UNIVERSIDADE
BEIRA INTERIOR

Robust Kalman Filtering in the Framework of Space Communications
(Versão Final Após Defesa)

Clara Braga Cardoso

Dissertação para obtenção do Grau de Mestre em
Engenharia Aeronáutica
(Mestrado Integrado)

Orientador: Prof. Doutor Kouamana Bousson

Abril de 2024.

Declaração de Integridade

Eu, Clara Braga Cardoso, que abaixo assino, estudante com o número de inscrição 40222 de Engenharia Aeronáutica da Faculdade de Engenharia, declaro ter desenvolvido o presente trabalho e elaborado o presente texto em total consonância com o **Código de Integridades da Universidade da Beira Interior**.

Mais concretamente afirmo não ter incorrido em qualquer das variedades de Fraude Académica, e que aqui declaro conhecer, que em particular atendi à exigida referenciação de frases, extratos, imagens e outras formas de trabalho intelectual, e assumindo assim na íntegra as responsabilidades da autoria.

Universidade da Beira Interior, Covilhã 3 de abril 2024

Dedication

To my parents, Alzira and Paulo, for all of their support, for their belief in me and for all their dedication to my education. I would not be at this stage of my life without their love and support.

For everything they did for me, this dissertation is my accomplishment as much it is theirs.

Thank you.

Acknowledgements

I would like to deeply thank my supervisor, Professor Kouamana Bousson, for this amazing opportunity to develop my work in an area I've been forever interested about, his dedication in helping me , all his teachings, opinions and critics that allowed me to finish my degree.

I would also like to give a big shout out to my friends that walked with me through this journey, suffering together is definitely better.

A big thank you to my boyfriend, Filipe, for his support, for his belief in me and for putting up with all the stressful moments that happened while working on this dissertation. I would like to thank his family as well.

A special thank you to my sister, Carlota, that always blindly believes in me and my abilities and all the joy she brings into my life.

I am also specially grateful for my parents, Paulo and Alzira, for all their support and given motivation in pursuing my academic career.

Finally, I would like to thank myself, I did an amazing job.

Resumo

Os satélites tornaram-se numa forma ideal de comunicação (sendo vitais em área remotas do mundo), navegação e monitorização do clima, tendo também outras funcionalidades e influencia no planeta.

Um satélite de comunicação age como uma estação de repetição de microondas que permite a dois ou mais utilizadores com um segmento terrestre para trocarem informação, funcionando como um canal/intermediário entre localidades distantes- as ondas de rádio 'viajam' por 'linha de vista' e podem ser obstruídas pela curvatura da Terra. Ainda assim, o interesse nas comunicações via satélite permanece devido as vantagens associadas a estes sistemas.

Um sistema de comunicação espacial divide-se em duas partes: o segmento espacial e o segmento terrestre. A utilização de sinais microondas permite uma comunicação móvel e sem fios, desde que o segmento na Terra esteja contido na zona de influência do satélite é possível estabelecer uma conexão efetiva entre o satélite e o utilizadores. Os satélites podem funcionar para qualquer região e podem cobrir áreas vastas.

Comummente, um satélite de comunicação tem uma órbita terreste baixa (LEO), com uma altitude igual ou inferior a 1500km, relativamente à Terra. Uma órbita real LEO distancia-se de uma órbita ideal (órbita Kepleriana) pois, para além da influência das forças gravitacionais entre os dois corpos, a trajetória é influencia por perturbações como a resistência aerodinâmica, a não-esfericidade do planeta Terra e a interação com o Sol e a Lua.

Numa trajetória LEO, a resistência aerodinâmica é a principal força não-gravitacional que age sobre o satélite: a força aplicada no satélite pelas moléculas do ar tem direção oposta ao vetor da velocidade o que pode levar a um aumento da velocidade do satélite.

A não-esfericidade da Terra deve-se á força centrífuga originada pela rotação do planeta. Isto combinado com anomalias de massa periféricas, como os continentes levam a que o centro de massa do planeta Terra é diferente do centro geométrico. Isto implica que a força da gravidade aplicada num corpo orbital não está direcionada a mesma. O segundo harmónico esférico, J_2, é um parâmetro adimensional que quantifica os maiores efeitos devidos à não-esfericidade do planete numa órbita.

Estas perturbações têm que ser contabilizadas nas equações de movimento do satélite (modelo dinâmico).

Outro ponto a ter em conta num sistema de comunicação espacial são os fatores que têm que ser considerados para obter um canal de comunicação: a frequência de operação, os ângulos de elevação, a localização geográfica do segmento terrestre e o clima.

Os meios de propagação têm um impacto em tempo real na operação, design e performance dos sistemas espaciais que operam em frequências de rádio e óticas. Portanto, para implementar um modelo de transmissão eficiente é necessário contabilizar com os efeitos de atenuação da: ionosfera, troposfera e efeitos de atenuação locais.

Como a trajetória do satélite pode alterar-se devido às perturbações referidas, é imperativo estimar e prever a órbita do satélite para saber os possíveis desvios de trajetória e quando corrigí-los, de modo a não perturbar as funções do satélite e os canais de transmissão. Atualmente, o filtro de Kalman é uma opção confiável para uma estimação estocástica a partir de medições ruidosas. Este filtro é um algoritmo recursivo capaz de incorporar toda a informação que recebe e conseguir obter os valores, neste caso, dos componentes de uma trajetória. Este algoritmo combina todos os dados disponíveis com todo o conhecimento que tinha *a priori* do sistema produzindo uma estimação com um erro estatisticamente minimizado.

As equações deste filtro dividem-se em dois grupos: equações *time update* e equações *measurement update* . As primeiras tem a função de projetar o estado e a covariância do erro atual para obter as estimações a priori para o próximo passo. As segundas são responsáveis para incorporar um valor medido com a estimação *a priori* visando obter uma estimação *a posteriori* melhorada.

No entanto, apesar da sua alta funcionalidade, a forma discreta do filtro de Kalman apresenta algumas limitações que podem a induzir a estimações com um erro maior.

A maior limitação do filtro é que a sua optimalidade existe quando o filtro é projetado para modelos lineares. No entanto, em aplicações reais, os modelos dinâmicos de um sistema contêm certas incertezas que levam a uma diminuição da performance do filtro clássico. Para contrapor isso, é necessário converter o sistema não-linear num sistema de equações lineares que mais se aproxime do modelo, contando com uma diminuição de precisão.

Para contrapor essas limitações, uma das opções disponíveis é o filtro de Kalman robusto. O algoritmo para a forma robusta do filtro consiste em dois passos: primeiramente, gera-se uma nova trajetória que contabilize essas incertezas, e um novo estimador é criado; depois, uma nova trajetória é filtrada recorrendo ao filtro de Kalman discreto. Esta forma do filtro robusto consiste em adicionar ruído artificialmente, forçando o algoritmo para compensar pelo ruído, e gerando novas estimações com o ruído praticamente eliminado.

Esta dissertação foca-se na aplicação do filtro a um satélite de comunicação e lida com as incertezas de um sistema com medições ruidosas. Apesar da aplicação clássica do filtro resultar numa estimação que quase elimina todo o ruído, procedeu-se a uma aplicação robusta para aprimorar a performance do filtro de Kalman.

Palavras-chave

Comunicação Espacial, Órbita terreste baixa (LEO), Satélite LEO, Estimação orbital, Sistemas não-lineares, Filtro de Kalman Discreto, Filtro de Kalman robusto

Abstract

A Low-Earth orbit (LEO) is the trajectory of a space vehicle relative to the Earth, with an altitude equal or below 1500 km. A real LEO trajectory is far from an ideal orbit (Keplerian Orbit), because in addition to the gravitational forces from both bodies, the orbit is influenced by perturbation like the oblateness of the Earth and the atmospheric drag. These perturbations alter the planned trajectory, making it imperative to estimate and predict a satellite trajectory to know its possible deviations and when to correct it.
Currently, the Kalman filter is a reliable option for a trajectory estimation. Although it is an algorithm simple to implement, it is necessary to know all the parameters of the system. So, the efficiency of the discrete Kalman filter depends on practically the full knowledge of the systems specifications: which, is pratically impossible in a aerospace context. This entails that it bust be implemented a Kalman filter dynamic enough to endure the simplication of the unknown parameters of the system without damage and increase the discrepancy between the real measured values and the estimated results from the filter.
This dissertation focus on the Kalman filter application to a communication satellite and how the uncertainties associated to a system with noisy measurements are handled. And, although tht classical form of the filter results in a estimation that very well decreases all the noise, it was also implemented a robust application to improve tha Kalman filter's performance.

Keywords

Space communications, Low-Earth Orbit, LEO satellite, Orbit estimation, Non-linear systems, Discrete Kalman filter, Robust Kalman filter

Contents

List of Figures

List of Tables

List of Acronyms

APSK	Amplitude and Phase Shift Keying
ATS	Application Technology Satellite
BSS	Broadcasting Satellite Service
CDMA	Code Division Multiple Access
CAI	Common Air Interface
CIR	Carrier-to-Interference Ratio
FSS	Fixed Satellite Service
FDMA	Frequency Division Multiple Access
FSK	Frequency Shift Keying
GMAT	General Mission Analysis Tool
GEO	Geosynchronous Orbit
ITU	International Telecommunication Union
LMS	Land Mobile Satellite
LOS	Line of Sight
LEO	Low-Earth Orbit
MEO	Middle-Earth Orbit
MSS	Mobile Satellite Service
PSK	Phase Shift Keying
SCC	Satellite Control Center
TDMA	Time Division Multiple Access
TT & C	Track, Telemetry and Command
UHF	Ultra-High Frequency
VHF	Very-High Frequency

Nomenclature

Symbol	Description
$\Gamma(.)$	Gamma Function
μ	Gravitational Parameter
μ	Robust Parameter
ν_k	Measurement Noise (at step k)
ν	True Anomaly
ρ	Density
σ	Standard Deviation
ω_c	Angular Frequency
ω	Argument of the Perigee
Ω	Right Ascension of the Ascending Node
a	Amplitude
a_D	Drag's Acceleration
a_e	Earth's Equatorial Radius
a	Semimajor Axis
A	Cross-Sectional Area
$\bar{A}$	Artificially Noise State Transition Matrix
A_d	Discretized State Transition Matrix
A	State Transition Matrix
B / D	Control Matrix
B_d/D_d	Discretized Control Matrix
C_D	Drag's Coefficient
C	Output Matrix
e	Eccentricity
E	Eccentric Anomaly
h_a	Apogee Height

h_p	Perigee Height
I / I_d	Identity Matrix
I_0	Modified Bessel Function
J_2	Jacobian Matrix
$_k$	State Estimate
K	Gain Matrix
K	Rician Factor
m	Mass
M	Mean Anomaly
P_k	Error Covariance Matrix (at step k)
P_0	Initial Error Covariance Matrix
Q	Process Noise Covariance Matrix
r	Signal Envelope
R	Measurement Error Covariance Matrix
t	Time
u_k	Control Input (at step k)
V	Velocity
x_k	State Vector
x,y,z	Relative Position Components in Cartesian Coordinates
$\dot{x}$, $\dot{y}$, $\dot{z}$	Relative Velocity Components in Cartesian Coordinates
$\ddot{x}$, $\ddot{y}$, $\ddot{z}$	Relative Acceleration Components in Cartesian Coordinates
r, θ, ϕ	Spherical Coordinates
$\dot{r}$, $\dot{\theta}$, $\dot{\phi}$	First Derivative of Spherical Coordinates
$\ddot{r}$, $\ddot{\theta}$, $\ddot{\phi}$	Second Derivative of Spherical Coordinates

Chapter 1. Introduction

In the past 70 years, space sciences and exploration resulted into great achievements, the inventions resulted from it have turned life on Earth more sustainable, easier and safer. Satellites have become a great mean for communicating around the world (being vital in remote areas), navigation, climate and environmental monitoring, having many other applications and influence in our daily lives. On account of all of those benefits and to constantly evolve our knowledge and technology it is necessary that we keep improving space vehicles including its communication systems.

A communications satellite acts as a microwave repeater station that allows two or more users with appropriate earth stations to exchange information, functioning as a channel/intermediate between widely separated locations - radio waves travel by line of sight, and can be obstructed by the Earth's curvature. The interest in relaying on satellite communications comes from the advantages associated with it [1].

The employ of microwave signals allows a wireless and mobile communication, as long as an Earth station is within the satellite's footprint it can be established an effective radio connection between the satellite and the user. Satellites can also work on any size region, being able to cover wide areas if designed properly, as well as having an ample frequency spectrum availability to apply. As referred above, satellite communications play a crucial role in modern life. Satellites are constantly used for a variety of applications, due to its assets such as the ability of covering a wide area and the use of radio telecommunication signals allowing a wireless communication. Other attractive features are the independence from terrestrial infrastructure and the rapid and low cost. installation of ground networks: as soon as the satellite is operational, a ground segment can easily be operated, not requiring an extensive ground construction like a terrestrial infrastructure would. Thanks to this last benefit, satellite communication has a key role in restoring communications after a disaster [1].

1.1 Communication Satellites Evolution and Components

1.1.1 Technological Evolution and Early Programs

This section chronicles a basic timeline of programs and experiments crucial to the developments of communication satellites. As it was referred above, the industry has grown massively, and now communication satellites now encompasses space technology, navigation, communication and information systems and the Internet. So its important to be familiar with past satellite programs to better understand today´s systems and services:

- SCORE: first communication by artificial satellite, powered by battery only. It relayed a recorded voice message, on a delayed basis, linking one Earth station to another. The relay performed on a 150 MHz uplink and 108 MHz downlink.

- ECHO (1 and 2): was the first experiment to evaluate communications relay by passive techniques. The satellites acted as passive reflectors for signals transmitted from Earth stations, the relays operated from 162 MHz to 2390 MHz frequencies and required large ground terminal antennas.

- COURIER: expanded SCORE delayed repeater technology and researched storeandforward and real-time capabilities on a low orbit. Operated on a 1.8 to 1.9 GHz uplink frequency and on a 1.7 to 1.8 downlink frequency. The satellite was the first to involve solar cells for power, and had two-watt output power tubes.

- WESTFORD: second artificial satellite to investigate communications relay by passive techniques. The communication was accomplished by reflection from dispersed dipole reflectors in an orbital belt. The fast development of active communications satellites reduced interest in passive techniques.

- TELSTAR (1 and 2): the satellites were launched into a low orbit and were the first active wideband communications satellites. The analog FM signals were relayed

with a 50 MHz bandwidth, and performed at frequencies of 6.4 GHz on the uplink and 4.2 GHz on the downlink. These frequencies initiated the 6/4 GHz C-band operation, responsible for a major portion of fixed satellite services (FSS). The TELSTAR 1 failed due to Van Allen belt radiation, leading to a redesign of ECHO 2 with radiation resistant transistors and to launch it to a higher orbit.

- RELAY (1 and 2): had two redundant repeaters with two 2MHz channels and on at 25 MHz. Performed with a 1725 MHz uplink and a 4160 MHz downlink. RELAY and TELSTAR proved that routine communications were successfully accomplished from orbiting satellites and connoted that satellite systems could share frequencies with terrestrial communication systems without interference.

- SYNCOM (1 to 3): established the first communications from a synchronous satellite. The satellites employed one 5MHz channel for one-way wideband transmission and two 500 kHz channels for two-way narrowband communication, and operated at frequencies of 7.4 GHz uplink and 1.8 GHz downlink and was the first to study orbital control principles for synchronous satellites.

- EARLYBIRD: first commercial operational synchronous communications satellite. The communication system performed at C-band with frequencies of 6.3 GHz uplink and 4.1 GHz downlink and had two 25MHz transponders.

- ATS-1 (Applications Technology Satellite 1): provided the first multiple access communications on a synchronous orbit. The system incorporated an electronically despun antenna with 18-dB gain and a 17^{0} beamwidth and operated at C-band (6.3 GHz uplink and 4.1 GHz downlink), with two 25MHZ repeaters. ATS-1 also had VHF (Very High frequency) links (149 MHz uplink and 136 MHz downlink) in order to evaluate the air to ground communications.

- ATS-3: Continued experimental techniques in the C and VHF bands, with multiple access communications and orbit control techniques. The satellite gave the first colour high-resolution pictures of Earth and was the first to be able to ‘cross-strap’

at C-band and VHF (the signal received at VHF could be transmitted to the ground at C-band).

- ATS-5: had a C-band communications subsystem and a L-band subsystem (1650 MHz uplink, 1550 MHz downlink) to research air to ground communication for navigation and air traffic control. ATS-5 was also designed with a millimeter wave experiment that operated at a 31.65 GHz uplink and 15.3 GHz downlink, to supply propagation data on the effects of the atmosphere on earth-space communications in this band of frequencies.

- ANIK A: first domestic commercial communications satellite, operated at C-band and had 12 transponders, each 36 MHz wide. It provided television distribution and single channel per carrier voice and data services.

- ATS-6: reached major advancements in communications satellite technology. After the launched, the satellite was relocated through the years in order to investigate new possible services. The satellite had eight communication systems for its different propagation experiments that had a frequency range from 860 MHz to 30 GHz. ATS-6 communication subsystem involved four receivers whose frequencies were: 1650 MHz (L-band), 2253 MHz (S-band), 5925-6425 MHz (Cband) and 13/18 GHz (K-band); and four transmitters with the follow frequencies: 860 MHz (L-band), 2063 MHz (S-band), 3953-4153 MHz (C-band), and 20/30 GHz (Ka-band).

These early programs and other events lead to a rapid growth of the industry beginning in the 60´s. The industry in the 70's focused on domestic satellite communications and the technological evolution in this decade brought to consideration the idea of regional satellite communications [9].

The 1980s saw the advent of new and innovative ways to pay for the high costs of satellite systems and services, like private networks (VSATs - very small antenna terminals) and private launch services [9].

The 1990s saw the development of mobile and personal communications services

via satellite. With the increasing data rate requirements, the industry moved to higher RF frequencies to support it as it was reached a saturation on the lower allocated frequency. On-board processing and more advanced techniques were introduced, turning communication satellite into a processing hub in the sky [9].

The rise of the new millennium came with new satellite services and we saw fast advances in non-geosynchronous orbit networks (for global cellular mobile communications).

1.2 Satellite Components

A communication satellite system can be divided in two major parts, the space segment and the ground segment. Fig. 1.1 is a scheme of the main elements of the space segment. After the spacecraft is placed in orbit, the control of the satellite during its lifetime in orbit is quite complex.

The TT &C stations (tracking, telemetry and command) monitors the satellite, by periodically tracking data via the tracking antenna to locate the satellite´s position to anticipate corrections in the position of the satellite. Any satellite can distort its position from the ideal orbit due to gravitational forces from Earth and also the pull of the sun and moon. This station includes separate communications equipment to test and monitor the performance and location of the satellite [1].

The satellite control center (SCC) accommodates the operator controls and equipment for data processing. These stations are commonly located in different places, and the satellite-related data is passed by either terrestrial or satellite data lines. It is common to add a second TT & C to improve the system's reliability.

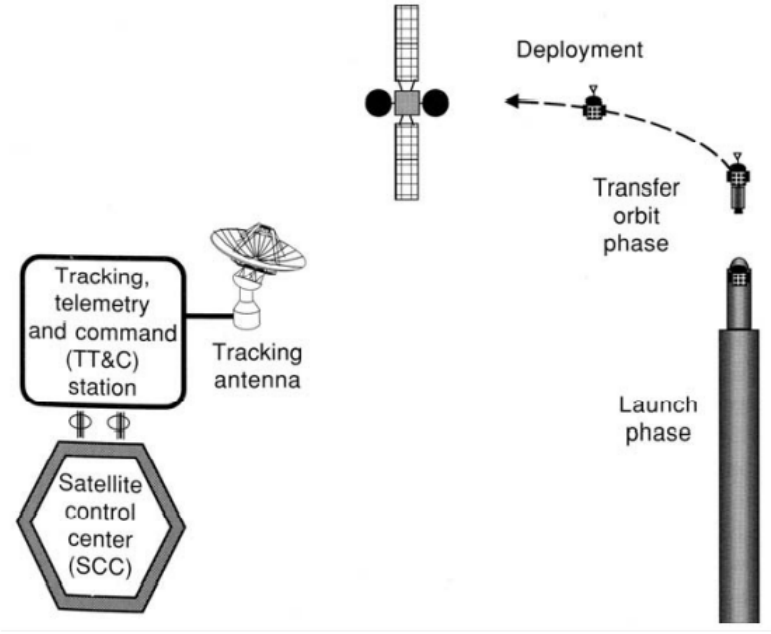

Figure 1.1: Components involved in the space segment. [1]

In the launch phase, it is imperative that the satellite is tracked as soon as it has been released from the last rocket stage. Prior to the placement of a geosynchronous equatorial orbit (GEO), a TT&C station can engage in transfer orbit manoeuvres, while being the station that commands the satellite to arrange its antennas and solar panels. Non-GEO can be launched directly into the desired orbit, refraining from the transfer orbit phase [1].

The ground segment is a diverse collection of facilities, users, and applications that gives access to the satellite repeater from Earth stations to answer communications needs. As it was mentioned before, instead of linking Earth station to Earth station, the links are established through the satellite repeater. An Earth Station can be located at a fixed point (on the ground), designated as a fixed satellite service (FSS), or can be mobile satellite service (MSS) operating on airplanes and ships [7].

Networks used for MSS applications share a common air interface (CAI), which is used for all the links related to and from the satellite. Satellites used in an FSS are able to support various applications within a common bandwidth [1].

Fig. 1.3 illustrates a ground segment, where the satellites network provides two-way interactive services to a variety of Earth stations. One of the larger stations is a control point for a partial part of the network or the complete network, acting as a network man-

agement center. The control Earth station is essential in networks with dynamic traffic flows and in remote unmanned sites. In a satellite communication service, the control station maintains a quality service by remote control, because the satellite is a common relay point for control messages.

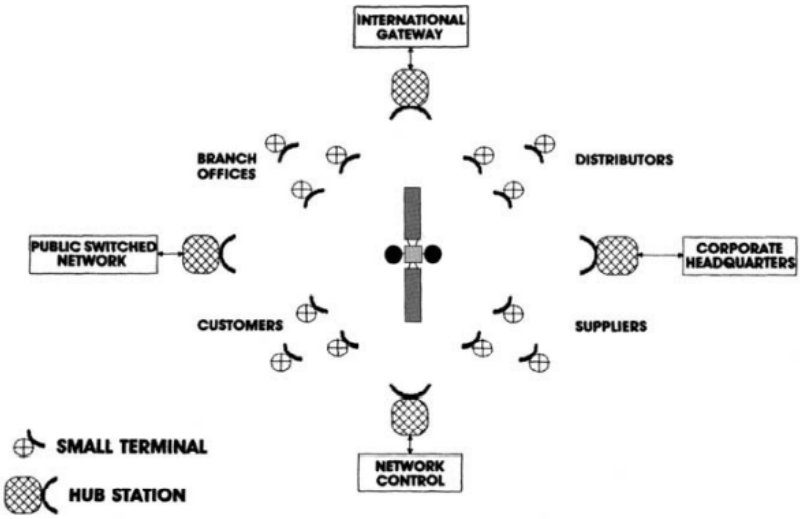

Figure 1.2: Ground segment scheme. [1]

1.2.1 Frequency Allocation

The frequency bands appealing to satellite communications lie above 100 MHz, as frequencies bellow this can be reflected or suffer bending from their path when reaching the ionosphere [7].

As it was refered in 1.1.1, satellite communications initially concentrated in the C-band as it presented the fewest propagation problems. With the consequential overcrowding on the band and the advances in the industry, higher frequencies were employed.

With a limited frequency spectrum, it became clear that it was necessary to ensure the proper use of the frequency bands. The International Telecommunication Union organized various guidelines for frequency allocation for various services and divided the world in three regions for this purpose:

Region	Areas	UHF	Microwave Freq.
1	Europe and Asia	862-870 MHz	2.446-2.454 GHz
2	North and South America	902-928 MHz	2.40-2.4835 GHz
3	Asia and the Pacific	950-956 MHz	2.427-2.470 GHz

Table 1.1: ITU regions and frequency allocation [7].

The ITU also classified the various satellite services, with fixed satellite services (FSS), mobile satellite services (MSS) and broadcasting satellite services (BSS) being the principal communication-related applications of satellites [7].

1.2.2 Payload

The payload is often considered the vital subsystem of any satellite, as it performs the satellite intended function. The basic payload in a communication satellite is the transponder, which acts as a receiver, amplifier and transmitter. A transponder is considered to be a microwave relay channel which also performs the function of frequency translation from the uplink frequency to the relatively lower downlink frequency. So is a combination of elements like antennas for transmit and receive functions, a subsystem of repeaters, frequency shifters and mixers, filters, low noise amplifiers and power amplifiers. A typical communication satellite has various transponders, each one generally operated in a different frequency band divided into slots [7].

The payload of a communication satellite is implemented in one of two general types of configurations; the frequency translation transponder, and the on-board processing transponder.

1.2.2.1 Frequency Translation Transponder

Frequency translation transponders process the uplink satellite signal by only their altering their amplitude and frequency; the modulation and the spectral shape of the signal are not affected [9].

This transponder incorporates and input filter, a low noise amplifier (LNA), down converter, input multiplexer, channel amplifiers, high power amplifiers and output demultiplexer. The uplink segment of the transponder comprises the input filter, the LNA and the down converter (which yields a fixed frequency translation corresponding to the

difference between the centre of the uplink and the downlink frequency bands) [9].

The complete bandwidth is divided into individual channels by the input multiplexer, where each filter is programmed to pass the bandwidth of a particular channel and reject the other parts. The output of each multiplexer filter is amplified by power amplifiers, that can be referred as travelling wave amplifiers. The output of all channels is merged in an output de-multiplexer and then reaches the transmitting antenna for down-beaming the signal on to Earth [9].

1.2.2.2 On-board Processing Transponder

On-board processing transponder (or regenerative transponder), as the name states, are those that perform on-board processing and the received signal is modified before re-transmission. On-bord processing improves the throughput and error performance as it restores the signal quality prior to retransmission. These transmitters use digital techniques like narrowband channel selection and routing, demodulation etc. to process the received signal [9].

Although this type of transponders are able to handle all multiple-access techniques, they cannot improve the transmission link, to reduce link noise or to improve the satellite link performance [9].

1.2.3 Antennae Subsystem

The antenna spacecraft systems transmit and receive signals, making it essential in improving the signal strength to allow amplification, processing and transmission. The performance of an antenna is evaluated according to various parameters, three of which are: antenna gain, antenna beamwidth and antenna sidelobes. The gain delineates the increase in strength seized by concentrating the radiowave energy by the subsystem. This parameter is expressed in decibels (dBi) above and isotropic antenna (radiates uniformly in all directions). The beamwidth measures the angle over which maximum gain occurs. The antenna sidelobes describes the gain in off-axis directions [7] [9].

A satellite communication system often employs linear dipole antennas, horn antennas, the parabolic reflector and the array antenna [9].

The linear dipole antenna radiates uniformly in all directions and are used mainly at VHF and UHF frequency bands for tracking, telemetry and command links. This type of antennas also present high relevance during launch operations and for satellites that operate with no attitude control or body stabilization [9].

Horn antennas are used when relatively wide beams are necessary, with frequencies above 4 GHz (i.e. a geosynchronous satellite). A horn is a flared waveguide that can offer gains that go up to 20 dBi and 10º (or higher) of beamwidth [9].

When required higher gains or smaller bandwidths, a reflector or array antenna is preferred. The parabolic reflector antenna is often used in satellite systems and usually operates with frequencies above 10 GHz. These antenna systems are generally beamed by horn antennas feeds at its focus. Parabolic reflectors reach gains higher than 25 dBi and 1º of beamwidths while working with C, Ku and Ka bands [9]. The C-band ranges from 4 to 8 GHz and is primarily used for satellite communications, for full-time satellite TV networks or raw satellite feeds. The Ka-band ranges from 26 to 40 GHz, where the uplink in satellite communications is in either the 27.5 GHz and 31 GHz bands, and targeting radars on military aircraft. And the Ku-band ranges from 12 to 18 GHz and is used for satellite communications as well. In Europe, the downlink is used from 10.7 GHz to 12.75 GHz for direct broadcast satellite services

An array antenna is a steerable, focused beam than can be formed by a combination of the radiation from several elements with dipoles, helices or horns. Beam forming is also reached by electronically phase shifting the signal of each element. Array antennas are of high interest in communication systems because the correct selection of the phase characteristics between the elements allows the control of direction and beamwidth without an actual physical movement in the antenna system [9].

1.2.4 Digital Modulation and Multiple Access

Digital modulation involves transferring the bit stream to an RF carrier. The process of digital modulation design in a satellite link includes generating a bit in the form of a pulse that meets certain criteria for bandwidth and shaping; transferring those shaped pulses onto a carrier; allowing for noise, and interference along the link; and providing a fitting demodulator capable of recovering the pulses from the carrier and correcting them up to reproduce a digital bit stream with minimum errors [1].

Digital modulation works by the basic characteristics of a pulse. As shown in Fig. 1.3,the pulse as a fixed duration of T seconds, which implies that pulses on a continuous stream would have a rate of $1/T$ bits per second. The bandwidth of the main lobe is proportional to the bit rate, 1/T, and the spectral shape of pulse amplitude given by,

$$\sin(\pi FT)/(\pi FT) \quad (1.1)$$

This means that a fixed pulse length produces maximum energy at zero, because for small angles $\sin(x) = x$ [1].

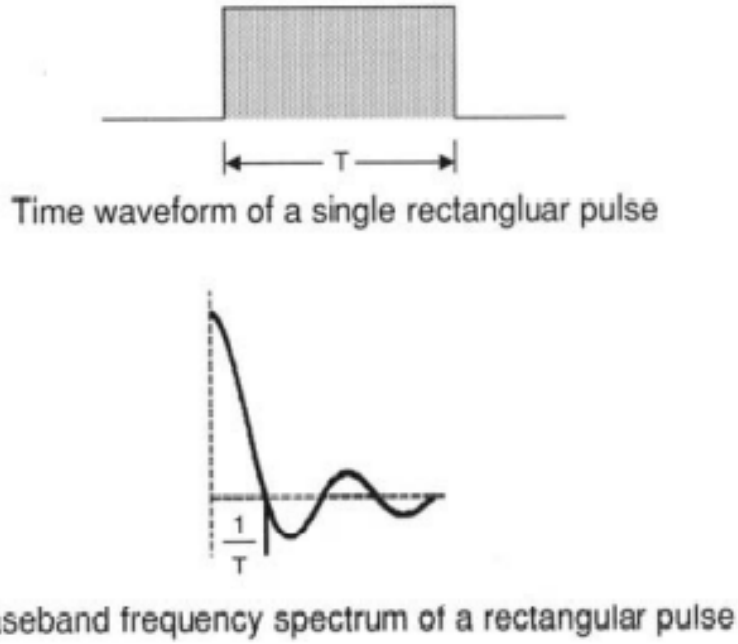

Figure 1.3: Characteristic of a pulse in time and frequency domain. [1]

Due to the bandwidth occupied by sidebands, the spectrum should be filtered to minimize the taken bandwidth. After filtering, the output spectrum is enclosed withing the main lobe's bandwidth. A stream of filtered pulses carries residual tails, which is referred as intersymbol interference (ISI) and increases the chance of errors. The effects of ISI errors can be lessened by aligning the tails so their zero crossing occur when the pulse main lobes reaches their peaks. This process of shaping can efficiently reduce the modulated bandwidth of a digital carrier and required RF bandwidth of a filtered pulse is expressed by $(1 + \alpha)$ times the bit rate (given in Hertz), where the roll-off factor, α, ranges from 0.2 to 0.4 [1].

1.2.4.1 Frequency Shift Keying, FSK

The first form of digital modulation used in communications was FSK. This modulation works by switching between two frequencies, which correlates to the '1' and '0' states. This causes FSK to consume twice the bandwidth of a phase shift keying (PSK) modulation, but the simplicity of FSK makes it usable in the modulator and the demodulator. Although, this plainness also means that FSK needs more signal power for the same performance when considering error rate [1].

1.2.4.2 Phase Shift Keying, PSK

PSK is the most common system to transmit and receive high-speed digital information when working with satellites. These technique modems perform at rates between 64 kbps and 1000 Mbps. Binary phase shift keying (BPSK) applies two phase states: 0^{o} and 180^{o}; and quadrature phase shift keying (QPSK) applies four states split into pairs: 0^{o} and 180^{o}, and 90^{o} and 270^{o}. QPSK transmits two bit per symbol and is efficient in bandwidth and power usage [1].

1.2.4.3 Amplitude and Phase Shift Keying, APSK

When in QPSK the occupied bandwidth exceeds what is available, it is necessary to rely on bandwidth efficient modulation (BEM). The most common BEM technique is amplitude and phase shift keying (APSK), in which the phase and the amplitude are assorted to carry the information bits. (for example, a 16APSK uses a shift between two amplitude states and up to 12 phase states) [1].

The phase-amplitude constellation has its main advantage in adaptive coding and modulation (ACM)- the modulation format and coding rate are altered automatically on the link to a more resistant combination at a lower data rate when the link goes through a high channel error rate. Because of this, it is necessary a higher carrier power vs. noise ratio (C/N) than those acceptable for PSK or FSK [1].

1.2.5 Multiple Access Methods

The last access allows various Earth Stations to transmit to the same satellite without

interference. RF carriers can be separated in frequency (FDMA), time (TDMA) or code (CDMA) [1].

1.2.5.1 Frequency Division Multiple Access (FDMA)

In FDMA, Earth stations must transmit their signals on different microwave frequencies and the modulation can't cause the carrier bandwidths to overlap, which means that various transmissions can be done independently at the same time. A limitation in FDMA is that the sum of the bandwidths of the carriers can´t exceed the satellite's available bandwidth [1].

Each Earth Station or terminal is assigned a different transmission frequency, which can be permanently fixed on changed by a manual command.

1.2.5.2 Time Division Multiple Access

In this method, Earth stations transmissions are on the same frequency, and each uses the complete bandwidth of the channel. To prevent interference caused by all the transmissions is necessary to synchronize the signals so they don't overlap in time. The signals are transmitted in burst of information, and Earth stations that rely on this method must have a system of timing and control of this transmissions. The preferable form of digital modulation for TDMA, is QPSK, considering this modulation allows to meet the compression and timing requirements [1].

There are disadvantages to a full transponder TDMA. The first is that is necessary to exist some form of compensation in the power system of the satellite to prevent that the pulses of direct current produced by the bursts don't affect the operation of the system. Another problem is the possibility of generation of sidebands in the transponder, which can cause interference in the transmission. Although, a proper modulation and repeater structure respond effectively to these concerns [1].

1.2.5.3 Code Division Multiple Access

CDMA is a method that guarantees better bandwidth usage and service quality when dealing with spectral congestion and interference. The basis of this system is to filter

different signals from different users, by the code of each transmission. Direct sequence spread spectrum works on a high-speed sequence of bits to spread the information. In DSSS, the code is a lengthy sequence of bits and it runs as many times the bit rate of the original information, which expands the bandwidth by the ratio of bit rates [1].

DSSS employs a computer-generated sequence of bits, called pseudo-random noise (PN) binary sequence, to assure incoherence with any delayed version of the transmission itself or any other PN sequence applicated in the CDMA network. Each signal contemplatesthe original message, which is multiplied by the PN sequence and then modulated using either BPSK or QSPK. Simultaneously, the bandwidth is multiplied, and as a result of this, the power spectral density is reduced. Multiple spread spectrum signals can be transmitted at the same time as long as the PN sequence are different or not synchronized [1].

1.3 Thesis' Objective

This study aims at reconstructuring the trajectory of a communication satellite by two algorithms: the Discrete and the Robust Kalman filters. This is done in order to compare the performances of the algorithms, as well as counteract the most evident limitations of the classical form of the filter.

The chosen satellite for the estimation is one of the back-ups satellites for the Iridium constellation and its data was analyzed through NASA's GMAT program.

For this dissertation, the filter's application is run using $Python^{TM}$.

1.4 Outline of the Dissertation

The current chapter gives a brief overview of everything related to communication satellites as well as its evolution.

Chapter 2 explains the dynamic model used to describe a space vehicle movement, the perturbations on a satellite's orbit and characterizes satellite communication channels.

Chapter 3 studies the Discrete and the Robust Kalman filters as well as relevant concepts to its comprehension.

Chapter 4 demonstrates the simulations performed for this study, the obtained results and compares the application of the estimation models.

Chapter 5 concludes the dissertation and suggests futur work possibilities.

Chapter 2. Model Description

2.1 Orbital Dynamic Model

Through Newton's laws, it is possible to derive the equations of motion of a spacecraft moving under the influence of a central body force field. The equation of motion of a Keplerian orbit can be given as [10]:

$$\frac{d\vec{r}^2}{dt^2} = -\frac{\mu\vec{r}}{r^3} \quad (2.1)$$

Where $\vec{r}$ is the position vector of the space body with its origin at the center of mass of the primary body and r is the norm of $\vec{r}$. The gravitational constant, μ , is defined by:

$$\mu = k^2(m_1 + m_2) \quad (2.2)$$

Where *k* represents the gravitational constant, m_1 is the mass of the primary body (in this case, the Earth) and m_2 is the mass of the second body/satellite [10].

As shown in Fig.2.1, *XYZ* is a rigid inertial frame of reference and $i_r i_\theta i_\varphi$ is rigid moving frame. The moving frame can move (translation and rotation) freely on its own accord or, in this case, it can be attached to a spacecraft. With $\vec{i_r}$, $\vec{i_\theta}$ and $\vec{i_\phi}$ being the unit vectors of the moving system. Kinematic quantities measured relative to the inertial frame are called absolute and those measured relative to the moving frame are called relative. The motion of the moving frame is arbitrary. However, because the moving frame is rigidly attached to a satellite, so that it translates and rotates with the object, then the frame is considered a body frame. Thus, a body frame has the same angular velocity as the body

to which it is bound [11].

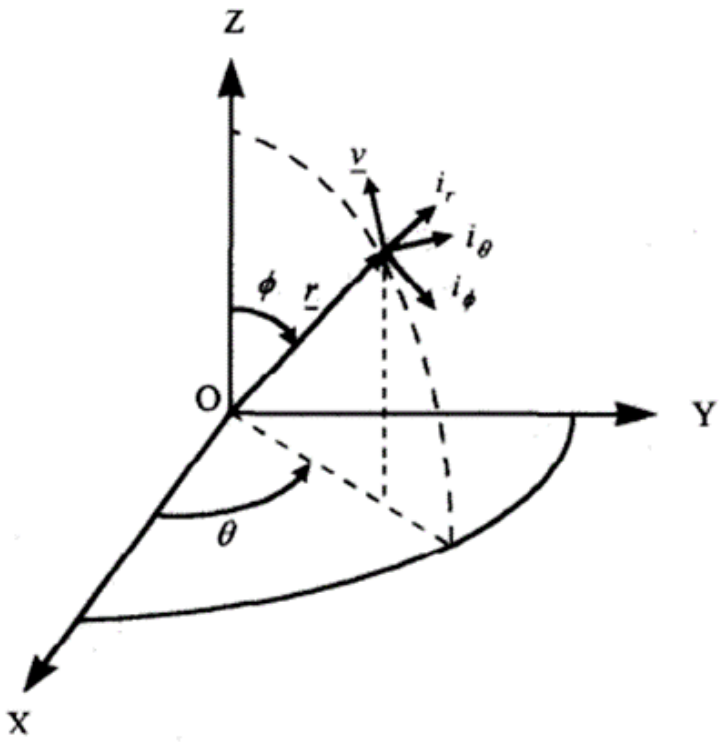

Figure 2.1: Motion of the spacecraft orbiting Earth in a spherical coordinate system. [2]

The equations of motion are extrapolated through the acceleration equations of the spacecraft orbiting the Earth in an inertial frame. Its scalar relations, in spherical coordinates, Fig. 2.1, are [2]:

$$\begin{cases} \ddot{r} - r\dot{\theta}^2 = -\frac{\mu}{r^2} \\ 2\dot{r}\dot{\theta} + r\ddot{\theta} = 0 \end{cases} \tag{2.3}$$

The unit vector from the central body to the satellite's position, $\vec{r}$, the orbital angular velocity, ω, and its orbital angular acceleration, $\dot{\omega}$, are, respectively [2]:

$$\vec{r} = \begin{bmatrix} r\sin\phi\cos\theta \\ r\sin\phi\sin\theta \\ r\cos\phi \end{bmatrix} \tag{2.4}$$

$$\omega = \begin{bmatrix} \dot{\phi}\sin\theta \\ -\dot{\phi}\cos\theta \\ \dot{\theta} \end{bmatrix} \quad (2.5)$$

$$\dot{\omega} = \begin{bmatrix} \ddot{\phi}\sin\theta + \dot{\theta}\dot{\phi}\cos\theta \\ -\ddot{\phi}\cos\theta + \dot{\theta}\dot{\phi}\sin\theta \\ \ddot{\theta} \end{bmatrix} \quad (2.6)$$

Substituting equations 2.4, 2.5 and 2.6 into 2.3 and solving for $\ddot{r}$, $\ddot{\theta}$ and $\ddot{\phi}$, gives us the equations of motion for a space vehicle revolving around the Earth :

$$\begin{cases} \ddot{r} = r\dot{\theta}^2(\sin\phi)^2 + r\dot{\phi}^2 - \frac{\mu}{r^2} + u_r \\ r\ddot{\theta}\sin\phi = -2\dot{r}\dot{\theta} - 2r\dot{\theta}\dot{\phi}\cos\phi \\ r\ddot{\phi} = -2\dot{r}\dot{\phi} + r\dot{\theta}^2\sin\phi\cos\phi \end{cases} \quad (2.7)$$

Where μ is the earth's gravitational constant, r is the radial distance of the spacecraft from the center of the Earth, θ is the angle measured from the X-axis in the XY-plane, and ϕ is the angle measured from Z-axis to the vector r; u_r is the thrust acceleration components relative to the direction. The satellite is modelled as a point mass [12].

As it was referred above, a space vehicle in an Earth orbit experiences perturbations due to atmospheric drag, solar radiation pressure, and Earth's oblateness (which creates different gravitational potentials). It is imperative to solve the problem of the effects of the second zonal harmonic, J_2, on an orbit. The perturbing acceleration due to the Earth's gravitational potentials are sufficient to determine the second zonal harmonic's major accelerations [12]:

$$V = \frac{\mu}{r}\{1 - J_2(\frac{a_e}{r})^2 P_2(\cos\phi)\} \quad (2.8)$$

$$a = \nabla V = \frac{\partial V}{\partial r}\vec{i_r} + \frac{1}{r\sin\phi}\frac{\partial V}{\partial \theta}\vec{i_\theta} + \frac{1}{r}\frac{\partial V}{\partial \phi}\vec{i_\phi} \quad (2.9)$$

Where μ is $3.986 \cdot 10^{14}\ m^3 \cdot s^{-2}$, J_2 is $1.082 \cdot 10^{-3}$, a_e, the equatorial radius of the Earth, is $6.378 \cdot 10^3\ m$ and $P_n(\cos\phi)$.

The equations of motion in 2.7 can be rewritten, in order to consider these perturbing effects as presented below, using equations 2.8 and 2.9 [12]:

$$\begin{cases} \ddot{r} = r\dot{\theta}^2(\sin\phi)^2 + r\dot{\phi}^2 - \frac{\mu}{r^2} + \frac{3}{2}\mu J_2 a_e^2 \frac{3(\cos\phi)^2 - 1}{r^4} + u_r \\ \ddot{\theta} = \frac{-2\dot{r}\dot{\theta}}{r} - 2\dot{\theta}\dot{\phi}\cot\phi \\ \ddot{\phi} = \frac{-2\dot{r}\dot{\phi}}{r} + \dot{\theta}^2\sin\phi\cos\phi + 3\mu J_2 \frac{a_e^2}{r^5}\cos\phi\sin\phi \end{cases} \quad (2.10)$$

2.2 Orbital Perturbations

The orbit in which the satellite operates is called the mission orbit. A Keplerian orbit is one in which the central body and satellite are the only two objects in the system, the only force applied is gravity, the central body is spherically symmetric and its mass is much greater than that of the satellite. Because the Keplerian orbit does not take into account non-gravitational forces or the interaction of third bodies, the perturbations caused by these elements change the calculated orbit into a real orbit. With that said, Keplerian orbits provide good approximation for spacecraft motion [13].

Effects that modify Keplerian orbits are divided into four categories: non-gravitational forces, third body interactions, non-spherical mass distributions and relativistic mechanics. For Earth's orbits, the perturbations are created by atmospheric drag (in LEO orbits), Earth's oblateness and Lunar- Solar effects at geosynchronous altitudes. In order to perform its mission correctly, these perturbations need to be corrected [13].

2.2.1 Atmospheric Drag

When the orbit's altitude is below 1500 km, atmospheric drag is the principal non-gravitational force acting on the satellite. The force that air molecules apply on the spacecraft is opposite to the direction of motion (velocity vector) and it can gradually decrease the mean orbit and speeds up the satellite. We can calculate drag's acceleration, a_D, by:

$$a_D = -(1/2)\rho V^2 (C_D A/m) i_v \quad (2.11)$$

Where ρ is the atmospheric density, V is the spacecraft´s velocity relative to the atmosphere, C_D is the drag coefficient, A is the cross-sectional area, m is the mass of the satellite and i_v is the unit vector of the satellite's velocity relative to the atmosphere [10].

The atmospheric density is computed via density models like Jacchia 71 and MSIS90. The second one is considered to be slightly more accurate model, and calculates the density as a function of the spacecraft's altitude, solar flux, Earth magnetic index, time of day, and the satellite's geocentric longitude/latitude [10].

To make-up for the drag effects, the spacecraft can perform a Hohmann transfer in order to return to the orbit's altitude when the semi-major axis reaches its lower limit [10]. The Hohmann transfer, represented in 2.2, applies an elliptical orbit to transfer the spacecraft from orbit (1) to (2). The propulsion system of the aircraft modifies the magnitude and direction of its velocity vector, and two burns are realized to perform the transfer. The first, is used to transfer the satellite from the initial orbit to the elliptical transfer orbit, where the perigee is the same as the radius of the initial orbit. The second burn is used to transfer the elliptical transfer orbit into the final circular orbit where the apogee becomes the radius of the final higher orbit.

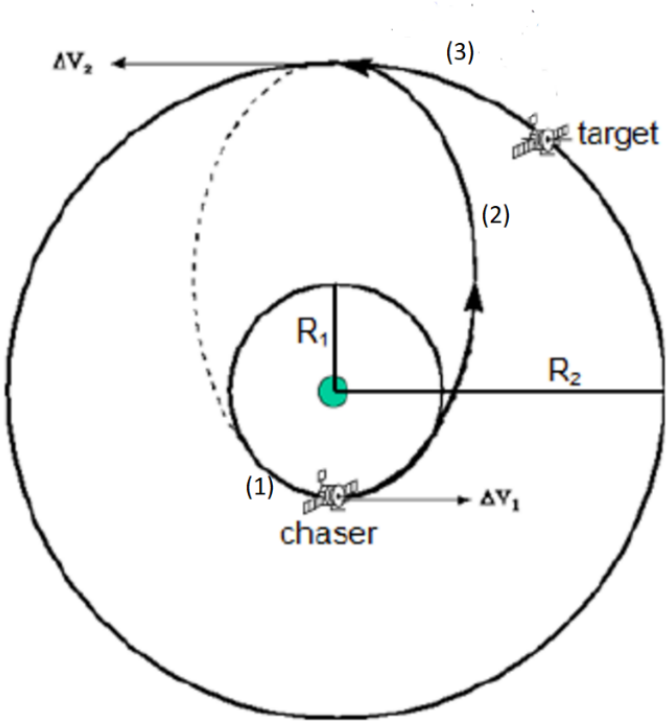

Figure 2.2: Hohmann Transfer Orbit. [3]

2.2.2 Earth's Oblateness

Although the Earth is similar to a sphere, the centrifugal force originated by the planet's rotation turns it into an oblate spheroid. This combined with peripheral mass anomalies like continents (2.3) implies that the planet's center of mass is different than its geometric center [13]. Consequently, the gravity force applied on an orbiting body is not directed towards it.

Oblateness causes a variation, called zonal variation, with the angular distance from the equator (latitude). The second zonal harmonic, J_2, is a dimensionless parameter that quantifies the major effects of a planet's oblateness on an orbit. For Earth, $J_2 \approx 1,082 \cdot 10^{-3}$ [10]. This shape causes the right ascending node (Ω) and the argument of the perigee (ω) to vary significantly in time as follows:

$$\dot{\Omega} = -\cos i \left[\frac{3}{2} \frac{\sqrt{\mu} J_2 R^2}{(1-e^2)^2 a^{7/2}} \right] \quad (2.12)$$

$$\dot{\omega} = -\left[\frac{3}{2}\frac{\sqrt{\mu}J_2R^2}{(1-e^2)^2a^{7/2}}\right]\left(\frac{5}{2}(\sin i)^2 - 2\right) \quad (2.13)$$

R is the radius of the planet; μ is the gravitational parameter; a is the semimajor axis; exp is the eccentricity of the orbit, and i is the orbit's inclination, $\dot{\Omega}$ and $\dot{\omega}$ are in deg/day [11].

This perturbation results in major impacts in the initial orbit due to the rotation of several degrees of the right ascension of the ascending node and the argument of the perigee per day [14].

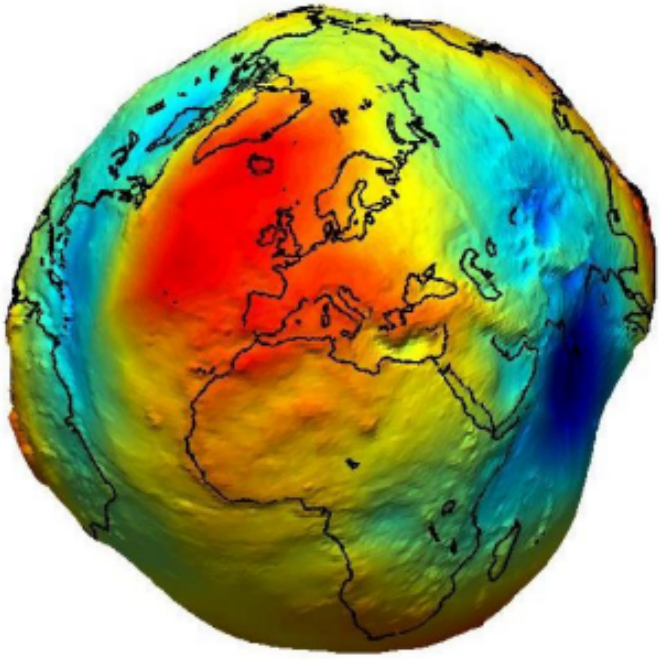

Figure 2.3: Earth's computer generated model. [4]

2.2.3 Third Body Interactions

The gravitational forces of the Sun and the Moon leads to periodic variations in all of the orbit elements. The Sun and the Moon apply an external torque to the orbits and cause the angular momentum vector to rotate. This effect is extremely small and in Leo orbits is dominated by the Earth's oblateness. Yet, the effect becomes relevant in geosynchronous

orbit and is the main reason for the need for stationkeeping in GEO [14].

For Keplerian orbits, the Lagrange planetary equations are used to determine the approximate rotation rates for the ascending node, $\dot{\Omega}$, and the argument of perigee, $\dot{\omega}$ due to the Sun and the Moon[14]:

$$\dot{\Omega}_{Moon} = -0.00338(\cos i)/n \quad (2.14)$$

$$\dot{\Omega}_{Sun} = -0.00154(\cos i)/n \quad (2.15)$$

$$\dot{\omega}_{Moon} = 0.00169(4 - 5\sin^2 i/n) \quad (2.16)$$

$$\dot{\omega}_{Sun} = 0.00077(4 - 5\sin^2 i/n) \quad (2.17)$$

where i is the orbit inclination, n is the number of daily orbit revolutions and, as already mentioned in section 2.2.2, $\dot{\Omega}$ and $\dot{\omega}$ are in deg/day. The equations ignore the variations caused by the change in orientation of the orbital plane with respect to the Moon's orbit and the ecliptic plane, being only average values [14].

2.3 Initial Conditions

Equations 2.7 and 2.10 can be reworked using state space notation that converts them from second order differential equations to first order differential equations and doubles the number of equations. This process assigns a new variable, x_n, for each variable from the spherical coordinate system:

$$\begin{cases} x_1 = r \\ x_2 = \dot{r} \\ x_3 = \theta \\ x_4 = \dot{\theta} \\ x_5 = \phi \\ x_6 = \dot{\phi} \end{cases} \tag{2.18}$$

Equations 2.19 are a state space form notation for equation 2.7, and equations 2.20 are the state space form of equations 2.10, as shown below:

$$\begin{cases} \dot{x_1} = x_2 \\ \dot{x_2} = x_1 x_4^2 (\sin x_5)^2 + x_1 x_6^2 - \frac{\mu}{x_1^2} + u_x 1 \\ \dot{x_3} = x_4 \\ \dot{x_4} = -\frac{2x_2 x_4}{x_1} - 2x_4 x_6 \cot x_5 + \frac{u_x 3}{x_1 \sin x_5} \\ \dot{x_5} = x_6 \\ \dot{x_6} = -\frac{2x_2 x_6}{x_1} + x_4^2 \sin x_5 \cos x_5 + \frac{u_x 5}{x_1} \end{cases} \tag{2.19}$$

$$\begin{cases} \dot{x_1} = x_2 \\ \dot{x_2} = x_1 x_4^2 (\sin x_5)^2 + x_1 x_6^2 - \frac{\mu}{x_1^2} + \frac{3}{2}\mu J_2 a_e^2 \frac{3(\cos x_5)^2 - 1}{x_1^4} \\ \dot{x_3} = x_4 \\ \dot{x_4} = -\frac{2x_2 x_4}{x_1} - 2x_4 x_6 \cot x_5 + \frac{u_x 3}{x_1 \sin x_5} \\ \dot{x_5} = x_6 \\ \dot{x_6} = -\frac{2x_2 x_6}{x_1} + x_4^2 \sin x_5 \cos x_5 + 3\mu J_2 \frac{a_e^2}{x_1^5} \cos x_6 \sin x_6 + \frac{u_x 5}{x_1} \end{cases} \tag{2.20}$$

In order to simulate and compare Keplerian orbit and the real orbit (which accounts for perturbations), it is imperative to find the initial conditions of the trajectory

of a satellite. For this, the chosen space vehicle is one of the recently sent backup satellites for the Iridium constellation.

On the 3rd of July of 2023 at 08:14:59 pm (UTC), the orbit data of the Iridium 2023-068W was [8]

Eccentriity, e	0.0002821
Inclination, i	1.5129 rad
Perigee height, h_p	627000 m
Apogee height, h_a	631000 m
Right ascension of the ascending node, Ω	2.1285 rad
Argument of the perigee, ω	1.6930 rad
Mean anomaly, M	4.5929 rad

Table 2.1: Iridium orbit data. [8]

To convert this data into spherical coordinates, NASA's program GMAT (General Mission Analysis Tool) was used. In GMAT, the input parameters are the semi major axis, the eccentricity, the inclination, the right ascension of the ascending node, the argument of the perigee and the true anomaly. Hence, it is necessary to calculate the semi major axis, a, the eccentric anomaly, E, and the true anomaly, υ. These missing parameters are computed through the astrodynamics equations below:

$$a = \frac{2a_e + h_p + h_a}{2} \quad (2.21)$$

$$M = E - e \sin E \quad (2.22)$$

$$\tan \frac{\upsilon}{2} = \sqrt{\frac{1+e}{1-e}} \tan \frac{E}{2} \quad (2.23)$$

The Earth's equatorial radius is 6.378×10^3 m, approximately. The orbital elements are shown below:

$$a = 7.007 \times 10^3 m \quad (2.24)$$

$$E = 4.5926rad \quad (2.25)$$

$$\upsilon = 4.5923rad \quad (2.26)$$

The orbit was simulated in GMAT taking in account the date and time referred above, the atmospheric model MSIS90 (considered to be the most accurate), the effects of J_2, and the Iridium's satllite mass (860 kg).

After the simulation's analysis, the initial conditions in spherical coordinates were found for the orbit of the Iridium 2023-068W:

r	7007236 m
$\dot{r}$	-2.1209 m/s
θ	2.1286 rad
$\dot{\theta}$	6.2258×10^{-5} rad/s
ϕ	1.5687 rad
$\dot{\phi}$	-1.0747×10^{-3} rad/s

Table 2.2: Initial Conditions for the Satellites's Orbit in Spherical Coordinates.

2.4 Characterisation and Channel Modelling

The performance of land mobile satellite (LMS) communication systems depends on numerous factors such as operating frequency, elevation angles, geographic location and climate [5].

The diverse nature of propagation environments has a huge impact on the real-time operation, design and performance assessment of satellite-terrestrial systems operating at radio frequencies ranging between 100 MHz and 100 GHz and optical frequencies. Therefore, a proper knowledge and modelling of the propagation channel is imperative for the performance of these systems, and to establish reliable communication links [5].

This chapter bestows an overview of the propagation impairments on LMS communication links, probability distributions describing these fading effects and channel models developed considering these probability distributions [5].

2.4.1 Propagation Impairments Impacting Satellite Communication Links

The use of satellite communication systems entails propagation environments for radio signals distinct from those in terrestrial radio systems. The radio waves propagating between an Earth station and satellites experience impairments including the effects of the ionosphere, the effects of the troposphere and local fading effects. The combination of these impairments on a satellite-Earth link can cause scattered fluctuations in amplitude, phase, angles of arrival, de-polarization of electromagnetic waves and shadowing which prompts the degradation of the signal quality and increase in the error rates of the communication links [5].

a) Ionospheric Effects

The effects of the ionosphere (an ionized section in the atmosphere extending from a height of 30 km to 1000 km) cause various impairments phenomena namely scintillation, polarization rotation, refraction, group delays and dispersion on the radio signals, with scintillation and polarization effects being the foremost concerns for satellite communications [5].

Ionospheric scintillations are deviations in the amplitude level, phase and angle of arrival of the received waves. The main effect of scintillation is fading that highly depends on the inhomogeneities of the ionosphere (rapid variation of local electron density). The effects of scintillation are significant at higher altitudes and at $\pm$ 20° around the earth's magnetic equator. These effects decrease with the increase in operating frequency and it has been observed that at the operating frequency of 4 GHz ionospheric scintillations can result in fades of several decibels (dBs) and duration between 1 to 10 seconds [5] [15].

The orthogonal polarization increases the spectral efficiency without increasing the bandwidth requirements. This technique has some limitations due to the depolarization of electromagnetic waves propagating through the atmosphere. The free electrons present in the ionosphere (due to ionization) interact with these waves under the influence of the Earth's magnetic field which results in the rotation of the plane of polarization of electromagnetic waves (known as Faraday rotation). The polarization rotation is significant for small percentages of time at frequencies 1 GHz or less and is much lower at higher fre-

quencies even in the regions of strong ionospheric impairments and low elevation angles [5].

b) Tropospheric Effects

The tropospheric is the lower portion of the Earth's atmosphere covering altitudes from the ground surface up to a height of about 15 km of the atmosphere. The impairments of this region include hydrometeors (rain, snow and clouds), fog and moisture in the atmosphere, gradient of temperature and sporadic structures of wind streams. The effects caused by these impairments on radio signals are rain attenuation, scintillation, depolarization, absorption, refraction, etc. The radio waves are degraded to varying degrees and depend on the geographic location, frequency and elevation angle. The tropospheric effects become significant when the operating frequency of the LMS communication link is greater than 1 GHz [15].

One of the major causes of attenuation for links operating at frequency bands higher than 10 GHz is rain. The attenuation in the received signal amplitude is due to absorption and scattering of the radio waves energy by raindrops and is measured as a function of rainfall rate and increases with the increase in the operating frequency, rainfall rate and low elevation angles [5].

The polarization of radio waves can be altered by raindrops and ice particles in the transmission path so that the power is transferred from the desired component to the undesired path, which leads to interference between two orthogonally polarized channels. When a linearly polarized wave goes through raindrops of non-spherical structure, the vertical component of the wave experiences less attenuation than that of the horizontal component. The differences in the amount of attenuation and phase shift endured by each of the wave components causes depolarization of the waves. Rain and ice depolarization have considerable impacts on LMS communication links for frequency bands above 12 GHz [5].

A radio wave will also experience a reduction in the received signal's amplitude level due to attenuation by atmospheric gases. The fading due to gases is related to the altitude above sea level, temperature, pressure, water vapour concentration and frequency. Molecular absorption is the main cause for signal attenuation due to gases and it occurs

due to the transformation of radio wave energy to thermal energy at some particular resonant frequency of the particles. The attenuation due to atmospheric gases is neglected at frequency bands below 10 GHz [5].

Scintillations in the troposphere occur due to heterogeneities in the refractive index of the atmosphere and can be severe to low elevation angles and frequency bands above 10 GHz [5].

c) Local Effects

In addition to the effects caused by the atmosphere, radio waves suffer from energy loss due to the varying propagation environments on the terrain (buildings, trees, vegetation, etc.). These obstructions cause different multipath effects: diffraction due to bending of the wave around buildings' edges, specular reflection of the signal from objects with dimensions greater than the wavelength, scattering by interaction of the waves with objects of uneven shapes, absorption through foliage, etc. The movement of a mobile Earth station results in rapid changes in the signal strength due to changes in phases. All these phenomena creates loss of the signal energy and degrade the performance of satellite-Earth communication link [5].

2.4.2 Probability Distribution Functions for Different Forms of Fading

There is an increase in complexity and uncertainty in characterization of transmission impairments on the LMS communication links due to the random and unpredictable nature of propagation environments as well as the decrease in the communication system reliability with the increase in operating frequency and at low elevation angles [5]. Hence, in order to assess the performance of these systems over fading channels, it is best to describe these in a stochastic method.

In general, signal fading is decomposed as large scale path loss, a medium slowly varying component that follows a lognormal distribution and small scale fading that follows Rayleigh or Rician distributions depending on the existence of the line of sight (LOS) path between the transmitter and the receiver. This section presents a concise overview of probability distribution functions used to model different fading effects on the LMS

communication links [5].

a) Rayleigh Distribution

In urban environments the transmitted signal arrives at the receiver through different multipath propagation mechanisms. The final signal at the receiver is the summation of diffuse multipath components defined by time-varying attenuations, different delays and phase shifts. With the increase of the number of paths, the sum approaches to complex Gaussian random variable having independent real and imaginary parts with zero mean and equal variance. The amplitude of the composite signal follows Rayleigh distribution and the phases of each component are uniformly distributed between 0 and 2π. The receiver signal, the real part, is written as [5]:

$$R_Ray = \sum_{i=1}^{n} a_i(t)\cos(\omega_c t + \theta_i(t)) \quad i = 0, 1, \ldots, n \quad (2.27)$$

where $a_i(t)$ is the amplitude, $\theta_i(t)$ is the phase of the i$^{\text{th}}$ multipath component and ω_c is the angular frequency of the carrier. The propability density function is expressed as:

$$P_Ray(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \quad r \geq 0 \quad (2.28)$$

where σ represents the standard deviation, and r is the envelope of the received signal [5].

b) Rician Distribution

When a LOS component is present between the transmitter and the receiver, the signal that reaches the receiver is expressed as the sum of one dominant vector and large number of independently fading uncorrelated multipath components with the amplitudes and phases uniformly distributed in the interval [0,2π]. The received signal is described by Rician/Rice distribution and is expressed as follows [5]:

$$R_Rice = C + \sum_{i=1}^{n} a_i(t)\cos(\omega_c t + \theta_i(t)) \quad i = 0, 1, \ldots, n \quad (2.29)$$

where C is a constant that defines the magnitude of the LOS signal between the transmitter and the receiver. The other parameter is the same as described in 2.4.2 (Rayleigh distribution). The pdf of the envelope of the received signal is characterized as follows [5]:

$$P_Rice(r) = \frac{r}{\sigma^2} e^{\frac{-(r^2+C^2)}{2\sigma^2}} I_0(rC/\sigma^2) \quad (2.30)$$

where I_0 is the modified Bessel function of first kind and zero order, and rC/σ^2 is the mean power of the LOS component. If there is no direct path (LOS) between the transmitter and the receiver ($C = 0$) the above equation reduces to a Rayleigh distribution. The ratio of the average specular power (LOS path) to the average fading power (multipath) over specular paths is designated as Rician factor, $a^2/2\sigma^2$, and is expressed in dBs [5].

c) Log-Normal Distribution

Signal power loss due to hindrance of objects of large dimensions and due to vegetation and foliage are weighty factors that cause scattering and absorption of radio waves. Subsequently, the power of the received signal varies in terms of the mean power predicted by the path loss. This variation is known as shadowing and is formulated as lognormally distributed. Shadowing creates holes in coverage areas and generates poor coverage and poor carrier-to-interference ratio (CIR) in different places. The received signal's envelope affected by shadowing presents a pdf that follows a lognormal distribution, which mathematical expression is [5]:

$$P_{\text{log normal}}(r) = \frac{1}{\sqrt{2\pi}\sigma r} e^{-\frac{1}{2}(\frac{(\ln r - \mu)^2}{2\sigma^2})} \quad (2.31)$$

where μ and σ are the mean and standard deviation of the shadowed component of the signal at the receiver, respectively.

d) Nakagami Distribution

As referred in the previous sections, the random fluctuations in the radio signals are categorized into two types of fading: shadowing and multipath fading. The composite shadow fading (direct path and multiplicative shadowing) can be modelled by lognormal distribution. An alternative to the lognormal distribution is the Nakagami distribution that can produce simple statistical models with great performance. The pdf of the received signal envelope using Nakagami distribution is expressed by,

$$P_r(r) = \frac{2}{\Gamma(m)}\left(\frac{m}{2\sigma^2}\right)^m r^{2m-1} e^{\left(\frac{-mr^2}{2\sigma^2}\right)} \quad (2.32)$$

where $\Gamma(.)$ is the Gamma function, $2\sigma^2 = E(r^2)$ represents the average mean power of the LOS component and $m \geq 1/2$ is the Nakagami-m parameter varying between $1/2$ and ∞. With the increase of m, the number of Gaussian random variables contributions increases and the probability of deep fades in the associated probability distribution function decreases. Small values of m corresponds to urban areas and large values are associated to open areas, intermediate values relate to suburban areas [5].

e) Suzuki Distribution

The Suzuki distribution is expressed as the product of Rayleigh distribution method and the lognormal process. Expressing a Rayleigh distributed random variable as α and another random variable β following lognormal distribution, ξ is a random variable described by the product of these independent variables, $\xi = \alpha\dot{\beta}$. The pdf of ξ is expressed by:

$$P_{\xi}(r) = \frac{r}{\sqrt{2\pi}\sigma_0^2\sigma_\mu}\int_0^\infty \frac{1}{r^3}e^{(-r^2/2z^2\sigma_0^2)}e^{(}-\frac{1}{2}(\frac{\ln r - m_\mu}{\sigma_\mu})^2\, dr \quad r \geq 0 \quad (2.33)$$

The Suzuki distribution can be considered to represent propagation environments when LOS component is absent in the received signal. [5]

2.4.3 Statistical Channel Models for Land-mobile-Satellite Communication Systems

Statistical methods and analysis are usually the most beneficial approaches for the characterization of transmission impairments and modelling of the LMS communication links due to the high complexity and unpredictability of the diverse nature of radio propagation paths. The statistical models for LMS channels are categorized as single state and multi-state models. The single state models are expressed by single statistical distributions and are considered for fixed satellite scenarios in which the channel statistics remain constant over the areas of interest. The muti-state models are used to express mobile conditions where channel statistics vary considerably over large areas for particular time intervals in irregular environments. This section briefly describes the statistical based channel models developed for satellites [5] [15].

a) Single-State Models

Loo-Model : The Loo model is a primitive statistical LMS channel model applicated for rural environments with shadowing due to roadside trees. The shadowing attenuation affecting the LOS signal is expressed by log-normal pdf and the diffuse multipath components are characterized by Rayleigh pdf. The resulting complex signal envelope is the sum of the correlated lognormal and Rayleigh methods. The pdf of the received signal envelope is given by [5]:

$$P(r) = \begin{cases} \frac{1}{r\sqrt{2\pi d_0}} e^{(\frac{(\ln r - \mu)^2}{2d_0})} & , r \gg \sqrt{b_0} \\ \frac{r}{b_0} e^{(-\frac{r^2}{2b_0})} & , r \ll \sqrt{b_0} \end{cases} \quad (2.34)$$

where μ and $\sqrt{d_0}$ are the mean and standard deviation, respectively. b_0 represents the aveerage scattered power due to multipath effects. If attenuation due to shadowing (log-normal distribution) is constant, then the pdf only yields in Rician distribution [5].

Corraza-Vatalaro Model: This model complies a combination of Rician and log-normal distribution that is used to model effects of shadowing on the LOS and diffuse components. The model is applicable for non-geostationary satellite channels and can be applied to different environments by adjusting the model parameters. The pdf of the received signal envelope is expressed as [5]:

$$P_(r)(r) = \int_0^\infty p(r/s p_S(s)\, dS \quad (2.35)$$

where $p(r/s)$ represents conditional pdf following Rician distribution conditioned on shadowing S:

$$p(r/s) = 2(K+1)\frac{r}{s^2} exp[-(K+1)\frac{r^2}{s^2} - K] I_0(2\frac{r}{s}\sqrt{K(K+1)}) \quad (2.36)$$

where K is the Rician factor and I_0 is zero order modified Bessel function of first kind. The pdf of lognormal shadowing S is expressed by:

$$P_S(S) = \frac{1}{\sqrt{2\pi} h\sigma S} exp[-\frac{1}{2}(\frac{\ln S - \mu}{h\sigma})^2] \quad (2.37)$$

where $h = {}^{\ln 10}/_{20}$, μ and $(h\sigma)^2$ are mean and variance of the normal variance associated, respectively. The received signal envelope is expressed as the product of two independent

methods, lognormal and Rician, with cumulative distribution function given by:

$$P_r(r_0) = P(r < r_0) = \int_0^\infty \frac{P_S(S)}{S} P_r(\frac{r}{S}), drdS = \quad (2.38)$$

$$1 - E_S\{Q(\sqrt{2K}, \frac{r_0}{S}\sqrt{2(K+1)})\} \quad (2.39)$$

where $E(.)$ represents the average with respect to S and Q is Marcum Q-function. [5]

Extended-Suzuki Model: An extension of the Suzuki model, for frequency non-selective satellite communication channels, considers that for most of the time a LOS component is presented in the received signal. This process is the product of Rician and lognormal probability distribution functions where inphase and quadrature components of Rician distribution can be mutually correlated and the LOS component is often shifted due to Doppler shift. We expressed the pdf of the extended Suzuki model process as [5]:

$$P_\xi(r) = \int_{-\infty}^{\infty} \frac{1}{|y|} P_\alpha\beta(ry, y), dy \quad (2.40)$$

where $P_{\alpha\beta}(x, y)$ represents the joint pdf of the independent Rice and lognormal methods $\alpha(t)$ and $\beta(t)$, and $x = {}^{r}/_{y}$ where y is the variable of integration. The pdfs of Rice and lognormal distributions can be used in equation 2.40 to obtain the following:

$$\begin{aligned} P_\xi(r) = \frac{r}{\sqrt{2\pi}\psi_0\sigma} \int_0^\infty \frac{1}{y^3} exp\{-[((\tfrac{r}{y})^2 + p^2)/2\psi_0]\} \cdot I_0({}^{rp}/_{y\psi_0}) \\ \cdot exp[-(^{-(\ln y - m)^2}/_{2\sigma^2})], dr \quad r \geq 0 \end{aligned} \quad (2.41)$$

where ψ_0 denotes the mean value of random variable x, m and μ are the mean and standard deviation of random variable y and p denotes de LOS component [5].

Xie-Fang Model: This model deals with the statistical modelling of propagation

characteristics in non-geostationary satellites communication systems. In these types of systems, either the mobile user or the satellite can move during a communication session and as result the received signals may fluctuate from time to time. The quality-of-service (QoS) degrades due to these random fluctuations in the received signal which are caused by various propagation impairments in the communication links[5].

As well as the movement of the receiver or the transmitter, satellite communications operating at low elevation angles and the use of small antennas inserts the probability of path blockage and multipath scattering components. This model characterizes fading as two independent random processes: short term fading and long term fading. Short term fading is modelled by a general form of Rice distribution and the long term fading is characterized by lognormal distribution. It is assumed that the amplitudes and phases of the scattering components that cause small scale fading resultant of superposition are correlated. The total electric field, expressed below, is the sum of multipath signals arriving at the receiver [5]:

$$E_{tot} = Ee^{j\psi} = \sum_{i=1}^{n} A_i e^{j\phi_i} \quad (2.42)$$

where n is the number of paths, A_i and ϕ_i denote the amplitude and phase of the i$^{\text{th}}$ path component, respectively. The pdf of the received signal envelope can be expressed as follows:

$$\begin{aligned} P_r(r) = {} & \frac{r}{\sqrt{S_1 S_2}} exp[-\frac{S_1 r^2 + S_2\alpha^2 + S_1\beta^2}{2S_1 S_2}] \\ & \times \frac{1}{2\pi}\int_0^{2\pi} exp[\frac{2S_2\alpha r\cos\theta + 2S_1\beta_r \sin\theta + (S_1+S_2)r^2(\cos\theta)^2}{2S_1 S_2}], d\theta \end{aligned} \quad (2.43)$$

and the pdf of the received signal power envelope is given by:

$$P_p(W) = \frac{1}{2\sqrt{S_1 S_2}} exp[-\frac{S_2 W + S_1\alpha^2 + S_2\beta^2}{2S_1 S_2}] \\ \frac{1}{2\pi}\int_0^{2\pi} exp[\frac{2S_2\alpha\sqrt{W}\cos\theta + 2S_1\beta\sqrt{W}\sin\theta + (S_1 - S_2)W(\cos\theta)^2}{2S_1 S_2}], d\theta \quad (2.44)$$

where the parameters S_1, S_2, α and β represent the variances and means of the Gaussian distributed real and imaginary parts of the received signal envelope, respectively, and W denotes the power of the received signal [5].

Abdi Model: This model relies on the Nakagami distribution to characterize the amplitude of the shadowed LOS signal and characterizes the multipath component of the total signal envelope through Rayleigh distribution. The model delivers mathematically precise closed-form expressions of the channel first order statistics like the signal envelope pdf, moment generating functions of the instantaneous power as well as the second order channel statistics such as level crossing rates and average fade durations. The pdf of the received signal envelope for the first order statistics of the model are given by [5]:

$$P_r(r) = (\frac{2b_0 m}{2b_0 m + \Omega})^m \cdot \frac{r}{b_0} exp(^{-r^2}/_{2b_0}) \cdot {}_1F_1(m, 1, \frac{\Omega r^2}{2b_0(2b_0 m + \Omega)}) \quad r > 0 \quad (2.45)$$

where $2b_0$ is the average power of the multipath component, Ω is the average power of the LOS component and ${}_1F_1(.)$ is the confluent hypergeometric function. [5]

b) Multi-State Models When considering nonstationary conditions, if the terminals move in a large area of a nonuniform environment, the received signal statistics might change significantly over the observation interval. Propagation characteristics of such environments are better characterized by multi-state models [5].

Markov models have been successfully employed to characterize fading channels. These models are characterized in terms of state probability and state probability transition matrices. Multi-state channel models define each state by a Markov method in terms of one of the single-state models referred on the previous section [5].

Lutz Model: Lutz's model is appropriately applied for the characterization of radio wave propagation in urban and suburban areas and is a statistical model based on two states: good state and bad state. The good state represents LOS condition where the received signal is Rice distributed, with Rice factor K. The bad state models the signal amplitude by following Rayleigh distribution with mean power $S_0 = \sigma^2$ that fluctuates with time. Another relevant parameter is time share of shadowing A. Therefore, the pdf of the received signal power is expressed as follows [5]:

$$p(S) = (1 - A) \cdot p_{Rice}(S) + A \int_0^{\infty} p_{Ray}(S/S_0) p_{LN}(S_0), dS_0 \quad (2.46)$$

The values of parameters A, K, means, variances and the associated probabilities were derived from measured data for various satellite elevations, antennas and environments [5].

Transitions between two states are defined by first order Markov method where transition from one state to the next only depends on the current state. As shown in Fig. 2.4 the probabilities P_{gb} and P_{bg} represent transition from state G (good state) to state B (bad) and vice-versa [5].

The transition probabilities P_{gb} and P_{bg} can be determined in terms of the average distances D_g and D_b, in meters, over which the system remains in the good and bad states, respectively:

$$P_{gb} = \upsilon R/D_g \quad P_{bg} = \upsilon R/D_b \quad (2.47)$$

where υ is the mobile speed, in m/s, R is the transmission data rate in bits per second. Because the sum of these probabilities equals to 1, $P_{gb} + P_{bg} = 1$, the time share of shadowing is obtained by [5]:

$$A = \frac{D_b}{D_b + D_g} \quad (2.48)$$

The parameter A, in this model, is independent of data rate and mobile speed. In different models, the time share of shadowing is obtained according to the available propagation conditions and parameters.

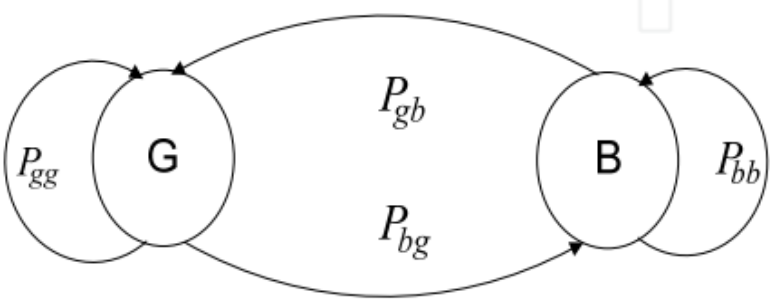

Figure 2.4: Two-state LMS channel statistical model. [5]

Three-State Model: This channel model is based on three states: clear or LOS state, the shadowing state and the blocked state. The clear state is described by Rice distribution, the shadowing state is characterized by Loo's pdf and the blocked state is described by Rayleigh distribution as shown in Fig. 2.5a), where a_1 represents the LOS component, a_2 denotes shadowing effects caused by foliage and a_3 represents blockage. Multipath contributions in the form of reflected waves from the ground are represented by b_1 and incoherently scattered components from the land obstructions are denoted by b_2.[5]. The pdf of the received signal envelope is given by the weighted linear combination of these distributions:

$$P_r(r) = MP_{Rice}(r) + LP_{Loo}(r) + NP_{Rayleigh}n(r) \quad (2.49)$$

where M, L and N are the time share of shadowing of Rician, Loo and Rayleigh distributions, respectively. The state transitions characteristics of the model were obtained using the Markov model, as shown in Fig. 2.5b). The state occurence probability functions P_A, P_B and P_C ($P_A + P_B P_C = 1$) are calculated as follows [5]:

$$P_A = (90 - \theta)^2 / a \quad (2.50)$$

where θ is the elevation angle of satellite ($10 \leq \theta \leq 90$) and 'a' is a constant with values:

$$a = \begin{cases} 7.0 \times 10^3 & \text{for urban areas} \\ 1.66 \times 10^4 & \text{for suburban areas} \end{cases} \quad (2.51)$$

$$P_B = \begin{cases} {}^{P_C}/_4 & \text{for urban areas} \\ 4P_C & \text{for suburban areas} \end{cases} \quad (2.52)$$

In LMS systems one satellite does not give a satisfactory coverage reliability with a high signal quality. Hence, to improve the system availability and signal quality it is desired to employ different satellite constellations: if a link with one satellite is broken by shadowing, another satellite should be available to maintain the connection and help reduce the outage probability. This statistical model provides analysis to improve signal quality and service availability by means of satellite diversity where at least two satellites in LEO/MEO orbit provide coverage area simultaneously in urban and suburban areas [5].

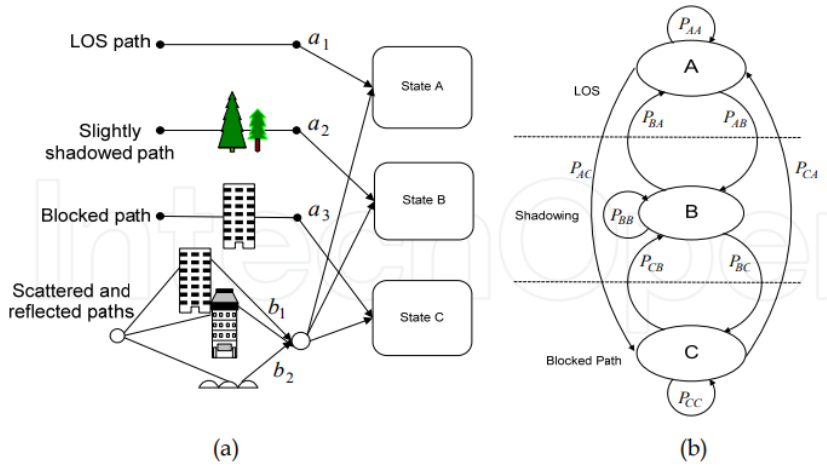

Figure 2.5: Three-state LMS statistical channel model (a). Propagation impairments (b). [5]

Five-State Model: This model is based on a modelling approach where two-state and three-state models are extended to five-state model under different time share of shadowing. Shadowing effects are split into three state: 'good' state denotes low shadowing, 'not good not bad' state describes moderate shadowing and 'bad' state characterizes heavy or complete shadowing. The 'good' state has two sub-states: clear LOS without shadowing and LOS state with low shadowing. Likewise, the 'bad' state has two-sub-states: heavily shadowed areas or blocked (completely shadowed) areas. A state transition might occur when the receiver is in low or high shadowing areas and can only take place from low and high shadowing conditions to moderate shadowing conditions but cannot directly occur between low and high shadowing environments [5].

In terms of the pdf, low shadowing is described by Rician distribution, moderate shadowing is represented by Loo's pdf and high shadowing conditions follow Rayleigh-lognormal distribution. The pdf of the received signal power is a combination of these distributions as shown below:

$$P(s) = X_1 P_{Rice1}(s) + X_2 P_{Rice2}(s) + X_3 P_{Loo}(s) + X_4 P_{Ray1}(s) + X_5 P_{Ray2}(s) \quad (2.53)$$

where X_i $(i = 1, ..., 5)$ are time shares of shadowing of the respective states S_i. The channel model is appropriate for urban and suburban areas [5].

Modelling Frequency Selective LMS Channel: If components of a signal propagating through different paths reach the receiver with delays significantly larger when compared to the bit duration, the signal will experience a significant amount of distortion across the information bandwidth, resulting in frequency selective fading. The response of a wideband channel model (or tapped-delay line model) under uncorrelated scattering can be written as:

$$h(t,\tau) = \sum_{i=1}^{N} a_i(t)\delta[t - \tau_i(t)]exp[j(2\pi f_{d,i} + \theta_i(t))] \quad (2.54)$$

where $a_i(t)$, $\tau_i(t)$, $f_{d,i}(t)$ and $\theta_i(t)$ denote the amplitude, delay, Doppler shift and phase of the i^{th} component of the received signal, respectively, and $\delta(t)$ represents the Dirac delta function.[5]

The parameters for the tapped-delay line model are extracted using extensive measurement data at L-band frequency for different applications and environments. The channel impulse response is defined into three components: the direct path, near echoes and far echoes. The delays τ_i ($i = 1, 2, ..., N$) of the taps are related to the delay of the direct path. The power of all taps is normalized to the power of the direct path and the amplitude distribution of the echoes is described by Rician or Rayleigh distribution, depending on the presence of LOS or non-LOS, respectively. The number of near echoes, N^n, in the location of the receiver is described by Poisson distribution with parameter λ and the corresponding delays ($i = 1, 2, ..., N$) defining near echoes follow exponential distribution with parameter b. The far echoes $N^f = N - N^n - 1$ are characterized by Poisson distribution. The amplitude distributions of the far echoes are described by Rayleigh distribution. [5]

Chapter 3. Kalman Filtering

The following chapter provides a complete literature review of the Kalman filter.

3.1 Stochastic System

In order to estimate an orbit, we consider the satellite as system, namely a stochastic system. A stochastic system has a random nature, one or more parts of the system has randomness associated with it. Unlike a deterministic system, a stochastic system does not always produce the same output for a given input. Due to the arbitrary aspect, the satellite is analysed based on probability theory [16].

A few elements of systems that can be stochastic in nature encompasses stochastic input, random time-delays, random noise disturbances, as well as stochastic dynamic processes [16].

Considering a physical system, it is imperative to develop a mathematical model that adequately represents some aspects of the system's behaviour. For this purpose, it is necessary to establish interrelationships among variables of interest, inputs and outputs to and from the system, through physical insights, fundamental laws and empirical testing [17]. Thus making it possible to analyse the system's structure and modes of response.

In order to monitor the system's behaviour, measurement instruments are assembled to output data signals proportional to certain variables of interest [17].

Deterministic systems and control theories present some limitations when it comes to correctly perform the system's analysis. Firstly, mathematical models are imperfect, because they only depict parameters of interest. The objective is to represent the critical modes of response, knowingly leaving many effects unmodelled [17].

Another defect of deterministic models is that dynamic systems, like the space seg-

ment, are influenced not only by our own control inputs but also by perturbations that are impossible to control or model deterministically. Besides that, it is essential to account for sensor error, because these are also subjected to its own dynamic, disturbances and noise [17].

To counter these problems, stochastic models essentially aim to include all uncertainties in a direct and practical manner; optimally estimate the quantities of interest, even if data obtained from sensors are incomplete and noise-corrupted; optimize the control of the system to perform correctly even with uncertain system descriptions, disturbances and incomplete data; and lastly, evaluate the performance capabilities of the estimation and control systems before and after they are implemented [17].

3.2 Probability and Statistics Concepts

3.2.1 Normal or Gaussian Distribution

Given a continuous random process $X \sim \mathcal{N}(\mu, \sigma^2)$ that is normally distributed with mean μ and variance σ^2 (standard deviation σ), the probability function for X is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} \quad (3.1)$$

for $-\infty < x < +\infty$. and is shown graphically in Fig. 3.1 [6].

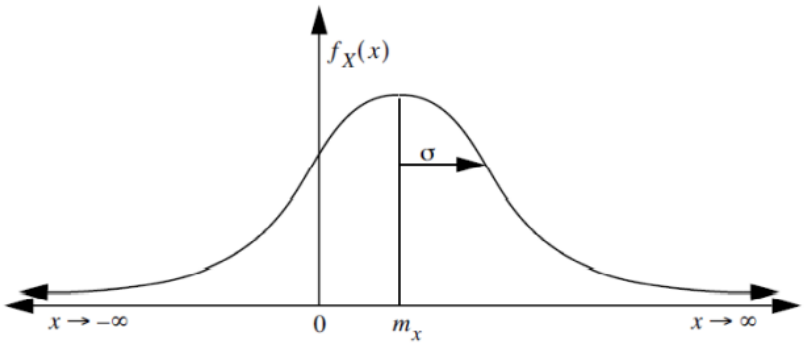

Figure 3.1: The Normal or Gaussian probability distribution function. [6]

Any linear function of a normally distributed random variable is also a normally dis-

tributed random variable. Particularly, if $X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y = aX + b$, then

$$X \sim \mathcal{N}(a\mu + b, a^2\sigma^2) \quad (3.2)$$

The probability density function for Y is described by [6]:

$$f_Y(y) = \frac{1}{\sqrt{2\pi a^2\sigma^2}} e^{-\frac{1}{2}\frac{(y-(a\mu+b)^2}{a^2\sigma^2}} \quad (3.3)$$

If $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are independent, then

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \quad (3.4)$$

and the density function is given by

$$f_X(x_1 + x_2) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{1}{2}\frac{(x-(\mu_1^2+\mu_2^2)}{\sigma_1^2+\sigma_2^2}} \quad (3.5)$$

3.2.2 Continuous Independence and Conditional Probability

Two continuous random variables X and Y are considered to be statiscally independent when [6]:

$$f_{XY}(x, y) = f_X(x) f_Y(y) \quad (3.6)$$

The probability density of the random variable Y in the presence of random variable Y, is called conditional probability is given by [6]:

$$f_{X|Y}(x,y) = \frac{f_{X|Y}(y)f_X(y)}{f_Y(y)} \quad (3.7)$$

Continuous-Discrete: Given a discrete variable X and a continuous variable Y, the discrete probability mass function for X conditioned on $Y = y$ is determined by [6]:

$$p_X(x|Y=y) = \frac{f_Y(y|X=x)p_X(x)}{\sum_z f_Y(y|X=z)p_X(z)} \quad (3.8)$$

3.2.3 Spartial-Spectral Signal Characteristics

Autocorrelation is a time-related characteristic of random signal (is correlation with itself over time). The autocorrelation of a random signal $X(t)$ for times t_1 and t_2 is defined by [6]:

$$R_X(t_1,t_2) = E[X(t_1)X(t_2)] \quad (3.9)$$

If the process is stationary (density does not vary in time) then the equation depends on the difference $\tau = t_1 - t_2$ and autocorrelation is defined by,

$$R_X(\tau) = E[X(t)X(t+\tau)] \quad (3.10)$$

Because autocorrelation is a function of time, it also has a spectral interpretation in the frequency domain. Hence, for a stationary process the temporal-spectral relationship (known as the *Wiener-Khinchine* relation) is given by [17]:

$$S_X(j\omega) = \mathfrak{F}[R_X(\tau)] = \int_{-\infty}^{\infty} R_X(\tau) e^{-j\omega\tau} d\tau \quad (3.11)$$

where $\mathfrak{F}[\cdot]$ is the Fourier transform and ω is the nu,ber of (2π) cycles per second. This function is called the power spectral density of the random signal and is of high importance because it shows the relationship between time and frequency spectrum representations of the same signal [6].

3.2.4 White Noise

A particular case of a random signal is the case where the autocorrelation function is a *dirac delta* finction $\delta(\tau)$ which has zero value everywhere except when $\tau = 0$, as shown below [6]:

$$R_X(\tau) = \begin{cases} Aif\tau = 0 \\ else0 \end{cases} \quad (3.12)$$

where A is the constant magnitude.

In this case where the autocorrelation is a 'spike', the Fourier transform results in a constant frequency spectrum, as showm in the graphics in Fig. 3.2. White noise is a noise with mean equal to zero and with non related values, where the functions R_X and S_X both have power at all frequencies in the spectrum, but are totally uncorrelated with itself at any time except at $\tau = 0$. So, white noise is considered as independent, any sample of the signal at one time is uncorrelated from a sample at any other time [6].

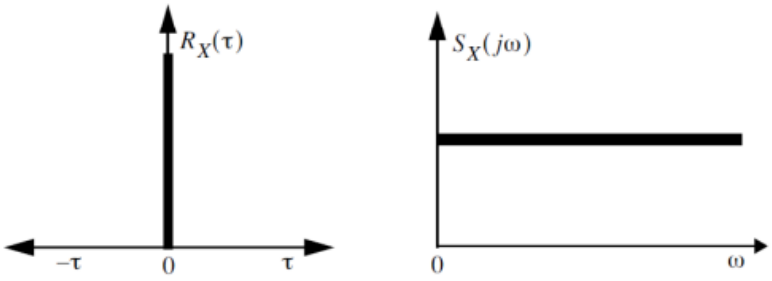

Figure 3.2: White noise in the time (left) and in the frequency domain (right). [6]

Often random signals are modelled as filtered white noise. This conveys that we could filter the output of a white noise hypothetical source to achieve a non-white source is band-limited in the frequency domain as well as more correlated in the time domain [6].

3.3 Kalman Filter

The most well-known and often-used mathematical tool for stochastic estimation from noisy measurements is known as the Kalman filter. This filter is an optimal recursive data processing mathematical algorithm that is able to incorporate all information that is provided to it. This means that it processes all attainable measurements, regardless of their precision, to estimate the value of variables of interest. This mathematical tool relies on [6]:

1. Knowledge of the system and measurement device dynamics;
2. The statistical description of the system noises, measurement errors and uncertainty in the dynamics models;
3. Any available information about initial conditions of the variable of interest.

The Kalman filter does not require all previous obtained data to be kept in storage and reprocessed every time a new measurement is taken. It combines all available measurement data with the prior knowledge of the system and measuring instruments, to produce an estimate of the values of the variables of interest so that the error is minimized statistically. That is why the Kalman filter is often preferred in comparison with other filters [6].

3.3.1 The Discrete Kalman Filter

The equations for the Kalman filter can be divided into two groups: *time update* equations and *measurement update* equations. The time update equations are in charge of projecting in time the current state and error covariance estimates to attain the *a priori* estimates for the next time step. The measurement update equations are responsible for the *feedback*, that is, they are meant to incorporate a new measurement into the *a priori* estimate to achieve an improved *a posteriori* estimate [6].

The figure below shows the discrete Kalman filter cycle. The *time update* equations projects the current state estimate and the *measurement update* adjusts the estimation by an actual measurement at that time [6].

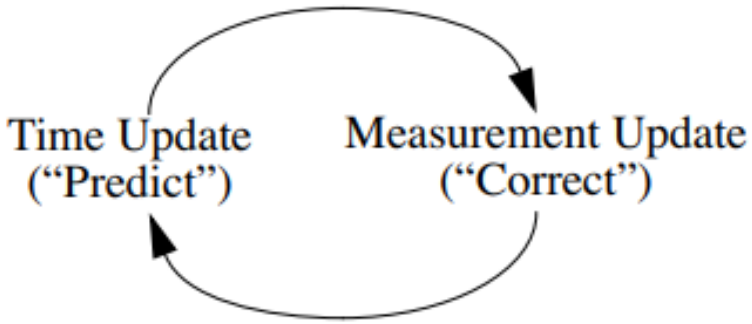

Figure 3.3: The Discrete Kalman Filter. [6]

The *time update* equations for a state $x\Re^n$ are:

$$\hat{x_k^-} = Ax^{k-1+Bu_k} \quad (3.13)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (3.14)$$

Where:

- $\hat{x_k^-} \in \Re^n$: is the state *a priori* estimate at step k;

- $\hat{x_k} \in \Re^n$: represents the state *a posteriori* estimate at step k;

- A: is a nxn matrix that relates the state at the previous time step $k-1$ to the current step k;

- B: $nx1$ matrix that links the optional control input u to the state x;

- u_k: optional control input defined in $\Re^1$;

- P_k^-: a *a priori* estimate error covariance:

- P_K: corresponds to the $aposteriori$ error covariance;

- Q: represents the process noise covariance matrix.

The *measurement update* equations for the discrete filter are[6]:

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (3.15)$$

$$\hat{x_k} = \hat{x_k^-} + K_k(z_k - H\hat{x_k^-}) \quad (3.16)$$

$$P_k = (I - K_k H) P_k^- \quad (3.17)$$

where H is a mxn matrix that relates the state to the measurement z_k.

The measurement $z \in \Re^m$ is calculated by :

$$z_k = H x_k + \nu_k \quad (3.18)$$

Where:

- ν_k: is the measurement noise;

- R: corresponds to the measurement error covariance;

- K: nxm matrix that represents the gain or blending factor that minimizes the *a posteriori* error covariance,

The difference $(z_k - H\hat{x_k^-})$ is named the residual or the measurement innovation and displays the variation between the predicted measurement $H\hat{x_k^-}$ and the actual measurement z_k. If the residual is zero it means that the two measurements are in full accord [6].

Firstly, it is necessary to compute the gain. K_k followed by the measure of the process to obtain z_k, allowing us to generate an *a posteriori* state estimates as well as an *a posteriori* error covariance estimate [6].

After each time and measurement update, the procedure is repeated with the previous $aposteriori$ estimates used to project new $apriori$ estimates. The recursive nature of the Kalman filter, as illustrated in Fig. 4.2, is one of its appealing aspects [6].

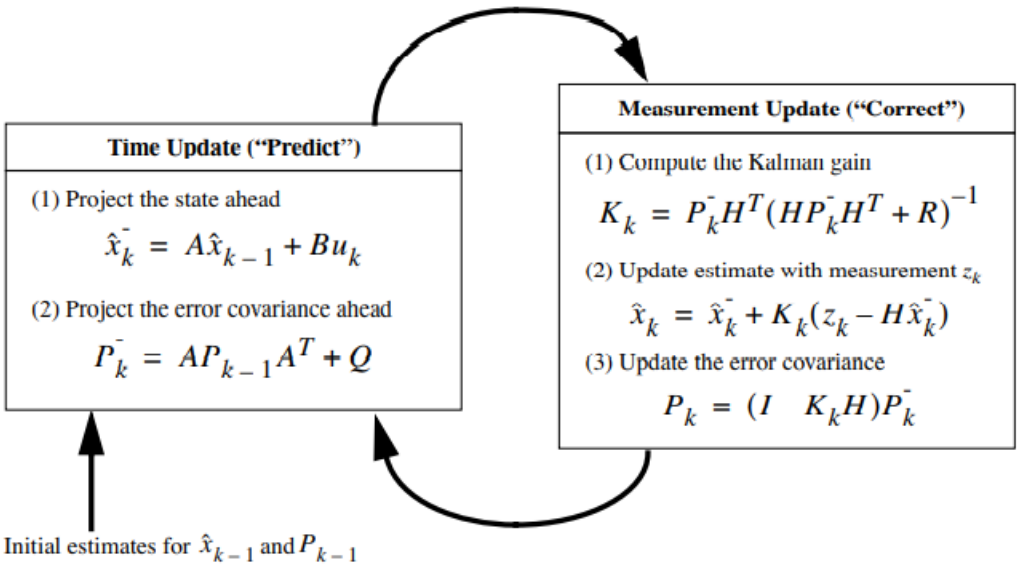

Figure 3.4: Complete operation of the Kalman filter. [6]

3.3.2 Covariance Matrices Assumptions

Choosing the assumed covariance matrices Q, R and P_0 is crucial to the estimation performance of a Kalman filter. The selection of P_0 as well as the initial state, x_0, affects the initial convergence of the filter. Although, in many applications, the effects of P_0 is not significant and is arbitrarily initialized to an identity matrix in order to simplify the filter [18].

More significantly, the effects of Q and R affect the overall performance of the filter. This is easily explainable as we see these matrices as weighting factors between the prediction equations and the measurement equations [18].

Considering a larger Q implies considering a larger uncertainty in the state equations, making the result of these equations less trustworthy, resulting in the filter correcting more with measurement update. A higher value for Q leads to a higher *a priori* error covariance P, which prompts a higher gain matrix, K [18].

Correspondingly, considering a larger R is equivalent to a larger uncertainty within the measurement, leading to trusting the measurement less, which means that the filter should correct less with the measurement update. A larger R results in smaller values of K, attributing less importance to the residuals [18].

In the implementation of the filter, the measurement noise covariance R is usually measured prior to the operation of the filter. The matrix can be determined by off-line sample measurements [6].

The selection of the process noise covariance has more complications as it is not possible to directly monitor the process we are estimating. The determination of Q can lead to acceptable results in a poor process model if it brings enough uncertainty into the process [6].

In a discrete Kalman filter, if Q and R are deemed constant in time, the error covariance P, and the gain factor, K, will stabilize quickly and then remain constant [6].

3.3.3 Discrete Kalman Filter Limitations

As any estimation process, the discrete Kalman filter presents some limitations that

can induce to erroneous estimations of a satellite trajectory.

The main limitation in the classical form of the filter is that its optimality holds only when the filter is projected to linear models (which is what it was designed for). However, in practical applications, dynamic models of the system (non-linear or *quasi*-linear models) contains uncertainties that generates performance degradation within the conventional filter [19]. To overcome this, it is imperative to convert the non-linear system into a system of linear equations that closely resemble the model, while accounting for a lesser filter's precision.

Another constraint resides in the selection of the matrices, mainly the transition matrix, A, and the observation matrix, H. In the conventional filter, these matrices have known and constant values, when, in reality, gaussian uncertainties are associated to them [20].

Another problem relates to the fact that the associated uncertainties are in nature white and gaussian, which often makes it impossible to observe, leading to a white noise approximation [6].

As referred in the previous subchapter, the Kalman filter heavily relies on the assumptions of the matrices which is a constraint of its own. If the chosen values for the transition matrix and the observation matrix are subpar, then the performance of the filter will be subpar as well. If chosen higher values, the filter will present considerable discrepancy between the measured and the estimated values, which will tend to increase at each iteration, affecting the filter's performance [6].

Lastly, the discrete Kalman filter is not prepared to correct estimation problem when running. A divergence phenomenon is easily created by an erroneous measured state, that consequently, generates an inaccurate state estimation leading to more wrong estimates.

3.3.4 Robust Kalman Filter

As referred in 3.3.3, one of the limitations of the classical form of the Kalman filter is that the filter is only optimal when projected to linear models, which makes necessary to convert a non-linear model into a linear system. Which, ultimately, reduces the filter precision.

In the linearization process, the measurement and the transition matrices are con-

sidered to be known and constant, a simplification that, in reality, is not factual because the model that we wish to filter has related gaussian uncertainties that vary in time and affect the state transition matrix, A, and the measurement matrix, H.

To account for those uncertainties, one of the available options is a robust form of the Kalman filter.

The algorithm for the robust filter consists of two steps: firstly, to generate a new trajectory that accounts for the uncertainties, a new state estimator is created; then, that new trajectory is filtered resorting to the discrete Kalman filter [21].

For the robust Kalman filter, the uncertainties are defined as gaussian, variable parameter in the A and H matrices, and the process noise covariance, Q, and the measurement error covariance, R, are considered known and constant [21].

In the robust estimator, the new model equations are:

$$\hat{x_{k+1}} = [A + \triangle A_k]x_k + u_k \quad (3.19)$$

$$z_k = [H + \triangle H_K] + \nu_K \quad (3.20)$$

Where A and H are the known and constants 'parts' of the state transition and measurement matrices, respectively, and $\triangle A_K$ and $\triangle H_k$ are unknown matrices that describe time-varying uncertainties and are of the form [21]:

$$[\frac{\triangle A_k}{\triangle H_K}] = [\frac{H_1}{H_2}]A_k E \quad (3.21)$$

Where E is a constant known matrix and F_k is an unknown matrix that satisfies [21]:

$$F_k^T \times F_k \leq I \quad (3.22)$$

and H_1 and H_2 are known and constant matrices that are related to and affect A and H [21].

As it is standard for the conventional Kalman filter, there are some assumptions

made for its robust form as well.

The system in eq. 3.19 is quadratically stable if exists a positive definitive matrix P that, for all values of $\triangle A_k$, satisfies

$$[A + \triangle A_k]^T P + P[A + \triangle A_k] < 0 \quad (3.23)$$

The process noise covariance, Q, and the measurement error covariance, R, matrices are conjectured as non-negative and the state transition matrix, A, is assumed as inverted [21].

With all assumptions satisfied, the new state estimator can be written as

$$\hat{x_{k+1}} = \hat{A}\hat{x_k} + K(z_k - H\hat{x_k}) \quad (3.24)$$

where $\hat{A}$ is the updated state transition matrix, K is the gain matrix and $\hat{H}$ is the updated measurement matrix [21].

In order to have a stable estimator, it is imperative to first solve two Riccati equations. Firstly in order to the error covariance matriz, $P = P^T > 0$ [21]:

$$A^T P A - P + A^T P \overline{Q}^{1/2} (I - \overline{Q}^{1/2} P \overline{Q}^{1/2})^{-1} \overline{Q}^{1/2} P A + \varepsilon (E^T E - \nu I) = 0 \quad (3.25)$$

With ε and ν being small and positive scalars and $\overline{Q}$ is the new noise covariance matrix calculated by [21]:

$$\overline{Q} = Q + \frac{1}{\varepsilon} H_1 H_1^T \quad (3.26)$$

With the first Riccati equation solved the updated matrices are calculated by,

$$\hat{A} = A + \overline{Q}(P^{-1}) - \overline{Q}^{-1} A \quad (3.27)$$

$$\hat{H} = H + \frac{1}{\varepsilon} H_2 H_1^T (P^{-1} - \overline{Q})^{-1} H_1 H_2^T \quad (3.28)$$

and the $\hat{Q}$ and $\hat{R}$ matrices are obtained by,

$$\hat{Q} = \overline{Q} + \overline{Q}(P^{-1} - \overline{Q})^{-1}\overline{Q} \quad (3.29)$$

$$\hat{R} = R + \frac{1}{\varepsilon} H_2 H_2^T + \frac{1}{\varepsilon^2} H_2 H_1^T (P^{-1} - \overline{Q})^{-1} H_1 H_2^T \quad (3.30)$$

The cross-variance matrix is

$$F = \frac{1}{\varepsilon}(I - \overline{Q}P)^{-1} H_1 H_2^T \quad (3.31)$$

With the matrices updated, the 2^{nd} Riccati equation is calculated in order to $U = U^T > 0$ [21]:

$$\hat{A}U\hat{A}^T - U - (\hat{A}U\hat{H}^T + F)(\hat{R} - \hat{H}U\hat{H}^T)^{-1}\hat{A}U\hat{H} + \hat{E}^T + \hat{Q} = 0 \quad (3.32)$$

Which then, allows to obtain the updated gain matrix [21]:

$$K = (\hat{A}U\hat{H}^T)(\hat{R} + \hat{H}U\hat{H})^{-1} \quad (3.33)$$

And, finally, the new estimator is deduced [21]:

$$\hat{x_k^-} = A\hat{x_{k-1}^+} K_k[y_k - H\hat{x_k^+}] \quad (3.34)$$

The robust estimator will generate new states that account for the uncertainties associated to the state transition ,A, and the measurement, H, matrices.

However, the states generated will still be disturbed by white noise. To counteract

this, it is necessary to apply the discrete Kalman filter to estimate states that to not have white noise. The robust filter relies on a robust estimator, to account for the uncertainties, and a conventional Kalman filter, to generate the trajectory created by the robust filter.

Chapter 4. Applications and Results

This chapter focuses on presenting the results from the application of discrete Kalman filter and the robust Kalman filter, which are described and developed in chapter 4.2, to the trajectory of the satellite Iridium 179 - one of the five spares satellites launched on May 2023 to provide further backup for the company's telecom satellites.

All the results, simulations and graphs were attained using the $Python^{TM}$ programming language.

4.1 Mathematical Discretization Model

This section presents the specific discrete algorithm used in the computational formulation. Considering the model:

$$\begin{cases} x_{k+1} = A_d x_k + D_d \omega_k \\ y_k = C x_k + \nu_k \end{cases} \quad (4.1)$$

where ω_k and ν_k are white, Gaussian noises.

The Kalman filter begins with state prediction steps, $\hat{x_k}$. The new state and respective covariance matrix,P_{k-1} are calculated based on the former state and covariance matrix at step $k-1$.

$$\hat{x_k} = A_{k-1} x_{k-1} \quad (4.2)$$

The *a priori* error covariance matrix, P_k can then be calculated based on its former value, P_{k-1} ,the A_{k-1}, the matrix D_{k-1} and the process noise covariance matrix, Q_{k-1}.

$$P_k = A_{k-1}P_{k-1}A_{k-1}^T + D_{k-1}Q_{k-1}D_{k-1}^T \quad (4.3)$$

What follows after P_k is a correction of the predicted state. This is done by the gain matrix, K_k, which composed of the error covariance matrix, P_k, the C_k matrix and the measurement error covariance matrix.

$$K_k = P_k C_k^T (C_k P_k C_k^T + R_k)^{-1} \quad (4.4)$$

Which allow us to then determine e new signal state estimate, $\hat{x_{k|k}}$:

$$\hat{x_{k|k}} = \hat{x_{k|k-1}} + K_k(y_k - C_K \hat{x_{k|k-1}}) \quad (4.5)$$

Finally, with all corrections done and errors minimized, a new error covariance matrix, P_k, is calculated:

$$P_{k|k} = (I - K_k C_k) P_{k|k-1} \quad (4.6)$$

4.2 *Python* Formulation: Discrete Kalman Filter

In the *Python* software the discrete model implemented was of the form:

$$\begin{cases} x_{k+1} = A_d x_k + D_d \omega_k \\ y_k = C x_k + \nu_k \end{cases} \quad (4.7)$$

Firstly, the matrices A, D and C, which are the control entries of the system, where defined as:

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.8)$$

$$D = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.9)$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.10)$$

The step, k, chosen of the application was:

$$k = \Delta t = 1s \quad (4.11)$$

With the entry data defined, the initial matrices of the discrete algorithm A_d and D_d are calculated, as defined in .

The matrix A_d is defined by :

$$A_d = e^{A\Delta t} \quad (4.12)$$

which results in:

$$A_d = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.13)$$

The D_d is determined as,

$$D_d = (I\Delta t + \frac{A(\Delta t^2)}{2!} + \frac{A^2(\Delta t^3)}{3!} + \frac{A^{N-1}(\Delta t)^N}{N!})\Gamma \quad (4.14)$$

with N,

$$\|\frac{A^{N-1}}{N!}\| < \varepsilon, with \varepsilon = 0.001 \quad (4.15)$$

For $N = 7$, the result for D_d is,

$$D_d = \begin{bmatrix} 1.0000248 & 1 & 1 \\ 1 & 1.0000248 & 1 \\ 1 & 1 & 1.0000248 \\ 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad (4.16)$$

The measurement error covariance are defined as,

$$R = \lambda Id(6) \quad (4.17)$$

where $\lambda \in [1, 10]$ and considering $\lambda = 10$.

And the process noise covariance matrix, Q_0, is calculated as

$$Q_0 = \lambda(3) \quad (4.18)$$

To define x_0, the first measurements obtained through GMAT,

$$x_0 = \begin{bmatrix} x \\ y \\ z \\ V_x \\ V_y \\ V_z \end{bmatrix} = \begin{bmatrix} -3709109 \\ 5945052 \\ 14921 \\ -360.992 \\ 246.544 \\ 7529.352 \end{bmatrix} \quad (4.19)$$

To determine the error covariance matrix, it was first defined the maximun error relative to the measurements of x_0 to then calculate σ_1, σ_2, σ_3, σ_4, σ_5 and σ_6,

$$\varepsilon_0^{(1)} = 15 \Rightarrow \sigma_1 = \frac{\varepsilon_0^{(1)}}{3} = \frac{15}{3} \quad (4.20)$$

$$\varepsilon_0^{(2)} = 15 \Rightarrow \sigma_2 = \frac{\varepsilon_0^{(2)}}{3} = \frac{15}{3} \quad (4.21)$$

$$\varepsilon_0^{(3)} = 40 \Rightarrow \sigma_3 = \frac{\varepsilon_0^{(3)}}{3} = \frac{40}{3} \quad (4.22)$$

$$\varepsilon_0^{(4)} = 3 \Rightarrow \sigma_4 = \frac{\varepsilon_0^{(4)}}{3} = \frac{3}{3} \quad (4.23)$$

$$\varepsilon_0^{(5)} = 3 \Rightarrow \sigma_5 = \frac{\varepsilon_0^{(5)}}{3} = \frac{3}{3} \quad (4.24)$$

$$\varepsilon_0^{(6)} = 3 \Rightarrow \sigma_6 = \frac{\varepsilon_0^{(6)}}{3} = \frac{3}{3} \quad (4.25)$$

So, the error covariance matrix, P_0, is determined as

$$P_0 = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_5^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_6^2 \end{bmatrix} = \begin{bmatrix} 25 & 0 & 0 & 0 & 0 & 0 \\ 0 & 25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 177.778 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.26}$$

The Kalman filter was then implemented, running through 1843 measurements, obtained through GMAT.

4.2.1 Results: Discrete Kalman Filter

As it was mentioned before, all the measurements are subject to noise and uncertainties. So, it is for important to account for those and add them to the measurements as to aim for realistic results. For this measurements, noise was associated to the values was assumed to follow a normal distribution of $\mathcal{N}(15, 0)$

The results for the position components estimation are shown. First, for the component x, Fig. 4.1:

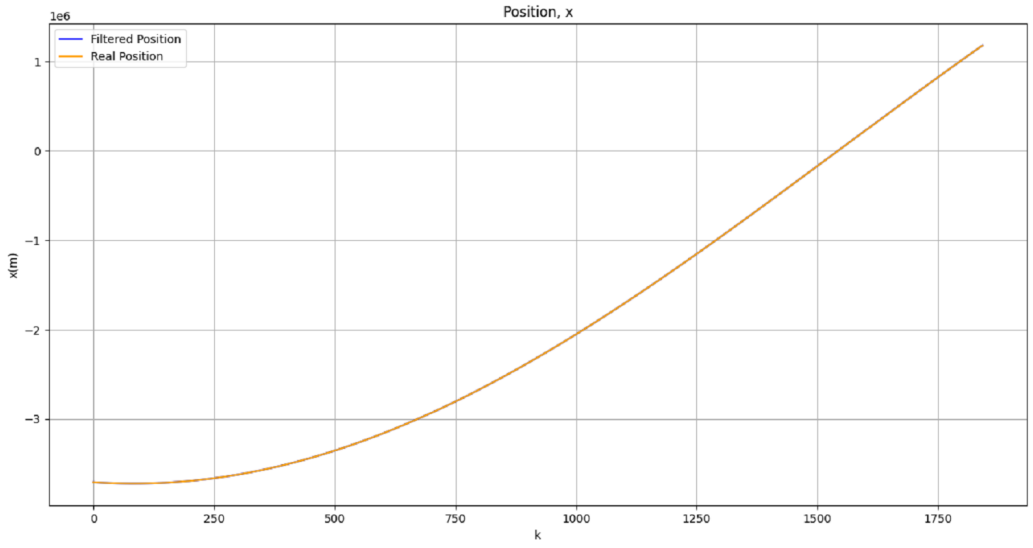

Figure 4.1: Satellite's estimation for the x component of the position along k.

The differences between the initial trajectory and the estimation cannot be seen with

an unaided sight. The higgest difference can be in by enhancing the graphic at $k \in [0, 1]s$.

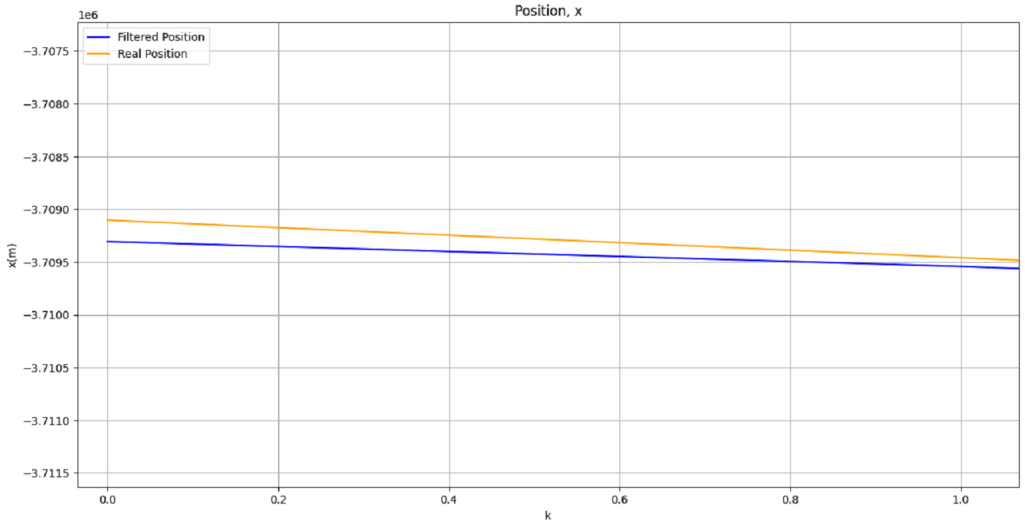

Figure 4.2: Enhancement of the graphic 4.1 at $k \in [0, 1]s$.

There is a difference of $223m$ at $k = 0$, for the x component, that rappidly approximates to the real values along k, with an average difference of $6.08m$ between the real and filtered trajectories along k.

The results for the coordinate y is represented in Fig. 4.3,

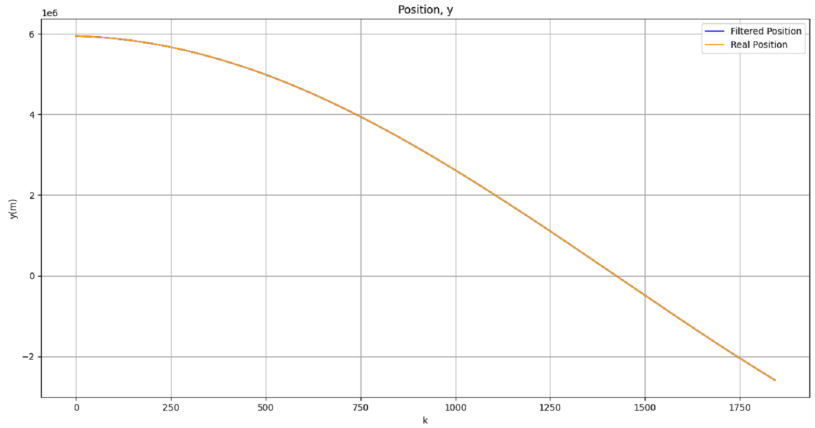

Figure 4.3: Satellite's estimation for the y component of the position along k.

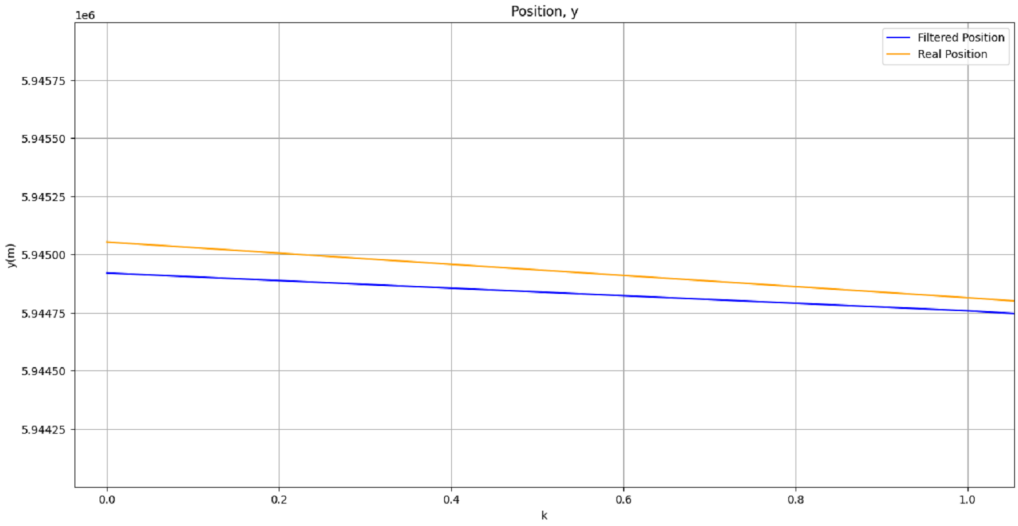

Figure 4.4: Enhancement of the graphic 4.3 at $k \in [0, 1]s$.

The same happens for the y component as for x. In Fig. 4.4, the enhancement at $k \in [0, 1]s$, the discrepancy between the trajectories has its higher point of $148, 3m$ when $k = 0$, with an average difference of $6.23m$ between the two orbits along k.

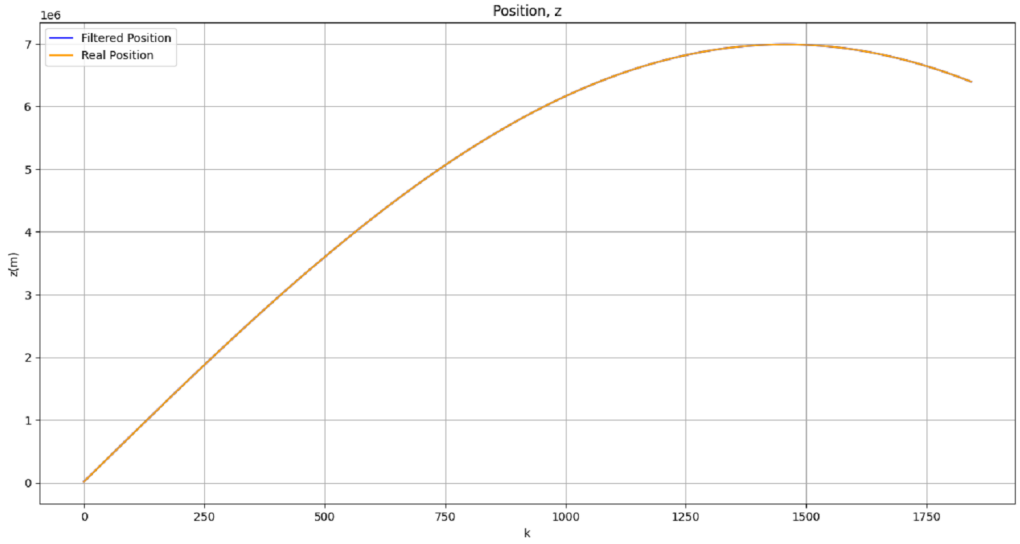

Figure 4.5: Satellite's estimation for the z component of the position along k.

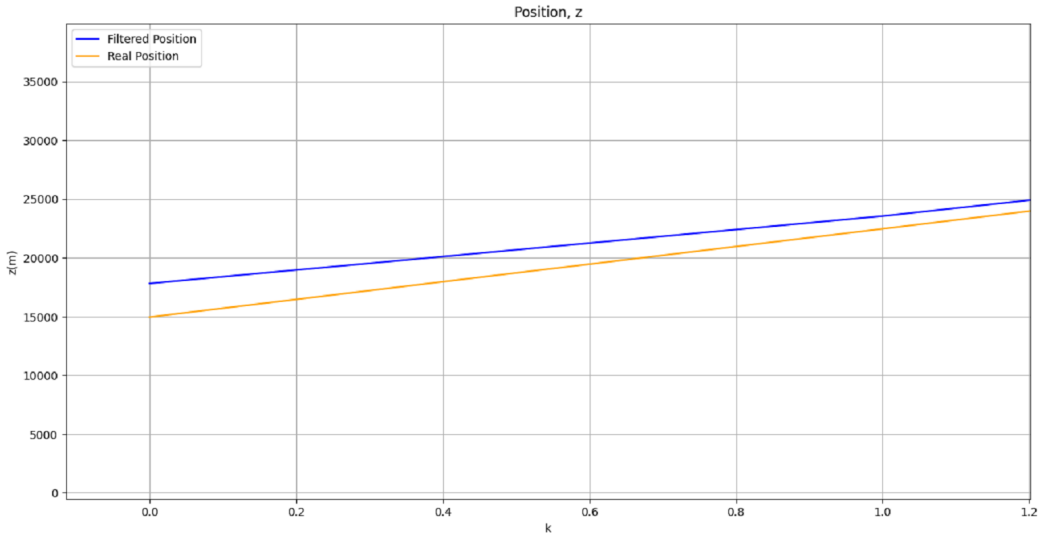

Figure 4.6: Enhancement of the graphic 4.5 at $k \in [0, 1.2]s$.

The results for the coordinate z also shows little difference between the filter's results and the measured values. Yet, at $k \in [0, 1.2]s$, it shows the biggest difference, of all component, of $3000m$, but also quickly being closer to the initial trajectory and reaching an average discrepancy of $22.9m$ along k.

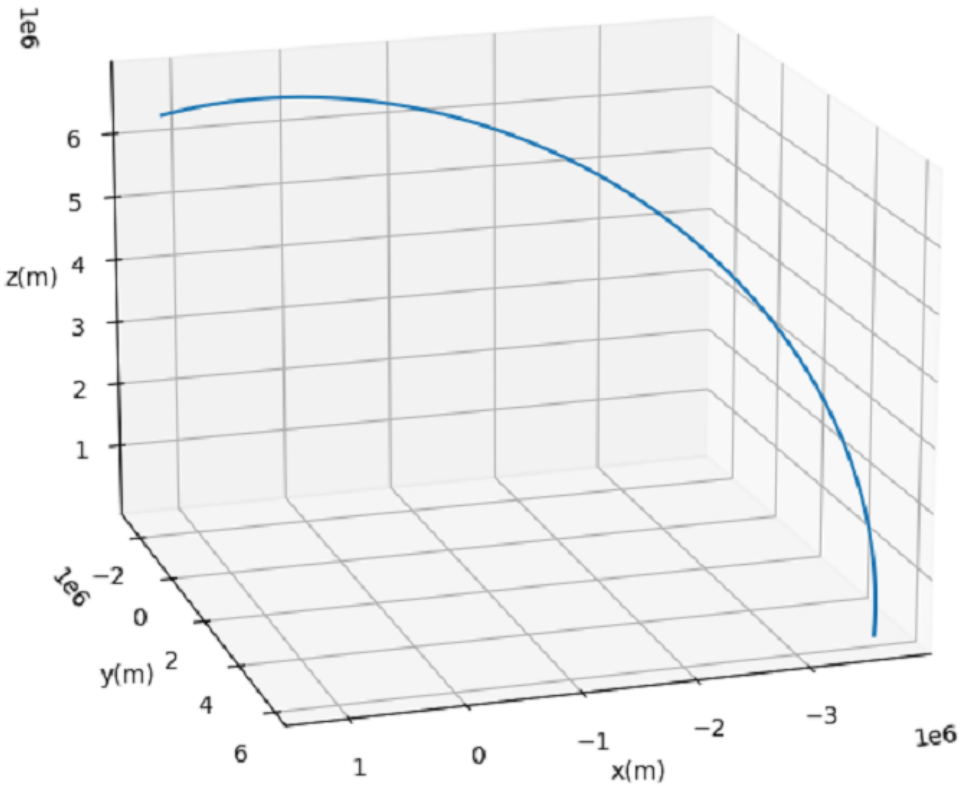

Figure 4.7: Satellite position estimation.

The estimated trajectory is shown in Fig. 4.7. Graphic's analysis verifies that after the implementation of the Kalman filter the results are akin to the real values.

This concludes that the filter's application decreases, and practically eliminates, the noise in the measurements, giving results of the Iridium satellite trajectory with a high level of accuracy.

As for the velocity relative to each components along k, although evolving differently it also obtained values with a high precision. For x, the results for the velocity component, V_y, are,

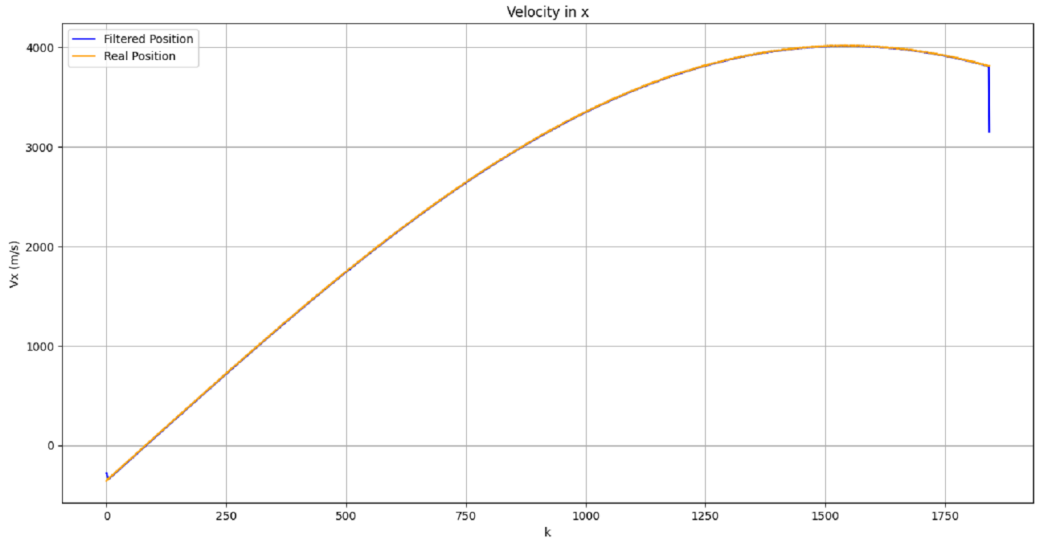

Figure 4.8: Satellite's estimation for the x component of the velocity along k.

Again, it is possibly to conclude that the real values and the filter's estimate are fairly similar. At $k \in [0, 300]s$, it is possible to see the trajectory behaviour closely. Where at $k = 0$, there is a difference of $32.3m/s$. For the x component of the velocity there was an average discrepancy of $4.52m/s$ along k.

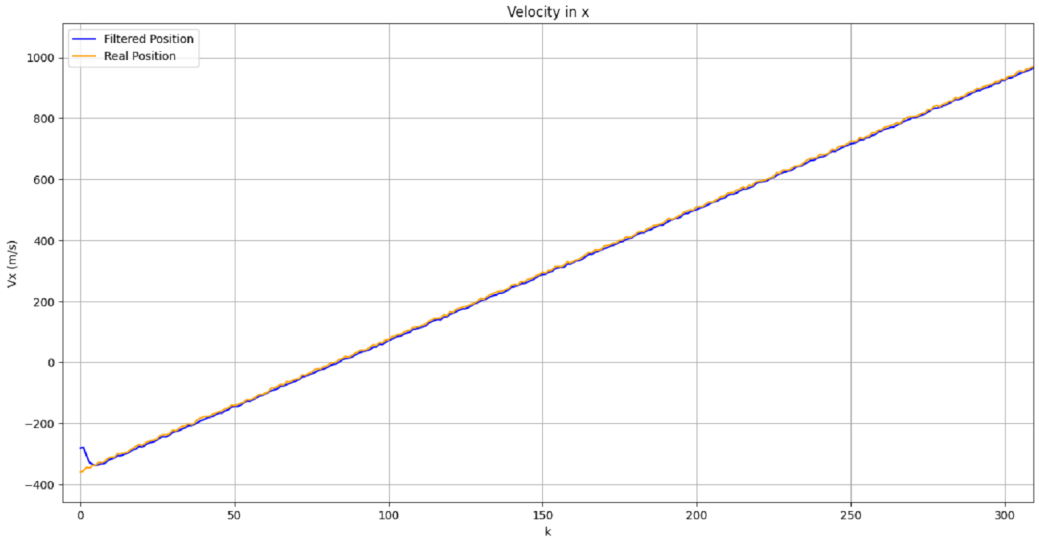

Figure 4.9: Enhancement of the satellite's estimation for the x component of the velocity at $k \in [0, 300]s$.

The results for the velocity in the y component, V_y, are, Fig. 4.10,

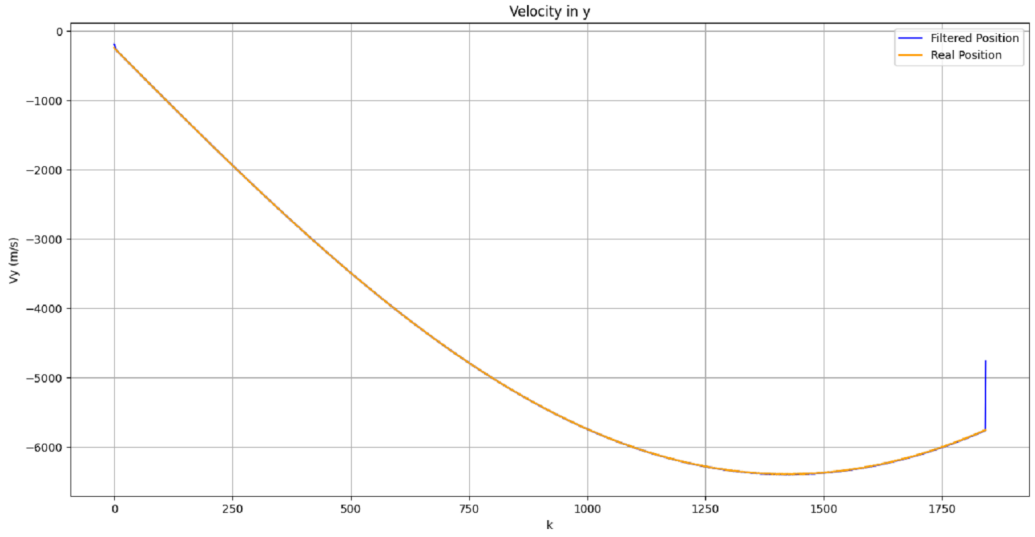

Figure 4.10: Satellite's estimation for the y component of the velocity along k.

The estimation along k shows that the application reaches repeatedly satisfactory results and always stabilizing closer to the initial values. At $k \in [0, 300]s$, Fig. 4.11, shows the highest disparity in the estimation, at $14m/s$. This difference stabilizes at around $5.04m/s$ along k.

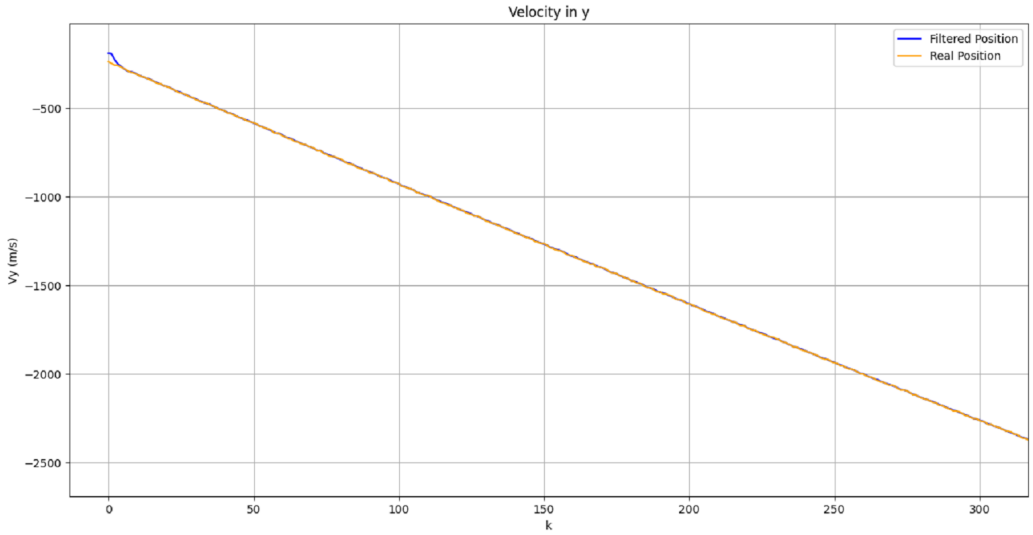

Figure 4.11: Enhancement of the satellite's estimation for the y component of the velocity at $k \in [0, 300]s$.

Finally, the results for the velocity z component were akin to the measured values, Fig. 4.12,

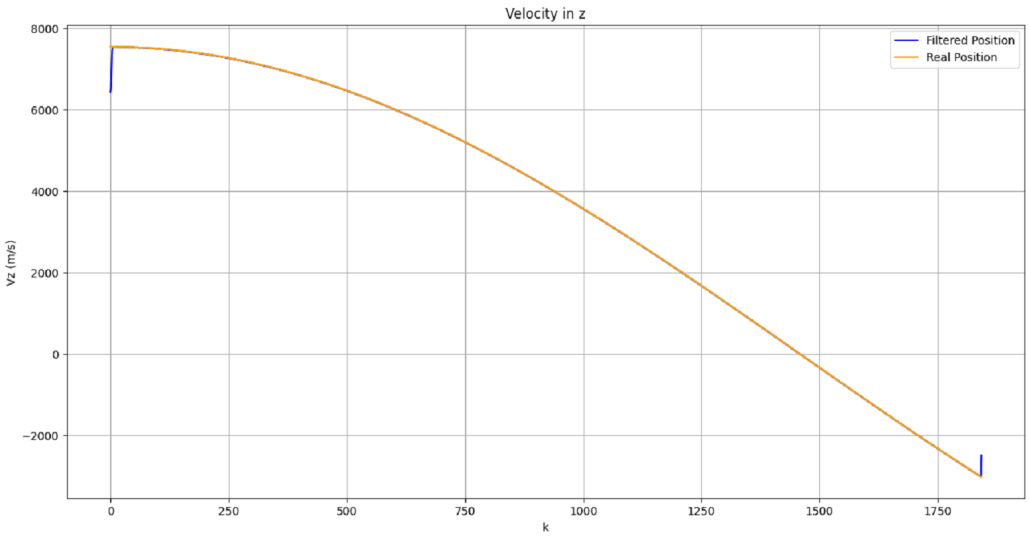

Figure 4.12: Satellite's estimation for the z component of the velocity along k.

Where at $k \in [0, 300]s$, Fig. 4.13, the estimation has the highest difference from the measurements at $445m/s$. Along all values of$k$ the average discrepancy between the real velocity and the filtered one was of $16.7m/s$

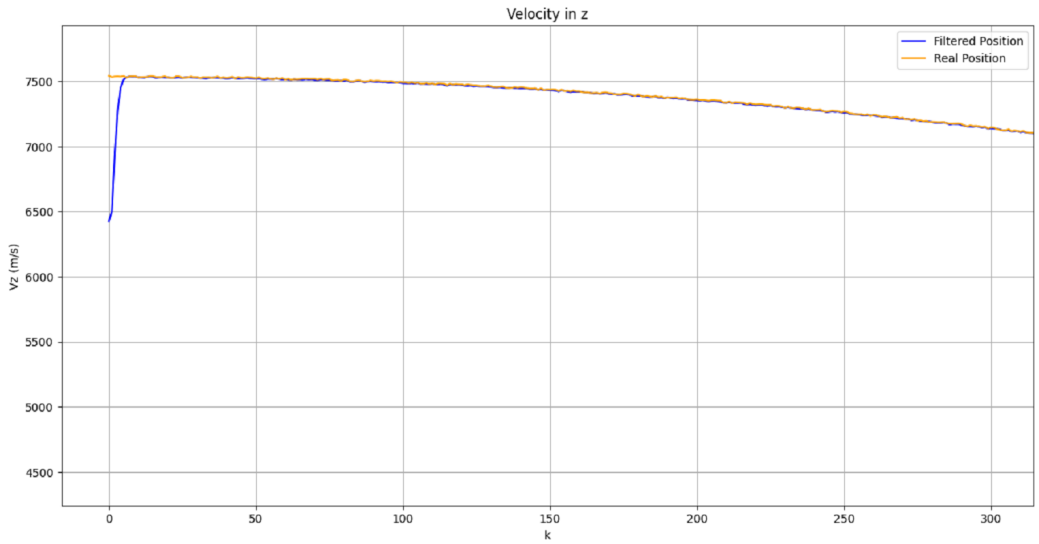

Figure 4.13: Enhancement of the satellite's estimation for the z component of the velocity at $k \in [0, 300]s$.

The filter's results for the velocity shown that the algorithm rapidly decreases the noise from the measurement. Meaning that the filter calculates a highly precise state estimation matrix, x_k.

The graphic in Fig. 4.14, shows the performance quantification of the conventional Kalman filter, by relating the error covariance matrix along k, P_k, and the first error covariance matrix, P_0. And although the error wasn't calculated, it is known that the higher the performance quantification the higher the estimation error.

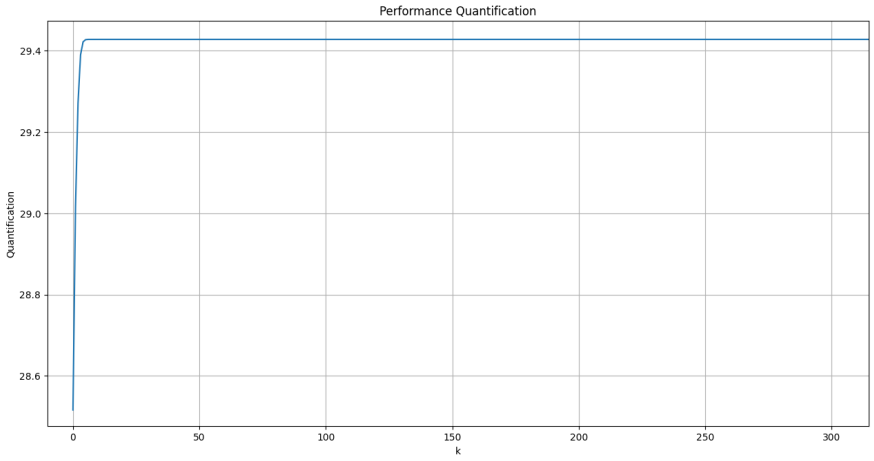

Figure 4.14: Performance quantification along the process.

The quantification of the performance of the estimation stabilized at 29.4, at $k = 9s$, This, again, infers that the overall estimation gives fairly good results.

4.2.2 Results: Robust Kalman filter

The robust application was all in all made by altering the state transiton matrix, $\bar{A}$, as the sum of the defined state transition matrix, eq. 4.8, and μI_d,

$$\bar{A} = A + \mu I_d = \quad (4.27)$$

where μ is a positive , $\mu > 0$,chosen scalar between $\mu \in [5, 100]$, and I_d is the identity matrix. The best results were obtained for $\mu = 15$.

$$x_{k+1} = (A + \mu I_d)x_k + D\omega_k = \quad (4.28)$$
$$= Ax_k + (\mu I_d x_k + D\omega_k) \quad (4.29)$$

What Eq. 4.29 shows is that the system is artificially disturbed in order to create more robustness (by increasing the eigenvalues of A). This means that the filter processes an even noisier system, forcing it to compensate for the noise and increasing the filter's performance and achieving results with pratically no effects from it.

The satellite's filtered position estimation were finely eliminated the noise applied in the measurements achieved a higher level of precision and robustness comparing to the already satisfactory results of the discrete form of the filter.

The estimation of the trajectory relative to the x component, Fig. 4.15, rapidly stabilized nearest to the initial trajectory,

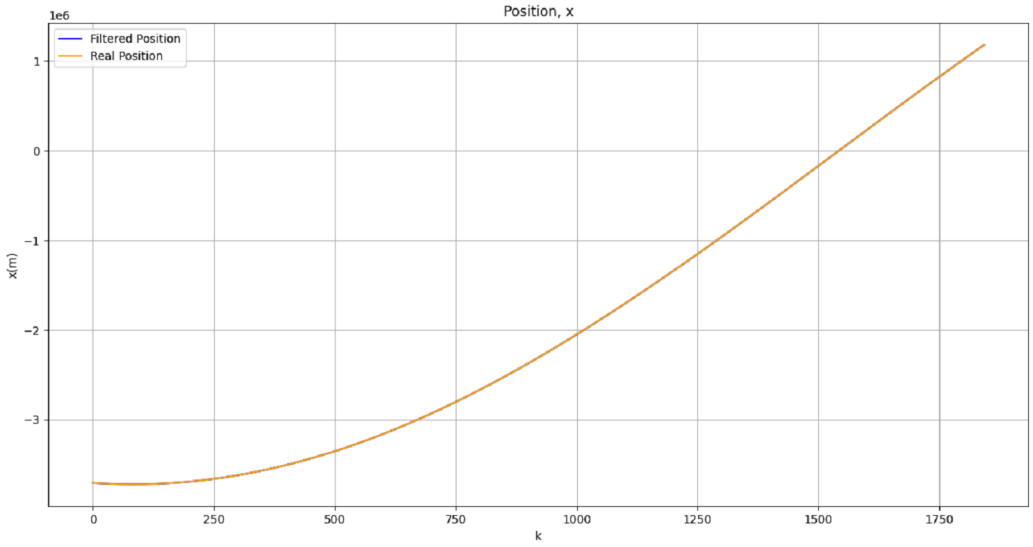

Figure 4.15: Satellite's robust estimation for the x component of the position along k.

The discrepancy was at a highest of $203.1m$ at $k \in [0; 1]s$, Fig. 4.16, and stabilized to an average discrepancy of at $5.02m$, along k.

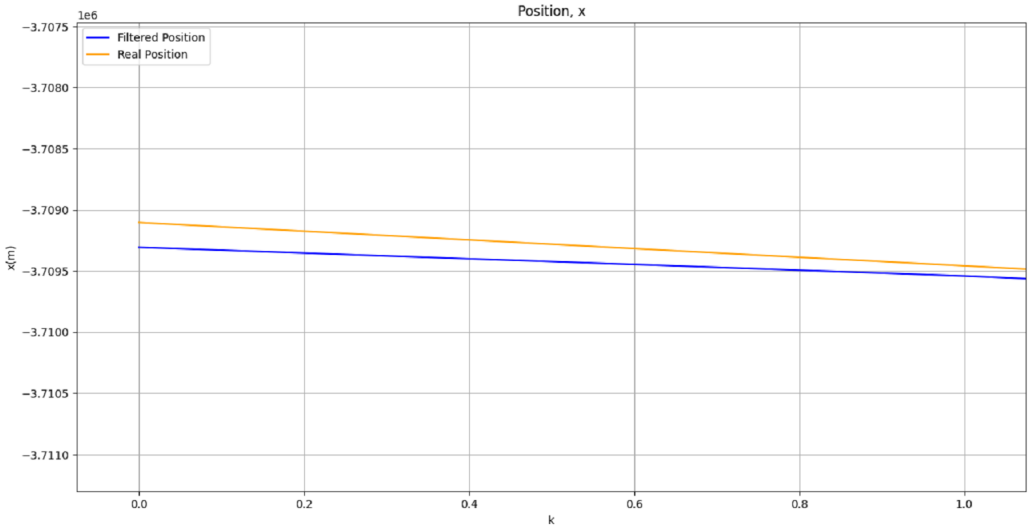

Figure 4.16: Enhancement of the graphic 4.15 at $k \in [0, 1]s$.

As it will be proven for all coordinates the disparity between the measured values and the estimation decreases to smaller values compared to those compilated in 4.2.1.

The results for the position along the y component were, Fig. 4.17,

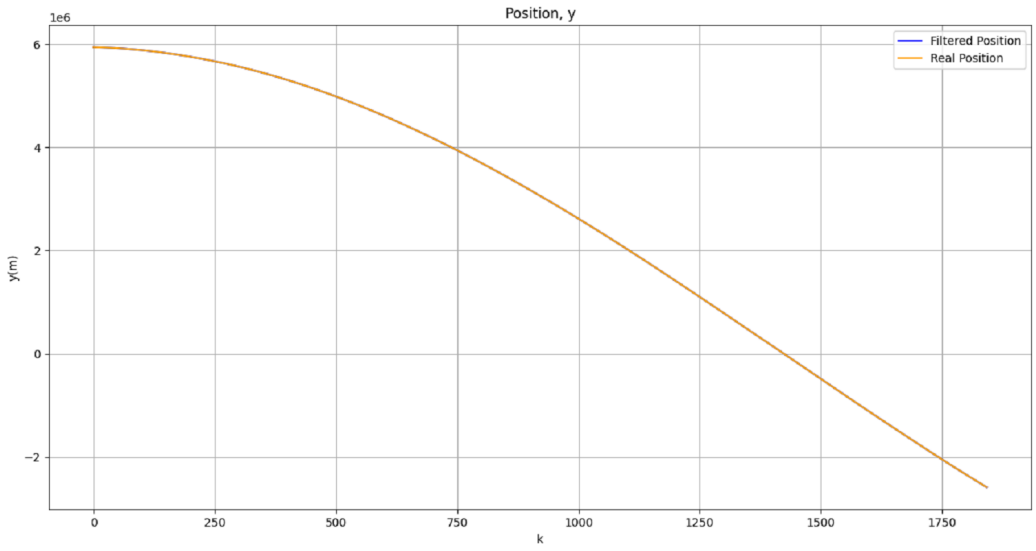

Figure 4.17: Satellite's robust estimation for the y component of the position along k.

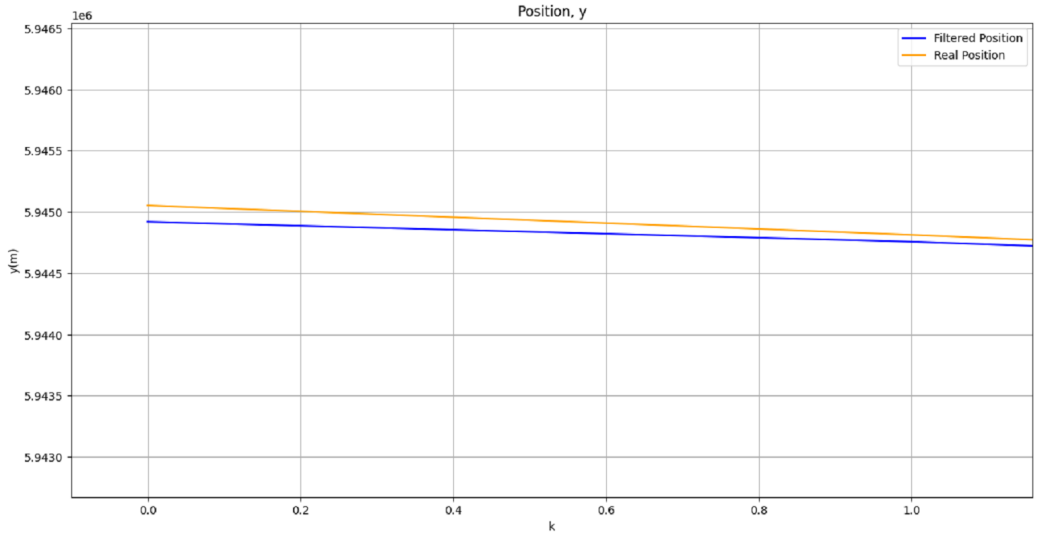

Figure 4.18: Enhancement of the graphic 4.17 at $k \in [0, 1]s$.

The robust application also gave favorable results for the position along y, with the highest difference being at $133m$, at $k \in [0; 1]s$, decreasing $15.3m$ compared to the classical algorithm. The average difference between the two trajectories in regards to the y component of the positon was of $3.24m$, along k.

As for z, the application results show again little difference between the two trajectories,

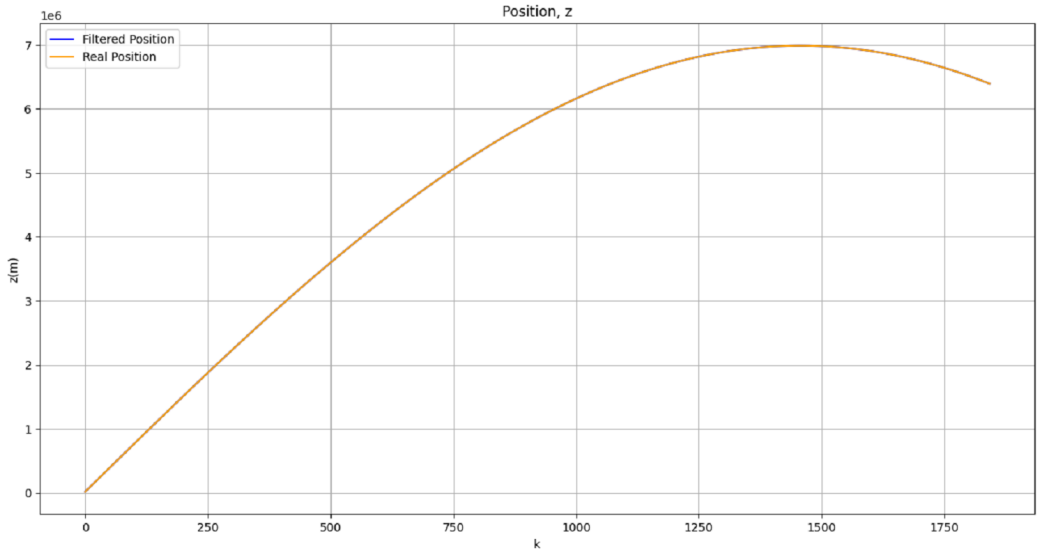

Figure 4.19: Satellite's robust estimation for the z component of the position along k.

At $k \in [0; 1]s$, Fig. 4.20, the robust estimation has a decrease of $150m$ in the variance of the position (with the highest discrepancy at $2850m$, at $k = 0$, for the z component). It is in this component where it is seen very clear differences between the two algorithms: specifically, the average discrepancy between the two trajectories is of $5.5m$ along k.

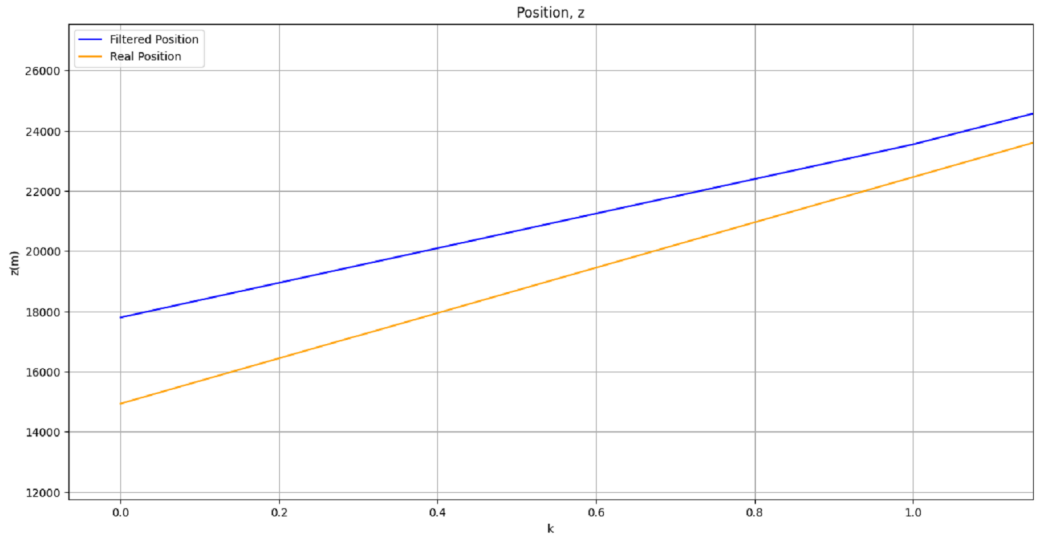

Figure 4.20: Enhancement of the graphic 4.19 at $k \in [0, 1]s$.

The estimated orbit, Fig. 4.21, demonstrated that the robust application has a high precision and robustness at conjecturing the satellite's position.

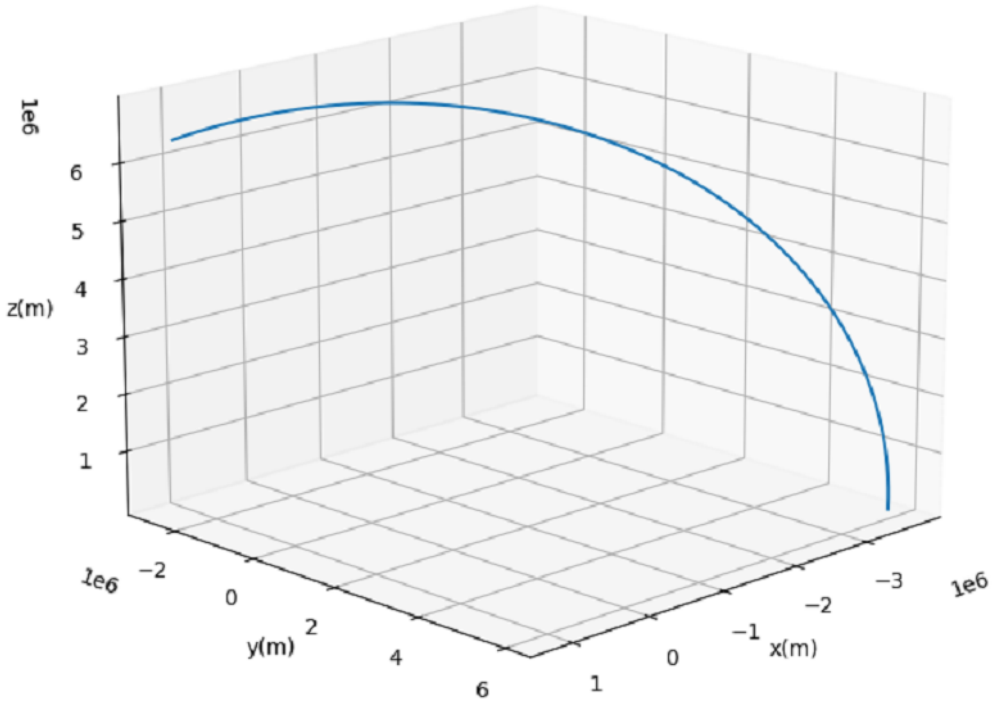

Figure 4.21: Satellite position robust estimation.

Concluding that, although the classical filter is well enough for a good estimation, it is reached a more precise estimation through the robust algorithm.

The same precision is achieved as well for the relative velocities. With the results for the velocity x component, V_x, being, Fig. 4.22,

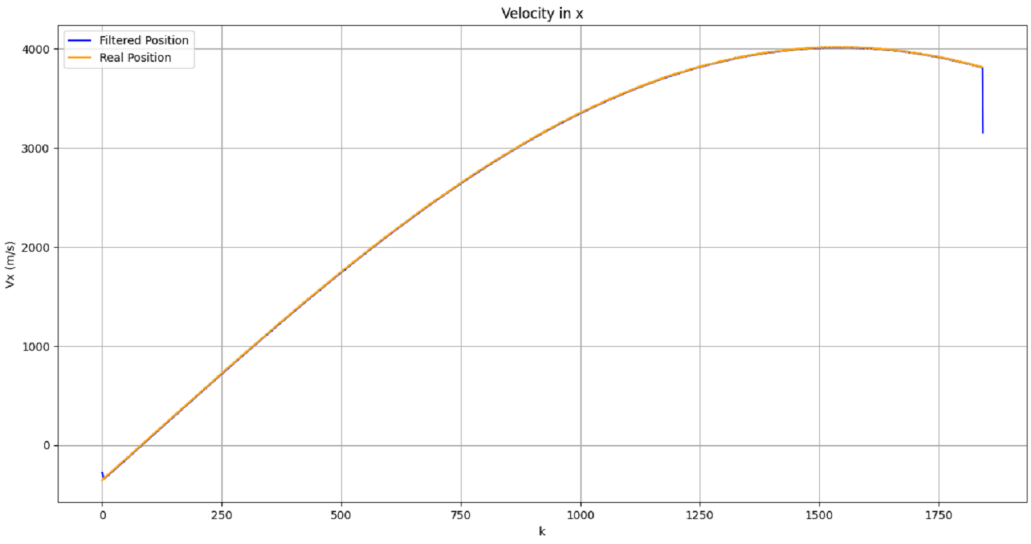

Figure 4.22: Satellite's robust estimation for the x component of the velocity along k.

The filter's results are, again, akin to the real values, showing a higher difference at $k \in [0, 300]s$ of $79m/s$, Fig. 4.23. Although the robust filter has a higher difference for V_x than the classical application, the robust algorithm manages to quickly stabilize, for a $2.86m/s$ average discrepancy along k, in the case for V_x.

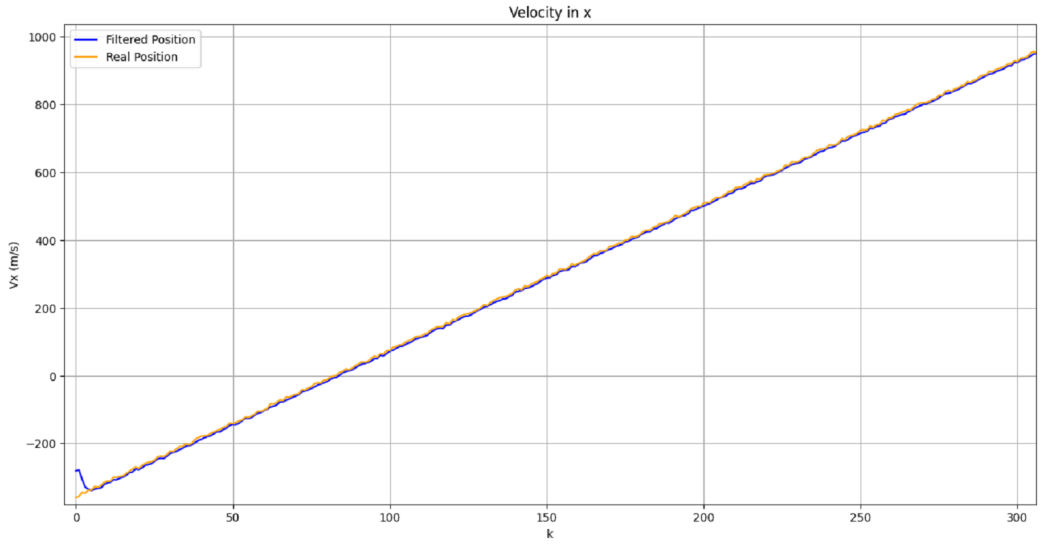

Figure 4.23: Enhancement of the satellite's estimation for the x component of the velocity at $k \in [0, 300]s$.

The velocity's estimation along to the y-axis, Fig. 4.24, presented a higher variance at $k = 0$, but presented an average difference of $4.32m/s$ along k.

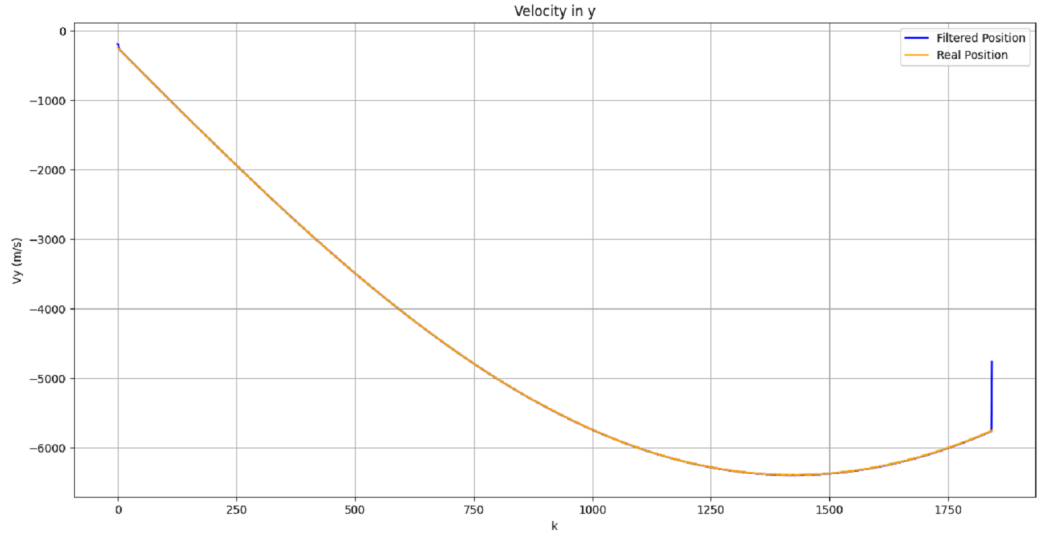

Figure 4.24: Satellite's robust estimation for the y component of the velocity along k.

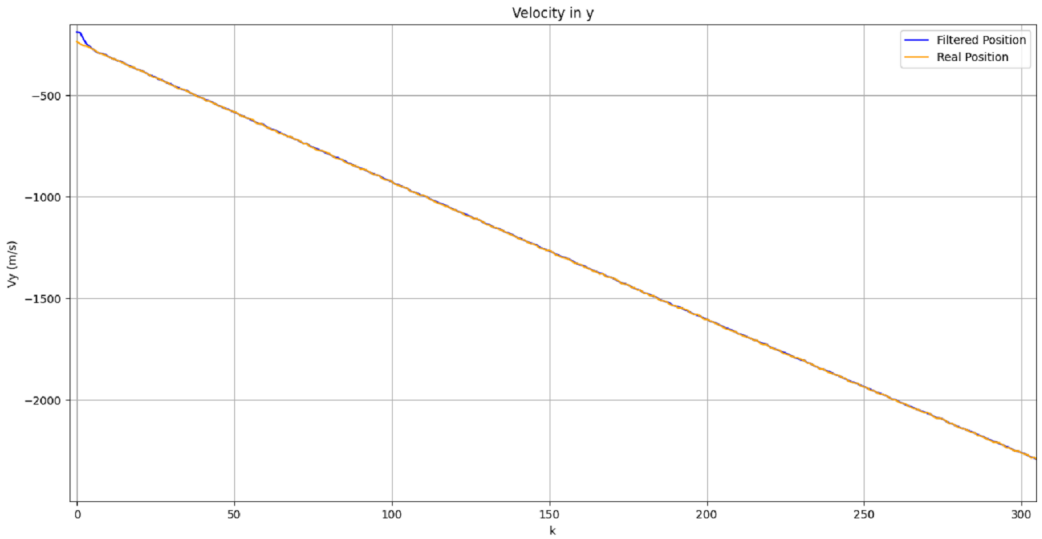

Figure 4.25: Enhancement of the satellite's estimation for the y component of the velocity at $k \in [0, 300]s$.

For the z component the values are expressed in Fig. 4.26, presenting an an average difference of $6.1m/s$ along k, Fig. 4.27, pratically eliminating all the added noise in the estimation. The estimation has a bigger discrepancy at $800m/s$.

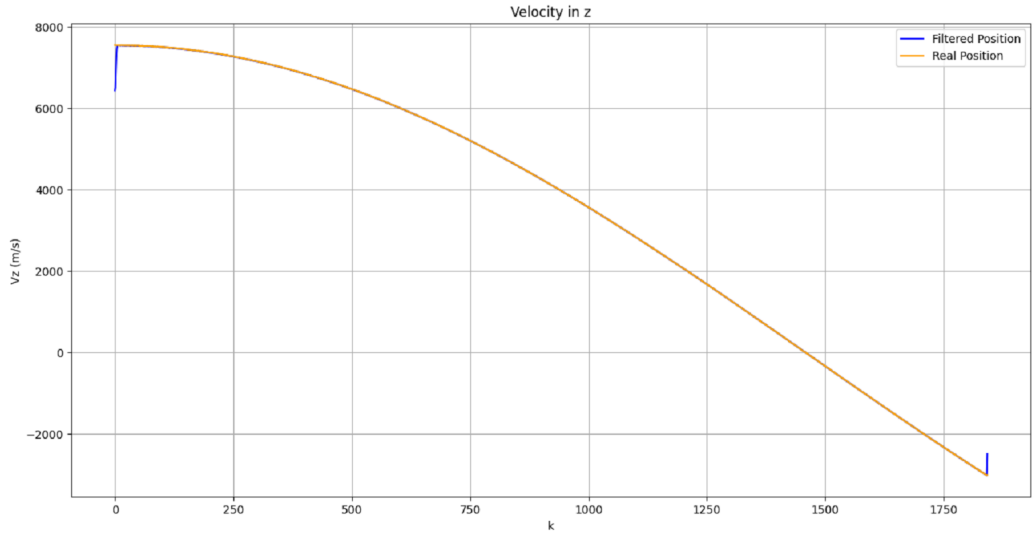

Figure 4.26: Satellite's robust estimation for the z component of the velocity along k.

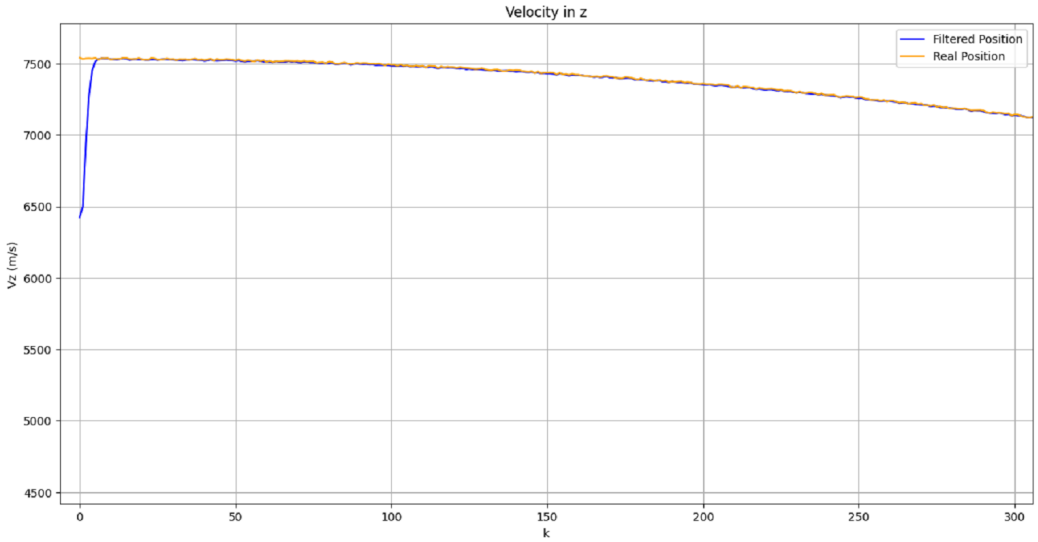

Figure 4.27: Enhancement of the satellite's estimation for the z component of the velocity at $k \in [0, 300]s$.

By also achieving more precise results for the state estimation matrix, x_k, it is inferred that the robust filter is an upgraded algorithm based on the discrete filter and despite the rigorous obtained results in the latter, it is imperative to always aim for optimal results.

This is also proven as the performance quantification, Fig. 4.28, related to the error covariance matrix, P_k, stabilizes at 18.5. Which means that the estimation error for the robust filter is lower than the one for the classical form.

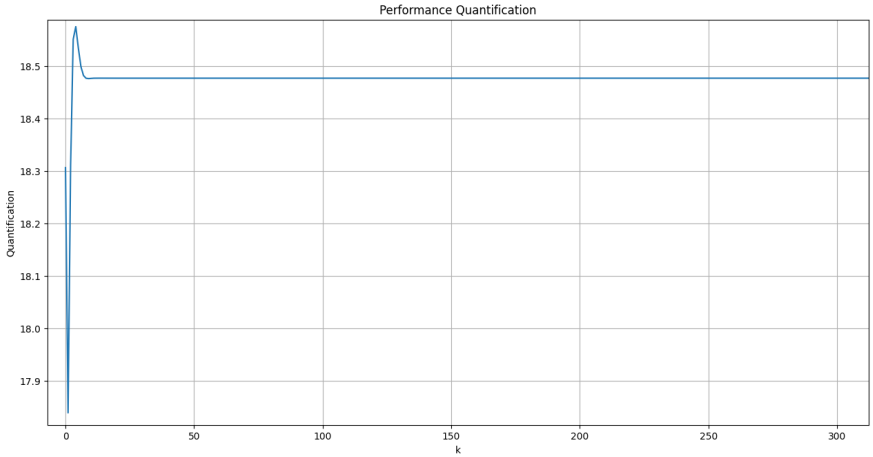

Figure 4.28: Performance quantification along the process.

The table in 4.29 shows the average difference, for all the components of x_k, between

the real and the filtered trajectories along k for both algorithms. It is, once again, inferable that for all components the robust Kalman filter has a higher precision in estimating trajectory values.

Components x_k	Results:Average Discrepancy	
	Conventional	Robust
x	6.08 m	5.02 m
y	6.23 m	3.24 m
z	22.9 m	5.5 m/s
V_x	4.52 m/s	2.86 m/s
V_y	5.04 m/s	4.32 m/s
V_z	16.7 m/s	7.03 m/s

Figure 4.29: Comparison of the average difference between the measured and the filtered values for each of the x_k components for both Kalman filters.

The rely on the Kalman filter for satellite trajectory estimation is understandable as it allow us to predict when it will be necessary to activate the control systems, especially for LEO orbits (like Iridium) as they suffer more orbital perturbations due to the effects from the phenomena explained in 2.2.

Chapter 5. Conclusion and Future Works

In this thesis, a LEO orbit was estimated based on noise-added measurements obtained through GMAT. The noise and uncertainties were accounted for because their existence is inevitable in any part of the process.

Currently, there are various methods and algorithms to predict orbits while accounting for the noise effects. This study focused on the classical and the robust form of the Kalman filter: with the latter giving more precise results by accounting for the uncertainties that can be present in the system of equations (which are not accounted for in the discrete Kalman filter).

The filter's simulation corroborated that the wrong implementation of, or even the lack of, the filter is enough to wrongly position the satellite and compromising its performance: this can lead to the input of a command that is not appropriate for the satellite's on time situation and its overall example (i.e. the use of the controller when this is not necessary). Nevertheless, the correct application results in a reliable and satisfactory estimation that practically eliminates the added noise.

Although both applications resulted in good estimations, the robust filter presented a more precise orbit reconstruction by linearizing the initial trajectory before actually applying the filter. In this thesis, the covariance matrices, Q and R, were considered constant along the process. In a more realistic context, these matrices vary, and were considered constant in this study in order to simplify and reduce the implementation time for both algorithms. For future works, it would be interesting to study how this variation along the application affects the two filters developed. Another intriguing possibility is reconstruction trajectories by combining two different variations of the Kalman filter, the robust and the adaptive estimation, into a new algorithm to check if it has a higher precision and the behaviour of the algorithm in time.

Lastly, both filters can be used for aeronautical and military applications.

Bibliography

[1] B. R. Elbert, *Introduction to Satellite Communication*, 3rd ed. Norwood MA: Artech House, 2008. xvii, 1, 5, 6, 7, 10, 11, 12, 13, 14

[2] M. Willis, K. T. Alfriend, and S. D'Amico, "Second-order solution for relative motion on eccentric orbits in curvilinear coordinates," *Proc. AAS/AIAA Astrodynamics Specialist Conference*, 2019. xvii, 18

[3] https://orbital-mechanics.space/_images/hohmann-transfer-orbit.svg, accessed: 14-Jun-2022. xvii, 22

[4] ESA - Space in Images, https://www.esa.int/ESA_Multimedia/Images/2004/10/The_Earth_s_gravity_field_geoid_as_it_will_be_seen_by_GOCE, Accessed: 18-Aug-2022. xvii, 23

[5] A. Mehmood and A. Mohammed, *Satellite Communications: Characterisation and Channel Modelling for Satellite Communication Systems*. Sweden: IntechOpen. xvii, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43

[6] G. Welch and G. Bishop, "An introduction to the kalman filter," *University of North Carolina*, Updated in 2006. xvii, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55

[7] A. K. Maini and V. Agrawal, *Satellite Technology: Principles and Applications*, 3rd ed. West Sussex: John Wiley Sons, 2014. xxi, 6, 7, 8, 9

[8] Orbiting Now, "Iridium/56730/2023/068w," https://orbit.ing-now.com/satellite/56730/2023-068w/iridium/, Accessed: 3-Jul-2023. xxi, 26

[9] L. J. Ippolito, *Satellite Communications Systems Engineering: Atmospheric Effects, Satellite Link Design and System Performance*, 2nd ed. West Sussex: John Wiley Sons, 2017. 4, 5, 8, 9, 10

[10] C. C. Chao, *Applied Orbit Mechanics*. California, USA: The Aerospace Press, 2005. 17, 21, 22

[11] H. Curtis, *Orbital Mechanics for Engineering Students*. Florida, USA: Elsevier Aerospace Engineering Series, 2005. 18, 23

[12] J. U. Park, “Orbital rendezvous using two-step sliding mode control,” *Aerospace Science and Technology, vol.3*. 19, 20

[13] J. R. Wertz, *Space Mission Engineering: The New SMAD, 2nd ed.* California, USA: Space Technology Library, 2011. 20, 22

[14] J. R. Wertz and W. J. Larson, *Space Mission Analysis and Design, 3rd ed.* California, USA: Space Technology Library, 1999. 23, 24

[15] M. Tropea and F. De Rango, *A Comprehensive Review of Channel Modeling for Land Mobile Satellite Communications*. Rende, Italy: Electronics, 2022. 28, 29, 34

[16] T. Chung, “What is a stochastic system,” Calthech, https://www.cds.caltech.edu/~murray/courses/cds101/fa02/faq/02-10-07_stochastic.html, accessed: 3-Dec-2022. 45

[17] P. S. Maybeck, *Stochastic Model, Estimation and Control*, vol. 1 ed. Ohio, USA: Academic Press, 1979. 45, 46, 48

[18] M. B. Rhudy, R. Salguero, and K. A. Holappa, “Kalman filtering tutorial for undergraduate students,” *International Journal of Computer Science Engineering Survey*, 2017, 1-18. 54

[19] L. Xie, Y. C. Soh, and C. E. de Souza, “Robust kalman filtering for uncertain discrete-time systems,” *IEEE Transactions on Automatic Control, Vol. 39*, 1994, pp. 1310-1314. 55

[20] W. Ding and C. Wang, J.and Rizos, “Improving adaptive kalman estimation in gps/ins,” *Integration the Journal of Navigation*, 2007, pp. 517-529. 55

[21] L. Xie and Y. C. Soh, “Robust kalman filtering for uncertain systems,” *Systems Control Letters, Vol. 22*, 1993, pp. 123-129. 56, 57, 58