Estimation of Parameter Errors in Inverse Problems. Determining Limb-Darkening Coefficients in Classical Eclipsing Systems

M. K. Abubekerov, N. Yu. Gostev, and A. M. Cherepashchuk

Sternberg Astronomical Institute, Moscow State University, Moscow, Russia
Received December 15, 2008; in final form, December 20, 2008

Abstract—The estimation of parameters and their errors is considered using the observed light curve of the eclipsing binary system YZ Cas as an example. The error intervals are calculated using the differential-correction and confidence-region methods. The error intervals and reliability of the methods are investigated, and the reliability of limb-darkening coefficients derived from the observed light curve analyzed. A new method for calculating parameter errors is proposed.

1. INTRODUCTION

A comparative analysis of various methods for estimating parameter errors for one-, two-, three-, and four-parameter functions and the observed light curve of the eclipsing binary system YZ Cas was carried out in [1] using numerical simulations. The numerical simulations yielded estimates of the unknown parameters of the problem and their error intervals, which were calculated using the differential-correction, Monte Carlo, and confidence-region methods. The error intervals and reliability of the various methods were studied.

The current study is a continuation of the analysis begun in [1]. In contrast to the current study, in [1], the trustworthiness of the methods used to estimate the error intervals were evaluated using an incomplete set of parameters. The analysis of the observed light curve for the binary system YZ Cas was based on three parameters: the radius of the primary and secondary components r_1 and r_2 and the orbital inclination of the binary i. In addition, the residual used depended on these three parameters and was minimized over the remaining, linear parameters, which included the limb-darkening coefficients for the primary and secondary stars x_1 and x_2 and the ratio of the brightnesses at the centers of the stellar disks, $I_0^{(1)}/I_0^{(2)}$. It is important that minimizing over linear parameters does not change the statistics of the problem, only decreases the number of degrees of freedom. This enabled us to analyze the errors in the non-linear parameters r_1, r_2, and i in a mathematically rigorous way. The error intervals were found as projections onto the r_1, r_2, and i parameter axes of the three-dimensional confidence region containing the values

of r_1, r_2, and i for which the residual was less than the quantile corresponding to a specified confidence level.

In the current study, in addition to the parameters listed above, the limb-darkening coefficients for the primary and secondary stars are unknown. We obtain fits to the observed light curve of YZ Cas using both linear and quadratic limb-darkening laws.

We focus on the reliability of the limb-darkening coefficients and their error intervals. This problem is especially relevant in connection with the appearance of high-accuracy photometric observations from satellites such as COROT and KEPLER, which will soon lead to the availability of massive databases of precise photometric observations. Therefore, it is important to know the reliability of estimates of photometric elements of both classical eclipsing binary systems and eclipsing systems with exoplanets at specified levels of accuracy.

As was emphasized in [1], in both the classical least-squares method (or the differential-correction method for non-linear problems) and the Monte Carlo method, the errors of unknown parameters are estimated in the framework of a normal distribution, which are obeyed by the central values of parameters, and searches for the true values of the parameters are carried out using statistics that arise from a normal distribution (such as χ_M^2 statistics, where M is the number of points in the observed light curve). However, it is most reasonable to calculate the errors of parameters using the same statistics that were used to search for the central values of the parameters themselves (for example, the statistics for a residual distributed as χ_M^2). Although the resulting parameter errors will be larger than the errors found using the

least-squares or Monte Carlo method, an approach based on the use of confidence regions provides the most trustworthy fits to observational data; it also avoids the need to separately consider "internal" and "external" errors in parameters.

2. MODEL FOR A CLASSICAL BINARY SYSTEM AND A BINARY SYSTEM WITH AN EXOPLANET

In view of the methodological character of our work, we consider a simple model with two spherical stars with thin atmospheres in circular orbits without any effects due to the proximity of the components. This model can easily be realized on modern computers, and makes it possible to obtain a large number of solutions of the problem with a comparatively small amount of computer time. This type of spherical-star model for a binary system is physically based on the case when the degree of Roche-lobe filling is small, $\mu < 0.5$. This occurs in the YZ Cas system, and also in most cases of observed eclipsing stars with exoplanets.

Our model considers the motion of the stellar disks projected onto the plane of the sky, i.e., perpendicular to the line of sight. Figure 1 shows the geometry of the stellar disks during eclipse. Here, r_1 and r_2 are the radii of the primary and secondary star (or of the star and planet), Δ is the distance between the centers of the stellar disks, and ρ and Ψ are polar coordinates of an arbitrary point on the surface of the disk of the primary (the coordinate origin is located at the geometrical center of the disk). The distance between the centers of the stellar disks is given by the expression

$$\Delta^2 = \cos^2 i + \sin^2 i \sin^2 \theta, \qquad (1)$$

where i is the orbital inclination of the binary system and θ is the current value of the orbital phase angle (the orbit of the system is circular with radius $a = 1$).

For the brightness distribution over the disk of each star, we used a linear limb-darkening law,

$$I(\rho) = I_0 \left(1 - x + x\sqrt{1 - \frac{\rho^2}{r^2}} \right) \qquad (2)$$

and also a quadratic limb-darkening law, which differs from the linear law in the presence of a non-zero quadratic limb-darkening coefficient y:

$$I(\rho) = I_0 \left[1 - x \left(1 - \sqrt{1 - \frac{\rho^2}{r^2}} \right) \right.$$
$$\left. - y \left(1 - \sqrt{1 - \frac{\rho^2}{r^2}} \right)^2 \right]. \qquad (3)$$

Here, ρ is the polar distance from the center of the stellar disk and r the radius of the stellar disk. The brightness was assumed to be zero at any point in the disk of the planet.

We denote $I_0^{(1)}$ and $I_0^{(2)}$ to be the brightnesses at the disk centers, x_1 and x_2 the limb-darkening coefficients, and y_1 and y_2 the quadratic limb-darkening coefficients of the primary and secondary stars.

The unknown parameters of the model for the two stars are r_1, r_2, i, $I_0^{(1)}$, $I_0^{(2)}$, x_1, x_2, and, in the case of the non-linear limb-darkening law, y_1 and y_2.

The unit length in our models is the distance between the stellar centers (or the centers of the star and planet), $a = 1$. There is no "third light" in the model. The light curve of the binary system is determined by the following three equations.

1. The total luminosity of the components, which describes the out-of-eclipse brightness:

$$2\pi \int_0^{r_1} I^{(1)}(\rho)\rho d\rho + 2\pi \int_0^{r_2} I^{(2)}(\xi)\xi d\xi = L^{full}. \qquad (4)$$

2. The loss of brightness in the first (secondary, in the case of YZCas) minimum, due to the eclipse of the smaller star by the larger companion:

$$L^{full} - L^{(1)}(\theta) = \iint_{S(\Delta)} I^{(2)}(\xi)dS, \qquad (5)$$

where $S(\Delta)$ is the area of overlap of the two disks.

3. The loss of brightness in the second (primary) minimum, due to the eclipse of the larger star by the smaller star:

$$L^{full} - L^{(2)}(\theta) = \iint_{S(\Delta)} I^{(1)}(\rho)dS. \qquad (6)$$

Equations (1), (4), (5), and (6) fully describe the observed light curve, and contain a set of parameters that depends on the model considered: r_1, r_2, i, $I_0^{(1)}$, $I_0^{(2)}$, x_1, x_2, y_1, y_2. Substituting the brightness distribution function approximated by the corresponding limb-darkening law (2) or (3) into the integral and carrying out the integration yields a system of non-linear, algebraic equations in the corresponding parameters.
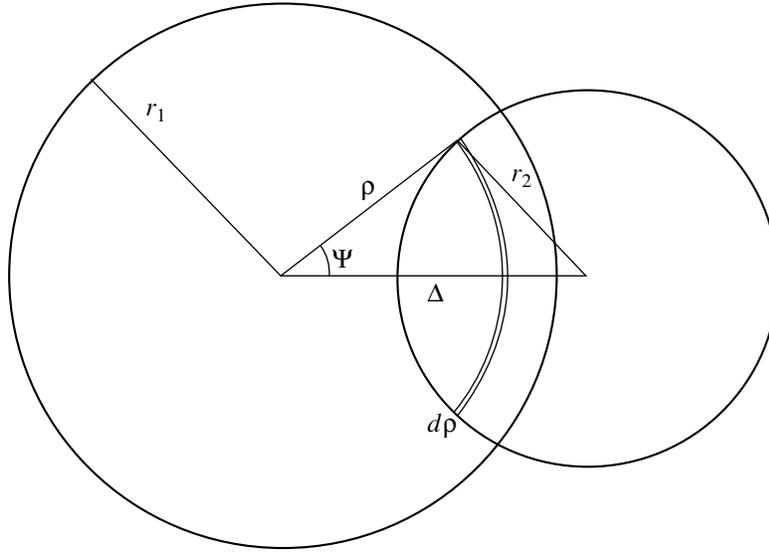
Fig. 1. Model for two eclipsing spherical stars, projected onto the plane of the sky.

3. DERIVATION OF THE COMPUTATIONAL FORMULAS

In the linear limb-darkening law, the brightness at a point in the stellar disk is

$$I(\rho) = I_0 \left(1 - x + x\sqrt{1 - \frac{\rho^2}{r^2}} \right),$$

where ρ is the distance of the point from the disk center, I_0 the brightness at the disk center, x the limb-darkening coefficient, and r the radius of the star.

Introducing the new parameters

$$\left. \begin{array}{c} X_0 = I_0(1 - x) \\ X_1 = I_0 x \end{array} \right\}, \qquad (7)$$

we find that

$$I(\rho) = \left(X_0 + X_1\sqrt{1 - \frac{\rho^2}{r^2}} \right). \qquad (8)$$

In the non-linear limb-darkening law, the brightness at a point in the stellar disk is

$$I(\rho) = I_0 \left[1 - x\left(1 - \sqrt{1 - \frac{\rho^2}{r^2}} \right) \right.$$
$$\left. -y\left(1 - \sqrt{1 - \frac{\rho^2}{r^2}} \right)^2 \right],$$

where ρ is the distance of the point from the disk center, I_0 the brightness at the disk center, x, y the

limb-darkening coefficients, and r the radius of the star. Introducing the new parameters

$$\left. \begin{array}{c} X_0 = I_0(1 - x - 2y) \\ X_1 = I_0(x + 2y) \\ X_2 = I_0 y \end{array} \right\}, \qquad (9)$$

we obtain

$$I(\rho) = \left(X_0 + X_1\sqrt{1 - \frac{\rho^2}{r^2}} + X_2\frac{\rho^2}{r^2} \right).$$

In these variables, the brightness at a point on the stellar disk for the non-linear limb-darkening law differs from the corresponding brightness for the linear law in the presence of an additional term containing X_2. This enables us to obtain a linear model by setting $X_2 = 0$ in the non-linear model. At orbital phase $\theta = 0$, component 1 eclipses component 2. The total brightness of the star in the non-linear model is

$$L^{(s)} = 2\pi \int_0^r I^{(s)}(\rho)\rho d\rho = X_0^{(s)}\pi r_s^2 + \frac{2}{3}X_1^{(s)}\pi r_s^2$$

$$+ \frac{1}{2}X_2^{(s)}\pi r_s^2, \qquad s = 1, 2.$$

The total brightness of the system of two stars in the non-linear model outside of eclipse is

$$L^{full} = L^{(1)} + L^{(2)} = X_0^{(1)}\pi r_1^2 \qquad (10)$$

$$+ \frac{2}{3}X_1^{(1)}\pi r_1^2 + \frac{1}{2}X_2^{(1)}\pi r_1^2 + X_0^{(2)}\pi r_2^2$$

$$+ \frac{2}{3}X_1^{(2)}\pi r_2^2 + \frac{1}{2}X_2^{(2)}\pi r_2^2.$$

To obtain a model for a system consisting of a star and planet, it is sufficient to set $L^{(2)} = 0$ (the planet does not radiate).

To universalize the external form of the computational formulas and the light-curve minima, and to decrease the number of equations needed, the eclipsing component (near component relative to a terrestrial observer) is ascribed the subscript n and the eclipsed component (farther component relative to a terrestrial observer) the subscript f. When directly computing the light-curve minima at orbital phases $-\pi/2 < \theta < \pi/2$ (or $\cos\theta > 0$), we must replace the variable r_1 with r_n, and the variable r_2 with r_f. At orbital phases $\cos\theta < 0$, we must carry out the opposite substitution, replacing the variable r_1 with r_f, and the variable r_2 with r_n.

In the new notation, the drop in brightness during eclipse is

$$L^{dec}(\Delta, r_f, r_n, X_0^{(f)}, X_1^{(f)}, X_2^{(f)}) \qquad (11)$$
$$= \iint_{S(\Delta)} I^{(f)}(\rho)dS,$$

where Δ is the distance between the disk centers and $S(\Delta)$ the area of overlap of the disks.

To compute the integral (11), as was done in [1], we introduce the functions

$$\mathcal{A}x \equiv \begin{cases} \pi, & x < -1, \\ \arccos x, & -1 \le x \le 1, \\ 0, & x > 1 \end{cases} \qquad (12)$$

and

$$\mathcal{Q}x \equiv \begin{cases} \sqrt{x}, & x \ge 0, \\ 0, & x < 0, \end{cases}$$

$$\Psi(\Delta, x, y) \equiv \mathcal{A}\left(\frac{x^2 + \Delta^2 - y^2}{2x\Delta}\right),$$

$$Q(\Delta, r_f, r_n)$$
$$\equiv \mathcal{Q}\left\{\left[r_f^2 - (\Delta - r_n)^2\right]\left[(\Delta + r_n)^2 - r_f^2\right]\right\},$$

together with a polar-coordinate system with its origin at the center of the disk of the eclipsed star and the polar angle φ measured from the center of the disk of the eclipsed component "f" to the center of the disk of the eclipsing component "n" (Fig. 1). We then have

$$L^{dec}(\Delta, r_f, r_n, X_0^{(f)}, X_1^{(f)}, X_2^{(f)}) \qquad (13)$$
$$= X_0^{(f)} L_0^{dec}(\Delta, r_f, r_n) + X_1^{(f)} L_1^{dec}(\Delta, r_f, r_n)$$
$$+ X_2^{(f)} L_2^{dec}(\Delta, r_f, r_n),$$

and the expressions for L_0^{dec} and L_1^{dec} obtained in [1]:

$$L_0^{dec}(\Delta, r_f, r_n) = \Psi(\Delta, r_f, r_n)r_f^2 \qquad (14)$$
$$+ \Psi(\Delta, r_n, r_f)r_n^2 - \frac{1}{2}Q(\Delta, r_f, r_n),$$

$$L_1^{dec}(\Delta, r_f, r_n) \qquad (15)$$

$$= \int_0^{r_f^2} \Psi(\Delta, \sqrt{\rho}, r_n)\sqrt{1 - \frac{\rho}{r_f^2}}d\rho.$$

Analogous to the way in which L_0^{dec} and L_1^{dec} were obtained in [1], we obtain for L_2^{dec}:

$$L_2^{dec}(\Delta, r_f, r_n) = 2\int_0^{r_f} \rho^3 \Psi(\Delta, \rho, r_n)d\rho \qquad (16)$$

$$= \Psi(\Delta, r_f, r_n)\frac{r_f^4}{2} + \frac{r_n^2}{2}\left(2\Delta^2 + r_n^2\right)\Psi(\Delta, r_n, r_f)$$
$$- \frac{1}{8}\left(\Delta^2 + 5r_n^2 + r_f^2\right)Q(\Delta, r_f, r_n).$$

The partial derivatives of L_2^{dec} have the form

$$\frac{\partial L_2^{dec}(\Delta, r_f, r_n)}{\partial \Delta} = 2\Delta r_n^2 \Psi(\Delta, r_n, r_f) \qquad (17)$$
$$- \frac{\Delta^2 + r_n^2 + r_f^2}{2\Delta}Q(\Delta, r_f, r_n),$$

$$\frac{\partial L_2^{dec}(\Delta, r_f, r_n)}{\partial r_f} = 2r_f^3 \Psi(\Delta, r_f, r_n)r_f, \qquad (18)$$

$$\frac{\partial L_2^{dec}(\Delta, r_f, r_n)}{\partial r_n} = 2r_n\left(\Delta^2 + r_n^2\right) \qquad (19)$$
$$\times \Psi(\Delta, r_n, r_f) - 2r_n Q(\Delta, r_f, r_n).$$

For a circular orbit, the distance between the disk centers Δ depends on the phase θ and orbital inclination i as

$$\Delta(\theta, i) = \sqrt{\cos^2 i + \sin^2 i \sin^2 \theta}. \qquad (20)$$

The brightness of the binary system depends on the phase θ as follows:

$$L(\theta, i, r_1, r_2, X_0^{(1)}, X_1^{(1)}, X_0^{(2)}, X_1^{(2)}) \qquad (21)$$
$$= L^{full} - \begin{cases} L^{dec}(\Delta(\theta, i), r_1, r_2, X_0^{(1)}, X_1^{(1)}, X_2^{(1)}), \\ \cos\theta < 0, \\ L^{dec}(\Delta(\theta, i), r_2, r_1, X_0^{(2)}, X_1^{(2)}, X_2^{(2)}), \\ \cos\theta > 0. \end{cases}$$

Introducing the functions

$$L_0^{(1)}(\theta, r_1, r_2, i) \qquad (22)$$

$$= \pi r_1^2 - \begin{cases} L_0^{dec}(\Delta(\theta, i), r_1, r_2), & \cos\theta < 0, \\ 0, & \cos\theta > 0, \end{cases}$$

$$L_1^{(1)}(\theta, r_1, r_2, i)$$

$$= \frac{2}{3}\pi r_1^2 - \begin{cases} L_1^{dec}(\Delta(\theta, i), r_1, r_2), & \cos\theta < 0, \\ 0, & \cos\theta > 0, \end{cases}$$

$$L_2^{(1)}(\theta, r_1, r_2, i)$$

$$= \frac{1}{2}\pi r_1^2 - \begin{cases} L_2^{dec}(\Delta(\theta, i), r_1, r_2), & \cos\theta < 0, \\ 0, & \cos\theta > 0, \end{cases}$$

$$L_0^{(2)}(\theta, r_1, r_2, i)$$

$$= \pi r_2^2 - \begin{cases} L_0^{dec}(\Delta(\theta, i), r_2, r_1), & \cos\theta > 0, \\ 0, & \cos\theta < 0, \end{cases}$$

$$L_1^{(2)}(\theta, r_1, r_2, i)$$

$$= \frac{2}{3}\pi r_2^2 - \begin{cases} L_1^{dec}(\Delta(\theta, i), r_2, r_1), & \cos\theta > 0, \\ 0, & \cos\theta < 0, \end{cases}$$

$$L_2^{(2)}(\theta, r_1, r_2, i)$$

$$= \frac{1}{2}\pi r_2^2 - \begin{cases} L_2^{dec}(\Delta(\theta, i), r_2, r_1), & \cos\theta > 0, \\ 0, & \cos\theta < 0, \end{cases}$$

we can write the total brightness as the linear combination

$$L(\theta, r_1, r_2, i, X_0^{(1)}, X_1^{(1)}, X_0^{(2)}, X_1^{(2)}) \quad (23)$$
$$= X_0^{(1)} L_0^{(1)}(\theta, r_1, r_2, i) + X_1^{(1)} L_1^{(1)}(\theta, r_1, r_2, i)$$
$$+ X_2^{(1)} L_2^{(1)}(\theta, r_1, r_2, i) + X_0^{(2)} L_0^{(2)}(\theta, r_1, r_2, i)$$
$$+ X_1^{(2)} L_1^{(2)}(\theta, r_1, r_2, i) + X_2^{(2)} L_2^{(2)}(\theta, r_1, r_2, i).$$

The partial derivatives of a function of the total brightness are expressed in terms of the partial derivatives of the functions L_0^{dec}, L_1^{dec}, and L_2^{dec} with respect to the variables Δ, r_f, and r_n and the partial derivatives of $\Delta(\theta, i)$ with respect to θ and i.

We will show that the function L_1^{dec} can be expressed in terms of the elliptical functions Π, E, and F and the function Θ, for whose calculation there exist efficient algorithms [2]:

$$L_1^{dec}(\Delta, r_1, r_2) = \frac{2\pi}{3}\Theta(r_n - \Delta)r_f^2 \quad (24)$$

$$+ \mathcal{Q}\left[\frac{1}{r_f^2 - (r_n - \Delta)^2}\right]$$

$$\times \left\{ \frac{2r_f^3(\Delta + r_n)}{3(\Delta - r_n)}\Pi\left(-\frac{4\Delta r_n}{(r_n - \Delta)^2}\right) \right.$$

$$\left. \frac{\Psi(\Delta, r_n, r_f)}{2}\middle| \frac{4\Delta r_n}{r_f^2 - (r_n - \Delta)^2}\right)$$

$$- \frac{2}{9}\left[3r_f(\Delta^2 - r_n^2)\right.$$

$$+ \frac{[r_f^2 - (r_n - \Delta)^2][(r_n + \Delta)^2 - r_f^2]}{r_f}\right]$$

$$\times F\left(\frac{\Psi(\Delta, r_n, r_f)}{2}\middle| \frac{4\Delta r_n}{r_f^2 - (r_n - \Delta)^2}\right)\right\}$$

$$+ \frac{2}{9r_f}\mathcal{Q}\left[r_f^2 - (r_n - \Delta)^2\right](7r_n^2 + \Delta^2 - 4r_f^2)$$

$$\times E\left(\frac{\Psi(\Delta, r_n, r_f)}{2}\middle| \frac{4\Delta r_n}{r_f^2 - (r_n - \Delta)^2}\right).$$

4. GENERAL RELATIONS FOR THE LEAST-SQUARES METHOD

Since our goal is to elucidate the relationship between various methods for estimating parameter errors, let us recall some relationships from mathematical statistics that are relevant to our calculations.

Let us consider a linear model specified by the sequential set of functions $g_0(\theta) \ldots g_P(\theta)$ and the linear functions expressed in terms of them[1] $f^{lin}(\theta, \alpha_1 \ldots \alpha_P)$, defined for real $\alpha_1 \ldots \alpha_P$ and θ from the set $\{\theta_1 \ldots \theta_M\}$:

$$f^{lin}(\theta, \alpha_1 \ldots \alpha_P) = g_0(\theta) + \sum_{p=1}^{P} g_p(\theta)\alpha_p. \quad (25)$$

Here, $\{\theta_1 \ldots \theta_M\}$ corresponds to the set of M points at which observations are conducted. Let us specify the vector $(\overline{\alpha}_1 \ldots \overline{\alpha}_P)^T$, corresponding to the set of true values of the physical quantities (here, M is the total number of observational points), and the vector $(w_1 \ldots w_M)^T$ representing the weighting coefficients, assuming that the matrix $A_{qp} = \sum_{m=1}^{M} g_q(\theta_m)g_p(\theta_m)w_m$ is non-degenerate. We also specify a vector representing the random observed quantities $\xi = (\xi_1 \ldots \xi_M)^T$, assumed to be statistically independent, i.e., $\text{cov}(\xi_i, \xi_j) = 0$ for $i \neq j$, and to be normally distributed with the mathematical expectation values $\text{M}(\xi_k) = f^{lin}(\theta_k, \overline{\alpha}_1 \ldots \overline{\alpha}_P)$. Moreover, we assume that the measurements of the random quantities $\xi_1 w_1 \ldots \xi_M w_M$ are equally accurate; i.e., $\sigma^2(\xi_1 w_1) = \sigma^2(\xi_2 w_2) = \ldots = \sigma^2(\xi_M w_M) = \varepsilon_0^2$, where $\sigma^2(\cdot)$ denotes the operation of calculating the

---

[1] The form f^{lin} [(25)] in which the first term is independent of the linear parameter is chosen here for convenience in practical use.

dispersion and ε_0^2 is called the unit-weight dispersion. Denoting $\sigma^2(\xi_m) = \sigma_m^2$, we find that $w_m = \varepsilon_0^2/\sigma_m^2$. We specify the *residual functional* for a given model by the expression

$$R^{lin}(\alpha_1 \ldots \alpha_P, \xi_1 \ldots \xi_M) \qquad (26)$$

$$= \sum_{m=1}^{M} \left[\xi_m - f^{lin}(\theta_m, \alpha_1 \ldots \alpha_P) \right]^2 w_m$$

$$= \sum_{m=1}^{M} \left(\xi_m - g_0(\theta_m) - \sum_{p=1}^{P} g_p(\theta_m)\alpha_p \right)^2 w_m.$$

The values of the parameters $\alpha_1^c(\xi) \ldots \alpha_P^c(\xi)$ yielding the minimum of the residual functional (26) for fixed $\xi_1 \ldots \xi_M$ (which we will call their central values), are obtained from the solution of the system of P linear equations

$$\frac{\partial R^{lin}(\alpha_1 \ldots \alpha_P)}{\partial \alpha_q} = 2B_q \qquad (27)$$

$$- 2\sum_{p=1}^{P} A_{qp}\alpha_p = 0, \quad q = 1 \ldots P,$$

where

$$A_{qp} = \sum_{m=1}^{M} g_q(\theta_m)g_p(\theta_m)w_m, \qquad (28)$$

$$B_q = \sum_{m=1}^{M} [\xi_m - g_0(\theta_m)]g_q(\theta_m)w_m.$$

These values are equal to

$$\alpha_p^c(\xi) = \sum_{q=1}^{P} A_{qp}^{\text{inv}} B_q \qquad (29)$$

$$= \sum_{m=1}^{M} [\xi_m - g_0(\theta_m)]w_m \sum_{q=1}^{P} A_{qp}^{\text{inv}} g_q(\theta_m),$$

$$p = 1 \ldots P,$$

where A_{qp}^{inv} are elements of the inverse matrix to A: $A_{qp}^{\text{inv}} \equiv (A^{-1})_{qp}$.

Thus, the central values of $\alpha_1^c(\xi) \ldots \alpha_P^c(\xi)$ are expressed as linear combinations of the $\xi_1 \ldots \xi_M$, and are consequently normally distributed. Their mathematical expectation values are $\overline{\alpha}_1 \ldots \overline{\alpha}_P$ [which is obtained if we take the mathematical expectations of both parts of (29) and substitute them into (25)]. Using the linearity of the covariance operation $\text{cov}(\cdot, \cdot)$,

the statistical independence of $\xi_1 \ldots \xi_M$, and the fact that $\text{cov}(\xi_i, \xi_i) = \sigma^2(\xi_i) = \varepsilon_0^2/w_i$, we find the covariance matrix of the parameters $\alpha_1^c \ldots \alpha_P^c$:

$$\text{cov}(\alpha_p^c(\xi), \alpha_q^c(\xi)) \qquad (30)$$

$$= \sum_{m=1}^{M} (\sigma^2(\xi_m)w_m^2 \sum_{i,j=1}^{P} A_{ip}^{\text{inv}} A_{jq}^{\text{inv}} g_i(\theta_m)g_j(\theta_m))$$

$$= \varepsilon_0^2 \sum_{i,j=1}^{P} (A_{ip}^{\text{inv}} A_{jq}^{\text{inv}} \sum_{m=1}^{M} w_m g_i(\theta_m)g_j(\theta_m) = \varepsilon_0^2 A_{pq}^{\text{inv}}).$$

The dispersions of the central values of $\alpha_1^c(\xi) \ldots \alpha_P^c(\xi)$ are the diagonal elements of the covariance matrix:

$$\sigma^2(\alpha_p^c(\xi)) = \varepsilon_0^2 A_{pp}^{\text{inv}}. \qquad (31)$$

In the model obtained via the linear substitution of parameters using the non-degenerate matrix C_p^i,

$$\alpha_p' = C_0 + \sum_{i=1}^{P} C_p^i \alpha_i,$$

the dispersions of the new parameters $\alpha_1' \ldots \alpha_P'$ are found from the formula

$$\sigma^2(\alpha_p'^c) = \sum_{i,j=1}^{P} C_p^i C_p^j \text{cov}(\alpha_i^c, \alpha_j^c). \qquad (32)$$

Knowing the dispersion of the central value of a parameter, we can construct the interval within which the true value of the parameter falls with a specified probability. For this, it is sufficient to note that it follows from the normal distribution for the central value of the parameter that

$$\mathbf{P}\left(\left|\alpha_p^c(\xi) - \bar{\alpha}_p\right| \leq \kappa(\gamma)\sigma(\alpha_p^c(\xi))\right) = \gamma, \qquad (33)$$

where $\mathbf{P}$ denotes the probability that a condition is satisfied; κ depends on a chosen probability (confidence level) γ, and is given by the root of the equation

$$\sqrt{\frac{2}{\pi}} \int_0^\kappa \exp\left(-\frac{t^2}{2}\right) dt = \gamma.$$

For example, the confidence level $\gamma = 0.6827$, 0.9545, and 0.9973 for $\kappa = 1$, 2, and 3, respectively (corresponding to 1, 2, or 3σ).

When ε_0 is not known (for example, as in the case of a real observed light curve), in place of (31),

we must find the errors using formulas obtained by replacing ε_0 in (31) and (30) with the quantity wq [4]:

$$v_0^2(\xi) = \frac{R^{lin}(\alpha_1^c(\xi) \ldots \alpha_P^c(\xi), \xi_1 \ldots \xi_M)}{M - P}, \quad (34)$$

called the rms estimated unit-weight dispersion. The resulting rms estimates for the dispersions (which are random quantities) are

$$\sigma_{est}^2(\alpha_p^c(\xi)) \equiv v_0^2(\xi) A_{pp}^{\mathrm{inv}}. \quad (35)$$

Here, v_0 is a random quantity, and $(\alpha_p^c(\xi) - \bar{\alpha}_p)/\sigma_{est}(\alpha_p^c(\xi))$ obeys a Student distribution with $M - P$ degrees of freedom. However, this is already close to a normal distribution for sufficiently large $M - P \gg 10$, and we can assume $\mathbf{P}(|\alpha_p^c(\xi) - \bar{\alpha}_p| \leq \kappa\sigma(\alpha_p^c(\xi))) \simeq \mathbf{P}(|\alpha_p^c(\xi) - \bar{\alpha}_p| \leq \kappa\sigma_{est}(\alpha_p^c(\xi)))$; i.e., the probability that the true value falls in an interval constructed by multiplying the rms estimate of the dispersion[2] by the corresponding coefficient $\kappa(\gamma)$ will be fairly close to γ.

Let us now consider a model specified by an arbitrary, in general non-linear, function $f(\theta, \beta_1 \ldots \beta_P)$, defined for $\theta \in \{\theta_1 \ldots \theta_M\}$ and for the vectors $(\beta_1 \ldots \beta_P)^T \in \mathcal{B}$, where $\mathcal{B}$ is some region of real, Euclidean space. We assume that $f(\theta, \beta_1 \ldots \beta_P)$ is differentiable with respect to $\beta_1 \ldots \beta_P$ over this entire region.

As for the linear model, we specify the $\overline{\beta}_1 \ldots \overline{\beta}_P$, $w_1 \ldots w_M$, the normally distributed random quantities $\xi_1 \ldots \xi_M$, the unit-weight dispersion ε_0, and the residual functional. The random quantities $\xi_1 \ldots \xi_M$ display a normal distribution, and

$$\mathrm{M}(\xi_k) = f(\theta_k, \overline{\beta}_1 \ldots \overline{\beta}_P), \quad (36)$$
$$\sigma^2(\xi_1 w_1) = \sigma^2(\xi_2 w_2) = \ldots = \sigma^2(\xi_M w_M) = \varepsilon_0^2,$$

where $\mathrm{M}(\xi_k)$ denotes the mathematical expectation of ξ_k and $\sigma^2(\cdot)$ the operation of finding the dispersion.

We also assume that the matrix

$$A_{qp}(\beta_1 \ldots \beta_P) \quad (37)$$
$$= \sum_{m=1}^{M} \frac{\partial f}{\partial \beta_q}(\theta_m, \beta_1 \ldots \beta_P)\frac{\partial f}{\partial \beta_p}(\theta_m, \beta_1 \ldots \beta_P)w_m$$

---

[2] Note that variation of the rms estimates of the dispersions (which are random quantities) from sample to sample, i.e., their scatter for various measurements of the same curve, can be appreciably larger than the difference between the Student and Gaussian distributions. This scatter is inversely proportional to the square root of the number of points in the light curve.

is non-degenerate for $(\beta_1 \ldots \beta_P)^T \in \mathcal{B}$. We will also denote the elements of the inverse matrix to (37) as $A_{qp}^{\mathrm{inv}}(\beta_1 \ldots \beta_P)$.

The residual functional is specified by the expression

$$R(\beta_1 \ldots \beta_P, \xi_1 \ldots \xi_M) \quad (38)$$
$$= \sum_{m=1}^{M} [\xi_m - f(\theta, \beta_1 \ldots \beta_P)]^2 w_m,$$

and we assume that, for fixed $\xi_1 \ldots \xi_M$, it is convex in the variables $\beta_1 \ldots \beta_P$ and reaches its minimum in the region of $\mathcal{B}$.

In the differential-correction method, the function f is replaced by a Taylor series expansion up to a linear term at the minimum of the residual functional. Dispersions found from a least-squares solution for the corresponding linear model are used as estimates of the dispersions of the optimal values $\beta_1 \ldots \beta_P$.

We denote $\beta_1^c(\xi) \ldots \beta_P^c(\xi)$ to be the parameter values (which we will call central values) providing the minimum of the residual function $R(\beta_1 \ldots \beta_P, \xi_1 \ldots \xi_M)$ for specified $\xi_1 \ldots \xi_M$, and set in (25), and then (28) and (29),

$$\alpha_p = \beta_p - \beta_p^c(\xi), \quad (39)$$
$$g_0(\theta) = f(\theta, \beta_1^c(\xi) \ldots \beta_P^c(\xi)),$$
$$g_p(\theta) = \frac{\partial f}{\partial \beta_p^c}(\theta, \beta_1^c(\xi) \ldots \beta_P^c(\xi)),$$

where $p = 1 \ldots P$. Then, the quantities

$$\mathrm{cov}_o(\beta_q^c(\xi), \beta_p^c(\xi)) \equiv \varepsilon_0^2 A_{qp}^{\mathrm{inv}}(\beta_1^c(\xi) \ldots \beta_P^c(\xi)) \quad (30')$$

and

$$\sigma_o^2(\beta_p^c(\xi)) \equiv \varepsilon_0^2 A_{pp}^{\mathrm{inv}}(\beta_1^c(\xi) \ldots \beta_P^c(\xi)), \quad (31')$$

obtained by substituting (39) into (30) and (31) for the covariances and dispersions of the central values in the linear model are taken as estimates of the covariances and dispersions $\beta_1^c(\xi) \ldots \beta_P^c(\xi)$. In the case of a real observed light curve, when the unit-weight dispersions are unknown, analogous to the linear case, we use in place of ε_0 the rms estimate of the unit-weight dispersion

$$v_0^2(\xi) = \frac{R(\beta_1^c(\xi) \ldots \beta_P^c(\xi), \xi_1 \ldots \xi_M)}{M - P} \quad (34')$$

and the corresponding approximate rms estimates of the parameter dispersions

$$\sigma_{est}^2(\beta_p^c(\xi)) \equiv v_0^2(\xi) A_{pp}^{\mathrm{inv}}(\beta_1^c(\xi) \ldots \beta_P^c(\xi)). \quad (35')$$

Using this approximation in the computations assumes that we can neglect variations in the derivatives of the function f in (37) calculated with the central parameter values when ξ is varied in the vicinity of its mathematical expectations.

It is obvious that, if the function $f(\theta, \beta_1 \ldots \beta_P)$ is linear in $\beta_1 \ldots \beta_P$, then $(31')-(35')$ coincide with $(31)-(35)$; i.e., in the case of a linear function, the differential-correction method is identical to the least-squares method. Note that the described methods make the assumption that the model used is fully correct, and statistics for a normal distribution are used to estimate the parameter errors.

In the model obtained via the reversible parameter substitution

$$\beta'_p = \phi_p(\beta_1 \ldots \beta_P), \qquad (40)$$

where ϕ are smooth functions and the matrices $\dfrac{\partial \phi_p(\beta_1 \ldots \beta_P)}{\partial \beta_i}$ are non-degenerate, we can obtain an expression for approximate estimates of the central-value dispersions for the new parameters $\beta'^c_1(\xi) \ldots \beta'^c_P(\xi)$ by replacing the right-hand side of (40) with a Taylor expansion up to the linear term at the point $(\beta^c_1(\xi) \ldots \beta^c_P(\xi))^T$, setting in formula (32) $C^i_p = \dfrac{\partial \phi_p(\beta^c_1(\xi) \ldots \beta^c_P(\xi))}{\partial \beta^c_i}$ and replacing there the covariances $\alpha^c_1(\xi) \ldots \alpha^c_P(\xi)$ with the corresponding approximate estimates of the covariances $(\beta^c_1(\xi) \ldots \beta^c_P(\xi))$:

$$\sigma^2_o(\beta'^c_p(\xi)) \qquad (41)$$

$$= \sum_{i,j=1}^{P} \frac{\partial \phi_p(\beta^c_1(\xi) \ldots \beta^c_P(\xi))}{\partial \beta^c_i} \frac{\partial \phi_p(\beta^c_1(\xi) \ldots \beta^c_P(\xi))}{\partial \beta^c_j}$$

$$\times \operatorname{cov}_o(\beta^c_i(\xi), \beta^c_j(\xi)).$$

We should make an important comment about the described least-squares method (and, in the case of a non-linear model, the differential-correction method). In these methods, the statistical hypothesis H of the adequacy of the model to describe the observational data is usually not verified, and instead it is immediately assumed that the model used is correct. In other words, when adopting the hypothesis H in this case, the possibility that the researcher could make an error of the second kind (the model is not correct, but is adopted in accordance with some statistical criterion) is excluded. Precisely because of this, and also due to the use of a simple statistical normal distribution when estimating the uncertainties in the unknown parameters [see (31), (35), (31'), (35')], it is possible to obtain relatively small parameter errors in the least-squares and differential-correction methods (and also in the Monte Carlo method).

In reality, it is not known a priori how appropriate the assumption of the correctness of the model is, since, as a rule, the number of observed points in the light curve is much larger than the number of unknown parameters, so that the search for the parameters of the problem is strongly overdetermined. Therefore, the search for the parameters and their uncertainties should be carried out jointly with a statistical verification of the hypothesis H of the adequacy of the model. Further, as was noted in [5, 8, 12], the acceptance or rejection of a hypothesis H according to some statistical criterion does not provide a final, logical basis for its ultimate acceptance or rejection (because of the statistical nature of the criterion). Four cases are possible.

1. The hypothesis H is correct and is accepted by the criterion.

2. The hypothesis H is incorrect and is rejected by the criterion.

3. The hypothesis H is correct, but is rejected by the criterion (an error of the first kind).

4. The hypothesis H is incorrect, but is accepted by the criterion (an error of the second kind).

There exists for a selected statistical criterion a non-negative probability to admit an error of the first kind, denoted $\alpha_0 \geq 0$. The quantity α_0 is called the significance level of the criterion. It is obvious that, in the case of an adequate model, the probability of case (1) is $\gamma = 1 - \alpha_0$.

For example, we can use the following criterion to verify a hypothesis H (for more detail, see [8, 12]). Let us choose some random quantity (statistic) $\Delta(\tilde{u})$ that depends on some experimental data $\tilde{u}$, with the distribution $\Delta(\tilde{u})$ known (for example, this is a χ^2_M distribution, where M is the number of observed points in a light curve). We specify a priori some significance level α_0 and calculate the number Δ_M (quantile) such that $P\{\Delta(\tilde{u}) > \Delta_0\} = \alpha_0$ for all parameter values of the model if the hypothesis H is correct. If we obtain that $\Delta(\tilde{u}) > \Delta_M$ for any values of parameters for a specific realization of the experimental data (a specific light curve) $\tilde{u}$, the hypothesis H is rejected. On the contrary, if $(\Delta(\tilde{u}) \leq \Delta_M)$, the hypothesis H is accepted. At the same time, we must not forget that the model is accepted, not because it is necessarily correct, but because there is no basis for it to be rejected.

In the best cases, the criterion that the minimum reduced chi-squared residual be close to unity is used when applying the least-squares, differential-correction, or Monte Carlo method to find parameters and their uncertainties: $\dfrac{\chi^2_{M-P}}{M-P} \simeq 1$, where P is the number of unknown parameters. However, here, no quantitative measure of the adequacy of the model to
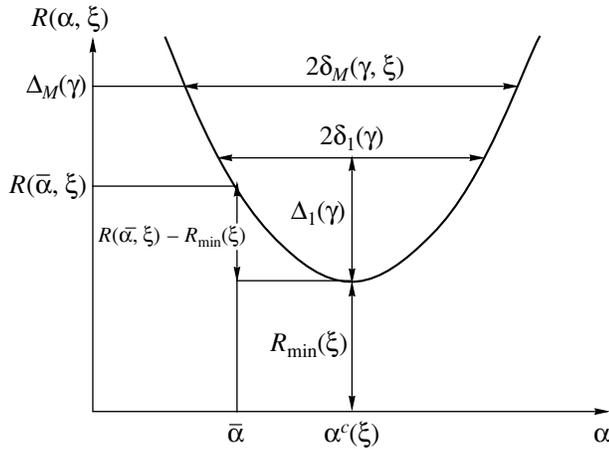
$R(\alpha, \xi)$

Fig. 2. A one-parameter problem. The dependence of the residual $R(\alpha, \xi)$ on the parameter α for a fixed vector comprising the random quantities ξ and the main statistical quantities. $R_{\min}(\xi)$ is the minimum residual in terms of α, $R(\alpha, \xi)$; $\Delta_P(\gamma)$ $(P = 1)$ the quantile of the χ^2_P $(P = 1)$ distribution; $\Delta_M(\gamma)$ the quantile of the χ^2_M distribution; $2\delta_1(\gamma)$ and $2\delta_M(\gamma, \xi)$ the widths of the confidence intervals obtained from methods using the statistics $R(\overline{\alpha}, \xi) - R_{\min}(\xi)$, $R(\overline{\alpha}, \xi)$ distributed as χ^2_P $(P = 1)$ and χ^2_M, respectively (see text for detail).

the observational data is considered, namely the significance level α_0 at which the model can be rejected. We have investigated this question here (see below).

5. ANALYTICAL RELATION BETWEEN THE ERROR INTERVALS OBTAINED USING STATISTICS DISTRIBUTED AS χ^2_M AND χ^2_P, AND THE DIFFERENTIAL-CORRECTION METHOD

In our previous study [1], we established via numerical simulations that, on average, the error interval obtained using the confidence-region method based on a statistic distributed as χ^2_M, where M is the number of measurement points (we will call this the χ^2_M method), exceeds the corresponding interval obtained using the differential-correction method by a factor of three to five. The error interval obtained using the χ^2_M method also exceeds the corresponding interval obtained using the confidence-region method based on a statistic distributed as χ^2_P, where P is the number of parameters being searched for (we will call this the χ^2_P method). The sizes of the confidence intervals obtained using the least-squares method with exact values of the dispersions and the χ^2_P method do not depend on the sample of random quantities, in contrast to the χ^2_M confidence intervals, which are themselves random quantities. We present below a

rigorous analytical relation between the error intervals for the above methods for a linear model. Namely, we derive the function $a_p(t)$ for the distribution of the ratio of the confidence interval for the χ^2_M method to its value for the χ^2_P method. In the one-dimensional problem $(P = 1)$, the χ^2_P confidence interval coincides with the corresponding interval for the least-squares method, while, for arbitrary P, these intervals are related by a well defined dependence (which will be presented below). We also study the behavior of $a_p(t)$ as a function of the number of observational points M and the selected confidence level γ.

Let us first consider the one-dimensional case. We consider the residual in a one-parameter linear model that depends quadratically on the one parameter α. We also assume for convenience that $\varepsilon_0 = 1$ (the statements below will remain valid for any ε_0). This residual can be written in the form (as follows from (26), where the terms that are quadratic in α do not contain the random quantities ξ)

$$R^{lin}(\alpha, \xi) = C[\alpha - \alpha^c(\xi)]^2 + R^{lin}_{\min}(\xi), \qquad (42)$$

where $R^{lin}_{\min}$ is its minimum value in terms of α, reached for $\alpha^c(\xi)$, and the coefficient C is expressed in a well defined way in terms of the specified parameters of the model (w and g), and does not depend on the sample of random quantities ξ.

The quantile $\Delta_n(\gamma)$ for $\gamma \in (0, 1)$ is given by

$$\chi^2_n(\Delta_n(\gamma)) = \gamma;$$

i.e., as an inverse function to χ^2_n.

Let us briefly recall the residual distribution. We showed in [1] that the residuals are distributed as

$$R^{lin}(\overline{\alpha}, \xi) - R^{lin}_{\min}(\xi) \sim \chi^2_1, \qquad (43)$$

$$R^{lin}(\overline{\alpha}, \xi) \sim \chi^2_M, \qquad (44)$$

$$R^{lin}_{\min}(\xi) \sim \chi^2_{M-1}, \qquad (45)$$

where $\overline{\alpha}$ is the true value of the parameter α, and the symbol $\sim$ denotes "distributed as." Figure 2 depicts the dependence of the residual $R(\alpha, \xi)$ on the parameter α for a fixed sample of random quantities ξ and its dependence on the main statistical parameters. The parabola representing this dependence in a linear model can shift in the upper half-plane for different ξ, while maintaining its shape, since the coefficient of the quadratic term in α does not depend on ξ.

The error interval obtained using the χ^2_1 method for a specified confidence level γ is defined as the distance between the roots of the quadratic equation in α:

$$C[\alpha - \alpha^c(\xi)]^2 + R^{lin}_{\min}(\xi) \qquad (46)$$
$$= R^{lin}_{\min}(\xi) + \Delta_1(\gamma).$$

We denote half this interval as $\delta_1(\gamma)$. It is obvious that $\delta_1(\gamma)$ obeys the relation

$$C\delta_1^2(\gamma) = \Delta_1(\gamma) \qquad (47)$$

and is independent of the sample of random quantities ξ. We thus have

$$C = \frac{\Delta_1(\gamma)}{\delta_1^2(\gamma)}. \qquad (48)$$

The error interval obtained using the χ_M^2 method for a specified confidence level γ is the distance between the roots of the quadratic equation in α:

$$C[\alpha - \alpha^c(\xi)]^2 + R_{\min}^{lin}(\xi) = \Delta_M(\gamma). \qquad (49)$$

We set $\delta_M(\gamma,\xi)$ equal to half this interval when this last equation has positive roots (i.e., the corresponding confidence set is not empty) and equal to zero otherwise. For positive values of $\delta_M(\gamma,\xi)$,

$$C\delta_M^2(\gamma,\xi) + R_{\min}^{lin}(\xi) = \Delta_M(\gamma). \qquad (50)$$

The quantity $\delta_M(\gamma,\xi)$ depends on the random quantities ξ and is itself a random quantity. Using (45) we obtain

$$\Delta_M(\gamma) - C\delta_M^2(\gamma,\xi) \sim \chi_{M-1}^2. \qquad (51)$$

This last expression indicates that

$$\mathbf{P}((\Delta_M(\gamma) - C\delta_M^2(\gamma,\xi)) < t) = \chi_{M-1}^2(t). \qquad (52)$$

We obtain from this last expression via straightforward linear and shift transformations in t and using (48)

$$\mathbf{P}\left(\frac{\delta_M^2(\gamma,\xi)}{\delta_1^2(\gamma)} > t\right) \qquad (53)$$

$$= \chi_{M-1}^2\left(\Delta_M(\gamma) - \Delta_1(\gamma)t\right).$$

It follows from (45) that the probability that $\delta_M(\gamma,\xi) > 0$ (i.e., that the confidence interval is nonzero for a specified confidence level γ), is $\chi_{M-1}^2(\Delta_M(\gamma))$.

The corresponding conditional probability is

$$\mathbf{P}\left(\delta_M(\gamma,\xi) > 0 \mid \frac{\delta_M(\gamma,\xi)}{\delta_1(\gamma)} > t\right) \qquad (54)$$

$$= \frac{\chi_{M-1}^2\left(\Delta_M(\gamma) - \Delta_1(\gamma)t^2\right)}{\chi_{M-1}^2(\Delta_M(\gamma))}$$

or

$$\mathbf{P}\left(\delta_M(\gamma,\xi) > 0 \mid \frac{\delta_M(\gamma,\xi)}{\delta_1(\gamma)} \leq t\right) \qquad (55)$$

$$= 1 - \frac{\chi_{M-1}^2\left(\Delta_M(\gamma) - \Delta_1(\gamma)t^2\right)}{\chi_{M-1}^2(\Delta_M(\gamma))}.$$

The right-hand side of this last expression is a function of the distribution of positive values of $\frac{\delta_M(\gamma,\xi)}{\delta_1(\gamma)}$, which is the ratio of the confidence interval for the χ_M^2 method to the corresponding interval obtained when the main statistic is distributed as χ_1^2, for a specified confidence level γ. The corresponding density of the distribution $a_1(t)$, obtained by differentiating the distribution function (with accuracy to within a normalization coefficient A_1), is equal to

$$a_1(t) = A_1 \begin{cases} e^{\frac{1}{2}(\Delta_1(\gamma)t^2 - \Delta_M(\gamma))}t\left(\Delta_M(\gamma) - \Delta_1(\gamma)t^2\right)^{\frac{M-3}{2}}, \\ \qquad 0 < t < \sqrt{\dfrac{\Delta_M(\gamma)}{\Delta_1(\gamma)}}, \\ 0, \qquad t \leq 0 \quad \text{or} \quad t \geq \sqrt{\dfrac{\Delta_M(\gamma)}{\Delta_1(\gamma)}}. \end{cases}$$

This function reaches its maximum when

$$t_{\max 1} = \sqrt{\frac{2 - M + \Delta_M(\gamma) + \sqrt{(2-M)^2 + \Delta_M(\gamma)(-2M + \Delta_M(\gamma) + 8)}}{2\Delta_1(\gamma)}}. \qquad (56)$$

Since the problem is one-dimensional, the confidence interval $\delta_1(\gamma)$ coincides with the interval obtained using the least-squares method for the corresponding confidence level.

The above can be generalized to the case of a

model with P parameters by obtaining the distribution of the ratio of the projections of the confidence regions determined using the χ_M^2 method and the χ_P^2 method. The distribution of the corresponding residuals is

$$R(\overline{\alpha_1}\ldots\overline{\alpha_P},\xi) - R_{\min}(\xi) \sim \chi_P^2, \qquad (57)$$

$$R(\overline{\alpha_1}\ldots\overline{\alpha_P},\xi) \sim \chi_M^2, \qquad (58)$$

$$R_{\min}(\xi) \sim \chi_{M-P}^2. \qquad (59)$$

Recall that, in the χ_P^2 and χ_M^2 methods, by confidence interval, we mean the projection of the confidence region obtained from the main corresponding statistic onto the axis for the corresponding parameter. This projection is found as the distance between the roots of the equation obtained by equating the residual minimized over all parameters except for one with the quantile for the selected confidence level.

Replacing throughout the previous derivations except in (43) and (44) $R(\alpha,\xi)$ with the P-parameter residual minimized over all parameters α except for one; and $\chi_1^2, \chi_{M-1}^2, \Delta_1(\gamma)$, and $\delta_1(\gamma)$ with $\chi_P^2, \chi_{M-P}^2, \Delta_P(\gamma)$, and $\delta_P(\gamma)$, we find that the distribution function is equal to

$$\mathbf{P}\left(\delta_M(\gamma) > 0 \Big| \frac{\delta_M(\gamma,\,\xi)}{\delta_P(\gamma)} \leq t\right) \qquad (60)$$

$$= 1 - \frac{\chi_{M-P}^2\left(\Delta_M(\gamma) - \Delta_P(\gamma)t^2\right)}{\chi_{M-P}^2(\Delta_M(\gamma))},$$

while the density of the distribution of positive values $\dfrac{\delta_M(\gamma,\,\xi)}{\delta_P(\gamma)}$ (with accuracy to within a normalization coefficient A_p) is

$$a_P(t) = A_p \begin{cases} e^{\frac{1}{2}\left(\Delta_P(\gamma)t^2 - \Delta_M(\gamma)\right)} t \left(\Delta_M(\gamma) - \Delta_P(\gamma)t^2\right)^{\frac{M-P-2}{2}}, & 0 < t < \sqrt{\dfrac{\Delta_M(\gamma)}{\Delta_P(\gamma)}}, \\ 0, & t \leq 0 \quad\text{or}\quad t \geq \sqrt{\dfrac{\Delta_M(\gamma)}{\Delta_P(\gamma)}}. \end{cases}$$

The maximum of this function is reached when

$$t_{\max P} = \sqrt{\frac{1 + P - M + \Delta_M(\gamma) + \sqrt{(1 + P - M)^2 + \Delta_M(\gamma)(2P - 2M - 2 + \Delta_M(\gamma) + 8)}}{2\Delta_P(\gamma)}}. \qquad (61)$$

Figure 3 shows the behavior of the density function $a_1(t)$ as a function of the confidence level γ for the χ_M^2 method. Recall that, in the one-parameter problem, the projection of the confidence region obtained using a statistic that is distributed as χ_P^2 coincides with the confidence interval obtained using the least-squares method. Figure 3 shows that the most probable value ratio of the confidence intervals obtained using the χ_M^2 method and the least-squares method depends on the chosen confidence level. For example, with $\gamma = 0.68$ and $\gamma = 0.99$, the most probable confidence interval obtained using the χ_M^2 method exceeds the confidence interval for the least-squares method by

about a factor of four and a factor of $\sim$2.1, respectively. These calculations were carried out for $M = 101$.

Further, we investigated the behavior of the maximum of $a_P(t)$ as a function of γ and the dimension of the χ_M^2 distribution. Figure 4 presents a set of these dependences for $M = 50, 100, 500,$ and 1000. Note that the dependence has a universal character for a linear model; i.e., it is independent of the coefficients of the linear function considered. Figure 4 shows that the excess of the χ_M^2 confidence intervals over the least-squares confidence intervals grows (comparatively weakly) with increasing dimension M of the χ_M^2 distribution (i.e., with the number of observational points M). It also follows from Fig. 4 that the ratio of
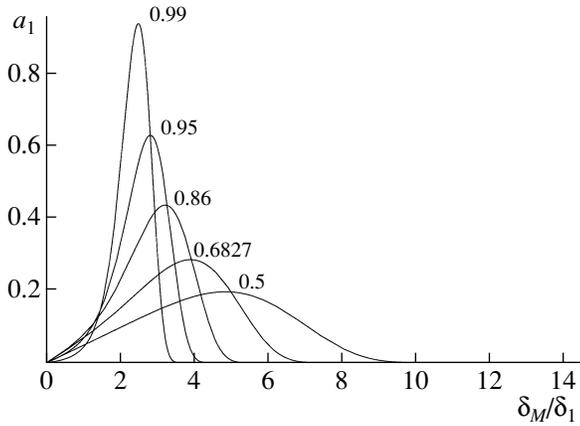
Fig. 3. Differential density of the distribution function for the ratio of the confidence interval $\Delta(\chi_M^2)$ (obtained using the χ_M^2 method with an adequate model) to the confidence interval obtained using the differential-correction method. The case of a one-dimensional function with $M = 101$ and $\gamma = 0.5, 0.6827, 0.86, 0.95, 0.99$ is shown.
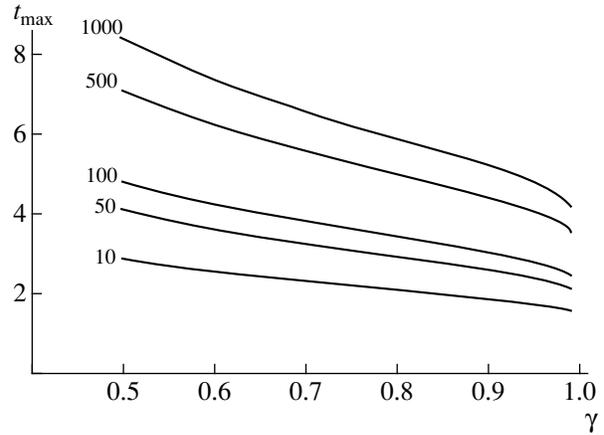

Fig. 4. Case of a one-parameter problem. Dependence of the maximum of the differential distribution function of the ratio of the confidence interval obtained using a main statistic distributed as χ_M^2 to the confidence interval obtained using the differential-correction method on the confidence level γ for numbers of observational points $M = 10, 50, 100, 500,$ and 1000. These data were obtained for a one-dimensional linear model function (the form of this function is not important; the dependence shown is universal).

the most probable χ_M^2 confidence interval to the least-squares confidence interval decreases with increasing γ. This dependence is shown in Fig. 5.

This raises the question of why the confidence intervals in the differential-correction method (which are identical to the intervals for the least squares method for linear models) and Monte Carlo method are appreciably smaller than those obtained for the χ_M^2 method. In the one-dimensional case ($P = 1$), the problem of projecting the confidence region onto the parameter axis does not arise, and the difference between the confidence intervals is clearly associated with two circumstances.

1. When using the χ_M^2 statistic (residual), there is no initial assumption that the model used is correct. In practice, this means that the surface of the residual functional (Fig. 2) for the χ_M^2 statistic for various realizations of the observational data can be shifted both along the parameter axis (in the "horizontal" direction) and along the axis for the residual functional (in the "vertical" direction). For some realizations of the observational data, the confidence region can even degenerate into an empty set (the number of such cases should be close to the adopted significance level, $\alpha = 1 - \gamma$). The surface of the residual functional (Fig. 2) for the χ_P^2 statistic for various realizations of the observational data can only be shifted along the parameter axis (in the "horizontal" direction). The corresponding confidence region never degenerates into an empty set.

2. For equal conditions, the distribution of χ_M^2 confidence intervals, obtained assuming that the model is adequate and using the χ_M^2 method is

broader (Fig. 3) than the normal distribution of the central parameter values used in the differential-correction method (in which the model is assumed to be adequate by definition). The characteristic parameters of these distributions are, in the former case, the most probable χ_M^2 error interval and, in the latter case, the root-mean-square (rms) estimate of a standard deviation σ multiplied by a coefficient corresponding to the selected confidence level (which coincides with the χ_P^2 interval in the one-dimensional case). The parameter t_{max} is the ratio of the first to the second of these parameters.

In the multi-parameter case ($P > 1$), the effect of having to project a multi-dimensional (in the spatial parameters) confidence region onto the parameter axes is added to the reasons indicated above. This must be done in order to make the transition from an "exact" confidence region to the values of the confidence intervals. Since the projection of the confidence region is replaced by its volumization as a parallelopiped whose volume is larger, the corresponding probability for overlapping the exact solution additionally grows.

Thus, when choosing the confidence level γ for the confidence intervals, we must understand that the probability P that they encompass the exact solution of the corresponding confidence region is $P < \gamma$ in the differential-correction or Monte Carlo methods, while the probability of encompassing the exact solution by the corresponding parallelopiped volumizing the
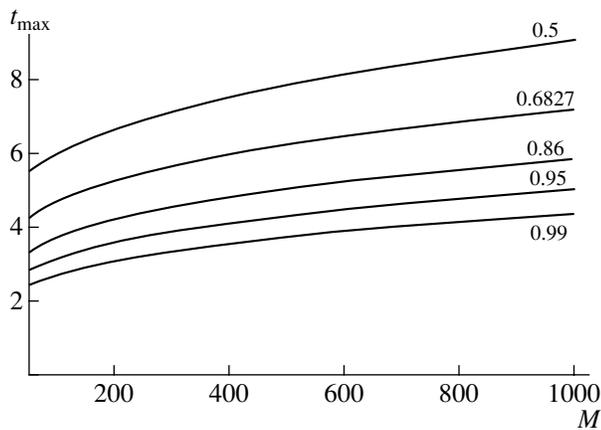
Fig. 5. The one-parameter problem. Dependence of the maximum of the differential distribution function for the ratio of the χ^2_M and differential-correction confidence intervals on the number of points M for specified confidence levels $\gamma = 0.5, 0.6827, 0.86, 0.95,$ and 0.99.
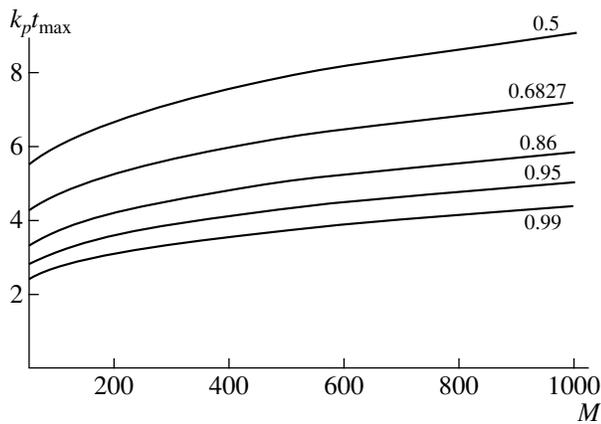

Fig. 6. Multi-parameter case. The quantity $k_P t_{max}$ for the multi-parameter problem as a function of the number of observational points M and of γ for $P = 10$. The confidence level is indicated next to the corresponding curve. It is striking that the probability that the true value of the parameter falls into the projection of the χ^2_M confidence region is higher than the specified confidence level when $\gamma = 0.6827$: it is equal to ~ 0.80. At the same time, the probability that the true solution falls into the error interval for the differential-correction method is equal to the specified confidence level, $\gamma = 0.6827$. See text for details.

"exact" confidence region is $P > \gamma$ in the χ^2_M method. In the χ^2_M method, the probability of encompassing the exact solution of the corresponding confidence region is $P = \gamma$.

In the case of a non-linear model, there is another reason why the error intervals obtained using the differential-correction method can be reduced: the application of a linearization procedure. However, in the model used by us, this factor does not appre-

ciably affect the results, as is indicated by the fact that the estimates of the dispersions of the parameters obtained using the differential-correction and Monte Carlo methods do not substantially differ. The appreciable difference between the intervals obtained using the different methods is due precisely to the use of different statistics, as well as differences in the assumptions concerning the correctness of the model.

5.1. New Method for Estimating the Parameter Errors

Thus, when the dependence of a function describing a model on the unknown parameters is linear, then (56) and (61) enable, based on the Monte Carlo or differential-correction method via multiplication by the coefficient t_{max} (or $k_P t_{max}$, see below), the derivation of the size of the error interval for an unknown parameter in the framework of the χ^2_M confidence regions; this distribution is the most probable among the cases in which the model is not rejected. Having obtained the size of the Monte Carlo or differential-correction error interval and convinced ourselves that the model is adequate in the framework of the χ^2_M method (for which it is sufficient to calculate the minimum corresponding residual), we can draw preliminary conclusions of a probabilistic nature about the size of the χ^2_M confidence interval, without turning directly to the more labor-intensive determination of the sizes of the projection of the confidence region onto the parameter axis.

We emphasize again that t_{max} is the most probable ratio of the projection of the χ^2_M and χ^2_P confidence regions. As was noted above, in the one-dimensional case, the projection of the χ^2_P confidence region is equal to the confidence integral yielded by the differential-correction method [1]. Therefore, the value of t_{max} for the one-dimensional case is the most probable ratio of the projection of the χ^2_M confidence region and the differential-correction confidence interval. Values of t_{max} for the one-dimensional problem are given in Table 1.

Table 1 shows that t_{max} differs appreciably from unity, reflecting the reduction in the parameter errors obtained using the differential-correction method (a normally distributed statistic).

Tables 2–5 present values of t_{max} for the multi-parameter case with $P = 3, 6, 10,$ and 50. In the multi-dimensional problem, the projection of the χ^2_P confidence region is not equal to the confidence intervals obtained using the differential-correction or Monte Carlo methods [1]. Therefore, for convenience in comparing the ratio of the projection of the χ^2_M confidence region to the differential-correction confidence interval, Table 6 presents the corresponding

Table 1. Parameter $t_{\max}$ for the one-parameter problem as a function of M and γ for $P = 1$. (The confidence interval obtained using the χ_P^2 method coincides with that for the differential-correction method)

M	γ										
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
10	8.49	5.20	3.98	3.32	2.89	2.58	2.33	2.11	1.89	1.76	1.58
20	10.4	6.21	4.68	3.87	3.35	2.98	2.68	2.42	2.16	2.01	1.79
30	11.7	6.89	5.16	4.25	3.67	3.25	2.92	2.64	2.35	2.18	1.94
40	12.7	7.41	5.53	4.54	3.92	3.47	3.11	2.81	2.50	2.31	2.05
50	13.5	7.84	5.84	4.79	4.12	3.64	3.27	2.95	2.62	2.42	2.15
60	14.1	8.22	6.10	5.00	4.30	3.80	3.41	3.07	2.73	2.52	2.23
70	14.7	8.54	6.34	5.18	4.45	3.93	3.53	3.17	2.82	2.60	2.31
80	15.3	8.84	6.55	5.35	4.59	4.06	3.63	3.27	2.90	2.68	2.37
90	15.8	9.11	6.74	5.50	4.72	4.17	3.73	3.36	2.98	2.75	2.43
100	16.2	9.35	6.92	5.64	4.84	4.27	3.82	3.44	3.05	2.81	2.49
200	19.5	11.1	8.20	6.67	5.70	5.02	4.49	4.03	3.57	3.29	2.90
300	21.7	12.3	9.06	7.36	6.29	5.53	4.94	4.44	3.93	3.61	3.18
400	23.3	13.3	9.73	7.89	6.74	5.93	5.29	4.75	4.20	3.87	3.40
500	24.7	14.0	10.3	8.33	7.11	6.25	5.58	5.01	4.43	4.08	3.58
600	25.9	14.7	10.8	8.71	7.44	6.54	5.83	5.23	4.63	4.26	3.74
700	26.9	15.3	11.2	9.05	7.72	6.78	6.06	5.43	4.80	4.41	3.88
800	27.9	15.8	11.6	9.35	7.98	7.01	6.25	5.61	4.96	4.56	4.00
900	28.7	16.3	11.9	9.63	8.21	7.21	6.43	5.77	5.10	4.69	4.12
1000	29.5	16.7	12.2	9.88	8.42	7.40	6.60	5.92	5.23	4.81	4.22
1500	32.7	18.5	13.5	10.9	9.31	8.17	7.29	6.53	5.77	5.30	4.65
2000	35.2	19.9	14.5	11.7	9.99	8.77	7.82	7.00	6.19	5.68	4.99
2500	37.2	21.0	15.3	12.4	10.6	9.26	8.26	7.39	6.53	6.00	5.27
3000	39.0	22.0	16.1	13.0	11.0	9.68	8.63	7.73	6.83	6.27	5.50
3500	40.5	22.9	16.7	13.5	11.5	10.1	8.97	8.03	7.09	6.52	5.71
4000	41.9	23.6	17.2	13.9	11.9	10.4	9.27	8.30	7.33	6.73	5.90
4500	43.2	24.4	17.8	14.3	12.2	10.7	9.54	8.54	7.54	6.93	6.08
5000	44.4	25.0	18.2	14.7	12.5	11.0	9.79	8.77	7.74	7.11	6.23
5500	45.4	25.6	18.7	15.1	12.8	11.2	10.0	8.98	7.93	7.28	6.38
6000	46.4	26.2	19.1	15.4	13.1	11.5	10.2	9.17	8.10	7.44	6.52
6500	47.4	26.7	19.5	15.7	13.4	11.7	10.4	9.35	8.26	7.58	6.65
7000	48.3	27.2	19.8	16.0	13.6	11.9	10.6	9.53	8.41	7.72	6.77
7500	49.1	27.7	20.2	16.3	13.9	12.1	10.8	9.69	8.56	7.86	6.89
8000	49.9	28.1	20.5	16.5	14.1	12.3	11.0	9.84	8.69	7.98	7.00
8500	50.7	28.6	20.8	16.8	14.3	12.5	11.2	9.99	8.82	8.10	7.10
9000	51.4	29.0	21.1	17.0	14.5	12.7	11.3	10.1	8.95	8.22	7.20
9500	52.1	29.4	21.4	17.3	14.7	12.9	11.5	10.3	9.07	8.33	7.30
10 000	52.8	29.7	21.7	17.5	14.9	13.0	11.6	10.4	9.19	8.43	7.39

Table 2. Parameter t_{max} for the multi-parameter problem as a function of M and γ for $P = 3$. (The quantity $k_P t_{max}$ must be used to translate to the confidence interval obtained using the differential-correction or Monte Carlo methods; the values of k_P were taken from Table 6)

M	γ										
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
10	1.71	1.60	1.55	1.51	1.48	1.45	1.42	1.39	1.35	1.32	1.28
20	1.93	1.78	1.71	1.67	1.64	1.61	1.59	1.56	1.52	1.48	1.43
30	2.10	1.92	1.84	1.79	1.76	1.73	1.71	1.68	1.63	1.60	1.54
40	2.24	2.03	1.94	1.89	1.86	1.83	1.80	1.77	1.72	1.69	1.62
50	2.36	2.13	2.03	1.98	1.94	1.91	1.88	1.85	1.80	1.76	1.69
60	2.46	2.21	2.11	2.05	2.01	1.98	1.95	1.92	1.87	1.83	1.75
70	2.56	2.29	2.18	2.11	2.07	2.04	2.01	1.98	1.93	1.89	1.81
80	2.64	2.36	2.24	2.17	2.13	2.10	2.07	2.03	1.98	1.94	1.86
90	2.71	2.42	2.30	2.23	2.18	2.15	2.12	2.08	2.03	1.99	1.90
100	2.78	2.48	2.35	2.28	2.23	2.20	2.16	2.13	2.07	2.03	1.95
200	3.30	2.91	2.74	2.65	2.59	2.55	2.51	2.47	2.41	2.36	2.26
300	3.64	3.20	3.01	2.90	2.84	2.79	2.75	2.70	2.64	2.58	2.47
400	3.91	3.43	3.22	3.10	3.03	2.98	2.93	2.88	2.81	2.75	2.63
500	4.13	3.62	3.39	3.27	3.19	3.13	3.09	3.03	2.96	2.90	2.77
600	4.32	3.78	3.54	3.41	3.33	3.27	3.22	3.16	3.09	3.02	2.89
700	4.49	3.93	3.68	3.54	3.45	3.39	3.34	3.28	3.20	3.13	2.99
800	4.64	4.05	3.80	3.65	3.56	3.50	3.44	3.38	3.30	3.23	3.09
900	4.78	4.17	3.90	3.75	3.66	3.59	3.54	3.48	3.39	3.32	3.18
1000	4.91	4.28	4.00	3.85	3.75	3.68	3.63	3.57	3.48	3.40	3.25
1500	5.43	4.73	4.42	4.24	4.13	4.06	3.99	3.93	3.83	3.75	3.58
2000	5.83	5.07	4.74	4.55	4.43	4.35	4.28	4.20	4.10	4.01	3.84
2500	6.17	5.36	5.00	4.80	4.67	4.58	4.51	4.44	4.33	4.23	4.05
3000	6.45	5.60	5.23	5.02	4.88	4.79	4.71	4.64	4.52	4.42	4.23
3500	6.71	5.82	5.43	5.21	5.07	4.97	4.89	4.81	4.70	4.59	4.39
4000	6.93	6.02	5.61	5.38	5.24	5.14	5.05	4.97	4.85	4.74	4.53
4500	7.14	6.19	5.78	5.54	5.39	5.29	5.20	5.11	4.99	4.88	4.66
5000	7.33	6.36	5.93	5.68	5.53	5.42	5.34	5.25	5.12	5.01	4.78
5500	7.51	6.51	6.07	5.82	5.66	5.55	5.46	5.37	5.24	5.13	4.90
6000	7.67	6.65	6.20	5.94	5.78	5.67	5.58	5.48	5.35	5.24	5.00
6500	7.83	6.78	6.32	6.06	5.90	5.78	5.69	5.59	5.46	5.34	5.10
7000	7.97	6.91	6.44	6.17	6.01	5.89	5.79	5.69	5.56	5.44	5.19
7500	8.11	7.03	6.55	6.28	6.11	5.99	5.89	5.79	5.65	5.53	5.28
8000	8.24	7.14	6.65	6.38	6.21	6.08	5.99	5.88	5.74	5.62	5.37
8500	8.37	7.25	6.75	6.48	6.30	6.18	6.07	5.97	5.83	5.70	5.45
9000	8.49	7.35	6.85	6.57	6.39	6.26	6.16	6.06	5.91	5.78	5.52
9500	8.60	7.45	6.94	6.66	6.47	6.35	6.24	6.14	5.99	5.86	5.60
10 000	8.72	7.55	7.03	6.74	6.56	6.43	6.32	6.21	6.06	5.93	5.67

Table 3. Parameter t_{max} for the multi-parameter problem as a function of M and γ for $P = 6$. (The quantity $k_P t_{max}$ must be used to translate to the confidence interval obtained using the differential-correction or Monte Carlo methods; the values of k_P were taken from Table 6)

M	γ										
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
10	1.23	1.22	1.21	1.20	1.19	1.18	1.17	1.16	1.15	1.14	1.13
20	1.22	1.24	1.25	1.26	1.27	1.27	1.27	1.27	1.27	1.26	1.24
30	1.26	1.28	1.30	1.32	1.33	1.34	1.35	1.35	1.35	1.34	1.33
40	1.31	1.32	1.34	1.36	1.38	1.39	1.40	1.41	1.41	1.41	1.39
50	1.35	1.36	1.39	1.41	1.43	1.44	1.45	1.46	1.47	1.46	1.45
60	1.40	1.40	1.42	1.45	1.47	1.48	1.50	1.51	1.52	1.51	1.50
70	1.43	1.44	1.46	1.48	1.50	1.52	1.54	1.55	1.56	1.56	1.54
80	1.47	1.47	1.49	1.51	1.54	1.56	1.57	1.59	1.60	1.60	1.58
90	1.50	1.50	1.52	1.54	1.57	1.59	1.61	1.62	1.63	1.63	1.62
100	1.54	1.53	1.55	1.57	1.60	1.62	1.64	1.65	1.67	1.66	1.65
200	1.78	1.75	1.76	1.79	1.82	1.84	1.87	1.89	1.91	1.91	1.90
300	1.94	1.91	1.92	1.94	1.97	2.00	2.03	2.06	2.08	2.08	2.07
400	2.08	2.03	2.04	2.06	2.09	2.12	2.16	2.19	2.21	2.22	2.20
500	2.19	2.14	2.14	2.16	2.19	2.23	2.26	2.30	2.32	2.33	2.31
600	2.28	2.23	2.23	2.25	2.28	2.32	2.35	2.39	2.42	2.42	2.41
700	2.37	2.30	2.31	2.33	2.36	2.40	2.44	2.47	2.50	2.51	2.49
800	2.44	2.38	2.38	2.40	2.43	2.47	2.51	2.55	2.58	2.59	2.57
900	2.51	2.44	2.44	2.46	2.50	2.54	2.58	2.61	2.65	2.66	2.64
1000	2.58	2.50	2.50	2.52	2.56	2.60	2.64	2.68	2.71	2.72	2.70
1500	2.84	2.75	2.74	2.77	2.80	2.85	2.89	2.94	2.98	2.99	2.97
2000	3.04	2.95	2.94	2.96	3.00	3.04	3.09	3.14	3.18	3.19	3.18
2500	3.21	3.11	3.09	3.12	3.16	3.20	3.26	3.31	3.35	3.37	3.35
3000	3.36	3.25	3.23	3.25	3.29	3.34	3.40	3.45	3.50	3.52	3.50
3500	3.49	3.37	3.35	3.37	3.42	3.47	3.52	3.58	3.63	3.65	3.63
4000	3.60	3.48	3.46	3.48	3.53	3.58	3.64	3.70	3.75	3.77	3.75
4500	3.71	3.58	3.56	3.58	3.63	3.68	3.74	3.80	3.86	3.87	3.85
5000	3.81	3.67	3.65	3.67	3.72	3.77	3.84	3.90	3.95	3.97	3.95
5500	3.90	3.76	3.74	3.76	3.80	3.86	3.92	3.99	4.05	4.06	4.04
6000	3.98	3.84	3.82	3.84	3.88	3.94	4.01	4.07	4.13	4.15	4.13
6500	4.06	3.91	3.89	3.91	3.96	4.02	4.08	4.15	4.21	4.23	4.21
7000	4.13	3.99	3.96	3.98	4.03	4.09	4.16	4.23	4.29	4.31	4.29
7500	4.20	4.05	4.03	4.05	4.10	4.16	4.23	4.30	4.36	4.38	4.36
8000	4.27	4.12	4.09	4.11	4.16	4.22	4.29	4.36	4.43	4.45	4.43
8500	4.34	4.18	4.15	4.17	4.22	4.29	4.36	4.43	4.49	4.51	4.49
9000	4.40	4.24	4.21	4.23	4.28	4.35	4.42	4.49	4.56	4.58	4.55
9500	4.46	4.29	4.27	4.29	4.34	4.40	4.48	4.55	4.62	4.64	4.61
10 000	4.51	4.35	4.32	4.34	4.39	4.46	4.53	4.61	4.67	4.70	4.67

ABUBEKEROV et al.

Table 4. Parameter t_{max} for the multi-parameter problem as a function of M and γ for $P = 10$. (The quantity $k_P t_{max}$ must be used to translate to the confidence interval obtained using the differential-correction or Monte Carlo methods; the values of k_P were taken from Table 6)

M	γ										
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
10	1.16	1.13	1.12	1.10	1.09	1.08	1.08	1.07	1.06	1.05	1.04
20	1.08	1.10	1.12	1.13	1.13	1.14	1.14	1.14	1.14	1.14	1.13
30	1.05	1.10	1.13	1.15	1.16	1.18	1.19	1.19	1.20	1.20	1.20
40	1.06	1.11	1.15	1.17	1.19	1.21	1.22	1.24	1.25	1.25	1.25
50	1.07	1.12	1.16	1.19	1.22	1.24	1.26	1.27	1.29	1.30	1.30
60	1.08	1.14	1.18	1.21	1.24	1.27	1.29	1.31	1.32	1.33	1.34
70	1.09	1.15	1.20	1.23	1.27	1.29	1.32	1.34	1.36	1.37	1.37
80	1.11	1.17	1.21	1.25	1.29	1.32	1.34	1.36	1.39	1.40	1.40
90	1.13	1.18	1.23	1.27	1.31	1.34	1.36	1.39	1.41	1.43	1.43
100	1.14	1.20	1.25	1.29	1.32	1.36	1.39	1.41	1.44	1.45	1.46
200	1.28	1.33	1.38	1.43	1.47	1.51	1.55	1.59	1.63	1.65	1.66
300	1.38	1.42	1.48	1.53	1.58	1.62	1.67	1.71	1.76	1.78	1.81
400	1.46	1.50	1.56	1.61	1.66	1.71	1.76	1.81	1.86	1.89	1.92
500	1.53	1.57	1.62	1.68	1.73	1.79	1.84	1.90	1.95	1.98	2.01
600	1.59	1.63	1.68	1.74	1.80	1.85	1.91	1.97	2.03	2.06	2.09
700	1.64	1.68	1.74	1.79	1.85	1.91	1.97	2.03	2.09	2.13	2.16
800	1.69	1.73	1.79	1.84	1.91	1.97	2.03	2.09	2.15	2.19	2.22
900	1.74	1.78	1.83	1.89	1.95	2.01	2.08	2.14	2.21	2.25	2.28
1000	1.78	1.82	1.87	1.93	2.00	2.06	2.12	2.19	2.26	2.30	2.34
1500	1.95	1.99	2.04	2.11	2.18	2.25	2.32	2.39	2.47	2.52	2.56
2000	2.08	2.12	2.18	2.24	2.32	2.39	2.47	2.55	2.64	2.69	2.73
2500	2.20	2.23	2.29	2.36	2.44	2.52	2.60	2.68	2.77	2.83	2.88
3000	2.29	2.33	2.39	2.46	2.54	2.62	2.71	2.80	2.89	2.95	3.00
3500	2.38	2.41	2.47	2.55	2.63	2.71	2.80	2.90	3.00	3.06	3.11
4000	2.46	2.49	2.55	2.63	2.71	2.80	2.89	2.99	3.09	3.15	3.21
4500	2.53	2.56	2.62	2.70	2.78	2.88	2.97	3.07	3.18	3.24	3.30
5000	2.59	2.62	2.69	2.77	2.85	2.95	3.04	3.15	3.26	3.32	3.39
5500	2.65	2.68	2.75	2.83	2.92	3.01	3.11	3.22	3.33	3.40	3.46
6000	2.71	2.74	2.80	2.89	2.98	3.07	3.18	3.28	3.40	3.47	3.54
6500	2.76	2.79	2.86	2.94	3.03	3.13	3.24	3.35	3.47	3.54	3.60
7000	2.81	2.84	2.91	2.99	3.09	3.19	3.29	3.41	3.53	3.60	3.67
7500	2.85	2.89	2.96	3.04	3.14	3.24	3.35	3.46	3.59	3.66	3.73
8000	2.90	2.93	3.00	3.09	3.18	3.29	3.40	3.51	3.64	3.72	3.79
8500	2.94	2.97	3.04	3.13	3.23	3.34	3.45	3.57	3.69	3.77	3.84
9000	2.98	3.01	3.09	3.17	3.27	3.38	3.49	3.61	3.75	3.82	3.90
9500	3.02	3.05	3.13	3.22	3.32	3.43	3.54	3.66	3.79	3.87	3.95
10 000	3.06	3.09	3.16	3.26	3.36	3.47	3.58	3.71	3.84	3.92	4.00

Table 5. Parameter $t_{\max}$ for the multi-parameter problem as a function of M and γ for $P = 50$. (The quantity $k_P t_{\max}$ must be used to translate to the confidence interval obtained using the differential-correction or Monte Carlo methods; the values of k_P were taken from Table 6)

M	γ										
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
50	1.03	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.01	1.01
60	1.01	1.02	1.02	1.02	1.02	1.02	1.02	1.02	1.03	1.03	1.03
70	1.00	1.01	1.02	1.02	1.02	1.03	1.03	1.03	1.04	1.04	1.04
80	0.991	1.00	1.01	1.02	1.03	1.03	1.03	1.04	1.04	1.05	1.05
90	0.982	1.00	1.01	1.02	1.03	1.03	1.04	1.05	1.05	1.06	1.07
100	0.973	0.996	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.08
200	0.913	0.967	1.00	1.03	1.05	1.07	1.08	1.10	1.12	1.14	1.16
300	0.881	0.953	0.998	1.03	1.06	1.09	1.11	1.14	1.17	1.19	1.22
400	0.864	0.947	1.00	1.04	1.08	1.11	1.14	1.17	1.21	1.24	1.27
500	0.856	0.945	1.00	1.05	1.09	1.13	1.16	1.20	1.24	1.27	1.32
600	0.852	0.946	1.01	1.06	1.11	1.15	1.18	1.22	1.27	1.30	1.35
700	0.852	0.948	1.02	1.07	1.12	1.16	1.20	1.25	1.30	1.33	1.39
800	0.854	0.952	1.02	1.08	1.13	1.18	1.22	1.27	1.32	1.36	1.42
900	0.857	0.956	1.03	1.09	1.14	1.19	1.24	1.29	1.34	1.38	1.44
1000	0.861	0.961	1.04	1.10	1.15	1.20	1.25	1.30	1.37	1.41	1.47
1500	0.886	0.989	1.07	1.14	1.20	1.26	1.32	1.38	1.45	1.50	1.58
2000	0.913	1.02	1.10	1.18	1.25	1.31	1.37	1.44	1.52	1.58	1.67
2500	0.939	1.04	1.13	1.21	1.28	1.35	1.42	1.49	1.58	1.65	1.74
3000	0.964	1.07	1.16	1.24	1.32	1.39	1.46	1.54	1.64	1.70	1.80
3500	0.987	1.09	1.18	1.27	1.35	1.42	1.50	1.58	1.68	1.75	1.86
4000	1.01	1.12	1.21	1.29	1.37	1.45	1.53	1.62	1.72	1.80	1.91
4500	1.03	1.14	1.23	1.32	1.40	1.48	1.57	1.66	1.76	1.84	1.95
5000	1.05	1.16	1.25	1.34	1.43	1.51	1.60	1.69	1.80	1.88	2.00
5500	1.06	1.17	1.27	1.36	1.45	1.53	1.62	1.72	1.83	1.92	2.04
6000	1.08	1.19	1.29	1.38	1.47	1.56	1.65	1.75	1.87	1.95	2.07
6500	1.10	1.21	1.31	1.40	1.49	1.58	1.67	1.77	1.90	1.98	2.11
7000	1.11	1.22	1.32	1.42	1.51	1.60	1.70	1.80	1.92	2.01	2.14
7500	1.13	1.24	1.34	1.44	1.53	1.62	1.72	1.82	1.95	2.04	2.18
8000	1.14	1.25	1.36	1.45	1.55	1.64	1.74	1.85	1.98	2.07	2.21
8500	1.15	1.27	1.37	1.47	1.56	1.66	1.76	1.87	2.00	2.09	2.24
9000	1.17	1.28	1.39	1.48	1.58	1.68	1.78	1.89	2.02	2.12	2.26
9500	1.18	1.30	1.40	1.50	1.60	1.70	1.80	1.91	2.05	2.14	2.29
10 000	1.19	1.31	1.41	1.51	1.61	1.71	1.82	1.93	2.07	2.17	2.32

Table 6. Recalculation coefficient $k_P(\gamma)$ for the parameter $t_{\max}$ from Tables 2−5 in the ratio of the projection of the confidence region obtained for the χ^2_M method to the confidence interval obtained for the differential-correction or Monte Carlo methods

P	γ										
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
3	6.08	3.96	3.10	2.61	2.28	2.04	1.85	1.68	1.52	1.43	1.31
6	11.8	6.92	5.08	4.08	3.43	2.96	2.59	2.28	1.98	1.81	1.59
10	17.6	9.81	7.00	5.49	4.53	3.85	3.31	2.86	2.43	2.18	1.87
50	48.9	25.4	17.3	13.1	10.4	8.56	7.14	5.95	4.83	4.19	3.39

values of the coefficient $k_P(\gamma) = \sqrt{\dfrac{\Delta_P(\gamma)}{\Delta_1(\gamma)}}$ [this expression follows from (47) and is analogous to the expression for $\delta_P(\gamma)$], which must be applied to the values of $t_{\max}$ from Tables 2−5. As is described for the one-dimensional case above, the quantity $k_P t_{\max}$ in the multi-dimensional case is the ratio of the projection of the χ^2_M confidence region to the differential-correction or Monte Carlo confidence interval.

Table 7 and Fig. 6 present values of $k_P t_{\max}$ for the case $P = 10$. The character of the numerical values of $k_P t_{\max}$ from Table 7 is very close to the values of $t_{\max}$ from Table 1, which shows the ratio of the χ^2_M to the differential-correction error intervals. A comparison of Fig. 5 and Fig. 6 illustrates this clearly. Thus, in the multi-parameter case, the numerical ratios of the χ^2_M and differential-correction error intervals remain as before (their exact values are obtained from (56) and (61); see also Tables 2−6).

Recall that, in the one-dimensional case, the probabilities that the true parameter value falls in the projection of the χ^2_P confidence region and in the differential-correction confidence interval are equal. In the multi-dimensional case, the probability that the true parameter value falls into the projection of the χ^2_P confidence region is higher than the corresponding probability for the differential-correction confidence interval.

Note that the results presented above were obtained for a linear model. The use of a χ^2_P statistic to construct the confidence intervals was based on the fact that the minimum residual is distributed as χ^2_{M-P}. However, this assertion is valid only for a model that depends linearly on its parameters. When this dependence is not linear, this assertion can be valid only asymptotically, as $M \to \infty$ [5, 8]. Therefore, the results obtained can also be applied to non-linear models, but only approximately, in an asymptotic sense.

5.2. Significance Level Corresponding to the Reduced Chi-Squared Statistic

The adequacy of a model to describe observational data is often judged based on the nearness to unity of the minimum value of the residual $\hat{R} \equiv R/(M - P)$, which is distributed as $\hat{\chi}^2_{M-P}$ (as is the random quantity $\dfrac{\xi}{M - P}$, if $\xi \sim \chi^2_{M-P}$), where P is the number of parameters over which the minimization is carried out (an alternative criterion is nearness of the residual $R_{\min}$, distributed as χ^2_{M-P}, to the value $M - P$). This omits the quantitative characteristic of the adequacy of the model, namely the significance level $\alpha_0 = 1 - \gamma$ at which the model can be rejected (recall that α_0 is the number of errors of the first kind that will occur if we reject the model). Figure 7 presents a graphical depiction of the solution of the equation
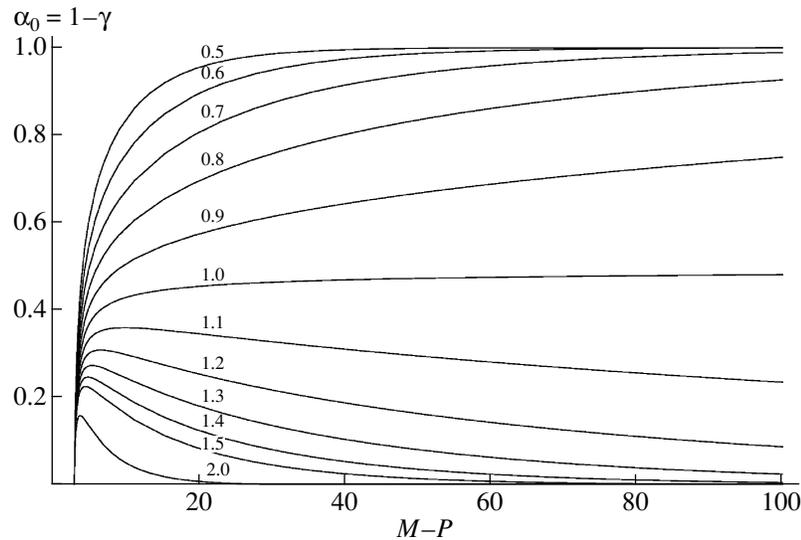
$$\frac{\Delta_{M-P}(1 - \alpha_0)}{M - P} = q$$

for the significance level α_0 as a function of the difference $M - P$ (where M is the number of observational points, P the number of unknown parameters, and $\Delta_{M-P}(\gamma)$ a function of the dependence of the quantile on the confidence level $\gamma = 1 - \alpha_0$) for a specified quantile q for the $\hat{\chi}^2_{M-P}$ distribution. The value of q was taken to be 0.5, 0.6, 0.7, 1.0, 1.1, 1.5, and 2.0. This dependence can be written in the explicit form

$$\alpha_0(M - P) = 1 - \chi^2_{M-P}(q(M - P)).$$

Figure 7 shows that α_0 begins to approach its asymptotic limit fairly rapidly (for $M \gtrsim 50$). For example, for $q = 1$, the asymptotic limit of α_0 is $\sim 50\%$; i.e., in this case, we will make an error of the first kind (rejecting a correct model) in $\sim 50\%$ of cases if we reject the model. Consequently, there is no serious basis to reject the model, and the model should therefore be accepted.

Table 7. Parameter $k_P t_{\max}$ for the multi-parameter problem as a function of M and γ for $P = 10$ (see also Fig. 6)

M	γ										
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.99
10	20.4	11.1	7.82	6.07	4.96	4.17	3.56	3.05	2.57	2.29	1.95
20	18.9	10.8	7.81	6.18	5.13	4.37	3.77	3.26	2.77	2.49	2.12
30	18.5	10.8	7.90	6.31	5.27	4.52	3.93	3.41	2.92	2.62	2.24
40	18.5	10.9	8.01	6.44	5.41	4.65	4.06	3.54	3.03	2.73	2.34
50	18.7	11.0	8.13	6.56	5.53	4.77	4.17	3.65	3.13	2.83	2.43
60	18.9	11.2	8.26	6.67	5.63	4.87	4.27	3.74	3.22	2.91	2.50
70	19.2	11.3	8.38	6.78	5.74	4.97	4.36	3.83	3.30	2.98	2.57
80	19.5	11.5	8.50	6.88	5.83	5.06	4.44	3.90	3.37	3.05	2.63
90	19.8	11.6	8.61	6.98	5.92	5.14	4.52	3.97	3.44	3.11	2.68
100	20.0	11.8	8.72	7.08	6.00	5.22	4.59	4.04	3.50	3.17	2.73
200	22.4	13.0	9.63	7.83	6.66	5.82	5.14	4.55	3.96	3.60	3.11
300	24.2	14.0	10.3	8.39	7.14	6.25	5.53	4.90	4.28	3.89	3.38
400	25.6	14.7	10.9	8.84	7.53	6.59	5.84	5.19	4.53	4.13	3.58
500	26.8	15.4	11.4	9.22	7.86	6.88	6.10	5.42	4.74	4.32	3.76
600	27.9	16.0	11.8	9.56	8.15	7.13	6.33	5.63	4.93	4.49	3.91
700	28.9	16.5	12.2	9.86	8.40	7.36	6.53	5.81	5.09	4.64	4.04
800	29.7	17.0	12.5	10.1	8.63	7.56	6.71	5.98	5.24	4.78	4.16
900	30.5	17.4	12.8	10.4	8.85	7.75	6.88	6.13	5.37	4.90	4.27
1000	31.2	17.8	13.1	10.6	9.04	7.92	7.04	6.27	5.49	5.02	4.37
1500	34.2	19.5	14.3	11.6	9.86	8.64	7.68	6.85	6.01	5.49	4.79
2000	36.6	20.8	15.2	12.3	10.5	9.20	8.18	7.30	6.41	5.86	5.11
2500	38.6	21.9	16.0	13.0	11.0	9.67	8.60	7.68	6.74	6.17	5.38
3000	40.2	22.8	16.7	13.5	11.5	10.1	8.97	8.00	7.03	6.44	5.62
3500	41.7	23.7	17.3	14.0	11.9	10.4	9.29	8.29	7.29	6.67	5.82
4000	43.1	24.4	17.8	14.4	12.3	10.8	9.58	8.55	7.52	6.88	6.01
4500	44.3	25.1	18.3	14.8	12.6	11.1	9.84	8.79	7.73	7.08	6.18
5000	45.5	25.7	18.8	15.2	12.9	11.3	10.1	9.00	7.92	7.25	6.33
5500	46.5	26.3	19.2	15.5	13.2	11.6	10.3	9.21	8.10	7.42	6.48
6000	47.5	26.9	19.6	15.8	13.5	11.8	10.5	9.40	8.27	7.57	6.61
6500	48.4	27.4	20.0	16.1	13.7	12.0	10.7	9.57	8.43	7.72	6.74
7000	49.3	27.9	20.3	16.4	14.0	12.3	10.9	9.74	8.58	7.86	6.86
7500	50.1	28.3	20.7	16.7	14.2	12.5	11.1	9.90	8.72	7.99	6.98
8000	50.9	28.8	21.0	17.0	14.4	12.6	11.3	10.1	8.85	8.11	7.09
8500	51.7	29.2	21.3	17.2	14.6	12.8	11.4	10.2	8.98	8.23	7.19
9000	52.4	29.6	21.6	17.4	14.8	13.0	11.6	10.3	9.10	8.34	7.29
9500	53.1	30.0	21.9	17.7	15.0	13.2	11.7	10.5	9.22	8.45	7.38
10 000	53.7	30.3	22.1	17.9	15.2	13.3	11.9	10.6	9.34	8.55	7.48

Fig. 7. Solution of the equation $\frac{\Delta_{M-P}(1-\alpha_0)}{M-P} = q$ for the significance level α_0 for $P = 3$ as a function of the number of observational points M for a specified quantile q of the reduced statistic $\hat{\chi}^2_{M-P}$. The value of q is indicated next to the corresponding curve. When the minimum value of the reduced chi-squared is equal to unity, the corresponding significance level α_0 approaches 0.5.

At the same time, when $q = 1.5$, the asymptotic limit for α_0 is already close to $\sim 0\%$. In this case, we will make an error of the first kind in $\sim 0\%$ of cases if we reject the model. Thus, here, there is no basis to accept the model, and the model should be rejected.

Thus, we conclude that, if the reduced chi-squared for the solution of an inverse parametric problem is $\gtrsim 1.5$, the model used is clearly inadequate to the observational data. The resulting values of the unknown parameters and their confidence intervals (errors) are "poor", and it is necessary to construct a better model.

6. FITTING THE OBSERVED LIGHT CURVE OF THE YZ Cas BINARY SYSTEM

Here, we analyze the normalized light curve of the eclipsing binary system YZ Cas obtained in a red filter ($\lambda = 6700$ Å) by Kron [6]. The observed light curve included $M = 42$ points, $\xi_1 \ldots \xi_{42}$. The dispersions of each brightness measurement in magnitudes were taken to be $(\sigma_m^{obs})^2 = 1.77 \times 10^{-6}$ [8], or $(\sigma_m^{obs})^2 = 1.5015 \times 10^{-6}$ on our (intensity) scale, taking into account that $L^{full} = 1$. The central value of each point in the light curve was obtained by averaging $N_m = 12$ points for $m = 1 \ldots 42$ [3, 6].

We supplemented the fits of the observed light curve of YZ Cas presented in [1] with fits for the limb-darkening coefficient assuming a linear limb-darkening law. Thus, our fitting of the light curve was carried out over six parameters: r_1, r_2, i, x_1, x_2,

$I_0^{(1)}/I_0^{(2)}$. The minimization of the residual R between the observed and theoretical light curves was carried out using a gradient-descent method. The central values and confidence intervals of the unknown parameters are presented in Table 8.

The calculation of the central values and error intervals of r_1, r_2, i in the differential-correction and confidence-region methods is described in detail in [1], and we do not present this information here. Note that we focus here specifically on the calculation of the central values and confidence intervals of the limb-darkening coefficients, together with verification of the reliability of the resulting confidence intervals. The errors in the limb-darkening coefficients in the differential-correction method were obtained using (41).

Since the use of statistics distributed as χ^2_P and χ^2_M requires knowledge of the unit-weight dispersion ε_0, which is not precisely known for a real observed light curve, we analyzed a model binary system whose phases $\theta_1 \ldots \theta_{42}$ coincided with the observed phases for YZ Cas. The true parameter values for the binary system were set equal to the central values obtained from the fit of the observed light curve using a linear limb-darkening model [1], i.e., $\bar{r}_1 = 0.14408$, $\bar{r}_2 = 0.07556$, $\bar{i} = 88.27°$, $x_1 = 0.2998$, $x_2 = 0.4071$. The unit-weight dispersion was set equal to the rms estimate v_0 obtained from this fitting of the observed light curve, $\varepsilon_0 = 0.003422904$. The weighting coefficients w_m and number of measurements at each
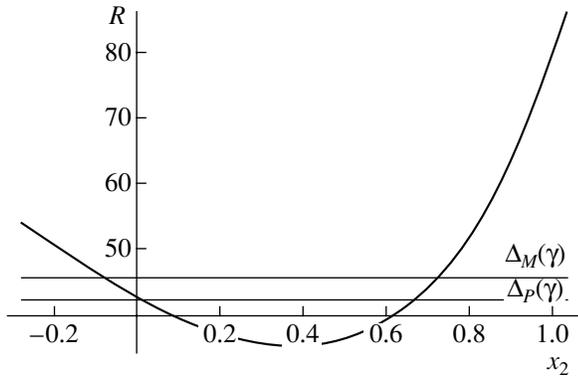
Fig. 8. Model observed light curve for the binary system YZ Cas. The residual R for the χ^2_M method ($M = 42$) for the limb-darkening coefficient x_2 of the secondary (star with a smaller radius), minimized over all remaining parameters ($r_1, r_2, i, x_1, I^1_0/I^2_0$). The rms estimated unit-weight dispersion of the light curve corresponds to $v_0 = 0.003423$. The straight lines show the quantiles for χ^2_M and χ^2_P distributions for the confidence level $\gamma = 0.6827$ (when using χ^2_P statistics, the quantity $\Delta_P(\gamma)$ is measured from the minimum residual).

from an analysis of another realization of the light curve (for example, obtained at another epoch).

Our numerical simulations showed that reliably determining the limb-darkening coefficient of the smaller star in the YZ Cas binary system requires increasing the accuracy of the observed light curve by approximately an order of magnitude. We perturbed the light curve constructed with the "true" parameter values, for which we adopted the central values from Table 8: $\bar{r}_1 = 0.14408$, $\bar{r}_2 = 0.07556$, $\bar{i} = 88.27°$, $x_1 = 0.2998$, $x_2 = 0.4071$. Here, the unit-weight dispersion root square was taken to be $\varepsilon_0 = $
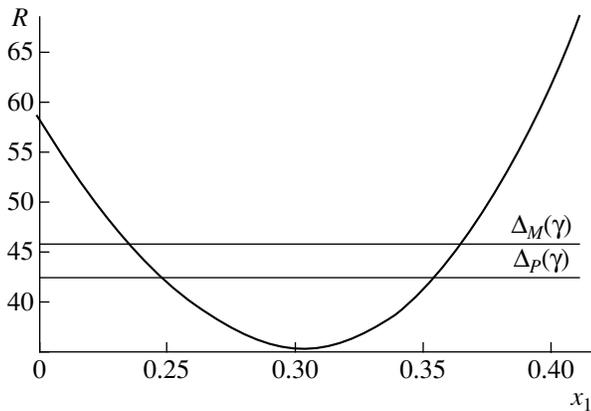

Fig. 9. Same as Fig. 8 for the primary component (star with the larger radius) and minimization over the parameters $r_1, r_2, i, x_2, I^1_0/I^2_0$.

$\sigma^{obs}_m/\sqrt{w_m} = 0.0003422904$ (where the weight of each point is $w_m = N_m = 12$). Further, we fit the resulting perturbed light curve.

Figure 10 presents the behavior of the residual obtained using the χ^2_M method to calculate the limb-darkening coefficient x_2. At the chosen confidence level, $\gamma = 0.6827$, the model is adequate to the observational data, with $x_2 = 0.4205 \pm 0.0399$. Thus, to reliably determine the limb-darkening coefficients in classical binaries whose smaller components have radii similar to that of the secondary component in YZ Cas, the root square from unit-weight dispersion must be close to $\varepsilon_0 \simeq 10^{-4}$. Such high-accuracy light curves can be obtained with the COROT and KEPLER orbiting photometric stations.

Analysis of the main minimum (annular eclipse) indicates that the limb-darkening coefficient of the larger star x_1 is found reliably with $\varepsilon_0 = 0.0034$: $x_1 = 0.3002 \pm 0.06488$ for the χ^2_M method, where this indicates the projection of the six-dimensional confidence region onto the x_1 axis.

6.1. Use of a Quadratic Limb-Darkening Law

We also fit the observed light curve assuming a quadratic limb-darkening law [see (3)]. Here, we obtained two different fits.

In the first case, the unknown parameters were r_1, r_2, i, x_1, x_2, y_1, and y_2; in the second, the coefficients x_1 and x_2 were fixed. We used the values of Van Hamme [10] for x_1 and x_2, and the limb-darkening law for the smaller component was assumed to be linear. The fitting amounted to searching for the parameters r_1, r_2, i, and y_1.

The results of this fitting are presented in Table 10. The errors were obtained using the differential-correction method (for the confidence level 0.6827). When fitting the observed light curve over the entire collection of parameters, the values of x_1, x_2, y_1, and
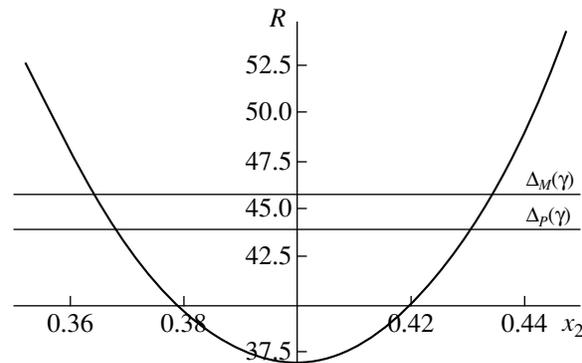

Fig. 10. Same as Fig. 8 for the perturbed light curve. See text for details.

Table 10. Results of fitting the observed light curve of YZ Cas using a quadratic limb-darkening law ($\varepsilon_0 = 0.003422904$)

r_1	r_2	i, deg	x_1	x_2	y_1
0.144324 ± 0.000308	0.753775 ± 0.000220	88.22 ± 0.099	0.331	0.383	-0.0134 ± 0.0100

Table 11. Half the error intervals for the limb-darkening coefficients for the YZ Cas components derived by fitting the model observed light curve using a linear limb-darkening law, and the number of times the "true" values fall in these intervals, $N_{\Delta x_1}$ and $N_{\Delta x_2}$. (N is the number of times the exact solution falls in the error interval. $\varepsilon_0 = 0.003422904$. The number of times the exact solution falls in the confidence region D is indicated in parantheses.)

Method	$N_{\Delta x_1}$	$x_1 \pm \Delta x_1$	$N_{\Delta x_2}$	$x_2 \pm \Delta x_2$	N
Differential-correction	6892	0.3052 ± 0.01942	6957	0.3759 ± 0.1184	4916
Monte Carlo	6899	0.3009 ± 0.02011	6884	0.4934 ± 0.1056	4871
Confidence region, χ_P^2	9722	0.3013 ± 0.05321	9842	0.3403 ± 0.3280	9634 (6898)
Confidence region, χ_M^2	8234	0.3002 ± 0.06488	8413	0.3213 ± 0.4015	8166 (6858)
Confidence region, $\mathbf{F}_{M,N-M}$	8112	0.3007 ± 0.0589	8306	0.3316 ± 0.3637	8060 (6827)

y_2 lie outside the physically reasonable range and are thus artefacts, although the best-fit light curve for these values fits the observed light curve well, and is very close to the light curve obtained for a linear limb-darkening law.

In the second case, $y_1 = -0.0134 \pm 0.0100$. The best-fit light curve is nearly identical (with accuracy to the third digit after the decimal point) to the best-fit light curve obtained using a linear limb-darkening law (Fig. 11).

Thus, with $\varepsilon_0 = 0.0034$, it is fully sufficient to restrict our analysis to a linear limb-darkening law when fitting the observed light curve. The corrections for non-linearity of this law are negligibly small. The test of reliability of the error intervals for the limb-darkening coefficients is satisfied for a linear law.

6.2. Test of Reliability of the Limb-Darkening Coefficients and Their Errors

We verified the trustworthiness of the error intervals for x_1 and x_2 calculated for each method. We carried out 10 000 perturbations of the model observed light curve (for $\varepsilon_0 = 0.003422904$) calculated with the "true" parameter values, taken to be the central values from Table 8: $\bar{r}_1 = 0.14408$, $\bar{r}_2 = 0.07556$, $\bar{i} = 88.27°$, $x_1 = 0.2998$, $x_2 = 0.4071$. We counted the number of times the "true" values fell in the corresponding confidence regions or their projections (Table 11).

The confidence intervals for x_1 and x_2 calculated as the projection onto the x_1, x_2 parameter axes of the six-dimensional confidence regions obtained using χ_P^2, χ_M^2, and Fisher statistics ($F_{M,N-M}$ in the case

with x_1, x_2 r_1, r_2, i, $I_0^{(1)}/I_0^{(2)}$, and $F_{M-3,N-M}$ in the case with r_1, r_2, i) by solving the corresponding inequalities for r_1, r_2, i, x_1, x_2, and $I_0^{(1)}/I_0^{(2)}$.

Table 11 shows that the number of times the exact solution falls in the individual error intervals for the differential-correction and Monte Carlo methods corresponds to the probability $\gamma = 68.2\%$, but the number of joint "hits" is less than this probability ($N \simeq 4900$, which is $\sim$30% lower than the value $N \simeq 6800$ corresponding to $\gamma = 0.6827$). In this case, the confidence intervals for x_1 and x_2 are projections of
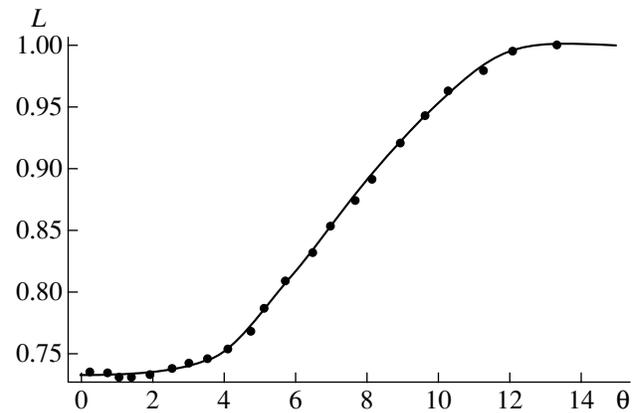

Fig. 11. Observed light curve for YZ Cas (annular eclipse) and the best-fit theoretical light curves obtained for linear and quadratic limb-darkening laws (the curves coincide with accuracy to within the third digit after the decimal place). The points show the observational data from [6].
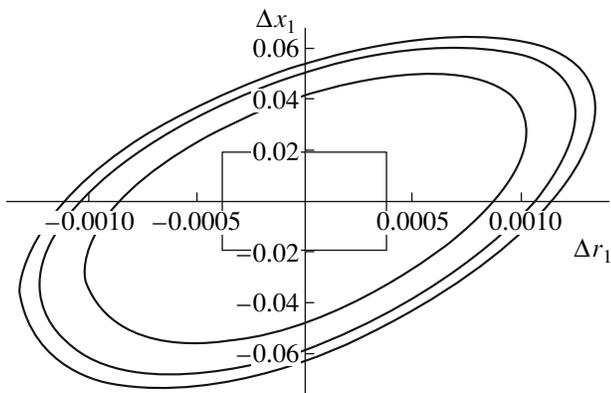
Fig. 12. Projection of the six-dimensional confidence region for the parameters r_1, r_2, i, x_1, x_2, and $I_0^{(1)}/I_0^{(2)}$ onto the plane of the parameters r_1, x_1, obtained by fitting the model observed light curve for YZ Cas using statistics distributed as χ_P^2 (small ellipse), the Fisher distribution $F_{M,N-M}$ (medium ellipse), and χ_M^2 (large ellipse). The rectangle in the center is the confidence region obtained using the differential-correction or Monte Carlo method. The deviations of the parameters from the central values are shown. The adopted unit-weight dispersion square root is $\varepsilon_0 = 0.003423$. The confidence intervals obtained using the differential-correction and Monte Carlo methods (using a normally distributed statistic) cover the exact values of the parameters with the specified probability, $\gamma = 0.68$, while the corresponding confidence region (central rectangle) covers the exact solution with a probability that is lower than γ, equal to 0.49. The confidence regions D obtained using χ_P^2, $F_{M,N-M}$, and χ_M^2 statistics cover the exact solution with probability $\gamma = 0.68$, while the corresponding confidence intervals (projections of the confidence region onto the r_1 and x_1 parameter axes) cover the exact parameter values with a probabilities higher than γ, namely 0.9904 and 0.9722 for the χ_P^2 statistic, 0.8456 and 0.8234 for the χ_M^2, and 0.8368 and 0.8112 for the $F_{M,N-M}$ statistic.

the six-dimensional confidence region for r_1, r_2, i, x_1, x_2, and $I_0^{(1)}/I_0^{(2)}$ onto the x_1 and x_2 parameter axes. Falling into the confidence region D fully corresponds to the specified confidence level $\gamma = 68.2\%$. Since the two-dimensional projection onto the (x_1, x_2) plane is an elongated "ellipse," number of coincidences with the projection of the "ellipse" onto the x_1 and x_2 parameter axes appreciably exceeds the confidence level $\gamma = 68.2\%$. The situation when the χ_P^2 method is used is analogous. Therefore, the parameter errors obtained using the confidence-area method correspond to a confidence level $>\gamma$, and are most preferred from the point of view of reliability. Figure 12 presents a comparison of the projected confidence regions for the parameters r_1 and x_1 obtained using various methods. Note that in our previous paper [1] small rectangles show, as in Fig. 12 of the given paper,
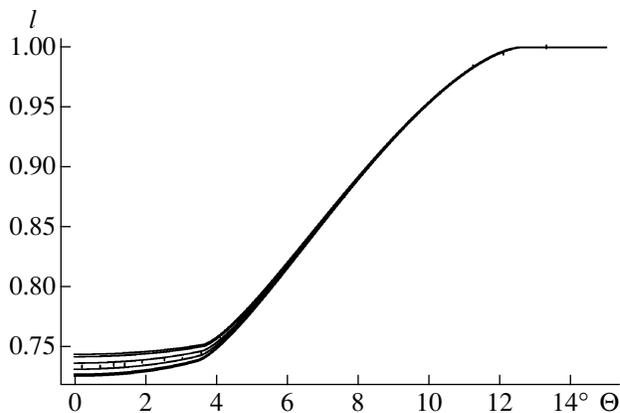

Fig. 13. Primary minimum (annular eclipse) of the model observed light curve of YZ Cas together with theoretical light curves obtained for the maximum and minimum values of r_1, r_2, x_1, and x_2 determined using the χ_M^2 method (outer curves), χ_P^2 method (middle curves), and differential-correction method (inner curves). The adopted unit-weight dispersion square root is $v_0 = 0.003423$.

the confidence regions obtained by the differential-correction method.

In spite of the fact that the most probable confidence interval obtained using the χ_M^2 method exceeds the corresponding interval obtained using the χ_P^2 method, the χ_P^2 confidence interval contains the true value of the parameter in $\sim99\%$ of cases, and the χ_M^2 confidence interval in $\sim85\%$ of cases (for a one-dimensional projection). This reflects the fact that the probability of falling into the confidence intervals obtained using the χ_M^2 method is determined largely, not by the most probable value, but by the character of the distribution of these intervals, which are themselves random quantities.

6.3. "Internal" and "External" Errors

Our results clearly demonstrate and explain the qualitative difference between the magnitudes of parameter errors derived using the differential-correction and confidence-region methods.

We constructed theoretical light curves using the upper and lower values for r_1, r_2, x_1, and x_2 from Table 9 and fixing the orbital inclination of the binary to be $i = 88.28°$ (Fig. 13; see also Fig. 12). The data in these theoretical light curves form an "error corridor", within which the observed values of the given realization of the light curve are located. The error corridor for the differential-correction method is fairly narrow, and some of the "observed" points in the given realization densely press against its walls.

The "observed" points in some realizations of the light curve may lie outside the indicated error corridor.

The analogous error corridor for r_1, r_2, x_1, and x_2 derived using the χ^2_M method (Fig. 13) is appreciably broader, and clearly allows for the statistical unpredictability of the light-curve realizations.

We emphasize again that, in adopting errors found with the differential-correction or Monte Carlo methods, we must bear in mind that these include only "internal" errors, which may not overlap with the "internal" errors of the parameters found from an analysis of another realization of the light curve (for example, obtained at a different epoch). The probability that the true parameters all fall into their error intervals in the differential-correction method is substantially lower (by a factor of $\sim 1.3 - 1.4$) than the probability for each error interval separately.

The basis for calling the errors determined using the differential-correction and Monte Carlo methods "internal" is that the "internal" statistical distribution of the derived central values of the parameters are used when computing these errors. The calculation of the parameter errors in the confidence-region method uses the "external" statistical distribution of the observed light-curve values, so that the errors yielded by this method can be considered to be "external".

7. DISCUSSION

Our computations have shown that trustworthy derivation of both limb-darkening coefficients based on the light curve for the eclipsing system YZ Cas is possible only if the accuracy of the individual measurements is $\varepsilon_0 \simeq 10^{-4}$ (on an intensity scale with $L_{full} = 1$). It is not possible to determine the limb-darkening coefficient for the smaller star reliably when $\varepsilon_0 \simeq 10^{-3}$. Not only the absolute accuracy of the individual brightness measurements for the system is important, but also the depth of the eclipse. For example, in the YZ Cas system, the depth of the secondary minimum (full eclipse) is $\sim 10\%$ of the total brightness of the system, which is fully sufficient to determine the limb-darkening coefficient with $\varepsilon_0 \simeq 10^{-4}$. However, if the depth of this eclipse were smaller, the required accuracy of the individual measurements would be higher. When fitting the primary minimum of the YZ Cas light curve (annular eclipse), we are able to reliable determine the linear limb-darkening coefficient $x_1 = 0.2917 \pm 0.1199$ for $\varepsilon_0 \simeq 0.0034$.

The light curves obtained using linear and quadratic limb-darkening laws were identical, indicating the smallness of non-linear limb-darkening effects. This makes it possible to fit the light curve reliably using a linear limb-darkening law when the accuracy of the light curve measurements is $\sim 10^{-3}$.

Table 11 shows that the individual error intervals for x_1 and x_2 derived using the Monte Carlo and differential-correction methods contain the true values in $\sim 68\%$ of cases, while the both parameters lie in their error intervals simultaneously in only $\sim 49\%$ of cases. This reflects the non-reliability of the differential-correction and Monte Carlo methods in terms of the simultaneous incidence of the unknown parameter values in their error intervals.

Recall that the confidence intervals for x_1 and x_2 in the χ^2_M method are projections of the six-dimensional region of the parameters r_1, r_2, i, x_1, x_2, and $I_0^{(1)}/I_0^{(2)}$ onto the x_1 and x_2 parameter axes. The probability that a parameter falls in the confidence region D corresponds fully to the specified confidence level $\gamma = 68.2\%$ (Table 11). Since the two-dimensional projection onto the (x_1, x_2) plane forms an elongated "ellipse," the number of "hits" in this projected "ellipse" appreciably exceeds the confidence level $\gamma = 68.2\%$. The situation for the χ^2_P method is analogous. Therefore, the parameter errors obtained using the confidence-region method correspond to $> \gamma$, and are preferable in terms of their trustworthiness.

For example, Table 1 shows that, in the case of a one-parameter model with $\gamma = 0.90$, with the number of observational points $M = 10 - 10\,000$, the parameter $t_{\max}$ varies from 1.89 to 9.19; i.e., this is the factor by which the error of an unknown parameter is reduced when derived using the Monte Carlo or differential-correction methods.

One of the important results of our analysis are the analytical relationships between the probable values of the confidence intervals obtained using the χ^2_M and the errors obtained using the differential-correction method for one-parameter and multi-parameter problems [see (56) and (61)]. These show that $t_{\max}$ (and $k_P t_{\max}$) decrease with growth in γ, while this parameter grows comparatively weakly with growth in M (Figs. 3–6).

We again note that formulas (56) and (61) (for a linear dependence of the function describing the model on the unknown parameters) can be used to obtain the error interval for an unknown parameter for the χ^2_M confidence-region method that is most probable among those for adequate models from the corresponding error interval in the Monte Carlo or differential-correction method via multiplication by the coefficient $t_{\max}$ or $k_P t_{\max}$ (see also Tables 1–7). After finding the error interval for the Monte Carlo or differential-correction method and verifying that the model is adequate in the χ^2_M method (for which it is sufficient to calculate the minimum of the corresponding residual), we can draw preliminary conclusions of a probabilistic nature about the size of the χ^2_M confidence interval, without turning directly to

the more labor-intensive determination of the sizes of the projections of the confidence region onto the parameter axes.

Briefly summarizing the results, the differential-correction and Monte Carlo methods yield estimates only of the "internal" parameter errors. At the same time, the confidence-region method provides estimates of the "external" errors of unknown parameters (see in this connection the study of Popper [11]).

Let us make one more important comment. The magnitude of the parameter errors obtained in the differential-correction, Monte Carlo, and χ_P^2 methods do not depend on the sample volume (number of points M in the light curve), and are stable in the sense that, the parameter errors for different realizations of a random process (light curve) with other conditions the same will differ only comparatively weakly (the corresponding differences are much less than 100%). These "good" properties of these methods are associated with the fact that they assume a priori that the model used is fully correct (the possibility of an error of the second kind is taken to be negligible).

In the case of the χ_M^2 and Fisher $F_{M,N-M}$ methods, there is no such *a priori* assumption that the model is fully correct, and the adequacy of the model is verified using the observtional data in the process of constructing the confidence regions. Therefore, in the χ_M^2 and $F_{M,N-M}$ methods, the resulting parameter errors depend on the sample volume M and vary strongly from one realization of the random process to another (for some realizations, the errors can even degenerate into an empty set). The dependence of the parameter errors in the χ_M^2 and $F_{M,N-M}$ methods on the sample volume M is very weak. For example, as follows from Table 1, when the number of points M is varied by a factor of two, from 100 to 200, the most probable value of relative error in the χ_M^2 method varies for $\gamma = 0.7$ by a factor of $\sim$17%; when the sample volume is varied by three orders of magnitude (from $M = 10$ to $M = 10\,000$), other conditions being equal, the most probable value of relative error varies by only a factor of $\sim$5. On the other hand, it is possible to obtain reliable estimates of parameter errors when a model is rejected at a fairly high significance level (corresponding to a reduced chi-squared that is close to unity) in the χ_M^2, $F_{M,N-M}$, and χ_P^2 methods.

8. CONCLUSION

We have shown that determining the errors of parameters in the differential-correction and Monte Carlo methods leads to a substantial reduction in the magnitude of the errors. These methods yield only the "internal" errors of the parameters for a given realization of an observed light curve. The parameter values for different realizations of a light curve can even not have common values in the error intervals derived using the differential-correction and Monte Carlo methods.

The comparatively small parameter errors obtained in the differential-correction and Monte Carlo methods is due to the application of the a priori assumption that the model is fully correct, and also the application of a simple, normally distributed statistic; at the same time, the solution of the inverse problem is searched for using another statistic—χ_M^2, which is only derived from a normal distribution. It is much more natural to search for the parameter errors using the same statistic that is applied to find the central parameter values, for example, the χ_M^2 statistic. It is desirable not to apply the artificial assumption that the model considered is fully correct; i.e., not to exclude the possibility of an error of the second kind. This requirement is satisfied by the χ_M^2 statistic [12]. Therefore, correct calculation of parameter errors requires the use of the confidence-region method with the χ_M^2 statistic, which makes it possible to judge the "external" errors of model parameters, as well as as the adequacy of the model itself.

Given the labor-intensive nature of computating the sizes of projections of a confidence region onto the parameter axes, we stress the importance of our proposed new method for estimating parameter errors (Tables 1−7, Figs. 4−6). Namely, in the case of a linear dependence of the model function on the unknown parameters, formulas (56) and (61) (see also Tables 1−7) can be used to make preliminary conclusions of a probabilistic nature about the sizes of the χ_M^2 confidence intervals based on the corresponding intervals for the Monte Carlo or differential-correction methods via multiplication by $t_{\max}$ or $k_P t_{\max}$, without the need for the more labor-intensive determination of the projections of the confidence region onto the parameter axes. If the confidence region is understood in an asymptotic sense, our new method for estimating parameter errors can also be applied to non-linear problems.

REFERENCES

1. M. K. Abubekerov, N. Yu. Gostev, and A. M. Cherepashchuk, Astron. Zh. **85**, 121 (2008) [Astron. Rep. **52**, 127 (2008)].
2. B. C. Carlson, arXiv:math/9409227v1 [math.CA] (1994).
3. A. V. Goncharskii, A. M. Cherepashchuk, and A. G. Yagola, *Ill-Posed Problems in Astrophysics* (Nauka, Moscow, 1985), p. 53 [in Russian].
4. B. M. Shchigolev, *Mathematical Processing of Observations* (Fizmatgiz, Moscow, 1962) [in Russian].
5. S. Wilks, *Mathematical Statistics* (Wiley, New York, 1962; Nauka, Moscow, 1967).
6. G. E. Kron, Astrophys. J. **96**, 173 (1942).
7. T. M. Brown, D. Charbonneau, R. L. Gilliland, et al., Astrophys. J. **552**, 699 (2001).
8. A. V. Goncharskii, S. Yu. Romanov, and A. M. Cherepashchuk, *Finite Parametric Inverse Problems in Astrophysics* (Mosc. Gos. Univ., Moscow, 1991) [in Russian].
9. J. Southworth, Mon. Not. R. Astron. Soc. **386**, 1644 (2008).
10. W. Van Hamme, Astron. J. **106**, 2096 (1993).
11. D. M. Popper, Astron. J. **89**, 132 (1984).
12. A. M. Cherepashchuk, Astron. Zh. **70**, 1157 (1993) [Astron. Rep. **37**, 585 (1993)].

Translated by D. Gabuzda