The Mass Distribution of Transit Exoplanets from the Mass–Radius Relationships: the Structurization within Planetary Systems

Article *in* Solar System Research · May 2021

DOI: 10.1134/S0038094621030084

CITATION
1

READS
10

7 authors, including:

Oleg Yakovlev
Special Astrophysical Observatory
14 PUBLICATIONS **30** CITATIONS
SEE PROFILE

Inna Shashkova
Space Research Institute
39 PUBLICATIONS **167** CITATIONS
SEE PROFILE

Jean-Loup Bertaux
LATMOS
1,035 PUBLICATIONS **31,279** CITATIONS
SEE PROFILE

Alexander Tavrov
Space Research Institute
79 PUBLICATIONS **308** CITATIONS
SEE PROFILE

ISSN 0038-0946, Solar System Research, 2021, Vol. 55, No. 3, pp. 200–217. © Pleiades Publishing, Inc., 2021.
Russian Text © The Author(s), 2021, published in Astronomicheskii Vestnik, 2021, Vol. 55, No. 3, pp. 213–231.

The Mass Distribution of Transit Exoplanets from the Mass–Radius Relationships: the Structurization within Planetary Systems

O. Ya. Yakovlev[a, b, *], A. E. Ivanova[a], V. I. Ananyeva[a], I. A. Shashkova[a], A. V. Yudaev[c], J.-L. Bertaux[d], and A. V. Tavrov[a, c]

[a] *Space Research Institute, Russian Academy of Sciences, Moscow, 117997 Russia*
[b] *Moscow State University, Moscow, 119992 Russia*
[c] *Moscow Institute of Physics and Technology, Dolgoprudnii, Moscow oblast, 141701 Russia*
[d] *LATMOS-IPSL, 78280 Guyancourt, France*
**e-mail: yko-v@yandex.ru*

Received October 21, 2020; revised December 18, 2020; accepted January 13, 2021

Abstract—Most transit exoplanets (85%) were discovered with the *Kepler* space telescope. However, the mass, which was measured mainly with the radial velocity method, is known only for ~15% of them. The mass of an exoplanet may be estimated by its radius from the statistical dependences based on the observational data, though no unambiguous interrelation between the mass and the radius of planets exists. Here, we calculate the earlier unknown masses of exoplanets from four statistical mass–radius relationships (Bashi et al., 2017; Chen and Kipping, 2017; Ning et al., 2018, and the averaged dependence derived) and added the results to the distribution of planets with known masses. The mass distributions of transit exoplanets obtained in this way are analyzed with taking into account the observational selection effect inherent in the transit method. The distributions are approximated by the power law $\partial N/\partial M \sim M^{\alpha}$, where the exponent ($\alpha < 0$) is determined by the maximum likelihood estimation for the samplings acquired with four mass–radius relationships: $\alpha = -2.12 \pm 0.03$, -2.09 ± 0.03, -1.94 ± 0.03, and -2.27 ± 0.04. Moreover, for one of these distributions, we determine the parameters of the power law, the exponent of which differs on three intervals (with the boundaries at 0.025, 0.28, and 1.34 Jupiter masses): -1.99, -0.62, and -2.88. We also conclude that there is no evidence of the interrelation between the mass of an exoplanet and its average distance to the host star (the structurization within planetary systems), if this distance is smaller than 1 AU; besides, the dependence of the exponent α on the considered mass interval is analyzed. The above estimates appertain to exoplanets detected by the space telescopes: Kepler Space Telescope and Transiting Exoplanet Survey Satellite (TESS) (these exoplanets compose group 1). The masses of the other transit exoplanets, which were detected by ground-based instruments, were known (they compose group 2). For the latter group, the exponent α is estimated at -2.21 ± 0.04. In general, the results of our analysis agree with those of the earlier statistical and theoretic studies. A key idea of the present paper is to apply the model interrelations between the mass and the radius of exoplanets to the analysis of the mass distribution of exoplanets on the basis of the recent data of observations.

Keywords: exoplanets, transit method, mass distribution, mass–radius relationship, statistical analysis
DOI: 10.1134/S0038094621030084

INTRODUCTION

With increasing the number of detected exoplanets, it has become possible to study the characteristics of exoplanets (their masses, radii, orbital parameters, etc.) in terms of statistics. The statistics of exoplanets is important for developing comparative planetology and verifying theoretical models of the formation and evolution of planetary systems (e.g., the planetary population synthesis model by Mordasini (2018)).

Along with the radius and the orbital parameters, the mass is one of the main characteristics of an exoplanet. The currently available papers, dealing with the mass distributions of exoplanets, are summarized in Table 1. In these papers, it is suggested that, to describe the mass distribution of exoplanets, one should use the power law $\partial N/\partial M \sim M^{\alpha}$, where the exponent α is negative and varies in dependence on the years of studies and the mass intervals.

In the early papers (see Table 1), the mass distributions of exoplanets were determined from the projective mass $M \times \sin i$, where i is the angle between the perpendicular to the orbital plane of the planet and the line of sight. In the column "M or ($M \times \sin i$)" of Table 1, it is indicated whether the corresponding estimates were obtained by the true masses M or the projective masses $M \times \sin i$. When analyzing the essentially non-

Table 1. Characteristics of the mass distributions of exoplanets (according to some papers)

Source	Exponent α in $\partial N/\partial M \sim M^{\alpha}$	M or $M\sin i$	Mass range, M_J^*	Number of planets (a survey)	Accounting for the observational selection
Marcy et al. (2005)	−1.05	$M\times\sin i$	0.02−15	104	–
Butler et al. (2006)	−1.16		<15	167	–
Cumming et al. (2008)	-1.31 ± 0.2		>0.3	182	Accounting for the survey completeness
Howard et al. (2010)	$-1.48^{+0.12}/_{-0.14}$		0.01−3.15	166	
Ivanova et al. (2019) Ananyeva et al. (2020a, 2020b)	-2.12 ± 0.12	M	0.02−13	328 (*Kepler*)	Accounting for the survey completeness: separation into two groups of instruments; accounting for the mass determination and the transit probability
	-1.9 ± 0.06		0.68−13	210 (Ground-based surveys and CoRoT)	
Mordasini (2018)	−1		0.09−5	(Modeling with the planetary population synthesis)	
	−2		<0.09, >5		

* Here after: M_J and R_J are the mass and the radius of Jupiter, respectively; and M_E and R_E are the mass and the radius of the Earth, respectively.

homogeneous data of observations, some authors took into account the observational selection caused by the characteristics of detecting instruments or surveys. In the last column of the table, it is indicated whether the corresponding observational or archive data were regularized and the observational selection was taken into account.

It was revealed in the first studies of the distribution of exoplanets over projective masses (obtained with the radial velocity method) that the exponent decreases in chronological order approximately from −1 to −1.5 (Marcy et al., 2005; Butler et al., 2006; Cumming et al., 2008; Howard et al., 2010). This is probably explained by the fact that with time the analysis has embraced more and more exoplanets and a more careful detuning from observational selection.

Ivanova et al. (2019) and Ananyeva et al. (2020a, 2020b) (thereafter, this series of papers is cited as Series IAA) estimated how the distribution of transit exoplanets over true masses depends on taking into consideration two factors of the observational selection—the factor of mass determination (to consider the planets with unknown masses) and the transit probability[1]. These authors found that the mass distribution, which is approximately described by the law $\partial N/\partial M \sim M^{-2}$ and does not change in dependence on the spectral class of host stars (the spectral classes F, G, K, and M[2] were considered). The power laws were derived for two cases: the data were analyzed generally, for all spectral classes of host stars together, and separately, for the classes F, G, K, and M. The corresponding histograms were approximated with the least-squares procedure, and the reliability of approximations was verified by the Kolmogorov–Smirnov test.

Ananyeva et al. (2020b) compared the mass distribution law obtained for exoplanets to the corresponding model distribution based on the planetary population synthesis (Mordasini, 2018). The later study yielded the mass distribution of exoplanets, which is different in three mass intervals. In two intervals—the masses are smaller than ~$0.1M_J$ (or $30M_E$) and larger than $5M_J$—it is described by the law $\partial N/\partial M \sim M^{-2}$, while the law $\partial N/\partial M \sim M^{-1}$ fits an intermediate interval of $(0.1-5)M_J$. When comparing the mass distributions of transit exoplanets, which were obtained from the NASA Exoplanet Archive (2019) and corrected for the observational selection effects (Series IAA), to the results by Mordasini (2018), it was noticed that the mass distributions exhibit the same trends. The discrepancy between the distributions in the intermediate mass range was explained by an insufficient number of long-period transit exoplanets discovered to date.

Here, we propose an alternative approach to the factor of mass determination considered in Series IAA. To account for ~75% of transit planets with unknown masses (their masses have not been measured yet), we consider the mass–radius relationship models[3] currently available for exoplanets (Bashi et al., 2017; Chen et al., 2017; Ning et al., 2018) and apply them to our analysis. We also examine the model, where the mass M is determined by the averaged observational data $M(R)$ obtained from the NASA Exoplanet Archive (2020) (hereafter, the Archive). To weaken the observational selection, the coefficient, which accounts for

[1] This is such a mutual position of the observer and the orbital plane of an exoplanet that the passage of an exoplanet across the disc of a host star, i.e., the transit, can be observed.

[2] For the spectral class M, the mass distribution was obtained from the analysis of data on the radial velocities (Tuomi et al., 2019).

[3] Hereafter, the mass–radius relationship means the functional relationship between the statistic mass and the statistic radius, which makes it possible to estimate one of these parameters from the other one for a statistically average case.

the transit probability, is calculated in the same way as that in the paper by Ananyeva et al. (2020b) (which follows Petigura et al. (2013)). However, we take into consideration all of known transit exoplanets from the Archive rather than only those with known masses. Moreover, in contrast to Series IAA, the distributions are considered as continuous, and the distribution law parameters are determined by the maximum likelihood estimation (Clauset et al., 2009).

Here, we examine in more detail the problem concerning probable observations of structurization of planets within planetary systems, which was preliminarily discussed by Ananyeva et al. (2020). Structurization is taken to mean that there is a statistical dependence of the mass of an exoplanet on the semimajor axis of its orbit in a planetary system. The structurization is observed in the Solar System: the planets of lower masses are closer to the star. To ascertain from the currently available data of observations whether the planetary systems are structured, it is necessary to compare the mass distributions of exoplanets determined before and after correcting for the transit-method selection.

The main purpose of this paper is to analyze the mass distribution law of transit exoplanets, for which the unmeasured masses of planets are determined from the mass–radius relationship. This allows us to verify the earlier estimates obtained in Series IAA.

The first six sections of the paper deal with the analysis of exoplanets detected by the space telescopes: Kepler Space Telescope and Transiting Exoplanet Survey Satellite (TESS); in the seventh section, the distributions for the other transit exoplanets, which were mainly detected by ground-based instruments and the masses of which are known, are considered. In the section "The distributions of transit exoplanets...," we briefly describe the results and the technique of accounting for two observational selection factors (Series IAA), one of which is used here, and the other one serves as a basis for the key idea of the present analysis—to use the mass–radius relationships. In the section "Parameters of exoplanets from the Archive," we systematize the data about exoplanets considered here, present the grounds for dividing exoplanets into two groups, and discuss some characteristic/specificity in the mass and radius distributions for these groups. In the section "The statistical mass–radius relationships...," the currently available mass–radius relationships are briefly described; and the dependences, which are used in the following to calculate the masses of exoplanets, are chosen. The distributions over masses (including the calculated masses), the power-law approximation, and their comparison to the estimates obtained in Series IAA are presented in the section "The mass distributions... comparison and analysis." Transit probability is taken into account in the next section, "Accounting for the transit probability...". The problem of structurization within planetary systems is also considered there. In the section "The lower boundary..." we present the results of verifying the hypothesis that the distributions obey the power law, when the minimal mass value in the considered interval is varied, and approximate two derived distributions by the broken power law, the parameters of which are compared to those obtained by Mordasini (2018).

THE DISTRIBUTIONS OF TRANSIT EXOPLANETS ACCOUNTING FOR THE OBSERVATIONAL SELECTION CAUSED BY THE MASS DETERMINATION AND THE TRANSIT PROBABILITY

In the recent papers of Series IAA, transit exoplanets known to the beginning of 2019 were divided into two groups depending on the instruments of observations: 2564 exoplanets detected by the *Kepler* space telescope (group 1) and 329 planets detected by the ground-based instruments and the COnvection ROtation and planetary Transits (CoRoT) space telescope (group 2). In total, the mass was known only for 539 out of 2893 planets; what is more, the mass was known for all planets of group 2, but only for 210 planets of group 1. The mass distributions were independently determined for the planets of each of the groups in the mass intervals $(0.02-13)M_J$ and $(0.68-13)M_J$, respectively.

To weaken the observational selection effect in the statistical analysis of group 1, the data on the radii of exoplanets were used. The considered mass and radius ranges were divided into equal, when expressed in logarithms, intervals ΔM and ΔR, respectively; and the histograms $N(M) = \partial N/\partial M$ and $N(R) = \partial N/\partial R$ were built. Further, in each of the radius intervals $\Delta R = i(R)$, the portion k of planets with known masses was determined:

$$k(R) = \frac{N_{\text{meas.mass}}(i(R))}{N_{\text{all}}(i(R))}, \quad (1)$$

where $N_{\text{meas.mass}}$ and N_{all} are the number of planets with known masses and the total number of planets in the interval *i,* respectively; $i(R)$ is the radius interval, within which the planet with the radius R falls.

In most cases, except the planets detected with the transit-timing variation (TTV) method, the mass of transit planets was subsequently determined by radial velocity measurements. To correct the exoplanet statistics based on the Archive, which is evidently distorted by the observational selection, the mass determination coefficient was introduced as a dependence $k(R)$ of the portion of planets with known masses on the planetary radius R. This distortion appears, because it is easier to determine the mass of large planets than that of small ones. The function $k(R)$ predominantly grows in a range of 0.02 to 1. For the overwhelming majority of small planets with radii of

~$0.1R_J$, the mass was not determined and $k \ll 1$; while for large planets with radii of $(1-13)R_J$, on the contrary, the mass was determined in many cases and $k \approx 1$. To each of the planets, the statistical weight $1/k(R)$ was assigned. Thus, in each of the radius intervals, the values of the mass were assigned to the exoplanets with unknown masses according to the distribution over the known masses of planets from this interval.

Moreover, following the paper by Petigura et al. (2013), the authors of Series IAA considered the observational selection factor, which accounts for the probability of observing a planet in transit, since only the planets whose orbital inclination does not differ much from 90° may be detected with the transit technique. The geometrical probability of observing a planet in transit p_{tr} (the probability that the observer's location allows the transit of a planet to be observed) is simplistically defined as a ratio of the star's radius r to the semimajor axis of the planetary orbit a: $p_{tr} \sim r/a$ (Winn et al., 2010). Since the transit probability is small, the observational selection results in an insufficient number of detected planets. The planets, which are not observed in transit, are taken into account hypothetically; and their masses are assumed to be the mass of a detected planet so that the correction leads to the increase of the statistical weight of each of the transit planets by k_1 times:

$$k_1 = \frac{a}{r}. \quad (2)$$

Thus, the authors of Series IAA analyzed the mass distribution of 393 planets and obtained the dependence $\partial N/\partial M \sim M^{-2}$, which was corrected relative to the observational selection by the coefficients k and k_1. For the statistical sampling of the *Kepler* planets (group 1), which includes a larger number of long-period planets than the sample of planets detected in ground-based and CoRoT observations (group 2), the characteristic trend in changing the number of planets in dependence on their mass was visually traced in the $(0.1-5)M_J$ interval. This resulted in growing the exponent α of the power law $\partial N/\partial M \sim M^{\alpha}$ to a value of -1 according to the theoretical model of the planetary population synthesis (Mordasini, 2018).

PARAMETERS OF EXOPLANETS FROM THE ARCHIVE

For this analysis, we use the data about exoplanets from the Archive (NASA Exoplanet Archive, 2020) as on July 2020[4], which contains 3169 confirmed exoplanets detected by the transit method. The radius was determined for all of the planets by light curves of the host stars during transit observations. In most cases, to determine the mass of a transit planet, it is necessary to perform additional spectroscopic measurements of the radial velocity of the host star. Since the accuracy of these measurements is low for small transit planets (with smaller masses), the mass was measured only for 858 exoplanets (27%). Moreover, for 22 planets among them, the mean relative deviation exceeds the mass value; and only the upper mass limit is known for 65 planets.

In this study, to determine the coefficient k_1, we use the semimajor axis of an orbit a and the radius of a star r in addition to the mass M and the radius R of an exoplanet. Since the Archive does not explicitly contain the value of a for 1478 exoplanets, we calculate the semimajor axes from the values of the orbital period P and the star's mass m. For this, we consider the known relationship, which describes the motion of a material point along the elliptic orbit in the central gravitational field with ignoring the planetary mass (Seager and Mallen-Ornelas, 2002):

$$a^3 = \frac{Gm}{4\pi^2} P^2, \quad (3)$$

where G is the gravitational constant.

The values of the above-listed parameters, as well as their limiting deviations, were taken from the extended table of the Archive[5], which contains the parameters for each of the exoplanets and their host stars according to the public sources. In this paper, we use the parameter values, which their smallest mean relative deviations correspond to.

In accordance with the approach by Series IAA, we divide all considered exoplanets into two groups. The updated Archive version contains the following objects that we take into account: 2773 exoplanets discovered with the *Kepler* and TESS space telescopes (group 1) and the other 396 planets discovered by 22 ground-based instruments and the space telescopes of the CoRoT and Hubble Space Telescope (HST) missions (group 2). For the whole collection of exoplanets and for groups 1 and 2, the mass distributions $N(M)$ covered by the Archive are shown in Fig. 1a. The corresponding mass–radius relationships $M(R)$ are presented in Fig. 1b.

Figure 1a shows that the mass distribution of all exoplanets (black solid line), not separated into groups 1 and 2, exhibits two clearly expressed maxima, the origins of which are mainly caused by the observational selection, i.e., the unequal coverage of exoplanets of groups 1 and 2 by the modern observational tools—the *Kepler* and TESS telescopes or ground-based surveys (as well as by the CoRoT and HST telescopes), respectively. The first maximum is in an interval of $[1.5, 4.4] \times 10^{-2}M_J$ (or $[4.8, 14]M_E$) and coincides with the maximum in the distribution of exoplanets of group 1; the

[4] https://github.com/yko-v/exoplanets2020

[5] https://exoplanetarchive.ipac.caltech.edu/docs/API_exomultpars_columns.html

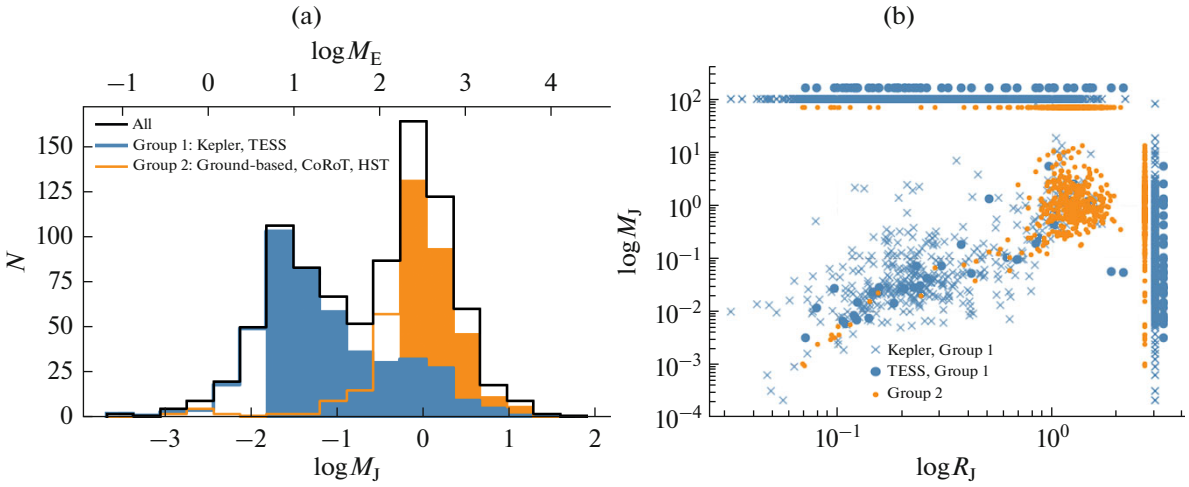

Fig. 1. (a) The distributions of the number of exoplanets over their masses $N(\log M)$ for all known transit exoplanets (black solid line) and two separate groups considered (blue and orange colors are for groups 1 and 2, respectively). The filled areas correspond to the mass intervals, within which the distribution law was determined. (b) The relationship between the mass and the radius of exoplanets expressed in the corresponding quantities for Jupiter, R_J and M_J, respectively. The transit exoplanets detected by different surveys are shown in the $\log(R)$–$\log(M)$ plane (the color cases correspond to those in panel (a)). On the top and on the right, there are projections of the distribution onto the corresponding axes.

second maximum is in an interval of $[0.54, 1.57]M_J$ and corresponds to the maximum for exoplanets of group 2. In particular, this suggests that the so-called desert of sub-Saturns (Ida and Lin, 2004; Mordasini et al., 2009)—the minimum in a mass interval of $[0.1, 0.3]M_J$ (or $[-1, -0.5]\log M_J$)—is most likely caused by the observational selection.

The mass intervals corresponding to complete observations of the planets in groups 1 and 2 are filled with colors in Fig. 1a. The observations with the *Kepler*[6] and TESS telescopes become incomplete starting from the exoplanets with masses smaller than $\approx 0.02M_J$. The ground-based surveys of transit exoplanets (and the CoRoT and HST observations) become incomplete starting from the exoplanets with masses smaller than $\approx 0.68M_J$ (Series IAA). In the mass intervals, within which the observations are complete, the mass dependences of the number of exoplanets $N(M)$ clearly follow a certain power function, the characteristics of which were later specified more exactly by analyzing subtler effects of the observational selection.

Figure 1b demonstrates that the space telescopes (*Kepler* and TESS) and ground-based instruments (including the CoRoT and HST) cover different values of the transit depth in dependence on the planetary radius. For example, the exoplanets of group 2 are mainly concentrated (94%) in a region within $[0.63, 13]M_J$ and $[0.1, 2]R_J$, while the wider mass and radius ranges correspond to the exoplanets of group 1.

In Fig. 2, the exoplanets of groups 1 and 2 are shown in the $\log M$–$\log R$ plane. The analysis of the distribution of exoplanets presented in the mass–radius plane suggests that the following features should be noted.

(1) There are two intervals, where the mass–radius interrelation behaves differently: for small solid planets with masses less than $0.1M_J$ and radii less than $0.1R_J$ ($-0.44\log R_J$) (Bashi et al., 2017), this relationship is more evident as compared to that for giant planets.

(2) The significant scattering in the planetary masses (by several orders of magnitude), corresponding to a small radius interval, is primarily caused by the differences in the composition for solid planets or the thermal flux from a host star for gaseous planets (Weiss et al., 2013).

(3) Since in a range of $[-4, 1.5]\log M$ the radii are estimated more accurately than the masses, it is more common to determine the relationship $R(M)$ than the inverse one $M(R)$.

(4) The measurement errors are nonhomogeneous, since the parameters of large planets are measured more reliably and their relative deviations are smaller than the corresponding quantities for planets with small masses.

[6] The *Kepler* observations become incomplete starting from the planets with a period exceeding 400 days and a radius smaller than $0.178R_J$ ($2R_E$) (Petigura et al., 2013). Most exoplanets with this radius (taken from the Archive or determined from the mass–radius relationship (see the section "The statistical mass–radius relationships...")) correspond to the masses in a range of $[0.01, 0.03]M_J$.

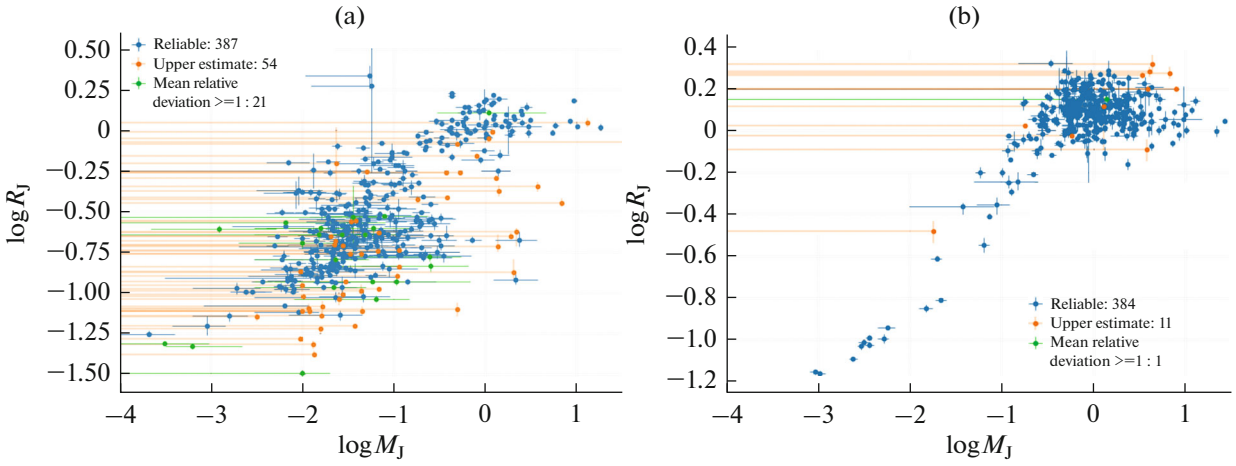

Fig. 2. Exoplanets of group 1 (a) and group 2 (b) are shown in the log$M$–log$R$ plane by dots with bars corresponding to the errors in the mass and radius measurements. The orange, green, and blue symbols indicate the exoplanets, for which only the upper mass estimates are specified, the exoplanets, the mean mass deviation in which exceeds the mass of a planet (see the section "Calculation technique" and the subsection "Calculations of the masses of exoplanets"), and the other exoplanets (with more reliable masses), respectively. The lines going out of the plot mean that the lower estimate of the mass is close or equal to zero (sometimes negative (see, Marcy et al., 2014)).

THE STATISTICAL MASS–RADIUS RELATIONSHIPS CHOSEN FOR ESTIMATING THE MASSES OF EXOPLANETS BY THEIR RADII

As has been already noted above (see the section "Parameters of exoplanets..."), the radius is known for all transit exoplanets, while the mass is known only for a small portion of them (858 out of 3169). To include as many as possible exoplanets into the mass distribution, we estimate the statistical values of unknown masses of exoplanets (2311) with the use of the mass–radius relationships (according to the models applied to the Archive observational data). The mass–radius relationships built on the basis of physical models (e.g., Zeng et al., 2016; Otegi et al., 2019) remain beyond the present analysis.

The relationship between the mass and the radius of a planet is determined by the density of its material. The density, in turn, depends on the composition of a planet and, for gaseous planets, on the environmental conditions (parameters of the orbit and the star). The mean density or the approximate composition of a planet cannot be determined directly from photometric measurements. Consequently, strictly speaking, the mass of an exoplanet cannot be determined from its radius (or vice versa). However, there is a correlation between the mass and the radius of planets, which makes it possible to estimate, at a first approximation, the statistical (probable or mean) mass of a planet by its radius from the observational data.

The currently available mass–radius relationships of exoplanets, which were determined from observational data (see a dozen papers cited and discussed by Chen and Kipping (2017) and Ning et al. (2018)), may be divided into parametric and nonparametric ones according to the models they use.

In the parametric models (see Table 2), the mass–radius relationship is approximated by a broken power-law $M \sim R^{\alpha}$ or $R \sim M^{\alpha}$. For this, two approaches are used: (i) the deterministic one, which defines a unique relationship between the mass and the radius (Weiss et al., 2013; Bashi et al., 2017), and (ii) the probabilistic one, according to which some mass distribution corresponds to a specified value of the radius and a target value of the mass is retrieved from this distribution (Wolfgang et al., 2016; Chen and Kipping, 2017).

In the nonparametric model described by Ning et al. (2018), the data of observations are approximated by the deterministic function from the space of the Bernstein basis polynomials. In addition to these models, the parametric mass–radius relationship derived by averaging the Archive data is considered (see the section "Calculation technique" and the subsection "Derivation of the *Averaged* mass–radius relationship").

Weiss et al. (2013) determined the mass–radius relationship $R(M)$ for 138 transit exoplanets (as confirmed to September 2012) by the least-squares method. Bashi et al. (2017) considered the parameters of 274 transit exoplanets (as confirmed to March 2016) also to find the mass–radius relationship $R(M)$. The latter research group used a more general approach—the least-squares method that accounts for the measurement errors in both parameters; and the inhomogeneity in the measurement errors and the mass distribution of planets was also taken into consideration. In these two papers (nos. 1 and 2 in Table 2) the mass–radius relationships are almost the same for the inter-

Table 2. Parametric models of the mass–radius relationship according to some papers

No.	Exponent α in $R(M) \sim M^{\alpha}$	Intervals by		Model name in the paper (the paper)
		mass, M_J	radius, R_J	
Deterministic				
1	0.53	<0.47	–	(Weiss et al., 2013)
	–0.04	≥0.47	–	
2	0.55	<0.39	<1.08	*Bashi17* (Bashi et al., 2017)
	0.01	[0.39, 13]	≥1.08	
Probabilistic*				
3	0.56**	–	≤0.71	(Wolfgang et al., 2016)
4	0.28	$<6 \times 10^{-3}$	–	*Chen17* (Chen and Kipping, 2017)
	0.59	$[6 \times 10^{-3}, 0.41]$	–	
	–0.04	[0.41, 13]	–	

*For probabilistic models, the mathematical expectations of α are presented; for model 5, the mathematical expectations of the boundary values of the mass intervals are shown.
**The exponent was obtained from the $M(R)$ dependence.

val of low-mass planets: $R \sim M^{0.53}$ and $R \sim M^{0.55}$, respectively. In the interval of massive planets ($M > 0.47M_J$ and $M > 0.39M_J$, respectively), which are mostly gaseous giants, the mass–radius relationship is weak and implicit ($R \sim M^{-0.04}$ and $R \sim M^{0.01}$, respectively), since the radius of planets substantially depends on the average thermal flux coming from a star to the planetary orbit (Weiss et al., 2013). Bashi et al. (2017) considered the transition point between these intervals as a free parameter, while Weiss et al. (2013) specified this point a priori, on the basis of observational data and physical considerations. Since the mass–radius relationship corresponding to model 2 (*Bashi17*) is based on more recent data and is more elaborate than model 1, we consider here only the *Bashi17* model.

In the probabilistic models (nos. 3 and 4 in Table 2), Bayesian hierarchical modeling was used (Wolfgang et al., 2016; Chen and Kipping, 2017). Model 3 is restricted by the largest radius value $0.71R_J$, while the radii of planets considered here reach $1.6R_J$. Consequently, model 3 will not be used here, since its range is insufficient for the present analysis.

Chen and Kipping (2017) modeled the mass–radius relationship in an interval of ~$[0.01, 100]R_J$, which is more than sufficient for the present analysis. In this model (called *Chen17*, no. 4 in Table 2), the mass range is divided into four intervals: $[10^{-3}, 6 \times 10^{-3}]M_J$ for Earth-like planets and super-Earths, $[6 \times 10^{-3}, 0.41]M_J$ for Neptunes, $[0.41, 83.8]M_J$ for Jupiters and brown dwarfs, and larger than $83.8M_J$ for stars (not shown in Table 2). Analogously to the models considered above, Chen and Kipping (2017) obtained the mass–radius relationship $R(M)$, but in their model the radius is a random function. Any specified mass value is related to a normally distributed radius value

$$R \sim N\left(\mu = CM^{\alpha}, \sigma_R\right), \quad (4)$$

where μ is the mathematical expectation in a form of the power function, σ_R is the standard deviation, and C is the constant. The mathematical expectation for the exponent α and the boundary values of the mass intervals is specified in Table 2.

Figure 3 presents the mass–radius relationships, according to which the masses of exoplanets of group 1 with unknown masses were calculated: the averaged model (hereafter, designated as *Averaged*), the *Bashi17* (Bashi et al., 2017) and *Chen17* (Chen and Kipping, 2017) models, and the *Ning18* nonparametric model (Ning et al., 2018).

THE MASS DISTRIBUTIONS OF TRANSIT EXOPLANETS OBTAINED WITH THE MASS–RADIUS MODELS: COMPARISON AND ANALYSIS

The distribution density function of the power law is determined[7] as

$$p(m) = Cm^{\alpha}, \quad C = \frac{(1+\alpha)}{m_{\min}^{1+\alpha}}, \quad \alpha < 0, \quad (5)$$

where C is the normalizing constant and α and $m_{\min}$ are the exponent and the minimal mass value in the considered interval of the distribution, respectively (Clauset et al., 2009).

[7] Or $p(m) = C_1 m^{-\alpha}$, $\alpha > 0$. To compare the exponent to the results of the other papers, the version specified in Eq. (5) has been accepted.

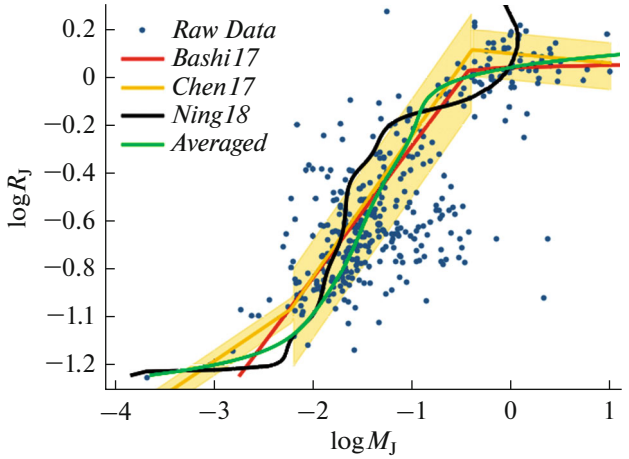

Fig. 3. The mass–radius relationships, namely, the *Bashi17*, *Chen17*, *Ning18*, and *Averaged* models, according to which the statistical masses of exoplanets (group 1) were calculated. They are shown in the log$M$–log$R$ plane together with the *Raw Data* (dots) from the Archive covering 386 exoplanets with measured masses $M \le 13 M_J$, which were detected by the *Kepler* and TESS space telescopes (group 1). For the *Chen17* model, the line shows the dependence for the mathematical expectation of the exponent in the filled area of $3\sigma_R$. The line designations are in the legend.

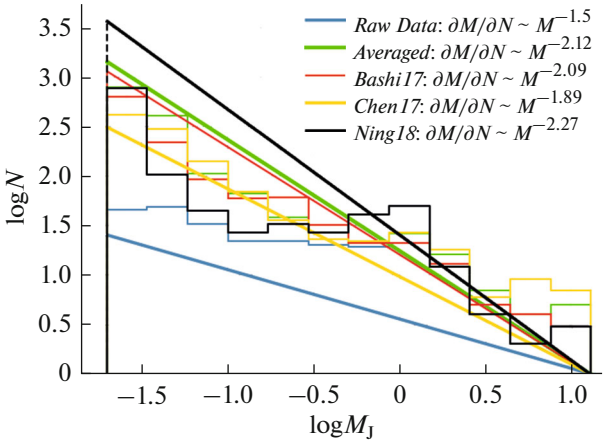

Fig. 4. Histograms[9] of the mass distributions of exoplanets of group 1 (the *Raw Data* sample contains the exoplanets discovered by the *Kepler* and TESS space telescopes according to the Archive) and the exoplanets from the samples obtained by adding the masses of exoplanets calculated from the mass–radius relationships to the *Raw Data* sample. These relationships are *Averaged*, *Bashi17*, *Chen17*, and *Ning18*. The mass interval is $[0.02, 13]M_J$. The types and colors of lines are specified in the legend. The slanting lines show the corresponding approximations by a power law with different power indices.

To estimate the parameter α and the corresponding standard deviation σ for the sampling $M = \{m_i\}$, $i = 1...n$ with a specified value of m_{min}, the maximum likelihood estimation (Clauset et al., 2009) is used:

$$\hat{\alpha} = 1 + n\left(\sum_{m_{min}}^{n} \ln\frac{m_i}{m_{min}}\right)^{-1}, \quad \hat{\sigma} = -\frac{\hat{\alpha}+1}{\sqrt{n}}.$$

For 2311 exoplanets with unknown masses (group 1, detected by the *Kepler* and TESS space telescopes) the masses were calculated according to the above-described $M(R)$ models 2 and 4 (see Table 2), the *Averaged* model (see the section "Calculation technique" and the subsection "Derivation of the *Averaged*..."), and the model by Ning et al. (2018). (The results of simulations are reported in the section "Calculation technique" and the subsection "Calculations of the masses...".) We analyzed the distributions of the following five samples of exoplanets of group 1: the *Raw Data* sample from the Archive and four samples of exoplanets, each of which contains the planets from the *Raw Data* sample and the remaining planets of group 1 with the masses calculated according to the *Bashi17* and *Chen17* models (nos. 2 and 4 in Table 2), the *Averaged* model, and the *Ning18* model. The planets with radii less than $0.178 R_J$ or periods longer than 400 days (1188 exoplanets) were removed from the samples, since the *Kepler* data are incomplete. Due to this, 1561 out of 2773 exoplanets of group 1 remain for consideration.

For each of these samples, the exponent α was estimated in a mass interval of $[0.02, 13]M_J$ (see Table 3 and Fig. 4). In these samples, the number of planets in the considered interval turned out to be different[8], which is caused by using the different mass–radius relationships.

To calculate the mass of exoplanets, the authors of Series IAA did not use the radius values, but corrected the mass distribution of planets from the Archive by the mass determination coefficient (see the section "The distributions of transit exoplanets..."). As a result, the corrected mass distribution of exoplanets corresponded to the power law with an exponent $\alpha_{19} = -1.90 \pm 0.06$, which coincides with the result of the *Chen17* sample and close to that of the *Bashi17* sample. The smallest residual corresponds to the *Bashi17* sample (0.1), while it is less than 12% for the *Averaged* and *Ning18* samples.

When verifying the considered samples for homogeneity in pairs (*Averaged* to *Bashi17*, *Averaged* to *Chen17*, etc.), the hypothesis that these samples belong to a single general collection is rejected, because different mass–radius relationships yield sta-

[8] In the samples with calculated masses, the number of mass values taken from the Archive does not coincide with that in the *Raw Data* sample, since the latter includes only sufficiently reliable values (see the section "The statistical mass–radius relationships...").

[9] In Fig. 4, the histograms are shown only for visualization of the considered distributions, the appearance of which depends on dividing into intervals; the power laws were obtained by the maximum likelihood estimation (6) rather than fitting.

Table 3. The exponent α of the law $\partial N/\partial M \sim M^{\alpha}$ for the considered samples in a mass interval of $(0.02-13)M_J$ and the residual as compared to a value of $\alpha_{19} = -1.90 \pm 0.06$ obtained in Series IAA

Sample of planets, model	$\hat{\alpha} \pm \hat{\sigma}$	$\min(\alpha - \alpha_{19})$	Number of planets in an interval of $[0.02, 13]M_J$	
			total	planets with the masses calculated from the mass–radius relationship
Raw Data	-1.50 ± 0.03	0.32	253	0
Averaged	-2.12 ± 0.03	0.13	1515	1256
Bashi17	-2.09 ± 0.03	0.10	1174	918
Chen17	-1.89 ± 0.03	–	1070	810
Ning18	-2.27 ± 0.04	0.27	1115	861

tistically different mass distributions of exoplanets. The consistency of the distributions with the power law is verified below in the section "The lower boundary...".

ACCOUNTING FOR THE TRANSIT PROBABILITY AS AN OBSERVATIONAL SELECTION FACTOR

As has been already discussed above (see the section "The distributions of transit exoplanets..."), the observational selection causes an insufficient number of detected exoplanets, since the probability of observing a planet in transit p_{tr} is low. The non-transiting planets (i.e., they do not induce the decrease in the brightness of a star as is seen from the Earth) cannot be detected with the transit method, though they constitute the majority of planets. To take them into account, the statistical weight of each transit planet can be multiplied by the factor $k_1 > 1$ (see Eq. (2)) (Petigura et al., 2013), i.e., it is assumed that, to each of the detected exoplanets, we may put into correspondence a set of planets with the same mass in the amount inversely proportional to p_{tr} (Series IAA).

In Fig. 5, we present the mass distributions of exoplanets obtained above from the $M(R)$ models (see the section "The statistical mass–radius relationships...") for the planets of group 1, which were discussed in the section "The mass distributions ... comparison and analysis". Two cases are considered: the observational selection is ignored or taken into account by the coefficient k_1.

For each of the samples, the Kolmogorov–Smirnov homogeneity test was performed before and after accounting for the selection (the coefficient k_1). From this verification, we may conclude that, for all of the considered samples except the *Chen17* one, the hypothesis on the same distribution law, describing the sample before and after including the selection into the procedure, is not rejected. For the *Chen17* sample, after accounting for the selection, the exponent α changes from -1.89 ± 0.03 (without k_1) to -1.94 ± 0.03, which coincides with $\alpha = -1.99 \pm 0.08$ obtained in Series IAA after correcting the distribution for the mass determination and the transit probability effects. Since the distributions for the considered samples (except *Chen17*) do not differ statistically, their power indices will also coincide (Table 3). Consequently, after accounting for the observational selection, the exponent for the *Bashi17* sample ($\alpha = -2.09 \pm 0.03$) also coincides with $\alpha = -1.99 \pm 0.08$, while the residual for the *Averaged* and *Ning18* samples decreases to 0.01 and 0.16, respectively, as compared to the values obtained with the selection ignored.

Based on the homogeneity test described above, we may conclude that the accounting for the observational selection of the transit method does not introduce statistically important changes into the mass distribution of exoplanets considered here (with a semimajor axis $a \le 1$ AU). The difference is observed only in a range of large masses $M > 3.16M_J$ ($0.5\log M_J$), where the number of planets is small (see Fig. 5). The correction did not induce the decrease in the number of planets in this range. For $M < 0.1M_J$, the distributions did not change in appearance, except for the *Ning18* sample.

When correcting the statistical distribution for the transit method selection (by the coefficient k_1, which is inversely proportional to the probability of observing a planet in transit), this distribution is enlarged by the planets with the masses that populate a small interval of the joint distribution $\partial^2 N/(\partial r \cdot \partial a)$ over the radius r of a star and the semimajor axis a of the orbit of an exoplanet according to the mass distribution of exoplanets observed in this interval. (For a small interval (r, a), the number of added exoplanets of a specified mass is proportional to the number of planets with this mass in this interval.) If the probability of detecting an exoplanet were dependent on its mass, the correction would substantially enhance the number of exoplanets with the masses that may be less likely to be found as compared to those whose detection probability is high. This would significantly change the mass distribution after correction. The detection probability $p_{tr} \sim r/a$ primarily depends on the semimajor axis, while the radius of a star varies less. Because of this, since no sta-

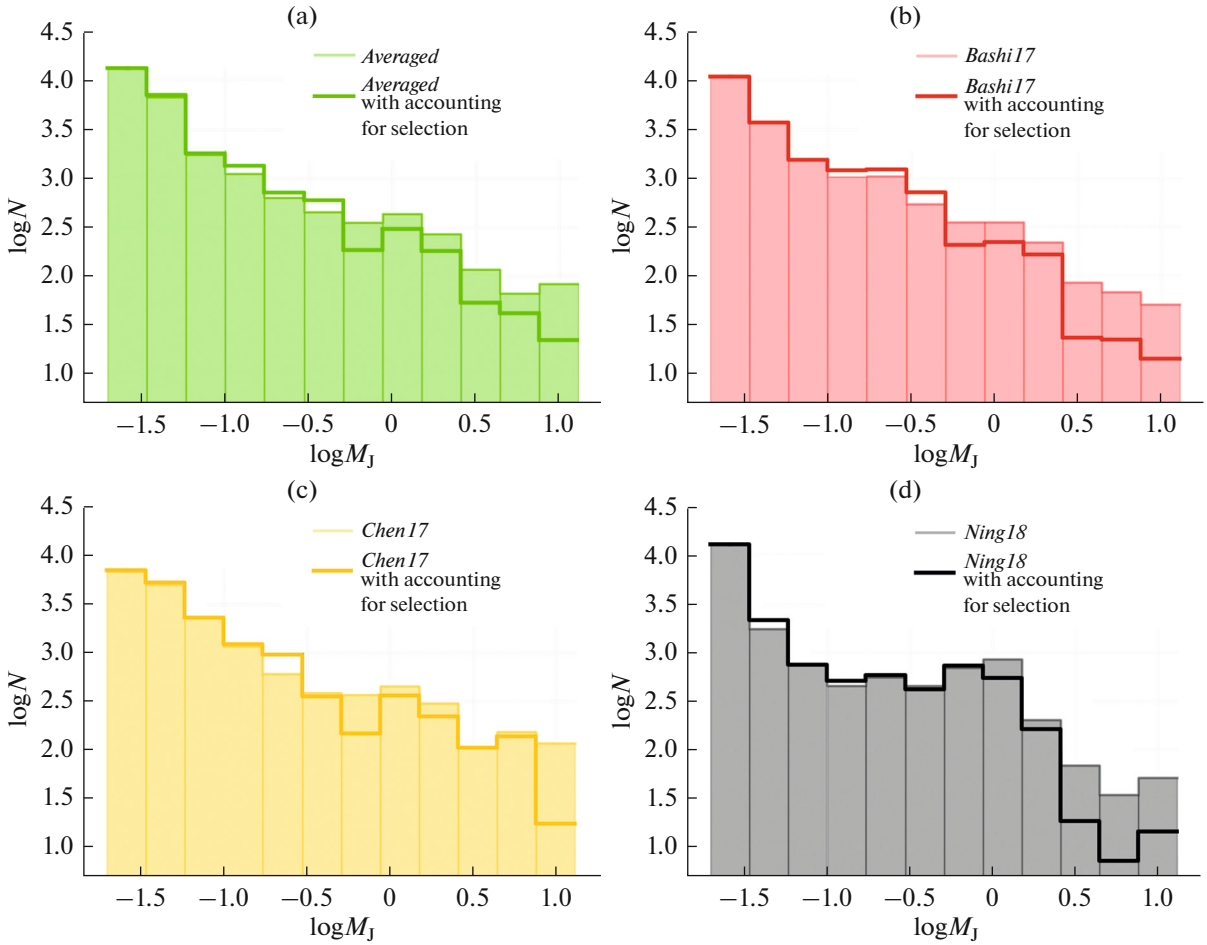

Fig. 5. Histograms of the mass distributions of exoplanets of group 1 (the exoplanets with a mass $M \le 13 M_J$ detected by the *Kepler* and TESS space telescopes), to which the exoplanets with statistical masses calculated from the mass–radius relationships are added. The panels correspond to the *Averaged* (a), *Bashi17* (b), *Chen17* (c), and *Ning18* (d) relationships. In each of the panels, two cases are presented: the transit probability is ignored (unfilled areas) or taken into account (filled areas).

tistical relationship between the mass and the semimajor axis of an exoplanet is observed, the dependence of the mass distribution of planets on their closeness to the star would indirectly indicate the structurization within planetary systems. The relative decrease in the number of massive planets after correction (which is barely discernible in Fig. 5) could be indicative of a larger probability of their detection and, consequently, the predominantly near-star location as compared to that of small-mass planets. However, the test for the homogeneity of distributions showed that, among the samples considered, only the *Chen17* model exhibits this difference. Thus, this test does not allow us to state any structurization for exoplanets, the semimajor axes of which are smaller than ≈1 AU. Due to limitations of the transit method, only few transit exoplanets were detected at larger distances from the stars, which is the reason for the statistical incompleteness in the detection of long-period planets.

Moreover, the probability of detecting an exoplanet with a specified semimajor orbital axis near large-sized stars is higher than that for smaller stars. In addition, since the stars of larger radii are less widespread, the probability of detecting exoplanets near these stars should be higher than that in the case of stars homogeneously distributed over radii. In other words, if the dependence of the occurrence of stars on their sizes is ignored, this circumstance cannot overestimate the probability of detecting exoplanets near large stars and underestimate it for small stars. In the sample, the exoplanets, which were detected with the *Kepler* and TESS telescopes, mainly orbit sunlike stars (for ≈90% of the considered stars, the radii are in a range of $r \in [0.4, 1.8] r_{sun}$). Consequently, the influence of this factor on the mass distribution is most likely insignificant, which allows us to ignore it in the present analysis.

THE LOWER BOUNDARY IN THE MASS DISTRIBUTIONS OF EXOPLANETS AND THE COMPARISON TO THE DISTRIBUTION BUILT BY THE PLANETARY POPULATION SYNTHESIS

There are two parameters in the mass distribution of exoplanets approximated by the power law: the power exponent α and the lower mass boundary $m_{\min}$ (see the section "The mass distributions ... comparison and analysis"). In the previous section "Accounting for the transit probability...", following the analysis performed for $m_{\min} = 0.02M_J$ (Series IAA) and keeping in mind the range of complete observations with the *Kepler* telescope (see the section "Parameters of exoplanets from the Archive"), we assumed $m_{\min} = 0.02M_J$. However, the obtained distributions may be approximated by both the power law with $m_{\min} > 0.02M_J$ and the broken power-law (Mordasini, 2018).

In Fig. 6, for each of the modeling samples considered (see the section "Accounting for the transit probability..."), we present the dependences $\alpha(m_{\min})$ for the power law exponent at the lower boundary of the considered mass interval. With the Kolmogorov–Smirnov criterion, the optimal values of the distribution parameters α and $m_{\min}$ were determined and the hypothesis on compliance of the samples' distributions with the power law for all values of $m_{\min}$ was verified (for the model results, see the section "Calculation technique" and the subsection "The power law compliance test"). The values of $\alpha(m_{\min})$, under which this hypothesis is not rejected, are shown by thick lines in the diagram.

In an interval of $[0.03, 0.15]M_J$ (or $[-1.5, -0.8]\log M_J$), the functions for the *Raw Data* and *Ning18* samples ($\alpha \approx -1.55$), as well as the functions for the *Averaged*, *Bashi17*, and *Chen17* samples (α varies from -2.0 to -1.7), exhibit a similar behavior. Further, when the lower boundary moves to $\sim 0.3M_J$ ($-0.5\log M_J$), $\alpha \in [-1.7, -2]$ for all of the samples. In an interval of $[0.3, 1.17]M_J$ (or $[-0.8, 0.69]\log M_J$), the range of α for the *Ning18* sample is $\alpha \in [-2, -3.25]$, which significantly differs from that for the other samples, $\alpha \in [-1.7, -2.5]$. For $m_{\min} > 1.17$, there is no difference between the *Raw Data* and *Ning18* samples, and the behavior of all functions is the same. When $m_{\min}$ grows, the number of planets in a sample substantially decreases. For $m_{\min} = 0.17M_J$ (or $-0.75\log M_J$), the number of planets varies in dependence on the sample from 101 to 169, among which 101 planets are from the Archive; while their number decreases to 31–57 in an interval of $[1, 13]M_J$.

In an interval of $[0.046, 0.061]M_J$ (or $[-1.34, -1.21]\log M_J$) the *Averaged* and *Bashi17* distributions are statistically indistinguishable, while the same can be said about the *Bashi17* and *Chen17* distributions in a range of $[0.093, 0.14]M_J$ (or $[-1.03, -0.85]\log M_J$). In an interval of $[0.058, 0.08]M_J$ (or $[-1.24, -1.10]\log M_J$), all of the *Averaged*, *Bashi17*, and *Chen17* samples are distributed according to the power law (except two small intervals for the *Bashi17* sample (see the section "Calculation technique" and the subsection "The power law compliance test")), while $\alpha \in [-1.88, -1.72]$.

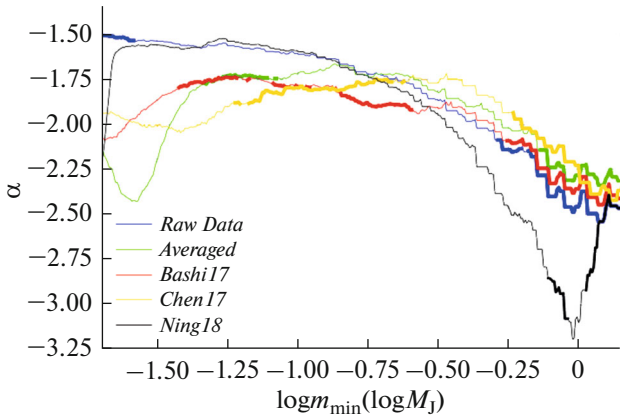

Fig. 6. The power law exponent in dependence on the lower boundary of the mass interval $\alpha(m_{\min})$ considered for the exoplanets from the sample of group 1 according to the Archive (the *Raw Data* contain exoplanets discovered by the *Kepler* and TESS space telescopes) and the exoplanets from the samples obtained by adding the masses of exoplanets calculated from the mass–radius relationships to the *Raw Data* sample. These relationships are *Averaged*, *Bashi17*, *Chen17*, and *Ning18*. The values of α, under which the distribution agrees with the power law according to the Kolmogorov–Smirnov test, are shown by thick lines.

The optimal parameters (i.e., those providing the agreement with the power law) obtained for the distributions of the considered samples (see the section "Calculation technique", the subsection "The power law compliance test", and Fig. 7) are presented in Table 4.

The result that the power law is better obeyed by the *Averaged, Bashi17,* and *Chen17* samples than the *Raw Data* and *Ning18* ones is also confirmed by the appearance of the distributions in Fig. 7. The *Raw Data* and *Ning18* distributions exhibit a substantially different behavior in an interval of approximately $[0.1, 0.56]M_J$ (or $[-1, -0.22]\log M_J$), due to which the hypothesis that these distributions are in compliance with the power law (Fig. 6) is true only in a range of massive planets (for the *Raw Data* sample, there is also a small interval of $m_{\min} \in [0.02, 0.03]M_J$). This form of the distribution law—a broken power-law with three regimes—was obtained by Mordasini (2018). Moreover, according to the Kolmogorov–Smirnov test, none of the distributions obeys the power law if $m_{\min} \in [0.31, 0.56]M_J$, which suggests in general that it is necessary to search for a more sophisticated distribution law, the exponent of which may differ on two or more intervals. According to the distribution obtained with the planetary population synthesis method (Mordas-

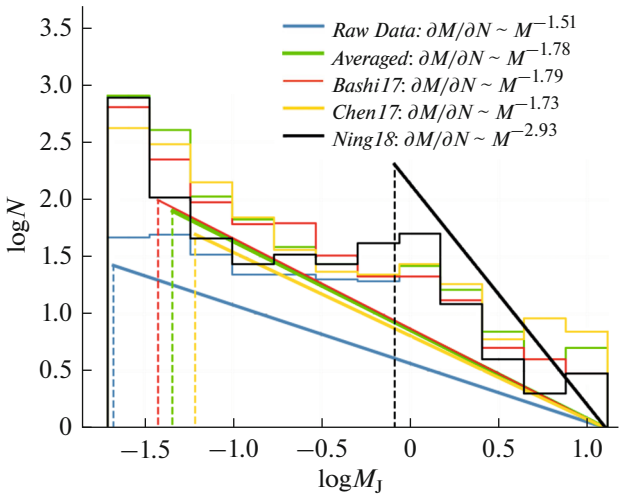

Fig. 7. Histograms of the mass distributions of the exoplanets of group 1 according to the Archive (the *Raw Data* sample contains the exoplanets discovered by the *Kepler* and TESS space telescopes) and the exoplanets from the samples obtained by adding the masses of exoplanets calculated from the mass–radius relationships to the *Raw Data* sample. These relationships are *Averaged*, *Bashi17*, *Chen17*, and *Ning18*. The designations of lines are in the legend. The vertical dashed lines indicate the minimal values m_{min} (specified in the inset), which correspond to the optimal value of α in the power law $\partial N/\partial M \sim M^{\alpha}$. The latter is shown by slanting lines on each of the intervals.

ini, 2018), we verified the approximation of the mass distribution of exoplanets by three power laws on three intervals.

When expressed mathematically, this law is defined by six parameters: the smallest value in the considered mass interval m_{min}; two transient points M_{T1} and M_{T2} specifying the boundaries of the intermediate interval; and the power indices α_1, α_2, and α_3 of the distribution law $\partial N/\partial M \sim M^{\alpha}$ on each of three intervals—[m_{min}, M_{T1}], [M_{T1}, M_{T2}], and [M_{T2}, 13M_J], respectively. With the maximum likelihood estimation (Eq. (6)) and by varying the parameters α_1, α_2, α_3, m_{min}, M_{T1}, and M_{T2}, we determined the optimal values of these parameters for the *Raw Data* and *Ning18* samples (see Table 5 and Fig. 8). These values provide the best agreement of the distributions with the power law (which is confirmed by the procedure described in the section "Calculation technique", the subsection "The power law compliance test").

Analogously to the result of Series IAA, we see that the planets with masses in an interval of $M > M_J$ are lacking in the both samples. The transient point M_{T1} of the *Ning18* distribution is closer than that of the *Raw Data* to M_{T1} in the distribution by Mordasini (2018) (*Mordasini18*), while the situation is inverse for the transient point M_{T2}. The power indices α_1 and α_2 (in the first and second intervals, respectively) coincide rather well with the compared laws for the *Ning18* and *Raw Data* distributions, respectively. In the third interval, the *Raw Data* distribution agrees with the *Mordasini18* one better than the *Ning18* distribution (the difference in the exponent is smaller). Though Mordasini (2018) obtained $\alpha_2 = -1$ in the second interval, this distribution, analogous to the *Ning18* one, exhibits the decrease (expressed in logarithm) in the distribution density with increasing mass. What is more, the difference (total and maximal) between the *Ning18* and *Mordasini18* discrete distribution functions is smaller than the analogous difference between the *Raw Data* and *Mordasini18* distributions. Thus, we may conclude that the *Ning18* distribution better agrees with the results by Mordasini (2018).

THE MASS DISTRIBUTION OF EXOPLANETS DETECTED WITH THE GROUND-BASED INSTRUMENTS AND THE COROT AND HUBBLE SPACE TELESCOPES

Since the masses of all 396 exoplanets of group 2 (detected with the ground-based instruments and the space telescopes of the CoRoT and HST missions) are known from the Archive data, no radius–mass relationships were applied to them. We consider 384 exoplanets of group 2, the masses of which are rather reliably known (see the section "Parameters of exoplanets from the Archive"); among them, 243 planets are in a considered interval of [0.68, 13]M_J. In the same way as for the previous samples, we determine the parameters of the power law that corresponds to the distribution of exoplanets of this group (group 2 in Fig. 9) for two cases: the observational selection is ignored and taken into account.

In Series IAA the distribution for the exoplanets sampled from the NASA Exoplanet Archive (2019)

Table 4. Optimal parameters of the distributions for the samples of exoplanets

Sample of planets, model	Exponent α in $\partial N/\partial M \sim M^{\alpha}$	Minimal value m_{min}, M_J	Number of planets in an interval of [m_{min}, 13]M_J	
			total	planets with calculated masses
Raw Data	-1.51 ± 0.03	0.021	248	0
Averaged	-1.78 ± 0.04	0.046	400	215
Bashi17	-1.79 ± 0.04	0.038	459	259
Chen17	-1.73 ± 0.04	0.062	333	176
Ning18	-2.93 ± 0.22	0.82	76	30

Table 5. The parameters of the broken power-law distribution for the *Raw Data* and *Ning18* samples and the parameters obtained by Mordasini (2018)

Sample name	Lower boundary m_{min}, M_J	Transient points, M_J		Exponent		
		M_{T1}	M_{T2}	α_1	α_2	α_3
Mordasini, 2018	–	0.094	5	–	–1	–2
Raw Data	0.025	0.28	1.35	–1.47	–0.97	–2.88
Ning18	0.025	0.16	1.04	–1.99	–0.62	–2.88

corresponded to the power law with an exponent of $\alpha = -2.12 \pm 0.12$ (called *Ground last* in Fig. 9) for $m_{min} = 0.68 M_J$, which is consistent with our result: $\alpha = -2.18 \pm 0.08$ for $m_{min} = 0.68 M_J$ or $\alpha = -2.22 \pm 0.08$ for $m_{min} = 0.71 M_J$ (the optimal case). The accounting for the observational selection of the transit method (see the section "Accounting for the transit probability...") does not change the analyzed distribution significantly, which follows from the Kolmogorov–Smirnov homogeneity test of the distributions ($\alpha = -2.21 \pm 0.04$ for $m_{min} = 0.68 M_J$ and $\alpha = -2.25 \pm 0.04$ for $m_{min} = 0.71 M_J$). This also agrees with the result of Series IAA for exoplanets of this group after accounting for the observational selection ($\alpha = -2.17 \pm 0.12$). The hypotheses on compliance of the values of m_{min} with the power law are accepted on the basis of the corresponding test (see the section "Calculation technique", the subsection "The power law compliance test").

CALCULATION TECHNIQUE

Derivation of the Averaged Mass–Radius Relationship

The *Averaged* relationship was obtained in the following way.

(1) The range of values of the radius is divided into n equal, when expressed in logarithms, intervals.

(2) In each of the intervals, the mean logarithmic values for the mass and the radius of planets are determined:

$$\log(M_{av\_i}) = \frac{1}{n_i}\sum_{k=1}^{n_i}\log(M_k),$$
$$\log(R_{av\_i}) = \frac{1}{n_i}\sum_{k=1}^{n_i}\log(R_k), \quad (6)$$

where R and M are the radius and the mass of planets with $R \in dR_i$.

(3) The obtained averaged values are approximated by the k-order polynomial

$$F_k(a, R) = \sum_{i=0}^{k} a_i (\log R_{av})^{k-i}, \quad (7)$$

where the coefficients a_i (see the result of the calculations in Table 6) are determined by the least-squares method.

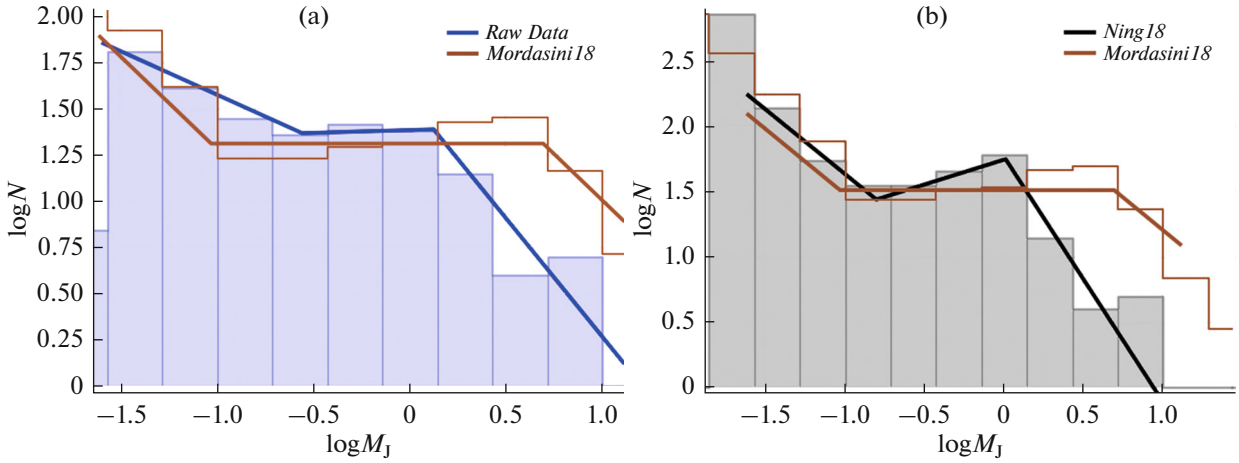

Fig. 8. Histograms of the mass distributions of exoplanets of group 1 and the corresponding broken power-laws of the *Mordasini18* distribution. (a) The *Raw Data* sample of exoplanets detected by the *Kepler* and TESS space telescopes according to the Archive data. (b) The exoplanets obtained by adding the planetary masses calculated from the mass–radius *Ning18* relationship to the *Raw Data* sample.

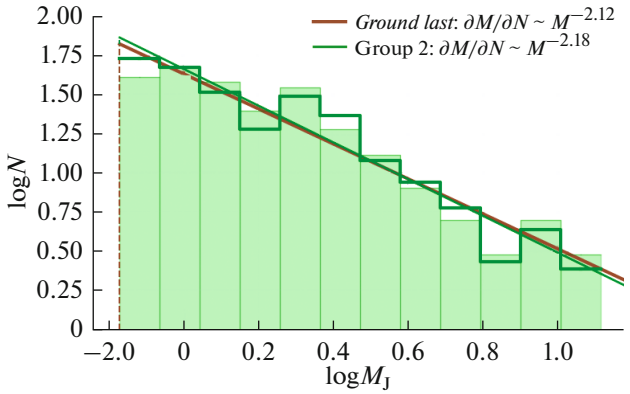

Fig. 9. Histograms of the mass distributions of exoplanets of group 2 according to the Archive data (365 planets with masses $M \in [0.1, 13]M_J$ detected with the ground-based instruments and the CoRoT and Hubble space telescopes). Two cases are presented: the transit probability is ignored (filled area) and taken into account (stepped line). The slanting lines show the distribution laws $\partial N/\partial M \sim M^{\alpha}$ obtained in Series IAA (*Ground last*) and the present study (Group *2*). The vertical dashed lines indicate the minimal values $m_{\min}$.

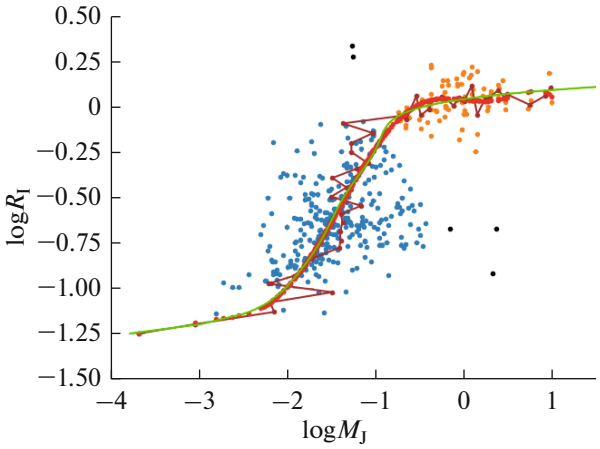

Fig. 10. The functions $M(R)$ (green line) and $R(M)$ (red dots) and the mean logarithmic values of the mass and the radius (brown dots connected by brown lines). The exoplanets of group 1 (the *Raw Data* sample contains 386 planets with a mass of $M \le 13M_J$ detected by the *Kepler* and TESS space telescopes according to the Archive) are shown by dots. The planets of small masses are divided by radius (blue dots) while the giants, by mass (orange dots). Five planets are considered to be outliers (black dots); and they were ignored in the analysis, since each of them produces a significant effect, if it is accounted for.

Estimation of the Masses of Exoplanets by Their Radii

Due to the data uncertainty, the masses of exoplanets of group 1, which were calculated by the models described in the section "The statistical mass–radius relationships...", should be corrected. The corrections of two types are introduced. For 21 planets, the masses of which are known, $M^{\Delta m^+}_{\Delta m^-}$ ($M_{\min} = M - \Delta m^-$ and $M_{\max} = M + \Delta m^+$), the mean relative deviation exceeds the mass value ($\delta M = (\Delta m^+ + \Delta m^-)/(2M) \ge 1$). Consequently, for them, the value M^* obtained from simulations under the condition $M^* \in [M_{\min}, M_{\max}]$ is accepted. If the calculated value does not fall into this interval, the closest value from this interval is assumed for M^*: $M = \min(|M^* - M_{\min}|, |M^* - M_{\max}|)$. Moreover, for 53 planets with masses $M < 13M_J$, for which only the upper estimate of the mass $M_{\max}$ is known, the smallest value among $M_{\max}$ and M^* is assumed: $M = \min(M_{\max}, M^*)$.

The above-mentioned 74 planets (21 plus 53) can be placed into a group, which is intermediate between the groups, containing planets with reliably known masses and unknown masses, respectively. The results of simulations are presented in Fig. 11. As is seen from the diagrams, for the same exoplanet of the intermediate group (the off-line points for the *Averaged* and *Bashi17* models), depending on the model, the calculated value or the value from the Archive will be used.

When analyzing the exoplanets of group 2, we ignored 12 exoplanets from the analogous intermediate group (for 11 out of these planets, the mean deviation exceeds the mass, while one planet has only an upper estimate of the mass).

The Power Law Compliance Test

The Kolmogorov–Smirnov statistics is defined as

$$D = \max\left|F_E(M) - F_T(M\,|\alpha, m_{\min})\right|, \qquad (8)$$

where $F_E(M)$ is the distribution function of the considered (empirical) sample $\mathbf{M} = \{m_i\}$, $i = 1...n$ and $F_T(M|\alpha, m_{\min})$ is the theoretical distribution function with parameters (α, $m_{\min}$) determined from the distribution of the sample $\mathbf{M}$ (Lemeshko, 2014). The hypothesis H_0 suggests that the empirical sample $\mathbf{M}$ obeys the theoretical power law $F_T(M|\alpha, m_{\min})$. This hypothesis is composite, since the parameters α and $m_{\min}$ are defined on the same sample, the compliance of which is verified. Consequently, the compliance criterion is not independent of the distribution (the statistical distribution law depends on the form of the distribution law, the method of estimating the parameters, and their values (see Lemeshko, 2014; Clauset et al., 2009)). In this case, we may verify the hypothesis by the distribution law generated for the statistics. For this, N artificial samples $\mathbf{M}_j = \{m_i\}$, $i = 1...n, j = 1...N$, which are similar to the analyzed one, are generated in the following way (Clauset et al., 2009):

(1) The number of masses n_{power} from the sample $\mathbf{M}$, for which $m \ge m_{\min}$, are determined.

(2) The values m_i are generated N times with the probability $p_{\text{power}} = n_T/n$; these values correspond to the power law distribution $F(\alpha, m_{\min})$ with the obtained parameters on the sample $\mathbf{M}$, and the value m_i is randomly chosen from the sample $\mathbf{M}$ on the

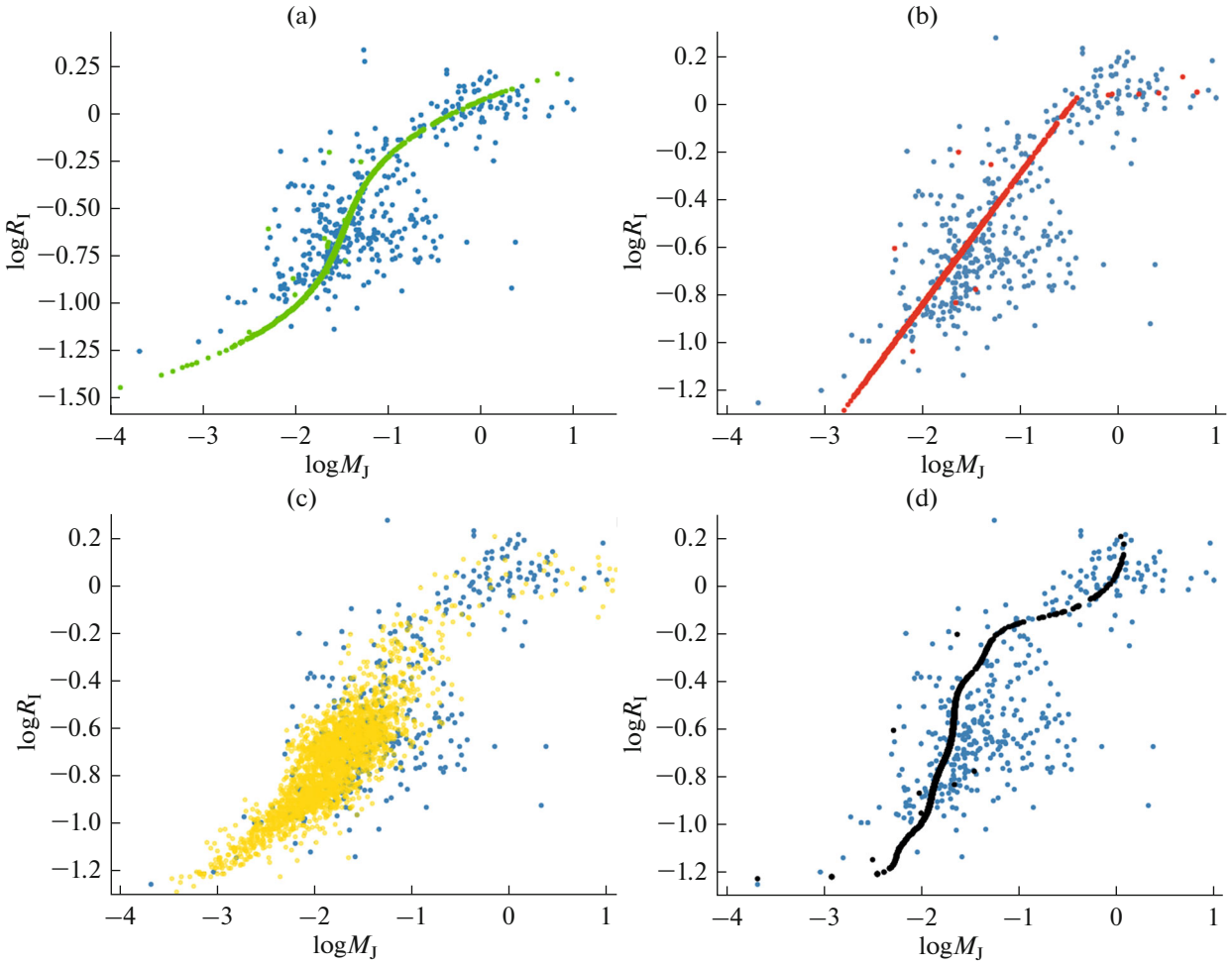

Fig. 11. The exoplanets of group 1 (the planets with a mass of $M \le 13 M_J$ detected by the *Kepler* and TESS space telescopes) (blue dots) are shown in the log$M$–log$R$ plane together with the exoplanets, the masses of which were calculated from the statistical mass–radius relationships according to the *Averaged* (a), *Bashi17* (b), *Chen17* (c), and *Ning18* (d) models.

semiopen interval $[M_{min}, m_{min})$ with the probability $p_{uniform} = 1 - p_{power}$.

After this, in the same way as it has been done for the empirical sample **M**, we determine the parameters α_j and m_{min_j} and the statistics D_j for each of the artificial samples $\mathbf{M}_j$. Further, we find the ratio of the number of artificial samples $N_{Dj > D}$, the distribution of which differs from the theoretical distribution law $F_j(\alpha_j, m_{min\_j})$ more than the distribution of the empirical sample differs from the theoretical one $F(\alpha, m_{min})$, to the total number N of artificial samples: $p = N_{Dj > D}/N$. The hypothesis H_0 is not rejected (i.e., it is assumed that the distribution obeys the power law), if this portion is larger than the critical value p_{cr}, which is specified beforehand. In other words, the hypothesis H_0 is accepted, if more than $p_{cr} \times 100\%$ of the similar artificial samples differ from the theoretical law more than the analyzed empirical sample. According to recommendations by Clauset et al. (2009), $p_{cr} = 0.1$ and $N = 2500$ are assumed.

Figure 12 presents the result of the power law compliance test performed by varying the lower boundary of the distribution for the analyzed samples. The characteristic values of m_{min}, at which the hypothesis was not rejected, are listed in Table 7. As an optimal value for the parameter α (see the section "The lower boundary..."), we assume the value corresponding to

Table 6. The coefficients a_i of the polynomial $Y(X)$

Dependence	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7
$R(M)$	-7.02×10^{-3}	6.67×10^{-2}	-1.87×10^{-1}	-3.69×10^{-3}	3.85×10^{-1}	-2.07×10^{-2}	5.01×10^{-2}	3.66×10^{-2}
$M(R)$	7.94	337.7	580.2	512.5	245.4	61.7	8.49	−0.42

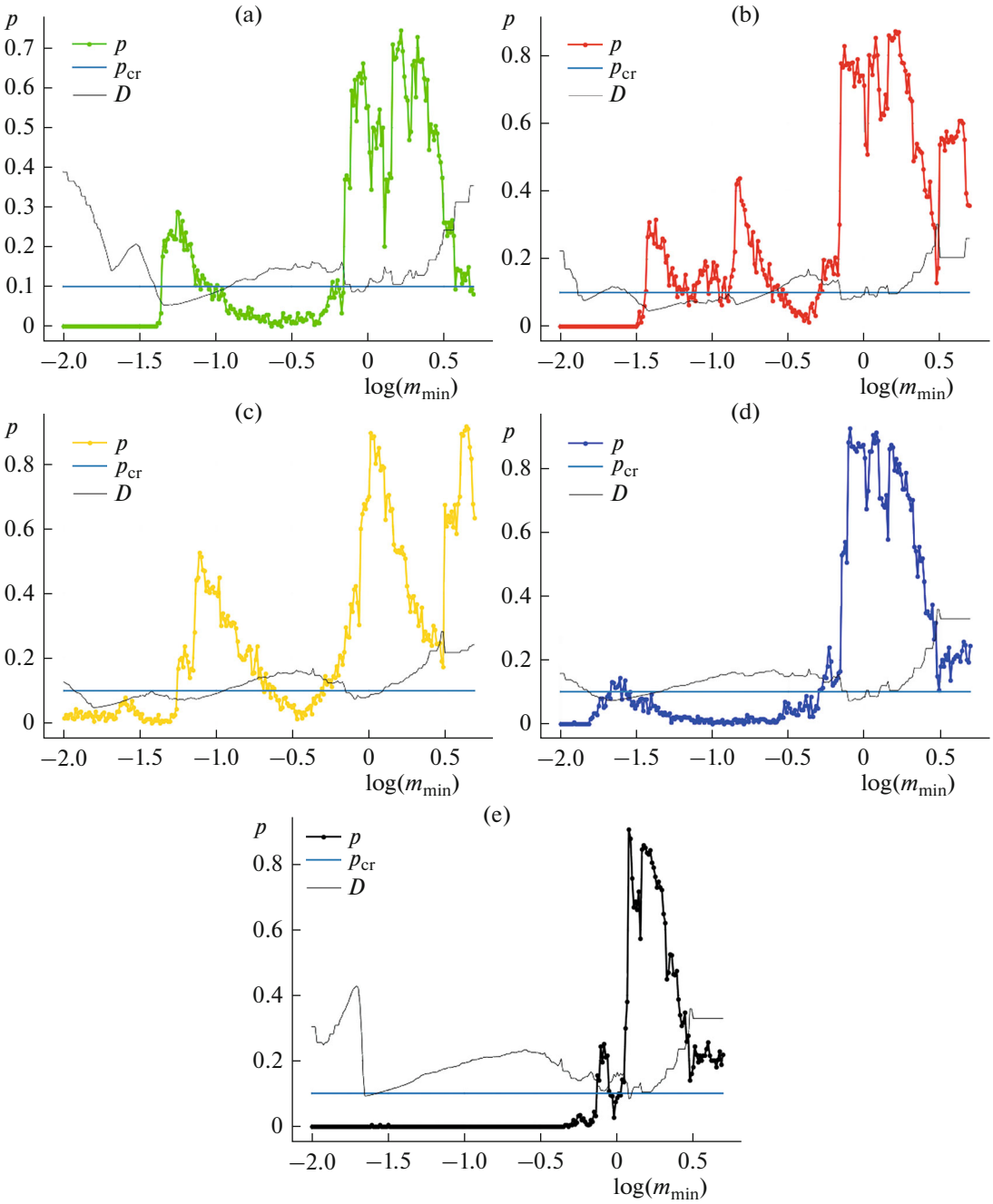

Fig. 12. The dependences of $p = N_{Dj > D}/N$ and D (Eq. (7)) on the minimal mass value in the considered interval for exoplanets of group 1: the *Raw Data* sample of exoplanets detected by the *Kepler* and TESS space telescopes according to the Archive data (a) and the samples obtained by enlarging the *Raw Data* by the planets with the masses calculated from the mass–radius relationships according to the *Averaged* (b), *Bashi17* (c), *Chen17* (d), and *Ning18* (e) models. The horizontal line indicates the value $p_{cr} = 0.1$.

Table 7. The values of m_{min}, for which the hypothesis on compliance with the power law is not rejected

Sample	The nearest to $0.02M_J$ interval	Minimal value $m_{min} < 1.8M_J$ (or $0.5\log(M_J)$)
Raw Data	[0.021, 0.031]	0.51
Averaged	[0.046, 0.087]	0.54
Bashi17	[0.038, 0.061]	0.54
Chen17	[0.058, 0.24]	0.56
Ning18	[0.75, 0.94]	1.28

the smallest criterion D (i.e., the best agreement with the power law) with the nearest interval of $m_{\min}$ to $0.02M_J$, for which $p_{cr} < p$.

CONCLUSIONS

The mass distributions of transit exoplanets were analyzed considering the peculiarities in the mass–radius relationships.

The distributions built from the masses calculated according to different mass–radius relationships generally correspond to those obtained in Series IAA, where the observational selection was corrected without using the mass–radius relationships. The agreement persisted after accounting for the transit probability.

Analogously to the result of Series IAA, accounting for the transit probability for exoplanets detected mainly by the ground-based telescopes does not introduce statistically important changes into the distribution on the considered mass interval. According to the present analysis, this result is also valid for the exoplanets detected by the *Kepler* and TESS space telescopes. However, it is not consistent with the conclusion of Series IAA about this group of planets. This disagreement may be explained by the variability and roughness of the models used to determine the statistical mass of an exoplanet by its radius.

The transit photometric method covers the exoplanets with a semimajor orbital axis smaller than 1 AU, and our analysis of the corresponding data yielded no evidence regarding the structurization in planetary systems (i.e., the interrelation between the mass of an exoplanet and its mean distance to a host star). We have arrived at this conclusion because all, except one, distributions of the considered samples (obtained from different mass–radius relationships) before and after accounting for the observational selection are statistically indistinguishable from each other.

The mass distribution of exoplanets, for which the unknown masses were calculated according to the averaged dependence $M(R)$ (*Averaged*), does not differ much from the distributions built according to the more complex *Bashi17* and *Chen17* models. The *Ning18* dependence $M(R)$ and the distribution obtained from this model substantially differ from the other dependences and distributions due to the behavior of the *Ning18* dependence and that of the other $M(R)$ dependences in the region corresponding to the giant planets.

For the *Ning18* and *Raw Data* samples, we constructed the broken power-law, the indices of which differ on three intervals. The distribution laws found were compared to the analogous law based on the planetary population synthesis analysis (Mordasini, 2018). The *Ning18* distribution, which was enlarged by adding the masses, agrees best of all (as compared to the other samples considered) with that by Mordasini (2018).

FUNDING

The study was supported by the Government and the Ministry of Education and Science of Russia (grant no. 075-15-2020-780 (N13.1902.21.0039)).

ACKNOWLEDGMENTS

In this paper we used the NASA Exoplanet Archive (2020), which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

REFERENCES

Ananyeva, V.I., Ivanova, A.E., Venkstern, A.A., Shashkova, I.A., Yudaev, A.V., Tavrov, A.V., Korablev, O.I., and Bertaux, J.-L., Mass distribution of exoplanets considering some observation selection effects in the transit detection technique, *Icarus,* 2020a, vol. 346, art. id. 113773.

Ananyeva, V.I., Ivanova, A.E., Venkstern, A.A., Tavrov, A.V., Korablev, O.I., and Bertaux, J.-L., The dependence of the mass distribution of exoplanets on the spectral class of host stars, *Sol. Syst. Res.,* 2020b, vol. 54, no. 3, pp. 175–186.

Bashi, D., Helled, R., Zucker, S., and Mordasini, C., Two empirical regimes of the planetary mass–radius relation, *Astron. Astrophys.,* 2017, vol. 604, art. id. A83.

Butler, R.P., Wright, J.T., Marcy, G.W., Fischer, D.A., Vogt, S.S., Tinney, C.G., Jones, H.R.A., Carter, B.D., Johnson, J.A., and McCarthy, C., Catalog of nearby exoplanets, *Astrophys. J.,* 2006, vol. 646, p. 505.

Chen, J. and Kipping, D., Probabilistic forecasting of the masses and radii of other worlds, *Astrophys. J.,* 2017, vol. 834, pp. 17–30.

Clauset, A., Shalizi, C., and Newman, M., Power-law distributions in empirical data, *SIAM Rev.,* 2009, vol. 51, no. 4, pp. 661–703.

Cumming, A., Butler, R.P., Marcy, G.W., Vogt, S.S., Wright, J.T., and Fischer, D.A., The Keck planet search: detectability and the minimum mass and orbital period distribution of extrasolar planets, *Publ. Astron. Soc. Pac.,* 2008, vol. 120, pp. 531–554.

Howard, A.W., Marcy, G.W., Johnson, J.A., Fischer, D.A., Wright, J.T., Isaacson, H., Valenti, J.A., Anderson, J., Lin, D.N., and Ida, S., The occurrence and mass distribution of close-in super-Earths, Neptunes, and Jupiters, *Science,* 2010, vol. 330, pp. 653–655.

Ida, S. and Lin, D. N. C., Toward a deterministic model of planetary formation. II. The formation and retention of gas giant planets around stars with a range of metallicities, *Astrophys. J.,* 2004, vol. 616, no. 1, p. 567.

Ivanova, A.E., Ananyeva, V.I., Venkstern, A.A., Shashkova, I.A., Yudaev, A.V., Tavrov, A.V., Korablev, O.I., Bertaux, J.-L., The mass distribution of transiting exoplanets corrected for observational selection effects, *Astron. Lett.,* 2019, vol. 45, no. 10, pp. 687–694.

Lemeshko, B.Yu., *Neparametricheskie kriterii soglasiya* (Nonparametric Goodness-of-Fit Tests), Novosibirsk: Novosibirsk Gos. Univ., 2014.

Marcy, G., Butler, R.P., Fischer, D., Vogt, S., Wright, J.T., Tinney, C.G., and Jones, H.R., Observed properties of exoplanets: masses, orbits, and metallicities, *Prog. Theor. Phys.,* 2005, vol. 158 Suppl., pp. 24–42.

Mordasini, C., Alibert, Y., Georgy, C., Dittkrist, K.-M., Klahr, H., and Henning, T., Characterization of exoplanets from their formation II: The planetary mass–radius relationship, *Astron. Astrophys.,* 2012, vol. 547, art. id. A112.

NASA Exoplanet Archive, 2020. Accessed July 25, 2020. https://doi.org/10.26133/NEA1

Ning, B., Wolfgang, A., and Ghosh, S., Predicting exoplanets mass and radius: A nonparametric approach, *Astrophys. J.,* 2018, vol. 869, no. 1.

Otegi, J.F., Bouchy, F., and Helled R., Revisited mass–radius relations for exoplanets below 120 Earth masses, *Astron. Astrophys.,* 2019, vol. 634, art. id. A43.

Petigura, E.A., Howard, A.W., and Marcy, G.W., Prevalence of Earth-size planets orbiting Sun-like stars, *Proc. Natl. Acad. Sci. USA,* 2013, vol. 110, no. 48, pp. 19273–19278.

Tuomi, M., Jones, H.R.A., Butler, R.P., Arriagada, P., Vogt, S.S., Burt, J., Laughlin, G., Holden, B., Shectman, S.A., Crane, J.D., Thompson, I., Keiser, S., Jenkins, J.S., Berdinas, Z., Diaz, M., Kiraga, M., and Barnes, J.R., Frequency of planets orbiting M dwarfs in the Solar neighbourhood, *Earth Planet. Astrophys.,* 2019. https://arxiv.org/abs/1906.04644.

Seager, S. and Mallén-Ornelas, G. A., Unique solution of planet and star parameters from an extrasolar planet transit light curve, *Astrophys. J.,* 2002, vol. 585, no. 2.

Weiss, L.M., Marcy, G.W., Rowe, J.F., Howard, A.W., Isaacson, H., Fortney, J.J., Miller, N., Demory, B.-O., Fischer, D.A., and Adams, E.R., The mass of KOI-94d and a relation for planet radius, mass, and incident flux, *Astrophys. J.,* 2013, vol. 768, no. 14.

Winn, J., Transits and occultations, *Earth and Planetary Astrophysics,* 2014. https://arxiv.org/abs/1001.2010v5.

Wolfgang, A., Rogers, L.A., and Ford, E.B., Probabilistic mass–radius relationship for sub-Neptune-sized planets, *Astrophys. J.,* 2016, vol. 825, no. 19.

Zeng, Li., Sasselov, D.D., and Jacobsen, S.B., Mass-radius relation for rocky planets based on PREM, *Astrophys. J.,* 2016, vol. 819, no. 2.

Translated by E. Petrova