PROCESSING TWO LINE ELEMENT SETS TO FACILITATE RE-ENTRY PREDICTION OF SPENT ROCKET BODIES FROM THE GEOSTATONARY TRANSFER ORBIT

Aleksander A. Lidtke[1]*, David J. Gondelach*[1]*, Roberto Armellin*[2] *Camilla Colombo*[1]
Hugh G. Lewis[1]*, Quirin Funke*[3]*, Tim Flohrer*[3]

1 - Astronautics Research Group, University of Southampton, Southampton, United Kingdom
2 - Departamento de Matemáticas y Computación, Universidad de La Rioja, Logroño, Spain
3 - European Space Operations Center, European Space Agency, Darmstadt, Germany

ABSTRACT

Predicting the re-entry of space objects enables the risk they pose to the ground population to be managed. The more accurate the re-entry forecast, the more cost-efficient risk mitigation measures can be put in place. However, at present, the only publicly available ephemerides (two line element sets, TLEs) should not be used for accurate re-entry prediction directly. They may contain erroneous state vectors, which need to be filtered out. Also, the object's physical parameters (ballistic and solar radiation pressure coefficients) need to be estimated to enable accurate propagation. These estimates are only valid between events that change object's physical properties, e.g. collisions and fragmentations. Thus, these events need to be identified amongst the TLEs. This paper presents the TLE analysis methodology, which enables outlying TLEs and space events to be identified. It is then demonstrated how various TLE filtering stages improve the accuracy of the TLE-based re-entry prediction.

Index Terms— Re-entry prediction, GTO, geostationary transfer orbit, rocket bodies, TLE, two line element set, filtering, outliers

1. INTRODUCTION

Upper stages of rockets are large objects, which contain components that are known to be able to survive atmospheric re-entry. Such surviving material, for example propellant tanks, will impact Earth's surface and might cause ground casualties. Being able to predict re-entry in advance enables mitigation measures to be implemented to reduce the risk to the ground population. However, at present, re-entry can be predicted with an accuracy of 2 to 28% of the remaining lifetime in orbit [1]. A "rule of thumb" relative re-entry prediction uncertainty of $\pm 20\%$ is recommended by Pardini and Anselmo [1].

Such low accuracy of re-entry prediction makes implementing efficient risk mitigation measures difficult because of the large area over which the spacecraft might re-enter [1]. In other words, current re-entry predictions do not provide information that is actionable because for every specific location at risk, e.g. a city, the probability that the satellite will impact this specific point is extremely low.

Generally, re-entry prediction is done by propagating an object until it reaches the altitude where the atmospheric break-up occurs, which is typically between 72 and 84 km. The main components of this approach are determining the object's initial orbit and accurately modelling the forces that act on it. The latter task requires object's physical parameters, e.g. mass, to be known or estimated.

The inaccuracies associated with re-entry prediction have several different origins, which are well reviewed by Pardini and Anselmo [2] and references therein. The largest source of uncertainty is associated with modelling of the acceleration due to atmospheric drag, a_D, an object of mass m, cross-sectional area A and drag coefficient C_D, flying through a fluid of density ρ at speed v, experiences. This acceleration is given in Eq. (1) [3]. The object's physical parameters in this equation, $C_D A/m$, are often combined into a single "ballistic coefficient", BC.

$$a_D = \frac{1}{2}\rho v^2 C_D \frac{A}{m}. \tag{1}$$

Based on Eq. (1), it can be deduced that the more accurately the object's parameters are known, the more accurate the force modelling and hence the re-entry prediction is. The same holds true for the acceleration due to solar radiation pressure, which is described by object's physical parameters that are often combined into the solar radiation pressure coefficient, SRPC [3]. The initial state of the object also has to be known and propagated as accurately as possible in order to correctly predict its position, and hence also the local atmospheric density ρ and orbital velocity v. Lastly, an accurate atmospheric model is also required to correctly model ρ.

Currently, two line element sets (TLEs) are the only publicly available data that can be used for re-entry prediction of a space object. However, there is a number of factors that, if unaddressed, could reduce the accuracy of re-entry prediction based on TLEs:

1. The quality of TLEs of an object is not homogeneous; sometimes TLEs of low quality or even belonging to a different object (outliers) are published.

2. Occasionally, the object's physical properties (BC or SRPC) or its orbit can be altered by collisions, fragmentations or space weather phenomena. Such space events render the TLEs of the object from before the event inapplicable to its new, changed state.

3. TLEs do not provide information on space object parameters, notably the BC and SRPC. TLEs only include the B^* parameter that accounts for combined atmospheric drag and solar radiation pressure forces, not BC and SRPC individually.

4. TLEs can only be propagated using the SGP4/SDP4 propagator. However, this propagator is based on the Brouwer theory and, therefore, only models the largest perturbations affecting a satellite. The many assumptions of the theory limit the accuracy of the resulting propagation and thus of the re-entry prediction.

5. TLEs are not supplied with uncertainty information, e.g. a covariance matrix. It is thus challenging to estimate the accuracy with which the re-entry is predicted based on these ephemerides.

In order to overcome these difficulties in TLE-based re-entry prediction, a multi-step procedure is proposed. The first step consists of analysing TLEs, with the goal of identifying outliers and space events (addresses points 1 and 2 on the above list). The filtered TLEs are then used to estimate the unknown spacecraft BC and SRPC, which are needed to accurately propagate the object's state (point 3). Different BC and SRPC are estimated between events in order to account for the fact that object's physical properties may change during e.g. collisions or explosions. The last step consists of performing orbit determination, in which the TLEs are used as pseudo-observations. A state obtained from the orbit determination can be propagated using any propagator, notably more accurate than SGP4/SDP4, and the associated uncertainty can be estimated (points 4 and 5). The functional flow of the re-entry prediction algorithm is shown in Fig. 1.

This paper presents the approach adopted to process the TLEs to improve the accuracy of re-entry prediction ("TLE filtering" block in Fig. 1). The remaining steps of the re-entry prediction algorithm are described by Gondelach et al. [4]. The re-entry prediction algorithm is tailored to rocket bodies (R/Bs) located in geostationary transfer orbits (GTO). A a set of 116 R/Bs will be used as a reference sample of such objects. The space surveillance catalogue numbers (SSCs) of these R/Bs are given by Gondelach et al. [4].

This TLE processing is based on methods previously employed to detect space weather events and manoeuvres, e.g. by Patera [5] or Song et al. [6]. However, developing a set of

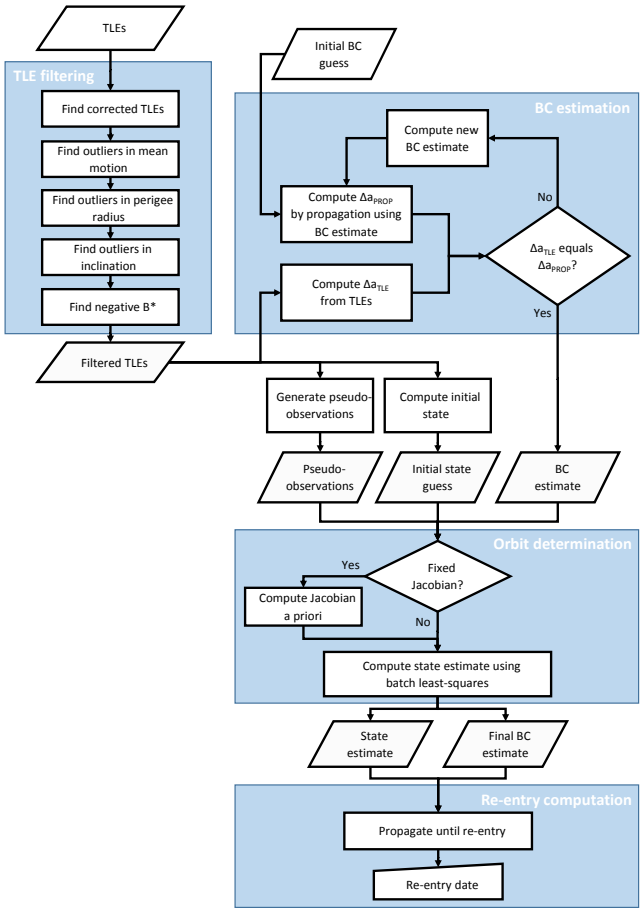

Fig. 1. Block diagram of the TLE-based re-entry prediction algorithm.

tools that allow outlying TLEs to be identified is a novel application of such algorithms. Details of the algorithm, which enables outlying TLEs to be identified in TLE time series of varying quality and generated in different phases of re-entry, are given. An approach to splitting the TLE time series into internally consistent sequences, where BC and SRPC estimates are believed valid, is described. The trade-off between the number of false positives and negatives, i.e. incorrectly identified and missed outliers, is emphasised and suitable algorithm settings are chosen. Lastly, the results of the TLE filtering algorithm are discussed in the context of the accuracy of the resulting re-entry prediction.

2. TWO LINE ELEMENT SET FILTERING

Details of every step of the TLE filtering algorithm from Fig. 1 will be described in turn. Finding outliers in mean motion is performed simultaneously with splitting of the TLE time series into internally consistent sequences. A sequence is defined as a set of TLEs that does not contain any events,

which might change object's physical properties (events are sequence endpoints). If a large time gap is present in the history of the TLEs for a given object, it cannot be reliably established whether the object's properties changed because TLEs cannot be reliably propagated from the beginning to the end of the time gap. Thus, large time gaps may also be endpoints of TLE sequences and a process for identifying those will be described.

2.1. Corrections

It is not uncommon for a TLE to be released soon after a previous one when the orbital elements in the TLE have been corrected [7]. Kelecy et al. filter TLEs so that only one element set is left in a 24-hour window, whereas Lemmens and Krag use half an orbital period [7, 8]. The time separations between consecutive TLEs of 116 R/Bs in GTO were computed and normalised w.r.t. the orbital period, computed using the mean motion in the more recent TLE. A subset of these data, where the time separation was at most five orbital periods, is shown in Fig. 2.

It can be noticed in Fig. 2 that most of the TLEs get updated at integer multiples of the orbital period. This supports the hypothesis that if a TLE is issued less than half an orbital period after the preceding one, it is a correction. If the update interval is between half of orbital period and one orbital period, the TLE could be an early release of the catalogue update. In the approach described herein, half an orbital period is taken as the correction threshold. This is to say, if two TLEs are separated by less time than half an orbital period, the more recent one is kept in the analysis and the older one is discarded. Should the more recent TLE be an outlier, it will be removed from the sample at a later stage of the filtering process.

TLE epochs being separated by integer multiples of orbital period is associated with the fact that, in the past, TLEs were released when objects were at their ascending nodes. However, as can be noted in Fig. 3, this no longer seems to be the case for most of the objects since approx. 2011. Had the objects been located at their ascending nodes at the epochs of their TLEs, the sum of the true anomaly and the argument of perigee would be equal to zero. This was the case for most of the TLEs published before 2011, but since then the spread in the position along the orbit at the TLE epoch is larger. This can also be seen in Fig. 4, which shows that the time spacing of TLEs published after the beginning of 2011 is more dispersed than before. However, there still appears to be a concentration of TLE updates at the integer multiples of orbital period, therefore the suggested threshold remains valid.

2.2. Time gaps

Large time gaps make it impossible to determine whether an event has taken place, i.e whether object's BC and/or SRPC

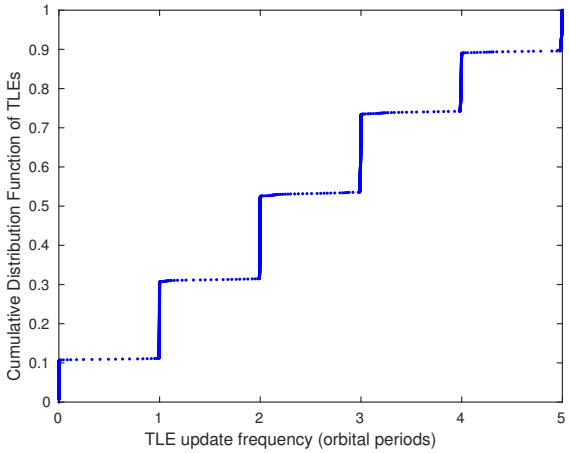

Fig. 2. Cumulative density function of the TLE update frequency for 116 example rocket bodies in GTO. Normalised to the orbital period computed from the more recent TLE. Only showing a subset of the data with update frequency of at most five orbital periods.

should be re-estimated to reflect the current object's parameters. This is because the state from the beginning of the large time gap cannot be reliably propagated until the end of the gap. If a TLE is both preceded and followed by a large time gap, it is deemed as an outlier.

It is unclear for how long a TLE can be accurately propagated. A TLE might cause automatic tracking of the object difficult if it is older than five days [9]. However, enhanced TLEs for highly-elliptical orbits (period greater than 225 minutes, eccentricity greater than 0.25) do not suffer almost any loss of accuracy c.f. numerically propagated states in the first week since the epoch of the element set [9]. Therefore, no validity period of the TLEs can be definitively established and the duration of the large time gaps has to be estimated.

Time spacing between TLEs of one object will vary, sometimes by tens of orbital revolutions. Such large time gaps may be present in the history of TLEs of an object with otherwise approximately constant TLE update frequency. Moreover, the TLE update frequency will vary from object to object. Thus, the threshold for the duration of large time gaps has to be computed for every object individually. Extraordinarily large time gaps, orders of magnitude longer than the typical TLE update interval for a given object, have been observed. Therefore, it was imperative that robust statistical metrics (e.g. median and median absolute deviation instead of mean and standard deviation) be used to reduce the effect of such outlying data points on the time gap threshold.

It was observed that the TLE update frequencies for most objects follow distributions, on which the extraordinarily large time gaps manifest themselves as tails. Thus, identifying the TLE time separation corresponding to the beginning

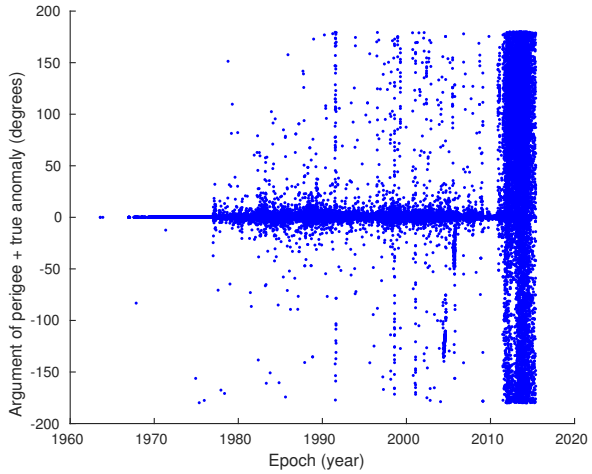

Fig. 3. Sum of the argument of perigee and the true anomaly of all the TLEs of the 116 GTO re-entry test cases at the TLE epochs.

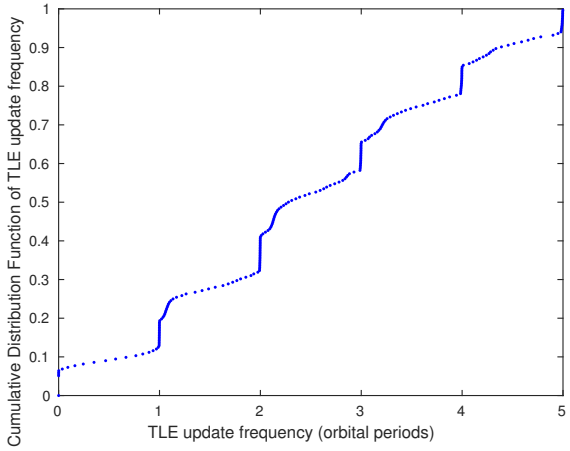

Fig. 4. Cumulative density function of the TLE update frequency for 116 example rocket bodies in GTO. Normalised to the orbital period computed from the more recent TLE. Data from Fig. 2 but only up to update frequency of five days and released after the beginning of 2011.

of this tail enables all large time gaps to be detected. This is advantageous c.f. using a threshold based on e.g. median plus a number of median absolute deviations (MADs), because it does not *a priori* assume that time gaps will be present; if the TLE time spacing distribution does not have a tail there are no unexpectedly large time gaps.

The tail of the distribution is found by first computing the time separations between the consecutive TLEs for a given object, as well as their median and MAD. Time separations in a percentile of choice are binned in equally spaced bins. The width of the bins is defined as the median time separation plus a number of MADs. The time separation of the first empty bin corresponds to the beginning of the tail of the distribution of the time separations, i.e. time separation at which the distribution stops being continuous.

The performance of this algorithm is shown in Fig. 5, where a large time gap in the TLEs was identified. The probability plot showing the distribution of the time separations between TLEs of this object, together with the identified large time gap duration, is shown in Fig. 6.

2.3. Outliers in mean motion and sequence endpoints

It was decided to detect events, which affect object's physical properties, by analysing the mean motion n contained in its TLEs because this parameter is proportional to object's orbital energy [3]. Mean motion and related semi-major axis were used to detect propulsive manoeuvres by Song et al. and Kelecy et al. [6, 8], and manoeuvres in the order of centimetres per second were detected in 95% of the cases [8]. Even though the objects nearing the re-entry are likely derelict, and so do not conduct manoeuvres, they interact with the atmosphere. Thus, if their physical properties change, so will the drag force acting on them, and consequently their orbital energies and n.

Single outlying TLEs and events are detected by sliding a window, containing a fixed number of TLEs, through the TLE time series. A polynomial of the chosen order is regressed through the TLEs contained in the window and the TLE following the window is examined ($i + 1$, where i is the last TLE in the window). Theil-Sen-Seigel robust linear regression or bisquare weighted least-squares regression are used in order to reduce the impact of the outliers, which may be present in the window, on the regressing function.

The threshold for identifying outliers and events is recomputed throughout the analysis [7, 10] rather than using e.g. a fixed number of standard deviations above the mean [5]. Recomputing the threshold is advantageous because the orbit will naturally evolve and the same difference between the TLE and the regressing function at one instant in time might signify an outlier, but otherwise it might correspond to a fast, natural evolution of the orbit. Also, the time spacing between the TLEs is not constant, as shown in section 2.1, and thus the differences between the regressing function and individual TLEs will vary.

The tolerance is set on a quantity called the relative threshold, T_R. T_R is the ratio between the difference between n of $i+1$ TLE and the regressing function (this difference is called the residual, Δ_A), and the change predicted between the i and $i+1$ TLE, Δ_P. T_R is given in Eq. (2) and the quantities used to define it are shown in Fig. 7.

$$T_R = \frac{\Delta_A}{\Delta_P}. \tag{2}$$

The predicted change, Δ_P, will be close to zero when the

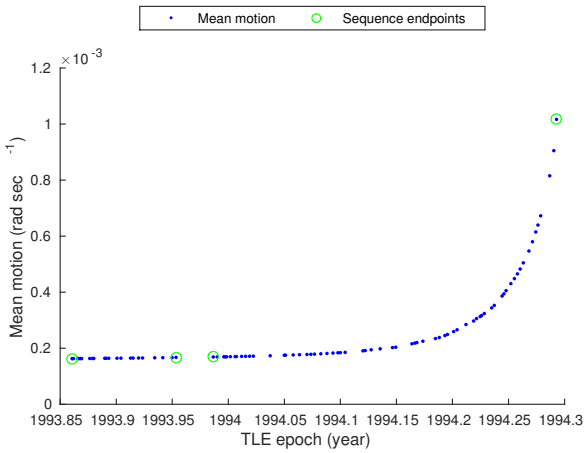

Fig. 5. Large time gap around year 1993.97 found for object 22906 by analysing data in the 95^{th} percentile by binning the time separations into bins with width equal to median plus one MAD.

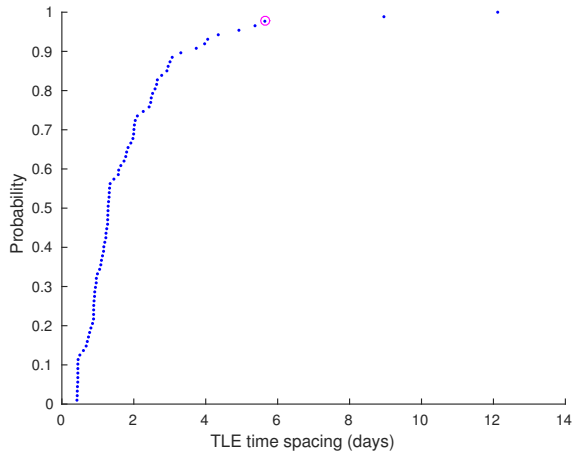

Fig. 6. Distribution of time separations between consecutive TLEs for object 22906. Beginning of the tail, identified by analysing data in the 95^{th} percentile by binning the time separations into bins with width equal to median plus one MAD, is indicated with a magenta circle.

mean motion does not vary significantly between the first and the last TLEs in the sliding window. This will cause T_R to be amplified and, occasionally, fall above the set tolerance, even if the TLE following the window is not an outlier or a beginning of an event. In order to avoid such false positives, or incorrectly marking TLEs as outliers, an absolute threshold, T_A, is used. The tolerances on both T_R and T_A need to be exceeded in order for a TLE to be classified as an outlier or a sequence endpoint. T_A is the previously defined residual ,Δ_A, normalised with the regression mean motion at the epoch of the last TLE in the window, $n_{REG}(t_i)$. Expression for T_A is given in Eq. (3) and the quantities used to define it are shown in Fig. 8. Due to the normalisation, T_A is in fact a relative quantity. However, it is used to set a threshold on the absolute value of n, while the relative threshold refers to difference in changes from the last TLE in the window to the next.

$$T_A = \frac{\Delta_A}{n_{REG}(t_i)}. \tag{3}$$

As the window is slid through the TLEs (index of the last TLE in the window, i, is incremented), T_R and T_A are computed at every $i+1$ TLE. If both T_R and T_A are above the set tolerances, $i+2$ TLE is investigated. If the relative difference between the mean motions n_{i+1} and n_{i+2} of the $i+1$ and $i+2$ TLEs, given in Eq. (4), is less than the absolute tolerance, then those two TLEs are considered consistent and a start of a new sequence is declared. This situation is shown in Fig. 9. Note that due to how the absolute tolerance is defined (to be used with T_A as well), Δ_{i+2} has to be normalised by the regression mean motion at the $i+1$ TLE, $n_{REG}(t_{i+1})$.

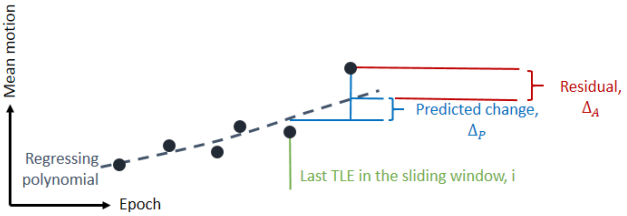

Fig. 7. Definition of the residual, and the predicted change between the regressing function at the epoch of the last TLE in the sliding window and the following TLE. The differences are computed using the value of the regressing function at the epoch of the last TLE in the window, not the value of that TLE. Used to define the relative threshold from Eq. (2).

$$\Delta_{i+2} = \frac{n_{i+2} - n_{i+1}}{n_{REG}(t_{i+1})}. \tag{4}$$

2.4. Outliers in eccentricity and inclination

Once the corrected TLEs are filtered out and the element sets with outlying mean motion are removed from the sample, some of the remaining TLEs may still have outlying orbital elements.

As far as the ballistic coefficient estimation is concerned (second step in the block diagram in Fig. 1), eccentricity e is particularly important because it affects the perigee altitude and thus also the drag force acting on the object. If eccentricity in the TLE is incorrect the BC will be estimated incorrectly, thus reducing the accuracy of re-entry prediction. Once

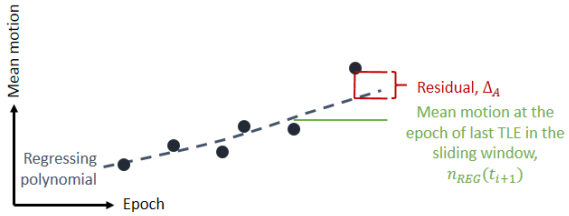

Fig. 8. Definition of the residual and the value of the regressing function at the epoch of the last TLE in the window, used to define the absolute threshold from Eq. (3).

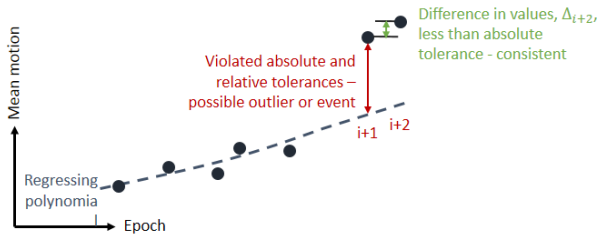

Fig. 9. Start of a new sequence after an event that affected the mean motion of the object.

the object parameters are estimated, orbit determination (OD) is performed (third step in Fig. 1). Individual TLEs (or orbital elements derived from them) are used as pseudo-observations in the OD process. It was found that TLEs outlying in inclination, i, may lead to poor estimates of the state, which is propagated until the re-entry, and consequently inaccurate re-entry predictions. Therefore, after filtering out outliers in n, TLEs with aberrant e and i also need to be removed from the sample before subsequent analyses are performed.

Instead of filtering TLEs based on eccentricity directly, perigee radius, r_P, is used instead. The two quantities are equivalent because $r_P = (1-e)a$, where a is the semi-major axis [3]. If the outliers in mean motion n are already filtered out, there should be no outliers in a present in the sample ($n = \left(\mu/a^3\right)^{1/2}$ [3]). Thus, any outliers in r_P will be present only due to outliers in e. It was found that filtering outliers in r_P is easier than using e directly due to scaling. Inclination extracted from the TLEs was used directly to filter out the outliers in this orbital element.

Outliers in inclination and eccentricity are identified within every internally consistent sequence found according to the algorithm described in Section 2.3, i.e. within sequences terminated by events and large time gaps. Because all the events have already been identified at this stage, a simpler filtering technique is used for i and r_P than for n. A window of a fixed length is slid through the time series of the given orbital element in every consistent sequence, and the median value in the window is computed. This median value is subtracted from the TLE in the middle of the window, thus converting the time series of TLE values to a time series of differences (data are detrended). Another window may be slid through the time series of differences, and MAD of the differences may be computed for the TLEs in this sliding window. Alternatively MAD may be computed from all the differences for a given object. Lastly, TLEs that have a given orbital element further than a fixed number of median absolute deviations away from the median are marked as outliers. This filtering process may be repeated a number of times with different settings for every orbital element to filter more outliers. The settings for the filtering in i and r_P are given in Table 1.

Table 1. Parameters of the algorithm used to identify outliers in inclination, i, and perigee radius, r_P. Showing lengths of the windows used to compute the median and MAD of the orbital element. Window lengths are given as the number of data points (TLEs) in the window. Where applicable, settings for every pass of the filtering algorithm are listed in sequence.

Filter setting	Orbital element	
	i	**r_P**
Window length - median	11	21,11
Window length - MAD	50	All TLEs,50
Number of MADs	12	15,15
Number of filter passes	1	2

2.5. B^* filtering

TLEs contain a B^* parameter that is proportional to the ballistic coefficient BC ($BC = 12.741621B^*$ [3]). It was observed that even 30 to 60 days before the actual re-entry, when the perigee altitude is low and the atmospheric drag is the most prevalent force disturbing the satellite, TLEs may contain negative B^* values. Negative B^*, corresponding to negative ballistic coefficient, signifies that the orbital energy of the object is increasing secularly. A situation like this is possible if the solar radiation pressure force acting on the object is strong. However, close to re-entry the atmospheric drag is stronger than the solar radiation pressure, therefore a long-term increase in the orbital energy is impossible (drag reduces the orbital energy). Only short period (less than orbital period), not secular increases in the orbital energy are possible in circumstances like this. Therefore, it was decided to filter out TLEs with negative B^* values because they would cause incorrect estimates of the ballistic coefficient.

3. FILTER TUNING

Settings for the correction, time, and the inclination and eccentricity filters, which are presented in Sections 2.1, 2.2 and 2.4, respectively, are the parameters that were found best in

this study. Arriving at the chosen settings for the correction filter is described in Section 2.1, whereas the cited values for the remaining three filters (time, e and i) were chosen based on a "trial-and-error"search. In particular, the e and i filters were tuned to remove the largest, most evident outliers in the TLE time series. Because the tuning of these filters' settings cannot be quantified, it will not be presented here. However, the mean motion filter was tuned rigorously because identification of events is crucial for the accuracy of re-entry prediction. Moreover, careful tuning of this filter was necessary in order to ensure that outliers are not misidentified as events and *vice versa*. The process of arriving at the optimal settings for the n filter will be described in this section.

3.1. Mean motion filter tuning

Manual analyses of quality of the TLEs of the sample GTO objects revealed that certain TLEs will be difficult to classify as outlying or correct, even to an operator. In order to provide a time series of mean motion with unambiguous outliers, thus enabling the performance of the filtering algorithm to be quantitatively assessed, a synthesised TLE time series was used.

To this end, the time series of the mean motion of the object 13025 (Ariane 1 R/B) was used because it was long (spanned 6.9 years) and contained a number of phases of high drag (rapidly increasing n) separated by phases of low drag (slowly varying n). These properties made it possible to test the filter for a number of gradients of n, denoted as $\dot{n}$. This gradient may affect the absolute value of the residual between the regressing function and the TLEs, as described in Section 2.3, and consequently also the filter performance. Therefore, testing the algorithm for a range of values of n and $\dot{n}$ was needed to find robust settings for the filter. TLE history of 13025 also contained several large time gaps, enabling the filter to be tested when restarting the sliding window analysis with no information about the preceding TLEs (no TLE from before a large time gap is used to filter after the gap). Otherwise, manual inspection of the TLEs of 13025 revealed it did not contain significant outliers, making its TLE time series trustworthy relatively to other investigated objects.

Corrected TLEs were removed from the 13025's time series and the resulting time series of n was smoothed using a running window with robust quadratic polynomial regression. Different outlier combinations, which included single, two or three consecutive outlying TLEs, were distributed in the time series. Multiple outlying TLEs of different magnitudes and directions (e.g. n increasing in one and decreasing in the following TLE), were found to be particularly prone to cause false positives and negatives, therefore such combinations of outliers were included in the data. The resulting time series with the simulated outlier locations is shown in Fig. 10. A number of outlier magnitudes, OM, was used to scale the n of the TLEs in the simulated outlier locations. To simulate

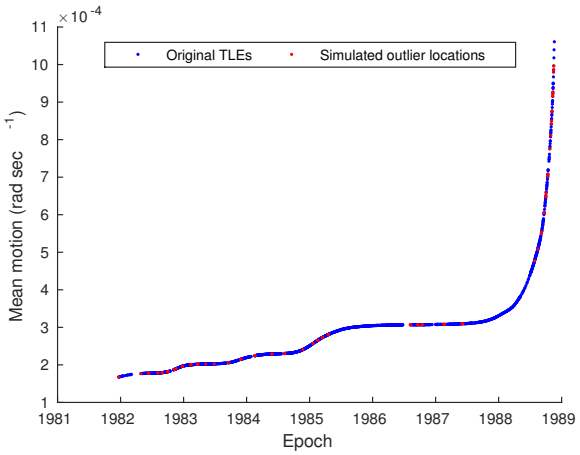

Fig. 10. Time series of the mean motion of object 13025 used to test the mean motion filtering algorithm. Locations, where the outliers were simulated, are marked with red. Details of outlier simulation are given in text.

an outlier, n of the TLE was scaled to an outlying n_o as $n_o = n + OM \times n$ for all simulated outlier magnitudes. The used outlier magnitudes were $\pm\{0.005, 0.01, 0.05, 0.1, 0.5, 0.75\}$.

The outlier detection algorithm was run on the TLE time series with all the outlier magnitudes. The automatically identified outliers were compared to the known outlier locations. In cases where the algorithm detected an outlier, which was not present in the data, a false positive (a false outlier), was recorded. Conversely, when the algorithm did not identify a simulated outlier, a false negative (a missed outlier), was recorded. The numbers of false positives and negatives for all outlier magnitudes were summed to produce a single metric for the given algorithm run. The runs were repeated with all the combinations of the investigated settings and three orders of three regressing polynomial, namely 1^{st} (Theil-Sen-Seigel), 3^{rd} and 5^{th}. The tested values of the relative and absolute tolerances, and window lengths are shown in Table 2.

The number of false positives and negatives produced by every tested combination of algorithm settings is shown in Fig. 11. No algorithm setting was able to produce zero false positives and negatives, i.e. perfect results. It appears that the investigated filter setting combinations spanned the entire spectrum of filter performance, because solutions with both many false positives and negatives, as well as solutions in between, were found for every polynomial order. Thus, it is unlikely that a combination of settings that yields perfect performance exists at all. Figure 12 shows the number of false positives and negatives for the Theil-Sen-Seigel regression. It can be noted that the larger the absolute and relative tolerances are, the fewer false positives, or false outliers, are identified. This is because the algorithm tolerates larger deviations from the regressing polynomial. However, large tolerances also give rise to many false negatives, i.e. missed

Table 2. Tested values of the absolute and relative tolerances (tolerances on T_A and T_R from Section 2.3) and window lengths that were tested with different regressing polynomial orders.

Filter setting	**Regression order**		
	$\mathbf{1^{st}}$	$\mathbf{3^{rd}}$	$\mathbf{5^{th}}$
Absolute tolerance	$\{5E{-}4, 1E{-}3, 5E{-}3, 1E{-}2, \ldots$ $2.5E{-}2, 5E{-}2, 1E{-}1, 2.5E{-}1, 5E{-}1\}$		
Relative tolerance	$\begin{bmatrix}0.01\\0.05\\0.10\\0.25\\\vdots\end{bmatrix}$ $\begin{bmatrix}0.50\\0.75\\1.00\\1.25\\1.50\end{bmatrix}$	$\begin{bmatrix}3.0\\3.5\\4.0\\5.0\\\vdots\end{bmatrix}$	$\begin{bmatrix}6.0\\7.0\\8.0\\9.0\\10.0\end{bmatrix}$
Window length	$\begin{bmatrix}3,4,\\5,10,\\15,20,\\25,30,\\50\end{bmatrix}$	$\begin{bmatrix}5,10,\\15,20,\\25,30,\\40,50,\\60\end{bmatrix}$	$\begin{bmatrix}7,10,\\15,20,\\25,30,\\40,50,\\60\end{bmatrix}$

outliers. The same behaviour is observed for all regressing polynomials and window lengths. A situation like this necessitates a trade-off between the number of false positives and negatives. In the context of re-entry prediction, false positives (false outliers) will reduce the number of TLEs that can be used, whereas false negatives (missed outliers) will deteriorate the accuracy of the results because subsequent analyses will use incorrect TLEs.

Because no perfect filter settings have been identified, a trade off study was performed. Points closest to the origin, which minimise the number of false positives and negatives, were chosen for every polynomial order from Fig. 11. The corresponding settings, together with the associated numbers of false and missed outliers, are presented in Table 3. The 3^{rd} order polynomial had the fewest false outliers. Generally, higher-order polynomials are better able to represent the variation of n with time, thus they result in fewer false outlier. However, this reduced number of false positives is associated with increased number of missed outliers. Missed outliers will deteriorate the accuracy of the re-entry prediction, thus this performance metric is deemed more important. Therefore, using 1^{st} order polynomial regression, which had the fewest missed outliers, was favoured.

Even with the best settings found here, the algorithm did not identify 101 outliers. The total number of outliers simulated in the study was 59 049, meaning that the algorithm correctly identified 99.83% of the simulated outliers. The settings from Table 3 could be further optimised by evaluating more combinations of the parameter values close to the best combination found so far. However, this was not performed because, in reality, other parts of the re-entry prediction algorithm will have a larger impact on deteriorating the prediction accuracy. Therefore, the settings from Table 3 for the 1^{st} order polynomial were used in the remainder of this study.

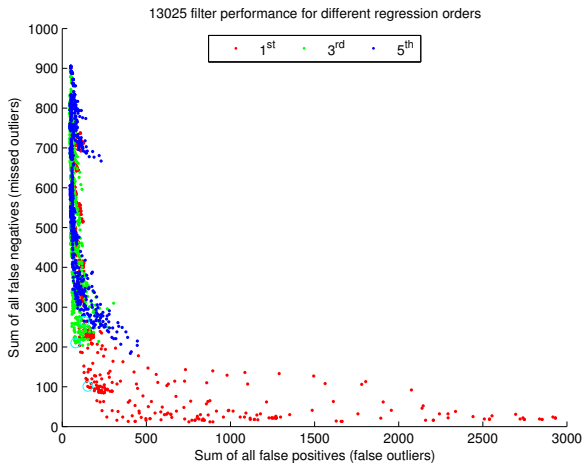

Fig. 11. Filter performance for all the combinations of settings from Table 2. Each dot represents one run of the algorithm, i.e. one combination of the settings for the given regression order. Points closest to the origin indicated with a cyan circle.

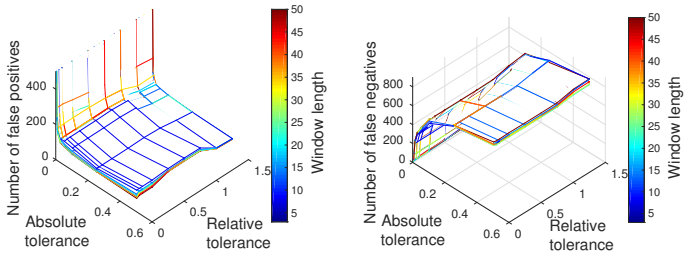

Fig. 12. The number of false positives (false outliers, left) and negatives (missed outleirs, right) identified with all the combinations of settings from Table 2 for 1^{st} order polynomial regression.

4. IMPACT OF TLE FILTERING ON RE-ENTRY PREDICTION

Re-entries of 92 test objects listed by Gondelach et al. [4] were predicted 30 days before the actual recorded re-entry date according to the algorithm given by Gondelach et al. [4] (TLE filtering, followed by BC and SRPC estimation, and orbit determination). All these objects have decayed already, therefore their re-entry epoch is made publicly available via www.space-track.org (these reference epochs are rounded to one day). This enables the discrepancy between the actual time that the object spent on-orbit since the prediction epoch, τ_A, and the predicted time on-orbit, τ_P, to be compared. The percent prediction error, $\delta\tau$, of τ_P relative to τ_A was computed as per Eq. (5) for different filtering levels. The investigated filtering levels and their acronyms are given

Table 3. Optimised parameters of the TLE filtering algorithm, together with the corresponding numbers of missed and false outliers (false negatives and positives, respectively).

Filter setting	Regression order		
	1^{st}	3^{rd}	5^{th}
Window length (no. TLEs)	5	20	40
Absolute tolerance	5E-3	1E-3	1E-3
Relative tolerance	5E-1	4.0E0	6.0E0
Performance metric			
Missed outliers	101	210	265
False outliers	152	80	131
Fraction of missed outliers	0.17%	0.36%	0.45%
Fraction of false outliers	0.26%	0.14%	0.22%

Table 4. Investigated TLE filtering levels and the acronyms by which they are referred to.

Filter acronym	Included filters
n, e, i, B^*	Correction, time, mean motion, perigee radius,inclination, B^*
n, e, i	Correction, time, mean motion, perigee radius,inclination
n, e	Correction, time, mean motion, perigee radius
n	Correction, time, mean motion
$NONE$	No filter

in Table 4.

$$\delta\tau = \frac{\tau_P - \tau_A}{\tau_A} \times 100. \quad (5)$$

Figure 13 shows the fraction of the sample that had a given re-entry prediction error when the raw, unfiltered TLEs were used as well as when full filtering from Table 4 was employed. 80% of the sample had $\delta\tau$ less than 14%. Up to that point, TLE filtering had little impact on the accuracy of the predictions, meaning that filtering was surplus. For three objects, however, maximum $\delta\tau$ of 50% was recorded if no filtering was performed. When the TLEs for these objects were filtered before predicting the re-entry, the error decreased to at most 37%. This shows that TLE filtering may not be necessary in all cases, however it is key to achieving consistent accuracy in the order of 10%.

Figure 14 shows the relative errors in predicting the re-entry epoch for four example R/Bs when different filtering stages were used. These objects exhibit typical types of behaviour that have been observed amongst the investigated sample. Note that the prediction for 25496 was performed 30 and 31 days before the actual re-entry.

Different filtering stages may improve the prediction accuracy by removing outliers. For example, filtering outliers in mean motion of 37949 improved the accuracy from over 50% to 14.3%. In certain cases, however, filtering TLEs may decrease the number of usable element sets and reduce the prediction accuracy, which was the case for 25051. In this case, the number of TLEs, which were used to predict the re-entry, reduced from three to one. These samples were small in both cases, which resulted in errors of $13.4 - 27.7\%$ for all filtering levels. Such small TLE samples do not, generally, result in accurate re-entry predictions [4]. Therefore, this reduction of accuracy with more filtering does not invalidate the approach to filtering because re-entry predictions based on too few data points are not reliable anyway.

The accuracy improvements brought by filtering may not always be visible because outliers are not always present in the used TLE set, e.g. for 22254 or for 25496 when the

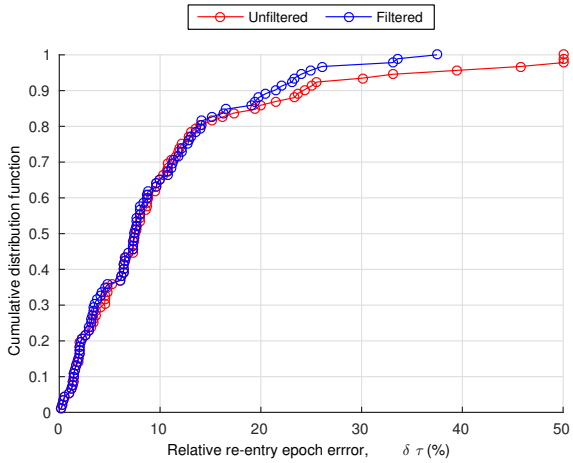

Fig. 13. Relative errors in predicting the re-entry epoch with full and no TLE filtering for 92 cases from Gondelach et al. [4], 30 days before the actual re-entry. The errors were capped at 50%.

prediction was performed 30 days before the actual re-entry (cyan series on Fig. 14). However, in the presence of outliers, in particular when an outlying TLE is being used for BC estimation, for example, filtering plays a key role, as shown in Fig. 13. This can be seen by examining the prediction accuracy for the same object, 25496, but done a day earlier, with a different TLE set (magenta series on Fig. 14). The TLEs of 25496 31 days before re-entry contained an outlier in inclination, which was causing the re-entry prediction to be accurate to 16.1%. However, when the outliers in i were filtered out, the accuracy improved to 9.3%. This shows that filtering does not always improve re-entry prediction accuracy. However, it is necessary to ensure that the accuracy is as high as possible, should any outliers be present in the data. This was shown for the n and i filters here, however it was observed for all the discussed orbital elements.

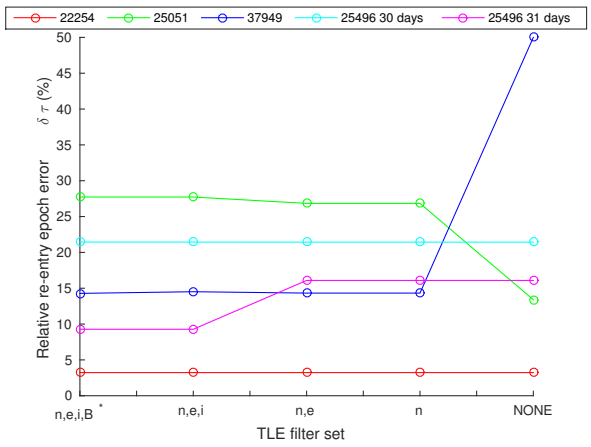

Fig. 14. Relative errors in predicting the re-entry epoch with different TLE filtering levels from Table 4 for four objects 30 days before the actual re-entry. The errors were capped at 50%. The prediction for 25496 was performed 30 and 31 days before the actual re-entry.

5. SUMMARY AND CONCLUSIONS

A multi-stage TLE filtering methodology was developed, and applicability of the filtering stages to improving the accuracy of re-entry prediction was discussed. It was observed that re-entry prediction accuracy may reduce due to TLE filtering because fewer TLEs are available to perform object's parameters estimation and orbit determination. However, it was shown that filtering out the outlying TLEs is vital in order to reduce the prediction errors. Thus, it is recommended to used filtered TLE time series, but without relying on a single prediction. It may prove useful to manually analyse the time series, however often it is difficult for a person to distinguish between a correct and an outlying TLE.

During the development of the TLE filtering algorithms, use of robust statistics and regression methods was found crucial. TLE time series may contain significant outliers, which render non-robust algorithms unreliable.

It was shown that the TLEs changed after 2011. Moreover, the TLE generation process changed further in 2013 [9]. Therefore, it is recommended to use different algorithms, or different algorithm settings, when working with TLEs generated before and after these years.

6. ACKNOWLEDGEMENTS

This work was performed in the framework of the European Space Agency project ITT AO/1-8155/15/D/SR, titled " Technology for Improving Re-Entry Predictions of European Upper Stages through Dedicated Observations".

The authors acknowledge the use of the IRIDIS High Performance Computing Facility at the University of Southampton in the completion of this work. Additionally, the Joint Space Operations Center is acknowledged for providing Vector Covariance Messages for 53 objects used for testing of the algorithms.

R. Armellin acknowledges the support received by the Marie Sklodowska-Curie grant 627111 (HOPT - Merging Lie perturbation theory and Taylor Differential algebra to address space debris challenges).

7. REFERENCES

[1] C. Pardini and L. Anselmo, "Reentry predictions of three massive uncontrolled spacecraft," in *23rd International Symposium on Space Flight Dynamics*, Pasadena, CA, USA, 2012.

[2] C. Pardini and L. Anselmo, "Re-entry predictions for uncontrolled satellites: results and challenges," in *6th IAASS Conference Safety is Not an Option*, Montreal, Canada, 2013.

[3] D Vallado, *Fundamentals of Astrodynamics and Applications; Fourth Edition*, Microcosm Press, 2013.

[4] D.J. Gondelach, A.A. Lidtke, R. Armellin, C. Colombo, H.G. Lewis, Q. Funke, and T. Flohrer, "Re-entry prediction of spent rocket bodies in gto," in *26th AAS/AIAA Space Flight Mechanics Meeting*, Napa, CA, USA, 2016, number AAS 16-240.

[5] R. Patera, "Space event detection method," *Journal of Spacecraft and Rockets*, vol. 45, pp. 554–559, 2008.

[6] W.D. Song, R.L. Wang, and J. Wang, "A simple and valid analysis method for orbit anomaly detection," *Advances in Space Research*, vol. 49, pp. 386–391, 2012.

[7] S. Lemmens and H. Krag, "Two-line-elements-based maneuver detection methods for satellites in low earth orbit," *Journal of Guidance, Control, and Dynamics*, vol. 37, pp. 861–868, 2014.

[8] T. Kelecy, D. Hall, K. Hamada, and D. Stocker, "Satellite maneuver detection using two-line element (tle) data," in *Advanced Maui Optical and Space Surveillence Technologies Conference*, Wailea, Maui, HI, USA, 2007.

[9] M.D. Hejduk, S.L. Casali, D.A. Cappellucci, N.L. Ericson, and D.E. Snow, "A catalogue-wide implementation of general perturbations orbit determination extrapolated from higher order orbital theory solutions," in *Spaceflight Mechanics Conference*, Kauai, HI, USA, 2013.

[10] J.D. Dolado-Perez, L.A. Garcia, A. Agueda Mate, and I. Llamas de la Sierra, "Opera: A tool for lifetime prediction based on orbit determination from tle data," in *24th International Symposium on Space Flight Dynamics*, Laurel, MD, USA, 2014.