universe

MDPI

Article

Early Recognition of Cometary Activity in Sky Survey Data

Gyula M. Szabó [1,2,*], Orsolya Patonai [1] and Andrii Maliuk [1]

[1] ELTE Eötvös Loránd University, Gothard Astrophysical Observatory, Szent Imre h. u. 112., H-9700 Szombathely, Hungary; orsolya.patonai@gmail.com (O.P.); maliuka@staff.elte.hu (A.M.)
[2] HUN-REN, Research Centre for Astronomy and Earth Sciences, Konkoly Observatory, MTA Centre of Excellence, Konkoly Thege Miklós út 15-17, H-1121 Budapest, Hungary
* Correspondence: szgy@gorhard.hu

Abstract

The recognition of cometary activity is a crucial step in the analysis of small bodies discovered by modern sky surveys. Detecting activity at the earliest stages of a newly discovered object can guide subsequent observations and help distinguish active comets from apparently inactive nuclei. Despite its practical importance, objective statistical criteria for detecting weak cometary activity in trailed survey images remain valuable. We present a bootstrap-calibrated matched-template framework for the statistical detection of cometary comae. The method employs a library of synthetic coma models based on elongated power-law surface-brightness profiles, convolved with a point-spread function (PSF) and a model of the object's apparent motion. The observed image is compared with these templates using a matched-template fitting procedure. We evaluate the detection performance through idealized simulations in which synthetic trailed-nucleus images are combined with spatially uncorrelated Gaussian noise. A parametric bootstrap under the null hypothesis of an inactive nucleus provides an empirical distribution of the detection statistic and a threshold corresponding to a user-selected false-alarm probability under the adopted model and noise assumptions. The simulations quantify the sensitivity to coma amplitude and morphology, while illustrating the limitations of parameter recovery in the presence of morphological degeneracies. We also apply the method to a publicly available Rubin Observatory image of the interstellar comet 3I/ATLAS. This case study demonstrates the application of the method to an observed image, but does not constitute a validation across realistic survey backgrounds or noise conditions. The framework provides a quantitative criterion for assessing evidence of cometary activity and, in non-detections, can support the derivation of upper limits under the adopted assumptions. Its application to individual survey observations requires appropriate characterization of the relevant image properties and calibration of the null model.

Keywords: comets; detection; techniques: imaging

Check for updates

Academic Editor: Yanxia Zhang

Received: 13 July 2026
Revised: 14 October 2026
Accepted: 22 October 2026
Published: 30 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Comets are among the most important messengers from the early epochs of our Solar System, preserving pristine material that has undergone minimal alteration since its formation [1]. In recent years, they have also been recognized as key links between distant planetary systems, as two of the three interstellar objects discovered so far—2I/Borisov and 3I/ATLAS—were of cometary nature [2,3].

Detecting activity in a comet as early as possible allows astronomers to characterize the onset of the activity process, and therefore to infer the physical state of the nucleus before large

amounts of dust and gas have been released [4]. Recognizing the earliest signs of activity makes it possible to probe the mechanisms that trigger dust and gas emission—such as the sublimation of super-volatile ices, phase transitions in amorphous water ice, or structural evolution of the surface crust—rather than observing only the later, fully developed phase of activity.

Long-term observations are particularly valuable, as they enable the reconstruction of the full temporal evolution of activity. A well-known example is 67P/Churyumov–Gerasimenko, where activity was detected at heliocentric distances of about 4.3 AU inbound [5]. More generally, cometary outbursts and extended activity have been observed at distances of 5–20 AU [6], and in extreme cases even beyond 20 AU, such as for comet Hale–Bopp [7]. These observations demonstrate that activity may begin at large heliocentric distances, where illumination is usually low and detection becomes challenging.

The onset of activity depends strongly on the temperature conditions in the surface or near-subsurface layers, which in turn are governed by heliocentric distance, thermal inertia, and the volatility of the dominant species. Different volatile components (e.g., CO, CO_2, and H_2O) sublimate or undergo phase transitions at different temperatures, so the distance at which a comet becomes active can vary widely depending on composition. Observations of distant activity are therefore essential for understanding the physical and chemical evolution of cometary nuclei and for constraining models of Solar System formation.

Large sky surveys such as ATLAS [8], Pan-STARRS [9], ZTF [10], NEOWISE [11], and the LSST [12] routinely discover small bodies while they are still in the early phases of activity. These discoveries underline the importance of reliably distinguishing between inactive asteroidal objects and weakly active comets, often on the basis of very subtle morphological differences in survey images.

The detection and characterization of faint cometary activity has long relied on photometric and morphological analysis of the coma. Surface brightness profiles have been widely used to identify the presence of extended emission and to separate nucleus and coma contributions [13,14]. Detailed analyses of dust distributions and trails have shown that radial and azimuthal brightness profiles encode information about emission mechanisms and particle dynamics [15,16]. Polarimetric and multi-filter observations have further demonstrated that subtle coma structures can be detected even at large heliocentric distances [17,18].

Photometric proxies have also played a central role in quantifying activity levels. The Afρ parameter introduced by A'Hearn et al. [19] provides a widely used measure of dust production that is largely independent of aperture size, and has been applied extensively to both short-period and distant comets [7,20]. Long-term photometric monitoring campaigns have demonstrated that systematic survey data can reveal the evolution of activity over time and across large samples of objects [21–23]. Reviews of transitional populations, such as Centaurs evolving into active comets, further emphasize the importance of sensitive photometric diagnostics for identifying the onset of activity [24].

Despite these advances, the detection of faint cometary activity remains challenging in the observing regime characteristic of modern wide-field surveys. Newly discovered small bodies are often recorded close to the limiting magnitude of the survey, where signal-to-noise ratios are low and morphological diagnostics become increasingly uncertain. At the same time, their apparent motion during the exposure frequently produces trailed point-spread functions rather than compact stellar images.

The traditional approach to identifying weak cometary activity is based on comparing the 1-D radial surface brightness profile of the moving object with that of the stellar point-spread function (PSF). This approach is effective when the coma produces a measurable broadening of the profile. However, for trailed objects the trailing itself redistributes the flux over a larger area, while the additional contribution of a weak coma becomes apparent primarily in the outermost regions of the profile, where the signal is already comparable to

the background noise (Figure 1). As a result, the distinction between an inactive nucleus and a marginally active comet becomes increasingly uncertain.

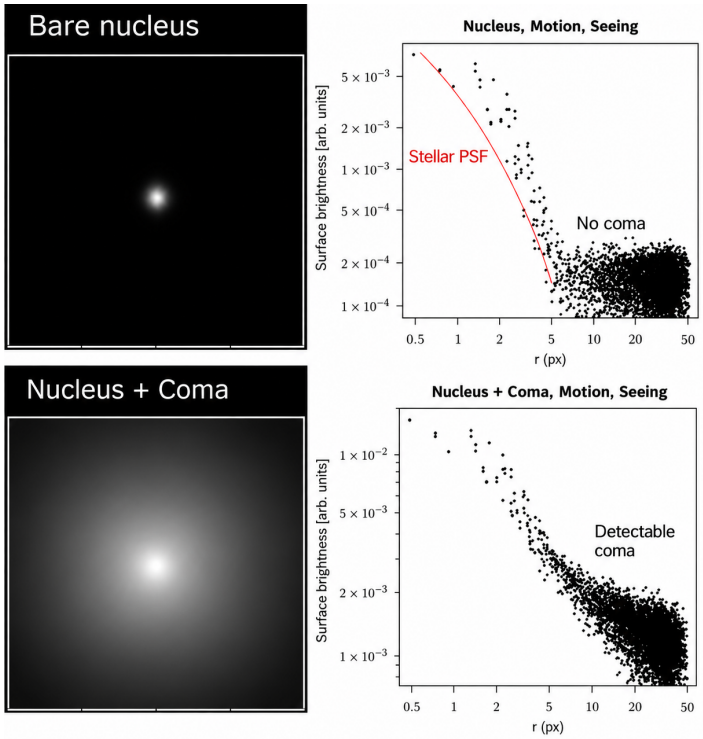


Figure 1. Simulated inactive (**top**) and weakly active (**bottom**) cometary nuclei. The stellar PSF is represented by a Moffat profile with a full width at half maximum (FWHM) of 2 pixels. The intrinsic coma has a radial power-law slope of ($\beta = -1$), is isotropic, and has a central peak intensity equal to that of the stellar PSF. Left: synthetic images; right: radial intensity distributions after motion and seeing. The red curve represents the stellar PSF. The figure illustrates a fundamental limitation of one-dimensional radial-profile methods: due to trailing motion, a faint coma becomes detectable only at large distances, where its surface brightness is already low. Intensity is shown in arbitrary units.

These difficulties are particularly important in large survey datasets, where thousands of moving objects may be detected during a single night and automated pipelines must identify promising comet candidates without human intervention. This motivates the development of a detection framework that exploits the full two-dimensional information content of the image rather than a one-dimensional radial profile, while providing an objective statistical decision criterion through bootstrap calibration.

Developing robust and quantitatively justified detection criteria is therefore essential. In particular, weak cometary activity must be distinguished from inactive nuclei in trailed, low-signal-to-noise survey images under realistic observing conditions. This requires an objective statistical framework that accounts simultaneously for the effects of the point-spread function, trailing, background, and noise.

2. Simulations

In this work, we introduce a forward-modeling framework for the statistical detection of faint cometary activity. Synthetic comet images matched to the observing conditions are used to calibrate a bootstrap-based hypothesis test, yielding empirical detection thresholds for cometary activity in survey images. The complete workflow of the proposed method is summarized in Figure 2.

To quantify the detection limits of the algorithm under controlled and statistically reproducible conditions, we constructed a forward-model simulation framework. The aim

of the simulations is not to reproduce the full physical complexity of cometary comae, but to generate a parametrized family of brightness distributions that capture the dominant observable features relevant for automated detection: centrally concentrated emission, controlled anisotropy, point-spread-function (PSF) smearing with motion trailing, and realistic stochastic background noise (Figure 3).

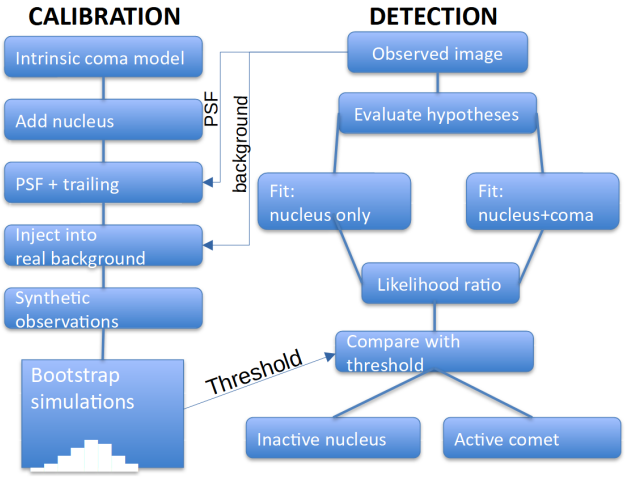


Figure 2. Workflow of the parametric-bootstrap-calibrated matched-template framework. The null distribution is obtained from synthetic images containing the trailed nucleus model and Gaussian noise. The detection threshold is determined from this distribution for a user-selected false-alarm probability and then applied to the observed image.

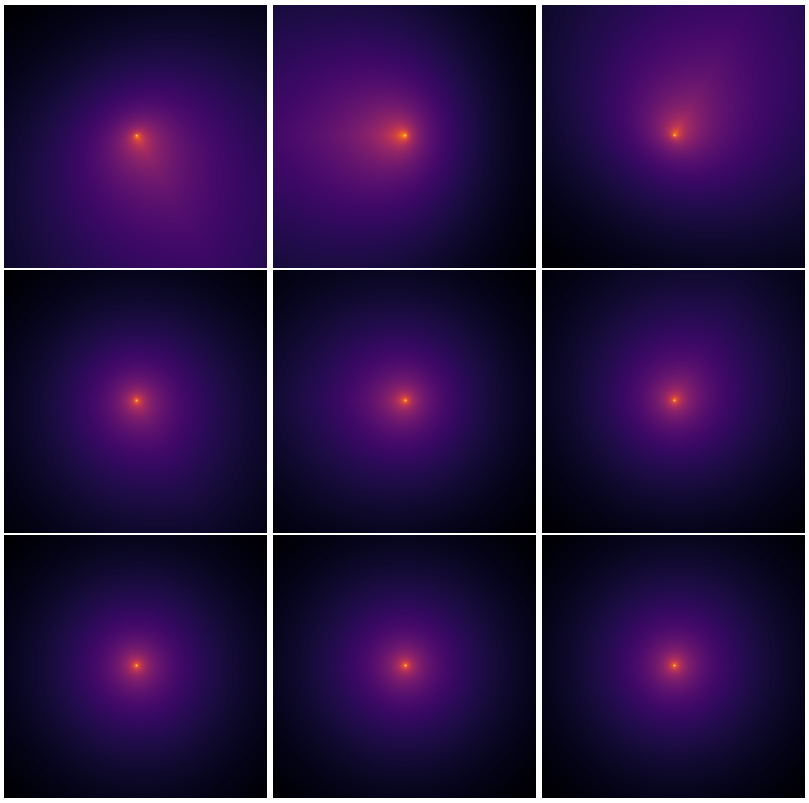


Figure 3. Examples of intrinsic coma models illustrating the effect of the angular modulation parameter (N) and the orientation angle (ϕ_0). Rows correspond to increasing values of (N = 2, 5, 12), while columns show different orientations ($\phi_0 = 60°, 180°, 300°$). All models have a radial power-law slope of ($\beta = -1.0$) and are normalized to unit peak intensity. No PSF convolution or motion trailing is applied.

2.1. Intrinsic Coma Model

The intrinsic coma surface brightness is modeled as a generalized power-law profile with optional angular modulation. Let

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2} \tag{1}$$

denote the projected cometocentric distance, and

$$\theta = \text{atan2}(y - y_0, x - x_0) \tag{2}$$

the polar angle. To introduce a preferred direction, the angular coordinate is rotated by an orientation angle ϕ_0,

$$\theta' = (\theta + \phi_0) \bmod 2\pi - \pi. \tag{3}$$

The surface brightness distribution within the coma is defined as an angular generalization of the classical steady-state model. In the standard description of isotropic, collisionless outflow, mass conservation implies a spatial density profile $n(r) \propto r^{-2}$, which leads to a projected surface brightness $I(\rho) \propto \rho^{-1}$ [19]. The present model replaces the fixed -1 slope with a generalized power-law index β, thereby allowing departures from ideal steady-state behavior.

In addition, we introduce a separable angular modulation term to account for anisotropic outflow, such as jets or preferential emission directions, inspired by the coma model in [18]. The adopted form

$$I(r, \theta') \propto (r + r_s)^{\beta} A(\theta')$$

may therefore be interpreted as a minimal extension of the classical steady-state coma model into a two-dimensional, anisotropic framework.

The adopted parametric coma model was defined as

$$I_{\text{coma}}(x, y) = C_{\text{coma}} (r + r_s)^{\beta} \left[1 + \left(\frac{\theta'}{N} \right)^2 \right], \tag{4}$$

where β is the radial power-law slope, N controls the angular scale of the modulation and inversely determines its strength (smaller N produces stronger anisotropy), ϕ_0 sets the orientation of the modulation, and r_s is a small smoothing scale preventing divergence at $r = 0$. The parameter C_{coma} scales the peak amplitude of the intrinsic coma profile relative to the unit peak intensity of the untrailed nucleus. Before applying motion blur, the intrinsic coma profile is normalized to unit peak intensity. The stellar PSF is also normalized to unit peak intensity, and the trailed PSF is not renormalized after convolution with the motion kernel. This convention keeps the amplitudes referenced to the untrailed peak intensity and avoids artificially increasing the peak intensity of a moving source after its flux has been spread by motion.

2.2. PSF Convolution and Trailing

Atmospheric seeing is modeled using a circular Moffat profile,

$$P_0(r) = \left[1 + \left(\frac{r}{\alpha} \right)^2 \right]^{-\beta_{\text{PSF}}}, \tag{5}$$

which is a parametric model of the stellar PSF at the position of the comet. Values of α and β_{PSF} can be found by any of the widespread PSF estimation processes. In our simulations, the PSF was represented by a circular Moffat profile with FWHM = 6 pixels (corresponding to

1.2″ at the Rubin Observatory LSST camera pixel scale of 0.2″ pixel^{-1}) and Moffat exponent $\beta_{\mathrm{Moffat}} = 2.5$. The scale parameter is calculated as

$$\alpha = \frac{\sigma}{\sqrt{2^{1/\beta_{\mathrm{Moffat}}} - 1}}, \qquad \sigma = \frac{\mathrm{FWHM}}{2.3548}.$$

The kernel is normalized to unit peak intensity.

During the exposure, the comet moves l pixels' length toward the ϑ position angle, due to its apparent motion. This results in a trail of the PSF in comet imaging (but not the stellar components) and can be described by convolving with a linear trailing kernel $T(l, \vartheta)$, producing an effective trailed PSF

$$P_{\mathrm{trail}} = P_0 * T(l, \vartheta), \tag{6}$$

where P_0 has unit peak intensity before convolution. The trailed PSF is not renormalized after convolution, so its peak intensity decreases as the source flux is spread along the trail.

2.3. The Nucleus

In the present formulation, the untrailed nuclear PSF has unit peak intensity, and its amplitude is fixed to unity before trailing.

The noiseless observed model image is therefore

$$I_{\mathrm{comet}} = C_{\mathrm{nuc}} P_{\mathrm{trail}} + C_{\mathrm{coma}} \left((r + r_s)^{\beta} \left[1 + \left(\frac{\theta'}{N} \right)^2 \right] \right) * P_{\mathrm{trail}} \tag{7}$$

where $C_{\mathrm{nuc}} = 1$ and all convolutions are performed numerically. (It should be recalled that the coma template is normalized to unit peak intensity before trailing, and the motion convolution is not followed by any peak or total-flux renormalization.)

2.4. Synthetic Observations

The coma parameters are sampled over the ranges

$$\beta \in [-1.8, -1.0], \qquad N \in \{2, 3, 5, 8, 12\}, \qquad \phi_0 = 0^\circ, 30^\circ, \ldots, 330^\circ,$$

covering weak to moderately active morphologies and a uniform distribution of position angles. This leads to a library of coma templates normalized to unit peak intensity before trailing, which we use during the model fitting. Sample images of the comet models in the library are presented in Figure 2.

All synthetic images were generated on a 401 × 401-pixel grid. The simulations are defined entirely in image coordinates: distances and the apparent-motion trail are expressed in pixels, and no angular pixel scale or exposure time is assumed. The adopted trail has a length of 6 pixels and is oriented eastward/westward in the image plane.

To build up synthetic observations, we add the nucleus model to the central PSF. We fix $C_{\mathrm{nuc}} = 1$ and vary only C_{coma}. Spatially uncorrelated Gaussian noise is added to each pixel,

$$D_{ij} = I_{ij} + \mathcal{N}(0, \sigma^2). \tag{8}$$

Throughout the present analysis we adopt fixed noise levels $\sigma = 0.05$ and $\sigma = 0.2$ (in normalized intensity units). This places the simulations in the regime relevant for marginal detections in wide-field survey data.

3. Detection Statistic and Evaluation Procedure

Here we propose a sensitive method to recognize the presence of an active coma and discern it from a trailed image of a bare nucleus. We adapt the general recipe of bootstrap-based statistical hypothesis testing to this specific problem.

Each simulated or observed image is evaluated against the full template library described in the previous section. The detection procedure assumes that the observed image D can be approximated as the superposition of a trailed nuclear component and a coma contribution represented by one of the template models. For each coma template T_k, the observed image is modeled as the linear combination

$$D(x, y) \approx A_k I_{\text{nuc}}(x, y) + F_k T_k(x, y), \tag{9}$$

where I_{nuc} is the normalized trailed nucleus model, T_k is the k-th coma template (after a convolution with a trailing PSF), and A_k and F_k are the fitted amplitudes of the nuclear and coma components, respectively. For a fixed template T_k, the amplitudes are determined through linear least-squares fitting within the circular fitting aperture. The value of the fitted coma amplitude is adopted as the detection score belonging to each tested coma component,

$$S_k = \max(F_k, 0), \tag{10}$$

where only templates with positive fitted coma amplitudes ($F_k > 0$) are considered as possible indicators of a non-zero coma. A negative fitted amplitude is interpreted as an indication of the absence of a coma contribution: under the null hypothesis, random fluctuations may yield a negative fitted coefficient, consistent with a stochastic realization of the zero-coma case. The final statistic assigned to the image is the amplitude of the coma component that returned with the largest amplitude parameter

$$S_{\max} = \max_k S_k. \tag{11}$$

The template producing the maximum statistic is retained as the best-matching coma morphology for the image. The purpose of this procedure is not to model the physical morphology of the coma, but to identify a heuristic detection threshold that is sensitive for the test of a presence of anything besides the trailed nucleus. Accordingly, neither the selected template nor its fitted amplitude will be interpreted physically. Instead, the result is treated as a mathematical indicator: among the coma templates in the model grid, it identifies the brightness distribution that yields the largest fitted amplitude. The method thus tests whether a template from the prescribed grid provides a sufficiently strong match to the observed image, rather than attempting to infer the physical properties inside the coma.

3.1. Empirical Threshold Calibration

The detection threshold was calibrated using a parametric bootstrap under the null hypothesis, generating images containing only the trailed nucleus model and Gaussian noise. The threshold was determined from the empirical null distribution as

$$S_{\text{thr}} = \text{Quantile}(\text{S}_{\text{max}}, \text{Confidence}),$$

incorporating template correlations, template multiplicity, and the associated look-elsewhere effect.

3.2. Detection Statistics

The detection performance of the matched-template procedure was evaluated using Monte Carlo simulations of synthetic cometary images generated over a range of coma amplitudes and morphological parameters.

For each simulated image, the maximum matched-filter response $S_{\max}$ was computed over the complete coma-template library. The synthetic cometary images were classified as detected when

$$S_{\max,\ \text{image}} > S_{\text{thr}}.$$

Figure 4 shows the resulting detection probability as a function of injected coma amplitude C. The transition between non-detection and nearly complete detection follows a characteristic sigmoidal form. A logistic model was fitted to the simulated detection fractions in order to characterize the effective survey sensitivity.

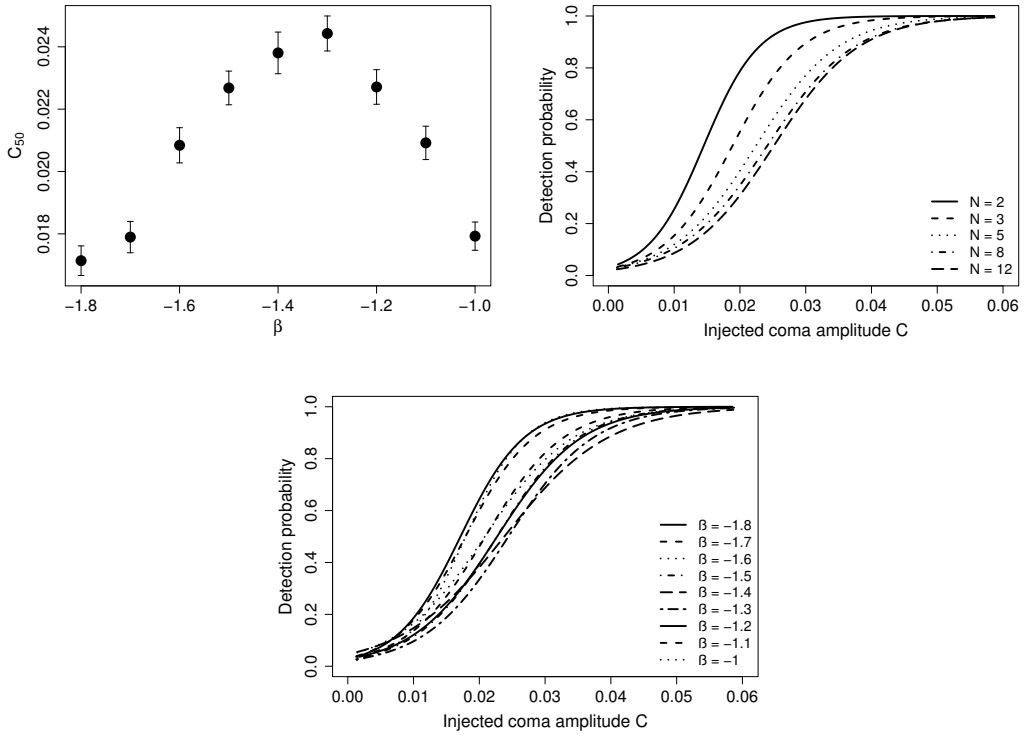


Figure 4. Detection efficiency analysis of the matched-template framework as a function of coma morphology parameters. Left: dependence of the 50% detection threshold (C_{50}) on the coma radial slope parameter β. Middle: empirical power curves obtained for different coma scale parameters N. Right: empirical power curves for different radial slope parameters β. The σ parameter of the Gaussian noise was 0.05 (i.e., 5% of the peak of the untrailed PSF component of the comet), and the applied confidence level was 98%. A total of 1000 simulations were calculated independently in each bin and all parameters were fitted at once.

The fitted 50% detection threshold is

$$C_{50} \approx 0.021,$$

indicating that intrinsic coma templates with peak amplitudes of approximately 2.1% of the untrailed nuclear peak intensity are detected with a probability of 50% under the adopted noise conditions.

The dependence of the detection efficiency on coma morphology was investigated by evaluating the detection curves separately for different template parameters. Figure 4 shows the detection probability as a function of C for different values of the radial profile slope parameter β and the scale parameter N, respectively.

More centrally concentrated coma models exhibit systematically lower detection limits, while extended diffuse comae require substantially larger amplitudes for reliable detection. In particular, the inferred C_{50} values increase monotonically with increasing N over the explored parameter grid. This trend reflects the dependence of the detection sensitivity on the adopted angular modulation and should not be interpreted as a direct dependence on the radial coma scale length.

These simulations demonstrate that the detection sensitivity of the matched-template framework depends not only on the coma amplitude itself, but also on the detailed mor-

phology of the coma model. Nevertheless, the transition between non-detection and high-completeness regimes remains well behaved across the explored parameter space.

3.3. Parameter Recovery

More compact coma morphologies are detected at systematically lower coma amplitudes, while extended and diffuse comae require stronger activity levels for reliable detection. This behavior is consistent with the expectations because compact comae preserve a higher central surface brightness after convolution with the trailed PSF, thereby producing a stronger matched-template response than diffuse emission with the same total coma amplitude.

In order to investigate the parameter-recovery properties of the matched-template framework, additional bootstrap simulations were performed for fixed input morphologies. For each realization, synthetic images were generated using the same underlying coma template while independent Gaussian noise realizations were added. The complete template library was then re-evaluated and likelihood shells were constructed from the relative residual statistic

$$\Delta\chi^2 = \chi^2 - \chi^2_{\min}.$$

4. Results

Figures 5 and 6 show corner plots of the parameter combinations providing acceptable fits for the high- and low-S/N simulations, respectively, based on the relative chi-square statistic.

In the high-S/N case, the acceptable parameter combinations are concentrated around the best-fitting solution, indicating that the adopted data and model constrain the parameters more tightly. The slope parameter β is recovered almost uniquely, while C, N, and ϕ exhibit only modest uncertainties.

Pairwise parameter distributions reveal strong correlations between C and β, with a weaker but still significant correlation involving the scale parameter N. Physically, more extended coma models require systematically larger amplitudes in order to reproduce the same central surface-brightness structure.

The dominant C–β correlation reflects the well-known difficulty of separating an unresolved nucleus from the steep inner coma after convolution with the instrumental PSF, which makes the reliable detection of the contribution of a faint compact coma more difficult in our methodology. A brighter unresolved nucleus can mimic a more centrally concentrated coma, whereas a steeper coma profile can compensate for a fainter nucleus. Consequently, a unique decomposition of the central brightness distribution is generally impossible unless the observations provide sufficient spatial resolution to resolve the innermost coma. This is a known limitation of comet studies, and HST-quality instruments are required for a reliable photometry of the nucleus of an active comet [13].

We also observe a moderate degeneracy between C or β and N, showing that a bias in N has some ability to compensate a less determined inner coma–nucleus separation. In contrast, the orientation angle ϕ exhibits almost no correlation with the remaining parameters. This indicates that the directional asymmetry of the coma is determined largely independently of the radial brightness structure and therefore represents the most robust morphological parameter against the nucleus–coma degeneracy.

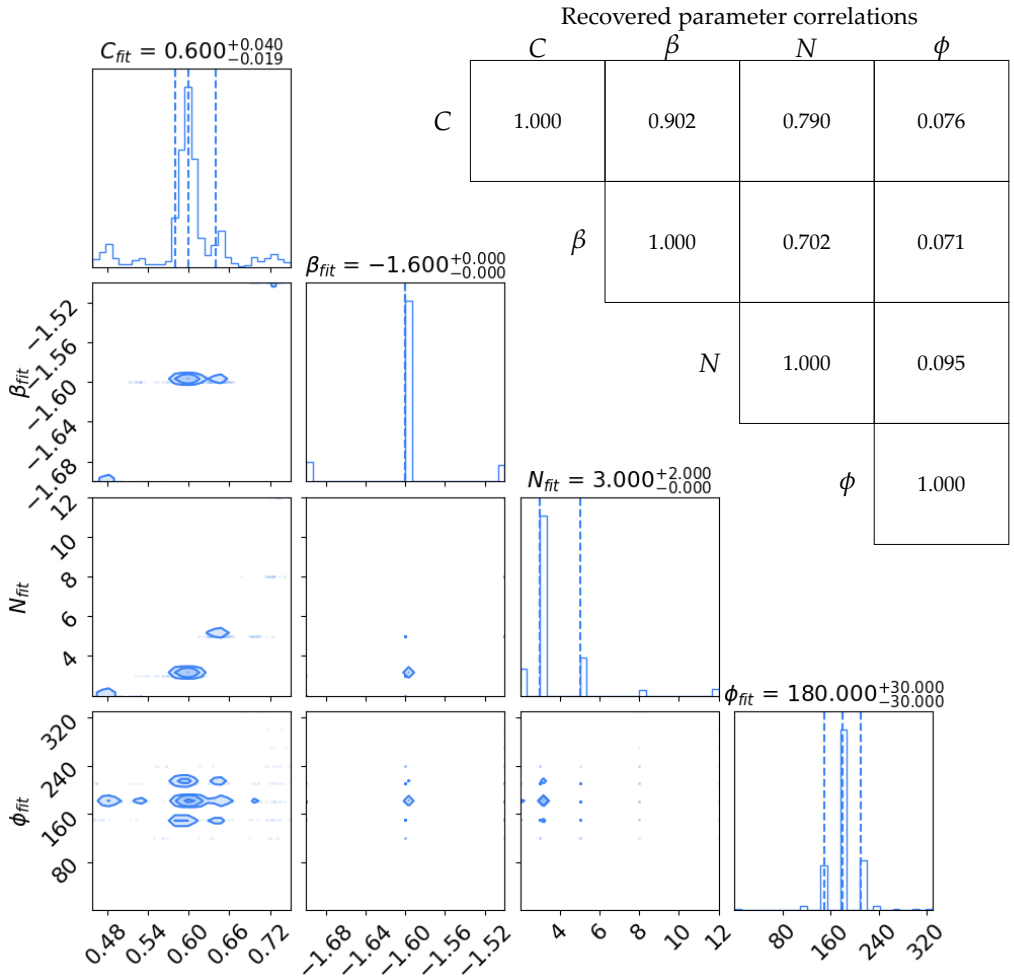


Figure 5. Corner plot of a comet with $C = 0.6$, $\beta = -1.6$, $N = 3$, $\phi = 180$, fitted with all templates that provided acceptable fits for the noise level at 98% confidence, based on the relative chi-square statistic $\Delta\chi^2 = \chi^2 - \chi^2_{\min}$. The diagonal panels show the one-dimensional parameter-space distributions, while the off-diagonal panels illustrate the pairwise relationships between parameters. Dashed lines indicate the best-fitting values and the intervals used to summarize the parameter ranges. The upper-right triangle reports the Pearson correlation coefficients between the recovered parameters, with unity on the diagonal. Owing to the discrete parameter grid adopted in the inversion, several parameter distributions appear as isolated clusters rather than continuous contours.

These parameter correlations reflect limitation of morphological studies of the inner coma in the presence of significant image blurring. However, these parameter correlations mainly limit the accuracy of morphological parameter recovery but have only a minor effect on the primary objective of the present method, namely the detection of ongoing cometary activity. As demonstrated in Figure 4, the activity decision remains robust even in the presence of substantial parameter covariance.

As the signal-to-noise ratio decreases, the ranges of acceptable parameter values broaden considerably. The uncertainties of C, N, and especially the orientation angle ϕ increase substantially, reflecting the reduced information content of the observations. This behavior is expected because the orientation is constrained primarily by low-contrast asymmetric structures in the outer coma, which are the first features to disappear as the image noise increases. Under such conditions, the bootstrap-derived detection threshold still provides a reliable criterion for distinguishing active and inactive objects. However, the available information is insufficient for a meaningful recovery of detailed coma morphology.

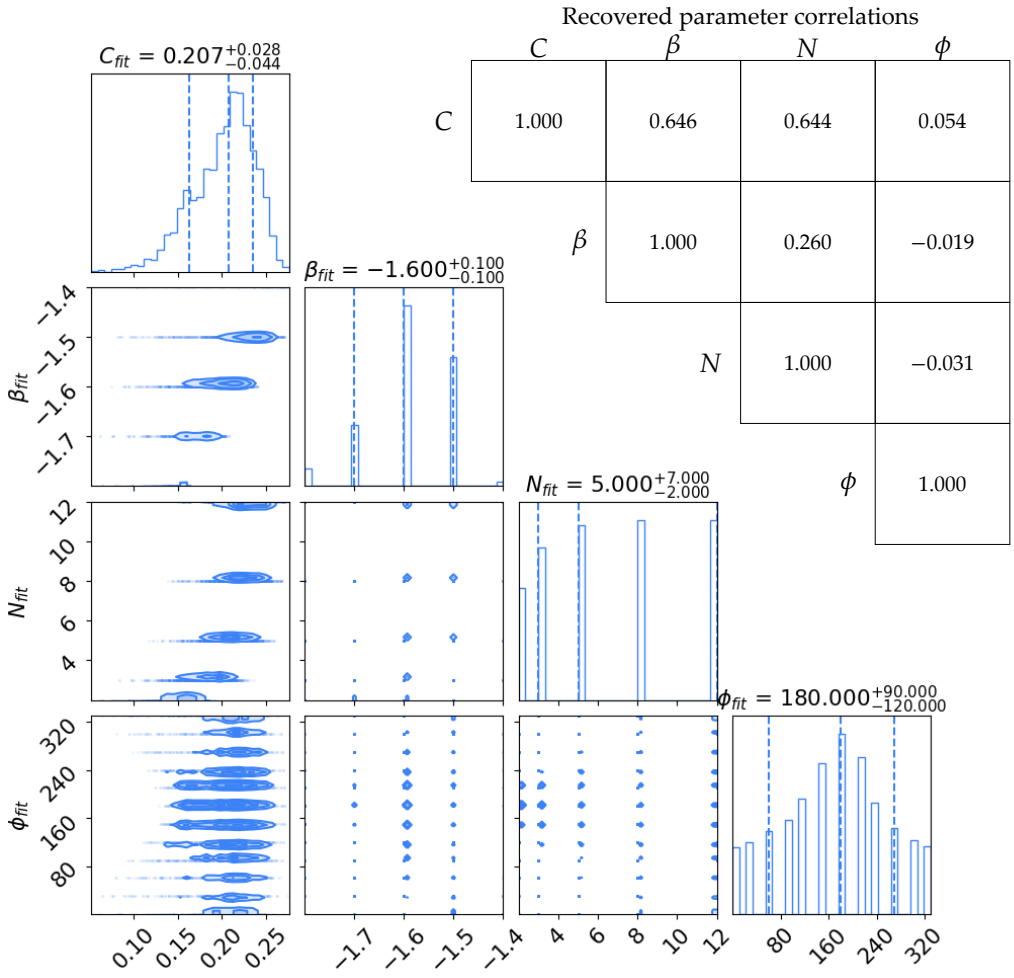


Figure 6. The same as Figure 5 but with $\sigma = 0.20$.

The simulations presented above also indicate that the bootstrap analysis should not be regarded as a generic calibration procedure. Instead, it should be performed individually for each observational dataset, using image properties that reproduce the actual observing conditions as closely as possible.

The calibration demonstrated in this work is a parametric bootstrap based on the adopted trailed-nucleus model and Gaussian noise. Its resulting false-alarm probability is conditional on these assumptions and should not be interpreted as a guarantee of the same performance under arbitrary real-survey backgrounds or noise properties. Adapting the null model to the observational characteristics of a particular survey would be a separate implementation step beyond the simulations presented here.

The use of randomly selected image cutouts also introduces an observationally realistic selection function. Bright stars and strongly contaminated regions are naturally avoided during the original comet detection process, and therefore restricting the bootstrap analysis to similarly clean image regions reproduces the effective detection conditions more faithfully than attempting to construct an unbiased artificial background.

The central conclusion of this work is that coma detection cannot be separated from the characterization of the complete observational process. The statistical significance of a detected coma is meaningful only if the bootstrap calibration reproduces the same PSF, image motion, background structure, and noise properties as the original observation. Under these conditions, the proposed matched-template bootstrap framework provides an

objective decision criterion for both positive coma detections and quantitative upper limits in non-detections.

5. Applications to LSST Data

We tested the proposed method on a single publicly available image of the interstellar comet 3I/ATLAS, obtained with LSST on 21 June 2025 [1]. The image was selected from the public gallery presented by Chandler et al. [3]. The comet lies in a locally clean sky region, without blending with field stars. We selected an isolated nearby star as the comparison source, located 24 pixels south and 3 pixels east of the comet's center.

The stellar PSF was modeled using a Moffat profile with $\beta = 2.5$, fitted to the comparison star. The image background and noise level were estimated from the full image. These quantities were then used consistently throughout the tests described above.

We first applied the radial-profile analysis in two configurations: comparison with the static stellar PSF and with a trailed stellar PSF. The static PSF is not a physically adequate representation of the observation, since 3I/ATLAS moved by approximately 2.5 pixels during the 30-s exposure [3]. Nevertheless, this simple comparison provides a rapid diagnostic of cometary activity. The observed and stellar radial profiles were statistically inconsistent, with $p = 0.0006$.

For a more appropriate comparison, we convolved the stellar PSF with a motion kernel corresponding to the comet's apparent displacement during the exposure. The observed radial profile remained inconsistent with the trailed stellar PSF, yielding $p = 0.0028$. The difference between these two results illustrates the information loss introduced by trailing, while showing that the radial-profile method can still identify evidence of activity. The profile comparisons are shown in Figure 7; the error bars represent the pixel-to-pixel intensity fluctuations.

We then applied the proposed method to reconstruct the coma parametrically. We searched the model grid and identified acceptable fits among models with $N = 2$, coma parameter $\beta = 1.3$, and elongation position angles between 180° and 230°. Across the ten best-fitting models, the fitted nucleus component contributed approximately 12–20% of the coma component, depending on the model configuration. The central intensity was dominated by the coma, while the extended wings of the radial intensity distribution were also reproduced by the coma component. Figure 8 compares the observed profile with the best-fitting parametric model, which reproduces the mean coma intensity closely.

Finally, we assessed the significance of the fitted coma component using a Monte Carlo null model based on the trailed stellar PSF, with no intrinsic coma. We generated 1000 synthetic images by adding independent, pixel-dependent noise with standard deviation

$$\sigma_{\mathrm{pixel}} = \sqrt{I_{\mathrm{model}} + B},$$

where I_{model} is the background-subtracted model intensity and B is the estimated background level. For each realization, we refitted the nucleus and coma coefficients using the same two-component fitting procedure as for the observed image. None of the 1000 realizations yielded a coma coefficient equal to or greater than the observed best-fit value. Applying the standard finite-simulation correction gives a Monte Carlo probability of $p_{\mathrm{MC}} = 0.0010$. Thus, under the adopted noise model, the observed coma component is not reproduced by the trailed stellar PSF alone.

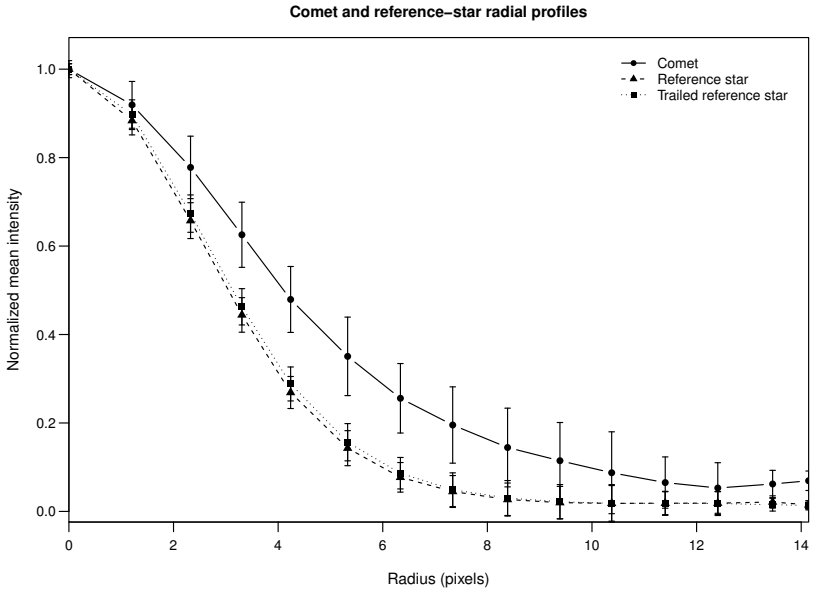


Figure 7. Comparison of the measured radial intensity profile of 3I/ATLAS observed by LSST on 21 June 2025 (solid line) with a stellar profile (dashed line) and a trailed stellar profile (dotted line).

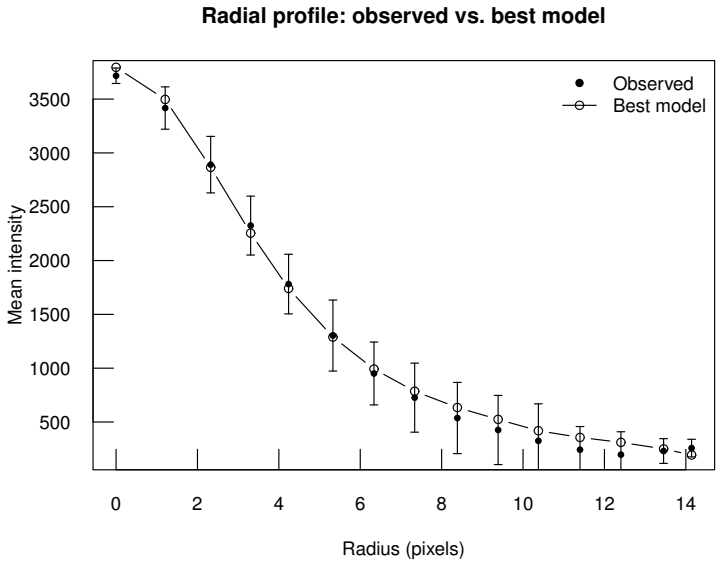


Figure 8. Comparison of the measured radial intensity profile of 3I/ATLAS observed by LSST on 21 June 2025 with the best-fitting synthetic coma model.

The largest coma coefficient obtained from the 1000 simulated coma-free images was only 0.018 times the value measured for 3I/ATLAS, while the standard deviation of the relative amplitudes was approximately 0.005. No realization required a coma component as strong as that inferred from the observed image. Assuming a Gaussian distribution for the relative amplitudes, the resulting Monte Carlo significance corresponds to $p_{\mathrm{MC}} = 0.0010$. The p values obtained using the different methods described above are compared in Table 1.

This probability is higher than that inferred from the chi-square comparison with the untrailed stellar PSF, but substantially lower than the probability obtained from the radial-profile analysis using the trailed PSF. This result demonstrates that two-dimensional image reconstruction can recover part of the information lost through trailing. By reconstructing the nucleus and coma components separately, using physically motivated parametric morphology and the known trailing pattern, the method exploits spatial information that is

not fully retained in the one-dimensional radial profiles. Nevertheless, trailing still affects the achievable sensitivity: it distributes the source flux over a larger number of pixels, thereby reducing the signal-to-noise ratio over the area required for the analysis.

Table 1. Comparison of *p*-values resulted by the compared methods.

Comparison/Method	Dim.	Physical Interpretation	*p*-Value
Comet vs. stellar PSF	1D	Non-plausible	0.0006
Comet vs. trailed stellar PSF	1D	Plausible as a trailed PSF	0.0028
Comet vs. trailed model	2D	Parametric reconstruction	0.0010

The numerical results obtained for 3I/ATLAS are not interpreted further in terms of the comet's morphology. Such an interpretation is beyond the scope of this study and, more importantly, is not an objective of the proposed method. The sole purpose of the method is to quantify and improve the sensitivity of the decision statistic used to establish whether an early comet identification exhibits detectable evidence of activity. The entire methodology is designed around this specific objective, rather than around the subsequent physical interpretation of the coma. Once activity has been securely detected, the interested reader is directed towards dedicated physical models of the coma and its evolution (e.g., Fulle et al. [25]; Marschall et al. [26]), which are better suited to a rigorous morphological and physical analysis.

Future Work

The grid-based approach adopted here is suitable for testing the proposed method and demonstrating its applicability. However, our experiments indicate that the parameter sampling used in the present grid—for example, position-angle steps of 60°—does not fully exploit the achievable sensitivity. A substantially denser grid would require a much larger template library, with a corresponding increase in computational and storage costs. We therefore suggest generating model templates on the fly or, alternatively, restricting the search to the vicinity of the best-fitting solutions and refining the parameters through numerical optimization of the parametric models. Such an approach could improve the sensitivity without requiring an exhaustive library of precomputed templates.

Even with these improvements, the method should not be regarded as a morphological analysis tool. Its purpose is specifically to determine whether an early comet identification contains statistically significant evidence of activity, including activity that is only marginally detectable. Detailed reconstruction and physical interpretation of the coma morphology remain separate tasks, to be addressed with dedicated physical models once the presence of activity has been established.

A further priority is the calibration of the method for specific survey observing conditions. The normalized noise levels adopted in the present simulations are intended for a controlled proof-of-concept analysis and do not correspond to predefined limiting magnitudes or signal-to-noise ratios in a particular survey. In the case of Rubin/LSST, a systematic assessment should investigate the dependence of detection efficiency on source brightness, PSF and seeing, apparent motion and trail length, background properties, coma-to-nucleus contrast, coma morphology, and observing band. Such a study would require dedicated simulations and tests on representative survey images, possibly followed by iterative validation within the LSST Solar System Science Collaboration. This calibration is beyond the scope of the present work and is an important objective for future development.

6. Conclusions and Summary

We presented a bootstrap-calibrated matched-template framework that transforms coma detection from a subjective visual assessment into an objective statistical decision with a well-defined confidence level. This question becomes particularly important when the activity is close to the detection limit, as is typical during the onset of cometary activity. In the era of modern wide-field sky surveys, such as the Zwicky Transient Facility (ZTF, [27]) and the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, [12]), enormous numbers of comet detections are becoming available, making objective and reproducible activity diagnostics increasingly important.

To address this problem, we developed a matched-template bootstrap procedure and evaluated it using idealized simulations with a trailed-nucleus model and spatially uncorrelated Gaussian noise. The application to a publicly available Rubin Observatory image of 3I/ATLAS provides an observational case study, while the adaptation of the calibration to realistic survey backgrounds and noise properties remains a separate implementation step. The workflow for coma detection under the adopted model and noise assumptions is summarized below.

The method demonstrated in this work uses a parametric-bootstrap procedure to calibrate the detection statistic under the null hypothesis. The workflow is as follows:

1. Construct the trailed nucleus model using the adopted PSF and apparent motion.
2. Generate null realizations containing the trailed nucleus model and Gaussian noise.
3. Evaluate each null realization using the same matched-template procedure as for the observed image.
4. Determine the detection threshold from the empirical null distribution for a user-selected false-alarm probability.
5. Apply the calibrated threshold to the observed image. A positive detection indicates that the observed score is inconsistent with the coma-free null distribution at the selected false-alarm level, under the adopted model and noise assumptions.

The method is specified through the mathematical formalism, model definitions, and statistical procedures presented in this paper, which are intended to enable independent implementation and reproduction of the analysis. As a future development, we plan to implement the method internally on the Rubin Science Platform. This implementation is not part of the present work and is not currently available as a public software package.

Author Contributions: Conceptualization, G.M.S.; methodology, G.M.S.; software, G.M.S. and O.P.; validation, A.M.; writing—original draft preparation, G.M.S.; writing—review and editing, G.M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Research, Development and Innovation Office (NKFIH, Hungary) through grants SNN-147362 and ADVANCED-153410, by the European Space Agency through PRODEX Experiment Agreements No. 4000137122 and No. 4000149203, and by the City of Szombathely. This material is based upon work supported in part by the National Science Foundation through Cooperative Agreements AST-1258333 and AST-2241526 and Cooperative Support Agreements AST-1202910 and 2211468 managed by the Association of Universities for Research in Astronomy (AURA), and the Department of Energy under Contract No. DE-AC02-76SF00515 with the SLAC National Accelerator Laboratory managed by Stanford University. Additional Rubin Observatory funding comes from private donations, grants to universities, and in-kind support from LSST-DA Institutional Members. Gy. M. Sz. gratefully acknowledges the support of the Fulbright Scholarship, the University of Washington for hosting a research stay, and Željko Ivezić for the inspiring idea that initiated this investigation.

Data Availability Statement: The mathematical formalism and statistical procedures required to implement the method are described in this article. No public software package or precomputed template library is provided with the present work. The test image used in this study is part of the publicly available Rubin image stamps for 3I/ATLAS dataset (https://doi.org/10.5281/zenodo.19244429), under image identifier `2025062100431-038-r.`

Conflicts of Interest: The authors declare no conflicts of interest.

Notes

[1] The image was selected from the public gallery presented by Chandler et al. [3] under image identifier: `2025062100431-038-r`

References

1. Bockelée-Morvan, D.; Crovisier, J.; Mumma, M.J.; Weaver, H.A. The composition of cometary volatiles. *Comets II* **2004**, 391–423. [CrossRef]
2. Guzik, P.; Drahus, M.; Rusek, K.; Waniak, W.; Cannizzaro, G.; Pastor-Marazuela, I. Initial characterization of interstellar comet 2I/Borisov. *Nat. Astron.* **2020**, *4*, 53–57. [CrossRef]
3. Chandler, C.O.; Bernardinelli, P.H.; Jurić, M.; Singh, D.; Hsieh, H.; Sullivan, I.; Jones, R.L.; Kurlander, J.; Vavilov, D.; Eggl, S.; et al. NSF-DOE Vera C. Rubin Observatory Observations of Interstellar Comet 3I/ATLAS (C/2025 N1). *Astrophys. J. Let.* **2026**, *1001*, L35. [CrossRef]
4. Meech, K.J.; Pittichová, J.; Bar-Nun, A.; Notesco, G.; Laufer, D.; Hainaut, O.R.; Lowry, S.C.; Yeomans, D.K.; Pitts, M. Activity of comets at large heliocentric distances pre-perihelion. *Icarus* **2009**, *201*, 719–739. [CrossRef]
5. Moreno, F.; Snodgrass, C.; Hainaut, O.; Tubiana, C.; Sierks, H.; Barbieri, C.; Lamy, P.L.; Rodrigo, R.; Koschny, D.; Rickman, H.; et al. The dust environment of comet 67P/Churyumov-Gerasimenko from Rosetta OSIRIS and VLT observations in the 4.5 to 2.9 AU heliocentric distance range inbound. *Astron. Astrophys.* **2016**, *587*, A155. [CrossRef]
6. Gronkowski, P.; Wesołowski, M. A Review of Cometary Outbursts at Large Heliocentric Distances. *Earth Moon Planets* **2016**, *119*, 23–33. [CrossRef]
7. Szabó, G.M.; Kiss, L.L.; Sárneczky, K. Cometary Activity at 25.7 AU: Hale-Bopp 11 Years after Perihelion. *Astrophys. J.* **2008**, 677, L121–L124. [CrossRef] [PubMed]
8. Tonry, J.L.; Denneau, L.; Heinze, A.N.; Stalder, B.; Smith, K.W.; Smartt, S.J.; Stubbs, C.W.; Weiland, H.J.; Rest, A. ATLAS: A High-cadence All-sky Survey System. *Publ. Astron. Soc. Pac.* **2018**, *130*, 064505. [CrossRef]
9. Chambers, K.C.; Magnier, E.A.; Metcalfe, N.; Flewelling, H.A.; Huber, M.E.; Waters, C.Z.; Denneau, L.; Draper, P.W.; Farrow, D.; Finkbeiner, D.P.; et al. The Pan-STARRS1 Surveys. *arXiv* **2016**, arXiv:1612.05560.
10. Bellm, E.C.; Kulkarni, S.R.; Barlow, T.; Feindt, U.; Graham, M.J.; Goobar, A.; Kupfer, T.; Ngeow, C.; Nugent, P.; Ofek, E.; et al. The Zwicky Transient Facility: Surveys and Scheduler. *Publ. Astron. Soc. Pac.* **2019**, *131*, 068003. [CrossRef]
11. Mainzer, A.; Bauer, J.; Cutri, R.M.; Grav, T.; Masiero, J.; Beck, R.; Clarkson, P.; Conrow, T.; Dailey, J.; Eisenhardt, P.; et al. Initial Performance of the NEOWISE Reactivation Mission. *Astrophys. J.* **2014**, *792*, 30.
12. Ivezić, Ž.; Kahn, S.M.; Tyson, J.A.; Abel, B.; Acosta, E.; Allsman, R.; Alonso, D.; AlSayyad, Y.; Anderson, S.F.; Andrew, J.; et al. LSST: From Science Drivers to Reference Design and Anticipated Data Products. *Astrophys. J.* **2019**, *873*, 111. [CrossRef]
13. Lamy, P.L.; Toth, I.; Weaver, H.A.; A'Hearn, M.F.; Jorda, L. Properties of the nuclei and comae of 13 ecliptic comets from Hubble Space Telescope snapshot observations. *Astron. Astrophys.* **2009**, *508*, 1045–1056. [CrossRef]
14. Jewitt, D.C.; Meech, K.J. Surface Brightness Profiles of 10 Comets. *Astrophys. J.* **1987**, 317, 992. [CrossRef] [PubMed]
15. Moreno, F. Monte Carlo Models of Comet Dust Tails Observed from the Ground. *Universe* **2022**, *8*, 366. [CrossRef]
16. Agarwal, J.; Müller, M.; Reach, W.T.; Sykes, M.V.; Boehnhardt, H.; Grün, E. The dust trail of Comet 67P/Churyumov-Gerasimenko between 2004 and 2006. *Icarus* **2010**, *207*, 992–1012. [CrossRef]
17. Ivanova, O.V.; Dlugach, J.M.; Afanasiev, V.L.; Reshetnyk, V.M.; Korsun, P.P. CCD polarimetry of distant comets C/2010 S1 (LINEAR) and C/2010 R1 (LINEAR) at the 6-m telescope of the SAO RAS. *Planet. Space Sci.* **2015**, 118, 199–210. [CrossRef]
18. Szabó, G.M.; Kiss, L.L.; Pál, A.; Kiss, C.; Sárneczky, K.; Juhász, A.; Hogerheijde, M.R. Evidence for Fresh Frost Layer on the Bare Nucleus of Comet Hale-Bopp at 32 AU Distance. *Astrophys. J.* **2012**, 761, 8. [CrossRef]
19. A'Hearn, M.F.; Schleicher, D.G.; Millis, R.L.; Feldman, P.D.; Thompson, D.T. Comet Bowell 1980b. *Astron. J.* **1984**, *89*, 579–591. [CrossRef] [PubMed]
20. Vereš, P.; Jedicke, R.; Denneau, L.; Wainscoat, R.; Holman, M.J.; Lin, H.W. Improved Asteroid Astrometry and Photometry with Trail Fitting. *Publ. Astron. Soc. Pac.* 124, 1197–1207. [CrossRef] [PubMed]
21. Gillan, A.F.; Fitzsimmons, A.; Denneau, L.; Siverd, R.J.; Smith, K.W.; Tonry, J.L.; Young, D.R. Dust Production Rates in Jupiter-family Comets: A Two Year Study with ATLAS Photometry. *Planet. Sci. J.* **2024**, *5*, 25. [CrossRef]

22. Mazzotta Epifani, E.; Dall'Ora, M.; Perna, D.; Palumbo, P.; Colangeli, L. The cometary activity of Centaur P/2004 A1 (LONEOS). *Mon. Not. R. Astron. Soc.* **2011**, *415*, 3097–3106. [CrossRef]
23. Hsieh, H.H.; Jewitt, D.; Lacerda, P.; Lowry, S.C.; Snodgrass, C. The return of activity in main-belt comet 133P/Elst-Pizarro. *Mon. Not. R. Astron. Soc.* **2010**, *403*, 363–377. [CrossRef]
24. Peixinho, N.; Thirouin, A.; Tegler, S.C.; Di Sisto, R.P.; Delsanti, A.; Guilbert-Lepoutre, A.; Bauer, J.G. From Centaurs to Comets: 40 years. In *The Trans-Neptunian Solar System*; Elsevier: Amsterdam, The Netherlands, pp. 307–329.
25. Fulle, M.; Blum, J.; Rotundi, A. CO-driven activity constrains the origin of comets. *Astron. Astrophys.* **2020**, *636*, L3. [CrossRef]
26. Marschall, R.; Markkanen, J.; Gerig, S.B.; Pinzón-Rodríguez, O.; Thomas, N.; Wu, J.S. The Dust-to-Gas Ratio, Size Distribution, and Dust Fall-Back Fraction of Comet 67P/Churyumov-Gerasimenko: Inferences From Linking the Optical and Dynamical Properties of the Inner Comae. *Front. Phys.* **2020**, *8*, 227. [CrossRef]
27. Bellm, E.C.; Kulkarni, S.R.; Graham, M.J.; Dekany, R.; Smith, R.M.; Riddle, R.; Masci, F.J.; Helou, G.; Prince, T.A.; Adams, S.M.; et al. The Zwicky Transient Facility: System Overview, Performance, and First Results. *Publ. Astron. Soc. Pac.* **2019**, *131*, 018002. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.