DEVELOPMENT OF SPECTRAL MODELS FOR ASSESSING SOIL FERTILITY STATUS IN VARIOUS AGROECOLOGICAL SUB-REGIONS OF PUNJAB

Thesis

Submitted to the Punjab Agricultural University
in partial fulfilment of requirements
for the degree of

MASTER OF SCIENCE

in

SOIL SCIENCE
(Minor Subject: Chemistry)

By

Saheed Garnaik
(L-2017-A-155-M)

Department of Soil Science
College of Agriculture

LUDHIANA – 141004

2019

CERTIFICATE I

This is to certify that the thesis entitled, "**Development of spectral models for assessing soil fertility status in various agroecological sub-regions of Punjab**" submitted for the degree of **M.Sc.** in the subject of **Soil Science** (Minor Subject: **Chemistry**) of Punjab Agricultural University, Ludhiana, is a bona fide research work carried out by **Saheed Garnaik (L-2017-A-155-M)** under my supervision and that no part of this thesis has been submitted for any other degree.

The assistance and help received during the course of investigations have been fully acknowledged.

_____

(Major advisor)
B.S. Sekhon
Editor Research
Directorate of Research
Punjab Agricultural University
Ludhiana

CERTIFICATE (II)

This is to certify that the thesis entitled, "**Development of spectral models for assessing soil fertility status in various agroecological sub-regions of Punjab**" submitted by **Saheed Garnaik (L-2017-A-155-M)** to Punjab Agricultural University, Ludhiana, in partial fulfilment of requirements for the degree of M.Sc. in the subject of Soil Science (Minor Subject**: Chemistry**) has been approved by the Student's advisory Committee along with Head of department.

_____
(Dr. B.S. Sekhon)
Major Advisor

_____
(Dr. Anil Sood)
External Examiner
Head
Agro-ecosystem and Crop Modelling Division
Punjab Remote Sensing Centre
PAU-Campus
Ludhiana - 141004 (Punjab), India

_____
(Dr. O.P. Choudhary)
Head of the Department

_____
(Dr. G.K. Sangha)
Dean Postgraduate Studies

ACKNOWLEDGEMENTS

At the very outset, with folded hands, I bow my head with reverence and dedicatedly accord my gratitude to the "Lord Jagannath", the merciful and compassionate whose grace glory and blessing gave me the courage in odd critical times for the successful completion of this degree.

Not only to fulfil a formality, but also to express the feeling of my heart, I put on record my deepest gratitude and profound indebtedness to my esteemed Major Advisor ***Dr. B.S. Sekhon****,* ***Editor Research****,* ***Directorate of Research*** *to help completing my research and degree successfully.*

I express my deep sense of gratitude to the respected members of my Advisory Committee, viz. ***Dr. Rajeev Sikka****, Sr. Soil Chemist,* ***Dr. Manpreet Kaur****, Asst. Prof. (Dept. of Chem),* ***Dr. Dhanwinder Singh****, Sr Soil Chemist (Dean PGS Nominee) and* ***Dr. O.P. Choudhary****, Head (Dept. of Soil Science) for their encouragement and guidance during the research work.*

I am also thankful to ***Dr. R.N. Sahoo****, Principal Scientist Div. of Agricultural Physics, IARI and* ***Dr. Manjeet Singh,*** *Head (Dept. of FM & PE) for his unfailing encouragement, valuable suggestions and moral support in present study.*

I am also thankful to faculties of ***O.U.A.T., Bhubaneswar*** *for their unfailing encouragement, valuable suggestions and moral support in present study.*

I express my respect, love and gratitude for my parents ***Mr. Arun Kumar Garnaik*** *and* ***Mrs. Pramila Pradhan****, I owe so much to both of them and acknowledge the gift of their guidance, kindness and support. I feel pleasure in acknowledging the warm affection and sincere wishes of my loving brother* ***Swaraj Salila Garnaik*** *his constant and unconditional support, cooperation and love.*

I would like to thank the supporting office and field staff of the Soil Science for their help during my research work. Special thanks to ***Shasikant sir, Samsher****,* ***Rajkumar****,* ***Omprakash, Lakhveer*** *and* ***Sunil ji, Manpreet, Tony and Jaskiran madam*** *for their cooperation and support in my research work.*

Indeed, the words at my command are inadequate to convey warm thanks to my friend ***Hari Sankar Nayak*** *for helping me through the period of my research.*

Pleasant company, ever willing help, regular encouragement and sweet memories of associates and my friends, ***Bhuyan****,* ***Dishri, Loky Bro, Ashutosh, Asish****,* ***Raavi****,* ***Pratik****,* ***Navneet****,* ***Prabhjot****,* ***Pratahvidya, Naina, Sonal****,* ***Samrat, Shikha, Kamaljit, Basit Raja bhai, Bhabani Mondal sir, Abinash bhai, Bhabani, Bibhuti, Srikanta, Sabya, Jaga, Arpan, Subrat, Kiran, Sradhaya, Manoj, Natha, Atul, Shubham, Bro, Aji Bhuta, Ani Baya, Bunty, Muna bhai, Tuna bhai, Sam babu, Sami, Pinu, Chinu, Cheru kaka, Tiki mausi, Aai, Raja, Roji*** *will always remain as a precious asset with me.*

Mr. Deepak Kumar *and* ***Mr. Beer Bahadur*** *(Ekta Computer Centre, Opp. P.A.U. Gate No.3, Ludhiana) deserve special thanks for bringing out this manuscript in its presentable form.*

Last but not least, I duly acknowledge my sincere thanks to all those who loved and cared for me. All may not be mentioned but none is forgotten. I feel proud to be a part of PAU, Ludhiana where I learnt a lot and spent some unforgettable moments of my life.

Date: _____

Place: Ludhiana **Saheed Garnaik**

Title of the thesis	:	Development of spectral models for assessing soil fertility status in various agroecological sub-regions of Punjab
Name of the student and admission number		Saheed Garnaik (L-2017-A-155-M)
Major subject	:	Soil Science
Minor subject	:	Chemistry
Degree to be awarded	:	M.Sc. (Soil Science)
Major advisor and designation		Dr. B.S. Sekhon Editor Research
Year of award of degree	:	2019
Total pages in thesis	:	98+ Annexures (vi)+ Vita
Name of the university	:	Punjab Agricultural University, Ludhiana-141004, Punjab, India.

ABSTRACT

The present study titled "Development of spectral models for assesing soil fertiity status in various agroecological sub-regions of Punjab" was carried out in five agroecological sub-regions, namely, Sub-mountainous Siwalik Hills (SSH), Northeastern Undulating (NEU), Piedmont and Alluvial Plain (PAP), Central Alluvial Plain (CAP) and Southwestern Alluvial Plain (SWAP). The objectives were to develop visible-near infrared (vis-NIR) models for predicting soil properties and to examine portability of spectral models across regions. GPS based soi fertility maps were prepared by using inverse distance weighted (IDW) interpolation and classification. Highest soil organic carbon (SOC) and sand content (mean 0.77 % and 74.2 %, respectively) in 0-15cm soil were observed in SSH, whereas highest nitrogen, phosphorous and, manganese contents (mean value 104.5 kg ha^{-1}, 40.8 kg ha^{-1}, 7.8 ppm, respectively) were found in NEU. The largest potassium, copper and silt levels (mean 173.1 kg ha^{-1}, 5.21 ppm, 22.4 %, respectively) were reported in PAP. The CAP represented highest calcium carbonate content (mean 1.36 %), while highest pH, electrical conductivity, iron and zinc (mean value 7.93, 0.95 dS/m, 29.5 ppm and 3.2 ppm, respectively) levels were observed in SWAP region. Spectral signatures in vis-NIR range were collected using ASD Spectroradiometer. Partial least square regression (PLSR) was emlpoyed to develop spectral models for assessing soil fertility attributes. Various model evaluation indices e.g. root mean square error (RMSE), ratio of performance to deviation (RPD) and coefficient of determination (R^2) were used to evaluate predictive performance of the PLSR models. The RMSE, RPD and R^2 values varied between 0.05 - 0.39, 0.32 – 1.73 and 0.05-0.69, respectively in SSH for various soil properties. Similarly, in NEU RMSE, RPD and R^2 values ranged between 0.05 – 0.39, 0.93 – 2.15 and 0.12 - 0.82, respectively. In PAP region, the RMSE, RPD and R^2 values varied between 0.07-2.90, 0.32-1.6 and 0.05-0.72, respectively. CAP region showed RMSE, RPD and R^2 values ranging between 0.03-0.43, 0.62-1.5 and 0.03-0.64, respectively. In case of SWAP region, RMSE, RPD and R^2 values ranged between 0.03-0.73, 0.72-1.5 and 0.05-0.68, respectively. Moreover, in common model, RMSE, RPD and R^2 values varied between 0.19-3.81, 0.16-0.98 and 0.01-0.40. A comparison among models suggested about their non-portability across regions. Based upon various performance indices, the $CaCO_3$ and sand predictions were reliable, whereas prediction of properties like SOC and Olsen-P was moderately reliable. The pH and EC, however, could not be predicted much accurately. The study suggested about exploring better statistical tools to enhance prediction performance of spectral models.

Key words: Model portability, Partial least squares regression, Soil fertility maps, Spectral library, Vis-NIR spectroscopy.

_____ _____

Signature of Major Advisor **Signature of the student**

ਖੋਜ ਦਾ ਸਿਰਲੇਖ	:	ਪੰਜਾਬ ਦੇ ਵੱਖ-ਵੱਖ ਖੇਤੀ ਮੌਸਮੀ ਖਿੱਤਿਆਂ ਵਿੱਚ ਮਿੱਟੀ ਦੇ ਉਪਜਾਊਪਨ ਦਾ ਆਂਕਲਨ ਕਰਨ ਲਈ ਸਪੈਕਟਰਲ ਮਾਡਲਾਂ ਦਾ ਨਿਰਮਾਣ
ਵਿਦਿਆਰਥੀ ਦਾ ਨਾਂ ਅਤੇ ਦਾਖ਼ਲਾ ਨੰਬਰ	:	ਸ਼ਹੀਦ ਗੜਨਾਇਕ (ਐੱਲ-2017-ਏ-155-ਐੱਮ)
ਪ੍ਰਮੁੱਖ ਵਿਸ਼ਾ	:	ਭੂਮੀ ਵਿਗਿਆਨ
ਸਹਿਯੋਗੀ ਵਿਸ਼ਾ	:	ਰਸਾਇਣ ਵਿਗਿਆਨ
ਮੁੱਖ ਸਲਾਹਕਾਰ ਦਾ ਨਾਂ ਅਤੇ ਅਹੁੱਦਾ	:	ਡਾ. ਬੀ.ਐੱਸ. ਸੇਖੋਂ ਸੰਪਾਦਕ ਖੋਜ
ਡਿਗਰੀ	:	ਐੱਮ.ਐੱਸ-ਸੀ.
ਡਿਗਰੀ ਨਾਲ ਸਨਮਾਨਿਤ ਕਰਨ ਦਾ ਸਾਲ	:	2019
ਖੋਜ ਪੱਤਰ ਵਿੱਚ ਕੁੱਲ ਪੰਨੇ	:	98 + ਅੰਤਿਕਾਵਾਂ (vi) + ਵੀਟਾ
ਯੂਨੀਵਰਸਿਟੀ ਦਾ ਨਾਮ	:	ਪੰਜਾਬ ਖੇਤੀਬਾੜੀ ਯੂਨੀਵਰਸਿਟੀ, ਲੁਧਿਆਣਾ-141 004 ਪੰਜਾਬ, ਭਾਰਤ ।

ਸਾਰ-ਅੰਸ਼

ਮੌਜੂਦਾ ਅਧਿਐਨ ''ਪੰਜਾਬ ਦੇ ਵੱਖ-ਵੱਖ ਖੇਤੀ ਮੌਸਮੀ ਖਿੱਤਿਆਂ (ਸਬ ਰਿਜਨ) ਵਿੱਚ ਮਿੱਟੀ ਦੇ ਉਪਜਾਊਪਨ ਦਾ ਆਂਕਲਨ ਕਰਨ ਲਈ ਸਪੈਕਟਰਲ ਮਾਡਲ ਦਾ ਨਿਰਮਾਣ'' ਸਿਰਲੇਖ ਹੇਠ ਪੰਜ ਖੇਤੀ ਮੌਸਮੀ ਖਿੱਤਿਆਂ ਨੀਮ ਪਹਾੜੀ ਸ਼ਿਵਾਲਿਕ ਖੇਤਰ (SSH), ਉੱਤਰੀ ਪੂਰਵੀ ਉੱਬੜ ਖਾਬੜ (NEU), ਪੀਡਮੋਂਟ ਅਤੇ ਜਲੌੜ ਪਲੈਨ (PAP), ਕੇਂਦਰੀ ਜਲੌੜ ਮੈਦਾਨ (CAP), ਦੱਖਣੀ-ਪੱਛਮੀ ਜਲੋੜ ਮੈਦਾਨ (SWAP) ਵਿੱਚ ਕੀਤਾ ਗਿਆ। ਇਸ ਅਧਿਐਨ ਦਾ ਉਦੇਸ਼ ਮਿੱਟੀ ਦੀਆਂ ਵਿਸ਼ੇਸ਼ਤਾਵਾਂ ਦਾ ਅਨੁਮਾਨ ਲਾਉਣ ਲਈ VIS-NIR ਮਾਡਲ ਬਣਾਉਣਾ ਅਤੇ ਇਹਨਾਂ ਖਿੱਤਿਆਂ ਵਿੱਚ ਮਾਡਲ ਦੀ ਪੋਰਟੇਬਿਲਿਟੀ ਦਾ ਨਿਰੀਖਣ ਕਰਨਾ ਸੀ। SSH ਵਿੱਚ 0-15 cm ਤਹਿ ਵਿੱਚ ਸਭ ਤੋਂ ਵੱਧ ਜੈਵਿਕ ਕਾਰਬਨ ਅਤੇ ਰੇਤ (0.77%, 74.2% ਕ੍ਰਮਵਾਰ) ਪਾਇਆ ਗਿਆ ਜਦਕਿ ਨਾਈਟ੍ਰੋਜਨ, ਫਾਸਫੋਰਸ ਅਤੇ ਮੈਂਗਨੀਜ ਮਾਤਰਾ (ਕ੍ਰਮਵਾਰ 104.5 kg/ha, 40, 81 kg/ha, ਅਤੇ 7.8 ppm) NEU ਵਿੱਚ ਪਾਈ ਗਈ। ਸਭ ਤੋਂ ਵੱਧ ਪੋਟਾਸ਼ੀਅਮ, ਕਾਪਰ ਅਤੇ ਸਿਲਟ (ਕ੍ਰਮਵਾਰ 173.1 kg/ha, 5.21 ppm, 22.4%) PAP ਵਿੱਚ ਪਾਈ ਗਈ। PAP ਵਿੱਚ ਪਾਈ ਗਈ। CAP ਵਿੱਚ ਸਭ ਤੋਂ ਵੱਧ ਕੈਲਸ਼ੀਅਮ ਕਾਰਬੋਨੇਟ ਮਾਤਰਾ (1.36%) ਜਦਕਿ ਸਭ ਤੋਂ ਵੱਧ pH, ਚਾਲਕਤਾ, ਲੋਹਾ ਅਤੇ ਜਿੰਕ (ਕ੍ਰਮਵਾਰ 7.93, 0.95 dS m^{-1}, 29.5 ppm ਅਤੇ 3.2 ppm) SWAP ਖੇਤਰ ਵਿੱਚ ਪਾਏ ਗਏ। GPS ਅਧਾਰਿਤ ਮਿੱਟੀ ਉਪਜਾਊਪਨ ਮੈਪ IDW ਇੰਟਰਪੋਲੇਸ਼ਨ ਅਤੇ ਵਰਗੀਕਰਨ ਰਾਹੀਂ ਬਣਾਏ ਗਏ। ਸਪੈਕਟਰਲ ਸਿਗਨੇਚਰ VIS-NIR ਵਿੱਚ ASD ਸਪੈਕਟੋਰੇਡੀਓਮੀਟਰ ਵਰਤ ਕੇ ਇਕੱਠੇ ਕੀਤੇ ਗਏ। ਮਿੱਟੀ ਉਪਜਾਊਪਨ ਦਾ ਆਂਕਲਨ ਕਰਨ ਲਈ ਸਪੈਕਟਰਲ ਮਾਡਲ ਬਣਾਉਣ ਲਈ PLSR ਦੀ ਵਰਤੋਂ ਕੀਤੀ ਗਈ। ਕਈ ਵੱਖ-ਵੱਖ ਮਾਡਲ ਨਿਰੀਖਣ ਸੂਚਕਾਂਕ ਜਿਵੇਂ RMSE, RPD ਅਤੇ R^2, NEU, RMSE, RPD ਅਤੇ R^2 ਵਿੱਚ ਵੱਖ-ਵੱਖ ਸਨ ਅਤੇ ਕ੍ਰਮਵਾਰ 0.05-0.39, 0.32-1.73 ਅਤੇ 0.05-0.69 ਦੇ ਵਿਚਕਾਰ ਸੀ। ਇਸੇ ਤਰ੍ਹਾਂ PAP ਖੇਤਰ ਵਿੱਚ RMSE, RPD ਅਤੇ R^2 ਮਾਨ 0.07-2.90, 0.32-1.6 ਅਤੇ 0.05-0.72 ਸੀ। CAP ਖੇਤਰ ਵਿੱਚ RMSE, RPD ਅਤੇ R^2 ਕ੍ਰਮਵਾਰ 0.03-0.73, 0.75-1.5 ਅਤੇ 0.05-0.68 ਸੀ। ਇਸਤੋਂ ਇਲਾਵਾ ਸਾਂਝੇ ਮਾਡਲ ਵਿੱਚ RMSE, RPD ਅਤੇ R^2 ਮਾਨ 0.19-3.81, 0.16-0.98 ਅਤੇ 6.01-0.40 ਦੇ ਵਿਚਕਾਰ ਸੀ। ਮਾਡਲਾਂ ਵਿਚਕਾਰ ਫਰਕ ਇਹਨਾਂ ਦੇ ਖਿੱਤਿਆਂ ਵਿੱਚ ਪੋਰਟੇਬਿਲਿਟੀ ਨਾ ਹੋਣ ਨੂੰ ਦਰਸਾਉਂਦਾ ਹੈ। ਕਈ ਵੱਖ-ਵੱਖ ਪ੍ਰਦਰਸ਼ਨ ਸੂਚਕਾਂਕ ਦੇ ਅਧਾਰ ਤੇ $CaCO_3$ ਅਤੇ ਰੇਤ ਦਾ ਆਂਕਲਨ ਭਰੋਸੇਯੋਗ ਸੀ ਜਦਕਿ ਜੈਵਿਕ ਕਾਰਬਨ ਅਤੇ ਫਾਸਫੋਰਸ ਦੀ ਭਰੋਸੇਯੋਗਤਾ ਦਰਮਿਆਨੀ ਸੀ। pH ਅਤੇ EC ਹਾਲਾਂਕਿ ਭਰੋਸੇਯੋਗਤਾ ਆਂਕਲਿਤ ਨਹੀਂ ਕੀਤੀਆਂ ਜਾ ਸਕੀਆਂ। ਅਧਿਐਨ ਦਰਸਾਉਂਦਾ ਹੈ ਕਿ, ਸਪੈਕਟਰਲ ਮਾਡਲ ਦੇ ਪ੍ਰਦਸ਼ਨ ਦੇ ਆਂਕਲਨ ਨੂੰ ਵਧਾਉਂਦਾ ਸਾਂਖਿਅਕੀ ਟੂਲ ਲੱਭਣ ਦੀ ਲੋੜ ਹੈ।

ਮੁੱਖ ਸ਼ਬਦ: ਮਾਡਲ ਪੇਸ਼ੀਨਗੋਈ, ਮਿੱਟੀ ਉਪਜਾਊਪਨ, VIS-NIR ਸਪੈਕਟਰੋਸਕੋਪੀ ।

ਮੁੱਖ ਸਲਾਹਕਾਰ ਦੇ ਹਸਤਾਖਰ **ਵਿੱਦਿਆਰਥੀ ਦੇ ਹਸਤਾਖਰ**

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER-I

INTRODUCTION

In order to achieve potential yields and to sustain soil health, continuous assessment of soil properties is the need of the hour. The ideal soil property assessment methods need to duly consider soil spatial variability and should be rapid, reproducible and cost effective. However, conventional soil testing protocols involve sampling a field at a few randomly selected locations. Conventional soil testing laboratories employ chemical methods to analyze these samples. These methods are laborious, time consuming and have environmental implications. Further, most of the soil testing laboratories do not test for physical and microbiological properties. A vivid example of the limitations of conventional soil testing laboratories comes from the huge costs and manpower involved in mega Soil Health Card Scheme launched by Government of India during 2015. The scheme aims at providing soil health cards to about 140 million landholdings in the country at a two-year interval. There are more than 1400 soil testing laboratories in the country. However, their annual capacity can cater to only 19.5 million soil samples. Landholdings are of variable size and one or two samples drawn from one farm unit cannot account for the spatial variability of soil fertility status at the field scale.

Soil testing methods that are rapid, capacious, relatively inexpensive, and which are able to duly consider spatial variability are required to ensure site-specific nutrient management for the larger cause of popularization of precision agriculture in the country. Precision agriculture is a site-specific management method explicitly taking within-field variation of soil and crop into consideration (Adamchuk *et al* 2004). It is a technology-driven system that provides spatial and temporal information (where, how much, and when to apply) about the application of farm inputs such as tillage, irrigation, fertilizers, pesticides, etc., in a field (Gebbers and Adamchuk 2010). Visible-Near Infrared (VIS-NIR, 350-2500nm) diffuse reflectance spectroscopy holds a great promise in honouring these conditions and can thus ensure effective site-specific nutrient management by providing high-density data at less cost. This technique involves directing the radiation in the VIS-NIR region onto the soil sample. Vis-NIR spectroscopy has grown as an analytical method due to its cost-efficiency, ease of handling, rapidity, minimal sample preparation, and the development of chemometrics (Davies 2005, Viscarra Rossel and Webster 2012). Depending upon the soil composition, radiation of a given wavelength gets absorbed, reflected or scattered to different degrees. The dependence of the extent of absorbance or reflectance on the soil composition forms the basis of prediction of soil property status. Spectral signature results from electrical transitions of atoms (vibrational stretching and bending of functional groups). A single scan of a soil sample can be used for prediction of multiple properties. It is possible to derive quantitative

information from the spectral information in addition to the qualitative information (Owen 2000). The spectral reflectance properties are dependent on soil mineral and organic constituents (biogeochemical components), particle size, roughness (geometrical optical scattering) and moisture content of the surface (Ben-Dor and Banin 1995). Different properties like pH, EC, phosphorous, carbon, nitrogen (Yang *et al* 2012), potassium, calcium, magnesium, sulphur, boron, sodium, manganese, nickel, zinc, chloride, copper (Viscarra Rossel *et al* 2006) have been reported to be successfully predicted by using this technique. Diffuse reflectance spectra of soil in the vis-NIR region lack specificity due to overlapping spectral behavior of soil constituents (Stenberg *et al* 2010). Thus, in order to extract the dependence of spectral behavior on soil composition suitable mathematical or statistical techniques need to be employed. Multivariate calibrations are often resorted to in order to obtain meaningful relationships between soil constituents or properties and spectra (Martens and Naes 1989). The most common calibration methods involve linear regression methods (Ben-Dor and Banin 1995, Dalal and Henry 1986) like partial least squares regression (PLSR), principal components regression (PCR) and stepwise multiple linear regression (SMLR). The PLSR technique finds the most common usage due to its characteristic of accounting for the variability in both response and predictor variables. Also, it is more interpretable and computationally faster (Stenberg *et al* 2010). Besides these linear regression techniques, certain data mining techniques like neural networks (NN, Daniel *et al* 2003), multivariate adaptive regression splines (MARS, Shepherd and Walsh 2002), boosted regression trees (Brown *et al* 2006), support vector machines (Mahmood and Babel 2013), and random forest (Hengel *et al* 2015) are finding increasing usage.

For accurate calibration and internal validation, an independent validation data set is produced to check the performance of the model. The calibration data set must consider the full variability expected in the full sample set and future unknown data sets where the validation data sets are independent. Besides, to handle spectral data selection of a suitable statistical approach, and proper choice of calibration and validation datasets is of vital importance. The choice of these sets affects the prediction performance of the developed models. General calibration models developed from diverse soil types may yield better prediction accuracies (Mahmood and Babel 2013). On the other hand, however, more generalized models based on datasets drawn from regional, national, continental or global scale may not accurately predict at local or field scale (Guerrero *et al* 2010).

Different agroecological regions are usually inhabited by various types of soils. For instance, five agroecological sub-regions of Punjab state are inhabited by around 17 benchmark soils (Raj-Kumar *et al* 2008). Also, these regions generally tend to support specific land use types. Keeping this in view, VIS-NIR models can be developed at different scales (agroecological sub-regions) and subsequently examined for their applicability across

various scales.

In this context, the present study was undertaken with the following objectives:

1. To develop VIS-NIR models for assessing soil fertility status of various agroecological sub-regions of Punjab state.
2. To examine the portability of spectral models across various agroecological sub-regions.

CHAPTER-II

REVIEW OF LITERATURE

Suitable literature is reviewed under the following headings:

2.1 Importance of reflectance spectroscopy in soil analysis

2.2 Spectral behaviour of soil physical attributes

2.3 Spectral behaviour of soil chemical attributes

2.4 Spectral behaviour of soil biological attributes

2.5 Different spectroscopic techniques

2.6 Pre-processing of spectral data

2.7 Statistical modelling techniques

2.8 Spectral library approach

2.1 Importance of reflectance spectroscopy in soil analysis

Soil is a heterogenous and complex system. It is a fundamental natural resource on which people rely for their production of food, fiber, and energy. Given the importance of soil, there is a need for regular monitoring to detect changes in its status so as to implement appropriate management in the event of degradation. It is not easy to assess physical, chemical and biological processes which are occurring inside the soil. Increasing concerns about soil health throughout the world became an issue for regulating nutrient inputs on the basis of soil testing. Adoption of site-specific nutrient management as part of precision agriculture can help to sustain soil health (Cassman 1999). Soil Health Card Scheme was implemented in India to provide soil test report to each farmer by 2020 (Govt. of India, 2015). However, inadequacy of soil testing laboratories in terms of count and capacity can come in the way of successful outcomes of this scheme. Traditional chemical-based soil analysis is laborious, hazardous and time consuming. As soil analysis by chemical method is time consuming and laborious, soil test-based fertilizer recommendation based upon conventional method has become a slow process. Spatial variability is another major concern (Geypens *et al* 2013). Modern spatial tools such as remote sensing, GPS (Global Positioning System) (Kennedy, 2002) and GIS (Geographical Information System) are gaining importance in precision farming (Burrough 1986) as these can help gauge spatial variability. Diffuse soil spectral reflectance is another promising sensor technology that can overcome the shortcomings of conventional soil testing method (Nanni and Dematte 2006), with respect to spatial variability and rapidity of analysis and can thus enhance adoption of precision farming. Visible near-infrared (VIS-NIR) reflectance spectroscopy can also be a viable alternative to ensure rapid and environmentally safe soil test results (Shepherd and Walsh 2007).

A wide range of research is being conducted across the world to develop proximal

soil sensing (PSS) techniques for their use in various applications. Sensors are highly advanced instruments providing quantitative results and can be regarded as time saving and economical than traditional laboratory analyses. These devices are portable, accurate and intelligent. Many devices may be used for proximal soil sensing (PSS), e.g. ion-sensitive field effect transistors (ISFETs) to measure soil pH and soil nutrients, or portable visible near-infrared (vis-NIR) spectrometers to measure various soil parameters like organic carbon content and mineral composition (Sahoo *et al* 2015).

Different soil factors like soil types, parent materials, soil and environmental factors, (e.g. water content, temperature, humidity, organic matter, topography and soil color) affect the performance of different sensors (Pirie *et al* 2005, Mulder *et al* 2011).

Basically, soil spectrum is generated by directing radiation having all relevant frequencies in the particular range onto the sample. Radiation tends to cause individual molecular bonds to vibrate, either by bending or stretching. These types of vibrations lead to absorption of light to different degrees with a specific energy quantum corresponding to the difference between two energy levels. As it is known the energy quantum is directly proportional to frequency, the produced characteristic shape can be analyzed (Stenberg *et al* 2010) and correlated with soil parameters. In the near infrared (NIR) region and the mid-infrared (MIR) region, the fundamental vibrations result in overtones and/or combinations. In the visible (vis) range (400–780 nm), absorption bands related to soil color are due to electron excitations, which assist the measurement of soil organic matter content (SOM) and moisture content However, in the NIR range, the overtones of OH and overtones and/or combinations of C-H. C-H, C-H, C-C, O-H minerals, and N-H are important for the detection of SOM, moisture content, clay minerals, and nitrogen content (Mouazen *et al* 2010).

2.2 Spectral behavior of soil physical attributes

We can assess the quality and function of soil through chemical and physical laboratory analysis. Rapid determination soil physical parameters is a crucial factor for soil environment monitoring and for identifying suitable land management practices. As the prediction of physical properties has relationship with some components *viz.* quartz, various types of clays, organic matter, carbonates and oxides, which are infrared active, these spectroscopic techniques are used for this purpose.

2.2.1 Soil texture

Relative proportion of sand**,** silt and clay in soil influences some crucial processes like soil water dynamics and soil porosity distribution (i.e. soil air and water), plant nutrition, microbial activity and nutrient mobility. Soil texture is an important variable because it plays a key role in determining soil degradation and water transport processes, controlling soil quality and its productivity, particularly.

Sand content is very well predicted by vis-NIR and MIR techniques (Curcio *et al*

2013). However, Cozzolino and Moron (2003) observed that sand and silt can be predicted very well by modified PLSR method (R^2 = 0.70, R^2 = 0.80 respectively). Also, predictability of sand and silt was very good (R^2 = 0.60, R^2 = 0.84) when PCR method was used (Chang *et al* 2001). Clay fraction has also been reported to be relatively accurately predicted (RMSE=5.8%, R^2=0.87) by reflectance spectroscopy tools (Viscarra Rossel *et al* 2006). However, clay can be predicted very well by using methods like PCR, PLSR and MARS as demonstrated by high R^2 values 0.87, 0.67, 0.86, and 0.78, respectively. (Janik and Skjemstad 1995, Chang *et al* 2001, Walvoort and McBratney 2001 and Shepherd and Walsh 2002).

2.2.2 Soil structure

Soil structure is the one of the most important physical parameters in relation to plant growth as it influences soil processes and properties like aeration, water retention, infiltration, compression and compaction. Askari *et al* (2013) found moderate prediction capability (RPD from 1.55-1.99, $R^2 \geq$ 0.64) of aggregate size distribution and penetration resistance. The geometric mean diameter (GMD) and the median aggregate size parameter showed excellent predictions, with ratio of performance deviation (RPD) values ranging from 1.99 to 2.28. In other studies (Sarathjith *et al* 2014; Janik *et al* 2009)., the RPD value for prediction of soil structure ranged from 1.36 to 1.72 (R^2 value ranges from 0.67-0.79 in MIR region), suggesting moderate prediction.

2.2.3 Soil water content

In the context of spectral analysis of soil properties, soil water content (WC) affects the accuracy of the visible (VIS) and near infrared (NIR) spectroscopic measurement of other soil properties like C, N, and other nutrients. A study was conducted to reduce the contribution of WC to VIS-NIR spectra by classifying soil spectra according to different WC groups (Morellos *et al* 2016). This type of classification has implications in improving the accuracy of prediction of other soil properties with calibration models established separately for each group of WCs. Partial least squares regression analysis and factorial discriminant analysis (FDA) were applied to the vis-NIR spectra to quantify WC and classify spectra into different WC groups, respectively. The PLSR for sample set of the single-field provided better estimation of WC (R^2= 0.98) than for the multiple-field sample set (R^2 = 0.88). Due to large variability in the multiple-field sample set, soils were successfully classified into three WC groups (Mouazen *et al* 2006).

Soil moisture is an important parameter in determining different moisture potentials like field capacity, permanent wilting point and hygroscopic moisture content, etc. Water is available to plant with varying degree of ease at these potential levels. Vis-NIR spectroscopy has been reported to predict moderately accurate (R^2 =0.60, RMSE=0.036) field capacity (Disla *et al* 2014). Also, Shibusawa *et al* (2001) reported that water content can be well predicted (R^2 = 0.66) by Vis-NIR spectroscopy.

2.2.4 Soil minerals

Viscarra Rossel *et al* (2009) compared mineral compositions quantified by qualitative X-ray diffraction (XRD) analysis with soil reflectance spectra and with the spectra of pure minerals. Continuous removal spectra techniques are used to quantify the soil mineral content. The analysis of soil mineral composition by vis–NIR was in good agreement with the results of this method and XRD analysis. The vis–NIR technique did not require sample preparation so it was less laborious than conventional XRD (Viscarra *et a*l 2010).

2.2.5 Specific surface area

PLS models developed by Knadel *et al* (2018) had good predictive ability for specific surface area as indicted by RPIQ (ratio of performance to interquartile range) of 1.7and for clay content (RPIQ = 1.6). The important wavelengths responsible for the predictions of SSA and clay were indicative of the organo-mineral content and interactions among them. Moreover, it included spectral response from iron oxides and minerals in relation with organic matter. It is reported to be due to masking effect of the non-complexed organic carbon on the mineral phases of some soils. However, Ben-Dor and Banin (1995) reported satisfactory prediction of specific surface area (R^2 = 0.70).

2.3 Spectral behaviour of soil chemical attributes

2.3.1 pH

Soil reaction (pH) is a key soil property which has a great role in nutrient dynamics and nutrient availability. Mohamdimonavar (2016) reported excellent prediction of pH by reflectance spectroscopy (RPD≥3.5, R^2 c ≥ 0.91). Similarly, in vis-NIR range, pH was predicted very well (R^2 = 0.70) (Shepherd and Walsh 2002; MARS method). However, Reeves and McCarty (2001) showed excellent prediction of pH in NIR region by using PLSR method (R^2 = 0.74).

2.3.2 Electrical conductivity

Electrical conductivity is a widely used parameter for describing soil salinity. However, Islam *et al* (2003) reported very poor prediction of EC (R^2 = 0.10) using principal components regression (PCR) method, while Janik et al (1998) showed a moderate prediction of EC (R^2 = 0.55) using PLSR method. Kodaira *et al* (2013), however, reported that EC was poorly predicted by both MIR (R^2 = 0.26) and NIR spectroscopy (R^2 = 0.57) and similar accuracy (R^2=0.60) was reported for vis-NIR. EC (Kodaira *et al* 2013). Electrical conductivity has also been reported to be well predicted (R^2 = 0.65) using SMLR method (Shibusawa *et al* 2001).

2.3.3 Soil organic carbon

The quality and quantity of soil organic carbon (SOC) and organic matter (SOM) determine the soil fertility status. Infrared analysis (near infrared range 780-2500 nm) is very well suited for SOM analysis because it is sensitive to the C-H, C-O and C-N functional

groups, present in the organic matter (Gomez *et al* 2008). Soil organic matter is most frequently estimated by vis-NIR (Stenberg *et al* 2010, Minansy *et al* 2011), although the MIR region has also been reported to be employed for C prediction by PLSR (Janik *et al* 1995). Du *et al* (2009) reported that MIR predictions for inorganic C, total C and SOM performed somewhat better (R^2=0.97, 0.93 and 0.94, respectively) in MIR vis-NIR (R^2 =0.83, 0.89 and 0.86).

2.3.4 Cation exchange capacity

As plant nutrient uptake depends upon exchange of ions, CEC is an important factor. MIR has been effective in predicting CEC with R^2 value of 0.88 (Janik *et al* 1998). By using vis-NIR, CEC can be predicted accurately. Good models (Mohamdimonavar 2016 and Mondal 2017) were obtained with moderate capability for prediction of particle size and exchangeable cations (RPD from 1.5 to 1.99, R^2 c $\geq$ 0.68 and RPD value 2.20 and R^2 value 0.78, respectively).

2.3.5 Primary nutrients

For plant growth and development, primary nutrients have great importance. Also, they are required by plant in bulk amount. Their rapid and timely assessment is the need of the hour. Diffused reflectance spectroscopy can be used very effectively for assessment of nitrogen, phosphorous and potassium.

Nitrogen is related with soil moisture content, soil organic matter, clay content, etc. which in turn influence the reflectance behaviour. Total nitrogen is very well predicted both by MIR (R2=0.90) and by Vis-NIR (R2=0.86, RMSEP = 0.071 and RPD = 1.96) as reported by Morellos *et al* (2016). However, Chang *et al* (2001) showed the good predictive ability of nitrogen by using PCR method (R^2 = 0.72). Similarly, total nitrogen was very well predicted (R^2 = 0.0.86) (Janik *et al* 1998, PLSR).

Phosphorous along with potassium has been reported to be estimated accurately by vis-NIR spectroscopy (Hu *et al* 2016). Total P is well predicted (R^2=0.75) in vis-NIR region (Minasny *et al* 2006). Mouazen and Kuang (2016) reported that both sorbed P and plant available P can be predicted well with Vis-NIR (with respective R^2 values 0.65 and 0.71). Daniel *et al* (2003) reported better prediction (R^2 = 0.81) of available phosphorus.

Potassium is strongly related to clay minerals like illite, montmorillonite and vermiculite which are spectrally active (Hu *et al* 2016). Exchangeable K is reported to be predicted (R^2 = 0.71) adequately by vis-NIR spectroscopy (Bellon-Maurel *et al* 2011). Similarly. Daniel *et al* (2003) demonstrated, good prediction of available potassium (R^2 = 0.80).

2.3.6 Secondary nutrients

Calcium and magnesium are required by plants in moderate amounts. Mayrink *et al* (2019) showed that PLSR models based on the diffuse reflectance of ion-exchange resins are

reliable for the fast and accurate prediction of these ions. R^2 value was found to be 0.85 and 0.88 for calcium and magnesium, respectively. Disla *et al* (2014) reported calcium (exchangeable) and magnesium (exchangeable) and calcium (total) and magnesium (total) coefficient of determination (R^2) values 0.80, 0.83, 0.78 and 0.61, respectively indicating thereby good prediction of these parameters.

Sulfur (S) is a secondary nutrient whose deficiency is a concern in Indian soils. It is associated with minerals and iron/aluminum oxides. Disla *et al* (2014) showed that total S can be predicted with good accuracy (R^2=0.84) by using vis-NIR spectroscopy.

2.3.7 Calcium carbonate

Viscarra Rossel *et al* (2006) demonstrated that there were significant relationships between clay and calcium carbonate ($CaCO_3$) contents and CR values computed were respectively at 2206 nm and 2341 nm as observed from reflectance measurements at the laboratory level with an ASD spectrophotometer. The coefficient of determination values ranged from 0.69 to 0.95 for $CaCO_3$. Also, Lagacherie *et al* (2008) reported good prediction of calcium carbonate content (R^2 = 0.61 to 0.86).

2.3.8 Micronutrients

Micronutrients are some essential nutrients that are required in very trace quantities for growth of plants and microorganisms. Mondal (2017) reported that micronutrient elements can be predicted with moderate accuracy by using vis-NIR spectroscopy (Reeves and Smith 2009). However, Janik *et al* (1998), Chang *et al* (2001) and Islam *et al* (2003) demonstrated poor to moderate prediction of micronutrients.

2.4 Spectral behaviour of soil biological attributes

Soil microbes play a very crucial role in major soil processes like nutrient mineralization, nutrient transformation and various physiological activities (Dick *et al* 1996). Assessment of soil microbiological properties is highly essential for monitoring soil health. Mondal *et al* (2017) reported RPD values 1.15 and 1.13 for dehydrogenase activity (DHA) and microbial biomass carbon (MBC). Poor prediction of model as opposed to substantial evidence in literature (Chang *et al* 2001) can be ascribed to range of factors like limited variability in spectra, relatively smaller sample size, use of entire spectra instead of selective regions. In addition, the samples collected under natural environment as opposed to those generated under controlled conditions may affect prediction accuracy. Reeves *et al* (2000) also reported poor prediction for dehydrogenase and urease.

Different spectral ranges like vis-NIR, NIR and MIR can generate good spectral models for assessing the microbial biomass carbon (Stenberg *et al* 2010) as evidenced by coefficient determination (R^2) values (R^2=0.82-0.84). Vis-NIR spectroscopy can be effectively used for estimating enzymatic activities (Dick *et al* 2013). Also, Soriano-Disla *et al* (2014) showed, soil biological properties can be predicted through vis-NIR spectroscopy.

2.5 Different spectroscopic techniques

2.5.1 Vis-NIR spectroscopy

Vis-NIR spectroscopy infers compositional and structural information of molecules at spectral wavelengths of 350–2500 nm. This spectroscopic method is a linear fusion of spectral signatures of its different components (Ben-Dor *et al* 2015). A small change in the constituents can make different spectra which can be captured by reflectance spectroscopy (Shepherd and Walsh 2002). Generally, different minerals like goethite, hematite, etc. (iron oxide bearing minerals) are strongly associated with visible spectroscopy. Similarly, soil organic matter (SOM) has a strong absorption band in 350 to 780 nm range (Stenberg 2010). NIR (Near-infrared) spectroscopy was used for faster moisture analysis in grain for first time (Ben-Gera and Norris 1968). Normally, NIR spectrum has weak overtones, combinations and analytical ability that depends on the repetitive and broad absorption with vibrational functional groups like C-H, O-H and N-H bonds (Lee *et al* 2009, 2010, Li *et al* 2002 and Xuemei and Jianshe 2013). In addition, it includes heavier groups like N-O, C-C, C-O, etc. and inorganic Fe-O, Al-O and Si-O bonds in minerals (Du and Zhou 2009). Moreover, NIR spectrum is affected by shape, size of the particles, pore space and arrangement of the particles which can interfere with the path length of incidental light. Spectrally active components of soil like soil carbon, minerals and functional groups play key roles in determination of soil properties (known as chemical chromophores). Similarly, soil structure and soil texture act as physical chromophores (Nduwamungu *et al* 2009).

A partial least square regression (PLSR) tool can be used to correlate the visible and near infrared spectra to estimated soil attributes (Lu *et al* 2013). However, Chang *et al* (2001) suggested categorizing prediction ability based on RPD (ratio of performance to deviation) and R-square values. In the proposed classification, excellent prediction i.e. category A has RPD > 2.0 and R^2 = 0.80-1.00 and category B has RPD 1.4-2.0 and R^2=0.5-0.80. They classified prediction for soil properties like total nitrogen, total carbon, silt, sand, CEC, Mehlich-3 extractable Ca, basal respiration rate, 1.5 MPa water; wet aggregation measures, clay, pH, biomass carbon, total respiration rate, Mehlich-3 extractable metals (iron, potassium, magnesium and manganese) under category B (RPD = 1.4-2.0, R^2 = 0.50-0.80) whereas category C (RPD < 1.4, R^2 <0.50) included NH_4OAc extractable Na and Mehlich-3 extractable phosphorous, zinc and copper and wet aggregates (2 and 0.25 mm). Around 4000 samples were tested to judge the accuracy of vis-NIR spectral models for soil characterization globally which reflected strong prediction for minerals like kaolinite, smectite, CEC, soil organic and inorganic carbon, iron and clay content (Brown *et al* 2006). Some other researcher found good prediction quality in estimation of heavy metals, macro and micro nutrients (Bertrand *et al* 2002, Cozzolino and Moron 2003) and C:N ratio and biological properties (Chodak *et al* 2007, Ludwig *et al* 2002).

In soil science research, diffuse reflectance spectroscopy has been used since the 1950s and 1960s (Bowers and Hanks 1965, Brooks 1952). However, its real importance became visible only during past 20 years with the establishment of good relationships between chemometrics and multivariate statistical techniques.

2.5.2 MIR spectroscopy

Mid infra-red (MIR) spectra can be divided into four regions like finger region (e.g. O-Si-O stretching and bending) from 1500-600 cm^{-1}, double bond (e.g. C=O, C=C and C=N) from 2000-1500 cm^{-1}, triple bond (e.g. C ≡O, C ≡C and C ≡N) from 2500-2000 $cm^{-}1$ and X-H stretching (e.g. O-H stretching) from 4000-2500 cm^{-1}. The mid infrared (MIR) models have been developed with higher prediction accuracy (R^2 = 0.62–0.85) for pH, organic carbon, clay, sand, CEC, and exchangeable Ca and Mg (Urselmans *et al* 2010).

Janik *et al* (1998) reported that pH, lime requirement, OC, CEC, clay, silt and sand contents, P and EC can be better predicted with MIR. The NIR could produce more accurate estimations for exchangeable Al and K. There were only substantial improvements in predictions of clay, silt and sand content using the combined VIS–NIR–MIR range. This work shows the potential of diffuse reflectance spectroscopy using the VIS, NIR and MIR for more efficient soil analysis and the acquisition of soil information. Overtones and combination bands, which also occur in the MIR overlap making qualitative and quantitative interpretations in the visible and NIR more difficult. The models which were developed using combined spectra were also found to predict pH, organic carbon, clay, sand, CEC, and exchangeable Ca and Mg with acceptable accuracy (R^2 = 0.59–0.79). Also, other researchers studied soil reflectance properties by using MIR spectra to compare with vis-NIR spectroscopy for better results. (Reeves and Smith 2009, Reeves 2010).

2.6 Pre-processing of spectral data

Various pre-processing methods have been used to transform soil spectra and remove noise due to light scattering. As a result, useful data are extracted for quantitative predictive models. Spectral pre-processing methods provide an important clue in multivariate calibration (Buddenbaum and Steffens 2012) and have enhanced the accuracy of predictive models (Bilgili *et al* 2010, 2011; Mashimbye *et al* 2012). Different kinds of spectral pre-processing include smoothing, normalization, scatter correction, continuum removal, and derivative algorithms. For example, first- and second-derivative absorption spectra have been used to enhance spectral information and demonstrate spectral changes during a decomposition process (Ben-Dor *et al* 1997). Savitzky–Golay (SG, Savitzky A and Golay 1964) smoothing was applied with derivatives to enhance the performance of models to predict soil organic carbon level (Stevens *et al* 2010) A common spectral pre-processing method for near-infrared spectroscopy for a spectral data set in a calibration model is the signal transformation. This transformation gives the best regression performance with comprehensive validation for all

possible variation in the data (Rinnan *et al* 2009).

Out of different types of pre-processing available, seven types are commonly used for studying soil properties, (1) raw spectra (R), (2) Savitzky–Golay (SG) smoothing, (3) first-derivative spectra with SG smoothing (FD-SG), (4) second-derivative spectra with SG smoothing (SD-SG), (5) continuum removed reflectance (CR), (6) standard normal variate and detrending (SNV-DT), (7) multiplicative scatter correction (MSC) (Buddenbaum and Steffens 2012).

2.7 Statistical modelling techniques

There are so many statistical tools those can be successfully utilized for developing prediction relationships between chemically tested data and spectral data sets.

2.7.1 Principal components regression (PCR)

Principal components regression (PCR) technique is used as a strategy whenever regression data are affected by multicollinearity. Multicollinearity is state in which a strong correlation is established among two or more prediction variables in a multiple regression model. PCR strives to search for some linear associations of variables that can be used to sum up the data without losing much information during the process (Ben-Dor and Banin 1995, Chang *et al* 2001).

2.7.2 Partial least square regression (PLSR)

PLSR technique is a promising and popular statistical tool used in chemometrics. Basically, the easy access to software and comparatively the principles representing the relationship between soil properties and spectral contribution to regression model are easily understood. This technique is more like PCR except PLSR tool identifies successive orthogonal factors that can maximize the covariance between predictor (X, reflectance spectra) and response variable (Y, laboratory measured soil properties). When predictor variables are highly collinear, PLSR technique develops the model effectively. One can get a few PLSR factors that draw the main variation in both predictor and response factors. PLSR can manage X and Y factors to delineate new components like latent variables, which are both orthogonal and weighted linear blends of X factors. Variables X and Y are centred by subtracting the column averages from each and every observation column before analysis. In case of PLSR, the possible number of latent variables is checked for over and under fitting data. Accurate fitting is attained by making cross validation, where the models are reformed every time by excluding small samples out of the calibration data set for use in the validation process unless all samples or groups are tested. The prediction mechanisms can be used in a spectral calibration model depending on the corresponding spectral interaction of the predominant soil chromophores (Vohland *et al* 2011). So, selection of the calibration method and its performance in modelling reflectance spectra is one of the main factors for calibration success (Mouazen *et al* 2010). PLSR is known for its simplicity and robustness (Farifteh *et al*

2007, Viscarra Rossel 2007 and Vasques *et al* 2008). Also, this approach helps the modelling of several response variables at a time, while effectively addressing strongly collinear and noisy predictor variables (Wold *et al* 2001). In case of complex and non-linear reflectance behaviour in soils PLSR becomes weaker (Vohland *et al* 2011).

2.7.3 Multivariate adaptive regression splines (MARS)

Besides linear regression method, multivariate adaptive regression splines (MARS), a nonparametric method that predicts complex nonlinear relationships among independent and dependent variables (Friedman 1991) has been effectively applied in different fields (Felicísimo *et al* 2012, Samui 2012) and generally exhibits higher performance results. MARS can be used to model soil properties, including soil OM and clay content, and is reported to provide better soil property estimations than those by PLSR method (Bilgili *et al* 2010).

2.7.4 Artificial neural network (ANN)

ANN attempts to develop a model which is compared with working of human brain and is analogous to the biological functioning mode of learning and data storing, comprising a dense network of connections between input data, neurons disposed in different hidden layers with parameters to be fitted, and output data. Previous knowledge of the relations between input and output data is not required. (deMelo *et al* 2015)

2.7.5. Support vector regression (SVR)

Support vector regression is a nonlinear method based on the statistical learning theory (Vapnik 1998). SVR has higher capacity of using fewer training data for learning in high dimensional feature space. Normally, it produces a set of linear equations to achieve the support vectors (Xuemei and Jianshe 2013). Moreover, support vector regression can develop model of linear and nonlinear relationships and is capable of solving calibration problems with high performance (Suykens and Vandewalle 1999). SVR has been successfully used for modelling soil OM and clay content based on reflectance spectroscopy (Viscarra Rossel and Behrens 2010). This statistical tool has been gaining broad application in soil spectroscopy, due to its advantages and high performance (Viscarra Rossel and Behrens 2010, Vohland *et al* 2011).

2.7.6 Boosted regression tree (BRT) / Boosted tree (BT)

BRT (Boosted regression trees) method compiles mixture of regression trees and a boosting technique to improve the predictive performance of multiple single models. Boosting technique is forward and step-wise procedure in which a subset of the data is randomly selected to randomly fit new tree models to minimize the loss function (Elith *et al* 2008). This may include a stochastic gradient boosting procedure which can make model performance better and ameliorate the risk of overfitting (Friedman 2002).

2.7.7 Random forest (RF)

The RF algorithm makes multiple trees without pruning. During training process, every tree is built based on a random sub set of the original data. However, a random subset of predictors is selected for each built tree (Breiman 2001). The productive use of bootstrap sampling in RF modeling employs the residual un-used subset (i.e., the out-of-bag data (OOB)) to be utilized for the determination of general errors. Random Forest estimations are the mean output of all aggregations. RF modeling needs three user-defined properties: the number of variables used to grow each tree (try), the number of trees in the forest (tree) and the minimum number of terminal nodes (node size) (Grimm *et al* 2008).

2.8 Spectral library approach

Spectral library approach is a holistic approach which includes generalized soil assessment results. Spectral library approach gives a consolidated framework which acts as a linkage between soil information and remote sensing information for further reference. A vast spectral library consisting of spectral signatures of different soil minerals is developed by USGS (United States Geological Services). In Indian context of soil science, the spectral library approach is not that common. However, there exists one study by National Bureau of Soil Survey and Land Use Planning (NBSS&LUP), Nagpur about agroecological sub regions of India which included reflectance libraries (128 soil samples) (Saxena *et al* 2005). Similarly, there was another study about swell and shrink behaviour of soil by employing spectral reflectance in central India that demonstrated good results (Srivastava *et al* 2004).

Moreover, global spectral library is developed by analyzing large number of data set which are diversified. The collected information must represent the land cover with relation to global geographical distribution. Global spectroscopic models are developed to predict individual soil properties and their uncertainties. To achieve this, unnecessary background noise from the spectra are removed. Therefore, a robust, prominent and a representative spectroscopic model can be derived considering spectra and laboratory analysis attributes (Brown 2007, Viscarra Rossel 2009, Viscarra Rossel *et al* 2015, 2016).

CHAPTER III

MATERIALS AND METHODS

3.1 Description of study sites

Punjab consists of five agroecological sub-regions namely Sub-mountain Siwalik Hills, Northeastern Undulating, Piedmont and Alluvial Plain, Central Alluvial Plain, and Southwestern Alluvial Plain (Fig 3.1) (Rajkumar *et al* 2008). These are broadly uniform with respect to shortcomings and potentials of the existing natural resources. Different kinds of land uses are seen in different agroecological sub regions. From each agroecological sub-region (Sub-mountain Siwalik Hills, Northeastern undulating, Piedmont and Alluvial Plain, Central Alluvial Plain, and Southwestern Alluvial Plain), 100 (total 500) georeferenced surface (0-15cm) soil samples were collected.

3.1.1 Sub-mountain Siwalik Hills

The Sub-mountain (Siwalik hills) agroecological sub-region covers only the North-eastern part of Punjab state adjoining Himachal Pradesh and Jammu and Kashmir. The soils of this sub region occur on moderately steep slopes.

3.1.2 Northeastern Undulating

The Northeastern Undulating agroecological sub-region comprises of recently sedimentary deposits of Siwalik hills. It occurs as low to moderately sloping and undulating land.

3.1.3 Piedmont and Alluvial Plain

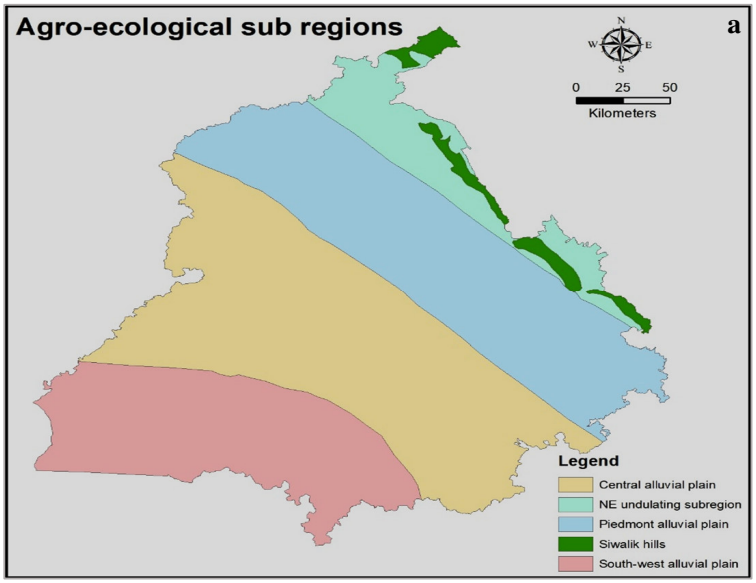
This agroecological sub-region comes between the foot hills of Siwalik hills and the river terraces of Punjab.

3.1.4 Central Alluvial Plain

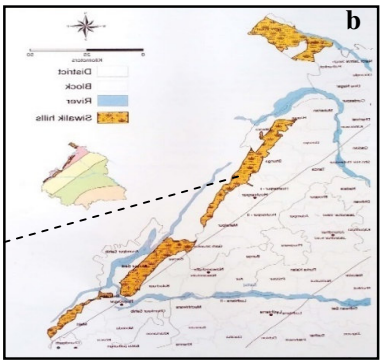
This agroecological sub-region covers central alluvial plain region of Punjab state. It is a part of the north Punjab plain, Ganga-Yamuna Doab and Rajasthan upland, hot, dry semi-arid eco sub region of India.

3.1.5 Southwestern Alluvial Plain

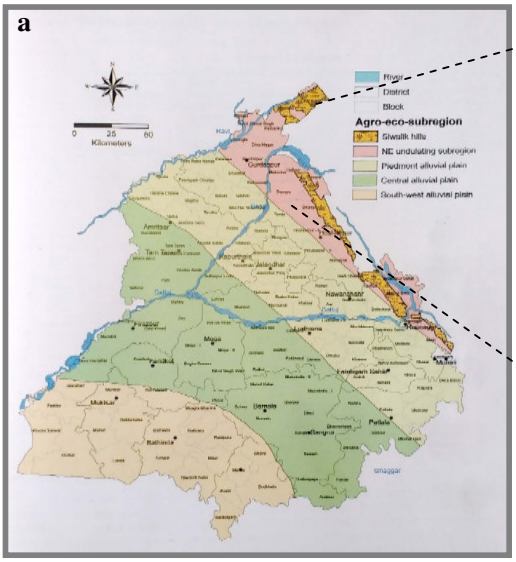
This alluvial plain is bordered by the desert area of Rajasthan to the south and covered by Sutlej. Sizeable part of the region is beset with limitations like salinity and waterlogging.

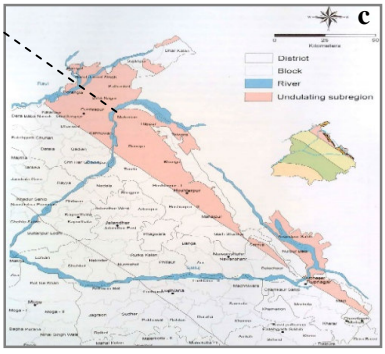


Location of study sites



Sub-mountain Siwalik Hills





Northeastern Undulating

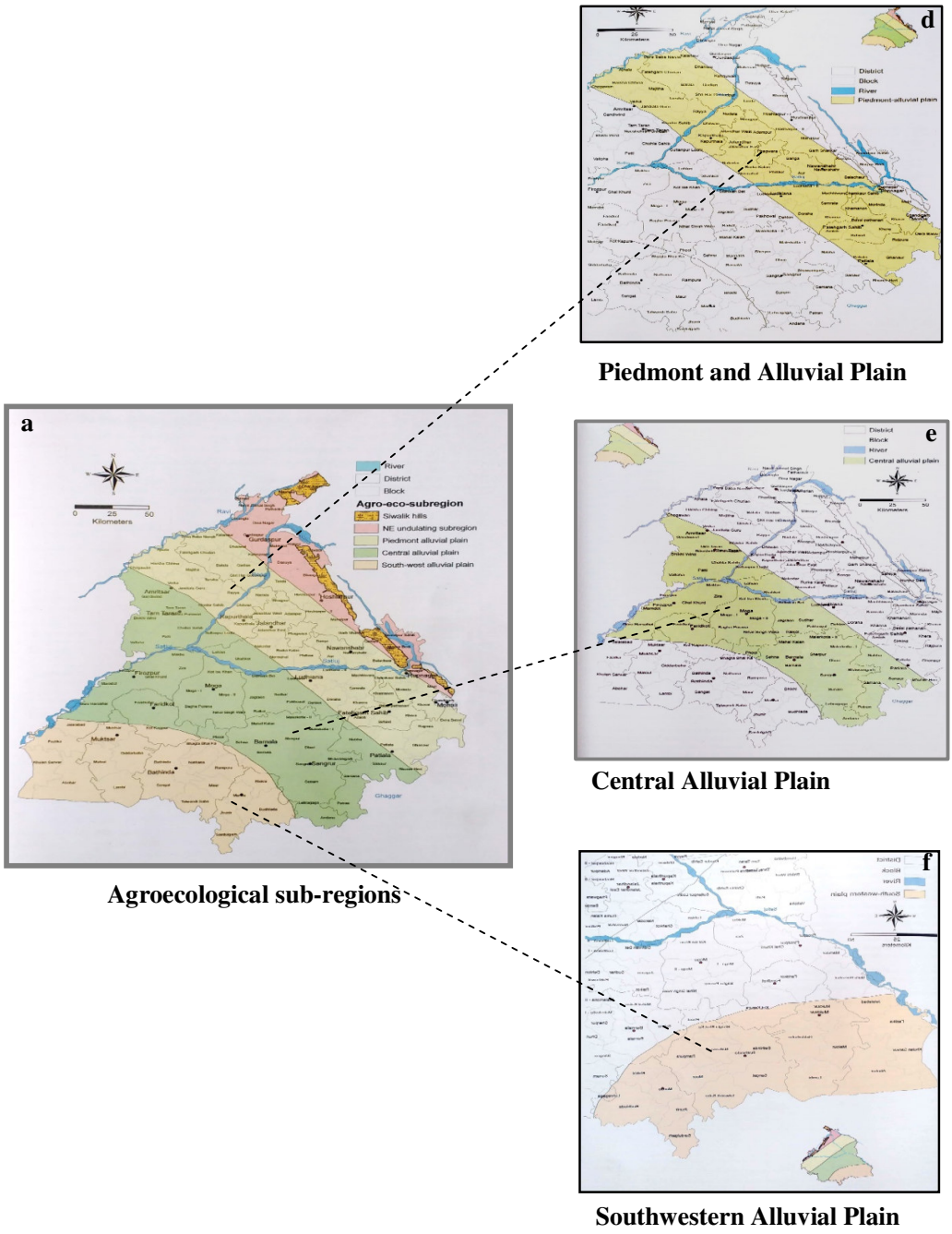


Fig. 3.1 (a-f) Maps showing study area involving various agroecological sub-regions of Punjab

(Rajkumar *et al* 2008)

3.2 Collection of soil samples

Geotagged surface soil samples (0-15 cm) were collected randomly by using stratified random sampling. From each agroecological sub-regions, 100 samples were collected

3.3 Preparation and processing of soil samples

Collected soil samples were air dried under shade. Different kind of pebbles, stones, concretions, etc. along with plant roots were removed. By using agate mortar large aggregates and clods were broken properly. Thereafter, each sample was passed through 2 mm sieve.

After processing, soil samples were divided into two sub samples. One part was used for traditional soil testing for determining physico-chemical parameters (soil texture, pH, EC, SOC, Available N, Available P, Available K, $CaCO_3$, DTPA extractable Fe, Mn, Cu, Zn).and other part was used for hyperspectral analysis by using spectroradiometer (Spectroscopic analysis, ASD FieldSpec® Pro FR, No. A110070).

3.4 Determination of soil parameters

The above-mentioned parameters were determined by following methods:

3.4.1 Soil textural analysis (Particle size distribution)

Soil textural analysis was done by following International pipette Method (Day 1965). Exactly 20 g of soil sample was taken in a 500 ml beaker. Then soil sample was treated with 30% hydrogen peroxide (30% H_2O_2). This was done to digest the organic matter present in soil sample by placing on a hot plate. Intermittent stirring with glass rod for preventing overflow of soil froths was performed. Then 50 ml of sodium hexametaphosphate was added to it along with considerable amount of distilled water dispersing soil particles. The suspension was then stirred for 10 minutes by the help of electric stirrer. After stirring, total sand particles were separated by using 0.2 mm sieve following wet sieving technique. These were collected by using a stoppered wash bottle carefully to reduce error. The silt and clay particles along with fine sand particles were transferred carefully into a 1000 ml cylinder. After that volume of the suspension was made up to 1 litre by adding distilled water and the components were mixed well by a plunger. The temperature of the suspension was recorded by dipping a thermometer. Then a 25 ml of suspension was pipetted out from a depth of 10 cm marked on the pipette after settlement of fine sand allowing requisite time (calculated from the table containing time and temperature relationship from Stokes' law) for silt+clay determination. After that plunging was carried out again. Again, thermometer was dipped inside the suspension and temperature was noted to determine time for pipetting for clay particles. Exactly 25 ml of suspension was pipetted out at prefixed time. The percent clay was calculated based upon the clay content in the suspension, obtained after specific time period and silt content was calculated by subtracting from percent silt+clay content. At last fine sand content was calculated by subtracting from silt+clay+coarse sand from 100 percent.

3.4.2 Soil chemical properties

3.4.2.1 Soil pH

The pH of soil suspension was determined by taking soil and distilled water in the ratio of 1:2. Soil suspension was continuously stirred by using a glass rod. 20 g of soil sample was taken in a 50 ml beaker. Intermittent shaking was done up to half an hour. Reading was taken on Elico LI 127 pH meter by following the potentiometric method. Before taking sample readings, pH meter was standardized by buffers like 4.0, 7.0 and 9.2 at 25^0 C. (Jackson 1973)

3.4.2.2 Electrical conductivity (EC)

The 1:2 soil water suspension, used for pH determination, was kept as such undisturbed for overnight. Then electrical conductivity of the supernatant of that suspension was measured by using conductivity meter. Before taking readings, instrument was calibrated with 0.1N KCl solution at 25^0 C and sample reading was taken at room temperature.

3.4.2.3 Soil organic carbon (SOC)

Soil organic carbon (SOC) was estimated by using rapid titration method (Walkley and Black, 1934). A standard amount of soil (1g) was taken and 5 ml of potassium dichromate (1N $K_2Cr_2O_7$) and 10 ml of conc. sulfuric acid (H_2SO_4) were added to generate heat to carry out oxidation reaction. After 30 minutes, when sample was cooled down excess potassium dichromate was determined by titration with ferrous ammonium sulfate (0.5N Fe $(NH_4)_2(SO)_2$). A pinch of sodium fluoride powder and diphenyl amine (as indicator) were added. Change of colour from purple to green was considered as end point. Oxidizable organic carbon was worked out from the amount of potassium dichromate used for oxidation.

3.4.2.4 Available nitrogen

Available nitrogen was estimated by alkaline potassium permanganate ($KMnO_4$) method of Subbiah and Asija (1956). An automated Kjeldahl (Kelplus) apparatus was used. Five g of soil was taken in a digestion tube in which 25 ml of potassium permanganate (0.32%) and 25 ml sodium hydroxide (2.5% NaOH) for making the alkaline condition were added. Separately, in a conical flask 5 ml of boric acid (4% boric acid and 20 ml mixed indicator (methyl red + bromocresol green) were mixed well and pH was adjusted between 4.1-4.5). When contents of boric acid flask became green (means evolved ammonia), it was titrated with N/200 sulfuric acid until the colour was changed from green to light pink. The volume of acid used for absorption was used to calculate available nitrogen.

3.4.2.5 Available phosphorous (Olsen-P)

Available phosphorous (P) in soil sample was estimated by following Olsen method (Olsen *et al* 1954). The extracting solution used was sodium bicarbonate (0.5N $NaHCO_3$) buffered to pH 8.5 to diminish the effect of calcium carbonate. A pinch of Darco G-60 was added to absorb the dispersed organic matter and to make the solution colourless. Then

sample was shaken for 30 minutes on an electrical shaker. After shaking, solution was filtered and 5 ml of filtrate was taken in a 25 ml volumetric flask. Then 0.5 ml of 5N H_2SO_4 was added and shaken up until evolution of CO_2 ceased. Four ml (4 ml) of reagent B was added to it. Reagent-A was made by mixing 12g of ammonium molybdate, 0.2908g of antimony potassium tartrate and 5N sulfuric acid. Reagent-B was made by taking 200 ml of reagent-A and 1.056 g of ascorbic acid. This solution was added to produce heteropoly complex, showing yellow colour which was turned into blue in the presence of ascorbic acid. After 10 minutes, blue colour started appearing. The intensity of blue colour was measured by spectrophotometer (Hach DR 3900) by setting the wavelength at 760 nm. Then the concentration of available P was computed from the standard curve of the phosphorous, prepared at the same time.

3.4.2.6 Available potassium

Determination of soil available potassium was carried out by the neutral normal ammonium acetate method (Merwin and Peech, 1950). A 5g soil sample was taken in a 150 ml conical flask and into it 25 ml of extracting reagent (1N CH_3COONH_4 pH 7.0) was added. Then sample was shaken constantly for 5 minutes in a mechanical electrical shaker and then filtered on a Whatman no. 1 filter paper. After it, 5 ml of extract was taken in a 25 ml volumetric flask and volume made up by distilled water. Then final reading for potassium was taken with a flame photometer. Before taking the readings, instrument was calibrated with a range of standard potassium solutions (5 ppm, 20 ppm and 100 ppm). (Elico CL 361)

3.4.2.7 Calcium carbonate ($CaCO_3$)

Calcium carbonate was estimated by rapid titration method (Puri, 1930). 10 g of air-dried soil sample was taken in a 250 ml conical flask. 100 ml of distilled water was added to it. Thereafter, 0.5 g of $CaSO_4.2H_2O$ and 10 ml of $AlCl_3.6H_2O$ were added into the solution. Then the solution was shaken thoroughly and 10 drops of each indicator (1 g bromothymol blue in 100 ml ethanol and 1g bromocresol green in 100 ml of ethanol) were added. The suspension was heated on a hot plate to bring to the boiling point and then the flask was removed for settlement of soil particles. Appearance of green colour indicated presence of $CaCO_3$ and that of golden-yellow colour the absence of $CaCO_3$. Whenever the supernatant was green coloured, it was titrated against 0.5N H_2SO_4. The suspension was brought to boil again and the above steps were repeated until a permanent golden-yellow colour appeared, which persisted for a few minutes on boiling and was allowed to settle for 1-2 minutes. Then volume of acid used was noted to calculate calcium carbonate.

3.4.2.8 DTPA extractable micronutrients (Fe, Mn, Cu and Zn)

Micronutrient cations (Fe, Mn, Cu and Zn) were determined by following the procedure given by Lindsay and Novell (1978). For extracting the micronutrients cations i.e. Fe, Mn, Cu and Zn, soil was treated with DTPA extracting solution (consisting of 0.005M

DTPA, 0.01M $CaCl_2$ $2H_2O$ and 0.1M TEA, buffered at 7.3) in 1:2 ratio and shaken for 2 hours in a temperature-controlled shaker (at 25^0 C). A 10 g soil sample was taken in a plastic bottle (150 ml) and into it 20 ml of DTPA reagent was added. Then it was shaken for two hours in a temperature controlled electrical shaker (25^0C and 180 rpm). After that it was filtered on a Whatman no. 42 filter paper in vial. Standards for cations like Fe, Mn, Cu and Zn were made and fed to atomic absorption spectrophotometer (Varian AAS-FS 240 model) for assessing the four micronutrient elements.

3.5 Classical statistical analysis

Statistical analyses were performed with SAS 9.3 (SAS (2011) Institute Inc. Cary, NC, USA) and R studio (R Core Team 2019). Descriptive statistics (mean, median, range, and coefficient of variation) were generated to characterize the distribution of different parameters. The data were subjected to suitable statistical analysis for evaluating departure from normality.

3.6 Soil fertility maps

Arc-GIS software (10.4) was used to make soil fertility maps. Inverse distance weighted (IDW) interpolation and classification were used to draw soil maps.

3.7 Spectroscopic analysis of soil properties

The second subsample (as explained earlier) was used for spectral analysis in visible-near infrared (Vis-NIR) range. Soil reflectance spectra (in the range 350-2500 nm) were recorded for each sample using a Fieldspec Pro spectroradiometer (Analytical Spectral Devices Inc., Colorado, USA). with a spectral sampling interval of 1 nm.

Air-dried soil samples were ground and passed through 2 mm sieve. Black polypropylene dishes were used to record spectra accurately. The dishes were overfilled with soil and excess soil was scraped off by using sharp blades to ensure the flat surface in level. Soil samples spectra were taken by using a contact probe which had constant illumination by the source fitted inside The soil spectra were taken inside a dark chamber (i.e. $\pm 25^0$C) in which each wall was painted with black colour so that all light would be absorbed as they acted like a black body with less interference from reflectance spectra.

An average of 30 spectra was recorded for each soil sample. The instrument was calibrated by taking 30 white reflectance spectra with white calibrated Spectralon (Labsphere, Sutton, NH). Reflectance readings for each wavelength band were expressed relative to the average of the white reference readings. Soil samples were studied by plotting the relative reflectance/absorbance values against different wavebands to detect the variation in overall spectral behaviour of various soil samples.

3.7.1 Processing of spectra

3.7.1.1 Noise removal

The noise in the spectral reflectance data *i.e.,* beyond 2200 nm and before 450 nm were removed before doing the derivative transformation. The prospectr package was used for noise removal. (Annexure-I)

3.7.1.2 Derivative transformation

The spectral data were then subjected to second derivative transformation by prospectr (Stevens and Ramirez-Lopez 2013) package in R software (R Core Team). The derivative transformation was computed by finite difference method. The advantages of derivative transformation are reduction of absorption overlapping, removal of multiplicative effect and likely increase in the predictive accuracy of complex datasets. (Annexure-II)

3.7.2 Statistical modelling technique

The collected soil samples of five agroecological sub regions with different land use systems were combined with the aim of producing a more robust hyperspectral model for soil fertility assessment. The pls package in R studio was used to fit the model by partial list square regression technique. To develop the spectral model for prediction of soil properties, the whole data set was randomly divided into two subsets viz. Training data set (66) and Test data set (34) in 2:1 ratio. The training data set was used to develop a prediction equation involving number of components such that the R^2 for calibration is maximum and RMSEP is minimum. The test data set was used to validate the predictive equation and to test the model accuracy. The Levene's test (p=0.402) for equality of variances (Levene 1960) and the student's 't' test (0.97) for equality of means in different parameter values between calibration and validation sets indicated that validation datasets were truly representative. The formula for the R^2, RMSEP and RPD are as below

$$R^2=1-\frac{\sum(Y_{pred}-Y_{meas})^2}{\sum(Y_i-Y_{meas})^2}$$

$$RMSEP=\sqrt{\frac{\sum(Y_{pred}-Y_{meas})^2}{n-1}}$$

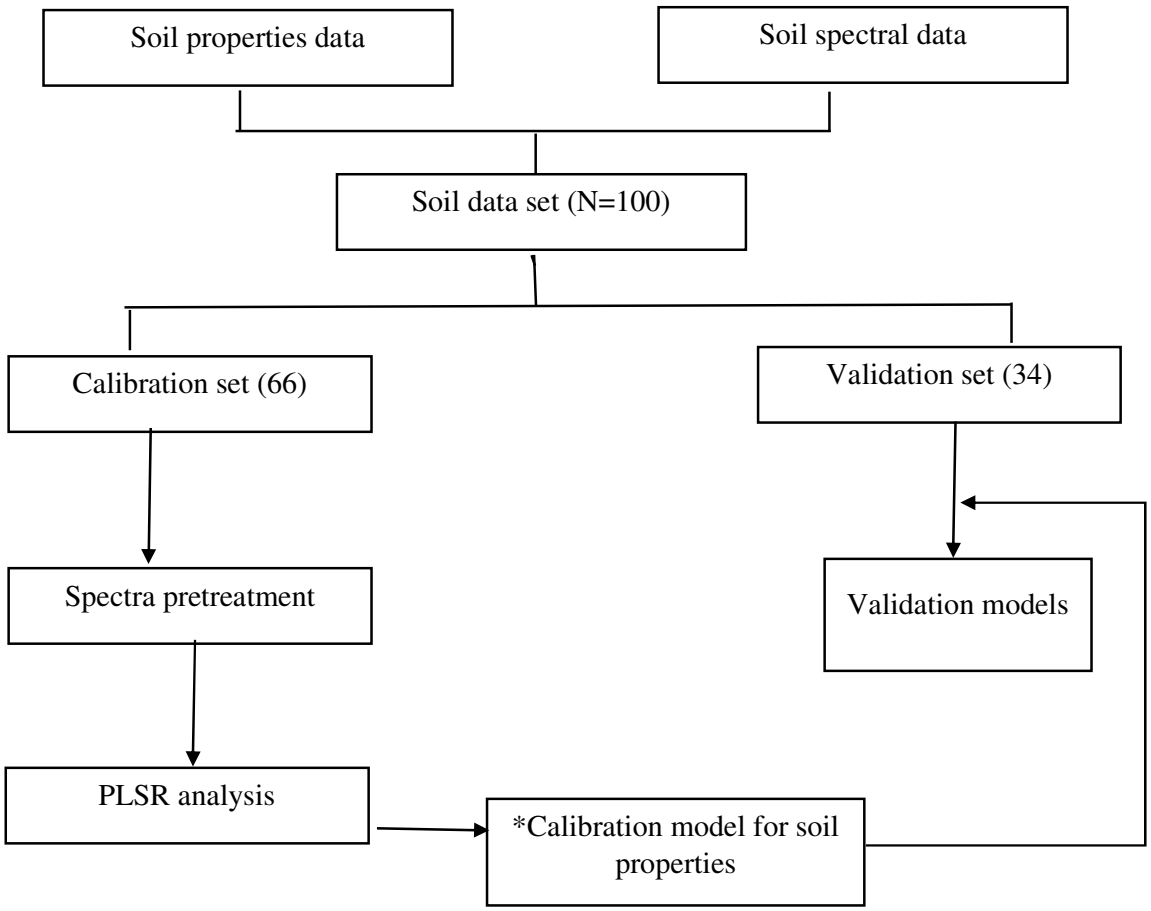
$RPD = SD_{val}. /RMSEP$

Where Y_{pred} =predicted values; Y_{mean} =mean of measured values; Y_{mean} = measured values; n= number of predicted or measured values with I = 1, 2…n; SD_{val} = standard deviation of measured values in the validation set; and RMSEP =root mean square error of prediction dataset.

The leave-one-out cross validation procedure was used to verify the predictive significance of each PLSR component and determine the number of factors (latent variables) to be retained in calibration model: one sample was left out of data set and the model was

calculated on remaining data points. The value of the left-out sample was then predicted and the residuals were computed. The process was repeated with another sample from the data set, and so on, until every sample had been left out once. The predictive ability of the cross-validation models was evaluated by the coefficient of determination (R^2) and the root mean square of prediction (RMSEP), mentioned before. The models were independently validated through validation dataset and the coefficient of determination and the root mean square error of validation were computed to check the performance of the model.

We followed the following flowchart to develop spectral model.



(*) Calibration models are presented in Annexure-II to Annexure-VI.

CHAPTER – IV

RESULTS AND DISCUSSION

Results of study are discussed under the following headings:

4.1 Distribution of soil fertility parameters under different agroecological regions

4.2 Modelling of hyperspectral data for prediction of soil properties

4.3 Portability of spectral models across various agroecological sub regions

4.1 Distribution of soil fertility parameters under different agroecological regions

In this experiment, different agroecological sub-regions, namely, Sub-mountain Siwalik Hills (SSH), Northeastern Undulating (NEU), Piedmont and Alluvial plain (PAP), Central Alluvial Plain (CAP) and Southwestern Alluvial Plain (SWAP) of Punjab were considered for studying various major physico-chemical parameters, such as soil reaction (pH), electrical conductivity (EC), soil organic carbon (SOC), soil available nitrogen (N), available phosphorus (Olsen-P), available potassium (K), calcium carbonate ($CaCO_3$), DTPA-extractable iron (Fe), DTPA-extractable manganese (Mn), DTPA-extractable zinc (Zn), DTPA-extractable copper (Cu), and sand, silt & clay contents. The important results are presented and discussed as under:

4.1.1 Soil reaction (pH)

The soil reaction values in concerned agroecological sub-regions were in the order: NEU<CAP<SSH<PAP<SWAP (Table 4.1 and Fig 4.1a). Soil pH values were concentrated around 7.44 and ranged from 6.65 to 7.71 with lower variation (CV= 5.4%) in SSH region, whereas in case of NEU region pH varied from 6.11 to 7.98 with an average of 7.04. Similarly, PAP region had pH values ranging from 6.9 to 8.28 with a mean of 7.56 (CV=4.9 %). However, CAP region pH varied (CV=2.4%) from 7.01 to 7.89 with a mean of 7.39. In case of SWAP region, soil pH had a mean value of 7.93 and ranged from 7.3 to 8.47 (CV=3.1 %).

In case of SSH region, pH had very poor correlation with other soil properties (Table 4.15). Moreover, pH had positive relationship with phosphorous (r = 0.243*) in NEU region (Table 4.16). However, it had a negative trend with calcium carbonate (r = -0.312**), and clay (r = -0.332**) content, whereas in CAP region the pH showed very poor correlation with other properties (Table 4.18). Keykha *et al* (2017) observed similar relationship between pH and calcium carbonate. In PAP region, EC (r = 0.651**), calcium carbonate (r = 0.295**), iron (r = 0.482**), manganese (0.377**), copper (0.474**), zinc (0.505**), sand (0.283**) and clay (r = 0.418**) influenced pH positively (Table 4.17), whereas nitrogen content (r = 0.236*) was significantly related to pH in SWAP region (Table 4.19). Zhu *et al* (2016) recorded similar trend in effect of pH on micronutrients. Regions rich in calcium carbonate,

showed generally higher pH (>7.5) values in accordance with the enhancing effect of calcium carbonate hydrolysis on soil pH. Strong correlations among EC, micronutrients and pH indicate basic nature of cations involved in salt composition in SWAP region. Huang *et al* (2016) observed similar trend in relationship between available P and soil pH.

4.1.2 Electrical conductivity (EC)

Electrical conductivity, the indicator of soil salinity is described in Table 4.2 and Fig. 4.1b. The trend was in the following order: NEU<SSH<PAP<CAP<SWAP. Electrical conductivity lay around the mean of 0.24 dS m^{-1} with a range from 0.12 to 0.51 dS m^{-1} (CV=38.2%) in SSH region. Similarly, in NEU region, EC values ranged from 0.14 to 0.19 dS m^{-1} with high variation (CV=59.1 %). However, in PAP region EC showed a mean of 0.55 dS m^{-1} with great variation (CV=90.8%). In CAP region, EC values varied from 0.11 to 2.04 dS m^{-1} (CV=81.2%) and concentrated around 0.59 dS m^{-1}. The SWAP region showed average EC value of 0.95 dS m^{-1} ranging from 0.02 dS m^{-1} to 2.42 dS m^{-1}. The trend in SWAP region is along expected lines as the soils in this area have accumulated salts over time owing to aridic nature of moisture regime.

The SSH region showed significant correlation of EC with nitrogen (r = 0.213*) in (Table 4.15). This can be ascribed to contribution of mineralized N. Similarly, in NEU region, phosphorous (r = -0.213*) was negatively correlated to soil electrical conductivity (Table 4.16) suggesting that ions like Ca may be predominant among active ions. However, iron (0.548**), manganese (0.269**), copper (0.514**), zinc (0.420**), sand (0.330**) showed positive relationship with electrical conductivity in PAP region (Table 4.17). In addition, electrical conductivity varied negatively with copper (r = -0.65**) in CAP region and did not have any significant tendency towards other parameters except clay (0.23*) (Table 4.18). Similarly, Table 4.19 clearly showed that phosphorous (0.235*), copper (0.263**) and clay (0.238*) were very well correlated with EC in SWAP region, suggesting thereby that relatively heavy textured soils might be harbouring more salts due to higher cation exchange capacity. Similarly, Serrano *et al* (2017) observed effect of fertilizer practices on salt distribution in soils.

4.1.3 Soil organic carbon (SOC)

A comparison of soil organic carbon under different agroecological sub-regions is presented in the Table 4.3 and Fig. 4.1c. Different agroecological sub regions exhibited the following order: SWAP<PAP<NEU<CAP<SSH. In SSH region the values stayed around the mean 0.77 with a coefficient of variation of 22.1 % (range 0.26 % to 1.13 %). Due to lack of intensive cropping, soil organic matter content is undisturbed in this region and is subject to minimal transformation. However, other regions like NEU and PAP regions had more variation (26.2 % and 36.6 % coefficient of variation, respectively). SOC content varied from

0.19 to 0.90 % and 0.23 to 1.01 % in NEU and PAP regions, respectively. CAP region exhibited a mean SOC content of 0.56 % ranging from 0.26 to 0.90 % (CV=27.2 %). The SOC content had a large variation (CV=32.5 %) with a range 0.18 to 0.83 % and 0.39 % mean in SWAP region.

The SOC had significant correlation with potassium (r = 0.217*) in SSH region (Table 4.15), suggesting thereby that decomposing plant liter contributed considerably to K content. In NEU region, SOC demonstrated significant association with nitrogen (r = 0.238*), iron (r = -0.206*), zinc (r = 0.220*), sand (-0.218*) and silt (r = 0.264**) (Table 4.16). However, SOC content varied in a positive way with potassium (r = 0.273**), calcium carbonate (r = 0.248*) and negatively with iron (r = -0.291**), manganese (-0.349**) in PAP region (Table 4.17). CAP region demonstrated negative correlation with zinc (r = -0.30**) and clay content (r = -0.26**) (Table 4.18). In SWAP region, phosphorous (r = 0.287**), manganese (r = 0.319**) and copper (r = 0.231**) were significantly correlated with SOC (Table 4.19). Also, Parat *et al* (2002) and Singh (2018) observed dependency of copper on soil organic carbon. Debicka *et al* (2016) observed similar trend in relationship between SOC and available P.

4.1.4 Available nitrogen

Alkaline potassium permanganate oxidizable nitrogen was noted as 91.9 kg ha^{-1} ranging from 31.4 to 235.2 kg ha^{-1} (CV=49.2 %) in SSH region. Similarly, in NEU region, soil N varied from 25.1 to 269.7 kg ha^{-1} with a mean of 104.5 kg ha^{-1} (CV= 41.40 %). In CAP region, an average 87.9 kg ha^{-1} ranging from 25.1 to 150.5 kg ha^{-1} (34.8 % coefficient of variation) was observed. However, in PAP region, available nitrogen content ranged from 41.4 to 345 kg ha^{-1} (mean of 89.7 kg ha^{-1}, CV=34.8 %). In SWAP region, a mean of 77.4 kg ha^{-1} (ranging from 15.68 to 244.6 kg ha^{-1}) was observed with very high variation (CV=48.66%) was noticed in SWAP region. Across various regions, available N exhibited the trend SWAP<CAP<PAP<SSH<NEU (Table 4.4 and Fig. 4.1d).

In SSH region, available N was significantly correlated with calcium carbonate (0.372**) and zinc (0.220*) (Table 4.15). This can be attributed to calcium carbonate and zinc content affecting nitrogen dynamics favourably. In NEU region, however, nitrogen was significantly correlated with silt (r = 0.208*) (Table 4.17). However, nitrogen was very poorly correlated with other properties in CAP (Table 4.18). In the same way, pH (r = 0.236*) and manganese (r = 0.339**) were significantly correlated with available nitrogen pool. (Table 4.19) in SWAP region. Mandal *et al* (2016) observed similar dependency of available nitrogen on pH.

4.1.5 Available phosphorous (Olsen-P)

A comparative evaluation of soil available phosphorous is presented in Table 4.5 and

Fig. 4.1e. Olsen-P under different agroecological sub-regions followed the following order: SSH<CAP<PAP<SWAP<NEU. The SSH region exhibited phosphorous content ranging from 1.2 to 29.6 kg ha^{-1} with a mean of 20.7 kg ha^{-1} (CV=24.1 %). Also, in CAP region, the available phosphorous content varied from 13.8 to 31.4 kg ha^{-1} with a mean of 22.4 kg ha^{-1} and low variation (CV=15.8 %). Similarly, Piedmont and Alluvial Plain region showed higher P content with a mean of 34.1 kg ha^{-1} and a range between 20.3 to 72.6 kg ha^{-1} with high variation (CV=29.2 %). The SWAP region also exhibited more variability (CV=22.9 %) ranging from 23.8 to 61.3 kg ha^{-1} and a mean of 36.2 kg ha^{-1}. Similarly, P content lay around a mean of 40.8 kg ha^{-1} with higher variability (CV=27.7 % and a range between 15 to 72.3 kg ha^{-1}) in NEU region.

Phosphorous content demonstrated poor association with other properties in SSH region (Table 4.15), while in NEU region Olsen-P was correlated negatively with pH (r = -0.243*), EC (r = -0.213*) and clay (r = -0.213*) (Table 4.16). However, PAP region exhibited association with sand (r = -0.215*) (Table 4.17). In SWAP region, pH (r = 0.236*) and manganese (r = 0.339**) showed significant correlation with available phosphorous (Table 4.19). This means, pH affects P availability in soil to plants. Bai *et al* (2017) also found similar relationship between available P and pH.

4.1.6 Available potassium (K)

Ammonium acetate extractable potassium under various agroecological sub regions is presented in Table. 4.6 and Fig. 4.1f. Trend in different agroecological sub regions was CAP<SSH<NEU<SWAP<PAP. The SSH region showed a mean of 143.9 kg ha^{-1} with high variation (CV % = 48.9) and a range from 27.8 kg ha^{1} to 457.3 kg ha^{-1}. Similarly, in NEU region, available potassium varied from 85.3 to 246.9 kg ha^{-1} with a mean of 167.7 kg ha^{-1} (CV=20.8 %). Available potassium ranged between 28.75 to 238.00 and a mean of 132.332kg ha^{-1} in CAP region (CV=30.1 %). However, SWAP showed a higher mean of 173.2 kg ha^{-1} and K content ranged from 76.3 to 277.6 kg ha^{-1} with (CV=29.8 %). Moreover, in potato prevalent region like PAP, available potassium content varied from 76.25 to 865 kg ha^{-1} with high variation (CV=67.7 %) and a mean of 173.2 kg ha^{-1}.

Generally, potassic fertilizers are applied to increase the potato tuber yield. In SSH, potassium had positive tendency towards SOC (r = 0.217*) (Table 4.15), whereas positive correlation with clay content (r = 0.272**) was observed in PAP region (Table 4.17). Also, in SWAP region potassium was poorly correlated with other properties (Table 4.19). Behera and Shukla (2015) observed similarly weaker correlations.

4.1.7 Calcium carbonate ($CaCO_3$)

A comparative evaluation of calcium carbonate carried out under various agroecological sub regions is displayed in Table 4.7 and Fig. 4.1g. It followed the order:

NEU<SSH<PAP<SWAP<CAP. In SSH, calcium carbonate varied from 0.3 to 1.7 % with a mean 0.84 % (CV=41.0 %). However, in NEU region the values concentrated around 0.80 % where range lay between 0.01 to 5.75 % with great variation (CV=111.6%). In PAP region, the $CaCO_3$ content stayed around 1.12 % and a range from 0.65 to 3.70 % with high variation (CV=46.9 %). Similarly, CAP region exhibited high variation (CV=49.9 %) ranging from 0.54 to 3.25 % and a mean of 1.36 %. However, SWAP region showed calcium carbonate content varying from 0.6 to 3 % with a mean 1.31 % and high variation (CV=39.1 %).

Pearson correlation (r) for calcium carbonate showed positive relationship with nitrogen (r = 0.372**) in SSH (Table 4.15). This implies calcium carbonate content in soil strongly affects the availability of nitrogen. Also, Hamdan *et al* (2017) found relationship between mineralizable nitrogen and carbonate. Similarly, in NEU region, it was related to copper significantly (r = 0.200*) (Table 4.16). Table 4.17 clearly depicted the direct relationship of sand (r = 0.209*) with soil calcium carbonate in Piedmont and Alluvial plain.

4.1.8 DTPA-extractable iron (Fe)

Diethylene triamine penta acetic acid (DTPA) extractable iron content is presented in Table 4.8 and Fig. 4.1h. Under different agroecological sub regions, iron content followed the following order SSH<NEU<PAP<CAP<SWAP. In SSH region, it varied from 2.4 to 15.4 ppm with a mean value of 5.3 ppm and higher variability (CV=51.9 %). Similarly, in NEU region, it ranged from 3.5 to 21.5 ppm with more variation (CV=37.3 %) and a mean of 9.6 ppm. However, PAP region showed lower variability (CV=61.4 %) ranging from 6.3 to 62.3 ppm and values stayed around 19.6 ppm. Also, CAP region witnessed average iron content of 14.7 ppm (CV=29.4 %) from 5.9 to 24.5 ppm. However, in SWAP region, the iron content varied from 5.5 to 67.1 ppm and a mean of 29.5 ppm with more variation (CV = 37.4 %).

In SSH, DTPA-Fe had a strong correlation with manganese (r = 0.359**) (Table 4.15). In contrary, in NEU region, iron had very poor correlation with other properties. However, manganese (0.369**), copper (r = 0.366**), zinc (0.202*) were significantly correlated with iron (Table 4.17). Also, Table 4.18 exhibited very poor correlation between iron and other properties in CAP (except potassium, r = 0.20*) as well as in SWAP (Table 4.19). This relationship is attributed to importance of soil mineralogy in regulating availability of micronutrients. Shuman (1986) found similar relationship between iron and zinc.

4.1.9 DTPA-extractable manganese (Mn)

DTPA-extractable manganese content is presented in Table 4.9 and Fig. 4.1i for different agroecological sub-regions. DTPA-extractable manganese content followed the trend: SSH<PAP<CAP<SWAP<NEU. SSH region showed a mean of 4.7 ppm, ranging from 1.3 to 11.3 ppm (CV %=43.3). Also, in PAP region, DTPA-manganese content varied from

2.1 to 10.2 ppm with more variation (CV=35.2 %) and a mean of 5.2 ppm. However, CAP region exhibited an average manganese content of 5.33 ppm ranging from 1.3 to 12.3 ppm and higher variability (CV=42.5%). Similarly, in SWAP region, values of manganese stayed around 6.5 ppm ranging between 2.6 and 14.5 ppm (CV %=36.6).

Moreover, Zinc (r = -0.241*) and silt (r = -0.203*), soil organic carbon (r = 0.319**), nitrogen (r = 0.339**) had significant correlations with manganese in SSH region (Table 4.15). Also, similar dependency of manganese on SOC was explained by Wang *et al* (2018). In case of NEU region, DTPA-Mn varied significantly with zinc (r = 0.313**) and silt content (r = -0.221*) (Table 4.16). However, Table 4.17 explained the dependence of DTPA-extractable manganese on pH (r = 0.505**), EC (r = 0.420**), iron (r = 0.202*), zinc (r = 0.223*) in PAP region, whereas Table 48 showed poor relationships between manganese and other properties in CAP region. Similarly, DTPA-Mn was related to properties like zinc (r = -0.241*) and silt (r = -0.313**) and clay (r = 0.215*) in SWAP region (Table 4.19). Such dependencies among micronutrients convey the importance of mineralogy in regulating micronutrient availability in these regions.

4.1.10 DTPA-extractable copper (Cu)

A comparative evaluation of DTPA-extractable copper content under different agroecological sub regions is placed in Table 4.10 and Fig. 4.1j. It followed the order: CAP<SSH<NEU<PAP<SWAP. In case of SSH region, copper showed a mean of 0.5 ppm with great variation (CV=168.4 %, range between 0.1 ppm and 6 ppm). However, PAP region exhibited a mean of 0.7 ppm ranging from 0.01 to 2.2 ppm (CV=59.8 %). Similarly, in CAP region, copper content stayed around 0.5 ppm ranging between 0.01 and 2.1 ppm and high variation (CV=95.6 %). However, SWAP region showed a mean of 0.95 ppm and a range from 0.04 to 3.2 ppm (CV=50.9 %). Also, copper content in NEU region ranged from 0.01 to 6.5 ppm and a mean of 0.7 ppm with high variation (CV=109.3 %).

Very poor relationship between pH and other properties in SSH region, whereas in NEU region, copper had a strong association with soil calcium carbonate (r = 0.200*) (Table 4.16). Also, pH (r = 0.474**), EC (0.514**), SOC (r = -0. 232*) and clay (r = -0.350**) had strong effects on copper availability in PAP region (Table 4.17). However, under CAP region, it varied inversely with electrical conductivity (r = -0.65**) and clay (r = 0.20*) (Table 4.18). Also, SOC (r = 0.231**) and phosphorous (r = 0.254*) had significant correlation with copper content in SWAP region (Table 4.19). This suggested copper availability is affected by high organic matter content. Parat *et al* (2002) showed similar dependence between copper and SOC. Hartley *et al* (2016) also explained effect of organic matter on micronutrient dynamics.

4.1.11 DTPA-extractable zinc (Zn)

DTPA-extractable zinc content under various agroecological sub regions is

represented in Table 4.11 and Fig. 4.1k. The trend was SSH<NEU<CAP<PAP<SWAP. The SSH region exhibited a range from 0.2 to 2.4 ppm with a mean 0.8 ppm (CV=51.9 %). However, in NEU region, zinc varied from 0.01 to 5.9 ppm and mean of 1.1 ppm with great variation (CV=81.7 %). Similarly, PAP region had average zinc content of 2 ppm ranging from 0.3 to 11.6 ppm (CV=95.3 %). Also, CAP region showed a higher variability (CV=50.2 %) ranging from 0.3 to 3.8 ppm and a mean of 1.8 ppm. Similarly, in SWAP region, zinc varied from 0.6 to 11.6 ppm with a mean of 3.2 ppm and high variation (CV=66.3 %).

In NEU region, soil organic carbon (r = 0.220*) and manganese (r = 0.313**) were strongly associated with zinc (Table 4.16). Also, PAP region indicated the positive dependence of zinc on pH (r = 0.505**), EC (r = 420**), iron (r = 0.202*), manganese (r = 0.223*), copper and (r = 0.322**). Similarly, in CAP region, zinc content varied significantly with soil organic carbon (r = -0.30**), potassium (r = 0.29**) (Table 4.18). Similarly, SWAP region showed strong association of DTPA-Zn with manganese (r = -0.241*), copper (r = -0.267**), sand (r = -0.343**) and silt (r = -0.287**) (Table 4.19). The trends are suggestive of influence of mineralogy in controlling micronutrient availability. Shuman (1986) observed similar dependence between iron and zinc. Hartley *et al* (2016) reported in general similar effect of organic matter on micronutrient dynamics.

4.1.12 Sand content

The sand content under different agroecological sub regions followed the order PAP<NEU<SWAP<CAP<SSH (Table 4.12 and Fig. 4.1l). Among all the agroecological sub regions, lowest sand content was shown by PAP region. It had a mean sand content of 66.4 % varying from 54 % to 76 % with low variation (CV=6.3 %). However, in NEU region, sand content ranged from 55 % to 76 % with a mean of 67.3 % (CV = 6.9 %). Similarly, SWAP region exhibited an average sand content of 68.9 % varying from 58.8 to 77 % with low variation (CV=5.2 %). In CAP region, mean sand content of 71 % ranging from 53 % to 80 % with higher variation (CV=10.8 %) was observed. Moreover, SSH region demonstrated highest mean sand content of 74.2 % ranging between 64 % and 82 % (CV=5.0 %).

In Sub mountain and Siwalik hills region, sand had no significant correlation with other soil properties (Table 4.15). However, sand content in northeastern undulating region (Table 4.16) was positively associated with pH (r = 0.251*), whereas soil organic carbon (r = -0.218*) and silt (r = -0.658**) were inversely related to sand content. In Piedmont and alluvial plain, sand had significant correlation with pH (r = 0.283**), EC (r = 0.330**), phosphorous (r = -0.215*) and calcium carbonate (r = 0.209*) (Table 4.17). In case of Central Alluvial Plain, sand content was not affected by other properties (Table 4.18), whereas Table 4.19 showed significant relationship of sand content with zinc (r = -0.343**), silt (r = -0.549**) and clay (r = -0.460**). With high amount of sand content, clay lessivage may

occur. Charles *et al* (2015) showed similar relationship of sand with silt.

4.1.13 Silt content

The silt content under various agroecological sub regions is presented in Table 4.13 and Fig. 4.1m. They showed the following order: SSH<CAP<SWAP<NEU<PAP. In SSH region, it varied from 8 to 24 % and mean of 16 % (CV=17.6 %). CAP region exhibited a mean of 17.8 % ranging from 10 to 27 % with higher variability (CV=18.9 %). NEU region showed a mean of 21.8 % ranging between 12% and 34 % (CV=22.3 %). However, SWAP represented lower variability (CV=15.0 %) of silt content (a range from 13 % to 29 % and mean of 21.36 %). The highest silt content was found in PAP region with a range from 16 % to 29 % (mean = 22.4 % and CV=13.1 %).

In northeastern undulating region, soil organic carbon (r = 0.264**), nitrogen (r = 0.208*), manganese (r = -0.221*), sand (r = -0.658**) and clay (r = -0.277**) were significantly associated with silt content. South western alluvial plain region portrayed the strong association of silt fraction with calcium carbonate (r = -0.203*), manganese (r = -0.313**), zinc (r = 0.287**) and sand (r = -0.549**) and clay (r = -0.460**) (Table 4.19).

4.1.14 Clay content

A comparison of clay content under different agroecological sub regions is depicted in Table 4.14 and Fig. 4.1n. Various agroecological sub-regions exhibited the following order in their clay values: SWAP<SSH<CAP<NEU<PAP. In SWAP region, clay content varied from 4 % to 14 % with more variability (CV=27.2 %) and mean of 9.3 %. However, SSH region showed a mean of 9.3 ppm varying from 3 % to 16 % (CV=21.3 %). Similarly, CAP region exhibited a mean of 10.8 % ranging between 8 % and 16 % (CV=16.2 %). Also, in NEU region, the values of clay varied from 6 % to 18 % with higher variability (CV=25.6 %) and a mean of 11.2 %. However, a mean of 10.7 % ranging from 4 to 18 % (CV=23.9 %) was observed in PAP region.

In SSH region, clay had significant but negative correlation with phosphorous (r = -0.239*), (Table 4.15). Similarly, in NEU region, pH (r = -0.332**), phosphorous (r = 0.213*) and silt (r = -0.277*) were associated with clay content (Table 4.16). However, pH (r = -0.418**), electrical conductivity (r = -0.414**), potassium (r = 0.272**), copper (r = -0.350**), zinc (r = -0.244*) and sand (r = -0.594**) were associated significantly with clay in PAP region (Table 4.17). Table 4.18 showed that electrical conductivity (r = 0.23*), soil organic carbon (r = -0.26**), potassium (r = 0.30**), copper (r = -0.20*), zinc (r = 0.25*) and silt content (r = -0.20*) were strongly associated with clay content in CAP region. However, clay content had a positive dependency on electrical conductivity (r = 0.238*) and manganese (r = 0.215*) in SWAP region (Table 4.19). This suggests that under various sub-regions, nutrient availability and salt index are highly dependent on clay fraction.

Table 4.1 Descriptive statistics of soil pH under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	7.44	0.40	6.73	7.71	5.4
NEU	7.04	0.29	6.11	7.98	4.2
PAP	7.56	0.37	6.90	8.28	4.9
CAP	7.39	0.18	7.01	7.89	2.4
SWAP	7.93	0.25	7.30	8.47	3.1

Table 4.2 Descriptive statistics of electrical conductivity (EC, dS m^{-1}) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	0.24	0.09	0.12	0.51	38.2
NEU	0.17	0.10	0.04	0.78	59.1
PAP	0.55	0.50	0.13	2.04	90.8
CAP	0.59	0.48	0.11	2.04	81.2
SWAP	0.95	0.47	0.02	2.42	49.6

Table 4.3 Descriptive statistics of soil organic carbon (%) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	0.77	0.17	0.26	1.13	22.1
NEU	0.54	0.14	0.19	0.90	26.2
PAP	0.49	0.18	0.23	1.01	36.6
CAP	0.56	0.15	0.26	0.90	27.2
SWAP	0.39	0.13	0.18	0.83	32.5

Table 4.4 Descriptive statistics of available nitrogen (kg ha^{-1}) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	91.9	45.2	31.4	235.2	49.2
NEU	104.5	43.3	25.1	269.7	41.4
PAP	89.7	44.0	41.4	345.0	49.0
CAP	87.9	30.6	25.1	150.5	34.8
SWAP	77.4	37.7	15.7	244.6	48.7

Table 4.5 Descriptive statistics of Olsen-P (kg ha^{-1}) under different agroecological sub regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	20.7	5.00	1.2	29.6	24.1
NEU	40.8	11.3	15.0	72.3	27.7
PAP	34.1	10.0	20.3	72.6	29.2
CAP	22.4	3.5	13.8	31.4	15.8
SWAP	36.2	8.3	23.8	61.3	22.8

Table 4.6 Descriptive statistics of available potassium (kg ha^{-1}) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	143.9	70.4	27.8	457.3	48.9
NEU	167.7	34.8	85.3	246.9	20.8
PAP	173.2	117.2	76.3	865.0	67.7
CAP	132.3	39.8	28.8	238.0	30.1
SWAP	159.3	47.5	74.8	277.6	29.8

Table 4.7 Descriptive statistics of soil calcium carbonate (%) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	0.84	0.35	0.30	1.70	41.0
NEU	0.80	0.89	0.01	5.75	111.6
PAP	1.12	0.52	0.65	3.70	46.9
CAP	1.36	0.68	0.54	3.25	49.9
SWAP	1.31	0.51	0.60	3.00	39.1

Table 4.8 Descriptive statistics of DTPA-extractable iron content (ppm) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	5.3	2.7	2.4	15.4	51.9
NEU	9.6	3.6	3.5	21.5	37.3
PAP	19.6	12.0	6.3	62.3	61.4
CAP	14.7	4.3	5.9	24.5	29.4
SWAP	29.5	11.0	5.5	67.1	37.4

Table 4.9 Descriptive statistics of DTPA-extractable manganese content (ppm) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	4.7	2.0	1.3	11.3	43.3
NEU	7.8	2.2	4.3	14.4	27.8
PAP	5.2	1.8	2.1	10.2	35.2
CAP	5.3	2.3	1.3	12.3	42.5
SWAP	6.5	2.4	2.6	14.5	36.6

Table 4.10 Descriptive statistics of DTPA-extractable copper content (ppm) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	0.5	0.9	0.1	6.0	168.4
NEU	0.7	0.7	0.01	6.5	109.3
PAP	5.2	1.8	2.1	10.2	35.2
CAP	0.5	0.5	0.01	2.1	95.1
SWAP	1.0	0.5	0.01	3.2	50.9

Table 4.11 Descriptive statistics of DTPA-extractable zinc content (ppm) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	0.8	0.4	0.2	2.4	51.8
NEU	1.1	0.9	0.01	5.9	81.7
PAP	2.0	1.9	0.3	11.6	95.3
CAP	1.8	0.9	0.3	3.8	50.2
SWAP	3.2	2.1	0.6	11.6	66.3

Table 4.12 Descriptive statistics of sand content (%) under different agroecological sub- regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	74.2	3.7	64.0	82.0	5.0
NEU	67.3	4.7	55.0	76.0	6.9
PAP	66.4	4.2	54.0	76.0	6.3
CAP	71.0	7.7	53.0	80.0	10.8
SWAP	68.9	3.6	58.8	77.0	5.2

Table 4.13 Descriptive statistics of silt content (%) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	16.0	2.8	8.0	24.0	17.6
NEU	21.8	4.9	12.0	34.0	22.3
PAP	22.4	3.4	16.7	29.0	15.2
CAP	17.8	3.4	10.0	27.0	18.9
SWAP	21.6	3.3	13.0	29.0	15.2

Table 4.14 Descriptive statistics of clay content (%) under different agroecological sub-regions

Agroecological sub-region (N=100)	Mean	Std Dev	Minimum	Maximum	CV (%)
SSH	10.0	2.1	3.0	16.0	21.3
NEU	11.2	2.9	6.0	18.0	25.6
PAP	10.7	2.6	4.0	18.0	23.9
CAP	10.8	1.8	8.0	16.2	16.2
SWAP	9.3	2.5	4.0	14.0	27.21

Table 4.15 Correlation among soil fertility parameters under Sub mountain Siwalik Hills (SSH)

SSH	pH	EC	SOC	N	P	K	CaCO3	Fe	Mn	Cu	Zn	Sand	Silt	Clay
pH	1.000													
EC	-0.169	1.000												
SOC	0.128	0.072	1.000											
N	0.089	0.213*	0.100	1.000										
P	0.092	-0.037	0.019	0.033	1.000									
K	0.177	0.151	0.217*	0.167	0.145	1.000								
CaCO3	0.082	-0.044	0.060	0.372**	-0.005	-0.029	1.000							
Fe	0.055	0.010	0.000	0.146	-0.110	0.065	0.021	1.000						
Mn	0.052	0.054	0.025	0.117	-0.050	0.157	-0.200	0.359**	1.000					
Cu	0.187	-0.042	-0.062	-0.114	-0.223	-0.011	-0.011	0.039	0.027	1.000				
Zn	0.167	0.002	0.065	0.220	-0.033	0.139	0.100	0.100	0.215	0.130	1.000			
Sand	0.090	-0.090	0.089	0.105	0.059	0.195	-0.041	0.124	0.192	0.092	0.144	1.000		
Silt	0.078	0.038	-0.137	-0.140	0.107	-0.172	-0.075	-0.128	-0.220	-0.096	-0.223	-0.748	1.000	
Clay	-0.123	0.008	-0.010	-0.125	-0.239	-0.175	0.124	-0.010	-0.096	0.073	-0.139	-0.688	0.244	1.000

***Significant at 5%**
****Significant at 1%**

Table 4.16 Correlation among soil fertility parameters under North Eastern Undulating region (NEU)

NEU	pH	EC	SOC	N	P	K	$CaCO_3$	Fe	Mn	Cu	Zn	Sand	Silt	Clay
pH	1.000													
EC	-0.182	1.000												
SOC	0.033	0.026	1.000											
N	0.139	0.073	0.238*	1.000										
P	0.243*	-0.213*	0.084	0.039	1.000									
K	-0.013	-0.040	0.020	0.152	-0.024	1.000								
$CaCO_3$	-0.312**	0.128	-0.057	0.015	0.011	0.100	1.000							
Fe	-0.292**	-0.002	-0.206*	-0.010	-0.073	0.074	0.111	1.000						
Mn	0.103	0.109	0.106	-0.007	0.093	-0.031	-0.002	-0.156	1.000					
Cu	-0.063	-0.039	0.099	-0.048	-0.096	-0.051	0.200*	0.056	0.002	1.000				
Zn	0.038	-0.036	0.220*	-0.079	0.126	0.022	0.020	-0.010	0.313**	-0.029	1.000			
Sand	0.251*	-0.163	-0.218*	-0.089	0.097	0.101	-0.146	0.096	0.089	-0.193	-0.042	1.000		
Silt	-0.120	0.081	0.264**	0.208*	-0.080	0.044	0.090	-0.015	-0.221*	0.147	-0.028	-0.658**	1.000	
Clay	-0.332**	0.089	-0.083	-0.162	-0.213*	-0.147	0.120	-0.002	0.061	-0.078	0.093	-0.247	-0.277**	1.000

***Significant at 5%**

****Significant at 1%**

Table 4.17 Correlation among soil fertility parameters under Piedmont and Alluvial Plain (PAP)

PAP	pH	EC	SOC	N	P	K	$CaCO_3$	Fe	Mn	Cu	Zn	Sand	Silt	Clay
pH	1.000													
EC	0.651**	1.000												
SOC	-0.312**	-0.194	1.000											
N	-0.101	-0.062*	0.239	1.000										
P	-0.023	-0.081	-0.079	0.420**	1.000									
K	-0.125	0.029	0.273**	0.150	0.034	1.000								
$CaCO_3$	0.295**	0.258**	0.248*	0.028	-0.092	-0.005	1.000							
Fe	0.482**	0.548**	-0.291**	0.039	0.053	0.068	0.141	1.000						
Mn	0.377**	0.269**	-0.349**	-0.064	-0.053	-0.169	0.056	0.369**	1.000					
Cu	0.474**	0.514**	-0.232*	-0.057	-0.071	-0.008	0.106	0.366**	0.131	1.000				
Zn	0.505**	0.420**	-0.186	0.001	0.039	-0.025	0.013	0.202*	0.223*	0.322**	1.000			
Sand	0.283**	0.330**	-0.069	-0.047	-0.215*	-0.012	0.209*	0.168	0.010	0.194	0.042	1.000		
Silt	-0.198	-0.026	0.031	-0.038	0.099	-0.095	-0.175	-0.022	-0.104	-0.134	0.052	-0.534**	1.000	
Clay	-0.418**	-0.414**	0.170	0.075	0.174	0.272**	-0.119	-0.319	-0.117	-0.350**	-0.244*	-0.594**	0.036	1.000

***Significant at 5%**

****Significant at 1%**

Table 4.18 Correlation among soil fertility parameters under Central Alluvial Plain (CAP)

CAP	pH	EC	SOC	N	P	K	$CaCO_3$	Fe	Mn	Cu	Zn	Sand	Silt	Clay
pH	1.00													
EC	-0.08	1.00												
SOC	0.07	-0.11	1.00											
N	-0.11	0.06	-0.11	1.00										
P	-0.12	-0.02	0.00	-0.04	1.00									
K	0.03	0.19	-0.16	0.07	-0.27**	1.00								
$CaCO_3$	0.11	0.02	0.07	-0.02	0.05	-0.11	1.00							
Fe	-0.09	0.06	-0.08	0.00	-0.16	0.20*	-0.14	1.00						
Mn	-0.01	0.04	-0.02	0.07	-0.10	-0.06	0.09	0.15	1.00					
Cu	0.18	-0.65**	0.00	-0.07	-0.10	-0.05	-0.15	0.18	0.01	1.00				
Zn	-0.17	0.03	-0.30**	-0.03	-0.02	0.29**	-0.13	0.27	0.05	0.16	1.00			
Sand	0.04	-0.02	0.16	-0.10	0.05	-0.10	0.12	-0.01	0.23	0.01	-0.16	1.00		
Silt	-0.03	-0.06	-0.06	-0.04	0.19	0.06	-0.06	0.03	-0.05	0.11	-0.06	-0.21*	1.00	
Clay	-0.08	0.23*	-0.26**	-0.05	-0.19	0.30**	-0.21*	0.05	-0.03	-0.20*	0.25*	-0.10	-0.20*	1.00

***Significant at 5%**
****Significant at 1%**

Table 4.19 Correlation among soil fertility parameters under South Western Alluvial Plain (SWAP)

SWAP	pH	EC	SOC	N	P	K	$CaCO_3$	Fe	Mn	Cu	Zn	Sand	Silt	Clay
pH	1													
EC	-0.044	1												
SOC	0.007	0.03	1											
N	0.236*	0.104	0.176	1										
P	0.022	0.235*	0.287**	0.143	1									
K	-0.104	-0.035	0.173	0.058	0.003	1								
$CaCO_3$	0.117	-0.026	0.035	0.135	0.044	0.045	1							
Fe	-0.063	-0.021	0.136	-0.024	0.031	0.01	0.158	1						
Mn	-0.013	-0.145	0.319**	0.339**	-0.016	0.116	0.116	0.082	1					
Cu	-0.145	0.263**	0.231**	0.143	0.254*	0.155	-0.059	0.196	0.108	1				
Zn	0.059	0.086	0.118	-0.179	0.064	-0.115	-0.104	0.115	-0.241*	-0.267**	1			
Sand	-0.156	-0.176	0.012	-0.06	0.018	0.09	0.109	-0.068	0.077	-0.015	-0.343**	1		
Silt	0.126	-0.043	-0.134	-0.111	0.063	-0.188	-0.203*	0.058	-0.313**	-0.099	0.287**	-0.549**	1	
Clay	-0.009	0.238*	-0.036	0.076	-0.084	0.194	0.05	0.021	0.215*	0.065	0.067	-0.460**	-0.196	1

***Significant at 5%**
****Significant at 1%**

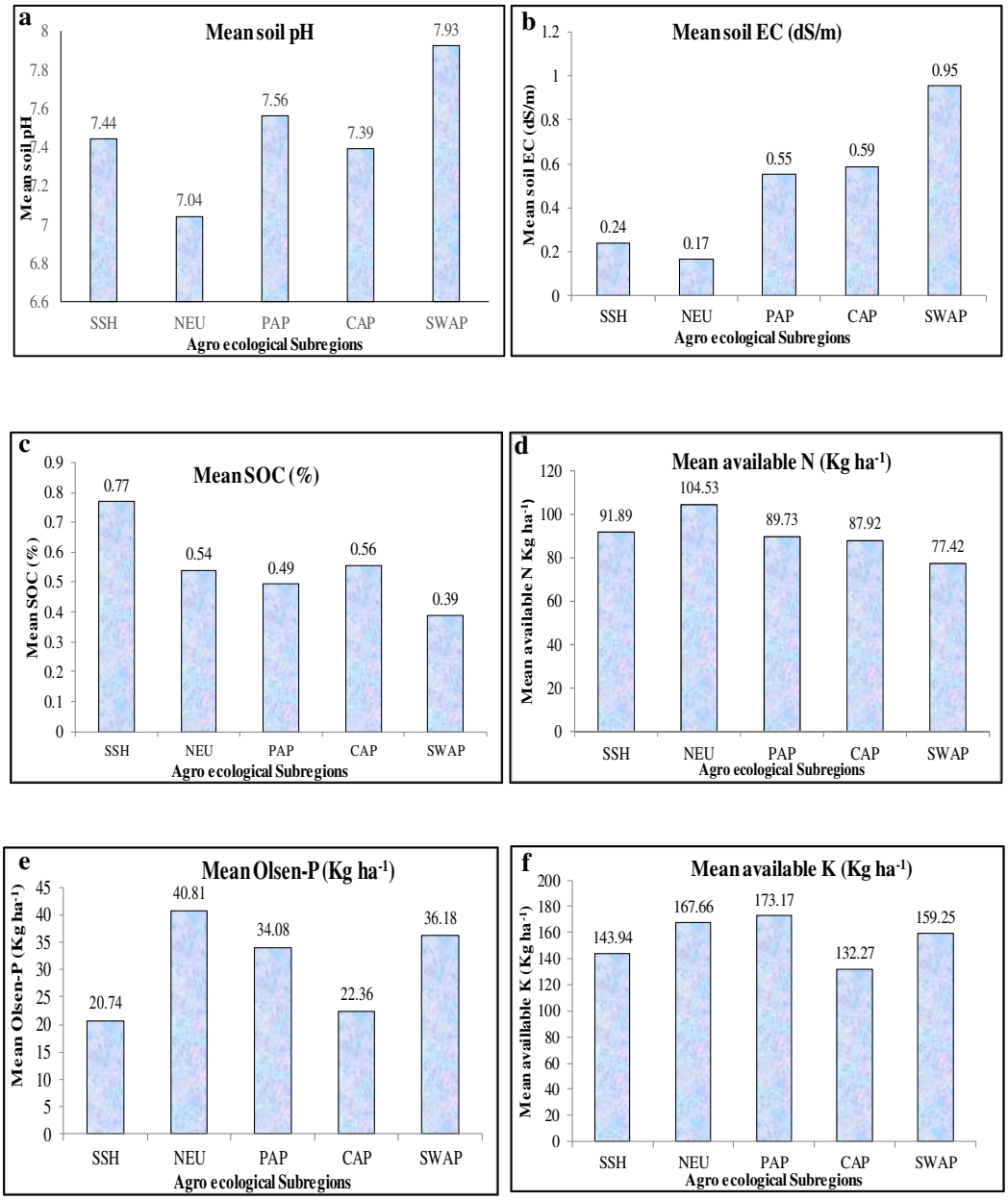


Fig. 4.1 Soil properties under different agroecological sub-regions (a) Soil pH (b) electrical conductivity (EC, dS m^{-1}) (c) Soil organic carbon content (%) (d) Available nitrogen content (kg ha^{-1}) (e) Olsen-phosphorous content (kg ha^{-1}) (f) Available potassium content (kg ha^{-1})

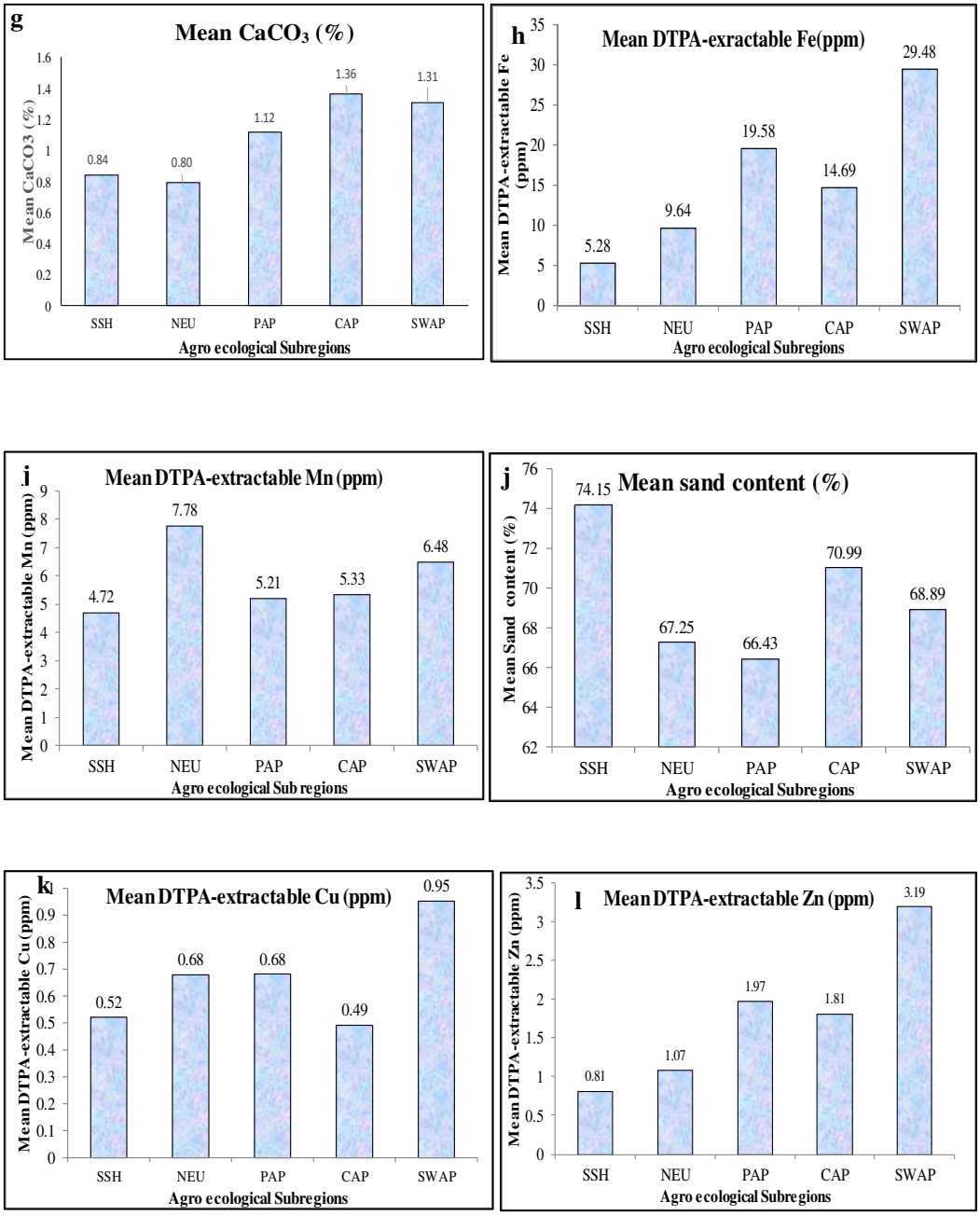


Fig. 4.1 Calcium carbonate and available micronutrient status of soils of different agroecological sub-regions of (g) Calcium carbonate content (%) (h) DTPA-extractable iron content (ppm) (i) DTPA-extractable manganese content (ppm) (j) DTPA-extractable copper content (ppm) (k) DTPA-extractable zinc content (ppm) (l) Sand content (%)

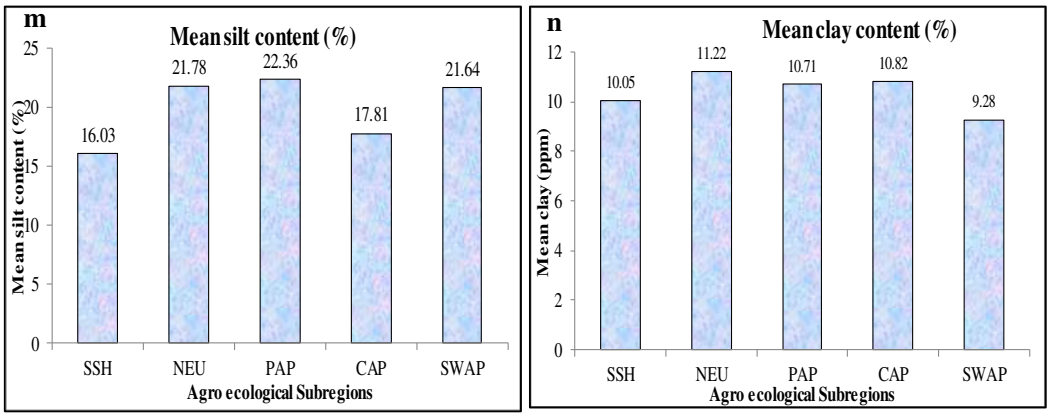


Fig. 4.1 Soil properties under different agroecological sub-regions of (m) Silt content (%) (n) Clay content (%)

4.1.15 Soil fertility maps

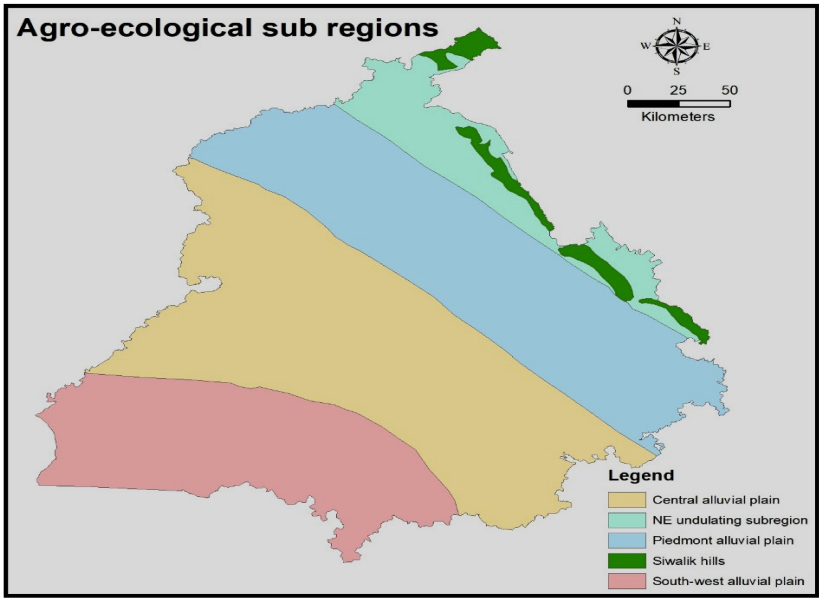


Fig. 4.2 Map showing agroecological sub-regions of Punjab

The soil fertility maps (Fig 4.3 to 4.16) were prepared to show soil fertility status across various agroecological sub regions in the state of Punjab by using ARC-GIS software. Different soil properties like pH, electrical conductivity, soil organic carbon, available nitrogen, phosphorous, potassium, calcium carbonate, DTPA-extractable iron, manganese, copper, zinc (chemical properties), sand, silt and clay (physical properties) were used for soil fertility maps.

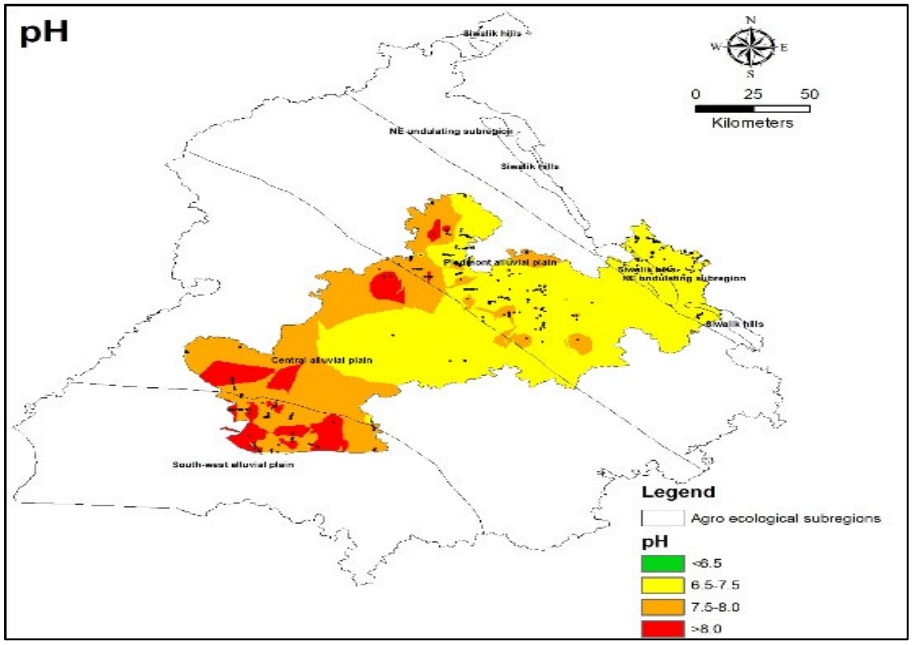


Fig. 4.3 Map showing pH under different agroecological sub-regions

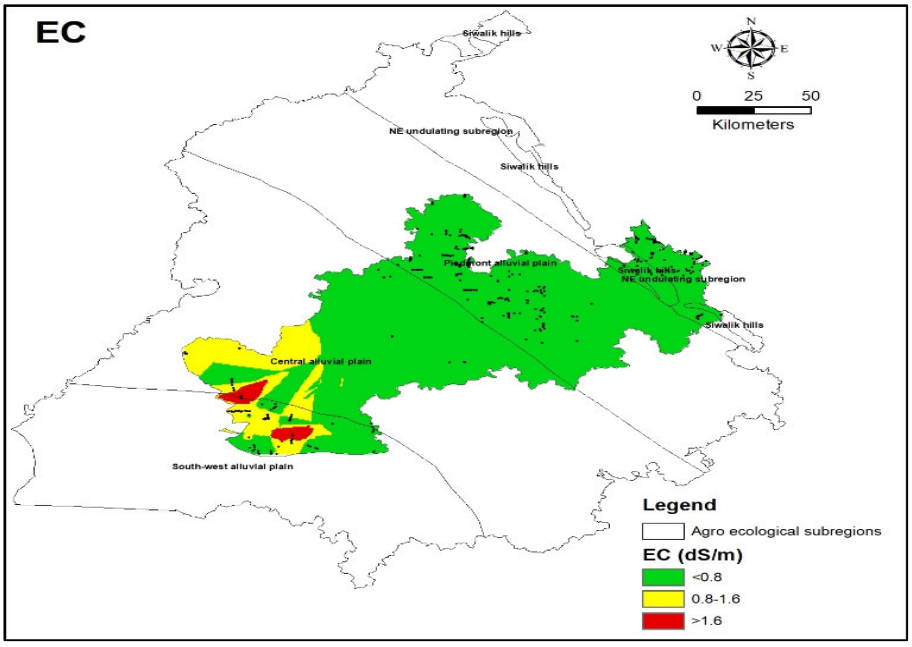


Fig. 4.4 Map showing electrical conductivity (EC dS m^{-1}) under different agroecological sub-regions

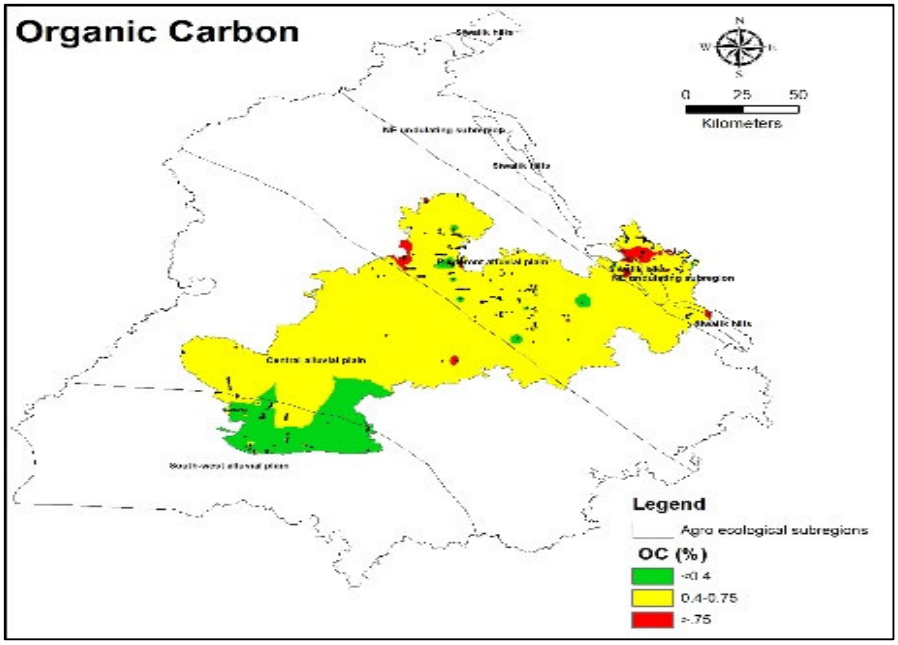


Fig. 4.5 Map showing soil organic carbon (%) under different agroecological sub- regions

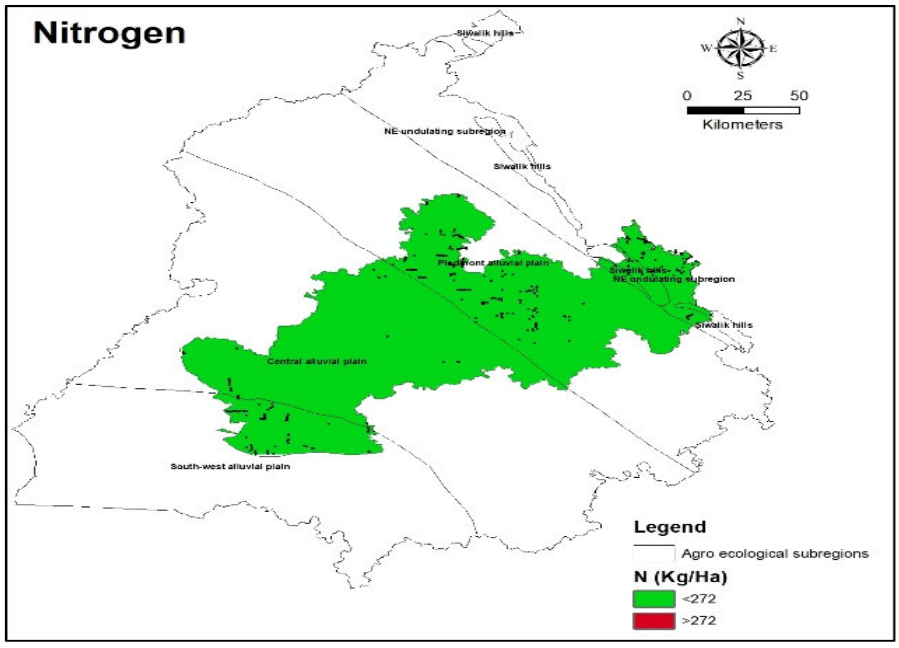


Fig. 4.6 Map showing soil available nitrogen (kg ha^{-1}) under different agroecological sub-regions

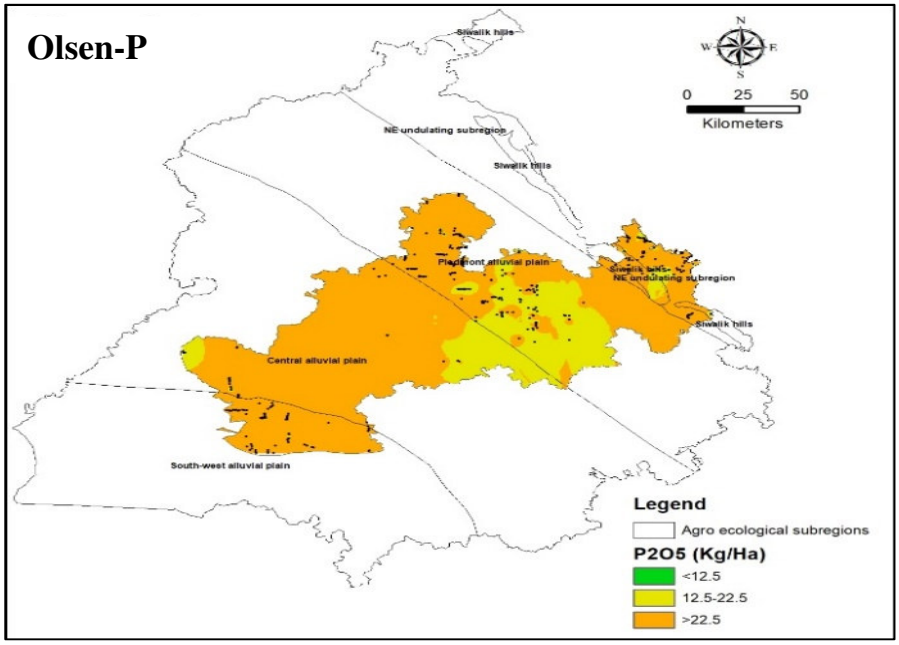


Fig. 4.7 Map showing Olsen-phosphorous (kg ha^{-1}) under different agroecological sub-regions

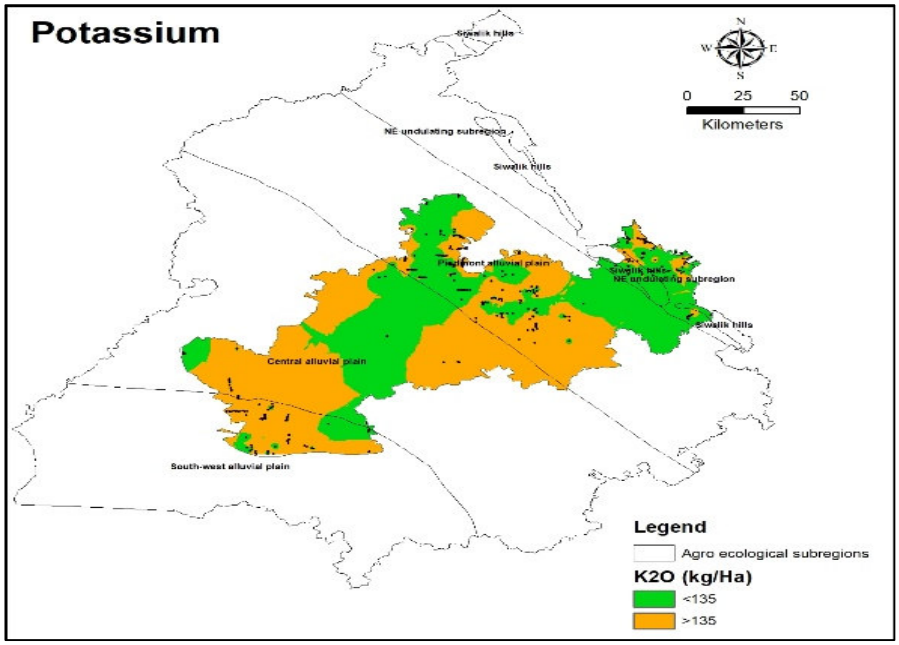


Fig. 4.8 Map showing soil available potassium (kg ha^{-1}) under different agroecological sub-regions

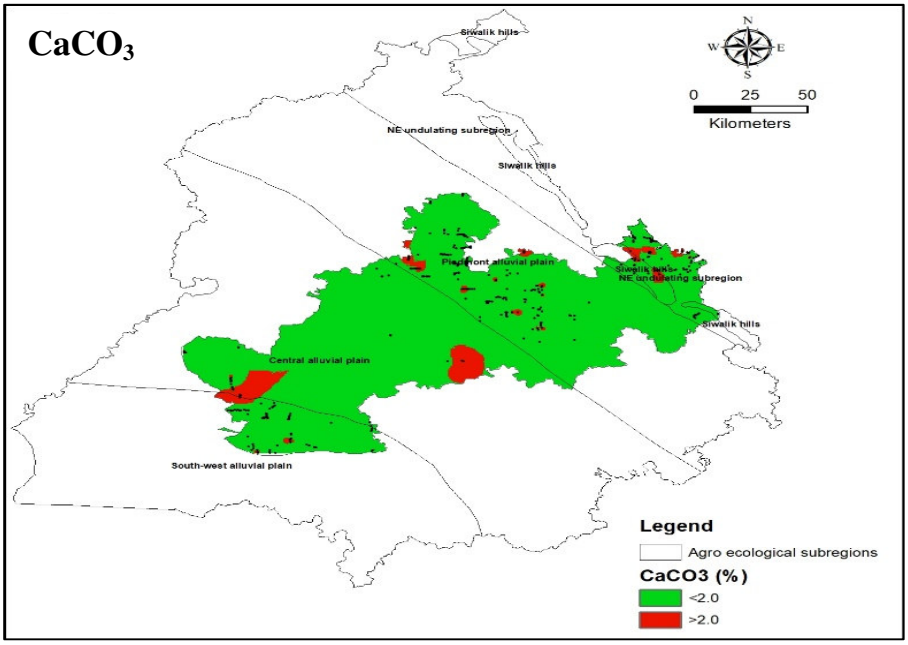


Fig. 4.9 Map showing soil calcium carbonate (%) under different agroecological sub-regions

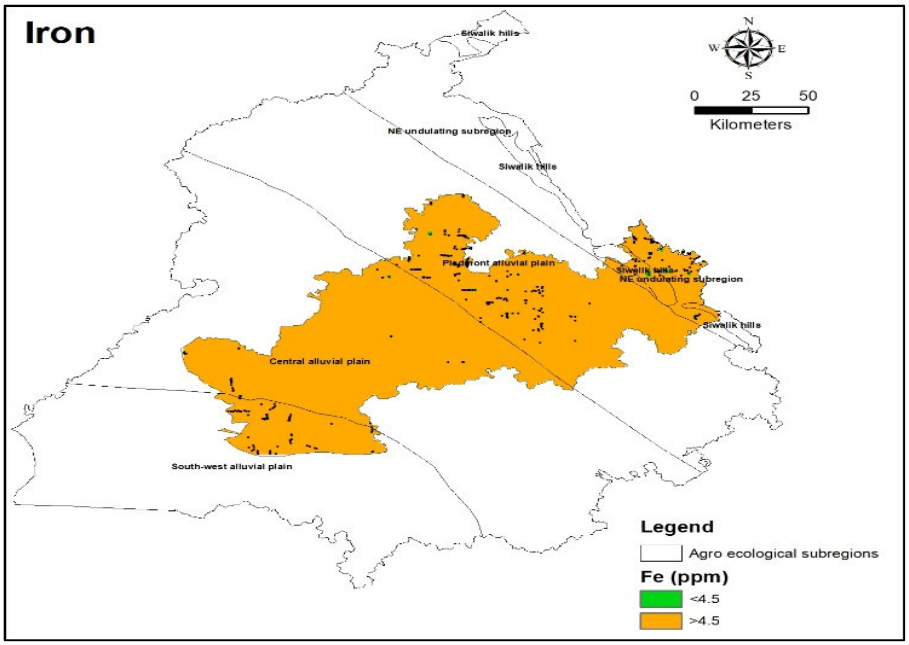


Fig. 4.10 Map showing DTPA-extractable iron content (ppm) under different agroecological sub-regions

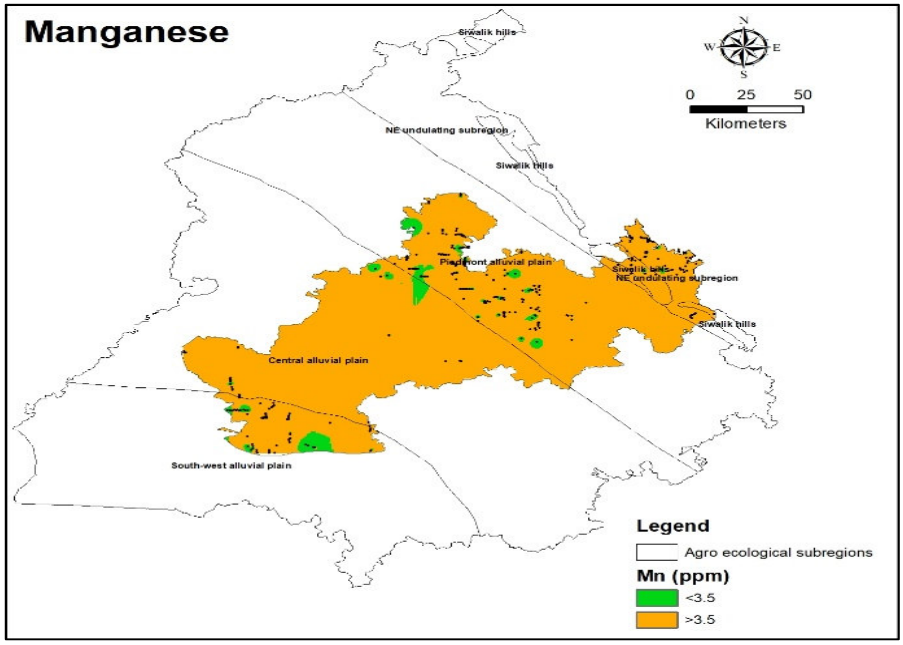


Fig. 4.11 Map showing DTPA-extractable manganese content (ppm) under different agroecological sub-regions

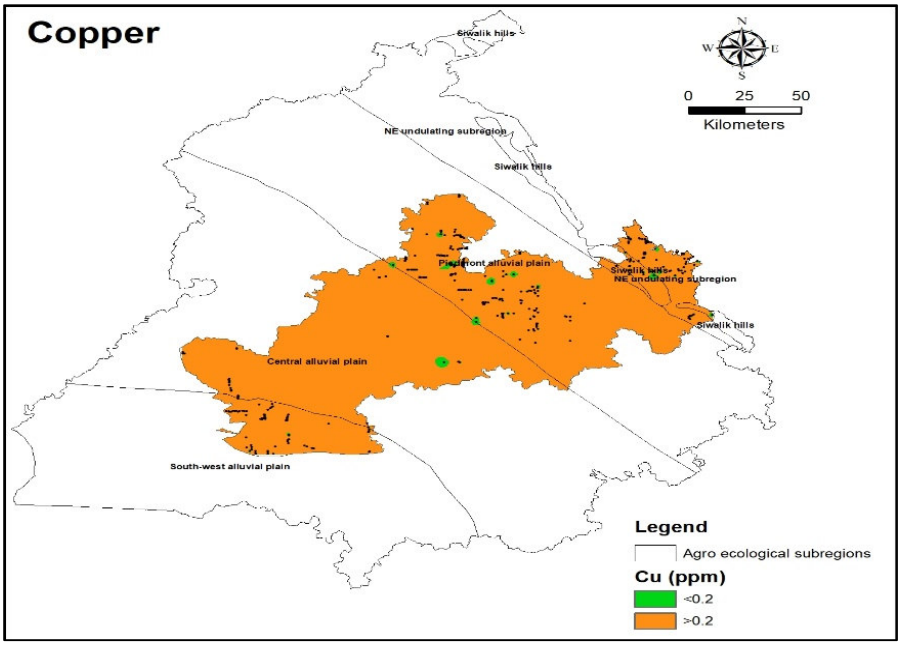


Fig. 4.12 Map showing DTPA-extractable copper content (ppm) under different agroecological sub-regions

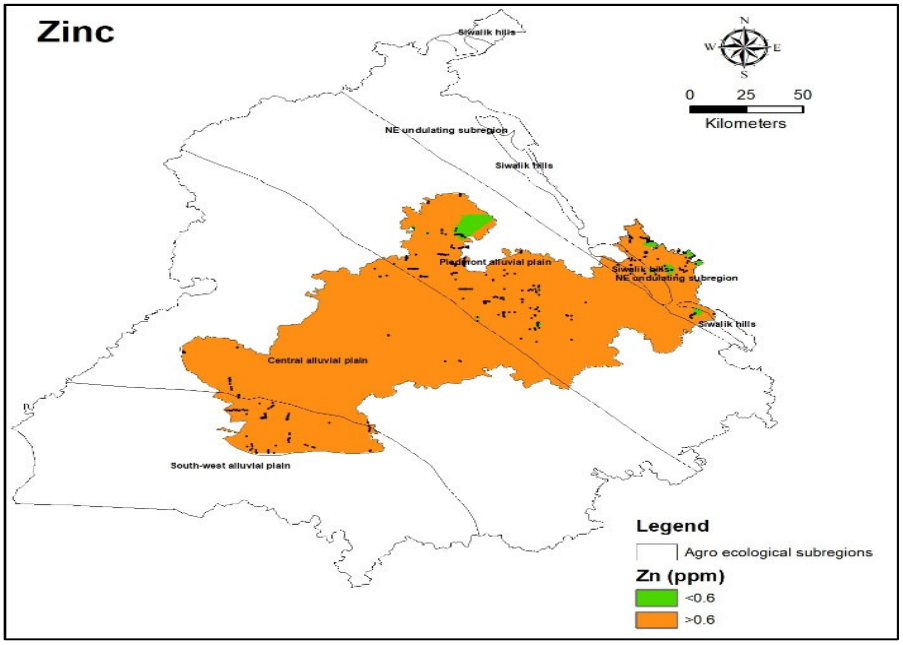


Fig. 4.13 Map showing DTPA-extractable zinc content (ppm) under different agroecological sub-regions

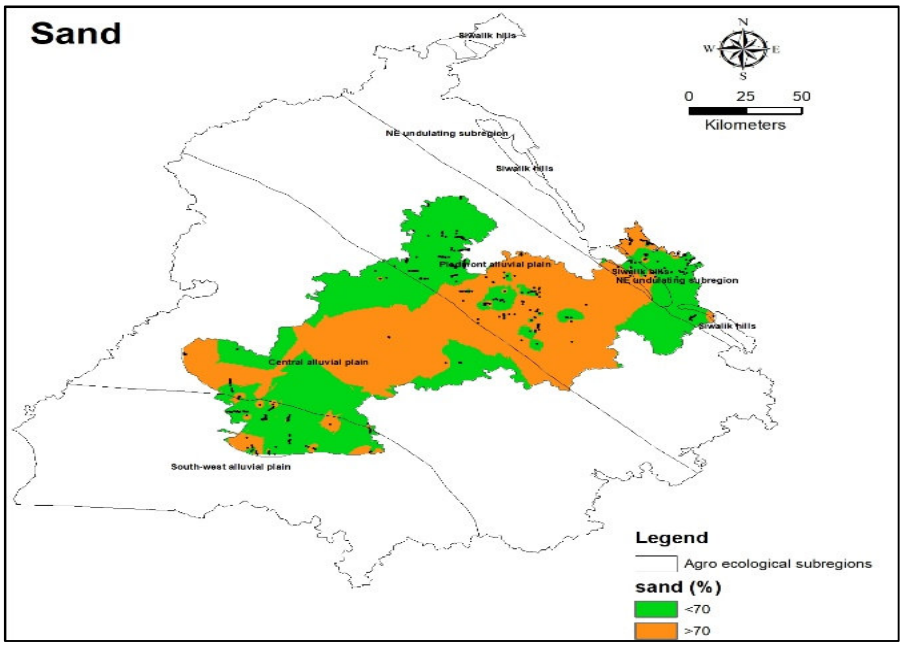


Fig. 4.14 Map showing sand content (%) under different agroecological sub-regions

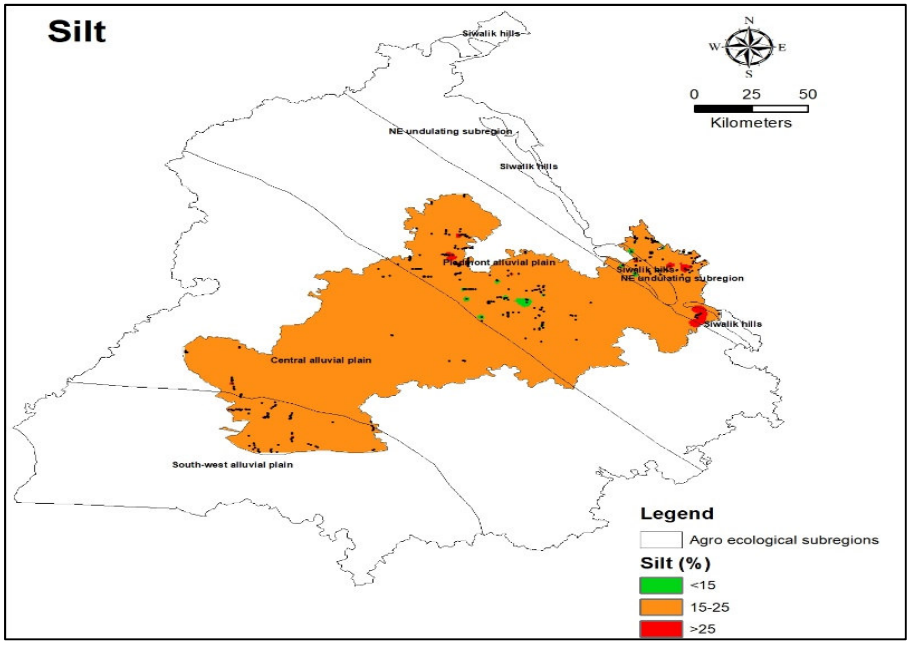


Fig. 4.15 Map showing silt content (%) under different agroecological sub-regions

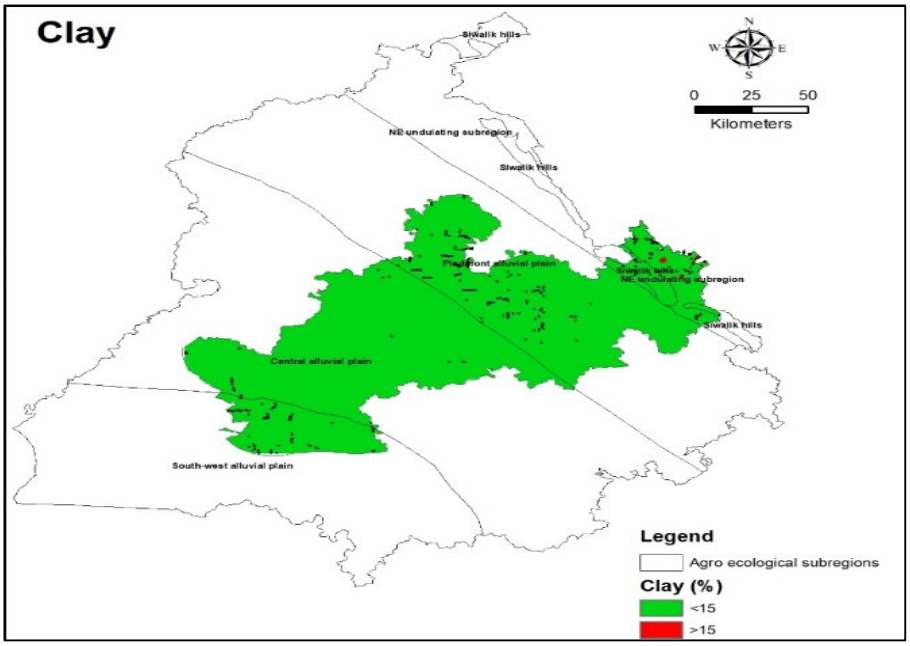
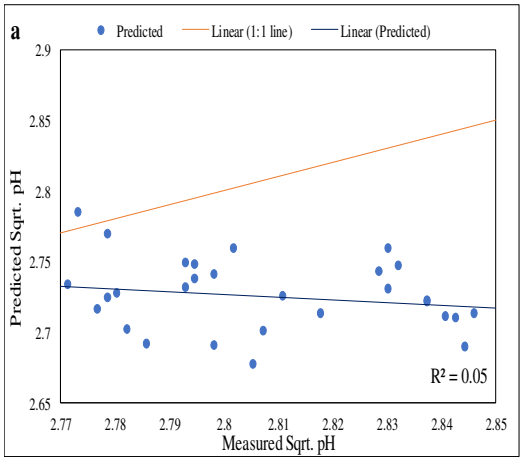
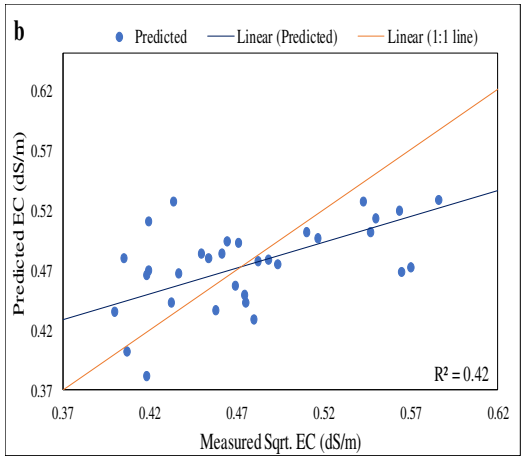


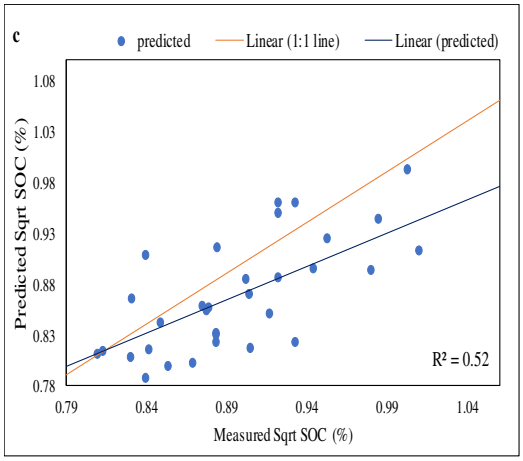
Fig. 4.16 Map showing clay content (%) under different agroecological sub-regions

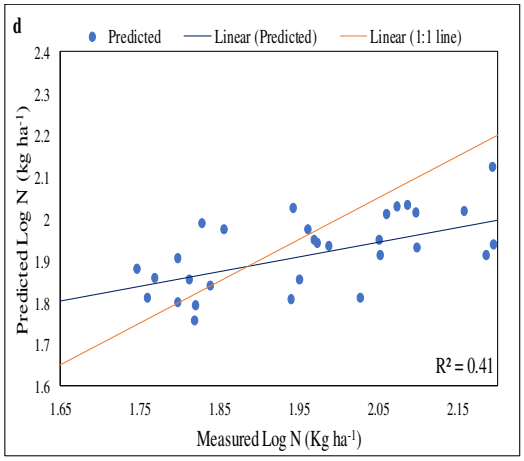
4.2 Modelling of hyperspectral data for prediction of soil properties

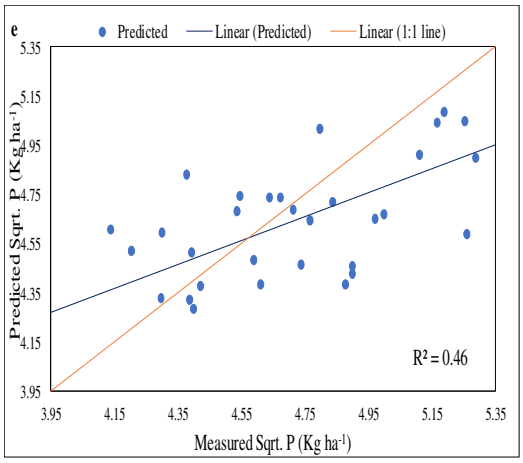
4.2.1 Sub-mountain Siwalik Hills

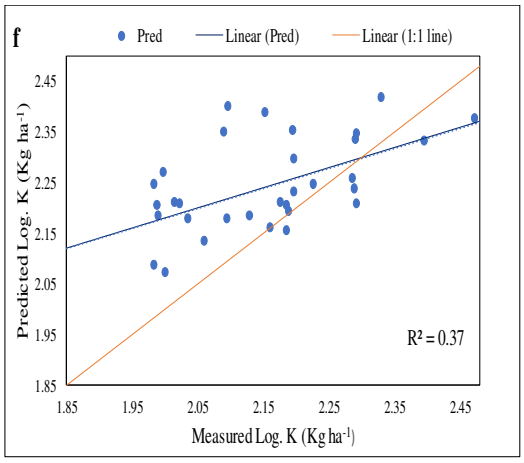


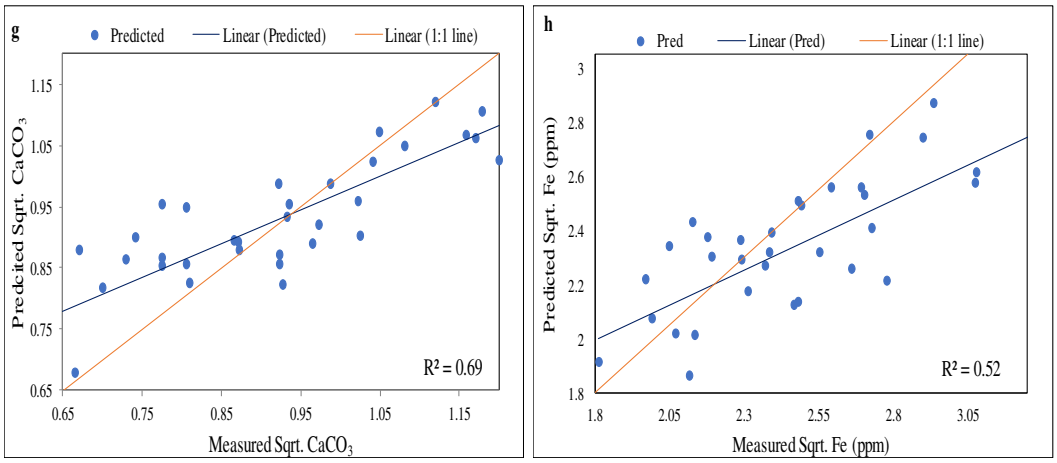


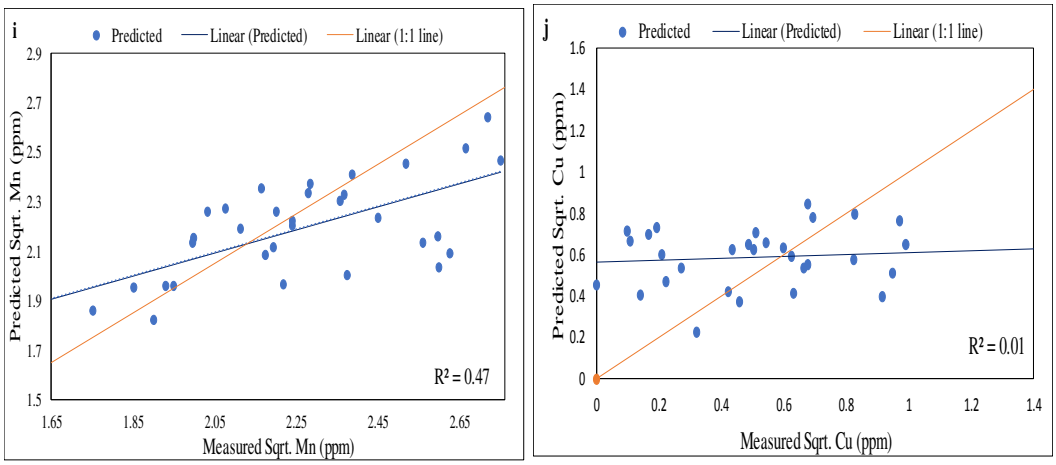


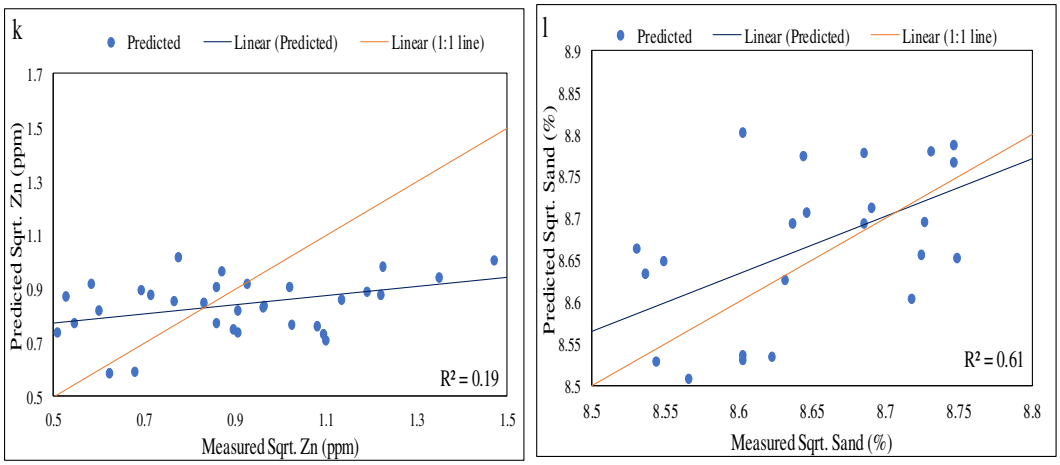







g
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. CaCO3
Measured Sqrt. CaCO3
R² = 0.69
h
Pred
Linear (Pred)
Linear (1:1 line)
Predicted Sqrt. Fe (ppm)
Measured Sqrt. Fe (ppm)
R² = 0.52


i
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Mn (ppm)
Measured Sqrt. Mn (ppm)
R² = 0.47
j
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Cu (ppm)
Measured Sqrt. Cu (ppm)
R² = 0.01


k
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Zn (ppm)
Measured Sqrt. Zn (ppm)
R² = 0.19
l
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Sand (%)
Measured Sqrt. Sand (%)
R² = 0.61

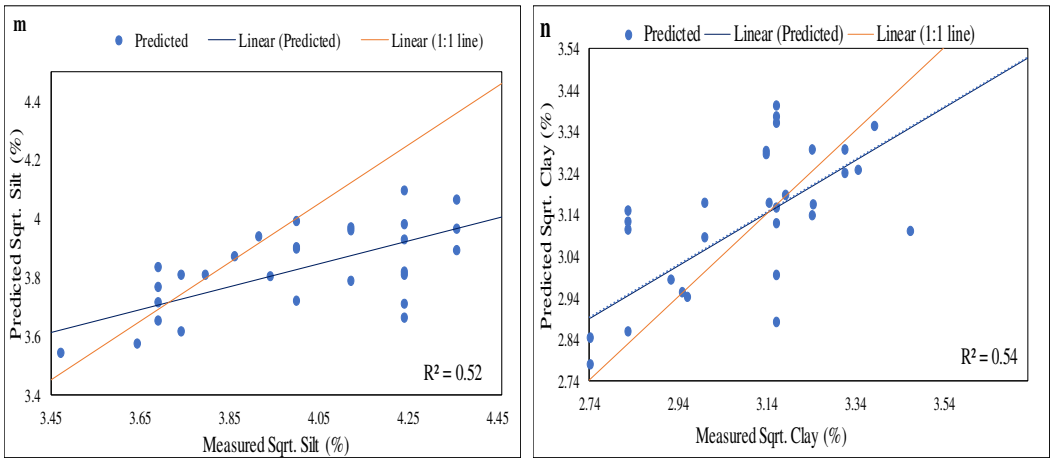


Fig. 4.17 Measured and predicted values, fitted linear line and 1:1 line of (a) Square root pH (b) Square root EC (c) Square root SOC (d) log N (e) Square root P (f) log K (g) Square root $CaCO_3$ (h) Square root Fe (i) Square root Mn (j) Square root Cu (k) Square root Zn (l) Square root Sand (m) Square root Silt (n) Square root Clay for Sub-mountain Siwalik hills.

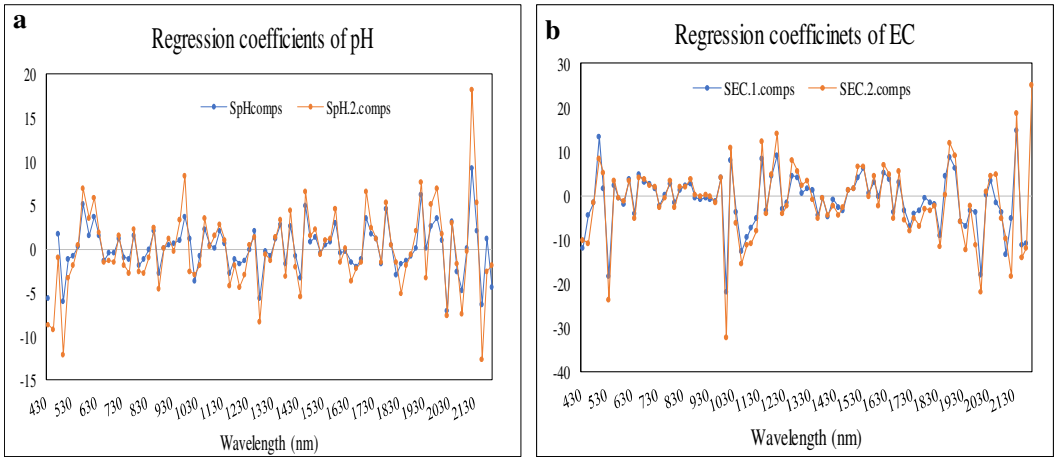
The standard deviation (SD), root mean square error (RMSE), ratio of performance deviation (RPD) and R^2 value ranged between 0.03-0.37, 0.05-0.34, 0.32-1.73 and 0.05-0.69, respectively among all the soil physical and chemical properties for the soil samples collected from SSH region. The R^2 value for the $CaCO_3$ (0.69), sand (0.61), silt (0.52), SOC (0.52) and clay (0.54) was, more than 0.5, whereas the R^2 value for DTPA extractable Cu and pH was very low. Alike the R^2 value, the RPD value was also observed higher for $CaCO_3$ (1.73), sand (1.58) and clay (1.43). Although, the RMSEP value was observed less for soil pH, the RPD and R^2 were also observed less. With the least variability in the soil pH, the model predicted higher pH for lower observed soil pH and lower pH for higher observed soil pH (Fig 4.1a). So, considering all the statistical parameters, we can recommend the use of developed spectral model for prediction of sand, $CaCO_3$ and clay content for sub-mountain Siwalik hill. The R^2 value, RMSE, SD and RPD for all parameter are presented in Table 4.20

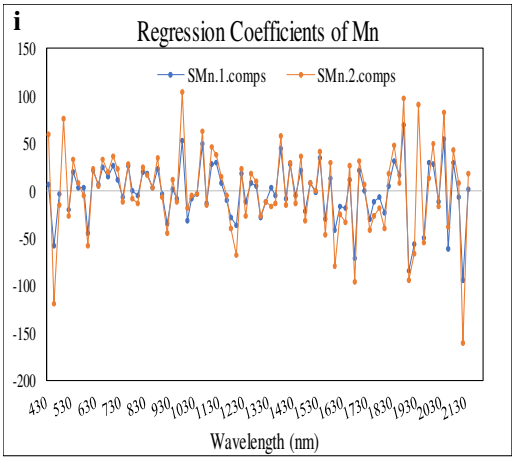
Table 4.20 Summary of statistics for the spectral models developed by partial least square regression for validation (test data set) of Sub-mountain Siwalik hill

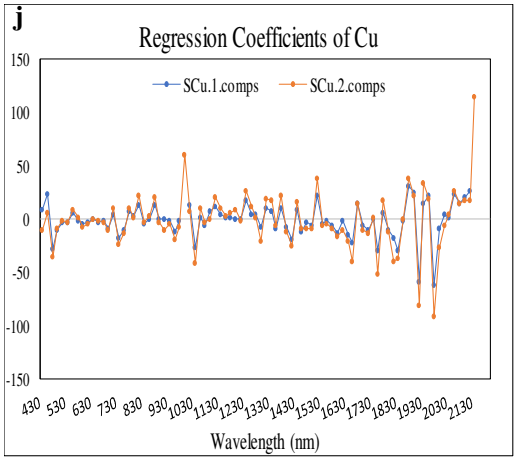
Soil property	Statistical parameter			
	SD	RMSE	RPD	R^2
pH*	0.03	0.05	0.32	0.05
EC* (dS/m)	0.06	0.12	1.33	0.42
SOC (%)	0.06	0.06	1.18	0.52
N# (kg ha^{-1})	0.17	0.15	1.20	0.41
P* (kg ha^{-1})	0.37	0.36	1.35	0.46
K# (kg ha^{-1})	0.15	0.06	1.04	0.37
$CaCO_3$ *(%)	0.16	0.14	1.73	0.69
DTPA-Fe* (ppm)	0.37	0.36	1.33	0.51
DTPA-Mn* (ppm)	0.30	0.28	1.35	0.47
DTPA-Cu* (ppm)	0.32	0.2	0.95	0.01
DTPA-Zn* (ppm)	0.27	0.28	1.08	0.19
Sand* (%)	0.13	0.19	1.58	0.61
Silt* (%)	0.30	0.39	1.10	0.52
Clay* (%)	0.24	0.24	1.43	0.54

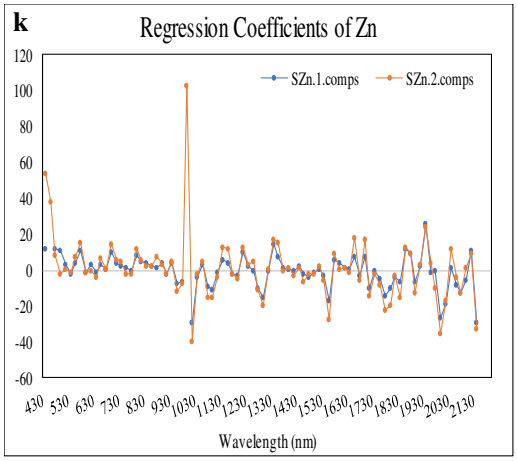
(*) Square root transformation
(#) Log transformation

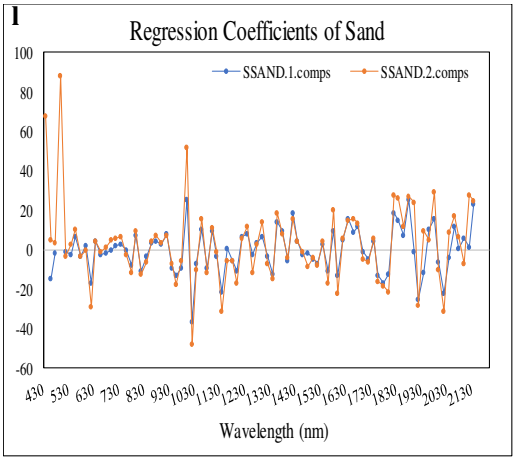
The regression coefficients of the fitted models for the Sub-mountain Siwalik Hills are presented below

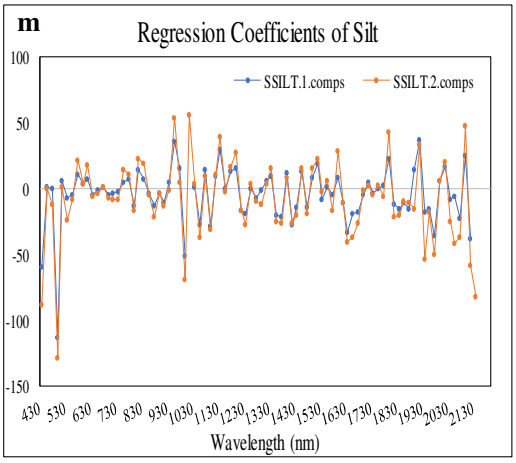


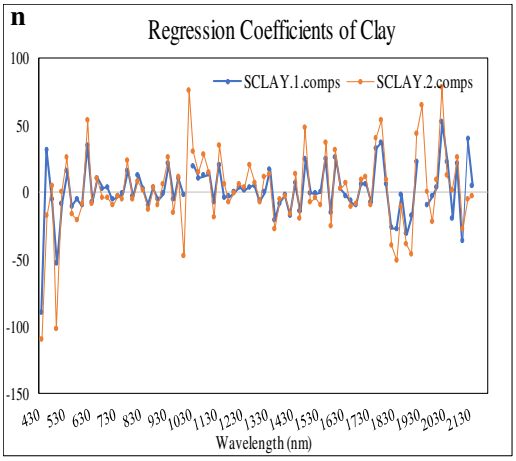

i
Regression Coefficients of Mn
SMn.1.comps
SMn.2.comps
Wavelength (nm)


j
Regression Coefficients of Cu
SCu.1.comps
SCu.2.comps
Wavelength (nm)


k
Regression Coefficients of Zn
SZn.1.comps
SZn.2.comps
Wavelength (nm)


l
Regression Coefficients of Sand
SSAND.1.comps
SSAND.2.comps
Wavelength (nm)


m
Regression Coefficients of Silt
SSILT.1.comps
SSILT.2.comps
Wavelength (nm)


n
Regression Coefficients of Clay
SCLAY.1.comps
SCLAY.2.comps
Wavelength (nm)

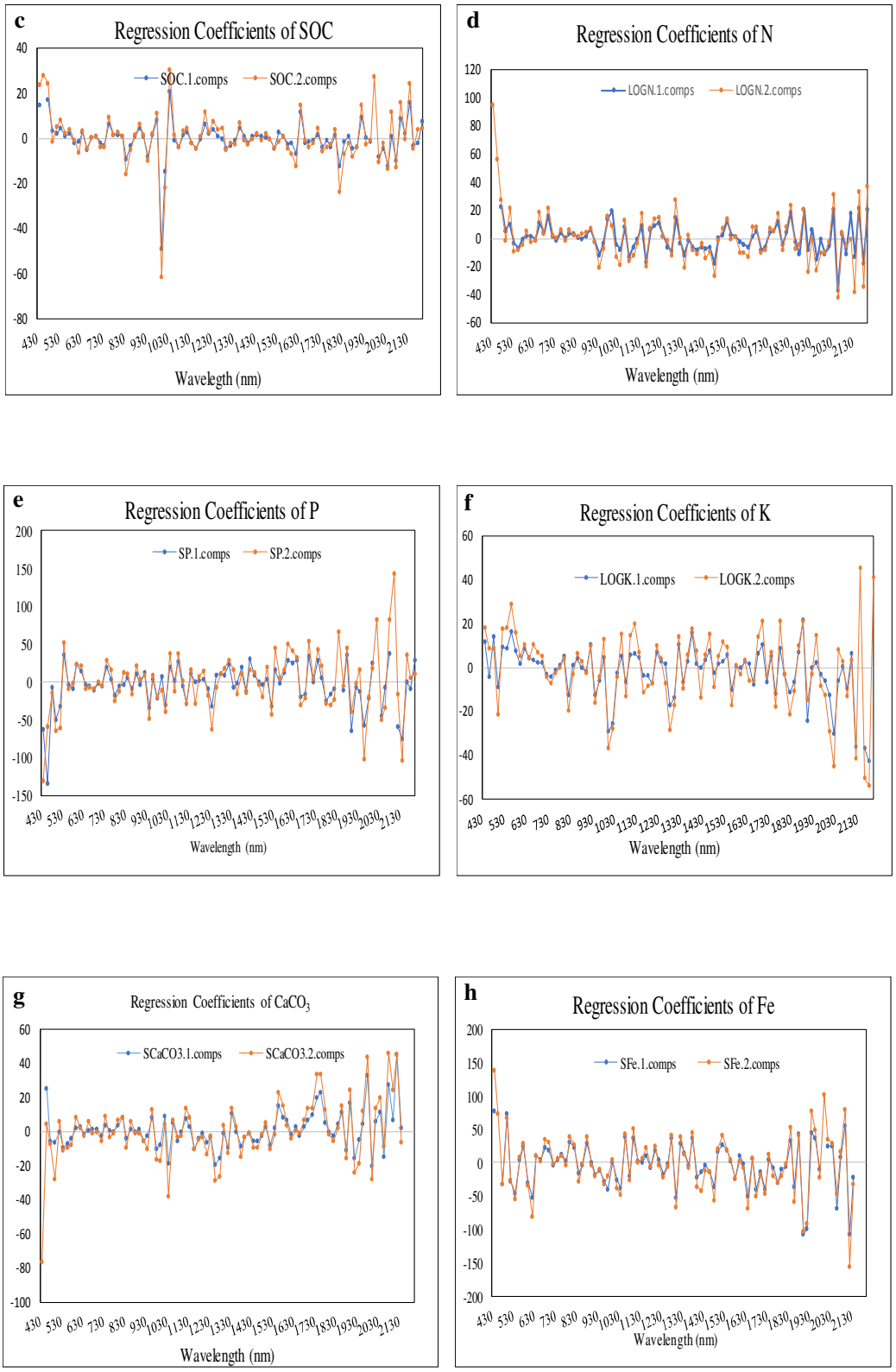
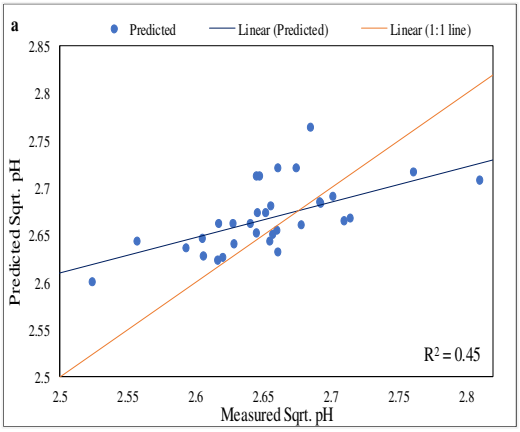
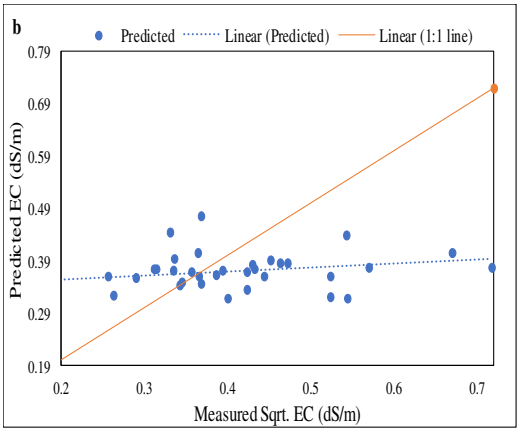
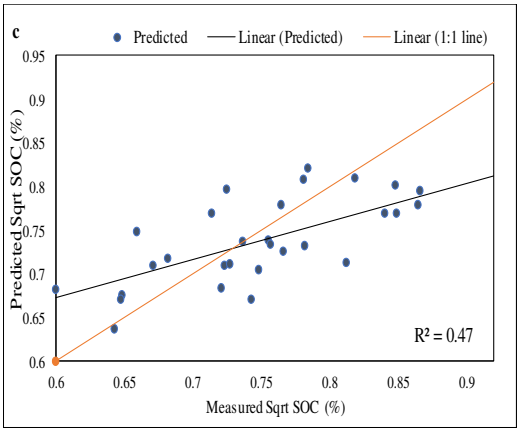


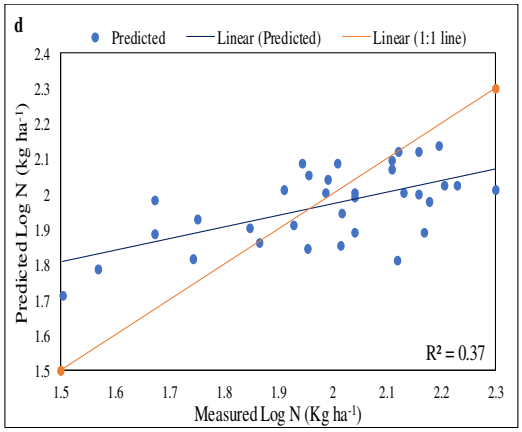
Fig. 4.18 Regression coefficients of (a) Soil pH (b) EC (c) SOC (d) Available N (e) Olsen-P (f) Available K (g) Soil calcium carbonate (h) DTPA-Fe (i) DTPA-Mn (j) DTPA-Cu (k) DTPA-Zn (l) Sand (m) Silt (n) Clay for Sub-mountain Siwalik Hills

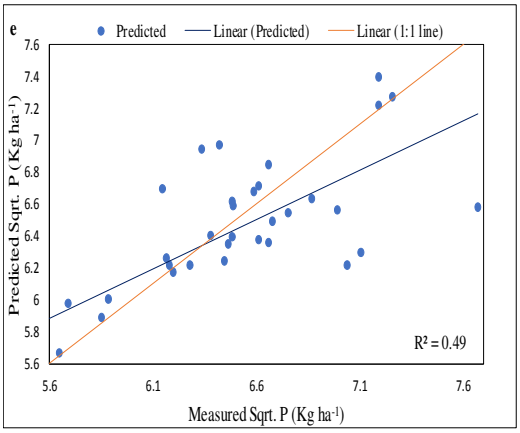
4.2.2 Northeastern Undulating

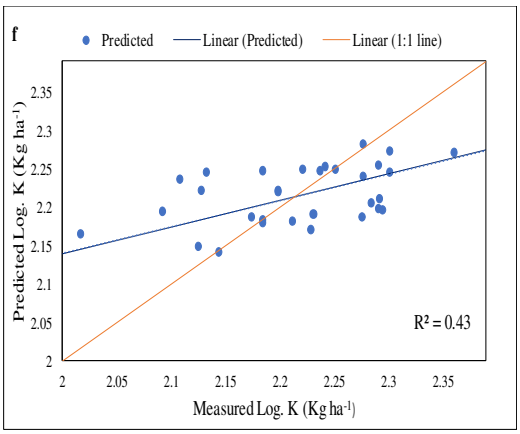


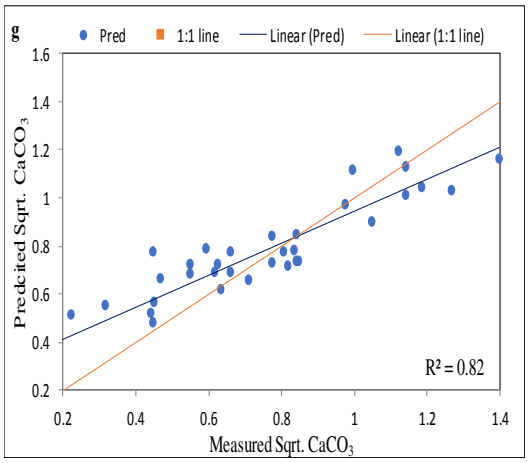


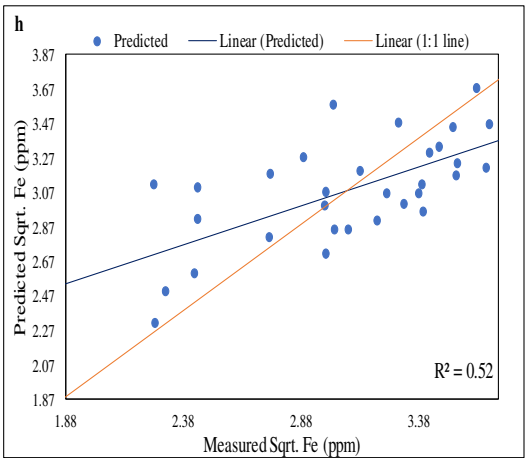


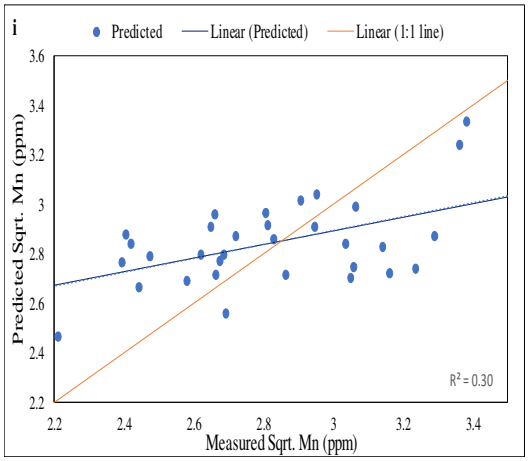


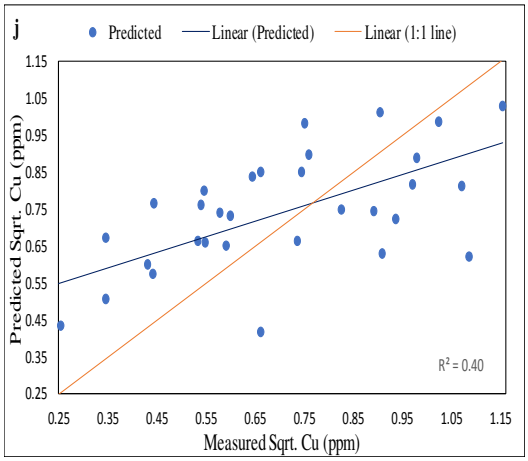


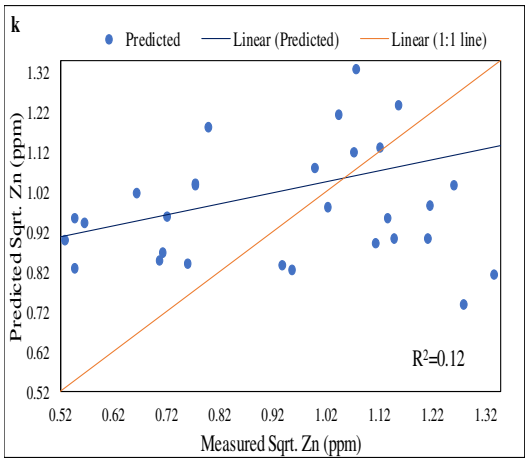


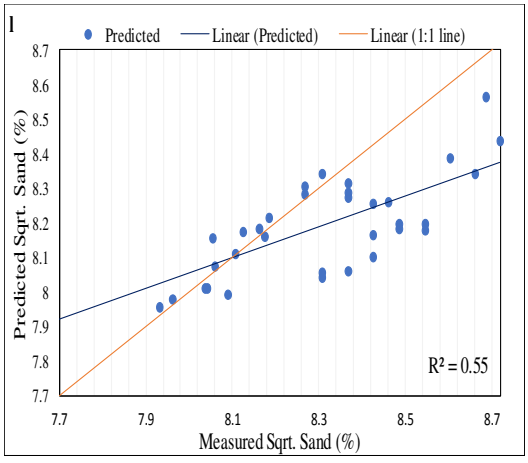

g
Pred
1:1 line
Linear (Pred)
Linear (1:1 line)
Predicted Sqrt. $CaCO_3$
Measured Sqrt. $CaCO_3$
$R^2 = 0.82$


h
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Fe (ppm)
Measured Sqrt. Fe (ppm)
$R^2 = 0.52$


i
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Mn (ppm)
Measured Sqrt. Mn (ppm)
$R^2 = 0.30$


j
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Cu (ppm)
Measured Sqrt. Cu (ppm)
$R^2 = 0.40$


k
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Zn (ppm)
Measured Sqrt. Zn (ppm)
R^2=0.12


l
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Sand (%)
Measured Sqrt. Sand (%)
$R^2 = 0.55$

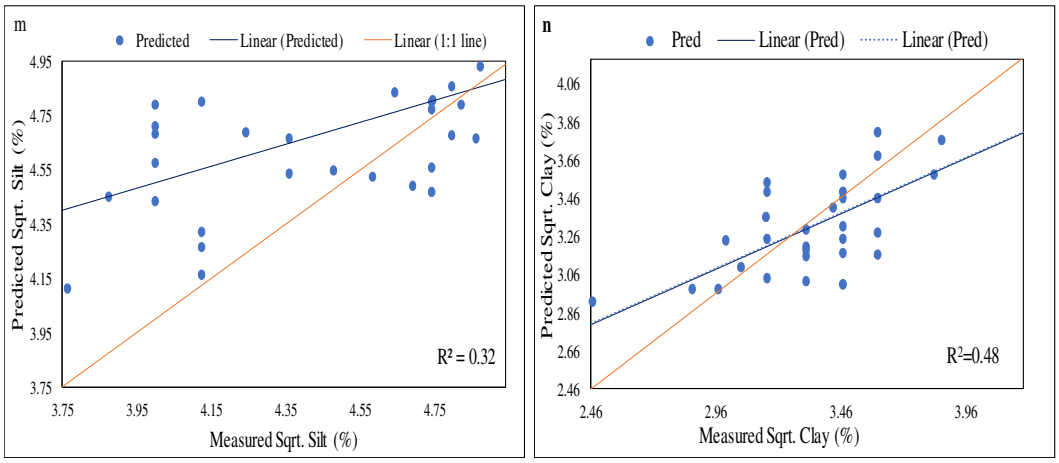


Fig. 4.19 Measured and predicted values, fitted linear line and 1:1 line of (a) Square root pH (b) Square root EC (c) Square root SOC (d) log N (e) Square root P (f) log K (g) Square root $CaCO_3$ (h) Square root Fe (i) Square root Mn (j) Square root Cu (k) Square root Zn (l) Square root Sand (m) Square root Silt (n) Square root Clay for Northeastern undulating.

The standard deviation (SD), root mean square error (RMSE), ratio of performance deviation (RPD) and R^2 value ranged between 0.07-0.52, 0.05-0.39, 0.93-2.15 and 0.12-0.82, respectively among all the soil physical and chemical properties for the soil samples collected from North eastern undulating plain (Fig 4.2a -4.2n). The R^2 value for the $CaCO_3$ (0.82), sand (0.55) and DTPA-Fe (0.82) were observed higher *i.e.,* more than 0.5, whereas the R^2 value for DTPA extractable Zn (0.12) and EC (0.12) was very low. Alike the R^2 value, the RPD value was also observed higher for $CaCO_3$ (2.15) and DTPA-Fe (1.45). Contrary to the poor prediction of pH in sub-mountain hill, a better RPD, R^2 and lower RMSE was observed for pH prediction in North eastern undulating plain. The lower RPD value for soil textural prediction (Table 4.2) was mainly because of larger variation in the soil texture in this region. The statistical parameters of validations are presented in Table 4.21. So, considering all the statistical parameters, developed spectral model can be used for prediction of sand, $CaCO_3$ and DTPA-extractable Fe content for North Eastern Undulating.

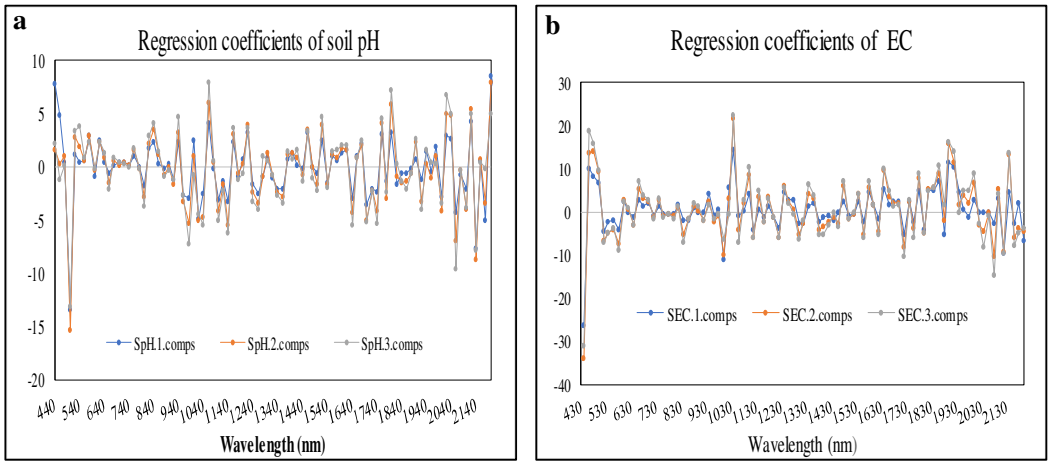
Table 4.21 Statistical parameters for validation of spectral models for predicting various soil properties of Northeastern Undulating

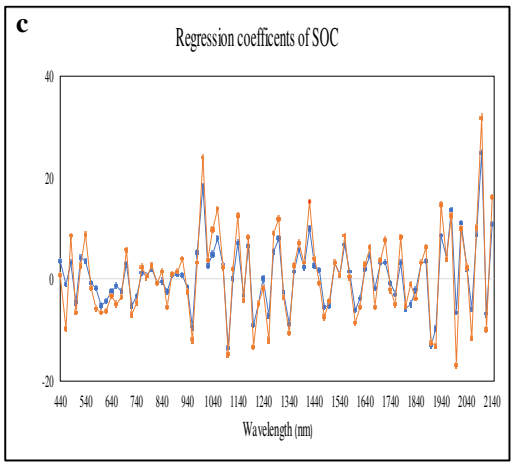
Soil property	Statistical parameter			
	SD	RMSE	RPD	R^2
pH*	0.07	0.05	1.3	0.45
EC* (dS/m)	0.11	0.12	0.97	0.12
SOC (%)	0.08	0.06	1.35	0.47
N# (kg ha^{-1})	0.19	0.15	1.26	0.37
P* (kg ha^{-1})	0.49	0.36	1.35	0.49
K# (kg ha^{-1})	0.08	0.06	1.32	0.43
$CaCO_3$ *(%)	0.30	0.14	2.15	0.82
DTPA-Fe* (ppm)	0.52	0.36	1.45	0.52
DTPA-Mn* (ppm)	0.34	0.28	1.21	0.3
DTPA-Cu* (ppm)	0.25	0.2	1.28	0.4
DTPA-Zn* (ppm)	0.26	0.28	0.94	0.12
Sand* (%)	0.24	0.19	1.23	0.55
Silt* (%)	0.37	0.39	0.93	0.32
Clay* (%)	0.32	0.24	1.36	0.48

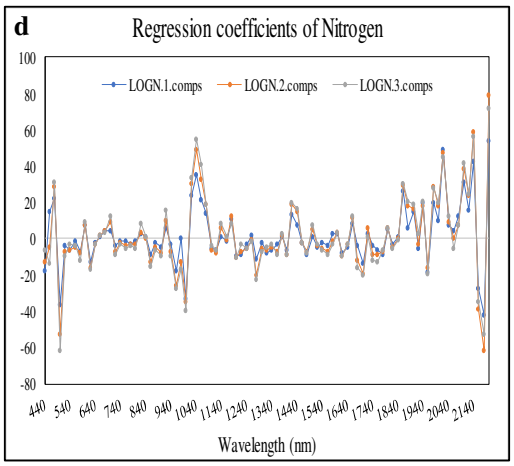
(*) Square root transformation

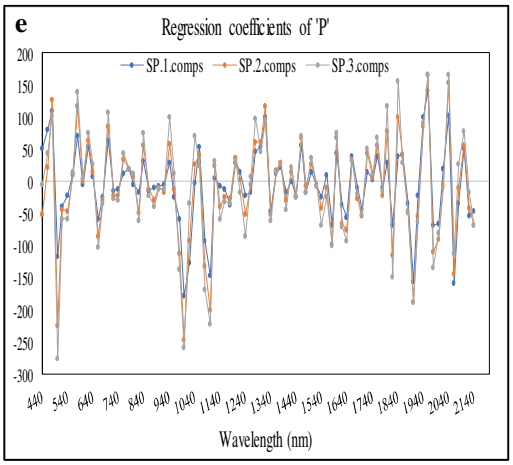
(#) Log transformation

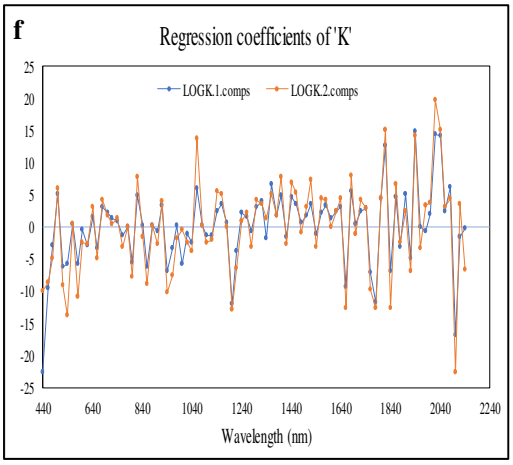
The regression coefficients of the fitted models for the Northeastern Undulating are presented below

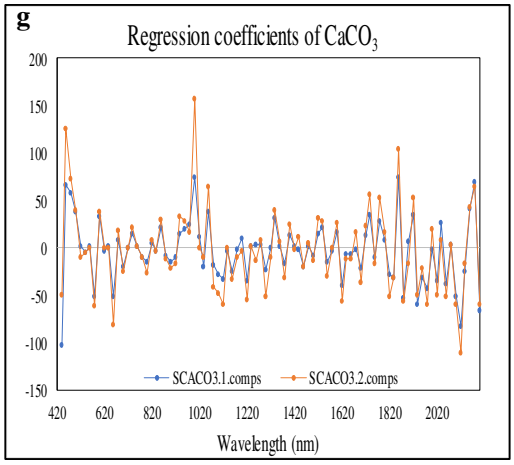


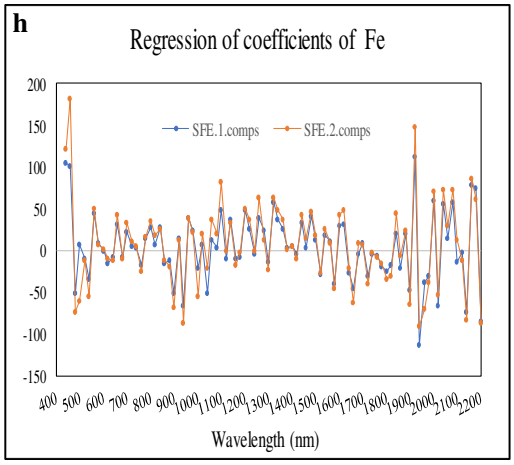

c
Regression coefficents of SOC
Wavelength (nm)


d
Regression coefficients of Nitrogen
LOGN.1.comps
LOGN.2.comps
LOGN.3.comps
Wavelength (nm)


e
Regression coefficients of 'P'
SP.1.comps
SP.2.comps
SP.3.comps
Wavelength (nm)


f
Regression coefficients of 'K'
LOGK.1.comps
LOGK.2.comps
Wavelength (nm)


g
Regression coefficients of CaCO3
SCACO3.1.comps
SCACO3.2.comps
Wavelength (nm)


h
Regression of coefficients of Fe
SFE.1.comps
SFE.2.comps
Wavelength (nm)

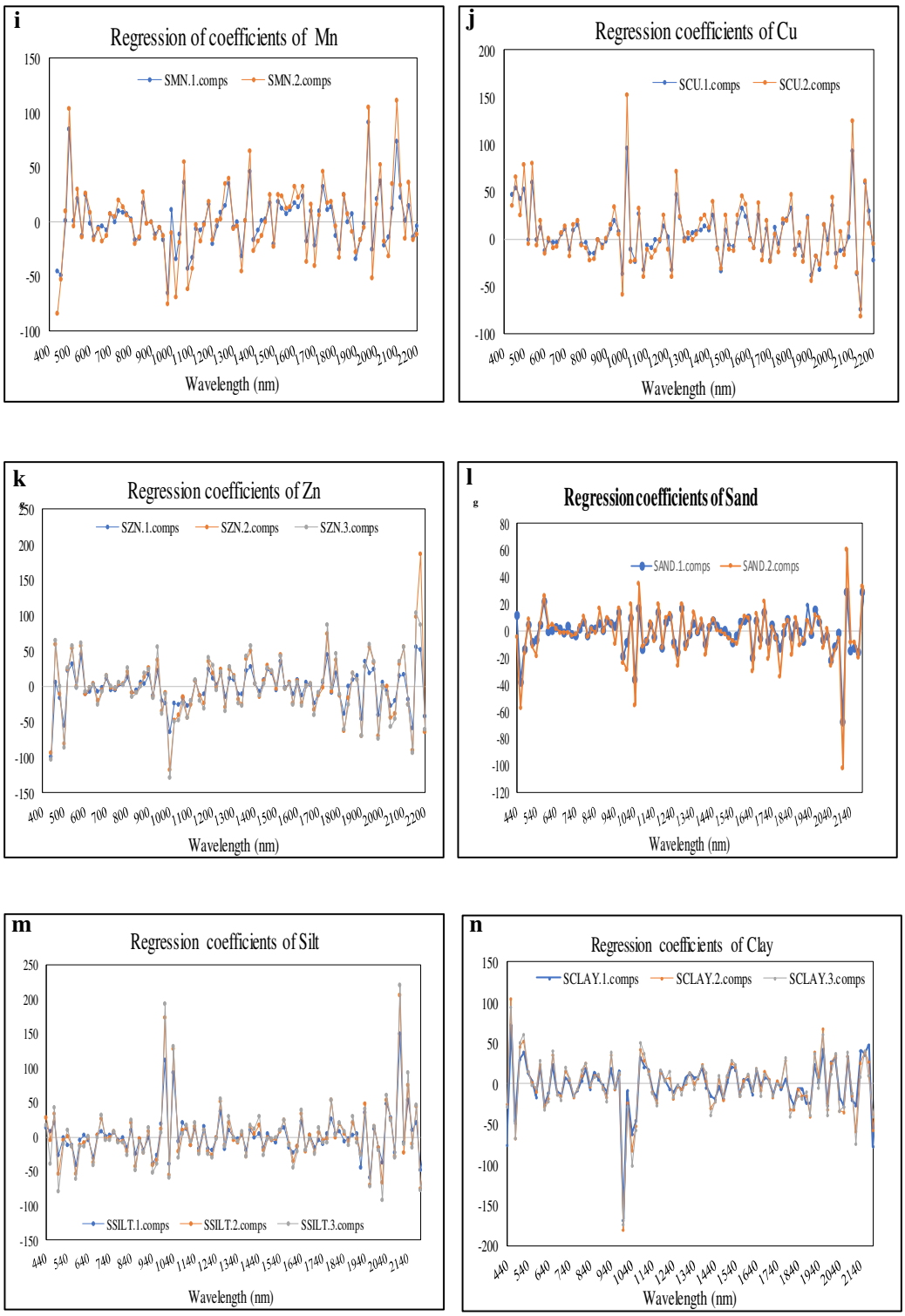
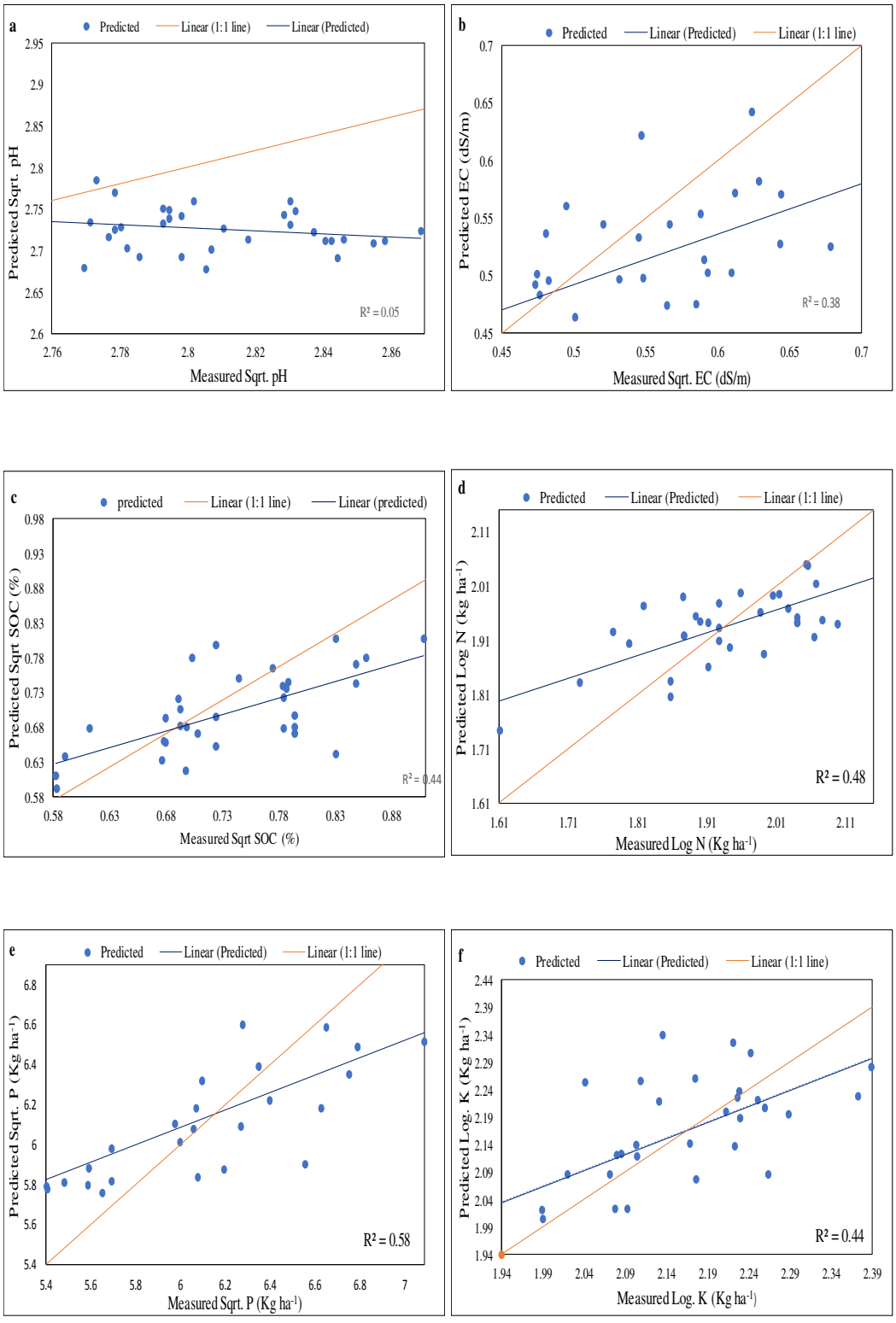
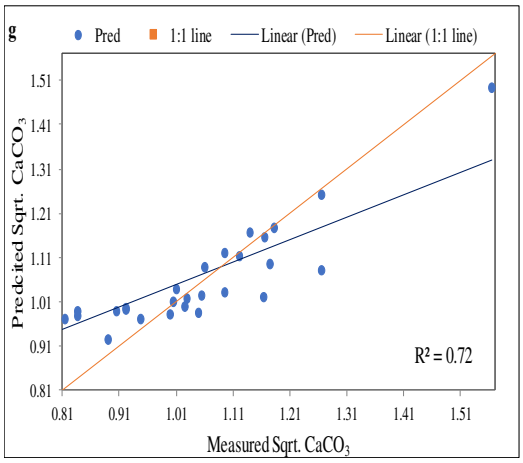
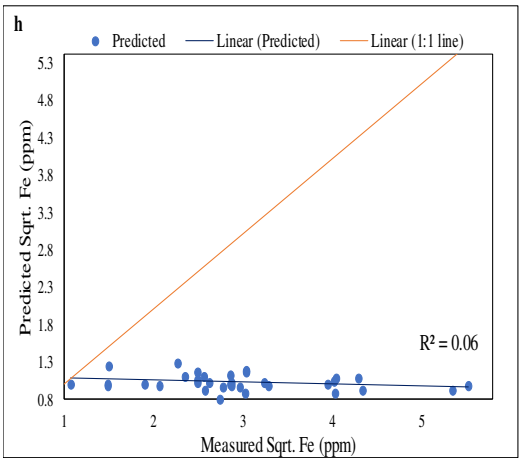


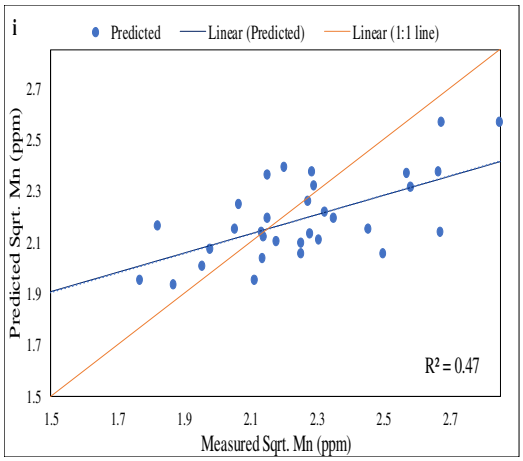
Fig. 4.20 Regression coefficients of (a) Soil pH (b) EC (c) SOC (d) Available N (e) Olsen-P (f) Available K (g) Soil calcium carbonate (h) DTPA-Fe (i) DTPA-Mn (j) DTPA-Cu (k) DTPA-Zn (l) Sand (m) Silt (n) Clay for Northeastern Undulating

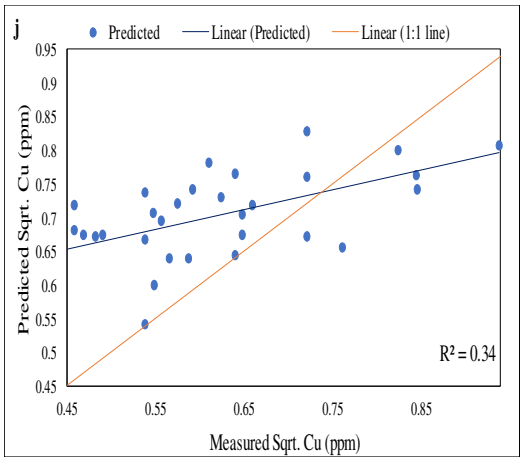
4.2.3 Piedmont and Alluvial Plain

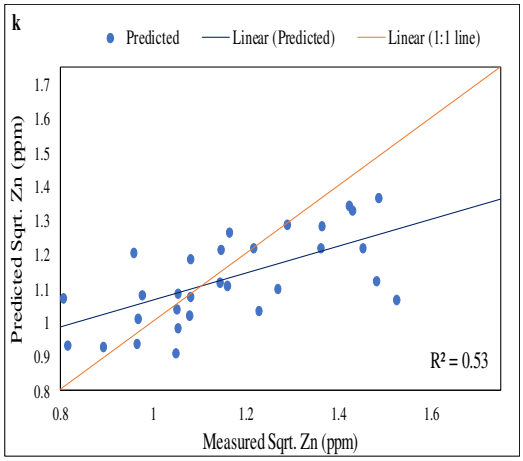


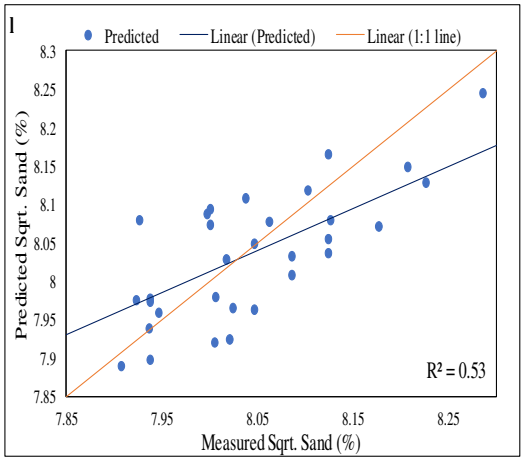

g
Pred
1:1 line
Linear (Pred)
Linear (1:1 line)
Predicted Sqrt. CaCO3
Measured Sqrt. CaCO3
R² = 0.72


h
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Fe (ppm)
Measured Sqrt. Fe (ppm)
R² = 0.06


i
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Mn (ppm)
Measured Sqrt. Mn (ppm)
R² = 0.47


j
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Cu (ppm)
Measured Sqrt. Cu (ppm)
R² = 0.34


k
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Zn (ppm)
Measured Sqrt. Zn (ppm)
R² = 0.53


l
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Sand (%)
Measured Sqrt. Sand (%)
R² = 0.53

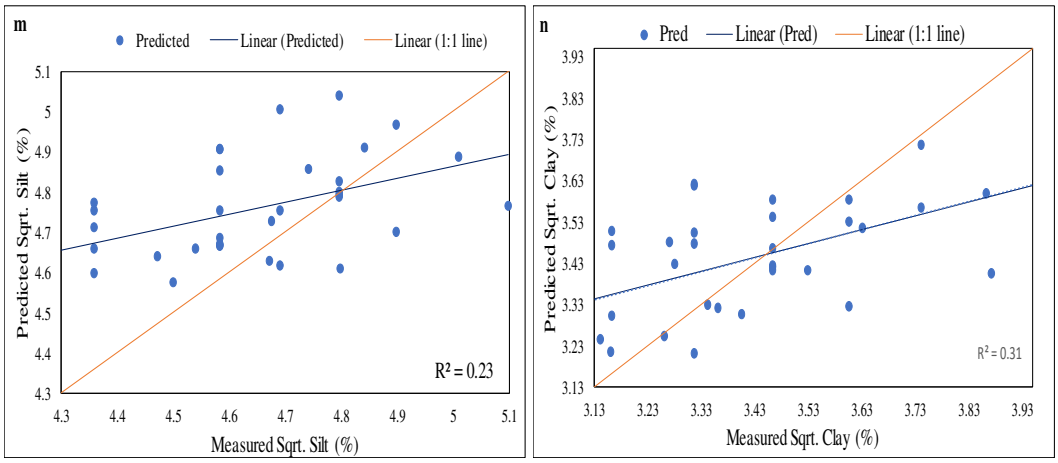


Fig. 4.21 Measured and predicted values, fitted linear line and 1:1 line of (a) Square root pH (b) Square root EC (c) Square root SOC (d) log N (e) Square root P (f) log K (g) Square root $CaCO_3$ (h) Square root Fe (i) Square root Mn (j) Square root Cu (k) Square root Zn (l) Square root Sand (m) Square root Silt (n) Square root Clay for Piedmont and Alluvial Plain.

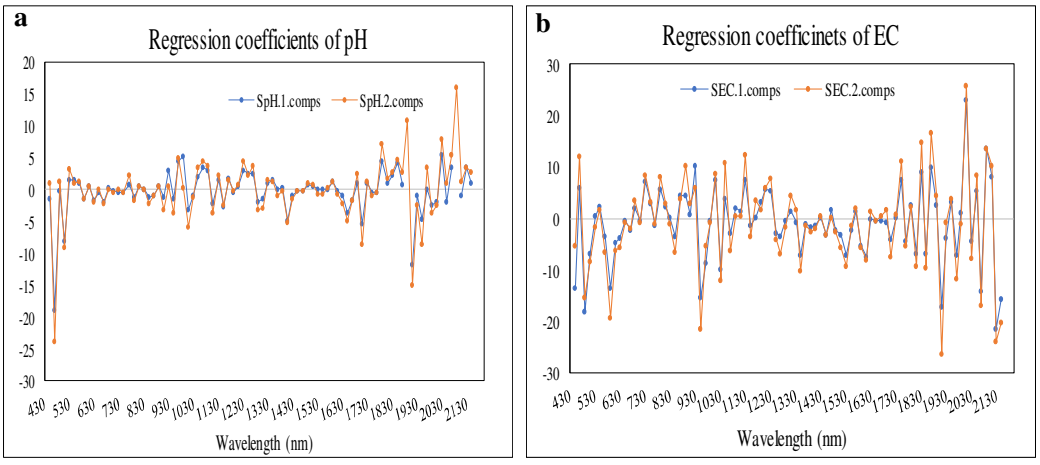
The standard deviation (SD), root mean square error (RMSE), ratio of performance deviation (RPD) and R^2 value ranged between 0.03-2.87, 0.07-2.90, 0.32-1.60 and 0.05-0.72, respectively, among all the soil physical and chemical properties for the soil samples collected from Piedmont and Alluvial Plain (Fig 4.3a-4.3n and Table 4.21). The R^2 value for the $CaCO_3$ (0.72), sand (0.53), phosphorus (0.58) and DTPA-Zn (0.52) was observed, more than 0.5, whereas the R^2 value for DTPA extractable Fe (0.06) and pH (0.05) was very low. Similar to the R^2 value, the RPD value was also observed higher for $CaCO_3$ (1.60), sand (1.46), available N (1.40), P (1.41) and DTPA-Zn (1.41). Contrary to the poor prediction of DTPA-extractable Zn in the NEU, a very good prediction of DTPA-extractable Zn was observed in the Piedmont and Alluvial plain. The DTPA-extractable Fe was not adequately predicted under NEU, which is very poorly predicted in Piedmont and Alluvial Plain. The pH in the Piedmont and alluvial plain was poorly predicted alike to the sub-mountain hill. So, considering all the statistical parameters, developed spectral model can be considered for prediction of sand, $CaCO_3$, DTPA-extractable Zn and available P content for soils of Piedmont and alluvial plain. The R^2 value, RMSE, SD and RPD for all parameter are presented in Table 4.22.

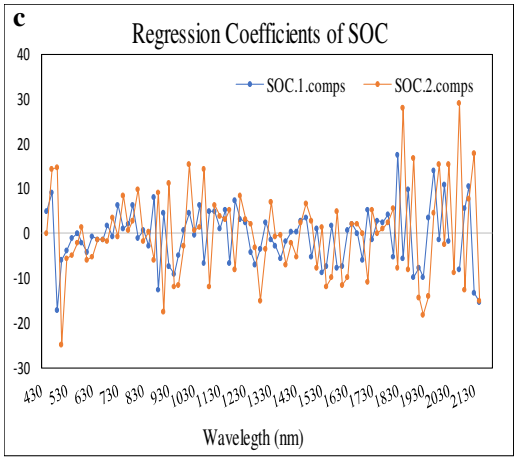
Table 4.22 Statistical parameters for validation of observed soil properties in Piedmont and Alluvial plain

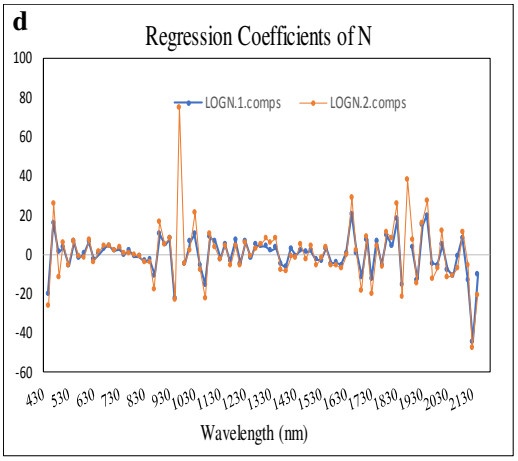
Soil property	Statistical parameter			
	SD	RMSE	RPD	R^2
pH*	0.03	0.09	0.32	0.05
EC* (dS/m)	0.08	0.07	1.12	0.38
SOC (%)	0.08	0.07	1.16	0.44
N# (kg ha^{-1})	0.12	0.08	1.40	0.48
P* (kg ha^{-1})	0.63	0.45	1.41	0.58
K# (kg ha^{-1})	0.13	0.10	1.30	0.44
$CaCO_3$ *(%)	0.20	0.13	1.60	0.72
DTPA-Fe* (ppm)	2.87	2.90	0.98	0.06
DTPA-Mn* (ppm)	0.30	0.23	1.31	0.47
DTPA-Cu* (ppm)	0.13	0.14	0.96	0.34
DTPA-Zn* (ppm)	0.25	0.18	1.41	0.53
Sand* (%)	0.11	0.076	1.46	0.53
Silt* (%)	0.20	0.21	0.95	0.23
Clay* (%)	0.23	0.19	1.21	0.31

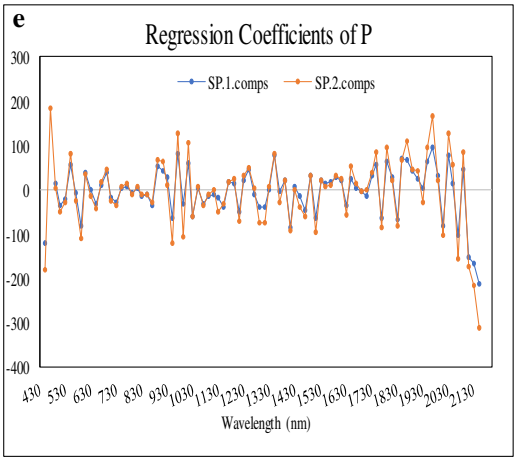
(*) Square root transformation
(#) Log transformation

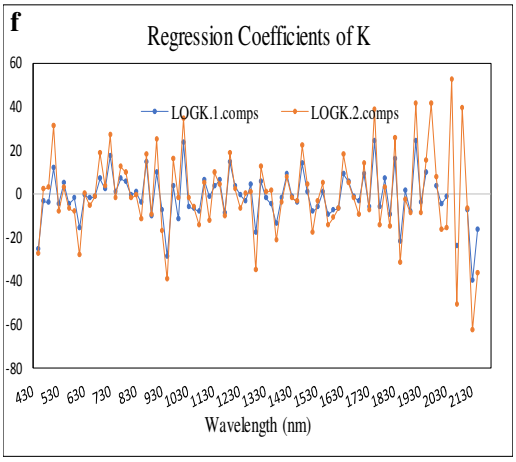
The regression coefficients of the fitted models for the Piedmont and Alluvial Plain are presented below

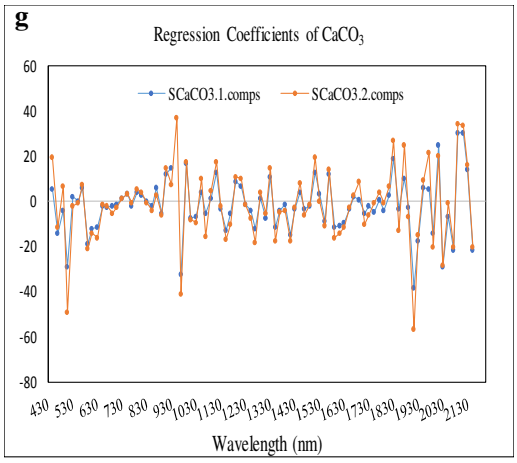


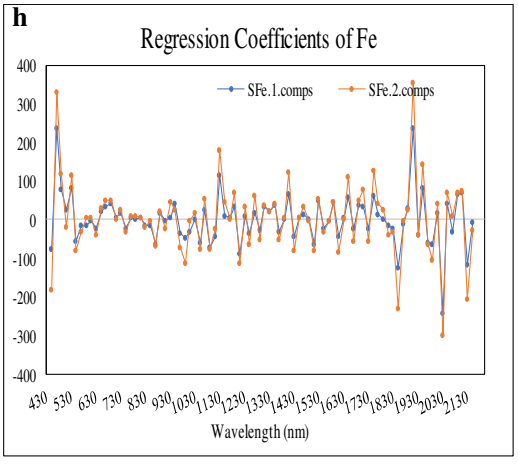

c
Regression Coefficients of SOC
SOC.1.comps
SOC.2.comps
Wavelength (nm)


d
Regression Coefficients of N
LOGN.1.comps
LOGN.2.comps
Wavelength (nm)


e
Regression Coefficients of P
SP.1.comps
SP.2.comps
Wavelength (nm)


f
Regression Coefficients of K
LOGK.1.comps
LOGK.2.comps
Wavelength (nm)


g
Regression Coefficients of CaCO3
SCaCO3.1.comps
SCaCO3.2.comps
Wavelength (nm)


h
Regression Coefficients of Fe
SFe.1.comps
SFe.2.comps
Wavelength (nm)

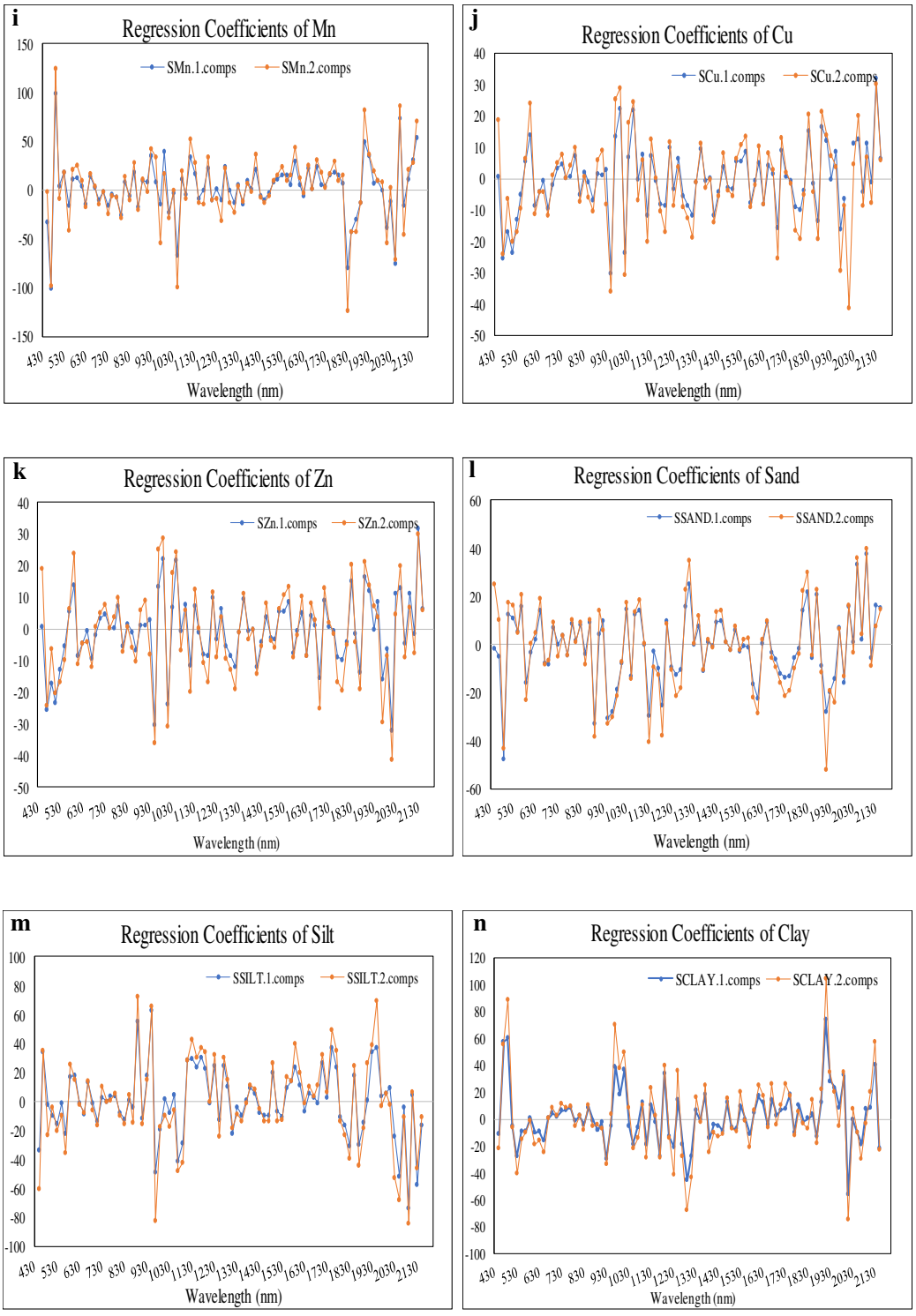
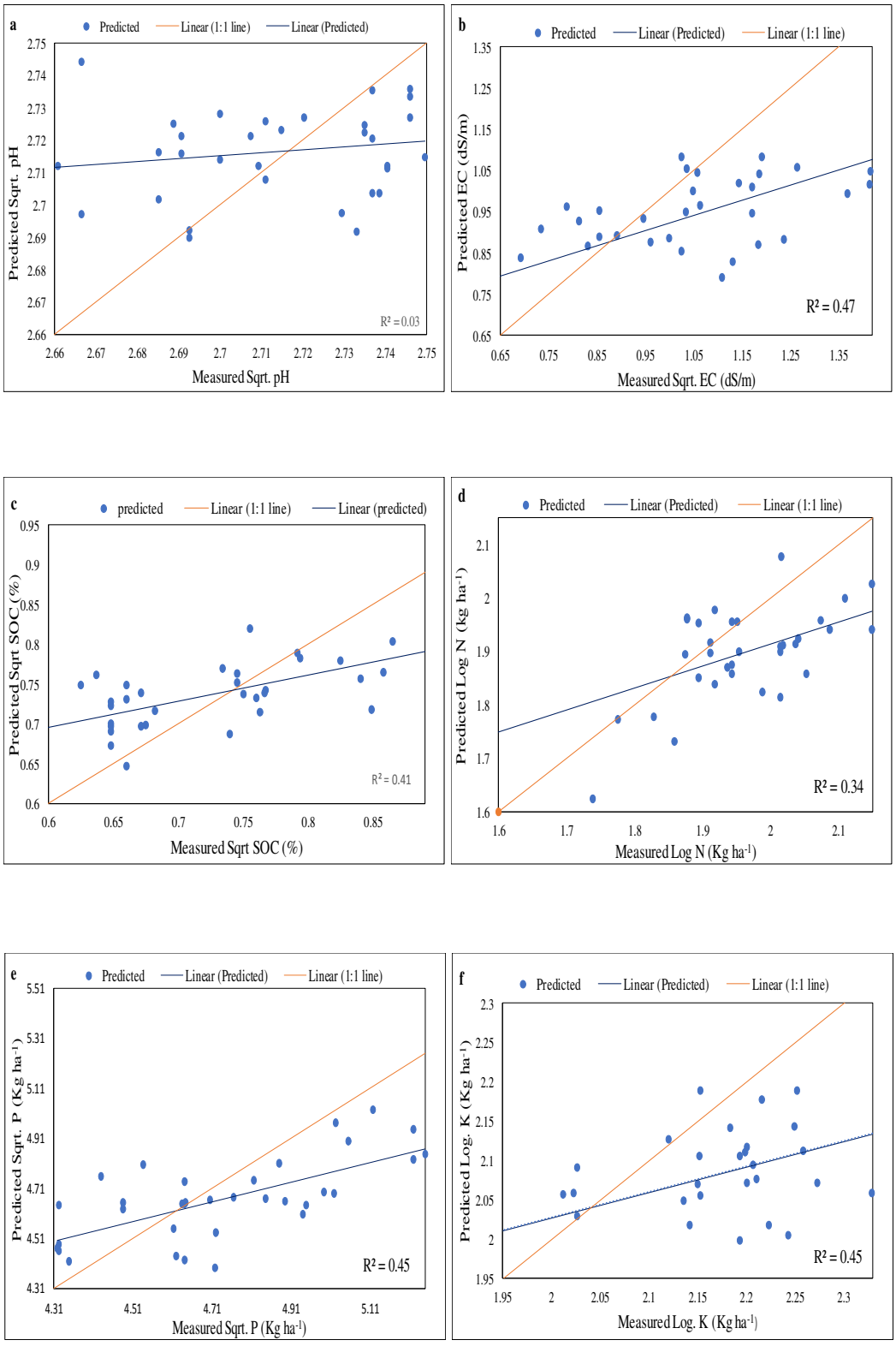
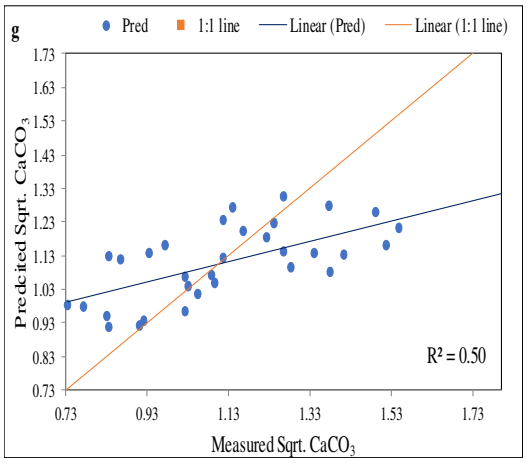
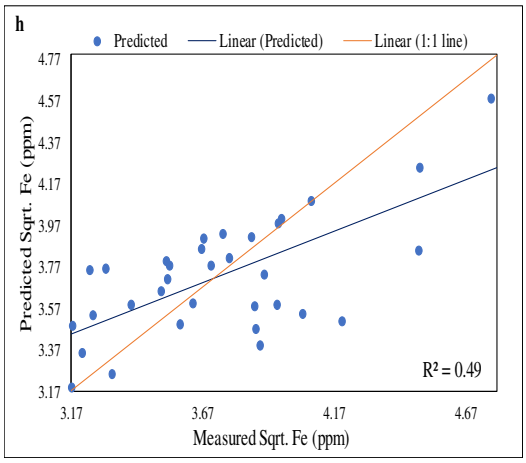


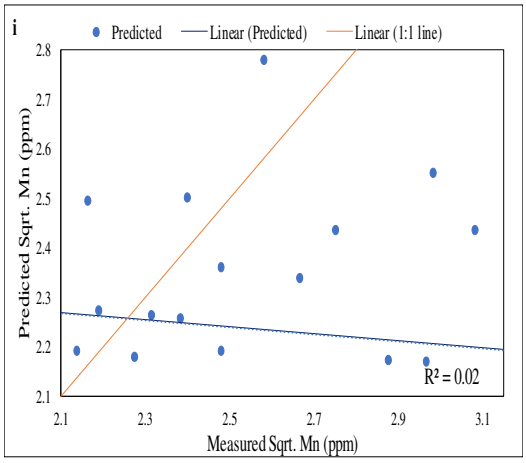
Fig. 4.22 Regression coefficients of (a) Soil pH (b) EC (c) SOC (d) Available N (e) Olsen-P (f) Available K (g) Soil calcium carbonate (h) DTPA-Fe (i) DTPA-Mn (j) DTPA-Cu (k) DTPA-Zn (l) Sand (m) Silt (n) Clay for Piedmont and Alluvial Plain

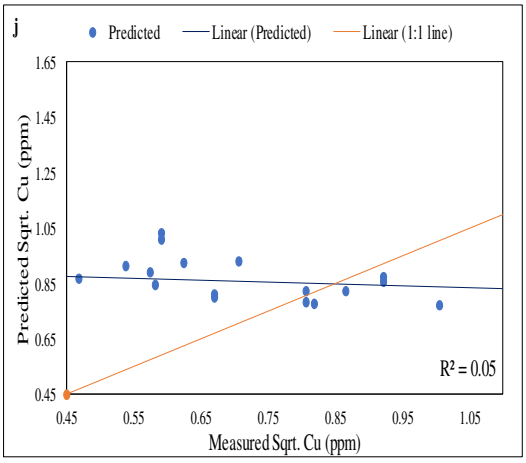
4.2.4 Central Alluvial Plain

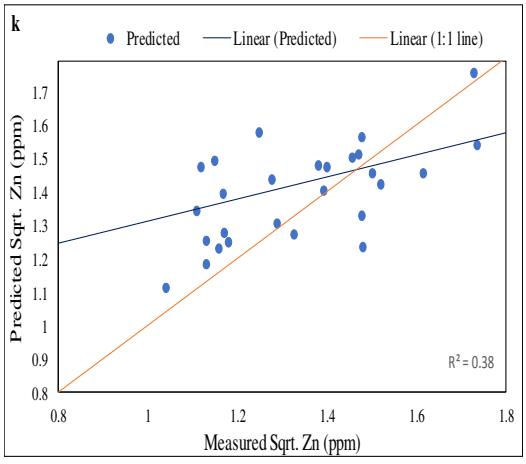


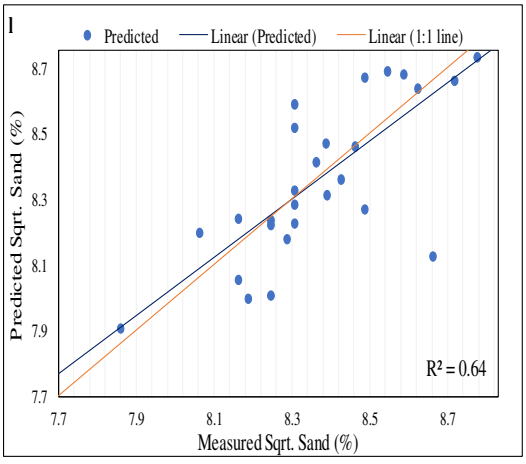

g
Pred
1:1 line
Linear (Pred)
Linear (1:1 line)
Predicted Sqrt. CaCO3
Measured Sqrt. CaCO3
R² = 0.50


h
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Fe (ppm)
Measured Sqrt. Fe (ppm)
R² = 0.49


i
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Mn (ppm)
Measured Sqrt. Mn (ppm)
R² = 0.02


j
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Cu (ppm)
Measured Sqrt. Cu (ppm)
R² = 0.05


k
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Zn (ppm)
Measured Sqrt. Zn (ppm)
R² = 0.38


l
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Sand (%)
Measured Sqrt. Sand (%)
R² = 0.64

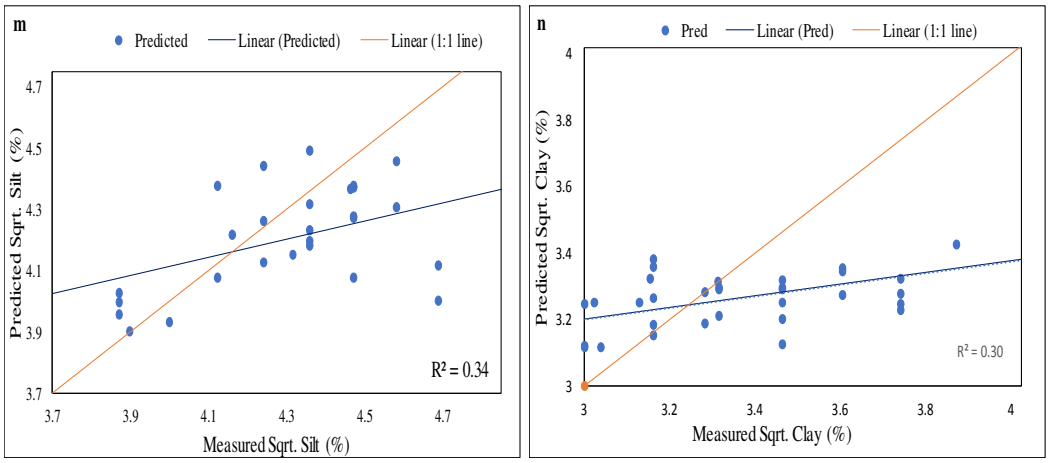


Fig. 4.23 Measured and predicted values, fitted linear line and 1:1 line of (a) Square root pH (b) Square root EC (c) Square root SOC (d) log N (e) Square root P (f) log K (g) Square root $CaCO_3$ (h) Square root Fe (i) Square root Mn (j) Square root Cu (k) Square root Zn (l) Square root Sand (m) Square root Silt (n) Square root Clay for Central Alluvial Plain.

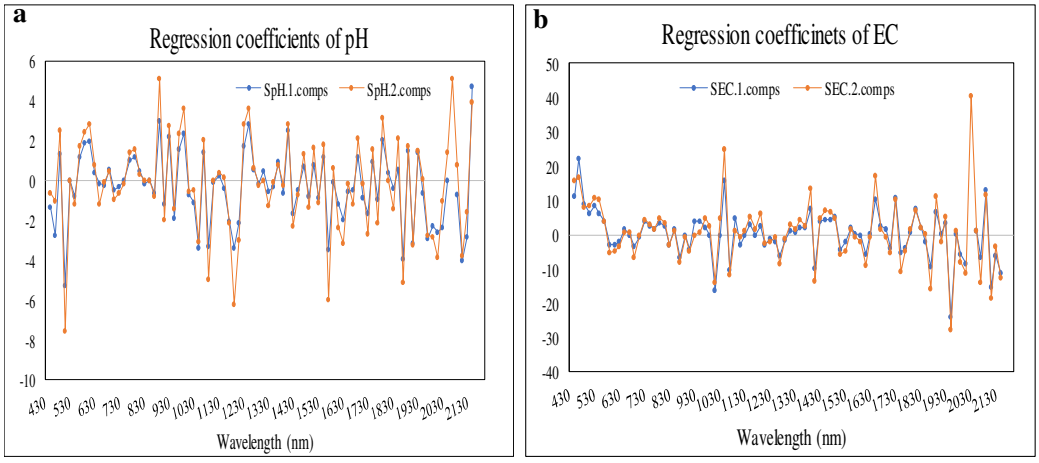
The standard deviation (SD), root mean square error (RMSE), ratio of performance deviation (RPD) and R^2 values ranged between 0.03-0.44, 0.03-0.43, 0.62-1.50 and 0.02-0.64, respectively, among all the soil physical and chemical properties for the soil samples collected from Central Alluvial plain (Fig 4.5a-4.5n and Table 4.5). The R^2 value for the $CaCO_3$ (0.5), sand (0.64), phosphorus (0.45) and DTPA-Fe (0.5) was observed to be more than 0.5, whereas the R^2 value for DTPA extractable Cu (0.06) and pH (0.05) was very low. The RPD in case of sand only was observed to be higher. Although a high R^2 was observed for DTPA-extractable Fe, the RMSE was high and equal to SD, resulting in a poor model for predicting Fe content. (Table 4.23)

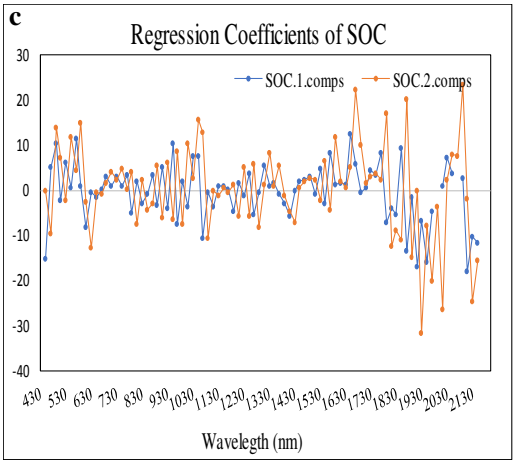
Table 4.23 Statistical parameter for validation of spectral models developed for predicting soil properties in Central Alluvial Plain

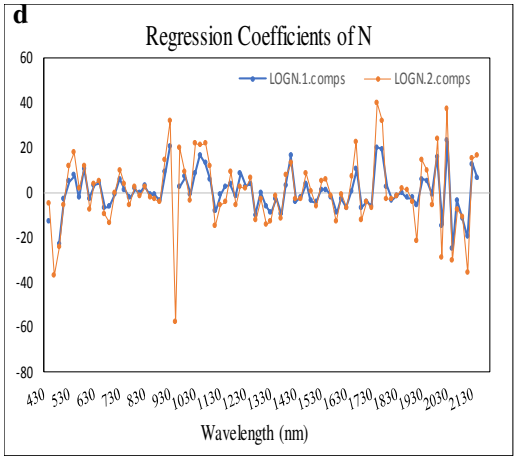
Soil property	Statistical parameter			
	SD	RMSE	RPD	R^2
pH*	0.027	0.028	0.94	0.03
EC* (dS/m)	0.24	0.20	1.20	0.47
SOC (%)	0.08	0.06	1.28	0.41
N# (kg ha^{-1})	0.13	0.12	1.10	0.34
P* (kg ha^{-1})	0.14	0.12	1.16	0.45
K# (kg ha^{-1})	0.14	0.12	1.16	0.45
$CaCO_3$ *(%)	0.27	0.21	1.31	0.5
DTPA-Fe* (ppm)	0.30	0.30	1.00	0.49
DTPA-Mn* (ppm)	0.44	0.34	1.3	0.46
DTPA-Cu* (ppm)	0.26	0.43	0.62	0.02
DTPA-Zn* (ppm)	0.36	0.35	1.02	0.26
Sand* (%)	0.29	0.19	1.50	0.64
Silt* (%)	0.36	0.29	1.20	0.34
Clay* (%)	0.27	0.26	1.07	0.30

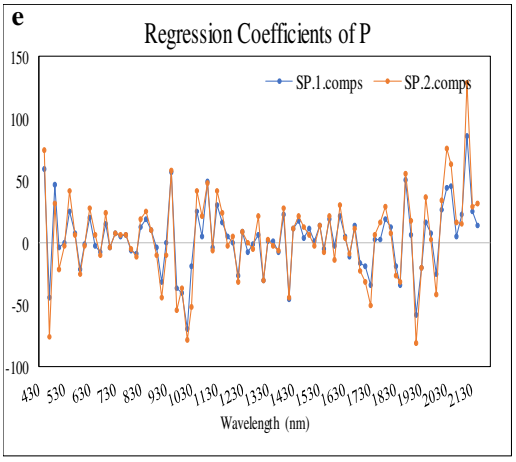
(*) Square root transformation
(#) Log transformation

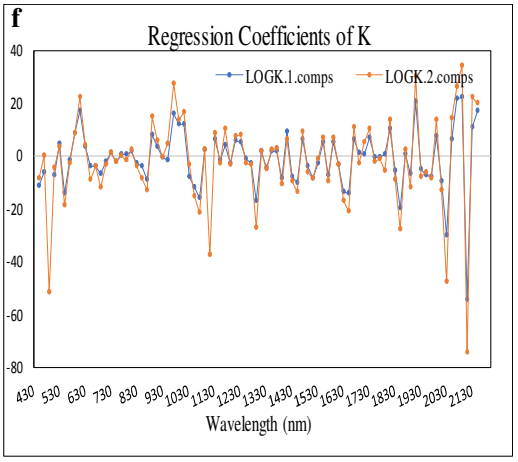
The regression coefficients of the fitted models for the Central Alluvial Plain are presented below

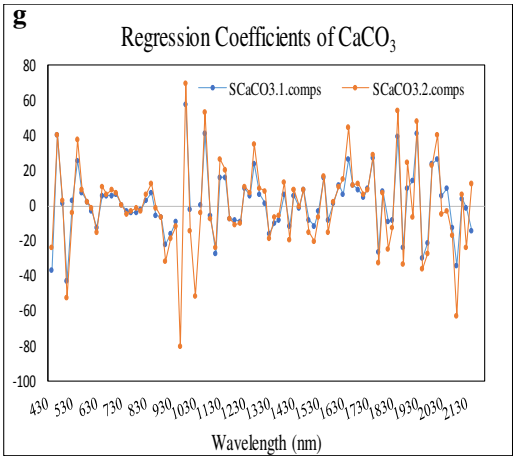


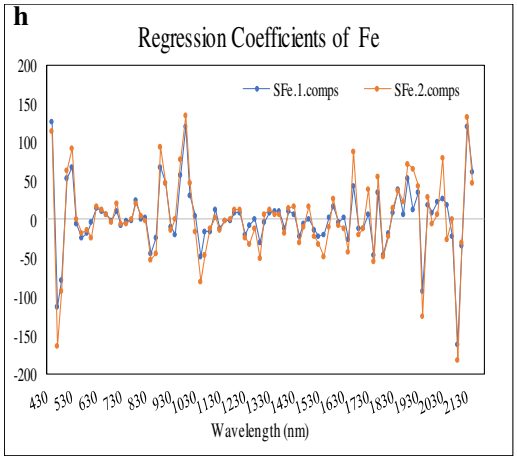

c
Regression Coefficients of SOC
SOC.1.comps
SOC.2.comps
Wavelength (nm)


d
Regression Coefficients of N
LOGN.1.comps
LOGN.2.comps
Wavelength (nm)


e
Regression Coefficients of P
SP.1.comps
SP.2.comps
Wavelength (nm)


f
Regression Coefficients of K
LOGK.1.comps
LOGK.2.comps
Wavelength (nm)


g
Regression Coefficients of CaCO3
SCaCO3.1.comps
SCaCO3.2.comps
Wavelength (nm)


h
Regression Coefficients of Fe
SFe.1.comps
SFe.2.comps
Wavelength (nm)

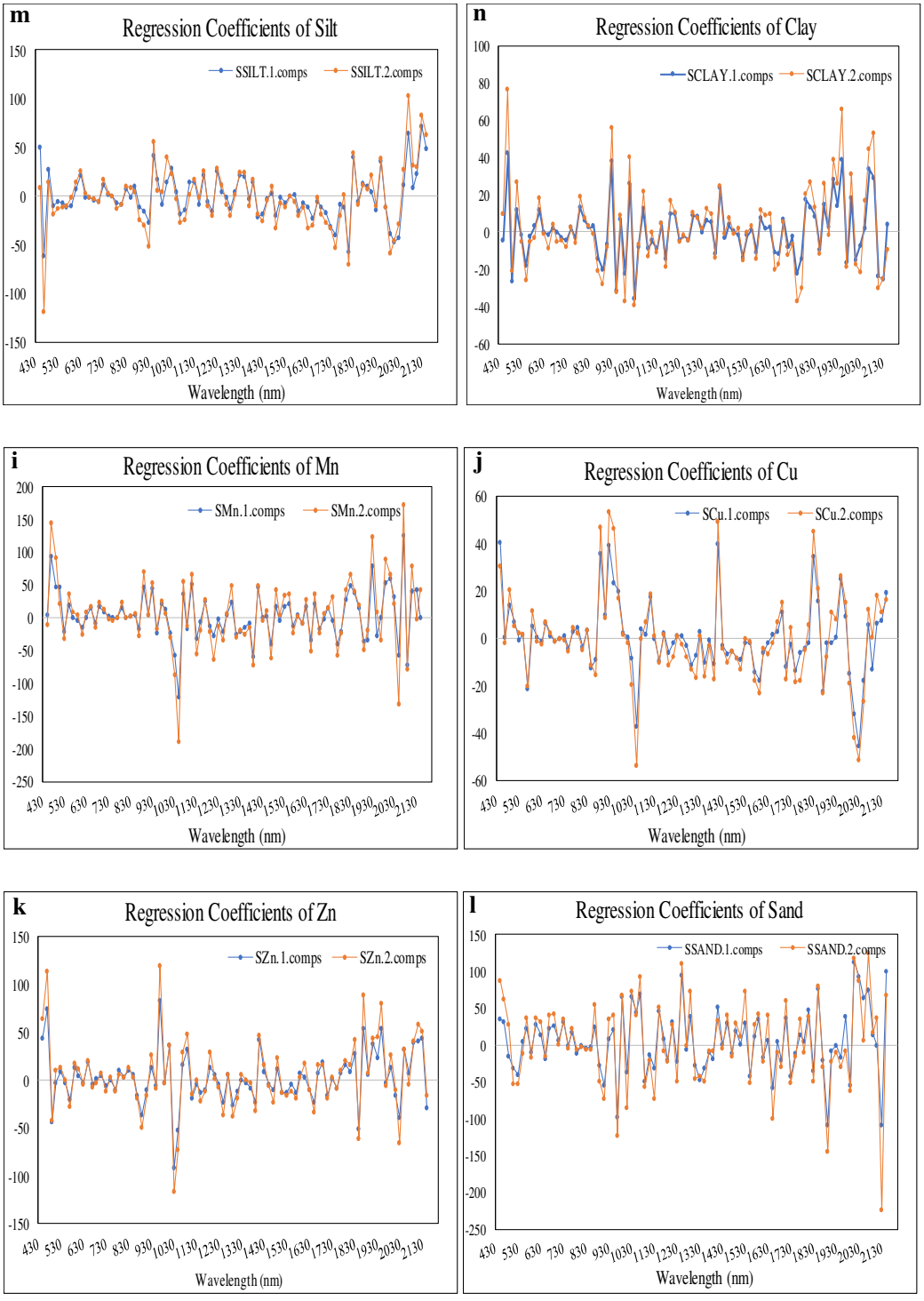
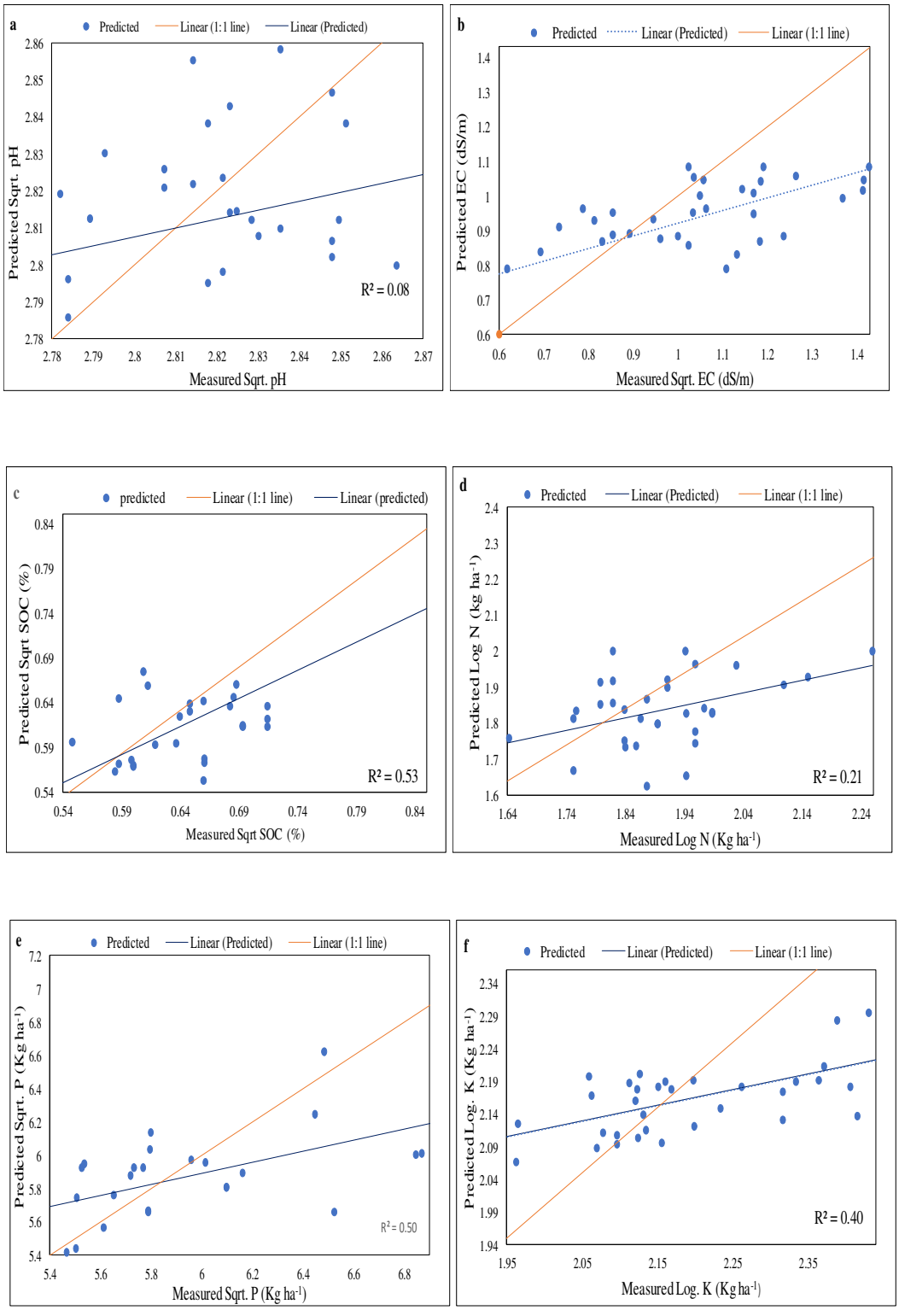
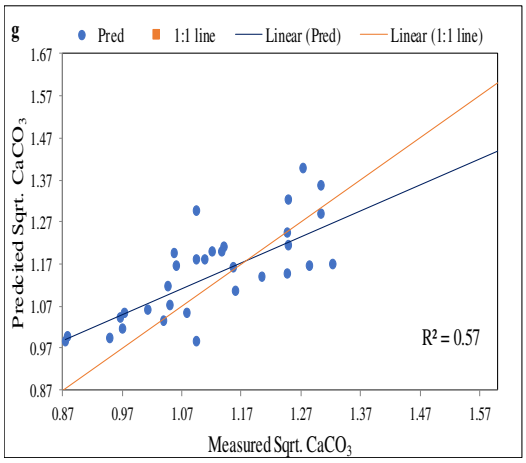
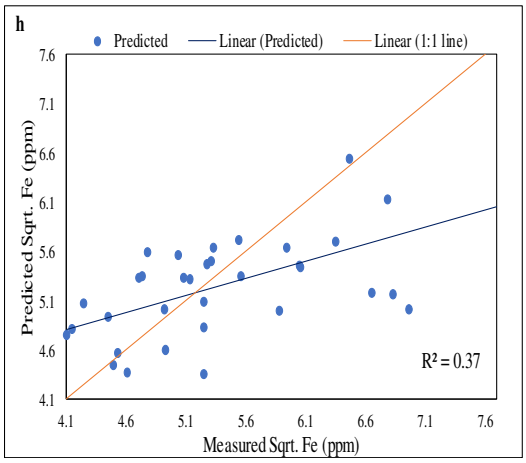


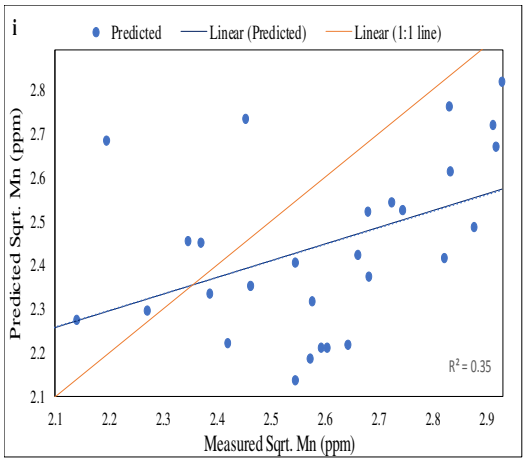
Fig. 4.24 Regression coefficients of (a) Soil pH (b) EC (c) SOC (d) Available N (e) Olsen-P (f) Available K (g) Soil calcium carbonate (h) DTPA-Fe (i) DTPA-Mn (j) DTPA-Cu (k) DTPA-Zn (l) Sand (m) Silt (n) Clay for Central Alluvial Plain

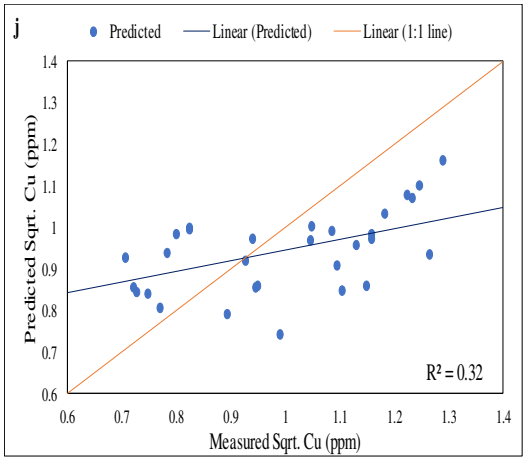
4.2.5 South Western Alluvial Plain

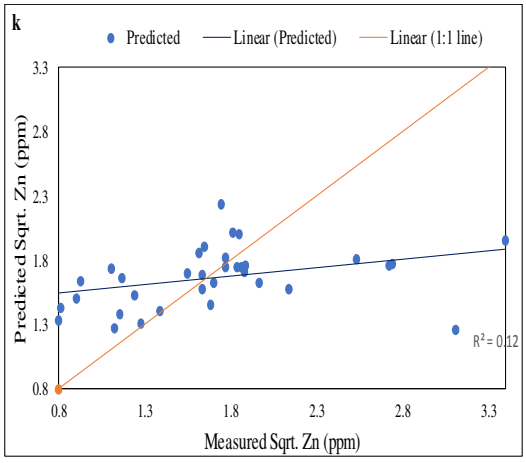


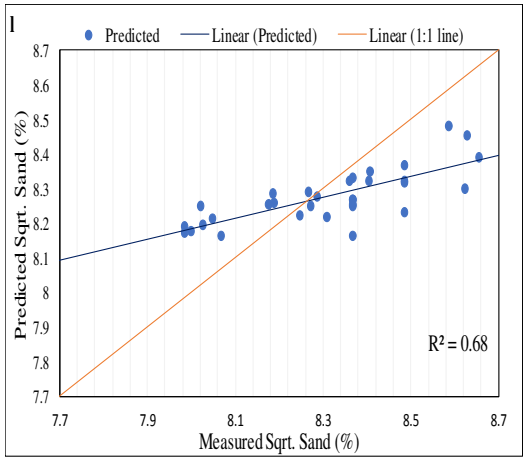

g
Pred
1:1 line
Linear (Pred)
Linear (1:1 line)
Predicted Sqrt. CaCO3
Measured Sqrt. CaCO3
R² = 0.57


h
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Fe (ppm)
Measured Sqrt. Fe (ppm)
R² = 0.37


i
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Mn (ppm)
Measured Sqrt. Mn (ppm)
R² = 0.35


j
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Cu (ppm)
Measured Sqrt. Cu (ppm)
R² = 0.32


k
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Zn (ppm)
Measured Sqrt. Zn (ppm)
R² = 0.12


l
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Sand (%)
Measured Sqrt. Sand (%)
R² = 0.68

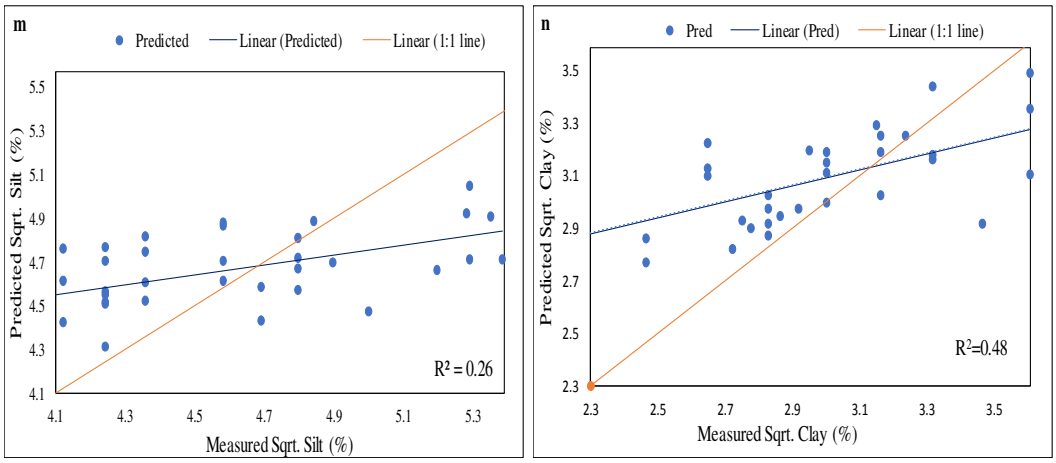


Fig. 4.25 Measured and predicted values, fitted linear line and 1:1 line of (a) Square root pH (b) Square root EC (c) Square root SOC (d) log N (e) Square root P (f) log K (g) Square root $CaCO_3$ (h) Square root Fe (i) Square root Mn (j) Square root Cu (k) Square root Zn (l) Square root Sand (m) Square root Silt (n) Square root Clay for South Western Alluvial Plain.

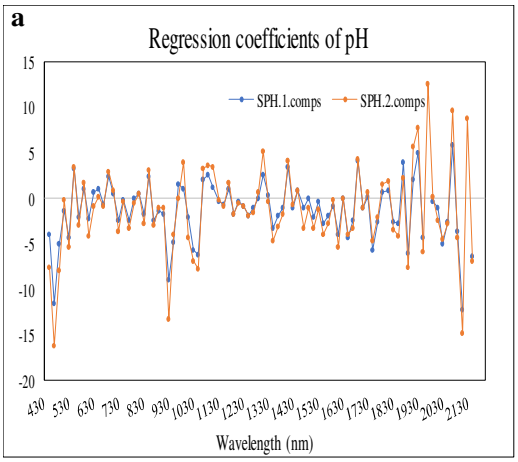
The standard deviation (SD), root mean square error (RMSE), ratio of performance deviation (RPD) and R^2 value ranged between 0.03-0.63, 0.03-0.73, 0.87-1.5 and 0.05-0.68, respectively among all the soil physical and chemical properties for the soil samples collected from South Western Alluvial Plain (Fig 4.4a-n) region. The R^2 value for the $CaCO_3$ (0.57), sand (0.68) and SOC (0.53) was observed to be, more than 0.5, whereas the R^2 value for DTPA extractable Zn (0.12), EC (0.05) and pH (0.08) was very low. Similar to the trend in R^2 value, the RPD value was also observed higher for $CaCO_3$ (1.5), sand (1.4) and SOC (1.31). All the four DTPA-extractable micronutrients along with pH and EC were very poorly predicted in South Western Alluvial Plain. Although the sand content was adequately predicted, the silt and clay contents were very poorly predicted. Keeping above in view, spectral models developed can be used for prediction of sand, $CaCO_3$ and clay content for South western Alluvial Plain. The R^2 value, RMSE, SD and RPD values for various soil properties are presented in Table 4.24.

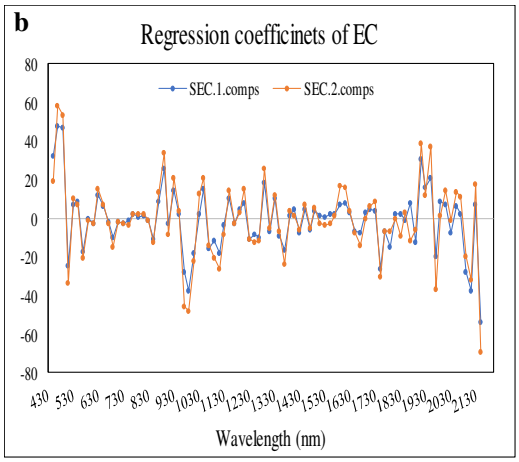
Table 4.24 Statistical parameters for validation of spectral models developed for predicting soil properties in South Western Alluvial Plain

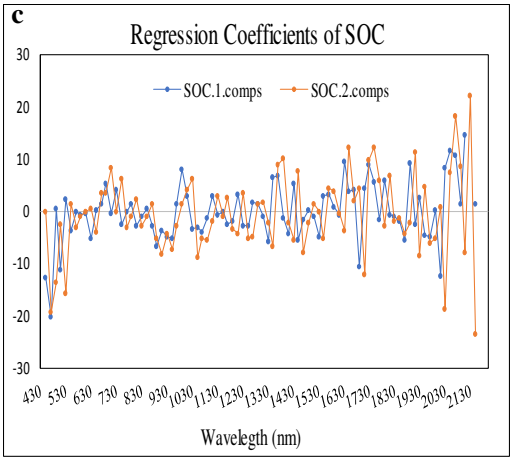
Soil property	Statistical parameter			
	SD	RMSE	RPD	R^2
pH*	0.03	0.033	0.87	0.08
EC* (dS/m)	0.23	0.20	1.20	0.05
SOC (%)	0.073	0.054	1.31	0.53
N# (kg ha^{-1})	0.13	0.14	0.95	0.21
P* (kg ha^{-1})	0.63	0.47	1.33	0.5
K# (kg ha^{-1})	0.14	0.12	1.21	0.4
$CaCO_3$ *(%)	0.15	0.10	1.5	0.57
DTPA-Fe* (ppm)	0.52	0.73	0.72	0.37
DTPA-Mn* (ppm)	0.32	0.27	1.20	0.35
DTPA-Cu* (ppm)	0.21	0.18	1.16	0.32
DTPA-Zn* (ppm)	0.62	0.58	1.10	0.12
Sand* (%)	0.23	0.17	1.40	0.68
Silt* (%)	0.41	0.35	1.16	0.26
Clay* (%)	0.37	0.32	1.17	0.48

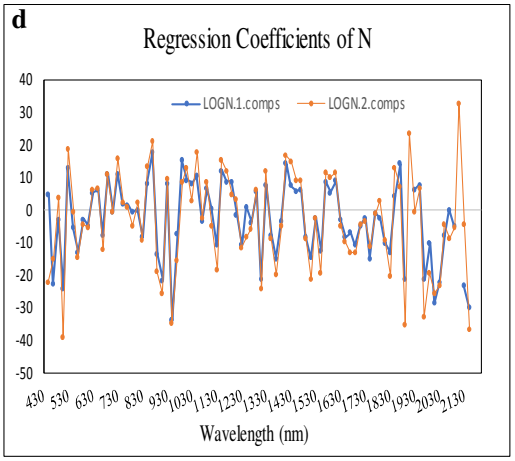
(*) Square root transformation
(#) Log transformation

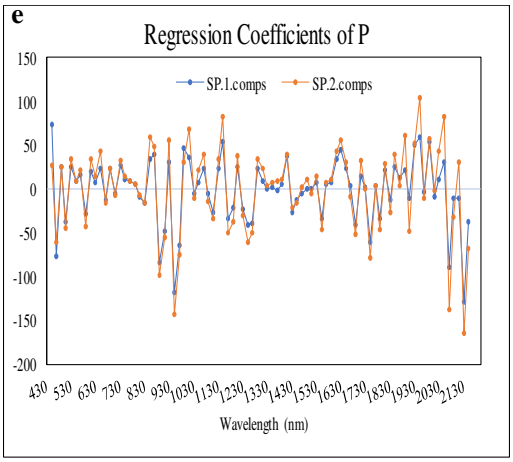
The regression coefficients of the fitted models for the South Western Alluvial Plain are presented below

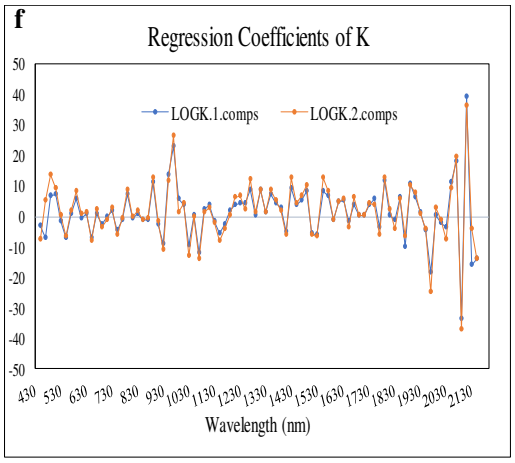


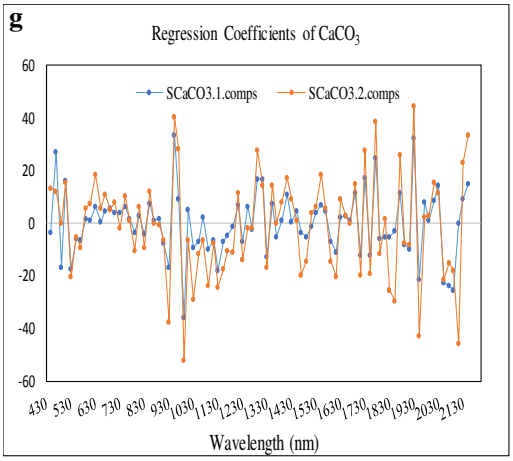


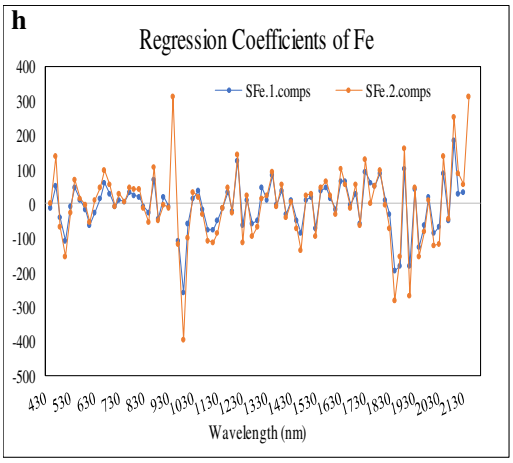

c
Regression Coefficients of SOC
SOC.1.comps
SOC.2.comps
Wavelength (nm)


d
Regression Coefficients of N
LOGN.1.comps
LOGN.2.comps
Wavelength (nm)


e
Regression Coefficients of P
SP.1.comps
SP.2.comps
Wavelength (nm)


f
Regression Coefficients of K
LOGK.1.comps
LOGK.2.comps
Wavelength (nm)


g
Regression Coefficients of CaCO3
SCaCO3.1.comps
SCaCO3.2.comps
Wavelength (nm)


h
Regression Coefficients of Fe
SFe.1.comps
SFe.2.comps
Wavelength (nm)

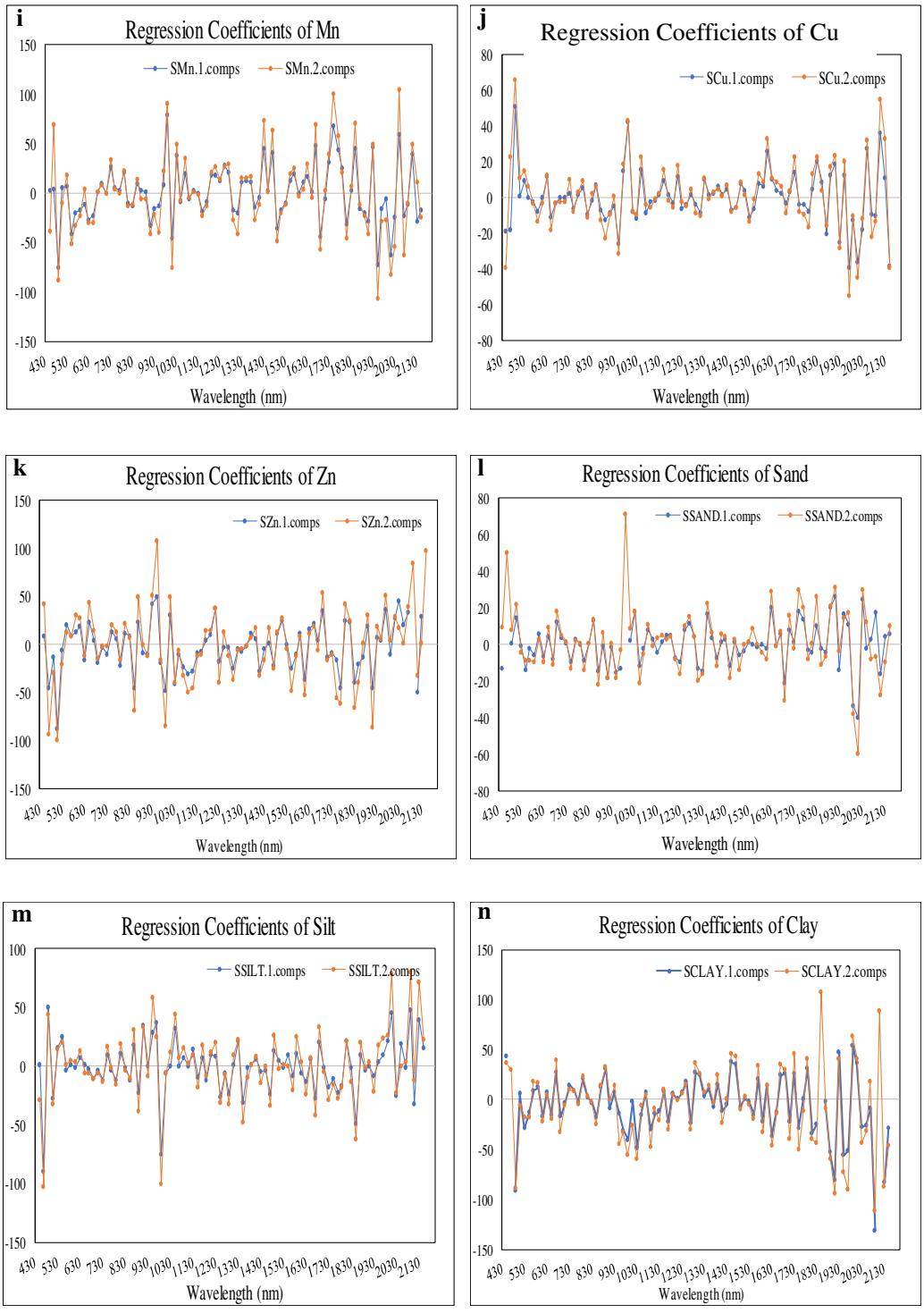
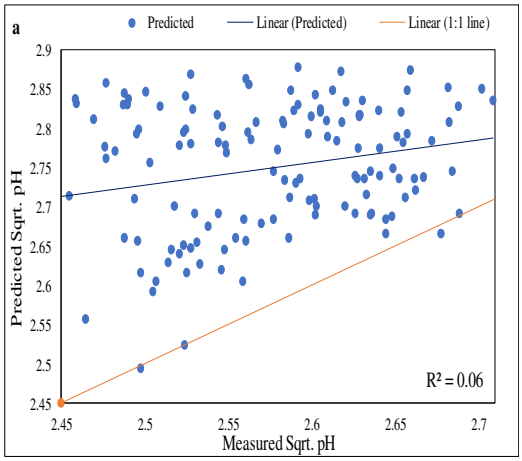
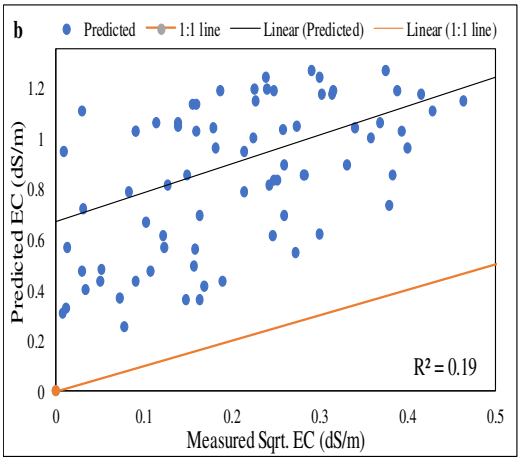
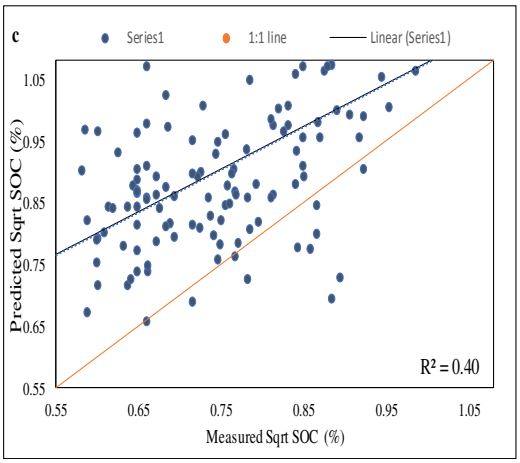


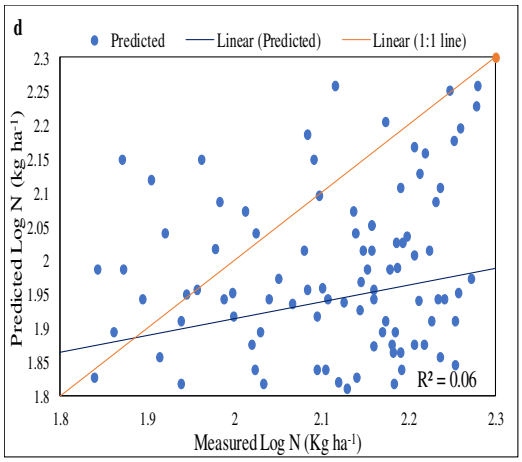
Fig. 4.26 Regression coefficients of (a) Soil pH (b) EC (c) SOC (d) Available N (e) Olsen-P (f) Available K (g) Soil calcium carbonate (h) DTPA-Fe (i) DTPA-Mn (j) DTPA-Cu (k) DTPA-Zn (l) Sand (m) Silt (n) Clay for South Western Alluvial Plain

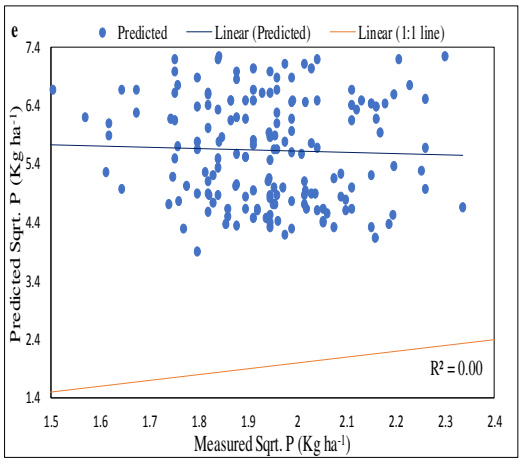
4.2.6 Common spectral models (for all agroecological subregions)

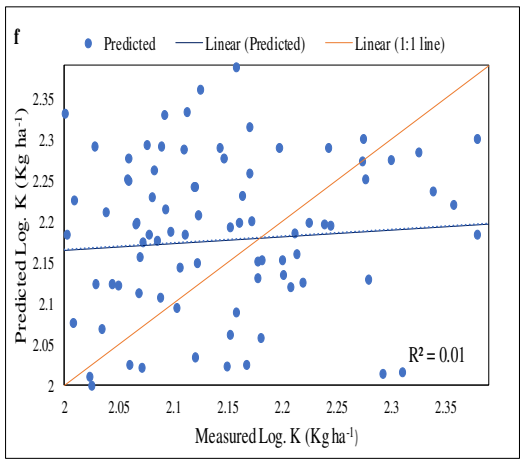


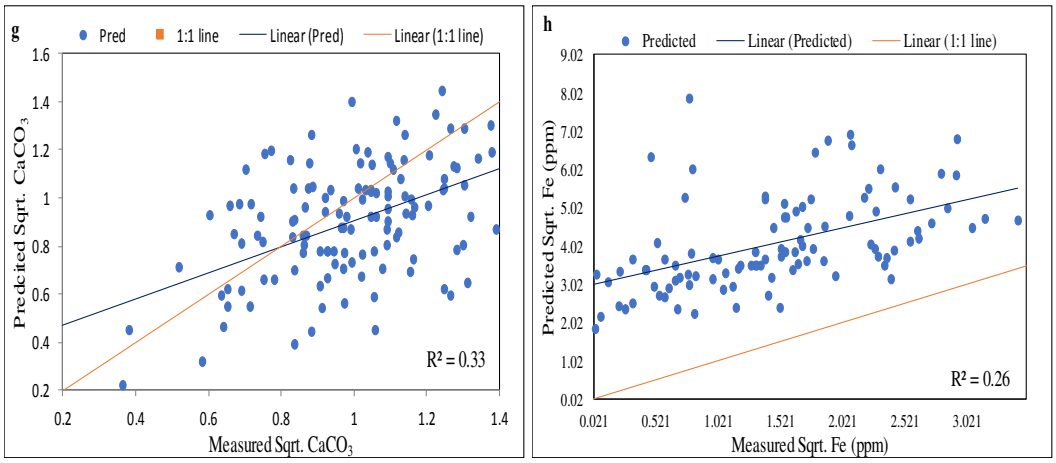


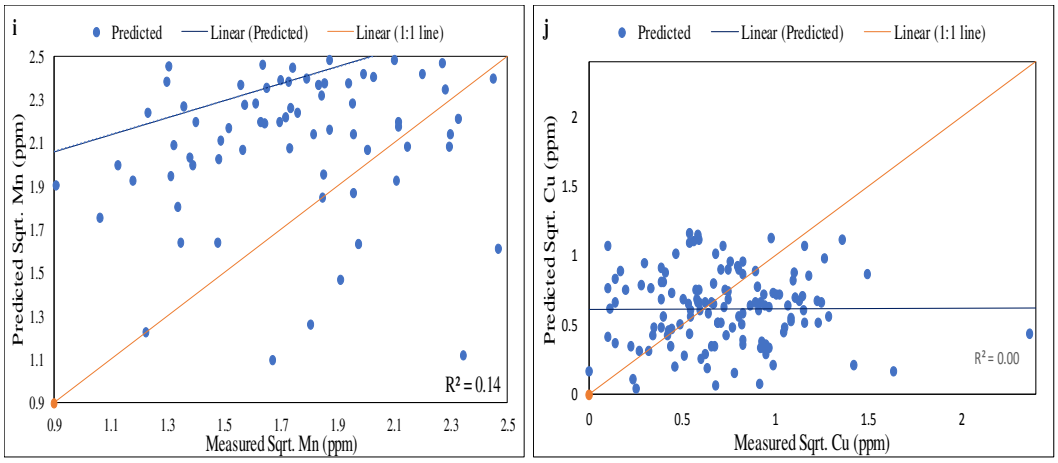


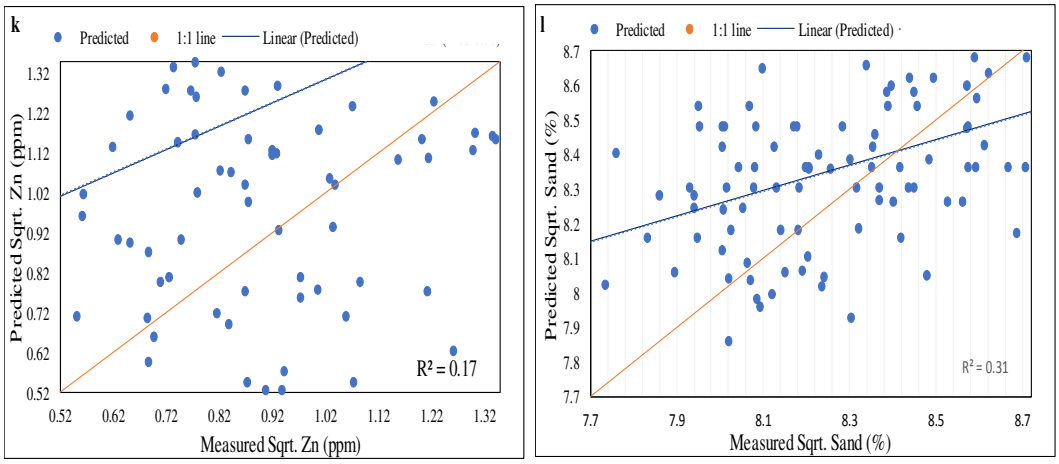







g
Pred
1:1 line
Linear (Pred)
Linear (1:1 line)
Predicted Sqrt. $CaCO_3$
Measured Sqrt. $CaCO_3$
$R^2 = 0.33$
h
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Fe (ppm)
Measured Sqrt. Fe (ppm)
$R^2 = 0.26$


i
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Mn (ppm)
Measured Sqrt. Mn (ppm)
$R^2 = 0.14$
j
Predicted
Linear (Predicted)
Linear (1:1 line)
Predicted Sqrt. Cu (ppm)
Measured Sqrt. Cu (ppm)
$R^2 = 0.00$


k
Predicted
1:1 line
Linear (Predicted)
Predicted Sqrt. Zn (ppm)
Measured Sqrt. Zn (ppm)
$R^2 = 0.17$
l
Predicted
1:1 line
Linear (Predicted)
Predicted Sqrt. Sand (%)
Measured Sqrt. Sand (%)
$R^2 = 0.31$

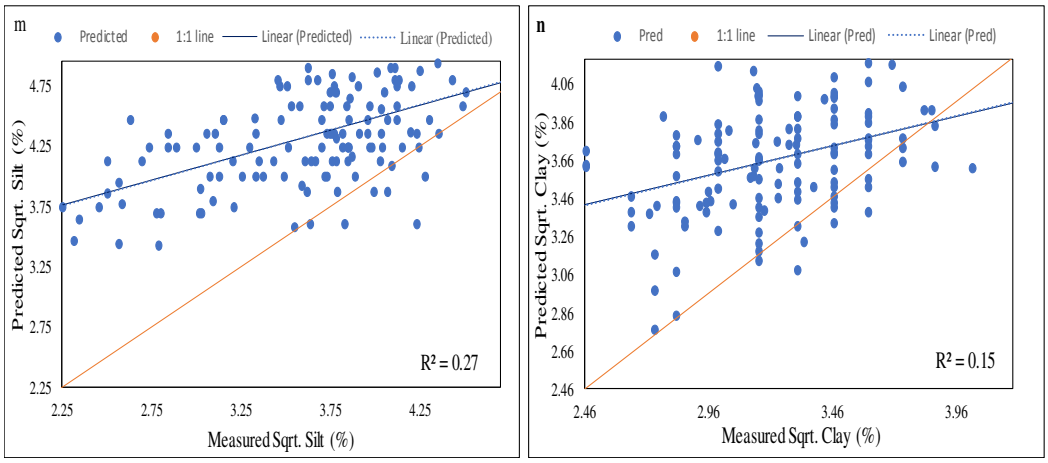


Fig. 4.27 Measured and predicted values, fitted linear line and 1:1 line of (a) Square root pH (b) Square root EC (c) Square root SOC (d) log N (e) Square root P (f) log K (g) Square root $CaCO_3$ (h) Square root Fe (i) Square root Mn (j) Square root Cu (k) Square root Zn (l) Square root Sand (m) Square root Silt (n) Square root Clay for Common spectral models.

The standard deviation (SD), root mean square error (RMSE), ratio of performance deviation (RPD) and R^2 values ranged between 0.07-0.50, 0.20-3.81, 0.04-0.98, 0.01-0.40, respectively among all the soil physical and chemical properties for the soil samples collected from five agroecological sub-regions (Fig 4.27 (a-n) and Table 4.25). The R^2 value for the SOC (0.40), $CaCO_3$ (0.33), sand (0.31), DTPA-Fe (0.26) and silt (0.27) was to be more than 0.25, whereas the R^2 value for phosphorous (0.01), potassium (0.01) and pH (0.01) was very low. Only the RPD for SOC and calcium carbonate were observed higher but it cannot be acceptable. Thus, we failed to develop a suitable model that could work across all regions. The R^2, RMSE, RPD, SD for all parameters are presented in the Table 4.25.

Table 4.25 Statistical parameters for validation of common model developed for predicting soil properties of any agroecological region

Soil property	Statistical parameter			
	SD	**RMSE**	**RPD**	**R^2**
pH*	0.07	0.20	0.34	0.01
EC *(dS/m)	0.12	0.75	0.16	0.19
SOC (%)	0.12	0.19	0.64	0.40
N# (kg ha^{-1})	0.16	0.28	0.57	0.06
P* (kg ha^{-1})	0.16	3.81	0.04	0.01
K# (kg ha^{-1})	0.15	0.20	0.71	0.01
$CaCO_3$ *(%)	0.25	0.26	0.98	0.33
DTPA-Fe* (ppm)	0.50	2.34	0.21	0.26
DTPA-Mn* (ppm)	0.44	0.67	0.65	0.14
DTPA-Cu* (ppm)	0.27	0.55	0.50	0.01
DTPA-Zn* (ppm)	0.27	0.47	0.57	0.17
Sand* (%)	0.27	0.34	0.79	0.31
Silt* (%)	0.42	0.85	0.50	0.27
Clay* (%)	0.36	0.57	0.62	0.15

(*) Square root transformation
(#) Log transformation

4.2.7 Sensitive spectral bands for various soil properties

Table 4.26 lists sensitive wavelength bands or wavelengths that contribute most to the prediction of a given soil property. Across various agroecological regions the most sensitive bands found for soil pH were 460 nm,510 nm, 1020 nm and 2120 nm, However, in case of electrical conductivity (EC dS/m) the wavelengths which contributed more in prediction were 1040 nm and 2150 nm. Different spectral bands like 460 nm, 470 nm, 490 nm, 1040 nm, 1910 nm, 2050 nm and 2110 nm affected the prediction of soil organic carbon (%) across different sub regions. For $KMnO_4$ oxidisable nitrogen, sensitive spectral bands were 460 nm, 910 nm, 1040 nm, 1660nm, 1860 nm and 2110 nm. Olsen-P prediction was affected by the spectral bands at 500 nm, 1000 nm, 1910 nm, 2020 nm and 2150 nm. In case of calcium carbonate ($CaCO_3$) estimation, wavelengths like 440 nm, 510 nm, 1030 nm, 2040 nm, 2120 nm, 2150 nm played a crucial role. Prediction of DTPA-extractable iron was affected by spectral bands like 460 nm, 950 nm, 1920 nm and 2150 nm. Also, for DTPA-extractable manganese important bands were 500 nm, 1050 nm, 1850 nm, 2070 nm and 2280 nm whereas in DTPA-copper estimation, wavelengths like 930 nm, 1000 nm, 1410 nm, 1830 nm, 1950 nm, 2030 nm and 2150 nm were important. However, for DTPA-extractable zinc, sensitive spectral bands were 460 nm, 550 nm, 1000 nm, 1940 nm and 2150 nm. In sand prediction, spectral bands like 460 nm, 500 nm, 990 nm, 1910 nm, 2010 nm, 2130 nm were detected. For silt estimation across various agroecological sub regions, 470 nm, 950 nm, 1660 nm, 1850 nm,

Table 4.26 Sensitive spectral bands for various soil properties

Soil property	Agroecological sub-region				
	SSH*	**NEU***	**PAP***	**CAP***	**SWAP***
pH	460,2020,2120	500, 1020, 2140	510, 1890, 2090	510, 2070	470,1970,2070 2130
EC (dS m^{-1})	1000,2140,2200	440, 1040, 2080	950, 2010, 2130	470, 2030, 2070	1050, 2150
SOC (%)	460,1040,1980	1000, 2000, 2100	490, 1830, 2050	1910, 1990, 2070	470, 1730, 2110, 2130
N (kg ha^{-1})	440,460,480, 2040, 2120	500, 1040, 2180	970, 1660, 1870	470, 950, 2020	910, 990, 1910, 2110
P (kg ha^{-1})	1840, 2020, 20802100	500, 1000, 1900	470, 1970, 2070	470, 1010, 1910, 2110	950, 2050, 2150
K (kg ha^{-1})	1880, 2140, 2200	1060, 1840, 2100	1010, 1970,2050	490, 1110, 1210	1970, 2070
$CaCO_3$ (%)	440, 2040	460, 100, 2120	510, 950, 2030	510, 970, 1030	950, 990, 2050, 2070, 2150
DTPA-Fe (ppm)	440, 460, 2000	480, 1920, 1940	1850, 1910, 2030	470, 1950, 2029	950, 2090, 2150
DTPA-Mn (ppm)	500, 1880, 1940	1960, 2100, 2280	1830, 1850, 1910	470, 1050, 1930, 2070	490, 990, 2010, 2110
DTPA-Cu (ppm)	1000, 1880 2140	1000, 2100, 2140	950, 2010, 2130	930, 1050, 1410, 1830, 1950, 2030	1930, 2010, 2110, 2150
DTPA-Zn (ppm)	460, 1000,1940	1000, 2140, 2180	950, 2010	470, 970, 1090	470, 550, 950, 1010, 2150
Sand (%)	440, 500, 1000	460, 1040, 2100	1910, 2050, 2090	2010, 2130, 2150	470, 990, 1910
Silt (%)	1020, 1660, 2140	1000, 2020, 2100	950, 2090, 2130	470, 1810, 1990	490, 1050, 1850, 2150
Clay (%)	1020, 1920, 2020	480, 1000, 2120	490, 1910, 2010	470, 950, 1010, 1950	470, 1990, 1850, 2130

` (*): SSH- Sub-mountain Siwalik Hills, NEU- North eastern Undulating, PAP- Piedmont and Alluvial Plain, CAP- Central Alluvial Plain, SWAP- South Western Alluvial Plain

1990 nm, 2020 nm and 2140 nm wavelengths were sensitive bands. Finally, clay estimation was affected by spectral bands at 480 nm, 1020 nm, 1850 nm, 1920 nm and 2130 nm across various agroecological sub regions. It can be concluded that different soil properties like soil pH, soil organic carbon, calcium carbonate, DTPA-extractable iron, manganese, zinc and sand are very well predicted in visible (VIS) as well as near infrared (NIR) regions. However, electrical conductivity, available phosphorous, potassium, silt and clay content were predicted better in near infrared region (NIR region).

4.3 Portability of spectral models across various agroecological sub regions

Preceding results clearly show that developed spectral models were highly region specific and thus lacked portability. However, some generalizations with respect to the predictability of a property can be made. Soil properties like calcium carbonate and sand content were very well predicted for most of the agroecological sub-regions. Fitted PLSR models had different predicting capabilities for different regions in case of DTPA-extractable micronutrients. Soil organic carbon and available phosphorous were moderately well predicted in all sub regions. Soil available nitrogen and potassium gave moderately accurate prediction except the model developed for south western alluvial plain. Soil pH and electrical conductivity (EC dS/m) were poorly predicted for most of the agroecological sub regions. In case of soil texture except sand content, silt and clay content were poorly predicted.

However, more complex statistical tools and other techniques like Artificial neural network (ANN), random forest (RF), boosted tree (BT) which are non-linear statistical techniques can be used to improve prediction. Artificial neural network can detect problems in models automatically and can solve the problems.

CHAPTER-V

SUMMARY

Various chemical methods are being used to assess nutrient supplying capacity. Precise analysis of soil physical properties like soil texture is often skipped to save both time and money. These methods, however, are generally laborious, time consuming and relatively uneconomical. As a result, for improving speed and economics of soil testing for fertility assessment purposes, the sampling quality is greatly compromised. Not many samples are collected from each farming unit. However, spatial variability at field scale is too considerable to ignore it. The ideal soil property assessment methods need to duly consider soil spatial variability and should be rapid, reproducible and cost effective. Further, most of the soil testing laboratories do not test for microbiological properties. Soil testing methods that are rapid, capacious, relatively inexpensive, and which are able to duly consider spatial variability are required to ensure site-specific nutrient management for the larger cause of popularization of precision agriculture in the country. Visible-Near Infrared (VIS-NIR, 350-2500nm) diffuse reflectance spectroscopy holds a great promise in honouring these conditions and thus ensuring effective site-specific nutrient management by providing high-density data at less cost. This technique involves directing the radiation in the VIS-NIR region onto the soil sample. A single scan of a soil sample can be used for prediction of multiple properties. Different agroecological sub-regions are usually inhabited by various types of soils. For instance, five agroecological sub-regions of Punjab state are inhabited by around 17 benchmark soils. Also, these regions generally tend to support specific ecosystems. Keeping this in view, VIS-NIR models can be developed at different scales (agroecological sub-regions) and subsequently examined for their applicability across various scales. The objectives of this studies were aligned with these researchable issues.

Geotagged surface soil samples (0-15 cm) were collected randomly by using stratified random sampling. From each agroecological sub region, 100 samples were collected. Collected soil samples were air dried under shade. Different kind of pebbles, stones, concretions, etc. along with plant roots were removed carefully. By using agate mortar, large aggregates and clods were broken properly. Then afterwards, each sample was passed through 2 mm (60 mesh) sieve. After processing, soil samples were divided into two sub samples. One part was used for traditional soil testing for determining physico-chemical parameters (soil texture, pH, electrical conductivity (EC), soil organic carbon (SOC), Available N, Available P, Available K, $CaCO_3$, DTPA extractable Fe, Mn, Cu, Zn).and other part was used for hyperspectral analysis by using spectroradiometer (ASD FieldSpec® Pro FR, No. A110070).

The prospectr package was used for the processing of the data like noise removal and second order derivative transformation. Finite approach for derivative and the leave one out

approach of cross validation was used. The pls package was used to fit the partial least square regression, where the processed spectral data were used as dependent variable and observed soil property was used as an independent variable. Number of components was so selected such that root mean square error (RMSE) was minimum and R^2 was maximum. Ratio of performance to deviation (RPD), root mean square error (RMSE), and standard deviation (SD) were some of the other statistical characteristics that were employed for model quality assessment.

The SD, RMSE, RPD and R^2 values ranged between 0.03-0.37, 0.05-0.34, 0.32-1.73 and 0.05-0.69, respectively among all the soil physical and chemical properties for the soil samples collected from Sub-mountain Siwalik Hills. Considering most of the statistical parameters, model developed for this region can be used for prediction of sand, $CaCO_3$ and clay content for Sub-mountain Siwalik hill. Similarly, the R^2 value for the $CaCO_3$ (0.82), sand (0.55) and DTPA-Fe (0.52) were observed higher in the Northeastern Undulating. With the RPD of more than 1.4 for $CaCO_3$ and DTPA-extractable Fe. Thus, the model developed for this region can be used for prediction of sand, $CaCO_3$ and Fe content. In case of South Western Alluvial Plain, the RPD value for sand and $CaCO_3$ was observed higher. All the four DTPA-extractable micronutrients were poorly predicted in case of samples collected from South Western Alluvial plain. Available N was also moderately well described in South Western Alluvial Plain. The amount of sand, $CaCO_3$, DTPA-extractable Zn and available P content was adequately predicted in case of Piedmont and Alluvial Plain. The RPD values for the available N (1.40) and P (1.41) were also observed relatively higher in this region, again suggesting the suitability of developed model for predicting these properties. Although a high R^2 was observed for DTPA-extractable Fe content, the RPD was relatively lower because of higher variation and error in case of data used for developing model for Central Alluvial Plain. The Cu and Zn levels also could not be predicted well in the case of Central Alluvial Plain. The study concludes that the prediction by spectral models is highly region specific particularly for micronutrients. The model fitted for a specific region cannot be used in some other region suggesting thereby lack of portability. When considering the amenability of various properties to prediction through spectral models, it can be concluded that calcium carbonate and sand can be reasonably well predicted. On the other hand, however, soil pH and EC cannot be predicted satisfactorily. However, review of the available literature evidences that there exist many models that predict these properties adequately and vice-versa. The study suggests that better statistical tools need be employed to improve performance of spectral models.

REFERENCES

Adamchuk V I, Hummel J W, Morgan M T and Upadhyaya S K (2004) On-the go soil sensors for precision agriculture. *Comput Electron Agric* **44**: 71–91.

Askari M S, Cui J, O'Rourke S M and Holden N M (2015) Evaluation of soil structural quality using VIS-NIR spectra. *Soil Tillage Res* **146**: 108-17

Bai J, Ye X, Jia J, Zhang G, Zhao Q, Cui B and Liu X (2017) Phosphorus sorption-desorption and effects of temperature, pH and salinity on phosphorus sorption in marsh soils from coastal wetlands with different flooding conditions. *Chemosphere* **188**: 677-88.

Behera S K and Shukla A K (2015) Spatial distribution of surface soil acidity, electrical conductivity, soil organic carbon content and exchangeable potassium, calcium and magnesium in some cropped acid soils of India. *Land Degrad Dev* **26**: 71-79.

Bellon-Maurel V and McBratney A (2011) Near Infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils- Critical view and research perspectives. *Soil Biol Biochem* **43**:1398-1410.

Ben-Dor E and Banin A (1995) Quantitative analysis of convolved thematic mapper spectra of soils in the visible near-infrared and shortwave-infrared spectral regions (0.4–2.5 μm). *Int J Rem Sens* **16**: 3509–28.

Ben-Dor E, Inbar Y and Chen Y (1997) The reflectance spectra of organic matter in the visible near-infrared and short-wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sens Environ* **15**: 1–15.

Ben-Dor E, Ong C and Lau I C (2015) Reflectance measurements of soils in the laboratory: standards and protocols. *Geoderma* **245**: 112–24.

Ben-Gera I and Norris K H (1968) Determination of moisture content in soybeans by direct spectrophotometry. *Israel J Agric Res* **18**: 124-32.

Bertrand I, Janik L J, Holloway R E, Armstrong R D and Mclaughlin M J (2002) The rapid assessment of concentrations and solid phase associations of macro- and micronutrients in alkaline soils by mid infrared diffuse reflectance spectroscopy. *Aust J Soil Res* **40**: 1339-56.

Bilgili A V, van Es H M, Akbas F, Durak A and Hively W D (2010) Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *J Arid Environ* **74**: 229–38.

Bilgili A V, Cullu M A, van Es H, Aydemir A and Aydemir S (2011) The use of hyperspectral visible and near infrared reflectance spectroscopy for the characterization of salt-affected soils in the Harran plain, Turkey. *Arid Land Res Manag* **25**: 19–37.

Bowers S and Hanks R (1965) Reflection of radiant energy from soils. *Soil Sci* **100** :130–38.

Breiman L (2001) Random forests. *Mach Learn* **45**: 5–32.

Brooks F A (1952) Atmospheric radiation and its reflection from the ground. *J Meteorol* **9**: 41–52.

Brown D J (2007) Using a global VNIR soil-spectral library for local soil characterization and landscape modeling in a 2nd-order Uganda watershed. *Geoderma* **140**: 444–53.

Brown D J, Shepherd K D, Walsh M G, Mays M D and Reinsch T G (2006) Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* **132**: 273–90.

Buddenbaum H and Steffens M (2012) The effects of spectral pretreatments on chemometric analyses of soil profiles using laboratory imaging spectroscopy. *Appl Environ Soil Sci* **14**: 1–12.

Burrough P A (1986) *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford: Oxford University Press.

Cassman K G (1999) Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. *Crop Sci* **32**: 1251–58.

Chang C W, Liard D A, Mausbach M J Hurburgh Jr C R (2001) Near-Infred reflectance spectroscopy-principal components regression analysis of soil properties. *Soil Sci Soc of Am J* **65**: 480-90.

Charles W W N, Liu J, Chen R and Xu J (2015) Physical and numerical modeling of an inclined three-layer (silt/gravelly sand/clay) capillary barrier cover system under extreme rainfall. *J Waste Manag* **38**: 210-21.

Chodak M, Niklinska M and Beese F (2007) Near-infrared spectroscopy for analysis of chemical and microbiological properties of forest soil organic horizons in a heavymetal-polluted area. *Biol Fertil Soils* **44**: 171–80.

Cozzolino D and Moron A (2003) The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. *J Agric Sci* **140**: 65– 71.

Curcio D, Ciraolo G, D'Asaro F and Minacapilli M (2013) Prediction of soil texture distributions using VNIR–SWIR reflectance spectroscopy. *Procedia Environ Sci* **19**: 494–503.

Dalal R C and Henry R J (1986) Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry. *J Soil Sci Soc Am* 50: 120–23.

Daniel K W, Tripathi N K and Honda K (2003) Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Aust J Soil Res* **41**: 47–59.

Davies T (2005) An introduction to near infrared spectroscopy. *NIR News* **16**: 9-11.

Day P R (1965) Particle fractionation and particle size analysis. In: Black C A et al (ed) Methods of soil analysis, Part 1. *Agron J* **43**: 1004-07.

Debicka M, Kocowicz A, Weber J and Jamroz E (2016) Organic matter effects on phosphorus sorption in sandy soils. *Arch Agron Soil Sci* **62**: 840-85.

deMelo T Mand Pedrollo O C (2015) Artificial Neural Networks for Estimating Soil Water Retention Curve Using Fitted and Measured Data. *Appl Environ Soil Sci* **20**: 15-16.

Dick R P, Breakwell D P and Jurco R F (1996) Soil enzyme activities and biodiversity measurements as integrative microbiological indicators. *Soil Sci Soc Am J* **49**: 247-71.

Dick W A, Thavamani B, Conley S, Blaisdell R and Sengupta A (2013) Prediction of β-glucosidase and β-glucosaminidase activities, soil organic C, and amino sugar N in a diverse population of soils using near infrared reflectance spectroscopy. *Soil Biol Biochem* **56**: 99–104.

Disla S J M, Janik L J, Viscarra Rossel R, MacDonald M L and McLaughlin M J (2014) The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. *Appl Spectrosc Rev* **49**: 139-86.

Du C and Zhou J (2009) Evalution of soil fertility using infrared spectroscopy: A review. *Environ Chem* **7**: 97-113.

Elith J, Leathwick J R and Hastie T (2008) A working guide to boosted regression trees. *J Anim Ecol* **77**: 802–13.

Farifteh J, Van der Meer F, Atzberger C and Carranza E J M (2007) Quantitative analysis of salt-affected soil reflectance spectra: a comparison of two adaptive methods (PLSR and ANN)**.** *Remote Sens Environ* **110**: 59-78.

Felicísimo Á M, Cuartero A, Remondo J and Quirós E (2012) Mapping landslide susceptibility with logistic regression, multiple adaptive regression splines, classification and regression trees, and maximum entropy methods: a comparative study. *Landslides* **10**: 175–89.

Friedman J H (1991) Multivariate adaptive regression splines. *Ann Stat* **19**: 1-67.

Friedman J H (2002) Stochastic gradient boosting. *Comput Stat Data Anal* **38**: 367–78.

Gebbers R and Adamchuk V I (2010) Precision agriculture and food security. *Sci Mag* **827**: 828–31.

Geypens M, Vanongeval L, Vogels N and J Meykens (2013) Spatial variability of agricultural fertility parameters in a gleyic podzol of Belgium*. Prec Agric* **1**: 319-26.

Gomez C, Viscarra Rossel R A and McBratney A B (2008) Soil organic carbon prediction by hyperspectral remote sensing and field vis–NIR spectroscopy: an Australian case study. *Geoderma* **146**: 403–11.

Grimm R, Behrens T, Märker M and Elsenbeer H (2008) Soil organic carbon concen-trations and stocks on Barro Colorado Island-digital soil mapping using RandomForests analysis. *Geoderma* **146**: 102–13.

Guerrero C, Zornoza R, Gomez I and Mataix-Beneyto J (2010) Spiking of NIR regional models using samples from target sites: effect of model size on prediction accuracy. *Geoderma* **158**: 66–77.

Hamdan N, Kavazanjian E, Rittmann B E and Karatas I (2017) Carbonate mineral precipitation for soil improvement through microbial denitrification. *Geomicrobiology* **34**: 139-46.

Hartley W, Riby P and Waterson J (2016) Effects of three different biochars on aggregate stability, organic carbon mobility and micronutrient bioavailability. *J Environ Manag* **181**: 770-78.

Hengel AV, Paisitkriangkrai S, Sherrah J and Janney P (2015) *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* Tech Bull, 7th edn. Pp 36-43. Workshops.

Hu G, Sudduth K A, He D, Myers D B and Nathan M V (2016) Soil phosphorus and potassium estimation by reflectance spectroscopy. *Trans ASABE* **59:** 97–105.

Huang J, Hu B, Qi K, Chen W, Pang X, Bao W, Tian G (2016) Effects of phosphorus addition on soil microbial biomass and community composition in a subalpine spruce plantation. *Eur J Soil Biol* **72**: 35-41.

Islam K, Singh B and McBratney A B (2003) Simultaneous estimation of various soil properties by ultra-violet, visible and near-infrared reflectance spectroscopy. *Aust J Soil Res* **41**: 1101– 14.

Jackson M L (1973) Soil chemical analysis-advanced course. A manual of methods useful for instruction and research in soil chemistry, physical chemistry, soil fertility and soil genesis. 2nd edition, Madison, US.

Janik L J and Skjemstad J O (1995) Characterisation and analysis of soils using mid-infrared partial least squares: II. Correlations with some laboratory data. *Aust J Soil Res* **33**: 637–50.

Janik L J, Merry R H and Skjemstad J O (1998) Can mid infrared diffuse reflectance analysis replace soil extractions? *Aust J Exp Agric* **38:** 681-96.

Ji W, Viscarra Rossel R A and Shi Z (2015) Accounting for the effects of water and the environment on proximally sensed Vis–NIR soil spectra and their calibrations. *Eur J Soil Sci* **66**: 555–65.

Kennedy A (2002) *The Global Positioning System and GIS*: An Introduction 2nd ed. *Taylor and Francis* pp345.

Keykha H A, Asadi A and Zareian M (2017) Environmental factors affecting the compressive strength of microbiologically induced calcite precipitation-treated soil. *Geomicrobiology* **34**: 65-71.

Knadel M, Arthur E and Weber P (2018) Soil Specific Surface Area Determination by Visible Near-Infrared Spectroscopy. *Soil Sci Soc Am J* **82**:1046–56.

Kodaira M and Shibusawa S (2013) Using a mobile real time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma* **199**: 64-79.

Lagacherie P, Baret F, Feret J B, Netto J M and Robbez-Masson J M (2008) Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sens Environ* **112**: 825–35.

Lee K S, Lee D H, Sudduth K A, Chung S O, Kitchen N R and Drummond S T (2009) Wavelength identification and diffuse reflectance estimation for surface and profile soil properties. *Trans ASABE* **52**: 683–95.

Lee K S, Lee D H, Sudduth K A, Drummond S T, Kitchen N R and Chung S (2010) Calibration methods for soil property estimation using reflectance spectroscopy. *Trans ASABE* **53**: 675–84.

Levene H (1960*) Robust tests for equality of variances. In: Contributions to probability and Statistics: Essay in honour of Harold Hotelling*. Standford, California: Standford University Press.

Li B, Morris J and Martin E B (2002) Model selection for partial least squares regression. *Intell Lab Syst* **64**: 79–89.

Lindsay W L and Norvell W A (1978) Development of DTPA soil test methods for zinc, iron, manganese and copper. *Soil Sci Soc Am J* **42**: 421-28.

Lu P, Wang L, Niu Z, Li L and Zhang W (2013) Prediction of soil properties using laboratory VIS–NIR spectroscopy and Hyperion imagery. *J Geochem Explor* **132**: 26–33.

Ludwig B, Khanna P K, Bauhus P and Hopmans P (2002) Near infrared spectroscopy of forest soils to determine chemical and biological properties related to soil sustainability. *Forest Ecol Manag* **171**: 121-32.

Mahmood R and Babel M (2013) Evaluation of SDSM developed by annual and monthly sub-models for downscaling temperature and precipitation in the Jhelum basin, Pakistan and India. *Theor and Appl Climatol* **133**: 27-44.

Mandal S, Thangarajan R, Bolan N S, Sarkar B, Khan N, Sik Y and Naidu R (2016) Biochar-induced concomitant decrease in ammonia volatilization and increase in nitrogen use efficiency by wheat. *Chemosphere* **142**: 120-27.

Martens H and Naes T (1989) Multivariate Calibration John Wiley & Sons, Chichester, UK. 419.

Mashimbye Z E, Cho M A, Nell J P, Clercq D E W P, van Niekerk A and Turner D P (2012. Model-based integrated methods for quantitative estimation of soil salinity from hyperspectral remote sensing data: a case study of selected south African soils. *Pedosphere* **22**: 640–49.

Mayrink G O, Valente D S M, Queiroz D M, Pinto F A Cand Teofilo R F (2019) Determination of chemical soil properties using diffuse reflectance and ion-exchange resins *Prec Agric* 20 :541–61.

Merwin H D and Peech M (1950) Exchangeability of sodium potassium in the sand, silt and clay fractions as influenced by the nature of complimentary exchangeable cations. *Soil Sci Soc Amer Proc* **15**: 125-28.

Minasny B, McBratney A B, Bellon-Maurel V, Roger J M, Gobrecht A, Ferrand L and Joalland S (2011) Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma* **167**: 118–24.

Mohamdimonavar H (2016) Determination of several soil properties based on ultraviolet, visible, and near-infrared reflectance spectroscopy. *Proc of The IRES Int Conf.* pp 978-93. Copenhagen, Denmark.

Mondal B P (2017) *Geostatistical assessment of spatial variability of soil fertility and its implications in calibration of visible-near infrared models*. M.Sc. thesis, Punjab

Agricultural University, Ludhiana, India.

Mondal B P, Sekhon B S, Sharma S, Singh M, Sahoo R N, Barman A, Sinha Y, Chattopadhyay A and Banerjee K (2017) VIS-NIR Reflectance spectroscopy for assessment of soil microbiological properties. *Int J Curr Microbiol App Sci* **6**: 719-28.

Morellos A, Pantazi X E, Moshou D, Alexandridis T, Whetton R, Tziotzios G, Wiebensohn J, Bill R and Mouazen A M (2016) Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. *Biosyst Eng* **152**: 104-16.

Mouazen A M, Karoui R, De Baerdemaeker J & Ramon H (2006) Characterization of Soil Water Content Using Measured Visible and Near Infrared Spectra. *Soil Sci Soc Am J* **70**: 1295.

Mouazen A M, Kuang B, De Baerdemaeker J and Ramon H (2010) Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma* **158**: 23–31.

Mouazen A M and Kuang B (2016). On-line visible and near infrared spectroscopy for in-field phosphorous management. *Soil Tillage Res* **155**: 471–77.

Mulder V L, de Bruin S, Scheapman M E and Mayr T R (2011) The Use Remote Sensing and Terrain mapping-A Review. *Geoderma* **162**: 1-19.

Nanni M R and Demattê (2006) Spectral reflectance methodology in comparison to traditional soil analysis. *Soil Sci Soc Am J* **70**: 393–407.

Nduwamungu C, Ziadi N, Trembley G F Parent L E and Thuries B (2009) Near-infrared reflectance spectroscopy prediction of soil properties: Effects of sample cups and preparation. *Soil Sci Soc Am J* **73***:* 1896-1903.

Olsen S R, Cole C V, Watanabe F S and Dean L A (1954) Estimation of available phosphorous in soils by extraction with sodium bicarbonate. *U S Department of Agriculture Circular* Pp: 939, USDA, Washington DC.

Owen A (2000) *Fundamentals of Modern UV-visible Spectroscopy*. Agilent Technologies.

Parat C, Chaussod R, Levenque J, Dousset S and Andreux F (2002) The relationship between copper accumulated in vineyard calcareous soils and soil organic matter and iron. *Eur J Soil Sci* **53**: 663-69.

Pirie A, Singh B and Islam K (2005) Ultra-violet, visible, near-infrared, and mid-infrared diffuse reflectance spectroscopic techniques to predict several soil properties. *Aust J Soil Res* **43**: 713-21.

Puri A N (1930) A new method for estimating total carbonates in soils. *Imp Agri. Res Pusa Bull*. **206**: 7.

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rajkumar, Singh B, Kaur B and Beri V (2008) Planning for precision farming in different

agro-ecological sub-regions of Punjab-role of natural resources in agricultural research, planning, and development and transfer of technology. Department of Soils, Punjab Agricultural University, Ludhiana.

Reeves J B, McCarty G W and Meisinger J J (2000) Near infrared reflectance spectroscopy for determination of biological activity in agricultural soils. *J Near Infrared Spectrosc* **8**: 25–34.

Reeves J B and McCarty G W (2001) Quantitative analysis of agricultural soils using near infrared reflectance spectroscopy and fibre-optic probe. *J Near Infrared Spectrosc* **9**: 25–34.

Reeves J B and Smith D B (2009) The potential of mid-and near-infrared diffuse reflectance spectroscopy for determining major-and trace-element concentrations in soils from a geochemical survey of North America. *Appl Geochem* **24**: 1472–81.

Reeves J B (2010) Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: where are we and what needs to be done? *Geoderma* **158**: 3–14.

Rinnan Å, Van Den Berg F and Engelsen S B (2009) Review of the most common pre-processing techniques for near-infrared spectra. *Anal Chem* **28**: 1201–22.

Sahoo R N, Ray S S and Manjunath K R (2015) Hyperspectral Remote Sensing In agriculture. *Curr Sci* **108**. 848-59.

Samui P (2012) Multivariate adaptive regression spline (mars) for prediction of elastic modulus of jointed rock mass. *Geotech Geol Eng* **31**: 249–53.

Sarathjith M C, Das B S, Vasava H B, Mohanty B, Sahadevan A S, Wani S P and Sahrawat K L (2014) Diffuse Reflectance Spectroscopic Approach for the Characterization of Soil Aggregate Size Distribution. *Soil Sci Soc Am J* **78**: 369.

SAS Institute (2013) SAS/STAT 9.3 User's guide. SAS Institute, Cary, NC, USA, ***P***. 8640.

Savitzky A and Golay M J E (1964) Smoothing and differentiation of data by simplified least square procedure. *Anal Chem* **36**: 1627-39.

Saxena R K, Srivastava R and Verma K S (2005) Spectral reflectance library of Indian soils. NBSS Report No. 835, NBBS&LUP, Nagpur.

Serrano J, da Silva J M, Shahidian S, Silva L L, A. Sousa A and Baptista F (2017) Differential vineyard fertilizer management based on nutrients spatio-temporal variability. *J Soil Sci Plant Nutr* **17**: 154-56.

Shepherd K D and Walsh M G (2002) Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci Soc Am J* **66:** 988–98.

Shepherd K D and Walsh M G (2007) Infrared Spectroscopy—Enabling an Evidence-Based Diagnostic Surveillance Approach to Agricultural and Environmental Management in Developing Countries. *J Near Infrared Spectrosc* **15**: 1-19.

Shibusawa S, Imade Anom S W, Sato S, Sasao A and Hirako S (2001) Soil mapping using the real-time soil spectrophotometer. *3rd Eur Conf on Precision Agriculture*, pp 497– 508. Agro Montpellier.

Shuman L M (1986) Effect of liming on the distribution of manganese, copper, iron and zinc among soil fractions. *Soil Sci Soc Am J* **50**: 1236-40.

Singh S (2018) *Evaluation of soil reflectance properties for calibrating phosphorous fertilizer use in potato (Solanum tuberosum L.*). M.Sc. thesis, Punjab Agricultural University, Ludhiana, India.

Soriano-Disla J S, Janik L J, Viscarra Rossel R A, Mcdonald L M and McLaughlin M J (2014) The performance of visible, near-infrared reflectance spectroscopy for prediction of soil physical, chemical and biological properties. *Appl Spectrosc Rev* **49**: 139-86.

Srivastava R, Prasad J and Saxena R K (2004) Spectral reflectance properties of some shrink-swell soils of central India as influenced by soil properties. *Agropedology* **14**: 45-54.

Stenberg B (2010) Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma* **158:** 15-22.

Stenberg B, Viscarra Rossel R A, Mouazen A M and Wetterlind J (2010) Visible and near infrared spectroscopy in soil science. *Adv Agron* **107**: 163-215.

Stevens A, Udelhoven T, Denis A, Tychon B, Lioy R, Hoffmann L and van Wesemael B (2010) Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. *Geoderma* **158**: 32–45.

Stevens A and Ramirez-Lopez L (2013) An introduction to prospectr package. R package Vignette R package version 0.1.3.

Subbiah B V and Asija G L (1956) A rapid procedure for the determination of available nitrogen in soils. *Curr Sci* **25**: 259-60.

Suykens J, Vandewalle J (1999) Least squares support vector machines classifiers. *Neural Process Lett*: 293–300.

Tahir S and Marschner P (2016) Clay addition to sandy soil - effect of clay concentration and ped size on microbial biomass and nutrient dynamics after addition of low C/N ratio residue. *J Soil Sci Plant Nutr* **16**: 325-28.

Urselmans T T, Vagen G T, Spaargaren O and Shepherd D K (2010) prediction of soil properties from globally distributed soil mid-infrared spectral library. *Soil Sci Soc Am J* **74**: 1792-99.

Vapnik V N (1998) Statistical Learning Theory**.** John Wiley and Sons Inc, New York.

Vasques G M, Grunwald S and Sickman J O (2008**)** Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra**.** *Geoderma*, **146**: 14-25.

Viscarra Rossel R A, Walvoort D J J, McBratney A B, Janik L J and Skjemstad J O (2006) Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **131**: 59–75.

Viscarra Rossel R A (2007) Robust modelling of soil diffuse reflectance spectra by bagging-partial least squares regression. *J Near Infrared Spectrosc* **15**: 39–47.

Viscarra Rossel R A (2009) The Soil Spectroscopy Group and the development of a global soil spectral library. *NIR News* **20**: 14–15.

Viscarra Rossel R A, Cattle S, Ortega A and Fouad Y (2009) In situ measurements of soil colour, mineral composition and clay content by vis–NIR spectroscopy. *Geoderma* **150**: 253–66.

Viscarra Rossel R A and Behrens T (2010) Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **158**: 46-54.

Viscarra Rossel R A, Bui E N, de Caritat P and McKenzie N J (2010) Mapping iron oxides and the color of Australian soil using visible–near infrared reflectance spectra. *J Geophys Res Earth Surf* **115**. 85-93.

Viscarra Rossel R A and Webster R (2012) Predicting soil properties from the Australian soil visible–near infrared spectroscopic database. *Eur J Soil Sci* **63**: 848–60.

Viscarra Rossel R A and Hicks W S (2015) Soil organic carbon and its fractions estimated by visible–near infrared transfer functions. *Eur J Soil Sci* **66**: 438–50.

Viscarra Rossel R A, Behrens T, Ben-Dor E, Brownd D J, Demattê J A M, Shepherd K D, Shi Z, Stenberg B , Stevens A, Adamchukj V, Aïchi H, Barthès B G, Bartholomeus H M, Bayer A D, Bernoux M, Böttcher K, Brodský L, Du C W, Chappell A, Fouads Y, Genot V, Gomez C, Grunwald S, Gubler A, Guerrerox C, Hedley C B, Knadel M, Morrás H J M, Nocita M, Ramirez-Lopez L, Roudier P, Rufasto Campos E M, Sanbornae P, Sellitto V M, Sudduth K A, Rawlins B G, Walter C, Winowiecki L A, Hongai S Y and Ji W (2016) A global spectral library to characterize the world's soil. *Earth-Sci Rev* **155**: 198–230.

Vohland M, Besold J, Hill J, Fründ H C (2011) Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy. *Geoderma* **166**: 198-205.

Walkley A and Black I A (1934) An examination of the Degtjareff method for determining the soil organic matter and a proposed modification of the chromic acid titration method. *Soil Sci* **37**: 29-38.

Walvoort D J J and McBratney A B (2001) Diffuse reflectance spectrometry as a proximal sensing tool for precision agriculture. *3rd Eur Conf on Precision Agriculture*, vol 1, pp 503– 07. Agro Montpellier.

Wang R, Wang X, Jiang Y, Cerdà A, Yin J, Liu H, Feng X, Shi Z, Dijkstra F A and Li M H (2018) Soil properties determine the elevational patterns of base cations and micronutrients in the plant–soil system up to the upper limits of trees and shrubs *Biogeosciences* **15**:1763-74.

Wold S, Trygg J, Berglund A and Antti H (2001) Some recent developments in PLS modeling. *Chemom Intell Lab Syst* **58**: 131–50.

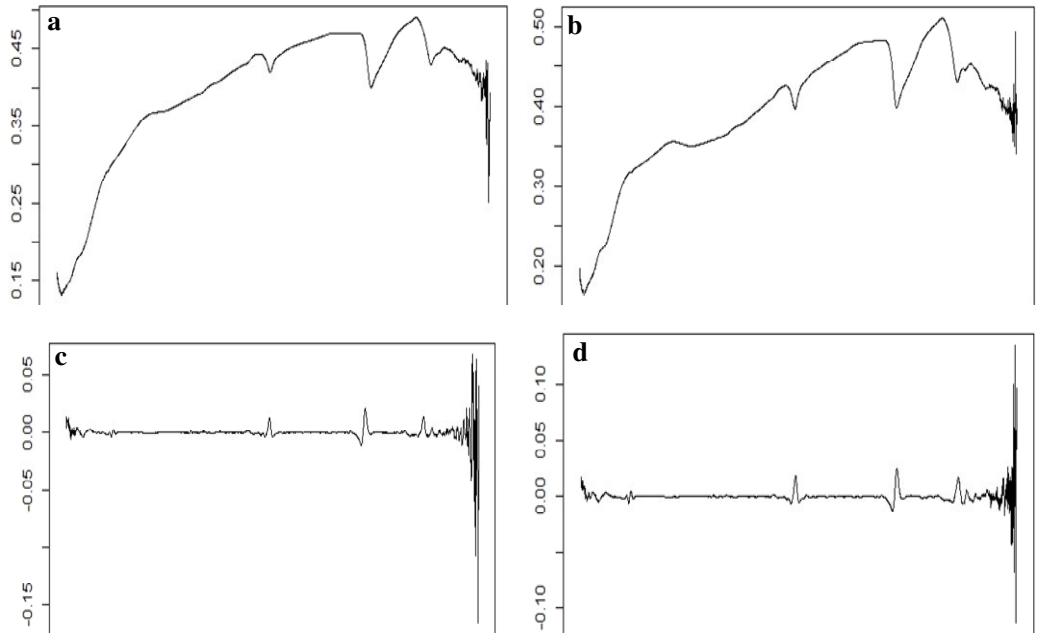
Xu S, Zhao Y, Wang M and Shi X (2018) Comparison of multivariate methods for estimating selected soil properties from intact soil cores of paddy fields by VIS–NIR spectroscopy. *Geoderma,* **310**: 29–43.

Xuemei L and Jianshe L (2013) Measurement of soil properties using visible and short wave-near infrared spectroscopy and multivariate calibration. *Measurement* **46**: 3808-14.

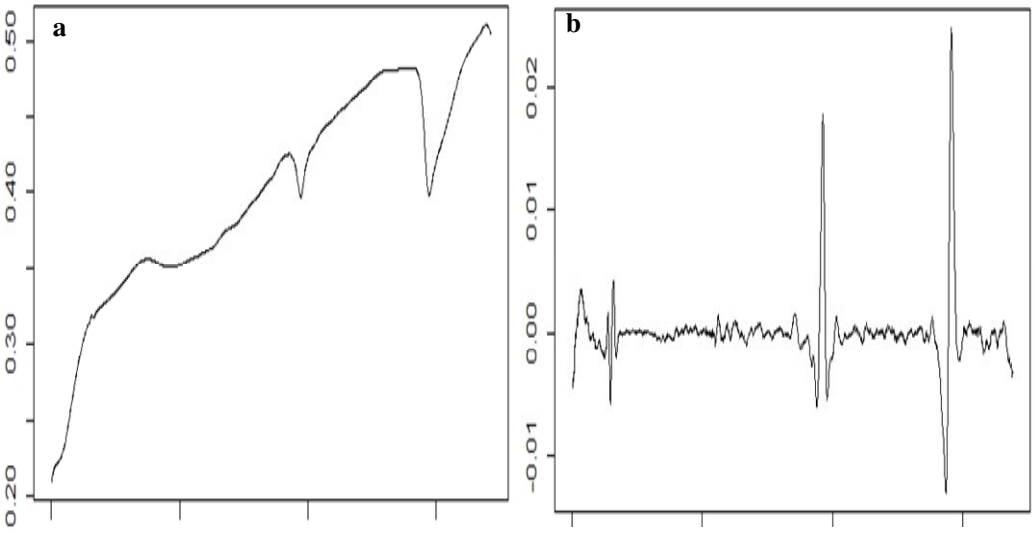
Yang X M, Xie H T, Drury C F, Reynolds W D, Yang J Y and Zhang X D (2012) Determination of organic carbon and nitrogen in particulate organic matter and particle size fractions of Brookston clay loam soil using infrared spectroscopy. *Eur J Soil Sci* **63**: 177–88.

Zhu H, Ying Zhao Y, Feng Nan F, Yonghong Duan Y and Rutian B R (2016) Relative influence of soil chemistry and topography on soil available micronutrients by structural equation modeling. *J Soil Sci Plant Nutr* **16**: 154-61.

ANNEXURE-I



Raw spectral data with noise (a, b) second order derivative smoothing (c, d)



(a) Noise removal of raw spectra (b) second order derivative smoothing

ANNEXURE-II

Summary of statistics for the spectral models developed by partial least square regression for calibration (train data set) of Sub-mountain Siwalik Hills

Soil property		Sub-mountain Siwalik Hills	
		RMSE	**R^2**
a	pH*	0.025	0.73
b	EC* (dS/m)	0.04	0.80
c	SOC (%)	0.06	0.75
d	N# (kg ha^{-1})	0.11	0.76
e	P* (kg ha^{-1})	0.26	0.76
f	K# (kg ha^{-1})	0.12	0.74
g	$CaCO_3$ *(%)	0.10	0.89
h	DTPA-Fe* (ppm)	0.22	0.82
i	DTPA-Mn* (ppm)	0.27	0.75
j	DTPA-Cu* (ppm)	0.13	0.78
k	DTPA-Zn* (ppm)	0.1	0.75
l	Sand* (%)	0.11	0.80
m	Silt* (%)	0.18	0.78
n	Clay* (%)	0.2	0.77

(*) Square root transformation
(#) Log transformation

ANNEXURE-III

Summary of statistics for the spectral models developed by partial least square regression for calibration (train data set) of Northeastern Undulating

Soil property		Northeastern Undulating (NEU)	
		RMSE	**R^2**
a	pH*	0.02	0.76
b	EC* (dS/m)	0.48	0.75
c	SOC (%)	0.05	0.81
d	N# (kg ha^{-1})	0.07	0.79
e	P* (kg ha^{-1})	0.45	0.77
f	K# (kg ha^{-1})	0.23	0.76
g	$CaCO_3$ *(%)	0.23	0.75
h	DTPA-Fe* (ppm)	0.21	0.77
i	DTPA-Mn* (ppm)	0.18	0.78
j	DTPA-Cu* (ppm)	0.16	0.87
k	DTPA-Zn* (ppm)	0.18	0.76
l	Sand* (%)	0.15	0.76
m	Silt* (%)	0.20	0.81
n	Clay* (%)	0.20	0.81

(*) Square root transformation
(#) Log transformation

ANNEXURE-IV

Summary of statistics for the spectral models developed by partial least square regression for calibration (train data set) of Piedmont and Alluvial Plain

Soil property		Piedmont and Alluvial Plain (PAP)	
		RMSE	**R^2**
a	pH*	0.53	0.79
b	EC *(dS/m)	0.53	0.79
c	SOC (%)	0.69	0.68
d	$N^{\#}$ (kg ha^{-1})	0.065	0.80
e	P* (kg ha^{-1})	0.35	0.78
f	$K^{\#}$ (kg ha^{-1})	0.12	0.70
g	$CaCO_3$ *(%)	0.92	0.43
h	DTPA-Fe* (ppm)	0.53	0.70
i	DTPA-Mn* (ppm)	0.067	0.78
j	DTPA-Cu* (ppm)	0.69	0.83
k	DTPA-Zn* (ppm)	0.16	0.77
l	Sand* (%)	0.10	0.80
m	Silt* (%)	0.12	0.81
n	Clay* (%)	0.15	0.77

(*) Square root transformation
(#) Log transformation

ANNEXURE-V

Summary of statistics for the spectral models developed by partial least square regression for calibration (train data set) of Central Alluvial Plain

Soil property		Central Alluvial Plain (CAP)	
		RMSE	**R^2**
a	pH*	0.15	0.78
b	EC *(dS/m)	0.045	0.83
c	SOC (%)	0.04	0.75
d	$N^{\#}$ (kg ha^{-1})	0.10	0.71
e	P* (kg ha^{-1})	0.16	0.83
f	$K^{\#}$ (kg ha^{-1})	0.078	0.76
g	$CaCO_3$ *(%)	0.13	0.81
h	DTPA-Fe* (ppm)	0.33	0.78
i	DTPA-Mn* (ppm)	0.24	0.70
j	DTPA-Cu* (ppm)	0.085	0.82
k	DTPA-Zn* (ppm)	0.16	0.77
l	Sand* (%)	0.28	0.81
m	Silt* (%)	0.17	0.79
n	Clay* (%)	0.10	0.75

(*) Square root transformation
(#) Log transformation

ANNEXURE-VI

Summary of statistics for the spectral models developed by partial least square regression for calibration (train data set) of South Western Alluvial Plain

Soil property		South Western Alluvial Plain (SWAP)	
		RMSE	**R^2**
a	pH*	0.22	0.79
b	EC *(dS/m)	0.10	0.83
c	SOC (%)	0.041	0.83
d	N# (kg ha^{-1})	0.099	0..81
e	P* (kg ha^{-1})	0.24	0.78
f	K# (kg ha^{-1})	0.46	0.86
g	$CaCO_3$ *(%)	0.12	0.75
h	DTPA-Fe* (ppm)	0.42	0.78
i	DTPA-Mn* (ppm)	0.27	0.76
j	DTPA-Cu* (ppm)	0.10	0.80
k	DTPA-Zn* (ppm)	0.23	0.77
l	Sand* (%)	0.11	0.77
m	Silt* (%)	0.18	0.78
n	Clay* (%)	0.23	0.80

(*) Square root transformation
(#) Log transformation

VITA

Name of the student	**:**	Saheed Garnaik
Father's name	**:**	Sh. Arun Kumar Garnaik
Mother's name	**:**	Smt. Pramila Pradhan
Nationality	**:**	Indian
Date of birth	**:**	3rd July, 1995
Permanent home address	**:**	At/PO-Balipatta, PS-Chhedipada, Dist-Angul, State-Odisha PIN-759124

EDUCATIONAL QUALIFICATION

Bachelor degree	**:**	B.Sc. Agriculture (Hons.)
University	**:**	Odisha University of Agriculture and Technology, Bhubaneswar, Odisha
Year of award	**:**	2017
OCPA	**:**	8.21/10.00
Master's degree	**:**	M.Sc. (Soil Science)
University	**:**	Punjab Agricultural University, Ludhiana
Year of award	**:**	2019
OCPA	**:**	7.84/10.00
Title of master's thesis	**:**	Development of spectral models for assesing soil fertiity status in various agroecological sub-regions of Punjab
Awards/Distinction/ Scholarship/Fellowship	**:**	University Merit Scholarship during B.Sc. and National Talent Scholarship during M.Sc.