CS289 Final Project Report - Using Machine Learning to Locate the Planets of the Solar System

I. Vouvakis Manousakis
Department of Civil Engineering
University of California, Berkeley
Berkeley, CA 94720
ioannis_vm@berkeley.edu

J. Lee
Department of Nuclear Engineering
University of California, Berkeley
Berkeley, CA 94720
jwonlee@berkeley.edu

Abstract

In this work, we have recovered the parameters used in the Ptolemaic geocentric model and the Keplerian model. First, the periodicity of Mars was explored with Fourier regression, and the parameters involved in the models were recovered using a combination of gradient descent and line search optimization techniques. Training and test datasets were prepared using Stellarium's remote control API. The results show that the geocentric model can make accurate predictions within a narrow time-span, whereas the heliocentric model generalizes much better.

1 Introduction

For the early deadline project, we had created a large dataset of synthetic astronomical observations of the planets of the solar system, the Sun, and the Moon. Among the recorded quantities were the right ascension RA, and the declination Dec of the considered celestial bodies. The data were collected for a period starting from -1500 BC and ending in 2020 AD. Time was expressed in Julian date (JD), and at least one measurement per celestial body was taken daily. Our goal for the final project was to create models that can accurately predict RA and Dec given a specific time t expressed in JD. We decided to use those quantities since they are essential components for predicting the azimuth and altitude of the celestial bodies in the sky given the longitude/latitude of the observer's location and the sidereal time (i.e., before one can predict the location in the sky, it is essential first to obtain RA and Dec).

All our code is hosted on GitHub, and can be accessed in the following address:
https://github.com/ioannis-vm/CS289_2020_ProjectS_TeamJupyter_Final

2 Method

We used three methods to derive a prediction model: A data lifting and regression approach, nonlinear model fitting using Ptolemy's geocentric model, and nonlinear model fitting using the modern Keplerian solar system model. The following subsections describe those approaches in more detail. For the first two approaches, we used synthetic measures for Mars. For the third, we repeated the process for all the other planets of the solar system.

Since the main goal of our work was to recover parameters that would have been used for the ancient Ptolemaic model, we used data corresponding to a period between 100AD - 110AD for the training and test sets. The Mars subset of the dataset consists of 3652 measurements of RA and Dec, and it was randomly split between 80% and 20% to make both training and test datasets. The entire dataset and the extracted training and test sets can be found in our GitHub repository. The same train and test sets were used by the three methods to evaluate their performance. Additionally, the entire dataset

32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

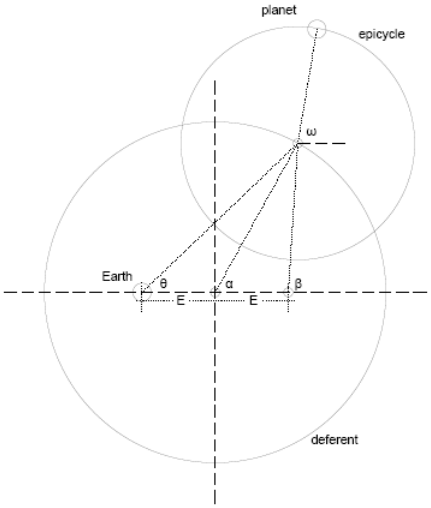

Figure 1: Geocentric Model.

was used to estimate the orbital elements of the planets of the solar system, based on the Keplerian model.

2.1 Regression using Fourier Features

The periodic nature of the celestial bodies' declination and right ascension naturally leads one to choose linear regression with Fourier features. In addition, using sparsity inducing algorithms together with Fourier features, may reveal important parameters such as the orbital periods of the celestial bodies.

We featurized the normalized time (Julian date) with $2n + 1$ features,

$$\left\{ \frac{1}{\sqrt{2n+1}} \right\}_{k=0} \left\{ \frac{\sqrt{2}\sin kx}{\sqrt{2n+1}} \right\}_{k=1}^{k=n} \text{and} \left\{ \frac{\sqrt{2}\cos kx}{\sqrt{2n+1}} \right\}_{k=1}^{k=n}$$

For this study, $n = 50$ was used. Then, to induce sparsity we solved the following Lasso regression problem:

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \frac{1}{2} \|\boldsymbol{y} - \boldsymbol{X}\boldsymbol{w}\|_2^2 + \lambda \|\boldsymbol{w}\|_1$$

For this study, $n = 100$ (201 features) and $\lambda = 0.001$ was used. Since the RA data involves sudden drops representing full revolution of Mars around the Earth, instead of directly using RA values, the cosine of RA was used for training and testing. After training, the absolute values of the learned weights were sorted to identify the most relevant features (frequencies).

2.2 Parameter recovery of the Ptolemaic model using Gradient Descent

Ptolemy's model of the solar system was used for thousands of years to approximate the location of celestial bodies. According to this model, planets travel on a circle called the *epicycle*, which itself orbits another circle called the *deferent* (Figure 1). The deferent's center does not coincide with the Earth but is shifted by a distance E. Opposite to Earth relative to the deferent's center, there is a point called the *equant*, from which the angle to the epicycle's center has a constant angular frequency Ω. The considered planet travels in a circular motion along the epicycle, with an angular frequency ω. The model also involves the angles α and θ, which are defined in the figure. The model can be mathematically formulated as follows:
For a given time t:

Calculate the location of the center of the epicycle on the deferent. Then, determine the location of the celestial body that travels on the epicycle, relative to the center of the epicycle.

$$\sin\alpha = \frac{\sin\left(\Omega t + \Phi\right)}{s - k \cdot \cos\left(\Omega t + \Phi\right)^2} \qquad X = \rho \cdot \cos\theta$$
$$Y = \rho \cdot \sin\theta$$
$$Z = 0$$

$$\theta = \tan^{-1}\left(\frac{\tan\left(\Omega t + \Phi\right)}{1 - \frac{2k}{k\cos\left(\Omega t + \Phi\right) - s}}\right) \qquad x_c = \frac{r}{R}\cos(\omega t + \phi)$$
$$y_c = \frac{r}{R}\sin\left(\omega t + \phi\right)$$
$$\rho = \frac{\sin\alpha}{\sin\theta} \qquad z_c = 0$$

where $s = \sqrt{1 - k^2 \sin\Omega t + \Phi}$. Φ is the phase angle, and $k = \frac{E}{R}$. $\frac{r}{R}$ is the ratio of the radii of the epicycle and the deferent, and ω, ϕ are the corresponding orbital frequency and phase angle related to the epicycle. To orient things in 3D space, the angles i_c, N_c, i_d, N_d are introduced, which define the rotation axis and the inclination of the deferent and epicycle.

$$X_{d,1} = X_d \qquad x_{c,1} = x_c$$
$$Y_{d,1} = Y_d\cos i_d + Z_d \sin i_d \qquad y_{c,1} = y_c \cos i_c + z_c \sin i_c$$
$$Z_{d,1} = -Y_d \sin i_d + Z_d \cos i_d \qquad z_{c,1} = -y_c \sin i_c + z_c \cos i_c$$
$$X_{d,2} = X_{d,1}\cos N_d + Y_{d,1}\sin N_d \qquad x_{c,2} = x_{c,1}\cos N_c + y_{c,1}\sin N_c$$
$$Y_{d,2} = -X_{d,1}\sin N_d + Y_{d,1}\cos N_d \qquad y_{c,2} = -x_{c,1}\sin N_c + y_{c,1}\cos N_c$$
$$Z_{d,2} = Z_{d,1} \qquad z_{c,2} = z_{c,1}$$

Having oriented the points in space, their rectangular coordinates are summed up. Finally, the right ascension and declination are given by the following expressions.

$$X_{tot} = X_{d,2} + x_{c,2} \qquad RA = \tan^{-1}\left(\frac{Y_{tot}}{X_{tot}}\right)$$
$$Y_{tot} = Y_{d,2} + y_{c,2}$$
$$Z_{tot} = z_{d,2} + z_{c,2} \qquad Dec = \tan^{-1}\left(\frac{Z_{tot}}{\sqrt{X_{tot}^2 + Y_{tot}^2}}\right)$$

The parameters of the model, therefore, are Ω, ω, Φ, ϕ, k, $\frac{r}{R}$, N_c, i_c, N_d, and i_d, and our goal was to learn those parameters, given the available $(t, RA - Dec)$ pairs.

2.3 Parameter recovery of the Keplerian model using Gradient Descent

The most accurate closed-form representation of the solar system that is known today is the Keplerian model. The mathematical formulation of this model is derived after simplifying assumptions that enable the analytical solution of the governing differential equations. One such simplifying assumption is that there are no external gravitational interactions, i.e., all positional equations are derived by solving a two-body problem. Due to the large number of equations involved, the model formulation will not be discussed herein, but interested readers are encouraged to look at our references(1) and our Python code, available at our GitHub repository. The parameters of the model involve some common parameters across all celestial bodies and some unique parameters related to Earth specifically. The common parameters are the longitude of the ascending node N, the inclination to the ecliptic i, the argument of perihelion W, the semi-major axis a, the eccentricity e, the mean anomaly M, whose rate of change is the angular frequency of the body's orbit $\frac{\partial M}{\partial t}$. Earth-specific parameters include the obliquity of the ecliptic ecl. Although most of the quantities mentioned above remain fixed over time in the standard formulation, we have allowed them to change over time by considering them as affine functions in time. The entire synthetic observation dataset was used for training. After obtaining a sufficiently accurate model for the Earth's orbit (using the RA-Dec observations of the Sun exclusively), the recovered parameters were subsequently used to train the models for the rest

Table 1: Ptolemaic model parameters

Parameter	Value	Parameter	Value
Ω	$+1.719 \cdot 10^{-2}$	i_d	$+4.018 \cdot 10^{-1}$
ω	$+9.139 \cdot 10^{-3}$	N_c	$+3.141 \cdot 10^{0}$
k	$-1.095 \cdot 10^{-1}$	i_c	$+4.254 \cdot 10^{-1}$
r/R	$+1.386 \cdot 10^{0}$	Φ	$-8.977 \cdot 10^{-1}$
N_d	$+3.271 \cdot 10^{0}$	ϕ	$-1.428 \cdot 10^{-1}$

of the celestial bodies, exclusively using their corresponding synthetic observations. In the end, we recovered expressions for the orbital elements of the considered celestial bodies that can be used to make accurate predictions of their RA and Dec at a given Julian date t.

3 Results

3.1 Regression and parameter recovery using Fourier Features

The Fourier LASSO regression model was implemented using the Scikit-learn Python package. As a result of using LASSO regression, only 19 weights were recovered out of 201 weights. The regression results on the Mars right ascension and declination data are shown in figure 3. The mean square errors were 2.85×10^{-1} and 8.73×10^{-3} for right ascension and declination, respectively. Since the periodic behavior of the training data, one can expect to relate the learned features with the orbital elements. Shown in figure 4 are the learned weights and corresponding Fourier feature periods. Note that the two most prominent feature periods, 1.89 and 1 years/cycle, are in good agreement with the orbit period of Mars and the Earth, 1.88 and 1 years.

3.2 Parameter recovery of the Ptolemaic model using Gradient Descent

The Ptolemaic model was implemented using Python's Pytorch computational package, and the parameters were learned using a combination of non gradient-based line-search and Gradient Descent. This alternation between methods had to be done due to the highly non-convex nature of the loss function with respect to the parameters that correspond to angular frequencies, Ω, and ω, which was preventing the algorithm from converging close to the global minimum. The training was done using the training set described in the methodology section, and the parameters learned are presented in table 1. It is worth noting that even though the epicycle and deferent were allowed to have distinct orientation angles, the orbit's resulting shape is very close to being planar. Figure 5 shows the data and the predicted values, where it is verified that the model's performance is acceptable. It is also interesting to examine the relationship of the learned parameters with the real world. As it can be seen in table 1, $\Omega = 0.01719\ \frac{rad}{day}$, which converted to days over a full cycle is $365.5\ days$, which matches with the period of the Earth's orbit. Similarly, the period derived from $\omega = 0.009139$ is $687.5\ days$, which matches with the period of the orbit of Mars. Furthermore, the ratio $r/R = 1.37$ closely resembles the ratio of the two planets' orbital semi-major axes, $\alpha_{Mars}/\alpha_{Earth} = 1.52$.

It was found, however, that the model does not perform well when doing extrapolation. Attempts to train the predictive model using more than 20 years result in high training error (more than $1.00 \cdot 10^{-2}$). In those cases, the model cannot adequately capture Mars' backward motion every time the phenomenon occurs. This shows that it is difficult for this model to be consistently accurate over periods of time longer than approximately 20 years. When training with a smaller time range, lower training error can be achieved, but the performance of the model deteriorates outside of the training time range, as is shown in figure 2.

3.3 Parameter recovery of the Keplerian model using Gradient Descent

The parameter recovery procedure for the Keplerian model was done in two steps. First, a model was built to recover the parameters of the Earth's orbit around the Sun, using all the available $(t, RA - Dec)$ pairs. Figure 7 shows samples of the residuals (the difference between the actual and the predicted value) for the model of the Earth, ranging over the entire dataset. In the figure,

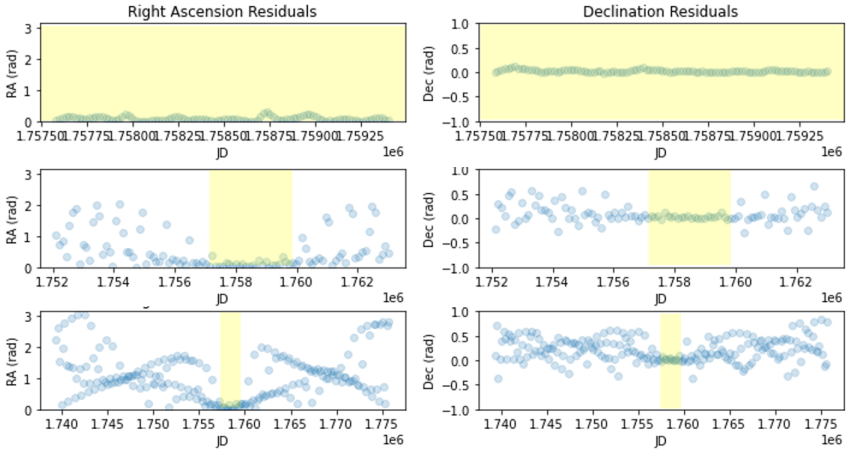

Figure 2: Poor extrapolation performance.

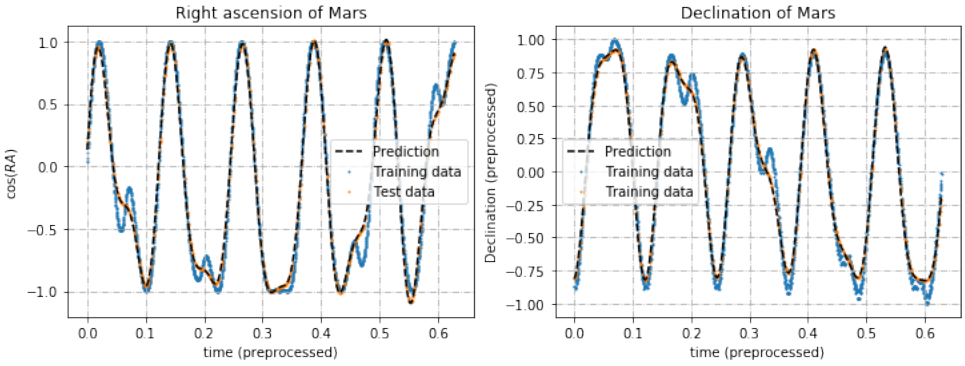

Figure 3: Fourier LASSO regression performance.

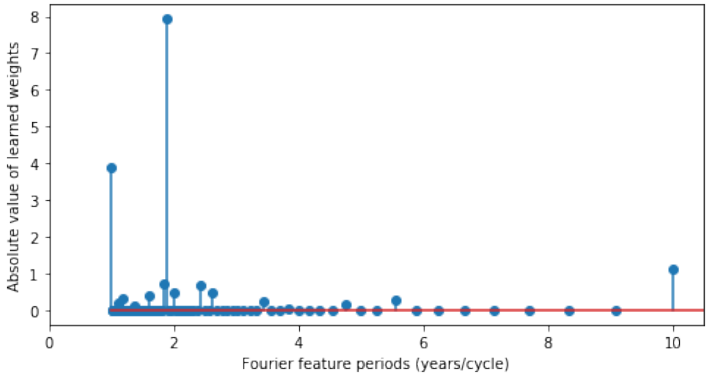

Figure 4: Learned weights of Fourier features. Note that the x-axis is periods of the Fourier features.

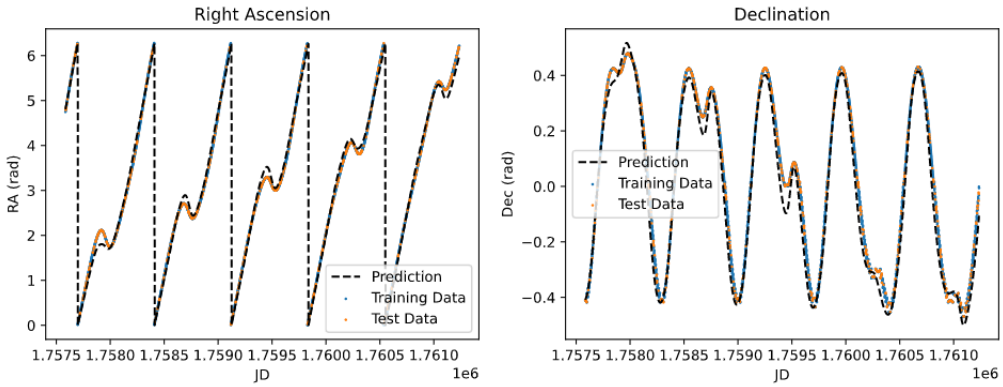

Figure 5: Geocentric model performance.

Table 2: Comparing mean squared error on the test set

Method	Test MSE
Geocentric	$8.107 \cdot 10^{-3}$
Heliocentric	$1.430 \cdot 10^{-3}$

we can see the presence of irreducible error that has accumulated towards the boundaries of the available data domain. A second model was created that treated those parameters as constants, aiming to learn the parameters associated with the rest of the planets (considering one planet at a time). Again, all available $(t, RA - Dec)$ pairs were used in this process. Just as described in the previous section, training had to be done in a series of alternating steps involving grid-search and gradient descent, due to the high non-convexity of the loss function in particular variables (see figure 9). The blue and orange dots correspond to the value of the parameter before and after the line-search step, respectively). At the very last gradient descent part of training, first-order time-varying terms were introduced to all the orbital element parameters, allowing them to change linearly over time, further elevating the prediction accuracy. The learned orbital elements are summarized in Table 3, calculated for January 1st, 2000. To make a fair comparison between the various methods considered, a separate analysis was made, explicitly considering Mars and the training-test set discussed in the previous sections. The test mean-squared error obtained is shown in table 2, and the performance of the model is shown in figure 8.

It is worth noting that even though the final mean-squared error on the training set when considering the entire dataset was at a low level (see table 3), the obtained orbital elements (especially angles associated with the orientation of the elliptical orbits such as the longitude of the ascending node N or the argument of perihelion W), do not match the values that are publicly available (2). The reason for that is that the shapes of the orbits are very close to being circular, and large changes in those quantities result in little change in the path of the planets in space. Other quantities, however, that have a much more direct impact on the location of the celestial bodies, such as the semi-major axis a, or the orbital period T were approximated much more accurately. The orbits of the Planets of the solar system based on the orbital elements derived from this process can be seen in figure 6, where a gray underlay corresponds to the actual orbits based on the official orbital elements. The figure justifies our claim that the deviation is small.

4 Conclusion

The goal of our work was to derive models that can predict the right ascension and declination of celestial bodies given time t, as those quantities are essential for locating the objects in the sky. Utilizing the periodicity of the data, lifting the input data using Fourier features and subsequent LASSO regression were performed. The learned weights revealed quantities that can be meaningfully interpreted. Higher accuracy levels were achieved by fitting models of the solar system to the data. The ancient Ptolemaic geocentric model was able to predict those quantities with high precision for a

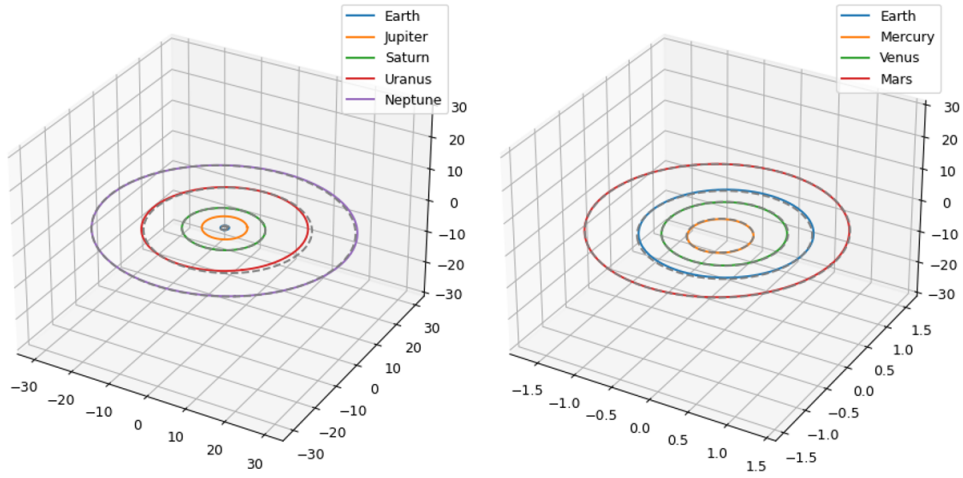

Figure 6: Obtained v.s. actual planetary orbits.

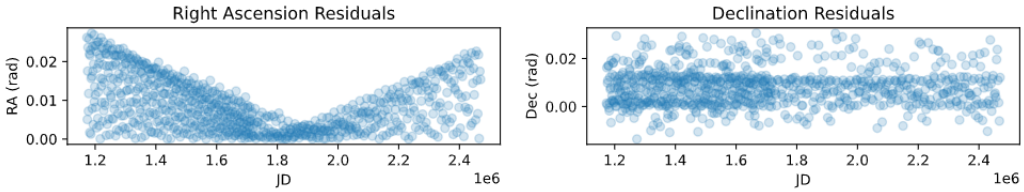

Figure 7: Residuals of the model of the Earth's orbit.

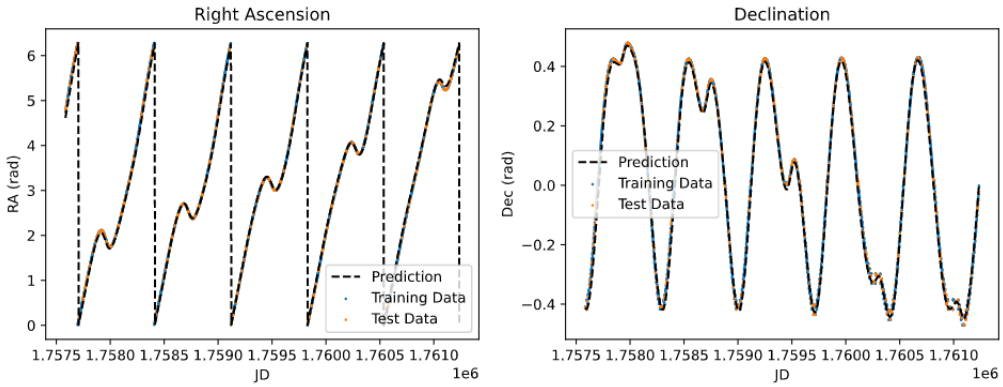

Figure 8: Keplerian model performance.

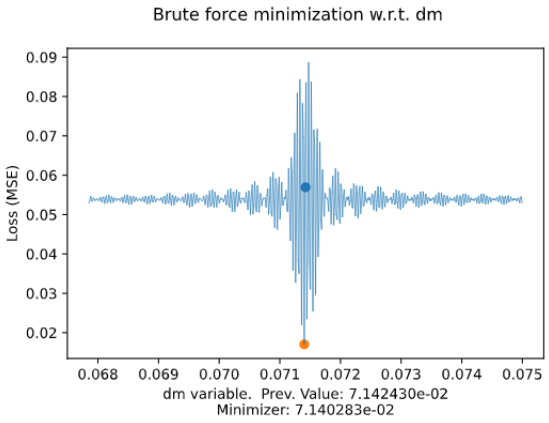

Figure 9: Non-convex loss function.

Table 3: Learned orbital elements

Planet	T (days)	a (AU)	e	i (Deg)	N (Deg)	$\bar{\omega}$ (Deg)	Training MSE
Earth	365.24	1.00	0.0149	0.00	-	280	9.05E-04
Mercury	87.970	0.385	0.0822	5.82	98.8	82.6	8.21E-04
Venus	224.72	0.718	0.00979	3.41	72.9	70.6	7.23E-04
Mars	686.93	1.52	0.0987	1.68	25.8	337	1.39E-03
Jupiter	4329.8	5.20	0.108	358	273	154	9.32E-04
Saturn	10745	9.52	0.0529	2.44	112	91.5	9.63E-04
Uranus	30708.4	19.0	0.001	0.214	359	351	1.63E-03
Neptune	60098.5	30.3	0.00149	358	323	342	1.11E-03

short time-span. The recovered parameters can be reasonably attributed to real-world characteristics of the solar system. The modern Keplerian model was finally used on all the available data. The recovered parameters allow us to define the relative positions of all the planets of the solar system and reconstruct their orbits in 3D space. We find this to be particularly surprising given that we used angular and temporal data exclusively, that can easily be measured from Earth.

Acknowledgments

We would like to thank the astronomer and friend, Christina Dimitreli, for providing us with valuable resources and feedback while we were working on this project.

References

[1] S. Paul, "How to compute planetary positions." [Online]. Available: https://stjarnhimlen.se/comp/ppcomp.html

[2] D. Williams, "Planetary Fact Sheet." [Online]. Available: https://nssdc.gsfc.nasa.gov/planetary/factsheet/

[3] M. El-Abbadi and O. M. Fathallah, Eds., *What happened to the ancient Library of Alexandria?*, ser. Library of the written word, The manuscript world. Leiden ; Boston: Brill, 2008, no. v. 3. v. 1, oCLC: ocn173808202.

[4] A. Jones, "Ptolemy's Ancient Planetary Observations," *Annals of Science*, vol. 63, no. 3, pp. 255–290, Jul. 2006. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/00033790600625427

[5] R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, "Discovering Physical Concepts with Neural Networks," *Physical Review Letters*, vol. 124, no. 1, p. 010508, Jan. 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.124.010508

[6] K. Brown, *Physics in space and time*, 2013, oCLC: 900825631.