Eelco Doornbos

Thermospheric Density and Wind Determination from Satellite Dynamics

Springer

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series "Springer Theses" brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student's supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today's younger generation of scientists.

Theses are accepted into the series by invited nomination only and must fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder.
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to scientists not expert in that particular field.

Eelco Doornbos

Thermospheric Density and Wind Determination from Satellite Dynamics

Doctoral Thesis accepted by
the Delft University of Technology, The Netherlands

Springer

Author
Dr. Eelco Doornbos
Faculty of Aerospace Engineering
Delft University of Technology
Kluyverweg 1
2629 HS Delft
The Netherlands
e-mail: e.n.doornbos@tudelft.nl

Supervisor
Dr. Ir. P. N. A. M. Visser
Faculty of Aerospace Engineering
Delft University of Technology
Kluyverweg 1
2629 HS Delft
The Netherlands
e-mail: p.n.a.m.visser@tudelft.nl

Parts of this thesis have been published in the following journal articles and conference papers:

Doornbos, E.N., H. Klinkrad, and P.N.A.M. Visser (2005), Atmospheric density calibration using satellite drag observations, *Advances in Space Research, 36*, 515–521, doi:10.1016/j.asr.2005.02.009

Doornbos, E.N., and H. Klinkrad (2006), Modelling of space weather effects on satellite drag, *Advances in Space Research, 37*, 1229–1239, doi:10.1016/j.asr.2005.04.097

Doornbos, Eelco, Heiner Klinkrad, Remko Scharroo, and Pieter Visser (2007), Thermosphere density model calibration in the orbit determination and prediction of ERS-2 and Envisat, in *Envisat Symposium, Montreux, Switzerland, 23–27 April, 2007, ESA SP–636*, edited by H. Lacoste

Doornbos, Eelco (2007), Thermosphere density model calibration, in *Space weather, research towards applications in Europe*, vol. 344 of *Astrophysics and Space Science Library*, edited by Jean Lilensten, Springer

Doornbos, Eelco, Heiner Klinkrad, and Pieter Visser (2008), Use of two-line element data for thermosphere neutral density model calibration, *Advances in Space Research, 41*(7), 1115–1122, doi:10.1016/j.asr.2006.12.025

Doornbos, Eelco, Jose van den IJssel, Hermann Lühr, Matthias Förster, and Georg Koppenwallner (2010), Neutral density and crosswind determination from arbitrarily oriented multiaxis accelerometers on satellites, *Journal of Spacecraft and Rockets, 47*(4), 580–589, doi:10.2514/1.48114.

Supervisor's Foreword

The thermosphere, regarded as the top level of the Earths atmosphere (100–1000 km altitude), is characterised by a high degree of variability. Despite of being investigated for several decades, our understanding of thermospheric dynamics is still incomplete. This is largely due to the lack of appropriate measurements which are sufficiently well distributed in time and space. At a few 100 km altitude the air density is twelve orders of magnitude lower than at the surface (less than any vacuum achievable in a laboratory). Even though, the remaining air exerts a significant force on satellites orbiting the Earth. This is mainly due to the high orbital velocity (7.5 km/s) and to the fact that the air drag is proportional to the square of the speed. For satellites in low-Earth orbit, air drag is the prime contributor to orbital perturbations. A proper characterisation of the atmospheric properties is therefore indispensable for precise orbit determination. This is in particular vital for monitoring the innumerable objects of space debris orbiting the Earth.

In this work it is outlined how thermospheric properties can be derived from the non-gravitational forces acting on satellites. First, the classical approach is described, which is based on interpreting orbital perturbations in terms of disturbance forces like air drag and radiation pressure. The advantage of this approach is that no active instruments on board are needed. Even passive objects in orbit can contribute to the results. The disadvantage is the coarse spatial resolution of that method. Much refined results are obtained when sensitive accelerometers are used on satellites for measuring the non-gravitational forces acting on the spacecraft. The great achievement of this work is the proper characterisation of the processes influencing the gas-surface interaction. Only after considering these processes in an adequate way, reliable values for air density and wind velocity can be retrieved. The developed algorithm has been applied successfully to the satellites CHAMP and GRACE.

These new and accurate datasets have been used for testing the reliability of commonly used atmospheric models. The comparison with measurements revealed the advantages and limitations of the different kinds of empirical models. This knowledge is of relevance for their operational application, for example, in orbit prediction software. The data obtained in this work show particularly clear the

inability of present atmospheric models to properly reflect the deep depletion of the thermospheric density during the extended solar activity minimum (2008–2009). Besides these more practical aspects the novel data set has also inspired a whole range of scientific studies. Quite prominent among them are investigations of tidal signatures in the upper atmosphere. Some of the observed tides are related to weather and climate phenomena. I am convinced that many more scientific investigations will follow using this unique set of thermospheric density and wind data.

Potsdam, June 2011   Prof. Dr. Hermann Lühr

Preface

Traditionally, the value of a Ph.D. thesis will be determined by its readers in terms of its scientific content. It is a collection of pages filled with words, equations, figures and tables, that hopefully conveys some useful knowledge for certain experts in some field of research. However, for the author it usually represents much more than that. It represents the end of a journey, the journey of the start of a career in science. If the author is fortunate, this has been an exciting journey. But perhaps like in most truly exciting endeavours, there may have been periods of difficulty as well. Much more than the words, equations, figures and tables contained in a thesis, it is the people encountered along this journey, with whom the excitement and difficulties could be shared, that represent the true value of the thesis to its author.

The journey that led to the publication of this work started back in 1999. I was an aerospace engineering student who had always been fascinated by space missions and by computer technology, but at the time I had no clear idea of a possible career path or even of a suitable thesis subject. Fortunately, professor Boudewijn Ambrosius knew just the right subject for me. He introduced me to Remko Scharroo, who was working on precise orbit determination of the ERS-1 and ERS-2 radar altimeter satellites. A follow-on mission, named Envisat, was under construction at the time. Envisat was going to be a massive satellite, with ten scientific instruments, covering many aspects of geophysical and environmental research.

Remko, being an expert on the radar altimeter instruments on these satellites, and having worked on the modelling of gravitational forces which perturb their orbits, gave me the assignment to work on an improved model of the so-called non-gravitational forces: atmospheric drag and radiation pressure. The prediction of these forces requires the 3D modelling of the satellite's external surfaces on the computer. The combination of computer graphics and satellites already made it an interesting enough project from my point of view at the time. Little did I know that I had not only found a great master's thesis topic, but that I had also set my first steps on the path to a career as a research scientist. The computer graphics were

nice, but the research that was required before and after that turned out to be the really interesting parts of the work.

Another very fortunate encounter occurred not much later, when I was given the opportunity to attend the International Astronautical Congress in Amsterdam. Among the wide variety of talks was one by Dr. Heiner Klinkrad, of the European Space Operations Centre ESOC in Darmstadt, Germany. His talk was on a new sophisticated software package for non-gravitational force modelling, called ANGARA, which would be perfect for use in my thesis work. Our meeting in Amsterdam led to a traineeship, partly at HTG, the developers of ANGARA, in Katlenburg-Lindau near Göttingen, and partly at ESOC, where I could implement and test the models in the orbit determination of ERS-2 and Envisat using the new orbit determination software, that was under development there at the time. The enthusiasm of my supervisors during that project (Remko Scharroo at TU Delft, Bent Fritsche and Georg Koppenwallner at HTG and Heiner Klinkrad and Rene Zandbergen at ESOC), as well as that of their colleagues (Dirk Kuijper at ESOC deserves special mention) provided the fuel to kindle my own enthusiasm for continuing this type of work after graduation.

After obtaining my master's degree, I gratefully accepted the invitation to continue working at Boudewijn Ambrosius' astrodynamics group in Delft, on several projects related to satellite radar altimetry and precise orbit determination. I substituted for Remko Scharroo on several of his running European projects, since he had moved on to live and work in the US. This fast introduction into the world of international cooperation on space projects turned out to be the ideal learning ground. I could immediately start to visit many international altimetry-related conferences and project meetings, where I quickly felt part of the vibrant satellite oceanography community. The approaches to multi-mission satellite data storage, processing and visualisation that were emerging in this field at that time, to which I could add my own small contributions, became inspirations for the way I set up the software system for my thermosphere density and wind research in later years. I have fond memories of the many encounters, conversations and collaborations with colleagues during this period, including Pascal Willis, Michiel Otten, Nikita Zelensky, Jean-Paul Berthias, John Ries, Patrick Vincent, Jérôme Benveniste, Richard Francis, David Cotton, Peter Challenor, Yves Menard, Phil Moore, John Lillibridge and many others.

While my work on the altimetry-related projects continued until well into 2008, my curiosity about possible improvements to the modelling of drag on satellites never went away, and I tried to combine these interests whenever possible. A next stage of the journey began with the invitation by Heiner Klinkrad, in 2002 and again in 2004, to work on projects concerning thermosphere density model calibration, in near real-time, for the space debris office at ESA/ESOC. The first project involved a feasibility study, and the second the delivery of a software implementation to ESOC, which was completed in early 2007. Since most of the research and programming work was done by myself, the feedback received from Heiner Klinkrad and Pieter Visser during the course of these projects was very welcome. The invitation, arranged by Heiner Klinkrad, to visit my US colleagues

in Colorado in 2005, funded through the European Office of Aerospace Research & Development with the help of Barrett Flake, proved extremely valuable to my future work as well. At this meeting I met Bruce Bowman, whose HASDM calibrated model was the inspiration for the ESOC assignment, and who was always willing to help me out by providing data and advice. I also met John Wise, who later invited me to come to the Air Force Research Lab in Massachusetts in 2007. There I presented my work and enjoyed the opportunity to have discussions with him and his colleagues, Frank Marcos, William Burke and Chin Lin.

My intention at that time was to turn the results of the two ESOC projects into a Ph.D. thesis, and the drafts of the first chapters of the current book were written during that summer. However, at about the same time, ESA released an invitation to tender for a study investigating the processing of density and wind information from accelerometer data, in preparation for their forthcoming Swarm mission. I had already worked with Dries Caluwaerts on this topic for his M.Sc. thesis, using data from the similar, but already operational CHAMP and GRACE satellites. He did a great job, and I had been looking forward to further expanding his work in the future. These satellites made direct measurements of the accelerations that we had been trying to model for years, using software such as ANGARA. Not only was it now possible to make detailed comparisons of these models with real data, these data could now also be used to gain detailed knowledge on the weakest parts in our models, in this case the thermospheric density and wind. I could not let the opportunity pass to write a proposal in response to the ESA invitation to tender, even though it meant putting the thesis-writing on hold for the time being.

The project that followed from this proposal allowed me to work closely with many colleagues for nearly 2 years. It was an intense period of collaboration for me. Pieter Visser, Jose van den IJssel and Tom van Helleputte at TU Delft, Georg Koppenwallner, Bent Fritsche and Nelli Eswein at HTG, Matthias Förster and Hermann Lühr at GFZ Helmholtz Centre Potsdam, David Rees at Hovemere and Michael Kern and Roger Haagmans at ESA/ESTEC deserve much credit for their contributions to this project, of which large parts can be found in this thesis. I learned a great deal by working together with these colleagues during this period. Additional encounters mainly happened at the bi-annual COSPAR conferences and at the ESA meetings that I attended during these years. Among the people with whom I could share my enthusiasm for my research, and who were always willing to answer my questions or share their results, were John Emmert, Douglas Drob, Marcin Pilinski, Ken and Mildred Moe, Kent Tobiska, Srinivas Bettadpur, Sean Bruinsma, Kathrin Häusler, Patricia Ritter and Huixin Liu.

By mid-2009, the ESA project was finalised, but it had kept me so busy that there was still no Ph.D. thesis, except for a couple of draft chapters dating back to 2007. On the one hand, I was very eager to make use of the momentum I had gained and continue with my research, to initiate new, thermosphere-related projects. Other projects were on the horizon as well, such as the orbit determination of the CryoSat-2 radar altimeter satellite. It was not easy to reach the decision that I had to either cancel or postpone such work, in order to be able to fully concentrate on the writing of this thesis. Once that decision was reached, the

writing itself was often not quite easy either, but in the end it is the result that counts, and that result is now in your hands.

With the thesis now completed, I would like to take this opportunity to first of all thank Boudewijn Ambrosius and Pieter Visser for offering me the freedom to grow professionally while working on such interesting projects, and also for providing me the time required to turn this thesis into something I can be proud of.

Sean Bruinsma, Matthias Förster, Hermann Lühr and Heiner Klinkrad deserve my gratitude for their suggestions for corrections in the final stages of preparation of this thesis. Both Pieter Visser and Erwin Mooij provided excellent feedback at earlier stages.

Many of my other colleagues in Delft deserve special mention as well, for providing such a pleasant environment to work in. First of all, Jose van den IJssel and Nacho Andres, with whom I shared my office (as well as coffee, tea and chocolate) for many years, but also Marc Naeije, Ejo Schrama, Bert Vermeersen, Ron Noomen, Karel Wakker, Relly van Wingaarden, Taco Broerse, Kartik Kumar, Jeroen Melman, Luuk van Barneveld, Hermes Jara Orue, Wouter van der Wal, Paolo Stocchi, Tom van Helleputte, Bert Wouters, Hugo Schotman and Remco Kroes.

Fortunately, there is also a life beyond work. Tom, Aynav, Kristina and Marieke, thank you for being such great friends throughout the years, and to Rachel, in addition, thanks for your encouragement and advice during the toughest part of this journey. To Bauke, Erik, Joost, Dennis, Coen, Robert and many other friends at DDS, thank you for the great times we had on and off the water.

Of course, special thanks go to my parents, sister and brother, and my extended family. It is always nice to come home to you.

And most of all, to Mieke, thank you for your gentle encouragement, your love and good humour, and for celebrating the near-finalisation of this thesis with me so many times in recent months.

Delft/Leiden, December 2010   Dr. Eelco Doornbos

Contents

Chapter 1
Introduction

Neutral density in the thermosphere is one of the most important variables to model for applications in solar-terrestrial physics and drag computations for satellite orbit determination. Thermospheric density varies over a wide range of spatial and temporal scales under the influence of the complex interactions between the Earth system and solar processes. Unfortunately, observation data on the thermosphere have always been quite sparse and measurement modelling uncertainties have often introduced a relatively high level of ambiguity. These characteristics have made the improvement of thermosphere density models quite a challenge in the past.

The availability of more powerful computers and new algorithms has made it possible to gather and process density information from the trajectory analyses of many space objects simultaneously. In addition, the accelerometer instruments on the CHAMP and GRACE satellite missions have provided density data at an unprecedented accuracy and resolution. Under favourable conditions, the data from these instruments have even enabled derivation of crosswind speeds.

Despite the fact that several basic data processing issues remain difficult to solve even with today's knowledge, data and computing power, such developments can be put to use to improve traditional empirical thermosphere models. The processing of spacecraft tracking, trajectory and acceleration data, as well as a basic strategy for adjusting density models using such observations, are the main topics of this thesis.

This introductory Chapter starts by providing an overview of the applications of thermosphere models in Sect. 1.1. The remaining sections of this introduction provide short overviews of topics which will be expanded upon in later Chapters. The various ways in which density and wind variations can be observed by satellites is introduced in Sect. 1.2. Section 1.3 provides a short discussion on the way this knowledge has been applied to build models of the thermosphere. This leads to the description of the research objective of this thesis in Sect. 1.4.

1.1 Why Study the Thermosphere?

Thermosphere density models are applied in scientific investigations as well as in many types of satellite orbit calculations, including re-entry prediction, manoeuvre planning for ground-track maintenance and precise orbit determination for Earth observation missions. The accuracy of density models therefore does not only influence scientific results, but for many satellite missions it also affects requirements for mission operations, tracking systems and propellant consumption. This section will provide descriptions of these fields of application.

Applications in Solar-Terrestrial Physics

Models and observations of the thermosphere are applied in scientific investigations of the solar-terrestrial environment, a region which is sometimes named geospace. This environment, as defined by Hargreaves [29], includes the upper part of the Earth's atmosphere, the outer part of the geomagnetic field and the solar emissions that affect them.

   The state of the neutral thermosphere, described by density, composition, temperature and wind velocity, depends in a complex way on both solar activity and geomagnetic activity. There are important interactions of the thermosphere with the Earth's ionosphere and magnetosphere [58, 63]. This has led to the development of coupled numerical models of this system. Our understanding of solar-terrestrial physics can be improved by comparing the output of such models with observational data, including data on thermosphere density and wind speed obtained from observations of satellite dynamics.

Applications in the Earth Sciences

For several Earth-observation missions, the accurate determination of the trajectory of the satellite is a requirement for fulfilling the mission's scientific objectives. Thermosphere density models are used to model the drag acceleration acting on the satellite.

   For example, satellite laser ranging SLR satellites, such as Stella, Starlette and Ajisai [71] allow the very accurate determination of station positions, which provide information on Earth-orientation and plate tectonics [54]. Another example is radar and laser altimetry [27], where the altitude of the satellite is required to relate the range measurements over changing water and ice levels to the terrestrial reference frame. Altimetry missions include ERS-1, ERS-2, Envisat, TOPEX-Poseidon, Jason, CryoSat and IceSat. After the spectacular improvement of gravity field modelling and the accompanying reduction of gravity-induced orbit error in the 1990s and

early 2000s, the relative importance of thermospheric density and drag on the orbit accuracy of these satellites has greatly increased [18, 21, 64].

In another application, known as synthetic aperture radar interferometry, or InSAR, the difference in position between two radar image acquisitions should be precisely known in order to study surface deformations [28, 50]. Data from the SAR instruments on ERS and Envisat have been widely used for this purpose.

The motion of satellites has also been analysed to recover information about the Earth's gravity field, by missions such as CHAMP [59] and GRACE [67]. These missions carry precise accelerometers. Any acceleration that is not due to gravity, such as atmospheric drag, will otherwise contaminate the gravity recovery.

Although these Earth observation missions and instruments were not specifically designed for the purpose, data from their precise tracking systems and accelerometers can be very useful for studying variations in the thermosphere. This will be extensively discussed in Chap. 4.

Applications in Space Mission Analysis and Operations

The atmospheric drag force causes all low Earth orbit objects to spiral downward, and eventually re-enter in the most dense layers of the atmosphere. This has profound consequences for many aspects of space mission analysis and operations.

To illustrate this, Fig. 1.1 shows photos of a specific class of space objects, known as Payload Assist Modules for Delta rocket launches (PAM- D). PAM-Ds have been used in a great number of satellite launches, including those of Global Positioning System (GPS) satellites. For these launches, the PAM- D initially remains in a highly elliptical orbit after separation from its satellite payload.

Because of the exponential decrease of atmospheric density with altitude, the drag force on a satellite in an elliptical orbit is strongest at its lowest point, the *perigee*. The drag force causes kinetic energy to be transformed into heat, causing a decrease of the size and ellipticity of the orbit, as measured by its *semi-major axis* and *eccentricity*. Figure 1.2 shows how the changing shape of the orbit of a PAM- D, that was catalogued as object 22659, has been observed during its lifetime. Chapter 3 will provide equations that describe these changes.

Close to the end of its lifetime, when the orbit became nearly circular, the decay of the orbit's semi-major axis quickly accelerated, to up to more than 30 km/day, as the density and drag forces along the trajectory rapidly increased. Occasionally, remnants of a PAM- D have been found on the ground after re-entry. The right-hand side photo in Fig. 1.1 shows such a fragment, belonging to object 22659, which landed in the Saudi Arabian desert.

As we will see in Chap. 4, the rate of change of the size of the orbit, as plotted in Fig. 1.2, is closely related to the density of the atmosphere, encountered by the satellite along its trajectory.
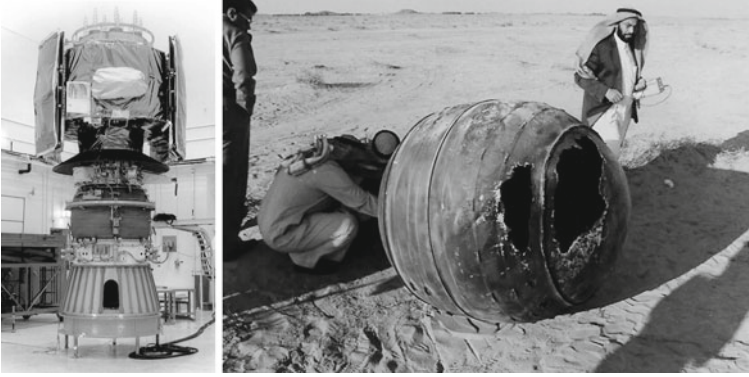
Fig. 1.1 *Left*: A GPS satellite mounted on a delta 2 payload assist module (PAM-D) before launch. Photo: air force space museum [16]. *Right*: Part of the motor casing of a PAM-D (object 22659), which re-entered the atmosphere and landed in the Saudi Arabian desert on January 12, 2001. Photo: KACST
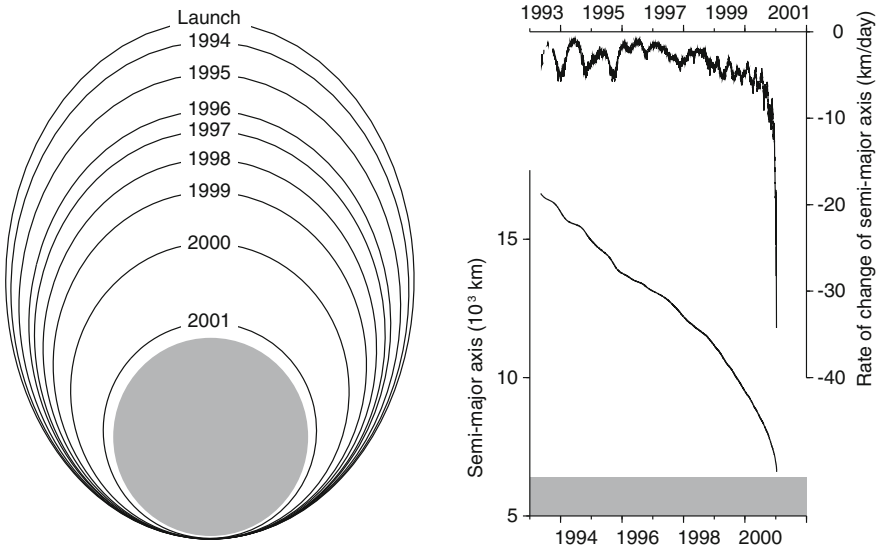
Fig. 1.2 *Left*: The changing shape of the orbit of object 22659, depicted right after its launch on May 13, 1993 and on Jan 1 of each year until its re-entry. *Right:* Time series of the semi-major axis and its rate of change for the object. Orbit data are derived from two-line elements

Orbiting Object Catalogue Maintenance

The United States Space Command has maintained a database of satellite orbit states, known as the Space Object Catalog, since the early days of the space age. In early 2010, the catalog contained over 36,000 distinct objects, of which over 15,000 were

still in orbit. The oldest still orbiting object in the catalogue is Vanguard 1, launched in March 1958. Most objects, such as object 22659 mentioned above, have re-entered.

The catalog is regularly updated using observations from the Space Surveillance Network (SSN), a globally distributed network of interferometer, radar and optical tracking systems. In order to be able to identify and keep track of objects as their orbits change over time, a model of their motion, as influenced by atmospheric drag, is required.

Russia operates its own operational space surveillance system, while France has an experimental system [38]. In recent years, plans for setting up a European operational space surveillance system have materialised as well [39].

Lifetime Analysis

Figure 1.2 shows that the rate of decay of the semi-major axis of object 22659 varied between approximately 2–8 km/day for most of its lifetime. This variation was mainly caused by changes in atmospheric density due to both changes in the orbit geometry with respect to atmospheric features and changes in density due to external causes. An accurate modelling of these effects is important for predicting the orbital lifetime of satellite missions [52, 72]. Of course, such predictions are required during the stage of mission design. Often, the predictions need to be revised after launch, in order to facilitate the planning of a mission extension or the decommissioning of the spacecraft.

Re-Entry Operations

Decommssioned low-Earth orbit objects ultimately re-enter into the denser layers of the Earth atmosphere, where they mostly burn up. Occasionally, objects are involved that are large and massive enough so that one or more remaining parts can reach the ground. Figure 1.1 already showed an example: the PAM- D with catalogue number 22659. Figure 1.3 shows what happened during the final week and hours of the orbital lifetime of this object, until it finally re-entered during the afternoon of January 12, 2001.

Because such re-entry events are relatively rare, and mostly happen over the oceans or uninhabited areas, there is not a particularly large risk for damage or injury [38]. However, sometimes re-entries involve spacecraft with very large masses. Extreme examples are 40,000 kg for Salyut-7 in 1991, 74,000 kg for Skylab-1 in 1979, and 135,000 kg for Mir in 2001. Some re-entry objects also carry hazardous payloads, parts of which can reach the ground, such as the nuclear reactors on Cosmos-954 and Cosmos-1402 in 1978 and 1983, respectively. For such objects, it is desired that re-entry trajectories are predicted, and if possible, controlled. Density fluctuations have a very large influence on the accuracy of trajectory predictions that are made during re-entry campaigns. Chapter 9 of Klinkrad [38] contains extensive information on this subject.
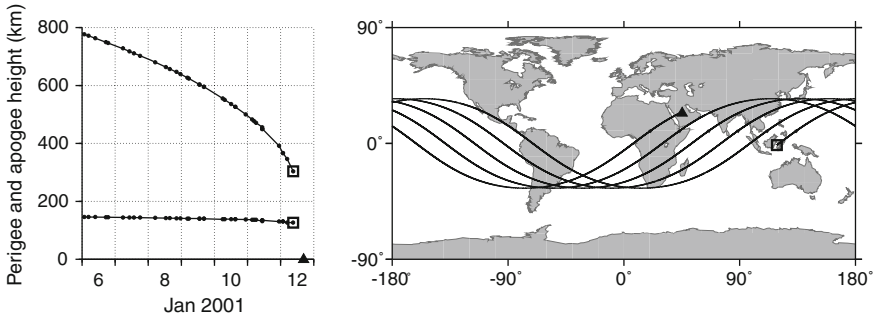
Fig. 1.3 *Left*: Graph of apogee and perigee heights for object 22659, derived from two-line element (TLE) data in the week prior to its re-entry. The small dots correspond to TLE epochs. *Right*: Map of the predicted orbital ground track between the final TLE and the re-entry location. In both figures, the last available TLE is indicated by an open square and the re-entry time and location is indicated by a triangle

Manoeuvre Planning for Orbit Maintenance

Some Earth-observation satellites are required to follow a certain repeating ground track pattern within a specified accuracy, or to remain within a certain altitude range. In routine operations, the orbit decay of these satellites is periodically compensated using thrusters to maintain the orbit altitude. The frequency of these orbit manoeuvres is dependent on the allowed offsets in altitude and ground track. Additional constraints are determined by the required periods of undisturbed payload operation.

As an example, ESA's ERS- 2 and Envisat missions are using near-circular Sun-synchronous orbits at an inclination of $98.52°$ and an altitude of approximately 800 km. The descending equator crossings of the orbit are kept at a constant mean local solar time, at 10:30 for ERS- 2 and 10:00 for Envisat. The ground track pattern for both satellites is identical, with ERS- 2 passing over the same Earth location exactly 30 min after Envisat. This ground track pattern, which is repeated after 501 orbits in exactly 35 days, must be maintained within a cross-track deadband of 1 km on either side to provide Earth-observation products which can be correlated over multiples of that time interval [61, 62].

At high latitudes, the drift in the cross-track direction is driven by luni-solar perturbations, while at equator crossings it is dominated by aerodynamic drag. If the orbital altitude is below the nominal altitude, the satellite orbital period is shorter than nominal. Since the ratio of the orbital period to the Earth's rotation period is then also lower than nominal, the ground track drifts to the east with respect to the nominal ground track. Likewise, when above the reference altitude, the ground track drifts to the west. Due to the cumulative effect of the semi-major axis offset on the drift of the orbit node, the difference between the true and reference geographic longitude at node crossings follows a parabolic pattern, as illustrated in Fig. 1.4.
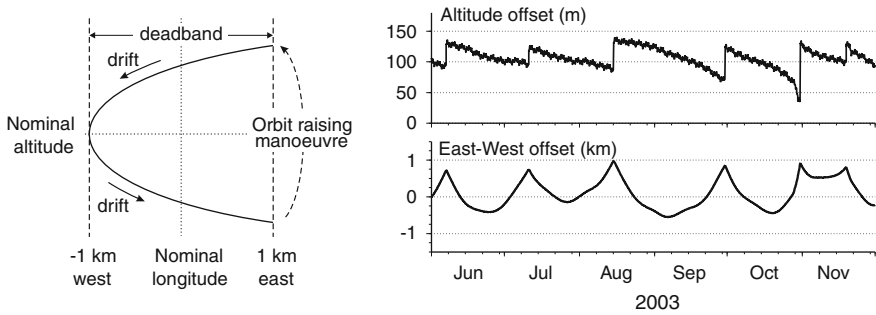
Fig. 1.4 *Left*: Schematic representation of the ground track drift at node crossings, as a function of altitude, for an orbit decaying under the influence of aerodynamic drag. *Right*: Orbit drift at the ascending node of Envisat, with respect to its reference orbit

In an ideal case, the parabola starts at an altitude of $H_{nom} + \Delta H$ at the eastern border of the longitude deadband, touching the western border for H_{nom}, and reaching the eastern border again with $H_{nom} - \Delta H$. At this point an orbit raising manoeuvre is performed to attain the initial $H_{nom} + \Delta H$ that leads to a maximum resident time within the given longitude deadband.

In reality, the time it takes to go through a complete longitude deadband cycle and the corresponding change in altitude depends on the level of thermospheric density. The time between manoeuvres can be as short as two weeks for high drag at solar maximum, one month for mean activity, and more than two months during the solar cycle minimum. The necessary altitude raises at the end of the cycle would be approximately 190, 90, and 50 m for high, medium, and low thermosphere density conditions. When unexpected peaks in solar and geomagnetic activity occur, the orbit decays much faster than expected. The western boundary of the deadband will not be reached in this case, and a new orbit raising manoeuvre is required much sooner than would have been necessary if the decay rate would have been better predicted. On the other hand, at low activities it takes a long time to drift through the parabola in Fig. 1.4, and there is a high risk of overshooting the Western boundary. Hence, safety margins are used which may considerably reduce the maximum possible time between maintenance manoeuvres and increase the required fuel budget. The right-hand side of Fig. 1.4 illustrates the sub-optimal orbit raising procedure for Envisat over several months in 2003, during which there were large unpredicted density fluctuations.

Manoeuvre Planning for Collision Avoidance

A very large number of orbiting objects has been classified as space debris in low Earth orbit. If such an object is large enough, it might cause catastrophic consequences in a collision with an operational satellite or manned space mission. In order

to prevent this, all objects larger than about 10 cm are tracked by the US Space Surveillance Network, and orbit information is distributed in the form of Two Line Element (TLE) sets (see also Sect. 1.2).

To avoid the possibly catastrophic consequences of a high-velocity collision in orbit, a special kind of manoeuvre planning is required. Knowledge of atmospheric density and its variations is essential for this purpose, since aerodynamic acceler-ations are the largest cause of uncertainty in trajectory predictions. Chapter 8 of Klinkrad [38] contains a detailed description of operational collision avoidance pro-cedures implemented at the European Space Operations Centre (ESOC), in Darm-stadt, Germany.

1.2 Satellite Observations of the Thermosphere

The derivation and use of satellite observations of thermosphere density is the main topic of this thesis. Chapter 4 will provide details of the three distinct data sources that will be used in this work: Two-Line Elements, precise satellite tracking and satellite accelerometers. This Section will provide context by providing an overview of these and other data types.

Satellite Drag from Orbit Tracking

The first measurements of thermospheric density were derived from optical and radio-tracking of satellite orbits, which were processed to provide measurements of orbital decay. From these density variations, changes in temperature and composition were derived based on theory. KingHele [37] gives a fascinating personal account of this research, early on in the space age. Availability of these observations e.g., Jacchia and Slowey [35] led to the development of the first thermosphere models by Harris and Priester [30] and Jacchia [33].

Orbit determination results based on radar tracking by the US Space Surveillance Network (SSN) have been made available in the form of Two-Line Element (TLE) sets with a latency of one day or less. TLEs are a way of describing and distributing satellite orbits using a minimum level of computer storage space. Historical TLE data, dating back to the early 1960s, is now easily accessible on the Internet. Although TLEs are produced for operational purposes, and not with scientific research of the thermosphere in mind, they nevertheless provide an easily accessible form of data on satellite drag. A major benefit of TLEs is the very large number of objects and orbits for which they are available.

King-Hele [36] discusses the benefits and pitfalls of using TLEs in combination with his analytical theory for orbits in the atmosphere. Picone et al. [57] provides an algorithm, suitable for automated analysis on modern computers, that is more

general and more powerful. It has been extensively applied in thermosphere research in recent years, and will be described in detail in Sect. 4.1.

A second type of drag data from satellite dynamics is the use of very precise tracking systems of geodetic satellites, already mentioned in the previous Section. The three major precise tracking systems in use today are satellite-to-satellite tracking using the GPS system [23], two-way Satellite Laser Ranging (SLR) from a network of laser stations to dedicated reflectors on satellites [54] and one-way Doppler radio tracking using the French network of DORIS beacons to satellites equipped with a DORIS receiver [68]. These tracking techniques facilitate an orbit accuracy of the order of a few centimetres, which is crucial for applications such as altimetry, gravity field determination, synthetic aperture radar interferometry and the determination of Earth-orientation parameters. The tracking data is used in orbit determination in order to estimate parameters, which have been selected to counteract uncertainties in force models, such as those caused by inaccuracies in the modelled thermosphere density. Section 4.3 outlines the procedure for the use of this data. Since satellite drag has mainly been considered as a nuisance parameter in the primary application of these instruments (e.g., [18]), the use of precise tracking for thermosphere investigations has been limited to only a few studies, such as those by Bruinsma [13], van den IJssel and Visser [32], and Willis et al. [71].

Accelerometers

From the 1960s to 1980s, accelerometer instruments for use in thermosphere modelling studies were flown on several US, Italian and French satellites. The MESA (Miniature Electrostatic Single-axis Accelerometer) experiment was flown on-board of eight satellites since 1968, including the Atmospheric Explorer C, D and E spacecraft [47]. A follow-on to MESA, the SETA (Satellite Electrostatic Triaxial Accelerometer) experiment [49] flew on several US Air Force missions during the first half of the 1980s. The Italian accelerometer-carrying San Marco satellites [9] were of spherical shape. Five San Marco satellites were launched into orbit between 1964 and 1988. Finally, the French CACTUS instrument flew on the CASTOR satellite, at approximately 30° inclination, in an elliptical orbit, allowing density measurement between 250 and 600 km altitude between May 1975 and February 1979 [5, 24]. CACTUS is a French acronym meaning ultrasensitive, three-axis, capacitive accelerometric transducer.

These instruments were generally operated only over relatively short time spans, and on elliptical orbits. The data from these missions were used over the years to study, among others, the accuracy of empirical density models [1, 24, 45, 47], thermospheric winds [46], the asymmetry of large-scale structures in the thermosphere [5], the midnight density maximum [2], and analyses of waves in the density data [3, 25].

The CHAMP and GRACE gravity field recovery missions (see Fig. 1.5) have been providing the scientific community with nearly continuous accelerometer observation data since August 2000 and April 2002, respectively. The CHAMP mission

Fig. 1.5 The CHAMP (*left*) and GRACE satellites (*right*) in the cleanroom. Photos courtesy of astrium GmbH

has ended with the satellite's re-entry in September 2010, ending a decade of near-continuous observation. The GRACE mission operations will likely be extended through 2015, at which point a follow-on mission, also likely to carry precision accelerometers, should take over.

Additional coverage and continuity of accelerometer data can be provided by ESA's future Swarm mission, a constellation of three satellites designed for studies of the Earth's magnetic field. Also of great interest is the ESA GOCE mission, launched in March 2009, and recently extended until the end of 2012. GOCE flies under drag-free control [14] at a fixed, very low altitude, just above 250 km. It is in a nearly sun-synchronous orbit, which causes the local time coverage to be rather limited. In addition, the drag-free control mode, required for its primary mission in gravity field recovery, is a complicating factor for density determination. Data processing for this purpose is likely to start in 2011, when actuation data of the drag-free system is scheduled to become available.

These satellites are in long-duration near-circular polar orbits, providing excellent coverage over latitude and solar activity. This is a huge improvement over the intermittent data of the earlier missions, even though those missions were dedicated to thermospheric studies. The missions also carry precise tracking systems in the form of GPS receivers and SLR (Satellite Laser Ranging) retroreflectors, which can be used for calibration of the accelerometer data. Especially the CHAMP accelerometer data have been used successfully to study atmospheric density [11, 12, 40, 41, 43, 44, 66] and wind patterns [66, 43, 26].

The fact that the CHAMP and GRACE missions have been largely overlapping in time (see Fig. 1.6) offers the opportunity to study the synergy of their data. The twin GRACE satellites and the single CHAMP satellite have formed a constellation of which the geometrical configuration constantly evolved over time. When the orbital planes were perpendicular, the instantaneous local solar time coverage was doubled compared to having just one mission. At times when the orbits were coplanar, it is more feasible to study relatively short spatial scale features, such as waves. Since
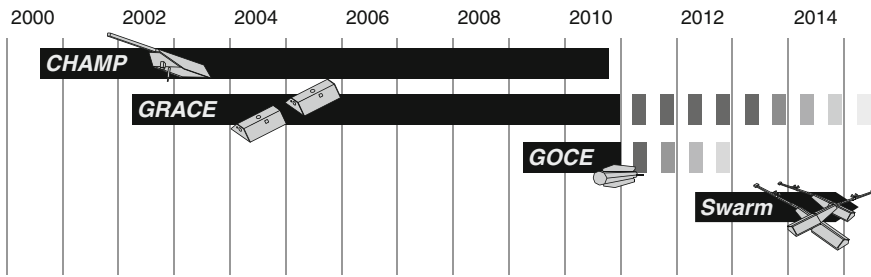
Fig. 1.6 Timeline for current and future low earth orbiting satellite missions carrying high-precision accelerometers

these satellites are polar orbiting, they offer unprecedented opportunities to study the complex dynamics of the thermosphere at high latitudes.

Mass Spectrometers

Missions such as the Atmospheric Explorer, Dynamics Explorer and San Marco series also carried mass spectrometers. These instruments provided the capability of measuring the number density of individual atomic and molecular species directly, although the calibration of these densities is not a trivial matter. Just like the accelerometers, the mass spectrometers delivered greatly improved temporal and spatial resolution, compared to satellite drag from orbit tracking at the time. Although mass spectrometer data was extensively used in building density models in the early 1980s, no new datasets of this type have become available in recent years. This data will remain outside the scope of this thesis.

Far Ultraviolet Imaging of Airglow

The TIMED (Thermosphere, Ionosphere, Mesosphere Energetics and Dynamics) mission, launched in December 2001, carries the GUVI (Global Ultraviolet Imager) instrument [53], as part of a suite of instruments to study the charged and neutral particles in the altitude range of 60–180 km, and the influence of solar radiation on this region. The GUVI instrument images the dayside airglow in the upper atmosphere in order to derive temperature and composition profiles. The SSULI (Special Sensor Ultraviolet Limb Imager) instrument [51], flown on DMSP (Defense Meteorological Satellite Program) satellites, is based on similar principles. The datasets from these missions will also remain outside of the scope of this thesis.

1.3 Thermosphere Models

The availability of observation data on density and wind in the thermosphere has led to the development of various models. A major distinction can be made between two types: physical models and empirical models, each with its own strengths and application domains.

Physical Models

Physical or theoretical models are based on the fundamental physical equations that govern the system. These models include so-called Thermosphere Global Circulation Models (TGCMs) and localised, high resolution models (e.g., [17, 55, 42]). Recent versions are usually part of larger coupled thermosphere-ionosphere-magnetosphere models. These models are able to represent the detailed response of thermospheric density, temperature, composition and winds to the diverse energy inputs that affect the thermosphere (see Sect. 2.1.3). The accuracy of the outputs are limited by the accuracy of the description of these energy inputs, the scope of the model and computational constraints. These factors need to be understood when interpreting physical model output. Compared to empirical models, physical models are computationally intensive. For all of these reasons, they require significantly more expert knowledge to run. This makes these models more suitable for scientific investigations of the behaviour of the thermosphere-ionosphere system than for routine applications in orbit determination.

Empirical Models

Empirical models, which will be described in more detail in Chap. 2, consist of a set of relatively simple functions representing only the major variations in number densities of the major atmospheric constituents, in the resulting total density, and usually also in temperature. They are therefore generally not able to represent features at short temporal or spatial scales. The models' functions contain parameters which have been fitted against a database of past observations, as a function of time, location and driving space weather proxies and indices. The use of empirical models can be seen as interpolating or extrapolating the database of past observations as a function of these input parameters. The predictive capabilities of all thermosphere models, whether physical or empirical in nature, are completely dependent on the prediction accuracy of the solar activity proxies and geomagnetic activity indices used.

   The most important examples of empirical models are those based on the work by Jacchia [7, 8, 34], the DTM series [4, 6, 10] and the MSIS series [31, 56]. All these

models were or are still widely used in orbit determination applications. The space science community has predominantly made use of the MSIS series of models.

Recent Challenges and Strategies

The Jacchia, MSIS and DTM models rely on the correlation between variations in density, temperature and composition with the $F_{10.7}$ and *ap* or *Kp* space weather proxies, in order to represent all stochastic atmospheric behaviour. Mainly because this correlation is limited, the model accuracy that can be achieved is restricted to around 15–30%, depending on altitude and activity levels [48, 60].

It seems unlikely that the complexities of upper atmospheric variations can be completely captured by a small set of solar and geomagnetic activity proxies. Likewise, it is clear that it will be very challenging to completely represent the state of the atmosphere at a given time with a thermosphere model that extrapolates a selection of past data. Because of these limitations in model formulation, investigations into the use of more historical data or more representative single proxies in the model generation (e.g., [10, 13, 56, 69]) have not resulted in a significant breakthrough in orbit determination and prediction applications.

Two strategies in recent years have been more successful in lowering empirical density model error. These are the simultaneous use of multiple solar activity proxies, employed by the Jacchia-Bowman models [70], discussed in the next Chapter, and assimilation of contemporaneous data in density model calibration, which will be described in detail in Chap. 6.

1.4 Research Objectives, Motivation and Scope

Section 1.1 described the need for accurate data and models of the thermosphere in both the space science and satellite orbit modelling communities. The work described in this thesis originated from three research projects funded by the European Space Agency (ESA), aimed at addressing these needs. The results from these projects were reorganised, and in some cases summarised or expanded, for their presentation in this thesis, with the aim to make the work stand on its own. The objectives of each of these projects by themselves will first be presented below, before formulating the overall research objective for the thesis work.

Objectives and Motivations of the ESA Studies

The first two ESA projects [19, 20] were initiated by the space debris office at the European Space Operations Centre (ESOC) in Darmstadt, Germany. These projects

were inspired by publications about the US Air Force Space Command's HASDM project (e.g., 15, [65]). The US researchers demonstrated the successful application of the concept of empirical density model calibration, in which density observations derived from satellite tracking data are used to compute corrections to empirical density models, at regular time intervals. This resulted in a very significant improvement of model accuracy. However, the calibrated HASDM model source code and coefficients themselves, as well as their underlying Space Surveillance Network tracking data, are not readily accessible to the international scientific and space operations community. The two ESA/ESOC projects were therefore aimed at investigating the possibility of running density model calibration using density data from publicly available sources. A requirement was that the calibration should be run in near real-time, in order to facilitate orbit predictions made for re-entry analyses and collision avoidance at ESOC. After a feasibility study, it was found that only densities derived from Two-Line Element orbit data offered the required spatial distribution and near real-time availability. In a follow-on project, the model calibration software was successfully implemented and delivered.

The third ESA project [22], funded by its General Studies Programme, and managed from its technology centre ESTEC in Noordwijk, The Netherlands, was motivated by the need to prepare for the data processing of the accelerometer instruments on board of the future Swarm satellite constellation. Swarm will enable the most detailed investigation yet of the Earth's magnetic field. As a secondary mission objective, the geomagnetic forcing of the thermosphere will be investigated using density and winds derived from the accelerometer measurements. During the course of this project, the accelerometer data processing of the CHAMP and GRACE missions, which shows many similarities, was critically reviewed. Significant shortcomings in the geometry and aerodynamic modelling of these satellites were identified, and where possible improved. In addition, the feasibility of using the resulting density data in the empirical model calibration, instead of or in addition to the TLE-data, was investigated.

Overall Research Objectives

The overall research objective for this thesis is formed from the combination and expansion of the objectives of these ESA studies. Its objective is to use data on the aerodynamic accelerations on satellites from publicly and freely accessible sources, and process it to obtain observations of thermospheric parameters that can then be applied in further investigations, including density calibration experiments.

These publicly available satellite data sets were usually not originally intended for thermosphere studies, and are therefore not distributed in an optimal manner for this purpose. Additional goals are therefore to shed light on the assumptions that are made in the data processing, to gain more insight into the accuracy and limitations of the resulting data, and to set up the data processing so that, given the limitations on input data and models, the resulting densities are as accurate as

practically possible. In addition, these investigations have led to recommendations for future satellite missions and data products that can be applied for investigations of the thermosphere.

Audience and Scope

The thesis is intended to be of value to those interested in the processing and application of satellite-derived data on the thermosphere. This work therefore brings together knowledge from the field of solar-terrestrial physics and from the field of satellite dynamics. Chapters 2 and 3 are extended introductions, discussing the governing equations and techniques from both these domains, which are applied in later Chapters.

Of course, during the course of the preparation and writing of the thesis, a decision had to be made to limit its scope. The description of the processing algorithms therefore start assuming preprocessed satellite data is available. For information on the design and functioning of the accelerometer instruments and tracking equipment, and the conversion of their raw output to measurements in the proper physical units, the reader is referred to various references.

A similar limitation was necessary at the other side of the processing chain, the output of density and wind data. The interpretation of this data, in terms of the governing solar-terrestrial physics, is rather minimal. Interpretations are only made in order to help validate the measurements and models, and to aid in assessing their limitations and accuracy. In addition, more attention is paid to the density measurements than to the wind data. However, many references are included throughout the text, to journal articles that describe the interpretation of both types of output data in terms of physical phenomena and processes. Some of the most recent of these references have made use of the newly developed accelerometer-processing algorithms described in this thesis.

References

1. Arduini C, Broglio L, Ponzi U, Laneve G, Marco S (1996) V drag balance neutral density compared to the models. Adv Space Res 18(9–10):351–360. doi:10.1016/0273-1177(96)-000695
2. Arduini C, Laneve G, Herrero FA (1997) Local time and altitude variation of equatorial thermosphere midnight density maximum (MDM): San Marco drag balance measurements. Geophys Res Lett 24(4):377–380. doi:10.1029/97GL00189
3. Arduini C, Laneve G, Ponzi U (1999) New insight on the internal waves in the equatorial thermosphere by the "s. marco 5" spacecraft data. Adv Space Res 24(11):1463–1472. doi:10.1016/S0273-1177(99)00707-3
4. Barlier F, Berger C, Falin JL, Kockarts G, Thuillier G (1978) A thermospheric model based on satellite drag data. Ann Geophys 34(1):9–24
5. Berger C, Barlier F (1981) Asymmetrical structure in the thermosphere during magnetic storms as deduced from the CACTUS accelerometer data. Adv Space Res 1(12):231–235

6. Berger C, Biancale R, Ill M, Barlier F (1998) Improvement of the empirical thermospheric model DTM: DTM-94—a comparative review of various temporal variations and prospects in space geodesy applications. J Geod 72(3):161–178

7. Bowman, Bruce R, Tobiska WK, Marcos FA (2006) A new empirical thermospheric density model JB2006 using new solar indices, in AIAA, AIAA, pp 2006–6166

8. Bowman, Bruce R, Tobiska WK, Marcos FA, Huang CY, Lin CS, Burke WJ (2008) A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices. In: AIAA/AAS Astrodynamics specialist conference and exhibit, August 2008, Honolulu, Hawaii, pp 18–21, AIAA 2008–6438

9. Broglio L, Ponzi U, Arduini C (1992) The San Marco 5 mission. Adv Space Res 13(1):165–184

10. Bruinsma S, Thuillier G, Barlier F (2003) The DTM-2000 empirical thermosphere model with new data assimilation and constraints at lower boundary: accuracy and properties. J Atmos Sol Terr Phys 65:1053–1070

11. Bruinsma S, Tamagnan D, Biancale R (2004) Atmospheric densities derived from CHAMP/STAR accelerometer observations. Planet Space Sci 52(4):297–312. doi:10.1016/j.pss.2003.11.004

12. Bruinsma S, Jeffrey M Forbes, R Steven Nerem, Xiaoli Zhang (2006) Thermosphere density response to the 20–21 November 2003 solar and geomagnetic storm from CHAMP and GRACE accelerometer data. J Geophys Res 111(A06303). doi:10.1029/2005JA011284

13. Bruinsma SL, Exertier P, Biancale R (1999) An assessment of new satellite total density data for improving upper atmosphere models. Planet Space Sci 47:1465–1473

14. Canuto, E, Bona B, Calafiore G, Indri M (2002) Drag free control for the european satellite GOCE, part I: modelling. In: Proceedings of the 41st IEEE conference on decision and control, Las Vegas, NV, December 2002, number WeA03–6

15. Casali SJ, Barker WN (2002) Dynamic calibration atmosphere (DCA) for the high accuracy satellite drag model (HASDM). In: AIAA/AAS astrodynamics specialist conference and exhibit, Monterey, California, AIAA, 5–8 Aug 2002, pp 2002–4888

16. Cleary MC (1994) The cape: military space operations. 45th SpaceWing History Office, 1971–1992

17. Crowley G (1991) Dynamics of the earth's thermosphere: a review. Rev Geophys 29(29 suppl):1143–1165

18. Doornbos E, Scharroo R, Klinkrad H, Zandbergen R, Fritsche B (2002) Improved modelling of surface forces in the orbit determination of ERS and envisat. Can J Remote Sens 28(4):535–543

19. Doornbos E (2004) Calibrated, high accuracy satellite drag model - ESOC contract 16643/02/D/HK(SC) final report. Faculty of Aerospace Engineering Delft University of Technology, Delft The Netherlands

20. Doornbos E (2006) NRTDM final report - near real-time density model (NRTDM) - ESOC contract 18576/04/D/HK(SC). Delft Institute for Earth-Oriented Space Research, The Netherlands

21. Doornbos E, Heiner K, Remko S, Pieter V (2007) Thermosphere density model calibration in the orbit determination and prediction of ERS-2 and Envisat. In: Lacoste H (ed) Envisat symposium montreux, Switzerland, 23–27 April 2007, ESA SP-636

22. Doornbos E, Matthias Förster Bent Fritsche, Tom van Helleputte, Jose van den IJssel, Georg Koppenwallner, Hermann Lühr, David Rees, Pieter Visser (2009) ESTEC contract 21022/07/NL/HE air density models derived from multi-satellite drag observations – final report. DEOS / TU Delft scientific report 01/2009, TU Delft

23. Dow JM, Neilan RE, Gendt G (2005) The international GPS service (IGS): celebrating the 10th anniversary and looking to the next decade. Adv Space Res 36(3):320–326

24. Falin JL, Kockarts G, Barlier F (1981) Densities from the CACTUS accelerometer as an external test of the validity of the thermospheric models. Adv Space Res 1(12):221–225

25. Forbes JM, Marcos FA, Kamalabadi F (1995) Wave structures in lower thermosphere density from satellite electrostatic triaxial accelerometer measurements. J Geophys Res 100(A8):14693–14701

26. Förster M, Haaland SE, Doornbos E (2011) Thermospheric vorticity at high geomagnetic latitudes from CHAMP data and its IMF dependence. Ann Geophys 29(1):181–186. doi:10.5194/angeo-29-181-2011
27. Fu Lee-Lueng, Cazenave A (eds) (2001) Satellite altimetry and earth sciences: A handbook of techniques and applications, Academic Press
28. Hanssen RF (2001) Radar interferometry, data interpretation and error analysis vol 2 of remote sensing and digital image processing. Springer, San Diego
29. Hargreaves JK (1992) The solar-terrestrial environment: an introduction to geospace - the science of the terrestrial upper atmosphere, ionosphere and magnetosphere vol 7 of cambridge atmospheric and space science series. Cambridge University Press, Cambridge
30. Harris I, Wolfgang P (1962) Time-dependent structure of the upper atmosphere. J Atmos Sci 19(4):286–301
31. Hedin AE (1987) MSIS-86 thermospheric model. J Geophys Res 92(A5):4649–4662
32. van den IJssel J, Visser P (2005) Determination of non-gravitational accelerations from GPS satellite-to-satellite tracking of CHAMP. Adv Space Res 36(3):418–423
33. Jacchia LG (1964) Static diffusion models of the upper atmosphere with empirical temperature profiles. Smithsonian astrophysical observatory special report 170
34. Jacchia LG (1972) Atmospheric models in the region from 110 to 2000 km. In: CIRA 1972: COSPAR international reference atmosphere 1972. Akademie-Verlag, Berlin, pp 227–338
35. Jacchia LG, Slowey J (1962) Accurate drag determinations for eight artificial satellites; atmospheric densities and temperatures. Smithsonian Astrophysical Observatory Special Report 100
36. King-Hele D (1987) Satellite orbits in an atmosphere, theory and applications. Blackie
37. King-Hele D (1992) A tapestry of orbits. Cambridge University Press, Cambridge
38. Klinkrad H (2006) Space debris, models and risk analysis, Springer, Berlin.
39. Klinkrad, Heiner, Richard Tremayne-Smith, Fernand Alby, Detlef Alwes (2008) Europe's eyes on the skies. the proposal for a European space surveillance system, ESA Bulletin, (133), Feb 2008
40. Liu H, Lühr H (2005) Strong disturbances of the upper thermospheric density due to magnetic storms: CHAMP observations. J Geophy Res 110(A09829). doi:10.1029/2004JA010908
41. Liu H, Lühr H, Henize V, Köhler W (2005) Global distribution of the thermospheric total mass density derived from CHAMP. J Geophys Res 110(A04301). doi:10.1029/2004JA010741
42. Liu H-L, Foster BT, Hagan ME, Mcinerney JM, Maute A, Qian L, Richmond AD, Roble RG, Solomon SC, Garcia RR, Kinnison D, Marsh DR, Smith AK, Richter J, Sassi F, Oberheide J (2010) Thermosphere extension of the Whole Atmosphere Community Climate Model. J Geophys Res 115(A12302). doi:10.1029/2010JA015586
43. Liu, H, Lühr H, Watanabe S, Köhler W, Henize V, Visser P (2006) Zonal winds in the equatorial upper thermosphere: decomposing the solar flux, geomagnetic activity, and seasonal dependencies. J Geophys Res 111(A07307). doi:10.1029/2005JA011415
44. Liu, H, Lühr H, Watanabe S (2007) Climatology of the equatorial thermospheric mass density anomaly. J Geophys Res 112(A05305). doi:10.1029/2006JA012199
45. Marcos FA (1990) Accuracy of atmospheric drag models at low satellite altitudes. Adv Space Res 10(3):417–422
46. Marcos FA, Forbes JM (1985) Thermospheric winds from the satellite electrostatic triaxial accelerometer system. J Geophys Res 90:6543–6552
47. Marcos FA, Garrett HB, Champion KSW, Forbes JM (1977) Density variations in the lower thermosphere from analysis of the AE-C accelerometer measurements. Planet Space Sci 25(5): 499–507
48. Marcos FA, Bass JN, Baker CR, Borer WS (1994) Neutral density models for aerospace applications. In: 32nd aerospace sciences meeting and exhibit, 10–13 Jan 1994 / Reno, NV
49. Marcos, FA, Swift ER (1982) Application of the satellite triaxial accelerometer experiment to atmospheric density and wind studies. Air Force Geophysics Laboratory, Air Force Systems Command, USAF

50. Massonnet D, Kurt LF (1998) Radar interferometry and its applications to changes in the earth's surface. Rev Geophys 36(4):441–500. doi:10.1029/97RG03139

51. McCoy RP, Kenneth FD, Gilbert GF, Stefan ET, Robert RM, Paul AR (1994) Special sensor ultraviolet limb imager: an ionospheric and neutral density profiler for the defense meteorological satellite program satellites. Opt Eng 33(2):423–429. doi:10.1117/12.155904

52. Owens JK, William VV, Niehuss KO, Minow J (2000) Space weather, earth's neutral upper atmosphere (thermosphere), and spacecraft orbital lifetime/dynamics. IEEE Trans Plasma Sci 28(6):1920–1930

53. Paxton LJ, Christensen AB , Humm DC, Ogorzalek BS, Pardoe CT, Morrison D, Weiss MB, Crain W, Lew PH, Mabry DJ, Goldsten JO, Gary SA, Persons DF, Harold MJ, Alvarez EB, Ercol CJ, Strickland DJ, Ching-I M (1999) Global ultraviolet imager (GUVI): measuring composition and energy inputs for the NASA thermosphere ionosphere mesosphere energetics and dynamics (TIMED) mission. Opt Spectrosc Tech Instrum Atmos Space Res III 3756(1):265–276. doi:10.1117/12.366380

54. Pearlman MR, Degnan JJ, Bosworth JM (2002) The international laser ranging service. Adv Space Res 30(2):135–143. doi:10.1016/S0273-1177(02)00277-6

55. Peymirat C, Richmond AD, Emery BA, Roble RG (1998) A magnetosphere thermosphere ionosphere electrodynamics general circulation model. J Geophys Res 103(A8): 17,467–17,477. doi:10.1029/98JA01235

56. Picone JM, AE Hedin, DP Drob, AC Aikin (2002) NRLMSISE-00 empirical model of the atmosphere: Statistical comparisons and scientific issues. J Geophys Res 107(A12). doi:10.1029/2002JA009430

57. Picone JM, JT Emmert, J Lean (2005) Thermospheric densities derived from spacecraft orbits - I. Accurate processing of two-line element sets. J Geophys Res 110(A03301). doi:10.1029/2004JA010585

58. Prölss GW (2004) Physics of the earth's space environment. Springer

59. Reigber Ch, Schwintzer P, Lühr H (1999) The CHAMP geopotential mission, in Bollettino di Geofisica Teoretica ed Applicata. In: Marson I, Sünkel H (eds) Proceedings of the 2nd joint meeting of the international gravity and the international geoid commission, Trieste 7–12 Sept–Dec 1998, ISSN 0006–6729, vol 40(3–4), pp 285–289

60. Rhoden EA, Forbes JM, Marcos FA (2000) The influence of geomagnetic and solar variabilities on lower thermosphere density. J Atmos Sol Terr Phys 62:999–1013

61. Rosengren M (2000) Orbit control of ERS-1, ERS-2 and ENVISAT to support SAR interferometry. In: ERS and Envisat symposium, Gothenburg, 16–20 Oct 2000, ESA SP–461

62. Rudolph A, Kuijper D, Ventimiglia L, Matatoros MAG, Bargellini P (2005) Envisat orbit control - philosophy, experience and challenge. In: Lacoste H , Ouwehand L (eds) Proceedings of the 2004 envisat and ERS symposium, Salzburg, Austria, ESA SP-572, 6–10 Sep 2004

63. Schunk RW, Nagy AF (2009) Ionospheres: physics, plasma physics and chemistry. 2nd edition

64. Sibthorpe AJ (2006) Precision non-conservative force modelling for low Earth orbiting spacecraft. PhD dissertation, Department of Geomatic Engineering, University College London

65. Storz MF, Bowman BR, Branson JI (2002) High accuracy satellite drag model (HASDM). In: AIAA/AAS astrodynamics specialist conference and exhibit, Monterey, California, AIAA, 5–8 Aug 2002, pp 2002–4886

66. Sutton EK, Forbes JM, Nerem RS (2005) Global thermospheric neutral density and wind response to the severe 2003 geomagnetic storms from CHAMP accelerometer data. J Geophys Res 110(A09S40). doi:10.1029/2004JA010985

67. Tapley BD, Bettadpur S, Watkins M, Reigber C (2004) The gravity recovery and climate experiment, mission overview and early results. Geophys Res Lett 31(L09607)

68. Tavernier G, Fagard H, Feissel-Vernier M, Le Bail K, Lemoine F, Noll C, Noomen R, Ries JC, Soudarin L, Valette JJ, Willis P (2006) The international DORIS service: genesis and early achievements. J Geod 80(8–1):403–417. doi:10.1007/s00190-006-0082-4

69. Thuillier G, Bruinsma S (2001) The Mg II index for upper atmosphere modelling. Ann Geophys 19(2):219–228

70. Tobiska WK, Bouwer SD, Bowman BR (2006) The development of new solar indices for use in thermospheric density modeling. In: AIAA, AIAA, pp 2006–6165
71. Willis P, Deleflie F, Barlier F, Bar-Sever YE, Romans LJ (2005) Effects of themosphere total density perturbations on LEO orbits during severe geomagnetic conditions (Oct-Nov 2003) using DORIS and SLR data. Adv Space Res 36:522–533. doi:10.1016/j.asr.2005.03.029
72. Woodburn J, Lynch S (2005) A numerical study of orbit lifetime. In: AIAA/AAS astrodynamics specialist conference and exhibit, Lake Tahoe, AAS 05-297

Chapter 2
Empirical Modelling of the Thermosphere

This chapter will describe the history, context, application and limitations of empirical thermosphere models. Section 2.1 will give an introduction to the atmospheric structure and variability. The observations of solar and geomagnetic activity that serve as drivers of thermosphere variability are presented in Sect. 2.2. Section 2.3 will provide a historical overview of model developments. The inner workings of the models, and the way they have been created, are described in Sect. 2.4. Section 2.5 provides an introduction to the concept of density model calibration, of which an implementation is described at the end of the thesis, in Chap. 6. This chapter ends with an overview of tools that are applied in the evaluation of density models and data sets in Sect. 2.6.

2.1 Atmospheric Structure and Variability

The Earth's atmosphere is often portrayed as a fragile, finite and thin layer of gases surrounding our planet. Since the number of atmospheric particles decreases exponentially with altitude, there is no definite boundary between the atmosphere and outer space. Even though the density at satellite altitudes (200 km and up) is at least a billion times lower than at sea level, the velocities of orbiting objects are so high that there is still a drag force that can often be measured by analysing satellite dynamics. This aerodynamic force can currently be derived from operational space surveillance satellite tracking orbit data for altitudes up to around 500 km and by using more specialised methods or equipment up to about 1500 km. Tracking measurements on the very first artificial satellites [50, 57] resulted in the identification of several important variations in density, which will be described in this Section.

Figure 2.1 shows how the main atmospheric layers are defined according to changes in the slope of the temperature profile. The maximum temperature is asymptotically reached in the topmost layer, the thermosphere. The level of the temperature in this region is highly dependent on the Sun, as will be explained in more detail
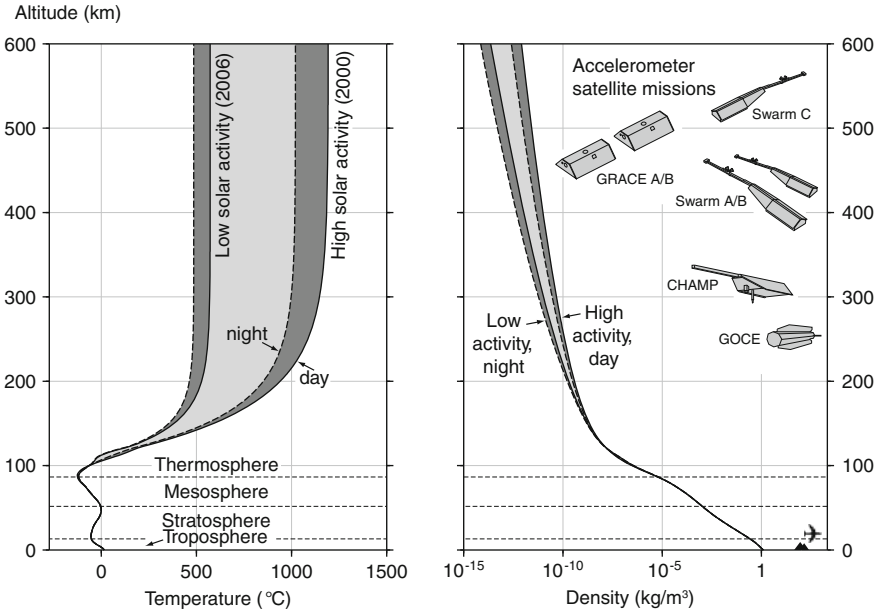
Altitude (km)

Fig. 2.1 Altitude profiles of atmospheric temperature (*left*) and density (*right*), according to the NRLMSISE-00 model, evaluated for 18:00 on July 15 2000 and 2006, over Delft, The Netherlands

below. The variations in temperature lead to variations in density, which can span several orders of magnitude.

2.1.1 Overview of Observed Density Variations

Figure 2.2 shows modelled density variations over an 11-year period, with two zoomed in views for 2000 and 2006. The NRLMSISE- 00 [76] model was evaluated at a single location: 400 km over Delft, The Netherlands. The major variations in density visible in this plot will be described below. A more detailed description of the drivers behind these variations will be provided in further sections of this chapter.

Diurnal and Seasonal Variations

First of all, uneven heating by the Sun of the spherical Earth causes the day-to-night variation in density, represented by the grey band in Fig. 2.2. This variation is also subject to seasonal changes, due to the movement of the sub-solar point as a result of the seasonally changing Sun-Earth geometry. In the thermosphere, the variation
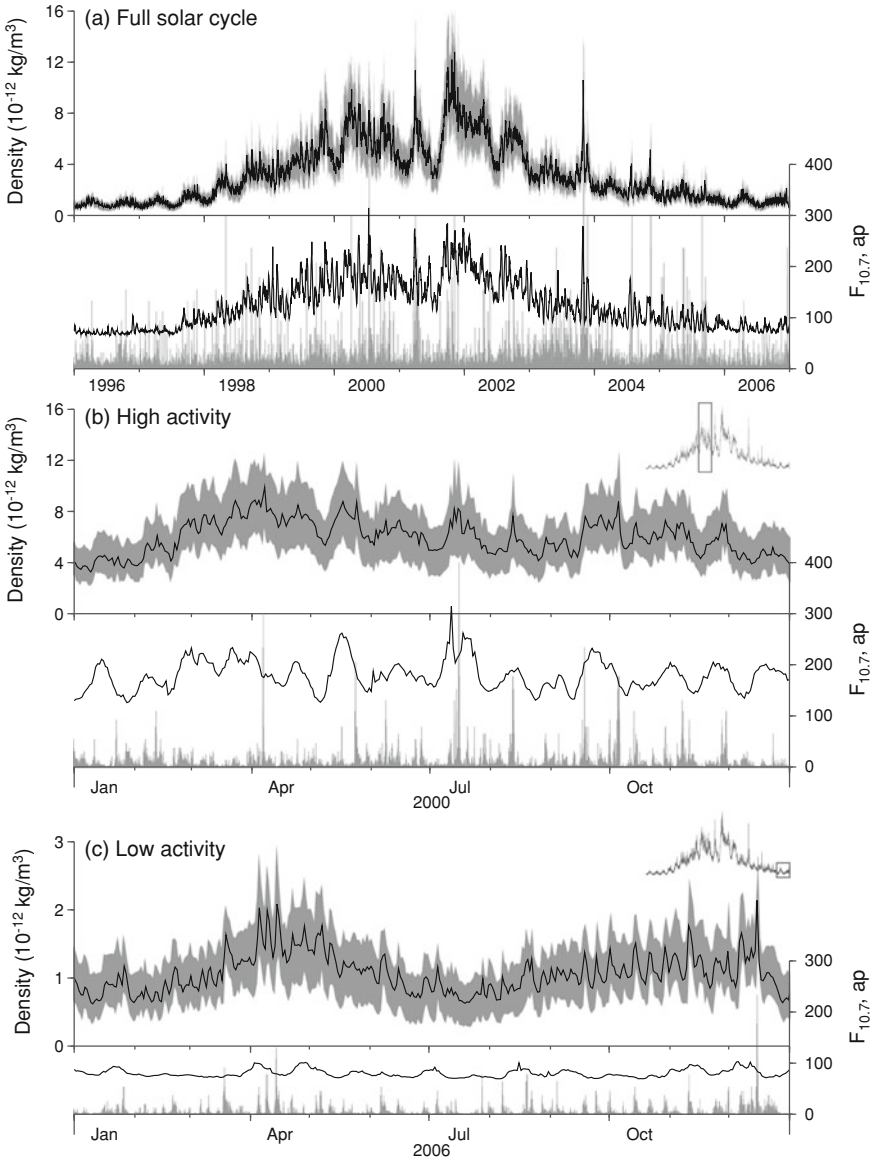
Fig. 2.2 Time series of modelled daily minimum and maximum (*grey area*) and average (*black line*) densities at 400 km over Delft. Time series of the solar and geomagnetic activity indices $F_{10.7}$ (*black line*) and ap (*grey spikes*), which drive the model, are plotted below each density time series

manifests itself in the form of a daytime density bulge, with a maximum following the sub-solar point with a time-lag of about two hours. In the horizontal plane, the density smoothly decreases to a minimum on the night side.

Because heating at the bottom of the thermosphere generally causes the entire thermosphere above this layer to expand, the day-night difference increases with altitude. It can be around a factor of 5 at 500 km and higher.

Solar Activity Variation

The second major variation in density is also related to the variation in heating of the thermosphere by solar radiation, but in this case it is not due to geometry but due to processes in the Sun. Although the level of solar radiation in the peak wavelengths of the visible spectrum is quite constant, the thermosphere absorbs only radiation at extreme ultraviolet (EUV) and X-ray wavelengths, which can be highly variable. The level of solar radiation at these wavelengths depends on the presence of active regions on the Earth-facing side of the Sun. Active regions are caused by local densifications in the magnetic field of the Sun. Sunspots often form in those areas.

During minima of the 11-year solar cycle period, there are fewer sunspots, and EUV output by the Sun is usually quite stable and at a low level. During solar cycle peak years however, the EUV output will be much higher and much more variable. This is due both to more rapid changes in the Sun's magnetic field, as well as due to the effect of solar rotation on the restricted view of the Sun's surface from Earth. After the 11-year cycle, the variability of received EUV radiation is most powerful at the period of the Sun's rotation, which is approximately 27 days.

The amplitude of the density variation due to variability in EUV radiation increases with altitude in the thermosphere. Between solar minimum and solar maximum, the density can vary by up to a factor of 20 at 500 km and higher.

The $F_{10.7}$ solar activity proxy (see Sect. 2.2.1) is used in current empirical models to drive this variation. This is apparent from the correlated behaviour of the black lines in Fig. 2.2.

Geomagnetic Activity Variation

Active regions are also related to the occurrence of solar flares and coronal mass ejections, which hurl great quantities of charged particles from the Sun into space. Near the Earth, these particles are mostly deflected by the Earth's magnetic field. Only a fraction can enter through the so-called polar cusps and via so-called reconnection processes at the dayside magnetopause, causing geomagnetic storms which are often accompanied by auroral displays. During these storms, a lot of additional energy is deposited in the polar thermosphere and ionosphere over a relatively short period of time. These sudden and short-lived events can cause density variations of up to one order of magnitude, dominating the normal diurnal variation. The relation of modelled density to peaks in the geomagnetic activity index *ap* (discussed in Sect. 2.2.2) is visible in Fig. 2.2. This relation will be shown in more detail in Sect. 2.2.3.

Semi-Annual Variation

The cause of the fourth major variation in thermospheric density has long been more elusive. The semi-annual variation, discovered by Paetzold and Zschorner [73], leads to density changes of about 30%, with maxima around April and October, although it can be quite irregular in both amplitude and phase [7]. During periods of low solar activity, it is the dominant cause of changes in density, as can be seen in Fig. 2.2. A possible physical explanation was proposed by Fuller-Rowell [34]. It involves an increased mixing of the major thermospheric species due to the uneven heating of the hemispheres at the solstices (the "thermospheric spoon"). This results in raised molecular oxygen and nitrogen and reduced atomic oxygen densities, which in turn lowers the density scale height and causes a compression of the atmosphere, independent of temperature variations.

2.1.2 Vertical Structure of the Atmosphere

By approximation, the pressure p, density ρ and temperature T of the Earth's atmosphere obey the ideal gas law,

$$\frac{p}{\rho} = \frac{RT}{M},$$

(2.1)

where R is the universal gas constant ($8.314472\,\mathrm{J\,K^{-1}\,mol^{-1}}$) and M is the molecular weight of the gas. The vertical structure of the Earth's atmosphere is largely governed by *hydrostatic equilibrium*, a balance between the pressure pushing the gas out into space, and gravity, pulling it towards the Earth's surface. The hydrostatic equation relates the decrease of pressure p with height z to the density ρ and gravitational acceleration g:

$$\frac{dp}{dz} = -\rho(z)g(z)$$

(2.2)

Combining these two equations, atmospheric pressure can be expressed as a function of altitude:

$$p(z) = p(0)e^{-\int_0^z \frac{1}{H(r)}dr}$$

(2.3)

where H is the pressure scale height:

$$H(z) = \frac{RT(z)}{g(z)M(z)}$$

(2.4)

The variation of density with height can be derived similarly, resulting in:

$$\rho(z) = \rho(0)e^{-\int_0^z \frac{1}{H^*(r)}dr},$$

(2.5)

with the density scale height H^* defined as

$$\frac{1}{H^*(z)} = \frac{1}{T(z)}\frac{dT(z)}{dz} + \frac{g(z)M(z)}{RT(z)} \tag{2.6}$$

For regions where temperature is constant, such as the upper thermosphere, the density scale height is equal to the pressure scale height. These scale heights, which depend on the temperature, represent the vertical distance over which the pressure or density decreases by a factor e. At sea level, the scale height is about 8.5 km, while in the thermosphere it is of the order of about 50 km.

Figure 2.1 shows the temperature and density profiles, according to the NRLMSISE- 00 model, for day and night conditions at very low and very high solar activity conditions. Below 100 km, a variety of heat sources and energy transport processes result in distinct temperature minima and maxima, which delineate the named layers of the atmosphere. With ascending altitude these are the *troposphere*, the *stratosphere* and the *mesosphere*. Our interest is in the top layer, the *thermosphere*, which is governed by the absorption of solar extreme-ultraviolet radiation, causing an asymptotically increasing temperature with altitude.

The Sun's radiation causes part of the gas particles in the thermosphere to become ionized, creating the Earth's *ionosphere*. Ionized particles and electrons tend to be greatly outnumbered by neutral particles, at least at altitudes below 800 km. Above that altitude, data on the composition of the thermosphere is very scarce. The presence of oxygen ions as a significant contributor to satellite drag is taken into account in the NRLMSISE- 00 model [76], but not in other models. In any case, the fact that charged particles are present in the upper atmosphere has a profound effect on the medium's electrical properties, which affect its heating and dynamics.

The *mean free path length* of the atmospheric gas is another important parameter. Below the thermosphere, the mean free path length is short, and there are many collisions between the gas molecules. This causes the constituents of the gas to remain well-mixed. This region is called the *turbosphere* or *homosphere*. In the *heterosphere*, collisions are much less frequent. The gases can separate through gaseous diffusion more rapidly than they are mixed by turbulence. According to Eq. 2.6, the density scale height of light-weight constituents is larger than for heavier constituents. Therefore, with increasing altitude, the major thermospheric species are molecular nitrogen (N_2), atomic oxygen (O), helium (He) and hydrogen (H), as can be seen in Fig. 2.3.

The *exosphere* is the region, above approximately 400–600 km (depending on the level of solar activity), from which particles may escape from the atmosphere. This occurs when the kinetic energy of the particle exceeds the gravitational binding energy, and if the mean free path length is long enough to prevent collisions with other particles.
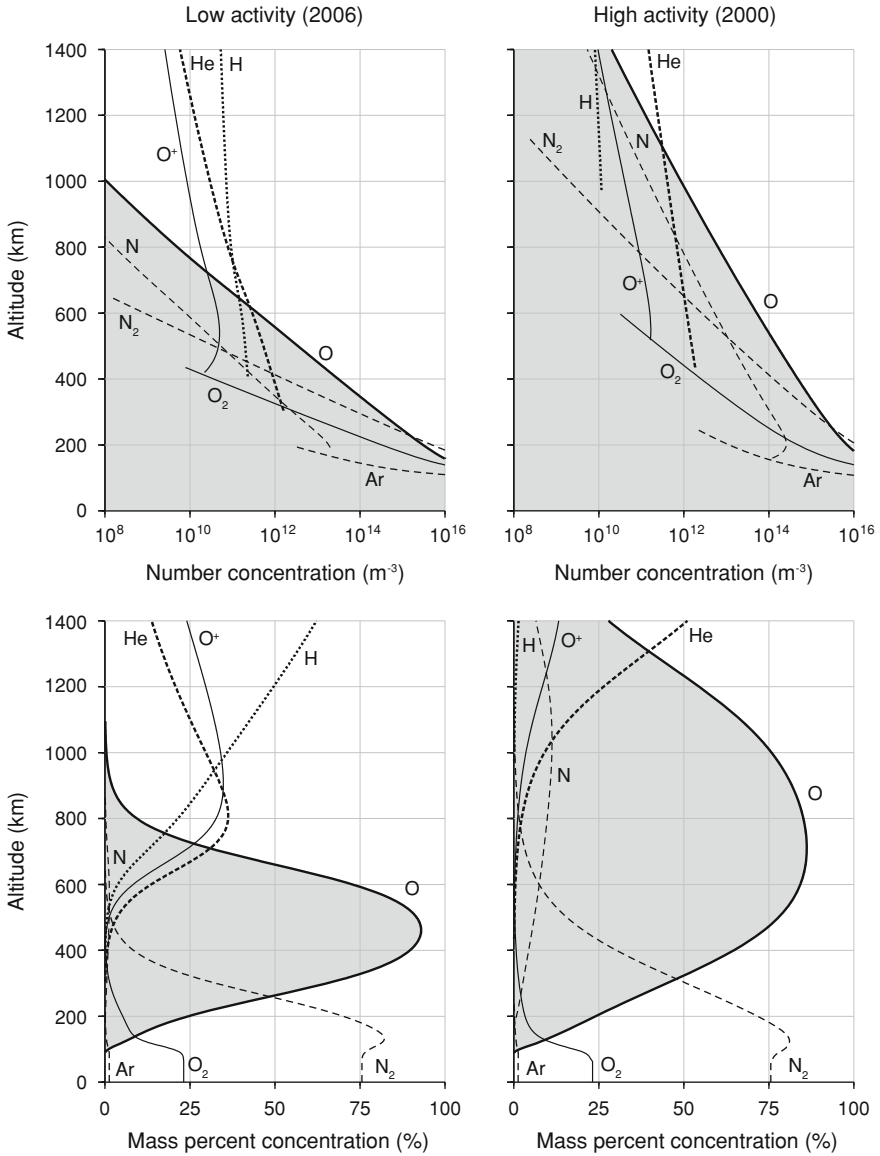
Fig. 2.3 Altitude profiles of atmospheric number concentrations (*top*) and mass percent concentrations (*bottom*), for low (*left*) and high (*right*) solar activity levels, according to the NRLMSISE- 00 model, evaluated over Delft, at 18:00 on July 15, 2000 and 2006

2.1.3 Energy Sources

The temperature profile of the atmosphere, discussed in the previous section, results from the balance between heat sources, loss processes and transport mechanisms. In the thermosphere, the most important sources of heat are solar EUV radiation, related to solar activity, and energetic charged particles, related to geomagnetic activity. Dissipation of tidal motions and gravity waves are minor heat sources, which are neglected in current empirical models. Heat is transported around and out of the thermosphere by conduction, radiation and convection.

For empirical modelling, the representation of density variations with solar and geomagnetic activity are of prime importance. They will be discussed in the next sections.

Solar EUV Radiation

The Sun emits electromagnetic radiation across a wide spectrum. At the wavelengths of visible light, where the solar irradiance is at its maximum, the emission is quite constant, but both at shorter wavelengths (ultraviolet) and longer wavelengths (radio), the level of emission is related to the level of activity of the Sun's magnetic field, and the interaction of this magnetic field with gases on the solar surface, the *photosphere*, and in its atmosphere, the *chromosphere* and *corona*. The emissions at the shortest wavelengths, of 170 nm and less, are responsible for heating the Earth's thermosphere. Radiation in this wavelength range is often designated extreme ultraviolet, or EUV radiation. Note that different publications might use slightly different definitions for the wavelength ranges in this part of the spectrum. The designations XUV (soft X-rays), EUV and FUV (Far Ultraviolet) may be encountered. In the following, we will avoid complications by simply using the term EUV for all solar radiation that heats the thermosphere.

The heating of the thermosphere by EUV radiation occurs through excitation, dissociation or ionization of the atoms or molecules (primarily O, O_2 and N_2). The excess energy of each photon is converted into kinetic energy of the reaction products. Rees [83] and Hargreaves [37] provide more detail about these processes.

Variations in EUV radiation are wavelength-dependent, as the shortest wavelengths are generally formed higher in the Sun's atmosphere and are more variable than longer wavelengths [59]. Radiation at different wavelengths is also generally absorbed at different altitudes in the Earth's atmosphere [78, 89]. Ideally, knowledge of the variability in irradiance over the entire UV and EUV spectrum is required in order to model the heating input in the atmosphere correctly.

The absorption of EUV radiation in the thermosphere results in a so-called diurnal bulge in temperature and density, on the daylight side of the Earth. For this reason, density variations in a horizontal plane are often mapped in a coordinate system of local solar time versus latitude (see Sect. 2.1.4). Local solar time (LST) is equivalent to the longitude coordinate on normal maps, but usually expressed in hours instead
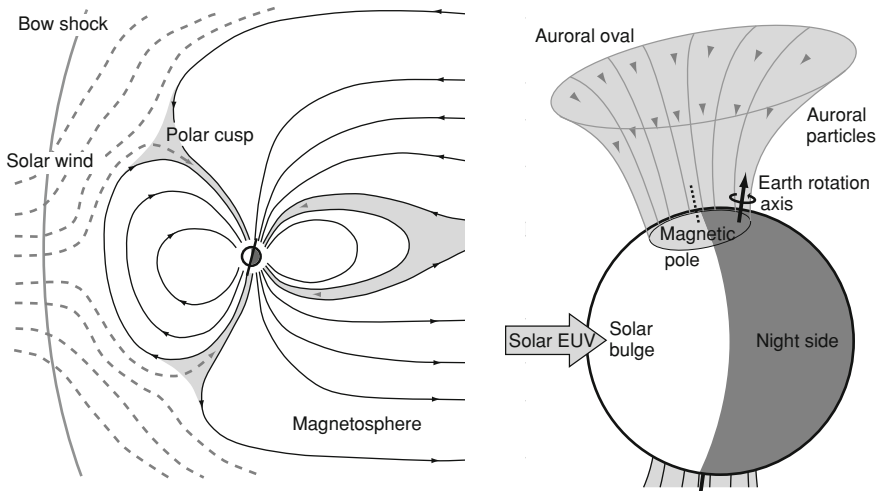
Fig. 2.4 Schematic views of the Earth's magnetosphere, the polar cusps and their relation with processes affecting thermosphere density variations

of degrees, and with the defining meridian, at 12 h (noon) LST, passing through the sub-solar point instead of through Greenwich. Due to thermal inertia, the diurnal bulge has its peak at around 14–15 h local solar time.

Solar Wind and Magnetospheric and Ionospheric Interactions

The Sun constantly emits a stream of particles out into space, called the *solar wind*. The solar wind consists mostly of electrons and protons, with a minor fraction of heavier ions [37]. The motion of these solar wind particles shapes the magnetic field of the Sun into the *interplanetary magnetic field*, which interacts with the terrestrial magnetic field, forming the Earth's *magnetosphere* (see Fig. 2.4). In effect, the terrestrial magnetic field shields the Earth from the solar wind. Solar wind particles can only enter the magnetosphere and upper atmosphere through the *polar cusps* and *reconnection processes* between the interplanetary magnetic field and the Earth's magnetic field.

During quiet times at the Sun, the solar wind originates mainly through *coronal holes*. These occur where magnetic field lines open up into interplanetary space. Sudden increases in the number of solar wind particles, and their speed, are related to *coronal mass ejections*, during which huge amounts of material are injected with high speed into the solar wind, and with *solar flares*, which are sudden short outbursts of electromagnetic energy and particles over a small area on the Sun. These phenomena are all closely related to the complexities of the Sun's magnetic field and occur more often during peak years of the 11-year solar activity cycle. The differences in speed between solar wind particles from various solar sources can cause interplanetary

shocks and compressions [37, 74]. The relationship between the Sun's magnetic field and these processes is the subject of ongoing investigations, with the aid of observatories on the ground and in space, including the SOHO [17] and STEREO [56] scientific satellite missions.

Near the Earth, the interactions between the solar wind, the magnetosphere and the ionosphere are complex. The solar wind particles are slowed down at the bow shock and are largely deflected around the magnetopause. They can only enter near the geomagnetic poles, i.e. approximately within the polar cap regions, where the Earth's magnetic field lines are open. Inside the magnetosphere, an extensive system of currents is present, as well as regions where charged particles are trapped. The system is complex, because the motion of the charged particles induces the currents and electric fields that in turn influence the magnetic field, which alters the motion of the charged particles. In the end, the processes on the Sun, which were briefly described above, can result in large disturbances at the Earth, called *geomagnetic storms* [e.g., 65, e.g., 35, e.g., 33, e.g., 94]. The physics of magnetospheric processes is a subject of intensive ongoing study. Moldwin [69], provides a basic introduction to the topic. Several space missions, including ESA's Cluster constellation [31], are helping to unravel the intricacies. These physical processes are studied by making use of extensive coupled thermosphere, ionosphere and magnetosphere models.

For the thermosphere, the most important aspect of this complex system is the mechanism in which energy is deposited. A relatively small part of this is due to particle precipitation. A larger contribution is due to Joule heating by ionospheric electric currents. The amount of energy deposited in the thermosphere through these two processes combined, normally amounts to about one fourth of the energy deposited by EUV radiation. After extraordinary large solar flares and coronal mass ejections, however, the energy deposited in the auroral zone can be up to twice that of the EUV energy [58].

Because there is an offset of the order of 10° between the dipole component of the Earth's magnetic field and its rotation axis (see Fig. 2.4), the location of maximum energy absorption in the auroral zone shows a diurnal shift when viewed in the local solar time versus geographic latitude reference frame. Variations in the thermosphere that are affected by geomagnetic activity are therefore often calculated and viewed in a system of geomagnetic coordinates.

2.1.4 Horizontal Structure

Now that the different energy inputs have been introduced, it is possible to have a look at how they affect variations in the thermosphere in a horizontal plane. Figures 2.5 and 2.6 show temperatures, densities and wind speeds from the NRLMSISE- 00 [76] and HWM- 93 [45] models, at 400 km altitude above the Earth's surface. Figure 2.5 shows the situation at low solar activity, when solar EUV radiation is the dominant energy source. Figure 2.6 shows the situation for the exact same time of day and day of year, only six years earlier. The factors influencing the diurnal and seasonal

variation are exactly the same, however the energy input is completely different. The $F_{10.7}$ proxy has a three times higher value in 2000, compared with 2006. Even more importantly, the snapshot of model values in Fig. 2.6 is taken at the peak of a major geomagnetic storm. How these circumstances affect the temperature, density and wind speeds will be explained in the following paragraphs.

Temperature and Density Variations at 400 km Altitude

The diurnal bulge, caused by EUV heating, is the dominant feature in the horizontal plane for temperature and density at low solar activity as seen in the top two frames of Fig. 2.5. The maximum density, at around $1.2 \ 10^{-12} \, \text{kg/m}^3$, is located several hours East of the sub-solar point. It is summer in the Northern hemisphere, and the maximum temperature there is around 900 K. The temperature minimum, of around 600 K is close to the South Pole.

At high solar and geomagnetic activity, Joule heating and particle precipitation in the auroral zones have become the dominant energy sources, so that the diurnal variation is hardly visible anymore. Temperatures have become twice as large as at solar minimum, and now there are maxima around both the Poles. The temperature minimum is located near the night-side equator.

The density at high solar activity is more than a factor 10 larger than at low activity. The density maximum has shifted towards the auroral zone, while there are now minima at mid-latitudes on the night-side.

These Figures are just snapshot images of the model at specific moments in time. Although the NRLMSISE-00 model shows increased variability during this geomagnetic storm, compared to low activity conditions, its spatial resolution is actually quite coarse when compared with density observations derived from accelerometer measurements, as will be shown in Chap. 5.

Horizontal Winds in the Thermosphere

The bottom two rows in Figs. 2.5 and 2.6 show horizontal wind speeds and directions at 400 km altitude, obtained from the HWM-93 model. This model outputs zonal and meridional wind speeds, with respect to a corotating atmosphere. The wind direction at the Poles is therefore undefined.

The third row in the Figures shows just the output of the model, which is a representation of the wind that could be measured by an observer on the ground. The bottom row adds the speed of the Earth's rotation to the output of the model. This is a representation of the wind that influences the motion of satellites. The corotation velocity is proportional to the cosine of the latitude, which goes from a maximum at the equator (483–502 m/s at altitudes of 250–500 km) to zero at the poles. Figure 2.5 shows that the HWM-93 output is below 200 m/s and is therefore quite small compared to the corotation velocity at low latitudes. The deviation from corotation becomes more important in the polar areas, and at high activity, as can
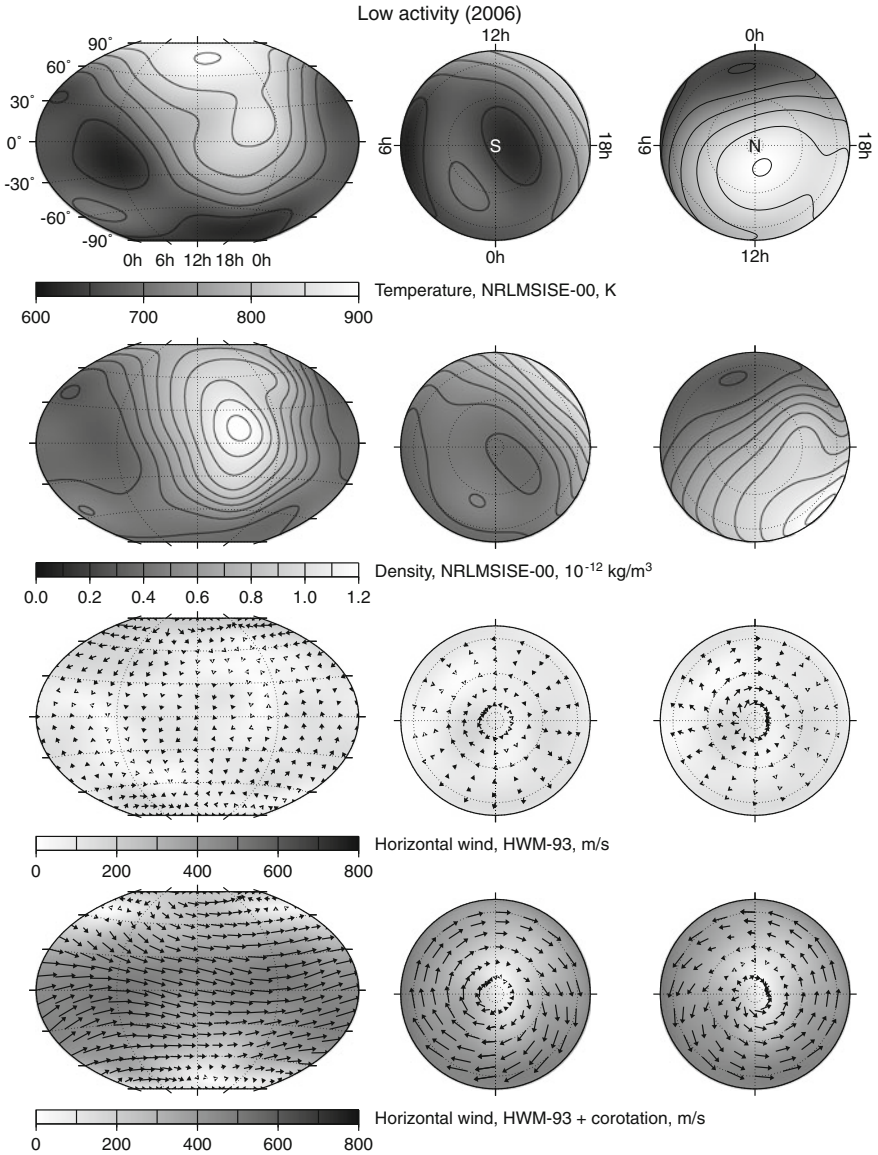
Fig. 2.5 Maps of modelled temperature, density and wind speeds at 400 km altitude, on July 15, 2006 at 18:00:00 UTC ($F_{10.7} = 70.2$, $ap = 3$). The map projections are, from *left* to *right*, a Winkel Tripel projection, and two orthographic hemispherical projections centered on the geographic South Pole and North Pole
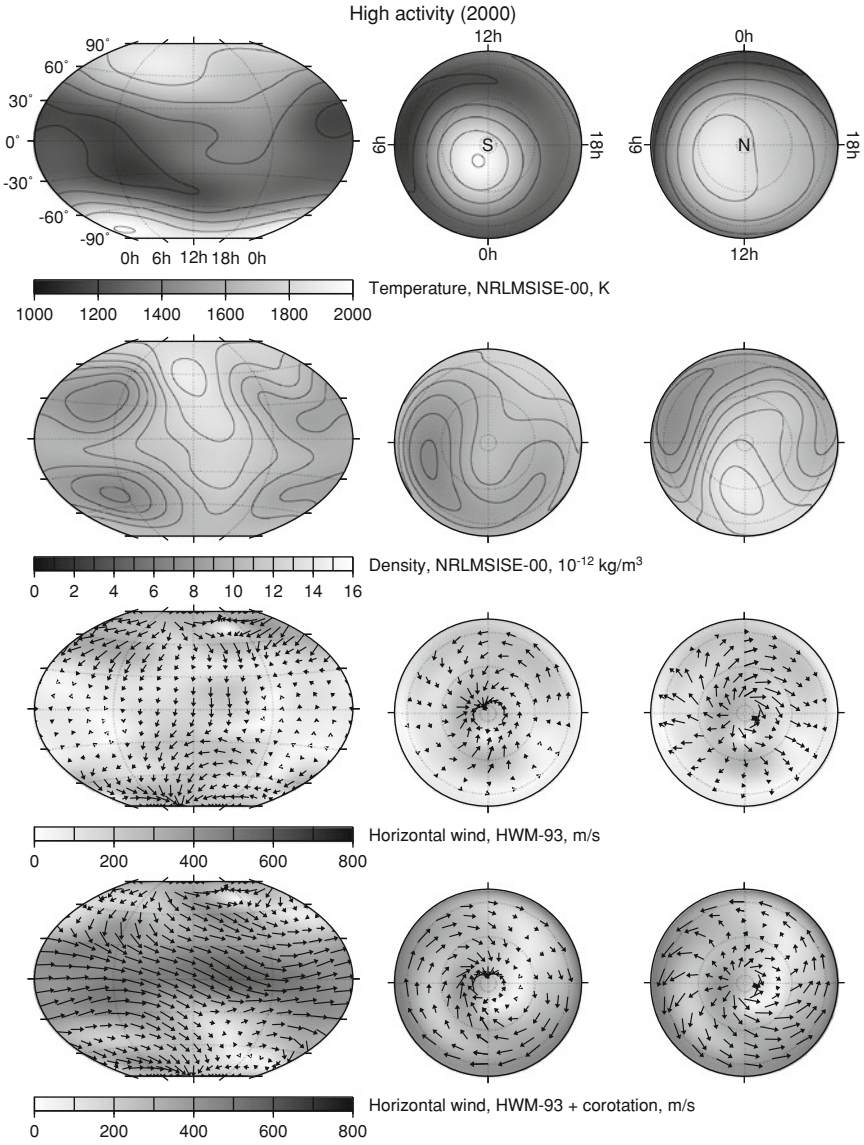
Fig. 2.6 Same as Fig. 2.5, but for July 15, 2000 at 18:00:00 UTC ($F_{10.7} = 213.1$, $ap = 400$)

be seen in Fig. 2.6. Unfortunately, the observation data at high latitudes on which models such as HWM- 93 are based, is quite sparse. Accelerometer-derived wind data, if properly calibrated, could be very helpful in that respect.

Altitude Dependence of the Horizontal Density Structure

Figures 2.1 and 2.3 have provided a representation of the variation of density with height at a single location, while Figs. 2.5 and 2.6 have shown the variations in the horizontal plane at a single altitude level of 400 km. Now, in Fig. 2.7, both views are combined, to get a more coherent picture of density variations in the thermosphere.

It is clear from the Figure that the situation at 400 km is quite representative for a large range in altitude. This is especially true at low solar activity, when the diurnal bulge is the dominant feature. However, at 100 km, *semi-diurnal tides* dominate. In this region, the behaviour of the thermosphere is tightly linked to that of the mesosphere, and density levels are actually lower during solar maximum than at solar minimum, according to the model.

The behaviour of total density above approximately 600–800 km is also quite different. There is a distinct density peak over the winter hemisphere. This *winter helium bulge* is caused by the atmospheric dynamics of lower altitudes [64], where helium is a minor constituent. The minor light-weight gas is dragged along with the circulation of the major gas flow at lower altitudes, but because of its large scale height, more helium is transported to the winter pole than is returned to the summer pole [13].

2.1.5 Temporal Variability

Up to this point, the figures in this Section have focussed on two distinct times with extreme low and high solar and geomagnetic activity conditions. To complete our overview of the major variations in the thermosphere, we go back to Fig. 2.2, which showed the magnitude of the day-to-day variability of density over the course of the a solar cycle, and zoomed in views for the years 2000 and 2006, as modelled by NRLMSISE- 00. During the year 2000, the approximately 27-day solar rotation period is a dominant feature in both $F_{10.7}$ and density. During the more quiet year of 2006, the semi-annual variation with maxima in April and October is more clearly visible, and the relative effect of occasional peaks in geomagnetic activity are larger. Note for example the three peaks in April 2006.

2.2 Observations, Proxies and Indices of Thermospheric Energy Inputs

In the previous Section, the large influence of processes on the Sun, in the magnetosphere and in the ionosphere on the thermosphere has been explained. Therefore, observations of these processes are very useful in order to study and model the

Low activity (2006)   High activity (2000)

2000 km
38.98 10^{-18} kg/m^3
176.06 10^{-18} kg/m^3

2000 km
0.13 10^{-15} kg/m^3
0.90 10^{-15} kg/m^3

1000 km
0.25 10^{-15} kg/m^3
2.23 10^{-15} kg/m^3

1000 km
5.61 10^{-15} kg/m^3
46.74 10^{-15} kg/m^3

800 km
0.64 10^{-15} kg/m^3
6.62 10^{-15} kg/m^3

800 km
38.11 10^{-15} kg/m^3
226.56 10^{-15} kg/m^3

700 km
1.25 10^{-15} kg/m^3
12.04 10^{-15} kg/m^3

700 km
0.12 10^{-12} kg/m^3
0.54 10^{-12} kg/m^3

600 km
3.83 10^{-15} kg/m^3
30.64 10^{-15} kg/m^3

600 km
0.45 10^{-12} kg/m^3
1.42 10^{-12} kg/m^3

500 km
25.80 10^{-15} kg/m^3
179.03 10^{-15} kg/m^3

500 km
1.70 10^{-12} kg/m^3
4.19 10^{-12} kg/m^3

400 km
0.29 10^{-12} kg/m^3
1.24 10^{-12} kg/m^3

400 km
6.79 10^{-12} kg/m^3
14.67 10^{-12} kg/m^3

300 km
4.60 10^{-12} kg/m^3
11.30 10^{-12} kg/m^3

300 km
34.77 10^{-12} kg/m^3
65.43 10^{-12} kg/m^3

200 km
0.13 10^{-9} kg/m^3
0.21 10^{-9} kg/m^3

200 km
0.31 10^{-9} kg/m^3
0.51 10^{-9} kg/m^3

100 km
0.28 10^{-6} kg/m^3
0.73 10^{-6} kg/m^3

100 km
0.20 10^{-6} kg/m^3
0.41 10^{-6} kg/m^3

Fig. 2.7 Maps of modelled density on July 15, 2006 at 18:00:00 UTC (*left*) and at the same date and time in 2000 (*right*). The locations of the absolute minimum and maximum values at each altitude are indicated by the white and *black triangles*, respectively. The corresponding values are provided alongside the maps. Contour lines are drawn at 20% intervals between the minimum and maximum.

thermosphere. These observations are often available in the form of *proxies* and *indices*. A proxy is an observation that is related to, and shows similar behaviour as the phenomena of interest, but that is easier to obtain, or for which the historical

record is more complete, than a direct observation. An index is a quantity that is derived from observations according to a pre-designed standard. Indices are often designed to summarize complex phenomena from multiple observations, facilitating the representation in models and the quick identification of disruptive events.

2.2.1 Solar EUV Observations and Proxies

Measurements of the Sun's radiative output at EUV wavelengths are difficult, as they need to be made by satellites above the absorbing atmosphere. To make matters worse, instrumentation sensitive to this radiation also degrades significantly under its influence [32, 59]. Measurements from satellites started in the 1960s, but they have only been available on a regular basis since the 1990s. These recent missions have also introduced calibration techniques to account for the effects of instrument degradation. Such missions include UARS [82], launched in 1991, SOHO [17], launched in 1995, TIMED [75], launched in 2001 and SORCE [85], launched in 2003.

It is important to note that investigations on the thermosphere before the 1990s had to be made without reliable direct and continuous observations of variations in EUV radiation, contemporaneous with the available data on density. Instead, investigators had to rely on observations of solar activity in parts of the spectrum that are available on the ground, such as the number of sunspots visible on the solar disc, or the level of radiation at radio-wavelengths, which can be used as proxies for the EUV radiation.

10.7-cm Radio Flux

Sunspot numbers have been recorded since their first observation by Galileo in 1610, but a better correlation with thermospheric density was found with $F_{10.7}$ [50], which is the amount of energy received from the Sun at 10.7 cm wavelength. $F_{10.7}$ measurements have been made daily by radio telescope, operated by the National Research Council at Ottawa from 1947 to 1991 and at Penticton, British Columbia from 1991 onwards. $F_{10.7}$ is always supplied in units of 10^4 Jansky (Jy) $= 10^{-22}$ W/m^2 /s. In plots and publications, the units are usually omitted, or designated simply as solar flux units (sfu). The proxy ranges from below 70 during solar minimum, to around 370 during extremely active days.

Early researchers such as Jacchia [50] first recognized the solar radiation influence on the thermosphere from the 27-day variations which were found in orbital drag time series. It was later found that a more slowly varying component of solar radiation was also present. The moving average of $F_{10.7}$ over three or four solar rotations (81 or 108 days), denoted $\bar{F}_{10.7}$ was used to represent this component.

The $F_{10.7}$ data are published in two forms: the *observed* fluxes are the actual measured values, while the *adjusted* fluxes are scaled to a standard distance of 1 AU, to compensate for the varying distance between the Earth and Sun through the year.

While the adjusted values give a better representation of the processes taking place on the Sun, the observed values better represent the EUV heating input in the upper atmosphere. Therefore, the latter should be used as input to the density models.

Mg II Core-to-Wing Ratio

It is now widely recognized that the use of $F_{10.7}$ and $\bar{F}_{10.7}$ as proxies in empirical density models has become a limiting factor in further improving their accuracy. With the availability of better satellite observations and empirical models of solar irradiances, researchers have started looking for alternative, more representative proxies, of which the most widely studied example is Mg II.

The use of the Mg II core-to-wing ratio (or just Mg II for short), was first proposed by Heath and Schlesinger [39], and is computed by taking the ratio of the intensity of the Mg II emission at 280 nm over that of the nearby line wings. The use of this ratio makes the index much less sensitive to instrument trends, since the wavelength dependence on degradation is limited. Mg II observations have been made from various satellites with different instrument characteristics. From these different datasets, composite time series have been constructed that often do not go further back than the 1990s. These time series correlate well with $F_{10.7}$. For density studies, the Mg II time series are often scaled using linear regression with $F_{10.7}$ so that they can be expressed in the same units.

Combinations of EUV Proxies

Of specific interest with regard to the use of EUV measurements for thermosphere modelling is the work by Dudok de Wit et al. [28], who performed a principal component analysis on EUV spectra recorded by the TIMED satellite. An important result of this analysis is that the full EUV spectrum can be reconstructed with high accuracy from just a small number of spectral lines. The work confirms that a single spectral line or proxy is clearly not sufficient to represent variability at the different wavelengths within the EUV spectrum that affect the thermosphere. This technique provides graphical representations of the similarity and dissimilarity of the behaviour of the various spectral lines and EUV proxies.

The use of a combination of proxies for representing EUV input in the thermosphere has been the strategy of Tobiska et al. [89] and Bowman and Tobiska [6] in preparation for the JB2006 density model. After testing several combinations, they arrived at the use of $F_{10.7}$, Mg II and S_{EUV}, and their 81-day averages. The S_{EUV} index is a measure of the integrated 26–34 nm emission, observed by SOHO. The Mg II and S_{EUV} measurements have been averaged to daily values and converted by linear regression to the same value range as $F_{10.7}$, resulting in the new indices $M_{10.7}$ and $S_{10.7}$. For the update of the Jacchia-Bowman model to its 2008 version, JB2008 [9], an additional solar activity index was added, designated $Y_{10.7}$. It is a composite index, designed to represent 0.1–0.8 nm X-ray emissions during solar maximum and

Lyman-α emissions during moderate and low solar activity. Separately from these modelling efforts, solar indices have recently been analysed by Guo et al. [36] using CHAMP accelerometer observations.

2.2.2 Geomagnetic Activity Observations and Indices

Satellite observations of charged particles from the solar wind and their interaction with the magnetosphere have been available since the late 1970s and early 1980s. These observations were made from NOAA's TIROS and POES satellites, as well as by the DMSP program [58]. More recently, observatories of the Sun, the magnetosphere and the interplanetary environment such as SOHO, ACE, Cluster and STEREO, have greatly increased our knowledge of these complex environments and their interactions.

These measurements have never been used in empirical thermosphere models though. Thermosphere models have traditionally made use of the *Kp* and *ap* geomagnetic indices, which are extensively described in the monograph by Mayaud [63], and the more accessible review article by Menvielle and Berthelier [66]. The information in this section draws heavily from these sources.

The *Kp* and *ap* Geomagnetic Indices

The *Kp* index is derived from measurements made at 11 geomagnetic observatories at mid-latitudes, through a rather intricate procedure, described by Menvielle and Berthelier [66]. At these stations, the horizontal magnetic components are separated in regular and irregular variations. The intensity of the irregular variations is expressed in a *K* index at each station, over each three-hour period.

This *Kp* index, as defined by Bartels et al. [3], ranges from 0 to 9, in discrete steps of one third, specified by plus and minus signs, as follows: 00, 0+, 1−, 10, 1+, . . ., to 90. The *K* indices from the 11 stations are standardized to avoid local time influences and averaged to form the planetary index *Kp*. Since *Kp* is a true index, it carries no physical units. It can be converted to the approximate amplitude of geomagnetic perturbation by the use of a conversion table, resulting in *ap* (planetary equivalent amplitude) values, expressed in units of nanoTesla (nT). These *ap* values should be multiplied by two to represent the magnitude of the observed variation in the magnetic field. Since the *ap* index is related directly to *Kp*, it also has a 3-hourly cadence, and can take no other values than those 27 listed in the *Kp* to *ap* conversion table. The average of all eight *ap* values in a UTC day is denoted as the *Ap* index. Time series of *Ap* can be useful for creating plots and quickly identifying active days, but the loss of temporal resolution with respect to *ap* makes the daily index less suitable for use in density models.

Alternative Geomagnetic Indices

Kp and related indices are widely used in many applications, including thermosphere models. However, by today's standards, they show some obvious shortcomings [66, 67]. Since the *Kp* network was set up at the start of the cold war period, most stations are located in Western Europe and North America. Relatively minor magnetic disturbances over one of these areas will therefore result in high values of the index, while disturbances of similar magnitude in uncovered areas might go largely unnoticed.

A new global network, with a much better geographical distribution, was set up after the International Geophysical Year (1957/1958), coinciding with the beginning of the space age. The *an*, *as* and *am* indices, all expressed in nT are derived from the *K* indices measured at these stations. The *an* index represents the northern hemisphere, *as* the southern, and *am* (where *m* stands for 'mondial') is the average of both hemispherical indices. The *am* index can therefore be considered an improved version of *ap*.

The *longitude sector geomagnetic indices*, or aλ indices, are *K*-derived indices which are provided for various pre-defined longitude sectors [68]. Five of these sectors, each of which contains two or more observing stations, are in the Northern hemisphere, and four are in the Southern. These indices are thus able to represent variations with longitude.

Other parameters, such as the *Dst* (*disturbance storm time*) and *PC* (*polar cap*) indices represent only certain parts of the geomagnetic activity, at low and high latitudes respectively, which are mainly driven by distinct current systems. See, for example, Campbell [14], Troshichev et al. [90], Asikainen et al. [1] and references therein. These two indices are available at an hourly or higher rate and therefore offer improved temporal resolution compared to the 3-hourly *K*-derived indices. Knipp et al. [58] provide an empirical model for the power input by Joule heating in the thermosphere, based on quadratic equations using these two indices.

2.2.3 Comparison of Proxies and Indices

Figure 2.8 shows a comparison of time series of several of the proxies and indices described in the previous Sections, along with densities derived from the accelerometer observations by CHAMP and GRACE. During the period of late October to the end of November 2003, the variations in solar activity were especially large, due to the presence of large active regions on the surface of the Sun (see Fig. 2.2 for context). Two of the largest geomagnetic storms of the solar cycle occurred during the last days of October, and on 20 and 21 November. Note that the behaviour of these indices during this time period is not necessarily indicative for other time periods.

The bottom panel in the plot shows the variation of the four daily solar radiation proxies that are provided for the JB2008 model. During this period, $F_{10.7}$, which is used by older models as well, shows the largest amplitude. The $M_{10.7}$ and $S_{10.7}$
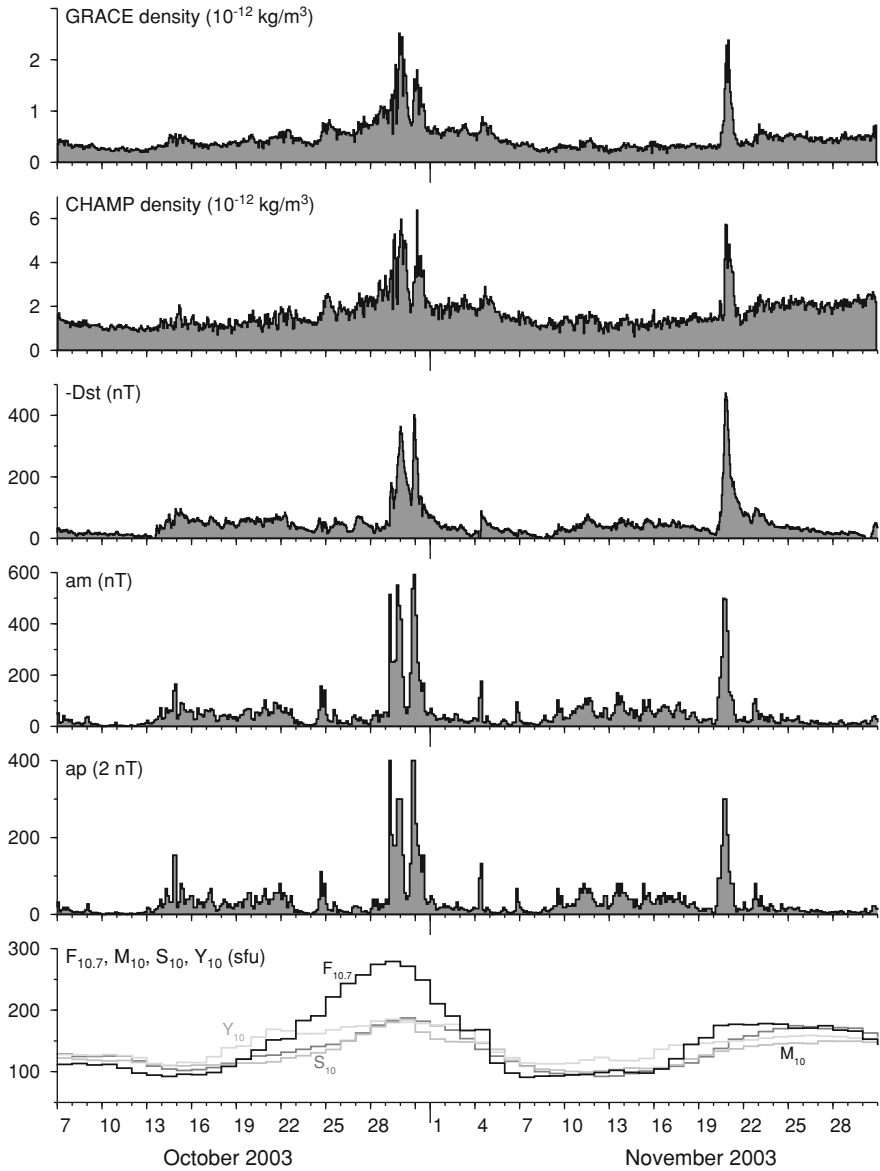
Fig. 2.8 Comparison of orbit-averaged densities from CHAMP and GRACE with geomagnetic indices and solar activity proxies

proxies, introduced with the JB2006 model, show similar values during low activity, but differences are considerable around the time of the storms. The $Y_{10.7}$ proxy, which was a new addition in the JB2008 model shows an earlier onset of the two periods of increased activity during these months.

Three other panels in Fig. 2.8 show the hourly *Dst* and the three-hourly *am* and *ap* indices. The *am* and *ap* profiles look quite similar. The main difference between these indices is in the improved global distribution of the observation network for the *am* index. The result of this is especially apparent in the relative magnitudes of the three peaks of the October storm. The *Dst* index, which returns predominantly negative values, has been multiplied by -1 here for easier comparison with the other indices. *Dst* provides more detail, because of its hourly rate, but it only represents the geomagnetic disturbances in the equatorial region. After major disturbances, *Dst* takes longer to recover than the *am* and *ap* values. Notice also how the disturbance on October 24 is not clearly present in the *Dst* plot, and how the maximum amplitude of the November storm exceeds that of the October storm in *Dst*, but not in *am* or *ap*.

From this example and the discussion in the previous Sections, it is clear that there is a wide variety of observations, proxies and indices available for research in thermospheric density and its connection to processes on the Sun. In the early 1960s, the combination of $F_{10.7}$ and *ap* was chosen to represent all non-periodic variations in density. This was a convenient choice, because long time series for these values were already easily available at the time. Empirical modellers have stuck to this choice until the recent introduction of new proxies and indices with the JB2006 and JB2008 models.

2.3 Empirical Density Models

The development of new density models has been largely driven by the availability of new sources of observation data. With the availability of more data, the models have become able to represent more subtle variations in density. Table 2.1 lists the most widely known models and their references. Jacchia, MSIS and DTM are the three major families of models, which have been under development since the 1960 and 1970s. Several models of these families will be evaluated and used in this research, and they are therefore introduced in the remainder of this chapter.

2.3.1 Early Models and Jacchia's Models

The first observations of orbital motions of satellites under the influence of drag in the late 1950s paved the way for the first generation of general empirical density models in the 1960s and early 1970s. Examples of these models include those by Harris and Priester [38] and Jacchia [51–53, 55]. The Jacchia-71 model was adopted as the COSPAR International Reference Atmosphere in 1972, and is therefore also known as CIRA- 72 [54]. Later models by Jacchia, such as those published in Jacchia [55], have not led to accuracy improvements with respect to this model [61]. The Jacchia-70 and -71 models have been used as the basis of the NASA Marshall Engineering Thermosphere (MET) models [49, 72], which have been applied in satellite lifetime

Table 2.1 List of empirical thermosphere models, their literature references, and the source of the FORTRAN code for the models that have been obtained and implemented for use and evaluation in this study

Model family	Model name	Reference	Source code
Jacchia	Jacchia-64	Jacchia [51]	
	Jacchia-70	Jacchia [52]	
	Jacchia-71	Jacchia [53]	
	Jacchia-Roberts-71	Roberts [84]	
	CIRA-72	Jacchia [54]	In reference
	Jacchia-77	Jacchia [55]	
	MET	Hickey [49]	
	MET 2.0	Owens [72]	
	JB2006	Bowman et al. [10]	Internet[1]
	JB2008	Bowman et al. [9]	Internet[2]
DTM	DTM	Barlier et al. [2]	In reference
	DTM-94	Berger et al. [4]	From author
	DTM-2000	Bruinsma et al. [11]	
MSIS	MSIS	Hedin et al. [43, 44]	
	MSIS-83	Hedin [40]	
	MSIS-86	Hedin [46]	Internet[3]
	CIRA-86	Hedin [41]	
	MSISE-90	Hedin [42]	
	NRLMSISE-00	Picone et al. [76]	Internet[3]
HWM	HWM-87	Hedin et al. [47]	
	HWM-90	Hedin et al. [48]	
	HWM-93	Hedin et al. [45]	Internet[4]
	HWM07	Drob et al. [27]	Internet[4]
	DWM07	Emmert et al. [30]	Internet[4]
	Harris-Priester	Harris and Priester [38]	
	TD88	Sehnal [86]	
	GOST	see Vallado [91, App. B]	
	GAMDM	Emmert and Picone [29]	In reference

[1] http://sol.spacenvironment.net/$\sim$ignorespacesjb2006/
[2] http://sol.spacenvironment.net/$\sim$ignorespacesjb2008/
[3] http://uap-www.nrl.navy.mil/models_web/msis/msis_home.html
[4] http://uap-www.nrl.navy.mil/models_web/hwm/hwm_home.html

predictions. More recently, Jacchia's models have been further improved under an initiative of the US Air Force Space Command, with new algorithms for computing the semi-annual density variation and the use of additional solar and geomagnetic activity proxies. This effort has resulted in the Jacchia-Bowman 2006 (JB2006) and 2008 (JB2008) models [9, 10].

2.3.2 DTM Models

The first of the DTM models was published in 1978 [2]. Like the Jacchia models, DTM was based solely on observations of satellite drag and neutral atmospheric temperatures, however it used a more generalised model formulation. A major update, named DTM- 94, was published by Berger et al. [4]. DTM- 94 is based on an extended satellite drag dataset for improved accuracy at solar minimum and solar maximum conditions. It included Dynamics Explorer 2 mass spectrometer and CACTUS accelerometer data as well. DTM model improvement efforts continued with the introduction of the Mg II solar EUV proxy [88] in DTM- 2000 [11] and incorporation of new data such as those from the CHAMP accelerometer [12].

2.3.3 MSIS and HWM Models

Starting in the late 1970s Alan Hedin created a new class of density models, named MSIS. His models [43, 44] were based solely on mass spectrometer and incoherent scatter radar observations. The main advantage of the use of these datasets over the drag-derived datasets is that they consist of independent observations of both temperature and number densities for the atmospheric constituents. The MSIS- 86 model replaced Jacchia-71 as the COSPAR International Reference Atmosphere thermosphere model, and is therefore also known as CIRA- 86 [41, 46]. An extension to the model was published as MSISE- 90 [42]. This model is identical to MSIS- 86 for the thermosphere region, but extends down to zero altitude.

At the end of the 1990s, development of the MSIS-class models was continued at the US Naval Research Laboratory (NRL). The resulting NRLMSISE- 00 model [76] includes additional mass spectrometer and incoherent scatter radar data, as well as accelerometer data and the DTM and Jacchia satellite orbit decay databases. An additional constituent, anomalous oxygen, was introduced in the model to investigate apparent discrepancies between the datasets at higher altitudes.

The HWM series of thermospheric horizontal wind models was created alongside the MSIS model series. The model subroutine accepts the same inputs as the MSIS series of density models and returns zonal and meridional wind speeds. The models are based on gradient winds from CIRA- 86 plus rocket soundings, incoherent scatter radar, MF radar and meteor radar data. Further development of the HWM model series is now also continued at the Naval Research Lab. The latest edition is HWM07 [27], which has a companion model for high geomagnetic activity, the Disturbance Wind Model DWM07 [30].

2.4 Empirical Model Implementations and Limitations

The empirical models considered in this chapter were all created by fitting a set of parametric equations to an underlying database of observations. The accuracy of the models is therefore dependant on both the strength of the database as well as the ability of the parametric equations to reproduce these data for interpolation and extrapolation. There are considerable differences between the number and complexity of the parametric equations used in the various models. Full details on the model formulations and generation can be found in the references provided in Table 2.1. However, a few details which are of importance for this study will be provided here.

2.4.1 Jacchia's Algorithm

In short, the Jacchia models use the exospheric temperature T_∞ as a defining parameter. This temperature is calculated for each location from the solar activity data, in combination with a model of the diurnal variation. Additional corrections are applied, depending on time and altitude. These include the semi-annual variation, a seasonal-latitudinal variation and an additional geomagnetic activity term. The exospheric temperature is then used as input for a temperature profile, which forms the basis for the integration of the diffusion equation.

The first step in the CIRA- 72 density computation is the calculation of the culmination point temperature. This temperature, denoted by T_c, is the night-time minimum of the global exospheric temperature field, when the geomagnetic index K_p is zero. It is a function of the solar activity proxy and its 81-day average, and three empirical constants (α, β and γ):

$$T_c = \alpha + \beta \bar{F}_{10.7} + \gamma(F_{10.7} - \bar{F}_{10.7}) \tag{2.7}$$

T_c is then multiplied with the diurnal variation function $D(t, \phi, lst)$. This function, which does not depend on the solar or geomagnetic activity, defines the shape of the diurnal bulge for the model, as shown in Fig. 2.9. It provides a smooth variation between the minimum temperature T_c and a maximum temperature of 1.3 times T_c to the East of the sub-solar point.

The calculation of the local exospheric temperature T_∞ is completed by adding a temperature correction for geomagnetic activity, ΔT_G

$$T_\infty(t, \phi, \lambda) = T_c(t) \, D(t, \phi, \lambda) + \Delta T_G(a_p) \tag{2.8}$$

Jacchia's temperature profile is further defined by a boundary condition of $T = 183 \, K$ and $dT/dh = 0$ at the minimum altitude of 90 km and an inflection point at an altitude of 125 km. The inflection point temperature T_x is a function of T_∞ and four empirical constants (a, b, c and k)
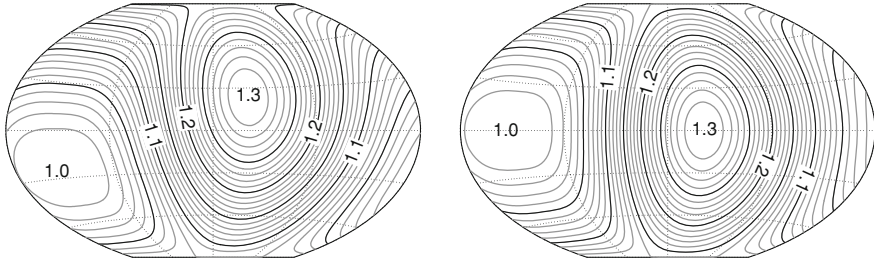
Fig. 2.9 The CIRA–72 diurnal temperature function at solstice (*left*) and equinox (*right*). The map is a Winkel-Tripel projection in local solar time versus latitude

$$T_x(T_\infty) = a + bT_\infty + c \, \exp(kT_\infty) \tag{2.9}$$

With the complete temperature profile now in place, the model computes the number density of each individual species by integrating the diffusion equation. The density output is completed by adding the empirical corrections for geomagnetic activity variation, the semi-annual variation and seasonal-latitudinal variation.

2.4.2 The DTM and MSIS Algorithms

In the DTM and MSIS models, the temperature profile with height is a function of the exospheric temperature. Just as in Jacchia's model, the exospheric temperature is determined from solar and geomagnetic activity proxies and indices. Unlike Jacchia's models, however, the density in the DTM and MSIS models is not derived from a direct integration with altitude over the temperature profile. Instead these models basically contain independent sub-models for the thermospheric temperature, T_∞, and number densities, n_i, of each of the constituents. The index i is used to iterate over each of the atmospheric constituents (H, He, O, etc.), and a set of model coefficients a_j (with $j = 0 \ldots N$) is defined for the temperature and densities of each of the constituents. We will use the notation and equations from the DTM- 94 model [4] in this Section. The formulation for the thermospheric part of the MSIS models is largely similar. The modelled thermospheric temperature and concentrations are expressed using a function $G(L)$ of the input parameter vector L as follows

$$T_\infty = a^0_{T_\infty}(1 + G_{T_\infty}(L)) \tag{2.10}$$

$$n_i(z) = a^0_i \exp(G_i(L)) f_i(z) \tag{2.11}$$

where $f_i(z)$ is the altitude law, based on the integration of the differential equation of diffusive equilibrium (see Berger et al. [4] for details).

The total density can then be computed from the concentrations by multiplying with the molecular mass m_i of the constituents and summing all constituents.

$$\rho_i = \frac{m_i n_i}{N_A} \tag{2.12}$$

$$\rho = \sum_i \rho_i \tag{2.13}$$

where $N_A = 6.022 \ 10^{26} \ \text{kmol}^{-1}$ is Avogadro's number.

The functions $G(L)$ can contain a large number of terms, for instance:

- Polynomial terms with the solar and geomagnetic activity proxies and indices,
- Spherical harmonic terms in geomagnetic latitude and geomagnetic local time,
- Periodic terms in day of year and time of day, and
- Cross terms of solar and geomagnetic activity with the periodic variations.

The coefficients for these terms have been determined in a least squares adjustment using a large database of density and temperature measurements. There is a large difference between the complexity of the DTM models and that of the MSIS models. This is mainly due to the fact that the MSIS models incorporate many more variations of density and temperature in the lower atmosphere. As an example, the DTM- 94 model is defined by 172 non-zero coefficients, while the NRLMSISE- 00 model requires approximately 2,200.

2.5 Density Model Calibration

The density observations that will be derived using the algorithms presented in Chap. 4 have many applications in space physics research. Ultimately, such research should result in updated, more accurate empirical and physical models of the thermosphere. This generally requires the availability of other types of data, as well as expertise from other disciplines, such as solar irradiance monitoring and thermosphere-ionosphere-magnetosphere coupling. These are exciting fields of research, but due to their extent they must largely remain outside the scope of the remainder of this thesis.

There is, however, a simpler and less time consuming way in which the density observations derived from satellite dynamics can contribute to the improvement of empirical density models and their applications in operational and precise orbit determination. This is done by assimilating the density data in an adjustment of the empirical model output. This process, which will be introduced below, has become known as empirical density model calibration. The output of the HASDM calibrated model will be used to evaluate density measurements in Chap. 5, while an independent implementation of the calibration technique, developed under contract for the European Space Operations Centre ESOC, will be the main subject of Chap. 6.

Traditional empirical density model
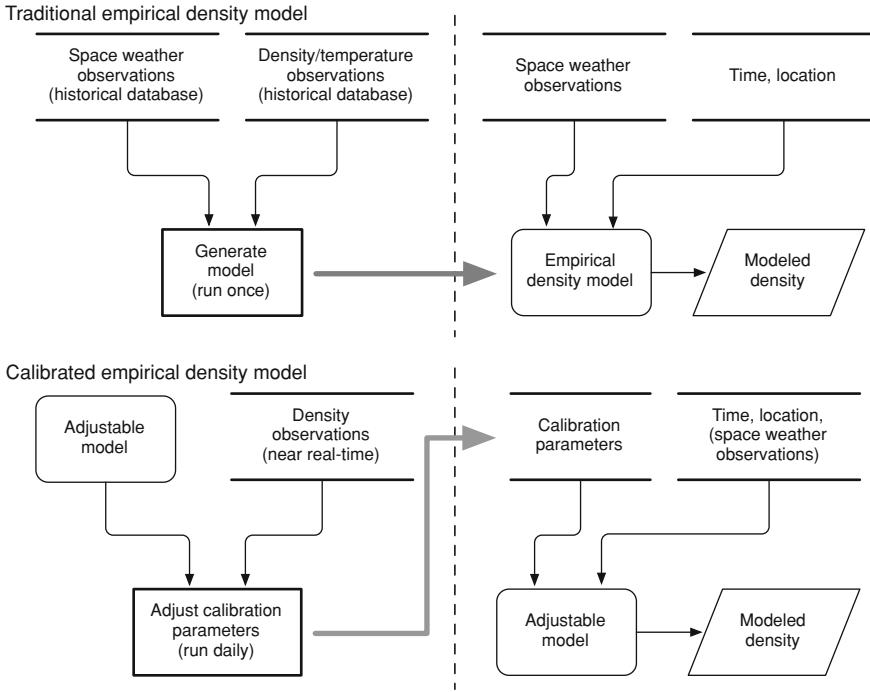
Fig. 2.10 Flow chart of traditional empirical density model creation and use (*top*), compared with that of a near real-time calibrated density model (*bottom*). The model generation or calibration procedure is displayed on the *left*, the use of the model in applications on the *right*

2.5.1 Approach

In order to calibrate an empirical density model, adjustable correction parameters that influence the model output have to be made available in the interface of the model's source code implementation. Optimal values for these correction parameters can be determined using a least-squares minimisation of density observation residuals. The estimated correction parameters can then correct or even replace the role of solar and geomagnetic activity indices in driving the irregular variations in the density model.

For use in operational applications, such as re-entry prediction and collision avoidance, the estimation of calibration parameters must be performed in near real-time, preferably at regular intervals, such as daily or 3-hourly. Figure 2.10 shows the flow chart of this process in the bottom panel, to be compared with the traditional one-time generation of a new version of an empirical model in the top panel.

Even though the original empirical model that is being adjusted still requires space weather proxies and indices as inputs, the calibration process should be able to absorb most of the shortcomings in the correlation between the proxies and indices and the density. The calibration parameters, estimated from satellite drag data, there-

fore largely replace the role of these space weather observations for the historical representation of the atmospheric state. Using this method, a significantly better level of accuracy than that of the uncalibrated empirical models can be reached.

The concept of empirical density model calibration was first published by Marcos et al. [62]. Shortly afterwards, two distinct density calibration initiatives were started by scientists in the space tracking community.

2.5.2 US-Russian Collaborative Density Calibration Project

A United States-Russian collaborative density calibration project involved the estimation of daily scale factors, varying linearly with altitude, which multiply the original model density to better represent orbit data. The analysis was performed for both the Russian GOST model [16] and the American NRLMSISE- 00 model [93]. Trajectory information for the calibration objects was taken from publicly available TLE data, although the efficient density-derivation algorithm by Picone et al. [77], discussed in Chap. 4, was not yet available to the researchers. Instead, the TLEs were processed into trajectories, which were used to estimate ballistic coefficients, containing information on the density model error.

A statistical analysis, investigating the possibility to predict correction parameters into the future has also been performed by this group [92]. As expected, the accuracy behaviour of predictions depends on the forecasting interval. While near-term results are significantly improved, forecasts over more than a few days can apparently no longer provide additional density accuracy over uncalibrated models.

2.5.3 The US Air Force Space Command HASDM Project

Another project, the United States Air Force Space Command's Dynamic Calibration Atmosphere (DCA) for the High-Accuracy Satellite Drag Model (HASDM) [87], used the tracking data from the Space Surveillance Network (SSN) directly, to adjust simultaneously the trajectories of around 75–80 calibration objects with spherical harmonic expansions of two temperature parameters from the Jacchia-70 thermosphere model [15]. Within Jacchia's models, these two temperatures define the vertical density profile completely. The spherical harmonic expansion allows corrections to the modelling of the diurnal variation. HASDM also includes a method for predicting the corrections three days into the future as a function of solar and geomagnetic indices.

A follow-on project, named Sapphire Dragon, was aimed at improving the prediction capabilities of HASDM through a series of enhancements, including an increase in the number of calibration objects and a more sophisticated use of various space weather proxies (*Bowman*, personal communication, 2006).

Unfortunately, the DCA calibrated model and the underlying precise SSN tracking data are not made publicly available. However, the HASDM model densities, computed along the CHAMP and GRACE tracks by Bruce Bowman, were kindly made available for evaluation. It is noteworthy that the density data derived from the SSN observations by Bowman et al. [8] have been used independently of HASDM as well. For instance, the data have been applied to study various modes of thermospheric forcing by Qian et al. [79–81].

2.5.4 ESA Funded Projects

Inspired by the first publications on the HASDM project, ESA's European Space Operations Centre ESOC funded a feasibility study [18, 26, 25] and software implementation [19, 20, 22] of thermosphere density model calibration using publicly available data. Although a number of different options were examined during the feasibility study, ultimately the choice was made to use TLE data, because of the near real-time data availability and the good spatial coverage if many objects are analysed simultaneously. The parameterisation schemes used by the US-Russian and HASDM projects were both implemented, adapted, and tested.

After the two ESOC projects were finalised, evaluation and improvement of the software have continued. First, the calibrated model was applied in the precise orbit determination of Earth observation satellites [21]. Later, the software was modified to use accelerometer-derived data from CHAMP and GRACE instead of TLE-derived data [23, 24].

Chapter 6 will cover the methods and a selection of the results from these investigations.

2.6 Tools for Evaluating Models and Observations

Further on in this thesis, the accuracy, resolution and other characteristics of thermosphere models and data sets will be evaluated. Several plotting and analysis methods that will aid in this endeavour are listed in the Sections below.

2.6.1 Plotting and Comparing Wide Ranges of Density Values

Density in the thermosphere ranges over many orders of magnitude, as a function of space and time (see Fig. 2.1). Sampling by different satellites in various orbits often results in a similarly wide-ranging data set. At first, this complicates the tasks of creating meaningful statistical comparisons and creating clear figures that retain visibility of interesting detail.

A straightforward and often applied solution is to use the logarithm of the densities to compress the range of values.

Another approach is to first convert the densities to equivalent exospheric temperatures, for example by inverting the CIRA- 72 algorithm for hydrostatic equilibrium. In this model, the exospheric temperature for a given epoch is modulated only by a fixed diurnal variation function. Therefore, temperatures derived for the same time, but at different locations from density observations of different satellites, can be compared with each other, and evaluated for consistency with this model. This straightforward link between temperature and density will form the basis of one of the density model calibration schemes, presented in Sect. 6.1.1. In a disturbed atmosphere, however, this equilibrium-based link between temperature and density can be too much of a simplification to allow for proper interpretation of the satellite data.

A third strategy to compress the range of values for further analysis is to normalise the density observations ρ_O by dividing them by equivalent model density values ρ_M, for the same time and location.

$$r = \frac{\rho_O}{\rho_M} \tag{2.14}$$

The resulting *density ratio* will generally fall in the range of about 0.5–2.0, depending on the combined error in the data and reference model. This much smaller range is suitable for incorporation in plots showing data from multiple satellites simultaneously, as will be done in Chap. 5. The ratios can also be used to either evaluate different density models using the same data set, or to identify outliers or erratic behaviour in contemporaneous data sets using the same model.

2.6.2 Density Evaluation Metrics

When evaluating density models, the deviation of the model with respect to an observation can be expressed as a percentage of the observation.

$$\epsilon = \frac{\rho_M - \rho_O}{\rho_O} \times 100\% \tag{2.15}$$

The Root-Mean-Square (RMS) of a set of these relative density error values will be used as a metric of density model accuracy. A possible disadvantage is that simple biases in the models or observations can not be distinguished from time-variable errors in this metric. It is therefore best suited when the mean with respect to the model has already been removed from the data. If this is not the case, a better approach is to use a two part metric, representing a mean and standard deviation. This leads us to the statistical distribution of density ratios.
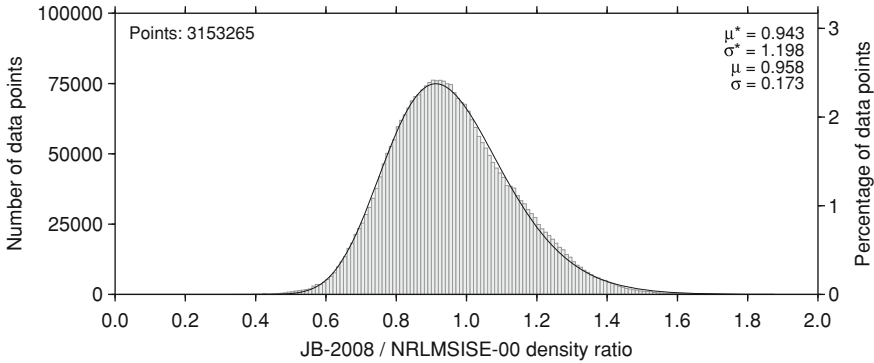
Fig. 2.11 Histogram showing the distribution of density ratios of two different models, evaluated along the CHAMP orbit during the year 2003. The log-normal distribution function, corresponding to the multiplicative mean μ^* and standard deviation σ^* of the data, is plotted with a *solid line*

2.6.3 Log-Normal Distribution of Density Ratios

Bezdèk [5] observed that the statistical relationship of the ratio of CACTUS accelerometer-derived densities and MSIS empirical model densities is consistent with the *log-normal distribution* [60]. This implies that the logarithm of the density ratios is approximately normal (Gaussian).

Figure 2.11 shows a histogram of density ratios for two empirical models. The data distribution is compared with an analytical log-normal distribution curve. Instead of the symmetrical bell-curve valid for Gaussian distributions, the log-normal distribution of the density ratios is skewed, with a shallower slope above the maximum and a steeper one below it. This is a consequence of the fact that the log-normal distribution is based on multiplication of random errors, while the Gaussian distribution is based on addition.

The Gaussian distribution is often used to characterise data sets with expressions such as $\mu \pm \sigma$, where the arithmetic mean μ and standard deviation σ are defining parameters of the distribution. These parameters can be used to express confidence intervals. There is an analogy for the log-normal distribution. Just as approximately 68.3% of the data lie within the mean, *plus or minus* one standard deviation, under a Gaussian bell-curve for normally distributed data, the same percentage of data lies within the log-normal mean μ^* *multiplied or divided by* the multiplicative standard deviation σ^* for log-normally distributed data. The minimum value of $\sigma^* = 1$ would indicate that the two density data sets are identical.

As we shall see in subsequent chapters, many (but not all) of the important error sources in deriving density data from satellite dynamics observations, such as the satellite frontal area and drag coefficient, are multiplicative in nature (see Sect. 3.2.1). Such errors scale the data, and thereby modify the log-normal mean. If the error is constant, the log-normal standard deviation will remain unaffected.

A comparison of density ratios using the Gaussian standard deviation can result in a misinterpretation of the results, because a scaling of the data set results in a modification of both the Gaussian mean and standard deviation.

2.6.4 Time Series, Binning and Gridding of Density Data

Further standard tools for analysing density models and observations, which will be used in subsequent chapters, include the display of time series, binning of data with respect to certain variables, computing statistics per bin and plotting the results in grids.

There are many variables which are interesting to use in plotting and/or binning density data or density ratios. We will discuss several important ones below. The reader is referred to Montenbruck and Gill [70] for additional details on orbital elements, time systems and coordinate systems.

In this thesis, time is expressed either with respect to the UTC or GPS systems. While UTC might be the preferred system for presentation of observation data, the CHAMP and GRACE data products are provided in GPS time. The difference between the two time systems, equal to the number of leap seconds introduced since January 1980, will be negligible in the plots.

Geodetic and geographic coordinates, including satellite altitudes, are computed with respect to the GRS 80 reference ellipsoid [71]. Geomagnetic coordinates are computed with respect to the IGRF dipole, based on the GEOPACK subroutine library by Dr. N. A. Tsyganenko.

As explained in Sect. 2.1.3, the local solar time is a measure of the longitudinal position of the satellite with respect to the sub-solar point. The related geocentric angle between the measurement location and the Sun can be useful in the analysis of density obtained near the perigee of elliptical orbits. It can be used to identify the position of the measurement with respect to the diurnal density bulge.

The *argument of latitude* ν is defined as the angle along the orbit from the ascending equator crossing. It is calculated by computing the Kepler elements of the osculating orbit, and summing the argument of perigee ω and true anomaly θ.

$$\nu = \omega + \theta \tag{2.16}$$

This variable is useful for plotting data from the CHAMP and GRACE accelerometer satellites, along their orbits. Values of $\nu = 0°$ and $\nu = 180°$ correspond to ascending and descending crossings of the equator. And since these missions have near polar orbit inclinations, values around $\nu = 90°$ and $\nu = 270°$ correspond to crossings of the North- and South-polar regions, respectively.

Finally, several analyses will show the relation of measurements with solar and geomagnetic activity levels, as represented by their proxies and indices described earlier in this chapter.

Examples of the analysis methods and plotting variables described above will be presented from Chap. 4 onwards. In the next chapter we will first focus on the

motion of the satellite in its orbit, and how it is affected by several forces, including the aerodynamic force.

References

1. Asikainen T, Maliniemi V, Mursula K (2010) Modeling the contribution of ring, tail and magnetopause currents to the corrected Dst index. J Geophys Res 115(A12203):16 doi:10.1029/2010JA015774
2. Barlier F, Berger C, Falin JL, Kockarts G, Thuillier G (1978) A thermospheric model based on satellite drag data. Ann Geophysique 34(1):9–24
3. Bartels J, Heck NH, Johnston HF (1939) Geomagnetic three-hour-range index measuring geomagnetic activity. J Geophys Res 44:411
4. Berger C, Biancale R, Ill M, Barlier F (1998) Improvement of the empirical thermospheric model DTM: DTM-94–a comparative review of various temporal variations and prospects in space geodesy applications. J Geod 72(3):161–178
5. Bezdèk A (2007) Lognormal distribution of the observed and modelled neutral thermospheric densities. Stud Geophys Geod 51(3):461–468 doi:10.1007/s11200-007-0027-2
6. Bowman BR, Tobiska WK (2006) Improvements in modelling thermospheric densities using new EUV and FUV solar indices. In: 16th AAS/AIAA space flight mechanics conference, Tampa Florida, pp 22–26, AAS 06-237, January 2006
7. Bowman BR (2004) The semiannual thermospheric density variation from 1970 to 2002 between 200–1100 km. In: AAS 04–174
8. Bowman BR, Marcos FA, Kendra MJ (2004) A method for computing accurate daily atmospheric density values from satellite drag data. In: 14th AAS/AIAA space flight mechanics conference, Maui Hawaii, February 8–12, 2004, AAS 04–173
9. Bowman BR, Tobiska WK, Marcos FA, Huang CY, Lin CS, Burke WJ (2008a) A new empirical thermospheric density model JB2008 using new solar and geomagnetic indices. In: AIAA/AAS Astrodynamics specialist conference and exhibit, pp 18–21 August 2008, Honolulu Hawaii, AIAA 2008–6438
10. Bowman BR, Tobiska WK, Marcos FA, Valladares C (2008) The JB2006 empirical thermospheric density model. J Atmos Sol Terr Phys 70(5):774–793 doi:10.1016/j.jastp.2007.10.002
11. Bruinsma S, Thuillier G, Barlier F (2003) The DTM-2000 empirical thermosphere model with new data assimilation and constraints at lower boundary accuracy and properties. J Atmos sol Terr Phys 65:1053–1070 doi:10.1016/S1364-6826(03)00137-8
12. Bruinsma S, Biancale R (2003) Total densities derived from accelerometer data. J Spacecraft Rockets 40(2):230–236
13. Cageao RP, Kerr RB (1984) Global distribution of helium in the upper atmosphere during solar minimum. Planet Space Sci 32(12):1523–1529 doi:10.1016/0032-0633(84)90019-9
14. Campbell WH (1996) Geomagnetic storms, the Dst ring-current myth and lognormal distributions. J Atmos Terr Phys 58(10):1171–1187 doi:10.1016/0021-916995001034
15. Casali SJ, Barker WN (2002) Dynamic calibration atmosphere (DCA) for the high accuracy satellite drag model (HASDM). In: AIAA/AAS astrodynamics specialist conference and exhibit, Monterey California, AIAA 2002–4888 5–8 August 2002
16. Cefola PJ, Nazarenko AI, Proulx RJ, Yurasov VS (2003) Atmospheric density correction using two line element sets as the observation data. In: AAS/AIAA astrodynamics specialists conference, Big Sky Montana, August 3–7, 2003, AAS 03–626
17. Domingo V, Fleck B, Poland AI (1995) The SOHO mission: an overview. Sol Phys 162 (1–2):1–37 doi:10.1007/BF00733425

18. Doornbos E (2004) Calibrated high accuracy satellite drag model - ESOC contract 16643/02/D/HK(SC) final report. Faculty of Aerospace Engineering Delft University of Technology, Delft The Netherlands
19. Doornbos E (2006) NRTDM final report - Near real-time density model (NRTDM) - ESOC contract 18576/04/D/HK(SC), Delft Institute for Earth-Oriented Space Research
20. Doornbos E (2007) Thermosphere density model calibration. In: Jean L (ed) Space weather, research towards applications in Europe, vol 344 of astrophysics and space science library, Springer
21. Doornbos E, Klinkrad H, Scharroo R, Visser P (2007) Thermosphere density model calibration in the orbit determination and prediction of ERS-2 and Envisat. In: Lacoste H (ed) Envisat symposium, Montreux, Switzerland, ESA SP-636, 23–27 April, 2007
22. Doornbos E, Klinkrad H, Visser P (2008) Use of two-line element data for thermosphere neutral density model calibration. Adv Space Res 41(7):1115–1122 doi:10.1016/j.asr.2006.12.025
23. Doornbos E, Förster M, Fritsche B, Van Helleputte T, Van den IJssel J, Koppenwallner G, Lühr H, Rees D, Visser P, Kern M (2009a) Air density models derived from multi-satellite drag observations. In: Proceedings of ESAs second swarm international science meeting, Potsdam Germany, 24–26 June 2009
24. Doornbos E, Förster M, Fritsche B, Van Helleputte T, Van den IJssel J, Koppenwallner G, Lühr H, Rees D, Visser P (2009b) ESTEC contract 21022/07/NL/HE Air density models derived from multi-satellite drag observations–final report, DEOS / TU Delft scientific report 01/2009, TU Delft
25. Doornbos EN, Klinkrad H (2006) Modelling of space weather effects on satellite drag. Adv Space Res 37:1229–1239 doi:10.1016/j.asr.2005.04.097
26. Doornbos EN, Klinkrad H, Visser PNAM (2005) Atmospheric density calibration using satellite drag observations. Adv Space Res 36:515–521 doi:10.1016/j.asr.2005.02.009
27. Drob DP, Emmert JT, Crowley G, Picone JM, Shepherd GG, Skinner W, Hays P, Niciejewski RJ, Larsen M, She CY, Meriwether JW, Hernandez G, Jarvis MJ, Sipler DP, Tepley CA, O'Brien MS, Bowman JR, Wu Q, Murayama Y, Kawamura S, Reid IM, Vincent RA (2008) An empirical model of the earth's horizontal wind fields: HWM07. J Geophys Res 113(A12304):18 doi:10.1029/2008JA013668
28. Dudok de wit T, Lilen sten J, Aboudarham J, Amblard P.O, Kretzchmar M (2005) Retrieving the solar EUV spectrum from a reduced set of spectral lines. Ann Geophys 23(9):3055–3069
29. Emmert JT, Picone JM (2010) Climatology of globally averaged thermospheric mass density. J Geophys Res 115(A09326):17 doi:10.1029/2010JA015298
30. Emmert JT, Drob DP, Shepherd GG, Hernandez G, Jarvis MJ, Meriwether JW, Niciejewski RJ, Sipler DP, Tepley CA (2008) DWM07 global empirical model of upper thermospheric storm-induced disturbance winds. J Geophys Res 113(A11319):113. doi:10.1029/2008JA013541
31. Escoubet CP, Schmidt R, Goldstein ML (1997) Cluster - science and mission overview. Space Sci Rev 79(1–2):11–32 doi:10.1023/A:1004923124586
32. Floyd L, Newmark J, Cook J, Herring L, McMullin D (2005) Solar EUV and UV spectral irradiances and solar indices. J Atmos Sol Terr Phys 3(15). doi:10.1016/j.jastp.2004.07.013
33. Forbes JM, Gonzalez R, Marcos FA, Revelle D, Parish H (1996) Magnetic storm response of lower thermosphere density. J Geophys Res 101(A2):2313–2319 doi:10.1029/95JA02721
34. Fuller-Rowell TJ (1998) The 'thermospheric spoon': a mechanism for the semiannual density variation. J Geophys Res 103(A3):3951–3956
35. Fuller-Rowell TJ, Codrescu MV, Moffett RJ, Quegan S (1994) Response of the thermosphere and ionosphere to geomagnetic storms. J Geophys Res 99(A3):3893–3914 doi:10.1029/93JA02015
36. Guo J, Wan W, Forbes JM, Sutton E, Nerem RS, Woods TN, Bruinsma S, Liu L (2007) Effects of solar variability on thermosphere density from CHAMP accelerometer data. J Geophys Res 112(A10308):13 doi:10.1029/2007JA012409

37. Hargreaves JK (1992) The solar-terrestrial environment: an introduction to geospace - the science of the terrestrial upper atmosphere ionosphere and magnetosphere vol 7 of cambridge atmospheric and space science series. Cambridge University Press, Cambridge
38. Harris I, Priester W (1962) Time-dependent structure of the upper atmosphere. J Atmos Sci 19(4):286–301
39. Heath DF, Schlesinger BM (1986) Mg 280-nm doublet as a monitor of changes in solar ultra-violet irradiance. J Geophys Res 91(D8):8672–8682
40. Hedin AE (1983) A revised thermospheric model based on mass spectrometer and incoherent scatter data - MSIS-83. J Geophys Res 88:10170–10188
41. Hedin AE (1988) The atmospheric model in the region 90 to 2000 km, Advances in Space Research, 8(5–6):(5)19–(5)25. In: COSPAR international reference atmosphere: 1986, Part I: Thermosphere models
42. Hedin AE (1991) Extension of the MSIS thermospheric model into the middle and lower atmosphere. J Geophys Res 96(A2):1159–1172
43. Hedin AE, Reber CA, Newton GP, Spencer NW, Salah JE, Evans JV, Kayser DC, Alcayde D, Bauer P, Cogger L (1977) A global thermospheric model based on mass spectrometer and incoherent scatter data MSIS. i - N2 density and temperature. J Geophys Res 82:2139–2147
44. Hedin AE, Reber CA, Newton GP, Spencer NW, Brinton HC, Mayr HG, Potter WE (1977) A global thermospheric model based on mass spectrometer and incoherent scatter data MSIS. II - composition. J Geophy Res 82:2148–2156
45. Hedin AE, Fleming EL, Manson AH, Schmidlin FJ, Avery SK, Clark RR, Franke SJ, Fraser GJ, Tsuda T, Vial F, Vincent RA (1996) Empirical wind model for the upper, middle and lower atmosphere. J Atmos Terr Phys 58(13):1421–1447 doi:10.1016/0021-916995001220
46. Hedin Alan E (1987) MSIS-86 thermospheric model. J Geophys Res 92(A5):4649–4662
47. Hedin Alan E et al (1988) Empirical global model of upper thermosphere winds based on atmosphere and dynamics explorer satellite data. J Geophys Res 93:9959–9978
48. Hedin Alan E et al (1991) Revised global model of thermosphere winds using satellite and groundbased observations. J Geophys Res 96(A5):7657–7688
49. Hickey MP (1988) The NASA Marshall engineering thermosphere model. NASA Marshall Space Flight Center, Alabama
50. Jacchia LG (1959) Two atmospheric effects in the orbital acceleration of artificial satellites. Nature 183(4660):526–527
51. Jacchia LG (1964) Static diffusion models of the upper atmosphere with empirical temperature profiles. Smithsonian astrophysical observatory special report 170
52. Jacchia LG (1971a) New static models of the thermosphere and exosphere with empirical temperature profiles. Smithsonian Astrophysical Observatory Special Report 313
53. Jacchia LG (1971b) Revised static models of the thermosphere and exosphere with empirical temperature profiles, Smithsonian Astrophysical Observatory Special Report 332
54. Jacchia LG (1972) Atmospheric models in the region from 110 to 2000 km. In: CIRA 1972: COSPAR International Reference Atmosphere. Akademie-Verlag, Berlin, pp 227–338
55. Jacchia LG (1977) Thermospheric temperature, density and composition: new models, Smithsonian Astrophysical Observatory Special Report 375
56. Kaiser ML (2005) The STEREO mission an overview. Adv Space Res 36(8):1438–1488 doi:10.1016/j.asr.2004.12.066
57. King-Hele D (1992) A tapestry of orbits. Cambridge University Press, Cambridge
58. Knipp DJ, Tobiska WK, Emery BA (2004) Direct and indirect thermospheric heating sources for solar cycles 21–23. Sol Phys 224(1–2):495–505 doi:10.1007/s11207-005-6393-4
59. Lean J (1991) Variations in the sun's radiative output. Rev Geophys 29(4):505–535 doi:10.1029/91RG01895
60. Limpert E, Stahel WA, Abbt M (2001) Log-normal distributions across the sciences: keys and clues. BioScience 51(5):341–352
61. Marcos FA (1990) Accuracy of atmospheric drag models at low satellite altitudes. Adv Space Res 10(3):417–422

62. Marcos FA, Kendra MJ, Griffin JM, Bass JN, Larson DR, Liu JJ (1998) Precision low earth orbit determination using atmospheric density callibration. J Astronaut Sci 46(4):395–409

63. Mayaud PN (1980) Derivation, meaning, and use of geomagnetic indices geophysical monograph 22. American Geophysical Union, Washington

64. Mayr HG, Harris I, Spencer NW (1978) Some properties of upper atmosphere dynamics. Rev Geophys Space Phys 16:539–565

65. McPherron RL, Baker DN (1993) Factors influencing the intensity of magnetospheric substorms. J Atmos Terr Phys 55(8):1091–1122 doi:10.1016/0021-9169(93)90040-6

66. Menvielle M, Berthelier A (1991) The K-derived planetary indices: description and availability. Rev Geophys 29(3):415–432

67. Menvielle M, Marchaudon A Geomagnetic indices in solar-terrestrial physics and space weather. In: Jean L (eds) Space weather research towards applications in Europe vol 344. Springer, Astrophysics and Space Science Library (2007)

68. Menvielle M, Paris J (2001) The aλ longitude sector geomagnetic indices. Contrib Geophys Geod 31(1):315–322

69. Moldwin M (2008) An introduction to space weather. Cambridge University Press, Cambridge

70. Montenbruk O, Gill E (2000) Satellite orbits - models, methods and applications. Springer-Verlag, Berlin

71. Moritz H (1980) Geodetic reference system 1980, Bulletin Geodesique

72. Owens JK (2002) NASA Marshall engineering thermosphere model - version 2 0 NASA/TM–2002–211786. Marshall Space Flight Center, Alabama

73. Paetzold HK, Zschorner H (1961) An annual and a semiannual variation of the upper air density. Pure Appl Geophys 48(1):85–92

74. de Pater I, Lissauer J (2001) Planetary sciences. Cambridge University Press, Cambridge

75. Paxton LJ, Christensen AB, Humm DC, Ogorzalek BS, Pardoe CT, Morrison D, Weiss MB, Crain W, Lew PH, Mabry DJ, Goldsten JO, Gary SA, Persons DF, Harold MJ, Alvarez B, Ercol CJ, Strickland DJ, Meng C-I (1999) Global ultraviolet imager (GUVI): measuring composition and energy inputs for the NASA thermosphere ionosphere mesosphere energetics and dynamics (TIMED) mission. Opt Spectrosc Tech Instrum Atmos Space Res III 3756(1): 265–276 doi:10.1117/12.366380

76. Picone JM, Hedin AE, Drob DP, Aikin AC (2002) NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues. J Geophys Res 107(A12):1468. doi:10.1029/2002JA009430

77. Picone JM, Emmert JT, Lean J (2005) Thermospheric densities derived from spacecraft orbits - I. Accurate processing of two-line element sets. J Geophys Res 110(A03301):19. doi:10.1029/2004JA010585

78. Prölss GW (2004) Physics of the Earth's Space Environment. Springer

79. Qian L, Solomon SC, Roble RG, Bowman BR, Marcos FA (2008) Thermospheric neutral density response to solar forcing. Adv Space Res 42(5):926–932 doi:10.1016/j.asr.2007.10.019

80. Qian L, Solomon SC, Kane TJ (2009) Seasonal variation of thermospheric density and composition. J Geophys Res 114(A01312). doi: 10.1029/2008JA013643

81. Qian L, Solomon SC, Mlynczak MG (2010) Model simulation of thermospheric response to recurrent geomagnetic forcing. J Geophys Res 115(A10301). doi: 10.1029/2010JA015309

82. Reber CA, Trevathan CE, McNeal RJ, Luther MR (1993) The upper atmospheric research satellite (UARS) mission. J Geophys Res 98(D6):10643–0647

83. Rees MH (1989) Physics and chemistry of the upper atmosphere cambridge atmospheric and space science series. Cambridge University Press, Cambridge

84. Roberts CE (1971) An analytical model for upper atmosphere densities based upon jacchia's 1970 models. Celestial Mech 4:368–377

85. Rottman G (2005) The SORCE mission. Sol Phys 230(1–2):7–25, doi:10.1007/s11207-005-8112-6

86. Sehnal L (1988) Thermospheric total density model TD. Bull Astron Institutes Czechoslovakia 39(2):120–127

87. Storz MF, Bowman BR, Branson JI, Casali SJ, Tobiska WK (2005) High accuracy satellite drag model (HASDM). Adv Space Res 36(12):2497–2505
88. Thuillier G, Bruinsma S (2001) The Mg II index for upper atmosphere modelling. Ann Geophys 19(2):219–228
89. Tobiska WK, Bouwer SD, Bowman BR (2006) The development of new solar indices for use in thermospheric density modeling. In: AIAA, AIAA 2006–6165
90. Troschichev O, Janzhura A, Stauning P (2006) Unified PCN and PCS indices: method of calculation, physical sense, and dependence on the IMF azimuthal and northward components. J Geophys Res 111(A05208):10. doi:10.1029/2005JA011402
91. Vallado DA (2001) Fundamentals of strodynamics and Applications. Second Edition: Microcosm Press, El Segundo, California
92. Yurasov VS, Nazarenko AI, Cefola PJ, Alfriend KT(2005a) Application of the ARIMA model to analyze and forecast the time series of density corrections for NRLMSIS-00. In: AAS/AIAA Astrodynamics specialist conference, Lake Tahoe, CA, August 7–11, 2005, AAS 05–256
93. Yurasov VS, Nazarenko AI, Cefola PJ, Alfriend KT (2005b) Density corrections for the NRLMSIS-00 atmosphere model. In: AAS/AIAA Space flight mechanics conference, Copper Mountain, Colorado, January 23–27, 2005, AAS 05–168
94. Zhang J, Richardson IG, Webb DF, Gopalswamy N, Huttunen E, Kasper JC, Nitta NV, Poomvises W, Thompson BJ, Wu C-C, Yashiro S, Zhukov AN (2007) Solar and interplanetary sources of major geomagnetic storms (Dst -100nT) during 1996–2005. J Geophys Res 112(A10102). doi:10.1029/2007JA012321

Chapter 3
Satellite Dynamics and Non-Gravitational Force Modelling

This chapter will discuss the link between the motion of satellites and thermospheric density and wind. Section 3.1 will start with a brief overview of relevant theory on satellite dynamics. Section 3.2 describes the basic equations used to model aerodynamic and radiation pressure accelerations on the satellite, as well as some often-used approximations. The influence of these accelerations on the shape of the orbit is derived using the perturbation equations in Sect. 3.3. The final part of this chapter describes the detailed force modelling of aerodynamics and radiation pressure for complex spacecraft, with Sect. 3.4 covering particle-surface interactions for flat panels and Sect. 3.5 acceleration computations for complex satellite geometries, combining many such panels.

3.1 Equations of Motion

Many textbooks on astrodynamics are available that provide the derivation of Kepler's laws from the equations of motion of the two-body problem (e.g., [32, 40, 41]). The discussion in this Section will be rather coarse compared to those texts, and will be strictly limited to the theory that is required to understand the processing algorithms presented later on in this thesis.

At first approximation, all satellite orbits are solutions of the equations of motion of the two-body problem in an inertial reference frame, in which the only acceleration on the satellite is the gravitational attraction by the Earth, which is assumed to be spherically symmetric.

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^2}\hat{\boldsymbol{r}} \qquad (3.1)$$

In this equation, $\boldsymbol{r}$ is the inertial position vector of the satellite and μ is the Earth's gravitational parameter, equal to $GM_\oplus$, the product of the gravitational constant and the mass of the Earth. The right-hand side of this equation simply represents Newton's law of gravitation, where the mass of the satellite is considered negligible
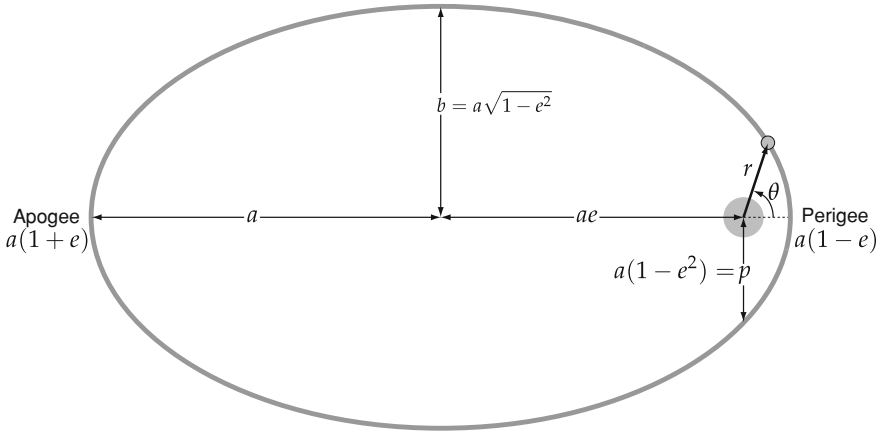
Fig. 3.1 Geometrical parameters of an elliptical satellite orbit

compared to the mass of the Earth. Manipulation of this equation, provided in any of the above-mentioned astrodynamics textbooks, reveals that the orbit obeys the equation of a conic section in polar coordinates r and θ, described by the parameters representing the semi-major axis a and eccentricity e.

$$r = \frac{a(1 - e^2)}{1 + e \cos \theta} \tag{3.2}$$

We will be concerned only with artificial Earth satellites for which $0 \leq e < 1$, so that Eq. (3.2) describes an elliptical orbit. Figure 3.1 shows the geometrical definition of the parameters of Eq. (3.2).

Kepler's third law follows after further manipulation of Eq. (3.1). It provides the relation between the semi-major axis a and the orbital period P for a given central body with gravitational parameter μ.

$$P = 2\pi \sqrt{\frac{a^3}{\mu}} \tag{3.3}$$

A closely related parameter, which will be useful further on, is the mean motion n, which expresses the number of orbital revolutions per unit of time.

$$n = \frac{2\pi}{P} = \sqrt{\frac{\mu}{a^3}} \tag{3.4}$$

The representation of the equations of motion in Eq. (3.1) is only valid when the Earth's gravity field is considered spherical-symmetric and there are no other forces acting on the satellite. In reality, the Earth's shape and its gravity field deviate from this spherical ideal, and there are various other forces which perturb the orbit.

Fortunately, these deviations are relatively small compared to the acceleration due to the spherical-symmetric component of the Earth's gravity field, and therefore the above equations are often used as first order approximations for true satellite orbits.

A more realistic representation of the equations of motion would be

$$\ddot{\boldsymbol{r}} = \ddot{\boldsymbol{r}}_{\mathrm{g}} + \ddot{\boldsymbol{r}}_{\mathrm{ng}} = -\frac{\mu}{r^2}\hat{\boldsymbol{r}} + \ddot{\boldsymbol{r}}_{\mathrm{eg}} + \ddot{\boldsymbol{r}}_{\mathrm{3b}} + \ddot{\boldsymbol{r}}_{\mathrm{gr}} + \ddot{\boldsymbol{r}}_{\mathrm{a}} + \ddot{\boldsymbol{r}}_{\mathrm{srp}} + \ddot{\boldsymbol{r}}_{\mathrm{erp}} + \ddot{\boldsymbol{r}}_{\mathrm{other}} \qquad (3.5)$$

where the additional acceleration terms on the right hand side can be grouped in two parts, the first of which represents gravitational accelerations ($\ddot{\boldsymbol{r}}_{\mathrm{g}}$). These include the central body gravitational attraction term, the effects of the irregularities of the Earth's gravity field ($\ddot{\boldsymbol{r}}_{\mathrm{eg}}$), the gravitational effects of third bodies such as the Sun, Moon and planets ($\ddot{\boldsymbol{r}}_{\mathrm{3b}}$) and the effects of general relativity ($\ddot{\boldsymbol{r}}_{\mathrm{gr}}$). The second group represents non-gravitational accelerations ($\ddot{\boldsymbol{r}}_{\mathrm{ng}}$), which are usually due to interactions between satellite surfaces and particles (photons, molecules, atoms) in its environment. These include aerodynamic accelerations ($\ddot{\boldsymbol{r}}_{\mathrm{a}}$), solar radiation pressure accelerations ($\ddot{\boldsymbol{r}}_{\mathrm{srp}}$), Earth albedo and infrared radiation pressure accelerations ($\ddot{\boldsymbol{r}}_{\mathrm{erp}}$) and other accelerations ($\ddot{\boldsymbol{r}}_{\mathrm{other}}$), for example due to thrusters, impacts with micro-meteoroids, electromagnetic interaction of charged spacecraft parts with the Earth's magnetic field, radiation emitted by the spacecraft due to heat or radio instrumentation, etc.

Computer models for many of these accelerations are provided in textbooks on orbit determination, such as Montenbruck and Gill [27] and Tapley et al. [39]. The aerodynamic and radiation pressure accelerations are crucial enough for our purposes to be described in more detail later on in this chapter. First, the aerodynamic acceleration and various simplifications that can be used in its computation will be introduced in more general terms in the next Section.

3.2 Basics of Non-Gravitational Acceleration Modelling

The aerodynamic acceleration $\ddot{\boldsymbol{r}}_{\mathrm{a}}$ is dependent on the level of density and velocity of winds in the thermosphere, and therefore provides the key for deriving information on these properties from satellite dynamics analysis. The radiation pressure acceleration $\ddot{\boldsymbol{r}}_{\mathrm{r}} = \ddot{\boldsymbol{r}}_{\mathrm{srp}} + \ddot{\boldsymbol{r}}_{\mathrm{erp}}$ is the most important perturbation that must be subtracted from observations of the total non-gravitational acceleration, to arrive at the aerodynamic acceleration observation. The modelling of both non-gravitational accelerations is quite similar, as will be described below.

3.2.1 Aerodynamic Acceleration and Drag Approximations

We will start by providing a full vector expression for the aerodynamic acceleration, and then afterwards derive the more widely used scalar approximations for the drag component, and its projection on the orbit plane.

Aerodynamicists customarily make use of dimensionless force coefficients in their calculations (e.g., [3]). We can define the vector

$$C_\text{a} \equiv \frac{F_\text{a}}{A_\text{ref} \frac{1}{2} \rho v_r^2} \tag{3.6}$$

by dividing the aerodynamic force F_a by a reference area A_ref and the dynamic pressure. The dynamic pressure is defined as $\frac{1}{2} \rho v_r^2$, where ρ is the atmospheric density and v_r is the magnitude of the velocity of the atmosphere, relative to the spacecraft body. The factor $\frac{1}{2}$ has no real significance for aerodynamics at satellite velocities, but is kept in this equation for compatibility with low-speed aerodynamics. In the context of satellite orbits, equations are usually expressed in terms of accelerations instead of forces, so that (3.6) is written as

$$C_\text{a} \equiv \frac{m \ddot{r}_\text{a}}{A_\text{ref} \frac{1}{2} \rho v_r^2} \tag{3.7}$$

This dimensionless force coefficient vector C_a is dependent on the geometry and orientation of the satellite shape with respect to the flow, the properties of the atoms and molecules that hit its surface, and the mode of interaction between these particles and the surface. For lower velocities, values for C_a can often be computed from observation data, by measuring or even controlling the quantities on the right-hand side of Eq. (3.6) in experiments. Unfortunately though, it is not possible to simulate the conditions of a satellite moving through the upper atmosphere in a wind tunnel or other type of laboratory facility. It will therefore be necessary to model C_a. The theory behind this modelling was widely investigated in the early years of the space age (e.g., [12, 34]), and due to the availability of additional data sets and increased computer processing power, has been met with renewed interest in recent years (e.g., [7, 15, 25, 31, 36]). The approach adopted for this thesis will be discussed in more detail in Sect. 3.4.

Our starting point for aerodynamic analysis will be the vector expression for the aerodynamic acceleration, derived directly from Eq. (3.7):

$$\ddot{r}_\text{a} = C_\text{a} \frac{A_\text{ref}}{m} \frac{1}{2} \rho v_\text{r}^2 \tag{3.8}$$

Projection of the aerodynamic acceleration on the relative velocity: Drag

Due to the nature of the interaction of the atmospheric particles with the satellite surfaces, the aerodynamic force acts mainly in the direction of the relative velocity, as will be demonstrated in Sect. 3.4. Furthermore, according to the perturbation equations (Sect. 3.3), the perpendicular force components (lift and side forces) are significantly less effective in changing the orbit geometry, compared to the along-track force component. The perpendicular components can therefore be safely ignored for many applications, and might even be zero due to symmetry, such as for spherical satellites. The aerodynamic acceleration is therefore often approximated by the drag

acceleration, which is defined as its projection on the relative velocity direction $\hat{\boldsymbol{v}}_r$, where the accent in the notation denotes the use of a unit vector:

$$\ddot{\boldsymbol{r}}_a \approx (\ddot{\boldsymbol{r}}_a \cdot \hat{\boldsymbol{v}}_r)\hat{\boldsymbol{v}}_r \equiv \ddot{\boldsymbol{r}}_D \tag{3.9}$$

Equation (3.8) then simplifies to

$$\ddot{\boldsymbol{r}}_D = C_D \frac{A_{\text{ref}}}{m} \frac{1}{2} \rho v_r^2 \hat{\boldsymbol{v}}_r \tag{3.10}$$

where the subscript D stands for drag. Note that while $\boldsymbol{C}_a$ in Eq. (3.8) is a vector, the drag coefficient

$$C_D \equiv \boldsymbol{C}_a \cdot \hat{\boldsymbol{v}}_r \tag{3.11}$$

is a scalar quantity, which can simplify analyses considerably.

Equation (3.10) is a reasonable simplification, and can be safely applied for most applications involving orbit determination or density derivation. Only in the precise processing of density and crosswind data from accelerometer observations, as explained in Sect. 4.2, will we require the additional effort of computing the small aerodynamic force components perpendicular to drag, according to Eq. (3.8).

Ballistic coefficient and the area to mass ratio

The ballistic coefficient it b, or its inverse B, defined by

$$B = \frac{1}{b} = C_D \frac{A_{\text{ref}}}{m} \tag{3.12}$$

can be substituted to reduce the number of symbols in equations. This is especially useful when the quantities on the right-hand side of (3.12) can be considered constants. Whether this is the case depends on the required accuracy of course, but is otherwise largely determined by the shape and attitude of the satellite, and to a lesser extent by the variation of temperature and composition along the orbit. These dependencies will be further discussed later on in this chapter.

The ballistic coefficient can be considered as a measure of the sensitivity of a space object to perturbations by the drag force. Satellites with a large area and a relatively small mass, such as balloon satellites, will experience a high drag. Because of the square-cube law, smaller satellites will have a larger area to mass ratio and will be more sensitive to drag accelerations than large satellites that are otherwise similarly constructed. Examples of such smaller, more drag-sensitive satellites are nano-, micro- and minisatellites, but also space debris fragments which have resulted from an on-orbit fragmentation of a larger satellite or rocket stage.

Relative velocity components

The relative velocity is the sum of contributions from the inertial velocity of the spacecraft in its orbit $\boldsymbol{v}_{r,i}$, the velocity caused by the corotating atmosphere $\boldsymbol{v}_{r,c}$, and the velocity of winds $\boldsymbol{v}_{r,w}$, with respect to an Earth-fixed atmosphere.

$$v_\mathrm{r} = v_\mathrm{r,i} + v_\mathrm{r,c} + v_\mathrm{r,w} \tag{3.13}$$

Note that in this thesis, to avoid unnecessary minus signs in aerodynamic calculations, relative velocities, denoted by the subscript r, will always refer to the velocity of the atmosphere with respect to the spacecraft. Only if the subscript r is omitted from a velocity variable, its opposite direction can be intended. An obvious example is the inertial satellite velocity

$$v = \dot{r} = -v_\mathrm{r,i}. \tag{3.14}$$

If required, an approximation of the wind with respect to the corotating atmosphere can be obtained from a model such as HWM07 (see Sect. 2.3.3). The corotation velocity can be easily and accurately computed from the Earth's angular rotation vector $\omega_\oplus = (0, 0, 0.7292115 \cdot 10^{-4})^T$ rad/s [17], and the satellite's position vector r.

$$v_\mathrm{r,c} = \omega_\oplus \times r \tag{3.15}$$

Projection of drag on the inertial velocity

A further simplification of Eq. (3.10) that will be useful is its projection on the inertial velocity, so that we will end up with a drag acceleration that is in its opposite direction. A corotation and wind correction factor F was introduced by King-Hele ([18], Sect. 2.5) and used by Picone et al. [29], so that

$$F v_\mathrm{r,i}^2 = v_\mathrm{r}^2 (\hat{v}_\mathrm{r} \cdot \hat{v}_\mathrm{r,i}) \tag{3.16}$$

resulting in

$$F = \frac{v_\mathrm{r}^2}{v_\mathrm{r,i}^2} (\hat{v}_\mathrm{r} \cdot \hat{v}_\mathrm{r,i}) \tag{3.17}$$

Making use of Eq. (3.14) the drag acceleration (3.10) can then be expressed as

$$\ddot{r}_\mathrm{D} \approx -C_\mathrm{D} \frac{A_\mathrm{ref}}{m} \frac{1}{2} \rho F v^2 \hat{v} \tag{3.18}$$

The benefit of this approximation is that the drag acceleration vector can be considered to lie completely in the orbital plane. This allows us to apply the in-plane perturbation equations from celestial mechanics, which will be discussed in Sect. 3.3.

3.2.2 Radiation Pressure Accelerations

A radiation pressure force coefficient vector can be defined, analogous to the aerodynamic force coefficient vector, by dividing the force by a reference area and pressure.

This results in the following basic expressions for the radiation pressure force coefficient and acceleration, analogous with the aerodynamics Eqs. (3.7) and (3.8)

$$C_{\mathrm{r}} \equiv \frac{m\ddot{\boldsymbol{r}}_{\mathrm{srp}}}{A_{\mathrm{ref}}\,P} \tag{3.19}$$

$$\ddot{\boldsymbol{r}}_{\mathrm{srp}} = C_{\mathrm{r}}\frac{A_{\mathrm{ref}}}{m}P. \tag{3.20}$$

The radiation pressure P is defined as

$$P = \frac{\Phi}{c} = \frac{\Delta E}{A\,\Delta t\,c}, \tag{3.21}$$

in which Φ is the energy flux and c the speed of light. In turn, the energy flux is defined as the amount of energy ΔE, in this case carried by photons, that passes through an area A during a time interval Δt.

For a satellite in orbit around the Earth, the Sun is of course the primary source of radiation, resulting in a pressure of $P \approx 4.56 \times 10^{-6}$ N/m^2 at the mean Sun-Earth distance. The precise value of P should be computed, as discussed in more detail further on, by taking into account the varying distance between the satellite and the Sun, as well as eclipses by the Earth and Moon, and refraction and absorption of sunlight by the Earth's atmosphere.

The Earth's surface and atmosphere form an important secondary source of radiation, both due to reflection (albedo) of sunlight and thermal (infrared) re-radiation of absorbed solar energy. The modelling of these secondary radiation pressure effects, as described in Doornbos et al. ([13], Sect. 5.2.8) will not be repeated in this thesis. They have little effect on density and wind derivation results, because the Earth radiation pressure accelerations are relatively small and act mainly in the radial direction, which is usually close to perpendicular to the along-track aerodynamic acceleration.

Projection of the radiation pressure acceleration on the sun-satellite vector

A useful approximation to Eq. (3.20), analogous to the drag Eq. (3.10), is to assume that the radiation pressure acceleration is in the direction of the vector $\boldsymbol{r}_{\odot s}$ from the light source, in this case the Sun, to the satellite.

$$\ddot{\boldsymbol{r}}_{\mathrm{r}} \approx (\ddot{\boldsymbol{r}}_{\mathrm{r}} \cdot \hat{\boldsymbol{r}}_{\odot s})\hat{\boldsymbol{r}}_{\odot s} \tag{3.22}$$

Equation (3.20) then simplifies to

$$\ddot{\boldsymbol{r}}_{\mathrm{r}} = C_{\mathrm{R}}\frac{A_{\mathrm{ref}}}{m}P\hat{\boldsymbol{r}}_{\odot s}, \tag{3.23}$$

where the scalar version of the radiation pressure coefficient is defined as

$$C_{\mathrm{R}} \equiv C_{\mathrm{r}} \cdot \hat{\boldsymbol{r}}_{\odot s}. \tag{3.24}$$

An inverse ballistic coefficient for radiation pressure can be defined as

$$B_R = C_R \frac{A_{\text{ref}}}{m}. \tag{3.25}$$

Eclipses

For a satellite in a low-Earth orbit, the Sun can often be eclipsed by the Earth. Less often, the Moon can cause eclipses as well. In order to model the influence of eclipses on the radiation pressure, a shadow function f_s is introduced. The shadow function is defined to have a value of one when the satellite is in full sunlight, and zero when the satellite is fully eclipsed. Because the Sun forms a solid disc of about 0.5°, the Earth and Moon cast a semi shadow region, or penumbra around the core shadow, or umbra. The discussion below will explain the calculation procedure for eclipses by the Earth. The calculation for eclipses by the Moon can of course be made in much the same way by substituting lunar coordinates for the Earth-centred coordinates, and by disregarding the atmospheric effects on the eclipse.

Different algorithms exist in order to determine whether a satellite is in eclipse [2, 27, 41]. An especially elaborate treatment of the problem can be found in the series of papers by Vokrouhlický et al. [42, 43, 45]. The method outlined below is provided in the ANGARA level 2 software [14]. It can be used to describe atmospheric effects on the semi-shadow boundaries. The basis of this method is a conical shadow model, where the Sun's disc is projected on an Earth radius vector which is perpendicular to the Sun-satellite line (see Fig. 3.2).

The subscript symbols $\oplus$ and $\odot$ are used to indicate positions with respect to the centres of the Earth and Sun, respectively, while the subscript s can be used to denote the satellite position if this is not already obvious. Let $\boldsymbol{r}$ and $\boldsymbol{r}_{\odot s}$ then denote the position vectors of the satellite with respect to the centres of the Earth and Sun.

The vector

$$\boldsymbol{r}_{ps} = \left(\hat{\boldsymbol{r}}_{\odot s} \cdot \boldsymbol{r}_{\oplus s} \right) \hat{\boldsymbol{r}}_{\odot s} \tag{3.26}$$

then defines the point p, which is the closest point to the Earth's centre on the Sun-satellite vector. This vector points from p to the satellite s in Fig. 3.2. The vector from the Earth's centre to p is

$$\boldsymbol{r}_{\oplus p} = \boldsymbol{r}_{\oplus s} - \boldsymbol{r}_{ps} \tag{3.27}$$

The height above the Earth's surface of a line connecting the center of the Sun with the satellite is then equal to

$$h_g = ||\boldsymbol{r}_{\oplus p}|| - R_\oplus \tag{3.28}$$

and the apparent radius of the solar disc projected on a plane through this point, perpendicular to the satellite-Sun vector, as indicated in Fig. 3.2, is
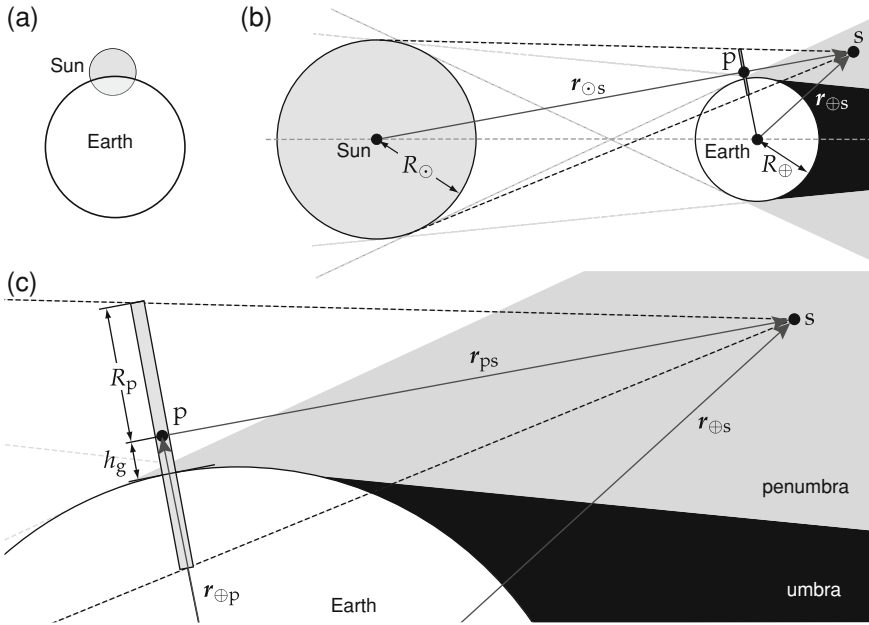
Fig. 3.2 The conical eclipse model, showing (**a**) a schematic view from the satellite, (**b**), (**c**) the geometric relationship between the Earth, Sun, the satellite at point s, and the projection of the Sun-satellite rays at the Earth closest point at point p

$$R_{\mathrm{p}} = \frac{|\boldsymbol{r}_{\mathrm{ps}}|}{|\boldsymbol{r}_{\odot\mathrm{s}}|} R_{\odot} \qquad (3.29)$$

In these equations, $R_{\oplus}$ and $R_{\odot}$ are the radii of the Earth and Sun, respectively. In the currently implemented version of the algorithm, these radii are considered constants. An extension which takes the oblateness of the Earth into account should improve the accuracy of the calculations [2, 45].

The geometric heights above the Earth's surface of the centre, top and bottom ray of light from the solar disc are $h_{\mathrm{c}} = h_{\mathrm{g}}$, $h_{\mathrm{t}} = h_{\mathrm{g}} + R_{\mathrm{p}}$ and $h_{\mathrm{b}} = h_{\mathrm{g}} - R_{\mathrm{p}}$, respectively. A negative value indicates that the beam is blocked by the Earth (see Fig. 3.2). The apparent height of these rays, as seen from the satellite will differ from these values, due to refraction in the atmosphere. As shown in Fig. 3.3, refraction will lead to a narrowing of the full eclipse cone. An elaborate discussion can be found in Vokrouhlický et al. [44].

An auxiliary parameter η can be introduced such that

$$\eta = \frac{h_c}{h_c - h_b}, \qquad (3.30)$$
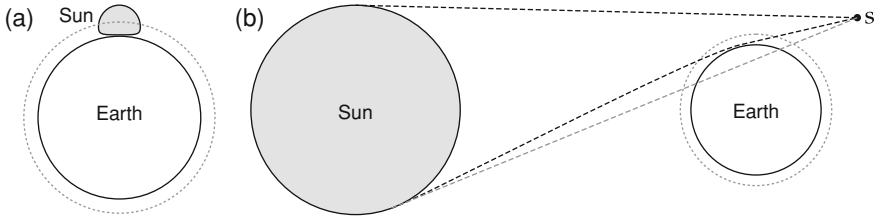
Fig. 3.3 (**a**) Distortion of the shape of the solar disc due to atmospheric refraction as seen from the satellite (*schematic*). (**b**) Eclipse transition geometry under the influence of atmospheric refraction

where

$$\eta < -1 \text{ if the satellite is in full eclipse}$$
$$-1 \le \eta \le 1 \text{ if the satellite is in partial eclipse}$$
$$\eta > 1 \text{ if the satellite is in full sunlight}$$

Two effects can be recognised that determine the value of the shadow function: geometric shadowing of part of the solar disc by the solid Earth and absorption of sunlight by the Earth's atmosphere. The shadow function can be expressed as the product of these two contributions.

$$f_s = f_g f_a \tag{3.31}$$

The geometric shadowing factor f_g indicates the fractional area of the solar disc that is blocked by the Earth. Its calculation can be found in Carrou [11] or Montenbruck and Gill [27], resulting in the following formula:

$$f_g = 1 - \frac{1}{\pi} \arccos{(\eta)} + \frac{\eta}{\pi}\sqrt{1 - \eta^2} \tag{3.32}$$

The amount of sunlight that is not absorbed in the Earth's atmosphere is represented by the factor f_a. Absorption of sunlight in the atmosphere will lead to an extension of the semi-shadow region [47].

The precise modelling of the effects of absorption, refraction and the elliptical shape of the Earth on the eclipse has not been investigated for this thesis research. Chaps. 4 and 5 will show, however, that these effects can be of significant importance for accurate determination of density and winds from accelerometer data.

Varying sun-earth distance

For a realistic solar radiation pressure calculation, the varying distance of the spacecraft from the Sun has to be taken into account. For sunlight at a distance of 1 astronomical unit (1AU $= 149, 597, 870, 660$ m) from the Sun, the radiation pressure has a value of $P_{1AU} = 4.56\,10^{-6}$ N/m^2. Taking the varying Sun-satellite distance $r_{\odot s}$ into account and applying the inverse square law as well as the shadow function from Eq. (3.31), the solar radiation pressure can be expressed as

$$P = \left(\frac{1\text{AU}}{r_{\odot s}}\right)^2 f_s P_{1\text{AU}}. \tag{3.33}$$

3.3 Perturbation Equations

We will now have a closer look at the effect on the orbit shape of the aerodynamic acceleration in the simplified form of Eq. (3.18). To do this, we will be making use of the perturbation equations, which are also known as Gauss's form of Lagrange's planetary equations, or the equations for the method of Gaussian variation of parameters (e.g., [41]).

The perturbation equations for the semi-major axis and eccentricity have been derived in various more or less intricate ways. The method of Burns [10] uses elementary principles of engineering mechanics, and is therefore perhaps the most insightful. Burns [10] first establishes the relationship between the orbital elements and the energy and angular momentum of the space object in its orbit. For unperturbed orbits, the energy and angular momentum are constants, but under the influence of perturbing accelerations, they will change.

To arrive at the perturbation equations, Burns [10] determines the time derivatives of the energy and angular momentum expressed in terms of the orbital elements. We will restrict ourselves to a and e. The change in energy over time should be the result of work done by the perturbing force, and the change of angular momentum over time should be the result of an applied perturbing moment. Substituting equations that express these relations, Burns [10] arrives at the classical expressions of the perturbation equations

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{2a^2}{\sqrt{\mu a(1-e^2)}}\{e\sin\theta\ddot{\boldsymbol{r}}\cdot\hat{\boldsymbol{u}}_{\mathrm{R}} + (1+e\cos\theta)\ddot{\boldsymbol{r}}\cdot\hat{\boldsymbol{u}}_{\mathrm{T}}\} \tag{3.34}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = \sqrt{\frac{a(1-e^2)}{\mu}}\left\{\sin\theta\ddot{\boldsymbol{r}}\cdot\hat{\boldsymbol{u}}_{\mathrm{R}} + \left(\cos\theta + \frac{e+\cos\theta}{1+e\cos\theta}\right)\ddot{\boldsymbol{r}}\cdot\hat{\boldsymbol{u}}_{\mathrm{T}}\right\}. \tag{3.35}$$

These equations provide the rate of change of the orbital elements as a function of accelerations projected on the radial and transverse directions, making use of the unit vectors $\hat{\boldsymbol{u}}_R$ and $\hat{\boldsymbol{u}}_T$, as defined in Fig. 3.4.

We will not have to take into account accelerations in the direction perpendicular to the orbit plane. Such accelerations will appear in the perturbation equations for other orbital elements (see Burns [10] and Vallado [41]), but only for those that affect the orientation of the orbit in space, and not the shape of the orbit. Therefore they do not have to be considered when analysing the effects of drag.

After transformation of the acceleration axes [18] to the along-track and in-plane normal direction, by making use of the flight path angle ψ (see Fig. 3.4), the perturbation equations can be expressed as
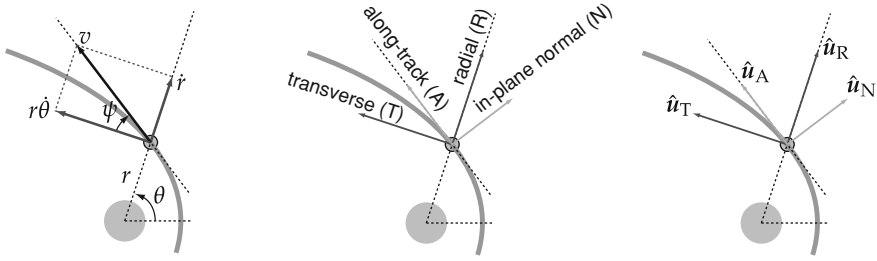
Fig. 3.4 Definitions of angles, axes and corresponding unit vectors in the orbit plane

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{2a^2 v}{\mu} \ddot{\boldsymbol{r}} \cdot \hat{\boldsymbol{u}}_A \tag{3.36}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{1}{v} \{ 2(e + \cos\theta) \ddot{\boldsymbol{r}} \cdot \hat{\boldsymbol{u}}_A - \frac{r}{a} \sin\theta \ddot{\boldsymbol{r}} \cdot \hat{\boldsymbol{u}}_N \}. \tag{3.37}$$

These equations show that the semi-major axis is only affected by along-track accelerations, such as drag, while the eccentricity changes under the influence of both along-track accelerations and in-plane normal accelerations.

Influence of in-plane drag and radiation pressure perturbations on the orbit

Substitution of the in-plane drag Eq. (3.18) and definition of the ballistic coefficient (3.12) in Eqs. (3.36) and (3.37), while making use of $\hat{\boldsymbol{u}}_A = \hat{\boldsymbol{v}}$, results in

$$\left.\frac{\mathrm{d}a}{\mathrm{d}t}\right|_a = -\frac{a^2}{\mu} B\rho F v^3 \tag{3.38}$$

$$\left.\frac{\mathrm{d}e}{\mathrm{d}t}\right|_a = -(e + \cos\theta) B\rho F v. \tag{3.39}$$

Equation (3.38) tells us that a drag acceleration should result in a continuous reduction of the semi-major axis. According to (3.39), the eccentricity of the orbit is reduced under the influence of drag around perigee (where $\theta = 0°$), and increased at apogee (at $\theta = 180°$). However, since for elliptical orbits, the density and velocity will be larger at perigee than at apogee, the decrease of the eccentricity at perigee will far outweigh the increase at apogee.

The resulting change of the orbit shape will be exactly as the observations in the example of Fig. 1.2. in the introduction already showed: the apogee height will be gradually reduced while the perigee height stays nearly the same. When the eccentricity approaches zero and the orbit gets close to circular, the semi-major axis will continue to decrease, causing the trajectory to spiral downwards until re-entry.

The radiation pressure acceleration (3.23) can be substituted into Eqs. (3.36) and (3.37) as well, resulting in

Fig. 3.5 Definition of the unit vectors used in the radiation pressure calculations for a flat panel
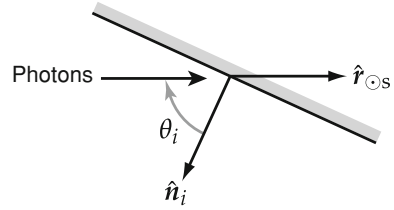
$$\frac{\mathrm{d}a}{\mathrm{d}t}\bigg|_{\mathrm{R}} = \frac{2a^2 v B_{\mathrm{R}} P}{\mu} \boldsymbol{r}_{\odot \mathrm{s}} \cdot \hat{\boldsymbol{u}}_{\mathrm{A}} \qquad (3.40)$$

$$\frac{\mathrm{d}e}{\mathrm{d}t}\bigg|_{\mathrm{R}} = \frac{B_{\mathrm{R}} P}{v} \{2(e + \cos\theta)\boldsymbol{r}_{\odot \mathrm{s}} \cdot \hat{\boldsymbol{u}}_{\mathrm{A}} - \frac{r}{a} \sin\theta \boldsymbol{r}_{\odot \mathrm{s}} \cdot \hat{\boldsymbol{u}}_{\mathrm{N}}\}. \qquad (3.41)$$

The main effect of the radiation pressure acceleration on the shape of the orbit is a harmonic variation at the orbital period. However, since P is modulated by eclipses, longer period and secular effects can occur as well, especially for satellites in highly eccentric orbits.

3.4 Particle Surface Interaction Modelling

A convenient way to compute non-gravitational accelerations on spacecraft of complex geometry with reasonable accuracy is to represent the spacecraft surfaces by a number of flat panels with index i, and combine the acceleration contributions of these panels. In this Section, we will present analytical equations for the force coefficient vector of flat panels, which can be derived from the exchange of energy and momentum between the particles and these panels. The radiation pressure modelling is described in Sect. 3.4.1 and the more complicated aerodynamic interaction in Sect. 3.4.2. How the contributions of these elementary panels can be combined into a model of the non-gravitational accelerations for a complex satellite geometry will be described afterwards, in Sect. 3.5.

3.4.1 Radiation Pressure for a One-Sided Panel

Figure 3.5 shows a flat panel, in combination with the direction of the incoming photons and the various unit vectors that will be used in the radiation pressure calculations below. As far as these force model computations are concerned, the flat panels are one-sided. In other words, the backside of each panel can be considered transparent for photons and atmospheric particles.

Absorption   Specular reflection   Diffuse reflection
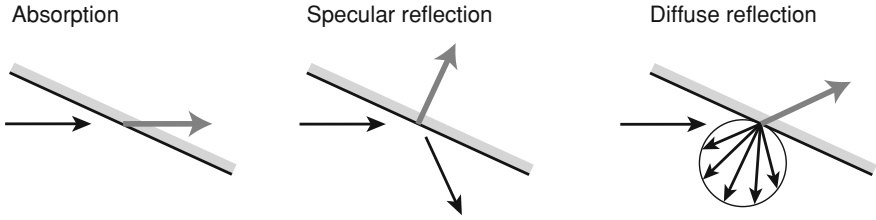
Fig. 3.6 The three basic types of interaction between photons (*black arrows*) and the panel surface, shown with the resulting radiation pressure force vectors (*grey arrows*)

First, we define γ_i as the cosine of the angle between the outward panel normal with index i, and the negative Sun-satellite direction vector.

$$\gamma_i = \cos \theta_i = -\hat{\boldsymbol{r}}_{\odot s} \cdot \hat{\boldsymbol{n}}_i \tag{3.42}$$

For negative γ_i, the panel is not illuminated, so that

$$\boldsymbol{C}_{\mathrm{r},i} = 0 \quad \text{for} \quad \gamma_i \leq 0 \tag{3.43}$$

We will apply a simple but effective model, in which, for positive values of γ_i, a photon arriving at the panel is either absorbed, reflected diffusely or reflected specularly, indicated by the subscripts a, d and s, respectively. Each of these modes of interaction corresponds to a different direction and magnitude of the radiation pressure force (see Fig. 3.6), which we will discuss in turn.

An absorbed photon transfers all its momentum to the panel. The resulting force is in the direction of the unit vector of the incoming particle direction, in this case of the vector from the Sun (or other light source) to the satellite $\hat{\boldsymbol{r}}_{\odot s}$. The force has a magnitude equal to the pressure times the frontal area of the panel $A_i \gamma_i$, so that the contribution to the acceleration of the satellite becomes

$$\ddot{\boldsymbol{r}}_{\mathrm{srp},\mathrm{a},i} = \frac{A_i \gamma_i}{m} P \hat{\boldsymbol{r}}_{\odot s} \quad \text{for} \quad \gamma_i > 0 \tag{3.44}$$

Following the definition of Eq. (3.19), the contribution to the radiation pressure force coefficient vector of the absorbed photons, on a panel with index i is then

$$\boldsymbol{C}_{\mathrm{r},\mathrm{a},i} = \hat{\boldsymbol{r}}_{\odot s} \frac{A_i \gamma_i}{A_{\mathrm{ref}}} \quad \text{for} \quad \gamma_i > 0 \tag{3.45}$$

Note that the value for the reference area A_{ref} that appears in this equation and similar ones below, should be an agreed value for the entire space vehicle. Its value is not related to the dimensions or orientation of the single panel. It appears in these equations just to make the force coefficient dimensionless.

On a shiny surface, particles will be reflected instead of absorbed. If the surface acts like a perfect mirror, the momentum component in the plane of the surface is the

same before and after the interaction, so that it does not contribute to the acceleration. However, the momentum component normal to the surface, which is proportional to γ_i, changes its sign. This results in

$$C_{\text{r,s},i} = 2\gamma_i \hat{\boldsymbol{n}}_i \frac{A_i \gamma_i}{A_{\text{ref}}} \quad \text{for} \quad \gamma_i > 0 \tag{3.46}$$

Diffuse reflection is caused by microscopic roughness of the surface, in which the photon can be considered to be temporarily captured. The direction of re-emission can then be considered independent of the direction of incidence, and in its most simple form, the probability of the re-emission direction can be described by Lambert's cosine law. The momentum exchange is calculated by adding the absorbed momentum of the photon, described by Eq. (3.45), to the momentum carried out by the same photon after it is re-emitted. After integration over the half space, the re-emitted photon flux is directed along the surface normal and carries a momentum equal to two-thirds that of the incoming photons, resulting in

$$C_{\text{r,d},i} = (\hat{\boldsymbol{r}}_{\odot \text{s}} - \frac{2}{3}\hat{\boldsymbol{n}}_i)\frac{A_i \gamma_i}{A_{\text{ref}}} \quad \text{for} \quad \gamma_i > 0 \tag{3.47}$$

Realistic satellite panel materials can be regarded as having a mixture of the optical properties described above. If p_{d} is the proportion of diffusely reflected photons and p_{s} the proportion of specularly reflected photons, then $(1 - p_{\text{s}} - p_{\text{d}})$ is the proportion of absorbed photons on a panel. Measured values for these optical properties of satellite materials are sometimes made available in technical documentation (e.g., [4]) or publications (e.g., [9]). Using these values, the radiation pressure coefficient of the panel can be expressed in vector form as

$$C_{\text{r},i} = \begin{cases} 0 & \text{for} \quad \gamma_i \leq 0 \\ \left[(1 - p_{\text{s}} - p_{\text{d}})\hat{\boldsymbol{r}}_{\odot \text{s}} + p_{\text{s}}2\gamma_i\hat{\boldsymbol{n}}_i + p_{\text{d}}\left(\hat{\boldsymbol{r}}_{\odot \text{s}} - \frac{2}{3}\hat{\boldsymbol{n}}_i\right)\right]\frac{A_i \gamma_i}{A_{\text{ref}}} & \text{for} \quad \gamma_i > 0 \end{cases} \tag{3.48}$$

3.4.2 Aerodynamics for a One-Sided Panel

The calculation of aerodynamic force coefficients on an elementary one-sided panel follows a similar pattern as the radiation pressure calculations in the previous Section. The photons causing the radiation pressure all travel at the speed of light, and in the same direction from the Sun to the satellite. However, the gas particles which cause the aerodynamic interaction have a thermal velocity component that might not be negligible compared to the satellite velocity, and which might well change during the interaction with the surface.

The description of the interaction of atmospheric particles with a surface element can be split into two distinct contributions: that of the incident particle flux, and that

Fig. 3.7 Definition of the
velocity components and
unit vectors used in the
aerodynamic calculations for
a flat panel
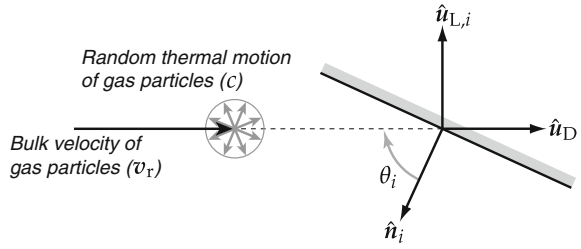
of the reflected or reemitted particle flux. Both contribute to the drag, however lift
and sideways aerodynamic forces are mainly generated by the reflected particle flux.

Incident particle flux

An exact description of the acceleration contribution by the incident particle flux
is possible, but it requires knowledge on the magnitude and direction of the rel-
ative velocity, v_r, with respect to the surface element, the atmospheric tempera-
ture T, and the relative mass concentrations, ρ_j/ρ, of the different particles species
($j = O_2, N_2, O, He, H, \ldots$) having different molecular masses m_j. These latter
parameters determine the velocity of the random motion c_j of the molecules and
atoms, which is to be superimposed on the bulk velocity v_r of the atmosphere with
respect to the satellite surface. Figure 3.7 illustrates this schematically, and also shows
the definition of unit vectors to be used in the aerodynamics equations below.

From gas kinetic theory (e.g., [5]), the most probable thermal velocity of the
molecules and atoms can be expressed as

$$c_{mp,j} = \sqrt{2\frac{k}{m_j}T},\tag{3.49}$$

where k is the Boltzmann constant.

An important parameter describing the flow characteristics is the speed ratio S of
the bulk velocity to the most probable thermal velocity

$$S_j = \frac{v_r}{c_{mp,j}}.\tag{3.50}$$

Early analyses of CHAMP and GRACE accelerometer data [8, 9, 35, 37] used
aerodynamic expressions by [19]. These are only valid for high speed ratios ($S \gg 1$)
and compact satellite shapes ($S\cos\theta_i > 1$) [19]. In the derivation of these equations,
the influence of the random thermal motion of the atoms and molecules on the
aerodynamic force could be ignored. Since the CHAMP and GRACE satellites have
elongated shapes, the data analysis using these equations resulted in lower drag
coefficients and higher densities with larger fluctuations. Sutton [36] presented an
update of the aerodynamic model, which showed a much improved fidelity of his
density data [38], a difference which was confirmed in our own processing.

Also in contrast to some previous analyses (e.g., [9]), it can be important to not use a mean molecular mass, but instead calculate the i-th panel's contribution to the aerodynamic force coefficient $\boldsymbol{C}_{a,i}$ as the mass-concentration-weighted sum of contributions by the various atmospheric constituents j.

$$\boldsymbol{C}_{a,i} = \sum_j \frac{\rho_j}{\rho} \boldsymbol{C}_{a,j,i} \tag{3.51}$$

As we shall see, the dependence of the contributions $\boldsymbol{C}_{a,j,i}$ on the molecular mass can be highly non-linear. When a small concentration of light-weight constituents (such as Helium) is present, this does not affect the mean molecular mass by much. However, these light-weight particles, because of their high thermal velocity, will have a higher collision rate with the satellite's side panels, that are oriented nearly parallel to the stream, than the heavier constituents (such as oxygen or nitrogen). This will result in significantly larger values of $\boldsymbol{C}_a$, especially for the elongated satellite shapes of the current accelerometer missions.

Reflected particle flux

A description of the reflected particle flux requires a model of the gas-surface inter-action, which specifies the angular distribution and energy flux of the reflected parti-cles. Unfortunately, experimental data on gas-surface interaction [16, 25] is limited to only a subset of the range of conditions under which current space accelerom-eters are making measurements. This puts an exact physical representation of this contribution to the aerodynamic force out of reach.

Ideally, information on the gas-surface interaction, as well as in-situ observa-tions of aerodynamic model parameters like the temperature T and concentrations c_j, should be measured by independent instruments on the accelerometer-carrying satellite. Since the current and planned accelerometer missions lack the required instrumentation, we have to rely on empirical atmosphere models such as NRLMSISE- 00, simplified gas-surface interaction models, and some educated guesses.

Such simplified gas-surface interaction models contain parameters like the energy flux accommodation coefficient α (e.g., [33]), which determines whether the particles retain their mean kinetic energy (for $\alpha = 0$) or acquire the temperature of the spacecraft surface T_{wall} (for $\alpha = 1$). This wall temperature often introduces another uncertainty in calculations on the gas-surface interaction. Often no measurements or models of this value are available. Fortunately, the sensitivity to this parameter is usually quite low.

Another possible gas-surface interaction parameter is the Maxwell coefficient σ, which determines the fraction of particles that leaves the surface in either a completely diffuse ($\sigma = 1$) or completely specular ($\sigma = 0$) angular distribution. This Maxwell model [20, 33] shows similarities with the interaction model for radiaton pressure discussed before and illustrated in Fig. 3.5. Of course, the principle of conservation

of mass requires that absorption of gas particles by the surfaces cannot occur like it can for photons.

Drag and lift coefficient equations

With the caveats concerning these simplifications in mind, the rarefied aerodynamic equations for flat panels, derived by Sentman [34] are currently seen as an appropriate choice for use in the processing of CHAMP and GRACE data. Sentman's equations take into account the random thermal motion of the incident particles, and assume a completely diffuse distribution of the reflected particle flux, with temperature T_{refl}. This is reasonably consistent with the limited data from in-orbit gas-surface interaction experiments [16, 23], which suggests that below about 500 km, the angular distribution is likely within a few percent of complete diffuse reemission ($\sigma \gtrsim 0.95$), and that the energy flux accommodation is quite high ($\alpha \gtrsim 0.8$). Moe et al. [24] introduced the energy flux accommodation coefficient α as a parameter in Sentman's equations. The implementation of this modification of Sentman's equations for accelerometer data processing is largely similar to the one recently published by Sutton [36]. It will be described briefly below.

The calculation procedure starts with some geometry, of which the symbols are explained in Fig. 3.7. The unit vector in the drag direction $\boldsymbol{u}_D$ is determined from the relative velocity vector.

$$\hat{\boldsymbol{u}}_{\mathrm{D}} = \frac{\boldsymbol{v}_{\mathrm{r}}}{||\boldsymbol{v}_{\mathrm{r}}||} \tag{3.52}$$

The unit vector for the lift and side force direction, $\hat{\boldsymbol{u}}_{\mathrm{L},i}$, is perpendicular to $\hat{\boldsymbol{u}}_{\mathrm{D}}$ and in the plane spanned by $\hat{\boldsymbol{n}}_i$ and $\hat{\boldsymbol{u}}_{\mathrm{D}}$.

$$\hat{\boldsymbol{u}}_{\mathrm{L},i} = -\frac{(\hat{\boldsymbol{u}}_{\mathrm{D}} \times \hat{\boldsymbol{n}}_i) \times \hat{\boldsymbol{u}}_D}{||(\hat{\boldsymbol{u}}_{\mathrm{D}} \times \hat{\boldsymbol{n}}_i) \times \hat{\boldsymbol{u}}_D||} \tag{3.53}$$

Sentman's equations make use of the cosine γ_i of the angle between the inward normal and the drag vector, and of the cosine l_i of the angle between the inward normal and the lift vector. These are easily determined using the inner products of these vectors.

$$\gamma_i = \cos(\theta_i) = -\hat{\boldsymbol{u}}_{\mathrm{D}} \cdot \hat{\boldsymbol{n}}_i, \quad l_i = -\hat{\boldsymbol{u}}_{\mathrm{L}} \cdot \hat{\boldsymbol{n}}_i \tag{3.54}$$

Using similar notation as Sutton [36] and Moe and Moe [25], Sentman's formulas for the drag coefficient and combined lift and side force coefficient are provided as

$$C_{\mathrm{D},i,j} = \left[\frac{P_{i,j}}{\sqrt{\pi}} + \gamma_i Q_j Z_{i,j} + \frac{\gamma_i}{2} \frac{v_{\mathrm{re}}}{v_{\mathrm{inc}}} (\gamma_i \sqrt{\pi} Z_{i,j} + P_{i,j}) \right] \frac{A_i}{A_{\mathrm{ref}}} \tag{3.55}$$

$$C_{\mathrm{L},i,j} = \left[l_i G_j Z_{i,j} + \frac{l_i}{2} \frac{v_{\mathrm{re}}}{v_{\mathrm{inc}}} (\gamma_i \sqrt{\pi} Z_{i,j} + P_{i,j}) \right] \frac{A_i}{A_{\mathrm{ref}}} \qquad (3.56)$$

where

$$G_j = \frac{1}{2S_j^2}, \quad P_{i,j} = \frac{1}{S_j} \exp(-\gamma_i^2 S_j^2), \quad Q_j = 1 + G_j, \quad Z_{i,j} = 1 + \mathrm{erf}(\gamma_i S_j)$$

$$(3.57)$$

The error function is defined as:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-y^2)\,\mathrm{d}y \qquad (3.58)$$

The last term in these equations is proportional to the ratio of the velocity of the re-emitted particles (v_{re}) to that of the incoming particles ($v_{\mathrm{inc}} = v_{\mathrm{r}}$). Koppenwallner [20] analysed the derivation of the expression by Moe and Moe [25] of this ratio as a function of the accommodation coefficient α and the wall temperature T_{w}, and proposed the following corrected version

$$\frac{v_{\mathrm{re}}}{v_{\mathrm{inc}}} = \sqrt{\frac{1}{2}\left[1 + \alpha\left(\frac{4RT_{\mathrm{w}}}{v_{\mathrm{inc}}^2} - 1\right)\right]} \qquad (3.59)$$

Aerodynamic calculation results

The result of Sentman's drag and lift coefficient equations is shown in Fig. 3.8, as a function of the incidence angle θ. Unless otherwise noted, the parameters used in the aerodynamic calculations for this and the next figures in this chapter are as follows: $A_{\mathrm{ref}} = 1\,\mathrm{m}^2$, $T_{\mathrm{w}} = 300\,\mathrm{K}$, $\alpha = 1$, $T = 1000\,\mathrm{K}$, $m_j = 16$(atomic oxygen) and $v_{\mathrm{r}} = 7600\,\mathrm{m/s}$, resulting in $S_j = 7.45$.

Notice that the lift coefficient is many times smaller than the drag coefficient. This is one of the reasons why the lift component can often be ignored in applications, as discussed in Sect. 3.2.1. At first approximation, the drag and lift coefficients vary with the cosine of the incidence angle, and the sine of twice that angle, respectively, corresponding to the values of γ and l from Eq. (3.54). For incidence angles near 90° however, where γ and l approach zero, the coefficients depart significantly from these sinusoidal curves. The reason for this is that the random thermal motion of the atmospheric particles superimposed on the relative velocity enables a fraction of these particles to interact with the front of the panel. This happens even when $\theta > 90°$. Without this thermal motion component, the force would be zero under that condition, comparable with the situation for radiation pressure, as described in Eq. (3.48). In contrast, in aerodynamic calculations, it is important not to automatically set the force to zero when the front of the panel is facing away from the bulk stream of particles.
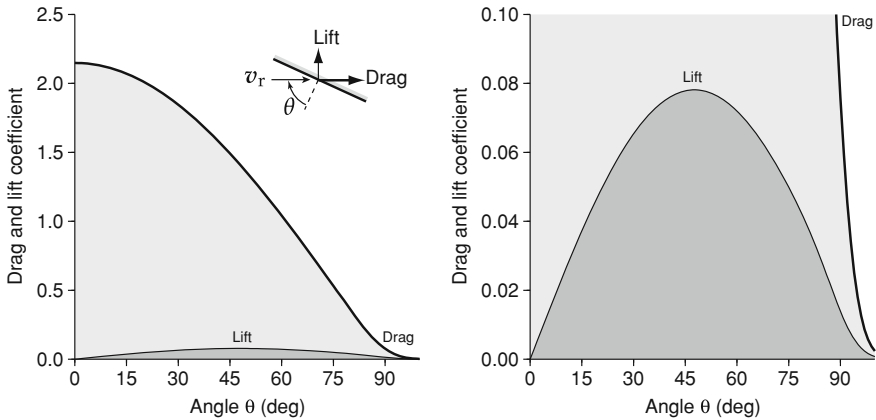
Fig. 3.8 Drag and lift coefficients, according to Sentman's equations for a one-sided flat panel, as a function of the incidence angle θ. The right-hand plot contains the same information as the left-hand plot, but with a scaled Y-axis to more clearly show the lift and values around $\theta \approx 90°$

3.4.3 Uncertainties in Aerodynamic Model Calculations

Several quantities, such as the atmospheric temperature and molecular mass are often not available as exact measurements, simultaneous with the drag measurement. They need to be obtained from an empirical model, in order to complete the aerodynamic calculation. Model errors could then lead to aerodynamic coefficient errors. Other quantities, such as those describing the gas-surface interaction, have not been very extensively measured and modeled at all. It is therefore important to study the sensitivity of the aerodynamic coefficient calculation to these quantities.

Aerodynamic sensitivity analysis for an elementary panel

The top row of Fig. 3.9 shows how the drag and lift coefficients are influenced by the atomic mass m_j of the atmospheric particles via the speed ratio S_j. The atmospheric temperature will have the inverse effect on the drag coefficient, as shown by Eq. (3.49). A higher temperature causes higher force coefficients. But for this discussion we will keep the temperature constant and focus on the particle mass.

The force coefficients are significantly higher for the very light-weight constituents hydrogen and helium. For the heavier constituents, the drag and lift coefficients asymptotically approach a constant value. This happens because, at a certain point, the thermal motion of these heavier particles is negligible compared to the satellite velocity. For this combination of particle mass, atmospheric temperature and satellite velocity, the flow can be considered *hyperthermal*. For the right-most figure ($\theta \approx 90°$), when the panel is nearly parallel to the velocity direction, the sensitivity to the mass (and temperature) of the atmospheric particles is much greater. The asymptotic value is zero in this case, but it will not be reached within the normal
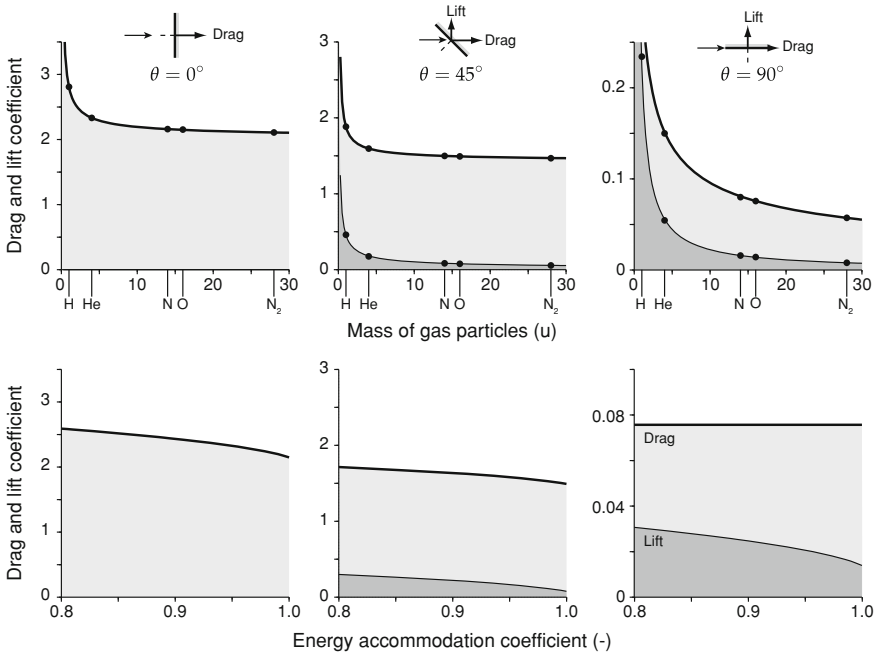
Fig. 3.9 Variation of drag and lift coefficients for a one-sided panel as a function of the mass of the gas particles (*top*) and of the energy accommodation coefficient (*bottom*), for three different incidence angles

range of particle masses, temperatures and satellite velocities. This means that the hyperthermal assumption is invalid for long and slender satellite shapes.

The bottom row of Fig. 3.9 shows the variation of the drag and lift coefficients with the energy accommodation coefficient α, which determines the velocity of the re-emitted atmospheric particles. The Figure shows that an uncertainty or error in the accommodation coefficient can easily lead to a significant error in the drag coefficient for panels perpendicular to the stream. It will also lead to much larger errors in the lift coefficient, for surfaces parallel to the stream. This large sensitivity of the lift force to the gas-surface interaction could perhaps be used in the design of future experiments for investigation of gas-surface interaction properties.

Choice of energy accommodation coefficient

Figure 3.9 demonstrated the importance of the choice of the value of the energy accommodation coefficient for the calculation of the aerodynamic force coefficient. Pardini et al. [28] have published on this issue, based on earlier work by Bowman and Moe [7]. Other recent works of relevance to this topic are the analyses of drag on spherical satellites by Moe and Bowman [22], and Bowman and Hrncir [6].

The underlying theory behind the use of the equations by Sentman [34] based on diffuse re-emission, and the energy accommodation coefficient as described by

Moe and Moe [22] is that, when exposed to an atmosphere rich in atomic oxygen, a thin layer of these oxygen atoms is thought to form on the solid spacecraft walls by physical adsorption or chemisorption. Leaving the scientific terms behind for a moment, it can be said that this oxygen layer is softer, at the atomic level, than the hard technical materials of the spacecraft wall, such as aluminium, glass, and thermal insulation sheets. This soft layer causes the atmospheric particles to bounce around in or stick on this surface layer, accommodating their energy to the wall temperature ($\alpha = 1$), and to "lose memory" of their initial traveling direction, causing a diffuse re-emission or reflection.

At higher altitudes and low solar activity, there will be less atomic oxygen, preventing the formation of the adsorbed layer. This would cause a reduction in the accommodation coefficient and most likely also a reduction of the diffuse re-emission. Numerical data on this change of re-emission direction, (implying an increase of quasi-specular reflection), is more scarce and complex than on the energy accommodation. Due to the increased modelling complexity it is ignored in the present analyses, and the diffuse re-emission assumption is kept.

The investigation of Pardini et al. [28] makes use of a density model to derive an observed average drag coefficient from the orbital decay observations for spherical and attitude-stabilised satellites, using Eq. (3.10). This *observed drag coefficient* is then considered valid for the selected level of solar activity and altitude for which the analysis took place. These observed drag coefficients are compared with drag coefficient calculation results, called *physical drag coefficients*, so that the energy accommodation coefficient can be obtained that makes the two identical.

The use of density models in these investigations [26, 28] implies the possibility of a circular reference. The drag and accommodation coefficient values found are dependent on the density model, and therefore on the choice of drag coefficient used in processing the drag data that was used to generate the density model. Unfortunately, this situation currently seems impossible to avoid, due to the lack of experiments that have independently measured both density and parameters related to the gas-surface interaction, in-situ, and over a wide-enough range of conditions.

With these caveats in mind, the accommodation coefficient values obtained by Pardini et al. [28] are presented as points in Fig. 3.10b. The data is provided in this paper in separate tables for solar minimum and solar maximum, and at discrete altitude levels, with the solar maximum data extending to higher altitudes (630 km) than at solar minimum (325 km). This representation is too coarse to apply the energy accommodation coefficients to other data, for example at intermediate levels of solar activity, or in between or beyond the provided altitude levels.

Pilinski et al. [30] therefore reanalysed the accommodation coefficient data of Pardini et al. [28] with the help of gas adsorption theory. The fraction θ of the surface area covered by adsorbed or chemisorbed gas particles can be expressed as a function of the pressure P in the form of a so-called *Langmuir isotherm* ([1], Equation XVII-5):

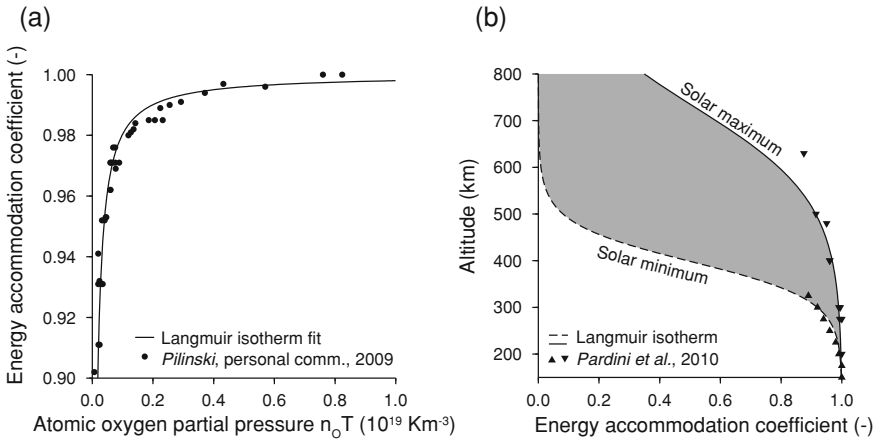$$\theta = \frac{bP}{1 + bP} \tag{3.60}$$

Fig. 3.10 Measured (*points*) and modelled (*lines*) energy accommodation coefficients, as a function of NRLMSISE- 00 atomic oxygen partial pressure (*left*) and of altitude and level of solar activity (*right*)

The constant b is related to the rates of condensation and evaporation per unit surface area. Pilinski et al. [30] investigated fitting the accommodation coefficient data using this function, assuming a linear relation between this data and the adsorbed surface fraction θ. They computed the partial pressure of atomic oxygen from the product of the number density and temperature $P_O = n_O T$, both taken from the NRLMSISE- 00 model. The constant molecular mass and gas constant are omitted in this pressure calculation by Pilinski et al. [30], which affects the value of the constant b. The resulting isotherm is plotted in Fig. 3.10a. We selected a value of $b = 5 \cdot 10^{-17}$. The function applied to representative solar minimum and maximum altitude profiles from the NRLMSISE- 00 model is plotted underneath the data in Fig. 3.10b.

Note that the data do not perfectly fit the isotherm, and that the isotherm extends far beyond the data at high altitude, especially at low solar activity. The function offers an alternative for just picking a fixed value for the energy accommodation coefficient to be used under all conditions. However, the validity of the use of the isotherm function for this purpose, and using the instantaneous and local NRLMSISE- 00 values as inputs remains to be investigated. Furthermore, the applicability of the function beyond the range of the original data points, as seen in Fig. 3.10b is questionable.

3.5 Modelling of Accelerations on Complex Satellite Shapes

There are various approaches by which the non-gravitational force computations for an elementary flat surface, as provided in the previous Sections, can be combined for a complete satellite. The two most widely used approaches will be discussed in turn below, after which some accuracy aspects and computation results will be discussed.

Panel models

A satellite panel model for non-gravitational acceleration computations consists in its simplest form of a set of $i = 1, \ldots, N$ panels, each having only an area A_i, an outward normal unit vector $\hat{\boldsymbol{n}}_i$ that defines its orientation in a satellite body-fixed frame, and a set of optical properties $p_{s,i}$ and $p_{d,i}$.

Such models are available in literature for the CHAMP [9] and GRACE [4] accelerometer satellites, consisting of 15 and 8 panels, respectively. These models do not provide information on the shape and relative position of each panel. Finer detail of the spacecraft surface, such as protruding antennae, star camera baffles, etc., are either neglected or handled by slightly adjusting the parameters of the larger flat surfaces.

For such panel models, the radiation pressure force coefficient vector for the entire satellite can be computed by simply summing the panel contributions, as follows:

$$\boldsymbol{C}_{\mathrm{r}} = \sum_i \boldsymbol{C}_{\mathrm{r},i} \tag{3.61}$$

For the aerodynamic force coefficient, a double summation is necessary, both over the satellite panels i, and over the atmospheric constituents j. In addition, the contributions of drag and lift were computed separately, and now need to be summed as well.

$$\boldsymbol{C}_{\mathrm{a}} = \sum_i \sum_j \frac{\rho_j}{\rho} (C_{\mathrm{D},i,j} \hat{\boldsymbol{u}}_{\mathrm{D}} + C_{\mathrm{L},i,j} \hat{\boldsymbol{u}}_{\mathrm{L},i}) \tag{3.62}$$

Note that the unit vector in the drag direction depends only on the relative velocity and is independent of the panel orientation, and therefore does not carry the index i. Depending on the orientation of the 2D panel within the 3D satellite model, the lift contribution of each panel, in the direction $\hat{\boldsymbol{u}}_{\mathrm{L},i}$ shall result in a combined aerodynamic lift and side force.

Full 3D models with shadowing and multiple reflections

More accurate geometry models, with an arbitrary number of precisely positioned and shaped panels, can be created in specialized software, such as ANGARA [14], based on CAD drawings of the satellites. ANGARA is able to calculate both radiation pressure and aerodynamic force coefficients using the same tools for building the satellite geometry. Similar techniques and software implementations are described in Ziebart [47] and Ziebart et al. [46] for radiation pressure effects and in Fuller and Tolson [15] for aerodynamics.

The modeled accelerations can then be calculated by evaluating the same analytical equations as for the simple panel models. In this case, sufficient information on the geometry is available to perform an analysis of shadowing of parts of the satellite by other parts [15]. Figure 3.11 shows ANGARA models for the GRACE and CHAMP satellites. Alternatively, a Monte Carlo test particle method [14] can be applied on the panellised representation of the satellite. The advantage of this
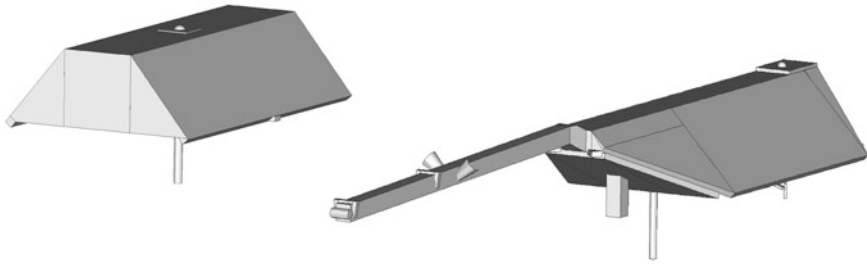
Fig. 3.11 Geometrical models of the GRACE and CHAMP satellites created in the ANGARA non-gravitational force modelling software

Table 3.1 Projected areas (in m^2) of CHAMP and GRACE as viewed along the spacecraft body-fixed axes

Satellite	Source	Reference	X		Y		Z	
CHAMP	CH- IT- DID- 0001	Lühr[Personal communication]	0.7425		3.1203		6.4442	
	ANGARA	Doornbos et al.[13]	0.7944	+7%	3.2446	+4%	6.6210	+3%
	Panels	Lühr et al. [21]	0.4701	-37%	3.3775	+8%	6.6899	+4%
	Panels	Bruinsma and Biancale [9]	0.6366	-14%	3.3775	+8%	6.6899	+4%
GRACE	CSR- GR- 03- 02	Bettadpur [4]	1.0013		2.4634		–	
	ANGARA	Doornbos et al.[13]	1.0433	+4%	2.5496	+3%	6.1530	–
	Panels	Bettadpur [4]	1.0013	0%	2.6383	+7%	6.2240	–

approach is that the contributions of multiple reflections of particles and shadowing of parts of the satellite by other parts will be automatically taken into account in the simulation.

Geometric modelling uncertainty

Before presenting aerodynamic calculation results for the complete satellite models of CHAMP and GRACE, it is important to first look at the geometrical descriptions of the satellites themselves.

During the course of this thesis work, it was found that the flow-projected area of the satellite, which varies as a function of the incidence angle of the particles, can be a major source of error in density data processing. To illustrate this, Table 3.1 lists values of projected areas for CHAMP and GRACE, as viewed along their principle body-fixed axes, taken from various sources and models.

The top row for each of the two missions in this Table consists of quoted values that were obtained from the agency responsible for the satellite's manufacturing and operations. These numbers are likely quite accurate, but they are available only for either two (GRACE) or all three (CHAMP) of the principle spacecraft body-fixed axes. These numbers are therefore not directly applicable in non-gravitational force modelling, but they can serve as a reference check for the areas of the fully
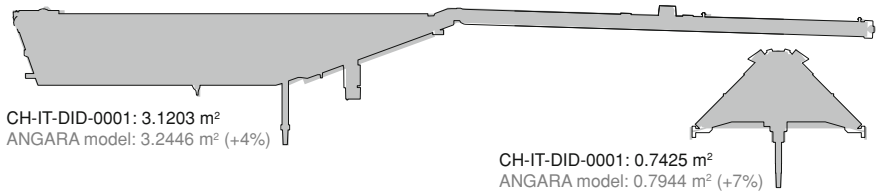
CH-IT-DID-0001: 3.1203 m²
ANGARA model: 3.2446 m² (+4%)

CH-IT-DID-0001: 0.7425 m²
ANGARA model: 0.7944 m² (+7%)

Fig. 3.12 Comparison of CHAMP frontal areas, acquired from the contractor's document (*outline*) and from the ANGARA model (*shaded area*)

three dimensional ANGARA and panel models. The second row for each satellite contains values from the detailed ANGARA models of Fig. 3.11, that were developed at HTG, Katlenburg-Lindau, as part of an ESA-funded project [13]. These models were carefully constructed on the computer, using detailed CAD drawings of the satellites. The third row for each of the satellites contains values from previously published simple panel models.

Discrepancies of the order of 4–8% are prevalent throughout the table, which gives an indication of the apparently inherent level of inaccuracy for these models. However, the much larger errors of 37% and 14% for the two CHAMP panel models along the X-direction are very striking. An explanation for this is that the frontal area contribution of the main panels of the satellite as viewed from this direction is relatively small, while protruding elements, such as the communications antenna, instruments, star camera baffles and attitude thrusters have a relatively large contribution to the area. It seems that these elements were not adequately taken into account in the creation of the panel models.

Figure 3.12 shows a more detailed comparison of the geometries of the CHAMP ANGARA model and from the CH- IT- DID- 0001 document [Lühr, personal communication]. It seems that the thickness of the inclined front/bottom panel of the satellite contributes significantly to the discrepancy in area. The shapes of the attitude control thrusters on both sides of the satellite, the star camera baffles halfway down the boom and the Overhauser magnetometer at the tip of the boom are smaller on the ANGARA model. Perhaps this is partly the case because covering thermal isolation materials were not present in the original CAD drawings. Unfortunately, the reference document was not available during the construction of the ANGARA model, otherwise, the model could have been more thoroughly checked and corrected.

With the help of these additional data, the panel models can be improved as well, to better match the actual dimensions of the satellite. Such an adjusted panel model [13] has been used for CHAMP in the remainder of this thesis, except if otherwise indicated. For GRACE, the panel model by Bettadpur [4] was used as the default model.
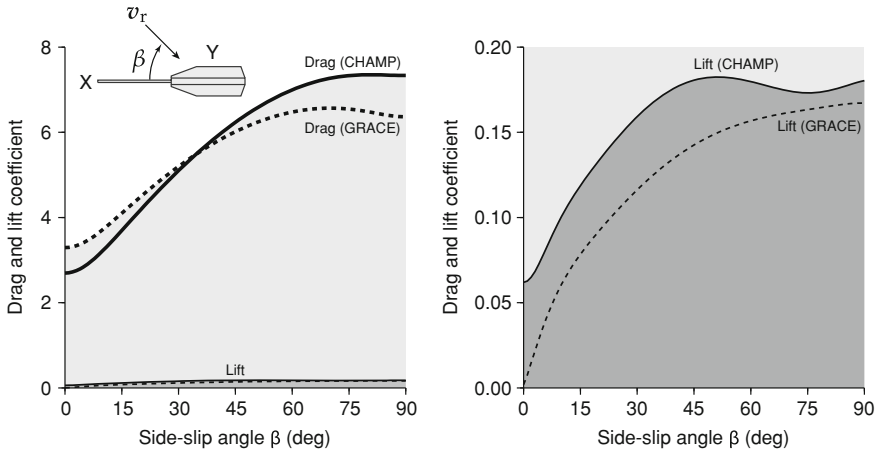
Fig. 3.13 Drag and lift coefficients for the complete CHAMP (*solid lines*) and GRACE (*dashed lines*) panel models, as a function of the side-slip angle β. A reference area of $1\,m^2$ was used for both satellites

Aerodynamic calculation results for CHAMP and GRACE

Figure 3.13 shows the drag and lift coefficients calculated for CHAMP and GRACE panel models, as a function of the side-slip angle β. This is the angle in the body-fixed X-Y plane, between the relative velocity vector and the X-axis.

A large part of the variation of the drag coefficient with β is due to the fact that the true frontal area $A_i \gamma_i$ of each panel changes with β, while the reference area A_{ref} remains fixed at $1\,m^2$ for this plot. Dividing the plotted values at $\beta = 0°$ and $\beta = 90°$ by the corresponding panel frontal areas, the drag coefficient with respect to these areas can be calculated. This results in values of about 3.6 and 2.4, respectively. This factor of about 1.5 is due to the additional drag of collisions of particles with the large panels that are parallel to the flow at $\beta = 0°$. While these panels do not contribute to the frontal area, they clearly do contribute to the aerodynamic acceleration.

A more elaborate illustration of the sensitivity of the drag coefficient to the environment and to the gas-surface interaction model is shown in Fig. 3.14. For this plot, the aerodynamic force coefficients have been evaluated using a CHAMP panel model at two orientations with respect to the flow: as an elongated shape (its nominal attitude), and sideways, as a compact shape. Again, Sentman's equations were used for the calculation, this time making use of the NRLMSISE- 00 model temperature and composition values, which were presented earlier in Figs. 2.1 and 2.3. The force coefficients were evaluated by making use of relative velocity values equivalent to the circular orbit velocity at each altitude, and at representative low and high solar activity conditions. The calculations were also repeated with three values of the accommodation coefficient: two fixed values ($\alpha = 1.0$ and $\alpha = 0.8$), and as a Langmuir isotherm function (3.60) of the NRLMSISE- 00 a tomic oxygen partial pressure.
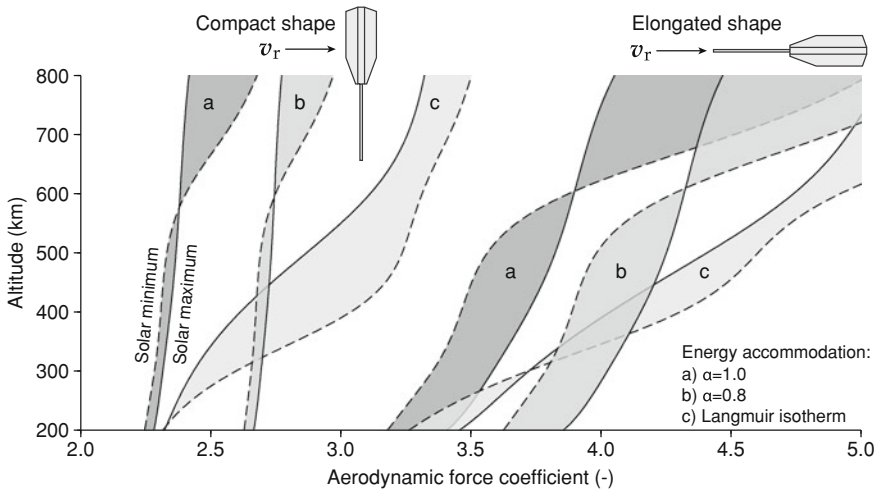
Fig. 3.14 Variation of the aerodynamic force coefficient as a function of altitude for two attitude modes, two levels of solar activity and three choices for the energy accommodation coefficient α, based on Sentman's equation using NRLMSISE composition and temperature values and a CHAMP panel model. The frontal areas of the panel model for each of the two shapes were used as the reference areas

Each of the curves shows increasing aerodynamic force coefficients with increasing altitude. For the curves with fixed energy accommodation coefficients, this increase is due to the increase in temperature, and, only at solar minimum above about 500 km, due to the increased abundance of light-weight helium. The sensitivity of the drag acceleration to these environmental conditions is clearly much larger for the elongated shape of the satellite at $\beta = 0°$, compared to the compact aerodynamic shape at $\beta = 90°$.

The use of a lower value of the energy accommodation coefficient α results in a shift to higher aerodynamic coefficient values. Analyses of drag on satellites [22, 28] (see also the discussion in Sect. 3.4.3) suggest that the true energy accommodation is not fixed but varies with the absorption of atomic oxygen on the surface, which likely decreases with altitude. When the accommodation coefficient is computed as the Langmuir isotherm function of atmospheric temperature and atomic oxygen number density (see Fig. 3.10), the force coefficient increases much faster with altitude. Note however, that most of the upper part of the curves marked with (c) in Fig. 3.14, are outside the range of the observation data used to derive the Langmuir isotherm fit of Fig. 3.10.

Unfortunately, exact and independent observations of the accommodation coefficient variation, and other parameters related to the gas-surface interaction, are scarce, especially at low activity and high altitude. This complicates the computation of an exact drag coefficient significantly. It is likely that the assumption of completely diffuse re-emission of atmospheric particles that was used in the derivation of Sent-

man's equations no longer holds at higher altitudes and lower solar activity levels, when the aerodynamic interaction with atomic oxygen is replaced by helium and hydrogen. This brings an additional level of uncertainty to the drag coefficient calculation, on top of the ones shown in Fig. 3.14. These facts are important to consider when the processing of density and wind data from satellite dynamics observations is discussed in the next chapter.

References

1. Adamson AW, Gast AP (1997) Physical chemistry of surfaces. 6th edn edn. Wiley, New York
2. Adhya S, Sibthorpe A, Ziebart M, Cross P (2004) Oblate earth eclipse state algorithm for low-earth orbiting satellites. J Spacecr Rockets 41(1):157–159
3. Anderson JD (2010) Fundamentals of aerodynamics 5th edn. McGraw-Hill, New York
4. Bettadpur S (2007) Gravity recovery and climate experiment product specification document (rev 4.5–Feb 20, 2007), GRACE 327–720/CSR–GR–03–02. Center for Space Research, The University of Texas, Austin
5. Bird GA (1994) Molecular gas dynamics and the direct simulation of gas flows. Oxford Engineering Science Oxford University Press, Oxford
6. Bowman BR, Hrncir S (2007) Drag coefficient variability at 100–300 km from the orbit decay analyses of rocket bodies. In: AIAA/AAS Astrodynamics Specialist Conference, August 19–23, 2007, Mackinac Island, Michigan, AAS 07-262
7. Bowman BR, Moe K (2005) Drag coefficient variability at 175–500 km from the orbit decay analyses of spheres. In: AAS/AIAA Astrodynamics Specialist Conference, Lake Tahoe, CA, Aug 7–11, 2005, AAS 05–257
8. Bruinsma S, Tamagnan D, Biancale R (2004) Atmospheric densities derived from CHAMP/STAR accelerometer observations. Planet Space Sci 52(4):297–312 doi:10.1016/j.pss.2003.11.004
9. Bruinsma S, Biancale R (2003) Total densities derived from accelerometer data. J Spacecr Rockets 40(2):230–236
10. Burns JA (1976) Elementary derivation of the perturbation equations of celestial mechanics. Am J Phy 44(10):944–949 doi:10.1119/1.10237
11. Carrou JP (ed.) (1995) Spaceflight dynamics part I, Cépaduès
12. Cook GE (1965) Satellite drag coefficients. Planet Space Sci 13(10):929–946 10.1016/0032-0633(65)90150-9
13. Doornbos E, Förster M, Fritsche B, van Helleputte T, van den IJssel J, Koppenwallner G, Lühr H, Rees D, and Visser P (2009) ESTEC contract 21022/07/NL/HE Air density models derived from multi-satellite drag observations – final report, DEOS / TU Delft scientific report 01/2009, TU Delft
14. Fritsche B, Ivanov M, Kashkovsky A, Koppenwallner G, Kudryavtsev A, Voskoboinikov U, Zhukova G (1998) Radiation pressure forces on complex spacecraft, final report, ESOC contract 11908/96/D/IM, HTG, Germany and ITAM, Russia
15. Fuller JD, Tolson RH (2009) Improved method for estimation of spacecraft free-molecular aerodynamic properties. J Spacecr Rockets 46(5):938–948 doi:10.2514/1.43205
16. Gregory JC, Peters PN (1987) A measurement of the angular distribution of 5 ev atomic oxygen scattered off a solid surface in earth orbit. In: Boffi V., Cercignani C. (eds) Proceedings of the 15th international symposium on rarefied gas dynamics. vol 1, pp. 644–656, B.G. Teubner, Stuttgart
17. Groten E (1999) Report of special commission 3 of IAG—fundamental constants. In: XXII IAG General Assembly

18. King-Hele D (1987) Satellite orbits in an atmosphere, theory and applications. Blackie
19. Koppenwallner G (2008) Comment on special section: new perspectives on the satellite drag environments of Earth, Mars and Venus. J Spacecr Rockets 45(6):1324–1326 doi:10.2514/1.37539
20. Koppenwallner G (2009) Energy accommodation coefficient and momentum transfer modeling. HTG–TN–08–11, HTG Katlenburg Lindau
21. Lühr H, Grunwaldt L, Förste Ch (2002) CHAMP Reference Systems. Transformations and Standards, CH-GFZ-RS-002
22. Moe K, Bowman B (2005) The effects of surface composition and treatment on drag coefficient of spherical satellites. In: AAS/AIAA Astrodynamics Specialist Conference, Lake Tahoe, CA, August 7–11, 2005, AAS 05–258
23. Moe K, Moe MM, Wallace SD (1998) Improved satellite drag coefficient calculations from orbital measurements of energy accomodation. J Spacecr Rockets 35(3):266–272
24. Moe K, Moe MM, Rice CJ (2004) Simultaneous analysis of multi-instrument satellite measurements of atmospheric density. J Spacecr Rockets 41(5):849–853
25. Moe K, Moe MM (2005) Gas–surface interactions and satellite drag coefficients. Planet Space Sci 53(8):793–801 doi:10.1016/j.pss.2005.03.005
26. Moe MM, Wallace SD, Moe K (1995) Recommended drag coefficients for aeronomic satellites in The upper mesosphere and lower thermosphere: a review of experiment and theory, geophysical monograph 87. Am Geophys Union, Washington
27. Montenbruck O, Gill E (2000) Satellite orbits—models, methods and applications. Springer
28. Pardini C, Anselmo L, Moe K, Moe MM (2010) Drag and energy accommodation coefficients during sunspot maximum. Adv Space Res 45(5):638–650 doi:10.1016/j.asr.2009.08.034
29. Picone JM, Emmert JT, Lean J (2005) Thermospheric densities derived from spacecraft orbits—I, Accurate processing of two-line element sets. J Geophys Res 110:A03301, doi:10.1029/2004JA010585
30. Pilinski MD, Argrow BM, Palo SE (2010) Semiempirical model for satellite energyaccommodation coefficients. J Spacecr Rockets 47(6):951 doi:10.2514/1.49330
31. Pilinski MD, Argrow BM, Palo SE (2011) Drag coefficients of satellites with concave geometries: Comparing models and observations. J Spacecr Rockets 48(2):312–324
32. Prussing JE, Conway BA (1993) Orbital Mechanics. Oxford University Press, Oxford
33. Schaaf SA, Chambré PL (1961) Flow of rarefied gases, Princeton Aeronautical Paperbacks. Princeton University Press, Princeton
34. Sentman LH (1961) Free molecule flow theory and its application to the determination of aerodynamic forces, LMSC–448514, Lockheed Missiles and Space Company
35. Sutton EK, Forbes JM, Nerem RS (2005) Global thermospheric neutral density and wind response to the severe 2003 geomagnetic storms from CHAMP accelerometer data. J Geophys Res 110: A09S40, doi:10.1029/2004JA010985
36. Sutton EK (2009) Normalized force coefficients for satellites with elongated shapes. J Spacecr Rockets 46(1):112–116 doi:10.2514/1.40940
37. Sutton EK, Nerem RS, Forbes JM (2007) Density and winds in the thermosphere deduced from accelerometer data. J Spacecr Rockets 44(6):1210–1219 doi:10.2514/1.28641
38. Sutton EK, Nerem RS, Forbes JM (2008) Reply by the authors to g. koppenwallner. J Spacecr Rockets 45(6):1328–1329
39. Tapley BD, Schutz BE , Born GH (2004) Statistical orbit determination. Academic, New York
40. Thomson WT (1986) Introduction to space dynamics, Dover
41. Vallado DA (2001) Fundamentals of astrodynamics and applications, 2nd edn. Microcosm Press, El Segundo
42. Vokrouhlický D, Farinella P, Mignard F (1993) Solar radiation pressure perturbations for earth satellites, i. a complete theory including penumbra transitions. Astron Astrophy 280:295–312
43. Vokrouhlický D, Farinella P, Mignard F (1994) Solar radiation pressure perturbations for Earth satellites, II. an approximate method to model penumbra transitions and their long-term orbital effects on LAGEOS. Astron Astrophys 285:333–343

44. Vokrouhlický D, Farinella P, Mignard F (1994) Solar radiation pressure perturbations for Earth satellites, III. global atmospheric phenomena and the albedo effect. Astron Astrophys 290: 324–334
45. Vokrouhlický D, Farinella P, Mignard F (1996) Solar radiation pressure perturbations for Earth satellites, IV. effects of the earth's polar flattening on the shadow structure and the penumbra transitions. Astron Astrophys 307:635–644
46. Ziebart M, Adhya S, Sibthorpe A, Edwards S, Cross P (2005) Combined radiation pressure and thermal modelling of complex satellites: algorithms and on-orbit tests. Adv Space Res 36(3):424–430 doi:10.1016/j.asr.2005.01.014
47. Ziebart M (2004) Generalized analytical solar radiation pressure modeling algorithm for space-craft of complex shape. J Spacecr Rockets 41(5):840–848

Chapter 4
Producing Density and Crosswind Data from Satellite Dynamics Observations

In this chapter, the tools and models discussed in the previous chapters will be used to derive density and wind data sets from satellite dynamics observations.

In Sect. 4.1, two-line element (TLE) data on orbits of satellites and space debris objects will be used to derive density data. The main advantages of this technique are its applicability to many space objects in different orbits simultaneously, the low latency of data availability and long time span of coverage into many years in the past. A major disadvantage is the coarse temporal resolution.

In Sect. 4.2 a newly developed algorithm is presented, which can be applied to process density and crosswind data from the precise accelerometers on satellites such as CHAMP and GRACE. These satellites deliver accurate data on thermospheric density and wind speed at a very high temporal resolution and with excellent global coverage.

To conclude, Sect. 4.3 provides a more concise discussion on how software for precise orbit determination can be applied to derive density data from precise tracking data of certain scientific and operational Earth observation satellites. Only a limited number of satellites with the right tracking instrumentation is available. However, the accuracy and temporal resolution will be higher than for the TLE-derived data. This method is not extensively applied in the scientific analysis in the remainder of this thesis. However, this option is also of importance for the calibration of accelerometer measurements.

4.1 Density Determination Using Two-Line Element Data

The United States Strategic Command (USSTRATCOM) distributes orbital information on a very large number of space objects, in the form of two-line element sets (TLEs). The TLE is an extremely compact format, designed for the constraints of the information infrastructure of the 1960s, consisting of just two lines of 69 characters. Figure 4.1 shows a single TLE as an example, identifying the Kepler elements and force model parameters contained in them.

Fig. 4.1 Example two-line element set with descriptions of the data fields (after http://www.space
-track.org). The TLE set shown is the first available after launch of object 22659, which was also
featured in Figs. 1.2 and 1.3

These Kepler elements and force model parameters have been estimated by fitting
an orbital model to space surveillance tracking data. For low Earth objects, this orbital
model is implemented in the Simplified General Perturbations 4 (SGP4) propagator.
SGP4 is described in detail by Hoots and Roehrich [36]. Many issues concerning
its history and use in combination with TLE data are provided by Vallado et al.
[82]. These two references also provide access to SGP4 source code. Vallado and
Crawford [81] provide more details on the propagator's accuracy, as well as source
code to use SGP4 to generate TLEs from other orbit ephemeris data.

Each of the two lines of a TLE contains the object identifier. The first TLE line
contains the epoch and several force model parameters. The second TLE line contains
fitted Kepler element data. The epoch is set at the last ascending equator crossing
before the end of the tracking data fit span.

The first and second time derivative of the mean motion were used to model drag
in an older version of the SGP model. In SGP4, these parameters are ignored, since
a more sophisticated drag model is used, in which the parameter $B*$ compensates
uncertainties in the ballistic coefficient of the satellite and in the density.

Note that the Kepler elements in the TLE do not represent the true orientation
and shape of the satellite orbit at the provided epoch, since part of the gravity-
induced orbit variations are modeled by the SGP4 propagator. The trajectories of
space objects that were fitted to the observation data can only be reconstructed by
the user by applying the SGP4 algorithm together with the TLE data.

The database of publicly available TLEs is updated continuously. For most objects
in orbits of interest for density derivation, updates are currently provided at least daily.
TLEs are also conveniently available backwards in time up to the early years of the
space age, although the update frequency and accuracy for the earlier data were lower
than the current level.

Figure 1.2 in the Introduction already demonstrated that there is a relation between
the orbit information contained in TLEs and thermosphere density variations along
the satellite's trajectory. Chapter 3 provided the mathematical description of this
relation. The data processing aspects required to complete the density derivation will
be provided in the following sections.
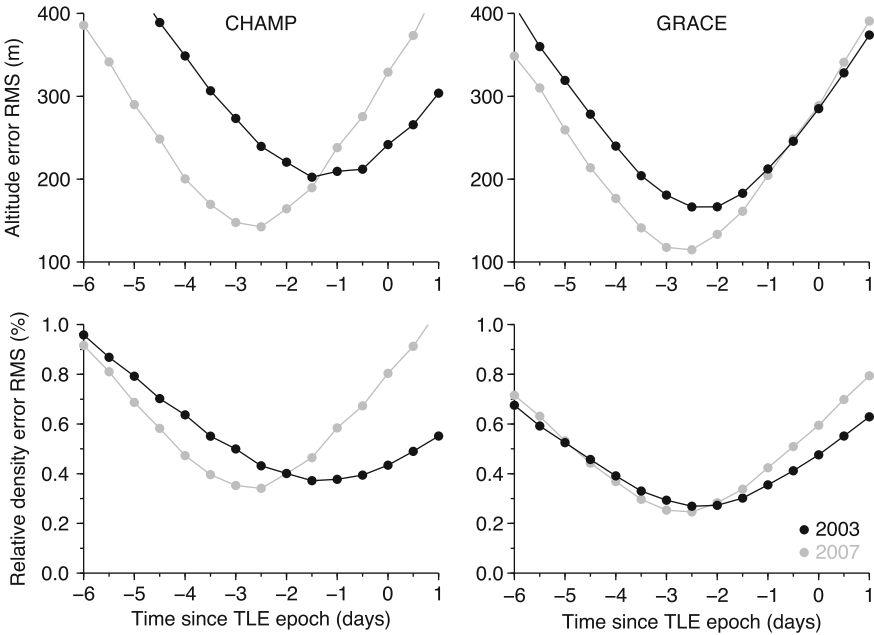
Fig. 4.2 RMS altitude error and orbit induced relative density error for CHAMP (*left*) and GRACE (*right*) TLEs

4.1.1 TLE Orbit Accuracy

Before describing the data processing strategy, it is worthwhile to have a closer look at the accuracy of the TLE orbit data, and its effect on the accuracy of density evaluation. This analysis was performed for the CHAMP and GRACE satellites, by comparing the output of the SGP4 algorithm with orbits computed using GPS tracking data. The GPS orbits used are the CHAMP rapid science orbits [42] and GRACE navigation solutions [11], which are accurate to the decimetre and few centimetre level, respectively. As we shall see, the TLE orbit errors are several orders of magnitude larger. The difference between the two orbits is therefore a good measure of the TLE orbit error.

Since the height component of the orbit is the most important for density studies, only this component is considered in the orbit error evaluation. The value for each evaluation epoch is evaluated for several surrounding TLEs. The orbit differences over two entire years, 2003 and 2007, free of orbit manoeuvres for both missions, are binned as a function of the time difference between the SGP4 evaluation epoch and the TLE epoch and plotted in the top frames of Fig. 4.2.

Densities from the NRLMSISE- 00 model were evaluated at each computed orbit position as well. The difference between the density at the TLE orbit position and the density at the precise orbit position are expressed as a percentage of the latter value. These data are processed in the same way as the orbit differences, and are plotted in

the bottom panels. They can be used as a metric to evaluate the effect of TLE orbit error on density model evaluation.

According to the Figure, the TLE orbits are most accurate when evaluated 1–3 days before the TLE epoch. The accuracy is then at a level of 100–200 m, depending on the satellite and on the year. These TLE orbit errors result in minimum RMS density errors of the order 0.25–0.40%. On either side of this optimal time offset, the orbit and density errors steadily increase.

This offset of the most accurate SGP4 evaluation epoch with respect to the TLE epoch can be explained by the observation that the fitted TLE orbit is likely to be most accurate where it is strongly constrained by equal amounts of tracking data on both sides, which is often at or near the centre of the tracking data interval. On the other hand, the TLE epoch is defined at the end of this tracking data span. Under this assumption, the tracking data interval length used at Air Force Space Command to determine the TLE parameters can be estimated at about double this offset, so of the order of 2–6 days. The differences between the 2003 and 2007 averages, as well as the differences between CHAMP and GRACE suggest that the length of this tracking data span might be changed from time to time, and varies from satellite to satellite. Unfortunately, data on the tracking data span is not supplied with the TLE data, otherwise this information could be used to select the optimal TLE for orbit evaluation from the TLE time series.

According to the Figure, the GRACE TLE orbits are more accurate than the CHAMP orbits, and for both missions the orbits are more accurate in 2007 than in 2003. This is likely caused by the lower, less variable and therefore easier to model drag perturbation forces on the higher altitude GRACE compared to CHAMP and at lower solar activity in 2007 compared with 2003. A contribution to this improvement due to possible changes made in the SSN observations or the TLE data processing strategy at Air Force Space Command in between these years cannot be excluded either.

In practical applications of the SGP4 algorithm in combination with TLE data, the time difference between the evaluation epoch and the TLE epoch will often not be near the optimum value. This is due to the uncertainty in the tracking span length and the finite epoch spacing between subsequent TLEs in a time series. Fortunately, as the bottom panels of the Figure show, the density error due to TLE orbit error stays below 1%, even if the time with respect to the optimal epoch amounts to several days.

This is true for CHAMP and GRACE, and may be considered valid for satellites experiencing similar or lower levels of drag as well. However, for satellites with very low perigees, rapid changes in atmospheric density have a larger effect on the orbit, which can not always be captured in the TLE-determination spanning multiple days. The TLE orbit errors and orbit-induced density errors are likely to be much larger for such objects, and this could cause problems in the application of these data for density studies. Unfortunately, precise orbit data for such objects were not available, making it difficult to further investigate this possible limitation of the use of TLE data for thermosphere studies at this time. The future analysis of CHAMP data during the final weeks or months before its re-entry in September 2010 will be helpful in this regard.

4.1.2 History of TLE Density Processing

King-Hele [41] described the processing of TLE data for investigations of the thermosphere, but his method, which involved interpretation using an analytical atmospheric model and various other approximations, was not widely employed. One application has been the investigation of long-term trends in the thermosphere by Keating et al. [40].

Pardini et al. [67] instead use an orbit computation software package [66] to fit, in a least-squares sense, a drag coefficient to the TLE-derived orbit decay of selected satellites, using a complete model of the forces on the satellite. Their approach does not result in observations of density, but instead uses the JB2006 model, to estimate the energy accommodation coefficient parameter used in satellite aerodynamic calculations (see Sect. 3.4.3). A similar approach based on orbit determination using TLE orbits as tracking data was used by a US-Russian group of researchers [12], investigating empirical density model calibration (see Sect. 2.5).

The publication of a more efficient TLE to density processing algorithm by Picone et al. [70] has led to a wider range of research applications. These have included studies of long-term change [22, 24], a validation of density data derived from dayside limb UV airglow observations [23], an analysis of solar forcing of the thermosphere [46] and the study of empirical density model calibration [15, 16], which will be discussed in Chap. 6. The Picone et al. [70] algorithm will be described in the remainder of this section, and its results will be discussed in the next chapters.

4.1.3 Density Derivation From TLE Mean Mean Motion Data

A summary of Picone et al. [70] will be provided below. The discussion will start by covering some particulars of the TLE data and SGP4 algorithm in more detail. The presentation of the equations used to derive density from these data is followed by examples of calculation results, and a discussion of data quality and data editing aspects.

General Perturbations and Special Perturbations Orbit Determination

Picone et al. [70] start their paper by discussing the distinction between *special perturbations* and *general perturbations* orbit determination techniques. The SGP4 propagator used in TLE processing is an implementation of the general perturbations technique, in which a low order analytical solution of the perturbed equations of motion is used, and fitted to the tracking data. In contrast, the special perturbations or precise orbit determination technique involves high order numerical integration of more detailed and highly parameterised force models, and using precise measurement models to process the tracking data. This method, which is discussed in

more detail in Sect. 4.3, can lead to significantly more accurate orbit and density results. On the other hand, this accuracy can only be reached for satellites equipped with suitable precise tracking instrumentation, as opposed to the ground-based radar tracking from the Space Surveillance Network employed in TLE production. In addition, the special perturbations approach requires a much larger investment in terms of force and measurement modelling effort and computation time, compared to general perturbations.

Osculating and Mean Orbital Elements

Another important distinction made in the introductory sections of Picone et al. [70] is that between osculating and mean orbital elements. Internally, the SGP4 propagator uses the analytic equations for low order orbit perturbations by Brouwer [3], and the atmospheric drag model by Lane and Cranford [43], to describe the continuously changing shape and orientation of the elliptical orbit. These are expressed in terms of osculating (or instantaneous) orbital elements, which can be directly converted to Cartesian position and velocity values at a desired epoch by applying the geometrical transformations between these two coordinate systems [63, 80].

The TLEs themselves, as shown in Fig. 4.1, contain fitted mean elements at epoch. In the fitting process, the periodic perturbations have been removed from these elements. This makes these values in the TLE data unsuitable for direct conversion to position and velocity values, either at arbitrary epochs or at the TLE epoch itself. The SGP4 propagator must always be applied to the TLE data to reconstruct the instantaneous position and velocity at its intended level of accuracy.

Derivation of the TLE to Density Processing Algorithm

Making use of the perturbation Eq. 3.38 for the semi-major axis, Picone et al. [70] show that, for the purposes of density derivation from TLEs, the focus should be on secular variations of the semi-major axis over multiple orbital revolutions. Furthermore, the paper extensively discusses the important distinction between the osculating semi-major axis a and the mean semi-major axis a_M, as defined by the SGP4 algorithm, from which contaminating perturbations due to the Earth's gravity field have been removed. This leaves only drag and radiation pressure to change the mean semi-major axis a_M. The change of this variable over time can therefore be expressed according to the sum of the perturbation Eqs. 3.38 and 3.40.

$$\frac{\mathrm{d}}{\mathrm{d}t} a_M \cong -\frac{a_M^2}{\mu} B\rho F v^3 + \frac{\mathrm{d}a}{\mathrm{d}t}\bigg|_R \tag{4.1}$$

Using Eq. 3.4, the mean semi-major axis can be converted to the mean mean motion n_M. This conversion is useful because n_M can be found directly in the TLE data.

The resulting differential equation is

$$\frac{\mathrm{d}}{\mathrm{d}t} n_\mathrm{M} \cong \frac{3}{2} \mu^{-\frac{2}{3}} n_\mathrm{M}^{\frac{1}{3}} B \rho F v^3 + \frac{\mathrm{d}n_\mathrm{M}}{\mathrm{d}t}\bigg|_\mathrm{R} \tag{4.2}$$

When the density is large enough, radiation pressure is negligible compared to drag, and the last term in these equations can be omitted. Since this is often the case for low orbiting satellites, the derivation in Picone et al. [70] is continued under this assumption. However, at higher altitudes and low solar activity, the radiation pressure acceleration will become a major limiting factor in the accurate application of the algorithm.

Going back to the derivation, Eq. 4.2 can now be numerically integrated over the time interval between a pair of TLEs, with the epochs t_i and t_k.

$$n_\mathrm{M}(t_k) - n_\mathrm{M}(t_i) \approx \frac{3}{2} \mu^{-\frac{2}{3}} \oint_{t_i}^{t_k} n_\mathrm{M}^{\frac{1}{3}} B \rho F v^3 \ \mathrm{d}t \tag{4.3}$$

From this equation, the density observable can be derived. This observable is the velocity-weighted average density along the object's trajectory over the time interval between t_i and t_k. It is assigned the timestamp $t_{ik} \equiv \frac{1}{2}(t_i + t_k)$, and is defined as

$$\bar{\rho}_\mathrm{O}(t_{ik}) \equiv \frac{\frac{2}{3} \mu^{\frac{2}{3}} [n_\mathrm{M}(t_k) - n_\mathrm{M}(t_i)]}{\oint_{t_i}^{t_k} n_\mathrm{M}^{\frac{1}{3}} B F v^3 \ \mathrm{d}t} = \frac{\oint_{t_i}^{t_k} n_\mathrm{M}^{\frac{1}{3}} B \rho F v^3 \ \mathrm{d}t}{\oint_{t_i}^{t_k} n_\mathrm{M}^{\frac{1}{3}} B F v^3 \ \mathrm{d}t} \tag{4.4}$$

Next, Picone et al. [70] approximate the integrals of the ballistic coefficient B and mean motion n_M over the time interval by constants. For the ballistic coefficient, either an estimated value, or a model calculation (as described in Chap. 3) can be applied. The result will be an average value $\bar{B}$, which is constant over the time interval. The mean mean motion $n_\mathrm{M}(t)$ during the time interval is approximated by the average of the two values at the endpoints.

$$\bar{n}_\mathrm{M}(t_{ik}) \approx \frac{n_\mathrm{M}(t_i) + n_\mathrm{M}(t_k)}{2} \tag{4.5}$$

These substitutions result in

$$\bar{\rho}_\mathrm{O}(t_{ik}) \cong \frac{\frac{2}{3} \mu^{\frac{2}{3}} [n_\mathrm{M}(t_k) - n_\mathrm{M}(t_i)]}{\left[\frac{1}{2} n_M(t_i) + \frac{1}{2} n_M(t_k)\right]^{\frac{1}{3}} \bar{B} \oint_{t_i}^{t_k} F v^3 \mathrm{d}t} \tag{4.6}$$

This equation can be evaluated using the two values of n_M which are conveniently available from the pair of TLEs, in combination with the integrated velocities obtained from using the TLE data with the SGP4 propagator.

An equivalent velocity weighted average model density along the object's trajectory over the time interval is defined as

$$\bar{\rho}_{\mathrm{M}}(t_{ik}) \equiv \frac{\oint_{t_i}^{t_k} \rho_{\mathrm{M}} F v^3 \, \mathrm{d}t}{\oint_{t_i}^{t_k} F v^3 \, \mathrm{d}t} \tag{4.7}$$

In the evaluation of this equation, the positions of the object obtained from the SGP4 propagator are used as inputs to the density model, in order to calculate the instantaneous densities $\rho_{\mathrm{M}}(t)$ at each time step required for the numerical computation of the integral.

4.1.4 Practical TLE Processing Considerations

A number of practical considerations has to be made when processing the TLE data into density observations using the algorithm described in the previous section. These concern the selection of suitable space objects, selection and editing of their TLEs, the choice of value for the ballistic coefficients $\bar{B}$ and the choice of the integration time interval. These will be shortly discussed below.

Selection and Editing of Suitable Objects and Their TLEs

Emmert [20] defined a comprehensive series of computational tests to select suitable space objects from the space tracking catalogue, and to edit out unsuitable data from TLE time series of specific objects. For some satellites, the TLE mean motion time series shows erratic variations, because apparently SSN tracking of another object is mistakenly used. The method is intended to filter out such data as well as data for which the object apparently has unsuitable variations in cross-sectional area or undergoes orbit manoeuvres.

Ballistic Coefficients for TLE Objects

For almost all remaining suitable TLE objects, accurate information on the object geometry and attitude is not easily available, which makes it impossible to calculate the aerodynamic coefficient in detail, as described in Sect. 3.5. This problem is circumvented by assuming that the ballistic coefficient is an object-specific constant. Based on a comparison between calculations for various objects at different levels of solar activity, the errors introduced by this assumption are at the level of a few percent [14].

A value for the ballistic coefficient can be determined for each satellite by computing a long-term data over model density ratio time series, and picking the ballistic coefficient which makes the mean of the ratios equal to 1.0. Of course, the resulting TLE-derived densities will then inherit any bias present in the model that is used.

Emmert [20] uses a more sophisticated approach, by calculating the ballistic coefficient for each space object based on the minimisation of density differences between

overlapping data of different objects. A single object, preferably one for which the ballistic coefficient can be calculated quite precisely, is chosen as a reference object. Emmert [20] chose the Starshine I sphere with a C_D of 2.1. However, Pilinski et al. [71] has pointed out that the Starshine spheres are not the most suitable objects for drag computations, due to the presence of attachment rings required for integration with the launch vehicle, protruding mirror plates, and cylindrical elements as part of the construction used to attach the two half-spherical elements. These deviations from a perfect aerodynamic sphere hinder a simple computation of cross-sectional area and drag coefficient. The inferred value used for the spherical area of Starshine I by Emmert [20] is about 5% smaller than the mean area of the more detailed satellite model evaluated by Pilinski et al. [71] ($A = 0.193\,\mathrm{m^2}$). Assuming that the reported mass ($m = 39.46\,\mathrm{kg}$) and assumed drag coefficient ($C_D = 2.1$) for Starshine I are correct, the TLE-derived densities derived by Emmert [20] will be scaled too large by the same amount.

Although the selection of values for the satellite area and mass might at first seem a rather easy part of the drag analysis problem, the example above shows that caution is required, even for seemingly trivial satellite shapes. Values provided in scientific literature and mission documentation for other satellites could well have similar, as yet undetected, issues with their accuracy.

The data selection, editing and ballistic coefficient computation schemes of Emmert [20] ensures a level of consistency within a large set of objects, allowing for the creation of a climatological analysis [21]. However, this scheme has not yet been implemented for the work presented in this thesis. Instead, objects and ballistic coefficients were chosen based on values computed by other analysts, which were found in earlier literature [1, 2, 22, 70].

Choice of Integration Time Interval

A further choice required for the application of the algorithm by Picone et al. [70] concerns the time interval between the two TLEs with epochs t_i and t_k, from which the mean motion data are used for the computation of the density observable, using Eq. 4.6.

The irregular time interval between epochs at which consecutive TLEs are computed and delivered by Air Force Space Command is currently less than one day for most low orbiting objects. However, as shown in Sect. 4.1.1, the tracking data fit span used to generate the TLEs is selected from two days or more prior to the TLE epoch. This means that there will be some overlap in the information content of multiple TLEs when their epochs are closer together in time than the length of this fit span.

Density observations derived from such closely spaced TLE pairs show noise-like variations, apparently related to the fact that observation, processing and other numerical errors in the mean motion data are larger than the drag-induced signal. On the other hand, since the observable represents an average density, a choice for long time intervals means that short-term fluctuations in density cannot be recovered, due to the smoothing inherent in this averaging process.
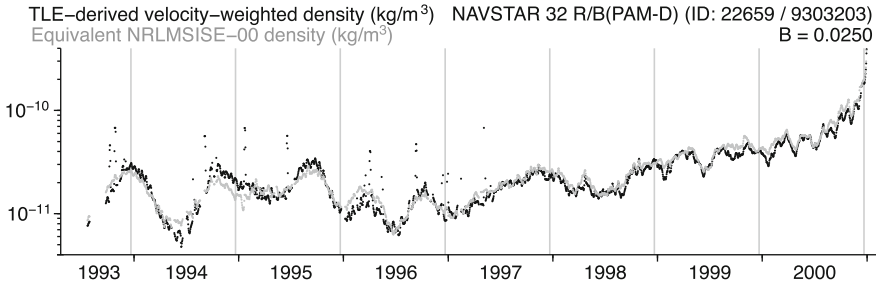
Fig. 4.3 Time series of TLE-derived velocity-weighted densities for the PAM- D object 22659, which was also featured in Figs. 1.1,1.2 and 1.2

Emmert [20] therefore selects for each TLE the first available later TLE for which the time interval is longer than three days. A maximum time interval between subsequent TLEs of 5 days is also imposed. However, such long delays between the availability of subsequent TLEs do not occur very frequently in the present data.

For this study, a different approach is used. The mean motion data from the TLEs is linearly interpolated and evaluated at regular time intervals, for instance for every integer day. This approach will be used to evaluate the influence of the integration time interval on the accuracy of the TLE-processing results in Sect. 5.3.

4.1.5 Example of TLE Density Processing

Figure 4.3 provides time series of computed values of $\bar{\rho}_O$ and $\bar{\rho}_M$, resulting from the evaluation of Eqs. 4.6 and 4.7. The values are computed for the PAM- D rocket stage that was featured earlier in Chap. 1. The general agreement between observation and model is clearly visible. Nevertheless, there are some clear outliers during the first years of the TLE-derived series, until mid-1997. These are likely due to erroneous TLE-data.

The densities derived from the TLE data will be further compared and analysed in the following chapters. In this chapter, we will move on to the presentation of other ways of deducing information on the aerodynamic interaction of the satellite with the neutral atmosphere from orbital dynamics data, starting with accelerometer data.

4.2 Density and Crosswind from Accelerometer Data

Accelerometers carried by low-Earth orbiters such as CHAMP, GRACE, GOCE and the future Swarm satellites, provide important data for improving our understanding of thermospheric density and winds. The CHAMP and GRACE missions were not

designed for studies of the thermosphere; they carry accelerometers in order to allow for the removal of non-gravitational signals from measured orbit perturbations due to inhomogeneities in the Earth's gravity field. Nevertheless, their application to thermosphere studies has resulted in density and wind data sets containing information at unprecedented levels of detail and coverage.

Analyses of accelerometer-derived density data sets resulted in the publication of a large number of scientific papers, on topics including the response to drivers such as solar EUV variability [32, 64], geomagnetic activity and storms [10, 26, 44, 45, 48, 54, 61, 74], substorms [73], Joule heating [75], solar flares [52], solar wind streams [47]; and on phenomena such as the equatorial mass density anomaly [49–51, 58], medium- to large-scale variability [7], upwelling in the cusp region [55, 72], travelling atmospheric disturbances [6, 8], solar terminator waves [27, 53, 62], atmospheric tides [28, 65] and wave structures at solar minimum [9]. These investigations generally made use of density data processed using algorithms published by Bruinsma and Biancale [5], Bruinsma et al. [4], Sutton et al. [74], Sutton et al. [76], Liu et al. [49] and Rentz and Lühr [72].

Investigations of wind results from the CHAMP mission have started to appear in recent years as well. These papers can be classified by their focus on either the equatorial to mid-latitudes [33, 34, 50, 53, 57] or the polar regions [30, 31, 56]. An important early paper on wind derivation from accelerometer data is that by Marcos and Forbes [59] who analysed triaxial accelerometer measurements from the SETA instrument flown on several spacecraft in the early 1980s. Recent publications on the derivation of winds from CHAMP data are provided by Liu et al. [50], Sutton et al. [75].

This section contributes to this research topic by presenting an improved, more generally applicable iterative density and wind derivation algorithm, recently published by Doornbos et al. [18]. The basic problem to be solved by such an algorithm is that the accelerometer delivers at most three orthogonal acceleration observations. However, there are at least four unknowns: Density and three orthogonal wind velocity components. The solution to this problem lies in the fact that the orbital velocity of the satellite is much greater than the wind velocity. The error in the total relative velocity, and therefore in the density, can be kept limited by assuming that the in-track wind speed is zero, or by applying an in-track wind value from an empirical model. Eliminating this unknown from the equation enables the determination of reasonably accurate density values, as well as wind speed components orthogonal to the in-track direction.

4.2.1 Accelerations and Velocities

Figure 4.4 provides a simple schematic view of the three vectors of importance for a density and wind retrieval algorithm: The observed and modelled aerodynamic accelerations a_{obs} and a_{mod} are shown originating in the centre of mass of the satellite. In addition, the relative velocity v_r of the atmosphere with respect to the spacecraft
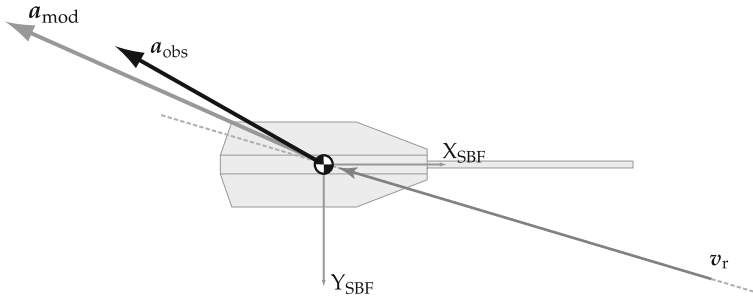
Fig. 4.4 Relative velocity, modelled and observed accelerations in the CHAMP spacecraft body-fixed (SBF) XY plane. CHAMP is viewed from the top

is shown. This quantity is partly observed and partly modelled, as will be explained below. Note that due to the asymmetrical shape of the satellite with respect to the flow, the acceleration is in general not exactly aligned with the relative velocity, as indicated by the dashed guide lines.

The modelled and observed aerodynamic acceleration vectors initially do not match in magnitude and direction. This is mainly due to the approximate character of the modelled density and wind speed. It is the purpose of the new algorithm to find those density and wind values, which, when replacing the original values, make these accelerations match.

Before more detailed descriptions of the previous and new algorithms are provided, the relationship between the parameters in Fig. 4.4 and the way in which they can be obtained from satellite observation data sets and models will be described. The description in the following sections will refer to the instruments and data products of the current generation of accelerometer missions (CHAMP and GRACE) and to the external models (atmospheric models, force models, etc.) that are currently available. The new algorithm is not limited by the use of these data sets and models, however. It can be applied just as well to equivalent data from historical or future accelerometer missions and using future improvements to external models.

Relative Velocity

The relative velocity of the atmosphere with respect to the spacecraft was described in Sect. 3.2.1 as the sum of contributions from the inertial velocity of the spacecraft in its orbit, the velocity caused by the corotating atmosphere, and the velocity of winds, with respect to an Earth-fixed atmosphere. For the accelerometer satellites, these can be expressed in satellite body-fixed (SBF) coordinates, as

$$\boldsymbol{v}_{\mathrm{r}} = -R_{\mathrm{ib}}\boldsymbol{v} + R_{\mathrm{ib}}(\boldsymbol{\omega}_{\oplus} \times \boldsymbol{r}) + R_{\mathrm{ib}}R_{\mathrm{ei}}R_{\mathrm{le}}\boldsymbol{v}_{\mathrm{w}}, \qquad (4.8)$$

in which the rotation matrix R_{ib} from the inertial to the satellite body-fixed frame is obtained from star camera observations; the inertial satellite position and velocity,

r and v, are obtained by precise orbit determination using tracking observations from the satellite's GPS receiver; $\omega_\oplus$ is the Earth's angular velocity vector. The wind velocity v_w is usually specified in local coordinates (East, North, Vertical). The transformation to satellite body-fixed coordinates requires the additional rotation matrices R_le from local to the Earth-fixed cartesian coordinate system, based on the satellite's latitude and longitude [63], and the rotation matrix R_ei from Earth-fixed to inertial coordinates, based on Earth orientation parameters [60]. The first two contributors to v_r, the orbit and corotation velocities, are known at a much higher accuracy than the wind velocity. If model values for v_w are required, these can be obtained from a wind model, such as [19].

In the description of the density and wind retrieval algorithms, we will use the notation $v_{\mathrm{r},0}$ or $v_{\mathrm{r},i=0}$ to indicate an initial guess of the relative velocity, by either neglecting winds or using a wind model. The notation $v_{\mathrm{r},i}$ will designate a relative velocity which already includes an accelerometer-derived wind component.

Observed Aerodynamic Acceleration

The observed aerodynamic acceleration a_obs is obtained from the raw accelerometer data after calibration and removal of non-aerodynamic acceleration signals. Details on the accelerometer instruments of CHAMP and GRACE, their performance and processing of raw data can be found in Touboul et al. [79]. The data is delivered to science users in the form of Level 2 products for CHAMP [29] and Level 1B products for GRACE [11].

Calibration is performed by multiplying the acceleration vector with a 3×3 diagonal scale factor matrix, and adding a bias vector:

$$a_\mathrm{cal} = S a_\mathrm{raw} + a_\mathrm{bias} \qquad (4.9)$$

The scale factors can often be considered nearly constant [78], while the bias is known to vary on timescales of days and more, under the influence of ageing effects and temperature variations. Changes to the satellite software or switches between the redundant on-board electronics parts can cause abrupt changes in the calibration parameters. The determination of the calibration parameters used in this study for the in-track accelerometer observations is described by Van Helleputte et al. [35], who made use of GPS observations. This method was found to be not sufficiently accurate for the cross-track accelerometer observations, for which an alternative method was applied, as discussed at the end of Sect. 4.2.4.

Various non-aerodynamic signals should be removed from the accelerometer data, including accelerations due to activity of cold gas thrusters for attitude control. If a set of two opposing thrusters is not perfectly balanced, as is often the case, they introduce a residual signal in the linear acceleration, besides the intended angular acceleration. Data around the activation times of these thrusters should therefore be removed. A less obvious example of accelerations that should be removed from the data are those due to mechanical forces caused by electrical current changes on the satellite [25].

Finally, modelled accelerations due to radiation pressure from the Sun $\boldsymbol{a}_{\mathrm{srp}}$, Earth albedo $\boldsymbol{a}_{\mathrm{alb}}$ and Earth infrared radiation $\boldsymbol{a}_{\mathrm{IR}}$ are computed and removed from the calibrated and edited accelerometer data $\boldsymbol{a}_{\mathrm{cal}}$, to arrive at the observed aerodynamic acceleration vector $\boldsymbol{a}_{\mathrm{obs}}$.

$$\boldsymbol{a}_{\mathrm{obs}} = \boldsymbol{a}_{\mathrm{cal}} - \boldsymbol{a}_{\mathrm{srp}} - \boldsymbol{a}_{\mathrm{alb}} - \boldsymbol{a}_{\mathrm{IR}} \tag{4.10}$$

Modelled Aerodynamic Acceleration

The modelled aerodynamic acceleration vector $\boldsymbol{a}_{\mathrm{mod}}$ was extensively discussed in the previous chapter. For the derivation of density and winds, the full vector form of Eq. 3.8 is used.

$$\boldsymbol{a}_{\mathrm{mod}} = \boldsymbol{C}_{\mathrm{a}} \frac{A_{\mathrm{ref}}}{m} \frac{1}{2} \rho v_{\mathrm{r}}^2 \tag{4.11}$$

The mass m can be obtained by subtracting from the satellite launch mass the amount of cold gas used, which is logged in the satellite's housekeeping data, and made available as a data product.

4.2.2 Direct and Iterative Processing Algorithms

Direct Algorithms for CHAMP and GRACE

Previously published algorithms made use of assumptions about the orientation of the accelerometer in space. For CHAMP and GRACE, the accelerometer instruments are carefully mounted near the satellite centre of mass, with an orientation so that their three axes can be considered perfectly aligned with the spacecraft body-fixed (SBF) axes. The spacecraft are under active attitude control, which keeps these axes within a few degrees of the orbit-fixed along-track, cross-track and radial directions (see Fig. 4.5). The relative orientation of these axes can be expressed in roll, pitch and yaw Euler angles.

Because these Euler angles are relatively small, the inertial orbital velocity of the satellite is kept closely aligned with the X_{SBF} axis. In Sect. 3.3, it was shown that accelerations in the velocity direction are the most effective in changing the orbital energy, and therefore have a much larger effect on the orbit than accelerations of similar magnitude in perpendicular directions. This means that the X_{SBF} axis of the accelerometer can be more accurately calibrated using positioning data from the GPS instrument [35] than the Y_{SBF} and Z_{SBF} axes, even without taking into account the larger measured signal. This consideration leads to an approach for density determination [4, 5, 76] where only the projection of the aerodynamic acceleration on the X_{SBF} axis is used, as shown schematically in Fig. 4.6. The density can then be solved directly from the X-component of the vector Eq. 4.11:
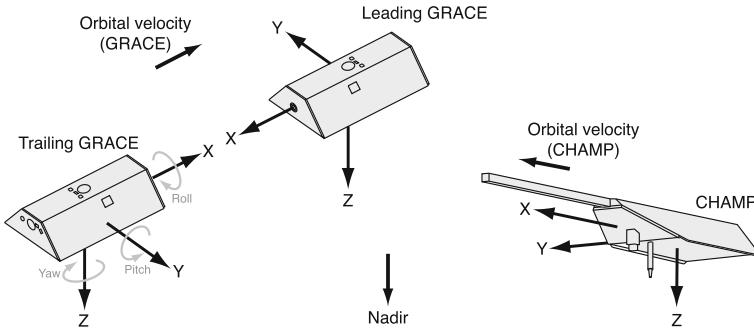
Fig. 4.5 Definition of spacecraft body-fixed (SBF) axes for GRACE and CHAMP. The nadir and velocity directions are shown for the satellite's nominal attitude
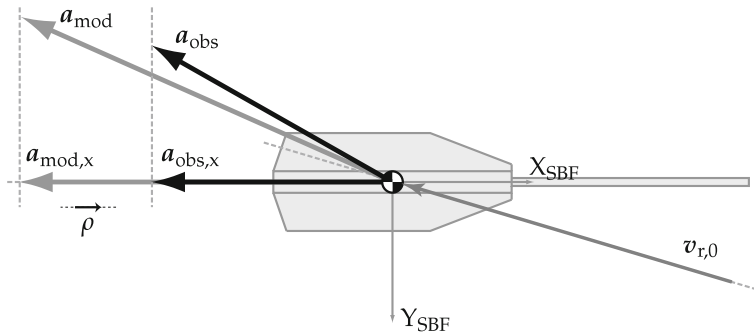
Fig. 4.6 Schematic representation of the determination of density from the projection of acceleration on the X_{SBF} axis

$$\rho = \frac{2m}{A_{\text{ref}} v_{\text{r},0}^2} \frac{a_{\text{obs,x}}}{C_{\text{a,x}}} \tag{4.12}$$

Information from the acceleration component in the Y_{SBF} direction, closely aligned with the cross-track direction, can be used to derive data on the wind speed in that direction. Sutton et al. [76] describe two approaches. In the first approach, C_{a} in Eq. 4.11 is expanded using the analytical equations for a panel model, using equations equivalent to (3.62). The resulting expansion is quadratic with respect to $v_{\text{w,y}}$, which can then be solved, resulting in an expression depending on $a_{\text{obs,y}}$ and ρ. Sutton names this approach the single axis method, even though information from both the X- and Y-axes is required, if ρ is to be substituted from Eq. 4.12.

The second approach is named the dual-axis method by Sutton et al. [76], and can be found in an earlier paper by Liu et al. [50] as well. The method requires that the lift and sideways forces are negligible, or are modelled and removed from the acceleration beforehand, so that only the observed acceleration due to drag $a_{\text{obs,D}}$ remains. The authors do not specify exactly how the lift and sideways forces should
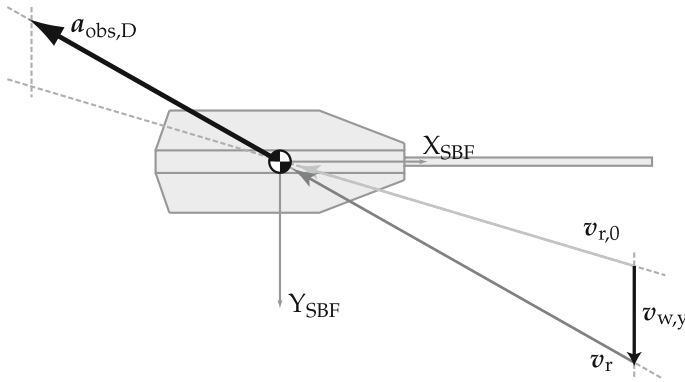
Fig. 4.7 Schematic representation of the determination of wind from the accelerometer Y_{SBF} axis

be modelled, but we have adopted the following approach: first, a new modelled aerodynamic acceleration is computed according to Eq. 4.11, now with the density from Eq. 4.12 and with the a-priori relative velocity $\boldsymbol{v}_{r,0}$ as inputs. This acceleration vector $\boldsymbol{a}_{\mathrm{mod}}$ can then be decomposed into a drag component, by projection on the relative velocity direction, and a perpendicular lift plus sideways force component, by subtraction of that drag component from the original modelled acceleration. In equations:

$$\boldsymbol{a}_{\mathrm{mod,D}} = (\boldsymbol{a}_{\mathrm{mod}} \cdot \hat{\boldsymbol{v}}_{r,0})\hat{\boldsymbol{v}}_{r,0}, \quad \boldsymbol{a}_{\mathrm{mod,L}} = \boldsymbol{a}_{\mathrm{mod}} - \boldsymbol{a}_{\mathrm{mod,D}} \tag{4.13}$$

The modelled lift plus sideways aerodynamic force $\boldsymbol{a}_{\mathrm{mod,L}}$ is then subtracted from the observed aerodynamic acceleration, to arrive at the observed drag.

$$\boldsymbol{a}_{\mathrm{obs,D}} = \boldsymbol{a}_{\mathrm{obs}} - \boldsymbol{a}_{\mathrm{mod,L}} \tag{4.14}$$

The velocity and drag acceleration are by definition in the same direction, so that the wind can be determined from a simple geometrical consideration (see Fig. 4.7). Expressed in the form of an equation, ρ and C_D disappear when the Y-component of Eq. 3.10 is divided by the X-component, and $v_{w,y}$ is solved for after substitution of Eq. 4.8, resulting in:

$$v_{w,y} = \frac{a_{\mathrm{obs,D,y}}}{a_{\mathrm{obs,D,x}}} v_{r,0,x} - v_{r,0,y} \tag{4.15}$$

A similar wind determination could in principle be performed for the Z_{SBF} axis. However, the aerodynamic acceleration in this direction is in general too small compared to errors in the instrument calibration, radiation pressure model and lift force model. In addition, on the CHAMP accelerometer, this Z_{SBF} component suffers from a malfunction which prevents the acquisition of accurate data [68].

Discussion of the Direct Algorithms

The schematic representations in Figs. 4.6 and 4.7 clearly show that when the angle between the relative velocity and the X_{SBF} axis gets larger, the errors in the density and wind increases. For the extreme case where this angle approaches 90 degrees, density values will approach zero, while the wind speed will go to infinity. In these figures, the angles are exaggerated for clarity, compared to the angles of CHAMP and GRACE under nominal attitude control. The roll, pitch and yaw Euler angles are kept within ± 1 degree for GRACE and ± 2 degrees for CHAMP. For the future Swarm mission, the attitude will likely be somewhat more loosely controlled, probably to within ± 4 degrees.

These attitude angles only determine the alignment of the body-fixed frame with the inertial velocity vector. The contributions to the relative velocity vector by the co-rotation of the atmosphere and thermosphere winds can be equally important. The atmospheric co-rotation velocity over the equator depends on the altitude, ranging from 483–502 m/s at 250–500 km. This increases the maximum angle between the relative velocity and the X_{SBF} axis by 3.6–3.8 degrees. The wind speed, which under most conditions is within the range of about 0–200 m/s, can reach peak velocities in the polar regions of up to 500–1,000 m/s [56, 30], causing the incidence angle to reach peak values of 8–10 degrees. In principle, the accuracy of the derived density and wind speed should be independent of these angles, but when using the direct approach this is not the case.

Another limiting factor of the direct algorithm results from the dependence of C_a on v_r. The methods use an initial value $v_{r,0}$, composed of the orbit and co-rotation velocity, and either neglect or model the in-track wind velocity. After the derivation of the cross-track wind $v_{w,y}$ however, there is a better estimate of the relative velocity:

$$v_{r,i=1} = v_{r,0} + v_{w,y} \qquad (4.16)$$

where the index i is an iteration counter. This new relative velocity leads to a new value of C_a (according to the equations in the previous chapter) and therefore to a new value of ρ. The change in C_a also leads to a change in the lift and sideways components of the aerodynamic acceleration, which are to be removed from a_{obs} to arrive at $a_{obs,D}$, yielding a new value for $v_{w,y}$. This chain of dependencies indicates that an iterative algorithm could be required to determine the density and wind speed with the highest possible accuracy.

Iterative Algorithm

This section presents an iterative algorithm, which avoids the restrictions and sources of error discussed in the previous section. Figure 4.8 illustrates schematically the principle of the algorithm in two steps. The goal of the algorithm is to make the modelled aerodynamic acceleration a_{mod} match the direction (top panel) and subse-
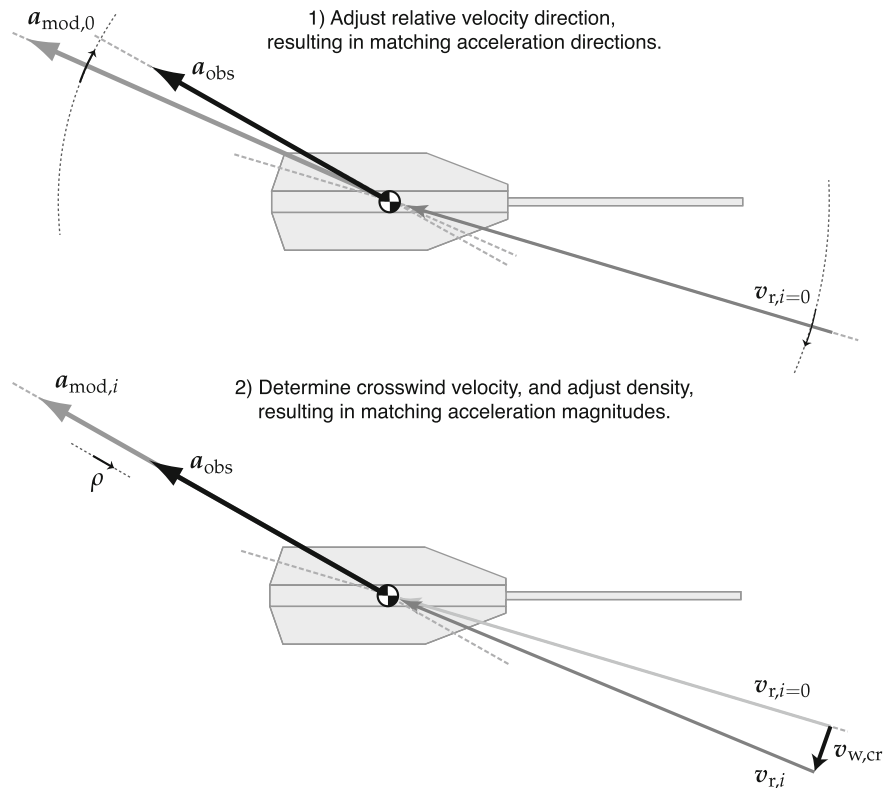
Fig. 4.8 Schematic overview of the iterative wind and density derivation algorithm for accelerometer satellites

quently the magnitude (bottom panel) of the aerodynamic acceleration observed by the accelerometer a_{obs}. This is achieved by first modifying the direction of the relative velocity vector v_{r}, without modifying its magnitude, until the modelled acceleration direction matches that of the observed acceleration. Subsequently the density ρ is modified, so that the magnitude of the accelerations matches.

The adjustment to the orientation is made by a rotation of the relative velocity about the local vertical direction, indicated by the unit vector $\hat{u}_{\mathrm{up}}$. The acceleration components projected on this direction will be set to zero. To simplify the notation, a prime is added to indicate this modification of the acceleration vectors, which is applied repeatedly:

$$a' = a - (a \cdot \hat{u}_{\mathrm{up}})\hat{u}_{\mathrm{up}} \tag{4.17}$$

We will use the sum of the orbital and co-rotation velocities as our a-priori relative velocity:

$$v_{\mathrm{r},i=0} = v_{\mathrm{o}} + v_{\mathrm{c}} \tag{4.18}$$

The possibility of including modelled in-track and vertical wind velocities in the algorithm computation will be discussed at the end of this section.

While modifying the direction of the velocity and modelled acceleration vectors, the magnitude of the acceleration is not of importance. We therefore make use of the unit vectors $\hat{a}'_{\text{obs}}$ and $\hat{a}'_{\text{mod}}$, and according to Eq. 4.11 substitute $\hat{C}_a$ for the latter. We can now define our measure of the acceleration direction residual:

$$d = \hat{a}'_{\text{obs}} - \hat{a}'_{\text{mod},i} = \hat{a}'_{\text{obs}} - \hat{C}'_{a,i}(v_{r,i}, \dots) \qquad (4.19)$$

In practice, if the magnitude of d is below a certain predefined threshold ϵ, convergence has been reached. Otherwise, another iteration is required. The convergence criterion is thus:

$$||d|| < \epsilon \qquad (4.20)$$

The unit vector representing the direction of the velocity adjustment for the current iteration is defined to be perpendicular to both the relative velocity and the rotation axis:

$$\hat{v}_{\text{adj},i} = \frac{v_{r,i} \times \hat{u}_{\text{up}}}{||v_{r,i} \times \hat{u}_{\text{up}}||} \qquad (4.21)$$

Next, to start our numerical differentiation, two relative velocity vectors are formed, which keep the magnitude of the unadjusted relative velocity, but which are rotated slightly in both directions with respect to the relative velocity of the current iteration:

$$v_r^+ = ||v_{r,i}|| \frac{v_{r,i} + \delta\hat{v}_{\text{adj},i}}{||v_{r,i} + \delta\hat{v}_{\text{adj},i}||}, \quad v_r^- = ||v_{r,i}|| \frac{v_{r,i} - \delta\hat{v}_{\text{adj},i}}{||v_{r,i} - \delta\hat{v}_{\text{adj},i}||} \qquad (4.22)$$

These modified relative velocities will result in modified modelled acceleration directions according to the equations in Chap. 3. The result from both rotation directions is substituted into Eq. 4.19:

$$d^+ = \hat{a}'_{\text{obs}} - \hat{C}'_a(v_r^+, \dots), \quad d^- = \hat{a}'_{\text{obs}} - \hat{C}'_a(v_r^-, \dots) \qquad (4.23)$$

The vector difference between the two velocity vectors is:

$$\Delta v_r = v_r^+ - v_r^-, \qquad (4.24)$$

and the effect of this velocity rotation on the acceleration direction residual is:

$$\Delta d = ||d^+|| - ||d^-|| \qquad (4.25)$$

Now, all the elements are in place to compute the next iteration of the relative velocity, which keeps the magnitude of the original velocity, but changes the direction.

$$\boldsymbol{v}_{\mathrm{r},i+1} = ||\boldsymbol{v}_{\mathrm{r},i}|| \frac{\boldsymbol{v}_{\mathrm{r},i} - d(\Delta \boldsymbol{v}_{\mathrm{r}}/\Delta d)}{||\boldsymbol{v}_{\mathrm{r},i} - d(\Delta \boldsymbol{v}_{\mathrm{r}}/\Delta d)||} \qquad (4.26)$$

At this point Eqs. 4.19 and 4.20 are reevaluated. If the convergence criterion of Eq. 4.20 is met, we can proceed computing the crosswind speed and mass density.

$$\boldsymbol{v}_{\mathrm{w,cr}} = \boldsymbol{v}_{\mathrm{r},i} - \boldsymbol{v}_{\mathrm{r},i=0} \qquad (4.27)$$

$$\rho = \frac{2m}{A_{\mathrm{ref}} v_{\mathrm{r},i}^2} \frac{||\boldsymbol{a}_{\mathrm{obs}}'||}{||\boldsymbol{C}_{\mathrm{a},i}'||} \qquad (4.28)$$

Modelling of in-Track and Vertical Winds

In the description of the algorithms above, we have not discussed the possible effect on the aerodynamics of wind components other than the cross-track component. Since we are interested in retrieving the crosswind $\boldsymbol{v}_{\mathrm{w,cr}}$ from the accelerometer data, a model value for this component should not be included in the a-priori relative velocity of Eq. 4.18. A model value for the in-track wind $\boldsymbol{v}_{\mathrm{w,it}}$, and the wind in the direction of the rotation axis $\boldsymbol{v}_{\mathrm{w,z}}$ could be applied in that equation however. These can be computed by projecting the full model wind on the unit vectors in these directions.

$$\boldsymbol{v}_{\mathrm{w,it}} = (\boldsymbol{v}_{\mathrm{w,mod}} \cdot \hat{\boldsymbol{v}}_{\mathrm{r}}) \hat{\boldsymbol{v}}_{\mathrm{r}} \qquad (4.29)$$

$$\boldsymbol{v}_{\mathrm{w,up}} = (\boldsymbol{v}_{\mathrm{w,mod}} \cdot \hat{\boldsymbol{u}}_{\mathrm{up}}) \hat{\boldsymbol{u}}_{\mathrm{up}} \qquad (4.30)$$

Since $\hat{\boldsymbol{v}}_{\mathrm{r}}$ changes its direction during the iterative process described in the previous section, Eq. 4.29 will have to be reevaluated and $\boldsymbol{v}_{\mathrm{r}}$ in Eqs. 4.22 and 4.26 will have to be updated after each iteration step.

The influence of the use of a model for the in-track wind on the density and crosswind data will be evaluated in Sect. 4.2.4

4.2.3 Example of CHAMP and GRACE density and wind processing

Figures 4.9 and 4.10 show examples of the inputs and outputs of the density and wind processing algorithms for two orbits of CHAMP and GRACE. The selected orbits are on April 5, 2005, when the two satellites were in nearly the same orbital plane, with local solar times of ascending and descending equator crossings at approximately 12:50 and 00:50, respectively. The time span for the GRACE data in Fig. 4.10 is shifted forward by 20 min with respect to the one for CHAMP in Fig. 4.9, so that the series for both satellites start near the ascending node. A geomagnetic storm

$(K_p = 7)$ occurred a few hours before the orbits that are shown here. CHAMP was at an approximate mean altitude of 360 km, while GRACE was at 470 km.

Velocity Inputs

Concentrating on the velocities first, it is clear that the X-direction for both spacecraft is dominated by the nearly constant orbital velocity. The velocities in the Y- and Z-directions are an order of magnitude smaller. In the Y-direction, the sinusoidal pattern of the velocity due to the corotating atmosphere is clearly visible, although the wind model and projection of the orbital velocity on this axis also show clear contributions. In the Z-direction, the projections of the corotation and horizontal wind are nearly zero. The orbital velocity projected on the Z-axis still results in a signal of up to about 100 m/s though. The behaviour of this component is determined by the satellite's pitch rotations. For CHAMP, the attitude thruster actuations are clearly visible as changes in slope. For GRACE, the pitch motion is controlled to keep accurate pointing between the twin satellites, resulting in a sinusoidal variation at the orbital period.

Acceleration Inputs

The top-right frames of Figs. 4.9 and 4.10 show both the observed and modelled accelerations. Note that observed accelerations in the Z-direction for CHAMP are missing from Fig. 4.9 though, because of a malfunction in the instrument.

The aerodynamic accelerations are dominant in the X-direction for both satellites, while in the Y-direction, this is true for CHAMP but not for GRACE. This is because the magnitude of the aerodynamic acceleration for GRACE is about one fifth of that of CHAMP, but the radiation pressure is about the same for both satellites.

Note that the modelled radiation pressure and total observed accelerations contain jumps at eclipse entries and exits. These jumps can be used to check the accuracy of the radiation pressure model and calibration scale factor. The observed aerodynamic acceleration, which is computed by taking the difference between these two, should not show such jumps, if eclipses are correctly modelled and the scale factors used in the accelerometer calibration are determined correctly.

Density Results

The density and crosswind results from both the direct and iterative algorithm are shown at the bottom of Figs. 4.9 and 4.10.

The density results for the direct and iterative algorithms are nearly identical. This indicates that the assumption of the direct algorithm that the accelerometer X-axis is closely aligned with the along-track axis is valid for the purpose of density determination. Despite the altitude difference of 110 km, the retrieved densities for

Fig. 4.9 Time series of relative velocity components (*left*) and acceleration components (*right*) in the body-fixed frame for CHAMP

the two satellites show many similar variations. The high-frequency variations in the data, which are not present in the NRLMSISE- 00 model, should be especially interesting for interpretation. A closer look at this data, and that of the surrounding days, will be provided in Chap. 5.

Crosswind Results

Note how for both satellites the crosswind pattern is quite different than that predicted by the HWM07 model. Especially the peak winds, encountered when crossing the
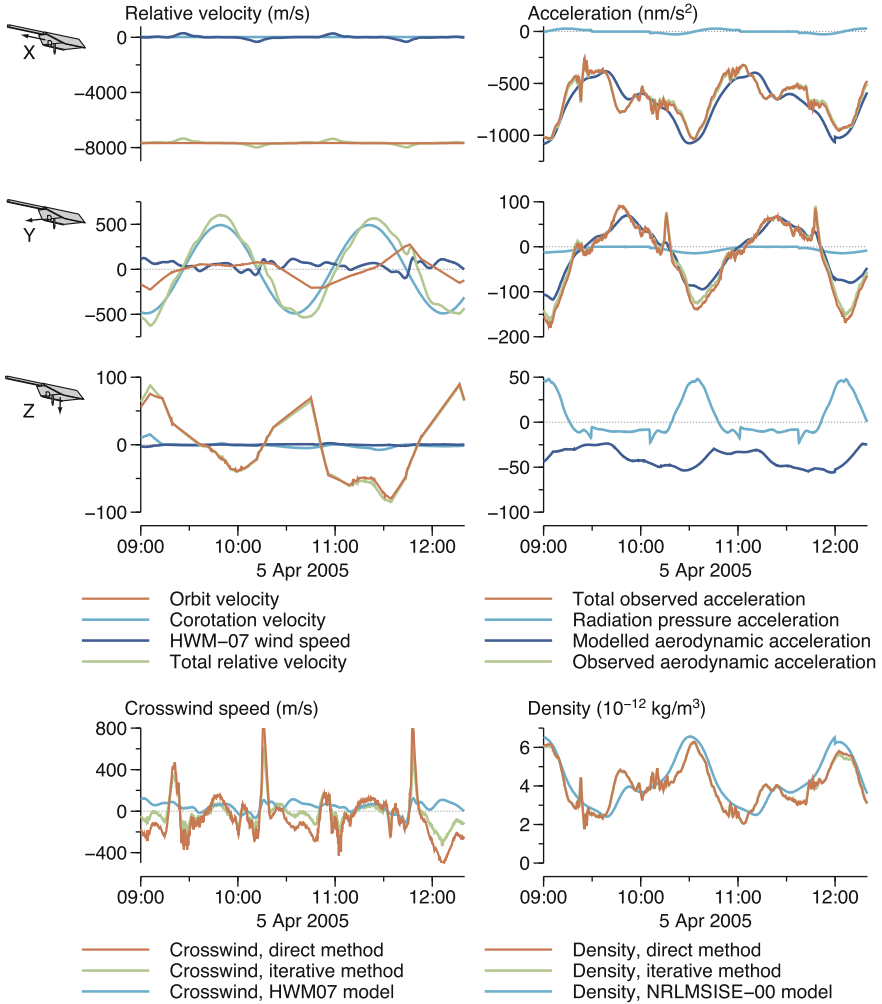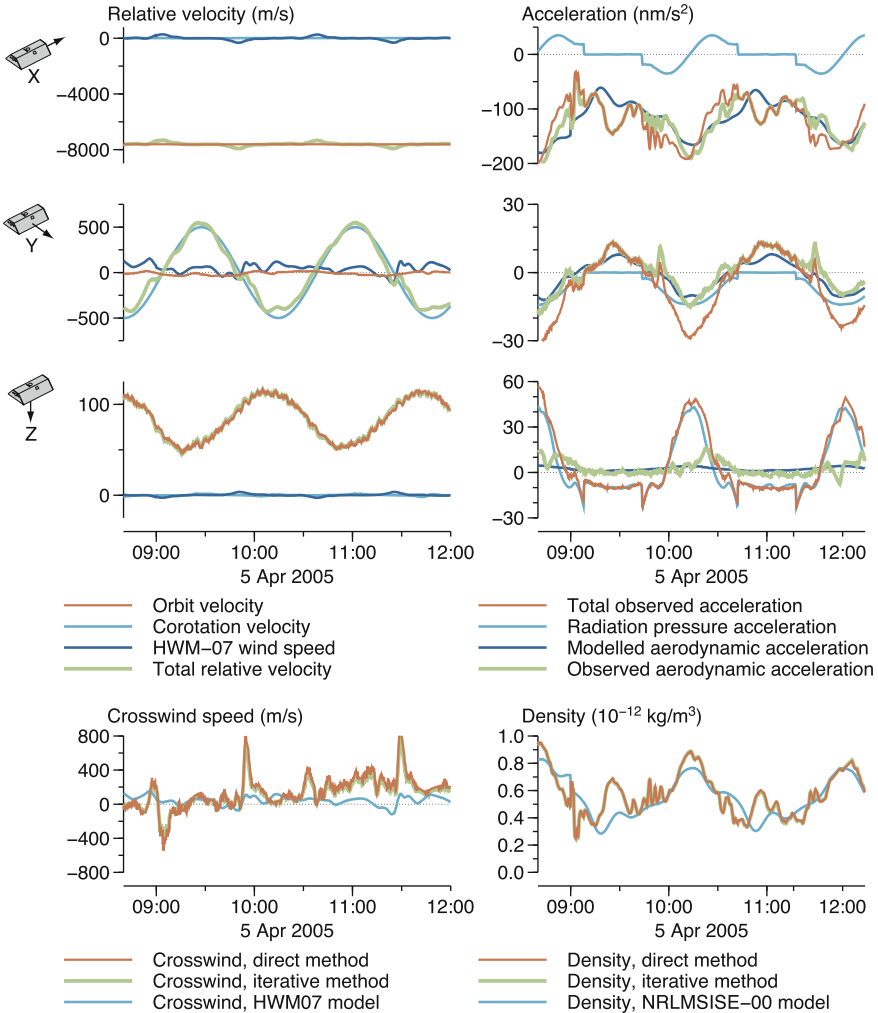
Fig. 4.10 Time series of relative velocity components (*left*) and acceleration components (*right*) in the body-fixed frame for GRACE-B

auroral zones, are much larger (up to 600 m/s). This discrepancy, which is certainly not specific to the two orbits shown here, will be further discussed in Chap. 5 as well.

The direct algorithm winds are, in general, even larger than those from the iterative algorithm. This difference is larger for CHAMP than for GRACE. This can be explained by the wider attitude deadband for CHAMP. This adversely influences the close alignment between spacecraft axes and orbit axes, which is required for the accuracy of the direct algorithm. The smaller crosswind outcome of the iterative algorithm is thought to be the more accurate.

The crosswind time series for GRACE shows a positive slope. This is most likely the result of a positive trend in the accelerometer Y-axis bias, which was not taken into account in the calibration.

Before proceeding with further interpretations of the density and wind output data in Chap. 5, the accuracy of the direct and iterative algorithm results, and their sensitivity to error sources, is evaluated using simulated data in the next section. Such an evaluation provides important insights into the reliability of the results.

4.2.4 Error Assessment Using Simulated Accelerometer Data

Simulated aerodynamic acceleration data were generated by applying Eq. 4.11, using modelled density ρ_m and wind $v_{w,m}$ values from the NRLMSISE-00 density model [69] and the HWM07 wind model [19]. Real CHAMP attitude and orbit data were used both in the generation of these simulated acceleration data and in the application of the retrieval algorithms. The accuracy of the algorithm results, comprising the density ρ and crosswind $v_{w,y}$ or $v_{w,cr}$ can then be tested by examining density and wind residuals r_ρ and r_v.

The density residuals are expressed as percentages, relative to the simulated density signal:

$$r_\rho = \frac{\rho - \rho_m}{\rho_m} \cdot 100\% \qquad (4.31)$$

The wind residuals are defined as the differences between the retrieved wind speed and the modelled wind speed's projection on the direction of the retrieved wind speed:

$$r_v = v_{w,y} - \boldsymbol{v}_{w,m} \cdot \hat{\boldsymbol{v}}_{w,y} \quad \text{or} \quad r_v = v_{w,cr} - \boldsymbol{v}_{w,m} \cdot \hat{\boldsymbol{v}}_{w,cr} \qquad (4.32)$$

Six cases for simulated errors are defined. In the first case, labeled "Identical input", the exact same models and input data were used in the retrieval algorithm that had been originally applied in the creation of the simulated accelerations. This allows for an assessment of the errors that are purely inherent in the algorithm.

However, it is also important to evaluate the algorithms under the influence of the uncertainties in the input data and models. In two further cases, named "X_{SBF} offset" and "Y_{SBF} offset", a value of 10 nm/s^2 was added to the simulated acceleration data in either direction, before applying the algorithms. For each single measurement, such an offset could be the result of an error in the instrument calibration, or due to errors in the removal of radiation pressure and attitude thruster accelerations. The magnitude of 10 nm/s^2 was found to be typical for the order of magnitude of such errors. However, each of these error sources comes with its own temporal variation of the acceleration error, and their effects can either add up or (partly) cancel each other out over time. Therefore, the constant 10 nm/s^2 offset introduced here should

be viewed as just a crude approximation, which should nevertheless give some idea of the sensitivity of the density and wind derivation algorithms. In the fourth test case, named "In-track wind", the HWM07 modelled wind in the in-track direction was neglected in the density and crosswind derivation algorithms.

Two further test cases are used to assess the effect of force model errors. The simulated aerodynamic accelerations were generated with a value of $\alpha = 0.93$ for the energy flux accommodation parameter in the aerodynamic model. In the "Energy accommodation" case, this value was changed to $\alpha = 0.88$ for the density and wind retrieval. This 5% difference can be used to represent one aspect of the inherent uncertainty in the gas-surface interaction modelling. The final test case, named "Panel model", is used to represent the uncertainty in the satellite geometry model for complexly shaped satellites such as CHAMP. Our own adjusted panellised geometrical model of the CHAMP satellite [17] used in the simulated data generation was replaced by an alternative one [5] in the density and wind retrieval. This replacement amounts to a reduction in frontal area of around 14% when the satellite is viewed along the X_{SBF} axis (front), and of 8% when viewed along the Y_{SBF} axis (side).

The statistics of the density retrieval residuals over the complete year 2004 (~ 3.15 million measurements), are presented in Table 4.1. During this year, CHAMP flew only in its nominal attitude mode for which the direct algorithm is applicable. The data in the table shows that the iterative algorithm leads to lower density residuals than the direct approach in the "Identical input" case. With the 10 nm/s^2 Y_{SBF} acceleration offset, this ranking is shifted however, because the iterative algorithm is sensitive to this acceleration component, while the direct algorithm is not. Both algorithms show an equal sensitivity to in-track wind errors. The direct algorithm seems slightly more sensitive to the force-model related errors though, if judged by the standard deviations. Note that for the force model related errors, the mean values are generally larger than the standard deviations, indicating that these accelerometer-derived density data are affected by mostly systematic errors. The data will therefore be more suitable to studies of relative changes in density than for use in modelling approaches which require absolute density values.

In the wind residual statistics, presented in Table 4.2, the advantage of the iterative algorithm over the direct approach is evident, certainly for the "Identical input" case. For both algorithms, the 10 nm/s^2 error introduced in the Y_{SBF} direction of the accelerations has a very large detrimental effect on the accuracy of the crosswind speed, leading to maximum errors of 915 m/s (iterative) and 1,283 m/s (direct). Such very large wind errors will occur in the real data processing when the aerodynamic acceleration signal in the Y_{SBF} direction is small compared to the instrument calibration, instrument noise and radiation pressure errors for that direction. These large wind errors are therefore prevalent at conditions of low density, such as at higher orbital altitudes and lower solar activity levels. For this reason, it is currently not possible to routinely acquire an accurate crosswind derivation from the GRACE satellites, which are at a higher altitude than CHAMP.

An important factor in this respect is the magnitude of cross-track radiation pressure accelerations and related acceleration errors. This magnitude is at a maximum when the satellite's orbital plane is near-perpendicular to the Sun-Earth vector (dawn-

Table 4.1 Statistics of the density retrieval residuals, in percentages of the density signal

Direct algorithm	Min.	Mean	Max.	RMS	Std.
Identical input	3.3	0.0	5.5	0.7	0.7
X_{SBF} offset	-17.4	-2.7	3.8	3.1	1.5
Y_{SBF} offset	-3.3	0.0	5.5	0.7	0.7
In-track wind	-10.8	0.2	14.3	2.2	2.1
Energy accomm.	-4.8	-1.7	3.7	1.9	0.7
Panel model	3.5	7.8	14.0	7.9	0.9
Iterative algorithm	Min.	Mean	Max.	RMS	std.
Identical input	-0.5	0.0	0.0	0.0	0.0
X_{SBF} offset	-17.4	-2.9	0.5	3.2	1.4
Y_{SBF} offset	-12.3	-0.3	4.5	1.6	1.6
In-track wind	-10.9	0.2	13.6	2.0	2.0
Energy accomm.	-2.1	-1.6	-0.7	1.6	0.2
Panel model	5.7	8.1	9.0	8.1	0.3

Table 4.2 Statistics of the wind retrieval residuals, in m/s

Direct algorithm	Min.	Mean	Max.	RMS	std.
Identical input	-225	-1	202	38	38
X_{SBF} offset	-226	-2	210	44	44
Y_{SBF} offset	-173	209	1,285	236	110
In-track wind	-228	-2	205	38	38
Energy accomm.	-258	-1	222	52	52
Panel model	-262	-2	224	54	54
Iterative algorithm	Min.	Mean	Max.	RMS	Std.
Identical input	-9	-0	10	1	1
X_{SBF} offset	-44	3	81	12	12
Y_{SBF} offset	-242	81	915	140	114
In-track wind	-9	-0	10	1	1
Energy accomm.	-55	-0	51	21	21
Panel model	-52	0	49	22	22

dusk orbit). We have also encountered particularly extreme wind errors around eclipse transitions. A small discrepancy between the modelled and true eclipse geometry will lead to short periods with an incorrect application or removal of the full modelled radiation pressure acceleration, which has a maximum magnitude of around 40 nm/s^2 for CHAMP, leading to wind errors far exceeding 1,000 m/s.

The results in Table 4.2 show that the crosswind derivation is practically insensitive to errors in the in-track wind, since the results are nearly the same as for the "Identical input" case. The other error sources also have only minor effects in comparison to the Y_{SBF} offset.

It should be clear that a large aerodynamic signal strength, a careful calibration of the accelerometer in the crosswind direction and an accurate modelling of cross-

track radiation pressure accelerations are a necessity for the derivation of accurate crosswind results.

The Y_{SBF} (cross-track) calibration for the CHAMP and GRACE satellites using GPS data is problematic [35], because of the relatively low acceleration signal, and because of the limited capability of accelerations in this direction to perturb the orbit, compared to along-track accelerations. Both limitations are the result of fundamental orbital dynamics in combination with the tight attitude control of the spacecraft, as discussed in Sect. 3.3 and the start of Sect. 4.2.2, respectively. We therefore adopted an alternative approach to calibrate CHAMP's Y_{SBF} accelerometer data, analogous to Sutton et al. [76]. We took density observations derived using the direct method from the X_{SBF} data, and combined these with the aerodynamic satellite model to arrive at simulated observations for the Y_{SBF} data. The accelerometer data for this direction was then calibrated by estimating the biases that minimize the difference between the data and these simulated observations. We have checked the reliability of these biases by comparing the local time variation of the zonal wind from ascending arcs with those from descending arcs and found no systematic difference. An error in the Y_{SBF} biases would influence the ascending and descending wind profiles in opposite directions.

4.3 Force Model Parameter Estimation From Precise Tracking Data

A third important technique for analysing orbital dynamics in terms of aerodynamics is *Precise orbit determination*, named *special perturbations orbit determination* by the space surveillance community. This approach was already briefly mentioned in Sect. 4.1.3. The technique involves numerically integrating the system of second order differential Eq. 3.5 from an initial position and velocity. Precise force models are applied to integrate the motion of the satellite, in order to obtain estimates of the positions and velocities over a certain time span, the *orbit arc*.

The length of this time span is usually between a few hours and a few days, depending on the temporal density of the tracking data (denser data enable shorter arcs), the quality of the models (better models allow longer arcs), as well as operational constraints.

Due to errors and uncertainties in the initial conditions, force models and measurement models, the initially integrated orbit likely does not fit well with tracking observations. The orbit estimate can then be iteratively improved by least-squares fitting the initial position, velocity and model parameters that result in minimal tracking data residuals. The procedure is explained in great detail by, e.g., Montenbruck and Gill [63] and Tapley et al. [77].

Information on the effect of atmospheric density on the orbit can be recovered in the same process. This involves estimation of carefully selected force model parameters, which will be designated with the letter *p* in the sections below.

Three parameterisation methods will be briefly described below, even though their results will not be directly analysed in the remainder of this work. The other two methods presented earlier in this chapter are easier to implement and result in a better spatial sampling of the data. Nevertheless, the force model parameter estimation method is important for accelerometer data calibration. It could also be an essential part of future thermosphere density studies, especially at altitudes above about 500 km, where the accuracy of TLE-derived data is insufficient and accelerometers are unavailable.

4.3.1 Estimating Aerodynamic Scale Factor Parameters

In the first parameterisation method, the computed aerodynamic acceleration is multiplied with an aerodynamic scale factor parameter p_a.

$$\ddot{\boldsymbol{r}}_a = p_a C_a \frac{A_{\text{ref}}}{m} \frac{1}{2} \rho_M v_r^2 \tag{4.33}$$

Assuming that errors in the measurements, measurement modelling and force models are negligible, the effect of the true density on the drag force will be close to that of the scale factor times the modelled density over the time span. The mean true density is then approximately equal to

$$\bar{\rho} \approx p_a \bar{\rho}_M \tag{4.34}$$

Note that the parameter p_a represents the data over model density ratio for a certain time interval, similar to those discussed in Sect. 2.6.1. The use of these ratios facilitates direct comparisons of the data between different methods and different objects.

Time-Dependent Estimation, Parameter Observability and Correlations

The estimated parameter, such as p_a in the equation above, does not necessarily have to be valid for the entire arc. When longer arcs are used, the force model can be adjusted using a time series of piecewise constant or piecewise linear parameters, each of which is valid for a certain *sub-arc*. The orbit analyst can therefore select the temporal resolution of the model adjustment.

This leads to the related topics of parameter observability and correlations between estimated parameters. In principle, the number of force model parameters that can be estimated is limited by the coverage of tracking observations. For modern techniques, such as GPS and DORIS this number will be very large. As more and more force model parameters are estimated this will almost certainly result in an improved least-squares fit of the orbit with the tracking observations. However, this does not
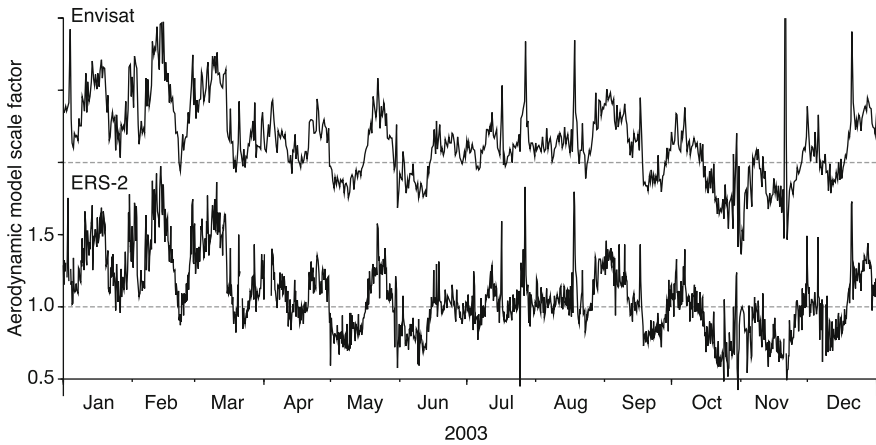
Fig. 4.11 Aerodynamic model scale factors, estimated from tracking data of ERS-2 and Envisat (offset by 1.0)

automatically mean that the estimated parameters are suitable for further processing into data on the thermospheric density and wind.

Over shorter time intervals, there is an increased chance that the parameters are affected by errors in the measurements or other force models, besides the aerodynamics. In addition, estimated parameters could be correlated, which means that adjustment of different parameters has a similar effect on the orbit fit. Solutions to these problems are either to eliminate certain estimation parameters, or constrain their solution to a-priori values. Neither solution is ideal, and interpretations of the results of parameter estimations should always be made while being aware of these possible shortcomings.

Examples of these limitations will be provided when other parameterisation schemes are discussed in the next sections. But first an example will be provided of the results of aerodynamic force model scale factor estimation.

ERS-2 and Envisat Example

Figure 4.11 shows results that were first presented in Doornbos et al. [15]. The time series of p_a shown in the Figure have been estimated for adjacent 6 h intervals, making use of SLR and DORIS tracking data for Envisat and of SLR tracking for ERS- 2. Accurate ANGARA models (see Sect. 3.5) were used to model the geometry and gas-surface interaction for both satellites [13]. NRLMSISE- 00 was used to model the density. Since radiation pressure and gravity forces can be fairly accurately modelled for both satellites, the plotted values mainly represent NRLMSISE- 00 density modelling errors, corresponding to Eq. 4.34.

The time series for ERS- 2 and Envisat are nearly identical, due to their similar orbits, which are near-circular with frozen eccentricity, sun-synchronous, near-polar and at an altitude of approximately 800 km. The only difference between the orbits is the right-ascension of the ascending node. ERS- 2 crosses the equator at 10:30/22:30 local solar time, while Envisat precedes it by 30 min at 10:00/22:00.

The Envisat time series contains less noise, which is due to the availability of the dense DORIS tracking data, compared to the more sparse SLR data which has been the only source of data that was used for ERS- 2. There are a few peaks distinct to each satellite, which are due to manoeuvres. Orbit arcs containing manoeuvres should therefore be eliminated before further analysis in terms of thermosphere density takes place. There are also some peaks, for instance in July and August, which are common to both ERS- 2 and Envisat. These probably indicate short-term disturbances in the atmosphere, which are not properly represented in the NRLMSISE- 00 density model.

4.3.2 Estimating Empirical Acceleration Parameters

A different estimation approach starts by eliminating the detailed non-gravitational force models altogether in the orbit determination. Instead, empirical accelerations $\ddot{r}_E$ are estimated directly. At a later time, these accelerations can then be processed much like calibrated accelerometer observations, as described in Sect. 4.2.

A basic parameterisation scheme for this method is to introduce piecewise constant empirical acceleration parameters p in the three orthogonal orbit-fixed directions, where the subscripts A, C and N signify the along-track, cross-track and in-plane normal directions, respectively (see Fig. 3.4).

$$\ddot{r}_a \approx p_A \hat{u}_A + p_C \hat{u}_C + p_N \hat{u}_N \tag{4.35}$$

Note however, that both in-plane acceleration components affect the shape of the orbit (see Sect. 3.3). These accelerations can therefore not easily be distinguished using tracking observations. The parameters p_A and p_N are therefore correlated. Usually, only the along-track component is estimated, since both its magnitude and effect on the orbit are generally much larger than the radial component. The radial acceleration in Eq. 4.35 is then set to zero.

The same caveat applies to a more elaborate 9-parameter estimation, which includes harmonic variations of the empirical acceleration as a function of the argument of latitude ν.

$$\begin{aligned} \ddot{r}_a \approx \ &(p_A + p_{A,c}\cos(\nu) + p_{A,s}\sin(\nu))\hat{u}_A \\ &+ (p_C + p_{C,c}\cos(\nu) + p_{C,s}\sin(\nu))\hat{u}_C \\ &+ (p_N + p_{N,c}\cos(\nu) + p_{N,s}\sin(\nu))\hat{u}_N \end{aligned} \tag{4.36}$$

These harmonic accelerations are resonant at the orbital period and are therefore effective at changing the orbital elements (see Sect. 3.3). This has proved to be a
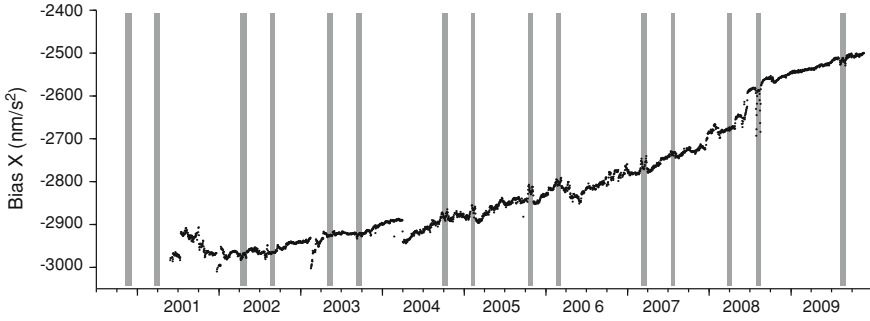
Fig. 4.12 Biases in the X-component of the CHAMP accelerometer data, estimated using GPS tracking data. The vertical grey bars indicate time periods without eclipses

highly useful strategy for fitting the integrated orbit to the tracking data. The true aerodynamic acceleration will exhibit such 1 cycle-per-revolution (CPR) accelerations as well, for instance due to the sampling of the diurnal density variation over a circular noon-midnight orbit. In many cases however, higher frequency variations are important in the aerodynamic accelerations as well. It is therefore more accurate, if the density of the tracking data allows it, to estimate piecewise constant or piecewise linear accelerations over intervals much shorter than the orbital period, in order to resolve the variation of the acceleration within the orbit. This strategy is applied by van den IJssel and Visser [37–39].

4.3.3 Estimating Accelerometer Calibration Parameters

The final force model parameterisation scheme discussed here is slightly different. It involves the use of accelerometer observations in the orbit determination process. As already mentioned in Sect. 4.2.1, the raw accelerometer observations often exhibit acceleration biases and scale errors. These can be solved by including them as force model parameters in the orbit determination, and disabling all non-gravitational force models in the software.

If the parameters p_s and p_b represent the scale factors and biases, respectively, and the indices x, y and z denote the three orthogonal axes of the accelerometer instrument, then the non-gravitational forces in the orbit determination software can be parameterised and implemented as follows:

$$\ddot{\boldsymbol{r}}_a \approx (p_{s,x}\ddot{r}_{raw,x} + p_{b,x})\hat{\boldsymbol{u}}_x + (p_{s,y}\ddot{r}_{raw,y} + p_{b,y})\hat{\boldsymbol{u}}_y + (p_{s,z}\ddot{r}_{raw,z} + p_{b,z})\hat{\boldsymbol{u}}_z \quad (4.37)$$

Note that the same caveat about the correlation between the along-track and radial acceleration components discussed above applies here. In addition, there also exists a correlation between the bias and the scale factor. According to the perturbation equations (Sect. 3.3), the effect of a constant acceleration component is much stronger

than that of variations about the mean acceleration. As a result, there is a tendency to arrive at values for the pair of scale factor and bias which results in the correct mean acceleration over the arc, but with errors in the individual parameters.

For CHAMP and GRACE, the scale factor is therefore considered a constant for each accelerometer axis, or its value is determined by combining observations over much longer time intervals. Details on such strategies are provided by Van Helleputte et al. [35].

Figure 4.12 shows the time series of calibration biases $p_{b,x}$, estimated from GPS data for CHAMP, as described by Van Helleputte et al. [35]. The scale factor was held fixed at a value of $p_{s,x} = 0.8332$. The time series shows a general trend, of the order of 50 nm/s^2/year, as well as large jumps, for example in mid 2001, early 2003 and early 2004. These jumps can be attributed to software updates and hardware switches onboard the satellite. Other excursions with similar signatures but different amplitudes occur at several of the periods where the satellite was in a dawn/dusk orbit and did not encounter eclipses, indicated by the grey bars in the plot. These jumps might be interpreted as signs that the accelerometer instrument is sensitive to variations in its thermal environment.

References

1. Bowman BR (2002) True satellite ballistic coefficient determination for HASDM. In: AIAA/AAS astrodynamics specialist conference and exhibit, 5–8 August 2002, Monterey, California, AIAA 2002–4887
2. Bowman BR, Moe K (2005) Drag coefficient variability at 175–500 km from the orbit decay analyses of spheres. In: AAS/AIAA astrodynamics specialist conference, 7–11 August 2005, Lake Tahoe, CA, AAS 05–257
3. Brouwer D (1959) Solution of the problem of artificial satellite theory without drag. Astron J 64(1274):378–397
4. Bruinsma S, Tamagnan D, Biancale R (2004) Atmospheric densities derived from CHAMP/STAR accelerometer observations. Planet Space Sci 52(4):297–312. doi:10.1016/j.pss.2003.11.004
5. Bruinsma S, Biancale R (2003) Total densities derived from accelerometer data. J Spacecraft Rockets 40(2):230–236
6. Bruinsma SL, Forbes JM (2007) Global observation of travelling atmospheric disturbances (TADs) in the thermosphere. Geophys Res Lett 34(L14103). doi:10.1029/2007GL030243
7. Bruinsma SL, Forbes JM (2008) Medium- to large-scale density variability as observed by CHAMP. Space Weather 6:S08002. doi:10.1029/2008SW000411
8. Bruinsma SL, Forbes JM (2009) Properties of traveling atmospheric disturbances (TADs) inferred from CHAMP accelerometer observations. Advan Space Res 43(3):369–376. doi:10.1016/j.asr.2008.10.031
9. Bruinsma SL, Forbes JM (2010) Anomalous behavior of the thermosphere during solar minimum observed by CHAMP and GRACE. J Geophys Res 115(A11323). doi:10.1029/2010JA015605
10. Burke WJ, Lin CS, Hagan MP, Huang CY, Weimer DR, Wise JO, Gentile LC, Marcos FA (2009) Storm time global thermosphere: a driven-dissipative thermodynamic system. J Geophys Res 114(A06306) doi:10.1029/2008JA013848

11. Case K, Kruizinga GLH, Wu SC (2004) GRACE level 1B data product user handbook. JPL D-22027

12. Cefola PJ, Nazarenko AI, Proulx RJ, Yurasov VS (2003) Atmospheric density correction using two line element sets as the observation data. In: AAS/AIAA astrodynamics specialists conference, 3–7 August 2003, Big Sky, Montana, AAS 03–626

13. Doornbos E, Scharroo R, Klinkrad H, Zandbergen R, Fritsche B (2002) Improved modelling of surface forces in the orbit determination of ERS and Envisat. Can J Remote Sens 28(4):535–543

14. Doornbos E (2006) NRTDM final report—Near Real-Time Density Model (NRTDM)—ESOC contract 18576/04/D/HK(SC). Delft Institute for Earth-Oriented Space Research

15. Doornbos E, Klinkrad H, Scharroo R, Visser P (2007) Thermosphere density model calibration in the orbit determination and prediction of ERS-2 and Envisat. In: Lacoste H (ed) Envisat symposium 23–27 April 2007. Montreux, Switzerland, ESA SP-636

16. Doornbos E, Klinkrad H, Visser P (2008) Use of two-line element data for thermosphere neutral density model calibration. Advan Space Res 41(7):1115–1122. doi:10.1016/j.asr.2006.12.025

17. Doornbos E, Förster M, Fritsche B, Helleputte T van, IJssel J van den, Koppenwallner G, Lühr H, Rees D, Visser P (2009) ESTEC contract 21022/07/NL/HE Air density models derived from multi-satellite drag observations—final report. DEOS / TU Delft scientific report 01/2009, TU Delft

18. Doornbos E, Van den IJsse J, Lühr H, Förster M, Koppenwallner G (2010) Neutral density and crosswind determination from arbitrarily oriented multiaxis accelerometers on satellites. J Spacecraft Rockets 47(4):580–589. doi:10.2514/1.48114

19. Drob DP, Emmert JT, Crowley G, Picone JM, Shepherd GG, Skinner W, Hays P, Niciejewski RJ, Larsen M, She CY, Meriwether JW, Hernandez G, Jarvis MJ, Sipler DP, Tepley CA, O'Brien MS, Bowman JR, Wu Q, Murayama Y, Kawamura S, Reid IM, Vincent RA (2008) An empirical model of the earth's horizontal wind fields: HWM07. J Geophys Res 113(A12304). doi:10.1029/2008JA013668

20. Emmert JT (2009) A long-term data set of globally averaged thermospheric total mass density. J Geophys Res 114(A06315). doi:10.1029/2009JA014102

21. Emmert JT, Picone JM (2010) Climatology of globally averaged thermospheric mass density. J Geophys Res 115(A09326). doi:10.1029/2010JA015298

22. Emmert JT, Picone JM, Lean JL, Knowles SH (2004) Global change in the thermosphere: Compelling evidence of a secular decrease in density. J Geophy Res 109(A2):A02301. doi:10.1029/2003JA010176

23. Emmert JT, Meier RR, Picone JM, Lean JL, Christensen AB (2006) Thermospheric density 2002–2004: TIMED/GUVI dayside limb observations and satellite drag. J Geophys Res 111(A10S16). doi:10.1029/2005JA011495

24. Emmert JT, Picone JM, Meier RR (2008) Thermospheric global average density trends, 1967–2007, derived from orbits of 5,000 near-earth objects. Geophys Res Lett 35(L05101). doi:10.1029/2007GL032809

25. Flury J, Bettadpur S, Tapley BD (2008) Precise accelerometry onboard the GRACE gravity field satellite mission. Advan Space Res 42(8):1414–1423. doi:10.1016/j.asr.2008.05.004

26. Forbes JM, Lu G, Bruinsma S, Nerem S, Zhang X (2005) Thermosphere density variations due to the 15–24 April 2002 solar events from CHAMP/STAR accelerometer measurements. J Geophys Res 110(A12S27). doi:10.1029/2004JA010856

27. Forbes JM, Bruinsma SL, Miyoshi Y, Fujiwara H (2008) A solar terminator wave in thermosphere neutral densities measured by the CHAMP satellite. Geophys Res Lett 35(L14802). doi:10.1029/2008GL034075

28. Forbes JM, Bruinsma SL, Zhang X, Oberheide J (2009) Surface-exosphere coupling due to thermal tides. Geophys Res Lett 36(L15812). doi:10.1029/2009GL038748

29. Förste Ch (2002) Format description: The CHAMP data format. CH-GFZ-FD-001

30. Förster M, Rentz S, Köhler W, Liu H, Haaland SE (2008) IMF dependence of high-latitude thermospheric wind pattern derived from CHAMP cross-track measurements. Ann Geophys 26(6):1581–1595

31. Förster M, Doornbos E, Van Hellepunte T , Haaland SE, Rentz S, Lühr H (2009) Magnetic forcing of the high-latitude upper atmosphere. In: Proceedings of ESA's second swarm international science meeting, 24–26 June 2009, Potsdam, Germany

32. Guo J, Wan W, Forbes JM, Sutton E, Nerem RS, Woods TN, Bruinsma S, Liu L (2007) Effects of solar variability on thermosphere density from CHAMP accelerometer data. J Geophys Res 112(A10308). doi:10.1029/2007JA012409

33. Häusler K, Lühr H (2009) Nonmigrating tidal signals in the upper thermospheric zonal wind at equatorial latitudes as observed by CHAMP. Ann Geophys 27(6):2643–2652

34. Häusler K, Lühr H, Rentz S, Köhler W (2007) A statistical analysis of longitudinal dependencies of upper thermospheric zonal winds at dip equator latitudes derived from CHAMP. J Atmos Solar-Terr Phys 69(12):1419–1430. doi:10.1016/j.jastp.2007.04.004

35. Helleputte T. van, Doornbos E, Visser P (2009) CHAMP and GRACE accelerometer calibration by GPS-based orbit determination. Advan Space Res 43(12):1890–1896. doi:10.1016/j.asr.2009.02.017

36. Hoots FR, Roehrich RL (1980) Spacetrack report no. 3: models for propagation of NORAD element sets. Aerospace Defense Center Peterson Air Force Base. Obtained from http://www.celestrak.com/

37. van den IJssel J, Visser P (2005) Determination of non-gravitational accelerations from GPS satellite-to-satellite tracking of CHAMP. Advan Space Res 36(3):418–423. doi:10.1016/j.asr.2005.01.107

38. van den IJssel J, Visser P (2007) Performance of GPS-based accelerometry: CHAMP and GRACE. Advan Space Res 39(10):1597–1603. doi:10.1016/j.asr.2006.12.027

39. van den IJssel J, Visser P (2010) Performance of gps-based accelerometry: a simulation experimen. Advan Space Res 45(2):225–238. doi:10.1016/j.asr.2009.09.012

40. Keating GM, Tolson RH, Bradford MS (2000) Evidence of long term global decline in the earth's thermospheric densities apparently related to anthropogenic effects. Geophys Res Lett 27(10):1523–1526

41. King-Hele D (1987) Satellite orbits in an atmosphere, theory and applications. Blackie, Glasgow

42. König R, Michalak G, Neumayer KH, Schmidt R, Zhu SY, Meixner H, Reigber C (2005) Recent developments in CHAMP orbit determination at GFZ. In: Earth observation with CHAMP, results from three years in orbit. pp 65–70

43. Lane MH, Cranford KH (1969) An improved analytical drag theory for the artificial satellite problem. In: Astrodynamics conference, 20–22 August 1969 AIAA. Princeton, NJ, number AIAA 69–925

44. Lathuillère C, Menvielle M (2010) Comparison of the observed and modeled low- to mid-latitude thermosphere response to magnetic activity: effects of solar cycle and disturbance time delay. Advan Space Res 45(9):1093–1100. doi:10.1016/j.asr.2009.08.016

45. Lathuillère C, Menvielle M, Marchaudon A, Bruinsma S (2008) A statistical study of the observed and modeled global thermosphere response to magnetic activity at middle and low latitudes. J Geophys Res 113(A07311). doi:10.1029/2007JA012991

46. Lean JL, Picone JM, Emmert JT, Moore G (2006) Thermospheric densities derived from spacecraft orbits: application to the Starshine satellites. J Geophys Res 111(A04301). doi:10.1029/2005JA011399

47. Lei J, Thayer JP, Forbes JM, Sutton EK, Nerem RS, Temmer M, Veronig AM (2008) Global thermospheric density variations caused by high-speed solar wind streams during the declining phase of solar cycle 23. J Geophys Res 113(A11303). doi:10.1029/2008JA013433

48. Liu H, Lühr H (2005) Strong disturbances of the upper thermospheric density due to magnetic storms: CHAMP observations. J Geophys Res 110(A09829). doi:10.1029/2004JA010908

49. Liu H, Lühr H, Henize V, Köhler W (2005) Global distribution of the thermospheric total mass density derived from CHAMP. J Geophys Res 110(A04301) doi:10.1029/2004JA010741

50. Liu H, Lühr H, Watanabe S, Köhler W, Henize V, Visser P (2006) Zonal winds in the equatorial upper thermosphere: decomposing the solar flux, geomagnetic activity, and seasonal dependencies. J Geophys Res 111(A07307). doi:10.1029/2005JA011415
51. Liu H, Lühr H, Watanabe S (2007a) Climatology of the equatorial thermospheric mass density anomaly. J Geophys Res 112(A05305). doi:10.1029/2006JA012199
52. Liu H, Lühr H, Watanabe S, Köhler W, Manoj C (2007b) Contrasting behavior of the thermosphere and ionosphere in response to the 28 Oct 2003 solar flare. J Geophys Res 112(A077305), doi:/10.1029/2007JA012313
53. Liu H, Lühr H, Watanabe S (2009) A solar terminator wave in thermospheric wind and density simultaneously observed by CHAMP. Geophys Res Lett 36(L10109). doi:10.1029/2009GL038165
54. Liu R, Lühr H, Doornbos E, Ma S-Y (2010) Thermospheric mass density variations during geomagnetic storms and a prediction model based on the merging electric field. Ann Geophys 28:1633–1645. doi:10.5194/angeo-28-1633-2010
55. Lühr H, Rother M, Köhler W, Ritter P, Grunwaldt L (2004) Thermospheric upwelling in the cusp region: evidence from CHAMP observations. Geophys Res Lett 31(6) doi:10.1029/2003GL019314
56. Lühr H., Rentz S., Ritter P., Liu H., Häusler K (2007a) Average thermospheric wind pattern over the polar regions, as observed by CHAMP. Ann Geophys 25(5):1093–1101
57. Lühr H, Häusler K, Stolle C (2007b) Longitudinal variation of F region electron density and thermospheric zonal wind caused by atmospheric tides. Geophys Res Lett 34(L16102). doi:10.1029/2007GL030639
58. Ma R, Xu J, Wang W, Lei J, Liu H-L, Maute A, Hagan ME (2010) Variations of the nighttime thermospheric mass density at low and middle latitudes. J Geophys Res 115(A12301). doi:10.1029/2010JA015784
59. Marcos FA, Forbes JM (1985) Thermospheric winds from the satellite electrostatic triaxial accelerometer system. J Geophys Res 90:6543–6552
60. McCarthy DD, Petit G (2003) IERS conventions (2003). IERS technical note, no. 32. International Earth Rotation and Reference Systems Service (IERS)
61. Menvielle M, Lathuillère C, Bruinsma S, Viereck R (2007) A new method for studying the thermospheric density variability derived from CHAMP/STAR accelerometer data for magnetically active conditions. Ann Geophys 25:1949–1958
62. Miyoshi Y, Fujiwara H, Forbes JM, Bruinsma SL (2009) Solar terminator wave and its relation to the atmospheric tide. J Geophys Res 114(A07303). doi:10.1029/2009JA014110
63. Montenbruck O, Gill E (2000) Satellite orbits models, methods and applications. Springer, Berlin
64. Müller S, Lühr H, Rentz S (2009) Solar and magnetospheric forcing of the low latitude thermospheric mass density as observed by CHAMP. Ann Geophys 27(5):2087–2099
65. Oberheide J, Forbes JM, Häusler K, Wu Q, Bruinsma SL (2009) Tropospheric tides from 80 to 400 km: propogation, interannual variability, and solar cycle effects. J Geophys Res 114(D00I05). doi:10.1029/2009JD012388
66. Pardini C, Anselmo L (1994) SATRAP: satellite reentry analysis program. Internal Report C94-17, Istituto CNUCE, CNR, Pisa, 30 Agosto 1994
67. Pardini C, Anselmo L, Moe K, Moe MM (2010) Drag and energy accommodation coefficients during sunspot maximum. Advan Space Res 45(5):638–650. doi:10.1016/j.asr.2009.08.034
68. Perosanz R (2003) On board evaluation of the STAR accelerometer. In: Christoph R et al.(eds) First CHAMP mission results for gravity magnetic and atmospheric studies, Springer, Berlin, pp 11–18
69. Picone JM, Hedin AE, Drob DP, Aikin AC (2002) NRLMSISE-00 empirical model of the atmosphere: statistical comparisons and scientific issues. J Geophys Res 107(A12). doi:10.1029/2002JA009430

70. Picone JM, Emmert JT, Lean J (2005) Thermospheric densities derived from space-craft orbits-I. Accurate processing of two-line element sets. J Geophys Res 110(A03301). doi:10.1029/2004JA010585
71. Pilinski MD, Argrow BM, Palo SE (2011) Drag coefficients of satellites with concave geome-tries: comparing models and observations. J Spacecraft Rockets 48(2):312–324
72. Rentz S, Lühr H (2008) Climatology of the cusp-related thermospheric mass density anomaly, as derived from CHAMP observations. Ann Geophys 26(9):2807–2823
73. Ritter P, Lühr H, Doornbos E (2010) Substorm-related thermospheric density and wind disturbances derived from CHAMP observations. Ann Geophys 28:1207–1220. doi:10.5194/angeo-28-1207-2010
74. Sutton EK, Forbes JM, Nerem RS (2005) Global thermospheric neutral density and wind response to the severe 2003 geomagnetic storms from CHAMP accelerometer data. J Geophys Res 110(A09S40). doi:10.1029/2004JA010985
75. Sutton EK, Forbes JM, Knipp DJ (2009) Rapid response of the thermosphere to variations in joule heating. J Geophys Res 114(A04319). doi:10.1029/2008JA013667
76. Sutton EK, Nerem RS, Forbes JM (2007) Density and winds in the thermosphere deduced from accelerometer data. J Spacecraft Rockets 44(6):1210–1219. doi:10.2514/1.28641
77. Tapley BD, Schutz BE, Born GH (2004) Statistical orbit determination. Elsevier Academic, New York
78. Tapley BD, Ries JC, Bettadpur S, Cheng M (2007) Neutral density measurements from the gravity recovery and climate experiment accelerometers. J Spacecraft Rockets 6(44):1220–1225. doi:10.2514/1.28843
79. Touboul P, Foulon B, Rodrigues M, Marque JP (2004) In orbit nano-g measurements, lessons for future space missions. Aerosp Sci Technol (8):431–441
80. Vallado DA (2001) Fundamentals of astrodynamics and applications, 2nd edn. Microcosm Press, El Segundo
81. Vallado DA, Crawford P (2008) SGP4 orbit determination. In: 2008 AIAA/AAS astrodynamics specialist conference
82. Vallado DA, Crawford P, Hujsak R, Kelso TS (2006) Revisiting spacetrack report ♯3. In: AIAA astrodynamics specialists conference and exhibit, AIAA 2006–6753

Chapter 5
Analysis and Interpretation of Density and Wind Data

The previous chapter already provided several figures and tables, illustrating results of the density and wind processing algorithms. These output data will now be analysed more closely. The various data sets will be plotted in various ways, compared with each other and with output from models, and interpreted in terms of geophysical signals and error sources.

The chapter starts with separate presentations of TLE- and accelerometer data. Section 5.1 shows time series of TLE-derived density data for several satellites. The chapter continues with an in-depth look at accelerometer-derived data in Sect. 5.2.1, by presenting an interesting case study, which puts the iterative accelerometer data processing algorithm to the test for the CHAMP satellite, taking advantage of a rare attitude manoeuvre. Section 5.2.2 gives a taste of geophysical features that can be observed in the accelerometer-derived density data, from both CHAMP and GRACE, over a slightly longer period, surrounding a geomagnetic storm. The chapter concludes with a statistical evaluation of the various density data sources and models, in Sect. 5.3.

5.1 Time Series of TLE-Derived Density Results

A single example of the density time series output of the TLE-processing algorithm by Picone et al. [23] was provided in the previous chapter as Fig. 4.3, for the space-debris object 22659 which was also used for illustrations in the introduction chapter. Similar results, for several more satellites, will be presented in this section, and discussed in somewhat more detail.

5.1.1 CHAMP and GRACE

Figure 5.1 shows data for the CHAMP and GRACEA satellites. These form an interesting case-study, since their TLE-derived data can be compared with the
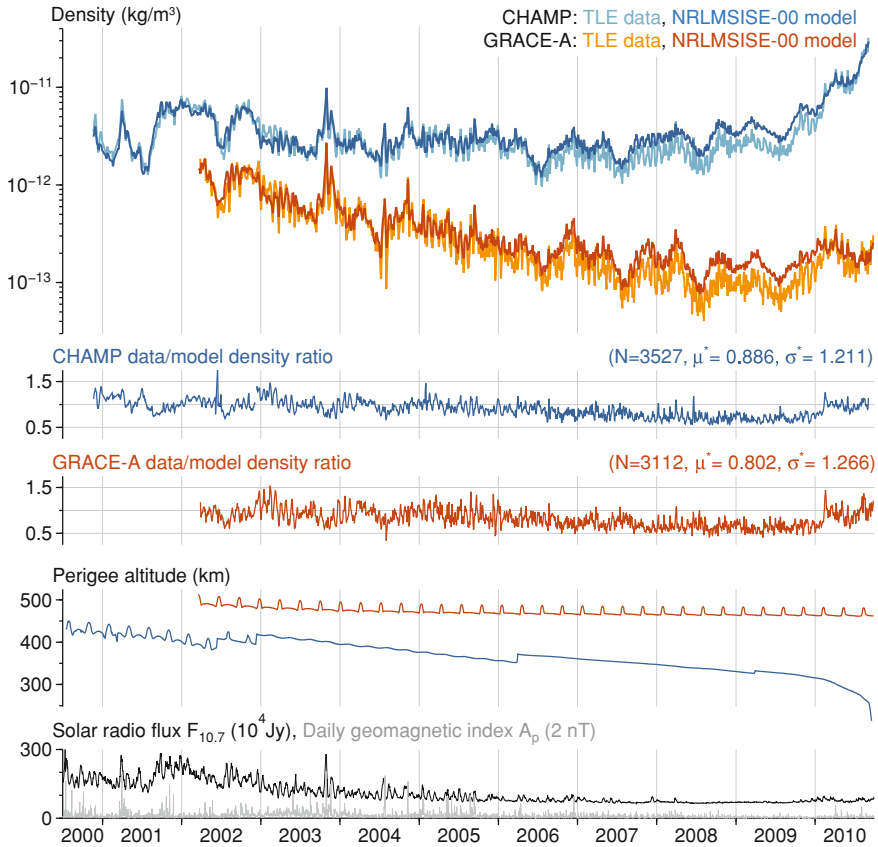
Fig. 5.1 Results of the TLE to density processing for the CHAMP and GRACE-A satellites

accelerometer-derived data. This comparison will be presented in Sect. 5.3, while this section will focus on features in the time-series data.

The available information on the geometry, controlled attitude, occasional orbit manoeuvres and related mass changes of CHAMP and GRACE can be used in a more complicated processing than that of most other TLE-objects, for which such information is not available. The TLE-processing for these satellites has therefore been performed twice: once using the exact same highly detailed input data and force models used in the accelerometer processing, and once using fixed ballistic coefficients, like for other TLE-objects. Figure 5.1 shows the results based on the fixed ballistic coefficient calculations.

The top frame of the Figure shows both the TLE-derived densities, as well as the equivalent model densities in a darker shade. The ratio of data over model density is plotted in the panel below it. The bottom two panels show the most important drivers

of density variations along the trajectories: the approximate altitudes of the satellites, and proxies for the level of solar- and geomagnetic activity.

Orbit evolution

CHAMP, which was launched earlier and at a slightly lower altitude (July 15, 2000, 454 km) than GRACE (March 17, 2002, 500 km), has re-entered on September 19, 2010, after slightly more than 10 years in orbit. Four orbit-raising manoeuvres, prolonging its lifetime, can be identified in the Figure: two in 2002, and one each in 2006 and 2009. Because of this evolution of the orbit altitude, which coincided with the declining phase of the solar cycle, the averaged density encountered by CHAMP has remained at the same order of magnitude, between 10^{-12} and 10^{-11} kg/m^3 throughout most of its lifetime. GRACE meanwhile, undergoing a considerably lower drag acceleration at its higher altitude, has remained at virtually the same height after the onset of low solar activity in 2003/2004. The slope in the density time-series as encountered by GRACE, amounting to a 10-fold decrease over its first 7 years in orbit, is therefore directly related to the effect on the thermosphere of the declining solar activity over the same period.

High and low solar activity discrepancy

For both satellites, there is a reasonably good agreement between observed and model density during the high solar activity years 2000–2003. During those years, periods of under- and overestimation by the NRLMSISE- 00 model alternate on timescales of weeks to months. At low solar activity, however, from mid-2006 to early 2010, the observed densities are consistently lower than the model predictions. This is in line with other observations of unusually low neutral density during the recent deep solar minimum, as reported by Emmert et al. [8]. These lower than predicted density levels seem to be directly related to unexpectedly low solar radiation measurements during the same period [28], which solar physicists are currently trying to understand [12].

Density biases

It should be noted that the overall biases between both sets of observations and the model is determined by the choice of the fixed ballistic coefficients used for both satellites. Since, in reality, the ballistic coefficients are variable and their computation is subject to uncertainties (see Sect. 3.4.3), the choice of these fixed values is rather arbitrary. A deeper investigation of the issue of data/model biases will be provided in Sect. 5.3.

Short-term variability in TLE-derived data

Another striking feature of the time-series in Fig. 5.1 is the increased amplitude of the short-term variability of the observations at low solar activity and at the higher altitude of GRACE. This 'noisiness' is not reflected in the model output, and its cause

is not yet fully investigated. When including TLE-results from other satellites, it is clear that this variability generally occurs for low drag situations (high altitude, low activity), and it is therefore likely related to errors in the mean motion data obtained from the TLEs of approximately fixed magnitude, which becomes significant when the signal strength is reduced.

Note that the high activity densities are strongly correlated with variations in solar and geomagnetic activity. At low solar activity, solar and geomagnetic activity are quite constant at a very low level. Therefore, during the years 2006–2009, the semi-annual variation, with density peaks around the equinoxes, is more clearly visible in the density time series for both satellites.

5.1.2 Data/Model Density Ratios for Many Objects

Figure 5.2 shows the variation of data/model density ratios for various satellites and space debris objects, including the data from CHAMP and GRACE from the previous plot.

We will, at first, focus here on Explorer 8, as a typical example of an object used for TLE-derived density analysis. Explorer 8 was launched in November 1960 in an orbit with perigee/apogee heights of about 416/2290 km, and an inclination of 50°, to study the composition of the neutral atmosphere and ionosphere. The eccentric orbit ensures a long lifetime. After 50 years in orbit, the perigee/apogee heights have become about 340/580 km. The satellite has long ceased active communications and is therefore no longer under control, but space surveillance tracking is still used for drag and thermosphere density studies, such as in this example. Orbit information on objects such as this one enables the analysis of density variations over the decades. The high eccentricity causes a very localised sampling of the density, near the orbit's perigee. This perigee location continually drifts over time covering all local solar times and all latitudes in between the limits set by the orbit inclination.

The similarity between the density time series of Explorer 8 and the other objects shows that, despite this localised sampling, the density model error is dominated by global phenomena, such as the level of solar activity and the semi-annual variation. An outlier in the observations for Explorer 8, perhaps due to a misidentification of the object in the TLE generation from SSN tracking data, is clearly visible in early 2000.

The time series in Fig. 5.2 are sorted by the average perigee altitude of the objects. The similarities between the time series illustrates the strong correlation of density model error for different satellites in different orbits. The Figure also shows that the amplitude of the variations of the ratios, and therefore of the model error, increases with increasing perigee altitude.
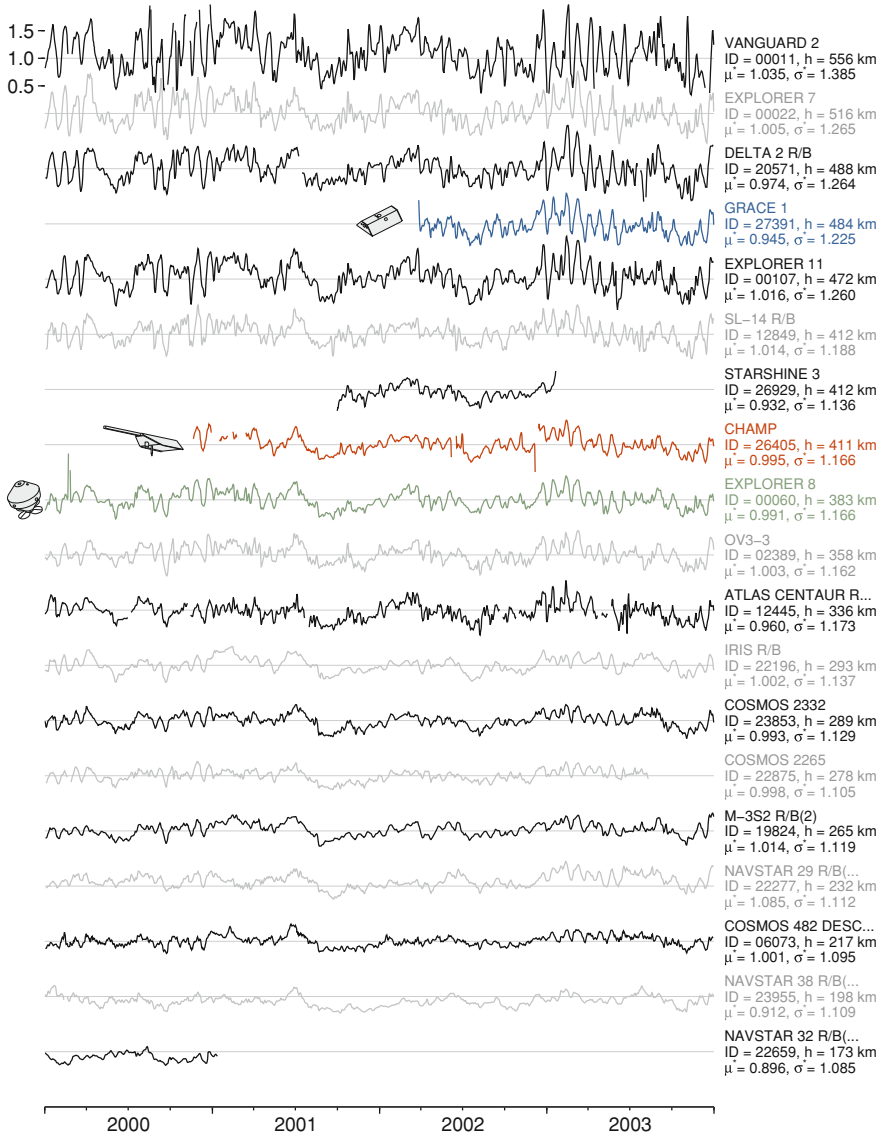
Fig. 5.2 Time series of TLE-derived velocity-weighted densities for various satellites and space debris objects, sorted by average perigee altitude over the time interval

5.2 Two Case Studies Using Accelerometer Data

Two case studies involving the interpretation of accelerometer-derived data will be presented below. Section 5.2.1 analyses a special manoeuvre made by CHAMP, that allows for an evaluation of the iterative density and wind processing algorithm.

Subsequently, a period of data in April 2005 is used in Sect. 5.2.2 to compare the results of both satellites with each other and with the output of empirical models.

5.2.1 CHAMP Flying Sideways

The CHAMP mission provides an interesting possibility for further assessment of the capabilities of the iterative algorithm. There have been two periods, on October 7 and 8, 2001, and on November 6, 2002, when CHAMP was commanded to fly sideways for about 11 h, or 7 orbital revolutions.

Acceleration data

Figure 5.3 shows the CHAMP calibrated accelerometer data with modelled radiation pressure accelerations removed, for three days around the November 6, 2002 event. Notice that the drag acceleration is in the Y_{SBF} direction during the sideways-flying period, and is approximately 3 times larger than the drag in the X_{SBF} direction during the surrounding nominal (forward-flying) period. This higher drag acceleration in the sideways-flying period is the result of the larger frontal area, but this is offset to some extent by a lower drag coefficient.

Basically, these two attitude modes present us with two completely different aerodynamic shapes with respect to the flow. In the forward-flying configuration, CHAMP is a long and slender shape, with only a small frontal area, but with its large solar panel and bottom surfaces oriented parallel to the particle stream. A considerable fraction of the aerodynamic acceleration on the shape in this orientation is due to collisions with these parallel surfaces, which cause an increase in the drag coefficient. In the sideways-flying configuration, CHAMP is a wide and short object, with most of the aerodynamic force caused by near-frontal collisions, and only a relatively small contribution of parallel surfaces.

Unfortunately, the satellite was not designed to operate in the sideways-flying mode for extended periods of time. Both the attitude control actuation and the thermal environment of the accelerometer instrument will have been quite different than during nominal operations on the surrounding days. The different orientation of the instrument also has consequences for the accuracy of the instrument calibration (see the discussion at the end of Sect. 4.2.4). These considerations are important to keep in mind in the interpretation of these data.

Results

Figure 5.4 shows the density and wind retrieval results from the iterative algorithm for three orbits, 24-h apart, on the sideways-flying day and the two surrounding days. The local solar time (LST) at the equator crossings was approximately 09:40 (ascending) and 21:40 (descending).
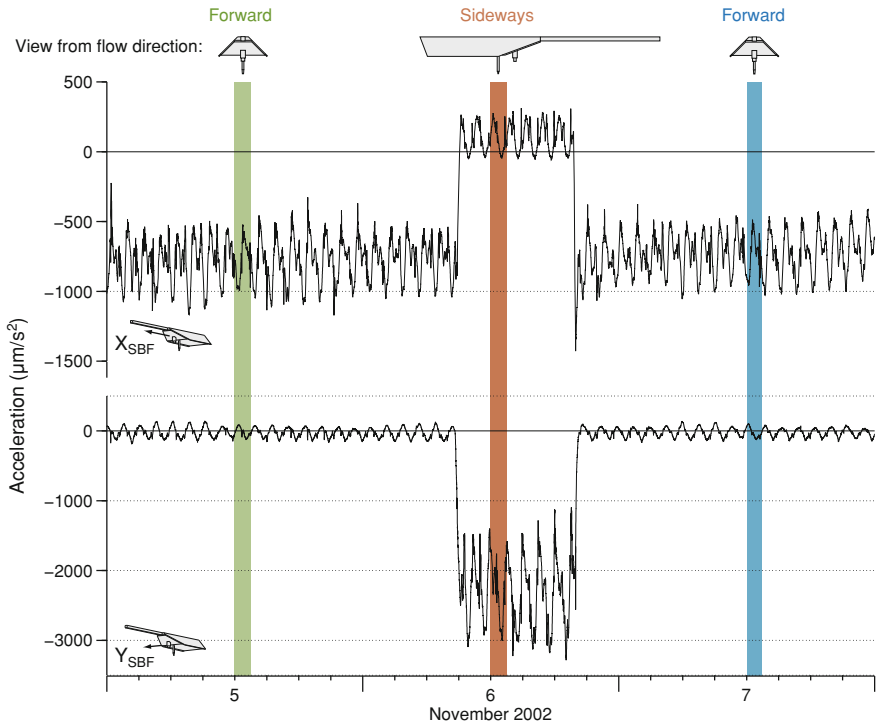
Fig. 5.3 Calibrated CHAMP accelerometer data with modelled radiation pressure removed, for three days in November 2002, surrounding the sideways-flying attitude period on November 6, 09:00–20:00 UTC. The shaded areas correspond to the three orbits of data starting at 12:00 UTC each day, which are used in Fig. 5.4

The 24-h interval between the three orbits was chosen so that each of the orbits crosses the auroral zones at approximately the same magnetic local time, thus reducing variability by keeping the sampling characteristics as similar as possible. Since proxies of solar EUV radiation and geomagnetic activity also show little variation during this period, the NRLMSIS and HWMSEVEN model curves for the orbits on the three different days show largely the same magnitudes and patterns.

Density

The large scale variations in density are determined by variations in orbital altitude, between 396/400 km at the evening/morning equator crossings and 417 km over the poles, and by thermospheric features such as the diurnal density bulge, causing the absolute maximum in density around the morning equator crossing. The density retrieved by CHAMP resembles this behaviour, but is in general around 15–25% lower than the model values. It is unlikely that such a large offset can be fully attributed to systematic errors in the accelerometer-derived data alone, so it should
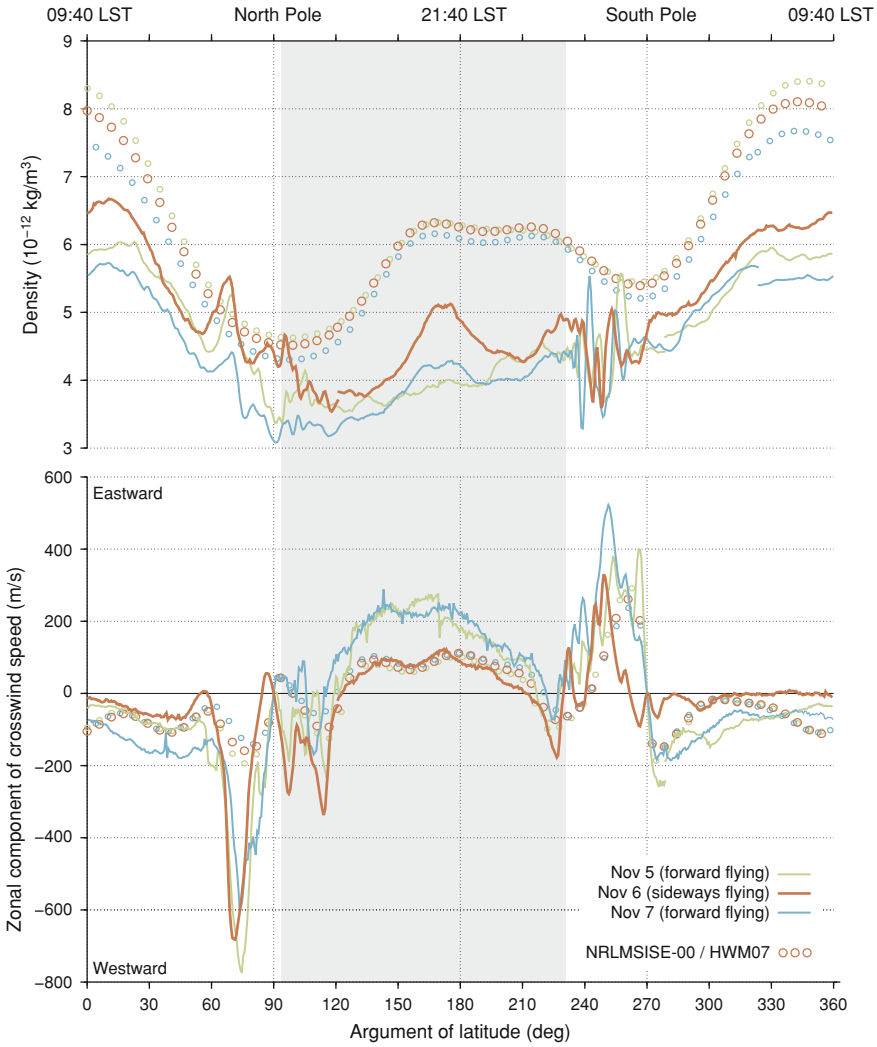
Fig. 5.4 Comparison of three orbits of density (*top*) and wind (*bottom*) data. Accelerometer-derived data are drawn using *solid lines*, while NRLMSISE-00 and HWM07 model data are drawn using *open circles*. The colours correspond to the data marked in the same colours in Fig. 5.3, for forward and sideways flying orbits. The shaded background indicates where the satellite was in eclipse

at least partly be attributed to a bias in the density model. Such an offset between model and observed density was seen in Fig. 5.1 as well. This topic will be covered in more detail in Sect. 5.3.

Another significant difference between the density model data and CHAMP results are the sharp spikes at high latitudes, especially in the Southern Hemisphere

evening sector, which could be related to upwelling in the polar cusp region [17, 24], and/or errors due to poorly modelled in-track wind variations.

The CHAMP density data from the sideways-flying period show largely the same features as the data for the two surrounding days, but they are on average slightly higher. This difference could be due to different errors in the satellite geometry and aerodynamic interaction modelling for the two attitude modes, but it is not possible to distinguish such an effect from the natural density variability over these days.

Wind

We shall now turn our attention to the zonal wind data in Fig. 5.4. The HWM07 model zonal winds at low- and mid-latitudes are at around 100 m/s, eastward in the evening and westward in the morning, away from the sub-solar density bulge. The HWM07 winds at high latitudes show much shorter wavelengths, related to localised energy inputs related to the geomagnetic field, with peaks of around 300 m/s eastward in the Southern Hemisphere and of around 200 m/s westwards in the Northern Hemisphere. The most obvious feature in the CHAMP-derived data is this Northern Hemisphere peak, which reaches an amplitude of over 700 m/s, nearly three times the model prediction. These strong westward winds blowing from the day to the night side across the pre-noon sector are common in the auroral region [11, 18]. The smaller Southern Hemisphere peak shows a better consistency between the HWM07 model and CHAMP-derived data than this Northern Hemisphere peak.

At low- and mid-latitudes, the HWM07 model predictions for the three consecutive days overlap to a large extent. It is therefore interesting to see that the CHAMP-derived zonal winds from the two days with a nominal orientation also show a large degree of consistency with each other, even though the maximum amplitudes are significantly higher than the model values. The wind observations derived for the sideways-flying period are inconsistent with the observations from the two surrounding days. They are more consistent in comparison with the model results, though, especially on the descending (evening) pass.

There are several reasons to believe that the wind data from the sideways-flying period are the most accurate, and that therefore the higher amplitudes at mid-latitude from the nominal data are an artefact of the data processing and do not indicate deficiencies in the HWM07 model. First of all, the compact satellite shape is less sensitive to errors in the aerodynamic and geometrical modelling than the elongated shape of the nominal configuration, as was seen in Fig. 3.14. Secondly, the larger frontal area when flying sideways results in a larger drag signal. This makes the wind derivation less susceptible to acceleration errors due to issues with the calibration and solar radiation pressure modelling. The fact that the evening wind data seem less noisy than the equivalent data on the surrounding days supports this reasoning. And thirdly, and perhaps a bit more tentatively, since the accelerometer axes for the along-track and cross-track axes are switched during the manoeuvre, the wind determination can benefit from the accurate GPS-derived calibration of the X_{SBF} axis, during the surrounding days. This axis is in the crucial crosswind direction during the sideways-flying period.

The difference between the low latitude crosswind data on the sideways-flying days and the surrounding days amounts to 50–150 m/s. If we assume that the sideways-flying crosswind data are accurate, and that the day-to-day natural variability of the low-latitude zonal wind is limited, this means that the wind errors in the nominal configuration are at this level as well. This conclusion is in line with the results from the simulations (Table 4.2).

Data from the other sideways-flying period in October 2001 lead to similar conclusions. These data are not presented here because the density and wind results are less clear than those for the 2002 event. This is mainly due to atypically large instrument calibration errors and the larger day-to-day density variability at that time. Further studies into the instrument calibration as well as into the geometrical and aerodynamic modelling are required to reconcile the wind data from both attitude modes.

5.2.2 Geomagnetic Storms Sampled by CHAMP and GRACE from Co-Planar Orbits

We will now move on to the analysis of samples of the accelerometer-derived density data from CHAMP and GRACE. Because of the much higher level of detail available in this data type, compared to the TLE-derived data shown in Sect. 5.1, a different way of plotting is required to facilitate the analysis.

In the previous chapter, Figs. 4.9 and 4.10 already showed examples of input and output of the accelerometer processing algorithms, for CHAMP and GRACE, respectively. Data from two orbits on April 5, 2005 were used in that example, because of the close alignment of the orbital planes of both missions at that time. In this section, a period of 11 days around the same time is selected for more elaborate comparisons of these observations with models.

Time vs. argument of latitude grids

Figure 5.5 shows the observed densities for both satellites, gridded as a function of GPS-time on the X-axis and the argument of latitude on the Y-axis (see Sect. 2.6.3). With each vertical line in the grid representing a single orbital revolution, this type of plot provides a compact but detailed view of the data. The dense 10-second temporal resolution of the accelerometer measurements is represented in the vertical direction, while the horizontal direction allows the tracking of density features from orbit to orbit, at a typical time scale of about 100 min.

Local solar time precession

Note that while universal time progresses forward along the Y-axis, the orbital precession for both satellites causes local solar time to move in the opposite direction. Local solar time values at the equator crossings are printed in white in the Figure.
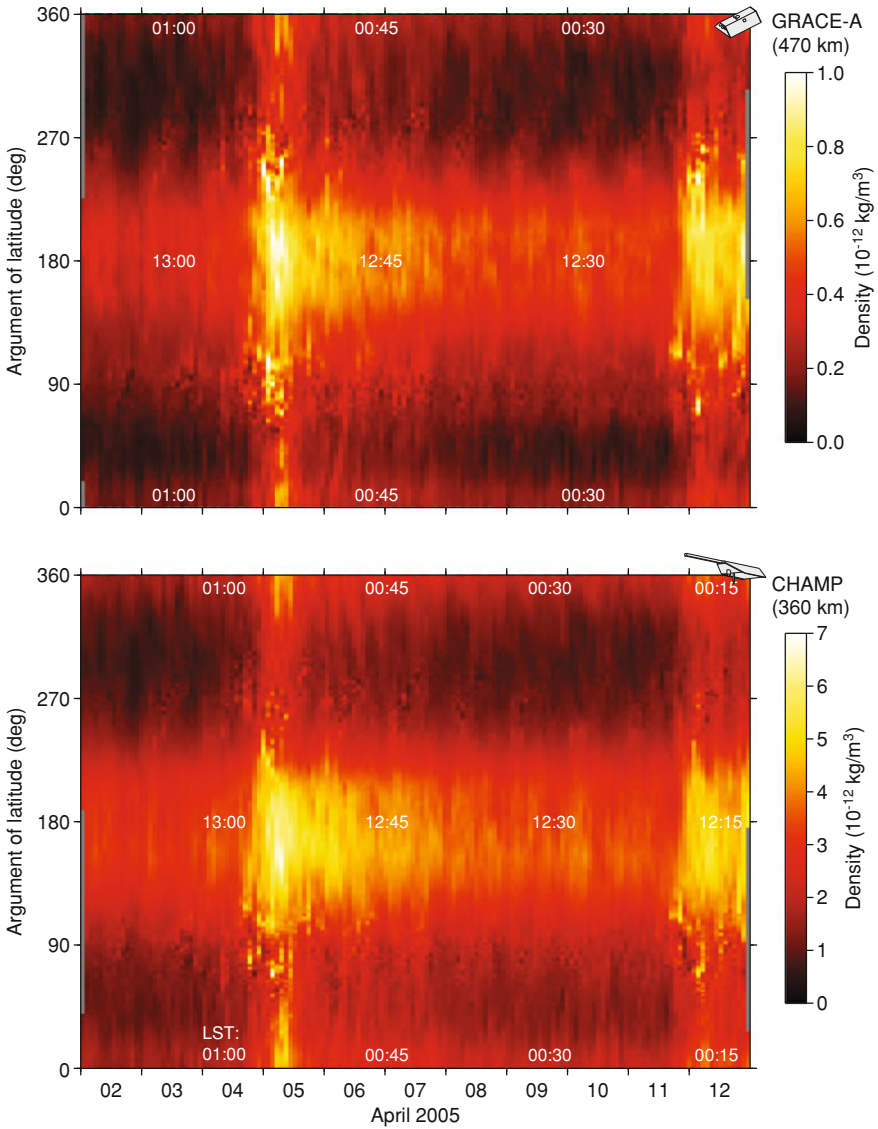
Fig. 5.5 Density along the tracks of GRACE-A (*top*) and CHAMP (*bottom*), for April 2–12, 2005. During these days, both missions were in nearly co-planar orbits

As can be seen, the local solar time at the ascending equator moves from about 13:15 to 12:15 during the 11 days for CHAMP. Because of its higher altitude, GRACE precesses just slightly slower. Due to this slight difference in the precession rates, the co-planar arrangement of the orbits of the two missions occurs quite rarely.
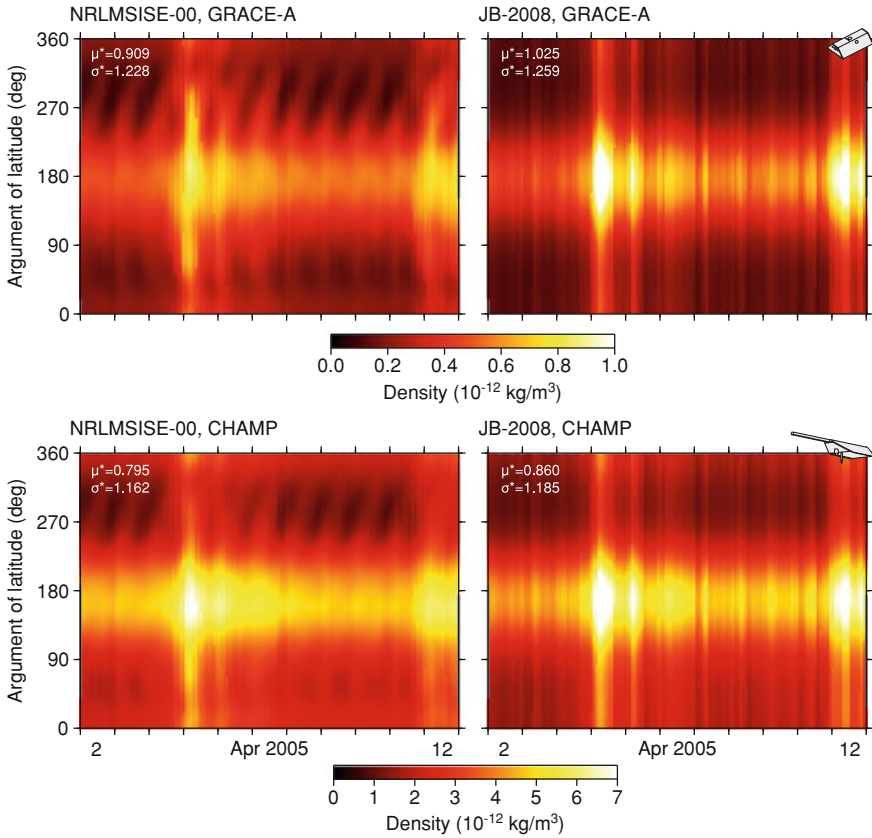
Fig. 5.6 NRLMSISE-00 and JB-2008 model densities evaluated along the GRACE-A tracks, equivalent to the observations of the previous figure

Equivalent model densities

Figure 5.6 provides the equivalent NRLMSIS- 00 and JB2008 model data for GRACE, sampled and gridded for the same times and locations. It is immediately obvious, by comparing this Figure with Fig. 5.5, that the observations contain much detail that is missing in the models. Some specific features will be pointed out below. The log-normal means μ^* and standard deviations σ^* of the data/model density ratios for the visualised data are printed in the top-left corner of the frame, for each of the models.

Diurnal density variation and polar region variability

Both satellites sample the diurnal density bulge near its maximum during their descending passes. The bulge is visible as the horizontal band of high density surrounding $180°$ argument of latitude, at the centre of each of the panels. This density maximum

is similarly represented in both the data and the models. Minimum densities in the data are reached at mid-latitudes on the night-side. The local maximum in density at the night side equator is clearly visible in the data and in the NRLMSISE- 00 model output, around 0 and 360 degrees argument of latitude, but not in the JB2008 model output. This local maximum could simply be caused by the local minimum altitude at the equator, or by a phenomena in the thermosphere-ionosphere system known as the midnight density maximum [1, 29]. The determination of the relative contribution of these two possible drivers would require a more sophisticated analysis than presented here.

The North and South Poles are crossed at 90° and 270°, respectively. The data plots show a high variability of density in these polar areas, possibly related to upwelling in the cusp regions [17], or density errors due to high in-track winds. This high frequency variability is not present in the empirical models. These show a much more smooth transition at the poles.

Geomagnetic storms

The selected period shows a classical example of a geomagnetic storm. The maximum value of the K_p index is 7, for the time interval of 00:00–03:00 UTC on April 5. The effects of the storm on density starts at high Northern latitudes late in the UTC day on April 4, and reaches its maximum intensity a few hours after the maximum geomagnetic activity is reached, in mid-morning on the 5th. It takes a couple of days before the density has returned to pre-storm levels, at which point a second storm occurs (maximum K_p of 5) on April 11/12.

When comparing Figs. 5.5 and 5.6, it is clear that the data contain much more detailed information on the development of the storm. Figure 5.5 shows several relatively short density peaks, occurring at different latitudes as the storm progresses. There might even be indications of travelling atmospheric disturbances [5] in the data. Figure 5.6 on the other hand, shows that the models were not made to represent such fine detail. The reason for this is that they rely on low degree and order spherical harmonic fits and energy input proxies with relatively coarse temporal resolutions.

For the JB2008 model, the enhancement of density occurs at all latitudes simultaneously. This model uses the hourly $D_s t$ index to drive the storm-time density variations. The NRLMSISE- 00 model, on the other hand, has a lower temporal resolution, because it uses the 3-hourly *ap* index. It does contain a more accurate representation of the influence of the geomagnetic field on the thermosphere, though. This is most clearly visible in the diagonal patterns occurring on the night-side, at the top and bottom of the plot in Fig. 5.6, corresponding nicely to the data in Fig. 5.5.

5.3 Density Ratio Statistics

The previous sections provided an overview of the contents of the TLE-derived data, as well as several detailed looks at accelerometer-derived data obtained under special conditions. In several of the plots, the log-normal mean μ^* and standard deviation
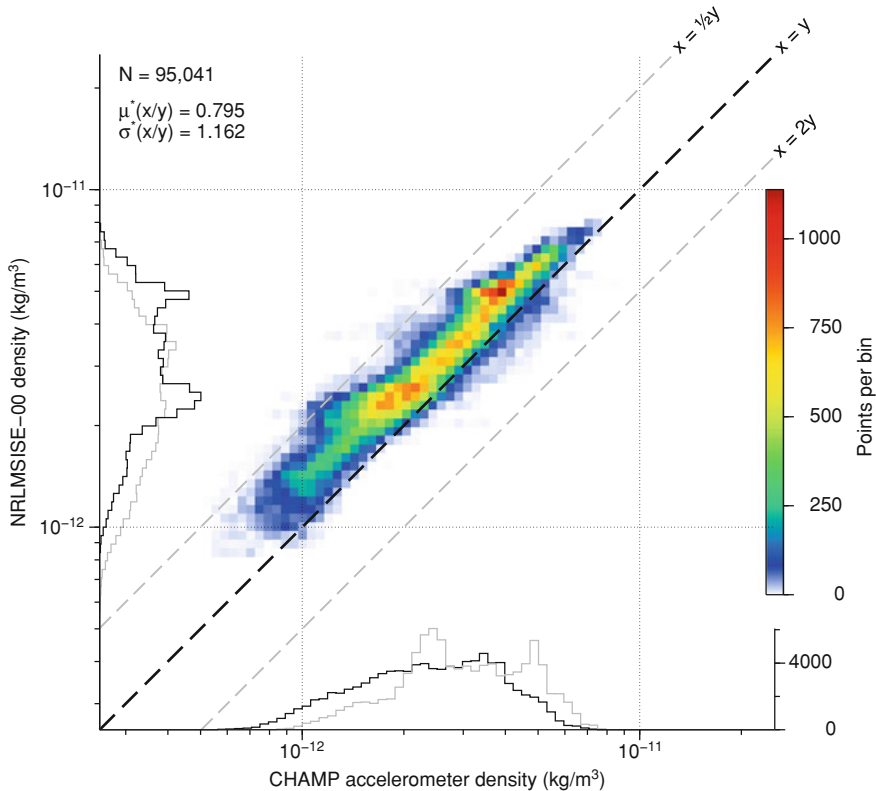
Fig. 5.7 Histogram showing the distribution of CHAMP observations and equivalent NRLMSISE-00 model densities in early April 2005

σ^* were provided to quantify how well the data compared to models. We will now continue to explore such statistics in more detail.

The statistics will be used for various purposes: Sect. 5.3.1 will start with an assessment of various empirical density models. Section 5.3.2 demonstrates how accelerometer processing choices affect density data accuracy and data/model biases. The accuracy of the TLE-density processing will be evaluated in Sect. 5.3.3. Finally, in Sect. 5.3.4, the large amounts of accelerometer-derived density data from CHAMP and GRACE- A will be compared with models in terms of various location coordinates, solar activity proxies and geomagnetic indices, in order to locate strong and weak parts in both data and models.

Before continuing, we will have one last look at the data from early April 2005, featured in Figs. 5.5 and 5.6. Figure 5.7 shows histograms of the CHAMP density data, with the accelerometer-derived data on the horizontal axis, and the equivalent NRLMSISE- 00 values on the vertical axis.

The line graphs at both axes show simple two-dimensional histograms of their corresponding data set in black. The data set corresponding to the other axis is plotted

Table 5.1 Log-normal statistics of accelerometer-derived over model density ratios for CHAMP and GRACE, comparing various models

Model	CHAMP		GRACE-A	
	μ^*	σ^*	μ^*	σ^*
CIRA-72	0.670	1.274	0.726	1.407
DTM-78	0.688	1.348	0.678	1.569
MSIS-86	0.706	1.284	0.744	1.406
DTM-94	0.705	1.358	0.724	1.556
NRLMSISE-00	0.710	1.274	0.753	1.394
Jacchia-Bowman 2006	0.731	1.248	0.831	1.388
Jacchia-Bowman 2008	0.789	1.237	0.910	1.377
HASDM(calibrated)	0.813	1.163	0.984	1.284

in gray, so that differences can be easily identified. In this case, for example, the model data clearly contains larger values than the observations, and it shows a more distinct two-peak structure in the histogram.

Figure 5.7 also displays a three-dimensional histogram in colour, showing how well the data and model match. For a perfect match, all points would be positioned along the central diagonal. The log-normal mean and standard deviation of the data/model density ratios have a direct relation to the cloud-like shape in this plot. The log-normal mean μ^* is a measure of the offset of the cloud of data points with respect to the central diagonal. The log-normal standard deviation σ^* is a measure of the width of the cloud, with respect to the diagonal. This visualisation of the distribution of data and model values is useful to keep in mind when interpreting the statistical results in the following subsections.

5.3.1 Evaluation of Empirical Density Models

Table 5.1 provides a comparison of the statistics for various models, all evaluated using the same accelerometer-derived density data sets from CHAMP and GRACE. All available data from the start of the missions up to and including 2009 were used. These statistics represent contributions of errors in both the data sets and in the models. Possible errors in the data set will be discussed in Sect. 5.3.2. At this point, it is most interesting to first have a look at the relative differences between the models, in these statistics. Figure 5.8 provides histograms corresponding to these data and model values for both satellites and three selected models.

It is clear that the calibrated HASDM model (see Sect. 2.5) is the most accurate for both CHAMP and GRACE, in terms of both the mean and standard deviation. This is expected, because HASDM is the only model in this comparison that assimilates independent contemporaneous satellite drag data, instead of relying solely on the correlation of historical density data sets with solar and geomagnetic proxies for extrapolation.
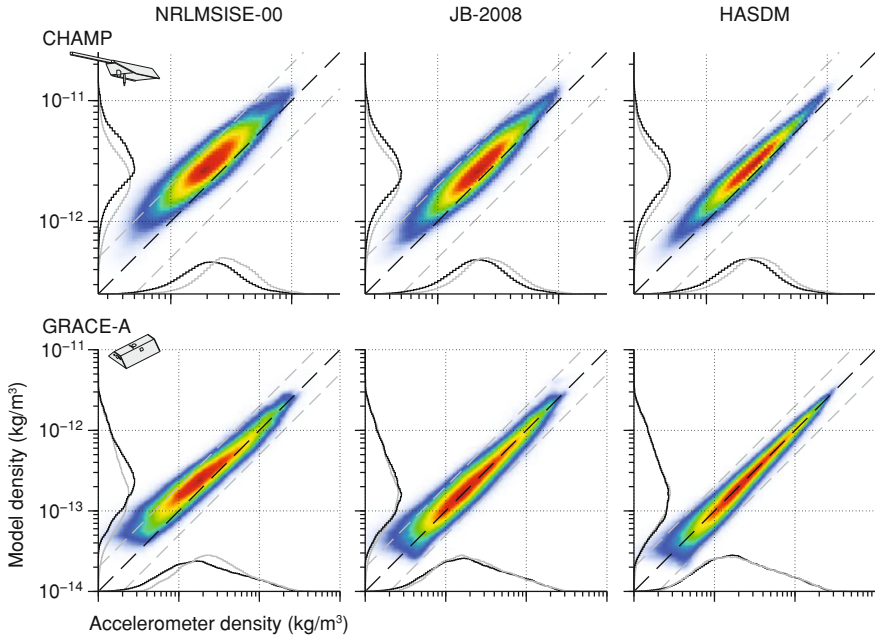
Fig. 5.8 Histograms showing the distribution of CHAMP and GRACE-A densities, with their equivalent model densities, for three different models. The X- and Y-axis scales are identical for each row

Traditional empirical models

When looking at the traditional empirical models, it is telling that until recently, subsequent generations within the same model family have not brought consistent substantial improvements. Only the recent JB2006 and JB2008 models, with their use of multiple EUV proxies, have brought a significantly improved accuracy, compared to their Jacchia ancestor model, which was very close to CIRA- 72.

It is striking that the log-normal mean values μ^* are significantly lower than 1 for all models, except for the GRACE- A/HASDM ratios. This indicates that, on average, the models return much higher densities than the accelerometer observations. A variety of reasons for this discrepancy, both in the models, and in the data, will be pointed out in this section and the next.

Influence of long-term trends in the thermosphere

One of the reasons for the higher model density values is that the models do not include long-term trends. It has been observed [7, 19] that thermospheric density decreases by 2–5% per decade, possibly due to increasing concentrations of greenhouse gases in the lower layers of the atmosphere [15, 25, 26]. This observed trend seems stronger at low solar activity and at higher altitudes. This observation is in

line with the fact that the offset of the data/model mean is smallest for models based on data contemporaneous with CHAMP and GRACE observations, such as Jacchia-Bowman and HASDM, and largest for the CIRA- 72 model based on data from the 1960s.

The difference in the values of μ^* between JB2008 and CIRA- 72 amounts to about 12% for CHAMP, and 18% for GRACE, corresponding to a trend of -3.0 % to -4.5 % per decade. This is in rough agreement with the density trends observed by Emmert et al. [7].

On top of this long-term trend, densities during the solar minimum years of 2006–2009 were exceptionally and unexpectedly low [8]. We have already observed this in Fig. 5.1. Our comparisons are for a large part based on data during that anomalous period. These density anomalies have recently been linked to anomalously low levels of extreme ultraviolet radiation, during the recent deep solar minimum [28].

Of course, this observed long-term trend and recent anomaly cannot explain the remaining discrepancy of up to about 20% between the accelerometer-derived observations and JB2008 and HASDM models, which are based on mostly contemporaneous observations. For that, we have to look at other sources of data/model biases.

Influence of ballistic coefficient and calibration choices by modellers

Another possible contribution to the large offset between data and model densities is the presence of errors in the ballistic coefficient modelling used in earlier analyses of satellite drag, on which these models were based.

For example, Jacchia used a fixed value for the drag coefficient of $C_D = 2.2$ in all of his analyses [14]. His papers and reports do not specifically mention the sources of the area-to-mass ratio values that he used, but there is a chance that there were unintentional errors in these values as well. Because it is easier to mistakenly forget to add the area of an antenna, baffle or attachment ring than to mistakenly add such an area in the calculation, and since a C_D of 2.2 is near the lower end of the range of possible drag coefficient values, it turns out that these errors tend to not average out. Instead, they tend to result in the use of ballistic coefficient values that are lower than the true values. This, in turn, leads to resulting density values that are higher than the true values.

Later modellers, such as Barlier et al. [2] and Hedin [13], have used independent observations. It is likely that they have run into uncertainties in determining accurate ballistic coefficients as well, or that they encountered similar issues in the calibration of drag and mass spectrometer data. Whether these modellers have, at some point, used Jacchia's models in order to tune the calibration and scaling of their density data sets, is not well documented, so this possibility cannot be ruled out.

The discussion above leads to the conclusion that, while variations in density can be quite well observed, it is considerably more difficult to determine absolute levels of density. Most density data sets and all past and current empirical models are therefore subject to biases with respect to the true density. When expressed as percentages, these biases can easily reach two-digit numbers. As we shall see further

on, such biases can have strong variations, correlated with location and solar activity levels.

While it is very difficult to determine the biases of these models with respect to the true density exactly, it is much easier to determine the relative biases between models and data sets, as shown in the figures and Tables in this chapter. It is likely that an analysis of such relative biases will lead to increased understanding of the underlying error sources. Understanding the extent and limitations of the influence of these error sources on the biases, and reducing such errors wherever possible, should lead to a reduction of the relative biases, hopefully in a manner such that convergence on the true absolute density is reached.

In that light, it is to be expected that the TLE- and accelerometer-derived data sets described in this thesis bring their own contributing error sources. These will be examined next.

5.3.2 Impact of Accelerometer Density Processing Choices

The problems in obtaining correct ballistic coefficient values, which were described above in the context of possible biases in historical empirical models, apply to the new TLE- and accelerometer-derived data sets as well. Details on such issues, concerning the processing of CHAMP and GRACE data, were already described in Chap. 3. The following important possible sources of error were identified:

- Lack of in-situ data on the atmospheric composition and temperature. These variables affect the random thermal motion of incoming particles;
- Limited knowledge of the gas-surface interaction, in particular concerning the extent of energy accommodation and diffuse or quasi-specular reflection of the atmospheric particles;
- Difficulty in the exact modelling of the satellite outer surface geometry for complex satellite shapes, even when detailed drawings are available.

Table 5.2 illustrates the effect on the log-normal mean and standard deviation of several choices that can be made in the accelerometer density processing. The ratios are all calculated with respect to the HASDM calibrated model data, evaluated by Bruce Bowman [personal communication, 2008, 2010] along the satellite tracks. This model turned out to be the most accurate choice in Table 5.1.

The baseline configuration, at the top row of Table 5.2, contains the same values as the bottom row of Table 5.1. For this configuration, a tuned panel model was applied, for which the frontal areas along the principal axes of the satellites matches exactly with the manufacturer specifications. The modification by Moe and Moe [21] of the aerodynamic equations by Sentman [27] (see Sect. 3.4.2) were applied for the baseline processing, assuming diffuse reflection with complete energy accommodation ($\alpha = 1.0$).

Two alternative processing choices were tested. For the second row of the Table, the energy accommodation coefficient was lowered to $\alpha = 0.8$. For the third row,

Table 5.2 Log-normal statistics of accelerometer-derived over HASDM calibrated model density ratios for CHAMP and GRACE, comparing various accelerometer processing choices

Model	α	CHAMP		GRACE-A	
		μ^*	σ^*	μ^*	σ^*
Baseline	1.0	0.813	1.163	0.984	1.284
Lower accommodation	0.8	0.757	1.163	0.870	1.285
Different panels	1.0	1.019	1.161	1.017	1.284

an alternative panel model was used. In the case of CHAMP, the panels as specified by Bruinsma and Biancale [4] were applied, resulting in a 14% lower frontal area, as seen along the satellite's X-axis, and a 8% larger area as seen along the Y-axis (see Table 3.1). For GRACE- A, the nadir-pointing boom containing the S-band antenna and magnetometer was omitted from the panel model [3], resulting in a 4.6% lower area. Both these modifications of the geometry representation in the data processing can be seen as a deterioration with respect to the true satellite geometry.

As can be seen in Table 5.2, the modifications made to the geometry and energy accommodation coefficient significantly changes the log-normal mean values, which ideally should be $\mu^* = 1.0$. This illustrates how easily biases can get introduced in the data. Unfortunately, the lowering of the accommodation coefficient from its maximum value of $\alpha = 1.0$ results in greater biases between data and model, indicating that the uncertainty in this parameter can not be used to explain the data/model bias. The use of the different panel models does remove much of the bias. However, the result is then based on a physically deteriorated representation of the satellite.

The effect of these modifications on the standard deviations is very small. However, in the case of the modified panel geometry for CHAMP, the reduction is unexpected. The difference looks hardly significant, but a possible explanation could be that the more unrealistic geometry model, which mainly reduces CHAMP's panel area perpendicular to the stream of atmospheric particles, compensates errors in the gas-surface interaction. This could indicate that the contribution to the aerodynamic force of the panels parallel to the stream is greater in reality than the aerodynamic calculations account for. As explained in Sect. 3.4.3, a higher temperature or a relatively larger contribution of Helium atoms results in a larger random thermal motion. This leads to an increased relative contribution of the parallel satellite wall panels to the aerodynamic force. The temperature and composition variables are currently obtained from the output of the NRLMSISE- 00 model, which likely contains temperature and composition biases, similar to the density biases that were already discussed. It remains to be investigated whether attempts at improving these temperature and composition inputs in the accelerometer processing also result in a lowering of the standard deviation of the data/HASDM ratio.
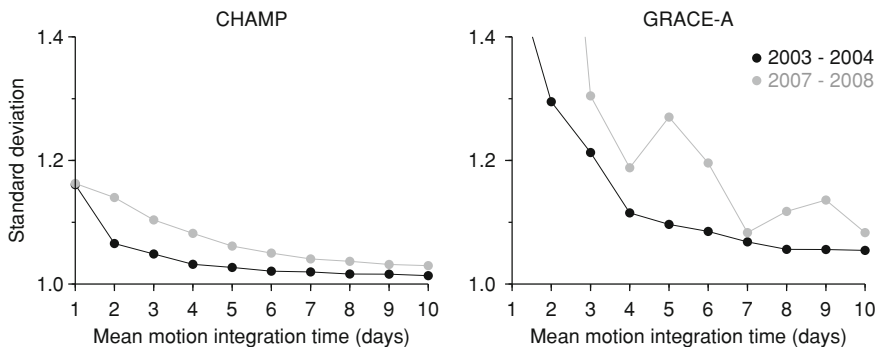
Fig. 5.9 Log-normal standard deviation of the TLE-derived over equivalent velocity-weighted average accelerometer-derived densities for CHAMP (*left*) and GRACE (*right*)

5.3.3 Comparison of Accelerometer and TLE-Derived Densities

We will now turn our attention to an accuracy assessment of the TLE data. In order to assess the influence of the integration interval length on the accuracy, we have analysed the log-normal standard deviation of TLE-derived over accelerometer-derived velocity-weighted mean densities. These values are computed according to the equations for $\bar{\rho}_O(t_{ik})$ and $\bar{\rho}_M(t_{ik})$ in Eqs. (4.6) and (4.7), respectively. For this test, the mean motion data from the TLE s of CHAMP and GRACE were linearly interpolated. Subsequently, the density values were computed for every day, using the interval $t_i = t_{ik} - l/2$ to $t_k = t_{ik} + l/2$, where the epoch time t_{ik} is at 12:00 UTC at the day in question and l is the integration time span, varying from 1 to 10 days. Days for which manoeuvres or data gaps were encountered in any of the 10 surrounding days were omitted from the analysis.

The results of this analysis are plotted in Fig. 5.9, showing selections of the data from 2003 to 2004 at high to moderate solar activity and from 2007 to 2008 at low solar activity. In general, the agreement between the two types of data increases with integration time interval length. A trade-off can therefore be made between accuracy and temporal resolution of the TLE-derived data. The CHAMP results are much more accurate than those for GRACE. This difference is likely related to the higher drag signal strength at CHAMP's lower altitude. The drag signal strength seems also to be the main driver for the differences between the high and low solar activity data.

Table 5.3 illustrate the relative distributions of 3-day averaged densities, derived from TLEs, the accelerometers, and the HASDM and NRLMSISE- 00 models, over all available years until the end of 2009. It is interesting to compare these values to the 10-second full rate data presented in the previous sections. The different selection and weighting of the data results in comparable, but slightly lower mean values for the ratios. The standard deviations are smaller for the averaged data, because high frequency density variations, such as those due to geomagnetic activity are largely averaged out. The reduction of the standard deviation for the accelerometer/HASDM

Table 5.3 Log-normal statistics for various density ratios. The densities are daily velocity-weighted average values, using an integration time interval of 3 days

Ratio	CHAMP		GRACE-A	
	μ^*	σ^*	μ^*	σ^*
TLE / NRLMSISE-00	0.671	1.235	0.699	1.384
TLE / HASDM	0.776	1.103	0.912	1.265
TLE / accelerometer	0.985	1.098	0.979	1.264
Accelerometer / HASDM	0.788	1.026	0.932	1.037

ratios to the level of 2–4% is especially striking. When discounting the ballistic coefficient related bias, these two sources of density data are in very good agreement.

The TLE-derived data for CHAMP and GRACE have standard deviations with respect to both the HASDM and accelerometer data of about 10 and 26%, respectively. The low accuracy for the GRACE data, which could be typical for TLE-objects at similar altitude and solar activity levels, might at first seem rather disappointing. However, when compared with the level of error in NRLMSISE- 00, it must be concluded that the data still contain valuable information for empirical model improvement.

The TLE-derived data errors that are implied by these statistics are rather higher than the short-term accuracy of ± 2% and long-term accuracy of 5–10%, quoted by Emmert [6] for his TLE-derived global density data set. There are two reasons why his data set may be more accurate. First of all, Emmert [6] averaged the density for several objects at similar altitudes, trading in spatial resolution to reduce the effect of random errors in single TLEs. Secondly, our data processing is based on strict 3-day integration intervals, while Emmert [6] used intervals of 3–6 day.

5.3.4 Statistics for Location and Activity Bins

So far, statistics of data/model ratios have been provided for all available data combined. In this section, the measurements were first sorted in bins, according to several binning variables, indicating measurement locations and activity levels. Subsequently, the log-normal mean and standard deviation were calculated for all the data in each bin. The binning was performed as a function of two variables at the same time, so that these statistics per bin could be plotted according to a colour scale, as the third dimension on a 2D map. The figures in this section provide such maps of log-normal mean and standard deviation values, for several combinations of the binning variables.

Latitude and longitude

Figure 5.10 provides a first example. In this Figure, the data are binned by longitude and latitude, so that the statistics can be plotted on a geographical map.
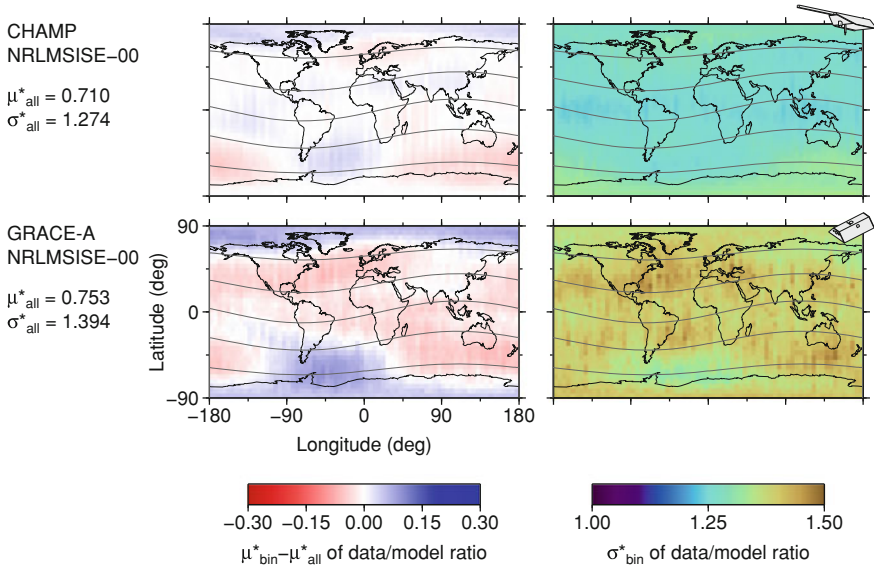
CHAMP
NRLMSISE–00

$\mu^*_{all} = 0.710$
$\sigma^*_{all} = 1.274$

GRACE-A
NRLMSISE–00

$\mu^*_{all} = 0.753$
$\sigma^*_{all} = 1.394$

Latitude (deg)

90

0

−90
−180   −90   0   90   180
Longitude (deg)

−0.30 −0.15 0.00 0.15 0.30
$\mu^*_{bin} - \mu^*_{all}$ of data/model ratio

1.00   1.25   1.50
σ^*_{bin} of data/model ratio

Fig. 5.10 Log-normal mean and standard deviations of the ratios of accelerometer derived densities over NRLMSISE-00 model densities, binned by latitude and longitude. The nearly horizontal curves indicate lines of equal geomagnetic latitude

The panels on the left-hand side of each of these plots show the offset of the mean of each bin (μ^*_{bin}), with respect to the global mean (μ^*_{all}), indicated in red and blue. The deviations for GRACE- A are somewhat stronger than for CHAMP. In general, the accelerometer-derived densities are higher near the poles and lower near the equator, with respect to the model, when the global bias is removed. There is relatively little structure in the longitudinal direction, except for a slight anomaly over the South Atlantic, which could be related to an influence on the density by the deformation of the geomagnetic field at that location. Such excursions from the dipole approximation of the geomagnetic field are not included in the model. This could perhaps become an interesting feature for closer study.

The right-hand side panels show the standard deviations of the ratios. CHAMP data seems in best agreement with the model at mid-to-low latitudes. GRACE- A, on the other hand, shows bands of increased data/model discrepancy on either side of the equator. This indicates that the quality of the accelerometer-derived density data could perhaps be improved by a more rigorous elimination of short spikes [9], which have been observed to concentrate at those locations as well [22].

Geomagnetic latitude and local solar time

Figures 5.11 and 5.12 show the data binned as a function of geomagnetic latitude and geomagnetic local solar time, for several models, and for CHAMP and GRACE- A, respectively. The patterns in these plots are much stronger than in the previous figure,
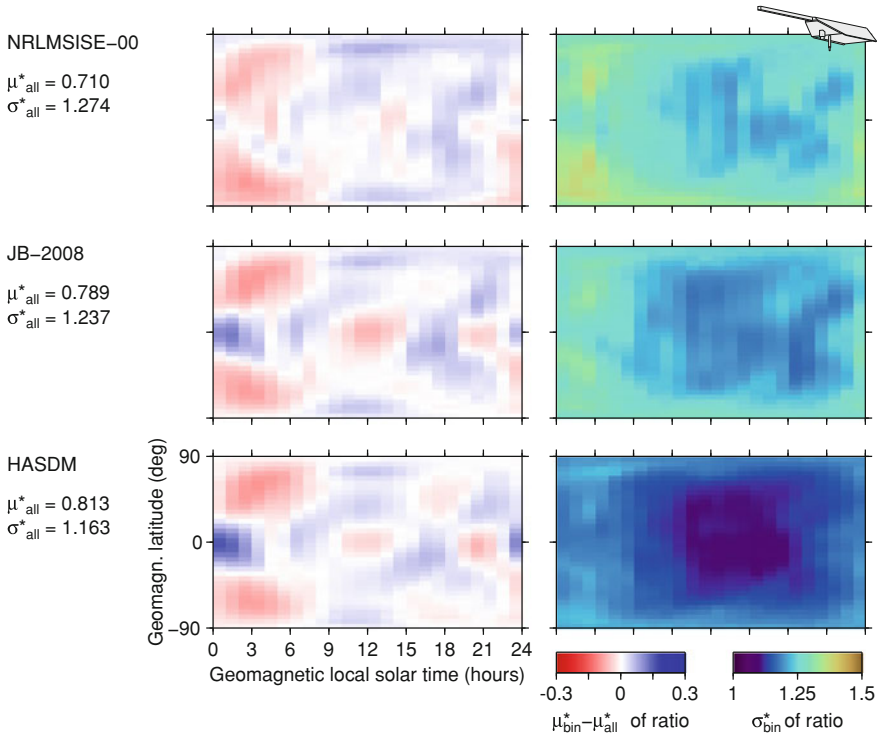
NRLMSISE−00

$\mu^*_{all} = 0.710$
$\sigma^*_{all} = 1.274$

JB−2008

$\mu^*_{all} = 0.789$
$\sigma^*_{all} = 1.237$

HASDM

$\mu^*_{all} = 0.813$
$\sigma^*_{all} = 1.163$

Geomagn. latitude (deg)

90

0

−90

0   3   6   9   12 15 18 21 24
Geomagnetic local solar time (hours)

-0.3   0   0.3   1   1.25   1.5
$\mu^*_{bin} - \mu^*_{all}$ of ratio σ^*_{bin} of ratio

Fig. 5.11 Log-normal mean and standard deviations of the ratios of CHAMP accelerometer densities over model densities, binned by geomagnetic latitude and local solar time

since these coordinates are more directly related to the geometry of the Earth and its atmosphere with respect to its primary heating source: the Sun, and with secondary heating sources, related to the magnetosphere and ionosphere.

There are several features in these figures that require attention. The standard deviation sections are the easiest to interpret. They all show minima around the early afternoon equator. This is the area of the diurnal bulge, containing maximum absolute densities, and a strong drag signal. The standard deviations are significantly higher during night and early morning, and around the poles. These areas contain low densities and drag signal, and large density variability, resulting in large relative density errors. Whether only the models, only the accelerometer-derived data, or both are weakest at these locations, is difficult to say.

The log-normal mean plots show strong and intriguing patterns, reaching quite large minimum and maximum values. The diagonal lines are likely related to solar terminator waves [10, 16, 20]. These local offsets with respect to the global mean reach the same order of magnitude as the global mean (or data/model bias) itself. It is intriguing to see that the model with the largest standard deviations, NRLMSISE- 00, has the weakest pattern in the mean. This model contains a relatively extended spherical harmonic expansion for the diurnal variation, which is apparently quite well
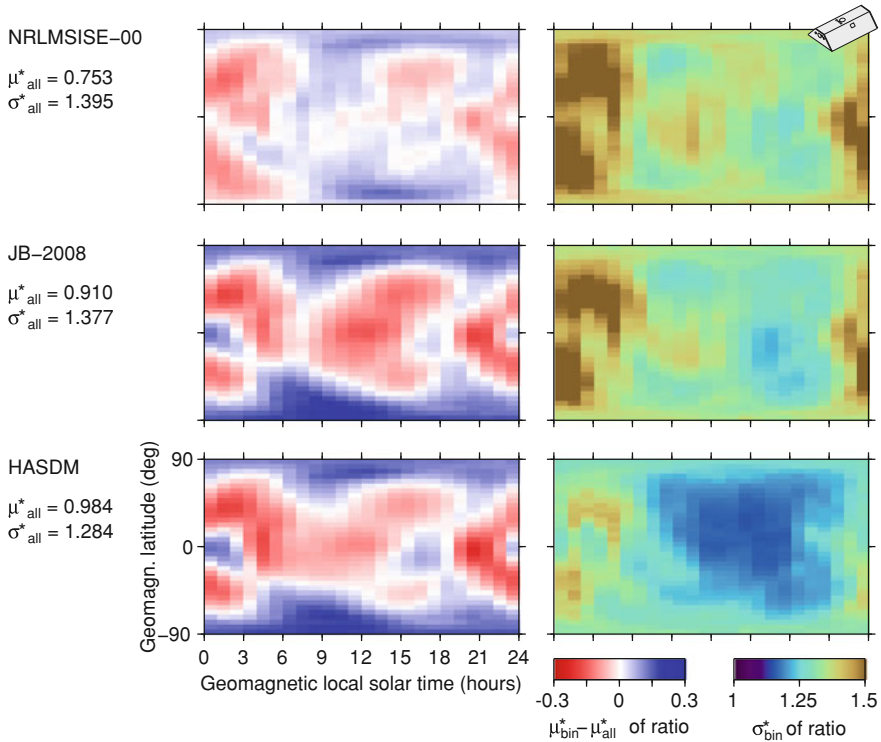
Fig. 5.12 Log-normal mean and standard deviations of the ratios of GRACE-A accelerometer densities over model densities, binned by geomagnetic latitude and local solar time

resolved by the underlying database. The JB2008 and HASDM models are both based on Jacchia's models, which include a much more simple diurnal variation function. In HASDM, a number of additional spherical harmonic parameters are estimated from contemporaneous satellite tracking data, in order to correct this diurnal function. Apparently, their effect tends to cancel over the long duration of the data sets used, and the underlying Jacchia-related pattern emerges.

For the GRACE- A data and the JB2008 and HASDM models, there is a distinct latitudinal pattern, with most of the underestimation of the data by the model around the equator, and a strong overestimation at the poles. Further analysis has shown that this distinction is only clearly visible under low solar activity conditions. It could therefore be related to weak observation data underlying these models, at high altitudes and low solar activity conditions.

Solar and geomagnetic activity

To conclude the analysis of the log-normal statistics, Fig. 5.13 shows the CHAMP and GRACE data, binned by $F_{10.7}$ and K_p, representing solar and geomagnetic activity, respectively. Bins containing no data are plotted in grey.
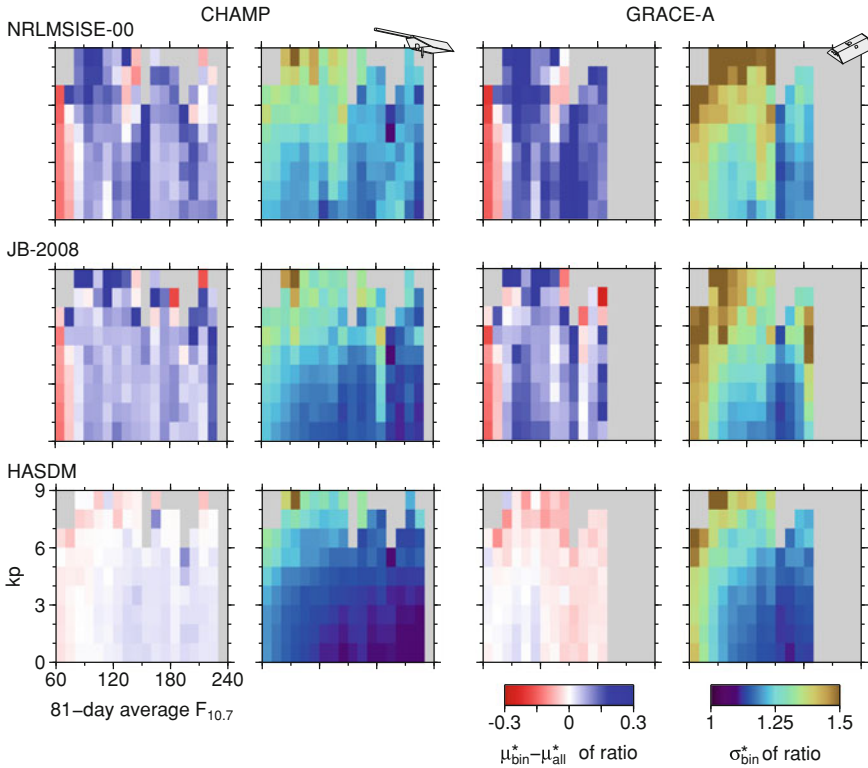
Fig. 5.13 Log-normal mean and standard deviations, binned by level of solar and geomagnetic activity, of CHAMP and GRACE-A data/model density ratios, for three different models

These panels show several features that have been encountered before. For instance, the large overestimation of density by the NRLMSISE- 00 and JB2008 models, at very low solar activity ($F_{10.7} < 75$) is clearly visible (see Fig. 5.1). For the HASDM model, the offsets of the bins with respect to the global mean are very small, indicating that the model calibration using Space Surveillance Network tracking data is successful in eliminating most solar EUV and geomagnetic activity-related errors.

The standard deviations also show a predictable pattern. The agreement between data and model is best at high solar activity, where the drag signal is strong, and at low geomagnetic activity, where short-term variability is low. This explanation is similar to the day/night and equator/pole differences encountered in the geomagnetic latitudes and local solar time plot, discussed earlier.

The patterns for both the means and standard deviations are most clear in the HASDM plots. The standard deviation panels for NRLMSISE- 00 and JB2008 show a more patchy pattern than for HASDM. These patches occur when bins contain data from only a small period of time, corresponding to these activity conditions. In

general, bins for low solar and geomagnetic activity contain more data than bins for high activity. The much cleaner pattern for HASDM is therefore an indication of the consistency of the calibrated model over all these conditions.

References

1. Arduini C, Laneve G, Herrero FA (1997) Local time and altitude variation of equatorial thermosphere midnight density maximum (MDM): San Marco drag balance measurements. Geophys Res Lett 24(4):377–380. doi:10.1029/97GL00189

2. Barlier F, Berger C, Falin JL, Kockarts G, Thuillier G (1978) A thermospheric model based on satellite drag data. Annales de Geophysique 34(1):9–24

3. Bettadpur S (2007) Gravity recovery and climate experiment product specification document (rev 4.5—february 20, 2007), GRACE 327–720/CSR-GR-03-02. Center for Space Research, The University of Texas at Austin

4. Bruinsma S, Biancale R (2003) Total densities derived from accelerometer data. J Spacecr Rocket 40(2):230–236

5. Bruinsma SL, Forbes JM (2007) Global observation of travelling atmospheric disturbances (TADs) in the thermosphere. Geophys Res Lett 34(L14103). doi:10.1029/2007GL030243

6. Emmert JT (2009) A long-term data set of globally averaged thermospheric total mass density. J Geophys Res 114(A06315). doi:10.1029/2009JA014102

7. Emmert JT, Picone JM, Meier RR (2008) Thermospheric global average density trends, 1967– 2007, derived from orbits of 5000 near-earth objects. Geophys Res Lett 35(L05101). doi:10.1029/2007GL032809

8. Emmert JT, Lean JL, Picone JM (2010) Record-low thermospheric density during the 2008 solar minimum. Geophys Res Lett 37(L12102). doi:10.1029/2010GL043671

9. Flury J, Bettadpur S, Tapley BD (2008) Precise accelerometry onboard the GRACE gravity field satellite mission. Adv Space Res 42(8):1414–1423. doi:10.1016/j.asr.2008.05.004

10. Forbes JM, Bruinsma SL, Miyoshi Y, Fujiwara H (2008) A solar terminator wave in thermosphere neutral densities measured by the CHAMP satellite. Geophys Res Lett 35(L14802). doi:10.1029/2008GL034075

11. Förster M, Rentz S, Köhler W, Liu H, Haaland SE (2008) IMF dependence of high-latitude thermospheric wind pattern derived from CHAMP cross-track measurements. Annales Geophysicae 26(6):1581–1595

12. Garcia RR (2010) Solar surprise? Nature 467:668–669. doi:10.1038/467688a

13. Hedin AE (1983) A revised thermospheric model based on mass spectrometer and incoherent scatter data—MSIS-83. J Geophys Res 88:10170–10188

14. Jacchia LG, Slowey J (1962) Accurate drag determinations for eight artificial satellites; atmospheric densities and temperatures. Smithsonian Astrophysical Observatory Special Report 100

15. Keating GM, Tolson RH, Bradford MS (2000) Evidence of long term global decline in the earth's thermospheric densities apparently related to anthropogenic effects. Geophys Res Lett 27(10):1523–1526

16. Liu H, Lühr H, Watanabe S (2009) A solar terminator wave in thermospheric wind and density simultaneously observed by CHAMP. Geophys Res Lett 36(L10109). doi:10.1029/2009GL038165

17. Lühr H, Rother M, Köhler W, Ritter P, Grunwaldt L (2004) Thermospheric upwelling in the cusp region: evidence from CHAMP observations. Geophys Res Lett 31(6). doi:10.1029/2003GL019314

18. Lühr H, Rentz S, Ritter P, Liu H, Häusler K (2007) Average thermospheric wind pattern over the polar regions, as observed by CHAMP. Annales Geophysicae 25(5):1093–1101

19. Marcos FA, Wise JO, Kendra MJ, Grossbard NJ, Bowman BR (2005) Detection of a long-term decrease in thermospheric neutral density. Geophys Res Lett 32(L04103). doi:10.1029/2004GL021269
20. Miyoshi Y, Fujiwara H, Forbes JM, Bruinsma SL (2009) Solar terminator wave and its relation to the atmospheric tide. J Geophys Res 114(A07303). doi:10.1029/2009JA014110
21. Moe K, Moe MM (2005) Gas-surface interactions and satellite drag coefficients. Planet Space Sci 53(8):793–801. doi:10.1016/j.pss.2005.03.005
22. Peterseim N, Schlicht A (2010) New investigations on twangs. In: GRACE accelerometer data. Poster presented at the Grace Science Team Meeting, Potsdam, Germany, 11–12 Nov 2010
23. Picone JM, Emmert JT, Lean J (2005) Thermospheric densities derived from spacecraft orbits-I. Accurate processing of two-line element sets. J Geophys Res 110(A03301). doi:10.1029/2004JA010585
24. Rentz S, Lühr H (2008) Climatology of the cusp-related thermospheric mass density anomaly, as derived from CHAMP observations. Annales Geophysicae 26(9):2807–2823
25. Rishbeth H, Roble RG (1992) Cooling of the upper atmosphere by enhanced greenhouse gases—modelling of thermospheric and ionospheric effects.. Planet Space Sci 40(7):1011–1026. doi:10.1016/0032-0633(92)90141-A
26. Roble RG, Dickinson RE (1989) How will changes in carbon dioxide and methane modify the mean structure of the mesosphere and thermosphere?. Geophys Res Lett 16:1441–1444
27. Sentman LH (1961) Free molecule flow theory and its application to the determination of aerodynamic forces, LMSC-448514. Lockheed Missiles & Space Company
28. Solomon SC, Woods TN, Didkovsky LV, Emmert JT (2010) Anomalously low solar extreme-ultraviolet irradiance and thermospheric density during solar minimum. Geophys Res Lett 37(L16103). doi:10.1029/2010GL044468
29. Spencer NW, Carignan GR, Mayr HG, Niemann HB, Theis RF, Wharton LE (1979) The midnight temperature maximum in the Earth's equatorial thermosphere. Geophys Res Lett 6:444–446

Chapter 6
Empirical Model Calibration

The concept of empirical density model calibration was introduced in Sect. 2.5. This chapter will provide some details on an implementation of this technique, as developed under contract for the European Space Operations Centre [7, 8, 11].

Section 6.1 gives a mathematical description of the approach, introducing two different calibration parameterisation schemes. Implementations of these parameterisations have been tested using density data derived from TLEs and accelerometer data. These tests and their results are discussed in Sects. 6.2 and 6.3, respectively.

6.1 Estimation, Model Parameterisation and Data Preparation

The estimation of density calibration parameters is based on least-squares adjustment, using the Levenberg-Marquardt algorithm [1, 18, 20], in order to minimise the difference between observed density values derived from satellite drag and their equivalent values computed using the parameterised model. The Levenberg-Marquardt algorithm interpolates between the method of gradient descent and the inverse-Hessian method. The algorithm therefore requires the parameterised model implementation to return the Jacobian matrix, containing the partial derivatives of the density with respect to the calibration parameters, as well as the density itself.

6.1.1 Parameterisation Methods

The next paragraphs will present the two ways of parameterising the density model that have been implemented and tested.

Height-Dependent Model Density Scale Factors

If the unadjusted empirical model density is designated ρ_m, then an adjusted model can be calculated by multiplying this value with a scale factor f. This scale factor f can be a function in three-dimensional space, expressed in height h above the Earth's
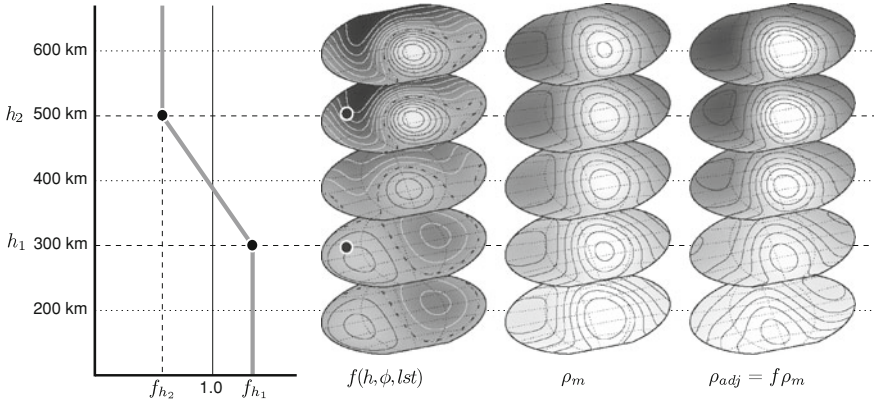
Fig. 6.1 Illustration of the height-dependent scale factor calibration scheme, using two height levels, at 300 and 500 km, with spherical harmonic expansions up to degree and order 1 and 2, at each height level, respectively

surface, latitude ϕ and local solar time $\lambda_\odot$.

$$\rho_{adj} = f(h, \phi, \lambda_\odot) \cdot \rho_m \tag{6.1}$$

The height variation of f is accomplished by a piecewise-linear interpolation between a set of N_{f_h} scale factors f_{h_i} at predefined reference heights h_i. Above and below the bottom- and top-most reference height (h_1 and h_N, respectively), the scale factor is kept constant.

$$f = \begin{cases} f_{h_1} & \text{if } h < h_1 \\ f_{h_i} + \frac{h - h_i}{h_{i+1} - h_i}(f_{h_{i+1}} - f_{h_i}) & \text{if } h_i \le h \le h_{i+1} \text{ for } i = 1, \dots, N_{f_h} - 1 \\ f_{h_N} & \text{if } h > h_N \end{cases}$$

$$\tag{6.2}$$

This equation is illustrated on the left-hand side of Fig. 6.1. If N_{f_h} equals one, there is no height-dependence in the correction, and the value of h_1 is irrelevant.

This scheme is an enhanced version of the method used by Yurasov et al. [23], in which a two-parameter correction corresponding to a scale factor with a slope in height was estimated. In their method, the density multiplication factor will reach unrealistic values at low and high altitudes. Therefore, such an adjustment could easily result in invalid density values above and below the perigee height range of the calibration data, even though the original unadjusted model could still return reasonable values there. Using Equation (6.2), this behaviour is avoided if the heights h_i are set within the height span of the calibration data.

The height levels offer the possibility of adding spatial resolution of the density correction function in the vertical direction. In order to accommodate spatial variations in the horizontal plane, each height-dependent scale factor f_{h_i} can be expanded in a set of spherical harmonics in latitude ϕ and local solar time $\lambda_\odot$. This allows

Table 6.1 Sets of spherical harmonic parameters used in the calibration tests, with their number of parameters N and names used in the discussion of these tests

Name	N	Spherical harmonic coefficients
Global	1	C_{00}
Offset	4	$C_{00}, C_{10}, C_{11}, S_{11}$
Zonal	4	$C_{00}, C_{10}, C_{20}, C_{30}$
Bulge	9	$C_{00}, C_{10}, C_{11}, S_{11}, C_{20}, C_{21}, S_{21}, C_{22}, S_{22}$

spatial features in horizontal planes around the Earth to be represented using the equation:

$$f_{h_i} = C_{00}^{h_i} + \sum_{n=1}^{N_{h_i}} \left[\sum_{m=0}^{n} C_{nm}^{h_i} P_{nm}(\sin\phi)\cos(m\lambda_\odot) + \sum_{m=1}^{n} S_{nm}^{h_i} P_{nm}(\sin\phi)\sin(m\lambda_\odot) \right]$$

(6.3)

The spherical harmonic coefficients $C_{nm}^{h_i}$ and $S_{nm}^{h_i}$ are the model parameters that are to be estimated. P_{nm} are the unnormalised associated Legendre functions of degree n and order m (e.g., [19]).

The low degree and order spherical harmonic coefficients have a clear physical meaning: C_{00} is a global scale factor. Its default value, for an unadjusted model, should be set to one. The higher degree and order coefficients are the ones that allow for variation around this mean in the horizontal plane, and their default value in an unadjusted model is zero. The three components of degree and order 1 allow for an offset with respect to the geocenter: C_{11} for the X-direction, S_{11} for Y and C_{10} for Z. Zonal coefficients, for which $m = 0$, can represent variations in latitude only: C_{20} introduces a flattening (ellipsoidal instead of spherical shape) and C_{30} a hemispherical asymmetry with respect to the equator (pear-shape).

Figure 6.1 illustrates the calibration scheme. In this specific example, a linear interpolation is used between two height levels, at 300 and 500 km altitude. At the lower level, a spherical harmonic expansion up to degree and order 1 is used, while at the higher level, where observation data is usually more abundant, the expansion is to degree and order 2. Using such low order spherical harmonics, the shape, position and amplitude of the diurnal density bulge can be easily modified. If data of sufficient coverage and accuracy are available, higher order spherical harmonics can be used to represent shorter wavelength density fluctuations.

A number of named sets of spherical harmonic parameters are listed in Table 6.1. These will be referred to later on in the chapter, when the effectiveness of the various sets of parameters is evaluated.

CIRA-72 temperature corrections

A second parameterisation method that was implemented involves the use of CIRA-72 temperature corrections. The scheme is based on the Dynamic Calibration Atmosphere (DCA) used in the HASDM project [3, 5].

Since the scheme requires a modification of the internals of the Jacchia model, we will refer to Sect. 2.4.1, which introduced the relation between several temperature parameters and the density output. As explained there, vertical density profiles at a given location are computed by integrating the hydrostatic and diffusion equations from the lower boundary conditions, based on a temperature profile with altitude that is a function of two pre-computed temperatures: T_x at 125 km, which is the temperature inflection point in the uncalibrated model, and T_∞ as the asymptotic maximum temperature at the top of the atmosphere.

In order to calibrate the model, corrections to these temperatures, ΔT_x and ΔT_∞, are both expanded in spherical harmonics, in a similar manner as the height-dependent scale factors of Equation (6.3). The set of spherical harmonic coefficients of these temperature corrections, $C_{nm}^{T_x}$, $S_{nm}^{T_x}$, $C_{nm}^{T_\infty}$ and $S_{nm}^{T_\infty}$, can then be estimated during the calibration procedure.

The resulting corrections ΔT_∞ and ΔT_x are added to the Equations (2.8) and (2.9), as follows:

$$T_\infty = T_c D + \Delta T_G + \Delta T_\infty \tag{6.4}$$

$$T_x = a + bT_\infty + c\exp(kT_\infty) + \Delta T_x \tag{6.5}$$

The complete temperature profile and local densities are then computed using the standard model formulations, as explained in Sect. 2.4.1. Figure 6.2 illustrates the influence of the calibration parameters on the temperature and density profiles of the CIRA-72 model. Note that the inflection point of the temperature profile moves to a lower altitude than 125 km, if the temperature T_x at that altitude is adjusted to a higher value, and vice versa, so that while we will keep the notation, technically the adjusted $T_x = T_{125\,km}$ is no longer the inflection point temperature.

Advantages and Disadvantages of the Two Calibration Methods

A theoretical disadvantage of the temperature correction method is that adjustments can only be made with two degrees of freedom in the vertical. In reality, due to sparseness of data, it has not been possible to test adjustments at more than two height levels in any case. A benefit of the temperature calibration scheme over the scale factor scheme is that the estimated parameters have a physical meaning in the model. The scale factors apply only to the total neutral density. Therefore, the temperature and composition results of the original model remain unaffected by the scale factors, and are no longer physically coherent with the scaled total density. In contrast, the corrected temperatures are at the basis of the CIRA-72 model, in which density and composition are computed based on physical principles from these temperatures.

Note however that the value of the corrected temperature T_x at 125 km should not be interpreted as representing the true temperature at this altitude. The uncorrected temperature parameter is closely tied to the lower boundary conditions of the CIRA-72 model, and its correction will be extrapolated from drag data above 200 km
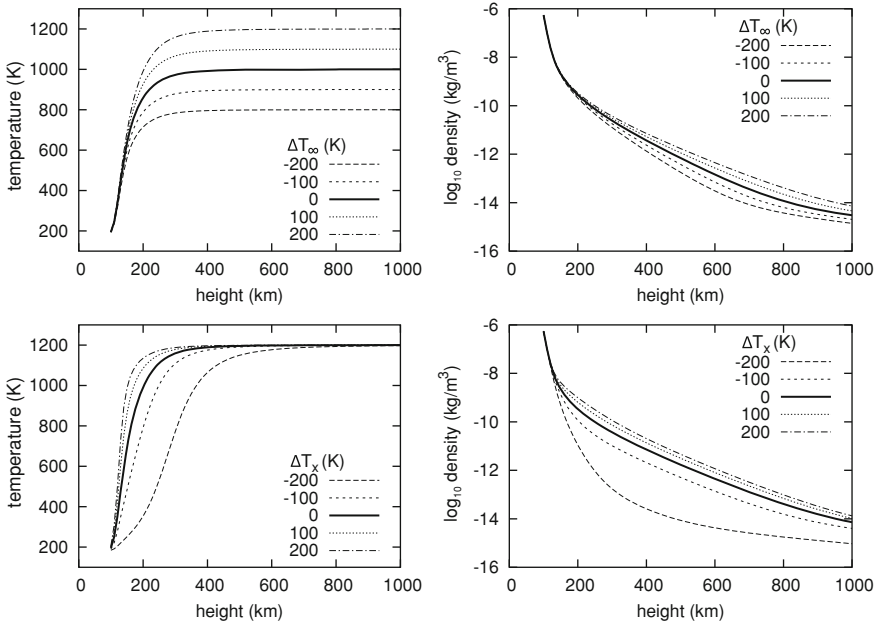
Fig. 6.2 Vertical profiles of density and temperature for various values of the calibration corrections for the inflection point and exospheric temperatures

through the temperature-density relations of the model, both of which will introduce considerable uncertainties.

6.1.2 Selection and Preparation of Density Data

The Levenberg-Marquardt algorithm is applied to minimise the difference between observed and equivalent model densities, divided by a weighting factor. The following sections will show how to arrive at these values, covering aspects of data selection, weighting and optimisation of computation time.

TLE-Derived Densities

The TLE-derived observed density data for calibration are obtained by evaluating Equation (4.6). Equation (4.7) can then be used to arrive at equivalent model values. Unfortunately, Equation (4.7) requires many calls to the density model per data point, because of the need to integrate the density over the satellite trajectory. In addition, the partial derivatives of the density with respect to the calibration parameters need to be integrated in a similar manner. This can be time consuming. Fortunately, the computations can be significantly speeded up by storing the intermediate calculation

Table 6.2 Characteristics of the selected validation objects, from left to right: NORAD identifier, SATCAT name, launch date, decay date, inclination, perigee and apogee height ranges (km) for the year 2000, and ballistic coefficient (kg/m^2)

ID	Name	Launch	Decay	i	h_p	h_a	B
22875	COSMOS 2265	1993/10/26	2003/08/11	82.9°	287–279	1239–1084	123.45
2389	OV3-3	1966/08/04	–	81.4°	350–347	2977–2899	57.14
4382	DFH-1	1970/04/24	–	68.4°	431–431	2143–2124	93.46
2611	OV1-10	1966/12/11	2002/11/30	93.4°	516–495	583–545	43.10
63	TIROS 2	1960/11/23	–	48.5°	535–523	592–574	69.93

results that are independent of the values of the calibration parameters, in computer memory or on disk.

Several aspects of the TLE data editing and selection process were already presented in Sect. 4.1.4. The next paragraph provides some further details on these aspects pertaining to the density model calibration tests.

An initial list of suitable calibration objects, used in the HASDM project, was kindly provided by Bruce Bowman (2005, personal communication). The densities were at first computed over the irregular integration time spans between pairs of TLE data. A minimum integration time span of half a day and a maximum of five days were used. Daily values for use in the daily density adjustment were calculated as an average of the collection of values valid for overlapping time spans. This introduces some smoothing in the process. Objects and time periods for which the effects of solar radiation pressure on the orbit are not negligible compared to drag, or where orbit differences between adjacent TLEs are excessively large, have been eliminated from further data processing. Ballistic coefficients were calculated under the assumption that the long-term average of the ratio of observed to modeled density ratios should equal one. The resulting ballistic coefficient values were generally found to be within 5% of values found in other research [2, 13, 17].

Depending on TLE data availability, either 48 or 49 objects, with a perigee range of approximately 180–550 km, were used to provide density estimates for each of the 366 days in 2000. These data were used as input to the 366 daily estimations of model parameters. The TLE-derived density data of five additional objects at different perigee altitudes were deliberately not used in the density calibration. Their data are used to independently assess the performance of the models. The validation satellites are listed in Table 6.2. In the presentation of results, they will be referred to by their name and approximate perigee height.

Accelerometer-Derived Densities

Accelerometer-derived densities as inputs to the calibration have been derived using Equation (4.12) for GRACE and Equation (4.28) for CHAMP. The accelerometer-derived data can be considered to consist of instantaneous measurements of the density, so they can simply be compared to the output of the parameterised density model evaluated at the same time and location.

Accelerometer data from CHAMP and GRACE are available at least every 10 s. In order to speed up the computation process when using accelerometer-derived data, it is beneficial to reduce the data amount, at the cost of some temporal and spatial resolution. This data reduction, either by averaging data into normal points, or by simply skipping data, can be tuned to match the temporal and spatial resolution that can be resolved by the chosen model parameterisation.

Data Weighting

Each data point derived from TLE s has been given a weight, which is inversely proportional to a fixed fraction (e.g. 10%) of the observed density. This ensures that high altitude and nighttime data gets sufficient relative weight in the adjustment. For these circumstances, the absolute density, and therefore also the density residual, is already small compared to data obtained at lower altitudes and during the daytime.

The data weighting for accelerometer data is based on an error analysis, similar to the one presented in Sect. 4.2.4. Details are available in Doornbos et al. [12]. Such more sophisticated data weighting schemes, based on an assessment of the relative accuracy of data obtained from different sources or conditions, is likely to be of benefit for the accuracy of the model calibrated using TLE data as well. This is foreseen to be part of future activities.

The combined adjustment of calibration parameters using both TLE- and accelerometer-derived data would also be of great interest, since the temporal and spatial resolutions of both methods could complement each other. This would, however, present an additional challenge regarding optimal relative data weighting and is left as a topic for future research.

6.2 Calibration Tests Using TLE Data

The evaluation of the density calibration procedure using TLE data consists of three parts. In the first part, we will look at time series of temperature adjustments. In the second part, the implementation of the estimation of spherical harmonic coefficients is checked, using maps of modeled density at 500 km altitude. In the third and final part, we will directly compare the performance of the different calibration schemes and number of spherical harmonic coefficients, using all five validation objects. In addition, the adjusted models are evaluated in the precise orbit determination of the ERS-2 satellite.

Time Series of Calibration Parameters and Density Ratios

A time series of estimated CIRA-72 temperature corrections is shown in Fig. 6.3, panel (b). Only the C_{00}, or global terms (see Table 6.1), of these corrections have been estimated. The variations can be compared with variations in $F_{10.7}$ and a_p in panel (a), which are the parameters used to determine the uncorrected temperatures in the CIRA-72 model. Large jumps in ΔT_{∞}, especially in July and August, coincide with

(a) Solar and geomagnetic activity indices

(b) Estimated CIRA-72 temperature corrections (degrees Kelvin)
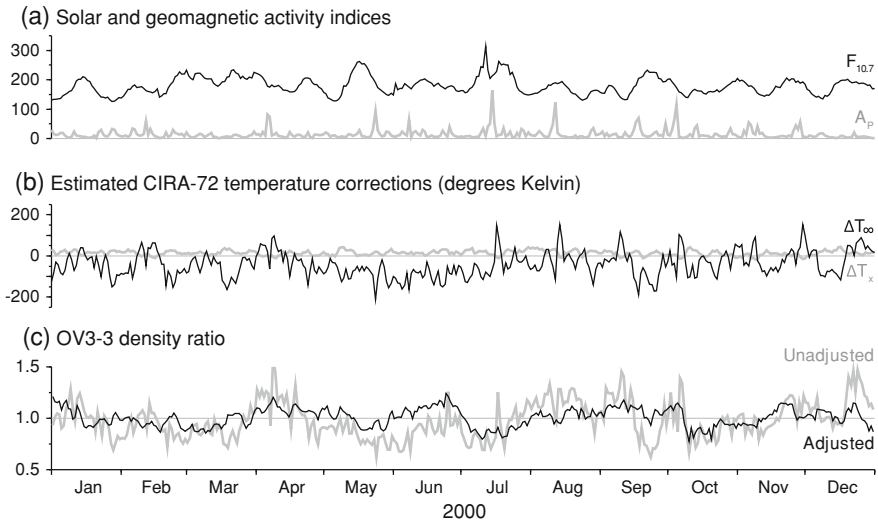
(c) OV3-3 density ratio

Fig. 6.3 Time-series of **a** $F_{10.7}$ solar flux and the A_p daily planetary geomagnetic index; **b** Estimated global CIRA-72 temperature corrections ΔT_x and ΔT_∞ in degrees Kelvin; and **c** Ratio of modeled over observed densities encountered by the OV3-3 satellite, for both the unadjusted CIRA-72 model, and the CIRA-72 model with the two global temperature corrections of panel (**b**) applied

large geomagnetic storms. Variations with a periodicity of approximately 27-days, similar to that in $F_{10.7}$ are also present. The temperature corrections therefore show variations that can be related to deficiencies in the representation of both geomagnetic and solar activity variations. Over the entire year, corrections to T_∞ are predominantly negative, while corrections to T_x are positive and much lower in amplitude. The two parameters often show an anti-correlated behaviour. The bottom panel of Fig. 6.3 makes a comparison of observed over modeled density ratios for the OV3-3 satellite, which is one of the validation objects listed in Table 6.2. The time series for the adjusted model clearly stays closer to the ideal ratio of 1 and shows less variability than for the unadjusted model. Indeed, the RMS relative density error was reduced from 19.8 to 9.2%. This indicates that the calibration was successful in lowering the model error.

Density Maps

Figure 6.4 shows the effect of adding higher degree and order spherical harmonic coefficients to the density calibration parameter estimation. The figure compares maps of modeled density at 500 km altitude for unadjusted and adjusted versions of CIRA-72 and MSIS. The spherical harmonic expansion in the adjusted model was up to degree and order 2 for T_∞ and f_{500}, respectively (corresponding to the keyword "bulge" in Table 6.1). Only a degree and order 0 (global) term was used for the low altitude parameters T_x and f_{300}. However, these parameters do not or hardly affect the density output at 500 km and above.
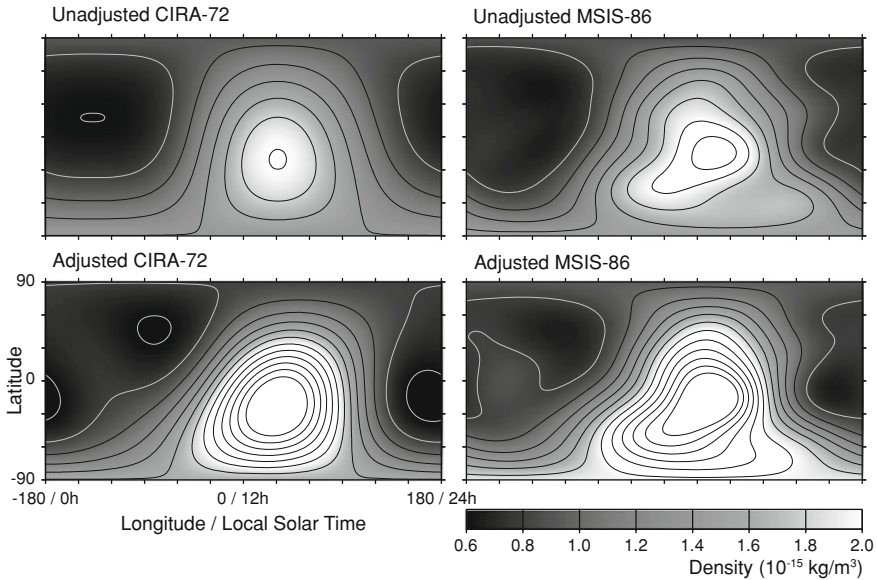
Fig. 6.4 Maps of modeled density at 500 km altitude at 12:00:00 UTC on January 16, 2000. The *top row* shows the output of traditional empirical models: CIRA-72 and MSIS-86. The *bottom row* shows the output of these models, calibrated using degree and order 2 spherical harmonics

A comparison of the top two frames of the figure shows that the MSIS-86 model shows a more detailed and asymmetrical picture of the diurnal bulge. This asymmetry had not yet been resolved from the rather sparse observation data at the time CIRA-72 was created. After adjustment however, both models show a much more similar picture. It appears that the asymmetrical shape of the diurnal bulge, which was already present in the MSIS model is indeed correct in this situation. It is also clear that both adjusted models show a much higher amplitude of the maximum day-time density than their unadjusted counterparts, even though the minimum density remains about the same for this epoch and altitude.

Comparison of Parameterisation Schemes

Table 6.3 contains a comparison of the performance of various density calibration schemes, in terms of the RMS of the relative density error (see Sect. 2.6.2), with respect to data from five validation satellites. The data from the validation satellites were not used in the adjustments of the models. The table contains three sections: several combinations of spherical harmonic expansions of scale factors have been applied to both the NRLMSIS-00 model and to CIRA-72, while the latter model has been tested with temperature corrections as well. The top row of each section contains the statistics of the unadjusted model.

Figure 6.5 contains a bar chart representation of the top and bottom rows of each section from this table. These relative RMS differences between TLE-derived and

Table 6.3 Comparison of adjusted empirical density models. N is the total number of estimated parameters (see Table 6.1). The other numbers show the RMS of the relative density errors, for the selected validation satellites over the year 2000. Lower values indicate better performance

| | | | COSMOS 2265 | OV3-3 | DFH-1 | OV1-10 | TIROS 2 |
			280 km	350 km	430 km	500 km	530 km
f_{300}	f_{500}	N	*NRLMSISE-00* adjusting scale factors f_{300} and f_{500}				
–	–	0	10.4	16.3	27.8	27.9	24.9
Global	–	1	9.4	8.3	16.3	11.3	11.3
Global	Global	2	5.4	8.1	14.9	7.9	9.1
Global	Zonal	5	5.4	7.8	14.2	9.6	9.2
Zonal	Zonal	8	11.6	8.4	13.8	10.0	9.3
Global	Bulge	10	5.3	7.3	9.9	11.0	9.1
f_{300}	f_{500}	N	*CIRA-72* adjusting scale factors f_{300} and f_{500}				
–	–	0	12.7	19.8	33.1	34.5	28.3
Global	–	1	11.5	9.6	18.7	12.9	11.1
Global	Global	2	7.7	9.8	17.1	7.6	11.5
Global	Zonal	5	7.5	9.4	16.1	7.7	10.8
Zonal	Zonal	8	16.0	9.7	15.7	8.4	10.7
Global	Bulge	10	7.4	8.3	10.3	8.9	9.8
ΔT_x	ΔT_∞	N	*CIRA-72* adjusting temperatures T_x and T_∞				
–	–	0	12.7	19.8	33.1	34.5	28.3
–	Global	1	9.0	9.5	16.7	8.9	9.3
Global	Global	2	8.1	9.2	15.6	7.9	11.5
Global	Zonal	5	7.1	8.7	14.3	7.7	10.7
Zonal	Zonal	8	9.3	8.9	13.9	8.6	10.4
Global	Bulge	10	6.9	8.7	11.5	8.0	9.5

modeled density values contain the contributions of the errors in the density model we are interested in, but they also contain any errors in the TLE to density processing. Based on analysis of our ballistic coefficient estimates and density adjustment results, these density observation errors are expected to be at a level of around 5% RMS at low altitude and high solar activity. At lower solar activity or higher altitudes, this error quickly increases (see Fig. 5.9). The limited temporal resolution of the TLE-derived data, as discussed in Sect. 4.1, should also be kept in mind when interpreting these numbers.

The RMS differences for the unadjusted models in Table 6.3 and Fig. 6.5 are between 10–13% at 280 km and 25–35% above 430 km, well above the estimate of the error in the validation observations. These numbers are therefore mainly an indication of the level of error in the unadjusted models. The more recent NRLMSIS-00 model performs better than CIRA-72 when both models are uncalibrated. The estimation of a single global calibration parameter per day, however, already improves the accuracy of both models by a great deal, to around 9–19% RMS. Adding the second global parameter, or increasing the number of spherical harmonic coefficients, brings the RMS error further down by more modest amounts. The global–bulge combination
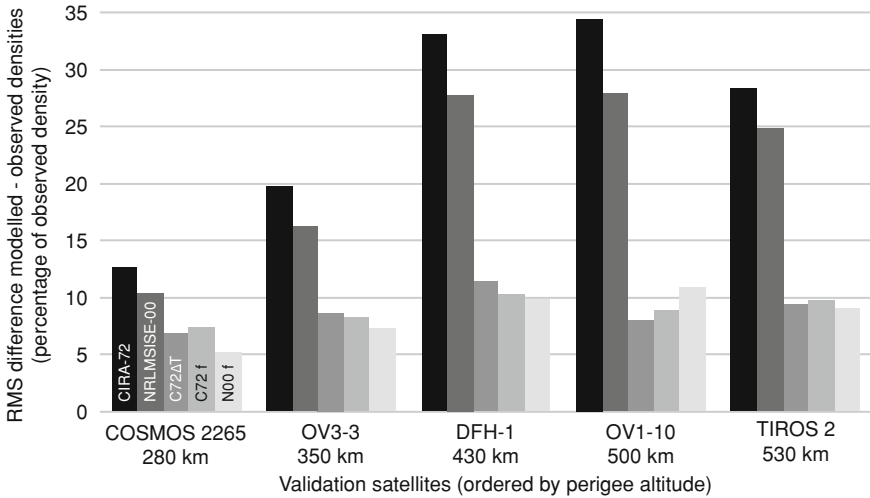
Fig. 6.5 *Bar graph* showing the increase in accuracy due to model calibration, using the data from Table 6.3. The calibrated models were adjusted using the global–bulge spherical harmonic combination of ten parameters representing temperature adjustments (ΔT) or scale factors (f)

is the most effective, bringing the RMS error down to 5–11% RMS, close to the estimated error level of the validation observations.

The zonal–zonal results, based on eight parameters, are generally worse than the global-zonal and global-bulge results, based on five and ten parameters respectively, This indicates that the introduction of additional spherical harmonics at the high altitude level (f_{500} or T_∞) is apparently more effective than at the lower level (f_{300} or T_x).

Since the two models are calibrated using the same data, it could be expected that the differences in remaining error between the calibrated models disappears as more parameters are estimated. This is not completely the case, indicating that the calibration parameters and data are not able to capture all density variations. The NRLMSIS-00 model still outperforms CIRA-72, even when ten parameters are adjusted for both models. It should be noted that the MSIS series of models contain much more detail and accuracy at a temporal and spatial resolution beyond that of the calibration data, compared to CIRA-72. This was already evident in Fig. 6.4.

The temperature correction method of CIRA-72, however, delivers slightly better results compared to the scale factor adjustment of the same model, at least when only a few coefficients are estimated. The addition of more spherical harmonic parameters reduces this difference.

Our results over the year 2000 can be compared to those presented for the HASDM test period during the first half of 2001 [4]. The HASDM results are based on 75–80 calibration objects for which Space Surveillance Network tracking data were processed. This has enabled a higher temporal resolution of corrections in HASDM than is available to us through the use of publicly available TLEs, which can be

Table 6.4 Comparison of gravity field models (top section) and density models (bottom), and their effect on the density error RMS for the year 2003, as well as the ERS-2 orbit error in the crossover RMS for cycle 81. The baseline models are indicated with asterisks (*)

Gravity model		Crossover RMS (cm)
JGM-3		9.04
DGM-E04		6.93
EIGEN-GRACE01S		6.59
EIGEN-CG03C*		6.47
EIGEN-GL04C		6.45
Density model	Relative density error RMS (%)	Crossover RMS (cm)
CIRA-72	30.8	7.03
DTM	21.8	6.73
MSIS	23.0	6.49
DTM-94	27.8	6.94
JB2006	23.3	7.06
NRLMSIS-00*	23.5	6.47
Calibrated NRLMSIS-00	17.1	6.37

considered a derived product of this tracking data with accuracy and temporal resolution limitations. The HASDM results presented by Bowman and Storz [4] show density errors at the 6–8% level across all heights from 200 to 800 km, which might be considered to be the best obtainable result using Space Surveillance data with this number of calibration objects.

Evaluation Using ERS-2 Precise Orbit Determination

An evaluation of the model calibration using the ERS-2 satellite, orbiting at about 800 km, was presented by Doornbos et al. [10]. A brief summary of the results will be given below. The evaluation was performed by implementing several empirical density models, including the "global/bulge" calibrated NRLMSIS model, in the ERS-2 orbit determination software. The orbit determination included the estimation of aerodynamic scale factors from the SLR tracking data, as explained in Sect. 4.3.

The evaluation of density error was made using SLR data from the year 2003, by estimating 6-hourly density scale factors, as explained in Sect. 4.3. Orbit determination arcs containing manoeuvres were eliminated. An evaluation of orbit error using radar altimeter crossover statistics was included as well, using only data from repeat cycle number 81, covering 35 days during January and February of 2003.

Table 6.4 shows that the calibrated NRLMSIS-00 model results in a lower RMS density error: 17.1% compared to 23.5% for the uncalibrated model, even when extrapolating to 800 km altitude. This improvement is considerable, since the TLE data used for calibration were observed over an altitude range of approximately 200–550 km. The temporal resolution of the STR-derived density scale factors, of about 6 h, is also higher than that of the TLE-derived data, which is another factor in explaining the somewhat higher residual RMS relative density error.

The 6-hourly scale factors are effective in removing the effects of long-term errors in the density models during the orbit determination process. It is likely that remaining orbit errors due to density error therefore represent mainly shorter-period errors, such as changes caused by variation in geomagnetic activity, as well as due to sampling of the diurnal density bulge over each orbit.

In order to study the effect of the choice of density model on the orbit quality of ERS-2, the radar altimeter data of this satellite were used as well. The altimeter instrument is intended to measure variations in sea and ice level. However, since the sea level usually does not change by more than several centimetres during periods of just a few days, this measurement can also be used as an independent form of tracking data, to evaluate the orbit error. Altimeter crossover differences are formed by differencing the two altimetric sea level measurements from ascending and descending passes which cross at the same location. Only crossovers over open ocean, with a maximum time difference of 5 days were used. The statistics are derived from 8531 crossover locations. These values contain the sum of the natural sea surface variability and a number of errors in the altimetric processing, of which the orbit accuracy is just one. These numbers can therefore be interpreted to evaluate the relative accuracy of different models used in orbit determination, if this is the only variable that is changed. Lower crossover RMS numbers indicate more accurate orbits. However, the crossover statistics do not give any absolute information on orbit accuracy, since the magnitudes of other contributions are not accurately known.

To provide context, the upper part of Table 6.4 shows a comparison between different gravity models, which have traditionally been the main driver for orbit accuracy improvement of altimeter satellites. The JGM-3 [22] and DGM-E04 [21] models were the state-of-the-art gravity models during the late 1990s, optimised for the TOPEX-Poseidon and ERS-1 missions, respectively. The three generations of EIGEN models are all based on GRACE gravity data. The EIGEN-CG03C model [16] was used in the density analysis. Subsequent generations of GRACEA-based gravity models only provide a modest improvement in ERS-2 orbit accuracy, which indicates that gravity model accuracy for this purpose is converging.

The differences in the altimeter crossover statistics for different density models, on the other hand, do not show such a clear development. The magnitude of the differences between the models indicate that in the GRACEA-era, improvements in density modelling have become more important than improvements in gravity modelling for altimetry satellites at this altitude. For the period under investigation, the calibrated NRLMSIS model shows slightly lower orbit errors in the crossovers than the uncalibrated model, which should result in improved sea and ice level data.

Conclusions of the Calibration Tests Using TLE Data

Calibration parameters that are estimated from TLE-derived density data can, to a large extent, compensate for the imperfections in traditional empirical density models at high solar activity.

An analysis of estimated temperature correction time series over the year 2000 shows that the exospheric temperature used in the CIRA-72 model is generally

too high and the temperature at 125 km too low over this period. Variations in the exospheric temperature correction reach amplitudes of several hundreds of degrees Kelvin, while the temperature at 125 km is adjusted by several tens of degrees Kelvin. A spherical harmonic expansion of the calibration parameters in local solar time and latitude can be used to improve the modelling of variations in the horizontal plane, such as the shape of the diurnal bulge.

Using only a single calibration factor per day already improves the RMS of daily integrated densities along the orbit of a validation object from a level of around 30% to below 16%. Further improvements, down to 5–10%, can be obtained by estimating more parameters at an additional height level and in a spherical harmonic expansion of latitude and local solar time.

The estimation of temperature corrections to the CIRA-72 model is preferred when temperature and constituent concentrations, consistent with total neutral density are required from the adjusted model. However, an alternative adjustment using scale factors applied to a more accurate base model, such as NRLMSISE-00, delivers slightly more accurate total densities.

The tests applied to the orbit determination of the ERS-2 satellite leads to the conclusion that the density model calibration using data from much lower altitudes (200–550 km) can still lead to significant density modelling improvements and modest orbit determination improvements at 800 km. This improved orbit determination accuracy has a direct positive effect on the accuracy of altimetric measurements of sea and ice level by European Earth observation satellites such as ERS-2, Envisat and CryoSat.

6.3 Calibration Tests Using Accelerometer Data

As mentioned in Sect. 2.5, the density calibration software was adapted to accept density data derived from accelerometer data, as part of an ESA-funded study [12]. CHAMP and GRACE data were used to test the software. The test results will be summarised in three parts. First, the time series of model calibration parameters will be analysed. Afterwards, the calibrated models are evaluated along the CHAMP satellite tracks during a short period surrounding the October 2003 geomagnetic storms. The last section presents calibration statistics for both the CHAMP and GRACE tracks, with both satellites used as either a calibration or validation satellite.

Time Series of Calibration Parameters

The simplest calibration is the application of a single density scale factor. This scale factor, estimated from the accelerometer data, is then applied to all the density model output. Such a calibration has been performed twice, one time using data from only the CHAMP satellite and one time using only the GRACE satellite data.

The resulting time series are visible in Fig. 6.6a. It is clear that the scale factor time series for both satellites are very similar, even though the satellites are at different
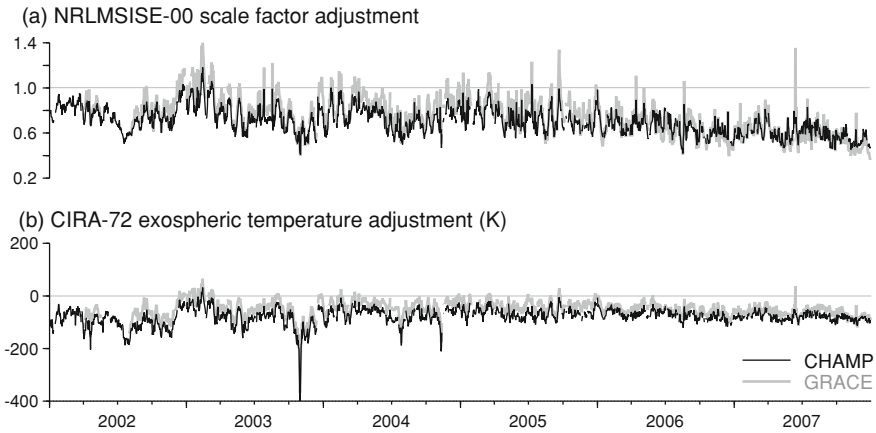
(a) NRLMSISE-00 scale factor adjustment

(b) CIRA-72 exospheric temperature adjustment (K)

Fig. 6.6 Estimated density model scale factors (C_{00}) for the NRLMSIS-00 model (**a**) and CIRA-72 exospheric temperature adjustments (**b**), comparing values derived solely from CHAMP data (*black line*) and GRACE data (*gray line*)

altitudes, and their local solar time coverage changes continually. This result is in line with the findings from the TLE-derived densities in the previous section. It indicates that there is a large global, time-varying error, which is largely due to the incorrect representation of solar EUV heating (through the $F_{10.7}$ proxy) in the empirical models.

The GRACEA-derived scale factors are a little higher than those from CHAMP, in the earlier years. However, for both satellites, the scale factors are generally below one, indicating that the NRLMSIS-00 model is scaled down, and was originally predicting densities that were too high. At high solar activity, there are large variations at approximately the solar rotation rate, also indicating that the model calibration compensates for inaccurate representation of the solar EUV radiation energy input in the thermosphere. At low solar activity, the scale factor becomes lower and lower. This might be related to the long-term thermospheric cooling trend, which in other analyses was shown to be strongest at solar minimum [14]. This unusual trend during the latest solar minimum was recently investigated in more detail by Emmert et al. [15].

A similar analysis can be made using the CIRA-72 temperature adjustments. Those results are visible in Fig. 6.6b. The discrepancy between the modelled and observed densities translate into an exospheric temperature that is about 100 K lower, than predicted by Jacchia's temperature Equations (2.7) and (2.8), based on $F_{10.7}$ and a_P. Note that this mean offset could be attributed to errors in the accelerometer processing, such as the satellite geometry model panel areas, energy accommodation coefficient, etc., as well as to errors in the density model. Note that an offset of the same magnitude is visible both at high solar activity (2002) and at low solar activity (2007). This is in contradiction to Emmert et al. [14, 15], who concluded that a larger temperature difference between data and models exists at solar minimum than
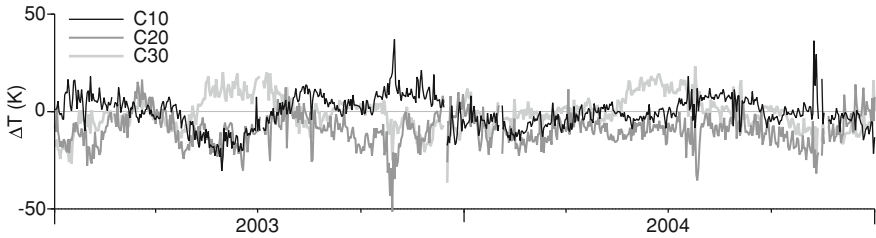
Fig. 6.7 Estimated zonal components, up to degree 3, of the CIRA-72 exospheric temperature adjustment, using the CHAMP and GRACE data combined

at solar maximum. However, their analysis was based on the NRLMSIS-00 model, not the CIRA-72 model, and they used TLE-derived drag data over a much longer time span, spanning multiple solar cycles. This difference between the two calibration results deserves further investigation.

A large variation of the temperature about this long-term mean can be observed. Its magnitude is about 50–100 K at high solar activity. This variability decreases at low solar activity. A very significant peak of up to −400 K is visible around the time of the October 2003 geomagnetic storms, suggesting that the CIRA-72 model was overestimating the temperature change due to the sudden large additional energy input at that time.

Figure 6.7 shows the exospheric temperature change time series for the CIRA-72 adjustment at the first, second and third order zonal components. Contrary to the previous plots, these values were estimated from the combined processing of CHAMP and GRACE data. The C_{10} and C_{30} components represent a North–South shift, the C_{20} component a shift from the poles to the equator or vice versa. The temperature corrections have an amplitude of about 20–30 K. Large peaks are visible around the October 2003 storm and other solar and geomagnetic activity related events. A larger part of the variation in these time series seems to be at longer wavelengths, such as at the annual and semi-annual periods. This indicates that the calibration adjusts for the inadequate (or non-existing) representation of the coupling between seasonal, local solar time and EUV radiation variation with the diurnal variation in the original CIRA-72 model.

Calibration Along CHAMP Tracks During the October 2003 Geomagnetic Storm

Figure 6.8 uses the geomagnetic storms at the end of October 2003 as an example to illustrate the model calibration. The top row of the figure shows the observed and uncalibrated model output, sampled along the CHAMP tracks, respectively. The bottom two rows show the output of several calibrated models. The titles for these panels list the base model used for calibration (NRLMSISE-00), the estimation interval (3 h), the spherical harmonic coefficients used as adjustment parameters (either a single scale factor, or an expansion in zonal harmonics up to degree 3),
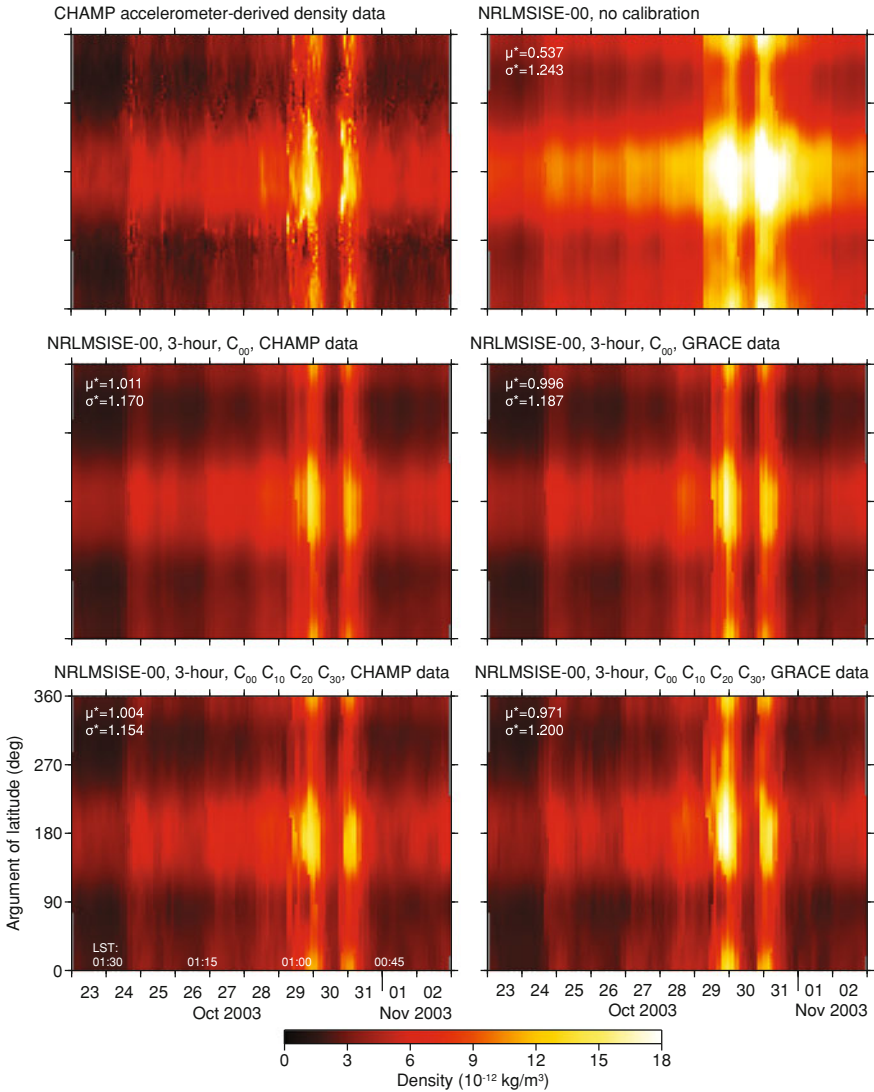
Fig. 6.8 Comparison of calibrated densities along CHAMP tracks, for the period around the October 2003 geomagnetic storms

and finally the source of density data used in the calibration (either CHAMP or GRACEA). The log-normal mean μ^* and standard deviation σ^*, included in each of the plots, are the statistics of the density ratios of the data (top-left) over each of the model values.

The figure shows that in general, as more calibration parameters are added, the model output more closely resembles the CHAMP observation data. While the origi-

Table 6.5 Statistics of calibrated density model output, evaluated along the CHAMP orbit, over the year 2003. N is the number of measurements used in computing the statistics. E is the RMS of the relative density errors (see Equation (2.15)), and μ^* and σ^* are the log-normal mean and standard deviations of the data/model density ratios

Base model	Parameters	Interval	N	E (%)	μ^*	σ^*
Uncalibrated						
NRLMSISE-00	–	–	3082486	47.4	0.759	1.251
JB-2008	–	–	3082486	35.0	0.803	1.193
HASDM calibrated model using SSN data						
HASDM	Custom	3-hourly	3082486	33.4	0.787	1.144
CHAMP data used for calibration						
NRLMSISE-00	C_{00}	Daily	3082474	14.1	1.029	1.152
NRLMSISE-00	C_{00}	3-hourly	3082472	13.0	1.024	1.138
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	Daily	3082478	13.4	1.026	1.145
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	3-hourly	3082472	11.2	1.018	1.119
GRACE-A data used for independent calibration						
NRLMSISE-00	C_{00}	Daily	2987441	17.8	0.949	1.169
NRLMSISE-00	C_{00}	3-hourly	2990681	17.7	0.941	1.160
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	Daily	2987441	19.1	0.938	1.176
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	3-hourly	2989601	19.1	0.925	1.164

nal, uncalibrated model in the top-right panel greatly overestimated the density during the storm, the calibration with a single scale factor brings the model and data to the same level, changing the log-normal mean from $\mu* = 0.537$ to 1.011 and 0.996 when calibrating with CHAMP or GRACE data, respectively. The addition of additional zonal harmonics in the calibration using CHAMP data further improves the fit between the model and data, as is to be expected. However, the adjustment using the independent GRACE data only results in a better fit when estimating the single scale factor. The addition of the three additional zonal harmonic coefficients changes the offset from the mean and the standard deviation slightly, from $\mu^* = 0.996$ to 0.971 and from $\sigma^* = 1.187$ to 1.200, as shown in the figure. This degradation is likely due to the different sampling of the diurnal bulge, due to the difference in local solar time of both satellite tracks at these dates. By calibrating using the zonal spherical harmonics, the different shape of the bulge as sensed by GRACEA at 01:00/13:00 LST, is imposed on the model evaluated for CHAMP at 04:00/16:00 LST.

Evaluation of the Calibration Along CHAMP and GRACE Tracks Over the Year 2003

For both CHAMP and GRACE-A, the various calibrated models have been evaluated using statistics over the entire year 2003 as well. The results are presented in Tables 6.5 and 6.6.

These tables contain four sections. The top two sections shows the statistics for the uncalibrated NRLMSIS-00 and JB2008 models, as well as for the independent HASDM data. These data, especially the uncalibrated NRLMSIS-00 statistics, serve

Table 6.6 Statistics of calibrated density model output, evaluated along the GRACE-A orbit, over the year 2003. The column headings are the same as in the previous table

Base model	Parameters	Interval	N	E (%)	μ^*	σ^*
Uncalibrated						
NRLMSISE-00	–	–	2313350	45.9	0.834	1.332
JB-2008	–	–	2313350	28.8	0.913	1.254
HASDM calibrated model using SSN data						
HASDM	Custom	3-hourly	2313350	22.4	0.909	1.186
GRACE-A data used for calibration						
NRLMSISE-00	C_{00}	Daily	2313339	18.8	1.052	1.211
NRLMSISE-00	C_{00}	3-hourly	2313339	17.4	1.044	1.193
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	Daily	2313339	16.5	1.040	1.185
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	3-hourly	2312521	13.5	1.027	1.149
CHAMP data used for independent calibration						
NRLMSISE-00	C_{00}	Daily	2300108	20.8	1.140	1.215
NRLMSISE-00	C_{00}	3-hourly	2299344	19.9	1.134	1.202
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	Daily	2307304	20.0	1.137	1.206
NRLMSISE-00	$C_{00}, C_{10}, C_{20}, C_{30}$	3-hourly	2299344	18.4	1.128	1.185

as a reference for the improvement that the calibration can bring. The evaluation of the HASDM model was included in order to be able to compare whether the calibration with accelerometer data from just one mission can compete with the calibration using many radar-tracked calibration objects.

The third section contains the results of the calibration using data from the same satellite as is used in the evaluation. This is, of course, not a fair evaluation of the benefits of the calibration, but these statistics are included in order to evaluate the limits of the parameterisation that are used. For example, the log-normal means of the data/model density ratios is within 2CHAMP and within 3exactly equal to one, because the data are weighted in the calibration, and not in the evaluation. The standard deviations decrease significantly with the introduction of additional calibration parameters, as expected. However, only in the case of the estimation of four zonal harmonic parameters over 3-hourly time intervals, do the calibrated models outperform HASDM. And this is when the same data are used for calibration and evaluation, while the HASDM calibration is done with independent data. This result serves as an excellent example of the validity of the HASDM data.

The fourth section in these two tables shows the results for the calibration with independent accelerometer data from the mission (CHAMP or GRACEA) that was not used in the evaluation. As expected, these are somewhat worse than in the case where the same data are used both for evaluation and calibration. Still, a very significant improvement over the uncalibrated NRLMSIS-00 model is seen.

In the case of the evaluation along the GRACEA track, the calibrated model using only CHAMP data results in only a slightly lower standard deviation (1.185) than the HASDM model (1.186), which is calibrated using many radar-tracked objects.

Nevertheless, this demonstrates the validity of the CHAMP data for improving thermospheric density modelling.

The GRACEA-calibrated model, evaluated along the CHAMP trajectory, on the other hand, does not exceed the performance of HASDM (1.164 vs. 1.144). This result is not unexpected, since it was shown in the previous chapters that density data from observations at altitudes above about 450 km, both from the GRACE accelerometers and from TLE s, are not as accurate as equivalent data obtained at altitudes comparable to that of CHAMP, due to the relatively low drag signal.

It would be interesting to evaluate the performance of CHAMP and GRACE calibrated models in the orbit determination of higher altitude satellites, such as ERS-2, Envisat and CryoSat. Perhaps the higher altitude GRACE data will bring in more value there. Another exciting possibility for future investigations is the calibration of density models using combined data from accelerometers and TLE data. As recommended by an earlier feasibility study for ESA/ESOC [6, 9], these two data types could complement each other. The accelerometer data provides high temporal and spatial resolution along a narrow satellite track, while TLE-derived densities could provide global spatial coverage, if enough calibration objects are available, albeit at a much lower temporal resolution. The cross-calibration of such data sets would be an interesting investigation by itself.

References

1. Bevington PR, Robinson DK (2003) Data reduction and error analysis for the physical sciences, 3rd edn. McGraw-Hill, New York
2. Bowman BR (2002) True satellite ballistic coefficient determination for HASDM. In: AIAA/AAS astrodynamics specialist conference and exhibit, 5–8 August 2002, Monterey, California AIAA 2002-4887
3. Bowman BR, Storz MF (2002) Time series analysis of HASDM thermospheric temperature and density corrections. In: AIAA/AAS astrodynamics specialist conference and exhibit, 5–8 August 2002, Monterey, California AIAA 2002-4890
4. Bowman BR, Storz MF et al (2003) High accuracy satellite drag model (HASDM) review. In: Lafontaine J (eds) Astrodynamics 2003 vol 116 of advances in the astronautical sciences. Univelt Inc, San Diego AAS 03-625
5. Casali SJ, Barker WN (2002) Dynamic calibration atmosphere (DCA) for the high accuracy satellite drag model (HASDM). In: AIAA/AAS astrodynamics specialist conference and exhibit, 5–8 August 2002, Monterey, California AIAA 2002-4888
6. Doornbos E (2004) Calibrated high accuracy satellite drag model—ESOC contract 16643/02/D/HK(SC) final report. Faculty of Aerospace Engineering Delft University of Technology, Delft, The Netherlands
7. Doornbos E (2006) NRTDM final report—Near Real-Time Density Model (NRTDM)—ESOC contract 18576/04/D/HK(SC). Delft Institute for Earth-Oriented Space Research
8. Doornbos E (2007) Thermosphere density model calibration in Space weather research towards applications in Europe. In: Lilensten J (ed) Astrophysics and space science library, vol 344. Springer, Dordrecht
9. Doornbos EN, Klinkrad H, Visser PNAM (2005) Atmospheric density calibration using satellite drag observations. Adv Space Res 36:515–521. doi:10.1016/j.asr.2005.02.009

10. Doornbos E, Klinkrad H, Scharroo R, Visser P (2007) Thermosphere density model calibration in the orbit determination and prediction of ERS-2 and Envisat. In: Lacoste H (ed) Envisat symposium Montreux, Switzerland, 23–27 April 2007, ESA SP-636
11. Doornbos E, Klinkrad H, Visser P (2008) Use of two-line element data for thermosphere neutral density model calibration. Adv Space Res 41(7):1115–1122. doi:10.1016/j.asr.2006.12.025
12. Doornbos E, Förster M, Fritsche B, van Helleputte T, van den IJssel J, Koppenwallner G, Lühr H, Rees D, Visser P (2009) ESTEC contract 21022/07/NL/HE air density models derived from multi-satellite drag observations—final report, DEOS / TU Delft scientific report 01/2009, TU Delft
13. Emmert JT, Meier RR, Picone JM, Lean JL, Christensen AB (2006) Thermospheric density 2002–2004: TIMED/GUVI dayside limb observations and satellite drag. J Geophys Res 111(A10S16). doi:10.1029/2005JA011495
14. Emmert JT, Picone JM, Meier RR (2008) Thermospheric global average density trends, 1967–2007, derived from orbits of 5000 near-earth objects. Geophys Res Lett 35(L05101). doi:10.1029/2007GL032809
15. Emmert JT, Lean JL, Picone JM (2010) Record-low thermospheric density during the 2008 solar minimum. Geophys Res Lett 37(L12102). doi:10.1029/2010GL043671
16. Förste C, Flechtner F, Schmidt R, Meyer U, Stubenvoll R, Barthelmes F, König R, Neumayer KH, Rothacher M, Reigber Ch, Biancale R, Bruinsma S, Lemoine J-M, Raimondo JC (2005) A new high resolution global gravity field model derived from combination of GRACE and CHAMP mission and altimetry/gravimetry surface gravity data. In: EGU general assembly 2005 Vienna, Austria, 24-29 April 2005
17. Lean JL, Picone JM, Emmert JT, Moore G (2006) Thermospheric densities derived from spacecraft orbits. Application to the Starshine satellites. J Geophys Res 111(A04301). doi:10.1029/2005JA011399
18. Moré JJ, Garbow BS, Hillstrom KE (1980) User guide for MINPACK-1 Argonne National Laboratory, Argonne, Illinois ANL-80-74
19. Olver FWJ, Lozier DW, Boisvert RF, Clark CW (eds) (2010) NIST handbook of mathematical functions. Cambridge University Press, Cambridge
20. Press WH, Teukolsky SA, Vetterling WT, Flannery BP (2007) Numerical recipes—the art of scientific computing, 3rd edn. Cambridge University Press, Cambridge
21. Scharroo R, Pieter V (1998) Precise orbit determination and gravity field improvement for the ERS satellites. J Geophys Res 103(C4):8113–8127
22. Tapley BD, Watkins MM, Ries JC, Davis GW, Eanes RJ, Poole SR, Rim HJ, Schutz BE, Shum CK, Nerem RS, Lerch FJ, Marshall JA, Klosko SM, Pavlis NK, Williamson RG (1996) The joint gravity model 3. J Geophys Res 101(B12):28,029–28,049. doi:10.1029/96JB01645
23. Yurasov VS, Nazarenko AI, Cefola PJ, Alfriend KT (2005) Density corrections for the NRLMSIS-00 atmosphere model. In: AAS/AIAA Space Flight mechanics conference. Copper Mountain, Colorado, January 23–27 2005, AAS 05-168

Chapter 7
Conclusions and Recommendations

In this final Chapter, the conclusions of the previous Chapters will be summarised, and recommendations will be given for improvements to the processing of data from current missions, as well as for the design of possible future missions.

7.1 Conclusions

The algorithms and investigations presented in this thesis have demonstrated that the accelerometer instruments on CHAMP and GRACE, as well as the TLE-data from many space debris objects, provide valuable observations of the thermosphere in various forms. These processing techniques and the resulting data have been used to improve our understanding of upper atmospheric processes and phenomena, and will continue to do so in the years to come.

The newly developed iterative algorithm, presented in Sect. 4.2.2, has been applied to derive information on the thermospheric wind speed, with a reduced error budget. The assessment of density and wind errors due to various errors in the input models and data, has shed light on the nature and extent of these errors. In the long-term, this is perhaps the most important contribution of this work, because it is valuable both for users of the data interested in geophysical interpretation of variations in density and wind, and for investigators wishing to further improve the data processing results.

Due to the largely global nature of observed density variations, combined with large systematic errors in existing empirical density models, assimilation of contemporaneous density data to calibrate such models will result in an immediate improvement of their accuracy. This can lead to improvements in applications such as precise orbit determination of Earth observation satellites.

7.1.1 Accelerometer-Derived Density and Wind Data

In contrast to previously published direct algorithms for deriving density and wind from accelerometer data, the iterative algorithm described in this thesis can be applied

in situations without a close alignment of the accelerometer axes with the orbit's along-track and cross-track directions. In fact, it can be used for arbitrary orientations of the accelerometer axes in space. The analysis using simulated CHAMP data in Sect. 4.2.4 shows that errors due to assumptions on the orientation are significantly reduced when the new algorithm is applied. However, this analysis also shows that errors in the instrument calibration and input models that are common in the use of both algorithm types lead to more significant density and wind errors than the errors in the algorithms.

The density data suffer mainly from systematic offsets (scaled densities, with respect to the truth), due to inaccurate information of the spacecraft geometry and gas-surface interaction. These issues are not so important for the wind data. Instead the accelerometer-derived winds are extremely sensitive to small errors in the calibration and radiation pressure modelling, especially when the density and aerodynamic acceleration are relatively small as well. Such errors distort the true patterns in the crosswind output in a complex manner, making the validation and interpretation of these data difficult.

Users of the current CHAMP and GRACE data (and possibly GOCE and Swarm data in the future as well), should be aware of the level and nature of such errors in density and wind. The investigation of these algorithms and their related error sources has led to recommendations for improvements of possible future dedicated accelerometer-carrying space missions for studies of the thermosphere, presented in Sect. 7.2.

Despite these systematic errors and uncertainties, the density and wind observations resulting from this study have proven useful for thermospheric modelling, and will likely continue to be useful for a long time.

7.1.2 Two-Line Element-Derived Density Data

Density data obtained from Two-Line Elements, according to the algorithm by Picone et al. [6] have proven to be a reliable and valuable source, that compares well with accelerometer-derived data. The disadvantage that the temporal resolution of the TLE-derived density data is very low, at three days or more, is offset by the advantage of the availability for a very large number of space objects. An important condition for this is the application of a robust data editing and ballistic coefficient determination scheme, such as the one implemented by Emmert [1].

The comparison between TLE-derived and accelerometer-derived density data for CHAMP and GRACE, presented in Fig. 5.9, has shown that the accuracy of the data degrades rapidly under low drag conditions, such as at higher altitudes and lower levels of solar activity. This non-uniform accuracy has not been sufficiently addressed in earlier publications on the topic, but it is important to keep in mind for users of the data.

7.1.3 Density Biases and Trends

When both types of air density measurements are compared with models based on similar measurements from the 1960s and 1970s, a large decrease in density, of the order of 20%, can be observed. The difference between these recent measurements and historical models is larger than the contribution of systematic errors in the satellite area and aerodynamic modelling, described above. Therefore, there must have been a decrease in the true density, which is the result of a downward trend in the thermospheric temperature. This thermospheric cooling trend, which has been studied in more detail by other researchers (e.g., [2]), is consistent with the effect of raised concentrations of greenhouse gases in the lower atmosphere [4].

Comparisons of measurements and models made at solar minimum, in 2008 and the surrounding period, show evidence of additional cooling, which is related to the extremely low solar activity at the time [3]. Such a deviation from the usual 11-year cycle in density and solar activity has not been observed since the beginning of the space age.

Future measurements, processed using the algorithms presented in this thesis, will be able to tell us whether the observed thermospheric cooling trend is continuing at the same rate, and whether the unexpected deviation of the pattern in the interaction between the Sun and the thermosphere around 2008 was a one-time event, or the start of a prolonged period of low solar activity. The latter possibility would have profound consequences on satellite lifetimes and the evolution of the population of space debris objects in low Earth orbits.

7.2 Recommendations for Future Missions

The instruments, tracking techniques and satellite missions employed in this thesis were, in most cases, not designed with the objective of studying the thermosphere in mind. Nevertheless, it is indisputable that the accelerometer instruments on the CHAMP and GRACE missions, and the various types of tracking data available for many other missions and space debris objects, have proven to be very valuable sources of information for the thermosphere modelling community.

The experience gained in the study of the CHAMP data in particular, and the development and analysis of the iterative algorithm presented in this thesis, have led to several recommendations for the development of possible future thermosphere missions, with the aim to reduce density and wind errors.

First of all, a compact and simple design of the satellite external shape, without protruding antennae, camera baffles, booms, etc., will reduce the uncertainty in geometrical and aerodynamic satellite modelling, which will result in a more reliable estimate of absolute density values. The availability of additional instruments on accelerometer missions, which could make contemporaneous in-situ measurements of the atmospheric temperature, molecular mass, in-track wind, and other

parameters important for gas-surface interaction, would increase the accuracy of the aerodynamic calculations required for the accelerometer processing, as discussed in Sect. 3.4.3. At the same time, such instruments would provide valuable data for atmospheric and aerodynamic modelling in general, which could in turn aid in a more accurate reprocessing of historical accelerometer datasets.

A large area-to-mass ratio of the satellite will increase the acceleration signal, which is especially beneficial for wind derivation. Flying at high solar activity and low altitude will help in that respect as well, but that will put limits on a mission's sampling characteristics. A high eccentricity orbit might aid in calibration and the separate fine-tuning of radiation pressure and aerodynamic satellite models, but again at the cost of the beneficial atmospheric sampling characteristics of circular orbits.

Finally, the example of CHAMP's sideways-flying periods, presented in Sect. 5.2.1, shows that a more versatile or more loosely defined attitude control of an irregularly shaped satellite will provide data that can be used to identify and possibly reduce density and crosswind errors. If the attitude control can be designed such that each of the three accelerometer axes can spend a sufficient amount of time, in turn, in the satellite flight direction, this could be beneficial for the instrument calibration using orbit tracking data, and reduce the crosswind error. The data processing of such a mission is possible using the iterative algorithm presented in this thesis.

7.3 Outlook for Further Research

Even though the CHAMP satellite has ended its 10 years in orbit with a fiery re-entry, the quality of its density and wind data, as well as the utilisation of these data, can still be improved. This is even more true for GRACE, which, at its higher altitude, requires an even more precise modelling and calibration, in order to reduce density and wind errors. Meanwhile, new data from the GRACE and GOCE missions, as well as fresh TLEs, keep arriving at a steady pace. The planned Swarm and GRACE follow-on missions bring with them the promise that satellite accelerometer data will remain available in abundance, and hopefully without significant interruptions, throughout the next decade. This Section lists several possible follow-on investigations, that might aid in getting the most out of these data.

We have seen that the largest errors in the density data originate in the modelling of the spacecraft geometry and gas-surface interaction. Since the creation of the current geometry models already required a significant investment in man-hours, making use of detailed satellite drawings, a further improvement of their fidelity likely requires an even larger investment. For example, for future missions it could be feasible to convert detailed CAD models of the satellites directly into force models, or to use 3D laser scanning techniques on the finished satellite in the clean-room.

As discussed in the previous Section, in order to really solve the problem of inaccuracies in modelling of the gas-surface interaction, even bigger investments are required, in the form of a newly designed satellite mission for investigations of the thermosphere and satellite aerodynamics, carrying additional instrumentation

beyond an accelerometer. But smaller improvements can already be made right now, for example by adjusting the temperature and composition output of the NRLMSISE- 00 model, to compensate for the fact that this model overestimates the energy input into the thermosphere, especially at low solar activity. The resulting density bias that was examined in Chap. 5 is most likely paired with similar temperature and composition biases, skewing the current aerodynamic calculations. The $F_{10.7}$ input to the NRLMSISE- 00 model can be adjusted, so that the density bias is removed, resulting in a temperature and composition output that is consistent with this level of density. Such adjusted temperature and composition values can then be applied in the density and wind algorithm, to see if a reduction in the standard deviations of density ratios occur.

Another option would be to utilise density data from satellites which are less sensitive to the settings of the gas-surface interaction parameters, such as the ANDE spheres [5]. If there are periods where such satellites sample the density at a similar altitude and local solar time as CHAMP or GRACE, the gas-surface interaction parameters can be adjusted so that the consistency between their density outputs is optimised.

In the derivation of crosswind speeds from accelerometer data, errors in the cross-track accelerometer calibration and solar radiation pressure modelling were identified as the largest error sources. Some improvements to both these problems are possible for the current missions as well. These improvements can be tested by looking at the consistency of the derived winds with empirical and physical wind model output, and by looking at the consistency between winds before, during and after special attitude manoeuvres, such as the one in November 2002, described for CHAMP.

The current calibration strategy is based on the estimation of daily biases, while keeping the scale factor fixed at an advertised value. In reality, the advertised value of the scale factor might not be the correct or optimal one. An optimised scale factor, or a time series of such scale factors, can possibly be determined by minimising the standard deviation of measured minus modeled accelerations.

A further proposed improvement concerns the accelerometer data preprocessing. The accelerometer data products used in this thesis have already been preprocessed before release for scientific use. The preprocessing step generally includes a reduction of the data rate. In the CHAMP data processing, acceleration spikes, due to thruster activations and other sources, have been removed from the high-rate data, before conversion to the low-rate data. This is a good situation for density and wind processing, since such spikes are not part of the aerodynamic acceleration signal. It is not a good situation for the calibration of the X-axis data using GPS tracking data, since these acceleration spikes likely have a real effect on the orbit, that is also present in the GPS data. The removal of the spikes should therefore result in an error in the calibration. For GRACE the situation is more or less reversed. The spikes are not removed before the data rate conversion in the preprocessing. However, the preprocessing also includes the application of a digital low-pass filter, which causes the thruster spikes to get spread out over a longer time duration, making them much more difficult to remove. In the ideal case, which should apply for future missions as well as the currently available ones, the lower-level high-rate data should be made

available to scientists, so that the appropriate digital filtering can be applied for calibration, while the spikes can be removed for further processing into density and wind information.

Finally, the modelling of solar radiation pressure can be improved in several ways as well. There are currently no efficient eclipse models available that include the effects of the Earth's flattening as well as atmospheric absorption and refraction. Such a model, employing ray-tracing techniques, would directly reduce wind errors near equator crossings. Its indirect effect on the accuracy of wind estimates, through the improvement of the acceleration calibration using models, might be even more important. A further option for improving the radiation pressure modelling would be to adjust the optical properties of the satellite surfaces, so that the resulting modelled acceleration better fits with observed accelerations. This will only lead to accurate results if data periods and acceleration components are used, for which the aerodynamic and Earth radiation pressure accelerations are very small compared to the solar radiation pressure acceleration.

References

1. Emmert JT (2009) A long-term data set of globally averaged thermospheric total mass density. J Geophys Res 114(A06315). doi:10.1029/2009JA014102
2. Emmert JT, Picone JM, Meier RR (2008) Thermospheric global average density trends, 1967–2007, derived from orbits of 5000 near-earth objects. Geophys Res Lett 35(L05101). doi:10.1029/2007GL032809
3. Emmert JT, Lean JL, Picone JM (2010) Record-low thermospheric density during the 2008 solar minimum. Geophys Res Lett 37(L12102). doi:10.1029/2010GL043671
4. Keating GM, Tolson RH, Bradford MS (2000) Evidence of long term global decline in the earth's thermospheric densities apparently related to anthropogenic effects. Geophys Res Lett 27(10):1523–1526
5. Nicholas A, Finne T, Galysh I, Davis M, Healy L (2010) The atmospheric neutral density experiment (ANDE), in 2009 NRL Review, Naval Research Laboratory, pp 138–142
6. Picone JM, Emmert JT, Lean J (2005) Thermospheric densities derived from spacecraft orbits - I. Accurate processing of two-line element sets. J Geophys Res 110(A03301). doi:10.1029/2004JA010585