MODELING OF THE IONOSPHERE'S DISTURBANCE USING DEEP LEARNING TECHNIQUES

İYONOSFER BOZULMALARININ DERİN ÖĞRENME TEKNİKLERİ KULLANILARAK MODELLENMESİ

RAHEM ABRI ZANGABAD

Assoc. DR. HARUN ARTUNER

Supervisor

Submitted to

Graduate School of Science and Engineering of Hacettepe University

as a Partial Fulfillment to the Requirements

for the Award of the Degree of Doctor of Philosophy

in Computer Engineering

2021

ABSTRACT

MODELING OF THE IONOSPHERE'S DISTURBANCE USING DEEP LEARNING TECHNIQUES

RAHEM ABRI ZANGABAD
Doctor of Philosophy, Computer Engineering Department
Supervisor: Assoc. Prof. Dr. Harun ARTUNER
September 2021, 158 pages

The ionosphere drives an essential role in the atmosphere and earth. Solar flares due to coronal mass ejection, seismic movements, and geomagnetic activity cause deviations in the ionosphere. The main parameter for investigating the structure of the ionosphere is Total Electron Content (TEC). This thesis converges the importance of ionosphere TEC data to evaluate seismic events. The dataset assessed in this thesis contains the ionospheric variability during moderate and severe earthquake events of varying strengths for 2012-2019 years in Chile station. TEC values obtained from GPS stations provide a powerful technique for analyzing the ionospheric response to earthquakes and solar storms. TEC data gathered from GPS stations (Dual-Frequency GPS receiver) is used to investigate the ionospheric variability through moderate and severe earthquakes. This thesis has three main contributions.

In the first contribution, our goal is to analyze the relations between earthquakes and TEC data. We concentrate on extracting features from earthquakes and classification over the ionospheric TEC data. In this phase, we do not focus on predicting earthquakes with previous days. The proposed model uses Deep Autoencoders to extract features from TEC data. As the ionospheric TEC data is a high-dimensional factor, reducing dimensionality to obtain

a compressed feature set is an essential step in the feature extraction phase. The collected features served as input to dense neural networks to perform classification. The classification model results are compared against the LDA(Linear Discriminant Analysis), SVM(Support Vector Machine) and Random Forest classifier models to evaluate the proposed model. The results report that the proposed model improves in distinguishing the earthquakes at an accuracy rate of about 0.94 in the target station zone.

In the second contribution, we propose a classification model to detect earthquakes in previous days. The LSTM methods handle this issue with the solution to short-term memory. The proposed models use LSTM-based (Long Short-Term Memory), deep learning models to classify earthquakes days by analyzing TEC values of the last seven days. The variant versions of the LSTM models are proposed to enhance the contribution of this research. The LSTM-Based classification models are compared against the SVM, LDA and Random Forest classifier models to evaluate the proposed models. The results reveal that the proposed models improve in detecting the earthquakes at an accuracy rate of about 78-80 and can be used as a successful tool for detecting earthquakes based on the previous days.

In the last contribution of this thesis, we develop a hybrid version of deep autoencoders and LSTM to detect earthquakes in previous days. This model proposes to improve the stacked LSTM-based earthquake classification introduced in the second model. The suggested model uses a deep autoencoder to derive beneficial features from ionospheric TEC data and perform Stacked LSTM to classify earthquakes days by analyzing TEC values of the last seven days. To analyze the contribution of the suggested DAE-STCK-LSTM model, we used Stacked-LSTM, LDA, and SVM classifiers. Our evaluation test results prove approximately 81-84 accuracy-based performance in the two test sets of the earthquakes, including moderate and severe earthquakes.

Keywords: Inosphere, Total Electron Content, Deep Autoencoder, Long Short-Term Memory, Linear Discriminant Analysis, Support Vector Machine, Random Forest

ÖZET

İYONOSFER BOZULMALARININ DERİN ÖĞRENME TEKNİKLERİ KULLANILARAK MODELLENMESİ

RAHEM ABRI ZANGABAD

Doktora, Bilgisayar Mühendisliği

Danışman: Doç. Dr. Harun ARTUNER

Eylül 2021, 158 sayfa

İyonosfer atmosferde ve dünyada önemli bir rol oynar. Koronal kütle atımı, sismik hareketler ve jeomanyetik aktiviteye bağlı güneş patlamaları iyonosferde sapmalara neden olur. İyonosferin yapısını araştırmak için ana parametre Toplam Elektron İçeriğidir (TEC).

Bu tez, sismik olayları değerlendirmek için iyonosfer TEC verisinin önemini araştırmaktadır. Bu tezde, 2012-2019 yılları arasında Şili istasyonundaki orta ve yüksek şiddetli depremler sonucunda ortaya çıkan iyonosfer anomalilerini içeren bir veri seti değerlendirilmiştir. GPS istasyonlarından elde edilen TEC değerleri, depremlere ve güneş fırtınalarına karşı iyonosferik tepkiyi analiz etmek için güçlü bir yöntemdir. GPS istasyonlarından (çift frekanslı GPS alıcısı) elde edilen TEC verileri, orta ve yüksek şiddetli depremlerin sebep olduğu iyonosferik anomalileri araştırmak için kullanılır. Bu tezin modeli üç ana bölümden oluşmaktadır.

İlk bölümdeki amaç, depremler ile TEC verileri arasındaki ilişkiyi analiz etmektir. Deprem verilerinden öznitelikler çıkarmaya ve iyonosferik TEC verisini sınıflandırmaya odaklanılmıştır. Bu aşamada depremleri önceki günlerin bilgisini kullanarak tahmin etmeye odaklanılmamıştır. Önerilen model, TEC verilerinden öznitelik çıkarmak için Derin Otomatik

Kodlayıcı kullanmaktadır. İyonosferik TEC verisi yüksek boyutlu olduğundan, öznitelik çıkarma aşamasında boyut azaltma işlemi yalın öznitelik seti elde etmek için kritik öneme sahiptir. Çıkarılan öznitelikler, sınıflandırma yapmak için kullanılan sinir ağlarına girdi olarak verilmektedir. Sınıflandırma modelinin sonuçlarını değerlendirmek için, LDA (Doğrusal Diskriminant Analizi), SVM(Destek Vektör Makinesi) ve Random Forest (Rastgele Orman) sınıflandırma modelleri ile kıyaslama yapılmıştır. Elde edilen sonuçlar, önerilen modelin belirlenen hedef istasyonda 0.94 doğruluk oranı ile depremleri ayırt edebildiğini göstermektedir.

İkinci bölümdeki amaç, geçmiş günlerdeki depremleri tespit etmek için bir sınıflandırma modeli önerilmiştir. LSTM yöntemleri, bu sorunu kısa süreli bellek(short-term memory) çözümü ile ele alır. Önerilen modeller, son yedi günün TEC değerlerini analiz ederek deprem günlerini sınıflandırmak için LSTM tabanlı (Uzun Kısa Süreli Bellek), derin öğrenme modellerini kullanır. LSTM modellerinin farklı versiyonları, bu araştırmanın doğruluğunu arttırmak için önerilmiştir. LSTM tabanlı sınıflandırma modelleri, önerilen modelleri değerlendirmek için SVM, LDA ve Random Forest sınıflandırıcı modelleri ile karşılaştırılır. Sonuçlar, önerilen modellerin depremleri yaklaşık yüzde 78-80 doğruluk oranında tespit etmede geliştiğini ve önceki günlere göre depremleri tespit etmek için başarılı bir araç olarak kullanılabileceğini ortaya koymaktadır.

Bu tezin son kısmında, önceki günlerde depremleri tespit etmek için derin otomatik kodlayıcıların(deep autoencoders) ve LSTM'nin hibrit bir versiyonu geliştirildi. Bu model, ikinci modelde tanıtılan yığın LSTM(stacked LSTM) tabanlı deprem sınıflandırmasını geliştirmeyi önermektedir. Önerilen model, iyonosferik TEC verisinden faydalı özellikler çıkarmak için derin bir otokodlayıcı kullanır ve son 7 günün TEC değerlerini analiz ederek deprem günlerini sınıflandırmak için yığın LSTM uygular. Önerilen DAE-STCK-LSTM modelinin amacını analiz etmek için Yığın-LSTM, LDA ve SVM sınıflandırıcıları kullanıldı. Değerlendirme test sonuçları, orta ve şiddetli depremler dahil, iki test kümesinde yaklaşık yüzde 81-84 doğruluk tabanlı(accuracy-based) performans göstermiştir.

Anahtar Kelimeler: İyonosfer, Toplam Elektron İçeriği, Derin Otomatik Kodlayıcı, Uzun Kısa Süreli Bellek, Doğrusal Diskriminant Analizi, Destek Vektör Makinesi, Rastgele Orman

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisors Assoc. Prof. Dr. Harun Artuner for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. They have always been there for me to provide encouragement and helpful feedback on my research. I would also like to thank the IONOLAB group, especially Prof.Dr. Feza Arıkan, for their valuable guidance throughout my studies. She provided me with the tools that I needed to choose the right direction and complete my dissertation.

Besides my supervisors, I would like to thank the rest of my thesis committee members, for their insightful comments and encouragement, but also for the hard question which incented me to widen my research from various perspectives.

Last but not least, I would like to thank my family, especially my dearest spouse Dr. Sara Abri, who has given me the greatest motivation to finish my work. To be honest, I would never have completed a Ph.D. without her. I sincerely appreciate his belief in me.

CONTENTS

Page

ABSTRACT i
ÖZET iii
ACKNOWLEDGEMENTS vi
CONTENTS vii
TABLES xi
FIGURES xiii
ABBREVIATIONS xiv
1. INTRODUCTION 1
1.1. Motivation 1
1.2. Contribution and Outline 3
1.2.1. Main Contributions 3
1.2.2. Thesis Outline 4
2. BACKGROUND INFORMATION 6
2.1. Atmosphere Layers Characteristics 6
2.1.1. Troposphere Layer 7
2.1.2. Stratosphere Layer 7
2.1.3. Mesosphere Layer 7
2.1.4. Thermosphere Layer 8
2.1.5. Exosphere Layer 8
2.2. Ionosphere Layers Characteristics 8
2.2.1. Ionosphere Layers 9
2.3. Deep Neural Networks 11
2.3.1. Design of the DNN Architecture 14
3. LITERATURE REVIEW 16
3.1. Scenario based Earthquake Prediction 18
3.2. Precursors based Earthquake Prediction using Machine Learning and Statistical Models 19

4. Ionosphere TEC Based Dataset 25
4.1. Dataset 25
4.2. Data Preparation 26
4.2.1. Similarity Learning Between Stations 27
5. Evaluation Metrics Used in Methodology 29
5.1. Related Terminology in Confusion Matrix 29
5.2. Accuracy Metric in Related Earthquake Classifier 29
5.3. Precision Metric in Related Earthquake Classifier 30
5.4. Recall Metric in Related Earthquake Classifier 30
5.5. Metric Evaluation in Related Earthquake Classifier 30
5.6. F1-score Metric in Related Earthquake Classifier 31
5.7. Receiver Operating Characteristic (ROC) Metric 31
5.8. Area Under the Curve (AUC) Metric 32
5.9. P-value Metric 33
5.10.Cross-validation 34
5.10.1. K-fold cross-validation 34
5.10.2. Stratified K-fold cross-validation 34
6. Interpretation of earthquakes based on ionosphere disturbances using deep neural network 36
6.1. Feature Extraction 36
6.2. Classification Method 39
6.3. Dataset for DAEclass Method 42
6.4. Evaluation Methodology 42
6.5. Evaluation Results 44
6.6. Conclusion for DAEclass model 50
7. LSTM-based deep learning methods for prediction of earthquakes using ionospheric data 51
7.1. LSTM-based Sequential Learning Models 51
7.2. Dataset For LSTM-Based Models 53
7.3. Evaluation Results 54

7.4. Conclusion for LSTM based models 62
8. An approach using Deep AutoEncoder and LSTM Neural Networks for prediction of earthquakes using ionospheric data 64
8.1. Deep Auto-Encoder Stacked-LSTM Model 64
8.1.1. Deep Autoencoder Sub-Model 65
8.1.2. Stacked LSTM Sub-Model 66
8.2. Dataset For DAE-STCK-LSTM Model 66
8.3. Evaluation Results 67
8.4. Conclusion for DAE-STCK-LSTM model 72
9. CONCLUSION 74

A Heat Maps of TEC based Data Related to IQUIQUE Station 77
B Heat Maps of TEC based Data Related to KARRATHA Station 95
C All Earthquake Detailed Information 113

REFERENCES 117

TABLES

Page

Table 4.1. Prepared data for experimentation. 27

Table 6.1. Detailed information about proposed network architecture. 41

Table 6.2. Prepared data for proposed DAEclass Model. 42

Table 6.3. Comparison of the DAEclass and other classification models using performance metrics. 44

Table 6.4. Detailed 10 fold comparison of DAEclass and LDA classification models using performance metrics. 46

Table 6.5. Detailed 10 fold comparison of SVM and Random Forest classification models using performance metrics. 47

Table 6.6. AUC values for the DAEclass and the LDA classifier. 49

Table 7.1. Prepared data for the proposed LSTM based models. 54

Table 7.2. Comparison of the proposed LSTM models other classification models using performance metrics. 55

Table 7.3. Detailed 10 fold comparison of the proposed LSTM models and SVM classification models using performance metrics in the EQs$\geq$4.5 dataset. 57

Table 7.4. Detailed 10 fold comparison of LDA and Random Forest classification models using performance metrics in the EQs$\geq$4.5 dataset. 58

Table 7.5. Detailed 10 fold comparison of the proposed LSTM models and SVM classification models using performance metrics in the EQs$\geq$5.0 dataset. 59

Table 7.6. Detailed 10 fold comparison of LDA and Random Forest classification models using performance metrics in the EQs$\geq$5.0 dataset. 60

Table 7.7. AUC values for the LSTM models and the SVM classifier. 62

Table 8.1. Prepared data for the proposed DAE-STCK-LSTM model. 66

Table 8.2. Comparison of the proposed DAE-STCK-LSTM models and other classification model using performance metrics. 67

Table 8.3. Detailed 10 fold comparison of the proposed DAE-STCK-LSTM model and other classification models using performance metrics in the EQs≥4.5 dataset. 69
Table 8.4. Detailed 10 fold comparison of the proposed DAE-STCK-LSTM model and other classification models using performance metrics in the EQs≥5.0 dataset. 70
Table 8.5. AUC values for the DAE-STCK-LSTM models and the other classification models. 72

FIGURES

Page

Figure 2.1. Layers of the atmosphere. 6
Figure 2.2. Layers of the ionosphere. 9
Figure 2.3. Detailed information of ionosphere layers. 10
Figure 2.4. Radio communication with the ionosphere. 11
Figure 2.5. Deep Neural Network layers. 12
Figure 2.6. Node from Neural Network. 13
Figure 4.1. The Iqqe station is located in Coordinate (Lat:-20.15, Lon:-70.13) Iquique in Chile and the karr station located in coordinate $(Lat : -20.85, Lon : 117.1)$ Karratha in the Pilbara region of Western Australia. 25
Figure 4.2. Iqqe and karr stations located in the same latitude from the two hemispheres of the east and west of the earth. 26
Figure 5.1. ROC table: detailed information. 31
Figure 5.2. AUC table: detailed information. 32
Figure 5.3. Four folds Stratified Cross Validation 35
Figure 6.1. A general overview of the proposed model. 36
Figure 6.2. A Deep Autoencoder with five layers. 38
Figure 6.3. The general architecture of the proposed DAEclass model. 40
Figure 6.4. Leaky ReLU chart 41
Figure 6.5. Performance metrics of the DAEclass and other classifier models based on two datasets (EQs≥ 4.5 and EQs≥ 5.0). 48
Figure 6.6. ROC curve of the DAEclass and the LDA classifier models based on two datasets (EQs$\geq$4.5 and EQs$\geq$5.0). 49
Figure 7.1. Structure of LSTM neural network cell. 52
Figure 7.2. Structure of Stacked LSTM. 53
Figure 7.3. Dataset structure used in the proposed LSTM based models. 54

Figure 7.4. Performance metrics of the LSTM models and the SVM classifier based on two datasets (EQs≥4.5 and EQs≥5.0). 61
Figure 7.5. ROC curve of the LSTM based models and the SVM classifier. 62
Figure 8.1. The general architecture of the proposed DAE-STCK-LSTM model. . 65
Figure 8.2. Performance metrics of the DAE-STCK-LSTM model and other classification model based on two datasets (EQs≥4.5 and EQs≥5.0). 71
Figure 8.3. ROC curve of the LSTM based models and the SVM classifier. 72

ABBREVIATIONS

TEC	**T**otal **E**lectron **C**ontent
TECU	**T**otal **E**lectron **C**ontent **U**nit
EQs	**E**arth **Q**uakes
DNN	**D**eep **N**eural **N**etwork
SVM	**S**upport **V**ector **M**achine
LDA	**L**inear **D**iscriminant **A**nalysis
RF	**R**andom **F**orest
LSTM	**L**ong **S**hort **T**erm **M**emory
GPS	**G**lobal **P**ositioning **S**ystem
ANN	**A**rtificial **N**eural **N**etwork
ReLU	**R**ectified **L**inear **U**nit
CNN	**C**onvolutional **N**eural **N**etwork
Iqqe	**I**quique
Karr	**K**arratha
ROC	**R**eceiver **O**perating **C**haracteristic
AUC	**A**rea **U**nder **C**urve
PCA	**P**rincipal **C**omponent **A**nalysis
Enc	**E**ncoder
Dcr	**D**ecoder
DAE	**D**eep **A**uto **E**ncoder
SGD	**S**tochastic **G**radient **D**escent
DAEclass	**D**eep **A**uto **E**ncoder **c**lassification
RNN	**R**ecurrent **N**eural **N**etwork
Bi-LSTM	**B**idirectional **L**ong **S**hort **T**erm **M**emory

1. INTRODUCTION

1.1. Motivation

As we gradually rise above the ground, we will encounter an atmospheric classification in terms of elevation, some of which will be of particular importance. The nature of molecules or ions that depend on the Earth's gravitational field changes the absorption of solar radiation and, therefore, temperature, density, and ionization. The boundaries of spatial layers are not constant, neither spatially nor temporally, because their interfering parameters are not fixed either. When moving away from the Earth's atmosphere, it is reached a layer from 48km to 965km (600mi) that is ionized to a plasma phase called the ionosphere layer.

A major measurable parameter that indicates a characteristic of the ionosphere is Total Electron Content (TEC). TEC provides an effective means to probe the structure of the ionosphere. In the literature, TEC is defined as the line integral of electron density along a ray path or as a measure of the total electrons along a ray path. The unit of TEC is given in TECU(TEC Unit), where 1 TECU = $10^{16}\ electron/m^2$ defined by Arikan et al. and Nayir et al. [1, 2]. The fluctuations and turbulences of the ionosphere layer can be captured adequately and efficiently by calculating and watching TEC Values. Over the past few decades, the Global Positioning System (GPS) provides a cost-effective explanation in calculating and assessing TEC and observing the ionospheric layer turbulences over a notable proportion of global landmass as conducted work [2].

Overall, the ionosphere layer's temporal and spatial variability directly relates to the earth's daily (each day) and annual rotation and the pattern of magnetic field lines of the geomagnetic dipole. As discussed in [3], earth's magnetic field even when no existence of geomagnetic activities is scarcely quiet. The standard intermittent alterations generate the dynamics of the quiet ionosphere. The ionosphere's quiet conditions are impacted by the variations in the geomagnetic and solar activity and seismicity. So, these consequences can cause disturbances in parameters like earthquakes.

Empirically, discovering precursory signals in strong earthquakes is the subject of most case studies on earthquake prediction in the literature. Nevertheless, the precursory signals through long time periods are not investigated. The main motivation of this thesis is to investigate models to analyze the relationship between earthquakes and ionospheric TEC data. The foremost aim of the recommended models is predicting an earthquake by analyzing previous days of the earthquake using ionospheric TEC data.

TEC values obtained from GPS stations are a powerful technique for analyzing the ionospheric response to earthquakes and solar storms. TEC data gathered from GPS stations (Dual-Frequency GPS receiver) is used to investigate the ionospheric variability through moderate and severe earthquakes. For this research, TEC data has been collected from two GPS stations. This data has been collected from the IONOLAB group (Hacettepe University of IONOLAB is an organization of electrical engineers to investigate hurdles of the ionosphere.) The ionospheric variability during moderate and severe earthquake events of varying strengths for 2012-2019 years is discussed in this thesis.

This thesis uses different deep learning techniques and algorithms to extract relations between TEC data. For this purpose, it aims to use supervised and unsupervised learning methods based on deep learning techniques. In the unsupervised learning methods, we use Deep Autoencoder and LDA method. On the other hand, we use deep, dense neural networks, Long short-term memory, and support vector machine in the supervised learning methods. Furthermore, as the TEC data related to days is large enough, we need a feature extraction mechanism to reduce valuable features. Therefore, classification models are developed to handle extracted features to predict earthquake days or quiet days in the dataset.

1.2. Contribution and Outline

1.2.1. Main Contributions

The contributions of the thesis are as bellows:

- To estimate the potential for predicting earthquakes, exploring the characteristics of the ionosphere to achieve a correct prediction process. For this purpose, TEC features are examined in several aspects.
- In this framework, new classification models are proposed using hybrid deep learning-based algorithms to extract the relations between TEC data and earthquakes. In this way, we can overcome the noise problem related to TEC data with deep autoencoders to extract the features for earthquake days and quiet days and perform classification to increase earthquake prediction accuracy.
- After feature extraction of TEC data, we use different algorithms such as deep dense neural networks or variants of LSTM to perform classification based on supervised learning methods for predicting earthquake days and quiet days.
- In the process of feature extraction and classification, we use two GPS stations TEC data for 2012-2019 years. The first station is located in coordinate $(Lat : -20.15, Lon : -70.13)$ Iquique in Chile and the second station is located in coordinate $(Lat : -20.85, Lon : 117.1)$ Karratha in the Pilbara region of Western Australia.
- Finally, our contribution in this thesis is divided into three main models. In the first model, we focus on interpreting earthquakes based on ionosphere disturbances using deep neural networks. In the second contribution, we propose a classification model to detect earthquakes in previous days using time series recurrent neural networks. Finally, In the last model, we develop a hybrid version of deep autoencoders and LSTM to detect earthquakes in previous days.

1.2.2. Thesis Outline

An outline of the chapters in this thesis includes an overview of the related works of ionosphere characteristics and anomalies, Scenario-based ionosphere anomalies, Earthquake scenarios based on ionosphere fluctuations, and machine learning algorithms for classification and feature extractions. Following that, three chapters detail our contributions and describe the foremost models related to predicting and classifying moderate and severe earthquakes based on the anomalies that occurred in the ionosphere layer. The organization of the thesis is as follows:

Chapter 2. describes the background information about atmosphere layers, especially ionosphere layers and their characteristics, machine learning methods, and deep learning algorithms based on supervised and unsupervised learning methods. Chapter 3. presents the related works on Scenario-based ionosphere irregularities, earthquakes based on ionosphere changes, and machine learning algorithms for classification ionospheric anomalies.

The dataset related to earthquakes and TEC values of GPS stations and preprocessing steps are presented in the Chapter 4.. Chapter 5. discusses learning parameters and evaluation metrics used for evaluation and helps better understand the metrics required to do a quality evaluation. Using the evaluation methodology, performances of our proposed models and the other classifier models are evaluated.

The models of predicting and classyfing are discussed in chapters 6., 7. and 8. These chapters include our main contributions about feature extraction, classification and predicting earthquake days. Every contribution in the previous chapters is organized to compare and examine the performance metrics of the proposed classification models.

In Chapter 6. , we concentrate on interpreting the relations between earthquake days and quiet days based on ionosphere disturbances using deep neural networks. Chapter 7. propose a classification model to detect earthquakes in previous days using time series LSTM models. Chapter 8. describes a hybrid version of deep autoencoders and LSTM to detect

earthquakes in previous days. Finally, in the last part of the thesis in Chapter 9., we includes the concluding and future works remarks.

2. BACKGROUND INFORMATION

This chapter focus on the background information about atmosphere layers and artificial intelligence based algorithms that are used in this thesis.

2.1. Atmosphere Layers Characteristics

As we gradually rise above the ground, we will encounter an atmospheric classification. Earth's atmosphere has a sequence of layers; every layer has its own specific characteristics. The atmosphere is divided into file layers, as depicted in the Figure 2.1.. These layers are identified as the troposphere, stratosphere, mesosphere, thermosphere, and exosphere.

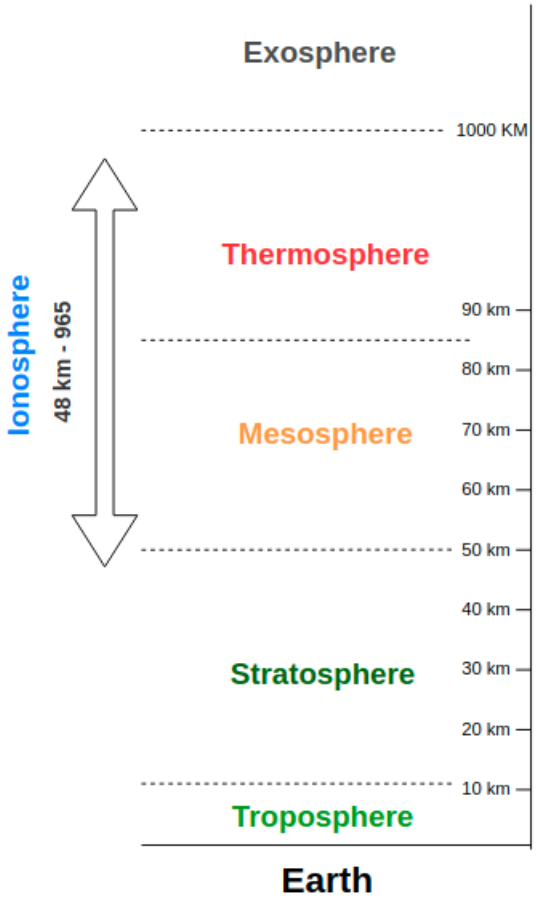


Figure 2.1. Layers of the atmosphere.

2.1.1. Troposphere Layer

The troposphere is the most inferior layer in the Earth's atmosphere. The troposphere is identified between the Earth's surface and about 12km above sea level. Approximately all kinds of clouds are observed in the troposphere layer, and nearly all weather happens inside this layer. This layer is the wettest layer between the atmosphere layers. The air in the troposphere layer consists of 78% nitrogen, 21% oxygen, and 1% of argon, water vapor, and carbon dioxide.

2.1.2. Stratosphere Layer

The layer directly above the troposphere is named Stratosphere. The borderline between the troposphere and the stratosphere is termed Tropopause. The second layer in the atmosphere is the stratosphere layer that starts is around 12 km (7.3 miles or 38,000 feet) above sea level at middle latitudes. However, the height of the starting point of the stratosphere varies with latitude and with the seasons. For instance, the start boundary of the stratosphere layer is decreased to about 11 km at the poles in winter. The stratosphere layer is ended at 50 km (31 miles). In contrast to the troposphere, where the temperature decreases when we move upwards, in the stratosphere the temperature increases as we move upwards.

2.1.3. Mesosphere Layer

Beyond the stratosphere is the mesosphere layer that stretches upward to a height of approximately 85 km (53 miles) over the sea level on Earth. Unlike the stratosphere, as it rises through the mesosphere, the temperature reduces once again. The coldest temperature in the Earth's atmosphere, approximately -90 ° C (-130 ° F), is the top of the mesosphere layer. The mesosphere layer is challenging to analyze, so more limited is comprehended about this layer than other layers in the atmosphere. For example, Satellite orbits are beyond the mesosphere and cannot undeviatingly estimate traits of this layer.

2.1.4. Thermosphere Layer

The layer above the mesosphere is the Thermosphere layer. The thermosphere layer absorbs son's High-energy X-rays and UV radiation. The amount of energy originating from the Sun is a dominant impact on the height of the top of this layer and the temperature within it. The thermosphere can be observed anyplace between 500 and 1,000 km (311 to 621 miles) above the sea level on Earth.

2.1.5. Exosphere Layer

Despite the fact that some specialists accept the thermosphere to be the uppermost layer of our atmosphere, others examine the exosphere to be the real upper layer of the Earth's atmosphere. The upper bound of the exosphere is no clear because the exosphere subsequently fades gone into space. Exobase is the name of the bottom of the exosphere. The lower bound of the exosphere is deviates.

2.2. Ionosphere Layers Characteristics

In the atmosphere layer classification, the ionosphere is not a distinguished layer like the other layers discussed above. Alternatively, the ionosphere layer is a range of precincts in parts of the mesosphere and thermosphere. In this layer, high-energy radiation from the Sun has beaten electrons unhitched from their origin atoms and molecules. The electrically charged atoms and molecules produced in this manner are called ions. As a collection, these areas are collectively named the ionosphere layer.

The Sun's high-energy X-rays and ultraviolet uniformly crash by gas molecules and atoms in Earth's atmosphere. As mentioned, some of these crashes beat electrons and generate electrically charged ions (atoms with missing electrons) and free electrons. These ions and electrons flow and act abnormally, electrically neutral atoms and molecules. In the literature, the mentioned ions and free electrons transpire at various altitudes, as a group, as the ionosphere.

2.2.1. Ionosphere Layers

There are three principal areas of the ionosphere, named the F layer, the E layer, and the D layer. Depend on the different seasons, and during the progression of a day, the regions(D, E, and F layers) have various boundaries. By way of explanation, the regions have not clear boundaries.

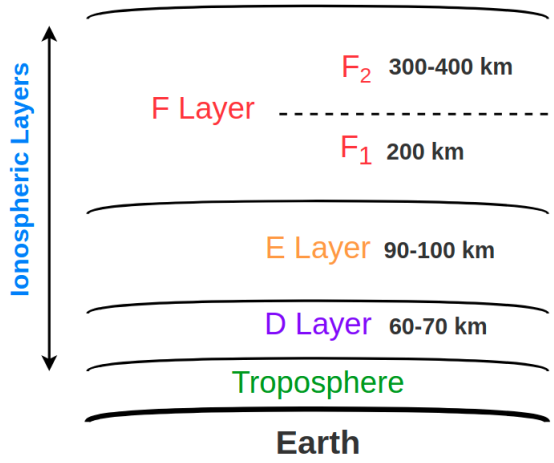


Figure 2.2. Layers of the ionosphere.

As depicted in the Figure 2.2. The lowest region in the ionosphere is the D layer that begins from 60 to 70 km (37 to 43 miles) above the sea level of the ground. The next layer after layer D is the E region or layer that is started at around 90-100 km (56-62 miles) and is ended approximately 120-150 km (75-93 miles). The next and the last region of the ionosphere is the F layer. This layer is the topmost area of the ionosphere that rises about 150 km and continues far higher, occasionally as high as 500 km over the Earth.

The mentioned sub-layers of the ionosphere layer are ionized areas installed within the standard atmospheric layers. As quoted, D region regularly appears in the uppermost portion of the mesosphere layer. The E region rises in the lower of the thermosphere layer, and the F region is located in the uppermost stretches of the thermosphere layer.

In different regions, the portion of ionized particles and the height are varied over time in various seasons. The ionosphere layer is entirely dissimilar in the daytime versus night. The

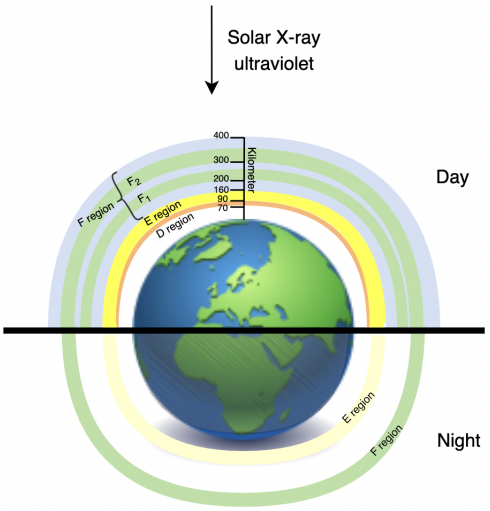


Figure 2.3. Detailed information of ionosphere layers.

ratio of ion production is higher than their elimination during the day as depicted Figure 2.3.. Its means that more ions are created throughout the day in the three regions. Due to the absence of high-energy X-rays and ultraviolet from the sunlight, the recombination process takes over, and the number of ions is dropped during the night. The D region vanishes completely at night, and the E region declines with the number of ions in that layer. As solar X-rays and UV light appearance, the D and E regions are repopulated with ions every morning. F layer usually divides into a lower F1 layer and an upper F2 layer during the day. however, the F region sticks almost throughout the night.

Radio communication systems frequently practiced the ionosphere to extend the range of their transmissions. Radio waves usually move in direct lines, so just a big transmission tower in terms of height can recognize the top of a receiver transmission tower. Consequently, the Earth's curvature restricts the range of radio transmissions to transmission stations that are not over the horizon. Nevertheless, parts of the frequencies of radio waves reflect off or jump the electrically charged particles in particular ionosphere regions. Before communication through satellites is popular, Radio communication systems used the advantage of this aspect, jumping radio waves off to the sky to extend the range of the signals.

As show in the Figure 2.4. [1] , due to the variations or removal of the regions between day and night in the ionosphere layer, Radio operators should consider the regular variations in the ionosphere layer to effectively benefit these mirroring reflections of radio waves.

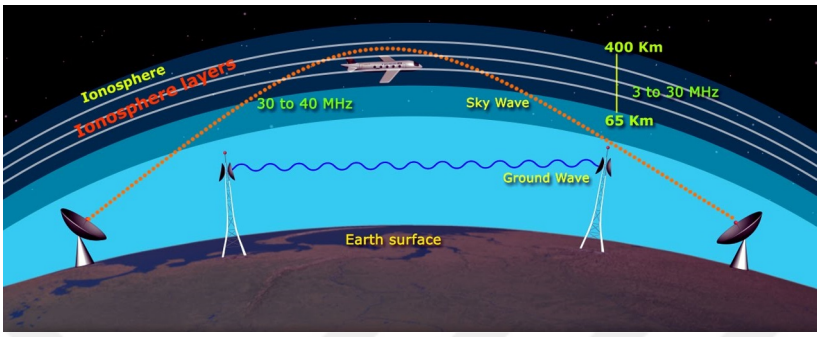


Figure 2.4. Radio communication with the ionosphere.

The ionosphere layers can reflect the signals as explained above, or the radio waves can be bend or absorb with regions in the ionosphere. Consequently, the particular behavior of radio signals depends on the frequency of the signal and the properties of the ionosphere layer involved.

Global Positioning System (GPS) satellites radio signals to arrange situations or locations, so the power of GPS signals be seriously decreased when the signals bend as they pass within ionosphere layers.Some of these signals depend on their frequency, can be absorbed entirely. Scientists regularly aim to produce calculator models of the fluctuations ionosphere so that radio operators in radio communications can predict disruptions. They monitor the constantly evolving ionosphere to handle radio waves in different directions, radar systems on the Earth, and on satellites.

2.3. Deep Neural Networks

Deep Learning is a subgroup of machine learning algorithms involved with methods inspired by the construction of the human brain named artificial neural networks. The theory of deep learning arose from the investigation of artificial neural networks (ANNs) [4]. ANNs have grown rapidly during the past few decades [5–9]. Nodes are small pieces of the neural

[1] Taken from https://qrznow.com/

network that are grouped into layers. The procedure is required to process layers of data between the input and output to solve a problem. In some problems, the more layer is needed to obtain the proper results, the deeper the network is supposed. It is named deep neural networks when we have more than two layers. In a deep neural network, there is the input layer that receives information, a number of hidden layers, and the output layer that stores estimable outcomes. It has been summarized the required information about deep neural networks for better understanding this thesis. This information has been separated as below.

- Theory of Deep Neural Network

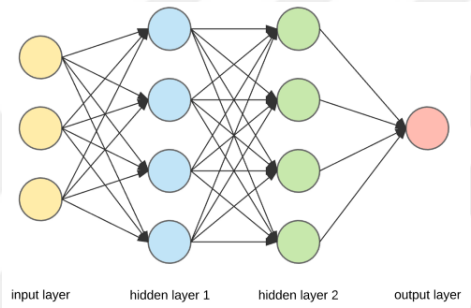


Figure 2.5. Deep Neural Network layers.

As illustrated in Figure 2.5., The Deep Neural Network is assembled from 3 types of layers. Input layer regularly is remarked as vector X. The nodes in the hidden layers represent the activation nodes and are usually stated as . The output layer is the values in the case of multiple predicted outputs. Each connection has a demanding weight that can be observed as the node's impact on the node from the next layer. Each node in the network looks as below.

For instance in a hidden layer node, all the input layer nodes are connected to it. As depicted in Figure 2.6., all nodes from previous layer are multiplied with their connection weight. The output is passed from "activation" function. The activation function decided the node to be active or not. The node with value 1 is named "bias" node.

- Activation Functions

The activation function determines if each node ought to be "activated" or not depend on the weighted sum. Activation functions are a crucial component of the architecture

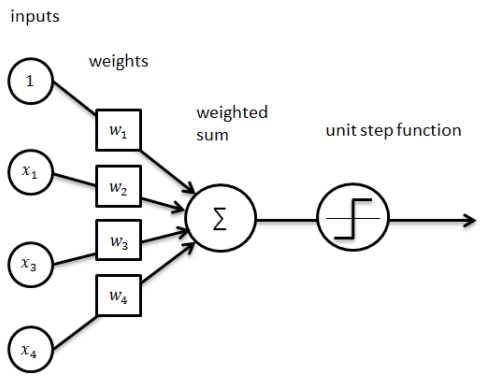


Figure 2.6. Node from Neural Network.

of a neural network. Successful training in the learning phase has a direct relation with the choice of correct activation function in the hidden layer. The activation function in the output layer should depend on the type of predictions the model has to perform. As a consequence, the activation functions in each layer should be chosen carefully to obtain the proper results in each problem. The activation functions that are used in thesis are Rectified Linear Unit (ReLU), Sigmoid and Softmax.

- Bias Node

  A bias node provides more varieties of weights to be deposited. Biases add a more harmonious description of the input space to the model's weights. It is true that ANNs can work without bias neurons. Nevertheless, they are usually added and counted as an essential element of the overall model. In summary, a bias value provides to shift the activation function to the left or right. The shifting of the activation functions promotes gaining a more desirable fit for the data.

- Loss Function

  The error or lost function describe the total of the error, difference between the predicted value and the original value. Loss functions are used to evaluate the error of the model; therefore, depending on the loss, the weights can be modified to decrease the loss on the next evaluation. As a deep neural network determines a mapping from the

input layer to the output layer, we need to choose a suitable loss function depends on the particular problem, such as classification or regression.

- Backpropagation Algorithm

  Backpropagation is the principle of the neural network training phase. It is a conventional process of training neural networks. This approach aims to fine-tune the weights using the loss rate achieved in the past iteration. The suitable adjusting of the weights permits the model to decrease loss rates and get a more reliable model.

  Backpropagation approach attends to estimate the gradient of an error function using all the weights in the network. As illustrated in Equation 1, it aims to decrease the cost function utilizing the optimal values for weights in order to calculate the partial derivative of cost function C(W). The calculated value is used to estimating the new weights (W) values for the network that reduce the cost function.

$$w_j = w_j - \alpha \frac{\partial}{\partial w_j} C(w) \tag{1}$$

  Where C(w), is the cost function using the optimal values for weight.

2.3.1. Design of the DNN Architecture

The machine learning algorithm can be possibly scaled up. These methods not only have been prosperous in such scale-up, but they have grown more robust[10]. For example, convolutional neural networks (CNN) [11] and recurrent neural networks [12] are the main subfields in deep learning neural networks (DNNs).

Deep neural networks are much more complex in terms of configuration. So there are a huge number of configuration hyperparameters and topologies during experimentation. In DNNs, accomplishment frequently depends on obtaining the proper architecture for the problem. Much of the contemporary literature in deep learning has concentrated on introducing various on-demanded designed architectures to new different problems [13–16].

The main problems of applying a DNN architecture are abstracted as three issues. The first is related to how we can design the components in the model. The second is how we can put the components together into network topology, and the last issue is how the set hyperparameters for the components in the leading architecture. These issues are needed to be optimized independently for every new problem. In conclusion, a novel idea of conceiving about such a design has begun to appear. Humans are held for the hand-level design of the architecture in these methods.

3. LITERATURE REVIEW

In this chapter, we review related works based on the conducted research. This section presents the related works for monitoring the ionospheric disturbances expresses that geomagnetic activity like storms and earthquakes can cause intense disorders in the electron density distribution and TEC. The satellite-based measurements from GPS stations have implemented a useful study to investigate the seismo-ionospheric anomalies as described in [17–19]. The articles [20, 21], investigated the ionospheric perturbation that storms can cause strong turbulence in the ionosphere's TEC values. Furthermore, the paper [22] reported earthquakes and seismic activity could cause the changes in electromagnetic signals and the chemical composition of the atmosphere in the lithosphere with the troposphere and ionosphere.

One of the initial research that proved the relationship between ionosphere fluctuations and seismic movement was published in [23]. This research analyzed preliminary data gathered from satellites that identified the consequences of an upcoming earthquake on the ionosphere layer. The research results revealed that the seismic flows cause some acute effects on the ionospheric layer, such as changing the attributes of ionospheric plasma in terms of heating, composition, and concentration.

Authors in [24], declared the consequences of a long-term investigation of fluctuations in the ionosphere in seismic movement regions before an earthquake. The research was aims to discover the foremost features of ionospheric precursors in earthquakes. The main goal was conceived to recognize features that could recognize earthquake precursors from changes that had different reasons. In this study, the critical frequency of the F2 layer of the ionosphere layer ($foF2$ parameter) is based. The results show that precede earthquakes can be recognized within five days and several hours before the earthquake. It means that earthquakes of magnitude five or higher can be detected. The main outcome was the determination that, given the measurement point and the location of the earthquake's epicenter, ionospheric precursors features will be alike for all consequent earthquakes with a near epicenter.

The research [25], related to the variances of $foF2$ were statistically validated. This study aims to extract event of irregularities in $foF2$ is related to a consequent earthquake and examines the effectiveness of $foF2$ irregularities as an earthquake precursor. In the statistical p-value test, $foF2$ irregularities ere earthquakes evets was discarded with p-value ≤ 0.0052. The test also verified the feasibility of the predictive model for earthquakes based on $foF2$ differences. The test also verified the feasibility of the predictive model for earthquakes based on $foF2$ differences. There are some studies related to $foF2$ values or the F2-layer critical frequency are exists to show that ionospheric precursors of earthquakes [26]. The value of $foF2$ depends on the geomagnetic and solar activity, seasons, and time of day. The value of $foF2$ rely on the geomagnetic and solar activity, seasons, and time of day.

Some researches are related to the defines the algorithms for detecting variations of $foF2$. For example, the authors in the papers [27, 28], proposed the matrices A_{ij}, which have 24 rows six columns, where the value of the i row and j column symbolizes the variation of $foF2$ by the middle value for the i hour of the j day. When the matrices are computed for six days before an earthquake, it turns out to be alike for earthquakes happening in the same areas. Therefore, the earthquakes are predicted when $foF2$ variation matrices are created for earlier earthquakes by the current matrix.

The writers in [28], also revealed the earlier defined approach that utilizes $foF2$ deviation matrices. In the newer research [29], they used $foF2$ deviation matrices with trivial adjustments. It is estimated ten days before and four days following an earthquake. It means the deviation matrices was calculated from the mean of 15 past values. This article is based on earthquakes in two regions of Italy and Greece. The results indicate the earthquake holds a pattern that repeats for a particular area.

The method suggested by the writers in [30], considers two ionosondes. The first one is in the earthquake preparation zone, and the other is about 500-700 km close to the first ionosonde (Outside that zone). Data are collected from these ionosondes and compared with each other in the earthquake days or the absence of any seismic movement. The result shows that the data will be extremely correlated in the lack of any seismic activity. In contrast, when an

earthquake happened, the correlation indicator is diminished. The main benefit of this work is that there is no necessity to process past data.

3.1. Scenario based Earthquake Prediction

Plenty of research like conducted works in [31–34] focus on statistical analyses about the experimental relationship between the ionospheric based irregularities and earthquakes. For example, the authors in [31] investigated statistically the TEC related to 20 strong (Magnitude $\geq$ 6.0) earthquakes in Taiwan during the four years (1999–2002). They find out abnormalities that indicate the TEC value are decreased within five days before the earthquakes through the other quiet days. Besides, Le et al. [32] illustrates the relation between ionospheric TEC based irregularities and strong earthquakes (Magnitude $\geq$ 6.0) during nine years (2002–2010). Moreover, they confirmed a high rate of abnormalities is happens in the ionosphere through earthquakes with Magnitude $\geq$ 7.0 and depth $\leq$ 20 km.

Recently Ulukavak, Pundhir, and Oikonomou et al. [35–37] have reported abnormal fluctuations in the ionospheric layer based on TEC values last days and several hours before severe or strong earthquakes. The TEC data has gathered from a GPS station in the earthquake zone. There are some uncertainty issues about generating such anomalies at the epicenter of the strong earthquakes. Tariq et al. [38], and Shah et al. [39] also stated direct relations between the ionosphere diversity and the existence of earthquakes. Their research demonstrated TEC values collected from the GPS receiver is fluctuated and increased before Magnitude $\geq$ 6.0 earthquakes occur during the long term of 1998–2014. As referred by authors in [32], severe or intense earthquakes according to their magnitudes were listed based on the ionospheric TEC irregularities.

In later times, the papers were presented to utilize other aspects of the ionosphere as earthquake precursors [40, 41]. Thus, for instance, the research [40], proves anomalies in TEC (total electron content). This research aims to extract anomalies by several days (2–9) before three earthquakes of magnitude $\geq$ 7.2. The authors in [41], just like in the earlier paper, a notable deviation was identified former to an earthquake. The results show $foF2$ parameter

would not be the exclusive earthquake feature. However, there is a demand for analytical support of observations that are only for singular events.

Following the conducted research in this field, we aim to analyze the prior days of the earthquakes using TEC values in the ionosphere layer and then classify the quiet and earthquake days in the target station zone. The main goal of this research is predicting the upcoming earthquakes in the prior days.

3.2. Precursors based Earthquake Prediction using Machine Learning and Statistical Models

Suggested approaches in the literature based on earthquake prediction can divide into two principal sections [42]. The first is model-based approaches that hypothesize a machine learning model or a statistical model to connected the earthquakes with seismic activities. The second is precursor-based approaches that focus on consideration of variations in the earthquake precursors.

The statistical models proposed in paper [43], that based on fault line strain to handle a periodicity in the earthquake. They believed that an earthquake depends on the strain collected close to a fault since a prior earthquake, and this manner reappears many times. The authors in [44], use Lucas and Fibonacci numbers in the happening times of previous earthquakes to forecast the expected earthquakes. In [45], it has been proposed a connection model that corresponds to happening times of earthquakes pattern throughout the fault zones. The older research is related to the empirical probabilistic model that has been introduced to predict the magnitude of forthcoming earthquakes [46].

The are plenty of researches using machine learning methods based on earthquake prediction by past seismic activity data like epicenter or magnitude [47–52]. Various machine learning approaches are used on past earthquakes to detect upcoming seismic activity based on seismic waveforms. These models are used random forest, k-nearest neighbors, support vector machine (SVM), and artificial networks. (Asencio-Cortés et al. [53]; Mahmoudi et

al. [54]; Moustra et al. [55]), In this research, we are focus on the most powerful Recurrent Neural Networks approaches such as LSTM-models to perform classification on quiet and earthquake days.

Some research depends on the measurement of variations in the earthquake precursors using the ionosphere layer fluctuations called precursor-based methods [56–60]. Some other precursor-based methods use various factors instead of the ionosphere layer, such as strange lights, unknown animal action, the chemical construction of underground water, radon gas emissions, and temperature [61–67].

The purpose of the author in [68] is to recommend possible methods between the electromagnetic state of the ionosphere and the seismic motion before the earthquake. The proposed model extracts the potential processes between the earth's ionosphere and the lithosphere. Gaseous radon outpouring ionizes the air in the atmosphere, leading to anomalies in the F2 layer ionospheric parameters, for example total electron (TEC), maximum ionization critical frequency (foF2) and maximum ionization height (hmF2). Another approach that depends on radon emissions is proposed in [69]. This work links the ionospheric anomalies to the Atmospheric Planetary Boundary Layer habits. The research believed that the lowest height of the upper boundary in the PBL causes the increase of radon gas that freed in the surface layer. This is the cause of the increase in the ionization quantity of the ionosphere layer at night. There are comprehensive examinations describing feasible correspondences between the noticeable irregularities in F2 layer ionospheric parameters and seismic movements. [27] study performed the statistical analysis on foF2 important frequencies estimated at Taiwan Island. It is explained ionospheric precursors to magnitudes greater or equal to four earthquakes between the time from 1978 to 1986.

There are additionally various researches that support anomaly detection approaches based on the ionospheric disturbances in ion temperatures, foF2, and TEC had been examined prior to severe earthquake nearby some country [70–81]. These studies based on machine learning and statistical models have been applied to impact seismic movements on the ionospheric layer irregularities, such as variation analysis and TEC difference or correlation investigation

among TEC and foF2 or various combinations of GPS receivers. The authors in [82] aim to perform objective research on the precursor detection of a TEC-based earthquake. The study reported similarity among ionospheric TEC irregularities and severe earthquakes using threshold TEC anomaly signals.

As mentioned, the machine learning algorithms are performed on ionospheric data to extract the relationship between earthquakes and them [83–89]. The research [90] was conducted to review the correlation between earthquakes and ionospheric magnetic field disturbances that supervised machine learning methods are used to identify active seismic areas from the magnetic field in the ionosphere. The model is based on Random Forest and K-Means to detect potential relationships among magnetic disturbances in the ionosphere and seismic movements. Another research related to earthquake precursor detection [85] applies correlation methods to identify the influence of severe earthquake movement on the ionosphere layer. The study aims to detect anomalies connected with geomagnetic storms and earthquakes that correlate with GPS stations. However, the method can be executed for severe earthquakes.

The research, based on genetic algorithms, is applied [86], to detect earthquake precursors, and another study [87], has been used Artificial Neural Network to detect earthquake precursors by TEC data. The reason of using machine learning techniques to detect earthquake precursors is to consider the TEC data of the learning pattern. In some scenarios utilizing machine learning, an irregularity on TEC data may happen before the earthquake. Applying machine learning techniques based on N-Model Articial Neural Network for detecting earthquake precursors has been made in Indonesia, particularly Sumatera [84].

The authors in [91] present a novel system that is called QuakeCast that detect short-term earthquake using global ionosphere TEC data. The proposed method using a classical logistic regression model and deep learning ConvLSTM autoencoder investigates whether signals foretell earthquakes in a TEC dataset of the ionosphere layer.

In the study [92], a gradient boosting-based technique called LightGBM is performed with a k-fold cross-validation test to discriminate electromagnetic pre-earthquake disturbances. The

study has been used the low-altitude satellite DEMETER for ionospheric disturbance investigation. The DEMETER reported various instances of ionospheric disturbances seen on the occasion of severe seismic phenomena. The conclusions of the research exhibit that the electromagnetic pre-earthquake data are discriminated adequately with the proposed method. Many investigations have been used DEMETER data to detect many perturbations before severe earthquakes happen. The authors in [93], used the DEMETER data to found many ionospheric perturbations before the Kii island earthquake with a magnitude of 7.3. The study [94] is related to the Wenchuan earthquake on 12 May 2008. The approach pronounced that the ion mass approached its lowest rates within three days before the Wenchuan earthquake.

Studies in [95] and [96], discovered plasma disturbance across the region of earthquakes and observed related plasma irregularities with electric field disturbances about 4-7 days before the earthquake. Ryu et al. [97] investigate the relationship between prior days of earthquake and ionospheric data and released that irregularities rose in the equatorial ionosphere approximately 25-30 days before the earthquake and sees the highest value in the eight days before the main earthquake. Based on the epicenter with the DEMETER data, the authors in the study [98] perform an analytical consequence of the boxplot model to obtain the outgrowths of the 1–6 days before the Wenchuan earthquake regarding ionosphere ion temperature and ion density fluctuations in the daytime and nighttime.

The study [99], and [100], investigated a statistical interpretation of plasma density fluctuations and noted a significant rise in the plasma density before some earthquake scenarios. Zhang et al. [101] mentioned that low-frequency electromagnetic changes are begun to emerge on a great range of latitudes and accomplished the highest after seven days before the earthquake. Ho et al. [102, 103] presented the impact concerning the abnormal enhancement of electron density, ion density, and ion temperature throughout the epicenter region of earthquakes. Additionally to the scenario-based earthquakes, many other investigations prove that the DEMETER ionospheric disturbances are valuable and fine-tuned for catching irregularities related to earthquakes [104–109].

The evidence in the literature has been confirmed that ionospheric disturbances are not only

affected by earthquakes. For instance, in some earthquake scenarios, ionospheric disturbances cannot be detected. However, some scholars worked to utilize statistical interpretation to eliminate the electromagnetic disruptions caused by non-earthquake origins. The authors in [110] and [111], work on a statistical model to examine the connection between seismic activities and equatorial irregularities recognized by DEMETER has been made applying equatorial plasma density and electric field measurement. He et al. [112, 113] present a statistical model to shows the irregularities increase when the magnitude is enhanced and the electron density raised in the nighttime close by the epicenter. In a similar study related to the ionospheric ion density, the authors in the [114], exhibits the ionospheric anomalies five days before an earthquake at 200 km from the epicenter. Other statistical investigations relating to the ion density peaks in the DEMETER data [115–119], confirmed that the turbulences are risen and then reduced continuously on the day of the earthquake.

As stated before, investigation on ionospheric irregularities before seismic activities usually depends on one or different explicit parameters, and complex consequences can be performed. Accordingly, the electromagnetic features of ionospheric irregularities before earthquakes need to be thoroughly investigated. Furthermore, numerous studies only refer to specific earthquakes, lack generality, and may represent various outcomes.

As mentioned, the ionospheric TEC data is a time series parameter. Ionospheric precursors regularly do not obey a systematic behavior. Consequently, it is commanding to generate machine learning-based methods for the detection of abnormal patterns in the investigation of earthquake precursors. A strong earthquake predicting approach depends on the particular earthquake parameters (magnitude, time, and location) gathered from the perceived ionosphere irregularities. The ionospheric TEC anomalies typically occur in the F-layer, E-layer, and D-layer and may be recognized until ten days before the earthquake [120]. For instance, the author in [121], develops an approach to identify the abnormal fluctuations of the TEC data by the severe Solomon earthquake utilizing genetic algorithm. It can be detected some irregular TEC based fluctuations on earthquake day.

In this thesis, we perform our purposes with an efficient differentiation interpretation of electromagnetic pre-earthquake disturbances on a large dataset of consecutive TEC data by handling machine learning and deep learning algorithms that are extensively applied in recent investigations of predicting earthquakes [122–126]. Furthermore, machine learning algorithms is suitable for finding out the various prominent features in distinguishing electromagnetic pre-earthquake disturbances.

4. Ionosphere TEC Based Dataset

This section has introduced the dataset related to earthquakes and TEC values of GPS stations. The dataset preprocessing steps are also discussed in the following.

4.1. Dataset

As mentioned in Section 1., TEC values obtained from GPS stations are a powerful technique for analyzing the ionospheric response to earthquakes and solar storms. TEC data gathered from GPS stations (Dual-Frequency GPS receiver) is used to investigate the ionospheric variability through moderate and severe earthquakes. For this research, TEC data has been collected from two GPS stations. This data has been collected from the IONOLAB group (Hacettepe University of IONOLAB is an organization of electrical engineers to investigate hurdles of the ionosphere.) [2]. The first station is located in coordinate $(Lat : -20.15, Lon : -70.13)$ Iquique in Chile represented in Figure 4.1.. The second station is located in coordinate $(Lat : -20.85, Lon : 117.1)$ Karratha in the Pilbara region of Western Australia represented in Figure 4.1.. The earthquake information is collected via (United States Geological Survey of Earthquakes) [3].

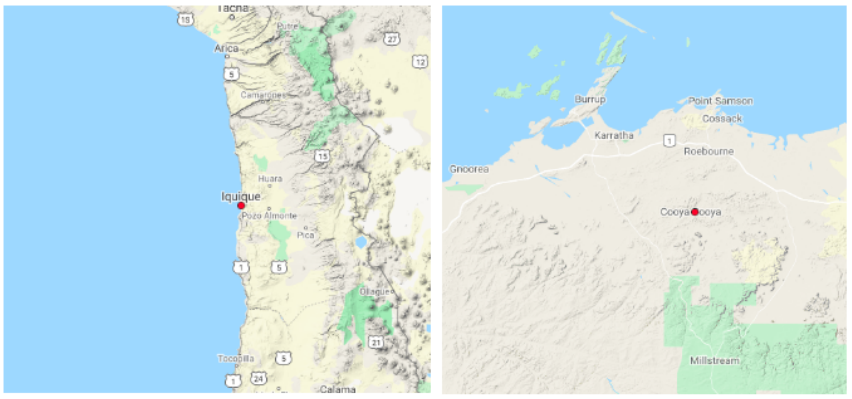


Figure 4.1. The Iqqe station is located in Coordinate (Lat:-20.15, Lon:-70.13) Iquique in Chile and the karr station located in coordinate $(Lat : -20.85, Lon : 117.1)$ Karratha in the Pilbara region of Western Australia.

[2] Available at http://ionolab.org/
[3] Available at https://earthquake.usgs.gov/

The ionospheric variability during moderate and severe earthquake events of varying strengths for 2012-2019 years is discussed in this paper. The data related to each station has been collected from 2012 to 2019. The data were separated day by day with 2880 TEC samples in a day. As illustrated in Figure 4.2., the stations are located in the same latitude from the two hemispheres of the east and west of the earth. Chile region has strong and severe earthquakes; however, the Karratha region is almost quiet in terms of earthquakes.

Figure 4.2. Iqqe and karr stations located in the same latitude from the two hemispheres of the east and west of the earth.

4.2. Data Preparation

In the machine learning concepts, the quality of the input dataset is so critical because there is a direct relationship between the quality of the input data and the performance of the trained model. Preprocessing of data is consists of data tuning techniques that modifying raw input data to an acceptable format. Ionospheric TEC data is often incomplete, inconsistent, and it is possible to contain many errors. Data preprocessing of this research is consists of three phases, including Cleaning, Transformation, and Reduction.

Since raw TEC data on specific days have missing and noisy values, the cleaning step attempts to eliminate missing values and make regressions with prior and later samples to smooth noisy values. Besides, the collected data is transformed into appropriate forms of mining. To perform the normalization, TEC values are scaled to small specified range (0,1.0). TEC data consists of 2800 samples in a day.

Table 4.1., illustrates detailed information about the ionospheric and the EQs(earthquakes) datasets. All earthquakes have gathered from stations that are placed in a radius of 250 km. The table shows the collected data as details.

Dataset Properties	Value
Number of days in uncleaned Dataset	2922
Number of days in cleaned Dataset	2571
Total number of earthquakes$\geq$4.5	141
Total number of earthquakes$\geq$5.0	91

Table 4.1. Prepared data for experimentation.

4.2.1. Similarity Learning Between Stations

The early studies have mentioned the ionosphere is affected by solar flares and other cosmic event factors. During the solar flares, X-ray fluxes are intensified that are identified as the cause of heightened ionization in the ionosphere. As this study focuses on ionospheric changes and the TEC variations during different earthquakes, it is needed to reduce the effect of solar flares and other cosmic events effect to recognize more correct earthquakes or geomagnetic activities.

As mentioned before, the stations are located at the same latitude from the two hemispheres of the east and west of the earth. Since the solar flares and other similar cosmic events affect the two hemispheres of the east and west of the earth, it can be calculated the similarity between two stations. The anomalies in the same days between the stations demonstrate the solar flares and other similar cosmic events. It can be supposed that have the same abnormalities due to the solar flares and the other cosmic events in the corresponding stations. It is required to calculate the similarity between the stations to decrease the effect of these anomalies. For this purpose, it is estimated the similarity between coincided days in each station in the dataset. Because of the nature of the dataset, it is considered cosine similarity to estimate the similarity between coincided days in each station. Cosine similarity is a metric between two non-zero vectors and is characterized by the cosine of the angles among the vectors. The cosine similarity is calculated as Equation 2.

$$\cos(x, y) = \frac{x.y}{\|x\| \, \|y\|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \sqrt{\sum_{i=1}^{n} y_i^2}} \tag{2}$$

Where x and y are the vectors of the TEC data related to each day in the stations. It is added Cosine similarity to ionospheric TEC data related to the Iquique station Data as a feature where the main aim is to interpret and detect earthquakes nearby Iquique Chile stations. As mentioned, we aim to extract similarities between stations to eliminate other effects except for earthquakes activities in the ionosphere. This research focuses on pre-earthquake disturbances on a large dataset of ionospheric TEC data by handling deep learning algorithms.

5. Evaluation Metrics Used in Methodology

This section discusses learning parameters and evaluation metrics used for evaluation and helps better understand the metrics required to do a quality evaluation. Model evaluation metrics are required to measure the performance of models. The choice of evaluation metrics are depends on the model.As the nature of data and proposed models, we use Confusion Matrix to explain the performance of classification models and calculate accuracy, precision, and recall as performance metrics.

5.1. Related Terminology in Confusion Matrix

- True positives (TP): The circumstances in which the model predicted correctly. The number of earthquakes that are predicted correctly in the classifier models. That means the classifier predicted the earthquake day, and it is true.
- True negatives (TN): The number of quiet days that are predicted as quiet days. In other words, the model predicted the quiet day, and it is true.
- False Positives (FP): The circumstances that the model predicted earthquake day; however, it is quite a day. It means the model predicted the earthquake day, and it is not true.
- False Negatives (FN): The number of earthquake days that are predicted as quiet days. Specifically, the model predicted the quiet day, and it is not true.

5.2. Accuracy Metric in Related Earthquake Classifier

Here accuracy is a common evaluation metric for classification models. It is defined as the number of correct classified earthquakes to the ratio of all classified days made as presented in Equation 3.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

5.3. Precision Metric in Related Earthquake Classifier

Precision is the ratio of correctly classified earthquake days to the total classified days as earthquake days. Precision is calculated as Equation 4.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

5.4. Recall Metric in Related Earthquake Classifier

The recall is the measure of our model correctly identifying earthquake days to the total of earthquake days. Recall is determined as Equation 5.

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

5.5. Metric Evaluation in Related Earthquake Classifier

Depending on the data and specifications of the classification model, it is needed to emphasize precision or recall. The nature of this research is the classification of ionosphere disturbances based on the earthquakes. The ionospheric disturbances may cause events other than earthquakes. Therefore in the classification model, the count of false positives is may be tolerable.

5.6. F1-score Metric in Related Earthquake Classifier

In generic classification models, between precision and recall values exist an association or trade-off. F1-score implies a measure of a test set efficiency that reflects both the precision and the recall to compute the score. The F1-score is the harmonious average of precision and recall in Equation 6.

$$F_1\text{–}score = 2 * \frac{precision * recall}{precision + recall} \tag{6}$$

5.7. Receiver Operating Characteristic (ROC) Metric

It can be beneficial to summarize each classifier's performance into a single measure to compare different classifiers. Receiver Operating Characteristic (ROC) curve is a common and standard measure with calculating the area under the curve to compare different classification models. As illustrated in Figure 5.2. the ROC curve depicted the relation between the true-positive rate and the false-positive rate. Classifiers that give curves closer to the top-left corner indicate a more trustworthy performance.

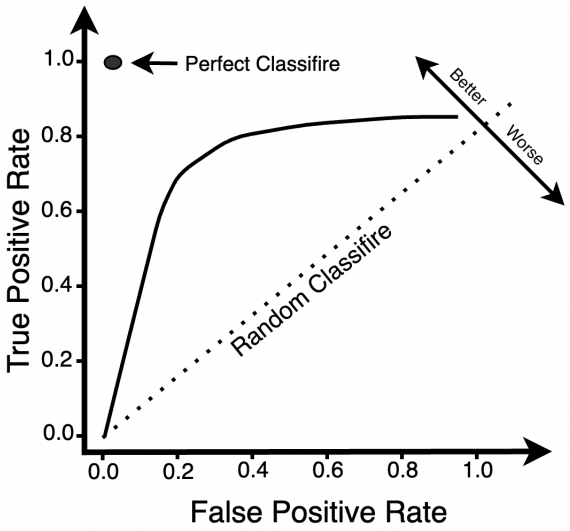


Figure 5.1. ROC table: detailed information.

In the ROC table, false positive rate and true positive rate (Sensitivity / Recall) are calculated as Equation 7 and 8.

$$True\ Positive\ Rate(TPR) = \frac{TP}{TP + FN} \tag{7}$$

$$False\ Positive\ Rate\ (FPR) = \frac{FP}{TN + FP} \tag{8}$$

5.8. Area Under the Curve (AUC) Metric

AUC (area under the curve) is an effective and combined measure of the true-positive and the false-positive rates for evaluating the intrinsic efficacy of the classification models. The larger AUC means diagnostic test is better in differentiating earthquake days from quiet days.

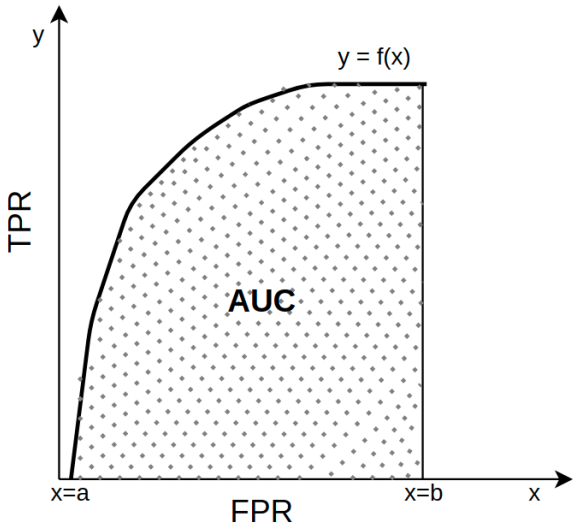


Figure 5.2. AUC table: detailed information.

The area under a curve among two points is obtained by performing a definite integral among the two points. To measure the AUC, $y = f(x)$ among $x = a$ and $x = b$, integrate $y = f(x)$. The area can be determined utilizing integration by given limits as illustrated in Equation 9.

$$AUC = \int_a^b f(x)dx \tag{9}$$

5.9. P-value Metric

The P-value is comprehended as the probability rate. It is described as the probability of obtaining one result representing the same observations or more significant than the actual observations. The P-value is recognized as the significance level in the hypothesis testing that describes the possibility of happening the offered event. The P-value is applied as an option to the denial point to present the most negligible significance of the null hypothesis. It is more substantial evidence in the alternative hypothesis if the P-value is small.

Ordinarily, the strength of analytical significance is frequently shown in P-value and the scale between 0 and 1. The P-value test is performed to decide whether the null hypothesis being tested can be rejected or not. The P-Value is the possibility of observing the effect(E) when the null hypothesis is true as Equation 10

$$P_Value = p(E \mid H_0) \tag{10}$$

The P-value table displays the hypothesis arguments:

- P-value $>$ 0.05 - This range shows the model is not statistically significant, and the null hypothesis is not rejected.
- P-value $<$ 0.05 - The result proves the model is statistically significant. Regularly, decline the null hypothesis in the alternative hypothesis.
- P-value $<$ 0.01 - The result confirms the proposed model is extremely significant and rejects the null hypothesis in the alternative hypothesis.

5.10. Cross-validation

For estimating the proposed classification methods and extracting machine learning algorithms' abilities, a statistical technique regularly used is Cross-validation. It is used to analyze and choose a proper model for a machine learning based predictive modeling issue. Cross-validation techniques are straightforward to implement, uncomplicated to understand, and usually have a lower bias than different techniques. As the earthquake data sample is limited count, to evaluate proposed classification models, it is needed to use K-fold cross-validation and Stratified k-fold cross-validation techniques.

5.10.1. K-fold cross-validation

K-fold cross-validation has been applied to handle the imbalanced classes problem. The k-fold cross-validation uses the repeated random sampling technique to evaluate model performance by dividing the data into n equal folds and assessing the performance of the model on each fold.

The value of n in the k-fold is determined such that every train/test group of earthquake samples is great enough to be statistically representational of the earthquake and ionospheric TEC dataset. In this thesis, the value for n is set to 10, which has been determined through experimentation to regularly result in the model ability assessment with low bias a modest variance.

5.10.2. Stratified K-fold cross-validation

The correct use of the k-fold cross-validation in an imbalanced class distribution problem requires to use of stratified k-fold cross-validation. Specifically, it can split randomly, although in such a way that keeps the same class distribution in each subset as depicted in the Figure **??**. For analysis of the days, a ten-fold cross-validation plan is used for testing. The dataset

is randomly divided into ten equal folds that each fold is an appropriate representative of the original dataset.

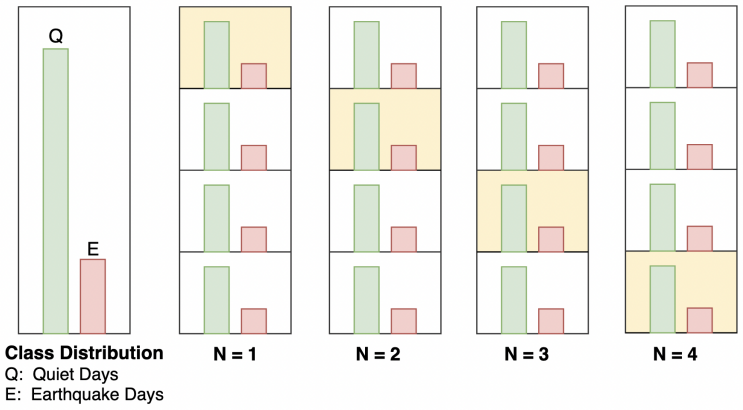


Figure 5.3. Four folds Stratified Cross Validation

6. Interpretation of earthquakes based on ionosphere disturbances using deep neural network

In this chapter, the architecture of the proposed model is presented. Figure 6.1. depicts a general overview of the proposed model. Due to the large amount of the TEC data related to each station, it is used deep neural networks. Since the main advantage of Deep Learning is the investigation and learning of tremendous amounts of supervised and unsupervised data, so it is valuable for Big Data Analytics, where data is mostly unlabeled. In the proposed model, unsupervised learning techniques like Deep Autoencoders and deep dense networks are used.

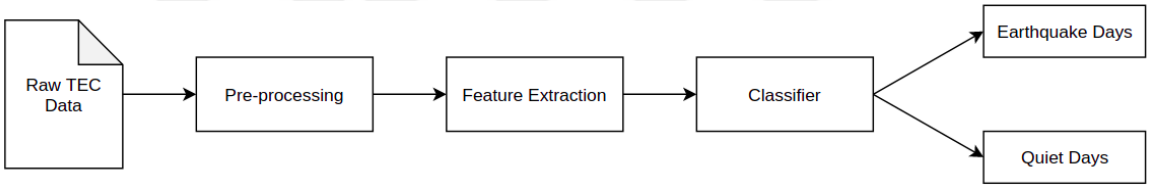


Figure 6.1. A general overview of the proposed model.

6.1. Feature Extraction

In the data-mining, learning large amounts of input data that include many attributes leads to impractical or impossible analysis, and the training phase may take long times. Data Reduction includes practical techniques to reduce the dataset's input samples without hazarding the original data's integrity. The fundamental feature extraction methods discussed by Liu et al. [127] are introduced on the basis of projection and data mapping in the original complex input space that has many dimensions to a new output space with lower dimensions by decreasing data loss. The typical popular projection methods are Principal Component Analysis (PCA), and Linear Discriminant Analysis (LDA) described by Khalid et al.[128, 129]. In the PCA method, via maximizing variance, the initial data are projected into its major directions. This method is categorized as an unsupervised method. In contrast, the LDA performs optimizing discriminating data from input classes to obtain a linear output space. The LDA

method is categorized as a supervised learning method. The main issue of the mentioned approaches is projecting linearly. Sharma et al. [130] offer non-linear kernel functions to reduce this issue.

As the ionospheric TEC data is a high-dimensional factor, reducing dimensionality to obtain a compressed feature set is considered as an essential step in the feature extraction phase. Theoretically, the traditional machine learning approaches be capable of acting on any amount of attributes. In contrast, these models with high-dimensional datasets always bring some obstacles such as over-fitting of the training set, high computational complexity, and the curse of dimensionality. For decreasing the input dimensions, autoencoders utilize neural networks by aiming to minimize reconstruction loss. Therefore, adding hidden layers to the autoencoders conducts to perform proper dimension reduction process. Deep Autoencoders have confirmed their effectiveness in detecting non-linear features across many complex problems.

Deep Autoencoders are multi-layer neural networks in which the wanted output is the input itself. In particular, an autoencoder learns a map from the input to itself within a pair of encoding and decoding steps. Autoencoders have latterly pointed in unsupervised methods for feature learning. It tries to learn a compressed description of the input while retaining the most important information.

$$\bar{X} = Dcr(Enc(X)) \tag{11}$$

Where X in Equation 11 is the input TEC data, Enc is an encoding map from the TEC data to the hidden layer, Dcr is a decoding map function of the code layer to the output layer, and $\bar{X}$ is the retrieved a similar version of the TEC data. The objective is to train Enc and Dcr to minimize the difference between X and $\bar{X}$. The encoder and the decoder functions (Enc and Dcr) are trained simultaneously to minimize the error of reconstruction of input from

hidden code nodes. Specifically, an autoencoder can be observed as an explanation for the optimization problems.

$$\min_{Dcr,Enc} \parallel X - Dcr(Enc(X)) \parallel \tag{12}$$

In Equation 12, $\parallel . \parallel$ is commonly considered to be the $\ell_2 - norm$.

Figure 6.2. depicts input, encoder, decoder and the output layer. The smaller reconstruction error indicates that the autoencoder can render a more compact vector as input vector. An autoencoder can be established with any activation function. In this research, it is used Sigmoid and ReLU as defined in Equations 13 and 14.

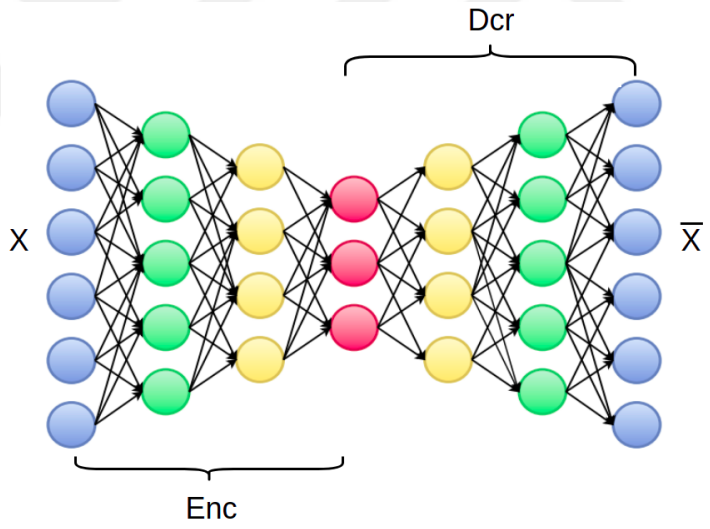


Figure 6.2. A Deep Autoencoder with five layers.

$$\varphi(z) = \frac{1}{1 + e^{-z}} \tag{13}$$

$$R(z) = max(0, z) \tag{14}$$

The output of the xth node in the ith layer is obtained sequentially from the output of the prior nodes in the previous layer as Equation 15 where $bias_x^i$ is bias scalar and N_i represents the number of nodes in the ith layer, $w_{x,n}^i$ is the weights which connect the xth node in the ith layer to the nth node in previous layer.

$$O_x^i = \varphi(net_x^i) = \varphi(\sum_{n=1}^{N_{i-1}} w_{x,n}^i O_n^{(i-1)} + bias_x^i) \tag{15}$$

Increasing the layers enhances the ability to learn more complicated patterns, evaluation time and network's complexity. As demonstrated in Equation, the matrix of all weights W are adjusted to minimize the mean square error across the training set.

$$\varepsilon = \sum_{i=1}^{N} \parallel X - Dcr(Enc(X)) \parallel^2 \tag{16}$$

6.2. Classification Method

After the feature extraction with Deep Autoencoder, it is needed to classifying these features. In the literature, several varieties of classification methods can be applied. The softmax regression is regularly applied in conjunction with deep neural networks. For this purpose, it is feasible to combine the classifier with the encoder function. Softmax (multinomial logistic) regression induces logistic regression. It calculates the probability of the ith class as in Equation 17.

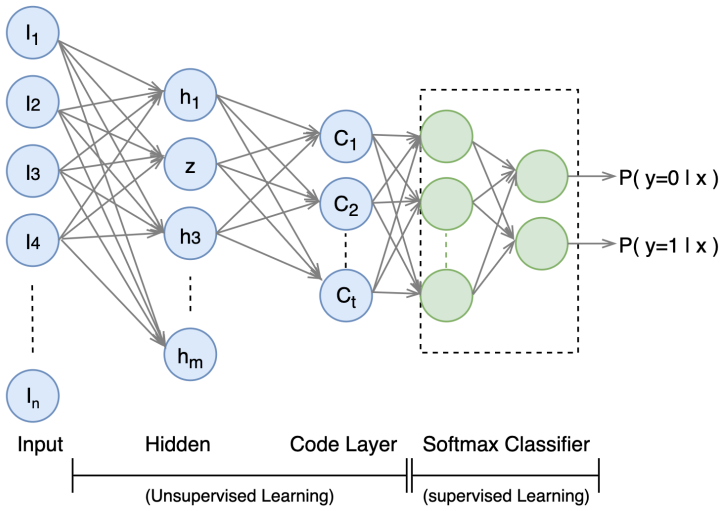


Figure 6.3. The general architecture of the proposed DAEclass model.

$$Prob_i = Prob(i|I) = \frac{e^{w_i I}}{\sum_{n=1}^{N} e^{w_i I}} \qquad i = 1, ..., N \tag{17}$$

Where w_i are training weights for the ith class and I is the input to the classifier. Then, the classification is performed by comparing $Prob_i$'s. The softmax can easily be combined to the encoder function and constitute the deep structure of the network.

Block diagram of the proposed autoencoders illustrated in Figure 6.3.. It is a consolidated model with unsupervised learning and supervised learning. Unsupervised learning is conducted with the encoder that has been trained before in the feature extraction phase, and the supervised model is a dense softmax classifier. Features at hidden layer two are decreased and compressed by minimizing the error rate using Equation 16. These features are then served as input to the next autoencoder layer, and the features at layer three (code layer) are fed to the softmax classifier, where perform classification with labeled data. The layers of Deep Autoencoder are trained independently in the feature extraction phase. Hence, the features are learned in an unsupervised manner, while classification is conducted as supervised.

Table 6.1. presented the detailed information about network architecture parameters in the deep autoencoder and classifier.

Property	Deep Autoencoder	Dense Classifier
Activation function	ReLU and Leaky ReLU	Sigmoid
Optimizer	Adam	SGD
Layer Count	3	2
Loss Function	BinaryCrossentropy	BinaryCrossentropy

Table 6.1. Detailed information about proposed network architecture.

As stated in the table, it used ReLU and Leaky ReLU in the deep autoencoder as activation functions. As illustrated in the Figure 6.4., instead of altogether zero in negative values, Leaky ReLU has a little slop. The reason for using Leaky ReLU instead of ReLU in some layers in the deep autoencoder is to retaliate against the "dying ReLU" dilemma and speeds up the training phase. Leaky ReLU raises the training time of the model. It has the "mean activation" be close to zero, and it does training quicker.

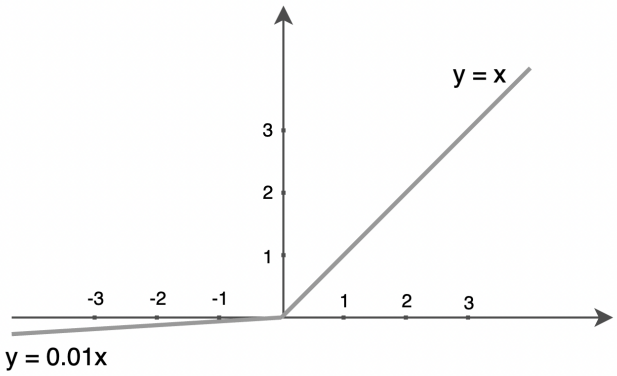


Figure 6.4. Leaky ReLU chart

The optimizer algorithm used in the deep autoencoder side is Adam, and in the classifier side is stochastic gradient descent (SGD). Adam is an optimizer that has been composed predominantly for deep neural networks algorithms. Adam optimizer estimates learning rates for various parameters as known as an adaptive learning rate approach. It is used estimations moments of the gradient to adaptively optimize the learning rate for every weight of the deep neural network.

Gradient descent is the favored approach to optimize neural networks algorithms and various different machine learning methods. Stochastic gradient descent is utilizing single records

to updates parameters. SGD optimizer requires forward and backward propagation for each record. So it is slow to coverage.

6.3. Dataset for DAEclass Method

The table shows the collected data as details such as the ratio of the train and test sets. The splitting of the dataset to train and test set is used to predict the performance of machine learning algorithms. As mentioned in Table 6.2., For a more considerate interpretation of the earthquakes, it has been divided the earthquakes to two subsets. The first is the set that contains moderate and severe earthquakes (EQs $\geq$4.5) and the second that contains only severe earthquakes (EQs $\geq$5.0).

Dataset Properties	Value
Train-Test Ratio	80%-20%
Number of days in Trainset	2057
Number of days in Testset	514
Number of EQs $\geq$4.5 in Trainset	113
Number of EQs $\geq$5.0 in Trainset	73
Number of EQs $\geq$4.5 in Testset	28
Number of EQs $\geq$5.0 in Testset	18

Table 6.2. Prepared data for proposed DAEclass Model.

6.4. Evaluation Methodology

This section discusses learning parameters and evaluation metrics used for evaluation and helps better understand the metrics required to do a quality evaluation.

The TEC values separated by days are classified with the proposed model in the method section. The proposed model is divided into two sub-models. The first sub-model extracts the features from days using a Deep Autoencoder. When the feature extraction phase is trained as an autoencoder, it produces a low dimensional representation that encodes a relevant topological structure of TEC features to reconstruct the high dimensional input. The second is a softmax dense network that is pinned to the encoder to perform classification tasks. In the

classifier, the hidden layers are constrained by the earthquake labels of the data. The network can also be two trained by mixing the autoencoder and classifier.

As stated in Table 6.2., it is apportioned the dataset into train and test sets, with an 80-20 split ratio. In the dataset, there is no proper proportion between the number of quiet days and the number of earthquake days. Therefore there are only 141 earthquakes (Magnitude $\geq$ 4.5) between 2012 to 2019. In the classification model, we have a small number of earthquake class. In the classification model, it has been a small number of earthquake classes, which causes a problem called the unbalanced classification in the model. The class imbalance problem occurs when a class is comparatively rare as compared with other classes. Many methods have been proposed for imbalanced classification by Lopez-Paz, and Makki et al. [131, 132], and some good results have been reported. For analysis of the days, a ten-fold stratified cross-validation plan is used for testing. The dataset is randomly divided into ten equal folds that each fold is an appropriate representative of the original dataset.

To examine the importance of the proposed combination model, we used LDA (Linear Discriminant Analysis) classifier based on the reduction of the features as conducted work by authors [133, 134]. Linear discriminant analysis is performed utilizing a max function M being a classification rule.

$$f_i(X) = \frac{1}{(2p)^{\frac{1}{2}} |\sum|^{\frac{1}{2}}} exp\left[-\frac{1}{2}(X-\mu_i)^T \sum^{-1}(X-\mu_i)\right] \tag{18}$$

$$M(X) = X^T \sum^{-1} \mu_i - \frac{1}{2}\mu_i^T \sum^{-1} \mu_i + log(p_i) \tag{19}$$

The conditional density of X in class i is $f_i(X)$, p_i is the probability of class i. It is supposed that the vector of features X is variable and it is distributed with mean vector μ_i and common

covariance matrix $\sum$. Next, $f_i(X)$ can be calculated as demonstrated in Equation 18. $M(X)$ discriminant function is computed as Equation 19 using the Bayes rule.

6.5. Evaluation Results

Using the evaluation methodology defined in Section 6.4., performances of our proposed model called DAEclass (Deep Autoencoder Classifier), LDA classifier, SVM (Support Vector Machine) classifier and RF(Random Forest) classifier are evaluated. This section is organized to compare and examine the performance metrics of the proposed classification model and other classifier.

EQs	Model	Accuracy	Precision	Recall	F1-Score
	DAEclass	0.932	0.563	0.753	0.644
≥4.5	LDA	0.894	0.303	0.637	0.408
	SVM	0.885	0.253	0.572	0.341
	RF	0.863	0.233	0.582	0.338
	DAEclass	0.964	0.615	0.836	0.707
≥5.0	LDA	0.946	0.357	0.718	0.474
	SVM	0.926	0.263	0.651	0.371
	RF	0.913	0.245	0.616	0.346

Table 6.3. Comparison of the DAEclass and other classification models using performance metrics.

Table 6.3. describes the median percentage of performance metrics like accuracy, recall, precision, and F1-score for the proposed DAEclass model and other classifiers based on two datasets. The first five rows illustrate the comparison between the proposed model and the other classifier models based on all earthquakes with dataset that its magnetite ≥ 4.5. Detailed information about ten-fold cross-validation related to the comparison of DAEclass and other classification models using performance metrics are presented in Table 6.4. and 6.5..

It is evident that the proposed model has higher performance in all metrics than the other classifiers. The proposed model boosted its precision and recall much more than the LDA classifier. The ionosphere layer disturbances reflect all cosmic and seismic activities. As the false positive count is high due to the imbalanced classes in the dataset, precision is lower

than recall and accuracy. The last five rows compare the proposed model and other models based on more strong earthquakes (magnetite-EQs ≥ 5.0). Since strong earthquakes affect the ionosphere more than moderate earthquakes, they are classified easily. The performance gap between the two models is smaller than EQs≥ 4.5 dataset.

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
DAEclass EQs≥4.5	**TP**	21	25	20	18	19	24	17	16	25	23
	FP	41	30	20	16	6	11	26	28	5	12
	FN	7	3	8	10	9	4	11	12	3	5
	TN	445	456	466	470	480	475	456	458	481	474
	Precision	0.33	0.45	0.5	0.52	0.76	0.68	0.39	0.36	0.83	0.65
	Recall	0.75	0.89	0.71	0.64	0.67	0.85	0.60	0.57	0.89	0.82
	Accuracy	0.90	0.93	0.94	0.94	0.97	0.97	0.92	0.92	0.98	0.96
LDA EQs≥4.5	**TP**	17	16	19	21	15	15	17	18	20	19
	FP	39	56	49	29	71	51	44	42	33	46
	FN	11	12	9	7	13	13	11	10	8	9
	TN	447	430	437	457	415	435	442	444	453	440
	Precision	0.30	0.22	0.27	0.42	0.17	0.22	0.27	0.3	0.37	0.29
	Recall	0.60	0.57	0.67	0.75	0.53	0.53	0.60	0.64	0.71	0.67
	Accuracy	0.90	0.86	0.88	0.92	0.83	0.87	0.89	0.89	0.92	0.89
DAEclass EQs≥5.0	**TP**	16	15	16	14	16	17	12	15	16	14
	FP	11	17	12	7	10	5	9	14	9	8
	FN	2	3	2	4	2	1	6	3	2	4
	TN	485	479	484	489	486	491	487	482	487	488
	Precision	0.59	0.46	0.57	0.66	0.61	0.77	0.57	0.51	0.64	0.63
	Recall	0.88	0.83	0.88	0.77	0.88	0.94	0.66	0.83	0.88	0.77
	Accuracy	0.97	0.96	0.97	0.97	0.97	0.98	0.97	0.96	0.97	0.97
LDA EQs≥5.0	**TP**	11	12	14	11	14	13	15	12	15	10
	FP	25	38	18	41	16	21	16	20	17	44
	FN	7	6	4	7	4	5	3	6	3	8
	TN	471	458	478	455	480	475	480	476	481	452
	Precision	0.30	0.24	0.43	0.21	0.46	0.38	0.48	0.37	0.46	0.18
	Recall	0.61	0.66	0.77	0.61	0.77	0.72	0.83	0.66	0.83	0.55
	Accuracy	0.93	0.91	0.95	0.90	0.96	0.94	0.96	0.94	0.96	0.89

Table 6.4. Detailed 10 fold comparison of DAEclass and LDA classification models using performance metrics.

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
SVM EQ≥4.5	**TP**	15	14	16	15	16	18	15	16	17	19
	FP	62	51	49	39	47	43	50	47	39	41
	FN	13	14	12	13	12	10	13	12	11	9
	TN	424	435	437	447	439	443	436	439	447	445
	Precision	0.19	0.21	0.24	0.27	0.25	0.29	0.23	0.25	0.30	0.31
	Recall	0.53	0.5	0.57	0.53	0.57	0.64	0.53	0.57	0.60	0.67
	Accuracy	0.85	0.87	0.88	0.89	0.88	0.89	0.87	0.88	0.90	0.90
RF EQ≥4.5	**TP**	16	15	19	16	15	19	16	15	16	17
	FP	58	54	50	48	56	54	51	49	47	47
	FN	12	13	9	12	13	9	12	13	12	11
	TN	428	432	436	438	430	432	437	437	439	439
	Precision	0.21	0.21	0.27	0.25	0.21	0.26	0.23	0.23	0.25	0.26
	Recall	0.57	0.53	0.67	0.57	0.53	0.67	0.57	0.53	0.57	0.60
	Accuracy	0.86	0.86	0.88	0.88	0.86	0.87	0.87	0.87	0.88	0.88
SVM EQ≥5.0	**TP**	10	10	12	13	11	9	11	12	15	14
	FP	34	45	28	35	28	24	29	34	28	36
	FN	8	8	6	5	7	9	7	6	3	4
	TN	462	451	468	461	468	472	467	462	468	460
	Precision	0.22	0.18	0.3	0.27	0.28	0.27	0.27	0.26	0.34	0.28
	Recall	0.55	0.55	0.66	0.72	0.61	0.5	0.61	0.66	0.83	0.77
	Accuracy	0.91	0.89	0.93	0.92	0.93	0.93	0.92	0.92	0.93	0.92
RF EQ≥5.0	**TP**	11	10	11	11	10	10	12	11	12	12
	FP	41	40	34	33	36	29	34	31	29	39
	FN	7	8	7	7	8	8	6	7	6	6
	TN	455	456	462	463	460	467	462	465	467	457
	Precision	0.21	0.2	0.24	0.25	0.21	0.25	0.26	0.26	0.29	0.23
	Recall	0.61	0.55	0.61	0.61	0.55	0.55	0.66	0.61	0.66	0.66
	Accuracy	0.90	0.90	0.92	0.92	0.91	0.92	0.92	0.92	0.93	0.91

Table 6.5. Detailed 10 fold comparison of SVM and Random Forest classification models using performance metrics.

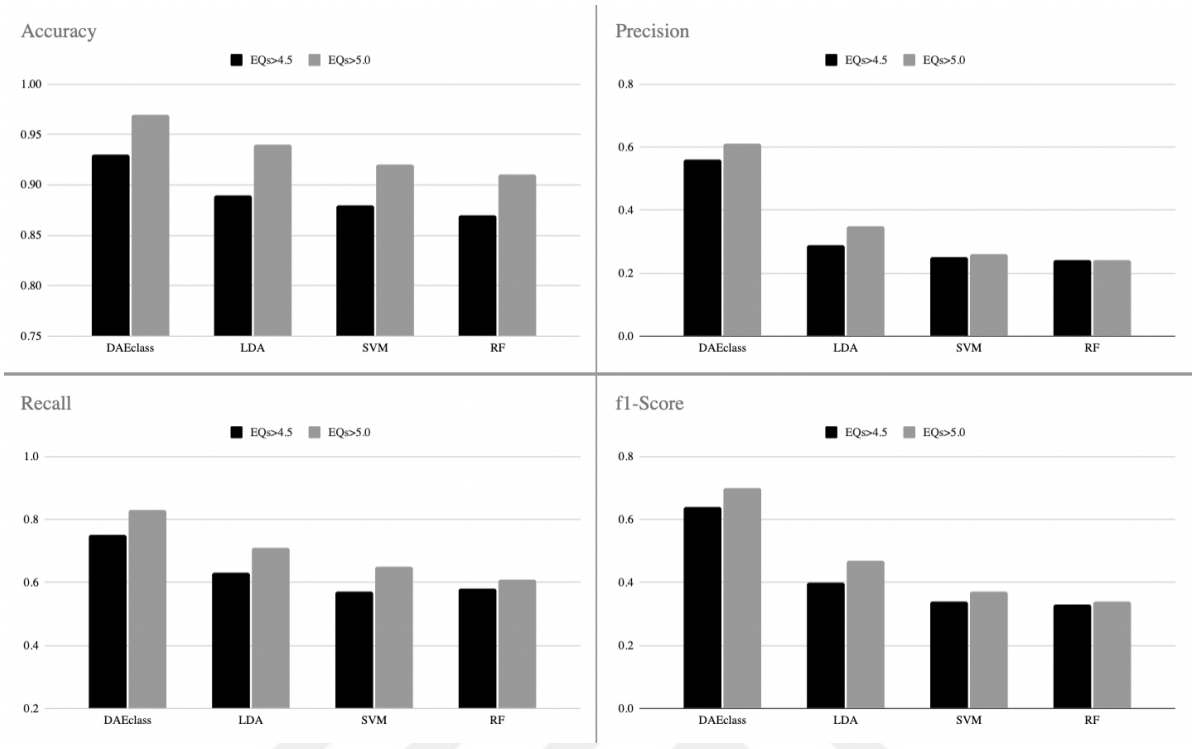


Figure 6.5. Performance metrics of the DAEclass and other classifier models based on two datasets (EQs$\geq$ 4.5 and EQs$\geq$ 5.0).

Figure 6.5. is the bar chart of performance metrics between four models in terms of two datasets (EQs$\geq$4.5 and EQs$\geq$5.0). It can be seen that in Figure 6.5. (A) the DAEclass's accuracy grows slightly in the two mentioned datasets. Precision and recall metrics are a useful measure of success in classification when the classes are imbalanced. Precision is a measure of classification result relevancy based on earthquakes and false alarms, while recall is a measure of the classifier's ability to find all the earthquake days. The comparison between datasets shows that the DAEclass is more stable than the LDA classifier and other classifiers in performance. Overall, Figure 6.5. (B)(C)(D) reveals the DAEclass approach notably increases the precision, recall, and F1-score. Consequently, the improved performance in the dataset EQs$\geq$4.5 is more extensive than the dataset EQs$\geq$5.0, and the DAEclass is more reliable in all earthquakes.

As mention in Chapter 5., Receiver Operating Characteristic (ROC) curve depicted the relation between the true-positive rate and the false-positive rate. Classifiers that give curves closer to the top-left corner indicate a more trustworthy performance. Figure 6.6. depicts the

ROC curve of the DAEclass and the LDA classifier models based on two datasets. As shown in the results, LDA is more accurate than SVM and Random Forest. So In the ROC table, we compare DAEclass and LDA classifier.

ROC curves related to the DAEclass positioned progressively closer to the upper left angle in ROC space, So DAEclass has the progressively more prominent discriminant capacity of earthquake classification. As depicted in the figure, the DAEclass, and the LDA classifier visually be compared simultaneously, and the results show the DAEclass is more efficient than the LDA classifier.

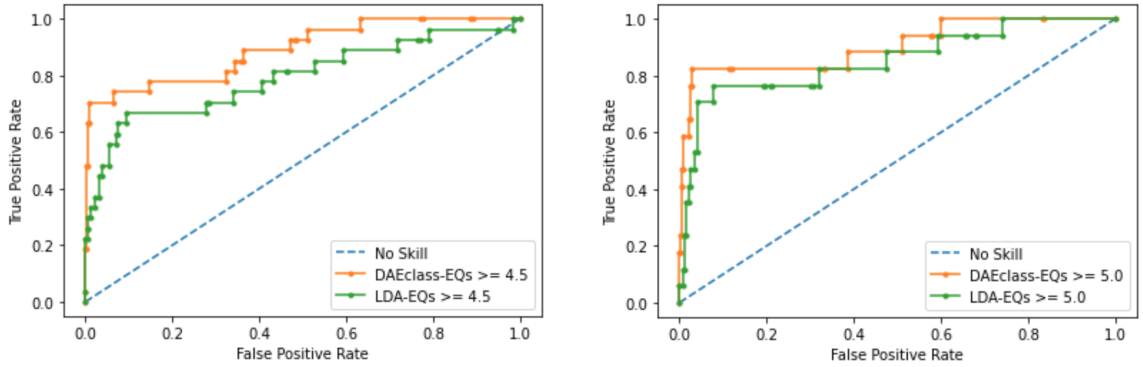


Figure 6.6. ROC curve of the DAEclass and the LDA classifier models based on two datasets (EQs≥4.5 and EQs≥5.0).

The results in Table 6.6. demonstrates AUC values of the model for each dataset. The value of AUC in the DAEclass-EQs≥4.5 is 0.89, while in the LDA-EQs≥4.5 classifier is 0.79, which means a keen improvement in the ROC metric. The DAEclass approach has a slight improvement against the LDA classifier on dataset EQs≥5.0 in AUC values from 0.85 to 0.91.

EQs	Model	ROC AUC
≥4.5	DAEclass	0.892
	LDA	0.794
≥5.0	DAEclass	0.914
	LDA	0.856

Table 6.6. AUC values for the DAEclass and the LDA classifier.

For further inspection of the proposed model, a statistical test is used to search for the level of significance and improvement. In the test result of the DAEclass model, the difference in performance over all metrics is significant $(p - value \ll 0.01)$.

6.6. Conclusion for DAEclass model

This research proposes an approach to interpret earthquakes based on ionospheric TEC values. Ionospheric TEC data has been collected from two GPS stations. Chile region has strong and severe earthquakes; however, the Karratha region is almost quiet in terms of earthquakes. The principal intention of this research is to attain the relationship between earthquakes and the ionosphere disturbances based on deep learning techniques concerning two sub-tasks of feature extraction and classification. In the first step, we focus on feature extraction of the TEC data using an unsupervised learning method. We design a Deep Autoencoder to extract useful information about the earthquake and quiet days. Then, in the next step, we utilize a supervised learning method to perform classification using a dense neural network. To investigate the contribution of the recommended combination model, we used LDA, SVM and Random Forest classifiers. Our results indicate approximately 90-94 accuracy-based performance in the two test sets of the earthquakes, including moderate and severe earthquakes. The proposed DAEclass has a more trustworthy and reliable performance than other models in terms of accuracy, recall, precision, F1-score, and ROC curve metrics.

Since in this model the main goal is extracting vital features of earthquakes using TEC values in the ionosphere layer and classifying the earthquake days in the target station zone, earthquake predicting is not the scope and priority of this model. In the fallowing chapters, we plan to work on a predictive system to predict ionosphere disturbances based on earthquake using earlier days of earthquakes using LSTM based deep learning models.

7. LSTM-based deep learning methods for prediction of earthquakes using ionospheric data

Our goal is to propose a classification model to detect earthquakes in previous days. The ionospheric variability during moderate and severe earthquake events of varying strengths for 2012-2019 years. The proposed models use LSTM-based (Long Short-Term Memory) deep learning models to classify earthquakes days by analyzing TEC values of the last days.

In this chapter, the architecture of the proposed LSTM based models are presented. Due to the large amount of the TEC data related to each station, it is used deep neural networks. As the TEC data fluctuations in prior days of the earthquakes play a critical role in the prediction phase, it is planned to implement a sequential learning approach. In the literature, the analysis of learning algorithms for sequential data is sequential learning. In our sequential learning approach, the sequential dependency between TEC data is analyzed at the algorithmic level. The sequence learning approaches are most generally based on sliding window methods and recurrent sliding windows. Sliding window-based methods disregard the relationship between data points inside the windows. However, model-based approaches, like the Markov chain defined on the data points that suppose sequential dependency between consecutive data.

7.1. LSTM-based Sequential Learning Models

Recurrent Neural Network (RNN) is a non-probabilistic model in that nodes satisfy the Markov-chain assumption. An RNN network predicts the output label given the sequence of the TEC data from the past. One of the more accurate variants of the RNN methods is Long Short-Term Memory Networks (LSTM) that learn long term dependencies in sequences data. Due to the TEC data sequence related to previous days is long enough, RNN's may leave out critical information from the beginning. The LSTM methods handle this issue with the solution to short-term memory. These methods can adjust the flow of information with internal mechanisms called gates.

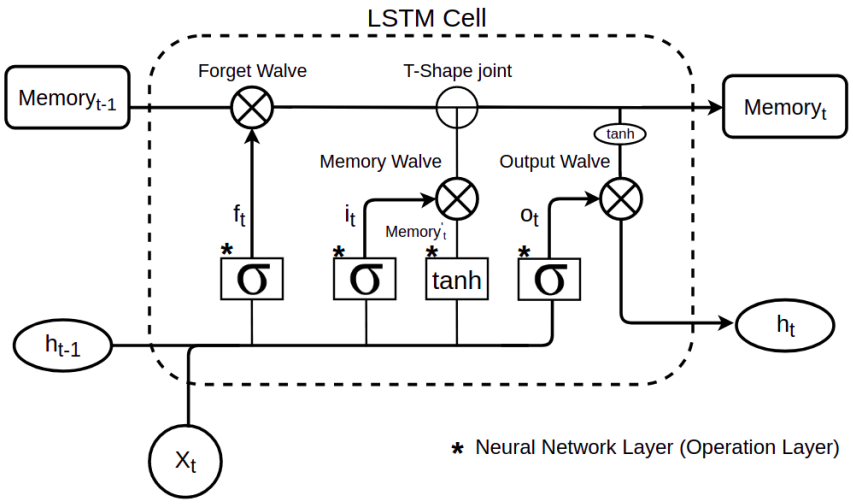


Figure 7.1. Structure of LSTM neural network cell.

Figure 7.1. demonstrates the typical structure of the LSTM neural network cells is configured mainly by gates. Input gate takes a new input TEC data, memory in (t-1) time step takes feedback from the LSTM cell's output in the last iteration. Deciding when to forget the output results is the task of the forget gate.

As illustrated in the LSTM cell formula in Equation20, $X = (x_1, x_2, ..., x_7)$ represents our TEC data separated by days. The proposed LSTM model is planned to use the previous seven days of the earthquakes and quiet days to perform prediction. The hidden state of memory cells is calculated in the Equation20 formulas.

$$
\begin{aligned}
i_t &= \sigma(x_t U^i + h_{t-1} W^i) \\
f_t &= (x_t U^f + h_{t-1} W^f) \\
o_t &= (x_t U^o + h_{t-1} W^o) \\
Memory'_t &= tanh(x_t U^g + h_{t-1} W^g) \\
Memory_t &= \sigma(f_t * Memory_{t-1} + i_t * Memory'_t) \\
h_t &= tanh(Memory_t * o_t)
\end{aligned}
\tag{20}
$$

Where i,f,o respectively represent the input gate, forget gate, and the output gates. g describes the candidate's internal state. W and U are the weight matrix and recurrent connections between the prior hidden and current hidden layers. $Memory'$ is a candidate hidden state memory based on the current input and prior hidden state. $Memory$ is a combination of the prior memory multiplied by the forget gate. σ stands for the standard sigmoid function defined in the previous chapter.

The variant versions of the LSTM models are used to enhance the contribution of this research. For this purpose, deep-bidirectional LSTM [135] and Stacked LSTM [136] are evaluated. In the Bi-LSTM model, two LSTMs are applied to the input data. The Bi-LSTM consists of two operation layers. The first is the forward layer that applied an LSTM on the input data, and the second is the reverse mode of the input data fed into the LSTM model called the backward layer. Stacked LSTMs were introduced by Pascanu, et al. [136] for solving complex sequence prediction problems.

The stacked LSTM structure presents in Figure 7.2. The Stacked LSTM model fed on TEC sequence input data that some critical memory is kept for further use. The model stack another dense layer that uses the earthquake labels of input data to perform classification.

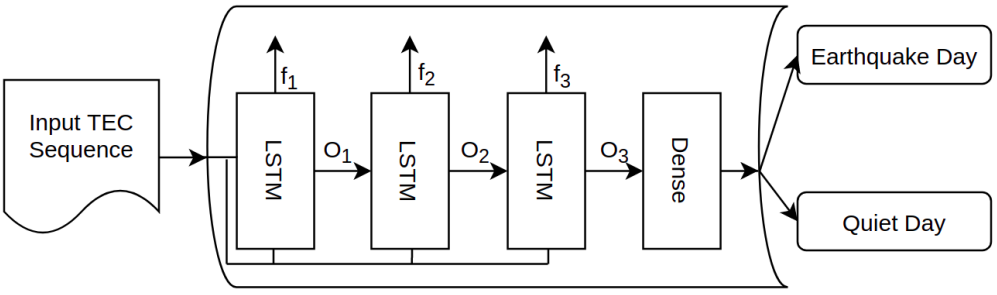


Figure 7.2. Structure of Stacked LSTM.

7.2. Dataset For LSTM-Based Models

As stated in Table 7.1., the dataset is distributed into train and test sets, with an 80-20 split rate. Therefore there are only 79 earthquakes (Magnitude $\geq$ 4.5) between 2012 to 2019. The

training set consists of 16 earthquakes events within two subsets of the earthquake set. On the contrary, there are only ten earthquakes in the test set.

Dataset Properties	**Value**
Train-Test Ratio	80%-20%
Total number of earthquakes	79
Number of EQs $\geq$4.5 in Testset	16
Number of EQs $\geq$5.0 in Testset	10

Table 7.1. Prepared data for the proposed LSTM based models.

The dataset is used in the proposed LSTM models is balanced. That means a fixed ratio exists between the number of going to be predicted quiet days and the number of predicted earthquake days. As depicted in the structure dataset in Figure 7.3., samples in the prediction process have eight continuous days that will predict the 8th day's event of an earthquake or quiet day with examining the previous seven days. Every row in the dataset has eight continuous days, and each day has 94 ionospheric TEC samples and a cosine similarity field.

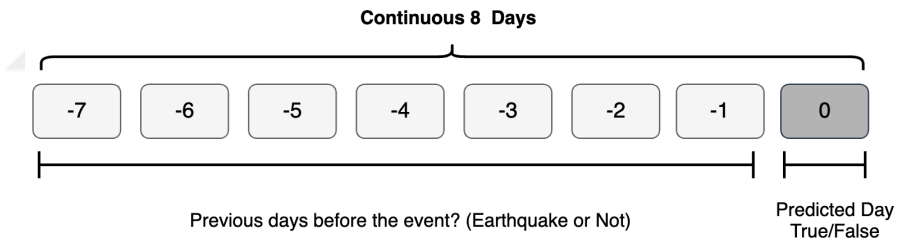


Figure 7.3. Dataset structure used in the proposed LSTM based models.

7.3. Evaluation Results

To examine the importance of the proposed LSTM based models, we used SVM (Support Vector Machine) classifier based on finding a hyperplane in an N-dimensional space. The most trustworthy hyperplane means the one with the largest margin between the quiet days and earthquake days. To For more contribution, it has been used various classifiers such as LDA and Random forest classifiers.

Using the evaluation methodology defined in Section 5., performances of our proposed LSTM models and the other classifiers model are evaluated. This section is organized to compare and examine the performance metrics of the classification models.

EQs	Model	Accuracy	Precision	Recall	F1-score
	LSTM	0.703	0.672	0.793	0.728
	Bi-LSTM	0.720	0.697	0.787	0.739
≥4.5	Stacked-LSTM	0.787	0.777	0.812	0.794
	SVM	0.659	0.640	0.756	0.693
	LDA	0.604	0.590	0.656	0.621
	RF	0.640	0.619	0.731	0.670
	LSTM	0.750	0.726	0.820	0.770
	Bi-LSTM	0.740	0.701	0.840	0.764
≥5.0	Stacked-LSTM	0.806	0.768	0.881	0.821
	SVM	0.695	0.678	0.760	0.716
	LDA	0.640	0.628	0.691	0.657
	RF	0.670	0.653	0.750	0.692

Table 7.2. Comparison of the proposed LSTM models other classification models using performance metrics.

Table 7.2. describes the median percentage of performance metrics like accuracy, recall, precision, and F1-score for the proposed LSTM based models and other classifiers based on two datasets.

The first six rows illustrate the comparison between the proposed models and other classifiers model based on all earthquakes with dataset that its magnetite ≥ 4.5. It is evident that the proposed models has higher performance in all metrics than the classifiers. The Stacked LSTM model boosted its precision and recall much more than the other classifiers. The ionosphere layer disturbances reflect all cosmic and seismic activities.As the false positive count is high in some scenarios in the dataset, precision is lower than recall and accuracy. The last six rows compare the LSTM models and the SVM, LDA and Random Forest classifiers based on more strong earthquakes (magnetite-EQs ≥ 5.0). Since strong earthquakes affect the ionosphere more than moderate earthquakes, they are classified efficiently.

Detailed information about ten-fold cross-validation related to the comparison of proposed LSTM models and other classification models using performance metrics are presented in tables 7.3.,7.4., 7.5. and 7.6..

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
LSTM	**TP**	13	14	13	14	14	11	12	13	10	13
	FP	5	7	6	8	7	7	5	6	5	6
	FN	3	2	3	2	2	5	4	3	6	3
	TN	11	9	10	8	9	9	11	10	11	10
EQs≥4.5											
	Precision	0.72	0.66	0.68	0.63	0.66	0.61	0.70	0.68	0.66	0.68
	Recall	0.81	0.87	0.81	0.87	0.87	0.68	0.75	0.81	0.62	0.81
	Accuracy	0.75	0.71	0.71	0.68	0.71	0.62	0.71	0.71	0.65	0.71
B-LSTM	**TP**	12	13	13	14	12	12	13	13	11	13
	FP	4	4	4	5	5	8	4	7	6	9
	FN	4	3	3	2	4	4	3	3	5	3
	TN	12	12	12	11	11	8	12	9	10	9
EQs≥4.5											
	Precision	0.75	0.76	0.76	0.73	0.70	0.6	0.76	0.65	0.64	0.59
	Recall	0.75	0.81	0.81	0.87	0.75	0.75	0.81	0.81	0.68	0.81
	Accuracy	0.75	0.78	0.78	0.78	0.71	0.62	0.78	0.68	0.65	0.64
Stacked-LSTM	**TP**	13	12	12	13	12	14	12	13	15	14
	FP	3	5	5	4	3	5	2	2	4	5
	FN	3	4	4	3	4	2	4	3	1	2
	TN	13	11	11	12	13	11	14	14	12	11
EQs≥4.5											
	Precision	0.81	0.70	0.70	0.76	0.8	0.73	0.85	0.86	0.78	0.73
	Recall	0.81	0.75	0.75	0.81	0.75	0.87	0.75	0.81	0.93	0.87
	Accuracy	0.81	0.71	0.71	0.78	0.78	0.78	0.81	0.84	0.84	0.78
SVM	**TP**	11	14	12	14	11	12	10	13	11	13
	FP	7	8	6	8	4	3	7	10	9	8
	FN	5	2	4	2	5	4	6	3	5	3
	TN	9	8	10	8	12	13	9	6	7	8
EQs≥4.5											
	Precision	0.61	0.63	0.66	0.63	0.73	0.8	0.58	0.56	0.55	0.61
	Recall	0.68	0.87	0.75	0.87	0.68	0.75	0.62	0.81	0.68	0.81
	Accuracy	0.62	0.68	0.68	0.68	0.71	0.78	0.59	0.59	0.56	065

Table 7.3. Detailed 10 fold comparison of the proposed LSTM models and SVM classification models using performance metrics in the EQs≥4.5 dataset.

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
LDA EQ≥4.5	**TP**	11	10	10	12	11	9	10	11	12	9
	FP	6	7	7	6	8	6	7	9	8	9
	FN	5	6	6	4	5	7	6	5	4	7
	TN	10	9	9	10	8	10	9	7	8	7
	Precision	0.64	0.58	0.58	0.66	0.57	0.6	0.58	0.55	0.6	0.5
	Recall	0.68	0.62	0.62	0.75	0.68	0.56	0.62	0.68	0.75	0.56
	Accuracy	0.65	0.59	0.59	0.68	0.59	0.59	0.59	0.56	0.625	0.5
RF EQ≥4.5	**TP**	12	13	11	12	10	12	11	12	11	13
	FP	7	7	6	8	9	7	7	6	9	6
	FN	4	3	5	4	6	4	5	4	5	3
	TN	9	9	10	8	7	9	9	10	7	10
	Precision	0.63	0.65	0.64	0.6	0.52	0.63	0.61	0.66	0.55	0.68
	Recall	0.75	0.81	0.68	0.75	0.62	0.75	0.68	0.75	0.68	0.81
	Accuracy	0.65	0.68	0.65	0.62	0.53	0.65	0.62	0.68	0.56	0.71

Table 7.4. Detailed 10 fold comparison of LDA and Random Forest classification models using performance metrics in the EQs≥4.5 dataset.

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
LSTM	**TP**	8	7	9	8	7	9	8	9	9	8
	FP	1	2	4	3	5	4	3	4	2	4
	FN	2	3	1	2	3	1	2	1	1	2
	TN	9	8	6	7	5	6	7	6	8	6
EQs≥5.0											
	Precision	0.88	0.77	0.69	0.72	0.58	0.69	0.72	0.69	0.81	0.66
	Recall	0.8	0.7	0.9	0.8	0.7	0.9	0.8	0.9	0.9	0.8
	Accuracy	0.85	0.75	0.75	0.75	0.6	0.75	0.75	0.75	0.85	0.7
B-LSTM	**TP**	9	8	8	9	7	8	9	9	9	8
	FP	3	5	5	3	4	4	3	3	2	4
	FN	1	2	2	1	3	2	1	1	1	2
	TN	7	5	5	7	6	6	7	7	8	6
EQs≥5.0											
	Precision	0.75	0.61	0.61	0.75	0.63	0.66	0.75	0.75	0.81	0.66
	Recall	0.9	0.8	0.8	0.9	0.7	0.8	0.9	0.9	0.9	0.8
	Accuracy	0.8	0.65	0.65	0.8	0.65	0.7	0.8	0.8	0.85	0.7
Stacked-LSTM	**TP**	10	8	9	9	8	9	10	9	9	8
	FP	1	3	3	3	3	2	3	2	3	4
	FN	0	2	1	1	2	1	0	1	2	2
	TN	9	7	7	7	7	8	7	8	7	6
EQs≥5.0											
	Precision	0.90	0.72	0.75	0.75	0.72	0.81	0.76	0.81	0.75	0.66
	Recall	1.0	0.8	0.9	0.9	0.8	0.9	1.0	0.9	0.81	0.8
	Accuracy	0.95	0.75	0.8	0.8	0.75	0.85	0.85	0.85	0.76	0.7
SVM	**TP**	8	7	8	7	7	8	7	8	8	8
	FP	4	3	3	3	4	3	3	3	7	4
	FN	2	3	2	3	3	2	3	2	2	2
	TN	6	7	7	7	6	7	7	7	3	6
EQs≥5.0											
	Precision	0.66	0.7	0.72	0.7	0.63	0.72	0.7	0.72	0.53	0.66
	Recall	0.8	0.7	0.8	0.7	0.7	0.8	0.7	0.8	0.8	0.8
	Accuracy	0.7	0.7	0.75	0.7	0.65	0.75	0.7	0.75	0.55	0.7

Table 7.5. Detailed 10 fold comparison of the proposed LSTM models and SVM classification models using performance metrics in the EQs≥5.0 dataset.

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
LDA	**TP**	7	7	6	8	7	6	7	6	8	7
	FP	4	5	3	4	4	5	4	3	4	5
	FN	3	3	4	2	3	4	3	4	2	3
	TN	6	5	7	6	6	5	6	7	6	5
EQ≥5.0											
	Precision	0.63	0.58	0.66	0.66	0.63	0.54	0.63	0.66	0.66	0.58
	Recall	0.7	0.7	0.6	0.8	0.7	0.6	0.7	0.6	0.8	0.7
	Accuracy	0.65	0.6	0.65	0.7	0.65	0.55	0.65	0.65	0.7	0.6
RF	**TP**	8	8	9	8	6	7	7	8	7	7
	FP	5	5	4	3	3	4	4	3	4	5
	FN	2	2	1	2	4	3	3	2	3	3
	TN	5	5	6	7	7	6	6	7	6	5
EQ≥5.0											
	Precision	0.61	0.61	0.69	0.72	0.66	0.63	0.63	0.72	0.63	0.58
	Recall	0.8	0.8	0.9	0.8	0.6	0.7	0.7	0.8	0.7	0.7
	Accuracy	0.65	0.65	0.75	0.75	0.65	0.65	0.65	0.75	0.65	0.6

Table 7.6. Detailed 10 fold comparison of LDA and Random Forest classification models using performance metrics in the EQs≥5.0 dataset.

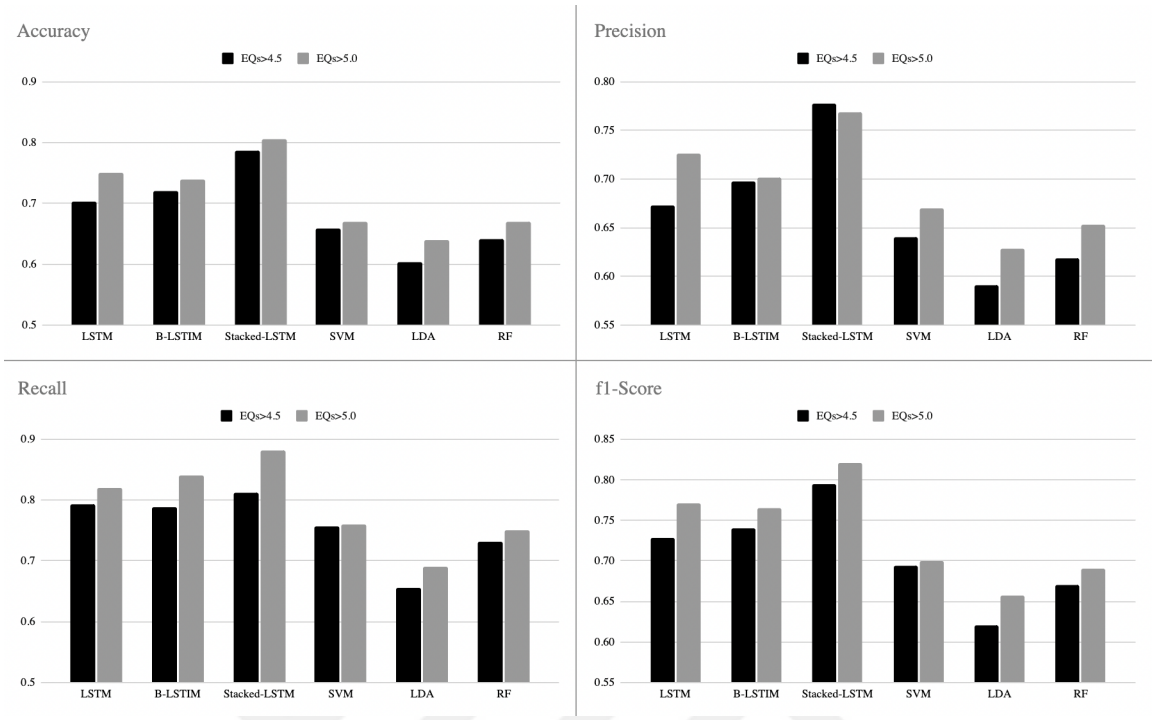


Figure 7.4. Performance metrics of the LSTM models and the SVM classifier based on two datasets (EQs≥4.5 and EQs≥5.0).

Figure 7.4. is the bar chart of performance metrics between the models in terms of two datasets (EQs≥4.5 and EQs≥5.0). It can be seen that in Figure 7.4. (Accuracy), the accuracy of the LSTM based grows slightly in the two mentioned datasets. Precision is a measure of classification result relevancy based on earthquakes and false alarms, while recall is a measure of the classifier's ability to find all the earthquake days. The comparison between datasets shows that the LSTM based models are more stable than the SVM and other classifiers in performance. Overall, Figure 7.4. reveals the Stacked-LSTM approach notably increases the precision, recall, and F1-score. Consequently, the improved performance in the dataset EQs≥4.5 is more extensive than the dataset EQs≥5.0, and the LSTM based models are more reliable in all earthquakes. As shown in the results, SVM is more accurate than Random Forest and LDA classifiers. So In the ROC table, we compare proposed LSTM models and SVM classifier. Figure 7.5. depicts the ROC curve of the LSTM based models and the SVM classifier model based on two datasets. ROC curves related to the Stacked-LSTM positioned progressively closer to the upper left angle in ROC space, So the Stacked-LSTM has the progressively more prominent discriminant capacity of earthquake classification. As

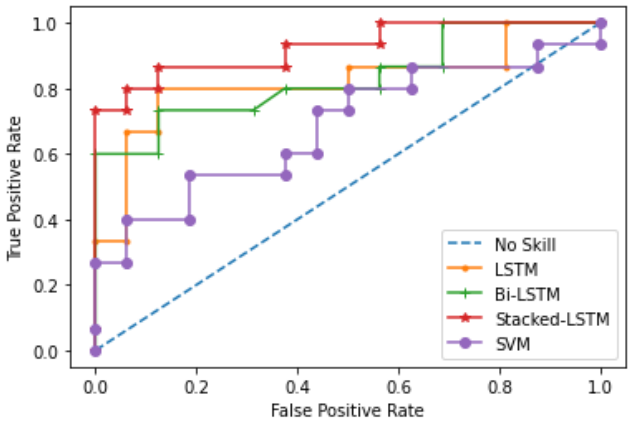


Figure 7.5. ROC curve of the LSTM based models and the SVM classifier.

depicted in the figure, the LSTM based models, and the SVM classifier visually be compared simultaneously, and the results show the LSTM based models is more efficient than the SVM classifier.

The results in Table 7.7. demonstrates AUC values of the models for each dataset. The value of AUC in the Stacked-LSTM is 0.92, while in the SVM classifier is 0.60, which means a keen improvement in the ROC metric.

Model	**ROC AUC**
LSTM	0.821
Bi-LSTM	0.831
Stacked-LSTM	0.925
SVM	0.603

Table 7.7. AUC values for the LSTM models and the SVM classifier.

The P-Value is the possibility of observing the effect(E) when the null hypothesis is true. In the test result of the Stacked-LSTM, the difference in performance over all metrics is significant $(p - value \ll 0.01)$.

7.4. Conclusion for LSTM based models

This research propose approaches to detect earthquakes based on ionospheric TEC values. We proposed a predictive models to predict ionosphere disturbances based on earthquake using earlier days of earthquakes using LSTM based deep learning models. Ionospheric TEC

data has been collected from two GPS stations. Chile region has strong and severe earthquakes; however, the Karratha region is almost quiet in terms of earthquakes. The principal intention of this research is to attain the relationship between earthquakes and the ionosphere disturbances based on deep learning techniques concerning LSTM based classification models. To investigate the contribution of the recommended LSTM models, we used SVM, LDA and Random Forest classifiers. Our results indicate approximately 78-80 accuracy-based performance in the two test sets of the earthquakes, including moderate and severe earthquakes. The proposed Stacked-LSTM has a more trustworthy and reliable performance than other models in terms of accuracy, recall, precision, F1-score, and ROC curve metrics.

8. An approach using Deep AutoEncoder and LSTM Neural Networks for prediction of earthquakes using ionospheric data

This section aims to improve the stacked LSTM-based earthquake classification model in the last chapter, which has better results than other LSTM based models.The proposed model similarly performs classification to detect earthquakes in previous days using ionospheric variability throughout moderate and severe earthquake events of varying powers for 2012-2019. The recommended models use a Deep autoencoder to extract useful features from ionospheric TEC data and perform Stacked LSTM to classify earthquakes days by analyzing TEC values of the last days.

This chapter presented the structure of the proposed hybrid model based on deep autoencoder and stacked LSTM as presented in Figure 8.1.. Due to the enormous amount of TEC data related to each station and day, it is required to perform a feature reduction process. In the classification phase, as the TEC data variations in prior days of the earthquakes play a significant role in the forecast phase, it is organized to implement a sequential learning approach.

8.1. Deep Auto-Encoder Stacked-LSTM Model

This model aims to combine the previous chapters' successful models. We propose the DAEclass model to interpret the earthquake days using ionospheric TEC data in the previous chapters. As DAEclass is flourishing in terms of feature reduction and extraction, we determined to use the deep autoencoder phase with this model. As proposed, the Stacked-LSTM has a more trustworthy and reliable performance than the other LSTM models and dense neural networks. It is approved to use the Stacked-LSTM model for the classification process.

8.1.1. Deep Autoencoder Sub-Model

As mentioned in Chapter 6., the ionospheric TEC data is a high-dimensional factor. Reducing dimensionality to obtain a compressed feature set is considered an essential step in the feature extraction phase. For interpretation of earthquakes, we designed a robust deep autoencoder to extract useful information from ionospheric TEC data each day. As stated in the 6., adding hidden layers to the autoencoders conducts to perform the proper dimension reduction process. Deep Autoencoders have confirmed their effectiveness in detecting non-linear features across many complex problems. In other words, an autoencoder gets a map from the input to itself within a pair of encoding and decoding steps. Autoencoders have latterly pointed in unsupervised methods for feature learning. It tries to learn a compressed description of the input while retaining the essential information.

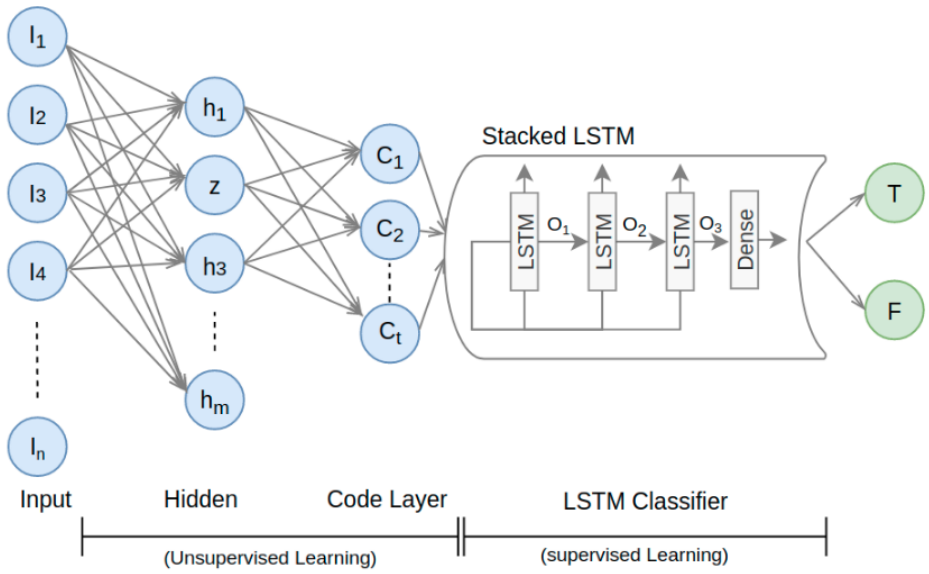


Figure 8.1. The general architecture of the proposed DAE-STCK-LSTM model.

The proposed hybrid model consists of deep autoencoder and the Stacked-LSTM model. As depicted in Figure 8.1., a deep autoencoder with two hidden layers is used in the unsupervised learning phase to extract features from ionospheric TEC data. The collected features served as input to LSTM network to perform classification.

8.1.2. Stacked LSTM Sub-Model

As specified in the chapter 7., Stacked LSTM is used for solving complex sequence prediction problems. The stacked LSTM structure presents in Figure 8.1. with three LSTM layers. The Stacked LSTM model fed on TEC sequence input data that critical memory is kept for further use. The model stack another dense layer that uses the earthquake labels of input data to perform classification.

Due to the TEC data sequence related to previous days is long enough, RNN's may leave out critical information from the beginning. The LSTM methods handle this issue with the solution to short-term memory.

8.2. Dataset For DAE-STCK-LSTM Model

As stated in Table 8.1., the dataset is distributed into train and test sets, with an 80-20 split rate. The proposed DAE-STCK-LSTM model has two sub-models in terms of unsupervised and supervised learning methods. Accordingly, each sub-model uses a distinct dataset for learning. The unsupervised phase uses 2057 days for learning deep autoencoder and uses Chapter 7. dataset for supervised phrase. There are only 79 earthquakes (Magnitude $\geq$ 4.5) between 2012 to 2019. The training set consists of 16 earthquakes events within two subsets of the earthquake set. On the contrary, there are only ten earthquakes in the test set.

Dataset Properties	**Value**
Train-Test Ratio	80%-20%
Total number of earthquakes	79
Number of days in Trainset in Unsupervides Learning	2057
Number of EQs $\geq$4.5 in Testset	16
Number of EQs $\geq$5.0 in Testset	10

Table 8.1. Prepared data for the proposed DAE-STCK-LSTM model.

Similary with 7. dataset, Samples in the prediction method have eight continuous days that will predict the 8th day's event of an earthquake or quiet day with analyzing the previous

seven days. Every row in the dataset has eight continuous days, and each day has 94 ionospheric TEC samples and a cosine similarity field.

8.3. Evaluation Results

For evaluating the proposed prediction model, we used Stacked-LSTM, SVM (Support Vector Machine) classifier, and LDA (Linear Discriminant Analysis) classifier. As mentioned in the previous chapters, the most trustworthy hyperplane means the largest margin between quiet and earthquake days.

Utilizing the evaluation methodology described in Section 5., performances of the proposed DAE-STCK-LSTM model and other classifier models are assessed. This section is prepared to analyze and measure the performance metrics of the classification models.

EQs	Model	Accuracy	Precision	Recall	F1-score
	DAE-STCK-LSTM	0.818	0.798	0.865	0.827
≥4.5	Stacked-LSTM	0.787	0.777	0.812	0.794
	LDA	0.627	0.614	0.706	0.657
	SVM	0.659	0.640	0.756	0.693
	DAE-STCK-LSTM	0.835	0.818	0.870	0.843
≥5.0	Stacked-LSTM	0.806	0.768	0.881	0.821
	LDA	0.661	0.664	0.687	0.676
	SVM	0.695	0.678	0.760	0.716

Table 8.2. Comparison of the proposed DAE-STCK-LSTM models and other classification model using performance metrics.

Table 8.2. represents the median rate of performance metrics like accuracy, recall, precision, and F1-score for the proposed DAE-STCK-LSTM model, Stacked-LSTM, LDA, and SVM classifier based on two datasets.

The first four rows demonstrate the comparison among the proposed model and other classifier models based on all earthquakes with the dataset that its magnetite ≥ 4.5. It is obvious that the DAE-STCK-LSTM and Stacked-LSTM model have higher performance in all metrics than the SVM and LDA classifier. The DAE-STCK-LSTM model raised its accuracy

and F1-score slightly more than the Stacked-LSTM classifier. The ionosphere layer irregularities reveal all cosmic and seismic movements. As the false positive count is high in more scenarios in the dataset, precision is lower than recall and accuracy.

The second four rows compare the DAE-STCK-LSTM model and other classifier models based on more powerful earthquakes (magnetite-EQs ≥ 5.0). As powerful earthquakes affect the ionosphere more than moderate earthquakes, they are classified efficiently.

We believed that the reason for the higher performance of DAE-STCK-LSTM in comparison with Stacked-LSTM is accurate feature extraction. As the ionospheric data is very high dimensional in terms of TEC data, it has noisy data and has other effects than earthquakes. The deep autoencoder reducts the data properly and extracts the variable features in the data that they reflect properly whole data. The other classifier like SVM and LDA have lower performance than LSTM models. They do not focus on the relationship between sequences in the data. However, the ionospheric TEC data is completely depending on the sequences in the dataset. It means that one day before the earthquake is more variable than the previous days of the earthquake day, and the sequence between days is more important.

More comprehensive results about ten-fold cross-validation related to the comparison of proposed DAE-STCK-LSTM model and other classification models using performance metrics are presented in Tables 8.3. and 8.4..

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
DAE -STCK -LSTM	**TP**	15	14	13	12	11	15	14	14	16	14
	FP	4	4	4	3	4	4	3	1	4	4
	FN	1	2	3	4	5	1	2	3	0	2
	TN	12	12	12	13	12	12	13	15	12	12
EQs>4.5	**Precision**	0.78	0.77	0.76	0.80	0.73	0.78	0.82	0.93	0.8	0.77
	Recall	0.93	0.87	0.81	0.75	0.68	0.93	0.87	0.82	1.00	0.87
	Accuracy	0.84	0.81	0.78	0.78	0.71	0.84	0.84	0.87	0.87	0.81
Stacked -LSTM	**TP**	13	12	12	13	12	14	12	13	15	14
	FP	3	5	5	4	3	5	2	2	4	5
	FN	3	4	4	3	4	2	4	3	1	2
	TN	13	11	11	12	13	11	14	14	12	11
EQs>4.5	**Precision**	0.81	0.70	0.70	0.76	0.80	0.73	0.85	0.86	0.78	0.73
	Recall	0.81	0.75	0.75	0.81	0.75	0.87	0.75	0.81	0.93	0.87
	Accuracy	0.81	0.71	0.71	0.78	0.78	0.78	0.81	0.84	0.84	0.78
LDA	**TP**	10	12	11	12	11	11	12	11	11	12
	FP	7	9	9	8	5	5	6	8	9	6
	FN	6	4	5	4	5	5	4	5	5	4
	TN	9	7	7	8	11	11	10	8	7	9
EQs>4.5	**Precision**	0.58	0.57	0.55	0.60	0.68	0.68	0.66	0.57	0.55	0.66
	Recall	0.62	0.75	0.68	0.75	0.68	0.68	0.75	0.68	0.68	0.75
	Accuracy	0.59	0.59	0.56	0.62	0.68	0.68	0.68	0.59	0.56	0.67
SVM	**TP**	11	14	12	14	11	12	10	13	11	13
	FP	7	8	6	8	4	3	7	10	9	8
	FN	5	2	4	2	5	4	6	3	5	3
	TN	9	8	10	8	12	13	9	6	7	8
EQs>4.5	**Precision**	0.61	0.63	0.66	0.63	0.73	0.80	0.58	0.56	0.55	0.61
	Recall	0.68	0.87	0.75	0.87	0.68	0.75	0.62	0.81	0.68	0.81
	Accuracy	0.62	0.68	0.68	0.68	0.71	0.78	0.59	0.59	0.56	0.65

Table 8.3. Detailed 10 fold comparison of the proposed DAE-STCK-LSTM model and other classification models using performance metrics in the EQs≥4.5 dataset.

Model	Count/ Results	*Ten Fold Cross Validation*									
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
DAE -STCK -LSTM	**TP**	8	9	9	10	9	8	8	9	8	9
	FP	1	3	2	2	3	1	1	2	3	2
	FN	2	1	1	0	1	2	2	1	2	1
	TN	9	7	8	8	7	9	9	8	7	8
EQs>5.0	**Precision**	0.88	0.75	0.81	0.83	0.75	0.88	0.88	0.80	0.72	0.81
	Recall	0.80	0.90	0.90	1.00	0.90	0.80	0.80	0.90	0.80	0.90
	Accuracy	0.85	0.8	0.85	0.90	0.80	0.85	0.85	0.85	0.75	0.85
Stacked -LSTM	**TP**	10	8	9	9	8	9	10	9	9	8
	FP	1	3	3	3	3	2	3	2	3	4
	FN	0	2	1	1	2	1	0	1	2	2
	TN	9	7	7	7	7	8	7	8	7	6
EQs>5.0	**Precision**	0.90	0.72	0.75	0.75	0.72	0.81	0.76	0.81	0.75	0.66
	Recall	1.00	0.80	0.90	0.90	0.80	0.90	1.00	0.90	0.81	0.80
	Accuracy	0.95	0.75	0.80	0.80	0.75	0.85	0.85	0.85	0.76	0.70
LDA	**TP**	7	7	7	6	8	7	6	7	7	6
	FP	4	4	2	5	4	4	3	4	2	3
	FN	3	3	2	4	2	3	4	3	3	4
	TN	6	6	3	5	6	6	7	6	8	7
EQs>5.0	**Precision**	0.63	0.63	0.77	0.54	0.66	0.63	0.66	0.63	0.77	0.66
	Recall	0.70	0.70	0.77	0.60	0.80	0.70	0.60	0.70	0.70	0.60
	Accuracy	0.65	0.65	0.71	0.55	0.70	0.65	0.65	0.65	0.75	0.65
SVM	**TP**	8	7	8	7	7	8	7	8	8	8
	FP	4	3	3	3	4	3	3	3	7	4
	FN	2	3	2	3	3	2	3	2	2	2
	TN	6	7	7	7	6	7	7	7	3	6
EQs>5.0	**Precision**	0.66	0.70	0.72	0.70	0.63	0.72	0.70	0.72	0.53	0.66
	Recall	0.80	0.70	0.80	0.70	0.70	0.80	0.70	0.80	0.80	0.80
	Accuracy	0.70	0.70	0.75	0.70	0.65	0.75	0.70	0.75	0.55	0.70

Table 8.4. Detailed 10 fold comparison of the proposed DAE-STCK-LSTM model and other classification models using performance metrics in the EQs≥5.0 dataset.

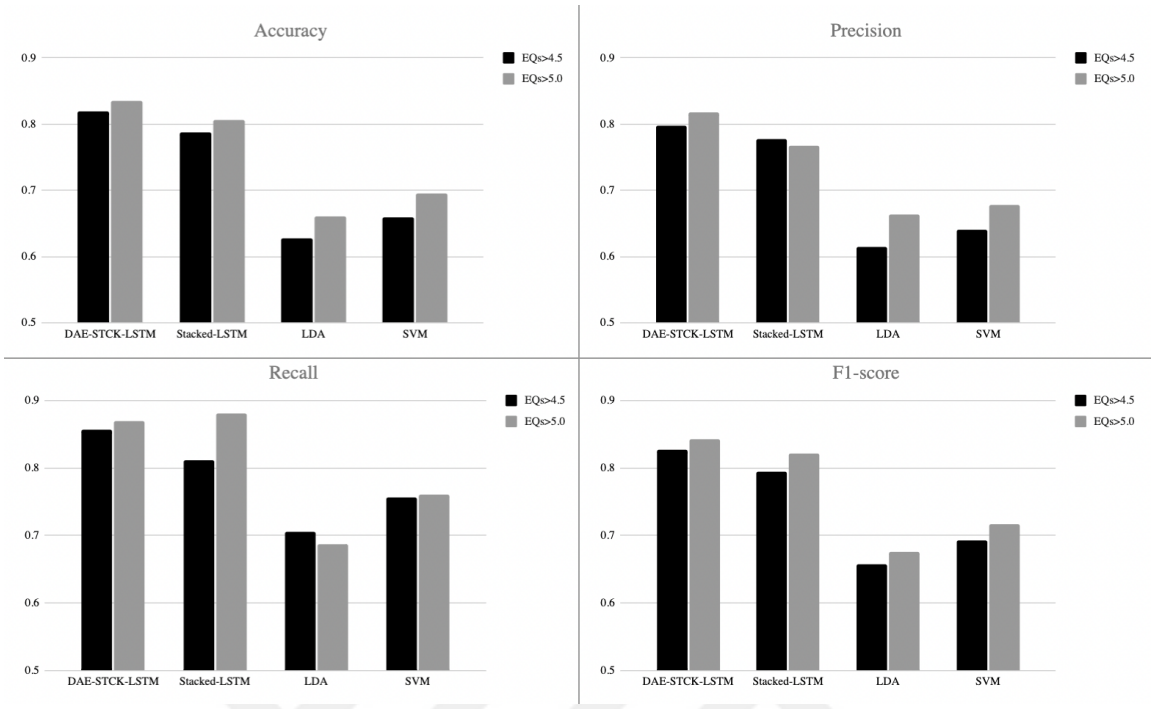


Figure 8.2. Performance metrics of the DAE-STCK-LSTM model and other classification model based on two datasets (EQs≥4.5 and EQs≥5.0).

Figure 8.2. is the bar chart of performance metrics between the models in terms of two datasets (EQs≥4.5 and EQs≥5.0). As depicted in Figure 7.4. (Accuracy), the accuracy of the DAE-STCK-LSTM based grows slightly in the two mentioned datasets in compared with Stacked-LSTM. Precision is a measure of classification result relevancy based on earthquakes and false alarms, while recall measures the classifier's capacity to obtain all the earthquake days. The comparison between datasets shows that the DAE-STCK-LSTM model is more beneficial than other classifiers in performance. Overall, Figure 8.2. explains the DAE-STCK-LSTM method prominently increases the precision, and F1-score. Similary with chapter 7., the enhanced performance in the dataset EQs≥4.5 is more comprehensive than the dataset EQs≥5.0, and the DAE-STCK-LSTM model are more accurate in all earthquakes than other LSTM-based and other classifiers.

The ROC curve of the DAE-STCK-LSTM model, Stacked-LSTM, LDA, and SVM classifiers based on two datasets are represented in Figure 8.3.. ROC curves related to the DAE-STCK-LSTM model are placed progressively closer to the upper left angle in ROC space. It

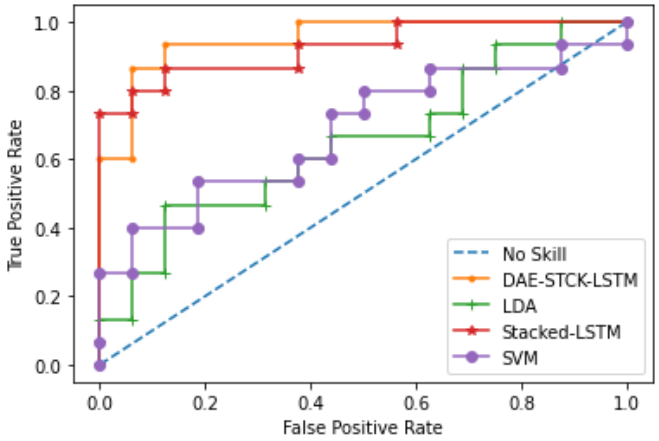


Figure 8.3. ROC curve of the LSTM based models and the SVM classifier.

is revealed that the DAE-STCK-LSTM has the progressively more prominent discriminant ability of earthquake prediction. As depicted in the figure, the DAE-STCK-LSTM model and the Stacked-LSTM are visually close together. However, The results show the DAE-STCK-LSTM model is more efficient than the Stacked-LSTM and other classifiers.

The AUC values of the proposed model and other classification models for each dataset are illustrated in Table 8.5.. The value of AUC in the DAE-STCK-LSTM is 0.92, while in the Stacked-LSTM classifier is 0.60, which means improvement in the ROC metric.

Model	**ROC AUC**
DAE-STCK-LSTM	0.951
Stacked-LSTM	0.925
LDA	0.651
SVM	0.683

Table 8.5. AUC values for the DAE-STCK-LSTM models and the other classification models.

In the test results of the statistical P-Value test, DAE-STCK-LSTM, the difference in performance in terms of all metrics is significant $(p - value \ll 0.01)$.

8.4. Conclusion for DAE-STCK-LSTM model

This section proposed a hybrid model called DAE-STCK-LSTM model to detect earthquakes based on ionospheric TEC values. The proposed model performs prediction of earthquakes

using earlier days of earthquakes using deep autoencoder and LSTM deep learning models. Ionospheric TEC data has been collected from two GPS stations. The Chile region has powerful and severe earthquakes; however, the Karratha region is almost quiet in terms of seismic activities. The primary aim of this model is to attain the relationship between earthquakes and ionosphere irregularities based on deep learning techniques concerning deep autoencoder and LSTM based classification models. To examine the contribution of the recommended DAE-STCK-LSTM model, we used Stacked-LSTM, LDA, and SVM classifiers. Our results show approximately 81-84 accuracy-based performance in the two test sets of the earthquakes, including moderate and severe earthquakes. The proposed DAE-STCK-LSTM model has a more accurate and stable performance than the other classification models in terms of accuracy, recall, precision, F1-score, and ROC curve metrics.

9. CONCLUSION

We have believed that monitoring the ionospheric irregularities reveals that geomagnetic activity like storms and earthquakes can cause significant disturbances in the electron density distribution and TEC values. The satellite-based measures from GPS stations have performed a beneficial study to investigate the seismo-ionospheric anomalies. TEC data collected from GPS stations (Dual-Frequency GPS receiver) is used to examine the ionospheric variability through moderate and severe earthquakes.

Discovering precursory signals in severe and moderate earthquakes using ionospheric TEC data to perform prediction is this study's subject. The primary motivation of this thesis is to examine models to analyze the relationship between earthquakes and ionospheric TEC data. The recommended models' principal aim is to predict an earthquake by analyzing previous days of the earthquake using ionospheric TEC data. Nevertheless, the precursory signals through long time periods are not investigated in the literature. The ionospheric variability during moderate and severe earthquake events of varying strengths for 2012-2019 years is discussed in this thesis.

This research utilizes various deep learning techniques and algorithms to extract relations between TEC data. For this purpose, it aims to use supervised and unsupervised learning methods based on deep learning techniques. In the unsupervised learning methods, we use Deep Autoencoder and LDA method. On the other hand, we use deep, dense neural networks, Long short-term memory, and support vector machine in the supervised learning methods.

Our contribution in this thesis is separated into three main models. In the first model, we focus on interpreting earthquakes based on ionosphere irregularities using deep neural networks that we called it DAEclass method. The primary aim of this model is to obtain the relationship between earthquakes and the ionosphere disturbances based on deep learning techniques concerning two sub-tasks of feature extraction and classification. In the feature extraction step, we focus on develop a Deep Autoencoder to extract valuable knowledge about the earthquake and quiet days. Then, in the classification step, we use a a dense neural

network to perform interpretation. To review the contribution of the recommended model, we applied LDA,SVM and Random Forest classifiers. Our results indicate approximately 90-94 accuracy-based performance in the two test sets of the earthquakes, including moderate and severe earthquakes. The proposed DAEclass has a more trustworthy and reliable performance than the LDA in terms of accuracy, recall, precision, F1-score, and ROC curve metrics.

The second contribution proposes a classification model that uses time series recurrent neural networks to predict earthquakes in previous days. As the TEC data changes in prior days of the earthquakes play an essential role in the prediction phase, it is planned to implement a sequential learning approach. In our sequential learning approach, the sequential dependency between TEC data is analyzed at the algorithmic level. Due to the TEC data sequence related to previous days is long enough, LSTM gets critical information of earlier days. The LSTM methods handle the problem with the solution to short-term memory. These methods can adjust the flow of information with internal mechanisms called gates. The variant versions of the LSTM models are proposed to enhance the contribution of this model. For this purpose, deep-bidirectional LSTM and Stacked LSTM are evaluated. In the Bi-LSTM model, two LSTMs are utilized in the input data. The Bi-LSTM consists of two operation layers. The first is the forward layer that applied an LSTM on the input data, and the second is the reverse mode of the input data fed into the LSTM model called the backward layer. Stacked LSTMs for solving complex sequence prediction problems. The Stacked LSTM model provided on TEC sequence input data that critical memory is kept for further use. The model stack another dense layer that uses the earthquake labels of input data to perform classification. To examine the contribution of the proposed LSTM models, we used SVM,LDA and Random Forest classifiers. Our results indicate approximately 78-80 accuracy-based performance in the two test sets of the earthquakes, including moderate and severe earthquakes.

Finally, In the last model, we develop a hybrid version of deep autoencoders and LSTM to detect earthquakes in previous days. This model aims to improve the stacked LSTM-based earthquake classification proposed in the second model. The recommended model uses a deep autoencoder to extract useful features from ionospheric TEC data and perform Stacked

LSTM to classify earthquakes days by analyzing TEC values of the last days. This model aims to achieve the relationship between earthquakes and ionosphere variations based on deep learning techniques concerning deep autoencoder and LSTM based classification models. To analyze the contribution of the recommended DAE-STCK-LSTM model, we used Stacked-LSTM, LDA, and SVM classifiers. Our evaluation test results prove approximately 81-84 accuracy-based performance in the two test sets of the earthquakes, including moderate and severe earthquakes. The proposed DAE-STCK-LSTM model has a more accurate and stable performance than the other classification models in terms of accuracy, recall, precision, F1-score, and ROC curve metrics.

Appendix A

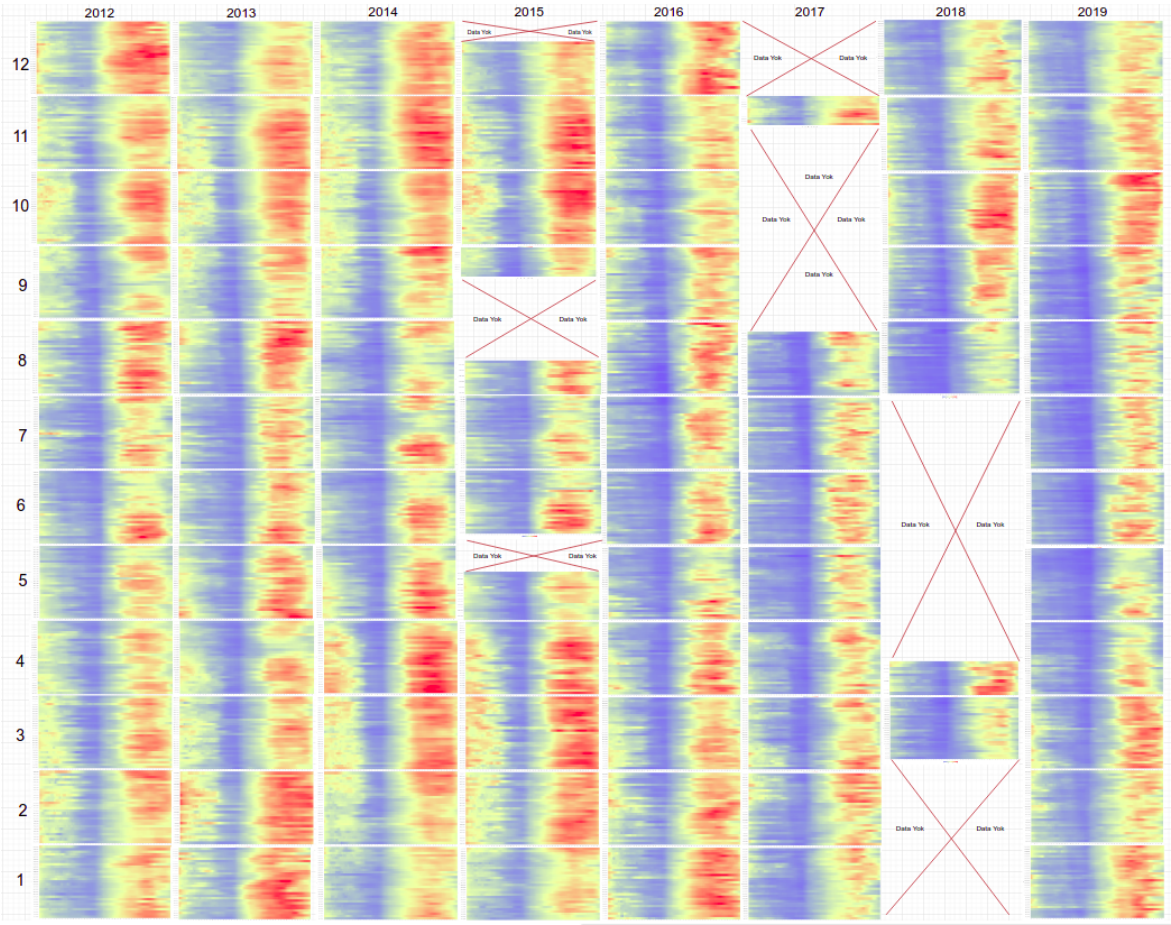
Heat Maps of TEC based Data Related to IQUIQUE Station

Appendix A depicted the dataset related to the IQUIQUE station and the earthquakes. The Iquique station is located in coordinate $(Lat : -20.15, Lon : -70.13)$ in Chile represented in the figure. TEC data gathered from GPS stations (Dual-Frequency GPS receiver) from the IONOLAB group (Hacettepe University of IONOLAB is an electrical engineer organization to investigate the ionosphere hurdles.) [1]. The earthquake information is collected via (United States Geological Survey of Earthquakes) [2]. All earthquakes in the dataset are marked with the red points in the pictures.

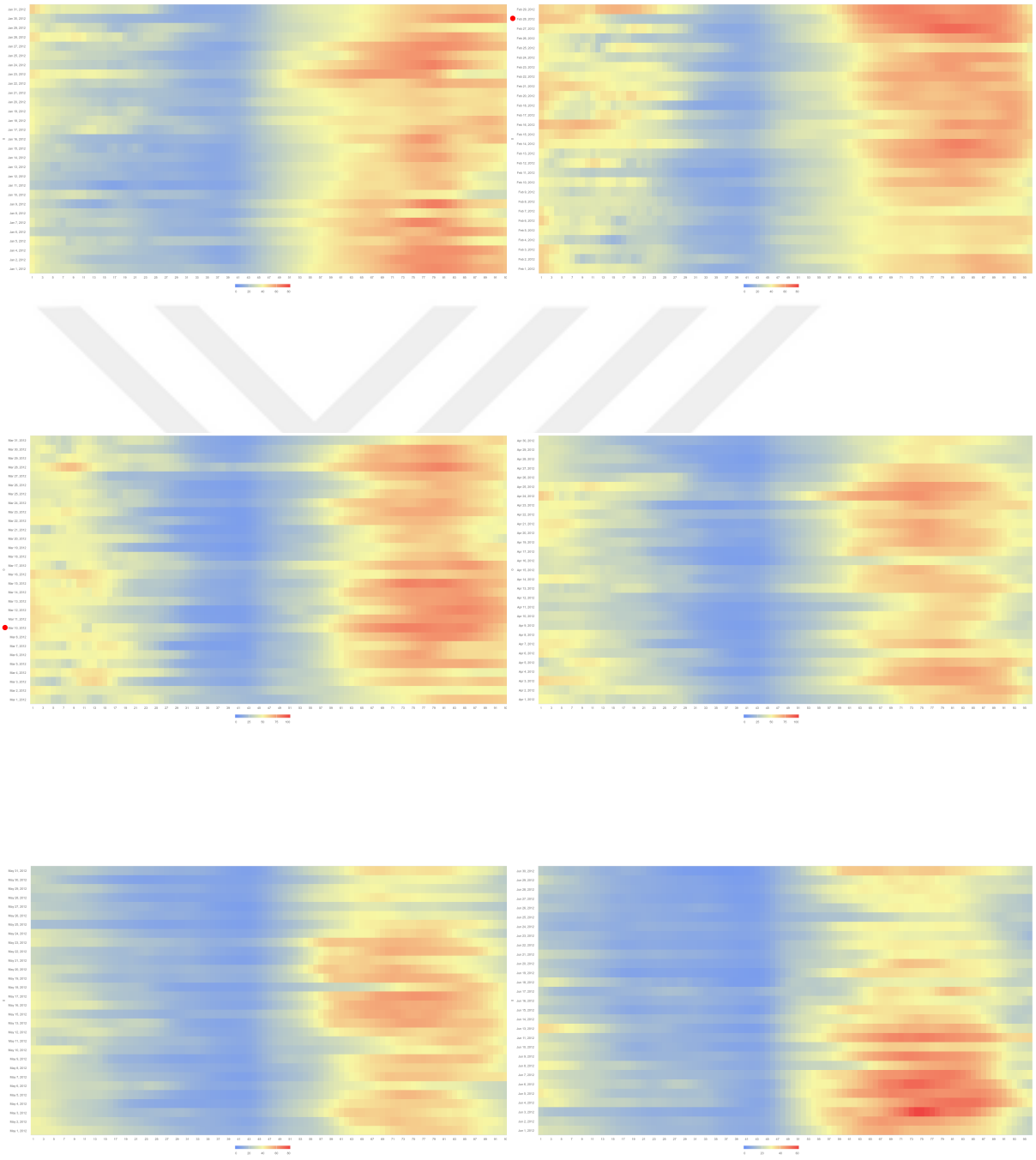
[1] Available at http://ionolab.org/
[2] Available at https://earthquake.usgs.gov/

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*

Overview of whole 2012-1019 Dataset related to the IQUIQUE GPS Station is represented as fallow:

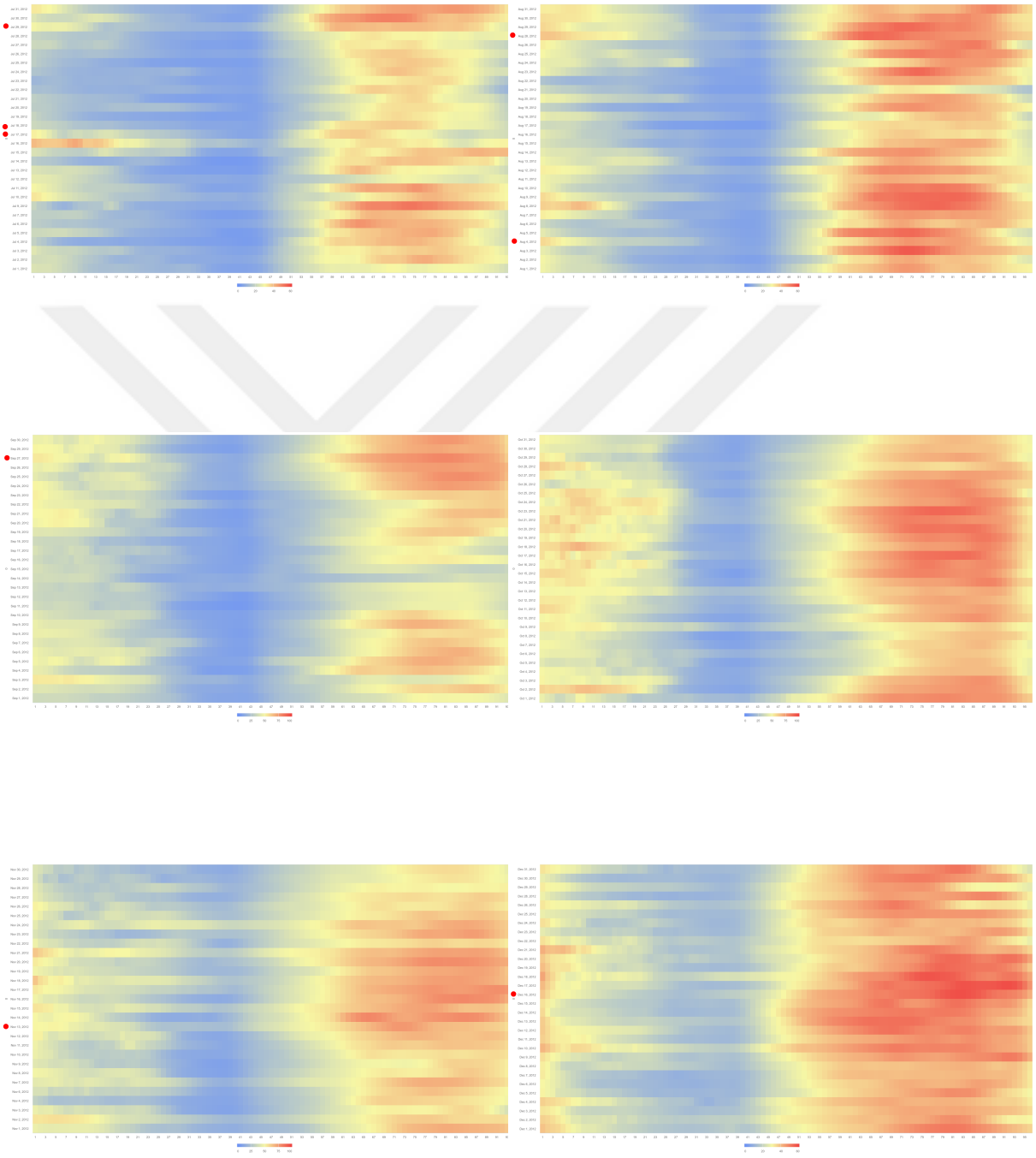


Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



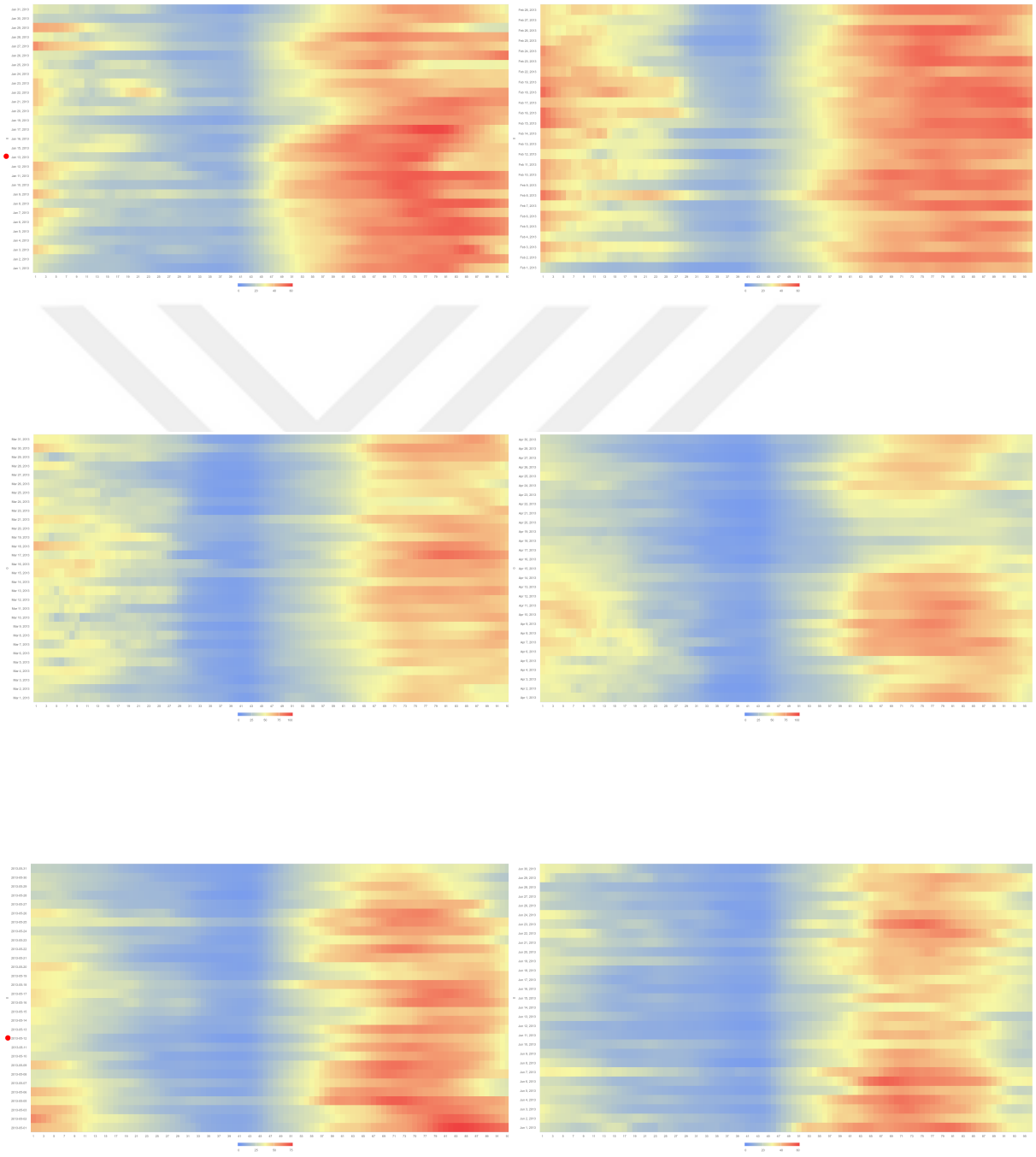
January 2012 to June 2012

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



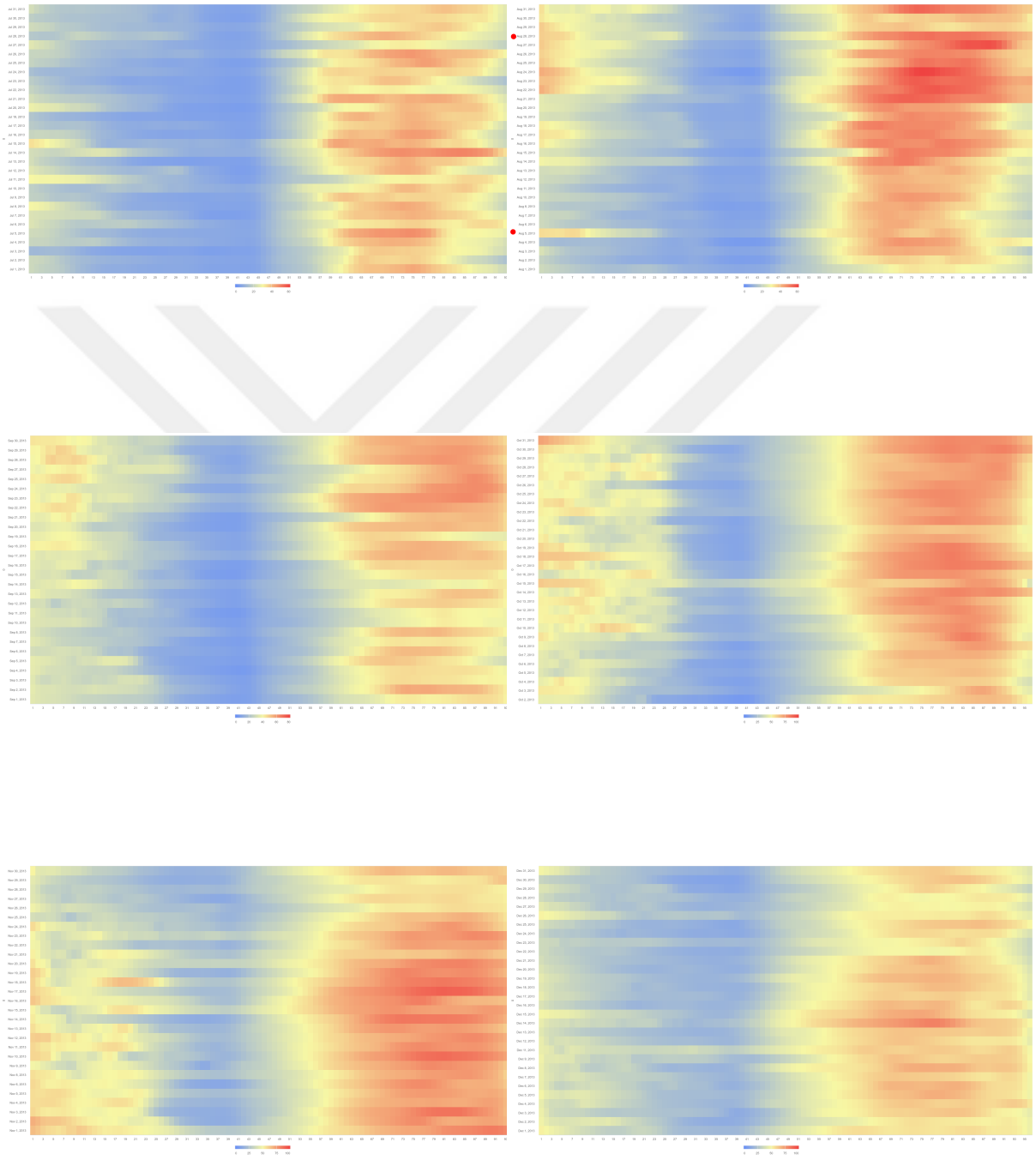
July 2012 to December 2012

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



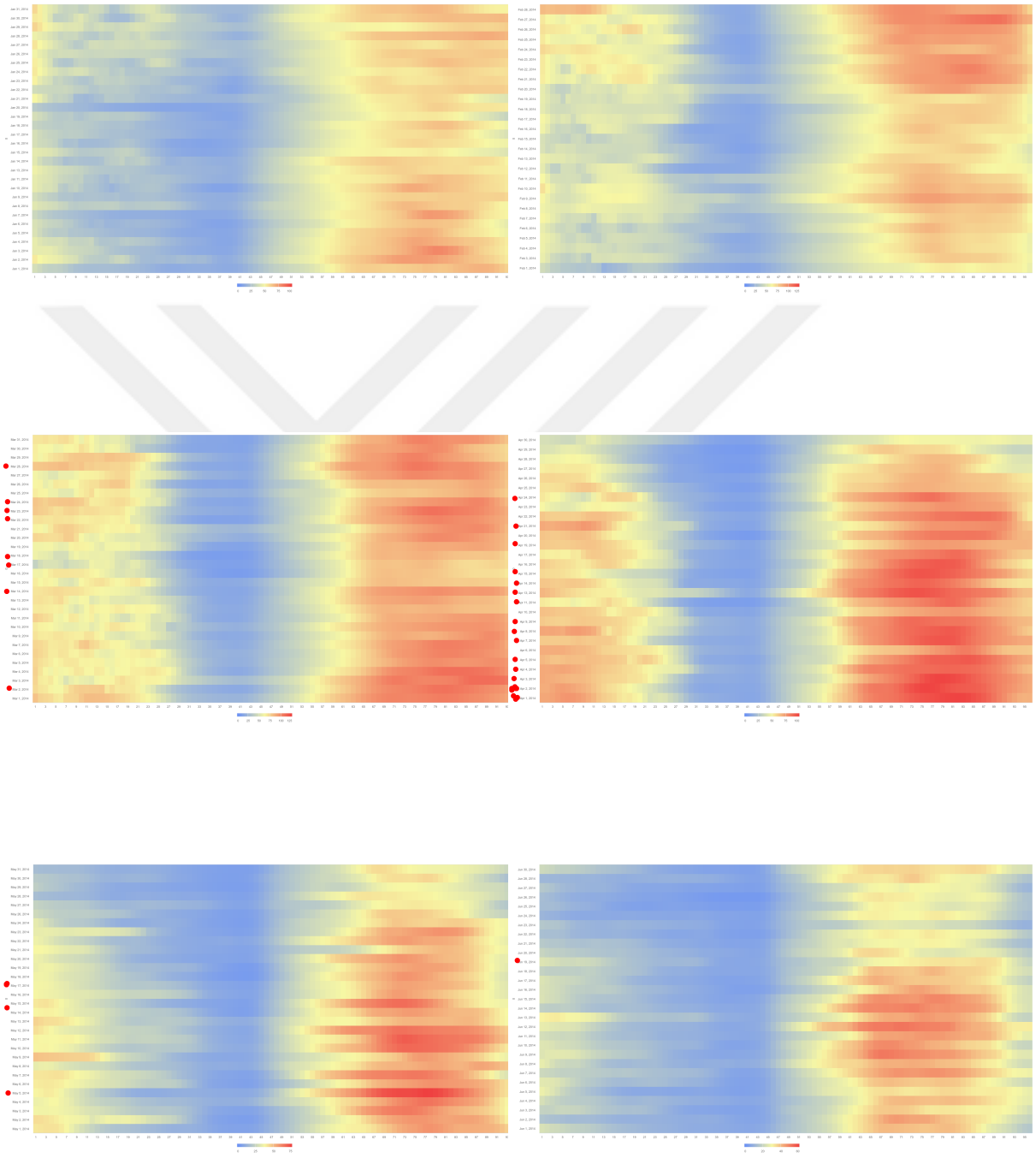
January 2013 to June 2013

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



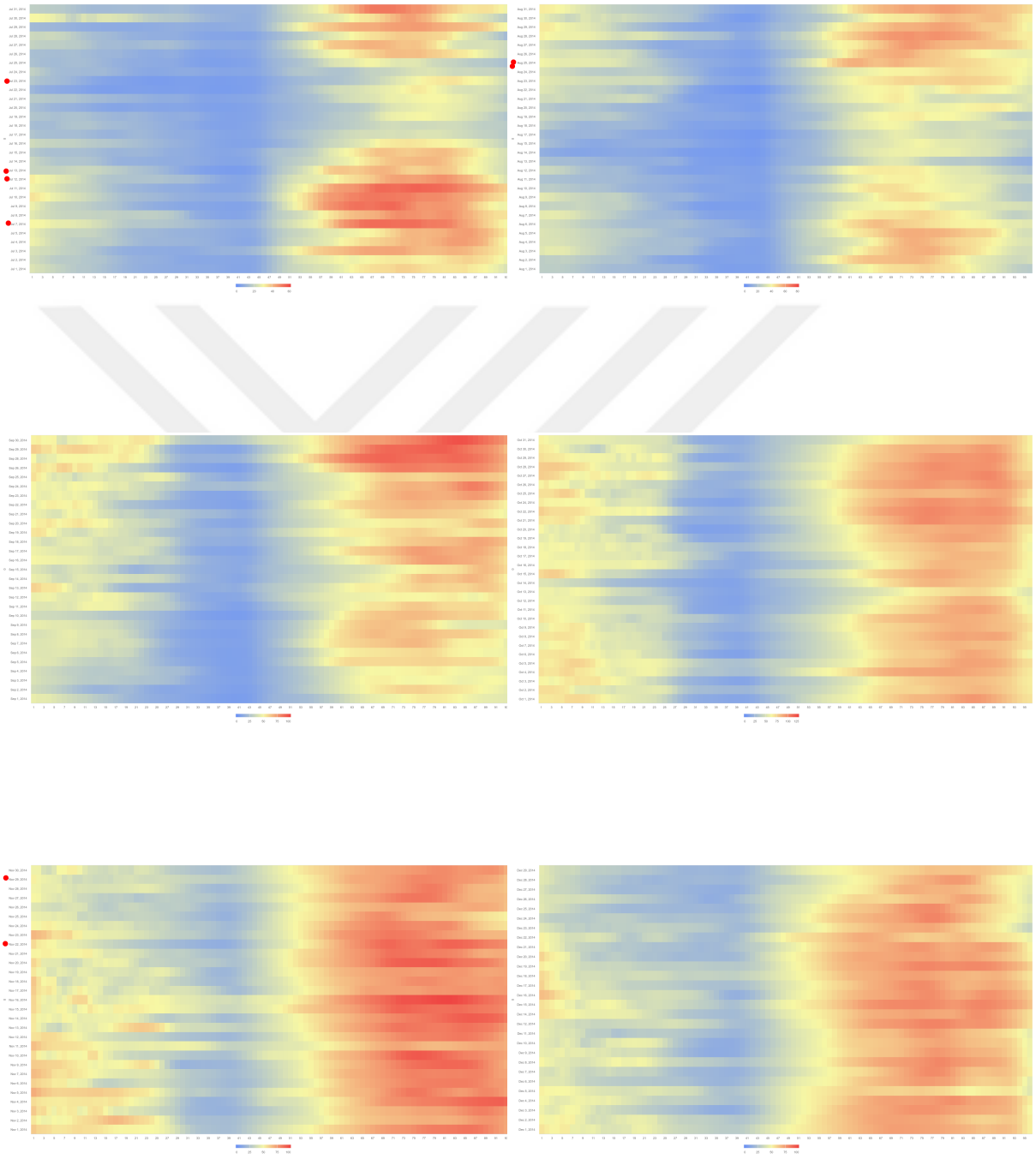
July 2013 to December 2013

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



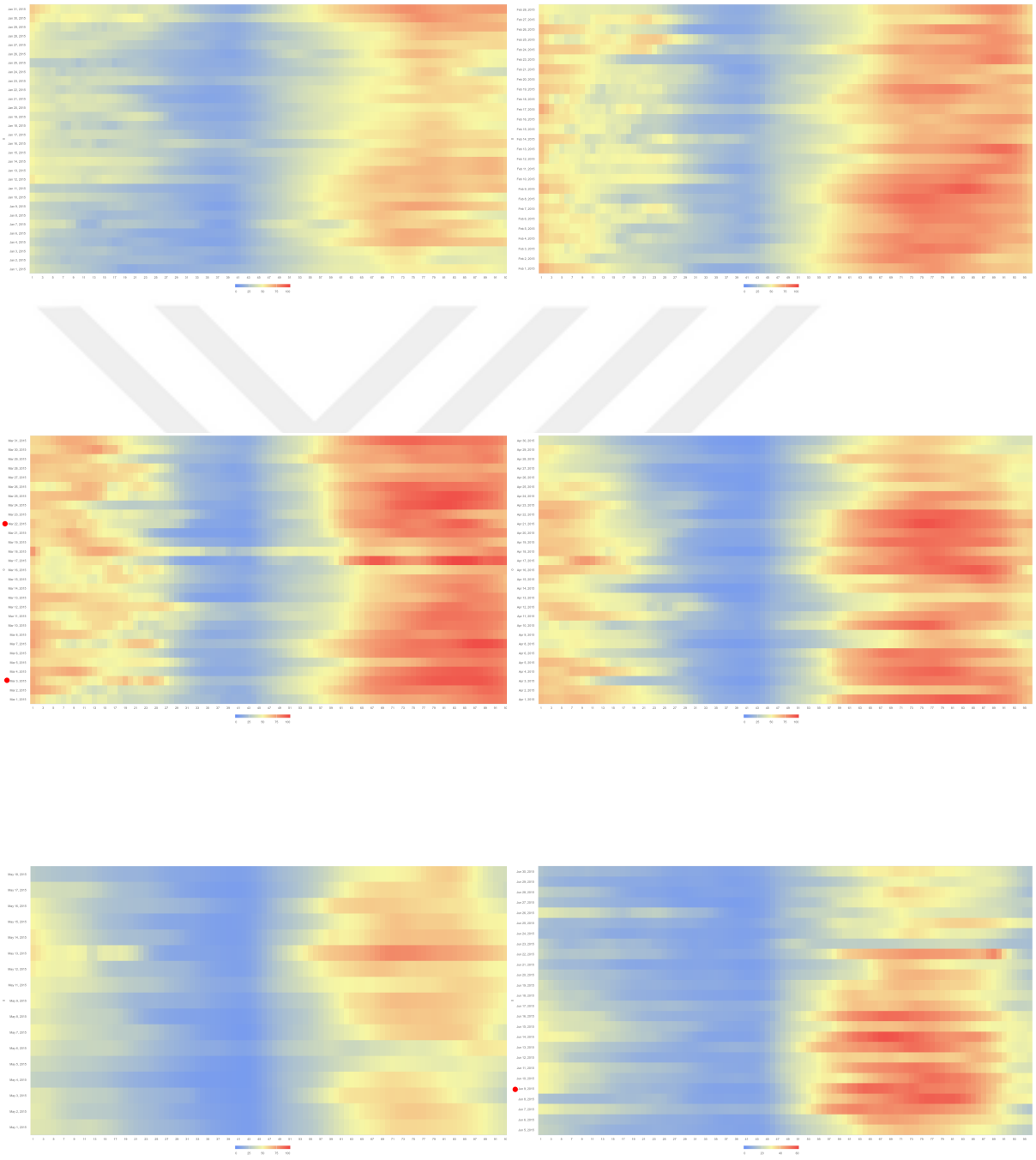
January 2014 to June 2014

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



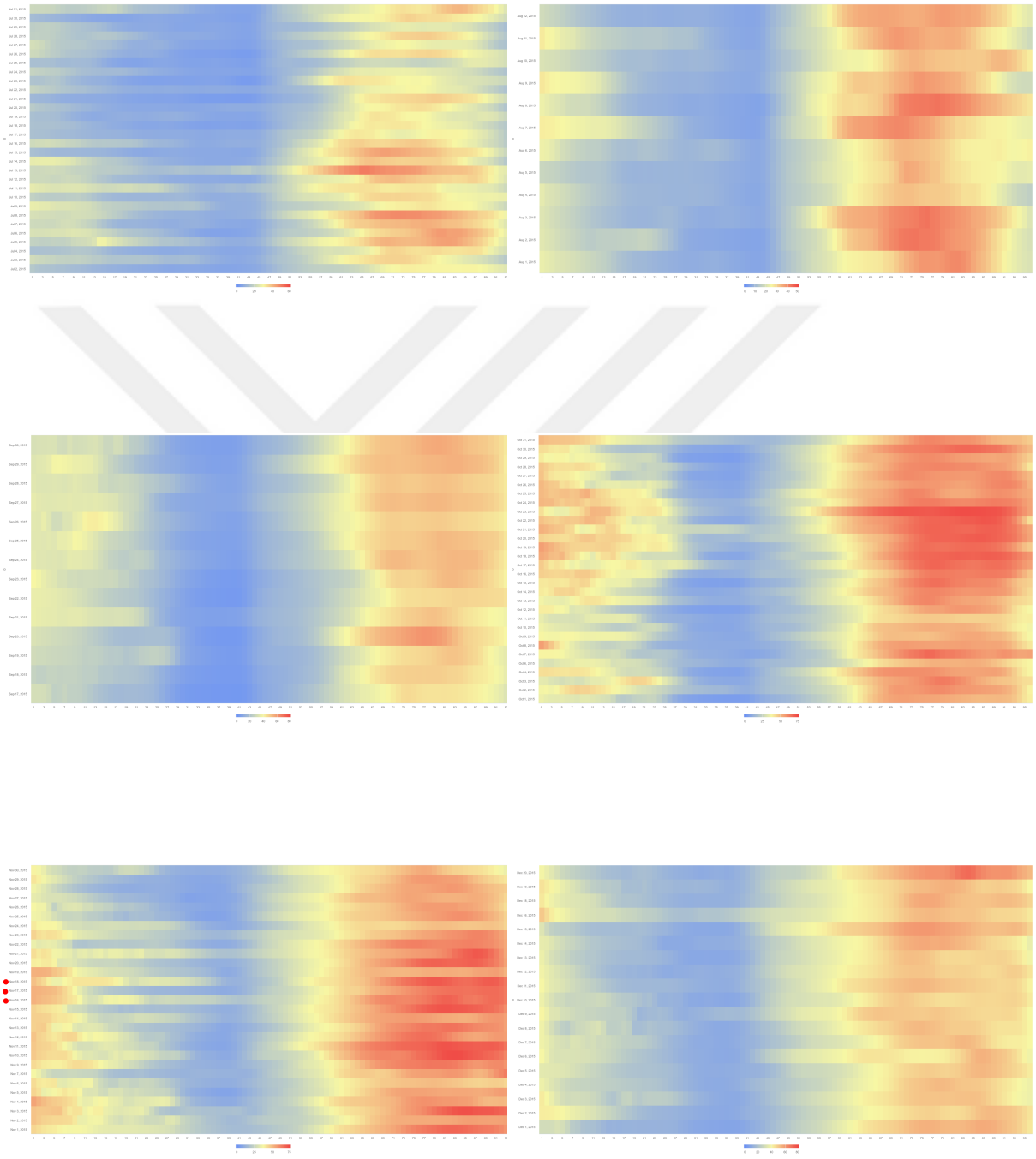
July 2014 to December 2014

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



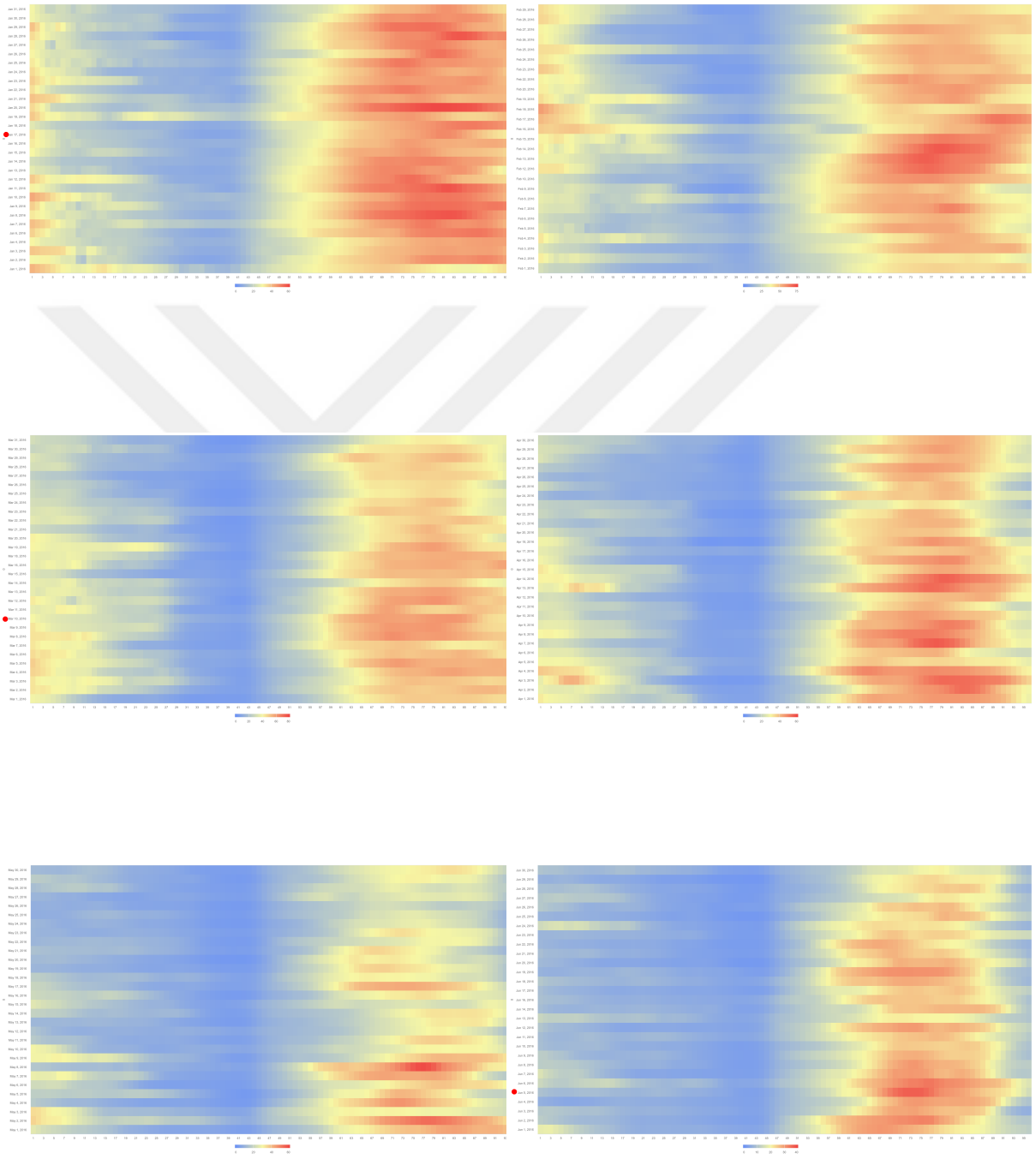
January 2015 to June 2015

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



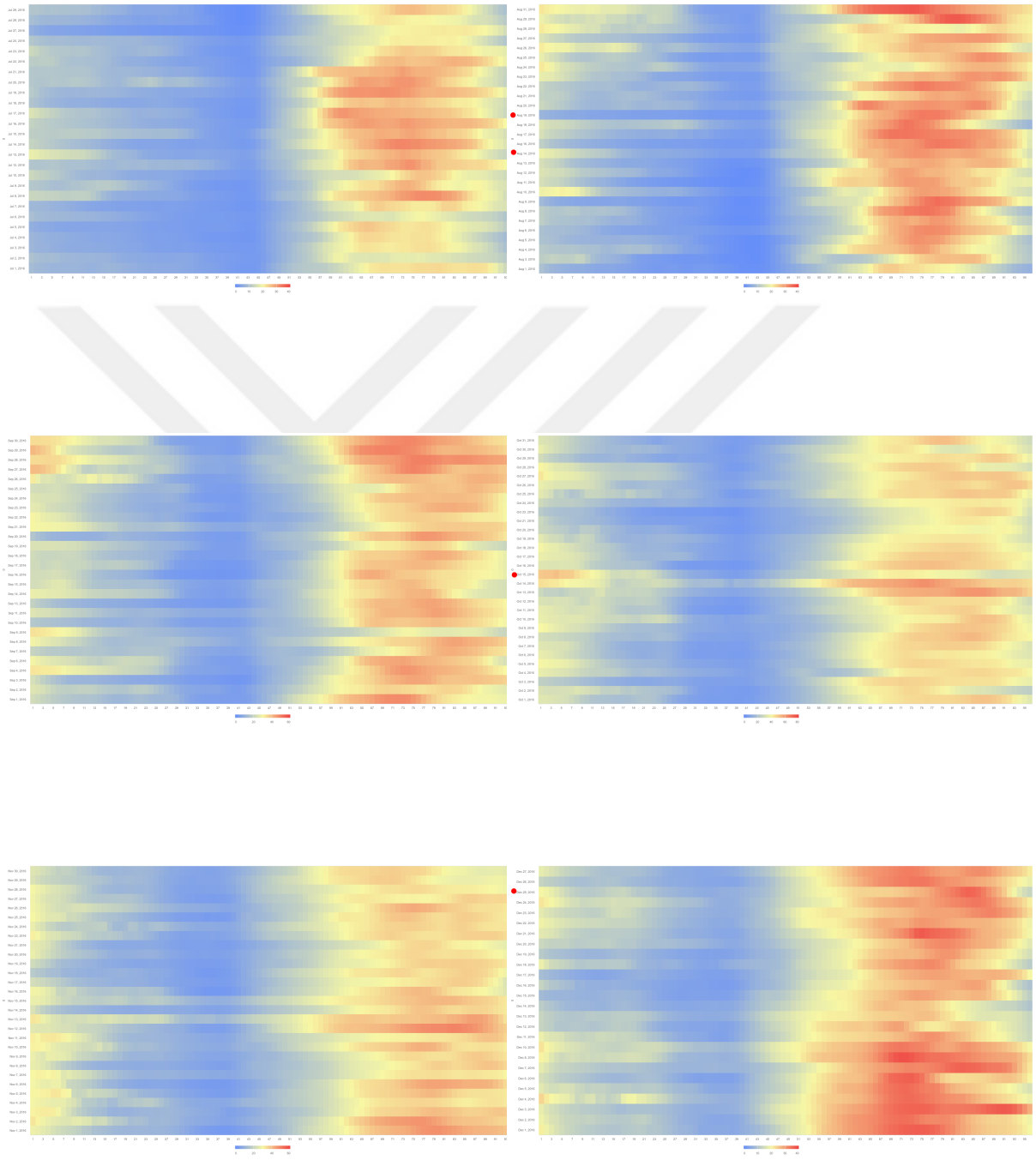
July 2015 to December 2015

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



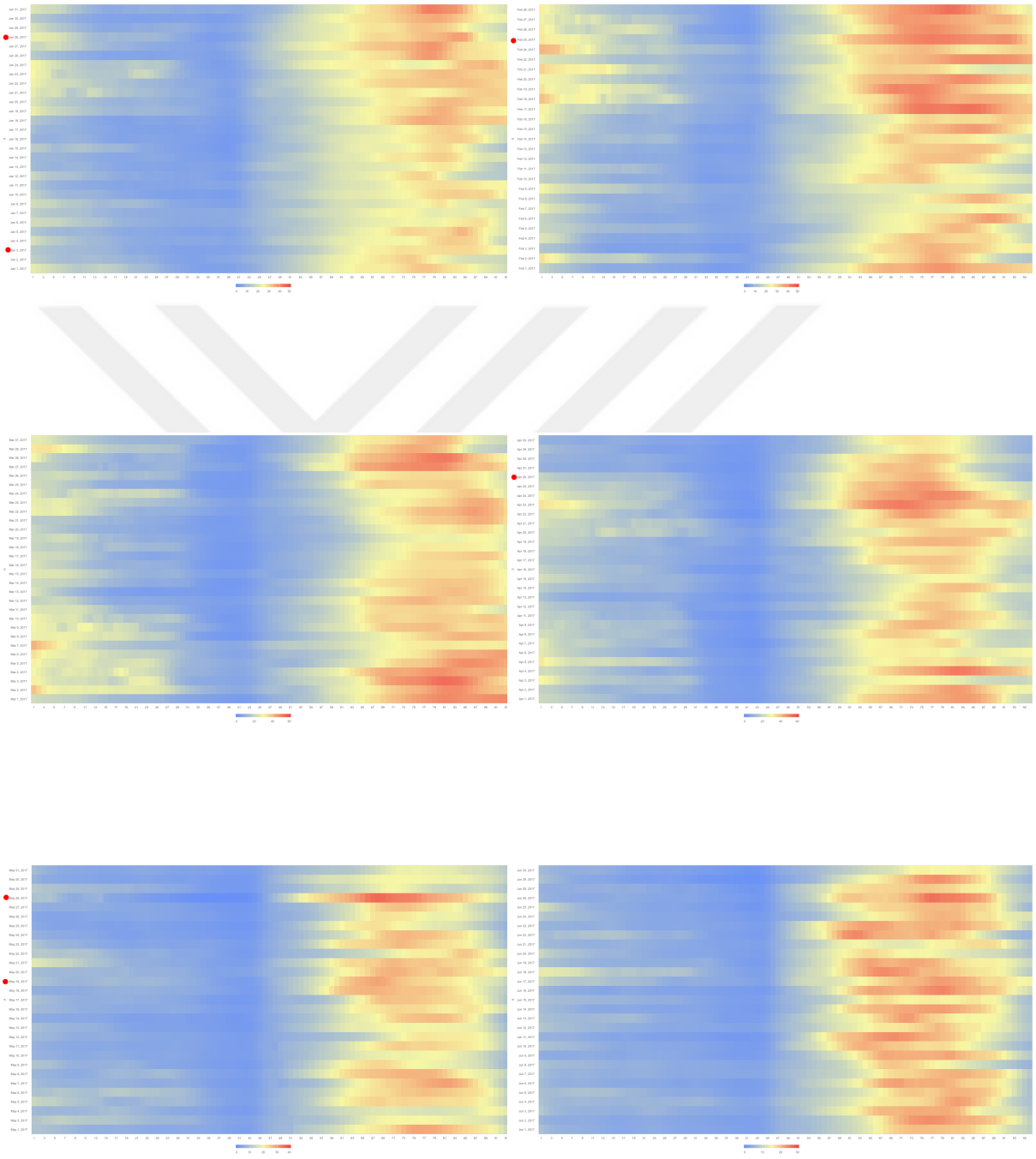
January 2016 to June 2016

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



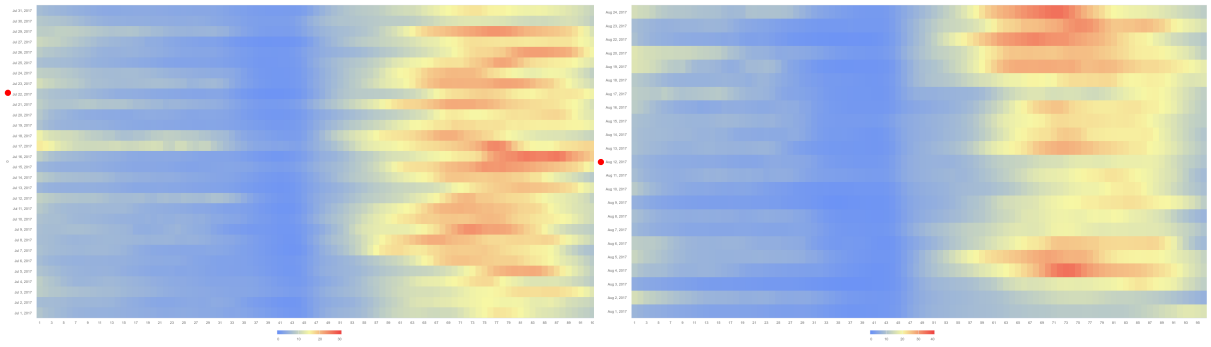
July 2016 to December 2016

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



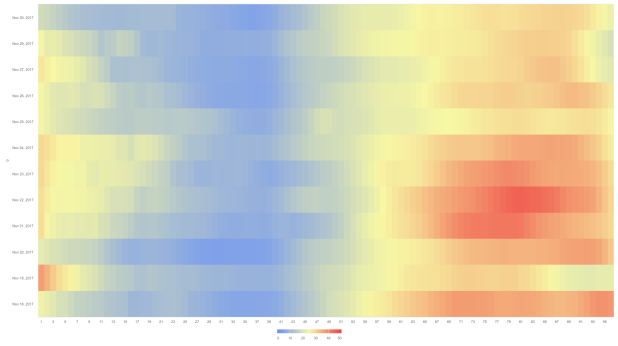
January 2017 to June 2017

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



No Data Available in
GPS Sation

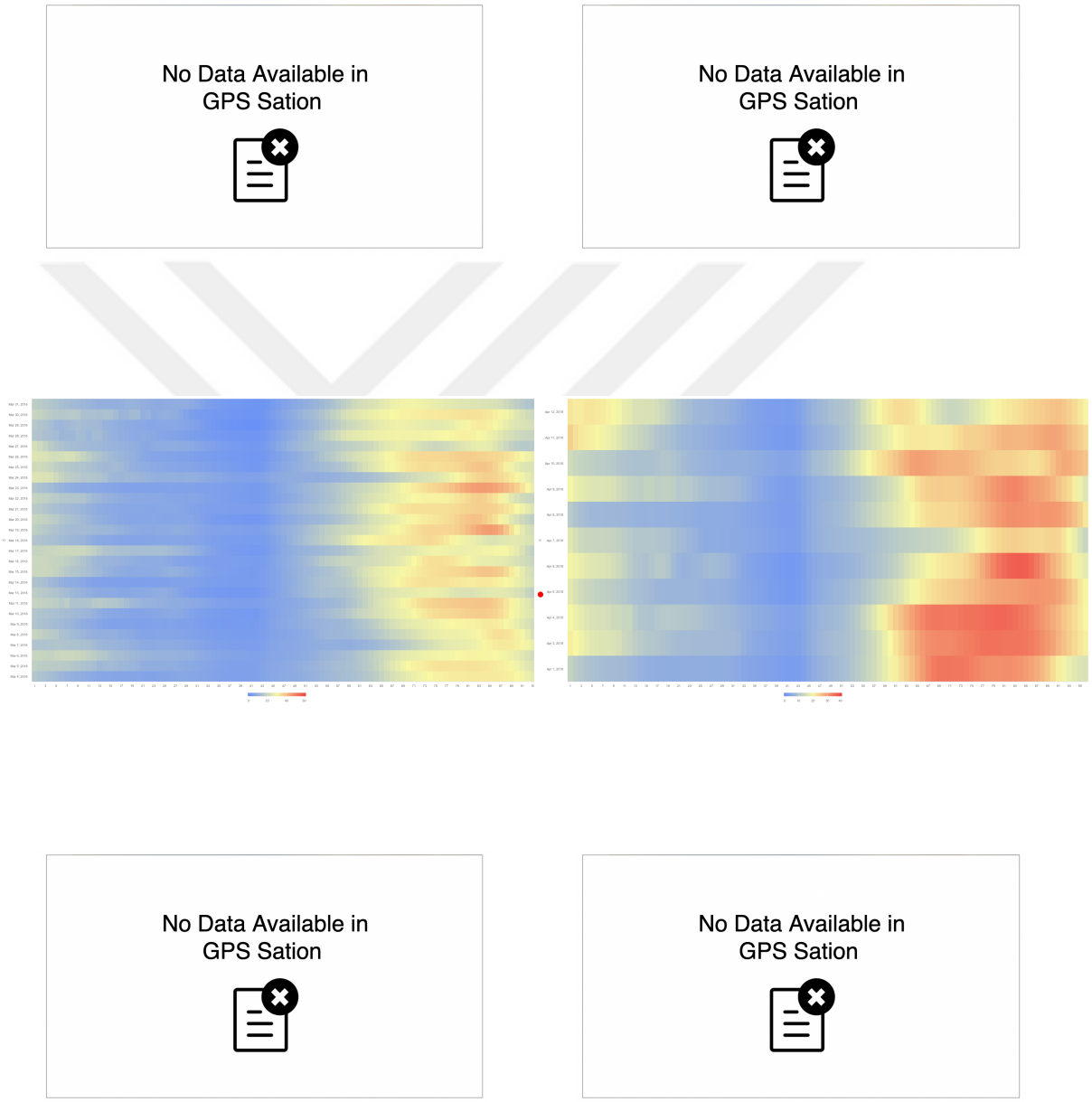
No Data Available in
GPS Sation



No Data Available in
GPS Sation

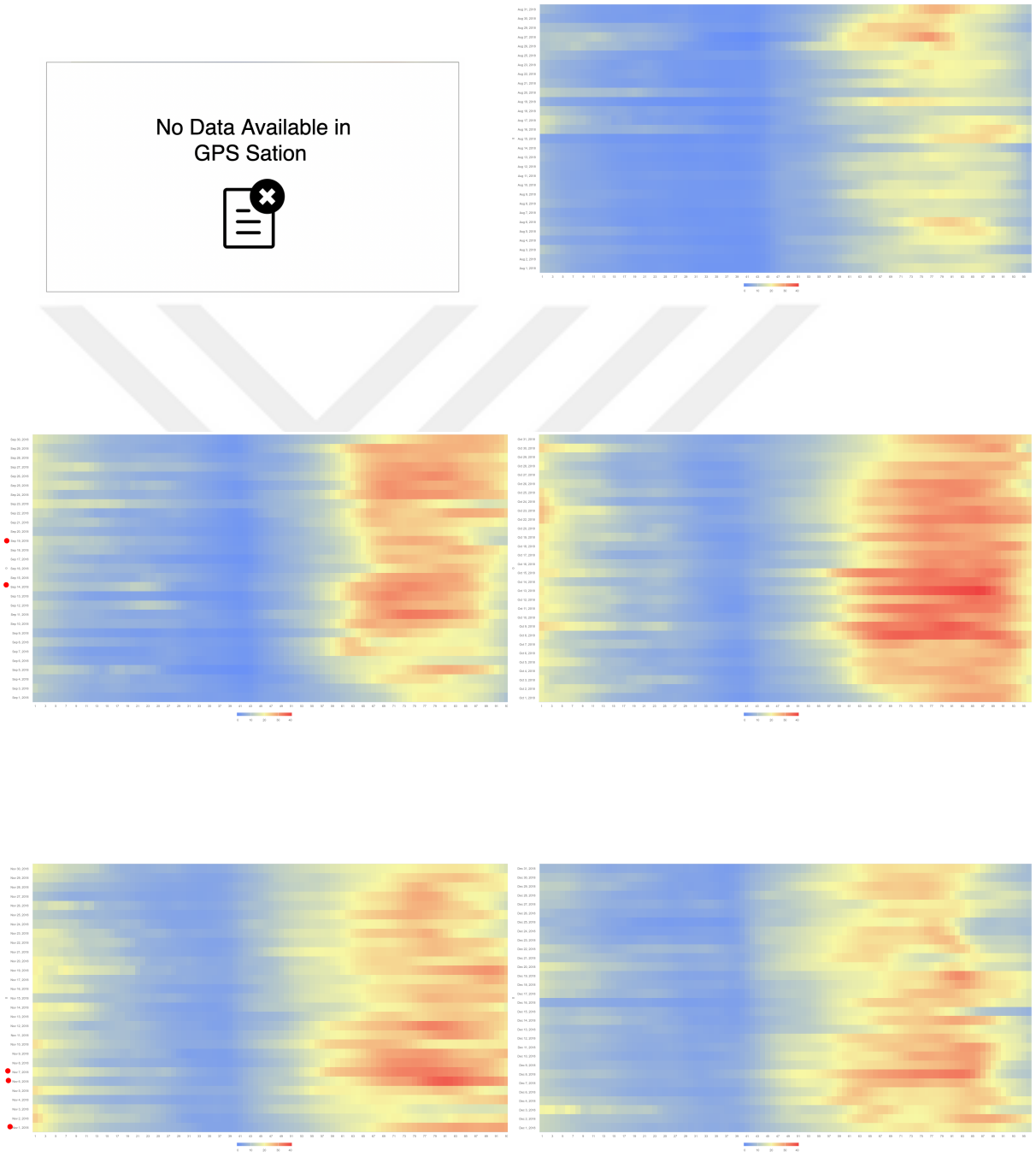
July 2017 to December 2017

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



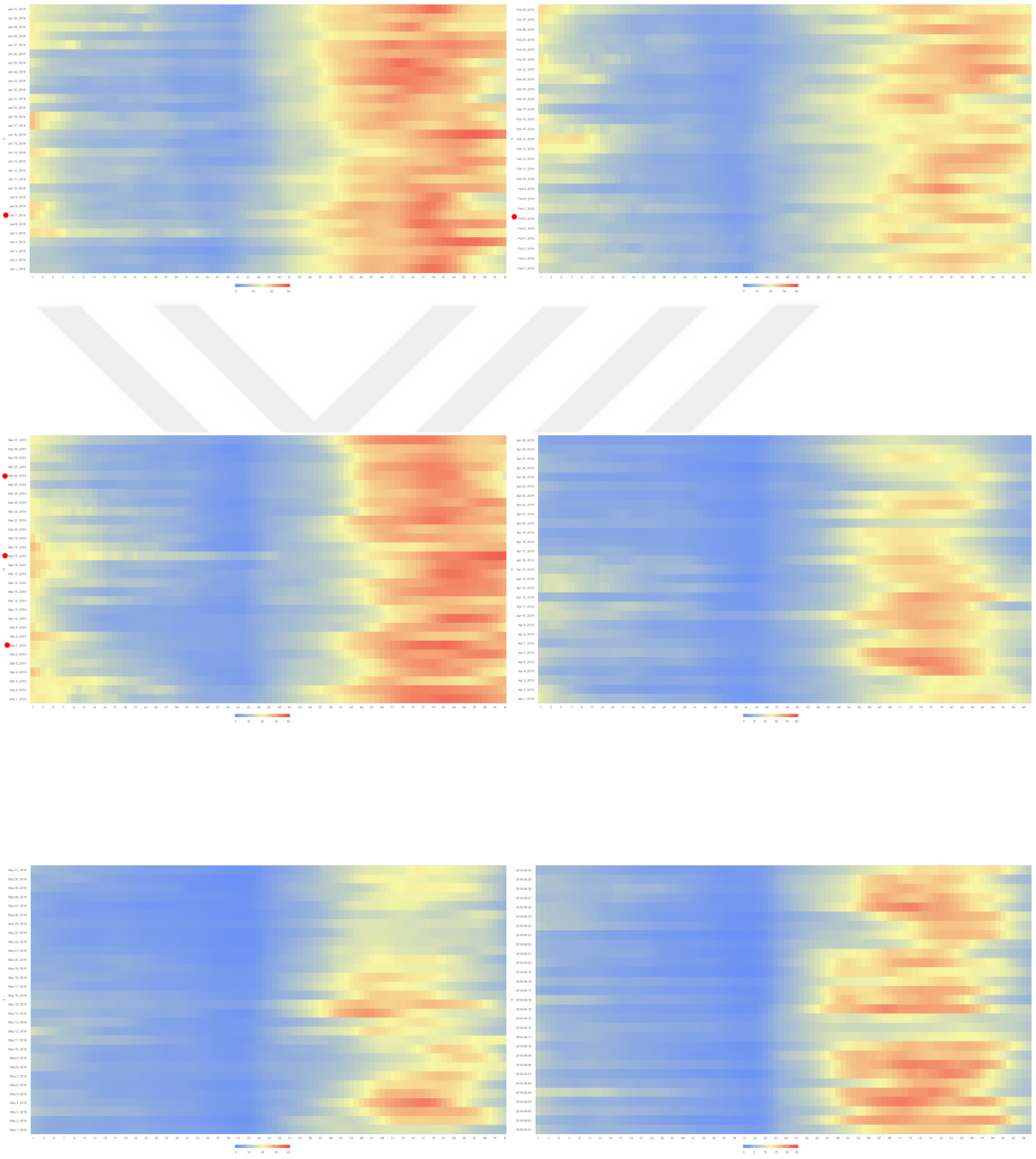
January 2018 to June 2018

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



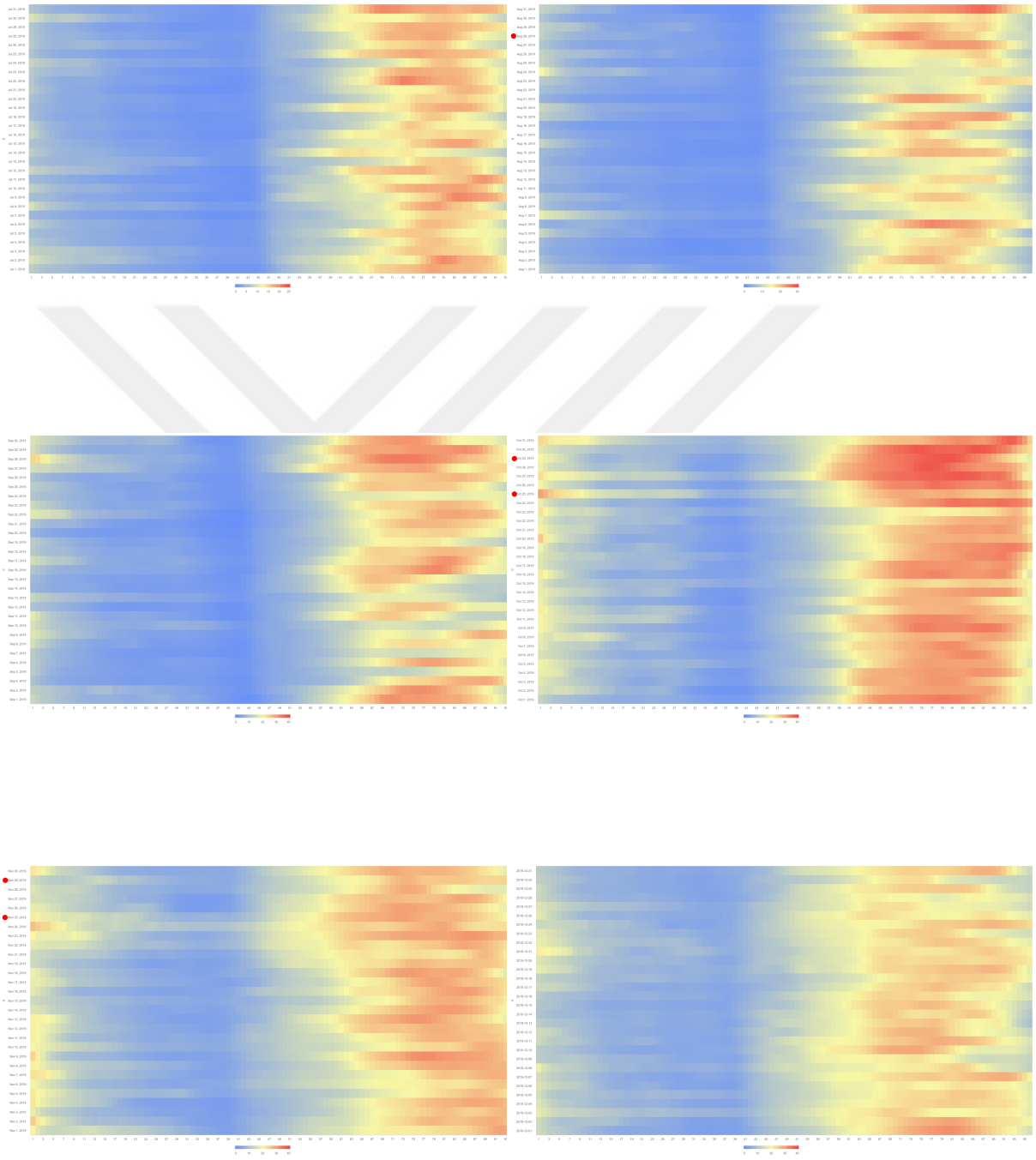
July 2018 to December 2018

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*



January 2019 to June 2019

Appendix A. *Heat Maps of TEC based Data Related to iqqe Station*

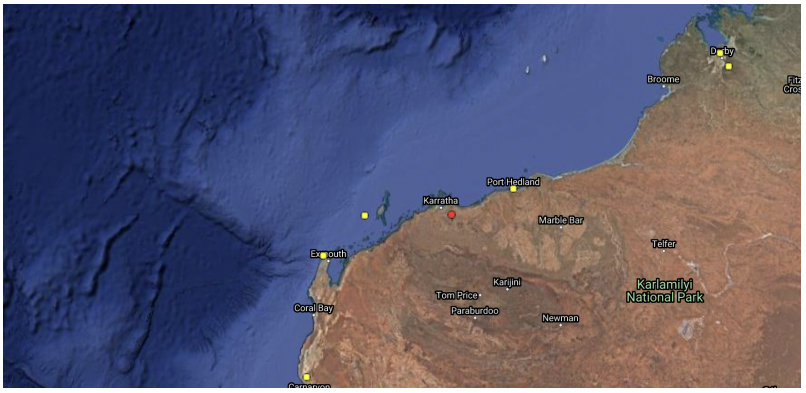


July 2019 to December 2019

Appendix B

Heat Maps of TEC based Data Related to KARRATHA Station

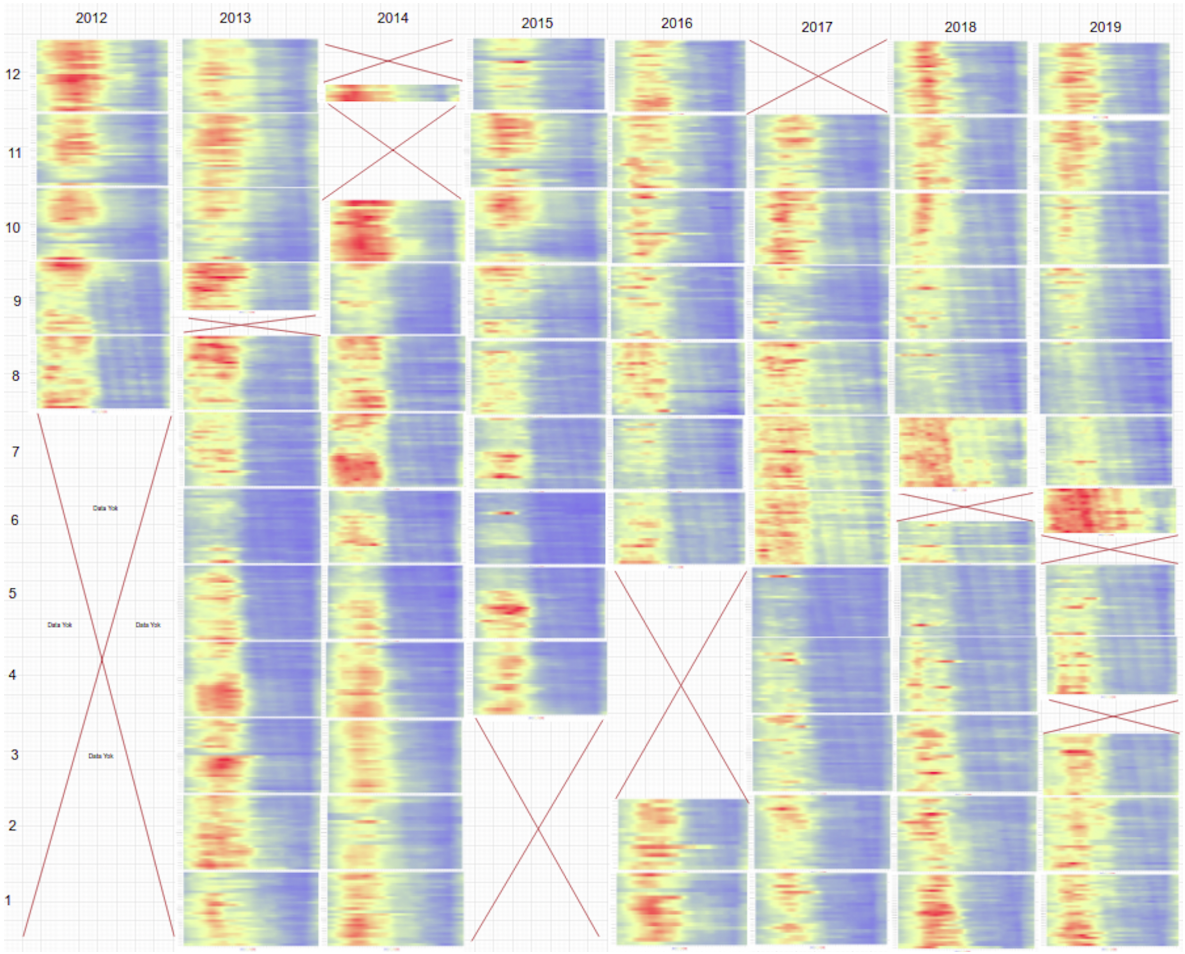
Appendix B depicted the dataset related to the KARRATHA station and the earthquakes. The KARRATHA station is located in coordinate $(Lat : -20.85, Lon : 117.1)$ the Pilbara region of Western Australia represented in the figure. TEC data gathered from GPS stations (Dual-Frequency GPS receiver) from the IONOLAB group (Hacettepe University of IONOLAB is an electrical engineer organization to investigate the ionosphere hurdles.) [1]. The earthquake information is collected via (United States Geological Survey of Earthquakes) [2]. All earthquakes in the dataset are marked with the red points in the pictures.



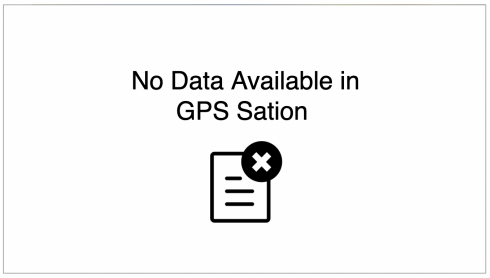
[1] Available at http://ionolab.org/
[2] Available at https://earthquake.usgs.gov/

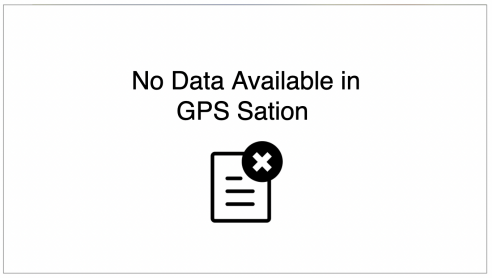
Appendix B. *Heat Maps of TEC based Data Related to karr Station*

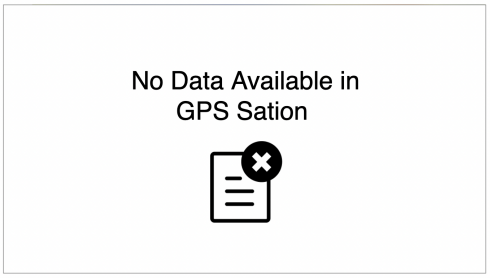
Overview of whole 2012-1019 Dataset related to the IQUIQUE GPS Station as fallow:

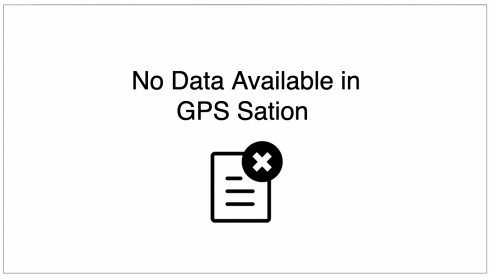


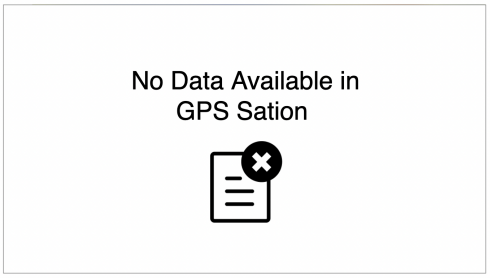
Appendix B. *Heat Maps of TEC based Data Related to karr Station*

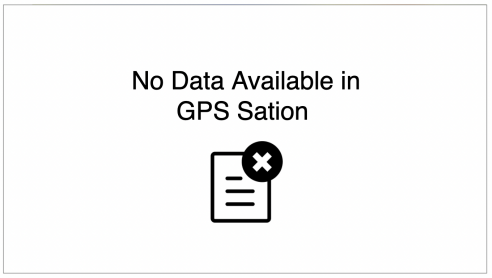






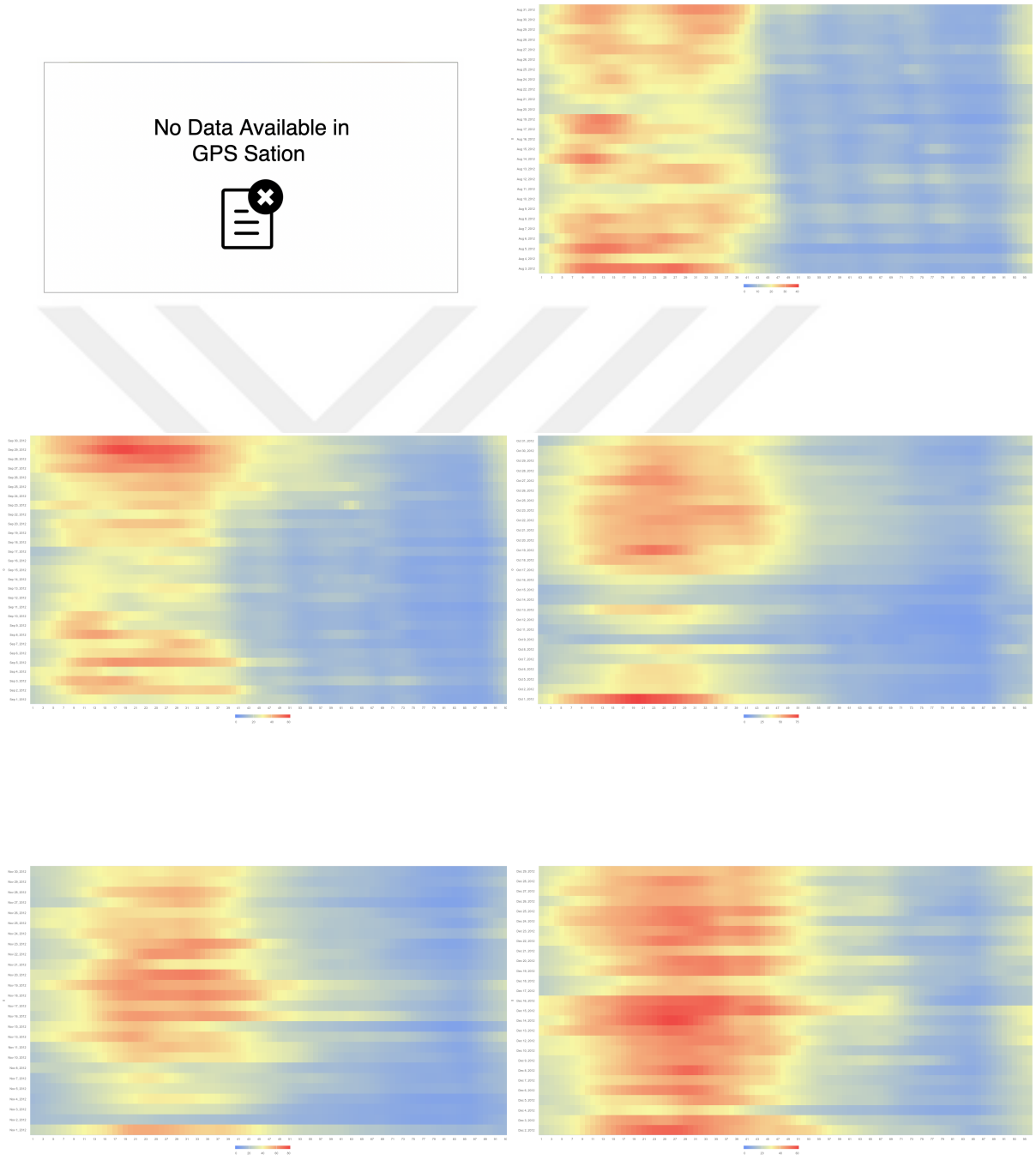






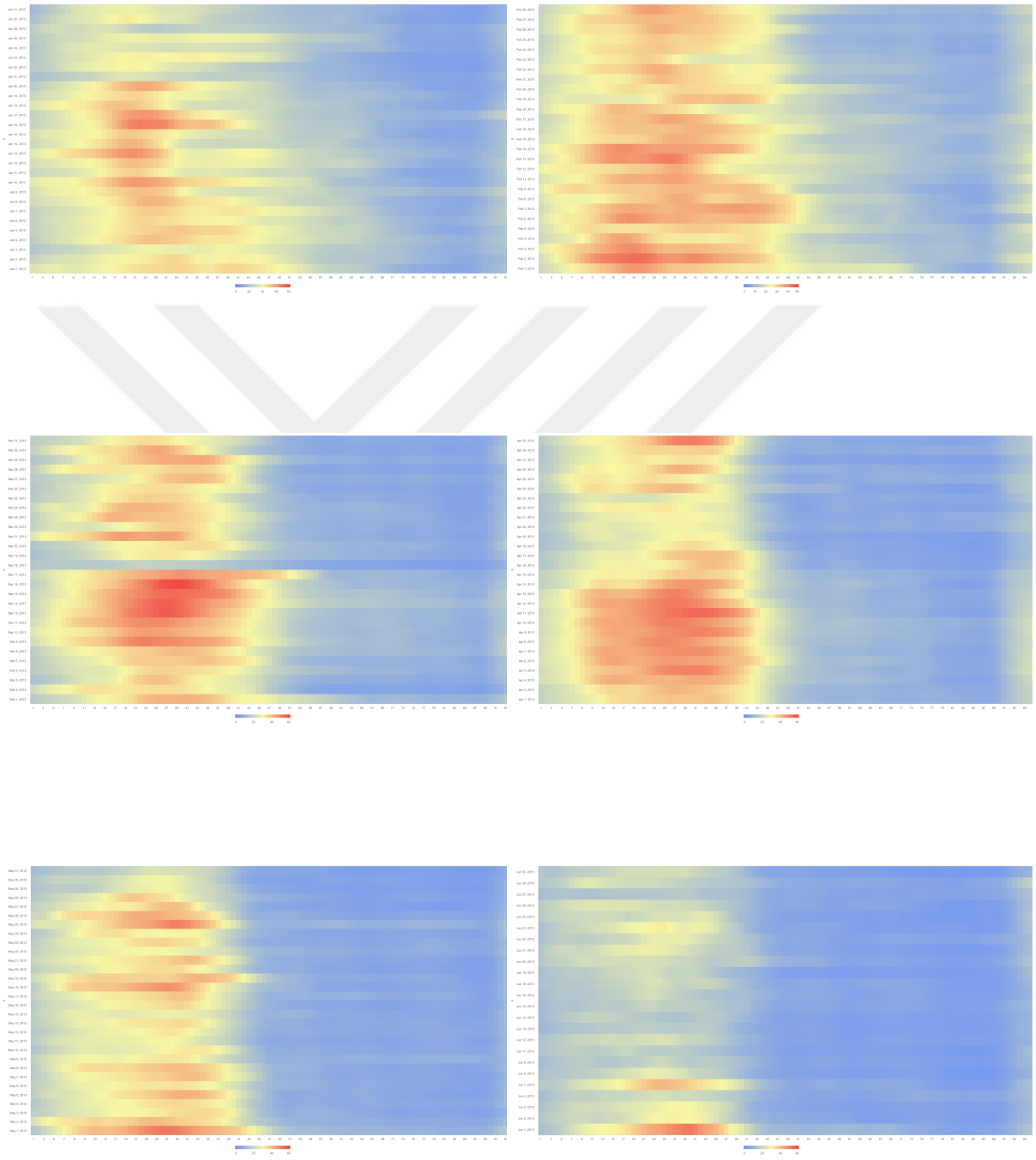
January 2012 to June 2012

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



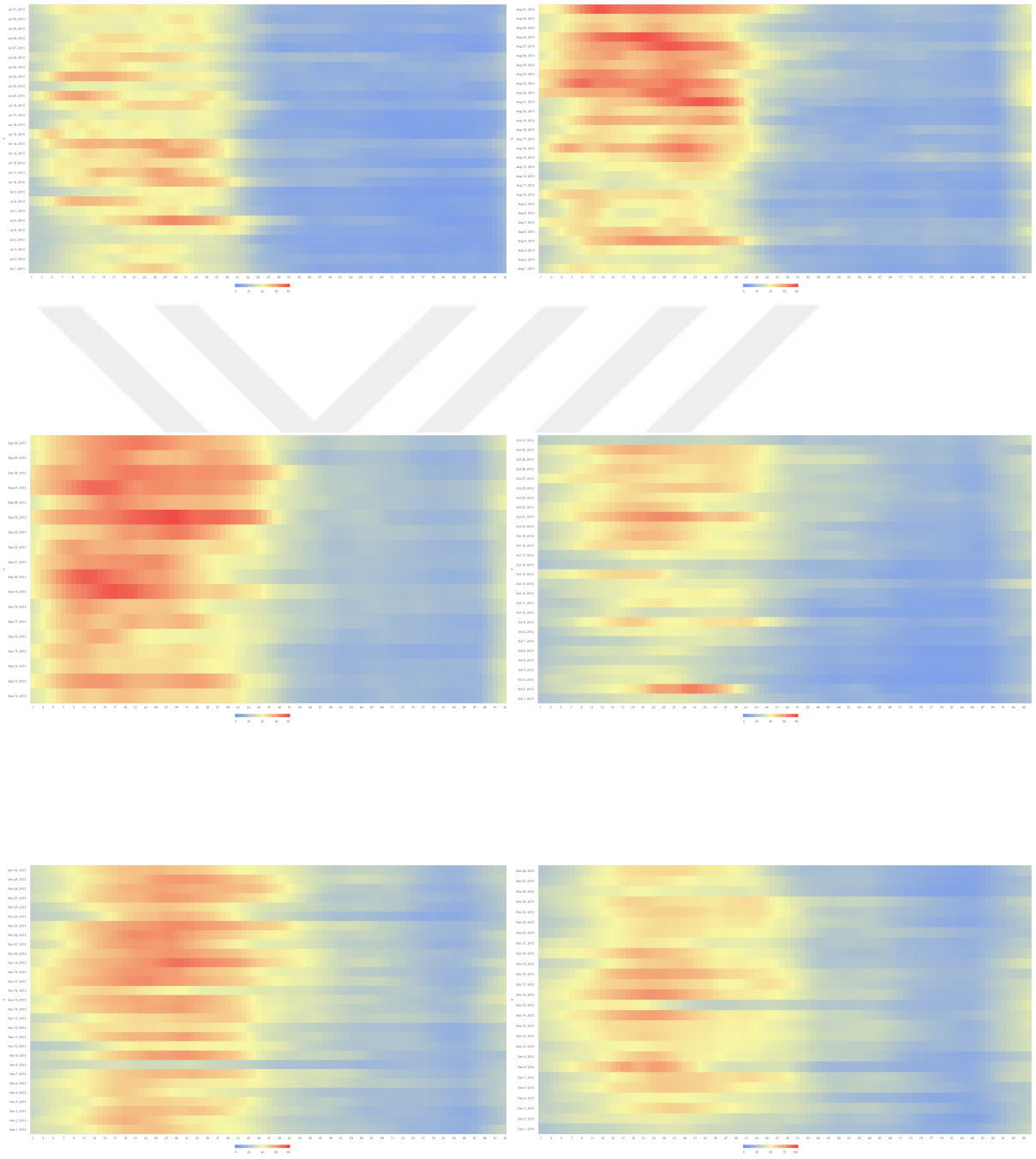
July 2012 to December 2012

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



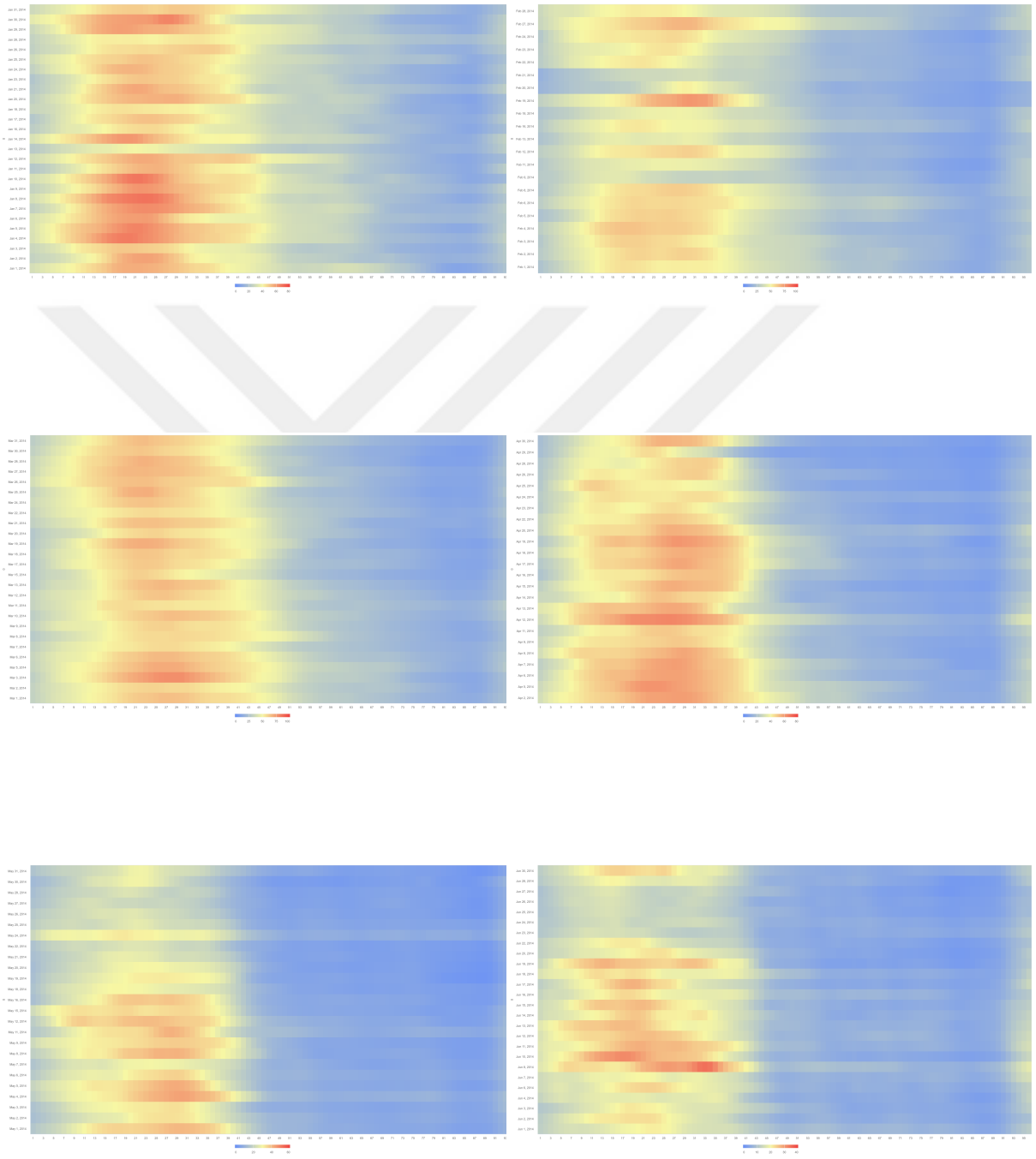
January 2013 to June 2013

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



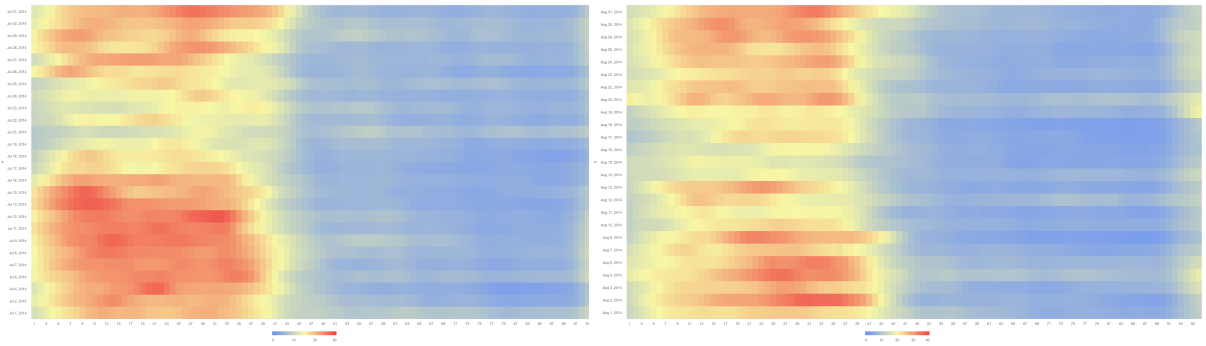
July 2013 to December 2013

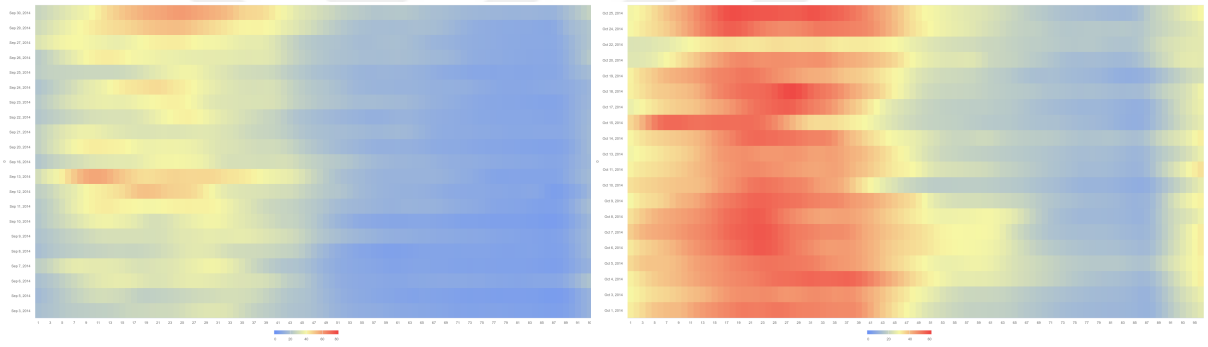
Appendix B. *Heat Maps of TEC based Data Related to karr Station*

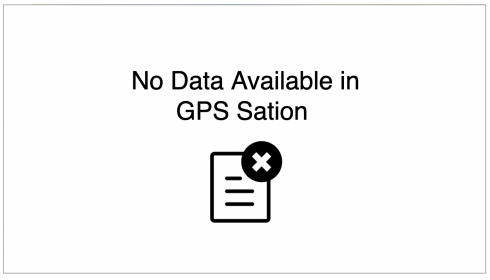


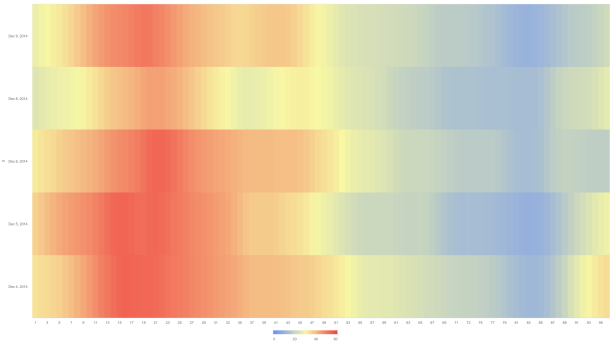
January 2014 to June 2014

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



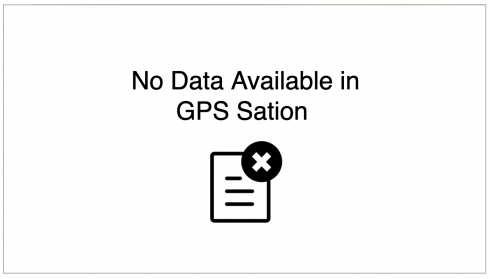


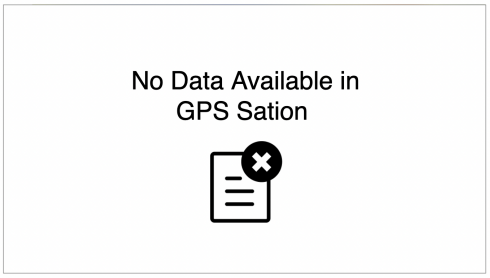


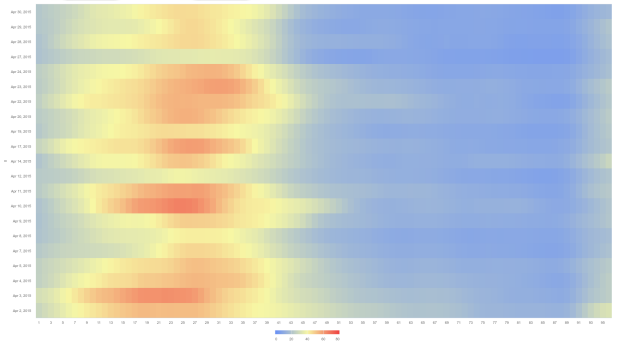


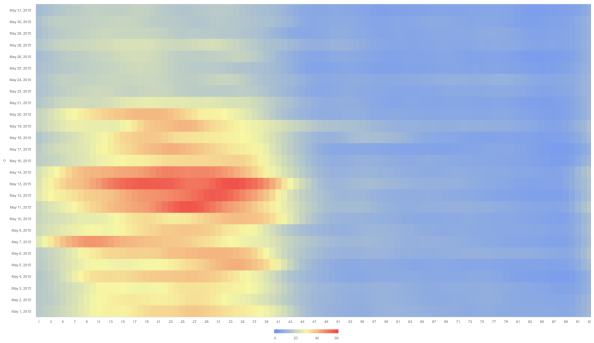
July 2014 to December 2014

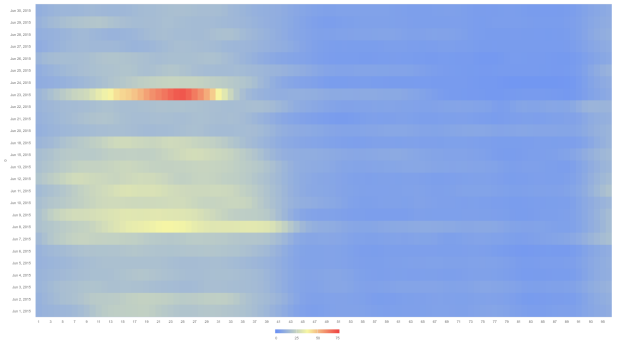
Appendix B. *Heat Maps of TEC based Data Related to karr Station*





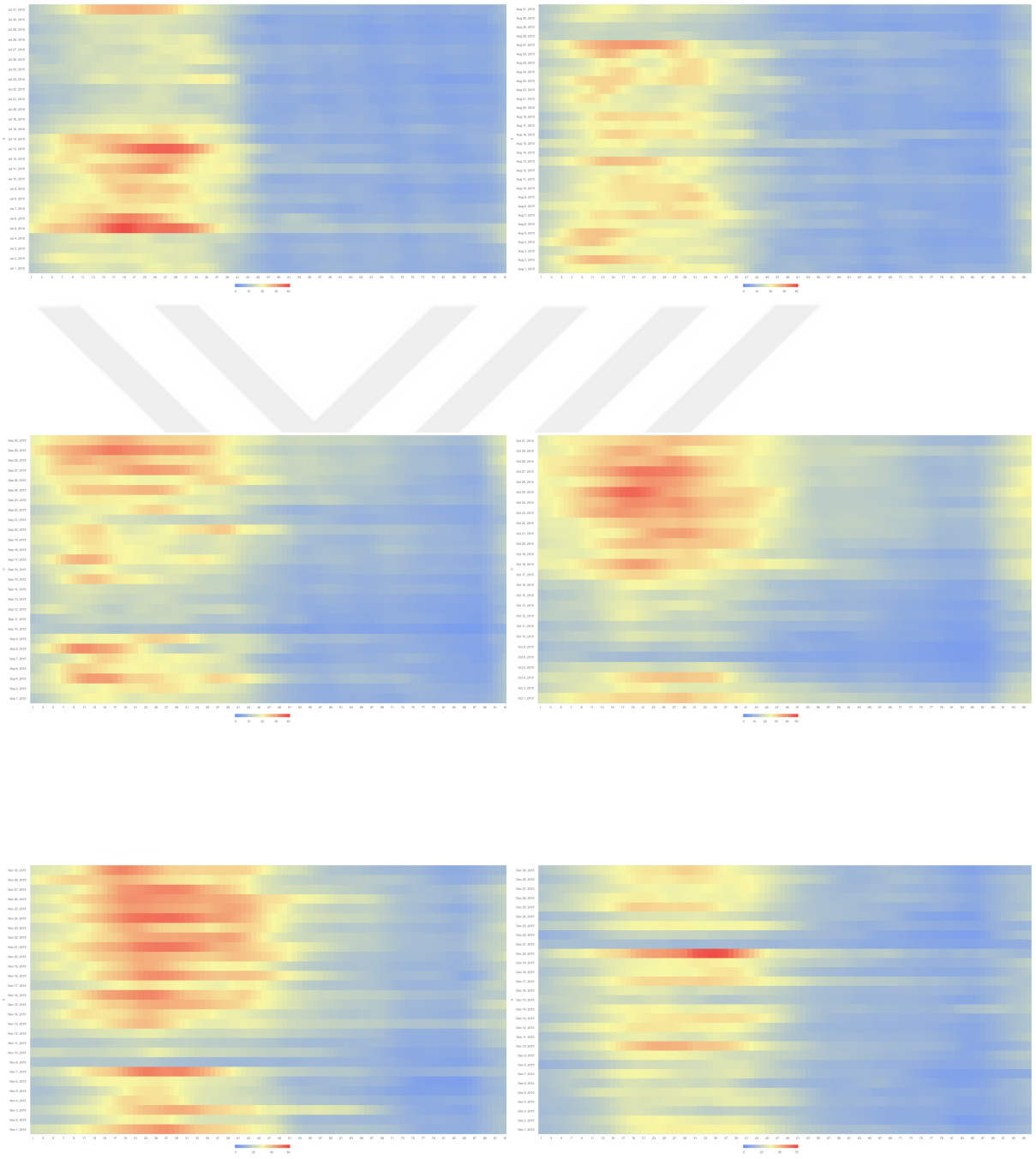






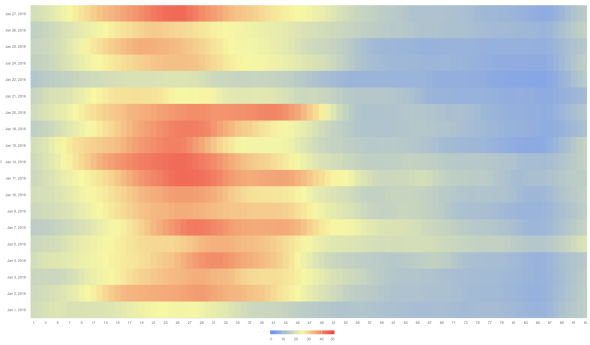
January 2015 to June 2015

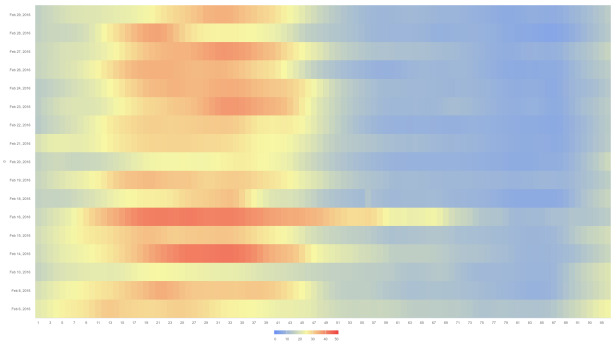
Appendix B. *Heat Maps of TEC based Data Related to karr Station*



July 2015 to December 2015

Appendix B. *Heat Maps of TEC based Data Related to karr Station*

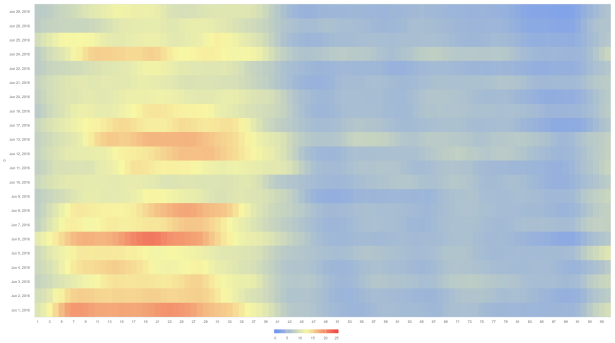




No Data Available in
GPS Sation

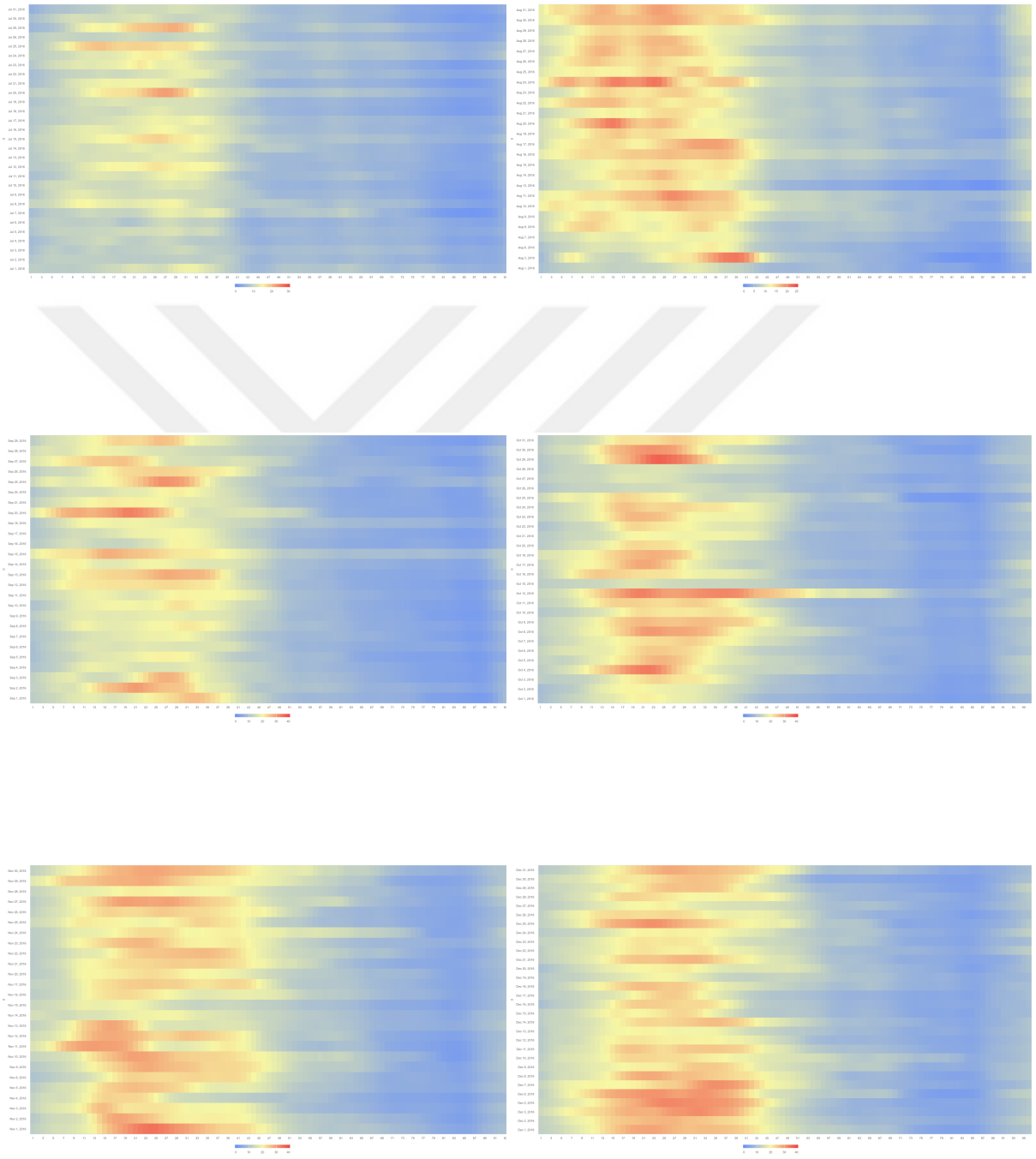
No Data Available in
GPS Sation

No Data Available in
GPS Sation



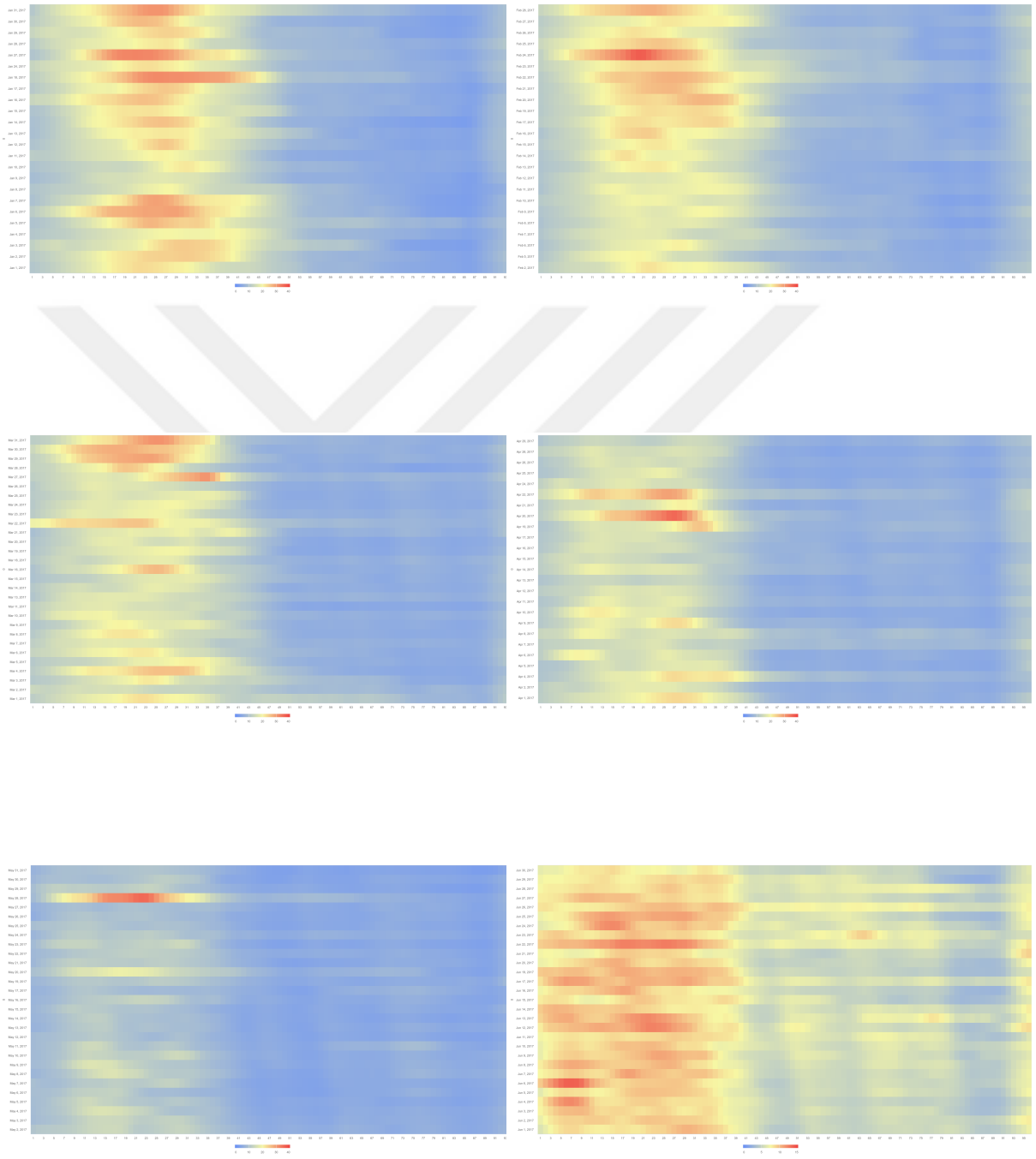
January 2016 to June 2016

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



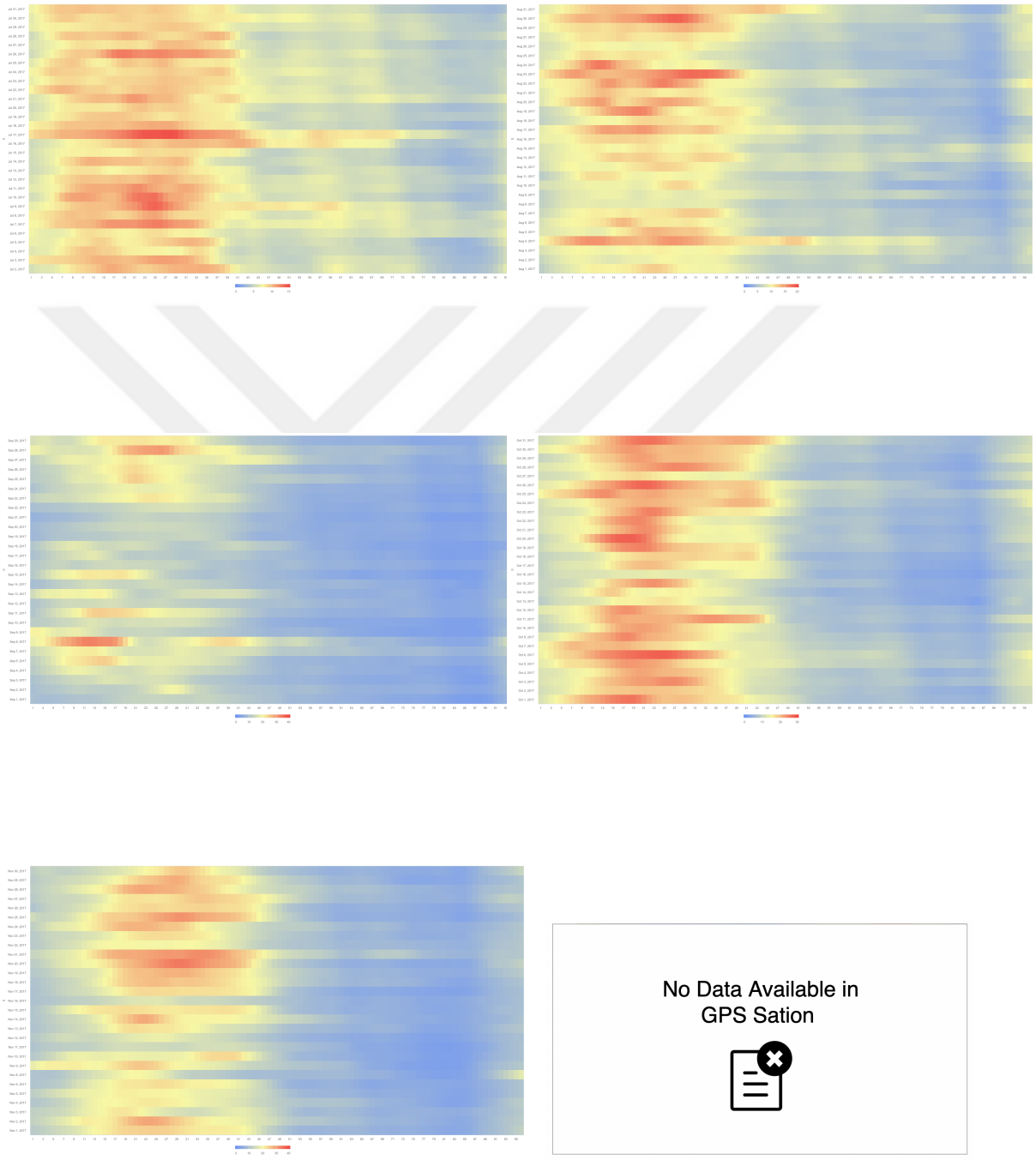
July 2016 to December 2016

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



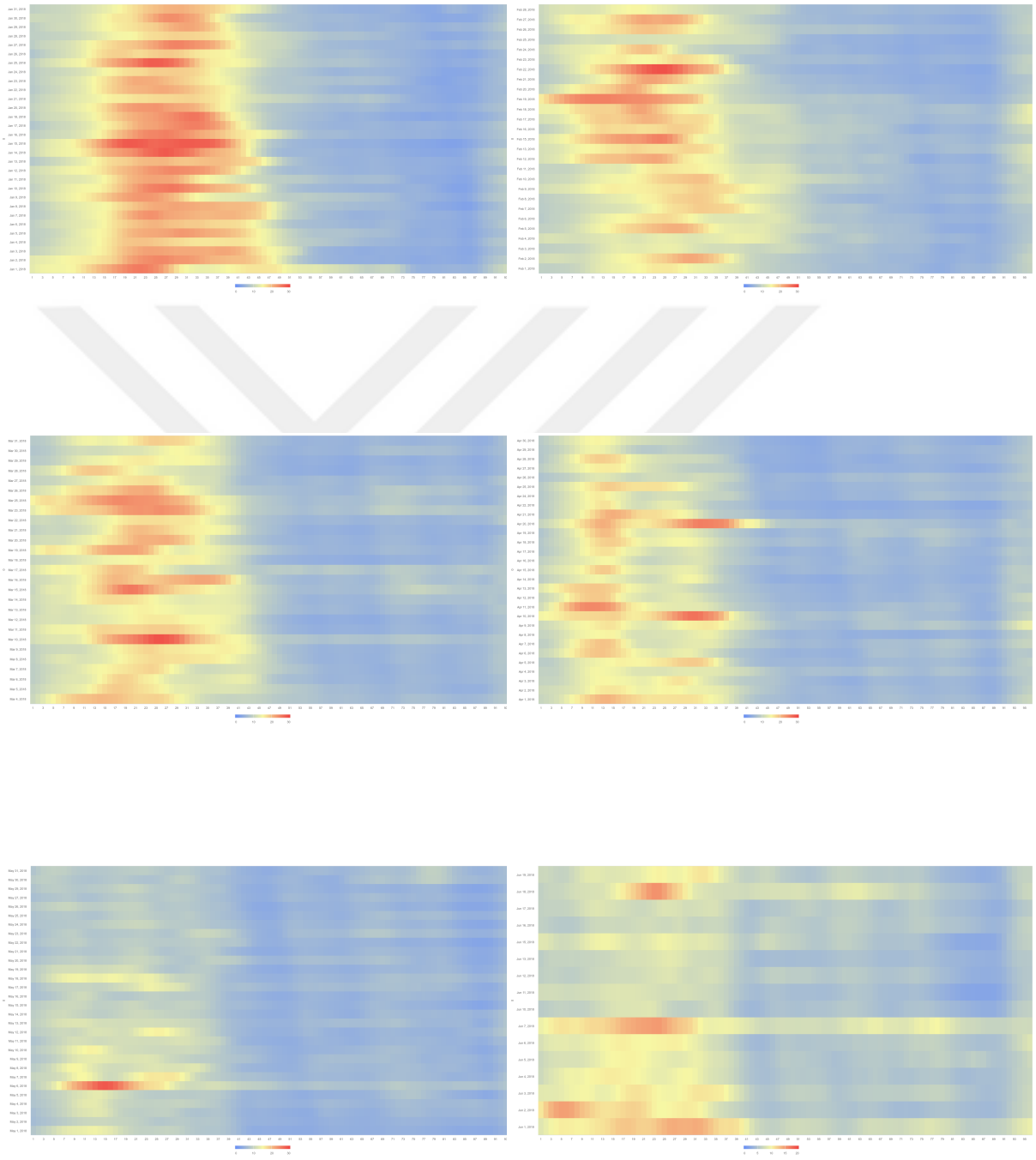
January 2017 to June 2017

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



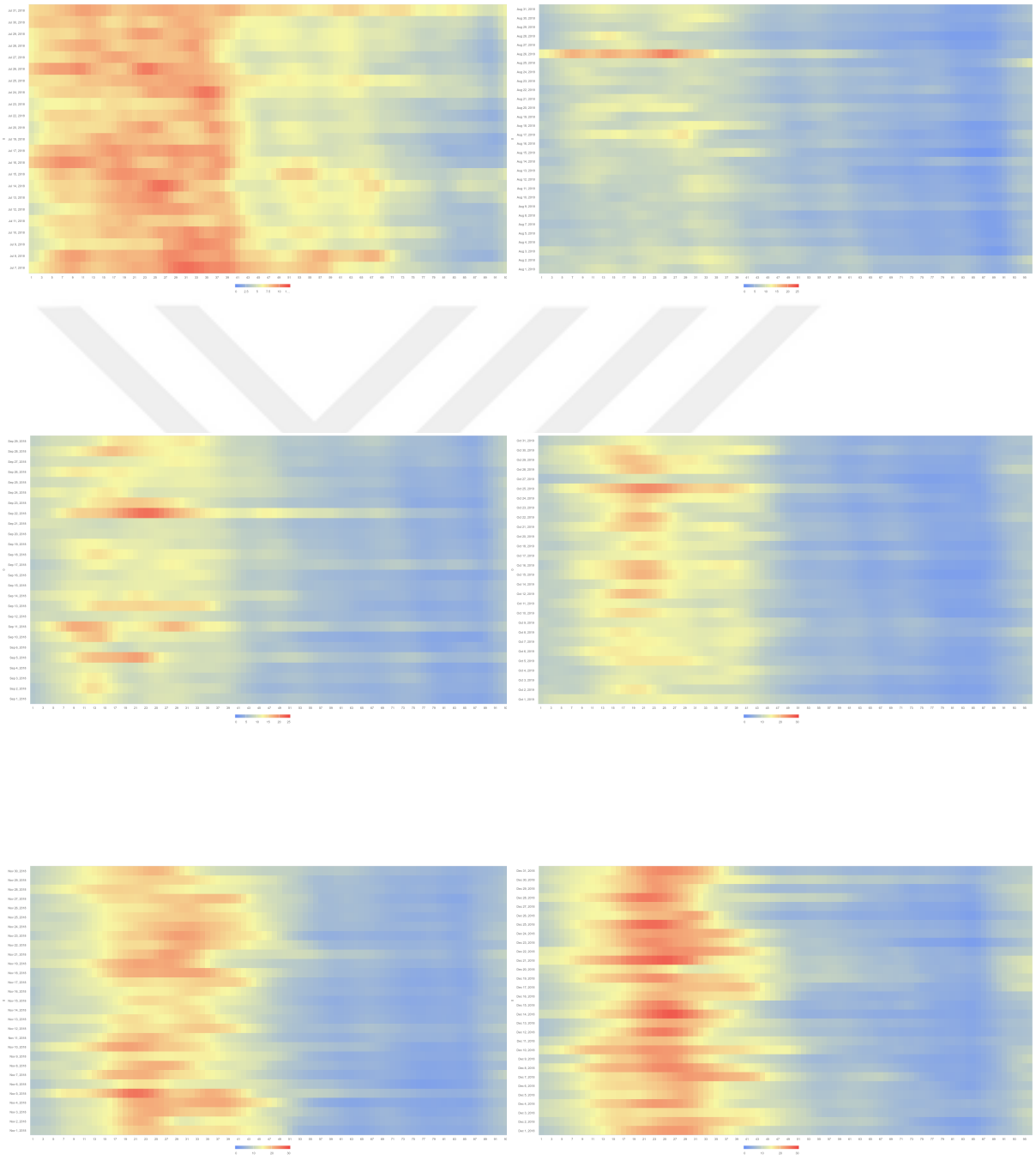
July 2017 to December 2017

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



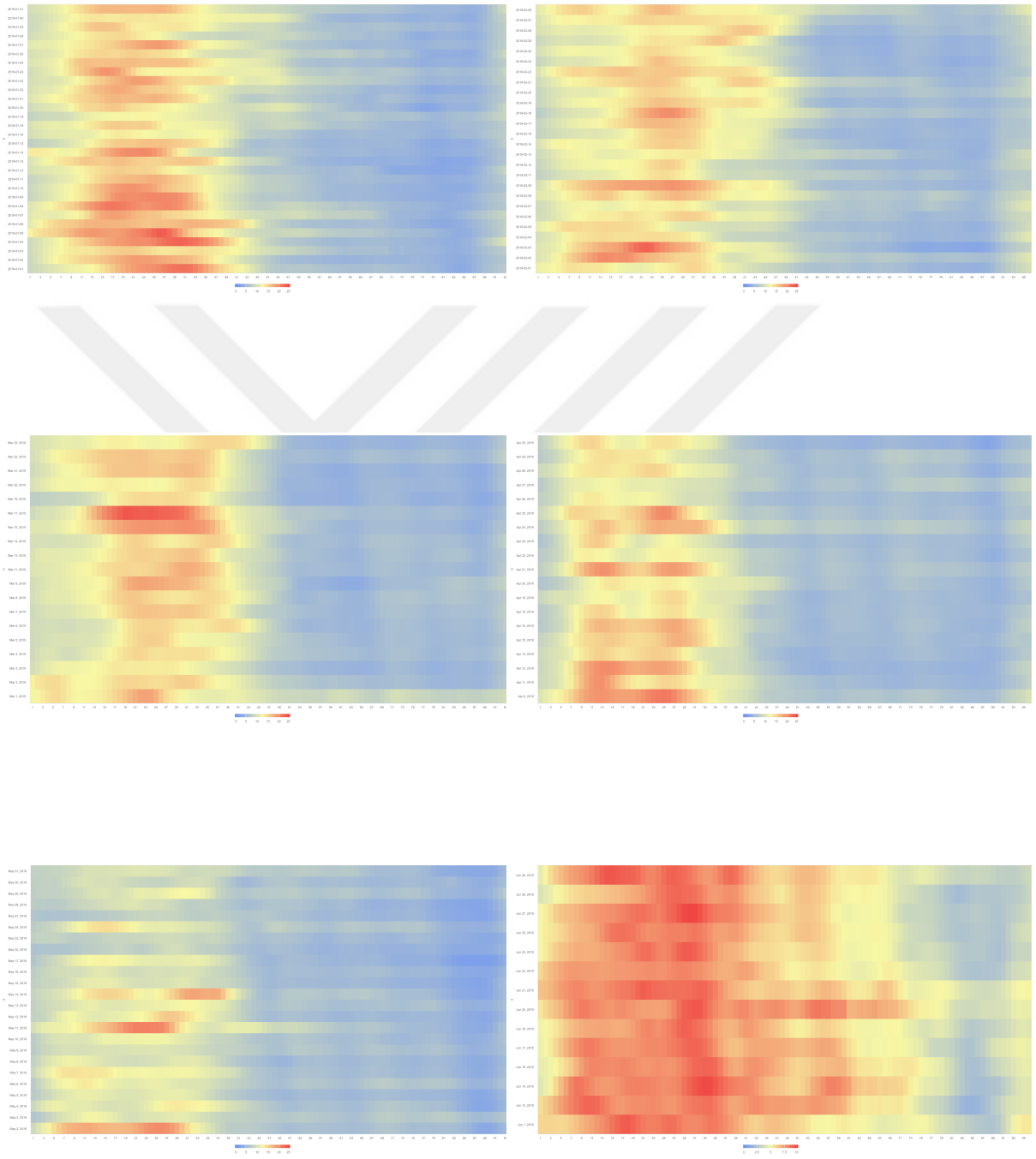
January 2018 to June 2018

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



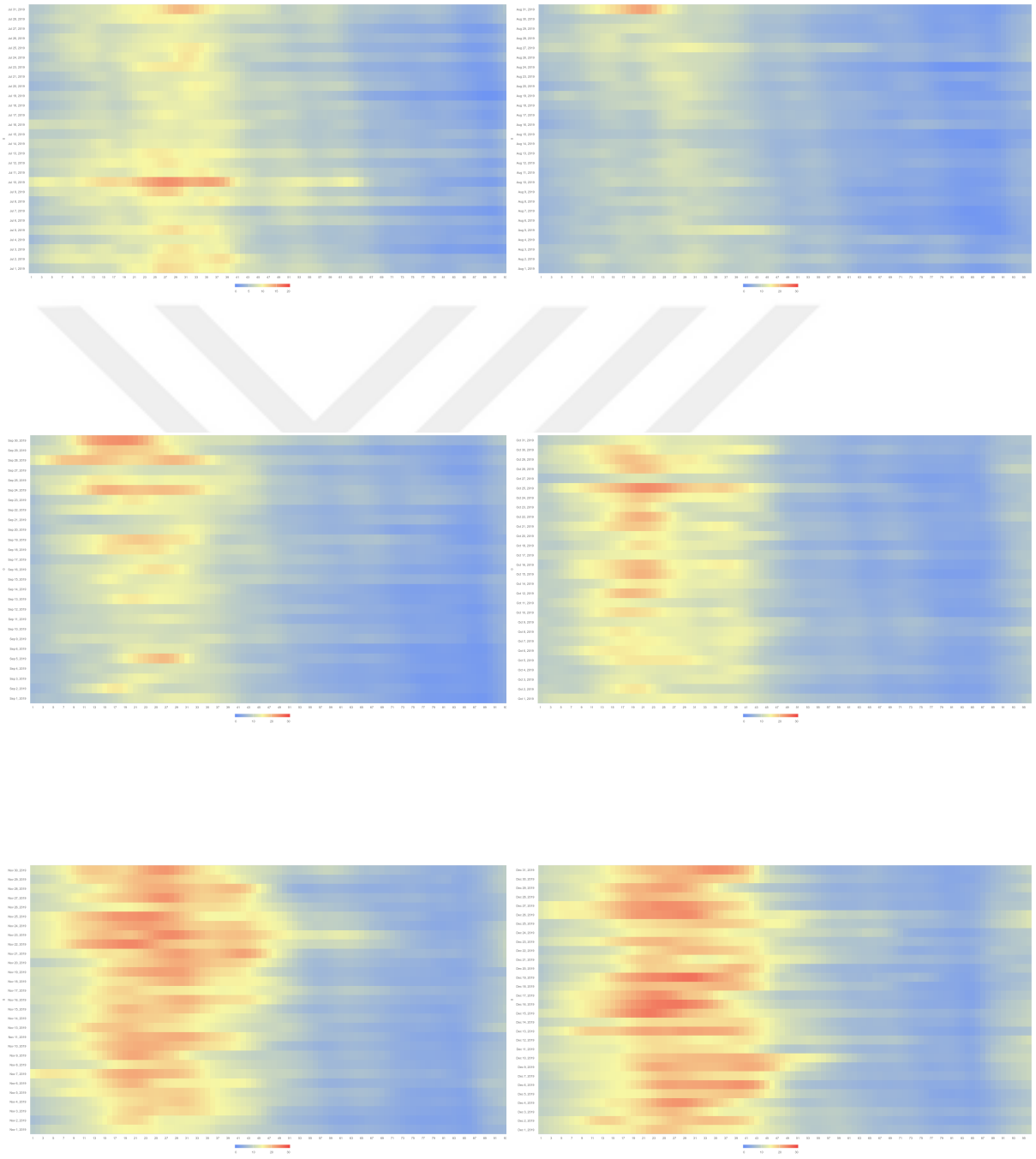
July 2018 to December 2018

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



January 2019 to June 2019

Appendix B. *Heat Maps of TEC based Data Related to karr Station*



July 2019 to December 2019

Appendix C

All Earthquake Detailed Information

All earthquake information is collected via (United States Geological Survey of Earthquakes) [1] related to the IQUIQUE Station.

Earthquake Time	Location		Earthquake Properties	
	Latitude	Longitude	Depth	Magnitude
2012-02-28T07:54:11.000Z	-21.0220	-69.1010	107.5	4.5
2012-03-10T02:26:57.000Z	-19.7380	-69.2500	101	5.3
2012-07-17T02:35:35.000Z	-20.2420	-69.1420	94	4.6
2012-07-18T18:24:08.680Z	-20.6930	-70.3650	10.9	5.1
2012-07-29T14:06:01.000Z	-18.6230	-69.4901	110.8	4.7
2012-08-04T20:35:08.030Z	-20.6441	-68.9290	108.1	4.6
2012-08-28T17:32:17.160Z	-18.8310	-68.9831	122.1	4.5
2012-09-27T16:39:00.000Z	-20.0760	-69.2541	100	4.7
2012-11-13T03:11:24.870Z	-20.4370	-68.8910	83.1	4.9
2012-12-16T20:39:24.000Z	-19.7961	-69.3871	112.6	4.9
2013-01-13T21:23:26.410Z	-20.0581	-69.0491	75.3	5.3
2013-05-12T12:51:53.000Z	-20.9960	-68.6731	132.6	5.2
2013-08-05T05:40:56.060Z	-20.1808	-70.7121	19.12	5.2
2013-08-28T17:56:29.310Z	-20.7472	-68.9566	111.39	4.7
2014-03-02T20:55:14.000Z	-19.1660	-69.3281	110.5	4.5
2014-03-14T15:31:18.000Z	-19.4861	-69.4820	95.6	4.5
2014-03-17T05:11:34.860Z	-20.0168	-70.8837	21	6.4
2014-03-17T05:19:34.290Z	-19.9635	-70.8478	10	5.0
2014-03-17T05:21:36.890Z	-19.9828	-70.7943	10	5.1
2014-03-17T08:32:35.840Z	-19.9851	-70.7581	8	5.2

[1] Available at https://earthquake.usgs.gov/

Appendix B. *Heat Maps of TEC based Data Related to karr Station*

Earthquake Time	**Location**		**Earthquake Properties**	
	Latitude	Longitude	Depth	Magnitude
2014-03-18T21:26:45.400Z	-19.9262	-70.7961	3.32	5.8
2014-03-18T21:33:06.350Z	-19.9829	-70.9143	9.14	5.1
2014-03-22T12:59:59.150Z	-19.7625	-70.8740	20	6.2
2014-03-22T13:14:58.400Z	-19.7701	-70.8804	24	5.2
2014-03-22T13:29:58.410Z	-19.7163	-70.9515	17	5.6
2014-03-22T22:14:58.050Z	-19.6953	-71.0063	21	5.0
2014-03-23T18:20:01.930Z	-19.6899	-70.8538	21	6.3
2014-03-23T20:23:03.290Z	-19.8774	-70.8631	16.56	5.1
2014-03-23T22:04:26.320Z	-19.8192	-70.7906	23.26	5.1
2014-03-24T11:26:39.400Z	-19.8239	-70.7651	22	5.7
2014-03-24T11:32:15.000Z	-19.7866	-70.8069	17.81	5.2
2014-03-24T11:40:43.470Z	-19.8288	-70.8824	15.19	5.6
2014-03-24T12:32:50.410Z	-19.7882	-70.8080	17.85	5.0
2014-03-24T15:45:31.140Z	-19.5932	-70.8207	17.05	5.7
2014-03-24T18:43:12.230Z	-19.6058	-70.7419	21.87	5.0
2014-03-25T00:15:12.710Z	-19.7936	-70.8261	15.39	5.2
2014-03-31T12:53:06.000Z	-19.5110	-69.1740	114.5	5.6
2014-04-01T23:46:47.260Z	-19.6097	-70.7691	25	8.2
2014-04-01T23:57:31.930Z	-20.0509	-70.5674	25.78	5.0
2014-04-01T23:57:58.790Z	-19.8927	-70.9455	28.42	6.9
2014-04-01T23:59:57.790Z	-19.4928	-70.1661	21.61	5.8
2014-04-02T00:03:13.710Z	-19.7831	-70.7191	20.72	5.7
2014-04-02T00:04:50.930Z	-19.5481	-70.7269	10	5.6
2014-04-02T00:06:43.900Z	-19.7029	-70.9445	10	5.6
2014-04-02T00:11:20.490Z	-19.6096	-70.9319	22.04	5.1
2014-04-02T00:14:20.850Z	-19.9564	-70.9946	10	5.3
2014-04-02T00:18:51.490Z	-19.9224	-70.7432	16.76	5.1
2014-04-02T00:24:45.420Z	-19.9147	-70.6776	11.39	5.2
2014-04-02T00:33:43.900Z	-20.2757	-70.6454	4.95	5.4
2014-04-02T00:37:49.150Z	-19.9898	-70.5676	20.1	5.4
2014-04-02T01:20:58.370Z	-19.5893	-71.0127	19.01	5.2
2014-04-02T01:29:41.420Z	-20.0390	-70.9675	12.12	5.1
2014-04-02T03:40:15.280Z	-19.9322	-70.9580	10	5.4
2014-04-02T04:19:47.290Z	-19.8463	-70.9937	10	5.2
2014-04-02T04:46:18.770Z	-20.0759	-70.8325	16.53	5.5
2014-04-02T05:02:52.630Z	-19.6889	-71.0027	10	5.4
2014-04-02T06:29:16.540Z	-20.1248	-70.8304	16	5.1
2014-04-02T08:25:52.310Z	-20.2046	-70.7654	17	5.0
2014-04-02T11:07:32.130Z	-20.0107	-70.9375	15	5.4
2014-04-02T11:11:34.680Z	-19.9884	-70.9525	17	5.1
2014-04-02T19:45:51.090Z	-20.3523	-70.4501	26.89	5.3

Appendix B. *Heat Maps of TEC based Data Related to karr Station*

Earthquake Time	**Location**		**Earthquake Properties**	
	Latitude	Longitude	Depth	Magnitude
2014-04-03T00:01:16.400Z	-19.6823	-71.1759	25.5	5.0
2014-04-03T01:58:30.530Z	-20.3113	-70.5756	24.07	6.5
2014-04-03T02:43:13.110Z	-20.5709	-70.4931	22.4	7.7
2014-04-03T02:56:05.060Z	-20.7468	-70.5310	10	5.5
2014-04-03T02:59:53.620Z	-20.5309	-70.4580	35	5.0
2014-04-03T03:11:14.470Z	-20.6403	-70.7332	10	5.5
2014-04-03T03:11:31.630Z	-20.5347	-70.5344	10	5.4
2014-04-03T04:17:55.520Z	-20.6173	-70.7220	10	5.2
2014-04-03T05:19:05.520Z	-20.5027	-70.5207	24.51	5.1
2014-04-03T05:26:15.700Z	-20.7969	-70.5865	25	6.4
2014-04-03T05:34:31.580Z	-20.4540	-70.5496	19.4	5.0
2014-04-03T05:51:44.500Z	-20.7591	-70.4171	29	5.3
2014-04-03T06:54:31.010Z	-20.6266	-70.6749	15.83	5.0
2014-04-03T09:23:22.040Z	-20.5231	-70.7154	19	5.4
2014-04-04T01:37:50.570Z	-20.6426	-70.6540	13.71	6.3
2014-04-05T04:05:03.290Z	-20.7510	-70.6692	22	5.0
2014-04-05T05:44:56.410Z	-20.1486	-70.5553	24.2	5.2
2014-04-07T13:43:21.010Z	-20.1273	-70.8513	8	5.7
2014-04-07T13:47:33.330Z	-20.1294	-70.7879	9.67	5.4
2014-04-07T14:03:43.490Z	-20.1239	-70.8984	10	5.1
2014-04-08T10:14:31.580Z	-20.5124	-70.9226	6	5.6
2014-04-09T11:14:43.910Z	-20.6010	-70.8026	9.9	5.2
2014-04-11T00:01:45.210Z	-20.6590	-70.6472	13.77	6.2
2014-04-11T12:00:51.640Z	-20.0730	-70.5558	22.3	5.0
2014-04-13T12:11:30.080Z	-20.5664	-70.7470	13.37	5.5
2014-04-14T05:56:17.220Z	-20.7773	-70.7541	14.88	5.1
2014-04-15T16:09:34.540Z	-20.2042	-70.7826	17.61	5.3
2014-04-15T16:21:17.040Z	-20.1473	-70.6938	15.44	5.1
2014-04-15T18:59:40.370Z	-20.1823	-70.8507	14.95	5.1
2014-04-19T20:54:42.280Z	-20.0283	-70.9196	10	5.8
2014-04-21T13:39:05.250Z	-19.6254	-71.0423	16.81	5.3
2014-04-24T10:26:47.250Z	-19.9896	-70.9706	10.52	5.1
2014-05-05T11:21:17.410Z	-20.2139	-70.7746	13.46	5.4
2014-05-14T05:51:47.560Z	-19.7058	-71.0580	10	5.2
2014-05-17T09:11:05.390Z	-19.9874	-70.8968	5.62	5.6
2014-06-19T09:38:36.080Z	-19.9749	-70.9451	11.3	5.7
2014-06-19T19:54:04.910Z	-19.8410	-70.8701	10.6	5.8
2014-07-07T22:20:25.000Z	-18.8990	-69.2950	125.8	4.5
2014-07-12T06:55:29.000Z	-19.7030	-69.1270	111.5	4.5
2014-07-13T20:54:14.560Z	-20.2585	-70.3475	33.12	5.5
2014-07-23T21:39:08.470Z	-20.2267	-68.6776	118.73	5.6

Appendix B. *Heat Maps of TEC based Data Related to karr Station*

Earthquake Time	**Location**		**Earthquake Properties**	
	Latitude	Longitude	Depth	Magnitude
2014-08-23T04:45:32.670Z	-20.1745	-69.0385	100	5.6
2014-08-24T13:55:13.230Z	-19.9407	-71.0466	10	4.7
2014-11-22T06:50:53.640Z	-20.0679	-71.1275	14.65	5.0
2014-11-29T14:18:07.930Z	-19.9973	-71.0688	6.07	5.3
2015-03-03T12:45:19.000Z	-20.3580	-69.1450	106.2	5.2
2015-03-24T22:46:52.000Z	-20.6801	-70.7850	15.5	5.2
2015-06-09T12:38:51.640Z	-20.2057	-70.8210	13.43	4.7
2015-06-09T13:09:47.740Z	-20.2671	-70.8474	10	4.5
2015-11-16T04:55:35.510Z	-18.7515	-69.7705	78.14	4.7
2015-11-17T05:37:01.000Z	-21.1540	-68.4860	130.8	4.7
2015-11-18T06:13:15.990Z	-19.3273	-69.2998	99.76	4.7
2016-01-17T15:28:38.980Z	-21.3044	-68.7630	90.29	5.0
2016-03-10T08:34:12.440Z	-18.8045	-69.6290	94.43	5.0
2016-06-05T15:05:33.750Z	-19.4083	-69.0325	97.17	5.1
2016-08-14T13:28:42.520Z	-20.2005	-69.1698	100	5.0
2016-08-19T05:59:23.890Z	-20.5989	-68.8989	109.8	4.7
2016-12-25T14:01:33.190Z	-19.8770	-68.6624	103.32	5.1
2017-01-03T13:41:52.400Z	-19.9071	-70.4091	43.13	5.2
2017-01-28T01:27:28.720Z	-19.9610	-70.0363	51.76	4.9
2017-02-25T02:31:24.480Z	-19.4876	-69.0941	108.46	5.2
2017-04-26T18:12:32.360Z	-19.7031	-69.3496	104.68	5.4
2017-05-14T19:44:25.160Z	-20.8287	-68.7110	116	5.7
2017-05-28T18:47:11.650Z	-21.1834	-68.6823	137.99	4.9
2017-07-22T06:29:49.660Z	-21.1751	-69.2364	39	5.3
2018-04-05T11:12:10.600Z	-20.3810	-70.5855	44	5.5
2018-09-14T10:50:37.900Z	-19.2439	-69.1901	100.53	4.6
2018-09-19T18:42:21.450Z	-19.5404	-70.3888	62.58	5.0
2018-11-01T22:19:51.690Z	-19.5827	-69.2656	102	6.2
2018-11-06T08:31:01.050Z	-19.7194	-69.4141	86.57	4.7
2018-11-07T04:06:41.060Z	-21.2617	-68.9007	114.8	4.5
2019-01-07T13:43:28.900Z	-19.5722	-68.9156	103.54	4.9
2019-03-07T05:50:36.650Z	-18.7902	-69.4234	107.98	4.9
2019-03-17T00:53:37.680Z	-20.3718	-68.4290	120.75	5.1
2019-03-25T10:52:40.475Z	-21.3137	-70.0005	72.03	5.0
2019-08-28T15:53:02.042Z	-20.5823	-70.3895	29.75	4.9
2019-10-25T16:29:37.841Z	-19.0541	-69.3486	145.68	4.6
2019-10-29T02:38:16.830Z	-19.5766	-69.3058	100.18	4.6
2019-10-29T09:05:55.278Z	-18.9629	-69.3633	110.67	4.5
2019-10-29T20:32:07.829Z	-21.2686	-68.9314	120.44	4.5

REFERENCES

[1] C. B. Erol Arikan, F. and O. Arikan. Regularized estimation of vertical total electron content from global positioning system data. In *Journal of Geophysical Research: Space Physics 108*. Advancing Earth and Space Science, **2003**.

[2] F. Arikan O. Arikan Nayir, H. and C. B. Erol. Total electron content estimation with reg-est. In *Journal of Geophysical Research: Space Physics 112*. Advancing Earth and Space Science, **2007**.

[3] Henry Rishbeth and Owen K. Garriott. Introduction to ionospheric physics. In *New York: Academic Press*. **1969**.

[4] Geoffrey E. Hinton and Ruslan R. Salakhutdinov. Reducing the dimensionality of data with neural networks. In *science 313*, pages 504–507. AAS, **2006**.

[5] Hongli Dong Zidong Wang Weijian Ren Hou, Nan and Fuad E. Alsaadi. Non-fragile state estimation for discrete markovian jumping neural networks. In *Neurocomputing 179*, pages 238–245. Elsevier, **2016**.

[6] Hongli Dong Zidong Wang Weijian Ren Yang, Fan and Fuad E. Alsaadi. A new approach to non-fragile state estimation for continuous neural networks with time-delays. In *Neurocomputing 179*, pages 205–211. Elsevier, **2016**.

[7] Hongli Dong Zidong Wang Weijian Ren Yu, Yajing and Fuad E. Alsaadi. Design of non-fragile state estimators for discrete time-delayed neural networks with parameter uncertainties. In *Neurocomputing 182*, pages 18–24. Elsevier, **2016**.

[8] Yuan Yuan and Fuchun Sun. Delay-dependent stability criteria for time-varying delay neural networks in the delta domain. In *Neurocomputing 125*, pages 17–21. Elsevier, **2014**.

[9] Lifeng Ma Zhang, Jie and Yurong Liu. Passivity analysis for discrete-time neural networks with mixed time-delays and randomly occurring quantization effects. In *Neurocomputing 216*, pages 657–665. Elsevier, **2016**.

[10] Jason Liang Elliot Meyerson Aditya Rawal Daniel Fink Olivier Francon Bala Raju Miikkulainen, Risto. Evolving deep neural networks. In *In Artificial intelligence in the age of neural networks and brain computing*, pages 293–312. Academic Press, **2019**.

[11] Léon Bottou Yoshua Bengio LeCun, Yann and Patrick Haffner. Gradient-based learning applied to document recognition. In *Proceedings of the IEEE 86*, pages 2278–2324. IEEE, **1998**.

[12] S. Hochreiter and J. Schmidhuber. Long short-term memory. In *Neural computation 9(8)*, pages 1735–1780. Neural Computation, **1997**.

[13] Vincent Vanhoucke Sergey Ioffe Jon Shlens Szegedy, Christian and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *In Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826. IEEE, **2016**.

[14] Sanjay Purushotham Kyunghyun Cho David Sontag Che, Zhengping and Yan Liu. Recurrent neural networks for multivariate time series with missing values. In *Scientific reports 8*, pages 1–12. Scientific Reports, **2018**.

[15] Xiangyu Zhang Shaoqing Ren He, Kaiming and Jian Sun. Identity mappings in deep residual networks. In *In European conference on computer vision*, pages 630–645. Springer, **2016**.

[16] Matthew Hausknecht Sudheendra Vijayanarasimhan Oriol Vinyals Rajat Monga Yue-Hei Ng, Joe and George Toderici. Beyond short snippets: Deep networks for video classification. In *In European conference on computer vision*, pages 4694–4702. IEEE, **2015**.

[17] Dimitar Ouzounov Alexander Karelin Pulinets, Sergey and D. Dmitry Davidenko. Lithosphere–atmosphere–ionosphere–magnetosphere coupling—a concept for pre-earthquake signals generation. In *Pre-Earthquake Processes: A*

Multidisciplinary Approach to Earthquake Prediction Studies, pages 79–99. Advancing Earth and Space Science, **2018**.

[18] Jinbin Cao Roberto Battiston Liuyuan Li Yuduan Ma Wenlong Liu Zeren Zhima Lanwei Wang Tao, Dan and Malcolm Wray Dunlop. Seismo-ionospheric anomalies in ionospheric tec and plasma density before the 17 july 2006 m7. 7 south of java earthquake. In *Annales Geophysicae*, pages 589–598. European Geosciences Union, **2017**.

[19] Daria Chaplygina and Natalia Grafeeva. Predicting earthquakes by anomalies in the ionosphere. In *Conference on Software Engineering and Information Management (SEIM2020)*, pages 34–38. **2020**.

[20] Wan Weixing Liu Libo Biqiang, Zhao and Mao Tian. Morphology in the total electron content under geomagnetic disturbed conditions: results from global ionosphere maps. In *Annales Geophysicae*, pages 1555–1568. European Geosciences Union, **2007**.

[21] Michel Parrot Sergey Pulinets Trigunait, Ankur and Feng Li. Variations of the ionospheric electron density during the bhuj seismic event. In *Annales Geophysicae*, pages 4123–4131. European Geosciences Union, **2004**.

[22] A. Leyva Contreras G. Bisiacchi-Giraldi Pulinets, S. A. and L. Ciraolo. Total electron content variations in the ionosphere before the colima, mexico, earthquake of 21 january 2003. In *Geofísica internacional 44*, pages 369–377. **2005**.

[23] S. Kuznetsov V. Oraevsky-M. Panasyuk Sergey Pulinets A. Biryukov, O. Grigoryan and V. Chmyrev. Effects from the earthquakes on the ionospheric heights (in russian). In *Space physics and ecology*, pages 92–115. **1996**.

[24] T.V. Gaivoronskaya S.A. Pulinets, A.D. Legen'ka and V.Kh. Depuev. Main phenomenological features of ionospheric precursors of strong earthquakes. In *Journal of Atmospheric and Solar-Terrestrial Physics*, pages 1337–1347. **2003**.

[25] Yi-Ben Tsai Yuh-Ing Chen, J. Liu and Chun-Shu Chen. Statistical tests for preearthquake ionospheric anomaly. In *Terrestrial, Atmospheric and Oceanic Sciences*, pages 385–396. **2004**.

[26] Vincenzo Romano-Bruno Zolesi Stamatis.S. Kouris, Kostas Polimeris and Ljiljana Cander. Within-the-hour variability: Levels and their probabilities. In *Annals of Geophysics*, pages 945–959. **2006**.

[27] A. Lomonosov-Valery Hegai Sergey Pulinets, K. Boyarchuk and J. Liu. Ionospheric precursors to earthquakes: A preliminary analysis of the fof2 critical frequencies at chung-li ground-based station for vertical sounding of the ionosphere (taiwan island). In *Geomagnetism and Aeronomy*, 42, pages 508–513. **2002**.

[28] Sergey Pulinets. Ionospheric precursors of earthquakes; recent advances in theory and practical applications. In *Terrestrial Atmospheric And Oceanic Sciences*, 15, pages 445–467. **2004**.

[29] Dmitry Davidenko and Sergey Pulinets. Deterministic variability of the ionosphere before strong (m 6) earthquakes in the regions of greece and italy based on many years of observation (in russian). In *Geomagnetism and Aeronomy*, 59, pages 529–544. **2019**.

[30] I. P Zubkov S. I. Miachkin V. I. Dobrovolsky. Estimation of the size of earthquake preparation zones. In *Pure and Applied Geophysics*, 117, pages 1025–1044. **1979**.

[31] Y. J. Chuo S. J. Shan Y. B. Tsai Y. I. Chen S. A. Pulinets Liu, Jann Yen and S. B. Yu. Pre-earthquake ionospheric anomalies registered by continuous gps tec measurements. In *Annales Geophysicae*, pages 1585–1593. European Geosciences Union, **2004**.

[32] Jann-Yenq Liu Le, Huijun and Libo Liu. A statistical analysis of ionospheric anomalies before 736 m6. 0+ earthquakes during 2002–2010. In *Journal of*

Geophysical Research: Space Physics 116. Advancing Earth and Space Science, **2011**.

[33] Y. I. Chen Y. J. Chuo Liu, JANN-YENQ and Chun-Shu Chen. A statistical investigation of preearthquake ionospheric anomaly. In *Journal of Geophysical Research: Space Physics 111*. Advancing Earth and Space Science, **2006**.

[34] C. H. Chen Y. I. Chen W. H. Yang K. I. Oyama Liu, J. Y. and K. W. Kuo. A statistical study of ionospheric earthquake precursors monitored by using equatorial ionization anomaly of gps tec in taiwan during 2001–2007. In *Journal of Asian Earth Sciences 39*, pages 76–80. Elsevier, **2010**.

[35] Mustafa Ulukavak and Mualla Yalcinkaya. Precursor analysis of ionospheric gps-tec variations before the 2010 m 7.2 baja california earthquake. In *Geomatics, Natural Hazards and Risk 8*, pages 295–308. Taylor and Francis, **2017**.

[36] Birbal Singh O. P. Singh Saral Kumar Gupta S. P. Karia Pundhir, Devbrat and K. N. Pathak. Study of ionospheric precursors using gps and gim-tec data related to earthquakes occurred on 16 april and 24 september, 2013 in pakistan region. In *Advances in Space Research 60*, pages 1978–1987. Elsevier, **2017**.

[37] Haris Haralambous Oikonomou, Christina and Buldan Muslim. Investigation of ionospheric tec precursors related to the m7. 8 nepal and m8. 3 chile earthquakes in 2015 based on spectral and statistical analysis. In *International conference on machine learning*, pages 97–116. Springer, **2016**.

[38] Munawar Shah M. Hernández-Pajares Tariq, M. Arslan and Talat Iqbal. Pre-earthquake ionospheric anomalies before three major earthquakes by gps-tec and gim-tec data during 2015–2017. In *Advances in Space Research 63*, pages 2088–2099. Elsevier, **2019**.

[39] Munawar Shah and Shuanggen Jin. Statistical characteristics of seismo-ionospheric gps tec disturbances prior to global mw$\geq$ 5.0 earthquakes (1998–2014). In *Journal of Geodynamics 92*, pages 42–49. Elsevier, **2015**.

[40] Tao Yu Guirong Xu Chunliang Xia, Qi Wang and Shaomin Yang. Variations of ionospheric total electron content before three strong earthquakes in the qinghai-tibet region. In *Advances in Space Research*, 47, pages 506–514. **2011**.

[41] L.P. Korsunova and Valery Hegai. Possible short-term precursors of strong crustal earthquakes in japan based on data from the ground stations of vertical ionospheric sounding. In *Geomagnetism and Aeronomy*, 58, pages 90–97. **2018**.

[42] Arikan O. Akyol, A.A. and F Arikan. A machine learning-based detection of earthquake precursors using ionospheric data. In *Radio Science*, 55, pages 1–21. IEEE, **2020**.

[43] R. Bendick and R. Bilham. Do weak global stresses synchronize earthquakes? In *Geophysical Research Letters 44*, 16, pages 8320–8327. **2017**.

[44] Gkasios M. Tselikas N.T. Boucouvalas, A.C. and G Drakatos. Modified-fibonacci-dual-lucas method for earthquake prediction. In *In Third International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2015)*, volume 9535, page 95351A. **2015**.

[45] Kannan and Suganth. Innovative mathematical model for earthquake prediction. In *Engineering Failure Analysis*, volume 41, pages 89–95. **2014**.

[46] B.C. Papazachos and C.A. Papaioannou. Long-term earthquake prediction in the aegean area based on a time and magnitude predictable model. In *Pure and applied geophysics*, volume 140, pages 593–612. **1993**.

[47] Rabinowitz N. Leonard G Last, M. Predicting the maximum earthquake magnitude from seismic data in israel and its neighboring countries. In *PloS one*, volume 11, pages 593–612. **2016**.

[48] Martínez-Álvarez F. Morales-Esteban A. Asencio-Cortés, G. and J. Reyes. A sensitivity study of seismicity indicators in supervised learning to improve earthquake prediction. In *Knowledge-Based Systems*, volume 101, pages 15–30. **2016**.

[49] Idris-A. Iqbal T. Asim, K. M. and F. Martínez-Álvarez. Earthquake prediction model using support vector regressor and hybrid neural networks. In *PloS one*, volume 13. **2018**.

[50] A. Panakkat and H. Adeli. Recurrent neural network for approximate earthquake time and location prediction using multiple seismicity indicators. In *Computer-Aided Civil and Infrastructure Engineering*, volume 24, pages 280–292. **2009**.

[51] Guo Y.-Yu L. Wang, Q. and P. Li. Earthquake prediction based on spatio-temporal data mining: An lstm network approach. In *IEEE Transactions on Emerging Topics in Computing*. IEEE, **2017**.

[52] Park-J. Ibrahim, M. A. and N. Athens. Earthquake warning system: Detecting earthquake precursor signals using deep neural networks. In *Technical Report CS 230*. Stanford University, **2018**.

[53] Asencio-Cortés and Gualberto. Improving earthquake prediction with principal component analysis: application to chile. In *International Conference on Hybrid Artificial Intelligence Systems*, pages 393–404. Springer, **2015**.

[54] Mohammad Ali Arjomand Masoud Rezaei Mahmoudi, Jamal and Mohammad Hossein Mohammadi. Predicting the earthquake magnitude using the multilayer perceptron neural network with two hidden layers. In *Civil engineering journal 2*, pages 1–12. **2016**.

[55] Marios Avraamides Moustra, Maria and Chris Christodoulou. Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals. In *Expert systems with applications 38*, pages 15032–15039. Elsevier, **2011**.

[56] Arikan F. Gulyaeva, T. L. and I. Stanislawska. Persistent long-term (1944–2015) ionosphere-magnetosphere associations at the area of intense seismic activity and beyond. In *Advances in Space Research*, volume 59, pages 1033–1040. **2017**.

[57] Arikan F. Karatay, S. and O. Arikan. Investigation of total electron content variability due to seismic and geomagnetic disturbances in the ionosphere. In *Radio Science*, volume 45, pages 1–12. **2010**.

[58] Zolotov O. V. Zakharenkova-I. E. Shagimuratov I. I. Namgaladze, A. A. and O. V. Martynenko. Ionospheric total electron content variations observed before earthquakes: Possible physical mechanism and modeling. In *arXiv preprint arXiv:0905.3313*. **2009**.

[59] Devi M. Ryu K.-Chen C. H. Liu J. Y. Oyama, K.-I. and H. Liu. Modifications of the ionosphere prior to large earthquakes: Report from the ionosphere precursor study group. In *Geoscience Letters*, volume 3, pages 1–6. **2016**.

[60] Mohd Ali Hasbi, A. M. and N. Misran. Ionospheric variations before some large earthquakes over sumatra. In *Natural Hazards and Earth System Sciences*, volume 11, pages 597–611. **2011**.

[61] J. Hartmann and J. K. Levy. Hydrogeological and gasgeochemical earthquake precursors—a review for application. In *Natural Hazards*, volume 34, pages 279–304. **2005**.

[62] G. Asteriadis and E. Livieratos. Pre-seismic responses of underground water level and temperature concerning a 4.8 magnitude earthquake in greece on october 20, 1988. In *Tectonophysics*, volume 170, pages 165–169. **1989**.

[63] Bella F. Della Monica G.-Ermini A. Improta S. Sgrigna V. Allegri, L. and P. F. Biagi. Radon and rilt anomalies detected before the irpinia (south italy) earthquake of november 23, 1980 at great distances from the epicenter. In *Geophysical Research Letters*, volume 10, pages 269–272. **1983**.

[64] C. Fidani. The earthquake lights (eql) of the 6 april 2009 aquila earthquake, in central italy. In *Natural Hazards and Earth System Sciences*, volume 10, pages 967–978. **2010**.

[65] R. A. Grant and T. Halliday. Predicting the unpredictable; evidence of pre-seismic anticipatory behaviour in the common toad. In *Journal of Zoology*, volume 281, pages 263–271. **2010**.

[66] Biagi P. F. Molchanov-O. A. Khatkevich Y. M. Tronin, A. A. and E. I. Gordeev. Temperature variations related to earthquakes from simultaneous observation at the ground stations and by satellites in kamchatka area. In *Physics and Chemistry of the Earth*, volume 29, pages 501–506. **2004**.

[67] Hayakawa M. Tronin, A. A. and O. A. Molchanov. Thermal ir satellite data application for earthquake research in japan and china. In *Journal of Geodynamics*, volume 33, pages 519–534. **2002**.

[68] Pulinets S. Romanov A. Romanov-A. Tsybulya K. Ouzounov, D. and D. Davidenko. Atmosphere-ionosphere response to the m9 tohoku earthquake revealed by multi-instrument space-borne and ground observations: Preliminary results. In *Earthquake Science*, volume 24, pages 557–564. **2011**.

[69] S. A. Pulinets and D. V. Davidenko. The nocturnal positive ionospheric anomaly of electron density as a short-term earthquake precursor and the possible physical mechanism of its formation. In *Geomagnetism and Aeronomy*, volume 58, pages 559–570. **2018**.

[70] Namgaladze A. A. Zolotov, O. V. and B. E. Prokhorov. Tpecific features of ionospheric total electron content variations in the periods of preparation of the earthquakes on march 11, 2011 (japan) and october 23, 2011 (turkey). In *Russian Journal of Physical Chemistry*, volume 7, pages 599–605. **2013**.

[71] Karia S. P. Yadav, K. S. and K. N. Pathak. Anomalous variation in gps based tec prior to the 5 earthquakes in 2009 and 2010. In *Positioning*, volume 6. **2015**.

[72] M. Akhoondzadeh. Decision tree, bagging and random forest methods detect tec seismo-ionospheric anomalies around the time of the chile, (mw = 8.8) earthquake of 27 february 2010. In *Advances in Space Research*, volume 57, pages 2464–2469. **2016**.

[73] D. V. Davidenko and S. A. Pulinets. Deterministic variability of the ionosphere on the eve of strong (m 6) earthquakes in the regions of greece and italy according to long-term measurements data. In *Geomagnetism and Aeronomy*, volume 59, pages 493–508. **2019**.

[74] Polimeris K. V. Kouris, S. S. and L. R. Cander. Specifications of tec variability. In *Advances in Space Research*, volume 37, pages 983–1004. **2006**.

[75] V. V. Plotkin. Gps detection of ionospheric perturbation before the 13 february 2001, el salvador earthquake. In *Natural Hazards and Earth System Science*, volume 3, pages 249–253. **1999**.

[76] Parrot M. Pulinets S. Trigunait, A. and F. Li. Variations of the ionospheric electron density during the bhuj seismic event. In *Annales Geophysicae*, volume 22, pages 4123–4131. **2004**.

[77] S. S. Kouris and D. N. Fotiadis. Ionospheric variability: A comparative statistical study. In *Advances in Space Research*, volume 29, pages 977–985. **2002**.

[78] Chen Y. I. Pulinets-S. A. Tsai Y. B. Liu, J. Y. and Y. J. Chuo. Seismo-ionospheric signatures prior to m 6.0 taiwan earthquakes. In *Geophysical Research Letters*, volume 27, pages 3113–3116. **2000**.

[79] Medhi A. Sarma A. J. D. S. Devi, M. and A. K. Barbara. Growth and inhibition of equatorial anomaly prior to an earthquake (eq): Case studies with total electron content (tec) data for major eqs of japan 2011 and indonesia 2012. In *Positioning*. **2013**.

[80] Guo J. Yue J. Yang Y. Li Z. Li, W. and D. Lu. Contrastive research of ionospheric precursor anomalies between calbuco volcanic eruption on april 23 and nepal earthquake on april 25, 2015. In *Advances in Space Research*, volume 57, pages 2141–2153. **2016**.

[81] Shi J. K. Wang X. Zhang, M. L. and S. M. Radicella. Ionospheric variability at low latitude station: Hainan, china. In *Advances in Space Research*, volume 34, pages 1860–1868. **2004**.

[82] Arikan O. Arikan F. Akyol, A. A. and M. N. Deviren. Investigation on the reliability of earthquake prediction based on ionospheric electron content variation. In *In Proceedings of the 16th International Conference on Information Fusion*, pages 1658–1663. IEEE, **2013**.

[83] Feza Arikan Arikan, Orhan and Ali Alp Akyol. Machine learning based detection of earthquake precursors using ionospheric data. In *42nd COSPAR Scientific Assembly*, volume 42, pages 1–4. **2018**.

[84] The Houw Liong Aji, Bernadus Anggo Seno and Buldan Muslim. Detection precursor of sumatra earthquake based on ionospheric total electron content anomalies using n-model articial neural network. In *International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, volume 42, pages 269–276. IEEE, **2017**.

[85] Buldan. Muslim. Pengujian teknik korelasi untuk deteksi pengaruh aktivitas gempa bumi besar pada ionosfer [examination of correlation technique for detecting the influence of great earthquake activities on ionosphere]. In *Jurnal Sains Dirgantara*, volume 12, pages 269–276. **2015**.

[86] M. Akhoondzadeh. Genetic algorithm for tec seismo-ionospheric anomalies detection around the time of the solomon (mw= 8.0) earthquake of 06 february 2013. In *Advances in Space Research*, volume 52, pages 581–590. **2013**.

[87] M. Akhoondzadeh. Investigation of gps-tec measurements using ann method indicating seismo-ionospheric anomalies around the time of the chile (mw= 8.2) earthquake of 01 april 2014. In *Advances in Space Research*, volume 54, pages 1768–1772. **2014**.

[88] Jyh-Woei. Lin. Taiwan'chi-chi earthquake precursor detection using nonlinear principal component analysis to multi-channel total electron content records. In *Journal of Earth Science*, volume 24, pages 244–253. Springer, **2013**.

[89] The Houw Liong Pattisahusiwa, Asis and Acep Purqon. A method for separating seismo-ionospheric tec outliers from heliogeomagnetic disturbances by using nu-svr. In *AIP Conference Proceedings*, volume 1677. **2015**.

[90] Chris Stewart Lorenzo Trenchi Rune Floberghagen Nowakowski, Artur and Iarla Kilbane-Dawe. A machine learning approach to investigate possible signals in the ionosphere related to earthquake activity. In *Geophysical Research Abstracts*, volume 21. **2019**.

[91] Jeffrey Liu Reid, Jessica and Bhavani Ananthabhotla. Quakecast, an earthquake forecasting system using ionospheric anomalies and machine learning. In *AGU Fall Meeting Abstracts*, volume 2020, pages 007–0016. **2020**.

[92] Cheng Long Huiyu Zhou Roberto Battiston Xuemin Zhang Xiong, Pan and Xuhui Shen. Identification of electromagnetic pre-earthquake perturbations from the demeter data by machine learning. In *Remote Sensing*, volume 12, page 3643. **2020**.

[93] Berthelier J.J. Lebreton J.P. Sauvaud J.A. Santolik O. Parrot, M. and J. Blecki. Examples of unusual ionospheric observations made by the demeter satellite over seismic regions. In *Phys. Chem. Earth Parts A/B/C*, 31, pages 486–495. **2006**.

[94] Shen X. Liu J. Ouyang X. Qian J. Zhang, X. and S. Zhao. Analysis of ionospheric plasma perturbations before wenchuan earthquake. In *Hazards Earth Syst*, 9, pages 1259–1266. **2009**.

[95] Parrot M. Błeçki, J. and R. Wronowski. Studies of the electromagnetic field variations in elf frequency range registered by demeter over the sichuan region prior to the 12 may 2008 earthquake. In *International Journal of Remote Sense*, 31, pages 3615–3629. **2010**.

[96] S. Sarkar and A.K. Gwal. Satellite monitoring of anomalous effects in the ionosphere related to the great wenchuan earthquake of may 12, 2008. In *Natural Hazards*, 55, pages 321–332. Springer, **2010**.

[97] Parrot M. Kim S.G. Jeong K.S. Chae J.S. Pulinets S. Ryu, K. and K.I. Oyama. Suspected seismo-ionospheric coupling observed by satellite measurements and gps tec related to the m7.9 wenchuan earthquake of 12 may 2008. In *Journal of Geophysical Research: Space Physics*, 119, pages 10305–10323. **2014**.

[98] Chen Y.I. Huang C.C. Parrot M. Shen X.H. Pulinets S.A. Yang Q.S. Liu, J.Y. and Y.Y. A Ho. A spatial analysis on seismo-ionospheric anomalies observed by demeter during the 2008 m8.0 wenchuan earthquake. In *Journal of Asian Earth Sciences*, 114, pages 414–419. **2015**.

[99] Kadirkamanathan V. Walker, S.N. and O.A. Pokhotelov. Changes in the ultra-low frequency wave field during the precursor phase to the sichuan earthquake: Demeter observations. In *Annals of Geophysics*, 31, pages 1597–1603. **2013**.

[100] Michel Parrot Píša, D. and O. Santolík. Ionospheric density variations recorded before the 2010 mw 8.8 earthquake in chile. In *Journal of Geophysical Research: Space Physics*, volume 116, pages 1597–1603. **2011**.

[101] J. D. Qian X. Y. Ouyang X. H. Shen J. A. Cai Zhang, X. M. and S. F. Zhao. Ionospheric electromagnetic disturbances observed on demeter satellite before

an earthquake of m 7.9 in chili. In *Progress in Geophysics*, volume 24, pages 1196–1203. **2009**.

[102] J-Y. Liu Michel Parrot Ho, Y-Y. and J-L. Pinçon. Temporal and spatial analyses on seismo-electric anomalies associated with the 27 february 2010 m= 8.8 chile earthquake observed by demeter satellite. In *Natural Hazards and Earth System Sciences*, volume 13, pages 3281–3289. **2013**.

[103] Hau-Kun Jhuang Yung-Chih Su Ho, Yi-Ying and Jann-Yenq Liu. Seismo-ionospheric anomalies in total electron content of the gim and electron density of demeter before the 27 february 2010 m8. 8 chile earthquake. In *Advances in Space Research*, volume 51, pages 2309–2315. **2013**.

[104] Said Gaci Louerguioui, Soraya and Naima Zaourar. Irregularities of the ionospheric plasma and the ulf electric components obtained from demeter satellite experiments above chile earthquake (27 february 2010). In *Arabian Journal of Geosciences*, volume 8, pages 2433–2441. **2015**.

[105] U. A. Mofiz and R. Battiston. Possible ion-acoustic soliton formation in the ionospheric perturbations observed on demeter before the 2007 pu'er earthquake. In *Earthquake Sciences*, volume 22, pages 257–262. **2009**.

[106] Dongmei Yang-Jiadong Qian He, Yufei and Michel Parrot. Anomaly of the ionospheric electron density close to earthquakes: Case studies of pu'er and wenchuan earthquakes. In *Earthquake Sciences*, volume 24, pages 549–555. **2011**.

[107] X. Zhang J. Liu-S. F. Zhao Shen, X. H. and G. P. Yuan. Analysis of the enhanced negative correlation between electron density and electron temperature related to earthquakes. In *Annales Geophysicae*, volume 33, pages 471–479. **2015**.

[108] Pietro Ubertini Piero Diego Igor Bertello Piersanti, Mirko and Maurizio Candidi. Electromagnetic field observations by the demeter satellite in connection

with the 2009 l'aquila earthquake. In *COSPAR Scientific Assembly*, volume 42, pages S–1. **2018**.

[109] Georgios C. Anagnostopoulos A. C. Iliopoulos G. P. Pavlos Athanasiou, M. A. and C. N. David. Enhanced ulf radiation observed by demeter two months around the strong 2010 haiti earthquake. In *Natural Hazards and Earth System Sciences*, volume 11, pages 1091–1098. **2011**.

[110] R. Nakamura M. Suzuki M. Hayakawa Hobara, Y. and Michel Parrot. Ionospheric perturbations observed by the low altitude satellite demeter and possible relation with seismicity. In *Journal of Atmospheric Electricity*, volume 33, pages 21–29. **2013**.

[111] E. Lee JangSoo Chae Michel Parrot Ryu, K. and S. Pulinets. Seismo-ionospheric coupling appearing as equatorial electron density enhancements observed via demeter electron density measurements. In *Journal of Geophysical Research: Space Physics*, volume 119, pages 8524–8542. **2014**.

[112] Dongmei Yang Rong Zhu Jiadong Qian He, Yufei and Michel Parrot. Variations of electron density and temperature in ionosphere based on the demeter isl data. In *Earthquake Science*, volume 23, pages 349–355. **2010**.

[113] D. Yang J. Qian He, Y. and Michel Parrot. Response of the ionospheric electron density to different types of seismic events. In *Natural Hazards and Earth System Sciences*, volume 11, pages 2173–2180. **2011**.

[114] Michel Parrot Yan, Rui and Jean-Louis Pinçon. Statistical study on variations of the ionospheric ion density observed by demeter and related to seismic activities. In *Journal of Geophysical Research: Space Physics*, volume 122, pages 12–421. **2017**.

[115] Michel. Parrot. Statistical analysis of the ion density measured by the satellite demeter in relation with the seismic activity. In *Earthquake Science*, volume 24, pages 513–521. **2011**.

[116] M. Li and Michel Parrot. Real time analysis of the ion density measured by the satellite demeter in relation with the seismic activity. In *Natural hazards and earth system sciences*, volume 12, pages 2957–2963. **2012**.

[117] Michel. Parrot. Statistical analysis of automatically detected ion density variations recorded by demeter and their relation to seismic activity. In *Annals of geophysics*, volume 55. **2012**.

[118] Mei Li and Michel Parrot. Statistical analysis of an ionospheric parameter as a base for earthquake prediction. In *Journal of Geophysical Research: Space Physics*, volume 118, pages 3731–3739. **2013**.

[119] Michel Parrot and Mei Li. Statistical analysis of the ionospheric density recorded by the demeter satellite during seismic activity. In *Pre-Earthquake Process*. **2018**.

[120] M. Akhoondzadeh. Anomalous tec variations associated with the powerful tohoku earthquake of 11 march 2011. In *Natural Hazards and Earth System Sciences*, volume 12, pages 1453–1462. **2012**.

[121] M. Akhoondzadeh. Genetic algorithm for tec seismo-ionospheric anomalies detection around the time of the solomon (mw= 8.0) earthquake of 06 february 2013. In *Advances in Space Research*, volume 52, pages 581–590. Elsevier, **2013**.

[122] Paul A. Johnson V. Maarten Bergen, Karianne J. and Gregory C. Beroza. Machine learning for data-driven discovery in solid earth geoscience. In *Science 363*, 6433. **2019**.

[123] Claudia Hulbert Rouet-Leduc, Bertrand and Paul A. Johnson. Continuous chatter of the cascadia subduction zone revealed by machine learning. In *Nature Geoscience*, volume 12, pages 75–79. **2019**.

[124] Bertrand Rouet-Leduc Paul A. Johnson Christopher X. Ren Jacques Rivière David C. Bolton Hulbert, Claudia and Chris Marone. Similarity of fast and slow earthquakes illuminated by machine learning. In *Nature Geoscience*, volume 12, pages 69–74. **2019**.

[125] Michaël Gharbi Perol, Thibaut and Marine Denolle. Convolutional neural network for earthquake detection and location. In *Science Advances*, volume 4, page e1700578. **2018**.

[126] Ossian O'Reilly Karianne J. Bergen Yoon, Clara E. and Gregory C. Beroza. Earthquake detection through computationally efficient similarity search. In *Science Advances*, volume 1, page e1501057. **2015**.

[127] Tehmina Khalil Khalid, Samina and Shamila Nasreen. A survey of feature selection and feature extraction techniques in machine learning. In *Science and Information Conference IEEE*, pages 372–378. IEEE, **2014**.

[128] Paul Harris Chris Brunsdon A. Stewart Fotheringham Demšar, Urška and Sean McLoone. Principal component analysis on spatial data: an overview. In *Annals of the Association of American Geographers 103*, pages 106–128. Taylor and Francis, **2013**.

[129] Alok Sharma and Kuldip K. Paliwal. Linear discriminant analysis for the small sample size problem: an overview. In *International Journal of Machine Learning and Cybernetics 6*, pages 443–454. Springer, **2015**.

[130] Suvrit Sra Alex Smola Zoubin Ghahramani Lopez-Paz, David and Bernhard Schölkopf. Randomized nonlinear component analysis. In *International conference on machine learning*, pages 1359–1367. MLResearchPress, **2014**.

[131] Zainab Assaghir Yehia Taher Rafiqul Haque Mohand-Said Hacid Makki, Sara and Hassan Zeineddine. An experimental study with imbalanced classification approaches for credit card fraud detection. In *IEEE Access 7*, pages 93010–93022. IEEE, **2019**.

[132] Qiong Chen Lin, Enlu and Xiaoming Qi. Deep reinforcement learning for imbalanced classification. In *Applied Intelligence*, pages 1–15. Springer, **2020**.

[133] Heung Ho Choi Chang Soo Moon Kim, Kang Soo and Chi Woong Mun. Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. In *Current applied physics 11*, pages 740–745. Elsevier, **2011**.

[134] Tarek Gaber Abdelhameed Ibrahim Tharwat, Alaa and Aboul Ella Hassanien. Linear discriminant analysis: A detailed tutorial. In *AI communications 30*, pages 169–190. **2017**.

[135] M. Schuster and K. K. Paliwalo. Bidirectional recurrent neural networks. In *IEEE Transactions on Signal Processing*, pages 2673–2681. IEEE, **1997**.

[136] Caglar Gulcehre Kyunghyun Cho Pascanu, Razvan and Yoshua Bengio. How to construct deep recurrent neural networks. In *arXiv preprint arXiv:1312.6026*. **2013**.