Perturbed Motion

David A. Vallado
Analytical Graphics Inc.

Contents

1.1 Basic definitions

Perturbations are deviations from a normal, idealized, or undisturbed motion. The actual motion will vary from an ideal undisturbed path (two-body) due to perturbations caused by other bodies (such as the Sun and Moon) and additional forces not considered in Keplerian motion (such as a non-spherical central body and drag).

It is important to know about *gradients*, *accelerations* (*specific forces*), and *functions*. A ***gradient*** is really a directional derivative which gives the rate of change of a *scalar function* in a particular direction (Kreyszig, [7]). It's a vector quantity and the ***del operator***,∇, designates the gradient process. The gradient gives an acceleration if the scalar function is a potential function related to a specific potential energy, such as the potential function of a central body's gravity field. I distinguish a potential function as

the negative of the potentil energy. Two conventions are "standard" in this area because many schools of thought have evolved over the last few decades. Brouwer and Clemence [3], Battin [2], Long et al. [10], and others express one of the two main approaches, in which the acceleration is the negative gradient of the potential function. This implies that positive work is done as the potential decreases. The other approach, used mainly by the geophysical community, holds an acceleration to be the positive gradient of the potential function [Lambeck [9], Kaula [6], Moritz and Mueller [13], Kaplan [5], Roy [14], and others]. Of course, both methods use potential functions that differ only by a minus sign; therefore, the results are identical! We'll follow the second method and place the sign change between the potential energy and the potential function. I'll also refer to the potential function instead of simply the *potential*, to avoid confusion with *potential energy*.

The distinction between a specific force (often used interchangeably with acceleration) and a potential is important because analysis of perturbations typically uses both concepts. It's common to analyze perturbations using a *disturbing function* and a *disturbing force*. The *disturbing force* simply expresses (in some coordinate system) the specific force (acceleration) that is perturbing the satellite's orbit. Non-conservative forces, such as the perturbing effects of drag and solar-radiation pressure, are usually modeled as a specific force ***Disturbing functions*** are simply the difference between perturbed and unperturbed potential functions. They model conservative forces that perturb the orbit, such as the central body's non-sphericity and third-body attractions.

A ***potential function*** is one way to mathematically characterize a conservative force, such as the gravitational potentialof a *spherical* central body ($U_{2\text{-}body} = \mu/r$). Some people distinguish a disturbing function from a disturbing potential by a minus sign. As mentioned earlier, considering the two to be equal is just as correct, as long as we maintain the correct sign convention. The potential function for an *aspherical* central body, U(sometimes referred to as the *anomalous potential*) includes the spherical potential ($U_{2\text{-}body}$) as the first term. The term *geopotential* is often used for this aspherical potential when the central body is the Earth.

Because we wish to examine the effect of perturbations on the orbital elements, we must characterize how they vary over time. Perturbations on orbital motion result in *secular* and *periodic* changes.

Secular changes in a particular element vary linearly over time, or in some cases, proportionally to some power of time, such as a quadratic. Secular terms grow with time, and errors in secular terms produce unbounded error growth. Secular terms are the primary contributor to the degradation of analytical theories over long time intervals. Although the dominant perturbing force for the Earth, J_2, results in all three types of effects, we can do a first-order approximation and approximate the main variations. We can also develop some higher-order solutions. ***Periodic*** changes are either *short-* or *long*-periodic, depending on the length of time required for an effect to repeat. Because so many definitions exist in the literature and in practice, I'll define each type.

Short-periodic effects typically repeat on the order of the satellite's period or less. ***Long-periodic*** effects have cycles considerably longer than one orbital period—typically one or two orders of magnitude longer. These long-periodic effects are often seen in the

motion of the node and perigee and can last from a few weeks to a month or more. This means a short-periodic effect for a satellite at an altitude of 400 km could vary with periods up to about 100 minutes, whereas a short-periodic effect for a geosynchronous satellite would be up to about 24 hours. Also, short-periodic variations occur when a fast variable (true anomaly, for instance) is present in the contributing perturbational effect.

We also distinguish certain orbital elements as either *fast* or *slow* variables, depending on their relative rate of change. ***Fast variables*** change a lot during one orbital revolution, even in the absence of perturbations. Examples are the mean, true, and eccentric anomalies, which all change 360°, or the cartesian coordinates, which also change dramatically in a single revolution. ***Slow variables*** (semimajor axis, eccentricity, inclination, node, argument of perigee) change very little during one orbital revolution. Perturbations cause these changes. Without perturbations, all the slow elements would remain constant. Fast variables would continue to change.

We can describe the perturbed motion of a satellite by an ordered series of position and velocity vectors. Consequently, at each point in time, we can use these vectors to find the orbital elements using two-body techniques. The corresponding position and velocity vectors define these ***osculating elements*** at any instant in time. "Osculate" comes from a Latin word meaning "to kiss." Thus, the osculating orbit kisses the trajectory at the prescribed instant. We define an ***osculating ellipse*** as the two-body orbit the satellite would follow if the perturbing forces were suddenly removed at that instant. Therefore, *each* point on the trajectory has a corresponding set of osculating elements. Osculating elements are the true time-varying orbital elements, and they include all periodic (long- and short-periodic) and secular effects. They represent the high-precision trajectory and are useful for highly accurate simulations, including real-time pointing and tracking operations.

In contrast, ***mean elements*** are "averaged" over some selected time (or an appropriate angle such as true anomaly), so they are relatively smoothly varying and do not chase the short-periodic variations. Notice that mean elements depend on some unspecified averaging interval of the time; the true, eccentric, or mean anomaly; or the longitude. Because there are many kinds of mean elements, it is important to understand how they are defined and used. Mean elements are most useful for long-range mission planning because they approximate the satellite's long-term behavior.

1.2 Forces

The accuracy of orbit determination largely depends on modeling of all physical forces affecting the motion of the Earth satellite or spacecraft in its orbital path through space. By far the largest effect is due to gravitation, usually followed by atmospheric drag, third body perturbations, solar radiation pressure effects and a suite of smaller effect such as tides, and several others. Vallado [16] shows the relative effect of various forces on several satellites at two different satellite altitudes. Figure 1.1 shows these quantitative effects of all physical forces in terms of positional differences for a 500 km altitude, 97.6° inclined satellite. Note that most of the effects like tides, third body forces and relativity are very small, but need to be taken into account when high precision is of

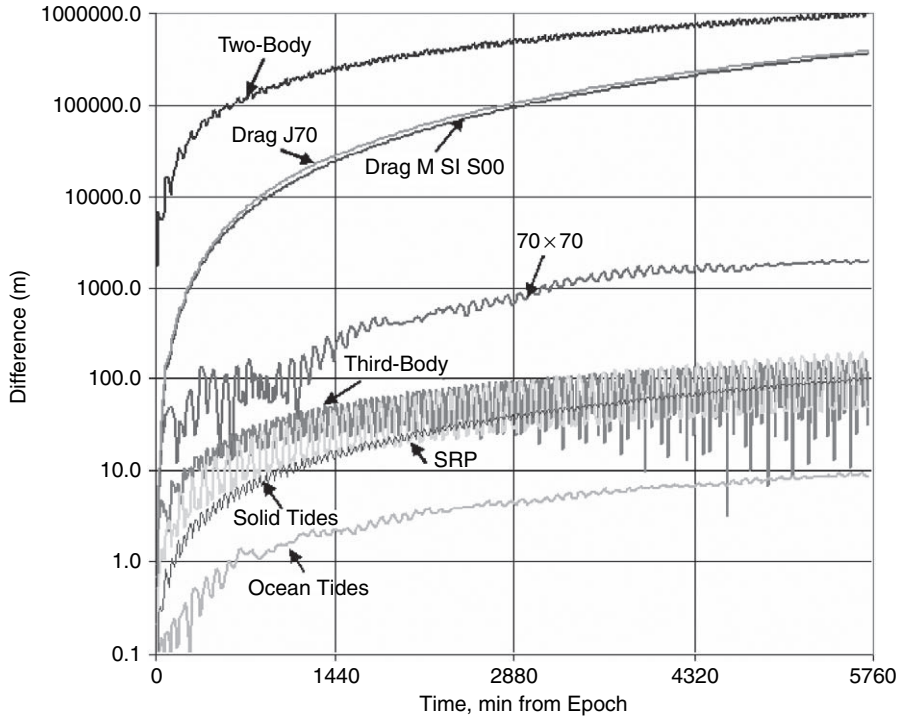

Fig. 1.1. Force Model Comparisons: This figure shows the positional difference over time (four days) when using various force models on the same initial state.

importance. The satellite parameters were chosen to illustrate force model effects. The coefficient of drag $c_{\mathrm{D}} = 2.2$, coefficient of solar radiation pressure $c_{\mathrm{R}} = 1.2$, and area to mass ratio $A/m = 0.04\,\mathrm{m}^2/\mathrm{kg}$. The simulation time, January 4, 2003, was chosen as the epoch to propagate as this was a moderate period of solar activity (solar flux $F_{10.7} \sim 140$).

1.3 Gravity

The general equation for the gravitational attraction uses a spherical harmonic potential equation in an Earth-centered, Earth-fixed reference frame of the form. The fundamental expression for Earth's gravitational potential acting on a satellite is usually given in the familiar form of Earth's geopotential with the origin at Earth's center of mass:

$$V = \frac{\mu}{r}\left[1 + \sum_{n=2}^{\infty}\sum_{m=0}^{n}\left(\frac{R_{\oplus}}{r}\right)^{n} P_{nm}(\sin\phi_{gcsat})(C_{nm}\cos m\lambda_{sat} + S_{nm}\sin m\lambda_{sat})\right] \tag{1.1}$$

Where $\mu =$ gravitational parameter, r is the satellite radius magnitude, ϕ_{gcsat} and λ_{sat} are the geographic coordinates of the satellite, $R_{\oplus}$ is the Earth radius, and C_{nm} and S_{nm} are the gravitational coefficients. Notice the presence of Legendre polynomials.

$$\begin{aligned} P_{nm}\left(\sin\phi_{gcsat}\right) &= \left(\cos\phi_{gcsat}\right)^{m}\frac{\mathrm{d}^{m}}{\mathrm{d}^{m}(\sin\phi_{gcsat})}P_{n}\left(\sin\phi_{gcsat}\right) \\ P_{n}\left(\sin\phi_{gcsat}\right) &= \frac{1}{2^{n}\ n!}\frac{\mathrm{d}^{m}}{\mathrm{d}^{m}(\sin\phi_{gcsat})}\left(\sin^{2}\phi_{gcsat}-1\right)^{n} \end{aligned} \tag{1.2}$$

For computational purposes, this expression is often used in the *normalized* form. This results from replacing P_{nm}, C_{nm}, and S_{nm} with $\bar{P}_{nm}$, C_{nm}, and $\bar{S}_{nm}$ where

$$\bar{P}_{nm} = \left[\frac{(2n+1)k(n-m)!}{(n+m)!}\right]^{\frac{1}{2}} P_{nm},$$

and

$$\begin{Bmatrix}\bar{C}_{nm}\\ \bar{S}_{nm}\end{Bmatrix} = \left[\frac{(n+m)!}{(2n+1)k(n-m)!}\right]^{\frac{1}{2}} \begin{Bmatrix}C_{nm}\\ S_{nm}\end{Bmatrix}, \tag{1.3}$$

with $k = 1$ for $m = 0$, and $k = 2$ for $m \neq 0$.

A Legendre function (polynomial or associated function) is referred to as a zonal harmonic when $m = 0$, sectorial harmonic when $m = n$, and tesseral harmonic when $m \neq n$.

When normalized coefficients are used, they must be used with the corresponding normalized associated Legendre function:

$$P_{nm} = \frac{\bar{P}_{nm}}{\Pi_{nm}} \tag{1.4}$$

such that $\bar{C}_{nm}\bar{P}_{nm} = C_{nm}P_{nm}$ and $\bar{S}_{nm}\bar{P}_{nm} = S_{nm}P_{nm}$ and the standard model is preserved. Computer software programs generally all use double precision values when converting these coefficients.

1.3.1 Earth Gravitational Models

The first attempts to standardize models of the Earth's gravitational field and the shape of the Earth were begun in 1961. A series of gravitational constants in the form of low degree and order spherical harmonic coefficients were published based on Sputnik, Vanguard, Explorer, and Transit satellite tracking data by special investigators within their respective sponsoring organizations. The first gravity models differed greatly primarily due to observational and computational limitations. As satellite tracking has become more commonplace and computing power has increased, there are still several gravitational models, but their differences are minimal for most applications. Currently there are several prevailing gravitational models being used within the scientific community for a variety of purposes. These models were determined from a wide range of measurement types,

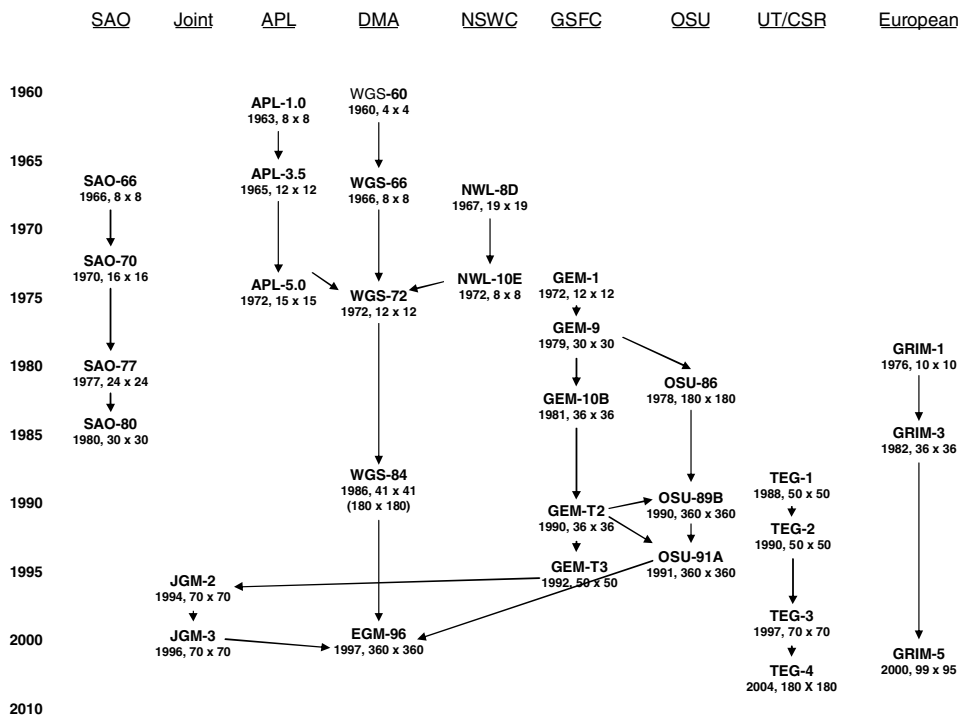

Fig. 1.2. Gravitational models: The *Joint Gravity Models* (JGM) come from Goddard Space Flight Center (GSFC), Ohio State University (OSU), University of Texas at Austin (UT), and the European communities. The *Earth Gravity Model* (EGM) combines the JGM work with Defense Mapping Agencies efforts. The *Goddard Earth Model*s (GEM) were produced annually beginning with GEM-1 in 1972. Even numbered models contain satellite and surface gravity data. Odd numbered models contain only satellite data. *Standard Earth* (SAO) and *Applied Physics Laboratory* (APL) models were among the first models. The basic information is from Vetter [18] and [19].

satellite inclinations and altitudes including surface gravity measurements and satellite altimetry data (Figure 1.2). From Vallado [15],

Many computational applications choose to truncate the gravitational field. While the rigorous approach requires the complete field, many applications use reduced gravity field orders to speed computational processing. Historically, there was some interest to truncate the gravity field for computational or program limitations. While this is often overlooked, some operational systems (AFSPC) often use a blanket 24×24 (for example) field for LEO orbits, rapidly truncating the gravity field as the orbits get higher. This may not be the best approach to accurately determine the orbit. Barker et al. [1] suggested a link in performance to the zonal truncation. Other studies have almost all examined the average behavior of the gravity field on the satellite orbit ephemeris. This may not tell the proper story for precise operations. Vallado [16] investigated the behavior of truncations for several satellites. One example is shown here for Japanese Earth Resources Satellite (JERS, about 500 km altitude circular orbit) (Figure 1.3).

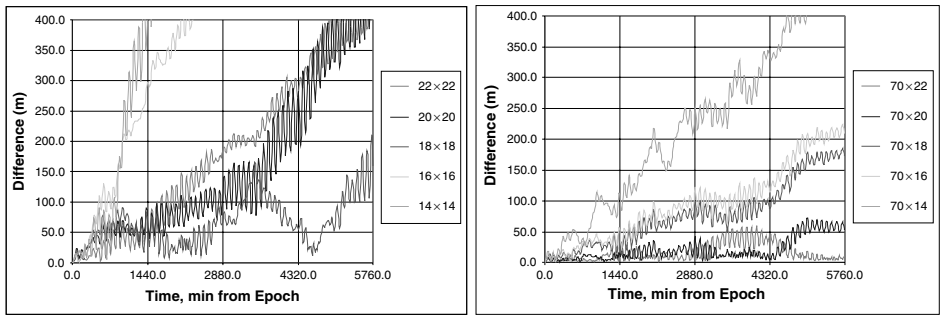

Fig. 1.3. Gravity field comparisons: Truncated gravity fields are compared to ephemeris runs for a complete EGM-96 70×70 field for a satellite at about 500 km altitude. The left plot is for a square gravity field. The right plot includes all the zonals in the truncations. The results do not always improve with a larger field (the differences for 22×22 are greater than 18×18 on the left, but the 70×22 is smaller than the 70×18 on right), but the accuracy generally improves as the non-square truncation is reduced.

Table 1.1
Fundamental Defining Parameters—EGM-96

Earth Semi-major Axis	$a = r_\oplus$	6378136.3 m
Flattening of the Earth	$1/f$	1.0/298.257
Angular Velocity of the Earth	$\omega_\oplus$	$7292115.8553 \times 10^{-11}$ rad/s
Earth's Gravitational Constant	GM, (μ)	$3.986004415 \times 10^5\,\mathrm{km}^3/\mathrm{s}^2$

Recognize that each time the gravity field is changed, the potential energy of the system changes, and an Orbit Determination (OD) process would produce a different state vector to reflect this change, based on the force models used during that evaluation. Although the most precise way to evaluate each force model would be to perform an OD on each individual case, the process would be unnecessarily long because we are only trying to establish the relative trends for each perturbation, not specific values for an individual case. As computers have become faster, the easiest approach is to simply use a complete gravity field (Table 1.1).

Models for gravitational perturbations are spherical harmonic expansions of the aspherical gravitational potential are in an Earth-centered, Earth-fixed reference frame.

1.4 Drag

The application of empirical atmospheric density models to astrodynamics in a real-world environment has been examined extensively since the launch of the first artificial satellites (Figure 1.4). Atmospheric density leads to significant drag effects for satellites below about 1000 km altitude, but its effects can be observed at altitudes well above this

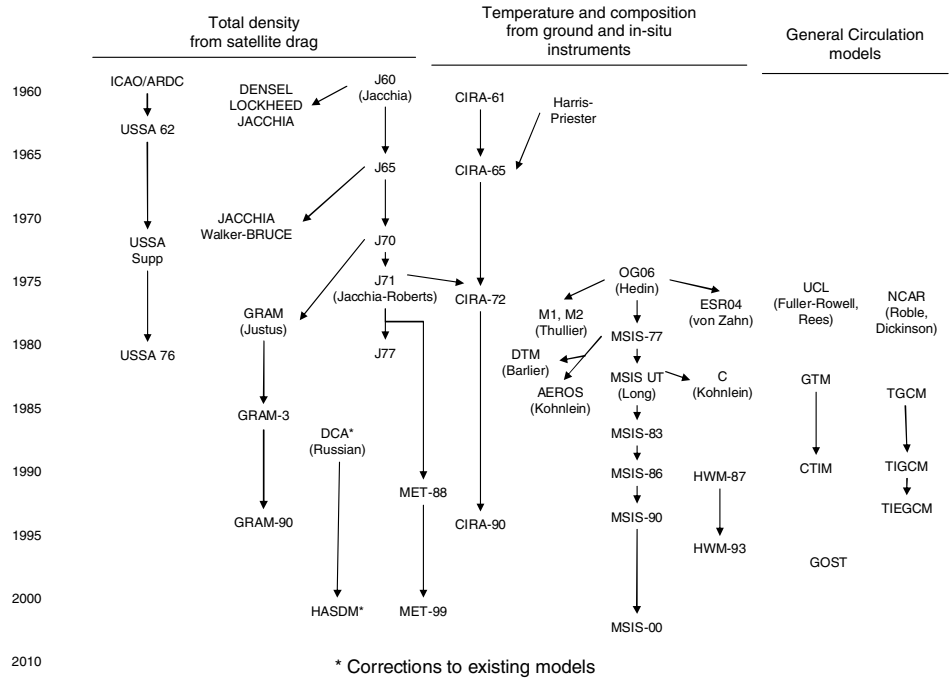

Fig. 1.4. Atmosphere Models: Notice the variety of models. Flow of information among the three overall categories is limited (Marcos, et al., 1993, 20). The main models in use today are the Standard Atmosphere, USSA76; variations of the Jacchia–Roberts, J71, J77, and GRAM90; COSPAR International Reference Atmosphere, CIRA90; Mass Spectrometer Incoherent Scatter, MSIS 00; Drag Temperature Model (DTM), Marshall Engineering Thermosphere (MET), the Russian GOST and general circulation models.

threshold. It is useful to review the basic acceleration equation. From Vallado [16], the following introduction and analysis is taken.

$$\vec{a}_{drag} = -\frac{1}{2}\rho\frac{c_D A}{m}v_{rel}^2\frac{\vec{v}_{rel}}{|\vec{v}_{rel}|} \tag{1.5}$$

ρ The density usually depends on the atmospheric model, Extreme Ultraviolet EUV, $F_{10.7}$, and geomagnetic indices a_p, prediction capability, atmospheric composition, etc. There is wide variability here, and many parameters that can cause significant changes. The popular parameters to examine today are the density and the exospheric temperatures. This single parameter represents the largest contribution to error in any orbit determination application.

c_D The coefficient of drag is related to the shape, but ultimately a difficult parameter to define. Gaposchkin [4] discusses that the c_D is affected by a complex interaction of reflection, molecular content, attitude, etc. It will vary, but typically not very much as the satellite materials usually remain constant.

A The cross-sectional area changes constantly (unless there is precise attitude control, or the satellite is a sphere). This variable can change by a factor of 10 or more depending on the specific satellite configuration. Macro models are often used for modeling solar pressure accelerations, but seldom if ever, for atmospheric drag.

m The mass is generally constant, but thrusting, ablation, etc., can change this quantity.

$\vec{v}_{rel}$ The velocity relative to the rotating atmosphere depends on the accuracy of the *a priori* estimate, and the results of any differential correction processes. Because it is generally large, and squared, it becomes a *very* important factor in the calculation of the acceleration.

The ballistic coefficient ($\mathrm{BC} = m/c_D A$ – a variation is the inverse of this in some systems) is generally used to lump the mass, area, and coefficient of drag values together. It *will* vary, sometimes by a large factor. Several initiatives are examining the time-rate of change for this parameter, but not looking at the variable area, and its effect in this combined factor. It is probably best not to model this parameter because it includes several other time-varying parameters that are perhaps better modeled separately.

There are numerous atmospheric drag models. Figure 1.4 lists some of the more popular models.

The primary inputs in any program are the atmospheric density (handled via a specified model), and the *BC*. The mass and cross-sectional area are usually well known, and an estimate of the drag coefficient permits reasonable approximations. The atmospheric models also vary depending on several factors, including the satellite orbit, intensity of the solar activity, and the geomagnetic activity. Vallado and Kelso [17] discuss the files needed to compile a seamless file for operations. They are available at http://celestrak.com/SpaceData.

Unlike any other force model, atmospheric drag receives extensive analysis and near-continual updates. The bottom line for drag (and to a lesser extent solar radiation pressure, as we will see shortly) is to have as many options and choices as possible. While the programming and certification tasks becomes more complicated, this non-conservative force is often the most difficult to match in ephemeris comparisons and having these options provides the user with a much greater ability to minimize differences with other programs.

There are three general observations that are important—the difference between atmospheric models, the variability that can result from treating the input data differently, and the actual implementation of an approach. Vallado [16] conducted a series of tests to determine the variability of different atmospheric models for a given satellite using a single flight dynamics program, and the differences resulting from the diverse treatment of the input solar weather data. The state vectors, epoch, *BC*, and solar radiation pressure coefficient ($m/c_r\ A_{sun}$) were held constant for all runs. The baseline used the Jacchia–Roberts atmospheric model. The simulations were run during a time of "average"

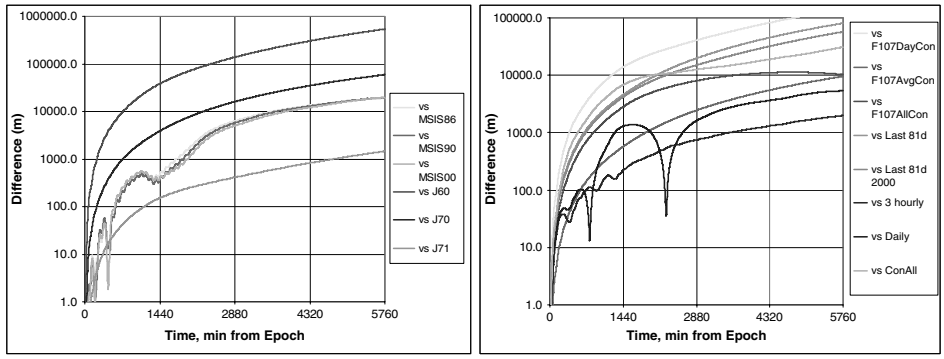

Fig. 1.5. Sample Atmospheric Drag Sensitivity: Positional differences are shown for JERS, about 500 km altitude and 97.6° inclination. Jacchia–Roberts drag is the baseline for all runs with 3-hourly interpolation. The left-hand graph shows the variations by simply selecting different atmospheric models. The right-hand graph shows the effect of various options for treating solar weather data. Specific options are discussed in the text. Note that the scales are the same, the relative effect of different models and solar data options are about the same, and any transient effects quickly disappear as the effect of drag overwhelms the contributions.

solar flux (January 4, 2003, $F_{10.7} \sim 140$). Minimum solar flux periods ($F_{10.7} \sim 70$) will show little difference. Maximum periods ($F_{10.7} \sim 220$) will show much larger excursions. Figure 1.5 shows the results for the JERS satellite, at about 500 km altitude and 97.6° inclination. Additional runs were performed with different satellites and as expected, the results were larger for lower and more eccentric orbits.

Most models as implemented in computer code, do not follow the exact technical derivation as defined in the literature. It is likely that none of the drag model implementations match the original technical definition. As a result, code contains numerous short cuts, and many additional features that may be the result of internal studies and information, but not the original work. This makes comparison of atmospheric models especially difficult.

Because atmospheric drag has perhaps the largest number of different models, defining an absolute standard is difficult to do, and would unnecessarily restrict research. There have been numerous studies to evaluate how well the atmospheric models perform, yet, no clear "winner" has ever emerged. Thus, we list models and present references that discuss the various merits of many of the models. An additional comment is necessary. Most models, as implemented in computer code, do not follow the exact technical derivation as defined in the literature. Numerous short cuts, and many additional features are included that may be the result of internal studies and information. This makes standardization of atmospheric models especially difficult.

For most of the simulations, the MSIS-86 and MSIS-90 models were quite close, as expected by the model descriptions. The Jacchia 1960 (J60) model appeared to be significantly different in all cases from the other models and J70 seemed to differ most from the J71 and JRob models. Because this chapter does not extensively examine comparisons with Precision Orbit Ephemerides (POEs), it is most important to come away with the overall level of variability within the different models. Essentially, if varying

atmospheric models show differences that are significantly larger than differences between flight dynamics programs using the "same" models, which is right? After examining these data, we conclude that neither are right. Primarily, this is due to the results shown on the right-hand side of Fig. 1.5 which are discussed next. Although each atmospheric model is carefully designed, the treatment of solar weather data by each program adds so much variability, coupled with the lack of independent references and availability of observational data for comprehensive evaluation makes it highly unlikely that one approach is definitive for all cases.

The following recommendations are set forth.

1. There should be an option to use either the last $F_{10.7}$ 81-day average, or the centered 81-day average. Atmospheric model descriptions generally cite a centered average, but this is impractical for many operational systems, and a trailing 81-day average is often used.
2. Using a_p should be seamless, but there is the possibility of difficulties for certain conversions of average values. There are discrete values for which a_p and k_p exist in the daily data. Thus, a program needs to be careful not to input a derived value that does not exist in the other scale. Inside a program, however, conversions may proceed without restriction to value. Consistency should be maintained with the atmospheric model.
3. The cubic splines routine discussed in Vallado and Kelso [17] should be used to interpolate geomagnetic indices.
4. The codes should treat all $F_{10.7}$ measurements at the time the measurement is actually taken. The offset (2000 UTC after May 31, 1991, 1700 UTC before) should be used with all $F_{10.7}$ and average $F_{10.7}$ values. Any model specific "day before", "6.7 hours before", etc., should be done with this offset in mind. There is not an established approach, yet it is a big factor (sometimes km level) in the comparisons.
5. The options for using a_p should be
   a. daily—just the daily values are interpolated. All 3-hourly values are ignored.
   b. 3-hourly—just the 3-hourly values are used. The daily values are ignored and there is no interpolation. This will produce step function discontinuities, but that could be useful in some programs.
   c. 3-hourly interp.—this should use the cubic splines from Vallado and Kelso (2005). It should produce the smoothest transitions from one time to the next while preserving the discrete values. The measurements should reproduce exactly at the measurement times (0000, 0300, 0600, etc. UTC), and be smooth in between.
6. The lag time for a_p values is somewhat fixed to 6.7 hours, but others have been proposed. Since it is a variable option, it would be prudent to have a means to change it, without recompiling the entire program.
7. The drag coefficient, area, and mass need to be included in state vector transmissions to permit increased accuracy in subsequent calculations.

Many sources state that the current atmospheric models introduce about a 15% error in the determination of atmospheric drag effects on a satellite. In fact, this is a combination of the inaccuracy of the predictions of the solar flux and geomagnetic indices, the imperfect nature of the mathematical models, imprecise information about the molecular interaction of the satellite and the atmospheric particles, and several others.

1.5 3-Body

Third body effects include the perturbations induced by the gravitational influence of the Sun, Moon, and the planets. These are also called n-body perturbations acting on the satellite. The contributions are computed using a point-mass equation. However, the Sun and Moon also include an indirect effect as an interaction between a point-mass perturbing object and an oblate earth. Thus the third-body perturbation includes both direct and indirect terms of point mass third-body perturbations.

The general form of the acceleration due to third-body forces is

$$\vec{a}_{3\text{-}body} = -\frac{G(m_E + m_{sat})}{r_{Esat}^3}\vec{r}_{Esat} + Gm_3\left(\frac{\vec{r}_{sat3}}{\vec{r}_{sat3}^{\,3}} - \frac{\vec{r}_{E3}}{\vec{r}_{E3}^{\,3}}\right) \tag{1.6}$$

Analytical and numerically generated models dominate astrodynamic programs. Many applications use the analytical approaches because they provide adequate accuracy. However, numerical routines often require the additional accuracy of the JPL models.

1.6 Solar radiation pressure

The force due to solar radiation pressure (SRP) rises when photons from the Sun impinge on a satellite surface and are absorbed (or reflected-specular and diffuse) thus transferring photon impulse to the satellite. In contrast to drag, the SRP force does not vary with altitude and its main effect is a slight change in the eccentricity and longitude of perigee. The effect of SRP is most notable for satellites with large solar panels like communications satellites and GPS and depends on its mass and surface area. In cases of geodetic precision orbits, complex modeling of the exposed satellite surfaces have to modeled usually using finite-element computer codes. This is the case with GPS where SRP represents an important force.

Vallado [16] provides a background for SRP and is included herein. Although not studied as extensively in the literature, it poses many of the same challenges as atmospheric drag, but has a significantly smaller effect than the other forces. Consider the basic equation.

$$\vec{a}_{srp} = -\rho_{SR}\frac{c_R A_{Sun}}{m}\frac{\vec{r}_{sat-Sun}}{|\vec{r}_{sat-Sun}|} \tag{1.7}$$

ρ_{SR} The incoming solar pressure depends on the time of year, and the intensity of the solar output. It is derived from the incoming solar flux and values of about 1358–1373 $\mathrm{W/m^2}$ are common.

c_R The coefficient of reflectivity indicates the absorptive and reflective properties of the material, and thus the susceptibility to incoming solar radiation.

A_{Sun} The cross-sectional area changes constantly (unless there is precise attitude control, or it is spherical). This variable can change iby a factor of 10 or more depending on the specific satellite configuration. Macro models are often used for geosynchronous satellites. This area is generally *not* the same as the cross-sectional area for drag.

m The mass is generally constant, but thrusting, ablation, etc., can change this quantity.

$r_{\text{sat}-\text{Sun}}$ The orientation of the force depends on the satellite–Sun vector—again a difference with atmospheric drag.

Despite the simple expression, accurate modeling of solar radiation pressure is challenging for several reasons. The major error sources are:

- Use of macro models/attitude—this is perhaps the largest difference between programs
- Use of differing shadow models (umbral/penumbral regions, cylindrical, none, etc.)
- Using a single value for the incoming solar luminosity, or equivalent flux at 1 AU
- Use of an effective Earth radius for shadow calculations (23 km additional altitude is common)—this approximates the effect of attenuation from the atmosphere
- Using different methods to account for seasonal variations in the solar pressure
- Not integrating to the exact points of arrival and departure at the shadow boundary
- Use of simplified treatment for the light-time travel from the Sun to the satellite (instantaneous (true), light delay to central body accounted for (app. to true), light delay to satellite (default))

A series of runs were made to determine the impact of each of these items on the results for a few selected satellites. Results are shown in Figure 1.6 for a nominal Global Positioning Satellite (GPS satellite).

1.7 Tides

Earth tidal effects on satellites are due to pole tides, ocean tides, and solid earth, tides. Most of the data that have resulted in a definitive model have come about within the last several years from satellites such as TOPEX and GRACE. The basis of the models for pole, solid earth, and ocean tide models can be found in IERS Conventions, with the latest update in McCarthy and Petit [12]. Tidal models do not enjoy the variety of the gravitational and atmospheric models yet, but there are several different approaches. These various models can be a factor if precise comparisons are desired. At this point in time, several models exist, and no clear "leader" has been recognized as the standard approach.

Pole tides define the rotational deformation of the pole due to an elastic earth. These are modeled by the C_{21} and S_{21} coefficients in the earth's potential.

Solid earth tidal contributions are computed as *corrections* to the spherical harmonics coefficients.

There are a wide variety of ocean tide models in existence that have been used since 1980 starting with the Swiderski hydrodynamic model. One current models is from the University of Texas, Center for Space Research (CSR) and are referred to by CSR4.0 which model the long wave-length characteristics. The early model was 1 degree by 1

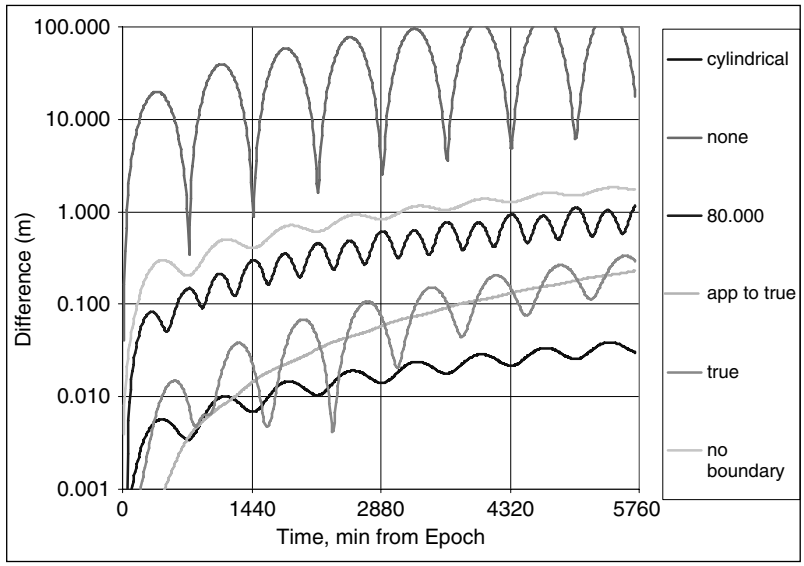

Fig. 1.6. Sample solar radiation pressure sensitivity. Positional differences are shown for a GPS satellite which is in eclipse. The baseline is a dual-cone (umbra/penumbra) shadow model. Using no shadow model (none) produces the largest differences. A simple cylindrical model introduces modest differences. Shadow boundary mitigation (no boundary) and the effective Earth size (23) contribute noticeable differences. The treatment of light travel time between the Sun and central body (app. to true) and instantaneous travel (true) produce smaller, but still detectable results.

degree model and the later model are 0.5×0.5 degree in extent. The IERS conventions describe another model. Most of these models have been developed since about 1995. The models use highly precise measurements from satellites such as JASON and TOPEX. Additional satellites such as GRACE and GOCE will contribute to a better future models.

1.8 Albedo

Albedo is the radiation pressure emitted from the Earth which causes a small perturbing force on a satellite. Although the effect of SRP is far larger, the effects of Earth's albedo can be comparable for certain configurations of orbits (e.g. sun-synchronous). The acceleration due to albedo is generally expressed in terms of a second degree zonal spherical harmonic model, and contributions from various Earth sectors are summed to determine the overall effect.

1.9 Other

As the accuracy of orbit determination and propagation increases, additional force models are included in analyses. In particular, applications using GPS data often must account for the [primarily] apsidal rotation caused by General Relativity. GPS signals

msut also be corrected for General Relativity, as well as atomic clock corrections. The effects of General Relativity are very small and only become important where high orbit precision at the cm level is needed. Satellites, such as the Gravity Probe B, will try to measure and quantify this effect in its verification and validation of Einstein's theory of General Relativity. Satellite thrusting can also be a significant perturbing force. Many satellites use maneuvers for mission operation and for orbit maintenance. The forces induced by these motor firings can be large and small. We do not describe these in any detail, but introduce the forces as something needed to be considered in mission planning and precision orbit determination modeling.

1.10 Propagating the orbit

There are several techniques to propagate an orbit. Generally, analytical, numerical, and semianalytical techniques encompass the potential choices. However, a primary technique to find analytical solutions, the variation of parameters, may be used in either analytical or numerical applications. The fundamental distinction is the use of position and velocity state vectors, or orbital elements as the elements of the state. Recalling the previous force model discussion, one can find benefits with each approach.

Direct integration where possible. Analytical methods are accurate and yield a quick solution; however, the series truncations may be difficult depending on the equations of motion. The numerical method is very accurate with the correct step size, but this determination may be tricky, and long propagations can still be time-consuming. The semianalytical technique combines the analytical and numerical approaches.

1.11 Analytical

General perturbation techniques replace the original equations of motion with an analytical approximation that captures the essential character of the motion over some limited time interval and which also permits analytical integration. Such methods rely on series expansions of the perturbing accelerations and are usually derived from variation of parameter equations which will be addressed shortly. In practice, we truncate the resulting expressions to allow simpler expressions in the theory. This trade-off speeds up computation but decreases accuracy. Unlike numerical techniques, analytical methods produce approximate, or "general" results that hold for some limited time interval and accept any initial input conditions. The quality of the solution degrades over time, but remember that the numerical solution also degrades—at different rates and for different reasons. Analytical techniques are generally more difficult to develop than numerical techniques, but they often lead to a better understanding of the perturbation source.

1.12 Numerical

Special perturbation techniques numerically integrate the equations of motion including all necessary perturbing accelerations. Because numerical integration is involved, we can

think of numerical formulations as producing a *specific*, or *special*, answer that is valid only for the given data (initial conditions and force-model parameters). To numerically integrate Cowell's formulation, we must have mathematical models for each perturbing force. The general form is usually taken as the following Albedo.

$$\vec{a} = \frac{\mu r}{r^3}\vec{r} + \vec{a}_{non\text{-}spherical} + \vec{a}_{drag} + \vec{a}_{3\text{-}body} + \vec{a}_{srp} + \vec{a}_{tides} + \vec{a}_{other} \tag{1.8}$$

Numerical integration may also be applied to the variation of parameter (VOP) equations, in which case a set of orbital elements is numerically integrated. To form an ephemeris, one then needs to convert the osculating orbital elements into the appropriate state vectors.

Although numerical methods can give very accurate results and often establish the "truth" in analyses, they suffer from their specificity, which keeps us from using them in a different problem. Thus, new data means new integration, which can add lengthy computing times. NASA began the first complex numerical integrations during the late 1960s and early 1970s. Personal computers now compute sufficiently fast enough to perform complex perturbation analyses using numerical techniques. However, numerical methods suffer from errors that build up with truncation and round-off due to fixed computer word-length. These errors can cause numerical solutions to degrade as the propagation interval lengthens.

1.13 Semianalytical

Semianalytical techniques combine the best features of numerical and analytical methods to get the best mix of accuracy and efficiency. The result can be a very accurate, relatively fast algorithm which applies to most situations. But semianalytical techniques vary widely. We choose a semianalytical technique mainly for its ability to handle varying orbital applications, its documentation, and the fidelity and the number of force models it includes. Most semianalytical techniques have improved accuracy and computational efficiency, but the availability of documentation (including very structured computer code) and flexibility are often important discriminators. We consider a technique semianalytical if it is not *entirely* analytical or numerical.

1.14 Variation of parameters

Most analytical, and some numerical solutions rely on the *variation of parameters* (VOP) form of the equations of motion originally developed by Euler and improved by Lagrange [8]. The overall process is called *variation of parameters* (VOP) because the orbital elements (the constant parameters in the two-body equations) are changing. Lagrange and Gauss both developed VOP methods to analyze perturbations—Lagrange's technique applies to conservative accelerations, whereas Gauss's technique also works for non-conservative accelerations. Depending on the orbital elements chosen, the form will differ. I will show a form for the classical orbital elements.

Using the VOP technique, we can analyze the effects of perturbations on specific orbital elements. This is very useful in mission planning and analysis. We want any theory to model as many perturbing forces as possible. Most operational analytical theories are limited to central body and drag. Analytical expressions for third-body and solar-radiation forces are far less common, mainly because their effects are much smaller for many orbits. Also, whenever accuracy requires us to use effects of third bodies and solar-radiation pressure, numerical integration is usually just as easy for all the perturbing forces.

1.15 Lagrangian VOP—conservative forces

The VOP method is a formulation of the equations of motion that are well-suited to perturbed, dynamical systems. The concept is based on the premise that we can use the solution for the unperturbed system to represent the solution of the perturbed system, provided that we can generalize the constants in the solution to be time-varying parameters. The unperturbed system is the two-body system, and it represents a collection of formulas that provide the position and velocity vectors at a desired time. Remember, these formulas depend only on the six orbital elements and time. In principle, however, we could use any set of constants of the unperturbed motion, including the initial position and velocity vectors. Time is related to the equations of motion through the conversions of mean, eccentric, and true anomaly.

The general theory for finding the rates of change of the osculating elements is known as the *Lagrange planetary equations of motion*, or simply the *Lagrangian VOP*, and is attributed to Lagrange because he was the first person to obtain these equations for all six orbital elements. He was concerned with the small disturbances on planetary motion about the Sun due to the gravitational attraction of the planets. He chose to model the disturbing acceleration due to this conservative perturbation as the gradient of a potential function. From Vallado [15],

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{2}{na}\frac{\partial R}{\partial M_o}$$

$$\frac{\mathrm{d}e}{\mathrm{d}t} = \frac{1-e^2}{na^2e}\frac{\partial R}{\partial M_o} - \frac{\sqrt{1-e^2}}{na^2e}\frac{\partial R}{\partial \omega}$$

$$\frac{\mathrm{d}i}{\mathrm{d}t} = \frac{1}{na^2\sqrt{1-e^2}\sin(i)}\left\{\cos(i)\frac{\partial R}{\partial \omega} - \frac{\partial R}{\partial \Omega}\right\}$$

$$\frac{\mathrm{d}\omega}{\mathrm{d}t} = \frac{\sqrt{1-e^2}}{na^2e}\frac{\partial R}{\partial e} - \frac{\cot(i)}{na^2\sqrt{1-e^2}}\frac{\partial R}{\partial i}$$

$$\frac{\mathrm{d}\Omega}{\mathrm{d}t} = \frac{1}{na^2\sqrt{1-e^2}\sin(i)}\frac{\partial R}{\partial i}$$

$$\frac{\mathrm{d}M_o}{\mathrm{d}t} = \frac{1-e^2}{na^2e}\frac{\partial R}{\partial e} - \frac{2}{na}\frac{\partial R}{\partial a}$$

1.16 Gaussian VOP—nonconservative forces

For many applications, it is convenient to express the rates of change of the elements explicitly in terms of the disturbing forces—actually acceleration (*specific* forces) to match the units in the equations. Gauss's form of VOP is advantageous for non-conservative forces because it is expressed directly from the disturbing acceleration. But it works equally well for conservative forces because the forces are simply gradients of the potential functions. It is also easy to visualize this representation because we're familiar with the concept of a force. Gauss's form of the VOP requires the partial derivatives of the elements with respect to the velocity. We must determine these for particular element sets. Gauss chose to develop the equations in the RSW system.[1] Let the components of the disturbing force (per unit mass) be along the radius vector, perpendicular to the *R*axis in the orbit plane in the *direction* of satellite motion, and normal to the orbit plane. The disturbing (specific) force become

$$\frac{da}{dt} = \frac{2}{n\sqrt{1-e^2}}\left(e\sin(v)F_R + \frac{p}{r}rF_S\right)$$

$$\frac{de}{dt} = \frac{\sqrt{1-e^2}}{na}\left(\sin(v)F_R + \left(\cos(v) + \frac{e+\cos(v)}{1+e\cos(v)}r\right)F_S\right)$$

$$\frac{di}{dt} = \frac{r\cos(u)F_W}{na^2\sqrt{1-e^2}}$$

$$\frac{d\Omega}{dt} = \frac{r\sin(u)F_W}{na^2\sqrt{1-e^2}\sin(i)}$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{na^2\sin(i)}\left\{-\cos(v)F_R + \sin(v)\left(1+\frac{r}{p}\right)F_S\right\} - \frac{r\cot(i)\sin(u)F_W}{h}$$

$$\frac{dM_o}{dt} = \frac{1}{na^2e}\{(p\cos(v) - -2er)F_R - (p+r)\sin(v)F_S\}$$

These VOP equations in classical orbital elements have some limitations. First, they are limited to eccentricities less than 1.0 because of the presence of the eccentricity in the denominator and in square roots. Also note that they suffer from the same singularities as Lagrange's form of the VOP equations because the singularities are due to the particular element set, not how the disturbing forces are characterized. The rate of change of Ω has $\sin(i)$ in the denominator. This causes the equation to be indeterminate for small inclinations. A similar problem exists for ω with small values of eccentricity. Thus, this particular set of equations is not recommended for small values of eccentricity or inclination.

[1] In the RSW system, the *R* axis is parallel to the position vector. Along-track displacements are normal to the position vector (along the *S* axis), but not necessarily aligned with the velocity vector. The *W* axis points in the instantaneous direction of the angular momentum vector.

1.17 Effect on orbits

Given the numerous techniques to analyze and account for the effects of perturbations, there are several different trends that can be noted for different satellite orbits. However, be aware that the results are specific to individual orbits. Also note that the increasing popularity of numerical methods has positive and negative effects. These techniques include all effects from the perturbations, but they do not indicate the source of dominant errors, and can thus be problematic for the satellite mission designer. We examine two major areas here—a J_2 secular perturbation, and the effects of combined forces resulting from numerical simulations.

1.18 J_2 Only

Due to its simplicity, studies are often conducted with only the secular effects of J_2 included. While this is true for many systems, the modern computer renders many of these analyses obsolete for precise studies. Still, the effects are illustrative of the effect on satellite orbits. Consider the nodal regression (Figure 1.7) and the apsidal rotation (Figure 1.8). These two effects are common for satellite mission planners, and they result from the secular effect of J_2.

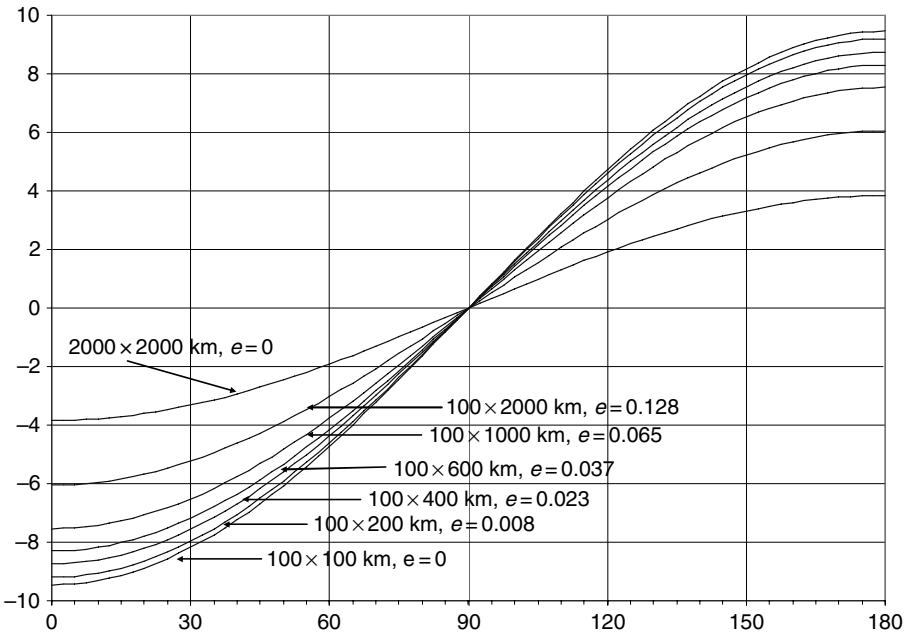

Fig. 1.7. Daily nodal regression. For the eccentric orbits, I used a perigee altitude of 100 km with the apogee values as indicated. The 100 × 2000 km orbit shows that the perturbing effect for each of the *eccentric* orbits would be smaller if the orbit were circular at that apogee altitude (2000 × 2000 km) [15].

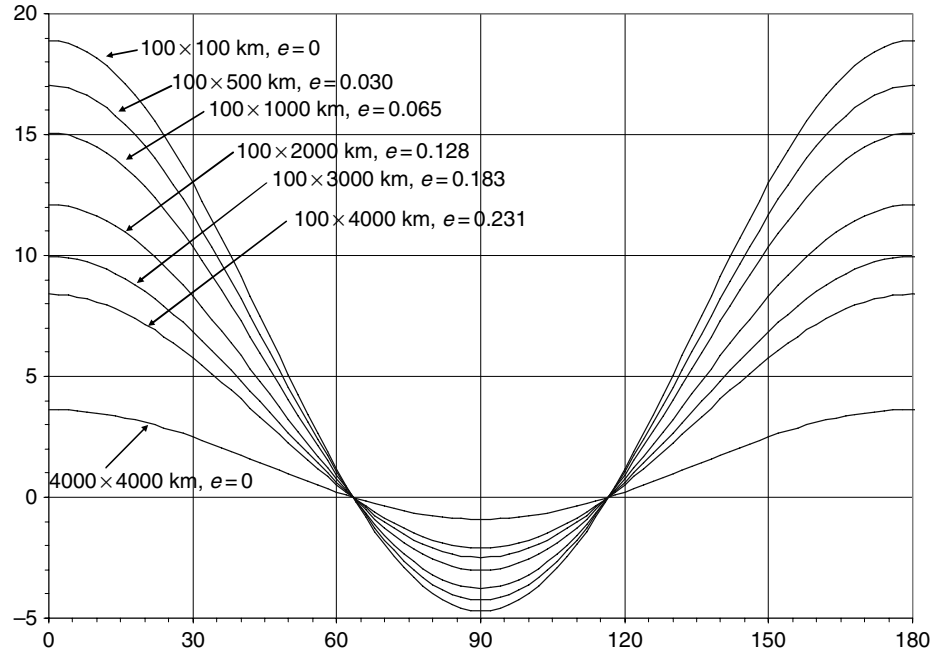

Fig. 1.8. Daily Apsidal Regression. As with nodal regression, circular orbits at an altitude (say, 4000 × 4000) would have a smaller daily change than an eccentric orbit with apogee at the same altitude (100 × 4000 km) [15].

1.19 Comparative force model effects

The effects of all physical forces affecting the motion of the earth satellite or spacecraft can be complex, and is often best handled by numerical integration. By far the largest effect is due to gravitation, usually followed by atmospheric drag, solar radiation pressure (SRP) effects and several other effects such as tides, third body perturbations and others. Vallado (2005) shows the relative effect of various forces on several satellites. Figure 1.9 reproduces these quantitative effects of all physical forces in terms of positional differences for two satellites—one is in an 800 by 110 km altitude, 50° inclined orbit, while the other is a geosynchronous satellite. Note that most of the effects like tides and third body forces and relativity are very small, but need to be taken into account when precision is of importance. The satellite parameters were chosen to illustrate force model effects. Each spacecraft parameter was held constant (coefficient of drag $c_D = 2.2$, coefficient of solar radiation pressure $c_R = 1.2$, area to mass ratio $A/m = 0.04\,m^2/kg$). The simulation time, January 4, 2003, was chosen as the epoch to propagate as this was a moderate period of solar activity (solar flux $F_{10.7} \sim 140$). The baseline for comparison in all cases was a 12 × 12 EGM-96 gravity field (degree and order, 12 zonal harmonics plus 12 sectoral terms). Except for the two-body (0 × 0) and 70 × 70 cases, all the other force model comparisons included a 12 × 12

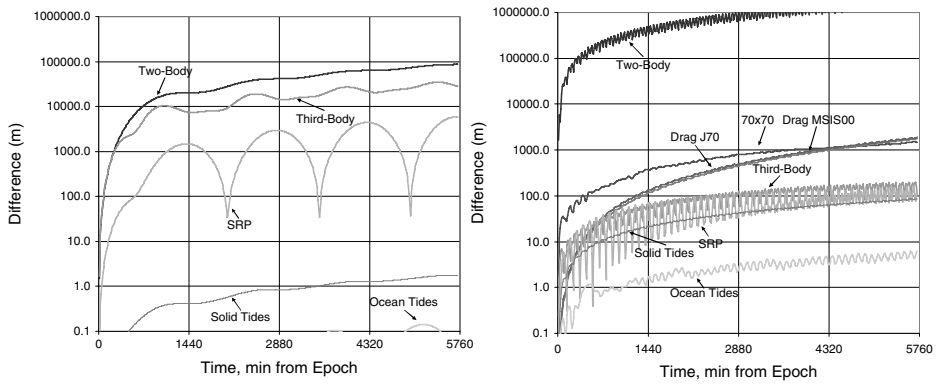

Fig. 1.9. Force model comparisons: This figure shows the positional difference over time (four days) when using various force models on the same initial state for a geosynchronous satellite. Each comparison is made with respect to a two-body ephemeris except for the gravity runs which compare to the nearest gravity case. Thus, "12×12" is a comparison of a 12×12 WGS84/EGM96 gravity field to a WGS84/EGM96 2×10 gravity field ephemeris, etc.

EGM-96 gravity field. Thus, for example, the "vs Drag JRob" case indicates a comparison of 12×12 EGM-96 gravity and a 12×12 EGM-96 gravity with Jacchia–Roberts drag.

In general, gravity was the largest single perturbation source, so additional tests were conducted to determine the sensitivity of this perturbation force. Atmospheric drag was generally second for lower orbits, but third-body effects were much higher for higher altitude satellites. Because the study results indicated the conservative forces could be matched to cm-level, no additional studies were performed on third-body forces. Drag was considered separately. It is important to note that these are prediction differences are based on the propagation of identical state vectors with differing acceleration models. A study of orbit determination accuracy using differing acceleration models would produce a very different set of results.

1.20 Conclusions

There are numerous forces that affect a satellite in orbit. Proper treatment of these forces is essential to proper mission planning and satellite operations. The selection of the type of propagation scheme and consistency with the chosen technique are important.

References

1. Barker, William N., S. J. Casali, and R. N. Wallner. (1995). The Accuracy of General Perturbations and Semianalytic Satellite Ephemeris Theories. Paper AAS-95-432 presented at the AAS/ AIAA Astrodynamics Specialist Conference. Halifax, Nova Scotia, Canada.
2. Battin, Richard H. (1987). *An Introduction to the Mathematics and Methods of Astrodynamics*. AIAA Education Series, New York.

3. Brouwer, Dirk, and G. M. Clemence. (1961). *Methods of Celestial Mechanics*. New York: Academic Press, Inc.
4. Gaposchkin, E. M. (1994). Calculation of Satellite Drag Coefficients. Technical Report 998. MIT Lincoln Laboratory, MA.
5. Kaplan, Marshall H. (1976). *Modern Spacecraft Dynamics and Control*. New York: John Wiley & Sons.
6. Kaula, William M. (1966). *Theory of Satellite Geodesy*. Waltham MA: Blaisdell Publishing Co.
7. Kreyzig, Erwin. (1983). *Advanced Engineering Mathematics*. 5th edn. New York: John Wiley Publishing.
8. Lagrange, J. L. (1873). *Collected Works*. Vol. 6. Paris: Gauthier-Villars.
9. Lambeck, Kurt. (1988). *Geophysical Geodesy*. Oxford: Clarendon Press.
10. Long, Anne C. et al. (1989). *Goddard Trajectory Determination System (GTDS) Mathematical Theory (Revision 1)*. FDD/552-89/001 and CSC/TR-89/6001. Goddard Space Flight Center: National Aeronautics and Space Administration.
11. Marcos, Frank A. et al. (1993). Satellite Drag Models: Current Status and Prospects. Paper AAS-93-621 presented at the AAS/AIAA Astrodynamics Specialist Conference. Victoria BC, Canada.
12. McCarthy, Dennis, and Gerard Petit. (2003). *IERS Technical Note #32*. U.S. Naval Observatory.
13. Moritz, Helmut, and I. Mueller. (1987). *Earth Rotation—Theory and Observation*. New York: Ungar Publishing Company.
14. Roy, Archie E. (1988). *Orbital Motion*. New York: John Wiley & Sons.
15. Vallado, David A. (2004). *Fundamentals of Astrodynamics and Applications*. 2nd edn, second printing. Microcosm, El Segundo, CA.
16. Vallado, David A. (2005). An Analysis of State Vector Propagation using Differing Flight Dynamics Programs. Paper AAS 05-199 presented at the AAS/AIAA Space Flight Mechanics Conference. Copper Mountain, CO.
17. Vallado, David A., and T. S. Kelso. (2005). Using EOP and Solar Weather Data for Real-time Operations. Paper USR 05-S7.3 presented at the US/Russian Space Surveillance Workshop, August 22–26, 2005. St Petersburg, Russia.
18. Vetter, Jerome R. (1994). The Evolution of Earth Gravity Models used in Astrodynamics. *APL Technical Digest*, John Hopkins. **15**(4), pp. 319–335.
19. Vetter, Jerome R. (2001). Private communication.