Accepted Manuscript

Added-value interfaces to asteroid photometric and spectroscopic data in the Gaia database

Johanna Torppa, Mikael Granvik, Antti Penttilä, Jukka Reitmaa, Violeta Tudose, Leena Pelttari, Karri Muinonen, Jorgo Bakker, Vicente Navarro, William O'Mullane

Added-value interfaces to asteroid photometric and spectroscopic data in the Gaia database

Johanna Torppa[a,*], Mikael Granvik[b,d], Antti Penttilä[b], Jukka Reitmaa[a],
Violeta Tudose[a], Leena Pelttari[a], Karri Muinonen[b], Jorgo Bakker[c], Vicente
Navarro[c], William O'Mullane[c]

[a]*Space Systems Finland, Kappelitie 6 B, 02200 Espoo, Finland. Emails:
johanna.t.torppa@gmail.com, jukka.reitmaa@gmail.com, violeta.tudose@ssf.fi,
marja-leena.pelttari@ssf.fi*
[b]*Department of Physics, P.O. Box 64, 00014 University of Helsinki, Finland. Emails:
mgranvik@iki.fi, antti.penttila@helsinki.fi, karri.muinonen@helsinki.fi*
[c]*European Space Astronomy Center, Camino bajo del Castillo, 28691, Villanueva de la
Cañada, Spain. Emails: jbakker@sciops.esa.int, vicente.navarro@esa.int, womullan@lsst.org*
[d]*Department of Computer Science, Electrical and Space Engineering, Luleå University of
Technology, Box 848, S-981 28 Kiruna, Sweden.*

Abstract

We present two added-value interfaces (AVIs) for analyzing photometric and
spectroscopic data observed by the Gaia satellite. The Gaia Added-Value Interface for Temporal Analysis (GAVITEA) is used to calculate an estimate for
the spin state and shape of an asteroid from its photometric data, and the Gaia
Added-Value Interface for Spectral Classification (GAVISC) provides tools to
define the taxonomic type and surface absorption coefficient based on spectroscopic asteroid data. Computations are mainly carried out using well-known
methods of asteroid data analysis but the AVIs also offer the possibility to test
novel methods that are specifically developed for analyzing temporally sparse
photometric data, typical for Gaia.

1. Introduction

The European Space Agency's (ESA) astrometric cornerstone mission Gaia
was launched in December 2013 from ESA's space port in Kourou in the French
Guyana (Gaia Collaboration et al., 2016b). After a commissioning phase it

---

*Corresponding author, johanna.t.torppa@gmail.com

started its science operations in August 2014. The first data release (DR1) in September 2016 did not contain data on solar system objects (SSOs; Gaia Collaboration et al., 2016a), but the subsequent data releases, starting with DR2, will contain gradually more extensive and more complex data sets on SSOs. Although Gaia is expected to discover tens of thousands new SSOs, Gaia's strength is in the extremely precise astrometry and photometry, and low-resolution spectroscopy of previously known SSOs. These data will enable numerous science investigations focusing on SSO properties such as asteroid shapes and spectral classification (Mignard et al., 2007). Gaia's original estimate for the cut-off magnitude is G=20.7, where G is comparable to Johnson-Cousins V magnitude (Jordi et al., 2010). This suggests that some hundreds of thousands of SSOs will be detected during the nominal 5-year Gaia survey (Mignard et al., 2007). The actual number will most likely be closer to half a million, because the limiting magnitude for transmission to ground is higher than expected. An extension to the nominal survey would not substantially increase the number of objects detected (unless the cut-off magnitude is increased) but will provide more frequent detection of each SSO. Studies concerning the performance of Gaia in terms of the number of observed asteroids and the quality of photometric and spectroscopic observations have been carried out by, for instance, Tanga and Mignard (2012), Delbo et al. (2012).

In 2006, the Gaia Data Processing and Analysis Consortium (DPAC) was formed and it was selected to provide the ground segment of Gaia. The role of the Gaia DPAC is to provide the software and computational infrastructure needed for the scientific processing of the Gaia data and the production of the Gaia catalogue (e.g., Cellino et al. (2015), Santana-Ros et al. (2015) and Delbo et al. (2012)). DPAC is composed of about 450 expert scientists across Europe with well-defined responsibilities. It is noteworthy that DPAC members have no privileged access to Gaia data apart from limited science verification studies that are published simultaneously (or near-simultaneously) with a particular data release. The lack of a proprietary period implies that everybody (that is, also a scientist outside DPAC) has identical opportunities to publish novel

2

results based on Gaia data as soon as it becomes public. To enable efficient data analysis and publication following each data release, the data-analysis software should be ready, well documented and in most cases close, if not integrated, to the Gaia database to allow rapid processing of large quantities of data. Software for analysis that require active expert interference should be available to the scientific community so that a large number of research groups can contribute to the data analysis.

The Gaia database, developed by DPAC, is limited to basic analysis, and there has been no opportunity for non-members to contribute to the analysis software. The Gaia Added-Value Interface (Gaia AVI) project was initiated by ESA in 2015 to establish a platform for added-value interfaces to Gaia data. Gaia AVIs allow users to analyze and visualize various types of Gaia data with tools that are not planned to be part of the Gaia database. In addition to the platform, called the Gaia added-value interface platform (GAVIP; http://gavip.esac.esa.int; Vagg et al., 2016), ESA also funded the development of three science AVIs. The programmatic goal for the science AVIs was to use them as case studies that would guide the development of GAVIP, so that it would evolve into flexible and user-friendly portal that would inspire other scientists to develop their own AVIs. In the future, Gaia AVIs can be provided by anyone with interest in further processing or analysis of Gaia data. Two of the science AVIs, developed by Space Systems Finland Ltd. and the University of Helsinki, are presented in this paper: Gaia Added Value Interface for Temporal Analysis (GAVITEA) is a tool for analysing the temporal variation of observed asteroid brightness in order to gather information on their spin state and shape, and Gaia Added Value Interface for Spectral Classification (GAVISC) is a tool for analysing asteroid spectra for defining their taxonomic type and surface absorption properties. The information provided by GAVITEA and GAVISC is complementary to that to be provided by DPAC. While DPAC will deliver information on asteroid shapes using tri-axial ellipsoids at the end of the Gaia mission, GAVITEA focuses on providing convex shape models with uncertainty estimates based on the limited data available earlier during the mission. Sim-

3

ilarly, while DPAC will provide a new spectral classification for asteroids for which there exists low-resolution spectra obtained by Gaia, GAVISC will provide the most likely spectral type for each asteroid in the Bus-DeMeo (B-DM) classification. Virtually all the observed asteroids can be analyzed with GAVITEA and GAVISC.

An alert service worth mentioning, developed in the framework of DPAC, is the Gaia Follow-Up Network for Solar System Objects (Gaia-FUN-SSO, https://gaiafunsso.imcce.fr/) developed as part of task DU459 of the Coordination Unit 4 (Object processing) of DPAC. Astrometric alerts are submitted to the Gaia-FUN-SSO from the Solar System short-term processing pipeline (Tanga et al., 2016) and can be used to schedule complementary observations to aid the analysis of Gaia data. This is useful especially considering GAVITEA; defining the rotation period for an asteroid using Gaia data alone is extremely tedious, since not a single full rotation cycle is covered but only single brightness observations about one month apart. Even one longer (hours) continuous sequence of observations would be helpful to constrain the rotation period and reduce the computation time.

In what follows we first describe the asteroid data that will eventually be available through the Gaia database. Then we continue with a description of the GAVITEA and GAVISC algorithms and their theoretical foundations, and describe the implementation in the Gaia AVI framework. Finally, we present our validation approach and validation results, and finish with conclusions including some potential future developments.

2. Gaia data of asteroids

Gaia DPAC will publish Gaia data in the Gaia archive and submit astrometry of asteroid discoveries to the Minor Planet Center. Submission to other repositories, such as PDS and asteroid light curve databases, will likely be up to members of the community at large. The Gaia data releases (DR; intermediate and final) will contain astrometric, photometric and spectroscopic data on

4

asteroids, although the exact schedule and contents for DRs have not yet been decided upon (apart from DR2 which will be released on April 25, 2018). ESA Cosmos web pages for the Gaia mission (https://www.cosmos.esa.int/web/gaia) are, at the time of writing this paper, the best public constantly updated source of information on the schedule of the data releases and contents of the Gaia archive. The astrometric data — uncertainty at the milliarcsecond (mas) level or even below for the brightest, non-saturated asteroids — will be the most revolutionary contribution from the Gaia mission in terms of accuracy. Most of the DPAC effort therefore also focuses on ensuring that the astrometric data can be exploited as extensively as possible.

Each astrometric measurement is associated with the corresponding measurement of the G-band flux (G stands for Gaia) in the astrometric CCDs. The G-band wavelength range is 0.33–1.05 µm and provides high-signal-to-noise-ratio photometry. While the photometry is less accurate than the astrometry relative to previous surveys, it is still expected to have a 1σ scatter of only millimags to few tens of millimags depending on the brightness of a stellar source with a spectral energy distribution similar to the Sun. The G-band data is thus very well suited for variability studies. During the estimated Gaia lifetime of five years, each asteroid will be observed at a wide range of observing geometries, providing good basis for spin state and shape determination. However, the total number of photometric observations depends on the asteroid's orbit and is not necessarily large (about 70 observations maximum, usually a few tens), and data do not cover full rotation cycles. Traditional lightcurve analysis techniques are not generally applicable for such data, but modified techniques must be used.

Spectro-photometry will be recorded with the blue photometer (BP; 0.33–0.68 µm) and the red photometer (RP; 0.64–1.05 µm). The BP and RP prisms spread the light from the targets over about 45 pixels each (with the wavelength range per pixel varying across the photometers), thus providing low-resolution spectroscopy well-suited for spectroscopic studies of asteroids.

5

3. Gaia Added-Value Interface for Temporal Analysis

Gaia Added-Value Interface for Temporal Analysis (GAVITEA) uses Gaia data on asteroid brightnesses to perform photometric lightcurve analysis for determining an asteroid's spin state and shape using three different algorithms. All the algorithms use the same basic approach of minimizing the difference between computed model brightness and the observed brightness. The main differences between the algorithms are related to the optimization approach and the parameterization of the shape model:

1. The algorithm for initial solution (A-IS) uses both systematic sampling and least-squares minimization to solve for the spin state and ellipsoid shape parameters.
2. The algorithm for ellipsoid solution (A-ES) performs statistical sampling of spin state and ellipsoid shape parameters.
3. The algorithm for convex solution (A-CS) performs statistical sampling of spin state and shape parameters of a convex spherical harmonics shape model.

In the following, we describe the model brightness computation which uses a common light scattering model for all the algorithms. In A-IS and A-ES the model brightness is computed analytically while in A-CS numerical integration is used. We also describe in detail the parameter optimization and other specific features for each algorithm.

3.1. Model brightness computation

3.1.1. Algorithms for initial and ellipsoid solutions

The shape model used in A-IS and A-ES is a triaxial ellipsoid. The disk integrated brightness $L(\alpha)$ of a Lommel-Seeliger scattering ellipsoid for a given observing geometry can be computed analytically following Muinonen and Lumme

6

(2015):

$$L(\alpha) = \frac{1}{8}\pi g F(\alpha) abc \frac{S_\odot S_\oplus}{S} \times \Big[\cos(\lambda' - \alpha') + \cos(\lambda') +$$
$$\sin(\lambda')\sin(\lambda' - \alpha') \times \log\left(\cot\left(\frac{\lambda'}{2}\right)\cot\left(\frac{\alpha' - \lambda'}{2}\right)\right)\Big], \quad (1)$$

where α is the phase angle, D is the size parameter, g is the geometric albedo, $F(\alpha)$ is the phase function and a,b,c are the principal axes of the ellipsoid. S, $S_\odot$, $S_\oplus$, α' and λ' are defined as follows:

$$S_\oplus = \sqrt{\mathbf{r}_{\mathrm{obs}}^T \, \mathbf{C} \, \mathbf{r}_{\mathrm{obs}}},$$
$$S_\odot = \sqrt{\mathbf{r}_{\mathrm{ill}}^T \, \mathbf{C} \, \mathbf{r}_{\mathrm{ill}}},$$
$$\cos\alpha' = \frac{\mathbf{r}_{\mathrm{ill}}^T \, \mathbf{C} \, \mathbf{r}_{\mathrm{obs}}}{S_\odot S_\oplus},$$
$$\sin\alpha' = \sqrt{1 - \cos^2\alpha'},$$
$$S = \sqrt{S_\odot^2 + S_\oplus^2 + 2S_\oplus S_\odot \cos\alpha'},$$
$$\cos\lambda' = \frac{S_\odot + S_\oplus \cos\alpha'}{S},$$
$$\sin\lambda' = \frac{S_\oplus \sin\alpha'}{S},$$

where r_{obs} and r_{ill} are the vectors from the asteroid to the observer and to the Sun, respectively, and

$$\mathbf{C} = \begin{bmatrix} a^{-2} & 0 & 0 \\ 0 & b^{-2} & 0 \\ 0 & 0 & c^{-2} \end{bmatrix}.$$

3.1.2. Algorithm for convex solution

In A-CS, the Gaussian curvature of the convex shape model is given as a spherical harmonics series (Muinonen et al., 2018). For numerical integration of model brightness, the surface is discretized using the quaternary triangular mesh (see, e.g., Dutton (1999)). For N_r rows in an octant, the number of facets is $4N_r^2 + 2$ and the number of vertices is $8N_r^2$. A sufficient value for N_r is in the range $N_r = 8...10$, producing no significant discretization artifacts in rotational

7

brightness variation. The model brightness $L(\alpha)$ for phase angle α is

$$L(\alpha) = \sum_{n=1}^{N_f} \left(\frac{1}{2}D\right)^2 A_n S_n, \tag{2}$$

where N_f is the number of facets, D is the size parameter, A_n is the facet area, and S_n is the scattering function value at facet n. Following the Lommel-Seeliger scattering function,

$$S_n = \frac{gF(\alpha)}{2} \frac{\cos\theta_\oplus \cos\theta_\odot}{\cos\theta_\oplus + \cos\theta_\odot}, \tag{3}$$

where $\theta_\oplus$ and $\theta_\odot$ are the angles between the facet normal vector and the directions to the observer and to the Sun, respectively. Integration weights for the polar angle are defined using Gauss-Legendre quadrature of order 64. The azimuthal angle is evenly divided in 180 parts for integration.

The facet areas are computed by multiplying the corresponding facet areas on a sphere by the Gaussian curvature, expressed as the exponential of associated Legendre polynomial

$$A_n = \sum_{m=1}^{m_{max}} \sum_{l=m}^{l_{max}} P_l^m(\theta)(a_{lm}\cos\phi + b_{lm}\sin\phi), \tag{4}$$

where ϕ and θ are the azimuthal and polar angles of the facet normal vector, a_{lm} and b_{lm} are the series coefficients, and P_l^m are the associated Legendre functions.

3.1.3. Phase function

The phase function $F(\alpha)$, common for all the algorithms, is modelled with the H, G_{12} function following Muinonen et al. (2010) as

$$F(\alpha) = G_1\phi_1(\alpha) + G_2\phi_2(\alpha) + (1 - G_1 - G_2)\phi_3(\alpha), \tag{5}$$

where $\phi_1(\alpha) = 1 - 6\alpha/\pi$ and $\phi_2(\alpha) = 1 - 9\alpha/5\pi$ for $\alpha < 7.5°$ and $\phi_3(\alpha) = 0$ for $\alpha > 30°$. The rest of the phase angle range is modelled using cubic splines fitted to empirically defined sets of values and their derivatives optimized using a representative sample of the asteroid population.

8

Parameters G_1 and G_2 are piecewise linear functions of G_{12}, determined using real observational data as:

$$G_1 = \left\{ \begin{array}{ll} 0.7527G_{12} + 0.06164 & \text{if } G_{12} < 0.2 \\ 0.9529G_{12} + 0.02162 & \text{if } G_{12} \geqslant 0.2 \end{array} \right\},$$

$$G_2 = \left\{ \begin{array}{ll} -0.9612G_{12} + 0.6270 & \text{if } G_{12} < 0.2 \\ -0.6125G_{12} + 0.5572 & \text{if } G_{12} \geqslant 0.2 \end{array} \right\}. \tag{6}$$

3.2. Algorithm descriptions

3.2.1. Algorithm for initial solution

A-IS approaches the spin solution from scratch, systematically sampling first the rotational period and then the spin axis, while fitting all the other model parameters. The shape model is a triaxial ellipsoid. A-IS is suitable for cases where an estimate of the spin state of the asteroid is not available or where the rotation period is poorly known. Existing estimates should be made use of, and they can be found (for 948 asteroids, as of April 18, 2018) in DAMIT (Database of Asteroid Models from Inversion Techniques, http://astro.troja.mff.cuni.cz/projects/asteroids3D/web.php, Ďurech et al. (2010). A-IS is divided in three stages for 1) period scanning, 2) pole scanning and 3) full model optimization.

Period scanning

The first stage of the A-IS is period scanning, where evenly spaced rotation periods are sampled while all the other model parameters are fitted. The expected minimum and maximum values for the period, $P_{\min}$ and $P_{\max}$, are given as input to the algorithm. The sampling step is computed by taking into account the total time span of the observations as well as the expected period value as

$$d_{\mathrm{per}} = sc_{\mathrm{per}} \, 0.5 \, P^2/\Delta T, \tag{7}$$

where sc_{per} is an additional input coefficient for adjusting the period sampling step, P is defined as the mean of $P_{\min}$ and $P_{\max}$, and ΔT is the total time span of the observations. In the GAVITEA implementation, also the maximum number of sampled periods N_{pmax} is provided to avoid extremely long computation

9

times. If the number of periods computed from Eq. 7 is greater than N_{pmax}, then the number of sampled periods is set to N_{pmax}, and d_{per} is computed as $\mathrm{int}((P_{\max} - P_{\min})/N_{\mathrm{pmax}})$. The goodness of each sampled period is estimated by fitting the other model parameters using simplex minimization. Initial model parameter values and parameter deviations are given realistic hardcoded values, except for the pole, for which a user defined number of initial values are generated using the quaternary triangular mesh (See Section 3.1.2).

Pole scanning

The second step in A-IS is the spin-axis scanning, where a user-defined number of pole longitudes and latitudes are sampled using the quaternary triangular mesh while fitting all the other model parameters using simplex minimization. Initial values for model parameters, apart from the pole coordinates, are obtained from the best-fit solution of period scanning. The goodness of the fit using each pole is evaluated using the residual root-mean-square (RMS) of the fit.

Full-model optimization

The third and last step in A-IS is the full-model optimization, where a single best-fit solution is found by fitting all the model parameters using simplex minimization, taking the initial model parameter values from the best-fit solution of the pole scanning step. The best-fit solution P_{a} from this step is used later in A-ES and A-CS.

3.2.2. Algorithm for ellipsoidal solution

A-ES approaches the spin and shape solution in a statistical manner using the best-fit result of A-IS as the initial solution. An ellipsoid shape model is used, and the model brightness is computed according to Muinonen and Lumme (2015) as in A-IS (Eq. 1).

A-ES is executed in three stages, each of which provides an independent statistical distribution for the model parameters: 1) generation of virtual solutions, 2) importance sampling and 3) Markov-chain Monte Carlo (MCMC) sampling. Virtual solutions are always generated and MCMC sampling is ex-

ecuted if importance-sampling solutions are not found. If neither importance-sampling nor MCMC solutions are found, virtual solutions are used as the result distribution of possible solutions.

Virtual solutions

Virtual solutions are obtained by finding the best-fit model parameters for virtual observations, that are generated by adding Gaussian noise to the original observations. The strength of the noise is provided as input to the algorithm and initial parameter vector for the simplex minimization is P_a from A-IS. Only solutions with acceptable fit to the data are stored as virtual solutions P_v. A sufficient number of virtual solutions can be used to present the distribution of possible model parameters.

Importance sampling

Virtual solutions are used to modify the parameter vector in importance sampling. At every iteration, a modified parameter vector P_p is obtained using equation

$$P_p = P_{per} + P_{v,k1} - P_{v,k2}, \tag{8}$$

where $k1$ and $k2$ refer to two randomly picked virtual solutions. The modified parameter vector P_p is accepted if $\chi^2 - N_{obs} < d\chi^2$, where χ^2 is the chi-square of the fit and $d\chi^2$ is provided as input. N_{obs} is the number of observations. The weight of each IS solution trial i is initialized as

$$w_i = e^{-(\chi_i^2 - N_{obs})/2}, \tag{9}$$

and updated at each unsuccessfull iteration as

$$w_i = w_i + e^{-(\chi_{i-1}^2 - N_{obs})/2}. \tag{10}$$

Markov-chain Monte Carlo

Similarly as for importance sampling, the proposed parameter vector in MCMC sampling is calculated using virtual solutions. The criterion for accepting the proposed solution in MCMC is

$$RAN < e^{-(\chi_i^2 - \chi_{i-1}^2)/2}. \tag{11}$$

11

The weights of the MCMC solutions are initialized for each new solution trial i as $w = 1$ and updated at each non successfull iteration as $w = w + 1$.

3.2.3. Algorithm for convex solution

Like A-ES, A-CS approaches the spin and shape solution in a statistical manner using the best-fit result of A-IS as the initial solution.

A-CS is executed in four stages. The first three are carried out following the description for A-ES, i.e., 1) generation of virtual solutions, 2) importance sampling, 3) Markov-chain Monte Carlo sampling. The only difference in these three steps in comparison to A-ES is the non-ellipsoidal convex shape model that requires numerical integration of the disk-integrated model brightness (Sec 3.1.2).

The fourth stage, Minkowski minimization, is needed in A-CS for generating the actual 3D model shape for visualization of the result as well as for checking if the model shape is realistic and not, for instance, extremely elongated or flat. The shape is constructed from the set of polyhedron facet areas, which are solved by fitting the computed model brightness (Eq. 2) to the observed brightness using the Levenberg-Marquardt minimization procedure. The solution for the Minkowski problem, originally formulated by Minkowski (1903) and later adjusted for asteroid photometric analysis by Kaasalainen et al. (1992) and Kaasalainen and Torppa (2001), uses the concept of so called mixed volume of two bodies E and F. The mixed volume $V(E,F)$ for a discretized convex surface can be written as

$$V(\mathrm{E}, \mathrm{F}) = \frac{1}{3} \sum_{j=1}^{N_f} l_j(\mathrm{E}) a_j(\mathrm{F}), \tag{12}$$

where $a_j(\mathrm{F})$ is the facet area of facet j on the surface of F and $l_j(\mathrm{E})$ is the distance of the corresponding facet j on the surface of E from the origin of E.

According to Minkowski, for a constant volume of body E, $V(\mathrm{E}, \mathrm{F})$ reaches its minimum when E and F refer to the same convex shape. Thus, by minimizing the mixed volume, the facet distances from the origin, corresponding to given facet areas, can be solved for. An analogous, and computationally easier way

12

to solve the problem is to maximize the volume of E, while keeping the mixed volume constant. For a discretized surface, the volume to be maximized is

$$V(E) = \frac{1}{3} \sum_{j=1}^{N_f} l_j A_j(\mathbf{l}),$$ (13)

where $A_j(\mathbf{l})$ is the area of the facet j as computed from $\mathbf{l}$. The constraint of keeping the mixed volume constant is fulfilled by projecting the gradients in the optimization procedure onto the constraint plane.

4. Gaia Added-Value Interface for Spectral Classification

Gaia Added-Value Interface for Spectral Classification (GAVISC) takes the reflectance measurements of an asteroid at different wavelengths as input, and performs spectral classification in the Bus-DeMeo (B-DM) system for the target using the algorithm SPEC-C. In addition, GAVISC inverts the spectral behaviour of the surface material's effective absorption coefficient using the SPEC-AC algorithm. Implementation is based on theoretical studies by Lindqvist et al. (2018), Martikainen et al. (2018) and Muinonen et al. (2009).

4.1. Spectral classification

The SPEC-C algorithm suggests a B-DM spectral type for the target asteroid. The taxonomy is presented in DeMeo et al. (2009), where it is defined for wavelengths ranging from 0.45 to 2.45 µm. Classifying asteroids observed with Gaia is not completely straightforward, because Gaia has 120 spectral bands from 0.33 to 1.05 µm, so the range is only partially overlapping with the B-DM system. For this reason, the classification provided by GAVISC should be treated as an estimate for a B-DM class.

There are other possibilities for taxonomic classification systems, most popular being the Tholen taxonomy (Tholen, 1984) and the Bus taxonomy (Bus, 1999). However, we preferred here the B-DM over these. The wavelengths used by the Tholen system (0.337–1.041 µm) are covering the Gaia wavelength bands almost completely. But, the Tholen system uses only seven bands on that range,

while Gaia would have 97 bands available. The Bus taxonomy is based on a more dense wavelength coverage, but then the upper limiting wavelength is 0.92 µm, which would force us to abandon 20 of the largest Gaia bands.

The first step in building the algorithm was to invert the original spectra from the 371 asteroids that were used to build the B-DM system from their principal-component presentation, and interpolate them at Gaia wavelengths. After this step, we have a 371×99 matrix that stores the 99 wavelengths from Gaia that overlap the B-DM wavelengths.

The B-DM taxonomy is built on the principal-component presentation of the 371 asteroid spectra, and consists of 24 taxonomic classes. Because the Gaia data are lacking the near-infrared wavelengths and is unable to distinguish certain classes, we limit our Gaia-B-DM taxonomy to only 13 main classes, joining the subclasses to their master complex, e.g., Cb, Cg, Cgh, and Ch to C. Also, since we now have the classes already defined, we choose to use the linear discriminant analysis (LDA) instead of the principal-component analysis (PCA). The methods are very similar, but LDA makes use of a training set with pre-defined classes, which is available in this case, whereas PCA is for cases with no pre-defined classification information.

4.1.1. Linear discriminant transform

Our training set for classification consists of all the B-DM data with classes A, B, C, D, K, L, O, Q, R, S, T, V, or X. For LDA, we need to compute from the training set within-class covariance matrices $\mathbf{S}_c$, their joined sum $\mathbf{W}$, between-classes covariance sum $\mathbf{B}$, class mean vectors $\mathbf{m}_c$, and total mean vector $\mathbf{m}$. The within-class and between-classes covariance sums are computed as

$$\mathbf{W} = \sum_c (n_c - 1)\mathbf{S}_c,$$
$$\mathbf{B} = \sum_c n_c(\mathbf{m}_c - \mathbf{m})^T(\mathbf{m}_c - \mathbf{m}), \tag{14}$$

where c iterates over the 13 classes, and n_c is the number of targets in class c.

From these we can form the LDA projection matrix with the help of eigen-

14

value decomposition as

$$\mathbf{L}\mathbf{\Lambda}\mathbf{L}^{-1} = \mathbf{W}^+\mathbf{B}, \tag{15}$$

where $^+$ is the Moore-Penrose pseudoinverse. Now the matrix $\mathbf{L}$ holds the eigen-vectors that project the spectral values to LDA-space. According to eigenvalue magnitudes, we select the three first LDA-dimensions to be used in classifica-tion as matrix $\mathbf{L}_3$. Finally, the LDA-projected value $\mathbf{x}$ of Gaia asteroid spectral observation $\mathbf{y}$ is

$$\mathbf{x} = (\mathbf{y} - \mathbf{m})\mathbf{L}_3. \tag{16}$$

4.1.2. Classification with naive Bayesian classifier

The final selection of the most probable taxonomic class for an observed Gaia spectrum $\mathbf{y}$ and its LDA-transformed version $\mathbf{x}$ is done using the so-called naive Bayesian classifier (NBC) with Gaussian classes. We can convert our training set into LDA-coordinates, as described above, and compute class means $\mathbf{m}_c^*$ and covariance matrices $\mathbf{S}_c^*$ in the LDA-space. The NBC treats every class as a (Gaussian) probability distribution, and our training set gives the estimated parameters to these distributions. The NBC probability for an observation $\mathbf{x}$ to belong to class c is therefore computed as

$$p_c = a_c \, \mathrm{f}(\mathbf{m}_c^*, \mathbf{S}_c^*), \tag{17}$$

where f is the probability density function of multivariate Gaussian distribution, and a_c is the a priori probability of class c, which is computed from the training set frequencies. The NBC probability will be computed for all the classes, and the most probable class is selected.

4.2. Absorption spectra

The reflectance (i.e., albedo, magnitude) of an object is a function of the optical and geometric (size, shape, packing) properties of the target's material. For visual wavelengths, it is the very surface of the object, the regolith layer, that affects the reflectance. The optical properties can be described by the refractive index of the regolith material, $m = n + ik$. The imaginary part of m governs

15

the absorption properties of the material that change with wavelength. The behaviour of k as a function of wavelength, i.e., the k spectrum, is characteristic to the reflecting material. By relating the k spectrum of an object with unknown surface material to k spectra of known materials, we can make an estimate of the composition of the target object's surface material.

The SPEC-AC algorithm inverts the observed spectral reflectance of the target into the k spectrum using pre-computed library of modelled reflectancies with given absorption coefficients. The range of k-values in the library go from 1.6×10^{-6} to 0.006, which will produce the geometric albedo values in the range from ~1 to 0.01. The light-scattering model that we use is the SIRIS code for geometrical optics with diffuse scatterers (Muinonen et al., 2009). In what follows, we describe the modeling procedure for creating the library.

SIRIS is able to compute scattering properties for a particle that is large compared to the wavelength and that can have internal diffuse scatterers in the geometrical optics/radiative transfer approximation. The particle geometry is modeled as a Gaussian random sphere with a given size and radial variance, and correlation function. For robust inversion, some model parameters must be fixed while the absorption coefficient k is estimated. The particle size we use is 200 μm with the relative Gaussian-sphere-radius variance 0.04 and the power-law correlation function index 3. Examples of these kinds of regolith grains are shown in Fig. 1. The real part of the refractive index, n, is fixed at 1.5, since it has only a small effect on the macroscopic reflectance spectra.
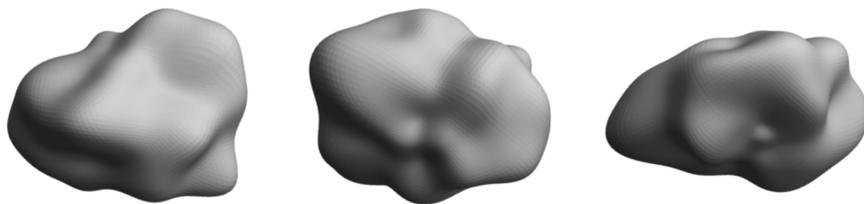
Figure 1: Three examples of Gaussian random sphere shapes used to model the regolith grains when inverting the absorption spectra.

We compute the modeled reflectance in two rounds. In the first round, the scattering properties, i.e., the single-scattering albedo ϖ, the extinction efficiency Q_{ext}, and the orientation-averaged Mueller matrix (i.e., scattering matrix, transformation matrix) $\mathbf{M}$, are computed for a single grain (See, e.g., van de Hulst, H.C., 1957). In the second round, the macroscopic asteroid is treated in SIRIS as an imaginary particle with refractive index of vacuum and with the grains computed in the first round as diffuse scatterers inside the large 'grain'. This is, effectively, a Monte Carlo radiative-transfer solution for the target consisting of these regolith-grain scatterers.

From the modeling described above we can build a library of reflectance for every Gaia wavelength with a range of absorption coefficients. The actual inversion that we do in the SPEC-AC algorithm is that we interpolate, for every wavelength, the k value that will produce the observed reflectance.

We note that for absorption inversion we need observations that are in the geometrical albedo scale, the normalized reflectance cannot give the proper scaling for absorption. As for now, the Gaia DR1 does not include asteroid observations. For that reason, it is a bit unclear in what units the Gaia asteroid spectra will be given in future data releases, and if there will be a Gaia albedo estimate for each object. We surpass this problem at this point by assigning an average albedo for the object using the abovementioned B-DM spectral classification and the albedo for the estimated spectral class from the NASA NEOWISE data (Mainzer et al., 2016, 2011, 2012; Masiero et al., 2014, 2017). In Fig. 2 we show the geometric albedo distribution for our subclasses of B-DM taxonomy.

5. Implementation

GAVITEA and GAVISC, like all other Gaia AVIs, are built on the Gaia Added-Value-Interface Platform (GAVIP), which provides a uniform environment for the user to manage the various AVIs and their associated data products. Both GAVISC and GAVITEA can be accessed also programmatically, or using the GAVIP client (http://user-manual.docs.gavip.science/GAVIP-Client.html).

17
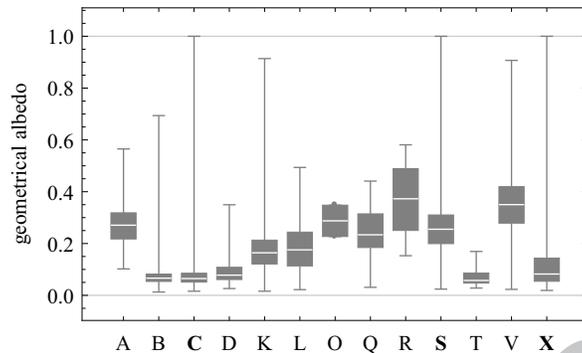
Figure 2: Geometric albedo from NEOWISE data for the taxonomic classes or complexes we use in the GAVISC SPEC-AC algorithm. Before using the data, we remove the smallest and largest 5 % of the values The box marks the 75 % quantile of the data with the median value in the middle, while the lines show the full spread of the (trimmed) data. We are using the mean of the data as the estimate for class albedo value.

Little design effort has been put towards programmatic use, however, and both AVIs are primarily intended to be used via the web GUI. The standard workflow that the user follows when using GAVISC or GAVITEA consists of four main steps: 1) selecting the asteroid to be analyzed, 2) providing input parameters for the computation, 3) submitting the computation task to GAVIP, 4) inspecting and downloading the results. Computations are scheduled by GAVIP and the total time from submitting the computation to achieving the results depends on the overall activity of GAVIP AVI users. The computation time estimates below refer to the actual GAVITEA and GAVISC computations once the computation is started by GAVIP.

The server component of both AVIs is implemented in Python 2.7, and the HMI (human-machine interface) as a web application using HTML5 and JavaScript (conforming mostly to ECMAScript 5). The use of Python version 2.7 (instead of the more recent Python 3.x was dictated initially by the availability of a GAVIP-provided container template, which included the older version of Python. Later on a container template with Python 3.x was made available, but the effort to port the AVI source code to Python 3.x was not deemed worth-

18

while. The Bokeh data-visualization library was utilized for visualization of the data products.

The server component exposes an interface for interacting over HTTP/1.1 (Hypertext Transfer Protocol version 1.1) conforming roughly to the established REST (Representational State Transfer) architectural software design principles. REST principles are typically utilized in web application client-server interactions, and require each client request to make use of a limited set of server-side operations (as opposed to allowing arbitrary remote procedure calls) and for all requests to contain all necessary information (e.g. authentication credentials) for the server to correctly process the request without having to refer to information in earlier requests. For a clear and detailed exposition of REST, see Fielding (2000).

While interacting with a so-called RESTful interface (conforms to REST principles) would be possible using static HTML5-pages only, JavaScript logic is added to the HMI to provide improved user experience, e.g., for form validation and interactivity.

The server component is divided as per GAVIP-mandated design into a frontend, for servicing user requests, and a backend, for asynchronous execution of added-value computation jobs. The frontend is implemented using the Django web framework, and the backend using Luigi task management framework for computation task dependency resolution and execution. Django is a software framework for writing web application server components in Python. Django provides e.g. basic HTTP request and response processing as well as handling of different types of message payloads, and facilitates developing RESTful web applications via an additional software framework (Django REST Framework).

5.1. GAVITEA

The scientific computations for GAVITEA are implemented in Fortran 90 instead of Python primarily for reasons of computational efficiency; further, earlier implementations of most of the functions were available for the developers in Fortran 77.

19

A-IS and A-ES computation of a single best-fit solution is fast (seconds), due to the analytical disk integrated brightness computation. However, especially for A-IS, the total computing time can easily become extremely long (days), if attention is not paid to input parameters, such as the sampled period range $[P_{min}, P_{max}]$ or the maximum number of sampled periods. A-CS computation time for a single solution is longer (minutes) due to numerical integration of the model brightness, and the total computation time depends directly on the number of output solutions.

Running either of A-ES or A-CS requires that the result of A-IS is available to be used as the initial solution. A-IS results can be computed in the current session prior to the other algorithms or, alternatively, the user can save time by using precomputed A-IS results if such have been saved in the user space or shared space.

Results are provided as text files as well as images. A-IS result is visualized as a plot showing the RMS of the fit for each scanned period for the period scanning phase (Fig. 3), and two plots showing the RMS of the fit for the scanned periods and the corresponding best fit shapes for the pole scanning phase (Fig. 4). A-ES and A-CS spin solution and A-ES shape solution are represented with similar plots as those for A-IS pole scanning, but show a different range and type of distribution (Fig. 5). A-CS shape results are visualized in an interactive 3D plot, where the user can select and rotate proposed asteroid solutions (Fig. 6). The 3D plot is implemented using WebGL technology, available in most modern web browsers.

The user can adjust computation parameters listed in Table 1.

5.2. GAVISC

The GAVISC implementation can be divided in algorithms for the data analysis described in Sec. 4, and in the AVI frontend and backend implementation. All the science algorithms are implemented using the Python numerical computation libraries NumPy and SciPy.

The interface to GAVISC is very simple, since no input parameters are given

20

Table 1: Input parameters for algorithm validation. In the "Values for validation" column, values on the top rows refer to validation with different data sets and the value ranges on the bottom row refer to validation with different input parameters. P is the correct rotation period of the simulated asteroid.

Param.	Description	Alg.	Values for validation
P_{min}	Minimum sampled period	A-IS	0.8P (0.1-0.9)P
P_{max}	Maximum sampled period	A-IS	1.2P (1.1-10)P
sc_{per}	Resolution coefficient for period sampling	A-IS	0.5 0.1-1.0
N_{pmax}	Maximum number of sampled periods	A-IS	400 100-500
$nPol_1$	Number of sampled poles for period sampling	A-IS	18 18-88
$nPol_2$	Number of sampled poles for pole sampling	A-IS	146 18-578
ca	Minimum allowed ellipsoid axis ratio c/a	A-IS A-ES	0.3 0.1-0.7
tol	Tolerance of simplex minimization	A-IS A-ES	0.001 0.0001-0.01
g	Geometric albedo	All	0.05 0.05
N_{mcv}	Number of virtual solutions	A-ES A-CS	1000 10-5000
N_{mcmc}	Number of MCMC solutions	A-ES A-CS	100 10-1000
σ_v	Virtual error for generating virtual data	A-ES A-CS	0.00001 0.00001-0.0001
vsc	Acceptance threshold for virtual error solutions	A-ES A-CS	5.0 0.5-20
N_r	Number of rows for the quaternary triangular mesh	A-CS	8 6-10
l_{min}	Minimum degree of spherical harmonics series	A-CS	0 0-2
l_{max}	Maximum degree of spherical harmonics series	A-CS	5 3-10

21
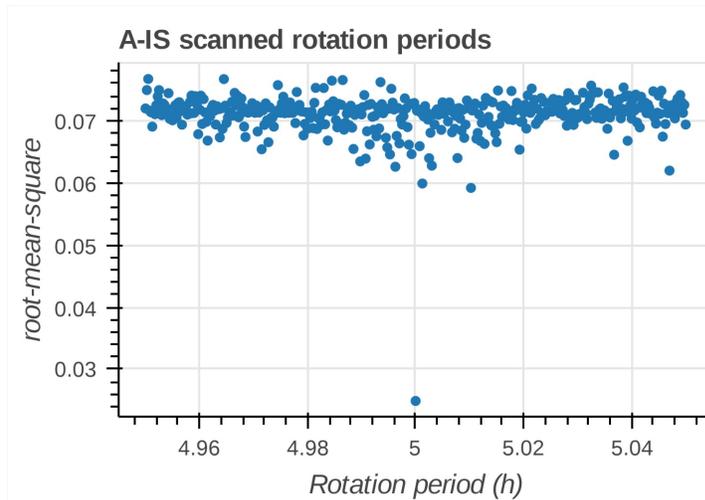
Figure 3: Period scanning solution from A-IS as RMS of the fit vs scanned period.

to the program. Results are provided as two plots, one of which shows the probability of the asteroid to belong to one of the taxonomic classes and the other showing the modelled absorption coefficient spectrum (see Fig. 7). Results are also provided in numerical form as text files.

6. Validation

Since there was no Gaia asteroid data available for the validation of the AVI implementations, with the exception of the Gaia image of the week (2017-04-24) spectra of four asteroids, we use synthetic data to test the AVIs. In this section, we present the algorithm validation procedure. Technical testing of the AVI implementation was carried out following standard procedures, but presenting the results is out of the scope of this paper.

6.1. GAVITEA

We tested the applicability of the GAVITEA algorithm to different input data as well as different input parameters. Observing geometries and epochs for the simulated test data were computed using the Gaia simulator developed at the Observatoire de la Cote d'Azur.

22
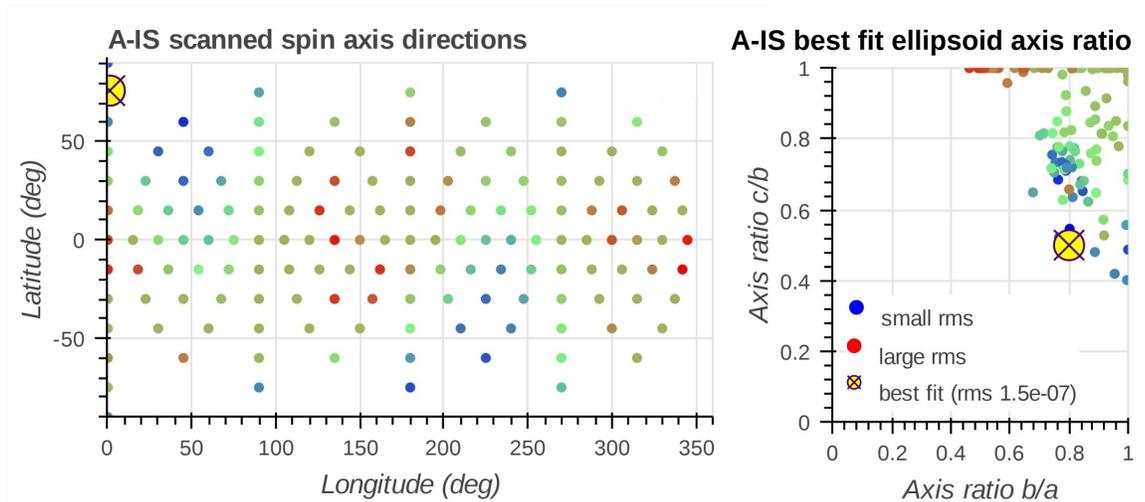
Figure 4: Pole scanning solutions from A-IS as spin axis longitude and latitude and the corresponding best fit ellipsoid axis ratios colour coded with RMS of the fit.

6.1.1. Algorithm validation using different data sets

For testing the robustness of each algorithm considering different kinds of targets and observing geometries, we generated 22 sparse Gaia-like photometric data sets. Twelve photometric sequences were generated using nonconvex shape models representing the known shapes of (951) Gaspra, (433) Eros and (216) Kleopatra, and ten for a few different ellipsoidal shapes. Different rotation periods and pole orientations were assigned to the targets. Random noise was added to the apparent magnitudes but other observation-related quantities (timing and geometry) were considered to be accurately known. Input parameters for the validation tests are given in Table 1, on the first row of column "Values for validation".

All the test cases for ellipsoid-model data produced rotation period within 1% of the correct one, spin axis direction within 5 degrees of the correct one and ellipsoid axis ratios within 5% of the simulated ellipsoid model.

All the test cases for A-IS and A-ES using nonconvex model data produced rotation periods within 5% of the correct one, spin axes within 10 degrees of the correct one and ellipsoid axis ratios within 40% of the dimension ratios of

23
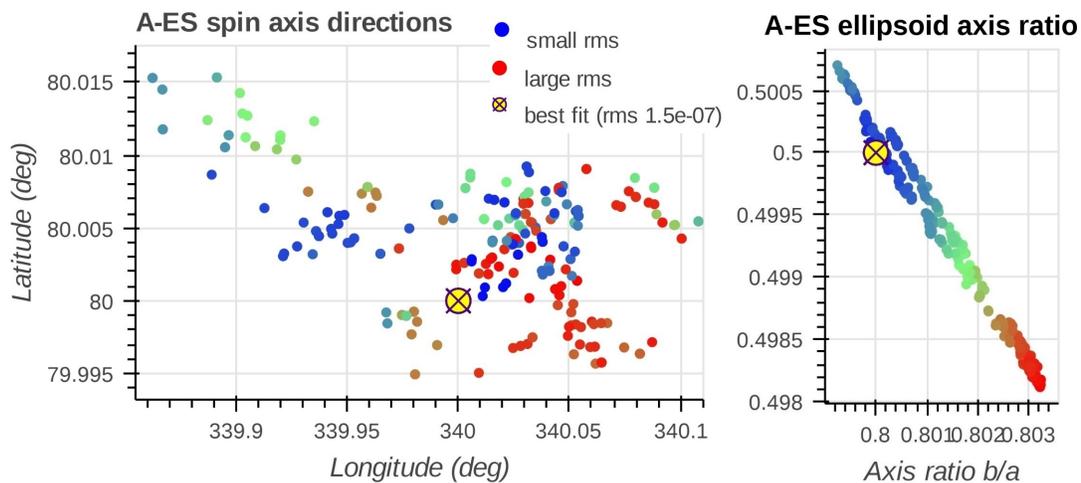
Figure 5: Spin axis for the accepted solutions from A-ES or A-CS and ellipsoid axis ratios for the accepted solutions from A-ES color coded with RMS of the fit.

the simulated nonconvex model.

All the test cases for A-CS using nonconvex model data produced rotation periods within 1% of the correct one, spin axes within 5 degrees of the correct one and axis dimensions within 5% of the dimensions of the simulated nonconvex model.

Figure 6: Shape solution from A-CS.

6.1.2. Validation using different input parameters

The effect of input parameters to the result and computation time was tested by performing multiple runs with varying input parameters for a single data set. Input parameters for the validation tests are given in Table 1, on the second row in the column "Values for validation".

A-IS

All the test cases for A-IS using ellipsoid model data produced rotation periods within 1% of the correct one, spin axes within 5 degrees of the correct one and ellipsoid axis ratios within 5% of the simulated ellipsoid model. All the test cases using nonconvex model data produce rotation periods within 5% of the correct one, spin axes within 10 degrees of the correct one and ellipsoid axis ratios within 40% of the dimension ratios of the simulated nonconvex model.

A-ES and A-CS

The validation results showed that A-ES and A-CS are sensitive to input parameters and not all the input parameter combinations provided the correct solution. The algorithm should thus be run using a number of different input parameter combinations to check which solution is the most stable one.

6.2. GAVISC

We validated the GAVISC algorithms using three Gaia-observed spectra, as well as with the leave-one-out procedure with the original B-DM data set (DeMeo et al., 2009).

6.2.1. Spectral classification SPEC-C

From the original B-DM data set, we will one-by-one leave each asteroid out from the training set, and build the classifier with the rest. Then, we will classify the asteroid left out from the training set, and record the result. We will do this for all the asteroids except for the one O- and the one R-type targets, since the classifier will need at least one asteroid from each class in the training set (see Table 2 for the classes and the number of objects in them). The results from this cross-validation are shown in Table 3.

25

Table 2: The 13 taxonomic classes or complexes from the Bus-DeMeo original data set and the number of objects in the classes, used here to validate the GAVISC SPEC-C algorithm. The complexes C, S, and X are marked with bold font to distinct them from single taxonomic classes.

Taxonomy	A	B	**C**	D	K	L	O	Q	R	**S**	T	V	**X**
#	6	4	45	16	16	22	1	8	1	199	4	17	32

The results of the abovementioned validation test show very good performance. In general, all the classes are found reasonably well except for the B-class where the success rate is only 25 %. Also, 16 % of the X-complex asteroids are misclassified as C-complex asteroids.

The Gaia asteroid spectral data is not published yet, and will most probably be published only in the final Gaia data release. However, the Gaia DPAC published preliminary spectra of four targets in the ESA Gaia 'Image of the Week' web site (Galluccio et al., 2017). We were able to get the numerical data from the authors for three of these — asteroids (19) Fortuna, (21) Lutetia, and (279) Thule. We ran the SPEC-C algorithm on these targets, and the results are shown in Table 4.

The results of the tests are quite promising. Asteroid (19) Fortuna has a classification G (Tholen) and Ch (B-DM) (Neese, 2010), and our algorithm suggests the C complex with 91% probability. Asteroid (21) Lutetia is classified as M (Tholen) and Xc (B-DM). Our algorithm gives the highest probability to the K class (73%). We would like to note, that the Gaia data actually has a subtle absorption band close to about 0.95 µm. The K class has an absorption band at ∼1 µm, while Xc has none. In that respect, the K-type classification with the Gaia data is understandable. Finally, (279) Thule should be a D-type (Tholen and B-DM), and is also classified as such by our SPEC-C with the probability of 97%.

26

Table 3: The confusion matrix from the leave-one-out cross-validation of the SPEC-C algorithm. In each row the asteroids from the corresponding class are classified, and the fractions of classes they are classified to are shown in the columns. For the classes O and R the results are not valid, since their is only one target of the type in the class, and thus they cannot be cross-validated.

		Fraction classified to, in %												
		A	B	**C**	D	K	L	O	Q	R	**S**	T	V	**X**
	A	83	0	0	0	0	17	0	0	0	0	0	0	0
	B	0	25	50	25	0	0	0	0	0	0	0	0	0
	C	0	0	91	0	0	0	0	0	0	0	0	0	9
	D	0	0	0	88	0	0	0	0	0	0	0	0	12
Original taxonomy	K	0	0	0	0	88	0	0	0	0	6	0	0	6
	L	0	0	0	5	0	68	0	0	0	14	0	0	14
	O	0	0	0	0	0	0	100	0	0	0	0	0	0
	Q	0	0	0	0	0	0	0	75	0	25	0	0	0
	R	0	0	0	0	0	0	0	0	100	0	0	0	0
	S	0	0	1	0	0	1	0	1	0	97	0	0	0
	T	0	0	0	0	0	0	0	0	0	0	75	0	25
	V	0	0	0	0	0	0	0	0	0	0	0	100	0
	X	0	0	16	3	6	3	0	0	0	0	3	0	69

6.2.2. Spectral absorption coefficient SPEC-AC

The validation of the SPEC-AC algorithm is quite difficult, since the absorption spectra of the asteroid's regolith material is commonly not known. Furthermore, the inversion result depends on the assumption of the grain size, for which the SPEC-AC uses a size of 200 µm. Since the surface-area-to-volume ratio of the material will affect the general mean albedo level, and the size parameter (i.e., size-wavelength ratio) to the slope of the absorption spectra, the size assumption will have an effect on the result. Despite the evident problems on the inversion of the absorption spectra, we consider it useful since similar tools do not, to our knowledge, exist, and even a preliminary modeling of the

27

Table 4: The results of the SPEC-C algorithm for three asteroids observed by Gaia.

	C	D	K	L	**X**
(19) Fortuna	**0.91**				0.09
(21) Lutetia			**0.73**	0.04	0.14
(279) Thule		0.97			0.02

absorption spectra could be useful.

We show the SPEC-AC results for asteroid (19) Fortuna using the published Gaia Image of the Week -data, see Fig. 7. The fit to the reflectance is more or less perfect, which is typical since this fit is what the algorithm is optimizing. The actual result, the absorption coefficient $k = \mathrm{Im}(m)$ behavior as a function of the wavelength, is shown in the lower sub-panel. In principle, as the reflectance increase, the k should decrease. On top of that, there is a particle-size effect where particle sizes (here, 200 µm) close to the wavelength will be efficient in both scattering and absorption, resulting in the positive slope in the k-spectra with wavelengths from 0.5 µm even though the reflectance spectra is quite constant.

7. Conclusions

We have constructed the first Gaia Added-Value Interfaces (Gaia AVIs) to asteroid data eventually available in the Gaia database. The Gaia AVI for temporal analysis (GAVITEA) solves the inverse problem of deriving the shape and spin state of an asteroid from photometric observations whereas the Gaia AVI for spectral classification (GAVISC) relies on spectral data to classify asteroids based on the scattering properties of their surface material and to derive their absorption spectra. The successful implementation of these two AVIs shows that the Gaia Added Value Interface Platform (GAVIP) provides a framework to build science applications upon. We hope that the users of Gaia data will see the benefit of first enabling their own science by developing their own AVI, and eventually sharing that AVI and a published description of it (such as a
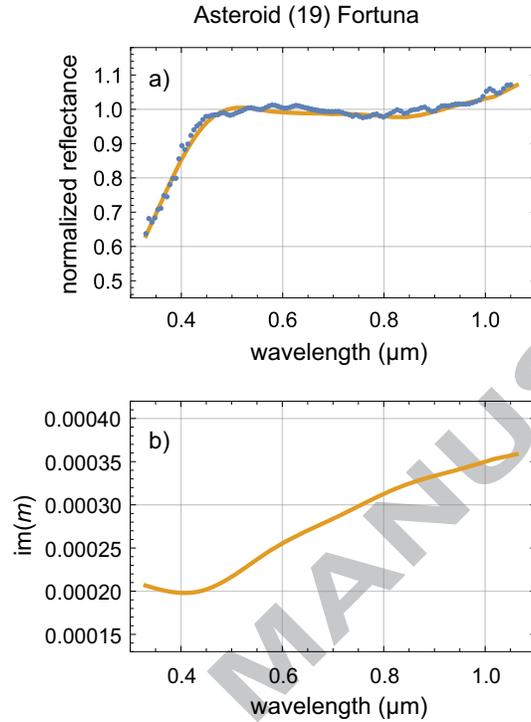
Figure 7: The reflectance spectra (in the upper panel a) and the absorption coefficient spectra (in the lower panel b) as a function of the wavelength for asteroid (19) Fortuna. In a), the blue dots are the Gaia observations, and the orange solid line is the fit. In b), the orange solid line is the inverted k-spectra of the asteroid's surface material.

Methods section in a journal paper) with the rest of the community.

Although the main reason for implementing GAVITEA and GAVISC was to guide the development of GAVIP, they do provide highly useful tools for analysing Gaia asteroid data. A few issues, however, remain to be solved for in further development of the algorithms and, hopefully, will be implemented in Gaia AVI updates.

First of all, at the time of writing this paper, there is no interface between the AVIs and the Gaia archive since there is yet no asteroid data available in the Gaia archive and the exact form of the data catalogue is still unclear. At Gaia DR2 (April 25th, 2018), asteroid data will be published making it possible

29

to construct the interface. Only application to the real Gaia asteroid data will reveal the true potential of GAVITEA and GAVISC.

Algorithm test results of A-CS and A-ES show that the results are somewhat sensitive to input parameters. The algorithms should be more extensively tested using various kinds of data sets, especially real observations, to be able to provide guidelines for AVI users for specifying suitable input parameters. Also a general estimate for the minimum required number of data points can be done based on more extensive use of GAVITEA. In any case, the user should test some number of input parameters to check whether the solution is stable.

When using GAVISC, it should be remembered that Gaia wavelengths do not span the range of Bus-DeMeo wavelength range, and some absorption features are not included in the analysis.

GAVISC and GAVITEA can be accessed via GAVIP Portal at http://gavip.esac.esa.int, where the user can register as a GAVIP user and find all the relevant information for using its functions. User manuals for GAVISC and GAVITEA specifically are available in each AVI. As long as real Gaia data is not available, GAVITEA and GAVISC can be tested with simulated data.

In terms of further development of other asteroid-specific Gaia AVIs we foresee interesting opportunities. Various processes — such as asteroid-asteroid collisions and spin-up due to thermal torques — can lead to a variation in the material properties across an asteroid's surface. These variations are valuable in that they allow direct information to be gained about, for example, asteroid interior composition and, partly, structure. Indirectly, the variations can be used to set constraints on the collision rates and the importance of thermal torques. We plan on implementing the so-called Gaia AVI for Temporal Spectral Classification (GAVITESC) prior to the publication of Gaia BP and RP data on asteroids. It would allow the user to search for spectral variation across asteroid surfaces using Gaia data by combining results from GAVISC and GAVITEA.

Acknowledgements

References

Bus, S., 1999. Compositional structure in the asteroid belt: Results of a spectroscopic survey. Ph.D. thesis, Massachusetts Institute of Technology.

Cellino, A., Muinonen, K., Hestroffer, D., Carbognani, A., 2015. Inversion of sparse photometric data of asteroid using triaxial ellipsoid shape models and a lommel-seeliger scattering law. Planetary and Space Science 118, 221–226.

Delbo, M., Gayon-Markt, J., Busso, G., Brown, A., Galluccio, L., Ordenovic, C., Bendjoya, P., Tanga, P., 2012. Asteroid spectroscopy with gaia. Planetary and Space Sciences 73 (1), 86–94.

DeMeo, F., Binzel, R., Slivan, S., Bus, S., 2009. An extension of the Bus asteroid taxonomy into the near-infrared. Icarus 202 (1), 160–180.

Ďurech, J., Sidorin, V., Kaasalainen, M., 2010. DAMIT: a database of asteroid models. Astronomy and Astrophysics 513, A46.

Dutton, G., 1999. Computational aspects of a quaternary triangular mesh. In: A hierarchical coordinate system for geoprocessing and cartography. Springer, Berlin, Heidelberg, pp. 41–70.

Fielding, R., 2000. Architectural styles and the design of network-based software architectures. Ph.D. thesis, University of California.

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Mignard, F., Drimmel, R., Babusiaux, C., Bailer-Jones, C. A. L.,

Bastian, U., et al., Nov. 2016a. Gaia Data Release 1. Summary of the astrometric, photometric, and survey properties. Astronomy & Astrophysics 595, A2.

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., Brown, A. G. A., Vallenari, A., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., Biermann, M., Evans, D. W., et al., Nov. 2016b. The Gaia mission. Astronomy & Astrophysics 595, A1.

Galluccio, L., DeAngeli, F., Delbo, M., Mignard, F., 2017. Image of the Week — Gaia reveals the composition of asteroids. https://www.cosmos.esa.int/web/gaia/iow_20170424, [Online; accessed 18-December-2017].

van de Hulst, H.C., 1957. Light scattering by small particles. Wiley, New York.

Jordi, C., Gebran, M., Carrasco, J. M., de Bruijne, J., Voss, H., Fabricius, C., Knude, J., Vallenari, A., Kohley, R., Mora, A., Nov. 2010. Gaia broad band photometry. A&A 523, A48.

Kaasalainen, M., Lamberg, L., Lumme, K., Bowell, T., 1992. Interpretation of lightcurves of atmosphereless bodies. i. general theory and new inversion schemes. Astronomy and Astrophysics 259, 318–32.

Kaasalainen, M., Torppa, J., 2001. Optimization methods for asteroid lightcurve inversion. i. shape determination. Icarus 153, 24.

Lindqvist, H., Martikainen, J., Räbinä, J., Penttilä, A., Muinonen, K., 2018. Ray optics in absorbing media with application to ice crystals at near-infrared wavelengths. Submitted to Journal of Quantitative Spectroscopy & Radiative Transfer.

Mainzer, A., Bauer, J., Cutri, R., Grav, T., Kramer, E., Masiero, J., Nugent, C., Sonnett, S., Stevenson, R., Wright, E., 2016. NEOWISE Diameters and Albedos V1.0. EAR-A-COMPIL-5-NEOWISEDIAM-V1.0. NASA Planetary Data System.

32

Mainzer, A., Grav, T., Bauer, J., Masiero, J., McMillan, R. S., Cutri, R. M., Walker, R., Wright, E., Eisenhardt, P., Tholen, D. J., Spahr, T., Jedicke, R., Denneau, L., DeBaun, E., Elsbury, D., Gautier, T., Gomillion, S., Hand, E., Mo, W., Watkins, J., Wilkins, A., Bryngelson, G. L., Molina, A. D. P., Desai, S., Camus, M. G., Hidalgo, S. L., Konstantopoulos, I., Larsen, J. A., Maleszewski, C., Malkan, M. A., Mauduit, J.-C., Mullan, B. L., Olszewski, E. W., Pforr, J., Saro, A., Scotti, J. V., Wasserman, L. H., 2011. Neowise observations of near-earth objects: Preliminary results. The Astrophysical Journal 743 (2), 156.

Mainzer, A., Grav, T., Masiero, J., Bauer, J., Cutri, R. M., McMillan, R. S., Nugent, C. R., Tholen, D., Walker, R., Wright, E. L., 2012. Physical parameters of asteroids estimated from the wise 3-band data and neowise post-cryogenic survey. The Astrophysical Journal Letters 760 (1), L12.

Martikainen, J., Penttilä, A., Gritsevich, M., Lindqvist, H., Muinonen, K., 2018. Spectral modeling of meteorites at uv-vis-nir wavelengths. Journal of Quantitative Spectroscopy & Radiative Transfer 204, 144–151.

Masiero, J. R., Grav, T., Mainzer, A. K., Nugent, C. R., Bauer, J. M., Stevenson, R., Sonnett, S., 2014. Main-belt asteroids with wise/neowise: Near-infrared albedos. The Astrophysical Journal 791 (2), 121.

Masiero, J. R., Nugent, C., Mainzer, A. K., Wright, E. L., Bauer, J. M., Cutri, R. M., Grav, T., Kramer, E., Sonnett, S., 2017. Neowise reactivation mission year three: Asteroid diameters and albedos. The Astronomical Journal 154 (4), 168.

Mignard, F., Cellino, A., Muinonen, K., Tanga, P., Delbó, M., Dell'Oro, A., Granvik, M., Hestroffer, D., Mouret, S., Thuillot, W., Virtanen, J., 2007. The Gaia mission: Expected applications to asteroid science. EM&P 101 (3–4), 97–125.

Minkowski, H., 1903. Volumen und oberfläche. Math. Annal. 57, 447–495.

Muinonen, K., Belskaya, I. N., Cellino, A., Delbò, M., Levasseur-Regourd, A.-C., Penttilä, A., Tedesco, E. F., Oct. 2010. A three-parameter magnitude phase function for asteroids. Icarus 209, 542–555.

Muinonen, K., Lumme, K., 2015. Disk-integrated brightness of a Lommel-Seeliger scattering ellipsoidal asteroid. A&A 584, A23.

Muinonen, K., Nousiainen, T., Lindqvist, H., Munoz, O., Videen, G., 2009. Light scattering by gaussian particles with internal inclusions and roughened surfaces using ray optics. Journal of Quantitative Spectroscopy & Radiative Transfer 110, 1628–1639.

Muinonen, K., Torppa, J., Wang, X., Cellino, A., 2018. Asteroid lightcurve inversion with bayesian inference. A&A in preparation.

Neese, C., 2010. Asteroid Taxonomy V6.0. EAR-A-5-DDR-TAXONOMY-V6.0. NASA Planetary Data System.

Santana-Ros, T., Bartczak, P., Michalowski, T., Tanga, P., Cellino, A., 2015. Testing the inversion of asteroids' gaia photometry combined with ground-based observations. MNRAS 450 (1), 333–341.

Tanga, P., Mignard, F., 2012. The solar system as seen by gaia: The asteroids and their accuracy budget. Planetary and Space Science 73 (1), 5–9.

Tanga, P., Mignard, F., Dell'Oro, A., Muinonen, K., Pauwels, T., Thuillot, W., Berthier, J., Cellino, A., Hestroffer, D., Petit, J.-M., Carry, B., David, P., Delbo, M., Fedorets, G., Galluccio, L., Granvik, M., Ordenovic, C., Pentikäinen, H., 2016. The daily processing of asteroid observations by gaia. Planetary and Space Science 123, 87–94.

Tholen, D., 1984. Asteroid taxonomy from cluster analysis of photometry. Ph.D. thesis, The university of Arizona.

Vagg, D., O'Callaghan, D., Ó hÓgáin, F., McBreen, S., Hanlon, L., Lynn, D., O'Mullane, W., 2016. GAVIP: a platform for Gaia data analysis. In: Chiozzi,

34

G., Guzman, J. C. (Eds.), Software and Cyberinfrastructure for Astronomy IV. Vol. 9913 of SPIE Astronomical telescopes + instrumentation. p. 19.