



Modeling the Vertical Distribution of the Milky Way's Flat Subsystem Objects

Igor' I. Nikiforov, Vadim A. Usik, and Angelina V. Veselova

Saint Petersburg State University, Universitetskij Prospekt 28, Staryj Peterhof, Saint Petersburg 198504, Russia; i.nikiforov@spbu.ru (IIN)
Received 2022 June 2; revised 2022 November 15; accepted 2022 November 29; published 2023 January 11

Abstract

This paper is an initial stage of consideration of the general problem of joint modeling of the vertical structure of a Galactic flat subsystem and the average surface of the disk of the Galaxy, taking into account the natural and measurement dispersions. We approximate the average surface of the Galactic disk in the region covered by the data with a general (polynomial) model and determine its parameters by minimizing the squared deviations of objects along the normal to the model surface. The smoothness of the model, i.e., its order n , is optimized. An outlier elimination algorithm is applied. The developed method allows us to simultaneously identify significant details of the Galactic warping and estimate the offset z_{\odot} of the Sun relative to the average (in general, non-flat) surface of the Galactic disk and the vertical scale of the object system under consideration for an arbitrary area of the disk covered by data. The method is applied to data on classical Cepheids. Significant local extremes of the average disk surface model were found based on Cepheid data: the minimum in the first Galactic quadrant and the maximum in the second. A well-known warp (lowering of the disk surface) in the third quadrant has been confirmed. The optimal order of the model describing all these warping details was found to be $n_o = 4$. The local (for a small neighborhood of the Sun, $n_o = 0$) estimate of $z_{\odot} = 28.1 \pm 6.1|_{\text{stat.}} \pm 1.3|_{\text{cal.}}$ pc is close to the non-local (taking into account warping, $n_o = 4$) $z_{\odot} = 27.1 \pm 8.8|_{\text{stat.}}^{+1.3|_{\text{cal.}}}$ pc (statistical and calibration uncertainties are indicated), which suggests that the proposed modeling method eliminates the influence of warping on the z_{\odot} estimate. However, the non-local estimate of the vertical standard deviation of Cepheids $\sigma_p = 132.0 \pm 3.7|_{\text{stat.}}^{+6.3|_{\text{cal.}}}$ pc differs significantly from the local $\sigma_p = 76.5 \pm 4.4|_{\text{stat.}}^{+3.6|_{\text{cal.}}}$ pc, which implies the need to introduce more complex models for the vertical distribution outside the Sun's vicinity.

Key words: Galaxy: disk – Galaxy: structure – Galaxy: fundamental parameters – methods: data analysis

1. Introduction

The vertical distribution of objects in various Galactic subsystems contains valuable information about the origin, evolution and dynamics of our Galaxy, so determining the characteristics of this distribution is an important task of Galactic Astronomy. The study of the vertical distribution may include consideration of many phenomena, but one of them should be taken into account necessarily—this is the offset z_{\odot} of the Sun relative to the plane of the Milky Way's disk toward the North Galactic Pole. Therefore, in the simplest case, modeling of the vertical distribution is reduced to determining the value of z_{\odot} and some dispersion parameter (standard deviation, scale height, etc.) that characterizes the scattering of subsystem objects relative to the average plane of the Galactic disk (usually relative to the midplane of this subsystem). The first estimate of $z_{\odot} = 13.5 \pm 1.7$ pc was obtained by van Tulder (1942) from the analysis of nearby stars. Subsequently, in many papers, the solar offset was determined by different methods for various objects and Galactic subsystems. Bland-Hawthorn & Gerhard (2016) adopted as the best (local)

estimate the result of Jurić et al. (2008) from the complete Sloan Digital Sky Survey (SDSS) photometric survey, $z_{\odot} = 25 \pm 5$ pc, which covers many other estimates.

However, the solar offset relative to the differently defined midplane of the disk does not seem to be described by a single value of z_{\odot} . For example, Bobylev & Bajkova (2016b) obtained significantly different results for reference objects (tracers) of different types: $z_{\odot} = 5.7 \pm 0.5$ pc for a sample of methanol masers, $z_{\odot} = 7.6 \pm 0.4$ pc for data on H II regions and $z_{\odot} = 10.1 \pm 0.5$ pc for data on giant molecular clouds; at the same time, Ferguson et al. (2017) derived values of $z_{\odot} = 14.9 \pm 0.5$ pc for a uniform selection of SDSS K and M dwarf stars and $z_{\odot} = 15.3 \pm 0.4$ pc for an expanded selection, Buckner & Froebrich (2014) found an estimate $z_{\odot} = 18.5 \pm 1.2$ pc for open clusters, and Majaess et al. (2009) obtained values of $z_{\odot} = 26 \pm 3$ pc for Cepheids. A comparison of these and other z_{\odot} estimates obtained in various studies (see, e.g., summaries in Yao et al. 2017; Skowron et al. 2019a) shows that the differences between these estimates cannot be explained only by statistical errors, with some estimates varying significantly, even

for objects of the same type (e.g., for open clusters and Cepheids). This shows that the discrepancies reflect not only the possible objective difference in the values of z_{\odot} between different types of objects (subsystems of the Galaxy), but also other factors in the problem.

In addition to the Z -offset of the Sun, the number of already established or potential factors affecting the results of modeling the vertical distribution of objects includes: (1) the warp of the Galactic disk, (2) the dependence of the values of the characteristics of the vertical distribution on the position on the disk for the selected Galactic subsystem (e.g., the flare of the Galactic disk), (3) the possible (and in the case of vertical dispersion, real) dependence of these characteristics on the type of Galactic subsystem, (4) the need to establish the functional type of the vertical distribution and its possible variations with the position on the disk and with the type of subsystem, as well as (5) taking into account the random uncertainty of heliocentric distances, systematically distorting the true vertical distribution. The problem in general (taking into account all these factors) has not yet been solved. Meanwhile, different combinations of these factors may be responsible for discrepancy of the results (in particular, of z_{\odot} estimates) in different papers. Subjective factors can also lead to this: the choice of the general appearance of the model of the average surface of the disk, possible mismatch of the distance scales used in different works, and the dependence of the results of modeling on the size and configuration of the disk area under consideration (the area covered by the data).

Despite the lack of a solution to the problem in general, some of these factors and their combinations were considered. The most important factor is the presence of a warp in the Milky Way's disk. The warp was noticed as soon as the observation data in the 21 cm line of neutral hydrogen appeared for the southern hemisphere (Burke 1957; Kerr 1957). Subsequent studies (Oort et al. 1958; see, e.g., Binney & Merrifield 1998 and Bland-Hawthorn & Gerhard 2016 reviews and references therein; Skowron et al. 2019a; Chrobáková et al. 2020; among others) have shown that a significant stellar/gas warp begins outside the solar circle, and in the inner Galaxy the disk is very close to flat, including on the far side of the disk (Minniti et al. 2021). Various data indicate that one part of the warped disk deviates from the plane of the inner disk toward the North Galactic Pole, the other deviates in the opposite direction.

Not taking into account the large-scale warp (if a plane parallel to the equator of the Galactic coordinate system $b = 0^{\circ}$ is taken as a model of the average surface of the Galactic disk) can significantly affect the estimates of the solar offset z_{\odot} and the vertical scales of flat subsystems (see, e.g., the dependence of these characteristics for planetary nebulae on the size of considered near-solar region in Bobylev & Bajkova 2017). One way to avoid this is to exclude the warp zone from consideration under the assumption that in the remaining area

of the disk its average surface is flat: restrictions are imposed on the selection of tracers, for example, by the heliocentric distances r (e.g., $r \lesssim 4$ kpc in Bobylev & Bajkova 2016a; $r \lesssim 4.5$ kpc in Bobylev & Bajkova 2016b), by the predicted maximum warp offsets (< 10 pc in Yao et al. 2017) and by the distance R to the axis of rotation of the Galaxy ($R < 7.0$ kpc in Reid et al. 2019). However, the exclusion of the warp zone requires the adoption of a specific warp model, and it is often taken as simple for this and other applications: the disk in the inner Galaxy ($R \leq R_w$) is considered undisturbed, and in the outer one ($R > R_w$) it is usually represented by a combination of a power dependence on R and a simple trigonometric function of the azimuthal coordinate (e.g., Binney & Merrifield 1998; Pohl et al. 2008; Xu et al. 2015; Yao et al. 2017; Romero-Gómez et al. 2019; Cheng et al. 2020; Mosenkov et al. 2021). At the same time, to describe the warp in the outer Galaxy, in most of its morphological studies, simple symmetric models with a limited set of parameters are used—the radius R_w at which the disk starts bending, the phase angle of the line-of-nodes and the maximum amplitude of the warp (see, e.g., Romero-Gómez et al. 2019 and references therein).

However, the warp is clearly more complicated. First, the inner part of the disk is not perfectly flat—there are corrugations on the scale of ~ 30 pc (Oort et al. 1958; Figure 3; Spicker & Feitzinger 1986, Figure 9.22). N -body simulations of the Milky Way interacting with a satellite similar to the Sagittarius dwarf galaxy show that repeated satellite passes can generate local ripples, including in the inner disk (Poggio et al. 2020, Figure 2). According to kinematics, the onset of the warp occurs at a guiding radius inside the solar circle, $R_g \lesssim 7$ kpc (Schönrich & Dehnen 2018), or even in the center of the Galaxy (Li et al. 2020). Second, the outer part of the warp is also not described by a simple model—there are manifestations of lopsidedness of the warp and twisting of its line-of-nodes (Romero-Gómez et al. 2019; Chrobáková et al. 2020); Xu et al. (2015) detected an oscillating asymmetry in the SDSS main sequence star counts on either side of the Galactic plane in the anticenter region, between longitudes of $110^{\circ} < l < 229^{\circ}$. In addition, the morphology and kinematics of the warp depend on the type/age of the tracers (e.g., Romero-Gómez et al. 2019; Chrobáková et al. 2020). Moreover, hydrodynamic modeling of the evolution of an ensemble of stars formed in the warp shows that only younger populations trace the warp detected by HI (Khachatryan et al. 2021) and that the influence of the bending waves excited by irregular gas inflow is most strongly manifested in the young populations (Khachatryan et al. 2022). This means that the warp model, universal for all disk subsystems of the Galaxy, can hardly be accepted.

Kinematic manifestations of the warp also indicate its asymmetry and complexity in general, as well as the dependence of its characteristics on the age of tracers (e.g.,

Romero-Gómez et al. 2019; Cheng et al. 2020; Li et al. 2020 and references in these works).

Based on the above, the exclusion of the warp zone as a method of eliminating biases in the vertical distribution parameters can only give a partial (local) solution to the problem, the accuracy of which depends on the details of the accepted warp model and on its realism in the case of the tracers under consideration and on assumptions about the boundaries of the warp-distorted area. All these assumptions can be sources of systematic errors. That is why it seems important to us to abandon simplified warp models and consider the most general analytical warp model describing all the significant structural features of the middle surface of the disk identified by the tracers under consideration. The method of excluding the warp zone is also unsuccessful due to the presence of the disk flaring, which begins at $R \gtrsim R_0$, where R_0 is the Galactic center distance (e.g., Reid et al. 2019; Mosenkov et al. 2021), since the dependence of the dispersion parameter on the accepted boundaries of the area “undisturbed” by the warp appears.

The warp is currently being actively explored in many ways. In particular, warp precession is actively discussed (e.g., Cheng et al. 2020). However, as noted by Poggio et al. (2020), the precession parameters depend on our knowledge of the shape of the warp and its differences for different stellar populations. In addition, Chrobáková & López-Corredoira (2021) even raise the question of the very existence of precession, since the application of a warp model inconsistent with the tracers used leads to a fictitious precession.

Detailed warp models are also important both for studying the dependence of z_\odot and vertical dispersion characteristics on the type/age of tracers, and for identifying the cause and dynamic nature of the warp of our Galaxy, which remain unclear (Binney & Merrifield 1998; Poggio et al. 2020; Khachatryan et al. 2021).

Note also that in the framework of an alternative approach applied by Mosenkov et al. (2021)—photometric 3D decomposition of the Milky Way taking into account flaring and warp—the parameters of the warp disk are poorly determined, since only a 2D map is considered, whereas for creating a reliable 3D model of the warp one needs to have a 3D distribution of stars in the Galaxy.

Despite the fact that the best solution would be to model the Z-distribution of objects taking into account all the factors mentioned at once, due to the complexity of the overall task, we focus in this paper on the task of constructing a detailed warp model with a minimum of assumptions. We will not consider the influence of random errors in the distance here (the selected data catalog allows this, see Section 3), as well as the disk flaring, since without taking into account errors in distances, the flaring parameters may turn out to be strongly biased.

2. Method

We will study the spatial distribution of objects in the heliocentric Cartesian coordinate system, which does not require taking any value of R_0 : X-axis is directed toward the Galactic center, Y-axis is toward the rotation of the Galaxy and Z-axis is toward the North Galactic Pole.

In order to free the warp model as much as possible from pre-accepted assumptions, we will consider as models the $\zeta_n(X, Y)$ polynomials, each of which is a Maclaurin series expansion in the solar neighborhood up to the n -th order of the average Z-coordinate of the Galactic disk as a function of position on the XY plane

$$\begin{aligned} \zeta_n(X, Y) &= \sum_{i=0}^n \sum_{j=0}^i z_{i-j,j} X^{i-j} Y^j \\ &= z_{00} + z_{10}X + z_{01}Y + z_{20}X^2 + z_{11}XY \\ &\quad + z_{02}Y^2 + \dots + z_{n0}X^n + \dots + z_{0n}Y^n, \\ \mathbf{z}_\zeta &= \{z_{ij}\}_0^n, \quad z_\odot = -z_{00}. \end{aligned} \quad (1)$$

Here \mathbf{z}_ζ is the vector of $M = (n+2)(n+1)/2$ parameters of the model. Function (1) represents the average surface of the Galactic disk, defined by the spatial distribution of objects (in the general case of matter) of the selected Galactic subsystem. The distance ρ from the object to the surface $\zeta_n(R, Z)$ along the normal to this surface will be considered as the value of the deviation of the object from the model average surface of the disk.

To obtain an estimate of the vector \mathbf{z}_ζ , we generally rely on the maximum likelihood method. In this paper, to simplify, we assume that ρ as a random variable is distributed according to a normal law with zero mean, that is, that the probability density of ρ has the form

$$f(\rho) = \frac{1}{\sigma_\rho \sqrt{2\pi}} e^{-\frac{\rho^2}{2\sigma_\rho^2}}, \quad (2)$$

where σ_ρ is the standard deviation. We assume that the value of σ_ρ is the same for all objects in the sample. However, in future work we propose to investigate and take into account the dependence of σ_ρ and/or other dispersion parameters on the Galactocentric distance. With probability density (2), the likelihood function L and logarithmic likelihood function \mathcal{L} are

$$\begin{aligned} L &= \prod_{i=1}^N \frac{1}{\sigma_\rho \sqrt{2\pi}} e^{-\frac{\rho_i^2}{2\sigma_\rho^2}}, \\ \mathcal{L} \equiv -\ln L &= \frac{N}{2} \ln(2\pi) + N \ln \sigma_\rho + \frac{1}{2} \sum_{i=1}^N \frac{\rho_i^2}{\sigma_\rho^2}, \\ \rho_i &= \rho_i(r_i, l_i, b_i; \mathbf{z}_\zeta), \end{aligned} \quad (3)$$

where N is the number of objects in the sample; r_i , l_i and b_i are the heliocentric distance, and galactic longitude and latitude of i -th object, correspondingly. Being a dispersion parameter under parameterization (2), σ_ρ is included in the general vector

of the problem parameters $\mathbf{z} = (\sigma_\rho, \mathbf{z}_\zeta) = (z_1, z_2, \dots, z_K)$, where $K = M + 1$. The minimum of the function \mathcal{L} gives estimates of the parameters, including the value of the standard deviation σ_ρ of objects across the disk, which represents the contributions of both the true (natural) vertical dispersion of objects and the random uncertainty of distance estimates (the latter contribution is negligible, see Section 5). Thus, under assumption (2), the maximum likelihood method was reduced to the nonlinear least squares method (Equation (3)). The resulting value of σ_ρ was multiplied by the coefficient $\sqrt{N/(N-M)}$ to obtain an unbiased estimate. The vector \mathbf{z}_{err} of mean parameter errors and the mean model prediction error $\sigma_{\zeta_n}(X, Y)$ were calculated based on the Hessian $H(\mathcal{L})$ with elements h_{ij} (Hudson 1964; Wall & Jenkins 2012)

$$\begin{aligned} h_{ij} &= \frac{\partial^2 \mathcal{L}^2}{\partial z_i \partial z_j}, \quad i = 1, \dots, K, \quad j = 1, \dots, K, \\ C &= [H(\mathcal{L})]^{-1}, \\ \mathbf{z}_{\text{err}} &= \{\sqrt{c_{ii}}\}_{i=1}^K, \\ \sigma_{\zeta_n}^2(X, Y) &= \{\text{grad}[\zeta_n(X, Y; \mathbf{z}_\zeta)]\}^T C' \text{grad}[\zeta_n(X, Y; \mathbf{z}_\zeta)], \end{aligned} \quad (4)$$

where c_{ii} are diagonal elements of the covariance matrix C , and C' is a submatrix of C that does not contain covariances involving σ_ρ ; here, an estimate of the vector \mathbf{z} obtained by minimizing \mathcal{L} is substituted for all values.

After finding the parameters, an outlier exclusion algorithm described in Nikiforov (2012) was applied to the sample of objects under consideration. This algorithm differs from the usual 3σ criterion in that it uses a variable exclusion limit, which increases as the number of objects increases.

In order to check how well the observed distribution of sample objects by deviations ρ agrees with the model probability density function (2), we use Pearson's chi-square test.

A priori choice of the order n of expansion (1) would lead to significant errors in all parameters, including z_{00} (see Section 4), so the value of n was also optimized using a simple algorithm. Models of the order $n = 0, 1, 2$ and so on are built sequentially. For each model of order n , the number of parameters z_{ij} of order n ($i + j = n$) whose estimates differ from zero at the significance level $\geq 2\sigma$ ($\sigma_{z_{ij}}/|z_{ij}| \leq 0.5$) is calculated. Then we find a model of the highest order such that it has at least one 2σ -significant parameter of the same order as the model. If the total number of significant parameters of the selected model is greater than the corresponding number for any lower-order model, then the selected model of order n is assumed to be optimal. Otherwise, a model of order $n - 1$ is considered as possibly optimal and is compared in the same way with models of lower orders in terms of the number of significant parameters. Then either the order $n - 1$ is assumed to be optimal, or a transition is made to the order $n - 2$, and so on until some order $n \geq 0$ is accepted as optimal, n_o . The

importance of choosing the correct model order will be illustrated in Section 4.

3. Data

We use the catalog of classical Cepheids by Berdnikov et al. (2000) in the version of Mel'nik et al. (2015), which provides data for 674 Cepheids from the General Catalogue of Variable Stars (e.g., Samus et al. 2017). The catalog is an updated version of the catalog of Cepheid parameters by Berdnikov et al. (2000). Observational data were obtained with 0.4–1 m telescopes of the Maidanak Observatory (Republic of Uzbekistan), Cerro Tololo and Las Campanas observatories (Chile), Cerro Armazones Observatory of the Catholic University (Chile) and South African Astronomical Observatory (see Dambis et al. 2015 for details). In particular, in order to reliably determine the distances, Cepheids discovered during CCD monitoring of the southern sky performed as a part of the All Sky Automated Survey (ASAS) project (Pojmanski 2002) were observed, therefore, a survey of Cepheids across the entire sky was performed. The peak of the distribution of Cepheids by the mean apparent magnitudes in the V band falls on the values of $\langle V \rangle = 12\text{--}13$ mag, and the limiting magnitude of the catalog is $\langle V \rangle = 15$ mag. The distances were obtained based on the period–luminosity relation in the K infrared band and interstellar-extinction law using the period–normal color ($B - V$) relation derived earlier (see Dambis et al. 2015).

The authors of the catalog did not specify the possible value of the distance modulus error. However, the analysis of these data in Veselova & Nikiforov (2020) showed that the nominal random errors of distance estimates given in Mel'nik et al. (2015) are small (the mean error of distance moduli $\sigma_d < 0.14^m$). Using this distance catalog gives us the opportunity to apply a simpler method that does not take into account random distance errors. In the future, we intend to use newer data covering more extensive areas where taking into account the uncertainty of distances within a more complex method is necessary. Recent versions of Berdnikov et al.'s catalog have been successfully used to study the Galactic structure. Based on the catalog, Mel'nik et al. (2015) identified signs of ring formations in the Galaxy, and Dambis et al. (2015) and Veselova & Nikiforov (2020) performed spatial modeling to determine the parameters of spiral arm segments.

According to the original distance scale calibration of the catalog the distance to the Large Magellanic Cloud (LMC) is $d_{\text{LMC}}^* = 18.25 \pm 0.05$ mag (Berdnikov et al. 2000). Modern LMC calibration is $d_{\text{LMC}} = 18.49 \pm 0.09$ mag (de Grijs et al. 2014), which leads to a correction factor c for the distances of the catalog used

$$\begin{aligned} c &= \frac{r + \Delta r}{r} = 10^{0.2\Delta d} = 1.117_{-0.050}^{+0.053}, \\ \Delta d &= d_{\text{LMC}} - d_{\text{LMC}}^*. \end{aligned} \quad (5)$$

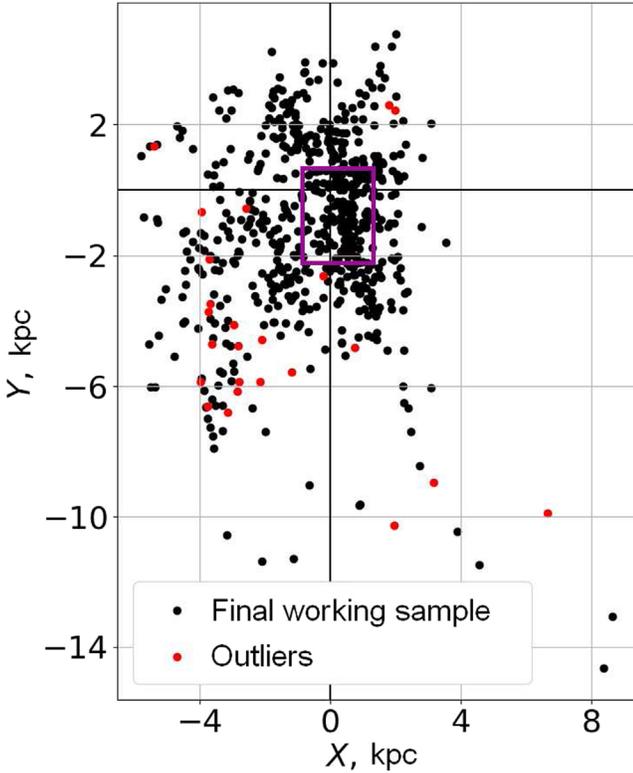


Figure 1. Objects of the final working sample and outliers projected onto the XY plane. The purple rectangle marks the boundary of the local sample (see text). X -axis is directed toward the Galactic center, while Y -axis is toward the rotation of the Galaxy. The Sun is placed at $X = 0$ kpc, $Y = 0$ kpc.

We analyzed the original catalog estimates of distances, and then adjusted the main results for the factor c .

Spatially isolated objects were manually excluded from the initial sample, i.e., those that are nominally located clearly far from the main group of catalog objects. The remaining objects, the totality of which we called the working sample, were used for calculations. After applying the outlier elimination algorithm (Nikiforov 2012) during the analysis of the working sample, we obtained the final working sample containing 615 objects (Figure 1).

We also identified a *local sample*, which is part of the working sample representing the largest neighborhood of the Sun with the property of relative completeness of the identification of objects of this type (classical Cepheids). Assuming a homogeneous distribution of objects projected onto the XY -plane (the Galactic plane) in the vicinity of the Sun, the number of objects projected onto a section of the XY -plane should be proportional to the area of this section. Usually, a set of concentric rings is used to select a local, close to complete, subsample, but the working sample has an asymmetry with respect to the origin of coordinates on the XY plane. For this reason, it was decided to use a set of rectangular

frames with borders homothetic to the rectangular border of the working sample: the lengths of each side of the frame vary according to a given scale factor, and the ratio of distances from the origin to the edges of the frames remain constant (Figure 2, left panel). In the case of a small frame width, the area of the frame will be proportional to the outer perimeter, so for a complete sample, the number of objects in the frame should increase linearly with the size of the frame (scale factor). On the right panel of Figure 2 the linear growth is observed within the first three bins (frames), which are marked in red; objects within these frames were taken as a local sample. The boundaries of this sample are shown in Figure 1 (purple rectangle). The sample size is 154. After excluding an outlier, the final local sample of 153 objects was obtained.

4. Results

The influence of the accepted order of the model ζ_n on the results can be illustrated by the example of the dependence of z_{00} on n for the final working sample (Figure 3). It can be seen that for small orders the estimates of z_{00} strongly depend on n , while for larger orders the estimates vary insignificantly.

The optimal order of the model ζ_n for the final working sample turned out to be $n_o = 4$. Estimates of model parameters are presented in Table 1, and significant estimates are expressed in bold. For the local and final local samples, n_o is 0. Final results for z_{\odot} and σ_{ρ} without and with the correction of distance scale (5) for the final working and final local samples are listed in Table 2. Comparison with Figure 3 shows that the choice of an underestimated order ($n < n_o$) of the model can lead to an obviously incorrect estimate of z_{\odot} (here at $n = 0, 2$).

The model surface $\zeta_4(X, Y)$ for the final working sample is depicted in Figure 4. Here the level

$$\zeta_4(X, Y) = z_{00}, \quad (6)$$

is represented by a black line. This line is the intersection of the model of the average surface of the disk with the nominal plane of the Galaxy XY . Line (6) corresponds to the line-of-nodes in simple models. Figure 4 affirms that in reality the line (6), skirting the areas of local extrema, is quite curved. By analogy with the line-of-nodes, the curve $\zeta_{n_o}(X, Y) = z_{00}$ can be called a “curve-of-nodes.” Of course, the constructed model $\zeta_4(X, Y)$ is real only for the area of the disk that is covered by the data used. Boundaries of areas, within which the mean error of the model is $\sigma_{\zeta_4}(X, Y) \leq \frac{1}{6}\sigma_{\rho} = 20$ pc, $\leq \frac{1}{4}\sigma_{\rho} = 40$ pc and $\leq \frac{1}{2}\sigma_{\rho} = 59$ pc shown in Figure 4 by dotted, dashed and dashed-dotted lines, respectively, give an idea of the applicability of the model ζ_4 depending on the specified level of its uncertainty. Here $\sigma_{\zeta_4}(X, Y)$ was calculated using the formula (4), and $\sigma_{\rho} = 119$ pc (see Table 1).

An alternative representation of the resulting model is displayed in Figure 5, which shows the extremum and boundary lines of the model surface $\zeta_4(X, Y)$ in projections

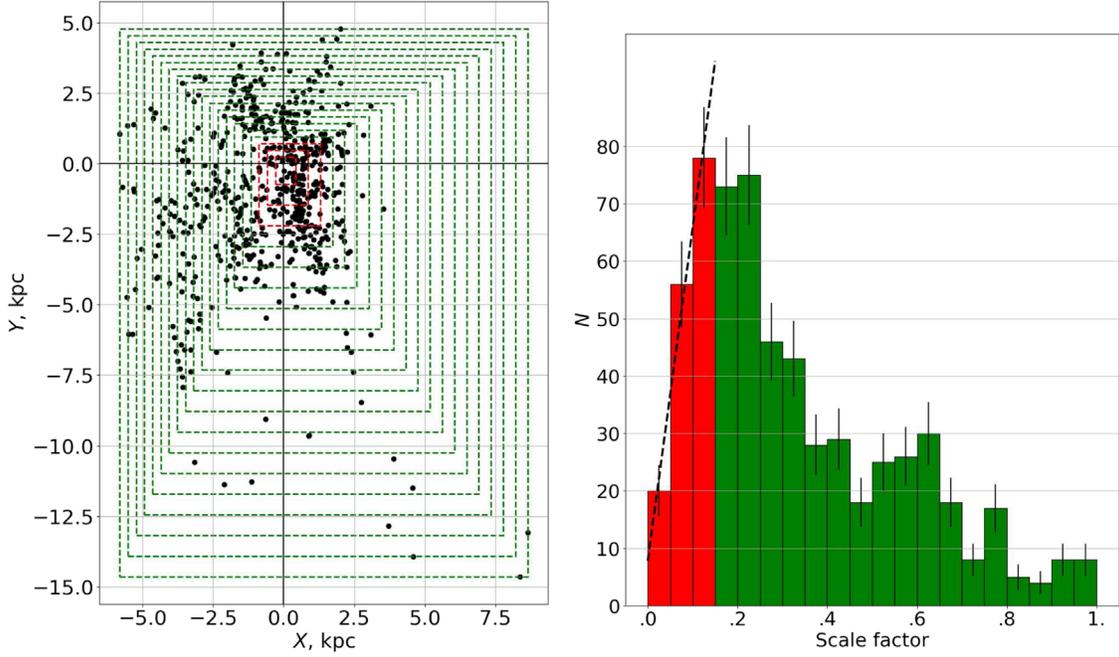


Figure 2. (Left) Homothetic boundaries of rectangular frames used to construct the local sample (see text), superimposed on the distribution of objects of the working sample in projection on the XY plane. (Right) Histogram of the distribution of objects of the working sample in rectangular frames. The red color corresponds to the frames within which the sample (local sample) can be considered complete; the dashed straight line shows the linear growth area of the number of objects in the first three bins.

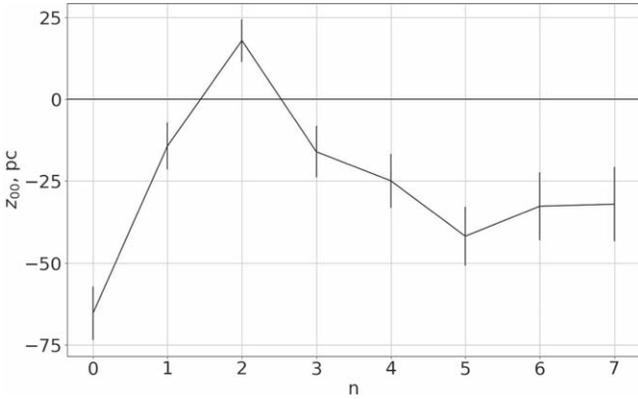


Figure 3. Dependence of z_{00} ($= -z_{\odot}$) on the model order n for the final working sample. Vertical bars indicate the uncertainties of estimates.

on the planes XZ and YZ in comparison with the Cepheids of the final working sample. The extremum lines (there may be several of them for each projection) are dependencies of Z -coordinates of points of local extremes of the surface $\zeta_4(X, Y)$ with fixed X — $\zeta_{\max}(X, Y)$, $\zeta_{\min}(X, Y)$ (Figure 5, top panel) or with fixed Y — $\zeta_{\max}(X, Y)$, $\zeta_{\min}(X, Y)$ (bottom panel). The boundary lines mark the values of $\zeta_4(X, Y)$ at the boundaries of the final working sample: $\zeta_4(X, Y_{\min})$ and $\zeta_4(X, Y_{\max})$ (top panel), and $\zeta_4(X_{\min}, Y)$ and $\zeta_4(X_{\max}, Y)$ (bottom panel), where

$X_{\min} = -5.8$ kpc, $X_{\max} = 8.6$ kpc, $Y_{\min} = -14.6$ kpc, $Y_{\max} = 4.8$ kpc. Figure 5 affirms that the extremum lines mainly fall on the areas covered by the data, and the boundary lines indicate edge approximation effects in the area $X \gtrsim 4$ kpc, $Y \gtrsim 0$ kpc.

The well-known general shape of the Galactic disk warp is clearly visible in Figure 4—in the first and second quadrants, the model surface as a whole rises above the XY plane; in the third and fourth quadrants, the surface decreases below this plane. In the area near $X \approx -3$ kpc, $Y \approx -6$ kpc, the decrease in the average surface is observed for almost all objects.

However, in addition, we found two extrema in the first and second quadrants that do not fit into simple warp models. The sections of the model with planes parallel to the XZ and YZ planes and passing through the points of extrema are depicted in Figure 6. The significance S of the extrema was estimated by the formula

$$S_{\zeta_4}(X, Y) = \frac{\zeta_4(X, Y) - z_{00}}{\sigma_{\zeta_4}(X, Y)}. \quad (7)$$

Table 3 shows the parameters of local extrema. The given values of S indicate that the local minimum and maximum are significant at the level of at least 2σ and 5σ , respectively.

In Figure 6 the slope of the model surface $\zeta_4(X, Y)$ to the nominal plane of the Galaxy XY is visible. Having calculated

Table 1

 Parameter Estimates for the Model ζ_n of the Average Surface of the Galactic Disk of Optimal Order $n_0 = 4$ Obtained for the Final Working Sample of Classical Cepheids

Parameter	Estimation	$\sigma_{z_i}/ z_i $	Parameter	Estimation	$\sigma_{z_i}/ z_i $
σ_ρ	0.1187 ± 0.0033	0.03	z_{21}	0.002 89 ± 0.00094	0.33
z_{00}	-0.0243 ± 0.0079	0.33	z_{12}	0.002 88 ± 0.00055	0.19
z_{10}	-0.0320 ± 0.0061	0.19	z_{03}	0.001 80 ± 0.00028	0.16
z_{01}	-0.0131 ± 0.0040	0.31	z_{40}	-0.00012 ± 0.00012	1.00
z_{20}	0.0055 ± 0.0025	0.45	z_{31}	0.000 13 ± 0.00014	1.08
z_{11}	0.0008 ± 0.0018	2.25	z_{22}	0.000 22 ± 0.00010	0.45
z_{02}	-0.0007 ± 0.0010	1.43	z_{13}	0.000 185 ± 0.000048	0.26
z_{30}	0.002 03 ± 0.00087	0.43	z_{04}	0.000 099 ± 0.000020	0.20

Note. Estimates that differ from zero at a significance level of at least 2σ are highlighted in bold. The standard deviation σ_ρ is given in kiloparsecs, while values of z_{ij} are in units of kpc^{1-i-j} .

Table 2

 Final Estimates of z_\odot and σ_ρ for the Optimal Models $\zeta_{n_0}(X, Y)$ Obtained for Two Samples of Cepheids

Sample	Original Distance Scale	Corrected Distance Scale
Final working sample ($N = 615$, $n_0 = 4$)	$z_\odot = 24.3 \pm 7.9$ pc $\sigma_\rho = 118.7 \pm 3.3$ pc	$z_\odot = 27.1 \pm 8.8 _{\text{stat.}}^{+1.3}_{-1.2} _{\text{cal.}}$ pc $\sigma_\rho = 132.0 \pm 3.7 _{\text{stat.}}^{+6.3}_{-5.9} _{\text{cal.}}$ pc
Final local sample ($N = 153$, $n_0 = 0$)	$z_\odot = 25.2 \pm 5.5$ pc $\sigma_\rho = 68.5 \pm 3.9$ pc	$z_\odot = 28.1 \pm 6.1 _{\text{stat.}} \pm 1.3 _{\text{cal.}}$ pc $\sigma_\rho = 76.5 \pm 4.4 _{\text{stat.}}^{+3.6}_{-3.4} _{\text{cal.}}$ pc

Note. The estimates are given in the original catalog distance scale and in the scale adjusted for calibration $d_{\text{LMC}} = 18.49 \pm 0.09$ mag (see text).

the distance ρ_\odot from the Sun to the surface $\zeta_4(X, Y)$ along the normal to the latter, we obtained an estimate of the angle of inclination of the local average surface of the Galactic disk to the plane XY of the galactic coordinate system

$$\gamma = \arccos\left(\frac{\rho_\odot}{z_\odot}\right) = 1^\circ 79^{+0.34}_{-0.33}. \quad (8)$$

The 1σ -uncertainty of the angle γ indicated here was found by the Monte Carlo method based on the results of processing 50 mock data catalogs.

According to Pearson's chi-square test, the probability density function $f(\rho)$ (2) does not fit data well (see Figure 7, left panel)—the probability of accepting the null hypothesis that the observed distribution has a probability density of the form (2) is less than 1%. This means the other functions $f(\rho)$ or combinations of them should be considered in the future. However, for the final local sample the chi-square test gives a probability of acceptance of the hypothesis (2) of about 30% (see Figure 7, right panel), i.e., the Gaussian distribution as a model for $f(\rho)$ should not be excluded from consideration.

Since in this paper we used a catalog (Mel'nik et al. 2015) based on observations in infrared bands (I_C and K ; see Dambis et al. 2015), this allows us to expect a low selection effect due to the absorption of light by dust in the Galactic disk. In particular, there should be no significant differential selection in the vertical direction, i.e., statistical sample bias to the north of the disk. Indeed, due to the position of the Sun above the average surface of the disk, a ray of light from an object located south of the average surface of the disk passes through a larger thickness of the disk and experiences greater light absorption compared to a northern object at the same distance from the average surface. Therefore, in principle, one would expect a sharper truncation of the observed distribution of deviations ρ from the model from the negative ρ side compared to the positive ones, i.e., greater detection of northern objects compared to southern ones. Asymmetry of this type in the distribution of deviations ρ does not really manifest itself in a noticeable way (Figure 7). In reality, for the working sample, instead of a sharper truncation on negative ρ , rather on the contrary, there is some deficit on $\rho \sim +0.2$ kpc (Figure 7, left panel). The standard deviation calculated for objects with $\rho > 0$ pc for any sample does not significantly exceed the standard deviation found for objects with $\rho < 0$ kpc (see Table 4). These results support the insignificance of north–south selection.

Note that the increase in the selection effect with distance from the Sun in the sense of incomplete sampling (Figure 2, right panel) does not in itself lead to bias, i.e., to systematic errors in the position of the *average* (model) surface of the disk and in estimation of the dispersion of objects relative to this surface. Classical Cepheids belong to population I and, being quite young objects— $\sim 10^7$ – 10^8 yr (see, e.g., Veselova & Nikiforov 2020)—represent a thin disk of the Galaxy. Moreover, the vertical deviation of the subsystem of classical Cepheids $\sigma_\rho \sim 130$ pc (this work) is significantly smaller than the average vertical scale $h_z = 300 \pm 50$ pc of the thin disk as a whole (Bland-Hawthorn & Gerhard 2016). This makes classical Cepheids a good tracer of the thin disk of the Galaxy:

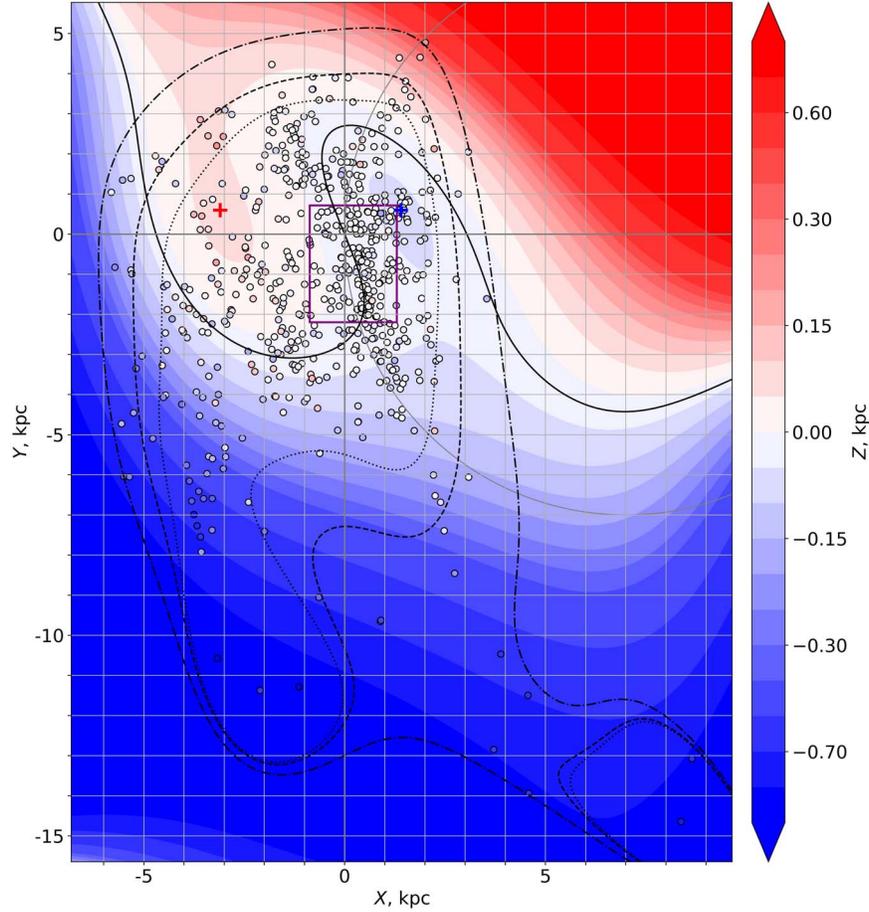


Figure 4. A map of the model ζ_4 of the average disk surface of optimal order, constructed from the objects of the final working sample. The color shows the value of the function $\zeta_4(X, Y)$. The objects are marked as colored circles with color representing the Z -coordinate. The black line represents the z_{00} -level, the red and blue crosses depict the local extrema (maximum and minimum, respectively), the purple rectangle signifies the boundary of the local sample, while the gray line corresponds to the solar circle $R = R_0 = 8$ kpc. The dotted, dashed and dashed-dotted lines are the boundaries of the applicable areas of the model σ_{ζ_4} , within which the mean error of the model is $\sigma_{\zeta_4}(X, Y) \leq \frac{1}{6}\sigma_p = 20$ pc, $\leq \frac{1}{4}\sigma_p = 40$ pc and $\leq \frac{1}{2}\sigma_p = 59$ pc, respectively.

even in areas where there are few of them, they still represent the position of the disk with a relatively small spread. In such areas the mathematical expectation of the average Z -coordinate of the Cepheids, \bar{Z} , remains equal to the true Z -coordinate of the average surface of the disk, and only the mean error of \bar{Z} estimate increases, i.e., in the case of our method, the mean error of the model σ_{ζ_4} increases (isolines of $\sigma_{\zeta_4}(X, Y)$ are shown on Figure 4).

On the other hand, Cepheids, concentrated toward spiral arms (see, e.g., Veselova & Nikiforov 2020), like other tracers of the spiral structure of the Galaxy, often have not a uniform, but a “patchy” distribution along the spiral arms (for example, Efremov 2011, Figure 1; Nikiforov & Veselova 2018, Figure 13; Reid et al. 2019, Figure 2; Veselova & Nikiforov 2020, Figure 5). This, as well as the growth of incomplete detection of objects and a decrease in the density of the disk to the periphery, leads to the fact that at large distances from the Sun,

gaps appear in the distribution of Cepheids, for example at $(X, Y) \sim (0, -7)$ kpc (Figure 4). Of course, in the areas of such gaps, the constructed model should be treated only as a smooth interpolation of the average trend between the areas represented by the data. However, when imposing a stricter restriction on the mean error of the model, e.g., $\sigma_{\zeta_4}(X, Y) \leq \frac{1}{6}\sigma_p = 20$ pc, the internal area of applicability of the model does not include most of these lacunae (see Figure 4). The relatively high accuracy of the model for the area in the lower right corner of Figure 4 is, of course, only formal, due to the fact that any sufficiently flexible model must pass through a few Cepheids in this area, located at some distance from the bulk of the sample objects. On the other hand, these and other Cepheids at large negative Y are in the general trend of a well-known decrease in the average surface of the disk in this area. So, in the region $Y < -10$ kpc, all Cepheids of the working sample are in

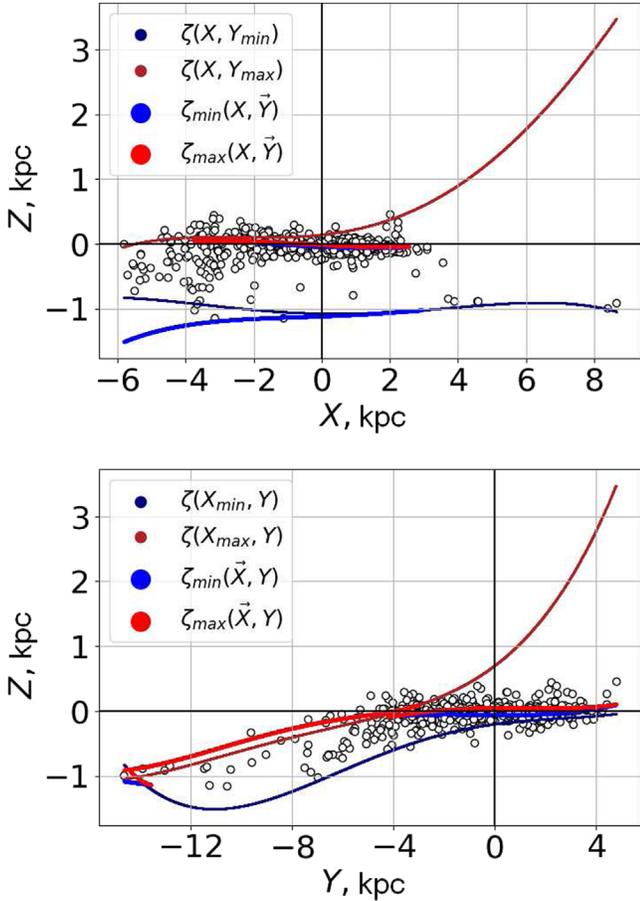


Figure 5. Extremum lines (thicker lines of lighter color) and boundary lines of the model surface $\zeta_4(X, Y)$ in comparison with Cepheids of the final working sample in the projection on the plane XZ (top panel) and YZ (bottom panel); see text. Local maxima and lines of the largest values of ζ_4 at the sample boundaries are displayed in red, while local minima and lines of the smallest boundary values are shown in blue.

the range $-1.2 \leq Z \leq -0.8$ kpc (Figure 5, bottom panel), which agrees well with the disk level in this area according to other data—e.g., $-2 \leq Z \leq 1$ kpc (Skowron et al. 2019b, from Cepheids), $\bar{Z} = -1.22$ kpc (Romero-Gómez et al. 2019, from red giant branch (RGB) stars) and $-1 \leq Z \leq -0.5$ kpc (Lemasle et al. 2022, from Cepheids). The vertical dispersion of objects at large negative Y is consistent with that for the rest of the working sample (Figure 5, bottom panel). At the same time, the use of the entire working sample, despite the gaps, does not create any fictitious ripples in the model of the average surface of the disk in the area $Y < -6$ kpc (Figures 4, 6). Thus, there was no reason to discard Cepheids at large negative Y . In addition, it was important to keep in the sample objects representing the decline of the disk surface in the III and IV quadrants in order to test the capabilities of the general model (1) within the framework of the proposed method to describe this well-known feature together with other possible

details of the disk surface (as it turned out, local extremes). Note that the value of the decrease in the disk level of $\Delta Z \sim 1$ kpc relative to the plane $Z = z_{00}$ is much larger than the standard deviation for Cepheids ($\sigma_\rho \sim 130$ pc), i.e., the downward trend is detected confidently, despite the incompleteness of the sample in this area.

We tested the algorithm used here for choosing the optimal order of the ζ_n model by the Monte Carlo method. As a model of the disk surface, the constructed model ζ_4 was adopted (Table 1) and 100 mock catalogs were generated with object deviations from $\zeta_4(X, Y)$ along the normal to this surface, distributed according to the law (2) with the value of σ_ρ indicated in Table 1. The results are displayed in Figure 8. They show that in most cases the order of the initial model is restored exactly, and the probability that the order of the model will be underestimated is less than 1%. These results suggest that, acting according to this algorithm, it is possible to obtain a model that does not fully reflect some details of the real disk structure, but it is unlikely to build an excessively complex model with fictitious details.

To check the stability of the results, we divided the final working sample (hereinafter, for short, “full sample”) into two parts: one part included objects with odd numbers in the sample list (we will call it the odd subsample), and the other objects with even numbers (even subsample). This separation is actually random. On the other hand, it keeps the relative population of data of different longitude intervals approximately the same for both subsamples and for the full sample, since in the catalog Mel’nik et al. (2015) objects are ordered by their names, i.e., mainly by the names of constellations. The latter is important if we want to check the reproducibility of the detected details on the relief of the disk surface. The calculations were repeated for each of the two independent subsamples. The results are presented in Tables 5–7 and in Figure 9. The optimal orders of the model ζ_n for both subsamples turned out to be the same and equal to n_0 for the full sample: $n_0 = 4$. The parameters of the model ζ_4 for even and odd subsamples and for the full sample within the error limits are consistent with each other (cf., Tables 1, 5, 6). At the same time, all significant parameters obtained from subsamples are also significant for the full sample. The characteristics of the local extremes for the odd subsample turned out to be similar to the characteristics for the full sample (cf., Tables 3, 7). For the even subsample, the model $\zeta_4(X, Y)$ does not formally have local extremes, but it has areas of depression and elevation (Figure 9, right panel), in position and amplitude close to those for models obtained from the full sample and odd subsample (Figure 4; Figure 9, left panel). The curves-of-nodes also turned out to be similar for all three samples in the area of applicability of the model (Figures 4, 9). Thus, the topology of the resulting model as a whole is preserved even when the sample is divided. At the same time, the drop in the significance of the

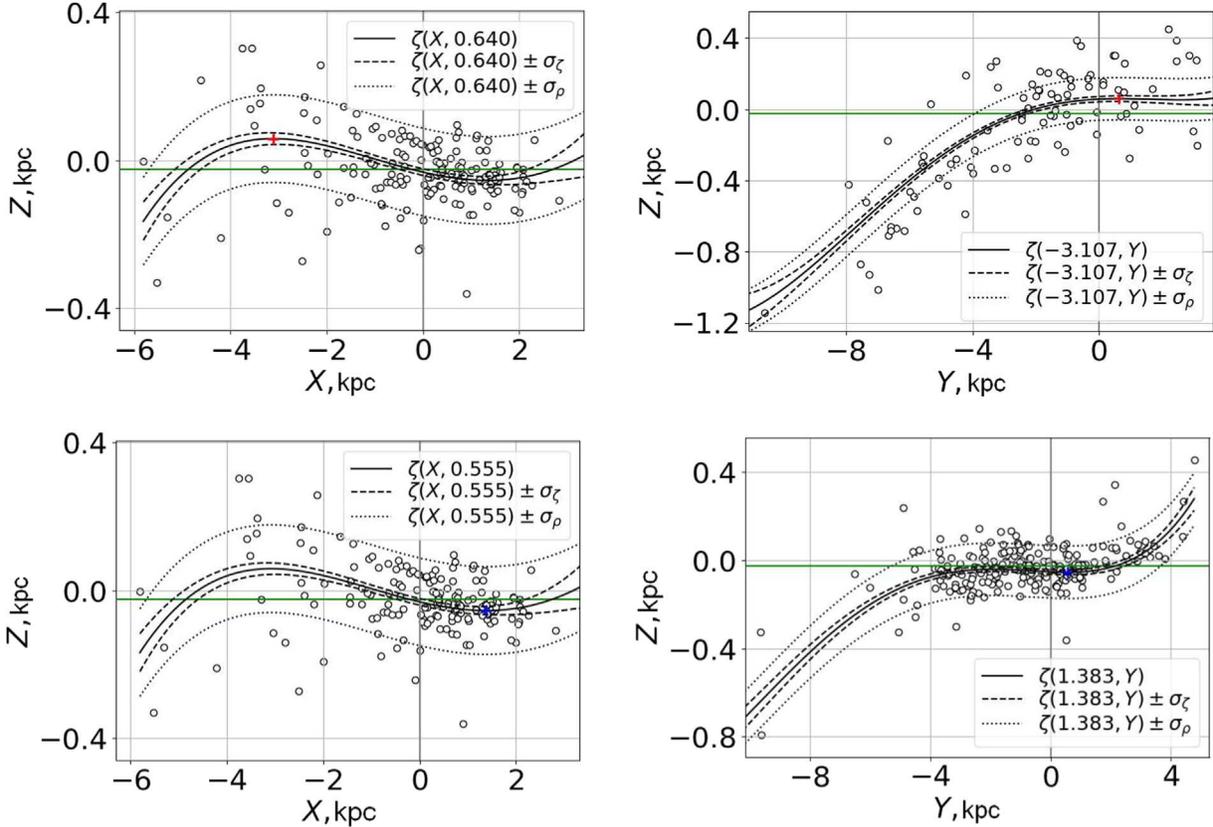


Figure 6. Sections of the model $\zeta_4(X, Y)$ of the middle surface of the disk passing through the local maximum (upper panels) and minimum (lower panels). Each panel features the confidence area for the model values $\zeta_4(X, Y) (\pm\sigma_\zeta)$ and the area of object deviations from the model $\pm 1\sigma_\rho$, as well as projections of the position of objects located in the ± 1 kpc band from the cut line. The green line indicates the position of the plane $Z = z_{00}$.

Table 3

Parameters of Local Extrema of the Optimal Model $\zeta_4(X, Y)$ Constructed for the Final Working Sample: The Cartesian Heliocentric Coordinates X, Y of the Extremum, the Value of $\zeta_4(X, Y)$ at the Extremum Point and the Significance Level $S_{\zeta_4}(X, Y)$ of the Extremum (the Number of σ)

Extremum	X (kpc)	Y (kpc)	$\zeta_4(X, Y)$ (pc)	$S_{\zeta_4}(X, Y)$
Local maximum	-3.1	0.6	58.4 ± 15.0	5.43
Local minimum	1.4	0.6	-55.0 ± 11.0	-2.81

details of the model surface $\zeta_4(X, Y)$ for subsamples confirms that dividing the sample into a larger number of parts is hardly meaningful.

5. Discussion

Consistency of local solar offset estimate $z_\odot = 28.1 \pm 6.1|_{\text{stat}}^{+1.3}_{-1.3}|_{\text{cal}}$ pc and global estimate $z_\odot = 27.1 \pm 8.8|_{\text{stat}}^{+1.3}_{-1.2}|_{\text{cal}}$ pc (Table 2) also implies that the proposed algorithm is valid and the model order is correct—when optimizing the order of the model, the value of the solar offset does not depend systematically on whether a large or small

neighborhood of the Sun is considered. Our estimates also agree with the best estimate of the solar offset $z_\odot = 25 \pm 5$ pc (Bland-Hawthorn & Gerhard 2016) and local estimates specifically for classical Cepheids (see below).

However, the estimate of $\sigma_\rho = 68.5 \pm 3.9$ pc obtained for the final local sample is inconsistent with the estimate of $\sigma_\rho = 118.7 \pm 3.3$ pc for the final working sample (in original distance scale, see Table 2). Since the estimates were obtained by optimizing the order of the model ζ_n , this mismatch should mainly be a consequence of a combination of the flaring and random errors in distances. Indeed, the observed deviation of objects relative to the model ζ_{n_0} is due to two factors: the *natural* (true, cosmic) dispersion of objects relative to the average surface of the Galactic disk and the measuring dispersion—the deviation of the observed positions of objects from their true positions due to random errors in distance moduli. The latter means that more distant objects have larger distance errors. On the other hand, the natural dispersion can grow to the periphery (the disk flaring). Indeed, at $X < 0$ kpc, the apparent spread of objects relative to the model increases somewhat (Figure 6, left panels). In order to separate the

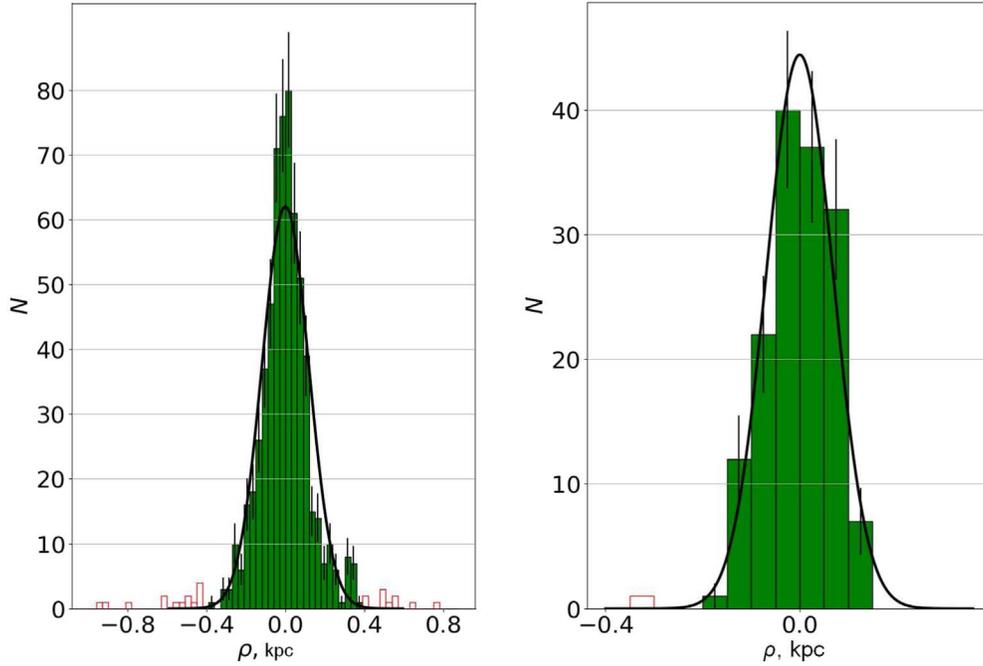


Figure 7. Observed (green columns) and model (black line) distributions of object deviations ρ along the normal to the model average surface for the working sample (left panel, for $n_o = 4$ and the model parameters given in Table 1) and local sample (right panel, for $n_o = 1$ and the parameters given in Table 2). Outliers are shown in red.

Table 4

Standard Deviations of Cepheids Across the Disk, Calculated for Objects of the Considered Samples with only Positive Deviations ρ from the Model, σ_ρ^+ , and with only Negative Ones, σ_ρ^-

Sample	σ_ρ^+ (kpc)	σ_ρ^- (kpc)
Working sample	0.1621 ± 0.0045	0.1641 ± 0.0046
Final working sample	0.1277 ± 0.0053	0.1162 ± 0.0048
Local sample	0.0765 ± 0.0043	0.0894 ± 0.0046
Final local sample	0.0757 ± 0.0043	0.0792 ± 0.0043

contributions of these two effects, a more complex version of the present method is required with direct consideration of random errors in distances. This will also provide more reliable results about flaring itself.

Note that the contribution of the measuring dispersion to the observed dispersion σ_ρ in this case is very small, as shown by the following approximate estimates. For photometric distances, their standard deviation due to measuring dispersion is $\sigma_r = \frac{\ln 10}{5} \sigma_d r$, where σ_d is the standard uncertainty of distance moduli. In the case of local sample, the main contribution of distance errors to the observed vertical standard deviation σ_Z (close to σ_ρ) occurs due to objects at high latitudes b . Then the assumption that the standard σ_r for *all* objects of the sample completely passes into vertical standard deviation gives an upper estimate for the contribution of the measuring dispersion to σ_Z : $\sigma_{Z,\text{mes}} < \sigma_r$. For the distance $r = 1\sigma_\rho = 76.5$ pc (see

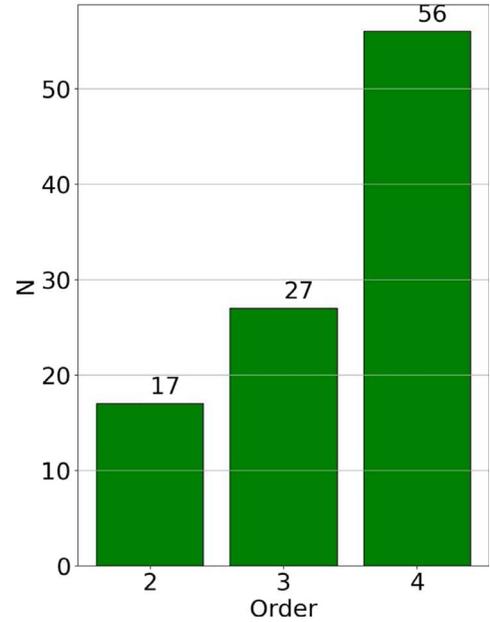


Figure 8. Distribution of optimal order values for 100 mock samples (see text).

Table 2) and $\sigma_d = 0.14^m$ (Veselova & Nikiforov 2020), this results in $\sigma_{Z,\text{mes}} < 4.9$ pc and the natural dispersion $\sigma_{Z,0} = \sqrt{\sigma_Z^2 - \sigma_{Z,\text{mes}}^2} > 76.3$ pc, i.e., the correction is no more than -0.2% .

Table 5
Same as in Table 1, but for the Odd Subsample of the Final Working Sample (see text). $n_o = 4$

Parameter	Estimation	$\sigma_{z_i}/ z_i $	Parameter	Estimation	$\sigma_{z_i}/ z_i $
σ_z	0.1218 ± 0.0050	0.04	z_{21}	0.0017 ± 0.0013	0.76
z_{00}	-0.026 ± 0.011	0.42	z_{12}	0.002 14 ± 0.00080	0.37
z_{10}	-0.0321 ± 0.0093	0.29	z_{03}	0.001 45 ± 0.00040	0.28
z_{01}	-0.0077 ± 0.0061	0.79	z_{40}	0.000 25 ± 0.00022	0.88
z_{20}	0.0074 ± 0.0039	0.53	z_{31}	0.000 26 ± 0.00021	0.81
z_{11}	-0.0002 ± 0.0026	13.00	z_{22}	-0.00008 ± 0.00018	2.25
z_{02}	-0.0013 ± 0.0014	1.08	z_{13}	0.000 133 ± 0.000065	0.49
z_{30}	0.0038 ± 0.0015	0.39	z_{04}	0.000 092 ± 0.000029	0.32

Table 6
Same as in Table 1, but for the Even Subsample of the Final Working Sample (see Text). $n_o = 4$

Parameter	Estimation	$\sigma_{z_i}/ z_i $	Parameter	Estimation	$\sigma_{z_i}/ z_i $
σ_z	0.1227 ± 0.0050	0.04	z_{21}	0.0037 ± 0.0015	0.41
z_{00}	-0.028 ± 0.011	0.39	z_{12}	0.002 49 ± 0.00086	0.35
z_{10}	-0.0279 ± 0.0097	0.35	z_{03}	0.002 10 ± 0.00044	0.21
z_{01}	-0.0167 ± 0.0060	0.36	z_{40}	-0.00030 ± 0.00020	0.67
z_{20}	0.0046 ± 0.0036	0.78	z_{31}	0.000 10 ± 0.00020	2.00
z_{11}	0.0011 ± 0.0028	2.55	z_{22}	0.000 35 ± 0.00016	0.46
z_{02}	0.0011 ± 0.0016	1.45	z_{13}	0.000 125 ± 0.000088	0.70
z_{30}	0.0008 ± 0.0014	1.75	z_{04}	0.000 097 ± 0.000033	0.34

Table 7

Same as in Table 3, but for the Odd Subsample of the Final Working Sample (see Text)

Extremum	X (kpc)	Y (kpc)	$\zeta_4(X, Y)$ (pc)	$S_{\zeta_4}(X, Y)$
Local maximum	-3.0	0.1	53.0 ± 20	4.01
Local minimum	1.0	1.0	-50.4 ± 15	-1.59

In the case of a working sample, most of the objects are located at small $|b|$: for characteristic distances $r = 3-6$ pc (see Figure 1) and $Z = 1\sigma_\rho = 132$ pc (Table 2) $|b| = 1^\circ.3-2^\circ.5$. For sample objects, the contribution of the measuring dispersion is $\sigma_{Z,\text{mes}} = \sigma_r \sin|b|$, $\tan b = Z/(r \cos b)$, then for small $|b|$, $b \approx Z/r$, $\sigma_{Z,\text{mes}} \approx \sigma_r |b| = \frac{\ln 10}{5} \sigma_d |Z|$. Then for $Z = 1\sigma_\rho = 132$ pc it turns out: $\sigma_{Z,\text{mes}} \approx 8.5$ pc, $\sigma_{Z,0} \approx 131.7$ pc, i.e., correction -0.2% . Thus, in both cases, the contribution of random distance uncertainty to the observed vertical dispersion is negligible.

From classical Cepheids on $r < 2$ kpc Majaess et al. (2009) obtained $z_\odot = 26 \pm 3$ pc and the scale height of $\leq 75 \pm 10$ pc which are consistent with our estimates (Table 2). Estimates of $z_\odot = (23-24) \pm 2$ pc and $\sigma_\rho = 76.4 \pm 1.8$ pc were found by Bobylev & Bajkova (2016a) for classical Cepheids according to the same version of the Berdnikov et al.'s catalog as in this work. Bobylev & Bajkova (2016a) considered the cylindrical region $r \leq 4$ pc. As they used the original calibration of the catalog we can compare these estimates with ours in the same calibration (Table 2). There is consistency with our estimates of

the solar offset for both final local and final working samples. However, the vertical scale estimates are consistent only in the case of final local sample. Exactly the same situation is for Cepheids-based estimates in Skowron et al. (2019a). The authors considered data on 2431 Cepheids, and the most part of the data was obtained by the OGLE-IV project. Skowron et al. obtained an estimate of the disk scale height of 73.5 ± 3.2 pc, so our estimate for the final local sample does not contradict this value. All this also points out the importance of taking into account the warp of the Galactic average disk surface in order to have the ability for proper consideration of all the data available. Otherwise, only local regions can be considered.

In addition, the fact that the estimates of z_\odot differ, as noted in Introduction, also suggests that the a priori assumption about the flat model of the Galaxy's disk should be limited. Indeed, according to our results such assumption might be made only for specific regions like the local one, i.e., close enough to the Sun. Moreover, it can be noticed that the value of z_\odot is less dependent on the Galactic disk warping than the value of σ_ρ . Based on what has been said, we can conclude that any a priori assumption about the Galactic disk warping must be carefully studied, especially when the vertical scale parameter is estimated.

The detected local extrema of the average surface of the disk may be manifestations of bending waves caused by interaction with the Sagittarius dwarf galaxy, in the form of local structures elongated in the azimuthal direction (Gómez et al. 2013; Figure 5; Laporte et al. 2019; Poggio et al. 2021; Figure 2), or by

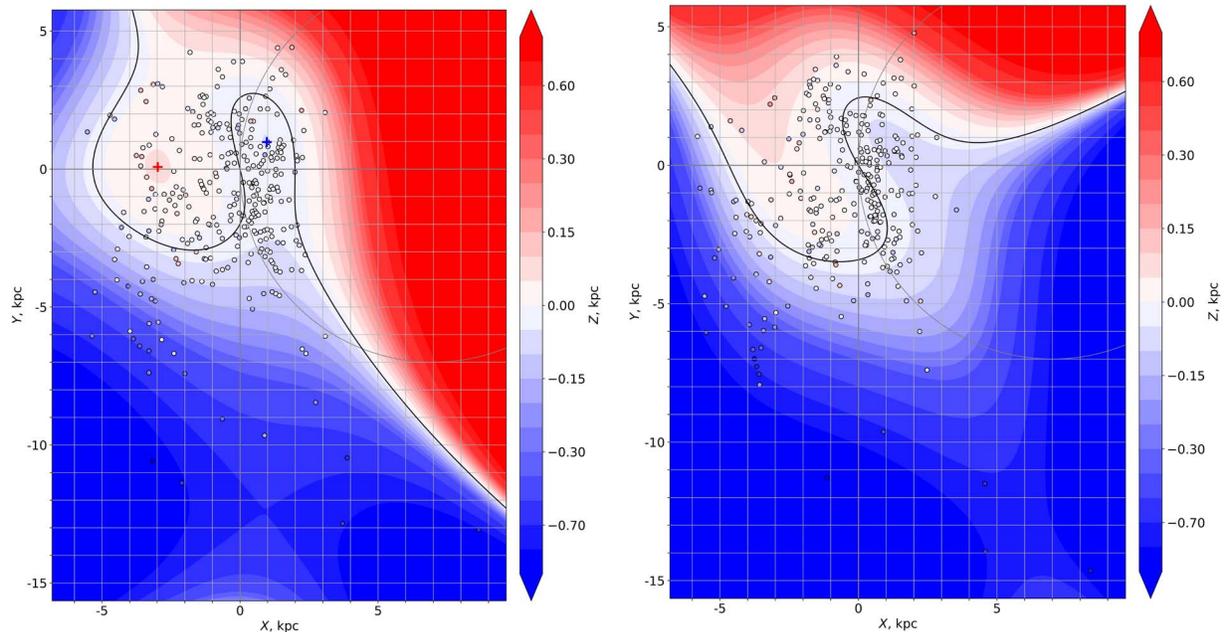


Figure 9. Maps of the models ζ_4 of the average disk surface of optimal order, constructed from the objects of independent odd (left panel) and even (right panel) subsamples of the final working sample (see text). The notation is the same as in Figure 4.

interaction with the LMC (Thulasidharan et al. 2022 and references therein).

The method used after testing on classical Cepheids can now be applied to other data (in particular, to Gaia data) and/or in other assumptions about the distribution function $f(\rho)$. The Gaia Data Release 2 (DR2) catalog was used in recent work by Ablimit et al. (2020) to obtain data on classical Cepheids. Despite the fact that the direct study of Galactic disk warping was not conducted in the work of these authors, according to the pictures plotted in the work on these data, the disk warping is clearly revealed. Unfortunately, the use of the current version of our method with these data as is will lead to significant biases mainly because of the dependence of distance uncertainty on distance, which will be significant due to the need to consider the large neighborhood of the Sun.

Taking into account the uncertainty of distances may also solve the problem of establishing the form of the vertical distribution law $f(\rho)$. Note that the analysis of the 2D distribution does not allow us to draw a definite conclusion about the functional form of this law (Mosenkov et al. 2021).

In the future, we plan to apply the proposed method in the variant of accounting for distance uncertainties to new databases.

Acknowledgments

We are grateful for the helpful comments from the anonymous referee.

References

- Ablimit, I., Zhao, G., Flynn, C., & Bird, S. A. 2020, *ApJ*, **895**, L12
- Berdnikov, L., Dambis, A., & Vozyakova, O. 2000, *A&AS*, **143**, 211
- Binney, J., & Merrifield, M. 1998, *Galactic Astronomy* (Princeton, NJ: Princeton Univ. Press)
- Bland-Hawthorn, J., & Gerhard, O. 2016, *ARA&A*, **54**, 529
- Bobylev, V., & Bajkova, A. 2016a, *Astron. Let.*, **42**, 1
- Bobylev, V., & Bajkova, A. 2016b, *Astron. Let.*, **42**, 182
- Bobylev, V., & Bajkova, A. 2017, *Astron. Let.*, **43**, 304
- Burke, B. F. 1957, *AJ*, **62**, 90
- Buckner, A. S. M., & Froebrich, D. 2014, *MNRAS*, **444**, 290
- Cheng, X., Anguiano, B., Majewski, S. R., et al. 2020, *ApJ*, **905**, 49
- Chrobáková, Ž., & López-Corrodoira, M. 2021, *ApJ*, **912**, 130
- Chrobáková, Ž., Nagy, R., & López-Corrodoira, M. 2020, *A&A*, **637**, 96
- Dambis, A. K., Berdnikov, L. N., Efremov, Yu. N., et al. 2015, *Astron. Lett.*, **41**, 489
- Efremov, Yu. N. 2011, *ARep*, **55**, 108
- de Grijs, R., Wicker, J., & Bono, G. 2014, *AJ*, **147**, 122
- Ferguson, D., Gardner, S., & Yanny, B. 2017, *ApJ*, **843**, 141
- Gómez, F. A., Minchev, I., O’Shea, B. W., et al. 2013, *MNRAS*, **429**, 159
- Hudson, D. J. 1964, *Statistics: Lectures on Elementary Statistics and Probability* (Geneva: CERN)
- Jurić, M., Ivezić, Ž., Brooks, A., Lupton, R., & Schlegel, D. 2008, *ApJ*, **673**, 864
- Kerr, F. J. 1957, *AJ*, **62**, 93
- Khachatryan, T., Beraldo e Silva, L., & Debattista, V. P. 2021, *MNRAS*, **508**, 2350
- Khachatryan, T., Beraldo e Silva, L., Debattista, V. P., & Daniel, K. J. 2022, *MNRAS*, **512**, 3500
- Laporte, C. F. P., Minchev, I., Johnston, K. V., & Gómez, F. A. 2019, *MNRAS*, **485**, 3134
- Lemasle, B., Lala, H. N., Kovtyukh, V., et al. 2022, *A&A*, **668**, 40
- Li, X.-Y., Huang, Y., Chen, B.-Q., et al. 2020, *ApJ*, **901**, 56
- Majaess, D. J., Turner, D. G., & Lane, D. J. 2009, *MNRAS*, **398**, 263
- Mel’nik, A., Rautiainen, P., Berdnikov, L., Dambis, A., & Rastorguev, A. 2015, *AN*, **336**, 70

- Minniti, J. H., Zoccali, M., Rojas-Arriagada, A., et al. 2021, *A&A*, **654**, A138
- Mosenkov, A. V., Savchenko, S. S., Smirnov, A. A., & Camps, P. 2021, *MNRAS*, **507**, 5246
- Nikiforov, I. I. 2012, *A&AT*, **27**, 537
- Nikiforov, I. I., & Veselova, A. V. 2018, *Astron. Lett.*, **44**, 81
- Oort, J., Kerr, F., & Westerhout, G. 1958, *MNRAS*, **118**, 379
- Poggio, E., Drimmel, R., Andrae, R., et al. 2020, *NatAst*, **4**, 590
- Poggio, E., Laporte, C. F. P., Johnston, K. V., et al. 2021, *MNRAS*, **508**, 541
- Pohl, M., Englmaier, P., & Bissantz, N. 2008, *AJ*, **677**, 283
- Pojmanski, G. 2002, *AcA*, **52**, 397
- Reid, M. J., Menten, K. M., Brunthaler, A., et al. 2019, *AJ*, **885**, 131
- Romero-Gómez, M., Mateu, C., Aguilar, L., Figueras, F., & Castro-Ginard, A. 2019, *A&A*, **627**, A150
- Samus, N. N., Kazarovets, E. V., Durlevich, O. V., et al. 2017, *ARep*, **61**, 80
- Schönrich, R., & Dehnen, W. 2018, *MNRAS*, **478**, 3809
- Skowron, D., Skowron, J., Mróz, P., et al. 2019a, *Sci*, **365**, 478
- Skowron, D., Skowron, J., Mróz, P., et al. 2019b, *AcA*, **69**, 305
- Spicker, J., & Feitzinger, J. V. 1986, *A&A*, **163**, 43
- Thulasidharan, L., D’Onghia, E., Poggio, E., et al. 2022, *A&A*, **660**, L12
- van Tulder, J. 1942, *BAN*, **9**, 315
- Veselova, A. V., & Nikiforov, I. I. 2020, *RAA*, **20**, 209
- Wall, J., & Jenkins, C. 2012, *Practical Statistics for Astronomers* (2nd edn.; New York: Cambridge Univ. Press)
- Xu, Y., Newberg, H. J., Carlin, J. L., et al. 2015, *ApJ*, **801**, 105
- Yao, J., Manchester, R., & Wang, N. 2017, *MNRAS*, **468**, 3289