DGK Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C | Dissertationen | Heft Nr. 924

Ganesh Lalgudi Gopalakrishnan

Ionosphere Electron Density Modelling Using a Constrained Optimization Approach

München 2024

Bayerische Akademie der Wissenschaften

ISSN 0065-5325 ISBN 978-3-7696-5336-6

Diese Arbeit ist gleichzeitig veröffentlicht in:
OPAC – Online Publikationen der Universitätsbibliothek Technische Universität München
https://mediatum.ub.tum.de/?id=1688144, München 2023

DGK Veröffentlichungen der DGK

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 924

Ionosphere Electron Density Modelling Using a Constrained Optimization Approach

Vollständiger Abdruck der von der TUM School of Engineering and Design
der Technischen Universität München
zur Erlangung eines
Doktors der Ingenieurwissenschaften (Dr.-Ing.)
genehmigten Dissertation

von

Ganesh Lalgudi Gopalakrishnan

München 2024

Bayerische Akademie der Wissenschaften

ISSN 0065-5325 ISBN 978-3-7696-5336-6

Diese Arbeit ist gleichzeitig veröffentlicht in:
OPAC – Online Publikationen der Universitätsbibliothek Technische Universität München
https://mediatum.ub.tum.de/?id=1688144, München 2023

Adresse der DGK:

Ausschuss Geodäsie der Bayerischen Akademie der Wissenschaften (DGK)

Alfons-Goppel-Straße 11 ● D – 80 539 München

Telefon +49 - 331 - 288 1685 ● E-Mail post@dgk.badw.de

http://www.dgk.badw.de

Prüfungskommission:

Vorsitzender: Prof. Dr.-Ing. Christoph Holst

Prüfer: apl. Prof. Dr.-Ing. Michael Schmidt
Prof. Dr.phil.nat. Urs Hugentobler
Prof. Dr. Manuel Hernández-Pajares (Universitat Politecnica de Catalunya, Barcelona, Spanien)

Tag der mündlichen Prüfung: 19.12.2022

ISSN 0065-5325 ISBN 978-3-7696-5336-6

Acknowledgements

To begin with, I would like to express my sincere gratitude to my supervisor Michael Schmidt for his supervision and invaluable suggestions on my dissertation. Without the discussions with him, this work would not have been finished. Special thanks are devoted to Florian Seitz for giving me a chance to work at Deutsches Geodätisches Forschungsinstitut der Technischen Universität München (DGFI-TUM). I am also very thankful to Urs Hugentobler and Manuel Hernández-Pajares for their feedback on this work. I would also like to thank all my colleagues at DGFI-TUM especially, the atmosphere modelling group. I will always remember the wonderful time spent during our technical discussions as well as the social events at DGFI-TUM.

Finally, I would like to thank my friends, my entire family in particular my son, my wife, my parents, my brother and his family for unconditional love as well as their continuous encouragement and support all through my life.

Abstract

The ionosphere is the region of the atmosphere extending from about 50 km above the Earth's surface gradually transitioning into the plasmasphere containing electrons and ions formed from the ionization process due to the solar wind, Extreme Ultraviolet (EUV) and X-ray radiation. Electron density is the most important ionospheric parameter in geodesy and its space and time dependent, i.e., four-dimensional (4D) magnitude needs to be accurately known, e.g. for positioning, and navigation applications. The ionosphere affects electromagnetic wave propagation and causes signal delays due to its dispersive behavior at specific frequencies. The objective of this thesis is the development of a 4D electron density model. The combined electron density of the individual D, E, F1, F2 layers and the plasmasphere is described using the multi-Chapman layer model which is parameterized by a set of so-called key parameters, namely, the peak density, the peak height and the scale height for the first four layers as well as the basis density and the scale height of the plasmasphere.

The spatial variation of the key parameters of this multi-layer representation is described by series expansions in tensor products of localizing B-spline basis functions. We use polynomial B-splines along latitude and trigonometric B-splines along longitude. The Chapman profile function is chosen to define the electron density along the altitude. The sets of coefficients of these basis functions are the unknown parameters to be estimated. This way, the electron density modeling (EDM) is setup as a parameter estimation problem. The global navigation satellite systems (GNSS) provide accurate information on the electron density distribution within the ionosphere in the form of Slant Total Electron Content (STEC), which is the integral of the electron density along the signal path between transmitter and receiver. The space-based GNSS radio occultation (RO) measurements allow the calculation of electron density profiles (EDPs) with high vertical resolution and global coverage. Ionospheric radio occultation (IRO) from the Formosat-3/COSMIC mission and the GRACE satellites as well as GNSS data are used as observations in this work. Additionally, electron density observations are generated from the vertical total electron content (VTEC) maps and vertical profile function from the IRI model.

The prior knowledge of the Chapman key parameters is expressed as constraints e.g., a non-negativity bound for peak density, minimum and maximum bounds for peak height. Due to the limited ability of the classical methods such as least squares to include inequality constraints on the unknown parameters, a novel constraint optimization approach is used in this work. Subsequently our global 4D EDM problem is transformed into an inequality constrained optimization problem (ICOP).

The Gauss-Markov model with B-spline coefficients as unknowns is combined with the equality and inequality constraints, to give the Lagrangian function, on which the optimality conditions are then applied to get the estimation model equations, also referred to as the Karush-Kuhn-Tucker (KKT) equations. This method involves transforming the inequality constraints to equalities using the Lagrange multiplier and the slack variables as additional unknown parameters. Together with B-spline coefficients they form the three set of unknown parameters. The KKT equations are then iteratively solved for the unknown parameters using an optimization algorithm called the interior point method. The modeled key parameters are validated using independent data from selected ionosonde stations, IRI model data (in the form of closed loop validation) as well as in-situ observations from Langmuir probe on board the SWARM satellites.

Zusammenfassung

Die Ionosphäre ist der Bereich der Atmosphäre ab ca. 50 km über der Erdoberfläche und geht allmählich in die Plasmasphäre über. Sie enthält Elektronen und Ionen, welche durch den Ionisierungsprozess infolge des Sonnenwindes, extremer Ultraviolett- (EUV) und Röntgenstrahlung entstehen. Die Elektronendichte ist der wichtigste Ionosphärenparameter für die Geodäsie. Ihre raum- und zeitabhängige, das heißt vierdimensionale (4D), Größe muss genau bekannt sein, beispielsweise für Positionsbestimmungs- und Navigationsanwendungen. Die Ionosphäre beeinflusst die Ausbreitung elektromagnetischer Wellen und verursacht aufgrund ihrer dispersiven Eigenschaften bei bestimmten Frequenzen Signalverzögerungen.

Das Ziel dieser Dissertation ist die Entwicklung eines 4D-Elektronendichtemodells. Die kombinierte Elektronendichte der einzelnen Layer D, E, F1 und F2 und der Plasmasphäre wird durch das Multi-Chapman-Layer-Modell beschrieben, welches durch einen Satz sogenannter Schlüsselparameter parametrisiert wird, nämlich die Maximaldichte, die Höhe der Maximaldichte und die Skalenhöhe der ersten vier Layer sowie die Basisdichte und die Skalenhöhe der Plasmasphäre. Die räumliche Variation der Schlüsselparameter dieser Multi-Layer-Repräsentation wird durch Reihenentwicklungen in Tensorprodukten lokalisierender B-Spline-Basisfunktionen beschrieben. Wir nutzen polynomiale B-Splines entlang der Breite und trigonometrische B-Splines entlang der Länge. Die Chapman-Profilfunktion wird zur Definition der Elektronendichte entlang der Höhe verwendet. Die Koeffizientensätze dieser Basisfunktionen sind die zu schätzenden Unbekannten. Die Elektronendichtemodellierung (EDM) wird somit als ein Parameterschätzungsproblem aufgesetzt. Die Globalen Navigationssatellitensysteme (GNSS) liefern genaue Informationen über die Elektronendichteverteilung in der Ionosphäre in der Form des Slant Total Electron Content (STEC), der das Integral der Elektronendichte entlang des Signalweges zwischen Sender und Empfänger ist. Die weltraumgestützten GNSS-Radio-Okkultationsmessungen (RO) ermöglichen die Berechnung von Elektronendichteprofilen (EDP) mit hoher vertikaler Auflösung und globaler Abdeckung. Ionosphärische Radio-Okkultationen (IRO) von der Formosat-3/COSMIC-Mission und den GRACE-Satelliten sowie GNSS-Daten werden in dieser Arbeit als Beobachtungen verwendet. Zusätzlich werden Beobachtungen der Elektronendichte aus Karten des Vertical Total Electron Content (VTEC) und der vertikalen Profilfunktion des IRI-Modells generiert.

Vorwissen über die Chapman-Schlüsselparameter wird in Form von Bedingungsgleichungen eingeführt, so beispielsweise eine Nicht-Negativitätsbedingung für die Maximaldichte und untere und obere Grenzwerte für die Höhe der Maximaldichte. Da klassische Verfahren wie der Ausgleichung nach der Methode den kleinsten Quadrate nur begrenzt in der Lage sind, Ungleichheitsbedingungen für die Unbekannten einzubeziehen, wird in dieser Arbeit ein neuartiger Optimierungsansatz für eine bedingte Ausgleichung verwendet. Hierfür wird das globale 4D-EDM-Problem in ein Optimierungsproblem mit Ungleichheitsbedingunen (Inequality Constrained Optimization Problem, ICOP) umgewandelt. Das Gauß-Markoff-Modell mit B-Spline-Koeffizienten als Unbekannte wird mit den Gleichheits- und Ungleichheitsbedingungen kombiniert, um die Lagrange-Funktion zu erhalten, auf die dann die Optimierungsbedingungen angewendet werden, um die Gleichungen für das Schätzmodell zu erhalten, die auch als Karush-Kuhn-Tucker-Gleichungen (KKT) bezeichnet werden. Diese Methode beinhaltet eine Transformation der Ungleichheitsbedingungen in Gleichheitsbedingungen unter Nutzung des Lagrange-Multiplikators und von Schlupfvariablen als zusätzliche Unbekannte. Sie bilden zusammen mit den B-Spline-Koeffizienten die drei unbekannten Parametersätze. Die KKT-Gleichungen werden dann iterativ nach den Unbekannten gelöst, wobei das Innere-Punkte-Verfahren als Optimierungsalgorithmus verwendet wird. Die modellierten Schlüsselparameter werden gegen

unabhängige Messungen ausgewählter Ionosondenstationen, IRI-Modelldaten (in Form einer Closed-Loop-Validierung) sowie In-situ-Beobachtungen der Langmuir-Sonde an Bord der SWARM-Satelliten validiert.

Contents

1 Introduction

The Sun's radiation has important impact on different components of the Earth system, for instance on the sub-components thermosphere, ionosphere and plasmasphere consisting of neutral and charged particles as well as plasma with different chemical compositions[1] and physical properties such as mass, temperature etc. The precise knowledge of the particle number density within these regions is important to many applications, e.g. positioning, navigation, satellite orbit determination, collision avoidance manoeuvres, trajectory forecasting of space debris and orbital mission lifetime predictions. Of particular interest in this work, is the ionosphere and plasmasphere where the neutral gas molecules dissociate into free electrons and ions.

> The number density of electrons is the most important key parameter to describe the state of the ionospheric plasma varying with latitude, longitude, altitude and time for geodetic applications.

Particularly, the four dimensional (4D) space and time dependent electron density needs to be accurately known for satellite navigation and telecommunication, since electromagnetic waves are strongly affected on their path through the plasmasphere and the ionosphere. The 4D electron density is a highly variable function: amongst others it reflects fluctuations with periods less than a few minutes, diurnal and seasonal variations, long-period changes following the solar cycle of 11 years as well as the impact of space weather events such as solar flares and coronal mass ejections (CME). If the 4D electron density would be known everywhere within the plasmasphere and the ionosphere at any time moment, we could correct every measurement of space-geodetic observation techniques such as the Global Navigation Satellite Systems (GNSS), in particular GPS, GLONASS and Galileo, Satellite Altimetry (SA), the French satellite tracking system Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS) or Very Long Baseline Interferometry (VLBI) for the plasmaspheric and ionospheric impact. Vice versa all these measurement techniques including the space-based GNSS technique of measuring Ionospheric Radio Occultations (IRO) provide valuable information on the state of the plasmasphere and the ionosphere and thus, for modeling the electron density.

1.1 Motivation and objectives of the thesis

The prediction of orbital position as well as velocity accuracy of satellites or space debris along a Low Earth Orbit (LEO) depends on the accuracy of the determination of the atmospheric drag, which in-turn depends on the neutral density; see Hastings (1995), Montenbruck and Gill (2001), Leonard et al. (2012), Vielberg et al. (2018). Due to a constant exchange of energy between the neutral and the charged particles during the ionization and recombination processes, a dynamic coupling exists between the thermosphere and the ionosphere[2]; see de Wit and Bruinsma (2017), Fuller-Rowell et al. (2018). The understanding of the coupling processes is critical to scientific and

[1] Nitrogen, Oxygen and other atmospheric gases have different concentration at various altitudes and owing to their respective ionization potential, they play a key role in the total electron density of each individual layer during the nominal and extreme space weather conditions; see e.g. Hanslmeier (2002).

[2] More generally there is a coupling between the lithosphere, the magnetosphere, the thermosphere and the ionosphere but for the sake of simplicity and relevance to this work, the coupling related aspects will not be discussed and instead referred to e.g. Liperovsky et al. (2008), Luehr et al. (2011)

commercial applications, e.g. accurate computation of the drag forces acting on LEO satellites and the space debris so that their trajectories are precisely determined (Leonard et al. (2012), Doornbos (2012b), Lühr et al. (2012)). Therefore, a high-accuracy and high-resolution global 4D electron density model is required so that it can be assimilated with e.g. a physical thermosphere model to yield improved estimates of neutral density.

Hence, the main objective of this thesis is to develop a 4D global electron density model (EDM) on the basis of multi-layer approach.

This problem is approached by considering the total electron density as the sum of the electron densities of the individual ionosphere layers[3] and using a novel constraint optimization algorithm for parameter estimation. The following sub-objectives are identified to accomplish the above primary objective:

- express an EDM as an inequality constrained optimization problem (ICOP) of the form according to Eq. (4.25)
- develop the inequality constrained optimization algorithm (ICOA) (The entire Chapter 4 is dedicated to the description of our developed ICOA)
- apply the ICOA from Chapter 4 to our global 4D electron density modelling (EDM) and perform an independent validation; see Chapter 5 and Chapter 7
- perform experiments on the developed EDM by varying the inequality constraint bounds; see Chapter 6 and Appendix C.1

1.2 Organization and structure of the thesis

For the scope of this work, unless otherwise stated, the following nomenclature and conventions apply:

- The phrase "*IRI model data*" refers to the Chapman key parameters obtained from the International Reference Ionosphere (IRI) model.
- The phrase "*performance of optimization algorithm*" refers to the resulting accuracy, a quantified measure of which will be shown in Chapters 6 and 7; see Eqs. (7.2), (7.3), (A.9) and (A.10).
- The phrase "*adverse space weather*" refers to either solar EUV[4] condition characterized by F10.7 $\geq$ 100 SFU or high geomagnetic[5] activity with Kp $\geq$ 5. On contrary, nominal ionosphere condition refers to F10.7 $<$ 100 SFU and Kp $<$ 5. It shall be noted that these thresholds are only meant subjectively for the scope of this work. In general there can be other definitions of extreme conditions as well.

As shown in Fig. 1.1, this thesis is a compilation of 8 chapters with each one focussing on a certain key aspect pertaining to the objectives mentioned above. The Chapter 2 introduces the ionosphere modelling and the role of the Chapman key parameters. It also provides the state of art in EDM and highlights the main novelties in this work. The observation techniques relevant for this work are introduced in Chapter 3. Chapter 4 is a major contribution to this work and consists of the theoritical and practical fundamentals of the ICOA and its application to EDM in Chapter 5 and its validation in Chapter 7. For validation we use simulated as well as ionosonde observations.

[3] The individual layers are described in detail in Chapter 2.

[4] Solar extreme ultra-violet (EUV) is a quantified measure of the input energy mainly absorbed by the Earth's atmosphere; see e.g. Schunk (1988), Hapgood (2017).

[5] Kp value of 5 to 9 denote NOAA Space Weather Scale Geomagnetic Storm Levels from G1 to G5 respectively. Source: https://www.swpc.noaa.gov/sites/default/files/images/u2/TheK-index.pdf accessed on 17 September 2022.

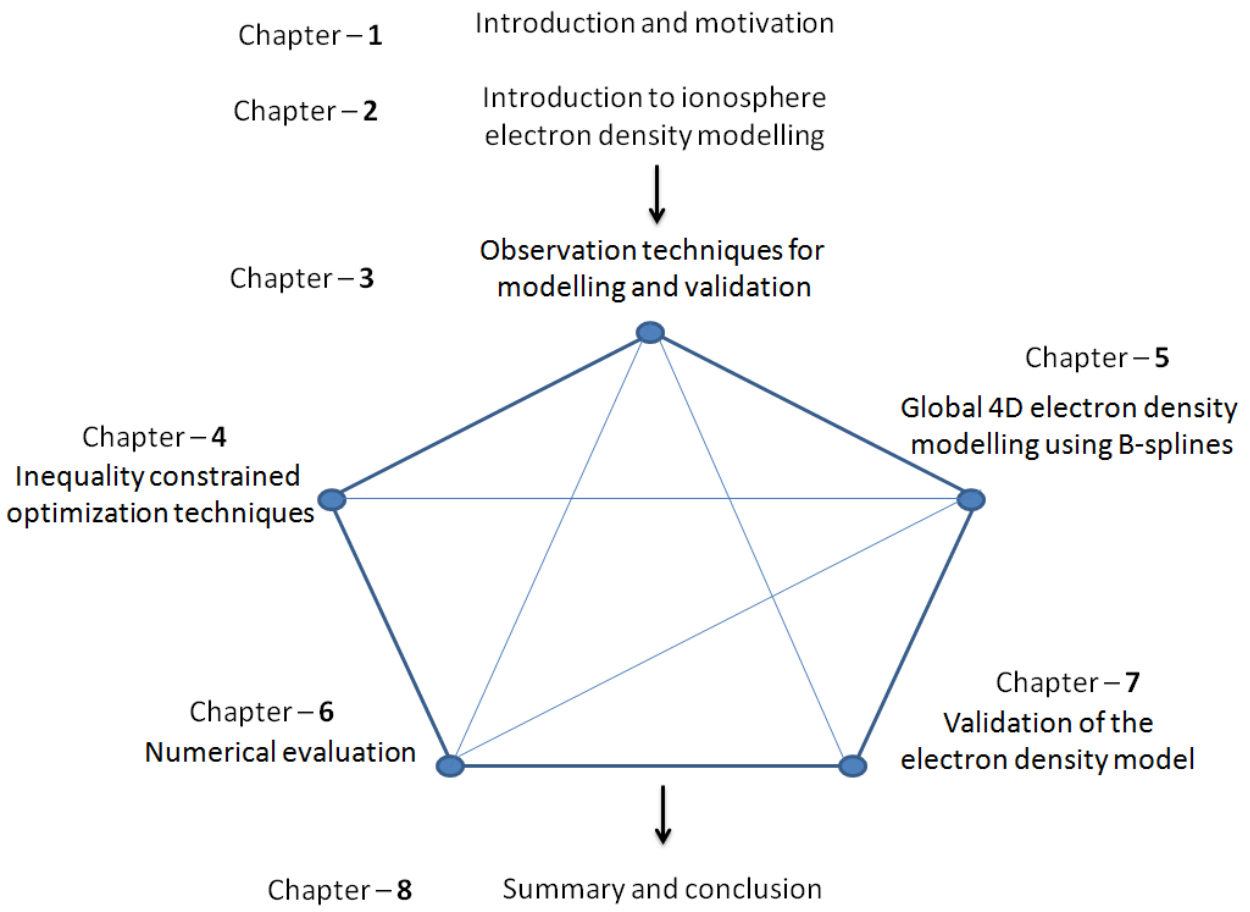


Figure 1.1: Outline of the thesis and the relation of the chapters to one another.

For the specific case of using IRI model to obtain the simulated data for validation, we show the performance of the ICOA using a closed loop validation (CLV). This aspect is substantiated in Chapters 6 and 7.

1.3 Acknowledgement to the Funding agency

This work was done in the scope of the project titled "Development of an operational prototype for the determination of the thermospheric density on the basis of a thermosphere-ionosphere coupling model" with the acronym TIK (original German title "Entwicklung eines operationellen Prototyps zur Bestimmung der thermosphärischen Dichte auf Basis eines Thermosphären-Ionosphären Kopplungsmodells"). The author would like to acknowledge the German Federal Ministry for Economic Affairs and Energy (BMWi) for sponsoring this project via the German Aerospace Center (DLR Space administration) under the research number: FKZ50LZ1605.

2 Introduction to ionosphere

G. Marconi[1] performed the first trans-atlantic radio transmission through the electrically conductive layer in the atmosphere to be later termed as the ionosphere. A few years later E.V. Appleton[2] published the theory of radio wave propagation in plasma describing the physical and the electrical properties of the ionosphere; see Appleton and Beynon (1940). It further led to the development of the theory of the ionospheric refractive index[3] as a function of the plasma frequency (Appleton and Beynon (1947)). Hence it can be appreciated that the subject of ionosphere research is at least as old and as important as the radiowave propagation itself. Since then, the era of ionosphere probing and space weather studies had already begun.

The ionosphere is colocated within the lower part of the thermosphere which is located between 50 km and 1000 km altitudes comprising of electrons and atoms; see Groves (1972), Doornbos (2012a). It absorbs the ultra-violet (UV) radiation and results in the dissociation of molecules as well as the ionization of atoms; see e.g. Schunk (1988), Hanslmeier (2002). Ionization is caused by both the solar radiation and the energetic particles. Accordingly they are called the photo-ionization and the particle-ionization, respectively, with the latter being dominant during adverse space weather conditions; see Rishbeth and Garriott (1969), Davies (1990).

This chapter is organized as follows: First, a brief summary of selected empirical models will be provided in Section 2.1. Subsequently, in Section 2.2, the three ionospheric key parameters, namely, (1) Slant Total Electron content (STEC), (2) Vertical Total Electron content (VTEC) and (3) the electron density will be introduced. A short introduction to the ionosphere stratification and a comprehensive list of previous relevant work on this topic will be presented in Section 2.3 and Section 2.4, respectively. In Section 2.5, some of the key challenges in EDM will be introduced and thereafter, in Section 2.6, the novelties introduced in this work will be listed.

As there are many textbooks, e.g. Schunk (1988), Hanslmeier (2002), Zolesi and Cander (2014), Jakowski (2017) and literature on this topic as summarized in Sections 2.4 and 2.5, the major parts of this thesis will focus on the development of ICOA as well as application to EDM and their numerical evaluations in the Chapters 4, 5 and 6 respectively.

2.1 Empirical models

Empirical models of the atmosphere are an important set of a tool suite[4] used by both the operational as well as the scientific communities. For the scope of this chapter, we will consider the following two categories of empirical models, namely, (1) the empirical thermosphere model (ETM) and (2) the empirical ionosphere model (EIM). Both the ETM as well as the EIM use proxy indices

[1] Guglielmo Marconi was an Italian inventor and electrical engineer, also regarded as the inventor of radio and shared the 1909 Nobel Prize in Physics.

[2] Sir Edward Victor Appleton was an English physicist and pioneer in radiophysics. He won the Nobel Prize in Physics in 1947 for his seminal work proving the existence of the ionosphere during experiments carried out in 1924.

[3] Refractive index is the ratio of the speed of light in a vacuum to that in a second medium of greater density

[4] A tool suite is a combination of many software sub-routines with or without a user-interface for operating the underlying software; e.g. an empirical model such as the IRI.

such as $F10.7$[5], Kp [6] for the characterization of space weather conditions. Selected space weather indices along with their description and temporal resolution are shown in Table 2.1. The relation between Kp and ap is shown in https://www.ngdc.noaa.gov/stp/GEOMAG/kp_ap.html. The ap index related information can be found in https://impc.dlr.de/affects/geomagnetic-indices/ae-index/index.html. The Dst index is an index of magnetic activity derived from a network of near-equatorial geomagnetic observatories that measures the intensity of the globally symmetrical equatorial electrojet (the "ring current"), as quoted from https://www.ngdc.noaa.gov/stp/geomag/dst.html. The F10.7 is measured in solar flux units (SFU) with 1 SFU = $10^{-22} Wm^{-2} Hz^{-1}$; see e.g. Maruyama (2010). The correlation between selected SWI and VTEC is graphically shown in Figure 2.1 and their respective correlation coefficients tabulated in Table 2.2, where only $F10.7$ shows a value exceeding 0.5. It is known that during adverse space weather conditions, there is a spatio-temporal variation of both the neutral as well as the electron density in the thermosphere-ionosphere system; see Richmond (1996), Jakowski et al. (2002). Empirical models are based on past data and describe an average climatology but are not expected to provide accurate model output values corresponding to adverse space weather input conditions. This can also be seen from the study of the estimated scale factors to the empirical thermosphere models based on SLR observations to spherical satellites by Panzetta et al. (2019), Zeitler et al. (2021).

Table 2.1: Description of selected SWI along with their temporal resolution

Index	Description	Temporal resolution
Kp	"**p**lanetarische **K**ennziffer" is obtained from averaging the indices from 13 mid-latitude observatories (no units)	3 hour
ap	ap is a related index obtained from Kp (whereas Kp is roughly logarithmic, ap is roughly linear). Ap is a running 24 hr average of ap	3 hour
AE	**A**uroral **E**lectrojet is designed to give measure of auroral zone magnetic activity. (units: nT)	1 minute
Dst	**D**isturbance **st**orm index is used to compare the intensity of geomagnetic storms based on the average value of the horizontal component of the Earth's magnetic field (units: nT)	1 hour
F10.7	Solar **f**lux emitted at 2800 MHz with wavelength of **10.7** cm (units: solar flux units (SFU))	8 hours

Generally, the ETMs are expected to provide neutral density and neutral temperature but not the ionosphere parameters such as the electron density or the electron temperature. Therefore dedicated EIMs e.g. IRI (Bilitza (2000)) or NeQuick (Di Giovanni and Radicella (1990)) are commonly used to obtain those or any other ionospheric parameters such as the total electron content (TEC), ion

[5]The F10.7 index has been measured consistently in Canada since 1947, first at Ottawa, Ontario; and then at the Penticton Radio Observatory in British Columbia, Canada.

[6]The estimated 3-hour planetary Kp-index is derived at the National Oceanic and Aerospace Administration (NOAA) Space Weather Prediction Center (SWPC) using data from the following ground-based magnetometers: Sitka, Alaska; Meanook, Canada; Ottawa, Canada; Fredericksburg, Virginia; Hartland, UK; Wingst, Germany; Niemegk, Germany; and Canberra, Australia. These data are made available thanks to the cooperative efforts between SWPC and data providers around the world, which currently include the U.S. Geological Survey, Natural Resources Canada (NRCAN), the British Geological Survey, the German Research Centre for Geosciences (GFZ), and Geoscience Australia. Quoted from source: https://www.swpc.noaa.gov/products/planetary-k-index accessed 17 September 2022.

temperature etc. Analogously, the EIMs also do not directly provide any thermosphere parameters. It shall be noted that both the ETMs as well as the EIMs are driven by similar (and in most cases the same) space weather indices; see e.g. Maruyama (2010), de Wit and Bruinsma (2017), Jackson et al. (2020).

Due to limitations with spatio-temporal resolution of especially the F10.7 and the Kp, as shown in Table 2.1, the empirical models cannot take into account the real-time atmosphere variability at a high temporal resolution[7]. Even though the thermospheric models are not used in this thesis, for the sake of completeness, we mention here that the two commonly used ones are the Jacchia-Bowman 2008 (JB2008) model (Bowman et al. (2008)) and the Naval Research Laboratory's Mass Spectrometer Incoherent Scatter Extension 2000 (NRLMSISE-2000) model (Picone et al. (2002)). On the other hand, the IRI model (version 2012) has been used by Limberger (2015) and Liang (2017) as well as in this work. The above mentioned empirical models have been developed from different measurement campaigns over many years and have been extensively validated as well; see e.g. Coiosson et al. (2005), Maruyama (2010), Altadill et al. (2013), Bjoland et al. (2016), Endeshaw (2020).

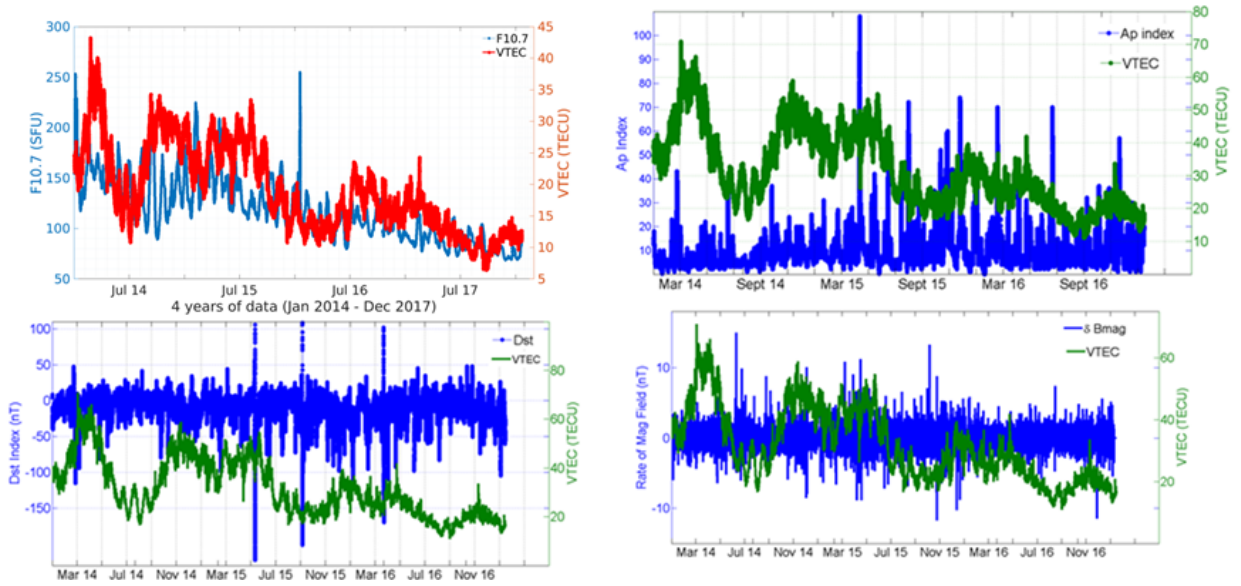


Figure 2.1: Comparison between four selected space weather indices (SWI) obtained from `http://omniweb.gsfc.nasa.gov` and 2 hour sampled VTEC time series data obtained from DGFI-TUM generated GIM over 2014 - 2016.

Table 2.2: Correlation coefficient between selected space weather indices (SWI) and VTEC over a period of three years from 2014 - 2016. The first three SWI are already defined in Table 2.1 and the last row contains the rate of magnetic field intensity denoted by δBmag. VTEC used in analysis was obtained from DGFI-TUM generated global ionospheric map (GIM).

Chosen SWI for correlation with VTEC	Correlation coefficient
F10.7	0.78
Ap	0.09
Dst	0.14
δBmag	0.11

[7]Typically the EIM run with a maximum spatial resolution of upto 1° along the latitude and longitude and a nominal temporal resolution of 30 minutes to 1 hour.

Contrary to the above mentioned empirical models, there are physical models which can provide both the neutral as well as the electron densities as direct model outputs. The physical models are based on the themodynamics, electrodynamics, optics laws and hence are considerably different[8] compared to empirical ones. One example is TIEGCM[9], which is a thermosphere-ionosphere coupled model (Richmond et al. (1992)). It uses the fluid convection[10] and momentum exchange processes[11] to calculate both electron and neutral density among the other TIEGCM output parameters (Richmond et al. (1992)). TIEGCM takes the interplanetary magnetic field (IMF) and the solar wind[12] parameters to calculate the high latitude electric potential using the Weimer model (Weimer (2005)) or the Heelis model (Heelis (2004)). Other similar physical models are DTM (Bruinsma (2015)) and CTIPe (Codrescu et al. (2012)). Although such models enable a detailed study of the atmospheric coupling processes but also have limitations with regard to the spatial coverage especially along the altitude range above 600 km or in the polar regions; see Richmond et al. (1992). For example, the TIEGCM version 1.92 has an altitude limitation of 590 km; see Qian et al. (2014). Depending on the chosen input scenario, the execution of the physical models can also be computationally exhaustive and configuring them with realistic atmospheric conditions remains challenging.

Even with the physical models, the thermospheric density is still determined on the basis of thermodynamic principles, but still not on the basis of input ionospheric parameters. Therefore despite having a variety of empirical and physical models, the ionosphere electron density modelling is still a highly sought-after topic of research in aeronomy and geodesy.

This is the primary motivation to develop a global 4D EDM at high spatial and temporal resolution. Although we will not discuss the thermospheric models any further in this thesis but the above discussion helps in the context, relevance as well as the need for a high accuracy global 4D EDM. The remaining chapters of this thesis will describe the development of our EDM in detail. However, before we go into the mathematical details on the observation modelling (see Chapter 3) and parameter estimation (see Chapter 4), the next section will introduce the three ionospheric key parameters[13].

2.2 Ionospheric key parameters

This section begins with the modelling of the electron density function

$$N_e(\varphi, \lambda, h, t) \tag{2.1}$$

depending on the latitude φ, the longitude λ, both defined in an Earth-centered coordinate system, on the height $h = r - R$ above the surface of a spherical Earth with a constant radius R and r being the radial distance and on the time t with high accuracy and high resolution. Since modelling a 4D function is rather challenging, for a long time the electron density was transferred into the 3D VTEC depending on the two spatial variables latitude and longitude as well as the time. Integrating

[8] Especially with regard to their functional procedures to generate an output parameter e.g. the neutral density or the dependencies of the different intermediate parameters e.g. electric and magnetic field intensities to the neutral density.

[9] TIEGCM is an acronym for Thermosphere Ionosphere Electrodynamics General circulation Model

[10] Convection is mass transfer due to the bulk motion of a fluid; see Richmond et al. (1992)

[11] A physical law that describes the transport of mass and pertaining to the conservation of momentum; see Richmond et al. (1992)

[12] The solar wind is a stream of charged particles released from the Sun consisting of the electrons, the protons and the alpha particles with kinetic energy between 0.5 and 10 keV; c.f. https://en.wikipedia.org/wiki/Solar_wind accessed on 27 March 2022

[13] It shall be made clear that the ionospheric key parameters are different from the Chapman key parameters. The latter will be presented in Chapter 5.

the electron density (2.1) along the ray-path between the transmitting satellite P^S and the receiver P_R the Slant Total Electron Content (STEC)

$$STEC(P^S, P_R, t) = \int_{P^S}^{P_R} N_e(\varphi, \lambda, h, t)\, ds \tag{2.2}$$

is obtained. Due to the dispersive property of the plasmasphere and the ionosphere the ray-path with the differential line element ds is depending on the carrier frequency f of the electromagnetic wave and deviates from the straight line between P^S and P_R. To simplify the notation and bring vector notation into play we introduce the 3D position vector $\mathbf{x} = r \cdot [\cos\varphi\cos\lambda, \cos\varphi\sin\lambda, \sin\varphi]^T$ and rewrite Eq. (2.2) as

$$STEC(\mathbf{x}^S, \mathbf{x}_R, t) = \int_{P^S}^{P_R} N_e(\mathbf{x}, t)\, ds\,. \tag{2.3}$$

Multiplying the right-hand side of Eq. (2.3) with the factor $\pm 40.3/f^2$ we obtain the ionospheric delay $d_{\text{ion}} = \frac{\pm 40.3}{f^2} \cdot STEC$. This relation is an approximation, valid for standard conditions of the geomagnetic field and signal frequencies f larger than 1 GHz. It neglects the effects of higher order which may reach up to 0.2 cm in zenith direction for GPS signals, depending on the sign it is valid for pseudorange measurements ($'+'$) or carrier phase observations ($'-'$). The conversion of STEC into VTEC is achieved by introducing an isotropic mapping function $M(z)$ depending on the zenith angle z. It performs the projection $VTEC(\mathbf{x}_{IPP}, t) = M(z) \cdot STEC(\mathbf{x}^S, \mathbf{x}_R, t)$ under the assumption that all electrons are concentrated in one thin spherical layer with constant height. This procedure is denoted as the single layer model (SLM) approach; the spatial reference for the VTEC observation is given by the ionospheric piercing point P_{IPP} with position vector $\mathbf{x}_{IPP}$ of the straight-line connection between the transmitting satellite and receiver station e.g. with the sphere of the SLM; see e.g., Schaer (1999).

2.3 Ionospheric stratification and Chapman key parameters

Although ionosphere layers and its vertical profile will be described in Chapter 5, a short introduction to the most important terminologies is made in this section in order to assist the reader. Ionosphere is stratified into the D–, the E–, the F_1–, the F_2–layer and the plasmasphere[14]. Each layer has its physical significance, due to the dominating atmospheric processes such as the ionization, the recombination and the overall energy balance of the ionosphere (see e.g. Chapman (1931)) and thus giving way to "physics-motivated" profile functions such as the Chapman and the Epstein functions; see Feltens (2007), Rawer (1988). A graphical and mathematical representation of profile function will be shown in Chapter 5; Fig. 5.1 and Eq. 5.2 respectively.

The D-, E- and F_1-layers are important for understanding of the geophysical phenomena such as electromagnetic radiation scatter and scintillation, especially the ionospheric sporadic E–layer (Es); see Smirnova et al. (1988), Arras et al. (2009), Zolesi and Cander (2014) and Tsai and Su (2018). It shall be noted that both the D– and the Es–layers are thin compared to the F_2–, the F_1–layers (see Berkner and Wells (1934), Bremer (1998)), as well as the magnitude of the D– and the E–layer peak electron densities are at least 2 order of magnitude lower compared to that of the F_2–layer under nominal ionosphere conditions; see Liang (2017).

For modelling the 4D electron density (2.1), we will introduce three Chapman key parameters, namely the peak density, the peak height and the scale height, with which the Chapman profile function for each layer is defined. The peak density corresponds to the maximum value and the peak height refers to the altitude corresponding to the peak density in an electron density profile. The scale height refers to the width of the profile around the peak height. All the three key parameters vary on spatio-temporal scales.

[14] As will be shown later in Chapter 5; Fig. 5.1

2.4 Previous relevant work

2.4.1 Relevance of B-splines to ionosphere modelling

Ionospheric parameters are often represented either as multi-dimensional series expansion in appropriate basis functions such as spherical harmonics (SH) (see e.g. Schaer (1999)), B-splines (BS); see e.g. Schmidt et al. (2015), or by discretization technique such as voxels (see e.g. Hernández-Pajares et al. (1999)). The specific choice among those three techniques is a trade-off among factors such as data distribution and representation of the underlying signal with a desired spectral resolution; see Schmidt (2011), Goss et al. (2019). SHs are one of the commonly used technique for the representation of 2D parameters, such as VTEC, on a sphere with the spatial resolution controlled by its degree and order, see Schaer (1999). It provides global support, i.e a localization in the spatial domain does not occur. Howe et al. (1998), Gao and Liu (2002) showed a 4D representation of ionosphere using SHs and empirical orthogonal functions (EOF). An overview of the SHs is given by Limberger (2015). Voxels are an alternate modelling approach, introduced in Hernández-Pajares et al. (1998a) with predefined cell width and height. The electron density within each voxel is held constant and has observation associated to it if either a measurement is directly available at that location or if a ray path from e.g. GNSS satellite crosses the voxel along the line of sight to a receiver. This approach was used for modelling VTEC in Hernández-Pajares et al. (1998b), Hernández-Pajares et al. (1999) and the F_2–layer peak electron density $N_m^{F_2}$ modelling by Gerzen et al. (2013). The path length within a given voxel, its spatial dimension (spatial resolution) and availability of the measurement in the proximity of its geometrical center are important considerations in the voxel approach. For further details we refer to Garcia Fernández (2004) in addition to the above mentioned references.

2.4.2 Legacy of B-splines and their advantages

At DGFI-TUM, and also in this thesis, the B-spline basis functions are used for ionosphere modelling, see Schmidt (2007), Schmidt et al. (2015), Limberger (2015), Liang et al. (2016), Erdogan et al. (2017), Goss et al. (2019), for two main reasons (1) they are characterized by their localizing feature (2) they can be used to generate a multi-scale representation (MSR); see Goss et al. (2019). An introduction to BS is given in Schumaker and Traas (1991), Lyche and Schumaker (2000) and they are characterized by their degree and level. For global modelling we use polynomial BS along the latitude and trigonometric BS along the longitude, as also used by Schmidt et al. (2015), Erdogan et al. (2017), Goss et al. (2019), Erdogan et al. (2020), Goss et al. (2020). BS and their application to our EDM will be discussed in Section 5.2. Compared to SHs, the most important advantage of BSs is that they provide localizing features but we need to define additional constraints on the geographic north and south poles.

More recently, a 3D BS representation of the ionosphere for precise positioning for the Australian region is described by Olivares-Pulido et al. (2019), where the product of three polynomial BS is set up using STEC observations. A similar approach with 3D (spatial) BS representation for electron density was also used before by Zeilhofer (2008). This particular approach of using BS for 3D increases the number of coefficients to be estimated, but is free from any assumption about the empirical distribution of electron density along the vertical. However with GNSS STEC observations, due to the short baseline distance between receivers in a regional network, Olivares-Pulido et al. (2019) mentioned that the tomographic method cannot provide a high vertical resolution.

2.5 Challenges in EDM

Earlier studies by Limberger (2015), Liang (2017) highlighted the two main challenges in EDM, which are (1) an inhomogenous and insufficient observation coverage and (2) significant correlations among

the Chapman key parameters. The former refers to data gaps in e.g. STEC observations due to the non-homogeneously distributed GNSS receiver stations and thus, leading to a non-uniform data distribution from the equator to the poles; see Erdogan et al. (2017), Goss et al. (2019). Furthermore, the northern and the southern hemisphere have a marked difference in the observation coverage as well. These problems are also relevant for the ionosonde observations which are only available over the landmasses. Ionospheric radio occultation (IRO) is another technique to obtain electron density measurements but those are limited by the availability of active occultation satellites, the altitudinal coverage, the need for observation pre-processing, and quality checks. In-situ observations from GNSS receiver on-board GRACE satellites or instruments such as the Langmuir probe on-board CHAMP provide only point measurements. These may be used as additional observations but not as standalone source for the global electron density modelling. Therefore, a background empirical ionospheric model is often used along with real observations to fill the above mentioned data gaps. The latter case of Chapman key parameters being mutually correlated is due to the electron-ion recombination and different coupling processes in the thermosphere-ionosphere system; see e.g., Stankov et al. (2003), Stankov. and Jakowski (2006), Belehaki et al. (2006). Statistical correlations between key parameters arise due to their observability and sensitivity to electron density[15]. This was shown by Limberger et al. (2013) and Limberger et al. (2014) with regard to regional ionosphere modelling considering only the F_2-layer in the South American region using a combination of real and simulated observations.

Limberger (2015), Liang (2017) also highlighted the need for stochastic modelling as well as inequality constraints on the Chapman key parameters. This thesis considers both these issues and especially the latter in detail[16]. Moreover, Liang (2017) mentioned that in global EDM, the unknown Chapman key parameters parameters are correlated, the estimation accuracy can be improved by imposing inequality constraints[17].

In addition to the aforementioned challenges, even if an observation model is setup, a robust method is still required to estimate the unknown parameters. Traditionally, the method of least squares has been used successfully in most circumstances. However, one of the disadvantages of unconstrained least squares is its inability to bound the magnitude of unknown parameters to within certain upper and lower limits.

Need for inequality constraints: an example case

Without going into detailed mathematical formalism adopted for denoting the essential variables, in the following paragraph, a hypothetical example problem arising in EDM is shown, which will demonstrate the need for applying inequality constraints.

Consider the problem of global modelling of the two Chapman key parameters[18], namely, (1) F_2 layer peak density $N_m^{F_2}$ and (2) F_2 layer peak height $h_m^{F_2}$. Furthermore, for the illustration of this example, let us also assume that the F_1 layer peak height $h_m^{F_1}$ is given and known a priori (e.g. from an empirical model). Let $\widehat{\boldsymbol{d}}_{N_m^{F_2}}$ and $\widehat{\boldsymbol{d}}_{h_m^{F_2}}$ denote the respective least squares estimates of the B-spline coefficients of the two above-mentioned Chapman key parameters. Without going into the details of transformation of the coefficients to key parameters (which will be described in Chapter 5; Eq. (5.5)), let $\widehat{\boldsymbol{N}}_m^{F_2}$ and $\widehat{\boldsymbol{h}}_m^{F_2}$ denote the two sets of estimated Chapman key parameter vectors respectively. Now one of the following cases may arise:

[15] or to one of its functional e.g. STEC

[16] Stochastic modelling is already well described in the PhD theses Limberger (2015) and Liang (2017) but the inequality constraints were not.

[17] As the inequality constraints partially decorrelate the unknown parameters by localizing them to within specific magnitude bounds.

[18] This example scenario was also widely used in both Limberger (2015) and Liang (2017).

- Case 1: A subset of the estimated coefficients are negative. Clearly, this is a physically unrealistic scenario because both the F_2 layer peak density as well as peak height are non-negative parameters.
- Case 2: All the estimated coefficients are positive, i.e $\widehat{\boldsymbol{d}}_{N_m^{F_2}} > \mathbf{0}$ and $\widehat{\boldsymbol{d}}_{h_m^{F_2}} > \mathbf{0}$ are satisfied[19], but a subset of the estimated coefficient vector $\widehat{\boldsymbol{d}}_{h_m^{F_2}}$ is physically unrealistic. It means that at $m \cdot n$ discrete latitude, longitude pairs $\{\varphi_i, \lambda_j\} \quad \forall \; i = \{1, \cdots, m\}, \quad \forall j = \{1, \cdots, n\}$, we have $\widehat{h}_m^{F_2}(\varphi_i, \lambda_j) < h_m^{F_1}(\varphi_i, \lambda_j)$. In other words, at locations φ_i, λ_j, the F_2 layer peak height is below that of the F_1 layer, which is physically unrealistic.

These two cases are just examples to illustrate the possible challenges when unconstrained least squares are used in EDM. Such problems could be overcome if the unknown Chapman key parameters are subject to physically realistic inequality constraint bounds:

- $N_m^{F_2}(\varphi, \lambda) > 0 \quad \forall \quad \varphi \in \boldsymbol{\varphi}, \; \lambda \in \boldsymbol{\lambda}$
- $h_m^{F_2}(\varphi, \lambda) > h_m^{F_1}(\varphi, \lambda) \quad \forall \quad \varphi \in \boldsymbol{\varphi}, \; \lambda \in \boldsymbol{\lambda}$,

where $\boldsymbol{\varphi}$, $\boldsymbol{\lambda}$ denote the latitude and longitude vectors respectively where the constraints are to be applied. This way the physical realism of the ionosphere layers are included in the EDM.

2.6 Proposed approach and novelties

The Chapman key parameters, namely, the peak electron density, the corresponding peak height and the scale height, of each layer are modeled by series expansions in terms of polynomial B-splines for latitude and trigonometric B-splines for longitude. The Chapman profile function is chosen to define the electron density along the altitude. This way, the EDM is set up as a problem of estimating the unknown coefficients of the Chapman key parameters. We estimate the unknown coefficients using the ICOA within which an optimization solver, such as the gradient descent or the Newton's method; see the Eqs. (4.36) and (4.48), is implemented. For demonstrating the proposed ICOA, a subset of the total key parameters from Eq. (5.4) would be chosen to demonstrate the application of the ICOA to our EDM. Simiar approaches were followed by Limberger et al. (2014), Liang et al. (2015b) and Tsai et al. (2016). While in the former two papers BS, in the latter case SHs are used for modelling the F_2-layer peak height and the peak density.

The main novelty introduced in this thesis is the development of ICOA and its appication to global 4D EDM. To the knowledge of the author, this approach has not yet been developed and applied to ionosphere modelling.

There are only a few relevant references where inequality constrained optimization is applied in geodesy, e.g. "geodetic levelling and datum definition" by Koch (1985), Koch (1988), non-negative variance component estimation (VCE) as well as zenith troposphere delay (ZTD) estimation in Roese-Koerner (2015). In the former two papers, the Lagrange multiplier (LM) and the slack variable (SV) concepts are already used (introduced and described in detail in Chapter 4) and the latter used the so called "Active-set methods"; see Boyd and Vandenberghe (2004), Nocedal and Wright (2006). Therefore, in this work the ICOA will be described in detail.

It shall be noted that generally an ICOP contains both the equality as well as inequality constraints. As a special case, if there are no inequality constraints, we have an equality constrained optimization problem. Although equality constraints alone can also be set up in our modelling problem, they are not sufficient to address the aforementioned challenges of 4D global EDM. Also it shall be noted that the solution of ICOP with only equality constraints do not necessarily require an ICOA with gradient descent or Newton's method. Infact, equality constraints can already be included alongside least squares; see Chapter 3 in Koch (1999).

[19] $\mathbf{0}$ denotes the vector of zeros with the same number of elements as that on the left hand side of the inequality considered in Case 2. Once again the author wishes to stress that the mathematical formalism and conventions for inequalities are introduced in Section 4.4.1

Inequality constraints, on the other hand, are relatively difficult to handle (see Nocedal and Wright (2006); Chapter 16) and optimization methods are used instead of least squares. Although there are software packages available based on the work of Lötstedt (1984), Gill et al. (1984), Coleman and Li (1996) and Mead and Renaut (2010) but their algorithmic details, assumptions and implementation along with the standard deviations are either not available or not explicitly transparent to the user. In many cases the algorithms are also not easily adaptable for large-dimensional problems (e.g. global 4D EDM). This is the main motivation to develop our own ICOA for the EDM. The following novelties within the ICOA development are noteworthy:

- defenition of the inequality constraints qualification concept; e.g. constraints being active or inactive; see definition 4.4.2 in Section 4.4
- description of the convexity of the ICO objective function as well as the feasible region; see definition 4.2.2 in Section 4.4.2
- transformation of the ICOP in a so-called standard-convex form (4.57) to a partial ICOP (p-ICOP); see Eqs. (4.65) and (4.66). As it will be shown in Chapter 4, the phrase partial in p-ICOP signifies that the inequality constraints have been partially transformed to equalities; see Eq. (4.65).
- describe the procedure for combining the unconstrained Gauss-Markov Model (GMM) (4.8) with inequality constraints; see Section 4.4; especially Eqs. (4.49), (4.50e) and (4.56).
- description of the necessary and sufficient optimality conditions (4.71) to (4.72e) when an ICOP has a feasible solution; see Section 4.5 for further details.

Keeping in view the aforementioned aspects, in this work we have (1) developed the EDM considering multiple ionosphere layers, (2) used 2D BS basis functions for the representation of the Chapman key parameters along latitude and longitude, (3) used the Chapman profile function for altitude dependency of electron density and (4) combining of Formosat-3/COSMIC and GRACE radio occultation as well as VTEC observations, as described in Section 3.2.

3 Space and ground based observation techniques

In Chapter 1, the motivation and rationale for electron density modelling was introduced. In Chapter 2 the ionospheric physics were described along with the spatio-temporal variability of the important parameters such as VTEC and the electron density.

> The goal of this chapter is to describe the most important space and ground based geodetic observation techniques for electron density modelling.

At the same time, their relative processing complexity and their utility for an independent validation will be discussed. A timeline of observation techniques along with the prevailing F10.7 space weather index is shown in Fig. 3.1. It gives an indication of the number of observation techniques available for electron density modelling at any given time frame.

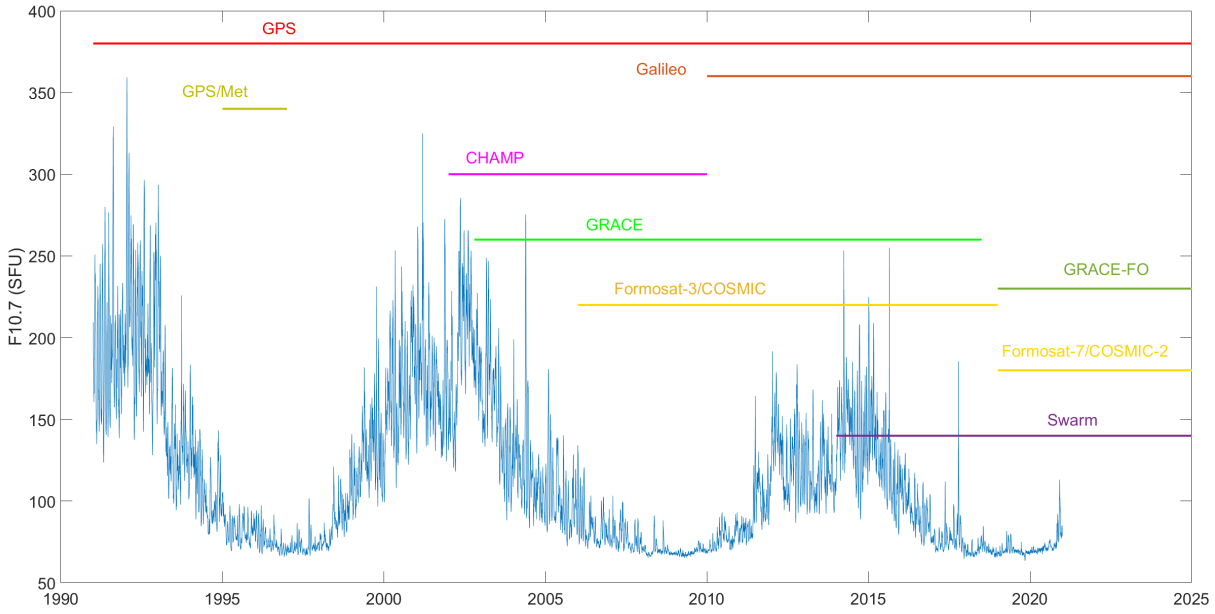


Figure 3.1: Timeline of observations from selected satellite missions as well as the F10.7 time series from 1990 to present.

This chapter is organized as follows:

- First, the GNSS observation equations and STEC computation will be developed and described.
- Subsequently, the radio occultation technique will be introduced and Formosat-3/COSMIC[1] (F3C), GRACE derived electron density profiles, as well as their spatial coverage, spatio-temporal resolution, data availability, pre-processing and quality control will be discussed.
- In-situ observations from Langmuir probe on-board the Swarm satellites, CHAMP and ground based ionosonde will be described. Especially, the ionosonde and in-situ observations will be used for an independent validation in Chapter 7.

[1] COSMIC, the Constellation Observing Systems for Meteorology, Ionosphere, and Climate

3.1 GNSS observations for electron density modelling

3.1.1 GNSS receiver and observation types

The use of GNSS for ionosphere modelling is an active area of research with continuous improvements in quality as well as availability of data products, e.g. VTEC; see Limberger et al. (2013), Liang et al. (2015a), Liang et al. (2015b), Erdogan et al. (2017), Goss et al. (2019), Erdogan et al. (2020), Goss et al. (2020). The growing number of receivers in ground networks (e.g. IGS) provide globally distributed measurements and, especially with the evolution of receivers capable of tracking multi-GNSS signals, the combination of multi-frequency GNSS observations is possible for global electron density modelling.

GNSS receivers are categorized based on the signal tracking technology (correlator design), constellation supported (single or multi-constellation[2]), low-cost or geodetic receivers or the number of tracked frequencies; see Eissfeller and Won (2017), Meurer and Antreich (2017) for a detailed categorization. Accordingly, the observation quality differs, which is one of the important considerations for electron density modelling.

Goss et al. (2019) show that spatial distribution of ground based GNSS observations is non-homogeneous and results in data gaps, which are in general undesirable for global modelling of any ionospheric parameter such as VTEC or the electron density. This problem is mitigated using a background ionospheric model, e.g IRI (Bilitza (2000)) or NeQuick (Di Giovanni and Radicella (1990)). In-situ GNSS measurements from LEO satellites can contribute as additional observations as well, with the only drawback that there would be a lack of observations except at the LEO altitude. Furthermore, GNSS also enables techniques such as radio occultation (RO); see Rocken et al. (2000), Hajj et al. (2002), which allow the study of vertical ionosphere structures[3]. In this work, terrestrial STEC measured by GNSS receivers is combined with radio occultation observations, to have an improved observation coverage of both the top as well as bottom-side ionosphere globally[4].

3.1.2 Relevance of multi-GNSS for ionosphere modelling

Two basic types of information are broadcasted by GNSS satellites, namely, the observation and the navigation data. The former contains pseudorange, carrier phase, doppler and signal-to-noise ratio measured on different frequencies. The latter contains satellite ephemeris, clock offset, ionospheric as well as health related information[5]. The rest of this section will describe one of the methods for ionospheric STEC computation using dual frequency GNSS pseudorange and phase observations. Agencies such as the international GNSS service (IGS) operate a network of receivers distributed worldwide and facilitate ionosphere modelling by providing station-specific RINEX files containing the so called the "GNSS observables" (see Dow et al. (2009), Beutler et al. (2009), Montenbruck et al. (2014b)).

Multi-GNSS enables two important and relevant aspects to this work, namely,

- Filling the data-gaps caused by the IPPs clustered from a single GNSS constellation
- The newer IRO missions e.g. Formosat-7/COSMIC-2 (F7C2), carry an improved occultation receiver capable of tracking multi-constellation and multi-freuency (MCMF) GNSS signals. This in-turn leads to a higher number of observations compared to that from F3C mission (see e.g. Fig. 3.4).

[2]When a receiver tracks more than one frequency and one constellation, it is capable of providing the so called "multi-frequency" (MF) and "multi-constellation" (MC) GNSS observations.

[3]Electron density variations with altitude.

[4]Generally top- and bottom-sides refer to above and below the F_2 layer peak height respectively.

[5]The broacast navigation file also contains the transmission group delay, which is also relevant for ionospheric modelling. However, the purpose of using the navigation file in this work is only to obtain satellite positions such that the ionospheric pierce point can be computed.

Therefore, this section describes the GNSS observables and their parameterization to compute STEC. At the same time, the increased processing complexity, relative to that from just one single constellation, will be highlighted.

For the sake of keeping the introductory chapters relatively short compared to the following chapters, the specific details of GNSS data processing can be found in Appendices D.

3.2 Radio occultation

The goal of this section is to introduce the ionosphere radio occultation (IRO) observation technique and summarize the most important electron density retrieval techniques from the literature. Subsequently, it is shown how the F3C profiles are obtained, pre-processed and used within the electron density modelling. As mentioned later in this section, the electron density profiles are directly obtained as profiles from the F3C data product and therefore the focus will be mainly on its pre-processing and quality control, as required for electron density modelling.

From satellite to terrestrial receivers, the GNSS signal travel paths are mostly along the radial direction helping in particular to determine the horizontal ionospheric variations[6]. An excellent geometrical scenario to resolve for the vertical stratification is when a signal is passing through the atmosphere rather horizontally while the receiver is located behind the Earth's limb; see e.g. Aragon-Angel et al. (2016). In this case, the signal source, i.e. the GNSS satellite, can reach an elevation angle below zero from the LEO receiver perspective. Depending on the LEO orbit height, the signal then passes the whole or partial ionosphere or even atmosphere.

GNSS phase observations are related, among others, to ionospheric electron density, see Eqs. (A.32), (A.35a) and (A.35b). A signal passing through the atmosphere[7] is refracted according to Snell's law[8] due to the vertical gradient of the refractive index (Jensen et al. (2004)).

Occultation

GNSS signals passing through the ionosphere will be refracted, resulting in the ray path bending from a straight line path. This is called an "occultation event" because the GNSS satellite is geometrically behind the Earth's limb from the field of view of the receiving LEO satellite.

The tracking of GNSS dual-frequency phase observations (A.24a) and (A.24b) in the occultation geometry during the so-called "occultation event" could result in the elevation angle of the observed satellite in the LEO receiver to be less than $0°$[9].

Retrieval of electron density

The retrieval of the electron density along the occultation tangent points (closest points on the occultation plane to Earth, shown in Fig. 3.2 and red dots in Fig. 3.3) require the processing of the respective signal bending angles, based on the physical principle that the refractivity N of the propagation medium is approximately proportional to N_e/f^2, where N_e and f denote the electron density and signal frequency respectively (Hernández-Pajares et al. (1998)).

[6] Measurements from low elevations may contribute to observe vertical structures, but increase measurement noise, induced by multipath.

[7] GNSS signals pass through the plasmasphere, the ionosphere and the bottom atmospheric layers. However, the focus on this work is mainly to quantify the occultation in the ionosphere.

[8] In the geometrical optics approximation, Snell's law states that the ratio of the sines of the angles of incidence and refraction is equivalent to the reciprocal of the ratio of the respective refractive indices.

[9] Occultation plane is defined by the intersection of radial vectors to GNSS and LEO satellite, their mutual orbit planes as well as their line of sight vectors; see Limberger (2015).

Furthermore, the dispersive and non-dispersive nature of the ionospheric and troposheric medium allows, their respective effects on the bending angle to be separated using the L1 and L2 phase combination, as shown in Eqs. (A.24a) and (A.24b) as also described by Tsai et al. (2011). This procedure is called the "IRO data inversion" and is part of a so called electron density profile retrieval process (see Jakowski et al. (2002) and Tsai et al. (2016)). The derived profiles contain discrete electron density observations that describe their distribution along the occultation tangent points and the relevant physical quantities, such as the maximum electron density $N_m^{F_2}$ or the corresponding peak height $h_m^{F_2}$, can be estimated. The physical concepts related to ionospheric radio occultation and electron density retrieval techniques are described in detail in Hernández-Pajares et al. (2009) and Limberger (2015). The geometrical situation of RO is schematically depicted in Fig. 3.3 for F3C as the LEO mission tracking GNSS signals.

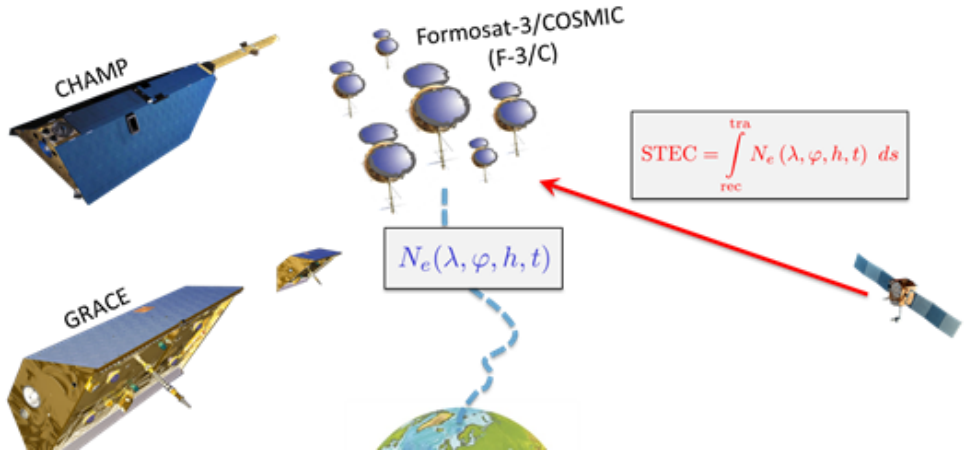


Figure 3.2: Electron density profile retrieval from F3C satellites providing the 4D electron density by STEC measured from GNSS satellites in view. Also shown are the GRACE and CHAMP satellites from which electron density profiles are retrieved and used in this work. Figure courtesy: Michael Schmidt, DGFI-TUM annual report 2019.

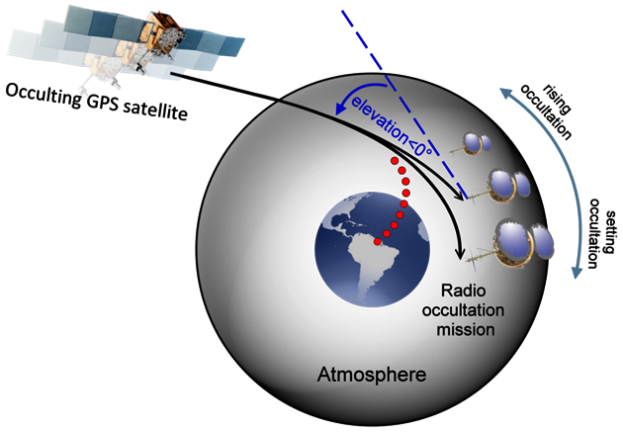


Figure 3.3: Example for an occultation event between a GNSS and F3C satellite. Shown in red are the occultation tangent points, as the relative geometry changes with time. Especially, the elevation angle of occulting GPS satellite from behind Earth's limb, resulting in a negative elevation angle is also shown. The direction of motion of LEO satellites resulting in rising or setting occultation is also shown. Figure courtesy: DGFI-TUM annual report 2019.

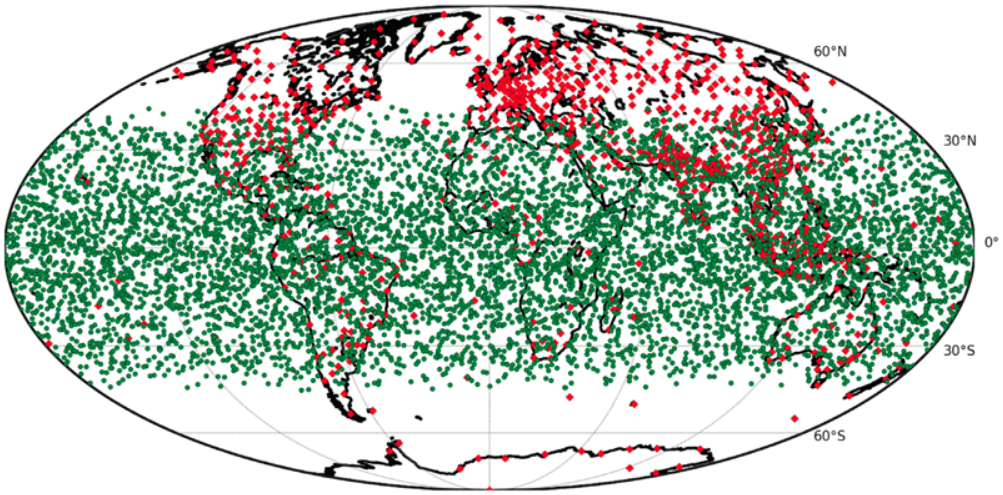


Figure 3.4: Location of the occultation points (in green) from Formosat-7/COSMIC-2 as on 27 Nov. 2019 and the locations of ionosonde stations (in red) operated by the national geophysical data centre (NGDC). Figure courtesy: Formosat-7/COSMIC 2 data portal.

3.2.1 Electron density profiles from Formosat-3/COSMIC

F3C is a bilateral collaborative mission of Taiwan and the United States (Rocken et al. 2000), a constellation of six LEO satellites launched in 2005. Its payload consists of GPS receivers, high-gain occultation and precise orbit determination (POD) antennas[10].

Electron density profiles are contained in the data product "ionPrf", which are computed by COSMIC Data Analysis and Archive Center (CDAAC)[11]. A detailed user manual (Schreiner et al. (2003)) is available with the list of parameters provided in the electron density profiles. A few examples of complete profiles, missing data, outliers and unusable profiles are shown along with the number of observations in each F3C profile in Fig. 3.5. Although the electron density measurements look like they are vertical profile but in fact they are located along the curved path of the occultation tangent points (TP). Electron density profiles from the occultation sensor on-board the GRACE satellites will also be used for EDM. The average number of observations per hour from GRACE, COSMIC and VTEC are shown in Fig. 3.6.

3.2.2 Pre-processing electron density profiles from CDAAC data

Before the "ionPrf" data is used for electron density modelling, a pre-processing step is performed for which, the following criteria have been used:

- **Minimum altitude** to be used in a profile is chosen at 80 km, which means only electron density values above 80 km are used. The specific threshold was chosen to match with the bottom altitude of the D-layer. It is known that the IRO measurements especially at lower altitudes (below 200 km) have poor quality due to the low GNSS signal to noise ratio as well as quality from electron density profile retrieval techniques (Tsai et al. (2011), Limberger (2015)). If higher altitudes are chosen as thresholds, then the important observations from the $D-$ and the $E-$ layer become unavailable. Since a multi-layer approach is used for electron density modelling, measurements from the bottom layers are important for this work as well.

[10] Detailed information for constellation system, mission objective and technique development can also be found in a special issue publication for the F3C mission in the Journal for Terrestrial, atmosphere and Ocean, Vol.11, No.1, March 2000.

[11] Access to the data is free but there is almost a full dependency on CDAAC for checking the data quality and availability.

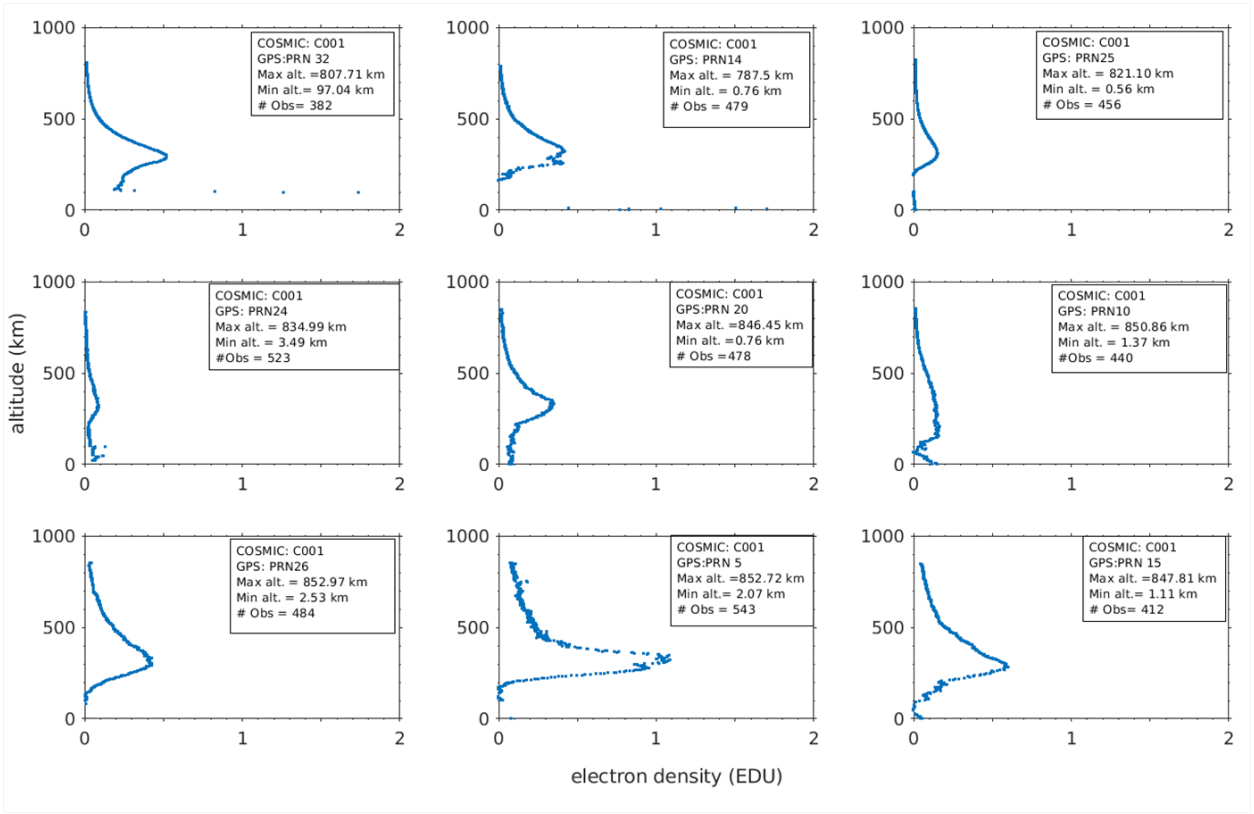


Figure 3.5: Electron density profiles retrieved from improved Abel transform. Data source: ionPrf data product from CDDAC, 12-March 2015.

- **Minimum number of observations** (the number of data points along a profile) shall be 150. This criteria was selected by analysing the average number of observations per profile over 12-14 March 2015 and one half of that average value (rounded to nearest 50) was chosen.
- **Minimum altitude dynamic range**, defined as the difference between the maximum and minimum altitude, shall be 200 km. This criteria was selected based on one-half of the average value of the altitude range, over F3C profiles during 12-14 March 2015.
- **Maximum allowed data gap** shall be 100 continuous data points. However, if these data-gaps occur in the F2 region (220 km $\leq$ altitude $\leq$ 440 km), then the complete profile is ignored. The threshold number 100 is chosen as 66% of the threshold minimum number of observations (which was 150).

The threshold values mentioned above are chosen as a fair balance between the observation quality and its availability. Strict thresholds lead to a relatively small number of IRO observations. Furthermore, an outlier detection is performed based on the following:

- The rate of change in electron density with altitude is computed for each profile. Upto 5 observations around any sample exceeding a threshold $|\frac{\Delta N_e}{\Delta h}| > 0.1$ EDU/km are removed from the profile[12]. It shall be noted that for an electron density sample i in any profile,

$$|\frac{\Delta N_e}{\Delta h}| = \frac{N_e(\varphi_{i+1}, \lambda_{i+1}, h_{i+1}) - N_e(\varphi_i, \lambda_i, h_i)}{\|h_{i+1} - h_i\|} \tag{3.1}$$

- A change in altitude exceeding 50 km between any two consecutive observations in any given profile are also removed. The value of 50 km was chosen based on Limberger (2015).

[12]Although F3C profiles are not vertical, yet the parameter $|\frac{\Delta N_e}{\Delta h}|$ is inversely proportional to the Chapman scale height.

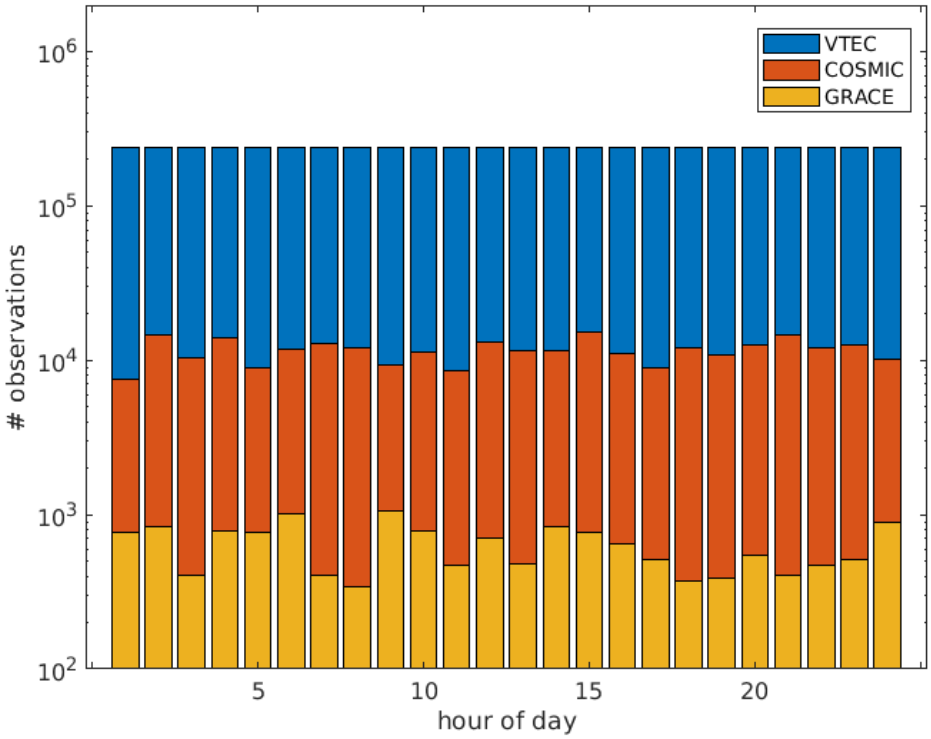


Figure 3.6: Average number of observations per hour from the three observation techniques between 12-14 March 2015.

- Measurements shall be available for within ±50 km around 350 km in a profile. This is ensured to allow a proper description of the profile shape in the F2 peak region. This threshold is also following Limberger (2015).
- The altitude corresponding to the F2 peak density in the profile is verified to be located within a physically reasonable range of 220 km - 440 km. This aspect is important considering the inequality constraints to be applied for the peak height, as described in numerical evaluations in Chapter 6[13].

3.3 Other geodetic observation techniques

Ionosonde

The working principle of an ionosonde is to transmit a radio pulse near to the vertical and to measure the travel time taken to receive its echo due to the reflection from the different ionosphere layers (see e.g. Hunsucker (1991), Belehaki et al. (2006) and McNamara et al. (2008)). As an ionosonde transmits signals of selected wavelengths[14], the critical frequencies[15] of the different ionosphere layers can be deduced from the characteristic "traces" of reflected signals, called an ionogram[16]. The traces contain the strength of the reflected signal, the frequency and phase offsets as well as the signal travel time.

[13] Additional diagnostics and exception checks were performed, as required, within the processing software developed for this work to ensure good measurement quality for electron density modelling.
[14] Corresponding to the so called "sweep frequencies" generally in the range of 0.1 – 30 MHz (Hunsucker (1991))
[15] The critical frequency of an atmospheric layer is the threshold at which the signal just passes through. Any lower, the signal would be reflected and a higher frequency than the critical frequency of a certain layer will not be reflected from that layer and it will pass to a relatively denser layer. (see e.g. Schunk (1988), Jakowski (2017)).
[16] Ionograms are 2-dimensional plot of the virtual height of the ionosphere along the vertical and frequency along the horizontal axis.

Especially, with the analysis of the phase of the received (reflected) signal, using an antenna array[17], it is possible to determine the temporal variations in the ionospheric electron density (Jakowski (2017)). Ionograms provide the critical frequency $f_o^{F_2}$, of the F_2 layer, from which $N_m^{F_2}$ can be derived (Hunsucker (1991)). In a similar manner, the Chapman key parameters of the D, E and F_1 layer can be obtained as well.

For this work, both the $N_m^{F_2}$ and $h_m^{F_2}$ are obtained from 17 ionosonde stations (see Table 3.1) belonging to the network of the National Geophysical Data Center (NGDC)[18].

Table 3.1: Ionosonde station used for validation

Station ID	Latitude, Longitude (deg.)	Location
GR13L	33.31 S, 26.52 E	Grahamstown, South Africa
EA036	37.09 N, 06.72 E	El Arenosillo, Spain
JR055	54.60 N, 13.40 E	Juliusruh, Germany
AS00Q	7.95 S,165.60 E	Ascension Island
AH223	23.00 N, 72.50 E	Ahmedabad, India
AT138	38.00 N, 23.50 E	Athens, Greece
EB040	40.80 N, 0.50 E	Roquetes, Spain
EG931	30.46 N, 86.55 W	Eglin, USA
FF051	51.70 N, 1.78 W	Fairford, UK
GU513	13.44 N,144.79 E	Guam, USA
HE13N	34.40 S, 19.25 E	Hermanus, South Africa
IC437	37.45 N,126.70 E	Incheon, South Korea
IF843	43.49 N,112.04 W	Idaho Falls, USA
IS141	41.00 N, 28.97 E	Istanbul, Turkey
JI91J	12.04 S, 77.04 W	Jicamarca, Peru
JJ433	33.48 N,126.49 E	Jeju Island, South Korea
LG178	78.22 N, 15.62 E	Longyearbyen, Norway

In-situ observations

Langmuir Probes (LP) are instruments used onboard satellites to measure the in-situ electron density and temperature of the plasma enountered along the orbit. It shall be noted that while CHAMP is a single satellite, Swarm is a constellation of three satellites as shown in Fig. 3.8. CHAMP and Swarm LP measurements (see Figures 3.7 and 3.9) will be used to validate the estimated electron density model values in Chapter 7. An example of the electron density from Swarm LP is shown in Fig. 3.11. For further details on ionospheric plasma density measurements by Swarm LP, their performance and limitations, see Knudsen et al. (2017), Diego et al. (2019), Catapano et al. (2020), and the corresponding one for CHAMP LP, see Cooke et al. (2003), Heise et al. (2005). The in-situ observations are directly used to validate the estimated key parameters, as will be shown in Fig. 7.19 with the result O_1. Swarm is also equipped with spaceborne GPS receiver and therefore an STEC data product (as example shown in Fig. 3.10) is made available from Swarm project office at

[17] Antennas alignment in a specific proximity to each other corresponding to the phase and wavelength of the received signal (Jakowski (2017)), also called a phase-array.

[18] Ionosonde data are obtained from the NGDC open source repository *ftp.ngdc.noaa.gov/ionosonde/*

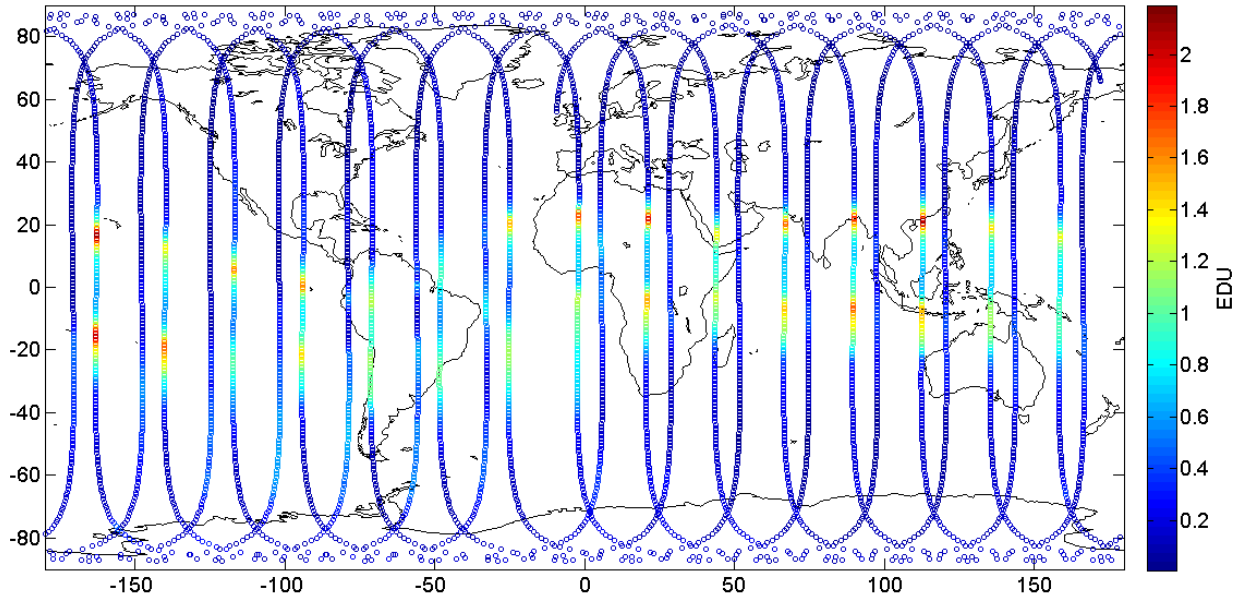


Figure 3.7: CHAMP electron density in-situ observations from Langmuir probe on 1 Aug. 2007. Courtesy GFZ Potsdam

ESA[19]. This in-turn can be used to validate the modelled key parameters, as will be shown in Fig. 7.19 with the result O_2.

Missions such as the Van Allen probes; see Lanzerotti (2013), Reeves et al. (2013), Mauk et al. (2013), Moldwin et al. (1995), whose science objective is the study and monitoring of the Earth's radiation belts, also provide electron densities at an altitude of 2 to 6 Earth radii. However, an initial assessment in this work reveals that the assumptions made for ionosphere and plasmasphere in the altitude range of 100 - 1000 km, and also the Chapman profile function in general, may not be appropriate to model the electron density at those altitudes and therefore Van Allen probe data was not used for this work. This aspect is substantiated further in Chapter 5.

Satellite altimetry, DORIS

Satellite altimetry (see Fu and Cheney (1995)) and Doppler Orbitography and Radio-positioning Integrated by Satellite or DORIS (see Auriol and Tourain (2010)) are already used for VTEC modelling at DGFI-TUM (see Dettmering et al. (2011b), Dettmering et al. (2011a),Dettmering et al. (2014), Erdogan et al. (2017), Erdogan et al. (2020)). As this VTEC data is used as observation within the separability approach, the DORIS and altimetry techniques are indirectly included in this work.

3.4 Summary of observation techniques

In this chapter we have introduced the STEC and the IRO observation techniques to be used for the global 4D EDM. Additionally, the ionosonde and in-situ observtaion techniques will be used for validation. Considering the relevance of GNSS in the computation of STEC as well as in the retrieval of electron density profiles from the F3C mission and the GRACE satellites, a detailed description of the observation equations, error sources including the phase ambiguity resolution has been provided (see Appendix D).

[19] https://www.esa.int/Applications/Observing_the_Earth/Swarm

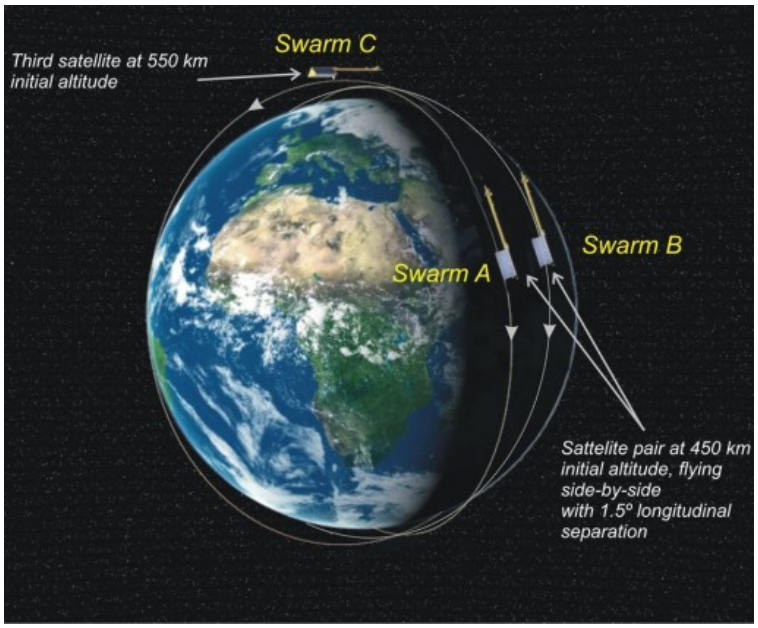


Figure 3.8: Swarm satellite orbit. Courtesy ESA

There are broadly two general categories of signal delay sources in the GNSS observation equations, namely,

- the dispersive or frequency dependent

  ionosphere delay, code multipath, the code, phase biases, phase ambiguities
- non-dispersive or frequency independent

  straight line true range, clock offsets and troposphere delay.

While the non-dispersive delays are eliminated in the GFLC, the remaining dispersive components e.g. DCBs are estimated along with remaining unknown parameters. Furthermore, the multipath effect is absorbed in the measurement error and the phase biases are absorbed by the integer ambiguity resolution. For the scope of this thesis, it is be assumed that the phase ambiguities are estimated as float and the focus will be on the applicability of STEC observations in the EDM.

The IRO has also been discussed and its advantages compared to the other techniques have been highlighted. Limb sounding is a complimentary technique compared to ground based, nadir pointing satellite payloads or along track pointed Langmuir probes, at least partially filling data gaps over the oceans and other regions such as deserts[20].

In summary, both STEC and IRO electron density retrieval require the GNSS observation equations; see Appendix (A.21a) to (A.21f). While the STEC is computed from ground GNSS receivers, the electron density profiles were obtained directly from CDDAC.

[20] As will be later specified, in order to mitigate the problem caused by sparse global distribution of IRO data, additional VTEC data from DGFI-TUM global ionospheric maps (Erdogan et al. (2017)) was also used for obtaining electron density observations using the separability approach.

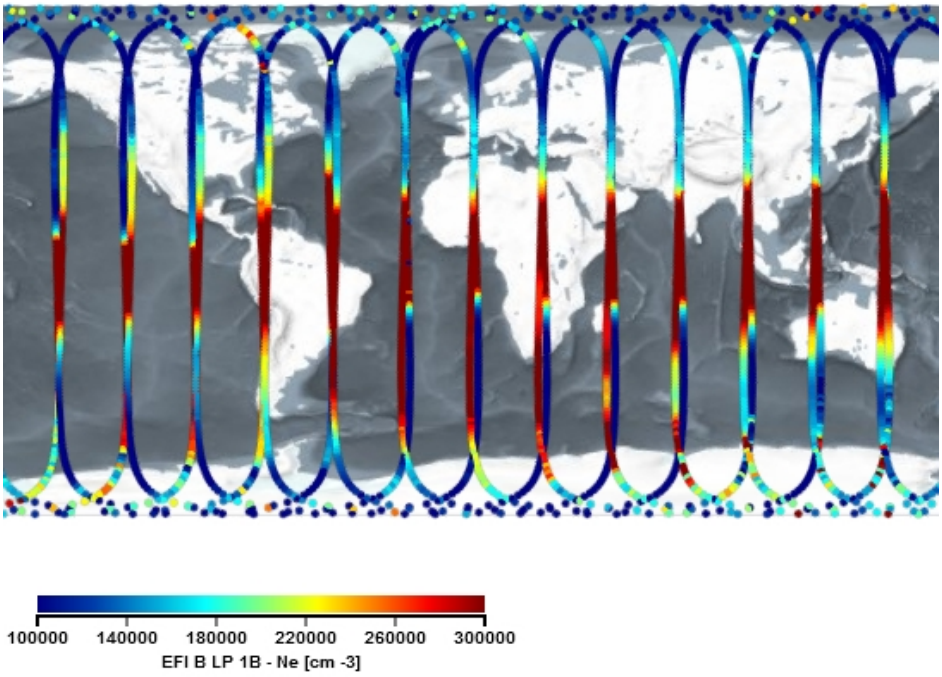


Figure 3.9: Swarm-B in-situ electron density on 12 March 2015 from on-board Langmuir probe. Data Courtesy: Swarm project office, ESA.

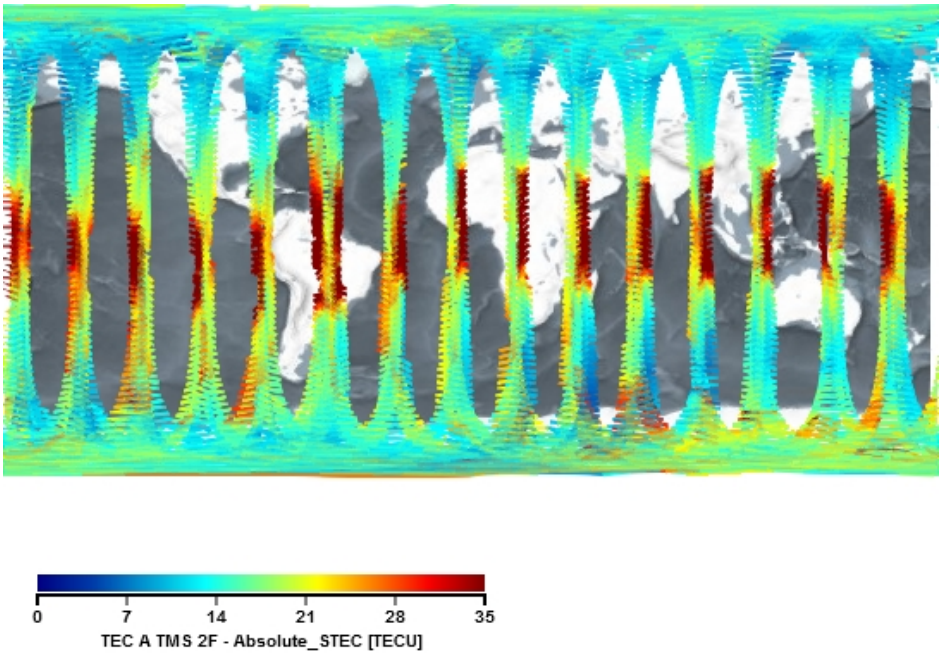


Figure 3.10: Swarm-A STEC measured on 12 March 2015 from on-board GNSS receiver. Data Courtesy: Swarm project office, ESA.

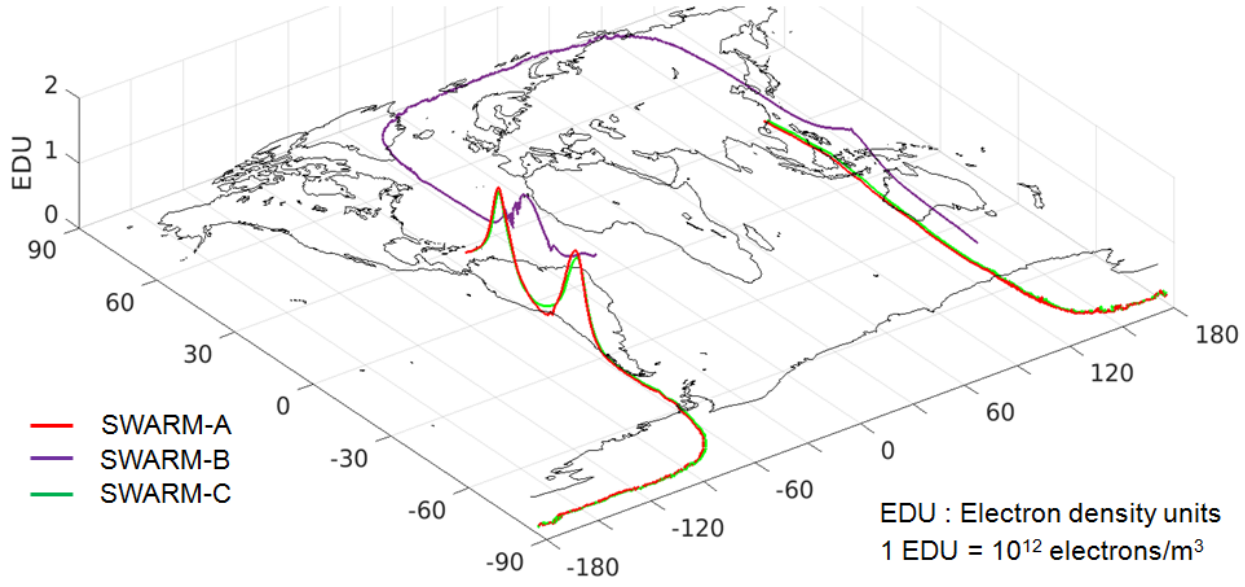


Figure 3.11: Electron density from in-situ Langmuir probe onboard Swarm satellite. Data source: ESA Earth observation data portal and Swarm project office.

4 Optimization techniques

Optimization methods are applied in situations when a certain function is to be minimized (e.g., mean square error) or maximized (e.g., likelihood function in probability, profit in business) with respect to the underlying parameters. If additional prior information about some or all of these parameters is available, then those can be expressed as upper or lower bounds or both (e.g., a non-negativity of physical parameters such as number of atoms, density) and then it is called an optimization problem with inequality constraints. Although such methods are commonly used in financial and engineering disciplines, their application to geodesy is relatively rare. Within this chapter, a detailed description of inequality constraints will be provided. As and when possible, analogies to physics are also presented, so that the mathematical equations behind constraint processing can be visualized in a practical sense. We will demonstrate example problems involving inequality constraints using graphical illustration and visualization before applying them to multi-dimensional parameters. Although inequality constraints are the primary focus in this chapter, we will show a more general formulation that includes equality constraints as well. Furthermore, few challenges are identified which currently limit the use of constraint optimization to real time geodetic applications (e.g. ionosphere modelling, orbit determination). These challenges are related to the search techniques used within the optimization algorithms and will be substantiated in this chapter. We also propose solutions for the identified challenges which will be eventually applied to ionospheric electron density modelling in the subsequent chapter.

This chapter is organized as follows: Section 4.1 introduces the Gauss-Markov model (GMM) for observation modelling and variance component estimation (VCE). Section 4.2 describes the fundamentals of optimization problems and begins with the definition of objective function and model parameter. It is followed by Section 4.3 that discusses the so-called "descent methods" along with the two fundamental algorithms, namely, gradient descent and the Newton's method. Section 4.4 introduces the inequality constraint formulation and inclusion in optimization problems. Section 4.5 describes the optimality conditions and their application to the Lagrangian function as well as the solution of the system of equation that follows. Section 4.6 discusses the impact of changing constraint functions on the model parameter in the form of sensitivity analysis. Section 4.7 discusses alternate methods to solve the same equations which were derived and solved in section 4.5. Wherever possible, a geometrical (or graphical) interpretation will be provided in order to compliment the algebraic results. In selected sections, important equations will be derived from fundamental concepts. These derivations are helpful in two particular ways: (1) they provide insights on the assumptions (if any) behind those equations and (2) they allow a clear and elegant interpretation of the estimated parameters. In the summary of the chapter, some best practices used during the implementation are highlighted. The focus of this thesis will be on the optimization of a quadratic objective function with linear equality and inequality constraints.

4.1 Observation modelling

Non-linear observation modelling is a thoroughly researched subject and Koch (1999), Rao and Toutenburg (1999), Seber (2015), discuss it in detail. Measurements are generally non-linear functions of the unknown parameters and the task of observation modelling is to establish this approximate functional relation. Let $\boldsymbol{y} = [y_1 \cdots y_n]^T$ be $n \times 1$ observation vector with $\boldsymbol{e} = [e_1 \cdots e_n]^T$ the associated $n \times 1$ measurement error vector and $\boldsymbol{\beta} = [\beta_1 \cdots \beta_u]^T$ the $u \times 1$ vector of unknown

parameters, then the non-linear relation f_i between each y_i, e_i and $\boldsymbol{\beta}$ can be written as

$$\begin{aligned} f_1(\beta_1, \cdots, \beta_u) &= y_1 + e_1 \\ f_2(\beta_1, \cdots, \beta_u) &= y_2 + e_2 \\ &\vdots \\ f_n(\beta_1, \cdots, \beta_u) &= y_n + e_n \end{aligned}, \tag{4.1}$$

where $f_i(\beta_1, \cdots, \beta_u)$ for $i = \{1, \cdots, n\}$ are differentiable non-linear real valued functions of the unknown parameters. The system of equations (4.1) can be linearized to a first order[1] using the Taylor series expansion about the initial value vector $\boldsymbol{\beta}_0 = [\beta_{1,0} \cdots \beta_{u,0}]^T$ to give

$$\begin{aligned} f_i(\boldsymbol{\beta}) &= f_i(\boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta}) \\ &\approx f_i(\boldsymbol{\beta}_0) + \left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_1}\right|_{\boldsymbol{\beta}_0} \Delta\beta_1 + \cdots + \left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_u}\right|_{\boldsymbol{\beta}_0} \Delta\beta_u \\ &\approx f_i(\boldsymbol{\beta}_0) + \left[\left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_1}\right|_{\boldsymbol{\beta}_0} \cdots \left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_u}\right|_{\boldsymbol{\beta}_0}\right] \cdot \Delta\boldsymbol{\beta} \\ &\approx f_i(\boldsymbol{\beta}_0) + \left(\nabla_{\boldsymbol{\beta}} f_i(\boldsymbol{\beta}_0)\right)^T \cdot \Delta\boldsymbol{\beta} \end{aligned}, \tag{4.2}$$

(see Koch (1999)) where $\Delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$, the operator $\nabla_{\boldsymbol{\beta}}$ denotes the gradient with respect to the elements of $\boldsymbol{\beta}$ and

$$\nabla_{\boldsymbol{\beta}} f_i(\boldsymbol{\beta}_0) = \left[\left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_1}\right|_{\boldsymbol{\beta}_0} \cdots \left.\frac{\partial f_i(\boldsymbol{\beta})}{\partial \beta_u}\right|_{\boldsymbol{\beta}_0}\right]^T \tag{4.3}$$

denotes the partial derivatives of the non-linear function f_i with respect to elements of $\boldsymbol{\beta}$. The truncation error involved in Eq. (4.2) is denoted implicitly. Eq. (4.2) can be rearranged as

$$\left(\nabla_{\boldsymbol{\beta}} f_i(\boldsymbol{\beta}_0)\right)^T \cdot \Delta\boldsymbol{\beta} = f_i(\boldsymbol{\beta}) - f_i(\boldsymbol{\beta}_0) \tag{4.4}$$

and collectively in matrix notation

$$\begin{bmatrix} \Delta f_1(\boldsymbol{\beta}) \\ \vdots \\ \Delta f_n(\boldsymbol{\beta}) \end{bmatrix} = \boldsymbol{A}\Delta\boldsymbol{\beta} \tag{4.5}$$

with

$$\Delta f_i(\boldsymbol{\beta}) = f_i(\boldsymbol{\beta}) - f_i(\boldsymbol{\beta}_0) \quad \text{and} \quad \boldsymbol{A} = \begin{pmatrix} \left.\frac{\partial f_1(\boldsymbol{\beta})}{\partial \beta_1}\right|_{\boldsymbol{\beta}_0} & \cdots & \left.\frac{\partial f_1(\boldsymbol{\beta})}{\partial \beta_u}\right|_{\boldsymbol{\beta}_0} \\ \left.\frac{\partial f_2(\boldsymbol{\beta})}{\partial \beta_1}\right|_{\boldsymbol{\beta}_0} & \cdots & \left.\frac{\partial f_2(\boldsymbol{\beta})}{\partial \beta_u}\right|_{\boldsymbol{\beta}_0} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial f_n(\boldsymbol{\beta})}{\partial \beta_1}\right|_{\boldsymbol{\beta}_0} & \cdots & \left.\frac{\partial f_n(\boldsymbol{\beta})}{\partial \beta_u}\right|_{\boldsymbol{\beta}_0} \end{pmatrix} \tag{4.6}$$

being the matrix of first derivatives, called the system design matrix (or the design matrix), evaluated at $\boldsymbol{\beta}_0$. The unknown parameter vector is iteratively computed by the so-called "update rule"

$$\boldsymbol{\beta}_{it+1} = \boldsymbol{\beta}_{it} + \Delta\widehat{\boldsymbol{\beta}}_{it}, \tag{4.7}$$

where $\Delta\widehat{\boldsymbol{\beta}}_{it}$ is an estimate of $\Delta\boldsymbol{\beta}$ from solving Eq. (4.5), at an iteration it. When $|\boldsymbol{\beta}_{it+1} - \boldsymbol{\beta}_{it}|$ becomes smaller than or equal to a predefined threshold[2], a convergence is said to be reached.

[1] The truncation error is omitted for better readability and accordingly the = symbol is replaced by ≈ in Eq. (4.2)
[2] In this work, it is chosen as 10^{-8}.

4.1.1 Gauss-Markov model

Following Koch (1999), without loss of generality[3], Eq. (4.5) will be redefined as the Gauss-Markov model (GMM)

$$\boldsymbol{y} + \boldsymbol{e} = \boldsymbol{A}\boldsymbol{\beta} \text{ with } D(\boldsymbol{y}) = \sigma^2 \boldsymbol{P}^{-1} \tag{4.8}$$

consisting of two parts, namely, the deterministic and the stochastic part, where $\boldsymbol{A}$ is the $n \times u$ design matrix, $\boldsymbol{\beta}$ is the $u \times 1$ vector of unknown parameters, $\boldsymbol{y}$ is the $n \times 1$ vector of observations, $\boldsymbol{e}$ is the $n \times 1$ vector of observation errors, D is the dispersion operator, $\boldsymbol{P}$ is a given positive definite observation weight matrix and σ^2 is the unknown variance factor. The linear relation in GMM (4.8) follows generally after a linearization process, as shown in Eq. (4.2).

Method of least squares

Since the main focus of this chapter is the optimization technique, the method of least squares is only briefly summarized. The subject of least squares has references dating back to the treatise written (in Latin) by Gauss[4] in 1820s (English translation is available in Stewart (1995)). For a detailed treatment of the linear, non-linear least squares, weighted and total least squares, the reader is referred to Osborne (1965), Gentleman (1976), White (1990), Schuh (1996), Koch (1999), Rao and Toutenburg (1999), Schuh (2003), Alkhatib and Schuh (2007), Moritz (2015).
From the GMM (4.8), the quadratic function

$$J(\boldsymbol{\beta}) = \frac{1}{\sigma^2} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\beta})^T \boldsymbol{P} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\beta}) \tag{4.9}$$

is minimized by setting its derivative with respect to $\boldsymbol{\beta}$ to zero,

$$\frac{1}{\sigma^2}(-2\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{y} + 2\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}\boldsymbol{\beta}) = \boldsymbol{0}, \tag{4.10}$$

leading to the normal equations

$$\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}\boldsymbol{\beta} = \boldsymbol{A}^T\boldsymbol{P}\boldsymbol{y} \tag{4.11}$$

from which an estimate $\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{y}$ of $\boldsymbol{\beta}$ along with its covariance

$$D(\widehat{\boldsymbol{\beta}}) = \sigma^2 \left(\boldsymbol{A}^T\boldsymbol{P}\boldsymbol{A}\right)^{-1} \tag{4.12}$$

can be obtained, where σ^2 is the unknown variance factor, assuming the matrix $\boldsymbol{A}$ is of full column rank (Koch (1999)). An estimate

$$\widehat{\sigma}^2 = \frac{\widehat{\boldsymbol{e}}^T\boldsymbol{P}\widehat{\boldsymbol{e}}}{n-u} \tag{4.13}$$

of σ^2 is obtained from the maximum likelihood method when $n > u$, where n and u are the number of observations and unknown parameters respectively and $\widehat{\boldsymbol{e}} = \boldsymbol{y} - \boldsymbol{A}\widehat{\boldsymbol{\beta}}$ is the vector of residuals.

4.1.2 Observation combination and variance component estimation

When n_{typ} different observation types are combined, a GMM

$$\boldsymbol{y}_o + \boldsymbol{e}_o = \boldsymbol{A}_o\boldsymbol{\beta} \quad \text{with } D(\boldsymbol{y}_o) = \sigma_o^2 \boldsymbol{P}_o^{-1} \tag{4.14}$$

[3] To avoid confusion, it is clarified that this change in convention from the previous section is for the ease of readability. The remaining parts of this chapter have additional Δ and ∇ operators applied to the parameter $\boldsymbol{\beta}$. So the idea is to start with a convention where $\boldsymbol{\beta}$ is already linearly related to $\boldsymbol{y}$.

[4] Johann Carl Friedrich Gauss was a German mathematician and physicist who made significant contributions to many fields in mathematics and science.

is established individually using the variance components σ_o^2 for each type denoted by the subscript $o \in \{1, \cdots, n_{typ}\}$. The prior information $\boldsymbol{\mu} = E\{\boldsymbol{\beta}\}$ is introduced as additional observations

$$\boldsymbol{\beta} = \boldsymbol{\mu} + \boldsymbol{e}_\mu \text{ with } D(\boldsymbol{\mu}) = \sigma_\mu^2 \boldsymbol{P}_\mu^{-1} \tag{4.15}$$

with the error vector $\boldsymbol{e}_\mu$, the unknown variance factor σ_μ^2 and a given positive definite weight matrix P_μ. If the different observations $\boldsymbol{y}_o$ and the prior information $\boldsymbol{\mu}$ are assumed to be independent, then following Koch and Kusche (2002), the combination of $n_{typ} + 1$ observation types yields the extended linear model

$$\begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_{typ}} \\ \boldsymbol{\mu} \end{bmatrix} + \begin{bmatrix} \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_{n_{typ}} \\ \boldsymbol{e}_\mu \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 \\ \vdots \\ \boldsymbol{A}_{n_{typ}} \\ \boldsymbol{I} \end{bmatrix} \boldsymbol{\beta} \tag{4.16}$$

with

$$D\left(\begin{bmatrix} \boldsymbol{y}_1 \\ \vdots \\ \boldsymbol{y}_{n_{typ}} \\ \boldsymbol{\mu} \end{bmatrix}\right) = \sigma_1^2 \boldsymbol{M}_1 + \cdots + \sigma_{n_{typ}}^2 \boldsymbol{M}_{n_{typ}} + \sigma_\mu^2 \boldsymbol{M}_\mu \tag{4.17}$$

where[5],

$$\begin{aligned} \boldsymbol{M}_1 &= \begin{bmatrix} \boldsymbol{P}_1^{-1} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}, \cdots, \boldsymbol{M}_{n_{typ}} = \begin{bmatrix} \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{P}_{n_{typ}}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \\ \boldsymbol{M}_\mu &= \begin{bmatrix} \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \ddots & \vdots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{P}_\mu^{-1} \end{bmatrix}. \end{aligned} \tag{4.18}$$

The normal equations from each observation type are combined to give

$$\left(\sum_{o=1}^{n_{typ}} \frac{1}{\sigma_o^2} \boldsymbol{A}_o^T \boldsymbol{P}_o \boldsymbol{A}_o + \frac{1}{\sigma_\mu^2} \boldsymbol{P}_\mu\right) \boldsymbol{\beta} = \sum_{o=1}^{n_{typ}} \frac{1}{\sigma_o^2} \boldsymbol{A}_o^T \boldsymbol{P}_o \boldsymbol{y}_o + \frac{1}{\sigma_\mu^2} \boldsymbol{P}_\mu \boldsymbol{\mu}. \tag{4.19}$$

When the different observation types are dependent, $D(\boldsymbol{y})$ is no longer a diagonal block matrix and unknown covariance matrices between different types of observations have to be taken into account and estimated as well (Koch (1999)). Since the weight matrices $\boldsymbol{P}_o$ are positive definite, the matrices $\boldsymbol{A}_o^T \boldsymbol{P}_o \boldsymbol{A}_o$ are at least positive semi-definite. Furthermore, the weight matrix $\boldsymbol{P}_\mu$ of the prior information is positive definite and therefore, the matrix of normal equations

$$\boldsymbol{N}_c = \frac{1}{\sigma_1^2} \boldsymbol{A}_1^T \boldsymbol{P}_1 \boldsymbol{A}_1 + \cdots + \frac{1}{\sigma_{n_{typ}}^2} \boldsymbol{A}_{n_{typ}}^T \boldsymbol{P}_{n_{typ}} \boldsymbol{A}_{n_{typ}} + \frac{1}{\sigma_\mu^2} \boldsymbol{P}_\mu \tag{4.20}$$

is invertible, when appropriate variance components σ_o^2, for $o = \{1, \cdots, n_{typ}\}$, and σ_μ^2 are chosen. The variance components can be chosen manually or estimated within the variance component estimation (VCE) procedure. Among a number of methods to estimate variance components (see Brockmann and Schuh (2010), Xu et al. (2006), and references therein), the method of Koch and Kusche (2002) is summarized below, which has also been used by Limberger (2015), Liang (2017), Erdogan et al. (2017), Goss et al. (2019), Goss et al. (2020), Erdogan et al. (2020), Liu et al. (2020). An estimate $\widehat{\sigma}_o^2$ of σ_o^2 and $\widehat{\sigma}_\mu^2$ of σ_μ^2 is obtained through an iterative procedure as follows:

$$\widehat{\sigma}_o^2 = \frac{\widehat{\boldsymbol{e}}_o^T \boldsymbol{P}_o \widehat{\boldsymbol{e}}_o}{r_o}, \text{ with } o \in \{1, \cdots, n_{typ}\}, \text{ and } \widehat{\boldsymbol{e}}_o = \boldsymbol{A}_o \widehat{\boldsymbol{\beta}} - \boldsymbol{y}_o \tag{4.21}$$

[5] In this case, $\boldsymbol{0}$ denotes a matrix with dimensions corresponding to $\boldsymbol{P}_1$, $\boldsymbol{P}_{n_{typ}}$ and $\boldsymbol{P}_\mu$ respectively. For improved readability, their dimension are not shown in subscript. Later in this chapter, such a convention will be followed.

and

$$\widehat{\sigma}_\mu^2 = \frac{\widehat{\boldsymbol{e}}_\mu^T \boldsymbol{P}_\mu \widehat{\boldsymbol{e}}_\mu}{r_\mu}, \text{ with } \widehat{\boldsymbol{e}}_\mu = \widehat{\boldsymbol{\beta}} - \boldsymbol{\mu}, \tag{4.22}$$

where $\widehat{\boldsymbol{e}}_0$ and $\widehat{\boldsymbol{e}}_\mu$ denote the residual vector corresponding to the observation type index $o \in \{1, 2, \cdots, n_{typ}\}$ and the prior information respectively. The partial redundancies (Koch (1999), Xu et al. (2006))

$$r_o = n_o - Tr(\frac{1}{\sigma_o^2} \boldsymbol{A}_o^T \boldsymbol{P}_o \boldsymbol{A}_o \boldsymbol{N}_c^{-1}) \tag{4.23a}$$

$$r_\mu = n_\mu - Tr(\frac{1}{\sigma_\mu^2} \boldsymbol{P}_\mu \boldsymbol{N}_c^{-1}) \tag{4.23b}$$

are the contributions of $\boldsymbol{y}_o$ and $\boldsymbol{\mu}$ to the total redundancy

$$r_t = \sum_{o=1}^{n_{typ}} r_o + r_\mu + r_c \tag{4.24}$$

of the GMM, where the symbol Tr denotes the trace of a matrix, n_o and n_μ refer to the total number of observations of the type "o" and the prior information, respectively. If there are no constraints in the GMM, then the number of constraints r_c can be omitted from Eq. (4.24). As can be seen from the Eqs. (4.21) to (4.23b), the estimated variance components are depending on the observation residuals computed from $\widehat{\boldsymbol{\beta}}$, which in turn depends on the variance components. Thus, the estimation of the variance components is performed iteratively. A stochastic trace estimator (Hutchinson (1990), Koch and Kusche (2002)) is used as alternate to avoid memory expensive inversion of normal equations in large dimensional problems.

The least squares method is one of the most commonly used technique for parameter estimation but when there are constraints on the unknown parameters, an optimization approach is more suitable, especially if those are inequalities. Equality constraints can be included alongside the least squares solution as described in Chapter 3 of Koch (1999). Inequality constraints, on the other hand, are relatively difficult to handle and advanced techniques are required (see Chapter 16 in Nocedal and Wright (2006)). Especially when the unknown parameters are highly correlated, the estimation accuracy can be improved by including prior information in the form of exact values of some of the parameters or approximate lower and upper bounds for others. Although there are software packages available based on the work of Lötstedt (1984), Gill et al. (1984), Coleman and Li (1996) and Mead and Renaut (2010) but their algorithmic details, assumptions and implementation along with the standard deviations are either not available or not explicitly transparent to the user. In many cases the algorithm are also not easily adaptable for large dimensional problems (e.g. global 4D electron density modelling). This is the main motivation to develop a constraint optimization approach for parameter estimation.

4.2 Introduction to optimization technique

Optimization is an important mathematical method used in modelling and analysis of physical systems (Nocedal and Wright (2006) pg. 2). According to Goldstine (1980), "*the mathematical formalism in optimization arrived with the development of variational calculus*[6]". In modern era, optimization is applied in science, engineering, finance, transportation, healthcare etc. and also many references are available such as Philip E. Gill and Wright (1981), Luenberger (1999), Nesterov et al. (2000), Boyd and Vandenberghe (2004), Nocedal and Wright (2006) and Bertsekas (2009). It is too detailed a subject to be completely described in this thesis and therefore, after a brief introduction, the focus will be on the specific aspect of inequality constrained optimization problems and their solution.

[6] Leonhard Euler is credited with the development of variational calculus. An excellent history on this topic is available in Koetsier (2007).

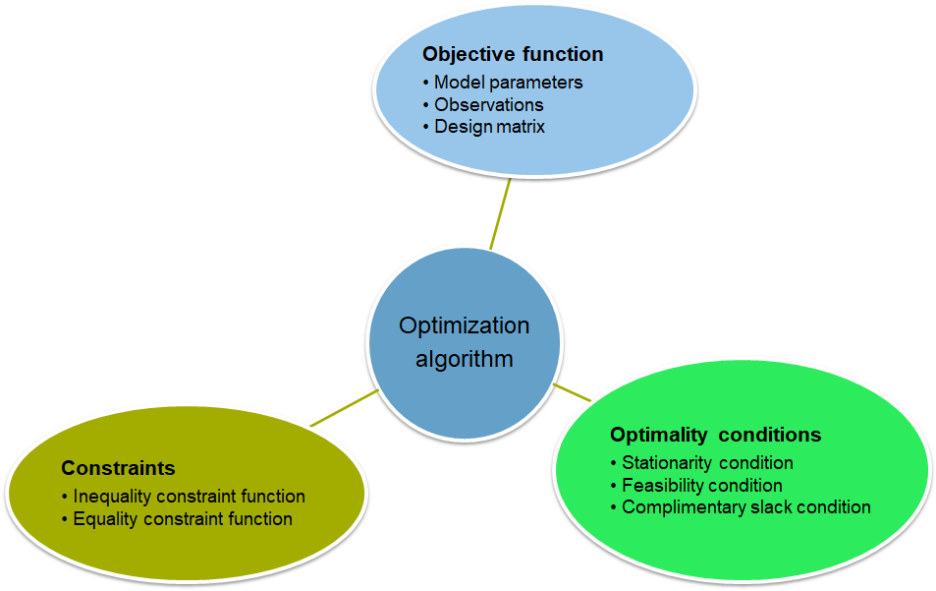


Figure 4.1: Main building blocks in optimization

The process of solving an inequality constrained optimization problem is divided into two sub-problems:

- Sub-problem 1: Feasibility, i.e., determining a solution that fulfills all constraints.
- Sub-problem 2: Optimality, i.e., determining the solution that minimizes the objective function.

These two sub-problems will be described in detail for a given set of objective and constraint functions. As shown in Fig. 4.1, there are four important building blocks in the optimization namely (1) the objective function along with the model parameter(s), (2) the constraints associated with the model parameters, (3) the optimality conditions and (4) optimization algorithm. The primary goal of this chapter, is to describe these items in detail. The objective function and model parameters are discussed in Section 4.3, constraints in Section 4.4, optimality conditions in Section 4.5. Optimization algorithms are discussed in two separate parts. The part covering the fundamentals common to most algorithms is discussed in Section 4.3 and more specific aspects in Section 4.5.2.

Optimization problem

An optimization problem $\mathcal{P}$

$$\begin{aligned} &\min_{\boldsymbol{\beta}} J(\boldsymbol{\beta}) \\ \text{subject to constraints}: \; &c_i(\boldsymbol{\beta}) = 0 \quad i \in \mathcal{E} \\ &c_i(\boldsymbol{\beta}) \leq 0 \quad i \in \mathcal{I} \end{aligned} \tag{4.25}$$

is defined with respect to the unknown model parameter vector $\boldsymbol{\beta}$, the objective function $J(\boldsymbol{\beta})$ and i^{th} constraint function $c_i(\boldsymbol{\beta})$ belonging to the sets $\mathcal{E}, \mathcal{I}$ for the equality and inequality constraints, respectively (see Fig. 4.2).

4.2.1 Objective function and model parameters

The first step towards the solution of an optimization problem is the definition of an objective function. It is an empirical function of the unknown model parameter that gives a quantitative measure of the optimization performance. Goal of an optimization problem is to estimate the

unknown model parameters, such that the objective function is optimized[7]. Specifically, it means the objective function is either minimized or maximized.

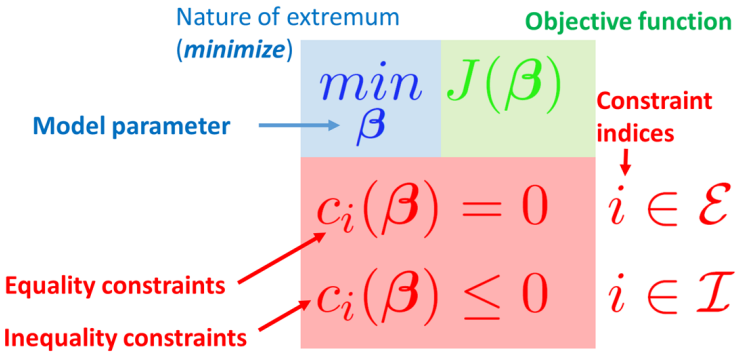


Figure 4.2: Optimization problem definition

Definition 4.2.1. *Objective function and model parameters*
For the minimization problem (4.25), the quadratic objective function (4.9) is used in this work, where $\boldsymbol{y}$, $\boldsymbol{A}$, $\boldsymbol{\beta}$ and $\boldsymbol{P}$ are defined in the Gauss-Markov model (4.8).
The objective function[8] $J(\boldsymbol{\beta})$ expresses how well the model parameter $\boldsymbol{\beta}$ fits the observations vector $\boldsymbol{y}$ in the Gauss-Markov model since it is the square of the errors. For a minimization problem, a smaller[9] value of the objective function indicates that the model parameters used in the GMM fits the observations well

Feasible region

Definition 4.2.2. The feasible region $\mathcal{F}_R$ for an optimization problem is the set of all possible and permissible solutions for the model parameters over which the objective function is to be optimized.

Optimal solution

Definition 4.2.3. The optimal solution to a problem is given by the model parameter corresponding to a minimum[a] value of the objective function over the feasible region. $\boldsymbol{\beta}_{opt}$ is an optimal solution to the problem $\mathcal{P}$ if

$$\boldsymbol{\beta}_{opt} \in \mathcal{F}_R \quad \text{and} \quad J(\boldsymbol{\beta}_{opt}) \leq J(\boldsymbol{\beta}) \ \forall \boldsymbol{\beta} \in \mathcal{F}_R. \tag{4.26}$$

[a] In this work the goal is to minimize the objective function but in general, objective functions can also be maximized.

In other words, for a minimization problem, the optimal solution corresponds to the model parameter $\boldsymbol{\beta}$ which results in the minimum value of the objective function over the entire feasible region. A feasible solution is any value of the model parameter satisfying $\boldsymbol{\beta} \in \mathcal{F}_R$. Examples of constraint and feasible region in one- and two- dimensions are shown in Figures 4.14 and 4.15 respectively. If a feasible region is empty, i.e. $\mathcal{F}_R = \emptyset$, where $\emptyset$ denotes an empty set, then $\mathcal{P}$ is called an infeasible problem. An important subset of optimization algorithms are exclusively dedicated for convex objective functions. Before going into further details, the terms convex set and functions are defined.

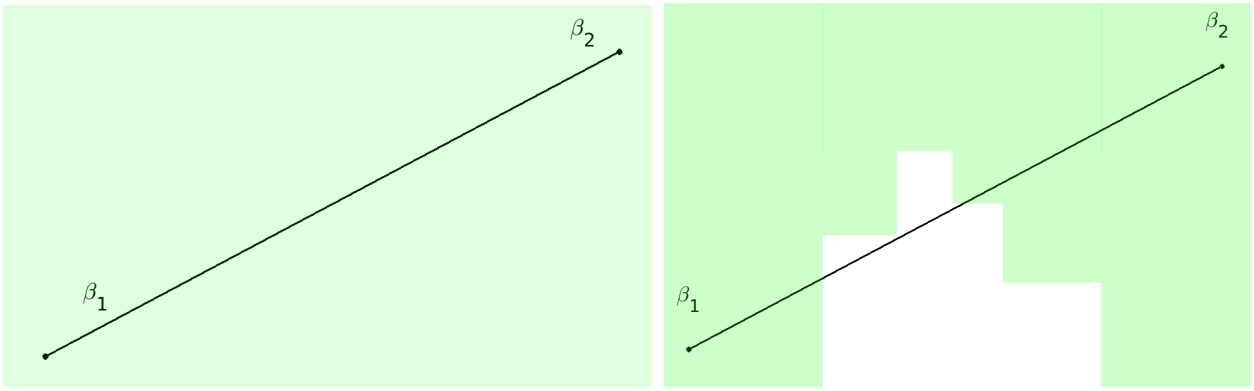
[7] For the scope of this thesis, the objective function is minimized unless otherwise specified.
[8] In some literature the objective is also called the loss or cost function.
[9] Here, it is only qualitatively defined.

Convex set

A set $\mathbb{C} \subseteq \mathbb{R}^u$ is convex if and only if

$$a_1\boldsymbol{\beta}_1 + (1 - a_1)\boldsymbol{\beta}_2 \subseteq \mathbb{C} \tag{4.27}$$

holds for any $a_1 \in [0, 1]$ such that $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{C}$ and $\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2$ (see Boyd and Vandenberghe (2004) p. 23). Geometrically, Eq. (4.27) means that the line segment, connecting $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, is entirely contained within the set $\mathbb{C}$, as shown in Fig. 4.3a. Conversely, a non-convex set is shown in see Fig. 4.3b.



(a) Convex set **(b)** Non-convex set

Figure 4.3: Example of convex and non-convex sets. Notice that the line segment on the left side is entirely contained within the green area whereas on the right it is not.

Convex function

A function that maps a non-empty convex set $\mathbb{C}$ from $\mathbb{R}^u$ to $\mathbb{R}$ is convex if and only if

$$\begin{gathered} f\left(a_1\boldsymbol{\beta}_1 + (1 - a_1)\boldsymbol{\beta}_2\right) \leq a_1 f\left(\boldsymbol{\beta}_1\right) + (1 - a_1) f\left(\boldsymbol{\beta}_2\right) \\ 0 \leq a_1 \leq 1, \ \forall \ \boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{C} \end{gathered} \tag{4.28}$$

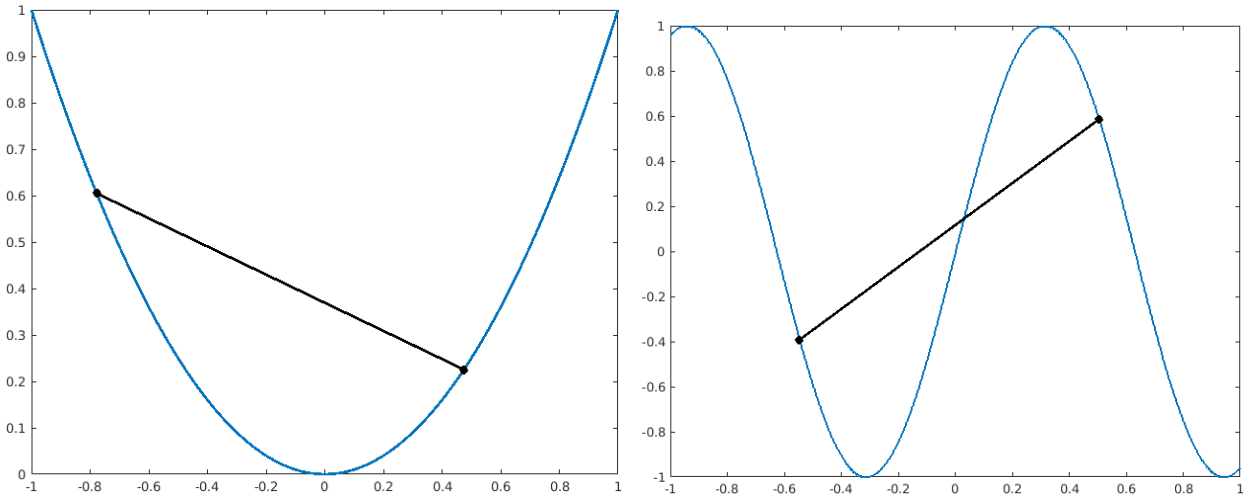
holds (see Boyd and Vandenberghe (2004) p. 67). Geometrically, Eq. (4.28) means that the line segment joining $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ is entirely contained "within" the graph of the function.

Although the phrase "within" is somewhat subjective, but specifically for the one-dimensional case, the concept is shown in Fig. 4.4a. By extension of Eq. (4.28), any linear function (e.g., a line or a plane) is also a convex function (Nocedal and Wright (2006) pg. 7).

Generally, optimization algorithms are used on a multi-dimensional space of the model parameters and they work on the principle of "*search and selection*". It means the model parameters are estimated iteratively by "search" and moving along a certain path of intermediate solutions. This point will be substantiated further in section 4.3.1.

When an objective function cannot be minimized any more without a violation of the constraints, a final solution is reached. It follows that when an objective function is optimized, the corresponding unknown parameters are referred to as the optimal solution.

This was a short introduction and the remaining part of this chapter will describe each of these specific aspects of the optimization problem and solution in detail.



(a) Convex function **(b)** Non-convex function

Figure 4.4: Example of convex and non-convex functions

4.2.2 Taxonomy of optimization

In the previous section, it has been mentioned that the specification of a constraint optimization problem involves defining the objective function, the model parameters and the constraints. For the scope of this thesis, GMM (4.8) is reformulated as an objective function (4.9). The optimization parameters are then the GMM model parameters and additional unknowns corresponding to the constraints. This aspect will be described in Section 4.4.2. Based on these three aspects, a brief taxonomy of optimization is presented.

Taxonomy refers to the different categorizations under which a given optimization problem may be classified (see Table 4.1). On the one hand it is important to identify the specific category to which a given problem belongs, so that an appropriate algorithm can be used for its solution. On the other hand, generating a complete taxonomy of all the known optimization problems is a difficult task because of two main reasons:

- A given problem may belong to two or more categories. For example a non-linear problem (NLP) can have either a single or multi-objective function. So there will be many interconnections between the different categories of the optimization problems. This makes the overall categorization complicated.
- Secondly, the subject of optimization is changing dynamically with newer developments, e.g. Salcedo-Sanz (2017), Rao and Waghmare (2017), where improvements are proposed to either existing optimization algorithms or their implicit assumptions. In both of these cases, the existing categorization changes.

Continuous and discrete optimization : In continuous optimization, the model parameters are allowed to take on any value within a range, in contrast to discrete optimization, where some or all of the variables may take only discrete values. Examples of discrete optimization are when model parameters are binary, integers, prime number etc.

Unconstrained and constrained optimizations : Constraints are empirical functions of the unknown optimization parameters. In an unconstrained optimiztaion, as the name suggests, there are no constraints on the model parameter. As detailed in this chapter, the solution approach with constraints is to augment them as additional penalty terms to the objective function.

Single- and multi-objective optimization : Most optimization problems have a single objective function but in some cases there is more than one objective function. The optimization parameters could be shared between the different objective functions, see e.g. Marler and Arora (2004). In

practice, a weighted combination of multiple objective functions is used to form a single objective function.

Table 4.1: Optimization Taxonomy

Classification based on the problem degree	LP, QP
Classification based on the number of objectives	single and multi-objective
Classification based on the constraints	unconstrained, constrained
Classification based on the optimization parameters	continuous, discrete

For a more detailed taxonomy along with a catalogue of the different algorithms, see Gropp and Moré (1997), Czyzyk et al. (1998), and Dolan (2001).

4.3 Structure of optimization algorithms

4.3.1 Fundamentals of descent method

Before describing an optimization algorithm in detail, the so-called "descent method" is introduced. It helps in understanding the common basic structure of the different optimization algorithms and also their specific weaknesses, which then lead to the need for the advanced optimization algorithms. As the name indicates, there is a tendency of the objective function to reduce[10], from an initial value, towards its minimum. The model parameter corresponding to the minimum value of the objective function are then the optimal solution (Boyd and Vandenberghe (2004) pg. 463).

Let $J(\boldsymbol{\beta}_0)$ and $J(\boldsymbol{\beta}_{it})$ denote the objective function at an initial value $\boldsymbol{\beta}_0$ and that at any iteration it, denoted as $\boldsymbol{\beta}_{it}$ respectively. The progressive reduction (or descent) in the objective function for iterations $it = \{1, 2, \cdots, p, \cdots it, \cdots\}$, generates a sequence

$$J(\boldsymbol{\beta}_0) \cdots J(\boldsymbol{\beta}_p) \cdots J(\boldsymbol{\beta}_{it}) \cdots$$

such that, it represents a decreasing trend. It shall be noted that, although desirable, it is not certain that the objective function will always decrease from one iteration to the next. In practice, there may be situations when the objective function remains nearly constant or even increases. The model parameters are updated as

$$\boldsymbol{\beta}_{it} = \boldsymbol{\beta}_{it-1} + t_{it-1} \Delta\boldsymbol{\beta}_{it-1}, \tag{4.29}$$

where t_{it} and $\Delta\boldsymbol{\beta}_{it}$ denote the step size and the descent direction respectively, at an iteration it. The terms in the update rule (4.29) are illustrated in Fig. 4.5. It shall be noted that although the step size is shown as a scalar in the update rule (4.29), in a more general case the individual components of the model parameter are updated using different step sizes. In such a case, the step size will be denoted by a diagonal matrix instead. This aspect will be revisited in Section 4.5.3. The descent direction

$$\Delta\boldsymbol{\beta}_{it-1} = \nu_{\Delta\boldsymbol{\beta}_{it-1}} \, \boldsymbol{u}_{\Delta\boldsymbol{\beta}_{it-1}} \tag{4.30}$$

can be decomposed into its magnitude

$$\nu_{\Delta\boldsymbol{\beta}_{it-1}} = \|\Delta\boldsymbol{\beta}_{it-1}\| \tag{4.31}$$

and the unit vector

$$\boldsymbol{u}_{\Delta\boldsymbol{\beta}_{it-1}} = \Delta\boldsymbol{\beta}_{it-1} / \|\Delta\boldsymbol{\beta}_{it-1}\|, \tag{4.32}$$

[10] and therefore the name "descend"

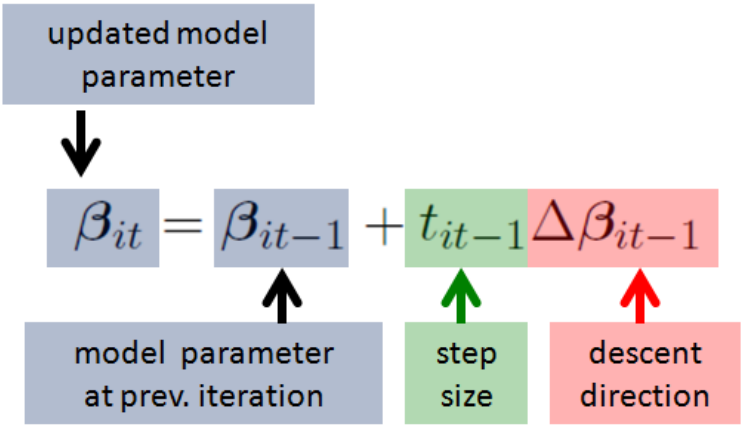


Figure 4.5: Update rule in descent method. The updated model parameter at an iteration it is computed using three terms namely, the model parameter, the step size and the descent direction at the previous iteration. More generally, different step sizes can be chosen for updating the model parameter, in which case, t_{it-1} will be replaced by a diagonal matrix. It will be shown later in section 4.5.3.

where $\|\cdot\|$ denotes the L_2 norm $\left(\|\boldsymbol{\beta}\| = \sqrt{\beta_1^2 + \cdots + \beta_u^2}\right)$. From a geometrical interpretation, Eq. (4.29) is the equation of the line segment $\boldsymbol{\beta}_{it}$ reached by traversing from its initial value $\boldsymbol{\beta}_{it-1}$ along the direction $\boldsymbol{u}_{\Delta\beta_{it-1}}$ by a magnitude of $t_{it-1}\nu_{\Delta\beta_{it-1}}$. The step size is an important parameter, with $0 \leq t_{it} \leq 1$ (see Boyd and Vandenberghe (2004), Nocedal and Wright (2006)), that controls by how much (in magnitude) the model parameter $\boldsymbol{\beta}$ shall be updated along the descent direction. For $t_{it} \approx 0$, the model parameter $\boldsymbol{\beta}_{it}$ is updated by an negligible magnitude and analogously, for $t_{it} = 1$, the model parameter $\boldsymbol{\beta}_{it}$ is updated by the "full" magnitude of the descent direction, in which case, it is also called "taking a full-step" along the descent direction. Convergence is reached and the optimization algorithm is said to have found $\boldsymbol{\beta}_{opt} = \boldsymbol{\beta}_{it}$, the optimal estimate of the model parameter $\boldsymbol{\beta}$ at an iteration it, when the condition

$$\|\boldsymbol{\beta}_{it} - \boldsymbol{\beta}_{it-1}\| \leq \eta_{pre} \tag{4.33}$$

is satisfied, where η_{pre} is a positive precision threshold[11].

Direction of decreasing gradient

Starting from an initial value $\boldsymbol{\beta}_0$ to the optimal solution $\boldsymbol{\beta}_{opt}$, the process of computing the successive intermediate solutions using the update rule (4.29) is called "traversal". Within the descent method, the direction of decreasing gradient of the objective function is determined at each iteration for traversing towards the optimal solution as shown in Fig. 4.6 and more specifically illustrated in the Fig. 4.7 using a two dimensional example with the axes labelled β_1 and β_2 as well as the contours of equal objective function values (iso-objective contours) depicted on the right side. The contours with larger radii have larger objective function values.

The example 2D objective function $J(\boldsymbol{\beta}) = \boldsymbol{\beta}^T\boldsymbol{\beta}$ shown in the Fig. 4.7 with convergence from an initial value $\boldsymbol{\beta}_0 = [10 \quad 10]^T$ to the final (optimal) value of $\boldsymbol{\beta} = [0 \quad 0]^T$. It can be seen in Fig. 4.8 that the objective function descent is steeper[12] for a larger step size or the number of iterations required is lower to achieve a desired precision. For e.g, the objective function reaches 10^{-11} after 30 iterations for a step size of 0.2, whereas for relatively smaller step sizes of 0.1 and 0.05, it takes 65 and 110 iterations respectively to reach the same level of precision. The objective function chosen in this example is convex and being a simple 2D function, the choice of the initial value was

[11] Typically chosen several orders of magnitude (10^{-8}) or smaller than the estimated parameter.
[12] Steep descent refers to the magnitude of change in objective function per iteration.

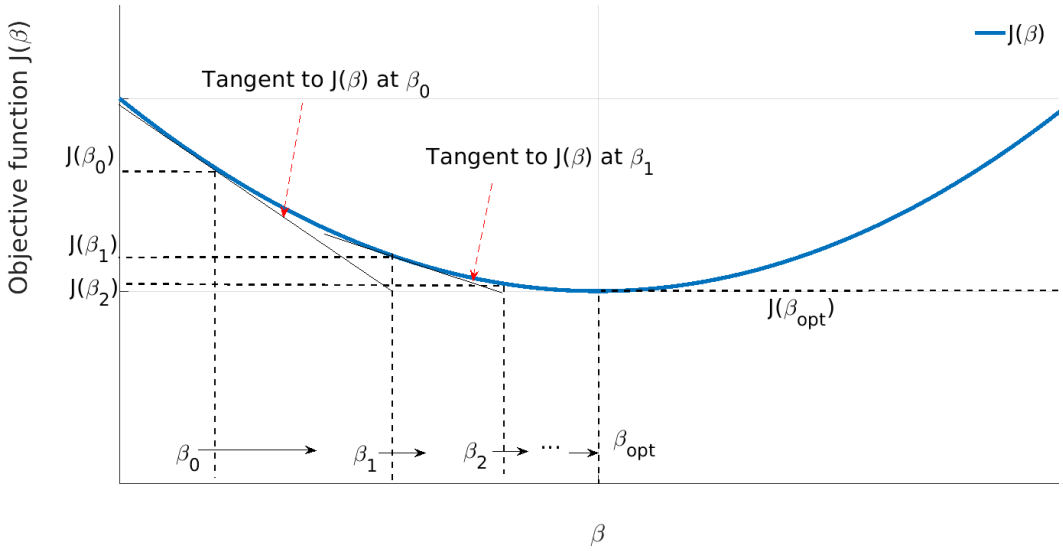


Figure 4.6: Descent method: Traversal from initial value β_0 to the optimal solution β_{opt}. The slope or tangent to the objective function at the initial solution provides a direction to traverse to the updated solution, e.g. from β_0 to β_1. At the objective function minimum, the slope is zero.

inconsequential. More generally, an optimal solution is sensitive to both the choice of initial value as well as the step size.

As shown in Fig. 4.9, at any intermediate solution $\boldsymbol{\beta}_{it} \in \mathcal{F}_R$, a tangent to the objective function (shown as dotted line) divides the entire feasible region into two sub-regions. On one side the objective function increases (sub-region I) and on the other, it decreases (sub-region D). The objective function gradient $\nabla J(\boldsymbol{\beta}_{it})$, at a given intermediate solution, is perpendicular to the local tangent and points in the direction of increasing objective function (black arrow in Fig. 4.9).

> The objective function $J(\boldsymbol{\beta})$ is a scalar function of the model parameter vector $\boldsymbol{\beta}$. Its gradient $\nabla J(\boldsymbol{\beta}_{it})$ is a vector. Furthermore, the second derivative of the objective function with respect to the model parameter will be a matrix. It will be used in Newton's method.

Moving any direction towards the sub-region I, increases the objective function. Therefore, in order to reach the minimum of the objective function, the descent direction $\Delta\boldsymbol{\beta}$ must, progressively (at each iteration), point towards the sub-region D. At a given intermediate solution $\boldsymbol{\beta}_{it}$, the computed descent direction $\Delta\boldsymbol{\beta}_{it}$ and the objective function gradient $\nabla J(\boldsymbol{\beta}_{it})$ must satisfy

$$\nabla J(\boldsymbol{\beta}_{it})^T \cdot \Delta\boldsymbol{\beta}_{it} < 0, \tag{4.34}$$

which means the objective function gradient[13] and the descent direction must form an obtuse[14] angle. As shown in Fig. 4.9, the descent direction along the blue arrow satisfies Eq. (4.34) but does not necessarily point towards the minimum. There is an ideal direction (green arrow) which is the negative of the objective function gradient, along which the it reduces in the steepest manner. On the right side in Fig. 4.9, on contrary, the direction indicated by the red arrow is not a valid descent direction, as it does not satisfy Eq. (4.34). This is an important and a fundamental result which will be utilized throughout the chapter.

The reason for an iterative procedure in the descent methods is because the initially estimated direction $\Delta\boldsymbol{\beta}_{it}$ (corresponding to the first iteration) may not directly lead to the global minimum. Therefore, given an intermediate solution $\boldsymbol{\beta}_{it}$ (or equivalently initial solution $\boldsymbol{\beta}_{it} = \boldsymbol{\beta}_0$) of the unknown model parameter, a two step approach is followed:

[13] $\nabla J(\boldsymbol{\beta}_{it})^T$ and $\Delta\boldsymbol{\beta}_{it}$ are generally not unit vectors.
[14] An obtuse angle is larger than 90 degrees but smaller than 180 degrees.

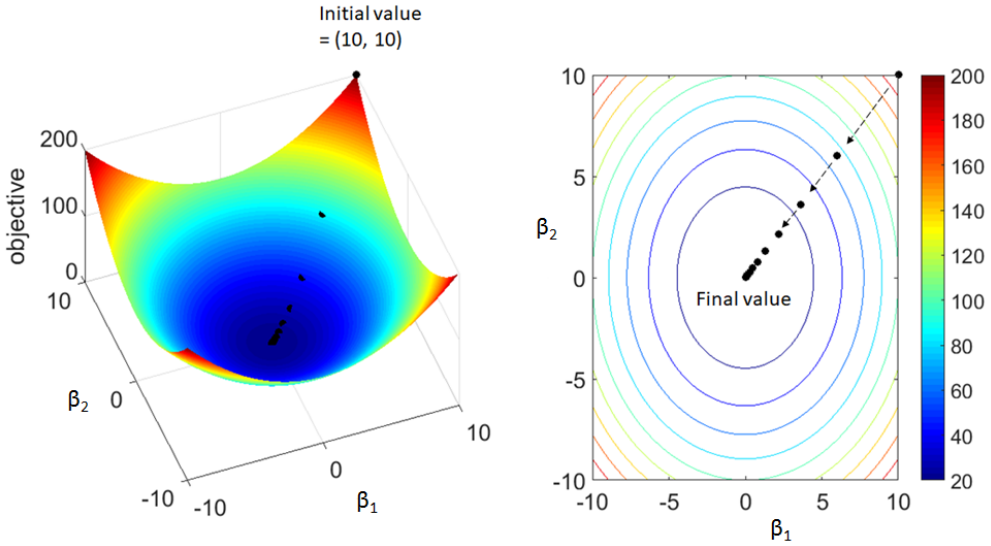


Figure 4.7: Descent method illustration. Starting from an initial value $[10\ 10]^T$, the descent method traverses towards the minima of the objective function and eventually reaches $J(\boldsymbol{\beta}) = 0$ at $\boldsymbol{\beta} = [0\ 0]^T$. The number of iterations required to reach the minima depends on the initial value and the step size. In this example, the objective function is convex and therefore the initial value is inconsequential for the optimal solution.

- Step 1: A descent direction $\Delta\boldsymbol{\beta}_{it}$ is determined that satisfies Eq. (4.34)
- Step 2: A step size t (omitting the iteration subscript) is determined and the unknown model parameters are updated as per Eq. (4.29).

The rest of this chapter will describe the methods for these two important steps in further detail. In practice, the above mentioned two steps are used with minor modifications by different optimization algorithms due to the need for performance improvements. In the next section, two fundamental algorithms, namely, the gradient descent (GD) and Newton's method (NM), are presented for the estimation of the descent direction. Both of these are building blocks in advanced optimization techniques used in this work. It will also be shown that several improvements to the basic gradient descent have been proposed in the last years.

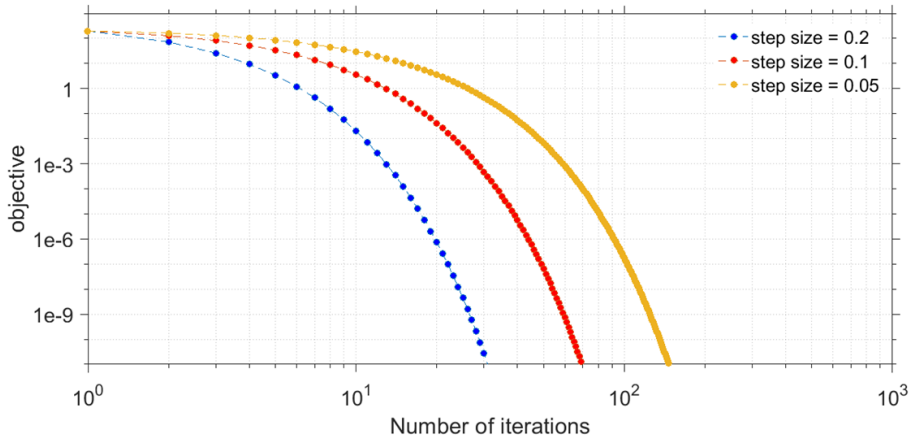


Figure 4.8: The number of iterations and the path taken by the intermediate solutions in BGD for three different step sizes.

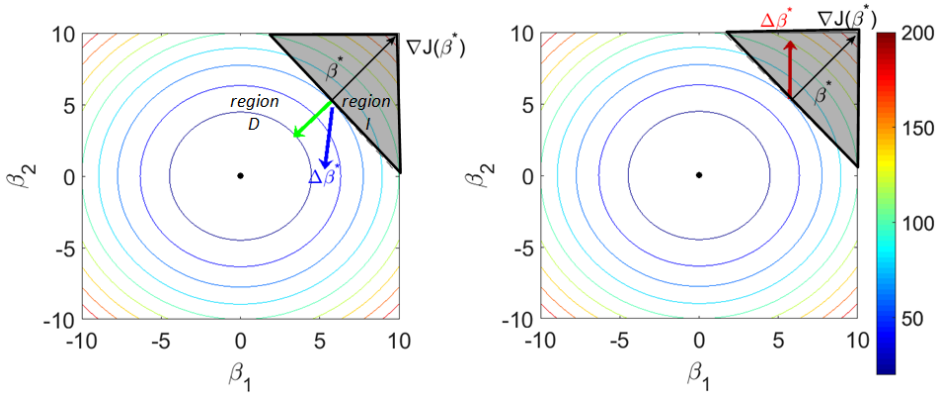


Figure 4.9: The objective function gradient and descent direction: Objective function iso-contours are plotted along the β_1 and β_2 axes. The objective function values increase with increase in $|\beta_1|$ and $|\beta_2|$. Although there are no constraints, but at a point $\boldsymbol{\beta}^*$, there are two possible directions (1) towards increasing objective function (indicated by the black arrow) into the "I" region painted in gray and (2) towards decreasing objective function to the region "D". On the left side, an appropriate descent direction $\Delta\boldsymbol{\beta}^*$ is illustrated at $\boldsymbol{\beta}^*$ that is pointed toward the "D" region whereas on the right side, it is inappropriately pointed in the "I" region.

4.3.2 Gradient descent method

The GD method was first suggested by Cauchy[15] in 1847 (Robbins (2007)) and its use for optimization problems was first studied by Curry (1944). Gradient descent works on the principle that a maximum decrease of the objective function occurs when moving along the negative gradient direction of the objective function (Boyd and Vandenberghe (2004) pg. 466, Nocedal and Wright (2006) pg. 101), just as shown in Eq. (4.34). Therefore, the search for an optimal solution begins with a first natural choice for the descent direction

$$\Delta\boldsymbol{\beta}_{it} = -\nabla J(\boldsymbol{\beta}_{it}), \tag{4.35}$$

the negative gradient of the objective function itself and substituting it in the update rule (4.29), leads to (c.f. Eq. (4.29))

$$\boldsymbol{\beta}_{it+1} = \boldsymbol{\beta}_{it} - t_{it}\ \nabla J(\boldsymbol{\beta}_{it}). \tag{4.36}$$

This algorithm is called the batch gradient descent (BGD) and its main advantage is the simpler update rule. However, its performance can be improved by considering the enhancements as described further.

Mini batch gradient descent

In BGD the objective function gradient is computed using the full observation vector $\boldsymbol{y}$ and especially for a quadratic program (QP) with a large number of observations[16] (e.g. $n_{obs} > 100,000$), there is a considerable complexity in every iteration. To mitigate this particular issue, a variant of BGD, called the mini batch gradient descent (MBGD) (see e.g. Khirirat et al. (2017)), uses only a subset of the total number of observations with which the objective function

$$\begin{aligned} J(\boldsymbol{\beta}_{it,MBGD})_{i:i+n} = &\left(\boldsymbol{y}_{i:i+n} - \boldsymbol{A}_{i:i+n}\ \boldsymbol{\beta}_{it,MBGD}\right)^T \boldsymbol{P}_{i:i+n} \left(\boldsymbol{y}_{i:i+n} - \boldsymbol{A}_{i:i+n}\ \boldsymbol{\beta}_{it,MBGD}\right) \\ &\text{for } i = \{1, 2, \cdots, n_{obs} - n\} \end{aligned} \tag{4.37}$$

[15] Baron Augustin-Louis Cauchy was a French mathematician, engineer, and physicist who made pioneering contributions to several branches of mathematics, including mathematical analysis and continuum mechanics.

[16] It shall be noted that a change in variable convention is introduced here. n_{obs} and n refer to the total number of observations and the mini batch size.

and its gradient are computed. Elements subscript $i:i+n$ in Eq. (4.37) denotes the mini batch subset with only $n+1$ from altogether n_{obs} observations. Specifically, $\boldsymbol{y}_{i:i+n}$ denotes the mini batch subset of observations from $\boldsymbol{y}$. Similarly, $\boldsymbol{A}_{i:i+n}$ denotes the sub-matrix corresponding to the i^{th} to $(i+n)^{th}$ row of the design matrix $\boldsymbol{A}$ and $\boldsymbol{P}_{i:i+n}$[17] denotes the sub-matrix corresponding to i^{th} to $(i+n)^{th}$ row and column of the weight matrix $\boldsymbol{P}$. The update rule in MBGD is

$$\begin{gathered}\boldsymbol{\beta}_{it,MBGD} = \boldsymbol{\beta}_{it-1,MBGD} - t_{it-1}\ \nabla J(\boldsymbol{\beta}_{it-1,MBGD})_{i:i+n} \\ \text{for } i = \{1,2,\cdots,n_{obs}-n\}\end{gathered}, \tag{4.38}$$

where the model parameter $\boldsymbol{\beta}_{it,MBGD}$ is updated along descent direction $\nabla J(\boldsymbol{\beta}_{it-1,MBGD})_{i:i+n}$, as illustrated in Fig. 4.10. The advantage of BGD (4.36) compared to MBGD (4.38) is that the intermediate solutions path is smooth but the disadvantage in BGD is that for each iteration the entire observation vector is used for computing the derivative of the objective function.

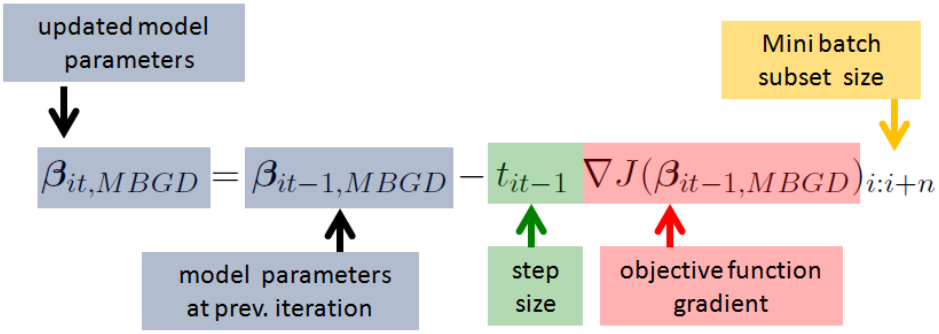


Figure 4.10: MBGD update rule: The model parameter is updated using same three terms as in the basic descent method. The only difference is the mini batch subset size used for the gradient computation.

Stochastic gradient descent

For a mini batch subset corresponding to $n = 1$, the method is called stochastic gradient descent (SGD), which was originally proposed by Robbins and Monro (1951) and is used extensively in solving optimization problems (see e.g. Bottou (2012), Bao (2020)). However, due to the development of CPUs/GPUs[18] capable of performing fast matrix operations, it is more common to use MBGD instead of SGD[19]. Both MBGD and SGD are particularly suitable for real time applications, where the model parameters have to be estimated from only a few observations.

Stochastic Gradient Descent with momentum

As described above, both MBGD and SGD compute an approximate gradient of the objective function using a mini batch. In other words, at each iteration of SGD, the traversal is not always along the optimal direction, due to the noise on the computed gradient of the objective function. "SGD with momentum" or SGDM (Polyak (1964), Qian (1999), Sutskever et al. (2013)) is an improved alternative to compute the objective function gradient. Compared to MBGD or SGD, the noise around the intermediate solution path is reduced by introducing a momentum parameter

[17] $\boldsymbol{P}_{i:i+n}$ must be positive definite
[18] CPU is the acronym for Central Processing Unit and GPU for Graphics Processing Unit.
[19] Due to the fact the performance is nearly the same when using a subset of $n \sim 1000$ compared to $n = 1$, sometimes MBGD is simply called SGD

γ_{mom} with $0 < \gamma_{mom} \leq 1$ such that the update rule in SGDM reads

$$\boldsymbol{\beta}_{it,SGDM} = \boldsymbol{\beta}_{it-1,SGDM} + t_{it-1}\ \boldsymbol{z}_{it,SGDM} \tag{4.39a}$$

$$\boldsymbol{z}_{it,SGDM} = \gamma_{mom}\boldsymbol{z}_{it-1,SGDM} - (1-\gamma_{mom})\nabla J(\boldsymbol{\beta}_{it-1,SGDM})_{i:i+n}. \tag{4.39b}$$

When the objective function gradient points almost in the same direction, progressively with iterations, the momentum factor γ_{mom} is set to a large value (≈ 0.9), thus increasing term $\gamma_{mom}\boldsymbol{z}_{it-1,SGDM}$. In contrast, γ_{mom} is set to relatively smaller value (≈ 0.5) to reduce the update when there is a considerable change in the gradient direction with iterations. This, in turn, results in two advantages, namely (1) convergence in fewer iterations compared to SGD due to the momentum gain of the updated model parameters and (2) reduced noise compared to MBGD or SGD. In practice, the value of γ_{mom} is initialized to 0.5 and progressively increased to 0.99 as convergence is reached (Driggs et al. (2020)). It can be seen that for $\gamma_{mom} = 0$, Eqs. (4.39a) and (4.39b) reduce to (4.38) or, in other words, SGDM reduces to MBGD. "*Momentum*" in SGDM comes from the analogy to physics, where a ball rolling down a slope gains momentum until an obstruction is encountered. Although γ_{mom} is called the momentum factor, actually it is analog to the coefficient of friction[20]. The two steps of SGDM are summarized below.

- Step 1: The momentum term $\boldsymbol{z}_{it,SGDM}$ is computed at an iteration it using $\boldsymbol{z}_{it-1,SGDM}$ and the objective function gradient, both from the previous iteration $it-1$, as well as the momentum factor γ_{mom} ; see Eq. (4.39b).
- Step 2: The unknown model parameter vector $\boldsymbol{\beta}_{it,SGDM}$ is updated using Eq. (4.39a).

Stochastic Gradient Descent - Nestrov

Another useful characteristic of the descent method is the ability to "predict" a direction, in which the objective function will start to increase (also called the "ascent" in contrast to the descent direction), so that the model parameter update is controlled. This method is abbreviated SGDN in the honour of Yurii Nestrov who first proposed it[21] (see Nesterov (1983)). In this method, an approximation at an iteration it of the intermediate solution

$$\boldsymbol{\beta}_{it,pred} = \boldsymbol{\beta}_{it-1,SGDN} - t_{it-1}\ \nabla J(\boldsymbol{\beta}_{it-1,SGDN})_{i:i+n} \tag{4.40}$$

is predicted using the model parameters and the objective function gradient both from the previous iteration. Accordingly the update rule becomes

$$\boldsymbol{z}_{it,SGDN} = \gamma_{mom}\boldsymbol{z}_{it-1,SGDN} - \nabla J(\boldsymbol{\beta}_{it,pred})_{i:i+n} \tag{4.41a}$$

$$\boldsymbol{\beta}_{it,SGDN} = \boldsymbol{\beta}_{it-1,SGDN} + t_{it-1}\ \boldsymbol{z}_{it,SGDN}. \tag{4.41b}$$

The three steps involved in SGDN are summarized below:

- Step 1: A predicted value $\boldsymbol{\beta}_{it,pred}$ of the model parameter vector is computed at iteration it ; see Eq. (4.40).
- Step 2: The momentum term $\boldsymbol{z}_{it,SGDN}$ is computed using that from previous iteration $\boldsymbol{z}_{it-1,SGDN}$ and the objective function gradient computed from the predicted value $\boldsymbol{\beta}_{it,pred}$ from Step 1 above using the momentum factor γ_{mom} ; see Eq. (4.41a).
- Step 3: Model parameter $\boldsymbol{\beta}_{it,SGDN}$ is updated; see Eq. (4.41b).

[20] A smoother surface has a smaller coefficient of friction compared to a rough surface. A ball rolling down a slope with small coefficient of friction will gain momentum and kinetic energy to reach the foot of the slope faster compared to that on a rough slope of same length

[21] SGDN is also called NAG, an acronym for Nestrov accelerated gradient

Table 4.2: Variants of the gradient descent algorithm

Gradient descent algorithm	Primary characteristics
BGD	Batch gradient descent: A basic version which uses the entire observation for objective function gradient computation
MBGD	Mini batch gradient descent: Objective function gradient is computed with a subset of observations.
SGD	Stochastic gradient descent: Objective function gradient computed with every single observation.
SGDM	SGD with Momentum: Improved convergence compared to SGD and reduces the noise in the path followed by the intermediate solution.
SGDN	SGD with Nestrov: Controls the model parameter update by its prediction, so that the traversal is within the feasible region.

Summary of gradient descent algorithms

Including the basic version (BGD), five variants of gradient descent algorithm discussed above are summarized in the Table 4.2. The step size parameter is used in all five algorithms mentioned above (BGD, SGD, SGDM and SGDN). However, it is a challenge setting it appropriately because for step size $t \approx 0$), the model parameters convergence will be slow and it may take a considerable time to achieve an acceptable objective function minimum. On the other hand, if $t \approx 1$, then the model parameters may diverge. This is only a qualitative assessment and a more specific technique for step size estimation will be presented in section 4.7.3. One way to mitigate this problem is to choose a different step size for each dimension, but that would not be practical when there are thousands or more dimensions. There are advanced algorithms "AdaGrad" (Duchi et al. (2011)), "Adam" (Kingma and Ba (2014)), "RMSProp" (Hinton (2012)) and "Adadelta" (Zeiler (2012)) based on gradient descent, which address such challenges.

4.3.3 Newton's method

Newton's method (NM) was first published in 1685 in "*A Treatise of Algebra both Historical and Practical*" (Wallis (1685)) and is an extremely powerful technique with quadratic convergence (see e.g. Knight and Adams (1975), Hestenes (1980) and Bartholomew-Biggs (2005)). It means the difference between a true solution and its approximation is squared (i.e. the number of accurate digits roughly doubles) at each iteration step. A second order Taylor series approximation of the objective function $J(\boldsymbol{\beta})$ about initial value $\boldsymbol{\beta}_0$, such that $\Delta\boldsymbol{\beta} = \boldsymbol{\beta} - \boldsymbol{\beta}_0$, is

$$J(\boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta}) = J(\boldsymbol{\beta}_0) + \nabla_{\boldsymbol{\beta}} J(\boldsymbol{\beta})^T \Delta\boldsymbol{\beta} + \frac{1}{2}\Delta\boldsymbol{\beta}^T \nabla^2_{\boldsymbol{\beta}} J(\boldsymbol{\beta})\Delta\boldsymbol{\beta}, \tag{4.42}$$

which is quadratic in $\Delta\boldsymbol{\beta}$ and $\nabla^2_{\boldsymbol{\beta}} J(\boldsymbol{\beta})$ denotes the Hessian matrix[22] of the objective function. Taking the first derivative of Eq. (4.42) with respect to $\Delta\boldsymbol{\beta}$ and equating it to zero gives

$$\nabla_{\boldsymbol{\beta}} J(\boldsymbol{\beta}) + \nabla^2_{\boldsymbol{\beta}} J(\boldsymbol{\beta})\Delta\boldsymbol{\beta} = \mathbf{0}_{u\times 1} \tag{4.43}$$

which in turn shows that $J(\boldsymbol{\beta}_0 + \Delta\boldsymbol{\beta})$ is minimized when

$$\Delta\boldsymbol{\beta} = -\nabla^2_{\boldsymbol{\beta}} J(\boldsymbol{\beta})^{-1} \nabla_{\boldsymbol{\beta}} J(\boldsymbol{\beta}). \tag{4.44}$$

[22] Matrix of second derivative of the objective function with respect to the model parameter.

Substituting this result in Eq. (4.42), the first order approximation can be written as

$$J(\beta_0 + \Delta\beta) = J(\beta_0) - \nabla_\beta J(\beta)^T \left(\nabla^2_\beta J(\beta)\right)^{-1} \nabla_\beta J(\beta) \tag{4.45}$$

or

$$J(\beta_0 + \Delta\beta) - J(\beta_0) = -\nabla_\beta J(\beta)^T \left(\nabla^2_\beta J(\beta)\right)^{-1} \nabla_\beta J(\beta). \tag{4.46}$$

Furthermore, when $\nabla^2_\beta J(\beta)$ is positive definite, the left hand side of Eq. (4.46) is negative or in

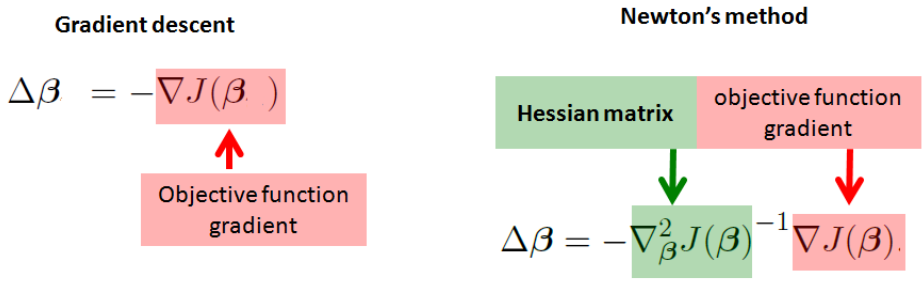


Figure 4.11: Comparison of the descent direction computed within the gradient descent and Newton's method. In GD, the descent direction is the negative gradient of objective function, whereas in NM, there is an additional Hessian matrix inverse term multiplied with the negative gradient of the obective function.

other words

$$J(\beta_0 + \Delta\beta) - J(\beta_0) < 0 \tag{4.47}$$

and therefore by moving a step $|\Delta\beta|$ along the descent direction, the objective function decreases. In the Newton's method, the update rule is defined, similar to as it was defined in the descent method (see Eq. (4.29)), as

$$\beta_{it+1} = \beta_{it} + t_{it}\Delta\beta_{it}, \tag{4.48}$$

where the value of t_{it} is set by default to 1. The main challenge in Newton's method is the Hessian matrix $\nabla^2_\beta J(\beta)$ from a computational perspective (see Boyd and Vandenberghe (2004) pg. 484, Nocedal and Wright (2006) pg. 274).

It has been shown that in both GD and NM the traversal is along the direction of the reducing gradient of the objective function. In fact, in any optimization algorithm, the estimation of the traversal direction is one of the important and possibly the most computationally expensive step. This is also referred to as the "*line search*" because the updated model parameter vector β_{it+1} will lie along the line segment $\beta_{it} + t_{it}\Delta\beta_{it}$ depending on the value of t_{it}, although in general, the line could be multi-dimensional. When $t_{it} = 1$, the update is said to be taking a "full Newton step". In the case of an unconstrained optimization (assuming convex objective function), a full Newton step is often used (Nocedal and Wright (2006) pg. 164), as shown in Fig. 4.12, but when there are inequality constraints, only relatively smaller steps ($t_{it} \ll 1$) along the descent direction[23] can be taken (Nocedal and Wright (2006) pg. 56).

Both GD and NM scale well for higher dimensional problems subject to the convexity of the objective function (Nocedal and Wright (2006) pg. 16). One of the challenges in the design of optimization algorithms is the estimation of the step size. Objective function values during the iterations are tested for identifying the need for adapting the step size. A comparison of the descent direction in the GD and NM methods is shown in Fig. 4.11. In the next section, the theory behind the inequality constrained optimization along with the definitions of some other important terms will be introduced.

[23] Descent direction computed using NM is also called Newton direction.

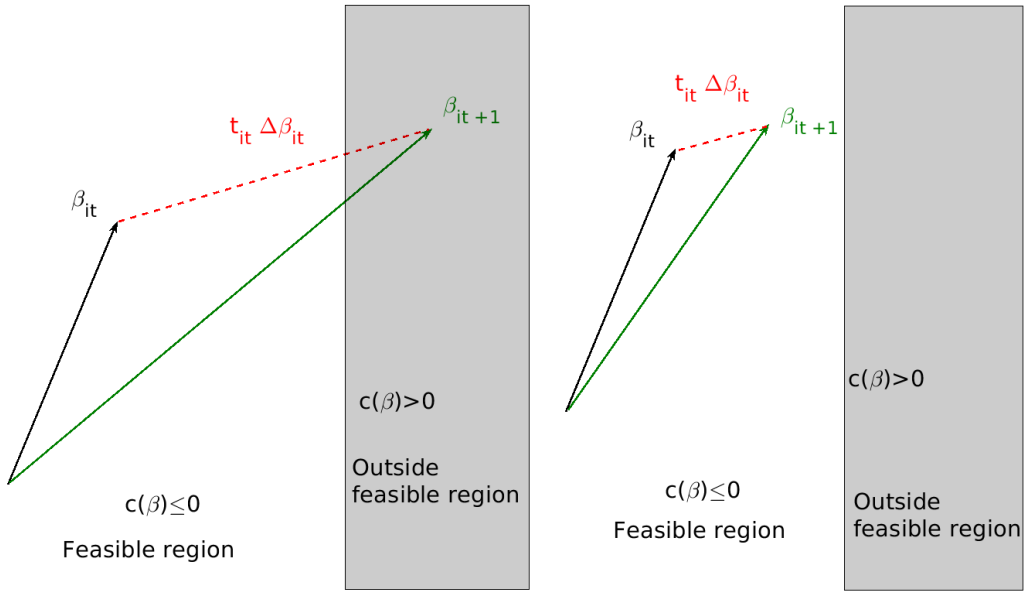


Figure 4.12: Illustration of step size t_{it} during parameter update from β_{it} to β_{it+1} along the descent direction $\Delta\beta_{it}$ in descent methods. The white region is where $c(\beta) \leq 0$ and in the gray region $c(\beta) > 0$. Two scenarios are depicted: On the left side, there is a constraint violation because the updated model parameter is outside the feasible region. On the right side, smaller step size keeps the parameter within the feasible region.

4.4 Inequality constraints

4.4.1 Introduction to inequalities

Mathematical inequalities are functional expressions with the following relations: (1) larger than $''>''$, (2) smaller than $''<''$, (3) larger than or equal to "$\geq$", (4) smaller than or equal to "$\leq$" or (5) "not equal to" $\neq$ (see e.g. Stephenson (1971), Cloud et al. (2014)). This section begins with an introduction to the scalar, vector and matrix inequality conventions used in this work. For scalars β_1 and β_2 such that $\beta_1 > \beta_2 \implies \beta_1$ is larger than β_2 and similarly, $\beta_1 < \beta_2 \implies \beta_1$ is smaller than β_2. Using a similar analogy, $\beta_1 \geq \beta_2$ denotes β_1 is larger than or equal to β_2 and $\beta_1 \leq \beta_2$ denotes β_1 is smaller than or equal to β_2. While scalar inequalities are commonly used and easily understood, a more formal convention is needed for vectors and matrices.

For vectors $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and the zero vector[24] denoted as $\mathbf{0}$, all three with the same dimensions,

$$\boldsymbol{\beta}_1 > \boldsymbol{\beta}_2 \implies \text{each element of } \boldsymbol{\beta}_1 \text{ is larger than that in } \boldsymbol{\beta}_2 \tag{4.49a}$$
$$\boldsymbol{\beta}_1 < \boldsymbol{\beta}_2 \implies \text{each element of } \boldsymbol{\beta}_1 \text{ is smaller than that in } \boldsymbol{\beta}_2 \tag{4.49b}$$
$$\boldsymbol{\beta}_1 \geq \boldsymbol{\beta}_2 \implies \text{each element of } \boldsymbol{\beta}_1 \text{ is larger than or equal to that in } \boldsymbol{\beta}_2 \tag{4.49c}$$
$$\boldsymbol{\beta}_1 \leq \boldsymbol{\beta}_2 \implies \text{each element of } \boldsymbol{\beta}_1 \text{ is smaller than or equal to that in } \boldsymbol{\beta}_2 \tag{4.49d}$$
$$\boldsymbol{\beta}_1 \neq \mathbf{0} \implies \text{at least one element of } \boldsymbol{\beta}_1 \text{ is non-zero.} \tag{4.49e}$$
$$\boldsymbol{\beta}_1 \neq \boldsymbol{\beta}_2 \implies \text{at least one element of } \boldsymbol{\beta}_1 \text{ is unequal to that in } \boldsymbol{\beta}_2. \tag{4.49f}$$
$$\boldsymbol{\beta}_1 \ll \boldsymbol{\beta}_2 \implies \text{at least one element of } \boldsymbol{\beta}_1 \text{ is much smaller than the corresponding one in } \boldsymbol{\beta}_2 \tag{4.49g}$$
$$\boldsymbol{\beta}_1 \gg \boldsymbol{\beta}_2 \implies \text{at least one element of } \boldsymbol{\beta}_1 \text{ is much larger than the corresponding one in } \boldsymbol{\beta}_2 \tag{4.49h}$$
$$\boldsymbol{\beta}_1 \approx \boldsymbol{\beta}_2 \implies \text{each element of } \boldsymbol{\beta}_1 \text{ is approximately equal to that in } \boldsymbol{\beta}_2 \tag{4.49i}$$

For a $u \times u$ square matrix $\boldsymbol{A}$, $\mathbf{0}_u$[25] and vector $\mathbf{0}$[26]

$$\boldsymbol{A} > \mathbf{0}_u \implies \boldsymbol{A} \text{ is positive definite} \quad \boldsymbol{\beta}^T \boldsymbol{A} \boldsymbol{\beta} > 0 \;\; ; \forall \;\; \boldsymbol{\beta} \neq \mathbf{0} \tag{4.50a}$$
$$\boldsymbol{A} \geq \mathbf{0}_u \implies \boldsymbol{A} \text{ is positive semi-definite} \quad \boldsymbol{\beta}^T \boldsymbol{A} \boldsymbol{\beta} \geq 0 \;\; ; \forall \;\; \boldsymbol{\beta} \tag{4.50b}$$
$$\boldsymbol{A} < \mathbf{0}_u \implies \boldsymbol{A} \text{ is negative definite} \quad \boldsymbol{\beta}^T \boldsymbol{A} \boldsymbol{\beta} < 0 \;\; ; \forall \;\; \boldsymbol{\beta} \neq \mathbf{0} \tag{4.50c}$$
$$\boldsymbol{A} \leq \mathbf{0}_u \implies \boldsymbol{A} \text{ is negative semi-definite} \quad \boldsymbol{\beta}^T \boldsymbol{A} \boldsymbol{\beta} \leq 0 \;\; ; \forall \;\; \boldsymbol{\beta} \tag{4.50d}$$
$$\boldsymbol{A} \nleq \mathbf{0}_u \text{ as well as } \boldsymbol{A} \ngeq \mathbf{0}_u \implies \boldsymbol{A} \text{ is indefinite.} \tag{4.50e}$$

Additionally for any m × n matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_2$,

$$\boldsymbol{A}_1 \neq \boldsymbol{A}_2 \implies \text{at least one element of } \boldsymbol{A}_1 \text{ is unequal to that in } \boldsymbol{A}_2. \tag{4.51}$$

Eq. (4.50e) is important because, unlike in scalars, the law of Trichotomy for real numbers (Rubin and Rubin (1985)) does not hold for matrices. Specifically, for a given pair of scalars β_1, β_2, only one of the three conditions

$$\beta_1 < \beta_2 \quad \text{or} \quad \beta_1 = \beta_2 \quad \text{or} \quad \beta_1 > \beta_2 \tag{4.52}$$

holds, whereas for matrices

$$\boldsymbol{A} \nleq \mathbf{0}_u \text{ does not } \implies \boldsymbol{A} > \mathbf{0}_u. \tag{4.53}$$

Testing for positive definiteness

A method of testing for positive definiteness of matrices, as given in Bernstein (2009), is summarized below. A square matrix

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1u} \\ a_{21} & a_{22} & \cdots & a_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ a_{u1} & a_{u2} & \cdots & a_{uu} \end{bmatrix}$$

[24] In this case $\mathbf{0}$ denotes a vector of all zeros having the same dimension as $\boldsymbol{\beta}_1$
[25] $\mathbf{0}_u$ denotes $u \times u$ square matrix of all zeros.
[26] $\mathbf{0}$ denotes a vector of zeros with the same dimension as $\boldsymbol{\beta}$.

is positive definite if the determinant of each of the principal minors

$$\boldsymbol{\mu}_1 = [a_{11}] \tag{4.54a}$$

$$\boldsymbol{\mu}_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \tag{4.54b}$$

$$\boldsymbol{\mu}_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{4.54c}$$

$$\boldsymbol{\mu}_u = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1u} \\ a_{21} & a_{22} & \cdots & a_{2u} \\ \vdots & \vdots & \ddots & \vdots \\ a_{u1} & a_{u2} & \cdots & a_{uu} \end{bmatrix} \tag{4.54d}$$

are positive, i.e

$$\det \boldsymbol{\mu}_i > 0 \quad \forall \; i \in [1, \cdots, u] \implies \boldsymbol{A} > \mathbf{0}_u. \tag{4.55}$$

Properties of matrix inequality

For $u \times u$ square matrices $\boldsymbol{A}_1, \boldsymbol{A}_2$,

$$\boldsymbol{A}_1 > \boldsymbol{A}_2 \implies \boldsymbol{A}_1 - \boldsymbol{A}_2 > \mathbf{0}_u \tag{4.56a}$$

$$\boldsymbol{A}_1 \geq \boldsymbol{A}_2 > \mathbf{0}_u \implies {\boldsymbol{A}_2}^{-1} \geq {\boldsymbol{A}_1}^{-1} > \mathbf{0}_u \tag{4.56b}$$

$$\boldsymbol{A}_1 > \mathbf{0}_u \text{ and } r > 0; r \in \mathbb{R} \implies r\boldsymbol{A}_1 > \mathbf{0}_u \tag{4.56c}$$

$$\boldsymbol{A}_1 \geq \mathbf{0}_u \text{ and } \boldsymbol{A}_2 \geq \mathbf{0}_u \implies \boldsymbol{A}_1 + \boldsymbol{A}_2 \geq \mathbf{0}_u \tag{4.56d}$$

$$\boldsymbol{A}_1 > \mathbf{0}_u \text{ and } \boldsymbol{A}_2 > \mathbf{0}_u \implies \boldsymbol{A}_1\boldsymbol{A}_2\boldsymbol{A}_1 > \mathbf{0}_u \text{ and } \boldsymbol{A}_2\boldsymbol{A}_1\boldsymbol{A}_2 > \mathbf{0}_u \tag{4.56e}$$

$$\boldsymbol{A}_1 > \mathbf{0}_u \text{ and } \boldsymbol{A}_2 > \mathbf{0}_u \implies \alpha\boldsymbol{A}_1 + (1-\alpha)\boldsymbol{A}_2 > \mathbf{0}_u \quad \forall \; \alpha \in [0, 1] \tag{4.56f}$$

Definitions of some important terms frequently used in optimization theory are introduced below.

Definition 4.4.1. *Constraint function*

Constraints are specific conditions, bounds or empirical relations between model parameters which restrict or limit their range of values. These can be both equality or inequality relations and accordingly are called equality or inequality constraints.

In this work, an inequality constraint function is represented as $c(\boldsymbol{\beta}) \leq 0$. When there are multiple inequality constraints, the system of constraint functions $c_j(\boldsymbol{\beta}) \leq 0$ are represented using the subscript j that refers to the set $\mathcal{I}$ of all inequality constraints. The feasible region $\mathcal{F}_R$ is determined by the constraints functions.

Definition 4.4.2. *Candidate solution*

A candidate solution is a member of the feasible region of a given problem (Boyd and Vandenberghe (2004) pg. 2, pg. 11, Nocedal and Wright (2006) pg. 20). It is not necessary that a candidate solution shall be the optimal or even an approximate solution to the optimization problem. The reason it is called as a "candidate solution" is because it is one of the many in the feasible region that satisfies the constraints.

Inequality constraint categorization

Following Arora (2017), an inequality constraint can either be active, γ-active, violated, or inactive. On the other hand, an equality constraint is either active or violated at a given value of the model parameter.

- Active constraint: An inequality constraint $c_j(\boldsymbol{\beta}) \leq 0$, for $j \in \mathcal{I}$, is said to be active at $\boldsymbol{\beta}_*$, if the equality relation $c_j(\boldsymbol{\beta}_*) = 0$ holds.
- Inactive constraint: An inequality constraint $c_j(\boldsymbol{\beta}) \leq 0$, for $j \in \mathcal{I}$, is said to be inactive at $\boldsymbol{\beta}_*$, if the inequality $c_j(\boldsymbol{\beta}_*) < 0$ holds.
- Violated constraint: An inequality constraint $c_j(\boldsymbol{\beta}) \leq 0$, for $j \in \mathcal{I}$, is said to be violated at $\boldsymbol{\beta}_-$, if $c_j(\boldsymbol{\beta}_-) \not\leq 0$ holds. An equality constraint $c_i(\boldsymbol{\beta}) = 0$, for $i \in \mathcal{E}$, is violated at $\boldsymbol{\beta}_-$, if $c_i(\boldsymbol{\beta}_-) \neq 0$ holds. An equality constraint is always either active or violated. For violated constraints, $\boldsymbol{\beta}_- \notin \mathcal{F}_R$ holds.
- γ-active inequality constraint: An inequality constraint $c_j(\boldsymbol{\beta}) \leq 0$, for $j \in \mathcal{I}$, is said to be γ-active at $\boldsymbol{\beta}_*$, if $c_j(\boldsymbol{\beta}_*) < 0$ but $c_j(\boldsymbol{\beta}_*) + \gamma = 0$ holds, such that $\gamma > 0$ and $\gamma \approx 0$. It indicates that a given model parameter is close to the constraint boundary on the feasible side with a margin γ. That is, the constraint $c_j(\boldsymbol{\beta})$ is strictly inactive (see definition 4.4.3) but it is also close to become active. The concept of γ-active constraints apply only to inequality constraints.

Definition 4.4.3. *Strict inequality*
The exclusive inequality part $c_j(\boldsymbol{\beta}) < 0$ of the relation $c_j(\boldsymbol{\beta}) \leq 0$, for $j \in \mathcal{I}$, is called the strict inequality.

Definition 4.4.4. *Linear matrix inequality*
A linear matrix inequality (LMI) is a mathematical relation used for the representation of a system of linear inequations (see e.g. Boyd and Vandenberghe (2004) pg. 38, Forst and Hoffmann (2010)). In this work, a LMI

$$\boldsymbol{B}\boldsymbol{\beta} \leq \boldsymbol{b}$$

is used for the representation of a system of linear inequality constraints, where $\boldsymbol{B}$ and $\boldsymbol{b}$ denote the constraint coefficient matrix and the constraint bound, respectively.

Definition 4.4.5. *Linear program*
A linear program (LP) is an optimization problem having a linear objective function and a feasible region determined by a linear equality and/or inequality constraints.

Definition 4.4.6. *Quadratic program*
A quadratic program (QP) is an optimization problem having a quadratic objective function and a feasible region determined by a linear or a quadratic equality and/or inequality constraints.

Standard form

Definition 4.4.7. A QP is said to be in standard form when the objective function is represented as a sum of a quadratic and a linear term. Additionally, following definition 4.4.4, the system of inequality constraints are represented in matrix form as a "smaller than or equal to" relation. The optimization problem $\mathcal{P}$ (4.25) in standard form reads

$$\begin{aligned} \min_{\boldsymbol{\beta}} \quad & \boldsymbol{\beta}^T \boldsymbol{Q} \boldsymbol{\beta} + \boldsymbol{\beta}^T \boldsymbol{q} \\ \text{subject to :} \quad & \boldsymbol{B}\boldsymbol{\beta} \leq \boldsymbol{b} \end{aligned}, \tag{4.57}$$

where the objective $J(\boldsymbol{\beta})$ from Eq. (4.25) is replaced by a specific quadratic function of $\boldsymbol{\beta}$ along with a given positive semi-definite matrix $\boldsymbol{Q}$ as well as a vector $\boldsymbol{q}$ that will be defined later in Eq. (4.85).

Definition 4.4.8. *Local and global minima*
A local minimum of an objective function $J(\boldsymbol{\beta})$ is defined as its value at a candidate solution $\boldsymbol{\beta}_*$ such that $J(\boldsymbol{\beta}_*) < J(\boldsymbol{\beta}_* \pm \delta\boldsymbol{\beta}_1)$ for $\delta\boldsymbol{\beta}_1 \approx \mathbf{0}$, but possibly $J(\boldsymbol{\beta}_*) > J(\boldsymbol{\beta}_* \pm \delta\boldsymbol{\beta}_2)$ for $\delta\boldsymbol{\beta}_2 \gg \mathbf{0}$. In other words, at a local minima, an objective function might have values smaller than that in its neighbourhood but possibly larger than those that are far away.

The global minimum of an objective function $J(\boldsymbol{\beta})$ is defined as its value at a candidate solution $\boldsymbol{\beta}_*$ such that $J(\boldsymbol{\beta}_*) < J(\boldsymbol{\beta}_* \pm \delta\boldsymbol{\beta}_1)$ for all $\delta\boldsymbol{\beta}_1 \in \mathcal{F}_R$. In other words, at the global minima an objective function has the smallest value compared to that anywhere else in the feasible region.

Constraint bound and constraint direction

Even though an inequality constraint is a scalar function of the model parameter vector, a direction is often associated to it. "Constraint direction" is not a standardized term but is nevertheless used in reference textbooks such as Boyd and Vandenberghe (2004), Forst and Hoffmann (2010), Nocedal and Wright (2006). The constraint direction is normal to the constraint bound hyperplane and points towards the interior of the feasible region. It will be shown using a graphical representation of constraints that their bounds, the direction as well as the the direction of the objective function gradient, all have an impact on the optimal solution.

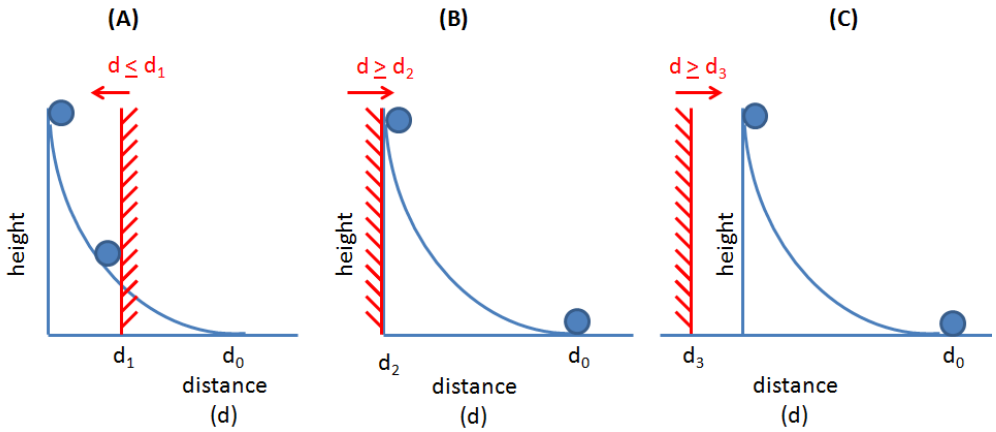


Figure 4.13: Analogy of a ball rolling down a curved path from an initial height. If there are no constraints the ball will eventually reach the final distance d_0 along the horizontal. Three examples illustrate the effect of a constraint or a barrier on the motion of the ball from its initial to its final height. (A): The ball hits a constraint at a distance $d = d_1$ (B) The ball starts rolling just where the constraint is active. The ball reaches $d = d_0$ (C) The presence of the constraint does not affect the ball at all and the ball once again reaches $d = d_0$.

In Fig. 4.13, three distinct examples are shown with a physics analogy, where the initial position of the ball can be compared with the initial value of the model parameter to be estimated. The barrier shown in red acts as an inequality constraint. The height and the horizontal distance are analogous to the objective function and the model parameter. The magnitude of the distance covered in unit time and the direction of motion of the ball correspond to the step length and the descend direction in an optimization algorithm. The final distance along the horizontal corresponds to the optimal solution and the curved path meant the convexity of the objective function. In case (A), the inequality constraint becomes binding at $d = d_1$, in case (B) the constraint is active at the initial value and thereafter remains inactive and in (C) the constraint remain inactive throughout. It can also be seen that both the direction of constraint and the initial value relative to the constraint play an important role in the optimum solution. When the constraint is pointing towards the negative gradient of the objective function, both the unconstrained and constrained optimum will be the same (depending on the initial value). On contrary, when the constraint points along the positive gradient of the objective function, then at some point, the constraint will become binding. In the next section, a mathematical formulation of equality and inequality constraints will be presented along with a complete description of their representation and processing within an optimization algorithm.

4.4.2 Constraint representation and optimization

Generally, there are two types of constraints, namely, the equality and the inequality ones

$$\begin{aligned} c_i(\boldsymbol{\beta}) &= 0 \quad i \in \mathcal{E} \\ c_j(\boldsymbol{\beta}) &\leq 0 \quad j \in \mathcal{I} \end{aligned} \tag{4.58}$$

that are denoted by their sets of indices $\mathcal{E}$ and $\mathcal{I}$ respectively, as in Eq. (4.25). This is a standard form representation where the constraint functions $c_i(\boldsymbol{\beta})$ or $c_j(\boldsymbol{\beta})$ are rearranged to get zero on the right hand side, as shown in Eq. (4.58); see Nocedal and Wright (2006). A constraint function divides the entire space of the model parameter into two sub-regions. Specifically, these are the:

- Feasible region,
$$\mathcal{F}_R = \{\boldsymbol{\beta} \mid c_i(\boldsymbol{\beta}) = 0,\ i \in \mathcal{E}\ ;\ c_j(\boldsymbol{\beta}) \leq 0,\ j \in \mathcal{I}\}\,, \tag{4.59}$$
i.e. the union of all possible model parameter values satisfying the constraints and the
- Infeasible region,
$$\mathcal{I}_R = \{\boldsymbol{\beta} \mid c_i(\boldsymbol{\beta}) \neq 0,\quad i \in \mathcal{E}\ ;\ c_j(\boldsymbol{\beta}) > 0,\quad j \in \mathcal{I}\}\,, \tag{4.60}$$
where the constraints are not satisfied.

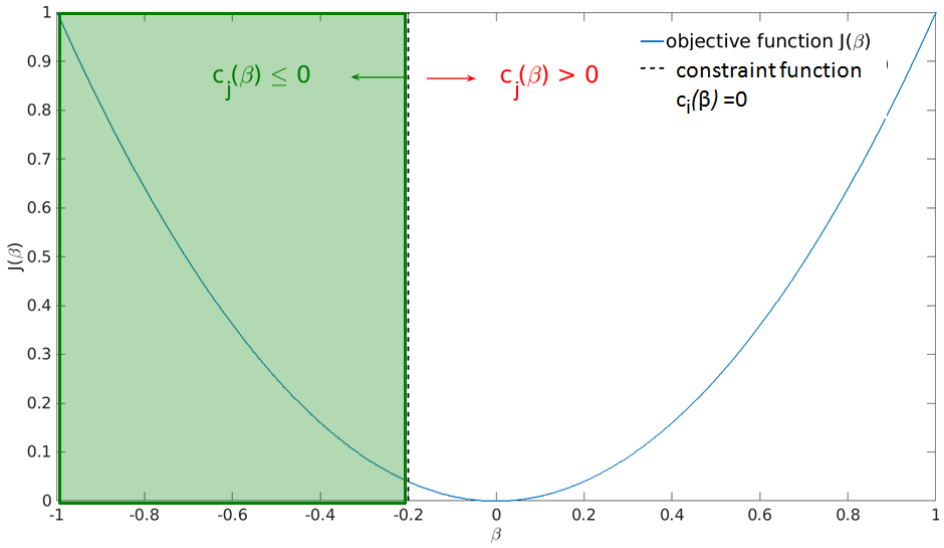


Figure 4.14: Constraint and feasible region in 1D: The dotted black line is the inequality constraint bound for $\beta \leq -0.2$. The green region $\{\beta \mid c_j(\beta) \leq 0\}$ is the feasible region and accordingly the green arrow points leftwards inside the feasible region $\mathcal{F}_R$, normal to the inequality constraint bound. In contrast, the white region is the infeasible region $\mathcal{I}_R : \{\beta \mid c_j(\beta) > 0\}$ pointed by the red arrow. It shall be noted that the entire parameter space is the union of $\mathcal{I}_R$ and $\mathcal{F}_R$. For the objective function $J(\beta)$, the constrained minimum is at $\beta = -0.2$ and the unconstrained minimum is at $\beta = 0$. In this case for the 1D objective function, the constraint function subscripts $i = j = 1$.

A boundary between the two sub-regions is where the condition $c_j(\boldsymbol{\beta}) = 0$ for $j \in \mathcal{I}$ prevails and the inequality constraints are said to be active. The set of active constraints

$$\mathcal{A} = \mathcal{E} \cup j \in \{\mathcal{I} \mid c_j(\boldsymbol{\beta}) = 0\} \tag{4.61}$$

is the union of the equality and inequality constraints $\{\mathcal{E}, \mathcal{I}\}$ for which $c_j(\boldsymbol{\beta}) = 0$. In the feasible region, the inequality constraints (4.58) are strictly satisfied and therefore the corresponding constraints are said to be inactive. It then follows from Eq. (4.61) that, at any infeasible solution, an inequality constraint is said to be violated, if $c_j(\boldsymbol{\beta}) > 0$, for $j \in \mathcal{I}$ holds. The feasible region

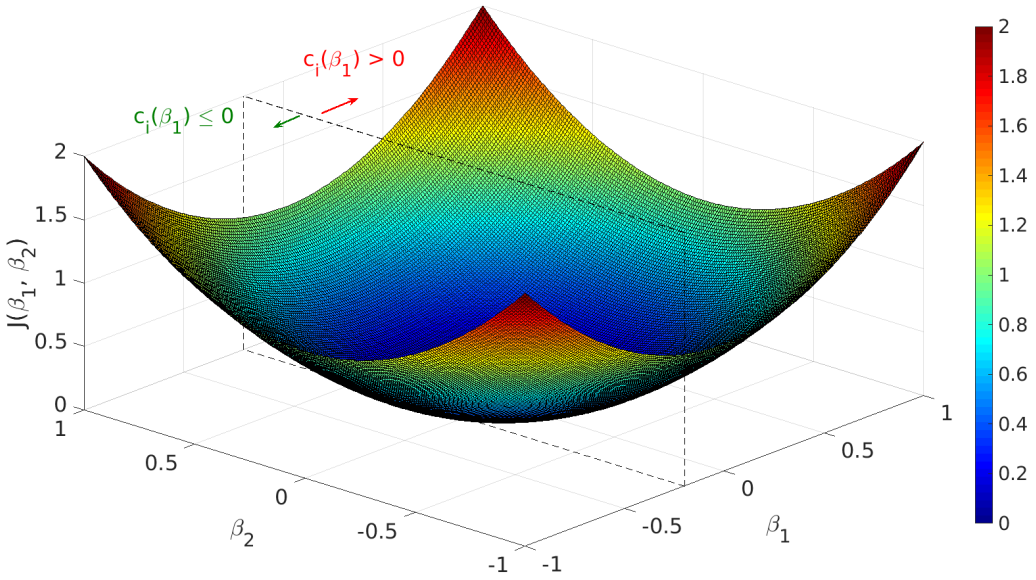


Figure 4.15: Constraint and feasible region in 2D: The dotted black plane is the inequality constraint bound for the function $c_i(\boldsymbol{\beta}) \leq 0$, specifically $\beta_1 \leq -0.2$. Left of the plane, $\beta_1 \leq -0.2$ is the feasible region $\mathcal{F}_R$ and accordingly the green arrow points leftward inside the feasible region, normal to the inequality constraint bound plane. In contrast, the infeasible region is $c_1(\beta_1) > 0$ pointed by the red arrow. For the objective function $J(\boldsymbol{\beta})$, the constrained minimum at $\boldsymbol{\beta} = [-0.2 \;\; 0]^T$ and the unconstrained minimum is at $\boldsymbol{\beta} = [0 \;\; 0]^T$.

can have open or closed boundaries and multiple linear constraints further restrict the feasible region in an optimization problem. Examples of multiple constraints and feasible region in one- and two-dimensions are shown in Figures 4.16 and 4.17 respectively. Only linear constraints are used in the scope of this work and geometrically, the constraints in 1D are straight lines, in 2D are planes and in multi-dimensions these are hyperplanes (see e.g. Nocedal and Wright (2006) pg. 327).

Some examples are used in this section to illustrate and describe the impact of constraints on the optimal solution. These examples are organized in the following manner:

- Four separate minimization problems are defined with linear objective functions (see Fig. 4.19).
- Two of those objective functions

$$J(\beta) = i + \beta$$
$$J(\beta) = i - \beta$$

  with different gradient directions are considered. Furthermore, objective function contours will be shown for different values of i.

- In both above cases, an unbounded and a bounded constraint function are considered. The unbounded constraint corresponds to a single inequality $c_1(\beta) \leq 0$ and thus the feasible region is open ended. When an additional constraint $c_2(\beta) \leq 0$ is added to the optimization problem, the feasible region becomes bounded. In other words, individually, the two constraint functions

$$c_1(\beta) : \beta - 8 \leq 0 \quad \text{with feasible region } \mathcal{F}_{R,c_1(\beta)} : -\infty \leq \beta \leq 8$$
$$c_2(\beta) : 2 - \beta \leq 0 \quad \text{with feasible region } \mathcal{F}_{R,c_2(\beta)} : 2 \leq \beta \leq \infty$$

  are both unbounded but when combined

$$c_1(\beta) \text{ and } c_2(\beta) \begin{cases} \beta - 8 \leq 0 \\ 2 - \beta \leq 0 \end{cases},$$

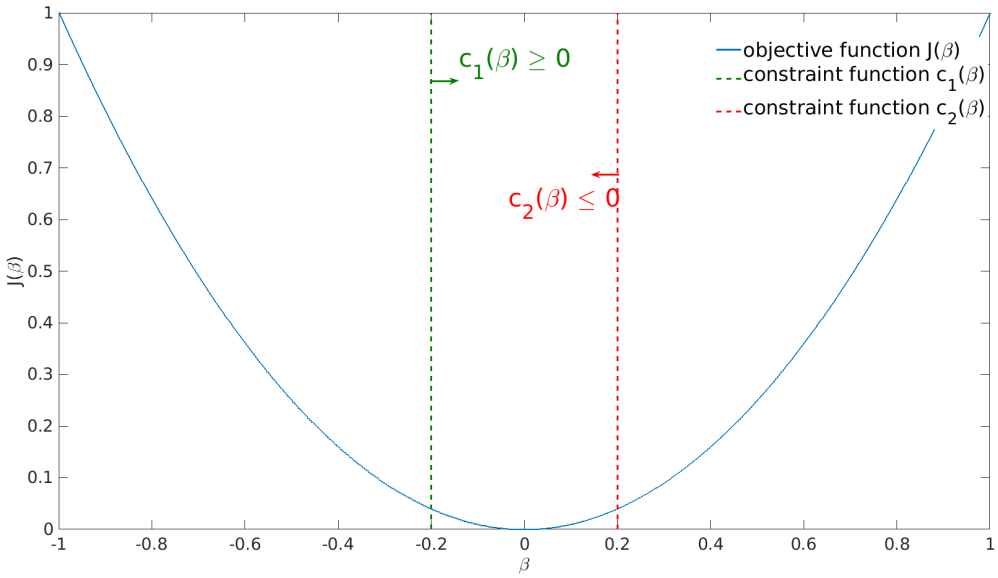


Figure 4.16: Multiple constraints and feasible region in 1D: The dotted green and red line are the inequality constraint bounds for the functions $c_1(\beta) \geq 0$ and $c_2(\beta) \leq 0$, specifically $\beta = -0.2$ and $\beta = 0.2$. For the inequality constraint, $c_1(\beta) \geq 0$, the green arrow points rightward inside its feasible region, normal to the constraint bound. Similarly, $\{\beta \mid c_2(\beta) \leq 0\}$ denotes the feasible region for the other constraint function. Within $-0.2 \leq \beta \leq 0.2$ both constraints are feasible. The two constraints point in opposite directions and therefore comprise a pair of bounded constraints. For the objective function $J(\beta)$, both the constrained and unconstrained minimum is at $\beta = 0$.

they become bounded with

$$\mathcal{F}_{R,c_1(\beta),c_2(\beta)} : 2 \leq \beta \leq 8$$

as their feasible region. It can be seen that the lower bound of $c_1(\beta)$ and the upper bound of $c_2(\beta)$ individually extend to $-\infty$ and ∞ respectively. Therefore, the two constraints are said to be pointing in opposite directions.

Subsequently, in a similar manner, four separate examples with quadratic objective functions are considered (see Figures 4.20 and 4.21).

- Two quadratic objective functions

$$J(\beta) = 3\beta^2 - i$$
$$J(\beta) = -3\beta^2 + i$$

with different gradient directions are considered.

- An unbounded and a bounded constraint function is considered for each of the two above cases. The unbounded constraint corresponds to a single inequality $c_1(\beta) \leq 0$. When an additional constraint $c_2(\beta) \leq 0$ is added to the optimization problem, the feasible region becomes bounded. Individually, the two constraint functions

$$c_1(\beta) : \beta - 0.8 \leq 0 \quad \text{with feasible region } \mathcal{F}_{R,c_1(\beta)} : -\infty \leq \beta \leq 0.8$$
$$c_2(\beta) : 0.2 - \beta \leq 0 \quad \text{with feasible region } \mathcal{F}_{R,c_2(\beta)} : 0.2 \leq \beta \leq \infty$$

are both unbounded but when combined

$$c_1(\beta) \text{ and } c_2(\beta) \begin{cases} \beta - 0.8 \leq 0 \\ 0.2 - \beta \leq 0 \end{cases},$$

they become bounded with $\mathcal{F}_{R,c_1(\beta),c_2(\beta)} : 0.2 \leq \beta \leq 0.8$ as their feasible region.

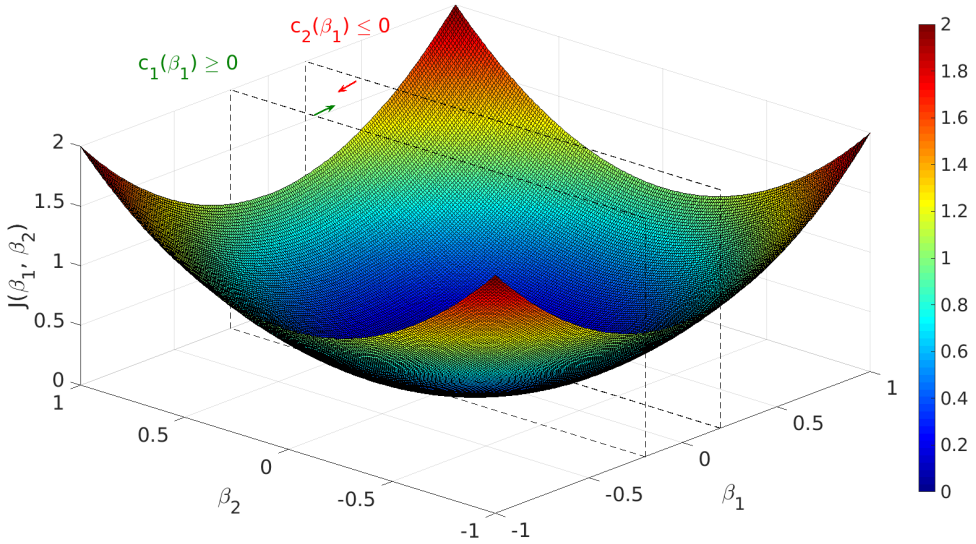


Figure 4.17: Multiple constraints and feasible region in 2D: The two planes $\beta_1 = -0.2$ and $\beta_1 = 0.2$ are the inequality constraint bounds for the functions $c_1(\beta_1) \geq 0$ and $c_2(\beta_1) \leq 0$. Within $-0.2 \leq \beta_1 \leq 0.2$ lies the feasible region. The two constraints point in opposite directions and therefore comprise a pair of bounded constraints along the β_1 axis but remain unbounded along β_2. For the objective function $J(\boldsymbol{\beta})$, both the constrained and unconstrained minimum is at $\boldsymbol{\beta} = [0 \;\; 0]^T$.

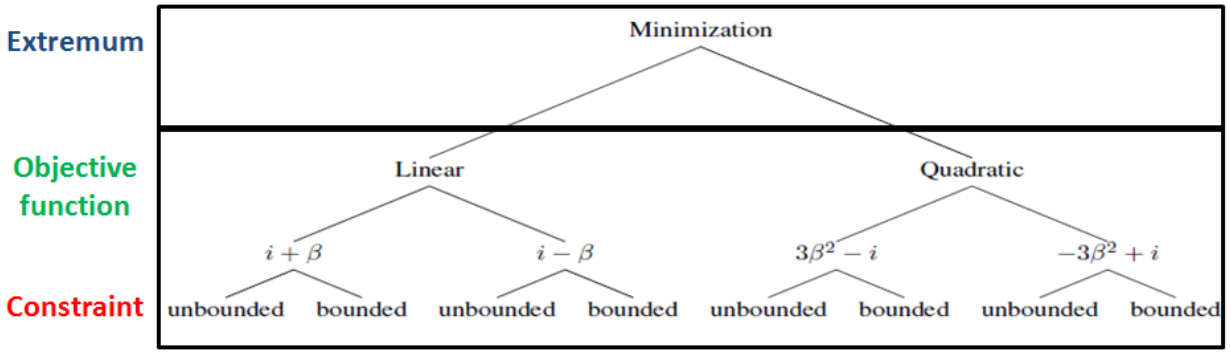


Figure 4.18: Hierarchy of the example minimization problem with linear and quadratic objective functions along with linear constraints

The hierarchy for the above mentioned examples is shown in Fig. 4.18. The examples are now described in detailed and the optimal solution will be worked out. The purpose of this example is to demonstrate that the optimal solution is sensitive to the nature of extrema (maxima or minima), the objective function and its gradient, constraint function direction, total number of constraints and their respective bounds. For the sake of completeness, both the unconstrained and constrained optimum will be indicated along side. This will additionally show that the constrained optimum can be same as an unconstrained one in selected cases. In other words, it will show if a given constraint is active or inactive.

Linear objective function with linear constraints

Fig. 4.19 shows an example of a linear objective function in β, subject to linear constraints, where the two left sub-figures have the objective function $J(\beta) = i + \beta$ and that on right represent the objective function $J(\beta) = i - \beta$. Each distinct objective function contour line, in blue, corresponds to a different value of i. In all the four sub-figures, the vertical and horizontal axes indicate the

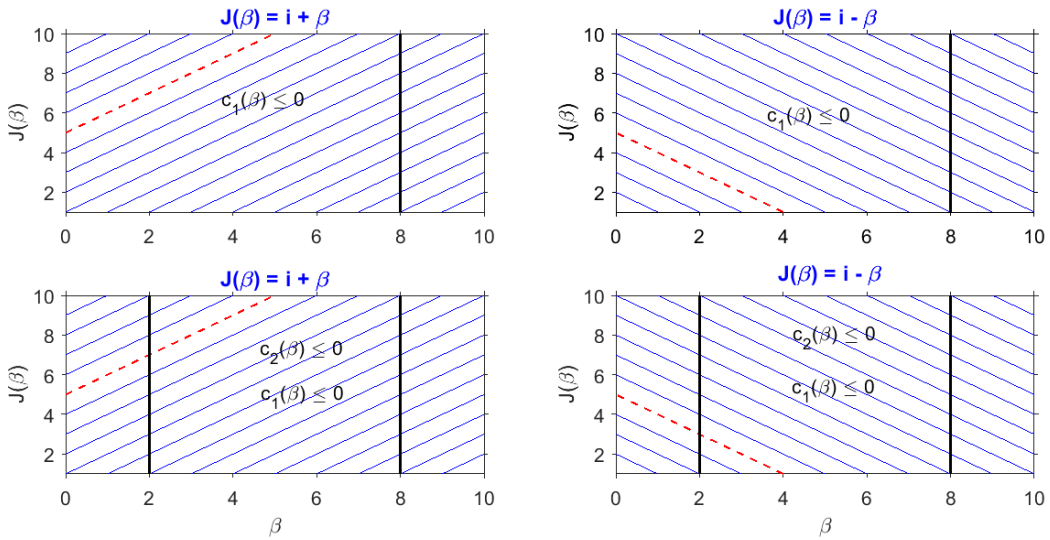


Figure 4.19: Example with two linear objective functions in β subject to linear constraints in 1D, where the two left sub-figures have the objective function $J(\beta) = i + \beta$ and that on right represent the objective function $J(\beta) = i - \beta$. Each distinct objective function contour line, in blue, corresponds to a different value of i. top-left: objective function $J(\beta) = i + \beta$ with one inequality constraint $c_1(\beta) \leq 0$, more specifically $\beta - 8 \leq 0$. The red dashed line refers to the objective function with $i = 5$. bottom-left: Objective function $J(\beta) = i + \beta$ with two constraints $c_1(\beta) \leq 0$ and $c_2(\beta) \leq 0$, more specifically $\beta - 8 \leq 0$ and $2 - \beta \leq 0$. The red dashed line refers to the objective function with $i = 5$. bottom-right: Objective function $J(\beta) = i - \beta$ with two constraints $c_1(\beta) \leq 0$ and $c_2(\beta) \leq 0$, more specifically $\beta - 8 \leq 0$ and $2 - \beta \leq 0$. The red dashed line refers to the objective function with $i = 5$. top-right: Objective function $J(\beta) = i - \beta$ with one constraint $c_1(\beta) \leq 0$, more specifically $\beta - 8 \leq 0$. The red dashed line refers to the objective function with $i = 5$.

objective function and the model parameter β respectively. As always, the objective function gradient points along the direction of increasing objective function.

In addition, there is also a difference between the two top and bottom sub-figures corresponding to the different number of inequality constraints, whose bounds are shown as the solid black line. In each of the four cases, the optimal value of β (optimal solution) is obtained through a graphical inspection. The four different sub-figures and their rationale is explained below.

In the top-left sub-figure of Fig. 4.19, with only one constraint, the optimization problem is unbounded in the direction of decreasing β and clearly, for any value of i, the objective function has a minimum at $\beta = -\infty$.

In the bottom-left sub-figure of Fig. 4.19, with the two constraints pointing in opposite directions, the optimization problem is bounded and for any value of i, the objective function has a minimum at $\beta = 2$. For any other value of β in the feasible region, the objective function has a relatively larger value.

In the top-right sub-figure of Fig. 4.19, with only one constraint, the optimization problem is unbounded but the difference with respect to the top-left is the opposite direction of the objective function gradient. While in top-left, the objective function increases, in top-right it decreases with increasing β. Therefore, for any value of i, the objective function has a minimum at $\beta = 8$. For any other value of β in the feasible region, the objective function has a relatively larger value.

It shall be noted that the inequality constraint $c_2(\beta) \leq 0$ is written as $2 - \beta \leq 0$ or equivalently[27] $\beta - 2 \geq 0$. For any value of i, the objective function has a minimum at $\beta = 8$. For any other value of

[27] Note that multiplying both sides by -1 results in a change of the inequality relation symbol.

β in the feasible region, objective function has a relatively larger value. Table 4.3 summarizes the optimal solution in each of the above four cases.

Table 4.3: Optimal solution for the four different examples of minimization problem with linear objective function and linear constraints shown in Fig. 4.19

Sub-figure in Fig. 4.19	Constrained optimum (minimum)	unconstrained optimum (minimum)
top-left	$-\infty$	$-\infty$
top-right	8	∞
bottom-left	2	$-\infty$
bottom-right	8	∞

Quadratic objective function with linear constraints

The minimization problem is now extended to the quadratic objective function $J(\beta) = 3\beta^2 - i$ for different values of i subject to linear constraints, as shown in Figures 4.20 and 4.21. Again four cases corresponding to different directions of the objective function and inequality constraints. In each of the four cases, both the constrained and unconstrained optimum (minimum) will be worked out by inspection.

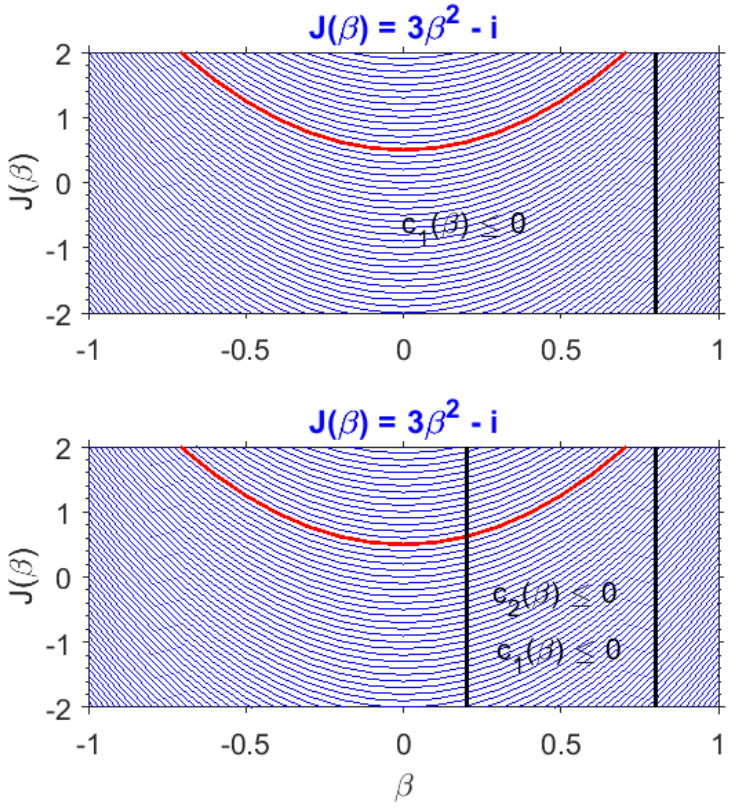


Figure 4.20: Example with quadratic objective function $J(\beta)$ and linear constraints in 1D. The top sub-figure shows the objective function $J(\beta) = 3\beta^2 - i$ with one inequality constraint $c_1(\beta) \leq 0$, more specifically $\beta - 0.8 \leq 0$. The red curve refers to the objective function with $i = -0.5$. The bottom sub-figure shows the objective function $J(\beta) = 3\beta^2 - i$ with two inequality constraints, $c_1(\beta) \leq 0$ and $c_2(\beta) \leq 0$, more specifically $\beta - 0.8 \leq 0$ and $0.2 - \beta \leq 0$. The red curve refers to the objective function with $i = -0.5$.

With only one inequality constraint, the optimization problem is unbounded in the direction of decreasing β and clearly, for any value of i, the objective function has a minimum at $\beta = 0$.

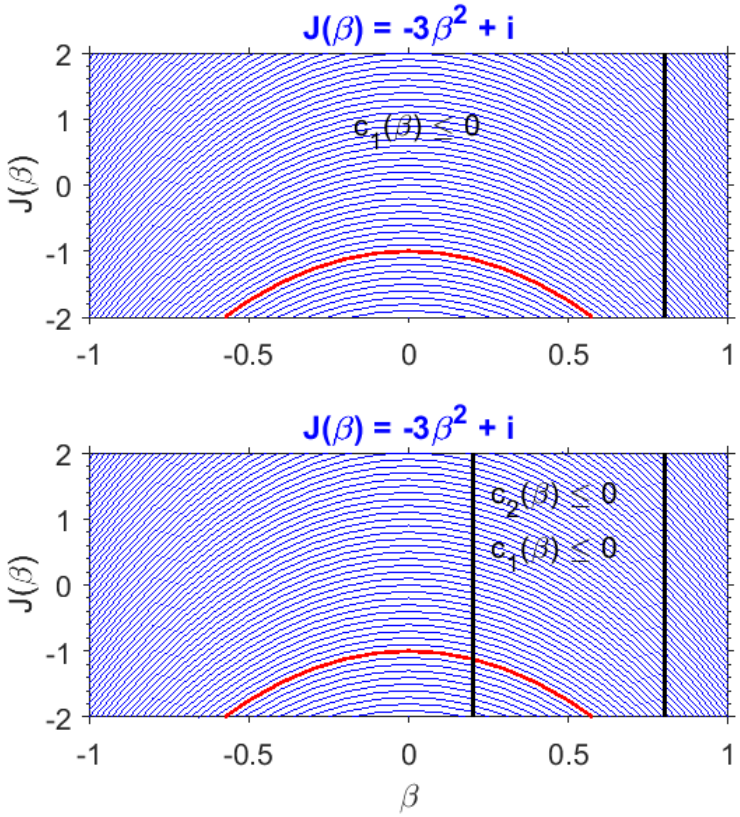


Figure 4.21: Example with quadratic objective function $J(\beta)$ and linear constraints in 1D. The top sub-figure shows the objective function $J(\beta) = -3\beta^2 + i$ with one constraint $c_1(\beta) \leq 0$, more specifically $\beta - 0.8 \leq 0$. The red curve corresponds to the objective function with $i = -1$. Bottom sub-figure shows the objective function $J(\beta) = -3\beta^2 + i$ with two constraints $c_1(\beta) \leq 0$ and $c_2(\beta) \leq 0$, more specifically $\beta - 0.8 \leq 0$ and $0.2 - \beta \leq 0$ respectively. The red curve corresponds to the objective function with $i = -1$. 4.20

With the two constraints pointing in opposite directions, the optimization problem is bounded and for any value of i, the objective function has a minimum at $\beta = 0.2$ as summarized in Table 4.4. For any other value of β in the feasible region, the objective function has a relatively larger value.

In order to show the impact of a change in the objective function gradient on the optimal solution, a modified objective function $J(\beta) = -3\beta^2 + i$ is now considered (see Fig. 4.21). Once again, the same linear inequality constraints $\beta - 0.8 \leq 0$ and $0.2 - \beta \leq 0$ are considered to analyse the optimal solution.

Table 4.4: Optimal solution for the two different examples of minimization problem with quadratic objective function.

Sub-figure of Fig. 4.20	constrained optimum	unconstrained optimum
Top	0	0
Bottom	0.2	0

With only one constraint, the optimization problem is unbounded but the difference with respect to the top sub-figure of Fig. 4.20 is the opposite direction of the objective function gradient. While in the former, the objective function increases, in the latter it decreases with increasing $|\beta|$. Therefore, in this case, for any value of i the objective function has a minimum at $\beta = -\infty$. For any other value

of β in the feasible region, objective function has a relatively larger value. The objective function considered here is non-convex and it shall be noted that in absence of constraints, there will be an inherent ambiguity in determining the objective function minimum. The objective function achieves minima at both $\pm\infty$.

With the two constraints pointing in opposite directions, the optimization problem is bounded and for any value of i, the objective function has a minimum at $\beta = 0.8$. For any other value of β in the feasible region, objective function has a relatively larger value as summarized in Table 4.5.

Table 4.5: Optimal solution for the four different examples of minimization problem with quadratic objective function shown in Fig. 4.21

Sub-figure of Fig. 4.21	constrained optimum	unconstrained optimum
Top sub-figure	$-\infty$	$\pm\infty$ (ambiguous)
Bottom sub-figure	0.8	$\pm\infty$ (ambiguous)

There is a definite relationship between a minimization and the corresponding maximization problem for a given objective function and constraints. It is interesting to analyse how the optimum solution discussed above will change, if we would consider a maximization instead of minimization problem to start with. Without repeating the procedure for inspection, the optimum (maximum) values are summarized in Tables 4.6, 4.7 and 4.8 for constrained and unconstrained solution.

Table 4.6: Optimal solution for the four different examples of maximization problem with linear objective function shown in Fig. 4.19

Sub-figure in Fig. 4.19	constrained optimum (maximum)	unconstrained optimum (maximum)
top-left	8	∞
top-right	$-\infty$	$-\infty$
bottom-left	8	∞
bottom-right	2	$-\infty$

Table 4.7: Optimal solution for the two different examples of maximization problem with quadratic objective function shown in Fig. 4.20

Sub-figure in Fig. 4.20	constrained optimum (maximum)	unconstrained optimum (maximum)
top	$-\infty$	$\pm\infty$ (ambiguous)
bottom	0.8	$\pm\infty$ (ambiguous)

Summary

To summarize, a linear and a quadratic objective function have been considered with two different objective function gradients each. Furthermore, an unbounded and a bounded constraint was considered for each of the objective functions. This resulted in a total of four sub-categories for linear and another four for quadratic objective function. With the help of these 8 optimization problem examples, it has been shown that the optimal solution is sensitive to the following aspects:

- Nature of extremum (minimum or maximum)
- Objective function gradient
- Number of constraints, its bounds and directions

The above examples were solved by a graphical inspection. In the following section, a rigorous mathematical technique will be presented for solving optimization problems with multi-dimensional model parameter and multiple constraints.

Table 4.8: Optimal solution for the two different examples of maximization problem with quadratic objective function shown in Fig. 4.21

Sub-figure in Fig. 4.21	constrained optimum (maximum)	unconstrained optimum (maximum)
top	0	0
bottom	0.2	0

Inequality constrained optimization

Generally, constrained optimization problems are solved by reducing them to unconstrained ones by adding penalty terms to the objective function, that include a penalty factor (more commonly called Lagrange multiplier) and the constraint function (Nocedal and Wright (2006) pg. 341). The sum of the "penalty term" and the objective function is called a Lagrangian[28] function. The phrase penalty signifies that adding any positive term will increase the Lagrangian function instead of minimizing it. If there are both equality and inequality constraints, they are considered as additional independent penalty terms and accordingly there are also separate Lagrange multipliers[29] The work of Roese-Koerner et al. (2012), Roese-Koerner and Schuh (2014), Roese-Koerner et al. (2014) and Roese-Koerner and Schuh (2016) provide a strong motivation to apply inequality constraints in optimization problems.

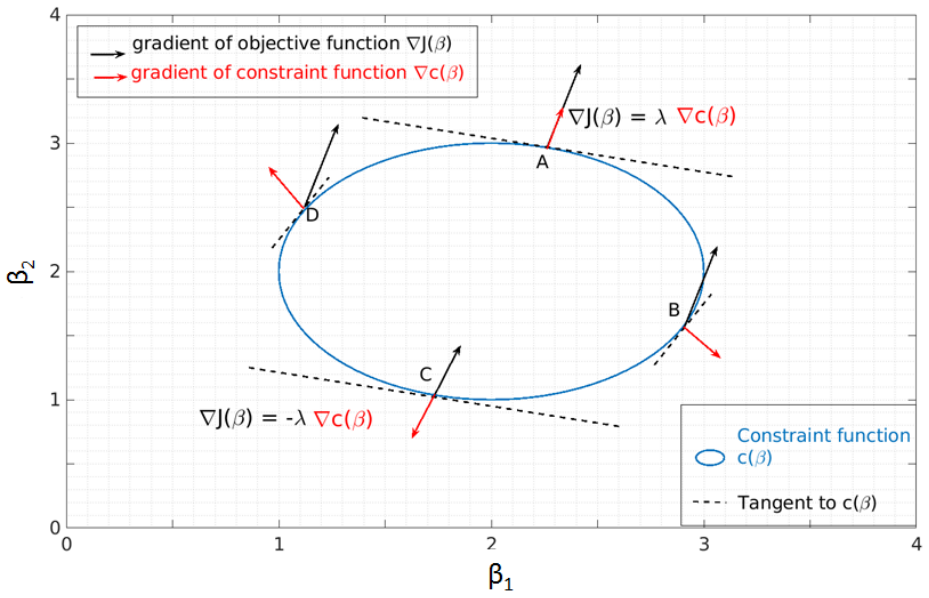


Figure 4.22: Gradient of the objective function and the constraint function. For the ease of readability the objective function contours are not shown in this figure. The tangent to the constraint drawn at the four points are also to be considered as tangent to the objective function.

[28] Joseph-Louis Lagrange, also reported as Giuseppe Luigi Lagrange or Lagrangia, was an Italian mathematician and astronomer, later naturalized French. He made significant contributions to the fields of analysis, number theory, and both classical and celestial mechanics.

[29] The Lagrange multiplier rule for problems with equality constraints appeared in 1797 in Lagrange's Theorie des fonctions analytiques (Fraser (2005)).

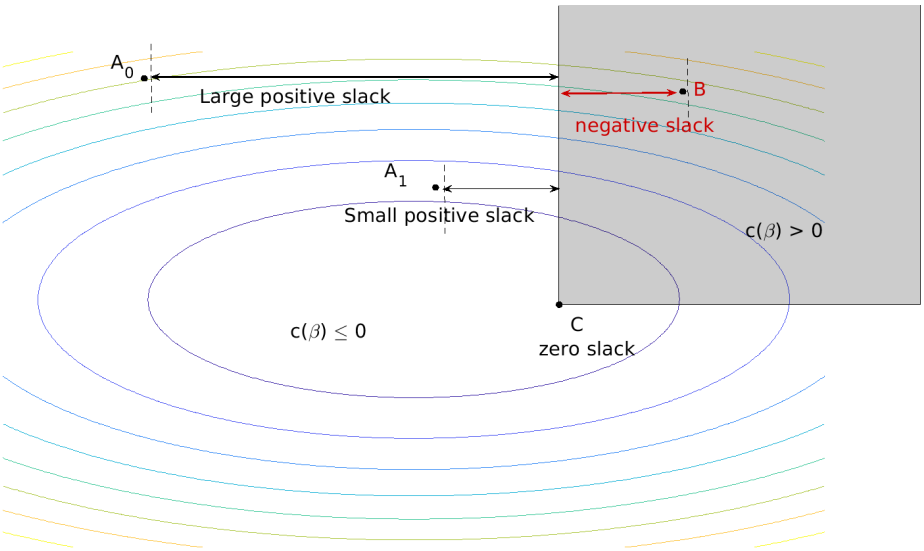


Figure 4.23: Slack variable representation: The white and gray area refer to the feasible and the infeasible region respectively.

Combining objective and constraint functions

The Lagrangian function

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}) = J(\boldsymbol{\beta}) + \sum_i \lambda_i c_i(\boldsymbol{\beta}) + \sum_j \lambda_j c_j(\boldsymbol{\beta}) \qquad i \in \mathcal{E} \ ; \ j \in \mathcal{I}, \tag{4.62}$$

combines the objective function $J(\boldsymbol{\beta})$ with multiple equality and inequality constraints along with their corresponding unknown Lagrange multipliers λ_i and λ_j.

An inequality constraint function $c_j(\boldsymbol{\beta}) \leq 0$ is transformed to an equality

$$c_j(\boldsymbol{\beta}) + \tilde{s}_{ineq,j} = 0 \quad \forall \quad j \in \mathcal{I}, \tag{4.63}$$

where $\tilde{s}_{ineq,j}$ is a slack variable associated with the j^{th} inequality constraint with the condition

$$\tilde{s}_{ineq,j} \geq 0 \quad \forall \quad j \in \mathcal{I}. \tag{4.64}$$

Two important aspects of this transformation are (1) the inequality constraint is no longer on the model parameter vector $\boldsymbol{\beta}$ but on the additional set of unknown parameters introduced i.e. the slack variable and (2) there is still a non-negativity constraint on the slack variable.

Optimization parameters

With $\tilde{\boldsymbol{s}}_{ineq} = \left[\tilde{s}_{ineq,1}, \cdots, \tilde{s}_{ineq,n_{ineq}}\right]^T$, the updated Lagrangian

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\lambda}, \tilde{\boldsymbol{s}}_{ineq}) &= J(\boldsymbol{\beta}) + \sum_{i \in \mathcal{E}} \lambda_i c_i(\boldsymbol{\beta}) + \sum_{j \in \mathcal{I}} \lambda_j \left(c_j(\boldsymbol{\beta}) + \tilde{s}_{ineq,j}\right) \\ &\text{subject to } \tilde{s}_{ineq,j} \geq 0 \ ; \ j \in \mathcal{I} \end{aligned} \tag{4.65}$$

becomes a function of the model parameters, Lagrange multipliers and slack variables, which are all-together called the optimization parameters.

Along with the slack variables, the Lagrange multipliers contain information about the sensitivity of the optimal value of the objective function with respect to changes in the constraints (Boyd and Vandenberghe (2004) pg. 249). Qualitatively, if a Lagrange multiplier corresponding to a constraint $c_i(\boldsymbol{\beta})$ is large, then it adds a corresponding large penalty on the value of the Lagrangian function. A more detailed sensitivity analysis of the Lagrange multiplier will be presented in Section 4.6.

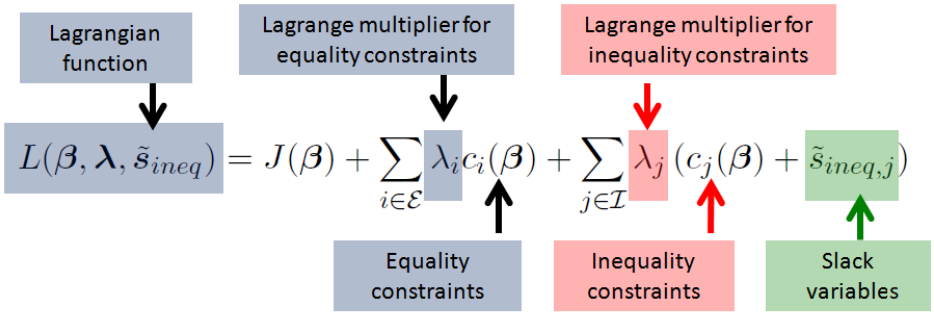


Figure 4.24: Lagrangian formulation with objective function, constraint functions and Lagrange multipliers

A slack variable $\tilde{s}_{ineq,j}$, corresponding to the inequality constraint function $c_j(\boldsymbol{\beta})$, is a "*margin*" to allow any given candidate solution to be moved in the feasible region until it becomes 0, at which the solution is at the boundary of the j^{th} constraint. This is illustrated in Fig. 4.23, where the intermediate solutions A_0 and A_1 are inside, B is outside and C is on the boundary of the feasible region. The slack can also be interpreted equivalently to the "Mahalabonis[30] distance" between an intermediate solution and the constraint function boundary. While the slack variable for the points A_0 and A_1 is positive, that for B is negative and for C is zero.

Furthermore, the partial inequality (4.65) on the slack variable is also removed using its square in the Lagrangian formulation and thereby removing the need to impose any further positive constraint on it. Accordingly, the Lagrangian function

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}, \tilde{\boldsymbol{s}}) = J(\boldsymbol{\beta}) + \sum_{i\in\mathcal{E}} \lambda_i c_i(\boldsymbol{\beta}) + \sum_{j\in\mathcal{I}} \lambda_j \left(c_j(\boldsymbol{\beta}) + \tilde{s}^2_{ineq,j}\right) \tag{4.66}$$

is free from any inequality constraints. The terms in the Lagrangian function are highlighted in Fig. 4.24 for better readability.

In this section, it has been shown that the objective function and the constraints can be combined using Lagrange multipliers. Especially for including inequality constraints, the concept of slack variable has been introduced. The Lagrangian function has been introduced which forms the basis for deriving the equations for estimating the model parameter. Specifically, the Lagrangian function will be minimized by applying the optimality conditions, as described in the next section.

4.5 Optimality conditions

A summary of the important results from extreme value theory is presented as an introductory part to understanding the origins of the optimality conditions. According to extreme value theorem (Protter and MorreyJr. (1970)), a continuous function $f(\boldsymbol{\beta})$, defined on a closed finite interval $[\boldsymbol{\beta}_a, \boldsymbol{\beta}_b]$, attains its extreme (minimum or maximum) value corresponding to an optimal solution $\boldsymbol{\beta}_{opt}$ with

$$\boldsymbol{\beta}_a \leq \boldsymbol{\beta}_{opt} \leq \boldsymbol{\beta}_b, \tag{4.67}$$

which means, $\boldsymbol{\beta}_{opt}$ could be either at the boundaries $\boldsymbol{\beta}_a$, $\boldsymbol{\beta}_b$ or at any "interior point" in between. Furthermore, when $f(\boldsymbol{\beta})$ is differentiable, then according to Fermat's theorem (Stavroudis (2006)),

[30] Prasanta Chandra Mahalanobis (29 June 1893 – 28 June 1972) was an Indian scientist and statistician

the stationarity condition,

$$\left.\frac{\partial f(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_{opt}} = \mathbf{0}_{u\times 1} \tag{4.68}$$

is satisfied at the extremum. However, it is only a necessary condition for the solution $\boldsymbol{\beta}_{opt}$ to be the optimal because $f(\boldsymbol{\beta})$ could have either a maximum or a minimum value. To decide upon the exact nature of the extremum, the second derivative needs to be checked, which if satisfies

$$\left.\frac{\partial^2 f(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_{opt}} > \mathbf{0}_u, \tag{4.69}$$

then $f(\boldsymbol{\beta}_{opt})$ is a minimum value and if

$$\left.\frac{\partial^2 f(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}^2}\right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_{opt}} < \mathbf{0}_u, \tag{4.70}$$

then $f(\boldsymbol{\beta}_{opt})$ is a maximum value. Another closely related and important result, shown by Meyer (1979), is that if condition (4.69) is satisfied for $\boldsymbol{\beta} \in [\boldsymbol{\beta}_a, \boldsymbol{\beta}_b]$, then $f(\boldsymbol{\beta})$ is said to be a convex function.

By extension of extreme value theory, following Protter and MorreyJr. (1970), Lang (1986) and Forst and Hoffmann (2010), every convex function has exactly one minimum value.

However, it is possible that in the neighbourhood of a minimum value, the objective function has almost constant values.

Out of the three aspects of the optimization (see Fig. 4.1), the objective function and the constraints have been described in the previous sections. It has been shown that boundaries of the feasible region are defined by the constraint bounds. However, in case of open-ended constraints[31], the feasible region is unbounded and could lead to a slow convergence or even divergence (Boyd and Vandenberghe (2004) pg. 466). The algorithm needs to traverse in the feasible region only until a set of necessary and sufficient conditions (also called the optimality conditions) are satisfied. These are the so-called Karush-Kuhn-Tucker (KKT) conditions (see Karush (1939), Kuhn and Tucker (1951)), also called first-order conditions[32] (Nocedal and Wright (2006) pg. 320). In the next sub-section, the KKT conditions will be introduced and described in detail.

4.5.1 Karush-Kuhn Tucker (KKT) conditions

Let $\boldsymbol{\beta}_*, \boldsymbol{\lambda}_*, \tilde{\boldsymbol{s}}_*$ be optimal solution to the optimization problem $\mathcal{P}$, then the following optimality conditions

$$\begin{aligned}
\nabla_{\boldsymbol{\beta}} L(\boldsymbol{\beta}_*, \boldsymbol{\lambda}_*, \tilde{\boldsymbol{s}}_*) &= \mathbf{0}_{u\times 1} \\
\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{\beta}_*, \boldsymbol{\lambda}_*, \tilde{\boldsymbol{s}}_*) &= \mathbf{0}_{(n_{ineq}+n_{eq})\times 1} \\
\nabla_{\tilde{\boldsymbol{s}}} L(\boldsymbol{\beta}_*, \boldsymbol{\lambda}_*, \tilde{\boldsymbol{s}}_*) &= \mathbf{0}_{(n_{ineq}+n_{eq})\times 1}
\end{aligned} \tag{4.71}$$

$$c_i(\boldsymbol{\beta}_*) = 0 \quad i \in \mathcal{E} \tag{4.72a}$$
$$c_j(\boldsymbol{\beta}_*) \leq 0 \quad j \in \mathcal{I} \tag{4.72b}$$
$$\lambda_{*,i} \geq 0 \quad i \in \mathcal{E} \;\; ; \;\; \lambda_{*,j} \geq 0 \quad j \in \mathcal{I} \tag{4.72c}$$
$$\lambda_{*,i}\; c_i(\boldsymbol{\beta}_*) = 0 \quad i \in \mathcal{E} \tag{4.72d}$$
$$\lambda_{*,j}\; c_j(\boldsymbol{\beta}_*) = 0 \quad j \in \mathcal{I} \tag{4.72e}$$

[31] Only with an upper or lower bound or no bound

[32] as they contain the first derivative of the Lagrangian function.

are satisfied. The first set of equations (4.71), is called the stationarity condition because at minimum value of the Lagrangian function, its gradient with respect to the model parameter $\boldsymbol{\beta}$ is zero and this relation allows computation of the corresponding optimal values of $\boldsymbol{\beta}$ (also called the stationary points). Analogously, the gradient of the Lagrangian function with respect to the Lagrange multiplier and the slack variable are also set to zero.

The second and the third condition (4.72a, 4.72b) state that the constraints must be satisfied at the optimal solution and therefore these are called the feasibility conditions. The fourth condition (4.72c) comes from the non-negativity of the Lagrange multipliers (see e.g. Nocedal and Wright (2006) pg. 330). Equations (4.72d) and (4.72e) are called the complimentary slack condition because both are so-called "linear complimentary equations", which will be described in this section; see e.g. Murty (1972), Fritsch (1985) and Cottle (2009) for more background on this topic. It shall be noted that the condition (4.72e) can also be equivalently represented as

$$\lambda_{*,j}\tilde{s}_{*,j} = 0 \quad j \in \mathcal{I} \tag{4.73}$$

due to the implicit relation

$$c_j(\boldsymbol{\beta}_*) + \tilde{s}_{*,j} = 0; \quad ;\tilde{s}_{*,j} \geq 0 \quad \forall \quad j \in \mathcal{I}.$$

Now the goal is to obtain the KKT equations from applying the conditions (4.71) through (4.72e). Beginning with the feasibility conditions (4.72a) and (4.72b), the system of constraint functions (4.58) is partitioned into the inequality

$$\boldsymbol{B}_{ineq}\boldsymbol{\beta} - \boldsymbol{b}_{ineq} \leq \mathbf{0}_{n_{ineq}\times 1} \tag{4.74}$$

and the equality

$$\boldsymbol{B}_{eq}\boldsymbol{\beta} - \boldsymbol{b}_{eq} = \mathbf{0}_{n_{eq}\times 1} \tag{4.75}$$

constraint functions in matrix form, where $\boldsymbol{B}_{ineq}$ and $\boldsymbol{B}_{eq}$ are the inequality and equality constraint coefficient matrices respectively. The vectors

$$\boldsymbol{b}_{ineq} = \begin{bmatrix} b_{ineq,1} & \cdots & b_{ineq,n_{ineq}} \end{bmatrix}^T \tag{4.76a}$$

$$\boldsymbol{b}_{eq} = \begin{bmatrix} b_{eq,1} & \cdots & b_{eq,n_{eq}} \end{bmatrix}^T \tag{4.76b}$$

denote the inequality and equality constraint bounds respectively. Furthermore, by introducing the slack variable

$$\tilde{\boldsymbol{s}}_{ineq} = \begin{bmatrix} \tilde{s}_{ineq,1} & \cdots & \tilde{s}_{ineq,n_{ineq}} \end{bmatrix}^T, \tag{4.77}$$

Eqs. (4.63) and (4.74) can be combined as

$$\boldsymbol{B}_{ineq}\boldsymbol{\beta} - \boldsymbol{b}_{ineq} + \tilde{\boldsymbol{s}}_{ineq} = \mathbf{0}_{n_{ineq}\times 1} \qquad \text{subject to } \tilde{\boldsymbol{s}}_{ineq} \geq \mathbf{0}_{n_{ineq}\times 1}. \tag{4.78}$$

In a more general way, the inequality and equality constraints (4.75) and (4.78) are combined and written as

$$\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{b} + \tilde{\boldsymbol{s}} = \mathbf{0}_{n_{ineq}+n_{eq}\times 1} \qquad \text{subject to } \tilde{\boldsymbol{s}} \geq \mathbf{0}_{n_{ineq}+n_{eq}\times 1}, \tag{4.79}$$

where

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_{ineq} \\ \boldsymbol{B}_{eq} \end{bmatrix} \text{ is the constraint coefficient matrix of dimension } (n_{ineq} + n_{eq}) \times u \tag{4.80a}$$

$$\boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_{ineq} \\ \boldsymbol{b}_{eq} \end{bmatrix} \text{ is the constraint bound vector of dimension} (n_{ineq} + n_{eq}) \times 1 \tag{4.80b}$$

$$\tilde{\boldsymbol{s}} = \begin{bmatrix} \tilde{\boldsymbol{s}}_{ineq} \\ \tilde{\boldsymbol{s}}_{eq} \end{bmatrix} \text{ is the cascaded slack vector of dimension} (n_{ineq} + n_{eq}) \times 1 \tag{4.80c}$$

$$\tilde{\boldsymbol{s}}_{eq} = \begin{bmatrix} 0 & \cdots & 0 \end{bmatrix}^T \text{ is a column vector of zeros of dimension } n_{eq} \times 1. \tag{4.80d}$$

The partial inequality on the slack variable, in Eq. (4.79), is removed by considering the vector transformation (see Nocedal and Wright (2006))

$$\boldsymbol{s} = \tilde{\boldsymbol{s}} \odot \tilde{\boldsymbol{s}}, \tag{4.81}$$

where $\odot$ denotes the element-wise vector multiplication operator. Using $\boldsymbol{s}$ instead of $\tilde{\boldsymbol{s}}$, in Eq. (4.79), eliminates the need for any explicit positivity constraint on the slack variable. Substituting $J(\boldsymbol{\beta})$ from Eq. (4.9) into Eq. (4.66) along with the transformed slack variable (4.81) leads to

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{s}) = \frac{1}{\sigma^2} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\beta})^T \boldsymbol{P} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{\beta}) + \boldsymbol{\lambda}^T (\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{b} + \boldsymbol{s}), \tag{4.82}$$

where the inequality and the equality constraint Lagrange multipliers

$$\boldsymbol{\lambda} = \begin{bmatrix} \boldsymbol{\lambda}_{ineq}^T & \boldsymbol{\lambda}_{eq}^T \end{bmatrix}^T \tag{4.83}$$

are brought together with the objective function[33]. It can be shown that using $\boldsymbol{s}$ instead of $\tilde{\boldsymbol{s}}$ in Eq. (4.82) does not affect the optimal solution. This is because any value of $\boldsymbol{\beta}$ which will minimize the Lagrangian function (4.82) will also minimize the corresponding function (4.65). The product on the right hand side of Eq. (4.82) yields,

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{s}) = \frac{1}{\sigma^2} \big[& \boldsymbol{y}^T \boldsymbol{P} \boldsymbol{y} - \boldsymbol{\beta}^T \boldsymbol{A}^T \boldsymbol{P} \boldsymbol{y} - \boldsymbol{y}^T \boldsymbol{P} \boldsymbol{A} \boldsymbol{\beta} + \boldsymbol{\beta}^T \boldsymbol{A}^T \boldsymbol{P} \boldsymbol{A} \boldsymbol{\beta} + \\ & \boldsymbol{\lambda}^T (\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{b} + \boldsymbol{s}) \big] \end{aligned} \tag{4.84}$$

and substituting

$$\boldsymbol{Q} = 2\boldsymbol{A}^T \boldsymbol{P} \boldsymbol{A} \quad \text{and} \quad \boldsymbol{q} = 2\boldsymbol{A}^T \boldsymbol{P} \boldsymbol{y} \tag{4.85}$$

for better readability, Eq. (4.84) becomes

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{s}) &= \frac{1}{\sigma^2} \left[\boldsymbol{y}^T \boldsymbol{P} \boldsymbol{y} - \frac{1}{2} \boldsymbol{\beta}^T \boldsymbol{q} - \frac{1}{2} \boldsymbol{q}^T \boldsymbol{\beta} + \frac{1}{2} \boldsymbol{\beta}^T \boldsymbol{Q} \boldsymbol{\beta} + \boldsymbol{\lambda}^T (\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{b} + \boldsymbol{s}) \right] \\ &= \frac{1}{\sigma^2} \left[\frac{1}{2} \boldsymbol{\beta}^T \boldsymbol{Q} \boldsymbol{\beta} - \left(\boldsymbol{q}^T - \boldsymbol{\lambda}^T \boldsymbol{B} \right) \boldsymbol{\beta} - \boldsymbol{\lambda}^T (\boldsymbol{b} - \boldsymbol{s}) + \boldsymbol{y}^T \boldsymbol{P} \boldsymbol{y} \right] \end{aligned}. \tag{4.86}$$

Following Eq. (4.71), the necessary conditions for the optimality are that the partial derivatives of Lagrangian satisfy

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = \boldsymbol{0}_{u \times 1} \tag{4.87a}$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = \boldsymbol{0}_{(n_{ineq}+n_{eq}) \times 1} \tag{4.87b}$$

$$\frac{\partial L}{\partial \tilde{\boldsymbol{s}}} = \boldsymbol{0}_{(n_{ineq}+n_{eq}) \times 1} \tag{4.87c}$$

and the corresponding sufficient conditions are that the three Hessian matrices satisfy

$$\frac{\partial^2 L}{\partial \boldsymbol{\beta}^2} \geq \boldsymbol{0}_u \tag{4.88a}$$

$$\frac{\partial^2 L}{\partial \boldsymbol{\lambda}^2} \geq \boldsymbol{0}_{n_{ineq}+n_{eq}} \tag{4.88b}$$

$$\frac{\partial^2 L}{\partial \tilde{\boldsymbol{s}}^2} \geq \boldsymbol{0}_{n_{ineq}+n_{eq}} \tag{4.88c}$$

[33] For the sake of completeness we mention that in the special case of slack variable $\boldsymbol{s} = \boldsymbol{0}$ and with Lagrange multiplier $\boldsymbol{\lambda} = \boldsymbol{0}$, the problem becomes an optimization with equality constraints.

or, in other words, the Hessian matrices must be positive semi-definite[34]. Applying the condition from both the first and second derivative of the Lagrangian function with respect to the model parameter $\boldsymbol{\beta}$ from Eqs.(4.87a) and (4.88a) together yields

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = \frac{1}{\sigma^2}\left(\boldsymbol{Q}\boldsymbol{\beta} - \boldsymbol{q} + \boldsymbol{B}^T\boldsymbol{\lambda}\right) \tag{4.89a}$$

$$\frac{\partial^2 L}{\partial \boldsymbol{\beta}^2} = \frac{1}{\sigma^2}\boldsymbol{Q}. \tag{4.89b}$$

From Eq. (4.85), it follows that the Hessian matrix with respect to $\boldsymbol{\beta}$ is at least positive semi-definite. In a similar manner, applying the conditions from both the first and second derivative of the Lagrangian function with respect to the Lagrange multiplier from Eq. (4.87b) and Eq. (4.88b) respectively yields[35]

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = \frac{1}{\sigma^2}\left(\boldsymbol{\beta}^T\boldsymbol{B}^T - \boldsymbol{b}^T + \boldsymbol{s}^T\right) \tag{4.90a}$$

$$\frac{\partial^2 L}{\partial \boldsymbol{\lambda}^2} = \mathbf{0}_{n_{ineq}+n_{eq}}. \tag{4.90b}$$

The Hessian matrix with respect to the Lagrange multipliers is zero and therefore also satisfies the sufficient condition (4.88b). The first derivative of the Lagrangian function with respect to the slack variable yields

$$\begin{aligned}\frac{\partial L}{\partial \tilde{\boldsymbol{s}}} &= \frac{\partial L}{\partial \boldsymbol{s}}\frac{\partial \boldsymbol{s}}{\partial \tilde{\boldsymbol{s}}} \\ &= \frac{1}{\sigma^2}\left(2\boldsymbol{\lambda} \odot \tilde{\boldsymbol{s}}\right) \\ &= \frac{2}{\sigma^2}\begin{bmatrix} \lambda_{ineq,1}\ \tilde{s}_{ineq,1} \\ \lambda_{ineq,2}\ \tilde{s}_{ineq,2} \\ \vdots \\ \lambda_{ineq,n_{ineq}}\ \tilde{s}_{ineq,n_{ineq}} \\ \\ \mathbf{0}_{n_{eq}\times 1} \end{bmatrix},\end{aligned} \tag{4.91}$$

where $\mathbf{0}_{n_{eq}\times 1}$ denotes a column vector of zeroes from the product of equality constraint Lagrange multipliers $\boldsymbol{\lambda}_{eq}$ (see Eq. (4.93b)) and the corresponding slack variable $\boldsymbol{s}_{eq}$. Equation (4.91) leads to the complimentary slack condition, as will be explained in the following section. For the sake of describing the following equations, we define "diag" as an operator which takes a $n \times 1$ vector as input and returns a $n \times n$ matrix with the elements of vector $\boldsymbol{\lambda}$ along its main diagnol. The second derivative of the Lagrangian function with respect to slack variable

$$\begin{aligned}\frac{\partial^2 L}{\partial \tilde{\boldsymbol{s}}^2} &= \frac{2}{\sigma^2}\ \mathrm{diag}(\boldsymbol{\lambda}) \\ &= \frac{2}{\sigma^2}\begin{bmatrix} \begin{bmatrix} \lambda_{ineq,1} & 0 & \cdots & 0 \\ 0 & \lambda_{ineq,2} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_{ineq,n_{ineq}} \end{bmatrix} & \mathbf{0}_{n_{ineq}\times n_{eq}} \\ \\ \mathbf{0}_{n_{eq}\times n_{ineq}} & \begin{bmatrix} \lambda_{eq,1} & 0 & \cdots & 0 \\ 0 & \lambda_{eq,2} & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_{eq,n_{eq}} \end{bmatrix} \end{bmatrix},\end{aligned} \tag{4.92}$$

[34]The notation of $\mathbf{0}_{n_{ineq}}$ is used to denote a square matrix of dimension $n_{ineq} \times n_{ineq}$ with all zeros
[35]It shall be noted that right hand side of Eq. (4.90a) is same as $\frac{1}{\sigma^2}\left(\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{b} + \boldsymbol{s}\right)$

being a diagonal matrix with non-negative Lagrange multipliers along the main diagonal, is also positive semi-definite, where

$$\boldsymbol{\lambda}_{ineq} = \begin{bmatrix} \lambda_{ineq,1} & \cdots & \lambda_{ineq,n_{ineq}} \end{bmatrix}^T \tag{4.93a}$$

$$\boldsymbol{\lambda}_{eq} = \begin{bmatrix} \lambda_{eq,1} & \cdots & \lambda_{eq,n_{eq}} \end{bmatrix}^T \tag{4.93b}$$

and "diag($\boldsymbol{\lambda}$)" refers to a diagonal matrix with the elements $\boldsymbol{\lambda}$ along its main diagonal.
In this section, extreme value theory has been used to describe the optimality conditions for the Lagrangian function. In the next section, the KKT equations will be derived and solved. Two specific approaches, namely, the Lemke's method and the interior point method will be presented along with their relative advantages and disadvantages. The specific challenges introduced due to the complimentary slack condition (4.72e) will be discussed as well.

4.5.2 KKT equations and solution

Now the focus will be on the estimation of the unknown optimization parameters by solving the system of equations (4.87a), (4.87b) and (4.87c). The first system of equation is obtained by setting the first derivative of the Lagrangian with respect to $\boldsymbol{\beta}$ to zero i.e. $\frac{\partial L}{\partial \boldsymbol{\beta}} = 0_{u\times 1}$, and using Eq. (4.89a) gives,

$$\boldsymbol{Q}\ \boldsymbol{\beta} - \boldsymbol{q} + \boldsymbol{B}^T\boldsymbol{\lambda} = \mathbf{0}_{u\times 1}. \tag{4.94}$$

The second system of equations is obtained by setting the gradient of the Lagrangian with respect to the Lagrange multiplier to zero i.e. $\frac{\partial L}{\partial \boldsymbol{\lambda}} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1}$ and using Eq. (4.90a) gives,

$$\boldsymbol{B}\boldsymbol{\beta} - \boldsymbol{b} + \boldsymbol{s} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1}. \tag{4.95}$$

The third system of equations is obtained by setting the gradient of the Lagrangian with respect to the slack variable to zero, i.e. $\frac{\partial L}{\partial \tilde{\boldsymbol{s}}} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1}$, and using Eq. (4.91) gives,

$$2\boldsymbol{\lambda} \odot \tilde{\boldsymbol{s}} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1}. \tag{4.96}$$

Furthermore, Eq. (4.96) leads to

$$\begin{bmatrix} \lambda_{ineq,1} \\ \vdots \\ \lambda_{ineq,n_{ineq}} \\ \lambda_{eq,1} \\ \vdots \\ \lambda_{eq,n_{eq}} \end{bmatrix} \odot \begin{bmatrix} \tilde{s}_{ineq,1} \\ \vdots \\ \tilde{s}_{ineq,n_{ineq}} \\ \tilde{s}_{eq,1} \\ \vdots \\ \tilde{s}_{eq,n_{eq}} \end{bmatrix} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1} \quad \Longrightarrow \quad \begin{bmatrix} \lambda_{ineq,1}\tilde{s}_{ineq,1} \\ \vdots \\ \lambda_{ineq,n_{ineq}}\tilde{s}_{ineq,n_{ineq}} \\ \lambda_{eq,1}\tilde{s}_{eq,1} \\ \vdots \\ \lambda_{eq,n_{eq}}\tilde{s}_{eq,n_{eq}} \end{bmatrix} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1}. \tag{4.97}$$

Adding the elements of the vector on both sides of Eq. (4.97) gives

$$\boldsymbol{\lambda}^T\tilde{\boldsymbol{s}} = 0 \tag{4.98}$$

which is another representation of KKT condition (4.72e).

> Eq. (4.97) is an important necessary condition that relates a Lagrange multiplier with its corresponding slack variable. It means that an element $\lambda_{ineq,j}$ of the Lagrange multiplier vector $\boldsymbol{\lambda}_{ineq}$ is zero unless the corresponding slack variable $\tilde{s}_{ineq,j}$ is zero, also meaning that the corresponding constraint is active (see the definition of active constraint 4.4.1).

More generally, solving for $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{s}}$ in problems of the type where

$$\boldsymbol{\lambda}^T\tilde{\boldsymbol{s}} = 0 \qquad \text{subject to } \boldsymbol{\lambda} \geq \mathbf{0}_{(n_{ineq}+n_{eq})\times 1} \ ; \ \tilde{\boldsymbol{s}} \geq \mathbf{0}_{(n_{ineq}+n_{eq})\times 1}$$

are called linear complimentary problem (LCP) (see e.g., Murty (1972), Fritsch (1985), Cottle (2009)). It has been already explained that in the minimization problem, the Lagrange multiplier shall be positive. Therefore, the consequence of Eq. (4.97) is that

$$\lambda_j \tilde{s}_j = 0 \quad \text{subject to } \tilde{s}_j \geq 0 \quad \forall \quad j \in \mathcal{I}, \tag{4.99}$$

which means the product of each pair of Lagrange multiplier and the associated slack variable must be zero. Particularly, the product of each pair of the inequality constraint Lagrange multipliers $\boldsymbol{\lambda}_{ineq}$ and the associated slack variable $\tilde{\boldsymbol{s}}_{ineq}$ must be zero. The slack variable for equality constraint problems are anyway zero ; see Eq. (4.80d).

Solution of KKT equations

There are different techniques to solve the KKT equations (see e.g. Borwein and Lewis (2000), Gass and Harris (2001), Roese-Koerner (2015)). In this section, the approach from Nocedal and Wright (2006) pg. 481, is described in detail. The basic idea of this approach is that the solution pair $\{\boldsymbol{\lambda}, \tilde{\boldsymbol{s}}\}$ is estimated iteratively using the method of centering and scaling, where $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{s}}$ are initialized as non-zero vector. The complimentary slack condition is satisfied as the optimization parameters converge. A tolerance to the pairwise product of slack variable and Lagrange multiplier

$$\lambda_j \tilde{s}_j = \sigma_{cen} \ \mu_{com} \quad \text{for } j \in \mathcal{I}, \tag{4.100}$$

allows for an efficient traversal in the feasible region using two additional given parameters, namely, the "complimentarity measure" μ_{com} and the "centering parameter" σ_{cen}. This approach has been used and recommended by Mehrotra (1992), Tütüncü et al. (2003), and Nocedal and Wright (2006). The factors μ_{com} and σ_{cen} are mainly responsible for the intermediate solution path taken by the optimization parameters (Nocedal and Wright (2006) pg. 483). Eq. (4.100) is written in matrix form

$$\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{e} - \mu_{com} \ \sigma_{cen}\boldsymbol{e} = \boldsymbol{0}_{(n_{ineq}+n_{eq})\times 1}, \tag{4.101}$$

where

$$\boldsymbol{S} = \begin{bmatrix} \boldsymbol{S}_{ineq} & \boldsymbol{0}_{n_{ineq}\times n_{eq}} \\ \boldsymbol{0}_{n_{eq}\times n_{ineq}} & \boldsymbol{S}_{eq} \end{bmatrix} \quad ; \quad \boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_{ineq} & \boldsymbol{0}_{n_{ineq}\times n_{eq}} \\ \boldsymbol{0}_{n_{eq}\times n_{ineq}} & \boldsymbol{\Lambda}_{eq} \end{bmatrix}$$

$$\boldsymbol{S}_{ineq} = \text{diag}(\tilde{\boldsymbol{s}}_{ineq}) = \begin{bmatrix} \tilde{s}_{ineq,1} & 0 & \cdots & 0 \\ 0 & \tilde{s}_{ineq,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \tilde{s}_{ineq,n_{ineq}} \end{bmatrix}$$

$$\boldsymbol{\Lambda}_{ineq} = \text{diag}(\boldsymbol{\lambda}_{ineq}) = \begin{bmatrix} \lambda_{ineq,1} & 0 & \cdots & 0 \\ 0 & \lambda_{ineq,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_{ineq,n_{ineq}} \end{bmatrix}$$

$$\boldsymbol{S}_{eq} = \text{diag}(\boldsymbol{s}_{eq}) = \begin{bmatrix} \tilde{\boldsymbol{s}}_{eq,1} & 0 & \cdots & 0 \\ 0 & \tilde{\boldsymbol{s}}_{eq,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \tilde{\boldsymbol{s}}_{eq,n_{eq}} \end{bmatrix}$$

$$\boldsymbol{\Lambda}_{eq} = \text{diag}(\boldsymbol{\lambda}_{eq}) = \begin{bmatrix} \lambda_{eq,1} & 0 & \cdots & 0 \\ 0 & \lambda_{eq,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_{eq,n_{eq}} \end{bmatrix} \qquad \boldsymbol{e} = \boldsymbol{1}_{(n_{ineq}+n_{eq})\times 1}.$$

The optimization parameters are estimated using either GD (already been discussed in detail in Section 4.3.2) or NM. In this section, the application of NM is shown to solve the KKT equations

$$\begin{aligned}
&\boldsymbol{Q}\ \boldsymbol{\beta} - \boldsymbol{q} + \boldsymbol{B}^T\boldsymbol{\lambda} = \mathbf{0}_{u\times 1} \leftarrow \text{ Stationarity condition} \\
&\boldsymbol{B}\ \boldsymbol{\beta} - \boldsymbol{b} + \boldsymbol{s} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1} \leftarrow \text{ Feasibility condition} \\
&\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{e} - \sigma_{cen}\mu_{com}\boldsymbol{e} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1} \leftarrow \text{ Complimentary slack condition .}
\end{aligned} \tag{4.102}$$

For better readability and consistency, we introduce $\tilde{\boldsymbol{\beta}}$

$$\tilde{\boldsymbol{\beta}} = \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda} \\ \tilde{\boldsymbol{s}} \end{bmatrix} \tag{4.103}$$

with the initial values

$$\tilde{\boldsymbol{\beta}}_0 = \begin{bmatrix} \boldsymbol{\beta}_0 \\ \boldsymbol{\lambda}_0 \\ \tilde{\boldsymbol{s}}_0 \end{bmatrix}, \tag{4.104}$$

so that the descend direction (correction to the optimization parameters)

$$\Delta\tilde{\boldsymbol{\beta}} = \begin{bmatrix} \Delta\boldsymbol{\beta} \\ \Delta\boldsymbol{\lambda} \\ \Delta\tilde{\boldsymbol{s}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\beta} - \boldsymbol{\beta}_0 \\ \boldsymbol{\lambda} - \boldsymbol{\lambda}_0 \\ \tilde{\boldsymbol{s}} - \tilde{\boldsymbol{s}}_0 \end{bmatrix} \tag{4.105}$$

can be estimated by solving

$$\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{B}^T & \mathbf{0}_u \\ \boldsymbol{B} & \mathbf{0}_{n_{ineq}+n_{eq}} & \boldsymbol{I} \\ \mathbf{0}_u & \boldsymbol{S}_0 & \boldsymbol{\Lambda}_0 \end{bmatrix} \begin{bmatrix} \Delta\boldsymbol{\beta} \\ \Delta\boldsymbol{\lambda} \\ \Delta\tilde{\boldsymbol{s}} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{Q}\boldsymbol{\beta}_0 + \boldsymbol{q} - \boldsymbol{B}^T\boldsymbol{\lambda}_0 \\ -\boldsymbol{B}\boldsymbol{\beta}_0 + \boldsymbol{b} - \boldsymbol{s}_0 \\ -\boldsymbol{S}_0\boldsymbol{\Lambda}_0\boldsymbol{e} + \sigma_{cen}\mu_{com}\boldsymbol{e} \end{bmatrix}, \tag{4.106}$$

where $\boldsymbol{Q}$, $\boldsymbol{B}$, $\boldsymbol{q}$ and $\boldsymbol{b}$ are known, $\boldsymbol{S}_0$ and $\boldsymbol{\Lambda}_0$ are evaluated at the current iterates $\tilde{\boldsymbol{s}}_0$ and $\boldsymbol{\lambda}_0$.

The above mentioned approach to solve the KKT equations is the essence of the so-called interior point method (IPM). It is an extension of the basic descent method (see Section 4.3.1) with advanced descent direction and step size estimation methods to ensure that the KKT conditions are satisfied (Nocedal and Wright (2006) pg. 480, Boyd and Vandenberghe (2004) pg. 561). Specifically, the most important modification compared to the basic descent method is the use of variable step size as well as the implicit solution of the linear complimentarity problem (LCP).

In summary, the system of KKT equations (4.102) are linearized about the initial values and solved using the NM. Positivity conditions on the Lagrange multiplier and slack variable lead to the development of linear complimentarity problem. It shall be noted that the performance of many IPM implementations differ considerably depending on the computational complexity of a given optimization problem, the nature of objective function (its convexity, dimensionality, etc.). Therefore, to make any general claim on the performance of IPM is not realistic. The reader is referred to Mehrotra (1992), Anderson et al. (1996), Billups and Ferris (1996), Potra and Wright (2000), Boyd and Vandenberghe (2004), Nocedal and Wright (2006), Mitchell et al. (2006), Forst and Hoffmann (2010) for different implementations and applications of IPM.

The next sub-section will focus on the two important sub-problems in solution of KKT equations, namely the estimation of the descent direction and the corresponding step size.

4.5.3 Descent direction and step size estimation

As introduced in the previous section, the step size is one of the parameters used to compute the updated values of the optimization parameters. In an unconstrained optimization problem, a ”full

step" can be taken (equivalently $\boldsymbol{T}_{it} = \boldsymbol{I}$) along the descent direction. However, the choice of the step size in presence of inequality constraints is not trivial, due to the risk of overshooting the feasible region or specifically violation of the complimentary slack condition (4.72c).

Optimization parameter is updated from iteration it to $it+1$

$$\begin{bmatrix}\boldsymbol{\beta}\\ \boldsymbol{\lambda}\\ \tilde{\boldsymbol{s}}\end{bmatrix}_{it+1} = \begin{bmatrix}\boldsymbol{\beta}\\ \boldsymbol{\lambda}\\ \tilde{\boldsymbol{s}}\end{bmatrix}_{it} + \boldsymbol{T}_{it}\begin{bmatrix}\Delta\boldsymbol{\beta}\\ \Delta\boldsymbol{\lambda}\\ \Delta\tilde{\boldsymbol{s}}\end{bmatrix}_{it} \tag{4.107}$$

until convergence, where

$$\begin{aligned}
\boldsymbol{T}_{it} &= \begin{bmatrix}\boldsymbol{T}_{\beta} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{T}_{\lambda_{ineq}} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{T}_{\lambda_{eq}} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{T}_{\tilde{s}_{ineq}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{T}_{\tilde{s}_{eq}}\end{bmatrix}\\
\boldsymbol{T}_{\beta} &= \begin{bmatrix}t_{\beta_1} & 0 & \cdots & 0\\ 0 & t_{\beta_2} & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \cdots & t_{\beta_u}\end{bmatrix} \quad \boldsymbol{T}_{\lambda_{ineq}} = \begin{bmatrix}t_{\lambda_{ineq,1}} & 0 & \cdots & 0\\ 0 & t_{\lambda_{ineq,2}} & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \cdots & t_{\lambda_{ineq,n_{ineq}}}\end{bmatrix}\\
\boldsymbol{T}_{\tilde{s}_{ineq}} &= \begin{bmatrix}t_{\tilde{s}_{ineq,1}} & 0 & \cdots & 0\\ 0 & t_{\tilde{s}_{ineq,2}} & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \cdots & t_{\tilde{s}_{ineq,n_{ineq}}}\end{bmatrix} \quad \boldsymbol{T}_{\lambda_{eq}} = \begin{bmatrix}t_{\lambda_{eq,1}} & 0 & \cdots & 0\\ 0 & t_{\lambda_{eq,2}} & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \cdots & t_{\lambda_{eq,n_{eq}}}\end{bmatrix}\\
\boldsymbol{T}_{\tilde{s}_{eq}} &= \begin{bmatrix}t_{\tilde{s}_{eq,1}} & 0 & \cdots & 0\\ 0 & t_{\tilde{s}_{eq,2}} & \cdots & 0\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \cdots & t_{\tilde{s}_{eq,n_{eq}}}\end{bmatrix}; \; \boldsymbol{T}_{\lambda} = \begin{bmatrix}\boldsymbol{T}_{\lambda_{ineq}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{T}_{\lambda_{eq}}\end{bmatrix}; \; \boldsymbol{T}_{s} = \begin{bmatrix}\boldsymbol{T}_{\tilde{s}_{ineq}} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{T}_{\tilde{s}_{eq}}\end{bmatrix}
\end{aligned} \tag{4.108}$$

such that the diagonal elements[36] of $\boldsymbol{T}_{it}$ are the respective step sizes used along $\boldsymbol{\beta}$, $\boldsymbol{\lambda}$ and $\tilde{\boldsymbol{s}}$ and satisfy

$$\begin{aligned}
\left[t_{\beta_1}, \cdots, t_{\beta_u}\right] &\in (0,1]\\
\left[t_{\lambda_{ineq,1}}, \cdots, t_{\lambda_{ineq,n_{ineq}}}, \cdots, t_{\lambda_{eq,n_{eq}}}\right] &\in (0,1]\\
\left[t_{\tilde{s}_{ineq,1}}, \cdots, t_{\tilde{s}_{ineq,n_{ineq}}}, \cdots, t_{\tilde{s}_{eq,n_{eq}}}\right] &\in (0,1].
\end{aligned}$$

Although the slack variable corresponding to equality constraints are included in $\boldsymbol{T}_{it}$, they can be omitted from Eq. (4.108). Since the non-negativity of the Lagrange multiplier and slack variable have to be checked in each iteration, the chosen step size along the descent direction is an important consideration. This is one of the main modifications compared to the classical Newton's method (see Section 4.3.3).

The descent direction $[\Delta\boldsymbol{\beta}^T \;\; \Delta\boldsymbol{\lambda}^T \;\; \Delta\tilde{\boldsymbol{s}}^T]^T$ is estimated by solving (4.106) with the initial value $\boldsymbol{T}_{it} = \boldsymbol{I}$ as a starting point for the simultaneous estimation of both the descent direction and step size. The step size estimation sub-problem is defined as follows. Estimate the largest diagonal elements of $\boldsymbol{T}_{it}$

$$\begin{aligned}
t_{\beta_j} \quad & j \in \{1, \cdots, u\}\\
t_{\lambda_k}, t_{s_k} \quad & k \in \{1, \cdots, n_{ineq} + n_{eq}\},
\end{aligned}$$

which when used in Eq. (4.107), results in Lagrange multiplier and slack variable satisfying the KKT conditions (4.72c) and (4.72d) or mathematically

$$\boldsymbol{T}_{it} = \text{maximize}\left\{t_{\beta_j}, t_{\lambda_k}, t_{\tilde{s}_k} \;\middle|\; (\tilde{\boldsymbol{s}} + \boldsymbol{T}_s\Delta\tilde{\boldsymbol{s}} \geq \boldsymbol{0}^{37}); (\boldsymbol{\lambda} + \boldsymbol{T}_{\lambda}\Delta\boldsymbol{\lambda} \geq \boldsymbol{0})\right\}, \tag{4.109}$$

[36] for readability, the subscripts of $\boldsymbol{0}$ matrices are not indicated.

such that $0 < \{t_{\beta_j}, t_{\lambda_k}, t_{\tilde{s}_k}\} \leq 1$; $j = \{1, \cdots, u\}, k = \{1, \cdots, n_{ineq} + n_{eq}\}$, where $\Delta\tilde{\boldsymbol{s}}$ and $\Delta\boldsymbol{\lambda}$ are the descent directions along the slack variable and the Lagrange multiplier. There are three sets of optimization parameters and in the most general case, different step size matrices corresponding to each of the three sets have to be chosen. However, during this work, it was found that using two step size matrices, i.e. one for the unknown model parameter $\boldsymbol{\beta}$ and the other for the Lagrange multiplier and slack variable $\{\boldsymbol{\lambda}, \tilde{\boldsymbol{s}}\}$ pair, was already sufficient. In other words $\boldsymbol{T}_\lambda$ and $\boldsymbol{T}_s$ are assumed to be equal and the reason for this assumption is that $\{\boldsymbol{\lambda}, \tilde{\boldsymbol{s}}\}$ together define the complimentary slack condition. These choices are made as part of the design aspects[38] of the algorithm. In general, the step sizes are not only different for the optimization parameters but also estimated in each iteration[39]. This aspect is one of the main challenges in adding more constraints or other unknown parameters to the modelling problem (see Nesterov and Nemirovskii (1994), Todd et al. (1998), Potra and Wright (2000), Halldórsson and Tütüncü (2003), Mitchell et al. (2006)). From the experience of this work, it was found that an approximate step size is sufficient for the algorithm. Having estimated the step size matrix, following Tütüncü et al. (2003), the complementarity measure

$$\mu_{com} = (\boldsymbol{\lambda} + \boldsymbol{T}_\lambda \Delta\boldsymbol{\lambda})^T (\tilde{\boldsymbol{s}} + \boldsymbol{T}_s \Delta\tilde{\boldsymbol{s}})/n_{ineq} \tag{4.110}$$

is computed, where n_{ineq} is the total number of inequality constraints. The iterative algorithm for combined step size and descent direction estimation is summarized below and the dataflow is shown in Fig. 4.25.

Step 1 : Set the initial values $\boldsymbol{\beta}_0$, $\tilde{\boldsymbol{s}}_0$, $\boldsymbol{\lambda}_0$ of the optimization parameters, the descent directions $\Delta\tilde{\boldsymbol{s}}$ and $\Delta\boldsymbol{\lambda}$ with $\tilde{\boldsymbol{s}}_0 > \boldsymbol{0}$, $\boldsymbol{\lambda}_0 > \boldsymbol{0}$.
Step 2 : Initialize the centering parameter σ_{cen} and use identity matrix as the initial step size; i.e. $\boldsymbol{T}_{it} = \boldsymbol{I}$. Compute μ_{com} using Eq. (4.110), this value of complimentarity measure is called $\mu_{com,full}$ because it corresponds to a "full step"[a].
Repeat (Step 3 to Step 7) until convergence:
Step 3 : Estimate $\Delta\boldsymbol{\beta}$, $\Delta\boldsymbol{\lambda}$ and $\Delta\tilde{\boldsymbol{s}}$ from (4.106) with initial values $\boldsymbol{\beta}_0$, $\boldsymbol{\lambda}_0$, $\tilde{\boldsymbol{s}}_0$, $\boldsymbol{T}_{it}$ and the tuning parameters μ_{com}, σ_{cen}.
Step 4 : Estimate the step size $\boldsymbol{T}_{it}$ from Eq. (4.109).
Step 5 : Compute the complimentarity measure μ_{com} using Eq. (4.110) with the updated of $\boldsymbol{T}_{it}$ from step 4.
Step 6 : Set centering parameter $\sigma_{cen} = (\mu_{com,full}/\mu_{com})^3$ (Tütüncü et al. (2003)).
Step 7 : Compute the updated optimization parameters with Eq. (4.107) using $\boldsymbol{T}_{it}$ computed in the step 4.

[a] $\boldsymbol{T}_{it} = \boldsymbol{I}$

The complexity in step size estimation depends on the dimensionality of the problem (number of optimization parameters) including the number of constraints.

4.6 Sensitivity of Lagrange multiplier

The goal of this section is to establish a relation between the gradients of the objective function and the constraint. This relationship is the basis for the Lagrangian formulation and allows interpretation

[38] For problems with smaller (e.g. 10 or even less) unknown parameters, a constant step size for all three set of parameters can be used.
[39] During implementation, it was found that at each iteration, it may not be practical and also not necessary to estimate the step size matrix. This depends on the progression of descent direction and the convergence in optimization algorithm.

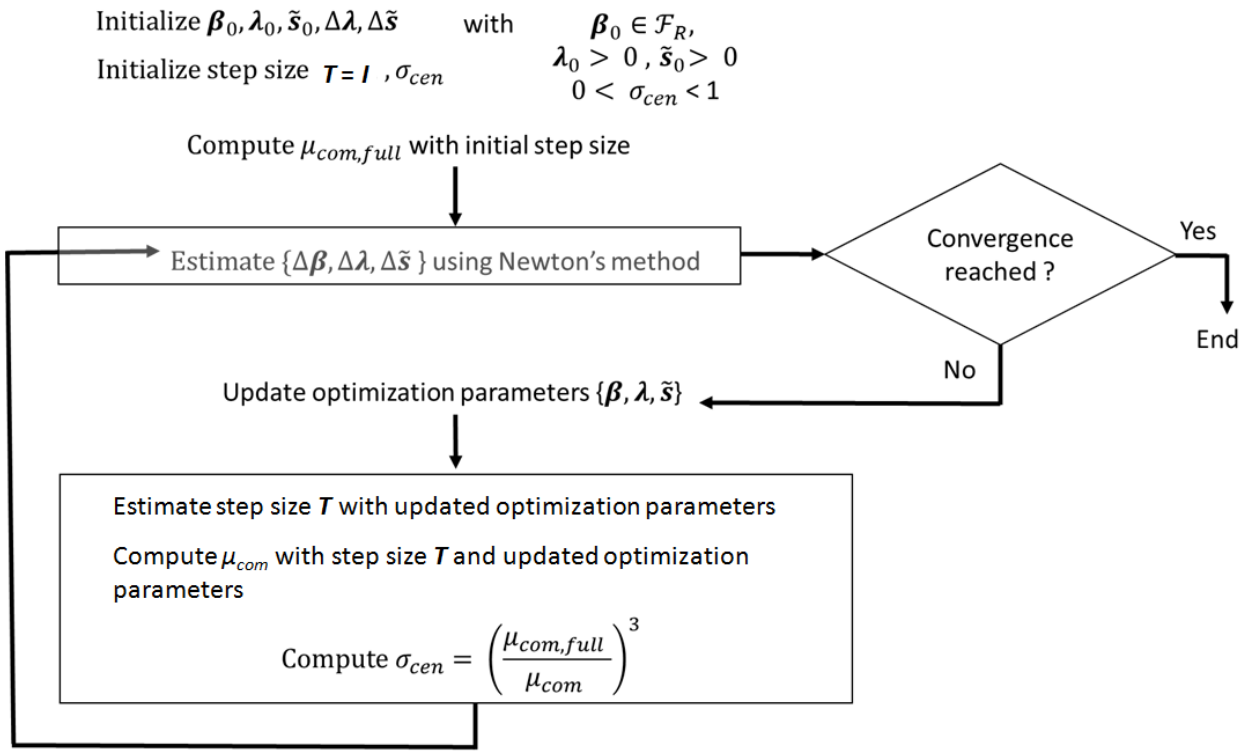


Figure 4.25: Dataflow in interior point method for simultaneous estimation of the step size and descent direction

of the Lagrange multiplier. For the ease of readability, we reproduce Eq. (4.25), where a general formulation for a constrained optimization problem $\mathcal{P}$, for minimization of the objective function

$$\begin{aligned} &\min_{\boldsymbol{\beta}} J(\boldsymbol{\beta}) \\ \text{subject to}: \; &c_i(\boldsymbol{\beta}) \leq 0 \quad i \in \mathcal{E} \\ &c_j(\boldsymbol{\beta}) \leq 0 \quad j \in \mathcal{I} \end{aligned}$$

was introduced. Using the linear approximation property of the total derivative (see e.g Knight and Adams (1975), Goldstine (1980), Koch (1999), Stavroudis (2006)), if

$$\Delta\boldsymbol{\beta} = \begin{bmatrix} \Delta\beta_1 & \Delta\beta_2 & \cdots & \Delta\beta_u \end{bmatrix}^T \approx \mathbf{0}_{u\times 1} \tag{4.111}$$

then

$$\left(\frac{dJ(\boldsymbol{\beta})}{d\boldsymbol{\beta}}\right)^T \Delta\boldsymbol{\beta} \approx J(\boldsymbol{\beta} + \Delta\boldsymbol{\beta}) - J(\boldsymbol{\beta}). \tag{4.112}$$

We begin with the total derivative of the objective function

$$\frac{dJ(\boldsymbol{\beta})}{d\boldsymbol{\beta}} = \begin{bmatrix} \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_1} & \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_2} & \ldots & \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_u}^T, \end{bmatrix} \tag{4.113}$$

which after rearranging to the differential form gives

$$\begin{aligned} dJ(\boldsymbol{\beta}) &= \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_1} d\beta_1 + \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_2} d\beta_2 + \cdots + \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_u} d\beta_u \\ &= \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_1} d\beta_1 + \sum_{q=2}^{u} \frac{\partial J(\boldsymbol{\beta})}{\partial\beta_q} d\beta_q. \end{aligned} \tag{4.114}$$

Similarly,

$$dc_k(\boldsymbol{\beta}) = \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_1} d\beta_1 + \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_2} d\beta_2 + \cdots + \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_u} d\beta_u \tag{4.115}$$

is the differential of k^{th} constraint function $\forall\ k \in \{\mathcal{E} \cup \mathcal{I}\}$. Two general cases for the constraints are possible. Firstly, for strict equality constraint $c_k(\boldsymbol{\beta}) = 0$, its derivative is also zero. Secondly, for the strict inequality constraint $c_k(\boldsymbol{\beta}) < 0$, the model parameter vector is in the feasible region, and therefore does not effect the constraints, which means the differential of the constraint in both cases is zero (see e.g. Borwein and Lewis (2000), Cloud et al. (2014)). Substituting $dc_k(\boldsymbol{\beta}) = 0$ in Eq. (4.115) gives

$$\frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_1} d\beta_1 + \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_2} d\beta_2 + \cdots + \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_u} d\beta_u = 0 \tag{4.116}$$

and further rearrangement leads to

$$\begin{aligned} d\beta_1 &= \frac{-1}{\frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_1}} \left[\frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_2} d\beta_2 + \cdots + \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_u} d\beta_u \right] \\ &= \frac{-1}{\frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_1}} \sum_{q=2}^{u} \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_q} d\beta_q. \end{aligned} \tag{4.117}$$

Substituting $d\beta_1$ from Eq. (4.117) into (4.114) gives,

$$dJ(\boldsymbol{\beta}) = \frac{\partial J(\boldsymbol{\beta})}{\partial \beta_1} \left(\frac{-1}{\frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_1}} \right) \left(\sum_{q=2}^{u} \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_q} d\beta_q \right) + \sum_{q=2}^{u} \frac{\partial J(\boldsymbol{\beta})}{\partial \beta_q} d\beta_q \tag{4.118}$$

and substituting

$$\lambda_k = -\frac{\frac{\partial J(\boldsymbol{\beta})}{\partial \beta_1}}{\frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_1}} \tag{4.119}$$

in Eq. (4.118) leads to

$$dJ(\boldsymbol{\beta}) = \lambda_k \sum_{q=2}^{u} \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_q} d\beta_q + \sum_{q=2}^{u} \frac{\partial J(\boldsymbol{\beta})}{\partial \beta_q} d\beta_q. \tag{4.120}$$

Applying first order optimality condition (4.71),

$$dJ(\boldsymbol{\beta})\Big|_{\boldsymbol{\beta}_{opt}} = 0 \tag{4.121}$$

at $\boldsymbol{\beta} = \boldsymbol{\beta}_{opt}$ on Eq. (4.120), gives

$$-\lambda_k \sum_{q=2}^{u} \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_q}\Big|_{\boldsymbol{\beta}_{opt}} d\beta_q = \sum_{q=2}^{u} \frac{\partial J(\boldsymbol{\beta})}{\partial \beta_q}\Big|_{\boldsymbol{\beta}_{opt}} d\beta_q \tag{4.122}$$

and further rearranging gives

$$\sum_{q=2}^{u} \frac{\partial J(\boldsymbol{\beta})}{\partial \beta_q}\Big|_{\boldsymbol{\beta}_{opt}} d\beta_q + \lambda_k \sum_{q=2}^{u} \frac{\partial c_k(\boldsymbol{\beta})}{\partial \beta_q}\Big|_{\boldsymbol{\beta}_{opt}} d\beta_q = 0 \tag{4.123}$$

or

$$\nabla_{\beta} J(\boldsymbol{\beta})\Big|_{\boldsymbol{\beta}_{opt}} + \lambda_k \nabla_{\beta} c_k(\boldsymbol{\beta})\Big|_{\boldsymbol{\beta}_{opt}} = \mathbf{0}_{u\times 1}, \tag{4.124}$$

which is the same as the derivative of the Lagrangian function with respect to the model parameter vector $\boldsymbol{\beta}$, using the Eqs. (4.66), (4.71) and (4.87a). It shows that the rate of change of the

k^{th} constraint function is related to that of the objective function through the corresponding k^{th} Lagrange multiplier. This analogy of Lagrange multiplier applies to both the equality and inequality constraints and has been already geometrically illustrated in Fig. 4.22. Another important aspect is the sensitivity of the objective function to changes in constraint. This can be described from Eq. (4.124) that can have three general possibilities, depending on the sign of Lagrange multiplier, as mentioned below:

Case 1: $\lambda_k = 0$: Any change in the constraint bound will not result in a corresponding change on the objective function.

Case 2: $\lambda_k > 0$: An increase in the constraint bound will result in a corresponding decrease in the objective function. For minimization problems, as considered in this thesis, this property is highly desirable and this is also the reason for applying the non-negative constraints on the Lagrange multiplier ; see Eq. (4.72c).

Case 3: $\lambda_k < 0$: An increase in the constraint bound will result in a corresponding increase of the objective function. In contrast to the case 2 above, this property would be highly desirable for maximization problems.

4.7 Alternative techniques in constraint optimization

In Section 4.5, the KKT equations have been derived and their solution based on the interior point method was presented. The goal of this section is two-fold:

- Two alternative approaches to the solution of KKT equations. The first one is based on a transformation of the KKT equations into a so-called linear complimentarity problem. Second alternative is based on the concept of duality theory, where the Lagrangian function (4.82) is transformed to its dual form and then the same optimality conditions, as shown in Section 4.5, will be applied.
- An alternative approach to step size estimation based on predicted optimization parameters using the concept of non-linear ARMA.

4.7.1 Linear complimentarity problem (LCP)

This section describes an alternate approach (based on Koch (1985)) to solving the KKT equations. At the end of this section, a comparison between this approach and the method of centering and scaling (Nocedal and Wright (2006) pg. 484) will be provided.

We begin this approach from the KKT equations,

$$\boldsymbol{Q}\ \boldsymbol{\beta} - \boldsymbol{q} + \boldsymbol{B}_{eq}^T\boldsymbol{\lambda}_{eq} + \boldsymbol{B}_{ineq}^T\boldsymbol{\lambda}_{ineq} = \mathbf{0}_{u\times 1} \tag{4.125a}$$

$$\boldsymbol{B}_{eq}\boldsymbol{\beta} - \boldsymbol{b}_{eq} + \boldsymbol{s}_{eq} = \mathbf{0}_{n_{eq}\times 1} \tag{4.125b}$$

$$\boldsymbol{B}_{ineq}\boldsymbol{\beta} - \boldsymbol{b}_{ineq} + \boldsymbol{s}_{ineq} = \mathbf{0}_{n_{ineq}\times 1}, \tag{4.125c}$$

with both equality and inequality constraints that will be formulated as LCP and subsequently, a solution method will be described as well. Following Koch (1985), the first two equations, i.e (4.125a) and (4.125b), are written in matrix form

$$\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{B}_{eq}^T \\ \boldsymbol{B}_{eq} & \mathbf{0}_{n_{eq}} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \boldsymbol{\lambda}_{eq} \end{bmatrix} = \begin{bmatrix} \boldsymbol{q} - \boldsymbol{B}_{ineq}^T\boldsymbol{\lambda}_{ineq} \\ \boldsymbol{b}_{eq} - \boldsymbol{s}_{eq} \end{bmatrix} \tag{4.126}$$

from which the resulting estimate of $\boldsymbol{\beta}$ will be substituted in (4.125c) or equivalently, elimination of $\boldsymbol{\beta}$ from the system of equations (4.125a) - (4.125c). Dimension of the matrices and vector used in this section are briefly summarized for ease of readability.

- $\boldsymbol{\beta}$ is the $u \times 1$ vector of unknown model parameters

- $\boldsymbol{Q}$ is the $u \times u$ matrix defined in Eq. (4.85)
- $\boldsymbol{q}$ is the $u \times 1$ vector defined in Eq. (4.85)
- $\boldsymbol{B}_{eq}$ is the $n_{eq} \times u$ equality constraint coefficient matrix introduced in Eq. (4.75)
- $\boldsymbol{\lambda}_{eq}$ is the $n_{eq} \times 1$ vector of equality constraint Lagrange multipliers defined in Eq. (4.93b)
- $\boldsymbol{b}_{eq}$ is the $n_{eq} \times 1$ vector of equality constraint bound vector defined in Eq. (4.76b)
- $\boldsymbol{B}_{ineq}$ is the $n_{ineq} \times u$ inequality constraint coefficient matrix introduced in Eq. (4.74)
- $\boldsymbol{\lambda}_{ineq}$ is the $n_{ineq} \times 1$ vector of inequality constraint Lagrange multiplier defined in Eq. (4.93a)
- $\boldsymbol{b}_{ineq}$ is the $n_{ineq} \times 1$ vector of inequality constraint bound vector defined in Eq. (4.76a)
- $\boldsymbol{s}_{ineq}$ is the $n_{ineq} \times 1$ inequality constraint slack variable vector defined in Eq. (4.77)
- $\boldsymbol{s}_{eq}$ is the $n_{eq} \times 1$ equality constraint slack variable vector defined in Eq. (4.80d).

An estimate

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\lambda}}_{eq} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{B}_{eq}^T \\ \boldsymbol{B}_{eq} & \boldsymbol{0}_{n_{eq}} \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{q} - \boldsymbol{B}_{ineq}^T \boldsymbol{\lambda}_{ineq} \\ \boldsymbol{b}_{eq} - \boldsymbol{s}_{eq} \end{bmatrix}, \tag{4.127}$$

of $\begin{bmatrix} \boldsymbol{\beta}^T & \boldsymbol{\lambda}_{eq}^T \end{bmatrix}^T$ is obtained from Eq. (4.126), where $\boldsymbol{s}_{eq}$ will be omitted in further steps from the right hand side vector since the slack variable for equality constraints are defined as zero ; see Eq. (4.80d). In Eq. (4.127), it is assumed that an inverse of the matrix $\begin{bmatrix} \boldsymbol{Q} & \boldsymbol{B}_{eq}^T \\ \boldsymbol{B}_{eq} & \boldsymbol{0}_{n_{eq}} \end{bmatrix}$ exists and if not, then a generalized inverse is applied (Koch (1985), Koch (1988)). Using the identity

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1}\boldsymbol{B}\left(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1} \\ -\left(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1}\boldsymbol{C}\boldsymbol{A}^{-1} & \left(\boldsymbol{D} - \boldsymbol{C}\boldsymbol{A}^{-1}\boldsymbol{B}\right)^{-1} \end{bmatrix} \tag{4.128}$$

for the inversion of block matrices (see e.g. Koch (1988)), Eq. (4.127) becomes

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\lambda}}_{eq} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}^{-1} - \boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\left(\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\right)^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1} & \boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\left(\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\right)^{-1} \\ \left(\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\right)^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1} & -\left(\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\right)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{q} - \boldsymbol{B}_{ineq}^T\boldsymbol{\lambda}_{ineq} \\ \boldsymbol{b}_{eq} \end{bmatrix}. \tag{4.129}$$

In the following part of this section, few additional variables are introduced to simplify the matrix expressions and also aimed for a better readability. Substituting the first additional variable

$$\boldsymbol{W} = \boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T \tag{4.130}$$

Eq. (4.129) yields

$$\begin{bmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\lambda}}_{eq} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}^{-1} - \boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1} & \boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\boldsymbol{W}^{-1} \\ \boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1} & -\boldsymbol{W}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{q} - \boldsymbol{B}_{ineq}^T\boldsymbol{\lambda}_{ineq} \\ \boldsymbol{b}_{eq} \end{bmatrix} \tag{4.131}$$

from which, both

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{Q}^{-1} - \boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right)\left(\boldsymbol{q} - \boldsymbol{B}_{ineq}^T\boldsymbol{\lambda}_{ineq}\right) + \boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^T\boldsymbol{W}^{-1}\boldsymbol{b}_{eq} \tag{4.132a}$$

$$\text{and } \widehat{\boldsymbol{\lambda}}_{eq} = \left(\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right)\left(\boldsymbol{q} - \boldsymbol{B}_{ineq}^T\boldsymbol{\lambda}_{ineq}\right) - \boldsymbol{W}^{-1}\boldsymbol{b}_{eq} \tag{4.132b}$$

are obtained. To recollect, the goal was to solve the system of equations (4.125a), (4.125b) and (4.125c). So far $\widehat{\boldsymbol{\beta}}$, as an estimate of $\boldsymbol{\beta}$, has been obtained from (4.125a), (4.125b) which will now be used to eliminate $\boldsymbol{\beta}$ from (4.125c). Substituting $\widehat{\boldsymbol{\beta}}$ from Eq. (4.132a) into (4.125c) yields

$$\begin{aligned}&\boldsymbol{B}_{ineq}\left[\left(\boldsymbol{Q}^{-1}-\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right)\left(\boldsymbol{q}-\boldsymbol{B}_{ineq}^{T}\boldsymbol{\lambda}_{ineq}\right)+\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq}\right]\\&-\boldsymbol{b}_{ineq}+\boldsymbol{s}_{ineq}=\boldsymbol{0}_{n_{ineq}\times 1},\end{aligned}\tag{4.133}$$

which when expanded, gives

$$\begin{aligned}&\boldsymbol{B}_{ineq}\left[\left(\boldsymbol{Q}^{-1}-\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right)\boldsymbol{q}\right]-\\&\boldsymbol{B}_{ineq}\left[\left(\boldsymbol{Q}^{-1}-\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right)\boldsymbol{B}_{ineq}^{T}\boldsymbol{\lambda}_{ineq}\right]+\\&\boldsymbol{B}_{ineq}\left[\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq}\right]-\boldsymbol{b}_{ineq}+\boldsymbol{s}_{ineq}=\boldsymbol{0}_{n_{ineq}\times 1}.\end{aligned}\tag{4.134}$$

Bringing the slack variable $\boldsymbol{s}_{ineq}$ to the left hand side and rearranging (4.134) yields

$$\begin{aligned}\boldsymbol{s}_{ineq}=&\boldsymbol{B}_{ineq}\left[\left(\boldsymbol{Q}^{-1}-\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right)\boldsymbol{B}_{ineq}^{T}\boldsymbol{\lambda}_{ineq}\right]+\boldsymbol{b}_{ineq}\\&-\boldsymbol{B}_{ineq}\left[\left(\boldsymbol{Q}^{-1}-\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right)\boldsymbol{q}\right]-\boldsymbol{B}_{ineq}\left[\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq}\right].\end{aligned}\tag{4.135}$$

As mentioned already, additional variables $\boldsymbol{\Psi}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\omega}$ are introduced to simplify the expression. The substitutions are shown step-by-step for the reader to follow. First substitution

$$\boldsymbol{\Psi}=\boldsymbol{Q}^{-1}-\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\tag{4.136}$$

in Eq. (4.135) yields

$$\boldsymbol{s}_{ineq}=\left(\boldsymbol{B}_{ineq}\boldsymbol{\Psi}\boldsymbol{B}_{ineq}^{T}\right)\boldsymbol{\lambda}_{ineq}+\left(\boldsymbol{b}_{ineq}-\boldsymbol{B}_{ineq}\boldsymbol{\Psi}\boldsymbol{q}-\boldsymbol{B}_{ineq}\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq}\right).\tag{4.137}$$

Further substituting

$$\boldsymbol{\Omega}=\boldsymbol{B}_{ineq}\boldsymbol{\Psi}\boldsymbol{B}_{ineq}^{T}\tag{4.138a}$$

$$\text{and }\boldsymbol{\omega}=\boldsymbol{b}_{ineq}-\boldsymbol{B}_{ineq}\boldsymbol{\Psi}\boldsymbol{q}-\boldsymbol{B}_{ineq}\boldsymbol{Q}^{-1}\boldsymbol{B}_{eq}^{T}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq}\tag{4.138b}$$

in Eq. (4.137) yields

$$\boldsymbol{s}_{ineq}=\boldsymbol{\Omega}\boldsymbol{\lambda}_{ineq}+\boldsymbol{\omega}\qquad\text{subject to }\boldsymbol{\lambda}_{ineq}\geq\boldsymbol{0}_{n_{ineq}\times 1}\tag{4.139}$$

thus transforming the system of KKT equations into a linear complimentarity problem (LCP).

Solution of LCP

The solution of the linear complementary problem in based on the method proposed by Lemke (1968) for quadratic programming. Following the solution approach presented in Koch (1985), the algorithm described below either terminates in a solution of the complementary problem or no solution exists, with the only necessary condition that $\boldsymbol{\Omega}$ shall be positive semi-definite (Lemke (1968)). Assuming that a feasible solution exists, the LCP (4.139) is written as

$$\begin{bmatrix}\boldsymbol{I}_{n_{ineq}} & -\boldsymbol{\Omega}\end{bmatrix}\begin{bmatrix}\boldsymbol{s}_{ineq}\\ \boldsymbol{\lambda}_{ineq}\end{bmatrix}=\boldsymbol{\omega},\tag{4.140}$$

where $\boldsymbol{I}_{n_{ineq}}$ is an $n_{ineq} \times n_{ineq}$ identity matrix. As shown in Koch (1988), the $2n_{ineq} \times 1$ vector of slack variable and Lagrange multiplier $\begin{bmatrix} \boldsymbol{s}_{ineq}^T & \boldsymbol{\lambda}_{ineq}^T \end{bmatrix}^T$ is divided into the basic $\begin{bmatrix} \boldsymbol{s}_{ineq,b}^T & \boldsymbol{\lambda}_{ineq,b}^T \end{bmatrix}^T$ and the non-basic $\begin{bmatrix} \boldsymbol{s}_{ineq,n}^T & \boldsymbol{\lambda}_{ineq,n}^T \end{bmatrix}^T$ variables respectively. The partition

$$\begin{bmatrix} \boldsymbol{s}_{ineq} \\ \boldsymbol{\lambda}_{ineq} \end{bmatrix} = \begin{bmatrix} \boldsymbol{s}_{ineq,b} \\ \boldsymbol{s}_{ineq,n} \\ \boldsymbol{\lambda}_{ineq,b} \\ \boldsymbol{\lambda}_{ineq,n} \end{bmatrix}$$

is done such that both the slack variable and the Lagrange multiplier

$$\begin{aligned} \boldsymbol{s}_{ineq} &= \begin{bmatrix} \boldsymbol{s}_{ineq,b}^T & \boldsymbol{s}_{ineq,n}^T \end{bmatrix}^T \\ \boldsymbol{\lambda}_{ineq} &= \begin{bmatrix} \boldsymbol{\lambda}_{ineq,b}^T & \boldsymbol{\lambda}_{ineq,n}^T \end{bmatrix}^T \end{aligned} \tag{4.141}$$

remain $n_{ineq} \times 1$ vectors, as well as the conditions

$$\begin{bmatrix} \boldsymbol{s}_{ineq,b} \\ \boldsymbol{\lambda}_{ineq,b} \end{bmatrix} \geq \boldsymbol{0}_{n_{ineq} \times 1} \tag{4.142a}$$

$$\begin{bmatrix} \boldsymbol{s}_{ineq,n} \\ \boldsymbol{\lambda}_{ineq,n} \end{bmatrix} \geq \boldsymbol{0}_{n_{ineq} \times 1} \tag{4.142b}$$

are satisfied. It shall be noted that the conditions (4.142a) and (4.142b) are due to the fact that the transformed slack variable (4.81) shall be positive and the inequality constraint Lagrange multiplier shall be non-negative for a minimization problem. Condition (4.142b) is itself satisfied by the definition of the non-basic variables, which are ignored or assumed zero (Lemke (1968)). Furthermore, $\boldsymbol{\Omega}$ is partitioned as

$$\begin{aligned} \boldsymbol{\Omega} &= \begin{bmatrix} \boldsymbol{\Omega}_b & \boldsymbol{\Omega}_n \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{B}_{ineq,b} \boldsymbol{\Psi} \boldsymbol{B}_{ineq,b}^T & \boldsymbol{B}_{ineq,b} \boldsymbol{\Psi} \boldsymbol{B}_{ineq,n}^T \\ \boldsymbol{B}_{ineq,n} \boldsymbol{\Psi} \boldsymbol{B}_{ineq,b}^T & \boldsymbol{B}_{ineq,n} \boldsymbol{\Psi} \boldsymbol{B}_{ineq,n}^T \end{bmatrix}, \end{aligned} \tag{4.143}$$

where

$$\boldsymbol{B}_{ineq} = \begin{bmatrix} \boldsymbol{B}_{ineq,b} \\ \boldsymbol{B}_{ineq,n} \end{bmatrix} \quad ; \quad \boldsymbol{B}_{eq} = \begin{bmatrix} \boldsymbol{B}_{eq,b} \\ \boldsymbol{B}_{eq,n} \end{bmatrix} \quad ; \quad \boldsymbol{b}_{eq} = \begin{bmatrix} \boldsymbol{b}_{eq,b} \\ \boldsymbol{b}_{eq,n} \end{bmatrix} \quad ; \quad \boldsymbol{b}_{ineq} = \begin{bmatrix} \boldsymbol{b}_{ineq,b} \\ \boldsymbol{b}_{ineq,n} \end{bmatrix}. \tag{4.144}$$

The LCP (4.140) is written as

$$\begin{bmatrix} \boldsymbol{I}_{n_{basic}}^{rect} & -\boldsymbol{\Omega}_b & \boldsymbol{I}_{n_{non-basic}}^{rect} & -\boldsymbol{\Omega}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{s}_{ineq,b} \\ \boldsymbol{\lambda}_{ineq,b} \\ \boldsymbol{s}_{ineq,n} \\ \boldsymbol{\lambda}_{ineq,n} \end{bmatrix} = \boldsymbol{\omega}, \tag{4.145}$$

where $\boldsymbol{I}_{n_{basic}}^{rect}$ and $\boldsymbol{I}_{n_{non-basic}}^{rect}$ are rectangular matrices[40]. As already mentioned, the non-basic variables are ignored by the pivot algorithm (see Lemke (1968), Koch (1985)) and the LCP (4.145) simplifies to

$$\begin{bmatrix} \boldsymbol{I}_{n_{basic}}^{rect} & -\boldsymbol{\Omega}_b \end{bmatrix} \begin{bmatrix} \boldsymbol{s}_{ineq,b} \\ \boldsymbol{\lambda}_{ineq,b} \end{bmatrix} = \boldsymbol{\omega} \tag{4.146}$$

[40]The name "rectangular matrices" is chosen for lack for a better name. Specifically, these are identity matrices padded with additional rows of zeroes to match the exact dimensions of the basic and non-basic variables. For example $\boldsymbol{I}_{n_{basic}}^{rect} = \begin{bmatrix} \boldsymbol{I}_{n_{basic}} \\ \boldsymbol{0}_{n_{basic}} \end{bmatrix}$ and $\boldsymbol{I}_{n_{non-basic}}^{rect} = \begin{bmatrix} \boldsymbol{I}_{n_{non-basic}} \\ \boldsymbol{0}_{n_{non-basic}} \end{bmatrix}$

and an estimate

$$\begin{bmatrix} \widehat{\boldsymbol{s}}_{ineq,b} \\ \widehat{\boldsymbol{\lambda}}_{ineq,b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}^{rect}_{n_{basic}} & -\boldsymbol{\Omega}_b \end{bmatrix}^{-1} \boldsymbol{\omega} \tag{4.147}$$

of $\begin{bmatrix} \boldsymbol{s}^T_{ineq,b} & \boldsymbol{\lambda}^T_{ineq,b} \end{bmatrix}^T$ is obtained. Using identity (4.128) to compute the inverse, Eq. (4.147) becomes

$$\begin{bmatrix} \widehat{\boldsymbol{s}}_{ineq,b} \\ \widehat{\boldsymbol{\lambda}}_{ineq,b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{n_{basic}} & -\boldsymbol{B}_{ineq,b}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}\left(\boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}\right)^{-1} \\ \boldsymbol{0}_{n_{non-basic}\times n_{basic}} & -\left(\boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}\right)^{-1} \end{bmatrix} \boldsymbol{\omega}. \tag{4.148}$$

Based on the substitutions (4.138a) and (4.138b), $\boldsymbol{\omega}$ is partitioned as

$$\boldsymbol{\omega} = \begin{bmatrix} \boldsymbol{b}_{ineq,b} - \boldsymbol{B}_{ineq,b}\boldsymbol{\Psi}\boldsymbol{q} - \boldsymbol{B}_{ineq,b}\boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq,b}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq,b} \\ \boldsymbol{b}_{ineq,n} - \boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{q} - \boldsymbol{B}_{ineq,n}\boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq,n}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq,n} \end{bmatrix} \tag{4.149}$$

and substituting in Eq. (4.148) gives

$$\begin{bmatrix} \widehat{\boldsymbol{s}}_{ineq,b} \\ \widehat{\boldsymbol{\lambda}}_{ineq,b} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{n_{basic}} & -\boldsymbol{B}_{ineq,b}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}\left(\boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}\right)^{-1} \\ \boldsymbol{0}_{n_{non-basic}\times n_{basic}} & -\left(\boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}\right)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_{ineq,b} - \boldsymbol{B}_{ineq,b}\boldsymbol{\Psi}\boldsymbol{q} - \boldsymbol{B}_{ineq,b}\boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq,b}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq,b} \\ \boldsymbol{b}_{ineq,n} - \boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{q} - \boldsymbol{B}_{ineq,n}\boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq,n}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq,n} \end{bmatrix}, \tag{4.150}$$

from which an estimate of the inequality constraint Lagrange multiplier

$$\widehat{\boldsymbol{\lambda}}_{ineq,b} = -\left(\boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}\right)^{-1} \left[\boldsymbol{b}_{ineq,n} - \boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{q} - \boldsymbol{B}_{ineq,n}\boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq,n}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq,n}\right] \tag{4.151}$$

is obtained. Furthermore, substituting $\widehat{\boldsymbol{\lambda}}_{ineq,b}$ from Eq. (4.151) into (4.132a), an estimate of the model parameter

$$\widehat{\boldsymbol{\beta}} = \left[\left(\boldsymbol{Q}^{-1} - \boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq}\boldsymbol{W}^{-1}\boldsymbol{B}_{eq}\boldsymbol{Q}^{-1}\right) \left(\boldsymbol{q} + \left(\boldsymbol{B}^T_{ineq}\boldsymbol{X}^{-1}\left(\boldsymbol{b}_{ineq,n} - \boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{q} - \boldsymbol{B}_{ineq,n}\boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq,n}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq,n}\right)\right)\right)\right] + \boldsymbol{Q}^{-1}\boldsymbol{B}^T_{eq}\boldsymbol{W}^{-1}\boldsymbol{b}_{eq} \tag{4.152}$$

is obtained with inequality and equality constraints, where

$$\boldsymbol{X} = \boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{B}^T_{ineq,b}.$$

Under the special case

$$\begin{bmatrix} \boldsymbol{b}_{eq,b} \\ \boldsymbol{b}_{eq,n} \end{bmatrix} = \boldsymbol{0}_{n_{eq}\times 1} \quad ; \quad \begin{bmatrix} \boldsymbol{b}_{ineq,b} \\ \boldsymbol{b}_{ineq,n} \end{bmatrix} = \boldsymbol{0}_{n_{ineq}\times 1} \quad ; \quad \begin{bmatrix} \boldsymbol{B}_{eq,b} \\ \boldsymbol{B}_{eq,n} \end{bmatrix} = \boldsymbol{0}_{n_{eq}} \quad ; \quad \begin{bmatrix} \boldsymbol{B}_{ineq,b} \\ \boldsymbol{B}_{ineq,n} \end{bmatrix} = \boldsymbol{0}_{n_{ineq}},$$

corresponding to no equality or inequality constraints, Eq. (4.152) reduces to unconstrained least squares estimate

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{Q}^{-1}\boldsymbol{q}. \tag{4.153}$$

As a summary of this approach for solving the LCP, two important points are noteworthy. We began with the first and second KKT equations (stationarity and feasibility condition). The model parameter vector $\boldsymbol{\beta}$ was eliminated from the system of equations and an estimate of Lagrange multiplier in Eq. (4.151) was obtained. Then, the Lagrange multiplier and slack variable are

partitioned into two types, namely, the basic and non-basic variables. The resulting Lagrange multiplier vector from the solution of LCP is substituted to obtain an estimate of the model parameter vector. It has also been shown that in absence of any equality or inequality constraints, the solutions reduces to that of least squares. From this perspective, there is an elegance to this approach and Koch (1985) is acknowledged for providing this LCP formulation as well as its solution. Compared to the method of centering and scaling (Nocedal and Wright (2006)), other obvious advantages of this approach are that no additional tuning parameters are required. The only two considerations are that (1) the matrix $\boldsymbol{B}_{ineq,n}\boldsymbol{\Psi}\boldsymbol{B}_{ineq,b}$ must be invertible and (2) prior information about the basic and non-basic variables must be given[41]. However, when it is not possible to identify in advance, which constraints are active or inactive, then the method of centering and scaling could be used instead.

4.7.2 Primal and dual variables

In this section, the concept of Lagrangian primal and dual variables is introduced. Duality is a principle according to which, an optimization problem is viewed from either of the two perspectives, namely, the "primal" or the "dual" (Nocedal and Wright (2006) pg. 304, Boyd and Vandenberghe (2004) pg. 255, Roese-Koerner (2015)). The transformation from Lagrangian primal to an equivalent dual is summarized below. For a detailed background on this topic, the reader is referred to Nocedal and Wright (2006) pg. 482. Once again, the starting point for the following discussion is the Lagrangian function

$$L_{primal}(\boldsymbol{\beta},\boldsymbol{\lambda},\boldsymbol{s}) = \frac{1}{\sigma^2}\left[\frac{1}{2}\boldsymbol{\beta}^T\boldsymbol{Q}\boldsymbol{\beta} - \left(\boldsymbol{q}^T - \boldsymbol{\lambda}^T\boldsymbol{B}\right)\boldsymbol{\beta} - \boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{s}) + \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y}\right]$$

in its primal form, reproduced from Eq. (4.84). The derivative of Lagrangian function with respect to the model parameter vector was already obtained to derive the first system of KKT equations

$$\boldsymbol{Q}\ \boldsymbol{\beta} - \boldsymbol{q} + \boldsymbol{B}^T\boldsymbol{\lambda} = \boldsymbol{0}_{u\times 1},$$

which is the same as Eq. (4.94) reproduced here and rearranged to express

$$\boldsymbol{\beta} = \boldsymbol{Q}^{-1}\left(\boldsymbol{q} - \boldsymbol{B}^T\boldsymbol{\lambda}\right) \tag{4.154}$$

as a function of the unknown Lagrange multiplier. The first step towards transformation into dual form is to eliminate the model parameter $\boldsymbol{\beta}$ from the primal form. Therefore, $\boldsymbol{\beta}$ from Eq. (4.154) is

[41] In evaluation of this method, the author wishes to clarify that equal number of basic and non-basic variables have been assumed.

substituted in primal Lagrangian form to give the dual form

$$L_{dual}(\boldsymbol{\lambda}, \boldsymbol{s}) = \frac{1}{\sigma^2}\left[\frac{1}{2}\left(\boldsymbol{Q}^{-1}\left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right)\right)^T\boldsymbol{Q}\left(\boldsymbol{Q}^{-1}\left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right)\right) - \left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right)^T\boldsymbol{Q}^{-1}\left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right) - \boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{s}) + \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y}\right] \tag{4.155a}$$

$$= \frac{1}{\sigma^2}\left[\frac{1}{2}\left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right)^T\boldsymbol{Q}^{-1}\boldsymbol{Q}\boldsymbol{Q}^{-1}\left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right) - \left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right)\boldsymbol{Q}^{-1}\left(\boldsymbol{q}-\boldsymbol{B}^T\boldsymbol{\lambda}\right) - \boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{s}) + \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y}\right] \tag{4.155b}$$

$$= \frac{1}{\sigma^2}\left[\frac{1}{2}\left(\boldsymbol{q}^T\boldsymbol{Q}^{-1}\boldsymbol{q} - \boldsymbol{\lambda}^T\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{q} - \boldsymbol{q}^T\boldsymbol{Q}^{-1}\boldsymbol{B}^T\boldsymbol{\lambda} + \boldsymbol{\lambda}^T\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{B}^T\boldsymbol{\lambda}\right) - \left(\boldsymbol{q}^T\boldsymbol{Q}^{-1}\boldsymbol{q} - \boldsymbol{\lambda}^T\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{q} - \boldsymbol{q}^T\boldsymbol{Q}^{-1}\boldsymbol{B}^T\boldsymbol{\lambda} + \boldsymbol{\lambda}^T\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{B}^T\boldsymbol{\lambda}\right) - \boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{s}) + \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y}\right] \tag{4.155c}$$

$$= \frac{1}{\sigma^2}\left[\frac{1}{2}\boldsymbol{q}^T\boldsymbol{Q}^{-1}\boldsymbol{q} - \boldsymbol{\lambda}^T\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{q} - \frac{1}{2}\boldsymbol{\lambda}^T\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{B}^T\boldsymbol{\lambda} - \boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{s}) + \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y}\right] \tag{4.155d}$$

$$= \frac{1}{\sigma^2}\left[\frac{-1}{2}\boldsymbol{\lambda}^T\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{B}^T\boldsymbol{\lambda} - \boldsymbol{\lambda}^T\left(\boldsymbol{B}\boldsymbol{Q}^{-1}\boldsymbol{q} - \boldsymbol{b} + \boldsymbol{s}\right) + \frac{1}{2}\boldsymbol{q}^T\boldsymbol{Q}^{-1}\boldsymbol{q} + \boldsymbol{y}^T\boldsymbol{P}\boldsymbol{y}\right] \tag{4.155e}$$

of the Lagrangian function. The main difference between the primal and dual representation is the number of unknowns (Boyd and Vandenberghe (2004) pg. 609, Nocedal and Wright (2006) pg. 482). The primal form is a function in $\boldsymbol{\beta}, \boldsymbol{\lambda}$ and $\boldsymbol{s}$ with $u + 2n_{ineq} + 2n_{eq}$ unknown parameters. The dual form is a function in $\boldsymbol{\lambda}$ and $\boldsymbol{s}$ with $2n_{ineq} + 2n_{eq}$ unknown parameters. Therefore, optimizing a dual Lagrangian function provides a lower bound with respect to the solution of the primal minimization problem (Nocedal and Wright (2006) pg. 343, Boyd and Vandenberghe (2004) pg. 223-224). Furthermore, if the primal form of Lagrangian function is positive, then the corresponding dual is negative (Roese-Koerner (2015)). Therefore, in order to estimate the model parameters from the dual form, the corresponding Lagrangian function must be maximized with respect to the Lagrange multiplier and slack variable. The difference in optimal values between the primal and dual problems is called the duality gap (Boyd and Vandenberghe (2004) pg. 224, Nocedal and Wright (2006) pg. 343, Roese-Koerner (2015)).

This concludes the discussion on the duality theory and the dual Lagrangian function is maximized subject to the KKT conditions to estimate the model parameter, in a similar manner as shown in section 4.5.1.

4.7.3 Alternate techniques for step size estimation

The importance of the step size in an optimization algorithm has been highlighted in the previous sections. With the evolving role of descent methods in optimization[42], there is a growing attention to it. The step size does not affect the performance in smaller dimensional problems (e.g. smaller than 10 unknown parameters) as much as it does in large dimensional problems (e.g. 1000 unknown parameters or more).

As already described in the previous section, the descend direction

$$\Delta\tilde{\boldsymbol{\beta}} = \begin{bmatrix}\Delta\boldsymbol{\beta}^T & \Delta\boldsymbol{\lambda}^T & \Delta\tilde{\boldsymbol{s}}^T\end{bmatrix}^T$$

[42] Evolving algorithms such as supervised learning and neural networks also use descent method in their so-called "training phase". In this context, the step size is also known as the "learning rate" (see e.g. Jacobs (1988), Wilson and Martinez (2001)).

is estimated from the system of equations (4.102), i.e.

$$\begin{aligned}
&\boldsymbol{Q}\ \boldsymbol{\beta} - \boldsymbol{q} + \boldsymbol{B}^T\boldsymbol{\lambda} = \mathbf{0}_{u\times 1} \leftarrow \text{ Stationarity condition}\\
&\boldsymbol{B}\ \boldsymbol{\beta} - \boldsymbol{b} + \boldsymbol{s} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1} \leftarrow \text{ Feasibility condition}\\
&\boldsymbol{S}\boldsymbol{\Lambda}\boldsymbol{e} - \sigma_{cen}\mu_{com}\boldsymbol{e} = \mathbf{0}_{(n_{ineq}+n_{eq})\times 1} \leftarrow \text{ Complimentary slack condition}
\end{aligned}$$

by linearizing at the initial values $\tilde{\boldsymbol{\beta}}_0 = \begin{bmatrix}\boldsymbol{\beta}_0^T & \boldsymbol{\lambda}_0^T & \tilde{\boldsymbol{s}}_0^T\end{bmatrix}^T$ thus yielding Eq. (4.106), reproduced here for the sake of better readability,

$$\begin{bmatrix}\boldsymbol{Q} & \boldsymbol{B}^T & \mathbf{0}_u\\ \boldsymbol{B} & \mathbf{0}_{n_{ineq}+n_{eq}} & \boldsymbol{I}\\ \mathbf{0}_u & \boldsymbol{S}_0 & \boldsymbol{\Lambda}_0\end{bmatrix}\begin{bmatrix}\Delta\boldsymbol{\beta}\\ \Delta\boldsymbol{\lambda}\\ \Delta\tilde{\boldsymbol{s}}\end{bmatrix} = \begin{bmatrix}-\boldsymbol{Q}\boldsymbol{\beta}_0 + \boldsymbol{q} - \boldsymbol{B}^T\boldsymbol{\lambda}_0\\ -\boldsymbol{B}\boldsymbol{\beta}_0 + \boldsymbol{b} - \boldsymbol{s}_0\\ -\boldsymbol{S}_0\boldsymbol{\Lambda}_0\boldsymbol{e} + \mu_{com}\sigma_{cen}\boldsymbol{e}\end{bmatrix}, \tag{4.156}$$

where $\boldsymbol{Q}$ and $\boldsymbol{B}$ are known matrices as well as, $\boldsymbol{q}$ and $\boldsymbol{b}$ are known vectors, $\boldsymbol{S}_0$ and $\boldsymbol{\Lambda}_0$ are evaluated at the current iterates $\tilde{\boldsymbol{s}}_0$ and $\boldsymbol{\lambda}_0$. Optimization parameters are updated from iteration it to $it+1$ using the update rule

$$\tilde{\boldsymbol{\beta}}_{it+1} = \tilde{\boldsymbol{\beta}}_{it} + \boldsymbol{T}_{it}\Delta\tilde{\boldsymbol{\beta}}_{it}, \tag{4.157}$$

where

$$\boldsymbol{T}_{it} = \begin{bmatrix}\boldsymbol{T}_\beta & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \boldsymbol{T}_{\lambda_{ineq}} & \mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \boldsymbol{T}_{\lambda_{eq}} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{T}_{\tilde{s}_{ineq}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{T}_{\tilde{s}_{eq}}\end{bmatrix} \tag{4.158}$$

such that the diagonal elements of $\boldsymbol{T}_{it}$ lie in interval $(0, 1]$.

Possibilities for the step size estimation

This section provides three options based on which the $\boldsymbol{T}_{it}$, as shown in Eq. (4.158), can be obtained to solve the KKT equations (4.156).

- Random choice of step size: For smaller dimensional problems (with 10 or fewer unknown parameters), a random step size could be chosen by trial and error until an acceptable value is determined. However, for large dimensional problems (more than 1000s of unknown parameters), a systematic procedure is required.
- Step size based on generalized cross validation (GCV) strategy: In this method, a pre-defined fixed number of step size values are chosen and with each of those, Eq. (4.102) is solved. The resulting accuracy (or error) of the estimated model parameters is then analysed against the chosen step sizes. Ideally, such an analysis should reveal a trend (or structure e.g a parabolic or exponential decay) with which an appropriate step size corresponding to minimum error is chosen. This is a more robust approach compared to choosing step size randomly but there are two challenges.
  - In order to obtain an unbiased estimation error for model parameters, independent observations (one that is not used for estimating the model parameter) must be used. In other words, the entire set of observation needs to be divided into two parts, namely, "estimation" and "cross-validation". In reality, the two cannot be completely independent because of correlations. It can be partially overcome by allowing randomized re-sampling of the observations to separately generate the estimation and validation observation sets. Nevertheless, some correlation would still remain. A similar approach was used in Liang (2017).
  - If the error in the estimated model parameters, obtained from a cross-validation set, does not show a deterministic trend (e.g. error is lower for smaller values along the diagnol elements of the step size matrix), then the step size cannot be chosen reliably.

- Choose step size based on predicted optimization parameters[43]: This approach is found to be the most promising for this work because it provides a systematic, reliable estimate of the step size with the least computational expense compared to the other two approaches. This approach has also the advantages that most of the computations have already been performed in the solution of the KKT equations and most importantly an estimate of the quality of the determined step size becomes available at little or no additional computational expense.

The third option from above is found to be the most suitable in context of global 4D electron density modelling, where the proposed optimization technique is to be applied (as described in the next chapter). This approach is substantiated below and begins with the update rule (4.107) (reproduced below for readability)

$$\tilde{\boldsymbol{\beta}}_{it+1} = \tilde{\boldsymbol{\beta}}_{it} + \boldsymbol{T}_{it}\Delta\tilde{\boldsymbol{\beta}}_{it},$$

which will be modified to

$$\tilde{\boldsymbol{\beta}}_{it+1} \approx \tilde{\boldsymbol{\beta}}_{it} + \boldsymbol{T}^{a}_{it}\Delta\tilde{\boldsymbol{\beta}}_{it-1} \tag{4.159}$$

to estimate $\boldsymbol{T}^{a}_{it}$, an approximate[44] step size parameter. First, the modification in Eq. (4.159) will be highlighted and then its justification will be given. Specifically, there are two changes:

- $\boldsymbol{T}_{it}$ is replaced with $\boldsymbol{T}^{a}_{it}$, an approximate step size to be determined.
- the descent direction $\Delta\tilde{\boldsymbol{\beta}}_{it-1}$ refers to the iteration $it-1$ in Eq. (4.159) instead of it as in Eq. (4.107). The justification for this modification is as follows: When an approximate step size has been estimated, it will be substituted back in Eq. (4.107), which in the first place, had been the basis for the step size estimation. This leads us to the parameter update being biased with that of the predicted value. Therefore, a change in Eq. (4.159) compared to Eq. (4.107) is necessary.

A predicted value of the optimization parameter $\tilde{\boldsymbol{\beta}}^{Pred}_{it+1}$ is substituted for $\tilde{\boldsymbol{\beta}}_{it+1}$ in Eq. (4.107) to give

$$\tilde{\boldsymbol{\beta}}^{Pred}_{it+1} - \tilde{\boldsymbol{\beta}}_{it} \approx \boldsymbol{T}^{a}_{it}\Delta\tilde{\boldsymbol{\beta}}_{it-1}, \tag{4.160}$$

so that the difference between the predicted optimization parameters and that at the current iteration becomes a linear function of the descent direction. Substituting $\Delta\tilde{\boldsymbol{\beta}}^{Pred}_{it+1} = \tilde{\boldsymbol{\beta}}^{Pred}_{it+1} - \tilde{\boldsymbol{\beta}}_{it}$ in Eq. (4.160) gives

$$\Delta\tilde{\boldsymbol{\beta}}^{Pred}_{it+1} \approx \boldsymbol{T}^{a}_{it}\Delta\tilde{\boldsymbol{\beta}}_{it-1} \tag{4.161}$$

from which, an estimate

$$\widehat{\boldsymbol{T}}^{a}_{it} = \boldsymbol{W}\boldsymbol{U}^{-1}, \tag{4.162}$$

of $\boldsymbol{T}^{a}_{it}$ is obtained[45], where the diagonal matrices

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{W}_{\beta} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{W}_{\lambda_{ineq}} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{W}_{\lambda_{eq}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{W}_{\tilde{s}_{ineq}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{W}_{\tilde{s}_{eq}} \end{bmatrix} \tag{4.163}$$

and

$$\boldsymbol{U} = \begin{bmatrix} \boldsymbol{U}_{\beta} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_{\lambda_{ineq}} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U}_{\lambda_{eq}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U}_{\tilde{s}_{ineq}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U}_{\tilde{s}_{eq}} \end{bmatrix} \tag{4.164}$$

[43] To the best knowledge of the author, this method has not been used until now by anybody
[44] Superscript "a" is to indicate that the step size is approximate
[45] The approximation symbol in (4.161) will assumed to be = for estimation of the step size.

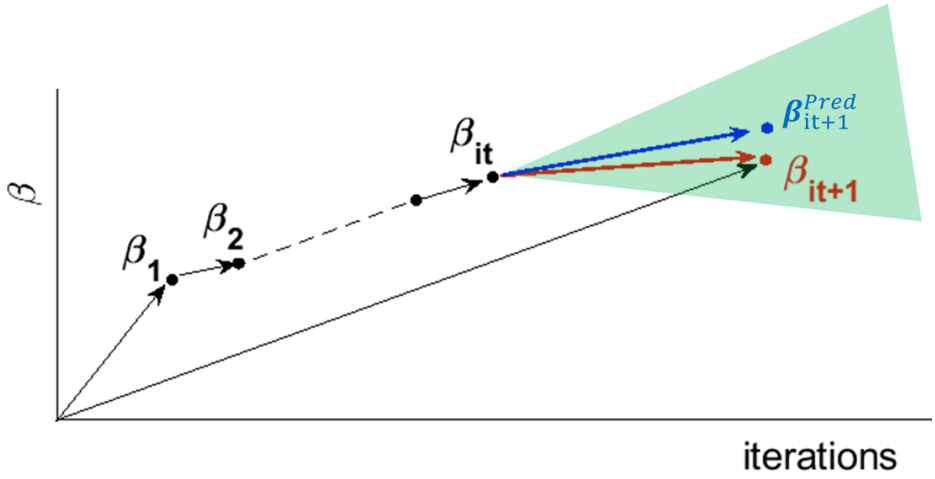


Figure 4.26: Geometrical interpretation of predicting the model parameter vector in step size estimation.

contain terms corresponding to the model parameter, Lagrange multiplier and the slack variable ; see Eqs. (4.165) to (4.170), where the $\boldsymbol{U}$ and $\boldsymbol{W}$ submatrices are as shown below:

$$\boldsymbol{U}_{\beta} = \begin{bmatrix} \Delta\beta_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\beta_u \end{bmatrix} \tag{4.165}$$

$$\boldsymbol{U}_{\lambda_{ineq}} = \begin{bmatrix} \Delta\lambda_{ineq,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\lambda_{ineq,n_{ineq}} \end{bmatrix} \; ; \; \boldsymbol{U}_{\lambda_{eq}} = \begin{bmatrix} \Delta\lambda_{eq,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\lambda_{eq,n_{eq}} \end{bmatrix} \tag{4.166}$$

$$\boldsymbol{U}_{\tilde{s}_{ineq}} = \begin{bmatrix} \Delta\tilde{s}_{ineq,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\tilde{s}_{ineq,n_{ineq}} \end{bmatrix} \; ; \; \boldsymbol{U}_{\tilde{s}_{eq}} = \begin{bmatrix} \Delta\tilde{s}_{eq,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\tilde{s}_{eq,n_{eq}} \end{bmatrix} \tag{4.167}$$

$$\boldsymbol{W}_{\beta} = \begin{bmatrix} \Delta\beta_1^{Pred} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\beta_u^{Pred} \end{bmatrix} \tag{4.168}$$

$$\boldsymbol{W}_{\lambda_{ineq}} = \begin{bmatrix} \Delta\lambda_{ineq,1}^{Pred} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\lambda_{ineq,n_{ineq}}^{Pred} \end{bmatrix} \; ; \; \boldsymbol{W}_{\lambda_{eq}} = \begin{bmatrix} \Delta\lambda_{eq,1}^{Pred} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\lambda_{eq,n_{eq}}^{Pred} \end{bmatrix} \tag{4.169}$$

$$\boldsymbol{W}_{\tilde{s}_{ineq}} = \begin{bmatrix} \Delta\tilde{s}_{ineq,1}^{Pred} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\tilde{s}_{ineq,n_{ineq}}^{Pred} \end{bmatrix} \; ; \; \boldsymbol{W}_{\tilde{s}_{eq}} = \begin{bmatrix} \Delta\tilde{s}_{eq,1}^{Pred} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta\tilde{s}_{eq,n_{eq}}^{Pred} \end{bmatrix}. \tag{4.170}$$

A geometrical interpretation of predicting the model parameter in context of step size estimation is illustrated in Fig. 4.26. Given the intermediate solutions $\{\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \cdots, \boldsymbol{\beta}_{\boldsymbol{it}}\}$ corresponding to iterations $\{1, 2, \cdots, it\}$, the goal is to obtain $\boldsymbol{\beta}_{it+1}^{Pred}$ (blue solid circle ●) within an uncertainty cone illustrated in green. The magnitude of the relative vector $\boldsymbol{\beta}_{it+1}^{Pred} - \boldsymbol{\beta}_{\boldsymbol{it}}$ (indicated in blue ⟶) is the step size used to compute the updated model parameter vector $\boldsymbol{\beta}_{it+1}$ (red solid circle ●). The magnitude of the relative vector $\boldsymbol{\beta}_{it+1} - \boldsymbol{\beta}_{\boldsymbol{it}}$ is indicated in red ⟶. Although the descent directions $\boldsymbol{\beta}_{it+1}$ and $\boldsymbol{\beta}_{it+1}^{Pred}$ are different, their magnitude is the same.

4.7.4 Prediction of optimization parameters

In the previous section, a technique was introduced for estimating the step-size based on the predicted value of the optimization parameters

$$\tilde{\boldsymbol{\beta}}^{Pred} = \begin{bmatrix} \boldsymbol{\beta}^{Pred} \\ \boldsymbol{\lambda}^{Pred} \\ \boldsymbol{s}^{Pred} \end{bmatrix}, \tag{4.171}$$

which include the predicted model parameter, the Lagrange multiplier as well as slack variable for inequality and equality constraints.

Assuming that a predicted value of the model parameter $\boldsymbol{\beta}^{Pred}$ is available (as it will be discussed in this section), the Lagrange multiplier is obtained indirectly by rearranging Eq. (4.102) in the form

$$\boldsymbol{B}^T \boldsymbol{\lambda}^{Pred} = \boldsymbol{q} - \boldsymbol{Q}\ \boldsymbol{\beta}^{Pred}, \tag{4.172}$$

from which an estimate

$$\widehat{\boldsymbol{\lambda}}^{Pred} = \left(\boldsymbol{B}\ \Sigma_{\boldsymbol{\beta}^{Pred}}\ \boldsymbol{B}^T\right)^{-1} \boldsymbol{B}\ \Sigma_{\boldsymbol{\beta}^{Pred}}\ \left(\boldsymbol{q} - \boldsymbol{Q}\boldsymbol{\beta}^{Pred}\right) \tag{4.173}$$

of $\boldsymbol{\lambda}^{Pred}$ is obtained, where $\Sigma_{\boldsymbol{\beta}^{Pred}}$ denotes the covariance matrix of $\boldsymbol{\beta}^{Pred}$. When predicted model parameter vector is in the feasible region, the corresponding slack variable

$$\boldsymbol{s}^{Pred} = \boldsymbol{b} - \boldsymbol{B}\boldsymbol{\beta}^{Pred} \tag{4.174}$$

becomes a direct measure of a "margin" between the constraint bound and the predicted parameter vector. Specifically, if $s_k^{Pred} > 0$, then the corresponding k^{th} constraint was inactive. This allows a reduction of parameterization in Eq. (4.173) because Lagrange multipliers need to be obtained for only those constraints which are binding (i.e. those that have an impact on the solution) and those corresponding to the other non-binding constraints can be set to zero. Thus two of the three optimization parameters in Eq. (4.171) are predicted.

The next section describes a method for the prediction of the model parameter vector $\boldsymbol{\beta}^{Pred}$, which was assumed to be known in Eqs. (4.173) and (4.174). Specifically, we are interested in those methods which are suitable for shorter prediction windows[46], such as ARMA processes. An example prediction problem is shown below:

Predict the model parameter vector at the iteration[47] $it+1$, given their intermediate values at $\{0, \cdots, it\}$ along with the corresponding descent directions and objective function values. In other words,

$$\begin{aligned} &\text{Predict } \boldsymbol{\beta}_{it+1} \\ &\text{given : } \boldsymbol{\beta}_0, \cdots, \boldsymbol{\beta}_{it} \quad ; \quad J(\boldsymbol{\beta}_0), \cdots, J(\boldsymbol{\beta}_{it}) \quad ; \quad \Delta\boldsymbol{\beta}_0, \cdots, \Delta\boldsymbol{\beta}_{it}. \end{aligned} \tag{4.175}$$

As described further, $\boldsymbol{\beta}^{Pred}$ is obtained using a non-linear ARMAX (N-ARMAX) method, with which an estimate of the step size, as described in the Eqs. (4.160) to (4.162), is obtained. The N-ARMAX method, described in the next section, is based on time series analysis of the model parameter vector and additional exogenous variables. An introduction to time series based time series data for geophysical applications is referred to Ulrych (1985), Noakes (1986) and Stoica and Nehorai (1989).

[46] Prediction window refers to the interval between a prediction instant and that of the last available data. It can be defined and measured in units of time or samples. Typically, the smallest prediction window is the sampling rate of the data (see e.g. Davis (1977), Priestley (1978)).

[47] The given data does not have to be intermediate values from iteration only. It could be time series data as well.

4.7.5 Non-linear ARMAX method

N-ARMAX is an acronym for **N**on-linear **A**uto-**R**egressive **M**oving **A**verage with e**X**ogenous variables. It has its origins from the basic ARMA technique (see e.g. Box and Jenkins (1990), Brockwell and Davis (1991)).

ARMA is a time series modelling technique to predict a certain stationary signal at a future time moment using its past samples (Priestley (1978)). However, there are several weaknesses when an ARMA process is used for modelling a non-linear function of time and possibly other parameters (Rosenblatt (1971)). An example is the ionospheric electron density, which is not only a non-linear function of time but also other physical parameters such as electron temperature, ion temperature, the density of energetic particles as well as the incident energy on the atmosphere, electric and magnetic field intensities etc. (see Schunk (1988), Hernández-Pajares et al. (1998), Menvielle et al. (2011)).

In this section, an introduction to ARMA is given and followed by the reasons for introducing the non-linear component (N-ARMAX) as well as the exogenous variables. An auto-regressive (AR) process of order p,

$$x_{t,AR} = \sum_{i=1}^{p} \phi_i x_{t-i,AR} + \epsilon_{t,AR}, \tag{4.176}$$

relates a certain stationary signal at time instant t, denoted as $x_{t,AR}$, to its past values $x_{t-i,AR}$ through coefficients ϕ_i for $i = \{1, \cdots, p\}$ corresponding to p discrete time lags and the random error $\epsilon_{t,AR}$. A moving average (MA) process of order q,

$$x_{t,MA} = \sum_{j=1}^{q} \theta_j \epsilon_{t-j,MA} + \epsilon_{t,MA}, \tag{4.177}$$

relates $x_{t,MA}$ to the prediction error[48] through the MA coefficients θ_j for $j = \{1, \cdots, q\}$ corresponding to q discrete time lags and random error $\epsilon_{t,MA}$. The combination of AR(p) and MA(q) processes, (4.176) and (4.177), yields an ARMA(p, q) process

$$x_{t,ARMA} = \sum_{i=1}^{p} \phi_i x_{t-i,AR} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j,MA} + \epsilon_{t,ARMA}, \tag{4.178}$$

where $\epsilon_{t,ARMA}$ is the random error. However, when the signal to be predicted depends not only on its time history but also on other parameters, then an ARMAX(p, q, m) process,

$$x_{t,ARMAX} = \sum_{i=1}^{p} \phi_i x_{t-i,AR} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j,MA} + \sum_{k=1}^{m} \omega_k u_{t-k,X} + \epsilon_{t,AX}, \tag{4.179}$$

with $\boldsymbol{\omega}$ the exogenous coefficients, $u_{t-k,X}$ the exogenous parameters at time instants $t-k$ are used. The subscript X in $u_{t-k,X}$ signifies that it is an exogenous variable. Applying a non-linear mapping $g_n(\cdot)$ to the ARMAX(p, q, m) process (4.179) with coefficient $\boldsymbol{\gamma}$ leads to N-ARMAX(p, q, m, r) process (see e.g. Chatfield and Pepper (1971), Davis (1977))

$$x_{t,N-ARMAX} = \sum_{n=1}^{r} \gamma_n \, g_n \left(\sum_{i=1}^{p} \phi_i x_{t-i,AR} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j,MA} + \sum_{k=1}^{m} \omega_k u_{t-k,X} \right) + \epsilon_{t,N-ARMAX}. \tag{4.180}$$

However, the exact non-linearity ($g_n(\cdot)$) between the inputs $x_{t-i,AR}$, $\epsilon_{t-j,MA}$, $u_{t-k,X}$ and the output $x_{t,N-ARMAX}$ is generally unknown. The orders p, q, m as well as the number, type of mapping

[48] Prediction error is defined as the difference between a predicted value at time instant t and its true value. Since the true value would not be available at the previous time instant, the recurrence relation (4.176) and (4.177) are used in estimation of AR and MA coefficients; see Brockwell and Davis (1991) and Box and Jenkins (1990).

functions are the essential specifications of a N-ARMAX process (Priestley (1978), Scargle (1981)). It shall be noted that additional of exogenous variables can be added to the N-ARMAX model at the cost of increased number of unknown coefficients. While trigonometric, exponential, logarithmic, polynomial, Gaussian are commonly used non-linear mapping functions (see Maravall (1983), McLeod and Li (1983)), specialized ones[49] (Clevert et al. (2016)), such as

$$g_{\text{GeLu}}(z) = \frac{1}{2} z \left(1 + \text{erf}\left(\frac{z}{\sqrt{2}}\right)\right) \tag{4.181a}$$

$$g_{\tanh}(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}, \tag{4.181b}$$

are also found in literature (see Hendrycks and Gimpel (2016b), Hendrycks and Gimpel (2016a)), where

$$z = \sum_{i=1}^{p} \phi_i x_{t-i,AR} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j,MA} + \sum_{k=1}^{m} \omega_k u_{t-k,X}. \tag{4.182}$$

The exact choice of a non-linear mapping function depends on the parameter to be predicted, the order of the N-ARMAX process, the auto-correlation properties of the parameter and the time window for prediction (Box and Jenkins (1990), Schuh et al. (2015)). In the next chapter, specific information about mapping functions used for ionospheric electron density modelling will be presented. Following are the main considerations of N-ARMAX modelling approach:

- Compared to ARMAX, the N-ARMAX approach requires at least r additional coefficients corresponding to the non-linear mapping functions.
- N-ARMAX model output depends on the exogenous variables, number and type of non-linear mapping functions.
- If a large order of non-linear coefficients $\boldsymbol{\gamma}$ (approximately 20% of the sum of the orders $p + q + m$ or larger) is considered, then there is a risk of "over-fitting";Ulrych (1985) Box and Jenkins (1990).
- Approaches like random forest (Breiman (2001)) or recurrent memory techniques (Williams and Zipser (1989)) are widely used prediction algorithms but are mostly developed for commercial applications (e.g. Face recognition and recommendation systems) for "classification" problem and are not directly applicable to geodesy without further detailed studies.
- N-ARMAX approach allows a larger level of "control and transparency" within the algorithm instead of using machine learning techniques as "black box". Control refers to the ability to configure the initial values, tuning parameters, step size in optimization algorithms. Transparency refers to the availability of reliable information regarding the mathematical equations within the algorithms.

This concludes the discussion of N-ARMAX process, a prospective method for prediction of the model parameter.

4.8 Summary of optimization techniques

We have reached the end of this chapter to summarize the most important points.

[49] GeLu: Gaussian error linear unit and tanh: Hyperbolic tangent function

As shown in Fig. 4.1, the four important aspects discussed in this chapter are (1) Objective function and model parameter (2) constraint function (3) optimality conditions and (4) optimization algorithm. It has been shown that there is a limiting value of the objective function, beyond which it cannot be minimized without violating the constraints. Specifically, the challenge is to identify and stay within the constraint boundaries or the feasible region $\mathcal{F}_R$. It has also been shown that the constraint bounds have a definitive impact on the performance of the optimization algorithm.

The following points are noteworthy in summary of the constraint optimization techniques:

- Most optimization algorithms are based on the descent method, where the descent direction is to be estimated[50].
- The convexity of the objective function and its sensitivity to initial values, convergence of the descent direction and the constraint bounds are important properties of a constraint optimization problem.
- The step size estimation for inequality constraint optimization problems is challenging, especially for a large dimensional (larger than 1000) model parameter vector and good[51] initial values (reasonably close to the optimal solution) are also needed to converge. The initial values used in this work were from within the feasible region.
- Tuning parameters such as the complimentarity measure, centering parameter and step size[52] are the available degrees of freedom. The performance of any algorithm cannot be assessed without trying different values for them. Therefore, a closed loop validation using simulated data is helpful in determining their approximate values.
- Providing initial values from the feasible region, is important for promoting constraint optimization techniques for near real time applications. They can be either obtained from an empirical model but specifically for Lagrange multipliers and slack variables, an initial "guess" is required.

Qualitatively, with respect to the difference between their upper and lower bounds, inequality constraints can be categorized in two types (1) relaxed and (2) critical. In the former, a smaller step size is considered and in latter case, a relatively larger feasible region results and accordingly, the time taken for convergence to the optimal solution can be considerable. In the next chapter, a specific quantitative assessment will be given towards the feasible region size and its impact on the optimal solution.

Optimality conditions allow checking a solution for stability and feasibility. Their reliability must be tested with additional quality parameters, such as standard deviation. In the next chapter, a Monte-Carlo approach will be described for assessing the quality of estimated model parameters. To the best knowledge of the author, there is no universal optimization technique but many different algorithms. Therefore, the formulation of an objective function as well as the choice of optimization algorithm are important.

From the perspective of software prototyping for scientific investigations, a faster convergence of the optimal solution is only one of the desired features. Other requirements such as the desired accuracy, availability of initial values, ability to recover from local non-convexity, step size estimation sub-problem, constraint qualification checks are a few other associated challenges in the optimization technique. Diagnostic checks are found useful, so that a warning may be issued, when the initial values are too far from being realistic. Trial evaluations are helpful to identify the critical parts

[50] More generally, both the descent direction as well as the step size are to be estimated. As a special case, a pre-calibrated step size could be used.

[51] *"Good"* in this context refers to an initial value from within the feasible region.

[52] Within the machine learning community, the step size parameter is also called the learning rate parameter and in general the tuning parameters are called hyperparameters. Some literature also use the name step-length. However, their context remains the same.

in an algorithm. When optimality conditions are not satisfied, some useful information about possible improvements can still be obtained. Furthermore, additional external diagnostic checks and sensitivity analysis can be complimented, for example with reduced model parameters, with another algorithm, changing constraint bounds or with a different subset of observations.

5 Global 4D electron density modelling

Fundamentals of the ionosphere, the observations techniques for electron density modelling (EDM) and the optimization approach have been described in the Chapters 2, 3 and 4 respectively. Specifically, this chapter will focus on the EDM using the BS basis functions, the EDM parametrization using the multi-layer Chapman[1] key parameters, defined in Eq. (5.4) and formulate the EDM as an inequality constrained optimization problem (ICOP) of the form (4.25). A system of constraints[2] will be established and expressed as a linear matrix inequality (LMI), analogous to the KKT system of equations (4.87a), (4.87b) and (4.87c). Furthermore, it will be shown that the inequality constraints on the Chapman key parameters are transformed to those on the corresponding BS coefficients. The unknown optimization parameters of the resulting ICOP are then estimated by solving the KKT system of equations, as already presented in Section 4.5.2; see Eqs. (4.94), (4.95) and (4.96). This chapter is organized as follows:

- An introduction to EDM and the term electron density profile (EDP) are presented in Section 5.1.
- The multi-layer approach to EDM is parameterized in Section 5.2.
- The GMM (5.19) and the equality as well as the inequality constraints in the form of linear matrix inequality (LMI) are combined together to form the Lagrangian function as described in Section 5.3.
- The inequality constraints imposed on the Chapman key parameters are equivalently transformed to those on the unknown B-spline coefficients; see Eq. (5.26) in Section 5.4.
- The solution of the unknown parameters of GMM (5.19), subject to the inequality constraints based on the ICOA described in Chapter 4. The solution of GMM (5.19) is described in Section 5.6.

Furthermore, the following general conventions are applicable for this chapter:

- The term "**optimization algorithm**" will refer to the complete set of KKT equations, that use either the gradient descent or the Newton's method to solve an inequality constraint optimization problem (ICOP) of the form (4.25).
- The term "**nominal ionospheric condition**" (or simply "nominal conditions") will refer to the ionospheric behaviour under moderate space weather conditions[3] as opposed to abnormalities during e.g. geomagnetic storms. Under nominal conditions, the ionosphere follows a rather predictable diurnal, monthly and seasonal climatology (see Hanslmeier (2002), Hapgood (2017)).
- **"Top-"** and **"bottom-side"** will generally refer to the region of ionosphere above and below the F_2 peak altitude respectively.

5.1 Background to electron density modelling

As described already in Chapter 2, the electron density is a 4D function of latitude, longitude, altitude and time. Of particular interest in this chapter is the variation of the electron density along

[1] Sydney Chapman was one of the pioneering geophysicists during the 20th century and also a pioneer in space physics (Akasofu (2020)). Chapman devoted his study on the transport property of gases and wrote "The Mathematical Theory of Non-Uniform Gases".

[2] Generally, containing both equality and inequality constraints

[3] e.g. charceterized by F10.7 index < 100.

the altitude at any given location and time. This is generally referred to as an electron density profile (EDP). More specifically, an EDP represents the number density (or distribution) of electrons along the vertical and is graphically visualized by plotting the electron density along vertical at a given latitude, longitude and time, as sketched in Fig. 5.1. The following are some of the important aspects related with an EDP:

- The study of the EDPs since the early 20th century (see Chapman (1931) revealed an approximate "bell-shape" at the different times of a day. Generally, the electron density increases upto the F_2 layer peak height and then decreases with altitude. The corresponding rate of change of the electron density in the bottom− and the top−sides respectively, are due to the ionization and the recombination processes (see e.g. Smirnova et al. (1988), Stankov. and Jakowski (2006)).
- Although theoretically an EDP can be defined from the surface of the Earth to any arbitrary large[4] altitude, it is practical in an electron density modelling problem (EDMP) to define fixed altitudinal limits[5]. In this work, these are denoted as h_{min} and h_{max} (see Section 5.2).
- As later described in Section 5.2, each EDP will be modelled as a sum of electron densities of the individual stratified $D-$, $E-$, F_1-, F_2- layers as well as the plasmasphere; see Eq. (5.1).
- As already mentioned, each EDP itself can be theoretically divided into two parts, namely the "top-" and the "bottom-side". The solar irradiance to the ionosphere-thermosphere system is absorbed in the top- and bottom-sides according to the exact concentration of the Nitrogen, Oxygen, Hydrogen, Helium and other gases compounds comprising of those elements, e.g. Nitrous Oxide (see Hargreaves (1992), Richmond et al. (1992), Picone et al. (2002), Luehr et al. (2011)).

The exact shape of an EDP, at any given location, generally varies[6] with the time as well as the space weather conditions (see e.g. Hanslmeier (2002), Stolle et al. (2013)).

In this work, the EDP is parametrized using the Chapman function[7], for which three so called "key parameters" namely, the peak density, the peak height and the scale height, for each individual layers will be introduced; see Eq. (5.4). Peak density at a location $\{\varphi, \lambda\}$, and a given time, corresponds to the maximum magnitude of the electron density and the corresponding altitude is referred to as the "peak height". The third key parameter "scale height", of a layer Q, refers to the width of the Q-layer EDP around its peak height. Generally, all the three key parameters vary on spatio-temporal scales[8] (see Limberger et al. (2013), Limberger et al. (2014), Tsai et al. (2016)).

[4] Generally upto infinity, but beyond 6 - 8 Earth radii altitude, there are not many sources of electron density measurements to understand their spatial distributions. As mentioned in Chapter 3, Van Allen probes orbit between 2 - 6 Earth radii provide in-situ observations.

[5] This helps in restricting the modelling to a reasonable altitude where there is an availablity of electron density measurements. Furthermore, as ionosphere transitions into plasmasphere, more realistic modelling assumptions will have to be made on the magnitude its scale height (see e.g. Stankov. and Jakowski (2006), Moldwin et al. (1995), Carpenter and Park (1973)).

[6] But follows a certain consistent underlying physical phenomena, as mentioned already.

[7] In other words, each EDP is modelled empirically using the Chapman profile function which was already introduced in Chapter 2.

[8] An interesting synergy exists between Chapman profile function (CPF) and normal distribution function (NDF), where the peak density, peak height and scale height of a CPF are analogous to the magnitude, mean and standard deviation of an NDF.

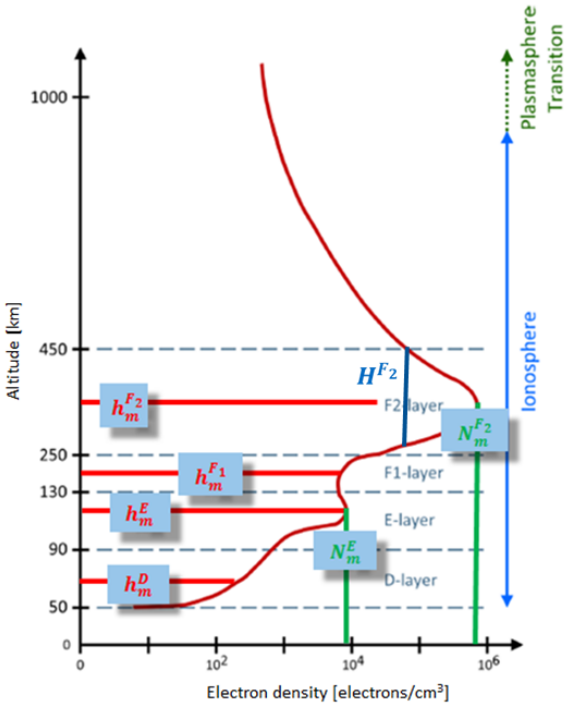


Figure 5.1: Electron density profile and key parameters of the ionosphere. Courtesy Limberger (2015). The behavior of the total electron density with altitude and magnitude of peak densities and peak height to expect in the different layers are shown. The electron density drops above the F_2-layer peak height as the ionosphere transitions into the plasmasphere.

5.2 Parametrization for multi-layer approach

Following Feltens (2007), a multi-layer model is defined as the sum of five EDPs, one for each layer. Accordingly, the electron density modelling[9] is posed as an additive model

$$\begin{aligned} n_e(h) &= n_e^D(h) + n_e^E(h) + n_e^{F1}(h) + n_e^{F2}(h) + n_e^P(h) \\ &= \sum_{Q=1}^{4} N_m^Q \, p^Q(h) + N_0^P p^P(h) \\ &Q{=}1 \text{ to } 4 \text{ correspond to the layers } \{D, E, F_1, F_2\}, \end{aligned} \tag{5.1}$$

which amounts to representing the EDP of each layer as a product of the peak density N_m^Q and a profile function $p^Q(h)$ as well as $p^P(h)$. n_e^P and N_0^P denote the total electron density and the basis electron density of the plasmasphere respectively. Limberger (2015) and Liang (2017) have also used the approach (5.1) in the scope of regional ionosphere modelling. The Chapman profile function,

$$n_e^Q(h) = N_m^Q \, \exp\left(\frac{1}{2}\left(1 - \frac{h - h_m^Q}{H^Q} - \exp\left(-\frac{h - h_m^Q}{H^Q}\right)\right)\right) \tag{5.2}$$

is used to describe the EDP for the Q–layer. Furthermore, plasmaspheric electron density

$$n_e^P(h) = N_0^P \, \exp\left(-\frac{|h - h_m^{F_2}|}{H^P}\right) \tag{5.3}$$

(Limberger (2015)) is characterized by the profile function $p^P(h) = \exp\left(-\frac{|h-h_m^{F_2}|}{H^P}\right)$ as a function of its scale height H^P. It shall be noted that the plasmaspheric profile function $p^P(h)$ is an exponential

[9] n_e is used instead of the more common N_e to be consistent with the convention of using bold small letters for vector and bold capital ones for matrices. Later on, electron density vector will be denoted as $\boldsymbol{n}_e$.

decay instead of a double exponential function used in $p^Q(h)$ for the other layers[10]. In this work, $p^Q(h)$ and $p^P(h)$ denote the empirical functions (5.2) and (5.3) respectively, defined within fixed altitude limits $\{h_{min}, \cdots, h_{max}\}$. An example of peak density N_m^Q, peak height h_m^Q and scale height H^Q are shown in Fig. 5.1. Having introduced the KPs and their relation to total electron density through (5.2) and (5.3), electron modelling problem in its most general form will now be formulated, which means in Eq. (5.1), the key parameters N_m^Q, and $p^Q(h)$, N_0^P and H^P are unknown as well as varying on spatio-temporal scales, i.e. $N_m^Q(\varphi, \lambda, t)$, $h_m^Q(\varphi, \lambda, t)$ and $H^Q(\varphi, \lambda, t)$ completely define their respective magnitudes at a given location and time. However, in this work, these parameters are modelled at each measurement time epoch in "batch-mode[11]" and therefore are considered as two-dimensional spatial functions. Furthermore, the spatial indices will be ignored for better readability, where it is implicitly understood[12]. Total electron density $n_e(\varphi, \lambda, h)$ and its functional such as STEC (see Appendix D.2) are expressed as empirical functionals of KPs. Accordingly, the multi-layer model (5.1) considering the Eqs. (5.2) and (5.3) contains the set of unknown key parameters

$$\mathcal{K} = \{N_m^{F_2}, h_m^{F_2}, H^{F_2}, N_m^{F_1}, h_m^{F_1}, H^{F_1}, N_m^E, h_m^E, H^E, N_m^D, h_m^D, H^D, N_0^P, H^P\} \tag{5.4}$$

of the five layer, where

- h is the height above the Earth in km
- N_m^D, h_m^D, H^D are the peak electron density, peak height and scale height of the D layer in elec./m^3, km, km respectively. It will be shown later that a more convenient unit of electron density 1 EDU = 10^{12} elec./m^3 will be applied.
- N_m^E, h_m^E, H^E are the peak electron density, peak height and scale height of the E layer in EDU, km, km respectively
- $N_m^{F_1}$, $h_m^{F_1}$, H^{F_1} are the peak electron density, peak height and scale height of the F_1 layer in EDU, km, km respectively
- $N_m^{F_2}$, $h_m^{F_2}$, H^{F_2} are the peak electron density, peak height and scale height of the F_2 layer in EDU, km, km respectively.

Therefore, a total of 14 key parameters would have to be estimated from the observations[13] at each time epoch. Since the key parameters in the set $\mathcal{K}$ describe the electron density of the ionosphere and the plasmasphere according to the Chapman profile function (5.2) and the exponential decay function (5.3), significant correlations exist between the estimated sets of B-spline coefficients $d_{k_1,k_2:\kappa_1}^{J_1,J_2}$ and $d_{k_1,k_2:\kappa_2}^{J_1,J_2}$ of various pairs $\{\kappa_1, \kappa_2\}$ of key parameters. Here the physical and statistical correlation needs to be distinguished. Whereas the latter, for instance, exists between all the parameters in the modelling problem, a physical correlation has to be considered between $N_m^{F_2}$ and $h_m^{F_2}$ (see e.g. Limberger et al. (2014)). Such correlations exist during both nominal as well as adverse space weather conditions (Private communication with Dr. Dieter Bilitza, 2020).

Each key parameter $\kappa \in \mathcal{K}_1$ in (5.4) is a spatial function of latitude, longitude $\{\varphi, \lambda\}$ at a given time. The set $\mathcal{K}$ of key parameters in (5.4) can be extended further, for example by choosing a different scale height for the "top-" and the "bottom-sides", i.e. H_{top}^Q, H_{bottom}^Q for each layer Q. As the Chapman key parameters $\kappa \in \mathcal{K}$ are spatio-temporal functions, an appropriate set of basis functions are sought for their global representation. Any two-dimensional (2D) spatial function, such as N_m^Q or h_m^Q to be modelled on a sphere is traditionally represented by series expansion in spherical harmonics (SH); see e.g. Schaer (1999). As already mentioned in Chapter 2, due to the challenges

[10] It shall be clarified that the Chapman key parameters peak density N_m^Q and scale height H^Q are traditionally denoted in capitalized first letters, whereas the peak height h_m^Q is denoted by small-caps letter. Although this is a slight deviation from the convention of using bold capital for matrices and bold small-caps for vectors, adopted in this work but it helps in maintaining the traditional convention seen in many of the cited literature.

[11] The term batch mode refers to the approach of estimating unknown parameters at each meausurement epoch.

[12] This helps in maintaining the readability of equations.

[13] electron density and STEC observations

in the representation of functions with a heterogenous data distribution (see Schmidt (2011)), in this work, B-spline basis functions are chosen becuase they are more suitable for the representation of multidimensional signals and for fitting surfaces to given measurements; see Schumaker and Traas (1991), Stollnitz et al. (1995), Lyche and Schumaker (2000), Koch (1999). Besides, B-splines also have a successful legacy of use in ionosphere modelling at DGFI-TUM[14]; see Zeilhofer (2008), Koch and Schmidt (2010), Limberger et al. (2013), Limberger et al. (2014), Liang et al. (2015a), Schmidt et al. (2015) Erdogan et al. (2017), Goss et al. (2019), Goss et al. (2020), Erdogan et al. (2020). As shown in Fig. 5.2, for modelling any key parameter $\kappa \in \mathcal{K}_1$, the B-spline series expansion

$$\kappa(\varphi,\lambda) + e_\kappa(\varphi,\lambda) = \sum_{k_1=0}^{K_{J_1}-1} \sum_{k_2=0}^{K_{J_2}-1} d_{k_1,k_2:\kappa}^{J_1,J_2} N_{k_1}^{d_N}(\varphi)\, T_{k_2}^{d_T}(\lambda) \tag{5.5}$$

is applied including the $K_{J_1} \cdot K_{J_2}$ unknown B-spline series coefficients[15] $d_{k_1,k_2:\kappa}^{J_1,J_2}$, with $K_{J_1} = 2^{J_1} + 2$, $K_{J_2} = 3 \cdot 2^{J_2}$ (see Schmidt et al. (2015)), the endpoint interpolating polynomial B-splines (PB) $N_{k_1}^{d_N}(\varphi)$, the trigonometric B-splines (TB) $T_{k_2}^{d_T}(\lambda)$ of degree d_N,d_T respectively (both d_N and d_T, are set to 2 for this work), J_1, J_2 are the B-spline levels, $k_1 = \{0,\cdots,K_{J_1}-1\}$, $k_2 = \{0,\cdots,K_{J_2}-1\}$ are the shifts and the truncation error e_κ; see Fig. 5.2. Detailed computation of PB and TB are already provided in Schmidt et al. (2015); pages 945-950, Limberger (2015), Goss et al. (2019) and therefore are not repeated here. The unknown B-spline coefficients $d_{k_1,k_2:\kappa}^{J_1,J_2}$ to be estimated from the observed functionals of the electron density, such as STEC or ionospheric radio occultation (IRO) measurements. The other advantages of BSs such as their localizing feature have been already introduced in Chapter 2 and described in detail in Schmidt et al. (2015). Although different B-spline levels J_1 and J_2 can be considered for the key parameters κ in set $\mathcal{K}_1$, for simplicity, same levels are chosen here, resulting in the same number of coefficients for each key parameter. More generally, the choice of B-spline levels J_1 and J_2 is following the sampling interval of the observations. For details on the relations between spectral content, sampling intervals and B-spline levels see Goss et al. (2019).

The key parameters in the set $\mathcal{K}$, and hence the corresponding B-spline coefficients $\boldsymbol{d}_\kappa$, have the following physical characteristics:

- The key parameters $\kappa \in \mathcal{K}$ are always positive and specifically $N_m^{F_2}(\varphi,\lambda) \neq 0 \;\; \forall \; \varphi \in \boldsymbol{\varphi}$ and $\lambda \in \boldsymbol{\lambda}$, where $\boldsymbol{\varphi}$, $\boldsymbol{\lambda}$ refer to the set of latitudes and logitudes of the observation locations. In this work, the $D-$ and the $E-$ layers as well as the plasmasphere are assumed to be existing during day as well as the night (see Oshiorenoya (2004)) but not the F_1 layer, i.e. $N_m^{F_1}(\varphi,\lambda,\boldsymbol{t}_N) = 0$ for time epochs $\boldsymbol{t}_N \in \{21:00-03:00\}$ hours[16]; see Berkner and Wells (1934), Bremer (1998). In other words, $N_m^D(\varphi,\lambda,t) \neq 0$, $N_m^E(\varphi,\lambda,t) \neq 0$, $N_m^{F_2}(\varphi,\lambda,t) \neq 0$ and $N_0^P(\varphi,\lambda,t) \neq 0 \;\; \forall \; t \in \boldsymbol{t}$, where $\boldsymbol{t}$ denotes the vector of measurement time epochs.
- Furthermore, based on magnitude of peak electron densities (Liang (2017))

$$N_m^{F_2}(\varphi,\lambda,t) \;>\; N_m^{F_1}(\varphi,\lambda,t) \;>\; N_m^{E}(\varphi,\lambda,t) \;>\; N_m^{D}(\varphi,\lambda,t) \tag{5.6}$$

  applies at those time epochs when all the layers are existing, i.e $t \notin \boldsymbol{t}_N$. Physical realism in manifested by eq. (5.6) because the number density of electrons in ionosphere layers are so arranged (see Liang (2017)). Once again, for the sake of readability, the time indices will be ignored and the Chapman key parameters are considered only as spatial functions henceforth.

[14] DGFI-TUM: Deutsches Geodatisches Forschungsinstitut, der Technische Universitat Muenchen, The hosting agency for the author

[15] Also sometimes referred to as B-spline coefficients

[16] At best, this was only a modelling assumption and a different night time definition could have been also chosen, if so desired.

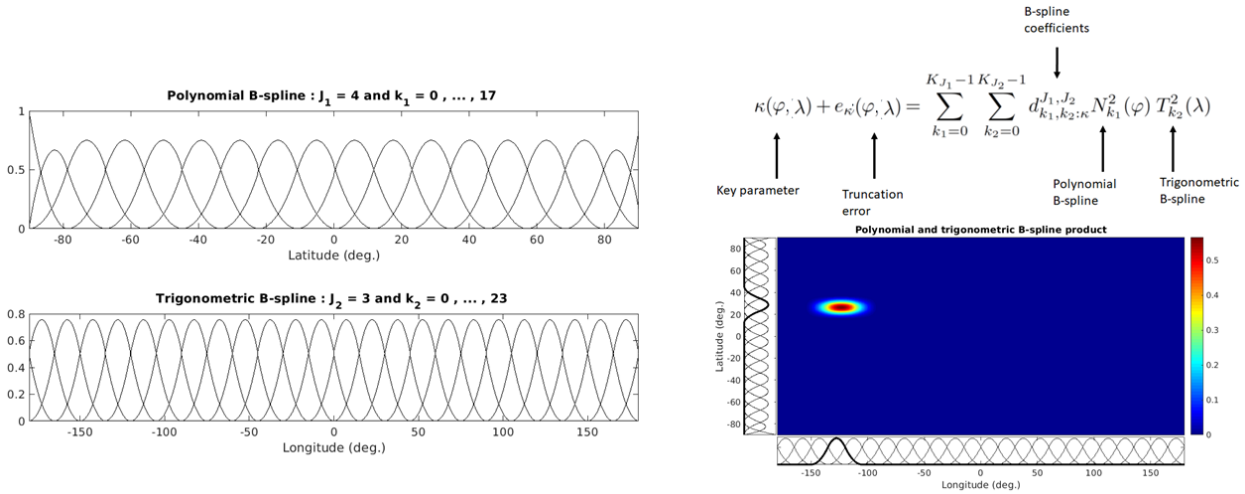


Figure 5.2: B-spline representation and tensor product. Key parameter κ (peak density, peak height or scale height) is represented in a B-spline series expansion and the corresponding unknown coefficients are estimated. Polynomial B-splines (top left) and trigonometric B-spline (bottom left) are shown for the chosen level $J_1 = 4$ and $J_2 = 3$ respectively. Accordingly, there are $K_{J_1} = 2^{J_1} + 2 = 18$ polynomial splines and $K_{J_2} = 3 \cdot 2^{J_2} = 24$ trigonometric splines, shown with indices $k_1 = 0, 1, \ldots, 17$ along the latitude and $k_2 = 0, 1, \ldots, 23$ along the longitude respectively. The product of polynomial and trigonometric B-splines is shown on right side.

- The key parameters exhibit a so called "physical realism"[17] with respect to the ionospheric stratification and the localization of the individual layers; i.e.

$$h_{min} < h_m^D(\varphi, \lambda, t) < h_m^E(\varphi, \lambda, t) < h_m^{F_1}(\varphi, \lambda, t) < h_m^{F_2}(\varphi, \lambda, t) < h_{max} \tag{5.7}$$

  applies at those time epochs t, such that $t \in \boldsymbol{t}$ and $t \notin \boldsymbol{t}_N$, i.e. when all the layers are existing.
- More importantly, the key parameters values lie within specific upper and lower bounds. This property is the basis for the formulation of inequality constraints; see e.g. Eqs. (5.8a) - (5.8d)

Therefore, the overarching goal of this work is to enable the use of above mentioned physically-realistic prior information in EDM by parametrizing those as inequality constraints on the Chapman key parameters $\kappa \in \mathcal{K}_1$. There exist different techniques to include equality constraints within a GMM (see e.g. Chapter 3 in Koch (1999)) but the novelty in this work is the development of a generalized approach for inclusion of both equality as well inequality constraints in an optimization problem for parameter estimation. Furthermore, these inequality constraints will be expressed as a so-called "linear matrix inequality" (LMI), as shown in (5.17).

The following example demonstrates the physical realism of (5.7) using examplary absolute upper bounds (AUB) and absolute lower bounds (ALB) for peak height, at a $\{\varphi, \lambda\}$, of the different ionosphere layers, yielding the inequalities

$$80 \text{ km} \leq h_m^D(\varphi, \lambda) \leq 100 \text{ km} \tag{5.8a}$$
$$108 \text{ km} \leq h_m^E(\varphi, \lambda) \leq 112 \text{ km} \tag{5.8b}$$
$$120 \text{ km} \leq h_m^{F_1}(\varphi, \lambda) \leq 200 \text{ km} \tag{5.8c}$$
$$220 \text{ km} \leq h_m^{F_2}(\varphi, \lambda) \leq 440 \text{ km}. \tag{5.8d}$$

[17] This implicitly refers to the general idea that the layers are stratified on the top of one another, with the respective peak heights of the D–, E–, F_1– and F_2–layer lying above the preceding layers.

Subtracting (5.8a) from (5.8b), (5.8b) from (5.8c), and (5.8c) from (5.8d), respectively lead to the three relative constraints

$$8 \text{ km} \leq h_m^E(\varphi,\lambda) - h_m^D(\varphi,\lambda) \leq 42 \text{ km} \tag{5.9a}$$

$$8 \text{ km} \leq h_m^{F_1}(\varphi,\lambda) - h_m^E(\varphi,\lambda) \leq 92 \text{ km} \tag{5.9b}$$

$$20 \text{ km} \leq h_m^{F_2}(\varphi,\lambda) - h_m^{F_1}(\varphi,\lambda) \leq 320 \text{ km}, \tag{5.9c}$$

with their relative upper bounds (RUB) and relative lower bounds (RLB) , which in-turn are linear inequalities as well. For e.g., the physical interpretation of (5.9c) is that the modelled $h_m^{F_2}$ shall be at least 20 km, and atmost 320 km, in altitude above $h_m^{F_1}$. Importantly, in the Eqs. (5.9a) to (5.9c), the left hand side RLBs are non-zero. It means, the peak heights of no two layers should ever coincide[18] and furthermore, these RLBs are positive as well, meaning that the physical realism of (5.7) is maintained[19]. It shall be noted that Eqs. (5.8a) to (5.9c) are only shown as example for the transformation and mathematical representation of physically realistic conditions (5.7). As will be mentioned in Section 5.2, only selected key parameters $\kappa \in \mathcal{K}_1$ are chosen. When a key parameter of a certain layer is considered known (or given), then those are used to construct the so called "equality constraints". For example, if the D, E and F_1 layer key parameters are not considered in the set $\mathcal{K}_1$, but instead in $\mathcal{K}_2$, then the inequalities (5.8a)-(5.8c) become equality constraints. In this way a combination of both the equality as well as inequality constraints are used in EDM. It shall be noted that the lower and upper bounds κ_u, κ_l $\forall$ $\kappa \in \mathcal{K}_1$, are chosen after an investigation using the IRI model.

Using the B-spline representation of the key parameters, EDMP (5.1) is transformed into an equivalent one where the B-spline coefficients are to be estimated. Analogous to Eq. (5.5), ignoring the truncation error, the key parameters $N_m^{F_2}$, $h_m^{F_2}$ and H^{F_2} at any $\{\varphi,\lambda\}$ at a given time are represented by their respective B-spline series

$$N_m^{F_2}(\varphi,\lambda) = \sum_{k_1=0}^{K_{J_1}-1} \sum_{k_2=0}^{K_{J_2}-1} d^{J_1,J_2}_{k_1,k_2:N_m^{F_2}} \quad N_{k_1}^{d_N}(\varphi)\; T_{k_2}^{d_T}(\lambda) \tag{5.10}$$

$$h_m^{F_2}(\varphi,\lambda) = \sum_{k_1=0}^{K_{J_1}-1} \sum_{k_2=0}^{K_{J_2}-1} d^{J_1,J_2}_{k_1,k_2:h_m^{F_2}} \quad N_{k_1}^{d_N}(\varphi)\; T_{k_2}^{d_T}(\lambda) \tag{5.11}$$

$$H^{F_2}(\varphi,\lambda) = \sum_{k_1=0}^{K_{J_1}-1} \sum_{k_2=0}^{K_{J_2}-1} d^{J_1,J_2}_{k_1,k_2:H^{F_2}} \quad N_{k_1}^{d_N}(\varphi)\; T_{k_2}^{d_T}(\lambda). \tag{5.12}$$

If desired, the key parameters of the other layers to be included in the modelling problem are also evaluated in a similar manner. Consequently, the joint estimation of the B-spline coefficients of all key parameters in the set $\mathcal{K}$ is very challenging. Alternately, a subset $\mathcal{K}_1 \subset \mathcal{K}$ is chosen for modelling and the remaining subset $\mathcal{K}_2 \subset \mathcal{K}$ from the complimentary set of $\mathcal{K}_1$ ($\mathcal{K}_2 = \mathcal{K}_1 \setminus \mathcal{K}$) are assumed given (e.g. from model values).

5.3 Formulation of constrained optimization problem

The upper and lower bound on key parameters $\kappa \in \mathcal{K}_1$ are formulated as inequalities and transformed equivalently into the corresponding B-spline coefficients so that estimated optimization parameters are within physically realistic limits of electron density and altitude. Consequently, the inequality

[18] Only as long as the constraint sare satisfied. If the constraints are violated, then the relative lower bounds could become zero or even negative.

[19] This is generally only true for nominal space weather conditions.

constrained optimization technique, as described in Chapter 4, is used to estimate the B-spline coefficients $d_{k_1,k_2:\kappa}^{J_1,J_2}$ of the selected key parameters $\kappa \in \mathcal{K}_1$ subject to the inequality constraints

$$\kappa_l(\varphi,\lambda) \leq \kappa(\varphi,\lambda) \leq \kappa_u(\varphi,\lambda), \tag{5.13}$$

where the lower bound functions $\kappa_l(\varphi,\lambda)$ and the upper bound functions $\kappa_u(\varphi,\lambda)$ are given and represent physically realistic limits, as shown in e.g. (5.7). By including an example peak density[20] constraint, in addition to peak height, the pair of inequalities

$$\begin{aligned} 220 \text{ km} \leq h_m^{F_2}(\varphi,\lambda) \leq 440 \text{ km} \\ 0.02 \text{ EDU} \leq N_m^{F_2}(\varphi,\lambda) \leq 2.2 \text{ EDU} \end{aligned} \tag{5.14}$$

shall be satisfied[21] simultaneously. Besides the inequality (5.13), the equality constraints

$$\kappa(\varphi,\lambda) = \kappa_e(\varphi,\lambda) \tag{5.15}$$

must hold for the remaining key parameters $\kappa \in \mathcal{K}_2$ with a given, physically realistic function $\kappa_e(\varphi,\lambda)$. Using the series expansion (5.5), the bound functions $\kappa_l(\varphi,\lambda)$, $\kappa_u(\varphi,\lambda)$ and $\kappa_e(\varphi,\lambda)$ are transformed into the lower and upper bounds on the corresponding B-spline coefficients $d_{k_1,k_2:\kappa_l}^{J_1,J_2}$, $d_{k_1,k_2:\kappa_u}^{J_1,J_2}$ and $d_{k_1,k_2:\kappa_e}^{J_1,J_2}$. The "bounded parameter space" enclosed by the constraint bounds κ_l and κ_u is the feasible region[22] $\mathcal{F}$ for the key parameters $\kappa \in \mathcal{K}_1$ (see Section 4.2.1). In Chapter 6, it will be shown that the lower and upper bounds for key parameters are varying with time.

To demonstrate the application of inequality constrained optimization to EDM, the subset

$$\mathcal{K}_1 = \{\kappa_1 = N_m^{F_2}, \kappa_2 = h_m^{F_2}, \kappa_3 = H^{F_2}, \kappa_4 = N_0^P, \kappa_5 = H^P\} \tag{5.16}$$

of the 5 key parameters is introduced, out of which three belong to the F_2-layer and the remaining two to the plasmasphere.

Analogous to (4.8), to set up a GMM for EDM, Eq. (5.1) is linearized by considering the partial derivative of Eqs. (5.2) and (5.3) with respect to the B-spline coefficients $d_{k_1,k_2:\kappa}^{J_1,J_2}$ of a selected key parameter $\kappa \in \mathcal{K}_1$. Collecting the B-spline coefficients in the vector $\boldsymbol{d}_\kappa = \left[d_{0,0;\kappa}^{J_1,J_2}, d_{0,1;\kappa}^{J_1,J_2}, \cdots, d_{K_{J_1}-1,K_{J_2}-1;\kappa}^{J_1,J_2}\right]^T$, decomposing it into the vector $\boldsymbol{d}_{k;0}$ of the initial values and the vector $\Delta\boldsymbol{d}_\kappa$ of the corrections, the inequality constraints (5.13) are expressed as a linear matrix inequality (LMI) $\boldsymbol{B}_\kappa \boldsymbol{d}_\kappa \leq \boldsymbol{b}_\kappa$ for $\kappa \in \mathcal{K}_1$, which upon substituting $\boldsymbol{d}_\kappa = \boldsymbol{d}_{\kappa;0} + \Delta\boldsymbol{d}_\kappa$, gives

$$\boldsymbol{B}_\kappa \left(\boldsymbol{d}_{\kappa;0} + \Delta\boldsymbol{d}_\kappa\right) \leq \boldsymbol{b}_\kappa \tag{5.17}$$

and upon further rearranging gives $\boldsymbol{B}_\kappa \Delta\boldsymbol{d}_\kappa \leq \boldsymbol{b}_\kappa - \boldsymbol{B}_\kappa \boldsymbol{d}_{\kappa;0}$ or compactly written as

$$\boldsymbol{B}_\kappa \Delta\boldsymbol{d}_\kappa \leq \Delta\boldsymbol{b}_\kappa, \tag{5.18}$$

where $\boldsymbol{B}_\kappa$ is a given constraint coefficient matrix and $\Delta\boldsymbol{b}_\kappa = \boldsymbol{b}_\kappa - \boldsymbol{B}_\kappa \boldsymbol{d}_{\kappa;0}$ is the constraint bound vector, which denotes the difference between a given bound and that computed from the initial or updated B-spline coefficients. The given bound $\boldsymbol{b}_\kappa$ means the constraint bound vector related to the lower bound function $\boldsymbol{\kappa}_l$ and the upper bound function $\boldsymbol{\kappa}_u$ of the key parameter κ.

To set up the estimation model, we introduce the vector $\boldsymbol{d}_\kappa = \left[\boldsymbol{d}_{\kappa_1}^T, \cdots, \boldsymbol{d}_{\kappa_5}^T\right]^T$ of the B-spline coefficients for all 5 key parameters κ_1 to κ_5, the corresponding vector $\boldsymbol{d}_{\kappa;0} = \left[\boldsymbol{d}_{\kappa_1;0}^T \quad \cdots \quad \boldsymbol{d}_{\kappa_5;0}^T\right]^T$ of the initial B-spline coefficient values, the vector $\Delta\boldsymbol{d}_\kappa = \left[\Delta\boldsymbol{d}_{\kappa_1}^T \quad \cdots \quad \Delta\boldsymbol{d}_{\kappa_5}^T\right]^T$ of the B-spline

[20] For convenience, the electron density units (EDU) is defined as 1 $EDU = 10^{12}$ elec/m^3

[21] This is just one example of the constraint bounds.

[22] Feasible region was introduced in Chapter 4; definition 4.2.2

coefficient corrections, the vector $\boldsymbol{b}_\kappa = \begin{bmatrix}\boldsymbol{b}_{\kappa_1}^T & \cdots & \boldsymbol{b}_{\kappa_5}^T\end{bmatrix}^T$ of the constraint bounds as well as the matrix $\boldsymbol{B}_\kappa = [\boldsymbol{B}_{\kappa_1} \cdots \boldsymbol{B}_{\kappa_5}]$ of the constraint coefficient matrices and the matrix $\boldsymbol{A}_\kappa = [\boldsymbol{A}_{\kappa_1} \cdots \boldsymbol{A}_{\kappa_5}]$ of the partial derivatives of the electron density with respect to the B-spline coefficients of the key parameters κ_1 to κ_5 (see Section 5.5). Furthermore, according to Eq. (5.15), the equality constraints $\boldsymbol{E}_\kappa \boldsymbol{d}_\kappa = \boldsymbol{e}_\kappa$ are defined, where $\boldsymbol{E}_\kappa$ is the equality constraint coefficient matrix and $\boldsymbol{e}_\kappa$ is the corresponding given value related to the function κ_e of the key parameter $\kappa \in \mathcal{K}_2$. In order to be consistent with the convention of representing matrices in bold captial letters and vectors with bold small caps letter, $\boldsymbol{n}_e$ will be used to refer to the vector of electron density[23]. Finally, the vector $\Delta\boldsymbol{n}_e = \boldsymbol{n}_e - \boldsymbol{n}_{e;0}$ of the differences between the observed electron density $\boldsymbol{n}_e$ and the vector $\boldsymbol{n}_{e;0}$ of the corresponding values computed from the vector $\boldsymbol{d}_0$ of the initial B-spline coefficient values and the vector $\boldsymbol{d}_\kappa$ for the key parameters are introduced. With the aforementioned matrices and vectors the GMM

$$\Delta\boldsymbol{n}_e + \boldsymbol{e}_{n_e} = \boldsymbol{A}\Delta\boldsymbol{d}_\kappa \quad \text{with} \quad \boldsymbol{D}(\Delta\boldsymbol{n}_e) = \sigma_0^2 \boldsymbol{P}^{-1} \tag{5.19}$$

is introduced, subject to the LMI $\boldsymbol{B}_\kappa \Delta\boldsymbol{d}_\kappa \leq \Delta\boldsymbol{b}_\kappa \quad \text{for} \quad \kappa \in \mathcal{K}_1$ and $\boldsymbol{E}_\kappa \boldsymbol{d}_\kappa = \boldsymbol{e}_\kappa \quad \text{for} \quad \kappa \in \mathcal{K}_2$ respectively. In the model part of (5.19), $\boldsymbol{e}_{n_e}$ denotes the random vector of the unknown measurement errors, $\boldsymbol{P}$ is a given positive definite observation weight matrix and σ_0^2 is the unknown variance factor. Thus, we have derived the GMM of EDMP in (5.19) analogous to (4.8). Depending on the observation techniques used, the design matrix $\boldsymbol{A}_\kappa$ shall be cascaded with the corresponding partial derivative sub-matrices. For example when STEC observations are considered in addition to IRO electron density, then the sub-matrices $\frac{\partial STEC}{\partial \boldsymbol{d}_\kappa}$, as shown in Appendix E.

5.4 Constraint transformation

It has been shown in the previous section that a 4D EDM problem (5.1) subject to physically realistic prior information, e.g. (5.7) are systematically transformed to a GMM (5.19). Implicitly, in this procedure, there is a transformation between the applied constraints on the Chapman key parameters $\kappa \in \mathcal{K}_1$ to that on the corresponding B-spline coefficients $\boldsymbol{d}_\kappa$. This transformation is an important design detail in EDM to ensure that the estimated B-spline coefficients also satisfy the KKT system of equations (4.87a), (4.87b) and (4.87c).

In this section, the above mentioned transformation of the inequality constraints on key parameters $\kappa \in \mathcal{K}_1$ to their corresponding B-spline coefficients will be described. Subsequently, the coefficient matrix $\boldsymbol{B}_\kappa$ and constraint bound vector $\boldsymbol{b}_\kappa$ will be derived.

The upper and lower bounds on the F_2-layer peak density and the peak height as examples of inequality constraints at a location is illustrated in Fig. 5.3, where the day and night side EDPs are shown alongside, indicating that the bounds are time dependent. As described before, the inequality constraints are applied on each key parameter κ in the set $\mathcal{K}_1$ and, in general, they are dependent on location and time (and hence, by extension on the space weather phenomena). In order to derive $\mathbf{B}_\kappa$, referring to (5.13) and taking a specific example of the lower and upper bounds

$$\begin{aligned}\kappa(\varphi,\lambda) &\geq \kappa_l(\varphi,\lambda)\\ \kappa(\varphi,\lambda) &\leq \kappa_u(\varphi,\lambda)\end{aligned} \tag{5.20}$$

at a certain location $\{\varphi,\lambda\}$ and time for the key parameter $\kappa \in \mathcal{K}_1$, the following transformation of constraints to B-spline coefficients will be shown (once again at a given time instant and therefore the time indices are ignored in this section). Multiplying the lower bound inequality (5.20) by -1 on both sides yields

$$-\kappa(\varphi,\lambda) \leq -\kappa_l(\varphi,\lambda), \tag{5.21}$$

[23] Although this is a small change from the earlier classical convention of using n_e to represent the electron density, it will serve to maintain the mathematical consistency.

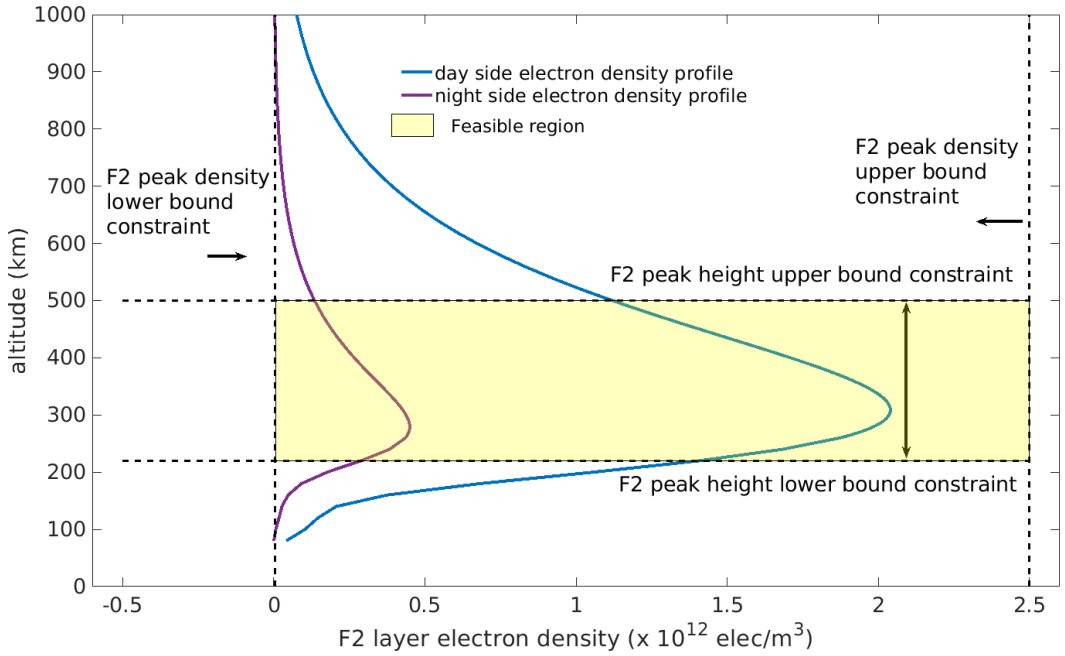


Figure 5.3: F_2-layer electron density profile with constraint bounds on the key parameters and feasible region.

to the convex standard form[24] and substituting in B-spline representation (5.5) gives

$$-\sum_{k_1=0}^{K_{J_1}-1}\sum_{k_2=0}^{K_{J_2}-1} d_{k_1,k_2:\kappa}^{J_1,J_2}\; N_{k_1}^{d_N}(\varphi)\; T_{k_2}^{d_T}(\lambda) \leq -\kappa_l(\varphi,\lambda). \tag{5.22}$$

The upper bound inequality $\kappa(\varphi,\lambda) \leq \kappa_u(\varphi,\lambda)$ in (5.20) is already in convex standard form and substituting (5.5) in (5.20) gives

$$\sum_{k_1=0}^{K_{J_1}-1}\sum_{k_2=0}^{K_{J_2}-1} d_{k_1,k_2:\kappa}^{J_1,J_2}\; N_{k_1}^{d_N}(\varphi)\; T_{k_2}^{d_T}(\lambda) \leq \kappa_u(\varphi,\lambda). \tag{5.23}$$

The product of polynomial $N_{k_1}^{d_N}(\varphi)$ and trigonometric B-splines $T_{k_1}^{d_T}(\lambda)$ at observation locations $\varphi \in \boldsymbol{\varphi}$ and $\lambda \in \boldsymbol{\lambda}$ is contained in $\boldsymbol{\Psi}_{NT}$ matrix, also referred as the tensor product matrix; see Schmidt (2015). Combining Eqs. (5.22) and (5.23), the linear matrix inequality (LMI)

$$\begin{aligned} -\boldsymbol{\Psi}_{NT}\; \boldsymbol{d}_\kappa &\leq -\boldsymbol{\kappa}_l \\ \boldsymbol{\Psi}_{NT}\; \boldsymbol{d}_\kappa &\leq \boldsymbol{\kappa}_u \end{aligned} \tag{5.24}$$

is obtained, which represents the global upper and lower bound constraints together. Comparing Eqs. (5.18), (5.17) and (5.24), the constraint coefficient matrix and constraint bound vector

$$\boldsymbol{B}_\kappa = \begin{bmatrix} -\boldsymbol{\Psi}_{NT} \\ \boldsymbol{\Psi}_{NT} \end{bmatrix} \qquad \boldsymbol{b}_\kappa = \begin{bmatrix} -\boldsymbol{\kappa}_l \\ \boldsymbol{\kappa}_u \end{bmatrix} \tag{5.25}$$

are obtained. Let $\boldsymbol{d}_{\kappa;l}$ and $\boldsymbol{d}_{\kappa;l}$ be the desired lower and upper bounds on the B-spline coefficient $\boldsymbol{d}_\kappa$, then using (5.24) gives

$$\begin{aligned} -\boldsymbol{\Psi}_{NT}\; \boldsymbol{d}_{\kappa;l} &= -\boldsymbol{\kappa}_l \\ \boldsymbol{\Psi}_{NT}\; \boldsymbol{d}_{\kappa;u} &= \boldsymbol{\kappa}_u \end{aligned} \tag{5.26}$$

[24] inequality expressed with a "less than or equal to" relation.

from which an estimate[25] $\widehat{\boldsymbol{d}}_{\kappa;l}$ and $\widehat{\boldsymbol{d}}_{\kappa;u}$ of the lower and upper bound constraint on the B-spline coefficients respectively are estimated. Thus, the constraints bounds $\boldsymbol{\kappa}_l$, $\boldsymbol{\kappa}_u$ have been transformed to that on the respective B-spline coefficients $\boldsymbol{d}_{\kappa;l}$, $\boldsymbol{d}_{\kappa;u}$. The unknown B-spline coefficient correction vector $\Delta\boldsymbol{d}_\kappa$ is estimated from the constraint optimization approach described in the Section 4.5.2.

5.5 Partial derivatives of electron density

To solve for the correction to the unknown B-spline coefficients $\Delta\boldsymbol{d}_\kappa$ in (5.19), the matrices $\boldsymbol{A}_\kappa$ and $\boldsymbol{B}_\kappa$ are needed. Computation of the latter was shown in Eq. (5.25) and in this section, the former, i.e. the partial derivative of electron density $n_e(\varphi,\lambda,h)$ at a certain location $\{\varphi,\lambda,h\}$ with respect to the B-spline coefficients $d^{J_1,J_2}_{k_1,k_2:\kappa}$ for $\kappa \in \mathcal{K}_1$, will be shown. For better readability, the location indices will not be used in this section. More generally, we are seeking to compute the matrices[26] $\frac{\partial \boldsymbol{n}_e}{\partial \boldsymbol{d}_\kappa}\ \forall\ \kappa \in \mathcal{K}_1$. The computation of an element with the row index corresponding to the observation $\{\varphi,\lambda\}$ as well as the column index $(k_1\cdot k_2)+1$, of the sub-matrices $\boldsymbol{A}_{\kappa_1}$ to $\boldsymbol{A}_{\kappa_5}$, will be shown in this section. Starting with the F_2 layer, the following three partial derivatives: $\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:N^{F_2}_m}}$, $\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:h^{F_2}_m}}$ and $\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:H^{F_2}}}$ are computed as follows:

$$\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:N^{F_2}_m}} = \frac{\partial n_e}{\partial N^{F_2}_m}\ \frac{\partial N^{F_2}_m}{\partial d^{J_1,J_2}_{k_1,k_2:N^{F_2}_m}} \tag{5.27}$$

which, using Eqs. (5.2) and (5.10), is further written as

$$\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:N^{F_2}_m}} = \exp\left(\frac{1}{2}\left(1-\frac{h-h^{F_2}_m}{H^{F_2}}-\exp\left(-\frac{h-h^{F_2}_m}{H^{F_2}}\right)\right)\right)\ \Psi_{NT;k_1,k_2}. \tag{5.28}$$

where $\Psi_{NT;k_1,k_2} = N^{d_N}_{k_1}(\varphi)\ T^{d_T}_{k_2}(\lambda)$ denotes an element, corresponding to the column index $(k_1\cdot k_2)+1$ and the row corresponding to the observation location $\{\varphi,\lambda\}$, of the $N_{obs}\times(K_{J_1}\cdot K_2)$ dimensional tensor product matrix $\boldsymbol{\Psi}_{NT}$, with N_{obs} being the number of observations at the given time epoch. Substituting

$$z = \frac{h-h^{F_2}_m}{H^{F_2}}, \tag{5.29}$$

and applying the chain rule of calculus, the partial derivative of electron density with respect to $h^{F_2}_m$ B-spline coefficients is computed as follows:

$$\begin{aligned}\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:h^{F_2}_m}} &= \frac{\partial n_e}{\partial z}\ \frac{\partial z}{\partial h^{F_2}_m}\ \frac{\partial h^{F_2}_m}{\partial d^{J_1,J_2}_{k_1,k_2:h^{F_2}_m}}\\ &= N^{F_2}_m\ \exp\left(\frac{1}{2}\left(1-z-e^{-z}\right)\right)\frac{(1-e^{-z})}{H^{F_2}}\ \Psi_{NT;k_1,k_2}.\\ &+\begin{cases}\frac{-N^P_0}{H^P}\exp\left(\frac{-(h-h^{F_2}_m)}{H^P}\right)\ \Psi_{NT;k_1,k_2}, & \text{if } h\geq h^{F_2}_m\\ \frac{N^P_0}{H^P}\exp\left(\frac{-(h-h^{F_2}_m)}{H^P}\right)\ \Psi_{NT;k_1,k_2}, & \text{otherwise.}\end{cases}\end{aligned} \tag{5.30}$$

[25] In this case, a least squares estimate is used.

[26] The derivative of a vector valued electron density $\boldsymbol{n}_e$ at a location (φ,λ) along the altitudes $(h_{min}-h_{max})$ with respect to vector valued B-spline coefficients $\boldsymbol{d}_\kappa$ yields the matrices $\boldsymbol{A}_{\kappa_1}$-$\boldsymbol{A}_{\kappa_5}$.

Using Eqs. (5.12), (5.29) the partial derivative of electron density with respect to H^{F_2} B-spline coefficients is evaluted as follows:

$$\begin{aligned}\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:H^{F_2}}} &= \frac{\partial n_e}{\partial z}\ \frac{\partial z}{\partial H^{F_2}}\ \frac{\partial H^{F_2}}{\partial d^{J_1,J_2}_{k_1,k_2:H^{F_2}}} \\ &= N_m^{F_2}\ \exp\left(\frac{1}{2}\left(1-z-e^{-z}\right)\right)\frac{\left(1-e^{-z}\right)}{\left(H^{F_2}\right)^2}\ \ \Psi_{NT;k_1,k_2}\end{aligned} \quad . \tag{5.31}$$

Using Eq. (5.3), the corresponding partial derivatives for plasmaspheric extension,

$$\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:N_0^P}} = \begin{cases} -\exp\left(\frac{-(h-h_m^{F_2})}{H^P}\right)\ \Psi_{NT;k_1,k_2}, & \text{if } h \geq h_m^{F_2} \\ \exp\left(\frac{-(h-h_m^{F_2})}{H^P}\right)\ \Psi_{NT;k_1,k_2}, & \text{otherwise} \end{cases} \tag{5.32}$$

$$\frac{\partial n_e}{\partial d^{J_1,J_2}_{k_1,k_2:H^P}} = N_0^P \begin{cases} -\exp\left(\frac{-(h-h_m^{F_2})}{H^P}\right)\frac{h-h_m^{F_2}}{(H^P)^2}\ \ \Psi_{NT;k_1,k_2}, & \text{if } h \geq h_m^{F_2} \\ \exp\left(\frac{-(h-h_m^{F_2})}{H^P}\right)\frac{h-h_m^{F_2}}{(H^P)^2}\ \ \Psi_{NT;k_1,k_2}, & \text{otherwise} \end{cases} \tag{5.33}$$

are additionally evaluated.

5.6 Solution of GMM

In the previous sections, the global 4D EDM is posed as an inequality constrained optimization problem (ICOP) with the B-spline coefficients as the unknown optimization parameters. The scope of this section will be to show the exact analogy of 4D EDM to the solution of ICOP using KKT equations (4.94), (4.95) and (4.96), described in Chapter 4. Section 4.4.2 already described the augmentation of inequality constraints into a Lagrangian function, thereby transforming an ICOP to a partial ICOP; see Eq. (4.65). Associated with each constraint on a key parameter κ are the additional unknown pair of Lagrange multiplier and slack variable. Cumulatively for the complete set of global constraints, these are denoted as the vectors $\boldsymbol{\lambda}$ and $\boldsymbol{s}$ respectively. As described in Chapter 4, the inequality constraints on KPs are transferred to an equivalent positivity constraint on the corresponding slack variables (see Eq. (4.79) in Section 4.4.2). Furthermore, it was also shown that using the squared slack variable transformation (4.66), eliminates that positivity constraint as well. The only remaining inequality constraint is on the Lagrange multiplier due to the KKT condition (4.72c); see Section 4.5.1. The unknown parameters of optimization problem (or the optimization parameters) are then estimated by solving the KKT equations, as described in Sections 4.5.2 and 4.5.3 in Eqs. (4.94)-(4.98), which are themselves not repeated again, but the necessary matrices $\boldsymbol{A}_\kappa$ and $\boldsymbol{B}_\kappa$ to solve the KKT equations were derived in Sections 5.3 and 5.4 respectively. The IRI model is used for the initialization of the B-spline coefficients $\boldsymbol{d}_{\kappa;0}$ for the chosen key parameters $\kappa \in \mathcal{K}_1$. It shall be noted that IRI does not provide the B-spline coefficients directly, instead these are estimated in 2 steps:

- Obtain the Chapman key parameters $\kappa \in \mathcal{K}_1$ from the IRI model and denoted as κ_{IRI}
- Initial B-spline coefficients $\boldsymbol{d}_{\kappa;0}$ are then estimated from Eq. (5.5) using κ_{IRI}.

This is an important step because, at the start of iteration in KKT solution ($it = 0$), the initial values $\boldsymbol{d}_{\kappa;0}$ are required to linearize the observation model; see Eq. (5.19) to estimate the corrections $\Delta\boldsymbol{d}_\kappa$. More generally, following the convention of Eq. (4.8), the observations $\boldsymbol{y}$ used for the estimation procedure are combined together with their respective given weight matrices $\boldsymbol{P}$ and the unknown variance factor σ_0^2. Thus far, the GMM (5.19) with both deterministic as well as stochastic models are derived and the equality, inequality constraints are also imposed.

The next processing step was to solve for the unknown model parameters of GMM by first forming, the objective and subsequently the Lagrangian function, analogous to Eq. (4.65), using the initial

B-spline coefficients $\boldsymbol{d}_{\kappa;0}$, initial variance factor σ_0^2 and the inequality constraints (5.17), (5.18). Applying KKT conditions (4.71) - (4.72d), the optimality system of equations are obtained and are then solved using Newton's method (see Sections 4.3.3, 4.5.3). This requires initial values for the Lagrange multiplier $\boldsymbol{\lambda}_0$, the slack variable $\boldsymbol{s}_0$ vectors and the step size matrix $\boldsymbol{T}_{it}$ from Eq. (4.109) at an iteration it to be used in Newton's method; see Fig. 4.25.

One of the challenges is in the computational complexity with iterating until an estimate of the step size is obtained. Specifically, this is quantified by the convergence of the iterative procedure, achieved when a given pair of step size $\boldsymbol{T}_{it}$ and descent direction $\Delta\boldsymbol{d}_\kappa$; see Eq. (4.29), satisfy the following two necessary conditions (hereafter referred to as NC_{opt}) (see e.g. Boyd and Vandenberghe (2004), Nocedal and Wright (2006)):

1. the updated B-spline coefficients
$$\widehat{\boldsymbol{d}}_\kappa = \boldsymbol{d}_{\kappa;0} + \Delta\boldsymbol{d}_\kappa \tag{5.34}$$
are within the feasible region $\mathcal{F}_R$ defined by constraints (see definition 4.2.2 in Chapter 4)
2. the updated Lagrange multiplier and slack variable vectors
$$\widehat{\boldsymbol{\lambda}} = \boldsymbol{\lambda}_0 + \Delta\boldsymbol{\lambda} \quad \text{and} \quad \widehat{\boldsymbol{s}} = \boldsymbol{s}_0 + \Delta\boldsymbol{s} \tag{5.35}$$
satisfy the KKT complimentary slack conditions; see Eqs. (4.72c), (4.72d).

Together these two conditions, i.e. NC_{opt} are checked to see if the estimated key parameters are within feasible region, in which case, the corresponding slack variables shall all be positive, i.e. $\widehat{\boldsymbol{s}} \geq \boldsymbol{0}$ and thus there would no constraint violations (more on this topic will be described in Chapter 6; Numerical evaluations). In some cases (as will be shown in Section 6.3, Appendix C.7 and Appendix C.8 during numerical evaluations E1 - E12), it has been observed that only one of the above two conditions NC_{opt} are met. Upon detailed investigation, those were found to be the specific case of only the need for improvement in the estimated Lagrange multipliers or the slack variables. Therefore, it was sufficient when the last computed values $\{\boldsymbol{\lambda},\ \boldsymbol{s}\}$ are recursively passed on to the step where optimality conditions (Section 4.5) are applied and KKT equations are obtained.

Hence, an iterative procedure is needed (and has been developed) until both the conditions in NC_{opt} are satisfied. The relative complexity between an itertaive procedure to compute a step size matrix as well as the variance components depend on the number of constraints and the convexity of the optimization problem. While the former is discussed in further detail in Chapter 6, the latter is dependant on the quality of observations; see Eq. (4.28) and Fig. 4.4. If both the conditions in NC_{opt} are satisfied, then the most recently estimated variance components $\widehat{\sigma}_0^2$ (based on VCE procedure, as described in Section 4.1.2) are passed recursively to the Lagrangian minimization procedure until convergence[27]. It has already been explained in Section 4.3.1 that due to the presence of inequality constraints, the convergence criteria does not apply if the condition (4.34) is violated[28].

For the scope of this chapter, the discussion were limited to estimated B-spline coefficient correction vector $\Delta\boldsymbol{d}_\kappa$, although more generally the correction to unknown parameters (see e.g. $\Delta\boldsymbol{\beta}$ in Eq. (4.107)) also additionally contains the corrections to Lagrange multipliers, slack variables as well as differential code biases (DCBs)[29]. To keep the discussion simpler and to the context of this chapter, $\Delta\boldsymbol{d}_\kappa$ shall be the desired descent direction vector.

The final estimated set of B-spline coefficients in Eq. (5.34), variance factor vector, the optimum step size matrix are further used in a Monte-Carlo procedure to obtain the standard deviations (or

[27] The convergence criteria is an important design criteria and has a significant impact on the accuracy of the modelled key parameters. For example a threshold of RMSE $\widehat{\boldsymbol{\beta}}_{it} = 10^{-18}$ for simulated observation dataset, for use in Eq. (4.7), was used for this work. For other datasets a value between 10^{-8} to 10^{-12} was used. These datasets are described in detail in Chapter 6, Table 13 and in Fig. 16.

[28] In fact this aspect has been identified as one of the challenges to overcome in possibly future work when real-time EDM shall be considered.

[29] DCBs are relevant only when GNSS STEC are considered in observations.

the quality estimate) for the Chapman key parameters. Examples of standard deviation maps will be shown in the discussion of numerical evaluations. The final estimated B-spline coefficients in Eq. (5.34) are used to calculate the Chapman key parameters globally as well as at selected ionosonde station locations for validation. An additional validation is performed by evaluating the electron density using the estimated key parameters and compared with along track in-situ observations of Swarm A, B, C satellites. These aspects will be discussed in detail in Chapter 7 during the validation of EDM.

5.7 Summary of global electron density modelling

We have reached the end of this chapter where the concept of inequality constrained optimization have been applied to EDM. Beginning with the stratification of the ionosphere into the D, E, F_1, F_2 layers and the plasmapshere, the electron density of each layer is parametrized by the Chapman key parameters, namely, (1) the peak density, (2) the peak height and (3) the scale height. Plasmaspheric electron density is parametrized using its basis density and the scale height. While the electron density of the former four layers are modelled using the Chapman profile function (5.2), that of the plasmasphere is modelled using an exponential decay function (5.3). More generally, the Chapman key parameters of each layer are contained in the set $\mathcal{K}$, from which a subset $\mathcal{K}_1$ was used for a demonstration of EDM. Each of the key parameters $\kappa \in \mathcal{K}_1$ are modelled using 2D B-splines spatially, thereby the corresponding B-spline coefficient correction about initial values become the unknown parameters of the GMM (5.19). The upper and lower bounds κ_u, κ_l respectively, on the the key parameters $\kappa \in \mathcal{K}_1$ lead to a system of linear inequalities which are accumulated in the LMI (5.17), such that together with the GMM, the 4D EDM is posed as a constrained optimization problem of the form (4.25). Specifically, the design matrix $\boldsymbol{A}_\kappa$ and the constraint coefficient matrix $\boldsymbol{B}_\kappa$ used for the estimation of the unknown B-spline coefficients were derived in this chapter. A detailed description of the transformation between constraints applied to Chapman key parameters and those equivalently transformed to B-spline coefficients was also provided. Finally, an analogy between the theoretical solution of constrained optimization problem (as described in Chapter 4) and that of EDM problem were described in Section 5.6.

6 Numerical evaluations

In the Chapters 4 and 5, the optimization algorithm and its applicability to the electron density modelling have been described respectively. In this chapter, numerical evaluations will be performed with this electron density model to analyse $N_m^{F_2}$, $h_m^{F_2}$ and the corresponding slack variables.

The following general conventions apply to this chapter:

- The term "F3C observations" refers to the electron density values derived from Formosat-3/COSMIC (F3C) occultation measurements (see Section 3.2.1).
- The term "GRACE observations" refers to the electron density values derived from the occultation measurements of the GRACE A/B satellites (see Section 3.2).
- The term "GNSS observations" refers to the STEC computed from GNSS pseudorange and carrier phase measurements (see Section D.1).
- The term "VTEC observations" refers to the electron density values computed using the separability approach (see Section 7.1).
- The term "model parameter" refers to the B-spline coefficient vector of the key parameters of the set $\mathcal{K}_1$ (see Section 4.2).
- The term "optimization parameter" refers to the composite vector of the unknown B-spline coefficients, the Lagrange multiplier and the slack variables (see Section 4.4.2).

This chapter focuses on two important output parameters, namely the slack variable and the deviations of a key parameter from the reference values. The slack variable provides information about the "qualification" of a given constraint (see Chapter 4, definition 4.4.1), whereas the deviation of an estimated key parameter $\widehat{\kappa}$ is defined as $\Delta\kappa = \widehat{\kappa} - \kappa_{ref}$ and that of the corresponding estimated B-spline coefficient vector as $\Delta\boldsymbol{d}_\kappa = \widehat{\boldsymbol{d}_\kappa} - \boldsymbol{d}_{\kappa_{ref}}$, where $\widehat{\kappa}$, $\widehat{\boldsymbol{d}_\kappa}$ are the estimated, κ_{ref} and $\boldsymbol{d}_{\kappa_{ref}}$ are the reference values of the key parameter and the estimated B-spline coefficient vector respectively.

Three observation datasets are used in this chapter, namely (1) DS1a, (2) DS2d and DS2a and (3) DS3. For evaluations using dataset DS1a, the deviation is computed with a known reference (κ_{ref}). The known reference itself is computed as described in Section 7.1. For the evaluations using the dataset DS2, the Chapman key parameters from an independent reference, e.g. dataset DS3, will be used for computing the deviation. The descriptions of the datasets are provided in Table 13 and in Fig. 16.

For the scope of this chapter, the following nine inputs (or the so-called "*independent variables*") are identified:

1. Number of ionosphere layers to be modelled
2. Inequality constraint bounds
3. Inequality constraint resolution[1]
4. B-spline levels of the key parameters
5. Total number of used observation techniques
6. Number of observations per technique
7. Initial conditions
8. Optimization algorithm specific parameters
9. Space weather conditions (year and month)

[1] In particular, spatial resolution is meant here. During the early investigations, constraint temporal resolution was not found to be a dominant factor. A constraint temporal resolution of 6 hours is used in this work.

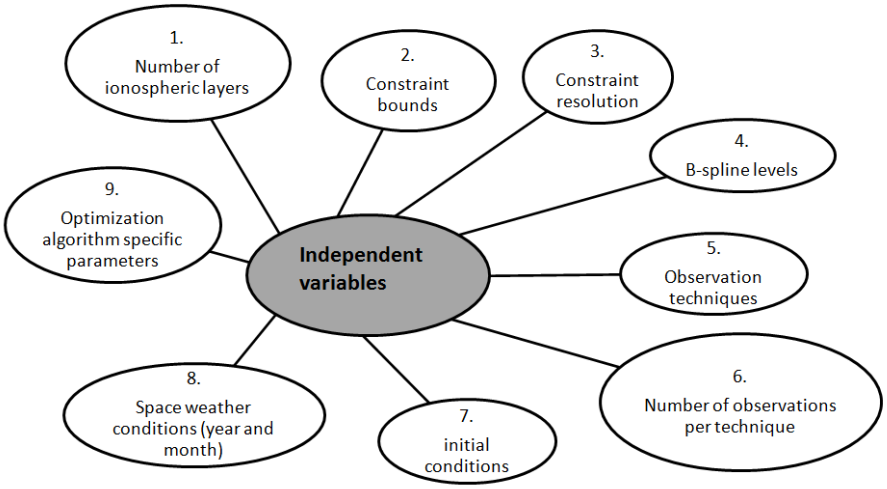


Figure 6.1: Independent variables in electron density modelling

These variables are shown in Fig. 6.1. In the next section we will describe and quantify each of the above variables, which could be varied, while keeping the other parameters fixed to evaluate the electron density model[2]. It shall be noted that the independent variables are closely linked to one of the three pillars illustrated in Fig. 4.1.

6.1 Independent variables

In the following, the independent variables are described and their nominal values quantified for the numerical evaluations.

Number of ionosphere layers

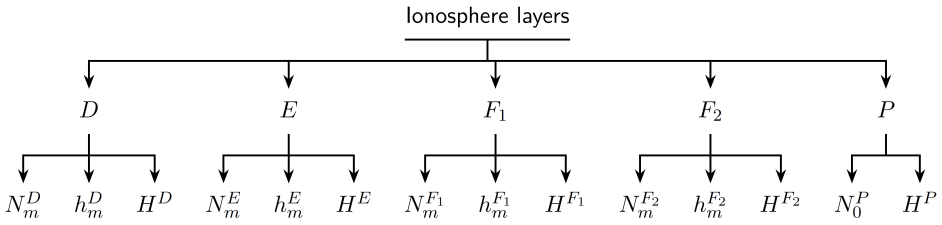


Figure 6.2: ionosphere layers and the Chapman key parameters

The outcome of an evaluation by varying the key parameters in $\mathcal{K}_1$ and $\mathcal{K}_2$, signifies of each of the individual key parameters. It shall be noted that the overall accuracy of the modelled key parameter depends on the quality of given values in the set $\mathcal{K}_2$ as well. Although among the 14 key parameters in $\mathcal{K}$, several permutations are possible, in the evaluations E1 to E8 the ionosphere layers D, E, F_1 and F_2 as well as the plasmasphere are considered. However, obtaining the given values for the scale height of the individual layers is a challenge because it is not a direct output of any empirical or physical ionosphere model. Indirect methods to model and charecterise the F_2 layer scale height can be found in Belehaki et al. (2006), Stankov. and Jakowski (2006), Altadill et al. (2013), Lei and Chuo (2014). The effect of varying the F_1- layer peak height and peak density at selected locations has been analysed in Liang (2017) and in this work the focus will be on varying the inequality constraint bounds imposed on the multi-Chapman key parameters.

[2]The number of permutations by varying each of the input variables is enormous and beyond the scope of this work.

As explained in Chapter 5, the key parameters of the ionosphere layers (see Fig. 6.2) are either estimated as unknown B-spline coefficients or treated as known values, e.g. given from a model. The first independent variable is the number of ionosphere layers considered in our modelling problem which is related to the total number of key parameters in the sets $\mathcal{K}_1$ and $\mathcal{K}_2$ together. Increasing the number of key parameters in $\mathcal{K}_1$ directly increases the number of B-spline coefficients, and consequently the total number of unknown parameters, to be estimated. Therefore, only a minimum number of key parameters shall be considered in $\mathcal{K}_1$ and reliable sources[3], e.g. the IRI model (Bilitza (2000)) or thermosphere-ionosphere coupled models (Richmond et al. (1992)), shall be used for values used in $\mathcal{K}_2$. In numerical evaluations, we select the 5 key parameters

$$\mathcal{K}_1 = \left\{\kappa_1 = N_m^{F_2},\quad \kappa_2 = h_m^{F_2},\quad \kappa_3 = H^{F_2},\quad \kappa_4 = N_0^{P},\quad \kappa_5 = H^{P}\right\} \tag{6.1}$$

to be modelled and the given values

$$\begin{aligned}\mathcal{K}_2 = \{&\kappa_6 = N_m^{F_1},\quad \kappa_7 = h_m^{F_1},\quad \kappa_8 = H^{F_1},\quad \kappa_9 = N_m^{E},\quad \kappa_{10} = h_m^{E},\quad \kappa_{11} = H^{E},\\ &\kappa_{12} = N_m^{D},\quad \kappa_{13} = h_m^{D},\quad \kappa_{14} = H^{D}\}\end{aligned} \tag{6.2}$$

obtained from IRI. As exception, the scale heights of the F_1-, $E-$ and the $D-$ layer are obtained

$$\kappa_8 = H^{F_1} = \frac{H^{F_2}}{10} \quad ; \quad \kappa_{11} = H^{E} = \frac{H^{F_2}}{10} \quad ; \quad \kappa_{14} = H^{D} = \frac{H^{F_2}}{10} \tag{6.3}$$

(see Limberger (2015)) assuming that the H^{F_2} are non-zero.

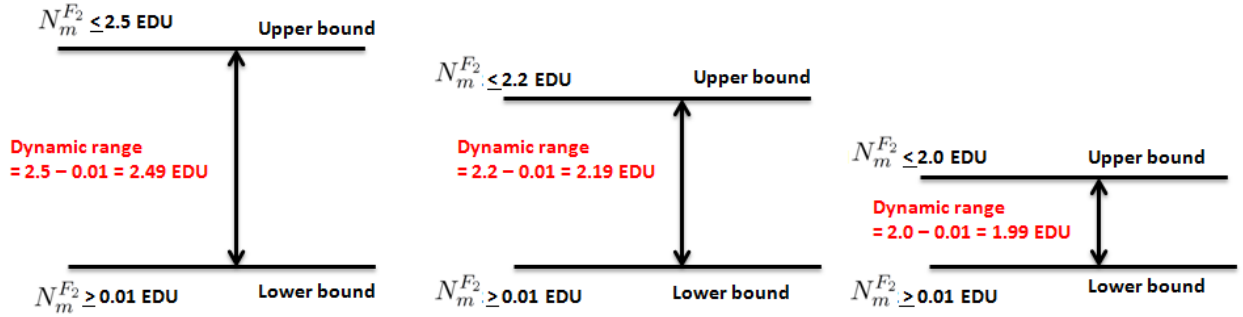


Figure 6.3: Illustration of the varying the upper bound constraint and the resulting dynamic range.

Constraint bounds

Constraint upper and lower bounds are the second set of independent variables. One of the main novelties of this work is the estimation B-spline coefficients subject to inequality constraints. Therefore,

> of particular interest in the numerical evaluations is to analyse the impact of varying the inequality constraint bounds on global 4D electron density modelling.

Specific aspects related to the analysis of the slack variables and constraint violations have received less attention in the literature and require extensive numerical evaluations. In contrast to a deterministic effect on the total electron density by varying the number of ionosphere layers, the effect of constraint bounds is not known apriori. More specifically, it is not known as to which of the constraints will be "active", "inactive" or "violated". Therefore, of particular interest in

[3] For the peak density and peak height of the F_1, E and D layers, approximate values are available from the ionosphere models, e.g. IRI, NeQuick.

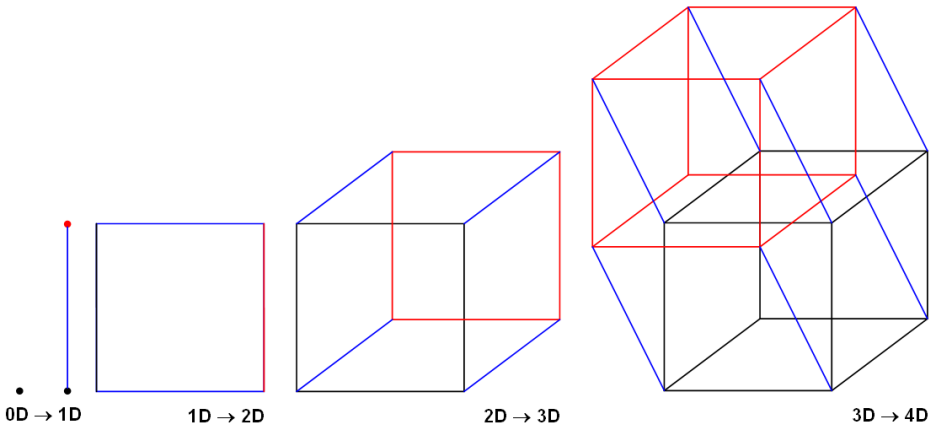


Figure 6.4: Hypercuboid topology from 1d to 4d. Source Wikipedia. Examples of hypercuboid in 0, 1, 2 and 3 dimensions are point, line, rectangle and cuboid respectively.

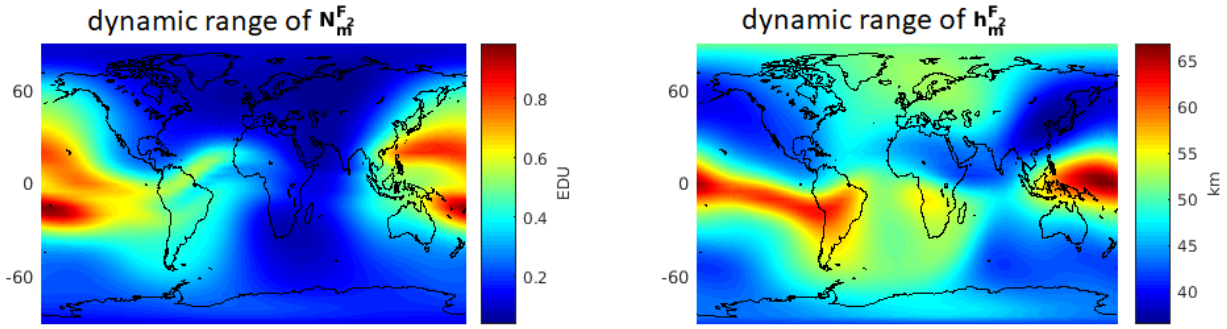


Figure 6.5: Dynamic range of F_2 layer peak density and peak height. As will be described in Appendix A the difference between the upper and the lower bound of a given constraint at a given location is shown in this figure. It means, for a given step size matrix in the ICOA, a larger value e.g. 0.8 EDU for $N_m^{F_2}$ and 65 km for $h_m^{F_2}$ denote that at the corresponding locations, the inequality constraints are imposed with a larger feasible region and thus allowing more number of candidate solutions to be considered for the Lagrangian minization.

this work is the analysis of the effect of the different inequality constraint bounds, which to the best knowledge of the author has not been studied yet. However, the optimization approach for inequality constrained problems in geodesy has been used for geodetic network adjustment (see Koch (1985), Koch (1988)), variance component estimation (Koch (1999)) and zenith tropospheric delay (Roese-Koerner (2015)).

Specifically, the constraint bounds for $N_m^{F_2}$ and $h_m^{F_2}$ will be varied in the sequel and this aspect will be discussed in the Section 6.2. An example of upper and lower constraint bounds on $h_m^{F_2}$ is shown in Fig. 6.6.

Constraint resolution

Constraints are applied on the Chapman key parameters globally at a specific spatial resolution and transformed to the corresponding B-spline coefficients, see Section 5.4; Eq. (5.26). Table 6.1 shows the number of Lagrange multipliers and slack variables to be estimated for the three different spatial resolutions of the constraint. Ideally, the resolution is paired with the B-spline levels of the individual key parameters (see Chapter 5, Eqs. (5.10), (5.11) and (5.12)). In this work, a 5° resolution along both the latitude and the longitude is used for the upper and lower constraint bounds.

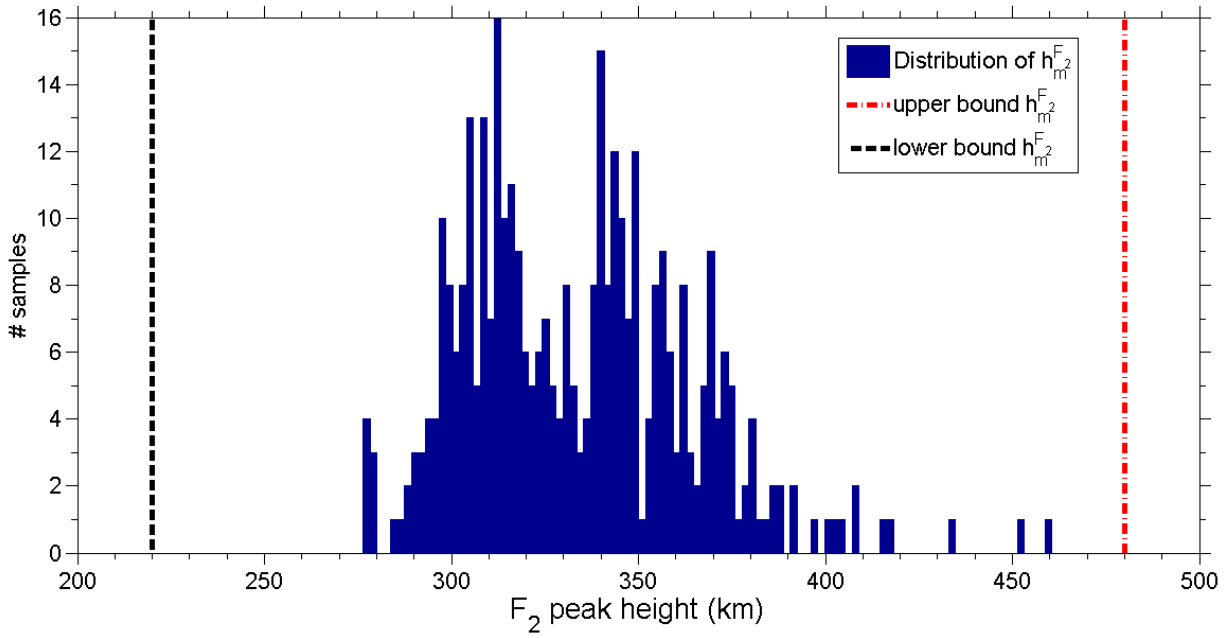


Figure 6.6: Global $h_m^{F_2}$ distribution with data from IRI model over the year 2014 along with the example constraint bounds. This is just an example to show the approximate bounds and the typical values of $h_m^{F_2}$ to be expected.

Table 6.1: Spatial resolution of constraints and the impact on the parameterization of the optimization problem.

Spatial resolution of constraints latitude × longitude(ř)	Lagrange multiplier and slack variables (#)
2.5 × 2.5	21608
2.5 × 5	10804
5 × 5	5702

B-spline levels

As mentioned in Section 5.1, the levels J_1 and J_2 define the total number $K_{J_1} = 2^{J_1} + 2$ and $K_{J_2} = 3 \cdot 2^{J_2}$ of polynomial and trigonometric B-spline functions as well as the shift parameters $k_1 = 0, 1, \cdots, K_{J_1} - 1$ and $k_2 = 0, 1, \cdots, K_{J_2} - 1$, their position in the latitude-longitude space respectively (see e.g. Schmidt (2011), Schmidt et al. (2015)). Appropriate values for the levels J_1 and J_2 are chosen from the average sampling intervals of the observations with respect to latitude and longitude (Schmidt et al. (2015), Goss et al. (2019)). The higher the numerical values for the levels, the more spline functions are included in Eq. (5.5) and thus finer are the signal structures which can be modelled (Zeilhofer (2008), Limberger (2015) and Liang (2017)). In the scope of this chapter, the level values $J_1 = 4$ and $J_2 = 3$) are used for all key parameters in the set $\mathcal{K}_1$. This convention for defining the B-spline levels was already described in Chapter 5.

Observation techniques

Space and ground based observation techniques (see Fig. 6.7), as introduced in Chapter 3, are used in this work for electron density modelling as well as for validation (to be later described in Chapter 7).

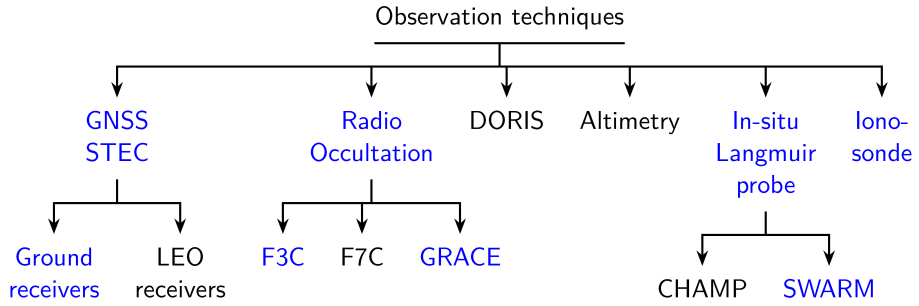


Figure 6.7: Observation techniques relevant for EDM. Those in blue font are the ones directly used within this work and the remaining ones in black font are not used. DORIS and altimetery techniques are indirectly used in this work due to the use of DGFI-TUM VTEC map which is generated already considerignt hese two techniques.

As already mentioned in Chapter 3, DORIS and altimetry techniques are used to generate VTEC global ionosphere maps (GIM)[4] (see Erdogan et al. (2017), Erdogan et al. (2020)), which in-turn is used to obtain the electron density observations in the separability approach (see Section 7.1).

For the numerical evaluations, three datasets are considered:

- Dataset-1 (DS1a): Simulated electron density for which the underlying Chapman key parameters are known. Procedure for the simulated data generation including realistic measurement noise[5] will be described in Chapter 7.
- Dataset-2 (DS2): A combination of GNSS STEC and electron density from F3C IRO, GRACE IRO, seperability approach for the time period between 08 - 14 March 2015. [6]
- Dataset-3[7] (DS3): Electron density, $N_m^{F_2}$ and $h_m^{F_2}$ from the ionosonde for the time period between 08 - 14 March 2015.

Number of observations per technique

The datasets can be configured to contain a different number of observations from each of the three techniques (GNSS STEC, IRO and VTEC separability approach). As described in Chapter 5, in the case of VTEC observations (from the separability approach), an altitude resolution δh = 20 km is used between 80 and 220 km, δh =5 km between 220 and 480 km and δh = 20 km between 480 and 1000 km, thus leading to a total of 89 discrete altitudes with a denser sampling in the F_2 layer compared to the others.

Initial condition or initial value

As mentioned in the Chapters 4 and 5, the choice of the initial values is an important consideration for electron density modelling using the optimization approach. The following options are investigated to obtain the approximate initial B-spline coefficients for the key parameters $\kappa \in \mathcal{K}_1$:

- Option 1 (Opt1): Use the Chapman key parameters from the IRI model to obtain the initial values for the B-spline coefficients.
- Option 2 (Opt2): Use the estimated B-spline coefficients from the previous (most recent) time epoch.

[4] GIM is one of the standard ionospheric VTEC data product from IGS analysis centers (see Feltens (2007) and Schaer (1999)).

[5] Dataset DS1a used for the numerical evaluations E9 - E12 meant simulated observations without noise.

[6] Details on GNSS receivers, satellites, pre-processing and the separability approach for computing the electron density from VTEC are described in Chapter 3.

[7] This dataset is only used from selected ionosonde stations for validation of the numerical evaluations. It will be used more comprehensively for independent validation in Chapter 7.

- Option 3 (Opt3): Use the B-spline coefficients (or key parameters) predicted[8] to the current time epoch.

There are two disadvantages of using Opt1: (1) The prior information from the most recently estimated B-spline coefficients are ignored and (2) the empirical ionospheric models (e.g. IRI, NeQuick) do not provide the scale height directly. Therefore, in this work a combination of Opt2 and Opt3 is preferred.

Optimization tuning parameters

There are three tuning parameters, namely, the "complimentarity measure" (μ_{com}); see Eq. (4.110), the "centering parameter" (σ_{cen}) and the "step-size", as introduced in the Section 4.5.3. Although included in the list of independent variables, the tuning parameters are not externally driven in the numerical evaluations for the two main reasons: (1) μ_{com} and σ_{cen} are computed using intermediate outputs (see Eq. (4.110) and Section 4.5.3) and hence are not completely independent and (2) as explained in Section 4.5.3, these two parameters are only responsible for the intermediate solution path (see Mehrotra (1992), Tütüncü et al. (2003), and Nocedal and Wright (2006)). The intermediate solution path of our optimization algorithm or the tuning parameters are not the main focus of this work[9] and therefore this aspect will not be discussed further.

The third tuning parameter is the step-size t_{it}; see Section 4.7.3 used within the gradient descent or Newton's method. Strictly speaking, the step-size is indeed an independent parameter in the optimization algorithms (see e.g. Gill et al. (1984), Jacobs (1988), Wilson and Martinez (2001), Boyd and Vandenberghe (2004), Nocedal and Wright (2006)) but[10] in this work it is computed based on the method described in the Section 4.7.3, Eq. (4.158). An example of the effect of a constant step-size on the optimal solution, with the gradient descent algorithm, is shown in Fig. 4.8.

Space weather condition

The last among the nine independent variables is the phase of the solar cycle (or in the time epoch of processing) chosen for electron density modelling[11]. Although a 11-year cycle is generally attributed for the sun-spots and their effect on the ionosphere (Hernández-Pajares et al. (1998), Aragon-Angel et al. (2016)), there are also short term variations in the order of days or even hours (see Jakowski et al. (1999), Förster and Jakowski (2000), Borries et al. (2009)). Therefore, a combination of the different periodic effects are included in the observations. Furthermore, the polar, mid and low latitudes are affected differently by the space weather events (see e.g. Hargreaves (1992), Jakowski (2017)).

These were the nine independent variables directly or indirectly required for electron density modelling, based on which the evaluations listed in Table 6.2 are performed.

In the next section, the focus will be on the analysis of the impact of constraint bounds on the estimated optimization parameters. Specifically, the following questions will be answered:

[8] Projected forward in time based on the last finite samples using an auto-regressive (AR) or an auto-regressive moving average (ARMA) model; see Section 4.7.4.

[9] The initial assessment performed during the development of our software prototype, the tuning parameters were identified and used throughout this work.

[10] Initial testing of the optimization algorithm involved less than 10 unknown parameters. However when extending to larger dimensional electron density modelling problems (more than 1000 unknown parameters), a more systematic way of step-size computation becomes a necessity.

[11] In order to avoid confusion, the time window of processing can be adapted to the exact obectives of any study, i.e the impact of solar maximum, minimum, etc. In case of simulated observations, the indices, e.g. F10.7, Dst etc. are used as proxy for quantifying the strength of a space weather event. However, when a particular events, e.g. the St. Patrick's day storm needs to be studied or analysed, the space weather conditions become fixed to the exact conditions observed at that particular time. It is only an independent variable when the performance of the optimization algorithm is to be tested on a variety of geomagnetic and space weather conditions.

Table 6.2: Outline of the numerical evaluations. The datasets are defined in Appendix C Table 13.

Evaluation identifier	independent parameter to be varied	observation dataset
E1	$N_m^{F_2}$ nominal lower and upper bounds	DS2d
E2	$N_m^{F_2}$ reduced upper bound compared to that in E1	DS2d
E3	$N_m^{F_2}$ increased lower bound compared to that in E1	DS2d
E4	$N_m^{F_2}$ nominal lower and upper bounds with different observation dataset from E1	DS2a
E5	$h_m^{F_2}$ nominal lower and upper bounds	DS2d
E6	$h_m^{F_2}$ reduced upper bound compared to that in E5	DS2d
E7	$h_m^{F_2}$ reduced lower bound compared to that in E5	DS2d
E8	$h_m^{F_2}$ nominal lower and upper bounds with different observation dataset than in E5	DS2a
E9 (E9a - E9r)	Fixed $N_m^{F_2}$ lower bound to 0.02 EDU and upper bound varied 1.8 - 3.2 EDU	DS1a
E10 (E10a - E10n)	Fixed $N_m^{F_2}$ upper bound to 2.8 EDU and lower bound varied 0.001 - 0.51 EDU	DS1a
E11 (E11a - E11d)	Fixed $h_m^{F_2}$ upper bound to 480 km and lower bound varied from 220 - 280 km	DS1a
E12 (E11a - E11d)	Fixed $h_m^{F_2}$ lower bound to 220 km and upper bound varied from 420 - 480 km	DS1a
(E1a - E1d)	Role of the bottom ionosphere layers	DS2d

- Are the optimization parameters sensitive to changes in the inequality constraints? If yes, how could they be quantified?
- Are the constraints active, inactive or violated corresponding to the changes in the corresponding inequality bounds?

6.2 Constraint quantification and violation

The system of inequality constraints

$$\boldsymbol{B}_{ineq}\boldsymbol{\beta} - \boldsymbol{b}_{ineq} \leq \mathbf{0}$$

as introduced in Eq. (4.74) with the upper and lower bounds of the key parameter will be used to describe the constraint quantification. There are three parameters in this inequality, namely (1) the inequality constraint bound vector $\boldsymbol{b}_{ineq}$, (2) the unknown parameter vector $\boldsymbol{\beta}$ and (3) the constraint coefficient matrix $\boldsymbol{B}_{ineq}$. In the context of this section, the numerical evaluations E1 to E12 (including 54 sub-evaluations) will be performed by varying the constraint bound vector[12] $\boldsymbol{b}_{ineq}$, while the other input parameters, described in Section 6.1 as well as inequality constraint coefficient matrix $\boldsymbol{B}_{ineq}$ will be kept constant. The only exception will be the evaluations E4 and E8, where a different observation dataset, compared to E1 and E5 will be used.

If a specific example of the F_2 layer peak density is considered, then for a given pair of lower and upper bounds $N_{m,l}^{F_2}$, $N_{m,u}^{F_2}$ at a point location $P(\{\varphi, \lambda\})$, the inequalities

$$N_{m,l}^{F_2}(\varphi, \lambda) \leq N_m^{F_2}(\varphi, \lambda) \leq N_{m,u}^{F_2}(\varphi, \lambda) \tag{6.4}$$

[12] Only the inequality constraint bounds are used in the numerical evaluations. The equality constraints are not changed.

hold. These are two separate inequalities

$$N_{m,l}^{F_2}(\varphi,\lambda) - N_m^{F_2}(\varphi,\lambda) \leq 0 \tag{6.5a}$$

$$N_{m,u}^{F_2}(\varphi,\lambda) - N_m^{F_2}(\varphi,\lambda) \geq 0, \tag{6.5b}$$

one each for the lower and the upper bound, and

$$-\left(N_{m,u}^{F_2}(\varphi,\lambda) - N_m^{F_2}(\varphi,\lambda)\right) \leq 0 \tag{6.6}$$

is obtained from (6.5), using the transformations introduced in Section 4.4.1, and expressed by Eq. (4.58).

> Constraint quantification refers to the process of defining numerical bounds, so that two or more constraints can be compared and their relative impact on the optimization problem be analysed.

In this work, a new[13] method for the constraint quantification will be introduced for the global 4D electron density modelling with the following two parameters, namely, (1) dynamic range (DR)[14] and (2) hypervolume[15]. Given the upper (κ_u) and lower (κ_l) bounds at $\{\varphi,\lambda\}$, the dynamic range is defined as

$$\mathcal{D}_\kappa^r(\varphi,\lambda) = \left|\kappa_u(\varphi,\lambda) - \kappa_l(\varphi,\lambda)\right|. \tag{6.7}$$

More generally, the quantification is extended to include global constraints at any $\{\varphi_m,\lambda_j\},\ \forall\ m = \{1,\ \cdots,\ \varphi_{max}\},\ j = \{1,\ \cdots,\ \lambda_{max}\}$, for a given set of key parameters $\kappa_p \in \mathcal{K}_1\ \ \forall\ p = \{1,\ \cdots,\ q\}$, where φ_{max} , λ_{max} denote the total numbers of latitude, longitude values respectively, resulting in a hypervolume[16]

$$\mathcal{V}(\varphi_m,\lambda_j) = \prod_{p=1}^{q} \breve{\mathcal{D}}_{\kappa_p}^r(\varphi_m,\lambda_j), \tag{6.8}$$

of the hypercuboid[17] formed by the edges corresponding to the normalized dynamic range

$$\breve{\mathcal{D}}_{\kappa_p}^r(\varphi_m,\lambda_j) = \frac{\mathcal{D}_{\kappa_p}^r(\varphi_m,\lambda_j)}{\mathcal{N}_{\kappa_p}}. \tag{6.9}$$

The normalization factor corresponding to the global maximum dynamic range

$$\mathcal{N}_\kappa = \max\left(\mathcal{D}_\kappa^r(\varphi_m,\lambda_j)\right) \quad \forall\ m = \{1,\ \cdots,\ \varphi_{max}\},\ j = \{1,\ \cdots,\ \lambda_{max}\}, \tag{6.10}$$

computed as a maximum of the differences between the upper (κ_u) and lower (κ_l) bounds for the parameter κ, where $L = \varphi_{max} \cdot \lambda_{max}$ is the total number of latitude-longitude pairs, at which the constraints are imposed. Furthermore, for each key parameter κ_p, an average dynamic range[18]

$$\bar{\mathcal{D}}_{\kappa_p}^r = \frac{1}{L}\sum_{m=1}^{\varphi_{max}}\sum_{j=1}^{\lambda_{max}} \breve{\mathcal{D}}_{\kappa_p}^r(\varphi_m,\lambda_j) \tag{6.11}$$

[13] To the best knowledge of author, inequality constraint quantification methods have not yet been used for global 4D electron density modelling.

[14] Use of the phrase "dynamic range" is from electrical enginnering where for example the dynamic range of an amplifer is specified as the difference between the largest and the lowest power levels in dB; see e.g. Smith (2007).

[15] The n-dimensional hypervolume was originally proposed by Hutschison (1957) as the space occupied by a set of n-independent axes and resulting in a geometrical shape.

[16] Extension of volume to multiple dimensions.

[17] For q dimensions, a hypercuboid has 2^q vertices, see Fig. 6.4.

[18] Averages can be computed similarly for the dynamic range (6.7) as well without normalization.

is the mean over the L global dynamic ranges of κ_p. Similarly,

$$\bar{\mathcal{V}} = \prod_{p=1}^{q} \bar{\mathcal{D}}^r_{\kappa_p} \tag{6.12}$$

gives the average hypervolume. Dynamic range has the same unit as the underlying key parameter itself and that of the hypervolume is key parameter units raised to the power q in exponent.

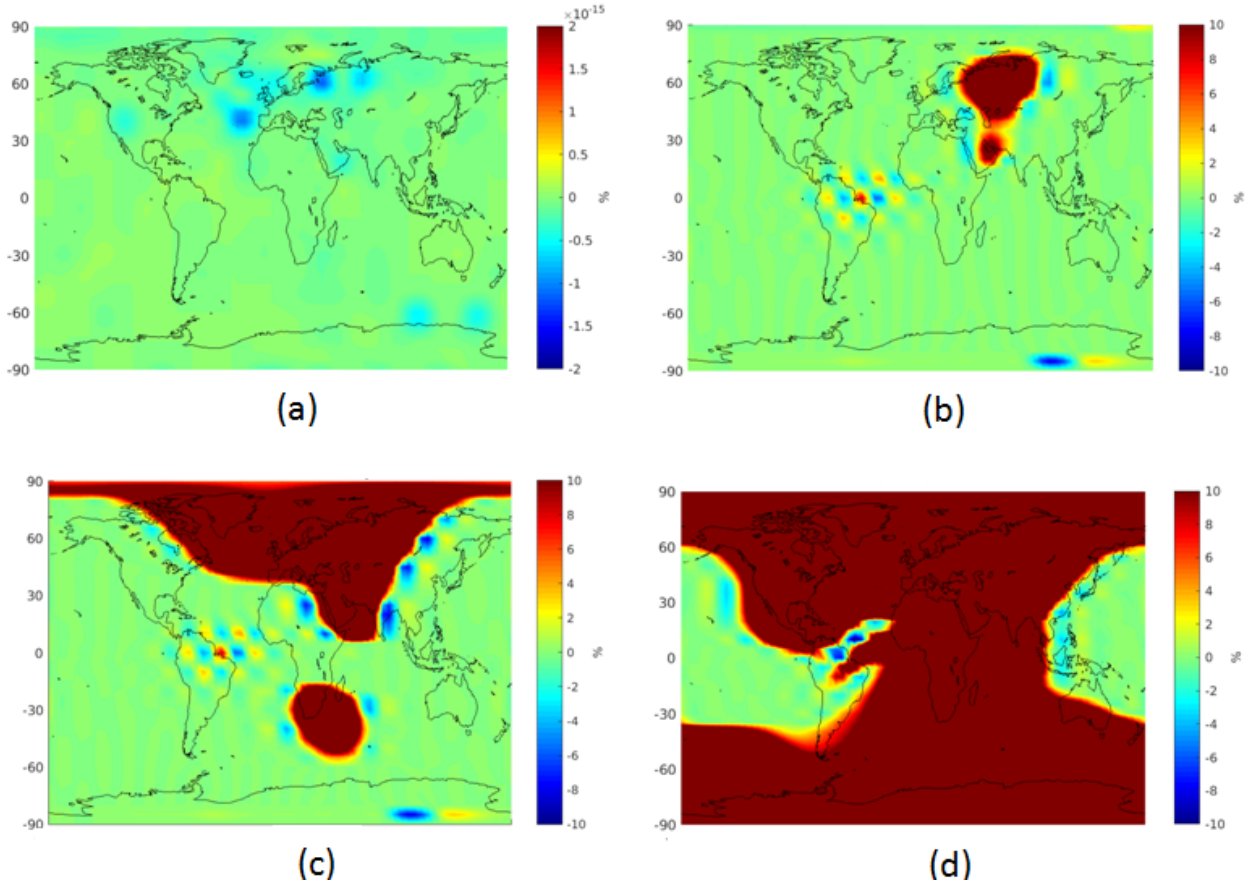


Figure 6.8: $N_m^{F_2}$ relative accuracy for four different constraint bounds. The dynamic range and slack are mentioned in the order (a) top-left: $N_m^{F_2} \geq 0.01$ EDU (b) top-right: $N_m^{F_2} \geq 0.11$ EDU (c) bottom-left: $N_m^{F_2} \geq 0.31$ EDU (d) bottom-right: $N_m^{F_2} \geq 0.51$ EDU. The dynamic range is reduced from 2.19 to 1.99 EDU and the electron density modelling problem becomes more critically constrained from (a) to (d).

The dynamic range (6.7) and the hypervolume (6.8) are both scalars, quantifying the size, and by extension, the number of "candidate solutions" (see Section 4.4.2) in the feasible region for a given step-size[19]. Figure 6.5 show the dynamic ranges of $N_m^{F_2}$ and $h_m^{F_2}$ used in the evaluation E1 corresponding to 12 March 2015 at 0 UT. It is shown in Figures 6.9 and 6.8 that the estimated optimization parameters are sensitive to changes in the dynamic range of the constraints.

To summarize, a constraint is quantified through the two parameters, namely the dynamic range (DR) and the hypervolume. While the former is defined for each pair of upper and lower bound of the key parameter $\kappa \in \mathcal{K}_1$, the latter is the product of the q individual normalized dynamic ranges of the key parameters in $\kappa \in \mathcal{K}_1$.

Constraint violation

As mentioned in Section 4.4.2, when it is not possible to determine an optimal solution within the

[19] Qualitatively, the normalized dynamic range $\breve{\mathcal{D}}^r_\kappa(\varphi_m, \lambda_j) \approx 1$ indicates a larger feasible region relative to $\breve{\mathcal{D}}^r_\kappa(\varphi_m, \lambda_j) \ll 1$.

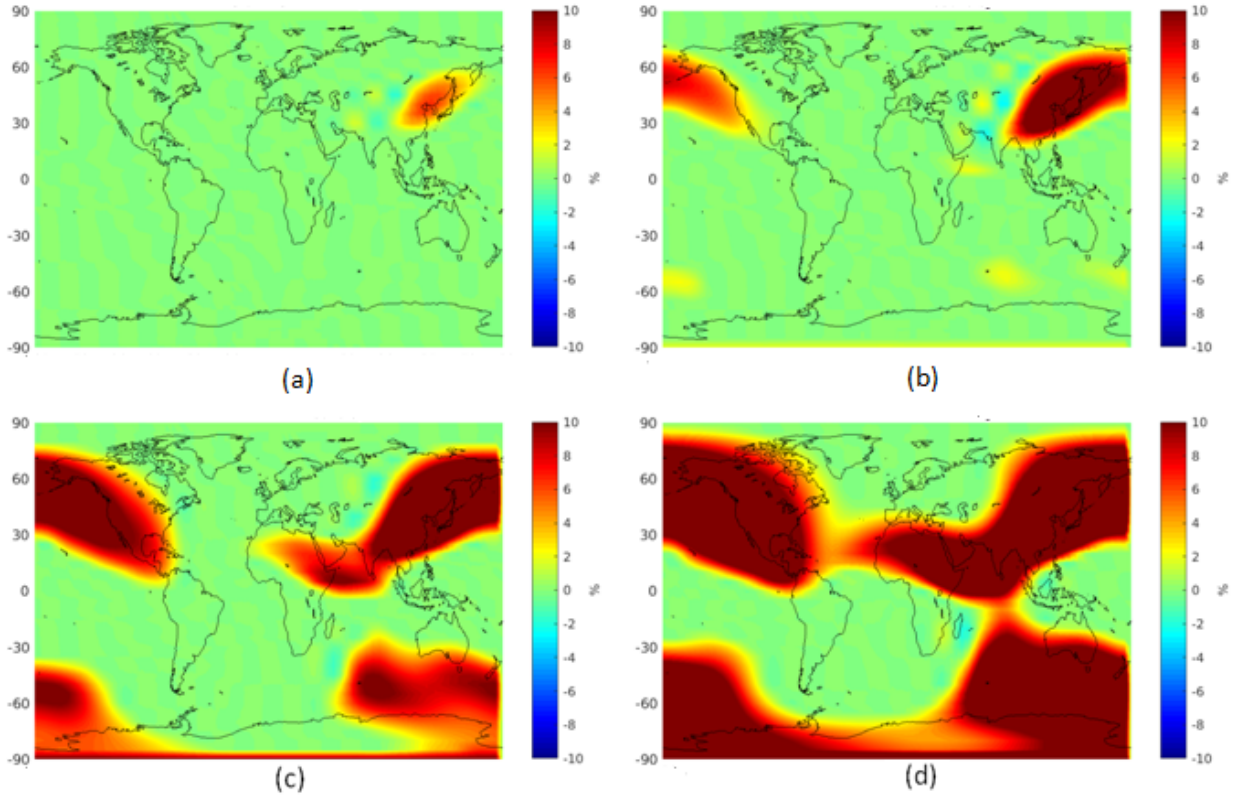


Figure 6.9: $h_m^{F_2}$ relative accuracy for four different constraint bounds. The dynamic range and slack are mentioned in the order (a) top-left: $h_m^{F_2} \geq 220$ km (b) top-right: $h_m^{F_2} \geq 240$ km (c) bottom-left: $h_m^{F_2} \geq 260$ km (d) bottom-right: $h_m^{F_2} \geq 280$ km. The corresponding dynamic range is reduced from 260 km in (a) to 200 km in (d) and the electron density modelling problem becomes more critically constrained.

feasible region, then we have a so called "infeasible" optimization problem (Boyd and Vandenberghe (2004)).

A constraint violation is the response of an optimization algorithm to an infeasible problem.

It shall be noted that the constraints violation are not allowed by the design of our interior-point method and hence are also not allowed in our constraint optimization problem. However, in certain undesired situations arising mostly due to poor choice of the tuning parameters, constraint violations occur when satisfying the KKT conditions is no longer possible for the given set of initial conditions, step-size, objective function and constraint bounds[20]. Specifically, the constraint violations occur[21] when the dynamic range of the constraints are reduced beyond a certain threshold. This is shown for $N_m^{F_2}$ in the evaluations E2, E3 (see Fig. 6.10).

[20] Qualitatively, other possible reasons for violation are initial value from infeasible region with large negative slack, larger step-size and lower normalized dynamic range of the constraints.

[21] The optimizations algorithms are developed in such a way that constraint violations are minimized. See the Lagrangian formulation in Section 4.4.2

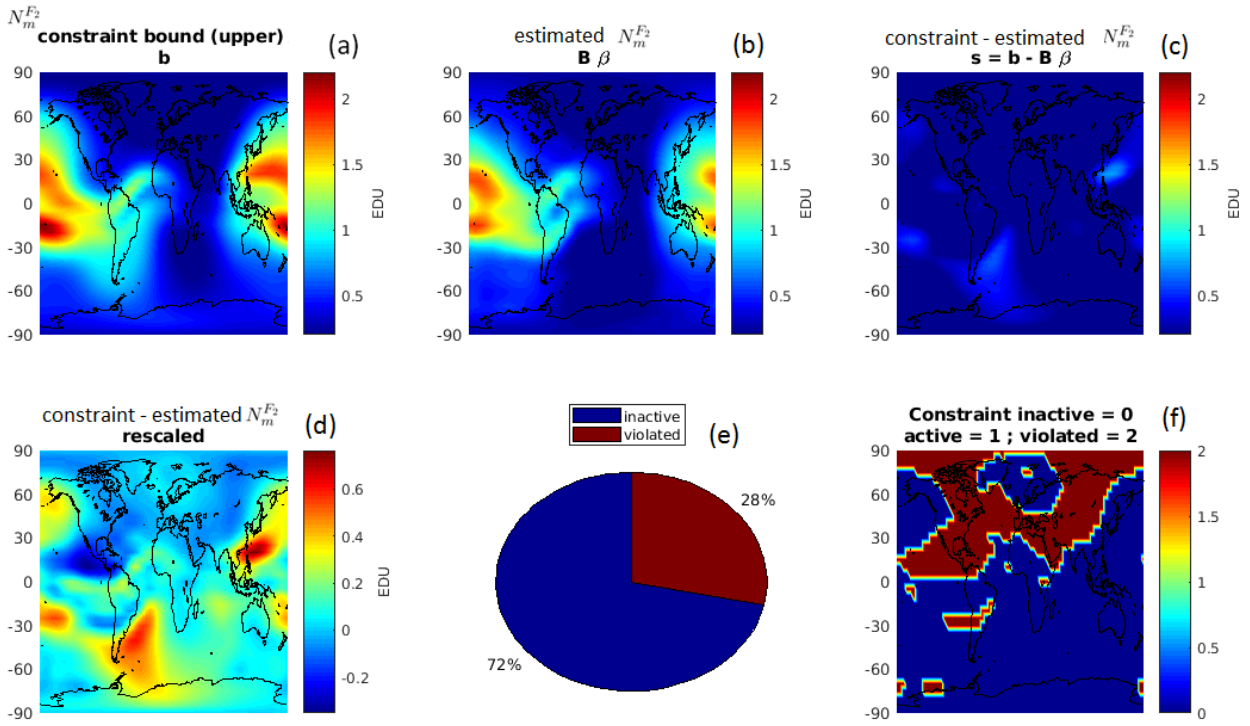


Figure 6.10: Evaluation E2: impact of $N_m^{F_2}$ upper bound constraint

A constraint $c(\boldsymbol{\beta}) \leq 0$ is violated at a candidate solution $\boldsymbol{\beta}^*$, if

$$c(\boldsymbol{\beta}^*) > 0$$

(see Chapter 4, definition 4.4.1) and, as a consequence, the slack variable becomes negative, leading to

$$\boldsymbol{B}_{ineq}\boldsymbol{\beta}^* - \boldsymbol{b}_{ineq} > 0, \text{in contrast to the Eqs. (4.74) and (4.75)}$$

There are three strategies to minimize the constraint violations :

1. Increasing the constraint upper bound (UB) or reducing the lower bound (LB), either of which increases its dynamic range (6.7).
2. Reducing the number of constraints in an optimization problem (This aspect is described in Appendix B: Constraint rationalization).
3. Reducing the step-size in the optimization algorithm.

To the best knowledge of the author, there is no general "rule-of-thumb" to select the precise number of constraints, their bounds or step-size and therefore, in this work, the slack variables, number of violated constraints etc., are analysed from the numerical evaluations. In this thesis, the contraint violations were identified and eliminated by increasing the dynamic range of the violated constraint. This requires a manual intervention to update the constraint bounds and re-execute the procedure with updated Lagrangian function (which is undesirable from an operational application). However the scope of this thesis was to understand the performance as well as the limitations of the chosen interior point method. Therefore it was considered valuable to not only desccribe the functional procedure, its advantages but also the potential problems arising from violations even if they are undesirable. Within a few experimental campaigns, we found the appropriate set of inidial consitions with which our electron density modelling problem (5.1) was set up without any

constraint violations at all. Selected campaigns to identify and eliminate constraint violations are included in this chapter and in Appendix C.

From the Sections 4.5 and 5.2, it follows that the minimization of the Lagrangian function is unconstrained with regard to the B-spline coefficients of the key parameters. In other words,

> the Lagrangian minimization is a least constraint violation problem (LCVP) (see Boyd and Vandenberghe (2004)).

A detailed theory on the constraint violation can be found in Bonnans J.F. (2000) and Ray et al. (2009). The next section will describe the numerical evaluations in detail. Particularly, the analysis will be based on the three output parameters:

- $N_m^{F_2}$ and $h_m^{F_2}$ slack variable
- $N_m^{F_2}$ and $h_m^{F_2}$ relative deviation (or accuracy)
- Percentage of $N_m^{F_2}$ and $h_m^{F_2}$ constraint violations.

Evaluations E1 - E8 and E9 - E12 will analyse the slack variable the RMS relative deviation respectively. In all evaluations, the percentage of violated constraints[22] will be shown.

Even though we have shown that a given constraint qualification can lead to violation, it shall be noted here that the main results of this thesis, i.e. the estimated key parameters which are used for the validation in Chapter 7, no violations were found. Only for the sake of completeness, to test the robust operation of our developed software as well as to understand the impact of the independent variables on the overall performance of ICOA, the concept of violations was brought in.

6.3 Evaluations E1 - E2

The goal of this section is to describe the evaluation E1 to E8 with respect to the different input dynamic ranges, hypervolume and to analyse the output slack variable. Especially, we are interested in determining the minimum slack variable (see summary of the slack variables in the evaluations E1 to E8 in Figures 3 and 8 presented later.). It signifies, the "Mahalabonis-distance" between the optimal solution and the constraint bound. Qualitatively, small positive values of the slack variable indicate that the solution comes close to being active[23]. If the slack variable is small negative, then it means the algorithm was not able to determine a solution within the feasible region. In this the optimization algorithm must be configured to do one of the following two:

1. Stop the algorithm and report no solution is possible
2. Accept the solution in the infeasible region or, in other words, allow a constraint violation but minimize it.

The second approach is comparatively practical and hence used more often (see e.g. Mehrotra (1992), Tütüncü et al. (2003)). It shall be noted that constraint violations must be avoided. As described before, we started with the minimization of the Lagrangian function or equivalently minimization of the constraint violation. This approach, although not ideal[24] but is more suited for many applications, including the electron density modelling. Also, in this work this approach

[22] For better readability, the % of violations are rounded off to the nearest integer.

[23] Furthermore, it also signifies that for the given step-size in the algorithm, this is the closest solution possible within the feasible region.

[24] The rationale for calling this approach "non-ideal but practical" implicitly involves a trade-off between having an acceptable solution with a tolerance and not having a solution at all. As it will be shown in our evaluations that a constraint violation is often associated with a relatively degraded solution accuracy compared to that for an inactive or active constraint. In our case of electron density modelling, constraint violations are not acceptable and hence the constraint bounds had to be carefully chosen for a given numerical evaluation. Furthermore, in those evaluations where a violation has been reported is only to demonstrate the consequence of a poor choice of constraint bound.

is followed. Naturally then, we are not only interested in the optimal solution but also in the associated slack variable to ensure that there had been no constraint violations. This aspect has received less attention in past and therefore motivates inclusion in numerical evaluations. At this point, it shall be noted

> the constraint violation is an undesired phenomena that as much as possible, shall be avoided at all the computation epochs.

Table 6.2 lists the goal of each evaluation. Furthermore, each evaluation will be summarized in a set of six panels. The evaluation results of E1 to E4, shown in Figs. 6.11, 6.10, 2 and 1, are organized as follows. The evaluations E1 and E2 will be shown in this chapter and the remaining evaluations are shown in Appendix C. In each case, the panel (a) on top left represents the constraint bound (either lower or upper) as defined in the optimization problem[25]. The remaining five panels (b) to (f) show the evaluation results. Specifically, in each case, the panel (b) shows the estimated key parameter from the estimated B-spline coefficients. Panels (c) and (d) show the slack variable, (e) and (f) show the proportion and locations of constraint violations, respectively.

6.3.1 Evaluation E1: Reference scenario

The evaluation E1 is considered as a reference scenario to analyse or compare the results of E2 to E4 and is performed with nominal values for the input parameters (see Table 6.4). The three input parameters ($\bar{\mathcal{D}}^r_{N_m^{F_2}}$, $\bar{\mathcal{D}}^r_{h_m^{F_2}}$ and $\bar{\mathcal{V}}$) and their corresponding values in the Table 6.4) are used to quantify the constraints. The input and output of evaluation E1 is summarized in Fig. 6.11, where the top panel shows the upper bound of input $N_m^{F_2}$ and the output $\widehat{N}_m^{F_2}$ calculated from the corresponding estimated B-spline coefficients in the panels (a) and (b) respectively. The top-right panel (c) shows the $N_m^{F_2}$ slack, or equivalently, the difference between the top-left and the top-middle panels (a) and (b) respectively, computed using

$$\widehat{\boldsymbol{s}} = \boldsymbol{b} - \boldsymbol{B}\widehat{\boldsymbol{\beta}}, \tag{6.13}$$

where $\boldsymbol{b}$, $\boldsymbol{B}$ and $\widehat{\boldsymbol{\beta}}$ denote the constraint bound, constraint coefficient matrix and the estimated B-spline coefficients respectively according to Eqs. (4.63), (4.74) and (4.78). The bottom left panel (d) in Fig. 6.11 shows the same $N_m^{F_2}$ slack variable (as the top-right (c)) with a modified scale and the positive values indicate that the KKT condition (4.72b) is satisfied at $\widehat{\boldsymbol{\beta}}$. The bottom panels (e) and (f) show the proportion of $N_m^{F_2}$ constraints and the corresponding regions, where those are either inactive, active or violated respectively. In this case, all $N_m^{F_2}$ constraints remained inactive and there are no violations (see Chapter 4, definition 4.4.1).

6.3.2 Evaluation E2: Varying $N_m^{F_2}$ upper bound

In E2, the goal is to analyse the impact of a reduced dynamic range of $N_m^{F_2}$, compared to that in E1, on the corresponding slack variable. As already discussed, the dynamic range can be controlled by the magnitudes of both upper as well as lower bound (see Fig. 6.3). Specifically in E2, the upper

[25] Although panel (a) shows only one of the two but both the upper or lower bound constraints are nevertheless imposed on the key parameters $\kappa \in \mathcal{K}_1$. When panel (a) shows the lower bound, then the associated slack variable (d) is also computed with respect to the lower bound. Similarly, if panel (a) shows the upper bound, then the associated slack variable (d) is also computed with respect to the upper bound.

Table 6.3: Common input variable or parameters for evaluations E1 to E8. The phrase 'norm. avg. DR' refers to the normalized average dynamic range. In order to fit the tables to within page width, the text is shortened.

Input variable or parameter	value
Key parameters in $\mathcal{K}_1$	see Eq. (6.1)
Key parameters in $\mathcal{K}_2$	see Eq. (6.2)
Modelling epoch	12 March 2015
Temporal resolution for the key parameters	1 hour
B-spline levels	{4,3} for key parameters in $\mathcal{K}_1$
Spatial resolution for the constraints	5 ° along latitude and longitude
Temporal resolution for the constraints	6 hours
Norm. avg. DR $\bar{\mathcal{D}}^r_{H^{F_2}}$	0.66
Norm. avg. DR $\bar{\mathcal{D}}^r_{N_0^P}$	0.5
Norm. avg. DR $\bar{\mathcal{D}}^r_{H^P}$	0.5

Table 6.4: Configuration for numerical evaluation E1.

	variable or parameter	value
Input	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.5
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.8
	average hypervolume $\bar{\mathcal{V}}$	0.066
Output	minimum $N_m^{F_2}$ slack	0.005 EDU
	% of $N_m^{F_2}$ constraints violated	0

bound of $N_m^{F_2}$ is reduced by 12%[26] compared to E1[27], with the other inputs remaining the same (see Table 6.5).

The reduced $N_m^{F_2}$ upper bound and that modelled from the corresponding estimated B-spline coefficients are shown in the top panel of Fig. 6.10 in the left and middle panels (a) and (b) respectively. The modelled $N_m^{F_2}$ exceeds (and thus violates) 28% of the $N_m^{F_2}$ upper bound constraints and a majority (> 85%) of these occur between the local-time 20:00 - 03:00 LT (night side)[28] of the northern hemisphere.

The physical interpretation of a constraint violation is described below. The atmospheric ionization is lower during the night[29] and accordingly the electron production, and hence $N_m^{F_2}$, is also reduced compared to that of the day time in the ionosphere (Bauer (1973), Oshiorenoya (2004)). In this as well as in all other evaluations, the upper or lower bounds are uniformly reduced or increased across the globe. Specifically in E2, $N_m^{F_2}$ upper bound is reduced by 12% relative to that in E1, where the minimum slack was already only 0.005 EDU (see Table 6.4), which also happened to be at the night side. Upon a closer examination of Fig. 6.10, it is concluded that in those regions, where the $N_m^{F_2}$ upper bound was γ−active ($\approx$ 0.1 EDU and see Chapter 4, definition 4.4.1), a uniform 12% reduction of the upper bound leads to the observed constraint violations. On the other hand, the

[26]The justification for decreasing the upper bound exactly by 12% is as follows: Different values of decreasing the upper bound were tried in the range of 5 to 20% and the value of 12% was then chosen based on the need for maintaining a consistency among the hypervolume in the evaluations E1 to E8; see Figs. 9 and 10. Other reasons were to have a fair tradeoff between the computational performance, time required for the execution and the memory requirements.

[27]In all numerical evaluations, the constraints are "uniformly increased or decreased" across the globe. For example a 12% reduction in dynamic range specifically refers to a uniform 12% reduction in bound of all constraints.

[28]Electron density is generally lower on the night side compared to the day side.

[29]It shall be noted that the day and night side are only qualitatively defined and not with respect to a fixed time of the day.

$N_m^{F_2}$ upper bound in E1 had a relatively larger associated slack at local times 06:00 - 18:00 (the day side) and therefore, despite a 12% reduction of the upper bound, there is less than 15% $N_m^{F_2}$ violations in E2.

The reported percentage of violations in each evaluation is an important aspect to understand the role of constraint bounds and the capability of the optimization algorithm as well as to the physical realism, the ionospheric dynamics introduced by the constraints in the electron density modelling. As already mentioned, a constraint violation is not acceptable and hence the constraint bounds will have to be carefully considered.

Table 6.5: Configuration for numerical evaluation E2.

	variable or parameter	value
Input	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.44 (12% lower UB than E1)
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.8 (same as that used in E1)
	average hypervolume $\bar{\mathcal{V}}$	0.058
Output	minimum $N_m^{F_2}$ slack	-0.35 EDU
	% of $N_m^{F_2}$ constraints violated	28

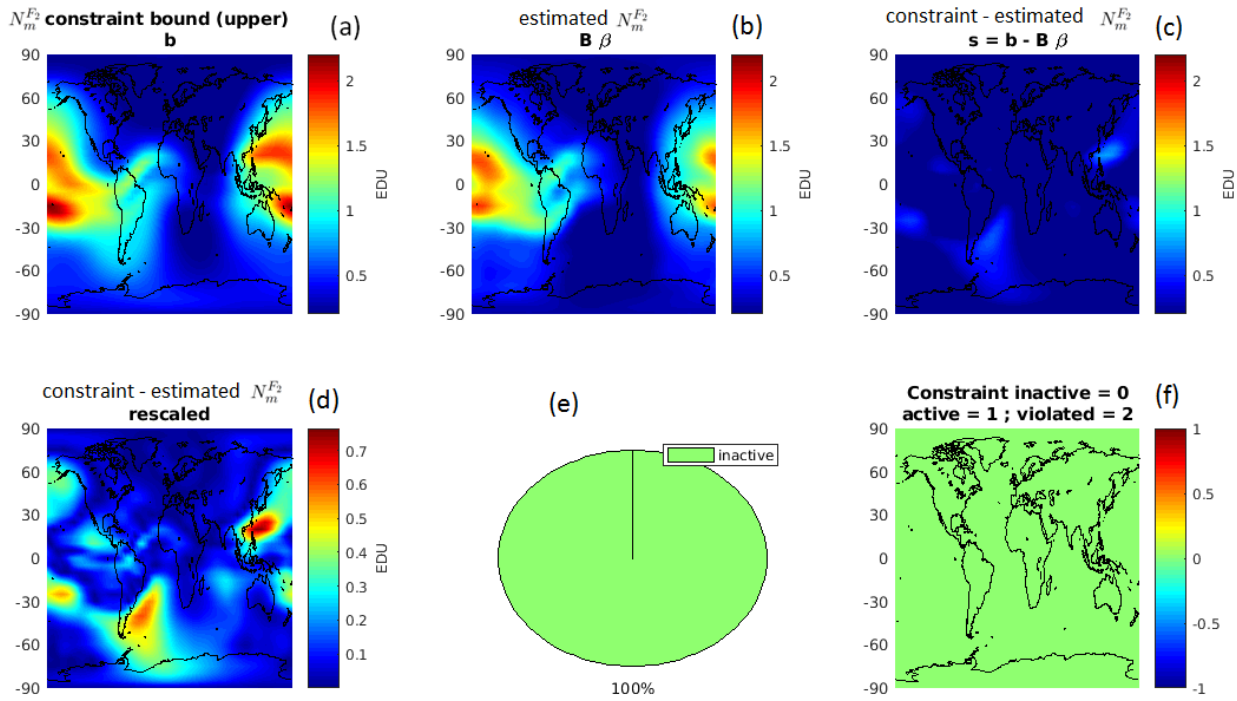


Figure 6.11: Evaluation E1: Slack and modelled $N_m^{F_2}$

6.4 Summary of numerical evaluations

Together with the evaluations shown in Appendix C, this chapter presented a set of 12 numerical evaluations (including 54 sub-evaluations). These have been performed based on the independent variables listed in Section 6.1 and their results are analysed. The evaluations E1 to E8 analysed the minimum value of the slack variable and the percentage of violations for the different nomalized dynamic ranges using two real observation datasets DS2d and DS2a. The evaluations E9 to E12 analyse the $N_m^{F_2}$ and $h_m^{F_2}$ relative deviation to the different upper and lower constraint bounds, based

on DS1d dataset. This allow for the determination of constraint bound values at which violations started to occur.

Being both time and memory consuming, the evaluation was limited to the investigation of the inequality constraints' impact on the electron density modelling. The following are the most important outcome of the numerical evaluations.

When multiple ionosphere layers are to be considered in electron density modelling problem, the total number of Chapman key parameters and hence the B-spline coefficients to be estimated increase[30]. However, the unavailability of realistic initial values of the scale height for those layers, especially the bottom layers is a challenge to application of our model. Also the imposition of inequality constraints on key parameters means that additional Lagrange multipliers and slack variables have to estimated together with the unknown B-spline coefficients.

Furthermore, the dynamic range of N_m^D, N_m^E constraints are lower relative to that of $N_m^{F_2}$. Specifically, it refers to the choice of constraint bounds. If the bounds on the key parameter κ_p are chosen "conservatively" (normalized dynamic range $\breve{\mathcal{D}}^r_{\kappa_p} \approx 1$), then the advantage of using the inequality constraints are lost, whereas if chosen too "critically" (normalized dynamic range $\breve{\mathcal{D}}^r_{\kappa_p} \ll 1$), then the estimated B-spline coefficients become inaccurate[31], as demonstrated in the evaluations E9 to E12.

> A large positive slack at the optimal solution indicates that the constraints remained inactive. In other words, the constraints were chosen rather "conservatively". On the other hand, a small positive slack at the optimal solution indicates that the constraints became $\gamma-$active.

It is not known apriori as to which inequality constraints will remain inactive or equivalently, if a given constraint $c(\boldsymbol{\beta})$ will become active or not. This is demonstrated in the evaluations and it helps in highlighting the two important phenomena (1) latitude and local time dependency of constraint violations and (2) those constraints which remain inactive are all "$\gamma-$active" with the 95^{th} percentile of $N_m^{F_2}$ and $h_m^{F_2}$ slack being equal to or smaller in magnitude than 0.2 EDU and 20 km respectively. Minimum value of the slack for each evaluation was analysed and its relevance described. This allows identification and rationalization of the inactive constraints for near real-time electron density modelling as well as to choose the constraint bounds such that the violations are minimized.

Although not shown, but in general, a reduction of the dynamic range of any one key parameter $\kappa_1 \in \mathcal{K}_1$ also affects the estimated B-spline coefficients of the other key parameters, $\kappa \in \mathcal{K}_1;\ \kappa \neq \kappa_1$, as well as their slack variables due to mutual correlations.

Although the negative slack is not allowed by the KKT conditions, but the Lagrangian minimization becomes an unconstrained optimization problem through a series of transformations (see Section 4.4.2).

Numerical evaluations E1 to E12 have focussed on the analysis of the slack variable and deviation (accuracy) of $N_m^{F_2}$ and $h_m^{F_2}$. The magnitude (and sign) of the resulting slack variable gives important information to improve the step-size and/or the initial value used, particularly when the constraints are $\gamma-$active or violated.

Analysis in this work reveals as to which constraints remain inactive in electron density modelling. In fact the KKT conditions are so designed that the constraint violations are penalized by positive Lagrange multiplier, thereby increasing the Lagrangian function. However, qualitatively, the descent direction tends to "pull-in" the solution to inside of the feasible region.

[30] compared to when only one layer (e.g. the F_2) is to be considered

[31] There would be no solution inside the feasible region when normalized dynamic range $\breve{\mathcal{D}}^r_{\kappa_p} \ll 1$

7 Validation of electron density modelling

The Chapters 4 and 5 have described the inequality constrained optimization algorithm (ICOA) and its application to the development of a global 4D EDM using B-spline basis functions respectively. The results from the numerical evaluation of the developed EDM has been presented in Chapter 6.

There are four main outputs of the developed electron density model, namely, the B-spline coefficients, the slack variables, the Lagrange multipliers and the differential code biases (DCBs).

The most important of these are the estimated B-spline coefficients corresponding to the Chapman key parameters in the set $\mathcal{K}_1$ (see Eq. (5.16)). The slack variables and the Lagrange multipliers are already qualitatively analysed in the numerical evaluations. The DCBs are related to the GNSS satellite and receiver signal tracking techniques (see Montenbruck and Hauschild (2013), Montenbruck et al. (2014a)) and are not validated in this work[1].

In this chapter, the following four methods are presented for the validation of the modelled Chapman key parameters[2]:

- Closed loop validation (see Section 7.1)
- Ionosonde based validation (see Section 7.2)
- GNSS STEC based validation (see Section 7.3)
- In-situ observations based validation (see Section 7.4), specifically, using Langmuir probe on-board CHAMP and Swarm satellites, as well as using Swarm STEC observations.

The following conventions shall be applied, in addition to those defined in the beginning of Chapter 6:

- The terms "ionosonde data" or "ionosonde observations" refer to the F_2 peak density and peak height obtained from the ionosonde stations, denoted as $N^{F_2}_{m;Ionosonde}$ and $h^{F_2}_{m;Ionosonde}$.
- "ICOA estimated B-spline coefficients" refer to the estimated values of the unknown B-spline coefficients corresponding to the key parameters in the set $\mathcal{K}_1$.
- "ICOA modelled key parameters" refer to the Chapman key parameters computed from the "ICOA estimated B-spline coefficients", denoted as $N^{F_2}_{m;ICOA}$ and $h^{F_2}_{m;ICOA}$.
- GNSS STEC observations refer to the STEC computed between a GNSS satellite and ground receiver (see Section D.2).
- "Swarm STEC" refers to the STEC measured between a GNSS satellite and the spaceborne GNSS receiver on-board the Swarm A, B and C satellites, denoted as $STEC_{Swarm}$.
- "In-situ observations" refer to the electron density measurements of the Langmuir probe on board Swarm A, B, C as well CHAMP. These are denoted as $N_{e;CLP}$, $N_{e;SLP}$.
- For the sake of readability, the multi-layer Chapman key parameters will be simply referred to as "Chapman key parameters" or equivalently as the "key parameters in the set $\mathcal{K}_1$".

[1] The DCBs are not direct geophysical parameters but become relevant due to their correlations to the GNSS observables described in Section D.1. More specifically, for this work, DCBs are not the primary parameters of interest and are assumed to be estimated reliably and not subject to validation.

[2] Validation of the modelled Chapman key parameters is also a direct validation of the overall global 4D electron density model.

It shall be noted that GNSS STEC computation was performed from the raw RINEX data within the scope of this work, as shown in Section D.2. STEC observations from GNSS are directly used in the dataset DS1d for the electron density modelling (see Table 13 and Fig. 16). On the other hand, Swarm STEC only captures the contribution of the ionosphere above the Swarm satellite altitude. Furthermore, those were directly obtained from ESA (see Stolle et al. (2013)) and not computed within the scope of this work. The only purpose of obtaining Swarm STEC in this work is for the validation of the modelled Chapman key parameters. This aspect will be described in detail in Sections 7.3 and 7.4.

In the first part of this chapter, the closed loop validation (CLV) technique and its dataflow will be presented. Subsequently, in the Sections 7.2 and 7.3, the validation with ionosonde and STEC observations will be described respectively. For the sake of completeness, Langmuir probe (see Section 7.4.1) and Swarm STEC (see Section 7.4.2) data will also be discussed. However, the latter two only are shown to be not sufficient for the validation of a global 4D EDM. Furthermore, they also require outlier detection and calibration[3]. Therefore, the CLV, the ionosonde, as well the GNSS STEC based validation are described in more detail relative to the $STEC_{Swarm}$ and Langmuir probe in-situ observations based ones.

7.1 Closed loop validation

Closed loop validation

In the context of this chapter, the closed loop validation (CLV) refers to the process, where a "known set" of multi-layer Chapman key parameters is available, with which the simulated electron density observations are generated. These are in-turn used as observations in a GMM to estimate the B-spline coefficients corresponding to the Chapman key parameters.

The known set or the "true" Chapman key parameters is obtained from a reliable source, e.g. the IRI model. Once the B-spline coefficients are estimated from the ICOA, the equivalent key parameters are validated against the "truth". Therefore, this approach is called the "closed loop validation". This approach allows a thorough analysis of the ICOA, its computational performance, the sensitivity to both the initial values of the optimization parameters, as well as to the different measurement noise levels. This type of validation has also been used in Limberger et al. (2013), Limberger et al. (2014) and Liang (2017). For this work, the IRI model[4] (Bilitza (2000)) is used with upto 5% random measurement noise, according to the F3C electron density quality reported in A.Anthes et al. (2000), Hsu et al. (2018).

Need for CLV

The development of a 4D EDM requires a global electron density observation coverage. However, despite using F3C, GRACE and GNSS STEC, there are still data gaps at any given time of a day (see Fig. 7.1). Therefore, one of the datasets (DS1a) used in the numerical evaluation, was generated from simulated electron density observations. Accordingly, the CLV approach has sufficient[5] observations for the estimation of the B-spline coefficients among the other optimization parameters $\boldsymbol{\beta}$ in the GMM (4.8) with additional inequality constraints. Furthermore, the CLV also has the advantage that it allows the determination of:

- the approximate initial values for the B-spline coefficients to be used in the optimization algorithm (see Section 4.3.1)

[3]The Swarm mission STEC measurements are used, as is from the ESA data portal (Stolle et al. (2013)), and no further quality check was made.
[4]IRI version 2012
[5]A full column rank of the design matrix $\boldsymbol{A}$, as defined in GMM (Eq. 4.8)

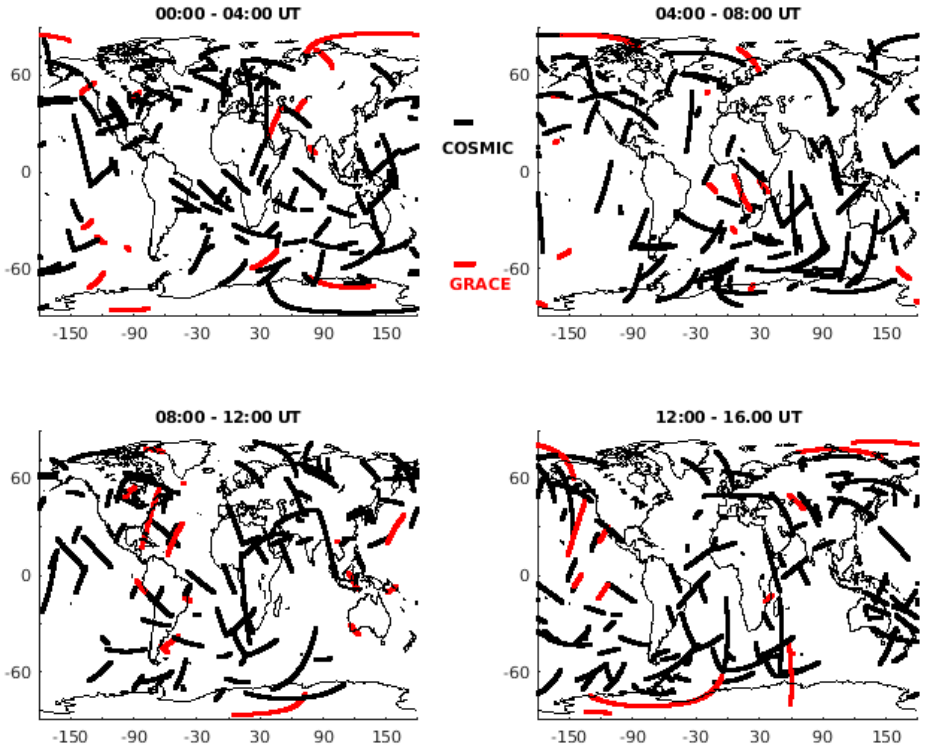


Figure 7.1: F3C and GRACE observation coverage in four different time windows.

- the step-size t_{it} to be used in the gradient descent method; see Eq. (4.29)
- the upper and lower constraint bounds κ_l, κ_u of the Chapman key parameter $\kappa \in \mathcal{K}_1$
- the two optimization algorithm tuning parameters, namely, the centering parameter σ_{cen} and the complimentarity measure μ_{com}; see Eq. (4.100).

In this section, the CLV procedure will be described. The numerical evaluations E9 - E12 (see Section C.9 - C.12) already provide the results with the dataset DS1d. There are two steps involved in the CLV and within each, there are three sub-steps. Therefore, Step 1 (with sub-steps 1a, 1b and 1c) and Step 2 (with sub-steps 2a, 2b and 2c) together describe the CLV procedure.

Step 1 of CLV

The objective of Step 1 is the generation of a "reference truth" of the multi-layer Chapman key parameters in the set $\mathcal{K}_1$

and achieved in the three sub-steps as follows:

- **Step 1a** : F_2-layer peak density $\kappa_1 = N_m^{F_2}(\varphi, \lambda)$, peak height $\kappa_2 = h_m^{F_2}(\varphi, \lambda)$, $\forall\ \varphi \in \boldsymbol{\varphi}, \lambda \in \boldsymbol{\lambda}$, are obtained from the IRI model and introduced into their corresponding B-spline series expansion (5.5). The scale height $\kappa_3 = H^{F_2}(\varphi, \lambda)$ is computed using the approximate slab thickness relation (Davies and Liu (1991)),

$$H^{F_2}(\varphi, \lambda) = \frac{VTEC(\varphi, \lambda)}{4.13\ N_m^{F_2}(\varphi, \lambda)}, \tag{7.1}$$

  see Limberger (2015).
- **Step 1b**: For the key parameters in $\mathcal{K}_1$, the equivalent B-spline coefficients $\widehat{d}_{k_1,k_2:\kappa}^{J_1,J_2}$, to be used as the "reference truth", are estimated from Eq. (5.5), using the known tensor product matrix as described in Section 5.2.

- **Step 1c** : The estimated B-spline coefficients $\widehat{d}^{J_1,J_2}_{k_1,k_2:\kappa}$ from Step 1b are substituted again in Eq. (5.5) to obtain the "true" key parameters $\widehat{\kappa}(\varphi,\lambda)$.

Step 1c "reference truth"

Thus, Step 1c provides the "true" values $\widehat{\kappa}(\varphi,\lambda)$ of the key parameters and Step 1b their corresponding B-spline coefficients $\widehat{d}^{J_1,J_2}_{k_1,k_2:\kappa}$.

In Step 2, $\widehat{\kappa}(\varphi,\lambda)$ will be further used to calculate the electron density $N_e(\varphi,\lambda,h)$ over a desired altitude range $h \in [h_{min}\ h_{max}]$ using the Chapman profile function (5.2). The B-spline coefficients and the Chapman key parameters in Step 2 are denoted using the "tilde" (e.g. $\tilde{d}$) to distinguish them from those in Step 1, denoted by the "hat" symbol (e.g. $\widehat{d}$). The three sub-steps in Step 2 are as follows:

- **Step 2a** : The estimated key parameters $\widehat{\kappa}(\varphi,\lambda)$ from Step 1c are used to generate electron density

$$N_e(\varphi,\lambda,h)\ \forall\ \varphi \in \boldsymbol{\varphi},\lambda\ \in \boldsymbol{\lambda} \text{ and } \boldsymbol{h} \in [h_{min}\ h_{max}]$$
$$\boldsymbol{\varphi} = [\varphi_{min} \cdots \varphi_{max}] \qquad \boldsymbol{\lambda} = [\lambda_{min} \cdots \lambda_{max}]$$

  using the Chapman profile function (5.2), as shown in Fig. 7.2. This results in a uniform global distribution of electron density observations with an altitude dependent vertical resolution[6] as already mentioned in Section 6.1 under the sub-title "Number of observations per technique". Clearly, the number of observations N_{obs} generated in the CLV approach depends on the resolution along the latitude, the longitude and the altitude, as well as the geographical coverage limits, i.e. whether a global or regional modelling is sought. Furthermore, N_{obs} also depends on the altitudinal limits[7] $h_{range} = |h_{max} - h_{min}|$, for a given altitudinal sampling resolution h_{step}.
- **Step 2b** : The electron density, $N_e(\varphi,\lambda,h)$, computed in the Step 2a, is used as observations in the GMM, (Eq. (4.8)).
- **Step 2c** : Finally, in this step, the KKT equations (4.94) - (4.96) are solved and the B-spline coefficients $\tilde{d}^{J_1,J_2}_{k_1,k_2:\kappa}$ estimated, as well as the corresponding Chapman key parameters $\tilde{\kappa}(\varphi,\lambda)$ are computed using Eq. (5.5).

Fig. 7.2 shows the overall data flow of the closed loop validation.

Number of optimization parameters N_β

The total number of optimization parameters N_β to be estimated depend on the number of key parameters in $\mathcal{K}_1$, their B-spline levels J_1, J_2 (see Section 5.2) and number of constraints.

More specifically, N_β also depands on $rmax$, $smax$, the number of GNSS receivers and satellites respectively, used in the STEC computation, as each receiver as well as GNSS satellite shall have one unknown differential code biases (DCB) per observable per frequency. For the sake of simplicity, the DCBs are not considered in the scope of this chapter and are assumed to have been estimated reliably. For CLV, 5 key parameters, as shown in Eq. (5.16), are considered in the set $\mathcal{K}_1$, with the same B-spline level (4, 3)[8].

[6] It shall be noted that instead of using the electron density from the IRI model directly as observations, the reference truth Chapman key parameters, $\widehat{\kappa}(\varphi,\lambda)$, are used to generate the electron density. This leads to a overall consistency in the CLV procedure.

[7] The absolute value operator in computing the h_{range} is not necessary, if $h_{max} > h_{min}$ is ensured within the electron density modelling problem definition. In the software prototype, it was used to catch any exceptions and therefore is shown in this way.

[8] This is just one example. More generally, larger B-spline levels allow a finer spatial scale representation of the ionospheric features (see Schmidt (2011), Erdogan et al. (2020)) but also increase the number of B-spline coefficients to be estimated.

Step 2b provides the estimated B-spline coefficients $\tilde{d}^{J_1,J_2}_{k_1,k_2:\kappa}$ and Step 2c transforms them into the respective key parameters $\tilde{\kappa}(\varphi,\lambda)$. To characterize the accuracy of the estimated Chapman key parameters, the deviation $\Delta\kappa$, of the key parameter κ, is defined as

$$\Delta\kappa(\varphi,\lambda) = \widehat{\kappa}(\varphi,\lambda) - \tilde{\kappa}(\varphi,\lambda). \tag{7.2}$$

It denotes the difference between the estimated key parameters from Step 2c and Step 1c. In a similar manner, the B-spline coefficient deviation Δd_κ is also defined

$$\Delta d_\kappa = \widehat{d}^{J_1,J_2}_{k_1,k_2:\kappa} - \tilde{d}^{J_1,J_2}_{k_1,k_2:\kappa}. \tag{7.3}$$

The relative deviation $\Delta_{rel}\kappa = \frac{\Delta\kappa(\varphi,\lambda)}{\widehat{\kappa}(\varphi,\lambda)}$ is also used to characterize the validation results relative to the reference truth $\widehat{\kappa}(\varphi,\lambda)$.

For the CLV, the elements of the subset $\mathcal{K}_2$ are obtained from the IRI model[9]. Furthermore, a consistent uniform 5 deg horizontal spatial resolution with altitudes between h_{min} = 100 km and h_{max} = 1000 km, is assumed. The altitude resolution h_{step} = 5 km in the F_2-layer (250 km $\leq h \leq$ 450 km) compared to h_{step} = 20 km in the remaining bottom ($h_{min} \leq h \leq$ 250 km) and topside (450 km $\leq h \leq h_{max}$) ionosphere, thus resulting in approximately 240000 electron density observations globally at a given time epoch.

In addition to the relative deviation $\Delta_{rel}N_m^{F_2}$, presented in Sections C.9, C.10, C.11 and C.12, Fig. 7.3 shows the relative deviation $\Delta_{rel}N_m^{F_2}$ obtained from the CLV with 0%, 1%, 3% and 5% noise added to the simulated electron density observations obtained in Step 2a. The measurement noise is added as a function of altitude, such that the lower altitudes ($h <$ 350 km) have a relatively larger noise[10]. RMS relative deviation of the three F_2 layer Chapman key parameters from the CLV is shown in Table 7.1. The random variations in the relative deviation in the order of 10^{-12}% for no noise and 10^{-3}% for the remaining three cases, in Table 7.1, Fig. 7.3, indicates that the optimization algorithm has converged effectively. Similar performance of the relative deviations $\Delta_{rel}N_m^{F_2}$, $\Delta_{rel}h_m^{F_2}$ were already reported in the numerical evaluations with the simulated dataset in Tables 8, 9 and 10. This is the most direct validation of the ICOA[11].

Table 7.1: Impact of noise on RMS relative deviation of the three F_2-layer key parameters

Noise level	$\Delta_{rel}N_m^{F_2}$ (%)	$\Delta_{rel}h_m^{F_2}$ (%)	$\Delta_{rel}H^{F_2}$ (%)
No noise	7.01×10^{-13}	7.25×10^{-13}	2.12×10^{-13}
1%	6.18×10^{-4}	7.10×10^{-4}	2.60×10^{-4}
3%	1.86×10^{-4}	1.10×10^{-4}	3.40×10^{-5}
5%	3.08×10^{-3}	3.30×10^{-4}	9.30×10^{-5}

The four rows in Table 7.1 correspond to four independent datasets with 0%, 1%, 3% and 5% measurement noise respectively[12]. As expected, the impact of increasing measurement noise is a corresponding increase in RMS relative deviation[13] as shown in Fig. 7.3. The optimization algorithm

[9] Specifically, the $D-$, $E-$ and F_1- peak density and peak height are modelled from IRI directly, whereas the corresponding scale heights are assumed to be one-half of that of F_2- layer.

[10] as generally observed in F3C, GRACE IRO profiles; see Tsai et al. (2016)

[11] Obtaining confidence in the design and development of ICOA from CLV is considered a big step forward in the potential future usage of inequality constraints in a more general class of parameter estimation problems in Geodesy.

[12] The measurement noise referred here is quantified as the percentage of the maximum electron density.

[13] As an exception, for the case of scale height, the increase in noise appears to decrease the RMS relative deviation but only marginally from 2.60×10^{-4} to 3.40×10^{-5}.

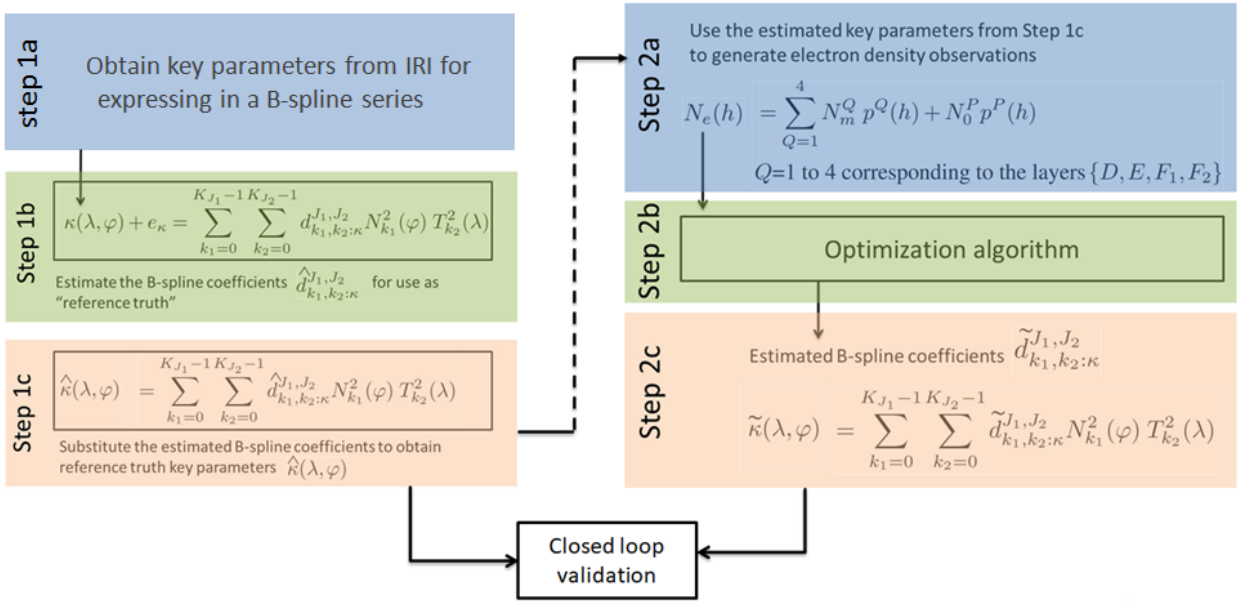


Figure 7.2: Generation of reference values for B-spline coefficients and multi-layer Chapman key parameters

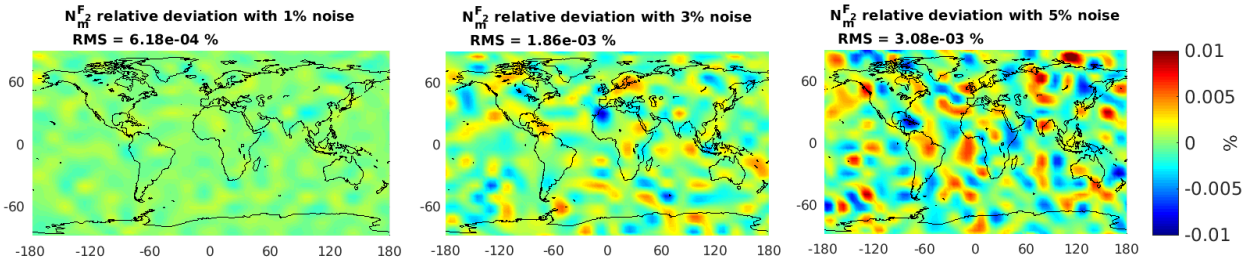


Figure 7.3: Impact of observation noise on relative deviation of peak density

(Gradient descent and Newton's method: see Sections 4.3.2, 4.3.3) was used in the same manner for the noisy observations as with the noise-free case[14].

7.2 Validation with ionosonde

In this section, the validation of the modelled key parameters in $\mathcal{K}_1$, using ionosonde data, will be presented. In contrast to the previous section, where simulated electron density observations were used, in this section the validation will performed on the B-spline coefficients estimated from IRO profiles derived from F3C, GRACE and GNSS STEC observations (dataset 1d)[15]. It shall be noted that the simulated and the real observations were used exactly in the same manner by the algorithm[16].

The period between 12 to 14 March 2015 is chosen for validation during which, the B-spline coefficients of the key parameters in $\mathcal{K}_1$, are estimated from dataset DS1d, at a 1 hour temporal resolution. During this time span, space weather activity was nominal with an average F10.7 index

[14]This is to signify that only the VCE procedure, as described in Section 4.1.2, was used and no additional parameters or techniques were introduced to smooth the effect of measurement noise in the optimization algorithm.

[15]This is just to clarify that the CLV used a synthetic dataset and ionosonde based validation will be based on real data. However, it is still to be noted that the real dataset used in this section still contain the electron density from VTEC with separability approach.

[16]It means the algorithm is not tailored for use by any specific type of dataset.

of 68.8 SFU[17]. Fig. 7.4 shows a comparison of the estimated $N_m^{F_2}$ by ICOA and that from IRI as well that from the GR13L ionosonde station.

Since ionosonde stations generally use different software and techniques, such as POLAN, ARTIST (see Titheridge (1985), McNamara et al. (1987), Galkin et al. (2008), Galkin and Reinisch (2008), Galkin et al. (2013)), to scale ionograms[18], systematic offsets arise (see Gilbert and Smith (1988)) and are not discussed further in this work.

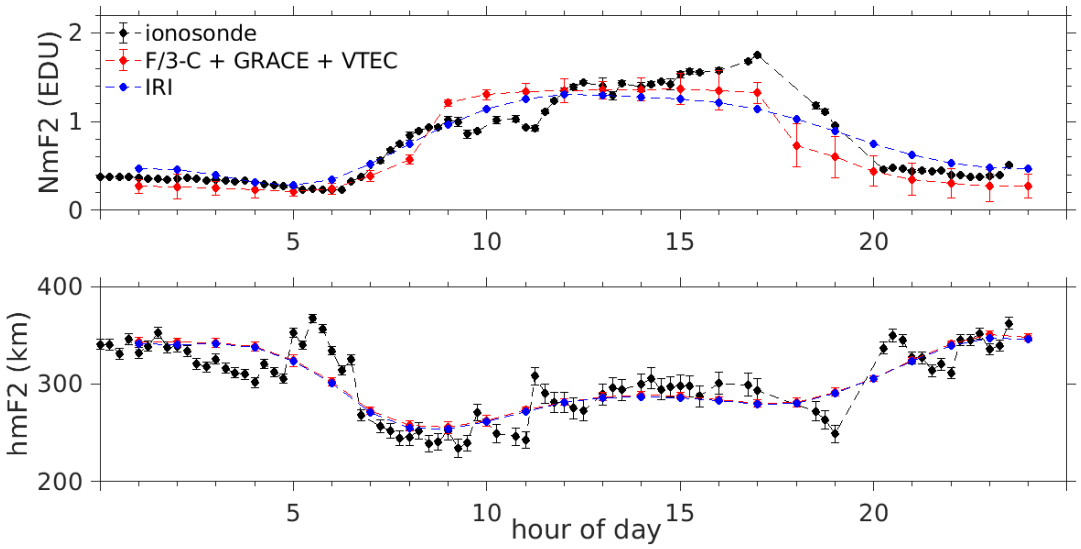


Figure 7.4: Estimated key parameters comparison with ionosonde station GR13L and IRI model

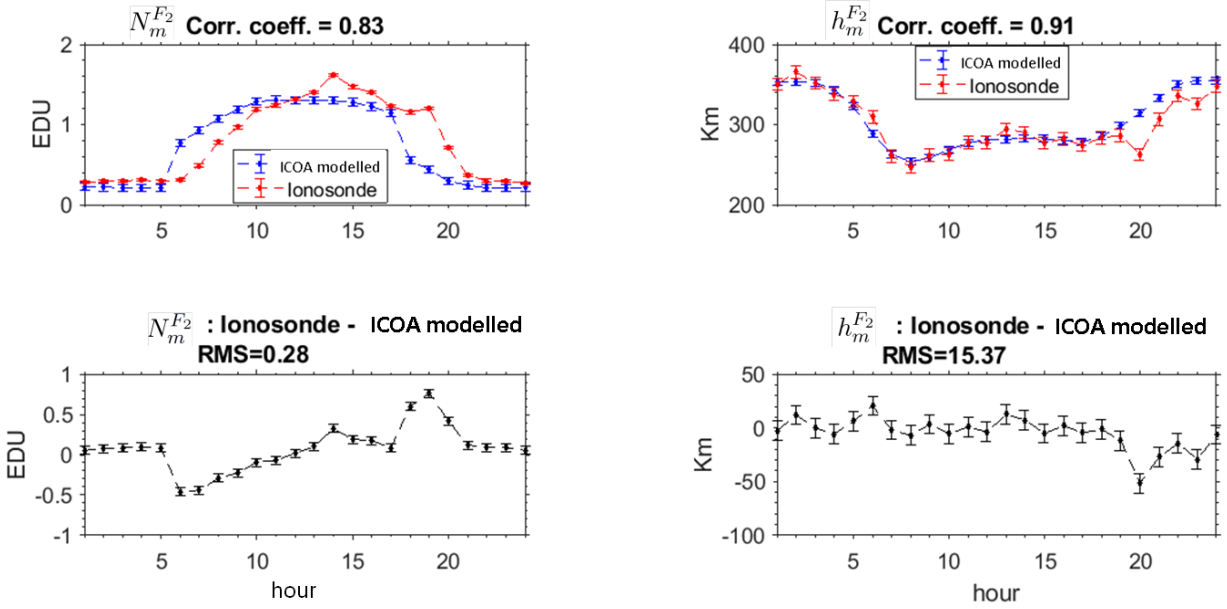


Figure 7.5: Estimated key parameters deviation from ionosonde station GR13L

Using the ICOA estimated B-spline coefficients, the estimated key parameters $N_{m;ICOA}^{F_2}$ and $h_{m;ICOA}^{F_2}$ are computed using the Eqs. (5.10) and (5.11) at the exact locations of the globally distributed

[17]SFU refers to Solar Flux Units

[18]Assumptions on optical properties of the ionosphere layers are made (see e.g. Wilkes (1997)) for the computation of $N_m^{F_2}$ from $f_0^{F_2}$.

Table 7.2: Deviations $\Delta N_m^{F_2}$ and $\Delta h_m^{F_2}$ and the corresponding relative deviations at selected ionosonde station locations along with their 3-day mean and standard deviations.

Station ID	$\Delta N_m^{F_2}$ (EDU)	$\Delta h_m^{F_2}$ (km)	$\Delta_{rel} N_m^{F_2}$ (%)	$\Delta_{rel} h_m^{F_2}$ (%)
GR13L	0.07 ± 0.16	-19.0 ± 9.28	10.45	6.40
EA036	0.02 ± 0.13	-0.98 ± 8.9	6.81	2.14
JR055	0.03 ± 0.18	-6.28 ± 9.44	9.54	3.80
AS00Q	0.04 ± 0.05	-18.0 ± 8.45	4.09	6.18
AH223	-0.15 ± 0.06	-3.23 ± 6.58	4.09	13.3
AT138	0.05 ± 0.12	10.34 ± 7.23	7.72	4.21
EB040	0.01 ± 0.11	21.35 ± 6.46	5.45	6.66
EG931	-0.11 ± 0.05	-10.87 ± 6.67	7.27	4.12
FF051	0.01 ± 0.11	-4.32 ± 12.61	5.45	4.04
GU513	-0.02 ± 0.10	3.76 ± 10.11	5.45	3.30
HE13N	0.04 ± 0.07	12.45 ± 4.67	4.54	4.09
IC437	0.04 ± 0.13	16.23 ± 7.87	7.70	5.73
IF843	-0.05 ± 0.11	-7.88 ± 9.67	7.27	4.11
IS141	-0.01 ± 0.15	-2.87 ± 4.70	7.27	1.69
JI91J	0.02 ± 0.08	14.11 ± 12.11	4.54	6.24
JJ433	0.06 ± 0.13	11.31 ± 3.25	8.63	3.46
LG178	0.12 ± 0.11	-8.56 ± 6.57	10.45	3.60

ionosonde stations; see Table 3.1. For the same time period, $N^{F_2}_{m;Ionosonde}$ and $N^{F_2}_{m;Ionosonde}$ are directly measured in those stations for the validation. The respective scale heights were not available from the ionosonde stations and therefore could not be validated. At the latitude and longitude $\varphi_i,\ \lambda_i$ of the ionosonde stations, indexed by the subscript $i\ \in 1, \cdots, N_{ionosonde}$, the deviations

$$\begin{aligned} \Delta N^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i) &= N^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i) - N^{F_2}_{m;ICOA}(\varphi_i, \lambda_i) \\ \Delta h^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i) &= h^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i) - h^{F_2}_{m;ICOA}(\varphi_i, \lambda_i) \end{aligned} \tag{7.4}$$

as well as the corresponding relative deviations

$$\begin{aligned} \Delta_{rel} N^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i) &= \frac{\Delta N^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i)}{N^{F_2}_{m;ICOA}(\varphi_i, \lambda_i)} \\ \Delta_{rel} h^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i) &= \frac{\Delta h^{F_2}_{m;Ionosonde}(\varphi_i, \lambda_i)}{h^{F_2}_{m;ICOA}(\varphi_i, \lambda_i)} \end{aligned} \tag{7.5}$$

are computed using the two sources of the key parameters. While the latter are directly obtained from the ionosonde stations, along with their standard deviations without any further calibration, the former ones are computed from the estimated B-spline coefficients using Eq. (5.10) and (5.11) and their corresponding standard deviations from a Monte-Carlo (MC) approach, in which the input electron density observations are varied by 1% in 100 MC simulation runs to obtain their ensemble mean and standard deviation. As shown in the Tables 7.2 and 7.3, $\Delta N^{F_2}_{m;Ionosonde}$ and $\Delta h^{F_2}_{m;Ionosonde}$ deviations are near-randomly distributed between ± 0.2 EDU, ± 30 km globally and

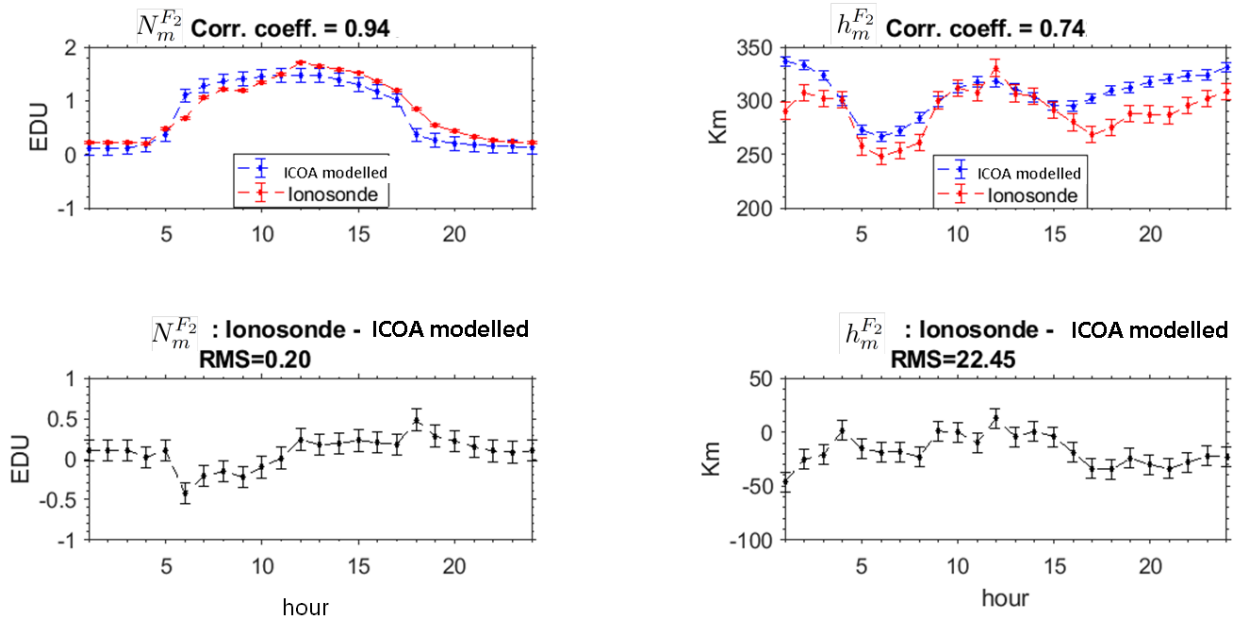


Figure 7.6: Estimated key parameters deviation from ionosonde station EA036.

without any systematic offsets at any one particular station or time of a day. McNamara et al. (2008), Lei and Chuo (2014) (and the references therein), Limberger et al. (2014) and Liang et al. (2015b) also report similar deviations. The statistics of peak density as well as peak height deviation and relative deviation are summarized in Tables 7.4, 7.5, 7.6, 7.7.

Station-wise analysis of the deviation

In this section, a station-specific analysis for the ionosonde stations EA0356 and GR13L are discussed. Generally, as it is seen in Fig. 7.4, the ionosonde captures the local $N_m^{F_2}$ variations, whereas the ICOA estimated $N_m^{F_2}$ follows a relatively smoother transition with time. This effect is due to the dominance of the IRI model though the seperability approach. Furthermore, the ionosonde data have a higher temporal resolution of 10 minutes compared to the 1 hour resolution of the estimated Chapman key parameters. Therefore, within the validation period, the key parameters from both are synchronized in time, by averaging of the ionosonde values and assigning to the one hour time epochs corresponding to the estimated key parameters. Due to the non-uniform sampling of the ionosonde data, the number of samples in each one hour averaging window is different.

For the GR13L station, the maximum value of $\Delta N_m^{F_2}$ is 0.49 ± 0.23 EDU, which is approximately 22% of the peak value, and that for $\Delta h_m^{F_2}$ reads -42.91 ± 7.97 km which is approximately 9% of the peak value, at this location. The mean deviation values for the two key parameters are 0.17 EDU and 14.3 km respectively over the validation period. For the EA036 station, the maximum $\Delta N_m^{F_2}$ is 0.20 ± 0.11 EDU, which is approximately 13% of the peak value, and that for $\Delta h_m^{F_2}$ is -21.33 ± 3.70 km, which is approximately 5% of the peak value at this location. The mean values of deviation for the two key parameters at EA036 are 0.07 EDU and 10.1 km respectively over the validation period. The deviation magnitudes are in similar in the other stations. The average deviations of the peak density and the peak height, during the validation period at various ionosonde stations, are shown in the Figures 7.8 and 7.9 respectively. A comparison of the estimated and ionosonde derived key parameters for stations GR13L, EA036 and AH223 are shown in Figures 7.5, 7.6 and 7.7 respectively. Furthermore, a noticable difference is also seen between the day and the night

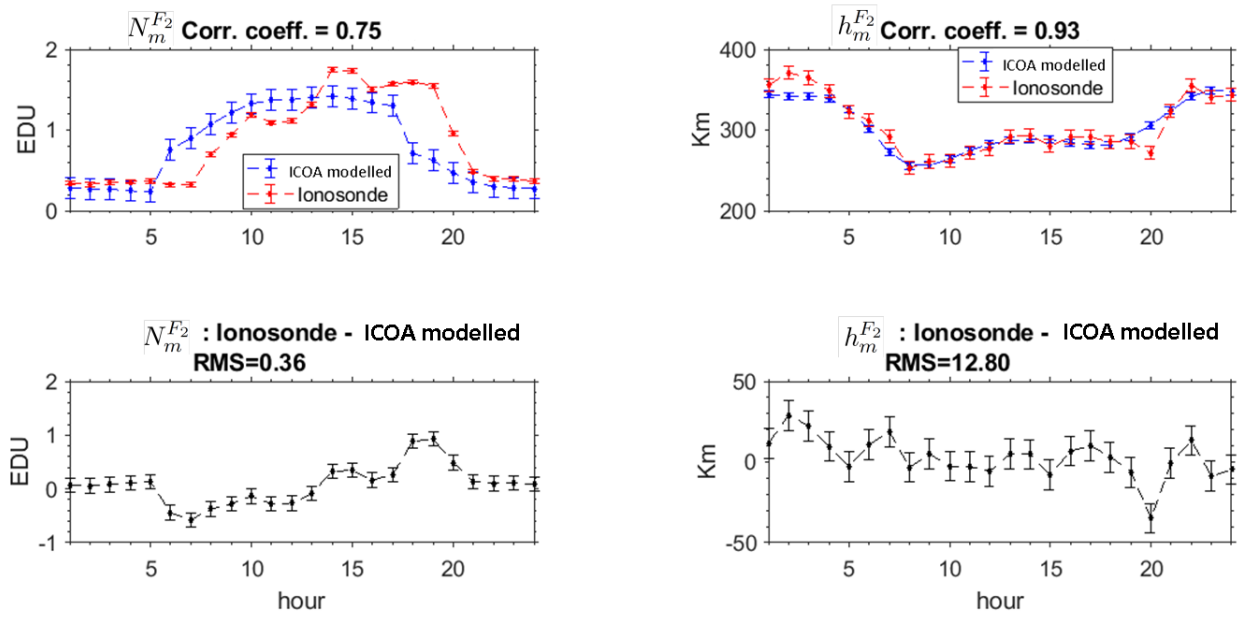


Figure 7.7: Estimated key parameters deviation from ionosonde station AH223

times, in $\Delta N^{F_2}_{m;Ionosonde}$ and $\Delta h^{F_2}_{m;Ionosonde}$. In this particular validation campaign, this diurnal trend can be attributed to a possible limiting accuracy of the given values in $\mathcal{K}_2$, especially at the night times. Similar diurnal patterns in deviations between IRI model and GPS derived TEC values were reported in Endeshaw (2020). Also, Bjoland et al. (2016) mention that the accuracy of the estimated electron density from IRI is higher at the altitudes around the peak height. This justifies the h_{step} = 5 km sampling for altitudes in the range 250 km $\leq h \leq$ 450 km to obtain the profile function for VTEC with separability approach.

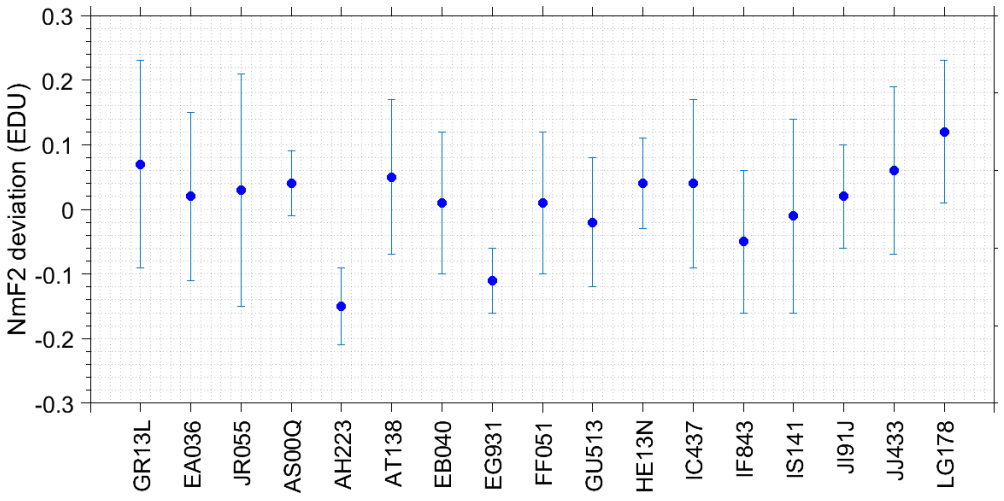


Figure 7.8: Average deviation $\Delta N_m^{F_2}$ during the validation period at different ionosonde stations

The exact magnitude of the Chapman key parameter deviations depends on several factors such as, the time of the day, season of the year, geographical location of ionosonde, the phase of the solar

Table 7.3: Average deviations $\Delta N_m^{F_2}$ and $\Delta h_m^{F_2}$ and the corresponding relative deviations at GR13L station during 12 to 14 March 2015

hour of day	$\Delta N_m^{F_2}$ (EDU)	$\Delta h_m^{F_2}$ (km)	$\Delta_{rel} N_m^{F_2}$ (%)	$\Delta_{rel} h_m^{F_2}$(%)
1	0.10 ± 0.21	-52.73 ± 5.40	13.0	13.3
2	0.05 ± 0.17	-42.70 ± 4.72	8.18	10.95
3	0.05 ± 0.11	-9.16 ± 5.68	6.82	3.30
4	0.00 ± 0.08	-10.96 ± 5.75	3.64	3.57
5	0.02 ± 0.06	-20.62 ± 7.04	3.64	6.43
6	0.18 ± 0.04	-9.14 ± 8.64	10.0	4.05
7	0.28 ± 0.04	-4.47 ± 10.03	14.5	3.33
8	0.32 ± 0.03	-6.11 ± 10.55	15.9	3.8
9	-0.07 ± 0.05	-6.67 ± 11.57	5.45	4.04
10	0.10 ± 0.02	-9.86 ± 11.88	5.41	4.52
11	0.09 ± 0.04	-26.58 ± 11.83	5.90	9.04
12	-0.01 ± 0.05	-15.58 ± 12.02	2.70	6.66
13	0.01 ± 0.02	-8.76 ± 12.18	1.36	4.76
14	0.01 ± 0.05	0.36 ± 11.04	2.72	2.85
15	0.13 ± 0.07	8.18 ± 11.65	10.07	4.68
17	0.40 ± 0.12	-21.12 ± 11.47	23.63	7.85
18	0.06 ± 0.21	-46.22 ± 9.60	12.27	13.09
19	-0.25 ± 0.26	-24.38 ± 7.68	22.72	7.61
20	-0.36 ± 0.28	-21.78 ± 6.32	29.09	6.42
21	0.13 ± 0.36	-30.64 ± 6.18	22.72	8.57
22	0.11 ± 0.27	-38.70 ± 6.29	18.18	10.71
23	0.08 ± 0.25	-19.85 ± 6.20	14.54	6.11
24	0.11 ± 0.22	-26.84 ± 6.13	15.44	7.85

cycle, sunspot number and other modelling assumptions[19] in the ICOA. Furthermore, Hunsucker (1991), McNamara et al. (2008) also mention the effect of diverse ionosonde instrument types, technologies, e.g. digisonde, dynasonde, incoherent scatter radar, to be considered while accounting for the deviations. There are other possible causes for deviation between the estimated and the ionosonde observed Chapman key parameters, namely:

- the B-spline levels used to model the Chapman key parameters
- the differences in temporal resolution of $N_m^{F_2}$, $h_m^{F_2}$ from ionosonde and the ICOA modelled key parameters
- the profile function[20] obtained from a climatological IRI model
- the differences arising from the station-specific ionogram scaling techniques
- the overall data quality of the ionosondes

[19] Such as the tuning parameters that include, step size (see Eq. (4.158)), complimentarity measure (Eq. (4.110)), centering parameter (see Step 6 of the solution of KKT system of equations in Section 4.5.3), the dimensions of the search hypervolume (6.8) etc.

[20] in VTEC with separability approach

Table 7.4: Mean F_2 peak density relative deviation at Grahamstown ionosonde station GR13L (33.3106°S, 26.5256°E) between 12 March and 14 March 2015

	relative deviation $\Delta_{rel}N^{F_2}_{m;Ionosonde}$		
Time window (UT)	12 March (%)	13 March (%)	14 March (%)
0 - 6	6.0	8.0	8.8
6 - 12	10.4	15.2	8.8
12 - 18	11.6	8.0	9.6
18 - 24	9.2	10.0	7.6

Table 7.5: Mean $N^{F_2}_m$ deviation at Grahamstown ionosonde station GR13L (33.3106°S, 26.5256°E) between 12 March and 14 March 2015

	Deviation $\Delta N^{F_2}_{m;Ionosonde}$		
Time window (UT)	12 March (EDU)	13 March (EDU)	14 March (EDU)
0 - 6	0.09 ± 0.06	0.13 ± 0.07	0.16 ± 0.06
6 - 12	0.06 ± 0.20	0.07 ± 0.31	0.04 ± 0.18
12 - 18	0.20 ± 0.09	0.07 ± 0.13	0.08 ± 0.16
18 - 24	-0.01 ± 0.22	-0.01 ± 0.24	-0.01 ± 0.18

- the truncation error of the B-splines representation
- the quality of GNSS STEC and IRO observations[21].

Therefore, it is concluded that the deviations (and hence by definition, also the relative deviation) reported in this section, as well as those summarized in Table 7.2, are within acceptable limits.

7.3 Validation with GNSS STEC

In this section, the STEC measured at two GNSS receiver stations KIRU (67.8558° N, 20.2253° E) and WTZR (49.1452° N, 12.8857° E), of the IGS network, are used for validation of the ICOA

[21]"IRO observations" refer to the electron density profiles obtained from F3C and GRACE.

Table 7.6: Mean $h^{F_2}_m$ deviation at ionosonde station Grahamstown GR13L (33.3106°S, 26.5256°E) between 12 March and 14 March 2015

	Deviation $\Delta h^{F_2}_{m;Ionosonde}$		
Time window (UT)	12 March (km)	13 March (km)	14 March (km)
0 - 6	-29.87 ± 7.66	-19.57 ± 6.87	-21.51 ± 9.56
6 - 12	-10.07 ± 6.13	-8.27 ± 7.33	-9.65 ± 10.40
12 - 18	-7.18 ± 6.98	-6.43 ± 7.10	-3.43 ± 8.50
18 - 24	-27.33 ± 7.31	-24.65 ± 8.19	-22.56 ± 7.08

Table 7.7: Mean $h_m^{F_2}$ relative deviation at ionosonde station Grahamstown GR13L (33.3106°S, 26.5256°E) between 12 March and 14 March 2015

	relative deviation $\Delta_{rel} h^{F_2}_{m;Ionosonde}$		
Time window (UT)	12 March (%)	13 March (%)	14 March (%)
0 - 6	8.5	6.0	7.06
6 - 12	3.68	3.54	4.54
12 - 18	3.21	3.07	2.71
18 - 24	7.87	7.46	6.73

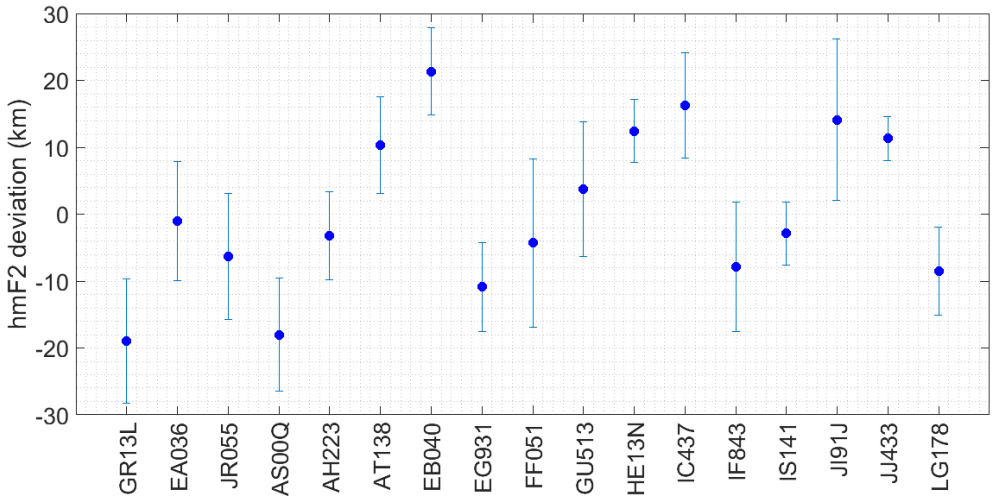


Figure 7.9: Average deviation $\Delta h_m^{F_2}$ during the validation period at differents ionosonde stations

modelled key parameters[22]. The main motivation to use these two stations is due to their higher quality of observations and superior receiver and antenna[23].

We begin with two datasets, namely, (1) the estimated B-spline coefficients and (2) the Swarm STEC. The validation approach will be to systematically transform the former to STEC, so that a comparison with the latter can be made, as shown in Fig. 7.14.

GNSS STEC as observations for validation

The first part of this validation procedure involves fetching the RINEX observation and navigation files (see Gurtner and Estey (2007)) and computing the STEC for each satellite-receiver line-of-sight (LoS) using the dual frequency linear combination of the GPS L1/L2 or L1/L5 and GALILEO E1/E5a or E1/E5b signal pairs. Although STEC can also be computed from single frequency, the use of dual frequency combinations are preferred due to a desired accuracy requirement of STEC in the order of 1-3 TECU. The single frequency approach is called the code-carrier divergence (CCD) method (e.g. Hofmann-Wellenhof et al. (2007)). One of the dual frequency combinations is called the "geometry-free" linear combination (GFLC), see Eq. (A.24a). An additional benefit of GFLC is that the probability of ambiguity fixing increases compared to the CCD method (see Verhagen and Teunissen (2006), Henkel (2010), Teunissen (2017)). This is very relevant for the quality of STEC as well as IRO observations. If only dual frequency phase observations are used for STEC

[22]These are the same Chapman key parameters as those in the set $\mathcal{K}_1$.

[23]KIRU station is managed by ESA-ESOC and in February 2013, a new Spetentrio Polarx4 receiver and Spetentrio chokering antenna were installed to extend the signal teracking capabilities to all four GNSS, namely GPS, Galileo, GLONASS and Beidou. Furthermore, at this latitude the ionospohere behavior is considered to be nominal; see Jakowski (2017). The WTZR station comprises of the JAVAD TRE3 DELTA receiver and LEICA AR antenna.

computation, then the measurement noise is in order of 10-20 cm (Henkel (2010)). However, in this work, both the phase and the pseudorange are used[24].

The second part of the data flow in Fig. 7.19 starts with the estimated B-spline coefficients $\widehat{d}^{J_1,J_2}_{k_1,k_2;\kappa}$, with which ICOA modelled key parameters, $\widehat{\kappa}_{ICOA}\ \forall\ \kappa\ \in \mathcal{K}_1$, are computed using Eq. (5.5). Substituting $\widehat{\kappa}_{ICOA}$, $\forall\ \kappa\ \in \mathcal{K}_1$ as well as $\kappa\ \in \mathcal{K}_2$, into Eqs. (5.1) - (5.2) give the estimated electron density $\widehat{N}_{e;ICOA}(\varphi,\lambda,h)$, where $h\ \in\ \{h_{min}\cdots h_{GNSS}\}$, which upon integrating along the vertical, provides the estimated VTEC,

$$\widehat{VTEC}_{ICOA}(\varphi_W,\lambda_W) = \int_{h_{min}}^{h_{GNSS}} \widehat{N}_{e;ICOA}\,(\varphi_W,\lambda_W,h)\,dh \tag{7.6a}$$

$$\widehat{VTEC}_{ICOA}(\varphi_K,\lambda_K) = \int_{h_{min}}^{h_{max}} \widehat{N}_{e;ICOA}\,(\varphi_K,\lambda_K,h)\,dh, \tag{7.6b}$$

at WTZR and KIRU station (φ_W,λ_W), (φ_K,λ_K), with position vectors $\mathbf{x}_W$, $\mathbf{x}_K$, respectively. Using an isotropic mapping function $M(z)$, corresponding to the zenith angles z^s_W and z^s_K, the modelled STEC, for the set of visible satellites, $s\ \in\{1,\ \cdots,\ smax\}$,

$$\widehat{STEC}_{ICOA}(\mathbf{x}_W,\mathbf{x}^s) = \frac{1}{M(z^s_W)}\widehat{VTEC}_{ICOA}(\varphi_W,\lambda_W) \tag{7.7a}$$

$$\widehat{STEC}_{ICOA}(\mathbf{x}_K,\mathbf{x}^s) = \frac{1}{M(z^s_K)}\widehat{VTEC}_{ICOA}(\varphi_K,\lambda_K) \tag{7.7b}$$

from the two stations are obtained, where $\mathbf{x}^s$ denotes the s^{th} GNSS satellite position vector. Substituting L1, L5 phase observations at KIRU, as well as WTZR, in Eqs. (A.33a) and (A.33b), as well as removing the s^{th} satellite and KIRU/WTZR receiver bias, gives $\widehat{STEC}(\mathbf{x}_W,\mathbf{x}^s)$ and $\widehat{STEC}(\mathbf{x}_K,\mathbf{x}^s)$, from which the STEC deviation

$$\Delta STEC_W(\mathbf{x}_W,\mathbf{x}^s) = \widehat{STEC}(\mathbf{x}_W,\mathbf{x}^s) - \widehat{STEC}_{ICOA}(\mathbf{x}_W,\mathbf{x}^s) \tag{7.8a}$$

$$\Delta STEC_K(\mathbf{x}_K,\mathbf{x}^s) = \widehat{STEC}(\mathbf{x}_K,\mathbf{x}^s) - \widehat{STEC}_{ICOA}(\mathbf{x}_K,\mathbf{x}^s) \tag{7.8b}$$

are computed. These are shown in the Figures 7.12 and 7.13 for selected GPS and Galileo satellites as observed from KIRU, and the same for WTZR in the Figures 7.10 and 7.11, respectively. The RMS STEC deviation of GPS L1C-L2W dual frequency phase combination, denoted by $\Delta STEC_{RMS,L1C-L2W}$, is in the order of 1 - 3 TECU, for all visible GPS satellites for the validation period. For the sake of completeness, the corresponding STEC deviations from GPS C1C and L1C single frequency, denoted by $\Delta STEC_{RMS,C1C-L1C}$, are also shown along side in Fig. 7.12. The relatively larger values of $\Delta STEC_{RMS,C1C-L1C}$ compared to $\Delta STEC_{RMS,L1C-L2W}$, show the superiority of the dual frequency combination in the validation, as already described in Section D.2. Similarly, $\Delta STEC_{RMS,L1C-L5Q}$ and $\Delta STEC_{RMS,L1C-L7Q}$ denote the corresponding performance with Galileo dual frequency, which is also in the order of 1 - 3 TECU, across the four satellites[25], as shown in Fig. 7.13.

The decision for number of GNSS receivers in EDM and validation are driven by computational complexity in the overall processing. Specifically, the main computational challenge with GNSS STEC observations, is the Gauss-Legendre quadrature integrations (see Limberger (2015)) performed for each satellite-receiver line of sight. Other considerations include the need for the estimation of additional unknown parameters[26], e.g. DCBs and the processing overhead from ambiguity resolution (see Section D.2). Although DCBs are not the primary parameters in electron density modelling,

[24] Pseudoranges are needed for carrier phase ambiguity resolution; see e.g. Hofmann-Wellenhof et al. (2012)

[25] In March 2015 time frame, there were only four operational Galileo satellites GSAT101, GSAT102, GSAT103 and GSAT104 with PRN 11, 12 19 and 20 respectively.

[26] If there are a total of s satellites and r receivers, then we have $s + r + s \cdot r$ additional unknown parameters.

Table 7.8: Summary of daily RMS STEC deviation with GPS L1-L2

Station	$\Delta STEC_{RMS,L1C-L2W}$ (TECU)			$\Delta STEC_{RMS,C1C-C2W}$ (TECU)		
	12-03	13-03	14-03	12-03	13-03	14-03
WTZR	1.82	2.01	1.73	2.22	2.61	2.31
KIRU	2.83	2.64	2.89	3.77	3.50	3.58

Table 7.9: Summary of daily RMS STEC deviation with Galileo E1-E5a, E1-E5b

Station	$\Delta STEC_{RMS,L1C-L5X}$ (TECU)			$\Delta STEC_{RMS,L1C-L7Q}$ (TECU)		
	12-03	13-03	14-03	12-03	13-03	14-03
WTZR	1.41	1.35	1.47	1.02	0.97	1.14
KIRU	1.23	1.08	1.26	1.09	1.00	1.19

but they still need to be estimated for the sake of consistency. Additionally, the phase biases are also relevant, especially when phase observations are used for STEC computation. The phase biases are also assumed to be reliably estimated and any unmodelled phase error is assumed to be absorbed by the float-ambiguities of the carrier phase. The GNSS data processing follows the methodology presented in Erdogan et al. (2017). The statistics from the GNSS STEC validation are summarized for GPS and Galileo in the Tables 7.8 and 7.9 respectively. The smaller absolute magnitudes of the deviation and relative deviation for STEC computed from Galileo is at least in part due to the lower measurement noise and thereby improved integer ambiguity resolution.

As shown in Section D.2, STEC computation requires understanding of GNSS observables on each signal type (e.g. C or P codes) per frequency $f \in \{L1,\ L2\ or\ L5\}$. A detailed description of the GNSS observable types (e.g. C1C, C2W) can be found in the RINEX 3.0 standard (Gurtner and Estey (2007)). There is one another important relation between DCB and VTEC modelling. Each global VTEC map (GIM) in the IONEX format contains a header with the receiver DCBs. The schematic dataflow in the STEC based validation is summarized in Fig. 7.19.

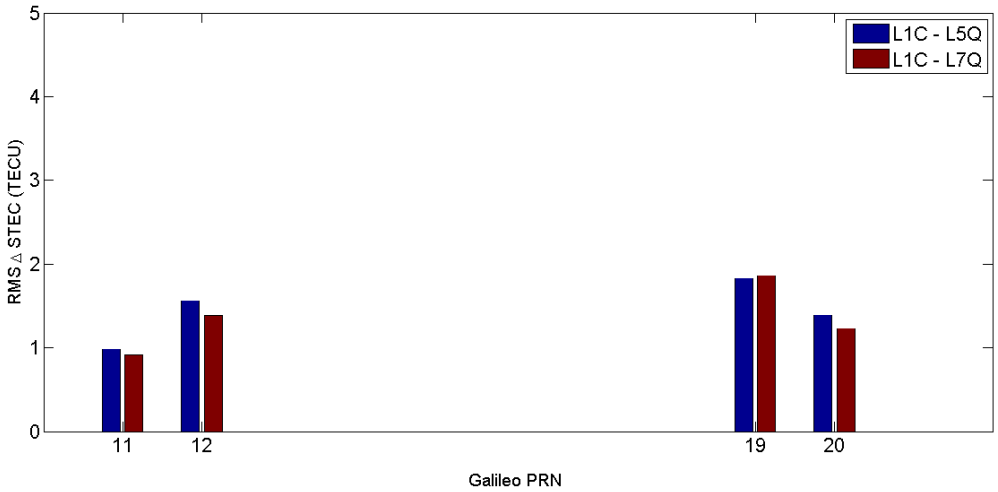


Figure 7.10: GPS ΔSTEC RMS at WTZR over the validation period between 12-14 March 2015.

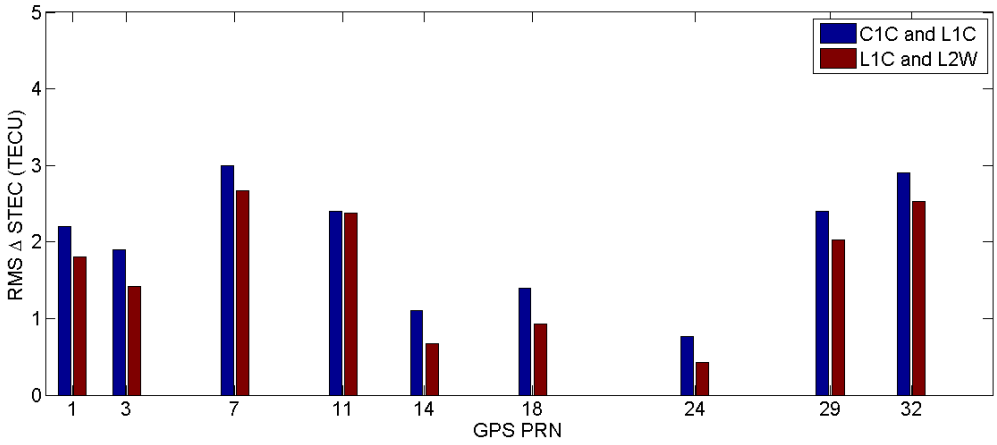


Figure 7.11: Galileo ΔSTEC RMS at WTZR over the validation period 12 to 14 March 2015.

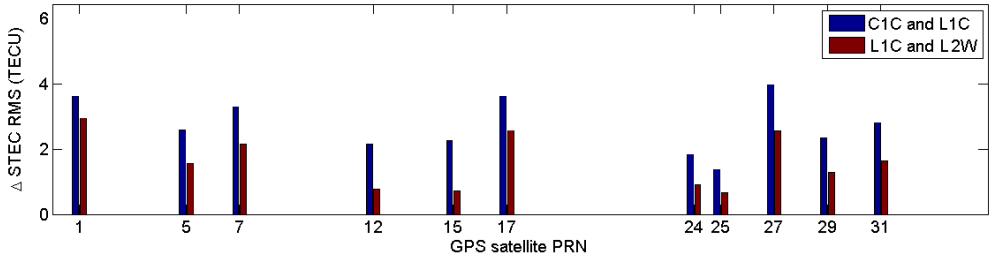


Figure 7.12: GPS ΔSTEC RMS at KIRU over the validation period between 12-14 March 2015.

7.4 In-situ observations based validation

An introduction of in-situ observations was already provided in Section 3.3. In this section, observations from Langmuir probe onboard CHAMP and Swarm will be used to validate the modelled electron density.

7.4.1 Validation with CHAMP, Swarm Langmuir probe

The Langmuir[27] probe (LP) is a scientific instrument used to determine plasma temperature, and density in the upper regions of the Earth's atmosphere (e.g. ionosphere). CHAMP and Swarm satellites are equipped with LP and their observations $N_{e;CLP}$ are obtained directly from CHAMP mission portal (see Reigber et al. (2002), Cooke et al. (2003)) between the time period 01 - 03 August 2007, during which the number of observations used for EDM are shown in Fig. 7.15. The computed CHAMP electron density deviation

$$\Delta N_{e;CLP}(\varphi_C, \lambda_C, h_C) = N_{e;CLP}(\varphi_C, \lambda_C, h_C) - \widehat{N}_{e;ICOA}(\varphi_C, \lambda_C, h_C) \tag{7.9}$$

are shown in Fig. 7.16. A clear trend along the lower latitudes can be seen. This aspect can be attributed to both LP instrumental biases as well as the B-spline levels used for modelling the key parameters $\kappa \in \mathcal{K}_1$.

Procedure for computing $\widehat{N}_{e;ICOA}$

[27] Irving Langmuir won the Nobel prize in Chemistry and made contributions in different fields of Physics during the 20^{th} century (Boyd (1965), Hopwood et al. (1993))

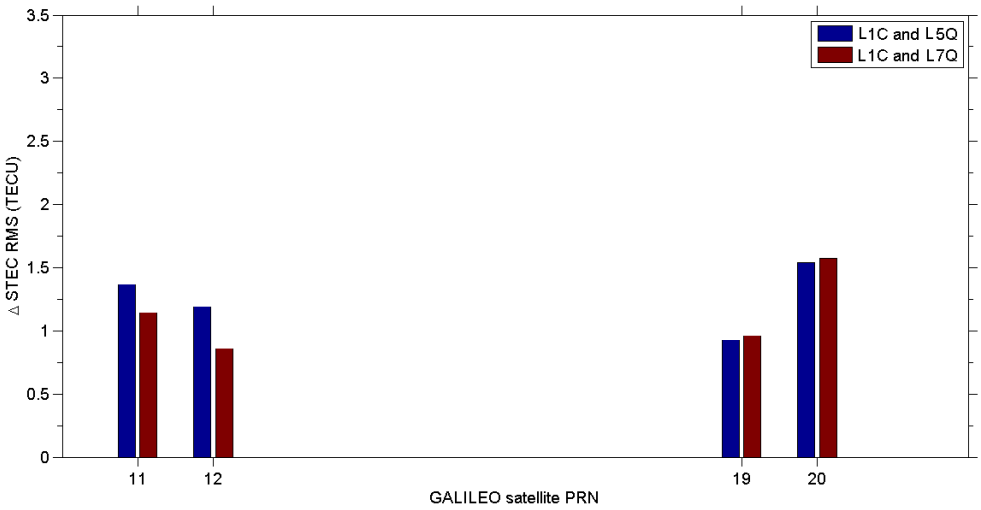


Figure 7.13: GALILEO ΔSTEC RMS at KIRU over the validation period between 12-14 March 2015.

The estimated B-spline coefficients $\widehat{d}^{J_1,J_2}_{k_1,k_2;\kappa}$ are first denoted as vector $\widehat{\boldsymbol{d}}_\kappa$ and substituted in Eq. (5.5) to yield the modelled key parameters vector

$$\widehat{\boldsymbol{\kappa}}_{ICOA} = \widehat{\boldsymbol{d}}_\kappa \boldsymbol{\Psi}_{NT}, \tag{7.10}$$

where $\boldsymbol{\Psi}_{NT}$ denotes the matrix of tensor products as described in Section 5.4 and the vector of modelled key parameter,

$$\widehat{\boldsymbol{\kappa}}_{ICOA} = \begin{bmatrix} \widehat{\boldsymbol{N}}^{F_2}_{m;ICOA} \\ \widehat{\boldsymbol{h}}^{F_2}_{m;ICOA} \\ \widehat{\boldsymbol{H}}^{F_2}_{ICOA} \\ \widehat{\boldsymbol{N}}^{P}_{0;ICOA} \\ \widehat{\boldsymbol{H}}^{P}_{ICOA} \end{bmatrix} \tag{7.11}$$

contains the individual Chapman key parameter vectors corresponding to $\kappa \in \mathcal{K}_1$. We apply Eq. (5.1) to compute the modelled electron density

$$\begin{aligned} \widehat{N}_{e;ICOA}(\varphi_C, \lambda_C, h_C) = \widehat{N}^{F_2}_e(\varphi_C, \lambda_C, h_C) + \widehat{N}^{F_1}_e(\varphi_C, \lambda_C, h_C) + \widehat{N}^{E}_e(\varphi_C, \lambda_C, h_C) \\ + \widehat{N}^{D}_e(\varphi_C, \lambda_C, h_C) + \widehat{N}^{P}_e(\varphi_C, \lambda_C, h_C), \end{aligned} \tag{7.12}$$

where the electron density of the individual F_2-, F_1-, $E-$, $D-$ layer as well as plasmasphere are computed as shown in Equations (7.12) - (7.13).

$$\begin{aligned} &\widehat{N}^{F_2}_e(\varphi_C, \lambda_C, h_C) = \widehat{N}^{F_2}_{m;ICOA}(\varphi_C, \lambda_C, h_C) \\ &\exp\left(\frac{1}{2}\left(1 - \frac{h_C - \widehat{h}^{F_2}_{m;ICOA}(\varphi_C, \lambda_C, h_C)}{\widehat{H}^{F_2}_{ICOA}(\varphi_C, \lambda_C, h_C)} - \exp\left(-\frac{h_C - \widehat{h}^{F_2}_{m;ICOA}(\varphi_C, \lambda_C, h_C)}{\widehat{H}^{F_2}_{ICOA}(\varphi_C, \lambda_C, h_C)}\right)\right)\right) \end{aligned} \tag{7.13}$$

$$\widehat{N}^{P}_e(\varphi_C, \lambda_C, h_C) = \widehat{N}^{P}_{0;ICOA}(\varphi_C, \lambda_C, h_C) \exp\left(-\left|\frac{h_C - \widehat{h}^{F_2}_{m;ICOA}(\varphi_C, \lambda_C, h_C)}{\widehat{H}^{P}_{ICOA}(\varphi_C, \lambda_C, h_C)}\right|\right) \tag{7.14}$$

$$\begin{aligned} &\widehat{N}^{F_1}_e(\varphi_C, \lambda_C, h_C) = N^{F_1}_m(\varphi_C, \lambda_C, h_C) \\ &\exp\left(\frac{1}{2}\left(1 - \frac{h_C - \widehat{h}^{F_1}_m(\varphi_C, \lambda_C, h_C)}{\widehat{H}^{F_1}(\varphi_C, \lambda_C, h_C)} - \exp\left(-\frac{h_C - \widehat{h}^{F_1}_m(\varphi_C, \lambda_C, h_C)}{\widehat{H}^{F_1}(\varphi_C, \lambda_C, h_C)}\right)\right)\right) \end{aligned} \tag{7.15}$$

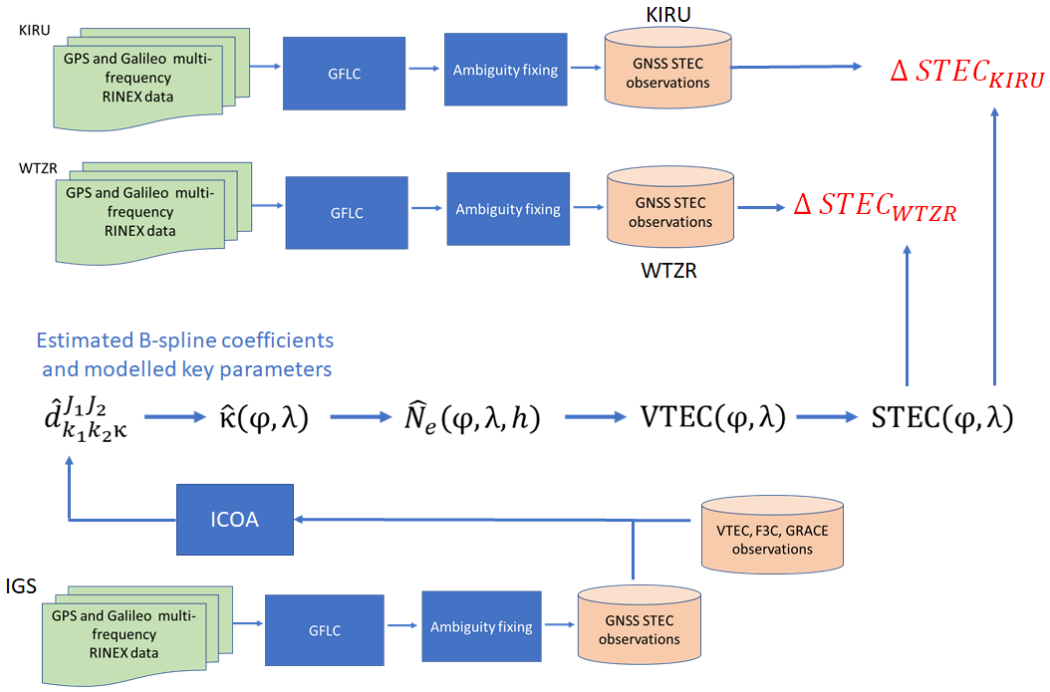


Figure 7.14: Dataflow of validation with GNSS STEC measured at KIRU and WTZR IGS receivers

$$\widehat{N}_e^E(\varphi_C,\lambda_C,h_C) = \widehat{N}_m^E(\varphi_C,\lambda_C,h_C) \exp\left(\frac{1}{2}\left(1 - \frac{h_C - \widehat{h}_m^E(\varphi_C,\lambda_C,h_C)}{\widehat{H}^E(\varphi_C,\lambda_C,h_C)} - \exp\left(-\frac{h_C - \widehat{h}_m^E(\varphi_C,\lambda_C,h_C)}{\widehat{H}^E(\varphi_C,\lambda_C,h_C)}\right)\right)\right) \tag{7.16}$$

$$\widehat{N}_e^D(\varphi_C,\lambda_C,h_C) = \widehat{N}_m^D(\varphi_C,\lambda_C,h_C) \exp\left(\frac{1}{2}\left(1 - \frac{h_C - \widehat{h}_m^D(\varphi_C,\lambda_C,h_C)}{\widehat{H}^D(\varphi_C,\lambda_C,h_C)} - \exp\left(-\frac{h_C - \widehat{h}_m^D(\varphi_C,\lambda_C,h_C)}{\widehat{H}^D(\varphi_C,\lambda_C,h_C)}\right)\right)\right) \tag{7.17}$$

Furthermore, Swarm Langmuir probe observations $N_{e;SLP}$ are obtained directly from Swarm mission data portal for the time period of 12 - 14 March 2015. Similar to the modelled electron density at CHAMP coordinates as described above, that at Swarm coordinates $\widehat{N}_{e;ICOA}(\varphi_S,\lambda_S,h_S)$ is determined from Equations (7.12) - (7.13). Accordingly, Swarm electron density deviation

$$\Delta N_{e;SLP}(\varphi_S,\lambda_S,h_S) = N_{e;SLP}(\varphi_S,\lambda_S,h_S) - \widehat{N}_{e;ICOA}(\varphi_S,\lambda_S,h_S) \tag{7.18}$$

with respect to Swarm LP observations as shown in Figures 7.17 and 7.18.

Both $\Delta N_{e;CLP}$ and $\Delta N_{e;SLP}$ depend on the quality of the respective LP observations as well as the modelled electron density. In-particular, the latter is dominated by the quality of the modelled F_2 peak density. The F_2 peak height and scale height are related to electron density through the double exponential functional relation (5.2), and therefore have a relatively lesser impact on the quality of $\widehat{N}_{e;ICOA}$.

7.4.2 Validation with Swarm STEC

In this section, STEC measured by GNSS receiver onboard the Swarm satellite are used for the validation of modelled key parameters. As the procedure for STEC computation has been already described before (see Sections D.2 and 7.3), those will not be repeated.

Once again, we begin with the estimated B-spline coefficients $\widehat{d}^{J_1,J_2}_{k_1,k_2;\kappa}$ and substituting those in Eq. (5.5), corresponding to the Swarm satellite coordinates (consistent with its position vector $\mathbf{x}_r$),

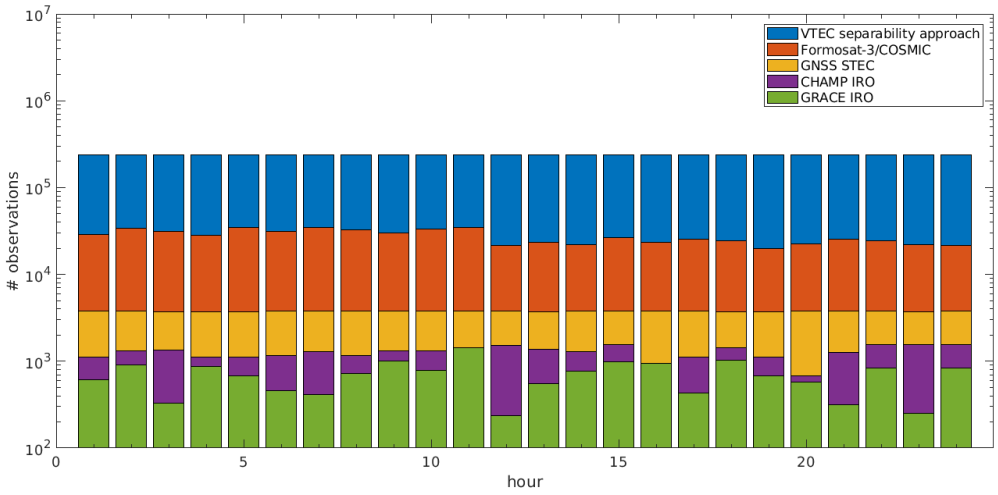


Figure 7.15: Average observations from different techniques between 01-03 August 2007. The time period of August 2007 is selected because F3C mission had more active satellites during this period as well as CHAMP was active. Therefore, it is interesting to see the relative number of electron density observations compared to the period of March 2015, for which the data distribution was shown in Chapter 3.

to obtain the modelled Chapman key parameters vector $\widehat{\boldsymbol{\kappa}}_{ICOA}$. The modelled electron density $\widehat{N}_{e;ICOA}(\varphi_S, \lambda_S, h_S)$ thus computed, are integrated as shown in Equations (7.6a)-(7.7b) to obtain $VTEC_{ICOA}$ an $STEC_{ICOA}$. With the above information, we are able to compute the Swarm STEC deviation

$$\Delta STEC_{Swarm}(\mathbf{x}_r, \mathbf{x}^s) = \widetilde{STEC}_{Swarm}(\mathbf{x}_r, \mathbf{x}^s) - STEC_{ICOA}(\mathbf{x}_r, \mathbf{x}^s), \tag{7.19}$$

computed at the Swarm position vector $\mathbf{x}_r$. The summary of validation is shown in Figures 7.10 and 7.11 for GPS and Galileo respectively.

Table 7.10: Summary of RMS deviation and RMS relative deviation between Swarm A,B,C and modelled electron density $\widehat{N}$ as well as $STEC_{ICOA}$ between 12-14 March 2015.

Satellite	RMS deviation Ne (EDU)	RMS deviation STEC (TECU)	RMS rel. deviation Ne (%)	RMS rel. deviation STEC (%)
Swarm-A	0.58	0.88	26.26	5.86
Swarm-B	0.31	0.67	14.04	4.46
Swarm-C	0.27	0.41	12.27	2.73

7.5 Summary of validation

We have reached the end of this chapter to summarize the most important points including on the validation of the modelled electron density. It has also been demonstrated that the estimated B-spline coefficients are systematically transformed to modelled Chapman key parameters and eventually to modelled electron density and/or modelled STEC.

In CLV, the electron densities of the $D-$, $E-$ and F_1-layers are computed using approximate values of the corresponding peak densities, peak height and scale heights. More specifically the $N_m^{F_1}$, $h_m^{F_1}$, N_m^E, h_m^E, N_m^D, h_m^D were taken from IRI and the scale height H^{F_1}, H^E, H^D are assumed to be given. These values are only approximate and may not reflect the real conditions observed at the ionosonde

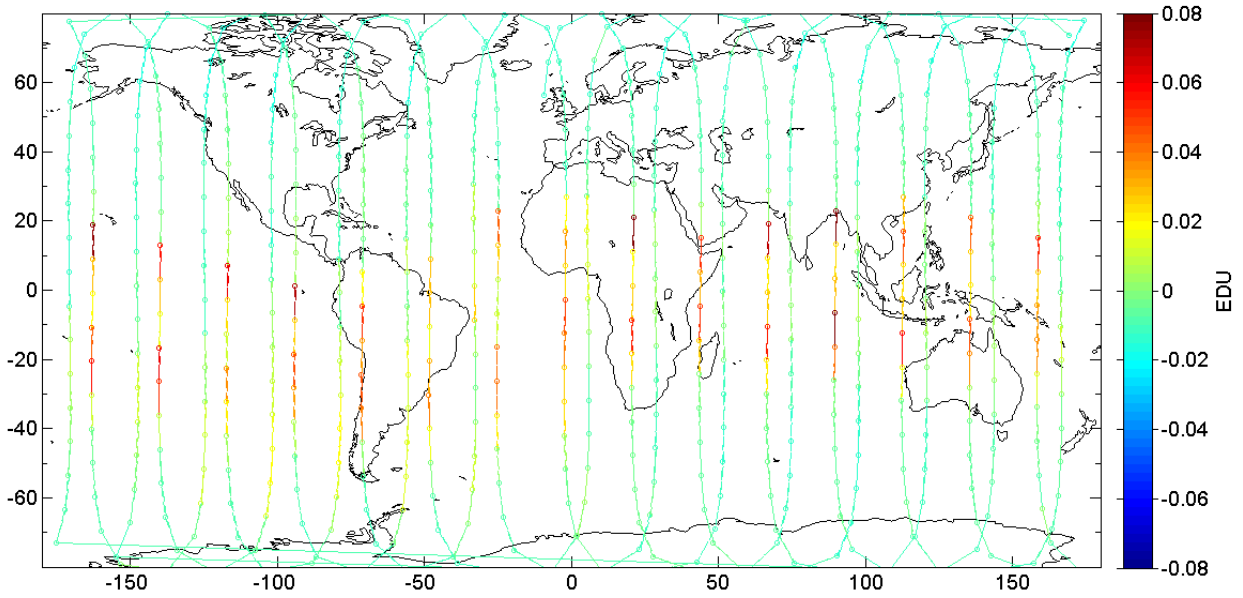


Figure 7.16: Comparison of CHAMP in-situ electron density observations from Langmuir probe on 1-Aug. 2007 with that from estimated B-spline coefficients of ionospheric key parameters. Data courtesy of LP GFZ Potsdam

station location. Furthermore, a constant scale height from 100 to 1000 km globally in the Chapman profile function (5.2) is assumed. In other words, H^{F_2}, H^{F_1}, H^E, H^D and H^P are assumed to be constant with altitude. This assumption may also need to be investigated in future work by considering altitude dependent scale height[28]. However, it is not trivial to fix the scale height prior values exactly in a modelling problem due to the dynamical temporal variations of electron density with altitude.

Ionosonde observations were used to validate the modelled electron density and an agreement in the order of $\approx 10\%$ has been determined for both the F_2 peak density as well as peak height. The absence of other ionosphere layers, such as the sporadic $E-$layer, see Arras et al. (2009) and Tsai and Su (2018), the modelling assumptions for a night-time ionosphere not containing the F_1 layer and a modelling limitation on the 1000 km upper altitude also lead to systematic deviations in the ionosonde based validation. Furthermore, an temporal interpolation on the modelled key parameters was necessary for the comparison with ionosonde. No other calibration was performed in this work on the ionosonde observations.

A detailed description of GNSS STEC based validation is provided using two IGS receivers KIRU and WTZR. The deviation and relative deviation computed with this technique has been summarized in Tables 7.8 and 7.9. the dataflow of GNSS STEC validation is shown in Figures 7.10 and 7.11. The main drawback of this approach is the need to determine the additional satellite-receiver biases which are directly correlated with STEC and therefore cannot be ignored. Furthermore the isotropic mapping function is at best an assumption of homogenous spatially distributed electron density. Despite these factor, an agreement to 1 - 3 TECU was found between GNSS and the modelled STEC.

Furthermore STEC measured by the Swarm GNSS receiver was used to validate the electron density. As already discussed, the main drawback of this approach was that only ionosphere above Swarm

[28] In this work, the Chapman scale height has been used which is constant with altitude at a given latitude and longitude.

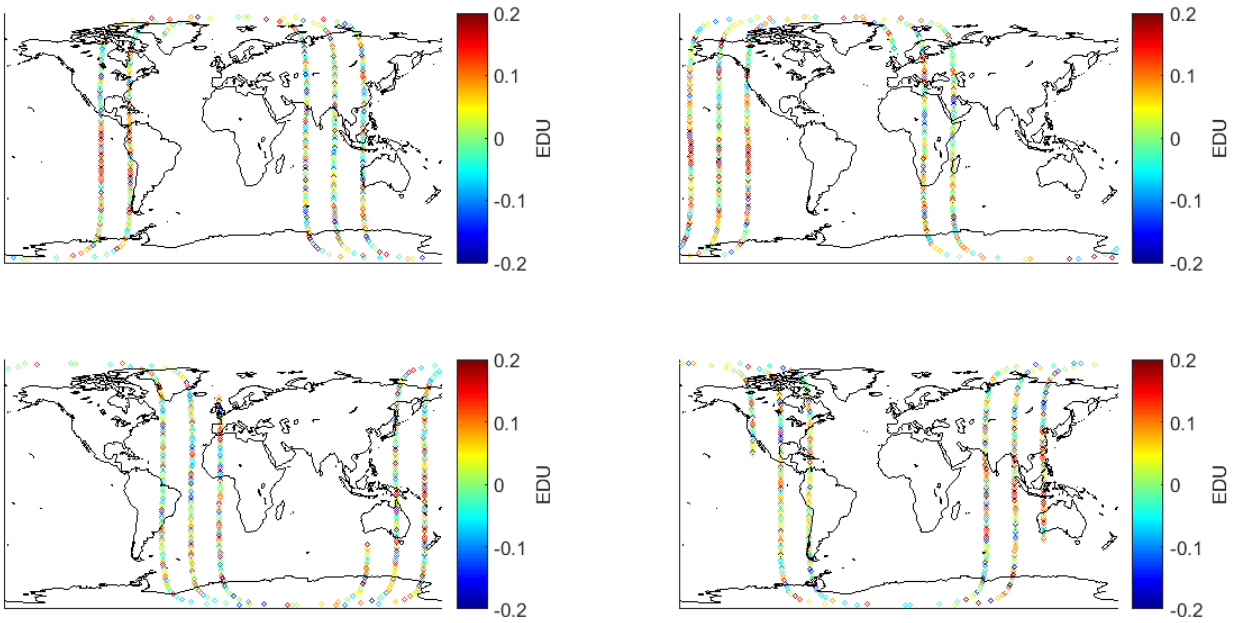


Figure 7.17: Deviation between in-situ and modelled electron density along track Swarm-A on 08 March 2015. Top left 00:00 - 04:00 UT, Top right 04:00 - 08:00 UT, Bottom left 08:00 - 12:00 UT, Bottom left 12:00 - 16:00 UT

altitude would be used and accordingly leads to a systematic offset (approximately 1 TECU), which was removed in the RMS computation as reported in Table 7.11.

A summary of statistics from the validation performed in this section is provided in Table 7.11. Validation has provided confidence that the ICOA has been successfully implemented and used for EDM.

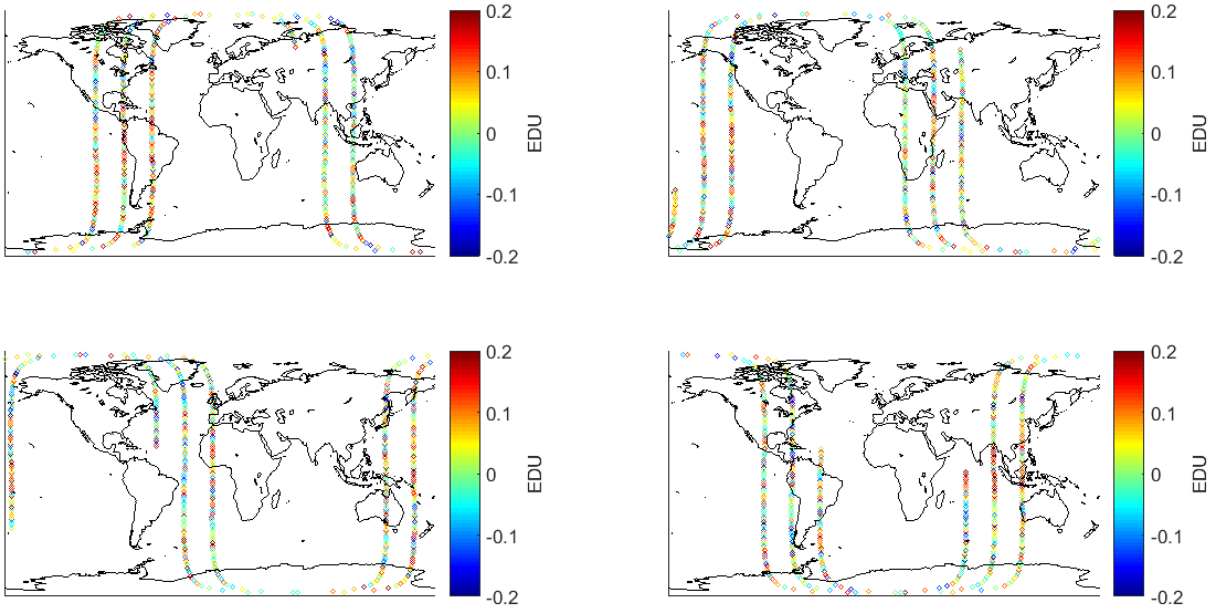


Figure 7.18: Deviation between in-situ and modelled electron density along track Swarm-C on 08 March 2015. Top left 00:00 - 04:00 UT, Top right 04:00 - 08:00 UT, Bottom left 08:00 - 12:00 UT, Bottom left 12:00 - 16:00 UT

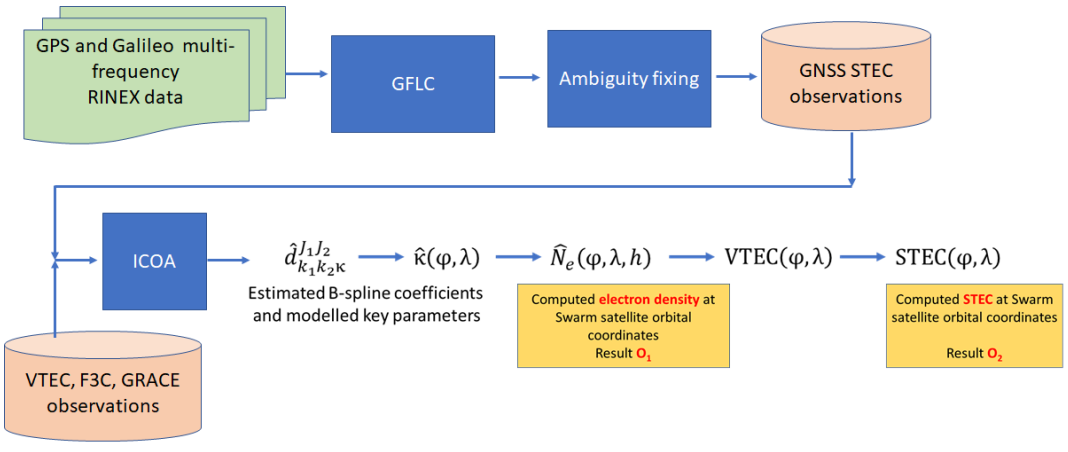


Figure 7.19: Dataflow of validation with Swarm STEC

Table 7.11: Overview of validation

Validation with	Num	Data resolution	Summary
Ionosonde	17 stations	30 min	RMS $\Delta N_m^{F_2}$ = 0.2 EDU RMS $\Delta h_m^{F_2}$ = 22 km RMS rel. $\Delta N_m^{F_2}$ = 9.8 % RMS rel. $\Delta h_m^{F_2}$ = 5 %
GNSS STEC	2 stations	10 min	RMS ΔSTEC = 0.83 TECU RMS rel. ΔSTEC = 5.53 %
Swarm STEC	A,B,C	10 min	RMS ΔSTEC = 0.83 TECU RMS rel. ΔSTEC = 5.53 % Bias = 4.8, = 2.3 and = 1.1 TECU
Swarm LP	A,B,C	10 min	RMS ΔN_e = 0.33 EDU RMS rel. ΔN_e = 15 % Bias = 1.6 EDU
CHAMP LP		10 min	RMS ΔN_e = 0.27 EDU RMS rel. ΔN_e = 12.27 % Bias = 0.31 EDU
Simulated data	–	1 hour	RMS $\Delta N_e < 10^{-3}$ EDU RMS rel. $\Delta N_e < 10^{-3}$ %

8 Summary conclusion and outlook

We have reached the end of this thesis and in this chapter we will summarize the most important outcomes as well as the outlook for the future of EDM using ICOA.

8.1 Summary

Electron density is the most important geodetic parameter in modelling the upper atmosphere, since it has a significant impact on many applications such as PPP and satellite navigation. Therefore, its precise modelling is essential for commercial as well as scientific applications including space weather studies. However, a limited number of observation techniques with a global coverage, especially along the altitude and the mutual correlation among the Chapman key parameters are the two main limitations. Hence there is a clear need to develop robust estimation techniques for the 4D global EDM. As a result, the ICOA has been developed in this thesis. The multi-layer Chapman approach has been used, where the electron density profiles of the different layers are defined by a set of altogether 14 key parameters. Depending on the chosen layers, their key parameters are then modelled using series expansions in terms of 2D tensor products of B-spline functions; the corresponding coefficients are then estimated using ICOA. This approach allows for modelling and estimating multiple Chapman key parameters simultaneously, thereby considering their mutual correlations. Furthermore, inequality constraints are applied to physically localize the selected key parameters to within physically realistic bounds. An independent validation with both the simulated and the ionosonde data was performed to assess the overall accuracy of ICOA.

Summary of findings

Some important findings from our investigations are summarized as follows:

- The developed ICOA described in the Chapter 4 has been successfully applied to our 4D global EDM.
- The relative deviation of the estimated F_2-layer peak height and peak density are approximately in the order of 5% and 10% respectively with respect to ionosonde observations. A summary of the validation results from other sources of data was shown in Table 7.11.
- Even though we have applied the ICOA to 4D global EDM, it can be adapted to regional EDM as well as to related applications e.g. global VTEC modelling.
- The main findings from the 54 numerical evaluations is that the estimated optimization parameters are generally sensitive to the choice of initial values $\boldsymbol{\beta}_0$ from being within or outside the feasible region $\mathcal{F}_R$, the total number as well as the spatial distribution and quality of the observations, the chosen lower and upper bound functions κ_l and κ_u as well as on the selected B-spline levels J_1, J_2 for the Chapman key parameter $\kappa \in \mathcal{K}_1$.
- Particularly, within the ICOA, the initial values $\boldsymbol{\beta}_0$ and the step size matrix $\boldsymbol{T}$ have a significant impact on the estimated optimization parameters. Therefore, a systematic procedure is required to obtain both $\boldsymbol{\beta}_0$ and $\boldsymbol{T}$. Whereas the initial values $\boldsymbol{\beta}_0$ are determined using an ARMA or N-ARMAX process, the step size matrix $\boldsymbol{T}$ was computed using the predicted value of the optimization parameter.
- The choice of the lower and the upper bound functions κ_l and κ_u are of particular importance when the key parameters from two or more ionosphere layers have to be considered in the set

$\mathcal{K}_1$. The bound functions are generally time varying and numerically in different orders of magnitude. As an example, while the peak density of the F_1-layer drops to zero at night, that of the F_2-layer and the plasmasphere still have a significant value; see e.g. Bremer (1998) and Berkner and Wells (1934). Therefore, if the F_1-layer key parameters are considered in the set $\mathcal{K}_1$, their corresponding constraint bounds need to be adapted with time. Accordingly, a balance between extremely restrictive or more relaxed constraints for both the upper and the lower bound functions is needed. Qualitatively, restrictive constraints are those characterized by a relatively smaller dynamic range and relaxed constraints have larger dynamic range. These aspects have been described using examples in Appendix A especially see Eq. (A.4).

- While the inequality constraint bounds on the key parameters can be held constant, we have shown that the bounds can also be defined as spatio-temporal functions; see Chapter 6; Figs. 14 and 15.
- We have also provided the uncertainty of the estimated key parameters using a Monte-Carlo approach.

8.2 Conclusion

By applying the ICOA to the 4D global EDM we have demonstrated that the estimated parameters lie within physically realistic limits, for example, non-negative values for the peak electron density and within an interval of 220 to 420 km the F_2 layer peak height.

The following conclusions are based on a total of 54 numerical evaluations. Furthermore, these conclusions can be directly applied as "best practices" in the subsequent studies so that the benefits of ICOA can be applied to real-time and near real-time ionosphere modelling activities.

1. A numerical value for the minimal set of the unknown Chapman key parameters $\kappa \in \mathcal{K}_1$ shall be estimated and the remaining key parameters $\kappa \in \mathcal{K}_2$ are assumed to be given with appropriate accuracy.
2. Limiting the total number of constraints: For each inequality constraint, an additional pair of unknown Lagrange multiplier and a slack variable has to be estimated. In this thesis, the upper and lower bound functions κ_u and κ_l are considered to be time-varying at 6 hour temporal resolution and are defined at a $5^\circ \times 5^\circ$ spatial resolution along latitude and longitude.
3. The choice of the feasible region: The dynamic range and the hypervolume in an ICOP is directly a function of the constraint bounds, which in-turn is related to the overall size of the feasible region $\mathcal{F}_R$. This problem is further related to two important aspects of the ICOP: (1) the choice of the constraint bound functions κ_l and κ_u and (2) the convexity of the objective function.
   a) To the best knowledge of the author, there is no specified procedure for the choice of the inequality constraint bound functions. Therefore in this work, the upper and the lower bound functions κ_u and κ_l were chosen based on the analysis of the maximum and minimum values of the Chapman key parameters from the IRI model values[1].
   b) A convexity check of the objective function is also recommended and has been used in this thesis; see Eq. (4.28)
4. Initial values: When the ICOA starts with initial values from within the feasible region, then the computation times are relatively lower compared to a completely random initialization.
5. In a sequential processing, the use of an effective forecast technique to determine the initial value of the unknown parameters based on Eqs. (4.171), (4.173) and (4.174) also helps in reducing the computation time.
6. Tuning parameters of the optimization algorithm: The step size matrix $\boldsymbol{T}$ in the search algorithm shall be chosen related to the number of ionosphere layers, the total number of

[1] A global IRI model simulation was analysed over a 10 year period (2005 - 2015) at 12 UT each day.

unknown parameters, and also the total number of constraints. We recommend the use of Eq. (4.109) or one of the methods described in Section 4.7.3.

7. Use of different B-spline levels for the different Chapman key parameters: Some of the Chapman key parameters are not spatially varying as much as some others. For example if the F_2- and the E-layers are simultaneously considered in $\mathcal{K}_1$, then the key parameters of the latter can be modelled with smaller B-spline levels J_1, J_2 compared to that of the F_2-layer. This is owing to the fact that the E-layer spatial decorrelation of the key parameters is smaller compared to that of the F_2-layer. This was checked in the IRI model during this thesis.

8.3 Outlook

While it is necessary to use VTEC data for EDM, its relatively higher number of observations compared to that of F3C and GRACE is due to a larger number of discrete altitudes used for defining the profile function. This situation is expected to change with more observations from the F7C2 mission[2] (see Fig. 3.4) and the availability of additional IRO observations, e.g. from the GRACE-FO mission as well as the satellites owned and operated by other private companies, e.g. SPIRE; see Bowler (2020). The F7C2, a constellation of 6 satellites to meet the RO data continuity (Yue et al. (2014)), is a follow-on mission to the F3C mission. An improved payload TriG GNSS-RO receiver with Galileo and GLONASS tracking capability, will produce a significantly higher spatial and temporal density of electron density profiles. The F7C2 constellation comprises 6 satellites at a 24^o inclination to enhance observations in the equatorial region (Hsu et al. (2018), Yue et al. (2014)). In future, the multi-constellation GNSS tracked by occultation sensors on-board LEO satellites will increase the number of observations, global coverage and will be a significant benefit for the 4D global EDM. The last of the F3C mission satellite was officially decommissioned on 1 May 2020[3].

In future, further investigations could be carried out to charecterize the performance of the ICOA for EDM with modelling different sets of key parameters $\kappa \in \mathcal{K}_1$, especially including the bottom layers. Furthermore, evaluation of our EDM with different B-spline levels J_1, J_2 for the individual key parameters $\kappa \in \mathcal{K}_1$ can also be performed. Since the scope of this thesis is the development of the ICOA and its application to EDM, the exact nature of mutual correlations between the key parameters $\kappa \in \mathcal{K}_1$ has not been discussed, which depends on the source and duration of data, space weather conditions and also the choice of the profile function $p(h)$. These aspects could be investigated further in future as well. In this thesis the key parameters of the $D-$, the $E-$ and the F_1-layer are obtained from the IRI 2012 model. It shall be noted that IRI also makes assumptions, such as the existence of the F_1 layer only within mid- and low-latitudes; see Bilitza (2000), Coiosson et al. (2005). Although these assumptions may be reasonable from a long term climatology perspective but we need improved models for these layers in our global 4D EDM.

If the ICOA is to be used within an operational product generation, then improvements to the computing platform shall also be considered. One such implementation could be the use of a distributed computing infrastructure (DCI). Although this work was performed on a desktop computer, but a DCI could be more efficient for an operational center targeting a low latency product e.g. real-time VTEC maps. Each of the dedicated tasks such as raw data download in near real time, pre-processing, constraint processing, ICOA based estimation of the Chapman key parameters, quality control and verification could be independently handled by a specific (and dedicated) computing infrastructure. High performance cluster computing facilities such as the one hosted by Leibniz-Rechnerzentrum in Garching near Munich and other cloud based hosting infrastructures are very relevant.

[2] COSMIC-2 data is already available since October 2019 but the analysis period chosen for this work is March 2015.
[3] $www.nesdis.noaa.gov/content/after-14-years-cosmicformosat-3-ends-service$

Bibliography

A.Anthes R., Rocken C., and Kuo Y.-H. (2000). *Applications of COSMIC to Meteorology and Climate*. Terrestrial, Atmospheric and Oceanic Sciences, 11(1):115–156, ISSN: 1017-0839.

Akasofu S.-I. (2020). *Sydney Chapman: A Biographical Sketch Based on the Book "Chapman Eighty, From His Friends"*. Perspectives of Earth and Space Scientists, 1(1):e2020CN000135, DOI: https://doi.org/10.1029/2020CN000135, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020CN000135. e2020CN000135 2020CN000135.

Alkhatib H. and Schuh W.-D. (2007). *Integration of the Monte Carlo covariance estimation strategy into tailored solution procedures for large-scale least squares problems*. Journal of Geodesy, 81(1):53–66.

Altadill D., Magdaleno S., Torta J., and Blanch E. (2013). *Global empirical models of the density peak height and of the equivalent scale height for quiet conditions*. Advances in Space Research, 52(10):1756–1769.

Anderson E. D., Gondzio J., Mészáros C., and Xu X. (1996). *Implementation of Interior-Point Methods for Large Scale Linear Programs*, pages 189–252. Springer US, Boston, MA, ISBN: 978-1-4613-3449-1, DOI: 10.1007/978-1-4613-3449-1_6, https://doi.org/10.1007/978-1-4613-3449-1_6.

Appleton E. V. and Beynon W. J. G. (1940). *The application of ionospheric data to radio-communication problems: Part I*. Proceedings of the Physical Society (1926-1948), 52(4):518.

Appleton E. V. and Beynon W. J. G. (1947). *The application of ionospheric data to radio communication problems: Part II*. Proceedings of the Physical Society (1926-1948), 59(1):58.

Aragon-Angel A., Limberger M., Hernandez-Pajares M., Altadill D., Erdogan E., and Schmidt M. (2016). *Evaluation of the Ionospheric F2 Characteristics Inferred from Radio Occultations Exploiting the Availability of FORMOSAT-3/COSMIC Data Over Half a Solar Cycle*. Proceedings of the 29th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2016), Portland, Oregon, September 2016, pp. 892-901.

Arora J. S. (2017). *Chapter 12 - Numerical Methods for Constrained Optimum Design*. In Arora J. S. (Ed.), *Introduction to Optimum Design (Fourth Edition)*, pages 511 – 553. Academic Press, Boston, fourth edition edition, ISBN: 978-0-12-800806-5, DOI: https://doi.org/10.1016/B978-0-12-800806-5.00012-3, http://www.sciencedirect.com/science/article/pii/B9780128008065000123.

Arras C., Jacobi C., and Wickert J. (2009). *Semidiurnal tidal signature in sporadic E occurrence rates derived from GPS radio occultation measurements at higher midlatitudes*. Annales Geophysicae, 27(6):2555–2563, DOI: 10.5194/angeo-27-2555-2009, https://www.ann-geophys.net/27/2555/2009/.

Auriol A. and Tourain C. (2010). *DORIS system: the new age*. Advances in space research, 46(12):1484–1496.

Bao F. (2020). *Stochastic gradient descent algorithm for stochastic optimization in solving analytic continuation problems*, volume 2. ISSN: A0000-0002, DOI: 10.3934/fods.2020001, http://aimsciences.org//article/id/7ccff998-12f7-4cd3-9da6-dde4fd4ab6e2.

Bartholomew-Biggs M. (2005). *The Newton Method*, pages 65–76. Springer US, Boston, MA, ISBN: 978-0-387-24149-4, DOI: 10.1007/0-387-24149-3_6, https://doi.org/10.1007/0-387-24149-3_6.

Bauer S. J. (1973). *Observed Properties of Planetary Ionospheres*. In *Physics of Planetary Ionospheres*, pages 178–203. Springer.

Belehaki A., Marinov P., Kutiev I., Jakowski N., and Stankov S. (2006). *Comparison of the topside ionosphere scale height determined by topside sounders model and bottomside digisonde profiles*. Advances in Space Research, 37(5):963 – 966.

Berkner L. V. and Wells H. W. (1934). *F-region ionosphere-investigations at low latitudes*. Terrestrial Magnetism and Atmospheric Electricity, 39(3):215–230, DOI: https://doi.org/10.1029/TE039i003p00215, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/TE039i003p00215.

Bernstein D. (2009). *Matrix Mathematics: Theory, Facts, and Formulas - Second Edition*. Princeton University Press, ISBN: 9781400833344, https://books.google.de/books?id=-c0NxJg4vHMC.

Bertsekas D. P. (2009). *Convex Optimization Theory*. Athena Scientific, Belmont, Massachusetts, ISBN: 978-1-886529-31-1.

Beutler G., Moore A. W., and Mueller I. I. (2009). *The international global navigation satellite systems service (IGS): development and achievements*. Journal of Geodesy, 83(3-4):297–307.

Bilitza D. (2000). *International Reference Ionosphere 2000*. Radio Science, 36(2):261–275, DOI: 10.1029/2000RS002432, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2000RS002432.

Bilitza D., McKinnell L., Reinisch B., and Fuller-Rowell T. (2011). *The international reference ionosphere today and in the future*. Journal of Geodesy, 85(12):909–920, ISSN: 1432-1394, DOI: 10.1007/s00190-010-0427-x, https://doi.org/10.1007/s00190-010-0427-x.

Billups S. C. and Ferris M. C. (1996). *Convergence of an Infeasible Interior-Point Algorithm from Arbitrary Positive Starting Points*. SIAM Journal on Optimization, 6(2):316–325, DOI: 10.1137/0806018, https://doi.org/10.1137/0806018.

Bjoland V. B., Løvhaug U. P., and Hoz C. L. (2016). *An evaluation of International Reference Ionosphere electron density in the polar cap and cusp using EISCAT Svalbard radar measurements*. Ann. Geophys., 34, 751–758, DOI: 10.5194/angeo-34-751-2016.

Blewitt G. (1989). *Carrier phase ambiguity resolution for the Global Positioning System applied to geodetic baselines up to 2000 km*. Journal of Geophysical Research: Solid Earth, 94(B8):10187–10203.

Bonnans J.F. S. A. (2000). *Stability and Sensitivity Analysis. In: Perturbation Analysis of Optimization Problems*.

Borries C., Jakowski N., and Wilken V. (2009). *Storm induced large scale TIDs observed in GPS derived TEC*. Annales Geophysicae, 27(4):1605–1612, DOI: 10.5194/angeo-27-1605-2009, https://www.ann-geophys.net/27/1605/2009/.

Borwein J. M. and Lewis A. S. (2000). *Inequality Constraints*, pages 15–32. Springer New York, New York, NY, ISBN: 978-1-4757-9859-3, DOI: 10.1007/978-1-4757-9859-3_2, https://doi.org/10.1007/978-1-4757-9859-3_2.

Bottou L. (2012). *Stochastic Gradient Descent Tricks*, pages 421–436. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-35289-8, DOI: 10.1007/978-3-642-35289-8_25, https://doi.org/10.1007/978-3-642-35289-8_25.

Bowler N. E. (2020). *An assessment of GNSS radio occultation data produced by Spire*. Quarterly Journal of the Royal Meteorological Society, 146(733):3772–3788.

Bowman B., Tobiska W. K., Marcos F., Huang C., Lin C., and Burke W. (2008). *A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices*. In *AIAA/AAS astrodynamics specialist conference and exhibit*, page 6438.

Box G. E. P. and Jenkins G. (1990). *Time Series Analysis, Forecasting and Control*. Holden-Day, Inc., USA, ISBN: 0816211043.

Boyd R. (1965). *An Introduction to Langmuir Probes for Space Research*. In *Introduction to Solar Terrestrial Relations*, pages 455–465. Springer.

Boyd S. and Vandenberghe L. (2004). *Convex Optimization*. Cambridge University Press, ISBN: 978-0-521-83378-3.

Braasch M. S. (1997). *Autocorrelation sidelobe considerations in the characterization of multipath errors*. IEEE Transactions on Aerospace and Electronic Systems, 33(1):290–295.

Breiman L. (2001). *Random forests*. Machine learning, 45(1):5–32.

Bremer J. (1998). *Trends in the ionospheric E and F regions over Europe*. Annales Geophysicae, 16(8):986–996, ISSN: 1432-0576, DOI: 10.1007/s00585-998-0986-9, https://doi.org/10.1007/s00585-998-0986-9.

Brillouin L. (2013). *Wave propagation and group velocity*, volume 8. Academic press.

Brockmann J. and Schuh W.-D. (2010). *Fast variance component estimation in GOCE data processing*. In *Gravity, Geoid and Earth Observation*, pages 185–193. Springer.

Brockwell P. J. and Davis R. A. (1991). *Stationary ARMA Processes*, pages 77–113. Springer New York, New York, NY, ISBN: 978-1-4419-0320-4, DOI: 10.1007/978-1-4419-0320-4_3, https://doi.org/10.1007/978-1-4419-0320-4_3.

Bruinsma S. (2015). *The DTM-2013 thermosphere model*. Journal of Space Weather and Space Climate, 5:A1.

Carpenter D. L. and Park C. G. (1973). *On what ionospheric workers should know about the plasmapause-plasmasphere*. Reviews of Geophysics, 11(1):133–154, DOI: https://doi.org/10.1029/RG011i001p00133, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG011i001p00133.

Catapano F., Buchert S., Coco I., Slominska E., Qamili E., Trenchi L., and Bouffard J. (2020). *Swarm Langmuir Probe measurements: analysis and characterization of the data quality*. In *EGU General Assembly Conference Abstracts*, page 13403.

Chapman S. (1931). *The absorption and dissociative or ionizing effect of monochromatic radiation in an atmosphere on a rotating earth*. Proceedings of the Physical Society, 43(1):26, http://stacks.iop.org/0959-5309/43/i=1/a=305.

Chatfield C. and Pepper M. (1971). *Time-series analysis: an example from geophysical data*. Journal of the Royal Statistical Society: Series C (Applied Statistics), 20(3):217–238.

Clevert D.-A., Unterthiner T., and Hochreiter S. (2016). *Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)*. CoRR, abs/1511.07289.

Cloud M. J., Drachman B. C., and Lebedev L. P. (2014). *Some Standard Inequalities*, pages 53–74. Springer International Publishing, Cham, ISBN: 978-3-319-05311-0, DOI: 10.1007/978-3-319-05311-0_3.

Codrescu M. V., Negrea C., Fedrizzi M., Fuller-Rowell T., Dobin A., Jakowsky N., Khalsa H., Matsuo T., and Maruyama N. (2012). *A real-time run of the Coupled Thermosphere Ionosphere Plasmasphere Electrodynamics (CTIPe) model*. Space Weather, 10(2).

Coiosson P., Radicella S., Leitinger R., and Nava B. (2005). *Topside electron density in IRI and NeQuick: Features and limitations*. J. Adv. Space Res., 37:937–942.

Coleman T. and Li Y. (1996). *A Reflective Newton Method for Minimizing a Quadratic Function Subject to Bounds on Some of the Variables*. SIAM Journal on Optimization, 6(4):1040–1058.

Cooke D., Roth C., Luehr H., and Jakowski N. (2003). *In situ plasma observation at 400 km on the CHAMP satellite*. In *EGS-AGU-EUG Joint Assembly*, page 13004.

Cottle R. W. (2009). *Linear complementarity problem*, pages 1873–1878. Springer US, Boston, MA, ISBN: 978-0-387-74759-0, DOI: 10.1007/978-0-387-74759-0_333, https://doi.org/10.1007/978-0-387-74759-0_333.

Curry H. B. (1944). *The Method of steepest descent for non-linear minimization problems*. Quarterly of Applied Mathematics, 2(3):258–261, ISSN: 0033569X, 15524485, http://www.jstor.org/stable/43633461.

Czyzyk J., Mesnier M. P., and Moré J. J. (1998). *The NEOS Server*. IEEE Journal on Computational Science and Engineering, 5(3):68 – 75.

Dach R., Brockmann E., Schaer S., Beutler G., Meindl M., Prange L., Bock H., Jäggi A., and Ostini L. (2009). *GNSS processing at CODE: status report*. Journal of Geodesy, 83(3-4):353–365.

Dalla Torre A. and Caporali A. (2015). *An analysis of intersystem biases for multi-GNSS positioning*. Gps Solutions, 19(2):297–307.

Davies K. (1990). *Ionospheric radio*. Number 31. IET.

Davies K. and Liu X. (1991). *Ionospheric slab thickness in middle and low latitudes.* Radio Science, 26(4):997–1005.

Davis R. E. (1977). *Techniques for statistical analysis and prediction of geophysical fluid systems.* Geophysical & Astrophysical Fluid Dynamics, 8(1):245–277.

de Wit T. D. and Bruinsma S. (2017). *The 30 cm radio flux as a solar proxy for thermosphere density modelling.* Journal of Space Weather and Space Climate, 7:A9.

Dettmering D., Heinkelmann R., and Schmidt M. (2011a). *Systematic differences between VTEC obtained by different space-geodetic techniques during CONT08.* Journal of Geodesy, 85(7):443, ISSN: 1432-1394, DOI: 10.1007/s00190-011-0473-z, https://doi.org/10.1007/s00190-011-0473-z.

Dettmering D., Limberger M., and Schmidt M. (2014). *Using DORIS measurements for modeling the vertical total electron content of the Earth's ionosphere.* Journal of Geodesy, 88(12):1131–1143.

Dettmering D., Schmidt M., Heinkelmann R., and Seitz M. (2011b). *Combination of different space-geodetic observations for regional ionosphere modeling.* Journal of Geodesy, 85(12):989–998, ISSN: 1432-1394, DOI: 10.1007/s00190-010-0423-1, https://doi.org/10.1007/s00190-010-0423-1.

Di Giovanni G. and Radicella S. M. (1990). *An analytical model of the electron density profile in the ionosphere.* Advances in Space Research, 10(11):27 – 30, ISSN: 0273-1177, DOI: https://doi.org/10.1016/0273-1177(90)90301-F, http://www.sciencedirect.com/science/article/pii/027311779090301F.

Diego P., Coco I., Bertello I., Candidi M., and Ubertini P. (2019). *Ionospheric plasma density measurements by swarm Langmuir probes: Limitations and possible corrections.* Annales Geophysicae Discussions, pages 1–15.

Dolan E. D. (2001). *The NEOS Server 4.0 Administrative Guide.* Technical Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory.

Doornbos E. (2012a). *Empirical modelling of the thermosphere.* In *Thermospheric density and wind determination from satellite dynamics*, pages 21–57. Springer.

Doornbos E. (2012b). *Thermospheric density and wind determination from satellite dynamics.* Springer Science & Business Media.

Dow J. M., Neilan R. E., and Rizos C. (2009). *The International GNSS Service in a changing landscape of Global Navigation Satellite Systems.* Journal of Geodesy, 83(3):191–198, ISSN: 1432-1394, DOI: 10.1007/s00190-008-0300-3, https://doi.org/10.1007/s00190-008-0300-3.

Driggs D., Ehrhardt M. J., and Schönlieb C.-B. (2020). *Accelerating variance-reduced stochastic gradient methods.* Mathematical Programming, ISSN: 1436-4646, DOI: 10.1007/s10107-020-01566-2, https://doi.org/10.1007/s10107-020-01566-2.

Duchi J., Hazan E., and Singer Y. (2011). *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization.* Journal of Machine Learning Research, page 2121–2159.

Eissfeller B. and Won J.-H. (2017). *Receiver Architecture*, pages 365–400. Springer International Publishing, Cham, ISBN: 978-3-319-42928-1, DOI: 10.1007/978-3-319-42928-1_13, https://doi.org/10.1007/978-3-319-42928-1_13.

Endeshaw L. (2020). *Testing and validating IRI-2016 model over Ethiopian ionosphere.* Astrophys Space Sci 365, 49, https://doi.org/10.1007/s10509-020-03761-1.

Erdogan E., Schmidt M., Goss A., Görres B., and Seitz F. (2020). *Adaptive Modeling of the Global Ionosphere Vertical Total Electron Content.* Remote Sensing, 12(11), DOI: 10.3390/rs12111822.

Erdogan E., Schmidt M., Seitz F., and Durmaz M. (2017). *Near real-time estimation of ionosphere vertical total electron content from GNSS satellites using B-splines in a Kalman filter.* Annales Geophysicae, 35(2):263–277, DOI: 10.5194/angeo-35-263-2017.

Feltens J. (2007). *Development of a new three-dimensional mathematical ionosphere model at European Space Agency/European Space Operations Centre.* Space Weather, 5(12), DOI: 10.1029/2006SW000294, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006SW000294.

Forst W. and Hoffmann D. (2010). Springer, New York, eBook Packages Mathematics and Statistics Mathematics and Statistics, ISBN: 978-0-387-78976-7, DOI: https://doi.org/10.1007/978-0-387-78977-4.

Förster M. and Jakowski N. (2000). *Geomagnetic Storm Effects on the Topside Ionosphere and Plasmasphere: A Compact Tutorial and New Results*. Surveys in Geophysics, 21(1):47–87, ISSN: 1573-0956, DOI: 10.1023/A:1006775125220, https://doi.org/10.1023/A:1006775125220.

Fraser C. (2005). *Joseph Louis Lagrange, Théorie des fonctions analytiques, first edition (1797)*. Landmark Writings in Western Mathematics 1640-1940, DOI: 10.1016/B978-044450871-3/50100-5.

Fritsch D. (1985). *Some Additional Informations on the Capacity of the Linear Complementarity Algorithm*. In Grafarend E. W. and Sansò F. (Eds.), *Optimization and Design of Geodetic Networks*, pages 169–184, Berlin, Heidelberg. Springer Berlin Heidelberg, ISBN: 978-3-642-70659-2.

Fu L.-L. and Cheney R. E. (1995). *Application of satellite altimetry to ocean circulation studies: 1987–1994*. Reviews of Geophysics, 33(S1):213–223.

Fuller-Rowell T., Emmert J., Fedrizzi M., Weimer D., Codrescu M. V., Pilinski M., Sutton E., Viereck R., Raeder J. J., and Doornbos E. (2018). *How might the thermosphere and ionosphere react to an extreme space weather event?* In *Extreme Events in Geospace*, pages 513–539. Elsevier.

Galkin I., Reinisch B. W., Huang X., and Khmyrov G. M. (2013). *Confidence score of ARTIST-5 ionogram autoscaling*. Ionosonde Network Advisory Group (INAG) Bulletin No, 73:1–7.

Galkin I. A., Khmyrov G. M., Kozlov A. V., Reinisch B. W., Huang X., and Paznukhov V. V. (2008). *The Artist 5*. In *AIP Conference Proceedings*, volume 974, pages 150–159. American Institute of Physics.

Galkin I. A. and Reinisch B. W. (2008). *The new ARTIST 5 for all digisondes*. Ionosonde Network Advisory Group Bulletin, 69(8):1–8.

Gao Y. and Liu Z. (2002). *Precise Ionosphere Modeling Using Regional GPS Network Data*. Journal of Global Positioning Systems, 1:18–24, DOI: 10.5081/jgps.1.1.18.

Garcia Fernández M. (2004). *Contributions to the 3D ionospheric sounding with GPS data*. Dissertation, Universitat Politècnica de Catalunya. Departament de Matematica Aplicada IV, Barcelona, ISBN: 8468881562.

Gass S. I. and Harris C. M. (2001). *Complementary slackness theorem*, pages 118–119. Springer US, New York, NY, ISBN: 978-1-4020-0611-1, DOI: 10.1007/1-4020-0611-X_140, https://doi.org/10.1007/1-4020-0611-X_140.

Ge M., Gendt G., Dick G., and Zhang F. (2005). *Improving carrier-phase ambiguity resolution in global GPS network solutions*. Journal of Geodesy, 79(1-3):103–110.

Gentleman W. M. (1976). *Solving Least Squares Problems (Charles L. Lawson and Richard J. Hanson)*. SIAM Review, 18(3):518–520, DOI: 10.1137/1018100.

Gerzen T., Jakowski N., Wilken V., and Hoque M. (2013). *Reconstruction of F2 layer peak electron density based on operational vertical Total Electron Content maps*. Annales Geophysicae, 31:1241–1249, DOI: 10.5194/angeo-31-1241-2013.

Gilbert J. D. and Smith R. W. (1988). *A comparison between the automatic ionogram scaling system ARTIST and the standard manual method*. Radio science, 23(06):968–974.

Gill P., Murray W., Saunders M., and Wright M. (1984). *Procedures for Optimization Problems with a Mixture of Bounds and General Linear Constraints*. ACM Trans. Math. Softw., 10(3):282–298, ISSN: 0098-3500, DOI: 10.1145/1271.1276, http://doi.acm.org/10.1145/1271.1276.

Goldstine H. H. (1980). *Lagrange and Legendre*, pages 110–150. Springer New York, New York, NY, ISBN: 978-1-4613-8106-8, DOI: 10.1007/978-1-4613-8106-8_3, https://doi.org/10.1007/978-1-4613-8106-8_3.

Goss, Schmidt M., Erdogan E., Görres B., and Seitz F. (2019). *High-resolution vertical total electron content maps based on multi-scale B-spline representations*. Annales Geophysicae, 37(4):699–717, DOI: 10.5194/angeo-37-699-2019, https://www.ann-geophys.net/37/699/2019/.

Goss A., Schmidt M., Erdogan E., and F S. (2020). *Global and Regional High-Resolution VTEC Modelling Using a Two-Step B-Spline Approach.* Remote Sens. Vol. 12, 1198., DOI: https://doi.org/10.3390/rs12071198.

Gropp W. and Moré J. J. (1997). *Optimization Environments and the NEOS Server.* In Buhman M. D. and Iserles A. (Eds.), *Approximation Theory and Optimization*, pages 167 – 182. Cambridge University Press.

Groves G. (1972). *Aspects of Atmosphere.* Nature, 239(5369):236–237.

Gulyaeva T. (1985). *Ionospheric electron density profiles at sunrise-sunset.* Advances in space research, 5(7):13–20.

Gulyaeva T. (1987). *Progress in ionospheric informatics based on electron-density profile analysis of ionograms.* Advances in space research, 7(6):39–48.

Gurtner W. and Estey L. (2007). *Rinex-the receiver independent exchange format-version 3.00.* Astronomical Institute, University of Bern and UNAVCO, Bolulder, Colorado.

Hajj G. A., Kursinski E., Romans L., Bertiger W., and Leroy S. (2002). *A technical description of atmospheric sounding by GPS occultation.* Journal of Atmospheric and Solar-Terrestrial Physics, 64(4):451–469.

Halldórsson B. V. and Tütüncü R. H. (2003). *An Interior-Point Method for a Class of Saddle-Point Problems.* Journal of Optimization Theory and Applications, 116(3):559–590, ISSN: 1573-2878, DOI: 10.1023/A:1023065319772, https://doi.org/10.1023/A:1023065319772.

Hanslmeier A. (2002). *The Ionosphere and Space Weather*, pages 169–192. Springer Netherlands, Dordrecht, ISBN: 978-0-306-48211-3, DOI: 10.1007/0-306-48211-8_9, https://doi.org/10.1007/0-306-48211-8_9.

Hapgood M. (2017). *Space Weather.* IOP Publishing, Bristol, ISBN: 78-0-7503-1372-8.

Hargreaves J. K. J. K. (1992). *The solar-terrestrial environment : an introduction to geospace–the science of the terrestrial upper atmosphere, ionosphere, and magnetosphere.* Cambridge [England] ; New York, NY, USA : Cambridge University Press, ISBN: 0521327482.

Hastings D. E. (1995). *A review of plasma interactions with spacecraft in low Earth orbit.* Journal of Geophysical Research: Space Physics, 100(A8):14457–14483, ISSN: 2156-2202, DOI: 10.1029/94JA03358, http://dx.doi.org/10.1029/94JA03358.

Hauschild A. (2017). *Combinations of Observations*, pages 583–604. Springer International Publishing, Cham, ISBN: 978-3-319-42928-1, DOI: 10.1007/978-3-319-42928-1_20, https://doi.org/10.1007/978-3-319-42928-1_20.

Heelis R. (2004). *Electrodynamics in the low and middle latitude ionosphere: A tutorial.* Journal of Atmospheric and Solar-Terrestrial Physics, 66(10):825–838.

Heise S., Jakowski N., and Cooke D. (2005). *Ionosphere/plasmasphere imaging based on GPS navigation measurements from CHAMP and SAC-C.* In *Earth Observation with CHAMP*, pages 471–476. Springer.

Hendrycks D. and Gimpel K. (2016a). *Bridging nonlinearities and stochastic regularizers with gaussian error linear units.*

Hendrycks D. and Gimpel K. (2016b). *Gaussian error linear units (gelus).* arXiv preprint arXiv:1606.08415.

Henkel P. (2010). *Reliable carrier phase positioning.* PhD thesis, Technische Universität München.

Hernández-Pajares M., Juan J., Sanz J., and Solé J. (1998). *Global observation of the ionospheric electronic response to solar events using ground and LEO GPS data.* Journal of Geophysical Research: Space Physics, 103(A9):20789–20796, ISSN: 2156-2202, DOI: 10.1029/98JA01272, http://dx.doi.org/10.1029/98JA01272.

Hernández-Pajares M., Juan J. M., Sanz J., Orus R., Garcia-Rigo A., Feltens J., Komjathy A., Schaer S. C., and Krankowski A. (2009). *The IGS VTEC maps: a reliable source of ionospheric information since 1998.* Journal of Geodesy, 83(3):263–275, ISSN: 1432-1394, DOI: 10.1007/s00190-008-0266-1, https://doi.org/10.1007/s00190-008-0266-1.

Hernández-Pajares M., Juan J., and Sanz J. (1999). *New approaches in global ionospheric determination using ground GPS data.* Journal of Atmospheric and Solar-Terrestrial Physics, 61(16):1237 – 1247, ISSN:

1364-6826, DOI: https://doi.org/10.1016/S1364-6826(99)00054-1, http://www.sciencedirect.com/science/article/pii/S1364682699000541.

Hernández-Pajares M., Juan J. M., Sanz J., and Solé J. G. (1998a). *Global observation of the ionospheric electronic response to solar events using ground and LEO GPS data*. Journal of Geophysical Research: Space Physics, 103(A9):20789–20796, DOI: 10.1029/98JA01272, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JA01272.

Hernández-Pajares M., Juan J. M., Sanz J., and Solé J. G. (1998b). *Global observation of the ionospheric electronic response to solar events using ground and LEO GPS data*. Journal of Geophysical Research: Space Physics, 103(A9):20789–20796, DOI: 10.1029/98JA01272, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98JA01272.

Hestenes M. R. (1980). *Newton's Method and the Gradient Method*, pages 1–80. Springer New York, New York, NY, ISBN: 978-1-4612-6048-6, DOI: 10.1007/978-1-4612-6048-6_1, https://doi.org/10.1007/978-1-4612-6048-6_1.

Hinton G. (2012). *Lecture Notes on computer science*.

Hoffmann P. and Thiele R. (2007). *Leonhard Euler and Russia*. In Bradley R. E. and Sandifer C. E. (Eds.), *Leonhard Euler: Life, Work and Legacy*, volume 5 of *Studies in the History and Philosophy of Mathematics*, pages 61 – 73. Elsevier, ISSN: 0928-2017, DOI: https://doi.org/10.1016/S0928-2017(07)80005-3, http://www.sciencedirect.com/science/article/pii/S0928201707800053.

Hofmann-Wellenhof B., Lichtenegger H., and Collins J. (2012). *Global positioning system: theory and practice*. Springer Science & Business Media.

Hofmann-Wellenhof B., Lichtenegger H., and Wasle E. (2007). *GNSS–global navigation satellite systems: GPS, GLONASS, Galileo, and more*. Springer Science & Business Media.

Hopwood J., Guarnieri C., Whitehair S., and Cuomo J. (1993). *Langmuir probe measurements of a radio frequency induction plasma*. Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, 11(1):152–156.

Howe B. M., Runciman K., and Secan J. A. (1998). *Tomography of the ionosphere: Four-dimensional simulations*. Radio Science, 33(1):109–128, DOI: 10.1029/97RS02615, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/97RS02615.

Hsu C.-T., Matsuo T., Yue X., Fang T.-W., Fuller-Rowell T., Ide K., and Liu J.-Y. (2018). *Assessment of the Impact of FORMOSAT-7/COSMIC-2 GNSS RO Observations on Midlatitude and Low-Latitude Ionosphere Specification: Observing System Simulation Experiments Using Ensemble Square Root Filter*. Journal of Geophysical Research: Space Physics, 123(3):2296–2314.

Hunsucker R. D. (1991). *Vertical Sounders - the Ionosonde*, pages 67–93. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-76257-4, DOI: 10.1007/978-3-642-76257-4_3, https://doi.org/10.1007/978-3-642-76257-4_3.

Hutchinson M. (1990). *A stochastic estimator of the trace of the influence matrix for laplacian smoothing splines*. Communications in Statistics - Simulation and Computation, 19(2):433–450, DOI: 10.1080/03610919008812866, https://doi.org/10.1080/03610919008812866.

Hutschison G. (1957). *Concluding Remarks*. Cold spring Harbour symposia on qualitative biology, 22:415–427, DOI: doi:10.1101/SQB.1957.022.01.039.

Jackson D. R., Bruinsma S., Negrin S., Stolle C., Budd C. J., Gonzalez R. D., Down E., Griffin D. J., Griffith M. J., Kervalishvili G., et al. (2020). *The Space Weather Atmosphere Models and Indices (SWAMI) project: Overview and first results*. Journal of Space Weather and Space Climate, 10:18.

Jacobs R. A. (1988). *Increased rates of convergence through learning rate adaptation*. Neural Networks, 1(4):295 – 307, ISSN: 0893-6080, DOI: https://doi.org/10.1016/0893-6080(88)90003-2, http://www.sciencedirect.com/science/article/pii/0893608088900032.

Jakowski N. (2017). *Ionosphere Monitoring*, pages 1139–1162. Springer International Publishing, Cham, ISBN: 978-3-319-42928-1, DOI: 10.1007/978-3-319-42928-1_39, http://dx.doi.org/10.1007/978-3-319-42928-1_39.

Jakowski N., Schlüter S., and Sardon E. (1999). *Total electron content of the ionosphere during the geomagnetic storm on 10 January 1997*. Journal of Atmospheric and Solar-Terrestrial Physics, 61(3):299 – 307, ISSN: 1364-6826, DOI: https://doi.org/10.1016/S1364-6826(98)00130-8, http://www.sciencedirect.com/science/article/pii/S1364682698001308.

Jakowski N., Wehrenpfennig A., Heise S., and Kutiev I. (2002). *Space weather effects on transionospheric radio wave propagation on 6 April 2000*. Acta Geodaetica et Geophysica Hungarica, 37(2):213–220, ISSN: 1587-1037, DOI: 10.1556/AGeod.37.2002.2-3.10, https://doi.org/10.1556/AGeod.37.2002.2-3.10.

Jekeli C. and Montenbruck O. (2017). *Time and Reference Systems*, pages 25–58. Springer International Publishing, Cham, ISBN: 978-3-319-42928-1, DOI: 10.1007/978-3-319-42928-1_2, https://doi.org/10.1007/978-3-319-42928-1_2.

Jensen A. S., Lohmann M. S., Nielsen A. S., and Benzon H.-H. (2004). *Geometrical optics phase matching of radio occultation signals*. Radio science, 39(3).

Karush W. (1939). *Minima of functions of several variables with inequalities as side conditions*. Master's thesis, Department of Mathematics, University of Chicago.

Kerker M. (1969). *The Scattering of Light (New York and London*.

Khirirat S., Feyzmahdavian H. R., and Johansson M. (2017). *Mini-batch gradient descent: Faster convergence under data sparsity*. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 2880–2887.

Kim D., Serrano L., and Langley R. (2006). *Phase wind-up analysis: Assessing real-time kinematic performance*. 17:58–64.

Kingma D. P. and Ba J. (2014). *Adam: A Method for Stochastic Optimization*.

Kleusberg A. and Teunissen P. J. (1996). *GPS for Geodesy*.

Knight B. and Adams R. (1975). *Newton's Method*, pages 67–71. Springer US, Boston, MA, ISBN: 978-1-4615-6594-9, DOI: 10.1007/978-1-4615-6594-9_10, https://doi.org/10.1007/978-1-4615-6594-9_10.

Knudsen D., Burchill J., Buchert S., Eriksson A., Gill R., Wahlund J.-E., Åhlén L., Smith M., and Moffat B. (2017). *Thermal ion imagers and Langmuir probes in the Swarm electric field instruments*. Journal of Geophysical Research: Space Physics, 122(2):2655–2673.

Koch K. R. (1985). *First Order Design: Optimization of the Configuration of a Network by Introducing Small Position Changes*. In Grafarend E. W. and Sansò F. (Eds.), *Optimization and Design of Geodetic Networks*, pages 56–73, Berlin, Heidelberg. Springer Berlin Heidelberg, ISBN: 978-3-642-70659-2.

Koch K. R. (1988). *Optimization of the configuration of geodetic networks*. Springer Verlag,, DOI: 10.1007/978-3-642-70659-2.

Koch K. R. (1999). *Parameter Estimation and Hypothesis Testing in Linear Models*. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-08461-4, DOI: 10.1007/978-3-662-03976-2, https://doi.org/10.1007/978-3-662-03976-2.

Koch K. R. and Kusche J. (2002). *Regularization of geopotential determination from satellite data by variance components*. J Geod, 76:259–268.

Koch K. R. and Schmidt M. (2010). *N-dimensional B-spline surface estimated by lofting for locally improving IRI*. Journal of Geodetic Science, 1(1), pp. 41-51., DOI: 10.2478/v10156-010-0006-3.

Koetsier T. (2007). *Euler and Kinematics*. In Bradley R. E. and Sandifer C. E. (Eds.), *Leonhard Euler: Life, Work and Legacy*, volume 5 of *Studies in the History and Philosophy of Mathematics*, pages 167 – 194. Elsevier, ISSN: 0928-2017, DOI: https://doi.org/10.1016/S0928-2017(07)80011-9, http://www.sciencedirect.com/science/article/pii/S0928201707800119.

Kuhn H. W. and Tucker A. W. (1951). *Nonlinear Programming*. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492, Berkeley, Calif. University of California Press.

Lang S. (1986). *The Derivative*, pages 57–116. Springer New York, New York, NY, ISBN: 978-1-4419-8532-3, DOI: 10.1007/978-1-4419-8532-3_3, https://doi.org/10.1007/978-1-4419-8532-3_3.

Langley R. B., Teunissen P. J. G., and Montenbruck O. (2017). *Introduction to GNSS*, pages 3–23. Springer International Publishing, Cham, ISBN: 978-3-319-42928-1, DOI: 10.1007/978-3-319-42928-1_1, https://doi.org/10.1007/978-3-319-42928-1_1.

Lanzerotti L. J. (2013). *Van Allen Probes Mission*. Space Weather, 11(4):133–133, DOI: https://doi.org/10.1002/swe.20037, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/swe.20037.

Lei J. and Chuo Y.-J. (2014). *Variations of Scale Height at F-Region Peak Based on Ionosonde Measurements during Solar Maximum over the Crest of Equatorial Ionization Anomaly Region*. The Scientific World Journal, Hindawi Publishing Corporation, DOI: 10.1155/2014/397402, http://www.sciencedirect.com/science/article/pii/0024379572900195.

Lemke C. E. (1968). *On complementary pivot theory*, pages 95–114. American Mathematical Society.

Leonard J. M., Forbes J. M., and Born G. H. (2012). *Impact of tidal density variability on orbital and reentry predictions*. Space Weather, 10(12):n/a–n/a, ISSN: 1542-7390, DOI: 10.1029/2012SW000842, http://dx.doi.org/10.1029/2012SW000842. S12003.

Liang (2017). *A regional physics-motivated electron density model of the ionosphere*. Dissertation, Technische Universität München, München.

Liang, Limberger M., M S., D D., and U H. (2016). *Combination of ground- and space-based GPS data for the determination of a multi-scale regional 4-D ionosphere model*. In: Rizos C., Willis P. (Eds.) IAG 150 Years, IAG Symposia, 143:751–758, ISSN: 0939-9585, DOI: 10.1007/1345_2015_25.

Liang W., Limberger M., Schmidt M., Dettmering D., and Hugentobler U. (2015a). *Combination of ground- and space-based GPS data for the determination of a multi-scale regional 4-D ionosphere model*. In *IAG 150 Years*, pages 751–758. Springer.

Liang W., Limberger M., Schmidt M., Dettmering D., Hugentobler U., Bilitza D., Jakowski N., Hoque M. M., Wilken V., and Gerzen T. (2015b). *Regional modeling of ionospheric peak parameters using GNSS data—An update for IRI*. Advances in Space Research, 55(8):1981–1993.

Limberger, Liang W., M S., D D., and U H. (2013). *Regional representation of F2 Chapman parameters based on electron density profiles*. In *Annales Geophysicae*, volume 31, pages 2215–2227. Copernicus GmbH.

Limberger M. (2015). *Ionosphere modeling from GPS radio occultations and complementary data based on B-splines*. Deutsche Geodätische Kommission, Reihe C, 755, Muenchen.

Limberger M., Liang, and Seitz (2014). *Correlation studies for B-spline modeled F2 Chapman parameters obtained from FORMOSAT-3/COSMIC data*. Annales Geophysicae, 32(12):1533–1545.

Liperovsky V., Pokhotelov O., Meister C.-V., and Liperovskaya E. (2008). *Physical models of coupling in the lithosphere-atmosphere-ionosphere system before earthquakes*. Geomagnetism and Aeronomy, 48(6):795–806.

Liu Q., Schmidt M., Pail R., and Willberg M. (2020). *Determination of the Regularization Parameter to Combine Heterogeneous Observations in Regional Gravity Field Modeling*. Remote Sensing, 12(10):1617, ISSN: 2072-4292, DOI: 10.3390/rs12101617, http://dx.doi.org/10.3390/rs12101617.

Lötstedt P. (1984). *Solving the minimal least squares problem subject to bounds on the variables*. BIT Numerical Mathematics, 24(2):205–224.

Luehr H., Liu H., Park J., and Mueller S. (2011). *New Aspects of the Coupling Between Thermosphere and Ionosphere, with Special regards to CHAMP Mission Results*, pages 303–316. Springer Netherlands, Dordrecht, ISBN: 978-94-007-0326-1, DOI: 10.1007/978-94-007-0326-1_22, http://dx.doi.org/10.1007/978-94-007-0326-1_22.

Luenberger D. G. (1999). *Practical optimization*. John Wiley and Sons, Inc., ISBN: 978-0-471-18117-0.

Lühr H., Park J., Ritter P., and Liu H. (2012). *In-situ CHAMP observation of ionosphere-thermosphere coupling*. Space Science Reviews, 168(1):237–260.

Lyche T. and Schumaker L. L. (2000). *A Multiresolution Tensor Spline Method for Fitting Functions on the Sphere*. SIAM Journal on Scientific Computing, 22(2):724–746, DOI: 10.1137/S1064827598344388, https://doi.org/10.1137/S1064827598344388.

Maravall A. (1983). *An application of nonlinear time series forecasting*. Journal of Business & Economic Statistics, 1(1):66–74.

Marler R. and Arora J. (2004). *Survey of multi-objective optimization methods for engineering*. Struct Multidisc Optim, 26(2):369–395, https://doi.org/10.1007/s00158-003-0368-6.

Maruyama T. (2010). *Solar proxies pertaining to empirical ionospheric total electron content models*. Journal of Geophysical Research: Space Physics, 115(A4), DOI: 10.1029/2009JA014890, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JA014890.

Mauk B., Fox N. J., Kanekal S., Kessel R., Sibeck D., and Ukhorskiy a. A. (2013). *Science objectives and rationale for the Radiation Belt Storm Probes mission*. Space Science Reviews, 179(1-4):3–27.

McGraw G. A. and Braasch M. S. (1999). *GNSS multipath mitigation using gated and high resolution correlator concepts*. In *Proceedings of the 1999 National Technical Meeting of The Institute of Navigation*, pages 333–342.

McLeod A. I. and Li W. K. (1983). *Diagnostic checking ARMA time series models using squared-residual autocorrelations*. Journal of time series analysis, 4(4):269–273.

McNamara L. F., Baker C. R., and Decker D. T. (2008). *Accuracy of USU-GAIM specifications of foF2 and M(3000)F2 for a worldwide distribution of ionosonde locations*. Radio Science, 43(1), DOI: 10.1029/2007RS003754, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2007RS003754.

McNamara L. F., Reinisch B. W., and Tang J. S. (1987). *Values of hmF2 deduced from automatically scaled ionograms*. Advances in space research, 7(6):53–56.

Mead J. L. and Renaut R. A. (2010). *Least squares problems with inequality constraints as quadratic constraints*. Linear Algebra and its Applications, 432(8):1936 – 1949, ISSN: 0024-3795, DOI: https://doi.org/10.1016/j.laa.2009.04.017, http://www.sciencedirect.com/science/article/pii/S0024379509002341. Special issue devoted to the 15th ILAS Conference at Cancun, Mexico, June 16-20, 2008.

Mehrotra S. (1992). *On the Implementation of a Primal-Dual Interior Point Method*. SIAM Journal on Optimization, 2(4):575–601, DOI: 10.1137/0802028, https://doi.org/10.1137/0802028.

Mendillo M., Rishbeth H., Roble R., and Wroten J. (2002). *Modelling F2-layer seasonal trends and day-to-day variability driven by coupling with the lower atmosphere*. Journal of Atmospheric and Solar-Terrestrial Physics, 64(18):1911–1931.

Menvielle M., Iyemori T., Marchaudon A., and Nosé M. (2011). *Geomagnetic Indices*, pages 183–228. Springer Netherlands, Dordrecht, ISBN: 978-90-481-9858-0, DOI: 10.1007/978-90-481-9858-0_8, https://doi.org/10.1007/978-90-481-9858-0_8.

Meurer M. and Antreich F. (2017). *Signals and Modulation*, pages 91–119. Springer International Publishing, Cham, ISBN: 978-3-319-42928-1, DOI: 10.1007/978-3-319-42928-1_4, https://doi.org/10.1007/978-3-319-42928-1_4.

Meyer R. M. (1979). *Convex Sets and Convex Functions*, pages 491–508. Springer New York, New York, NY, ISBN: 978-1-4613-8072-6, DOI: 10.1007/978-1-4613-8072-6_18, https://doi.org/10.1007/978-1-4613-8072-6_18.

Misra P. and Enge P. (2006). *Global Positioning System: signals, measurements and performance second edition*. Global Positioning System: Signals, Measurements And Performance Second Editions, 206:43.

Mitchell J. E., Farwell K., and Ramsden D. (2006). *Interior Point Methods for Large-Scale Linear Programming*, pages 3–25. Springer US, Boston, MA, ISBN: 978-0-387-30165-5.

Moldwin M. B., Thomsen M. F., Bame S. J., McComas D., and Reeves G. D. (1995). *The fine-scale structure of the outer plasmasphere*. Journal of Geophysical Research: Space Physics, 100(A5):8021–8029.

Montenbruck O., Garcia-Fernandez M., Yoon Y., Schön S., and Jäggi A. (2009). *Antenna phase center calibration for precise positioning of LEO satellites*. GPS solutions, 13(1):23.

Montenbruck O. and Gill E. (2001). *Book Review: Satellite orbits: models, methods and applications/Springer, 2000*. The Observatory, 121:182.

Montenbruck O. and Hauschild A. (2013). *Code Biases in Multi-GNSS Point Positioning*. Proceedings of the ION NTM.

Montenbruck O., Hauschild A., and Steigenberger P. (2014a). *Differential code bias estimation using multi-GNSS observations and global ionosphere maps*. Navigation: Journal of the Institute of Navigation, 61(3):191–201.

Montenbruck O., Steigenberger P., Khachikyan R., Weber G., Langley R., Mervart L., and Hugentobler U. (2014b). *IGS-MGEX: Preparing the Ground for Multi-Constellation GNSS Science*. Inside GNSS, 9(1):42–49.

Moritz H. (2015). *Classical Physical Geodesy*, pages 253–289. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-54551-1, DOI: 10.1007/978-3-642-54551-1_6, https://doi.org/10.1007/978-3-642-54551-1_6.

Murty K. G. (1972). *On the number of solutions to the complementarity problem and spanning properties of complementary cones*. Linear Algebra and its Applications, 5(1):65 – 108, ISSN: 0024-3795, DOI: https://doi.org/10.1016/0024-3795(72)90019-5, http://www.sciencedirect.com/science/article/pii/0024379572900195.

Nesterov Y. (1983). *A method for unconstrained convex minimization problem with the rate of convergence* $o(1/k^2)$.

Nesterov Y. and Nemirovskii A. (1994). *Interior-point polynomial algorithms in convex programming*. SIAM.

Nesterov Y., Wolkowicz H., and Ye Y. (2000). *Semidefinite Programming Relaxations of Nonconvex Quadratic Optimization*, pages 361–419. Springer US, Boston, MA, ISBN: 978-1-4615-4381-7, DOI: 10.1007/978-1-4615-4381-7_13.

Noakes D. J. (1986). *APPLIED TIME SERIES MODELLING AND FORECASTING*.

Nocedal J. and Wright S. J. (2006). *Numerical Optimization*. Springer, second edition edition, ISBN: 978-0-387-30303-1.

Obrou O., Bilitza D., Adeniyi J., and Radicella S. (2003). *Equatorial F2-layer peak height and correlation with vertical ion drift and M (3000) F2*. Advances in Space Research, 31(3):513–520.

Odijk D. and Teunissen P. J. (2013a). *Characterization of between-receiver GPS-Galileo inter-system biases and their effect on mixed ambiguity resolution*. GPS solutions, 17(4):521–533.

Odijk D. and Teunissen P. J. (2013b). *Estimation of differential inter-system biases between the overlapping frequencies of GPS, Galileo, BeiDou and QZSS*. In *4th International colloquium scientific and fundamental aspects of the Galileo programme*, pages 4–6.

Olivares-Pulido G., Terkildsen M., and Arsov, K. e. a. (2019). *A 4D tomographic ionospheric model to support PPP-RTK*. J Geod, 93:1673–1683.

Osborne E. E. (1965). *Smallest Least Squares Solutions of Linear Equations*. Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis, 2(2):300–307, DOI: 10.1137/0702024, https://doi.org/10.1137/0702024.

Oshiorenoya A. E. (2004). *Plasma in the ionosphere: Ionization and recombination*. Space Physics P, 5.

Panzetta F., Bloßfeld M., Erdogan E., Rudenko S., Schmidt M., and Müller H. (2019). *Towards thermospheric density estimation from SLR observations of LEO satellites: a case study with ANDE-Pollux satellite*. Journal of Geodesy, 93(3):353–368.

Parkinson M., Dyson P., Monselesan D., and Morris R. (1998). *On the role of electric field direction in the formation of sporadic E-layers in the southern polar cap ionosphere*. Journal of atmospheric and solar-terrestrial physics, 60(4):471–491.

Philip E. Gill W. M. and Wright M. H. (1981). *Practical optimization, , 1981. No. of pages.* ISBN: 978-1-611975-59-8.

Picone J. M., Hedin A. E., Drob D. P., and Aikin A. C. (2002). *NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues.* Journal of Geophysical Research: Space Physics, 107(A12):SIA 15–1–SIA 15–16, ISSN: 2156-2202, DOI: 10.1029/2002JA009430, http://dx.doi.org/10.1029/2002JA009430. 1468.

Polyak B. T. (1964). *Some methods of speeding up the convergence of iteration methods.* Computational Mathematics and Mathematical Physics, 5(4):1–17.

Potra F. A. and Wright S. J. (2000). *Interior-point methods.* Journal of Computational and Applied Mathematics, 124(1):281 – 302, ISSN: 0377-0427. Numerical Analysis 2000. Vol. IV: Optimization and Nonlinear Equations.

Prange L., Orliac E., Dach R., Arnold D., Beutler G., Schaer S., and Jäggi A. (2017). *CODE's five-system orbit and clock solution—the challenges of multi-GNSS data analysis.* Journal of geodesy, 91(4):345–360.

Priestley M. (1978). *Non-linear models in time series analysis.* Journal of the Royal Statistical Society. Series D (The Statistician), 27(3/4):159–176.

Protter M. H. and MorreyJr. C. B. (1970). *A first course in real analysis.* Springer, ISBN: 978-1-4615-9992-0, DOI: https://doi.org/10.1007/978-1-4615-9990-6.

Qian L., Burns A. G., Emery B. A., Foster B., Lu G., Maute A., Richmond A. D., Roble R. G., Solomon S. C., and Wang W. (2014). *The NCAR TIE-GCM: A community model of the coupled thermosphere/ionosphere system.* Modeling the ionosphere–thermosphere system, pages 73–83.

Qian N. (1999). *On the momentum term in gradient descent learning algorithms.* Neural networks : the official journal of the International Neural Network Society, 12(1):145—151, ISSN: 0893-6080, DOI: 10.1016/s0893-6080(98)00116-6, https://doi.org/10.1016/s0893-6080(98)00116-6.

Rao C. R. and Toutenburg H. (1999). *Linear Models Least Squares and Alternatives.* Springer-Verlag New York, ISBN: 978-0-387-98848-1, DOI: https://doi.org/10.1007/b98889.

Rao R. V. and Waghmare G. (2017). *A new optimization algorithm for solving complex constrained design optimization problems.* Engineering Optimization, 49(1):60–83.

Rawer K. (1988). *Synthesis of ionospheric electron density profiles with Epstein functions.* Advances in Space Research, 8(4):191 – 199, ISSN: 0273-1177, DOI: https://doi.org/10.1016/0273-1177(88)90239-6, http://www.sciencedirect.com/science/article/pii/0273117788902396.

Rawer K. and Argence E. (1954). *Origin of the Ionospheric E Layer.* Physical Review, 94(2):253.

Ray J. K. (2000). *Mitigation of GPS code and carrier phase multipath effects using a multi-antenna system.* University of Calgary.

Ray T., Singh H., Isaacs A., and Smith W. (2009). *Infeasibility Driven Evolutionary Algorithm for Constrained Optimization*, volume 198, pages 145–165. ISBN: 978-3-642-00618-0, DOI: 10.1007/978-3-642-00619-7_7.

Reeves G., Spence H. E., Henderson M., Morley S., Friedel R., Funsten H., Baker D., Kanekal S., Blake J., Fennell J., et al. (2013). *Electron acceleration in the heart of the Van Allen radiation belts.* Science, 341(6149):991–994.

Reigber C., Lühr H., and Schwintzer P. (2002). *CHAMP mission status.* Advances in Space Research, 30(2):129 – 134, ISSN: 0273-1177, DOI: https://doi.org/10.1016/S0273-1177(02)00276-4, http://www.sciencedirect.com/science/article/pii/S0273117702002764.

Richmond A. D. (1996). *Space weather research prompts study of ionosphere and upper atmospheric electrodynamics.* Eos, Transactions American Geophysical Union, 77(11):101–104, DOI: 10.1029/96EO00066, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96EO00066.

Richmond A. D., Ridley E. C., and Roble R. G. (1992). *A thermosphere/ionosphere general circulation model with coupled electrodynamics.* Geophysical Research Letters, 19(6):601–604, ISSN: 1944-8007, DOI: 10.1029/92GL00401, http://dx.doi.org/10.1029/92GL00401.

Rishbeth H. (1968). *On explaining the behavior of the ionospheric F region*. Reviews of Geophysics, 6(1):33–71.

Rishbeth H. (1998). *How the thermospheric circulation affects the ionospheric F2-layer*. Journal of Atmospheric and Solar-Terrestrial Physics, 60(14):1385–1402.

Rishbeth H. and Barron D. (1960). *Equilibrium electron distributions in the ionospheric F2-layer*. Journal of Atmospheric and Terrestrial Physics, 18(2-3):234–252.

Rishbeth H. and Garriott O. K. (1969). *Introduction to ionospheric physics*. Introduction to ionospheric physics.

Robbins (2007). *A Stochastic Approximation Method*. Annals of Mathematical Statistics, 22:400–407.

Robbins H. and Monro S. (1951). *A Stochastic Approximation Method*. Annals of Mathematical Statistics, 22(3):400–407, DOI: 10.1214/aoms/1177729586, https://doi.org/10.1214/aoms/1177729586.

Rocken C., Ying-Hwa K., Schreiner W. S., Hunt D., Sokolovskiy S., and McCormick C. (2000). *COSMIC system description*. Terrestrial Atmospheric and Oceanic Sciences, 11(1):21–52.

Roese-Koerner (2015). *Convex Optimization for Inequality Constrained Adjustment Problems*. PhD dissertation, Schriftenreihe des Instituts fuer Geodaesie und Geoinformation.

Roese-Koerner L., B D., N S., and D S. W. (2012). *A stochastic framework for inequality constrained estimation*. Journal of Geodesy, 86(11):1005–1018, ISSN: 1432-1394, DOI: 10.1007/s00190-012-0560-9, https://doi.org/10.1007/s00190-012-0560-9.

Roese-Koerner L., Devaraju B., and Schuh W. D. (2014). *Convex optimization under inequality constraints in rank-deficient systems*. Journal of Geodesy, 88(5):415–426, ISSN: 1432-1394, DOI: 10.1007/s00190-014-0692-1, https://doi.org/10.1007/s00190-014-0692-1.

Roese-Koerner L. and Schuh W. D. (2014). *Convex optimization under inequality constraints in rank-deficient systems*. Journal of geodesy, 88(5):415–426.

Roese-Koerner L. and Schuh W. D. (2016). *Effects of Different Objective Functions in Inequality Constrained and Rank-Deficient Least-Squares Problems*. In *VIII Hotine-Marussi Symposium on Mathematical Geodesy*, pages 325–331, Cham. Springer International Publishing, ISBN: 978-3-319-30530-1.

Rosenblatt M. (1971). *Nonlinear Representations in Terms of Independent Random Variables*, pages 161–182. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-65238-7, DOI: 10.1007/978-3-642-65238-7_6.

Rubin H. and Rubin J. E. (1985). *The Law of the Trichotomy*, volume 116 of *Studies in Logic and the Foundations of Mathematics*, pages 21 – 29. Elsevier, ISSN: 0049-237X, DOI: https://doi.org/10.1016/S0049-237X(08)70295-7.

Salcedo-Sanz S. (2017). *A review on the coral reefs optimization algorithm: new development lines and current applications*. Progress in Artificial Intelligence, 6(1):1–15.

Scargle J. D. (1981). *Studies in astronomical time series analysis. I-Modeling random processes in the time domain*. The Astrophysical Journal Supplement Series, 45:1–71.

Schaer S. (1999). *Mapping and predicting Earth's ionosphere using Global Positioning System*. PhD dissertation, Astronomical Institute, University of Bern.

Schmid R., Rothacher M., Thaller D., and Steigenberger P. (2005). *Absolute phase center corrections of satellite and receiver antennas*. GPS solutions, 9(4):283–293.

Schmidt M. (2007). *Wavelet modelling in support of IRI*. Advances in Space Research, 39(5):932 – 940, ISSN: 0273-1177, DOI: https://doi.org/10.1016/j.asr.2006.09.030, http://www.sciencedirect.com/science/article/pii/S0273117706005837.

Schmidt M. (2011). *Comparison of spherical harmonic and B spline models for the vertical total electron content*. Radio Science, 46(6), ISSN: 1944-799X, DOI: 10.1029/2010RS004609, http://dx.doi.org/10.1029/2010RS004609. RS0D11.

Schmidt M. (2015). *Inverse Modeling in Ionospheric Research*. In *2015 SIAM Conference on Mathematical and Computational Issues, Stanford Univ., Palo Alto, USA*.

Schmidt M., Dettmering D., and Seitz F. (2015). *Using B-Spline Expansions for Ionosphere Modeling*, pages 1–40. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-642-27793-1, DOI: 10.1007/978-3-642-27793-1_80-1, http://dx.doi.org/10.1007/978-3-642-27793-1_80-1.

Schreiner B., Hunt D., Rocken C., and Sokolovskiy S. (2003). *Radio Occultation Data Processing at the COSMIC Data Analysis and Archival Center (CDAAC)*, pages 536–544. Springer Berlin Heidelberg, Berlin, Heidelberg, ISBN: 978-3-540-38366-6, DOI: 10.1007/978-3-540-38366-6_73, https://doi.org/10.1007/978-3-540-38366-6_73.

Schuh, Brockmann J.-M., and Kargoll B. (2015). *Correlation analysis for long time series by robustly estimated autoregressive stochastic processes*. EGUGA, page 13050.

Schuh W.-D. (1996). *Least squares adjustment of high degree spherical harmonics*. In *Inverse Methods*, pages 276–283. Springer.

Schuh W.-D. (2003). *The processing of band-limited measurements; filtering techniques in the least squares context and in the presence of data gaps*. Space Science Reviews, 108(1-2):67–78.

Schumaker L. L. and Traas C. (1991). *Fitting scattered data on spherelike surfaces using tensor products of trigonometric and polynomial splines*. Numerische Mathematik, 60(1):133–144.

Schunk R. W. (1988). *A Mathematical Model of the Middle and High Latitude Ionosphere*, pages 255–303. Birkhäuser Basel, Basel, ISBN: 978-3-0348-6532-6, DOI: 10.1007/978-3-0348-6532-6_4, http://dx.doi.org/10.1007/978-3-0348-6532-6_4.

Seber G. (2015). *The Linear Model and Hypothesis A General Unifying Theory*. Springer International Publishing Switzerland, Cham Packages Mathematics and Statistics, ISBN: 978-3-319-21929-5, DOI: https://doi.org/10.1007/978-3-319-21930-1.

Silvestrin P., Bagge R., Bonnedal M., Carlstrom A., Christensen J., Hagg M., Lindgren T., and Zangerl F. (2000). *Spaceborne GNSS radio occultation instrumentation for operational applications*. In *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, pages 872–880.

Smirnova N., Ogloblina O., and Vlaskov V. (1988). *Modelling of the lower ionosphere*. Springer pure and applied geophysics, https://doi.org/10.1007/BF00879817.

Smith J. O. (2007). *Mathematics of the Discrete Fourier Transform (DFT)*. Stanford University Press.

Stankov, Jakowski N., Heise S., Muhtarov P., Kutiev I., and Warnant R. (2003). *A new method for reconstruction of the vertical electron density distribution in the upper ionosphere and plasmasphere*. Journal of Geophysical Research: Space Physics, 108(A5):n/a–n/a, ISSN: 2156-2202, DOI: 10.1029/2002JA009570, http://dx.doi.org/10.1029/2002JA009570. 1164.

Stankov. S. and Jakowski N. (2006). *Topside ionospheric scale height analysis and modelling based on radio occultation measurements*. Journal of Atmospheric and Solar-Terrestrial Physics, 68(2):134 – 162.

Stavroudis O. (2006). *Fermat's Principle and the Variational Calculus*, pages 1–13. John Wiley and Sons, Ltd, ISBN: 978-3-527-60817-1, DOI: 10.1002/3527608176.

Steigenberger P., Fritsche M., Dach R., Schmid R., Montenbruck O., Uhlemann M., and Prange L. (2016). *Estimation of satellite antenna phase center offsets for Galileo*. Journal of geodesy, 90(8):773–785.

Steigenberger P., Hugentobler U., Loyer S., Perosanz F., Prange L., Dach R., Uhlemann M., Gendt G., and Montenbruck O. (2015). *Galileo orbit and clock quality of the IGS Multi-GNSS Experiment*. Advances in space research, 55(1):269–281.

Stephenson G. (1971). *Inequalities and Optimal Problems in Mathematics and the Sciences*. A Longman text. Longman, ISBN: 978-0-582-44423-2, https://books.google.de/books?id=1mcZAQAAIAAJ.

Stewart G. W. (1995). *Theory of the Combination of Observations Least Subject to Errors, Part One, Part Two, Supplement*. SIAM, DOI: 10.1137/1.9781611971248.ch1.

Stoica P. and Nehorai A. (1989). *On multistep prediction error methods for time series models*. Journal of Forecasting, 8(4):357–368.

Stolle C., Floberghagen R., Lühr H., Maus S., Knudsen D. J., Alken P., Doornbos E., Hamilton B., Thomson A. W. P., and Visser P. N. (2013). *Space Weather opportunities from the Swarm mission including near real time applications.* Earth, Planets and Space, 65(11):17, ISSN: 1880-5981, DOI: 10.5047/eps.2013.10.002, https://doi.org/10.5047/eps.2013.10.002.

Stollnitz E. J., DeRose A., and Salesin D. H. (1995). *Wavelets for computer graphics: a primer. 1.* IEEE computer graphics and applications, 15(3):76–84.

Sutskever I., Martens J., Dahl G., and Hinton G. (2013). *On the importance of initialization and momentum in deep learning.* In Dasgupta S. and McAllester D. (Eds.), *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 1139–1147, Atlanta, Georgia, USA. PMLR.

Teunissen P. (1995). *The least-square ambiguity decor-ˆ iation adjustment: A method for fast GPS ambigu-ity estimation.* J. Geodesy, 70:65–82.

Teunissen P. and Odijk D. (2003). *Rank-defect integer estimation and phase-only modernized GPS ambiguity resolution.* Journal of Geodesy, 76(9-10):523–535.

Teunissen P. J. (2017). *Carrier Phase Integer Ambiguity Resolution*, pages 661–685. Springer International Publishing, Cham, ISBN: 978-3-319-42928-1, DOI: 10.1007/978-3-319-42928-1_23, https://doi.org/10.1007/978-3-319-42928-1_23.

Titheridge J. (1985). *Ionogram analysis with the generalised program POLAN, Rep. UAG-93, World Data Cent. A for Sol.* Terr. Phys., Boulder, Colo.

Todd M. J., Toh K. C., and Tütüncü R. H. (1998). *On the Nesterov–Todd Direction in Semidefinite Programming.* SIAM Journal on Optimization, 8(3):769–796, DOI: 10.1137/S105262349630060X, https://doi.org/10.1137/S105262349630060X.

Tsai L.-C., Cheng K.-C., and Liu C. (2011). *GPS radio occultation measurements on ionospheric electron density from low Earth orbit.* Journal of Geodesy, 85(12):941–948.

Tsai L. C. and Su (2018). *Global morphology of ionospheric sporadic E layer from the FormoSat-3/COSMIC GPS radio occultation experiment.* GPS Solutions, 22, https://doi.org/10.1007/s10291-018-0782-2.

Tsai L. C., Su S., and C L. (2016). *Ionospheric electron density profiling and modeling of COSMIC follow-on simulations.* J Geod, 90:129–142.

Tütüncü R. H., Toh K. C., and Todd M. J. (2003). *Solving semidefinite-quadratic-linear programs using SDPT3.* Mathematical Programming, 95(2):189–217, ISSN: 1436-4646, DOI: 10.1007/s10107-002-0347-5, https://doi.org/10.1007/s10107-002-0347-5.

Ulrych T. J. (1985). *Spectral Analysis and Time Series Models: A Geophysical Perspective.* In *Maximum-Entropy and Bayesian Methods in Inverse Problems*, pages 243–272. Springer.

Verhagen S. (2005). *The GNSS Integer Ambiguities: Estimation and Validation, TU Delft.* Delft University of Technology.

Verhagen S. and Teunissen P. (2006). *New global navigation satellite system ambiguity resolution method compared to existing approaches.* Journal of Guidance, Control, and Dynamics, 29(4):981–991.

Vielberg K., Forootan E., Lück C., Löcher A., Kusche J., and Börger K. (2018). *Comparison of accelerometer data calibration methods used in thermospheric neutral density estimation.* In *Annales Geophysicae*, volume 36, pages 761–779. Copernicus GmbH.

Wallis J. (1685). *A treatise of algebra, both historical and practical.* Philosophical Transactions of the Royal Society of London, 15(173):1095–1106, DOI: 10.1098/rstl.1685.0053, https://royalsocietypublishing.org/doi/abs/10.1098/rstl.1685.0053.

Weimer D. (2005). *Improved ionospheric electrodynamic models and application to calculating Joule heating rates.* Journal of Geophysical Research: Space Physics, 110(A5).

White H. (1990). *Least Squares*, pages 118–125. Palgrave Macmillan UK, London, ISBN: 978-1-349-20865-4, DOI: 10.1007/978-1-349-20865-4_15, https://doi.org/10.1007/978-1-349-20865-4_15.

Wilkes M. V. (1997). *Sir Edward Appleton and early ionosphere research.* Notes and Records of the Royal Society of London, 51(2):281–290.

Williams R. J. and Zipser D. (1989). *A Learning Algorithm for Continually Running Fully Recurrent Neural Networks.* Neural Comput., 1(2):270–280, ISSN: 0899-7667, DOI: 10.1162/neco.1989.1.2.270, https://doi.org/10.1162/neco.1989.1.2.270.

Wilson D. R. and Martinez T. R. (2001). *The need for small learning rates on large problems.* In *IJCNN'01. International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*, volume 1, pages 115–119 vol.1. DOI: 10.1109/IJCNN.2001.939002.

Xu P., Shen Y., Fukuda Y., and Liu Y. (2006). *Variance Component Estimation in Linear Inverse Ill-posed Models.* Journal of Geodesy, 80:69–81, DOI: 10.1007/s00190-006-0032-1.

Yue X., Schreiner W. S., Pedatella N., Anthes R. A., Mannucci A. J., Straus P. R., and Liu J.-Y. (2014). *Space Weather Observations by GNSS Radio Occultation: From FORMOSAT-3/COSMIC to FORMOSAT-7/COSMIC-2.* Space Weather, 12(11):616–621, DOI: 10.1002/2014SW001133, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014SW001133.

Yue X., Schreiner W. S., Pedatella N. M., and Kuo Y.-H. (2016). *Characterizing GPS radio occultation loss of lock due to ionospheric weather.* Space Weather, 14(4):285–299.

Zeiler M. D. (2012). *ADADELTA: An Adaptive Learning Rate Method.*

Zeilhofer C. (2008). *Multi-dimensional B-spline Modeling of Spatio-temporal Ionospheric Signals.* Deutsche Geodätische Kommission, Reihe A, 123, Muenchen, ISSN: 0938-2836.

Zeitler L., Corbin A., Vielberg K., Rudenko S., Löcher A., Bloßfeld M., Schmidt M., and Kusche J. (2021). *Scale factors of the thermospheric density: A comparison of satellite laser ranging and accelerometer solutions.* Journal of Geophysical Research: Space Physics, 126(12):e2021JA029708.

Zolesi B. and Cander L. R. (2014). *Ionospheric Prediction and Forecasting.* Springer Geophysics, DOI: 10.1007/978-3-642-38430-1_2.

Appendices

A Dynamic range and hypervolume

Dynamic range computation : example

Given a set of constraints at $\{\varphi, \lambda\}$

$$\begin{aligned} 0.02 \leq N_m^{F_2}(\varphi,\lambda) &\leq 2.5 \text{ EDU} \\ 220 \leq h_m^{F_2}(\varphi,\lambda) &\leq 480 \text{ km} \\ 5 \leq H^{F_2}(\varphi,\lambda) &\leq 180 \text{ km}, \end{aligned} \tag{A.1}$$

the three corresponding dynamic ranges are computed as

$$\begin{aligned} \mathcal{D}^r_{N_m^{F_2}}(\varphi,\lambda) &= |2.5 - 0.02| = 2.48 \text{ EDU} \\ \mathcal{D}^r_{h_m^{F_2}}(\varphi,\lambda) &= |480 - 220| = 260 \text{ km} \\ \mathcal{D}^r_{H^{F_2}}(\varphi,\lambda) &= |180 - 5| = 175 \text{ km}. \end{aligned} \tag{A.2}$$

with the corresponding normalized dynamic ranges

$$\begin{aligned} \breve{\mathcal{D}}^r_{N_m^{F_2}}(\varphi,\lambda) &= 0.992 \text{ and } \mathcal{N}_{N_m^{F_2}} = 2.5 \text{ EDU} \\ \breve{\mathcal{D}}^r_{h_m^{F_2}}(\varphi,\lambda) &= 0.928 \text{ and } \mathcal{N}_{h_m^{F_2}} = 280 \text{ km} \\ \breve{\mathcal{D}}^r_{H^{F_2}}(\varphi,\lambda) &= 0.921 \text{ and } \mathcal{N}_{H^{F_2}} = 190 \text{ km} \end{aligned} \tag{A.3}$$

resulting in a hypervolume $\mathcal{V}(\varphi,\lambda) = 0.02389$.

It follows that:

- for a given pair of upper and lower bound constraint on any one key parameter $\kappa \in \mathcal{K}_1$ at $\{\varphi,\lambda\}$, there is one corresponding dynamic range $\mathcal{D}^r_\kappa(\varphi,\lambda)$.
- for a given pair of upper and lower bound constraint on any one key parameter $\kappa \in \mathcal{K}_1$ at each of the $L = \varphi_{max} \cdot \lambda_{max}$ latitude-longitude pairs, there are L corresponding dynamic ranges $\mathcal{D}^r_\kappa(\varphi_m,\lambda_j) \;\; \forall \; m = \{1, \cdots, \varphi_{max}\}; \; j = \{1, \cdots, \lambda_{max}\}$.
- for a given pair of upper and lower bound constraint on q key parameters κ_p at each of the L latitude-longitude pairs, theer are $L \cdot q$ corresponding dynamic ranges $\mathcal{D}^r_{\kappa_p}(\varphi_m,\lambda_j)$ $\forall \; p = \{1, \cdots, q\}; \; m = \{1, \cdots, \varphi_{max}\}; \; j = \{1, \cdots, \lambda_{max}\}$ and L hypervolumes $\mathcal{V}(\varphi_m,\lambda_j)$.

B Constraints rationalization

Qualitatively, it has been shown in the numerical evaluations (see section 6.3) that the constraints with a large dynamic range, result in a corresponding large positive slack. In other words, this means, overly-conservative contraint bounds were imposed on those key parameters. For near real-time electron density modelling, the number of constraints and unknown parameters need to be chosen to satisfy the operational requirements on the spatial and temporal resolution, accuracy, computational performance, continuity during space weather events etc. For example, consider the constraint

$$220 \leq h_m^{F_2}(\varphi,\lambda) \leq 480 \;\; \text{km} \quad \text{for } \varphi = 10^\circ E, \lambda = 0^\circ \text{ at } 14:00 \tag{A.4}$$

is conservative because at the day-time equatorial latitude, the $h_m^{F_2}$ will be larger than 220 km. Fig. 6.6 shows a distribution of IRI model derived $h_m^{F_2}$ at $\varphi = 10^\circ E, \lambda = 0^\circ$, under nominal ionospheric

conditions over a period of 1 year[4] and, in retrospect, the dynamic range of constraints, such as (A.4), could be reduced in the electron density modelling problem. However, if the associated $h_m^{F_2}$ slack variable was only marginally positive (for e.g. between 0 - 20 km), despite being inactive, the constraint (A.4) is then regarded as $\gamma-$active (see chapter 4, definition 4.4.1).

> Rationalization is the process of identifying and reducing inactive constraints in the optimization problem.

Upper or lower bound to be rationalized ?

Consider the key parameter $N_m^{F_2}$, for which a larger than nominal value ($N_m^{F_2} \geq 2.5$ EDU) is possible, e.g. during severe space weather events, but a negative value (i.e. $N_m^{F_2} < 0$) is not. Therefore, an inequality constraint on the lower bound of $N_m^{F_2}$ is necessary for the for physical realism. Similarly, if both F_1 and F_2 layer peak heights are to be modelled, then the set $\mathcal{K}_1$ would include both the $h_m^{F_2}$ and $h_m^{F_1}$. However, from a physical perspective, the F_2 layer should be localized at altitudes above that of the F_1 or in other words, the inequality

$$h_m^{F_2} > h_m^{F_1} \tag{A.5}$$

holds. Therefore, the lower bound of $h_m^{F_2}$ and the upper bound of $h_m^{F_1}$ layers become important considerations for rationalization, without which the physical realism (A.5) could be violated.

The following possibilities to reduce the computational time may be considered in the subsequent studies and in near real time applications.

1. A minimal set of unknown Chapman key parameters to be modelled and the other remaining key parameters must be given with a high accuracy.
2. Limiting the total number of constraints: For each constraint, a pair of Lagrange multiplier and slack variable need to be added to the unknown parameters.
3. Choice of feasible region: A large feasible region would mean evaluation of a large number of potential candidate solutions. This problem is further related to two important aspects of the problem: choice of constraint bounds and the convexity of the objective function.
   - A dedicated sub-routine is recommended for determining the constraint bounds.
   - A convexity check of the objective function is also recommended.
4. Initial values: When the optimization algorithms starts with initial values from within the feasible region, then the computation times are relatively lower compared to a completely random initialization.
5. In an epoch-by-epoch sequential processing, the use of effective prediction technique (described in section 4.7.4) to determine the initial value for the unknown parameters also reduce the computation time.
6. Use of different B-spline levels for the key parameters: Some of the parameters are not changing spatially as much as others. For e.g. if F_2 and E layers are simultaneously considered, then the key parameters of latter can be modelled with relatively smaller B-spline levels compared to that of the former. This is owing to the smaller E layer spatial decorrelation of the key parameters compared to that of the F_2 layer (see Bilitza et al. (2011)).

[4] Measurements at 14:00 local time at the given latitude and longitude for the year 2014.

C Numerical evaluations E3 - E12

C.1 Evaluation E3: : Varying $N_m^{F_2}$ lower bound

In order to allow better readability, the evaluations E3 - E12 were removed from the main body of this thesis and put to this Appendix. As a general remark, the figure captions presented in this chapter will not provide the interpretation details. Instead the detailed interpretation of the presented results will be included directly in the respective sections of this chapter where the numerical evaluation is introduced, defined and subsequently the results are tabulated as well as discussed. This is mainly to avoid long sentences as well as repetitions in the figure captions. Furthermore, when multiple results are presented together in one figure, the phrase "panel" refers to one specific sub-figure within the figure.

In E3, the goal is to continue from Chapter 6 and analyse the impact of a reduced $N_m^{F_2}$ dynamic range on the corresponding slack variable, with the lower bound of $N_m^{F_2}$ increased by 5.5%[5] compared to E1[6] (see Table 1).

The increased $N_m^{F_2}$ lower bound and that estimated from the corresponding B-spline coefficients are shown in the top-left and mid panels (a) and (b) of Fig. 1 respectively. The estimated $N_m^{F_2}$ is lower than (and thus violates) the corresponding lower bound for 12% of the $N_m^{F_2}$ constraints and >70% of these occur in the night side.

Table 1: Configuration for numerical evaluation E3.

	variable or parameter	value
Input	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.47 (5.5% higher LB than E1)
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.8 (same as that used in E1)
	average hypervolume $\bar{\mathcal{V}}$	0.0620
Output	minimum $N_m^{F_2}$ slack	-0.23 EDU
	% of $N_m^{F_2}$ constraints violated	12

C.2 Evaluation E4: Using DS2a observation dataset

In E4, the goal is to analyse the impact of using a different observation dataset on the $N_m^{F_2}$ slack variable. The total number of unknown parameters is the same but the observation dataset DS2d is used in E1 to E3, whereas DS2a is used in E4 respectively (see Fig. 16 and Table 13). The other inputs including the normalized average dynamic range and the hypervolume are the same as in E1 (see Table 2).

The left top panel of Fig. 2 shows the $N_m^{F_2}$ lower bound and that calculated from the corresponding estimated B-spline coefficients in the mid panels respectively. The resulting $N_m^{F_2}$ slack is shown on the panels (c) and (d). The difference between the lower bound $N_m^{F_2}$ and that estimated is the "negative of slack variable" (see Eq. (5.22), where the constraint was multiplied by -1 to transform it to the convex form). Therefore, all $N_m^{F_2}$ slack variables are still positive and the bottom mid (e) and right (f) panels show that all $N_m^{F_2}$ constraints remain inactive and hence no violations occur. At the same time, the global minimum of the $N_m^{F_2}$ slack variable was similar (0.005 and 0.01 EDU respectively) between E1 and E4 despite using different observation datasets[7].

[5] The justification for increasing the lower bound exactly by 5.5% is as follows: Different values of increasing the lower bound were tried in the range of 1 to 10% and the value of 5.5% was then chosen based on the need for maintaining a consistency among the hypervolume in the evaluations E1 to E8; see Figs. 9 and 10

[6] Keeping the other inputs same as in E1

[7] In order to make a representative assessment, E1 results in Fig. 6.11(a) showed the upper bound and in panel (a) of Fig. 6.10 the corresponding lower bound. It would be duplicate to include both the upper as well as the lower bound constraint maps in all the evaluations.

Table 2: Configuration for numerical evaluation E4.

	variable or parameter	value
Input	observation	dataset DS2a
	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.5 (same as that in E1)
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.80 (same as that in E1)
	average hypervolume $\bar{\mathcal{V}}$	0.066
Output	minimum $N_m^{F_2}$ slack	0.02 EDU
	% of $N_m^{F_2}$ constraints violated	0

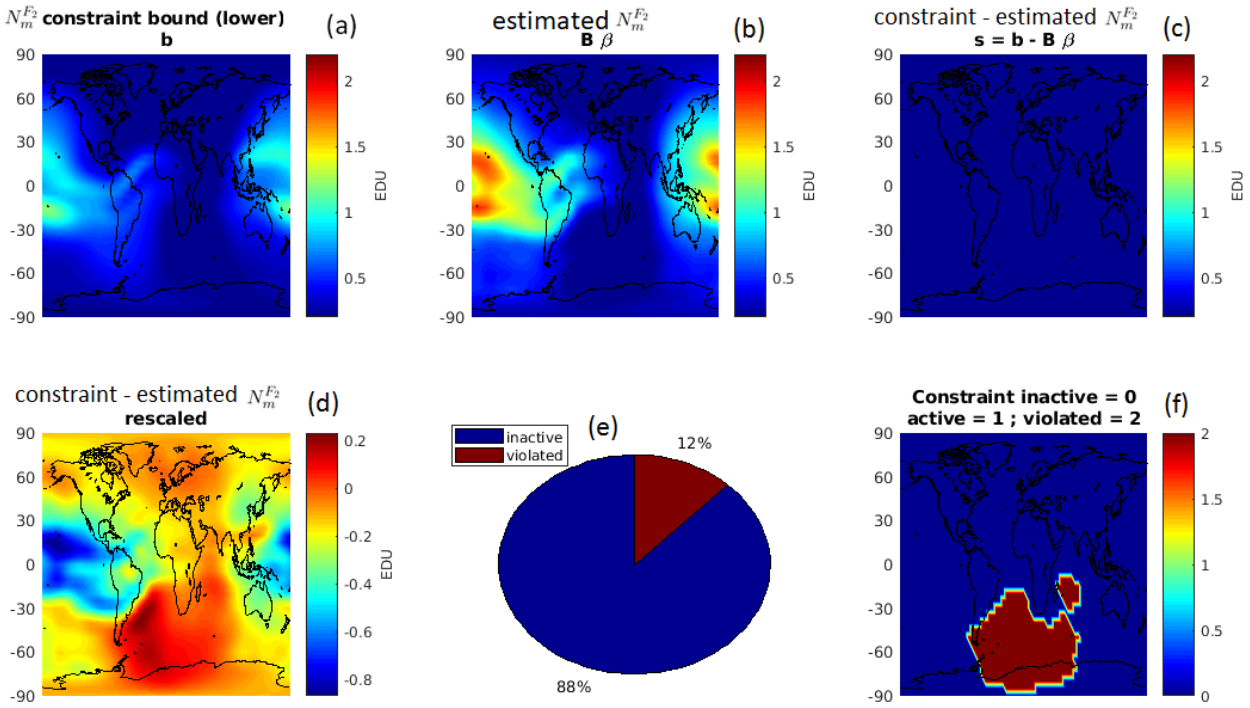


Figure 1: Evaluation E3 : impact of $N_m^{F_2}$ lower bound constraint

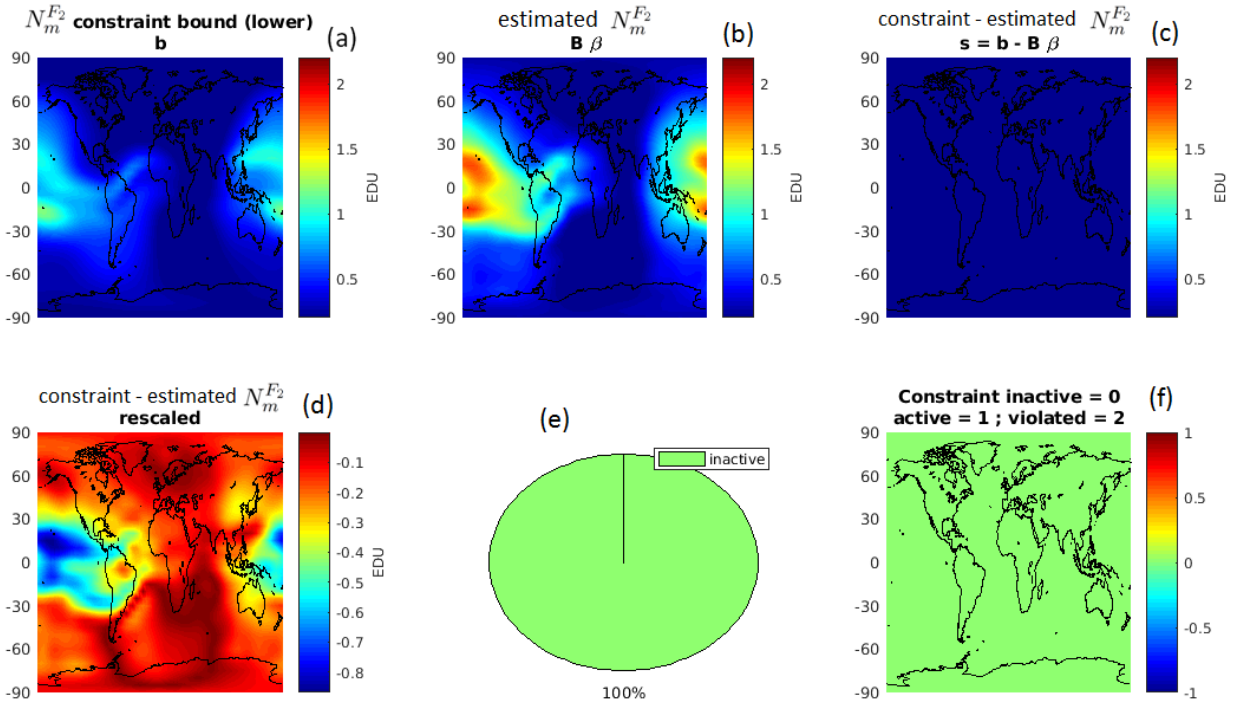


Figure 2: Evaluation E4: Slack and modelled $N_m^{F_2}$ with observation dataset DS2a

Summary of evaluations E1 - E4

There are two possibilities for reducing the dynamic range (1) by reducing the upper bound and (2) by increasing the lower bound.

The former is shown in E2, where the upper bound was reduced by 12% and the latter in E3, where the lower bound was increased by 5.5%. In both cases, a reduction in the average $N_m^{F_2}$ dynamic range, compared to that in E1, results in the corresponding constraint violations (see Fig 9). The $N_m^{F_2}$ slack values for E1 to E4 are are shown in Fig. 3.
In the evaluations E6 and E7, in the following section, $h_m^{F_2}$ constraint bounds will be varied with respect to a reference scenario E5.

C.3 Evaluation E5: Reference scenario

The inputs for E5 is exactly the same as those for E1 as repeated in 3 for easier readability. In order to distinguish between the $N_m^{F_2}$ and the $h_m^{F_2}$ analysis, the evaluation E5 is separately included here. Specifically, now the results of $h_m^{F_2}$ slack will be shown. The top panel of Fig. 4 shows the $h_m^{F_2}$ upper bound and that calculated from the corresponding estimated B-spline coefficients in the left and mid panels (a) and (b) respectively. The top-right sub-figure (c) shows the $h_m^{F_2}$ slack, or equivalently the difference between the top left and top mid panels (a) and (b) respectively, computed using Eq. (6.13). The bottom panel left (d) shows the same slack variable (as the top-right panel (c)) but in a modified scale. It is seen that all $h_m^{F_2}$ slack variables are positive, indicating that the KKT condition (4.72b) is satisfied. The bottom mid (e) and right (f) panels show that all $h_m^{F_2}$ constraints are inactive and there are no violations.

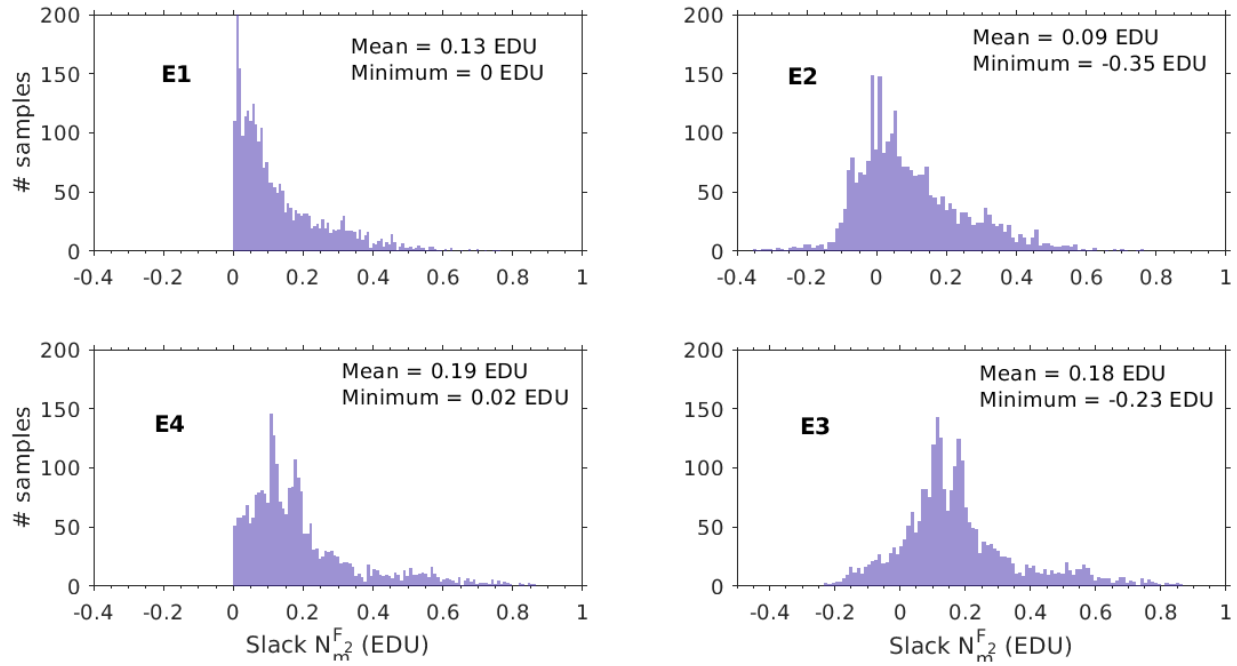


Figure 3: $N_m^{F_2}$ slack for numerical evaluations E1 to E4. Note that in E1, the minimum slack was 0.005 EDU, as shown in Table 6.4, which is rounded off to 0 EDU in the legend.

Table 3: Configuration for numerical evaluation E5.

	variable or parameter	value
Input	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.5 (same as that in E1)
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.8 (same as that in E1)
	average hypervolume $\bar{\mathcal{V}}$	0.066
Output	minimum $h_m^{F_2}$ slack	0.02 km
	% of $h_m^{F_2}$ constraints violated	0

C.4 Evaluation E6: Varying $h_m^{F_2}$ upper bound

In E6, the goal is to analyse the impact of a reduced dynamic range of $h_m^{F_2}$ on the corresponding slack variable. Specifically, the upper bound of $h_m^{F_2}$ is reduced by 15%[8] compared to E5 with the other inputs remaining same as in evaluation E5 (see Table 4).

The reduced $h_m^{F_2}$ upper bound and that calculated from the corresponding estimated B-spline coefficients are shown in Fig. 5 in the left and mid panels (a) and (b). The calculated $h_m^{F_2}$ exceeds (and thus violates) 18% of the corresponding $h_m^{F_2}$ upper bound constraints.

C.5 Evaluation E7: Varying $h_m^{F_2}$ lower bound

In the evaluation E7, the goal is, once again, to analyse the impact of the reduced dynamic range of $h_m^{F_2}$ on the corresponding slack variable. Specifically, the lower bound of $h_m^{F_2}$ is increased by 15% compared to E5, with the other inputs remaining the same as in E5 (see Table 5).

The $h_m^{F_2}$ lower bound and that calculated from the corresponding estimated B-spline coefficients are shown in Fig. 6 in the left and mid panels (a) and (b). The estimated $h_m^{F_2}$ is lower than (and thus violates) 5% of the corresponding lower bound constraints by a maximum of 22 km and >75%

[8]The justification for choosing the reduction of 15% in $h_m^{F_2}$ UB is similar to that presented in the evaluation E2.

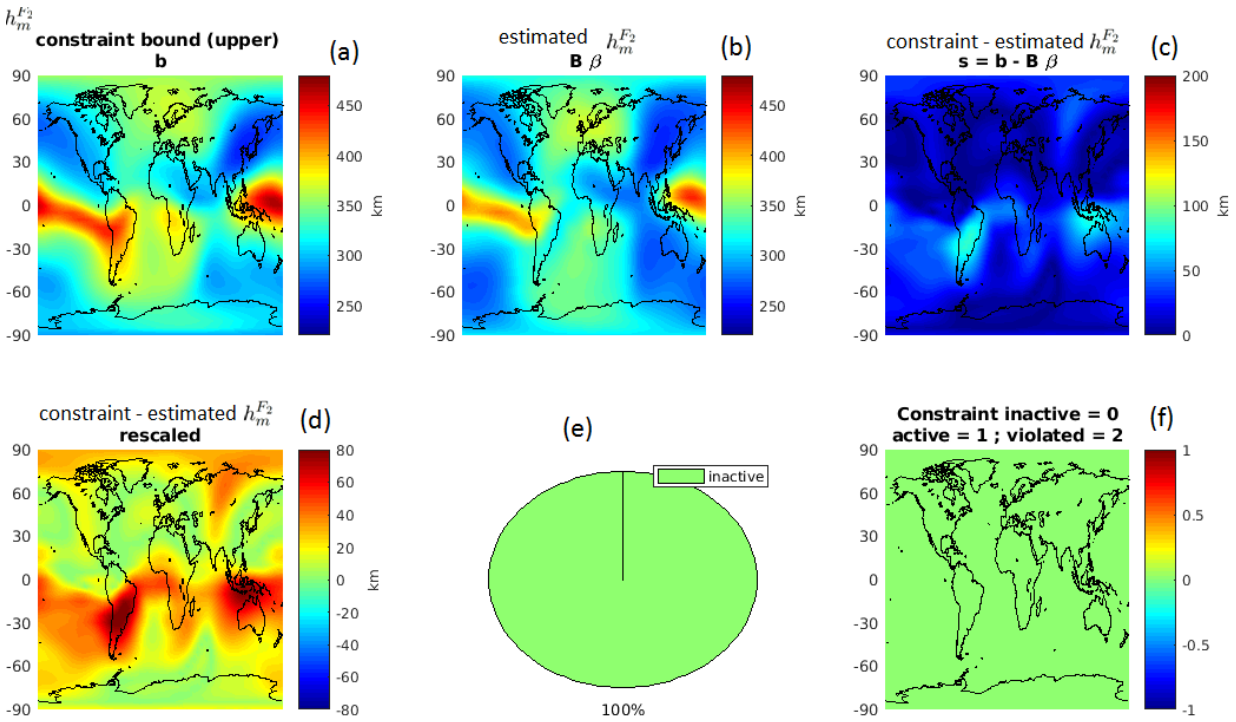


Figure 4: Evaluation E5: Slack and modelled $h_m^{F_2}$

Table 4: Configuration for numerical evaluation E6.

	variable or parameter	value
Input	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.5 (same as that in E5)
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.72 (15% lower UB than that in E5)
	average hypervolume $\bar{\mathcal{V}}$	0.056
Output	minimum $h_m^{F_2}$ slack	-19.9 km
	% of $h_m^{F_2}$ constraints violated	18

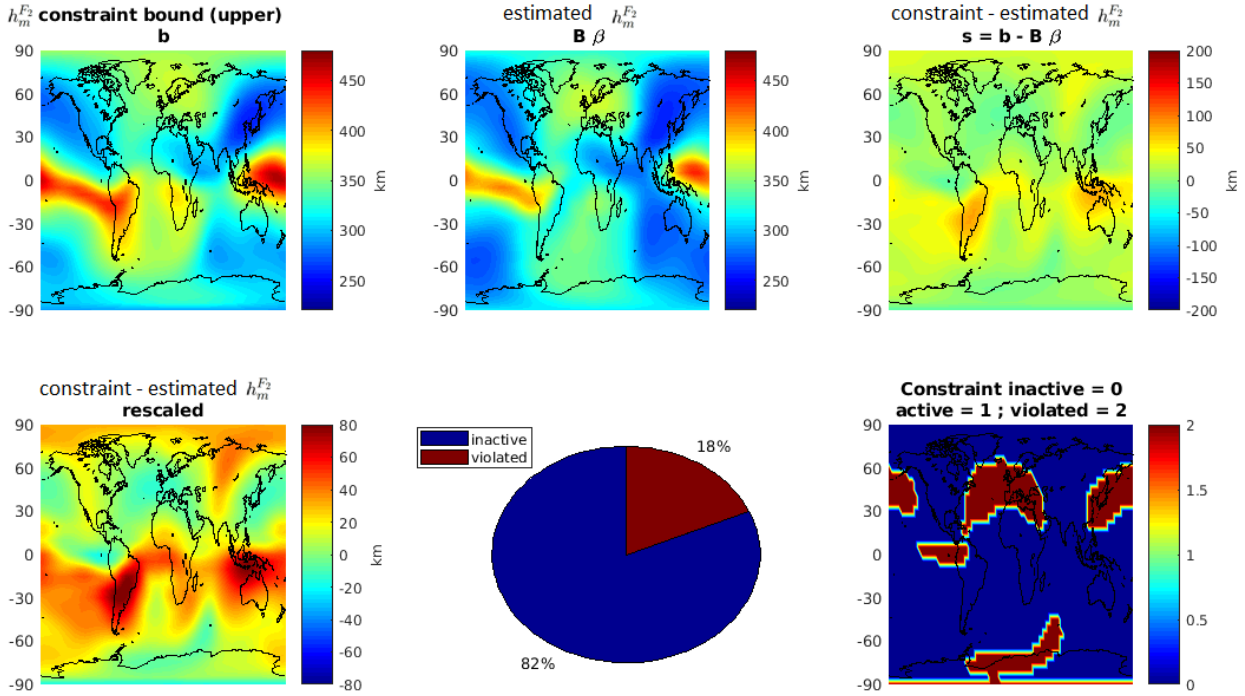


Figure 5: Evaluation E6: Impact on slack and modelled $h_m^{F_2}$ due to varying upper bound.

of these occur in the equatorial region between 0 - 30° South during the 05:00 - 07:00 and 17:00 - 19:00 (around the local sunrise and sunset times).

The physical interpretation of the violations in E7 is described below. Physical and chemical processes, such as ionoization, recombination, field aligned currents, thermospheric winds are involved in the steady-state[9] settlement of the electrons at a given altitude (see e.g. Gulyaeva (1985)). More generally, the typical daily behaviour of $h_m^{F_2}$ includes a reducing trend from the midnight to sunrise (approximately between midnight and 06:00 local time) and from the afternoon around 14:00 to sunset (approximately 18:00) (see e.g. Rishbeth (1998), Altadill et al. (2013)). When the $h_m^{F_2}$ lower bound in E7 is increased globally by 15%, with respect to the corresponding values in E5, then the modelled $h_m^{F_2}$ around sunrise and sunset are violated. Similar to E3, where the $N_m^{F_2}$ lower bound was increased by 5%, in E7 the corresponding 15% increase is also seen to be "excessive", with respect to the nominal conditions of E5. On the other hand, generally from sunrise to noon, due to the increase in ionization (see e.g. Gulyaeva (1987), Obrou et al. (2003)), $h_m^{F_2}$ increases[10]. Therefore, a 15% increase in the $h_m^{F_2}$ lower bound with respect to E5 causes qualitatively less violations during those times compared to the night.

Table 5: Configuration for numerical evaluation E7.

	variable or parameter	value
Input	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.5 (same as that in E5)
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.72 (15% larger LB than that in E5)
	average hypervolume $\bar{\mathcal{V}}$	0.056
Output	minimum $h_m^{F_2}$ slack	-25.91 km
	% of $h_m^{F_2}$ constraints violated	5

[9]Steady-state is defined as a quasi-equilibrium between gravitational, electrodynamic and hydrostatic forces acting on the charged particles, such as electrons. (see e.g. Rishbeth (1998))

[10]at different rates though at the different latitudes.

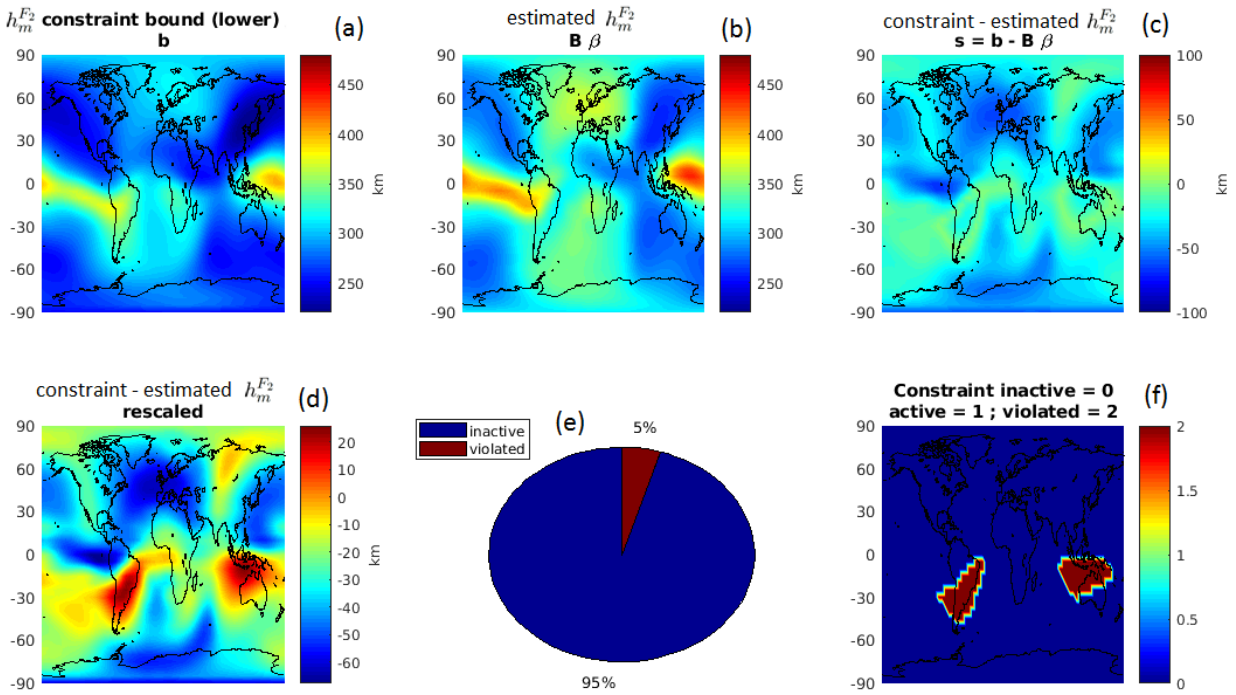


Figure 6: Evaluation E7: Impact on slack and modelled $h_m^{F_2}$ due to varying upper bound.

C.6 Evaluation E8: Using DS2a observation dataset

In the evaluation E8, the goal is to analyse the impact of using the DS2a observation dataset on the $h_m^{F_2}$ slack variable. The other inputs are the same as in E5 (see Table 6).

The top panel of Fig. 7 shows the $h_m^{F_2}$ lower bound and that calculated from the corresponding estimated B-spline coefficients in the left and mid panels (a) and (b) respectively. It is seen that the KKT condition (4.72b) is satisfied for this solution, shown by the positive $h_m^{F_2}$ slack variable in the panels (c) and (d). The difference between the $h_m^{F_2}$ lower bound and that estimated is the "negative of slack variable" (as explained already in evaluation E3 and also explained in Eq. (5.22)). Therefore, all $h_m^{F_2}$ slack variables are positive and the bottom middle (e) and right (f) panels show that all $h_m^{F_2}$ constraints remained inactive.

Table 6: Configuration for numerical evaluation E8.

	variable or parameter	value
Input	Observation dataset	DS2a
	norm. avg. DR $\bar{\mathcal{D}}^r_{N_m^{F_2}}$	0.5 (same as that in E5)
	norm. avg. DR $\bar{\mathcal{D}}^r_{h_m^{F_2}}$	0.8 (same as that in E5)
	average hypervolume $\bar{\mathcal{V}}$	0.066
Output	minimum $h_m^{F_2}$ slack	0.02 km
	% of $h_m^{F_2}$ constraints violated	0

Summary of evaluations E5-E8

The $h_m^{F_2}$ slack variable is shown to be sensitive to the corresponding lower and upper bounds in the evaluations E6 and E7. Similar to the case of $N_m^{F_2}$, the impact of a different observation dataset was not found to be significant. In both E5 and E8, the minimum $h_m^{F_2}$ slack was 0.02 km. The $h_m^{F_2}$ slack values in E5 to E8 are shown in Fig. 8.

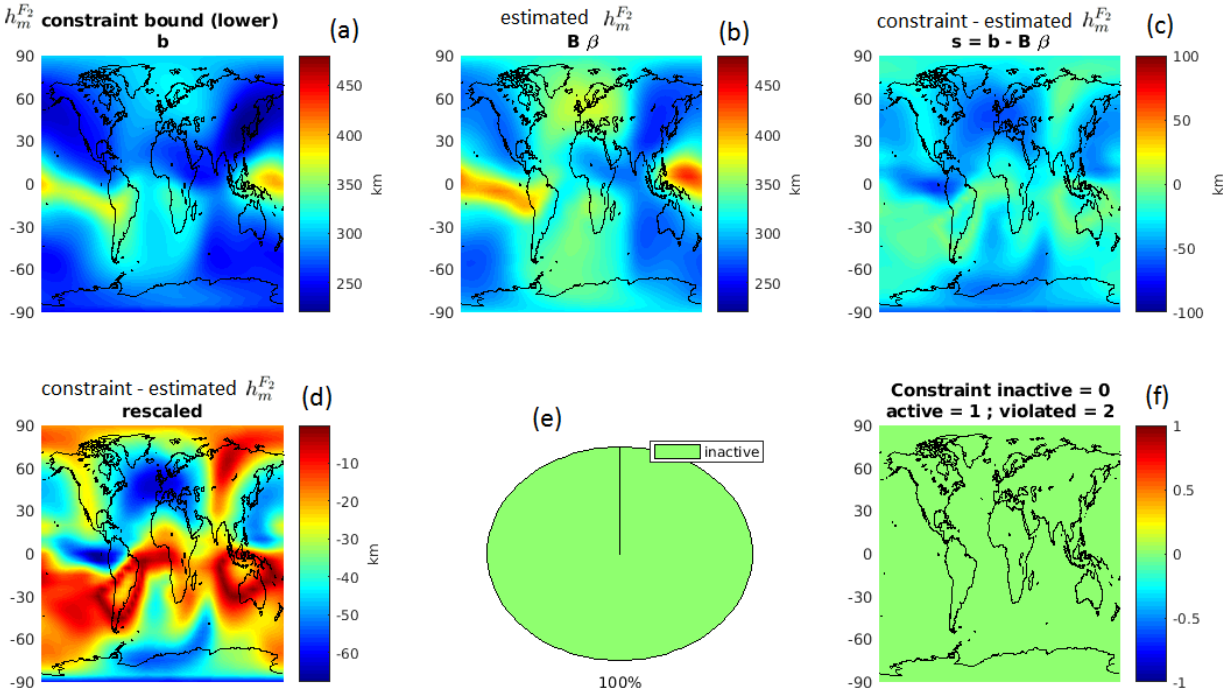


Figure 7: Evaluation E8: Slack and modelled $h_m^{F_2}$ with observation dataset DS2a.

C.7 Sub-evaluations E1a-E1d: Accuracy assessment and the impact of ionospheric layers

There are two goals of the evaluations E1a to E1d:

1. To analyse the accuracy of the estimated$N_m^{F_2}$ in evaluation E1.
2. To analyse the impact of the ionospheric bottom ($D-$, $E-$, F_1-) layers in the electron density modelling.

The evaluations E1 to E3 and E5 to E7 have used the DS2d observation dataset, for which there is no reference (or "truth") to compute the deviation[11] of the estimated key parameter, unlike with the simulated dataset DS1a. Therefore, an independent observation dataset from ionosonde stations (DS3) is used to determine the accuracy of the estimated$N_m^{F_2}$ with that obtained from four selected ionosonde stations. The estimated $N_m^{F_2}$ B-spline coefficients in the evaluation E1 are used to compute the estimated $N_m^{F_2}$ at four ionosonde station locations at 15:00 local time using Eq. (5.5). The difference (or deviation) of the estimated$N_m^{F_2}$ with respect to that obtained from the ionosonde stations, between 14:00 - 16:00, is computed and its mean over three consecutive days (10 - 12 March 2015) is shown in Fig. 11.

The sub-evaluations, E1a to E1d, corresponding to the inclusion of the ionospheric bottom layers (F_1-, $E-$ and $D-$) are described below. As already described in Eq. (5.16) the Chapman key parameters are partitioned into the two sets $\mathcal{K}_1$ and $\mathcal{K}_2$. The former contains the key parameters corresponding to the F_2 layer and plasmasphere to be modelled and the latter contains the given values of the remaining parameters. It has also been mentioned that the N_m^Q and h_m^Q, corresponding to the layer Q, in $\mathcal{K}_2$ are obtained from reliable sources (e.g. IRI model) and the respective scale heights of each layer from Eq. (6.3). In the evaluations E1a to E1c, the impact of given values for the F_1-, the $E-$ and the $D-$ layer given values in the subset $\mathcal{K}_2$ will be demonstrated.

As in all other evaluations, 5 key parameters in (6.1) are to be estimated and the known given values of the remaining key parameters in (6.2) are used.

[11] a measure of accuracy, defined in Eq. (A.9)

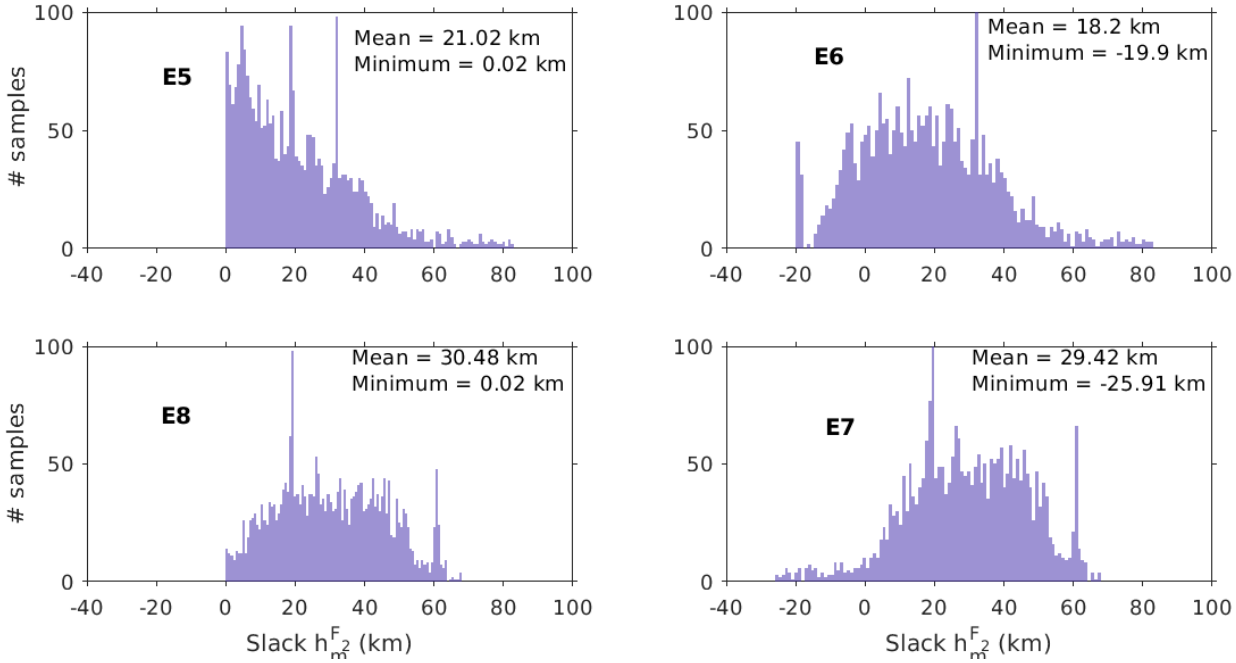


Figure 8: $h_m^{F_2}$ slack for the numerical evaluations E5 to E8

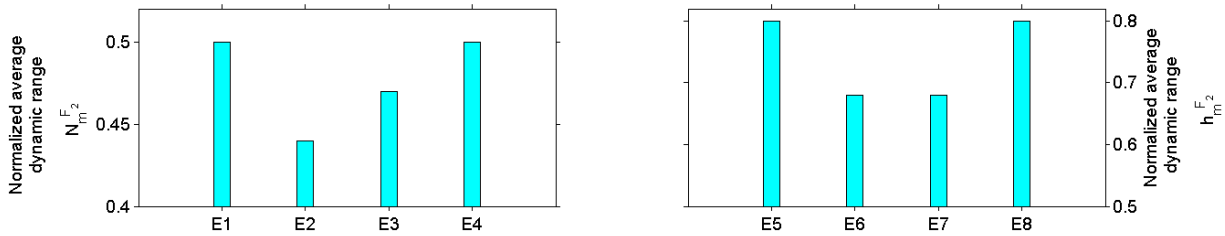


Figure 9: Average $N_m^{F_2}$ and $h_m^{F_2}$ dynamic ranges used in evaluations E1 - E8.

Evaluation E1a: F_2 layer and plasmasphere
In E1a, only the F_2 layer and plasmasphere are considered by ignoring the remaining layers and accordingly, it follows:

$$N_m^{F_1}(\varphi_m, \lambda_j) = 0 \qquad N_m^{E}(\varphi_m, \lambda_j) = 0 \qquad N_m^{D}(\varphi_m, \lambda_j) = 0 \tag{A.6}$$

are set in $\mathcal{K}_2$, $\forall\ m = \{1, \cdots, \varphi_{max}\}$; $j = \{1, \cdots, \lambda_{max}\}$. By setting the peak density of the F_1-, $E-$ and $D-$ layers to zero, they do not contribute to the total electron density of (5.1) [12].

Evaluation E1b: F_2-, F_1- layer and plasmasphere
In E1b, the F_1 layer is considered in addition to that in E1a, but ignoring the key parameters of the two remaining two ($E-$ and $D-$) layers and accordingly, we set

$$N_m^{E}(\varphi_m, \lambda_j) = 0 \qquad N_m^{D}(\varphi_m, \lambda_j) = 0, \tag{A.7}$$

in $\mathcal{K}_2$, $\forall\ m = \{1, \cdots, \varphi_{max}\}$; $j = \{1, \cdots, \lambda_{max}\}$. It shall be noted that the $N_m^{F_1}$ and $h_m^{F_1}$ are directly obtained from the IRI model for this work, whereas the corresponding H^{F_1} is assumed to be 10% of H^{F_2}[13].

[12] In other words, the ionosphere is assumed to be comprised of only F_2 layer and plasmasphere.
[13] This is only an approximation, for lack of a more precise information. Similar assumptions were also made by Liang et al. (2016)

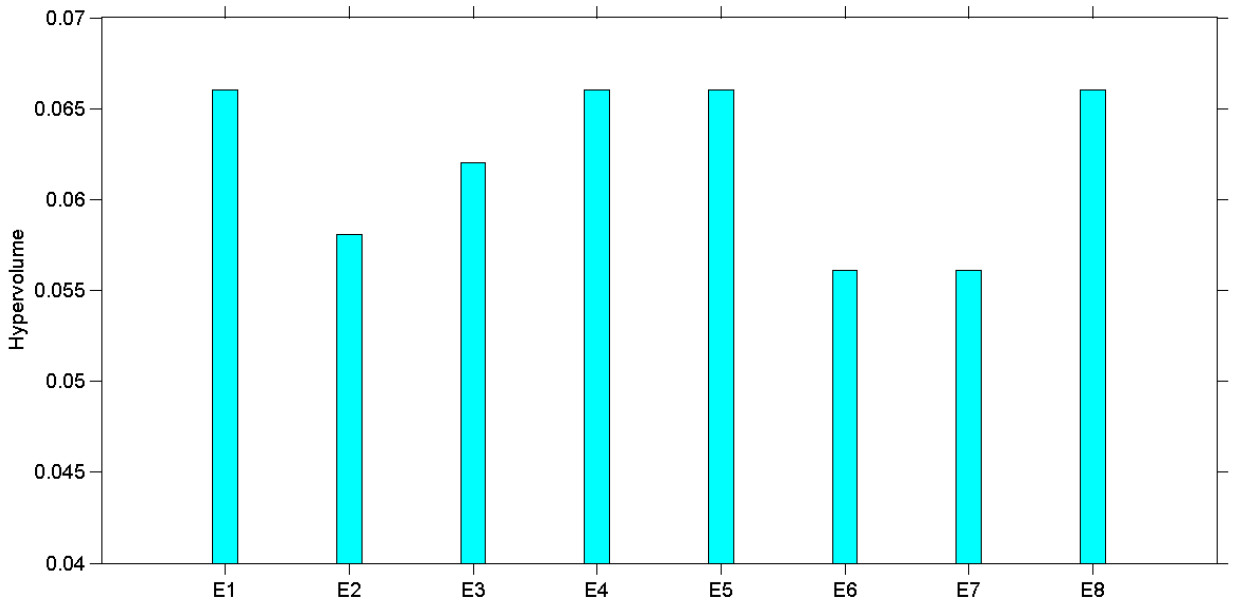


Figure 10: Hypervolume computed from the average dynamic ranges in evaluations E1 - E8.

Evaluation E1c: F_2–, F_1–, E– layer and plasmasphere
In E1c, the F_1– and the E– layer are considered in addition to that in E1a, but still ignoring the D layer key parameters and accordingly, in $\mathcal{K}_2$ we set

$$N_m^D(\varphi_m, \lambda_j) = 0 \tag{A.8}$$

$\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. It shall be noted that the $N_m^{F_1}$, $h_m^{F_1}$, N_m^E and h_m^E are directly obtained from the IRI model, whereas the corresponding H^{F_1} and H^E are both assumed to be 10% of H^{F_2}[14].

Evaluation E1d: F_2–, F_1–, E–, D– layer and plasmasphere[15]
In E1d, the F_1–, E–, D– layers are considered in the given values of parameters in $\mathcal{K}_2$ (see Eq. (6.2)) and are obtained from the IRI model. It shall be noted that the H^{F_1}, H^E and H^D are assumed to be 10% of H^{F_2}[16].

Summary of E1a - E1d
Setting the given values of N_m^D, N_m^E and $N_m^{F_1}$ in $\mathcal{K}_2$ to zero, is equivalent to ignoring the corresponding D–, E–, F_1– layers in electron density modelling, the impact of which is approximately 0.012 EDU ($\approx 0.5\%$) between E1a and E1d, consistently among the four ionosonde stations (see Fig. 11). It indicates that the deviation between the estimated and the ionosonde derived $N_m^{F_2}$ is not only due to ignoring the ionospheric bottom layers but also (possibly) due to other effects, such as:

- the thermosphere-ionosphere coupling, especially in the bottom layers (see Mendillo et al. (2002), Luehr et al. (2011))
- field aligned current (FAC), sporadic-E (Es) layer, equitorial electrojets etc. (see Rawer and Argence (1954), Rishbeth and Barron (1960), Rishbeth (1968), Parkinson et al. (1998))
- differences in the maximum vertical penetration of the signals from ionosonde stations (see e.g. Hunsucker (1991), Lei and Chuo (2014), McNamara et al. (2008)).

[14] This is only an approximation, for lack of a more realistic information, especially for the E-layer. Similar assumptions were also made by Liang et al. (2016)

[15] This is considered to be the most realistic case of modelling the ionospheric and plasmaspheric layers.

[16] This is only an approximation, for lack of a more realistic information, especially for the D and E-layers. Similar assumptions were also made by Liang et al. (2016). In future, when there would be accurate models of the bottom layers, these assumptions should be relaxed.

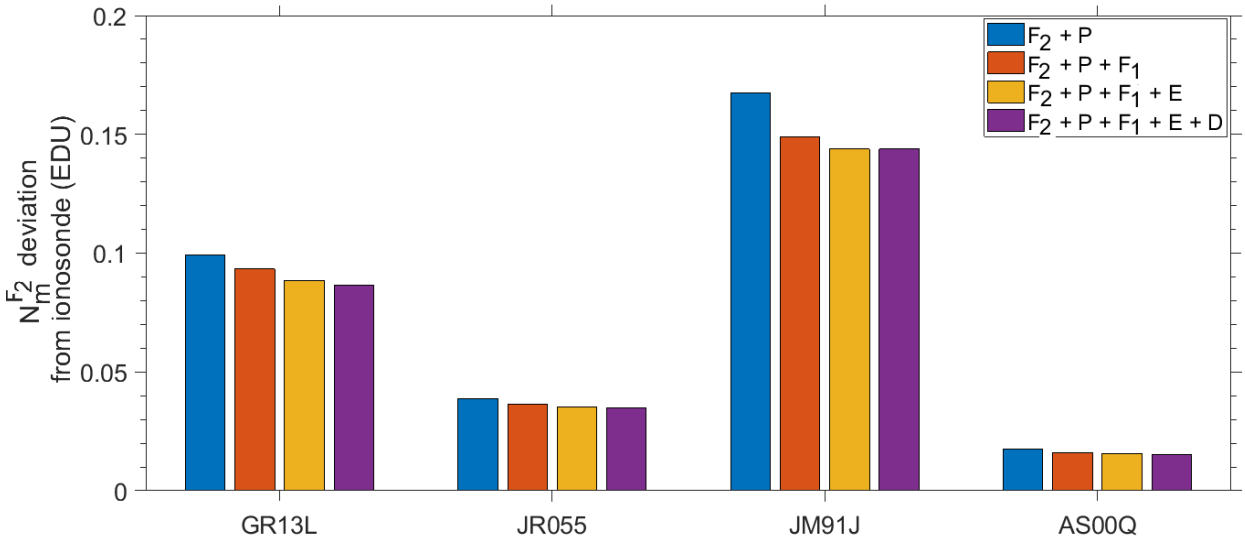


Figure 11: Effect of considering the given values of F_1, E and D layers on $N_m^{F_2}$ deviation with respect to four ionosonde stations and mean over 10-12 March 2015 at 15:00 hours.

In the further evaluations E9 - E12, the dynamic range will be progressively varied and a threshold $N_m^{F_2}$ and $h_m^{F_2}$ bounds is determined, beyond which any constraint violations occur.

C.8 Varying constraint bounds with simulated data

In the evaluations E1 to E8, the constraints and dynamic ranges of the key parameters, $\kappa \in \mathcal{K}_1$, have been defined at a $5° \times 5°$ spatial resolution globally (see Figs. 6.5, 14, 15). In contrast, a constant global upper and lower bound will be applied in the evaluations E9 to E12.

Firstly, in E9, a constant lower bound for $N_m^{F_2}$ of 0.02 EDU is applied, while progressively reducing the corresponding upper bound from 3.2 to 1.8 EDU. Subsequently in E10, a constant upper bound for $N_m^{F_2}$ of 2.8 EDU is applied, while progressively increasing the corresponding lower bound from 0.001 to 0.51 EDU. Accordingly, for each pair of upper and lower bounds, there exist sub-evaluations E9a to E9r and E10a to E10n, as shown in the Figs. 13, 10.

Similarly, in the evaluations E11 and E12, the $h_m^{F_2}$ constraint bounds will be varied. In the sub-evaluations E11a to E11d, a constant lower bound of 220 km for $h_m^{F_2}$ is applied, while progressively reducing the corresponding upper bound from 480 to 420 km. Subsequently, in E12, a constant upper bound of 480 km for $h_m^{F_2}$ is applied, while progressively increasing its lower bound from 220 to 280 km.

In all the four evaluations E9 to E12, the observation dataset DS1a is used and the number of key parameters from $\mathcal{K}_1$ are estimated. The advantage of the simulated observation DS1a is that a reference (κ_{ref}) is available, with which the deviation

$$\Delta\kappa(\varphi_m, \lambda_j) = \widehat{\kappa}(\varphi_m, \lambda_j) - \kappa_{ref}(\varphi_m, \lambda_j) \tag{A.9}$$

and the relative deviation

$$\Delta\kappa_{rel}(\varphi_m, \lambda_j) = \frac{\Delta\kappa(\varphi_m, \lambda_j)}{\kappa_{ref}(\varphi_m, \lambda_j)} \tag{A.10}$$

are computed, where $\widehat{\kappa}$ denotes the estimated key parameter from the estimated B-spline coefficients $\forall\ m = \{1, \cdots, \varphi_{max}\};\ j = \{1, \cdots, \lambda_{max}\}$. The relative deviation (A.10) is a dimensionless parameter and its global RMS is a measure of accuracy of the estimated key parameter $\kappa \in \mathcal{K}_1$, for e.g. as shown for $N_m^{F_2}$ in Tables 8 and 9.

Qualitatively, the feasible region in E9a, E10a have a larger hypervolume compared to E9b, E10b respectively and it progressively reduces from E9a to E9r and E10a to E10n (see Fig. 13). As expected, both the RMS relative deviation and constraint violations increase, as summarized in Tables 8, 9.

C.9 Evaluation E9: Varying $N_m^{F_2}$ upper bound

The evaluations E9a - E9r, corresponding to the different $N_m^{F_2}$ upper bounds, are listed in Table 8 along with the resulting RMS relative deviation and the percentage of constraint violations. As examples, the upper and lower bound constraints of the three selected evaluations (E9e, E9h and E9k) are shown below. It shall be noted that the constraint bounds on other key parameters, $\kappa \in \mathcal{K}_1 \;;\; \kappa \neq N_m^{F_2}$, are not modified in E9 and their values are the same as used in the evaluation E1.

Evaluation E9e

In E9e, the constraints

$$\begin{aligned} 0.02 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.8 \text{ EDU} \\ 220 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.11}$$

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. Compared to nominal space weather conditions, where $N_m^{F_2}$ reaches a global maximum of approximately 2.5 EDU (Bilitza et al. (2011)), a constraint bound of 2.8 EDU yields a relatively large dynamic range[17] and thus the $N_m^{F_2}$ constraints (A.11) remain inactive.

Evaluation E9h

In E9h, the constraints

$$\begin{aligned} 0.02 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.5 \text{ EDU} \\ 220 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.12}$$

are applied $\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. Compared to E9a - E9g, the dynamic range of $N_m^{F_2}$ constraints are reduced and, as explained in E9e, 2.5 EDU is approximately the global maximum $N_m^{F_2}$ under nominal ionospheric conditions[18] and thus becomes a critical threshold, at which 12% of constraints (A.12) are violated.

Evaluation E9k

In E9k, the constraints

$$\begin{aligned} 0.02 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.2 \text{ EDU} \\ 220 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.13}$$

are applied $\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. Compared to E9a - E9j, the dynamic range of $N_m^{F_2}$ constraints are reduced and as explained already in E9h, the bound of 2.2 EDU is lower than the threshold at which 15% of constraints are violated.

An unconstrained optimization solution is additionally computed (included in the column titled "Unc.") along with the selected sub-evaluations E9e, E9g and E9k in Table 8. As the upper bound is reduced from 3.2 EDU, a significant reduction[19] in the relative deviation, occurs at the critical threshold of 2.5 EDU. The lowerst RMS relative deviation corresponds to the $N_m^{F_2}$ upper bound of 2.6 EDU in E9g. From E9a to E9g, although the $N_m^{F_2}$ dynamic range is progressively reduced, the optima is still inside the feasible region, which is no longer the case for the evaluations E9h - E9r, where the RMS relative deviation as well as the number of constraint violations increase.

[17] Larger dynamic range also means that the constraints are relatively "conservative".

[18] Charecterized by nominal space weather indices such as Kp and F10.7. Typical daily values are 1 < Kp < 4 and 50 < F10.7 < 100 solar flux units (SFU).

[19] by 13 orders of magnitude

Table 7: Summary of the three specific evaluations for varying $N_m^{F_2}$ upper bound

	Unc	E9e	E9g	E9k
$N_m^{F_2}$ upper bound (EDU)	-	2.8	2.5	2.2
$N_m^{F_2}$ lower bound (EDU)	-	0.01	0.01	0.01
$N_m^{F_2}$ dynamic range (EDU)	-	2.79	2.49	2.19
$h_m^{F_2}$ upper bound (km)	-	480	480	480
$h_m^{F_2}$ lower bound (km)	-	220	220	220
$h_m^{F_2}$ dynamic range (km)	-	260	260	260
H^{F_2} upper bound (km)	-	180	180	180
H^{F_2} lower bound (km)	-	5	5	5
H^{F_2} dynamic range (km)	-	175	175	175
# iterations	15	15	15	15
# $N_m^{F_2}$ constraints	-	5402	5402	5402
# $h_m^{F_2}$ constraints	-	5402	5402	5402
# H^{F_2} constraints	-	5402	5402	5402
Observations source	DS1a	DS1a	DS1a	DS1a
# observations	240389	240389	240389	240389
# unknowns	2160	56180	56180	56180
RMS relative deviation (%)	7.2×10^{-11}	1.12×10^{-14}	6.17×10^{-17}	3.63×10^{-4}

Table 8: Global RMS and mean relative deviation as well as constraint violations for evaluations in E9.

Evaluation	upper bound (EDU)	RMS relative deviation (%)	mean relative deviation (%)	constraint violations (%)
E9a	3.2	1.17×10^{-13}	-3.76×10^{-13}	0
E9b	3.1	6.57×10^{-13}	-2.96×10^{-13}	0
E9c	3.0	2.61×10^{-13}	-1.24×10^{-13}	0
E9d	2.9	6.60×10^{-13}	7.43×10^{-14}	0
E9e	2.8	1.12×10^{-14}	-3.96×10^{-14}	0
E9f	2.7	1.35×10^{-14}	-5.24×10^{-15}	0
E9g	2.6	6.17×10^{-17}	-1.11×10^{-17}	0
E9h	2.5	2.25×10^{-4}	-2.50×10^{-6}	12
E9i	2.4	1.52×10^{-3}	-3.63×10^{-5}	12
E9j	2.3	3.73×10^{-3}	-1.13×10^{-4}	15
E9k	2.2	3.63×10^{-4}	-2.05×10^{-4}	15
E9l	2.1	9.08×10^{-3}	-2.97×10^{-4}	16
E9m	2.0	1.18×10^{-2}	-3.90×10^{-4}	18
E9n	1.98	1.23×10^{-2}	4.18×10^{-4}	21
E9o	1.94	1.35×10^{-2}	-4.85×10^{-4}	21
E9p	1.90	1.47×10^{-2}	-5.71×10^{-4}	21
E9q	1.86	1.60×10^{-2}	-6.75×10^{-4}	22
E9r	1.82	1.75×10^{-2}	-8.00×10^{-4}	22

C.10 Evaluation E10: Varying $N_m^{F_2}$ lower bound

Table 9 lists the evaluations E10a - E10n, corresponding to the different $N_m^{F_2}$ lower bounds, along with the resulting RMS relative deviation and the percentage of $N_m^{F_2}$ constraint violations. As examples, the upper and lower bound constraints of three selected sub-evaluations (E10b, E10d and E10j) are shown below. Once again, as with E9, the constraint bounds on the key parameters, $\kappa \in \mathcal{K}_1 \;;\; \kappa \neq N_m^{F_2}$, are not modified and their values are the same as those used in the evaluation E1.

Evaluation E10l

In E10l, the constraints

$$\begin{aligned} 0.11 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.8 \text{ EDU} \\ 220 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.14}$$

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. As the $N_m^{F_2}$ global minimum value[20] is approximately 0.05 EDU, under nominal space weather conditions (Bilitza et al. (2011)), for a lower bound of 0.11 EDU, the dynamic range in E10l is lower compared to that in E10a - E10k. Accordingly, 7% of the $N_m^{F_2}$ constraints in (A.14) are violated.

Evaluation E10m

In E10m, the constraints

$$\begin{aligned} 0.31 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.8 \text{ EDU} \\ 220 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.15}$$

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. Compared to E10a - E10l, the $N_m^{F_2}$ dynamic range is reduced in E10m (for a lower bound of 0.31 EDU), at which 21% of the corresponding $N_m^{F_2}$ constraints in (A.15) are violated.

Evaluation E10n

In E10n, constraints

$$\begin{aligned} 0.51 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.8 \text{ EDU} \\ 220 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.16}$$

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. Compared to E10a - E10m, $N_m^{F_2}$ dynamic range is reduced in E10n (due to the "large[21]" lower bound of 0.51 EDU), at which 39% of $N_m^{F_2}$ constraints in (A.16) are violated (see Fig. 6.8).

As the $N_m^{F_2}$ lower bound is increased from 0.001 EDU, a significant change[22] in RMS relative deviation, is only seen for a lower bound of 0.08 EDU or larger. The lowest RMS relative deviation results in E10g, at a lower bound of 0.06 EDU.

C.11 Evaluation E11: Varying $h_m^{F_2}$ lower bound

Evaluation E11b

In E11b, the constraints

$$\begin{aligned} 0.02 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.8 \text{ EDU} \\ 240 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.17}$$

[20] Minimum value occurs at night time and varies during a solar cycle and also exhibits an annual climatology in IRI model.

[21] Compared to a nominal lower bound of 0.02 EDU in E10a, 0.51 EDU is considered as extremely large.

[22] by 10 orders of magnitude

Table 9: Summary of the three specific evaluations for varying $N_m^{F_2}$ lower bound

	Unc.	E10l	E10m	E10n
$N_m^{F_2}$ upper bound (EDU)	-	3	3	3
$N_m^{F_2}$ lower bound (EDU)	-	0.11	0.31	0.51
$N_m^{F_2}$ dynamic range (EDU)	-	2.89	2.69	2.49
$h_m^{F_2}$ upper bound (km)	-	480	480	480
$h_m^{F_2}$ lower bound (km)	-	220	220	220
$h_m^{F_2}$ dynamic range (km)	-	260	260	260
H^{F_2} upper bound (km)	-	180	180	180
H^{F_2} lower bound (km)	-	5	5	5
H^{F_2} dynamic range (km)	-	175	175	175
# iterations	15	15	15	15
# $N_m^{F_2}$ constraints	-	5402	5402	5402
# $h_m^{F_2}$ constraints	-	5402	5402	5402
# H^{F_2} constraints	-	5402	5402	5402
Observation dataset	DS1a	DS1a	DS1a	DS1a
# observations	240389	240389	240389	240389
# unknowns	2160	56180	56180	56180
$N_m^{F_2}$ RMS relative deviation (%)	4.12×10^{-11}	1.83×10^{-2}	7.68	44.3

Table 10: RMS and mean relative deviation as well as constraint violations for various lower bound of $N_m^{F_2}$.

Evaluation	lower bound (EDU)	RMS relative deviation (%)	mean relative deviation (%)	constraint violations (%)
E10a	0.001	2.61×10^{-12}	-1.24×10^{-13}	0
E10b	0.01	1.07×10^{-12}	-5.26×10^{-13}	0
E10c	0.02	2.98×10^{-12}	-1.36×10^{-12}	0
E10d	0.03	1.22×10^{-13}	-5.11×10^{-14}	0
E10e	0.04	5.42×10^{-13}	-2.17×10^{-14}	0
E10f	0.05	6.87×10^{-16}	-2.40×10^{-16}	0
E10g	0.06	3.43×10^{-16}	-1.46×10^{-16}	0
E10h	0.07	1.03×10^{-15}	-1.24×10^{-16}	0
E10i	0.08	3.39×10^{-4}	8.25×10^{-6}	3
E10j	0.09	3.08×10^{-3}	1.50×10^{-4}	3
E10k	0.1	9.56×10^{-3}	6.99×10^{-4}	7
E10l	0.11	1.83×10^{-2}	1.89×10^{-3}	7
E10m	0.31	7.68	8.14	21
E10n	0.51	44.3	46.2	39

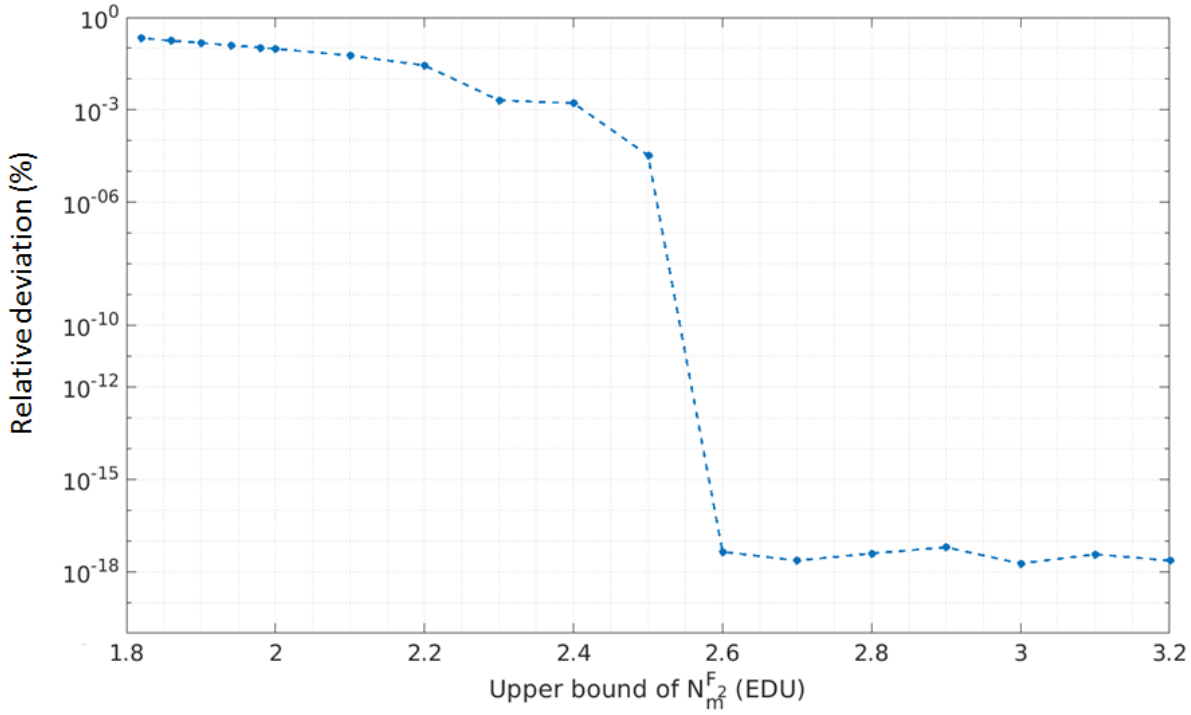


Figure 12: RMS relative deviation for different upper bounds of $N_m^{F_2}$

Table 11: RMS and mean relative deviation as well as constraint violations for various lower bound of $h_m^{F_2}$.

Evaluation	lower bound (km)	RMS relative deviation (%)	mean relative deviation (%)	constraint violations (%)
E11a	220	6.11×10^{-16}	-3.47×10^{-15}	0
E11b	240	7.44×10^{-7}	-1.60×10^{-6}	1
E11c	260	3.31×10^{-6}	-5.66×10^{-6}	4
E11d	280	2.23×10^{-5}	-1.01×10^{-3}	10

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}$; $j = \{1, \cdots, \lambda_{max}\}$. As the $h_m^{F_2}$ reaches a minimum of approximately 240 km[23] under nominal space weather conditions (Bilitza et al. (2011)), a bound of 240 km yields a lower dynamic range (relative to E11a) at which 1% of $h_m^{F_2}$ constraints in (A.17) are violated.

Evaluation E11c

In E11c, the constraints

$$\begin{aligned} 0.02 \leq N_m^{F_2}(\varphi_m, \lambda_j) \leq 2.8 \text{ EDU} \\ 260 \leq h_m^{F_2}(\varphi_m, \lambda_j) \leq 480 \text{ km} \end{aligned} \tag{A.18}$$

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}$; $j = \{1, \cdots, \lambda_{max}\}$. Following the analogy described in E11b, a bound of 260 km yields a lower dynamic range (relative to E11a, E11b), at which 4% of $h_m^{F_2}$ constraints in (A.18) are violated (see Fig. 6.9). The relative deviation statistics as well as the percentage of total number of $h_m^{F_2}$ constraints that were violated as a result of the varying the $h_m^{F_2}$ lower bounds in evaluations E11a - E11d are summarized in Table 11. Similarly, the relative deviation statistics as well as the percentage of total number of $h_m^{F_2}$ constraints that were violated as a result of the varying the $h_m^{F_2}$ upper bounds in evaluations E11a - E11d are summarized in Table 12.

[23] A minimum value of approximately 240 km is reached in night-time in IRI model. In addition to the time of day, the minimum $h_m^{F_2}$ also depands on the input solar flux (F10.7) and follows an annual climatology.

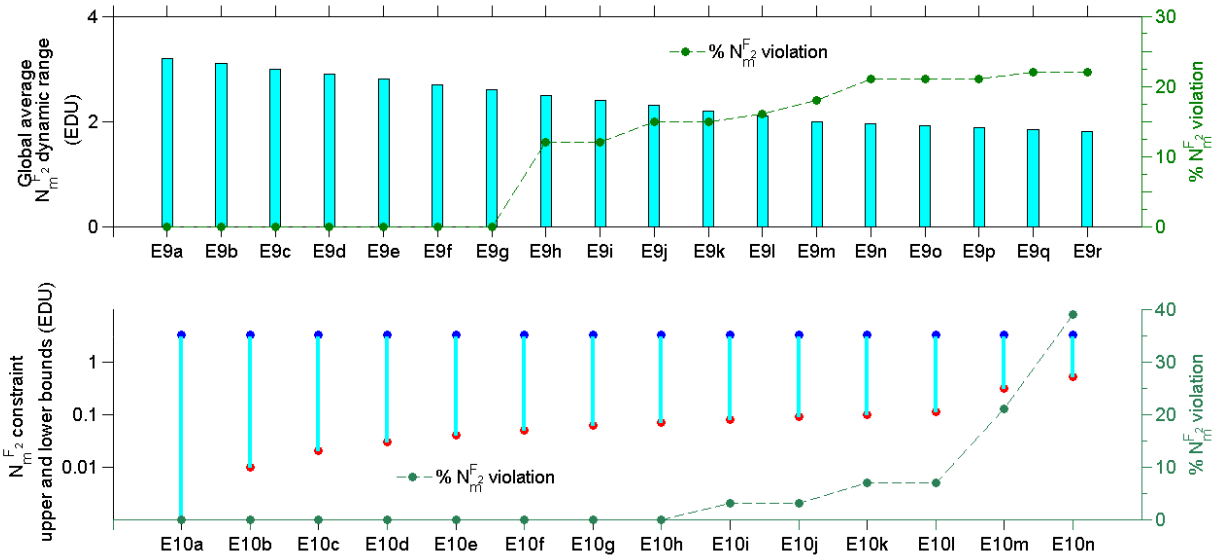


Figure 13: Top panel: Left vertical axis shows the global average $N_m^{F_2}$ dynamic range used as input for the constraints in sub-evaluations E9a - E9r. Bottom panel: $N_m^{F_2}$ upper and lower constraint bounds used in sub-evaluations E10a - E10l. The left vertical axis is in logarithmic scale to especially show the variation in lower bounds. In both the top and bottom panels the resulting % of constraint violation are shown on the right side vertical axis.

Table 12: RMS and mean relative deviation as well as constraint violations for various upper bound of $h_m^{F_2}$.

Evaluation	upper bound (km)	RMS relative deviation (%)	mean relative deviation (%)	constraint violations (%)
E12a	480	3.87×10^{-14}	-2.11×10^{-15}	0
E12b	460	4.61×10^{-16}	-2.10×10^{-17}	0
E12c	440	6.10×10^{-6}	-1.16×10^{-5}	3
E12d	420	5.43×10^{-5}	-5.88×10^{-5}	8

C.12 Evaluation E12: Varying $h_m^{F_2}$ upper bound

Evaluation E12a

In E12a, the constraints

$$\begin{aligned} 0.02 \le N_m^{F_2}(\varphi_m, \lambda_j) \le 2.8 \text{ EDU} \\ 220 \le h_m^{F_2}(\varphi_m, \lambda_j) \le 480 \text{ km} \end{aligned} \tag{A.19}$$

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}$; $j = \{1, \cdots, \lambda_{max}\}$. As $h_m^{F_2}$ reaches a maximum of approximately 450 km[24] under nominal space weather conditions (Bilitza et al. (2011)), a bound of 480 km yields a relatively larger dynamic range and thus the $h_m^{F_2}$ constraints in (A.19) remain inactive.

Evaluation E12c

In E12c, constraints

$$\begin{aligned} 0.02 \le N_m^{F_2}(\varphi_m, \lambda_j) \le 2.8 \text{ EDU} \\ 220 \le h_m^{F_2}(\varphi_m, \lambda_j) \le 440 \text{ km} \end{aligned} \tag{A.20}$$

[24] A maximum value of approximately 450 km is reached in day-time around 12:00 - 14:00 local time betweem 30° N and 30° S latitudes in IRI model. Specifically, maximum $h_m^{F_2}$ depands on the input solar flux as well as follows an annual climatology.

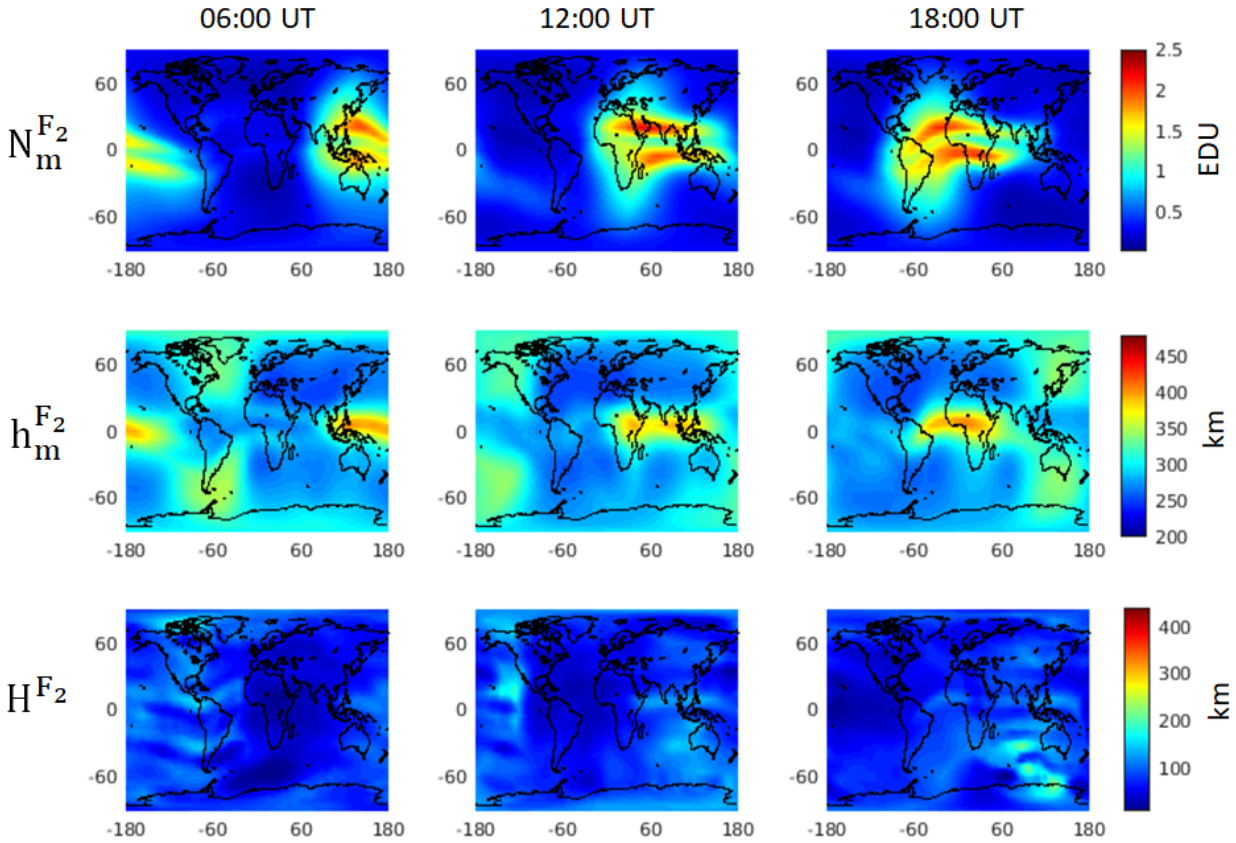


Figure 14: Global lower bound constraints on $N_m^{F_2}$, $h_m^{F_2}$ and H^{F_2} at three different times of day (06:00, 12:00 and 18:00 UT).

are applied, $\forall\ m = \{1, \cdots, \varphi_{max}\}\ ;\ j = \{1, \cdots, \lambda_{max}\}$. Following the analogy of the maximum $h_m^{F_2}$ described in E12a, a bound of 440 km is lower than the critical threshold as well as yields a relatively lower dynamic range, at which 3% of $h_m^{F_2}$ constraints in (A.19) are violated.

Table 13: Dataset with different number of observations in each techniques.

Dataset	VTEC sepr.	GNSS	Formosat-3/COSMIC	GRACE
DS2a	240389	7500	3153	822
DS2b	60000	7500	3153	822
DS2c	120000	7500	3153	822
DS2d	240389	150000	3153	822
DS2e	60000	150000	3153	822
DS2f	120000	150000	3153	822

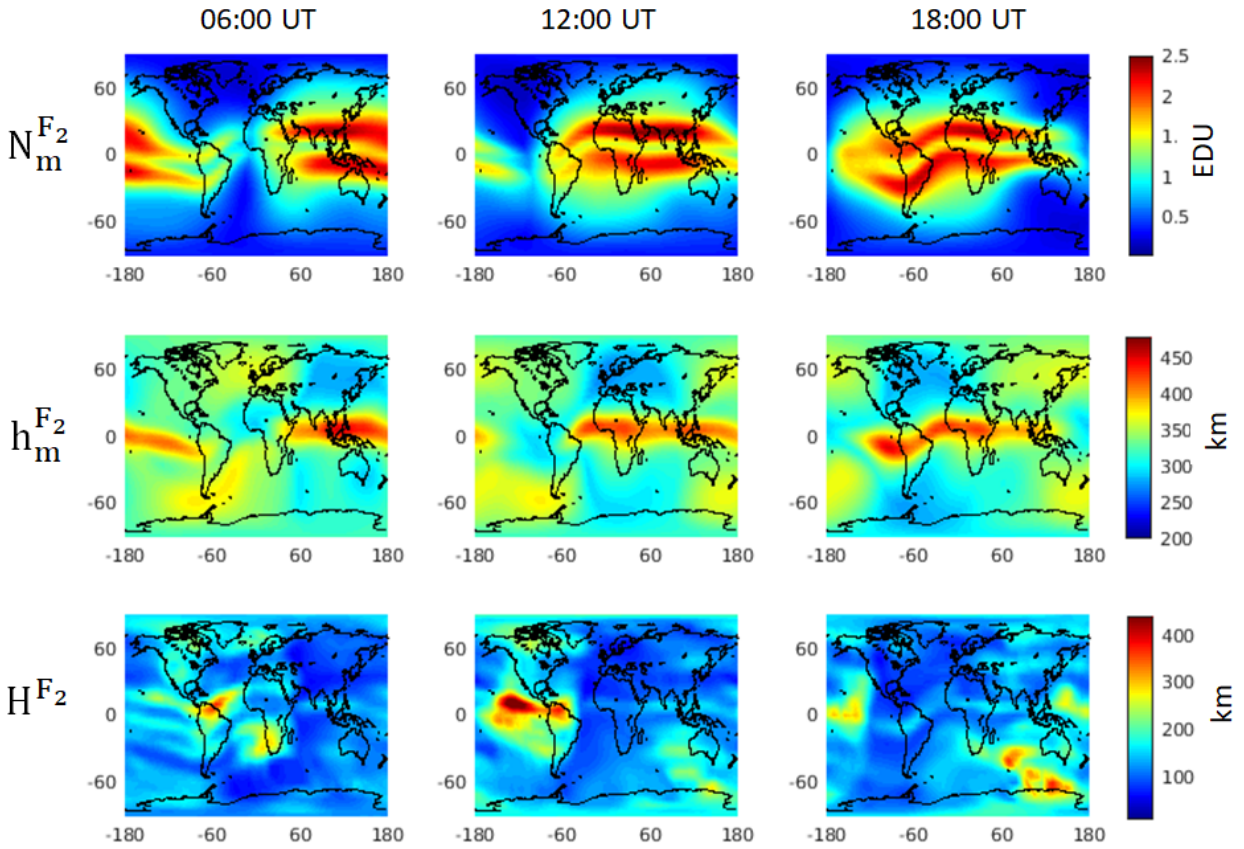


Figure 15: Global upper bound constraints on $N_m^{F_2}$, $h_m^{F_2}$ and H^{F_2} at three different times of day (06:00, 12:00 and 18:00 UT).

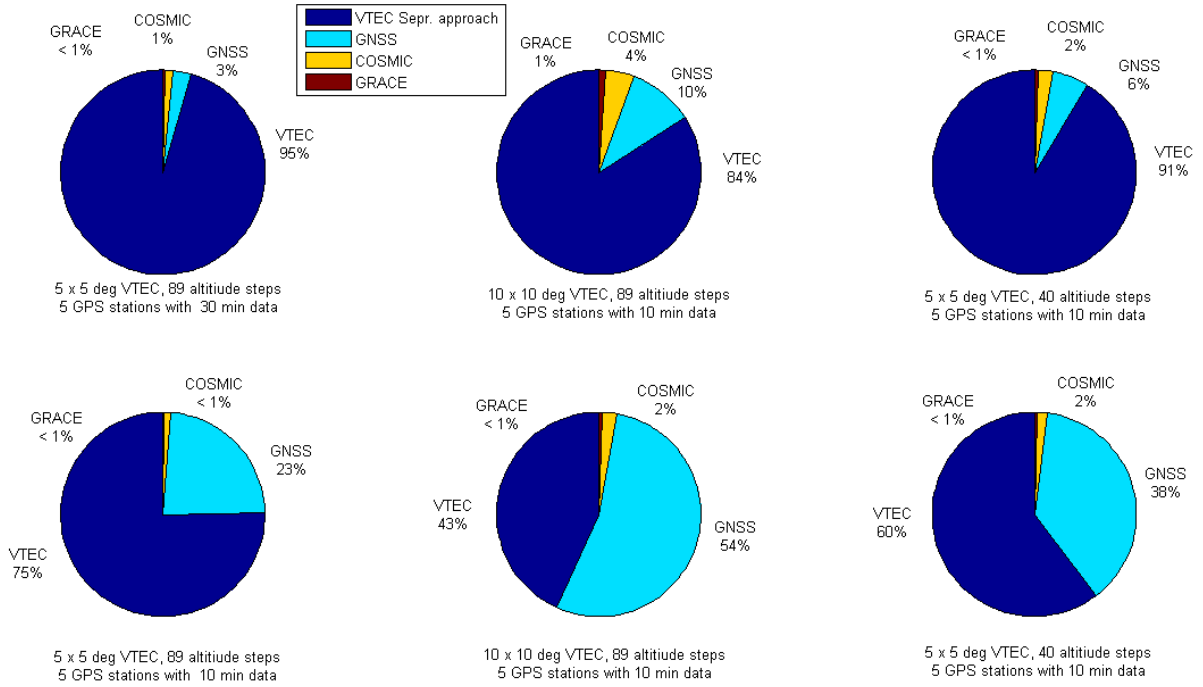


Figure 16: Varying the relative number of observations from different techniques.

D GNSS data processing

GNSS error sources

The main goal of this section is to describe GNSS data processing procedure, particularly relevant for elecctron density modelling. Inherently, it involves the study of the various GNSS error sources[25], especially those which are relevant and have an impact on both the STEC as well as radio occultation observations quality. The remaining error sources are mentioned only for the sake of completeness. The characteristics of the GNSS error sources allow the selection of appropriate variance factors to be used in the stochastic modelling (see Sections 4.1.1 and 4.1.2).

Signal-in-space error

The total contribution of error due to the satellite orbit and the clock are together referred to as the signal-in-space error (SISE), which is an approximate figure of merit (FoM) of the space segment for any given GNSS constellation. A constellation-wide SISE of 80 cm and 50 cm, for GPS and Galileo respectively, is reported by Montenbruck et al. (2014b). The superior passive hydrogen maser (PHM) clock in Galileo, compared to the Rubidium clocks of GPS, is an important enabler for precise STEC measurements (Steigenberger et al. (2015), Prange et al. (2017)).

Propagation medium errors

The ionosphere and the troposphere are the two main propagation media related effects in GNSS observations. While the modulating signal[26] is delayed in time, thereby resulting in a slower group velocity[27] compared to that in free-space, the corresponding phase velocity[28] is advanced. These are denoted by the positive sign on the first order ionosphere delay term in pseudorange and a corresponding negative one in the phase observation equations ; see Eqs. A.21a)-(A.21f). As the central topic of this chapter, STEC will be derived from GNSS observations to be further used for electron density modelling. This is described in detail in Appendix D.2.

Other error sources

There is a multipath effect on the pseudorange observations (A.21a), (A.21c) and (A.21e) due to specular and diffuse reflections[29] of the received GNSS signals due to obstructions in the proximity of the receiver antenna (Braasch (1997)). The multipath on carrier phase is comparatively negligible (see Misra and Enge (2006), Hoffmann and Thiele (2007)) and is not considered in the scope of this work. There are two additional effects on phase observations, which are the antenna phase center offset (Schmid et al. (2005), Steigenberger et al. (2016)) and the phase wind up effect (Kim et al. (2006)). The former is caused by the difference between a satellite's geometrical center of mass and a frequency specific signal transmission point (Montenbruck et al. (2009)). The latter is caused by a relative rotation of the satellite antenna to keep pointing towards the Earth and the solar panels towards the Sun. This causes a change from a nominal right hand circular polarization (RHCP) of the transmitted signals at the receiving antenna[30]. Further details about these error sources can be found in, e.g. Hoffmann and Thiele (2007).

[25] Also called GNSS error budget.

[26] Pseudorandom code is modulated on GNSS carrier signal.

[27] Velocity with which the overall envelope of a modulated signal's amplitude propagates (see e.g. Brillouin (2013)).

[28] Velocity at which the phase of any single frequency component of the wave travels (see e.g. Brillouin (2013)).

[29] In this context, specular reflection is when a GNSS signal is reflected from a smooth surface at a definite angle, and diffuse reflection, that results from a rough surfaces that tends to scatter the reflected signal in all directions; see e.g. Kerker (1969)

[30] When an electromagnetic wave is reflected, its plane of polarization undergoes a rotation. If the antenna and signal are not aligned in polairzation, the received signal would have a relatively smaller signal-to-noise ratio. This aspect is an important consideration for the quality of STEC and IRO observations.

It is important to highlight the role of pseudorange multipath towards both STEC as well as the IRO observation quality[31]. IGS receivers are equipped with high-gain antenna as well as robust receiver signal tracking algorithms targeted towards multipath mitigation (see e.g. Braasch (1997), McGraw and Braasch (1999), Ray (2000)). On the other hand, spaceborne occultation receivers onboard LEO satellites have only minimal specular reflections from any obstruction to the direct line of sight to GNSS satellites but have relatively less robust signal tracking capabilities compared to ground stations (Silvestrin et al. (2000)). Accordingly, the quality of derived STEC and IRO measurements depend on the nature of receiver. Although the multipath term in pseudorange (A.21a), (A.21c) is not negligible[32] but, without specific details on receiver signal processing, such as the use of multipath limiting antenna (MLA) or special receiver tracking techniques (Braasch (1997), McGraw and Braasch (1999)), they are assumed to be absorbed by the pseudorange measurement noise. In this work, the multipath effects on STEC are partially mitigated by applying a 10° elevation angle cut-off.

D.1 GNSS observation equation

The goal of this section is to introduce the GNSS observation equations and the STEC derivation for EDM. Due to the differences in signal types[33], time reference[34], carrier frequencies, satellite orbit and biases, the observation equations (A.21a) - (A.21f) are specific for different GNSS constellations as well as carrier frequencies. For the sake of consistency, the observations are converted metric units[35]. Pseudorange (P) and phase (L) observations, on three frequencies, L1, L2 and L5, denoted by the subscript 1, 2 and 5, read[36] (see e.g. Hofmann-Wellenhof et al. (2007), Hauschild (2017)),

$$P^s_{i,1} + e^s_{P_i,1} = \rho^s_i + c\,(\delta t^s - \delta t_i) + T^s_i + I^s_{i,1} + M^s_{i,1} + b^P_{i,1} + b^{P,s}_1 \tag{A.21a}$$

$$L^s_{i,1} + e^s_{L_i,1} = \rho^s_i + c\,(\delta t^s - \delta t_i) + T^s_i - I^s_{i,1} + \lambda_1 N^s_{i,1} + b^L_{i,1} + b^{L,s}_1 \tag{A.21b}$$

$$P^s_{i,2} + e^s_{P_i,2} = \rho^s_i + c\,(\delta t^s - \delta t_i) + T^s_i + I^s_{i,2} + M^s_{i,2} + b^P_{i,2} + b^{P,s}_2 \tag{A.21c}$$

$$L^s_{i,2} + e^s_{L_i,2} = \rho^s_i + c\,(\delta t^s - \delta t_i) + T^s_i - I^s_{i,2} + \lambda_2 N^s_{i,2} + b^L_{i,2} + b^{L,s}_2 \tag{A.21d}$$

$$P^s_{i,5} + e^s_{P_i,5} = \rho^s_i + c\,(\delta t^s - \delta t_i) + T^s_i + I^s_{i,5} + M^s_{i,5} + b^P_{i,5} + b^s_5 \tag{A.21e}$$

$$L^s_{i,5} + e^s_{L_i,5} = \rho^s_i + c\,(\delta t^s - \delta t_i) + T^s_i - I^s_{i,5} + \lambda_5 N^s_{i,5} + b^L_{i,5} + b^{L,s}_5, \tag{A.21f}$$

where

- the subscript i and superscript s denote the receiver and satellite indices, the frequency dependent terms are denoted using the subscript $\{1, 2, 5\}$ corresponding to GPS $\{L1, L2, L5\}$ frequencies respectively.
- c is the speed of light in vaccum
- $\lambda_j \quad \forall \; j \; \in \{1, 2, 5\}$ is the wavelength corresponding to the frequency $f_j = \frac{c}{\lambda_j}$. Specifically, $\lambda_1 \approx 19$ cm, $\lambda_2 \approx 24$ cm and $\lambda_5 \approx 80$ cm
- $P^s_{i,j}$ is the pseudorange observation of satellite s to receiver i on frequency j (m)

[31] Even though phase observations are used for STEC modelling, the pseudorange multipath has an impact on the success and accuracy of ambiguity fixing.
[32] It could be larger tha 2 m in the range domain, especially for elevation angles below 40 deg.
[33] e.g. P, C or M codes: "P" denotes precision codes used on GPS L2, "C" denotes the legacy coarse-acquision codes used in almost all GNSS systems. "M" is reserved for mititary code and were not used in this work.
[34] Each GNSS has its own "time reference" and it is necessary to refer the GPS and Galileo observables to their respective system times or equivalently consider an inter-system offset; see e.g. Jekeli and Montenbruck (2017).
[35] Units of distance
[36] For description in this section, GPS triple frequency is chosen but other systems and frequencies can be chosen in a similar manner.

- $e^s_{i,j}$ is the measurement error on the pseudorange measurement of satellite s to receiver i on frequency j (m)
- ρ^s_i is the true range or geometrical distance between satellite s to receiver i (m)
- δt^s is the clock offset satellite s with respect to a common system time, e.g. the GPS time (GPST)[37] (m)
- δt_i is the clock offset of receiver i with respect to the system time (m)
- $I^s_{i,j}$ is the ionosphere delay along the ray-path from satellite s to receiver i on frequency j (m)
- $M^s_{i,j}$ is the multipath effect of receiver i on satellite s and frequency j (m)
- T^s_i is the troposphere delay along the ray-path from satellite s to receiver i (m)
- $N^s_{i,j}$ is the phase ambiguity of receiver i to satellite s on frequency j
- $b^P_{i,j}$ is the code bias of receiver i on frequency j. Superscript "P" denotes the bias is on pseudorange (or code) (m)
- $b^{P,s}_j$ is the code bias of satellite s on frequency j (m)
- $b^L_{i,j}$ is the phase bias of receiver i on frequency j. Superscript "L" denotes the bias is on phase (m)
- $b^{L,s}_j$ is the phase bias of satellite s on frequency j (m).

Generally, each error source in the observation equations is unknown[38] as well as time-varying. In this work, the time index in the observation equation is ignored because the STEC derivation will be described for a given time instant, unless specified otherwise. The ionospheric delay on L2 and L5 frequencies, compared to that in L1, is scaled

$$I^s_{i,2} = \gamma_{L1,L2}\ I^s_{i,1} \quad \text{and} \quad I^s_{i,5} = \gamma_{L1,L5}\ I^s_{i,1}, \tag{A.22}$$

by the frequency squared ratios

$$\gamma_{L1,L2} = \frac{f_1^2}{f_2^2} \quad \text{and} \quad \gamma_{L1,L5} = \frac{f_1^2}{f_5^2} \tag{A.23}$$

respectively.

D.2 STEC derivation for EDM

The goal of this section is to describe the STEC derivation procedure from the GNSS observation equations (A.21a) to (A.21f). This section is divided into two parts: Firstly, three possibilities for combining GNSS observations will be described, including the dual frequency geometry-free linear combination (GFLC) along with their relative advantages and disadvantages. In the second part,

[37] Each system has a definition of its own system time.

[38] More specifically, the signal time delay from each error source is unknown. But these are converted to units of distance, as mentioned above. Therefore, the GNSS error sources modelled for STEC computation are also referred to in units of distances (in meters).

a procedure for carrier phase ambiguity resolution will be described and its impact on STEC as well as IRO observation quality will be highlighted. In STEC computation, the GNSS observations (A.21a)-(A.21f) are combined such that the ionospheric delay term is retained[39].

Geometry free linear combination

In this work, a dual frequency geometry-free linear combination (GFLC)

$$L^s_{i,1} - L^s_{i,2} + e^s_{i,L_{12}} = -I^s_{i,1}(1-\gamma_{L1,L2}) + \lambda_1 N^s_{i,1} - \lambda_2 N^s_{i,2} + b^L_{i,12} + b^{L,s}_{12} \tag{A.24a}$$

$$L^s_{i,1} - L^s_{i,5} + e^s_{i,L_{15}} = -I^s_{i,1}(1-\gamma_{L1,L5}) + \lambda_1 N^s_{i,1} - \lambda_5 N^s_{i,5} + b^L_{i,15} + b^{L,s}_{15} \tag{A.24b}$$

$$P^s_{i,1} - P^s_{i,2} + e^s_{i,P_{12}} = I^s_{i,1}(1-\gamma_{L1,L2}) + \Delta M^s_{i,12} + b^P_{i,12} + b^{P,s}_{12} \tag{A.24c}$$

$$P^s_{i,1} - P^s_{i,5} + e^s_{i,P_{15}} = I^s_{i,1}(1-\gamma_{L1,L5}) + \Delta M^s_{i,15} + b^P_{i,15} + b^{P,s}_{15} \tag{A.24d}$$

has been used, following Limberger (2015), Erdogan et al. (2017) and Erdogan et al. (2020). Although from Eqs. (A.21a) to (A.21f), three separate GFLC are possible, the L1-L2 and L1-L5 are the commonly used ones[40] (see Henkel (2010), Montenbruck et al. (2014a)). This is due to the simultaneous tracking capabilities and multi-band antenna availability for GNSS receivers. GFLC removes the non-dispersive terms completely[41], where

$$\begin{aligned}
L^s_{i,12} &= L^s_{i,1} - L^s_{i,2} \\
P^s_{i,12} &= P^s_{i,1} - P^s_{i,2} \\
L^s_{i,15} &= L^s_{i,1} - L^s_{i,5} \\
P^s_{i,15} &= P^s_{i,1} - P^s_{i,5} \\
e^s_{i,P_{12}} &= e^s_{i,P_1} - e^s_{i,P_2} \\
e^s_{i,P_{15}} &= e^s_{i,P_1} - e^s_{i,P_5} \\
e^s_{i,L_{12}} &= e^s_{i,L_1} - e^s_{i,L_2} \\
e^s_{i,L_{15}} &= e^s_{i,L_1} - e^s_{i,L_5}
\end{aligned} \tag{A.25}$$

will be used for simplification. The L1, L2 receiver and satellite differential code biases are denoted as

$$\begin{aligned}
b^P_{i,12} &= b^P_{i,1} - b^P_{i,2} \\
b^{P,s}_{12} &= b^{P,s}_{1} - b^{P,s}_{2}
\end{aligned} \tag{A.26}$$

and the corresponding differential phase biases as

$$\begin{aligned}
b^L_{i,12} &= b^L_{i,1} - b^L_{i,2} \\
b^{L,s}_{12} &= b^{L,s}_{1} - b^{L,s}_{2}.
\end{aligned} \tag{A.27}$$

Analogously, the L1-L5 differential code or phase biases The L1, L2 receiver and satellite differential code biases are denoted as

$$\begin{aligned}
b^P_{i,15} &= b^P_{i,1} - b^P_{i,5} \\
b^{P,s}_{15} &= b^{P,s}_{1} - b^{P,s}_{5}
\end{aligned} \tag{A.28}$$

[39] Implicitly, it means that the other error sources are eliminated to the extent possible. However, as shown in this section, not all error sources would be entirely eliminated.

[40] As a convention to indicate the signal frequency L1 L2 and L5 is used here but the same applies to Galileo where it is customary to denote the frequencies by E1, E5a and E5b (see e.g. Langley et al. (2017)), Meurer and Antreich (2017)).

[41] Therefore, the system time-scale offsets (Jekeli and Montenbruck (2017)) causing an inter-system bias (ISB) (see Odijk and Teunissen (2013a)) are not considered in this work[42].

and the corresponding differential phase biases as

$$\begin{aligned} b^{L}_{i,15} &= b^{L}_{i,1} - b^{L}_{i,5} \\ b^{L,s}_{15} &= b^{L,s}_{1} - b^{L,s}_{5}. \end{aligned} \tag{A.29}$$

are also defined in a similar manner. The terms $\Delta M^{s}_{i,12}$ and $\Delta M^{s}_{i,15}$ are the differential multipath delays on L1-L2 and L1-L5 respectively.

If the differential multipath $\Delta M^{s}_{i,12}$ is absorbed by the measurement error $e^{s}_{i,P_{12}}$ and the phase ambiguities $N^{s}_{i,1}$, $N^{s}_{i,2}$ and $N^{s}_{i,5}$ are fixed to the correct integers, thereby absorbing the six phase biases $b^{L}_{i,1}$, $b^{L}_{i,2}$, $b^{L}_{i,5}$, $b^{L,s}_{1}$, $b^{L,s}_{2}$ and $b^{L,s}_{5}$, then the only remaining unknown parameters on the right-side of Eqs. (A.24a) to (A.24d) are the ionosphere delay $I^{s}_{i,1}$ and the four differential code biases (DCB) $b^{P,s}_{12}$, $b^{P,s}_{15}$, $b^{P}_{i,12}$, $b^{P}_{i,15}$. These five parameters are estimated together from GFLC observations $P^{s}_{i,12}$, $L^{s}_{i,12}$, $P^{s}_{i,15}$ and $L^{s}_{i,15}$.

There exists the integral relation (A.32) between STEC and electron density in (A.32). Following Kleusberg and Teunissen (1996), a relation between STEC and ionospheric delay on L1 or L5

$$\begin{aligned} I^{s}_{i,1} &= \alpha_1 \; STEC(\mathbf{x}_i, \mathbf{x}^s) \\ I^{s}_{i,5} &= \alpha_5 \; STEC(\mathbf{x}_i, \mathbf{x}^s), \end{aligned} \tag{A.30}$$

along the $\mathbf{x}^s$ $\mathbf{x}_i$ line-of-sight is established, where α_1 and α_5 are the frequency dependent factors for L1-L2 and L1-L5 combination respectively, computed as

$$\begin{aligned} \alpha_1 &= \frac{40.3 \; (1 - \gamma_{L1,L2})}{f_1^2} \cdot 10^{16} \quad \text{m/TECU} \\ \alpha_5 &= \frac{40.3 \; (1 - \gamma_{L1,L5})}{f_1^2} \cdot 10^{16} \quad \text{m/TECU}, \end{aligned} \tag{A.31}$$

(Kleusberg and Teunissen (1996)). Furthermore, given the GFLC observations, a direct relation between STEC and satellite as well as receiver biases has also been shown in (A.24a) to (A.24d). Therefore, combining Eqs. (2.2), (A.24a)-(A.24d) and (A.30), an empirical relation between the biases, STEC and electron density is established as shown in (A.32).

At a given time instant of STEC computation, corresponding to receiver and satellite indices $i \in \{1, \; ..., \; rmax\}$ and $s \in \{1, \; ..., \; smax\}$ respectively, the STEC observation equation reads,

$$STEC(\mathbf{x}_i, \mathbf{x}^s) + e_{STEC}(\mathbf{x}_i, \mathbf{x}^s) = \int_{\mathbf{x}^s}^{\mathbf{x}_i} N_e(\varphi, \lambda, h) \; ds, \tag{A.32}$$

where $STEC(\mathbf{x}_i, \mathbf{x}^s)$ is the observed slant total electron content from a satellite s to a receiver i, with position vectors $\mathbf{x}^s$, $\mathbf{x}_i$ and observation error $e_{STEC}(\mathbf{x}_i, \mathbf{x}^s)$ respectively. Substituting (A.30) in (A.24a) - (A.24d) along with Eqs. (A.25) to (A.31), the observation equation

$$L^{s}_{i,12} = \alpha_1 \; STEC(\mathbf{x}_i, \mathbf{x}^s) + CPB^{s}_{i,12} + e^{s}_{i,L_{12}} \tag{A.33a}$$

$$L^{s}_{i,15} = \alpha_1 \; STEC(\mathbf{x}_i, \mathbf{x}^s) + CPB^{s}_{i,15} + e^{s}_{i,L_{15}} \tag{A.33b}$$

are obtained. The carrier phase bias

$$\begin{aligned} CPB_{i,12} &= b^{L}_{i,12} + b^{L,s}_{12} + \lambda_1 N^{s}_{i,1} - \lambda_2 N^{s}_{i,2} \\ CPB_{i,15} &= b^{L}_{i,15} + b^{L,s}_{15} + \lambda_1 N^{s}_{i,1} - \lambda_5 N^{s}_{i,5} \end{aligned} \tag{A.34}$$

is the combined term containing the receiver differential code bias, the satellite differential code bias and the difference of ambiguities on the two frequency pairs,i.e. L1, L2 or L1, L5. Similarly,

the linear combination[43] of the pseudorange observations give

$$P^s_{i,12} = \alpha_1 \; STEC(\mathbf{x}_i, \mathbf{x}^s) + DCB^s_{i,12} + e^s_{i,P_{12}} \tag{A.35a}$$
$$P^s_{i,15} = \alpha_2 \; STEC(\mathbf{x}_i, \mathbf{x}^s) + DCB^s_{i,15} + e^s_{i,P_{15}}. \tag{A.35b}$$

Since the ambiguity term is assumed to be constant, the CPB can be determined by averaging the $k = 1 \cdots, L$ differences of $P^s_{i,12}$ and $L^s_{i,12}$. The terms $DCB^s_{i,12} = b^{P,s}_{12}$ and $DCB^s_{i,12} = b^{P,s}_{12}$ are introduced for the sake of consistency to the previous works by Limberger (2015), Liang (2017), Erdogan et al. (2017) and Goss et al. (2020). This procedure is called phase levelling:

$$\frac{1}{L}\sum_{k=1}^{L}\left(L^s_{i,12} - P^s_{i,12}\right)_k = \frac{1}{L}\langle L^s_{i,12} - P^s_{i,12}\rangle_{arc}. \tag{A.36}$$

The averaged difference for GPS L1, L2 can be written as

$$\langle L^s_{i,12} - P^s_{i,12}\rangle_{arc} = CPB_{i,12} - DCB^s_{i,12} + e^s_{i,L_{12}} - e^s_{i,L_{12}} \tag{A.37}$$

where $DCB^s_{i,12}$ is contant along the continuous satellite arc and the error quantity $(e^s_{i,L_{12}} - e^s_{i,L_{12}})$ has an expectation value zero. The final phase levelled GPS L1, L2 linear combination of phase observations reads

$$L^s_{i,12} - \langle L^s_{i,12} - P^s_{i,12}\rangle_{arc} = \alpha_1 \; STEC(\mathbf{x}_i, \mathbf{x}^s) + DCB^s_{i,12} + e^s_{i,L_{12}} \tag{A.38}$$

Even though the observable conventions used are slightly different, the Eqs. (A.33a) to (A.38) are the same as used by Limberger (2015), Liang (2017), Erdogan et al. (2017) and Goss et al. (2020). The accuracy of STEC from (A.38) in-turn depends on the accuracy of fixing ambiguities $N^s_{i,1}$, $N^s_{i,2}$ to the respective correct integer values (Henkel (2010)).

Ambiguity resolution

The ambiguity term is one of the most relevant terms in the Eqs. (A.21b), (A.21d) and (A.21f) with regard to quality of the observations. Ambiguity arises due to an offset between the transmitted and received signal phase by an unknown integer number of carrier cycles. While the fractional part of phase is continuously being tracked by the receiver, an additional unknown parameter for the phase ambiguity has to be considered in the phase observation equations (A.21b), (A.21d) and (A.21f) for each satellite - receiver pair. There is a direct impact of cycle slips, due to high dynamics receivers (e.g. LEO occultation receivers), on ambiguity resolution. Yue et al. (2016) reports a contamination of ≈ 23% of IRO observations due to cycle slips, thus degrading the IRO observation quality as well as the quantity.

In summary of STEC computation, the following points are noteworthy:

- STEC has been computed from L1-L2 and/or L1-L5 GFLC and contains the unknown biases as shown in (A.33a) and (A.33b).
- Differential code biases are unknown parameters to be estimated together with other unknown parameters within the EDM
- Phase biases are absorbed by integer ambiguities and any additional unmodelled biases are estimated together with other unknown parameters within the EDM.

Ambiguity resolution is one of the most important steps in processing GNSS phase observations and is a detailed subject by itself; see Dach et al. (2009), Prange et al. (2017). In this work, carrier levelling is used for the ambiguity resolution and has been found to be satisfactory for ionosphere

[43] An obvious question that arises is "Why is there differential code bias in STEC computed from GFLC phase observations"? The answer is that the code biases are entering (A.32) through phase levelling and thus the STEC computed from phase observations, contains code biases.

modelling by Limberger (2015), Liang (2017), Erdogan et al. (2017), Erdogan et al. (2020). For a detailed analysis on other ambiguity resolution and validation methods, the reader is referred to e.g., Blewitt (1989), Teunissen (1995), Ge et al. (2005), Teunissen and Odijk (2003), Verhagen (2005), Verhagen and Teunissen (2006), Teunissen (2017).

E GNSS STEC for modelling the B-spline coefficients

In Section D.2, the observation equations for computing $\widetilde{STEC}(\mathbf{x}_W, \mathbf{x}^s)$ has been described. In this section, the procedure for including $\widetilde{STEC}(\mathbf{x}_W, \mathbf{x}^s)$ in EDM will be described. Specifically, this section will provide a GMM for the estimation of the unknown B-spline coefficients as well as the bias parameters.

For the sake of simplicity, the discussion will refer to L1-L2 biases but the same apply also any other dual frequency, e.g. to GPS L1-L5 or Galileo E1-E5a or E1-E5b combinations. R_i, S_j denote the receiver and satellite indices with respective position vectors $\boldsymbol{x}_i$ and $\boldsymbol{x}^j$. Previously, $STEC(\boldsymbol{x}_r, \boldsymbol{x}^s)$ was used to denote STEC between receiver r and satellite s. For the sake of consistency in the representation of STEC and DCBs, the following convention for STEC will be applied and $STEC(\mathbf{x_r}, \mathbf{x^s})$ will be replaced with $STEC(R_i, S_j)$.

The first step will be the grouping of the bias related terms into satellite and receiver specific ones, so that they could be consistently represented in the resulting GMM. GPS L1-L2 receiver DCB is denoted as

$$DCB_{R_i} = b^P_{i,12}, \tag{A.39}$$

and that of the satellite as

$$DCB_{S_j} = b^{P,j}_{12}, \tag{A.40}$$

without the explicit frequency subscript. Analogously,

$$DPB_{R_i} = b^L_{i,12} \tag{A.41}$$

and

$$DPB_{S_j} = b^{L,j}_{12} \tag{A.42}$$

denote the differential phase biases for receiver i and satellite j respectively. More generally, the differential code bias (DCB) vector[44]

$$\boldsymbol{DCB} = \begin{bmatrix} \boldsymbol{DCB}_{R_i} \\ \boldsymbol{DCB}_{S_j} \end{bmatrix} \tag{A.43}$$

denotes the accumulated biases in the STEC computation, where,

$$\begin{aligned} \boldsymbol{DCB}_{R_i} &= \begin{bmatrix} DCB_{R_i} & DPB_{R_i} \end{bmatrix}^T \ \forall\ i \in \{1, \cdots, rmax\} \\ \boldsymbol{DCB}_{S_j} &= \begin{bmatrix} DCB_{S_j} & DPB_{S_j} \end{bmatrix}^T \ \forall\ j \in \{1, \cdots, smax\} \end{aligned} \tag{A.44}$$

Substituting the Eqs. (A.39) to (A.43) into Eq. (A.32) gives the observation equation

$$STEC(R_i, S_j) + e_{STEC}(R_i, S_j) = \int_{S_j}^{R_i} N_e(\varphi, \lambda, h) ds + \boldsymbol{DCB}_{R_i}\ \mathbf{1} + \boldsymbol{DCB}_{S_j}\ \mathbf{1}, \tag{A.45}$$

where $\mathbf{1} = \begin{bmatrix} 1 & 1 \end{bmatrix}$. As described already in Chapter 4 and 5, the unknown B-spline coefficients are the main unknown parameters. In this section, the unknown parameters are extended to include the biases, and hence along with their corresponding initial values, are denoted by

$$\boldsymbol{\beta}_{STEC} = \begin{bmatrix} \boldsymbol{d}^T_\kappa & \boldsymbol{DCB}^T \end{bmatrix}^T \tag{A.46a}$$

$$\boldsymbol{\beta}_{STEC;0} = \begin{bmatrix} \boldsymbol{d}^T_{\kappa;0} & \boldsymbol{DCB}^T_0 \end{bmatrix}^T. \tag{A.46b}$$

[44] A change in convention to denote STEC and biases of receivers and satellite using the subscript R_i and S_j is necessary in this section for better readability.

Using the STEC computed from the initial values (A.46b),

$$STEC_0(R_i, S_j) = \int_{S_j}^{R_i} N_e(\varphi, \lambda, h) ds \Big|_{\boldsymbol{\beta}_{STEC,0}}, \tag{A.47}$$

the corrections

$$\begin{aligned} \Delta \boldsymbol{d}_\kappa &= \boldsymbol{d}_\kappa - \boldsymbol{d}_{\kappa;0} \\ \Delta \boldsymbol{DCB}_{R_i} &= \boldsymbol{DCB}_{R_i} - \boldsymbol{DCB}_{R_i;0} \\ \Delta \boldsymbol{DCB}_{S_j} &= \boldsymbol{DCB}_{S_j} - \boldsymbol{DCB}_{S_j;0} \end{aligned} \tag{A.48}$$

respectively, the linearized STEC observation equation for a single receiver-satellite (R_i, S_j) line of sight reads

$$\begin{aligned} STEC(R_i, S_j) = {} & STEC_0(R_i, S_j) + \boldsymbol{DCB}_{R_i;0}\ \mathbf{1} + \boldsymbol{DCB}_{S_j;0}\ \mathbf{1} \\ & + \frac{\partial STEC(R_i, S_j)}{\partial \boldsymbol{d}_\kappa}\Big|_{\boldsymbol{\beta}_{STEC;0}} \Delta \boldsymbol{d}_\kappa \\ & + \frac{\partial STEC(R_i, S_j)}{\partial \boldsymbol{DCB}_{R_i}}\Big|_{\boldsymbol{\beta}_{STEC;0}} \Delta \boldsymbol{DCB}_{R_i} \\ & + \frac{\partial STEC(R_i, S_j)}{\partial \boldsymbol{DCB}_{S_j}}\Big|_{\boldsymbol{\beta}_{STEC;0}} \Delta \boldsymbol{DCB}_{S_j}. \end{aligned} \tag{A.49}$$

Similarly, a system of equations for multiple receiver-satellite $(R_i, S_j \ \forall \ i \in \{1, \cdots, rmax\}$ and $j \in \{1, \cdots, smax\})$ line of sights are formed. Since the B-spline coefficients are collected together for multiple key parameters, corresponding to the set $\mathcal{K}_1$, the term corresponding to the partial derivative of STEC with respect to the B-spline coefficients is written as (omitting the indices (R_i, S_j) for better readability)

$$\frac{\partial STEC}{\partial \boldsymbol{d}_\kappa}\Big|_{\boldsymbol{\beta}_{STEC;0}} = \left[\left(\frac{\partial STEC}{\partial \boldsymbol{d}_{\kappa_1}}\right)^T \quad \left(\frac{\partial STEC}{\partial \boldsymbol{d}_{\kappa_2}}\right)^T \quad \cdots \quad \left(\frac{\partial STEC}{\partial \boldsymbol{d}_{\kappa_p}}\right)^T\right]\Big|_{\boldsymbol{\beta}_{STEC;0}} \tag{A.50}$$

and using the chain rule of calculus,

$$\begin{aligned} & \frac{\partial STEC}{\partial \boldsymbol{d}_\kappa}\Big|_{\boldsymbol{\beta}_{STEC;0}} = \\ & \left[\left(\frac{\partial STEC}{\partial \kappa_1}\frac{\partial \kappa_1}{\partial \boldsymbol{d}_{\kappa_1}}\right)^T \quad \left(\frac{\partial STEC}{\partial \kappa_2}\frac{\partial \kappa_2}{\partial \boldsymbol{d}_{\kappa_2}}\right)^T \quad \cdots \quad \left(\frac{\partial STEC}{\partial \kappa_p}\frac{\partial \kappa_p}{\partial \boldsymbol{d}_{\kappa_p}}\right)^T\right]\Big|_{\boldsymbol{\beta}_{STEC;0}}. \end{aligned} \tag{A.51}$$

The terms $\frac{\partial \kappa_q}{\partial \boldsymbol{d}_{\kappa_q}} \ \forall \ q = \{1, \cdots, p\}$ are the tensor products of the polynomial and trigonometric B-splines denoted as $N(\varphi)T(\lambda)$. Accordingly, Eq. (A.49) is written as

$$\begin{aligned} STEC(R_i, S_j) = {} & \int_{S_j}^{R_i} N_e(\varphi, \lambda, h)\Big|_{\boldsymbol{d}_{\kappa,0}} ds \\ & + \int_{S_j}^{R_i} \left(\sum_{q=1}^{p} \frac{\partial N_e(\varphi, \lambda, h)}{\partial \boldsymbol{d}_{\kappa_q}}\Big|_{\boldsymbol{d}_{\kappa_q,0}} \Delta \boldsymbol{d}_{\kappa_q}\right) N(\varphi)T(\lambda) ds \\ & + \left[\boldsymbol{DCB}_{R_i;0}\ \mathbf{1} + \frac{\partial STEC(R_i, S_j)}{\partial \boldsymbol{DCB}_{R_i}}\Big|_{\boldsymbol{\beta}_{STEC;0}} \Delta \boldsymbol{DCB}_{R_i}\right] \\ & + \left[\boldsymbol{DCB}_{S_j;0}\ \mathbf{1} + \frac{\partial STEC(R_i, S_j)}{\partial \boldsymbol{DCB}_{S_j}}\Big|_{\boldsymbol{\beta}_{STEC;0}} \Delta \boldsymbol{DCB}_{S_j}\right]. \end{aligned} \tag{A.52}$$

Substituting Eq. (A.47) into (A.52) gives,

$$\begin{aligned} STEC(R_i,S_j) &= STEC_0(R_i,S_j) \\ &+ \int_{S_j}^{R_i} \left(\sum_{q=1}^{p} \left. \frac{\partial N_e(\varphi,\lambda,h)}{\partial \boldsymbol{d}_{\kappa_q}} \right|_{\boldsymbol{d}_{\kappa_q,0}} \Delta \boldsymbol{d}_{\kappa_q} \right) N(\varphi) T(\lambda) ds \\ &+ \left[\boldsymbol{DCB}_{R_i;0}\ \mathbf{1} + \left. \frac{\partial STEC(R_i,S_j)}{\partial \boldsymbol{DCB}_{R_i}} \right|_{\beta_{STEC;0}} \Delta \boldsymbol{DCB}_{R_i} \right] \\ &+ \left[\boldsymbol{DCB}_{S_j;0}\ \mathbf{1} + \left. \frac{\partial STEC(R_i,S_j)}{\partial \boldsymbol{DCB}_{S_j}} \right|_{\beta_{STEC;0}} \Delta \boldsymbol{DCB}_{S_j} \right], \end{aligned} \tag{A.53}$$

and further rearranging leads to

$$\begin{aligned} &STEC(R_i,S_j) - \left(STEC_0(R_i,S_j) + \boldsymbol{DCB}_{R_i;0}\ \mathbf{1} + \boldsymbol{DCB}_{S_j;0}\ \mathbf{1} \right) \\ &\quad = \int_{S_j}^{R_i} \left(\sum_{q=1}^{p} \left. \frac{\partial N_e(\varphi,\lambda,h)}{\partial \boldsymbol{d}_{\kappa_q}} \right|_{\boldsymbol{d}_{\kappa_q,0}} \Delta \boldsymbol{d}_{\kappa_q} \right) N(\varphi) T(\lambda) ds \\ &\quad + \left[\left. \frac{\partial STEC(R_i,S_j)}{\partial \boldsymbol{DCB}_{R_i}} \right|_{\beta_{STEC;0}} \Delta \boldsymbol{DCB}_{R_i} \right] \\ &\quad + \left[\left. \frac{\partial STEC(R_i,S_j)}{\partial \boldsymbol{DCB}_{S_j}} \right|_{\beta_{STEC;0}} \Delta \boldsymbol{DCB}_{S_j} \right] \end{aligned} . \tag{A.54}$$

Substituting

$$\Delta STEC(R_i,S_j) = STEC(R_i,S_j) - \left(STEC_0(R_i,S_j) + \boldsymbol{DCB}_{R_i;0}\ \mathbf{1} + \boldsymbol{DCB}_{S_j;0}\ \mathbf{1} \right), \tag{A.55}$$

$$\begin{bmatrix} y_{STEC_1} \\ \vdots \\ y_{STEC_{N_{LOS,max}}} \end{bmatrix} = \begin{bmatrix} STEC(R_1,S_1) \\ \vdots \\ STEC(R_{rmax},S_{smax}) \end{bmatrix} \tag{A.56}$$

and

$$\begin{bmatrix} e_{STEC_1} \\ \vdots \\ e_{STEC_{N_{LOS,max}}} \end{bmatrix} = \begin{bmatrix} e_{STEC}(R_1,S_1) \\ \vdots \\ e_{STEC}R_{rmax},S_{smax} \end{bmatrix} \tag{A.57}$$

for better readability, where $N_{LOS,max} = rmax{\cdot}smax$ is the actual number of LoS ray paths. Following the convention of using the vector variable $\boldsymbol{y}$ and $\boldsymbol{e}$ to denote observations and measurement errors respectively, the corrected STEC observations vector reads

$$\Delta \boldsymbol{y}_{STEC} = \begin{bmatrix} \Delta y_{STEC_1} & \cdots & \Delta y_{STEC_{N_{LOS,max}}} \end{bmatrix}^T \tag{A.58a}$$

$$= \begin{bmatrix} \left(y_{STEC_1} - y_{STEC_1}|_{\beta_{STEC;0}} \right) & \cdots & \left(y_{STEC_{N_{LOS,max}}} - y_{STEC_{N_{LOS,max}}}|_{\beta_{STEC;0}} \right) \end{bmatrix}^T \tag{A.58b}$$

and the corresponding measurement error vector

$$\boldsymbol{e}_{STEC} = \begin{bmatrix} e_{STEC_1} & \cdots & e_{STEC_{N_{LOS,max}}} \end{bmatrix}^T . \tag{A.59}$$

The corrections

$$\Delta \boldsymbol{\beta}_{STEC} = \begin{bmatrix} \Delta \boldsymbol{d}_{\kappa}^T & \Delta \boldsymbol{DCB}^T \end{bmatrix}^T , \tag{A.60}$$

where

$$\Delta \boldsymbol{DCB} = \left[\Delta \boldsymbol{DCB}_{R_1} \quad \cdots \quad \Delta \boldsymbol{DCB}_{R_{rmax}} \quad \Delta \boldsymbol{DCB}_{S_1} \quad \cdots \quad \Delta \boldsymbol{DCB}_{S_{smax}}\right]^T, \tag{A.61}$$

to the unknown parameters $\boldsymbol{\beta}_{STEC} = \left[\boldsymbol{d}_\kappa^T \quad \boldsymbol{DCB}^T\right]^T$, are computed about initial values $\boldsymbol{\beta}_{STEC}$ such that

$$\boldsymbol{DCB}_0 = \left[\boldsymbol{DCB}_{R_1;0} \quad \cdots \quad \boldsymbol{DCB}_{R_{rmax};0} \quad \boldsymbol{DCB}_{S_1;0} \quad \cdots \quad \boldsymbol{DCB}_{S_{smax};0}\right]^T. \tag{A.62}$$

With above mentioned matrices and vectors, we get the Gauss-Markov model

$$\boldsymbol{A}_{STEC}\ \Delta\boldsymbol{\beta}_{STEC} = \Delta\boldsymbol{y}_{STEC} + \boldsymbol{e}_{STEC} \quad \text{with} \quad D(\Delta\boldsymbol{y}_{STEC}) = \sigma^2_{STEC}\boldsymbol{P}^{-1}_{STEC}, \tag{A.63}$$

where

$$\boldsymbol{A}_{STEC} = \left[\boldsymbol{A}_{STEC,\boldsymbol{d}_\kappa} \quad \boldsymbol{A}_{STEC,\boldsymbol{DCB},Rx} \quad \boldsymbol{A}_{STEC,\boldsymbol{DCB},Sat}\right] \tag{A.64}$$

such that

$$\boldsymbol{A}_{STEC,\boldsymbol{d}_\kappa} = \begin{pmatrix} \frac{\partial STEC(R_1,S_1)}{\partial \boldsymbol{d}_{\kappa_1}} & \cdots & \frac{\partial STEC(R_1,S_1)}{\partial \boldsymbol{d}_{\kappa_p}} \\ \vdots & \ddots & \vdots \\ \frac{\partial STEC(R_{rmax},S_{smax})}{\partial \boldsymbol{d}_{\kappa_1}} & \cdots & \frac{\partial STEC(R_{rmax},S_{smax})}{\partial \boldsymbol{d}_{\kappa_p}} \end{pmatrix} \tag{A.65}$$

$$\boldsymbol{A}_{STEC,\boldsymbol{DCB},Rx} = \begin{pmatrix} \frac{\partial STEC(R_1,S_1)}{\partial \boldsymbol{DCB}_{R_1}} & \cdots & \frac{\partial STEC(R_1,S_1)}{\partial \boldsymbol{DCB}_{R_{rmax}}} \\ \vdots & \ddots & \vdots \\ \frac{\partial STEC(R_{rmax},S_{smax})}{\partial \boldsymbol{DCB}_{R_1}} & \cdots & \frac{\partial STEC(R_{rmax},S_{smax})}{\partial \boldsymbol{DCB}_{R_{rmax}}} \end{pmatrix} \tag{A.66}$$

$$\boldsymbol{A}_{STEC,\boldsymbol{DCB},Sat} = \begin{pmatrix} \frac{\partial STEC(R_1,S_1)}{\partial \boldsymbol{DCB}_{S_1}} & \cdots & \frac{\partial STEC(R_1,S_1)}{\partial \boldsymbol{DCB}_{S_{smax}}} \\ \vdots & \ddots & \vdots \\ \frac{\partial STEC(R_{rmax},S_{smax})}{\partial \boldsymbol{DCB}_{S_1}} & \cdots & \frac{\partial STEC(R_{rmax},S_{smax})}{\partial \boldsymbol{DCB}_{S_{smax}}} \end{pmatrix}. \tag{A.67}$$

Both the receiver and satellite DCBs are assumed to be constant for 6 hours and estimated along with the B-spline coefficients of the key parameters. The stochastic model of the STEC observations is defined by the covariance matrix $D(\Delta\boldsymbol{y}_{STEC})$, which includes the given positive definite weight matrix $\boldsymbol{P}_{STEC}$ and the unknown variance factor σ^2_{STEC}.
In order to avoid rank deficiency and separate the receiver and satellite DCBs from each other, an additional zero-mean constraint,

$$\begin{aligned} b_{12}^{P,s_1} + \cdots + b_{12}^{P,s_{total}} &= 0 \\ b_{12}^{L,s_1} + \cdots + b_{12}^{L,s_{total}} &= 0 \end{aligned} \tag{A.68}$$

is applied over the total number of operational satellites s_{total} for a given constellation[45] (see Schaer (1999)).

[45] For the satellites not in view from a given receiver station, the corresponding DCB values from IGS are used instead. It shall be further noted that at all times $smax \leq s_{total}$ is satisfied.