,[56] classification,[57] and climate change.[58] Random forest is defined as a collection of weak learners having a tree structured interface with uniformly distributed random vectors where each tree provides a prediction for the resultant variable. While generating the group of weak learners based on the bootstrap of the data, the overall calculation converges on an optimal result avoiding the issues of general parametric, classification, and regression statistics. Bootstrap of training, independent variables (M) at each node, and retaining the variables that provided the most useful information sum up the significance of RF algorithm for both regression and classification purposes. To assess the information and purity of the node, Gini entropy index is used.

The algorithm for the prediction of a new variable can be explained as

$$\hat{f}_{rf}^{B}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x), \tag{5}$$

where $\hat{f}_{rf}^{B}(x)$ is the bootstrapped RF function for predicting x, T_b is the RF tree, and B is the number of bootstraps.

2.4.3 *Boosted regression trees*

BRT is one of the most popular machine learning techniques that is widely used for SDM. It is based on improving the performance of a model by fitting multiple models and combine them for prediction. Broadly, BRT uses two algorithms, first one is the regression trees based on classification and regression of input parameters and the second builds the boosting and combines the collection of multiple models. Tree-based models divide the predictors into small clusters using a series of rules to identify the region having the most homogeneous response during prediction. The regression tree fits the mean response from predictors in any specific region. As per Ref. 59, the best way to fit a decision tree for growing a large tree and then trimming it by eliminating the weakest links identified through cross-validation. Decision trees are widely used because they are easy to implement, visualize, and one of the most flexible algorithms, especially for species distribution modeling.[60] At the same time, boosting technique is used for improving the accuracy of the model based on its background architecture of finding and averaging multiple rules of thumb rather than a single rule.[61] While other techniques include bagging, stacking, averaging, and merging the results from multiple models, boosting works as sequential models based on forward and stagewise procedures. The AdaBoost boosting algorithm is used to determine species distribution by fitting the predictors using sequential iteration technique.

2.4.4 *Generalized linear model*

GLM is also termed as an extension of the classical linear regression model, where the transformation is achieved to get a normal distribution for the dependent variable. In a GLM-based model, predictor variables are used to calibrate the model, and the link function is selected based on the statistical distribution of dependent variables. GLM being a parametric function is not optimized using the least-square method, rather it uses the maximum likelihood method for model optimization.[62–64] GLM model has a set of distribution function including binomial, Poisson, gamma, etc. in which gamma distribution having link function $f(x) = \mu$, where μ is the mean value of predictor variables, is used for establishing the relation between the predictors.

In GLM, the predictors $X_i (i = 1, 2, \ldots, n)$ are joined together to get a linear predictor (LP), which is related to $\mu = E(\gamma)$, where γ is the response variable, to the link function $f()$, which is represented as

$$f(E(\gamma)) = LP = \propto + X^T \beta, \tag{6}$$

where $\propto$ is the intercept, $X^T = X_1, X_2, \ldots, X_n$ is the vector of p variable, and $\beta = (\beta_1, \beta_2, \ldots, \beta_n)$ is the regression coefficient.

Therefore, for i'th observation,

$$g(\mu) = \propto + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_n X_{in}. \tag{7}$$

The error rate within the model is resolved using the least square algorithm[65] during the model fitting.

2.5 *Model Validation*

In any regression and classification-based machine learning model, the estimation of model performance and classification accuracy is an important task. Therefore, to evaluate the overall accuracy of the distribution modeling, AUC parameter evaluation matric is used for all four machine learning algorithms. Primarily, the data are divided into training and testing sets, in which one-third data are assigned for the model validation, and rest is allocated for model development. AUC is widely used for model validation, especially in binary classification models. AUC is a threshold-independent evaluation metric that validates the model performance at various discrimination thresholds. At each discrimination threshold, the true positive rate (TPR), also known as the probability of detection or sensitivity and the false positive rate (FPR), also known as the probability of false alarm, is estimated and plotted against each other. The TPR and FPR for each point (x, y) are plotted together to get the final AUC curve. It is also explained as

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100, \tag{8}$$

whereas specificity is calculated as

$$\mathrm{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \times 100, \tag{9}$$

$$\mathrm{FPR} = 1 - \mathrm{Specificity}, \tag{10}$$

where TP is the true positive, and FP is the false positive values. Specificity defines the true negative rate, whereas TPR calculates the percentage of correctly predicted values. The AUC value varies from 0 to 1, where the value closer to 1 shows the accurate classification, and the values close to 0 denote poor classification accuracy.

3 Results and Discussion

The geographical distribution of a particular species is dependent upon the regional climatic, topographical, and land cover conditions. To build a more relevant SDM, the predictor variable should directly influence the existence and growth of the species. The occurrence data for the target species, *Rhododendron arboreum*, are collected within the Garhwal Division of Uttarakhand state, where the topography and climatic conditions are complex. To establish a relationship between species occurrence and its regional climatic condition supporting its existence and growth, 16 predictor variables are considered input parameters to different machine learning-based SDMs. The yearly trend is analyzed for each input parameter so the generated relation can be widely accepted and irrespective of any short-term bias caused by local weather conditions. The machine learning algorithms used in developing SDMs are Maxent, GLM, RF, and BDT. The algorithms are intercompared and validated using statistical evaluation matrices.

3.1 *Input Variables*

Several satellite-based input variables are acquired for modeling the species distribution. With better spatial and temporal resolutions, products from the sensors such as MODIS, Sentinel-5p, GPM, and ECOSTRESS are widely used in analyzing the regional and global ecology. Presently, MODIS products, namely NDVI, EVI, LAI, and fPAR, are used in the SDM algorithms. These variables indicate the information related to the surface reflectance, productivity, energy transfer by the vegetation, water cycle processes, and other biophysical and biochemical properties of the vegetation, and the spatial resolution of these data varies from 250 to 500 m. The products from ECOSTRESS, namely ET and LST, is responsible for providing the information related to plant water consumption and regional temperature levels, the thing that makes the ECOSTRESS products more valuable is their spatial resolution, it provides data at 70 m spatial resolution and has comparatively better temporal revisit. Sentinel-5p being one of the most recent sensors that provide data of different atmospheric gases, including the greenhouse gases at global with a spatial resolution of 0.01 arc degree, which is better than other satellite sensors currently active. Sentinel-5p provides a wide range of atmospheric data, but the data that highly influence *Rhododendron arboreum* are taken as model inputs and are, namely, ozone, nitrogen dioxide, albedo, carbon monoxide, sulfur dioxide, and AAI. Precipitation is one of the most important parameters for SDM acquired from GPM at a spatial resolution of 0.1 arc degree, whereas DEM is used to consider the topography. As the Himalayas is made up of different biomes, this study area has five biomes as listed in Fig. 1, and the biome information is also used as a predictor variable to the SDM.

All the predictor variables are shown in Fig. 2, which is the yearly average to maintain the temporal consistency, and all are resampled so they can match each other on the pixel level. The estimated NDVI and EVI values are varying from −0.08 to 0.84 and −0.11 to 0.48, respectively, in which it was found that the southern part of the study area is having high vegetation content in the tropical evergreen and deciduous forest than the northern part where the temperate forests dominate the biome. NDVI is used to identify the green vegetation, and EVI can enhance the vegetation signal by reducing the canopy background noises. ET also supported NDVI and EVI results, as it varies from 3.34 to 37.2 $\mathrm{kg/m^2}$ where the maximum values are observed around the boundary of tropical deciduous and tropical and subtropical conifer forest shoeing the latent heat flux coming from the earth surface. Also, the fPAR and LAI values are highly correlated with ET and varying from 0 to 0.87 and 0 to 5.48, respectively. The LST is varying from 259.16 to 300.89 K, the value of LST is higher in the southern part, and it is gradually decreasing in the northern direction due to increase in the elevation range, which is in between 416 and 7801 m, that shows the elevation drop in the region and its impact on LST. Precipitation is also very low in the northern part, as it almost covered with snow throughout the year, whereas the middle and southern regions have significant average rainfall in between 0.049 and 0.188 mm/h, higher values are seen in the Pithoragarh region, where the forest is dominated by tropical evergreen biome. The sentinel-5p-based parameters are also provided significant information regarding the regional climate condition throughout the year. Water vapor is one of the major greenhouse gases found to be higher in the highly forested regions and lower in the higher altitudes. The range of observed water vapor is between 151.86 and 1933.73 $\mathrm{mol/m^2}$. The quantity of water vapor has a direct impact on plant growth and photosynthesis. The ozone layer is found to be higher in tropical forests and lesser in the higher altitude showing the thinner atmosphere in the northern part, it is varying from 0.1207 to 0.1258 $\mathrm{mol/m^2}$. The higher nitrogen dioxide value affects the plant growth, and currently, it is ranging from 4.6e-005 to $6.2e-005$ $\mathrm{mol/m^2}$ in which the higher values are found in the southwestern part where the forest density is low. The albedo and AAI have shown similar results with an observed value between 0.21 to 0.83 and −1.85 to 0.04, respectively. Albedo and AAI are higher in the higher altitudes and low values in forest area due to their high absorption by the dense vegetation. Carbon monoxide varies from 0.014 to 0.036 $\mathrm{mol/m^2}$ with higher values over the forested region, whereas sulfur dioxide varies between −0.00055 and 0.00061 $\mathrm{mol/m^2}$. Both carbon monoxide and sulfur dioxide are major greenhouse gases and have an impact on plant growth and distribution.

All the input parameters have major significance over the distribution of species. Although increasing the parameter may improve the accuracy, the model will become more

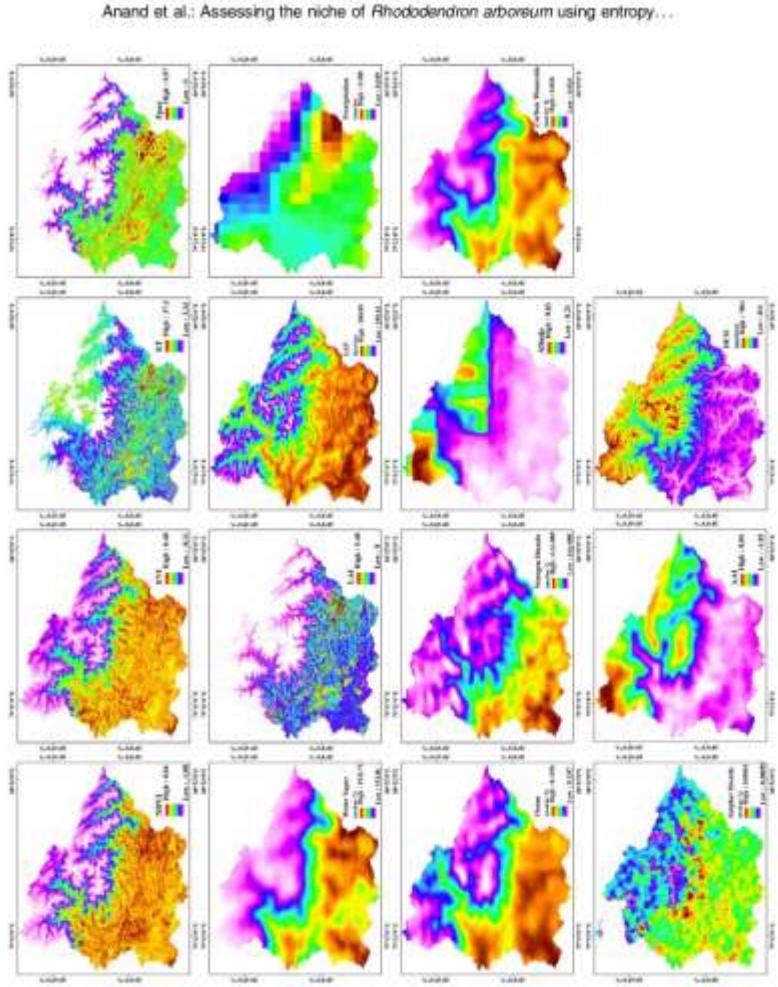


Fig. 2 Input variables for SDM.

computationally complex and parametric bias will also increase. Therefore, the predictor variable is limited to 16 in this study so a more robust model is designed. Rather than considering the satellite data during the sampling period, an overall trend is used to generate the mean value of each predictor variable. The trend of the predictors holds great importance as it demonstrates the change in climatic and land use condition throughout the year. The yearly trend of each variable is shown in Fig. 3. NDVI, EVI, and ET have demonstrated a similar trend, as the values are minimum in January, which linearly increased until the monsoon and gradually decreased in the winter. fPAR and LAI have similar trends as the maximum value is observed before and after the monsoon season when the insolation is on its peak. Precipitation and water vapor also followed a similar curve where the value is high in the monsoon period, and LST gradually increases until the monsoon and then starts decreasing. Atmospheric variables also have different responses to the local weather, and NO_2, albedo, and AAI values are high during monsoon, whereas O_3 is highest during January and gradually decreases the entire year. SO_2 value is lowest during May to September and progressively increases in winter.

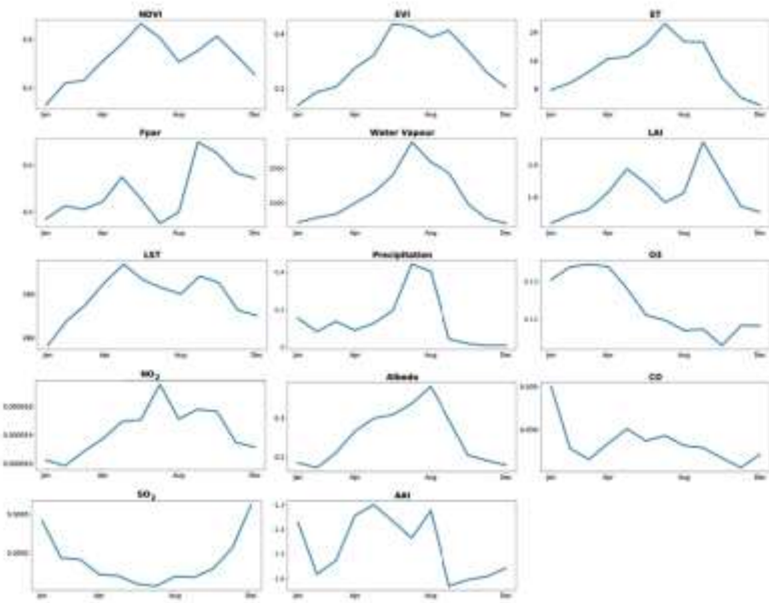


Fig. 3 Trend analysis of input variables for year 2020.

3.2 *Species Distribution*

The distribution of *Rhododendron arboreum* is predicted using four machine learning and regression-based algorithms considering 16 climatic and environmental datasets as input parameters. The prediction has been made for a spatial resolution of 100 m as shown in Fig. 4. As GLM is based on a regression-based linear modeling approach, the prediction given by GLM is showing overestimation with the higher probability of *Rhododendron arboreum* occurrence in the central region between tropical and subtropical coniferous and temperate conifer forest. The distribution predicted by Maxent is largely covering the mountains' tails and dominantly occurring in the central and southern parts of the study area. A similar result is shown by the BRT algorithm in which tree-based relation is generated within the input parameters. BRT is establishing major presence of *Rhododendron arboreum* in the southern part of the study area and some distributed patches on the northern side. On the other hand, RF vastly underestimated the prediction and has only shown higher probability of *Rhododendron arboreum* occurrences around the center of the study area and some occurrences on the south side. Overall, it is observed that GLM is overestimating the species distribution prediction, and RF is underestimating the same. But Maxent and BRT are showing promising results. The distribution of *Rhododendron arboreum* is largely found in the central and southern parts of the study area, and a higher probability can be seen near the tails of high topographic mountains.

3.3 *Model Validation*

The AUC curve is used as an evaluation metric to validate the prediction made to model the distribution of *Rhododendron arboreum* using four different machine learning and regression algorithms, as shown in Fig. 5. As AUC value varies from 0 to 1, with the value nearer to 1 is showing the high probability that the species is present. The maximum AUC value is recorded by Maxent and is 0.871, which shows that Maxent is the most promising machine learning model to assess species distribution on regional scale. After that, 0.835 AUC is recorded for GLM, which is also a considerable value in modeling species distribution, but the overestimation of the

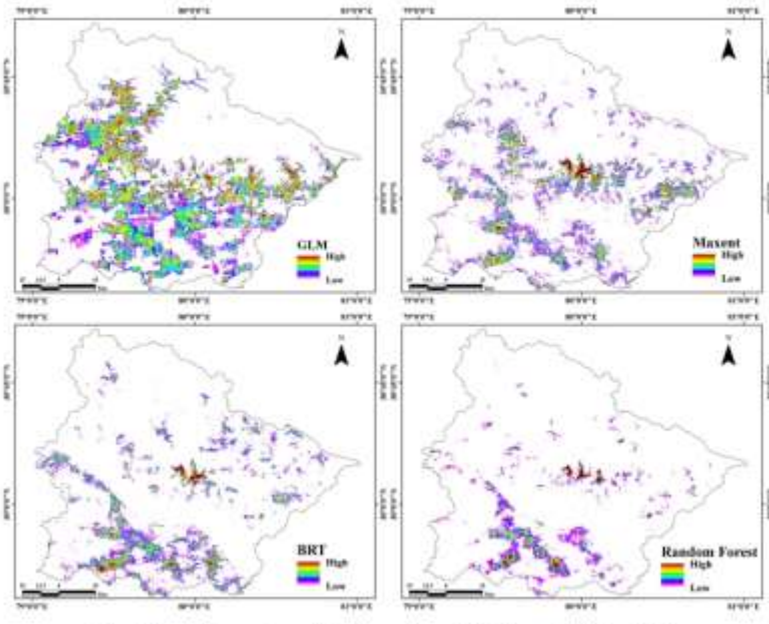


Fig. 4 Predicted species distribution using GLM, Maxent, BRT, and RF.

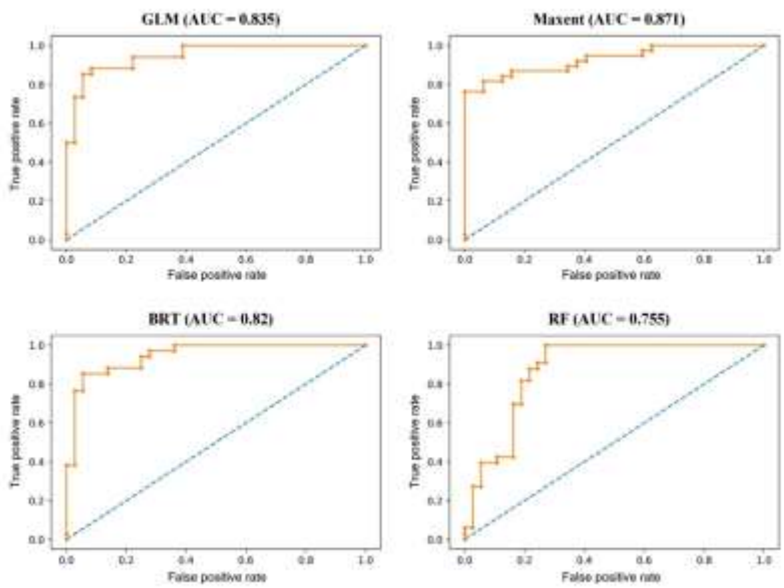


Fig. 5 AUC curve for each SDMs.

species in GLM is something that is more concerning. Apart from that, BRT has given an AUC value of 0.82 and according to its prediction, it has provided more precise values than even Maxent at some places. An underestimation is observed in RF with an AUC of 0.755, making it the worst-performing machine learning model among the four. Overall, it is observed that Maxent is performing well, and BRT has also shown promising results for modeling the distribution of *Rhododendron arboreum*.

The previous studies, such as Reiss et al.,[66] compared several models, such as MAXENT, RF, and SVM (support vector machine) for species distribution modeling and revealed that they have similar predictive performance. When compared with the other model such as BIOCLIM through their AUC, they found that values are significantly higher than BIOCLIM. In another study by Tsoar et al.,[67] they confirmed that Mahalanobis distance can even predict better than BIOCLIM and DOMAIN. In the study by Elith and Graham et al.,[68] they compered the performance of MAXENT with BIOCLIM and DOMAIN and pointed out that MAXENT gives a significantly higher predictive performance than the later. In Ref. 69, the authors divided the SDMs into two categories; at first they included the best performing one with higher stability such as Mahalanobis distance, RF, MAXENT, and SVM, whereas the second category composed of low stability one with lower performance such as BIOCLIM and DOMAIN. In other study by Giovanelli et al.,[70] they also confirmed the better performance of MAXENT as well as SVM for species distribution modeling and concluded that both SVM and MAXENT can be used. Overall, the above-mentioned studies indicate the superior predictive accuracy of MAXENT in SDM and recommended it for further use. The findings also revealed that the varying performance and stability of SDMs can be linked to changing environmental variables and climatic conditions. The results of this study are in agreement with the previous studies as mentioned above and hence can be used for prediction of *Rhododendron* sp. in the Himalayan region.

3.4 *Future Perspective and Challenges*

SDMs can predict the distribution of species on a regional scale and if well calibrated then for a global scale as well. Several models help in getting an insight into species distribution as well as establish linear and nonlinear interactions between the predictors, but they still lack establishing ecological theories and predefined assumptions. The confined understanding of species response to bioclimatic variables and limited statistical approaches bring error to the SDM. But with the advancement in spatial datasets and statistical algorithms, the prediction accuracy is continuously improving. The availability of satellite images providing bioclimatic, ecological, and ecohydrological responses reduced the uncertainty related to the satellite-based biotic interactions. References 24 and 71 listed several actions to counter uncertainties, including the continuation of ecological and biological research that focuses on biotic interactions, regular and systematic collection of species occurrence data, temporal validation of retrieval models, selection of predictors, and algorithm improvements based on different climatic scenarios.

The integration of bioclimatic and atmospheric variables provides an exciting option for defining the consequences of global climate change on species distribution, especially for future scenarios. But, not all SDMs are optimally suited for predicting the species distribution based on various predictors. The earlier SDMs, namely BIOCLIM and DOMAIN, were based on the predefined hypothesis and did not have the option to integrate atmospheric gases and other satellite products. These early-stage models also lacked establishing complex nonlinear relation between the predictors. So, the data-driven models, namely GLM and BRT, are introduced to model the distribution based on observational data and substantially to integrate ecological hypothesis. These models are based on the observed realized niche and limited to the *in situ* observations and predictor variables. When combined with a certain degree of ecological knowledge, the data-driven models will act as the process-based approach in which the prediction will be more accurate and equally supported by the ecological hypothesis. Fot iterative models such as Maxent and RF, when supplied with process-based predictors, the generated prediction will have better accuracy, and the nonlinear relation will be more precise.

Despite ongoing improvements in SDM algorithms and the satellite-based predictors, the prediction is still influenced by the degree of uncertainty based on the biotic behavior of species and its interactions with changing climate and regional biota. Therefore, it is recommended to use a process-based approach in modeling the distribution of species, which allows prediction beyond the observational data. Machine learning algorithms prove their importance in modeling process-based prediction, and with regular upgradation in the knowledge of species interaction, SDMs are continuously improving. However, additional research still needs to focus on the ecological understanding of species, ecological theories, and combining observational data rather than concentrating only on a data-driven approach. In the last few decades, with the high computing system and database technologies, now big datasets are available to users for deriving efficient outcomes.[72] With the advancement in deep learning, AI and data mining methods have entered a new age[8] that can help in analysis of high dimensional datasets with high accuracy, which now provides an enormous possibility in species distribution mapping also.[73]

4 Conclusion

In this study, we attempt to establish the relation between species occurrence data and their respective environmental predictor variables. The yearly tread of each parameter is analyzed to observe the variations throughout the year and a pixelwise mean value is calculated to be used in the SDM. Machine learning algorithms, namely Maxent, BRT, RF, and GLM, are implemented to establish the relation between predictor variables. The AUC-based performance evaluation matric is generated, and it is found that Maxent is performing better than others with an AUC of 0.871, also BRT has shown promising results with AUC = 0.82, but the GLM and RF are found to be overestimating and underestimating, respectively. The machine learning algorithms performed significantly well, and remote sensing data proved to be a vital source of information in ecological studies, which is continuously improving with regular upgradation in satellite data and algorithms. Although the SDMs are providing better results on regional studies but still lack in explaining the ecological background, the true meaning of their prediction and boundary conditions is a topic of research for future. To summarize, if SDMs are to be a standard tool, the background should be supported by good ecological understanding, and they give a new direction to the future research opportunities in SDM.

Acknowledgments

The authors are thankful to the National Mission for Himalayan Studies (NMHS), G.B. Pant National Institute of Himalayan Environment (NIHE) for the necessary financial assistance and support throughout. The authors declare no conflict of interest.

References

1. T. H. J. F. E. Booth, "Species distribution modelling tools and databases to assist managing forests under climate change," *For. Ecol. Manage.* **430**, 196–203 (2018).
2. C. D. Wilson, D. Roberts, and N. J. B. C. Reid, "Applying species distribution modelling to identify areas of high conservation value for endangered species: a case study using *Margaritifera* (L.)," *Biol. Conserv.* **144**(2), 821–829 (2011).
3. R. G. Mateo et al., "Do stacked species distribution models reflect altitudinal diversity patterns?" *PLoS One* **7**(3), e32586 (2012).
4. R. Kadmon, O. Farber, and A. J. E. A. Danin, "A systematic analysis of factors affecting the performance of climatic envelope models," *Ecol. Appl.* **13**(3), 853–867 (2003).
5. M. S. Wisz et al., "Effects of sample size on the performance of species distribution models," *Divers. Distrib.* **14**(5), 763–773 (2008).
6. N. Jamali et al., "Estimating the depth of anesthesia during the induction by a novel adaptive neuro-fuzzy inference system: a case study," *Neural Process. Lett.* **53**(1), 131–175 (2021).
7. M. Radmanesh and M. Ebadi, "A local mesh-less collocation method for solving a class of time-dependent fractional integral equations: 2D fractional evolution equation," *Eng. Anal. Boundary Elements* **113**, 372–381 (2020).

8. S. Fouladi et al., "Efficient deep neural networks for classification of COVID-19 based on CT images: virtualization via software defined radio," *Comput. Commun.* **176**, 234–248 (2021).
9. H. Rafieipour et al., "Study of genes associated with Parkinson disease using feature selection," *J. Bioeng. Res.* **2**(4), 1–12 (2020).
10. F. Heydarpour et al., "Solving an optimal control problem of cancer treatment by artificial neural networks," *Int. J. Interact. Multimedia Artif. Intell.* **6**(4), 18–25 (2020).
11. A. Altaher et al., "Using multi-inception CNN for face emotion recognition," *J. Bioeng. Res.* **3**(1), 1–12 (2020).
12. M. F. Udvardy, "Notes on the ecological concepts of habitat, biotope and niche," *Ecology* **40**, 725–728 (1959).
13. R. K. M. Malhi et al., "Synergistic evaluation of Sentinel 1 and 2 for biomass estimation in a tropical forest of India," *Adv. Space Res.* (2021).
14. R. K. M. Malhi et al., "An integrated spatiotemporal pattern analysis model to assess and predict the degradation of protected forest areas," *ISPRS Int. J. Geo-Inf.* **9**(9), 530 (2020).
15. J. R. Busby, "BIOCLIM-a bioclimate analysis and prediction system," *Plant Prot. Q.* **6**, 8–9 (1991).
16. G. Carpenter, A. Gillison, and J. Winter, "DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals," *Biodivers. Conserv.* **2**(6), 667–680 (1993).
17. J. Elith, J. R. Leathwick, and T. Hastie, "A working guide to boosted regression trees," *J. Anim. Ecol.* **77**(4), 802–813 (2008).
18. J. Peters et al., "Random forests as a tool for ecohydrological distribution modelling," *Ecol. Modell.* **207**(2-4), 304–318 (2007).
19. J. Franklin, *Mapping Species Distributions: Spatial Inference and Prediction*, Cambridge University Press (2010).
20. S. J. Phillips and M. Dudík, "Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation," *Ecography* **31**(2), 161–175 (2008).
21. J. Leathwick, J. Elith, and T. Hastie, "Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions," *Ecol. Modell.* **199**(2), 188–196 (2006).
22. M. Bobrowski et al., "Searching for ecology in species distribution models in the Himalayas," *Ecol. Modell.* **458**, 109693 (2021).
23. A. Guisan and C. Rahbek, "SESAM-A new framework for predicting spatio-temporal patterns of species assemblages: Integrating macroecological and species distribution models," *J. Biogeogr.* **38**, 1433–1444 (2011).
24. J. Elith and J. R. Leathwick, "Species distribution models: ecological explanation and prediction across space and time," *Annu. Rev. Ecol. Evol. Syst.* **40**, 677–697 (2009).
25. E. Saupe et al., "Variation in niche and distribution model performance: the need for a priori assessment of key causal factors," *Ecol. Modell.* **237**, 11–22 (2012).
26. L. R. D. A. Carneiro et al., "Limitations to the use of species-distribution models for environmental-impact assessments in the Amazon," *PLoS One* **11**(1), e0146543 (2016).
27. R. G. Pearson and T. P. J. G. E. Dawson, "Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful?" *Global Ecol. Biogeogr.* **12**(5), 361–371 (2003).
28. R. K. M. Malhi et al., "Synergetic use of in situ and hyperspectral data for mapping species diversity and above ground biomass in Shoolpaneshwar Wildlife Sanctuary, Gujarat," *Trop. Ecol.* **61**(1), 106–115 (2020).
29. P. Singh et al., "Delineation of ground water potential zone and site suitability of rainwater harvesting structures using remote sensing and *in-situ* geophysical measurements," in *Advances in Remote Sensing For Natural Resource Monitoring*, Vol. 1, P. C. Pandey and L. K. Sharma Eds., John Wiley & Sons Ltd. (2020).
30. P. C. Pandey, A. Anand, and P. K. Srivastava, "Spatial distribution of mangrove forest species and biomass assessment using field inventory and earth observation hyperspectral data," *Biodiver. Conserv.* **28**(8), 2143–2162 (2019).

31. R. J. Hijmans et al., "Very high resolution interpolated climate surfaces for global land areas," *Int. J. Climatol.: J. R. Meteorol. Soc.* **25**(15), 1965–1978 (2005).
32. E. Kalnay et al., "The NCEP/NCAR 40-year reanalysis project," *Bull. Am. Meteorol. Soc.* **77**(3), 437–472 (1996).
33. F. Molteni et al., "The ECMWF ensemble prediction system: methodology and validation," *Quart. J. R. Meteorol. Soc.* **122**(529), 73–119 (1996).
34. G. Negi et al., "Impact of climate change on the western Himalayan mountain ecosystems: an overview," *Trop. Ecol.* **53**(3), 345–356 (2012).
35. J. Salick, Z. Fang, and A. Byg, "Eastern Himalayan alpine plant ecology, Tibetan ethnobotany, and climate change," *Global Environ. Change* **19**(2), 147–155 (2009).
36. T. Yao et al., "δ 18 O record and temperature change over the past 100 years in ice cores on the Tibetan Plateau," *Sci. China Ser. D* **49**(1), 1–9 (2006).
37. IPCC, "Summary for policymakers," in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York (2014).
38. U. B. Shrestha and K. S. Bawa, "Impact of climate change on potential distribution of Chinese caterpillar fungus (Ophiocordyceps sinensis) in Nepal Himalaya," *PLoS One* **9**(9), e106405 (2014).
39. A. Fischer, M. Blaschke, and C. Bässler, "Altitudinal gradients in biodiversity research: the state of the art and future perspectives under climate change aspects," *Waldökol. Landschaft. Nat.* **11**, 35–47 (2011).
40. A. S. Jump, T. J. Huang, and C. H. Chou, "Rapid altitudinal migration of mountain plants in Taiwan and its implications for high altitude biodiversity," *Ecography* **35**(3), 204–210 (2012).
41. P. Kumar, "Assessment of impact of climate change on Rhododendrons in Sikkim Himalayas using Maxent modelling: limitations and challenges," *Biodivers. Conserv.* **21**(5), 1251–1266 (2012).
42. S. N. Veera et al., "Prediction of upslope movement of *Rhododendron arboreum* in Western Himalaya," *Trop. Ecol.* **60**, 518–524 (2020).
43. S. Ranjitkar et al., "Separation of the bioclimatic spaces of Himalayan tree rhododendron species predicted by ensemble suitability models," *Global Ecol. Conserv.* **1**, 2–12 (2014).
44. P. Rawat et al., "Review on *Rhododendron arboreum*-a magical tree," *Orient. Pharm. Exp. Med.* **17**(4), 297–308 (2017).
45. P. K. Sonar et al., "Isolation, characterization and activity of the flowers of *Rhododendron arboreum* (Ericaceae)," *E-J. Chem.* **9**(2), 631–636 (2012).
46. G. Secretariat, "GBIF backbone taxonomy," Checklist Dataset [cited 2017 Nov 14], Vol. 10 (2017).
47. J. Veefkind et al., "TROPOMI on the ESA Sentinel-5 precursor: a GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications," *Remote Sens. Environ.* **120**, 70–83 (2012).
48. J. B. Fisher et al., "ECOSTRESS: NASA's next generation mission to measure evapotranspiration from the International Space Station," *Water Resour. Res.* **56**(4), e2019WR026058 (2020).
49. J. J. Van Zyl, "The shuttle radar topography mission (SRTM): a breakthrough in remote sensing of topography," *Acta Astron.* **48**(5–12), 559–565 (2001).
50. T. Hengl et al., "Global mapping of potential natural vegetation: an assessment of machine learning algorithms for estimating land potential," *PeerJ.* **6**, e5457 (2018).
51. C. E. Shannon, "A mathematical theory of communication (concluded)," *Bell. Syst. Tech. J* **27**, 379–423 (1948).
52. E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.* **106**(4), 620 (1957).
53. E. Moreno-Amat et al., "Impact of model complexity on cross-temporal transferability in Maxent species distribution models: an assessment using paleobotanical data," *Ecol. Modell.* **312**, 308–317 (2015).
54. L. Breiman, "Random forests," *Mach. Learn.* **45**(1), 5–32 (2001).

55. J. Mascaro et al., "A tale of two "forests": random forest machine learning aids tropical forest carbon mapping," *PLoS One* **9**(1), e85993 (2014).
56. J. S. Evans et al., "Modeling species distribution and change using random forest," in *Predictive Species and Habitat Modeling in Landscape Ecology*, C. A. Drew, Y. F. Wiersma, and F. Huettmann, Eds., pp. 139–159, Springer, New York (2011).
57. A. Liaw and M. Wiener, "Classification and regression by random forest," *R. News* **2**(3), 18–22 (2002).
58. H. Hashimoto et al., "High-resolution mapping of daily climate variables by aggregating multiple spatial data sets with the random forest algorithm over the conterminous United States," *Int. J. Climatol.* **39**(6), 2964–2983 (2019).
59. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer Science & Business Media (2009).
60. G. De'ath and K. E. Fabricius, "Classification and regression trees: a powerful yet simple technique for ecological data analysis," *Ecology* **81**(11), 3178–3192 (2000).
61. R. E. Schapire, "The boosting approach to machine learning: an overview," in *Nonlinear Estimation and Classification*, D. D. Denison et al., Eds., pp. 149–171, Springer, New York (2003).
62. P. McCullagh and J. A. Nelder, *Generalized Linear Models*, Routledge (2019).
63. F. E. Harrell, Jr, *Regression Modeling Strategies: with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis*, Springer (2015).
64. A. Anand, S. K. Singh, and S. Kanga, "Estimating the change in forest cover density and predicting NDVI for West Singhbhum using linear regression," *Int. J. Environ. Rehabil. Conserv.* **9**, 193–203 (2018).
65. A. Davison, "Biometrika centenary: theory and general methodology," *Biometrika* **88**, 13–52 (2001).
66. H. Reiss et al., "Species distribution modelling of marine benthos: a North Sea case study," *Mar. Ecol. Prog. Ser.* **442**, 71–86 (2011).
67. A. Tsoar et al., "A comparative evaluation of presence-only methods for modelling species distribution," *Divers. Distrib.* **13**(4), 397–405 (2007).
68. J. Elith et al., "Novel methods improve prediction of species' distributions from occurrence data," *Ecography* **29**(2), 129–151 (2006).
69. R.-Y. Duan et al., "The predictive performance and stability of six species distribution models," *PloS One* **9**(11), e112764 (2014).
70. J. G. Giovanelli et al., "Modeling a spatially restricted distribution in the Neotropics: how the size of calibration area affects the performance of five presence-only methods," *Ecol. Modell.* **221**(2), 215–224 (2010).
71. S. J. Sinclair, M. D. White, and G. R. Newell, "How useful are species distribution models for managing biodiversity under future climates?" *Ecol. Soc.* **15**(1), 8 (2010).
72. M. Rostami, K. Berahmand, and S. Forouzandeh, "A novel method of constrained feature selection by the measurement of pairwise constraints uncertainty," *J. Big Data* **7**(1), 83 (2020).
73. A. Anand et al., "Integrating multi-sensors data for species distribution mapping using deep learning and envelope models," *Remote Sens.* **13**(16), 3284 (2021).

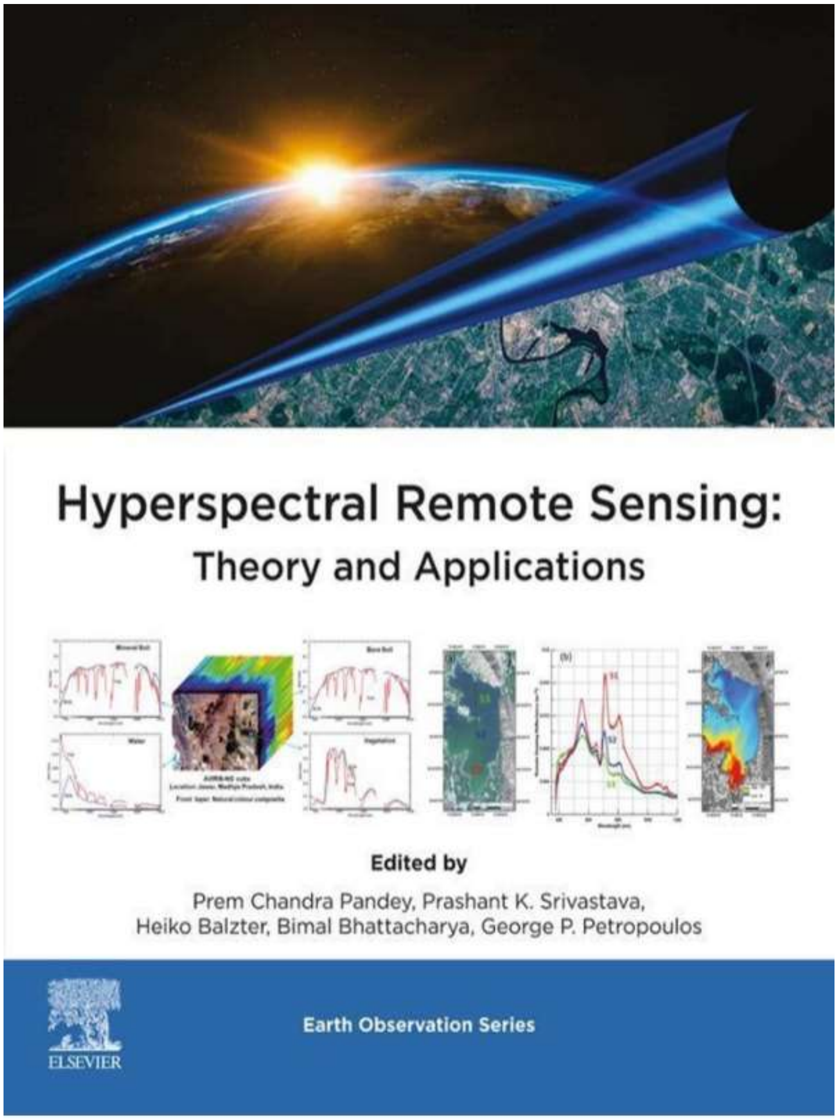
Biographies of the authors are not available.

Journal of Environmental Management

Seeing from space makes sense: Newly developed phenological and earth observation variables accurately map species distributions over Himalaya

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research Article
Keywords:	Alpine species distributions, Niche overlapping, Machine learning algorithms, Remote sensing, Mountains, Conservation and management
Corresponding Author:	Prashant Kumar Srivastava, PhD Banaras Hindu University INDIA
First Author:	K V Satish, PhD
Order of Authors:	K V Satish, PhD
	Vikas Dugesar
	Manish Pandey, PhD
	Prashant Kumar Srivastava, PhD
	Dalbeer Singh Pharswan, PhD
	Zishan Ahmad Wani
Abstract:	Topical advances in earth observation have enabled spatially explicit mapping of current species' fundamental niche limits that can be used for nature conservation and management applications. This study investigates the possibility of applying functional variables retrieved from MODIS onboard sensor data to map the most realistic species distribution limits of two alpine treeline species, namely Betula utilis and Rhododendron campanulatum over the Himalayan biodiversity hotspot. In this study, we have developed novel earth observation variables (EOVs) from MODIS products, an asset to the present investigation. To identify the effective and ecologically significant predictor variable combinations, we built and compared four different models, i.e., bioclimatic model (BCM), remote sensing model (RSM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, RSM, and HM were developed and tested for the first time in this study. For each model, we assessed the congruence of predictions and made pairwise comparisons to assess the performance of models. Models including EOVs were competitive for study species, and models without EOVs had considerably poor model performance and explanatory strength. Among three machine learning algorithms tested (artificial neural networks, generalised boosting model, and maximum entropy), maximum entropy produced the most promising predictions for BCM, PhenoM, RSM, and HM. Area under curve (AUC) and true skill statics (TSS) estimates for the highest ranked models were high for B. utilis (BCM = AUC 0.91, TSS 0.62; PhenoM = AUC 0.9, TSS 0.68; RSM = AUC 0.93, TSS 0.68; HM = AUC 0.94, TSS 0.72) and R. campanulatum (BCM = AUC 0.96, TSS 0.67; PhenoM = AUC 0.95, TSS 0.62; RSM = AUC 0.94, TSS 0.64, HM = AUC 0.95, TSS 0.62). The overall investigation revealed the competency of EOVs in the accurate prediction of species' fundamental niches. Additionally, although conventional BCM maps have been generated for two focal species, most fail to match the ground reality. A principal component analysis of environmental spaces disclosed that niches of focus species substantially overlapped each other. In contrast, we demonstrate that the use of satellite onboard sensors' biotic and abiotic variables with species occurrence data can provide precision and resolution at a scale that is relevant ecologically and at the operational scale of most conservation and management actions.


Hyperspectral Remote Sensing:
Theory and Applications
Edited by
Prem Chandra Pandey, Prashant K. Srivastava,
Heiko Balzter, Bimal Bhattacharya, George P. Petropoulos
Earth Observation Series
ELSEVIER

8

Hyperspectral remote sensing in precision agriculture: present status, challenges, and future trends

Prachi Singh[1], Prem Chandra Pandey[2], George P. Petropoulos[3], Andrew Pavlides[4], Prashant K. Srivastava[1,5], Nikos Koutsias[6], Khidir Abdala Kwal Deng[7], Yangson Bao[7]

[1]REMOTE SENSING LABORATORY, INSTITUTE OF ENVIRONMENT AND SUSTAINABLE DEVELOPMENT, BANARAS HINDU UNIVERSITY, VARANASI, INDIA
[2]CENTER FOR ENVIRONMENTAL SCIENCES & ENGINEERING, SCHOOL OF NATURAL SCIENCES, SHIV NADAR UNIVERSITY, GREATER NOIDA, INDIA
[3]DEPARTMENT OF GEOGRAPHY, HAROKOPIO UNIVERSITY OF ATHENS, ATHENS, GREECE
[4]SCHOOL OF MINERAL RESOURCES ENGINEERING, TECHNICAL UNIVERSITY OF CRETE, CHANIA, GREECE
[5]DST-MAHAMANA CENTRE FOR EXCELLENCE IN CLIMATE CHANGE RESEARCH, BANARAS HINDU UNIVERSITY, VARANASI, INDIA
[6]DEPARTMENT OF ENVIRONMENTAL ENGINEERING, UNIVERSITY OF PATRAS, AGRINIO, GREECE
[7]COLLABORATIVE INNOVATION CENTER ON FORECAST AND EVALUATION OF METEOROLOGICAL DISASTERS, NANJING, UNIVERSITY OF INFORMATION SCIENCE & TECHNOLOGY, NANJING, P.R. CHINA

8.1 Introduction

Precision agriculture (PA) is the science of improving crop yields and assisting management decisions using high technology sensor and analysis tools. PA is a new concept adopted throughout the world to increase production, reduce labor time, and ensure the effective management of fertilizers and irrigation processes. It uses a large amount of data and information to improve the use of agricultural resources, yields, and the quality of crops (Mulla, 2013). PA is an advanced innovation and optimized field level management strategy used in agriculture that aims to improve the productivity of resources on agriculture fields. Thus PA is a new advanced method in which farmers provide optimized inputs such as water and fertilizer to enhance productivity, quality, and yield (Gebbers and Adamchuk, 2010). It requires a huge

Hyperspectral Remote Sensing. DOI: https://doi.org/10.1016/B978-0-08-102894-0.00009-7

2

Spectroradiometry: Types, Data Collection, and Processing

Prem Chandra Pandey[1], Manish Kumar Pandey[2], Ayushi Gupta[2], Prachi Singh[2], and Prashant K. Srivastava[2]
[1]Center for Environmental Sciences & Engineering, School of Natural Sciences, Shiv Nadar University, Greater Noida, Uttar Pradesh, India
[2]Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, Uttar Pradesh, India

2.1 Introduction

Spectroradiometry is the evaluation and analysis of light energy at discrete wavelengths within the electromagnetic spectrum. It can be evaluated over the complete spectrum or within a specific wavelength band. The popularity of sensors due to advancement of technologies and emergence of the Internet of Things (IoT) founds its application in hyperspectral sensors in remote sensing for field spectroscopy, as well as in several fields like medical informatics (Pandey and Subbiah 2016a; Pandey and Subbiah 2017; Pandey and Subbiah 2018) and Internet of Vehicles (IoV) (Pandey and Subbiah 2016b), etc. Field spectroscopy is the procedure of measuring the reflectance properties of vegetation, soil, rock, or water bodies in the natural environment, preferably under solar illumination. Spectroscopy is the study of the interaction between electromagnetic radiation (EMR) and matter. In remote sensing, EMR is of the utmost importance, as it carries information about the composition of the object or the nature of the processes occurring within it. Spectroscopy is the branch of science involved in laboratory and field investigations of energy–matter interactions. The instrument used to measure the spectra are referred to as spectrometers, which is for quantitative measurements, and spectroscope, which is for an image. The spectroscopists tend to base their studies around absorption spectra, and they present absorption spectra with units of frequency. Srivastava et al. (2016) explored optical and microwave satellite for soil moisture deficit estimation. Srivastava, et al. (2020a) applied ANN-based sensitivity analysis for chlorophyll prediction using hyperspectral data. Srivastava et al. (2020b) revisited origin, importance and future scope of hyperspectral remote sensing. Recently, Pandey et al (2019a) provided an opportunity to consider again from a different perspective about remote sensing data, to know about dimensions, characteristics of remotely sensed data for its applications.

In remote sensing, the focus is on energy reflected or emitted from the object. In remote sensing, spectra are plotted against the wavelength. A hyperspectral signature of a certain object or surface cover (e.g. vegetation or rock type) consists of a collection of reflectance from a surface having a number of continuous spectral bands with narrow bandwidth. This collection is known as a hyperspectral database or spectral library. This collection can be utilized further in the retrieval and validation of biochemical and biophysical properties of the objects against the measured observations from ground instruments or airborne or spaceborne sensors. A number of applications are developed using spectral libraries ranging from vegetation classification, plant species classification and pigmentation, soil composition and moisture content, to crop infestation (Nasarudin and Shafri 2011; Zhang et al. 2011; Darvishsefat et al. 2011; Adam et al. 2010; Abdel-Rahman and Ahmed 2008; De Castro et al. 2012; Ray et al. 2010; Ferri et al. 2004; Cheng et al. 2007; Stagakis et al. 2012; Kurz et al. 2012, Pandey et al 2014, Pandey et al. 2019b). Another use of these libraries is in the calibration, validation, and simulation of remote sensing imagery obtained from visible and near-infrared sensors ranging from 0.4 to 3 nm. Due to the high dimensionality of the hyperspectral data, interpretation of this data is a tedious task. The instruments used to measure the hyperspectral signature of surface cover are known as spectroradiometers. The biggest advantage of the spectroradiometer is its non-destructive sampling of radiation from the surface, which enables the measurement of spectral signatures under any physical condition and configuration. Feather in the cap is its simpler and easier mode of operation and data collection, as mentioned by Hueni and Tuohy (2006). Calvin and Kruse (2010) utilized reflectance spectroscopy in exploring the mineral deposits and geothermal systems in the Great Basin of the United States.

The spectral library should offer easy access and navigation to the user, which is provided through the augmentation of the spectral data by metadata and storing it in an organized way as specified by Hueni et al. (2009). Examples of some hyperspectral databases are SPECCHIO, as given by Bojinski et al. (2003), and SpectraProc DB, as given by Hueni and Tuohy (2006). SPECCHIO holds spectra from several spectroradiometers while SpectraProc DB contains data from ASD FieldSpec Pro. As per Pfitzner et al. (2006), the sharing of spectral signatures is not advisable due to the difference in methods of collection as well as sampling environment conditions.

Appendix 2 List of Trainings/ Workshops/ Seminars with details of trained resources and dissemination material and Proceedings

Table 13 List of Trainings/ Workshops/ Seminars

S. No.	Type of Activities	Title of the activity	Held on	Activity Intended for	No. of Participants/Trained
1	Workshop	2nd National Workshop on Techniques in Hyperspectral Data Analysis and Processing	27th January - 31st January 2020	Training to P.G and PhD candidates & Govt./autonomous bodies/PSU/ Industry and NGOs	42
2	Workshop (Online)	Online Workshop on Species Distribution Modelling Using R	15th -16th March 2021	Training to P.G and, Researchers/ PhD candidates	28
3	Workshop (Online)	One-week Online Workshop on QGIS, Image Processing and Species Distribution Modelling	30th August – 04th September 2021	Training to P.G and, Researchers/ PhD candidates	33
4	Lecture	A Six Lecture	22nd-24th	Training and	15

	Course	Course on Data Assimilation	March 2022	knowledge transfer to Researchers/ PhD candidates	
5	Seminar	Graduate Seminar on Natural Resource Management	17th -18th March 2021	Training and knowledge transfer to Researchers/ PhD candidates	63

Table 14 Participants Details

2nd National Workshop on Techniques in Hyperspectral data Analysis and Processing (27th -31st January 2020)				
Sr. No.	Name	Sex	Address	Mobile No
1	Prachi Chachondhia	F	Anand Agriculture University, Anand, Gujarat	9452749622
2	Bhagyashree Choudhary	F	National Environment Engineering Research Institute, Nagpur, Maharashtra	9096416034
3	Amit Kumar Tiwari	M	IESD, BHU	9807831618
4	Neeraj Kumar	M	Sam Higginbottom university of agriculture technology & sciences, Prayagraj, Uttar Pradesh	9452440052
5	Subhadip Datta	M	Vidyasagar University, Midnapore, West Bengal	8348703617

6	Debarati Bera	F	Vidyasagar University, Midnapore, West Bengal	9838602286
7	Tanushree Basu Roy	F	Vidyasagar University, Midnapore, West Bengal	9046738044
8	Atul Srivastava	M	University of Allahabad	7081259162
9	Gaurav Mishra	M	NIT Raipur	9452338539
10	Abhishek Adhikari	M	IIT Rorkee	9411168123
11	Chandni Pradhan	F	Institute of Agricultural Sciences, BHU	8895566546
12	Shubham Srivastava	M	IESD, BHU	8267920746
13	Srishti Kumari	F	IESD, BHU	9102021143
14	Pushkar Goyal	M	Sam Higginbottom university of agriculture technology & sciences, Prayagraj, Uttar Pradesh	8299739744
15	Neeraj Kumar Maurya	M	IESD, BHU	8387882663
16	Kapil Bisht	M	G B Pant National Institute of Himalayan Environment and Sustainable Development	9627694404
17	Renu Rawal	F	G B Pant National Institute of Himalayan Environment and Sustainable Development	7579461254
18	Koppineedi Veera Satish	M	G B Pant National Institute of Himalayan	9491257581

			Environment and Sustainable Development	
19	Bodi Surya Pratap Chandra Kishore	M	Central university of Jharkhand	9491257581
20	Neha Thapiyal	F	G B Pant National Institute of Himalayan Environment and Sustainable Development	9568210078
21	Shahbaz Ahmad	M	Central university of Jharkhand	7870737534
22	Sunil Joshi	M	G B Pant National Institute of Himalayan Environment and Sustainable Development	9675442890
23	Manoj Kumar Patley	M	G B Pant National Institute of Himalayan Environment and Sustainable Development	9491398372
24	Juby Thomas	M	DRDO, DTRL	8148490784
25	Dwijendra Pandey	M	MCG, DTU Delhi	8574020036
26	Puja Bhojak	F	G B Pant National Institute of Himalayan Environment and Sustainable Development	9557966146
27	Adluri Anjaneyulu	M	Mangrove Foundation	9966522335
28	Vinay kumar singh	M	IIT(BHU)	9936239355
29	Seema	F	IAS, BHU	9454185104
30	Rimjhim Rao	F	DTU, Delhi	7597850745

31	Urvashi Nag	F	CSIO, Chandigarh	7696066623
32	Babban S. Bansood	M	CSIO, Chandigarh	9888520252
33	Suparn Pathak	M	NRSC	0291-2796381
34	Anand Singh Dinesh	M	IESD, BHU	9450557770
35	Surbhi Barnwal	F	IIT, Khragpur	9304005883
36	Maneeb Masood	M	IIT, BHU	9565148584
37	Himanshu Mishra	M	VIAET	9956140381
38	Billy Graham Ram	M	VIAET	9161802082
39	Abhishek Ranjan	M	VIAET	9581692408
40	Sahil	M	CSIR, Chandigarh	9646326804
41	Priyanshi Tiwari	F	BHU	7753064856
42	Jaya Prajapati	F	BHU	9977671884

2nd National Workshop on Techniques in Hyperspectral Data Analysis and Processing

CHIEF PATRON
Prof. Rakesh Bhatnagar
Vice Chancellor - BHU

PATRONS
Prof. A S Raghubanshi
Director- IESD, BHU, Varanasi

Prof. P. K. Jain
Director- IIT BHU

CONVENER
Prof. R K Mall
IESD, BHU, Varanasi

ORGANIZING SECRETARY
Dr. Prashant K. Srivastava
IESD, BHU, Varanasi

Prof. Rajendra Prasad
Dept. of Physics,
IIT-BHU, Varanasi

TREASURER
Dr. Vishal Prasad
IESD, BHU , Varanasi

ORGANIZING COMMITTEE
Prof. G. S. Singh
Prof. Kavita Shah
Dr. J P Verma
Dr. Sunita Verma
Dr. V. Mishra
Dr. Rajiv Pratap Singh
Dr. P C Abhilash
Dr. T Banerjee
Dr. Kirpa Ram
Dr. Sudhakar Srivastava

"THE GREATEST THREAT TO OUR PLANET IS THE BELIEF THAT SOMEONE ELSE WILL SAVE IT!"
ROBERT SWAN

Registration Form

Name: Dr./Mr./Ms./Mrs.

Designation:

Date of Birth:

College/Organization Name:

Address for correspondence:

Pin Code:

Mobile Number:

Email:

Payment Details:

DD No.:

Bank Name:

Amount :

Date:

Weather Accommodation Required

YES NO

Area of Interest:

Why do you want to attend this workshop

Signature: Date:

*Travel and accommodation charges will have to be borne by participants.
Contact email: workshopiesd@gmail.com
Contact Number: +91- 7571927744

2nd National Workshop on
Techniques in Hyperspectral Data Analysis and Processing
27th January – 31st January 2020

Jointly Organized By

Institute of Environment & Sustainable Development, Banaras Hindu University, Varanasi

Indian Institute of Technology, Banaras Hindu University, Varanasi

Figure 43 Brochure HYPERTECH

ABOUT INSTITUTE

The Institute of Environment & Sustainable Development is dedicated to a better understanding of critical scientific and social issues related to sustainable development goals through guided research. The institute is making its mark at the national and international platforms in a number of frontier areas of Environment Science like hydrometeorology & agrometeorology, climate research, Remote sensing and GIS, modelling air pollution, water resources etc.

The Department of Physics (IIT BHU) established in 1985, is a centre of repute for quality research and training in Physics. The Department offers an excellent research programme in the field of Solar Physics, Space & Planetary Physics, Astrophysics, Fibre Optics, Photonics, Condensed Matter Physics, Biophysics, Materials sciences, Remote Sensing, Quantum Information, and Renewable Energy.

ABOUT WORKSHOP

From past decade hyper-spectral imaging has emerged as a powerful technology in remote sensing and a widely used in research areas. It provides the basic data to undertake inventory of land as well as other resources though contiguous bands.

Recent innovations in Hyper-spectral Imaging offer promising techniques to achieve sustainability. The main aim of this workshop is to unveil the advances in remote sensing through lectures, demonstrations and hands-on training that helps the participants gain an understanding to deal with spatial data using remote sensing tools for various applications. This workshop will also cover the demonstrations and hands-on training on software like R, ENVI, and PRISM for image processing.

WORKSHOP TOPICS

- Introduction to Remote Sensing and Hyper-Spectral Imaging
- Basics of R and Hyperspectral Image Processing
- Introduction to ENVI and Hyperspectral Image Processing for Land use/Land cover and vegetation monitoring
- Hyper-Spectral Radiometer Calibration, data collection and spectra processing
- Introduction to PRISM and hands-on in mineral identification using Hyper-spectral data

CATEGORY OF DELEGATES	Amount (INR)
PG Students	₹ 2000
Research /Ph.D. Scholar	₹ 3500
Govt./Autonomous Bodies /PSU /Industry/NGO	₹ 6000

The payment can be made through cash/electronic transfer/DD/Cheque. DD/Cheque and completely filled registration form along with CV should be sent to **Dr. P. K. Srivastava**, Organizing Secretary, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi at the email id: **workshopiesd@gmail.com**. Bank details for the electronic transfer are given below:

Bank Details
Account Name: Organizing Secretary Hypertech workshop IESD
Account No.: 27790200000612
IFSC Code: BARB0BHUVAR
Bank Name: Bank of Baroda (BHU, Varanasi)

IMPORTANT DATES

Registration deadline	5th January, 2020
Confirmation of Participation	10th January, 2020

*Registration seats are limited (30 Seats, On First Come First Serve Basis)

ACCOMMODATION

Accommodation charges has to be borne by the participants themselves. Guest house in BHU campus will be booked on request subject to availability and on payment basis.

ABOUT VARANASI

Varanasi, the spiritual capital of India, is believed to be the most ancient surviving city of the globe. Legend has it that the first Siva Jyotirlinga emanated from the earth in this city and flared straight into the sky. This is why **Varanasi** is also known as Kashi or **"The City of Light"**. The boat rides on the Ganga, visiting the different temples and exploring the vibrant city is made double fun for visitors.

For any enquiry or further information please contact
Organizing Secretary
2nd National Workshop on Techniques in Hyperspectral Data Analysis and Processing
Banaras Hindu University & IIT (BHU),
Varanasi-221005.
Mobile No: +91-7571927744
Email id: : workshopiesd@gmail.com

Figure 44 Brochure HYPERTECH

Figure 45 Snaps of the workshop

Summary of the workshop

Institute of Environment and Sustainable Development, Banaras Hindu University in collaboration with the Department of Physics, IIT-BHU has organized a 2^{nd} National workshop on "Techniques in Hyperspectral Data Analysis and Processing". The event started with the garlanding of the bust of Mahamana Madan Mohan Malviya Ji followed by lamp lightening and Kulgeet recitation. The workshop was inaugurated in Senate Hall, Swatantrata Bhawan by the Chief Guest of the program, Prof. V. K. Gaur, Honorary Emeritus Scientist, CSIR-4PI, Bangalore. Prof. Gaur has emphasized the importance of Hyperspectral Technologies in the medical diagnostics, precision agriculture as well as in keeping the cultural heritage. The Guest of Honour Prof. P. K. Nag, Ex-DG, General Surveyor of India and Ex-VC, MGKVP has provided an overview of the benefits on hyperspectral data over multispectral data.

The workshop has received participation from renowned universities and Institutes across India. The participants were Scientists from CSIR-CSIO and ISRO along with GIS specialist from Mangrove Foundation and PhD, MTech students from IITs, NITs, DRDOs, CSIR, Central/State Universities and other institutes of national importance. The workshop started with a lecture on "**Basics of Remote Sensing" by Prof. K N P Raju, Geography, BHU** who has delivered a glimpse of the much-required fundamentals of remote sensing. The next session was delivered by our eminent alumni, **Dr. Lokesh Kumar Sinha, Director, DTRL-DRDO.** on **"Hyperspectral Remote Sensing".** In his lecture, he has highlighted the importance of hyperspectral remote sensing in the fields ranging from Atmosphere, Defense, Ecology, and Geology. "**The Potential and Scope of Hyperspectral Remote Sensing Applications with Special Reference to Agriculture"** was delivered by **Dr. Bimal K Bhattacharya, Sr Scientist, SAC ISRO.** He highlighted the importance of hyperspectral remote sensing in the area of agriculture and other fields as well as suggested the students on the available data sets in the area of hyperspectral remote sensing**.** The lecture was followed by **a lecture on "an Introduction to Software for Hyperspectral Data Analysis and Processing".** This session was taken by Dr. P K. Srivastava, IESD, BHU.

The second day was started by Prof M. L. Khan, an eminent name in the Forest Ecology who delivered his talk on "**Forest Across Climates and Biomes: An Indian Perspectives**. He has given the participants, a perspective on the Impact of climate and biomass on forestry. The next lecture was on **"Tools and Techniques used in Ecological Analysis of Forest"** which was taken by Dr. Purabi Saikia, CUJ, Ranchi. She has enriched the knowledge of participants on the tools used as well as the methodologies required in the analysis of the forest. The third lecture was taken by **Prof. Mukund Behera, IIT Kharagpur** on "**Multispectral Remote Sensing in Forest Monitoring**". He has given an insight into the importance of multispectral remote sensing in forest mapping and monitoring. The next lecture was taken by Dr. Amit Kumar, CUJ, Ranchi on **"Hyperspectral Remote Sensing for Forest Monitoring".** The use of hyperspectral remote sensing in forest mapping and monitoring was discussed by him in brief. The **"Hands-on Training in Hyperspectral Data Analysis for Vegetation Properties"** was taken by Dr. P. C. Pandey, SNU and Dr. Ramandeep K Malhi, IESD, BHU. Both of them have given the participants a healthy dose of hyperspectral data processing and analysis using several tools.

The third day started with a discussion on "**Disruptive Innovations in Precision Agriculture" by Prof. J Adinarayana, Head, CSRE-IITB.** He has given an insight into the Internet of Things and Artificial Intelligence in Precision Agriculture. A lecture on **"Hyperspectral Remote Sensing Application in Agriculture"** was taken by **Prof. R N Sahoo, Principal Scientist, IARI, New Delhi.).** He offered a detailed insight into the application of hyperspectral remote sensing in the diverse fields of agriculture from many prospective such as soil contents in the prediction of soil health etc. A "**Hands-on Training on Hyperspectral spectroradiometer" was taken by (Mr A. Purshottaman, Manager, Electrotek, Chennai.** He has trained the participants in capturing the dark current, white reference as well as reflectance using the spectroradiometer. He has also trained the participants in data analysis and processing of reflectance data. Several plant species, as well as soil samples, were taken for capturing the reflectance data. The last lecture of the third day was taken by **Prof. Chalapathi Rao, Geology, BHU** on

“**Introduction to Mineralogy”.** He has given the participants an exposure to x-ray crystallography in the identification as well as characterization of various minerals.

The fourth day started with the demonstration of an Imaging spectrometer by **Mr A. Purshottaman.** The most awaited lecture of the workshop was started in the afternoon session **by Prof. Ramkrishnan Desikan, Earth Sciences, IIT Bombay.** His lecture topic was **“Imaging Spectroscopy of Minerals”.** He has started by linking all the prerequisites required in hyperspectral remote sensing and beaded all the associated concepts beautifully. The day ended with the resolution of the queries. He has also laid the foundation of Imaging Spectrometers to the participants and briefed them about its usage in the medical field apart from the usual field of Geology.

The final day started with a “**Hands-on Training on PRISM for Mineral Identification**” by Dr. P. K. Srivastava**.** The training program has provided the participants with a platform where they have developed their skills using PRISM for mineral identification using some well-known datasets. The last lecture of the workshop was taken by **Dr. Manika Gupta, University of Delhi** on **“Introduction to R language and its use in Image Processing”.** Detailed, hands-on experience in hyperspectral data analysis was provided to the participants in this session using some well-known satellite images.

Online Workshop on Species Distribution Modelling Using R

Table 15 Participants Details

Online Workshop on Species Distribution Modelling Using R 15th – 16th March, 2021					
Name		Designation	Institute	Mail	Contact
Mohammad Suhail Meer	M	PhD Scholar	Sathyabama Institute of science and technology	mohdsuhailrs@gmail.com	9906749497/ 8828086552
Ruquia Gulzar	F	PhD Scholar	University of Kashmir	mirruquia1@gmail.com	6005500528
Satya Narayan Mishra	M	M.Sc.	IESD, BHU	mishrasatyanarayan15@gmail.com	9783331004
Rakhi Mahto	F	PhD Scholar	IAS, BHU	rakhianubau@gmail.com	9801728973
HARSHITA RANI AHIRWAR	F	JRF	IESD, BHU	aharshita66.mceccr@gmail.com	7047060529
Vibha Verma	F	M.Sc.	IESD, BHU	vibhaverma1997@gmail.com	7379719870
Nidhi Sharma	F	M.Sc.	IESD, BHU	nadda.nidhi@gmail.com	8278787782
Garima Shukla	F	PhD Scholar	NIO	garimashukla.nio@gmail.com	7237833963
Dr. Anand S Sahadevan	M	Scientist	Space Applications Centre, ISRO	anandss.isro@gmail.com	7907305993
Srishty Gwal	F	PhD Scholar	IIRS,	srishtigwal01@gmail.com	7376642533

Sheikh Aneaus	F	PhD Scholar	Kashmir University	sheikh.aneaus19@gmail.com	7006078213
Smeeksha	F	M.Sc.	IESD, BHU	sameekshaone1@gmail.com	7667989837
Snigdha Dey	F	M.Sc.	IESD, BHU	snigdhadey53@gmail.com	8876954505
Shivangi Singh	F	M.Sc.	IESD, BHU	shivangi2102.as@gmail.com	9044125551
Akash Pradhan	M	M.Sc.	IESD, BHU	padhanakash@gmail.com	8789453972
Unnati Raghuwanshi	F	M.Sc.	IESD, BHU	uraghuwanshi31@gmail.com	9399686931
Shashank Tripathi	M	M.Sc.	IESD, BHU	shashankt47@gmail.com	8299750689
Apurwa Raj	F	M.Sc.	SGVU	apurwarajsingh1995@gmail.com	7357613446
Charu Pundir	F	JRF	GBPIHED	charupundir0514@gmail.com	6397924757
Gurulingappa	F	JRF	CSIR - 4PI	gurucsir7@gmail.com	9535555831
Gajendra Kumar	M	Research Scholar	Central university of Jharkhand	gajendrachaurasiya@gmail.com	9936903761
SHIVAM AGRAHARI	M	M.Sc.	IESD, BHU	shivambhu011995@gmail.com	9511471722
SATYAJIT NAYAK	M	M.Sc.	IESD, BHU	nayaksatyajit198@gmail.com	7008517783
ANIKET PRAKASH	M	Research Scholar	Central university of Jharkhand	aniketprakash95@gmail.com	7004880492
Amit Kumar Tiwari	M	Research Scholar	IESD, BHU	amit.iesd@gmail.com	9807831618
Srishti Srivastava	F	M.Sc.	IESD, BHU	srishtisrivastava510@gmail.com	9102021143
Aditi Shukla	F	M.Sc.	IAS, BHU	shuklaaditi089@gmail.com	8770959552

Shubham Srivastava	M	M.Sc.	IESD, BHU	shubham.imdbst@gmail.com	8267920746

Figure 46 Brochure SDM

Figure 47 Brochure SDM

Figure 48 Snaps of the workshop

Summary of the Workshop

Institute of Environment and Sustainable Development, Banaras Hindu University has organized a Online Workshop on Species Distribution Modelling Using R. The primary focus of the workshop is in Modelling species' environmental requirements and mapping their distributions through space and time that constitute important aspects of many biological analyses, particularly in support of conservation and management interventions. Species distribution models (SDM) use known locations of a species and information on environmental conditions to predict the species distributions. SDM is also known under other names including climate envelope-modelling, habitat modelling, and (environmental or ecological) niche-modelling. SDM aims to estimate the similarity of the conditions at any site to the conditions at the locations of known occurrence (and perhaps of non-occurrence) of a phenomenon. A common application of this method is to predict species ranges with climate data as predictors. SDM use a variety of algorithms to estimate relationships between species locations and environmental conditions, predict and map habitat suitability. SDMs represent a set of popular techniques for interpolating and extrapolating species distributions based on quantitative or rule-based models. This workshop will also cover the demonstrations and hands-on training on software R. The following steps would be covered (1) locations of occurrence of a species (or other phenomena) would be compiled; (2) values of environmental predictor variables (such as climate) at these locations would be extracted from spatial databases; (3) the environmental values would be used to fit a model to estimate similarity to the sites of occurrence, or another measure such as the abundance of the species; (4) The model would be used to predict the variable of interest across the region of interest (and perhaps for a future or past climate). The chief Guest, Prof. Vinod Kumar Gaur Sir, Honorary Scientist, CSIR Fourth Paradigm Institute, Bangalore, Guest of Honor Prof. P. S Roy Sir, Senior Fellow, World Resources Institute India and Visiting Fellow CEAOS, University of Hyderabad, Prof M. L. Khan Sir, a distinguished expert in Forest Ecology, Biodiversity Conservation, Eco-restoration and Natural Resource Management from Dr. Harisingh Gour Central University, Sagar, MP, Er. Kireet Kumar Sir, Coordinator NMHS and Scientist G, G.B. Pant National Institute of Himalayan Environment, Kosi-Katarmal, Almora, Uttarakhand, India, Dr. K Chandra Sekar Sir, Scientist E and Dr. Irfan Rashid Sir, Assistant Professor, University of Kashmir were present during the workshop.

Table 16 Participants Details

One-week Online Workshop on QGIS, Image Processing and Species Distribution Modelling 30th August – 04th September 2021			
S.No.	Name	Mail Id	Phone No
1	Nakhshab Moiz	nakhshabm2@gmail.com	8081324298
2	PRIYANKA SINGH	ps.priyankasingh26996@gmail.com	7976213586
3	Akanksha Pandey.	akanksha8299@gmail.com	7599033511, 6383797231
4	Amit Maurya	amitkr.maurya22@gmail.com	8127958954 8303836059
5	Rainy Sonkar	rainysvns@gmail.com	9140717625
6	Sameksha Johari	sameeksha.johri@gmail.com	8090755971
7	Aritra Roy	royaritra215@gmail.com	
8	Abhay Kumar Pandey	Aykupy777@gmail.com	9456332325, 9458084296
9	Sonam Mohinee	sonammohinee01@gmail.com	8933836610
10	Srishti Kumari	srishtisrivastava510@gmail.com	9102021143 7488133602
11	Shubham Srivastava	shubham.imdbst@gmail.com	8267920746
12	Nidhi Sharma	nadda.nidhi@gmail.com	8278787782
13	Jadev Chauhan	jaidevchauhan678@gmail.com	8126211560
14	Kavita Khatri	kabitakhatri000@gmail.com	8954465311
15	Bhawna Negi	negibhawna20@gmail.com	8650818748
16	Gokul Gopal Iyer	iyergokulgopal@gmail.com	9140447870
17	Archana Fartyal	archanafartyal2008@gmail.com	9690242981

18	Rajendra S Rawat	rajendrarawat1994@gmail.com	9720975462
19	Sarika Sarika	sarika05091994@gmail.com	8934955211
20	Ankita Tiwari	ankitatiwari87392@gmail.com	8795589554
21	Anveshika Singh	anveythakur1999@gmail.com	8577861202
22	Ismita Singh	ismitasingh280@gmail.com	9336244697
23	Swati Suman	swatigarg1708@gmail.com	950363128
24	Varsha Pandey	varshu.pandey07@gmail.com	7839265717
25	K.V. Satish	dr.kvsatish@yahoo.com	9491257581
26	Prachi Singh	prachisngh246@gmail.com	8318991665
27	Vikas Dugesar	dugesarvikas@yahoo.com	804488671
28	Rishabh Singh	er.rishabh.singh@outlook.com	9935203687
29	Dileep K. Gupta	dileepgupta85@gmail.com	8887737516
30	Ayushi Gupta	ayushiguptaiilm@gmail.com	8795334643
31	Shashank Srivastava	shashanksrivastava123.vns@gmail.com	6394434666
32	Manish K Pandey	pandey.manish@live.com	7388140147
31	Rahul Jaiswal	rj642034@gmail.com	8802997328
32	Megha Paul	meghapaul340@gmail.com	
33	Prashant Narayan	pnarayan@allduniv.ac.in	09450309247

Online workshop on QGIS, Image Processing and Species Distribution Modelling and Inauguration of SHRIME-17th-20th August 2021.

Figure 49 Brochure

Figure 50 Brochure

Figure 51 Snaps of the workshop and Inauguration of SHRIME

Summary of the workshop and Inauguration of SHRIME

The Spectral tool of Himalayan Rare, Invasive, Medicinal and Economical (SHRIME) Plant Species was Inaugurated by the Chief Guest, Prof. V. K. Gaur, Honorary Scientist, CSIR Fourth Paradigm Institute, Bangalore on 1st September 2021 at Institute of Environment and Sustainable Development. This tool is developed under the project funded by the National Mission on Himalayan Studies, GBPNIHE, Almora for monitoring the high-value plants in the Himalayan region, their sustainable utilization, and the conservation of related resources. The ultimate goal of this tool is to provide India's first hyperspectral library providing various aspects of Himalayan species in terms of their location, importance, conservation status and other remote sensing parameters including vegetation indices, thus, this tool offers a dynamic platform for basic analysis of biophysical and hyperspectral aspects of Himalayan plants. The primary focus of the workshop is capacity building in Modelling species' environmental constraints and mapping their distributions through space and time that constitute important insights for setting conservation and management priorities from species to landscape level. Organizing Secretary Dr Prashant K. Srivastava has introduced the SHRIME tool and discussed the importance of Hyperspectral Remote Sensing.

Prof. V.K. Gaur has provided an overview of the Ecosystem's Design and the need for hyperspectral vision. The role of atmospheric flux, energy dissipation, dynamics of ice sheets, the future pathways of climate change was emphasized in his lecture. He has motivated young researchers to take upcoming challenges due to climate change.

The Guest of Honor, Er. Kireet Kumar Sir has discussed the National Mission on Himalayan Studies and focused on the importance of the Restoration of the Himalayan Ecosystem. He has also shared the motivation behind the conservation of Himalayan species and associated biodiversity.

Media Coverage

अपना प्रदेश जागरण

विलुप्त हो रहे दुर्लभ औषधीय पौधों पर नजर रखेगी डिजिटल लाइब्रेरी

जागरण विशेष

सेटेलाइट बताएगा कहां, किस हाल में है पौधे

वाराणसी अंचल/पूर्वांचल जागरण

हिमालयी पौधों की खोज में मददगार होगी डिजिटल लाइब्रेरी

बीएचयू में देश की पहली हाइपर स्पेक्ट्रल डिजिटल लाइब्रेरी की लांचिंग आज

जागरण संवाददाता, वाराणसी : देश की पहली हाइपर स्पेक्ट्रल डिजिटल लाइब्रेरी की लांचिंग बुधवार को बीएचयू में होगी। इसके माध्यम से उन पौधों की पहचान एवं वास्तविक स्थान की खोज आसान होगा जो औषधीय गुणों से भरपूर हैं और अक्सर ही हिमालयी क्षेत्र में पाए जाते हैं। इस लाइब्रेरी को तैयार करने वाले पर्यावरण एवं धारणीय विकास संस्थान बीएचयू के डा. प्रशांत कुमार श्रीवास्तव ने बताया कि एपयुक्त इस लाइब्रेरी को श्राइम नाम दिया गया है। बताया कि लांचिंग कार्यक्रम का उद्घाटन सीएसआइआर फोर्थ पैराडिग्म इंस्टीट्यूट बेंगलुरु के विज्ञानी प्रो. विनोद कुमार गौड़ व जीबी पंत नेशनल इंस्टीट्यूट आफ हिमालयन इन्वायरमेंट के निदेशक व विज्ञानी किरीत कुमार करेंगे।

Figure 52 Media Coverage

A Six Lecture Series on Data Assimilation-22nd-24th March 2022.

Figure 53 Snaps of the workshop

CHIEF PATRON
Prof. Sudhir K Jain
Vice Chancellor - BHU

PATRON
Prof. A S Raghubanshi
Director - IESD, BHU, Varanasi

Convener
Prof. R. K. Mall

ORGANIZING SECRETARY Convener
Dr. Prashant K. Srivastava
IESD, BHU, Varanasi

ORGANIZING COMMITTEE
Prof. G. S. Singh
Prof. Kavita Shah
Dr. J P Verma
Dr. Sunita Verma
Dr. V. Mishra
Dr. Rajiv Pratap Singh
Dr. P C Abhilash
Dr. T Banerjee
Dr. Kirpa Ram
Dr. Sudhakar Srivastava
Dr. Vishal Prasad

"THE GREATEST THREAT TO OUR PLANET IS THE BELIEF THAT SOMEONE ELSE WILL SAVE IT"

ROBERT SWAN

Registration Form

Name: Dr./Mr./Ms./Mrs. ____

Designation: ____

Date of Birth: ____

College/Organization Name: ____

Address for correspondence: ____

Pin Code: ____

Mobile Number: ____

Email: ____

Payment Details: ____

DD No.: ____

Bank Name: ____

Amount: ____

Date: ____

Weather Accommodation Required:

YES NO

Area of Interest: ____

Why do you want to attend this workshop ____

Signature: ____ Date: ____

Contact email:

Contact Number:

A Six Lecture Course on

Data Assimilation

22-24 March, 2022

Organized By

DST - Mahamana Centre of Excellence in Climate Change Research (MCECCR), Institute of Environment and Sustainable Development, BHU, Varanasi

Figure 54 Brochure

A Six lecture series on Data assimilation was conducted by Prof. V. K. Gaur, Honorary Scientist, CSIR Fourth Paradigm Institute, Bangalore between 22nd -24th March 2022 at Institute of Environment and Sustainable Development. Data Assimilation is a powerful technique which has been widely applied in investigations of the atmosphere, ocean, and land surface. It combines observation data and the underlying dynamical principles governing the system to provide an estimate of the state of the system which is better than could be obtained using just the data or the model alone. The purpose of data assimilation is to determine a best possible atmospheric state using observations and short-range forecasts. Data assimilation is typically a sequential time-stepping procedure, in which a previous model forecast is compared with newly received observations, the model state is then updated to reflect the observations, a new forecast is initiated, and so on. The series was distributed in following lectures.

Table 17 Lecture Details

Lecture 1. What is Data Assimilation	**Lecture 2.** Mathematical Probability Distribution Functions and their numerical description (Rapid Review)
Lecture 3. Linear Systems and Vector Spaces	**Lecture 4.** Optimal Interpolation a) Scalar quantities b) Vector quantities and Kalman Gain c) Bayesian Inversion d) Homework and Hands on work
Lecture 5. Variational Methods of Data Assimilation a) The Variational Method – Minimization of a Cost Function b) Wherefore the Cost Function c) The 3-D and 4-D VAR methods	**Lecture 6.** Data Assimilation Exemples

d) Basics of the Adjoint Method	

Table 18 Participants Details

Graduate Seminar on Natural Resource Management" 17th – 18th March 2021			
S.No.	**NAME**	**EMAIL ID**	**PHONE NUMBER**
1	Kajal Singh	kajalsingh200997@gmail.com	9650119344
2	Akanksha Joshi	akanksha19066@gmail.com	8853845810 7651886774
3	Akanksha Yadav	yadavakanksha2008@gmail.com	9452692195
4	Ankit Devoanshi	ankit054200@gmail.com	8604240918
5	Jayant Karwadiya	jayantkarwadiya@gmail.com	8319650422
6	Ankita Gupta	ankita051296@gmail.com	7982455028 9455541955
7	Snigdha Dey	snigdhadey53@gmail.com	
8	Sirgha Parasher	dparashergeg@gmail.com	7839275979
9	Vibha Verma	vibhaverma1997@gmail.com	7985237219
10	Astha Singh	astha.singh12198@gmail.com	9450650198
11	Hricha Mishra	richarkt918@gmail.com	95407 16162
12	Sameeksha	sameekshaone1@gmail.com	

13	Ayush Lepcha	royalshwansh.sl@gmail.com	9990157018
14	Sudeepa Kumari	sudeepakumaripathak@yahoo.com	
15	Shubhra Pandey	shubhrashubh3@gmail.com	
16	Sankhadip Chakrabarty	shankhadip2711@gmail.com	7980866942
17	Shivam Agrahari	shivambhu011995@gmail.com	9511471722
18	Chandan das	chandandas9873@gmail.com	9773808278
19	Kanika Aggarwal	kanika14.aggarwal@gmail.com	
20	Unnati Raghuwanshi	uraghuwanshi31@gmail.com	9399686931
21	Gokul Gopal Iyer	iyergokulgopal@gmail.com	9140447870
22	Akash pradhan	pradhanakash1234@gmail.com	8789453972
23	Satyajit Nayak	nayaksatyajit198@gmail.com	7008517783
24	Vartika Jain	jainvartika.vj@gmail.com	9336066900
25	Praveen kumar	praveenkumarbhu1995@gmail.com	7275661591
26	Swati Kumari	swatikumari030198@gmail.com	9026295696
27	Nidhi Tiwari	tiwarinidhi246810@gmail.com	8318835936
28	Navneet Shukla		
29	Gitanjalee Kumari	kumarigita1997@gmail.com	

30	Sreejita Dutta	duttasreejita2018@gmail.com	8599909485
31	Tripti Panday	Triptipanday2104@gmail.com	9559494749
32	Priyanka singh	priyankasbhu@gmail.com	
33	Shraddha Pandey	shraddhapandey32@gmail.com	
34	Akanksha Gupta	akanshagupta26662@gmail.com	
35	Harshita Singh	harshitasingh4255@gmail.com	
36	Shashank Tripathi	shashankt47@gmail.com	
37	Prem Nisha	nishaprem666@gmail.com	
38	Komal Kumari	komalkumari382@gmail.com	
39	Srishti	srishtisrivastava510@gmail.com	
40	Shubham Srivastava	shubham.imdbst@gmail.com	
41	Pragati Yadav	shaluyadav0759@gmail.com	
42	Akanksha Pandey	akanksha8299@gmail.com	
43	Sanket Shankar Mishra	sanketshanker@gmail.com	
44	Rahul Jaiswal	rj642034@gmail.com	
45	Kavita Shukla	kavitashukla493@gmail.com	
46	Vishwani Sharma	vishwani011999@gmail.com	

47	Richa Pandey	richapandey0108@gmail.com	
48	Monika Mahajan	monikamahajan678@gmail.com	
49	Prachi Singh	Prachisngh246@gmail.com	
50	Shweta	syvns55555@gmail.com	
51	Anam Naheed	khananamnaheed@gmail.com	
52.	Satyanarayan Mishra	mishrasatyanarayan15@gmail.com	
53.	Tsering Lamoo	tsinglamoo@gmail.com	
54.	Shivangi Singh	shivangi2102.as@gmail.com	
55.	Shraddha Pandey	shraddhapandey32@gmail.com	
56.	Ayushi Gupta	ayushiguptaiilm@gmail.com	
57.	Nidhi Tyagi	nidhityagi1001@gmail.com	
58.	Sinha Sahab	sahab.sinha@gmail.com	
59.	Ibha Suhani	ibha.suhani@gmail.com	
60.	Ishaan Dev	ishandev2901@gmail.com	
61.	Ankita Tiwari	ankitatiwari87392@gmail.com	
62	Neeraj K Maurya	Oo7nkm@gmail.com	
63.	Mahendra Kumar		

<u>Operational and Maintenance Manual, Pre-flight check list, and user manual for Flir and Micasense and Indoor training manuals</u>

Figure 55 Snaps of the training

Figure 56 Snaps of the training 1

Figure 57 Snaps of the training 2

Figure 58 Snaps of the training 3

Figure 59 Snaps of Instrument

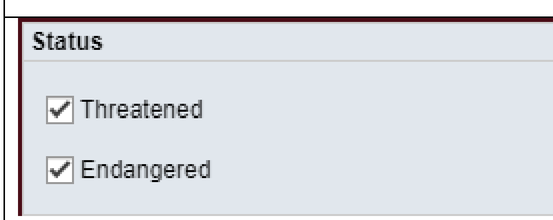
Figure 60 UAV Flight in the Field (Munsiyari)

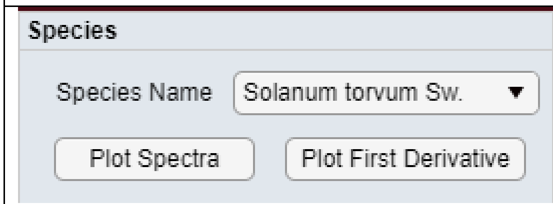
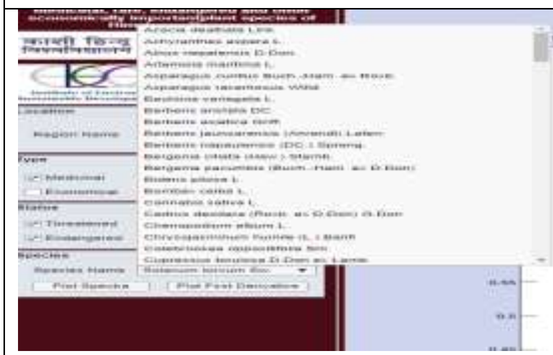
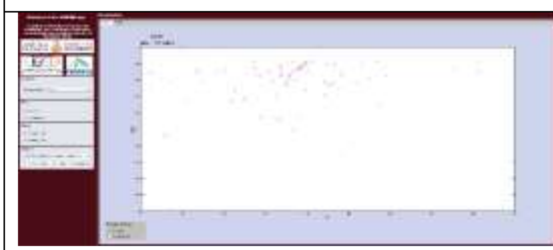
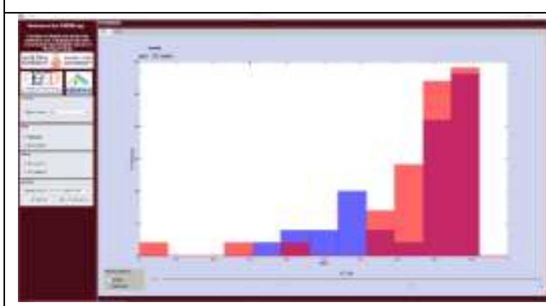
Appendix 3 New Database generated (Types)

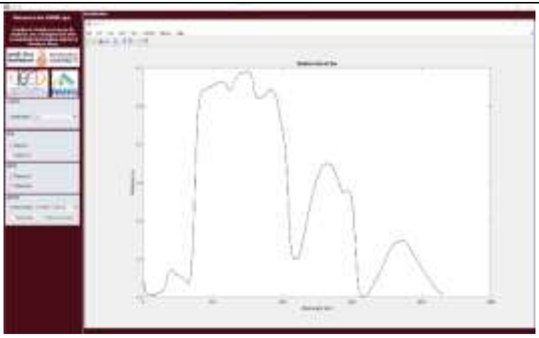
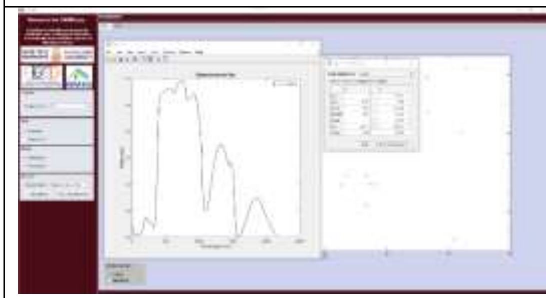
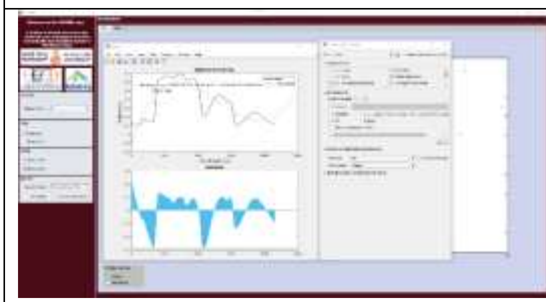
1. **SHRIME: -** The Spectral tool of Himalayan Rare, Invasive, Medicinal and Economical (SHRIME) Plant Species is created as a deliverable of the objective 2. This MATLAB tool could be utilized to create a spectral library in terms of reflectance and first derivative for the Himalayan plant species and a toolbox to be installed to extend its functionality further. This tool is developed for the visualization, export and pre-processing of hyperspectral data, which is crucial for monitoring the high-value plants in the Himalayan region, their sustainable utilization, and the conservation of related resources. The ultimate goal of this tool is in offering a dynamic platform for reflectance spectrum and derivative spectrum generation of various plant species, creation of the spectral library, application of basic fitting operations on the spectra's generated, understanding the relationship between the chlorophyll content with vegetation indices and a database of the various Himalayan species in terms of location, type, status and other parameters.

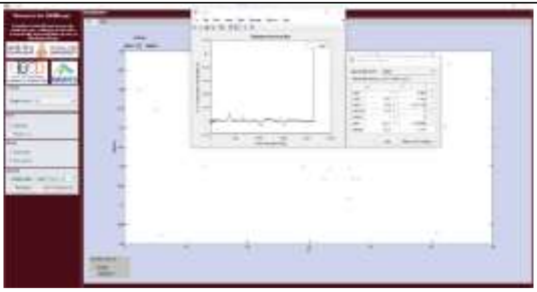
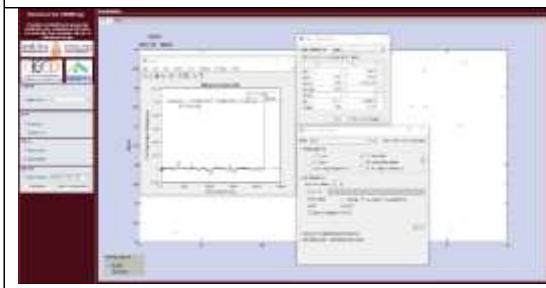
   The software with the database in a DVD has been attached with this FTR for the use of NMHS. The functionalities and a detailed description are given in table 19.

Table 19 List of functionalities and a detailed description

Functionalities	Description
	Selection of the Himalayan climatic regions is conducted using this feature. The regions to select are Tropical, Sub-Tropical, Alpine, Temperate or All. The type of species is selected using the check box. Currently, two types of species are specified. Medicinal or Economical.
	Selection of the Status of the species as per IUCN guidelines is made using the check box. The status currently being taken into consideration is Threatened and Endangered. At least one of the status types is mandatory; otherwise, an

	error message would be popped.
	Selection of the species for plotting the spectra and first derivative is conducted using these buttons. Plot Spectra and Plot first derivative would open a new window where the spectra and first derivative would be available for basic fitting operations.
	The list of species available, from which the selection of the species for plotting the spectra and first derivative could be made, is assessed from here.
	The relationship between the vegetation indices, NDVI and MSAVI with chlorophyll content is depicted here using a scatter plot.
	The Quantitative details of the species having NDVI or MSAVI values are represented here using a Histogram plot. The representation could be made by selecting only Medicinal or Economical species or both.
	The representation when both, Medicinal and Economical species are selected.

	The bin width of the histogram would be adjusted using this feature.
	The details of the species are represented here in tabular format. The species could be sorted based on their name, regions, status or type.
	After selecting the species and pressing the Plot Spectra button, spectra of the selected species popped. Further operations are also possible such as descriptive statistics of the species and basic fitting operations.
	The descriptive statistics of the species in terms of mean, median, mode, min value, max value etc., are represented here.
	Basic fitting operations with the generated spectra are represented here. Fittings in terms of linear, quadratic, polynomial, fit results in terms of R^2, the norm of residuals etc., Error Estimate, Interpolate/Extrapolate Data is available here.

	The descriptive statistics of the first derivative of the species in terms of mean, median, mode, min value, max value etc., are represented here.
	Basic fitting operations with the generated First derivative plot are represented here. Fittings in terms of linear, quadratic, polynomial, fit results in terms of R^2, the norm of residuals etc., Error Estimate, Interpolate/Extrapolate Data is available here.

The tool can be launched by clicking shrime.exe and adding it as a toolbox in the MATLAB working environment. The steps are given in Figures 61 to 64.

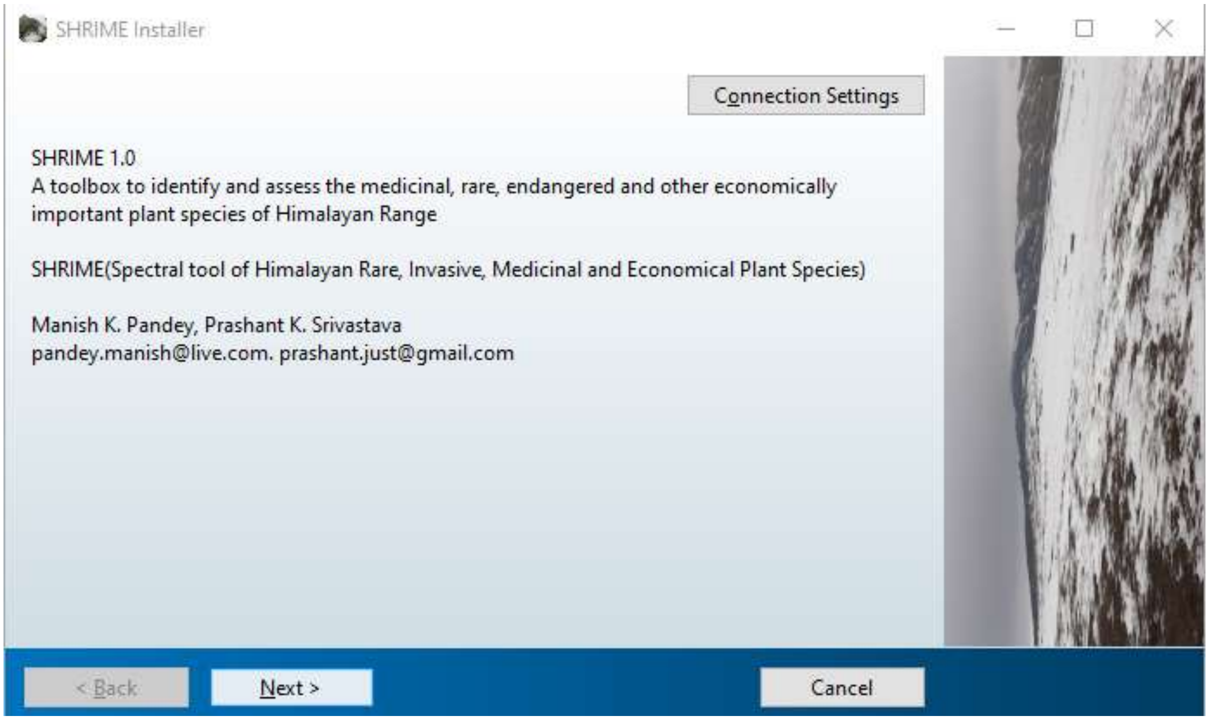


Figure 61 Information of the tool

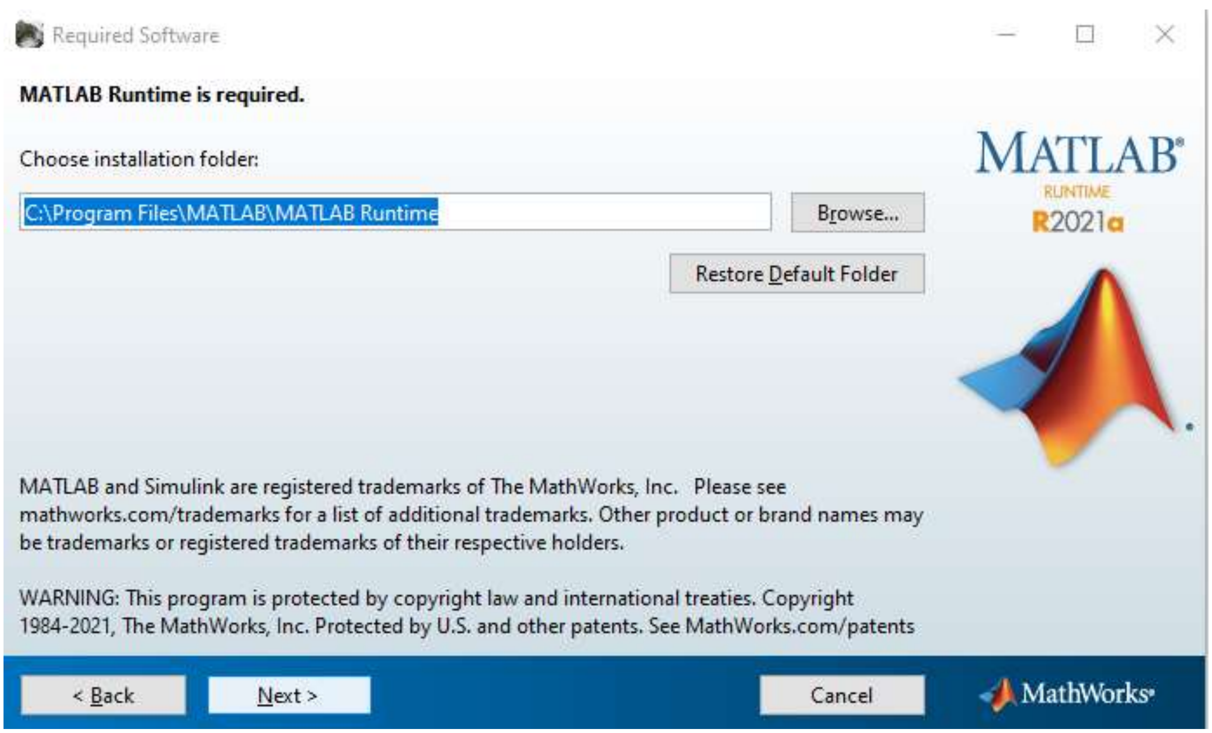


Figure 62 Selection of the Installation Location

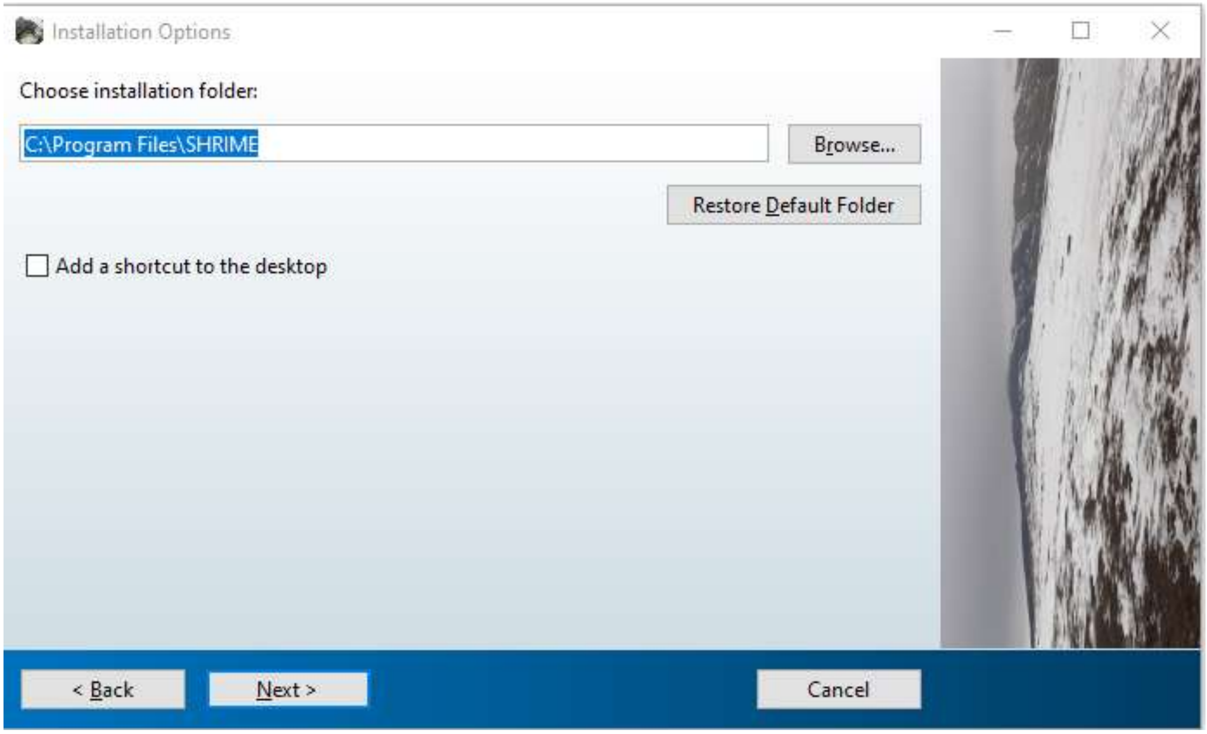


Figure 63 Installation Options

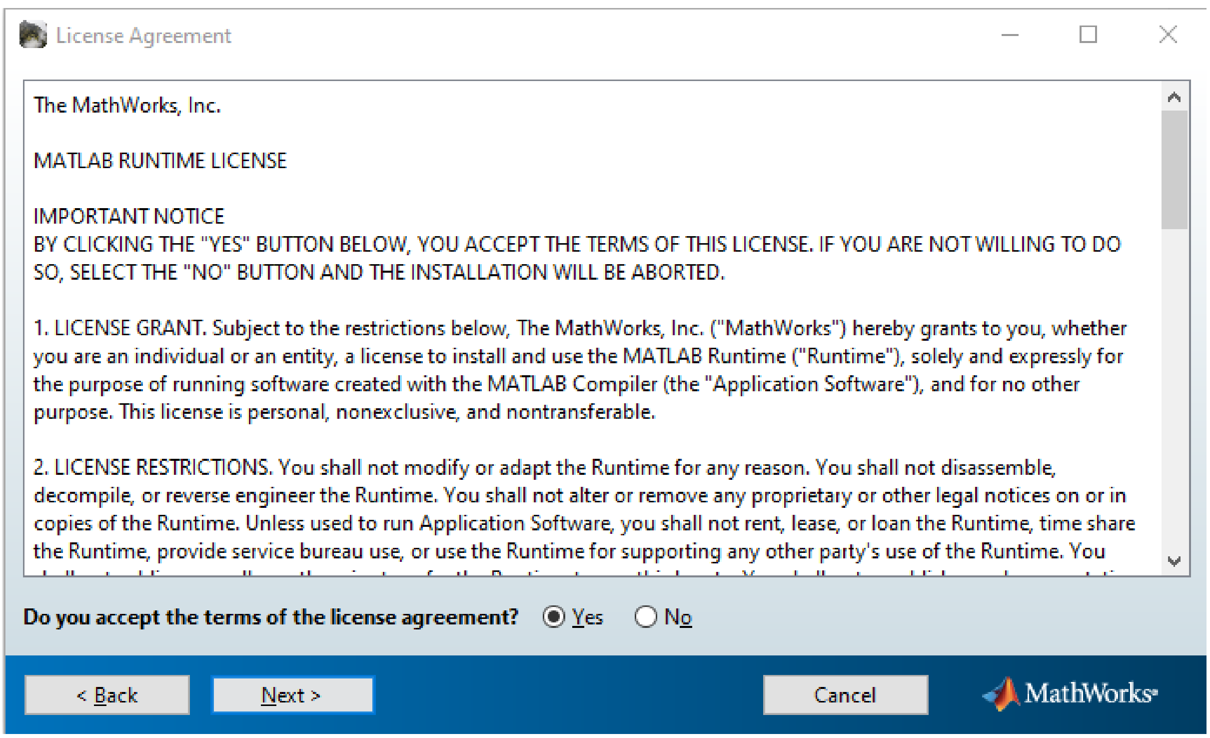


Figure 64 Licence Agreement

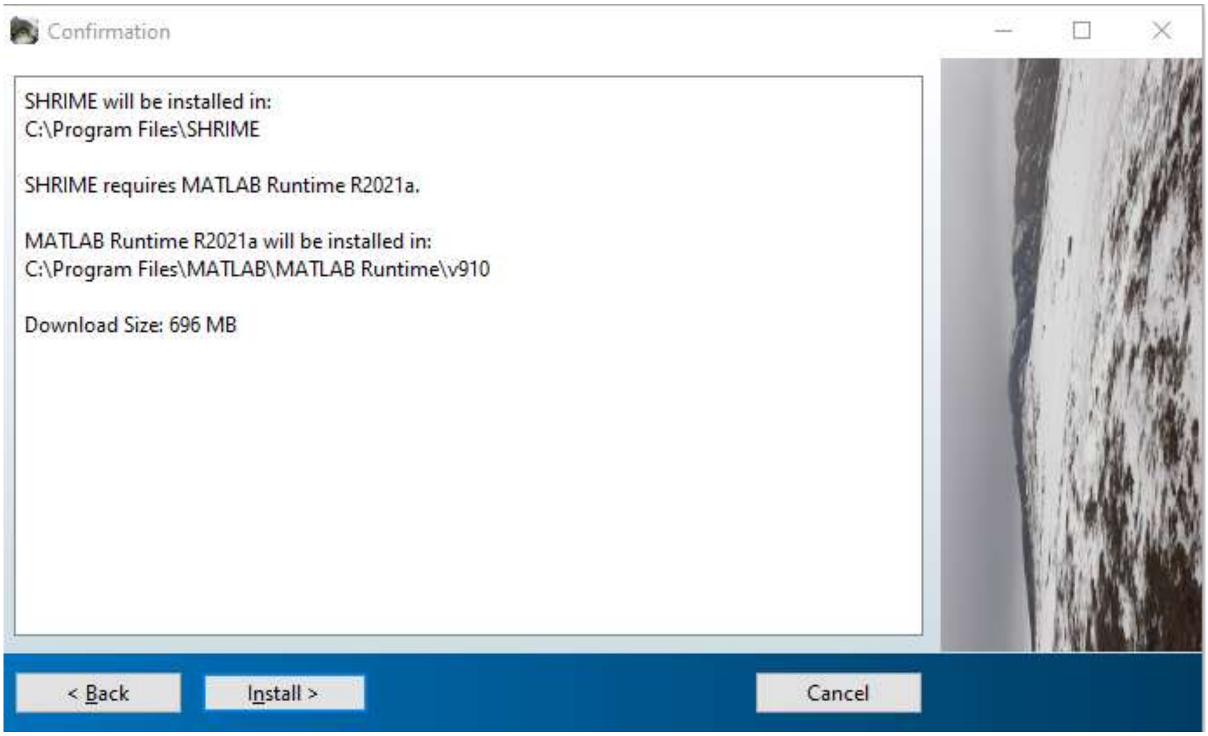


Figure 65 Confirmation Page

2. **Novel EOVs** database for species distribution modelling developed under the project

Table 20 Databases created for species distribution modelling

S. No.	**Novel variable full name**	**Description/ explanation of the variable**	**Temporal Scale**
1	Annual ET range	Mean ET of warmest month - Mean ET of coldest month	Annual
2	Annual mean ET	Average ET for each month, and then average these results over twelve months	Annual
3	Mean ET of coldest month	Monthly mean of daily ET	Month
4	Mean ET of coldest quarter	Quarterly mean of monthly ET	Quarter
5	Mean ET of warmest quarter	Quarterly mean of monthly ET	Quarter
6	Mean ET of warmest month	Monthly mean of daily ET	Month
7	Annual EVI range	Mean EVI of warmest month - Mean	Annual

		EVI of coldest month	
8	Annual mean EVI	Average EVI for each month, and then average these results over twelve months	Annual
9	Mean EVI of coldest month	Monthly mean of daily EVI	Month
10	Mean EVI of coldest quarter	Quarterly mean of monthly EVI	Quarter
11	Mean EVI of warmest quarter	Quarterly mean of monthly EVI	Quarter
12	Mean EVI of warmest month	Monthly mean of daily EVI	Month
13	Annual GPP range	Mean GPP of warmest month - Mean GPP of coldest month	Annual
14	Annual mean GPP	Average GPP for each month, and then average these results over twelve months	Annual
15	Mean GPP of coldest month	Monthly mean of daily GPP	Month
16	Mean GPP of coldest quarter	Quarterly mean of monthly GPP	Quarter
17	Mean GPP of warmest quarter	Quarterly mean of monthly GPP	Quarter
18	Mean GPP of warmest month	Monthly mean of daily GPP	Month
19	Annual LAI range	Mean LAI of warmest month - Mean LAI of coldest month	Annual
20	Annual mean LAI	Average LAI for each month, and then average these results over twelve months	Annual
21	Mean LAI of coldest month	Monthly mean of daily LAI	Month
22	Mean LAI of coldest quarter	Quarterly mean of monthly LAI	Quarter
23	Mean LAI of warmest quarter	Quarterly mean of monthly LAI	Quarter
24	Mean LAI of	Monthly mean of daily LAI	Month

	warmest month		
25	Annual Mean LST	average LST for each month, and then average these results over twelve months	Annual
26	Annual Mean Diurnal Range LST	(Mean of monthly (max temp - min temp))	Annual
27	Isothermality LST	100 * Mean Diurnal Range/ LST Annual Range	Variation
28	LST Seasonality (Standard Deviation)	(Standard deviation * 100)	Variation
29	LST Seasonality (CV)	The temperature Coefficient of Variation (CV) is the ratio of the standard deviation of the monthly mean LSTs to the mean of the monthly LSTs	Variation
30	Max LST of Warmest Month	Bio 5 = max ({Tmax1, ..., Tmax12}) monthly mean of daily maximum LST (°C)	Month
31	Min LST of Coldest Month	Bio 6 = min ({Tmin1, ..., Tmin12}) monthly mean of daily minimum LST (°C)	Month
32	LST Annual Range	(Max LST of Warmest Month - Min LST of Coldest Month)	Annual
33	Mean LST of Warmest Quarter	index provides mean LST during the warmest three months of the year	Quarter
34	Mean LST of Coldest Quarter	index provides mean LST during the coldest three months of the year	Quarter

Appendix 4 IHR States covered: Uttarakhand and Jammu & Kashmir

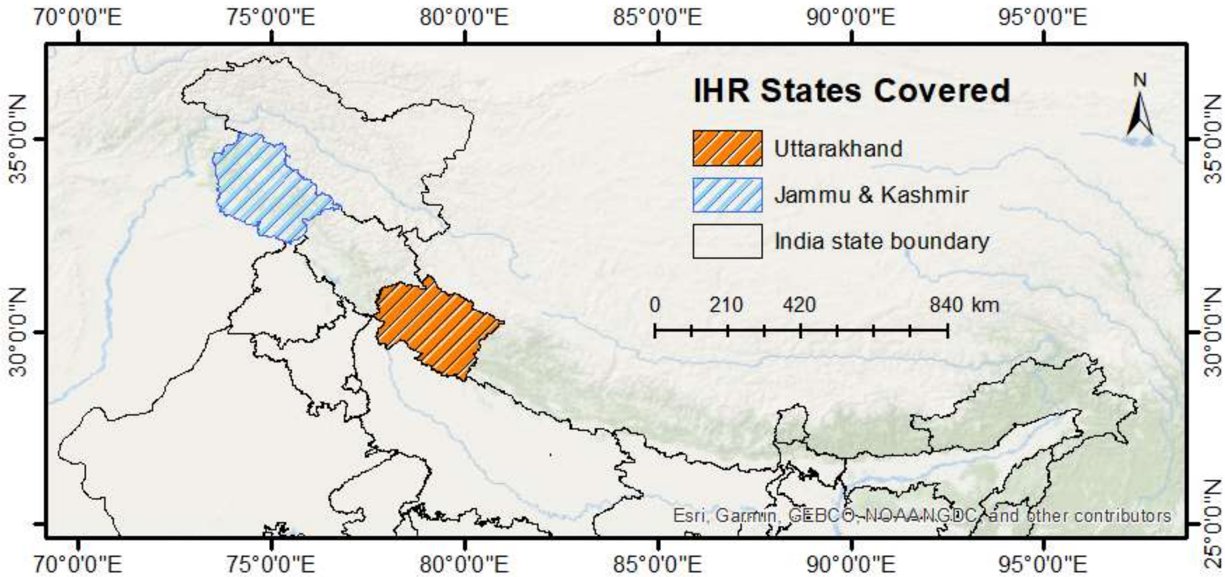


Figure 66 IHR States covered: Uttarakhand and Jammu & Kashmir

Appendix 5 Details of Technology/Methodologies/Models Developed

1. **SHRIME: -** The Spectral tool of Himalayan Rare, Invasive, Medicinal and Economical (SHRIME) Plant Species is created as a deliverable of the objective 2. This MATLAB tool could be utilized to create a spectral library in terms of reflectance and first derivative for the Himalayan plant species and a toolbox to be installed to extend its functionality further. This tool is developed for the visualization, export and pre-processing of hyperspectral data, which is crucial for monitoring the high-value plants in the Himalayan region, their sustainable utilization, and the conservation of related resources. The ultimate goal of this tool is in offering a dynamic platform for reflectance spectrum and derivative spectrum generation of various plant species, creation of the spectral library, application of basic fitting operations on the spectra's generated, understanding the relationship between the chlorophyll content with vegetation indices and a database of the various Himalayan species in terms of location, type, status and other parameters.
2. Fine-scale space-time distribution maps were generated for the species Betula utilis D.Don, Rhododendron campanulatum, Taxus wallichiana, Rhododendron arboretum and Pyracantha crenulata D.Don.
3. The species distribution limits of two alpine treeline species, namely Betula utilis D.Don and Rhododendron campanulatum over the Himalayan biodiversity hotspot were conducted. Novel Earth Observation Variables (NEOVs) was developed to identify the effective and ecologically significant NEOVs combinations, using four different models, i.e., bioclimatic model (BCM), biophysical model (BiophyM), phenology model (PhenoM), and hybrid model (HM), of which PhenoM, BiophyM, and HM were developed and tested for the first time in this study. This not only has helped in the assessment of the prevalent condition but would also help in the identification of the practices required for safeguarding the medicinal and threatened plant species.
4. Retrieval of biochemical parameters through lab-based methods as well as through Automated Radiative Transfer Models Operator (ARTMO) for retrieval of various vegetation parameters for the species Taxus wallichiana.
5. For the retrieval of biochemical parameters, various indices were generated and three different types of filters, Average Mean, Savitzky Golay, and Fast Fourier Transform (FFT) one at a time followed by feature selection on each denoised spectra were applied to smooth the spectra for better estimation of biochemical parameters.
6. Fine-grained species distributions were modelled using Sentinel-2, Sentinel-5P, Moderate Resolution Imaging Spectroradiometer (MODIS), ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) and Shuttle Radar Topography Mission (SRTM) along with other climatic variables over the sampling regions.

7. Several machine learning algorithms were incorporated to establish the relation between physical and climatic variables to estimate the probability distribution of species.

8. The phenology assessment was carried out for two decades to the effect of an elevational gradient, temperature, and precipitation on the start of the season (SOS) and end of the season (EOS) in major forest types of the Kumaon region of the western Himalaya using MODIS NDVI time series data (2001-2019).

9. The potential distribution of Rhododendron arboreum, a medicinal plant species found within the foothills of the Himalayas was also evaluated using Conventional Species Distribution Models (SDM) namely BIOCLIM and Maxent, a Machine Learning variant as well as Convolutional Neural Network (CNN).

10. Sample Extraction and Solution preparation

For TC, 1 g of crushed leaves was deflated with hexane using sonication. The hexane portions were discarded, and aliquots of methanol were concentrated using a rotary evaporator, extracted in chloroform, then dried under reduced pressure using a rotary evaporator, and then re-dispersed in methanol (1 mL).

11. HPLC processing

Taxoid standard paclitaxel (Sigma, St. Louis, USA) as used as a standard in HPLC for quantification. The working solution of paclitaxel was prepared from standard methanol. The UV-DAD scanned acquisitions of Taxol were performed at 230 nm.

The percentage of Taxol was calculated using Equation.

$$\text{Taxol (\%) Content} = \frac{\text{Ar.}_{\text{sample}} \text{ x Conc.}_{\text{std.}}\left({}^{\text{mg}}/_{\text{ml}}\right)}{\text{Ar.}_{\text{std.}} \text{ *1000 x Conc.}_{\text{sample}}\left({}^{\text{g}}/_{\text{ml}}\right)} \text{ x100}$$

where Ar_{std} and Ar_{sample} are the areas under the peak associated with the standard or reference and sample taxoid, respectively, and $Conc._{sample}$ and $Conc._{std}$ are the concentrations of the sample and reference taxoid, respectively.

12. Spectral data pre-processing

Data pre-processing is a crucial step. It has been stated that a key issue of applying filters for pre-processing is to allow the smoothing techniques to match the scale of the spectral features of interest. This led us to use filtering techniques namely Savitzky–Golay smoothing, Mean filter smoothing, and Fast fourier transform (FFT) (Figure 1). The transformed filtered spectra were applied with feature selection algorithm to facilitate the extraction of "useful" information from hyperspectral data. Absorption features in reflectance spectra are enhanced using derivative spectroscopy.

Adding the derivatives as features in the identification process optimizes and minimizes the number of bands required to achieve acceptable results due to larger JM distances. Using wavelengths selected data and spectra recorded data using spectroradiometer both were used in model called ARTMO (Automated Radiative Transfer Models Operator*)*. ARTMO generates new modelled values which were correlated with the real field values.

13. Statistical unfolding approach to understand influencing factors for Taxol content through Absorption based indices

In this study the spectra from spectroradiometer were processed using feature selection algorithm continuum removal to develop absorption-based indices. The PRISM software applies continuum removal to each spectrum and derives the spectral feature parameters (e.g., centre, depth, width, area, etc.). PRISM performs continuum removal twice and reports the feature parameters by the following steps (1) the initial set of continuum endpoints; and (2) an automatically adjusted set of continuum endpoints.

These indices were along with previously developed reflection-based indices were tested with ground measured taxol, chlorophyll, phenolic content, and land surface temperature parameters were put to various statistical analysis to check the robustness of developed indices.

14. Identification and Quantification of flavanols in *Rhododendron arboreum* by Mass Spectrometry through

14.1 Sample Extraction and Solution preparation

The plant material (2 g) was extracted three times with Methanol (MeOH) under reflux. The extracted solution was filtered and evaporated under reduced pressure on a rotary evaporator to give the MeOH extract. The extract was suspended in water, fractionated with ethyl acetate three times, and the ethyl acetate soluble part was concentrated in vacuum to yield the fraction suspended in HPLC grade MeOH.

14.2 Chromatographic fingerprinting

14.2.1 Chromatographic method development and System Suitability

The processed extract was then analysed in Accquity Waters ultra-performance liquid chromatography (UPLC) system (Waters Corp., Milford, MA, USA) equipped with a C18 column

(100 × 2.1 mm, particle size 1.7 mm; Waters Accquity) at fixed wavelength of 254nm for the UV detector.

Solvent system with water (solvent A) and ACN (solvent B) as the mobile phase. The gradient programming was: 0-1 min, 85% A: 15%, At 4 min 80% A: 20% B; Between 4-7 min, mobile phase changes to 70% A: 30% B. HPLC was performed at a flow rate of 0.3mL/min. Electrospray ionization (ESI) source was operated under the following parameters: electrospray voltage, 4.0 kV; capillary temperature, 275◦C; sheath gas, N_2. The instrument was operated with Empower Service Release 2

14.2.2 High Resolution Mass Spectrometry (HRMS)

UPLC method was applied for HRMS processing with the same column and conditions. The chromatogram obtained using HRMS when analysed using Mass Hunter software. Simultaneously *Rhododendron* compound library was prepared from all the previous research papers on genus *Rhododendron*. Library file generated was used as an input in Mass Hunter software. The generated hits of the compounds in the list were then shortlisted on the basis of statistical score and ppm error for each peak.