Ground-based telescope data processing methods: Frequency domain vs. Time domain postprint

Authors: Xiong Qiang, Li Tanda

Date: 2026-03-03T15:53:35+00:00

Abstract

Time-domain observations from ground-based telescopes are frequently characterized by data gaps due to the alternation of day and night and weather conditions, 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 processing time-domain data with gaps and their applicability in asteroseismology, this study employs the Lomb-Scargle algorithm and the Inpainting interpolation method as frequency-domain approaches, alongside the Gaussian Process (GP) method as a time-domain approach. These methods were used to analyze simulated time-domain photometric data featuring solar-like oscillations with duty cycles ranging from 0.20 to 0.50. The results demonstrate that the Gaussian Process method performs best in terms of both accuracy and stability in recovering true values, outperforming the Lomb-Scargle and Inpainting methods. The Inpainting method may introduce a significant number of spurious signals when processing low duty cycle data, interfering with signal measurement. Therefore, the Gaussian Process method is the preferred choice for analyzing low duty cycle data from ground-based telescopes, followed by the Lomb-Scargle method, while the Inpainting method is not recommended.

Full Text

Preamble

Vol. 67, No. 1

2026 年 1 月

ACTA ASTRONOMICA SNICA Vol. 67 No. 1 Jan., 2026 doi: 10.15940/j.cnki.0001-5245.2026.01.002

Data Processing Methods for Ground-Based Telescopes: Frequency Domain vs. Time Domain

XIONG Qiang[1], LI Tanda[1,2†] ([1] School of Physics and Astronomy, Beijing Normal University, Beijing 100875) ([2] Institute for Frontier in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206)

Abstract

Ground-based astronomical observations are inherently subject to various noise sources and atmospheric interference, necessitating robust data processing techniques to extract scientifically meaningful signals. This paper provides a comprehensive comparative analysis of data processing methods for ground-based telescopes, focusing on the distinction between frequency-domain and time-domain approaches. We evaluate the theoretical foundations, computational efficiency, and practical applications of both paradigms. Frequency-domain methods, primarily utilizing Fourier transforms, excel in identifying periodic signals and characterizing stationary noise components. Conversely, time-domain methods are essential for analyzing transient phenomena and non-stationary signals where temporal resolution is critical. By examining specific case studies from current ground-based surveys, we discuss the advantages and limitations of each approach in the context of modern large-scale astronomical data. Our analysis suggests that while frequency-domain techniques remain indispensable for spectral analysis, the increasing demand for real-time transient detection is driving significant advancements in time-domain algorithms and machine learning-based processing pipelines.

Keywords: ground-based telescopes, data processing, frequency domain, time domain, signal-to-noise ratio, astronomical algorithms

摘要地面望远镜的时域观测因昼夜交替和天气影响常导致数据空缺，占空比较低(典型值约为 0.30)，对时

This has a significant impact on time-domain astronomical research. To compare the performance and applicability of frequency-domain and time-domain analysis methods in asteroseismology when dealing with gapped time-domain data, this study evaluates the Lomb-Scargle algorithm and Inpainting interpolation (as frequency-domain methods) against Gaussian Process (GP) regression (as a time-domain method). The analysis was conducted using simulated photometric time-series data exhibiting solar-like oscillations with duty cycles ranging from 0.20 to 0.50. The results demonstrate that the Gaussian Process method performs best in terms of accuracy and stability when recovering true values, outperforming both the Lomb-Scargle and Inpainting methods. Furthermore, the Inpainting method may introduce significant spurious signals when processing low duty cycle data, thereby interfering with signal measurement. Consequently, the Gaussian Process method is the preferred choice for analyzing low duty cycle data from ground-based telescopes, followed by the Lomb-Scargle method, while the Inpainting method is not recommended.

Keywords: methods: data analysis, stars: oscillations, stars: solar-type **CLC number:** P141; **Document code:** A

1 引言

Introduction

High-precision time-domain astronomy is currently a focal point and frontier of astronomical research, encompassing fields such as asteroseismology [?], exoplanet detection [?], and stellar rotation measurements [?]. However, high-precision and continuous photometric data typically originate from space-based telescopes, such as *Kepler* [?] and TESS (Transiting Exoplanet Survey Satellite) [?]. While these space missions have significantly advanced their respective fields, their data possess inherent limitations. For instance, the observation duration of *Kepler* is approximately 4 years; for long-period variable stars with pulsation periods ranging from several days to years, the frequency resolution remains insufficient.

In this context, data from ground-based telescopes continue to hold irreplaceable value. Projects such as OGLE (Optical Gravitational Lensing Experiment) [?], MACHO (Massive Compact Halo Objects project) [?], and ASAS-SN (All-Sky Automated Survey for SuperNovae) [?] typically offer much longer observation baselines. This significantly improves frequency resolution, providing a unique advantage in the study of long-period variables and helping to compensate for the limitations of space telescopes.

Although ground-based telescopes can provide long-baseline time-domain data, they are constrained by weather conditions and the diurnal cycle. Consequently, large gaps often appear in the data, leading to a very low duty cycle for single-station observations. Here, the duty cycle refers to the ratio of the time an object is actively observed (e.g., during a periodic phenomenon) to the total observation duration. For long-baseline projects like OGLE, the typical duty cycle is approximately 0.30. Multi-station joint observations can help increase the duty cycle to 0.60–0.95, depending on the number of stations and their geographical distribution.

These gaps have a significant impact on frequency-domain analysis. When performing a Fourier transform on observed data with gaps, a window function is introduced. The spectrum of this window function convolves with the true spectrum of the original signal, producing sidelobe effects that cause spectral energy to leak from the main peak. This spectral leakage makes signal measurement difficult. Therefore, effectively addressing data gaps is of paramount importance.

There are three primary strategies for handling this issue: 1. Directly using the Lomb-Scargle algorithm [?, ?], which is specifically designed for non-equidistantly sampled data. 2. Processing the gaps before performing frequency-domain analysis, such as the approach by Hekker et al. [?] who directly removed gaps, or Pires et al. [?] who interpolated gaps based on existing data points. 3. Using Gaussian Processes (GP) [?] to perform analysis directly in the time domain. For example, Pereira et al. [?] used Gaussian processes to model granulation and oscillations in the time domain, finding that this approach effectively handles data gaps. Similar conclusions were reached by Hey et al. [?], who processed sparse TESS data using Gaussian processes in the time domain, demonstrating the utility of GP in analyzing asteroseismic data containing gaps.

This paper focuses on simulated ground-based telescope data with duty cycles ranging from 0.20 to 0.50. We process the data using both frequency-domain and time-domain analysis methods and systematically evaluate their performance. Specifically, we employ three methods: the Lomb-Scargle algorithm and the Inpainting method for frequency-domain analysis, and the Gaussian Process method, which is conducted entirely in the time domain.

2.1 数据

In this study, we utilize simulated light curve data for solar-like oscillating stars. Solar-like oscillations refer to the vibration mechanism randomly excited by turbulent convection in the outer layers of a star, identical to the excitation mechanism of the Sun; hence, these stars are termed solar-like oscillators. Due to the characteristics of stochastic excitation and damping, solar-like oscillations exhibit frequency instability, as well as random variations in amplitude and phase. Consequently, the characteristic frequencies possess a certain degree of dispersion, making them highly sensitive to the effects of data gaps, which imposes higher requirements on data processing methods. The simulated data consist of three components: granulation, oscillations, and white noise. Granulation is an observational feature describing the convective cell structures on the surface of solar-like stars, specifically referring to the fine light and dark structures caused by convection. Its characteristics in the power spectrum can be approximated by a Harvey profile [?], namely:

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

where $P(f)$ represents the power density, f is the frequency, f_{gran} is the characteristic frequency, and A_{gran} is the amplitude corresponding to the characteristic frequency. Solar-like oscillations are standing waves of pressure modes (p-modes) or mixed pressure-gravity modes, which are damped by near-surface convection. They are excited within a specific frequency range, and their amplitudes generally follow a bell-shaped distribution. In the power spectrum, they can be approximated by a Lorentzian profile:

$$P(f) = \frac{A(\Gamma f_0)^2}{(f^2 - f_0^2)^2 + (\Gamma f)^2}$$

where f_0 represents the central frequency of each mode, Γ represents the full width at half maximum (FWHM), and A represents the amplitude at the central frequency. We configured three modes described by Eq. (2) to simulate p-mode oscillations (hereafter referred to as the "Signal"). The true parameter values used in our setup are shown in , where the subscripts "left" , "central" , and "right" correspond to the left, middle, and right signals, respectively, based on their central frequencies f_0.

Parameters configured in the light curve and their corresponding true values

We introduced normally distributed white noise into the data with a mean of 0 and a variance of 7 ppm, i.e., $N(0, 7^2)$, to simulate the signal-to-noise ratio of actual ground-based data as closely as possible. Referencing data from the ground-based OGLE telescope, we set the total duration of the light curve to approximately 30 years, ranging from 0 to 11,000 days, with a sampling rate of approximately 2 days, totaling 5,000 data points. An example of the generated light curve and its corresponding power spectrum is shown in [Figure 1: see original paper].

To generate gaps similar to those found in real data, it is necessary to construct a window function that conforms to actual conditions, defined as:

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

(Actual error controlled within 1%).

[Figure 1: see original paper] An example of the generated light curve and its corresponding power spectrum, with each part labeled in the legend.

To this end, we also referenced data from the ground-based OGLE telescope. The procedure for creating a window with a specific duty cycle is as follows: (1) First, data from a real star with a high duty cycle was obtained from the OGLE-III database, as shown in [Figure 2: see original paper] (a), from which the observation window was derived, as shown in [Figure 2: see original paper] (b); (2) The window obtained in step (1) was repeated sequentially until it extended to the same length as the initial light curve data (11,000 days); (3) To simulate more complex scenarios, we introduced one large gap of approximately 1,000 days and three small gaps of 200 days into the window. The resulting initial window is shown in [Figure 2: see original paper] (c); (4) Within the window obtained in step (3), a position was randomly selected as a starting index, and the subsequent 100 window values were set to 0 or 1. This process was repeated until the duty cycle of the window reached the specified target value.

[Figure 2: see original paper] (a) Light curve data of a specific target (ID: OGLE-LMC-LPV-35662) from the ground-based OGLE telescope; (b) The corresponding observation window; (c) The initial window finally used in this paper.

Fig. 2 (a) Light curve data of a specific target (ID: OGLE- LMC-LPV-35662) from the ground-based OGLE telescope; (b) the corresponding observational window; (c) the initial window ultimately used in this work.

By combining the initially generated light curve data with window functions of specific duty cycles, we can produce gapped light curve datasets that satisfy the required duty cycle constraints.

We generated light curve data with duty cycles ranging from 0.20 to 0.50. For each duty cycle, we selected two light curves and labeled them as "good" or "bad" based on whether the primary oscillation signal could be identified through visual inspection of the power spectrum obtained via the Lomb-Scargle algorithm (detailed below). The purpose of this classification is to compare the performance of different measurement methods when the signal undergoes significant degradation due to data sampling patterns, white noise, and window function variations. [Figure 3: see original paper] illustrates all the light curve data utilized in this study.

[Figure 3: see original paper] All light curve data used in this paper, where the left panels represent data labeled as "good" and the right panels represent data labeled as "bad."

Fig. 3 All the light curve data used in this study, with the left panel showing data flagged as "good" and the right panel showing data flagged as "bad" .

2.2.1 Lomb-Scargle

Unlike the Fast Fourier Transform (FFT), which is only applicable to equally spaced and uniformly sampled data, the Lomb-Scargle algorithm is capable of processing non-uniformly sampled data. Its fundamental principle involves fitting a sine wave to the data in the time domain for every candidate frequency. If a sine wave at a specific frequency fits the data well (resulting in small residuals), the corresponding amplitude in the frequency domain will be large. Consequently, the Lomb-Scargle algorithm can effectively handle gaps within a dataset. In cases where the data contains no gaps, the results produced by Lomb-Scargle are identical to those obtained via the Fast Fourier Transform.

[Figure 1: see original paper]

The figure above illustrates the performance of the algorithm under varying duty cycles and data quality flags. As shown in the relative flux measurements over time, the algorithm maintains its ability to identify periodic signals even as the duty cycle decreases from 0.50 to 0.20. The comparison between "good" and "bad" flags demonstrates how the method distinguishes between reliable signals and noise or artifacts in the time-series data.

2.2.2 Inpainting

This method aims to recover the ideal light curve by utilizing the ground truth values and the observation windows present in the data. By applying a specific transformation matrix to the light curve data containing gaps, the approach employs an iterative process to find a solution where most of the transformed coefficients are zero, thereby identifying the sparsest solution.

For example, the characteristics of a time-series dataset featuring a single sinusoidal variation can be reflected through coefficients in the Fourier space. In this case, the corresponding transformation matrix is the Fourier transform, as most coefficients in the Fourier domain are zero. A detailed description of this method can be found in Ref. [?]. Figure 4 [Figure 4: see original paper] illustrates the results of applying this method to a light curve with a duty cycle of 0.30 (labeled as "good"), along with the power spectrum obtained using the corresponding Lomb-Scargle algorithm.

Figure 4. Results of the Inpainting method applied to light curve data with a duty cycle of 0.30 and labeled as "good." The upper panel shows the light curve, and the lower panel shows the corresponding power spectrum.

Fig. 4 The results of the Inpainting method applied to light curve data with a duty cycle of 0.30 and flagged as "good".

The upper panel shows the light curve, while the lower panel presents the corresponding power spectrum.

Gaussian Process (GP) is a non-parametric Bayesian statistical method widely applied in the field of machine learning [?]. A Gaussian Process is a stochastic process that assumes the output value corresponding to each input point (such as time or wavelength) is a normally distributed random variable. Furthermore, any finite subset of these random variables follows a multivariate Gaussian distribution. The distribution of a Gaussian Process is jointly described by a mean function and a covariance function, defined as follows:

In this study, the mean function $m(x)$ is set to 0, representing our initial guess of the target function in the absence of any observational data. The covariance function $k(x, x')$, also known as the kernel function, describes the similarity between any two points in the input space. Different kernel functions determine different prior distributions for the function. If an appropriate kernel function is selected and its parameters are effectively optimized, the parameters of interest can be obtained directly in the time domain through the Gaussian Process.

We utilize the Simple Harmonic Oscillator (SHO) term from the CELERITE2 library [?] as the kernel function to construct the Gaussian Process model. This kernel function describes a damped simple harmonic motion process, which shares similar physical characteristics with solar-like oscillations and thus satisfies our requirements for model construction. Its corresponding power spectrum is:

Where S_0 is the scaling factor of the system' s power spectral density, representing the total power or energy intensity of the system; ω_0 is the characteristic angular frequency; and Q is the quality factor, which describes the magnitude of the damping. Among these three parameters, when $Q = 1/\sqrt{2}$, this equation corresponds exactly to Equation (1) and can be used for modeling the granulation component. Note that in Equation (5), the free parameters are S_0, ω_0, and Q (excluding granulation). Therefore, it is necessary to establish conversion relationships between these parameters and those in Equation (1) and Equation (2). Following Pereira et al. [?], the conversion relationships are:

Where S_0 and ω_0 are as defined in Equation (5), and these are physical quantities related to the granulation.

2.3 拟合

The fitting for the Lomb-Scargle and Inpainting methods is performed in the frequency domain. [Figure 1: see original paper] (a) and (b) illustrate the relative flux and power density for a duty cycle of 0.30, categorized as "good" data. The Gaussian Process (GP) is defined as:

$$F(x) \sim \mathcal{GP}[m(x); k(x, x')]$$

where x and x' are input variables, $m(x)$ is the mean function, and $k(x, x')$ is the covariance function. The power spectral density of the SHO (Shot Noise Oscillator) kernel is given by:

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

where S_0, ω_0, and Q are the power scale, undamped frequency, and quality factor, respectively. At $\omega = \omega_0$, the power is $S(\omega_0) = \sqrt{2/\pi} S_0 Q^2$. For the granulation component, we set $Q = 1/\sqrt{2}$, which yields $S_0 \omega_0 Q = A_{\text{gran}} = \sqrt{1/2\pi} S_{\text{gran}}$. Defining $f_{\text{gran}} = 2\pi/\omega_{\text{gran}}$, $f_0 = 2\pi/\omega_0$, $\Gamma = \omega_0/2\pi Q$, and $A = 4 S_0 Q^2$, the granulation is described by Equation (1) and the oscillation by Equation (2). A constant term σ is added to account for white noise. The complete fitting equation is as follows:

Figure 5 shows the fitting results of the three methods for data with a duty cycle of 0.50, categorized as "good" (top panel) and "poor" (bottom panel). The variables are defined as previously described, with subscripts corresponding to the left, middle, and right signals. The Gaussian Process fitting is conducted in the time domain. We utilize one SHO term to represent the granulation component and a superposition of three standard SHO terms to account for the oscillation components, while incorporating a constant term σ to model the white noise.

In this study, all three methods utilize the same uniform distribution priors. The prior for the white noise is set to σ, while the remaining input parameters and their corresponding priors are detailed in .

Fitting input parameters and corresponding priors for the Lomb-Scargle, Inpainting, and Gaussian Process methods. [Figure 5: see original paper] Fitting results of the Lomb-Scargle method (gray dashed line), Inpainting method (gray dash-dotted line), and Gaussian Process method (gray dotted line) for data with a duty cycle of 0.50, labeled as "good" (top) and "poor" (bottom). The black solid line represents the power spectrum corresponding to the true parameters.

Parameter Granulation [0.01, 0.2] [0.3, 0.5] [0.5, 0.7] [0.7, 0.9] Priors [0, 7] [0, 6] [0, 6] [0, 6] [0.01, 0.2] [0.01, 0.2] [0.01, 0.2]. We utilize PyMC3 [**?**] to perform the inference via Markov Chain Monte Carlo (MCMC) sampling.

方法在参数空间中采样以获得参数的后验分布，

The initial positions of the parameters are randomly generated from a prior distribution. We selected two chains and first performed 1,000 tuning steps to optimize the sampler' s hyperparameters (such as step size and acceptance rate), thereby improving sampling efficiency and quality. Subsequently, we executed 1,000 draws to obtain the posterior distribution of the parameters. We utilized the Gelman-Rubin convergence diagnostic statistic ($\hat{R}$) [**?**] to assess parameter convergence. When $\hat{R} < 1.05$, it indicates that the parameter distributions across the two chains are extremely similar, at which point the parameters are considered to have converged. We take the median of the posterior distribution as the best estimate for each parameter. [Figure 5: see original paper] displays the fitting results for data with a duty cycle of 0.50 flagged as "good" (upper panel) and "bad" (lower panel), obtained using the Lomb-Scargle method (gray dashed line), the inpainting method (gray dash-dotted line), and the Gaussian process method (gray dotted line). The black line represents the power spectrum corresponding to the true parameters.

It is noteworthy that the Gaussian process method operates in the time domain. [Figure 6: see original paper] provides a schematic representation of the Gaussian process method applied to the light curves shown in [Figure 5: see original paper]. To more clearly illustrate the details within the figure, only the data from 3,000–8,000 days and their corresponding fitting results are selected for display.

3 结果

This study employs the Lomb-Scargle algorithm, the Inpainting method, and Gaussian Process (GP) techniques to analyze data with duty cycles ranging from 0.20 to 0.50, categorized by "good" and "bad" flags.

分析，测定了米粒组织的特征频率和振幅以及振

The center frequency, amplitude, and full width at half maximum (FWHM) of the dynamic signal. All measurements

结果列于附录．参数的最佳估计值由后验分布的

2. Methodology and Model Fitting

The power spectrum density (PSD) of the stellar signal is modeled using a multi-component semi-empirical formula. Specifically, the background noise and granulation signals are represented by the following expression:

$$W_P(f) = W + \frac{2A_{\text{gran}}}{1 + (f/f_{\text{gran}})^4} + \sum_{i=1}^{3} \frac{3A_i(\gamma_i f_{0,i})^2}{(f^2 - f_{0,i}^2)^2 + (\gamma_i f)^2} \tag{7}$$

where $i = 1, 2, 3$ corresponds to different oscillation modes or physical components. The quality factor is defined as $Q = 1/\sqrt{2}$. The parameters for the fit, including the white noise level W, the granulation amplitude A_{gran}, and the characteristic frequency f_{gran}, are typically estimated within the logarithmic range of $[0, 5]$ for frequency f (measured in μHz) and power density A (measured in $\text{ppm}^2 \cdot \mu\text{Hz}^{-1}$).

[Figure 1: see original paper]

Figure 1 illustrates the comparative performance of different data recovery methods under varying observational conditions. Subfigure (a) displays a "good" flag scenario with a duty cycle of 0.50, while subfigure (b) shows a "bad" flag scenario. The plots compare the true power spectrum against results obtained via the Lomb-Scargle periodogram, Inpainting techniques, and Gaussian Process (GP) regression. The power density is plotted against frequency, demonstrating how each method reconstructs the signal across the μHz range.

As noted by Xiong Qiang et al. in "Ground-based Telescope Data Processing Methods: Frequency Domain vs. Time Domain," the statistical results are reported using the median of the distribution, while the associated uncertainties are expressed using the Median Absolute Deviation (MAD).

3.1 频率

The standard deviations (hereafter referred to as standard deviation) calculated from the data are 9.8%, 11.2%, and 26.1%, respectively. [Figure 7: see original paper] illustrates the frequency measurement bias for the three methods across duty cycles ranging from 0.20 to 0.50, normalized as a ratio relative to the ground truth.

The upper panel corresponds to data labeled as "good," while the lower panel corresponds to data labeled as "poor." The black dashed line represents the reference position where the ratio equals zero, and the gray dashed lines indicate various reference ranges for the ratios. For the data labeled as "good," the reference values for the granulation, left signal, center signal, and right signal are provided. For the data labeled as "poor," the corresponding values are denoted as I. Regarding the measurement of the characteristic frequency of granulation for the "good" data: in terms of accuracy, the Gaussian Process (GP) method outperforms both the Lomb-Scargle and Inpainting methods, with measurement bias remaining below 20% across all duty cycles.

Overall, the Lomb-Scargle method performs better than the Inpainting method; the latter exhibits a bias approaching 90% at a duty cycle of 0.2. In terms of stability, the Lomb-Scargle method is slightly superior to the Gaussian Process method and significantly outperforms the Inpainting method. [Figure 6: see original paper] shows the direct fitting results of the Gaussian Process method for light curve data labeled as "good" with a duty cycle of 0.50. The black dots represent the original light curve data, the light gray line represents the predicted results at each time point, and the dark gray shaded area indicates the 68% confidence interval of the predictions.

Fig. 6 Direct fitting results of the Gaussian process method applied to light curve data with a duty cycle of 0.50 and flagged as "good". The black dots represent the original light curve data, the lightgray line indicates the predicted values at each time t, and the gray shaded region represents the 68% confidence interval of the predictions.

[Figure 7: see original paper] Fig. 7 The measurement deviations of granulation frequency (ν_{gran}) and signal frequency (ν_0) relative to the true values, normalized as a ratio of the true values, using the Lomb-Scargle method (circles), Inpainting method (squares), and Gaussian Process method (triangles) under different duty cycle conditions.

(2) Measurement of the central frequency of oscillation signals: In terms of accuracy, the Gaussian Process (GP) method performs best across all duty cycles, with measurement deviations consistently below 5%, significantly outperforming both the Lomb-Scargle and Inpainting methods. Meanwhile, the Lomb-Scargle method proves superior to the Inpainting method at all duty cycles, with the difference being particularly pronounced at a duty cycle of 0.20. Regarding stability, the standard deviations for the GP method across the left, central, and right signals are 2.5%, 0.8%, and 1.3%, respectively. This overall performance is superior to that of the Lomb-Scargle method (4.4%, 0.8%, and 1.5%) and the Inpainting method (12.2%, 3.0%, and 2.3%). Notably, while the stability of the Lomb-Scargle method is comparable to the GP method for the central signal, it remains slightly inferior in other cases, whereas the Inpainting method consistently exhibits the lowest stability.

II. For data flagged as "bad":

(1) Measurement of the characteristic frequency of granulation: In terms of accuracy, the GP method demonstrates the best performance, with measurement deviations remaining below 10% across all duty cycles. The Lomb-Scargle method outperforms the Inpainting method, showing a significant advantage especially at a duty cycle of 0.20. In terms of stability,

the GP method is superior to both the Lomb-Scargle and Inpainting methods, with standard deviations of 6.2%, 8.7%, and 35.3%, respectively.

(2) Measurement of the central frequency of oscillation signals: Regarding accuracy, the GP method performs best across all signals; specifically, the measurement deviation for the left signal is consistently below 10%, within 2% for the central signal, and within 1% or 2% for the right signal, significantly outperforming the other methods.

Overall, the Lomb-Scargle method outperforms the Inpainting method, with the sole exception of the right signal at duty cycles of 0.30 and 0.40, where the Inpainting method is slightly superior. In terms of stability, the Gaussian Process method···

方法的标准差最小，在左、中、右信号下分别为

2.3%, 0.6%, and 0.9%, respectively, representing the best overall performance. The Lomb-Scargle method yielded a standard deviation of 17.8% for the left signal, which was significantly better than the Inpainting method (24.4%); however, it performed the worst for the middle signal (1.2%) and the right signal (2.5%). In comparison, the Inpainting method produced standard deviations of 3.0% and 1.2% for these two cases. Overall, the Gaussian Process method consistently outperformed the other two methods in terms of stability.

In summary, the Gaussian Process method consistently demonstrated the highest accuracy and stability in frequency measurement, regardless of whether the data were labeled as "good" or "poor." The Lomb-Scargle method generally outperformed the Inpainting method, with its advantages being particularly prominent at duty cycles of 0.20 and 0.30.

3.2 振幅

[Figure 8: see original paper] illustrates the measurement deviations of the amplitudes for the three methods across duty cycles ranging from 0.20 to 0.50, normalized as a ratio relative to the ground truth values. The upper panel corresponds to the data labeled as "good," while the lower panel corresponds to the data labeled as "poor." The black dashed line indicates a ratio of 0, and the gray dashed lines mark the reference values for the ratios. Specifically, for the "good" data, the reference values for the granulation, left signal, center signal, and right signal are 0.2, 0.5, 0.2, and 0.4, respectively. For the "poor" data, these values are 0.2, 0.1, 0.3, and 0.5, respectively.

Figure 8. Similar to Figure 7, but displaying the measurement deviations of the amplitude (A) relative to the ground truth, normalized as a ratio relative to the ground truth values.

Fig. 8 Similar to Figure 7, but showing the normalized deviations of amplitude ( ) measurements relative to the true values.

I. For data labeled as "good" : (1) Measurement of granulation feature amplitudes: In terms of accuracy, the Gaussian Process (GP) method achieved measurement biases within 20%, outperforming both the Lomb-Scargle and Inpainting methods. Although the Inpainting method performed better than the Lomb-Scargle method at duty cycles of 0.3, 0.4, and 0.5—with biases remaining within 20%—it exhibited significantly larger biases at a duty cycle of 0.20. Regarding stability, the GP method demonstrated the best performance with a standard deviation of 4.0%, while the Lomb-Scargle and Inpainting methods showed standard deviations of 8.2% and 70.1%, respectively.

(2) Measurement of oscillation signal amplitudes: In terms of accuracy, the GP method performed best across all signals. For the left signal, measurement biases were mostly within 50% and fell below 10% at duty cycles of 0.30 and 0.50. For the central signal, measurement biases were consistently below 20%, while for the right signal, biases remained within 40%. Overall, the Lomb-Scargle method outperformed the Inpainting method, although the Inpainting method was slightly superior at a duty cycle of 0.40 for both the left and central signals. In terms of stability, the Lomb-Scargle method achieved the lowest standard deviation for the left signal (26.8%), slightly better than the GP method (27.8%). However, the GP method was significantly superior for the central signal (8.9% vs. 14.4%) and the right signal (23.0% vs. 29.2%). The Inpainting method exhibited the highest standard deviation in all cases, recorded at 63.9%, 20.0%, and 61.0%, respectively. Overall, the Gaussian Process method performed best in most scenarios and demonstrated high stability.

II. For data labeled as "bad" :

(1) Measurement of granulation feature amplitudes: In terms of accuracy, the Gaussian Process method outperformed both the Lomb-Scargle and Inpainting methods. While the Inpainting method was superior to the Lomb-Scargle method at duty cycles of 0.30, 0.40, and 0.50, its bias increased significantly at a duty cycle of 0.20, reaching nearly 170%. Regarding stability, the GP method showed the best performance with a standard deviation of 8.8%. The Lomb-Scargle method outperformed the Inpainting method, with standard deviations of 14.5% and 89.7%, respectively.

(2) Measurement of oscillation signal amplitudes: In terms of accuracy, among the three…

方法在左、右信号下均存在较大偏差，测量结果

The deviations from the ground truth generally exceed 30%, with the deviation for the left signal even exceeding [percentage missing]%. Only the Lomb-Scargle method achieved a deviation of less than 10% when the duty cycle of the right signal was 0.50. For the middle signal, the Gaussian Process (GP) method exhibited the smallest overall measurement deviation, except at a duty cycle of 0.50. Additionally, the Inpainting method slightly outperformed the Lomb-Scargle method when the duty cycle of the middle signal was 0.20. In terms of stability, the Lomb-Scargle method yielded the lowest standard deviation for the left signal (8.7%), performing better than both the Inpainting method (13.2%) and the Gaussian Process method (14.0%).

For the middle signal, the Gaussian Process method demonstrated the best performance with the smallest standard deviation (4.2%), while the Inpainting method (27.2%) outperformed the Lomb-Scargle method (34.2%). For the right signal, the Gaussian Process method maintained the smallest standard deviation (6.0%), followed by the Lomb-Scargle method (31.2%), while the Inpainting method exhibited the largest standard deviation (46.8%).

In summary, for the amplitude measurement of granulation tissue, the Gaussian Process method demonstrates higher accuracy and stability in both "good" and "poor" quality data, offering distinct advantages over the Lomb-Scargle and Inpainting methods. Regarding the amplitude measurement of vibration signals, the Gaussian Process method outperforms both the Lomb-Scargle and Inpainting methods in terms of accuracy and stability for "good" data. For "poor" data, while the accuracy of all three methods is suboptimal, the Gaussian Process method remains the most superior in terms of stability.

3.3 半高全宽

[Figure 9: see original paper] illustrates the measurement deviations of the Full Width at Half Maximum (FWHM) for three methods across duty cycles ranging from 0.20 to 0.50, normalized as a ratio relative to the ground truth. The upper panel corresponds to data labeled as "good," while the lower panel corresponds to data labeled as "poor." The black dashed line represents the true value, and the gray dashed lines indicate reference ratios. Specifically, for the "good" data, the reference ratios for the left, middle, and right signals are 0.5, 0.2, and 0.2, respectively; for the "poor" data, these values are 0.5, 0.5, and 0.5.

I. For data labeled as "good": In terms of accuracy, all three methods performed poorly overall for the left signal. Among them, the Inpainting method exhibited the largest deviation at a duty cycle of 0.20, exceeding 200%. In contrast, the Gaussian Process (GP) method demonstrated superior accuracy compared to the other two methods, particularly at a duty cycle of 0.30, where its deviation was approximately 20%.

For the middle signal, all three methods showed significant deviations at a duty cycle of 0.20. However, at duty cycles of 0.30, 0.40, and 0.50, the GP method performed best, with measurement deviations maintained within 10%. This was followed by the Lomb-Scargle method (with deviations within 20%), while the Inpainting method exhibited the most significant errors. For the right signal, the GP method outperformed the others overall, with measurement deviations of less than 20% across all duty cycles. Furthermore, at duty cycles of 0.20 and 0.30, the Lomb-Scargle method yielded significantly smaller measurement deviations than the Inpainting method. Regarding stability, the GP method performed best, exhibiting the smallest standard deviations for the left and right signals (42.7% and 15.0%, respectively). The Lomb-Scargle method followed (88.0% and 23.8%), while the Inpainting method showed the highest standard deviations (145.4% and 59.3%). However, for the middle signal, the Inpainting method achieved the lowest standard deviation (18.3%), outperforming both the Lomb-Scargle method (40.0%) and the GP method (48.9%).

II. For data labeled as "poor" : In terms of accuracy, all three methods exhibited large deviations across all signals, resulting in low overall accuracy. In terms of stability, the GP method performed best, yielding the smallest standard deviations of 31.7%, 31.3%, and 51.3% for the left, middle, and right signals, respectively. The Inpainting method showed larger standard deviations (46.1%, 53.9%, and 81.6%), while the Lomb-Scargle method exhibited the poorest stability, with standard deviations reaching 97.5%, 78.8%, and 92.7%.

Overall, the Gaussian Process method demonstrated the highest stability across all cases, whereas the Lomb-Scargle method exhibited the lowest stability. [Figure 9: see original paper] Similar to Figure 7, but displaying the measurement deviations of the signal' s Full Width at Half Maximum (FWHM) normalized as a ratio relative to the true value.

Fig. 9 Similar to Figure 7, but showing the normalized deviations of the signal' s full width at half maximum ( ) relative to the true values.

In summary, regarding the measurement of the full width at half maximum (FWHM) of vibration signals, the Gaussian Process (GP) method demonstrates the best overall performance in terms of accuracy and stability for data labeled as "Good," outperforming both the Lomb-Scargle and Inpainting methods. For data labeled as "Poor," while all three methods exhibit relatively low accuracy, the Gaussian Process method maintains superior stability compared to the Lomb-Scargle and Inpainting approaches.

3.4 假信号问题

Regarding the Inpainting method, we found that it tends to introduce a significant number of spurious signals when processing data with low duty cycles (e.g., 0.30). As shown in [Figure 10: see original paper], the upper panel displays the spectrum obtained from the original data using the Lomb-Scargle algorithm (Original), while the lower panel shows the spectrum after processing with the Inpainting method; arrows indicate the frequency positions of the true signals. In the original spectrum, the signals are significantly weakened due to the presence of data gaps, making them nearly impossible to identify. Conversely, in the spectrum processed by the Inpainting method, multiple false peaks appear near the true signals. These spurious peaks significantly interfere with the measurement of real signals and may lead to incorrect frequency determinations, thereby affecting the accurate derivation of stellar physical parameters. Consequently, we do not recommend employing the Inpainting method for processing ground-based telescope data with low duty cycles.

[Figure 10: see original paper] Results of the Inpainting method applied to a light curve with a duty cycle of 0.30, where arrows indicate the frequency positions of the true signals.

Fig. 10 Results of the Inpainting method applied to light curve data with a duty cycle of 0.30, where the arrows indicate the frequency positions of the true signals. [Figure 10: see original paper] Xiong Qiang et al.: Ground-based Telescope Data Processing Methods: Frequency Domain vs. Time Domain

4 结论

Due to the frequent presence of significant gaps in ground-based telescope data, measuring real signals—particularly when using frequency-domain analysis methods—presents a substantial challenge. This paper evaluates the performance of three distinct methods applied to simulated ground-based telescope data with duty cycles ranging from 0.20 to 0.50 and characteristic solar-like oscillations: the frequency-domain-based Lomb-Scargle algorithm [**?**, **?**], the Inpainting method [**?**], and the time-domain-based Gaussian Process (GP) method [**?**]. Through a categorical analysis of data quality, we find that the Gaussian Process method demonstrates the best overall performance in measuring frequency, amplitude, and full width at half maximum (FWHM). Its results are more stable and more closely align with the ground-truth values. In contrast, while the Lomb-Scargle method maintains acceptable accuracy for some low-duty-cycle data, its stability is relatively poor. The Inpainting method performed the worst overall, exhibiting low accuracy and stability while frequently introducing spurious signals, particularly at duty cycles of 0.20 and 0.30.

Consequently, for ground-based telescope data with low duty cycles, the time-domain Gaussian Process method demonstrates superior robustness and applicability. It emerges as the preferred method for measuring periodic signals characterized by irregular amplitudes and phases, as well as stochastic excitation and damping—such as stellar oscillations, rotational modulation, and granulation motion induced by convection.

References

Koch D G, Borucki W J, Basri G, et al. ApJ, 2010, 713:

Ricker G R, Winn J N, Vanderspek R, et al. JATIS, 2015, 1: 014003 Udalski A, Szymanski M, Kaluzny J, et al. AcA, 1992, 42: 253 Alcock C, Allsman R A, Alves D R, et al. ApJ, 2000, 542: 281 Shappee B J, Prieto J L, Grupe D, et al. ApJ, 2014, 788: 48 Lomb N R. Ap&SS, 1976, 39: 447 Scargle J D. ApJ, 1982, 263: 835 Hekker S, Barban C, Baudin F, et al. A&A, 2010, 520:

Pires S, Mathur S, García R A, et al. A&A, 2015, 574:

Williams C K I, Rasmussen C E. Gaussian Processes for Machine Learning. Cambridge, MA: MIT Press, 2006:

Pereira F, Campante T L, Cunha M S, et al. MNRAS, 2019, 489: 5764 Hey D, Huber D, Ong J, et al. aXiv, 2024: 2403.02489 Harvey J. Future Missions in Solar, Heliospheric & Space Plasma Physics. Noordwijk, NL: ESA Scientific & Technical Publications Branch, 1985, 235: 199 Foreman-Mackey D, Agol E, Ambikasaran S, et al.

ApJ, 2017, 154: 220 Foreman-Mackey D. RNAAS, 2018, 2: 31 Salvatier J, Wiecki T V, Fonnesbeck C. PeerJ, 2016, 2:

Yu J, Huber D, Bedding T R, et al. ApJ, 2018, 236: 42 Seager S, Deming D. ARA&A, 2010, 48: 631 Gelman A, Meng X L, Stern H. Stat Sinica, 1996, 6:

Kraft R P. ApJ, 1967, 150: 551

Processing Ground-Based Asteroseismic Photometric 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 conducted with ground-based telescopes are frequently interrupted by the day-night cycle and fluctuating weather conditions. These factors result in significant data gaps and a relatively low duty cycle (typically around 0.30), which poses a substantial challenge to time-domain astronomical research. To evaluate the performance of frequency-domain and time-domain analysis methods in processing gapped data and to determine their suitability for asteroseismology, this study employs the Lomb-Scargle algorithm and the Inpainting interpolation method as representative frequency-domain approaches, alongside the Gaussian Process (GP) method as a time-domain approach. These methods were applied to simulated light curves exhibiting solar-like oscillations with duty cycles ranging from 0.20 to 0.50. Our results indicate that the Gaussian Process method outperforms both the Lomb-Scargle and Inpainting methods in terms of accuracy and stability when recovering true physical values. Notably, the Inpainting method tends to introduce significant spurious signals when applied to low-duty-cycle data, leading to potential distortions in measurements. Consequently, the Gaussian Process method is the preferred choice for analyzing low-duty-cycle data from ground-based telescopes, followed by the Lomb-Scargle method; the Inpainting method is not recommended for such applications.

Key words: methods: data analysis, stars: oscillations, stars: solar-type

The fitting results of the Lomb-Scargle method, Inpainting method, and Gaussian process (GP) method for data flagged as "good" and "bad."

Method

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

The table above summarizes the extracted features for various duty cycles using both Lomb-Scargle periodogram analysis and granulation signal modeling. The parameters include characteristic frequencies and power spectral density estimates, represented as $f(\text{ppm}^2\text{Hz}^{-1})$. Values in parentheses indicate the uncertainty in the last significant digit. For instance, at different duty cycles, the central granulation signal frequencies were measured at 234.9 ± 3.4, 221.0 ± 3.4, 257.3 ± 3.4, and 296.4 ± 3.4, respectively.

Comparative Analysis of Frequency and Time Domain Methods

In the study of stellar oscillations and granulation signals, the choice between frequency-domain and time-domain processing is critical for ground-based telescope data. Frequency-domain methods, such as the Lomb-Scargle periodogram, are particularly robust for detecting periodicities in unevenly sampled data, which is a common occurrence in ground-based observations due to day-night cycles and weather interruptions. As shown in the feature extraction results, the Lomb-Scargle approach provides consistent estimates for duty cycles ranging from approximately 0.12 to 0.81.

But time-domain analysis offers complementary insights, especially when characterizing stochastic processes like granulation. By modeling the signal directly in the time domain, researchers can better account for the colored noise profiles inherent in stellar intensity fluctuations. The transition from low duty cycles (e.g., 0.121) to high duty cycles (e.g., 0.813) demonstrates how the precision of the signal recovery improves, with the granulation signal power and frequency shifts becoming more clearly defined as the temporal coverage increases.

The data suggests that while frequency-domain methods excel at identifying dominant modes, time-domain modeling is essential for capturing the underlying physical parameters of the granulation signal. The integration of both approaches allows for a more comprehensive understanding of the stellar signals, mitigating the artifacts introduced by the window function of ground-based observations. Future work will focus on optimizing these algorithms to handle the increased data rates expected from next-generation ground-based survey telescopes.

Method

Continued

Results of Power Spectral Density Analysis

The following table presents the comparative results of the power spectral density (PSD) analysis, contrasting the Lomb-Scargle periodogram method with the Inpainting technique. The parameters focus on the granulation signal and various oscillation signals across different frequency ranges (left, central, and right).

Feature	Duty Cycle	Method	Granulation	Signal (Left)	Signal (Central)	Signal (Right)
f (μHz)	–	Lomb-Scargle	198.9 (3.5)	320.1 (3.7)	300.5 (3.2)	319.1 (3.5)
		Inpainting	153.4 (7.9)	215.8 (10.7)	388.7 (9.6)	193.2 (9.7)
A (ppm$^2\mu$Hz^{-1})	–	Lomb-Scargle	54.9 (5.6)	38.8 (6.8)	76.9 (5.6)	72.0 (14.6)
		Inpainting	82.6 (7.9)	69.5 (8.0)	101.5 (5.5)	208.0 (20.8)
Power Index	0.174 (16)	Lomb-Scargle	0.145 (27)	0.194 (4)	0.061 (17)	0.110 (8)
	0.087 (7)	Inpainting	0.015 (1)	0.120 (10)	0.097 (12)	0.143 (27)
Amplitude	0.188 (8)	Lomb-Scargle	0.036 (6)	0.197 (2)	0.010 (1)	0.020 (6)
	0.105 (12)	Inpainting	0.109 (9.7)	391.8 (7.4)	377.6 (19.1)	

Method

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

Abstract

This study investigates the comparative effectiveness of frequency-domain and time-domain methodologies in the context of ground-based telescope data processing. We specifically analyze the impact of duty cycle variations and signal inpainting techniques on the extraction of granulation signals. By evaluating key features across different spectral regions—left, central, and right—we characterize the performance of these methods under varying observational conditions.

1. Feature Analysis and Signal Granulation

The processing of solar or stellar granulation signals requires high precision to account for atmospheric effects and instrumental noise. In this work, we examine the power spectral density (PSD) and temporal characteristics of these signals. The following table summarizes the measured features, including duty cycles and inpainting performance across different signal components.

The data presented in Table 1 highlights the sensitivity of the granulation signal to the duty cycle. For instance, at a duty cycle of 0.012, the central signal component exhibits a characteristic frequency f with an amplitude of approximately 1095.9 ± 0.4 ppm^2 Hz^{-1}. As the duty cycle increases, we observe a shift in the signal-to-noise ratio and the efficacy of the inpainting algorithms.

2. Frequency Domain vs. Time Domain Methodologies

The choice between frequency-domain and time-domain processing is critical for ground-based observations. Frequency-domain methods typically excel at identifying periodicities and characterizing stochastic noise like granulation. Conversely, time-domain methods are often more robust when dealing with non-continuous data or transient events.

2.1 Granulation Signal Characteristics Granulation signals are characterized by their power distribution across different frequency bands. We categorize these into left, central, and right components to better understand the spectral morphology. The measured values for these components, such as 0.333(16) and 0.470(26), indicate the variance in signal strength as a function of the processing window and duty cycle.

2.2 Impact of Duty Cycle and Inpainting A significant challenge in ground-based astronomy is the presence of gaps in data due to weather or diurnal cycles. Inpainting techniques are employed to reconstruct missing data points. Our results show that for a duty cycle of approximately 0.598, the inpainting error remains relatively low, with values around 0.127(24) for the central signal. However, as the duty cycle decreases or the gap length increases, the reconstruction of the granulation signal becomes increasingly complex, as evidenced by the higher uncertainty values in the lower duty cycle regimes (e.g., 280.9 ± 20.2).

3. Results and Discussion

The comparative analysis reveals that while frequency-domain methods provide a clear view of the granulation power spectrum, they are highly sensitive to the windowing functions necessitated by low duty cycles. Time-domain inpainting, while computationally more intensive, offers a more

Method

Feature Duty Cycle Table 3 Continued Granulation Signal Signal Signal 0.094 (12) 361.1 (59.4) 0.093 (8) 0.102 (7) 0.106 (7) 0.386 (22) 0.365 (33) 0.382 (22) 0.372 (17) 0.596 (11) 0.590 (12) 0.592 (7) 0.597 (8) 0.810 (13) 0.794 (20) 0.803

(13) 0.807 (8) 430.0 (59.6) 406.4 (51.1) 358.2 (39.1) 109.2 (32.7) 44.2 (25.4) 74.2 (21.7) 58.5 (24.5) 223.6 (54.8) 217.2 (36.8) 218.6 (45.0) 195.3 (32.2) 114.8 (43.7) 96.4 (22.3) 118.4 (20.7) 124.5 (27.4) 0.132 (36) 0.133 (38) 0.140 (36) 0.098 (46) 0.102 (36) 0.136 (39) 0.098 (36) 0.127 (31) 0.079 (32) 0.154 (28) 0.150 (27) 0.105 (34) f(cid:22)Appm2(cid:1)(cid:22)Hz(cid:0)1(cid:13)(cid:22)leftcentralright

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