# Ground-based Telescope Data Processing Methods: Frequency Domain VS. Time Domain

**Authors:** Xiong Qiang, Li Tanda, Xiong Qiang

**Date:** 2025-04-24T00:00:00+00:00

## Abstract

Ground-based time-domain observations are often plagued by data gaps due to day-night cycles and weather effects, resulting in low duty cycles (typical values around 0.30), which significantly impacts time-domain astronomical research. To compare the performance of frequency-domain and time-domain analysis methods in handling such gapped time-domain data and their applicability in asteroseismology, we employed the Lomb-Scargle algorithm and Inpainting interpolation as frequency-domain approaches, and Gaussian Processes (GP) as a time-domain approach, to analyze simulated time-domain photometric data exhibiting solar-like oscillation characteristics with duty cycles ranging from 0.20 to 0.50. The results demonstrate that the Gaussian Processes method achieves the best performance in both accuracy and stability for restoring true values, outperforming both the Lomb-Scargle and Inpainting methods. The Inpainting method may introduce numerous spurious signals when processing low duty cycle data, causing interference with signal measurements. Therefore, Gaussian Processes is the preferred method for analyzing low duty cycle data from ground-based telescopes, Lomb-Scargle is the second choice, while Inpainting is not recommended.

## Full Text

## Preamble

**Vol. No.**
**Acta Astronomica Sinica**

### Processing Ground-Based Telescope Data: Frequency Domain vs. Time Domain

XIONG Qiang[1], LI Tan-da[1,2]†
(1 School of Physics and Astronomy, Beijing Normal University, Beijing 100875)

(2 Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206)

## Abstract

Time-domain observations from ground-based telescopes are frequently affected by the day-night cycle and weather conditions, resulting in data gaps and low duty cycles (typically around 0.30), which significantly impacts time-domain astronomical research. To compare the performance and asteroseismic applicability of frequency-domain versus time-domain analysis methods for handling such gapped data, we employed the Lomb-Scargle algorithm and Inpainting interpolation as frequency-domain approaches, and Gaussian Processes (GP) as a time-domain approach to analyze simulated photometric data exhibiting solar-like oscillations with duty cycles ranging from 0.20 to 0.50. The results demonstrate that the Gaussian Process method achieves the highest accuracy and stability in recovering true values, outperforming both Lomb-Scargle and Inpainting methods. The Inpainting method, particularly when applied to low-duty-cycle data, tends to introduce numerous false signals that interfere with measurements. Therefore, the Gaussian Process method is the preferred choice for analyzing low-duty-cycle ground-based telescope data, followed by the Lomb-Scargle method, while the Inpainting method is not recommended.

**Keywords:** methods: data analysis, time-domain astronomy, stars: oscillations, stars: solar-type

## 1. Introduction

High-precision time-domain astronomical research represents a frontier and active area in modern astrophysics, encompassing asteroseismology [?], exoplanet detection [?], and stellar rotation measurements [?]. However, high-precision, continuous photometric data typically originate from space telescopes such as Kepler [?] and TESS (Transiting Exoplanet Survey Satellite) [?]. Despite their transformative impact, these space-based missions have limitations. For instance, Kepler' s approximately four-year observation baseline provides insufficient frequency resolution for studying long-period variables with pulsation cycles spanning days to years. In this context, ground-based telescope data remain invaluable. Projects like OGLE (Optical Gravitational Lensing Experiment) [?], MACHO (Massive Compact Halo Objects project) [?], and ASAS-SN (All-Sky Automated Survey for SuperNovae) [?] offer substantially longer observational baselines, dramatically improving frequency resolution for long-period variable studies and complementing space telescopes.

Although ground-based telescopes provide long-baseline time-domain data, weather conditions and the day-night cycle inevitably create numerous gaps, yielding low duty cycles for single-site observations. Here, duty cycle refers to the fraction of time a celestial object is in an active state (e.g., exhibiting periodic phenomena) relative to the total observation period. For long-term projects like OGLE, typical duty cycles are approximately 0.30. Multi-site campaigns can improve this to 0.60–0.95 depending on the number and geographic distribution of stations. These gaps profoundly affect frequency-domain analyses. Fourier transforms of gapped data introduce windowing effects, convolving the window function' s spectrum with the true signal spectrum and generating sidelobes that spread spectral power away from main peaks, complicating signal measurement. Three main strategies address this challenge: (1) direct use of the Lomb-Scargle algorithm [?, ?] designed for unevenly sampled data; (2) gap preprocessing before frequency analysis, such as gap removal [?] or interpolation using existing data points [?]; and (3) time-domain analysis using Gaussian Processes (GP) [?], as demonstrated by Pereira et al. [?] who modeled granulation and oscillations directly in the time domain, finding it robust against gaps. Similar conclusions were reached by Hey et al. [?], who applied GP to sparse TESS data, highlighting its potential for asteroseismic analysis of gapped data.

This study focuses on simulated ground-based telescope data with duty cycles of 0.20–0.50, systematically evaluating frequency-domain and time-domain analysis approaches. We employ three specific methods: the Lomb-Scargle algorithm and Inpainting for frequency-domain analysis, and Gaussian Processes for pure time-domain analysis.

## 2. Methods

### 2.1 Data

We utilize simulated light curves of solar-like oscillating stars. Solar-like oscillations are stochastically excited by turbulent convection in the outer layers, sharing the Sun' s excitation mechanism. Their stochastic nature—random excitation and damping—produces unstable frequencies with random amplitude and phase variations, making them particularly sensitive to data gaps and demanding robust analysis methods. Our simulated data comprise three components: granulation, oscillations, and white noise.

Granulation describes the surface convection cell patterns in solar-like stars, appearing as fine bright and dark structures. Its power spectrum can be approximated by the Harvey profile [?]:

$$S(f) = \frac{2A_{\text{gran}}}{1 + (f/f_{\text{gran}})^4}$$

where $f_{\text{gran}}$ is the characteristic frequency and $A_{\text{gran}}$ is the amplitude at that frequency.

Solar-like oscillations are standing pressure modes (p-modes) or mixed pressure-gravity modes damped by near-surface convection. They are excited within a specific frequency range with roughly bell-shaped amplitudes and can be approximated in the power spectrum by Lorentzian profiles:

$$S(f) = \frac{A(\gamma f_0)^2}{(f^2 - f_0^2)^2 + (\gamma f)^2}$$

where $f_0$ is the central frequency of each mode, $\gamma$ is the full width at half maximum (FWHM), and $A$ is the amplitude at the central frequency. We configured three modes following Equation (2) to simulate pressure-mode oscillations (hereafter "signals" ). The true parameter values are listed in Table 1 , where subscripts "left," "central," and "right" correspond to signals based on their $f_0$ values.

We added normally distributed white noise with mean 0 and variance 72 ppm ($N(0, 72)$) to simulate realistic ground-based signal-to-noise ratios. Following OGLE data, we set the total light curve duration to approximately 30 years (0-11,000 days) with a sampling rate of about 2 days, yielding 5,000 data points. An example light curve and its power spectrum are shown in Figure 1 [Figure 1: see original paper].

To generate realistic gaps, we constructed observation windows as:

$$\text{window}(t) = \begin{cases} 1 & \text{if data exists at time } t \\ 0 & \text{if no data at time } t \end{cases}$$

We followed OGLE data patterns through this procedure: (1) obtained high-duty-cycle data for a real star from the OGLE-III database (Figure 2a [Figure 2: see original paper]), extracted its observation window (Figure 2b); (2) repeated this window to match our 11,000-day baseline; (3) introduced one large (~1000-day) and three small (~200-day) gaps to simulate complex scenarios (Figure 2c); (4) randomly selected start indices and set subsequent 100-point segments to 0 or 1 until achieving the target duty cycle ($±$1% tolerance). Combining these windows with our synthetic light curves produced gapped data at specified duty cycles.

We generated light curves with duty cycles from 0.20 to 0.50. For each duty cycle, we selected two curves based on whether oscillation signals were visually identifiable in Lomb-Scargle periodograms, labeling them Flag=good and Flag=bad. This distinction allows comparison of method performance under varying data arrangements, noise realizations, and window effects. All light curves used in this study are shown in Figure 3 [Figure 3: see original paper].

### 2.2 Analysis Methods

**2.2.1 Lomb-Scargle** Unlike the Fast Fourier Transform (FFT), which requires evenly spaced data, the Lomb-Scargle algorithm handles non-uniform sampling by fitting sinusoids to the data in the time domain for each frequency. Better fits (smaller residuals) produce larger amplitude values in the frequency domain, enabling it to tolerate data gaps. For gap-free data, Lomb-Scargle and FFT yield identical results.

**2.2.2 Inpainting** This method aims to recover an idealized light curve using observed values and the observation window. It applies a specific transform matrix to gapped data and iteratively seeks the sparsest solution (most zero coefficients). For pure sinusoidal variations, Fourier space provides sparse representation, making Fourier transform the appropriate transform matrix. Detailed descriptions appear in [**?**]. Figure 4 [Figure 4: see original paper] shows Inpainting results for a duty cycle 0.30, Flag=good light curve and its subsequent Lomb-Scargle power spectrum.

**2.2.3 Gaussian Processes** Gaussian Processes are non-parametric Bayesian statistical methods widely used in machine learning [**?**]. A GP is a stochastic process where outputs at input points (e.g., time or wavelength) are normally distributed random variables, with any finite subset following a multivariate Gaussian distribution. The process is defined by mean and covariance functions:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

where $m(x)$ is the mean function (set to 0 in this work) and $k(x, x')$ is the covariance (kernel) function describing similarity between input points. With appropriate kernel selection and parameter optimization, GP can directly recover target parameters in the time domain.

We used the SHO (Simple Harmonic Oscillator) kernel from the CELERITE library [**?**, **?**], which describes damped harmonic oscillation—physically analogous to solar-like oscillations. Its power spectral density is:

$$S(\omega) = \sqrt{\frac{2}{\pi}} \frac{S_0 \omega_0^4}{(\omega_0^2 - \omega^2)^2 + \omega_0^2 \omega^2 / Q^2}$$

where $S_0$ scales the power, $\omega_0$ is the characteristic angular frequency, and $Q$ is the quality factor describing damping. The parameters relate to those in Equations (1) and (2) as [**?**]:

$$A_{\text{gran}} = \frac{1}{2\pi} S_{\text{gran}}; \quad f_{\text{gran}} = \frac{2\pi}{\omega_{\text{gran}}}; \quad f_0 = \frac{2\pi}{\omega_0}; \quad \gamma = \frac{\omega_0}{2\pi Q}; \quad A = 4 S_0 Q^2$$

where $S_{\text{gran}}$ and $\omega_{\text{gran}}$ correspond to the $Q = 1/\sqrt{2}$ case in Equation (5).

### 2.3 Fitting

Lomb-Scargle and Inpainting fitting occur in the frequency domain. Granulation follows Equation (1), oscillations follow Equation (2), and a constant term ($W$) accounts for white noise:

$$\text{Model} = \frac{2A_{\text{gran}}}{1 + (f/f_{\text{gran}})^4} + \sum_i \frac{A_i(\gamma_i f_{i,0})^2}{(f^2 - f_{i,0}^2)^2 + (\gamma_i f)^2} + W$$

Gaussian Process fitting occurs in the time domain using one $Q = 1/\sqrt{2}$ SHO term for granulation, three standard SHO terms for oscillations, and a constant ($W$) for white noise.

All three methods use identical uniform priors: $\ln W \sim [0, 5]$, with other parameters and priors listed in Table 2 . We sampled posterior distributions using PYMC3 [**?**] with Markov Chain Monte Carlo (MCMC). Two chains were run with 1,000 tuning steps (optimizing sampler hyperparameters) followed by 1,000 draws. Convergence was assessed via the Gelman-Rubin statistic ($\hat{R}$); $\hat{R} \approx 1$ indicates convergence, with posterior medians serving as best estimates. Figure 5 [Figure 5: see original paper] shows fits for duty cycle 0.50, Flag=good data. For GP, Figure 6 [Figure 6: see original paper] illustrates the time-domain fit (3000–8000 days shown for clarity).

## 3. Results

We measured granulation frequencies/amplitudes and oscillation central frequencies, amplitudes, and FWHMs for duty cycles 0.20–0.50 with both Flag=good and Flag=bad classifications. All results are compiled in the Appendix. Best estimates are posterior medians with uncertainties expressed as median absolute deviations.

### 3.1 Frequency

Figure 7 [Figure 7: see original paper] shows normalized frequency measurement deviations relative to true values for all methods across duty cycles. Upper and lower panels correspond to Flag=good and Flag=bad data, respectively. Reference lines mark zero deviation (black) and tolerance ranges (gray): $±$0.2, $±$0.05, $±$0.01, $±$0.02 for granulation, left, central, and right signals (Flag=good); $±$0.1, $±$0.1, $±$0.02, $±$0.01 for Flag=bad.

**Flag=good data:** 1. **Granulation frequency:** GP shows highest accuracy (deviations <20% across all duty cycles) and stability (standard deviation 9.8%), outperforming Lomb-Scargle (11.2%) and Inpainting (26.1%). Lomb-Scargle consistently outperforms Inpainting, which approaches 90% deviation at duty cycle 0.20. 2. **Oscillation frequencies:** GP achieves the best accuracy (deviations <5% for all signals) and stability (standard deviations: 2.5%, 0.8%, 1.3% for left, central, right signals). Lomb-Scargle outperforms Inpainting, especially at duty cycle 0.20, though its stability is slightly inferior to GP (4.4%, 0.8%, 1.5%). Inpainting shows the poorest stability (12.2%, 3.0%, 2.3%).

**Flag=bad data:** 1. **Granulation frequency:** GP maintains superior accuracy (<10% deviation) and stability (6.2% standard deviation) versus Lomb-Scargle (8.7%) and Inpainting (35.3%). 2. **Oscillation frequencies:** GP again performs best, with left signal deviations <10%, central <2%, and right <1-2%. Lomb-Scargle generally outperforms Inpainting except for the right signal at duty cycles 0.30 and 0.40. GP demonstrates optimal stability (2.3%, 0.6%, 0.9%), while Lomb-Scargle shows poor stability for the left signal (17.8%) but better than Inpainting for central (1.2% vs 3.0%) and right (2.5% vs 1.2%) signals.

Overall, GP consistently delivers superior accuracy and stability for frequency measurements across both data quality classifications, with Lomb-Scargle outperforming Inpainting.

### 3.2 Amplitude

Figure 8 [Figure 8: see original paper] presents normalized amplitude deviations. Reference tolerances are $±$0.2, $±$0.5, $±$0.2, $±$0.4 for Flag=good and $±$0.2, -0.5, $±$0.3, $±$0.5 for Flag=bad.

**Flag=good data:** 1. **Granulation amplitude:** GP shows highest accuracy (deviations <20%) and stability (4.0% standard deviation). While Inpainting outperforms Lomb-Scargle at duty cycles 0.3–0.5, it exhibits large deviations at 0.20. Lomb-Scargle and Inpainting show poorer stability (8.2% and 70.1%). 2. **Oscillation amplitudes:** GP achieves best accuracy (left: <50%, often <10% at 0.30 and 0.50; central: <20%; right: <40%) and overall stability. Lomb-Scargle generally outperforms Inpainting but is inferior to GP for central (14.4% vs 8.9%) and right (29.2% vs 23.0%) signals. Inpainting shows the worst stability (63.9%, 20.0%, 61.0%).

**Flag=bad data:** 1. **Granulation amplitude:** GP maintains superior accuracy (<20%) and stability (8.8%). Inpainting outperforms Lomb-Scargle at 0.30–0.50 but shows catastrophic deviations (~170%) at 0.20. 2. **Oscillation amplitudes:** All methods show large deviations (>30%) for left and right signals, with only Lomb-Scargle achieving <10% for the right signal at duty cycle 0.50. For the central signal, GP shows smallest deviations (<30% except -35% at 0.50). Stability analysis reveals mixed results: Lomb-Scargle performs best for the left signal (8.7% vs 13.2% and 14.0%), while GP excels for central (4.2%) and right (6.0%) signals.

In summary, GP demonstrates superior accuracy and stability for granulation amplitude measurements across both data types. For oscillation amplitudes, GP performs best for Flag=good data, while all methods struggle with accuracy for Flag=bad data, though GP maintains the best stability.

### 3.3 Full Width at Half Maximum

Figure 9 [Figure 9: see original paper] shows normalized FWHM ($\gamma$) deviations. Reference tolerances are $±$0.5, $±$0.2, $±$0.2 for Flag=good and 0.5, $±$0.5, $±$0.5 for Flag=bad.

**Flag=good data:** All methods perform poorly for the left signal, with Inpainting exceeding 200% deviation at duty cycle 0.20. GP shows superior accuracy, particularly for the central signal (<10% deviation at 0.30–0.50) and right signal (<20% across all duty cycles). Stability analysis reveals GP performs best for left (42.7%) and right (15.0%) signals, while Inpainting shows lowest standard deviation for the central signal (18.3% vs 40.0% and 48.9%).

**Flag=bad data:** All methods exhibit large deviations and low overall accuracy. However, GP demonstrates optimal stability across all signals (31.7%, 31.3%, 51.3%), while Lomb-Scargle shows the poorest stability (97.5%, 78.8%, 92.7%).

Overall, GP provides the best stability for FWHM measurements, with superior accuracy for Flag=good data.

### 3.4 False Signals

The Inpainting method introduces substantial false signals when applied to low-duty-cycle data (e.g., 0.30). As shown in Figure 10 [Figure 10: see original paper], the original Lomb-Scargle spectrum displays severely suppressed, barely identifiable signals due to gaps. After Inpainting, multiple spurious peaks appear near the true frequencies (marked by arrows), interfering with signal identification and potentially leading to erroneous stellar parameter determinations. Consequently, Inpainting is not recommended for low-duty-cycle ground-based telescope data.

## 4. Conclusion

Ground-based telescope data frequently contain numerous gaps that challenge signal measurement, particularly for frequency-domain analyses. This study compared three methods—Lomb-Scargle [**?**, **?**], Inpainting [**?**], and Gaussian Processes [**?**]—on simulated solar-like oscillation data with duty cycles of 0.20–0.50. By analyzing data quality classifications, we find that Gaussian Processes consistently deliver the most accurate and stable measurements of frequency, amplitude, and FWHM, yielding results closest to true values. Lomb-Scargle achieves reasonable accuracy for some low-duty-cycle data but suffers from poor stability, while Inpainting performs worst overall, exhibiting weak accuracy and stability and introducing false signals, particularly at duty cycles of 0.20 and 0.30.

For low-duty-cycle ground-based telescope data, the time-domain Gaussian Process method demonstrates exceptional robustness and applicability, making it the preferred approach for measuring periodic signals with irregular amplitudes and phases, such as stellar oscillations, rotational modulation, and convection-driven granulation. Lomb-Scargle serves as a secondary option, while Inpainting is not recommended.

## References

[1] Yu J, Huber D, Bedding T R, et al. ApJS, 2018, 236: 42
[2] Seager S, Deming D. ARA&A, 2010, 48: 631-672
[3] Kraft R P. ApJ, 1967, 150: 551
[4] Koch D G, Borucki W J, Basri G, et al. ApJL, 2010, 713: L79
[5] Ricker G R, Winn J N, Vanderspek R, et al. JATIS, 2015, 1: 014003-014003
[6] Udalski A, Szymanski M, Kaluzny J, et al. AcA, 1992, 42: 253-284
[7] Alcock C, Allsman R A, Alves D R, et al. ApJ, 2000, 542: 281
[8] Shappee B J, Prieto J L, Grupe D, et al. ApJ, 2014, 788: 48
[9] Lomb N R. Ap&SS, 1976, 39: 447-462
[10] Scargle J D. ApJ, 1982, 263: 835-853
[11] Hekker S, Barban C, Baudin F, et al. A&A, 2010, 520: A60
[12] Pires S, Mathur S, García R A, et al. A&A, 2015, 574: A18
[13] Williams C K I, Rasmussen C E. Gaussian processes for machine learning. Cambridge, MA: MIT Press, 2006
[14] Pereira F, Campante T L, Cunha M S, et al. MNRAS, 2019, 489: 5764-5774
[15] Hey D, Huber D, Ong J, et al. arXiv:2403.02489, 2024
[16] Harvey J. Future missions in solar, heliospheric & space plasma physics. Noordwijk, NL: ESA Scientific & Technical Publications Branch, 1985, 235: 199
[17] Foreman-Mackey D, Agol E, Ambikasaran S, et al. ApJ, 2017, 154: 220
[18] Foreman-Mackey D. arXiv:1801.10156, 2018
[19] Salvatier J, Wiecki T V, Fonnesbeck C. PeerJ, 2016, 2: e55
[20] Gelman A, Meng X L, Stern H. STAT SINICA, 1996, 6: 733-760

## Appendix: Fitting Results

Table 3 presents fitting results for Flag=good and Flag=bad data across all methods. Values are medians with median absolute deviations in parentheses.

**Table 3:** Fitting results of the Lomb-Scargle, Inpainting, and Gaussian process methods for Flag=good and Flag=bad data.

*Note: Figure translations are in progress. See original paper for figures.*

*Source: ChinaXiv —Machine translation. Verify with original.*