Spacecraft Collision Avoidance

by

Charles Bussy-Virat

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Atmospheric, Oceanic and Space Sciences)
in The University of Michigan
2017

Doctoral Committee:

Professor Aaron J. Ridley, Chair
Senior Engineer Joel W. Getchius
Associate Professor Anouck Girard
Professor Mark B. Moldwin
Associate Professor Shasha Zou

To my sister and my parents.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor Prof. Aaron Ridley. On the very first week of my arrival in the US, you welcomed me and offered to do research with you. You've then shown a constant support until the very end of my Ph.D., by spending a tremendous amount of time teaching me about programming, about the ionosphere, and more generally about doing research. Your tireless efforts in correcting my scientific papers are proof of your dedication to helping me. Most of all, you gave me the freedom to work on the projects that I was the most interested in. You made this Ph.D. an exciting and unforgettable moment of my student life. For these few reasons, and so many more, I am deeply grateful to you.

I would like to show my immense gratitude to Joel Getchius for his strong support with the development of SpOCK. To put it in simple words, the research I conducted in collision avoidance would have probably not even happened without your help. But most of all, thank you for always being available to teach me about astrodynamics and sharing your great experience in this field.

I am sincerely grateful to Prof. Mark Moldwin. Mark, thank you for sharing with me your great experience and for guiding me during my Ph.D.. I think I can speak for all the CLASP students by saying that we all strongly appreciate your constant support, advice, and generosity to all of us in the department.

I would like to express my sincere thanks to Prof. Shasha Zou. Shasha, thank you for your insight and your expertise in thermosphere and ionosphere that greatly contributed to the manuscript.

I wish to thank Prof. Anouck Girard. Anouck, thank you for serving as the cognate member to my dissertation committee. Thank you too for your assistance in the writing of the thesis, and for sharing your expertise in orbital mechanics.

I would like to thank Prof. Chris Ruf for giving me the opportunity to work on the CYGNSS mission. It is a great honor for me to be part of the team. I am also grateful to all the members of the CYGNSS team, in particular Derek Starkenburg, Damen Provost, Linda Chadwick, Scott Gleason, Darren McKague, Derek Posselt, Kenny Carlsen, Mary Morris, Stephen Musko, Timothy Butler, Brian Gilchrist, Barbara Lupi, Randi Rose, and Kyle Nave.

Thank you to all the staff at CLASP. In particular, the IT personal who helped me with any kind of problems related to my computer. Thank you to Sandra Pytlinski, Jan Beltran, and Laura Hopkins for their considerable effort and help during my entire Ph.D..

Thank you to all my officemates with who I shared great times in the office. In particular thank you Jie Zhu, Xiangyun Zhang, Lizz Ultee (almost an officemate!), Lucas Luizzo, Garima Malhotra, Emily Judd, Sergio Vidal-Luengo, Zihan Wang, and Jiaen Ren for making the office a fun place to work in.

Thank you Tom Heine, first for your help in the 582 class project. Then, thank you for giving me the opportunity to contribute to the SCION proposal. Most of all, thank you for being a great friend, and for our very interesting discussions and great moments at the SmallSat conference. I wish we had gone to the conference earlier in our PhDs so we could have met sooner.

Thank you Nick Perlongo, Nathan Boll, and Casey Steuer for being such great friends. These years together are simply unforgettable. My only regrets are that they had to end so soon and that we ended up being spread all over the US so quickly. I have no doubt though that our friendships will go on and that so many more great times are to come. Casey, I promise I will always keep my French accent, and if by

some miracle I lose it, I'll make sure to always fake it for you. Nathan, thank you for constantly surprising me with your talents in French and your deep knowledge of our culture and gestures. I have no doubt that you are ready to come back to France, just don't forget the wine and cheese. Nick, your constant positive attitude and your presence since the very first day were an immense support. Your weird and always unexpected jokes made every single of my days at the office.

A huge thank to Nicolas Fougere and Ashley Kryscynski. Nico, you have been the longest time friend since I've arrived in Ann Arbor. I'm glad that you got to stay until the end of my Ph.D. and that we shared, and are going to share, so many great times together. You were my Nicois sunshine in my grey Parisian sky all these years. Moufle, your constant optimism, friendship, and unstoppable humor have been one of the most helpful supports of my years in Ann Arbor, and I can't wait to speak French with you.

Thank you to Lauren Stroner and Marine Besson for all the great weekends in Chicago. Having you guys a few hours away was one of the greatest supports of my stay in Ann Arbor. We set the bar really high for the next city!

Thank you to Kristin Miller. Kristin, of everything that I've lived during these years, our story is the best thing that happened to me. Thank you for being who you are, and thank you for being here every single and possible time I needed you. I can't imagine what I would have missed in my life if I had not gone to this party the very first week of my arrival in Ann Arbor.

Thank you to all my friends in France. In particular, thank you to my friends from Supaero (to name a few, Raphael d'Agostin, Maxime Maurice, Hugo Weissbart, Benjamin Godard, Josephine Wolff, Kevin Goburdhun), from Henri IV (Prepa, lycee and college), and from school.

Nathalie Royneau, Jules Moreau de Balasy, Alexandre Jegu, and Raphael Chen, thank you for being the proofs that friendships can last a lifetime. Distance, or any

other obstacle, will never weaken what I consider to be the best friendship stories of my life.

Finally, I'd like to thank my family for the constant support every day of my life. I've a particular thought for my grandparents, Isabelle, Bob, and Pierre for being always here for me. Vadim, thank you for coming to this world right on time to give me so many moments of pure happiness for the last year of my Ph.D.. Papa, Maman, Juliette, it's one of these rare situations where words are simply not strong enough. Papa, Maman, thank you for giving me the wings that allow me to now fly on my own. Juliette, thank you not only for always believing in my dream, but also for giving me the courage to pursue it. Thanks to you, it might one day become a reality.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF ABBREVIATIONS

ACPL Accepted Collision Probability Level

AGI Analytical Graphics, Inc

ANCAS Alfano-Negron Close Approach Software

CA Conjunction Assessment

CAESAR Conjunction Analysis and Evaluation Service, Alerts and Recommendations

CARA Conjunction Assessment Risk Analysis

CDM Conjunction Data Message

CHAMP CHAllenging Mini-satellite Payload

CIR Corotating Interaction Region

CME Coronal Mass Ejection

CNES Centre National d'Etudes Spatiales

CYGNSS CYclone Global Navigation Satellite System

DCA Distance of Close Approach

DDM Delay Doppler Map

DDMI Delay Doppler Mapping Instrument

ECEF Earth Centered Earth Fixed

ECI Earth Centered Inertial

EGM96 Earth Gravitational Model 1996

ESA European Space Agency

EUV Extreme UltraViolet

FDDM Full Delay Doppler Mapping

GEO Geosynchronous Earth Orbit

GITM Global Ionosphere Thermosphere Model

GMAT General Mission Analysis Tool

GPS Global Positioning System

GSFC Goddard Space Flight Center

GSOC German Space Operations Center

GUI Graphical User Interface

HAF Hakamada-Akasofu-Fry

HASDM High Accuracy Satellite Drag Model

HEO High Elliptical Orbit

HPOP High Precision Orbit Propagator

HSE High Speed Event

IMF Interplanetary Magnetic Field

ISS International Space Station

JAT Java Astrodynamics Toolkit

JSpOC Joint Space Operations Center

LEO Low Earth Orbit

LHCP Left-Hand Circular Polarization

LTAN Local Time of the Ascending Node

LVLH Local Horizontal Local Vertical

MEO Medium Earth Orbit

MOC Mission Operation Center

NASA National Aeronautics and Space Administration

NOAA National Oceanic and Atmospheric Administration

NPV Negative Predicted Value

NRLMSISE Naval Research Laboratory Mass Spectrometer Incoherent Scatter Radar Extended

NRMS Normalized Root Mean Squared

OD Orbit Determination

OSRA One Solar Rotation Ago

O/O Owner/Operator

P_c Probability of collision

PDF Probability Distribution Function

PPV Positive Predicted Value

RAAN Right Ascension of the Ascending Node

RK4 Fourth Order Runge-Kutta

RMS Root Mean Squared

SDA Space Data Association

SDO Space Debris Office

SIFT Storm Intersection Forecast Tool

SOC Science Operation Center

SpOCK Spacecraft Orbital Characterization Kit

SSA Space Situational Awareness

STK Satellite Took Kit

SWPC Space Weather Prediction Center

SWRI SouthWest Research Institute

TAD Traveling Atmospheric Disturbances

TCA Time of Close Approach

TIEGCM Thermosphere-Ionosphere Electrodynamics Global Circulation Model

TLE Two Line Elements

UTC Universal Time Coordinated

UV UltraViolet

WGS84 World Geodetic System 1984

WSA Wang-Sheeley-Arge

ABSTRACT

Spacecraft Collision Avoidance

by

Charles Bussy-Virat

Chair: Prof. Aaron J. Ridley

The rapid increase of the number of objects in orbit around the Earth poses a serious threat to operational spacecraft and astronauts. In order to effectively avoid collisions, mission operators need to assess the risk of collision between the satellite and any other object whose orbit is likely to approach its trajectory. Several algorithms predict the probability of collision but have limitations that impair the accuracy of the prediction. An important limitation is that uncertainties in the atmospheric density are usually not taken into account in the propagation of the covariance matrix from current epoch to closest approach time.

The Spacecraft Orbital Characterization Kit (SpOCK) was developed to accurately predict the positions and velocities of spacecraft. The central capability of SpOCK is a high accuracy numerical propagator of spacecraft orbits and computations of ancillary parameters. The numerical integration uses a comprehensive modeling of the dynamics of spacecraft in orbit that includes all the perturbing forces that a spacecraft is subject to in orbit. In particular, the atmospheric density is modeled by thermospheric models to allow for an accurate representation of the atmospheric drag. SpOCK predicts the probability of collision between two orbiting objects taking

into account the uncertainties in the atmospheric density. Monte Carlo procedures are used to perturb the initial position and velocity of the primary and secondary spacecraft from their covariance matrices. Developed in C, SpOCK supports parallelism to quickly assess the risk of collision so it can be used operationally in real time.

The upper atmosphere of the Earth is strongly driven by the solar activity. In particular, abrupt transitions from slow to fast solar wind cause important disturbances of the atmospheric density, hence of the drag acceleration that spacecraft are subject to. The Probability Distribution Function (PDF) model was developed to predict the solar wind speed five days in advance. In particular, the PDF model is able to predict rapid enhancements in the solar wind speed. It was found that 60% of the positive predictions were correct, while 91% of the negative predictions were correct, and 20% to 33% of the peaks in the speed were found by the model. Ensemble forecasts provide the forecasters with an estimation of the uncertainty in the prediction, which can be used to derive uncertainties in the atmospheric density and in the drag acceleration.

The dissertation then demonstrates that uncertainties in the atmospheric density result in large uncertainties in the prediction of the probability of collision. As an example, the effects of a geomagnetic storm on the probability of collision are illustrated. The research aims at providing tools and analyses that help understand and predict the effects of uncertainties in the atmospheric density on the probability of collision. The ultimate motivation is to support mission operators in making the correct decision with regard to a potential collision avoidance maneuver by providing an uncertainty on the prediction of the probability of collision instead of a single value. This approach can help avoid performing unnecessary costly maneuvers, while making sure that the risk of collision is fully evaluated.

CHAPTER I

Introduction

This dissertation focuses on predicting the risk of collision between operational spacecraft and objects in orbit around the Earth. More specifically, it studies the effects of uncertainties in the atmospheric density on the probability of collision.

1.1 Collision avoidance

1.1.1 History of collision avoidance

The population of objects in orbit around the Earth has dramatically increased in the past decade. In April 2005, the National Aeronautics and Space Administration (NASA) performed its first collision avoidance maneuver on a robotic spacecraft (Terra satellite). Two years later, the Chinese satellite Fengyun-1C was voluntarily destroyed, causing the largest increase in debris in space history ($\sim$ 3,000 objects larger than 10 cm). On February 10th, 2009, the collision between operational communication satellite Iridium 33 and retired Russian communication satellite Cosmos 2251 generated 2,000 debris larger than 10 cm and thousands smaller pieces at an altitude of 800 km. It is now quite common that the International Space Station (ISS) needs to be moved to avoid collisions. In 2015, four collision avoidance maneuvers and one "shelter-in-Soyuz" procedure were performed by the ISS (*Liou*, 2016; *Newman*, 2016). The "shelter-in-Soyuz" procedure is a safety protocol that requires the three

astronauts on-board to seek emergency shelter in the Soyuz capsule, docked to the space station, to prepare for a getaway.

There are currently 23,000 objects orbiting the Earth that are at least 10 cm wide. Among them, $\sim$ 5,000 are spacecraft (1,500 are still operational), $\sim$ 5,000 are mission related debris and rocket bodies, $\sim$ 13,000 are fragmentation debris (some of these resulted from the Iridium 33 collision with Cosmos 2251 and the destruction of Fengyun-1C) (*Liou*, 2016). In addition to these relatively big objects, it is estimated that an extra half million objects larger than 1 cm and 100 million objects larger than 1 mm are in orbit. Even submillimiter objects represent a threat for humans and satellites because of the very high speed that orbiting objects have relative to each other: the average relative velocity of satellites orbiting at a Low Earth Orbit (LEO) ($h \sim$ few hundred kilometers) is $\sim$10 km/s and the average relative velocity of satellites orbiting at the Geosynchronous Earth Orbit (GEO) ($h \sim 36,000$ km) is 0.5 km/s. Collisions with small untracked debris are also thought to cause many anomalies on satellites.

An important issue is that only objects greater than 10 cm are big enough to be tracked in LEO orbit and only objects greater than 1 m are tracked in GEO orbits. Passive techniques, like shielding, are used against smaller debris or against meteoroids that cannot be tracked either (*Council et al.*, 2011). For example, the ISS is shielded to resist debris smaller than 1 cm. Another concern is that debris stay in orbit for a very long time. Due to friction with the atmosphere, debris at relatively low altitude (under $\sim$ 700 km) re-enter the atmosphere, but sometimes only after decades. However, objects at higher altitude, particularly in GEO orbits, never re-enter the atmosphere and therefore, since launches continue, they accumulate with time. The GEO orbit regime is widely used for telecommunications and Earth observations, as well as space science. If no orbit maintenance maneuver is performed on a GEO satellite, gravitational perturbations from the Sun and the Moon make its inclination

oscillate around the Equator with an amplitude of 15° (*Allan and Cook*, 1964). This can pose a threat for operational GEO satellites that orbit at the Equator as the inactive satellites drift back through the GEO belt (*Lee and Yoola Hwang*, 2012; *Vallado and McClain*, 2007). A solution is to increase the altitude by a few hundreds km at the end of the spacecraft life, moving the satellite into a "graveyard orbit".

Figure 1.1 shows the evolution of the number of objects since the first satellite was launched in 1957. The dots represent space objects in LEO and GEO orbits (*Liou*, 2016). Note that only tracked objects cataloged in the Space Surveillance Network are shown, so a large number of small debris (< 10 cm) are not represented. The number of objects has been increasing dramatically over the last few decades: debris keep accumulating in orbit, dangerously increasing the probability of collision with operational satellites, and humans (*Liou*, 2016; *Council et al.*, 2011; *Aida et al.*, 2009; *Aida and Kirschner*, 2010; *Aida et al.*, 2015).

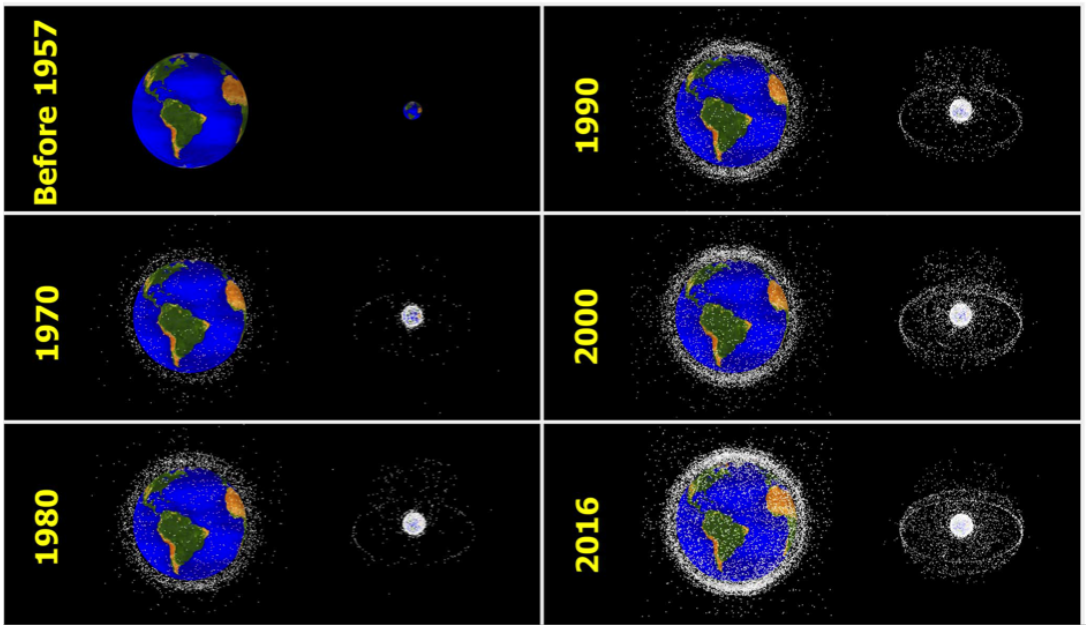

Figure 1.1: Evolution of the number of tracked objects since the first satellite in orbit in 1957. In each panel, the Earth on the left shows debris in LEO orbits, and the Earth on the right shows debris in GEO orbits (*Liou*, 2016).

In the next section, a review of the current methods used to avoid collisions

between operational satellites and objects in orbit is presented. First, the different steps followed by mission operators to avoid collisions are detailed. Then, the typical approach used to predict the probability of collision between two space objects is explained. Finally, examples of collision avoidance maneuvers that are commonly carried out are given.

1.1.2 Procedures for collision avoidance

Collision avoidance requires the knowledge of the position and velocity of all objects in orbit. Occasionally, satellite mission operators can keep track of their satellites with good accuracy using Global Positioning System (GPS) data, but the trajectories of all the other orbiting objects are much harder to follow. Space Situational Awareness (SSA) refers to the effort of gathering and updating the trajectories of natural and man-made orbiting objects (*Stoll et al.*, 2013). Two organizations support the SSA process worldwide: the Joint Space Operations Center (JSpOC) and the Space Data Association (SDA). SDA was established for satellite operators to share the most up-to-date satellite data (*Stoll et al.*, 2013). JSpOC is currently the single 24/7 global provider used in collision avoidance due to the accuracy and timeliness of the available information (*Aida et al.*, 2015).

Although the procedure for collision avoidance can vary between satellite operators, the overall scheme is similar, and is presented in this section.

1.1.2.1 Terminology

- State: vector that represents any kind of information about the satellite. It usually includes the position and velocity, and sometimes the drag and solar radiation coefficients used in the computation of atmospheric drag and solar radiation pressure accelerations.

- Covariance matrix: matrix that represents the uncertainties in the state, as well

as how the elements are correlated with each other. The i-th diagonal element corresponds to the variance of the i-th parameter (square of the standard deviation). The off-diagonal element i, j is proportional to the correlation coefficient between the i-th and the j-th elements.

- Primary spacecraft/satellite: the operational satellite the Owner/Operator (O/O) wants to prevent from collisions.
- Secondary object: any space object in orbit that can potentially encounter the primary spacecraft. Example: debris or operational satellite.
- Assessor: entity that assesses the risk of collision.

1.1.2.2 Procedure steps

The flow diagram in Figure 1.2 describes the overall process used in collision avoidance. There are three main steps in the collision avoidance process:

1. Conjunction Assessment (CA);
2. Risk analysis;
3. Mitigation.

Conjunction assessment

The goal of the conjunction assessment is to identify any possible close approach between the primary spacecraft and any space object in orbit. This can be an operational or a non-operational satellite or a piece of orbital debris. It requires the maintenance of a catalog of all space objects. JSpOC tracks more than 23,000 objects to evaluate any risk of collision with operating satellites (*Laporte*, 2014).

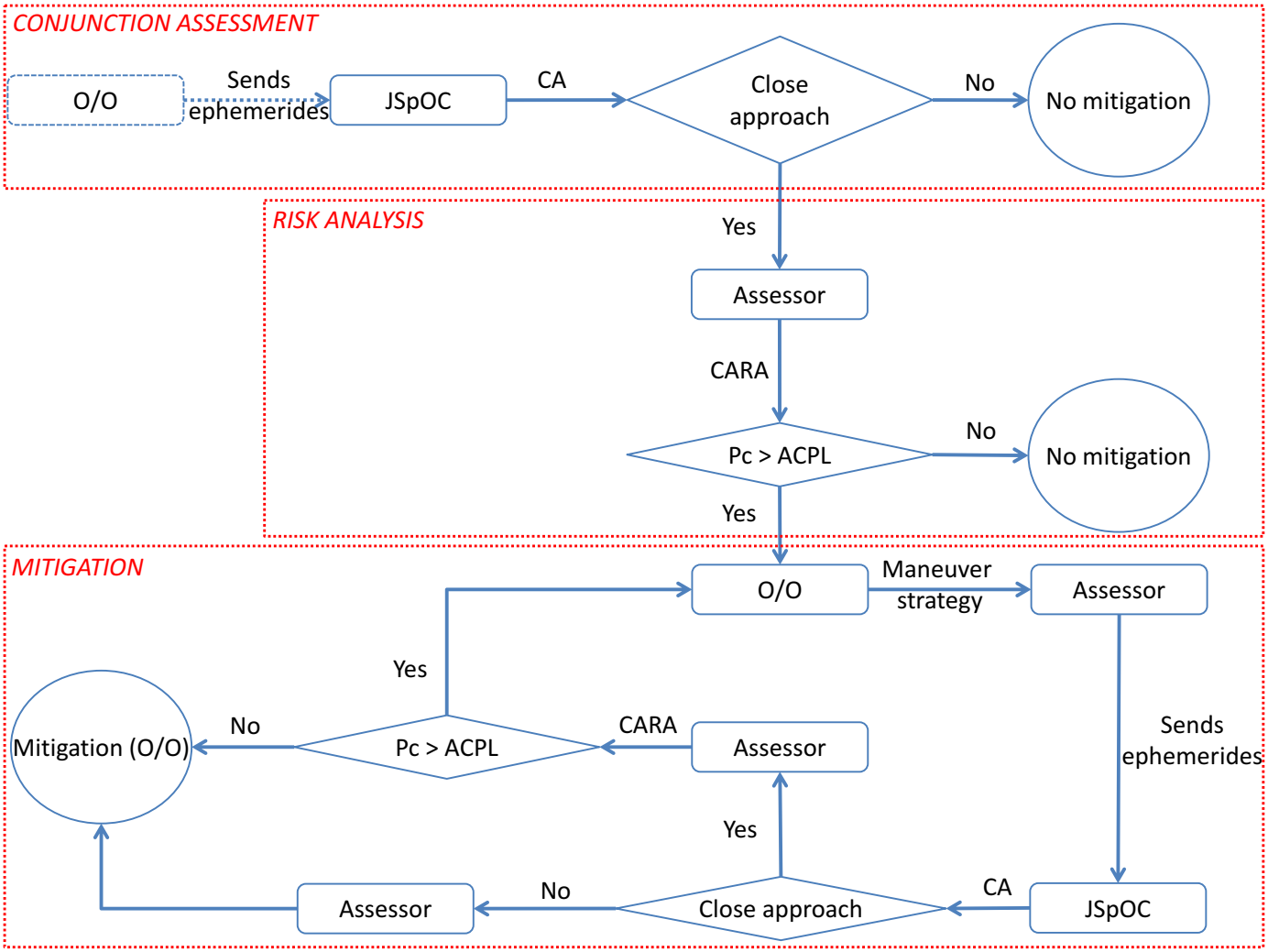

Figure 1.2: Flow diagram showing the procedures for predicting, assessing, and mitigating a risk of collision.

The ideal approach that an O/O would follow to identify any potential close approach with its satellite is to regularly send to JSpOC its latest ephemeris. This implies that the O/O has a tracking system, such as an optical telescope, a radar, or a GPS on-board the spacecraft. This is not always possible as not every mission has the capacity to track the spacecraft, in which case the O/O must rely solely on the tracking conducted by JSpOC. Therefore, this step is indicated as a dotted box in Figure 1.2.

From the information of the latest ephemeris, JSpOC looks for any close approach

with the cataloged objects. A close approach is defined as any distance between the primary and secondary object less than 1 km, and with a radial distance less than 200 m for LEO orbits (for Medium Earth Orbit (MEO) and GEO orbits the overall distance must be smaller than 10 km) (*Aida and Kirschner*, 2010; *Laporte*, 2014). JSpOC screens for any close approach in the following 7 days for LEO orbits and 10 days for GEO/High Elliptical Orbit (HEO) orbits (*Lauri K. Newman*, 2016). It first filters out objects that cannot possibly encounter the primary satellite by comparing the maximum perigee with the minimum apogee of the two orbits. If this distance is greater than a given threshold, then the secondary object does not represent any threat so it is ruled out. Otherwise, an analytical method is used to determine the closest point between the two elliptical orbits. An example of such method, the Alfano-Negron Close Approach Software (ANCAS), is presented by *Alfano and Thorne* (1994) and will be explained in a further section. If this distance is smaller than a given threshold, it generates ephemerides for both objects. It then considers a screening volume box around the primary satellite and looks for any penetration of this box by the secondary object, in which case a conjunction is possible to occur. JSpOC then generates a Conjunction Data Message (CDM) that includes the following information about the closest approach (*Hejduk and Frigm*, 2015):

- Time of Close Approach (TCA);
- States of both objects at TCA;
- Covariance matrices of both objects at TCA;
- Encounter information: miss distance, relative speed, TCA;
- Orbit Determination (OD) information (force model settings, OD quality).

The CDM is sent to the assessor to give alert of a possible encounter.

CDMs are usually not sufficient for the mission operator to make a decision regarding a potential collision avoidance maneuver. Indeed, CDMs do not provide any direct recommendation to perform a maneuver, do not take into account possible maintenance maneuvers the mission operator plans or have recently executed, and do not take into account the operational constraints of the mission operator (*Laporte*, 2014). More importantly, they do not provide any level of confidence in the prediction. Therefore, the next step of the collision avoidance process consists in determining the probability of collision associated with the close approach predicted by JSpOC.

Risk analysis

When the assessor receives a CDM from JSpOC, they assess the level of risk associated with the conjunction. This step is called the Conjunction Assessment Risk Analysis (CARA).

The entity that plays the role of the assessor depends on the satellite operator. NASA Goddard Space Flight Center (GSFC) provides CARA to $\sim$ 65 operational satellites including NASA unmanned operational assets, United State Government assets, and international partner assets (*Newman*, 2016). Not all mission operators use the services of NASA-GSFC to perform risk analysis. For example, RapidEye, a CubeSat constellation mission, uses the services from the European Space Agency (ESA)'s Space Debris Office (SDO) to evaluate the risk associated with the close approach (*Stoll et al.*, 2013, 2011). The French Space Agency, Centre National d'Etudes Spatiales (CNES), uses their own assessor Conjunction Analysis and Evaluation Service, Alerts and Recommendations (CAESAR) (*Laporte*, 2014). The German Space Operations Center (GSOC) also performs CARA for the eight satellites that they monitor (6 in LEO, 2 in GEO) (*Aida et al.*, 2009; *Aida and Kirschner*, 2010; *Aida et al.*, 2015).

To evaluate the risk, the assessor calculates the probability of collision P_c based on the covariance matrices and the states of both objects at TCA. There are several methods to calculate P_c. The common approach is presented here.

The covariance matrix of each of the two objects is represented by an ellipsoid, with an orientation given by the principal axes of the covariance matrix. These two ellipsoids are illustrated in Figure 1.3.

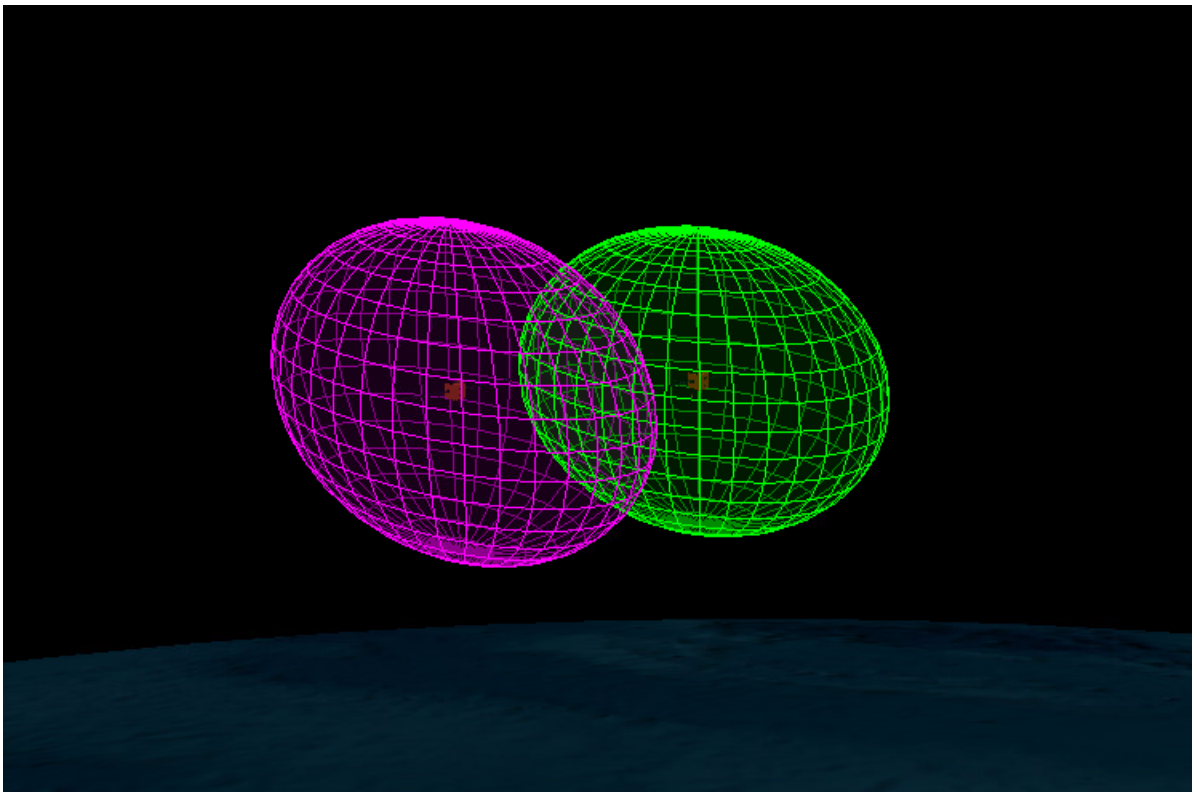

Figure 1.3: Ellipsoids surrounding the primary and secondary spacecraft to represent the uncertainties in the position of each object (STK - Analytical Graphics, Inc.).

The volume of intersection of the two ellipsoids represents the probability of collision of the two objects. Because computing the intersection volume implies calculating a three-dimensional integral, which can be computationally intense and complex, a few assumptions are usually made (*Sanchez-Ortiz et al.*, 2015):

- the orbital uncertainties of the primary object are not correlated with the orbital

uncertainties of the secondary object;

- the uncertainty in velocity is negligible;
- the uncertainty in position for both objects is constant during the encounter; and
- the relative motion can be considered rectilinear.

The last two assumptions are valid if the duration of the encounter is short enough (< 500 s) or equivalently if the relative velocity between the primary and secondary objects is high enough (> 10 m/s) (*Hejduk and Frigm*, 2015). Because the relative motion is assumed to be rectilinear, the 3D problem can be reduced to a 2D problem treated in the plane normal to the relative velocity vector, called the conjunction plane. Because the two covariance matrices are assumed to be uncorrelated, they can be combined in one covariance matrix: $C = C_1 + C_2$ (*Chan*, 2008). The combined covariance matrix is assumed to be both centered at the position of the secondary object, and constant during the encounter.

The position uncertainty is usually represented by a Gaussian three-dimensional probability function (*Sanchez-Ortiz et al.*, 2015; *Akella and Alfriend*, 2000; *Alfriend et al.*, 1999; *Foster*, 2001):

$$p_i(\vec{\delta r}) = \frac{1}{\sqrt{(2\pi)^3 \det(C_i)}} e^{-\frac{1}{2}\vec{\delta r}^T C_i^{-1} \vec{\delta r}} \tag{1.1}$$

where $\vec{\delta r}$ is the distance vector of the object i from its nominal position ($i = 1$ or 2), and C_i the covariance matrix associated with the uncertainties in the position of object i.

The probability P_c is then expressed as (*Sanchez-Ortiz et al.*, 2015):

$$P_c = \frac{1}{2\pi\sqrt{\det(C)}} \int_{-R}^{R} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} e^{-\frac{1}{2}\vec{r}^T C^{-1}\vec{r}} dz dx \tag{1.2}$$

where R is the sum of the two object radii, $\vec{r}$ the vector between the point of interest in the conjunction plane and each point (x, z) in the circle of radius R (integration area).

Another well known method to calculate the probability of collision relies on Monte Carlo simulations. This technique will be explained in more details in Chapter II.

Once the probability of collision is calculated, its value is compared against the defined Accepted Collision Probability Level (ACPL) (*Sanchez-Ortiz et al.*, 2015). This level represents the probability of collision threshold above which a maneuver should be performed. It is defined by the spacecraft operator and depends on the level of risk the mission is able to accept. A large value of ACPL involves a low number of maneuvers but it implies that the operator is ready to accept higher risks. On the other hand, if ACPL is set to a very low value, the number of maneuvers could be large. As explained in a further section, maneuvers require fuel so a large number of maneuvers could cost more fuel than the related fuel budget allocated for collision avoidance (*Sanchez-Ortiz et al.*, 2015). The value of ACPL also depends on the uncertainty of the orbital data. If the uncertainty is large, the probability of collision is less likely to get large values so a high value of ACPL (10^{-3} for example) is associated with a very low number of encounters. However, if the orbital data are more accurate, the probability of collision is more likely to reach higher values and therefore a high number of encounters could be predicted to occur (*Sanchez-Ortiz et al.*, 2015). Consequently, ACPL should be set to a lower value if the orbital data are not very accurate (e.g., from Two Line Elementss (TLEs)) than if it is accurate

(e.g., from CDMs), in which case a higher value of ACPL is often advised. Other factors that influence the optimum value of ACPL are the population at the altitude of the primary spacecraft, and the object size of the primary spacecraft.

Mitigation

If the assessor determines that the probability of collision is higher than the level of accepted risk, it alerts the O/O and recommends a maneuver (*Laporte*, 2014). The O/O plans a maneuver and sends its characteristics to the assessor, who then predicts the new orbit resulting from the maneuver and sends its prediction to JSpOC. Based on these new ephemerides, JSpOC screens the orbits against their internal catalog to check if the maneuver does not increase the risk of collision with other objects, and sends the result of the screening to the assessor. If the risk of collision is reduced by the maneuver (either no close approach has been detected by JSpOC or the new probability of collision is lower than ACPL), then the O/O is advised by the assessor to carry it out (*Stoll et al.*, 2013). Otherwise, the O/O needs to plan another maneuver strategy. One possible decision is to not perform any maneuver if no solution can be found to decrease the risk of collision (option not shown in Figure 1.2).

1.1.3 Maneuver strategies

The O/O is responsible for making the final maneuver decision and for executing it (*Newman*, 2016). A common maneuver is to increase the separation in the radial direction between the two objects by giving a thrust in the in-track direction half an orbit before the close approach (*Aida et al.*, 2015; *Lee and Yoola Hwang*, 2012). Table 1.1 (*Stoll et al.*, 2011; *Walter*, 2008) shows the effect of a thrusting on the orbital elements in the in-track, cross-track, and radial directions. A burn in any direction changes the orbit and can therefore be potentially used to increase the separation

Table 1.1: Effects of thrusting maneuvers in the three directions on the orbital elements (*Stoll et al.*, 2011; *Walter*, 2008).

Effects	In-track thrust	Radial thrust	Cross-track thrust
semi-major axis	X		
eccentricity	X	X	
inclination			X
RAAN			X
argument of perigee	X	X	X

distance between the two spacecraft. Although a combination of burns in all three directions is possible (*Stoll et al.*, 2011), in-track thrusting is usually recommended, for different reasons (*Chan*, 2002, 2008; *Aida and Kirschner*, 2010; *Aida et al.*, 2015). First, it is much simpler to execute operationally. Burns in the cross-track or radial directions are possible but the operation is generally more complicated to execute (*Chan*, 2002, 2008). A separation is achieved in a shorter period with a maneuver in the in-track direction than with a maneuver in the out-of-plane direction. Moreover, cross-track burns change the inclination and the Right Ascension of the Ascending Node (RAAN) of the orbit, which is disadvantageous for Sun-synchronous orbits. Another reason is that the orbit prediction is usually more accurate in the radial direction (*Aida and Kirschner*, 2010). Finally, cross-track thrusting use more fuel than in-track burns (*Stoll et al.*, 2011).

After the maneuver, a complementary maneuver is necessary to bring the satellite back to the nominal orbit, if needed. Indeed, the mission objectives often involve constraints on the orbit. This is particularly true for Sun-Synchronous orbits, repeat-groundtrack (for example with the Aqua, Aura, and Terra spacecraft (*McKinley*, 2008)), minimum altitude variation orbits, and for GEO orbits (*Vallado and McClain*, 2007).

Performing a maneuver is not always necessary, though. Maneuvers consume fuel, are risky, complicated, and can also be disruptive to mission operations and to zero gravity experiments such as the ones conducted in the ISS. An alternative is to change

the execution epoch or size of a regular orbit maneuver (*Aida and Kirschner*, 2010). This option is usually preferred as it is more fuel efficient, less complicated and less risky than operating an avoidance maneuver. However, it is not always available as the timing of the initially planned maneuver does not necessarily coincide with the time of the predicted close approach.

1.1.4 Challenges in predicting collisions for Low Earth Orbits

The altitude regime below about 500 km involves important challenges for collision avoidance. At these altitudes, the main force acting on spacecraft and debris is atmospheric drag (after gravity) and is a perturbing force particularly hard to model and predict.

The drag acceleration $\boldsymbol{a_{drag}}$ of a simple surface is represented by (*Vallado and McClain*, 2007):

$$\boldsymbol{a}_{\text{drag}} = -\frac{1}{2}\frac{C_D A}{m}\rho v_{rel}^2 \frac{\boldsymbol{v_{rel}}}{v_{rel}} \tag{1.3}$$

where C_D, A, and m are the drag coefficient, area projected towards the velocity vector and mass of the surface respectively, and $\boldsymbol{v_{rel}}$ is the satellite velocity with respect to the moving atmosphere of density ρ. Equation 4.2 demonstrates the dependence of the drag acceleration on the thermospheric density ρ. Under 500 km, the thermosphere is too dense to be neglected in accurate orbit calculations. It is the most complex parameter to estimate, the reasons of which are now detailed.

1.2 Challenge: predicting the atmospheric density

Although the neutral mass density at altitudes where LEO satellites orbit is small (a few grams per cubic km on average at 400 km), it is enough to create drag on orbiting objects and considerably modify their motion. This is an important consideration in the planning of the satellite mission, particularly in predicting the lifetime of the spacecraft and collision avoidance with other space objects (*Emmert*, 2015).

Many thermosphere models attempt to estimate the thermospheric density. The High Accuracy Satellite Drag Model (HASDM) (*Storz et al.*, 2005) was developed in 2002 by the Air Force Space Battlelab to improve satellite trajectory prediction accuracy by analyzing the effect of drag on trajectories of LEO satellites. It has been combined with empirical density model Jacchia-Bowman 2008 (*Bowman et al.*, 2008b), which is an improved version of the 2006 version that is based on Jacchia's diffusion equations, in the Jacchia-Bowman-HASDM 2009 (*D. Pachura*, 2016). This is currently used at JSpOC. Other examples of thermosphere models include the Naval Research Laboratory Mass Spectrometer Incoherent Scatter Radar Extended (NRLMSISE)-00 Model 2001 (*Picone et al.*, 2002), the Global Ionosphere Thermosphere Model (GITM) (*Ridley et al.*, 2006), and the Thermosphere-Ionosphere Electrodynamics Global Circulation Model (TIEGCM) (*Roble and Ridley*, 1994).

There are two main reasons why the thermospheric density is particularly hard to model. The first one is that the atmosphere between 100 and 500 km altitude is strongly coupled to the space environment. This system is very complex and the response of the density of this coupling is very challenging to estimate. The second reason is that the inputs to this coupled system, mainly linked to solar activity, are themselves quite difficult to predict. This double challenge makes it very hard to predict the density at LEO orbits, and hence the drag acceleration. In this section, these two reasons are described more fully.

1.2.1 Thermosphere coupling with the space environment

The thermosphere

The Earth's thermosphere is the upper layer of the atmosphere between 90 and about 600 km. Figure 1.4 shows the altitude profiles of the neutral densities in the thermosphere (*Emmert*, 2015). The primary constituents of the thermosphere are molecules of nitrogen (N_2), oxygen (O_2), and atomic oxygen (O) (*Schunk and Nagy*, 2009). The total neutral density drops exponentially with altitude, with a scale height that varies from 25 km (quiet solar activity) to 75 km (strong solar activity) (*Emmert*, 2015). This vertical variation is associated with zonal and meridional variations that vary with season and activity level. The global average density reaches maxima at the equinoxes and minima in July and January (*Emmert*, 2015). The amplitude of the semiannual (19%) and annual (13%) variations increase with height and solar activity (*Bowman et al.*, 2008a; *Emmert and Picone*, 2010; *Lei et al.*, 2012).

The neutral density is highly driven by the interaction of the thermosphere with the Sun, the magnetosphere, the ionosphere, and the lower atmosphere. These interactions result in the heating of the atmosphere, which causes it to expand and increase the density at LEO altitudes. This implies that to predict the thermospheric density, models not only need to predict the inputs to this strongly coupled system, but also need to accurately model the response of the thermosphere to these inputs. This is one of the main reasons why predicting atmospheric drag is so challenging. In this section, the couplings of the thermosphere with the Sun, the magnetosphere, the ionosphere and the lower atmosphere are first presented. A particular attention is given to the impact of strong solar events on the thermospheric density as they cause important disturbances by dramatically increasing the effects of these couplings.

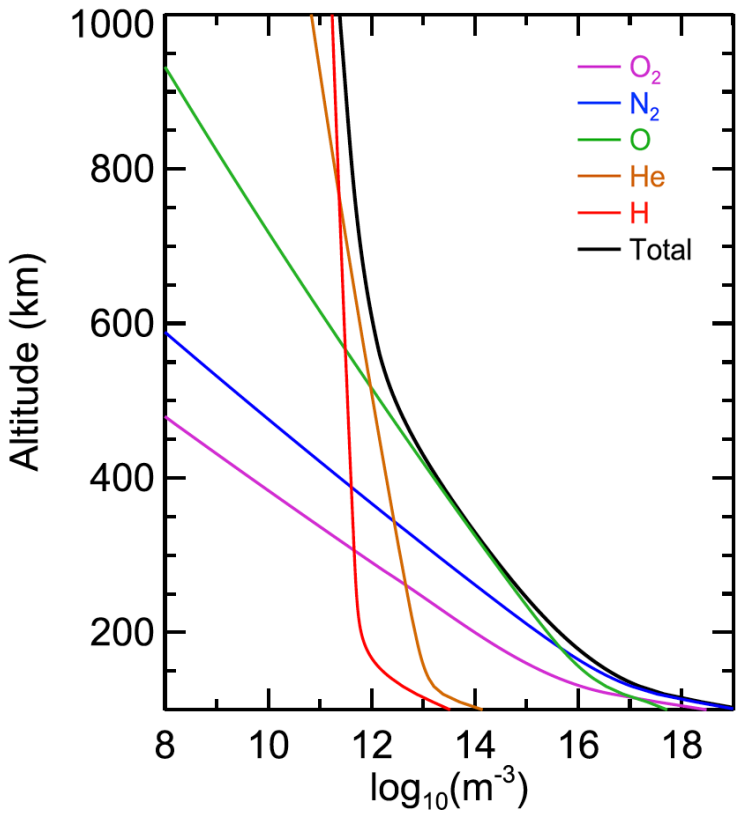

Figure 1.4: Density profiles of the main constituents of the thermosphere (*Emmert*, 2015).

1.2.1.1 Coupling with the Sun

The atmosphere is heated via the absorption of Extreme UltraViolet (EUV) radiation. The decrease in intensity, dI, of the photon flux, I, due to the absorption by a layer dz of the atmosphere is given as:

$$dI = -In_i(z)\sigma_i dz \tag{1.4}$$

where σ_i is the cross absorption cross section and $n_i(z)$ the neutral density of constituent i at the altitude z where the absorption occurs (*Schunk and Nagy*, 2009). The Sun is assumed to be at the zenith (90° overhead). The integration of Equation 1.4 from the top of the atmosphere to an altitude, z, results in a vertical neutral

density profile of constituent i expressed as (*Schunk and Nagy*, 2009):

$$n_i(z) = n_i(z_0)\frac{T(z_0)}{T(z)}\exp\left(\int_{z_0}^{z}\frac{dz'}{H_i(z')}\right) \tag{1.5}$$

where z_0 is an altitude reference, T the temperature of the atmosphere, and H_i the neutral gas scale height of constituent i. H_i is defined as (*Schunk and Nagy*, 2009):

$$H_i(z) = \frac{kT(z)}{m_i g(z)} \tag{1.6}$$

where k is the Boltzmann constant ($k = 1.38 \times 10^{-23}$ m^2 kg s^{-2} K^{-1}), m_i the mass of the neutral species i, and g the gravitational acceleration. One can show that the optical depth τ, defined as the integration of the density along a vertical column times the absorption cross section σ_i for wavelength λ ($\tau = \int_z^\infty \sum_i n_i(z)\sigma_i dz$) can be written as (*Schunk and Nagy*, 2009):

$$\tau(z) = \sum_i n_i(z)\sigma_i H_i(z_0) \tag{1.7}$$

In the case of a multispecies absorption of a radiation coming from a zenith angle θ, Equation 1.7 is generalized as:

$$\tau(z) = \frac{1}{\cos\theta}\sum_i n_i(z)\sigma_i H_i(z_0) \tag{1.8}$$

The optical depth is directly proportional to the scale height H_i, itself proportional to the temperature T. By absorbing the energy of the solar EUV radiation, the thermosphere is heated. Figure 1.5 shows the temperature profile of the thermosphere at solar maximum and solar minimum, derived from the density model NRLMSISE-00. Because the solar flux in the EUV and X-ray wavelengths increases during solar maximum, the thermosphere gets heated more at solar maximum than at solar minimum

by about a factor 2 (*Emmert*, 2015).

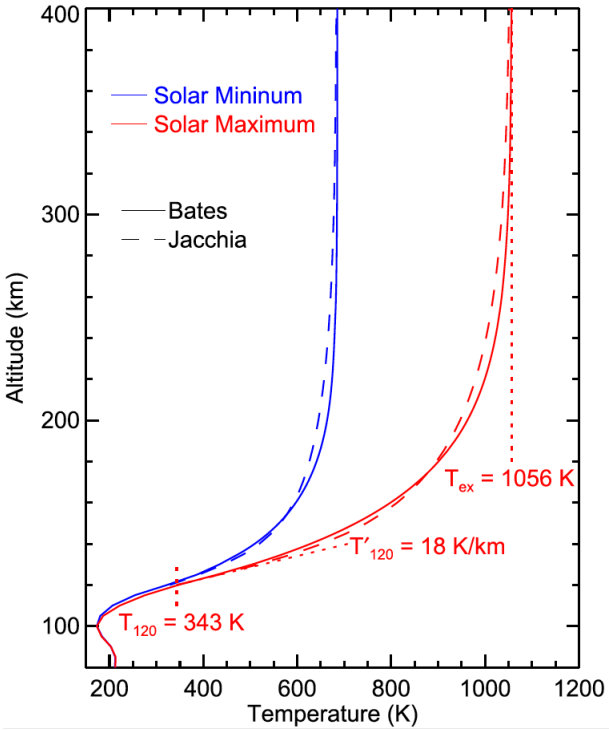

Figure 1.5: Temperature profiles at solar maximum and solar minimum modeled by NRLMSISE-00 (*Emmert*, 2015).

Another effect of solar EUV radiation is photoionization of neutrals. When a photon encounters a particle in the thermosphere (O, O_2, N_2), it can ionize it. The photoionization reaction can be written as: $X + h\nu \rightarrow X^+ + e^-$, where X is the molecule of the atmosphere being ionized, h is the Plank constant ($h = 6.63 \times 10^{-34}$ $m^2.kg.s^{-1}$), and ν is the frequency of the solar radiation. X^+ is the resulting ion from the photoionization and e^- is called a photoelectron. Photoelectrons are very energetic because the solar radiation typically has more energy than the ionization threshold of the particle. The excess of energy goes into kinetic energy of the photoelectrons that heat the thermosphere by collisions with the neutrals. Therefore, collisions of neutrals by photoelectrons are another source of thermospheric heating resulting from the solar EUV radiation. In addition, when the ion is lost, most often through dissociative recombinations, or $X_2^+ + e^- \rightarrow X + X$, energy is typically given

off and absorbed by the particles, which heats the thermosphere.

1.2.1.2 Coupling with ionosphere-magnetosphere system

Ionosphere

The ionosphere is the region of the upper atmosphere where the presence of electrons and ions become increasingly important. It is arranged in five layers (*Gombosi*, 1998). In the D region (60-90 km) and the E region (90-140 km), the dominant ions are O_2^+, NO^+, and N_2^+. Chemical reaction time scales are extremely short, such that the chemical source of ions, mainly photoionization due to EUV and X-rays, is balanced by losses due to dissociative recombinations. Collisions between charged particles are not important because the E region plasma is weakly ionized (the total ion density of 10^5 cm^{-3} and the neutral density is greater than 10^{11} cm^{-3}). In the F_1 region (140-200 km), the dominant ion is NO^+, with O^+ becoming dominant in the upper F_1 region ($>$ 180 km). The F_2 region (200-500 km) is where the ion density maximum occurs: at 300 km, the ion density peaks at $\sim 10^6$ cm^{-3} with a predominance of O^+ ions. This region is characterized by a balance between chemical loss processes and plasma transport, such as ambipolar diffusion and wind-induced drifts along the magnetic field lines, as well as electrodynamic drifts across the magnetic fields lines ($\boldsymbol{v_{drift}} \propto \boldsymbol{E} \times \boldsymbol{B}$). The dominant neutral is atomic oxygen O, with a density on the order of 10^8 cm^3. In the topside ionosphere ($>$ 500 km), the O^+ ion density drops and the contribution of the lighter helium and hydrogen ions increases (*Schunk and Nagy*, 2009; *Gombosi*, 1998).

The coupling between the thermosphere and the ionosphere has a strong impact on the thermospheric density. The ionosphere heats the thermosphere through collisions between ions and neutrals. Joule heating is due to the presence of electric currents

in the ionosphere that interact with the magnetospheric electric fields mapping to the ionosphere. It is another important source of heating, particularly during strong solar events, as it will be shown in a further section. Finally, exothermic chemical reactions are also one of the most important sources of heating of the thermosphere (*Richmond and Thayer*, 2000; *Vasyliūnas and Song*, 2005).

Magnetosphere

The magnetosphere is the region of space surrounding the Earth where the influence of its magnetic field controls the properties of charged particles. It extends to about 10 Earth radii in the sunward direction, and to a couple ofon hundred Earth radii in the antisunward direction (*Gombosi*, 1998). Because the solar wind (plasma material emanating from the solar corona) is supersonic and super Alfvenic, a shock forms on the dayside to slow down the solar wind and divert the charged particles around the magnetic field, called the bow shock. The magnetopause is the boundary between the magnetosphere and the shocked solar wind, and is defined as where the solar dynamic pressure balances the Earth's magnetic pressure. The region between the bow shock and the magnetopause is called the magnetosheath. The magnetic field lines on the night side stretch along the solar wind flow direction, forming the magnetotail. Figure 1.6 represents these different structures in the magnetosphere (*Gombosi*, 1998).

Although the Earth's magnetosphere acts like a shield against the solar wind, charged particles from the solar wind can still penetrate the magnetopause through magnetic reconnection on the day side. On the night side, solar wind particles can enter the magnetosphere and then also travel along magnetic field lines and collide with the upper atmosphere creating the auroral oval, which is a band a few degrees wide located about 20° equatorward from the poles during quiet times, moving

to lower latitudes during strong solar events (*Schunk and Nagy*, 2009). Figure 1.6 shows the solar wind particles that can penetrate the magnetopause and drift along the magnetic field lines (blue) to enter the Earth upper atmosphere in the auroral oval. This phenomenon, called particle precipitation, produces aurora by excitation of atmospheric particles, that then return to lower energy states by emitting light at altitudes between 100 and 250 km. Similar to the solar EUV ionization, particle precipitation deposits energy and causes ionization to occur. Particle precipitation is a direct source of heating by collision with neutrals and an indirect source of heating due to the resulting chemical heating and ion-neutral interactions, such as frictional heating and ion-neutral heat transfer (*Rees et al.*, 1983).

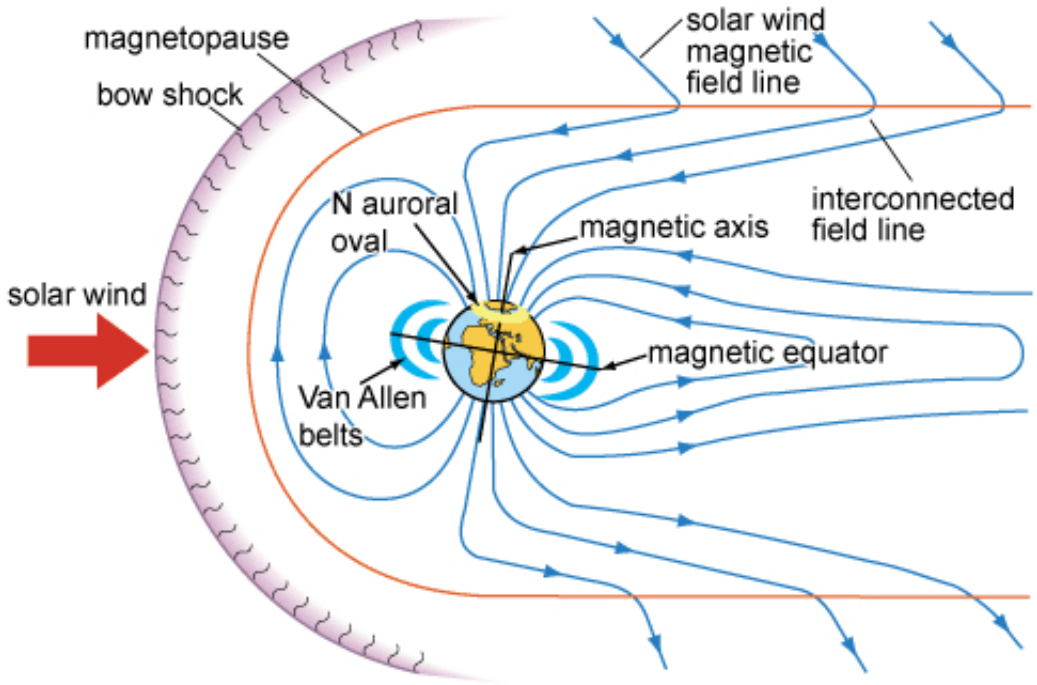

Figure 1.6: Structure of the Earth's magnetosphere and auroral oval where particle precipitation produces aurora and heats the upper atmosphere (`http://www.hk-phy.org/iq/aurora/aurora_e.html`).

1.2.1.3 Solar storms and substorms

Coronal mass ejections and solar flares

The Sun can eject large amounts of hot and ionized gas (10^{12} to 10^{13} kg) into the interplanetary medium, called Coronal Mass Ejections (CMEs). The ejection speeds range from 50 km/s to 2,000 km/s. In the regions from where the CMEs emanate, the Sun's magnetic field lines are closed, constraining the plasma from expanding outwards. These regions are usually found near the solar equator (*Hundhausen*, 1999; *Gombosi*, 1998).

The basic magnetic structure of the Sun is dipolar. The solar wind blows outward. When the solar wind moves along a magnetic field line of the Sun, it accelerates to faster speeds, but equatorial field lines inhibit the motion so the flow is slower near the equator. Sometimes, faster speed regions move to lower latitude and get mixed with slower speeds plasma. These interactions are called Corotating Interaction Regions (CIRs). Sometimes, the magnetic field traps plasma and the "bubble" expands greatly and can be explosively released, resulting in a flare (energy by light) and a CME (energy by plasma and magnetic field). Flares are typically in the wavelength of EUV and X-rays, which ultimately drive ion density and temperature change in the Earth's atmosphere. CMEs can have strong magnetic fields in the core (magnetic cloud) and form a shock in the front.

When a solar storm, a CME, or a CIR encounters the Earth, it creates a geomagnetic storm (*Tsurutani et al.*, 2006). In the next section, the mechanism of geomagnetic storms and their impact on the Earth's thermosphere are presented.

Impact of strong solar events on the thermosphere

The initial compression of the magnetopause is called the sudden commencement. When the Interplanetary Magnetic Field (IMF), an extension of the solar magnetic field into the interplanetary medium, has a large southward component it triggers

magnetic reconnection on the dayside of the magnetopause. Magnetic reconnection is a rearrangement of the Earth's magnetic field lines that results in a conversion of magnetic energy to thermal energy and particle acceleration. Magnetic reconnection at the dayside of the magnetopause enables the solar wind particles to penetrate into the Earth's magnetosphere, which ultimately increases the particle precipitation in the Earth's atmosphere. Moreover, the magnetospheric and ionospheric electric fields intensify, which increase the Joule and particle heating in the thermosphere. While solar EUV absorption is the main source of thermospheric heating during solar quiet times and is centered on the dayside at low latitudes, Joule heating dominates during strong solar events and is concentrated at high latitudes near the auroral zone. The heating that happens at high latitudes results in density changes that can be an order of magnitude in intensity and that spread rapidly across the globe, such that the density increases everywhere within 3-6 hours of the storm commencement. This intensification of energy input into the Earth's upper atmosphere corresponds to the main phase of the geomagnetic storm and lasts from a few hours to a day. The recovery phase, during which the thermospheric conditions go back to a steady state, can take several days. The thermosphere loses energy from radiative cooling by NO, whose density is enhanced during the increased auroral precipitation (*Gombosi*, 1998; *Schunk and Nagy*, 2009).

A number of publications has studied the variations in thermospheric density resulting from geomagnetic storms. For example, it was found that CIRs can cause density increases by 75% on average with periodicities of 9 and 13.5 days (*Lei et al.*, 2011). These periodicities are a result of the 27 day rotation of the Sun, which is consistent with the fact that CIRs emanate from coronal holes that rotate with the Sun. Figure 1.7 shows the density variations in daytime and at night resulting from a CIR hitting the Earth. The density increases by more than 75% in less than a day.

It then takes about a week to recover to its nominal value. *Liu et al.* (2012) showed that the density response to a CIR is larger during solar minimum, at nighttime, and in the summer.

Solar flares also drive density increases in the thermosphere. Using GITM, *Pawlowski and Ridley* (2011) showed that the dayside density response due to a solar flare depends linearly on the integrated input energy. A 15% density increase followed by the excitation of gravity waves resulting from two solar flares in 2003 and 2004 was reported in *Pawlowski and Ridley* (2008). A 30% to 60% increase at low to mid-latitudes in the density response to EUV solar flux enhancement resulting from a solar flare was reported in *Sutton et al.* (2006).

Forbes et al. (2005) analyzed the effect of three coronal mass ejections on the density of the thermosphere and found density perturbations of $\sim 20\%$ resulting from the constructive interferences of Traveling Atmospheric Disturbancess (TADs) (gravity waves generated by localized heating resulting from geomagnetic storms) at the equator. *Bruinsma et al.* (2006) and *Sutton et al.* (2005) studied the response to severe solar and geomagnetic storms. Both noted a rapid response of the thermospheric density ($\sim$ a few hours) at all latitudes with enhancements by more than three times the density. *Bruinsma and Forbes* (2007) calculated the density enhancements at the equator during sudden increases in geomagnetic activity and reported up to 800% density increases.

1.2.1.4 Conclusion

The thermosphere is strongly coupled to the surrounding space environment. It is coupled to the lower atmosphere through gravity waves and tides, to the ionosphere through Joule heating, chemical reactions, and collisions with ions, to the magnetosphere through particle precipitation and electrodynamic coupling, and to the Sun through EUV absorption and collisions with photoelectrons. When the solar activity

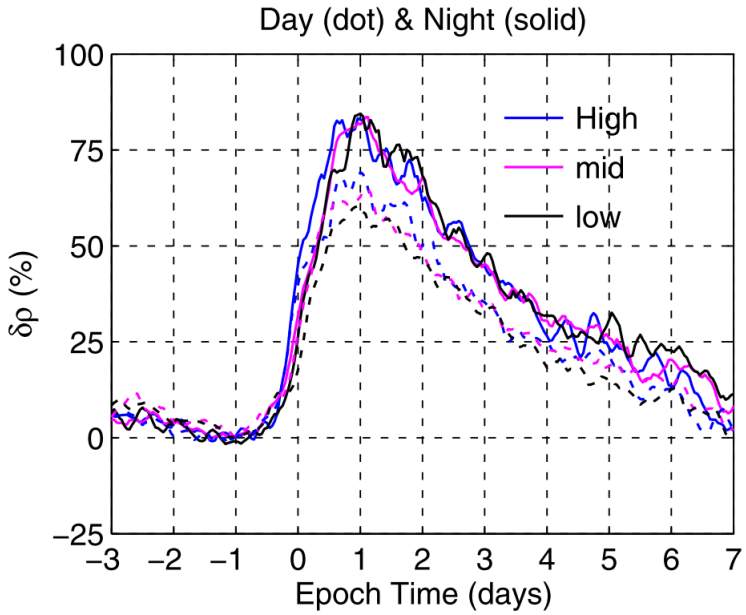

Figure 1.7: Density variations in daytime and at night resulting from a CIR hitting the Earth(*Lei et al.*, 2011).

is low, the absorption of solar EUV is the main source of heating of the thermosphere and the main cause of density variations. During strong solar events, the enhancement of Joule heating resulting from an increase of the ionospheric and magnetospheric electric currents, and the intensification of particle precipitation from the solar wind along magnetic field lines mapping into the thermosphere at high latitudes, become dominant and can highly perturb the thermospheric density. This is why the response of the thermospheric density to the space environment is very complex, and the main reason why atmospheric drag is hard to model. Additionally, the inputs to this coupled system are themselves hard to predict.

1.2.2 Predictions of the solar activity

The space environment of the Earth is mainly driven by solar activity, more specifically, the solar flux, the solar wind velocity and IMF (mainly the B_z component, perpendicular to the ecliptic plane). Space weather refers to the study of

these parameters as well as their impact on the Earth's magnetosphere, ionosphere, and thermosphere. One of the basic problems in space weather forecasting is that the drivers (solar wind speed and IMF) are only measured about 1 h before they affect the environment. In order to allow for adequate planning for some members of the commercial, military, or civilian communities, reliable long-term space weather forecasts are needed (*Wright et al.*, 1995; *Bussy-Virat and Ridley*, 2014).

Direct measurements of the drivers are not always possible. For example, since the upper atmosphere absorbs EUV radiation, measurements of solar EUV can not be made from the ground, but only from rockets and satellites. Instead, proxies are used to model such drivers, such as F10.7. This is the solar radio flux at a wavelength of 10.7 cm and it is commonly used to model EUV irradiance (*Emmert*, 2015). Moreover, the perturbations of the Earth's magnetic field resulting from its interaction with the IMF and the solar wind particles are described by magnetic activity indices such as Kp and Ap. These indices can be derived from magnetic perturbations directly measured from ground stations on Earth. The current models that predict these proxies are first presented. The multiple models that have been developed to predict the solar wind velocity will then be reviewed.

The B_z component plays an important role in auroral activity but, because of its very large variability, it is very hard to predict. There is no model that is currently able to forecast it with a reasonable accuracy.

1.2.2.1 Predictions of F10.7 and Ap

The rotation period of the Sun depends on the latitude: it is 25 days at the equator and 35 days at the poles. On average, it is about 27.2 days. Consequently, the same active regions of the Sun are directed toward the Earth roughly every 27 days. This implies that a 27 day periodicity should also be observed in the solar radiation flux.

On a larger time scale, the 11 year change in the Sun's magnetic activity induces a 11 year periodicity in the EUV flux. Figure 1.8 (left) shows the variations of a 27 day average of F10.7 from 1995 until 2016. The 11 year periodicity is evident, with solar minima in 1997 and 2008 ($\overline{F_{10.7}}_{27d} \sim 70$), and solar maxima in 2002 ($\overline{F_{10.7}}_{27d} > 200$) and 2013 $\overline{F_{10.7}}_{27d} \sim 150$. Figure 1.8 (right) demonstrates the 27 day periodicity of the daily average F10.7 over five months in 2014.

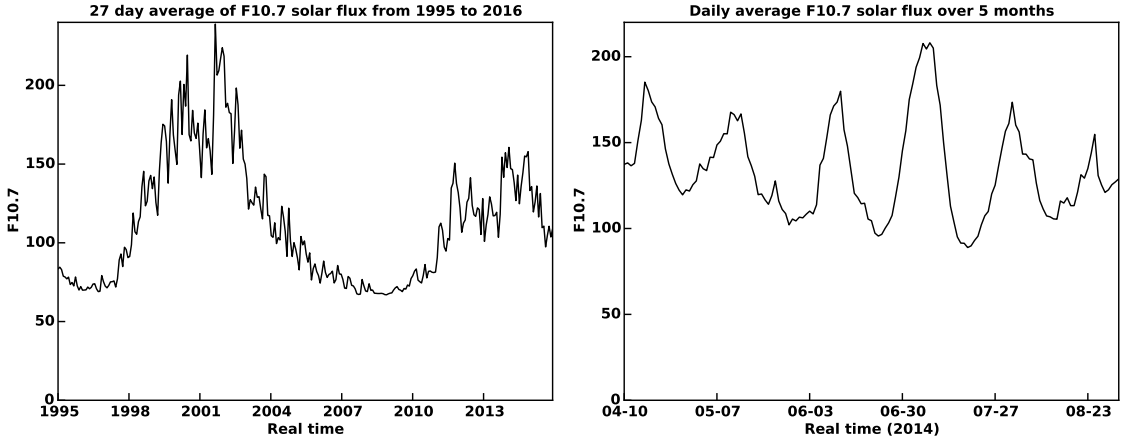

Figure 1.8: 27 day average of F10.7 solar flux from 1995 to 2016 (left); daily average F10.7 solar flux over 5 months (right) (Omniweb data).

Many models exist to predict the solar flux F10.7, and the geomagnetic indices Kp and Ap.

The Space Weather Prediction Center (SWPC) at the National Oceanic and Atmospheric Administration (NOAA) provides a 27-day forecast of the F10.7 radio flux and the geomagnetic indices Kp and Ap (`http://www.swpc.noaa.gov/`). These predictions are based on the persistence of patterns in the solar and geomagnetic activity from one solar rotation to the next.

The SOLAR2000 model (*Tobiska et al.*, 2000; *Tobiska*, 2003; *Tobiska and Bouwer*, 2006) uses different techniques to predict the F10.7 flux depending on the time scale of the predictions. Neural-network algorithms are used for short term forecast (1-72

hour). Interplanetary hydrogen backscatter measurements provide a proxy of solar far-side Lyman that can be used to forecast solar irradiance for 14-28 day forecasts. For 1-6 month forecasts, Fast Fourier Transforms of non-stationary periodicities are used to estimate the solar irradiance. Finally, statistical methods predict the F10.7 radiation flux for very long time scales (several years).

Other models include the SVR model (*Huang et al.*, 2009), which uses neural network techniques, and the Support Vector Machine (SVM) (*He et al.*, 2008). Linear-regression techniques (*McNish and Lincoln*, 1949) use relationships between F10.7 and the sunspot numbers. Precursor models have shown the best performance for long term predictions (several years), such as the solar dynamo model (*Schatten et al.*, 1978).

1.2.2.2 Prediction of the solar wind speed

Many predictive empirical and physics-based solar wind speed models have been created, most of them coronal or heliospheric. One of the most commonly used is the Wang-Sheeley-Arge (WSA) Model, currently used at the SWPC at NOAA. *Wang and N.R. Sheeley* (1992) found a correlation between the solar wind speed and the inverse of the divergence rate of the magnetic field in the corona. Based on this relationship, they created a model to predict the solar wind speed by extrapolating observations of the photospheric magnetic field into the corona. *Arge and Pizzo* (2000) improved and tested its performance by comparing the predictions to data collected by the Wind satellite for 3 years (1994-1997), in particular for the 1996 solar minimum. The average fractional deviation and the correlation were found to be equal to 0.15 and 0.4, respectively. *Arge et al.* (2003) proposed a new relationship for the solar wind speed as a function of the magnetic field expansion factor of open coronal field lines and of the minimum angular separation at the photosphere between an open field foot point and its nearest coronal-hole boundary. Another longer-term validation study

conducted by *Owens et al.* (2005) showed that the root-mean-square error between the model and the data is better at solar maximum than at solar minimum. However, the large-scale structure is better predicted during solar minimum than during solar maximum. The WSA Model has also been compared in previous studies to the Persistence Model (which keeps the velocity constant at its current value). *MacNeice* (2009a) validated the ability of the WSA Model to forecast specific solar events, while *MacNeice* (2009b) showed that the WSA Model provides better predictions than the Persistence Model clearly only after 2 days.

The Hakamada-Akasofu-Fry (HAF) Model (currently used at the Air Force Weather Agency) is a solar wind speed prediction model. *Fry et al.* (2007) presented results of the solar wind forecast based on this model. It particularly distinguished simulations of the ambient solar wind and simulations of event-driven solar wind (for example, CMEs). Predictions of shocks following solar events such as CMEs are detailed in *Fry et al.* (2001) and *Smith et al.* (2004). *Norquist and Meeks* (2010) compared the predictions between the WSA and the HAF Models (5 day forecasts over 6 years of Solar Cycle 23). The WSA gave slightly better predictions than the HAF Model for the speed of the solar wind. For instance, correlations between the model and the data are slightly better for the WSA Model (0.42) than for the HAF Model (0.32). However, both models under-represent the temporal variability of speed of the solar wind: the standard deviation is smaller in the models than it is in the data, by about 15%.

Toth and Odstrcil (1996) conducted a comparison of methods for simulations of MHD Models, one of which was the ENLIL Model, which enables numerical modeling of solar wind structures and disturbances in 3-D (*Odstrcil*, 2003). An idea proposed in the past was to couple models. For instance, the ENLIL heliospheric

model was also coupled to the Magnetohydrodynamics Around Sphere (MAS) coronal model (also called the CORHEL model) to predict solar wind parameters based on solar and coronal structures (*Odstrcil et al.*, 2004). *Owens et al.* (2008) also made a comprehensive study of the coupling of kinematic, empirical, and MHD Models. The MAS/ENLIL and WSA/ENLIL Models were compared in *Lee et al.* (2009) to measurements from the ACE satellite during the declining phase of Solar Cycle 23. However, this was a study of a more scientific nature, and it did not test the forecasting ability. In particular, it revealed a good agreement for the general large-scale solar wind structures but not for the CME or shock associated with active regions.

Other methods based on observations of the chromosphere have been created. For instance, a hybrid intelligent system uses magnetic field observations and combines the potential field model and an artificial neural network to give prediction of the daily solar wind speed (*Wintoft and Lundstedt*, 1997). The correlation between predictions from the model and the data between 1976 and 1994 varies from 0.2 to 0.5, depending on the year. *Robbins et al.* (2006) presented another model to predict solar wind speed related to geomagnetic events. The model was based on the location and the size of coronal holes. It differs from the WSA Model in particular because it does not need a full magnetic synoptic map but only the image of one coronal hole to give predictions. The linear correlation was 0.38 for the 11 years of comparison between the model and data. The Pch method (*Luo et al.*, 2008) was based on a correlation between the speed and the brightness of the solar EUV emission (characterizing the brightness and the area of a coronal hole). *Leamon and McIntosh* (2007) also presented predictions based on the structure of the chromosphere.

Another method, the support-vector-regression algorithm, was applied in *Liu et al.* (2007) to predict the value of solar wind speed. The comparison of the predictions to

the data shows very good results, but the predictions are only 1 to 3 h ahead of real time.

Finally, *Owens et al.* (2013) made a comprehensive study of the 27 day periodicity of the solar wind parameters and presented a possible way to make predictions based on this periodicity. In particular, they showed that the correlation between two solar rotations is very good for the speed during solar minimum or the declining phase of a solar cycle. However, this is not the case during solar maximum, where no clear correlation was found. The study also explained how such a model can represent a benchmark for other space weather forecast models.

1.2.3 Summary

In this section, the two main reasons why the atmospheric density is hard to predict have been explained. First, the thermosphere, where LEO satellites orbit, is a very complex system coupled to its surrounding environment: the lower atmosphere, the ionosphere, the magnetosphere, and the Sun. The response of the density to this coupling is not perfectly known so thermospheric models can only determine the density at the position of the satellite with a limited accuracy, usually around 80%. Moreover, the inputs of this coupling system that are mainly linked to the solar activity are really hard to predict. Strong solar events such as CMEs and solar flares, can completely disturb the thermosphere and cause large and abrupt variations of the density, resulting in important perturbations of the drag acceleration. For these reasons, the most complex and challenging factor to predict and model in the drag equation is the atmospheric density.

1.3 Contributions of the thesis

Assessing the risk of collision between two spacecraft in the next days involves the prediction of their positions and velocities. In that regard, the Spacecraft Orbital Characterization Kit (SpOCK) was developed. Spacecraft are subject to perturbing forces such as gravity, atmospheric drag, third and fourth body gravitational perturbations (mainly from the Sun and the Moon), Sun and Earth radiation pressure, and tides. The numerical integration of SpOCK software uses a comprehensive modeling of the dynamics of spacecraft in orbit by taking into account these perturbing forces. Specifically, the non-spherical portion of the mass distribution of the Earth is modeled with a decomposition of spherical harmonics for the gravitational potential. Moreover, SpOCK uses thermospheric models (NRLMSISE-00 and GITM) to derive the atmospheric density at the position of the spacecraft, allowing for an accurate representation of the atmospheric drag. The fundamental function of any orbit propagator is the accurate prediction of orbits, but SpOCK also includes a variety of functionalities that can be used for mission design and analysis. For example, it allows the computation of solar power and the coverage of ground stations. SpOCK supports parallelism and is therefore well suited for ensemble, Monte Carlo, or satellite constellation analysis. The algorithm, as well as a validation and demonstrations of the different functionalities are presented in Chapter II.

Uncertainties in the atmospheric drag can result in important errors in the predictions of the positions and velocities of satellites orbiting at low altitude (< 600 km), causing a real challenge for collision avoidance. These uncertainties are mainly induced by uncertainties in the neutral density of the upper atmosphere. As this region of the atmosphere is strongly coupled with the Sun, predicting the solar activity, and being able to quantify the uncertainties on that forecast, are important steps of the overall process of determining the drag acceleration. The Probability Distribution Function (PDF) model, a data based empirical model, was developed in

that regard, and is described in Chapter III. Although focusing on the prediction of the solar wind speed, the technique can be generalized to the prediction of the solar flux, the Interplanetary Magnetic Field, the proton density and temperature. The PDF model predicts the solar wind speed five days in advance and the uncertainty interval associated with the main prediction. Ensemble forecasts provide the forecasters with an estimation of the uncertainty in the prediction and with an estimation of the possible spread in the characteristics. Transitions from slow to fast solar wind represent a particular concern because they result in the strongest disturbances of the atmospheric density, hence the drag acceleration. These transitions are observed as peaks in the solar wind speed. The PDF model predicts several features of the peak in speed associated with fast transitions, including the amplitude of the peak as well as the time when the maximum occurs. A single value for the prediction of the magnitude of the speed and the time when the increase occurs is associated with an estimation of the error on the peak prediction. Specifically, the uncertainty on the predicted amplitude and in the time of the peak provides the forecast community with an interval of uncertainty on the prediction. The first version of the PDF model focused on predictions of the background solar wind speed, which corresponds to times when the solar activity is quiet or moderate. This version is introduced in the first section of Chapter III. The improvement of the PDF model for predictions of such transitions led to the second version of the model that is presented in the second section of Chapter III. In particular, comparisons of the predictions with the observations made by the Advanced Composition Explorer satellite are made, as well as comparison of the accuracy of the model with the previous studies that predicted such transitions, in particular those made by the WSA model (*MacNeice*, 2009a,b; *Owens et al.*, 2005, 2008; *Emmons et al.*, 2013).

Collision risk assessment is usually performed without taking into account the uncertainties in the atmospheric density. However, neglecting the large uncertainties

in the forecast of the solar activity and in the coupling of the upper atmosphere with the Earth's space environment (i.e., the ionosphere and the magnetosphere) can result in important errors in the prediction of the probability of collision. In particular, strong and abrupt solar events such as CMEs, solar flares, and CIRs can completely change the outcome of a prediction on the risk of a collision. SpOCK offers an approach of taking into account these considerations for the calculation of the probability of collision. The algorithm, presented in Chapter IV, relies on Monte Carlo procedures that allow for an accurate prediction of the probability of collision. Uncertainties in the predictions of the solar flux F10.7 and the geomagnetic index Ap are included in the process so that a correct modeling of the uncertainties in the forecast of the atmospheric density allows for a more realistic determination of the probability of collision. In particular, the effects of these uncertainties on the risk of collision are presented. The effects of geomagnetic storms on the probability of collision are then illustrated with an example of a parallel conjunction between two spacecraft two days after epoch. On March 17th, 2015, a strong storm hit the upper atmosphere, with values of the geomagnetic index Ap exceeding 200. Density enhancements by almost 50% at 400 km strongly increased the drag, modifying the orbits of the satellites. The effects on the risk of collision are presented. The algorithm developed can be used in real-time, and can provide mission operators with a better estimation of the risk of collision.

CHAPTER II

Orbit propagation

The Spacecraft Orbital Characterization Kit (SpOCK) is presented. This software enables mission analysts to accurately model the motion of spacecraft in orbit. It also allows the computation of solar power, the coverage of ground stations, and the assessment of risk of collision between two space objects. SpOCK supports parallelism and is therefore well suited for ensemble, Monte Carlo, or satellite constellation analysis. The algorithm, a validation and demonstrations of the different functionalities are presented. Chapter IV will introduce in detail the collision risk assessment algorithm and some of its applications.

2.1 Introduction

Space mission design and analysis requires the use of software tools to adequately address the desired requirements of a particular mission. A key element of such tools is the ability to design and analyze the candidate orbits and the spacecraft's environment along those orbits. While there exist several capable commercial off the shelf software solutions (Satellite Took Kit (STK) from Analytical Graphics, Inc (AGI), and Free Flyer from a.i. Solutions), and even a few open source ones (Java Astrodynamics Toolkit (JAT) (http://jat.sourceforge.net/), and the General Mission Analysis Tool (GMAT) (http://opensource.gsfc.nasa.gov/projects/GMAT/index.php)), such solutions may not be adequate for all interested parties. Civilian, defense, and scientific requirements in such readily available software are as variable as the number of missions serving each of these fields. These readily available solutions address the requirements of their largest user base (justifiably so), sometimes at the expense of specialized, infrequent users. Other issues one may encounter with such software packages include the need to obtain expensive licenses or extensions (plug-ins), lack of customization, and lack of cross-platform support.

Therefore, SpOCK, a feature rich mission design tool created specifically to address the needs of space and atmospheric physicists was developed. SpOCK, which is distributed in C, is better described as a software framework that enables physicists to study past, present, and future space missions as necessary. The central capability of SpOCK is a high accuracy numerical propagator of spacecraft orbits and computations of ancillary parameters. The modular code base allows users to quickly and efficiently implement new features and enhance current capabilities as well as write simulations for any number of scenarios. SpOCK supports parallelism and is therefore well suited for ensemble, Monte Carlo, or satellite constellation analysis.

In this chapter, numerical methods are reviewed and the specific capabilities of SpOCK are described in further details. We then describe a few examples of analyses

that have been performed using SpOCK are provided, including the examination of coverage of tropical storms by the CYclone Global Navigation Satellite System (CYGNSS). Finally, we review some validation data of SpOCK. SpOCK is validated by comparing different computations with the STK from AGI.

2.2 Methodology

In this section, the algorithm used to integrate the spacecraft trajectory in SpOCK is presented. Then, a description of the different features and their implementation is given.

2.2.1 Numerical methodology

The fundamental function of any satellite mission analysis tool is the accurate prediction of orbits. Spacecraft are subject to perturbing forces such as gravity, atmospheric drag, third and fourth body gravitational perturbations (mainly from the Sun and the Moon), Sun and Earth radiation pressure, and tides. Gravity is clearly the largest force that the satellites encounter, while the importance of the rest is quite dependent on the altitude of the satellite. Once the forces are summed, the acceleration is determined. From this, the velocity and position are derived through numerical integration using modern numerical techniques.

2.2.1.1 Dynamic model: perturbing forces

Gravity model

The gravity itself is a complex force to model. The Earth is not a perfect sphere so the distribution of its mass is non-spherical. This implies that the gravitational force varies as a function of the position. A good approach that takes into account the non-spherical portion of the mass distribution of the Earth is to represent the gravitational potential (U) as a decomposition of spherical harmonics (*Vallado and McClain*, 2007):

$$U = \frac{\mu}{r}\left[1 + \sum_{l=2}^{\infty}\sum_{m=0}^{l}\left(\frac{R_{\oplus}}{r}\right)^{l} P_{l,m}\left[\sin\left(\phi_{gc_{sat}}\right)\right]\left\{C_{l,m}\cos\left(m\lambda_{sat}\right) + S_{l,m}\sin\left(m\lambda_{sat}\right)\right\}\right] \tag{2.1}$$

where r is the distance of the satellite from the center of the Earth, λ_{sat} and $\phi_{gc_{sat}}$

are the longitude and geocentric latitude of the satellite, $P_{l,m}$ are the Legendre functions, $C_{l,m}$ and $S_{l,m}$ are the gravitational coefficients, l and m are the degree and order of the decomposition, and μ and $R_\oplus$ are the gravitational parameter ($\mu = 398,600.442km^3/s^2$) and mean equatorial radius of the Earth ($= 6,378.137km$), as defined in the World Geodetic System 1984 (WGS84) (*National Imagery and Mapping Agency*, 2000). The coefficients $C_{l,m}$ and $S_{l,m}$ used in the propagator are taken from the Earth Gravitational Model 1996 (EGM96) (*Lemoine et al.*, 1997). The gravitational potential of a simple sphere is the first term of the right hand side ($\frac{\mu}{r}$). The double summation describes the perturbation due to the non-spherical distribution of the Earth mass. This spherical harmonic decomposition breaks into three categories (*Vallado and McClain*, 2007):

- zonal harmonics, that represent bands of latitude ($m = 0$);
- sectorial harmonics, that represent bands of longitude ($l = m$);
- tesseral harmonics, that represent tiles ($l \neq m \neq 0$).

The strongest perturbation is due to the first zonal harmonic $C_{2,0}$, more commonly noted J_2 ($J_2 = 0.0010826267$), which corresponds to the equatorial bulge due to the Earth's rotation causing it to be oblate.

Atmospheric drag

For Low Earth Orbit (LEO) satellites below $\sim$ 500 km, atmospheric drag is the main force acting on a satellite, after gravity, and cannot be neglected for accurate orbit calculations. The drag acceleration $\boldsymbol{a_{drag}}$ of a simple surface is represented by (*Vallado and McClain*, 2007):

$$\boldsymbol{a}_{\text{drag}} = -\frac{1}{2}\frac{C_D A}{m}\rho v_{rel}^2 \frac{\boldsymbol{v_{rel}}}{v_{rel}} \tag{2.2}$$

where C_D, A, and m are the drag coefficient, area projected towards the velocity vector (discussed below) and mass of the surface respectively, and $\boldsymbol{v_{rel}}$ is the satellite velocity with respect to the moving atmosphere of density ρ.

Atmospheric drag is hard to model, mainly because the thermospheric density is very complex to model. It is driven by many phenomema that are themselves hard to predict, such as the solar X-Ray and EUV fluxes, aurora, high latitude Joule heating, which are often proxied by indices such as F10.7, Ap, Kp, and Hemispheric Power. Even if the drivers are well known, deriving the density is still a challenge for the scientific community. Many thermosphere models attempt to model it. The High Accuracy Satellite Drag Model (*Storz et al.*, 2005) was developed in 2002 by the Air Force Space Battlelab to improve satellite trajectory prediction accuracy, by analyzing the effect of drag on trajectories of LEO satellites. It has been combined to the empirical density model Jacchia-Bowman 2008 (*Bowman et al.*, 2008b), an improved version of the Jacchia-Bowman 2006 model that is based on Jacchia's diffusion equations, in the Jacchia-Bowman-HASDM 2009 (*Newman et al.*, 2015), and is currently used at the Joint Space Operations Center (JSpOC). Other examples of thermosphere models include the Naval Research Laboratory Mass Spectrometer Incoherent Scatter Radar Extended (NRLMSISE) model (*Picone et al.*, 2002), the Global Ionosphere Thermosphere Model (GITM) (*Ridley et al.*, 2006), and the Thermosphere-Ionosphere Electrodynamics Global Circulation Model (TIEGCM) (*Roble and Ridley*, 1994).

Third body perturbation

Gravitational perturbations by a third body (Sun and Moon) are also modeled in the propagator. The accelerations due to these perturbations can be written as (*Vallado and McClain*, 2007):

$$\boldsymbol{a}_{\text{3rd body}} = \mu_3 \left(\frac{\boldsymbol{r}_{\text{sat, 3}}}{r^3_{\text{sat, 3}}} - \frac{\boldsymbol{r}_{\text{Earth, 3}}}{r^3_{\text{Earth, 3}}} \right) \tag{2.3}$$

where μ_3 is the gravitational parameter of the third body (Sun or Moon), $\boldsymbol{r}_{\text{sat, 3}}$ the vector from the satellite to the third body, and $\boldsymbol{r}_{\text{Earth, 3}}$ the vector from the Earth to the third body. The first term represents the gravitational force of the third body on the satellite. The second term, called the indirect effect, corresponds to the gravitational force of the third body on the Earth (*Vallado and McClain*, 2007). The National Aeronautics and Space Administration (NASA) SPICE Toolkit (`naif.jpl.nasa.gov`) is used to calculate the position of the Sun and the Moon at every time step of the propagation.

Solar radiation pressure

Finally, the Sun's radiation pressure is taken into account in the propagator with the acceleration described in Equation 2.4 (*Wyatt*, 1961).

$$\boldsymbol{a}_{\text{srp}} = -\frac{C_r A}{m} \frac{L_{\text{Sun}}}{4 * \pi * c * r^3_{\text{sat, Sun}}} \boldsymbol{r}_{\text{sat, Sun}} \tag{2.4}$$

where A is the cross sectional area as seen by the Sun (discussed below), m the mass of the satellite, C_r the solar radiation coefficient, L_{Sun} the luminosity of the Sun ($L_{\text{Sun}} = 3.823 \times 10^{26}$ W), c the speed of light ($c = 299,792.458$ km/s), and $\boldsymbol{r}_{\text{sat, Sun}}$ the vector from the satellite to the Sun.

Combining forces

For each perturbing force computed independently, Figure 2.1 shows the error made on the position after one day of propagation by neglecting the different forces at different altitudes. For example, the black circles represent the difference in position after one day between a propagation modeling the atmospheric drag in a two body spherical gravitational potential and a propagation only modeling a two body spherical gravitational potential, without atmospheric drag. A similar approach was

followed for the solar radiation pressure (magenta diamonds), the third body perturbations from the Moon and the Sun (green stars), and the main perturbing gravitational effect due to the Earth equatorial bulge (coefficient J2, blue triangles). A propagation modeling the gravity potential as a spherical harmonics decomposition of order and degree 20 was compared to the propagation modeling only the equatorial bulge. The difference in position between both propagation represents the error made by neglecting terms of order greater than 2 in the spherical harmonics decomposition presented in Equation 2.1. It is represented by the red squares in Figure 2.1.

As the altitude increases, the error made by neglecting the atmospheric drag decreases. For altitudes higher than 1,000 km, the thermospheric density is so low that the drag can be neglected, since the error is approximately smaller than the round off error on the position. The error made by neglecting the equatorial bulge (blue triangles) decreases with the altitude, as does the error made by neglecting the higher order spherical harmonics (red squares). However, both the errors made by neglecting the third body (green stars) and solar radiation pressure (magenta diamonds) increase with altitude. Since the distance to the Sun on the day side decreases as the altitude increases, the solar radiation pressure effects get more important. The third body effects get also more important with the altitude because the term of amplitude $r_{\text{sat, 3}}^{-2}$ in Equation 2.3 varies more as the satellite orbits the Earth. Overall, the most important term is the equatorial bulge of the Earth.

Note that these results depend highly on many factors, such as the cross sectional area with respect to the Sun and with respect to the velocity, the mass of the spacecraft, the solar activity, the drag, and solar radiation pressure coefficients. Therefore, the relative contribution of each perturbing force changes as well. For example, if the area/mass ratio increases, the solar radiation pressure increases and can have more effect than the third body perturbation (the ratio used in Figure 2.1 is 0.01 m^2/kg). However, the variation of the error with the altitude stays the same as in Figure 2.1.

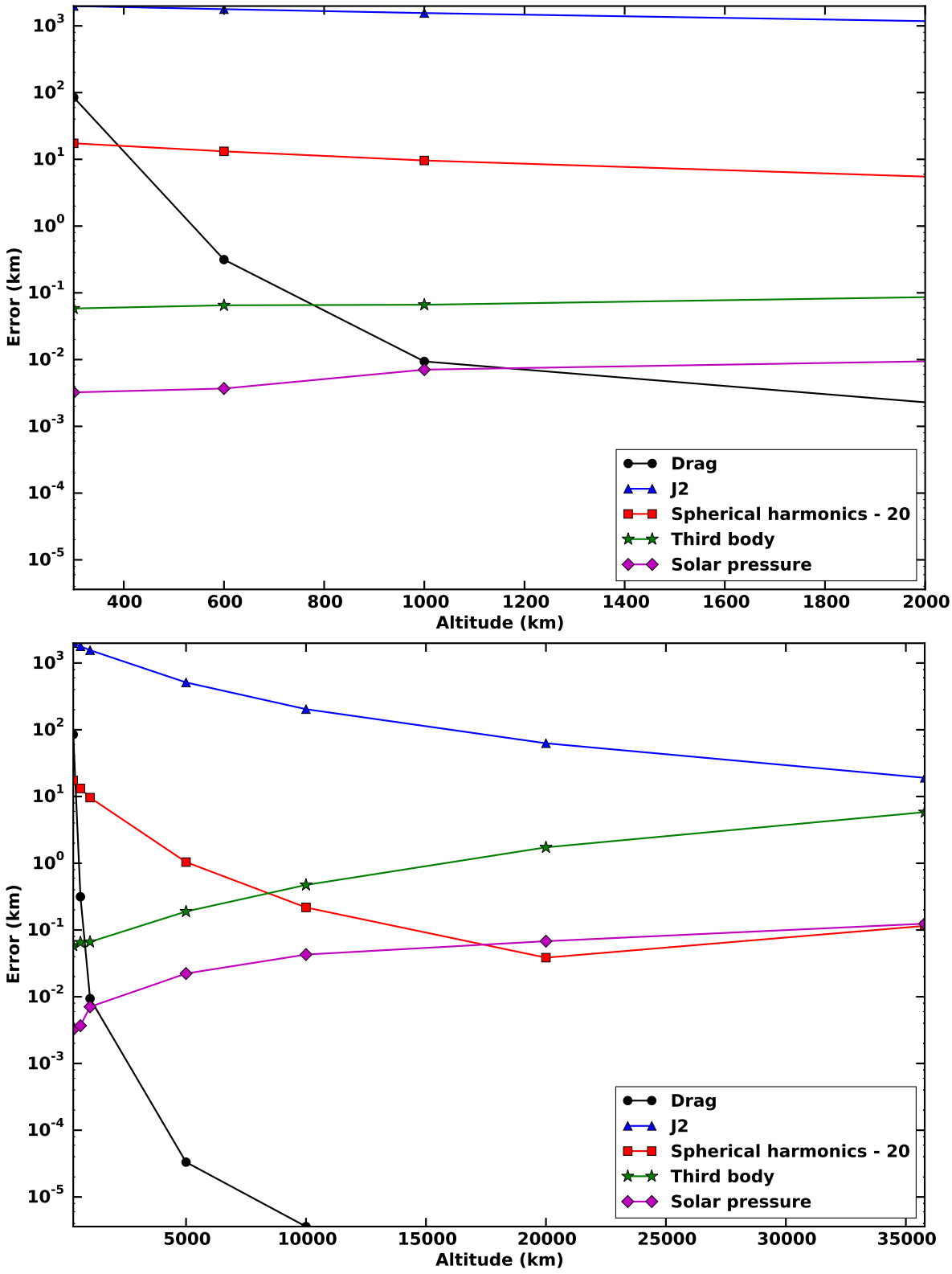

Figure 2.1: Position error after one day as a function of the altitude for different perturbing forces. Top is from 300 to 2000 km, bottom is from 300 to 36,000 km.

2.2.1.2 Numerical integration

The flow diagram in Figure 2.2 shows the structure of SpOCK's algorithm. Here, the integration block is presented (center of the diagram). The inputs and outputs of SpOCK are presented in a further section.

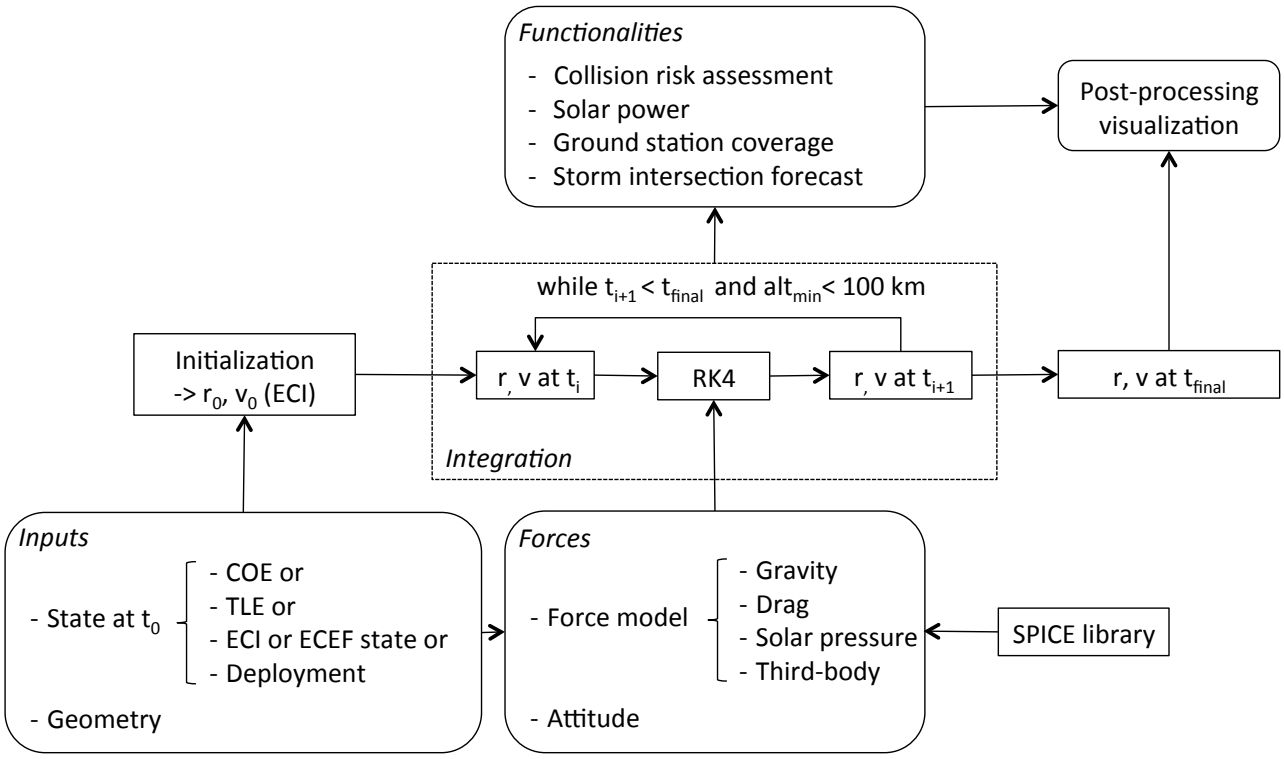

Figure 2.2: Flow diagram of SpOCK and its functionalities.

The total acceleration computed at each time step of the simulation is the sum of the accelerations due to gravity, drag, third body perturbations, and solar radiation pressure. The propagator then uses a Fourth Order Runge-Kutta method with a fixed time step to integrate the acceleration at each time step of the simulation. Once the simulation reaches the final epoch or the spacecraft reaches an altitude below 100 km,

implying reentry into the Earth's atmosphere, the propagation stops. The time step is chosen by the user at initialization. Figure 2.3 shows the error made on the position as a function of the integrator step size. SpOCK was run for one day with a circular orbit at 5,000 km with different time steps. The position after one day was compared with the position calculated with a time step of one second. The error varied linearly in logarithmic scale, which reflects an exponential increase in error with the time step. The error increases to hundreds and thousands of kilometers as the time step reaches several minutes. Errors due to neglecting the different perturbing forces reported in Figure 2.1 are also shown (with a time step of 10 s for reference). Neglecting any of the perturbing forces results in a higher error than the error caused by a time step of 10 s compared to a time step of 1 s. This underlines the impact of perturbing forces on the dynamics, even atmospheric drag, which is very weak at 5,000 km because of the very low atmospheric density.

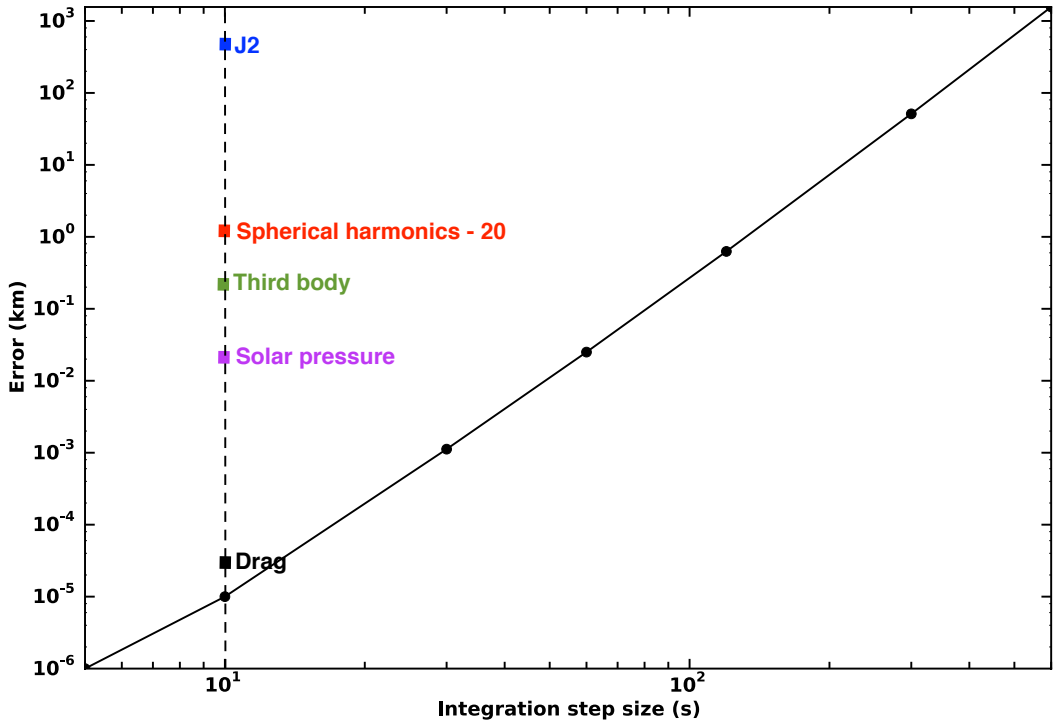

Figure 2.3: Position error after one day as a function of the integration step size for a 5,000 km orbit. Also shown is the error from neglecting perturbing forces in the same orbit, taking a 10 seconds time step. Note that errors for time steps smaller than 1 s are not represented here. As the step size approaches 0, the error is expected to increase because of truncation effects.

2.2.2 Functionalities of SpOCK

The flow diagram in Figure 2.2 illustrates the functionalities of SpOCK (top center), as well as the inputs the user needs to set up to initialize the program (bottom left). In this section, these will be introduced first. Then, the different features will be presented.

2.2.2.1 Initialization of SpOCK

There are several inputs the user needs to set up for a simulation:

- initial state: position and velocity of the spacecraft at the start epoch;

- geometry: a description of the geometry of the spacecraft;
- attitude: the orientation of the spacecraft during the propagation; and
- force model: which forces to include in the dynamics.

This section summarizes the different options the user has to implement each of these inputs.

Orbit initialization

To initialize the orbit, the user has different options. The most simple one is to use osculating orbital elements: apogee altitude, inclination, argument of perigee, Right Ascension of the Ascending Node (RAAN), true anomaly, and eccentricity. The user can also initialize the orbit with Two Line Elements (TLE). One can also give the initial state (position and velocity) in the Earth Centered Earth Fixed (ECEF) reference system, or in the Earth Centered Inertial (ECI) reference system.

Sometimes satellites are deployed from a single launch vehicle, such as what is done for secondary payloads, satellites from the International Space Station (ISS), or constellation missions such as CYGNSS. Satellites are ejected from a same position with different speed and angle, which slightly changes the parameters of their orbit (in particular the inclination and the semi-major axis). For mission planning and collision avoidance, it is important to be able to know what the relative motion is between spacecraft. SpOCK is able to initialize the position and the velocity of the satellites from a deployment module, by specifying the initial position of the module (with osculating orbital elements, that transformed in SpOCK to initial positions and velocities in ECI coordinates) and the angle and speed of ejection of each satellite in the Local Horizontal Local Vertical (LVLH) reference system.

Spacecraft geometry and attitude

To accurately model the perturbing forces (atmospheric drag and solar radiation pressure) and compute the power generated by the solar panels, the geometry and the attitude of the spacecraft need to be accurately characterized in the simulation.

To represent the geometry of the spacecraft, the satellite is divided up into surfaces in an input file. For each surface, the following are specified:

- the normal to the surface expressed in the spacecraft body reference system;
- the area of the surface;
- the total solar cell area of the surface;
- the drag coefficient and the standard deviation around this value if ensembles of the drag coefficient are run; and
- the solar radiation coefficient.

The user has different options to characterize the attitude of the spacecraft. The satellite can have one of its faces be constantly nadir or Sun pointing, or the satellite can be rotating at a constant rate in the body reference frame. Moreover, since a constant cross section area is not relevant for a spacecraft with a varying attitude, SpOCK allows input of the attitude as a function of time and it determines the cross section area in the ram direction for drag, and in the Sun direction for power and solar pressure. As expressed in Equations 4.2 and 2.4, the drag and solar radiation pressure accelerations can be accurately modeled by SpOCK with the specification of the geometry and the attitude of the spacecraft.

Force model

As explained in section 2.2.1.1, several perturbing forces need to be considered in the propagation to accurately model the trajectory of the spacecraft.

The propagator allows a decomposition of the gravitational potential with an order and degree of 360, guaranteeing an accurate representation of the non-spherical portion of the mass distribution of the Earth. Within the input file, the user can select the maximum order of gravity to include. The higher the order, the more accurate the gravitational potential. Figure 2.4 (top) shows the maximum difference in position over one day of propagation for a LEO orbit at 300 km, between a simulation that uses an order N and a reference simulation using order 40. N varied from 2 (only including the equatorial bulge represented by the J2 term), to 40. The case with a perfect spherical potential (i.e., neglecting the J2 term) was not considered as it leads to significantly larger errors in position than an aspherical Earth. One can notice that the difference in position is very large when considering only the J2 term (> 16 km) compared to higher orders, for which the error drops to values well below 1 km. Although the general trend of the error is to decrease as the order increases, one can notice the non-linear effect of increasing the order of the gravity model.

Modeling more orders is computationally intensive and the user may not need the accuracy for the particular task. Figure 2.4 (bottom) shows the run time as a function of the gravity model order. This figure demonstrates the rapid increase of computational time with the level of accuracy of the aspherical potential. While the error only decreases slightly for orders greater than 30, the run time increases to values up to 4 times higher for order above 30. Therefore, the user needs to make a compromise between accuracy and computational time. For most applications, orders below 20 are sufficient, particularly for satellites orbiting at higher altitudes because the influence of the asphericity of the Earth decreases, as shown below.

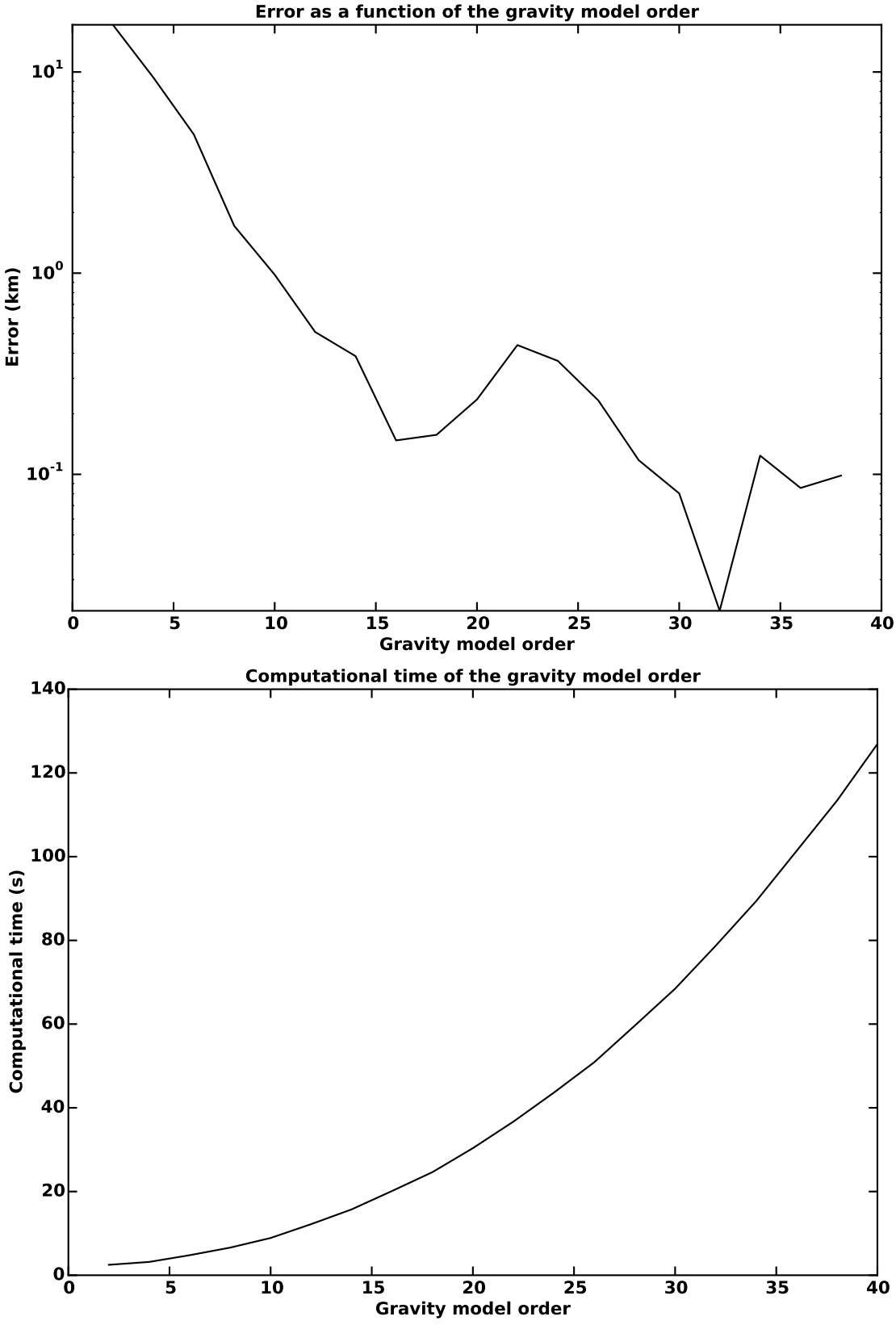

Figure 2.4: Position error (top) and computational time (bottom) after one day of propagation as a function of the order of the gravitational potential spherical harmonics decomposition used.

For LEO orbits, the atmospheric drag is also a perturbing force that is relatively difficult to model accurately. Because the thermospheric density is highly variable, representing it with a constant value in Equation 4.2 leads to important errors when modeling the drag acceleration. The propagator offers the possibility to use either the NRLMSISE or the GITM model to calculate the density at the satellite position, under different driving conditions. If the thermospheric NRLMSISE is used, the solar flux F10.7 and geomagnetic index Ap can be specified as constant values during the propagation, or as time-varying values. For this latter option, SpOCK downloads the data from Omniweb (`http://omniweb.gsfc.nasa.gov/`). If the simulation includes the future, predictions from the Space Weather Prediction Center (SWPC) from the National Oceanic and Atmospheric Administration (NOAA) are used. The users can also input their own file with the values of F10.7 and Ap as a function of time. The NRLMSISE density model is called at each time step of the propagation using the time-varying values of F10.7 and Ap. This allows the propagator to accurately model the density, and consequently the drag acceleration.

Each of these perturbing forces (drag, solar radiation pressure, Sun and Moon gravitational perturbation) can be easily turned on or off independently. This provides users with the ability to rapidly run a coarse simulation to explore mission planning or run with a highly accurate determination of the position for collision avoidance calculations.

2.2.2.2 Features

As with any propagator, SpOCK predicts the positions of satellites. It also has other features, such as computing the solar power generated by solar panels, predicting the position of specular points between two satellites (Earth surface reflection points of the signal from a transmitting satellite to a target satellite), assessing the coverage by ground stations, and assessing the risk of collision with another object in orbit.

This section describes some of these features. The collision risk assessment tool will be presented in detail in Chapter IV.

Solar power

A crucial component of any satellite mission is determining the amount of power generated by solar panels. SpOCK has a feature to allow this, given a user specified geometry of the solar panels in the body frame of reference, as well as the efficiency of the solar cells. The propagator calculates the solar power SP generated by these surfaces as a function of time when the satellite is in sunlight:

$$SP = \sum_{i=1}^{N_s} f A_i \eta_i \cos\phi_i \tag{2.5}$$

where i is the index for the solar panel with an area A_i, N_s the total number of solar panels, f the solar flux ($f = 1358$ W/m^2), ϕ_i the angle between the normal to the solar panel and the direction of the Sun, and η_i the solar cell efficiency. Examples of solar power computations are presented in Section 2.3.3.

Specular points

A specular point is the location on the surface of the Earth of the reflection of a signal sent by one satellite to a target satellite, as shown in Figure 2.5. As a remote sensing technique, this can be used to calculate sea surface heights, ionospheric total electron content, or the root mean squared slope of the ocean that can then be related to the surface wind speed.

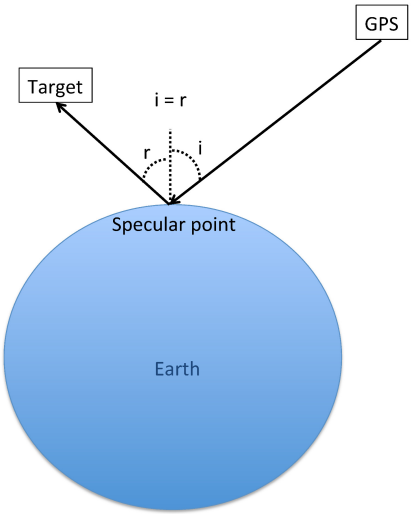

Figure 2.5: Geometry of a specular point.

This technology is used on the CYGNSS mission. The CYGNSS mission, launched in December 2016, consists of eight satellites that measure the wind speeds within tropical storms using reflected Global Positioning System (GPS) signals off the ocean surface. Being able to plan in advance when the specular points between the CYGNSS and GPS spacecraft will be in a storm or cyclone is an important requirement of the mission. SpOCK is used to predict the position of the specular points between the GPS and CYGNSS constellations. A demonstration of this feature is presented in Section 2.3.1.

Ground station coverage

Another feature available in the propagator is the calculation of when a satellite is visible to a ground station. This is important for mission design and mission planning for the transmission of data between the satellite and the ground. For a set of ground stations and minimum elevation angles (representing the cones in which the satellite has to be in order to communicate with the ground station), the propagator returns the times when the satellites are in sight of the ground stations. Examples

are presented in Section 2.3.2.

Collision risk assessment

SpOCK can assess the risk of collisions with other space objects in orbit (operational satellites or debris). Monte Carlo procedures are used to perturb the epoch state (position and velocity) of the primary and secondary spacecraft from the given covariance matrices. If a close approach is found (defined as a relative distance smaller than a given threshold), it calculates the number of samples from each of the two sets of satellite ensembles that are spaced out by less than the sum of both object radii. This situation is recorded as a collision. The probability of collision is equal to the total number of collisions divided by the number of samples used in the ensemble simulation.

In addition to modeling uncertainties in the initial state, SpOCK can also model uncertainties in the thermospheric density, the drag coefficient, and the attitude. To model the uncertainty in the density, SpOCK uses the predictions of the solar activity and adds a random Gaussian variability to the values. The standard deviation of the Gaussian distributions is based on an average of the error in the activity level predictions compared to actual observations over the last five years.

Modeling the uncertainty in the attitude consists in having thousands of samples that drift with a random angular velocity from a nominal attitude for a given time before going back to the nominal attitude. This enables the simulation of the attitude determination and control system of satellites that randomly drift from a nominal attitude.

Finally, to model the uncertainty in the drag coefficient, the user can either add a dimension to the state covariance matrix that includes the variance of the drag coefficient, or directly specify the standard deviation on the drag coefficient of each surface.

This complete modeling of the uncertainties (state, thermospheric density, drag coefficient, attitude) allows the assessment of the risk of collision by taking into account both errors in the initial state and in the dynamics of the system, which improves the accuracy of the risk assessment calculation. Determining the risk of collision with a Monte Carlo technique is computationally intensive, since one has to consider a large number of ensemble members in order to produce accurate results of the probability of collision as it is most often a very low number. For example, NASA recommends performing a collision avoidance maneuver if the probability of collision exceeds a value of 10^{-4}. To be more computationally efficient, SpOCK runs the ensemble members in parallel, which allows the risk assessment to be performed in only a few hours.

Chapter IV will explore the effects of uncertainties in the thermospheric density on the probability of collision.

2.3 Demonstration of capabilities

2.3.1 Storm Intersection Forecast Tool (SIFT)

The main objective of the CYGNSS mission is to measure wind speeds in tropical storms. Measuring the reflected power of GPS signals off the surface of the ocean, the Delay Doppler Mapping Instrument (DDMI) infers the surface roughness to derive the wind speed. The points where the signals sent by the GPS are reflected back to the CYGNSS satellites are called specular points, as explained in Section 2.2.2.2.

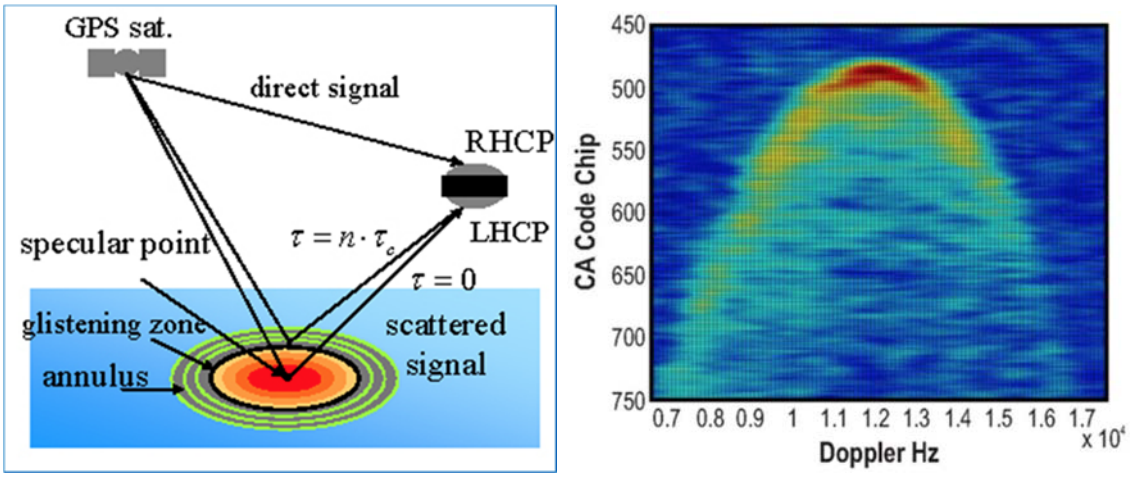

Figure 2.6: GPS signal propagation and scattering geometries (left); Example of a Delay Doppler Map (right) (*Clarizia*, 2015).

Figure 2.6 shows the geometry of a GPS signal scattered off the ocean surface and reflected to the CYGNSS Left-Hand Circular Polarization (LHCP) antenna. The strength of the reflected signal contains information about the roughness of the ocean as a function of delay and Doppler, that can then be related to the wind speed around the specular point. An example of a Delay Doppler Map (DDM) is given in Figure 2.6. Typically, the winds over the ocean do not have much structure over scales of 25-100 km, allowing the decimation of the DDM to an area directly around the specular point. Within a tropical cyclone, however, the wind speed can gave significant struc-

tures, which makes the information contained within the DDM further away from the specular point valuable. However, the problem is that this requires more information to be transmitted to the ground stations, requiring more operations cost. Therefore, the mission would like to limit the times of larger DDM transmissions to only time periods when flying directly over the cyclone. When this situation occurs, the satellites switch from the standard mode to the Full Delay Doppler Mapping (FDDM) mode. The Mission Operation Center (MOC), based at the SouthWest Research Institute (SWRI) in Boulder, Colorado, does not constantly communicate with the satellites. Consequently, the MOC sends the commands a day and half in advance to tell the observatories when to switch to the FFDM, which creates the need for a tool that predicts when individual CYGNSS observations will fly over the cyclone. These considerations constrain the MOC to anticipate when the GPS and CYGNSS satellites align in a way that results in specular points being within the cyclone, and requires an accurate prediction of the positions and velocities of the two constellations in order to predict the positions of the specular points.

SpOCK can predict when the specular points will be within the path of the cyclone. Because a perfect prediction of the storm trajectory is impossible, NOAA regularly publishes predictions of cyclones path and size, along with uncertainties associated with these predictions (`http://manati.star.nesdis.noaa.gov/TC_cone_info/`). To know when the specular points have a good chance to be within the storm, the tool predicts the probability of each specular point being within the cyclone one to two days in advance. From the most recent CYGNSS and GPS positions and velocities information, as well as the inputs necessary to propagate the orbits (predicted attitude of the observatories and latest predictions of solar activity), SpOCK calculates the position of the specular points for the next two days. It then combines these results with the NOAA prediction of the storm trajectory to predict when the

specular points will be in the cyclone. A Graphical User Interface (GUI) was implemented by the CYGNSS science team to allow for quick and easy visualization of the trajectories of the observatories, the specular points, and the cyclones. This interface enables the Science Operation Center (SOC) at the University of Michigan to visualize animations of these trajectories in order to select times when the specular points intersect the storm. The SOC requests that the MOC in Boulder, Colorado send commands to the CYGNSS satellites to switch the DDMI to the Full Delay Doppler Mapping mode for these particular times. The GUI also shows the positions of the three ground stations (Chile, Australia, and Hawaii) to which the observatories can downlink the data measured during the over-flight, so that the full DDMs can be retrieved quickly to allow for rapid specification of the winds through the hurricane.

Figure 2.7 shows an example of specular point trajectories that intersect the path of a cyclone, from the SIFT GUI. Specular point paths by one CYGNSS observatory for three consecutive orbits predicted one day ahead are shown. Two tracks of specular points are predicted to overlap with the cyclone trajectory during the first fly-by of CYGNSS, so the SOC would send a request to the MOC to turn on the FDDM for this particular observatory over-flight.

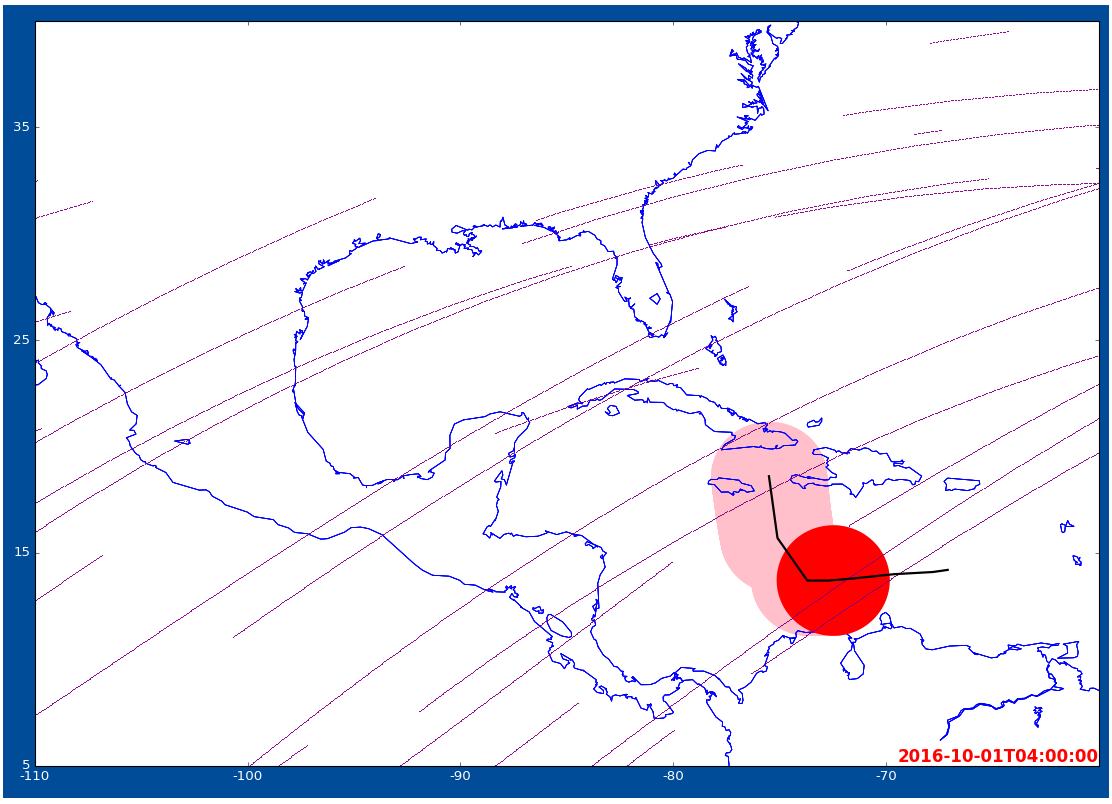

Figure 2.7: Specular points path over five hours (3 passes of a CYGNSS satellite) near the cyclone. The dark red disk represents the current position of the cyclone (2016-10-01T04:00:00), and the light red disk its predicted trajectory. The black line represents the trajectory (past and future) determined by NOAA and the shaded red cone includes the 34 knots radii of the cyclone. The hurricane is predicted to move North.

2.3.2 Coverage of ground stations

SpOCK can compute the times in which a satellite can communicate with particular ground stations. The user specifies the position and minimum elevation angle of each ground station, and SpOCK computes the azimuth and elevation angles in both the ground station and the spacecraft reference systems, as well as the range. The output file includes these four angles and the range for every second of the propagation. A summary report is also created that shows the access start and end times when the spacecraft flies over the ground station, in the cone defined by the minimum elevation angle.

Figure 2.8 shows an example of the coverage of three ground stations over one day by one of the eight CYGNSS satellites: South Point (HI, USA), Santiago (Chile), and Western Australia (Australia). The minimum elevation angle used for each station is 5°. Over the day, CYGNSS flies over each station 5 times, with durations of about 8 minutes. This type of figure is automatically generated by SpOCK when the user chooses to compute the coverage of ground stations.

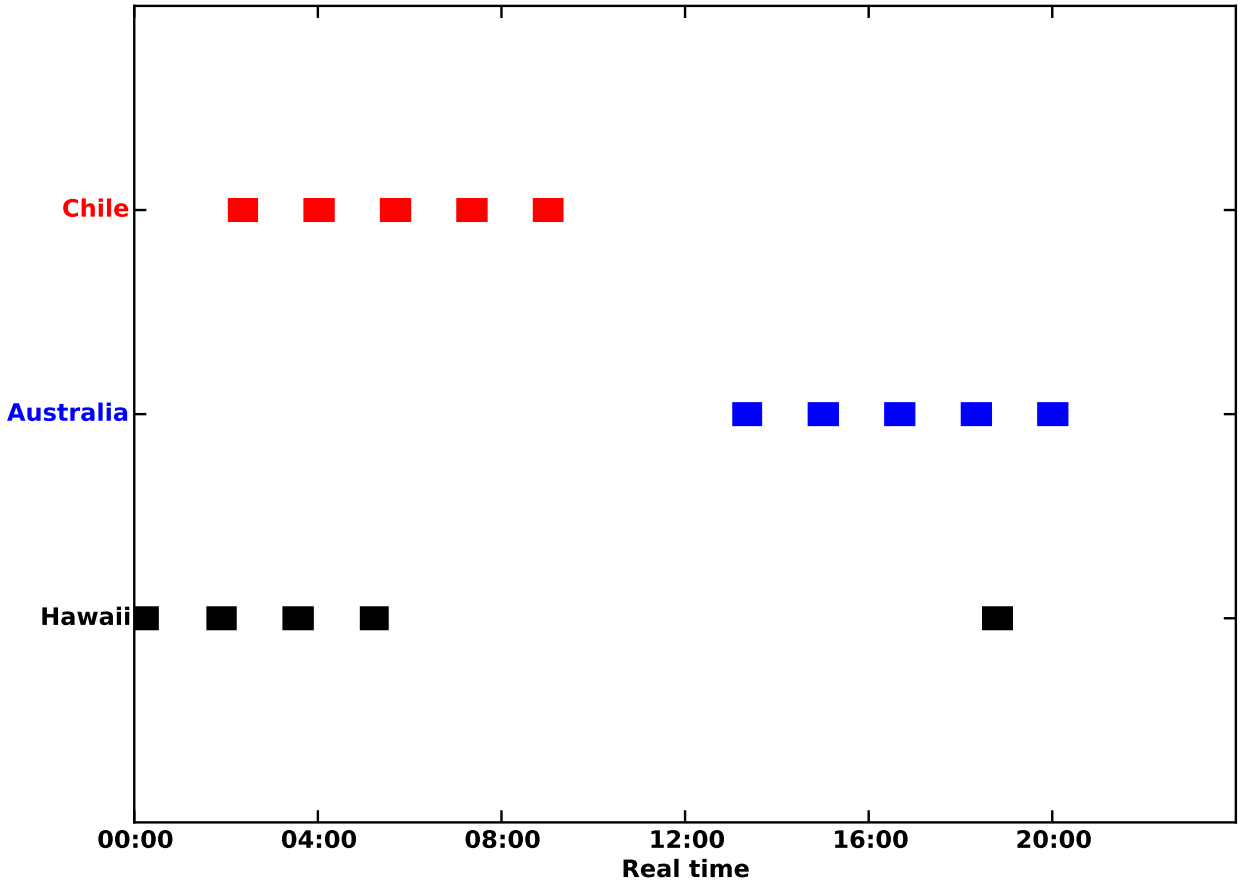

Figure 2.8: Example of the coverage of 3 ground stations by CYGNSS - Access times over a day.

2.3.3 Solar power

In this section, a few examples of computation of solar power by SpOCK are presented. SpOCK can be used for mission design to investigate the optimum spacecraft geometry to meet the power requirements of the mission, once the orbital parameters,

such as inclination, altitude, and attitude, have been selected.

Year long simulations are sometimes necessary to ensure that the solar panels will generate enough power during the entire mission, since the precession of the orbital planes with respect to the Sun, as well as seasonal effects, are quite important. The equatorial bulge of the Earth increases the gravitational attraction and introduces a force component toward the equator. This force induces a torque that rotates the angular momentum of the orbit around the rotational axis of the Earth, causing the satellite's orbital nodes to precess westward (for prograde orbits, and eastward for retrograde orbits). The zonal gravitational harmonic coefficient J_2 represents this bulge. A mathematical equation can be derived to evaluate the precession rate of the nodes (*Vallado and McClain*, 2007):

$$\dot{\Omega} = -\frac{3}{2} R_{\oplus}^2 J_2 \sqrt{\frac{\mu}{a^7}} \frac{\cos i}{(1-e^2)^2} \tag{2.6}$$

where i is the inclination, a the semi-major axis, and e the eccentricity of the orbit. For example, if the orbit is circular at 510 km with an inclination of 35° (such as the orbit of CYGNSS), the nodes will precess with an angular rate of 6.23°/day, which corresponds to 6.32 revolutions of the ascending node during a year, with respect to a fixed position in the ECI frame.

Because of the rotation of the Earth around the Sun, the relative position of the Sun in the ECI frame changes by 360° in a year. Since this motion is eastward and the precession of the orbital plane is westward (for prograde orbits), then the Local Time of the Ascending Node (LTAN) (the angle between the Earth-Sun vector and the ascending node) evolves faster than the precession rate: LTAN = $\dot{\Omega}$ + 360/365. With the same example, the ascending node makes 7.32 revolutions with respect to the Earth-Sun direction. Figure 2.9 shows an example, as viewed from the North Pole.

Seven orbits are represented, each separated by a day. Since a complete revolution of the ascending node with respect to the Sun takes $365/7.32 \sim 50$ days, these 7 orbits correspond to 15% of a full revolution of the ascending node. As time goes, the orbit precesses westward, or equivalently in the direction of decreasing local time.

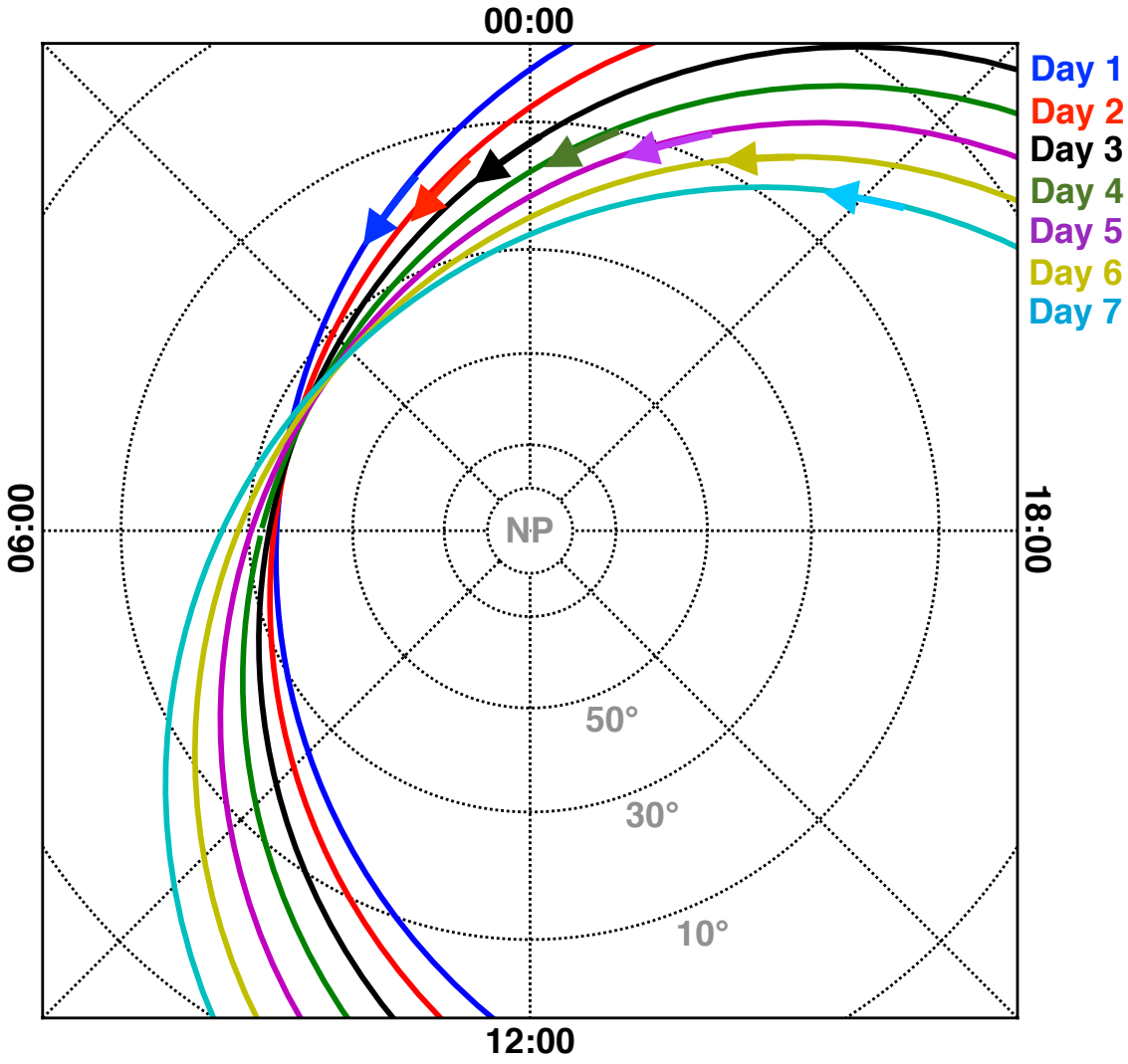

Figure 2.9: Precession of the ascending node with respect to the Sun over 7 days - example with CYGNSS's orbital parameters. The arrow on each orbit represents the direction of motion of the satellite (eastward as the orbit is prograde).

The precession of the orbit is important to take into account when computing the

power because it changes the orientation of the orbit with respect to the Sun. For example, the orbit-average power of a 3-axis stabilized, nadir facing 1U CubeSat orbiting at an inclination of 90° is shown in Figure 2.10 (top), with the LTAN indicated at the local minima and maxima. Only the zenith face of the CubeSat is assumed to have solar panels. Because the orbit is polar, it does not precess in the inertial frame. The variation of the orbit-average power is due to the rotation of the Earth around the Sun. In other words, the LTAN changes by 360° in exactly one year. The maxima of the orbit-average power in Figure 2.10 (top) are reached for a noon-to-midnight orbit because the angle between the Sun direction and the normal to the solar panel is the smallest (0°). However, when the orbit is dawn-to-dusk, the CubeSat moves near the terminator and the Sun light grazes the solar panel for the entire orbit. In this example, the dawn-dusk orbit happens at the solstices. Therefore, because the Earth is tilted by $\sim 23°$ with respect to the ecliptic plane, the angle between the orbit and the Sun direction is not 90° but 67°, which explains why the zenith facing solar panels still receive some sunlight.

As explained in Section 2.2.2.1, the user can change the geometry of the spacecraft by modifying the normal vector in the Local Vertical Local Horizontal frame for each surface. This can be used in mission planning when trying to determine the geometrical configuration that optimizes the solar power. To give an example, two 1U wings extending from the zenith panels in the cross-track directions fully covered with solar cells were added to the previous 1U CubeSat, and their angle with respect to the nadir and horizontal cross-track was varied. For instance, if the angle was 90°, it means that the wings were folded all the way to the port and starboard faces of the CubeSat. An angle of 0° means that the wings were in the same plane as the zenith face. Figure 2.10 (bottom) shows the orbit-average power of four 1U CubeSats with wings that make different angles with the local body horizontal: 0° (black), 30°

(green), 60° (red), and 90° (blue). For flat wings (0°, black line), the shape was the same as in Figure 2.10 (top) but the power was multiplied by a factor 3, since the two wings had a solar cell area equal to the zenith face. As the wings were more tilted (30°, 60°, 90°), the power for the noon-to-midnight orbit decreased. When the wings were completely folded (90°), they did not get any sunlight from the Sun in the noon-midnight plane since the Sun vector was in the orbital plane. Therefore, the total power received by the CubeSat was similar to the power of the CubeSat without any wings in Figure 2.10 (top) (1.2 W) for only the times when the orbit was in the noon-midnight plane (3-21 and 9-21). However, when the orbit was dawn-to-dusk, the flat wings (0°) received less power than the tilted wings because the Sun direction was almost perpendicular to the orbital plane, allowing sunlight to fall on one of the panels constantly. Therefore, the more tilted the wings were, the more power they received from the Sun.

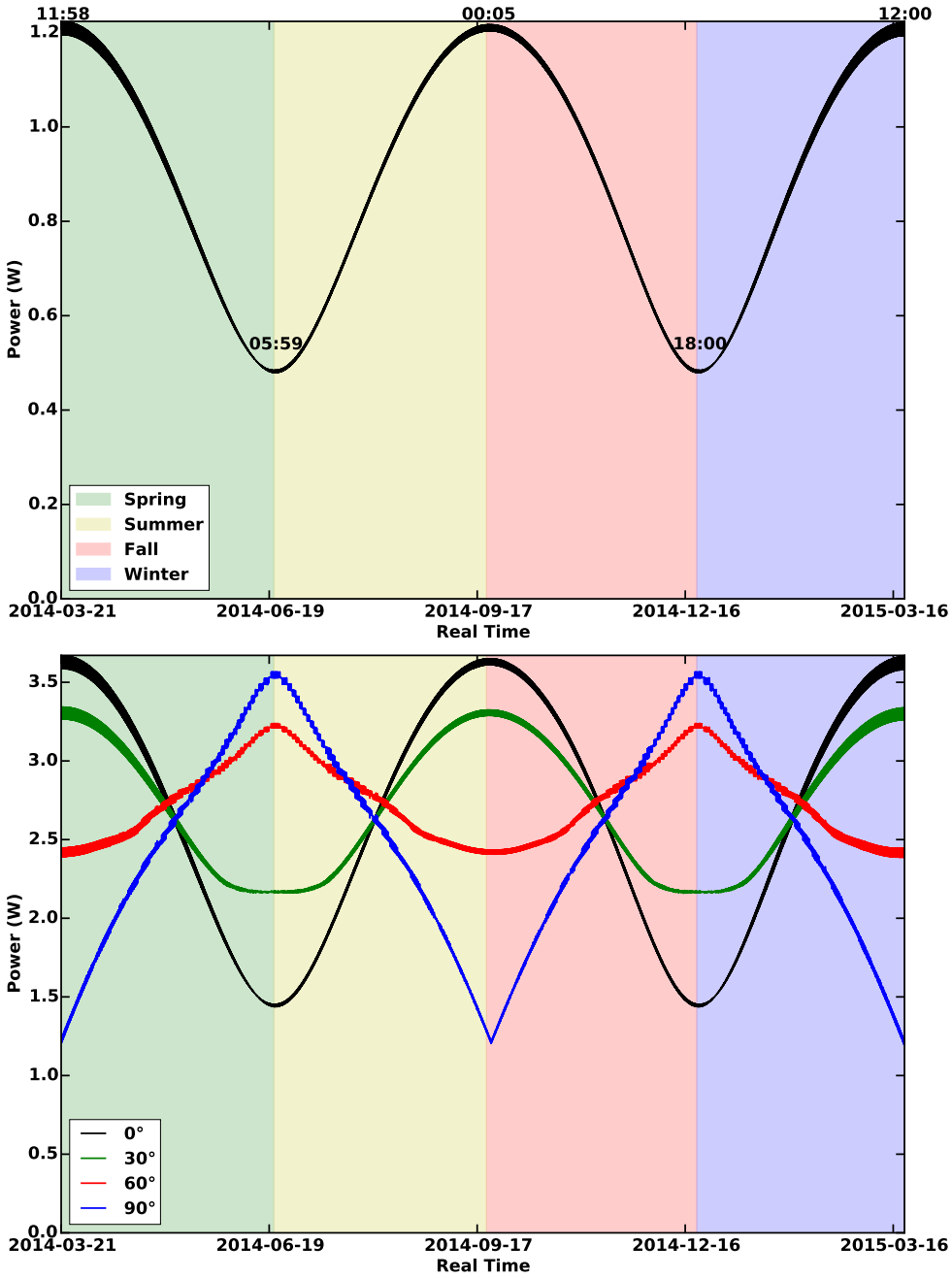

Figure 2.10: Orbit-average power of a 1U CubeSat in a polar orbit (top); orbit-average power of a 1U CubeSat in a polar orbit with wings at different angles (bottom).

Although simplified here, this kind of analysis is often made in mission design when the geometry of the spacecraft must be optimized to maximize the solar power over the entire mission. The propagator offers the possibility of making such analysis

by only modifying a simple geometry file.

Table 2.1: Osculating elements for three test case orbits.

	a (km)	i (°)	Arg of Perigee (°)	RAAN (°)	True anomaly (°)	e
LEO	6758.524129	51.667871	151.908712	45.649594	32.718375	0
HEO	26553.376213	63.400000	270.000001	330.214160	359.999998	0.740969
GEO	42164.166418	0	103.882708	0	226.368596	0

2.4 Comparison with other tools

In order to validate the propagator, results from the High Precision Orbit Propagator (HPOP) within STK from AGI were compared to SpOCK results. Specifically, the position and velocity, ground coverage, and solar power were compared.

2.4.1 Inertial position and velocity

Vallado and McClain (2007) recommended considering different types of orbits in order to fully validate a propagator, in particular because the magnitude of the perturbing forces vary dramatically with altitude. Therefore, ephemerides were compared between HPOP and SpOCK for three different orbits, following a similar approach to *Gaylor et al.* (2006), so that results could be compared with their study: LEO (altitude = 380 km), High Elliptical Orbit (HEO) (eccentricity = 0.74, semi-major axis = 26,553 km), and Geosynchronous Earth Orbit (GEO) (altitude = 35,786 km). Initial osculating elements of the orbits are presented in Table 2.1. As it will be shown, the differences between SpOCK and HPOP are similar to the ones between JAT and HPOP as presented in *Gaylor et al.* (2006).

To validate the force models, each perturbing force was run independently:

1. two-body spherical Earth;

2. two-body spherical Earth with atmospheric drag using NRLMSISE with a constant F10.7 of 100, a constant Ap of 15 (Kp = 3 in HPOP), drag coefficient = 2, and area/mass = 0.01 m^2/kg;

3. two-body spherical Earth with third-body perturbations (Sun and Moon);

4. two-body spherical Earth with solar radiation pressure using solar radiation pressure coefficient = 1, area/mass = 0.01 m^2/kg; and

5. spherical harmonic gravity model using the WGS84 - EGM96 model with degree and order equal to 20.

For each of these orbits, satellites were propagated for one day with HPOP and SpOCK with an integration time step of 5 seconds and ephemerides were generated every 60 seconds. The numerical integration is made using a fourth-order Runge-Kutta method in STK and SpOCK.

Table 2.2 shows the maximum position and velocity differences over the entire propagation for the LEO orbit. The two-body spherical Earth with and without the third-body perturbation from the Moon and the Sun show maximum differences of 12 millimeters between HPOP and SpOCK. When the spherical harmonic gravity model was used, both propagators agreed to within 2 meters, and 5 meters when the solar radiation pressure was taken into account. The largest difference occured when the atmospheric drag was computed, with a 9.5 meter maximum difference. This may be due to the fact that the thermospheric density is very complex to model. NRLMSISE is highly dependent on inputs such as F10.7 and Ap (or Kp) and the exact way the thermospheric model was used in STK is not well documented, so errors might come from the difficulty in perfectly matching the atmospheric models. In particular, STK uses the geomagnetic index Kp while SpOCK uses Ap. A value of 3 for Kp is converted into a value of 15 for Ap (`http://www.ngdc.noaa.gov/stp/GEOMAG/kp_ap.html`). However, this is not an exact solution, which implies that a difference in the thermospheric density can possibly arise from this conversion. For example, if Ap were equal to 14 (16) instead of 15, the maximum position difference with HPOP increased to 150 m (125 m), which demonstrates the high sensitivity of the density model, and therefore the propagation, to its inputs. Because the altitude was very low (380 km), the density was quite important, so it is consistent that the

Table 2.2: Maximum position and velocity differences over one day with different perturbing forces for the LEO orbit.

Perturbing force	Position (m)	Velocity (m/s)
Two-body spherical Earth	0.012	0.001
Two-body spherical Earth & drag	9.554	0.011
Two-body spherical Earth & third-body	0.012	0.001
Two-body spherical Earth & solar pressure	5.382	0.005
WGS84 - EGM96	2.011	0.003

Table 2.3: Maximum position and velocity differences over one day with different perturbing forces for the HEO orbit.

Perturbing force	Position (m)	Velocity (m/s)
Two-body spherical Earth	0.063	0.001
Two-body spherical Earth & drag	0.250	0.001
Two-body spherical Earth & third-body	0.063	0.001
Two-body spherical Earth & solar pressure	1.838	0.002
WGS84 - EGM96	8.319	0.006

uncertainty on the atmospheric drag was the main source of errors.

Table 2.3 shows the maximum position and velocity differences over the entire propagation for the HEO orbit. This time, the most important source of differences between both propagators was the gravity model, accounting for a difference of 8.3 meters after a day. When the other forces were run independently, differences of less than a meter were found, except for the solar radiation pressure, which resulted in a maximum difference of 1.8 meters.

Table 2.4 shows the maximum position and velocity differences over the day for the GEO orbit. In every case, the difference in positions was well under a meter, with the maximum difference occurring when the solar pressure was used as it was the main perturbing force at this very high altitude.

For each configuration, the maximum difference in velocity was a few millimeters per second.

Table 2.4: Maximum position and velocity differences over one day with different perturbing forces for the GEO orbit.

Perturbing force	Position (m)	Velocity (m/s)
Two-body spherical Earth	0.076	0
Two-body spherical Earth & drag	0.076	0
Two-body spherical Earth & third-body	0.075	0.001
Two-body spherical Earth & solar pressure	0.714	0.001
WGS84 - EGM96	0.098	0.001

2.4.2 Coverage of ground stations

The ground station coverage by a CYGNSS satellite over a day computed by SpOCK and presented in Figure 2.8 was compared against the coverage derived in STK. The access times were very similar, with maximum differences smaller than a second, so the blocks symbolizing the accesses in Figure 2.8 overlap with the ones from the coverage computed by STK. A 1 s difference in access time translates into an elevation difference in angle of the spacecraft as seen from the ground station of only $\tan^{-1}\left[7.6/\left(R_{\oplus}+500\right)\right]=0.06^{\circ}$ for a satellite orbiting at 500 km with a speed of 7.6 km/s.

2.4.3 Solar power

In this section, the solar power computed for CYGNSS over one orbit by SpOCK is compared to calculations made in STK. The geometry of CYGNSS is illustrated in Figure 2.11. The spacecraft included 0.0612 m^2 solar cells on each the ram and wake faces, as well as 0.558 m^2 of cells on the zenith face.

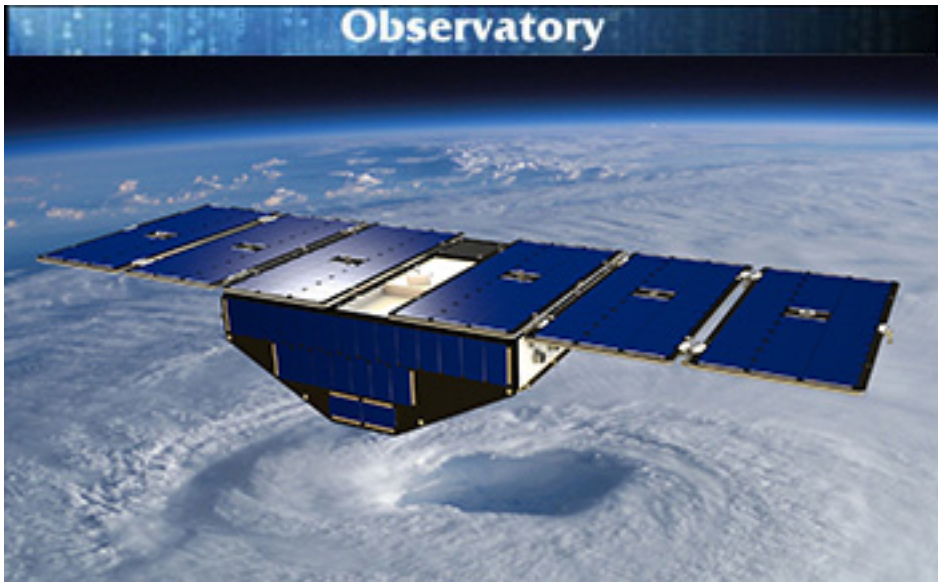

Figure 2.11: Artistic representation of one of the eight CYGNSS satellites (`http://clasp-research.engin.umich.edu/`).

The circular orbit was initialized at 500 km, with an inclination of 35°, and a RAAN, true anomaly and argument of perigee of 0°. Figure 2.12 shows the power generated by the three solar panels (ram, wake, zenith) for a nadir pointing configuration. The dots represent the power computed by STK, and the solid lines the power computed by SpOCK. Figure 2.12 shows a very good agreement between both power computations, with a Normalized Root Mean Square Error of 0.13 % (defined as $\sqrt{\frac{\overline{\left(P_{\mathrm{SpOCK}}-P_{\mathrm{STK}}\right)^2}}{\overline{P_{\mathrm{STK}}^2}}} \times 100$, where $\overline{u}$ represents the temporal average of the quantity u), showing that the SpOCK power calculation tool is very consistent with the STK tool.

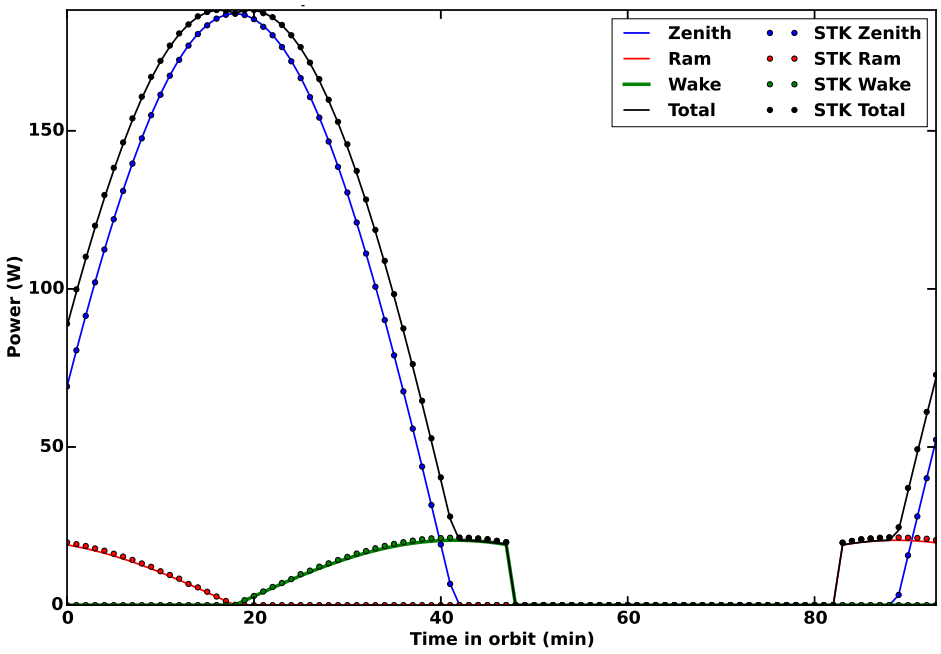

Figure 2.12: CYGNSS power over one orbit computed by SpOCK (solid lines) and STK (dots).

2.5 Summary

The numerical integration of SpOCK uses a comprehensive modeling of the dynamics of spacecraft in orbit by taking into account the perturbing forces acting on the satellite. Specifically, the non-spherical portion of the mass distribution of the Earth is modeled with a decomposition of spherical harmonics for the gravitational potential. Moreover, SpOCK uses thermospheric models (NRLMSISE and GITM) to derive the atmospheric density at the position of the spacecraft, allowing for an accurate representation of the atmospheric drag. The errors resulting from neglecting perturbing forces and from integrating the trajectories with a too large time step were presented.

In addition to modeling the motion of satellites, the different functionalities of SpOCK were introduced. For instance, the coverage of the three ground stations that communicate with the CYGNSS observatories was computed with the determination of the access times over a day. The precession of the ascending node due to the asphericity of the Earth was demonstrated over a 7 day simulation. The orbit-average solar power of a 1U CubeSat in a polar orbit with wings at different angles was computed to study the influence of the spacecraft geometry on the power generated by the solar cells of the satellite. The influences of seasons and of the precession of the nodes were shown too. Finally, a demonstration of the SIFT was made. This algorithm predicts the intersection of the CYGNSS observatories' specular points trajectories with the path of tropical storms forecast by NOAA. It provides support to the CYGNSS mission operation center in sending the commands a day and half in advance to tell the observatories when to switch to the higher resolution mode in order to take more accurate measurements of the winds in the cyclone.

SpOCK was validated by comparing the inertial positions and velocities with results from the High Precision Orbit Propagator within the Satellite Took Kit from Analytical Graphics, Inc.. Specifically, the ephemerides after one day of propaga-

tion were compared for three different orbits: Low Earth Orbit (altitude = 380 km), High Elliptical Orbit (eccentricity = 0.74, semi-major axis = 26,553 km), and Geostationary Orbit (altitude = 35,786 km). The complete force model was validated by evaluating the perturbing forces independently. The comparison showed sub-meter differences between SpOCK and HPOP positions, except for the validation of the drag model in the LEO configuration for which a difference of 10 meters was found, probably due to a difference in modeling the atmospheric density. Differences in velocities did not exceed a few mm/s.

Collision risk assessment has been implemented in SpOCK and will be presented in Chapter IV.

CHAPTER III

Predictions of the solar wind speed

3.1 The Probability Distribution Function model (version 1)

Uncertainties in the atmospheric drag can result in important errors in the predictions of the positions and velocities of satellites orbiting at low altitude (< 600 km), posing a real challenge for collision avoidance. These uncertainties are mainly induced by uncertainties in the neutral density of the upper atmosphere, which are driven by the solar activity, in particular the solar wind speed. The Probability Distribution Function (PDF) model was developed to predict the solar wind speed five days in advance, and to assess the uncertainty on the prediction. The first version of the PDF model focused on predictions of the background solar wind speed, which corresponds to times when the solar activity is quiet or moderate. The methodology is presented, as well as results of comparisons with the Wang-Sheeley-Arge (WSA) model.

3.1.1 Methodology

There are two concepts that drive the PDF model. The first is the idea that the solar drivers do not randomly change from hour to hour. For example, it has been shown that persistence is one of the best models of the solar wind speed for short-term predictions (e.g., *Norquist and Meeks*, 2010). For the two first days of prediction,

keeping the speed constant at its current value is one of the best estimators for the value of the speed. This implies that the solar wind speed changes relatively slowly, such that, if the speed at the present time is known, then the speed in an hour from now will most likely be quite similar to the value now. This idea can be formalized into a probability-based model, where PDFs can be derived for the solar wind speed, given the current speed and the prediction horizon (each hour from the current time until 5 days into the future). Through exploring the behaviour of the solar wind speed over times of many 10s of hours, it was noted that the prediction can be separated into two different groups, differentiated by whether the speed had been increasing or decreasing over the previous 12 hours.

The second idea of the PDF model is based on the approximately 27-day rotation of the Sun. The rotation period depends on the latitude: it is 25 days at the equator and 35 days at the poles. On average, it is about 27.2 days. Therefore, the same active regions of the Sun are directed towards the Earth every 27 days. This implies that a 27-day periodicity should also be observed in the solar wind speed, the Interplanetary Magnetic Field (IMF), and the solar-radiation flux. A comprehensive study of this 27-day periodicity has been detailed by *Ram et al.* (2010). That study highlighted a strong 27-day periodicity, but also signatures of periodicities corresponding to 13.5, 9, and 6.8 days. Given the 27-day periodicity of the Sun, a model theoretically can be created in which the solar wind-speed prediction can be based on the solar wind speed 27 days ago. PDFs can therefore be created given the present speed and the speed in approximately 27 days from now.

The biggest problem with this technique is that the solar structures are not perfectly periodic. Two consecutive solar rotations can lead to very different speeds of the solar wind, as shown in Figure 3.1. The Pearson correlation between these two speeds is 0.2 (all the correlation numbers that we present in this study correspond to Pearson correlations). The Pearson correlation, a number between -1 (anti-correlated) and 1

(correlated), quantifies the similarities between the shapes of two curves. The problem with a cross-correlation comparison is that the two series could be quite steady, with small variations on top of a large baseline, and the cross correlation could return a value of almost zero, even though the baselines are identical. This is because the cross correlation focuses only on the variations, which could be a minor part of the signal. With the solar wind speed, the baseline can sometimes be quite large, compared to the variations. A root-mean-squared (Root Mean Squared (RMS)) difference can also be used to quantify the differences between two series. For an RMS, if baselines are similar, even though the (small) variations are different, the result will be close to 0. A normalized root-mean-squared difference divides the RMS by the mean value of one of the series, so that one can judge the relative difference between the series. For example, an Normalized Root Mean Squared (NRMS) of 0 still implies perfect agreement, but an NRMS of 0.8 (which is the value calculated for the two speeds in Figure 3.1) implies that one series is different from the other by an offset of 80%. The mathematical expression is:

$$\text{NRMS} = \sqrt{\frac{\overline{(u-v)^2}}{\overline{u^2}}} \tag{3.1}$$

where u and v are two different series and the symbol $\overline{u}$ represents the mean of the series u.

It was often found that the lag between consecutive solar rotations is not exactly 27 days. Indeed, this is often the case. The optimum lag could vary between approximately 22 and 32 days ago. Table 3.1 shows that when the lag was forced to exactly 27 days, the correlation and NRMS are relatively poor. However, when the lag was allowed to adjust to optimize the correlation or the NRMS, these values were dramatically improved. The correlation and NRMS were calculated for 6,000 5-day periods of solar wind speed during 1995-2011.

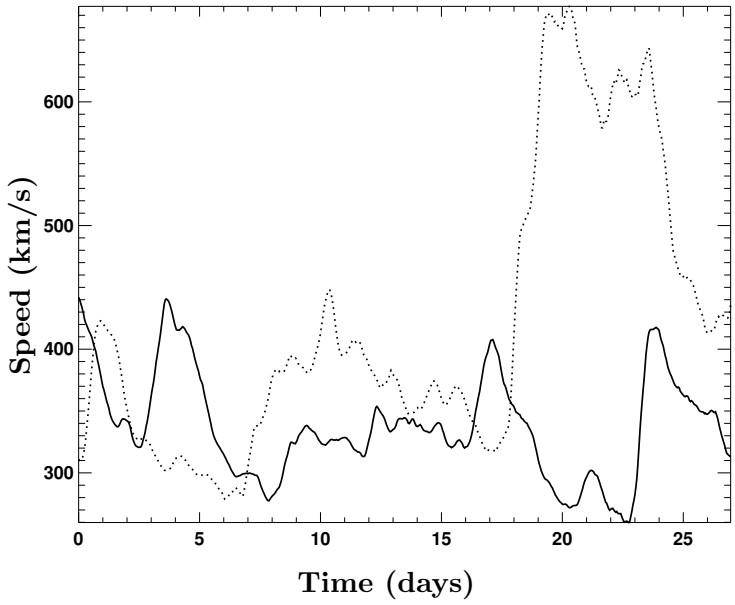

Figure 3.1: Example of the solar wind speed during two rotations that correspond to a bad correlation.

Table 3.1: The mean correlation and NRMS between approximately 6,000 consecutive solar rotations with an absolute 27-day lag and with allowing the lag to be adjusted until an optimum is found.

Condition	Correlation	NRMS
27-day	0.22	0.19
27±5 day	0.89	0.09

The general idea for forecasting the solar wind speed i hours into the future would be to use the data from one solar rotation ago (OSRA), or 27 days plus an optimum lag ago (denoted $t_{OSRA} = t_{now} - 27$ days + lag). This optimum lag could be found using the last, for example, 3 days of solar wind data, compared to the solar wind speed approximately 27 to 30 days ago. The expectation was that using the lag t_{OSRA}, the predictions of the solar wind speed would improve. Instead, the opposite was found. Using the lag of exactly 27 days ($t_{27} = t_{now} - 27$ days) produced the best results. Figure 3.2 illustrates this. The red lines show the correlation (left) and NRMS (right) between the 3-day period just before the current time and the period 0-3 days before t_{OSRA} days ago, where the lag was allowed to vary to optimize the

result up to the value on the X-axis. For example, at an x-axis value of 1 day, the lag was allowed to shift to any value between ±1 day (for a total shift of between 26-28 days). For a value of 4 days, the lag was allowed to shift to any value between ±4 days (for a total shift of between 23-31 days). With the knowledge that the lag is not exactly 27 days, but a bit more or less than this value, the red lines show the expected behaviour: as the window size of comparison is opened more and more, the correlation and NRMS improve. The values of the red lines at the 0 and 5-day marks are shown in Table 3.1. The blue lines are the comparisons between the subsequent three days, using the optimized lag determined for the past three days. At lag=0, the red and blue lines have the same values, since t_{27} was used for all times compared. When the lag was allowed to increase, the comparisons for the subsequent days become worse. This means that, statistically, the optimum lag for the last three days is not the optimum lag for the next three days. The optimized lag changes rapidly, which makes prediction of the solar wind speed using the previous solar rotation extremely difficult.

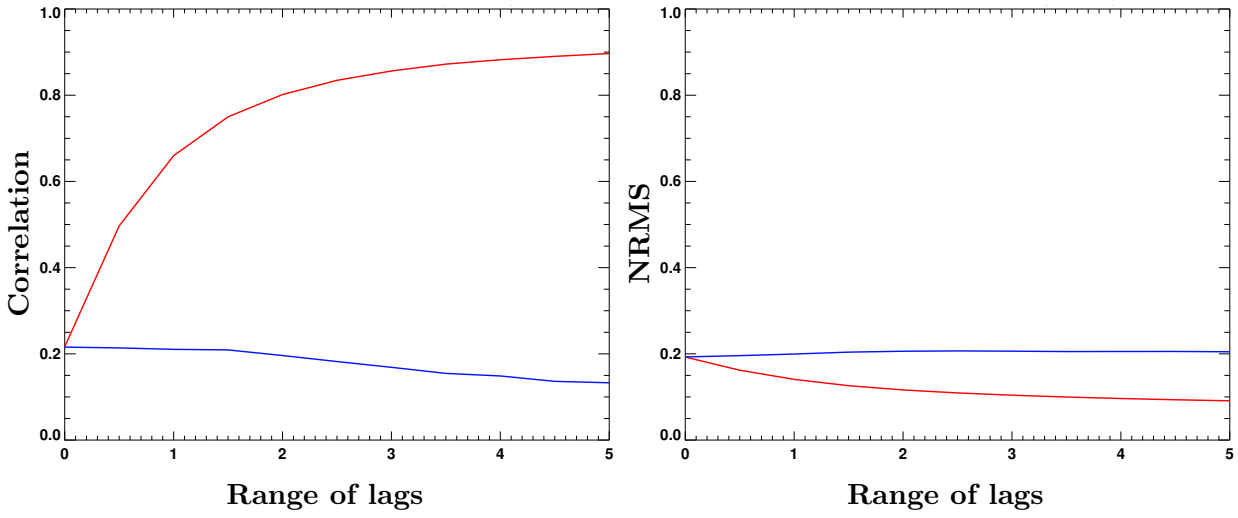

Figure 3.2: Dropoff in the correlation (left) and the NRMS (right) of the predictions. With a correlation, values close to 1 are better, while with an NRMS, values close to 0 are best.

This is the main issue that the PDF model has to face. It implies that if the optimum lag can not be predicted, then the method can not use the past solar rotation in an optimized way for the prediction of the following five days. At this time, it is unknown how to predict the optimum lag. Currently, as described below, the lag is calculated by using a temporally weighted NRMS comparison, where the current time is weighted the strongest and the data from three days ago is weighted the least. This allows the delay to be optimized for the time now, instead of an average of the last three days. Attempting to determine this lag will be the focus of future research and should greatly improve the forecast ability of this type of model. With the current PDF model, there is a greater reliance on the PDFs based on the current solar wind speed and trend.

Construction of the PDFs

The first set of PDFs are simply the distribution of speeds each hour for the next 120 hours, given the speed now and the slope of the 12 preceding hours. For example, if the current solar wind speed is 450 ± 10 km/s and if the mean of the speed of the 12 past hours is greater than the speed now (meaning that the speeds are decreasing), then the distribution of speeds for 1, 10, and 24 hours after the current time is given by the left plot in Figure 3.3. The right plot shows the same thing, but for increasing solar wind speed. The most important thing to notice about the progression of PDFs is that the peak decreases in intensity and that the width of the distribution increases with time. This means that, as the forecast horizon becomes longer, there is a larger number of speeds that could happen and that the percentage of likeliness that the most probable speed will happen decreases. Furthermore, the peak of the PDFs in the left plot move towards slower speeds. This is consistent with the fact that the speed is decreasing. In the right plot, the peak in the PDFs also decrease in speed, but the tails at high solar wind speeds become larger, which shows that there is a

population of speeds that can be significantly larger than the current 450 km/s speed. These PDFs can be compared to a PDF of all of the solar wind speeds over the entire model time period, shown in Figure 3.4 (the analysis includes the Coronal Mass Ejections (CMEs) that occurred during the time period). While these plots are on different scales, it can be seen that the global solar wind distribution peaks at about 9.5%, while the 24-hour PDF peaks at approximately 20% for the decreasing-speed case (left plot) and 15% for the increasing-speed case (right plot). This indicates that the 24-hour PDF is still more useful for predictions than simply using a PDF of all of the solar wind.

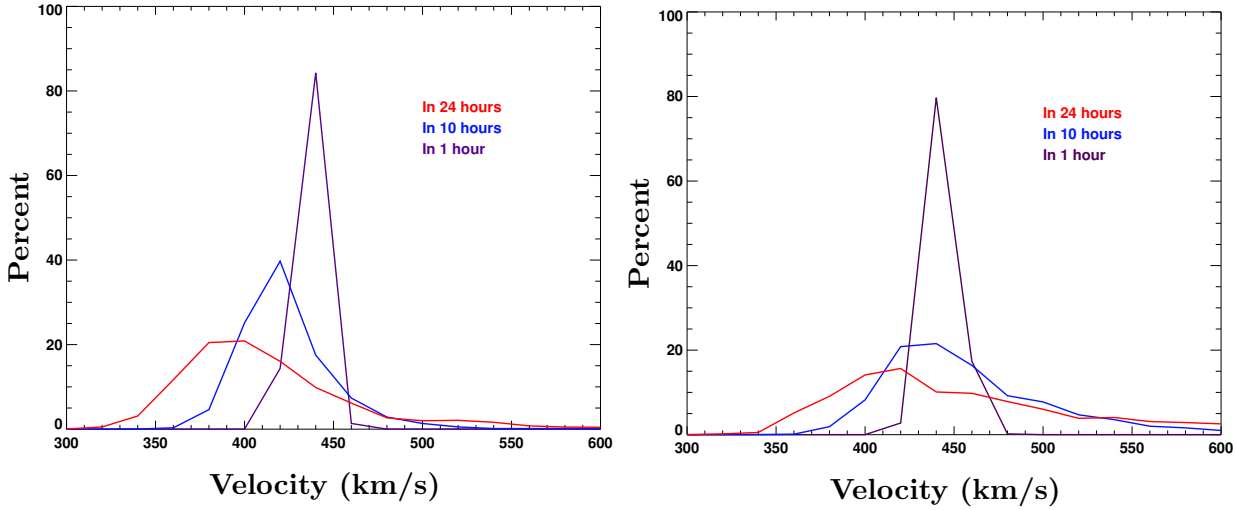

Figure 3.3: Examples of three PDFs (1, 10 and 24 hours from the current time) from P1 based on the current speed being 450 km/s and the speed decreasing (left) and increasing (right) over the last 12 hours.

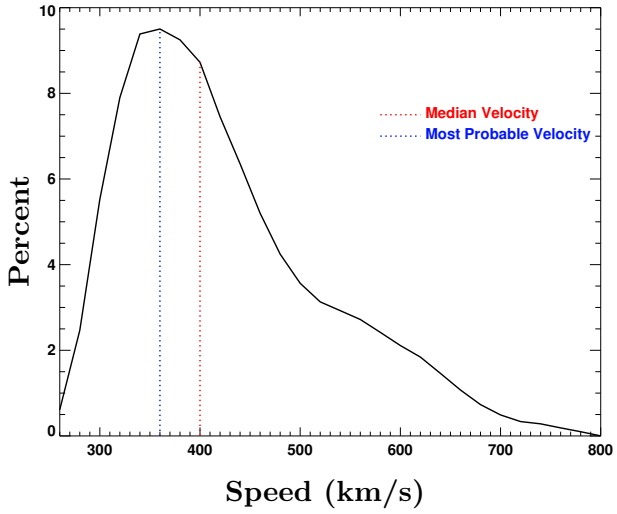

Figure 3.4: The probability distribution function of the solar wind speed over 1995-2011 from 260-800 km/s, with the median and most probable values indicated.

The PDFs were created for bins from 260-800 km/s with bin sizes of 20 km/s (i.e., bin center ±10km/s). The prediction horizon stretches from one to 120 hours, and the data is separated into increasing or decreasing solar wind speeds. This set of PDFs will be called P1. Figure 3.5 gives examples of predictions made using only the P1 PDFs, given the current solar wind speed (Time=0) and the direction of the slope of the solar wind speed over the previous 12 hours. The top two plots and the left plot in the middle row show examples in which the solar wind speed was decreasing, while the right plot in the middle row and the bottom plots show times in which the speed was increasing. The majority of the time, the predicted speed is within the red curves, which indicate the 10% and 90% levels of the PDF. The solar wind-decreasing cases appear to be better predicted than the solar wind-increasing cases, which is to be expected, given that the solar wind-decreasing cases have narrower and taller PDFs .

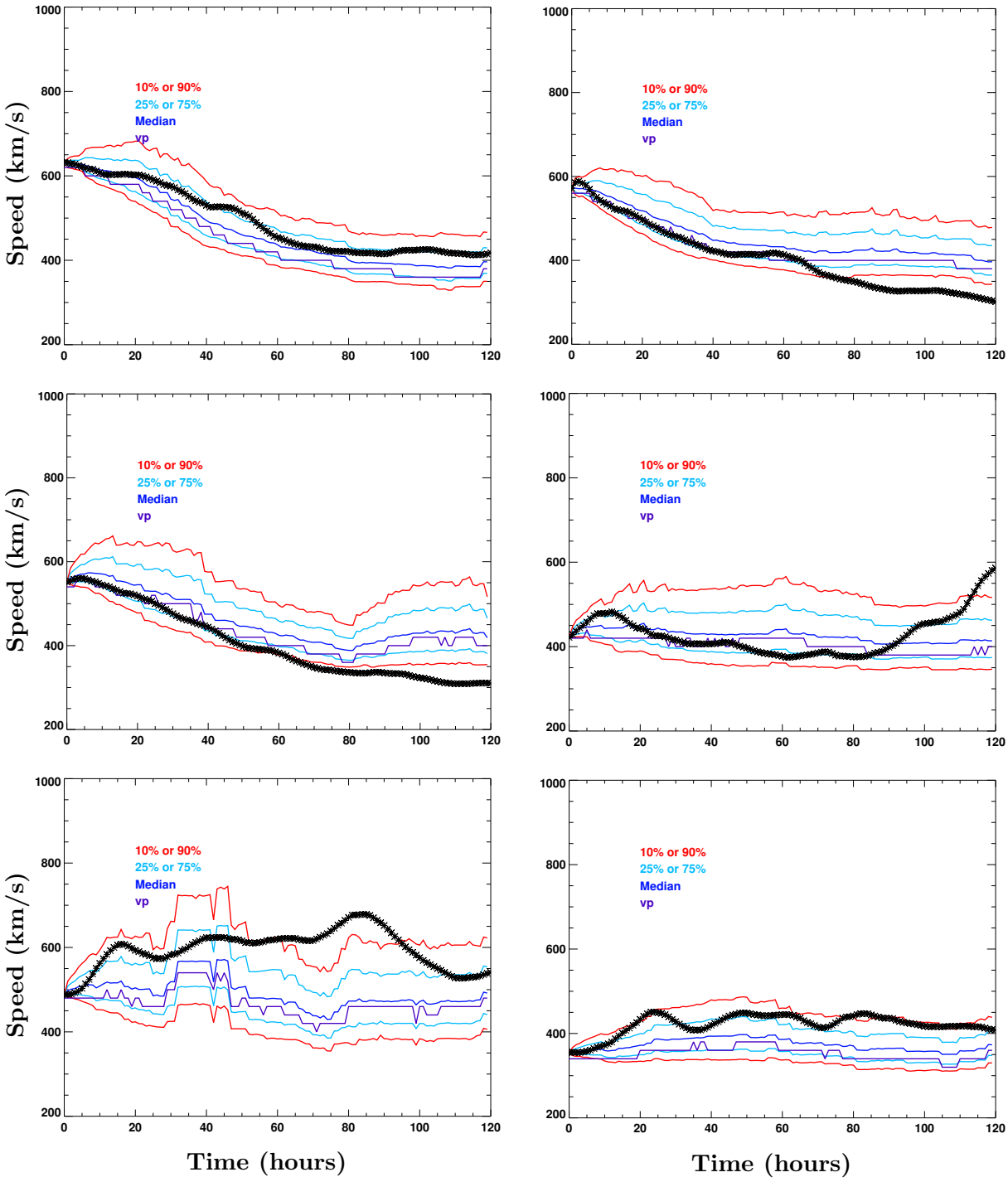

Figure 3.5: Six examples of solar wind speed predictions out to 120 hours using the P1 PDFs. In these cases, the actual solar wind data is shown as black stars. The predictions are indicated by the colored lines, with the most probable and median values indicated by the purple and dark blue lines. The light blue lines indicate the 25th and 75th percentiles, while the red lines indicate the 10th and 90th percentiles.

These plots illustrate one of the advantages of using PDFs to determine the predictions–a range (or uncertainties) of the predictions can be determined. Furthermore, an ensemble of different solar wind-speed prediction scenarios could be generated to allow for ensemble forecasting of the near-Earth space environment. Using either of these allows the forecaster to have more information about the prediction than simply the value and the past performance.

From the P1 PDFs, the median solar wind speed (50% of the speeds are below/above this speed) and the most probable speed (speed that corresponds to the peak in the PDF, which is typically a bit lower than the median speed), as demonstrated in Figure 3.4, can be determined. The median speed (M1) and the most probable speed (VP1), as determined by the set of PDFs P1, can be used to generate a typical single-value predictive model of the solar wind speed.

The second set of PDFs will be called P2. They are based on the 27-day periodicity of the solar wind speed. For each hour of solar wind data from 1995-2011, the optimum lag was found to the following solar rotation. An NRMS comparison with the previous five days was used to determine the optimum lag. The solar wind speed at t_{OSRA} was then noted. PDFs from 260-800 km/s with bin sizes of 20 km/s (i.e., bin center $\pm$10km/s) were created. The change of speed was not included in this case, nor were there predictions beyond the value at the optimum lag time. Therefore there are significantly fewer PDFs in this set. As with PDF P1, the median speeds (M2) and the most probable speeds (VP2) were determined from PDFs in P2.

To summarize, if a prediction of the solar wind speed in i hours is desired, two PDFs are available. P1 gives directly the distribution of speeds in i hours, given the current speed and the trend in speed. In order to use P2, a lag needs to be determined from "exactly" one solar rotation ago, or approximately 27 days - i hours ago. Mathematically, this can be thought of as $t_{OSRA,i} = t_{now} - 27$ days + lag + i hours. The solar wind speed at $t_{OSRA,i}$ determines which PDF P2 should be used to

predict the solar wind speed in i hours from now. A slightly different set of PDFs ($P2_{27}$) can be created simply based on a 27-day delay, instead of finding the optimized delay. This time can be referred to mathematically as: $t_{27,i} = t_{now} - 27$ days $+ i$ hours.

Construction of the pre-models

We introduce the following notations:

- $v_{pred,i}$: the speed to be predicted in i hours from the current time;
- $v_{pdf,i}$: the speed determined by one of the PDFs for i hours from the current time (or, more specifically, $v_{P1,i}$ and $v_{P2,i}$);
- $v_{OSRA,i}$: the actual speed one solar rotation ago (optimized), plus i hours; and
- $v_{27,i}$: the actual speed 27 days ago, plus i hours.

An important point is that $v_{OSRA,i}$ and $v_{27,i}$ are actual solar wind speeds and are not derived from PDFs, with the difference being the use of an optimized lag ($v_{OSRA,i}$) or an exact lag of 27 days ($v_{27,i}$).

A simple way to take into account the two speeds $v_{pdf,i}$ ($v_{P1,i}$ or $v_{P2,i}$) and $v_{OSRA,i}$ (or $v_{27,i}$) in the model is to use the expression:

$$v_{pred,i} = a \times v_{pdf,i} + b \times v_{OSRA,i} \tag{3.2}$$

where a and b are parameters that can varied (with $a + b = 1$) to optimize the prediction ability. This model would select a speed from a PDF and an actual speed from approximately 27 days ago in an optimized way.

To summarize, there are four different decisions that can be made to build the model:

1. whether the PDF that is used is based on the current solar wind speed (P1) or the solar wind speed approximately 27 days ago (P2);

2. whether the median or most probable value from the PDFs is used;

3. whether the solar wind speed from the previous rotation is determined using an optimized lag or not; and

4. the value of a and b in Equation 3.2, which determines the reliance on either the predicted speed based on the PDFs ($a = 1$) or the actual speed approximately 27 days ago ($b = 1$).

For example, to predict $v_{pred,i}$, the following parameters could be used:

1. $v_{PDF,i}$ is $v_{P1,i}$;

2. the median of $v_{P1,i}$ is used;

3. $v_{OSRA,i}$ uses an optimized lag; and

4. $a = 0.7$ and $b = 0.3$.

The Median 1 model is the pre-model that uses the median speed determined by P1; the VP 1 model uses the most probable speed determined by P1; the Median 2 model uses the median speed determined by P2; and the VP 2 model uses the most probable speed determined by P2.

Finally, two simple models are presented for comparison: the Persistence model, which predicts that the solar wind speed will be constant over the next five days at the current value; and the OSRA model, which simply predicts the solar wind speed for the next five days will be exactly as it was 27-22 days ago. The OSRA model is equivalent to $b = 1$ and not using an optimized lag above.

3.1.2 Results and discussion

Results of the Different Pre-models

For each of the pre-models and models, the difference was calculated between the actual speed and predicted speed as a function of the prediction horizon (i) and model: $\Delta V_{t,i}^2 = \frac{(v_{pred,i} - v_{data,i})^2}{v_{data,i}^2}$, where $v_{data,i}$ was the real solar wind speed and $v_{pred,i}$ was the prediction, both of which were i hours from the "current" time. The $\Delta V_{t,i}$ was calculated for every hour between the current time and 120 hours into the future (i.e., i=1 to 120), and was done for 6,000 different times between 1995 and 2010 (each time moving one day forward, t=1995 to 2010). The square root of the mean over t (the 6000 times) of $\Delta V_{t,i}$ was calculated, to get:

$$\text{NRMS}_{model,i} = \sqrt{\overline{\Delta V_{t,i}}} \tag{3.3}$$

This is not exactly the same expression as the NRMS as described by Equation 3.1, since the normalization is done before the mean as opposed to after the mean.

As described above, the model can be either one of the pre-models (Median 1, VP 1, Median 2, VP 2), the Persistence model, or the OSRA model. The pre-models can be combined with the solar wind speed from one solar rotation ago through the use of the parameters a and b, which can vary between 0 and 1 (with $a + b = 1$). When combining the pre-models with the solar wind speed from the last solar rotation, an optimum lag can be included or a lag of exactly 27 days can be used. When a=1 and b=0, the prediction is based solely on the PDFs – either the PDFs given the solar wind speed now and the trend (P1) or the PDFs based on the solar wind one solar rotation ago (P2). When a=0 and b=1, the prediction is based solely on the actual solar wind speed from (approximately, depending on the method) 27 days ago. Note that if no optimized lag is used and $b = 1$, the OSRA model is derived. When the values of a and b are between 0 and 1, there is a blending of the techniques.

Figures 3.6 and 3.7 show the $\mathrm{NRMS}_{model,i}$ for many different models with different a and b values. The $\mathrm{NRMS}_{model,i}$ for the persistence model is also indicated as a dotted line on each plot. On the left side are the models with an optimum lag included in $v_{OSRA,i}$, and on the right side are the models without an optimum lag included in $v_{27,i}$.

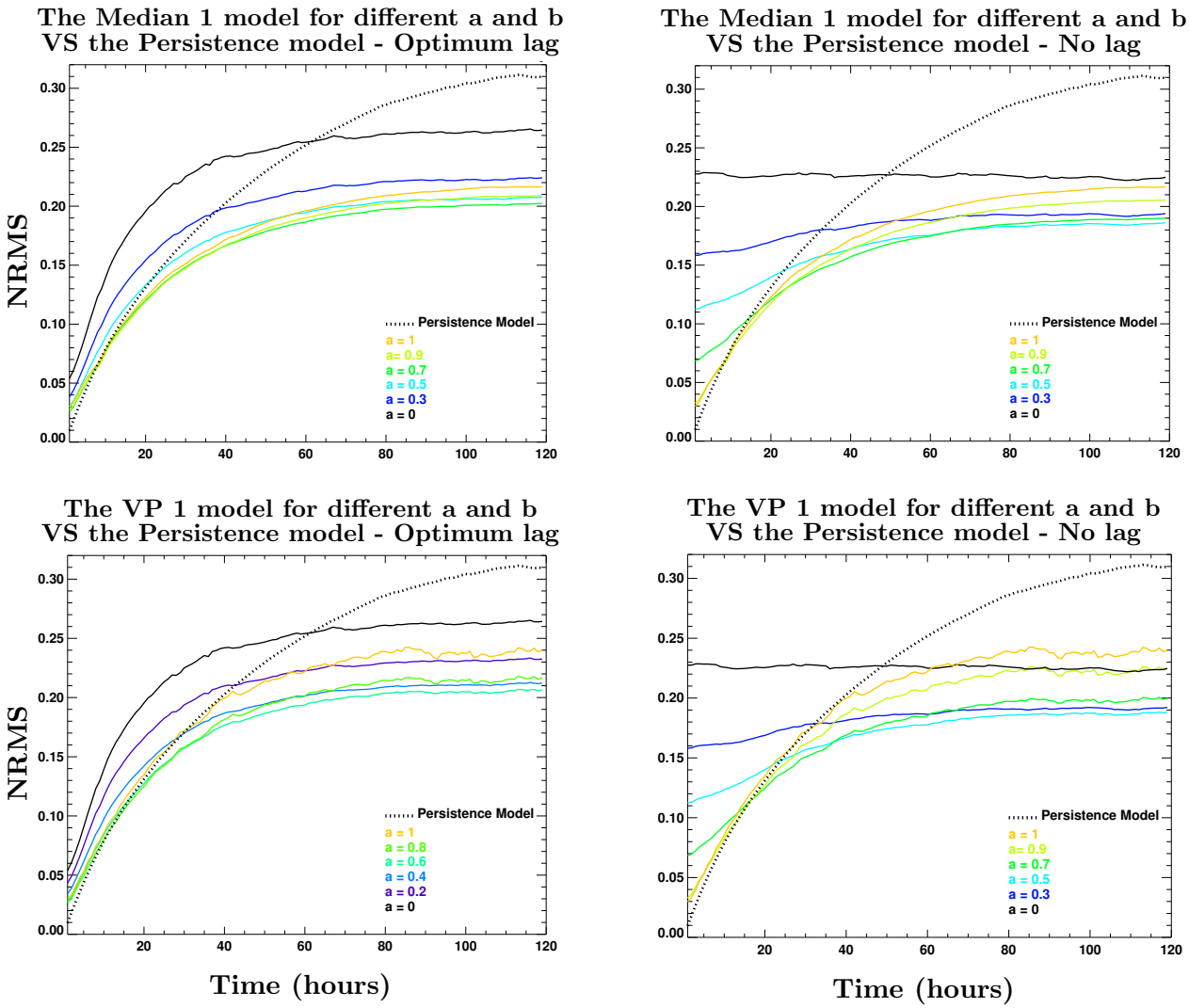

Figure 3.6: The different pre-models based on a linear combination of the P1 PDFs and the solar wind speed one solar rotation ago compared to the Persistence model.

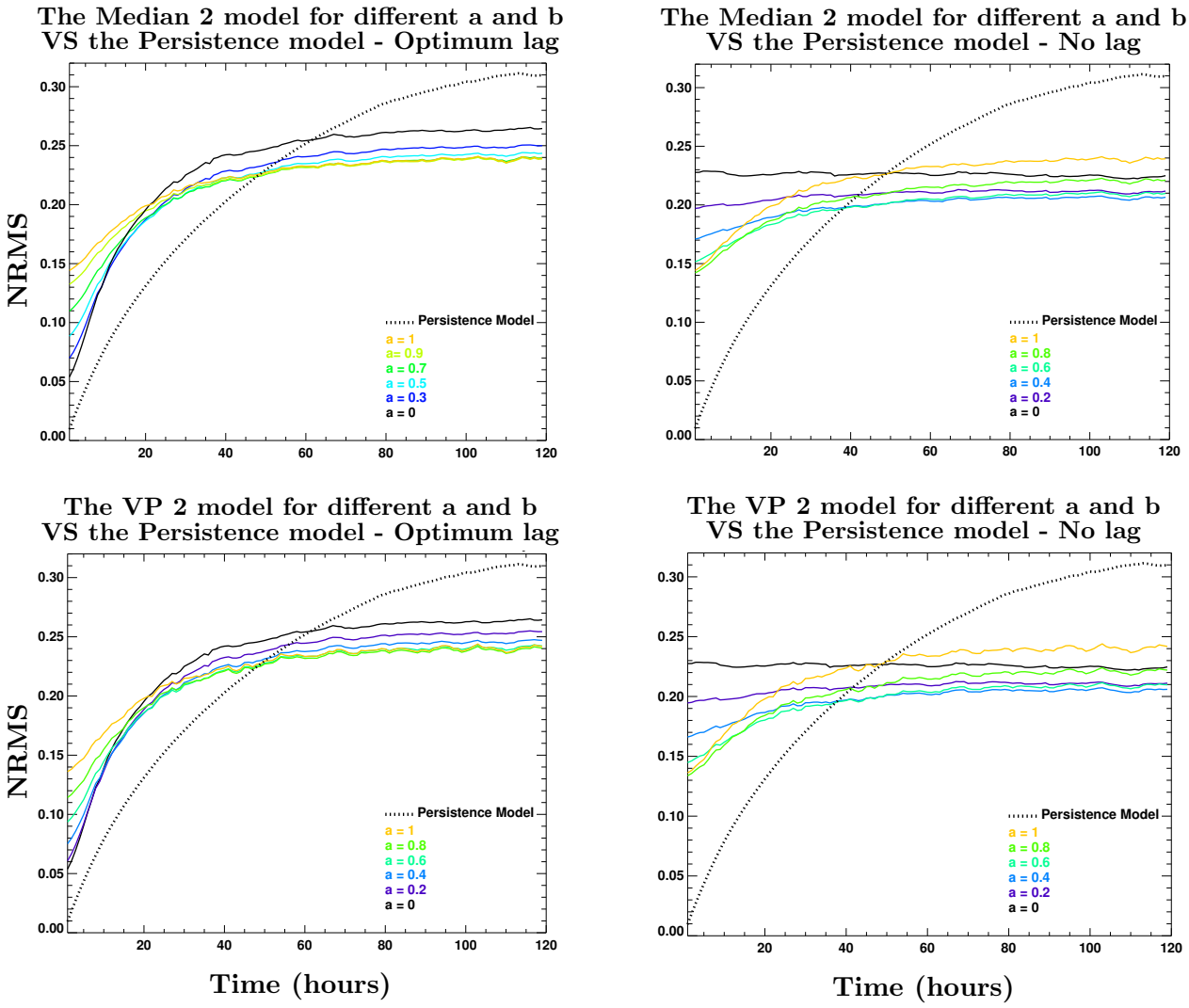

Figure 3.7: The different premodels based on a linear combination of the P2 PDFs and the solar wind speed one solar rotation ago compared to the Persistence model.

The first observation that can be made is that all of the models tend to have a lower $\text{NRMS}_{model,i}$ value than the Persistence model after approximately one to three days. The cases on the right (exactly 27-day lag) have models (a close to 0, or black and blue lines) that have quite poor results for low prediction horizons. This is because the lag might often be far from a perfect 27 days, such that taking the value exactly 27 days ago is often quite a poor choice. The models that use an optimized lag

tend to perform better for shorter prediction horizons, but the models with an exactly 27-day lag tend to perform better for longer prediction horizons, as evidenced by the black line being lower in the right plots than in the left plots. This is because the optimized lag calculation is good for the present time, but rapidly becomes useless, as described above and in Figure 3.2.

Additionally, independent of the lag or which speed is chosen (median vs. most probable), neither the prediction based on the PDFs alone (a=1) nor the prediction based on the previous solar rotation alone (b=1) is optimum. It is a linear combination of the two that provides the best performance. Further, the optimum linear combination changes with time, such that low prediction horizons tend to be predicted best with a closer to 1, while later times tend to be predicted best with a and b close to 0.5.

A subtle feature to note is that the Median values are better than the most probable values over the entire range of prediction horizons. There is almost always a consistent bias between the most probable value and the median value, with the median being larger, as indicated by Figure 3.4.

Finally, none of the models give better predictions than the Persistence model for the first seven hours. The models that are almost as good as the Persistence model are the Median 1 and the VP 1 models (with an optimized lag) with a close to one. All these results mean that the PDF model has to be a combination of the different pre-models and the Persistence model in order to have the best performance.

The Completed PDF model

As described above, the PDF model has to be a combination of the different pre-models and the Persistence model. More precisely, the PDF model predicts the speed in i hours using the pre-models and the Persistence model as follows:

- for $1 \leq i \leq 7$ hours, Persistence model;

- for $i = 8$ hours, Median 1 with an optimized lag and $a = 0.9$;
- for $9 \leq i \leq 12$, Median 1 with an optimized lag and $a = 0.8$;
- for $13 \leq i \leq 17$, Median 1 with 27-day lag and $a = 0.9$;
- for $18 \leq i \leq 32$, Median 1 with 27-day lag and $a = 0.8$;
- for $33 \leq i \leq 51$, Median 1 with 27-day lag and $a = 0.7$;
- for $52 \leq i \leq 89$, Median 1 with 27-day lag and $a = 0.6$;
- for $90 \leq i \leq 120$, Median 1 with 27-day lag and $a = 0.5$.

Using this formulation, the $\text{NRMS}_{model,i}$ of the prediction is minimized. Figure 3.8 illustrates the $\text{NRMS}_{model,i}$ of the final PDF model using this combination of the pre-models and the Persistence model, as well as comparing the PDF model to the Persistence model and the OSRA model. After seven hours, the PDF model gives better predictions than the Persistence model. Additionally, the difference between the NRMS of the PDF model and the Persistence model increases as the time goes by. After two days, the PDF model gives considerably better predictions, and after five days the $\text{NRMS}_{model,i}$ is equal to 0.19 for the PDF model and is equal to 0.30 for the Persistence model. Further, the PDF model is better than the OSRA model all of the time. The Persistence and PDF models are both better than the OSRA model until about 45 hours, at which time the OSRA model becomes better than the Persistence model. The OSRA and PDF models converge to be within approximately 0.03 of each other by the 120-hour prediction horizon, with the PDF model being slightly better.

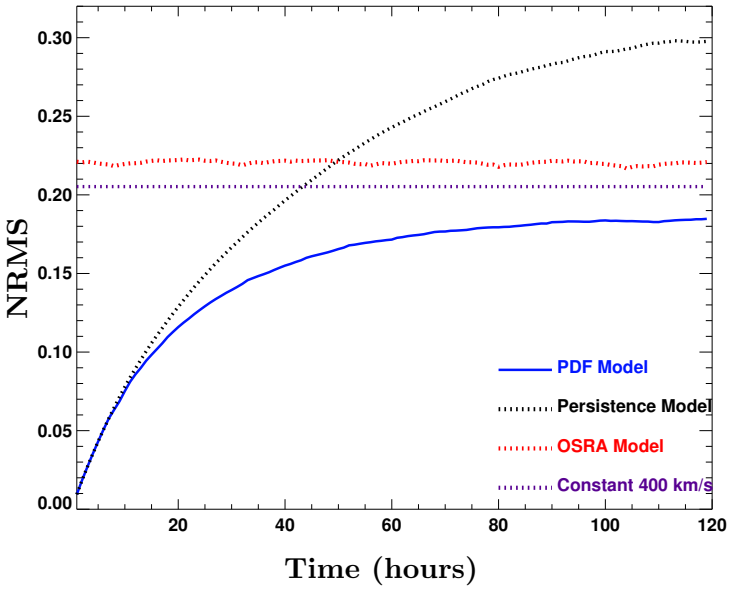

Figure 3.8: A Normalized root-mean-squared comparison between the actual solar wind speed and the PDF model, the Persistence model, the OSRA model, and a constant solar wind value of 400 km/s over the 16 years of this study.

Figure 3.9 shows that the same trend is true for the cross correlation also. The correlation for the PDF model is higher than for the Persistence model after about twelve hours all the way out to 120 hours. The study by *Arge and Pizzo* (2000) showed that the correlation for the WSA model for a five-day prediction horizon was about 0.4, while Figure 3.9 shows that the PDF model has a correlation of about 0.52 for the same prediction horizon, which is quite comparable.

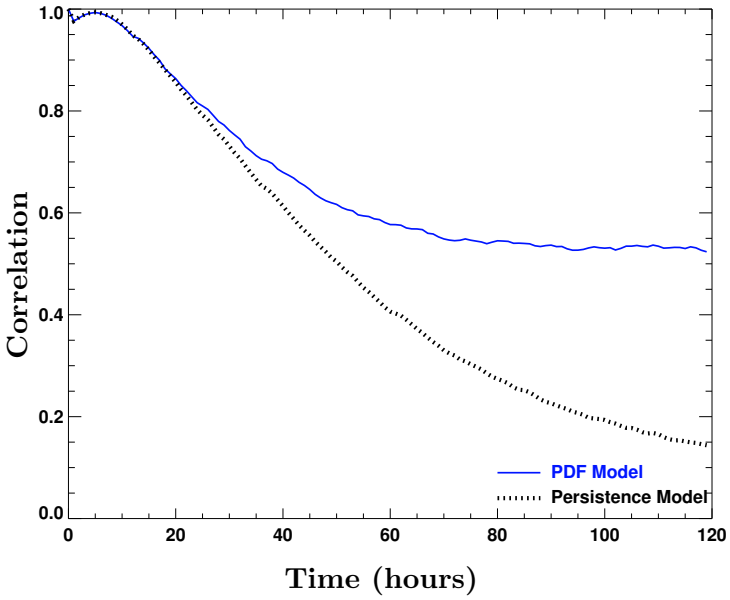

Figure 3.9: Correlation between the PDF model and the observations (blue) and between the Persistence model and the observations (black dash line) over the 16 years of this study.

One fact that might be surprising in this study is that the set of PDFs P2 (i.e., PDFs based on the solar wind approximately 27 days ago) are not used in the final model. The main reason for this is illustrated in Figure 3.2: the determination of the lag is not optimal and is quite difficult to determine accurately. This means that we do not know exactly where to look back 27 days ago to find the time that will best match the solar wind speed in a few days from current time. Therefore, the speed that is used to determine which P2 PDF is not optimal. In other words, the inaccuracy in the lag results in using the wrong P2. If the time delay could be determined accurately, then the technique could go back in time 27 days plus the optimum lag, then move forward i hours and determine the right PDF P2 to use, considering the speed at that time. This bias, not present in the prediction made by the set of P1 PDFs, makes the predictions by P2 worse than the ones by P1.

Two examples of predictions using the PDF model are provided in Figure 3.10. These two examples were chosen to show that even during a strongly increasing (left)

or decreasing (right) solar wind speed, the predictions follow the data quite well. In the first case (left), the solar wind speed increases from about 260 km/s to 450 km/s over the five days. The PDF model roughly captures the increase, although not the exact details of the smaller-scale variability in the speed. In the second case (right), the predictions are very close to the data for the two first days. However, important differences appear around three days and then decrease around five days. The solar wind speed observed by ACE has been averaged using a running average over 11 hours to reflect the trend of the variation and filter the high frequency variability of the solar wind speed.

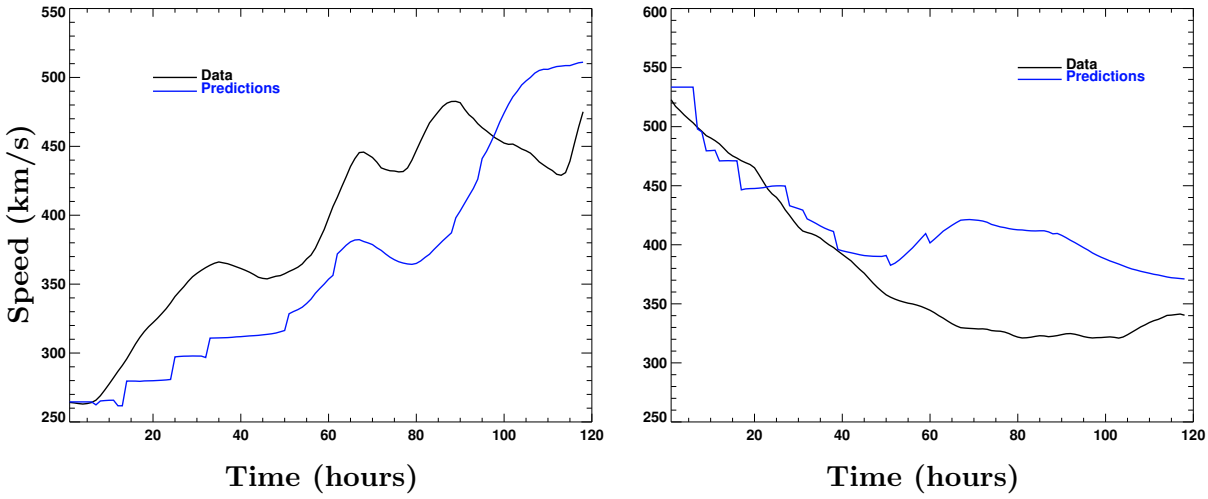

Figure 3.10: Two examples of solar wind predictions based on the PDF model with the prediction in blue and the data in black.

Predicting the solar wind speed during solar maximum is more challenging than during solar minimum. Figure 3.11 shows the variation of the normalized difference in speeds between the PDF model and the measurements as a function of time. As illustrated in this figure, the error in the 24-hour ahead prediction was higher during the solar maximum in 2002 than it was during the two last solar minima, in 1996 and 2008.

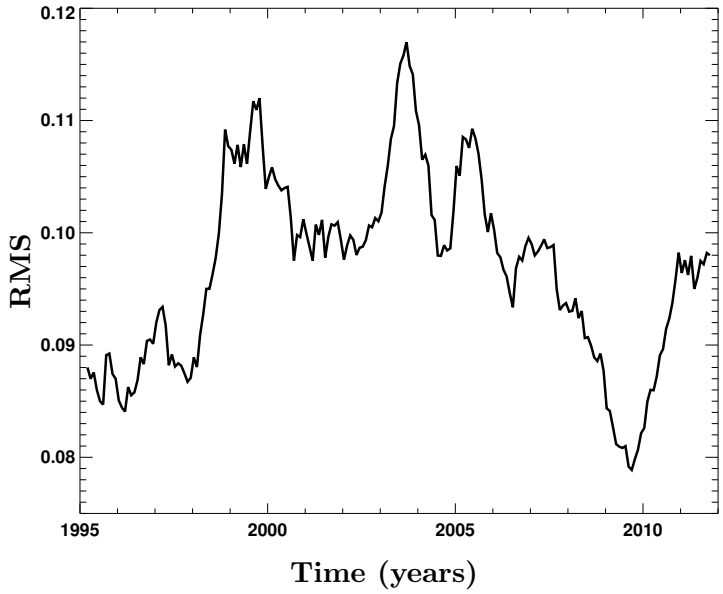

Figure 3.11: The RMS over the two last solar cycles for 24 hours ahead predictions.

The fact that the errors in the predictions made by the PDF model are higher during a solar maximum may be a consequence of two different phenomena: (1) as detailed in *Owens et al.* (2013), no clear correlation in speed between two solar rotations was found during solar maximum, whereas during solar minimum, the variability of the solar wind speed is much more periodic. This implies that using data from one solar rotation ago during solar maximum may not be the best idea. (2) During solar maximum, there are more impulsive events, which are not really accounted for in the PDF model. These impulsive events, such as CMEs, make it so the current speed (preceding the CME) is not a good indicator of the future speed (during the CME). In addition, unless two CMEs occured 27 days apart, the past solar rotation would not be a good indicator of the solar wind speed. This fact shows a weakness in the PDF model: it does not have the ability to predict impulsive events, which is an active area of research.

Figure 3.12 illustrates the variability of the lag (calculated with a one day step) over both a long (solar cycle) and a short (two solar rotations) time scales. There is a great deal of variation in the lag, even on a day-to-day basis. However, it is

interesting to note that this variability increases at the solar maximum and decreases at the solar minimum, making the predictions based on the current lag easier during solar minimum (discussed below). Figure 3.13 quantifies this by showing the monthly standard deviation of the lag as a function of time. The standard deviation is shown to have a minimum in 2007-2008 and a maximum in 2000-2002.

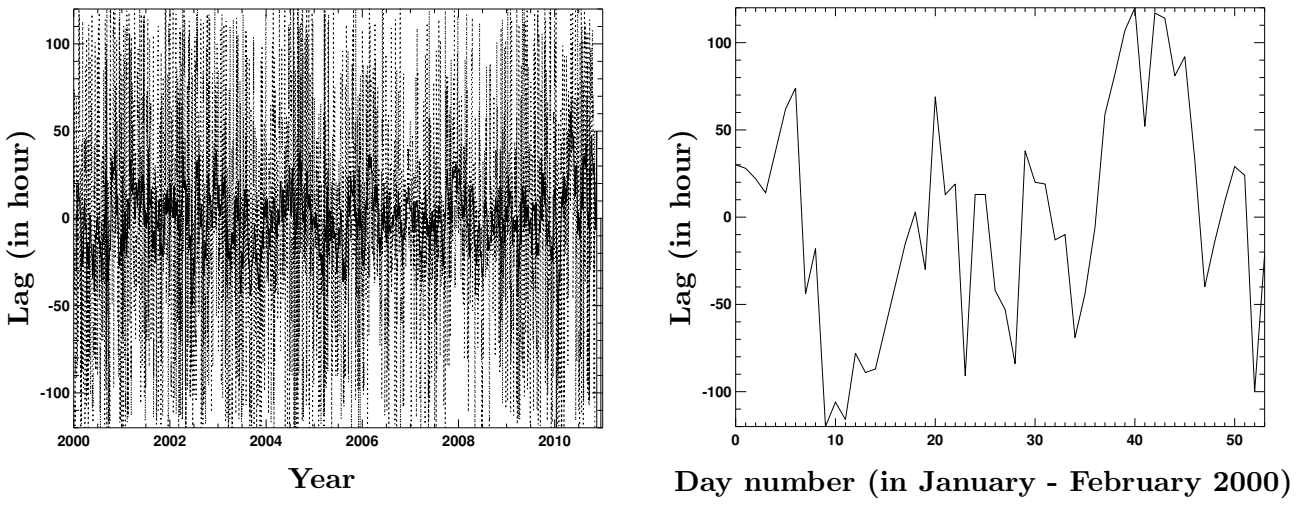

Figure 3.12: The optimum lag between solar rotations over 11 years (left) and over two solar rotations (right).

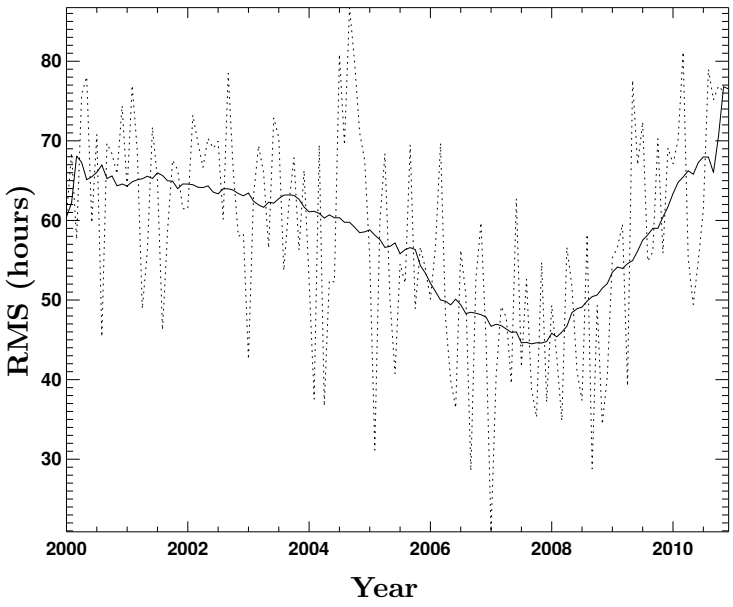

Figure 3.13: The dotted line shows the standard deviation of the lag during each month. The solid line shows a 13 month running average of the standard deviation, so the trend can be determined.

Because the variability of the lag during a solar minimum seems to be less than during a solar maximum (Figure 3.13), the predictions based on the previous solar rotation should be more relevant during a quiet period of solar activity. This is verified in Figure 3.14 where a and b were calculated for a period corresponding to a solar maximum and a period corresponding to a solar minimum (only a is plotted, and b is simply $1 - a$). Recall that a is the weight of the current solar wind speed, while b is the weight of the last solar rotation. One can notice that in 2008, the value of a is decreased for all prediction horizons (after seven hours since the seven first hours correspond to the Persistence model), showing that using the last solar rotation provides a more accurate solution. However, a is increased in 2002, reflecting the high variability of the solar speed during such a period, and therefore predictions that have to rely mostly on the predictions based on the current solar wind speed.

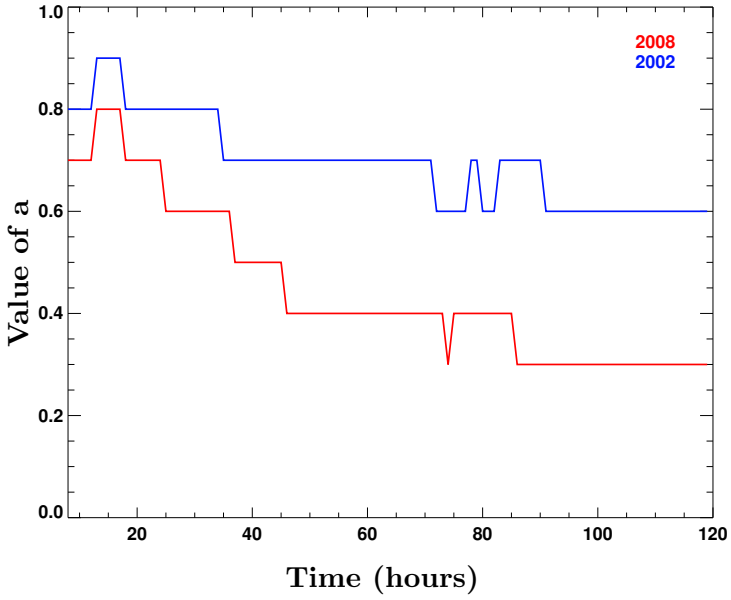

Figure 3.14: Coefficient a in the PDF model: during a solar maximum (blue) and during a solar minimum (red). Coefficient b is equal to $1 - a$.

While it is impossible to estimate the optimum lag at this time, if the optimum lag was able to be forecast, the predictions would dramatically improve. Figure 3.15 compares the NRMS of predictions that use an optimum lag to the NRMS of predictions made by the OSRA model, with only a 27-day lag. The error is more than four times lower with an optimum lag. Moreover, the corresponding NRMS is much lower than any other model: the difference between the prediction and the data is about 5%, which means that for a typical solar wind speed of 400 km/s, the error would be around 20 km/s only. This shows how important finding an optimum lag is, and points to a need to have studies explore how to determine this optimum lag.

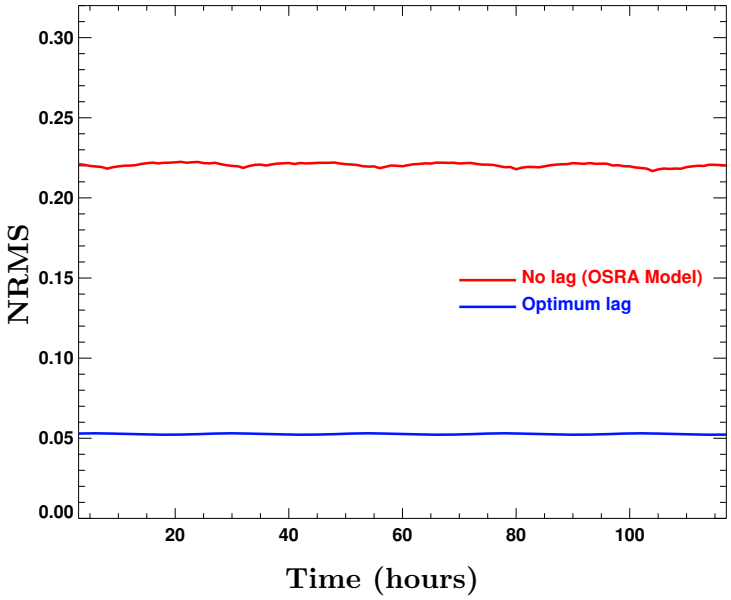

Figure 3.15: The NRMS if the lag was optimum compared to the NRMS without any lag.

One idea that is being explored for the next generation of the PDF model will be to look at structures on the Sun, such as sunspots, and compare them to the same structures one solar rotation ago. This will enable us to find the lag so that the two sets of structures are superimposed on each other the best. This lag may be the one to take into account for the predictions in three days, time for the solar wind to travel from the Sun to the Earth.

Direct Comparison to the WSA model

The PDF model has been directly compared to the WSA model. The WSA model has been run for two whole years, 2008 and 2011, to make one to five days ahead predictions that have been compared to the same predictions made using the PDF model.

Examples of predictions using the PDF and WSA models are shown in Figure 3.16. They correspond to one to five days ahead predictions for June 10 - July 10, 2011. Two main observations can be made. First, the one-day ahead predictions using the

PDF model get the sudden increase in speeds around the 25th of June (even if they underestimate the increase for the first two days), while the WSA model does not. Then, both the PDF and WSA models miss the variability of the solar wind speed under short time scales. The average structure of the speed is conserved but the high frequency variability is lost.

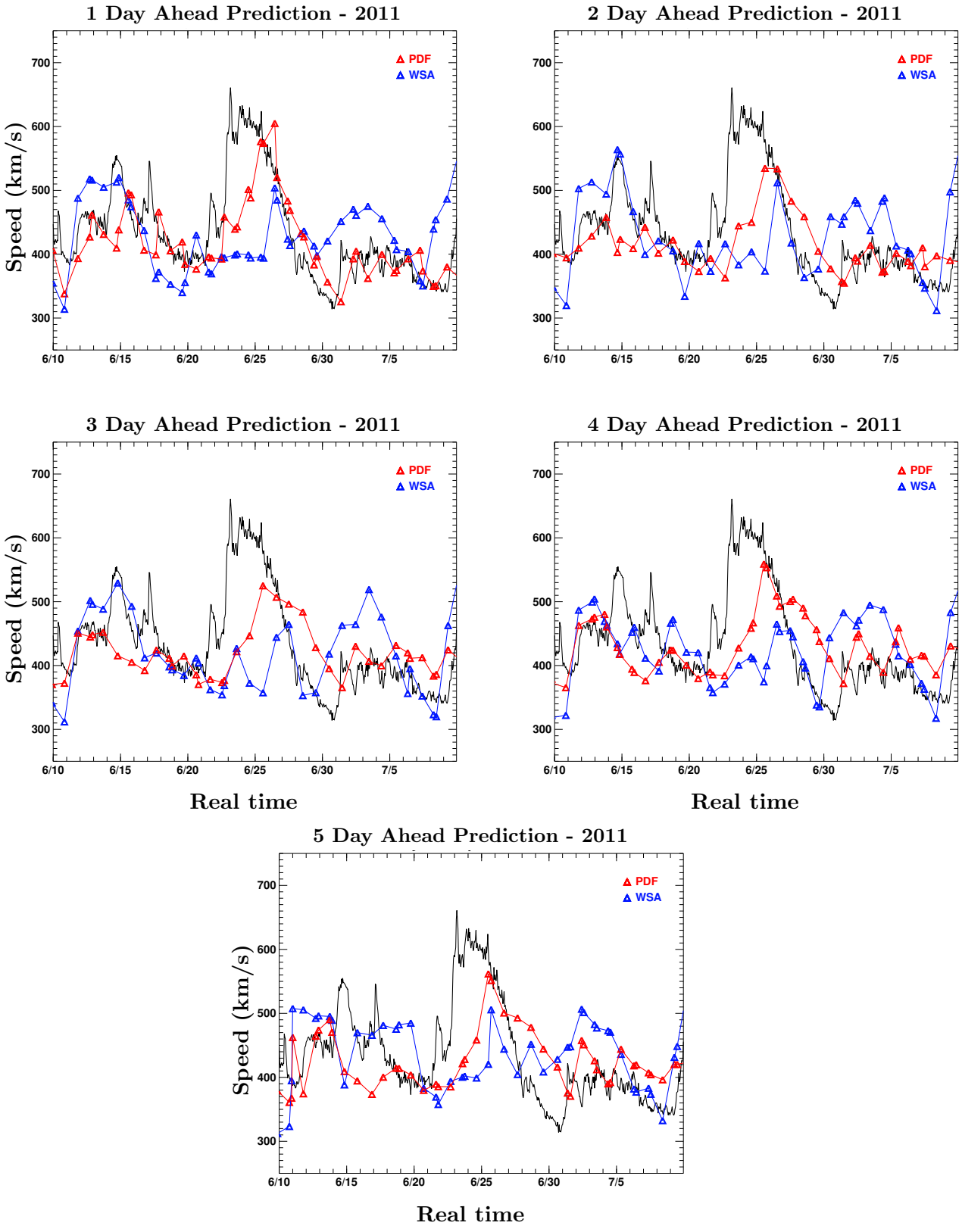

Figure 3.16: Example of predictions made by the PDF and the WSA models for June 10 - July 10, 2011. The actual data is shown as a solid black line, while predictions (one or two per day) by the PDF and WSA models are plotted in red and blue, respectively. One - through five - day ahead predictions are shown.

The RMS difference between each model and the observations by the ACE satellite were calculated. The results are presented in Table 3.2. The year of 2011 is right between a solar minimum (2008) and a solar maximum ($\sim$ 2013). Looking at the PDF model column of this table, the RMS starts off small on the first day, then grows to a maximum value on day 3, at which time it asymptotes. This is consistent with Figure 3.8. Additionally, the Persistence model has a similar trend, but asymptotes to a higher level, also consistent with Figure 3.8. However, the WSA model has RMS errors that are approximately constant for all prediction horizons. These values are close to the value that the PDF model asymptotes to, indicating that the PDF model is most likely slightly better than the WSA for the first (approximately) two days, then is quite similar to the WSA model. This is consistent with the study made by *MacNiece* (2009), which showed that persistence is better than the WSA model for the first two days of prediction. Table 3.2 also shows that for the last solar minimum that occured in 2008, the PDF model performs better than the WSA model: the RMS is less by about 15 km/s for every prediction horizon. These results are consistent with the study by *Owens et al.* (2005), which found that the Root Mean Square Error between the WSA model and the data is better at solar maximum than at solar minimum.

Table 3.2: The RMS between the PDF, WSA, and Persistence models and the observations by ACE for 2008 and 2011.

Prediction Horizons	RMS PDF (km/s)		RMS WSA (km/s)		RMS Persistence (km/s)	
	2008	2011	2008	2011	2008	2011
1-day ahead	83	66	107	88	79	68
2-day ahead	93	83	105	84	119	100
3-day ahead	90	89	103	88	147	118
4-day ahead	88	88	105	91	163	128
5-day ahead	90	88	107	90	173	130

3.1.3 Conclusion

From this study a few different conclusions can be drawn:

1. The solar wind speed is quite periodic, but the period is not always exactly 27 days; it changes on a day-to-day basis and can vary between approximately ± 5 days of the mean 27-day period. This makes it quite difficult to predict the next five days using data from the previous solar rotation. Indeed, using an optimized lag is best for the first 12 hours of prediction; afterwards, using a straight 27 days for the lag is best. Moreover, keeping the speed constantly equal to the median speed of the solar wind, $400 km/s$, gives a constant NRMS better than the NRMS calculated from the last solar rotation without any lag.

2. The solar wind speed typically changes quite slowly, such that using the current solar wind speed is the optimum prediction for the first seven hours.

3. The current solar wind speed, as well as the trend in the solar wind speed (speeding up or slowing down) allows creation of probability distribution functions for the solar wind speed i hours into the future. The width of the PDF increases as time goes on, while the peak in the PDF decreases. The PDFs are narrower and taller than the distribution function of the solar wind in general, which means that these PDFs are more useful for predictions, than just assuming the solar wind distribution.

4. These PDFs can be used to generate ensemble prediction scenarios for the solar wind speed or to assign an uncertainty on the prediction based solely on the median speed.

5. Using a linear combination of the medians of the PDFs based on the current solar wind speed and trend as well as the actual solar wind speed from approximately 27 days ago allows the best predictions. This linear combination is

highly weighted towards using the current value of the solar wind for low prediction horizons and using both roughly evenly for larger prediction horizons. It is shown that this linear weighting can change over the solar cycle, with more weighting on the previous rotation during solar minimum conditions.

6. The final PDF model performs equal to or better than the Persistence model for all times up to a five-day prediction. The further out the prediction, the better the PDF model does compared to the Persistence model. After five days, the difference in the NRMS of 0.11 between the two models reflects an improvement of the accuracy of the prediction of 40 km/s for a typical solar wind velocity of 400 km/s. The model also performs better than simply taking data from 27 days ago, although after approximately three days, the differences between the two levels off, with the PDF model then being slightly better than simply using the data from 27 days ago.

7. The comparison of the PDF model to the WSA model predictions made for 2011 showed that the final PDF model performs better than the WSA model for one-day ahead predictions. For longer prediction horizons, both models perform about the same. For 2008, the last solar minimum, the PDF model performs better than the WSA model for all prediction horizons, with a 15 km/s difference in the accuracy of the predictions.

8. The predictions with the PDF model give better results during the last two solar minima (in 1996 and 2008) than during the solar maximum in 2002.

9. If the lag between solar rotations was predicted more accurately, then it would improve the predictions of the PDF model. The current errors vary from 10% (predictions in one day) to 20% (predictions in 5 days), but with a perfect knowledge of the lag, they could decrease to values around 5%. It is unclear how to determine the perfect lag, though.

The main flaw of the first version of the PDF model is that it is not able to predict transitions from slow to high solar wind speed. The cause is that $v_{pdf,i}$ is based only on the current speed and the slope in the twelve previous hours. This study showed that statistically the majority of times when the slope of the solar wind speed is positive, the speed starts to decrease after a few hours. What the PDF model misses is that when the slope is above certain levels, the speed rises to a peak above 500 km/s and is considered a high speed event. Therefore, when a shock reaches the Earth, the PDF model does not predict the speed reaching above the high speed event cutoff of 500 km/s.

Consequently, the PDF model was improved to predict rapid increases in speed. Since CMEs often happen randomly and are not often associated with a similar event that occurred roughly twenty-seven days ago, the One Solar Rotation Ago (OSRA) model should not be used for the prediction of such transitions. Indeed, even when an event repeats over one rotation, this study showed that the time between the two events is not exactly one solar rotation but that there is an unpredictable and highly variable lag. Therefore, version 2 focuses on predictions using only PDFs, which means that $a = 1$ and $b = 0$ in Equation 3.2.

3.2 The Probability Distribution Function model (version 2)

3.2.1 Current models of solar wind speed peak predictions

The space weather forecasting community has a particular interest in the prediction of transitions from slow to fast solar wind (*Wright et al.*, 1995). A strong increase in the speed leads to enhanced geomagnetic activity that affects the near-Earth space environment. Such transitions can be associated with shocks or high speed streams, which can be a consequence of a CME (*MacNeice*, 2009a).

Many models have been created and refined to predict the magnitude of the solar wind speed increase and its arrival time at the Earth. The Interplanetary Shock Propagation Model (ISPM) (*Smith and Dryer*, 1990) and the Shock Time of Arrival (STOA) model (*Dryer*, 1974) estimate the propagation time of the shocks from the Sun to the Earth and are based on type II meter wave burst data. The Hakamada-Akasofu-Fry (HAF) kinematic solar wind model has the ability to predict the speed, the solar wind density, and the IMF. *Fry et al.* (2001) compared the predictions of shock arrival times made by HAF with the predictions made by the IPSM and the STOA models for 36 events between 1997 and 1999. They defined the shock arrival time as the time when the ram pressure rose above the background ram pressure by a certain threshold. They calculated the contingency table for the 36 events and showed that HAF (version 1) had comparable skills to the two shock models. *Smith et al.* (2009) evaluated the skills of HAF (version 2) during the declining phase of Solar Cycle 23 (2002 to 2006) for predictions of shock arrival times and compared them with predictions during the ascending phase (1997 to 2000, *Fry et al.* (2003)) and the maximum of the solar cycle (2000 to 2002, *McKenna-Lawlor et al.* (2006)). They predicted particular events and compared the shock arrival time and the strength (given by the ratio of the maximum pressure and background pressure) of the predicted shocks to the observed shock at 1 AU. Comparing the number of hits, misses, and false posi-

tives, they concluded that the predictions were a bit improved for the ascending and maximum conditions, as opposed to the solar minimum conditions. *McKenna-Lawlor et al.* (2006) compared predictions of 173 shocks during the maximum phase of Solar Cycle 23 by STOA, ISPM, and the HAF model (version 2). The accuracy of a model can be defined as the sum of hits (i.e., the number of times the model correctly identified shocks) and correct nulls (i.e., the number of times the model correctly identified non-shock time periods) divided by the total number of predictions, times 100. In that study, the STOA, ISPM, and HAF v-2 models received scores of 54%, 60%, and 52%, respectively, for a hit window (size of the window in which the event was allowed to be observed in) of $\pm$24 hours. The number of events classed as hits decreased when the window size was $\pm$12 hours, so the accuracies were lower: 44%, 54%, and 42%, respectively. For a window size larger than $\pm$24 hours, the results were not relevant, because the shock might not have been related to the specific solar event. However, the accuracy is not the best measure to assess the ability of a model in predicting rare events: a high accuracy could mainly be due to a large number of correct nulls, which is not crucial for forecasters, who are mostly interested in hits. The threat score, defined as the number of hits divided by the sum of hits, misses (i.e., the model did not predict an event that actually occurred), and false positives (i.e., the model predicted an event that actually did not occur), is a more suitable measure. Similar to the accuracy, it does not take into account the number of correct nulls (*Buizza*, 2001).

Vandegriff et al. (2005) introduced a model that predicted shock arrival times. The method relied on the energetic particle intensity data. It was able to predict the shocks 24 hours in advance with an accuracy of 8.9 hours, and 12 hours in advance with an accuracy of 4.6 hours.

Kim et al. (2010) developed a model to predict the occurrence of geomagnetic storms and their strengths, which were represented by the Dst index. The prediction

was based on the CME parameters (southward orientation of the magnetic field in the CME source region, asymmetry of the CME shape, and CME source location). They found that the predictions of the Dst index were far from the observed index, but the model showed good skill at predicting when a geoeffective event may occur, based on when the Dst index calculated by the model became less than -50 nT.

Finally, the WSA model was tested for the predictions of high-speed enhancements. These correspond to abrupt transitions from slow to fast solar wind. *MacNeice* (2009a) presented an algorithm to characterize this type of transition and tested the performance of the WSA model in predicting these events. He answered the question: "[I]f the WSA model predicts or does not predict a high speed event in the next 24 hours, what is the probability that a high speed event will or will not occur?" (*MacNeice*, 2009a) and found that 29% of the positive forecasts and 89% of the negative forecasts were accurate, showing an improvement compared with the study by *Owens et al.* (2005). *MacNeice* (2009b) provided a similar study but used other sources of magnetograms. Averaging over the three sources of magnetograms, he found that a positive/negative forecast was correct 17%/94% of the time.

Forecasters need to be able to understand the level of confidence they can place in the predictions. A single value for the prediction of the magnitude of the speed and the time when the increase occurs needs to be associated with an estimation of the error made on the prediction. Ensemble forecasts can quantify this uncertainty. Few studies have been published regarding this issue. The WSA model, coupled to the ENLIL model, has been used to start to perform ensemble modeling of the background solar wind at the Community Coordinated Modeling Center since 2009 by taking the average of the two models' output for each of the solar wind parameters (*Taktakishvili et al.*, 2009). *Emmons et al.* (2013) coupled the WSA, ENLIL and CONED models to create ensemble forecasts for 15 CMEs. They compared the predicted distributions of the propagation time and the Kp index (obtained from 100 sets of CME cone

parameters derived from the CONED model) with the observations and found that out of the 15 events, 5 (33%) had propagation times that were within the ensemble average plus or minus one standard deviation, and 8 (53%) had propagation times within the range of the ensemble, while 10 (67%) had maximum Kps that were within the range of the ensemble. *Mays et al.* (2015) found that the arrival time was within the range of the ensemble for 8 out of 17 events (47%). They also showed that the accuracy of the predictions depends on the CME input parameters and that the ensembles do not sample a wide enough spread in these parameters. *Cash et al.* (2015) used an ensemble approach to study the effect of the input parameters. They considered the initial speed, the angular width, the direction, and the ambient solar wind background and put in evidence their influence on the predicted arrival time of a CME that occured in July 2012. Goddard's Space Weather Laboratory received support in 2012 to implement such ensemble forecasts.

This study presents an improvement of the PDF model (*Bussy-Virat and Ridley*, 2014) (Section 3.1) for predictions of transitions from slow to fast solar wind. The older model did not have this ability as it was based on all solar wind data over 15 years rather than focused on only solar wind speed increases. The model now is able to predict these transitions when there is an increase in the slope of the solar wind speed and performs better when transitions are associated with strong variations in the magnitude of the IMF or the density. This includes predictions of shocks. The way that this study characterized such transitions was very similar to the one presented in *Owens et al.* (2005) and *MacNeice* (2009a). The model presented here predicts several features of the peak in speed, including the amplitude of the peak as well as the time when the maximum occurs. It also provides ensembles of predictions of each peak, which enables forecasters to have an estimation of the uncertainty in the prediction, or the possible spread in the characteristics. We compare the predictions with the observations made by the Advanced Composition Explorer satellite and compare the

accuracy of the model with the previous studies that predicted such transitions, in particular those made by the WSA model (*MacNeice*, 2009a,b; *Owens et al.*, 2005, 2008; *Emmons et al.*, 2013).

3.2.2 Methodology

3.2.2.1 Identification of high speed events

An algorithm was created to identify transitions from slow to fast solar wind, named high speed enhancements (or high speed events). It is similar to the one presented in *MacNeice* (2009a) because it detects any sharp transition above a certain speed gradient. Figure 3.17 (left) shows an example of a single peak and key times within the peak for reference. The peak-finding algorithm can be summarized as:

1. The solar wind speed is averaged with an 11-hour running average to filter out the high-frequency variability of the solar wind.

2. The algorithm flags any increase in speed by more than 50 km/s in less than 24 hours, noting the time in which this occurs as $t_{beginning}$.

3. The end of the peak is identified as the time when the speed decreased by 50 km/s in less than 36 hours and is noted as t_{end}. This less sharp gradient is a consequence of a study (not shown here) that showed that the slope is typically less sharp in the descending phase of the peak than in its ascending phase.

4. The time of the maximum of the peak is noted as t_{peak}.

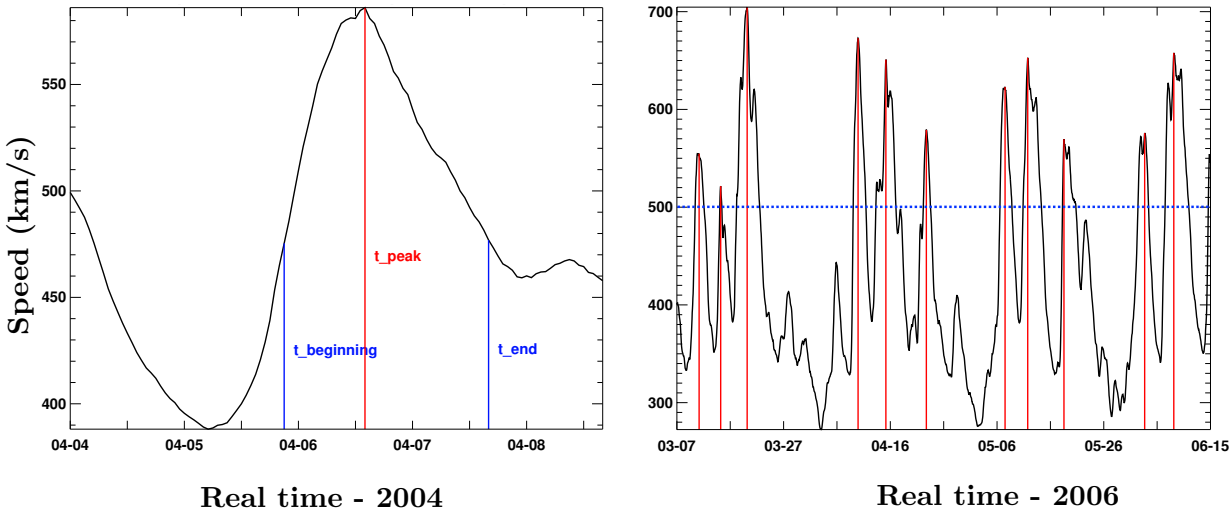

Figure 3.17: Example of the detection of a peak in April 2004 using a minimum gradient of 70 km/s in 24 hours (left); example of the detection of peaks in March-June of 2006 adding a threshold at 500 km/s for the maximum value of the peaks (right).

The algorithm allows for any modification in the choice of these gradients (50 km/s in 24 hours can be changed to 70 km/s in 30 hours, for example) as well as setting a threshold on the minimum value of the maximum speed. Examples of peaks found by the automated algorithm are shown on Figure 3.17 (right). Vertical red lines are drawn when peaks are found by the algorithm using a minimum gradient of 50 km/s in 24 hours and adding a threshold at 500 km/s on the minimum value of the peak. We can notice that the algorithm finds every peak in this three-month period.

For this study, peaks with gradients larger than 50 km/s in 24 hours were investigated.

3.2.2.2 New Probability Distribution Function model

Forecasters need to know when the high speed events will occur as well as the magnitudes. In this regard, several approaches were investigated. The first idea was to relate the amplitude of the peak (defined as the difference between the minimum

and maximum speed in the peak) to the maximum slope in the ascending phase of the peak (noted s_{max}). Figure 3.18 shows the distribution of the amplitude of 761 peaks detected by the algorithm as a function of the maximum slope. The slope is calculated with a step of two hours:

$$\text{slope}_i = \frac{v_{i+1} - v_{i-1}}{2} \tag{3.4}$$

$$s_{max} = \max([\text{slope}_0, \text{slope}_1, \text{slope}_2, ..., \text{slope}_n]) \tag{3.5}$$

where slope_i is the slope i hours after the beginning of the peak (in km/s/hour), n is the number of hours until the maximum of the peak, and $v_{i\pm1}$ are the speeds $i \pm 1$ hours after the beginning of the peak.

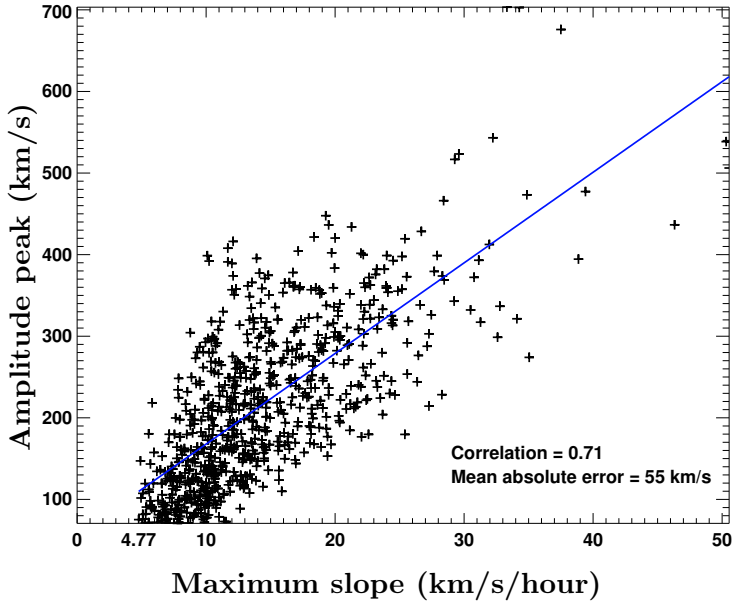

Figure 3.18: Amplitude of a peak in function of the maximum slope.

The correlation between the maximum slope and the amplitude was 0.71 with a mean absolute error between the data and a linear fit of 55 km/s. Distribution functions of the amplitude were made for each bin representing a range of slopes of 2 km/s/hour. Note that the minimum value on the x-axis of Figure 3.18 is 4.77

km/s/hour: every peak found by the algorithm between 1995 and 2012 had a maximum slope greater than 4.77 km/s/hour. The PDF model uses these distributions for slopes below 10 km/s/hour. Although the exact value of 10 km/s/hour is arbitrary, for slopes above 10 km/s/hour, events get more rare so not enough data can be used to make such accurate distributions. Therefore, the PDF model uses Gaussian distributions with a mean value equal to the amplitude given by the linear fit (blue curve in Figure 3.18) and a full width at half maximum of 110 km/s. Figure 3.19 shows the distribution of the peak amplitudes given a maximum slope of 6 km/s/hour (left) and 20 km/s/hour (right). From these distributions, the median and the deviations from this median (10%, 25%, 75%, and 90%) were determined. In the first case, the distribution of the predicted amplitudes has a median at 88 km/s, and in the second case the median is at 275 km/s. This higher value can be explained by the fact that the maximum slope detected in the ascending phase is almost four times higher than in the first case. When the model detects that a peak is occurring, the maximum slope is determined and the predicted peak amplitude (noted Δv) of the solar wind speed is derived, given Figure 3.18 and the distribution functions described above, to specify the uncertainty on the predicted amplitude. This method is used to determine the speed of the peak (v_{peak}) and the uncertainty in the speed of the peak.

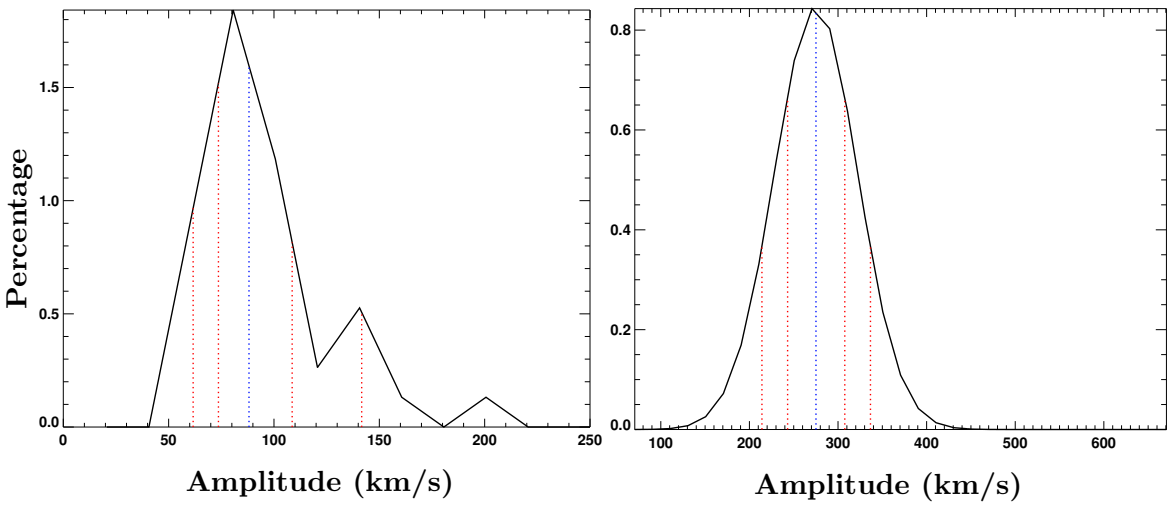

Figure 3.19: Example of a PDF (left) and a Gaussian (right) distribution of the amplitude of the peak for a maximum slope of 6 km/s/hour (left) and 20 km/s/hour (right): the red dash lines represent the 10%, 25%, 75%, and 90% quartiles, the blue dash line represents the median.

The time of the peak is important and can be deduced in two different ways. The first relies on the solar wind speed continuing to increase at the maximum measured rate to the peak amplitude, as described above. The time (further noted $t_{75\%}$, the meaning of which will be described below) can thus be derived by dividing the amplitude by the slope s_{max}: $t_{75\%} = t_{current} + (v_{peak} - v_{current})/s_{max}$. However, it was noted that this time occurs earlier than the actual time of the peak. The speed typically increases quickly, then the gradient decreases significantly, resulting in a peak that lasts longer than expected if only the maximum slope is considered. Therefore, a second method was explored to find the actual time of the peak.

This method relies on the PDF of the time it takes for the speed to increase from the beginning of the peak (at $t_{beginning}$) to the maximum of the peak (at t_{peak}). A study not presented here showed that although there is a relationship between the amplitude of the peak and the maximum slope (Figure 3.18), there is not a clear relationship between $t_{peak} - t_{beginning}$ and the maximum slope. The distribution of times for the 761 peaks is shown in Figure 3.20. It is sharp enough and the most

probable value is high enough for the prediction based on this method to be valuable. The median of the distribution is 23 hours, meaning that 50% of the time, the speed will take less than 23 hours to increase from the speed at $t_{beginning}$ to the speed at t_{peak}. Therefore, the predicted time from the beginning to the maximum of the peak (noted Δt) is specified as 23 hours, with an uncertainty distribution as specified in Figure 3.20. This method is used to determine the time of the peak (t_{peak}) and the uncertainty in the time of the peak.

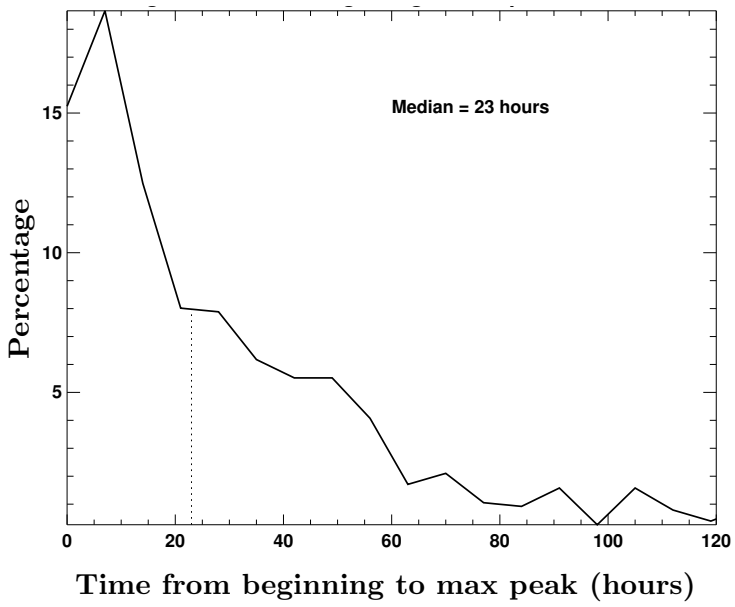

Figure 3.20: Histogram of the time from the beginning of the peak to its maximum.

Combining the two methods described above and in Figure 3.21, v_{peak} ($= v_{current} + \Delta v$) occurs at t_{peak} ($= t_{current} + \Delta t$). When the rapid rise ends (at $t_{75\%}$), the peak is assumed to have reached 75% of its maximum speed: $v_{75\%} = v_{peak} - 0.25 * (v_{peak} - v_{current})$. Two polynomial expressions of order 3 link the speeds $v_{current}$ to $v_{75\%}$ (at $t_{75\%}$) and $v_{75\%}$ to v_{peak} (at t_{peak}), so that the slopes before and after $t_{75\%}$ are continuous and the slope at t_{peak} is 0, which is the definition of a maximum. Specifically, the following algorithm is used:

1. The maximum slope s_{max} is calculated in the ascending phase (from $v_{data,min}$ to

$v_{current}$). Using Figure 3.18 and the distribution functions described in the first method, the PDF model derives Δv to find v_{peak}: $v_{peak} = v_{data,min} + \Delta v$.

2. The time and speed when the rapid rise ends are calculated from v_{peak} and s_{max}: $t_{75\%} = t_{current} + (v_{peak} - v_{current})/s_{max}$ and $v_{75\%} = v_{peak} - 0.25 * (v_{peak} - v_{current})$. The first polynomial expression calculates the speeds that occur between $t_{current}$ and $t_{75\%}$ (blue dashed line in Figure 3.21).

3. The time t_{peak} when v_{peak} is reached is derived from $t_{beginning}$ (increase by 50 km/s in less than 24 hours) and the PDF of Figure 3.20. The second polynomial expression calculates the speeds that occur between $t_{75\%}$ and t_{peak} (red dashed line in Figure 3.21).

4. Once the ascending phase has been predicted, the model uses the methodology described in *Bussy-Virat and Ridley* (2014) (Section 3.1) to predict the descending phase of the peak (black dashed line).

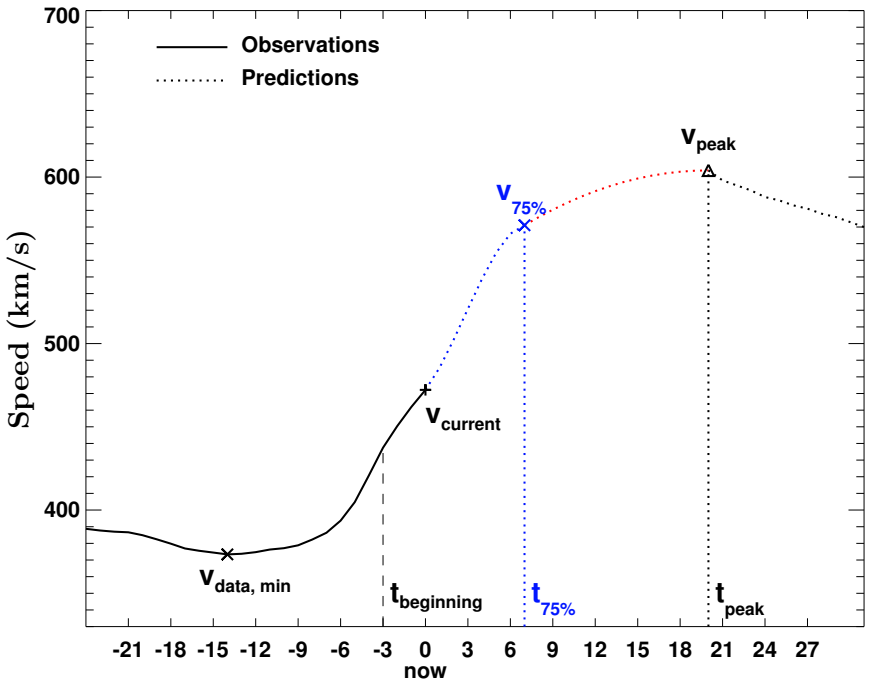

Figure 3.21: Calculation of the speeds from the current time to the time the speed is predicted to reach its maximum value.

Examples of predictions made by the model are shown in Figure 3.22. For example, in the top left graph the actual speed (black line) increased from 440 km/s to 600 km/s in less than 15 hours. The peak was predicted (blue line) to occur in 24 hours from the start time with an amplitude of 110 km/s (above the current solar wind speed), which corresponded to an underestimation of the peak speed. The descending phase of the actual speed followed the predicted descending phase very closely. Moreover, the actual speed was in between the 25% and 75% quartiles 108 hours out of the 120 hour prediction horizon and in between the 10% and 90% quartiles 115 hours out of the 120 hours. The other examples illustrate the variety of results that the PDF model predicts, as well as the actual solar wind speed data.

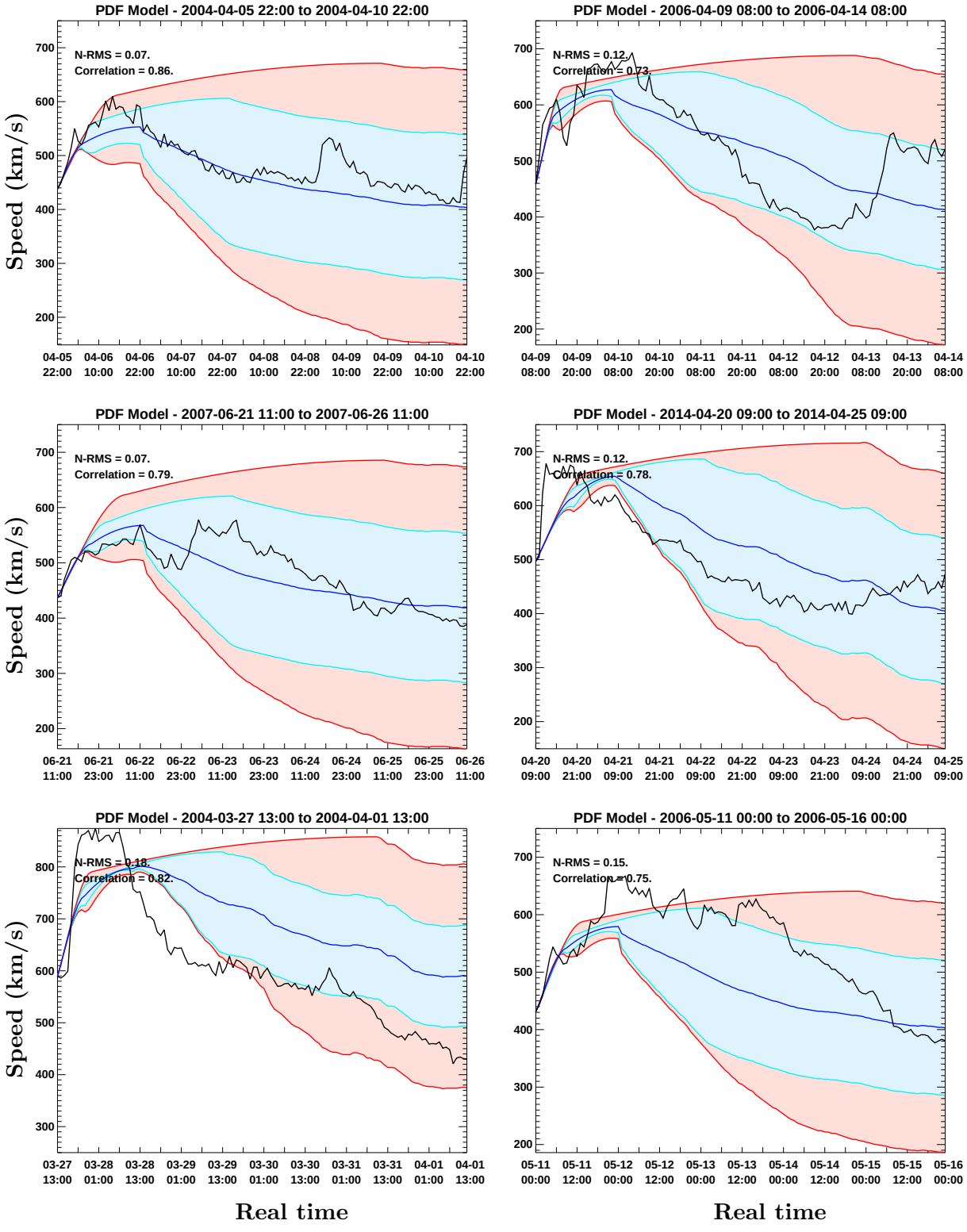

Figure 3.22: Examples of predictions with the PDF model version 2. The black line is the actual speed; the blue line is the median of the PDFs; the light blue lines are the 25% and 75% quartiles (interquartile range); the red lines are the 10% and 90% quartiles.

3.2.2.3 Comparison with peaks in the Interplanetary Magnetic Field and the density

A sudden large increase in the speed could be due to either a high speed stream or could be the result of a shock in front of a CME that reaches the Earth environment. In such events, the IMF magnitude and density should also increase before the velocity peaks, since speed increases can act as a "snow plow" on the density and IMF. The same algorithm that characterized a peak in the speed was also applied to find peaks in the IMF magnitude and the density. The thresholds on the IMF and density gradients were 4 nT/day and 8 cm^{-3}/day, respectively. A threshold was added on the minimum value for the peaks in IMF and density: 6 nT and 10 cm^{-3}, respectively (i.e., any peak under 6 nT and 10 cm^{-3} were not considered). Once peaks in the solar wind density and IMF magnitude were found, they were compared with peaks in the solar wind speed.

Figure 3.23 (left) shows the histogram of the lags between the maxima of a peak in the IMF magnitude and the maxima in the speed, with a bin size of 3 hours. The maximum in the speed typically occurs 10-30 hours after the maximum of the peak in IMF magnitude. These results imply that if the algorithm finds a peak in the IMF magnitude then it can be expected that it will find a peak in the speed several hours after. Specifically, the most probable time delay is 11 hours, while the median delay is 20 hours. A similar analysis shows that peaks in density occur before peaks in speed as well.

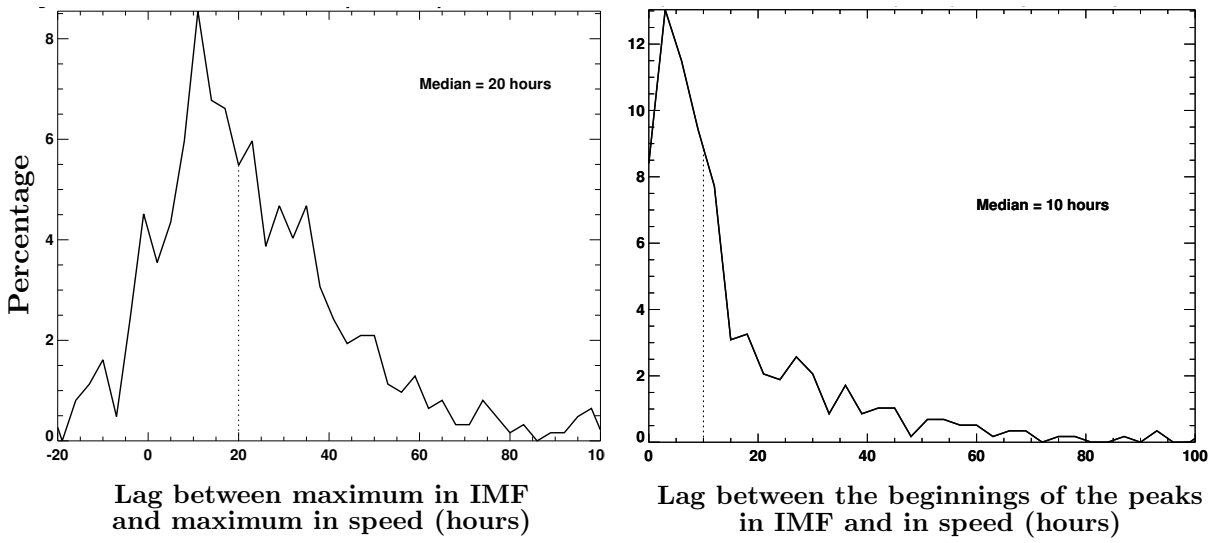

Figure 3.23: Histogram of lag between a peak in IMF and a peak in the speed (left). Histogram of lag between the beginning of a peak in the IMF and the beginning of a peak in the speed (right).

Figure 3.23 (left) also indicates that peaks in the speed can be found more than 40 hours after the peak in the IMF magnitude. However, this might be due to random peaks that are not related to the shock. To reduce the effect of limiting the peak in speed to being above a given threshold, the times when both peaks (IMF and speed) start to increase can be compared by determining the lag $t_{beginning,speed} - t_{beginning,IMF}$, where $t_{beginning}$ is the time in which the gradient in IMF magnitude or speed has crossed a given threshold. The histogram of these time delays is shown in Figure 3.23 (right). The median value of the histogram is now at 10 hours but the curve drops at 15 hours such that 80% of events have delays less than 30 hours. This means that most of the time, a rapid increase in the IMF magnitude will be followed by a rapid increase in the speed in the following day (most likely in the next 10 hours). Any increase in speed beyond one day most likely cannot be attributed to the shock but may correspond to a random increase in speed.

Figure 3.24 (left) shows an example of a high speed event that occurred after a peak in the IMF. At $t_{beginning,speed}$ (marked by a vertical blue line), the IMF magnitude

(solid red line) had already increased by large values. However, Figure 3.24 (right) shows an example where the peak in the IMF occurred after the peak in the speed. As illustrated in Figure 3.23 (left), this situation happens much more rarely.

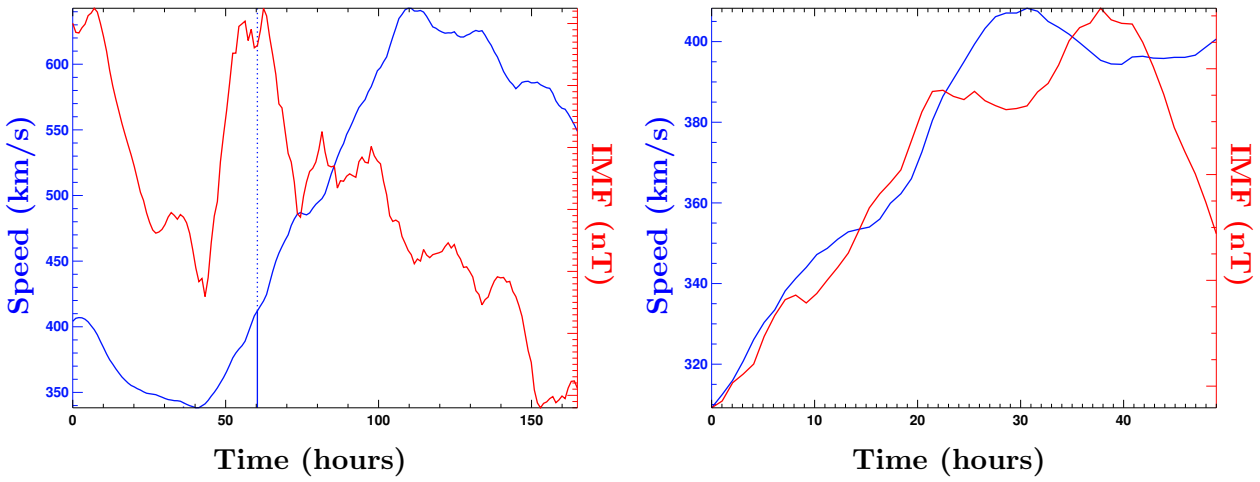

Figure 3.24: Example of a peak in the speed after a peak in the IMF (left); the peak in IMF occurs after the peak in the speed (right).

3.2.2.4 Assessment of the new Probability Distribution Function model

To assess the ability of the new PDF model to predict high speed events, contingency tables were calculated. A typical contingency table presents the number of events that were correctly predicted, or true positives (i.e., the model predicted an event and there actually was an event), the number of false positives (i.e., the model predicted an event that actually did not occur), the number of missed events, or pure misses (i.e., the model did not predict an event that actually occurred), and the number of correct non-events, or true negatives (i.e., the model predicted there would be no event and there actually was no event).

The sensitivity of a model represents the probability that an event that actually occurred was correctly predicted by the model, and is defined as:

$$\text{sensitivity} = \frac{\#\text{ true positives}}{\#\text{ events}} = \frac{\#\text{ true positives}}{\#\text{ true positives } + \#\text{ pure misses}} \tag{3.6}$$

The positive predicted value (PPV) is the probability that a predicted event actually occurred, and is defined as:

$$\text{PPV} = \frac{\text{\# true positives}}{\text{\# predictions of an event}} = \frac{\text{\# true positives}}{\text{\# true positives } + \text{ \# false positives}} \quad (3.7)$$

The negative predicted value (NPV) is the probability that a negative forecast was correct, and is defined as:

$$\text{NPV} = \frac{\text{\# true negatives}}{\text{\# predictions of a non-event}} = \frac{\text{\# true negatives}}{\text{\# true negatives } + \text{ \# pure misses}} \quad (3.8)$$

For the three numbers, the ideal value is 100%.

To get a complete view of the quality of the model, all three of these numbers should be calculated and compared. For example, a model that does not ever predict any peaks never leads to any false positives, but has 0 true positive so the sensitivity is 0%. A model that always predicts a peak in the next 24 hours would have a sensitivity of 100% but its PPV would tend to 0% since it would have a very large number of false positives. This is why forecasters are interested in models that present a good balance between these numbers so that they are able to give correct predictions (positive and negative) as often as possible. Finally, the false alarm rate (= 1 - PPV) is a measure that forecasters are also interested in, as it corresponds to the probability that a predicted event did not actually occur.

3.2.3 Results and discussion

The improved PDF model was run between 1995 and 2012 to make one-day forward predictions, moving by steps of 3 hours between each of the predictions. This enabled the testing of the ability of the model to predict the speed at any phase of the high speed event. For each prediction, a similar question was asked as the one by *MacNeice* (2009a,b): "[I]f the model predicts / does not predict a high speed event in

the next 24 hours, does a high speed event actually occur?" However, answering this question gives only the PPV and the NPV. To determine the sensitivity, a further question must be asked: "[I]f there is actually a high speed event, what is the probability that the model predicted it?" To answer these questions, we used the definition of the high speed event given in the previous section. Specifically, an event, predicted (v_{pred}) or observed (v_{data}), is categorized as a high speed event if the speed now or in the next 24 hours has a gradient greater than 50 km/s/day and goes over 500 km/s. If one of these two conditions is not met then it is not a high speed event.

Figure 3.25 illustrates how this analysis was made. The six graphs represent predictions of the same (observed) high speed event at different times, moving by steps of three hours between each prediction: 13 UT, 16 UT, 19 UT, 22 UT, 01 UT (next day), and 04 UT (next day). At 13 UT and 16 UT, the event had not yet started so the slope was not higher than the threshold at 4.77 km/s/hour. According to the distributions from Figure 3.18, every peak found by the algorithm between 1995 and 2012 had a maximum slope greater than 4.77 km/s/hour. Therefore, no peak was predicted and the two predictions were pure misses (bottom left of the contingency table). At 19 UT, the current speed had increased sufficiently for the PDF model to predict a peak, but the gradient was not high enough for the prediction of the maximum speed to be close to the actual peak. Therefore, the PDF model, although predicting a peak, underestimated the amplitude of the increase, and was a pure miss as well. At 22 UT, 01 UT (next day), and 04 UT (next day), the PDF model correctly predicted the event: both the amplitude and the time of the peak closely match the actual speed, and the descending phase follows the shape of the actual speed. However, recall that the event is categorized as a high speed event if the current speed is below 500 km/s and increase to a value above 500 km/s. This implies that only the predictions at 22 UT and 01 UT (next day) are taken into account in the contingency table and counted as true positives (top left of the contingency table).

Ignoring predictions when the current speed is already higher than this threshold ensures that the assessment of the model reflects a meaningful forecast quality, and not simply observational assist. To conclude, of the six times, three were pure misses, two were true positives, and the last event was ignored because the current speed had already increased at a value higher than 500 km/s so the forecast was not meaningful.

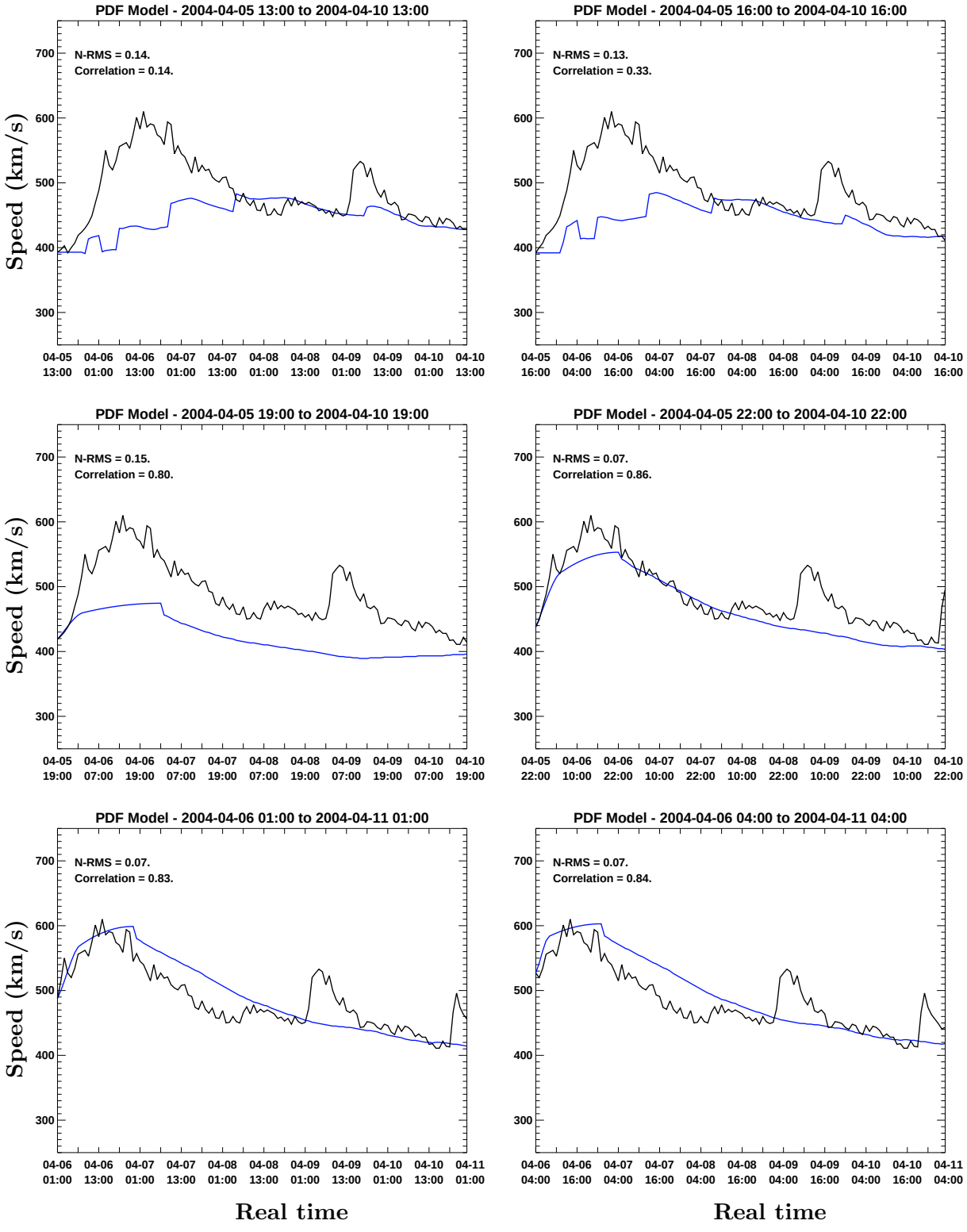

Figure 3.25: Examples of predictions with the PDF model version 2. The black line is the actual speed; the blue line is the median of the PDFs.

Table 3.3: Contingency table for any types of high speed events (HSE: high speed event.)

		Actual solar wind	
		HSE	NO HSE
PDF predictions	HSE	798	523
	NO HSE	3,125	33,401

3.2.3.1 Predictions of high speed events

The contingency table for the results of the predictions from 1995-2012 data is presented in Table 3.3. Out of the 3,923 events that occurred, 798 were predicted by the PDF model (note that several predictions can correspond to the same high speed event but at different times of the ascending phase of the peak, as shown in Figure 3.25). This corresponds to a sensitivity of 20.3%. There were 523 false positives, which implies that the PPV was 60.4%. 3,125 of the 36,526 negative predictions were incorrect, which means that the NPV was 91.4%, so 8.6% were pure misses. If the past performance is an indication of future performance, a positive forecast by the PDF model will be correct 60.4% of the time, while a negative forecast will be correct 91.4% of the time. The PDF model should be able to predict 20.3% of the high speed events.

On the other hand, the performance of the PDF model changes in the presence of IMF and density peaks. While the model itself does not change at all, if an operator of the model sees a peak in the IMF and density, the operator can be more confident in the PDF model. This is shown by considering only time periods that include peaks in both the IMF and the density. The results of these events are shown in the contingency table in Table 3.4.

A positive forecast was correct 69.5% of the time, while a negative forecast was correct 72.0% of the time. The PDF model was able to predict 33.6% of the high speed events. Although the NPV decreased (91.4% to 72.0%), the PDF model missed fewer high speed events (the sensitivity increased from 20.3% to 33.6%) and led to

Table 3.4: Contingency table for high speed events associated with IMF and density peaks (HSE: high speed event).

		Actual solar wind	
		HSE	NO HSE
PDF predictions	HSE	137	60
	NO HSE	271	696

fewer false positives under these conditions (the PPV increased from 60.4% to 69.5%).

3.2.3.2 Direct comparison with the Wang-Sheeley-Arge model

MacNeice (2009b) provided results of high speed event predictions made by the WSA model. He found that a positive forecast was correct 17% (= PPV) of the time and that a negative forecast was correct 94% (= NPV) of the time. However, we chose to make a direct comparison between the PDF and the WSA models using the same definition of the high speed event described earlier in the text for both models, since it was unclear if exactly the same things were being compared in the *MacNeice* (2009b) study as here. In addition to comparing the PPV and the NPV of the two models, their sensitivities were also compared, which is of equal importance for the space weather forecasting community as the PPV and the NPV. The WSA model was run for one year in 2011 to make one day ahead predictions at a temporal resolution of 24 hours. To link the current speed to the speed 24 hours ahead and to provide the same temporal resolution as the PDF model, a cubic spline interpolation was used.

Multiple parameters that characterize a high speed event were used to explore their influence on the model results: the gradient in 24 hours (noted gd, equal to 50 km/s/day so far) and the threshold (noted v_{th}, equal to 500 km/s so far) between slow to fast solar wind. Four sets of values were chosen from different combinations of v_{th} and gd: $v_{th} = 500$ km/s and $v_{th} = 600$ km/s with $gd = 50$ km/s/day and $gd = 70$ km/s/day. As *MacNeice* (2009a) concludes, the algorithm that defines a high speed event can have a significant influence on the results. It is therefore important to apply

the selection criteria and evaluation matrix to both models.

The sensitivity, the PPV, and the NPV for both the WSA and the newly modified PDF model are presented in Table 3.5. *MacNeice* (2009b) used a value for gd of 50 km/s/day and a value for v_{th} of 500 km/s. For the corresponding column of Table 3.5, a sensitivity of 22%, a PPV of 17.9% (corresponding to a false alarm rate of 82.1%), and an NPV of 90.5% for the WSA model were found, very similar to values found in *MacNeice* (2009b). For the PDF model, the sensitivity, the PPV, the false alarm rate, and the NPV were 19.1%, 66.7%, 33.3%, and 91.1%, respectively, close to the values found for the study over 17 years. The sensitivity of the WSA model (22%) is slightly better than the sensitivity of the PDF model (19.1%). However, the WSA model leads to two to three times as many false positives than the PDF model (i.e., the false alarm rate is 2.5 times higher for the WSA model). The PDF model leads to four times more true positives than the WSA model (i.e., the PPV is 4 times higher for the PDF model). Both NPV are comparable, which means that when either model says that there is not an event, it is true about 91% of the time.

Changing gd to 70 km/s/day (leaving v_{th} as 500 km/s) does not change the results by much. It decreases the sensitivity and increases the false alarm rate by about four percentage points for both models, implying that both models have slightly more false positives.

However, when changing v_{th} from 500 km/s to 600 km/s (keeping gd at 50 km/s/day, for example), the sensitivity of the WSA model drops by a factor of two. The false alarm rate of both models is also affected, particularly for the PDF model, in which this increased v_{th} leads to twice as many false positives (i.e., the false alarm rate increases by a factor of two). The NPV is better for both models. This may be due to the fact that since v_{th} has increased, there are fewer events that are counted as high speed events in the observations, so the number of true negatives increases for the same number of pure misses. These results indicate that the PDF model most

Table 3.5: Comparison with the WSA model for different sets of values for gd and v_{th}. The first line corresponds to the value of gd followed by the value of v_{th}: gd / v_{th} (FA: False Alarm Rate).

	50 / 500		70 / 500		50 / 600		70 / 600	
	PDF	WSA	PDF	WSA	PDF	WSA	PDF	WSA
Sensitivity (in %)	19.1	22.0	15.5	17.8	20.9	10.5	18.8	10.0
PPV (FAR) (in %)	66.7 (33.3)	17.9 (82.1)	62.3 (37.7)	13.6 (86.4)	31.0 (69)	12.7 (87.3)	29.4 (70.6)	11.4 (88.6)
NPV (in %)	91.1	90.5	91.9	91.2	97.3	97.0	97.5	97.2

likely overpredicts the solar wind speeds during a peak.

Whatever the value of gd and v_{th} are, the PDF model appears to be more reliable than the WSA model for predictions of high speed events. Because the false alarm rate is smaller for the PDF model, the PDF model leads to fewer false positives than the WSA model. If the threshold for the fast solar wind speed is set at 600 km/s, then the number of misses is also smaller for the PDF model. This can be seen by a higher (by a factor ~ 2) sensitivity. The NPV, which represents the ability of a model to predict correctly the absence of an event, is comparable for each set of gd and v_{th}. Finally, the PPV and NPV found in this study are consistent with those described in *MacNeice* (2009a,b).

3.2.3.3 Ensemble predictions of high speed events

Figure 3.22 shows examples of ensemble predictions of high speed events made by the PDF model. For example, the top left graph shows that the actual speed (black solid line) stays in between the 25% and 75% quartiles for about 108 hours and in between the 10% and 90% quartiles for about 115 hours. This means that most of the time, the PDF model manages to predict the uncertainty interval associated with the main prediction. However, these intervals by themselves are not sufficient to judge if they are relevant. For example, if the widths of the probability distributions the model uses were of the order of 1,000 km/s then the actual speed would always be in between the 25% and 75% quartiles, but the prediction of the uncertainty would not be useful for the forecasters because of the huge uncertainty interval ($\pm$ 500 km/s) associated with the prediction. Therefore, it is important to look at the width of these uncertainty intervals. On the six different predictions shown in Figure 3.22, the half widths are, on average, about 100 km/s for the 25%-75% quartiles and 150 km/s for the 10%-90% quartiles.

The average time the observations last between the 25% and 75% quartiles and

the 10% and 90% quartiles were calculated, as well as the average 25%-75% and 10%-90% widths over all the predictions made between 1995 and 2012. The results were classified according to whether the prediction was a true positive, a false positive, or a pure miss.

Out of 120 hours, the observations were in between the 25%-75% quartiles for 58 hours for true positives, 68 hours for false positives, and 26 hours for pure misses. These numbers increased to, respectively, 88 hours, 86 hours, and 55 hours for 10%-90% quartiles. The smaller time for the pure miss in both cases may be explained by the fact that the model did not predict a high speed event for these times, so the associated speed distributions from the PDF model were narrower. This is confirmed by the measurements of the half-widths: 53 km/s and 95 km/s for the 25%-75% and 10%-90% quartiles, respectively, when the prediction was a miss. When it was a true positive and a false positive, the average widths of the 25%-75% and 10%-90% quartiles were around 145 km/s and 230 km/s, respectively.

3.2.4 Conclusion

An algorithm was created to identify transitions from slow to fast solar wind as any increase in the speed by more than 50 km/s in less than 24 hours. This enabled the construction of new PDFs for the prediction of high speed events. To predict the peak value of the speed, PDFs of the amplitude of the peak were calculated as a function of the maximum slope before the peak. The distribution of the time it takes for the speed to increase from the beginning of the peak to the maximum of the peak was used to predict the time of the peak. From these, predictions of the solar wind speed including peaks could be made.

Contingency tables were calculated to assess the ability of the new PDF model to predict high speed events by answering the following question: "[I]f the model predicts / does not predict a high speed event in the next 24 hours, does a high

speed event actually occur?" It was found that 60.4% of the positive predictions were correct, while 91.4% of the negative predictions were correct. 20.3% of the peaks in the speed are found by the model. The percentage increases to 33.6% when there is an associated peak in both the solar wind density and IMF magnitude before the increase in the solar wind speed.

A direct comparison with the WSA model showed that the number of false positives is more than three times smaller for the PDF model compared with the WSA model. The sensitivity (probability that an actual occurring event is correctly predicted by the model) is similar in both models, as is the NPV (probability that a negative forecast is correct). However, when the peak reaches very high speeds ($>$ 600 km/s), the PDF model misses fewer events than the WSA model, but tends to have more false positives than when the speed is lower.

Finally, ensemble predictions of high speed events by the PDF model provides the forecast community with an interval of uncertainty on the prediction. The study showed that on average, the observations were in between the 25%-75% quartiles $\sim$ 60 hours out of 120 hours and in between the 10%-90% quartiles $\sim$ 85 hours out of 120 hours.

This study was published in *Bussy-Virat and Ridley* (2016).

CHAPTER IV

Effects of density uncertainties on the probability of collision

The rapid increase of the number of objects in orbit around the Earth poses a serious threat to operational spacecraft and astronauts. In order to effectively avoid collisions, mission operators need to assess the risk of collision between the satellite and any other object whose orbit is likely to approach its trajectory. Several algorithms predict the probability of collision but have limitations that impair the accuracy of the prediction. An important limitation is that uncertainties in the atmospheric density are usually not taken into account in the propagation of the covariance matrix from current epoch to closest approach time. The atmosphere between 100 km and 500 km is strongly driven by solar and magnetospheric activity. Therefore, uncertainties in the drivers directly relate to uncertainties in the neutral density, hence in the drag acceleration. This results in important considerations for the prediction of LEO orbits, especially for the determination of the probability of collision. This chapter shows how uncertainties in the atmospheric density can cause significant differences in the probability of collision and presents an algorithm that takes these uncertainties into account to more accurately assess the risk of collision. As an example, the effects of a geomagnetic storm on the probability of collision are illustrated.

4.1 Introduction

The population of objects in orbit around the Earth has dramatically increased in the past decade. In April 2005, NASA performed its first collision avoidance maneuver on a robotic spacecraft (Terra satellite). Two years later, the Chinese satellite Fengyun-1C was destroyed, causing the largest increase in debris in space history (about 3,000 objects larger than 10 cm). On February 10th, 2009, the collision between the operational communication satellite Iridium 33 and the retired Russian communication satellite Cosmos 2251 generated 2,000 debris larger than 10 cm, with many thousands of smaller pieces at an altitude of 800 km. In 2015, four collision avoidance maneuvers and one "shelter-in-Soyuz" procedure were performed by the ISS (*Liou*, 2016; *Newman*, 2016).

Collision avoidance requires the knowledge of the position and velocity of all objects in orbit. Some satellite mission operators can keep track of their satellites quite accurately using GPS data, but the trajectories of all other orbiting objects are harder to follow. Space Situational Awareness (SSA) refers to the effort of gathering and updating the trajectories of natural and man-made orbiting objects (*Stoll et al.*, 2013). Two organizations support the SSA process: JSpOC and the Space Data Association (SDA). SDA was established for satellite operators to share the most up-to-date satellite data (*Stoll et al.*, 2013). JSpOC is currently the single full time global provider of object positions used in collision avoidance due to the accuracy and timeliness of the available information (*Aida et al.*, 2015). It tracks more than 23,000 objects to evaluate the risk of collision with operating satellites by looking for any close approach with cataloged objects. If a potential conjunction is predicted by JSpOC, a Conjunction Data Message (CDM) is generated, which includes information about the close approach to alert the mission operator of a possible encounter. To evaluate the risk of collision, the probability of collision P_c is calculated based on the covariance matrices and the states of both objects at the Time of Close Approach (TCA).

This step is called the Conjunction Assessment Risk Analysis (CARA).

In order to perform CARA, the probability of collision must be accurately determined. This is done by comparing the covariance matrices of the two objects at the time of closest approach. A covariance matrix represents the uncertainties in the state of the object. The i-th diagonal element corresponds to the variance of the i-th parameter (square of the standard deviation). The off-diagonal element i, j is proportional to the correlation coefficient between the i-th and the j-th elements. The covariance matrices of each of the two objects are represented by an ellipsoid, with an orientation given by the principal axes of the covariance matrix. These two ellipsoids are illustrated in Figure 4.1.

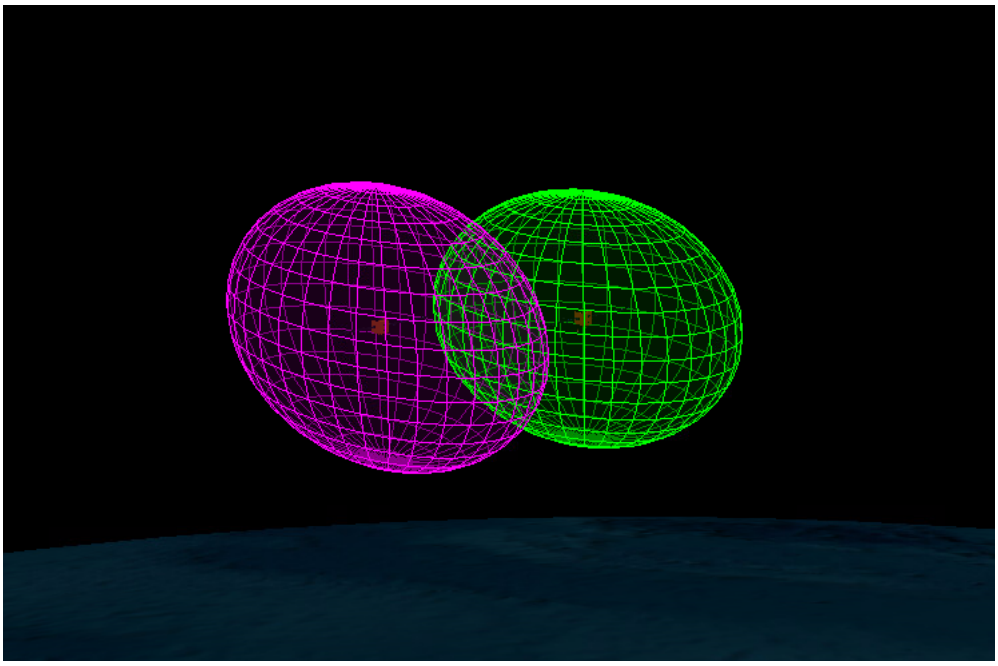

Figure 4.1: Ellipsoids surrounding the primary and secondary spacecraft to represent the uncertainties in the positions of each object (STK - AGI).

The volume of intersection of the two ellipsoids represents the probability of collision of the two objects. Computing this intersection volume implies calculating a three-dimensional integral, which can be computationally intense and complex. Therefore, the relative motion is assumed to be rectilinear so the derivation can be

reduced to a two-dimensional integral in the plane normal to the relative velocity vector, called the conjunction plane (*Hejduk and Frigm*, 2015; *Sanchez-Ortiz et al.*, 2015). Moreover, the orbit uncertainties of the primary object are assumed to be uncorrelated with the orbit uncertainties of the secondary object so the two covariance matrices can be combined in a single covariance matrix (*Chan*, 2008). Finally, the position uncertainties of the two objects are assumed constant during the encounter so the combined covariance matrix is constant during the close approach. These assumptions are valid if the duration of the encounter is short enough (< 500 s) or equivalently if the relative velocity between the primary and secondary objects is high enough (> 10 m/s) (*Hejduk and Frigm*, 2015). The probability of collision, P_c, is then expressed as (*Sanchez-Ortiz et al.*, 2015; *Akella and Alfriend*, 2000; *Alfriend et al.*, 1999; *Foster*, 2001):

$$P_c = \frac{1}{2\pi\sqrt{\det(C)}} \int_{-R}^{R} \int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} e^{-\frac{1}{2}\vec{r}^T C^{-1}\vec{r}} dz dx \tag{4.1}$$

where R is the sum of the two object radii, $\vec{r}$ the vector between the point of interest in the conjunction plane and each point (x, z) is in the circle of radius R (integration area).

Others have modified this technique. *Patera* (2001) used the symmetric form of the probability density to reduce the 2D integral to a 1D contour integral. The contour corresponds to the perimeter of the 2D integration area. This method provides an easier numerical implementation and is computationally faster. *Patera* (2005) used numerical quadrature techniques to transform the contour integral to a 1D angular integral by shifting the origin of the coordinate system. *Alfano* (2005a) reduced the 2D integral into a 1D integral using the error function and analyzed it applying the Simpson's one-third rule.

Although these methods are numerically easy to implement and computationally

efficient, they rely on multiple assumptions that do not necessarily hold in every conjunction case. *Alfano* (2009) assessed Patera's and Alfano's methods for linear relative motion and showed important differences with the baseline Monte Carlo method, particularly when the relative motion assumption does not hold, where errors up to 60% where found. Other methods, that do not assume linear relative motion (Adjoining Cylinders's method (*Alfano*, 2005b), Bundled Parallelepipeds's method (*Alfano*, 2005b), and method of Voxels (*Alfano*, 2006)) also resulted in large errors because of limiting assumptions. Monte Carlo procedures do not require the relative motion to be linear or the covariance matrices to be constant during the encounter. Therefore, the collision risk assessment is more accurate. The main drawback of the Monte Carlo approach is that it is computationally intensive. This method is presented in detail in Section 4.2.

Alfriend et al. (1999) investigated the sensitivity of the probability of collision to errors in the covariance matrix and to the encounter geometry. The study underlined the fact that although the position covariance at epoch is accurate, the velocity covariance is too optimistic because it assumes the dynamic model is perfect. Specifically, the uncertainties in the atmospheric density are usually not taken into account. As a result, the position uncertainties at the time of close approach are too optimistic by about an order of magnitude. *Alfriend et al.* (1999) showed that small errors in the covariance matrix can cause important changes in P_c.

At LEO altitudes, one the main forces acting on spacecraft and debris is atmospheric drag and is a perturbing force particularly hard to model and predict. The drag acceleration $\boldsymbol{a_{drag}}$ of a simple surface is represented by (*Vallado and McClain*, 2007):

$$\boldsymbol{a}_{\text{drag}} = -\frac{1}{2}\frac{C_D A}{m}\rho v_{rel}^2 \frac{\boldsymbol{v_{rel}}}{v_{rel}} \tag{4.2}$$

where C_D, A, and m are the drag coefficient, area projected towards the velocity vector and mass of the surface respectively, and $\boldsymbol{v_{rel}}$ is the satellite velocity with respect to

the moving atmosphere of density ρ. Under 500 km, the thermosphere is too dense to be neglected in accurate orbit calculations. The atmosphere above about 100 km is strongly coupled to the space environment. This system is very complex and the response of the density to driving from the ionosphere, magnetosphere, and Sun is very challenging to estimate. Moreover, the drivers of entire near-Earth space environment, mainly linked to the solar activity, are themselves difficult to predict. Proxies are used to model these drivers, such as F10.7. This is the solar radio flux at a wavelength of 10.7 cm and it is commonly used to model Extreme UltraViolet (EUV) irradiance (*Emmert*, 2015). The perturbations of the Earth's magnetic field resulting from its interaction with the IMF and the solar wind particles are described by magnetic activity indices such as Kp, Ap, Dst, and AE. These indices can be derived from magnetic perturbations directly measured from stations on Earth.

Very few studies have analyzed the impact of uncertainties in the atmospheric drag, despite the fact that it is the largest source of errors for LEO orbits (*Storz et al.*, 2005; *Emmert et al.*, 2016). By modeling the errors in EUV ten-day forecasts with a Brownian motion process, *Emmert et al.* (2014, 2016) derived an analytical equation that relates the uncertainties in F10.7 to uncertainties in the in-track position and found that the in-track position errors grow with time as $\sim t^5$. Although an analytical equation saves a lot of computational time, it does not provide a solution accurate enough for collision risk assessment. In particular, the solution assumes that the only perturbing force is atmospheric drag. However, neglecting the other perturbations can lead to important errors, in particular if neglecting the asphericity of the Earth (Figure 2.1). In addition, the effects of errors in the prediction of the solar proxies increase dramatically in the presence of a solar storm. *D. Pachura* (2016) showed that a miss prediction of an event leads to important changes in the probability of collision, up to a few orders of magnitude.

This chapter first presents an improvement of SpOCK to accurately predict the

probability of collision between two space objects using Monte Carlo procedures. The algorithm models all perturbing forces and takes into account uncertainties in both F10.7 and Ap. The goal is to show how uncertainties in solar driver predictions result in important errors in the probability of collision. Additionally, the effects of a miss prediction of a geomagnetic storm on the probability of collision are demonstrated in an example. The algorithm developed can be used in real-time, and can provide mission operators with a better estimation of the risk of collision.

4.2 Methodology

The central capability of SpOCK is a high accuracy numerical propagator of spacecraft orbits using a comprehensive model of the dynamics of spacecraft in orbit, including the asphericity of the Earth, atmospheric drag acceleration, solar radiation pressure and gravitational perturbations from the Moon and the Sun. Specifically, the non-spherical portion of the mass distribution of the Earth is modeled with a decomposition in spherical harmonics of the gravitational potential. Thermospheric models (NRLMSISE (*Picone et al.*, 2002) and GITM (*Ridley et al.*, 2006)) are implemented in SpOCK to derive the atmospheric density at the position of the spacecraft, allowing for an accurate representation of the atmospheric drag. In addition to modeling the motion of satellites, SpOCK includes several functionalities, such as the coverage of ground stations, the computation of solar power, and the storm intersection forecast tool developed for the CYGNSS mission. Developed in C, SpOCK supports parallelism and is therefore well suited for ensemble, Monte Carlo, or satellite constellation analysis. The algorithm and its different functionalities were presented in Chapter II.

4.2.1 Conjunction Assessment Risk Analysis algorithm

SpOCK can assess the risk of collision with other space objects in orbit (operational satellites or debris). Monte Carlo procedures are used to perturb the initial epoch state (position and velocity) of the primary and secondary spacecraft from the covariance matrices. The algorithm does not have to make any assumptions about the relative motion, the uncertainties in motion, or the covariances during the time span.

SpOCK reads in an input file with a similar format as a CDM generated by JSpOC. More specifically, this input file includes the state (position and velocity) and the covariance matrices of both space objects (noted $\overline{O}_1$ and $\overline{O}_2$) in the ECI coordinate system (6×6 dimension). The sum of the two object radii, $D_{\text{collision}}$,

and the threshold distance under which a close approach is flagged, D_{CA}, are also indicated in the input file, as well as the number of ensembles used in the Monte Carlo procedures, noted N_e.

The flow diagram in Figure 4.2 illustrates the process to evaluate the risk of collision that leads to the determination of the probability of collision between the two space objects.

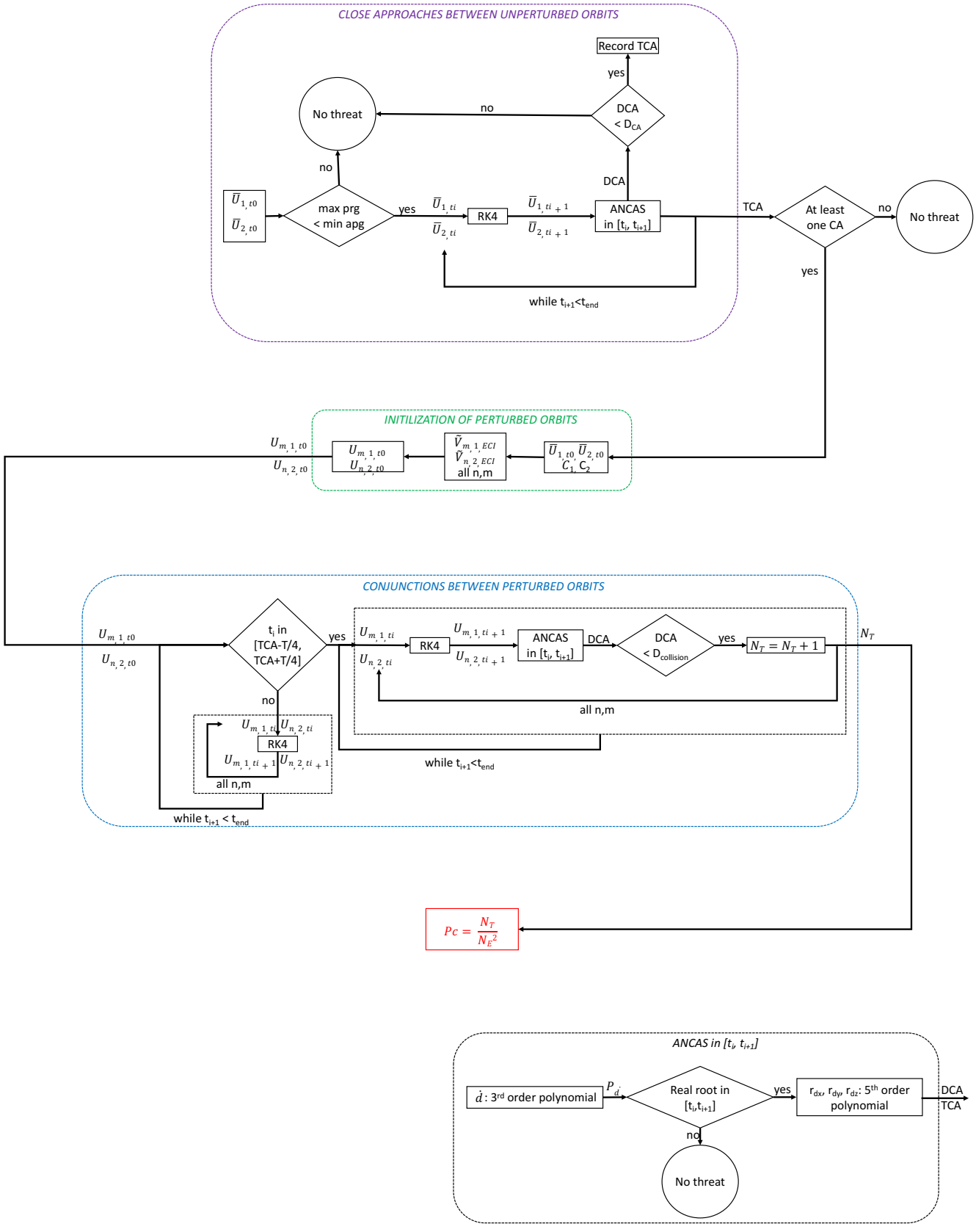

Figure 4.2: Flow diagram of SpOCK's CARA algorithm.

Close approach between the unperturbed orbits

Close approaches are screened between the two unperturbed orbits $\overline{O}_1$ and $\overline{O}_2$ (purple diagram at the top in Figure 4.2). The initial states of $\overline{O}_1$ and $\overline{O}_2$ are noted $\overline{U}_{1,t_0}$ and $\overline{U}_{2,t_0}$. A first filter rules out the possibility of an encounter if the perigee of the higher object's orbit is greater than the apogee of the lower object's orbit. However, if this is not the case, the altitudes cross at some point, and the secondary object represents a potential threat. In that case, the two orbits are propagated for screening for potential close approaches. The propagator uses a Fourth Order Runge-Kutta (RK4) method with a fixed time step dt to integrate the acceleration at each time step t_i of the simulation. A close approach is defined as any minimum in the distance between the two objects smaller than D_{CA}. To determine if a minimum occurs in each interval $[t_i, t_{i+1}]$, SpOCK uses a similar algorithm as Alfano-Negron Close Approach Software (ANCAS) (*Alfano*, 1994). The flow diagram for this algorithm is presented in a separated block from the rest of the flow diagram in Figure 4.2 (black block at bottom of the figure). First, it looks for the existence of a minimum in the distance between the objects by modeling the time derivative of the distance between the two objects by a third-order polynomial, P_d, and assessing if any real root t_{root} exist in the interval $[t_i, t_{i+1}]$. The additional condition $\frac{dP_d(t_{\mathrm{root}})}{dt} > 0$ ensures the root corresponds to a minimum and not a maximum (not represented in Figure 4.2). If a root is found in the interval $[t_i, t_{i+1}]$, the algorithm then determines the Distance of Close Approach (DCA) at t_{root} (now noted as TCA). If the relative vector between the secondary and the primary objects is noted $\boldsymbol{r_d} = \boldsymbol{r_2} - \boldsymbol{r_1}$, SpOCK models the components $r_{d,x}$, $r_{d,y}$, and $r_{d,z}$ by fifth-order polynomials $P_{r_{d,x}}$, $P_{r_{d,y}}$, and $P_{r_{d,z}}$. The distance at close approach is then expressed as:

$$\mathrm{DCA} = \sqrt{P_{r_{d,x}}^2 + P_{r_{d,y}}^2 + P_{r_{d,z}}^2} \tag{4.3}$$

If DCA $< D_{\text{CA}}$ then the situation is recorded as a close approach.

This operation is repeated for each interval $[t_i, t_{i+1}]$ of the propagation.

Initialization of perturbed orbits

If a close approach between the two unperturbed orbits is found, there is a potential risk for a collision. In that case, the Monte Carlo process is initialized (green diagram second to top of Figure 4.2). The initialization consists in perturbing the ECI states of the primary and secondary objects ($\overline{U}_{1,t_0}$ and $\overline{U}_{2,t_0}$). First, the covariance matrices (noted C_1 for the primary object and C_2 for the secondary object) are diagonalized and the two sets of six eigenvalues derived. For each set of eigenvalue, N_e random vectors (6×1 for the position and the velocity) are generated following a Normal distribution centered around 0 with a standard deviation equal to the square root of the eigenvalue of the covariance matrix:

$$\widetilde{V}_{m,j} = \begin{pmatrix} \text{randn}\left(0, \sqrt{\lambda_{1,j}}\right) \\ \text{randn}\left(0, \sqrt{\lambda_{2,j}}\right) \\ \text{randn}\left(0, \sqrt{\lambda_{3,j}}\right) \\ \text{randn}\left(0, \sqrt{\lambda_{4,j}}\right) \\ \text{randn}\left(0, \sqrt{\lambda_{5,j}}\right) \\ \text{randn}\left(0, \sqrt{\lambda_{6,j}}\right) \end{pmatrix}$$

where $\widetilde{V}_{m,j}$ represents the m^{th} random vector ($m = 1, ..., N_e$) associated with the primary ($j = 1$) or secondary ($j = 2$) object, and $\lambda_{i,j}$ the i^{th} eigenvalue ($i = 1, ..., 6$) of the covariance matrix C_j.

Each vector $\widetilde{V}_{m,j}$ is then converted back to the ECI coordinate system ($\widetilde{V}_{m,j,ECI}$) with the rotation matrix used for the diagonalization of the covariance matrix.

Finally, each perturbation $\widetilde{V}_{m,j,ECI}$ is added to the unperturbed ECI state to

generate N_e perturbed states $U_{m,1}$ and $U_{n,2}$:

$$U_{m,1,t_0} = \overline{U}_{1,t_0} + \widetilde{V}_{m,1,ECI} \tag{4.4}$$

$$U_{n,2,t_0} = \overline{U}_{2,t_0} + \widetilde{V}_{n,2,ECI} \tag{4.5}$$

Each ensemble member initialized as U_{m,j,t_0} is now noted $O_{m,j}$.

4.2.1.1 Conjunctions between the perturbed orbits

Once the $2N_e$ perturbed orbits are initialized, SpOCK propagates them (blue block third to top in Figure 4.2). During the propagation, it screens for any conjunction between a perturbed ensemble $O_{m,1}$ $(m = 1, ..., N_e)$ of the primary object and a perturbed ensemble $O_{n,2}$ $(n = 1, ..., N_e)$ of the secondary object. The algorithm is the same as the one to find a close approach between the unperturbed orbits $\overline{O}_1$ and $\overline{O}_2$. However, it now uses the sum of the two object radii $D_{\text{collision}}$ as the minimum distance under which a conjunction is recorded. Therefore, SpOCK first looks for the existence of a minimum distance in each interval $[t_i, t_{i+1}]$, in which case it then calculates the minimum distance as in Equation 4.3. If this distance is smaller than $D_{\text{collision}}$, the situation is recorded as a collision. SpOCK repeats this operation for each combination in the set (m, n) resulting in N_e^2 comparisons. The total number of collisions found is noted N_T.

To be computationally efficient, conjunctions between each perturbed orbit $O_{m,1}$ and $O_{n,2}$ are not screened in every interval $[t_i, t_{i+1}]$ of the propagation, but only in a time spanning the unperturbed close approach determined in the first step of CARA. This time span is equal to half an orbital period T to ensure that no collision is missed between the perturbed orbits. If several close approaches have been found, then SpOCK applies this algorithm for each interval $[\text{TCA} - T/4, \text{TCA} + T/4]$.

Derivation of the probability of collision

Once the $2N_e$ orbits are propagated and screened for collisions, the probability of collision is calculated as the ratio between the total number of collisions N_T divided by the total number of possible scenarios N_e^2 (red block fourth to top in Figure 4.2):

$$P_c = \frac{N_T}{N_e^2} \tag{4.6}$$

Dagum et al. (2000) recommends that the total number of possible scenarios, $N(= N_e^2)$, must be greater than a lower bound that is a function of the absolute error ϵ, the level of confidence $1 - \delta$, and the expected value of the probability of collision P_T:

$$N > \frac{4\,(e-2)\,[(1-P_T)\,P_T]}{\epsilon^2} \ln \frac{2}{\delta} \tag{4.7}$$

providing the condition:

$$\epsilon < (1 - P_T)\,P_T \tag{4.8}$$

For example, if the expected probability of collision is 10^{-4}, then the total number of possible scenarios must be greater than 1.06 billion to meet a 1% relative accuracy ($\epsilon = 10^{-6}$) with a 95% confidence level ($\delta = 0.05$). The condition in Equation in 4.8 is easily met because P_T is not near 1 (*Alfano*, 2009). This implies that the total number of ensemble members must be greater than 32,600 ($N_e > \sqrt{1.06 \times 10^9}$). To be more computationally efficient, SpOCK runs the ensemble members in parallel, allowing the risk assessment of a two day scenario to be performed in an hour using 200 cores.

4.2.2 Modeling uncertainties in F10.7 and Ap

The uncertainty in the thermospheric density modeling is generally of the order of 20% (*Alfriend et al.*, 1999). This is when considering a perfect knowledge of the

inputs used in the thermosphere models, such as the solar flux F10.7, Ap, or Dst. However, predicting the solar drivers a few days ahead adds more uncertainties in the prediction of the thermospheric density. The large variability of the solar activity makes the task even harder. Geomagnetic activity driven by solar flares, CMEs, and Corotating Interaction Regions (CIRs) cause the most important disturbances of the neutral density. For example, it was found that CIRs can cause density increases by 75% on average (*Lei et al.*, 2011). A 30% to 60% increase at low to mid-latitudes in the density response to EUV solar flux enhancement resulting from a solar flare was reported in *Sutton et al.* (2006) and *Pawlowski and Ridley* (2011). *Bruinsma et al.* (2006) and *Sutton et al.* (2005) studied the response to severe geomagnetic storms. Both noted a rapid response of the thermospheric density ($\sim$ a few hours) at all latitudes with enhancements by more than three times the density. *Bruinsma and Forbes* (2007) reported up to 800% density enhancements at the equator during sudden increases in geomagnetic activity.

Despite these considerations, collision risk assessment is usually performed by neglecting the uncertainties in the atmospheric density. At LEO altitudes, where atmospheric drag is the dominant perturbation, this can result in important errors in the probability of collision.

SpOCK primarily uses the NRLMSISE density model to derive the density at the position of the spacecraft at each time step of the propagation. The solar inputs of NRLMSISE are the solar flux F10.7 and the geomagnetic index Ap. These are predicted in real time for the following 45 days at SWPC from NOAA. To model uncertainties in the predictions of F10.7 and Ap, SpOCK compares historical 3 day predictions by SWPC (`http://www.swpc.noaa.gov/`) with actual measured values. For each forecast time (1 day, 2 days, 3 days ahead), the difference between the predicted index and the measured value is computed ($P_{\text{predicted}} - P_{\text{measured}}$ where P is F10.7 or Ap), from which distributions are derived. The 10%, 20%, 30%, 40%, 50%

Table 4.1: Quartiles of the distributions of the difference between the predicted and actual F10.7.

	10%	20%	30%	40%	50%	60%	70%	80%	90%
1 day	-4.	-3.	-2.	-0.2	0.	1.	2.	3.	4.8
2 days	-6.	-4.	-2.7	-1.	0.	1.	2.7	4.8	7.
3 days	-8.	-5.	-3.	-2.	0.	1.	3.	5.	8.

Table 4.2: Quartiles of the distributions of the difference between the predicted and actual Ap.

	10%	20%	30%	40%	50%	60%	70%	80%	90%
1 day	-7.	-3.	-2.	0.	1.	2.	3.	6.	9.8
2 days	-9.	-4.	-2.	0.	1.	2.	3.	5.	8.9
3 days	-9.	-5.	-2.	0.	1.	1.	3.	5.	8.

(median), 60%, 70%, 80%, and 90% quartiles are reported in Tables 4.1 (F10.7) and 4.2 (Ap).

The quartile values are used to model the uncertainties in the predictions of F10.7 and Ap. For example, the predictions of F10.7 and Ap on November 26th, 2016 by SWPC are reported in Table 4.3. The uncertainties on these predictions are taken from the quartile values of Tables 4.1 and 4.2. While SWPC predicted F10.7 to be 83 on November 27, the 20% quartile of the distribution of difference between the prediction and the observations being -3 (second column of Table 4.1), so the 20% quartile of the predicted distribution of F10.7 is 80. The same reasoning is valid for each quartile, resulting in a distribution of predicted F10.7 on November 27 with quartiles equal to the values in Tables 4.1 and 4.2, instead of the single nominal value predicted by SWPC. Applying this approach for each prediction horizon (1 day, 2

Table 4.3: Predictions of F10.7 and Ap by SWPC on November 26th, 2016.

	F10.7	Ap
Nov. 26	82	15
Nov. 27	83	10
Nov. 28	83	8
Nov. 29	81	8

days, and 3 days ahead), SpOCK generates a series of predictions for F10.7 and Ap 3 days ahead. These are shown in Figure 4.3.

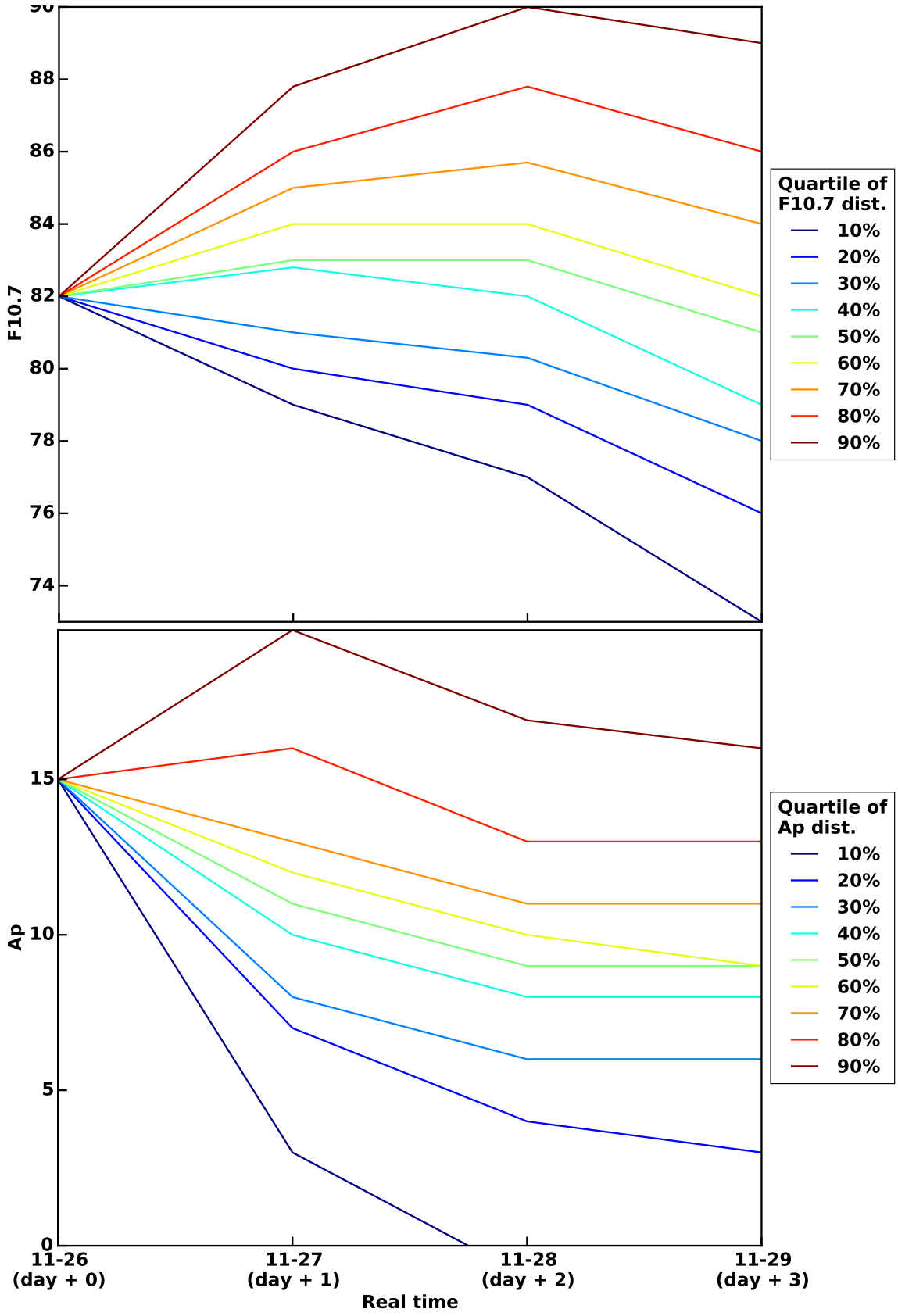

Figure 4.3: Prediction of the solar flux F10.7 (top) and geomagnetic index Ap (bottom) as a function of forecast time from the 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% quartiles of the distributions of the error in historical forecast.

The density modeled with NRLMSISE from these different solar activity conditions is shown in Figure 4.4. SpOCK assumes that the different quartiles of F10.7 and Ap are related, such that when, for example, the 20% quartile for F10.7 is used, the 20% quartile for Ap is also used. While this may not be true, with more research being needed on this, it bounds the problem. The 90% quartile of the distribution 3 days ahead is $\sim$ 50% higher than the 10% quartile. This means that if SWPC repeated the same predictions a large number of times and that these predictions were compared to the actual values, 80% of the predictions would result in differences in density that are less than $\sim$50%, which is large compared to the $\sim$20% uncertainty in the initial density, described by *Alfriend et al.* (1999).

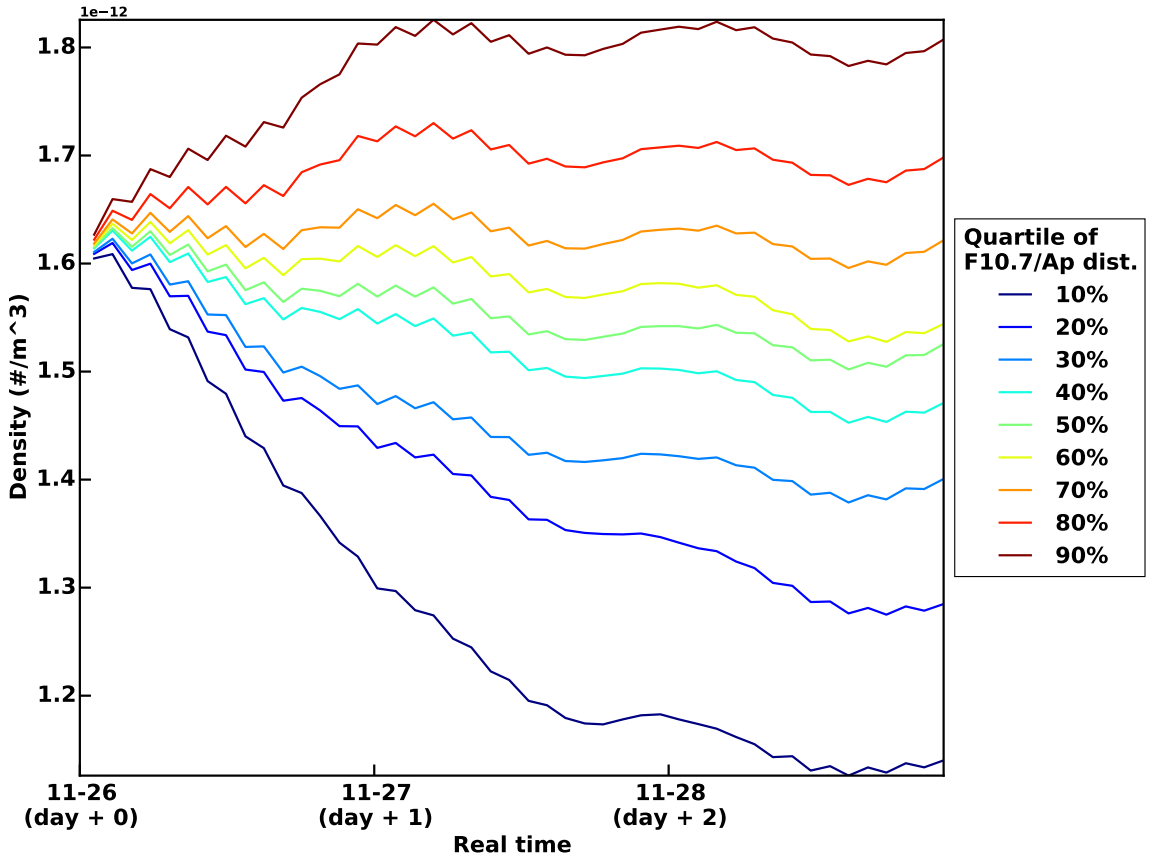

Figure 4.4: Atmospheric density at the position of the spacecraft as a function of forecast time modeled by NRLMSISE from the quartiles of the distributions of the error in historical F10.7 and Ap forecast (shown in Figure 4.3).

SpOCK uses real time predictions from SWPC and adds uncertainties to assess the risk of collision under different solar conditions. SpOCK propagates different ensembles, consisting of thousands of members, through different thermospheres. Each individual ensemble is propagated through NRLMSISE driven by the same drivers, so the members can be realistically compared to each other. Different ensembles represent different density drivers, such that the dependence of the probability of collision on the density drivers can be explored. These simulations are run in parallel, which enables a quick evaluation of the range of possible values of the predicted probability of collisions, instead of one unique value as it is currently done in a typical conjunction risk assessment analysis. Examples are shown in Section 4.4.

4.2.3 Modeling other uncertainties

When assessing the true risk of collision, all uncertainties must be taken into account. Uncertainties in the initial position and velocity are important because of the limited availability and accuracy of tracking the object. The thermospheric density is the main source of uncertainties in the propagation of the spacecraft but, as shown in Equation 4.2, it is not the only parameter of the drag acceleration. Uncertainties in the drag coefficient and the cross section area are not negligible either.

Drag coefficient

The coefficient of drag C_D represents the transfer in momentum between the molecules of the atmosphere and the surfaces of the satellite. It is a function of the satellite shape and attitude, of the atmospheric conditions (temperature and composition), and of the properties of the satellites surfaces. It usually decreases with the altitude (*Moe and Moe*, 2005; *Horsley*, 2012).

In order to model uncertainties in the drag coefficient, SpOCK initializes the properties of each surface of the ensemble members with drag coefficients following

Gaussian distributions centered around the nominal values with standard deviations specified as an extra dimension of the covariance matrices. Figure 4.5 shows an example of the distribution of 5,000 ensemble members corresponding to a nominal value of Cd of 2.2 with a standard deviation of 0.2.

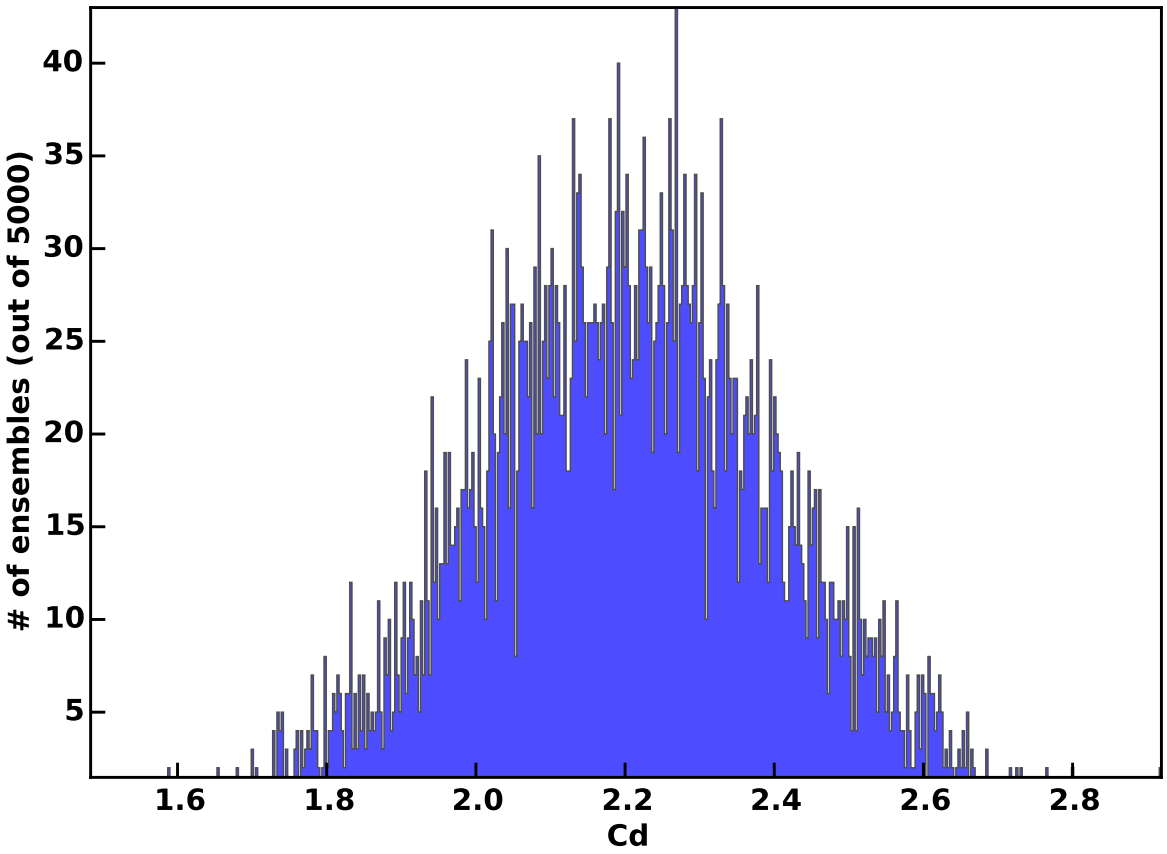

Figure 4.5: Drag coefficient distribution for 5,000 samples representing an uncertainty of 0.4 around a nominal value of 2.2.

Attitude

Modeling the uncertainty in the attitude for a controlled satellite that has a known attitude control uncertainty consists in having ensemble members that drift with a random angular velocity from a nominal attitude for a given time before going back to the nominal attitude. This enables the simulation of the attitude determination and control system of satellites that randomly drift from a nominal controlled atti-

tude. Additionally, objects that are uncontrolled can be simulated by setting nominal tumbling rates around the three axes, with uncertainties around those nominal rates also specified.

Orbital elements

The uncertainties in the initial positions and velocities given in the covariance matrices are usually formulated in ECI coordinates. However, they can also be expressed as classical orbital elements (altitude of the apogee, inclination, eccentricity, true anomaly, RAAN, and argument of perigee) to allow the user to initialize the orbits with orbital elements rather than with an ECI state.

Although these different uncertainties need to be taken into account in a realistic collision risk assessment, this study focuses on exclusively modeling uncertainties in the initial positions and velocities, as well as in the thermospheric density (described in Section 4.2.2).

Table 4.4: Comparison of the cumulative probability of collision between SpOCK and *Alfano* (2009).

	N	SpOCK P_c	Alfano P_c	ΔP_c / P_c * 100
Case 1	10,176,100	0.221528	0.217467140	1.87
Case 6	23,040,000	0.004199	0.004300500	2.36
Case 7	659,462,400	0.000163	0.000161462	0.95

4.3 Validation

To validate the conjunction assessment risk analysis algorithm in SpOCK, the cumulative probabilities of collision computed for three different cases were compared with results from *Alfano* (2009). More specifically, two LEO configurations (cases 6 and 7 in *Alfano* (2009)) and one GEO configuration (case 1 in *Alfano* (2009)) were analyzed. The initial states and covariance matrices, included in *Alfano* (2009), were used to initialize CARA in SpOCK. *Alfano* (2009) uses a similar CARA algorithm as SpOCK's. In all three cases, the time of close approach is 2 days after the epoch. The dynamic model uses a spherical Earth with no perturbations. The goal here is to validate the algorithm for computation of the probability of collision. SpOCK's propagator and its dynamic model were validated in Chapter II.

Table 4.4 summarizes the results for the cumulative probability of collision. The relative difference is smaller than 3% for the three cases, proving that both algorithms are in a good agreement. The small differences might be due to the fact that *Alfano* (2009) uses an analytical equation to propagate the covariance matrices from current epoch to time of closest approach.

Figure 4.6 shows the cumulative distribution functions computed by SpOCK for cases 1 (top), 6 (center), and 7 (bottom). Similar figures are available in *Alfano* (2009) but are not shown here because they correspond to probabilities determined with a different method than the Monte Carlo approach. Therefore, a direct comparison of the cumulative probability as a function of time is not possible. However, the trends

of the figures in *Alfano* (2009) are very similar to the ones presented in Figure 4.6. Case 1 represent a nonlinear relative motion that causes the probability to increase again a few hours after the first close approach. Cases 6 and 7 correspond to a linear relative motion where the probability of collision starts growing a few minutes before the unperturbed close approach, with a faster increase for case 6.

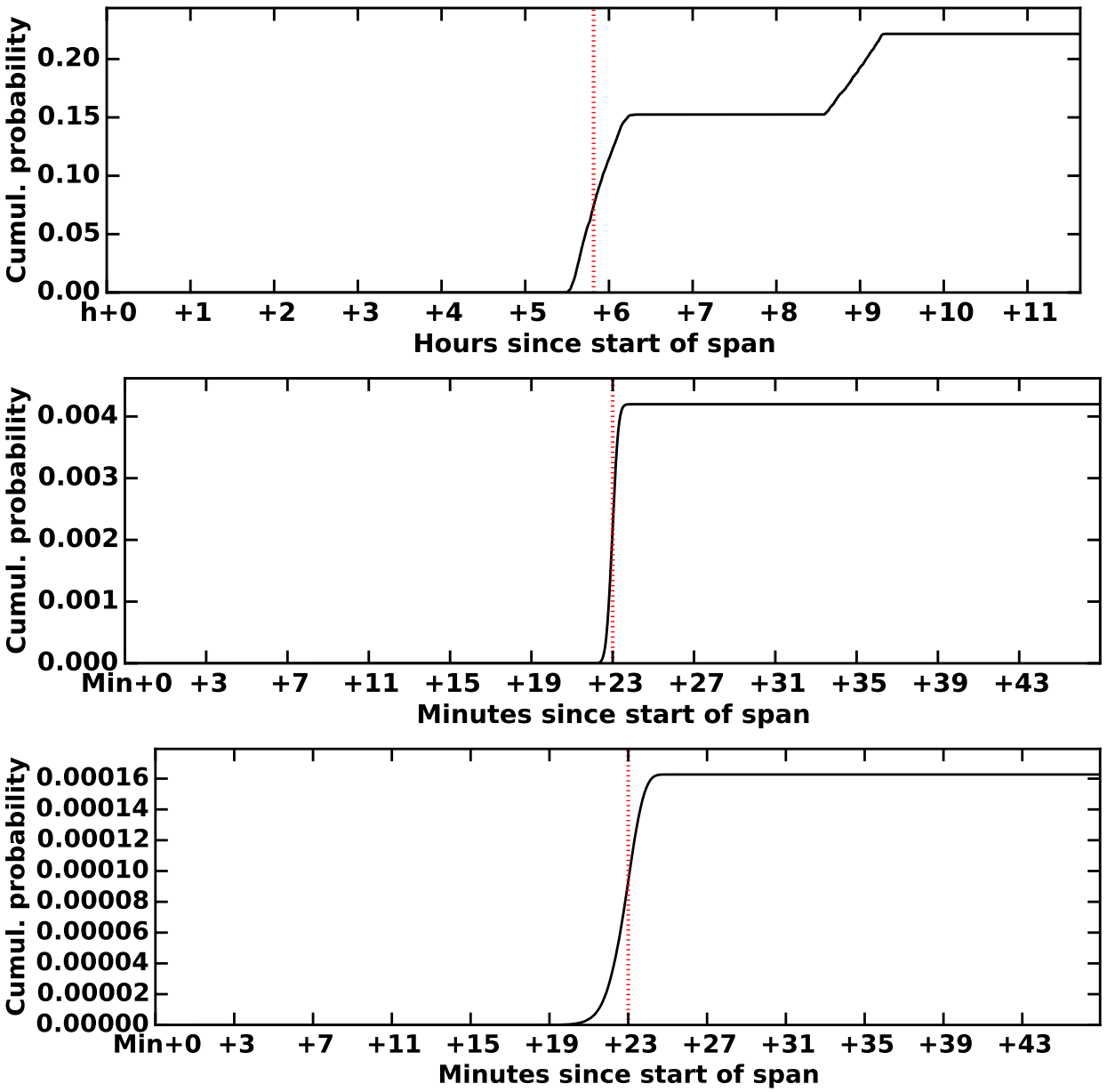

Figure 4.6: Cumulative probability of collision during the half orbit spanning the unperturbed close approach. The vertical red dashed line represents the time of close approach between the two unperturbed orbits. The initial states and covariance matrices are taken from *Alfano* (2009): top is case 1, center is case 6, and bottom is case 7.

4.4 Results and discussion

4.4.1 Effects of atmospheric drag on the probability of collision

To understand the effects of atmospheric drag on the probability of collision, a collision between two satellites with slightly different velocities is considered. Both spacecraft start at their respective initial position and Satellite 2 is assumed to have the higher speed of the two satellites. The distance between the initial positions and the conjunction point are respectively noted D1 and D2.

If the density of the atmosphere is actually higher than predicted, the increase in atmospheric drag will have more effect on Satellite 2 than on Satellite 1 because Satellite 2 moves faster. In other words, Satellite 2's altitude will decrease faster than Satellite 1's, so Satellite 2's speed will increase more than Satellite 1's. Therefore, Satellite 2 will reach the distance D2 before Satellite 1 reaches the distance D1. Because of this differential change in speed, Satellite 2 will get to the conjunction point before Satellite 1 so the two satellites will potentially not collide anymore. Similarly, two spacecraft that are not predicted to collide could actually collide if the density is different from the prediction.

In the Monte Carlo method, a large number of such situations are considered. The total number of collisions in the Monte Carlo procedure can therefore directly depend on the density of the atmosphere in which the spacecraft orbit. In other words, uncertainties in the predictions of the thermospheric density have an effect on the probability of collision.

4.4.2 Effects of uncertainties in the thermospheric density predictions on the probability of collision

In Section 4.2.2, the methodology to model uncertainties in the atmospheric density resulting from uncertainties in the predictions of the indices F10.7 and Ap was

explained. In this section, an example of a possible collision is described to illustrate how the density can affect the probability of collision.

Encounter geometry

SpOCK was run to assess the risk of collision between two hypothetical objects using the algorithm presented in Section 4.2. Both spacecraft orbit at 400 km with a 30° and 60° inclination respectively. The RAAN is 0° for Satellite 1 and 30° for Satellite 2. The ECI states and covariance matrices at the initial epoch (12 am Universal Time Coordinated (UTC) on November 26th, 2016) are included in Appendix A. A first collision risk assessment was performed using the median values of the F10.7 and Ap predictions of Figure 4.3 (50% quartile in green). 50,000 ensemble members were propagated for each of the two satellites so the total number of possible encounters is 2.5 billion. According to Dagum's Equation 4.7, this number of samples allows the assessment of a probability of collision as low as 4×10^{-5} with a 1% relative accuracy and a 95% confidence level. The orbits are propagated with a 10 second time step and the gravitational perturbation due to the equatorial bulge ($J2$) is taken into account. The density is modeled using the predictions of F10.7 and Ap by SWPC shown in Table 4.3. This scenario represents a situation where the mission operator wants to assess the risk of collision on November 26th for the following three days. The two orbits are illustrated in Figure 4.7. The green line represents the orbit of Satellite 1 (30° inclination) and the magenta line represents the orbit of Satellite 2 (60° inclination). The closest approach is predicted to occur over the Pacific Ocean.

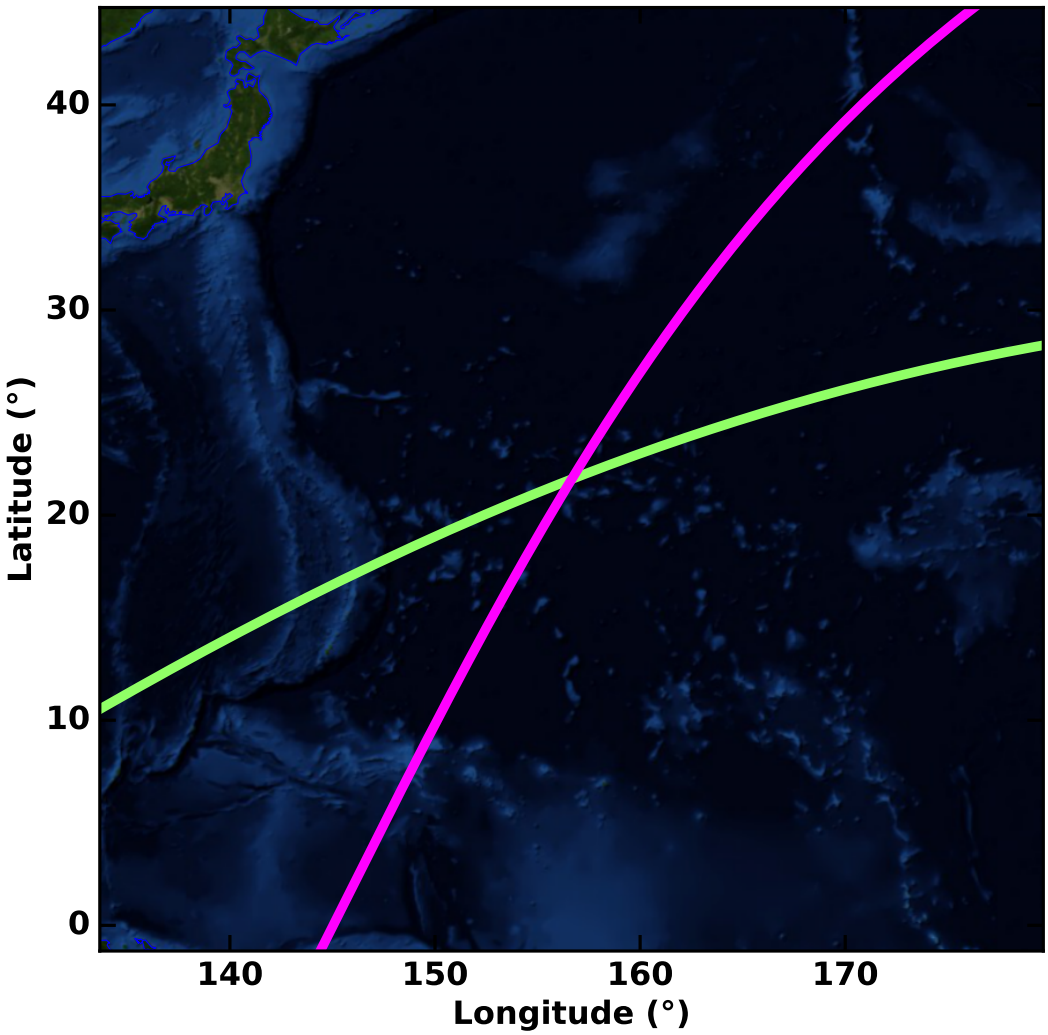

Figure 4.7: 2D visualization of the two orbits. Satellite 1 (green line) orbits at 400 km with a 30° inclination and a 0° RAAN. Satellite 2 (magenta line) orbits at 400 km with a 60° inclination and a 30° RAAN.

The relative distance over the first two days between the two unperturbed orbits is presented in Figure 4.8. The time of closest approach is predicted ∼36 hours after epoch, at noon on November 27th.

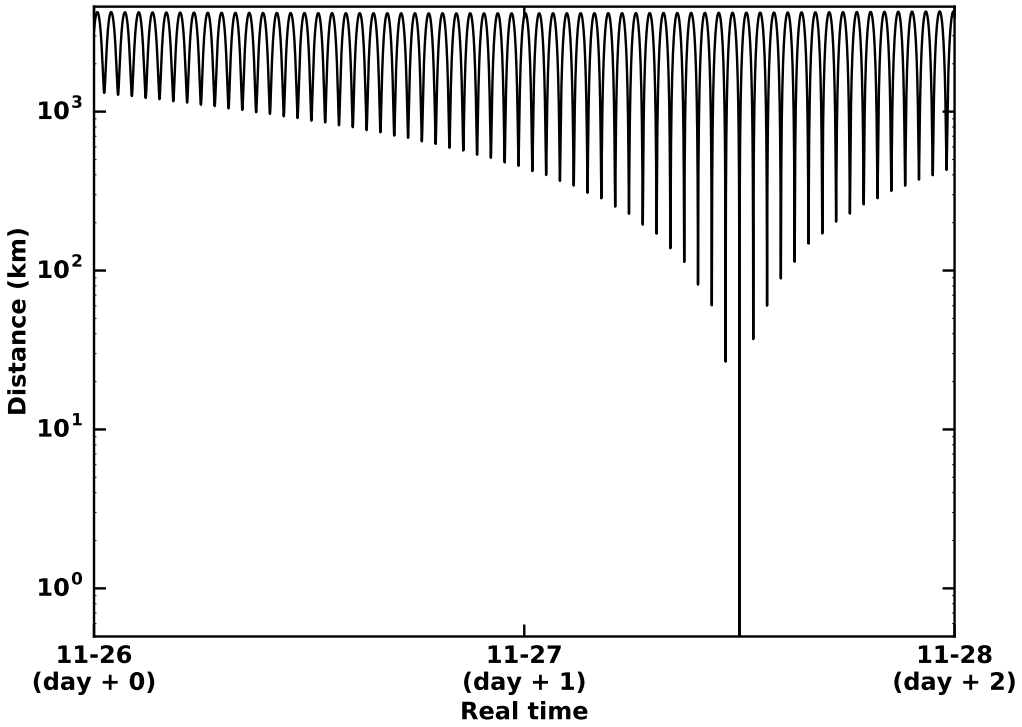

Figure 4.8: Distance between the two unperturbed orbits over the first two days of the propagation. The close approach occurs on November 27th at 12 pm.

The geometry of encounter is illustrated in Figure 4.9. Three 10 s time steps are represented around the time of closest approach. Object 1 is represented by a square (smaller inclination) and Object 2 by a circle (higher inclination). The time of closest approach is predicted to be at 12:00:00.104, at $\sim 21.75°$N latitude.

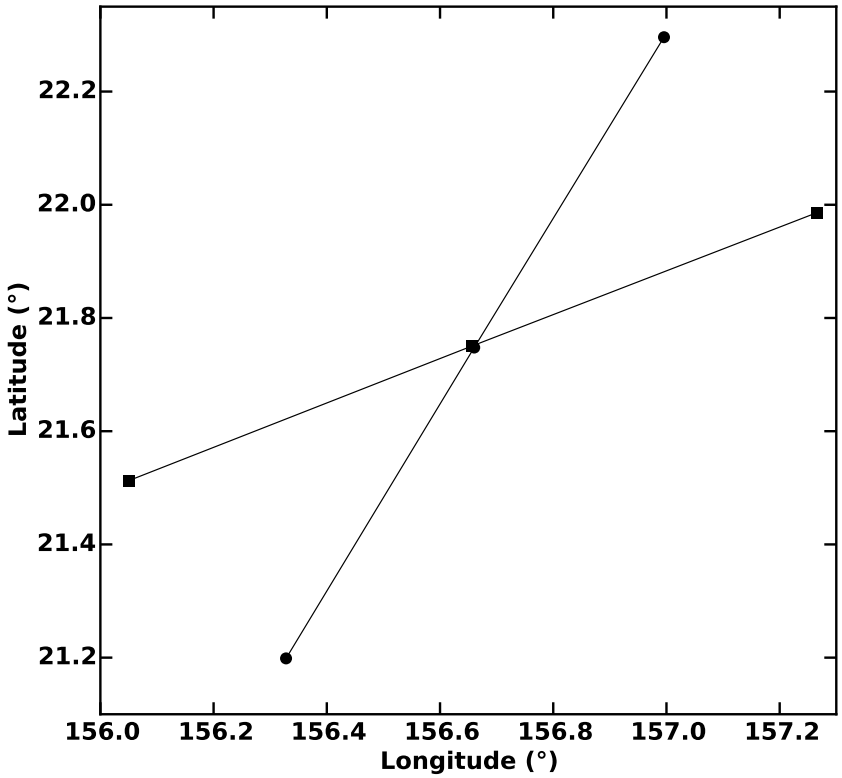

Figure 4.9: Geometry of encounter. Object 1 (30° inclination) is represented as square and Object 2 (60° inclination) as a circle.

Collision risk assessment

The minimum distance under which a close approach between the two unperturbed orbits was flagged was 10 km. Therefore, if the two spacecraft $\overline{O}_1$ and $\overline{O}_2$ got close to each other by less than 10 km then SpOCK's Monte Carlo algorithm presented in Section 4.2.1 assessed the risk of collision for this close approach. The sum of the two spacecraft radii was 1.3 m, implying that any situation with a distance between an object $O_{m,1}$ and an object $O_{n,2}$ smaller than 1.3 m was recorded as a collision.

The probability of collision was 1.051×10^{-4}, right above the threshold for a collision avoidance maneuver used by NASA (10^{-4}). The distribution of the time of closest approach for each collision between a perturbed object $O_{m,1}$ and a perturbed object $O_{n,2}$ is shown in Figure 4.10. The 50% width of the distribution (defined the

difference between the 75% and the 25% quartiles) is 164 ms, which represents the uncertainty in the time of closest approach. This uncertainty is particularly small because of the encounter geometry. For example, if the inclinations of the two orbits were similar, the velocities at conjunction would be almost parallel to each other. Since uncertainties are usually greater in the along-track direction, the range of values for the time of close approaches increases for such a parallel conjunction. This will be shown in Section 4.4.3.

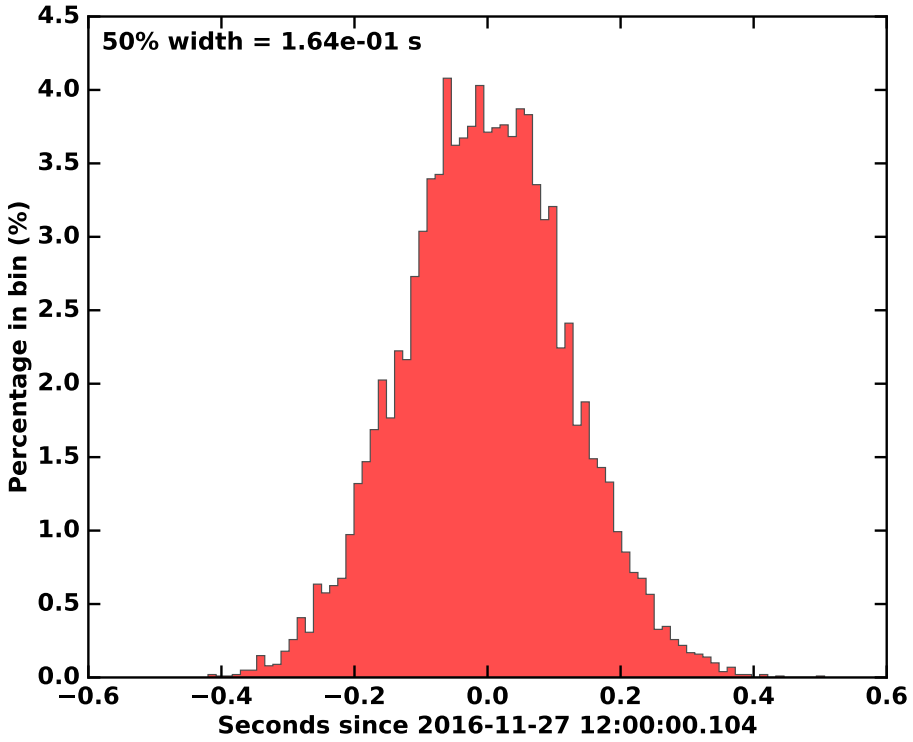

Figure 4.10: Distribution of the time of close approach for all conjunctions.

To demonstrate the effects of uncertainties in the thermospheric density on the probability of collision, the risk of collision was assessed with the different densities shown in Figure 4.4. More specifically, nine cases were evaluated for collision risk assessment. Each case corresponds to the probability of collision between the two objects orbiting in an atmosphere whose density is represented by one of the nine quartiles of Figure 4.4.

Table 4.5: Time and distance of close approach for different atmospheric density scenarios.

Quartile of the predicted density distribution	TCA	DCA (m)
10%	2016-11-27 12:00:00.311	125
20%	2016-11-27 12:00:00.213	106
30%	2016-11-27 12:00:00.182	91.2
40%	2016-11-27 12:00:00.124	64.3
50%	2016-11-27 12:00:00.104	58.0
60%	2016-11-27 12:00:00.075	42.6
70%	2016-11-27 12:00:00.047	26.7
80%	2016-11-27 11:59:59.989	1.61
90%	2016-11-27 11:59:59.914	38.2

The time and distance of close approach between the two unperturbed orbits, i.e., using the nominal orbital parameters, for each scenario are represented in Table 4.5. The unperturbed orbits are not predicted to collide because the distances at closest approach are greater than 1.3 m. But they are flagged to be investigated due to the distances being less than 10 km.

The higher the density, the more drag there is on the spacecraft, which leads to a lower orbital altitude and thus a higher orbital speed, resulting in earlier closest approaches. This can be noticed in Table 4.5 where the time of close approach gradually decreases from the 10% quartile to the 90% quartile by $\sim$ 400 ms.

Figure 4.11 shows the positions of the two spacecraft on November 27th at 12:00:00 for the nominal orbit condition for the nine different thermospheric conditions. This corresponds to the time step preceding the close approach for all quartile scenarios, except for the 80% and 90% quartiles cases for which it corresponds to the time step following the encounter. The increasing delay between this snapshot and the encounter from the 70% quartile scenario to the 10% quartile appears clearly in Figure 4.11. Because the closest approach for the 80% and 90% quartile cases occurred before the time of this snapshot, the red and purple markers are located beyond the conjunction point (not represented in the Figure for clarity).

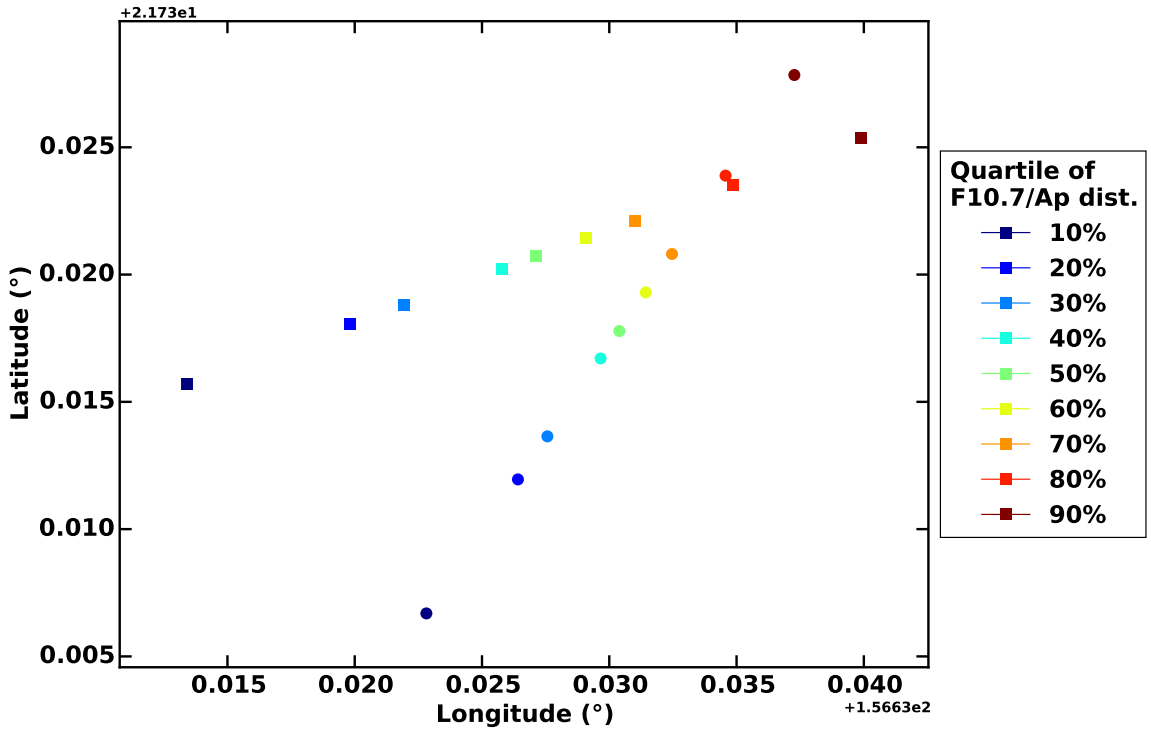

Figure 4.11: Snapshot of the unperturbed orbits $\overline{O}_1$ (square) and $\overline{O}_2$ (circle) at 12:00:00 pm on November 27th, 2016. The two objects are about to collide for the 10%, 20%, 30%, 40%, 50%, 60%, and 70% quartile scenarios and have just collided for the 80% and 90% quartile scenarios.

Added to the uncertainty of the time of closest approach between ensembles of a same scenario (160 ms, Figure 4.10), the total uncertainty on the time of close approach is $\sim$ 560 ms. This is a small uncertainty and it does not cause any concern for operations because maneuvers are planned hours in advance. However, this short uncertainty is due to the encounter geometry, where the 2 objects approach with non parallel velocities. For parallel conjunctions, this uncertainty grows dramatically, as it will be shown in Section 4.4.3.

Although the effects of density uncertainties on the time of closest approach are small for this geometrical configuration, the effects on the probability of collision are much more important. This is shown in Figure 4.12. Risks of collisions are usually assessed using a single prediction of the thermospheric density, ignoring any

uncertainties in the forecast. Figure 4.12 demonstrates the errors resulting from such an approach. The x axis corresponds to the time of closest approach between the two unperturbed orbits from Table 4.5, and the y axis to the probability of collision associated with the close approach. The horizontal red dashed line corresponds to the probability of collision threshold (10^{-4}) above which a maneuver is typically recommended.

The scenario corresponding to a thermospheric density predicted from the median distribution is represented in green and results in a probability of collision of 1.051×10^{-4}, as shown previously. The maximum probability of collision is obtained with the 70% quartile scenario (1.129×10^{-4}). Although the 40%, 50%, 60%, 70%, and 80% quartile cases lead to probabilities of collision above the maneuver threshold, the 10%, 20%, 30%, and 90% quartile scenarios result in probabilities of collision lower than this threshold, reaching values as low as 6.575×10^{-5} and $\sim 9 \times 10^{-5}$. This corresponds to a relative difference of 25-30% with the probability of collision from the 60% and 70% quartile scenarios.

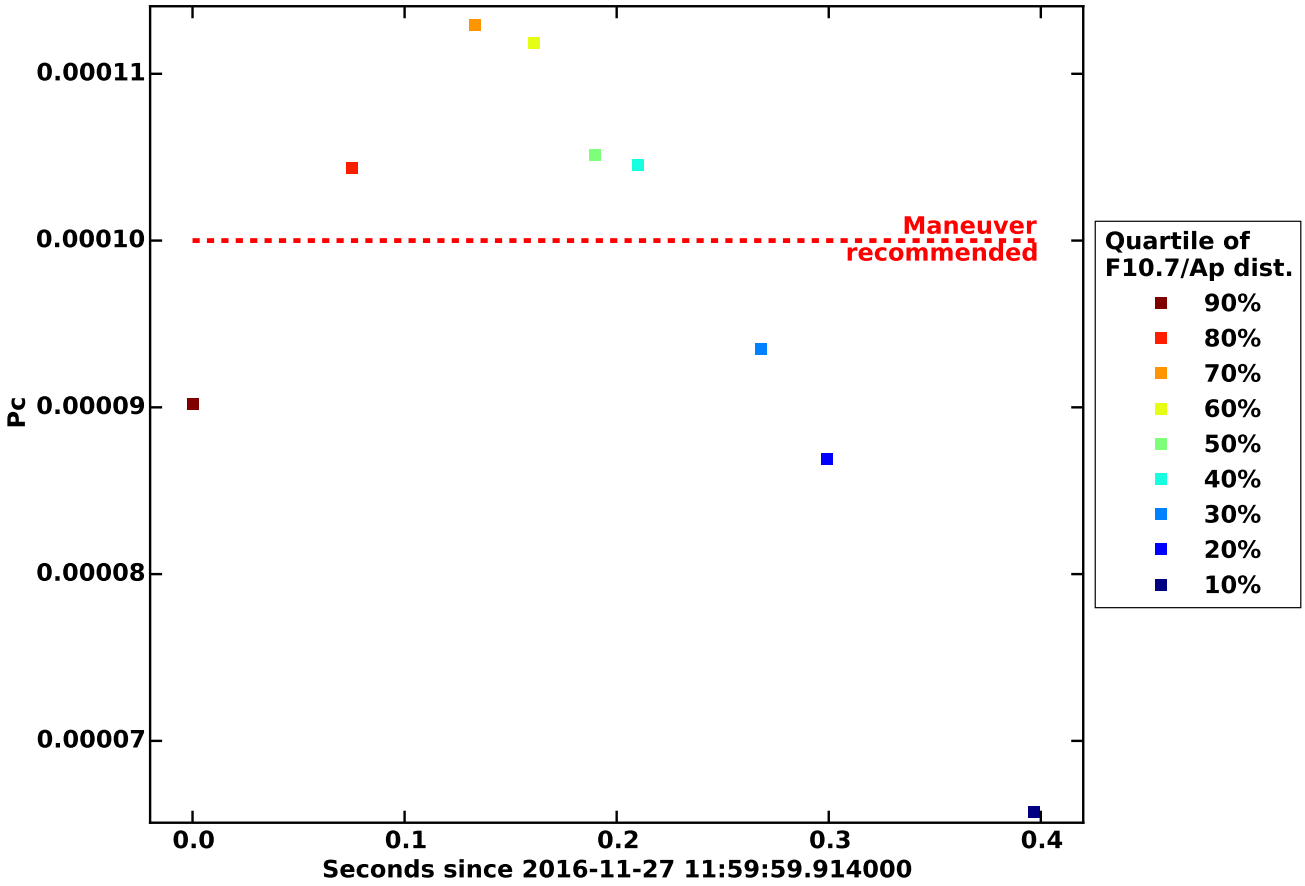

Figure 4.12: Probability of collision as a function of unperturbed time of closest approach for different solar activity conditions.

This analysis shows that uncertainties in the prediction of the thermospheric density can lead to important errors on the probability of collision if only the nominal prediction is considered for the collision risk assessment, even if the conjunction occurs only 36 hours after the epoch. More specifically, it proves that although 40% of the possible scenarios for the predicted density (40% to 80% quartiles) lead to probabilities of collision above the collision avoidance maneuver threshold, there is also high probability that the actual density is such that the probability of collision is lower than this threshold. This is represented by the 10%, 20%, 30%, and 90% quartile scenarios, which in other words imply that there is a 30% chance (all scenario with densities below the density of the 30% quartile) added to a 10% chance (all sce-

nario which density is above the density of the 90% quartile) that the probability of collision is $\sim$ 25-30% lower than expected.

4.4.3 Probability of collision errors due to a miss prediction of a solar storm

In this section, a different problem is analyzed. The goal in the previous section was to study the effects of uncertainties in the prediction of the atmospheric density on the probability of collision. This section studies the impact of a geomagnetic storm on the probability of collision.

For this analysis, the geomagnetic storm occurring on March 17th, 2015 was studied. Two orbits are propagated for two days, starting on March 16th. The encounter geometry corresponds to a parallel path conjunction: both orbit are at an inclination of 45° at 400 km. Object 2's eccentricity is slightly higher than Object 1's eccentricity (0.00002 and 0.00001 respectively). This is a particularly interesting configuration as the assumption of short time of encounter used in many collision risk assessment algorithms does not hold for this parallel conjunction geometry.

The geomagnetic index Ap was reported to reach values as high as 200 on March 17. The orbit averaged density modeled at the position of the satellites using NRLMSISE is shown in Figure 4.13. The red line shows the density modeled from the actual F10.7 and Ap and the blue line represents the density modeled with keeping F10.7 and Ap constant to the initial value on March 16th.

To study the effect of the storm on the probability of collision, two simulations were made. The first one used the actual solar activity conditions to model the atmospheric density (red line in Figure 4.13). The second one used a constant solar activity for the two days of the simulations (blue line in Figure 4.13). This simulation corresponded to a situation without a geomagnetic storm. For both runs, the risk of collision was assessed by SpOCK to compare the probability of collision and the time

of conjunction.

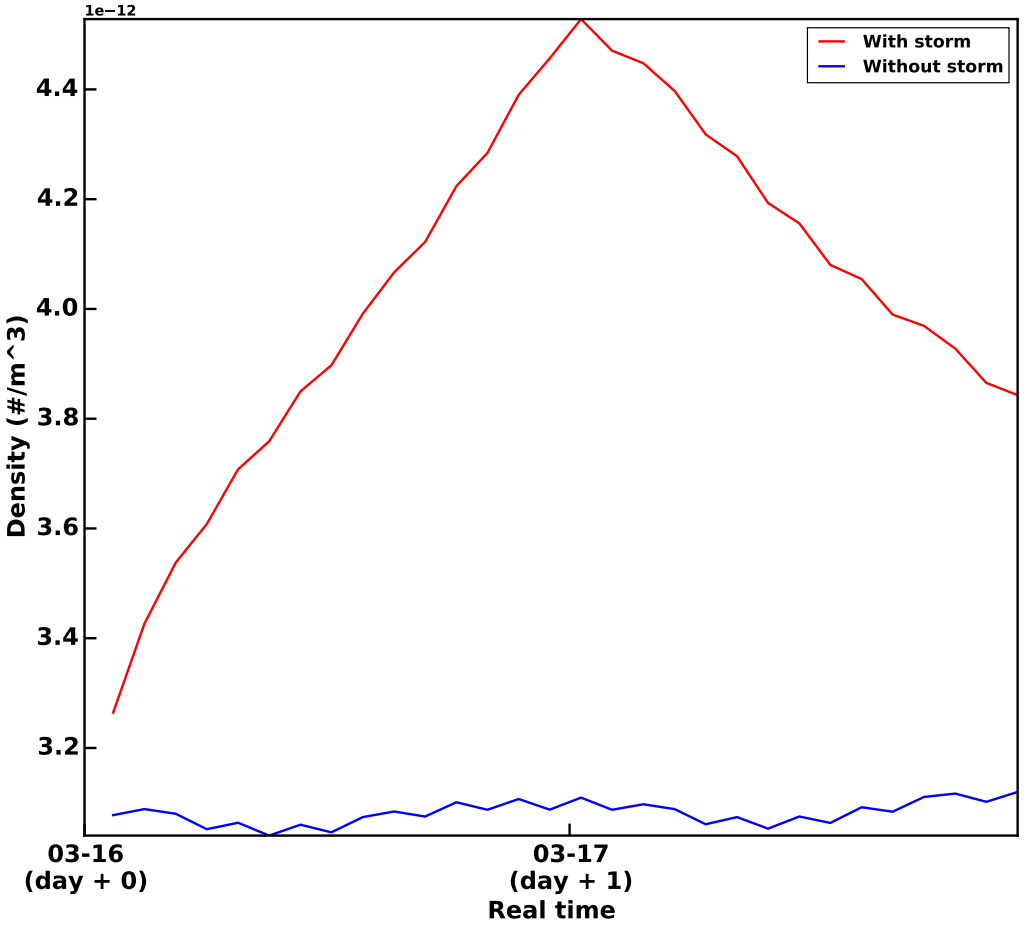

Figure 4.13: Orbit average density at the position of the spacecraft with the geomagnetic storm (red) and without the geomagnetic storm (blue).

Figure 4.14 shows the cumulative probability distribution function as a function of time for both simulations, defined as the total number of recorded collisions as a function of time, divided by the total number of possible conjunctions. Consequently, the total cumulative probability of collision at the end of the span is equal to the probability of collision N_T/N used so far in the study. The red line corresponds to the simulation with the geomagnetic storm, and the blue line without it. The vertical dashed lines represent the times of closest approach between the unperturbed orbits, which are also indicated in the top left corner of the graph.

The conjunction occurs sooner in the presence of the storm by ~ 2 seconds. This is due to a stronger drag because of the increase in density resulting from the effects of the geomagnetic storm on the thermosphere. Therefore, the satellites orbit with higher velocities and the close approach occurs sooner than without the storm.

The probability of collision in the absence of the storm is slightly above the maneuver threshold (1.13×10^{-4}). However, with the storm, the probability of collision is 8.80×10^{-5}, which is under the maneuver threshold. This $\sim 30\%$ difference in the probability of collision not only demonstrates the important effect of the storm on the probability of collision, but it also shows that if on March 16th, the space weather models had not predicted the storm to occur a day later, the predicted probability of collision would have been right above the maneuver threshold so the mission operator would have probably been advised to perform a collision avoidance maneuver. This miss prediction of the storm would have led to a false alert, which means that the maneuver would have unnecessarily been carried out. This approach is not specific to this storm but can be applied to any storm and shows not only that a storm can greatly change the probability of collision but can also result in making the wrong decision with regard to a collision avoidance maneuver.

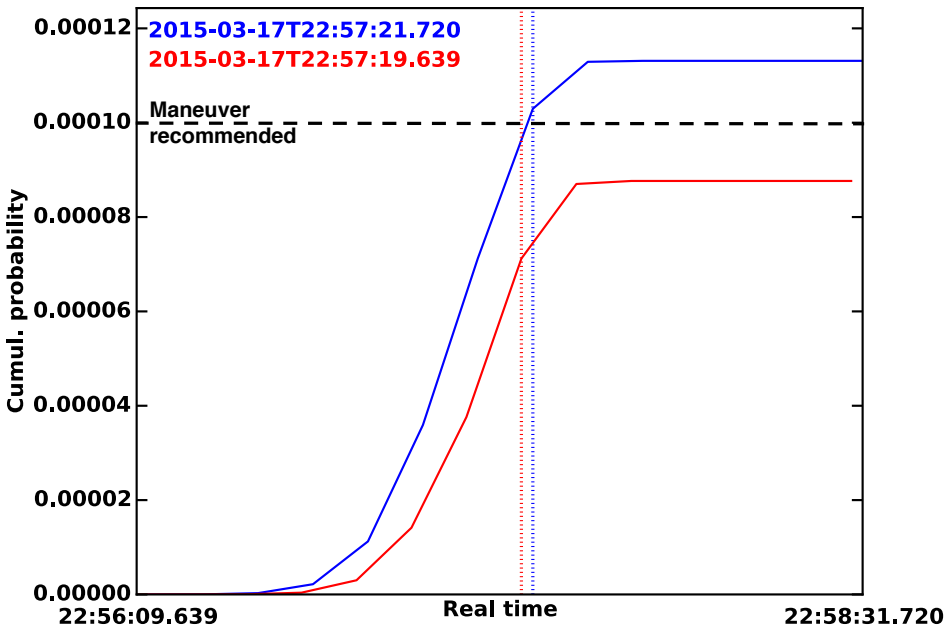

Figure 4.14: Cumulative probability of collision as a function of time during the encounter. The scenario with the storm is in red and without the storm in blue. The vertical dashed lines represent the times of closest approach between the unperturbed orbits (indicated on the top left corner). The collision avoidance maneuver threshold (10^{-4}) is represented with a black dashed line.

The influence of the minimum distance of collision on the relative change in probability of collision resulting from the effects of the storm on the thermospheric density is shown in Table 4.6. The effects of the geomagnetic storm increase with the decrease of the minimum distance of collision: the relative difference in the probability of collision increases with the decrease of the minimum distance (so does the absolute difference, not shown in Table 4.6).

The storm perturbed the orbits by increasing the distance at closest approach between two objects in some cases, and by decreasing the distance at closest approach between two objects in other cases, as explained in Section 4.4.1. The decrease in probability of collision shows that taking into account all possible conjunctions, the

storm overall decreased the total number of encounters in this particular example. In other words, there are more cases where it increased the distance at closest approach from a value below the minimum distance (without the storm) to a value above the minimum distance (with the storm) than cases where it decreased the distance at closest approach from a value above the minimum distance (without the storm) to a value below the minimum distance (with the storm).

The fact that the probability decreases more as a result of lowering the minimum distance of collision in the presence of a storm means that the number of conjunctions decreases more with the storm than without the storm if the minimum distance of collision is smaller. To understand why, consider a situation where the storm increased the distance at closest approach from 1.5 m (without the storm) to 4 m (with the storm). The situation is recorded as a collision with and without the storm if the minimum distance of collision if 5 m. If the minimum distance is now 2 m, this situation is still recorded as a collision without the storm but it is not recorded anymore as a collision with the storm. In other words, the total number of collisions without the storm is still the same with the new threshold but it is smaller by one conjunction in the presence of the storm.

Consequently, for lower probability of collisions (10^{-5}, 10^{-6} for instance), the relative difference is expected to be even greater. An application of this study is that there is often an unknown on the minimum distance of collision to set for the evaluation of the probability of collision. Recall that this distance corresponds to the sum of the two object radii only if the objects are spherical, which is not the case most of the time, particularly if the objects are satellites. Satellites can also have tethers or booms, such as the CHAllenging Mini-satellite Payload (CHAMP) satellite. Therefore, the attitude of the satellite has a direct influence on the minimum distance of collision. Unfortunately, a perfect knowledge of the attitude is impossible, resulting in uncertainties on the minimum distance of collision. This analysis showed that the

Table 4.6: Probability of collision with and without the geomagnetic storm for different values of the minimum distance of collision.

Min dist. collision (m)	P_c with storm	P_c without storm	Relative difference
5	0.001701	0.001559	9.1%
2	0.000338	0.000276	22.5%
1.2	0.000113	0.000088	28.4%
1.0	0.000073	0.000056	30.4%

effects of a miss prediction of a storm can vary with the attitude of the spacecraft at the time of closest approach.

4.5 Conclusion

The Conjunction Assessment Risk Analysis algorithm developed in SpOCK was first presented. It uses Monte Carlo procedures to predict the probability of collision from the covariance matrices of the two objects at epoch. SpOCK propagates ensemble members that represent small perturbations of the initial positions and velocities by modeling the perturbing forces such as the perturbations of the gravitational potential due to the asphericity and non-uniform mass distribution of the Earth, the atmospheric drag, the solar radiation pressure and the gravitational perturbations from the Moon and the Sun. More specifically, it uses a thermosphere model, NRLMSISE or GITM, to accurately model the atmospheric density at the position of the spacecraft, hence the drag acceleration. While propagating the perturbed orbits, SpOCK screens for collisions between all ensemble members. Specifically, it interpolates the minimum distance between two objects with fifth order polynomials. If this distance is smaller than the sum of the two object radii, it records the situation as a conjunction. After repeating this operation for all ensemble members, SpOCK divides the total number of encounters by the total number of cases, usually at least hundred of millions, which gives the probability of collision.

Although atmospheric drag is one of the main perturbing forces for LEO orbits (representing 90% to 95% of the force in the in-track direction), uncertainties in the density are usually not taken into account in collision risk assessment algorithms. Important uncertainties rely on the atmospheric density because the coupling of the upper atmosphere with the Earth space environment (the ionosphere and the magnetosphere) is not well understood, and because the solar activity, the main driver of this coupled system, is itself really hard to predict. CMEs, solar flares, CIRs cause important and hardly predictable disturbances of the atmospheric density. Therefore, ignoring all these effects results in important errors in the prediction of the probability of collision. After introducing the approach for taking into account the uncertainties

in the solar activity for the determination of the probability of collision, the effects of these uncertainties on the risk of collision were studied. More specifically, an example demonstrated that errors up to 30% are likely to be made if the density goes below the 30% quartile of the predicted distribution or above the 90% quartile. This large range of possible values for the density makes it very likely that the actual probability of collision deviates from the single expected value if only the nominal prediction of the density is used to assess the risk of collision. The example showed that this can result in mission operators making the wrong decision with regard to collision avoidance maneuvers.

Finally, the effects of geomagnetic storms on the probability of collision were illustrated with an example of a parallel conjunction between two spacecraft two days after epoch. On March 17th, 2015, a strong storm hit the upper atmosphere, with values of the geomagnetic index Ap exceeding 200. Density enhancements by almost 50% at 400 km strongly increased the drag, modifying the orbits of the satellites. The effects on the risk of collision led to differences up to 30% in the probability of collision, compared to a situation without a storm. This relative difference was proved to increase as the minimum distance of collision decreases. This analysis demonstrates the importance of predicting geomagnetic storms. In this example, the miss prediction of the geomagnetic storm resulted in a false alert because the probability of collision without the storm was above the threshold for a collision avoidance maneuver.

CHAPTER V

Conclusion

Atmospheric drag is the most important perturbing force for LEO orbits (after gravity). Consequently, large uncertainties in this force result in large uncertainties in the positions and velocities of satellites, and cannot be neglected for high accuracy orbit propagations. However, the risk of collision between space objects is often assessed by neglecting the uncertainties in the atmospheric density. This thesis presented an algorithm that takes into account uncertainties in the atmospheric density for predictions of the risk of collision.

The Probability Distribution Function model was developed to predict the solar wind speed 5 days in advance. The solar wind speed is a driver of the Earth's upper atmosphere and has an influence on its neutral density, which in turn affects the drag on spacecraft. Therefore, predicting the solar wind speed is an important and necessary step in predicting the atmospheric drag. In addition to predicting the solar wind speed, the PDF model quantifies the uncertainties associated with the main prediction, which can be directly converted into uncertainties in the atmospheric density, hence into uncertainties in the drag acceleration. The most important conclusions that can be drawn from the PDF model are (Chapter III):

1. The first version of the PDF model focused on the prediction of the background solar wind speed, which corresponds to times when the solar activity is quiet or

moderate.

2. The PDF model uses a linear combination of probability distribution functions based on the current solar wind speed and trend with the speed from approximately 27 days ago. These PDFs can be used to generate ensemble prediction scenarios for the solar wind speed or to assign an uncertainty on the prediction based solely on the median speed.

3. The PDF Model performs better than the WSA Model for one-day ahead predictions. For longer prediction horizons, both models perform about the same. For 2008, the last solar minimum, the PDF Model performs better than the WSA Model for all prediction horizons, with a 15 km/s difference in the accuracy of the predictions.

4. The second version of the PDF Model allows for the prediction of high speed events, which are particularly important because they cause the most important disturbances of the atmospheric density, consequently of the drag acceleration.

5. To predict the peak value of the speed, PDFs of the amplitude of the peak were calculated as a function of the maximum slope before the peak. The distribution of the time it takes for the speed to increase from the beginning of the peak to the maximum of the peak was used to predict the time of the peak.

6. It was found that 60.4% of the positive predictions (prediction that a peak will occur) were correct, while 91.4% of the negative predictions (prediction that a peak will not occur) were correct. 20.3% of the peaks in the speed are found by the model. The percentage increases to 33.6% when there is an associated peak in both the solar wind density and IMF magnitude before the increase in the solar wind speed. The number of false positives is more than three times smaller for the PDF model compared with the WSA model.

7. Ensemble predictions of high speed events by the PDF model provide the forecast community with an interval of uncertainty on the prediction. The study showed that on average, the observations were in between the 25%-75% quartiles $\sim$ 60 hours out of 120 hours and in between the 10%-90% quartiles $\sim$ 85 hours out of 120 hours. This can in turn give an interval of uncertainty on the prediction of the atmospheric density, to ultimately quantify uncertainties in the drag acceleration.

To predict the trajectories of spacecraft orbiting the Earth under the influence of atmospheric drag, the Spacecraft Orbital Characterization Kit was developed. This tool enables a highly accurate prediction of satellites motions. The numerical integration of SpOCK uses a comprehensive modeling of the dynamics of spacecraft in orbit by taking into account the perturbing forces acting on the satellite. Specifically, the non-spherical portion of the mass distribution of the Earth is modeled with a decomposition of spherical harmonics for the gravitational potential. Moreover, SpOCK uses thermospheric models (NRLMSISE and GITM) to derive the atmospheric density at the position of the spacecraft, allowing for an accurate representation of the atmospheric drag. Results showed the error resulting from neglecting perturbing forces and from integrating the trajectories with time steps that were too large. In addition to modeling the motion of satellites, the different functionalities of SpOCK were presented. For instance, the coverage of the three ground stations that communicate with the CYGNSS observatories was computed with the determination of the access times over a day. The precession of the ascending node due to the asphericity of the Earth was demonstrated over a 7 day simulation. The orbit-average solar power of a 1U CubeSat in a polar orbit with wings at different angles was derived to study the influence of the spacecraft geometry on the solar power generated by the solar cells of the satellite. The influence of seasons and of the precession of the nodes were shown too. Finally, a demonstration of SIFT was made. This algorithm pre-

dicts the intersection of the CYGNSS observatories' specular points trajectories with the path of tropical storms forecast by NOAA. It provides support to the CYGNSS mission operation center in sending the commands a day and half in advance to tell the observatories when to switch to the higher resolution mode in order to take more accurate measurements of the winds in the cyclones. SpOCK was validated by comparing the inertial positions and velocities with results from HPOP within STK from AGI. Specifically, the ephemerides after one day of propagation were compared for three different orbits: Low Earth Orbit (altitude = 380 km), High Elliptical Orbit (eccentricity = 0.74, semi-major axis = 26,553 km), and Geostationary Orbit (altitude = 35,786 km). The complete force model was validated by evaluating the perturbing forces independently. The comparison showed sub-meter differences between SpOCK's and HPOP's positions, except for the validation of the drag model in the LEO configuration for which a difference of 10 meters was found, probably due to a difference in the modeling of the density between the two propagators. Differences in velocities did not exceed several mm/s (Chapter II).

Initially designed for mission design and analysis, SpOCK was then improved for the assessment of the risk of collision between space objects. As the population of objects in orbit around the Earth dramatically increased in the past decade, the risk of collisions with operational satellites and astronauts has become a serious threat, highlighting the importance of accurately predicting when such collisions are likely to occur. The Conjunction Assessment Risk Analysis algorithm developed in SpOCK was first presented in Chapter IV. Monte Carlo procedures are used to predict the probability of collision from the covariance matrices of the two objects at the initial epoch. SpOCK propagates ensemble members, which represent small perturbations of the initial positions and velocities. While propagating the perturbed orbits, SpOCK screens for collisions between all ensemble members. Specifically, it calculates the minimum distance between two objects with fifth order polynomials. If this distance

is smaller than the sum of the two object radii, it records the situation as a conjunction. After repeating this operation for all ensemble members, SpOCK divides the total number of encounters by the total number of cases, usually at least hundreds of millions, which gives the probability of collision. The highly populated altitude regime below 800 km represents a particular challenge because atmospheric drag considerably perturbs the trajectories of objects. Large uncertainties in the drag acceleration are the results of important uncertainties in the atmospheric density. Ignoring the effects of uncertainties in the atmospheric density results in important errors in the prediction of the probability of collision. Despite these considerations, the risk of collision is usually assessed by neglecting the uncertainties in the atmospheric density. After introducing the approach for taking into account uncertainties in the atmospheric density in the determination of the probability of collision, their effects on the risk of collision were studied. An example showed errors up to 30% in the probability of collision if the density would go below the 30% quartile of the predicted distribution or above the 90% quartile. This proved that the actual probability of collision can deviate from the single expected value using solely the nominal prediction of the density. The example demonstrated that this approach can result in mission operators making the wrong decision with regard to collision avoidance maneuvers. Finally, the effects of geomagnetic storms on the probability of collision were illustrated with an example of a parallel conjunction two days after epoch. On March 17th, 2015, a strong storm hit the upper atmosphere, with values of the geomagnetic index Ap exceeding 200. Density enhancements by almost 50% at 400 km strongly increased the drag, modifying the orbits of the satellites. The effects on the risk of collision led to differences up to 30% in the probability of collision, compared to a situation without a storm. This relative difference was proved to increase as the minimum distance of collision decreases. This analysis demonstrated the importance of predicting geomagnetic storms. In this example, the miss prediction of the geomagnetic storm

resulted in a false alert because the probability of collision without the storm was above the threshold for a collision avoidance maneuver (Chapter IV).

The ultimate goal of this research is for SpOCK to support mission operators in making the correct decision with regard to a potential collision avoidance maneuver. The thesis demonstrated that uncertainties in the atmospheric density result in large errors in the prediction of the probability of collision. Further studies will investigate the dependence of the encounter geometry on the effects of such uncertainties. Finally, SpOCK considers the objects to be spherical when evaluating the probability of collision. The algorithm will be improved to take into account the exact geometries of the satellites so that the determination of the probability of collision is more accurate.

APPENDIX

APPENDIX A

Initial states and covariance matrices

Distances are expressed in meters, and time in seconds.

$$\overline{U}_{1,t_0} = \begin{pmatrix} 202780.692 \\ -5873336.722 \\ -3356137.888 \\ 7644.314213 \\ 540.915272 \\ -493.855868 \end{pmatrix} \tag{A.1}$$

$$C_1 = \begin{pmatrix} 4.744089478916300000000 & -1.258327906777000000000 & -1.258327906777000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ -1.258327906777000000000 & 6.127955260541900000000 & 2.127955260541900000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ -1.258327906777000000000 & 2.127955260541900000000 & 6.127955260541900000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000001000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000001 & 0.000001000000000000000 & -0.000000000000000000001 \\ 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000001 & -0.000000000000000000001 & 0.000001000000000000000 \end{pmatrix} \tag{A.2}$$

$$\overline{U}_{2,t_0} = \begin{pmatrix} 558232.011 \\ -3649646.379 \\ -5675244.936 \\ 6569.617992 \\ 3593.547352 \\ -1664.494700 \end{pmatrix} \tag{A.3}$$

$$C_2 = \begin{pmatrix} 4.743951262471500000000 & -1.258255000046000000000 & -1.258207926532000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ -1.258255000046000000000 & 6.128103983286500000000 & 2.128024367275000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ -1.258207926532000000000 & 2.128024367275000000000 & 6.127944754242000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000001000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 \\ 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000001000000000000000 & -0.000000000000000000002 \\ 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & 0.000000000000000000000 & -0.000000000000000000002 & 0.000001000000000000000 \end{pmatrix} \tag{A.4}$$

BIBLIOGRAPHY

BIBLIOGRAPHY

Aida, S., and M. Kirschner (2010), Collision risk assessment and operational experiences for leo satellites at gsoc.

Aida, S., T. Patzelft, L. Leushacke, M. Kirschner, and R. Kiehling (2009), Monitoring and Mitigation of Close Proximities in Low Earth Orbit, in *21th International Symposium on Space Flight Dynamics*, CNES, Toulouse, France.

Aida, S., M. Kirschner, and F. Meissner (2015), Collision risk assessment and mitigation strategy for the gsoc geo satellites.

Akella, M. R., and K. T. Alfriend (2000), Probability of collision between space objects, *Journal of Guidance, Control, and Dynamics*, *23*(5), 769–772.

Alfano, S. (1994), Determining satellite close approaches, *Journal of the Astronautical Sciences*, *42*, 143–152.

Alfano, S. (2005a), A numerical implementation of spherical object collision probability, *Journal of the Astronautical Sciences*, *53*(1), 103.

Alfano, S. (2005b), Beta conjunction analysis tool, *AAS/AIAA Astrodynamics Specialist Conference*.

Alfano, S. (2006), Addressing nonlinear relative motion for spacecraft collision probability, in *AAS/AIAA Astrodynamics Specialist Conference, Keystone, CO., Paper*, vol. 6760.

Alfano, S. (2009), Satellite conjunction monte carlo analysis, *AAS Spaceflight Mechanics Mtg, Pittsburgh, PA., Paper*, pp. 09–233.

Alfano, S., and J. D. Thorne (1994), Circle-to-circle constant-thrust orbit raising, *Journal of the Astronautical Sciences*, *42*(1), 35–45.

Alfriend, K. T., M. R. Akella, J. Frisbee, J. L. Foster, D.-J. Lee, and M. Wilkins (1999), Probability of collision error analysis, *Space Debris*, *1*(1), 21–35.

Allan, R. R., and G. E. Cook (1964), The Long-Period Motion of the Plane of a Distant Circular Orbit, *Proceedings of the Royal Society of London Series A*, *280*, 97–109, doi:10.1098/rspa.1964.0133.

Arge, C., and V. Pizzo (2000), Improvement in the prediction of solar wind conditions using near-real time solar magnetic field updates, *Journal of Geophysical Research*, *105*, 10,465–10,479.

Arge, C. N., D. Odstrcil, V. J. Pizzo, and L. R. Mayer (2003), Improved method for specifying solar wind speed near the sun, *AIP Conf. Proc.*, *679*, 190–193.

Bowman, B. R., W. K. Tobiska, and M. J. Kendra (2008a), The thermospheric semi-annual density response to solar euv heating, *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(11), 1482–1496.

Bowman, B. R., W. K. Tobiska, F. A. Marcos, and C. Valladares (2008b), The jb2006 empirical thermospheric density model, *Journal of Atmospheric and Solar-Terrestrial Physics*, *70*(5), 774–793.

Bruinsma, S., J. M. Forbes, R. S. Nerem, and X. Zhang (2006), Thermosphere density response to the 20-21 November 2003 solar and geomagnetic storm from CHAMP and GRACE accelerometer data, *Journal of Geophysical Research (Space Physics)*, *111*, A06303, doi:10.1029/2005JA011284.

Bruinsma, S. L., and J. M. Forbes (2007), Global observation of traveling atmospheric disturbances (tads) in the thermosphere, *Geophysical Research Letters*, *34*(14).

Buizza, R. (2001), Accuracy and Potential Economic Value of Categorical and Probabilistic Forecasts of Discrete Events, *American Meteorological Society*, doi: http://dx.doi.org/10.1175/1520-0493(2001)129⟨2329:AAPEVO⟩2.0.CO;2.

Bussy-Virat, C. D., and A. J. Ridley (2014), Predictions of the solar wind speed by the probability distribution function model, *Space Weather*, *12*, 337–353, doi: 10.1002/2014SW001051.

Bussy-Virat, C. D., and A. J. Ridley (2016), Twenty-four hour predictions of the solar wind speed peaks by the probability distribution function model, *Space Weather*, *14*, 861–873, doi:10.1002/2016SW001437.

Cash, M. D., D. A. Biesecker, V. Pizzo, C. A. de Koning, G. Millward, C. N. Arge, C. J. Henney, and D. Odstrcil (2015), Ensemble modeling of the 23 july 2012 coronal mass ejection, *Space Weather*, *13*(10), 611–625, doi:10.1002/2015SW001232, 2015SW001232.

Chan, F. (2008), *Spacecraft collision probability*, Aerospace Press.

Chan, K. (2002), Spacecraft maneuvers to mitigate potential collision threats, in *AAS/AIAA Astrodynamics Specialist Conference, Monterey, CA*.

Clarizia, M. P. (2015), Cyclone Glocal Navigation Satellite System - Algorithm Theoretical Basis Document, Level 2 Wind Speed Retrieval, *Technical Note*.

Council, N. R., et al. (2011), Limiting future collision risk to spacecraft: An assessment of nasa's meteoroid and orbital debris programs.

D. Pachura, M. D. H. (2016), Conjunction Assessment Late-Notice High-Interest Event Investigation: Space Weather Aspects, *NASA robotic CARA*.

Dagum, P., R. Karp, M. Luby, and S. Ross (2000), An optimal algorithm for monte carlo estimation, *SIAM Journal on computing*, *29*(5), 1484–1496.

Dryer, M. (1974), Interplanetary shock waves generated by solar flares, *Space Science Reviews*, *15*, 403–468.

Emmert, J. (2015), Thermospheric mass density: A review, *Advances in Space Research*, *56*(5), 773–824.

Emmert, J., and J. Picone (2010), Climatology of globally averaged thermospheric mass density, *Journal of Geophysical Research: Space Physics*, *115*(A9).

Emmert, J., J. Byers, H. Warren, and A. Segerman (2014), Propagation of forecast errors from the sun to leo trajectories: How does drag uncertainty affect conjunction frequency?, *Tech. rep.*, DTIC Document.

Emmert, J., H. Warren, A. Segerman, J. Byers, and J. Picone (2016), Propagation of atmospheric density errors to satellite orbits, *Advances in Space Research*.

Emmons, D., A. Acebal, A. Pulkkinen, A. Taktakishvili, P. MacNeice, and D. Odstrcil (2013), Ensemble forecasting of coronal mass ejections using the WSA-ENLIL with CONED Model, *Space Weather*, *11*, 95–106, doi:10.1002/swe.20019.

Forbes, J. M., G. Lu, S. Bruinsma, S. Nerem, and X. Zhang (2005), Thermosphere density variations due to the 15-24 April 2002 solar events from CHAMP/STAR accelerometer measurements, *Journal of Geophysical Research (Space Physics)*, *110*, A12S27, doi:10.1029/2004JA010856.

Foster, J. (2001), The analytical basis for debris avoidance operations for the international space station and space shuttle, *Orbital Debris Quarterly News*, *6*(2), 11.

Fry, C., W. Sun, C. Deehr, M. Dryer, Z. Smith, S.-I. Akasofu, M. Tokumaru, and M. Kojima (2001), Improvements to the HAF solar wind model for space weather predictions, *Journal of Geophysical Research*, *106*, 20,985–21,001.

Fry, C., M. Dryer, Z. Smith, W. Sun, C. Deehr, and S.-I. Akasofu (2003), Forecasting solar wind structures and shock arrival times using an ensemble of models, *Journal of Geophysical Research*, *108*.

Fry, C., T. Detman, M. Dryer, Z. Smith, W. Sun, C. Deehr, S.-I. Akasofu, C.-C. Wu, and S. McKenna-Lawlor (2007), Real-time solar wind forecasting: Capabilities and challenges, *Journal of Atmospheric and Solar-Terrestrial Physics*, *69*, 109–115.

Gaylor, D., R. Page III, and K. Bradley (2006), Testing of the Java Astrodynamics Toolkit Propagator, *American Institute of Aeronautics and Astronautics*, *AIAA 2006-6754*.

Gombosi, T. I. (1998), *Physics of the space environment*, Cambridge University Press.

He, H., H. Wang, Z. Du, R. Li, Y. Cui, L. Zhang, and Y. He (2008), Solar activity prediction studies and services in naoc, *Advances in Space Research*, *42*(9), 1450–1456.

Hejduk, M., and R. Frigm (2015), Collision Avoidance Short Course - Part I: Theory, *NASA robotic CARA*.

Horsley, M. (2012), Satellite re-entry modeling and uncertainty quantification, in *Advanced Maui Optical and Space Surveillance Technologies Conference*, vol. 1, p. 52.

Huang, C., D.-D. Liu, and J.-S. Wang (2009), Forecast daily indices of solar activity, f10. 7, using support vector regression method, *Research in Astronomy and Astrophysics*, *9*(6), 694.

Hundhausen, A. (1999), *Coronal Mass Ejections*, pp. 143–200, Springer New York, New York, NY.

Kim, R., et al. (2010), An empirical model for prediction of geomagnetic storms using initially observed CME parameters at the Sun, *Journal of Geophysical Research*, *115*.

Laporte, F. (2014), JAC Software, Solving Conjunction Assessment Issues, in *Advanced Maui Optical and Space Surveillance Technologies Conference*, p. E4.

Lauri K. Newman, M. D. H. L. C. J. (2016), Operational implementation of a pc uncertainty construct for conjunction assessment risk analysis.

Leamon, R. J., and S. McIntosh (2007), Empirical solar wind forcasting from the chromosphere, *The Astrophysical Journal*, *659*, 738–742.

Lee, B.-S., and B.-Y. K. Yoola Hwang, Hae-Yeon Kim (2012), Geo satellite collision avoidance maneuver strategy against inclined gso satellite.

Lee, C., J. Luhmann, D. Odstrcil, P. MacNeice, I. de Pater, P. Riley, and C. Arge (2009), The Solar Wind at 1 AU During the Declining Phase of Solar Cycle 23: Comparison of 3D Numerical Model Results with Observations, *Solar Physics*, *254*, 155–183.

Lei, J., J. P. Thayer, W. Wang, and R. L. McPherron (2011), Impact of cir storms on thermosphere density variability during the solar minimum of 2008, *Solar Physics*, *274*(1-2), 427–437.

Lei, J., T. Matsuo, X. Dou, E. Sutton, and X. Luan (2012), Annual and semiannual variations of thermospheric density: Eof analysis of champ and grace data, *Journal of Geophysical Research: Space Physics*, *117*(A1).

Lemoine, F. G., et al. (1997), *The Development of the NASA GSFC and NIMA Joint Geopotential Model*, pp. 461–469, Springer Berlin Heidelberg, Berlin, Heidelberg.

Liou, J.-C. (2016), The Orbital Debris Problem, *Space Tech Conference*.

Liu, D. D., C. Huang, J. Y. Lu, and J. S. Wang (2007), The hourly average solar wind velocity prediction based on support vector regression method, *Monthly Notices of the Royal Astronomical Society*, *413*, 28772882.

Liu, J., L. Liu, B. Zhao, J. Lei, J. P. Thayer, and R. L. McPherron (2012), Superposed epoch analyses of thermospheric response to cirs: Solar cycle and seasonal dependencies, *Journal of Geophysical Research: Space Physics*, *117*(A9).

Luo, B., Q. Zhong, S. Liu, and J. Gong (2008), A new forecasting index for solar wind velocity based on EIT 284 A Observations, *Solar Physics*, *250*, 159–170.

MacNeice, P. (2009a), Validation of community models: Identifying events in space weather model timelines, *Space Weather*, *7*, S06,004.

MacNeice, P. (2009b), Validation of community models: 2. Development of a baseline using the Wang-Sheeley-Arge model, *Space Weather*, *7*, S12,002.

MacNiece, P. (2009), Validation of community models: 2. Development of a baseline using the Wang-Sheeley-Arge model, *Space Weather*, *7*, S12,002.

Mays, M. L., et al. (2015), Ensemble Modeling of CMEs Using the WSA-ENLIL+Cone Model, *Solar Physics*, *290*, 1775–1814, doi:10.1007/s11207-015-0692-1.

McKenna-Lawlor, S. M. P., M. Dryer, M. D. Kartalev, Z. Smith, C. D. Fry, W. Sun, C. S. Deehr, K. Kecskemety, and K. Kudela (2006), Near real-time predictions of the arrival at Earth of flare-related shocks during Solar Cycle 23, *Journal of Geophysical Research (Space Physics)*, *111*, A11103, doi:10.1029/2005JA011162.

McKinley, D. (2008), Maneuver planning for conjunction risk mitigation with ground-track control requirements.

McNish, A., and J. Lincoln (1949), Prediction of sunspot numbers, *Eos, Transactions American Geophysical Union*, *30*(5), 673–685.

Moe, K., and M. M. Moe (2005), Gas surface interactions and satellite drag coefficients, *Planetary Space Science*, *53*, 793–801, doi:10.1016/j.pss.2005.03.005.

National Imagery and Mapping Agency (2000), Department of defense world geodetic system 1984: its definition and relationships with local geodetic systems, *Tech. Rep. TR8350.2*, National Imagery and Mapping Agency, St. Louis, MO, USA.

Newman, L., R. Besser, and M. Hejduk (2015), Predicting Space Weather Effects on Close Approach Events, in *Advanced Maui Optical and Space Surveillance Technologies Conference*, p. 39.

Newman, L. K. (2016), NASA Conjunction Assessment Risk Analysis Approach, *NASA robotic CARA*.

Norquist, D., and W. Meeks (2010), A comparative verification of forecasts from two operational solar wind models, *Space Weather*, *8*, 512,005.

Odstrcil, D. (2003), Modeling 3-D solar wind structure, *Adv. Space Res.*, *32*, 497–506.

Odstrcil, D., V. J. Pizzo, J. A. Linker, P. Riley, R. Lionello, and Z. Mikic (2004), Initial coupling of coronal and heliospheric numerical magnetohydrodynamic codes, *Journal of Atmospheric and Solar-Terrestrial Physics*, *66*, 13111320.

Owens, M., C. Arge, H. Spence, and A. Prembroke (2005), An event-based approach to validating solar wind speed predictions: High-speed enhancements in the Wang-Sheeley-Arge model, *Journal of Geophysical Research*, *110*, A12,105.

Owens, M. J., H. E. Spence, S. McGregor, W. J. Hughes, J. M. Quinn, C. N. Arge, P. Riley, J. Linker, and D. Odstrcil (2008), Metrics for solar wind prediction models: Comparison of empirical, hybrid, and physics-based schemes with 8 years of L1 observations, *Space Weather*, *6*, S08,001.

Owens, M. J., R. Challen, J. Methven, E. Henley, and D. R. Jackson (2013), A 27 day persistence model of near-Earth solar wind conditions: A long lead-time forecast and a benchmark for dynamical models, *Space Weather*, *11*, 225–236.

Patera, R. P. (2001), General method for calculating satellite collision probability, *Journal of Guidance, Control, and Dynamics*, *24*(4), 716–722.

Patera, R. P. (2005), Calculating collision probability for arbitrary space vehicle shapes via numerical quadrature, *Journal of guidance, control, and dynamics*, *28*(6), 1326–1328.

Pawlowski, D. J., and A. J. Ridley (2008), Modeling the thermospheric response to solar flares, *Journal of Geophysical Research: Space Physics*, *113*(A10).

Pawlowski, D. J., and A. J. Ridley (2011), The effects of different solar flare characteristics on the global thermosphere, *Journal of Atmospheric and Solar-Terrestrial Physics*, *73*(13), 1840–1848.

Picone, J., A. Hedin, D. P. Drob, and A. Aikin (2002), Nrlmsise-00 empirical model of the atmosphere: Statistical comparisons and scientific issues, *Journal of Geophysical Research: Space Physics*, *107*(A12).

Ram, S. T., C. H. Liu, and S. Su (2010), Periodic solar wind forcing due to recurrent coronal holes during 19962009 and its impact on earths geomagnetic and ionospheric properties during the extreme solar minimum, *Journal of Geophysical Research*, *115*, A12,340.

Rees, M. H., B. A. Emery, R. G. Roble, and K. Stamnes (1983), Neutral and ion gas heating by auroral electron precipitation, *Journal of Geophysical Research: Space Physics*, *88*(A8), 6289–6300, doi:10.1029/JA088iA08p06289.

Richmond, A., and J. Thayer (2000), Ionospheric electrodynamics: A tutorial, *Magnetospheric Current Systems*, pp. 131–146.

Ridley, A., Y. Deng, and G. Toth (2006), The global ionosphere–thermosphere model, *Journal of Atmospheric and Solar-Terrestrial Physics*, *68*(8), 839–864.

Robbins, S., C. Henney, and J. Harvey (2006), Solar wind forecasting with coronal holes, *Solar Physics*, *233*, 265–276.

Roble, R., and E. Ridley (1994), A thermosphere-ionosphere-mesosphere-electrodynamics general circulation model (time-gcm): Equinox solar cycle minimum simulations (30–500 km), *Geophysical Research Letters*, *21*(6), 417–420.

Sanchez-Ortiz, N., R. Domnguez-Gonzlez, H. Krag, and T. Flohrer (2015), Impact on mission design due to collision avoidance operations based on tle or csm information, *Acta Astronautica*, *116*(Complete), 368–381, doi:10.1016/j.actaastro.2015.04.017.

Schatten, K. H., P. H. Scherrer, L. Svalgaard, and J. M. Wilcox (1978), Using dynamo theory to predict the sunspot number during solar cycle 21, *Geophysical Research Letters*, *5*(5), 411–414.

Schunk, R., and A. Nagy (2009), *Ionospheres: physics, plasma physics, and chemistry*, Cambridge university press.

Smith, Z., and M. Dryer (1990), MHD study of temporal and spatial evolution of simulated interplanetary shocks in the ecliptic plane within 1 AU, *Solar Physics*, *1291*, 387–405.

Smith, Z., T. Detman, M. Dryer, C. Fry, C.-C. Wu, W. Sun, and C. Deehr (2004), A verification method for space weather forecasting models using solar data to predict arrivals of interplanetary shocks at Earth, *IEEE Transactions of Plasma Science*, *32*, 14981505.

Smith, Z., M. Dryer, S. McKenna-Lawlor, C. Fry, C. Deehr, and W. Sun (2009), Operational validation of HAFv2s predictions of interplanetary shock arrivals at Earth: Declining phase of Solar Cycle 23, *Journal of Geophysical Research*, *114*.

Stoll, E., R. Schulze, B. DSouza, and M. Oxfort (2011), The impact of collision avoidance maneuvers on satellite constellation management, in *proc. Of European Space Surveillance Conference, Madrid, Spain*.

Stoll, E., B. D'Souza, B. B. Virgili, K. Merz, and H. Krag (2013), Operational collision avoidance of small satellite missions, in *Aerospace Conference, 2013 IEEE*, pp. 1–11, doi:10.1109/AERO.2013.6496955.

Storz, M. F., B. R. Bowman, M. J. I. Branson, S. J. Casali, and W. K. Tobiska (2005), High accuracy satellite drag model (hasdm), *Advances in Space Research*, *36*(12), 2497–2505.

Sutton, E. K., J. M. Forbes, and R. S. Nerem (2005), Global thermospheric neutral density and wind response to the severe 2003 geomagnetic storms from CHAMP accelerometer data, *Journal of Geophysical Research (Space Physics)*, *110*, A09S40, doi:10.1029/2004JA010985.

Sutton, E. K., J. M. Forbes, R. S. Nerem, and T. N. Woods (2006), Neutral density response to the solar flares of October and November, 2003, *Geophysical Research Letters*, *33*, L22101, doi:10.1029/2006GL027737.

Taktakishvili, A., P. J. MacNeice, and M. Hesse (2009), Ensemble Modeling of the Inner Heliosphere Using WSA-ENLIL and SWMF Models, *AGU Fall Meeting Abstracts*, p. B1533.

Tobiska, W. K. (2003), © forecasting of space environment parameters for satellite and ground system operations.

Tobiska, W. K., and S. D. Bouwer (2006), New developments in solar2000 for space research and operations, *Advances in Space Research*, *37*(2), 347–358.

Tobiska, W. K., T. Woods, F. Eparvier, R. Viereck, L. Floyd, D. Bouwer, G. Rottman, and O. White (2000), The solar2000 empirical solar irradiance model and forecast tool, *Journal of Atmospheric and Solar-Terrestrial Physics*, *62*(14), 1233–1250.

Toth, G., and D. Odstrcil (1996), Comparison of some flux corrected transport and total variation diminishing numerical schemes for hydrodynamic and magnetohydrodynamic problems, *Journal of Computational Physics*, *128*, 82–100.

Tsurutani, B. T., et al. (2006), Corotating solar wind streams and recurrent geomagnetic activity: A review, *Journal of Geophysical Research: Space Physics*, *111*(A7).

Vallado, D., and W. McClain (2007), *Fundamentals of Astrodynamics and Applications, Third Edition*, Space Technology Library, Microcosm Press and Springer.

Vandegriff, J., K. Wagstaff, G. Ho, and J. Plauger (2005), Forecasting space weather: Predicting interplanetary shocks using neural networks, *Advances in Space Research*, *36*, 2323–2327, doi:10.1016/j.asr.2004.09.022.

Vasyliūnas, V. M., and P. Song (2005), Meaning of ionospheric joule heating, *Journal of Geophysical Research: Space Physics*, *110*(A2).

Walter, U. (2008), Astronautics.

Wang, Y.-M., and J. N.R. Sheeley (1992), On potential field models of the solar corona, *The Astrophysical Journal*, *392*, 310–319.

Wintoft, P., and H. Lundstedt (1997), Prediction of daily average solar wind velocity from solar magnetic field observations using hybrid intelligent systems, *Phys. Chem. Earth*, *22*, 617–622.

Wright, J. M., T. J. Lennon, R. W. Corell, N. A. Ostenso, W. T. Huntress, J. F. Devine, P. Crowley, and J. B. Harrison (1995), The National Space Weather Program - Strategic Plan, *National Space Weather Program Council*, pp. FCM–P30–1995.

Wyatt, S. P. (1961), The Effect of Radiation Pressure on the Secular Acceleration of Satellites, *SAO Special Report*, *60*.