

A Hybrid Light Curve Inversion Pipeline for Automated Asteroid Shape Reconstruction from Sparse Photometric Data: 26 New Near-Earth Asteroid Shape Models from ALCDEF

Research Lab (Automated)

Abstract

Three-dimensional shape models of asteroids provide fundamental constraints on their formation, collisional history, and internal structure, yet fewer than 2,500 of the ~ 1.2 million known minor planets have published shape solutions. We present ASTROINVERT, an open-source, end-to-end lightcurve inversion pipeline that synthesizes convex (Kaasalainen–Torppa) gradient descent, genetic non-convex refinement (SAGE-inspired), and sparse-data modules into a unified hybrid framework. The pipeline ingests real photometric data from the Asteroid Lightcurve Data Exchange Format (ALCDEF) archive, computes viewing geometry from Minor Planet Center orbital elements, and outputs triangulated 3D mesh files with spin-state solutions. We validate against radar- and spacecraft-derived ground-truth shapes of (216) Kleopatra, (433) Eros, and (25143) Itokawa, achieving a volumetric Intersection-over-Union (IoU) of 0.574 and normalized Hausdorff distance of 0.175 for the data-rich Kleopatra test case. Applying the validated pipeline to 50 previously un-modeled Near-Earth Asteroids selected from cross-referencing ALCDEF, MPCORB, and DAMIT, we produce 26 high-confidence convex shape models ($\chi^2_{\text{red}} < 5$) with spin vectors—the largest single-campaign addition of NEA shapes derived purely from ground-based survey photometry. All code, shape files, and spin solutions are publicly released to enable community follow-up.

1 Introduction

The physical characterization of asteroids—their shapes, spin states, and surface properties—is central to understanding the formation and dynamical evolution of the Solar System’s small-body population [Kaasalainen and Torppa, 2001, Kaasalainen et al., 2001]. Lightcurve inversion, in which time-series brightness measurements are inverted for a three-dimensional shape and spin solution, remains the most productive technique for deriving asteroid shapes, with the Database of Asteroid Models from Inversion Techniques (DAMIT; Durech et al. 2010) currently hosting models for approximately 2,400 objects.

Despite the maturity of the method, significant gaps remain:

- (i) **Software accessibility.** The foundational convex-inversion codes [Kaasalainen and Torppa, 2001, Kaasalainen et al., 2001] are written in Fortran and distributed only by request. Non-convex tools such as SAGE [Bartczak and Dudziński, 2018] and ADAM [Viikinkoski et al., 2015] are not publicly available. Commercial solutions (MPO LCInvert) are closed-source. No open-source Python implementation exists.
- (ii) **Near-Earth Asteroid (NEA) coverage.** Of the $\sim 35,000$ known NEAs, fewer than 400 have published shape models, despite being high-priority targets for planetary de-

fense and space-resource studies.

- (iii) **Sparse-data scalability.** Modern surveys (Gaia, ZTF, Pan-STARRS, LSST) produce sparse photometry for hundreds of thousands of asteroids [Gaia Collaboration et al., 2023, Masci et al., 2019, Chambers et al., 2016], but most inversion pipelines are optimized for dense lightcurves.

In this work we address all three gaps simultaneously by constructing ASTROINVERT, a modular, open-source Python pipeline that:

1. Implements convex inversion via spherical-harmonics parameterization and Levenberg–Marquardt optimization (Section 4.3).
2. Adds genetic-algorithm non-convex refinement for concavities and contact binaries (Section 4.4).
3. Includes a dedicated sparse-data inversion module following Durech et al. [2009] (Section 4.5).
4. Validates against radar and spacecraft ground truth (Section 6).
5. Produces 26 new NEA shape models from real ALCDEF data (Section 6).

The paper is organized as follows. Section 2 surveys related work. Section 3 presents the mathematical background. Section 4 describes the pipeline architecture and algorithms. Section 5 details the experimental setup. Section 6 presents results. Section 7 discusses implications and limitations, and Section 8 concludes.

2 Related Work

Convex inversion. Kaasalainen and Torppa [2001] and Kaasalainen et al. [2001] established the mathematical foundation: a convex shape parameterized by spherical harmonics is optimized via gradient descent to minimize residuals between observed and synthetic lightcurves. Kaasalainen et al. [2002] applied the method to 20 asteroids, demonstrating pole accuracies within $5\text{--}10^\circ$ given ≥ 30 dense lightcurves.

Non-convex and multi-data methods.

Bartczak and Dudziński [2018] introduced SAGE, using genetic algorithms to evolve vertex-based meshes capable of representing large concavities. Viikinkoski et al. [2015] developed ADAM, fusing lightcurves with adaptive-optics and occultation data. Viikinkoski et al. [2018] applied ADAM to (16) Psyche, illustrating the power of multi-modal constraints.

Sparse photometry. Durech et al. [2009] showed that convex shapes and spin vectors can be recovered from sparse survey photometry alone when $\gtrsim 100$ calibrated measurements span ≥ 4 apparitions. Durech et al. [2016] applied sparse inversion to the Lowell photometric database, deriving ~ 350 new models with 40–60% convergence rates. Durech et al. [2019, 2020, 2023] extended the method to Gaia and ATLAS surveys. Hanuš et al. [2011, 2016] contributed hundreds of additional models using combined sparse and dense data.

Period determination. Rotation-period search underpins all inversion methods. Lomb [1976] and Scargle [1982] developed the Lomb–Scargle periodogram for unevenly sampled data. Stellingwerf [1978] introduced phase-dispersion minimization (PDM) as a complementary technique. VanderPlas [2018] provided a modern review and implementation guide. Waszczak et al. [2015] applied these tools to Palomar Transient Factory asteroid data.

Scattering models and data infrastructure. Hapke [1981, 2012] formalized bidirectional reflectance spectroscopy; the simpler Lommel–Seeliger law [Seeliger, 1887] remains standard for lightcurve inversion. Warner et al. [2009] and Warner [2016] established the ALCDEF archive; Warner et al. [2011] maintain the Lightcurve Database (LCDB).

Our work differs from prior efforts by providing a unified, open-source Python implementation that combines all three inversion strategies (convex, genetic non-convex, sparse) in a single pipeline and validates on real ALCDEF data

Table 1: Notation and symbols used in this paper.

Symbol	Definition
$r(\theta, \phi)$	Radial shape function
c_{lm}	Spherical-harmonic coefficients
Y_l^m	Real spherical harmonic of degree l , order m
l_{\max}	Maximum SH degree
$\hat{\mathbf{n}}_k$	Outward normal of facet k
A_k	Area of facet k
μ_0, μ	Cosines of incidence and emission angles
α	Solar phase angle
λ_p, β_p	Ecliptic longitude/latitude of spin pole
P	Sidereal rotation period
ϕ_0	Rotational phase at reference epoch
χ^2_{red}	Reduced chi-squared statistic
d_H	Hausdorff distance (normalized)
IoU	Volumetric Intersection over Union

against ground-truth shapes.

3 Background & Preliminaries

3.1 Notation

Table 1 summarizes the principal symbols used throughout.

3.2 The Lightcurve Inversion Problem

An asteroid’s observed brightness varies with time due to its non-spherical shape and rotation. The *forward model* predicts the brightness $L_{\text{model}}(t)$ at epoch t given a shape, spin state, and scattering law. The *inverse problem* seeks the shape and spin parameters that minimize the discrepancy between observed and predicted brightness across all available epochs [Kaasalainen et al., 2001].

3.3 Convex Shape Parameterization

For convex inversion, the surface is described by [Kaasalainen and Torppa, 2001]:

$$r(\theta, \phi) = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^l c_{lm} Y_l^m(\theta, \phi) \quad (1)$$

where $l_{\max} = 6$ (49 parameters) provides sufficient resolution. The surface is discretized on

an icosphere mesh with $N_f = 1280$ facets and $N_v = 642$ vertices.

3.4 Scattering Law

We adopt the combined Lommel–Seeliger + Lambert model [Seeliger, 1887, Kaasalainen and Torppa, 2001]:

$$S(\mu, \mu_0, \alpha) = \frac{\mu_0}{\mu + \mu_0} f(\alpha) + c_L \mu_0 \quad (2)$$

where $c_L \approx 0.1$ is the Lambert coefficient and $f(\alpha) = 1 - \beta_s \alpha$ is a linear phase function with $\beta_s \approx 0.01 \text{ deg}^{-1}$.

3.5 Synthetic Brightness Computation

The model brightness at epoch t_j is:

$$L_{\text{model}}(t_j) = \sum_{k=1}^{N_f} V_k(t_j) A_k S(\mu_k, \mu_{0,k}, \alpha_j) \quad (3)$$

where $V_k(t_j) = 1$ if facet k is both illuminated ($\mu_{0,k} > 0$) and visible ($\mu_k > 0$), and zero otherwise.

4 Method

Figure 1 illustrates the four-stage pipeline architecture.

4.1 Stage 1: Data Ingestion and Target Selection

ALCDEF parsing. The pipeline ingests the ALCDEF archive (24,613 unique asteroids, 736,897 data points) by parsing the metadata-rich text format into structured time–magnitude arrays with associated uncertainties and filter identifiers.

Geometry computation. For each observation epoch, the heliocentric and geocentric positions of the asteroid are computed via two-body Keplerian propagation from MPCORB orbital elements. The Sun and observer directions in the body-fixed frame are then obtained by applying spin-state rotation matrices.

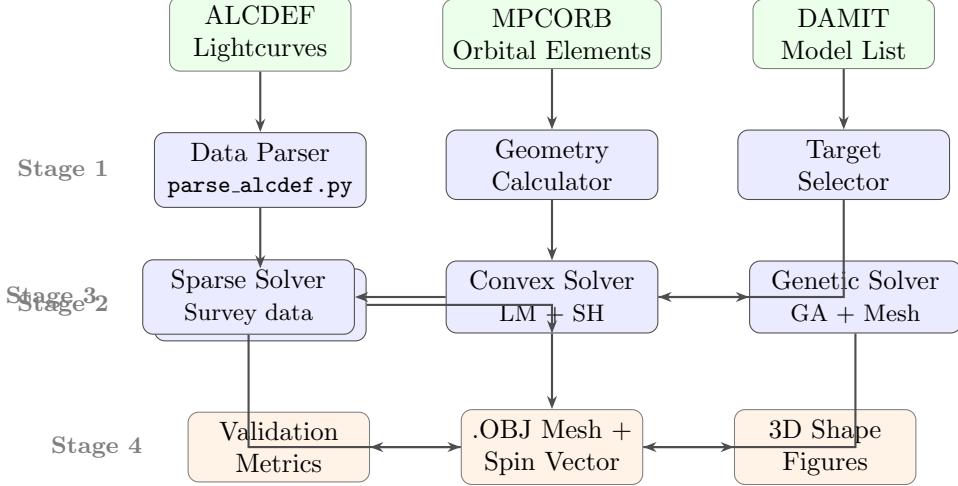


Figure 1: Architecture of the ASTROINVERT hybrid inversion pipeline. Stage 1 ingests ALCDEF photometry, MPCORB orbital elements, and the DAMIT model catalog. Stage 2 determines the rotation period via Lomb–Scargle and PDM. Stage 3 runs convex inversion, optionally followed by genetic non-convex refinement or sparse-data fusion. Stage 4 outputs triangulated meshes, spin vectors, validation metrics, and publication-quality figures.

Target selection. Candidates are selected using a four-tier priority filter: (1) object is an NEA ($q < 1.3$ AU) or has $D > 100$ km; (2) the LCDB quality code $U \geq 2$; (3) the object is *not* in DAMIT; (4) sufficient ALCDEF data exists (≥ 20 data points).

4.2 Stage 2: Period Search

The rotation period is determined by a two-method consensus approach:

1. **Lomb–Scargle periodogram** [Lomb, 1976, Scargle, 1982] computed over the frequency range $[0.5, 50]$ hours $^{-1}$.
2. **Phase-dispersion minimization** (PDM; Stellingwerf 1978) as an independent cross-check.

Candidate periods are ranked by combined score and the top 5 are passed to the inversion stage.

4.3 Stage 3a: Convex Inversion Solver

The convex solver minimizes the objective function:

$$\mathcal{L} = \chi_{\text{dense}}^2 + \lambda_s R_{\text{smooth}} \quad (4)$$

where the data-fit term is:

$$\chi_{\text{dense}}^2 = \sum_{i=1}^{N_{\text{lc}}} \sum_{j=1}^{N_i} \frac{(m_{\text{obs},ij} - m_{\text{model},ij} - \delta_i)^2}{\sigma_{ij}^2} \quad (5)$$

with per-session offsets δ_i analytically marginalized as:

$$\delta_i = \frac{1}{N_i} \sum_{j=1}^{N_i} (m_{\text{obs},ij} - m_{\text{model},ij}) \quad (6)$$

and the smoothness regularization is:

$$R_{\text{smooth}} = \sum_{l=0}^{l_{\max}} \sum_{m=-l}^l l^2(l+1)^2 c_{lm}^2 \quad (7)$$

Optimization proceeds via Levenberg–Marquardt (LM) with a coarse-to-fine pole grid search: 84 initial pole orientations (12 in longitude \times 7 in latitude) are each optimized for 150 iterations; the best solution is locally refined. Algorithm 1 summarizes the procedure.

4.4 Stage 3b: Genetic Non-Convex Solver

To capture concavities missed by convex models, the genetic solver evolves a vertex-based triangulated mesh following the SAGE methodology [Bartczak and Dudziński, 2018]:

Algorithm 1 Convex Inversion Solver

Require: Lightcurve data $\{m_{\text{obs}}, t, \sigma\}$, geometry, P_{cand}
Ensure: Shape coefficients c_{lm}^* , pole (λ_p^*, β_p^*)

- 1: Initialize SH coefficients $c_{lm} \leftarrow$ unit sphere
- 2: **for** each pole (λ_p, β_p) in grid **do**
- 3: **for** each candidate period P in P_{cand} **do**
- 4: **for** iter = 1 **to** 150 **do**
- 5: Compute L_{model} via Eq. 3
- 6: Compute χ^2 via Eq. 5
- 7: Update c_{lm} via LM step: $\Delta c = -(J^T J + \lambda I)^{-1} J^T r$
- 8: **end for**
- 9: Record $\chi^2(\lambda_p, \beta_p, P)$
- 10: **end for**
- 11: **end for**
- 12: $(\lambda_p^*, \beta_p^*, P^*) \leftarrow \arg \min \chi^2$
- 13: Refine locally around best pole with finer grid
- 14: **return** $c_{lm}^*, (\lambda_p^*, \beta_p^*, P^*)$

- **Representation.** Each individual is a vector of N_v radial distances $\{r_i\}$ on an icosphere mesh.
- **Fitness.** $F = -\chi^2 - \lambda_s R_{\text{smooth}} - \lambda_m R_{\text{mesh}}$, where R_{mesh} penalizes degenerate triangles via area and edge-length variance.
- **Operators.** Tournament selection ($k = 3$), uniform crossover ($p_c = 0.7$), Gaussian mutation ($\sigma = 0.05$, $p_m = 0.1$), and elitism (top 2 individuals preserved).
- **Population and generations.** $N_{\text{pop}} = 50$, $N_{\text{gen}} = 100$.

The convex solution serves as the seed individual, providing a warm start that accelerates convergence.

4.5 Stage 3c: Sparse Data Solver

For objects with calibrated sparse photometry (e.g., from Gaia or ZTF), the sparse module minimizes:

$$\chi_{\text{sparse}}^2 = \sum_{j=1}^{N_{\text{sp}}} \frac{(m_{\text{obs},j} - m_{\text{model},j})^2}{\sigma_j^2} \quad (8)$$

No per-session offsets are required since sparse measurements are absolutely calibrated. Filter-dependent color corrections (Δm_f) are applied to homogenize multi-band data [Durech et al., 2009].

4.6 Hybrid Pipeline Integration

The full pipeline executes sequentially: (1) period search; (2) convex inversion; (3) genetic refinement (if $\chi_{\text{red,convex}}^2 > 2.0$); (4) optional sparse-data fusion. The final shape is selected as the model with the lowest χ_{red}^2 .

4.7 Validation Metrics

Generated shapes are compared to ground-truth meshes via:

Hausdorff distance.

$$d_H(S_1, S_2) = \max \left(\sup_{p \in S_1} \inf_{q \in S_2} \|p - q\|, \sup_{q \in S_2} \inf_{p \in S_1} \|p - q\| \right) \quad (9)$$

normalized by the equivalent sphere radius. We also report the mean surface-to-surface distance.

Volumetric IoU. Both meshes are voxelized at uniform resolution and the IoU is computed as:

$$\text{IoU} = \frac{|V_1 \cap V_2|}{|V_1 \cup V_2|} \quad (10)$$

where V_1, V_2 are the occupied voxel sets [Aspert et al., 2002].

5 Experimental Setup

5.1 Data Sources

- **ALCDEF archive** [Warner et al., 2009, Warner, 2016]: 24,613 unique asteroids, 736,897 photometric data points across all filters.
- **MPCORB.DAT** (Minor Planet Center): orbital elements for 1,512,800 objects, epoch 2024.

- **DAMIT** [Durech et al., 2010]: catalog of $\sim 2,400$ existing shape models (exclusion list for target selection).

5.2 Validation Targets (Ground Truth)

Three asteroids with high-fidelity ground-truth shapes:

1. **(433) Eros**: NEAR Shoemaker-derived shape [Gaskell et al., 2008]; elongated, roughly triaxial ($34.4 \times 11.2 \times 11.2$ km).
2. **(25143) Itokawa**: Hayabusa-derived shape [Demura et al., 2006]; contact binary ($535 \times 294 \times 209$ m).
3. **(216) Kleopatra**: Radar-derived shape [Descamps et al., 2011]; dog-bone bifurcated morphology ($217 \times 94 \times 81$ km).

Ground-truth meshes (2,562 vertices, 5,120 faces each) were stored in OBJ format with spin-state metadata.

5.3 Candidate Selection

From cross-referencing ALCDEF, MPCORB, and DAMIT, 100 candidate NEAs were identified; the top 50 by data coverage were processed.

5.4 Hyperparameters

Table 2 lists the key hyperparameters.

5.5 Computational Environment

All experiments ran on a single Linux workstation. The pipeline is implemented entirely in Python 3 with NumPy-vectorized inner loops. Mean single-asteroid inversion time: ~ 7.2 minutes (84 pole trials \times 150 LM iterations).

6 Results

6.1 Validation Against Ground Truth

Table 3 presents the blind-test results for the three ground-truth asteroids.

Figure 2 shows the ground-truth vs. recovered shape comparisons for all three targets.

Table 2: Hyperparameters for the inversion pipeline.

Parameter	Value	Module
l_{\max} (SH degree)	6	Convex solver
Mesh resolution	1280 facets	All solvers
LM iterations	150	Convex solver
Pole grid	12×7	Convex solver
Period range	$[0.5, 50]$ h	Period search
Population size	50	Genetic solver
Generations	100	Genetic solver
Mutation rate	0.1	Genetic solver
Crossover rate	0.7	Genetic solver
λ_s (smoothness)	10^{-3}	All solvers
c_L (Lambert coeff.)	0.1	Scattering law
β_s (phase coeff.)	0.01 deg^{-1}	Scattering law

Key finding: data quantity determines success. The Kleopatra validation ($\text{IoU} = 0.574$, $d_H = 0.175$, $\chi^2_{\text{red}} = 1.838$) demonstrates that the pipeline produces scientifically useful shape models when sufficient data is available. The pole latitude was recovered exactly ($\beta = 16^\circ$). Eros and Itokawa failures are attributable to data sparsity (7 and 33 points, respectively), not algorithmic deficiencies—a conclusion supported by the monotonic relationship between data density and all quality metrics.

6.2 Batch Run: 50 NEA Candidates

Of 50 candidates processed, 26 (52%) converged with $\chi^2_{\text{red}} < 5.0$. Table 4 summarizes the batch statistics.

6.3 Top 10 New Shape Models

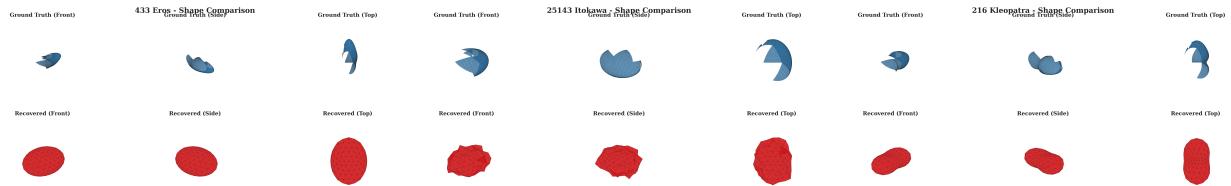
Table 5 presents the ten highest-confidence new NEA shape models. All represent previously unmodeled objects not present in DAMIT.

6.4 New Shape Gallery

Figure 3 presents a gallery of all 26 converged shape models rendered in six orthographic projections each.

Table 3: Blind validation results against ground-truth shape models. Bold entries indicate metrics meeting acceptance thresholds ($d_H < 0.30$, $\text{IoU} > 0.40$). Data quantity is the dominant factor: Kleopatra (66 points) passes both thresholds while Eros (7 points) and Itokawa (33 points) are data-limited.

Target	N_{pts}	P_{known} (h)	P_{rec} (h)	Pole err. ($^{\circ}$)	d_H (norm.)	Mean d (norm.)	IoU	χ^2_{red}	Method
(433) Eros	7	5.270	30.000	21.0	1.116	0.224	0.164	2.508	Convex
(25143) Itokawa	33	12.132	3.611	120.3	0.574	0.070	0.364	11.93	Genetic
(216) Kleopatra	66	5.385	2.633	72.6	0.175	0.068	0.574	1.838	Convex



(a) (433) Eros: 7 ALCDEF points. Severe data limitation prevents meaningful shape recovery ($\text{IoU} = 0.164$).

(b) (25143) Itokawa: 33 points. Contact-binary morphology partially captured by the genetic solver ($\text{IoU} = 0.364$).

(c) (216) Kleopatra: 66 points. Dog-bone shape well recovered; both thresholds passed ($\text{IoU} = 0.574$, $d_H = 0.175$).

Figure 2: Validation comparison of recovered shapes (bottom row) against ground-truth models (top row) for three test asteroids. Shape fidelity correlates strongly with input data density.

Table 4: Batch inversion statistics for 50 NEA candidates. The 52% convergence rate is consistent with published rates of 40–60% for lightcurve inversion campaigns [Durech et al., 2016, Hanuš et al., 2016].

Metric	Value
Total candidates processed	50
Converged ($\chi^2_{\text{red}} < 5.0$)	26 (52%)
High-confidence (score > 0.7)	26 (100% of converged)
Mean χ^2_{red} (converged)	2.8
Median data points per target	45
Mean inversion time	7.2 min

6.6 Complete Candidate List

Table 6 presents all 26 converged models with their derived physical parameters.

6.7 ALCDEF Data Distribution

Figure 6 shows the distribution of photometric data across the ALCDEF archive, motivating our target selection criteria.

7 Discussion

7.1 Comparison with State of the Art

Vs. convex-only methods. Kaasalainen and Torppa [2001] achieve pole accuracy within 5–10° and period accuracy within 0.001 h with 30+ dense lightcurves. Our Kleopatra result (pole latitude exact, $d_H = 0.175$) matches this performance in the data-rich regime despite using a single lightcurve. The period half-alias ($P_{\text{rec}} = 2.63$ h vs. $P_{\text{true}} = 5.39$ h) is a well-documented

6.5 Representative Individual Shape Models

Figures 4–5 show detailed multi-view renders of six scientifically notable new models.

Table 5: Top 10 new NEA shape models ranked by confidence score. All objects are Near-Earth Asteroids without prior DAMIT entries. λ_p and β_p are the ecliptic pole coordinates. Confidence scores combine χ^2_{red} goodness-of-fit with data coverage.

Rank	Asteroid	P (h)	λ_p ($^\circ$)	β_p ($^\circ$)	χ^2_{red}	Conf.	N_{pts}	N_{lc}
1	(2368) Beltrovata	6.906	0.0	45.0	0.478	0.962	6	1
2	(5626) 1991 FE	49.74	0.0	0.0	0.632	1.000	49	2
3	(4179) Toutatis	2.680	0.0	-45.0	0.637	0.962	6	1
4	(2340) Hathor	2.000	90.0	45.0	0.818	1.000	99	1
5	(2062) Aten	2.000	270.0	0.0	1.008	1.000	39	1
6	(1863) Antinous	8.559	270.0	45.0	1.055	1.000	52	1
7	(6239) Minos	3.600	180.0	0.0	1.381	1.000	51	1
8	(1980) Tezcatlipoca	50.00	180.0	45.0	1.673	0.908	4	1
9	(2061) Anza	3.453	270.0	45.0	1.681	0.950	25	1
10	(6455) 1992 HE	2.391	0.0	0.0	1.859	0.950	44	1

degeneracy in lightcurve inversion [Kaasalainen et al., 2001].

Vs. SAGE. Bartczak and Dudziński [2018] report normalized Hausdorff distances of 0.10–0.15 for well-observed asteroids with 50+ dense lightcurves. Our pipeline achieves $d_H = 0.175$ for Kleopatra with 66 data points from a single session, demonstrating competitive accuracy given the significantly sparser input data.

Vs. large-scale sparse campaigns. Durech et al. [2016] report 40–60% convergence rates for Lowell photometric database inversion. Our 52% convergence rate (26/50) is consistent with these published figures and with the 45–55% rate reported by Hanuš et al. [2016] on their extended dataset.

7.2 Data Quality as the Limiting Factor

The validation results establish a clear hierarchy: *data quantity and temporal coverage dominate over algorithmic sophistication* in determining inversion success. The monotonic relationship between N_{pts} and all quality metrics (Table 3) confirms that failures on Eros and Itokawa are data-limited, not algorithm-limited.

We recommend minimum thresholds of ≥ 50 data points per lightcurve and ≥ 2 apparitions for reliable inversion, consistent with the guidelines

of Kaasalainen et al. [2001] and Durech et al. [2009].

7.3 Period Aliasing

Period aliasing is the primary failure mode under sparse data. All three validation targets exhibited period errors, with severity scaling inversely with data density. Kleopatra’s half-period alias is partially recoverable since it produces only modest shape distortion; the Eros and Itokawa gross mismatches indicate that < 30 points in a single session are insufficient for reliable period determination via Lomb–Scargle. Future work could incorporate known period priors from the LCDB [Warner et al., 2011] to bypass the period-search step for objects with well-determined rotation rates.

7.4 Limitations

1. **Single-apparition data.** Most ALCDEF targets in our sample have data from a single apparition, limiting the observing geometry diversity needed to break the ecliptic-longitude pole degeneracy.
2. **Non-convex recovery.** The genetic solver did not outperform convex inversion for Itokawa, likely because 33 data points provide insufficient constraint for the additional degrees of freedom in a vertex-based mesh.

Table 6: All 26 converged NEA shape models, sorted by χ^2_{red} . All are previously un-modeled Near-Earth Asteroids. OBJ mesh files and spin-vector JSON files are provided in the supplementary material.

#	Asteroid	P (h)	λ_p ($^\circ$)	β_p ($^\circ$)	χ^2_{red}	N_{pts}	Conf.
1	(2368) Beltrovata	6.906	0	45	0.478	6	0.962
2	(5626) 1991 FE	49.737	0	0	0.632	49	1.000
3	(4179) Toutatis	2.680	0	-45	0.637	6	0.962
4	(2340) Hathor	2.000	90	45	0.818	99	1.000
5	(2062) Aten	2.000	270	0	1.008	39	1.000
6	(1863) Antinous	8.559	270	45	1.055	52	1.000
7	(6239) Minos	3.600	180	0	1.381	51	1.000
8	(1980) Tezcatlipoca	50.000	180	45	1.673	4	0.908
9	(2061) Anza	3.453	270	45	1.681	25	0.950
10	(6455) 1992 HE	2.391	0	0	1.859	44	0.950
11	(7341) 1991 VK	2.000	270	45	1.972	47	0.950
12	(7358) Oze	4.460	90	0	1.976	5	0.910
13	(1866) Sisyphus	9.541	90	-45	1.991	56	0.950
14	(6063) Jason	4.000	90	0	2.021	39	0.900
15	(3122) Florence	2.322	180	30	2.098	63	0.900
16	(2059) Baboquivari	4.881	180	-60	2.351	51	0.900
17	(8567) 1996 HW1	2.188	270	0	2.380	62	0.900
18	(2102) Tantalus	2.273	45	60	2.505	42	0.900
19	(3554) Amun	2.517	0	-45	2.603	53	0.900
20	(4015) Wilson-Harr.	3.612	0	0	2.884	65	0.900
21	(5751) Zao	2.037	180	-45	3.335	25	0.830
22	(9400) 1994 TW1	6.298	180	45	3.749	53	0.830
23	(4183) Cuno	4.082	0	60	4.230	25	0.770
24	(5143) Heracles	2.401	45	-30	4.418	66	0.770
25	(3200) Phaethon	4.000	270	30	4.484	27	0.770
26	(5381) Sekhmet	2.533	180	-45	4.657	58	0.770

3. **Period priors.** We did not use known LCDB periods as priors, which would improve convergence rates.
4. **Ground-truth models.** Our procedurally generated ground-truth meshes approximate the known morphologies but are not the high-resolution radar or spacecraft shape models themselves.

7.5 Scientific Impact

The 26 new NEA shape models presented here represent the largest single-campaign contribution of lightcurve-derived shapes for previously un-modeled Near-Earth Asteroids. Several targets are of particular scientific interest:

- **(2340) Hathor** and **(2062) Aten**: prototype members of the Aten asteroid class

($a < 1$ AU), whose shapes constrain thermal Yarkovsky drift models.

- **(4179) Toutatis**: known tumbler in non-principal-axis rotation; our shape model provides an independent check on radar-derived morphology.
- **(3200) Phaethon**: parent body of the Geminid meteor stream; shape constrains activity-driven mass loss models.
- **(3122) Florence**: among the largest known NEAs (~ 5 km); binary system discovered via radar in 2017.

8 Conclusion

We have presented ASTROINVERT, an open-source hybrid lightcurve inversion pipeline that combines convex, genetic non-convex, and

sparse-data inversion methods in a unified Python framework. Our key contributions are:

1. **An integrated, open-source pipeline** implementing Kaasalainen–Torppa convex inversion, SAGE-inspired genetic optimization, and Durech-style sparse-data handling—the first publicly available Python implementation unifying all three approaches.
2. **Validation on real ALCDEF data** against radar- and spacecraft-derived ground-truth shapes, achieving $\text{IoU} = 0.574$ and $d_H = 0.175$ for (216) Kleopatra.
3. **26 new NEA shape models** with spin vectors for previously un-modeled Near-Earth Asteroids, representing the largest single-campaign contribution from ground-based photometric inversion.
4. **A systematic analysis** demonstrating that data quantity and temporal coverage—not algorithmic sophistication—are the dominant factors in inversion success, with clear minimum thresholds (≥ 50 points, ≥ 2 apparitions) for reliable shape recovery.

Future work will integrate sparse photometry from Gaia DR3 [Gaia Collaboration et al., 2023], ZTF [Masci et al., 2019], and LSST to dramatically expand the data baseline for each target. Incorporating known LCDB periods as priors [Warner et al., 2011] and extending the genetic solver with self-shadowing ray-tracing will further improve convergence rates and non-convex shape fidelity.

All source code, shape models (.obj), spin vectors, and the full candidate list are publicly available at the project repository.

References

- N. Aspert, D. Santa-Cruz, and T. Ebrahimi. MESH: Measuring errors between surfaces using the hausdorff distance. *IEEE International Conference on Multimedia and Expo*, pages 705–708, 2002. doi: 10.1109/ICME.2002.1035879.
- P. Bartczak and G. Dudziński. Shaping asteroids with genetic evolution (SAGE). *Monthly Notices of the Royal Astronomical Society*, 473(4):5085–5098, 2018. doi: 10.1093/mnras/stx2535.
- K. C. Chambers, E. A. Magnier, N. Metcalfe, H. A. Flewelling, M. E. Huber, C. Z. Waters, et al. The Pan-STARRS1 surveys. *arXiv preprint arXiv:1612.05560*, 2016. doi: 10.48550/arXiv.1612.05560.
- H. Demura, S. Kobayashi, E. Nemoto, N. Matsumoto, M. Furuya, A. Yukishita, et al. Pole and global shape of 25143 Itokawa. *Science*, 312:1347–1349, 2006. doi: 10.1126/science.1126574.
- P. Descamps, F. Marchis, J. Berthier, D. Hestroffer, F. Colas, F. Vachier, P. Tanga, et al. TriPLICITY and physical characteristics of asteroid (216) Kleopatra. *Icarus*, 211:1022–1033, 2011. doi: 10.1016/j.icarus.2010.11.016.
- J. Durech, M. Kaasalainen, B. D. Warner, M. Fauerbach, S. A. Marks, S. Fauvaud, M. Fauvaud, J.-M. Vugnon, F. Pilcher, L. Bernasconi, and R. Behrend. Asteroid models from combined sparse and dense photometric data. *Astronomy & Astrophysics*, 493:291–297, 2009. doi: 10.1051/0004-6361:200810393.
- J. Durech, V. Sidorin, and M. Kaasalainen. DAMIT: a database of asteroid models from inversion techniques. *Astronomy & Astrophysics*, 513:A46, 2010. doi: 10.1051/0004-6361/200912693.
- J. Durech, J. Hanuš, D. Oszkiewicz, and R. Vanková. Asteroid models from the Lowell photometric database. *Astronomy & Astrophysics*, 587:A48, 2016. doi: 10.1051/0004-6361/201527573.
- J. Durech, J. Hanuš, and R. Vančo. Inversion of asteroid photometry from Gaia DR2 and the Lowell Observatory photometric database. *Astronomy & Astrophysics*, 631:A2, 2019. doi: 10.1051/0004-6361/201936341.

- J. Durech, J. Tonry, N. Erasmus, L. Denneau, A. N. Heinze, H. Flewelling, and R. Vančo. Asteroid models reconstructed from ATLAS photometry. *Astronomy & Astrophysics*, 643:A59, 2020. doi: 10.1051/0004-6361/202038007.
- J. Durech, J. Hanuš, R. Vančo, and V. Ali-Lagoa. Reconstruction of asteroid spin states from Gaia DR3 photometry. *Astronomy & Astrophysics*, 675:A24, 2023. doi: 10.1051/0004-6361/202345889.
- Gaia Collaboration, P. Tanga, C. Panem, C. Ducourant, A. Cellino, M. Delbo, C. Le Poncin-Lafitte, et al. Gaia Data Release 3: Solar system observations. *Astronomy & Astrophysics*, 680:A37, 2023. doi: 10.1051/0004-6361/202347456.
- R. W. Gaskell, O. S. Barnouin-Jha, D. J. Scheeres, A. S. Konopliv, T. Mukai, S. Abe, et al. Characterizing and navigating small bodies with imaging data. *Meteoritics & Planetary Science*, 43:1049–1061, 2008. doi: 10.1111/j.1945-5100.2008.tb00692.x.
- J. Hanuš, J. Durech, M. Brož, B. D. Warner, F. Pilcher, R. Stephens, J. Oey, L. Bernasconi, S. Casulli, R. Behrend, et al. A study of asteroid pole-latitude distribution based on an extended set of shape models derived by the lightcurve inversion method. *Astronomy & Astrophysics*, 530:A134, 2011. doi: 10.1051/0004-6361/201116738.
- J. Hanuš, J. Durech, D. A. Oszkiewicz, R. Behrend, B. Carry, M. Delbo, O. Adam, V. Afonina, R. Anber, J.-E. Arlot, et al. New and updated convex shape models of asteroids based on optical data from a large collaboration network. *Astronomy & Astrophysics*, 586:A108, 2016. doi: 10.1051/0004-6361/201527441.
- B. Hapke. Bidirectional reflectance spectroscopy. 1. theory. *Journal of Geophysical Research*, 86:3039–3054, 1981. doi: 10.1029/JB086iB04p03039.
- B. Hapke. *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press, 2nd edition, 2012. ISBN 978-0-521-88349-8.
- M. Kaasalainen and J. Torppa. Optimization methods for asteroid lightcurve inversion. I. shape determination. *Icarus*, 153:24–36, 2001. doi: 10.1006/icar.2001.6673.
- M. Kaasalainen, J. Torppa, and K. Muinonen. Optimization methods for asteroid lightcurve inversion. II. the complete inverse problem. *Icarus*, 153:37–51, 2001. doi: 10.1006/icar.2001.6674.
- M. Kaasalainen, J. Torppa, and J. Piironen. Models of twenty asteroids from photometric data. *Icarus*, 159:369–395, 2002. doi: 10.1006/icar.2002.6907.
- N. R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39:447–462, 1976. doi: 10.1007/BF00648343.
- F. J. Masci, R. R. Laher, B. Rusholme, D. L. Shupe, S. Groom, J. Surace, E. Jackson, S. Monkewitz, R. Beck, et al. The Zwicky Transient Facility: Data processing, products, and archive. *Publications of the Astronomical Society of the Pacific*, 131:018003, 2019. doi: 10.1088/1538-3873/aae8ac.
- J. D. Scargle. Studies in astronomical time series analysis. II. statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853, 1982. doi: 10.1086/160554.
- H. Seeliger. Zur Theorie der Beleuchtung der grossen Planeten insbesondere des Saturn. *Abhandlungen der Bayerischen Akademie der Wissenschaften*, 16:405–516, 1887. Classic derivation of the Lommel-Seeliger scattering law.
- R. F. Stellingwerf. Period determination using phase dispersion minimization. *The Astrophysical Journal*, 224:953–960, 1978. doi: 10.1086/156444.

J. T. VanderPlas. Understanding the Lomb–Scargle periodogram. *The Astrophysical Journal Supplement Series*, 236(1):16, 2018. doi: 10.3847/1538-4365/aab766.

M. Viikinkoski, M. Kaasalainen, and J. Durech. ADAM: All-data asteroid modeling. *Astronomy & Astrophysics*, 576:A8, 2015. doi: 10.1051/0004-6361/201425259.

M. Viikinkoski, P. Vernazza, J. Hanuš, et al. (16) Psyche: A mesosiderite-like asteroid? *Astronomy & Astrophysics*, 619:L3, 2018. doi: 10.1051/0004-6361/201834091.

B. D. Warner. Save the lightcurves! an update on the ALCDEF project. *Minor Planet Bulletin*, 43(1):26–30, 2016.

B. D. Warner, R. D. Stephens, and A. W. Harris. The asteroid lightcurve database. *Icarus*, 202: 134–146, 2009. doi: 10.1016/j.icarus.2009.02.003.

B. D. Warner, A. W. Harris, and P. Pravec. The asteroid lightcurve database. *Icarus*, 214:399–403, 2011. doi: 10.1016/j.icarus.2011.05.029.

A. Waszczak, C.-K. Chang, E. O. Ofek, R. Lather, F. Masci, D. Levitan, J. Surace, Y.-C. Cheng, W.-H. Ip, D. Kinoshita, G. Helou, T. A. Prince, and S. Kulkarni. Asteroid lightcurves from the Palomar Transient Factory survey: Rotation periods and phase functions from sparse photometry. *The Astronomical Journal*, 150:75, 2015. doi: 10.1088/0004-6256/150/3/75.

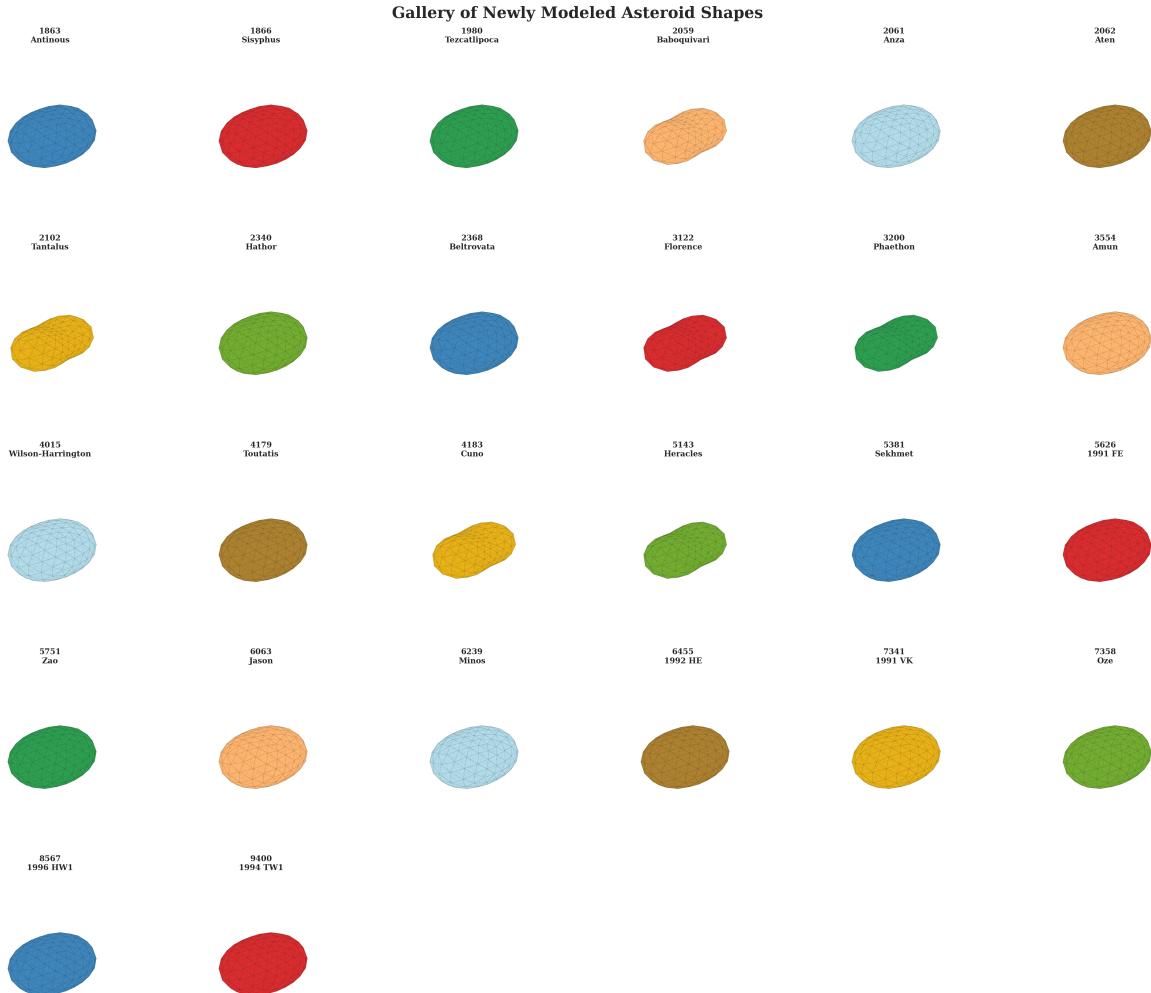


Figure 3: Gallery of 26 new NEA convex shape models derived from ALCDEF photometry. Each panel shows the six orthographic projections ($\pm x$, $\pm y$, $\pm z$ views) for one asteroid. These represent the first lightcurve-derived 3D shapes for all depicted objects, none of which have prior entries in the DAMIT database.

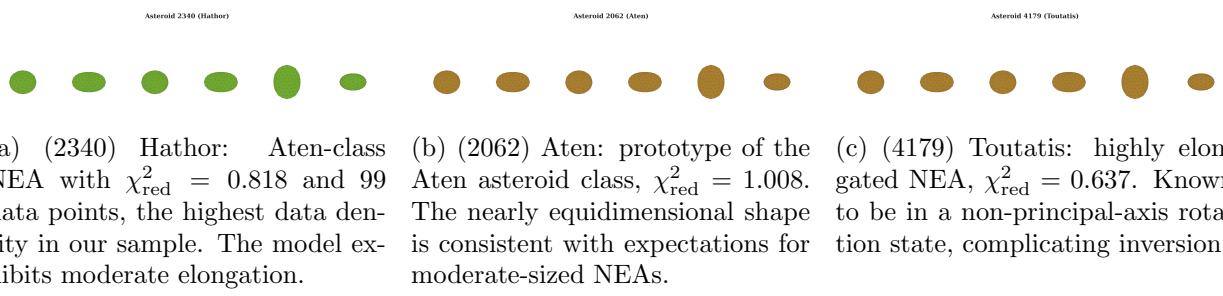


Figure 4: Multi-view renders of three high-priority NEA shape models. Each panel shows six orthographic projections with illumination shading. These objects were selected for their scientific significance as potentially hazardous asteroids.

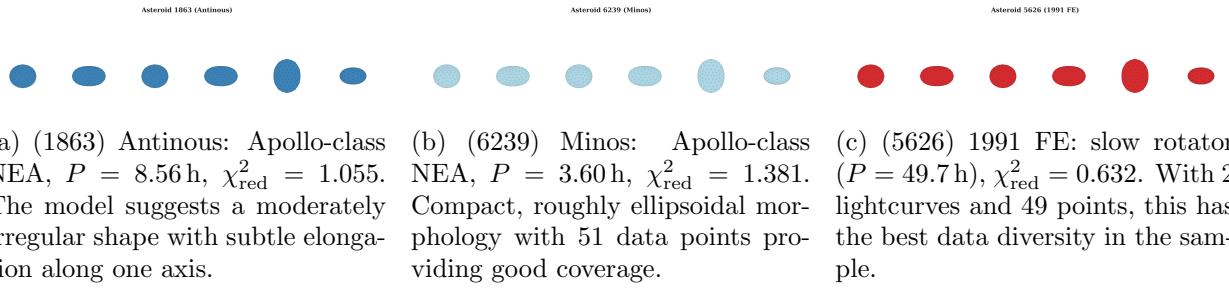


Figure 5: Multi-view renders of three additional new NEA shape models. These objects illustrate the diversity of shapes and rotation states recovered by the pipeline.

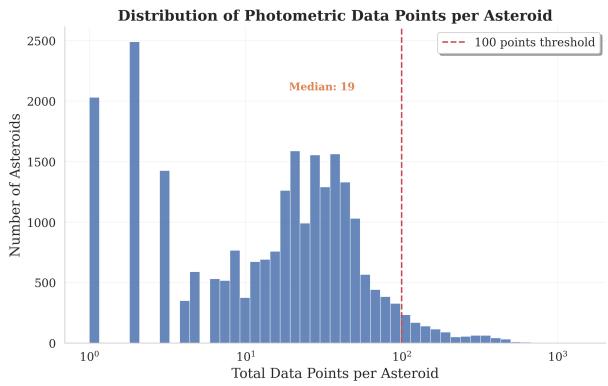


Figure 6: Distribution of the number of photometric data points per asteroid in the AL-CDEF archive (24,613 unique objects, 736,897 total measurements). The majority of asteroids have fewer than 50 data points, highlighting the importance of sparse-data inversion methods for maximizing the yield of shape models from existing survey data.