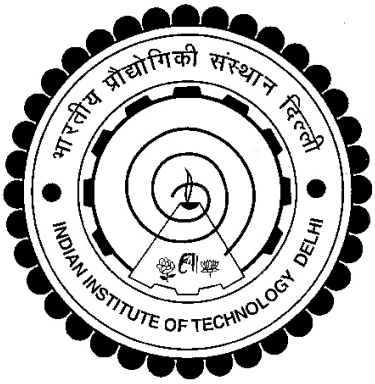
INTEGRATING SATELLITE REMOTE SENSING WITH SPECTRAL UNMIXING AND ENSEMBLE TECHNIQUES TO QUANTIFY SUB-PIXEL LAND COVER HETEROGENEITY AND IMPROVE HYDROLOGICAL MODELLING

NITESH PATIDAR



DEPARTMENT OF CIVIL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY DELHI

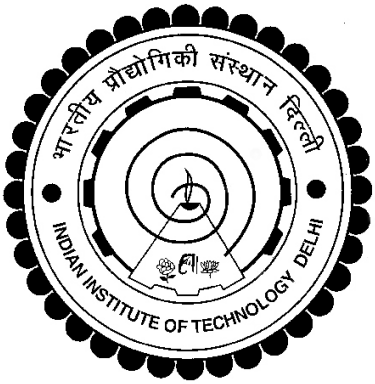
OCTOBER 2019

INTEGRATING SATELLITE REMOTE SENSING WITH SPECTRAL UNMIXING AND ENSEMBLE TECHNIQUES TO QUANTIFY SUB-PIXEL LAND COVER HETEROGENEITY AND IMPROVE HYDROLOGICAL MODELLING

by

NITESH PATIDAR
DEPARTMENT OF CIVIL ENGINEERING

Submitted
in fulfilment of the requirements of the degree of Doctor of Philosophy
to the



INDIAN INSTITUTE OF TECHNOLOGY DELHI

OCTOBER 2019

CERTIFICATE

This is to certify that the thesis entitled, "**Integrating Satellite Remote Sensing with Spectral Unmixing and Ensemble Techniques to Quantify Sub-Pixel Land Cover Heterogeneity and Improve Hydrological Modelling**" being submitted by **Mr. Nitesh Patidar** to the Indian Institute of Technology Delhi, for the award of degree of Doctor of Philosophy is a bonafide record of research work carried out by him under my supervision and guidance. The thesis, in my opinion has reached the requisite standard, fulfilling the requirements for the award of degree of ***Doctor of Philosophy***. The research report and results presented in this thesis have not been submitted, in part or full, to any University or Institute for the award of any degree or diploma.

Dr. Ashok K. Keshari
Professor
Department of Civil Engineering
Indian Institute of Technology Delhi
Hauz Khas, New Delhi-110016, India

ACKNOWLEDGEMENTS

I would like to express my sincere thanks and deepest gratitude to my supervisor Professor Ashok K. Keshari for his consistent advice and valuable guidance throughout my PhD. I very much appreciate his persistence, kindness and encouragement. My gratitude also extends to the Student Research Committee (SRC) members Prof. G. Tiwari (Chairperson), Dr. C. T. Dhanya (Departmental Expert) and Dr. S. Dey (Concerned Area Expert) for their valuable comments and suggestions.

I gratefully acknowledge scholarly discussions in our Simulation Lab with Dr. Basant Yadav, Mr. Gopinadh Rongali, Mr. Shushobhit Chaudhary, Mr. Sameer Arora, Mr. Amarsinh Landage, Dr. Mulue Sewinet and Dr. Veenarsi. I wish to thank all of my friends, specially Rohit Namdeo, Arvind Yadav, JD Sharma, Ajay Patel and Bhupendra Ghodki, for their warm friendship.

I am indebted to my parents Mr. G. L. Patidar and Mrs. Meera Patidar for their blessings, love and support. I also thank my sisters and brother for their love and encouragement. Finally, I owe heartfelt thanks to my wife Mrs. Durga Patidar and daughter Stuti Patidar for their love and support.

Nitesh Patidar

ABSTRACT

The appropriate management of water resources draws paramount significance in the backdrop of the ever increasing water demands for agricultural, industrial and domestic uses. The hydrological models are important tools for water resources management and for simulating the effects of various natural and anthropogenic changes, such as climate and land cover change. Integration of the data obtained from satellite remote sensing, such as land cover, precipitation, topography and soil moisture, enhances the utility of these models for simulating hydrological components on a finer spatio-temporal scale. Keeping the above in view, this study aimed at developing new techniques for improving land cover classification in heterogeneous land covers and integrating them with the hydrological model to improve hydrological simulations.

In the present study, an ensemble model is developed for improving the sub-pixel classification by combining the outputs of three different techniques, namely Linear Spectral Mixture Analysis (LSMA), Multi-Layer Perceptron (MLP) and Support Vector Regression (SVR). The ensemble model utilizes a multi-model ensemble approach, named Bayesian Model Averaging (BMA). A comparative evaluation is performed first to identify the most accurate and feasible model among various LSMA models, namely Multiple Endmember Spectral Mixture Analysis (MESMA), Normalized Spectral Mixture Analysis (NSMA), Pre-Screened and Normalized MESMA (PNMESMA) and Spatially Adaptive Spectral Mixture Analysis (SASMA). The PNMESMA is selected as an ensemble member model considering its higher accuracy and the lower computational burden. The comparison of the ensemble member models reveals that the SVR is more robust than the other member models as it produces the highest overall accuracy (84%) and the κ value (0.73). The developed ensemble model ranked higher in terms of accuracy as compared to the member models in all type of

land covers. The overall accuracy is improved by approximately 3% as compared to the best performing member model.

A spectral unmixing method is developed in the present study to investigate the annual dynamics of impervious surface at the sub-pixel level using time series analysis of the Landsat images. The developed method integrates temporal contextual information into the Normalized MESMA (NMESMA) to improve the separation between the spectrally similar land covers. A temporal filtering algorithm is also developed to improve the consistency of impervious surface between the years. The developed method has been tested in the part of the National Capital Region (NCR), India, to investigate the annual dynamics of impervious surface from 1992 to 2017. The accuracy of the developed method has been tested by comparing the estimated impervious surface fraction with the reference fractions obtained from the high resolution (~1 m) image of the OrbitView satellite. The developed method estimates very precise sub-pixel fractions of impervious surface. The mean overall accuracy is observed to be 89.57%. The improved accuracy is achieved by reducing the confusion between the bare soil and impervious surface using the temporal contextual information. The annual impervious surface fractions derived using the developed method indicates that the impervious surface in the NCR, India, has increased significantly during the past 26 years. Moreover, the urban growth rate was considerably high between 2000 and 2008 when compared to the other periods. Out of the total study area, i.e. 3986 km^2, approximately 377 km^2 area was observed to be impervious in 1992, which has increased to approximately 708 km^2 in 2017.

Further, the developed sub-pixel classification technique is integrated with a hydrological model to improve hydrological simulations in a heterogeneous region. A physically based distributed hydrological model, named WetSpass, is used to simulate annual water balance components, namely, evapotranspiration, runoff and groundwater recharge. To demonstrate the potential of sub-pixel land cover data for improved hydrological modelling, a

detailed comparison is carried out between the simulated hydrological components obtained from the fraction (sub-pixel) based parameterization approach and the traditional land use (per-pixel) based approach. Results show that the aggregation of land cover information within the raster cell in the per-pixel approach leads to overestimation of runoff by 10% and groundwater recharge by 7.7%, and underestimation of evapotranspiration by 6.5%. In addition, the annual sub-pixel land cover data obtained from the developed method is utilized to assess the impact of urbanization on groundwater recharge in the part of NCR, India, from 1992 to 2014. The annual groundwater recharge has decreased considerably in the areas where the impervious surface has increased. The total annual recharge decreased from ~550 Mm^3 to ~531 Mm^3 due to an increase of impervious surface from ~366 km^2 to ~684 km^2 between 1994 and 2012 in the study area. The use of sub-pixel land cover data in hydrological modelling can help in reducing the uncertainty and can significantly improve the reliability of hydrological simulations.

सारांश

कृषि, औद्योगिक और घरेलू उपयोगों के लिए बढ़ती पानी की मांगों के कारण जल संसाधनों का उपयुक्त प्रबंधन अत्यंत आवश्यक है। हयड्रोलॉजिकल मॉडल जल संसाधन प्रबंधन और विभिन्न प्राकृतिक और मानवजनित परिवर्तनों के प्रभावों का अनुकरण करने के लिए महत्वपूर्ण हैं, जैसे जलवायु और भूमि परिवर्तन। सैटेलाइट रिमोट सेंसिंग से प्राप्त आंकड़ों का एकीकरण, जैसे भूमि आवरण, वर्षा, स्थलाकृति और मिट्टी की नमी, इन मॉडलों की उपयोगिता को हाइड्रोलॉजिकल घटकों के अनुकरण के लिए बढ़ाती है। उपरोक्त बातों को ध्यान में रखते हुए, इस अध्ययन का उद्देश्य विषम भूमि आवरण में भूमि आवरण वर्गीकरण विधि में सुधार के लिए नई तकनीकों को विकसित करना और उन्हें हाइड्रोलॉजिकल सिमुलेशन में सुधार के लिए हाइड्रोलॉजिकल मॉडल के साथ एकीकृत करना है।

वर्तमान अध्ययन में, एक एन्सेम्बल मॉडल तीन अलग-अलग तकनीकों के आउटपुट को मिलाकर उप-पिक्सेल वर्गीकरण में सुधार करने के लिए विकसित किया गया है, जैसेकि लीनियर स्पेक्ट्रल मिक्सचर एनालिसिस (एलएसएमए), मल्टी-लेयर परसेप्ट्रॉन (एमएलपी) और सपोर्ट वेक्टर रिग्रेशन (एसवीआर)। एन्सेम्बल मॉडल बेयेशियन मॉडल एवरेजिंग (बीएमए) का उपयोग करता है। विभिन्न एलएसएमए मॉडलों के बीच सबसे सटीक और व्यवहार्य मॉडल की पहचान करने के लिए सबसे पहले एक तुलनात्मक मूल्यांकन किया गया है, जिसमें शामिल है, मल्टीपल एंडेमबर स्पेक्ट्रल मिक्सचर एनालिसिस (एमईएसएमए), नॉर्मलाइज्ड स्पेक्ट्रल मिक्सचर एनालिसिस (एनएसएमए), प्री-स्क्रीन्ड और नॉर्मलाइज्ड एमईएसएमए (PNMESMA) और स्पेसियल एडाप्टिव स्पेक्ट्रल मिक्सचर एनालिसिस (SASMA)। PNMESMA को इसकी उच्च सटीकता और कम कम्प्यूटेशनल बोझ को देखते हुए एक सदस्य मॉडल के रूप में एन्सेम्बल मॉडलिंग के लिए चुना गया है। सदस्य मॉडल की तुलना से पता चलता है कि एसवीआर अन्य सदस्य मॉडल की तुलना में अच्छा है क्योंकि यह उच्चतम एक्यूरेसी (84%) और k (0.73) देता है। सभी प्रकार के लैंड कवर में सदस्य मॉडल की तुलना में

विकसित एन्सेम्बल मॉडल सटीकता के मामले में उच्च स्थान पर है। सबसे अच्छा रिजल्ट देने वाले सदस्य मॉडल की तुलना में एन्सेम्बल मॉडल 3% ज्यादा एक्यूरेसी देता है|

लैंडसैट इमेजेस के समय श्रृंखला विश्लेषण का उपयोग करके उप-पिक्सेल स्तर पर अभेद्य सतह की वार्षिक स्केल पर जांच करने के लिए वर्तमान अध्ययन में एक स्पेक्ट्रल अनमिक्सिंग विधि विकसित की गई है। विकसित पद्धति वर्णक्रमीय-संदर्भ जानकारी को नॉर्मलाइज्ड MESMA (NMESMA) में स्पेक्ट्रल रूप से समान भूमि आवरणों के बीच पृथक्करण को बेहतर बनाने के लिए इंटेग्रट करती है। एक टेम्पोरल फ़िल्टरिंग एल्गोरिथ्म भी विकसित किया गया है जो एस्टिमेटेड अभेद्य क्षेत्रफल की कंसिस्टेंसी में सुधार करता है। सन 1992 से 2017 तक अभेद्य सतह की वार्षिक क्षेत्रफल की जांच करने के लिए भारत के राष्ट्रीय राजधानी क्षेत्र (एनसीआर) के हिस्से में विकसित विधि का परीक्षण किया गया है। ऑर्बिटव्यू उपग्रह की उच्च रिज़ॉल्यूशन (~1 मीटर) छवि से प्राप्त इनफार्मेशन का उपयोग करके विकसित विधि का अवलोकन किया गया है। विकसित विधि अभेद्य सतह के बहुत सटीक उप-पिक्सेल अंशों का अनुमान लगाती है। औसत समग्र सटीकता 89.57% देखी गई है। विकसित विधि का उपयोग करके प्राप्त की गई वार्षिक अभेद्य सतह यह इंगित करता है कि पिछले 26 वर्षों के दौरान एनसीआर, भारत में अभेद्य सतह में काफी वृद्धि हुई है। इसके अलावा, शहरी विकास दर अन्य अवधि की तुलना में 2000 और 2008 के बीच काफी अधिक थी। कुल अध्ययन क्षेत्र में से, यानी 3986 किमी^{2}, लगभग 377 किमी^{2} क्षेत्र 1992 में अभेद्य था, जो 2017 में बढ़कर लगभग 708 किमी^{2} हो गया है।

विकसित उप-पिक्सेल वर्गीकरण तकनीक हाइड्रोलॉजिकल सिमुलेशन को बेहतर बनाने के लिए एक हाइड्रोलॉजिकल मॉडल के साथ एकीकृत की गई है। इसमें, WetSpass नामक मॉडल का उपयोग करके वार्षिक जल संतुलन घटकों, जैसे वाष्पीकरण और भूजल पुनर्भरण का आकलन किया गया है| बेहतर हाइड्रोलॉजिकल मॉडलिंग के लिए उप-पिक्सेल लैंड कवर डेटा की क्षमता को प्रदर्शित करने के लिए, अंश (उप-पिक्सेल) आधारित पैरामीटराइजेशन विधि और पारंपरिक प्रति-पिक्सेल आधारित विधि से प्राप्त घटकों के बीच एक विस्तृत तुलना की गई है। परिणाम बताते हैं कि प्रति-पिक्सेल विधि में पिक्सेल

के भीतर भूमि कवर की जानकारी के एकत्रीकरण से 10% तक रनाफ और 7% तक भूजल पुनर्भरण का ओवरस्टिमशन पाया गया, और 6.5% तक वाष्पीकरण कम पाया गया। इसके अलावा, 1992 से 2014 तक एनसीआर, भारत के हिस्से में भूजल पुनर्भरण पर शहरीकरण के प्रभाव का आकलन करने के लिए विकसित विधि से प्राप्त वार्षिक उप-पिक्सेल लैंड कवर डेटा का उपयोग किया गया है। उन क्षेत्रों में वार्षिक भूजल पुनर्भरण में काफी कमी आई है जहां अभेद्य सतह बढ़ गई है। अध्ययन क्षेत्र में 1994 से 2012 के बीच ~366 किमी^{2} से ~684 किमी^{2} तक की अभेद सतह की वृद्धि के कारण कुल वार्षिक रिचार्ज ~550 एमएम^3 से ~531 एमएम^3 तक घट गया। हाइड्रोलॉजिकल मॉडलिंग में उप-पिक्सेल लैंड कवर डेटा का उपयोग अनिश्चितता को कम करने में मदद कर सकता है और हाइड्रोलॉजिकल सिमुलेशन की विश्वसनीयता में काफी सुधार कर सकता है।

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

R_b	Reflectance of pixel in band *b*
f_i	Fraction of endmember *I* in a pixel
e_b	Residual error of band *b* in a pixel
$R_{i,b}$	Reflectance of endmember *i* in band *b*
rRMSE	Reflectance Root Mean Square Error
$\overline{R_b}$	Normalized reflectance of pixel in band *b*
$\overline{R_{i,b}}$	Normalized reflectance endmember *i* in band *b*
$R_{m,n,k}$	Synthetic endmember spectrum of pixel (*m*,*n*) for endmember *k*
$R_{i,j,k}$	Spectrum of the endmember candidate (*i*,*j*) for endmember *k*
l	Size of moving window
R	Set of real numbers
$\mathbf{R}^n$	Real number space of dimension n
$\boldsymbol{x}$	Input vector in Support Vector Regression
$\boldsymbol{t}$	Output variable in Support Vector Regression
$\boldsymbol{\omega}$	Weight vector in the feature space
$\mathbf{x}$	Input matrix
b	Bias
C	Regularization parameter
ξ^- and ξ^+	Slack variables representing upper and lower bounds on the outputs of the system
L_ε	ε -insensitive loss function
γ	Kernel parameter

ε	ε -insensitive loss function parameter
$p(\boldsymbol{\theta})$	Prior probability density function
$p(\boldsymbol{\theta}\|y_{\mathrm{T}})$	Posterior probability density function
y_{T}	Evidentiary data or training data
$\boldsymbol{\theta}$	Parameter vector
$l(\boldsymbol{\theta})$	Likelihood function
c_i	Estimated land cover fraction from model i
w_i	Bayesian weight
σ_i^2	Variance
A_r	Modelled class areas
$\overline{A_c}$	Reference class areas
$P_{r,c}$	Element of the confusion matrix $\mathbf{P}$ at row r and column c
P_0	Observed proportion of agreement
P_e	Expected proportion of agreement
$\overline{f_i}$	Reference fractions in validation sample i
f_i	Modelled fractions in validation sample i
κ	Kappa coefficient
ET	Evapotranspiration
S	Surface runoff
R	Groundwater recharge
a	Area of the respective land cover class
P	Precipitation

I	Interception fraction
T_v	Actual transpiration from vegetated surface
$S_{v\text{-}pot}$	Potential surface runoff from vegetated surface
C_{sv}	Surface runoff coefficient for vegetated areas
S_v	Surface runoff from vegetation
C_{Hor}	Coefficient for parameterizing the Hortonian overland flow equation
T_{rv}	Reference transpiration of a vegetated surface
E_o	Potential evaporation from open water
$f(\theta)$	Function of soil moisture content
ET_v	Actual evapotranspiration
E_s	Evaporation from the bare soil

LIST OF ACRONYMS

ACICA	Abundance Characteristic-Based Independent Component Analysis
AGWA	Automated Geospatial Watershed Assessment tool
ANN	Artificial Neural Network
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AVHRR	Advanced Very High Resolution Radiometer
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
BCI	Biophysical Composition Index
BMA	Bayesian Model Averaging
BP	Back Propagation
BRT	Boosted Regression Trees
CGWB	Central Ground Water Board
CMESMA	Composite Multiple Endmember Spectral Mixture Analysis
CRHM	Cold Regions Hydrological Model
CWT	Continuous Wavelet Transform
DEM	Digital Elevation Models
DHSVM	Distribute Hydrology Soil vegetation Model
DN	Digital Number
ELM	Extreme Learning Machine
EM	Ensemble Model
EMM	Ensemble Member Model
EROS	Earth Resources Observation and Science
ETM+	Enhanced Thematic Mapper Plus

FCC	False Colour Composite
IDW	Inverse Distance Weighting
IMD	Indian Meteorological Department
IS	Impervious Surface
ISODATA	Iterative Self-Organizing Data Analysis Technique
LAI	Leaf Area Index
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LISS III	Linear Imaging and Self-Scanning Sensor
LSMA	Linear Spectral Mixture Analysis
MESMA	Multiple Endmember Spectral Mixture Analysis
MK	Mann-Kendall
MLP	Multi-layer Perceptron
MLP	Multi-Layer Perceptron
MNDWI	Modified Normalized Difference Water Index
MNF	Minimum Noise Fraction
MODIS	Moderate Resolution Imaging Spectroradiometer
NBSS and LUP	National Bureau of Soil Survey and Land Use Planning
NDVI	Normalized Difference Vegetation Index
NMESMA	Normalized Multiple Endmember Spectral Mixture Analysis
NOAA	National Oceanic and Atmospheric Administration
NSMA	Normalized Spectral Mixture Analysis
NSPI	Neighbourhood Similar Pixel Interpolator
OLI	Operational Land Imager
PDF	Probability Density Function

PNMESMA	Pre-screened and Normalized Multiple Endmember Spectral Mixture Analysis
RF	Random Forests
RFR	Random Forest Regression
RMSE	Root Mean Square Error
SASMA	Spatially Adaptive Spectral Mixture Analysis
SLC	Scan-line Corrector
SOM	Self-Organizing Map
SPOT	Systeme Probatoire d'Observation de la Terre
SRTM	Shuttle Radar Topography Mission
SVR	Support Vector Regression
SWAT	Soil and Water Assessment Tool
SWIR	Short-wave Infrared
TM	Thematic Mapper
TOPMODEL	TOPography based hydrological MODEL
VIC	Variable Infiltration Capacity model
VNIR	Visible and Near-infrared
WetSpass	Water and Energy Transfer between Soil, Plants and Atmosphere under quasi Steady State
XWT	Cross-wavelet Transform