Durations of Moon Phases

Time Series Project

Chia-Yu Wei, Ming-Wei Kuo

2025-04-25

1. Introduction

The Moon's orbital cycle is one of the most ancient and observable natural rhythms, dividing each month into four primary phases: New Moon, First Quarter, Full Moon, and Last Quarter. While it is commonly assumed that the durations between these phases are evenly spaced—each representing one-fourth of the lunar cycle—astronomical data suggest that this symmetry does not hold precisely in reality.

This project investigates the durations between consecutive lunar phases using time series analysis to evaluate whether these durations are stable or exhibit meaningful variation. We also examine the total duration of each full lunar cycle to identify underlying trends and long-term fluctuations.

Importantly, the Moon's cycle plays a central role in many traditional calendars, especially in Eastern cultures such as the Chinese, Korean, and Islamic lunar calendars, where months are defined by the timing of the Moon's phases. Understanding the behavior of these durations is not only of scientific interest but also has cultural and historical significance.

To explore this, we employ statistical modeling techniques such as ARIMA forecasting, Fourier transform-based spectral analysis, and residual diagnostics. These tools help us detect periodic components, assess stationarity, and build models for forecasting future lunar behavior based on observed trends

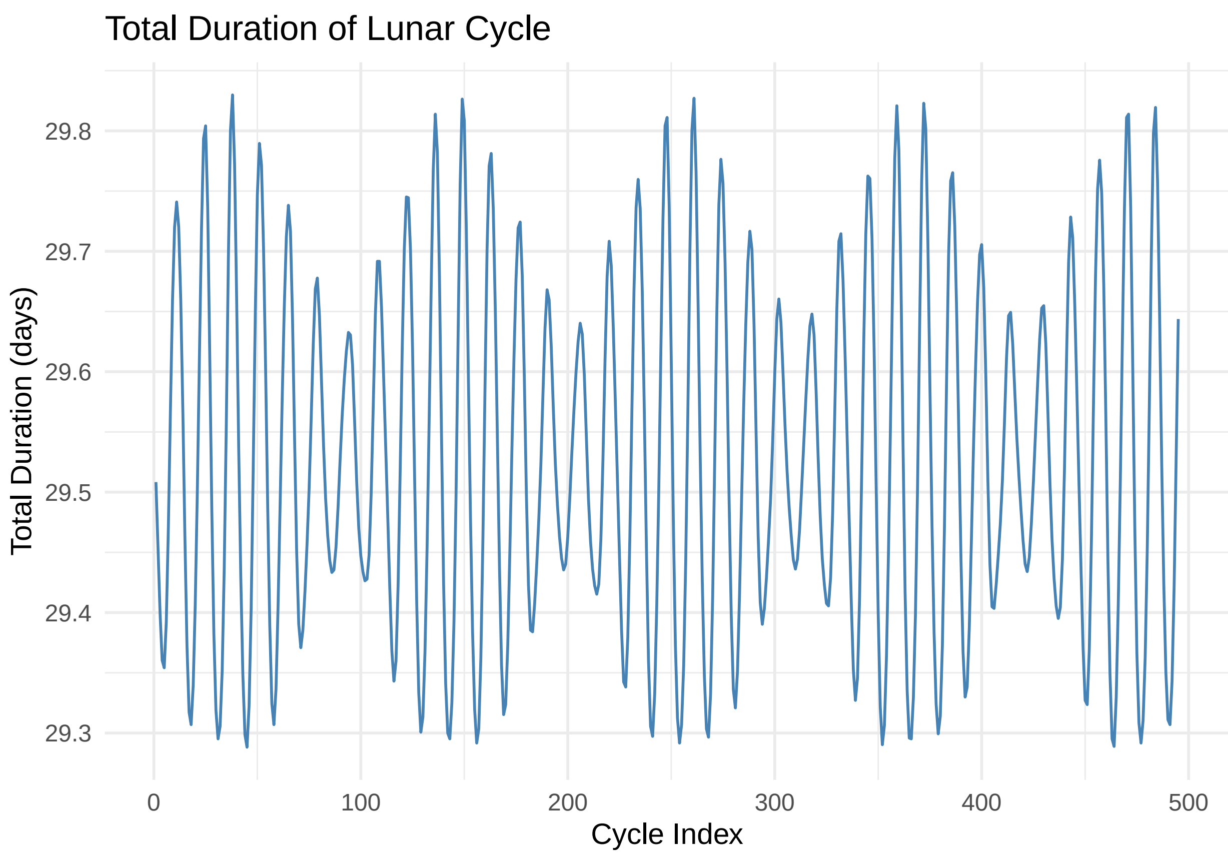
2. Exploratory Data Analysis

The data used in this project was obtained from Astropixels, a website maintained by astronomer Fred Espenak. Astropixels provides detailed astronomical data including lunar phases, eclipses, planetary events, and other celestial phenomena. It also features astrophotography and downloadable ephemeris resources. For this project, we specifically used the Moon Phase Calendar to analyze the durations between lunar phases.The data includes the date and time of the four main stages form 1970 to 2021. We calculated the durations between stages by ourselves.

In this project, we define each time interval between lunar phases in terms of **days**, which is the standard unit used across all time series plots and calculations. And the *cycle* refers to a **full lunar cycle**, which includes all four major lunar phases: New Moon → First Quarter → Full Moon → Last Quarter → back to New Moon.

2.1 Total Duration of Lunar Cycle

Let's take a look at Total Duration first, we can see the total duration subtly rises and falls with certain pattern, which creates a beating pattern. Also, We observe a larger-scale fluctuation, occurring roughly every 100 cycles, forming a clear beating pattern over time.



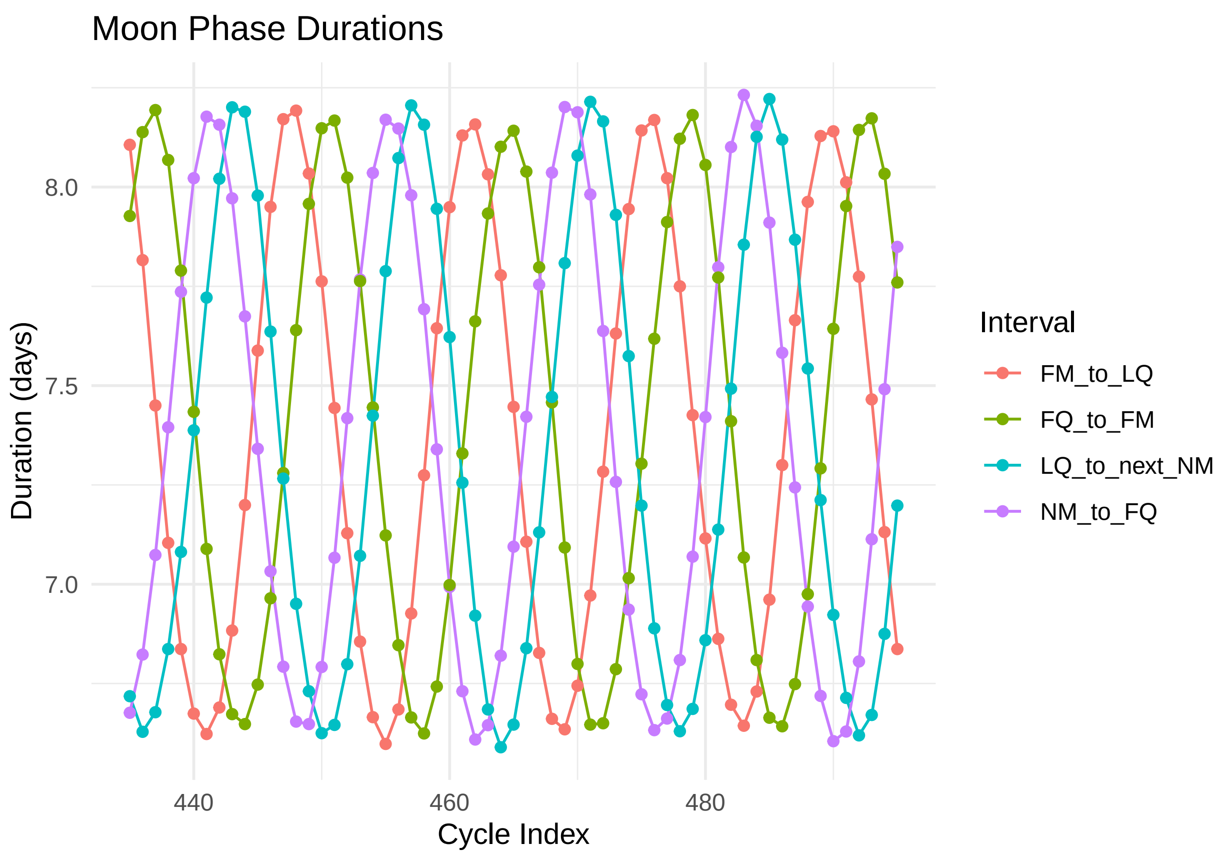
2.1 The durations for each period

Next, we compared the average durations of each phase. They are all around 7.38 days.

Period	Average Duration (days)
New Moon → First Quarter	7.381
First Quarter → Full Moon	7.387
Full Moon → Last Quarter	7.383
Last Quarter → New Moon	7.378
Total Duration	29.529

Although they are all around 7.38 days. But they are not equal at the same time, it appears to follow a structured wave, likely influenced by the Moon's elliptical orbit.

We observe a consistent wave-like fluctuation pattern across the four lunar phase durations: `NM_to_FQ`, `FQ_to_FM`, `FM_to_LQ`, and `LQ_to_next_NM`. Notably, the peaks and troughs appear to cycle in a fixed sequence. For instance, when `FM_to_LQ` increases, it is typically followed by an increase in `FQ_to_FM`, then `NM_to_FQ`, and finally `LQ_to_next_NM`. This repeating order suggests that the durations are not change independently.


Moon Phase Durations
Duration (days)
8.0
7.5
7.0
440
460
480
Cycle Index
Interval
FM_to_LQ
FQ_to_FM
LQ_to_next_NM
NM_to_FQ

3. Statistical Testing

3.1 One-way ANOVA

To statistically assess whether the time durations between each pair of consecutive lunar phases are significantly different, we performed a one-way ANOVA (Analysis of Variance) test. The dataset was first reshaped into a stacked format, combining the four phase-to-phase duration series. The ANOVA test was then applied to compare the means across these four groups.

The null hypothesis of this test assumes that all four phase durations have the same mean. A significant result (p-value < 0.05) would indicate that at least one of the phase durations differs significantly from the others, providing statistical evidence against the naive assumption that the lunar cycle is divided into equal quarters.

```
##               Df Sum Sq Mean Sq F value Pr(>F)
## ind            3    0.0 0.00708   0.023  0.995
## Residuals   1976  612.5 0.30997
```

3.2 Stationary Check

We chose one fo the segment to test stationarity, and got the conclusion that it is stationary because the p-value is smaller than 0.05.

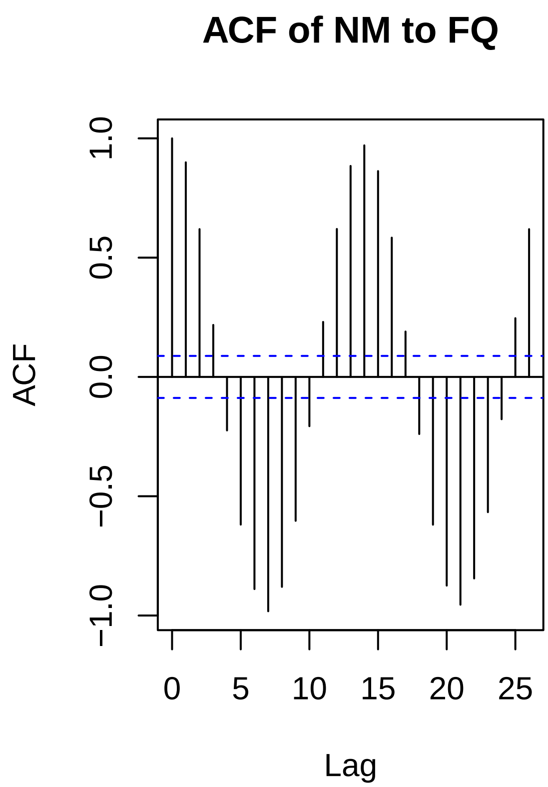
```
##
##  Augmented Dickey-Fuller Test
##
## data:  ts_nm_fq
## Dickey-Fuller = -12.852, Lag order = 7, p-value = 0.01
## alternative hypothesis: stationary
```

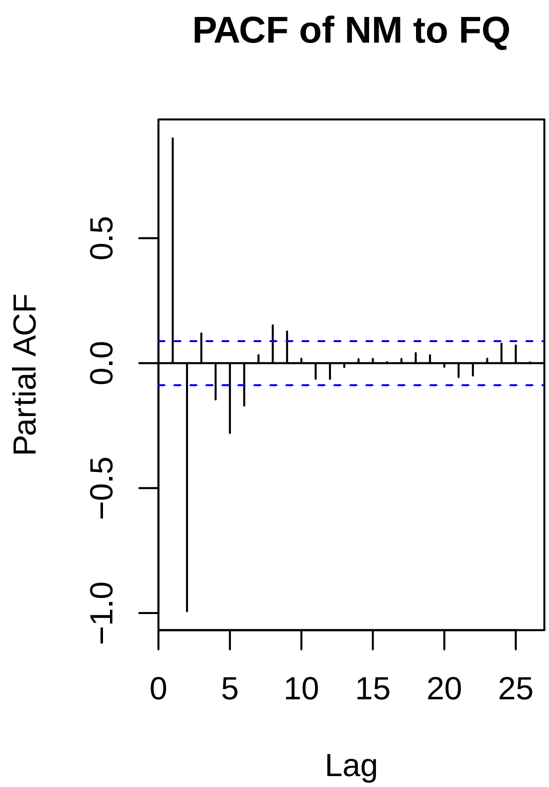
4. Time Series Analysis

4.1 ACF/PACF:

We focused on the NM to FQ series and examined its autocorrelation.

- The **ACF** plot shows significant spikes, which means that each value in the series is influenced not just by its immediate past, but by several previous values.
- The **PACF** plot also shows several significant spikes, especially at lower lags, which suggests a possible AR component in the data.





4.2 ARIMA Model

Firstly, We tried to fit the model. We tried AR term from 0 to 15, MA parameter from 0 to 5 to get the best model, and the best-fit model based on AIC was ARIMA(13, 0, 5) with AIC = -4585.761.

Secondly, we applied Ljung-Box test to test whether residuals from a time series model are independent (i.e., white noise).

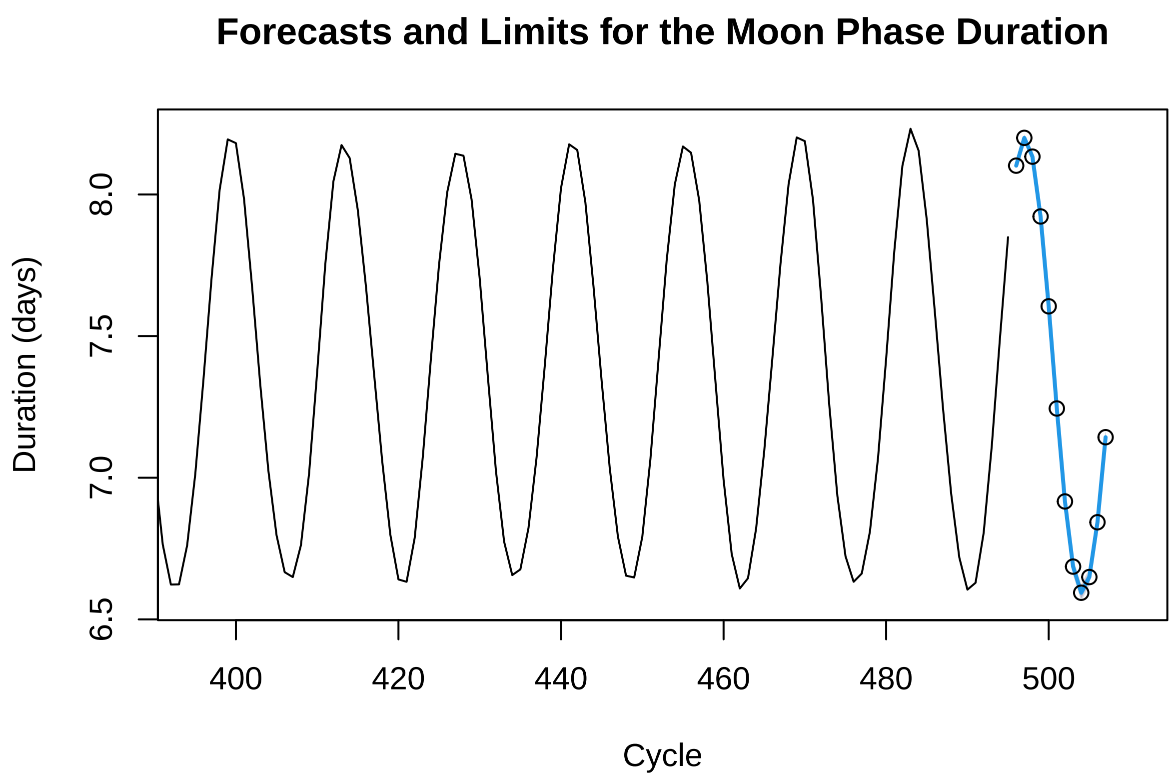
Residuals still show significant autocorrelation for all model we tried!

5. Forecast

5.1 Next one year forecast

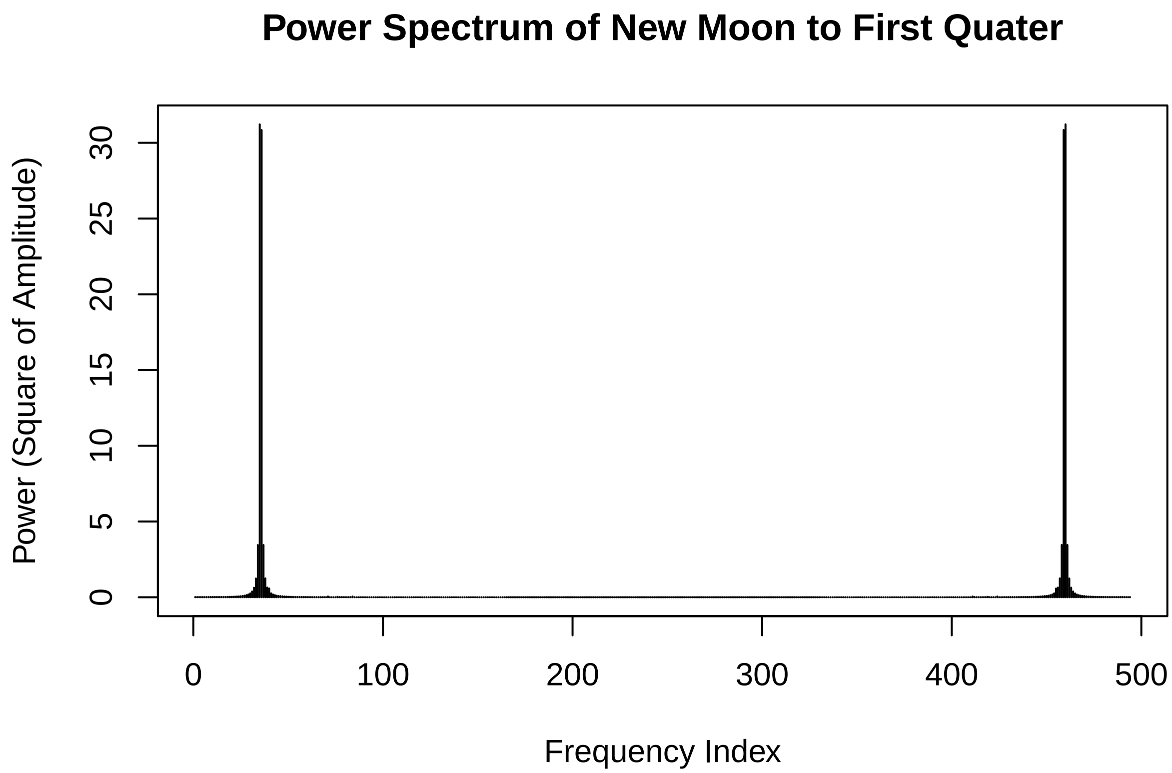
Then we used the fitted ARIMA(13, 0, 5) model to forecast the next 12 lunar cycles. The blue portion on the right represents the 12-step-ahead forecast, with confidence intervals shown as a shaded area. We don't see shaded area because the confidence interval is to narrow. We can observe that the forecast continues the smooth periodic trend, which matches the underlying moon phase cycle very well. The confidence interval is pretty narrow, indicating low variance and strong model confidence.

Overall, this result shows that the ARIMA model effectively captures the moon's cyclic duration behavior and can reasonably forecast future phase lengths.



5.2 Fourier Transform

To explore deeper periodic patterns, we applied a Fourier Transform. The power spectrum reveals a strong peak at frequency index 35.5.



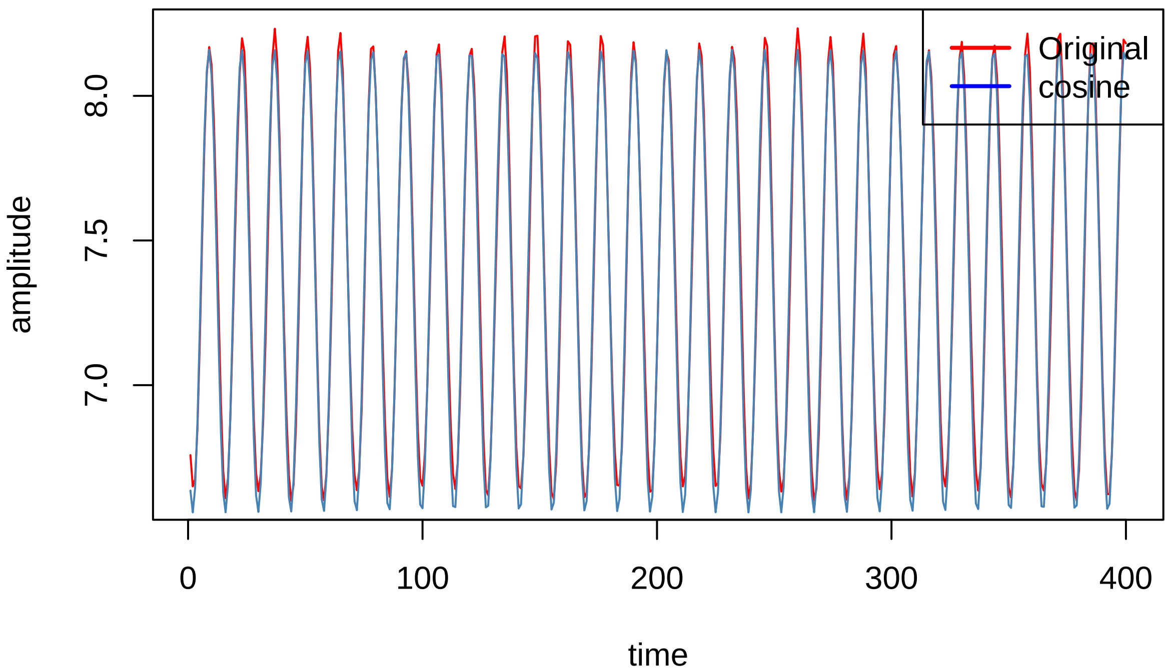
Estimated Period:

$$\frac{N}{index} = \frac{495}{35.5} = 13.94$$

This suggests a repeating cycle approximately every 13.94 observations. Our data set is structured by lunar phase segments (with four segments per full cycle). Thus, this result implies that the pattern repeats approximately every **14** full lunar cycles. This aligns with the moon's orbital mechanics and longer-period irregularities.

We then reconstructed a cosine wave with the original data. As shown in the plot, the cosine wave closely matches the pattern in the original series. This confirms that the data follows a strong and stable periodic structure, which aligns with the moon's orbit and explains the success of AR terms like lag 13 in our earlier ARIMA model.

$$y = 7.38 + 0.8 * cos(\frac{2\pi t(35.5)}{495})$$



amplitude
8.0
7.5
7.0
0
100
200
300
400
time
Original
cosine

6. Conclusion

Durations are not equal.

The average durations of each phase differ slightly (all around ~7.38 days), but they fluctuate in a cyclical and coordinated manner. This confirms that the durations are not fixed and deviate from the "ideal" quarter-moon expectations.

Total cycle is consistent (~29.5 days)
Despite fluctuations within each phase, the total lunar cycle duration remains remarkably consistent over time.

Phase durations are interdependent.
There is a clear wave pattern in which the four durations rise and fall in a fixed sequence, suggesting compensation behavior — when one phase duration increases, others adjust to maintain the total cycle length.

ARIMA models detect structure but not fully explain residuals.
An ARIMA(13, 0, 5) model best fits the NM_to_FQ series based on AIC, but Ljung-Box tests showed that residuals still contain autocorrelation, meaning the model cannot capture all the dynamics.

Fourier analysis reveals a strong periodic signal.
A dominant frequency at index 36 indicates a recurring cycle every ~14.1 observations (or ~411 days), matching the observed oscillation in moon phase durations. This supports the presence of long-term cyclic behavior likely driven by astronomical mechanics, such as the moon's elliptical orbit and varying speed.

The durations between moon phases are not only unequal but also follow a stable, rhythmic oscillation. The moon's non-uniform motion leads to redistributions of time across each segment of the cycle, and this redistribution is reflected in the data's periodic structure. While ARIMA models help capture short-term fluctuations, frequency-domain analysis (FFT) reveals the deeper, long-term cyclical patterns that standard models may overlook.

7. Discussion

This study explores the lunar phase durations using time series analysis techniques, revealing both short-term variations and long-term cyclical patterns in the data. The Fourier analysis identified a dominant periodic component with an estimated cycle length of approximately 14 lunar months. This aligns with astronomical expectations, as similar periodic behaviors are known to emerge due to the Moon's elliptical orbit and complex gravitational interactions.

Although our ARIMA and SARIMA models captured the main short-term fluctuations and produced reasonable forecasts, the residual diagnostics—particularly the Ljung-Box test—suggest remaining autocorrelations, implying that more complex or nonlinear models might better explain the dynamics.

In addition, while the phase durations appear to compensate each other over time to maintain a consistent full cycle (~29.5 days), the total cycle length itself shows subtle long-term fluctuations with beating patterns, possibly related to longer periodic influences.

7.1 Limitations

1. The dataset only covers a limited historical period and is restricted to four-phase transition timestamps.
2. Astronomical factors such as nodal precession or perigee cycles were not explicitly modeled.

7.2 Future Work

1. Extend the analysis to the total lunar cycle and include Fourier terms directly in SARIMA modeling.
2. Investigate nonlinear models (e.g., TAR or GARCH) for better capturing residual dynamics.
3. Incorporate astronomical simulations or external covariates to enrich the time series modeling framework

Reference

Petersen, Fred. *Catalog of Moon Phases*. Retrieved April 2025, from https://astropixels.com/ephemeris/phasescat/phasescat.html