

Characterizing stellar flares in Ariel targets: activity analysis and transit contamination

G. Galletta^{1,2★}, S. Colombo³ and G. Micela³

¹*Dipartimento di Fisica e Chimica, Università di Palermo, Via Archirafi 36, 90128, Palermo, Italy*

²*Blue Skies Space Italia S.R.L., Via Vincenzo Monti 16, 20123, Milano, Italy*

³*INAF, Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, 90134, Palermo, Italy*

Accepted 2026 August 17. Received 2026 July 8; in original form 2026 February 23

ABSTRACT

Stellar flares are sudden releases of magnetic energy that can distort exoplanet transit photometry and transmission spectroscopy, biasing planet radius estimates, transit timings, and atmospheric characterization. Understanding flare activity in Ariel targets is therefore essential to identify stars where flares may compromise observations and to characterize the radiation environment affecting atmospheric escape and photochemistry. We analysed 290 Ariel target stars using *TESS* light curves. Flares were identified via iterative Gaussian process detrending, and their energy distributions were modelled with two-segment power laws. We performed injection–recovery tests by adding synthetic flares to detrended light curves and running the full pipeline to quantify completeness and detection biases. We detected 15 857 flares across 1638 *TESS* sectors, with 2–86 events per sector. We defined a normalized flare index GF.01 to compare activity across stellar luminosities. Near 3% of the sample exhibits enhanced flare activity (GF.01 > 1). AU Mic and HD 28109 show a high likelihood of flare contamination during transit observations. GF.01 correlates negatively with stellar bolometric luminosity, indicating higher relative flare output in lower-luminosity stars. AU Mic is an extreme case: four of five observed transits of AU Mic b are affected by flares, consistent with statistical expectations. We validate the framework by comparing predicted flare-contamination probabilities with observed flare occurrences in a representative subset of transits, finding agreement within uncertainties. These results confirm that energetic flares can significantly impact transit observations and provide quantitative guidance for Ariel target selection and analysis strategies.

Key words: Data Methods – stars – flare – transit contamination.

1 INTRODUCTION

The brightness of solar-type and smaller stars is intrinsically variable on a wide range of temporal and energetic scales, from short, impulsive events lasting seconds to long-term magnetic cycles spanning years–decades (e.g. R. R. Radick et al. 1990). This variability arises from multiple, physically distinct processes, including surface convection and the emergence of photospheric spots driven by the stellar dynamo (e.g. J.-F. Donati et al. 2008), quasi-periodic activity cycles (e.g. S. V. Jeffers et al. 2023), and impulsive releases of magnetic energy through reconnection (e.g. A. F. Kowalski 2024). In some systems, interactions between the star and close-in planets can further modulate activity, producing detectable signatures that differ from purely stellar-driven behaviour (e.g. P. W. Cauley et al. 2019).

Among stellar activity phenomena, flares are particularly important: they are sudden, large releases of magnetic energy that produce intense bursts of electromagnetic radiation across the spectrum on short time-scales (solar analogues were discussed early on by H. Carmichael 1964). Because flares probe the processes of magnetic-field generation and dissipation, their occurrence rates, energies, and temporal profiles provide direct constraints on dynamo operation and magnetic reconnection in stars of different types and ages (e.g. J.-F. Donati et al. 2008; A. Lanza 2009).

Stellar variability, and flares in particular, has multiple direct implications for exoplanet science. Surface heterogeneities, including spots, faculae, and flares, can distort both transit photometry and transmission spectroscopy, potentially biasing inferred planet radii, transit timings, and atmospheric signatures if these effects are not properly accounted for (P. Ballerini et al. 2012). This highlights the importance of incorporating stellar activity models and flare statistics when interpreting exoplanet observations, especially for active stars. In particular, both occulted and unocculted stellar heterogeneities can imprint wavelength-dependent signals on transmission

★ E-mail: gabriele.galletta@inaf.it

spectra, giving rise to the Transit Light Source Effect (TLSE), which can mimic or obscure planetary atmospheric features (B. Rackham et al. 2017; B. V. Rackham et al. 2018). For example, simulations by M. Oshagh et al. (2013) show that stellar spots may lead to underestimates of the planet radius by $\sim 4\%$, errors of transit duration of similar magnitude, and even transit-timing variations (TTVs) up to ~ 200 s, all derivable from anomalies in the light curve caused by occulted or non-occulted spots. Similarly, stellar flares occurring during transits can distort light curves and contaminate transmission spectra, potentially leading to incorrect estimates of transit parameters or misinterpretation of atmospheric features if not properly modelled (e.g. W. S. Howard et al. 2023).

Beyond their observational impact, flares can also physically affect exoplanetary atmospheres. High-energy radiation and particle fluxes associated with flares can change atmospheric chemistry (e.g. ozone and other key species), drive enhanced photochemistry, and contribute to atmospheric escape, thereby modifying the long-term habitability and observability of planetary atmospheres (e.g. A. P. Buccino et al. 2007; A. Segura et al. 2010; C. P. Johnstone 2016). For these reasons, quantifying flare activity is essential both to interpret exoplanet observations reliably and to assess the environmental impact of stellar activity on planets.

A substantial body of work has characterized stellar flare statistics using space-based photometric surveys. Early large-scale flare studies with the *Kepler* mission revealed that flare frequency distributions (FFDs) follow power-law behaviours spanning several orders of magnitude in energy, and that flare rates strongly depend on stellar rotation and spectral type (T. Shibayama et al. 2013; J. R. Davenport et al. 2014). These studies demonstrated that both Sun-like and low-mass stars can produce rare but extremely energetic superflares, establishing a statistical framework for flare occurrence. More recently, the *TESS* mission has significantly expanded flare studies to a much larger and more diverse stellar sample. Systematic analyses have provided flare catalogues for hundreds to thousands of stars, confirming the general power-law behaviour of FFDs while extending them to different stellar populations and ages (W. S. Howard et al. 2019; M. N. Günther et al. 2020; E. Ilin et al. 2021; W. S. Howard 2022). In particular, M-dwarf stars have been shown to exhibit the highest flare rates and energies, making them key targets for studies of stellar magnetic activity and its impact on exoplanet environments (J. R. Davenport et al. 2014; M. N. Günther et al. 2020; E. Ilin et al. 2021). These observational results provide the empirical basis for modelling flare occurrence rates and for interpreting flare contamination in exoplanet transit and atmospheric studies. These studies motivate the need for homogeneous flare analyses across stellar samples, such as the one presented in this work.

The Atmospheric Remote-sensing Infrared Exoplanet Large-survey (Ariel) is the fourth medium-class mission (M4) in ESA's Cosmic Vision programme, scheduled for launch in 2031 (G. Tinetti et al. 2016, 2018, 2022). It is the first space observatory specifically designed to perform a large and homogeneous survey of exoplanet atmospheres, aiming to observe about one thousand transiting planets across a wide range of stellar types and orbital configurations. By combining transit, eclipse, and phase-curve spectroscopy, Ariel will provide precise measurements of atmospheric composition, thermal structure, and cloud properties, thereby enabling a comparative understanding of planetary formation and evolution beyond the Solar system.

Understanding the flaring activity of stellar targets is critical for the success of the Ariel mission since stellar flares can severely impact the quality of atmospheric signals and therefore need to be properly studied, both to mitigate their effects and to exploit them as a diagnostic tool for star–planet interactions. The analysis of stellar flares serves multiple purposes: it allows the identification and exclusion of problematic targets where frequent flaring could compromise transit observations; it provides insights into the planetary radiation environment, with direct implications for atmospheric escape and photochemistry; it enables the inclusion of magnetically active stars in dedicated investigations of star–planet magnetic interactions; and it supports the development of specific analysis techniques to model or remove flare signatures from light curves, ensuring more reliable atmospheric inference.

Building on the methodology originally developed by S. Colombo et al. (2022), who introduced a flare-fitting framework applied to individual active stars such as DS Tuc and AU Mic, G. Galletta et al. (2025) extended this approach to a volume-limited sample of all M-type stars within 10 pc, enabling a systematic characterization of flare activity at the population level. In this study, we further adopt and apply the same flare analysis framework to a different and substantially larger set of stellar targets selected from the Ariel target catalogue, in order to investigate stellar activity across a broader range of spectral types and luminosities. This analysis is performed using *TESS* light curves (G. R. Ricker et al. 2016).

The paper is structured as follows. Section 2 describes the sample, the method used to analyse flare properties, and the definition of the flare energy index. Section 3 shows the results obtained from the analysis. Section 4 presents the discussion of the results in the context of previous works and their implications. In the end in Section 5 summarizes our conclusions.

2 METHODS

In this section, we describe the methodology adopted to analyse the flare activity of the stellar sample. In Section 2.1 we outline the selection criteria and characteristics of the sample. In Section 2.2 we detail the procedure used to identify individual flares and to measure their main physical properties, such as amplitude, duration, and total emitted energy. In Section 2.3 is presented a validation test of the flare detection and characterization pipeline, used to assess completeness and potential systematic effects. Finally, in Section 2.4, we introduce the Flare Energy Index (GF.01), a quantitative parameter defined to characterize the overall flare activity of each star and to enable comparison across different stellar populations.

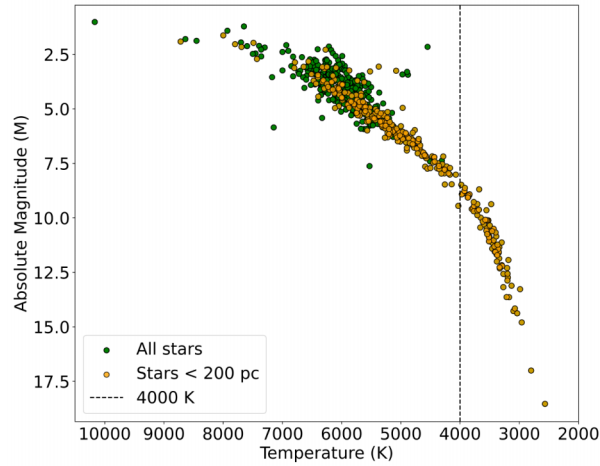


Figure 1. HR diagram of confirmed Ariel targets with *TESS* light curves (green), compared to our sample, which comprises all targets within 200 pc with available *TESS* light curves (290 targets total, orange). The comparison highlights that the 200 pc subsample spans the same temperature–luminosity range, ensuring representativeness of the full sample. The black dotted line represent the upper temperature limit of M dwarfs.

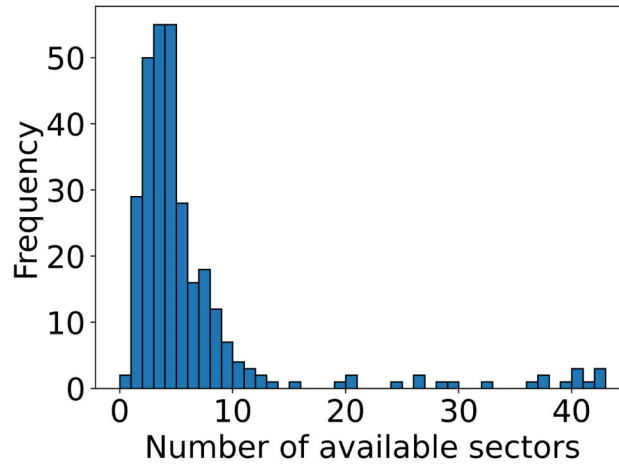


Figure 2. Number of *TESS* sectors available for the 290 targets in our sample.

2.1 The sample

This study is based on the Ariel Mission Candidate Sample (MCS) released in July 2024, which includes confirmed exoplanets only. The Ariel MCS is based on the Ariel target list methodology described by B. Edwards & G. Tinetti (2022) and developed using the ArielRad performance simulator (L. V. Mugnai et al. 2020). The 2024 version of the catalogue also builds upon the procedures and selection criteria first introduced in B. Edwards et al. (2019), ensuring a consistent and optimized target selection for atmospheric characterization within the Ariel mission framework. The 2024 MCS catalogue includes:

- (i) 722 confirmed planets
- (ii) 657 stars (303 stars < 200 pc)
- (iii) 612 stars with available *TESS* light curves
- (iv) 290 stars with *TESS* light curves and distance < 200 pc

For this flare analysis, we focus on the subsample of 290 nearby stars. This analysis is particularly relevant to mission planning: knowing the expected flare probability allows for estimating the number of visits required to reach a given signal-to-noise ratio in the presence of flares, as well as for identifying targets that may need specific data processing.

To contextualize the sample of stars within 200 pc used for the flare analysis, in Fig. 1 we compare their distribution in the Hertzsprung–Russell (HR) diagram with that of the entire sample of 612 confirmed Ariel targets with available *TESS* light curves. Fig. 2 shows the distribution of *TESS* sectors available per star. A larger number of observed sectors improves our ability to constrain the flare distribution and increases the likelihood of detecting rare, high-energy events. Most of the targets in our sample have 4 sectors,

while only 20 are observed in more than 15 sectors. The method used to analyse flares and distribution is based on G. Galletta et al. (2025).

To retrieve the *TESS* data, we used the Lightcurve package (Lightcurve Collaboration 2018) using as search parameters a cadence of 120 s and a search radius of 0.5 arcsec. Specifically, we searched for light curves with 120 s cadence data from the SPOC (Science Processing Operations Center) pipeline. The target columns retrieved for analysis were *TIME* and *SAP_FLUX*, where *SAP_FLUX* refers to the Simple Aperture Photometry flux, i.e. the raw flux extracted by summing the calibrated pixel values within a pre-defined aperture, without correction for instrumental systematics or contamination from nearby sources. In addition, we apply the standard *TESS* quality flags and remove all cadences with non-zero quality indicators to exclude known instrumental artefacts and anomalous data points, following standard *TESS* data reduction practices.

2.2 Measuring flare properties

Following the method used in S. Colombo et al. (2022) and G. Galletta et al. (2025) we identify flares and then for each one, we derived the amplitudes, time-scales, and total energy emitted. To isolate flare signals, we first remove long-term stellar variability from the light curves. This is achieved using an iterative Gaussian Process (GP) regression with a SHOTerm kernel which is a term representing a stochastically driven, damped harmonic oscillator (D. Foreman-Mackey et al. 2017), which is well suited to describe correlated variability on time-scales longer than typical flare durations. The SHOTerm kernel is parametrized by three hyperparameters (amplitude, characteristic period, and damping time-scale), with the adopted priors defined in the range $\sigma \in [10^{-5}, 10^5]$, $P \in [1, 50]$, and $\tau \in [10^3, 10^4]$, where all quantities are expressed in units of the light-curve time axis (2-min *TESS* cadences, corresponding to 120 s). These choices ensure that the Gaussian Process captures variability on time-scales of ~ 1 –10 d, while preserving short-duration flare signals, which typically occur on time-scales of minutes to hours. The lower bound on the characteristic time-scale is therefore well separated from the typical flare duration range adopted in this work, preventing the GP from absorbing impulsive flare-like events.

At each iteration, we compute a Maximum A Posteriori (MAP) GP model and subtract it from the data to obtain residuals. A 3σ clipping procedure is then applied to remove outliers, and the GP is re-fitted to the cleaned data. This process is repeated until the number of outliers is consistent with statistical expectations.

The final GP model produced at the end of this iterative sigma-clipping procedure is therefore constrained to model only the long-term variability in the light curve. Since flare-like events are removed during the iterative sigma-clipping, the GP is not able to reproduce short impulsive signals, ensuring that flares are preserved in the final residual curve (RC), which is obtained by subtracting the final GP model from the original light curve.

Short-term events are identified in the RC as significant local maxima above a 3σ threshold, computed from the statistical properties of the residuals. These events are characterized by a rapid rise, a slower decay, durations of minutes to hours, and peak fluxes exceeding the standard deviation of the residual light curve by at least 3σ , allowing us to distinguish them from residual systematics and noise.

Each flare is modelled using a two-part exponential function, consisting of an increasing exponential for the rise phase and a decaying exponential for the decay phase, joined at the flare peak. This model captures flare morphology and provides robust estimates of amplitudes and characteristic time-scales.

For each star, we construct the cumulative flare energy distribution, representing the number of flares above a given energy. These distributions are fitted using a two-segment power law in logarithmic space. The two segments are separated by an energy break corresponding to a change in slope.

We perform piecewise linear fitting in logarithmic space using a bootstrap resampling approach to estimate the uncertainties on both the slopes and the breakpoint energy. If fewer than six flares are detected for a given star, only a single power-law fit is applied, as the data do not allow a reliable determination of a breakpoint. This condition applies to 316 out of 1638 analysed sectors. Additionally, in cases where fewer than two flares are present above the energy break (66 out of 1638 sectors), only the first slope is retained because we are unable to fit a slope to only one point. Fig. 3 shows both example of the single- and the two-segment fits, respectively using sectors of WASP-189 and AU Mic, the black dashed lines represent the fit and the red dashed vertical line represents the break point.

2.3 Validation test

To evaluate the completeness of the flare detection pipeline and quantify possible biases in the inferred FFDs, we performed injection-recovery simulations using synthetic flare events. Artificial flares were injected into light curves constructed directly from the observed *TESS* photometry of AU Mic, preserving the original temporal sampling, data gaps, and instrumental systematics. This ensures that the simulations fully retain the intrinsic red-noise properties of the stellar variability as well as all observational effects present in the real data set.

Synthetic flares were generated by randomly sampling flare amplitudes and characteristic time-scales over broad parameter ranges representative of the observed flare population. In particular, flare energies were sampled over the interval 10^{31} – 10^{35} erg using a logarithmically spaced grid, while flare durations were sampled between ~ 100 and ~ 2000 s. The injected flare model follows the same two-component exponential model adopted for the flare fitting procedure. The injected flare model is defined as a piecewise exponential

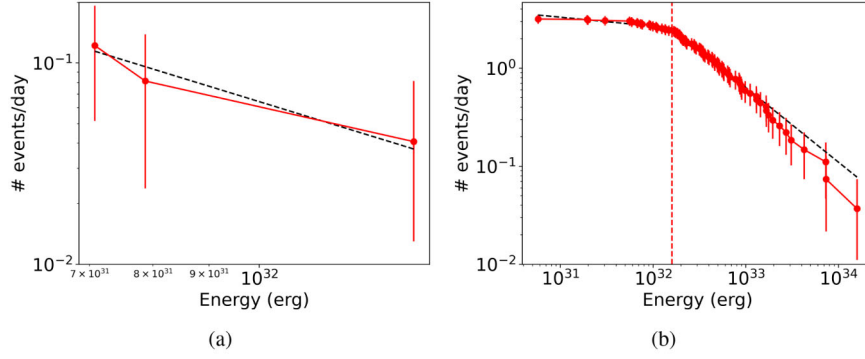


Figure 3. (a) Example of one-segment slope of a sector of WASP-189. (b) Example of two-segment slope of a sector of AU Mic. The red line with error bars is the observed cumulative curve of the star in a sector, the black dashed lines represent the fit and the red dashed vertical line (only in two-segment) represents the break point.

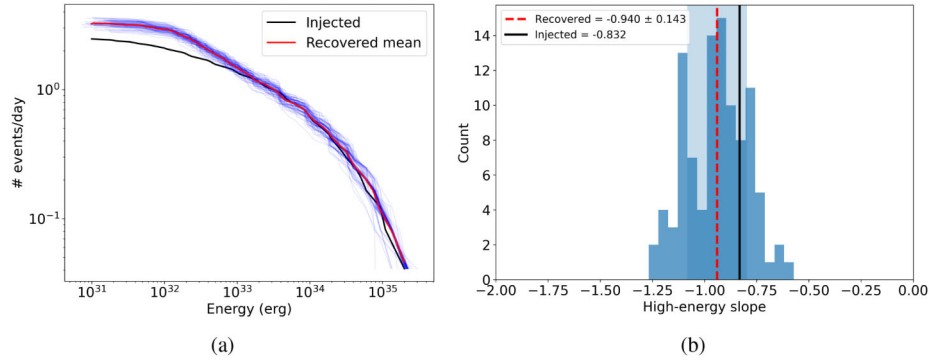


Figure 4. Injection–recovery validation tests of the flare detection pipeline. (a) Comparison between the FFD of injected and recovered synthetic flares. The injected FFD is shown as a black curve, while the recovered distributions from 100 realizations are displayed in blue. The red curve represents the mean recovered FFD. The comparison reveals that the recovered distribution shows an excess of low-energy events, whereas high-energy events are recovered with high efficiency. (b) Distribution of recovered power-law slopes obtained from 100 synthetic realizations with an input slope of -0.832 . The vertical black line indicates the injected slope, while the red dashed line and shaded region show the mean recovered value and its error.

function describing a rapid rise followed by a slower decay:

$$F(t) = \begin{cases} Ae^{\left(\frac{t-t_0}{\tau_r}\right)}, & t \leq t_0 \\ Ae^{\left(-\frac{t-t_0}{\tau_d}\right)}, & t > t_0 \end{cases} \quad (1)$$

where A is the flare amplitude, t_0 is the peak time, τ_r is the rise time-scale, and τ_d is the decay time-scale. The model is evaluated within a finite temporal window of $\pm 5\tau$ around the flare peak to reproduce the typical observable duration of stellar flares in *TESS* light curves.

The rise and decay components are constrained such that $\tau_r : \tau_d = 1 : 10$, corresponding to a total time-scale $\tau = \tau_r + \tau_d$. This choice ensures that the decay phase is systematically longer than the rise phase ($\tau_d > \tau_r$), enforcing physically realistic flare asymmetry consistent with observed stellar flare profiles.

A total of 60 synthetic flare events were injected at random times within the valid observing windows and analysed using the same reduction and detection pipeline adopted throughout this work.

Across the 100 bootstrap realizations, the number of recovered flares ranges from 68 to 89, with a mean value of ~ 79 recovered events out of the 60 injected flares. Fig. 4(a) shows the comparison between the FFDs of the injected and recovered samples. The injected population is shown as a black curve, while the recovered distributions from individual realizations are displayed in blue, and their mean trend is shown in red. We can see that the recovered FFD exhibits an excess of low-energy events. A visual inspection of the detected flares indicates that this excess is mainly produced by noise spikes that occasionally satisfy the adopted 3σ detection criterion and are therefore classified as flare candidates. At higher energies, however, the recovered FFD closely follows the injected distribution, demonstrating that the detection pipeline reliably reproduces the intrinsic flare population in the regime where the signal-to-noise ratio is sufficiently high. This behavior further justifies the choice adopted throughout this work of restricting the analysis to the high-energy portion of the FFD.

Fig. 4(b) shows the distribution of recovered power-law slopes obtained from 100 synthetic flare populations with an injected slope of -0.832 . The distribution peaks close to the injected value, with a mean recovered slope of -0.940 ± 0.143 . This indicates that

the adopted methodology does not introduce significant systematic biases in the determination of the high-energy slope and that the observed offset remains small compared with the statistical uncertainties.

For this reason, throughout the analysis we adopt a two-segment power-law model and derive the GF.01 index (see Section 2.4) exclusively from the high-energy slope. The injection–recovery simulations show that this portion of the FFD provides the most reliable characterization of the flare population, whereas the low-energy regime is more strongly affected by contamination from spurious detections generated by noise fluctuations near the detection threshold.

2.4 The flare energy index (GF.01)

To quantify flare activity in a normalized way, we adopt the GF.01 index, introduced in G. Galletta et al. (2025), which quantifies the flare activity normalized to the stellar bolometric luminosity. The index represents the energy level at which the flare frequency equals 0.1 flares per day. This index is derived from the fitted power-law parameters (slope and intercept) and the star’s bolometric luminosity. GF.01 allows comparison across stars of different luminosities by normalizing the flare frequency–energy relation. The GF.01 is given by the following equation:

$$\text{GF.01} = \frac{0.1 - q}{m} / L_{\text{bol}}, \quad (2)$$

where m is the slope of the high-energy segment of the two-segment flare frequency distribution (i.e. the portion unaffected by biases at low energies), q is the intercept, L_{bol} is the bolometric luminosity of the star [we used the Virtual Observatory SED Analyser (VOSA) to compute the bolometric luminosity of the stars (C. Rodrigo et al. 2020)], and 0.1 indicates the characteristic energy of flares occurring on average every 0.1 d.

GF.01 is computed independently for each sector of a given star, rather than on a star-averaged flare distribution. As a result, multiple GF.01 values may be associated with the same star. We compute GF.01 only for sectors with well-constrained two-segment fits. Sectors in which fewer than six flares are detected, and for which only a single power-law fit is available, are excluded from the analysis, as the absence of a well-defined high-energy slope prevents a reliable determination of the index. The use of the high-energy slope ensures that the derived values are not biased, which predominantly affects the low-energy regime of the flare frequency distribution (see Section 2.3).

Uncertainties on GF.01 are estimated through standard error propagation from the uncertainties on the fitted parameters. Assuming uncertainties σ_m , σ_q , and $\sigma_{L_{\text{bol}}}$, the uncertainty on GF.01 is given by:

$$\sigma_{\text{GF.01}} = \text{GF.01} \cdot \sqrt{\left(\frac{\sigma_q}{0.1 - q}\right)^2 + \left(\frac{\sigma_m}{m}\right)^2 + \left(\frac{\sigma_{L_{\text{bol}}}}{L_{\text{bol}}}\right)^2}, \quad (3)$$

where the uncertainties on m and q are derived from bootstrap resampling of the flare frequency distribution fits.

This normalization accounts for differences in stellar luminosity, thereby enabling an assessment of how efficiently a star converts its available energy into flare energy. A higher GF.01 value indicates that, on average, the flares are more energetic, whereas a lower GF.01 value corresponds to less energetic flares.

3 RESULTS

In this section, we present the main results of our analysis. Overall, our flare-search pipeline identified a total of 15 857 flare events across 290 analysed stars and 1638 *TESS* sectors. The number of detected flares per sector ranges from a minimum of 2 up to a maximum of 86 events. In Section 3.1 we report the outcomes of the slope analysis, which provides information on the shape and energy distribution of the detected flares. In Section 3.2 we discuss the behaviour of GF.01 across the stellar sample and its dependence on stellar parameters. In Section 3.3 we evaluate the probability that stellar flares occur during planetary transits, using the FFDs to estimate the expected occurrence of energetic events during Ariel observations and to assess their potential impact on exoplanet detection and characterization.

3.1 Slope analysis

The FFDs of the analysed stars are constructed as cumulative occurrence rates of flares above a given energy and are modelled with a two-segment power-law fit following the methodology described in G. Galletta et al. (2025). The two-segment model consists of two power laws joined at a break energy, and it is introduced to separate the energy range where observational biases dominate from the range where the flare occurrence can be robustly characterized.

At low energies, the observed distribution is strongly affected by detection biases: smaller flares have lower amplitudes and are progressively more difficult to detect above the photometric noise floor, leading to a flattening of the measured FFD. In this regime, the low-energy slope derived from the fit does not necessarily represent the intrinsic flare production physics of the star. This first segment therefore serves mainly to identify the onset of biases at low energies. In contrast, the high-energy segment corresponds to the part of the distribution where the sample is more complete with respect to detection sensitivity, flares above the break energy are reliably

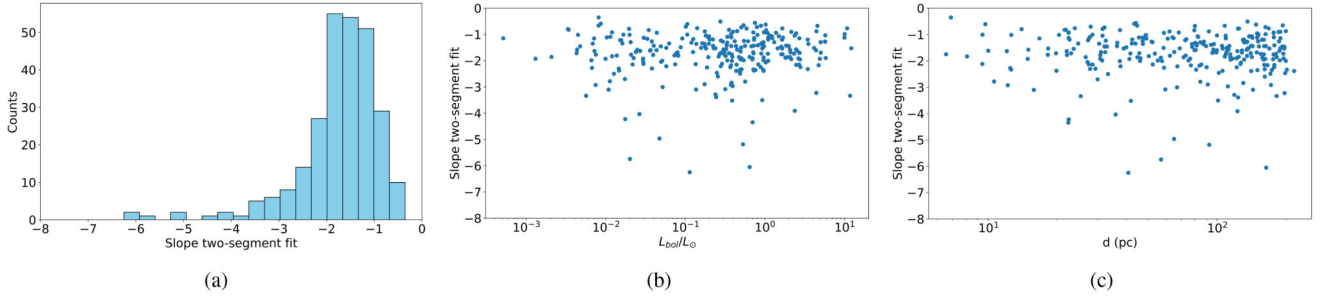


Figure 5. (a) Distribution of the high-energy slopes derived from the two-segment power-law fits of the flare frequency distributions (FFDs). The distribution shows a clear peak around -1.8 , with a mean value of -1.87 , a median of -1.59 , and a standard deviation of 1.55 , indicating a broad but well-defined range of flare activity across the sample. (b) High-energy slope versus stellar bolometric luminosity (L_{bol}). No clear correlation is observed, although more luminous stars tend to show a reduced dispersion in slope values, suggesting a more homogeneous flare energy distribution in this regime. (c) High-energy slope versus stellar distance. No significant trend is detected across the explored distance range, indicating that the derived slopes are not strongly affected by distance-dependent observational biases.

detected across the full data set. The high-energy slope of the FFD therefore provides a more robust characterization of the intrinsic flare occurrence rate, and can be interpreted in terms of the underlying magnetic energy-release processes. A steeper high-energy slope indicates that energetic flares are relatively rare compared to lower energy ones, whereas a shallower slope implies a relatively higher incidence of more energetic flares.

The high-energy slope is particularly relevant because it determines the occurrence rate of energetic flares, which are the events most likely to be detectable and to significantly affect planetary transit observations. Energetic flares produce large and rapid increases in stellar flux, often comparable to or exceeding the transit depth, and can therefore alter key transit parameters such as the measured transit depth, duration, and ingress/egress shape if they occur during a transit. Because the FFD quantifies the rate of flares above a given energy threshold, the high-energy power-law index allows us to estimate the expected number of flares with energies $E \geq E_{\text{threshold}}$ occurring within a finite time window, such as the duration of a planetary transit. If the slope is flatter, energetic events are relatively more frequent, increasing the likelihood that a flare will coincide with an Ariel observation and contaminate the light curve. In contrast, steeper slopes indicate that high-energy events are rare, lowering the probability of strong flare contamination but still allowing moderate activity to affect transit depth measurements.

Fig. 5 presents the results of the two-segment power-law fit to the FFDs of the analysed stars. Fig. 5(a) shows the distribution of the high-energy slopes derived from the fits, which displays a clear peak around a value of approximately -1.8 , representing the most common behaviour within the sample. The mean high-energy slope is -1.87 , with a standard deviation of 1.55 and a median value of -1.59 . These values are broadly consistent with previous studies of stellar FFDs, which have shown that flare energies follow power-law statistics over several orders of magnitude (e.g. S. L. Hawley et al. 2014; M. J. Aschwanden & M. Güdel 2021). In particular, large-scale analyses based on *Kepler* data, such as the Kepler Catalogue of Stellar Flares (J. R. Davenport 2016), demonstrate that stellar flare populations are well described by power laws, although the inferred slopes depend on stellar type, activity level, and completeness effects. Studies of M dwarfs, including S. L. Hawley et al. (2014), show that FFDs are consistent with power-law behaviour at high energies, with deviations at low energies primarily due to observational biases. More recent analyses of *TESS* data, such as A. Rajpurohit et al. (2025), find cumulative FFD slopes corresponding to high-energy slopes close to -2 , in agreement with previous *Kepler* results. Additional work on active M dwarfs in large *TESS* samples (e.g. B. K. Capistrant & J. Dittmann 2026) confirms that typical slopes lie in the range -1.5 to -2.0 , although significant star-to-star scatter is observed due to intrinsic variability and detection biases. From a theoretical perspective, such power-law distributions are naturally explained by self-organized criticality models of magnetic reconnection, which predict scale-free flare energy release processes (M. J. Aschwanden & M. Güdel 2021). Overall, the average slope derived in this work (-1.87) lies well within the range reported in the literature and is consistent with results obtained for both individual active stars and large stellar samples, including those presented in G. Galletta et al. (2025). This agreement indicates that the flare energy distributions in the analysed sample follow the same statistical behaviour observed in previous studies, supporting the robustness of the adopted detection and fitting methodology.

Fig. 5(b) shows the high-energy slope as a function of the stellar bolometric luminosity (L_{bol}). No clear correlation between the two quantities is observed. However, stars with higher bolometric luminosities tend to exhibit a smaller dispersion in slope values compared to less luminous stars, suggesting a more homogeneous flare energy distribution within this subset of the sample. Finally, Fig. 5(c) presents the high-energy slope as a function of stellar distance. No significant trend is observed across the explored distance range. Any apparent variation is likely dominated by observational effects, such as changes in detection sensitivity with distance, rather than reflecting an intrinsic physical dependence.

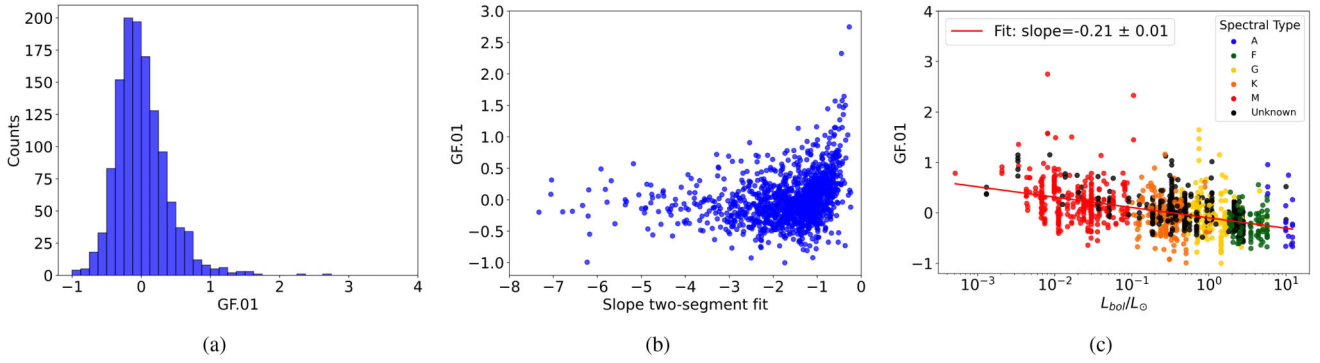


Figure 6. (a) Distribution of the GF.01 index, which represents the flare energy corresponding to a frequency of 0.1 flares per day, normalized to the stellar bolometric luminosity. (b) GF.01 versus the high-energy slope of the two-segment flare frequency distribution, illustrating the relationship between flare energetics and the power-law behaviour of the distribution. (c) GF.01 versus stellar bolometric luminosity (L_{bol}), showing a clear anticorrelation between flare activity and stellar energy output. The red solid line indicates the best-fitting linear relation in logarithmic space. The points are colour-coded according to their spectral type classification (A, F, G, K, M), while black points correspond to sources for which no reliable spectral type could be retrieved.

3.2 GF.01 analysis

The GF.01 allows us to identify which stars are intrinsically more flare-active relative to their energy budget, making it a valuable diagnostic for evaluating the potential impact of flares on planetary environments and on the correct measurements of the Ariel observations. By studying the distribution and correlations of GF.01, we can compare the activity of the Ariel sample with the sample of nearby M dwarfs and if their different behaviours is linked to selection effects, stellar type, or distance. In particular, combining GF.01 with the slopes of the FFDs provides insight into the relative occurrence of high-energy events, while the comparison with L_{bol} tests whether the scaling of activity with luminosity remains consistent across diverse stellar spectral types. This framework is useful to estimate the probability of flare contamination during planetary transits and to better characterize the stellar environments probed by Ariel.

Fig. 6(a) shows the overall distribution of GF.01, which displays a clear peak around zero, indicating that most stars exhibit relatively low or moderate flare activity compared to their bolometric output. A high-value tail is also visible, highlighting the presence of a small subset of particularly active stars.

Fig. 6(b) presents the comparison between GF.01 and the high-energy power-law slope derived from the two-segment model. Each data point corresponds to a single *TESS* sector rather than to an individual star, since both the slope and GF.01 are computed independently for each sector. The majority of data points cluster around $\text{GF.01} \approx 0$ and slope values between -2 and -1 , suggesting that stars with intermediate flare activity tend to exhibit steep slopes, corresponding to a rapid decline in flare frequency with increasing energy.

Finally, Fig. 6(c) shows the relation between GF.01 and stellar bolometric luminosity (L_{bol}), the points are colour-coded according to their spectral type classification (A, F, G, K, M), while black points correspond to sources for which no reliable spectral type could be retrieved, the red solid line represents the fitted model. A clear negative relation emerges, confirming that less luminous stars display higher relative flare activity. To quantify this behaviour, we performed a linear fit in logarithmic space, obtaining a slope of $m = -0.206 \pm 0.011$ and an intercept of $q = -0.100 \pm 0.011$. The fit yields a χ^2 value of 0.110, indicating a statistically consistent description of the observed trend within the data uncertainties. It is worth noting that GF.01 is computed using only high-energy flares, for which the detection completeness is significantly improved. This reduces the impact of observational biases related to flare detectability, ensuring that the derived trend is primarily driven by intrinsically robust flare events. This behaviour aligns with the findings of G. Galletta et al. (2025) for M-dwarf stars, where a similar trend was interpreted as evidence that magnetic activity represents a more significant fraction of the total energy output in lower luminosity stars.

3.3 Flare probability

The primary goal of this analysis is to quantify the likelihood that an Ariel transit observation is affected by stellar flares.

Before performing the flare analysis, all planetary transits were removed from the *TESS* light curves using literature ephemerides. For each planet, we adopted the published transit epoch, orbital period, and transit duration, and propagated the linear ephemeris across the *TESS* observing window. Transit mid-times were computed iteratively starting from the reference epoch and stepping forward and backward in time using the orbital period to cover the full time span of the observations. For each predicted transit, we removed data points within a window centred on the transit mid-time and extending to half the transit duration on either side. To account for uncertainties in the ephemerides and possible timing drifts over time, we applied an additional buffer by enlarging the exclusion window by 10% of the transit duration. This ensures that ingress and egress phases, as well as potential deviations from the predicted timing, are

fully excluded from the analysis. This step is essential to avoid spurious flare detections caused by imperfect transit modelling causing residuals at ingress and egress, which could otherwise artificially inflate the inferred flare rates. The flare statistics are thus derived exclusively from out-of-transit stellar variability.

The probability estimation relies on the FFD of each star, modelled as a two-segment power law. In this work we assume that the energetic slope is an intrinsic estimate of the flare frequency. Using these fits, we estimate the expected rate of flares above a given energy threshold, expressed as a fraction of the stellar bolometric luminosity, $E_{\text{cut}} = f \cdot L_{\text{bol}}$. Two thresholds are considered in this work: a conservative threshold corresponding to very energetic events ($f = 1.0$, 100% of L_{bol}) and a lower threshold representative of less intense but still potentially disruptive flares ($f = 0.1$, 10% of L_{bol}).

To link the flare statistics to the observing conditions of Ariel, we define a temporal window representative of a typical transit visit. This window is set equal to twice the planetary transit duration, effectively assuming that flares occurring up to 50% before ingress or 50% after egress may still affect the observation through baseline distortions or biased detrending. An implicit assumption of this approach is that individual flare durations are generally shorter than the adopted window, which is justified for the majority of optical flares observed in the sample. Under this assumption, the probability of overlap between a flare and the transit window can be treated using a Poisson formalism.

For each star and energy threshold, the flare occurrence rate inferred from the appropriate segment of the FFD is multiplied by the observing window to obtain the expected number of flares that can contaminate the Ariel observation,

$$\lambda = \text{rate} \times t_{\text{window}}. \quad (4)$$

The probability of observing at least one flare during a single transit visit is then given by

$$P_{\text{flare}} = (1 - e^{-\lambda}) \times 100. \quad (5)$$

Only stars with well-constrained two-segment FFD fits and with a measured minimum flare energy lower than the adopted threshold are included in this calculation. This criterion is satisfied by 184 stars when adopting a threshold of 10% L_{bol} , and by 229 stars when adopting a threshold of 100% L_{bol} . This ensures that the probability estimates rely on an interpolation anchored to the observed flare distribution rather than on an extrapolation beyond the energy range covered by the data.

In order to assess the robustness of the flare probability estimates, we quantified the level of agreement between the flare occurrence rates inferred from the fitted FFDs and those directly measured from the observed cumulative flare distributions. For each star and for each observing sector, we computed the relative percentage difference between the slope-predicted flare frequency at the adopted energy threshold, f_{slope} , and the observed frequency derived from the cumulative flare distribution, f_{obs} , evaluated at the same energy cut. The percentage difference is defined as

$$\Delta \% = \frac{|f_{\text{slope}} - f_{\text{obs}}|}{f_{\text{obs}}} \times 100, \quad (6)$$

and provides a direct diagnostic of how well the slope reproduces the empirical flare statistics in the energy range relevant for the probability calculation.

The observed frequency f_{obs} is obtained by interpolating the cumulative flare distribution in logarithmic energy space around the selected energy threshold, ensuring that the comparison is not biased by the discrete sampling of flare energies. The frequency f_{slope} is computed from the slope of the fitted FFD (below or above the break energy), consistently with the procedure adopted for the flare probability estimation. This comparison is performed independently for each sector, allowing us to capture possible sector-to-sector variations due to limited statistics, instrumental effects, or intrinsic stellar variability.

Large values of $\Delta\%$ indicate cases in which the FFD fit significantly overestimates or underestimates the observed flare rate at the relevant energy scale. In the context of transit contamination probabilities, an overestimation of the flare rate directly translates into an overestimated probability of flare occurrence during a transit visit. To limit the impact of such biases, we adopt a conservative quality criterion and exclude individual sectors for which $\Delta\% \geq 30\%$ from the probability analysis. This threshold represents a compromise between retaining a statistically meaningful sample and avoiding sectors where the inferred probabilities would be dominated by poorly constrained or unreliable extrapolations of the FFD.

Applying this criterion results in the exclusion of 103 sectors out of a total of 831 analysed sectors ($\approx 12\%$) for the more stringent threshold of 1.0 L_{bol} . For the lower threshold of 0.1 L_{bol} , fewer sectors are excluded: 53 out of 744 analysed sectors ($\approx 7\%$). These distributions of $\Delta\%$ for the two thresholds are shown in Fig. 7 (threshold = 1) and Fig. 8 (threshold = 0.1), highlighting that the fraction of excluded sectors is limited, indicating that the quality cut removes only the most discrepant sectors while preserving the majority of the data.

The effect of this selection depends on the number of available sectors per star.

Globally 0 stars are completely discarded from the final sample for either threshold. However, the number of available sectors is reduced for 11 stars in the 0.1 L_{bol} case and for 18 stars in the 1 L_{bol} case. The rejected sectors are spread along different stars, with no obvious clustering of bad sectors on few objects. Table 1 summarizes the number of excluded sectors relative to the total number of available sectors for stars with more than 1 excluded sector.

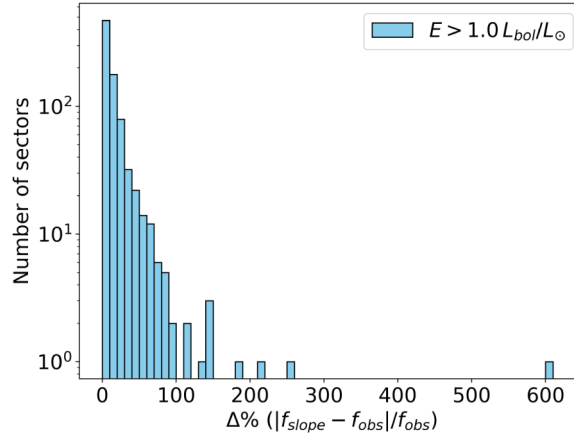


Figure 7. Distribution of the percentage difference $\Delta\%$ between the flare occurrence rates predicted from the fitted FFDs and those directly measured from the cumulative observed flare distributions, evaluated at an energy threshold of $1.0, L_{\text{bol}}$. The quantity $\Delta\%$ quantifies the agreement between the model-based slope extrapolation and the empirical flare rate in each *TESS* sector, with larger values indicating sectors where the FFD fit significantly over or underestimates the observed flare frequency. The distribution shows that most sectors are in good agreement with 12% (103/831 sectors) exceeding the 30% quality cut applied in the analysis.

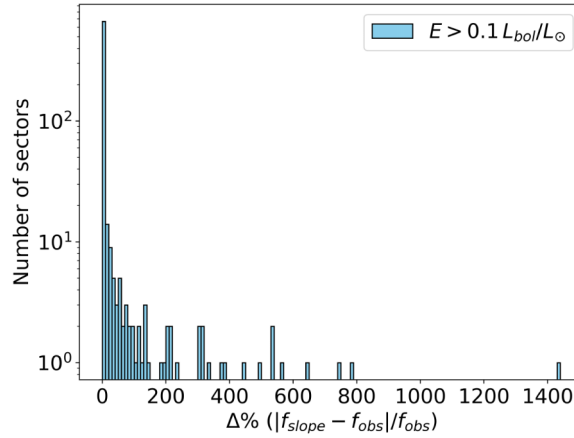


Figure 8. Distribution of the percentage difference $\Delta\%$ between flare occurrence rates predicted from the fitted FFDs and those directly measured from the cumulative observed flare distributions, evaluated at an energy threshold of $0.1, L_{\text{bol}}$. The metric $\Delta\%$ provides a direct sector-by-sector diagnostic of the reliability of the FFD-based extrapolation used in the flare probability framework. As in the higher threshold case, most sectors show good agreement between predicted and observed flare rates, with 7% (53/744 sectors) exceeding the 30% quality cut applied in the analysis.

This approach ensures that the final flare probability estimates are based on sectors where the fitted FFD provides a faithful representation of the observed flare activity, while minimizing the risk of artificially inflating the contamination probabilities for Ariel transit observations. By explicitly quantifying and controlling the slope–data discrepancies, we improve the reliability of the inferred probabilities and strengthen the statistical foundation of the results presented in Fig. 9.

Fig. 9 shows the distribution of flare probabilities for the two energy thresholds. For the most energetic events ($E_{\text{cut}} = 1L_{\text{bol}}$), the vast majority of stars exhibit probabilities between 0% and 5%, indicating that such extreme flares are unlikely to occur during a single transit observation. When $E_{\text{cut}} = 0.1L_{\text{bol}}$, the distribution broadens and a small subset of stars reaches substantially higher probabilities, exceeding 40% in some cases. These objects, including AU Mic and HD 28109, are among the most active stars in the sample and contribute significantly to the high-probability tail.

Overall, this analysis shows that while most Ariel targets are unlikely to experience strong flare contamination during a single transit observation, a small number of highly active stars represent a non-negligible risk. The subset of targets with a flare contamination probability exceeding 20% for the $0.1 L_{\text{bol}}$ threshold is reported in Table 2. These targets may require special attention in scheduling, data reduction, or interpretation, as stellar flares could represent a dominant source of astrophysical noise during transit observations. For the more stringent threshold of $1 L_{\text{bol}}$, only AU Mic exceeds this level.

Finally, it is worth noting that flare contamination is not limited to total emitted energy value. The duration and amplitude of individual events play a crucial role in determining their impact on transit observations, as short, intense flares may affect light curves

Table 1. Stars with more than one excluded sector ($\Delta\% \geq 30\%$) for the two thresholds. Left: threshold = $1 L_{\text{bol}}$. Right: threshold = $0.1 L_{\text{bol}}$.

Star ($1 L_{\text{bol}}$)	$\text{Sect}_{\text{excl}}/\text{Sect}_{\text{tot}}$	Star ($0.1 L_{\text{bol}}$)	$\text{Sect}_{\text{excl}}/\text{Sect}_{\text{tot}}$
WASP-62	9/25	HD23472	3/10
TOI-1670	6/33	TOI-1670	3/33
TOI-201	5/27	TOI-500	3/7
HD28109	4/34	TOI-700	3/23
DS Tuc A	4/5	HD110113	2/3
L98-59	3/20	HD191939	2/21
TOI-1450A	3/30	HD28109	2/34
HD23472	2/10	TOI-1246	2/32
K2-266	2/3	TOI-1450A	2/30
KELT-20	2/6	TOI-481	2/16
GJ357	2/4	WASP-62	2/25
HD191939	2/21	–	–
TOI-1759	2/8	–	–
HD108236	2/3	–	–
TOI-481	2/16	–	–
TOI-500	2/7	–	–
TOI-700	2/23	–	–
TOI-712	2/24	–	–

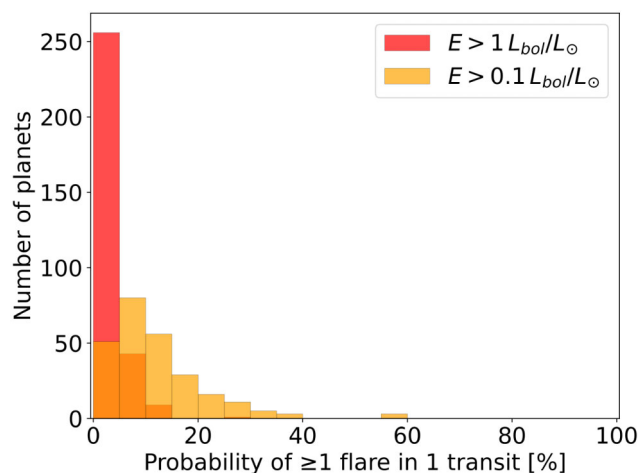


Figure 9. Distribution of flare probability during Ariel visit. The orange are the stars with the probability to have flare of the 10% of energy of the L_{bol} of the star, the red are the stars with the probability to have flare of the 100% of energy of the L_{bol} .

differently than longer, moderate events. Although correlations between flare energy, duration, and amplitude are expected, a detailed exploration of these relationships lies beyond the scope of this work and will be addressed in future studies.

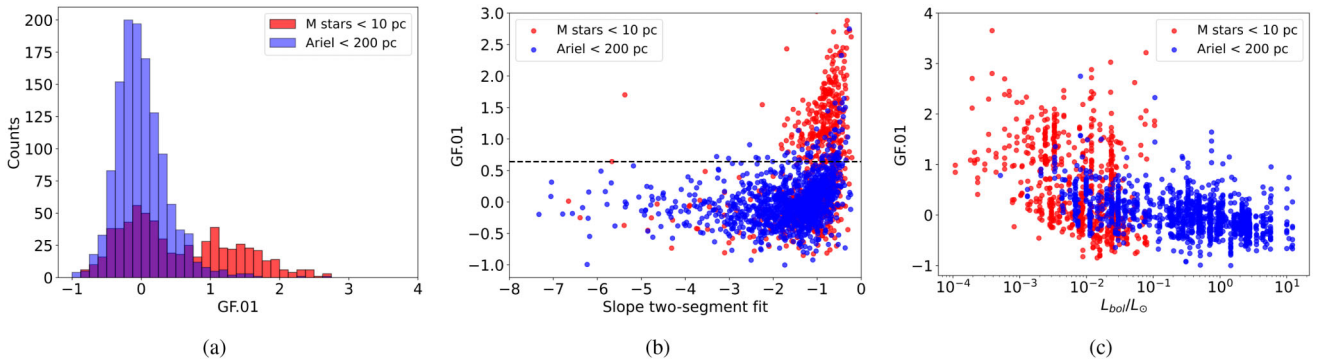
Table 1 listing all systems for which a flare contamination probability could be computed. The table reports the stellar bolometric luminosity, the high-energy FFD slope, the index GF.01, and the flare contamination probabilities for the two adopted energy thresholds (1.0 and $0.1 L_{\text{bol}}$). The table is publicly available at <https://zenodo.org/records/18773384>.

4 DISCUSSION

In this section, we extend the analysis by comparing the results obtained for the Ariel target sample with those derived from other stellar populations and by examining specific case studies that provide additional physical context and validation of our statistical framework. In Section 4.1, we compare the Ariel targets with the M-dwarf sample of G. Galletta et al. (2025), highlighting differences in flare activity and statistical trends across stellar populations. In Section 4.2, we first focus on AU Mic and DS Tuc as extreme cases of stellar activity, examining in detail their flare properties and the occurrence of flares within planetary transit events. We then extend the analysis to a broader and more representative sample of targets, selected to reflect the typical contamination regime of our full data set. For this larger sample, we perform a systematic comparison between predicted flare contamination probabilities and direct *TESS* observations of transits. Taken together, these two complementary approaches allow us to validate the statistical flare probability framework both in extreme conditions and in the typical regime, demonstrating its ability to reproduce, within statistical uncertainties, the flare contamination rates observed in real transit data.

Table 2. High-risk targets with flare contamination probability $> 20\%$ for a threshold of $0.1 L_{\text{bol}}$.

Target	Probability (%)	Target	Probability (%)
HD28109 d	59.1	TOI-257 b	26.8
HD28109 c	57.4	DS Tuc A b	26.4
AU Mic b	57.0	TOI-444 b	26.4
HD63935 c	37.7	HD106315 c	26.4
HD1397 b	36.4	TOI-1136 b	25.6
TOI-4641 b	35.3	K2-266 c	25.5
HIP67522 b	32.9	HD63935 b	25.4
KELT-11 b	32.6	WASP-99 b	25.4
TOI-1670 d	32.3	TOI-2443 b	24.8
HD221416 b	32.0	TOI-712 d	24.2
V1298Tau b	30.7	HIP29442 b	24.0
HD332231 b	29.7	HD202772A b	23.8
HAT-P-2 b	28.9	HD89345 b	23.6
HD191939 d	26.9	TOI-561 e	22.9
TOI-199 b	22.8	TOI-712 e	22.7
KELT-2A b	22.1	TOI-1136 d	21.6
TOI-1136 e	21.2	HD191939 c	21.1
TOI-1246 b	21.1	K2-266 b	21.0
TOI-622 b	20.8	TOI-561 d	20.8

**Figure 10.** (a) Distribution of the GF.01 index, defined as the flare energy corresponding to a frequency of 0.1 flares per day and normalized to the stellar bolometric luminosity. (b) GF.01 versus the high-energy slope of the two-segment flare frequency distribution. (c) GF.01 versus stellar bolometric luminosity (L_{bol}). In all panels, the blue points represent stars from the Ariel sample within 200 pc, while the red points correspond to M-dwarf stars from the sample of G. Galletta et al. (2025) within 10 pc. This comparison highlights differences in flare activity properties between nearby low-mass stars and the broader Ariel target sample.

4.1 Comparison with M stars sample

To better understand how the flare properties of the Ariel targets compare with those of well-studied nearby active late-type stars, we performed a direct comparison between the Ariel sample and the M-dwarf sample analysed by G. Galletta et al. (2025). This comparison allows us to assess whether the trends observed in the local population of cool, magnetically active stars persist in the more diverse and distant set of stars selected for the Ariel mission.

Fig. 10 presents the distributions and correlations of the GF.01 index for the two samples: the Ariel targets (blue, stars within 200 pc) and the M-dwarf sample (red, stars within 10 pc). The Ariel sample spans a wide range of spectral types and effective temperatures, whereas the nearby M-dwarf sample includes only cool stars below ~ 4000 K. This distinction is clearly visible in Fig. 1, where the dashed vertical black line marks the upper temperature boundary of M dwarfs. The Ariel stars extend beyond this regime. Furthermore, the Ariel target selection might introduce a bias against highly active stars, since strong activity complicates the detection and characterization of planets. This could partially explain the absence of very active stars in the Ariel data set compared to the M-dwarf one. Fig. 10(a) shows the distribution of the GF.01 index. The Ariel sample exhibits a single peak around zero with a modest positive tail. In contrast, the M-dwarf sample reveals a bimodal distribution, with one population centred near zero and another at higher GF.01 values, corresponding to a subset of highly active stars. This second peak, evident in the M dwarfs, is absent in the Ariel sample possibly due to selection effects or intrinsic differences in stellar types.

Fig. 10(b) compares GF.01 with the high-energy slope from the two-segment power-law fit. Both samples follow a general trend in which higher flare activity, as measured by GF.01, corresponds to flatter slopes indicating that energetic flares become more frequent as overall activity increases. However the M-dwarf sample spans a wider range of GF.01 values and includes a greater number of highly active stars. The Ariel sample, instead, remains more tightly clustered around $\text{GF.01} \approx 0$.

Finally, Fig. 10(c) displays the relation between GF.01 and the stellar bolometric luminosity (L_{bol}). As discussed for the Ariel sample (see Section 3.2), a negative trend is observed, with less luminous stars tending to exhibit higher relative flare activity. For the M-dwarf sample, we also performed a linear fit, obtaining a slope of $m = -0.559 \pm 0.058$ and an intercept of $q = -0.608 \pm 0.128$. The fit yields a χ^2 value of 0.689, indicating an overall good agreement between the model and the observed distribution, indicating a statistically significant anticorrelation, steeper than the one observed in the Ariel data set and in agreement with the results of G. Galletta et al. (2025).

This behaviour demonstrates that the relationship between stellar luminosity and relative flare strength extends beyond the nearby M dwarfs to the broader, more diverse Ariel target set. Overall, while the Ariel stars show a narrower range of activity levels, the consistency of these trends suggests that the fundamental connection between stellar energy output and magnetic activity is the same across different stellar populations.

4.2 Extreme cases and method validation

AU Mic emerges as an extreme case within our sample. For a flare energy threshold of $0.1 L_{\text{bol}}$, the estimated probability of observing at least one flare during a single transit event reaches approximately 57%, making AU Mic the most flare-active target in our analysis. This exceptionally high probability reflects the intense and persistent magnetic activity of the star and places it at the upper end of the activity distribution among the analysed targets.

The flare behaviour of AU Mic has been extensively characterized in previous studies (A. D. Feinstein et al. 2022; E. Ilin & K. Poppenhäger 2022; I. I. Tristan et al. 2023), notably by S. Colombo et al. (2022), who presented its cumulative FFD and the statistical properties of flare amplitudes and decay times. Using the methodology developed by G. Galletta et al. (2025), the present analysis achieves an increased sensitivity to low-energy flare events compared to previous approaches. This is primarily due to the flare-detection and fitting strategy, which is optimized to identify weak and short-lived events that would otherwise be difficult to detect. In particular, while previous studies (e.g. E. Martioli et al. 2021; S. Colombo et al. 2022) typically probe flare energies down to $\sim 10^{31}$ erg, our analysis reaches energies of $\sim 10^{30}$ erg. However, we emphasize that this low-energy regime corresponds to the biased-affected part of the flare frequency distribution, as also shown by our injection–recovery tests, and is therefore not used in the derivation of the main physical results of this work. This extension toward lower energies reflects an improved sensitivity of the detection pipeline, rather than a direct increase in the intrinsic flare occurrence rate.

However, while the enhanced detection of low-energy flares improves the completeness of the flare sample, these events do not play a dominant role in the context of this work. In particular, they do not significantly contribute to the determination of the high-energy segment of the FFD, which is described by the second power-law slope and is the key quantity used to characterize the overall flare activity level relevant for planetary transit contamination. The second slope is primarily constrained by energetic flares above the break energy, and therefore remains largely unaffected by the improved sensitivity at low energies.

At high energies, AU Mic exhibits a remarkably high flare occurrence rate. From our analysis, the star produces about three flares per day with energies greater than or equal to 10^{32} erg, placing it among the most active known M dwarfs. This behaviour is comparable to that observed in benchmark active stars such as EV Lac, which shows a similar flare rate at comparable energy thresholds (G. Galletta et al. 2025). Such high-frequency, energetic flare activity indicates an efficient and sustained release of magnetic energy, consistent with AU Mic’s young age and the presence of a powerful magnetic dynamo.

The statistical properties of the flare population further reinforce this picture. In our analysis, the flares span a wide range of amplitudes and decay times, with events lasting up to approximately 1500 s. The distribution of amplitudes and time-scales is broadly consistent with that observed in other highly active M dwarfs (G. Galletta et al. 2025), although AU Mic tends to exhibit slightly shorter decay times than EV Lac. The presence of long-duration and energetic flares is particularly relevant for transit observations, as such events can overlap with planetary transits and significantly bias photometric and spectroscopic measurements.

Overall, AU Mic combines extremely high flare occurrence rates with energetic and, in some cases, long-duration events. Quantitatively, AU Mic exhibits an average slope of -0.606 ± 0.156 , together with a high flare rate indicator of $\text{GF0.1} = 1.885 \pm 0.311$, placing it at the extreme end of the activity distribution within our sample.

A further validation of the extreme flare activity inferred from the statistical analysis is provided by a direct inspection of the available transit observations of AU Mic. The star was observed by *TESS* in two sectors, Sector 1 and Sector 27, during which a total of five planetary transits were observed. Remarkably, four out of these five transits are affected by at least one detectable stellar flare, corresponding to an empirical contamination rate of 80%. This observational result is fully consistent with the high flare probability estimated from the FFD and provides an independent, event-level confirmation of the statistical predictions.

Examples of flare contamination during *TESS* transit observations are shown in Fig. 11. Fig. 11(a) illustrates a case for AU Mic b, where a flare occurs near the centre of the transit. As already discussed, AU Mic represents an extreme and highly active system, in which strong flare contamination is relatively frequent. In this configuration, a flare overlapping the transit core produces a temporary

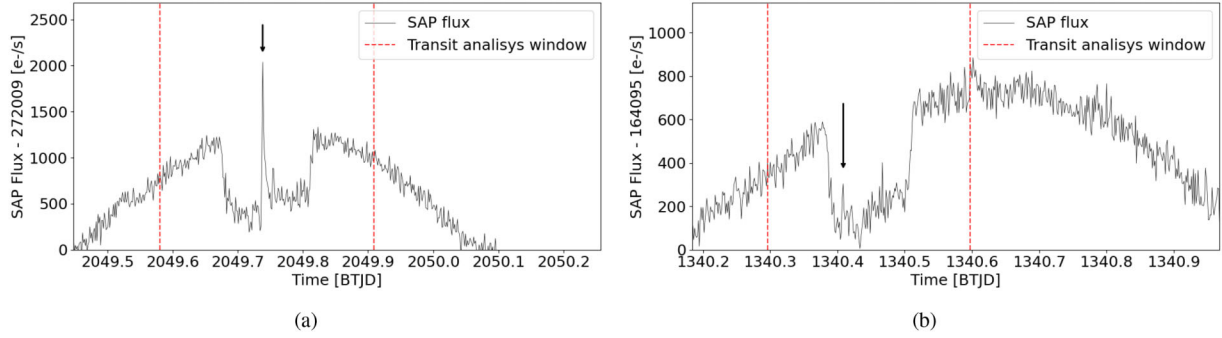


Figure 11. Examples of stellar flare contamination during planetary transits observed by *TESS*. (a) Stellar flare occurring during a transit of AU Mic b, near the centre of the transit. (b) Stellar flare occurring within the transit window of DS Tuc A b. In both panels, the red dashed vertical lines indicate the time interval adopted for the transit analysis, defined as the nominal transit duration extended by 50% before ingress and 50% after egress. The arrow marks the event identified as a stellar flare.

Table 3. Comparison between predicted and observed flare contamination in planetary transit observations. The table lists systems with at least 10 observed transits and predicted flare contamination probabilities in the 20–30% range. For each system, we report the number of analysed transits, the expected and observed numbers of flare-contaminated transits, the model-predicted contamination probability per transit, and the observed contamination fraction derived from visual inspection of the *TESS* light curves.

Planet	# transit	# f_{expected}	# f_{observed}	P_{model} (%)	P_{observed} (%)
TOI-199 b	10	2	4 ± 2.0	23	40
TOI-712 e	11	3	2 ± 1.4	23	18
HD63935 b	11	3	1 ± 1.0	25	9
TOI-1136 b	14	4	3 ± 1.7	26	21
TOI-712 d	15	4	3 ± 1.7	24	20
HD191939 d	18	5	4 ± 2.0	27	22
Kelt-11 b	22	7	9 ± 3.0	33	41
HD191939 c	24	5	6 ± 2.4	21	25
Wasp-99 b	25	6	6 ± 2.4	25	24
HAT-P-2 b	25	7	7 ± 2.6	29	28
TOI-622 b	31	7	10 ± 3.2	21	32
Kelt-2A b	35	8	12 ± 3.5	22	34
TOI-1246 b	50	11	5 ± 2.2	21	10
HD202772A b	52	12	12 ± 3.5	24	23
Total		84	84 ± 9.2		

increase in stellar flux that partially fills in the transit signal, leading to an underestimation of the transit depth and, consequently, of the inferred planet-to-star radius ratio.

However, the data set also includes events that are more representative of the average flare contamination expected for the Ariel sample. An example is shown in Fig. 11(b) for DS Tuc A b, where a flare occurs within the transit analysis window but does not dominate the light curve as strongly as in AU Mic. This case illustrates a more typical scenario, in which flares contribute additional noise or mild distortions to the transit profile rather than completely altering its shape.

These examples highlight the diversity of flare–transit configurations present in the sample, ranging from extreme cases such as AU Mic to more moderate events. While strong contamination remains rare, lower level flare activity within the transit window is not uncommon and should be accounted for in the analysis and interpretation of high-precision transit observations.

Beyond these illustrative cases, we also carried out a further check aimed at validating the flare contamination probabilities derived in this work against the available *TESS* observations. In particular, we selected a subset of systems characterized by a sufficiently large number of observed transits (at least 10) and intermediate predicted flare contamination probabilities in the range 20–30%, about 10% of the ARIEL sample. For these targets, we directly inspected all available transit windows in the *TESS* light curves, searching for the presence of flare-like events occurring during the transit intervals as defined in our analysis.

This procedure allowed us to obtain an empirical estimate of the fraction of transits affected by flaring activity, which we then compared with the probabilities predicted by our model. The results of this comparison are summarized in Table 3, where we report both the expected number of flare-contaminated transits predicted by the model and the number of transits in which a flare was visually identified in the *TESS* light curves. Overall, we find that the predicted and observed values are broadly consistent within statistical uncertainties. Importantly, when considering the full sample, the total number of observed flare-contaminated transits is the same to

the total number expected from the model, with both yielding 84 events. This agreement provides an important consistency check for our methodology, supporting the reliability of the flare contamination estimate beyond the extreme cases discussed previously, such as AU Mic. In particular, it indicates that the statistical approach used to infer flare occurrence during transit windows is able to reproduce, within the expected statistical scatter, the overall flare contamination rate observed in real *TESS* transit data, even if deviations are naturally expected for individual planetary systems.

This effect is particularly critical because it directly biases one of the fundamental planetary parameters. In addition, stellar flares are intrinsically chromatic, with a spectral energy distribution that is significantly bluer than the quiescent stellar photosphere. As a result, the apparent transit depth becomes wavelength-dependent, potentially mimicking or masking atmospheric signatures in transmission spectroscopy. In multiband or spectroscopic observations, such chromatic contamination can introduce spurious slopes or features in the transmission spectrum, which may be incorrectly interpreted as evidence for atmospheric scattering, molecular absorption, or clouds.

5 CONCLUSIONS

In this work, we analysed a sample of 290 Ariel target stars using *TESS* light curves to characterize stellar flare activity and its potential impact on transit observations. Building upon the methodology developed by S. Colombo et al. (2022) and updated by G. Galletta et al. (2025), we identified individual flares, modelled their cumulative energy distributions with two-segment power laws, and derived high-energy slopes as indicators of stellar activity levels. Overall, the analysis led to the detection of 15 857 flare events across 1638 analysed *TESS* sectors, with the number of detected flares per sector ranging from 2 to 86 events. The distribution of high-energy slopes across the sample shows a clear peak around -1.8 , with a mean of -1.87 and a standard deviation of 1.55 . Slopes ≤ -1 indicate that the flare energy distribution is dominated by low-energy events, implying that the overall flare activity is governed by a large number of frequent, low-energy flares rather than rare, high-energy ones. This behaviour is typically associated with stars exhibiting relatively high quiescent brightness levels, where the magnetic energy release occurs more efficiently through numerous small-scale reconnection events, as explained in G. Galletta et al. (2025).

By combining slope and energy break information with transit durations, we computed the probability of observing at least one flare, above an energy threshold, during a single transit for each planet. This analysis provides a robust and quantitative framework for assessing the likelihood of flare contamination, which is critical for planning Ariel observations and ensuring reliable atmospheric characterization.

Near 3% of the sample exhibits enhanced flare activity ($GF_{0.1} > 1$). AU Mic and HD 28109, shows a high probability of flare contamination during transit observations. In the case of AU Mic, this is driven by its intrinsically high flare rate. In contrast, for HD 28109 the elevated probability is primarily due to the long duration of its transits (exceeding 10 h), which increases the likelihood of overlapping flare events despite its lower intrinsic activity level. A fraction of about 10% of the sample has a contamination probability in 20%–30% range. A practical implication of this analysis concerns the observing strategy required to mitigate flare contamination. If p denotes the probability that a transit is affected by at least one flare, the probability of obtaining a flare-free (‘clean’) transit is given by $(1 - p)$. The number of observed transits required to secure a desired number N_{clean} of uncontaminated events therefore scales as:

$$N_{\text{obs}} \simeq \frac{N_{\text{clean}}}{1 - p}. \quad (7)$$

This relation implies that the observational cost increases rapidly as the flare contamination probability increases. For example, if $p = 0.25$, about 25% of transits are affected by flares, and the required number of observations increases by a factor of ~ 1.33 . For $p = 0.5$, the required observing time doubles, while for $p = 0.75$ it increases by a factor of ~ 4 , making such targets significantly more expensive in terms of observing time.

Therefore, the flare contamination probability derived in this work provides a quantitative criterion for target prioritization. Systems with high p may require either a significant increase in allocated observing time or dedicated analysis strategies to model and correct for flare contamination.

Overall, our study demonstrates that detailed flare characterization enables reliable estimation of flare occurrence within transits, allowing for improved target selection, observation planning, and interpretation of exoplanet atmospheric signals, particularly for magnetically active low-mass stars.

ACKNOWLEDGEMENTS

GM acknowledges support from the European Union – Next Generation EU through the grant no. 2022J7ZFRA – Exo-planetary Cloudy Atmospheres and Stellar High energy (Exo-CASH) funded by MUR–PRIN 2022 and the ASI-INAF agreement 2021-5-HH.2-2024.

This research made use of Lightcurve, a Python package for *Kepler* and *TESS* data analysis (Lightcurve Collaboration 2018).

This publication makes use of VOSA, developed under the Spanish Virtual Observatory (<https://svo.cab.inta-csic.es>) project funded by MCIN/AEI/10.13039/501100011033/ through grant PID2020-112949GB-I00. VOSA has been partially updated by using funding from the European Union’s Horizon 2020 Research and Innovation Programme, under Grant Agreement n°776403 (EXOPLANETS-A).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY

The data underlying this study are based on publicly available light curves from *TESS*, which can be accessed through the MAST archive (<https://mast.stsci.edu>).

The bolometric luminosities, two-segment FFD fit parameters, GF.01 activity indices, and the derived flare contamination probabilities for all Ariel targets analysed in this work are available in machine-readable format at <https://zenodo.org/records/18773384>.

All additional data products generated during this study, including intermediate flare lists and sector-level measurements, are available from the corresponding author upon reasonable request.

REFERENCES

- Aschwanden M. J., Güdel M., 2021, *ApJ*, 910, 41
- Ballerini P., Micela G., Lanza A., Pagano I., 2012, *A&A*, 539, A140
- Buccino A. P., Lemarchand G. A., Mauas P. J., 2007, *Icarus*, 192, 582
- Capistrant B. K., Dittmann J., 2026, *ApJ*, 998, 173
- Carmichael H., 1964, in Hess W. N., ed., Proc. AAS-NASA Symp.: The Physics of Solar Flares. NASA (Special Publ.), Greenbelt, p. 451
- Cauley P. W., Shkolnik E. L., Llama J., Lanza A. F., 2019, *Nat. Astron.*, 3, 1128
- Colombo S., Petralia A., Micela G., 2022, *A&A*, 661, A148
- Davenport J. R., 2016, *ApJ*, 829, 23
- Davenport J. R. et al., 2014, *ApJ*, 797, 122
- Donati J.-F. et al., 2008, *MNRAS*, 385, 1179
- Edwards B., Tinetti G., 2022, *AJ*, 164, 15
- Edwards B., Mugnai L., Tinetti G., Pascale E., Sarkar S., 2019, *AJ*, 157, 242
- Feinstein A. D. et al., 2022, *AJ*, 164, 110
- Foreman-Mackey D., Agol E., Ambikasaran S., Angus R., 2017, *AJ*, 154, 220
- Galletta G., Colombo S., Prisinzano L., Micela G., 2025, *A&A*, 698, A180
- Günther M. N. et al., 2020, *AJ*, 159, 60
- Hawley S. L., Davenport J. R., Kowalski A. F., Wisniewski J. P., Hebb L., Deitrick R., Hilton E. J., 2014, *ApJ*, 797, 121
- Howard W. S., 2022, *MNRAS*, 512, L60
- Howard W. S., Corbett H., Law N. M., Ratzloff J. K., Glazier A., Fors O., Ser D. d., Haislip J., 2019, *ApJ*, 881, 9
- Howard W. S. et al., 2023, *ApJ*, 959, 64
- Ilin E., Poppenhäger K., 2022, *MNRAS*, 513, 4579
- Ilin E., Schmidt S. J., Poppenhäger K., Davenport J. R., Kristiansen M. H., Omohundro M., 2021, *A&A*, 645, A42
- Jeffers S. V., Kiefer R., Metcalfe T. S., 2023, *Space Sci. Rev.*, 219, 54
- Johnstone C. P., 2016, *Proc. IAU Symp. S328: Living Around Active Stars*, 12. Cambridge Univ. Press, Cambridge, p. 168
- Kowalski A. F., 2024, *Living Rev. Sol. Phys.*, 21, 1
- Lanza A., 2009, *A&A*, 505, 339
- Lightcurve Collaboration, 2018, Astrophysics Source Code Library, record ascl:1812.013
- Martoli E., Hébrard G., Correia A., Laskar J., Des Etangs A. L., 2021, *A&A*, 649, A177
- Mugnai L. V., Pascale E., Edwards B., Papageorgiou A., Sarkar S., 2020, *Exp. Astron.*, 50, 303
- Oshagh M., Santos N. C., Boisse I., Boue G., Montalto M., Dumusque X., Haghighipour N., 2013, *A&A*, 556, A19
- Rackham B. et al., 2017, *ApJ*, 834, 151
- Rackham B. V., Apai D., Giampapa M. S., 2018, *ApJ*, 853, 122
- Radick R. R., Lockwood G., Baliunas S. L., 1990, *Science*, 247, 39
- Rajpurohit A., Kumar V., Srivastava M., Labadie L., Rajpurohit K., Fernández-Trincado J., 2025, *A&A*, 704, A154
- Ricker G. R. et al., 2016, in MacEwen H. A. et al., eds, Proc. SPIE Conf. Ser. Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 767
- Rodrigo C., Bayo Arán A., Solano E., Cortés-Contreras M., 2020, in XIV. 0 Scientific Meeting (Virtual) of the Spanish Astronomical Society. p. 181. Available at: <https://www.sea-astronomia.es/reunion-cientifica-2020>
- Segura A., Walkowicz L. M., Meadows V., Kasting J., Hawley S., 2010, *Astrobiology*, 10, 751
- Shibayama T. et al., 2013, *ApJS*, 209, 5
- Tinetti G. et al., 2016, in MacEwen H. A. et al., eds, Proc. SPIE Conf. Ser. Vol. 9904, Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 658
- Tinetti G. et al., 2018, *Exp. Astron.*, 46, 135
- Tinetti G., Eccleston P., Lueftinger T., Salvignol J.-C., Fahmy S., Alves de Oliveira C., 2022, in *Ariel: Enabling Planetary Science Across Light-Years, Europlanet Science Congress 2022, Granada, Spain*. p. EPSC2022–1114,
- Tristan I. I. et al., 2023, *ApJ*, 951, 33

SUPPLEMENTARY MATERIAL

Supplementary data are available at [RASTI](#) online.

suppl_data

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.