Forward modeling will never be perfect. The shape model will never be able to account for every wrinkle in the MLI coating, and similarly, the surface model as represented by the bidirectional reflectance distribution function (BRDF) will never be identical to that found on actual space objects due to space weathering and inherent differences in the same material from lot to lot. The key to making estimation filters work is thus to continually improve the shape and surface models of the space object, and to intelligently and completely account for all remaining error sources in the models and how these manifest in the observations. This paper uses a multiple hypothesis filtering scheme to determine the attitude of a model geosynchronous satellite with possible surface modeling errors to compare the estimates made when each hypothesis is either propagated by an unscented Kalman filter (UKF) or an unscented Schmidt Kalman filter (USKF). It is shown that the USKF, which incorporates the influence of the known modeling errors in so-called “consider” terms, provides a better estimate and a more realistic uncertainty on the final attitude estimate.

1. INTRODUCTION

The utility of using brightness (radiometric flux intensity) measurements to determine a space object (SO)’s attitude and shape has been long established [1]. Ref. [2] first demonstrated how these brightness measurements can also be used in an estimation filter, such as the unscented Kalman filter (UKF) [3,4], to estimate the attitude and angular rates of a SO. Multi-bandpass observations, fusion with angles measurements, multiple hypothesis testing, and simultaneous estimation with position, velocity, attitude, angular rates, surface parameters and mass have also been demonstrated [5,6,7]. Modeling uncertainties, however, introduce a systematic error in the UKF’s resulting final state. A process noise can be introduced into the UKF to mitigate the effects of modeling uncertainty, but the required level of process noise necessary is not known a-priori, given the optimal process noise is possibly variable with time. For example, an uncertainty in the surface reflectance of a particular component of the SO will only affect the filter’s performance when the attitude and observing conditions are such that the component is illuminated and visible to the observer. The unscented Schmidt Kalman filter (USKF) [8,9] instead utilizes the known modeling uncertainties within the construct of the filter itself.

2. DETAILS OF APPROACH

In this paper, the UKF [3,4] with no process noise is compared to the USKF [8,9]. In both filters, a number of parameters associated with the scenario are estimated and contained in a state vector. The state vector can consist of only those parameters being estimated, as in the UKF, or also include so-called “consider” parameters, as in the USKF.

\[
\hat{x}_{\text{UKF}} = \begin{bmatrix} x^1 & \ldots & x^n \end{bmatrix}^T
\]

\[
\hat{z}_{\text{USKF}} = \begin{bmatrix} x^1 & \ldots & x^n & c^1 & \ldots & c^n \end{bmatrix}^T
\]

where \( x \) are the \( n_x \) estimated parameters, and \( c \) are the \( n_c \) considered parameters and \( \hat{z} \) is used to distinguish the USKF’s state that includes the consider parameters from the UKF’s state that does not. The state has an associated covariance.

\[
P_{\text{UKF}} = P_{\text{ss}}
\]
\[
P_{USKF} = P = \begin{bmatrix} P_{xx} & P_{xc} \\ P_{cx} & P_{cc} \end{bmatrix}
\]  
(2b)

In both filters, the state parameters are propagated forward in time with a dynamics, or state, equation.

\[
\hat{x}_k = f(\hat{x}_{k-1}, w_k)
\]  
(3a)

\[
\hat{z}_k = f(\hat{x}_{k-1}, v_k)
\]  
(3b)

where \(w_k\) is the process noise vector and (3) shows the state going from the a posteriori state at time step \(k-1\) to the a priori state at time step \(k\). The general measurement function used in both filters is

\[
\hat{y}_k = h(\hat{x}_k, v_k)
\]  
(4a)

\[
\hat{y}_k = h(\hat{z}_k, v_k)
\]  
(4b)

where \(v_k\) is the measurement noise vector and \(y\) is a vector representing the \(m\) observations.

The UKF and USKF proceed identically with the generation of sigma points, propagation forward in time using (3), generation of the predicted observations using (4), and comparison to the actual observations. Whereas the UKF has \(2n+1\) sigma points, the USKF has \(2(n+n_1)+1\) sigma points. The significant difference occurs in the update step. The state update

\[
\hat{x}^i_k = \hat{x}_k + K_x(\bar{y}_k - \hat{y}_k)
\]  
(5a)

\[
\hat{z}^i_k = \begin{bmatrix} \hat{x}^i_k \\ \hat{c}^i_k \end{bmatrix} + \begin{bmatrix} K_x \\ 0 \end{bmatrix}(\bar{y}_k - \hat{y}_k)
\]  
(5b)

does not update the consider parameters in the USKF (5b), while similarly the covariance update

\[
P_{xx}^{i+} = P_{xx}^{i-} + K_x P_{yx} K_x^T
\]  
(6a)

\[
P_{xx}^{i-} = \begin{bmatrix} P_{xx}^{i+} & P_{xc}^i \\ P_{cx}^i & P_{cc} \end{bmatrix} - \begin{bmatrix} K_x P_{yx} & K_x P_{yx} K_x^T \\ K_x P_{yx} K_x^T & 0 \end{bmatrix}
\]  
(6b)

does not update the portion of the covariance associated with the consider parameters (6b). The effect of the consider parameters are associated with the influence of the \(2n-1\) extra sigma points on the calculation of \(P_{yy}, P_{xy}\), and \(K_x\) in the filter. It is these sigma points that were generated using the uncertainty in the consider parameters. In (5) and (6), the Kalman gain \(K\) for the USKF is represented by

\[
K_z = \begin{bmatrix} K_x \\ K_c \end{bmatrix}
\]  
(7)

and the actual observations at time step \(k\) is \(\bar{y}_k\).

In this paper, the UKF and USKF are first compared directly, and then compared using multiple hypotheses with different initial states processed simultaneously and weighted according to how well they predict the observations. Each hypotheses is a separate instantiation of the UKF or USKF. Generally, this architecture is that of the Unscented Particle Filter (UPF) first proposed by van der Merwe et al. [10]. As such, “particles” and “hypotheses” represent the same thing. The resampling step of the UPF is not utilized, and thus this UPF implementation is essentially identical to Multiple Model Adaptive Estimation (MMAE) [11,12]. The conditional probability for the i-th hypothesis and k-th time step is calculated using

\[
p_k^i = (2\pi)^{m/2} |P_{yy,k}^i|^{1/2} e^{-1/2(y_i^i - \bar{y}_k)^T P_{yy,k}^{-1} (y_i^i - \bar{y}_k)}
\]  
(8)

The likelihood function that expresses the weight for the i-th hypothesis at the k-th time step is then
\[ w_k^i = \frac{p_k^i}{\sum_{i=1}^{N} p_k^i} \]  

(9)

where \( N \) is the total number of hypotheses. It is the final composite state and covariance, which are a weighted composite of all hypotheses that are compared to the truth and with each other.

3. DETAILS OF SCENARIO

The scenario used in this paper is that of a satellite at geosynchronous Earth orbit (GEO) whose attitude is such that its primary axis (z-axis) is pointing at a particular latitude and longitude on the surface of the Earth, and its secondary axis (x-axis) is pointing as close to the velocity vector of its orbit as possible. The model satellite has solar panels that project along the third direction (±y-axis). The solar panels have a ±5° E-W tilt from exact primary axis pointing, and for the sake of simplicity, do not track the Sun. The model satellite is shown in Fig. 1.

![Model Satellite](image)

Figure 1 – Model Satellite

The satellite is placed in the near-geostationary orbit of SSN27715 using the historic two-line element (TLE) from 2011 Dec 29 to generate the 6DOF position/velocity at 2012 Jan 01 06:00:00 UT. This corresponds to \( \ell = -140^\circ \), and three different truth values will be examined corresponding to the satellite pointing at French Frigate Shoals (\( \lambda = 23.7^\circ, \ell = -166.1^\circ \)), Pago Pago, American Samoa (\( \lambda = -14.3^\circ, \ell = -170.7^\circ \)), and Monument, Colorado (\( \lambda = 39.1^\circ, \ell = -104.8^\circ \)).

For simplicity, all surfaces of the model satellite are represented by the same Bidirectional Reflectance Distribution Function (BRDF). The specific BRDF is the Cook-Torrance BRDF [13] with microfacet slope set to 0.15, and reflectance (both diffuse and specular) set to 0.6 and the diffuse fraction set to 0.3 for both wavebands. Details of this BRDF model and how it compares to the Ashikhmin-Shirley BRDF [14] using similar nomenclature can be found in Ref. [15]. In this scenario, the state vector of (1) is

\[
\hat{x}_{UKF} = \begin{bmatrix} g_1 & g_2 & g_3 \end{bmatrix}^T
\]

(10a)

\[
\hat{z}_{USKF} = \begin{bmatrix} g_1 & g_2 & g_3 & p_m & p_d & p_d & p_d \end{bmatrix}^T
\]

(10b)

where \( g \) is a three parameter set of generalized Rodrigues parameters (GRPs) that define the local attitude error of the sigma points from a quaternion \( q \) that defines the global mean attitude at which the satellite is pointing [16], \( p_m \) is an unconstrained proxy value for the microfacet slope, \( p_d \) is an unconstrained proxy value for the reflectance and \( p_d \) is an unconstrained proxy for the diffuse fraction. The reflectance and diffuse fraction have different values for the two different wavebands (B and V). Proxy values are used to ensure the sigma points generated by the unscented transform within the filter remain physical and are calculated using

\[ p_m = \ln(m), \quad m = \exp(p_m) \]  

(11a)
\[
\rho = \frac{1}{2} \ln \left( \frac{\rho}{1 - \rho} \right), \quad \rho = \frac{1}{2} (\tanh(p_\rho) + 1) \\
\]

\[
p_\rho = \frac{1}{2} \ln \left( \frac{\rho}{1 - \rho} \right), \quad \rho = \frac{1}{2} (\tanh(p_\rho) + 1) \\
\]

The angular rates for each sigma point are set to maintain an attitude that is pointing at a single point on the Earth’s surface. Note also that it is the global quaternion that is saved from step to step to represent the attitude and not the GRP. The initial uncertainties used to build the initial covariance are \(2^\circ\) for Roll and Pitch in the equivalent Euler angle attitude representation, which corresponds to approximately \(11^\circ\) in latitude and longitude on the surface of the Earth, and \(0.2^\circ\) for Yaw in the equivalent Euler angle attitude representation, representing how close the x-axis is pointing, and eight different uncertainties in the BRDF parameters corresponding to reflectance and diffuse fraction uncertainty values of 0.001, 0.025, 0.050, 0.075, 0.100, 0.125, 0.150, and 0.175 and a corresponding uncertainty in the microfacet slope. For a given case, the BRDF parameters used in the measurement model is not the truth but consistent with the truth and associated uncertainties.

The satellite is observed in two wavebands from the location of the Advanced Electro-Optic System (AEOS) on Maui, Hawaii for forty-nine measurements starting at 06:00:00 UT with 600s time step (8 hours). The measurement vector is

\[
\ddot{y} = \begin{bmatrix} B & V \end{bmatrix}
\]

Two different measurement noise values are used corresponding to 0.05 and 0.10 magnitudes respectively. The truth light curves for the B and V wavebands for the 0.050 reflectance/diffuse fraction uncertainty are shown in Fig. 2. The total number of cases examined is 3 (pointing attitudes) \(\times\) 8 (reflectance/diffuse fraction uncertainties) \(\times\) 2 (measurement noise values) = 48 cases, and each case has its own set of truth light curves.

![Figure 2](image_url)

Three analyses are accomplished. In the first, a single case is chosen for a Monte Carlo analysis using random initial values for the attitude, and surface parameters consistent with the truth and uncertainty. This is the direct comparison of a single instantiation of the UKF and USKF. In the second, for each of the 48 cases, the UKF and
USKF are run with the same truth measurements and the same eighteen initial hypotheses. Each hypothesis corresponds to different attitudes pointing at the latitudes and longitudes as shown in Fig. 3 and the same initial reflectance/diffuse fraction values consistent with the uncertainty. This is the multiple hypothesis testing scenario.

In the final analysis, the previous multiple hypothesis testing scenario is repeated, but with a different BRDF model used as the measurement function in the estimation filter to that which created the truth observations. Specifically, the Ashikhmin-Shirley BRDF [14] is used and a total of six different cases are examined (3 (pointing attitudes) × 2 (measurement noise values)).

The Ashikhmin-Shirley BRDF uses an exponential factor instead of a microfacet slope, a non-Lambertian diffuse component, and a different definition for the diffuse and specular reflectances. As a result, the reflectance and diffuse fraction parameters in the Ashikhmin-Shirley BRDF model needed to achieve the best match to a light curve generated by the Cook-Torrance BRDF model are also different.

The relation between the Ashikhmin-Shirley exponential factor and the Cook-Torrance microfacet slope is given approximately by equating the specular terms in each model when \( \hat{\mathbf{N}} = \hat{\mathbf{L}} = \hat{\mathbf{V}} \) (the illumination direction and observation direction are equal to the facet normal). This yields

\[
n = \frac{2}{m^2} - 1
\]  

(13)

The resulting match between the specular components of the two models is nearly exact for near normal incidence of both the illumination angle and observation angle, and becomes slightly different for larger angles as the influence of differences in the Fresnel reflection approximations within each model are manifested.

Approximate relations between the Ashikhmin-Shirley reflectance and diffuse fraction and the Cook-Torrance reflectance and diffuse fraction can be derived by setting the total diffuse reflectance for each over the entire hemisphere with normal illumination equal and the specular reflectance for each equal

\[
d_{CT}\rho_{CT} = d_{AS}\rho_{AS} \frac{31}{32}(1-(1-d_{AS})\rho_{AS})
\]  

(14a)

\[
(1-d_{CT})\rho_{CT} = (1-d_{AS})\rho_{AS}
\]  

(14b)

Solving for the Ashikhmin-Shirley parameters in terms of the Cook-Torrance parameters yields:

\[
\rho_{AS} = \rho_{CT} \left( \frac{d_{CT}}{1-\left(1-d_{CT}\rho_{CT}\right)} \right)
\]  

(15a)

\[
d_{AS} = \frac{1}{1 + \frac{31}{32}(1-(1-d_{CT})\rho_{CT})(1-d_{CT})/d_{CT}}
\]  

(15b)
Unfortunately, Eq. (15a) does not always produce a physical result for high Cook-Torrance reflectance values. In other words, there are some Lambertian reflectance values that cannot be duplicated on average using the non-Lambertian diffuse term in the Ashikhmin-Shirley model! Thus, to get as close as possible, the Ashikhmin-Shirley reflectance is set to a value slightly less than 1.0 and the appropriate diffuse fraction is calculated using (from Eq. 14b):

\[
d_{\text{AS}} = 1 - \frac{(1 - d_{\text{CT}})\rho_{\text{CT}}}{\rho_{\text{AS}}}
\]  

(16)

The resulting match between the diffuse components of the two models is good on average, but becomes significantly different for large illumination and/or observation angles. For a diffuse reflection dominated light curve, a very large uncertainty in the BRDF parameters would thus be required. For a specular reflection dominated light curve, a relatively small uncertainty is necessary. For this particular analysis, where specular reflections are present in each light curve, an uncertainty of 0.2 in the reflectance and diffuse fraction parameters was used.

4. RESULTS

Figure 4 shows the results for five different runs of the case where the satellite is pointing at Monument, the model uncertainty is 0.050, and the measurement noise is 0.05 magnitudes. The black lines are the state values as a function of time, the green lines are the ±3σ values as determined from the diagonal terms of the covariance. For visualization purposes, the global quaternion and GRP uncertainties at each time step have been converted to Euler angle equivalents in Fig. 4.

As can be seen in Fig. 4, UKF uncertainties are unrealistic and often the filter’s estimate is outside the ±3σ bounds. The USKF, however, always includes the estimate within bounds (which have been inflated by the inclusion of the consider terms).

Figures 5 and 6 show the multiple hypotheses testing scenario results in Roll, Pitch, and Yaw offset from truth for two particular cases. Black dots indicate the initial state and red dots indicate the final state for each hypothesis connected by a gray line. The truth is indicated by a blue diamond (at origin), and the final composite estimate as calculated by a weighted mean of all the hypotheses is represented by a green dot. The corresponding 1σ, 2σ, and 3σ uncertainties related to the Gaussian mixture probability density function (pdf) as calculated from the mean and covariance of all the hypotheses are represented by green lines. Specifically, the pdf is calculated for a 200 × 200 point grid ±20° from the origin (truth). The values are then sorted in descending order and summed cumulatively to
determine the pdf value corresponding to when the fraction 0.6827, 0.9545, and 0.9973 is reached. Thus, 68.27% of the pdf is contained within the 1σ contour, 95.45% of the pdf within the 2σ contour, and 99.73% of the pdf within the 3σ contour.

Figure 5 – UKF (left) and USKF (right) results of the Monument, Colorado, 0.001 uncertainty, and 0.10 measurement noise case.

Figure 6 – UKF (left) and USKF (right) results of the French Frigate Shoals, 0.100 uncertainty, and 0.05 measurement noise case.

In the case depicted in Fig. 5, the reflectance/diffuse fraction uncertainty is so small, that the final estimate between the UKF and the USKF are nearly identical. In the case depicted in Fig. 6, however, there is a significant difference between the UKF and USKF. In fact, the UKF not only exhibits a large systematic error in the final estimate, it also significantly underestimates the uncertainty. The USKF provides a reasonable estimate and associated uncertainty. In both cases, the resulting pdf contours can be reasonably approximated by a single Gaussian. In fact, in only a single case for the UKF do the resulting contours exhibit a more complex bimodal distribution.

The results of all 48 cases are plotted in Fig. 7. The left hand plot displays the Mahalanobis distance \( \left( \sqrt{\left( \hat{x}_i^* - \bar{x} \right)^T P_{x,d,k}^{-1} \left( \hat{x}_i^* - \bar{x} \right)} \right) \) where \( \bar{x} \) is the truth state (in Euler angles), \( \hat{x}_i^* \) is the final composite state (in Euler angles), and \( P_{x,d,k}^* \) is the final composite covariance as a function of reflectance/diffuse fraction uncertainty.
Although the Mahalanobis distance is meant for Euclidian parameters and not angular parameters as applied here, because the pdf of the final Gaussian mixture distribution can be approximated well with the final composite covariance and the magnitude of the covariances are relatively small, the Mahalanobis distance is an adequate measure. In the left hand plot, the red line corresponds to a Mahalanobis distance of three while the blue and green lines are the trends of the UKF and USKF respectively (note that the y-axis is logarithmic, so these are exponential trends). Of significance is that the green line stays well below the red line while the blue line goes significantly above the red line indicating the UKF significantly underestimates the actual uncertainty in the final state while the USKF provides a reasonable estimate of the uncertainty. The right hand plot displays the great circle distance (in degrees) between the truth and the estimate also as a function of uncertainty. The red line is at 3°, the nominal separation between initial hypotheses, and the blue and green lines are the trends of the UKF and USKF respectively. Of significance is the fact that the green line is consistently below the blue line indicating the USKF provides a final state closer to the truth on average than the UKF. In both plots, French Frigate Shoals results are in red, Pago Pago results are in green, Monument results are in blue, UKF with 0.05 magnitude measurement noise are circles, UKF with 0.10 magnitude measurement noise are squares, USKF with 0.05 magnitude measurement noise are pluses, and USKF with 0.10 magnitude measurement noise are asterisks.

Figure 8 plots the difference between the great circle angular distances to truth for the UKF minus the great circle angular distances to truth for the USKF. French Frigate Shoals results are in red, Pago Pago results are in green, Monument results are in blue, 0.05 magnitude measurement noise are pluses, and 0.10 magnitude measurement noise are asterisks. The red line is at zero, and the black line is the linear trend. Positive values correspond to the USKF performing better while negative values correspond to the UKF performing better. It is evident that the USKF outperforms the UKF and performs better with increasing uncertainty.
For the analysis where a different BRDF model was used in the estimation filter than the model used to generate the observations, the Mahalanobis distance and great circle distance for the six different cases are plotted in Fig. 9 using the same symbols as in Fig. 7. As before, the USKF provides better accuracy and a more realistic uncertainty in the estimated quantities.

5. CONCLUSIONS

In this paper it is shown that the unscented Schmidt Kalman filter (USKF), which incorporates the influence of the known modeling errors in so-called “consider” terms, provides a better estimate and a more realistic uncertainty on the final attitude estimate for this Earth-pointing geosynchronous satellite than the unscented Kalman filter (UKF). Managing model uncertainties is key in ensuring the filter estimates and associated uncertainties are accurate and reliable.
ACKNOWLEDGEMENTS

This work is sponsored by a Small Business Innovative Research (SBIR) contract FA9451-13-C-0093 and contract FA9453-14-C-0204, both through the Air Force Research Laboratory (AFRL). The authors would also like to acknowledge useful discussions with Marcus Holzinger of Georgia Institute of Technology, and Kris Hamada and David Sheppard of Integrity Applications Incorporated-Pacific Defense Solutions.

REFERENCES


