A shadow function model based on perspective projection and atmospheric effect for satellites in eclipse

Zhen Li, Marek Ziebart, Santosh Bhattarai, David Harrison

Abstract

Accurate Solar Radiation Pressure (SRP) modelling is critical for correctly describing the dynamics of satellites. A shadow function is a unitless quantity varying between 0 and 1 to scale the solar radiation flux at a satellite’s location during eclipses. Errors in modelling shadow function lead to inaccuracy in SRP that degrades the orbit quality. Shadow function modelling requires solutions to a geometrical problem (Earth’s oblateness) and a physical problem (atmospheric effects). This study presents a new shadow function model (PPM$_{atm}$) which uses a perspective projection based approach to solve the geometrical problem rigorously and a linear function to describe the reduction of solar radiation flux due to atmospheric effects. GRACE (Gravity Recovery And Climate Experiment) satellites carry accelerometers that record variations of non-conservative forces, which reveal the variations of shadow function during eclipses. In this study, the PPM$_{atm}$ is validated using accelerometer observations of the GRACE-A satellite. Test results show that the PPM$_{atm}$ is closer to the variations in accelerometer observations than the widely used SECM (Spherical Earth Conical Model). Taking the accelerometer observations derived shadow function as the “truth”, the relative error in PPM$_{atm}$ is -0.79% while the SECM 11.07%. The influence of the PPM$_{atm}$ is also shown in orbit prediction for Galileo satellites. Compared with the SECM, the PPM$_{atm}$ can reduce the radial orbit error RMS by 5.6 cm over a 7-day prediction. The impacts of the errors in shadow function modelling on the orbit remain to be systematic and should be mitigated in long-term orbit prediction.

Keywords: eclipse; SRP shadow function; orbit quality in eclipse; GRACE; GNSS

1. Introduction

Eclipses occur for both celestial bodies and artificial satellites. Solar eclipses reduce the solar irradiance at a satellite’s location because all or part of the Sun is blocked by the occulting body. For artificial Earth satellites, the occulting body can be the Earth or the moon. Eclipses caused by the Earth are far more frequent than those caused by the moon (Vallado, 2007; Srivastava et al, 2015). However, lunar eclipses do affect the orbit of artificial satellites (Khalil and Abd, 2011; Song and Kim, 2016; Zhang et al, 2018).

Three conditions define the eclipse state: full phase, penumbra, and umbra. Under conditions of full phase, the Sun is fully visible to the satellite, the only factor influencing the energy incident on the satellite is the satellite–Sun distance. The shadow function is the ratio of solar radiation incident on a satellite during penumbra to that at full phase. The value of the shadow function $F_s$ for the three eclipse states is shown in Tab. 1. Orbital trajectory calculations rely upon modelling forces accurately. Ignoring the variations in solar flux caused by shadow function effects will result in orbit errors at some level (Vallado, 2007; Montenbruck and Gill, 2005). This is particularly important for missions that require cm-level orbit accuracy (Sibthorpe et al, 2011; Montenbruck et al, 2013; Solano, 2014; Robertson, 2015; Guo et al, 2016; Kucharski et al, 2017).

Table 1: The value of a shadow function

<table>
<thead>
<tr>
<th>eclipse state</th>
<th>shadow function ($F_s$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>umbra</td>
<td>0</td>
</tr>
<tr>
<td>penumbra</td>
<td>$0 &lt; F_s &lt; 1$</td>
</tr>
<tr>
<td>full phase</td>
<td>1</td>
</tr>
</tbody>
</table>

Conventional approaches to shadow function modelling use a spherical Earth assumption. There are currently two approaches that are widely used: the SECM (Spherical Earth Conical Model) and the CYM (Cylinder Model) (Longo and Rickman, 1995; Neta and Vallado, 1997; Montenbruck et al, 2005; Vallado, 2007). As shown in Fig. 1, the CYM only describes the full phase and umbra, thus, it can not be used in shadow function modelling. The SECM (shown in Fig. 2) is able to deal with the full-phase, penumbra and umbra. It uses the ratio of the unblocked solar disk area to the area of the whole solar disk to calculate the shadow function (Montenbruck et al, 2005). However, the real Earth is closer to an ellipsoid than a sphere, it is more accurate to model the shadow function with an ellipsoidal Earth. In addition, the solar radiation can interact with the Earth’s at-
mosphere when it goes through. Therefore, a logical step in improving shadow function accuracy would be to consider the Earth’s oblateness—a geometrical effect—and refraction, scattering and absorption by the atmosphere—which are physical effects.

An elegant approach to determine the eclipse states (full phase, penumbra and umbra) for an ellipsoidal Earth can be found in Adhya et al (2004). However, they did not propose a method to model the shadow function. Although a shadow function reducing from 1 (at full phase) to 0 (at umbra) is used in Adhya (2005), the Earth was modelled as a sphere in the research. This method used a distance from the satellite to the boundary of penumbra as a measure to calculate the value of shadow function. It is difficult to consider atmospheric effects in this method because atmospheric effects cannot be included in a “distance”. In the aspect of shadow function modelling, the approach from Adhya, 2005 has similar performance as the SECM, based on the fact that they both use spherical Earth assumption and ignore the atmospheric effects. There are also other approaches that consider the Earth’s oblateness in eclipse modelling (Ferraz Mello, 1972; Longo and Rickman, 1995; Neta and Vallado, 1997; Ismail et al, 2015). These approaches do not deal with the shadow function problem. However, they do predict penumbra passage durations.

There are a few studies considering the atmospheric effects in modelling the shadow function. Kabelac (1985) studied the impacts of improved shadow function on the orbital elements by considering Earth’s atmosphere but with a spherical Earth. Additionally, a theory of Earth penumbra transition effects can be found in Vokrouhlicky et al (1993). This method assumes the atmosphere is a sphere and makes approximations in the atmospheric absorption and scattering processes. Robertson (2015) proposed a physical shadow function model for SRP modelling during penumbra. In his study, he assumes the Earth is a sphere but adjusts the radius of the Earth according to the position of the satellite so that the spherical surface fits the ellipsoidal Earth locally. Both the methods in Vokrouhlicky et al (1993) and Robertson (2015) introduced complex physics (including several physical processes like scattering, atmospheric absorption and refraction) in modelling the atmospheric effects of the shadow function. The work from Vokrouhlicky et al (1993) is highly theoretical and the model from Robertson (2015) was only tested on the GRACE mission.

In this study, a combination of the perspective projection algorithm (Oswald et al, 1982) and the linear reduction of the solar radiation flux in Earth’s atmosphere is used to build a new shadow function model (the PPM_atm). The geometrical problem is solved by the perspective projection using rigorous mathematics. The physical problem is dealt with an assumption that the radiation reduction in the Earth’s atmosphere is evenly linear.

In this paper, the modelling methods and the limitations of several widely used shadow function models are described. Then, the new perspective projection based shadow function with atmospheric effects (PPM_atm) is introduced. Finally, the PPM_atm model is validated with the GRACE (Gravity Recovery And Climate Experiment) accelerometer observations (Taylor et al, 2004) following the method proposed by Robertson (2015). The improvement in orbit prediction for the Galileo satellites using the PPM_atm is also shown.

![Figure 1: The CYM (spherical Earth cylindrical model), it can only describe full phase and umbra states](image1)

![Figure 2: The SECM (spherical Earth conical model), it can describe all the possible eclipse states including annular umbra](image2)

2. Method

Let the satellite centre of mass be a viewpoint and place an image plane between the satellite and the Sun. The normal to the image plane lies in Sun–satellite direction. The Earth (an ellipsoid) and the Sun (a sphere) are projected onto the image plane using perspective projection. Perspective projection is a 3D projection that makes objects appear as they would effectively to the human eye. That is, objects in the distance appear smaller than those are close under perspective projection.

The overlapping area between the Sun’s image and the Earth’s image can be computed according to their perspective projections. The amount of unblocked area of the solar disk on the image plane is a measure of the shadow function. Thus, the key to the geometrical problem is to work out how much of the solar disk is not blocked by the Earth’s image.

In addition, when the solar disk is in the projection of the Earth’s atmosphere, the brightness of solar disk will be varying due to the solar radiation reduction caused by atmospheric effects. In this study, a linear function is used to describe the radiation reduction caused by such atmospheric effects.

2.1. Basic geometrical relations and coordinate systems

Several coordinate systems are defined to clarify the geometrical relations between the Sun, the Earth and the satellite. The ECEF (Earth Centred Earth Fixed) frame is chosen as the global system. We define the ISF (Image Space Frame) to deal
with calculations on the image plane. Figure 3 shows the definition of ISF and the geometrical relations between the Sun, the Earth, and the satellite. The image plane is at a distance of $\gamma$ from the satellite. We define the origin (denoted as $o$ in ECEF) of the ISF to be the projection of the Sun’s centre (PSC) on the image plane, $z$-axis (denoted as a unit vector $n$ in ECEF) of the ISF to be along the vector from the Sun to the satellite, $x$-axis (denoted as a unit vector $u$ in ECEF) of the ISF to be from the origin to the projection of Earth’s mass centre (PEC), $y$-axis (denoted as a unit vector $v$ in ECEF) lies in the image plane and completes the right-hand coordinate system.

The unit vector $u$ along the positive $x$-axis of the ISF can be worked out from the geometrical relation between the PSC, PEC and the satellite.

$$u = \frac{r - (n \cdot r) n}{\|r - (n \cdot r) n\|}$$  \hspace{1cm} (4)

The definition of $u$ does not exist when the Earth is between the satellite and the Sun, and the Earth’s centre is on the satellite–Sun line. However, in this situation, the satellite is either in full phase or umbra. It will be discussed later in section 2.2.

The unit vector along the positive $y$-axis of the ISF is

$$v = n \times u$$  \hspace{1cm} (5)

The coordinates of the ISF origin are

$$o = r - \gamma n$$  \hspace{1cm} (6)

2.2. Existence of the Earth’s image on the projection plane

The solution to the geometrical problem is worked out according to the perspective projection. However, the Earth’s image on the projection plane may not exist when the projection plane is in front of the Earth. Therefore, before starting the calculation of the perspective projection, the condition of the existence of the Earth’s image has to be checked.

We define a plane including the centre of mass of the Sun, the satellite and the Earth. In this plane, we draw a line through the centres of mass of the satellite and the Sun (see line $S_1S_2$ in Fig. 4). There exist two planes which are perpendicular to the line and are tangent to the Earth (planes $S_1P_1$ and $S_2P_2$ in Fig. 4, $S_1$ and $S_2$ are the intersections between the line and the two planes). The sun-side of line $S_1S_2$ is defined as the front. There are three possibilities for the position of the satellite on line $S_1S_2$: (1) the satellite is in front of $S_1$; (2) the satellite is between $S_1$ and $S_2$; (3) the satellite is behind $S_2$. The first case is full phase. However, for the latter two cases, the perspective projection algorithm can be used. The keys to distinguish amongst the three possibilities is the positions of $S_1$ and $S_2$. They are determined as follows:

Let the coordinates of $P_1$ and $P_2$ be $x_{P_i}, i \in 1, 2$. Let the coordinates of $S_1$ and $S_2$ be $x_{S_i}, i \in 1, 2$. The parametric representation of $x_{S_i}$ is

$$x_{S_i} = r - t n$$  \hspace{1cm} (7)

where $t$ is the unknown parameter.

Planes $S_iP_i$ ($i \in 1, 2$) are tangent to the Earth ellipsoid and touch the Earth at $P_i$ and $P_2$, thus, the normal to the Earth at $P_i$ are perpendicular to planes $S_iP_i$. Meanwhile, line $S_1S_2$ (which is also in the direction of vector $n$) is perpendicular to planes $S_iP_i$. Therefore, the normal to the Earth at $P_i$ are parallel to vector $n$, which yields

$$Ax_{P_i} = sn$$  \hspace{1cm} (8)

where $s$ is a parameter for the parallel lines condition ($s \neq 0$). $A$ is a matrix describing the Earth ellipsoid (see equation 1).
Vector $n$ is perpendicular to the planes $S_i P_i$, therefore, $n$ is also perpendicular to the vectors from point $S_i$ to point $P_i$ that lie in plane $S_i P_i$. This yields

$$\left( x_{P_i} - x_{S_i} \right)^T n = 0, \ i \in 1, 2 \tag{9}$$

$P_i$ are on the surface of Earth ellipsoid, therefore, equation 8 can be substituted into equation 1. That yields

$$s^2 n^T A^{-1} n = 1 \tag{10}$$

There are two solutions of parameter $s$ to equation 10. They describe the positions of $P_1$ and $P_2$ separately. Thus, the positions of $P_1$ and $P_2$ can be obtained:

$$x_{P_i} = \pm s A^{-1} n \tag{11}$$

Substituting equation 11 and equation 7 into equation 9 yields

$$r^T n - |n^T n| n = \pm s n^T A^{-1} n \tag{12}$$

From equation 10, we see

$$n^T A^{-1} n = \frac{1}{s^2} \tag{13}$$

Thus, parameter $t$ can be solved by

$$t = \frac{1}{|n^T n|} \left( r^T n + \frac{1}{s} \right) \tag{14}$$

Parameters $s$ and $t$ both have two solutions which correspond to the position of $S_1$ and $S_2$. Once the positions of $S_1$ and $S_2$ are calculated, those three cases can be distinguished by the variation range of the satellite–Sun distance. We define the distance from the Sun to $S_1$ as $d_{S_1}$, the distance from the Sun to $S_2$ as $d_{S_2}$, and the distance from the Sun to the satellite as $d$. Let us image the satellite carries a camera, which is pointing towards the Sun. In the first case, the satellite cannot "see" the Earth, and it is denoted as "no image" in the following description. In the second case, the satellite only "sees" part of the Earth. Thus, it is described as "partial image". In the third case, the satellite can "see" the whole Earth, and it is denoted as "full image".

$$\begin{cases} \text{no image}, & d \leq d_{S_1} \\ \text{partial image}, & d_{S_1} < d < d_{S_2} \\ \text{full image}, & d \geq d_{S_2} \end{cases} \tag{15}$$

Note that for the first case, the satellite is in full phase. However, for the second and third cases, the eclipse state still needs to be confirmed by the following perspective projection calculation.

2.3. Perspective projection of the Earth and the Sun

The satellite is staring at the centre of the Sun, and the Sun is assumed to be a sphere in this study, thus, the perspective projection of the Sun on the image plane is always a circle. The equations for the projections of the Earth and the Sun on the image plane are worked out by treating the Earth as an ellipsoid and the Sun a sphere.

For the situation when the ellipsoid is in front of the viewpoint (the full image case), Eberly (1999) proposed a method to calculate the equation of the perspective projection of the ellipsoid. However, the approach needs to be extended to the partial image situation.

In general, the visible boundary of an ellipsoid is composed of points on the surface that have normals perpendicular to the line of sight from the satellite to the boundary points. Let $g$ be a unit vector which starts from the satellite to arbitrary boundary point (i.e. a vector in the line of sight). The coordinates of arbitrary points on the visible boundary are described in a parametric equation:

$$x_a = r + pg \tag{16}$$

where $p$ is a parameter representing the distance from the satellite to the boundary point. Because the satellite cannot be on the Earth’s surface, parameter $p \neq 0$.

Necessarily all the boundary points are on the surface of the Earth. Their coordinates also satisfy the equation of the Earth ellipsoid. Substituting equation 16 into equation 1 yields

$$p^2 g^T A g + p (g^T A r + r^T A g) + r^T A r = 1 \tag{17}$$

Equation 17 is a quadratic with respect to parameter $p$. Because the vectors in line of sight are tangent to the Earth’s surface, the line of sight vectors only have one intersection with the Earth. That means the discriminant of the quadratic (equation 17) is 0.

$$(g^T A r)^2 - g^T A (r^T A r - 1) = 0 \tag{18}$$

Equation 18 can also be expressed as

$$g^T M g = 0 \tag{19}$$

where

$$M = A r^T A - (r^T A r - 1) A$$
Equation 19 is the condition that all of the vectors in line of sight \( g \) should satisfy.

The points on the visible boundary, the projections of these points on the image plane and the viewpoint (the mass centre of the satellite) are on the line of sight. Let \( x_b \) be the coordinates of the projection of the boundary points on the image plane, then, the vector \( x_b - r \) is at the same direction as \( g \), and it should satisfy equation 19. That is

\[
(x_b - r)^T M (x_b - r) = 0 \quad (20)
\]

Because the projections are two dimensional in the ISF (see definition at section 2.1), the coordinates of the projections can also be described in the following form

\[
x_b = o + \alpha u + \beta v \quad (21)
\]

where \( u \) and \( v \) are the unit vectors of two ISF axes, \( o \) is the origin of the ISF, \( \alpha \) and \( \beta \) are the two dimensional coordinates of these projection points in ISF. Substituting equation 21 and equation 6 into equation 19 yields

\[
k_0 \alpha^2 + 2k_1 \alpha \beta + k_2 \beta^2 + k_3 \alpha + k_4 \beta + k_5 = 0 \quad (22)
\]

where

\[
\begin{align*}
k_0 &= u^T M u \\
k_1 &= u^T M v \\
k_2 &= v^T M v \\
k_3 &= -2 \gamma u^T M n \\
k_4 &= -2 \gamma v^T M n \\
k_5 &= \gamma^2 n^T M n
\end{align*}
\]

In equation 22, \( \gamma \) is the distance from the satellite to the image plane. Equation 22 describes the projection of the Earth on the image plane. This quadratic is in a general form of conical curves. Equation 22 can also be written in the following form:

\[
\chi^T B \chi + \delta^T \chi + k_5 = 0 \quad (23)
\]

where

\[
\begin{align*}
\chi &= \left( \begin{array}{c} \alpha \\ \beta \end{array} \right) \\
\delta &= \left( \begin{array}{c} k_3 \\ k_4 \end{array} \right) \\
B &= \left( \begin{array}{cc} k_0 & k_1 \\ k_1 & k_2 \end{array} \right)
\end{align*}
\]

The shape of the conical curve is determined by the determinant of matrix \( B \).

\[
\begin{align*}
\text{ellipse, } &\|B\| > 0 \\
\text{parabola, } &\|B\| = 0 \\
\text{hyperbola, } &\|B\| < 0
\end{align*}
\]

The parabola case is an instantaneous state in the variation from an ellipse to a hyperbola, it is not discussed in this study. However, neither equation 23 nor equation 22 is in its standard form of conical curves. For the ellipse and hyperbola cases, a rotation and translation are applied to equation 23 to achieve the standard form. The eigen decomposition of the real symmetric matrix \( B \) is

\[
B = Q D Q^T = Q D Q^{-1} \quad (25)
\]

where \( D \) is a two dimensional diagonal matrix whose diagonal elements are the corresponding eigenvalues (i.e. \( D = \text{diag} \{ \lambda_1, \lambda_2 \} \)), \( Q \) is an orthogonal matrix whose \( p^\text{th} \) column is the \( p^\text{th} \) eigenvector.

Let \( \phi = Q^T \chi + \frac{1}{2} D^{-1} Q^T \delta \), equation 23 becomes

\[
\phi^T D \phi = \frac{1}{4} \delta^T B^{-1} \delta - k_5 \quad (26)
\]

Equation 26 is in the standard form of conical curves and describes the shape of the projection of the Earth on the image plane.

However, for the Sun (a sphere), the projection on the image plane is a circle. The radius of the circle \( R_0 \) is given

\[
R_0 = \frac{\gamma R_s}{\|r - r_s\|} \quad (27)
\]

That is

\[
R_0 = \frac{\gamma R_s}{\|r - r_s\|} \quad (28)
\]

In order to describe the circle and the conical curve in the same coordinate system, the equation of the circle has to be translated and rotated (The rotation transformation has no effect to the equation of the circle because the origin of the coordinate system before transformation lies at the centre of the circle). After performing the same transformation as the conical curve, the equation of the circle is

\[
(\omega - \frac{1}{2} D^{-1} Q^T \delta)^T (\omega - \frac{1}{2} D^{-1} Q^T \delta) = R_0^2 \quad (29)
\]

where \( \omega \) is 2-vector of coordinates of any point on the projected circle after the transformation.

2.4. Determining the intersections

The intersections between projections of the Earth and the Sun can be determined by combining their equations together. Equation 26 and equation 29 form two simultaneous equations.

\[
\begin{align*}
(x - \frac{1}{2} D^{-1} Q^T \delta)^T (x - \frac{1}{2} D^{-1} Q^T \delta) &= R_0^2 \\
x^T Dx &= \frac{1}{4} \delta^T B^{-1} \delta - k_5
\end{align*}
\]

where \( x \) contains the coordinates of the intersections between the circle and the conical curve.

The above equation 30 is expressed in the elements of the vectors.

\[
\begin{align*}
(x - \frac{1}{2} \lambda_1)^2 + (y - \frac{1}{2} \lambda_2)^2 &= R_0^2 \\
(\lambda_1 x^2 + \lambda_2 y^2) &= \Omega - k_5
\end{align*}
\]

(31)
where \( \Omega = \frac{k_2 t_x^2 - 2k_1 k_3 k_4 k_5}{4(k_0 k_2 - t_x^2)} \); \( D^{-1} Q^T \delta = (t_x, t_y)^T \); \( x = (x, y)^T \).

In order to solve equation 31, variables \( x \) and \( y \) are related to the unknown parameter \( \eta \):

\[
\begin{align*}
\frac{1}{1 + \eta^2} R_0 + \frac{1}{2} t_x & = x \\
\frac{2\eta}{1 + \eta^2} R_0 + \frac{1}{2} t_y & = y
\end{align*}
\] (32)

Substitute equation 32 into equation 31, a quartic with respect to parameter \( \eta \) is obtained. The values of \( x \) and \( y \) are determined by solving for \( \eta \).

\[
a \eta^4 + b \eta^3 + c \eta^2 + d \eta + e = 0 \] (33)

where

\[
\begin{align*}
a & = \lambda_1 (R_0 - \frac{1}{2} t_x)^2 + \frac{1}{4} \lambda_2 t_y^2 - \Omega + k_5 \\
b & = 2 \lambda_2 R_0 t_y \\
c & = \lambda_1 \left( \frac{1}{2} t_x^2 - 2R_0^2 \right) + \lambda_2 \left( \frac{1}{2} t_y^2 + 4R_0^2 \right) - 2\Omega + 2k_5 \\
d & = 2 \lambda_2 R_0 \eta \\
e & = \lambda_1 (R_0 + \frac{1}{2} t_x)^2 + \frac{1}{4} \lambda_2 t_y^2 - \Omega + k_5
\end{align*}
\]

The quartic can be solved with Ferrari’s method. The detailed algorithm can be found in Nonweiler (1968). Once parameter \( \eta \) is obtained, the coordinates of the intersections between the Sun’s image and the Earth’s image can be computed using equation 32. Solutions of the quartic will yield 4 roots (complex or real), only the real solutions are chosen.

2.5. Calculate the overlapping area

In the penumbra case, part of the solar disk is blocked by the Earth’s image. The total area of the solar disk is split into the bright part (denoted as \( A_b \)) and the shaded part (denoted as \( A_s \)). i.e.

\[
A_b + A_s = \pi R_0^2
\] (34)

The area calculation is different for the hyperbolic and elliptical image of the Earth. For both the ellipse case and the hyperbola case, the calculation processes are separated according to the condition that if the centre of the Sun’s image is inside the Earth’s image or not. The condition that the Sun’s image is inside the Earth’s image is given by

\[
\begin{align*}
\text{ellipse:} & \quad \frac{\Omega}{\Omega - k_5} \leq 1 \\
\text{hyperbola:} & \quad \frac{\Omega}{\Omega - k_5} > 1
\end{align*}
\] (35)

The centre of the circle is on the ellipse or the hyperbola when \( \frac{\Omega}{\Omega - k_5} = 1 \).

Figure 5 and Fig. 6 show the two situations for the ellipse and hyperbola cases separately. The area of the bright
part $A_{b,e,in}$ and the area of the shadow part $A_{s,e,in}$ of the ellipse situation in Fig. 5(a) are computed by
\[
\begin{align*}
A_{b,e,in} &= S_{Q_1 Q_2 O_s} - S_{\triangle O_s Q_1 Q_2} - S_1 \\
A_{s,e,in} &= \pi R_0^2 - A_{b,e,in}
\end{align*}
\tag{36}
\]
where $S_{Q_1 Q_2 O_s}$ : the area of the circular sector $Q_1 Q_2 O_s$, $S_{\triangle O_s Q_1 Q_2}$ : the area of the triangle $O_s Q_1 Q_2$, $S_1$ : the area of the elliptical arch as shown in Fig. 5. Its calculation method can be found in Hughes and Chraibi (2012). The area of the bright part $A_{b,e,out}$ and the area of the shadow part $A_{s,e,out}$ of the ellipse situation in Fig. 5(b) are computed by
\[
\begin{align*}
A_{b,e,out} &= \pi R_0^2 - A_{s,e,out} \\
A_{s,e,out} &= S_{Q_1 Q_2 O_s} - S_{\triangle O_s Q_1 Q_2} + S_1
\end{align*}
\tag{37}
\]
The area of the bright part $A_{b,h,in}$ and the area of the shadow part $A_{s,h,in}$ of the hyperbola situation in Fig. 6(a) are computed by
\[
\begin{align*}
A_{b,h,in} &= S_{Q_1 Q_2 O_s} - S_{\triangle O_s Q_1 Q_2} - S_2 \\
A_{s,h,in} &= \pi R_0^2 - A_{b,h,in}
\end{align*}
\tag{38}
\]
where $S_{Q_1 Q_2 O_s}$ and $S_{\triangle O_s Q_1 Q_2}$ have the same meaning as equation 36, $S_2$ is the area of elliptical arch as shown in Fig. 6.
The area of the bright part $A_{b,h,out}$ and the area of the shadow part $A_{s,h,out}$ of the hyperbola situation in Fig. 6(b) are computed by
\[
\begin{align*}
A_{b,h,out} &= \pi R_0^2 - A_{s,h,out} \\
A_{s,h,out} &= S_{Q_1 Q_2 O_s} - S_{\triangle O_s Q_1 Q_2} + S_2
\end{align*}
\tag{39}
\]

2.6. Atmospheric effects and shadow function

The atmospheric effects cause a reduction in the solar radiation at a satellite’s location. There exists a complex physical model (including atmospheric refraction, scattering and absorption) to describe the atmospheric effects on the shadow function (Robertson, 2015). The interactions between the solar radiation and the atmosphere above the stratosphere (at about 50 km altitude) are negligibly small for SRP modelling (Robertson, 2015). In this study, we provide a simple linear function to model the solar radiation reduction in the atmosphere. We assume the TOA (Top Of Atmosphere) is an ellipsoid that wraps the solid Earth. The relative positions between the projections of the Sun, the Earth and the TOA are known by applying the perspective projection algorithm to the solid Earth and the TOA ellipsoid separately.

A function $f(h)$ that describes the radiation reduction coefficient is defined. This function computes the ratio between the solar radiation at a depth $h$ in the atmosphere and the radiation outside of the atmosphere. The thickness $H_0$ (as shown in Fig. 7) of the atmosphere on the image is different depending on the position of the satellite. For any point that has a distance $h$ to the solid Earth in the direction from the centre of the Earth $O_e$ to the centre of the Sun $O_s$ on the image, $f(h)$ is given by
\[
f(h) = (\mu_2 - \mu_1) \frac{h}{H_0} + \mu_1 \tag{40}
\]
where $\mu_1$ is the radiation reduction coefficient at the boundary of the solid Earth and $\mu_2$ is the radiation reduction coefficient at the boundary of the atmosphere. Thus, $\mu_1 = 0$ and $\mu_2 = 1$.

Figure 7: The relative positions of the projections for the Sun, the Earth and the TOA. In this figure, the Sun’s image is totally inside the atmosphere. $O_s$ is the centre of the Earth’s image, $O_e$ is the centre of the Sun’s image. The point (with a distance of $h$ to the boundary of the solid Earth’s image) that for the calculation of the radiation reduction lies exactly at the centre of the Sun’s image, but it could be anywhere (inside the atmosphere’s image) on line $O_s O_e$. $H_0$ is the thickness of the atmosphere’s image.

Figure 8 shows 5 possible relative positions between the solar disk, the image of the Earth, and the image of the atmosphere (Vokrouhlicky et al, 1993, 1994). Case a shows the full phase with a shadow function value 1.0 while case e shows the umbra with a shadow function value 0.0.

Cases b, c and d are considered to be in the penumbra, and the details are shown in Fig. 9. $G_1$ and $G_2$ are the points on the lower boundary and upper boundary of the Sun’s image in the atmosphere separately. The average solar radiation reduction coefficient $f(h_G)$ for the area of the solar disk that is in the atmosphere is the average of radiation reduction coefficient at $G_1$ and $G_2$. That is
\[
f(h_G) = \frac{f(h_{G_1}) + f(h_{G_2})}{2} \tag{41}
\]
where $h_{G_1}$ and $h_{G_2}$ are the distances from the boundary of the solid Earth’s image to the point $G_1$ and $G_2$ separately.

The shadow function $F_s$ is a ratio between the solar radiation flux in the penumbra and that in full phase. Because the shadow function in full phase is defined to be 1.0, the solar radiation flux in full phase can be represented by $\pi R_0^2$ in this study. The calculation of the radiation flux that goes through the solar disk in penumbra is calculated by the average radiation reduction coefficient of the solar disk times the area that is inside the atmosphere.

**case a**
The satellite can see the full solar disk and the solar disk is not blocked by the image of the atmosphere. The shadow function is
\[
F_s = 1 \tag{42}
\]
case b
In case b, part of the solar disk has entered the atmosphere’s image as shown in Fig. 9. The area of the solar disk that is inside the atmosphere is $A_{s,b}$, and the area outside of the atmosphere’s image is $A_{b,b}$. For the area inside the atmosphere, the solar radiation flux is $f(h_{G_1})A_{s,b}$. For the area outside of the atmosphere, the solar radiation flux is $A_{b,b}$. In this case, because $G_2$ is at the boundary of the atmosphere, the solar radiation reduction coefficient at $G_2$ is $\mu_2=1$. The shadow function is calculated by

$$F_s = \frac{0.5(f(h_{G_1}) + \mu_2)A_{s,b} + A_{b,b}}{\pi R_0^2}$$ (43)

case c
In case c, the whole solar disk is inside the atmosphere’s image. The shadow function is computed by

$$F_s = \frac{f(h_{G_1}) + f(h_{G_2})}{2}$$ (44)

case d
In case d, part of the solar disk is blocked by the solid Earth’s image, part of the solar disk is inside the atmosphere’s image. The area of the solar disk that is blocked by the Earth’s image is $A_{s,d}$ while the area that is in the atmosphere is $A_{b,d}$. In Fig. 9, $G_1$ lies on the boundary of the solid Earth’s image, the solar radiation reduction coefficient is $\mu_1=0$. Thus, the shadow function is computed by

$$F_s = \frac{0.5(f(h_{G_2}) + \mu_1)A_{s,d}}{\pi R_0^2}$$ (45)

case e
In case e, the whole solar disk is blocked by the solid Earth, the shadow function is given

$$F_s = 0$$ (46)

2.7. Algorithm flow and source code publication

The algorithm flow is given to implement the PPM_atm conveniently. Figure 10 describes the whole algorithm of the model PPM_atm. The inputs of the model are the position of the Sun and the satellite, and the output is the shadow function value ranging between 0 and 1. The core parts of the model are the perspective projection and the atmospheric effects (the 5 cases in the algorithm). Figure 11 shows the flow chart of the perspective projection algorithm. The outputs of the perspective projection are the area of the solar disk that is not blocked by the Earth’s image and the coordinates of the intersections between a circle and the conical curve. The perspective projection algorithm is applied to both the TOA and the Earth ellipsoid (the Sun is a sphere, the perspective projection of the Sun is a circle, and it does not need this perspective projection).

In order to make a better contribution to the community, the working code of the model PPM_atm is published on github at this link: [http://github.com/whulizhen/PPM_atm_shadow_function.git](http://github.com/whulizhen/PPM_atm_shadow_function.git).
3. Model Validation

The model validation applies two methods. In the first, the modelled shadow function is compared to the GRACE accelerometer measurements—essentially as a consistency check, although we treat the GRACE data as the “truth”. Secondly, we consider the impacts of the model on Galileo orbit prediction. The GRACE spacecraft (Tapley et al., 2004) carry accelerometers to record the non-gravitational forces and thruster events. The transition from full phase to umbra causes variations in SRP on the GRACE spacecraft, thus, the accelerometer measurements can reveal the variation of the shadow function. GRACE level-1B products record all the available measurements of the accelerometers (Case et al., 2010). Therefore, the level-1B accelerometer data (calibrated using the a priori calibration parameters) (Bettadpur, 2009) are used as a baseline to validate the accuracy of the shadow function.

In the tests, we introduce 5 different shadow function models (the PPM, the PPM\textsubscript{atm}, the SECM, the SECM\textsubscript{atm}, and the shadow function derived from the accelerometer observations). The PPM and the PPM\textsubscript{atm} are developed in this study. The only difference between them is that PPM does not contain the atmospheric effects. The SECM\textsubscript{atm} is modified from the SECM which considers the atmospheric effects but using spherical Earth model.

The impacts of the shadow function models on the orbit of the Galileo satellites are shown in orbit prediction during eclipse seasons. In precise orbit determination, the errors in the force models may get absorbed by other estimated parameters. The orbit prediction is sensitive to the force models. Thus, we avoid using any empirical parameters in orbit prediction.

3.1. The validation of the shadow function

The acceleration derived shadow function $F_{s,\text{acc}}$ is obtained by normalising the accelerometer observations with equation 47. In this way, the shadow function $F_{s,\text{acc}}$ gets 0 at umbra while gets 1 at full phase.

$$F_{s,\text{acc}} = \frac{\zeta_t - \zeta_0}{\zeta_1 - \zeta_0}$$  \hspace{1cm} (47)

where $\zeta_t$ is the acceleration at time $t$, $\zeta_0$ is the acceleration at umbra and $\zeta_1$ is the acceleration at full phase.

There are 16 eclipse events (30 penumbra transitions) on Jan 20th, 2007 for GRACE-A spacecraft as shown in Tab. 2. All 30 penumbra transitions contain clear signals of shadow function and are used to produce $F_{s,\text{acc}}$. Figure 12 and Fig. 13 show the comparisons of different shadow function models for one eclipse event on Jan 20th, 2007 for the GRACE-A spacecraft. The SECM assumes the Earth is a sphere (with radius 6371 km) while the PPM considers the oblateness of the Earth, and uses the WGS-84 ellipsoid to describe the Earth. The shadow function models with atmospheric effects are tagged as “PPM\textsubscript{atm}” and “SECM\textsubscript{atm}”. All the different shadow function models are compared with $F_{s,\text{acc}}$, which is derived from accelerometer measurements of GRACE-A satellite.

Table 3 shows the statistics of the differences between different shadow function models and $F_{s,\text{acc}}$ for the 30 penumbra
Figure 12: The variations of shadow function from full phase to umbra for GRACE-A satellite on Jan 20th, 2007. The acceleration derived shadow function \( F_{\text{acc}} \) is tagged as “acc”. The “PPM” model considers Earth’s oblateness while the “SECM” uses a spherical Earth. “PPM\(_{\text{atm}}\)” and “SECM\(_{\text{atm}}\)” both consider the atmospheric effect.

Figure 13: The variations of shadow factor from umbra to full phase for GRACE-A satellite on Jan 20th, 2007.

Table 2: The penumbra entry/exit time (in GPS time) for GRACE-A satellite on 20th January, 2007 (time format: H:M:S).

<table>
<thead>
<tr>
<th>event #</th>
<th>entry1</th>
<th>exit1</th>
<th>entry2</th>
<th>exit2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01:27:25</td>
<td>01:27:55</td>
<td>02:03:21</td>
<td>02:03:51</td>
</tr>
<tr>
<td>2</td>
<td>03:01:24</td>
<td>03:01:53</td>
<td>03:37:22</td>
<td>03:37:50</td>
</tr>
<tr>
<td>3</td>
<td>04:35:24</td>
<td>04:35:52</td>
<td>05:11:21</td>
<td>05:11:49</td>
</tr>
<tr>
<td>4</td>
<td>06:09:22</td>
<td>06:09:55</td>
<td>06:45:20</td>
<td>06:45:48</td>
</tr>
<tr>
<td>6</td>
<td>09:17:22</td>
<td>09:17:49</td>
<td>09:53:19</td>
<td>09:53:46</td>
</tr>
<tr>
<td>11</td>
<td>17:07:17</td>
<td>17:07:45</td>
<td>17:43:14</td>
<td>17:43:42</td>
</tr>
</tbody>
</table>

Table 3: The error statistics of different shadow function models based on the comparison with \( F_{\text{acc}} \) for GRACE-A (30 penumbra transitions)

<table>
<thead>
<tr>
<th>Shadow function model</th>
<th>mean</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPM</td>
<td>0.2314</td>
<td>0.2417</td>
</tr>
<tr>
<td>PPM(_{\text{atm}})</td>
<td>-0.0079</td>
<td>0.0220</td>
</tr>
<tr>
<td>SECM</td>
<td>0.1107</td>
<td>0.1741</td>
</tr>
<tr>
<td>SECM(_{\text{atm}})</td>
<td>-0.1350</td>
<td>0.0899</td>
</tr>
</tbody>
</table>

transitions. As shown in Tab. 3, if the \( F_{\text{acc}} \) is assumed to be the “truth”, the error of the SECM has an average of 0.1107 and standard deviation of 0.1741, the error of the PPM has an average of 0.2314 and standard deviation of 0.2417, the error of the SECM\(_{\text{atm}}\) has an average of -0.1350 and standard deviation of 0.0899, the errors of the PPM\(_{\text{atm}}\) has an average of -0.0079 and standard deviation of 0.022. This indicates that PPM\(_{\text{atm}}\) model outperforms the closest competitor (the SECM) by approximately a factor of 10 in the mean, a factor of 8 in the standard deviation.

There are no observations for the shadow function because the Galileo satellites do not carry accelerometers. However, we can compare between the SECM, the SECM\(_{\text{atm}}\), the PPM, and the PPM\(_{\text{atm}}\) model. Figure 14 and Fig. 15 show the variation of shadow function for the Galileo E11 satellite during one eclipse event on January 11th, 2015.

There are 18 eclipse events (36 penumbra transitions) for Galileo E11 satellite from 6th to 16th January, 2015. Results from GRACE-A show that the PPM\(_{\text{atm}}\) is the closest to the accelerometer observations, it is reasonable to treat the PPM\(_{\text{atm}}\) as the “truth”. The error statistics of the SECM, the PPM, and the SECM\(_{\text{atm}}\) for these 36 penumbra transitions are shown in Tab. 4. The error of the SECM has an average of 0.0528 and standard deviation of 0.0306; the error of the PPM has an average of 0.0657 and standard deviation of 0.0374; the error of the SECM\(_{\text{atm}}\) has an average of -0.0131 and standard deviation of 0.0071.

Figure 14: The variations of shadow function from full phase to umbra for Galileo E11 satellite on Jan 11th, 2015.
Table 4: The error statistics of different shadow function models based on the comparison with the PPM_atm for Galileo E11 (with 36 penumbra transitions)

<table>
<thead>
<tr>
<th>Shadow function model</th>
<th>mean</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPM</td>
<td>0.0657</td>
<td>0.0374</td>
</tr>
<tr>
<td>SECM</td>
<td>0.0528</td>
<td>0.0306</td>
</tr>
<tr>
<td>SECM_atm</td>
<td>-0.0131</td>
<td>0.0071</td>
</tr>
</tbody>
</table>

3.2. The impact on orbit prediction

In this study, the impacts of different shadow function models on the orbit are shown by orbit prediction. The Galileo satellites in 3 different orbital planes are chosen as test objects during their eclipse seasons in 2016. The tested shadow function models include the PPM, the SECM, and the PPM_atm.

In the orbit prediction tests, the force models used include SRP (Li et al, 2018), TRR (Thermal Re-radiation Pressure), ERP (Earth Radiation Pressure) (Li et al, 2017) and Antenna Thrust (the transmitted power is found in Steigenberger et al (2018)). The nominal attitude is applied during eclipses. All the tests use the same attitude law during eclipses, only the shadow function model is different.

Precise orbit products from the CODE (Centre for Orbit Determination in Europe), which is one of the International GNSS service analysis centres (Dow et al, 2009) are used as a reference. The initial state (position and velocity) is obtained by orbit fitting using the first 2 hours’ precise orbit. No empirical parameters are used in the orbit fitting. After obtaining the initial position and velocity in the orbit fitting, the orbit prediction is performed for the next 1 day, 3 days, and 7 days. These “orbit fitting and prediction” processes are performed over the eclipse seasons during 2016. Table 5 shows the satellites information and time span used in the data processing.

Table 6 shows the RMS (Root Mean Square error) in radial orbit error after 7 days’ prediction with the SECM, the PPM and the PPM_atm [m]

<table>
<thead>
<tr>
<th>Satellites</th>
<th>SECM</th>
<th>PPM</th>
<th>PPM_atm</th>
</tr>
</thead>
<tbody>
<tr>
<td>E11</td>
<td>1.984</td>
<td>1.991</td>
<td>1.931</td>
</tr>
<tr>
<td>E22</td>
<td>2.196</td>
<td>2.206</td>
<td>2.144</td>
</tr>
<tr>
<td>E08</td>
<td>1.396</td>
<td>1.407</td>
<td>1.320</td>
</tr>
<tr>
<td>E19</td>
<td>3.731</td>
<td>3.742</td>
<td>3.665</td>
</tr>
<tr>
<td>E24</td>
<td>1.212</td>
<td>1.217</td>
<td>1.182</td>
</tr>
<tr>
<td>E30</td>
<td>2.342</td>
<td>2.351</td>
<td>2.286</td>
</tr>
<tr>
<td>average</td>
<td>2.144</td>
<td>2.152</td>
<td>2.088</td>
</tr>
</tbody>
</table>

Both Fig. 16 and Fig. 17 show that the orbital errors in the cross track due to the mismodelling of the shadow function are negligibly small, even the 7-day’s orbit prediction RMS in the cross track is less than 1 mm. This is due to the fact that the Sun elevation over orbital plane is low during eclipse seasons, which makes the SRP mainly in the radial and along track directions.

4. Discussion and Conclusions

This study presents an improved shadow function considering both the Earth’s oblateness and atmospheric effects. Perspective projection is used in combination with an ellipsoidal Earth model to solve the geometrical component of the problem. A linear reduction of the solar radiation due to the atmospheric effect is applied and combined with the perspective projection approach to produce a shadow function for Earth-orbiting satellites during eclipses.

A shadow function model enables the user to predict both the time and the magnitude of solar radiation pressure induced accelerations during penumbra transitions. Tests using the GRACE-A accelerometer dataset show the PPM_atm follows the pattern value in Fig. 16 means the RMS of the PPM is larger than that of PPM_atm. The difference in the RMS of 1-day, 3-day, 5-day and 7-day’s orbit prediction error between PPM and PPM_atm are 3.26 mm, 22.0 mm, 36.57 mm, 64.81 mm in the radial direction, 6.12 mm, 40.85 mm, 113.99 mm and 224.03 mm in the along track separately.

Figure 17 shows the difference in the RMS of orbit prediction error between the SECM and the PPM. The positive value in Fig. 17 shows the RMS of the PPM is larger than that of the SECM. The difference in the RMS of 1-day, 3-day, 5-day and 7-day’s orbit prediction error between PPM and SECM are 1.28 mm, 4.35 mm, 7.16 mm, 8.11 mm in the radial direction, 2.57 mm, 38.60 mm, 89.96 mm and 158.31 mm in the along track separately.
of measured accelerations more closely than those predicted by existing conventional approaches. The improvement in accuracy is essentially a factor of 10 (see Table 3). Although a simple atmospheric effect model is included in this study, its benefit to improve the accuracy of the shadow function has been proved. Comparing with the complex physical model developed and validated for GRACE satellites (Robertson, 2015), the PPM_atm in this study is simple and easier to implement.

Tests on both the GRACE-A and Galileo satellites show the greater improvement in accuracy derives from including the atmospheric effects as opposed to the effects of the Earth’s oblateness. The orbit prediction results are quantified and visualised in Fig. 16 and Fig. 17. The improvements are consistent over 1, 3, 5, and 7 day predicted arcs.

The errors due to ignoring the atmospheric effects are larger for the GRACE-A than the Galileo satellites. The errors in the PPM reveal the atmospheric effects. The PPM has 23.14% and 6.57% errors for the GRACE-A satellite and the Galileo satellites respectively as shown in Tab. 3 and Tab. 4. This is due to the fact that the GRACE satellite has a lower altitude (around 500 km) than the Galileo satellites (around 23,000 km).

One interesting point is that the PPM is worse than the SECM if the atmospheric effect is not included. This is probably due to the cancellation of the errors caused by the Earth’s oblateness and the atmospheric effects. This hypothesis can be supported by the analyses of the orbit prediction and the shadow function model comparisons shown in Tab. 3 and Tab. 4. The average error of the SECM is smaller than that of the PPM. The PPM model computes the Earth’s oblateness in rigorous mathematics but does not include the atmospheric effects. The SECM ignores the atmospheric effects and also has errors in modelling the Earth’s oblateness. The errors due to ignoring the Earth’s oblateness can cancel the errors due to ignoring the atmospheric effect. This is the reason why the SECM has better accuracy than the PPM.

This paper focuses on the shadow function modelling due to solar eclipses where the Sun is blocked by the Earth. The newly developed PPM_atm accounts for both the Earth’s oblateness and the atmospheric effects. It should be noted that the modelling method can be applied generally and it could be used to deal with a multi-body eclipse situation. In further work, it will be worth investigating the impacts of PPM_atm on the accuracy of different types of orbit. In geodesy, the solution to geodetic parameters usually rely on cm-level orbit products, the errors in the orbit will affect the accuracy of the geodetic parameters. It will also be worth exploring the impact of PPM_atm on these geodetic parameters.

Acknowledgements

This work is jointly funded by University College London and China Scholarship Council. The IGS (International GNSS services) and the GRACE mission are greatly acknowledged for providing the precise orbit and accelerometer observations. The authors thank four anonymous reviewers for their constructive comments. Special thanks to Dr Paul Groves for the careful review and constructive suggestions.