ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

EVALUATING PERFORMANCE OF DIFFERENT REMOTE SENSING TECHNIQUES AND VARIOUS INTERPOLATION APPROACHES FOR SOIL SALINITY ASSESSMENT

Ph.D. THESIS

Taha GORJI

Department of Applied Informatics

Geographical Information Technologies Programme

AUGUST 2021

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

EVALUATING PERFORMANCE OF DIFFERENT REMOTE SENSING TECHNIQUES AND VARIOUS INTERPOLATION APPROACHES FOR SOIL SALINITY ASSESSMENT

Ph.D. THESIS

Taha GORJI
(706162003)

Department of Applied Informatics

Geographical Information Technologies Programme

Thesis Advisor: Prof. Dr. Ayşe Gül TANIK

AUGUST 2021

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

TOPRAK TUZLULUĞU DEĞERLENDİRMESİ İÇİN FARKLI UZAKTAN ALGILAMA TEKNİKLERİ VE ÇEŞİTLİ INTERPOLASYON YAKLAŞIMLARININ PERFORMANSININ DEĞERLENDİRİLMESİ

DOKTORA TEZİ

Taha GORJI
(706162003)

Bilişim Uygulamaları Anabilim Dalı

Coğrafi Bilgi Teknolojileri Programı

Tez Danışmanı: Prof. Dr. Ayşe Gül TANIK

AĞUSTOS 2021

Taha Gorji, a Ph.D. student of ITU Graduate School student ID 706162003, successfully defended the thesis entitled "EVALUATING PERFORMANCE OF DIFFERENT REMOTE SENSING TECHNIQUES AND VARIOUS INTERPOLATION APPROACHES FOR SOIL SALINITY ASSESSMENT"
which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Ayşe Gül TANIK** Istanbul Technical University	.....
Jury Members :	**Prof. Dr. Elif SERTEL** Istanbul Technical University	.....
	Prof. Dr. Füsun BALIK ŞANLI Yıldız Technical University	.....
	Prof. Dr. Şinasi KAYA Istanbul Technical University	.....
	Prof. Dr. Bülent BAYRAM Yıldız Technical University	.....

Date of Submission: 7 June 2021
Date of Defense: 9 August 2021

To my family,

FOREWORD

I would like to express my appreciation to all contributors whose cooperation and assistance helped me throughout my Ph.D. Study. First, I would like to thank to my supervisor Prof. Dr. Ayşe Gül TANIK who really gave me a lot of helpful advice during my study. Besides, special thanks to Prof. Dr. Elif SERTEL for both guiding me immensely to complete this research and for providing me scholarship to support my education costs. In addition, I am so thankful to Prof. Dr. Füsun BALIK ŞANLI for providing relevant suggestions, which lead to better progress in my thesis. I would like to thank especially to Dr. Nikou HAMZEHPOUR who provided EC ground data belonging to west playas of Urmia Lake and Bonab Region and she supported me by giving detailed information on salinity problems of the regions. Thanks go to all Professors for their help inspiring me to convey the research and for their assistance during my academic studies. I am honored to have a chance to work alongside with knowledgeable and highly experienced professors like them.
Finally, I would like to express my sincere thanks to my family for their kind encouragement to complete my Ph.D. study.

June 2021

Taha GORJI
GIS Specialist

TABLE OF CONTENTS

Page

ABBREVIATIONS

AC : Atmospheric Corrections
ASTER : Advanced Space borne Thermal Emission and Reflection Radiome
ANN : Artificial Neural Network
BI : Brightness Index
DTA : Decision-tree Analysis
DNs : Digital Numbers
EC : Electrical Conductivity
ESP : Exchangeable Sodium Percentage
ETM : Enhanced Thematic Mapper
EU : European Union
EVI : Enhanced Vegetation Index
FAO : Food and Agriculture Organization
GIS : Geographical Information System
LSU : Linear Spectral Unmixing
MARS : Multivariate Adaptive Regression Splines
MODIS : Moderate Resolution Imaging Spectroradiometer
NDSI : Normalized Differential Salinity Index
NDVI : Normalized Difference Vegetation Index
NIR : Near-infrared
NASA : National Aeronautics and Space Administration
OLI : Operational Land Imager
PCA : Principal Component Analysis
PLSR : Partial Least Square Regression
RS : Remote Sensing
SI : Salinity Index
SWIR : Shortwave Infrared
TIRS : Thermal Infrared Sensor
TM : Thematic Mapper
UTM : Universal Transverse Mercator
WGS : World Geodetic System

LIST OF TABLES

Page

LIST OF FIGURES

Page

EVALUATING PERFORMANCE OF DIFFERENT REMOTE SENSING TECHNIQUES AND VARIOUS INTERPOLATION APPROACHES FOR SOIL SALINITY ASSESSMENT

SUMMARY

Soil salinization is one of the drastic environmental phenomena due to its adverse effects on land productivity, plant growth and sustainable development especially in arid and semi-arid regions of the world. As population is growing fast, the demand for supplying food is increasing; despite, plenty of arable land is abandoned due to primary and secondary soil salinization. Among primary sources of soil salinization, natural factors such as existence of parent material in soil structure, closeness of salty groundwater table to surface, weathering of the parent rock and sea water intrusion intensify soil salinity occurrence. In terms of secondary sources of soil salinization, irrigating agricultural land with water rich in salt, land clearing and using fertilizer containing nitrogen and potassium salts exacerbate salt accumulation in soil can be addressed. In many nations flood irrigating agriculture and lack of relevant drainage systems has caused environmental disturbances such as waterlogging, salinization, and depletion and pollution of water supplies and as a result it increased concern about the sustainability of irrigated agriculture. Indeed, exerting traditional irrigation approaches leads to acceleration of salt accumulation and water logging in soil. Accordingly, it is essential to monitor soil salinity on local, regional and global scale to track spatial variation of salt-affected soils particularly in places, which are more prone to soil salinization. Spatio-temporal soil salinity mapping is remarkably significant to support management strategies for soil related applications. Knowledge of spatiotemporal variation and probability of reoccurrence of salt-affected lands is critical to our understanding of land degradation and for planning effective remediation strategies in face of future climatic uncertainties. However, traditional approaches used for tracking the temporal and spatial distribution of soil salinity/sodicity are extensively localized, making estimations on a global scale very tough. projecting more soil salinity detecting and mapping along with monitoring spatial and temporal variation of salt-affected lands is necessary for taking relevant and prompt decisions to enhance the management practices and provide solutions to overcome or diminish soil salinity issues. Every soil salinity assessment requires two fundamental steps. Initially, it is essential to detect the areas where salts are accumulated and concentrated in the soil profile. In the next step, seasonal monitoring of the temporal and spatial alteration of salt-affected lands is required. In order to estimate periodical changes of soil salinity in large scale regions, it is essential to utilize rapid, fast and economical approaches. In that sense, Remote Sensing (RS) technologies, machine-learning algorithms and Geographical Information Systems (GIS) provide cost-effective, non-destructive, qualitative and quantitative spatio-temporal information on soil salinity changes.

This research aimed to evaluate the performance of various RS techniques and interpolation methods for soil salinity mapping in three different geographical

locations. Moreover, as a novelty to the study, a new soil salinity index was derived from visible and NIR bands, and it was applied for soil salinity mapping in all the three selected locations suffering from salinity. Capability of this new index was firstly compared with two other commonly used salinity indices independently. Then, it was adopted in combination with other indices as an input variable in Cubist model.

West and southeast playas (Bonab Region) of Urmia Lake were the two selected places in Iran, and Tuz Lake in Turkeywas the other case study area selected for completing the analysis of this research and especially for testing the performance of the new index.

RS algorithms, GIS techniques, modelling and machine-learning methods highly contributed to generating various salinity maps despite limited knowledge and information about field measurement data. In this research, for each case study, different soil salinity maps with six classes including none-saline with Electrical Conductivity (EC) value of 0-2 dS/m, slightly saline with EC value of 2-4 dS/m, moderately saline with EC value of 4-8 dS/m, highly saline with EC value of 8-16 dS/m were used together with two new classes representing extremely saline soils with EC value of 16-32 dS/m and above 32 dS/m were produced within this study for evaluating the case study areas in different years via various methodologies.

This study initially focused on preparing relevant raw data including ground EC data and their corresponding visible and NIR band pixel values for each of the three case studies separately. Then, several arithmetic operations of bands by using trial and error checking has been tested for determining the best combination which could differentiate extremely saline soil from none saline soils. In the second phase of this study, the generated soil salinity index was utilized for producing soil salinity maps in each of the geographical locations. In addition, after applying several soil salinity indices, two commonly used salinity indices has been selected and applied for all the three case studies to compare their soil salinity maps with the maps produced from the new soil salinity index. Regression analysis results indicated that soil salinity maps generated by new SI demonstrate acceptable results with model R^2 values similar to model R^2 values of other indices in all the three case studies. In addition different combination of SI images derived from Landsat-8 OLI were adopted as input variable in Cubist model; after running the model in each of three studies, Cubist selected new SI as the main parameter for defining the criteria of the rules.

In parallel to technical analysis of soil salinity mapping on the mentioned case studies, application of RS data and several algorithms to assess soil salinity in different case studies were reviewed. As a result, relevant information on soil salinity detection including novel soil salinity mapping methods, sensing techniques, RS data and main causes of soil salinity for each case study were achieved and summarized in the form of a database. In this review, sensing approaches were classified based on obtaining information methods on land surfaces including airborne photogrammetry, satellite images, ground measurements and laboratory analysis.
Overview of studies depicts that soil salinity mapping is mostly conducted by utilizing multispectral RS data in combination with simultaneous field measured data. According to the literature review that was completed in this survey, we reach to this fact that selecting an index derived from multispectral RS data for a case study

is significantly dependent to the characteristics of the study area and there is no salinity index that demonstrate best results for all geographical locations.
Inspecting various case studies indicate that both primary and secondary salinization can be contemplated as sources of soil salinity. Despite, exploring studies which are performed in arid and semi-arid regions of the globe depicts that anthropological factors not only exacerbate soil salinization specifically in agricultural lands; but, also the adverse effects of human-induced activities has worsen natural causes of soil salinization.

TOPRAK TUZLULUĞU DEĞERLENDİRMESİ İÇİN FARKLI UZAKTAN ALGILAMA TEKNİKLERİ VE ÇEŞİTLİ İNTERPOLASYON YAKLAŞIMLARININ PERFORMANSININ DEĞERLENDİRİLMESİ

ÖZET

Özellikle dünyanın kurak ve yarı-kurak bölgelerinde; arazi verimliliği, üretim, bitki yetişmesi ve sürdürülebilir büyüme üzerinde yarattığı olumsuz etkiler neticesinde, toprak tuzluluğu en radikal çevre olaylarının başında gelmektedir. Ülkelerin nüfusunun artmasına bağlı olarak besin kaynaklarına olan talebin büyümesine rağmen, birçok ekilebilir ve dikilebilir arazi birinci ve ikinci derece toprak tuzluluğuna terk edilmiş durumdadır. Toprak tuzluluğunun ana sebepleri arasında; toprak yapısında ana madde bulunması, yataktan yüzeye kadar tuzlu yeraltı sularına yakınlık, ana maddenin bulunduğu kayanın aşınması ve deniz suyu girişimi toprak tuzluluğunu artıran en önemli doğal sebeplerdir. Toprak tuzluluğunu şiddetlendiren ikincil nedenler arasında ise; tarım arazilerinin tuz bakımından zengin sularla sulanması, arazi temizlenmesi ve azot ve potasyum tuzları barındıran suni gübre kullanımı gibi sebepler sayılabilmektedir.

Birçok ülkede geleneksel salma sulama düzenindeki tarım ve ilgili drenaj sistemlerinin eksikliği aşırı su basması, toprağın tuzlanma ve su kaynaklarının tükenmesi ve kirlenmesi gibi çevresel problemlere neden olmuştur. Sonuç olarak bu durum sulu tarımın sürdürülebilirliği konusundaki endişeleri de artmıştır. Aslında, geleneksel sulama sistemlerinde ısrar edilmesi toprakta tuzun birikmesine sebep olduğu gibi toprağın da su emmesine yol açmaktadır. Buna bağlı olarak, özellikle toprağın tuzlanmaya elverişli olduğu yerlerde; yerel, bölgesel ve küresel ölçekte tuz tarafından etkilenen toprakların tuzluluk oranlarının takip edilmesi ve arazi üzerinde gösterdiği değişik tuzluluk oranlarının takip edilmesi vazgeçilmez bir unsurdur. Yer ve zaman bazlı toprak tuzluluk haritalaması, toprağa bağlı olan uygulamaların yönetiminin desteklenmesi konusunda mühimdir.

Mekansal-zamansal değişim (varyasyon) bilgisi ve tuzdan etkilenen arazilerin tekrar oluşma olasılığı, arazi bozulmasını anlamak ve gelecekteki iklimsel belirsizlikler karşısında etkili iyileştirme stratejileri planlamamız için kritik öneme sahiptir. Bununla birlikte, toprak tuzluluğunun/sodikliğinin zamansal ve mekansal dağılımını izlemek için kullanılan geleneksel yaklaşımlar geniş ölçüde yerelleştirilmiştir ve bu da küresel ölçekte tahminleri çok zor hale getirmektedir. Tuzdan etkilenen arazilerin mekansal ve zamansal değişimlerinin izlenmesi ile birlikte daha fazla toprak tuzluluğu tespit ve haritalama projelendirmesi, yönetim uygulamalarını geliştirmek ve toprak tuzluluğu sorunlarının üstesinden gelmek veya azaltmak için çözümler sağlamakla ilgili hızlı kararlar almak için gereklidir. Toprak tuzluluğu değerlendirmesi iki temel adım gerektirmektedir. Öncelikle toprak profilinde tuzların biriktiği ve yoğunlaştığı alanların tespit edilmesi esastır. Bir sonraki adımda, tuzdan etkilenen arazilerin zamansal ve mekansal değişiminin mevsimsel olarak izlenmesi gerekmektedir. Büyük ölçekli bölgelerde toprak tuzluluğunun dönemsel

değişimlerini tahmin etmek için hızlı ve ekonomik yaklaşımların kullanılması esastır. Bu bağlamda, Uzaktan Algılama (UA) teknikleri, yapay zekâ algoritmaları ve Coğrafi Bilgi Sistemleri (CBS) toprak tuzluluğu değişikliklerinin ekonomik, hızlı, nitelik ve nicelik itibariyle zengin, zaman ve yer bakımından bilgileri sağlamakla yükümlüdür.

Toprak tuzluluğunun topraktan numuneler alınarak, laboratuvarlarda elektriksel iletkenlik deneyleri ile saptanmasının yanı sıra özellikle son yıllarda topraktaki tuzluluğu belirlemek, izlemek ve haritalandırmak amacıyla, uydu görüntülerinden de yararlanılmaktadır. Özellikle cokluspektral sensörler bu amaca hizmet etmektedir. Bu modern teknolojik araçlar arasında, Landsat Tematik Haritalama (TM), Landsat Cokluspectral Tarama Sistemi (MSS), Landsat7, Landsat8, Geliştirilmiş Landsat Tematik Haritalama (ETM), SPOT, Gelişmiş Uydu Bağlantılı Termal Emisyon ve Yansıyan Görüntü Radiome (ASTER), IKONOS, MODIS ve Hindistan Uzaktan Algılama sistemi (IRS) sayılabilir.

Bu araştırma, toprak tuzluluğunun haritalanması konusunda kullanılan çeşitli uzaktan algılama (UA) tekniklerinin ve interpolasyon yaklaşımlarının performansını, üç değişik coğrafi konumda değerlendirmektedir. Ek olarak; bu çalışmaya bir yenilik getirmesi amacıyla, görünür ve yakın infrared (NIR) bantlardan elde edilen yeni bir toprak tuzluluk indeksi geliştirilmiştir. Geliştirilen bu indeks; tuzluluktan mustarip olan seçilmiş bu üç bölgede toprak tuzluluk oranının haritalanması bakımından kullanılmıştır. İlk olarak, geliştirilmiş olan yeni indeksin kabiliyetleri yaygın kullanılan kullanılan diğer iki tuzluluk indeksiyle bağımsız olarak karşılaştırılmıştır. Ardından, Cubist modelinde bir girdi değişkeni olarak diğer endekslerle birlikte çalıştırılmıştır.

Yeni indeksin performansının test edilmesi ve bu araştırmanın tamamlanması için seçilen yerlerin ikisi İran'da bulunan Urmiye Gölü'nün batı ve güneydoğu sahilleridir (Bonab Bölgesi). Türkiye'den Tuz Gölü bu araştırmanın diğer bölümünü oluşturmaktadır.

Her ne kadar arazi ölçümleri konusunda bilgi ve imkânlar kısıtlı olsa da; uzaktan algılama algoritmaları, coğrafi bilgi sistemleri ve yapay zekâ yaklaşımları çeşitli tuzluluk haritalarının oluşturulmasına yardımcı olmuştur. Bu çalışmada yer alan her vaka çalışmasında, birbirinden değişik altı kademeli toprak tuzluluğu haritalanmıştır. İlk kademede tuzlu olmayan ve iletkenlik değeri 0-2 dS/m arasında değişen toprak bulunmaktadır. İkinci kademede ise az tuzlu toprak, yani iletkenlik değeri 2-4 dS/m toprak bulunmaktadır. Üçüncü kademe ise orta derecede tuzlu toprak, yani iletkenlik değeri 4-8 dS/m toprak bulunmaktadır. Dördüncü kademe ise yüksek derecede tuzlu toprak bulunup, iletkenlik değeri 8-16 dS/m arasında bulunmaktadır. Bu dört kademeye ek olarak, vaka çalışmasına konu edilen bölgelerin farklı yıllarda ve farklı yaklaşımlarda değerlendirilmesi adına, aşırı derecede tuzlu olan topraklar için iki yeni kademe eklenmiştir. Bu yeni kademelerin iletkenlik oranları 16-32 dS/m ve 32 dS/m üzeri olarak belirtilmiştir.

Bu çalışma ilk bakışta, her 3 uygulama çalışmasının her biri için ayrı ayrı olmak üzere, toprak iletkenlik değerleri ve bu değerlerin NIR bandı piksel değerleri de dâhil olmak üzere, ham verilerin hazırlanması üzerine yoğunlaşmıştır. Aşırı derecede tuzlu toprakların tuzsuz topraklardan ayrıştırılması için en iyi kombinasyonların bulunması

adına, birçok bant aritmetik çalışması deneme-yanılma yöntemiyle gerçekleştirilmiştir. Bu çalışmanın ikinci aşamasında ise, geliştirilen toprak tuzluluk indeksi her bir coğrafi bölgenin tuzluluk haritasının çıkarılmasında kullanılmıştır. Buna ek olarak, birçok toprak tuzluluk indeksinin tatbik edilmesi sonrasında, yoğunlukla kullanılan iki indeks seçilmiştir. Seçilen bu iki indeksin uygulanmasıyla ortaya çıkan toprak tuzluluk haritalarının, yeni geliştirilmiş toprak tuzluluk indeksiyle oluşturulan tuzluluk haritaları ile karşılaştırılması adına, seçilen bu iki indeks her üç seçili örnek alanlarda da uygulanmıştır. Yapılan regresyon analizi sonucuna göre, her üç çalışmada da yeni toprak tuzluluk indeksine göre hazırlanmış haritalar, diğer indekslerin ortaya koyduğu R^2 değerlerine benzer ve kabul edilebilir R^2 değerleri ortaya koymuştur. Daha da ileri gitmek gerekirse; Landsat-8 OLI'den gelen farklı toprak tuzluluk görüntüleri kombinasyonları Cubist modelde giriş değişkeni olarak kullanılmak üzere kabul edilmiştir. Modelin her üç örnek uygulama alanlarında çalıştırılması sonucunda, Cubist; kriterlerin tanımlanması çerçevesinde, yeni toprak tuzluluk indeksi ana parametre olarak seçilmiştir.

Yukarıda bahsi geçen çalışmalardaki toprak tuzluluk haritalamasının teknik analizlerine paralel olarak, toprak tuzluluğunun her uygulama alanında değerlendirilmesi adına, UA verileri ve birçok algoritmaların uygulaması tekrar gözden geçirilmiştir. Sonuç olarak, yeni toprak tuzluluk haritalama yöntemleri, algılama teknikleri, UA verileri ve her bir vaka çalışması için toprak tuzluluğunun ana nedenleri gibi toprak tuzluluk tespiti ile ilgili bilgiler elde edilmiş ve veri tabanı şeklinde özetlenmiştir. Bu incelemede, hava fotogrametrisi, uydu görüntüsü, yersel ölçümler ve laboratuvar analizi gibi kara yüzeyine ilişkin bilgi edinme yöntemleri baz alınarak, algılama yaklaşımları sınıflandırılmıştır.

Süregelen çalışmalar, toprak tuzluluk haritalamasının çoğunlukla eş zamanlı sahada ölçülen verilerle birlikte multispektral RS verileri kullanılarak gerçekleştirildiğini göstermektedir. Tez kapsamında tamamlanan literatür taramasına göre, belirlenmiş bir alanda multispektral RS verilerinden türetilen bir indeks seçilmesinin çalışma alanının özelliklerine önemli ölçüde bağlı olduğu ve tüm coğrafi konumlar için en iyi sonuçları gösteren bir tuzluluk indeksi olmadığı gerçeğine ulaşıyoruz. Çeşitli uygulama çalışmalarının incelenmesi, hem birincil hem de ikincil tuzlanmanın toprak tuzluluğu kaynakları olarak düşünülebileceğini göstermektedir. Buna rağmen, dünyanın kurak ve yarı kurak bölgelerinde yapılan keşif çalışmaları, antropolojik faktörlerin sadece özellikle tarım arazilerinde toprak tuzlanmasını şiddetlendirmekle kalmayıp; aynı zamanda insan kaynaklı faaliyetlerin olumsuz etkileri de toprak tuzlanmasının doğal nedenlerini kötüleştirmiştir.

1. INTRODUCTION

Spatio-temporal monitoring of soil richness is becoming highly important for especially arid and semi-arid countries and regions of the world as a result of rise in concerns of cultivating agricultural products for rapidly increasing population (Eisele et al., 2015). Soil salinization is one of the major environmental issues accelerating land degradation processes, which in turn, deteriorate soil fertility and negatively impacts agricultural production. Indeed, agricultural lands containing rich and fertile soil as main sources of worldwide food production are under pressure resulting from both natural and human-induced soil salinization (Gorji et al., 2017). Accumulation of salt minerals in soil results in loss of structure, soil dispersion, soil compaction and crust formation (Metternicht and Zinck, 2003).

Soil salinity also negatively affects stability of soil. Aeration, biological activity and the growth of crops, movement and storage of water in soil profile and erosion are indeed among the parameters that are directly dependent on cohesion of the soil aggregates that lead to excess amount of salt accumulation in the soil profile. In turn, it decreases the capability of soil to retain the heterogeneous arrangement of solid and void space. Figure 1.1. defines the form and stability properties of soil structure.

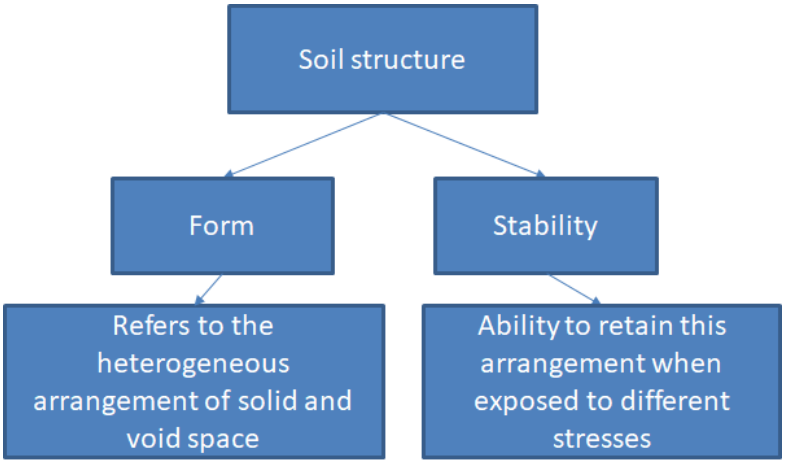

Figure 1.1 : Form and stability of soil structure.

Gorji et al. (2015) conferred that old-fashioned/conventional irrigation approaches exacerbate soil salinization and degrade soil fertility particularly when there is lack of and/or poor drainage system in agricultural lands. Mapping and monitoring soil salinity alteration is an essential issue for predicting natural disasters like desertification and for diminishing severe economic and social consequences in especially arid and semi-arid regions of the world. Because of high spatial and temporal changes of soil salinity, information on both the spatial and temporal changes of soil salinity is required to mitigate with the negative impacts of this significant phenomenon on biomass production (Song et al., 2016). So, timely monitoring and mapping of soil salinity have become remarkably important in order to overcome this issue (Gorji et al., 2017). Frequent and dynamic monitoring of soil salinization is necessary to generate quantitative spatial and temporal data for land reclamation as referred by Peng et al. (2019). Within the last two decades, remote sensing (RS), geographical information systems (GIS), modelling and machine-learning techniques have outperformed the traditional methods of soil salinity mapping. Detection of salt-affected lands has enhanced from qualitative to quantitative mapping due to large area coverage, multiple spectral information and approximately constant observation via RS systems. Salt-affected lands can be detected from remotely sensed data either directly on exposed soils or indirectly through vegetation and plants. In the regions containing densely vegetated soils, using vegetation indices for soil salinity mapping leads to obtaining much promising results, whereas applying soil salinity indices can be considered as a relevant approach in the case of bare lands or soils with low scattered vegetation. Various soil salinity indices derived from simple or complicated spectral band ratio combinations of RS data have been utilized to detect spatio-temporal variation of salt-affected lands in several case studies (Elhag and Bahrawi, 2017). Different arithmetic combination visible and near infrared (NIR) bands generated promising soil salinity indices in most studies (Liu et al., 2018). Normally, extremely saline soils indicate relatively higher spectral reflectance in the visible and near-infrared regions of the spectrum than non-saline soils. Sentinel, SPOT, IKONOS, ASTER, Landsat series, IRS and MODIS are among the commonly used multispectral sensors that provide possibilities for temporal, fast and economic monitoring and mapping of soil salinity since 1990s (Allbed and Kumar, 2013).

1.1 Objective of the Thesis

There were four main goals for conducting this research:

- Generating a new soil salinity index by trial and error and through examining various arithmetic combination of visible and NIR bands derived from Landsat-8 OLI
- Comparing performance of the new soil salinity index with the commonly known two other salinity indices by utilizing Cubist model.
- Combining geo-statistical methods with RS data and Cubist model for soil salinity mapping.
- Reviewing application of RS data and several algorithms to assess soil salinity mapping approaches in different case studies.

2. LITERATURE REVIEW

In this section, firstly major concepts about soil salinity are discussed. The basic definition of soil salinity, primary and secondary soil salinization sources, negative impacts of soil salinization in the environment, and spectral behavior of saline soils are mentioned. In addition, information on global importance of soil salinity and necessity of tracking its changes are given. In the next phase, application of RS data and several algorithms to assess soil salinity mapping approaches in different case studies are summarized in the form of a database. Studies are sorted according to their survey data from the most recent ones in the year 2021 to the oldest one which were conducted in year 2014. Chronological listing of studies presents technological developments of soil salinity mapping approaches over time.

2.1 Major Concepts about Soil Salinity

Soil salinity is an extensive environmental issue in those geographical areas of the world that have higher evaporation rate and lower rainfall in comparison with other regions.

Inspecting Figure 2.1 indicates that soil salinity is a major socio-economical threat to nations locating in arid and semi-arid regions of the world. The map categorized salt-affected lands based on their likelihood of containing EC values above 4 dS/m between years 1980 and 2018. Here, the likelihood is a dimensionless index with a range of 0 to 1 that is measured by dividing number of years that a specific place face EC value above 4dS/m to the total number of observation years in the same geographical location.

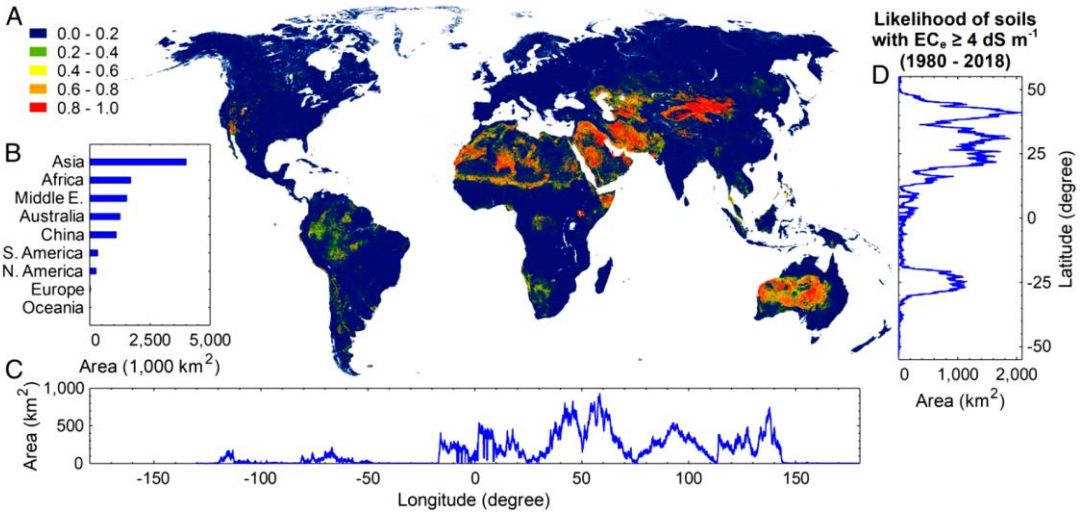

Figure 2.1 : Global distribution of salt-affected soils (Hassani et al., 2020).

The map illustrates more than 4000 km^2 of lands in Asia, around 2000 km^2 of lands in Africa and nearly 1500 km^2 of lands in the Middle East have been faced with soil salinity problem with likelihood index above 0.8.

As it is shown in Figure 2.2, all continents in the world are negatively impacted by soil salinization. Ivushkin et al. (2019) discussed that approximately 1 billion hectares of land is adversely affected by soil salinization and there is rising trend in more than 100 countries in various continents. If relevant decisions, measures and management techniques such as monitoring soil salinity via RS, integrating agroforestry, applying bio-fertilizers, leaching maintenance, tillage practices and setting appropriate drainage systems would not be take into place rapidly, condition would become even more complex and much more fertile land will be salinized (Machado and Serralheiro, 2017).

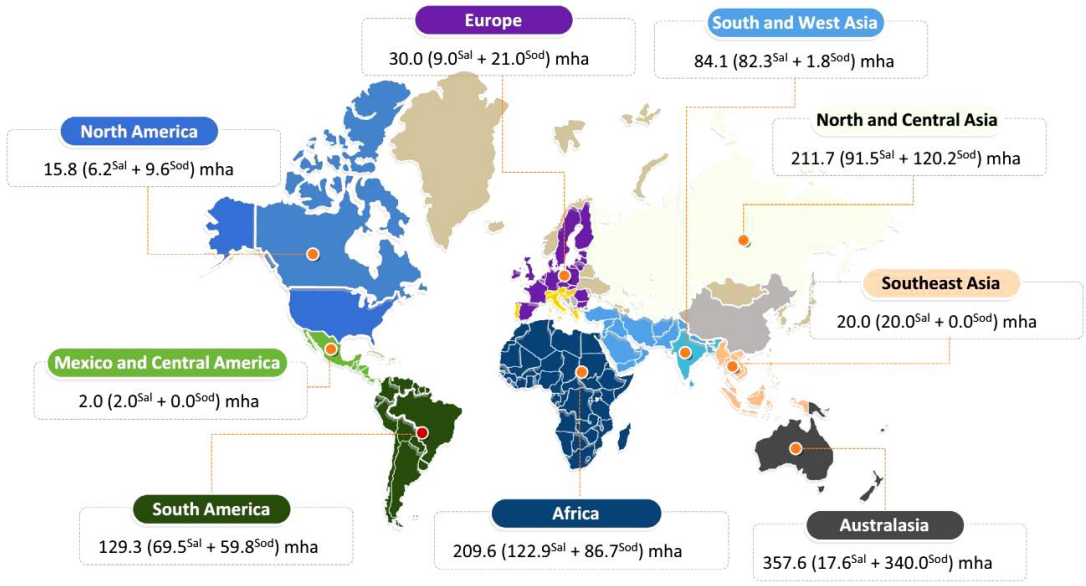

Figure 2.2 : Total area of lands affected by salinity and sodicity in various continents (Sahab et al., 2020).

Presence of high amount of soluble salts (cations and anions) in the soil profile either emerged from a natural source or caused by a human-induced factor leads to soil salinization. Cations including (Na^+), potassium (K^+), magnesium (Mg^{2+}), and calcium (Ca^{2+}), besides anions such as chloride (Cl^-), sulfate (SO_4^{2-}), and carbonate in the form of bicarbonate (HCO_3^-) are the most common elements found in water and soil. Electrical conductivity (EC) with the unit of dS/m is an index that is commonly used for measuring soil salinity. Most of the plants can survive in the soil including EC range of 0-4 dS/m. Despite, many crops lose their normal yield at highly saline soils with EC values ranging from 8 dS/m to 16 dS/m, and even almost disappear in extremely saline soils with EC value above 16 dS/m. In addition, moderately saline soils with ECs between 4 dS/m to 8 dS/m bring yield decline for various crops (Munns, 2005).

There are two main sources of soil salinization namely primary (natural) and secondary (human-induced) sources. Primary minerals and stones with excess amount of cations and anions are major natural sources of salt. Indeed, wind and water are the main sources of weathering parent materials and transforming the salt into the soil profile. In the dry season, high temperature leads to rapid evaporation of water from soil, and as a result, remaining salt will be accumulated on the surface and in subsurface soil. In arid regions due to lack of precipitation, there is no sufficient water to leach down the accumulated salts, and this condition exacerbate arable land soil

salinization. Moreover, topography and geographical location of lands can be counted as other natural factors for soil salinization. Topography impacts on closeness of salty groundwater table to land surface. In addition, arable lands that are located in the vicinity of saline lakes are more vulnerable to soil salinization as a consequence of saline water intrusion and floods.

Secondary soil salinization appears due to man-made interventions. Among 230 million ha of irrigated land, nearly 45 million ha are salt-affected soils due to secondary salinization (Koohafkan, 2008). Traditional irrigation techniques (flood irrigation with salty water), old cultivation methods and wrong agricultural practices combined with irrelevant decisions on the sequences of crop cultivation lead to soil salinization in arable lands. Failure in washing out salts from soil profile and plant root zone due to inadequate drainage systems is known as another problem which exacerbate salt accumulation. Deforestation, intrusion of saline and polluted industrial and domestic wastewater into the soil profile, overgrazing and urbanizing arable lands increase the probability of groundwater salinization which further leads to surface soil salinization (Shrestha and Farshad, 2009). Additionally, excess usage of chemical fertilizers for agriculture is another human-induced source of salinization. Salt-affected agricultural lands have lower productivity for growing crops and plants since soluble salt in the soil profile prohibits majority of plants seed germination and diminish osmotic potential of soil water which results in inadequacy of plants to uptake water from the root zone. Despite, different plants and crops has shown distinctive sensitivity and tolerance to each salinity class (Bahtt et al., 2008). Indeed, soil salinity classes are categorized based on effects of saline soil on plant growth. It is tested that none saline class with EC range of 0 to 2 dS/m salinity effects are negligible for all plant types. On the contrary, in extremely saline soil with the EC range above 16 dS/m only some highly tolerant crops can survive. In Spain, scientists compared ranges of tolerance to EC data for 16 species along the River Guadiamar. The study depicted that halophilic plants; namely, *Sarcocornia fruticosa*, *Scirpus maritimus compactus* and *Arthrocnemum macrostachyum* survived in extremely saline soil with EC value of 55 dS/m in some cases as depicted by Mercado et al. (2012). Table 2.1 demonstrates average, minimum and maximum EC values that some halophilic plants could survive in various environmental conditions.

Table 2.1 : Minimum, maximum and average EC values that some halophilic species can tolerate (Mercado et al., 2012).

Plant Name	Minimum EC (dS/m)	Average EC (dS/m	Maximum EC (dS/m)
Scirpus maritimus compactus	5	14	55
Sarcocornia fruticosa	5	25	55
Arthrocnemum macrostachyum	5	26	55
Phragmites australis	8	16	38
Suaeda vera	5	16	28
Spartina densiflora	9	17	38

Halophyte trees are more salt-tolerant because of their high capability of ion transport system potential and osmotic adjustment which is totally different from glycophytes. Planting halophyte trees supports rehabilitation of salt-affected lands. Besides, these trees have functions in recycling saline agricultural wastewater and they can be used as grains for animal feeding systems (Glenn et al., 2016).

In terms of common crops that are harvested seasonally and support agricultural food production, sensitivity to EC value is much higher than halophilic species. Information on effects of various ranges of EC on different crops and their corresponding yield loss are given in Table 2.2. As an example, crop yield can be decreased by 50% for onion when EC is 4.3 dS/m.

Table 2.2 : Crop yield loss in various EC values (Horneck et al., 2007).

Crop	Expected yield reduction (%)			
	None	10%	25%	50%
	(EC), dS/m			
Barley	8.0	10.0	13.0	18.0
Wheat	6.0	7.4	9.5	13.0
Sugar beet	4.0	4.1	6.8	9.6
Alfalfa	2.0	3.4	5.4	8.8
Potato	1.7	2.5	3.8	5.9
Corn(grain)	1.7	2.5	3.8	5.9
Onion	1.2	1.8	2.8	4.3
Beans	1.0	1.5	2.3	3.6
Apples, Pears	1.7	2.3	3.3	4.8
Strawberries	1.0	1.3	1.8	2.5
Sudan grass	2.8	5.1	8.6	14.0
Grapes	1.5	2.5	4.1	6.7
Broccoli	2.8	3.9	5.5	8.2
Cucumbers	2.5	3.3	4.4	6.3

Reducing drinking and irrigation water quality is another negative effect of soil salinity which results in economic, social and environmental issues. It has also adverse effect on ecological health of rivers and it threats biodiversity by loss of different habitats. Increasing flood risk is another threat of soil salinization especially in arid regions of the world. Since salts in the soil profile prohibit penetration of surface water into the soil, it accelerates rate of surface run-off and also washes valuable soil minerals and nutrients that leads to more soil degradation.

It is expected that soil salinization will be expanded and will adversely affect much more arable lands with forthcoming climate change scenarios such as increasing temperature, sea level rise, and impact on coastal regions that will consequently accelerate evaporation, depletion of soil moisture, and rise of soil salinization.

Thus, projecting more soil salinity mapping and monitoring along with tracking changes of salt-affected lands is necessary for taking relevant and prompt decisions to alter the management approaches and provide solutions to overcome or lessen soil salinity problems. There are two main steps for soil salinity monitoring. Firstly, it is necessary to detect the places where salts are accumulated and concentrated in the soil profile. In the next phase, tracking of the temporal and spatial alteration of salt-affected lands is essential. In order to estimate periodical changes of soil salinity in large scale regions, rapid, fast and economical approaches are required.

Polous et al. (2011) discussed soil salinization phenomena and referred that it is widely extended and dispersed in almost all continents even Antarctica. Soil salinization is increasing in many nations especially in arid regions such as Iraq,

Egypt, India, Pakistan, Iran, Argentina and Syria. EC values differ in various regions. In fact, there are three main factors including local climatic conditions, environmental characteristics, and agricultural management practices that affect the severity of soil salinization.

Soil salinity mapping by applying only old approaches such as laboratory analyses and field surveys is not possible since dispersion rate of salt-affected lands in large scale areas are accelerating around the world. Contrarily, machine-learning approaches, RS techniques, GIS analysis tools, and modeling can be applied for continuous tracking of the progress of this phenomenon by their fast computing and predicting capabilities (Lhissou et al., 2014). Coordinating RS data with field measured EC data can be entirely functional for accurate soil salinity mapping. Indeed, measurement is a complement parameter for RS data to apply different predicting models and analysis.

Knowledge about spectral behavior of saline soil support soil salinity detection by applying optical RS. Salt crusts can be detected by utilizing visible and NIR bands of spectrum. Despite, direct sensing of salt-affected lands is functional only when there is accumulated salt on surface soil. The brightness and number of sensing bands decrease, while salt concentration in surface soil reduces (Farifteh et al., 2008). Direct sensing of slightly or moderately saline soils is not applicable due to this fact that other soil minerals and components are combined with the salt and they alter the spectral behavior of sensing surface (Allbed and Kumar, 2013). Utilizing spectral behavior of crops and plants can support indirect soil salinity mapping especially in cases where there is densely vegetated areas or when salt accumulation on the surface soil is limited (Weiss et al., 2016).

Gorji et al. (2020) generated spectral reflectance curves of various soil salinity classes by utilizing Landsat-8 OLI image of 16-bit data from western part of Urmia Lake, Iran. Based on the produced graph that is shown in Figure 2.3, it is apparent that soil salinity classes can be categorized and differentiated from each other by the help of RS visible bands especially the red band. On the other hand, in the NIR and shortwave spectral regions, spectral reflectance values of each salinity class indicate similar values. Thus, it is difficult to differentiate saline classes by using these bands.

Non-saline and slightly saline soil classes demonstrate higher reflectance values in comparison with other classes when it comes to shortwave infrared region of spectrum. Comparable reflectance behavior and similar trend were also obtained for Sentinel-2A data.

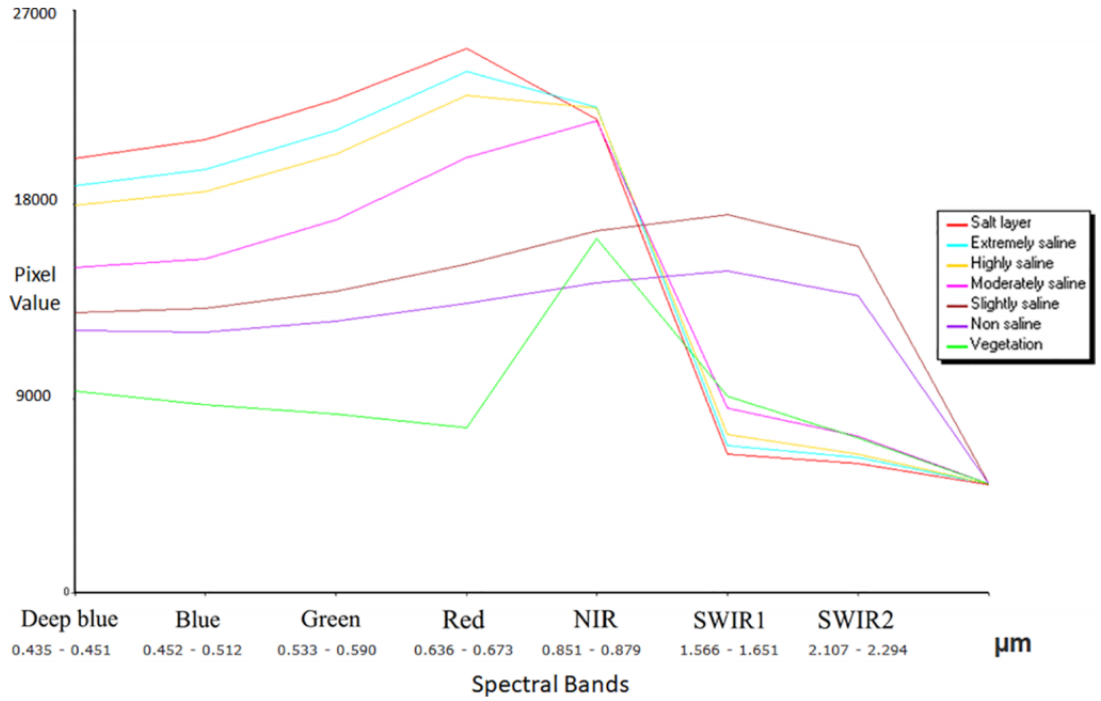

Figure 2.3 : Spectral reflectance curves of soil salinity classes based on Landsat-8 OLI (Gorji et al., 2020).

Typically, higher reflectance values are obtained from salty crusts in comparison with non-saline soils. Due to this fact their surface are smoother than cultivated non- saline soil surfaces (Metternicht and Zinck, 2003).

In literature, scientists use soil salinity classes that is introduced by the United States (USA) Salinity Laboratory in 1954. These salinity classes are expressed in deci Siemens per meter (dS/m) at 25 °C (Brown, 1954). Table 2.3 illustrates soil salinity classes in terms of electrical conductivity (EC) and corresponding effects of each individual salinity class on plants.

Table 2.3 : Effects of each soil salinity class on plants (Weiss et al., 2016).

Salinity class	EC (dS/m)	Salinity effects on plants
Non-saline	0-2	Negligible effects
Slightly saline	2-4	Yield loss for very sensitive crops
Moderately saline	4-8	Yield restriction for many plants
Highly saline	8-16	Only tolerant crops can survive
Extremely saline	>16	Only a few very halophytes can resist

2.2 Summary of Recent RS Approaches for Soil Salinity Mapping

Broad surveys applying RS data for detecting and mapping soil salinity has been conducted over the recent years, mainly with support of multispectral sensors. These incorporate Sentinel-2A, Landsat data series including (Landsat Thematic Mapper (TM), Landsat Multispectral Scanner System (MSS), Landsat 7, Landsat-8 OLI, Landsat Enhanced Thematic Mapper (ETM), Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER), SPOT, MODIS, Indian Remote Sensing (IRS) and IKONOS (Gorji et al., 2015).

In this research, RS based studies on soil salinity are overviewed and summarized according to their mapping techniques, RS data type, location, spatial extent, sensing approaches together with the reason of salinity for each case study. Sensing approaches were categorized according to acquiring information methods on land surfaces including airborne satellite images, photogrammetry, ground measurements and laboratory analysis. Overview of studies depicts that soil salinity mapping is mostly conducted by utilizing multispectral RS data in combination with simultaneous field measured data.

Selection of relevant articles published from January 1, 2014 up to March 30, 2021 is achieved by searching soil salinity as main keyword in web of science engine. Then, other keywords including remote sensing, satellite data, mapping methods and GIS were utilized for advance filtering of articles.

In Table A.1 that is inserted in the annex section, summary of the reviewed case studies are listed regarding to their publication years from the most recent to the oldest to demonstrate the progress and enhancement of RS methods through time. Several typical mapping techniques such as applying common place salinity and vegetation indices, correlation and regression analysis, decision tree classification (DTC),

principal component analysis (PCA), partial least square regression (PLSR), maximum likelihood classification have been extensively used in the past and still they are also preferred in the recent studies. Within the last decade, new soil salinity indices were generated from arithmetic combination of spectral bands, and this trend of innovating new soil salinity and vegetation indices is still in progress. In addition, several advanced models and classification algorithms like Cubist Model, random forest (RF) regression models, support vector machine (SVM), neural network model, and some other recently established models have been widely in use in the recent years (Gorji et al., 2019).

The literature review over various case studies indicates that most of researches on soil salinity mapping have been conducted in a local spatial extent and some are applied over regional scale. The overview revealed that soil salinity monitoring carried out in both developed and developing countries. Hence, threats of converting fertile lands to salt affected areas are becoming a more universal concern and many nations are attempting to govern this environmental phenomenon.

The overview of case studies proves that soil salinity mapping is fundamental for efficient monitoring, organizing and managing agricultural activities in salt-affected soils. In some field scale case studies for enhancing crop managements and agricultural productivity, temporal montoring has conducted by localized mapping. By this way most proper irrigation and soil management practices in various crop zones can be applied. On the other hand for regional scale case studies temporal monitoring is applied to track in progress salinization/desalinization over time and to guarantee timely delineation of crop management zones.

In terms of data that is applied in each research, the review exposes that multispectral sensors including IKONOS, MODIS, IRS, Huan Jing (HJ)-1A, Quickbird, ASTER, Landsat series, WorldView2 (WV2) and SPOT have been used for exploring soil salinity studies with the aim of detecting, monitoring and mapping saline soils. It can be interpreted that Landsat series have been widely used in comparison to other multispectral data regarding the fact that this data is freely available in the global sense and spanning from 1972 till today.

The overview indicates that most of the selected case studies are located in arid and semi-arid regions of the world like Middle East countries, India, China and United States. Indeed, preserving agricultural lands and food supply for rapidly increasing population in these regions is highly troubled.

Inspecting various case studies indicate that both primary and secondary salinization can be contemplated as sources of soil salinity. Despite, exploring studies which are performed in arid and semiarid regions of the globe depicts that anthropological factors not only exacerbate soil salinization specifically in agricultural lands but also the adverse effects of human induced activities has worsen natural causes of soil salinization. For instance, rising global temperature due to excess use of energy by many nations result in rapid evaporation and loss of soil moisture which leads to salinization specifically in arid regions.

The literature review proved that accurate soil salinity estimation requires updated and temporal data from different sectors which in many developing countries are not easy to produce.

The overview of case studies indicated that most of the researchers detect soil salinity directly or indirectly through various spectral band combinations specifically by utilizing salinity indices that derived from multispectral data rather than hyperspectral data. Despite, the effectiveness is impeded by the spatial and spectral resolutions of the multispectral images, plant coverage, agricultural practices, atmospheric effects, etc.Indeed, the multispectral imagery only produces information on soil surface and not from the soil profile since sensors obtain only reflectance data of surface soil. The restrains of remote sensing for obtaining data from the soil profile can be overcome by integrating remote and proximal sensing data with soil surveys and sampling. Accumulation of salt minerals on surface soil and through the soil profile particularly at the top 0 to 30 cm is a dynamic process as dissolved salts are transported by water and surface particles can be moved by wind. Therefore, true timing for soil sampling is highly significant for soil salinity assessment studies .

2.2.1 Principal Components Analysis (PCA)

Literature proves that soil salinity assessment specifically in regional scale requires large or massive data sets with several variables. PCA is commonly and frequently utilized in many researches to decrease the number of variables remarkably while still keeping much of the information in the original data set. PCA is apparently the most commonly and widely used dimension-reduction technique for this sort of analysis. It is an algorithm that converts dimensional image data into a new set of images (components) with orthogonal shape. The calculated statistical data of components

including variance/covariance matrices, average, standard deviation, correlation coefficients are used in this method. It should be considered that the first component always contain major information from the original image. PCA is typically a dimension-reduction method that applied broadly to assess and interpret a dataset with many variables. In different studies scientists apply PCA prior to main modeling in order to obtain a significant compact representation of a dataset. In fact, the whole dataset can be summarized in terms of some principal components in place of original lots of variables. Then the components can be applied in several various purposes.

First principal component is a linear combination of original predictor variables which include the maximum variance in the data set. It regulates the direction of highest variability in the data. In each preprocessing the higher the variability obtained in the first component, the larger the information captured by component. In fact, no other component can have variability greater than the first principal component.

2.2.2 Classification and Regression Trees (CART)

CART is a specific technique of producing decision rules to describe the difference among clusters of observations and categorize the class of new observations.

It is an algorithm that estimates both qualitative and quantitative parameters by applying training samples (Hateffard et al., 2019). Classification and Regression Trees (CART) is a structure of binary nodes that separate entities into two sub-groups at each node. Abundant training samples are required since rapid division of entities into sub-groups make training samples smaller in each step (Bittencourt and Clarke, 2004). In this method, variables are divided to independent and dependent ones. Indeed, dependent variables are predicted by utilizing independent variables that are considered as measurements and analysis. Classification trees or regression trees are representing dependent variables in this methodology (Choubin et al., 2018).

In this method, initially learning sample data are divided into subsets in order to find best node questions. Parallel to this step, an iterative computer procedure is defined and all possible splits are processed. There are several advantages of utilizing this technique. Some analysis techniques require selection of variables prior to main process; but, CART will automatically select the most appropriate ones from a given subset of variables constituting a learning sample. So, model will select true splits by

itself even if learning samples include inappropriate information because of measurement errors. Moreover, CART can process datasets with complex structures containing any combination of continuous and categorical data. These advantages of CART will be supportive for any researcher to establish models without dependency to a specific class of data and enables scientists to obtain more realistic effects and improve prediction accuracy.

2.2.3 Linear Spectral Unmixing (LSU)

As a commonly used RS method, LSU is applied for accurate prediction of various endmembers with their spectral signatures and fractional abundances. Studies utilizing this approach indicated that in some conditions while there is variability in spectral signature, poor spatial resolution, large data size, mixing of materials at various scales, not availability of pure endmember utilizing LSU can be a challenging task. LSU presume that the pixel is a linear mixture of its constituents and spectral variation is caused by a limited number of surface materials such as water, vegetation, soil and shadow.

Land cover classification is normally based on categorizing individual pixels of satellite images and allocating each pixel to a class. Despite sometimes in one pixel one may observe more than one land cover (for instance; vegetation, soil and water). LSU is a functional method to predict the fractions of various land cover classes for a given pixel. 'End members' that are pure reflectance of corresponding classes are required to apply this method. It is considered that each pixel is representing linear combination of all end members in the image frame (Ramak et al., 2015).

2.2.4 Decision-tree Analysis (DTA)

It is a supervised non-parametric learning approach used for classification and regression in order to generate a model that estimates the value of a target variable by learning simple decision rules derived from the data structure. DTA is a prediction mapping model that is supportive for soil salinity detection over large areas. In this method, those environmental variables that significantly impact on spectral behavior of saline soils could be integrated in the classification procedure. Information on environmental parameters like terrain and landform maps, geological maps could

considerably improve the accuracy of soil salinity. Regarding to data analytics, DTA is a kind of technique that contains conditional 'control' structures to categorize data. A decision tree initiates at a single point and then divides in two or more directions. Each direction suggests various possible outcomes, including a set of decisions and chance events until a final outcome is accomplished. Since DTA splits large and complex data into more applicable segments, it is highly useful for machine-learning and data analytics.

2.2.5 Random Forest (RF)

It is one of the major machine-learning techniques that is functioning according to CART (Hoa et al., 2019; Liu et al., 2019; Wang et al., 2019; Fathizad et al., 2020). RF is generated based on incorporation of multiple decision trees, which have no correlations with each other and each tree has its own evaluation process (Wang et al., 2019). Both class variables and regression variables can be predicted by utilizing this method (Li et al., 2019). As an ensemble algorithm, RF trains various decision trees in parallel with bootstrapping followed by aggregation, jointly referred as bagging. Bootstrapping specifies that several individual decision trees are trained in parallel on various subsets of the training dataset applying different subsets of available features. It quarantines that each particular decision tree in RF is distinct, which diminishes the total variance of the RF classifier. In the final step, RF classifier integrates the decisions of individual trees. Therefore, RF classifier depicts relevant generalization. Due to high accuracy and absence of overfitting issues, RF outperforms most of the other classification approaches. It is also more dynamic for the selection of training samples and noise in training dataset. Another advantage of RF is that unlike DT classifier, it does not require feature scaling.

2.2.6 Inverted Gaussian Function (IG)

IG is an iterative likelihood distribution to obtain unity as default for both mean and scale. It is more relevant to model a case study containing numerically extensive values since the tails of the distribution decline more slowly than the normal distribution. It is applied to measure salt content in surface soil assuming the variations in spectral reflectance. In IG model, variables of saline soil spectral reflectance are

typically responsive to both overall and maximum reflectance positions. Maximum reflectance value changes to shorter wavelengths, while soil salinity level increases from slightly saline to extremely saline (Farifteh et al., 2010).

2.2.7 Support Vector Regression (SVR)

Support Vector Machine (SVM) is one of the commonly used machine-learning algorithms and a kernel-based structure method called SVR is derived from SVM for specific analysis. SVR has various variables including gamma and cost value and also linear, polynomial and radial basis function (RBF) as kernel type parameters (Gorelick et al., 2017).

It is highly contributive for modeling the complex nonlinear relationships in the multi-dimensional or hyper-dimensional feature space and predicts the linear dependency of the variables to be predicted on the predictive covariates by fitting an optimal approximating hyperplane to the training data. For linearly non approximal problems, the training data are implicitly mapped by a kernel function with regularization into a higher dimensional space, wherein the new data distribution enables a better fitting of a linear hyperplane that appears nonlinear in the original feature space (van der Linden et al., 2014). Selecting the parameters of the kernel function , regularization and loss function is a significant procedure for users who apply SVR algorithms for their studies.

2.2.8 Cubist Model

Cubist is a program for producing rule-based predictive models from input data. As a rule–based model, Cubist is a structure of trees that contain linear regression models and can predict numerical values. These models are built according to predictors applied in previous splits. In addition, intermediate linear models are embedded at each step of the tree. A prediction is made using the linear regression model at the terminal node of the tree, but is “smoothed” by taking into account the prediction from the linear model in the previous node of the tree. The tree is reduced to a set of rules, which initially are paths from the top of the tree to the bottom. The rules have the structure of a boolean statement (if[], then[], else[]), one response for when true, and

an alternative response for not true. Indeed, the rules split data into similar classes which can then be conveniently examined with linear regression (Smith et al., 2020).
In each Cubist model there are two essential extensions namely; "Data" that is the input variables for establishing model and "Names" that is the description of application's attributes. The primary algorithm contains two steps. The first step establishes a set of rules that divides the data into smaller subsets. The second part of the algorithm applies a regression model to these smaller subsets to arrive at a prediction.

2.2.9 Partial Least Square Regression (PLSR)

It is a multivariate statistical algorithm that provides an environment to compare multiple response variables with multiple explanatory variables. PLSR contains broad class of methods for modeling relations between sets of observed variables by means of latent variables. It includes regression and classification structures as well as dimension reduction algorithms and modeling tools. The fundamental hypothesis of all PLSR methods are that the observed data is generated by a system or process which is driven by a small number none directly observed or measured variables. When there are lots of predictor variables, and all of the parameters are considerably correlated with each other, PLSR is used to model the most response variable. Regarding to inferential competency of PLSR method, it can be applied to model the relationship between measured spectral reflectance values and salt concentrations in soil profile (Farifteh et al., 2007). Visible and NIR spectral reflectance values are the most useful input data in PLSR method that can be applied for the accurate prediction of accumulated salts in the soil (Sidike et al., 2014). (Sidike, Zhao, & Wen, 2014).

3. STUDY AREAS AND DATA USED

This study examined measured data from three different regions located in the vicinity of Urmia Lake in Iran (West playas and Bonab Region) and around Tuz Lake in Turkey.

3.1 Tuz Lake in Turkey

As the second largest lake in Turkey, it is located among the boundaries of three provinces; namely, Aksaray, Konya and Ankara. Tuz Lake cannot be considered a deep lake since the deepest zones has approximate depth of 0.3-0.5 m. Wadis are the main natural sources that feed the lake and due to high evaporation they are shrinking or absolutely get dry during summer months. This lake supplies 55% of Turkey's salt demand. In 14 September 2000, it has been asserted as a special environmental protection zone. In terms of animal's habitats and protection of biological diversity, it is considered as one of Turkey's richest lakes (Yagmur et al., 2021). In this basin 85 various species of births, fifteen mammal species and 38 endemic plant species are existed (TVKGM, 2014). Economic condition of this region is highly dependent on salt production, tourism, agriculture, and livestock breeding (Mergen and Karacaoglu, 2015). The lake basin occupies an area of approximately 1.500 km^2 with an altitude of 905 m (Ucan and Dursun, 2009).

Figure 3.1 illustrates the geographical location of the region with a Landsat-8 OLI Image (Path 177 Row 33 Acquired 18 April 2020)

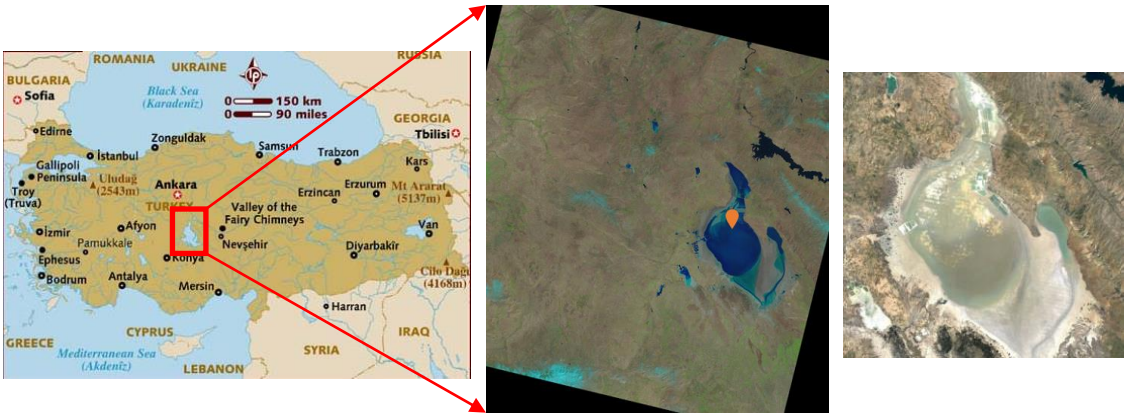

Figure 3.1 : Geographical location of Tuz Lake in Turkey

Geographical location of all the 322 measured field samples are demonstrated in Figure 3.2. Totally, 33 EC measurements were selected from a total number of 322 soil samples which were measured by the State officials within the context of a National State Project (Tuz Lake, 2004) in the period of May–July 2002. The sample selection was done considering homogeneous distribution of samples with different EC ranges around the lake. For generating soil salinity maps, Landsat 5 image was acquired in a simultaneous time with measured samples in June 2002 from United States Geological Survey (USGS) web site.

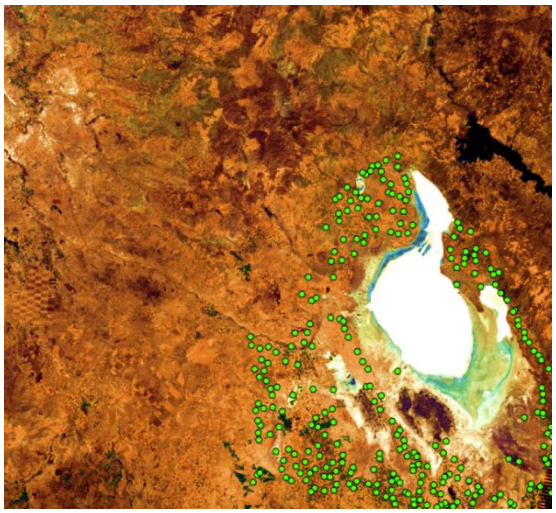

Figure 3.2 : Demonstration of all soil samples that were collected around the Tuz Lake.

3.2 Urmia Lake in Iran

With a surface area of around 5.000 km^2, Urmia Lake is known as one of the largest saline lakes of the world (Ghalibaf and Moussavi, 2014). For this research, two different case study areas around Urmia Lake Basin that is located at the northwestern part of Iran were examined. The first area is located in west playas of basin and the second one belongs to Bonab Region that is placed in the southeastern part of the lake.

Regarding to environmental characteristics of the region, precipitation commonly occurs from October to April with a mean annual rate of around 360 mm and there is seasonal temperature variation in this semi-arid region (Hamzehpour et al., 2014). In the vicinity of the lake, there is no significant elevation variation and the region has mostly flat topography. As one of the major industrial and agricultural provinces of Iran, this region has experienced remarkable population growth and urban development within the past few decades (Taghipour et al., 2013). Groundwater salinization due to lake shrinkage and appearance of salty dust storms from accumulated salt on the surface soil are considered as the two major environmental problems in this region (Hamzehpour and Bogaert, 2017).

3.2.1 West Playas of Urmia Lake

The selected region for analysis in West playas of Urmia Lake with an area of approximately 18 km^2 is demonstrated in Figure 3.3. Both primary and secondary causes of salinization including presence of saline playa deposits and excess use of poor quality irrigation water have intensified soil salinization problems in the agricultural lands of the basin (Hamzehpour et al., 2014).

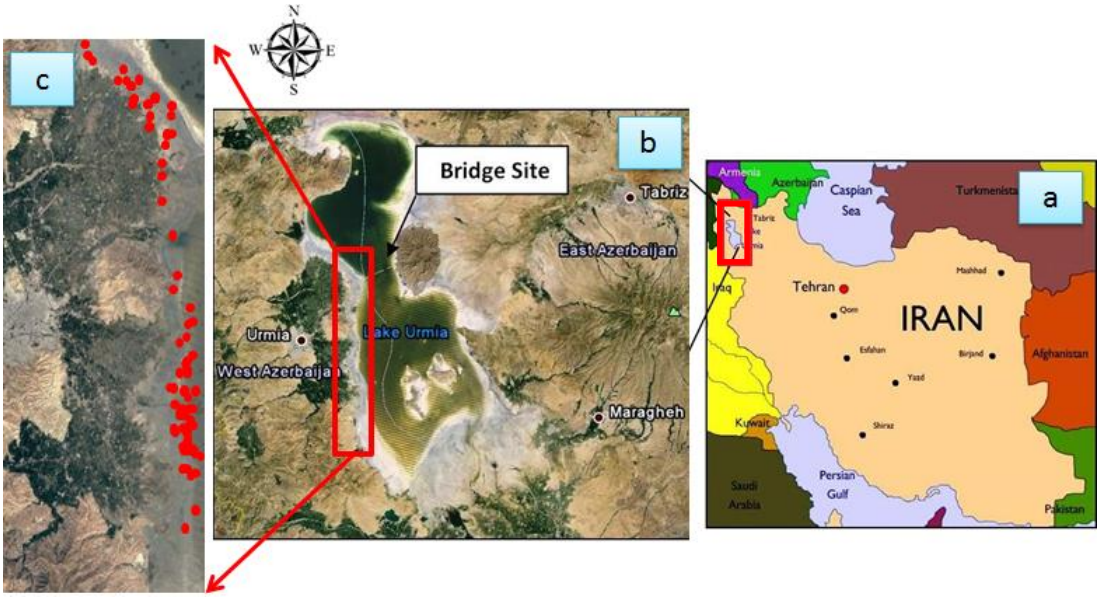

Figure 3.3 : (a) Geographical location of Urmia lake in Iran, (b) General view of study area in west playas of Urmia lake, (c) soil sampling locations

For this case study, satellite images and field measurements were both used for preparing the soil salinity maps. Between 02–12 October 2018, totally 133 soil samples were collected from the top 20 cm of barren land including rare halophilic plants. One Sentinel-2A image of 14 October 2018 and one Landsat-8 OLI image dated 13 October 2018 were acquired by downloading from the United States Geological Survey (USGS) website. The field measurements were collected from the study area of 18 km^2 at the end of the dry season where there was significant salt accumulation on top soil. 1:2.5 soil- to- water suspension method was used for measuring their EC values of soil samples. All ground measurements data for Urmia Basin and Bonab Region were provided by a scientific team under supervision of Dr. Nikou Hamzehpour from the University of Maragheh, Iran.

3.2.2 Bonab Region

Bonab Region with elevation ranges between 1270 to 1300 m is located in southeastern part of Urmia Lake. Figure 3.4 depicts geographical location of this region which is placed at intervals of 45^o 58’ 41’’ to 46 ^{o} 02’ 35’’ E longitudes and 37 ^{o} 20’ 21’’ to 37 ^{o} 16’ 18’’ N latitudes. In this region, due to continuous shrinkage of the lake area, huge amount of salt was accumulated on the soil surface and the playas

were becoming extremely susceptible to wind erosion (Rahmati and Hamzehpour, 2017).

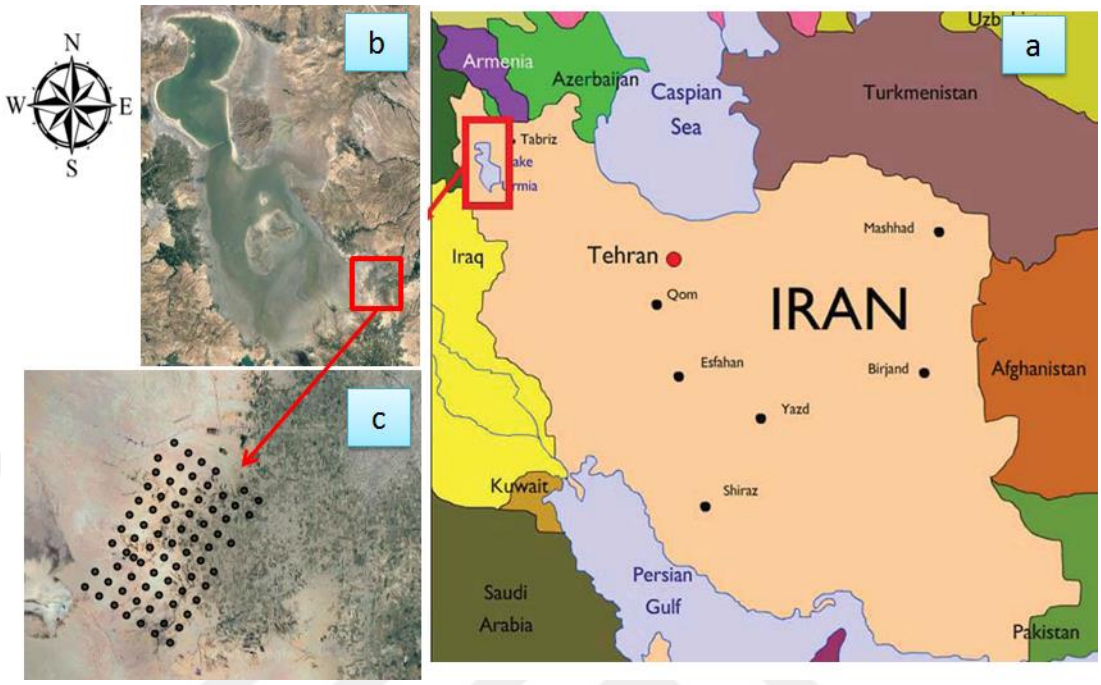

Figure 3.4 : (a) Geographical location of Urmia lake in Iran, (b) General view of Bonab plain, (c) soil sampling locations

Totally 77 soil samples were measured in Autumn 2014 on a rectangle area which covers the 40 km2 of the study area. Saturation paste soil electrical conductivity (ECe) was measured at a depth of 0-25 cm in the laboratory using a Jenway 4510 bench conductivity meter.

One Landsat-8 OLI satellite image, which was acquired on 15 September 2014, was utilized. For matching coordinates of geographical information obtained from different sources, Universal Transverse Mercator (UTM) projection system with Zone 38 N and World Geodetic System (WGS-1984) datum were used. Landsat-8 OLI data was selected since it was approximately simultaneous with sampling time and there was no relevant data from Sentinel- 2A imagery.

3.2.3 EC Measurements

Analysis based on EC of soil is a practical approach to measure soil salinity. For this aim, several methods have been developed and the suitability of a given approach depends on the intended use of the measurement, the soil water content, the

established methodology in the area, and the availability of specialized equipment. Indeed, the most conventional methods for providing ground EC data are manual methods for both field and laboratory measurements. Despite some studies integrate the information that obtains from proximal sensors with laboratory data. Proximal sensors refer to sensors that obtain data from the soil when they are in contact with the soil or within the 2 m distance with surface soil.

Sensors types that are currently used for crop and land management applications are electrical resistivity sensors, dielectric sensors, and electromagnetic induction (EMI) sensors. EC measurement with sensors enables continuous or regular monitoring at different locations. Dielectric sensors are useful for on-going monitoring at different depths and specific locations. EMI sensors are useful for mapping spatial variation.

Conventional soil salinity detection necessitated geo-referenced field sampling and laboratory analysis for calculating the electrical conductivity of soil saturation extract . Other rapid approaches have been commonly used such as us of RS imagery and geophysical methods (EM38) for salinity diagnostics. The EM38 is quite applicable for agricultural salinity assessment. It is a quick and mobile method for measuring bulk soil electrical conductivity. It provides 1.5-m and 0.75-m depth of examination of the vertical and horizontal modes, respectively.

EMI sensors are lightweight, compact, non-invasive, and non-destructive instruments that are capable of making rapid field EC measurement without contacting with soil. Moreover, they can characterize relatively large soil volume in different directions and even in stony soils. Despite there are also some disadvantages in the measurement procedure. Soil temperature should be measured at different depths and calibrate with the reference temperature. In addition, EC value can be affected by metallic objects closer than 1m and EC measurements are limited to soil moisture between 0.5–1 of field-capacity .

Accurate soil salinity mapping via RS data requires relatively equal number of EC measured data for each of different soil salinity classes. Generally, extremely saline soils and none-saline soils can be detected easier than the other three classes including slightly saline soils, moderately saline soils and highly saline soils. Uneven distribution of EC ground data for different soil salinity classes result in difficulties of categorizing boundaries of each salinity class.

In this research, as it is shown in Table 3.1, EC ground measurements of Urmia Lake Basin are distributed evenly for all salinity classes. Despite, soil salinity mapping for other two case studies were completed under data scarce conditions.

Table 3.1 : EC ranges and number of model samples in each salinity class

Study area	Total samples	Number of modelling samples	Number of validation samples	EC ranges	Number of model samples in each salinity class					
					0-2 dS/m	2-4 dS/m	4-8 dS/m	8-16 dS/m	16-32 dS/m	32< dS/m
Bonab Region	77	51	14	0-107	19	10	7	8	4	3
Tuz Lake	322	33	7	0-25	10	7	6	7	3	0
Urmia Lake	133	92	21	0-150	12	15	14	17	16	18

3.3 Satellite Images

In this section, brief information including history of Landsat-8 OLI, Landsat Thematic Mapper 4 and 5 (TM4 and TM5) and Sentinel-2A satellites are given. Moreover spectral and spatial information of each sensor are shown in separate tables. In addition summary of their application in different fields of studies are explained.

3.3.1 LANDSAT-8 OLI, Landsat Thematic Mapper 4 and 5 (TM4 and TM5)

As one of the significant NASA earth-observation Plans, Landsat satellite was launched to space in 1972. As a sensor flying on Landsat 4 and 5 satellites, Landsat Thematic Mapper (TM) had provided data for many years. In 1982, Landsat 4 was launched and after providing data for more than 10 years its operation was completed in 1993. After that, between years 1984 to 2011, Landsat 5 was launched and provided numerous RS data continuously. Indeed, Landsat series were launched for providing various satellite images in different studies including military, climate, agriculture,

forestry, geology and land-use planning. Continuous image obtaining schedule is known as the main advantages of Landsat satellite images. Indeed, its temporal resolution is 16 days. Moreover, its freely accessible achieves history that contains data from year 1982, support multi-temporal mapping surveys for plenty of analyses. On the other hand, its 30 m spatial resolution can be counted as a disadvantage since it restrains providing detail mapping.

Bands, wavelength and resolution of Landsat TM4 and TM5 are shown in Table 3.2

Table 3.2 : Spectral and spatial information about Landsat TM4 and TM5.

Landsat 4-5 Spectral Bands	Wavelength (micrometres)	Resolution (m)
Band1-Blue	0.45-0.52	30
Band2-Green	0.52-0.60	30
Band 3- Red	0.63-0.69	30
Band 4- Near-Infrared	0.76-0.90	30
Band 5- Near-Infrared	1.55-1.75	30
Band 6- Thermal	10.40-12.50	120
Band 7- Mid-Infrared	2.08-2.35	30

In 2013, Landsat-8 OLI satellite with its two main sensors, OLI and TIRS, was successfully launched.

OLI with its nine spectral bands including visible, NIR, SWIR, Panchromatic, Cirrus and Thermal Infrared collecting images covering wide areas of the earth's landscape, and providing relevant resolution to distinguish features like urban buldings, roads, rails, agricultural lands, water surfaces, vegetations and other surfaces. Landsat-8 OLI spectral bands with their corresponding wavelength in micrometers and spatial resolutions in meter are depicted in Table 3.3.

Table 3.3 : Spectral and spatial information about Landsat-8 OLI.

Landsat-8 OLI Spectral Bands	Wavelength (micrometers)	Resolution (m)
Band 1 - Coastal aerosol	0.43 - 0.45	30
Band 2 - Blue	0.45 - 0.51	30
Band 3 - Green	0.53 - 0.59	30
Band 4 - Red	0.64 - 0.67	30
Band 5 - Near Infrared (NIR)	0.85 - 0.88	30
Band 6 - SWIR 1	1.57 - 1.65	30
Band 7 - SWIR 2	2.11 - 2.29	30
Band 8 - Panchromatic	0.50 - 0.68	15
Band 9 - Cirrus	1.36 - 1.38	30
Band 10 - Thermal Infrared	10.60 - 11.19	100 (30)
Band 11 - Thermal Infrared	11.50 - 12.51	100 (30)

3.3.2 Sentinel-2A

In June 2015, as the first optical earth observation satellite in the European Copernicus Program, Sentinel-2A was successfully sent off from the spaceport in Kourou, French Guiana. Obtained satellite images from this sensor are providing relevant information on generic land cover, land use and change detection maps. Characteristics of Sentinel-2A spectral bands with their corresponding wavelengths are shown in Table 3.4.

Table 3. 4 : Spectral and spatial information about Sentinel-2A.

Sentinel-2A Spectral Bands	Central Wavelength (micrometres)	Resolution (m)
Band1-Coastal aerosol	0.443	60
Band2-Blue	0.49	10
Band3-Green	0.56	10
Band 4- Red	0.66	10
Band 5- vegetation red edge	0.7	20
Band 6- vegetation red edge	0.74	20
Band 7- vegetation red edge	0.78	20
Band 8-NIR	0.84	10
Band 8A- vegetation red edge	0.86	20
Band 9- water vaper	0.94	60
Band 10- SWIR-cirrus	1.37	60
Band 11- SWIR	1.61	20
Band 12- SWIR	2.19	20

4. METHODOLOGY

The methodology of this research includes two parts. In the first part, steps of generating the new salinity index by utilizing relevant EC data and their corresponding visible and NIR band pixel values are explained. In the second part of the research, explanation about how the new generated soil salinity index is utilized for producing soil salinity maps in each of the geographical locations is given. Moreover, after applying several soil salinity indices, two commonly used salinity indices have been selected and applied for all the three case studies in order to compare their output soil salinity maps with the maps produced from the new soil salinity index. Finally, all three salinity indices along with individual visible and NIR spectral bands were adopted in Cubist model for generating soil salinity maps in the three case study areas.

4.1 Generating the New Soil Salinity Index

Typically, extremely and highly salt-affected lands can be mapped by all kinds of soil salinity indices. In the regions with accumulated salts on surface soil, even there is no need to apply any index since salt crust are easily visible due to their bright reflectance. Despite, detecting slightly saline and moderately saline soils required accurate testing of different soil salinity indices in order to obtain optimized salinity maps. In this research, spectral behavior of non -saline regions compared with extremely saline locations to investigate which arithmetic combination of spectral bands provide highest gap between SI values of non-saline ones and SI values of extremely saline soils. Indeed, an appropriate salinity index should have strong capability to show a large gap between SI value of non- saline soils and extremely saline soils. According to this fact, various arithmetic operations were tested to check how one can increase the gap to an acceptable value. In the first step, information on visible and NIR band pixel values were extracted for the three case studies separately. Then, descriptive statistics of all bands in each area was calculated. In this step, training samples from non-saline ones and extremely saline ones was selected in order

to compare their descriptive statistics with each other. Comparison and trial and error examinations proved two important facts during the generation of the new index. Firstly, multiplication of blue and red band provided a relatively high pixel values for extremely salt-affected soil and this combination was highly useful for salt detection. Due to this fact, multiplication of blue and red band was used as numerator of the equation since it had directly proportional relationship with extremity of salinity. Secondly, in all the analyses absolute value of difference between NIR and red band (|Red-NIR|) gave approximately highest integer value for all non-saline samples and it put forth nearly the lowest integer value for extremely saline samples. Considering this fact, (|Red-NIR|) was used as denominator of the new index equation since bigger value in the denominator of the equation resulted in lower salinity index value for non-saline soils. The new salinity index generated within this research is given below;

New salinity index $= (\text{Band Blue} * \text{Band Red}) / \text{Abs}(\text{Band NIR} - \text{Band Red})$

According to Figure 4.1, spectral reflectance value of 8 extremely saline samples ranging between 76 to 123 dS/m were compared with 8 slightly saline soil samples ranging between 1.3 to 6.2 dS/m. Based on this spectral behavior, it was found that the gap between NIR spectral values with visible band spectral values increased while the trend shifted from extremely saline samples to slightly saline ones. In the generated new index, the difference between NIR band and red band has been taken into consideration because the gap between these two was minimum for extremely saline ones. Figure 4.2 demonstrates visible and NIR spectral band values of 8 extremely saline samples and with 8 slightly saline soil samples. The graph shows high spectral values in all the 4 individual bands for extremely saline soils. Despite, one can observe lower values in visible bands for slightly saline soils. Regarding to spectral behavior of slightly saline soils, it was seen that a sharp rise in spectral values occurred when the trend shifted from visible bands towards NIR band. According to the red and green dot shape arrows on the graph, the slope of trend between red and NIR bands for slightly saline soils was steeper than extremely saline soils. Same analysis has been repeated for Bonab Region and Tuz Lake that are shown in Figure 4.3, Figure 4.4 and Figure 4.5, respectively. Similar results for these two regions contributed significant information for generating the new index. Indeed, these information gained from the

two other regions highly supported the mathematical logic for generating the new index.

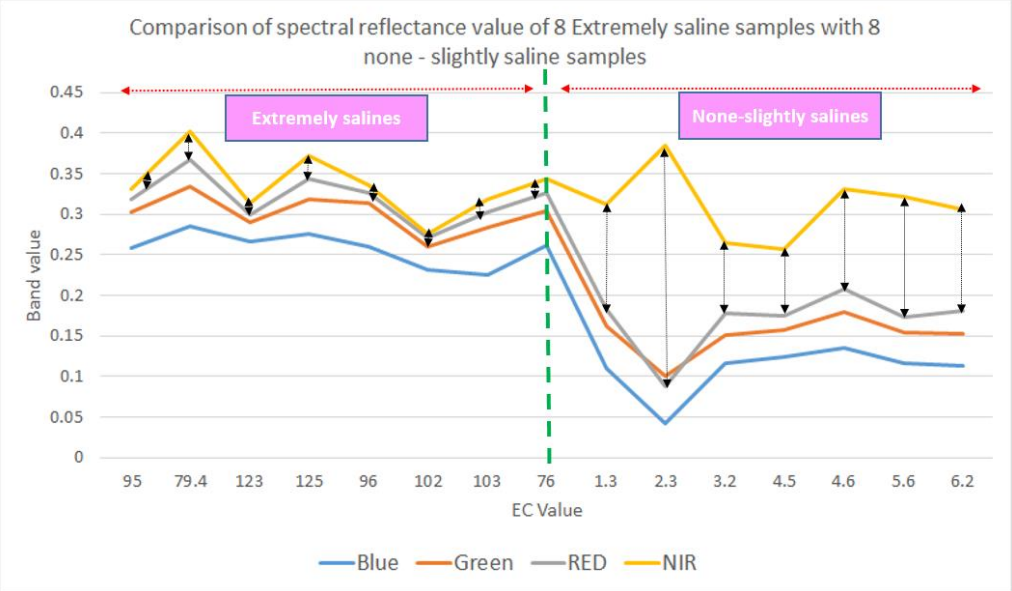

Figure 4.1 : Comparing spectral behavior of extremely saline soils with non-saline soils located in West playas of Urmia Lake.

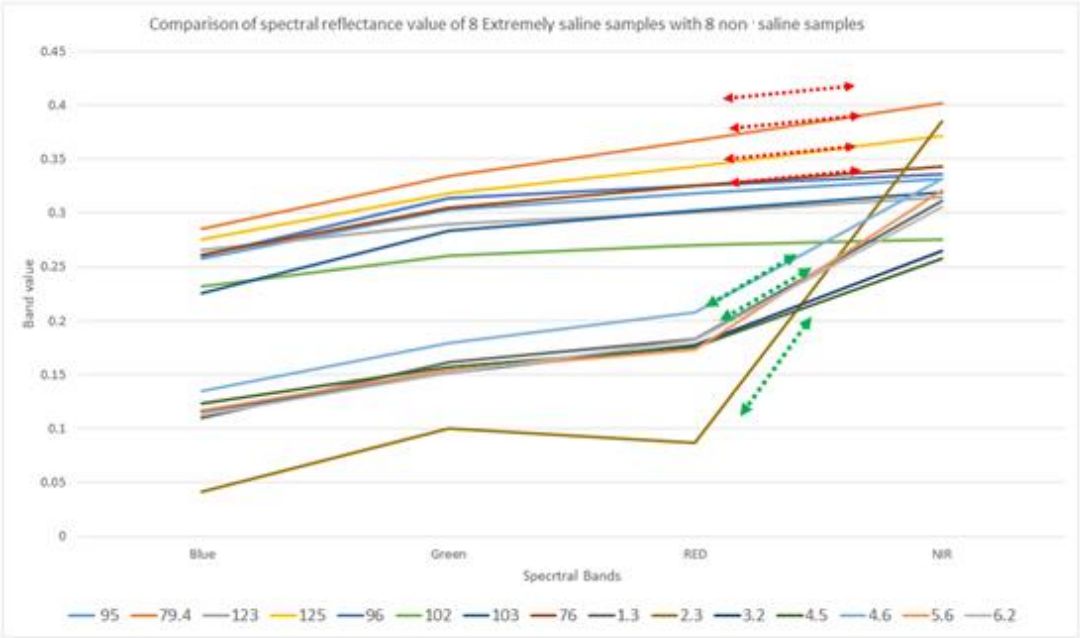

Figure 4.2 : Comparing visible and NIR spectral reflectance value of saline soils with slightly-saline soils located in West playas of Urmia Lake.

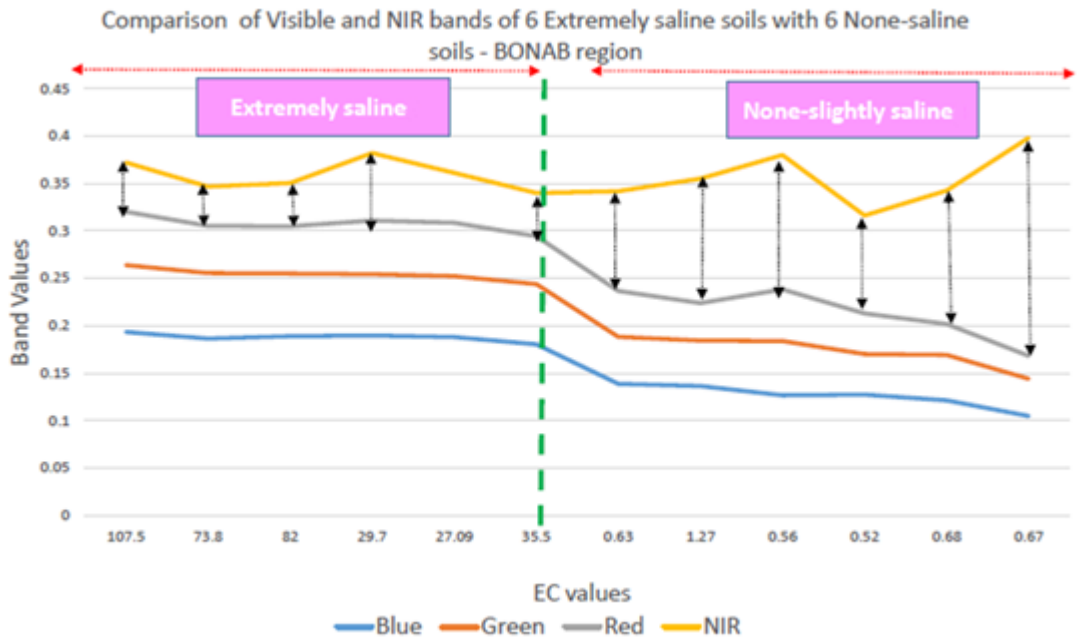

Figure 4.3 : Comparing spectral behavior of extremely saline soils with non-saline soils located in Bonab Region.

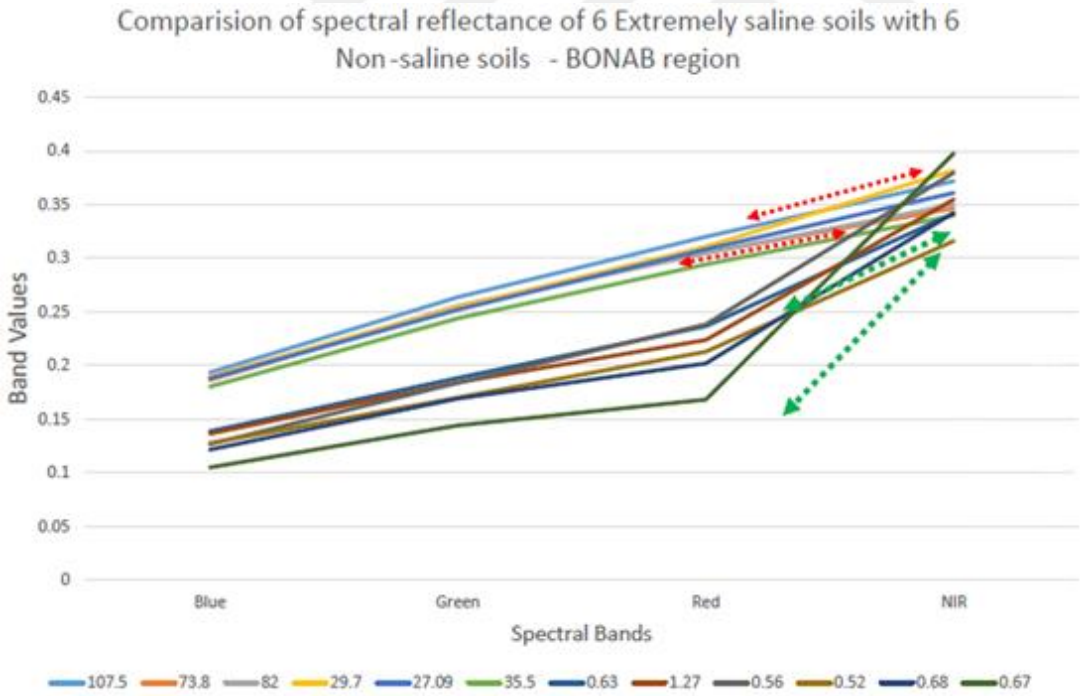

Figure 4.4 : Comparing visible and NIR spectral reflectance value of saline soils with slightly-saline soils located in Bonab Region.

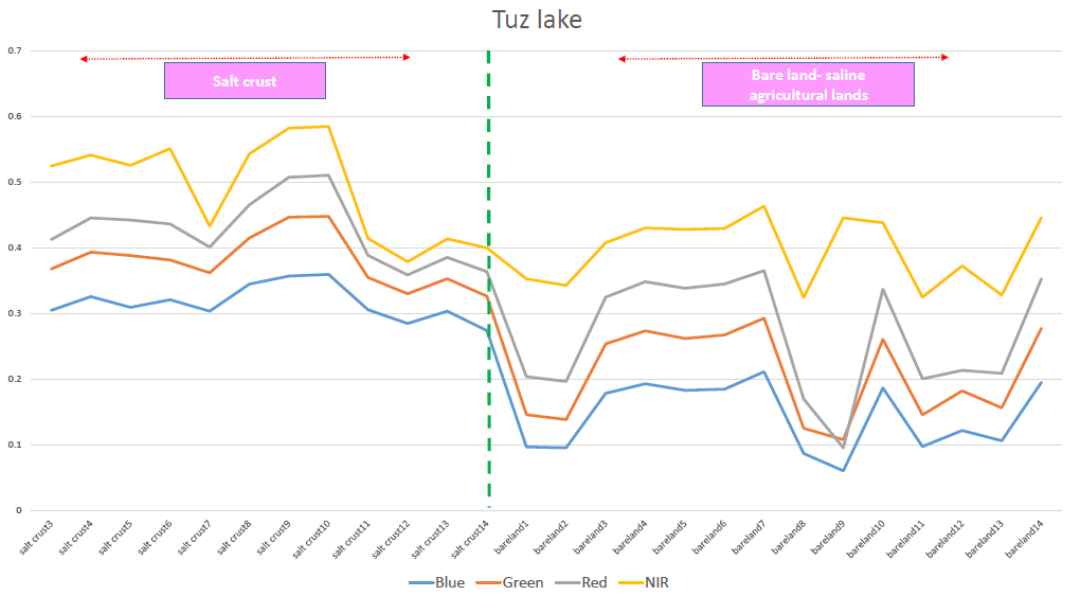

Figure 4.5 : Comparing spectral behavior of extremely saline soils with non-saline soils located in Tuz Lake.

Table 4.1 : Soil salinity indices used for generating salinity maps.

Index	Description
SI 1=$(B1 * B3)/B2$	B1=Band 1 :Blue , B2=Band 2:Green , B3= Band 3:Red
SI 2 = $\sqrt{B1 * B3}$	B1=Band 1:Blue , B3= Band 3 : Red
New Index = $(B1 * B3)/\lvert B4 - B3\rvert$	B1=Band 1:Blue , B3= Band 3 : Red, B4= Band 4 : NIR

For correlating RS data with EC measurements, linear and multiple regression analyses have commonly been used (Allbed and Kumar, 2013). For this research, initially, around 7 salinity indices that are shown in Table 4.2 were tested to find the acceptable correlation between spectral values of indices as independent variables with EC values as dependent variable. Two salinity indices that demonstrated satisfactory results were chosen to compare their performance with the new generated salinity index. Most of the other indices revealed almost acceptable correlation with EC values. Regarding to each case study, in the first step, index calculation and regression analysis was applied for all the three indices. In the next step, level- slicing technique was applied for defining corresponding index range of each soil salinity class.

Table 4.2 : Seven preliminary used soil salinity indices for generating salinity maps.

Index	Description
SI 1=$(B1 * B3)/B2$	B1=Band 1 :Blue , B2=Band 2:Green , B3= Band 3:Red
SI 2 = $\sqrt{B1 * B3}$	B1=Band 1:Blue , B3= Band 3 : Red
SI 3==$(B2 + B3)/2$	B2=Band 2:Green , B3= Band 3 : Red
SI 4 = $\sqrt{B2 * B3}$	B2=Band 2:Green , B3= Band 3 : Red
SI 5== $B1/B3$	B1=Band 1:Blue , B3= Band 3 : Red
SI 6=$(B3 * B4)/B2$	B4=Band 4 :NIR , B2=Band 2:Green , B3= Band 3:Red
SI 7=2*B2-5*(B3+B4)	B4=Band 4 :NIR , B2=Band 2:Green , B3= Band 3:Red

4.2 Adopting Soil Salinity Indices as Input Data in Cubist Model

After examining soil salinity mapping performance of each index individually, Cubist model was applied to investigate which soil salinity index as an independent input variable can contribute more for generating soil salinity maps in the selected case studies.

4.2.1 Preparing Data for Cubist Model

First of all, in each case study EC was considered as target attribute or dependent variable. Other attributes including new index, SI1, SI2, individual band values and other available soil properties were considered as independent attributes that provided information for estimating the target attribute. Indeed, Cubist Model's function is finding how to predict the EC target value in terms of its attribute values. Cubist performs this function by generating a model including one or more rules, where each rule is a conjunction of conditions associated with a linear expression. In other words,

if a case satisfies all the conditions, then, the linear expression is relevant for estimating the target value.

Every Cubist application has two required extensions including "Names" that is the description of application's attributes and "Data "that is the input variables for generating the model. In addition, there is also another optional extension called "test" that includes unseen cases used to test the model.

Data file provides information on the training cases that Cubist will apply to generate a model. The entry for each case consists of one or more lines that give the values for all explicitly-defined attributes. Values are separated by commas and the entry for each case is optionally completed by a period. Models produced by Cubist are assessed on the training data from which they were constructed. In the output file, three parameters including average error, relative error and correlation coefficient are demonstrated as indicators for evaluating the training data. The average error is the typical degree to which a series of observations that are inaccurate with respect to an absolute observation. The relative error is the ratio of the average error to the error that would result from estimating the mean value; for acceptable models, this value must be below 1. The correlation coefficient calculates the degree of equivalency between the absolute values of the target attribute and those values estimated by the model. In this research we utilized Cubist, since it is a multiple sub-function linear tree model and uses a looping partitioning of the predictor variable space (Ziqiang et al., 2017). It considers divide-and-conquer approach and function to minimize the intra subset variation at each node. Cubist models take the form: if [conditions] then [linear model] approach. If the predictor variables associated with an observation satisfy a set of conditions, the linear model is used to predict the response. The advantage of the condition set in each rule is that they enable interactions to be handled automatically by allowing different linear models to capture the local linearity in various parts of the predictor variable space (Peng et al., 2019).

5. RESULTS AND DISCUSSION

In this research, two main types of salinity maps were produced. Salinity maps were initially generated by applying regression analysis to correlate new index, SI1and SI2 values with measured data in all the three case studies. Table 5.1 illustrates the summary result for regression analysis of the case studies using Landsat-8 OLI data.

Table 5.1 : Results of regression analyses for case studies using Landsat-8 OLI data.

Study area	Salinity index	Total model samples	Model R^2	Model RMSE	Model MAE	Total validation samples	Validation R^2
West playas Urmia	SI1	92	0.72	10.21	8.34	21	0.61
	SI2		0.73	10.11	8.12		0.63
	New SI		0.75	9.95	8.11		0.64
Bonab Region	SI1	51	0.68	11.52	9.32	14	0.62
	SI2		0.71	11.23	9.12		0.64
	New SI		0.74	10.98	8.94		0.66
Tuz Lake	SI1	33	0.82	9.78	8.23	7	0.68
	SI2		0.78	10.12	8.56		0.64
	New SI		0.81	9.84	8.31		0.62

Regression analysis results indicated that soil salinity maps generated by new SI exert acceptable results with model R^2 values similar to model R^2 values of other indices in all the three case studies.

Figure 5.1, Figure 5.2 and Figure 5.3 depicts linear regression results between ground measurement EC values and SI values for Bonab Region, Tuz Lake and Urmia Lake regions, respectively.

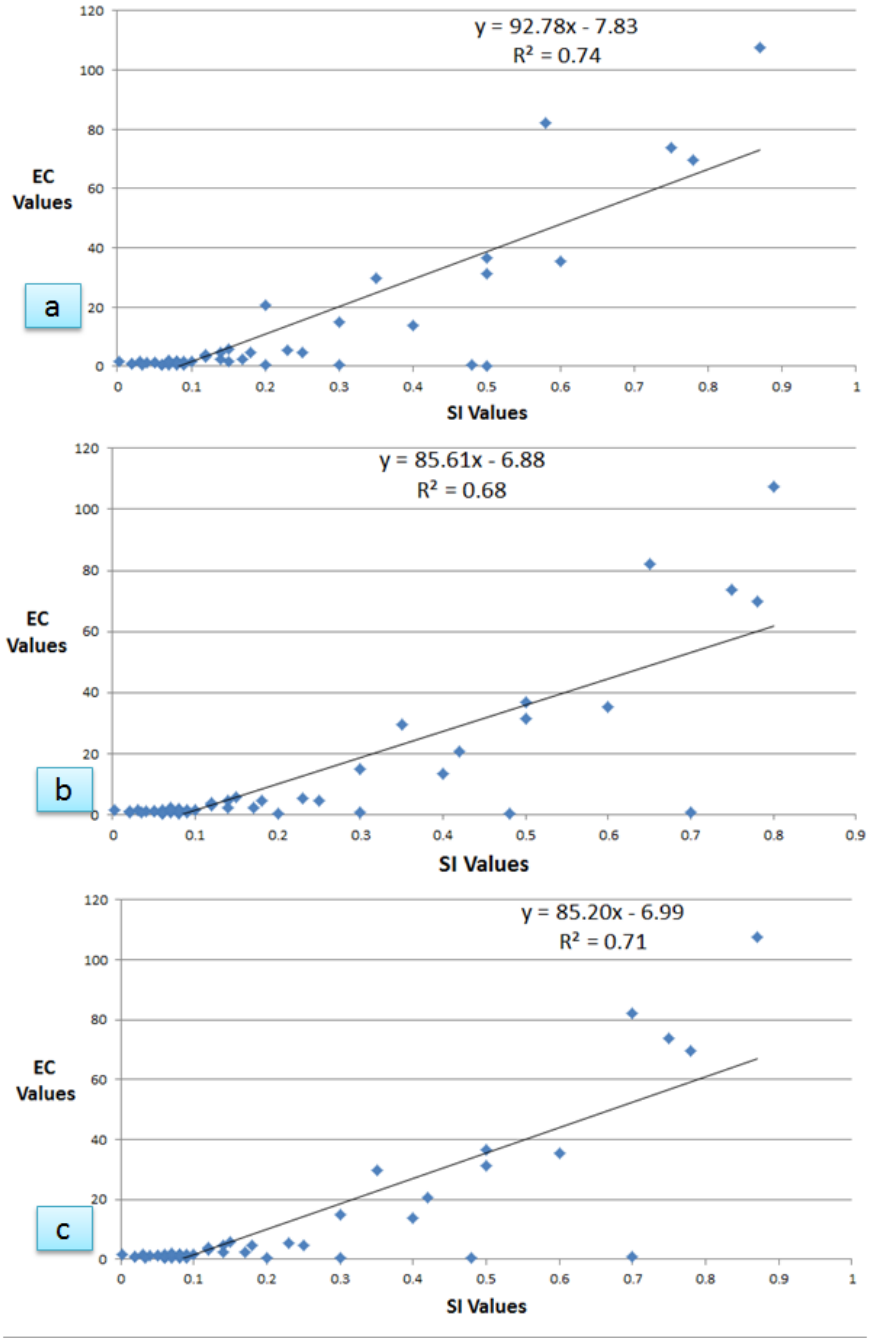

Figure 5.1 : Linear regression results between ground measurement EC values and SI values (a) new SI, (b) SI2, (c) SI1for Bonab Region.

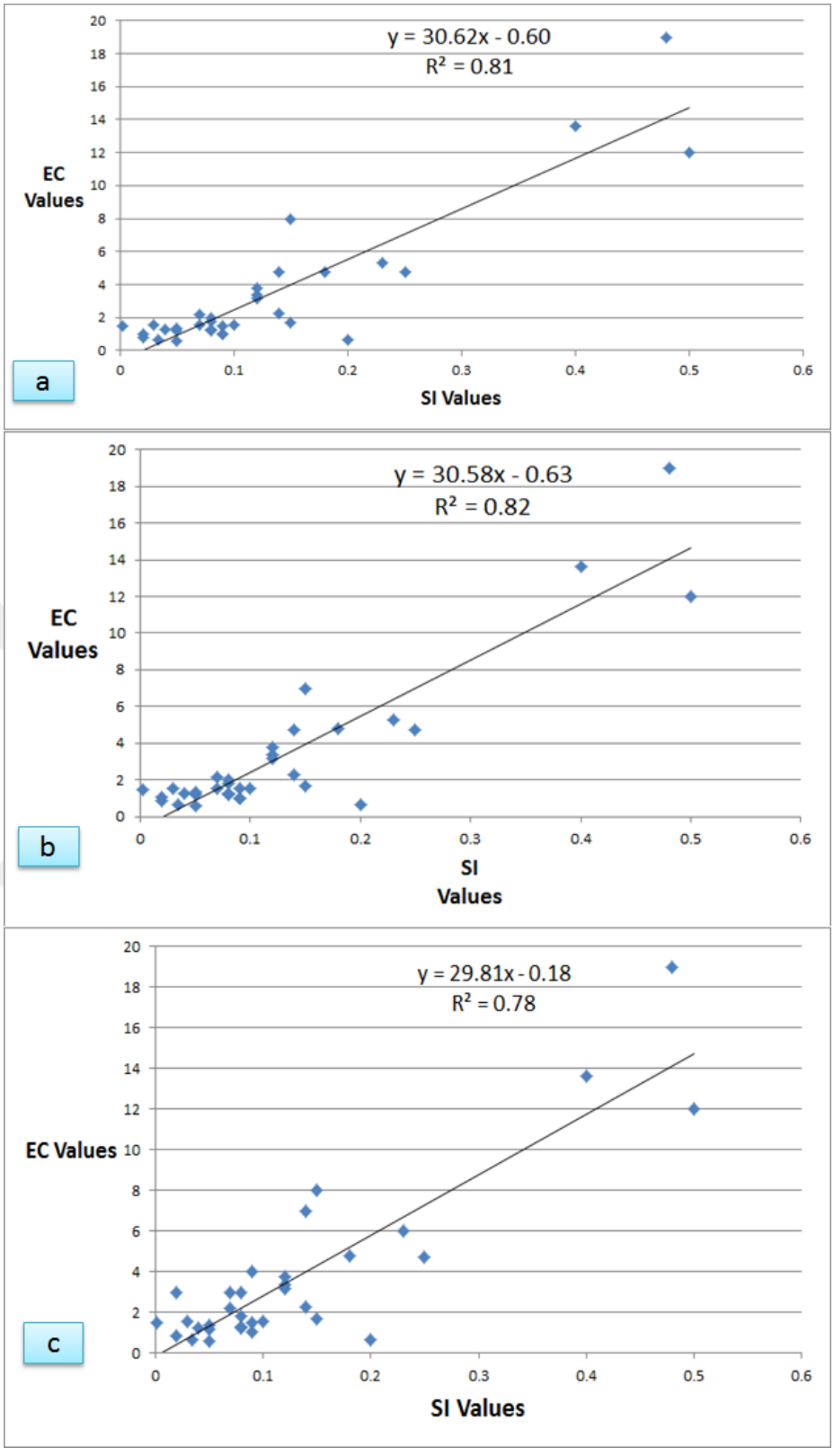

Figure 5.2 : Linear regression results between ground measurement EC values and SI values (a) new SI, (b) SI1, (c) SI2 for Tuz Lake.

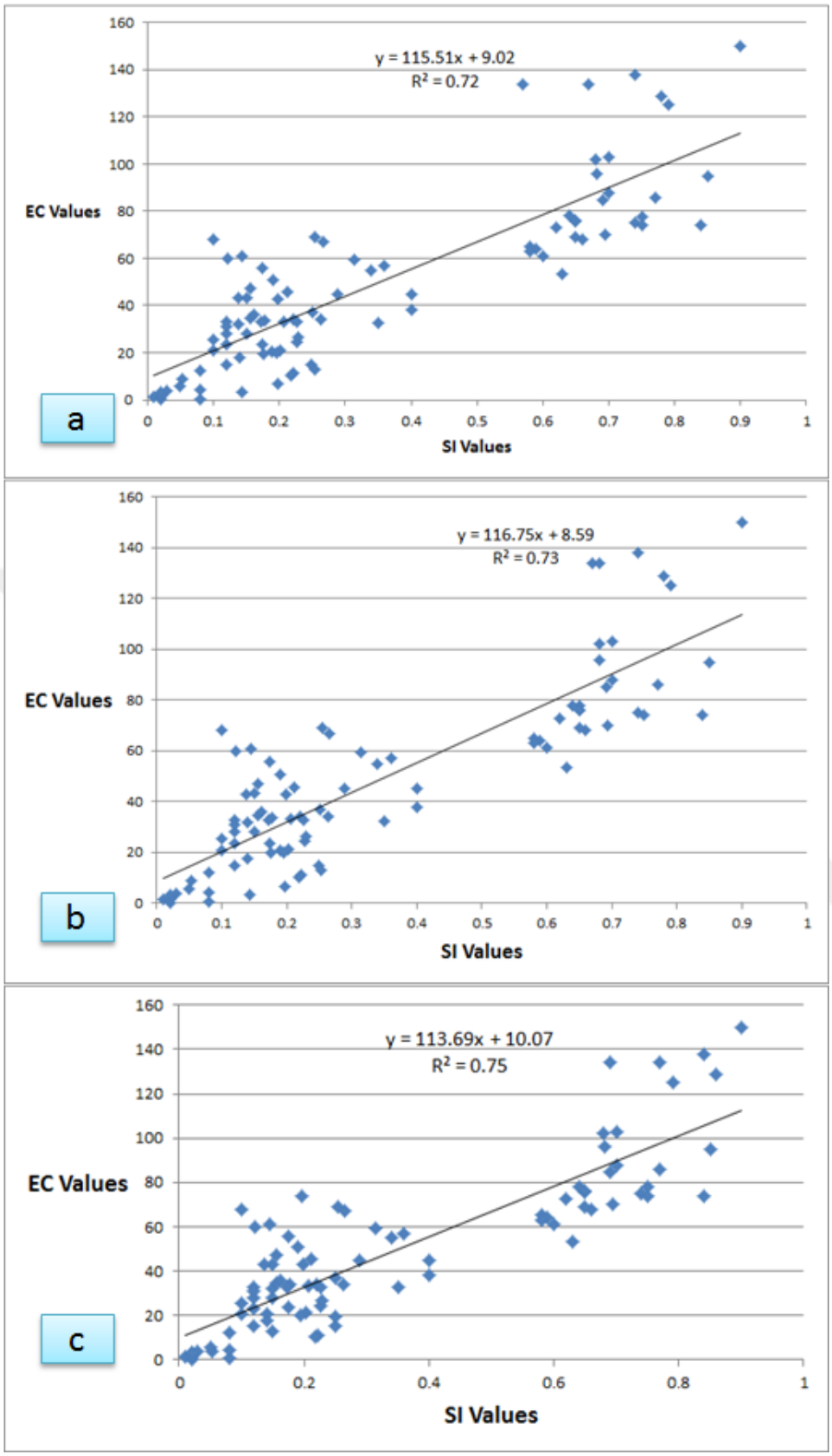

Figure 5.3 : Linear regression results between ground measurement EC values and SI values (a) SI1, (b) SI2, (c) new SI for Urmia Lake.

The scatter plots derived from regression analysis results indicated that most of the data are clustered in the SI value ranges between 0 to 0.5 for all the three indices specifically for Tuz Lake and Bonab case studies since most of samples were belong

to slightly and moderately saline soils. For each case study, R^2 values and the generated model equation are approximately similar to each other and due to this fact, boundaries of SI values for categorizing each salinity class are approximately similar for all three each indices.

Figure 5.4. depicts spatial distribution of different classes of salt-affected soils in the Bonab Region. Relatively, similar pattern can be seen in all the six maps generated. Extremity of soil salinity increased in soils that were located closer to the lake. All SI images illustrated non-saline -slightly saline soils where there were more agricultural lands placed.

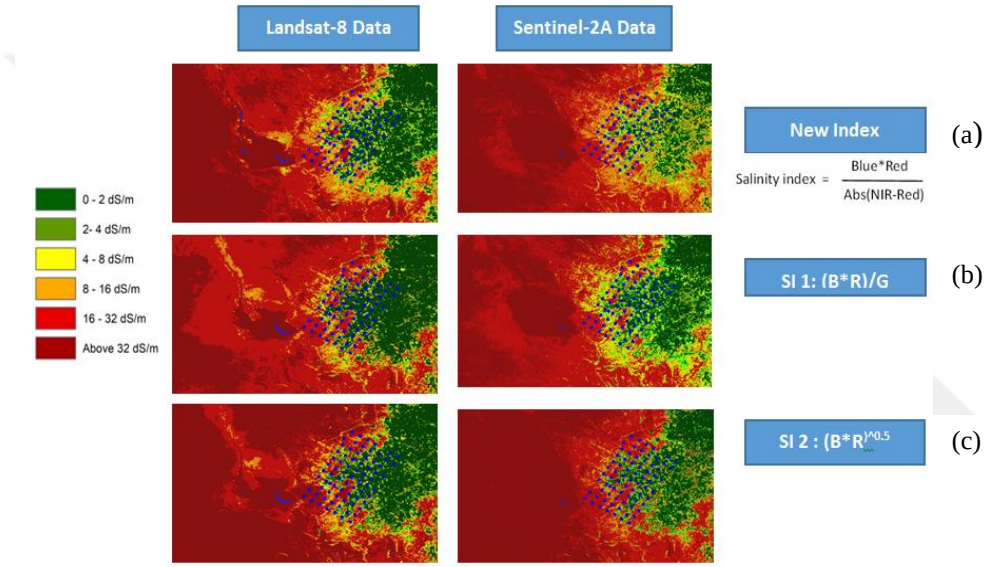

Figure 5.4 : Soil salinity maps generated by (a)-utilizing new generated soil salinity index, (b)-SI1and (c)-SI2 derived from Landsat-8 OLI and Sentinel-2A data in the Bonab Region.

Regarding to West playas of Urmia Lake Basin, Figure 5.5 shows soil salinity maps produced by Landsat-8 OLI data using the three salinity indices. All maps showed relatively similar results for non-saline–slightly saline classes. In addition, salt crusts located in the vicinity of the lake were detected clearly using all the three indices. Despite, there were differences in the performance of indices for categorizing moderately and highly salt-affected lands. SI1 detected extremely saline soils in the middle and below parts of urban areas; however, new index salinity map showed these regions with highly saline soils, and SI2 map illustrated moderately saline soils in these regions. Generally, the maps derived from Landsat-8 OLI data categorized boundaries of highly saline soils from extremely saline soils better than the maps

derived from Sentinel-2A data. Despite, the salinity maps derived from Sentinel-2A provided more detail information for slightly and moderatey saline soils since the spatial resolution of Sentinel-2A sensor outperform spatial reolution of Landsat-8 OLI sensor.

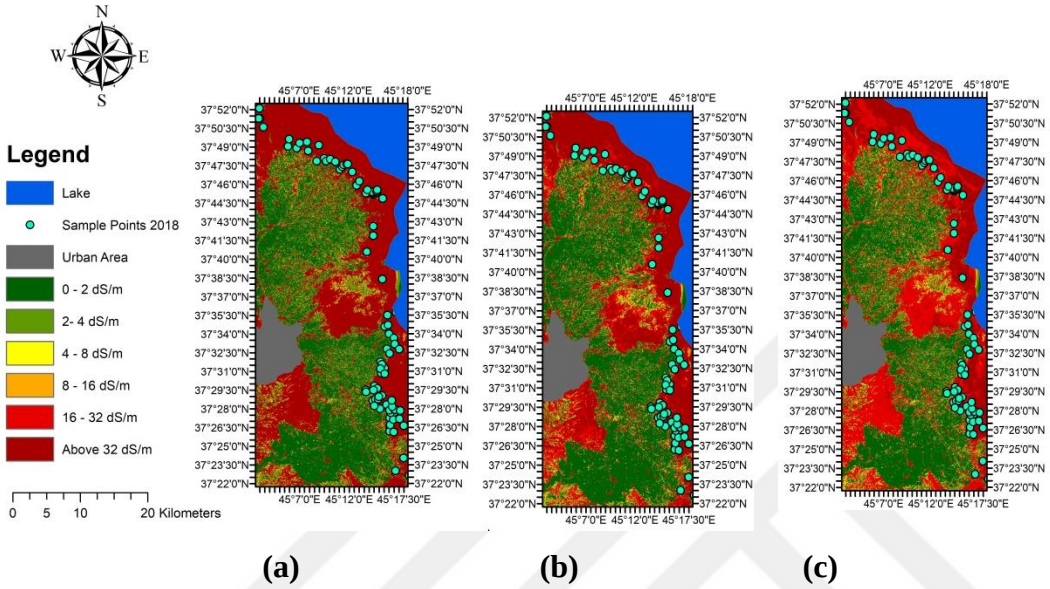

Figure 5.5 : Soil salinity maps generated by utilizing (a)- SI1, (b)- new generated soil salinity index, and (c)- SI2 index derived from Landsat-8 data in the West playas of Urmia Lake Basin.

Regarding to Tuz Lake Basin, Figure 5.6 illustrates soil salinity maps generated by applying SI1, SI2 and the new index. Since the range of EC ground measurement data varies between 0 to 25 dS/m for this case study, and as there is data deficiency for extremely saline soils, categorizing boundaries of extremely saline soils including ranges of (16-32 dS/m) and above 32 dS/m will be difficult. Moreover, most of the EC data for all 322 samples lay between 0 to 4 dS/m. In this research, 40 samples from all salinity classes that were placed in different geographical locations around the lake were selected in order to estimate the spatial distribution of saline soils relatively in an accurate way under data scarce condition. Soil salinity maps derived from SI2 and the new index depicts similar results. Despite the map derived from SI1 has inaccuracy in estimation of salinity level in the upper and middle parts of the lake. Generally, all maps show that there are much none- saline soils in the upper part of the lake in comparison with other places around the lake.

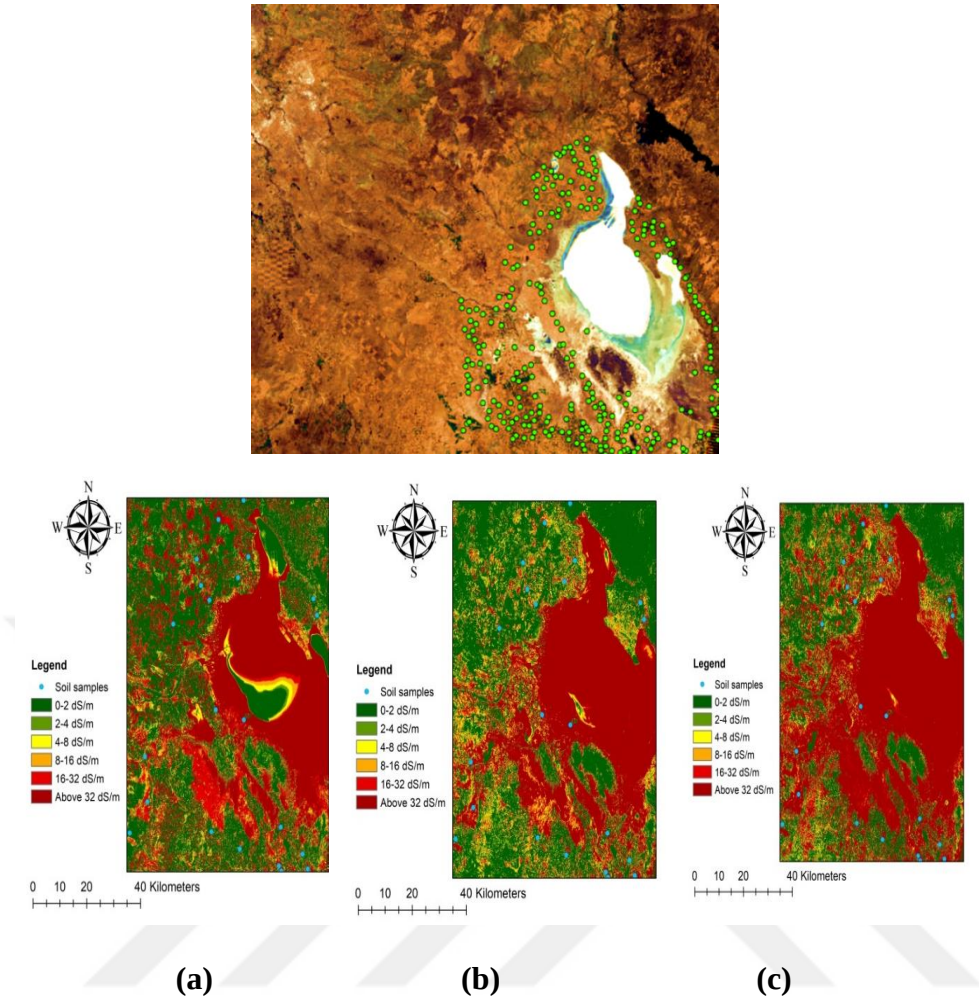

Figure 5.6 : Soil salinity maps generated by utilizing (a)- SI1, (b)- new generated soil salinity index, and (c)- SI2 index in Tuz Lake Basin.

After comparing capability of the new generated index with the two other commonly used salinity indices for producing soil salinity maps in the three different geographical locations, Cubist Model was used for generating more maps in order to evaluate the performance of the new index by means of another way. The output maps were generated as a result of adopting geo-statistical and RS data used as input variables in the model.

Since ordinary kriging raster data was generated according to interpolation analysis, the new map scales were resized to the boundary of the selected sample points. Output map and generated rules of Cubist Model using 3 salinity indices and 3 ordinary kriging raster data including (SAR, Ca, Mg) soil properties in the Bonab Region is depicted in Figure 5.7.

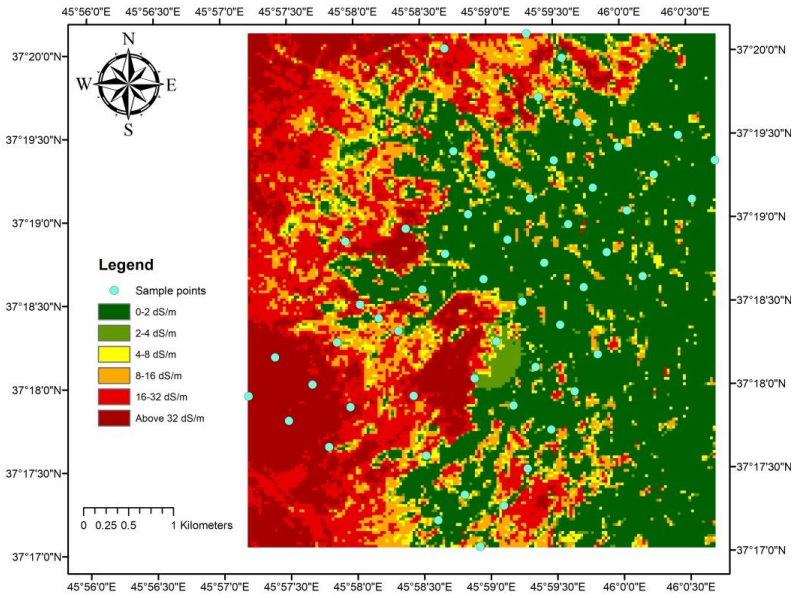

Figure 5.7 : Output map of Cubist model using 6 attributes including 3 SI and 3 ordinary kriging raster data (SAR, Ca, Mg) in the Bonab Region.

In this case, the Cubist Model utilized new index as a criteria parameter for generating the rules. The code in Figure 5.8 shows two different equations for EC prediction. Three attributes including (SAR, New SI and (Ca+Mg)) out of 6 input attributes were selected by Cubist model for generating the soil salinity map.

```
    Target attribute `EC'
Read 54 cases (6 attributes) from bonab.data
Model:
  Rule 1: [34 cases, mean 1.775, range 0.33 to 5.8, est err 1.055]
    if
        New SI <= 0.44
    then
        EC = 3.51 - 0.1 (Ca+Mg)
  Rule 2: [20 cases, mean 28.535, range 1.18 to 107.5, est err 11.441]
    if
        New SI > 0.44
    then
        EC = -29.523 + 60.8 New SI + 1.23 SAR
Evaluation on training data (54 cases):
    Average  |error|                3.721
    Relative |error|                0.25
    Correlation coefficient         0.94
        Attribute usage:
          Conds  Model
          100%    37%    New SI
                  63%    (Ca+Mg)
                  37%    SAR
```

Figure 5.8 : Rules of Cubist model using 6 attributes including 3 SI and 3 ordinary kriging raster data (SAR, Ca, Mg) in the Bonab Region.

The same approach was re-examined by adding one more attribute (ordinary kriging raster data for Na), and it is seen that it highly affected the results. Figure 5.9 which is the output map of Cubist model using 3 salinity indices and 4 ordinary kriging raster data including (Na, SAR, Ca, Mg) in the Bonab Region shows the study area as more saline in comparison with the previous map. That situation was mostly due to high Na values which are one of the significant indicators of soil salinity.

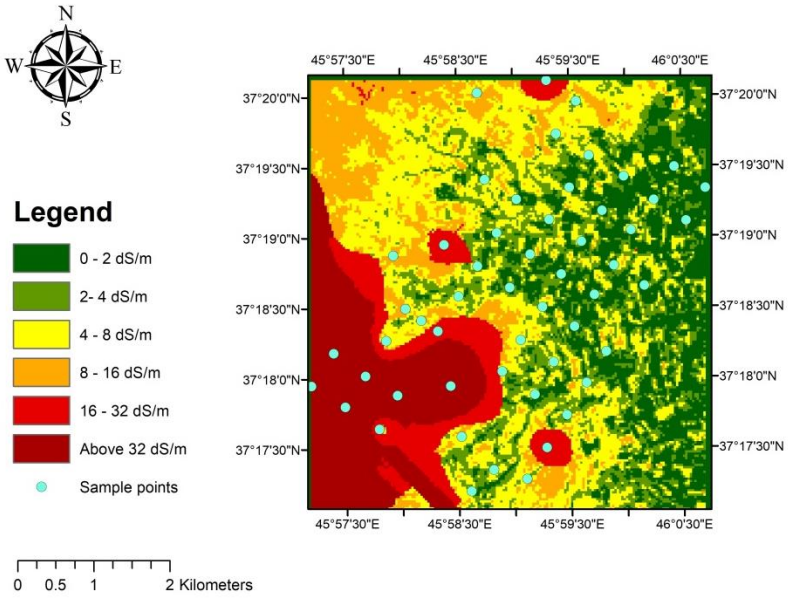

Figure 5.9 : Output map of Cubist model using 3 salinity indices and 4 ordinary kriging raster data including (Na, SAR, Ca, Mg) in the Bonab Region.

In this case Cubist model utilized Na as conditional parameter for generating the rules since there is high correlation between Na and EC parameters according to their ground measurement data.

The output map is highly impacted by Na raster kriging map since Na attribute is used both as conditional parameter and also as the main parameter in the equations.

Figure 5.10 illustrates the Rules of Cubist model using 3 salinity indices and 4 ordinary kriging raster data including (Na, SAR, Ca, Mg) in the Bonab Region.

```
    Target attribute `EC'

Read 54 cases (7 attributes) from bonab.data

Model:

  Rule 1: [44 cases, mean 2.767, range 0.33 to 18.56, est err 2.106]

    if
        Na <= 17.87
    then
        EC = 5.837 + 13.5 New SI + 0.035 Na + 0.24 SAR - 53 SI1

  Rule 2: [10 cases, mean 50.93, range 20.8 to 107.5, est err 20.843]

    if
        Na > 17.87
    then
        EC = 25.663 + 0.179 Na


Evaluation on training data (54 cases):

    Average  |error|                4.213
    Relative |error|                0.28
    Correlation coefficient         0.92


        Attribute usage:
          Conds  Model

          100%   100%    Na
                  81%    SAR
                  81%    New SI
                  81%    SI1
```

Figure 5.10 : Rules of Cubist model using 3 salinity indices and 4 ordinary kriging raster data including (Na, SAR, Ca, Mg) in the Bonab Region

Analysis in the Bonab Region continued by generating one more soil salinity map. This time geo-statistical data were excluded and only 8 remotely sensed data including 3 salinity indices, visible bands, NIR and SWI were used. In this investigation, the output result as shown in Figure 5.11 indicated that Cubist model selected 3 variables including the new index, green band and red band as the main attributes for generating the rules.

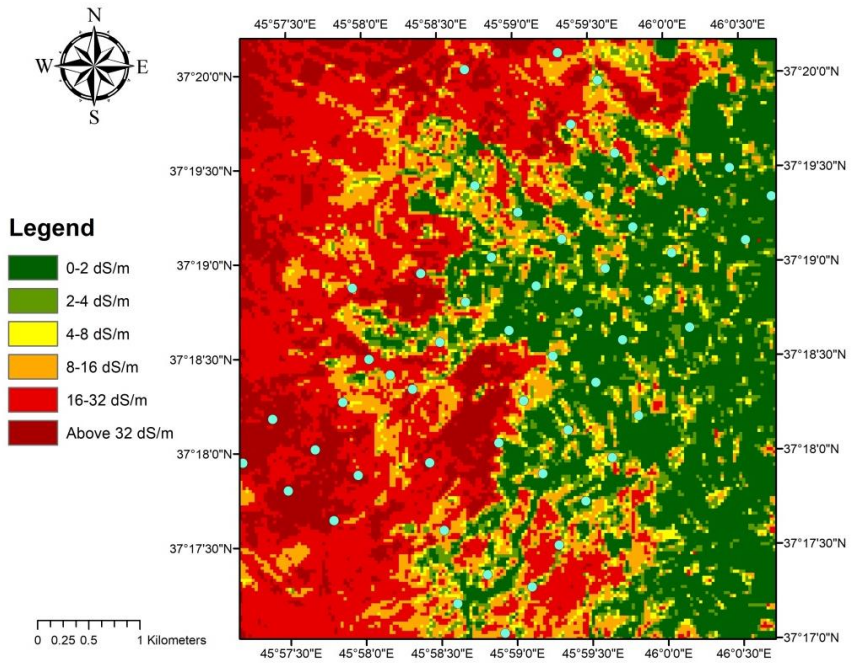

Figure 5.11 : Output map of Cubist model using 3 salinity indices and 5 spectral bands including (visible, NIR and SWI) in the Bonab Region.

Similar to the first analysis using Cubist model, in this case again, new SI is used as a conditional parameter for generating the rules. The condition for this case (New SI>0.44 and New SI<0.44) is exactly the same with the condition of the first analysis. This is due to the fact that excluding Na from the input attributes result in the selection of the new SI as conditional parameter, and also as the main parameter in EC estimation equations.

Figure 5.12 demonstrates rules of Cubist model using 3 salinity indices and 5 spectral bands including (visible, NIR and SWI) in the Bonab Region.

```
    Target attribute `EC'
Read 54 cases (8 attributes) from bonab.data
Model:
  Rule 1: [35 cases, mean 1.763, range 0.33 to 5.8, est err 0.989]
    if
        New SI <= 0.4467154
    then
        EC = -2.24 + 111 green - 74 red
  Rule 2: [19 cases, mean 29.965, range 1.18 to 107.5, est err 15.241]
    if
        New SI > 0.4467154
    then
        EC = -45.761 + 91 New SI
Evaluation on training data (54 cases):
    Average  |error|                4.881
    Relative |error|                0.33
    Correlation coefficient         0.90
        Attribute usage:
          Conds  Model
          100%    35%    New SI
                  65%    green
                  65%    red
```

Figure 5.12 : Rules of Cubist model using 3 salinity indices and 5 spectral bands including (visible, NIR and SWI) in the Bonab Region.

Regarding to West playas of Urmia Lake, several tests with various input attributes to Cubist model was performed. The most acceptable result is depicted in Figure 5.13. It is shown that the output soil salinity map that was generated by Cubist Model using only 7 RS data including 3 indices together with visible and NIR bands as input attributes. In this analysis, the output result indicated that Cubist selected 6 variables including 3 indices and visible bands for generating the rules.

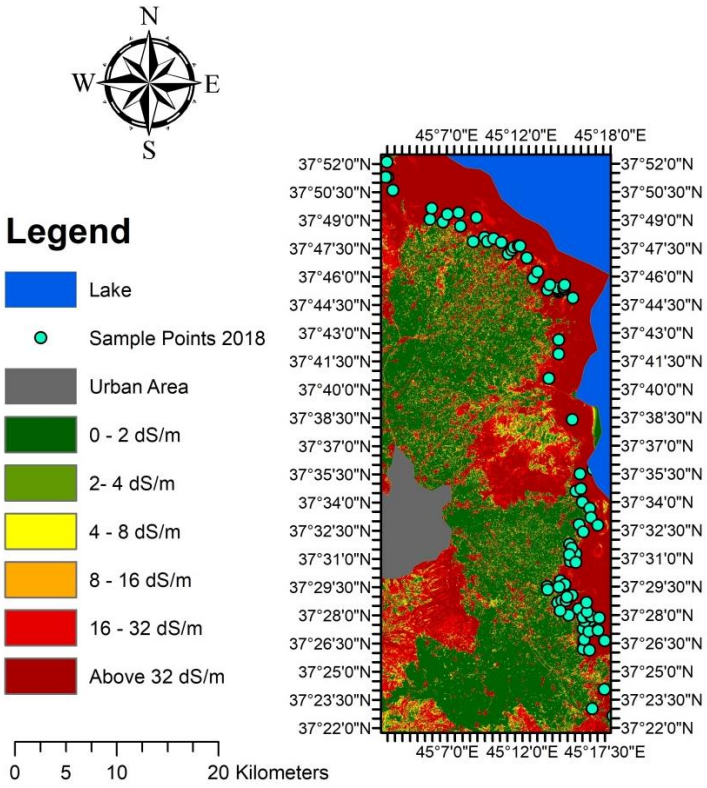

Figure 5.13 : Output map of Cubist model using 3 salinity indices and Landsat-8 OLI visible and NIR bands in the West playas of Urmia Lake.

In this case, the output soil salinity map categorized different salinity levels similar to the output map derived from the simple linear regression model run between new index and EC data. Cubist analysis using 7 attributes generate only one equation for EC calculation. In some cases, the model does not select any conditional parameter for generating rules, and the estimation will be similar to multiple regression analysis by utilizing only one equation.

In this analysis, Cubist model applied all input variables except NIR data for generating the equation; since, it could not find relevant correlation between NIR spectral data and EC values.

Figure 5.14 shows rules of Cubist model using 3 salinity indices and Landsat-8 OLI visible and NIR bands in the West playas of Urmia Lake.

```
    Target attribute `EC'

Read 92 cases (7 attributes) from urmiawest.data

Model:

  Rule 1: [92 cases, mean 40.75, range 0.3 to 143, est err 17.93]

        EC = 45.91 + 7654 SI2 - 3111 SI1 - 2663 Green - 1127 Red - 823 Blue
             + 20.4 New SI


Evaluation on training data (92 cases):

    Average  |error|                15.40
    Relative |error|                 0.58
    Correlation coefficient          0.81


        Attribute usage:
          Conds  Model

                 100%    Blue
                 100%    Green
                 100%    Red
                 100%    New SI
                 100%    SI1
                 100%    SI2
```

Figure 5.14 : Rules of Cubist model using 3 salinity indices and Landsat-8 OLI visible and NIR bands in the West playas of Urmia Lake.

Further analysis was performed by combining geo-statistical data with RS data in the Cubist Model. As it is shown in Figure 5.15, these maps were not reliable as interpolation analysis for these widely distributed sample points did not work well.

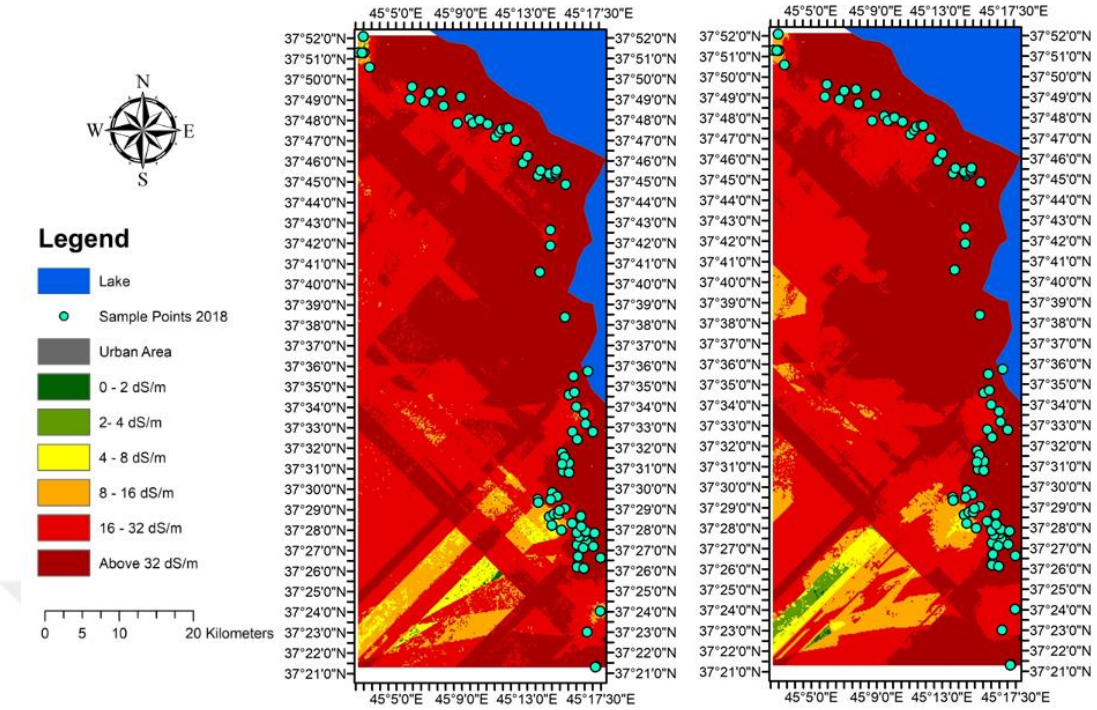

Figure 5.15 : Output map of Cubist model using 3 salinity indices and geo-statistical data in the West playas of Urmia Lake.

As a solution, a set of sample points that were located in one homogeneous region (samples that were distributed logically in the vicinity of each other) were selected, and then, new maps with new scales were generated as shown in Figure 5.16.

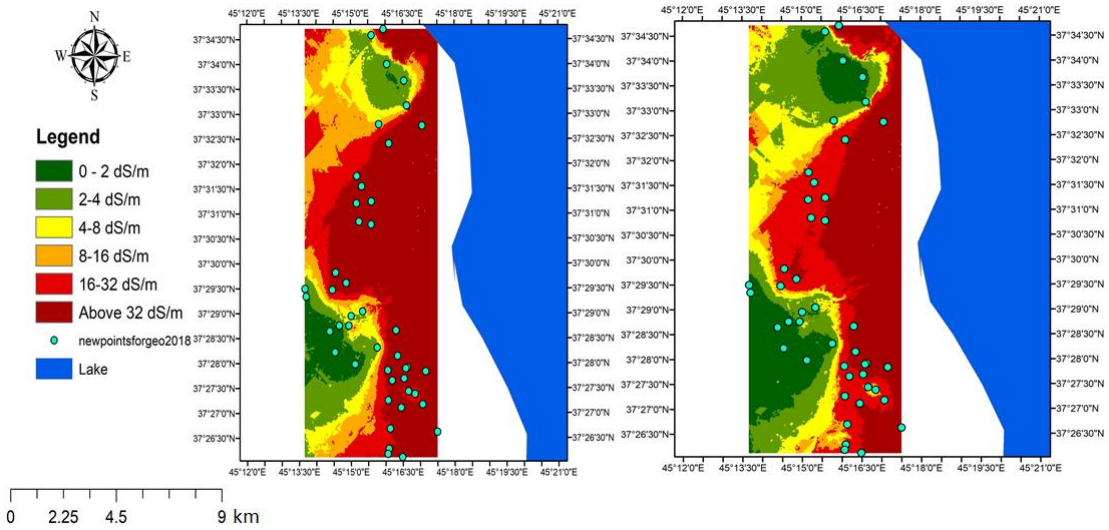

Figure 5.16 : Output map of Cubist model using 3 salinity indices and different ordinary kriging raster data of soil parameters in the West playas of Urmia Lake.

In terms of Tuz Lake Basin, two further analyses were performed by using the Cubist model. Firstly, 3 soil salinity indices and band blue as 4 RS input data were adopted in the model. Figure 5.17 and Figure 5.18 illustrate the output map and the rules that is generated by Cubist Model, respectively. In this analysis, Cubist selected 3 indices for generating the rules, and among them, new SI was chosen as defining the criteria of the rules.

The output soil salinity maps that were generated by a set of sample points including those which were located in one homogeneous region, estimated different salinity levels more accurately than the previous analysis. This is due to the fact that raster data derived from samples in a homogeneous region provided relevant prediction on spatial distribution of saline soils in the region.

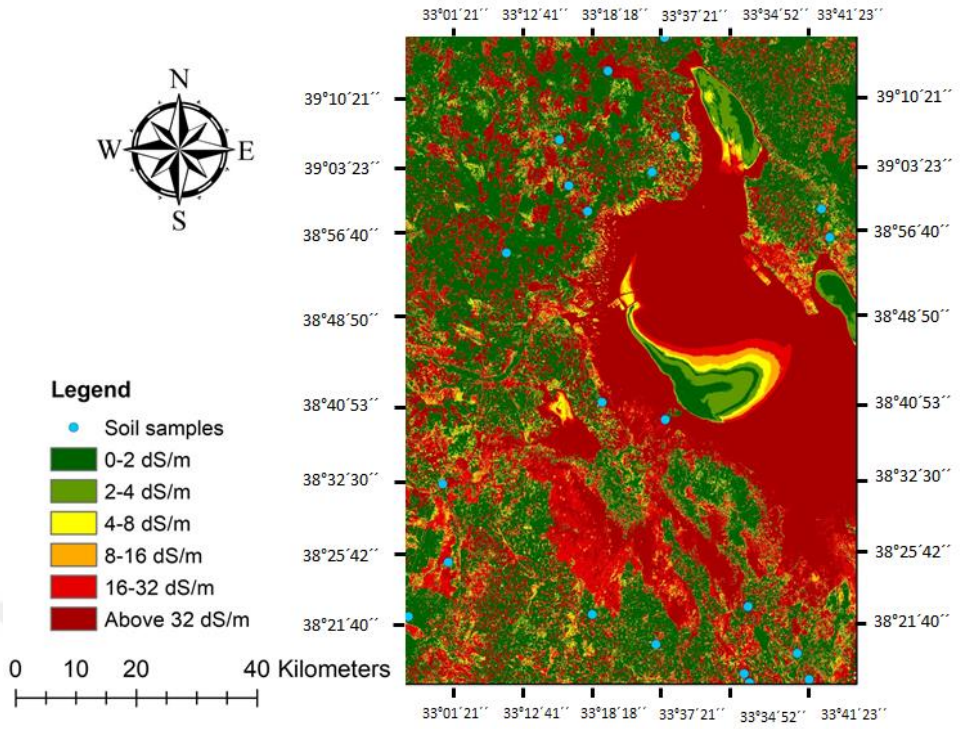

Figure 5.17 : Output map of Cubist model using 3 salinity indices and blue band in the Tuz Lake Basin.

In this analysis, new index is applied as a conditional attribute for generating the rules and all three indices were selected by Cubist model as variables for estimating the EC value in two different equations.

```
    Target attribute `EC'
Read 24 cases (4 attributes) from Tuzlake.data

Model:

  Rule 1: [6 cases, mean 13.4583, range 0.46 to 31.47, est err 2.0654]

    if
        NewSI > 0.3611808
    then
        EC = -21.2357 + 82.8 NewSI - 31.8 SI1 + 5.5 SI2

  Rule 2: [18 cases, mean 1.6697, range 0.45 to 5.4, est err 1.0478]

    if
        NewSI <= 0.3611808
    then
        EC = -1.9435 + 28.3 SI2 - 19 SI1 + 12.2 NewSI


Evaluation on training data (24 cases):

    Average  |error|                0.8753
    Relative |error|                  0.17
    Correlation coefficient          0.99


        Attribute usage:
          Conds  Model

          100%   100%    NewSI
                 100%    SI1
                 100%    SI2
```

Figure 5.18 : Rules of Cubist model using 3 salinity indices and blue band in the Tuz Lake Basin.

In the second analysis, combination of RS data (3 salinity index images) and geo-statistical data (ordinary kriging raster data of other soil properties) were adopted as 6 different attributes to Cubist model. Figure 5.19 is output map of Cubist model using 3 RS data and 3 ordinary kriging raster data in the Tuz Lake Basin.

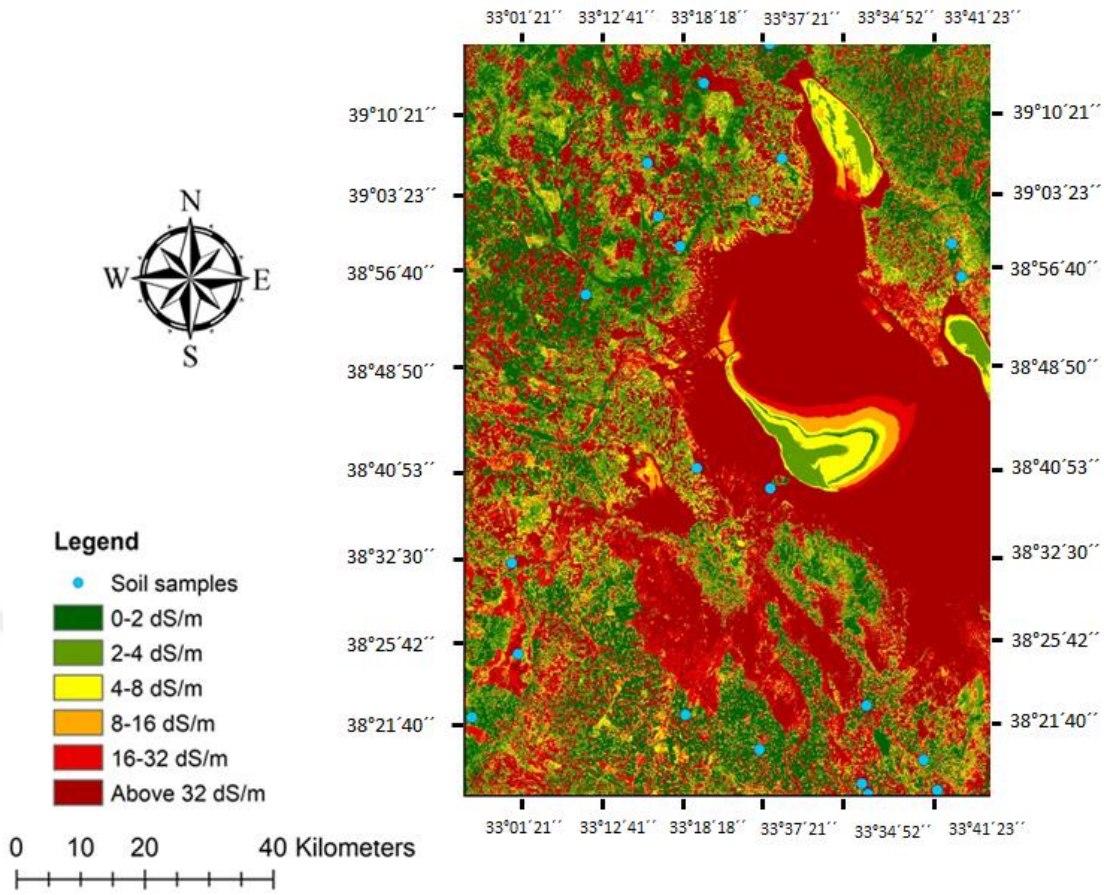

Figure 5.19 : Output map of Cubist model using 3 RS data and 3 ordinary kriging raster data in the Tuz Lake Basin.

Adding 3 more input data to Cubist model changed the attributes for calculating the EC value, but not the conditional variable that was applied for generating the rules. In this analysis, again new index is selected with same criteria (New SI>0.36 , New SI<0.36) as a conditional parameter.

In this investigation, as it is illustrated in Figure 5.20, new SI and SI1 as RS data, and ordinary kriging raster data of ESP as a soil property were selected by the model for generating the map.

```
Read 24 cases (6 attributes) from Tuzlake.data

Model:

  Rule 1: [6 cases, mean 13.4583, range 0.46 to 31.47, est err 1.7021]

    if
        NewSI > 0.3611808
    then
        EC = -19.8722 + 82.3 NewSI - 31.3 SI1 + 0.034 ESP

  Rule 2: [18 cases, mean 1.6697, range 0.45 to 5.4, est err 0.9914]

    if
        NewSI <= 0.3611808
    then
        EC = 0.4984 + 76.6 NewSI - 69.9 SI1 + 0.225 ESP


Evaluation on training data (24 cases):

    Average  |error|                0.6150
    Relative |error|                  0.12
    Correlation coefficient          0.99


        Attribute usage:
          Conds  Model

          100%   100%    NewSI
                 100%    SI1
                 100%    ESP
```

Figure 5.20 : Output map of Cubist model using 3 RS data and 3 ordinary kriging raster data in the Tuz Lake Basin.

6. CONCLUSIONS AND RECOMMENDATIONS

Due to high spatial and temporal variability of soil salinity in especially the semi-arid and arid regions of the world, it can be considered as one of the major environmental issues accelerating the land degradation processes. As a result, it provokes severe economic and social consequences in the affected regions. Rapid population growth and insufficient arable lands for producing adequate amount of food and fabric to humans force many nations to apply optimized solutions and methods for diminishing adverse effects of soil salinity. Despite, this phenomenon is not as debatable as environmental problems such as climate change, global warming, air pollution, water pollution and deforestation among scientists and decision makers; but, it should not be underestimated. If soil salinization expands in future with this pace, several countries will most probably experience difficulties in supplying adequate amount of food for their population.

Recently, many researches focused on differentiating salinized soil and non-salinized soil, qualitatively analyzing the spatial distribution of soil salinity and monitoring the dynamics of soil salinity. RS, GIS, machine-learning and modelling have become the commonly used technological tools to map soil salinity due to their rapid and economic coverage of large area which is extremely significant both from the agricultural and environmental perspectives. In this research, exploring main reasons of soil salinization in various case studies demonstrates that both natural and anthropogenic factors can negatively impact the soil fertility. Evaluating researches which are conducted especially in the arid and semi-arid regions of the world indicates that low precipitation rates and excessive evapotranspiration combined with the presence of parent material and minerals including salt cations and anions are the most general natural sources of soil salinization. Despite, those places which are located in the vicinity of coastal zones, salty seawater intrusion is considered as the main parameter for soil salinization. Regarding secondary salinization, it is investigated that old-fashioned (traditional) farming methods, poor agricultural irrigation practices and

irrelevant drainage systems result in soil salinization, and also cause rising of groundwater table. In addition, by reviewing and summarizing various case studies related to soil salinization mapping techniques between years 2014 to 2021, advances of several RS techniques , modeling approaches, soil salinity indices and geo-statistical methods for detecting salt-affected lands were assessed and then 3 case studies were selected for comparing the performance of a new generated soil salinity index with two other commonly used salinity indices and examining their capability for detecting salt-affected soils. In addition, Cubist model was utilized for generating more accurate maps since all the 3 indices along with visible and NIR spectral bands were adopted as input data to this model. Series of analysis indicated that the new generated index has capability of detecting salt-affected soil in the selected case study areas and the salinity maps produced by the new index were as acceptable as those maps produced by the other two indices. Therefore by this study, a promising new index will be introduced to the scientific platform on soil salinity detection and mapping. Despite, in the literature several soil salinity indices have already been derived from visible bands of multispectral data including Landsat series, Sentinel, IKONIS, Spot series, MODIS and other sensors. According to the literature review that was completed in this survey, we reach to this fact that selecting an index for a case study is significantly dependent on the characteristics of the study area and there is no fixed/single salinity index that demonstrate best results for all geographical locations.

Overview of various case studies depicts that RS approaches are not sufficiently applicable to detect surface salt-affected soils in the absence of white salt crust, and also it is not possible to estimate the salinity level in the root zone. In order to overcome this limitation, procedures for validation and calibration of RS data with ground truth soil salinity data should be defined precisely. Moreover, developing guidelines on harmonized standard approach for ground measurement, mapping and monitoring soil salinity can result in more reliable predictions.

There are always many challenges and problems for producing soil salinity map in various researches, specifically those that are conducted in the developing countries. Lack of required input data at required scale, frequent lack of soil and geological maps, inadequacy of knowledge about irrigation water quality and ground water level, lack of information on correlated environmental parameters are all main challenges of

soil salinity modeling in developing countries. Simply, we may call this situation as data scarce condition.

Professional soil salinity modeling in field scale requires calculating water and salt balance by developing a model for water dynamics and solute transport. Despite, in both local and regional scale mapping, RS data contributes more in comparison with field scale mapping. In large scale studies, identification of the most appropriate spectral bands and then examining their combinations to generate a robust spectral index can be supportive for better estimation of salt-affected lands. Moreover, developing approaches for automatic processing and obtaining information from long-term RS data via google earth engine (GEE) and machine-learning techniques can enhance the accuracy and speed of analysis.

In addition, integrating multiple source data including RS data, terrain attributes extracted from digital elevation model (DEM), land use information, meteorological data, water quality and soil data are essential for accurate soil salinity modeling in critical arid regions of the globe considering different climate change scenarios.

In terms of RS restrains for soil salinity mapping, radiometric measurements and ground observations depicts that abundance and diversity of salts, along with soil moisture, soil smoothness or roughness and soil color are the major parameters affecting salt reflectance. Soil surface smoothness or roughness is affected by salt minerals. Non-saline cultivated lands are commonly rougher than salty crusts, and due to this fact, there is lower spectral reflectance in the visible and NIR bands for non-saline agricultural lands in comparison with salt crusts.

Spectral contrast with other surface elements, soil moisture content, salt mineralogy and covering percentage are among major factors affecting considerably identification of saline soils via optical RS. Indeed, salt crusts and highly saline regions are simply and quickly identifiable, despite low salinity levels and initial stages of salinization regularly cause detection inaccuracies.

Appropriate timing for soil salinity detection via RS data is highly significant. Literature review demonstrates that most of the researches have obtained the related RS data and EC ground measurement data at the end of the dry season, as salts dissolve during the rainy season. Good timing for passive remote sensing data acquisition must take into account that salt identification is easiest at the end of the dry season, as salts dissolve during the rainy season.

Shifting from simple spatial detection of salt-affected areas to the monitoring of seasonal changes in salinity necessitates minimum two identification dates along a given time span. As accumulated salts on surface soil strongly change with seasons, time series of RS data should be obtained in the same periods of the year to be comparable, preferably at the end of the dry season, if passive remote sensors are used. Moreover, geo-referencing and co-registration of multi temporal data are necessary to provide possibility of tracking ground sites over time and enable researchers to compare RS data with ancillary data.

Recent soil salinity mapping and prediction methods can benefit from the synergistic usage of airborne and satellite-borne sensors covering different regions of the spectrum. The reclamation, rehabilitation, and monitoring of saline lands need not only knowledge about mineralogy and variety of salts, but also require severity levels of salinity. This cannot be accomplished from RS data alone and needs complete integration of remote sensing, ground, and laboratory data.

In soil salinity studies, it is the analyst challenge to recognize the most appropriate salinity indicators for a specific area, so that the relevant ground measurements and remote sensing techniques can be applied to achieve information in an accurate, rapid and economical manner. For better soil salinity assessment in a specific area, researchers can utilize data on rainfall, topography, soil type and any other accessible spatial information to determine the locations that are most vulnerable to salinization, and then predict similar regions that may be at risk.

The key to favorably estimate the soil salt content using spectral variables is to select an appropriate mathematical regression model. Algorithms such as MLR, PLSR, and BP neural networks have been commonly used in the inversion and modeling of soil component contents. Despite, in the recent studies, application of machine-learning methods has provided the ability of autonomous learning and can solve the problem of complex nonlinear function approximation in soil salinization monitoring.

Modern approaches are introducing h new promotions to guarantee effective, appropriate, and sustainable saline soil management. Among the available management methods, some including crop selection, genetic enhancement, agroforestry and seed preparation focus on improving plant properties condition. The other approaches such as application of chemical amendments, biochar, fertilizers and vermicomposting contribute for the improvement of soil properties.

Generating a new soil salinity index and examining various models for soil salinity assessment requires periodical and relevant sampling from different places of each case study. This research had some limitation for ground data sampling. Further site studies are suggested for providing more data in order to better evaluate the performance of the new index that has been developed within the context of this thesis in future. Another limitation of this research linked with lack of accessibility to environmental factors of each case study. Deep study and knowledge about environmental characteristics of each geographical location may contribute producing better covariates and parameters for modeling and predicting EC values.

By this study, a promising new index will be introduced to the scientific platform on soil salinity detection and mapping. Despite, in the literature several soil salinity indices derived from visible bands of multispectral data including Landsat series, Sentinel, IKONIS, Spot series, MODIS and other sensors.

According to the literature review that was completed in this survey, we reach to this fact that selecting an index for a case study is significantly dependent to the characteristics of the study area and there is no salinity index that demonstrate best results for all geographical locations.

It is recommended to determine relevant policies and practices in future for monitoring and finding all natural causes of soil salinization in both Urmia Lake and Tuz Lake in order to lessen the adverse effects of soil salinization. These projects can contain periodical and disciplined tracking of soil salinity changes in the vicinity of lakes, seasonal field sampling and correspondingly soil salinity mapping in the area of collected samples by examining new machine-learning and modeling approaches, setting plans for cultivating halophilic plants in lands which are extremely saline since it leads to reclamation of saline soils. Shifting application of traditional irrigation approaches to modern irrigation techniques and establishing convenient drainage systems, might be another practice for diminishing negative effects of traditional methods such as flooding irrigation which makes soils more saline.

REFERENCES

Abbas, A., Khan, S., Hussain, N., Hanjra, M. a., & Akbar, S. (2013). Characterizing soil salinity in irrigated agriculture using a remote sensing approach. *Physics and Chemistry of the Earth, Parts A/B/C*, *55*: 43–52.

Abdullah AYM, Biswas RK, Chowdhury AI, Billah SM. (2019) Modeling soil salinity using direct and indirect measurement techniques: A comparative analysis. *Environ Dev 29*:67–80.

Abuelgasim A, Ammad R. (2019) Mapping soil salinity in arid and semi-arid regions using Landsat-8 OLI OLI satellite data. *Remote Sens Appl Soc Environ 13*:415–425.

Allbed, A. and Kumar, L. (2013). Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: a review. *Advances in Remote Sensing 02*(04):373–85.

Celleri C, Zapperi G, González Trilla G, Pratolongo P. (2019) Assessing the capability of broadband indices derived from Landsat-8 OLI Operational Land Imager to monitor above ground biomass and salinity in semiarid saline environments of the Bahía Blanca Estuary, Argentina. *Int J Remote Sens 40*:4817–4838.

Chen H, Zhao G, Li Y, Wang D, Ma Y. (2019) Monitoring the seasonal dynamics of soil salinization in the Yellow River delta of China using Landsat data. *Nat Hazards Earth Syst Sci 19*:1499–1508.

Chi Y, Sun J, Liu W, Wang J, Zhao M. (2019) Mapping coastal wetland soil salinity in different seasons using an improved comprehensive land surface factor system. *Ecol Indic 107*:105517.

Davis E, Wang C, Dow K. (2019) Comparing Sentinel-2 MSI and Landsat-8 OLI OLI in soil salinity detection: a case study of agricultural lands in coastal North Carolina. *Int J Remote Sens 40*:6134–6153.

Dong F, Tang Y, Xing X, Liu Z, Xing L. (2019) Formation and evolution of soil salinization in Shouguang city based on PMS and OLI/TM sensors. *Water (Switzerland) 11*:1–15.

Didi S, Housni FE, Bracamontes del Toro H, Najine A. (2019) Mapping of soil salinity using the Landsat-8 OLI Image and direct field measurements: a case study of the Tadla Plain, Morocco. *J Indian Soc Remote Sens 47*:1235–1243.

El hafyani M, Essahlaoui A, El baghdadi M, Teodoro AC, Mohajane M, El hmaidi A, El ouali A. (2019) Modeling and mapping of soil salinity in Tafilalet plain (Morocco). *Arab J Geosci 12*: 231-236.

Elnaggar, A. a., & Noller, J. S. (2009). Application of Remote-sensing Data and Decision-Tree Analysis to Mapping Salt-Affected Soils over Large Areas. *Remote Sensing*, *2*(1), 151–165.

Eisele, A. Chabrillat, S. Hecker, C. Hewson, R. Lau ,I. Rogass, Ch. Segl, K. John, T. Udelhoven, T. Hostert, P. and Kaufmann, H. (2015). Advantages Using the Thermal Infrared (TIR) to Detect and Quantify Semi-Arid Soil Properties. *Remote Sensing of Environment163*:296–311.

Elhag, M. and Jarbou, A. Bahrawi. (2017). Soil Salinity Mapping and Hydrological Drought Indices Assessment in Arid Environments Based on Remote Sensing Techniques. Geosci. Instrum. Method. *Data Syst. 6*: 149–158.

Fathizad H, Ali Hakimzadeh Ardakani M, Sodaiezadeh H, Kerry R, Taghizadeh-Mehrjardi R. (2020) Investigation of the spatial and temporal variation of soil salinity using random forests in the central desert of Iran. *Geoderma 365*:114233.

Fazeli Sangani M, Namdar Khojasteh D, Owens G. (2019) Dataset characteristics influence the performance of different interpolation methods for soil salinity spatial mapping. *Environ Monit Assess 191*: 124-129.

Fernández-Buces, N., Siebe, C., Cram, S., & Palacio, J. L. (2006). Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former lake Texcoco, Mexico. *Journal of Arid Environments*, *65*(4), 644–667.

Gorji, T. Sertel, E. and Tanik, A. (2017). Monitoring Soil Salinity via Remote Sensing Technology under Data Scarce Conditions :A Case Study from Turkey. *Ecological Indicators 74*:384–91.

Gorji, T. Sertel, E. and Tanik, A. (2015). Soil Salinity Prediction, Monitoring and Mapping Using Modern Technologies. *Procedia Earth and Planetary Science 15*:507–12.

Gorji, T., Yıldırım, A., Hamzehpour, N., Tanık, A., Sertel, E. (2020). Soil Salinity Analysis of Urmia Lake Basin using Landsat-8 OLI and Sentinel-2A based Spectral Indices and Electrical Conductivity Measurements, *Ecological Indicators, 112* (106173).

Gorji, T., Yıldırım, A., Tanık, A., Sertel, E. (2019). Remote sensing approaches and mapping methods for monitoring soil salinity under different climate regimes, *International Journal of Environment and Geoinformatics (IJEGEO), 6*(1):33-49.

Guo B, Han B, Yang F, Fan Y, Jiang L, Chen S, Yang W, Gong R, Liang T. (2019a) Salinization information extraction model based on VI–SI feature space combinations in the Yellow River Delta based on Landsat-8 OLI image. Geomatics, *Nat Hazards Risk 10*:1863–1878.

Guo B, Yang F, Fan Y, Han B, Chen S, Yang W. (2019b) Dynamic monitoring of soil salinization in Yellow River Delta utilizing MSAVI–SI feature space models with Landsat images. *Environ Earth Sci 78*:1–10.

Guo B, Yang F, Han B, Fan Y, Chen S, Yang W, Jiang L .(2019c) A model for the rapid monitoring of soil salinization in the Yellow River Delta using Landsat-8 OLI imagery based on VI-SI feature space. *Remote Sens Lett 10*:796–805.

Guo Y, Zhou Y, Zhou l qing, Liu T, Wang l gang, Ceng Y Zheng, HE J, Zeng qing. (2019) Using proximal sensor data for soil salinity management and mapping. *J Integr Agric 18*:340–349.

Han L, Liu D, Cheng G, Zhang G, Wang L. (2019) Spatial distribution and genesis of salt on the saline playa at Qehan Lake, Inner Mongolia, China. *Catena, 177*:22–30.

Habibi V, Ahmadi H, Jafari M, Moeini A. (2021) Mapping soil salinity using a combined spectral and topographical indices with artificial neural network. *PLOS ONE, 16*(5): e0228494.

Hassani A, Azapagic A, Shokri N. (2020) Predicting long-term dynamics of soil salinity and sodicity on a global scale. *Earth atmospheric and planetary sciences, 117*(52):33017-33027.

Hoa PV, Giang NV, Binh NA, Hai LVH, Pham TD, Hasanlou M, Bui DT. (2019) Soil salinity mapping using SAR Sentinel-1 data and advanced machine-learning algorithms: A case study at Ben Tre Province of the Mekong River Delta (Vietnam). *Remote Sens 11*:1–21.

Hu J, Peng J, Zhou Y, Xu D, Zhao R, Jiang Q, Fu T, Wang F, Shi Z. (2019) Quantitative estimation of soil salinity using UAV-borne hyperspectral and satellite multispectral images. *Remote Sens .11*(7), 736.

Huang Y, Li Z, Ye H, Zhang S, Zhuo Z, Xing A, Huang Y. (2019) Mapping soil electrical conductivity using ordinary kriging combined with bBack-propagation network. *Chinese Geogr Sci 29*:270–282.

Ivushkin K, Bartholomeus H, Bregt AK, Pulatov A, Kempen B, de Sousa L. (2019) Global mapping of soil salinity change. *Remote Sens Environ 231*:111260.

Jiang H, Shu H. (2019) Optical remote-sensing data based research on detecting soil salinity at different depth in an arid-area oasis, Xinjiang, China. *Earth Sci Informatics 12*:43–56.

Kahaer Y, Tashpolat N. (2019) Estimating Salt concentrations based on optimized spectral indices in soils with regional heterogeneity. *Journal of Spectroscopy Article ID 2402749*, 15 pages.

Kasim N, Maihemuti B, Sawut R, Abliz A, Dong C, Abdumutallip M. (2020) Quantitative estimation of soil salinization in an arid region of the keriya oasis based on multidimensional modeling. *Water , 12*(3), 880.

Kilic, O., & Kilic, A. M. (2010). Salt crust mineralogy and geochemical evolution of the Salt Lake (Tuz Gölü), Turkey, *Scientific Research and Essays 5(11)*, 1317–1324.

Lhissou, R., El, A., & Chokmani, K. (2014). Mapping soil salinity in irrigated land using optical remote sensing data, *Eurasian Journal of Soil Science 3*, 82–88.

Liang, S., Member, S., Fang, H., & Chen, M. (2001). Atmospheric Correction of Landsat ETM + Land Surface Imagery — Part I : Methods, *Earth Sci Informatics 39*(11), 2490–2498.

Li Z, Li Y, Xing A, Zhuo Z, Zhang S, Zhang Y, Huang Y. (2019) Spatial prediction of soil salinity in a semiarid Oasis: environmental sensitive variable selection and model comparison. *Chinese Geogr Sci 29*:784–797.

Liu Y, Zhang F, Wang C, Wu S, Liu J, Xu A, Pan K, Pan X. (2019) Estimating the soil salinity over partially vegetated surfaces from multispectral remote sensing image using non-negative matrix factorization. *Geoderma 354*:113887.

Masoud AA, Koike K, Atwia MG, El-Horiny MM, Gemail KS. (2019) Mapping soil salinity using spectral mixture analysis of Landsat-8 OLI images to identify factors influencing salinization in an arid region. *Int J Appl Earth Obs Geoinf 83*:101944.

Machado, R.M.A., Serralheiro, R.P. (2017). Soil salinity: effect on vegetable crop growth. Management practices to prevent and mitigate soil salinization. *Horticulturae 3*, 30.

McKinney R, Hanson A, Johnson R, Charpentier M. (2019) Seasonal variation in apparent conductivity and soil salinity at two Narragansett Bay, RI salt marshes. *PeerJ 12*:1–21.

Metternicht, G., and J. Zinck. (2003). Remote Sensing of Soil Salinity: Potentials and Constraints. *Remote Sensing of Environment 85*(1):1–20.

Mergen, O., & Karacaoglu, C. (2015). Tuz Lake Special Environment Protection Area , Central Anatolia , Turkey : The EUNIS Habitat Classification and Habitat Change Detection between 1987 and 2007, *EKOLOJI, 24*(95), 1–9.

Mohammadifar,A., Gholami, H., Golzari,SH., Colins, A.L. (2021). Spatial modelling of soil salinity: deep or shallow learning models?*Environmental Science and Pollution Research*, *021*-13503-7.

Nawar, S., Buddenbaum, H., Hill, J., & Kozak, J. (2014). Modeling and Mapping of Soil Salinity with Reflectance Spectroscopy and Landsat Data Using Two Quantitative Methods (PLSR and MARS). *Remote Sensing*, *6*(11), 10813–10834.

Nguyen KA, Liou YA, Tran HP, Hoang PP, Nguyen TH (2020) Soil salinity assessment by using near-infrared channel and Vegetation Soil Salinity Index derived from Landsat-8 OLI data: a case study in the Tra Vinh Province, Mekong Delta, Vietnam. *Prog Earth Planet Sci 7*:1–16.

Odongo, V. O., Hamm, N. A. S., & Milton, E. J. (2014). Spatio-Temporal Assessment of Tuz Gölü, Turkey as a Potential Radiometric Vicarious Calibration Site. *Remote Sensing 6(3) :* 2494–2513. https://doi.org/10.3390/rs6032494

Peng, J. Biswas, A. Jiang, Q. Ruiying, Zh. Hu, J. Hu, B. and Shi, Zh. (2019). Estimating Soil Salinity from Remote Sensing and Terrain Data in Southern Xinjiang Province, China. *Geoderma 337*:1309–19.

Pouladi N, Jafarzadeh AA, Shahbazi F, Ghorbani MA (2019) Design and implementation of a hybrid MLP-FFA model for soil salinity prediction. *Environ Earth Sci 78*:159.

Racetin I, Krtalic A, Srzic V, Zovko M (2020) Characterization of short-term salinity fluctuations in the Neretva River Delta situated in the southern Adriatic Croatia using Landsat-5 TM. *Ecol Indic 110*:105924.

Ramak, R., Valadan Zouj, M. J., & Mojaradi, B. (2015). Improving Linear Spectral Unmixing Through Local Endmember Detection. ISPRS - International Archives of the Photogrammetry, *Remote Sensing and Spatial Information Sciences 2*:, 177–181.

Ren D, Wei B, Xu X, Engel B, Li G, Huang Q, Xiong Y, Huang G (2019) Analyzing spatiotemporal characteristics of soil salinity in arid irrigated agro-ecosystems using integrated approaches. *Geoderma 356*: 113935.

Sahab S, Suhani I, Srivastava V, Chauhan P, Singh R, Prasad V (2020) Potential risk assessment of soil salinity to agroecosystem sustainability: current status and management strategies. *Science of the Total Environment. 764*:144164.

Scudiero, E., Skaggs, T. H., & Corwin, D. L. (2014). Regional scale soil salinity evaluation using Landsat 7, western San Joaquin Valley, California, USA. *Geoderma Regional 2*(3): 82–90.

Sidike, A., Zhao, S., & Wen, Y. (2014). Estimating soil salinity in Pingluo County of China using QuickBird data and soil reflectance spectra. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 156–175.

Smith, C., Karunaratne, S., Badenhorst, P., Cogan, N., Spangenberg, G., & Smith, K. (2020). Machine-learning algorithms to predict forage nutritive value of in situ perennial ryegrass plants using hyperspectral canopy reflectance data. *Remote Sensing*, *12*(6), 1–15.

Solangi KA, Siyal AA, Wu Y, Abbasi B, Solangi F, Lakhiar IA, Zhou G (2019) An assessment of the spatial and temporal distribution of soil salinity in combination with field and satellite data: A case study in Sujawal district. *Agronomy* , *9*(12), 869.

Song JH, Shi XY, Cui J, Zhu YQ, Wang HJ (2019) Assessment of the accuracy of a soil salinity model for shallow groundwater areas in Xinjiang based on electromagnetic induction. *Appl Ecol Environ Res 17*:10037–10057.

Song, Ch. Ren, H. and Chong, H (2016). Estimating soil salinity in the Yellow River Delta, Eastern China - an Integrated Approach Using Spectral and Terrain Indices with the Generalized Additive Model. Pedosphere: *An International Journal 26*(5):626–35.

Taghadosi MM, Hasanlou M, Eftekhari K (2019) Retrieval of soil salinity from Sentinel-2 multispectral imagery. *Eur J Remote Sens 52*:138–154.

Taghizadeh R, Schmidth, K, Toomanian, N, Heung B, Behrens, T , Mosavi ,A, S.Band, SH, Chakan, A, Fathabadi, A, Scholten, Th (2021) Improving the spatial prediction of soil salinity in arid regions using wavelet transformation and support vector regression models . *Geoderma 383*:114793.

Tran T V., Tran DX, Myint SW, Huang C ying, Pham H V., Luu TH, Vo TMT (2019) Examining spatiotemporal salinity dynamics in the Mekong River Delta using Landsat time series imagery and a spatial regression approach. *Sci Total Environ 687*:1087–1097.

Thiam S., Villamor G., Faye., L (2021) Monitoring land use and soil salinity changes in coastal landscape: a case study from Senegal. *Environ Monit Assess. 193*:259

Tuz Lake Report. (2004) Final Report of the Project on the Determination of Alternative Agricultural Practices and Quality Classification of Soil and Land at Tuz Lake Special Protection Area, TR Ministry of Environment and Forestry General Directorate of Special Environmental Protection Agency & Ankara Research Institute under the General Directorate of Village Affairs, October 2004, Ankara, 391 pages.

Ucan, H. N., & Dursun, S. (2009). Environmental Problems of Tuz Lake (Konya-Turkey) #, 4(April), *J. Int. Environmental Application & Science 4*(2):231–233.

Van der Linden SL, Leiserowitz AA, Feinberg GD, Maibach EW. (2015) The Scientific Consensus on Climate Change as a Gateway Belief: Experimental Evidence. *PLoS ONE 10*(2): e0118489.

Wang F, Shi Z, Biswas A, Yang S, Ding J (2020a) Multi-algorithm comparison for predicting soil salinity. *Geoderma 365*:114211.

Wang F, Yang S, Yang W, Yang X, Jianli D (2019a) Comparison of machine-learning algorithms for soil salinity predictions in three dryland oases located in Xinjiang Uyghur Autonomous Region (XJUAR) of China. *Eur J Remote Sens 52*:256–276.

Wang J, Ding J, Yu D, Ma X, Zhang Z, Ge X, Teng D, Li X, Liang J, Lizaga I, Chen X, Yuan L, Guo Y (2019b) Capability of Sentinel-2 MSI data for monitoring and mapping of soil salinity in dry and wet seasons in the Ebinur Lake region, Xinjiang, China. *Geoderma 353*:172–187.

Wang J, Ding J, Yu D, Teng D, He B, Chen X, Ge X, Zhang Z, Wang Y, Yang X, Shi T, Su F (2020b) Machine-learning-based detection of soil salinity in an arid desert region, Northwest China: A comparison between Landsat-8 OLI and Sentinel-2 MSI. *Sci Total Environ 707*:136092.

Wang J, Sun Q, Shang J, Zhang J, Wu F, Zhou G, Dai Q (2020c) A new approach for estimating soil salinity using a low-cost soil sensor in situ: A case study in saline regions of China's east coast. *remote sensing. 2*(2), 239

Wang S, Chen Y, Wang M, Li J (2019c) Performance comparison of machine-learning algorithms for estimating the soil salinity of salt-affected soil using field spectral data. *Remote Sens 11*:1–26.

Wei Y, Shi Z, Biswas A, Yang S, Ding J, Wang F (2020) Updated information on soil salinity in a typical oasis agroecosystem and desert-oasis ecotone: Case study conducted along the Tarim River, China. *Sci Total Environ 716*:135387.

Wu, W., Mhaimeed, A. S., Al-Shafie, W. M., Ziadat, F., Dhehibi, B., Nangia, V., & De Pauw, E. (2014). Mapping soil salinity changes using remote sensing in Central Iraq. *Geoderma Regional, 2(3)*: 21–31.

Yagmur, N., Bilgilioglu, B., Dervisoglu, A., Musaoglu, N., Tanik, A. (2021). Long and Short-Term Assessment of Surface Area Changes in Saline and Freshwater Lakes via Remote Sensing, *Water and Environment Journal, 35*: 107-122.

Yang J, Zhao J, Zhu G, Wang Y, Ma X, Wang J, Guo H, Zhang Y (2020) Soil salinization in the oasis areas of downstream inland rivers —Case Study: Minqin oasis. *Quat Int 537*:69–78.

Yang RM, Guo WW (2019) Using sentinel-1 imagery for soil salinity prediction under the condition of coastal restoration. IEEE J Sel Top Appl Earth Obs *Remote Sens 12*:1482–1488.

Yahiaoui, I., Douaoui, A., Zhang, Q., & Ziane, A. (2015). Soil salinity prediction in the Lower Cheliff plain (Algeria) based on remote sensing and topographic feature analysis. *Journal of Arid Land, 7*(6): 794–805.

Zare S, Fallah Shamsi SR, Abtahi SA (2019) Weakly-coupled geo-statistical mapping of soil salinity to Stepwise Multiple Linear Regression of MODIS spectral image products. *J African Earth Sci 152*:101–114.

Zhang S, Zhao G (2019) A harmonious satellite-unmanned aerial vehicle-ground measurement inversion method for monitoring salinity in coastal saline soil. *Remote sensing 58*:126-133.

Zhang X, Huang B (2019) Prediction of soil salinity with soil-reflected spectra: A comparison of two regression methods. *Sci Rep 9*:1–8.

Zhang, T.-T., Qi, J.-G., Gao, Y., Ouyang, Z.-T., Zeng, S.-L., & Zhao, B. (2015). Detecting soil salinity with MODIS time series VI data. *Ecological Indicators, 52*, 480–489.

Ziqiang, M., Zhou, Sh., Yin Zh., Junfeng X., Wu Y., Yuanyuan Y., A.(2017). spatial data mining algorithm for downscaling TMPA 3B43 V7 data over the Qinghai–Tibet Plateau with the effects of systematic anomalies removed, *Remote Sensing of Environment,200*:378-395.

APPENDICES

APPENDIX A: Summary of various soil salinity mapping methods utilizing different sensing approaches

Table A.1 : Summary of various soil salinity mapping methods utilizing different sensing approaches

Published Year	Study Area	Spatial Extent	Sensing Approach				Satellite data	Analysis Methods
			Satellite Images (SI)	Aerial İmage (AI)	Field Measurement (FM)	Laboratory Analysis (LA)		
2021	Jaghin basin, Iran		*		*		Landsat-8 OLI	Deep learning (DL) models machine learning (ML) models Cubist Model
2021	Senegal		*		*		Landsat series	Combination of Salinity and vegetation indices Supervised classification and intensity analysis
2021	Qom plain, Iran		*		*		Landsat-8 OLI	Utilizing spectral indices of salinity, vegetation, topography, and drainage , Artificial Neural Network
2021	Isfahan Province, Iran		*		*		Landsat-8 OLI	Support vector regression integrated with wavelet transformation derived from a remote sensing data.
2021	Jaffna Peninsula, Sri Lanka		*		*	*	Landsat-8 OLI Sentinel-2A	Soil salinity indices PLSR
2020	Yucheng County,China	local (990.7 km^2)	*		*		Landsat	Principal Component Analysis (PCA) transformation, Tasseled Cap (TC) transformation, and Optimal Band Combination (OBC)
2020	Manas River Basin, China	local (3.35 × 10^4 km^2)	*		*	*	Landsat-8 OLI	PCR, PLSR and MLR modeling methods
2020	Urmia Lake, Iran		*		*	*	Landsat-8 OLI,Sentinel-2A	Regression Analysis
2020	Yazd-Ardakan plain, Iran	482,900 ha	*		*	*	Landsat MSS, ETM + , TM, OLI	Random Forest
2020	Ebinur Lake,CHINA		*		*	*	Landsat-8 OLI,Sentinel-2 MSI	Cubist Model
2020	Neretva River Delta,Croatia		*		*		Landsat-5 TM	Utilizing vegetation, moisture and salinity indices. Regression Analysis
2020	Minqin County , China		*		*		Landsat TM-OLI	trend analysis, stability analysis

Table A.1 (Continued) : Summary of various soil salinity mapping methods utilizing different sensing approaches

Published Year	Study Area	Spatial Extent	Sensing Approach				Satellite data	Analysis Methods
			SI	AI	FM	LA		
2020	Tra Vinh Province, Mekong Delta, Vietnam	local	*		*		Landsat-8 OLI	Utilizing salinity and environmental indices. Regression Analysis.
2020	CHINA	local			*		-	Utilizing Wifi POGO proximal sensor.
2020	Tarim River, China	572,858 km2	*		*		Landsant-8 OLI	Cubist and Random Forest
2020	China		*		*	*	Landsant-8 OLI	Utilizing Algorithms
2020	China		*		*		WorldView-2 (WV-2)	Partial Least-Squares Regression
2019	China		*		*	*	HuanJing-1 (HJ-1)	utilizing indices , IDW, Krigging
2019	Sujawal, Pakistan		*		*	*	Landsat5, Landsat-8 OLI	Utilizing indices , maximum likelihood classification algorithm
2019	Northwest China				*	*	-	partial least-squares regression (PLSR) model, principle component regression(PCR) model
2019	Shizuishan City, China				*	*	-	Random Forest Regression (RFR), Support Vector Regression (SVR), gradient-boosted regression tree (GBRT), Multilayer perceptron regression (MLPR), and Least angle regression (Lars)
2019	Yinchuan Plain, China	6967.80 km2	*		*	*	Landsat-8 OLI	Multiple linear regression model, Geographical weighted regression model, Random forest regression model
2019	Dongying City, China		*		*	*	Landsat 7 ETM+, Landsat-8 OLI	stepwise multiple linear regression, neural network and support vector machine method
2019	Xinjiang, China	local	*	*	*	*	GF-2	Random Forest Regression

Table A.1 (Continued) : Summary of various soil salinity mapping methods utilizing different sensing approaches

Published Year	Study Area	Spatial Extent	Sensing Approach				Satellite data	Analysis Methods
			SI	AI	FM	LA		
2019	China		*		*		Landsat-8 OLI	(MFR), PLSR, IDW
2019	Jiangsu Province,China	local (1,540 km^2)	*		*	*	Landsat 5	support vector machine (LS-SVM), back propagation neural network (BPNN), and random forest (RF)
2019	Egypt		*		*	*	Landsat ETM +, Landsat-8 OLI	Utilizing indices, linear spectral unmixing (LSU), and mixture tuned matched filtering (MTMF)
2019	Xinjiang, China		*		*	*	Sentinel-2A -MSI	Partial Least Squares Regression (PLSR)
2019	Xinjiang, China				*	*	–	PLSR
2019	China		*		*	*	Landsat-8 OLI	Utilizing vegetation and salinity indices
2019	China		*	*	*	*	Sentinel-2A	Multivariate linear regression, Second-order least squares regression
2019	Argentina		*		*	*	Landsat-8 OLI	Utilizing vegetation and salinity indices
2019	Morocco		*		*	*	Landsat-8 OLI	Regression Analysis
2019	China		*		*		Landsat-8 OLI	Multivariable linear regression
2019	Bangladesh		*		*		Landsat	Principal Component Analysis
2019	Iran		*		*	*	Landsat-8 OLI	MLP and hybrid MLP-FFA models
2019	China		*		*	*	Landsat-8 OLI	Cubist and Partial least squares regression
2019	Morocco		*		*	*	Landsat-8 OLI	Regression Analysis
2019	Vietnam	2360.2 km2	*		*	*	Landsat TM,ETM+,OLI	Regression Analysis
2019	China		*		*	*	Landsat 5 TM, Landsat-8 OLI	Utilizing indices
2019	North Carolina, USA		*		*	*	Sentinel-2 MSI, Landsat-8 OLI	Ordinary least squares regression
2019	China		–		*		–	Fuzzy k-means (FKM) algorithm

Table A.1 (Continued) : Summary of various soil salinity mapping methods utilizing different sensing approaches

Published Year	Study Area	Spatial Extent	Sensing Approach				Satellite data	Analysis Methods
			SI	AI	FM	LA		
2019	China		*		*	*	Landsat-8 OLI	Ordinary kriging (OK),back-propagation network (BP) and regression kriging (RK)
2018	Iran	local-2.3×10^4 km^2	*		*		Landsat TM/ETM+/OLI	support vector machine ,multi-criteria evaluation (MCE) procedure,cellular-automata (CA) Markov chain (CA-Markov) model
2018	Junggar Basin, China	local	*		*		MODIS	Empirical model decomposition method, linear and Random Forest (RF) regression models
2018	Ebinur Lake Wetland National Nature Reserve (ELWNNR)	local (2670.8 km^2)	*		*	*	Landsat OLI, Huanjing (HJ) 1-B CCD	Develop optimal band Difference Index (DI), Ratio Index (RI), and Normalization Index (NDI) algorithms , Bootstrap-BP neural network model
2018	Pompenillo, Grañén, Spain	local (45.2 ha)	*		*	*	Landsat 5 TM	Regression analysis and ordinary kriging (OK)
2018	Zhangye Oasis, Northwest China	local		*	*		Airborne hyperspectral data	Forced Invariance Approach, Generalized Linear Model (GLM)-(generalization of ordinary linear regression)
2018	Inner Mongolia of China	local (0.81 km2)	*		*		Landsat 5, Landsat-8 OLI	PLSR model, principal components analysis (PCA), canonical correlation analysis, linear regression analysis method
2018	Timpaki, Greece	local (50 km^2)	*		*		WV2 images, Landsat-8 OLI	Regression analysis, Utilizing indices, Principal Component Analysis (PCA)
2017	South Khorasan, Iran	local-2487 km2	*		*		Landsat ETM+	Interpretation and calculation of salinity index from satellite data (Normalized Difference Salinity Index- NDSI) Regression analysis
2017	Oran watershed, Algeria	local-1878 km2	*		*		Landsat-8 OLI, TIRS	RBF interpolation, regression fitting model
2017	Dharmapuri,India	local (430 km2)	*		*		Landsat 7-8 (ETM+ and OLI), RADAR (SAR)	cross--correlation analysis, simple linear regression analysis
2017	Wadi ad-Dawasir,Saudi Arabia	local	*				Landsat-8 OLI	Principal component analysis (PCA),artificial neural network (ANN) analyses

Table A.1 (Continued) : Summary of various soil salinity mapping methods utilizing different sensing approaches

Published Year	Study Area	Spatial Extent	Sensing Approach				Satellite data	Analysis Methods
			SI	AI	FM	LA		
2017	Turkey	local-1.500 km2	*		*		Landsat-5 TM, Landsat-8	Linear regression analysis
2017	Biskra area, Algeria	local-5000 km2	*		*		Landsat 5 , Landsat-8 OLI	Decision tree classification (DTC), Principle component analysis , Utilizing spectral indices
2017	Iran	local	*		*		Landsat7 ETM+	PC analysis;MNF transformation;utilizing PPI, and nDV algorithms; regression analysis
2016	Yellow River Delta, China	local	*			*	Landsat-4/5 TM,Landsat-7 ETM+,Landsat-8 OLI	PLSR (Partial Least Square Regression) model
2016	Yellow River Delta,China	local	*			*	Landsat TM	The generalized additive model (GAM) with using spectral and terrain indices model - NDVI , RVI , DVI
2016	Morocco	local	*		*		Landsat 5 TM and Landsat-8 OLI	Utilizing the SoilSalinity Spectral Indices (SSSI)-(OLI-SI index,NDVI)-,simple linear regression between sampled soil EC
2016	Northern Cape, South Africa	local	*		*		WorldView2 (WV2) sensor	principle component analysis (PCA) components.Regression analyses classification and regression tree (CART) modelling
2016	Vaalharts and Breede River ,South Africa	local	*		*		SPOT-5	Regression modelling (stepwise linear regression, partial least squares regression, curve fit regression modelling) and supervised classification (maximum likelihood, nearest neighbor, decision tree analysis, support vector machine (SVM) and random forests)
2016	Çukurova,Turkey	local	*		*		Landsat TM/ETM	Object based classification, Vegetation indices),Multi-Linear regression (MLR) and SLR, Radial basis function technique
2016	Western Desert, Egypt	local	*		*		Landsat-8 (OLI)	regression analysis, utilizing satellite image
2016	Northeast China	local	*		*	*	Landsat-8 OLI	Correlation analysis regression model
2016	Bangladesh	local	*		*		Landsat 7 ETM+	Regression analysis, Utilizing salinity indices
2016	Morocco	local	*		*		EO-1 ALI	Utilizing salinity indices

Table A.1 (Continued) : Summary of various soil salinity mapping methods utilizing different sensing approaches

Published Year	Study Area	Spatial Extent	Sensing Approach				Satellite data	Analysis Methods
			SI	AI	FM	LA		
2016	Saudi Arabia	local	*		*	*	Landsat-8 OLI	regression correlation / Linear regression model
2016	Yellow River Delta	local	*		*		Landsat 7- Landsat-8 OLI	Multiple linear regression, kriging
2015	Gabes, Tunisia	local	*		*	*	Landsat-8	Partial least square regression (PLSR) method; utilizing spectral salinity indices
2015	Algeria	local	*		*		Landsat ETM+	Multiple linear regression model
2015	East Nile Delta Region, Egypt	local	*		*		Landsat-8 OLI, ASTER GDEM	Principal component analysis, utilizing salinity indices
2015	Brazil	local	*		*	*	OLI/Landsat-8 ,Hyperion	Principal component analysis (PCA), support vector machine (SVM), regression analysis
2015	Keriya River Basin, Northwestern China	local	*		*		Landsat ETM+, PALSAR and Radarsat-2	Support vector machine (SVM) classification, decision tree (DT) classifier
2015	Xinjiang, China	local	*		*	*	Hyperion	Maximum likelihood, k-nearest neighbor, support vector machine (SVM), neural network
2015	Nile Delta, Egypt	local	*		*		Landsat 5, 7 ,8-ASTER	Supervised and unsupervised classification (maximum likelihood classification), Principal Components Analysis (PCA), and indices
2015	West Texas, USA	local	*		*		Landsat-8 OLI	Regression and correlation analysis
2015	Yellow River Delta, China	local	*		*		Landsat 7 ETM+, EO-1 ALI and Landsat-8 OLI	Partial least square regression (PLSR) model
2015	Yellow River, China	local	*		*		MODIS	Correlation and regression analysis
2015	Southern Japan	local	*		*		Landsat 5	Utilizing NDVI index
2014	Kuqa Oasis, China	local	*		*	*	Landsat-TM	Regression kriging interpolation, Spectral index regression,
2014	Mesopotamia, Iraq	local	*		*		Landsat ETM+ and MODIS	Multiple linear regression analysis
2014	El-Tina Plain, Egypt	local	*		*		Landsat7 ETM+ , Landsat 5	PLSR and MARS
2014	San Joaquin Valley, California, USA	local	*		*		MODIS	Utilizing time-series vegetation indices

CURRICULUM VITAE

Name Surname: Taha Gorji

M.Sc.: Istanbul Technical University**,** Department of Environmental Engineering, Environmental Science and Engineering Program

B.Sc.: Science and Culture University, Department of Industrial Engineering, Tehran, Iran

PUBLICATIONS/PRESENTATIONS ON THE THESIS

Gorji, T., Yıldırım, A., Hamzehpour, N., Tanık, A., Sertel, E. (2020). Soil Salinity Analysis of Urmia Lake Basin using Landsat-8 OLI and Sentinel-2A based Spectral Indices and Electrical Conductivity Measurements, Ecological Indicators, 112 (106173). (Article)

Gorji, T., Alganci, U., Sertel, E., Tanik, A. (2019). Comparing Two Different Spatial Interpolation Approaches to Characterize Spatial Variability of Soil Properties in Tuz Lake Basin- Turkey, Fresenius Environmental Bulletin, 28 (2): 986-992. (Article)

Gorji, T., Yildirim, A., Sertel, E., Tanik, A. (2019). Remote sensing approaches and mapping methods for monitoring soil salinity under different climate regimes, International Journal of Environment and Geoinformatics (IJEGEO), 6 (1): 33-49. (Article)

Yıldırım, A., **Gorji, T**., Hamzehpour, N., Sertel, E., Tanık, A. (2019). Comparison of Different Soil Salinity Indices Derived from Sentinel-2A Images, International Symposium on Applied Geoinformatics (ISAG-2019), 7-9, Proceedings, Vol 1, No.1, p. 230-233.

Gorji, T., Yıldırım, A., Hamzehpour, N., Sertel, E., Tanık, A. (2019). Mapping Soil Salinity by Using Landsat-8OLI Imagery and Regression Analysis over Bonab Country of East Azerbaijan Province in Iran, International Symposium on Applied Geoinformatics (ISAG-2019), 7-9 November 2019, Istanbul, Turkey, Proceedings, Vol 1, No.1, p. 504-507. (Conference paper)

Gorji, T., Yildirim, A., Hamzehpour, N., Sertel, E., Tanik, A. (2019). Characterizing Spatial Variability in Urmia Lake Basin by Applying Geo-Statistical Methods, 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics), 16-19 July 2019, Istanbul, Proceedings, 978-1-7281-2116-1/19/$31.00 ©2019 IEEE. (Conference paper)