



# Deliverable



## H2020 COMPET-05-2015 project ”Small Bodies: Near And Far (SBNAF)”

**Topic:** COMPET-05-2015 - Scientific exploitation of astrophysics, comets, and planetary data

**Project Title:** Small Bodies Near and Far (SBNAF)

**Proposal No:** 687378 - SBNAF - RIA

**Duration:** Apr 1, 2016 - Mar 31, 2019

<b>WP</b>	WP6, Synergies from ground and space
<b>Del.No</b>	D6.7
<b>Title</b>	Quality assessment system for the models
<b>Lead Beneficiary</b>	UAM
<b>Nature</b>	Report
<b>Dissemination Level</b>	Public
<b>Est. Del. Date</b>	31 March 2018
<b>Version</b>	1.0
<b>Date</b>	31 March 2017
<b>Lead Author</b>	Grzegorz Dudziński, g.dudzinski@amu.edu.pl

### WP6, Synergies from ground and space

**Objectives:** To combine observational data from space and ground, from remote disk-integrated data and disk-resolved data from interplanetary missions to obtain (validated) high-quality model solutions for a wide range of applications: improvement of the scientific understanding, answering key questions for the reconstruction of minor body properties, calibration aspects, support for Gaia density determination, Hayabusa-2 target characterization and operational support, tools and methods for applications to large object samples.

## Description of deliverable

D6.7: Quality assessment system for the models [24, public, report]

– Established (and applied) quality system for all our SBNAF sample targets

### 1 Introduction

In this deliverable we summarise the contents of a publication ([Bartczak and Dudziński, in prep.](#)) where we propose a procedure to assess the quality and uncertainty of asteroid models.

The term *asteroid model*, or just *model*, denotes a set of parameters describing the shape, spin axis orientation and scattering law used to determine the amount of light reflected by the surface. The last two sets of parameters are explicitly defined, but shape parameters are not unique – one can describe the same surface using different notations and numbers (e.g. voxels, spherical harmonics, etc). In this work we always refer to the shapes or surfaces as composed of triangular facets defined on a mesh of vertices in 3D space.

Measurement uncertainty and the error propagation are inseparable parts of physical studies and of any scientific inquiry. Any model that tries to explain the observed phenomena without proper uncertainty assessment is worthless as we are unable to judge the quality of the solution and put constraints on properties calculated based on that model. One can judge the quality of a model by referring to the fit to the observational data ( $\chi^2$  test or *RMSD* values), but as the models get more complex and sophisticated trying to explain multiple entangled phenomena at once, one number proves to be insufficient to test their robustness – especially when the quality, type and abundance of observations can influence the solution, as is definitely the case in the field of asteroid shape modelling.

Determining an asteroid model is an inverse problem and those are usually undeterministic, i.e. many different solutions can explain the same data at comparable level. It is due to the fact that the number of parameters describing the model can be very high compared to the amount of information contained in the observational data. Additionally, the model uncertainty arises due to plethora of inescapable factors like detector noise, human error, non-uniform data coverage and observational biases, assumptions (e.g. homogeneous albedo and mass distribution, principal axis rotation), simplifications of underlying physics (e.g. using simplistic light scattering law), numerical rounding errors and approximations, etc. The list has to be completed with the assumptions, simplifications, algorithms and bugs (which are an intrinsic part of any computer code) connected with particular modelling technique that constitute additional sources of uncertainty.

It is very tempting to evaluate an asteroid model's quality from a set of observations used in the modelling process. There is no doubt that the quality and quantity (e.g. signal to noise ratio, number of lightcurves, number of points per lightcurve, etc.) have a major influence on the solution as well as does the distribution of apparitions (entangled with the spin axis orientation) that define geometries in which an asteroid was observed. Having little to do with the uncertainty – thus quality – of the models, judgement of the dataset can only be used to foresee a potential success of finding unique solution.

As of today, asteroid shape models are published without any information about the uncertainty (with the exception of some of the 3-axial ellipsoid models) The best we can hope for is the family of solutions acquired from many runs of the modelling. Analysis of such a family can give some insight into the robustness of the model, particularly when it comes to the spin axis orientation. Nonetheless, this approach is hard to standardize and does not yield the results that would allow for inter-method model comparison. Uniform scanning of a parameter space near the best solution is also not utilized here.

The volume of an asteroid shape model, when combined with the size and mass estimates, is essential in the density determination – one of the most important physical parameters that allow us to peek inside asteroids and make conclusions about their inner structure. The shape model uncertainty propagates to volume, and analogously to any other property determined from the model, affecting every conclusion drawn for a population of small bodies under scrutiny. There definitely exists a compelling need for a procedure of quality

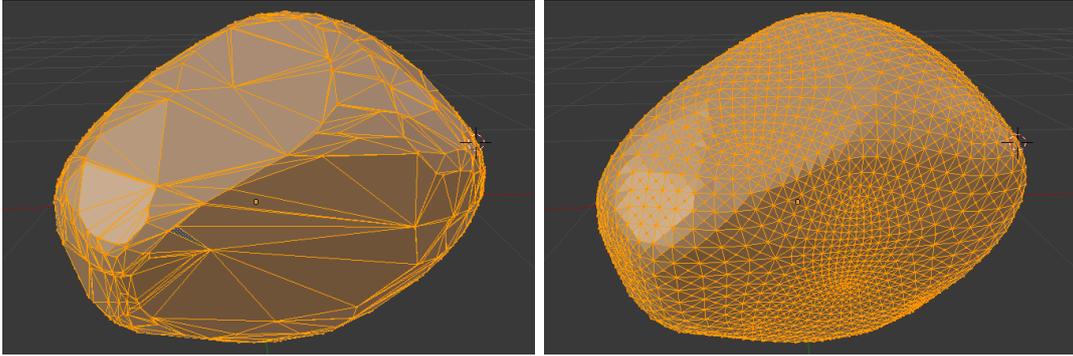


Figure 1: Example shape before (left) and after (right) the remeshing operation.

and uncertainty assessment of asteroid models.

## 2 Requirements

We require the uncertainty assessment to be modelling-technique independent. The uncertainty has to be expressed in numerical values, convenient for use and visualisation. Large majority of available asteroid models are based on the relative photometry alone and for that reason our focus on this type of data. Extension of the method taking other types of observations into account (e.g. delay-Doppler images, stellar occultations, adaptive optics imaging) can be addressed at the later stage building upon single-data-type scheme.

The main idea behind uncertainty assessment is to create clones of a nominal model introducing changes in its parameters and accepting or rejecting them based on the fit to the observations. Resulting parameter's uncertainties are derived from accepted clones' statistics. We test the clones only against the observations that were used during the modelling process; the same holds for the scattering law.

## 3 Uncertainty assessment procedure

### 3.1 Remesh

When dealing with the models from different methods the need for unifying the shape representations arises. Making small changes in models' parameters is the key component of the method described here. To assure the uniformity in model's shape changes (guaranteeing good statistics), vertices need to be evenly distributed. That also ensures consistent triangle sizes and the lack of spike-like or huge ones. The resulting shape stays the same (and produces the same lightcurves, which is of our main concern here), but is much easier to work with (Fig. 1). The new mesh is created by calculating the intersection points between model's surface and equally distributed set of rays originating in the model's center. This simple scheme is sufficient for most asteroid shapes although some modifications would need to be done for more complex ones, e.g. binary asteroids. Model's reference frame is left intact, i.e. we do not recalculate its center nor apply any rotations. A remeshed model is denoted as a *reference* or *nominal* model.

### 3.2 Confidence level of the nominal model

Before the cloning of the nominal model can begin a confidence level of the nominal model needs to be established. Later, basing on it, the decision to accept or reject a clone is made. For a set of lightcurves from which a model was derived during a modelling process, a root-mean-square deviation is calculated

$$RMSD_{ref} = \sqrt{\frac{\sum_i^N (O_i - C_i)^2}{N}} \quad (1)$$

where  $O_i$  and  $C_i$  denote the observed and corresponding synthetic photometric points and  $N$  denotes the total number of them. Next, a standard error of  $RMSD_{ref}$  distribution in a form:

$$\mathcal{E} = \frac{RMSD_{ref}}{\sqrt{N - n}} \quad (2)$$

with  $n$  degrees of freedom (i.e. number of parameters describing the model) is used to establish a test for clones. Accepted ones need to satisfy the following equation:

$$RMSD_{clone} \leq RMSD_{ref} + \mathcal{E}. \quad (3)$$

In summary, we are interested in clones that have the same level of goodness of the fit to the lightcurves lying within a confidence level established by  $\mathcal{E}$ .

### 3.3 Cloning the shape

There are many ways to represent and change the 3D shapes and surfaces. We decided to work with meshes of vertices and triangular faces defined on them – a representation commonly used in computer graphics applications. To ensure that vertices positions and triangle sizes distribution are as uniform as possible we start with the largest platonic solid consisting of 12 vertices and 20 faces and then use [Catmull and Clark \(1978\)](#) surface subdivision algorithm to get denser meshes. After 2 iterations we get 242 vertices, and 3842 after 2 more. The latter mesh size is detailed enough to represent complex shapes and is suitable for asteroid models, whereas the former offers sufficient number of parameters to introduce meaningful changes while not being computationally unfeasible. By varying 242 parameters and then subdividing the surface we can change the resulting more detailed shape in a convenient way.

For greater flexibility we also control generated hills and concavities by manipulating congruent vertices. If  $\alpha$  is an angle between two vertices, then the formula

$$\exp(-\alpha/\alpha_{max})^{2k}, \quad (4)$$

where  $\alpha_{max} \in [0^\circ, 90^\circ]$  and  $k \in \{1, 2, 3\}$  are random numbers, will generate hills of different width and flatness when applied to vertex distance from the model's centre.

To demonstrate a cloning procedure let us consider a spherical shape and no lightcurve criteria.

Changes to a vertex position are generated randomly and should have uniform distribution. Because vertices are not independent from each other, special care has to be taken to achieve that. About  $2 \times 10^6$  clones are generated and in this population a certain magnitude of a vertex position change appears fixed number of times. With this criterion, when

analysing the probability of a vertex having given amplitude of change in length we get the same distribution for each vertex. In other words, the variances of probability density distributions for a vertex position changes differ by only 0.1%.

The population of vertex changes described above needs to be created only once. The uniformity of the population ensures that no biases will be present when applied to models of asteroids, which is done by moving vertices of a model along surface normal direction accordingly. Now we also introduce the lightcurve criterion. For each of  $1.4 \times 10^6$  clones the synthetic lightcurves are created doubling the photometric observations for calculating *RMSD* of the fit. Next, only the clones which satisfy eq. 3 are selected producing population of accepted clones  $\mathcal{A}$ .

A reference model's  $i$ -th vertex serves as an expected value  $\mu_i$ . From  $i$ -th vertex positions from all of the accepted clones the probability density distribution (PDD) is created. The probability is normalized so that

$$\sum_{j=1}^{\bar{\mathcal{A}}} p_j = 1,$$

where  $p_j$  is  $j$ -th probability bin for  $i$ -th vertex and  $\bar{\mathcal{A}}$  is the cardinality of the set of accepted clones  $\mathcal{A}$ .

The uncertainty of a vertex position is assessed from PDD, which is expected to be non-symmetrical. Therefore we analyse the changes of vertex position above and below the nominal value separately. This is justified by the fact that hills and concavities have different influence on the lightcurves. The variance of PDD is considered as the vertex uncertainty and can be expressed formally as follows:

$$u_i^{\pm} = \sum_j^{\bar{\mathcal{A}}} (x_j - \mu_i)^2 p_j, \quad (5)$$

where  $x_j$  is a position of  $i$ -th vertex in  $j$ -th clone,  $\mu_i$  is a reference model's  $i$ -th vertex position and  $p_j$  is a probability. The '+' and '-' upper index sections PDD above and below  $\mu_i$  value.

### 3.4 Pole position

Anticipated enormous computation time needed to test the pole solution for all of the clones combining the effects of shape and pole uncertainty would make the presented method unusable in practice. That said, we check the spin axis orientation only for the reference shape. The effects of shape, pole and rotational period uncertainties can be analysed separately or be combined later on.

To test spin axis solution we perform a scan of different longitude and latitude coordinates  $\lambda, \beta$  covering a hemisphere with nominal solution at its pole.

### 3.5 Rotational period

Periodogram for the nominal model is made by scanning the proximity of nominal period. Period discrepancy at the level

$$\varepsilon = \frac{RMSD_{\text{best}}}{\sqrt{N - n}}, \quad (6)$$

where  $N$  is the number of photometric data points and  $n$  is the number of degrees of freedom of the model, is interpreted as the rotational period uncertainty.

Also, we keep track of the periods for each clone (which changes when surface is modified because the period is fitted to minimise the RMSD of synthetic and observed lightcurves). If the dispersion of rotational periods from  $\mathcal{A}$  is larger than the uncertainty indicated by nominal model periodogram, the former is used.

## 4 Uncertainty propagation

### 4.1 Volume

Asteroid shape models from lightcurve inversion are dimensionless. By default, the length of the longest vector of a model  $R_{max}$  is 1. Various scaling methods (e.g. fit to stellar occultation chords, thermophysical modelling) provide the scale  $S$  by which every vertex of the model can be multiplied to create a model with independent dimensions expressed in physical units. Scale, in that sense, can be understood as  $R_{max}$ .

The volume of the scaled model can be expressed as

$$V_m = V' S^3 \pm u(V_m), \quad (7)$$

where  $V'$  is a dimensionless volume of unscaled model and  $u(V_m)$  is volume uncertainty. Uncertainties for the scale  $S$  and volume  $V'$  have different origins: the former comes from a scaling method while the latter from the model's clones. The resulting uncertainty of the volume  $u(V_m)$  carries both uncertainties and, assuming they are uncorrelated, the following equation is used:

$$u(V_m) = \sqrt{\left(\frac{\partial V_m}{\partial S}\right)^2 u^2(S) + \left(\frac{\partial V_m}{\partial V'}\right)^2 u^2(V')} = \sqrt{9S^4 V'^2 u^2(S) + S^6 u^2(V')}. \quad (8)$$

where partial derivatives serve as sensitivity coefficients describing how the volume uncertainty is dependent on each component (Taylor, 1997).

### 4.2 Phase of rotation

Rotational period is calculated from the model's synthetic lightcurves fit to observations and its uncertainty propagates to rotational phase. The rotational phase uncertainty should stay on the same level inside the interval between the first and the last observation because the period was established using all the available lightcurves at once. By going outside this time interval rotational phase uncertainty increases linearly with time from the last observation. To achieve that in one formula we put the reference epoch of the model right in the middle of observing time span, i.e.  $JD_0 = (JD_{last} + JD_{first})/2$ , and then calculate

$$u(\gamma) = \frac{2\pi}{P^2} \left[ |u(P)\Delta t| + \frac{1}{2}u(P)\Delta T_{obs} + \max(-|u(P)\Delta t| + \frac{1}{2}u(P)\Delta T_{obs}, 0) \right] + u(\gamma_0) \quad (9)$$

where  $P$  is a rotational period,  $\gamma$  is a rotational phase,  $\gamma_0$  is a phase for reference epoch  $JD_0$ ,  $\Delta T_{obs} = JD_{last} - JD_{first}$  is an observations' time span,  $\Delta t = t - JD_0$  is a distance

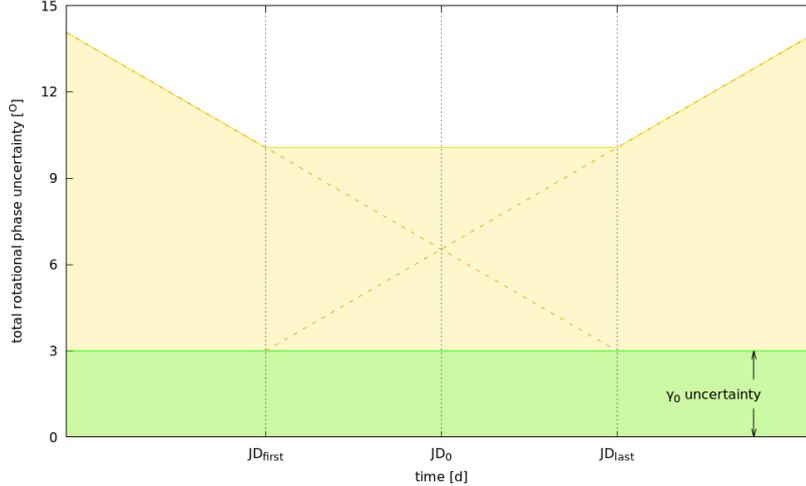


Figure 2: The plot of total rotation phase uncertainty. In this example  $\Delta T_{obs} = 38.5$  years, the uncertainty of rotation phase for reference epoch  $JD_0$  is  $u(\gamma_0) = 3^\circ$  and the period uncertainty  $u(P) = 1.5 \times 10^{-6} h$ .

in time from the reference epoch, and  $u(x)$  denotes uncertainty of  $x$ . The *max* function returns larger of the two arguments. An example plot of this function can be seen in Fig. 2.

The  $\gamma_0$  uncertainty (i.e. uncertainty of the rotational phase angle at  $JD_0$ ) is also connected with period uncertainty and data points density in lightcurves.  $u(\gamma_0)$  comes from the discrepancy of  $\gamma_0$  among the clones population. For each clone the period and  $\gamma_0$  are calculated separately. The nature of  $\gamma_0$  uncertainty is however different from the rotational phase uncertainty mainly because it does not accumulate with time. These values have to be added together in order to facilitate total rotational phase uncertainty.

### 4.3 Scaling the model based on stellar occultations

Stellar occultation observations, when projected on the Earth's surface, can be used as an almost-direct measurement of asteroid silhouette size. The right observer location, precise timing measurements, a relatively good model of the Earth surface and precise orbit of occulting object are needed. Fitting the model to observations is done by projecting the models' silhouette on Earth surface for the time of occultation event. Then, points on the edge of the model's silhouette are compared with data points from occultation using weighted RMSD value

$$RMSD = \sqrt{\frac{1}{\sum w_i} \frac{\sum_i^N w_i (Sx_i - y_i)^2}{N}}, \quad (10)$$

where  $x_i$ ,  $y_i$  and  $w_i$  represent the model's vertex position, the occultation event position and the vertex position weight (established based on position uncertainty from the model's clones or occultation timing error) of  $i$ -th vertex along corresponding occultation chord;  $S$  is a scale of the model and  $N$  is the number of occultation data points.

The scale  $S$ , which minimises the RMSD, has to be found. To evaluate the goodness of the fit, i.e. give the value of the scale uncertainty, the RMSD vs  $S$  space need to be explored. The scale uncertainty  $u(S)$  is calculated as the range of  $S$  values at the level

$$\varepsilon = \frac{RMSD_{\text{best}}}{\sqrt{N-1}}. \quad (11)$$

The number of degrees of freedom is set to 1 as we fit the scale alone. Values obtained from this method enter Eq. 8 to account for total volume uncertainty of the model.

## 5 Example

The test body was a modified ellipsoid with  $a/b = 1.5$  and  $b/c = 1.14$  and two craters, but the lightcurves that were used in the uncertainty assessment were of the unmodified ellipsoid. We wanted to test whether the craters – treated as fictitious features of the body – will be detected as the surface areas with large uncertainty. Moreover, gaps in the observations were expected to influence the uncertainty as well. In summary, modified ellipsoid with craters was being altered during the process and compared with incomplete lightcurves of the ellipsoid.

### 5.1 Lightcurves

Synthetic lightcurves were not complete – they covered 3/8 of the rotational period. In the body’s reference frame the observer was situated along longitudes from  $270^\circ$  to  $45^\circ$ , marked in blue in Fig. 3. Yellow colour marks the parts of the body that happened to be perpendicular to the observer during the observations, i.e. the visible edge of the body. The choice of longitudes’ range was motivated by the positions of the craters which were desired to be covered in the lightcurves.

There was no Gaussian noise added to the photometric data points and the error in the magnitude was 0. There were 8 lightcurves in total evenly distributed on the orbit (Fig. 4). The target and the observer orbits were circular with semi-major axes  $a_{\text{target}} = 3$  and  $a_{\text{obs}} = 1$ ; apparitions were one year apart. The phase angle in all the observations was  $\alpha = 18.5^\circ$ . The rotational period of the body was  $P = 4.12345\text{h}$  and pole orientation was  $\lambda = 0^\circ$ ,  $\beta = 45^\circ$ . For such pole latitude the whole body is visible for the observer if the target’s and observer’s orbits are coplanar and the apparitions are evenly distributed. Had  $\beta$  been  $90^\circ$  all the lightcurves would have looked the same and having more than one lightcurve only serves the rotational period estimation. For  $\beta = 0^\circ$  the amplitude of two of the lightcurves would drop to 0 mag for any clone, making these lightcurves worthless from the uncertainty assessment point of view.

### 5.2 Results

The projections (Fig. 5) show relative values of uncertainty compared with the  $R_{\text{max}}$  of the model. The values ranged from  $-0.6\%$  to  $-10\%$  for negative uncertainty, and from  $0.4\%$  to  $25\%$  for positive uncertainty. There were 388925 clones out of  $1.7 \times 10^6$  that satisfied Eq. 3.

The unscaled volume could be also computed in relation to the unit volume or volume of a unit sphere. The comparison would be less biased when it comes to elongated bodies.

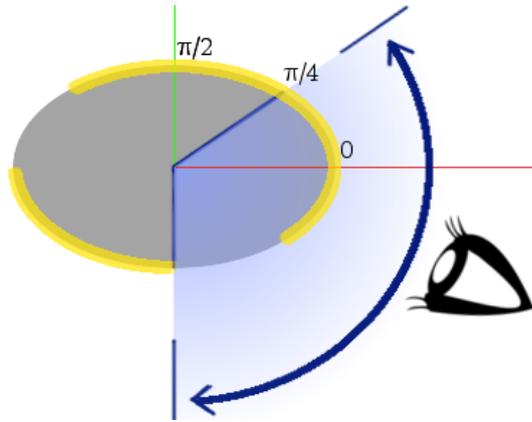


Figure 3: A scheme of observational geometries used in the example. Description in text.

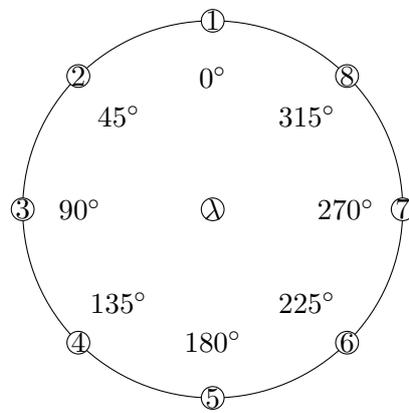


Figure 4: The image shows the distribution of apparitions (with the observer in the center of the graph) in heliocentric reference frame. Description in text.

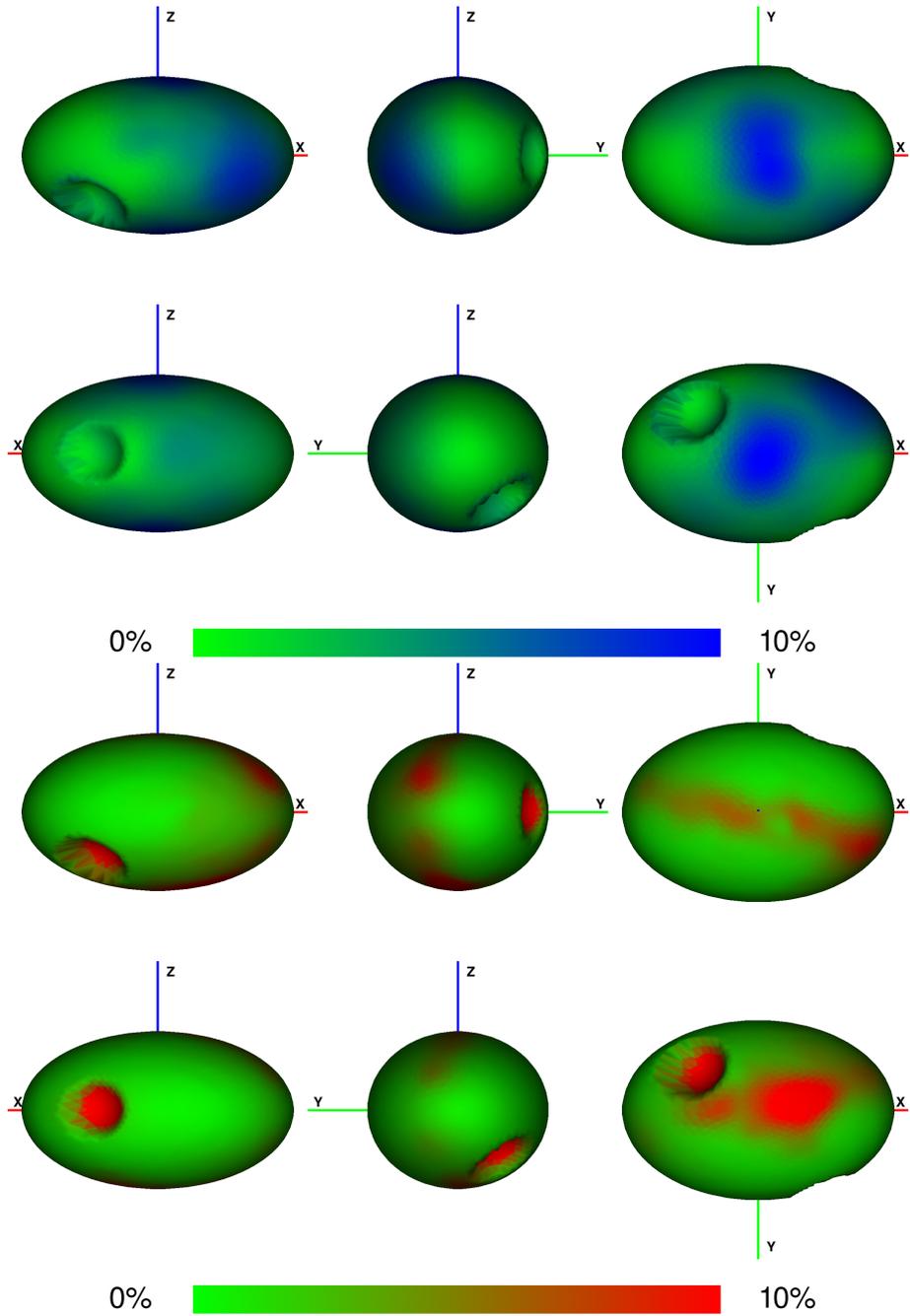


Figure 5: Projections of the target with colour-coded uncertainty information. The uncertainty values come from the variance of probability density distribution (PDD) calculated for each vertex. Top figure shows uncertainty due to changes that create concavities (blue colour) while the bottom shows the hills (red). The scale of the values is the same for the two uncertainty maps, the brightest blue/red showing the values  $\geq 10\%$  of the maximal value. The craters have uncertainty of 25% which corresponds with the differences from the ideal ellipsoid in surface position.

Other type of scaling (e.g. in reference to projected surface area of a sphere) might be used as well depending on the needs.

### 5.2.1 Craters

The biggest positive uncertainty was correlated with the presence of the craters. The uncertainty of 25% corresponds with the difference between modified and unmodified ellipsoids in radial direction. This shows that the surface inside the craters was allowed to move outwards without changing the lightcurves meaningfully until it reached the level of the unmodified ellipsoid surface. Going further changes the silhouette of the body influencing the lightcurves more dramatically and leading to higher RMSD values.

### 5.2.2 Poles

The areas near the poles of the body are very uncertain especially in the negative direction. It was the case in every uncertainty assessment we performed so far. It seems that creating dents near the poles do not change the lightcurves much. Except for the craters, it is the most uncertain part of the body.

### 5.2.3 Surface coverage

The *geometric scattering law* can be treated as a first approximation of any, more sophisticated, scattering law. The area of the body's projection on the observer's field of view has by far the biggest influence on the amount of reflected light. Fig. 3 shows the positions of the observer in the body's reference frame that were used in this example. The yellow parts mark the surface elements perpendicular to the observer – the ones that would change the amount of light the most if moved (especially outwards). Even though observer saw  $135^\circ$  of the surface (if we count the triangles that were facing the observer, i.e. angle between their surface normals and direction to the observer were equal to  $0^\circ$ ), almost whole body was represented in lightcurves. The parts that were not have much larger uncertainty (above 5%). Strange shape of the uncertainty *ribbons* is due to other effects connected with the actual distribution of apparitions on the orbit and the inclination of the pole. Also, the discrepancy between the modified and unmodified ellipsoids' lightcurves (that define the confidence level) allow for surface changes not only away from the craters and unobserved parts.

### 5.2.4 Volume

One of the final goals is to obtain the volume of the model with the error bars. We search for the largest and the smallest body among the accepted clones. From that we establish the range of volumes of the asteroid model. For our example we obtained  $V' = 1.578_{-0.046}^{+0.013}$ , or +0.8% and -2.9% with respect to the nominal model volume percentage-wise.

### 5.2.5 Rotational period and zero-phase

The rotational period of the test body was  $P = 4.123450$  h, while the range of period values found in the population of accepted clones was  $[4.12344854, 4.12345224]$  h, and  $\Delta P = 3.7 \times 10^{-6}$  h. In another test (not shown here) with full lightcurves the range

was [4.12344993, 4.12345039] h,  $\Delta P = 4.6 \times 10^{-7}$  h. The period range larger by one order of magnitude is therefore most probably due to incompleteness of the lightcurves.

The uncertainty of the rotational phase at  $JD_0$  was  $\Delta\gamma_0 = 5^\circ$ , which is acceptable having in mind the density of data points in lightcurves (data point every  $4^\circ$ ), and observational gaps. In full-lightcurve example  $\Delta\gamma_0$  was near  $1^\circ$ .

## 6 Quality assessment of SBNAF targets

Calculating asteroid models' uncertainties is computationally expensive and time consuming. Due to the delay of the AMU computer cluster purchase (about 1 yr.) there was not sufficient amount of time to develop the method in full and assess the quality of SBNAF target models. This process, however, will start promptly as the method is about to reach usable and robust state very soon.

Quality assessment needs to be done case by case and depends on the available data and goals for specific objects. Some cases might require more focus on specific set of parameters (pole, period or shape). At this stage we are prepared to make a statistically solid assessment of the quality of specific object properties.

## References

P. Bartczak and G. Dudziński. in prep.

E. Catmull and J. Clark. Recursively generated B-spline surfaces on arbitrary topological meshes. *Computer-Aided Design*, 10:350–355, 1978.

John R. Taylor. *An Introduction To Error Analysis*. University Science Books, 2 edition, 1997.