

Time Series Modeling of Lifespans of Earth Satellites

Venkata Jaipal Reddy BATTHULA

Department of Information Science, University of Arkansas at Little Rock
Little Rock, AR 72204 USA

Richard S. SEGALL

Department of Computer & Information Technology, Arkansas State University
State University, AR 72467 USA

Daniel BERLEANT

Department of Information Science, University of Arkansas at Little Rock
Little Rock, AR 72204 USA

Venkat KODALI

CARTI, 8901 Carti Way
Little Rock, AR 72205 USA

Michael HOWELL

Department of Information Science, University of Arkansas at Little Rock
Little Rock, AR 72204 USA

Hyacinthe ABOUDJA

Department of Computer Science, Oklahoma City University
Oklahoma City, OK 73106 USA

ABSTRACT¹

Numerous countries have satellites for various needs like communication, weather forecasting and security. The first satellite was launched into space in 1957. Thousands of satellites' lifetimes have already ended but many of these satellites are still in orbit although no longer functional. Nevertheless, satellite technology is improving compared with the past. We seek to understand trends in improvements to satellites related to their lifespans better, using a new dataset that has not been available before for this article, as part of a project that also analyzes other relevant data sets.

The main aim of this paper is to analyze the satellite data with data visualization tools and find some significant patterns in the data. The paper explores the average lifespan of satellites over the period of artificial satellites, 1957 to present.

Keywords: Lifespan, Time Series, Satellites

1. INTRODUCTION & BACKGROUND

Satellites are launched for various purposes and users of the satellites are now spread across the world. The

activities dependent on satellites have increased greatly over time and they now have a major impact on everyday life. A satellite for our purposes is a technological device which was placed into orbit for a purpose. These are known more fully as artificial satellites. Since satellites are launched into space to do a specific task, they can be classified by their functionality (Wikipedia [11], 2019).

This paper is focused on data analysis of satellite lifespans. The dataset used for analysis is discussed below. The objective is to develop data visualization charts and modeling techniques, identify trends in the satellite data, and thereby better understand satellite lifespans. Previous and related work on this research by the co-authors has been provided in Batthula ([1], 2019), Berleant et al. ([2], 2019), Duggirala et al. ([3], 2018), Hall et al. ([4], 2017), Howell et al. ([5] [6], 2019; [7], 2020), Kodali et al. ([8], 2019; [9], 2018).

2. SATELLITE DATA SET

The satellite database of the Union of Concerned Scientists (UCS) ([12], 2019) consists of approximately 2000 working satellites currently in Earth Orbit. The Union of Concerned Scientists dataset gets updated three times each year.

Satellite catalog data may be downloaded from **planet4589.org**. The site provides the latest Satellite Catalog data and releases Space Report current issues.

¹Acknowledgement is made to Peng-Hung Tsai of University of Arkansas at Little Rock (UALR) for editorial review of final submission.

This site provides statistics about satellites, the master satellite list, and a geostationary satellite log. We also used data we call “New Death Data” herein which contains different kinds of values in the Death Date field such as those shown in Table 1. This necessitated the need for data cleansing before any data modeling could be performed.

Table 1: New Death Data records showing different kinds of values in the Death Date field.

Satellite Catalog	Death Date
S000002	1957 Oct 27
S000003	1957 Nov 9
S000022	1961 Feb
S000063	1962 May
S000116	1972?
S000567	1963?
S000965	* U
S025336	*
S025338	*
S025478	* U
S023726	* E
S024653	* ?

3. MODELING METHODOLOGY

3.1 Data Cleansing

According to Wikipedia ([10], 2019), “Data cleansing or data cleaning is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table or database and refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data.”

For cleansing the data, Python code was written a sample of which is shown in Table A in the Appendix of this paper.

- If we have month but not the day of death, then a set of imputation rules were used.
- If launch was in a different month, then estimate day of death as the middle day of the month.
- If launch was in the same month, then estimated death is the middle of the rest of the month.

If we have year but not the month or day, or not even the year is provided, we did not use the record in the analysis.

Sample Python code for the above methods of data imputation is shown as Table B in the Appendix.

3.2 Data Analysis of New Death Data with Imputed Records

After completing data cleaning and data merging, the output file was useful for analysis using the Minitab 19 statistical and data visualization software.

The cleaned output file with imputed data records was used for analysis. It contained 5610 records from the New Death Data dataset which included 767 imputed records. All analysis related to the lifetimes of satellites was performed using satellite lifetimes calculated as the difference between Death Date and Launch Date (Batthula, 2019 [1]):

$$\text{Lifetime} = \text{Death Date} - \text{Launch Date.}$$

3.2.1 Linear Trend Modeling: Figures 1 and 2 below illustrate the modeling of the data using time series analysis with trend analysis. These were generated using Minitab 19 with the fitted model $Y(t) = -1.298 + 0.12913t$. This model yields a Mean Absolute Percent Error (MAPE) of 73.29% which indicates a poor fit to a linear trend. Figure 2 illustrates several plots of the residuals between the fitted model values and actual data values.

The residual plots illustrate a dichotomous grouping for these residuals with linear decreasing approximation for the first half followed by erratic dispersion of the fits, indicating a not good fit to the linear trend model shown in red diagonal in Figure 1.

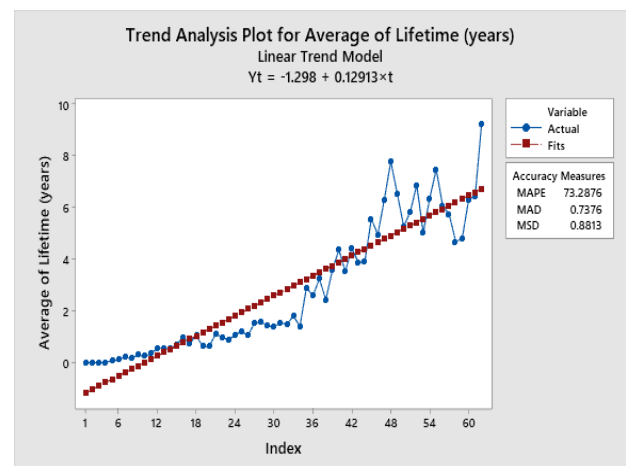


Figure 1: Linear Trend Model for Yearly Average Satellite Lifetimes. X-axis Shows Year of Satellite Death Starting from 1957 (Indicated as Index 1).

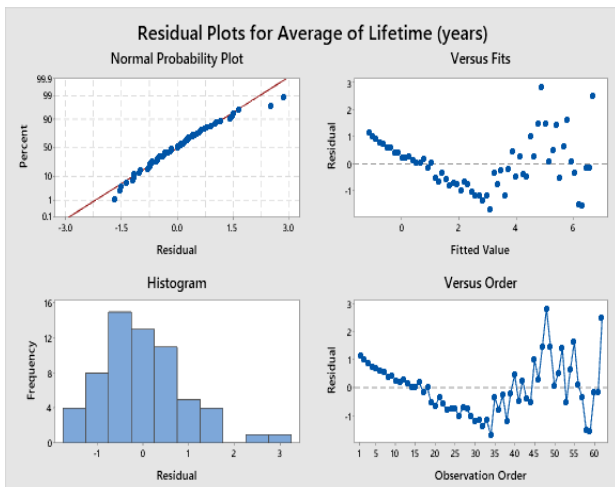


Figure 2: Residual Plots for Average of Lifetimes (by year) using the Linear Trend Model of Fig. 1.

3.2.2 3-year Moving Average Analysis: To calculate the 3 point moving averages from a list of numbers, follow these steps: 1) Add up the first 3 numbers in the list and divide answer by 3. 2) Store this answer as the first 3 point moving average. 3) Give that moving average the index of 3. 4) Repeat from step 1) with the 2nd through 4th numbers giving the result the index of 3. 5) Continue repeating with the 3rd through 5th numbers, then the 4th through 6th, etc., up to the (n-2) through nth. Going through the 62 data values, we note that the MAPE (Mean Absolute Percent Error) is now significantly reduced from 73.28% in Fig. 1 to 22.72% for non-centered data, and 14.53% using centered data for the 3-year moving averaged data as shown in Figure 3. If the model was tweaked so that each average of three was used as an estimate of the middle of the three, the resulting error calculations would likely improve slightly. The moving average provides an indication of the level of year-to-year variability, such as noise, in the data.

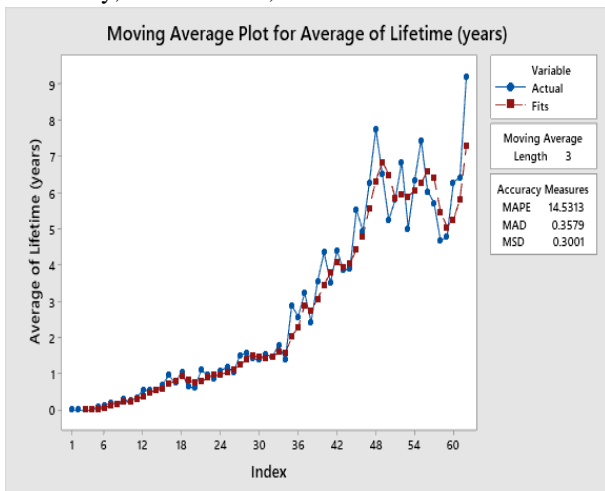


Figure 3: 3-Year Moving Average Plot for Average of Lifetimes by Year.

The residual plots for the 3-year moving average values as shown in Figure 4 illustrate a funnel type cone that envelops the Versus Fits and Versus Order graphs' data with some clustering at the earlier fitted values of the Versus Fits plot and insignificant residuals for the first 30 in the Versus Order plot.

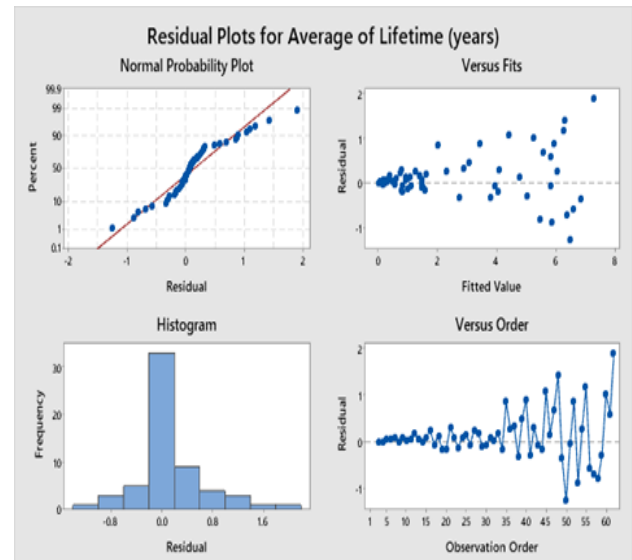


Figure 4: Residual Plots for 3-Year Moving Averages of Yearly Averages of Satellite Lifespans.

3.2.3 Single Exponential Smoothing Modeling:

The basic formula for simple (single) exponential smoothing uses a weighted moving average with exponentially decreasing weights as given by:

$$S_t = \alpha y_{t-1} + (1 - \alpha) S_{t-1}$$

where t = time period,

α = the smoothing constant, a value from 0 to 1,

y_{t-1} is actual data value of previous period $t-1$, and S_{t-1} is the exponential smoothed value of previous time period $t-1$.

When α is closer to zero, smoothing happens more slowly. The best value for α is the one that results in the smallest Mean Squared Error (MSE).

Using Single Exponential Smoothing Modeling with smoothing constant $\alpha = 0.744$ and data scaled on the Y-axis as days instead of years, the MAPE (Mean Absolute Percent Error) is a relatively low 19.9% as shown in Figure 5 below with residual plots for yearly averages of satellite lifetimes as shown in Figure 6.

We anticipate that refining this model by performing a detrending operation as an initial step will result in further reduction in the MAPE, and plan this for a follow-up study. This technique is called Forecast Including Trend (FIT) and is the sum of an exponentially smoothed forecast average with an exponentially smoothed trend.

The residual plots of Versus Fits in Figure 6 show a cluster for about the first quarter of the fitted values followed by a funnel of residuals indicated goodness of fit is decreasing for later values.

The change in granularity of data from years to days in addition to use of single exponential smoothing yielded a relatively low value of the MAPE (Mean Absolute Percent Error). Measuring lifespan with a granularity of years is less accurate because two lifetimes with the same lifespan in years can have quite different lifespans in days.

For example, if one satellite is launched in January and dies in December of the next year, and another satellite is launched in December and dies in January of the next year, they both have lifespans of one year, measured in years, but if measured in days one has a lifespan many times longer than the other.

A partial autocorrelation analysis of the data points in Figure 5 is shown in Figure 7.

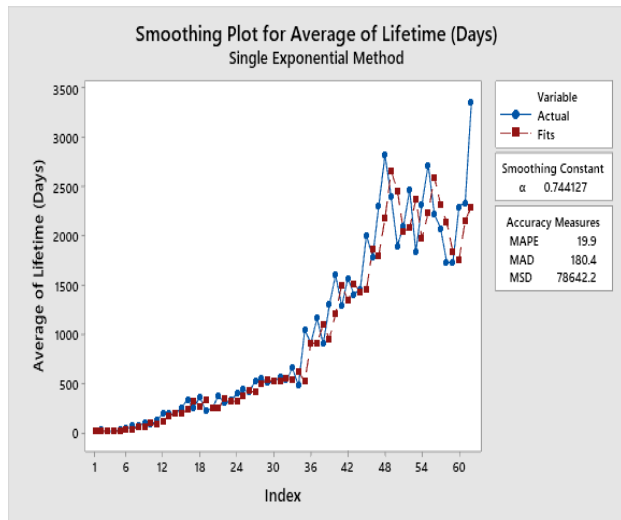


Figure 5: Single Exponential Smoothing Plot for Yearly Averages of Lifetimes Measured in Days.

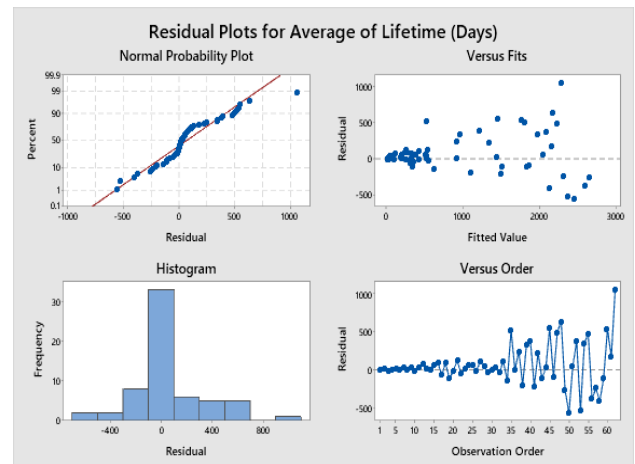


Figure 6: Residual Plots for Yearly Averages of Lifetimes Measured in Days Using Single Exponential Smoothing Modeling.

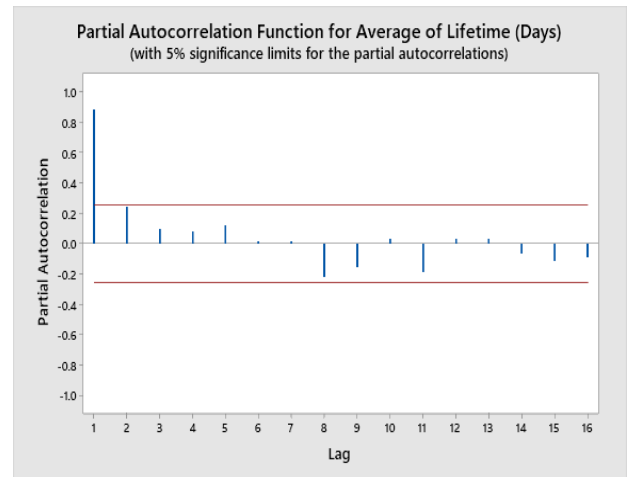


Figure 7: Partial Autocorrelation Function for Yearly Average of Satellite Lifetimes Measure in Days.

3.2.4 Growth Curve Model for Trend Analysis: Figure 8 shows the use of a Growth Curve Model for trend analysis of yearly averages of satellite lifetimes with a MAPE (Mean Absolute Percent Error) of 45%. The Growth Curve Model to which it was fitted is

$$Y_t = 48.006 * (1.07813)^t$$

Figure 9 shows the residual plots for yearly averages of lifetimes measured in days using the Growth Curve Model.

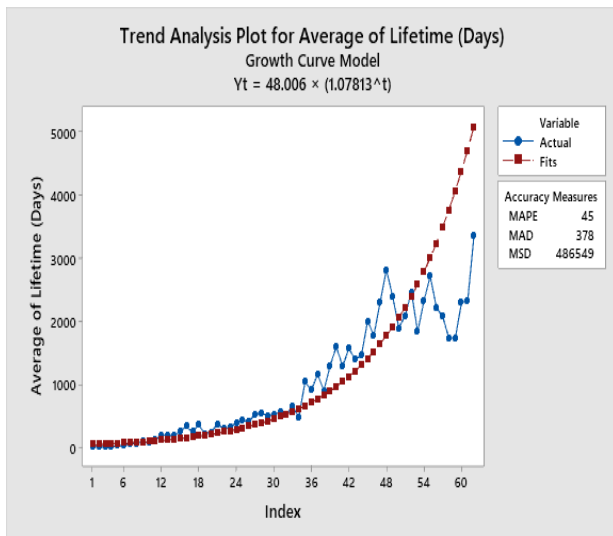


Figure 8: Growth Curve Model for Trend Analysis for Yearly Averages of Lifetimes Measured in Days.

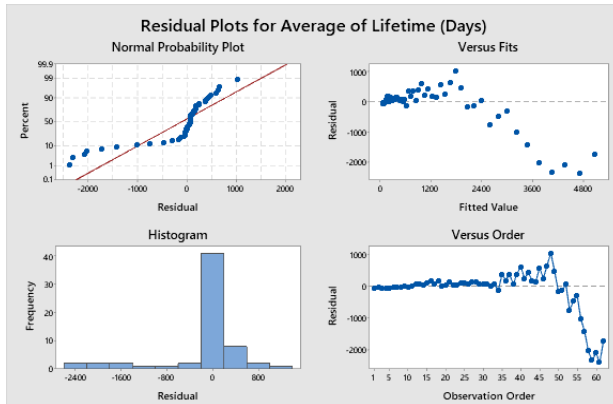


Figure 9: Residual Plots for Average of Lifetimes Using the Fitted Growth Curve Model.

4. CONCLUSIONS

This paper focused on yearly and daily averages for lifetimes of artificial satellites. The lifetimes were calculated based on satellite death date. The lifespan generally increased with respect to death date, but the tested models did not fit exactly.

Measuring lifespans with a granularity of days provides more accurate results than using a granularity of years. This may be due to the previously mentioned issue of some satellites launching late in one year and dying shortly thereafter in the next year.

The linear trend modeling did not yield a good fit to the data. The modeling was substantially improved using 3-year moving averaging and single exponential smoothing.

The use of a Growth Curve Model provided a close fit for the first half of the data and less for the second half. This implies that the satellite launches for these two time periods are not the same, perhaps due to different and new types of missions as the technological age advances with new and different reasons for these launches.

The future directions of this research include the use of additional data and different methods for analysis that give better insights into future satellite lifespans.

5. ACKNOWLEDGEMENTS

The authors thank J. McDowell for kindly providing satellite lifespan data, and Peng-Hung Tsai of University of Arkansas at Little Rock (UALR) for editorial review of final submission.

6. APPENDIX

Appendix Table A: Sample Python code used for data cleansing of data

```
## Reading-in New-Death-data
Death_dates=pd.read_excel('New Death
Data.xlsx')
New Death_data.head()

## Eliminating rows having Death_Month as
'U', 'E', '?'
New Death_data_1=Death_dates.loc[(New
Death_data ['Death_Month'].notnull())&
(~New Death_data
['Death_Month'].isin(['U','E','?'])),:]
New Death_data_1.head()
```

```
## Stripping extra spaces
New Death_data_1['Death_Month']= New
Death_data_1['Death_Month'].str.strip()

## Stripping Special characters
def strip_special(special):
return
special.replace("?", "").replace("Arp", 'Apr')

New Death data_1['Death_Month']=
```

```
New Death
data_1['Death_Month'].apply(strip_special)
```

Appendix Table B: Sample Code in Python for Data Imputation

```
# Imputing Day value in Death Dates
M=[]
for index,row in df.iterrows():
```

```

if pd.isnull(row['Death_Day'])==True:
if (row['Launch_Date'].year ==
row['Death_Year'])&(row['Launch_Month']
==row['Death_Month']):
M.append(round((30-
row['Launch_Date'].day)/2))
else:
M.append(15)
else:
M.append(row['Death_Day'])

```

6. REFERENCES

- [1] Batthula, V.J.R., **Lifespan Analysis of Earth Satellites**, Master of Science Thesis, Department of Information Science, University of Arkansas at Little Rock (UALR), Donaghey College of Engineering and Information Technology, December 2019. 48 pp.
- [2] Berleant, D., Kodali, V., Segall, R., Aboudja, H., & Howell, M.; Moore's Law, Wright's Law and the Countdown to Exponential Space, **The Space Review**: January 7, 2019. Available at <http://www.thespacereview.com/article/3632/1>.
- [3] Duggirala, R., Kodali, V., Segall, R., Berleant, D. & Aboudja, H.; Steps Toward Measuring World Space Exploration Activity, **Presentation at 102nd Annual Meeting of the Arkansas Academy of Sciences**, Jonesboro, AR, April 7, 2018.
- [4] Hall, C., Berleant, D., Segall, R. S., and Lu, S.; Steps Toward Quantifying Advancement in Space Exploration, **Proceedings of the 20th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2017)**, Volume II, WMSCI 2017, Orlando, FL, July 8-11, 2017, pp. 106-112.
- [5] Howell, M., Kodali, V., Kreinovich, V., Aboudja, H., Batthula, V. J. R., Segall, R. & Berleant, D., Trends in Technology Developmet in the US and USSR During the Space Race, **The Space Review**, December 2, 2019. Available at <https://www.thespacereview.com/article/3840/1>.
- [6] Howell, M., Kodali, V., Segall, R., Aboudja, H., and Berleant, D., Moore's Law and Space Exploration: New Insights and Next Steps, **Presentation at Arkansas Academy of Sciences 103rd Annual Meeting**, Hendrix College, Conway, AR, March 29-30, 2019.
- [7] Howell, M., Kodali, V., Segall, R., Aboudja, H., and Berleant, D., Moore's Law and Space Exploration: New Insights and Next Steps, **Journal of Arkansas Academy of Science**, Vol. 73, Article 6, pp. 13-17.2020. Available at <https://scholarworks.uark.edu/jaas/vol73/iss1/6>.
- [8] Kodali, V., Duggirala, R., Segall, R., Aboudja, A., and Berleant, D.; Travel to Extraterrestrial Bodies Over Time: Some Exploratory Analyses of Mission Data. **Journal of Arkansas Academy of Science**, Vol. 72, pp. 58-66, 2019. Available at <https://scholarworks.uark.edu/jaas/vol72/iss1/13>.
- [9] Kodali, V., **Earth Satellite Data Analysis**, Master of Science Thesis, Department of Information Science, University of Arkansas at Little Rock (UALR), Donaghey College of Engineering and Information Technology, May 2018.
- [10] Wikipedia. **Data Cleansing**. Retrieved December 1, 2019 from https://en.wikipedia.org/wiki/Data_cleansing.
- [11] Wikipedia. **List of Satellites**. Retrieved December 1, 2019 from https://en.wikipedia.org/wiki/Category:Lists_of_satellites.
- [12] Union of Concerned Scientists (UCS). Satellite Database currently orbiting earth. Retrieved December 1, 2019 from <https://www.ucsusa.org/resources/satellite-database>.