Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **00,** 1 (2019)
Advance Access publication 2019 March 02

doi:10.1093/mnras/stz615

Predicting extragalactic distance errors using Bayesian inference in multimeasurement catalogues

Germán Chaparro-Molano [1]★ Juan Carlos Cuervo,[1] Oscar Alberto Restrepo Gaitán[1,2] and Sergio Torres Arzayús[3]

[1] *Vicerrectoría de Investigación, Universidad ECCI, Bogotá, Colombia*
[2] *Radio Astronomy Instrumentation Group, Universidad de Chile, Santiago de Chile, Chile*
[3] *Centro Internacional de Física, Bogotá, Colombia*

Accepted 2019 February 26. Received 2019 January 17; in original form 2018 May 3

ABSTRACT

We propose the use of robust, Bayesian methods for estimating extragalactic distance errors in multimeasurement catalogues. We seek to improve upon the more commonly used frequentist propagation-of-error methods, as they fail to explain both the scatter between different measurements and the effects of skewness in the metric distance probability distribution. For individual galaxies, the most transparent way to assess the variance of redshift independent distances is to directly sample the posterior probability distribution obtained from the mixture of reported measurements. However, sampling the posterior can be cumbersome for catalogue-wide precision cosmology applications. We compare the performance of frequentist methods versus our proposed measures for estimating the true variance of the metric distance probability distribution. We provide pre-computed distance error data tables for galaxies in three catalogues: NED-D, HyperLEDA, and Cosmicflows-3. Additionally, we develop a Bayesian model that considers systematic and random effects in the estimation of errors for Tully–Fisher (TF) relation derived distances in NED-D. We validate this model with a Bayesian *p*-value computed using the Freeman–Tukey discrepancy measure as a posterior predictive check. We are then able to predict distance errors for 884 galaxies in the NED-D catalogue and 203 galaxies in the HyperLEDA catalogue that do not report TF distance modulus errors. Our goal is that our estimated and predicted errors are used in catalogue-wide applications that require acknowledging the true variance of extragalactic distance measurements.

Key words: methods: data analysis – methods: statistical – astronomical data bases: miscellaneous – catalogues – galaxies: distances and redshifts – galaxies: statistics.

1 INTRODUCTION

Understanding the uncertainties in redshift-independent extragalactic distance measurements is absolutely necessary before reporting statistically sound conclusions regarding the structure of the local universe (Nasonova & Karachentsev 2011; Courtois et al. 2012; Ma, Taylor & Scott 2013; Sorce et al. 2014; Springob et al. 2014; Said et al. 2016; Kourkchi & Tully 2017), large-scale structure (McClure & Dyer 2007; Javanmardi & Kroupa 2017; Roman & Trujillo 2017; Jesus et al. 2018; Torres & Cuervo 2018), and events like transient gravitational wave detections (White, Daw & Dhillon 2011). Hubble constant estimations have been using increasingly sophisticated statistical tools for primary distance determination methods, such as Type Ia supernova (SNIa; Barris & Tonry 2004; Rubin et al. 2015; Dhawan, Jha & Leibundgut 2018), Cepheids (Humphreys et al. 2013), or both (Riess et al. 2016). Although most estimates of the Hubble constant use Cepheid calibration for calibrating secondary methods (Tully & Pierce 2000; Freedman et al. 2001; Freedman & Madore 2010), Mould & Sakai (2008) have explored changes in Hubble constant estimation using the Tully–Fisher (TF) relation without Cepheid calibration. Secondary methods for extragalactic distance determination like the TF relation or the Fundamental Plane (FP) have recently become more precise thanks to increasing volumes of data from surveys like 6dF (Springob et al. 2014) and Two Micron All-Sky Survey (2MASS; Jarrett et al. 2000; Springob et al. 2007) together with *Spitzer* data (Sorce et al. 2013), along with improved statistical methods (Obreschkow & Meyer 2013).

As of 2018, three multimeasurement catalogues including a substantial amount of redshift-independent extragalactic distance measurements have been released: HyperLEDA (Makarov et al.

★ E-mail: germancho@gmail.com

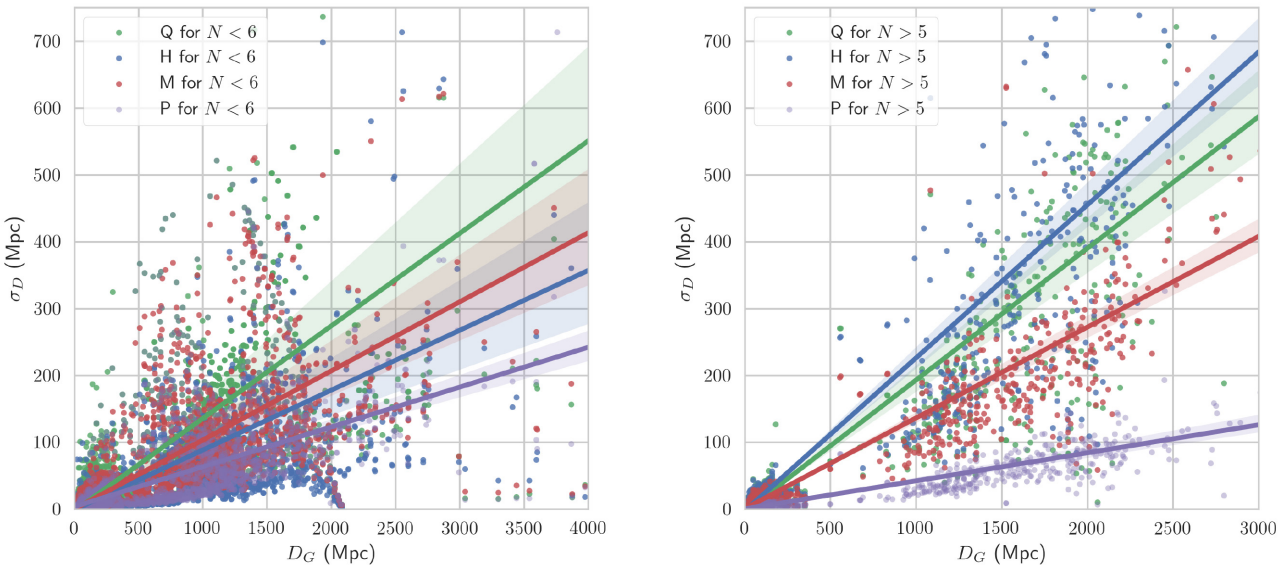

Figure 1. Estimated extragalactic distance errors versus median extragalactic distance for galaxies with $N < 6$ (left) and $N > 5$ (right) redshift-independent distance measurements in NED-D according to the H, M, Q, and P error models (explained in the text), showing linear regressions and confidence intervals computed using the SEABORN.REGPLOT PYTHON function.

2014), NED-D (Mazzarella & Team 2007; Steer et al. 2017), and Cosmicflows-3 (Tully, Courtois & Sorce 2016). HyperLEDA includes a homogenized catalogue for extragalactic distances in the nearby universe, with 12 866 distance measurements for 518 galaxies to date. NED-D is the NASA/IPAC Extragalactic Distance catalogue of redshift-independent distances, which compiles 326 850 distance measurements for 183 062 galaxies in its 2018 version. Here, ∼1800 galaxies (∼1 per cent) have more than 13 distance measurements, and 180 galaxies (∼0.1 per cent) have distance measurements using more than six different methods. Cosmicflows-3 is the most up-to-date catalogue, which reports distance measurements for 10 616 galaxies (all of which include errors) using up to four distance determination methods, calibrated with supernova luminosities. However, unlike HyperLEDA or NED-D, Cosmicflows-3 only reports the latest distance measurement for each method. In HyperLEDA, NED-D, and Cosmicflows-3 errors are reported as one standard deviation from the reported distance modulus. Treatment of errors for combining distance moduli across methods or across measurements is suggested by Mazzarella & Team (2007) and Tully et al. (2016) to be based on weighted estimates such as the uncertainty of the weighted mean, albeit with caution partly due to the heterogeneous origin of the compiled data and partly due to Malmquist bias. In the case of NED-D, this is additionally complicated by the fact that many errors are not reported or are reported as zero. In fact, the TF relation method has the largest number of galaxies with non-reported distance modulus errors (884 to date). Even though extragalactic distances measured using the TF relation were originally reported to have a relative error in distance modulus of 10–20 per cent (Tully & Fisher 1977), we consider that this conservative estimate can be improved upon by using a predictive model based on the distance error of galaxies that use the same distance determination method. This requires a robust estimation of the variance of extragalactic distances based on the available data.

For many galaxies in all three catalogues, the random error for each distance modulus measurement ϵ_i (for $i = 1, ..., N$, for N distance measurements per galaxy) is not representative of the scatter across measurements, even when considering the same method for determining distances. In addition, distance modulus distributions for each measurement (which are assumed to be Gaussian) are transformed to lognormal distributions in metric distance space. This can introduce a significant bias in peculiar velocity studies for large-scale structure studies (Watkins & Feldman 2015). We improve upon previous methods by robustly estimating the underlying variance across measurements and distance determination methods. To do this, we measure the 84th and 16th percentiles, and the median absolute deviation of the bootstrap-sampled posterior probability distribution of each extragalactic distance (Chaparro Molano et al. 2018). We compare our results to other more commonly used frequentist methods, such as the weighted estimates mentioned above, and we produce pre-computed data tables for the three catalogues mentioned above, which can be found in the repository for this work at `https://github.com/saint-germain/errorprediction`. We then perform a Bayesian analysis of the systematics and randomness of the computed errors in the NED-D catalogue for TF relation derived distances. From this analysis we build predictive models for the estimation of errors and evaluate them by performing posterior predictive checks using a discrepancy measure-derived Bayesian 'p-value' (Gelman, Meng & Stern 1996). Furthermore, we make predictions for the 884 galaxies in the NED-D catalogue and the 203 galaxies in the HyperLEDA catalogue whose distances were measured using the TF relation but have non-reported errors. Inference based on Bayesian posterior predictive checks has been advocated for in Gelman (2003) and Chambert, Rotella & Higgs (2014).

We organize this paper as follows. In Section 2, we talk about the posterior distribution of distance for individual galaxies and set up methods for measuring its variance. In Section 3, we make a comparison between the proposed variance estimation methods, and in Section 4, we propose and evaluate predictive Bayesian models for two robust methods of error estimation, and we summarize our work in Section 5. The Appendix includes a description and brief analysis of extragalactic distance error data tables pre-computed

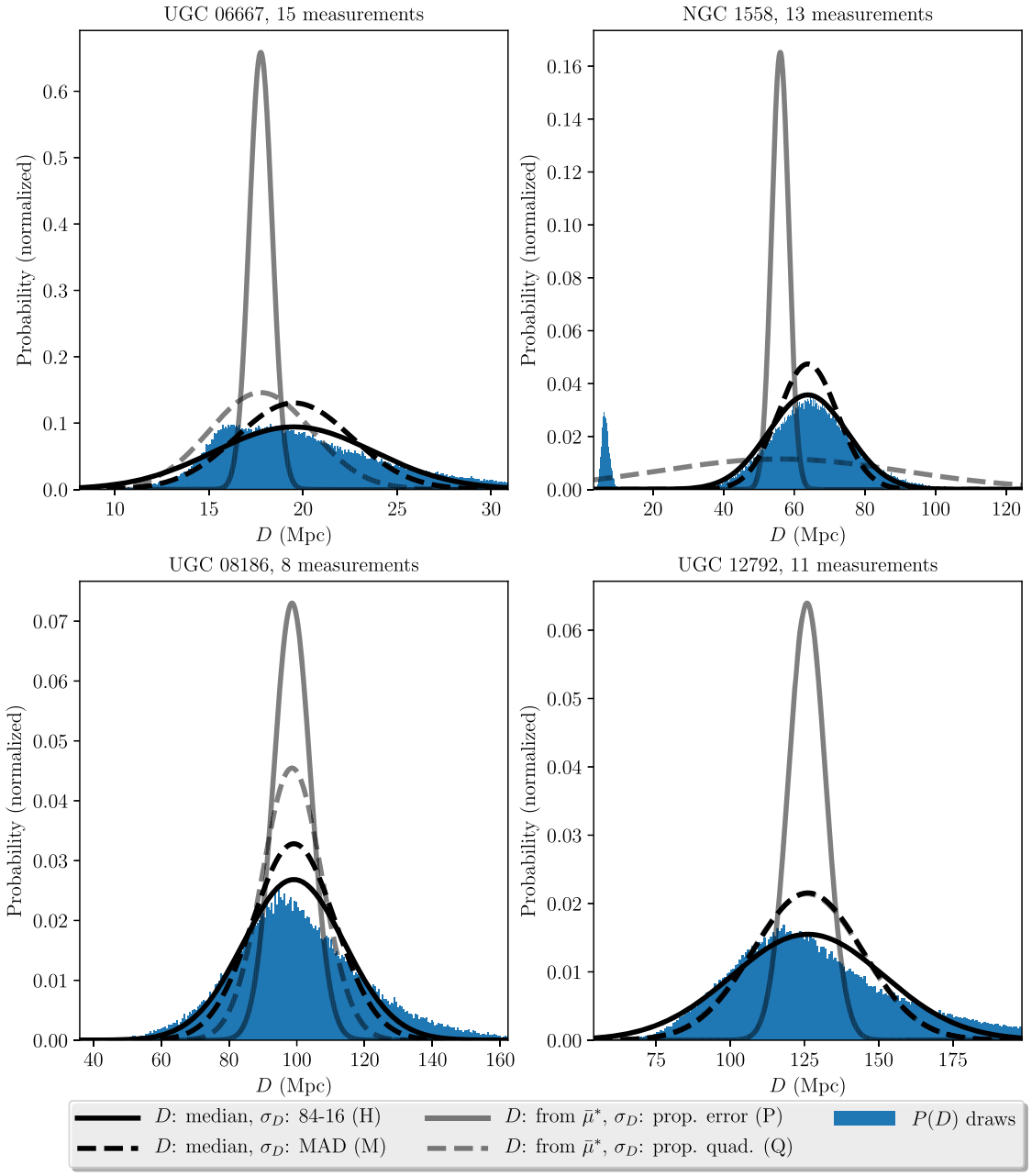

Figure 2. Comparison of four examples of extragalactic distance posterior distribution draws (10 000 per measurement) and modelled distributions for UGC 06667, NGC 1558, UGC 08186, and UGC 12792 using the Tully–Fisher (TF) method for distance determination in NED-D. The methods used for approximating the posterior distribution (H, M, P, and Q) are described in the text.

with the methods described in this paper for the HyperLEDA, Cosmicflows-3, and NED-D catalogues.

2 ESTIMATION OF EXTRAGALACTIC DISTANCE ERRORS

The best approach to consider the effects of random and systematic errors in catalogue-wide, multimethod distance analyses is to directly sample the posterior probability distribution of each extragalactic distance. This can be achieved by drawing distance modulus samples from $P(\mu)$, which is the unweighted mixture of normal distributions corresponding to each distance modulus measurement μ_i,

$$\mu \sim \sum_i^N \mathcal{N}(\mu_i, \epsilon_i^2),$$

and then converting to metric distance,

$$D = 10^{\frac{\mu}{5}+1}.$$

Therefore,

$$D_\mathrm{G} \sim \sum_i^N \mathrm{lognormal}(M_i, \sigma_{M_i}^2). \quad (1)$$

Here $M_i = \ln D_i$ and $\sigma_{M_i} = \epsilon_i \ln 10$.

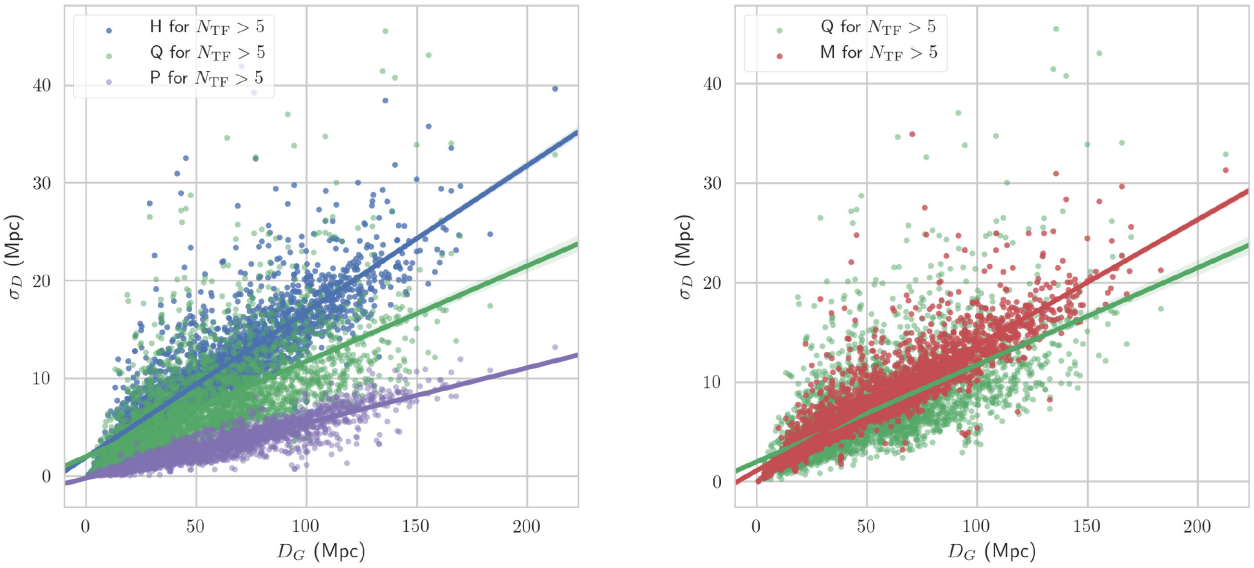

Figure 3. Estimated extragalactic distance errors versus median extragalactic distance for galaxies with more than five TF distance measurements in NED-D according to the H, Q, and P (left) and Q and M (right) error models, showing linear regressions and confidence intervals computed using the SEABORN.REGPLOT PYTHON function.

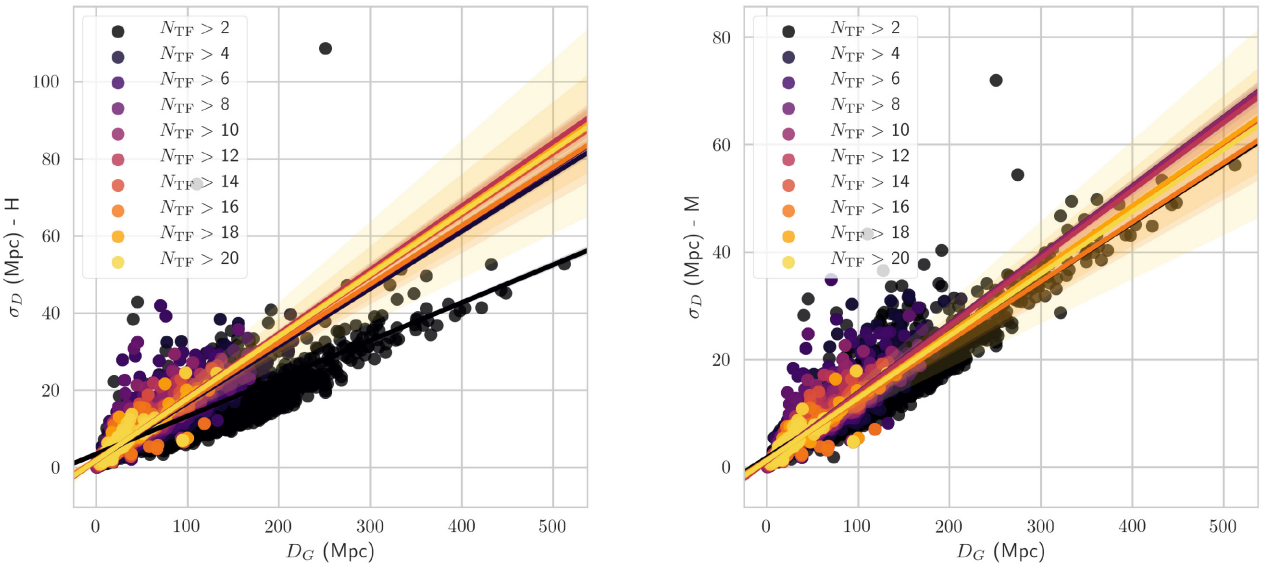

Figure 4. Estimated extragalactic distance errors versus median extragalactic distance using the TF method for distance determination in NED-D, according to the H error model (left) and the M error model (right) showing linear regressions and confidence intervals computed using the SEABORN.REGPLOT PYTHON function.

2.1 Estimating the variance of $P(D_G)$

Although directly sampling the distribution of D_G is the most transparent way to acknowledge the true variance of distance measurements, it is not a very efficient way to achieve a standardized treatment of errors. One simple measure of the variance of D_G that acknowledges the possible skewness of the distribution is to take the median 16th and 84th percentile of 10 000 bootstrap samples of the distribution of D_G, e.g. one bootstrap sample corresponds to N draws, one from each reported measurement. In our pre-computed error tables we report these quantities as `Dmin` and `Dmax`, respectively.

It can be even more convenient to treat each extragalactic metric distance D_G as a normal random variable with a single-valued σ_D as a measure of the uncertainty in the estimation of an extragalactic distance,

$$D_G \sim \mathcal{N}(D, \sigma_D^2). \quad (2)$$

For this reason we compare four methods for estimating the D, σ_D pair. Two of these methods (H, M) use robust measures of the distribution of each extragalactic distance, and the other two (P, Q) use measures based on propagation of errors.

Methods H and M, which are the methods we propose to robustly estimate σ_D in equation (2), are based on measuring the median

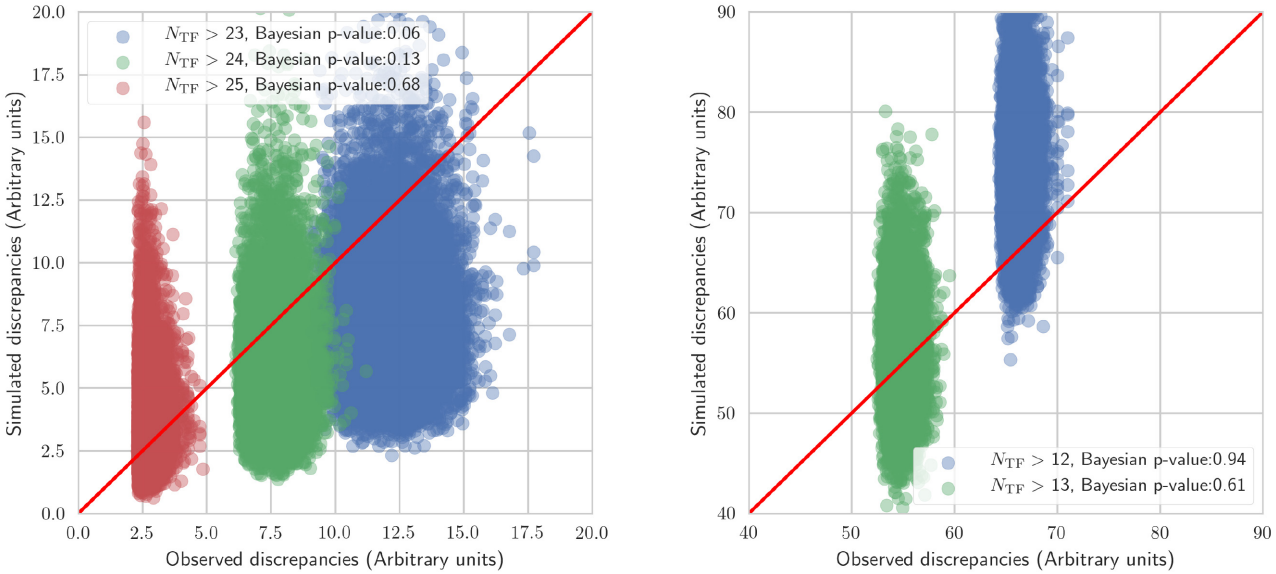

Figure 5. Discrepancy plot for the Bayesian quadrature model (equation 9) based on errors estimated using method H for $N_{TF} > 23, 24, 25$ (left) and using method M for $N_{TF} > 12, 13$ (right).

and variance of repeated bootstrap samples from the distribution of D_G (equation 1) as mentioned in the previous section. Method H takes D as the median of the bootstrap samples and σ_D as the half-distance (H) between the 84th and 16th percentiles of 10 000 bootstrap samples. We consider this to be the method that most faithfully measures the variance regardless of the shape of the posterior distribution. Method M takes D as the median of the bootstrap samples and σ_D as the median absolute deviation (MAD) of the bootstrap samples. This method is better suited for avoiding the effects of outliers.

The other two methods (P, Q) considered here are based on commonly used frequentist estimates of the distance error. Method P consists on calculating D from the weighted mean distance modulus $\bar{\mu}^*$ with weights $w_i = \epsilon_i^{-2}$. σ_D is calculated by propagation (P) of measurement errors, i.e. from the uncertainty of the weighted mean (Tully et al. 2016),

$$\sigma_D^P = 0.461\, \bar{D}^* \left(\sum_i^N w_i \right)^{-1/2} . \quad (3)$$

Method P does not take into account the scatter in distance measurements for single galaxies, which is why it can be convenient to calculate σ_D as the sum in quadrature (Q) of the propagated uncertainty of the weighted mean and the propagated unbiased weighted sample variance σ_D^*:

$$\sigma_D^Q = \left[\left(\sigma_D^P\right)^2 + \left(\sigma_D^*\right)^2 \right]^{1/2} . \quad (4)$$

Here σ_D^* is calculated as (Brugger 1969)

$$\sigma_D^* = 0.461\, \bar{D}^* \sqrt{\frac{N}{N-1.5} \frac{\sum_i^N w_i (\mu_i - \bar{\mu}^*)^2}{\sum_i^N w_i}} . \quad (5)$$

If the non-robust P and Q methods were truly representative of the variance of the distribution of D_G, they should yield similar results as the H or M methods. The following section shows that this is not the case.

3 COMPARISON OF DISTANCE ERROR ESTIMATION METHODS

In this section, we focus on NED-D distance measurements since it is the largest of the three catalogues considered here. A full discussion of our error estimation method applied to multimethod measurements in the HyperLEDA, NED-D, and Cosmicflows-3 is given in the Appendix. A repository for this work, including the pre-computed error tables for the HyperLEDA, NED-D, and Cosmicflows-3, is located at `https://github.com/saint-germain/errorprediction`. From here on, when we mention distance measurements in the NED-D catalogue, we will be excluding from our analysis measurements that require the target redshift to calculate the distance, as indicated in the `redshift (z)` field.

For galaxies with a number of distance measurements between 2 and 5 (Fig. 1, left), errors estimated with the quadrature (Q), and median absolute deviation (M) methods show a linear trend with similar slopes that overpredict the variance with respect to the half 84th–16th percentile distance (H) method, whereas the propagation (P) method tends to underpredict the errors. Furthermore, errors estimated using the Q method show a larger dispersion around the linear trend than the H and M methods. Fig. 1 (right) shows that the P and Q methods underpredict errors for galaxies with more than five distance measurements.

3.1 Distance errors in Tully–Fisher relation derived measurements

Even though our analysis for error estimation can be used to combine distance measurements using different methods for single galaxies, we think that due to method-intrinsic systematics it is more appropriate to separate the analysis by method. Without loss of generality, we now focus on galaxies whose distances have been measured using the TF method in the NED-D catalogue because it is the method with the largest number of galaxies without reported measurement errors (884) in the data base.

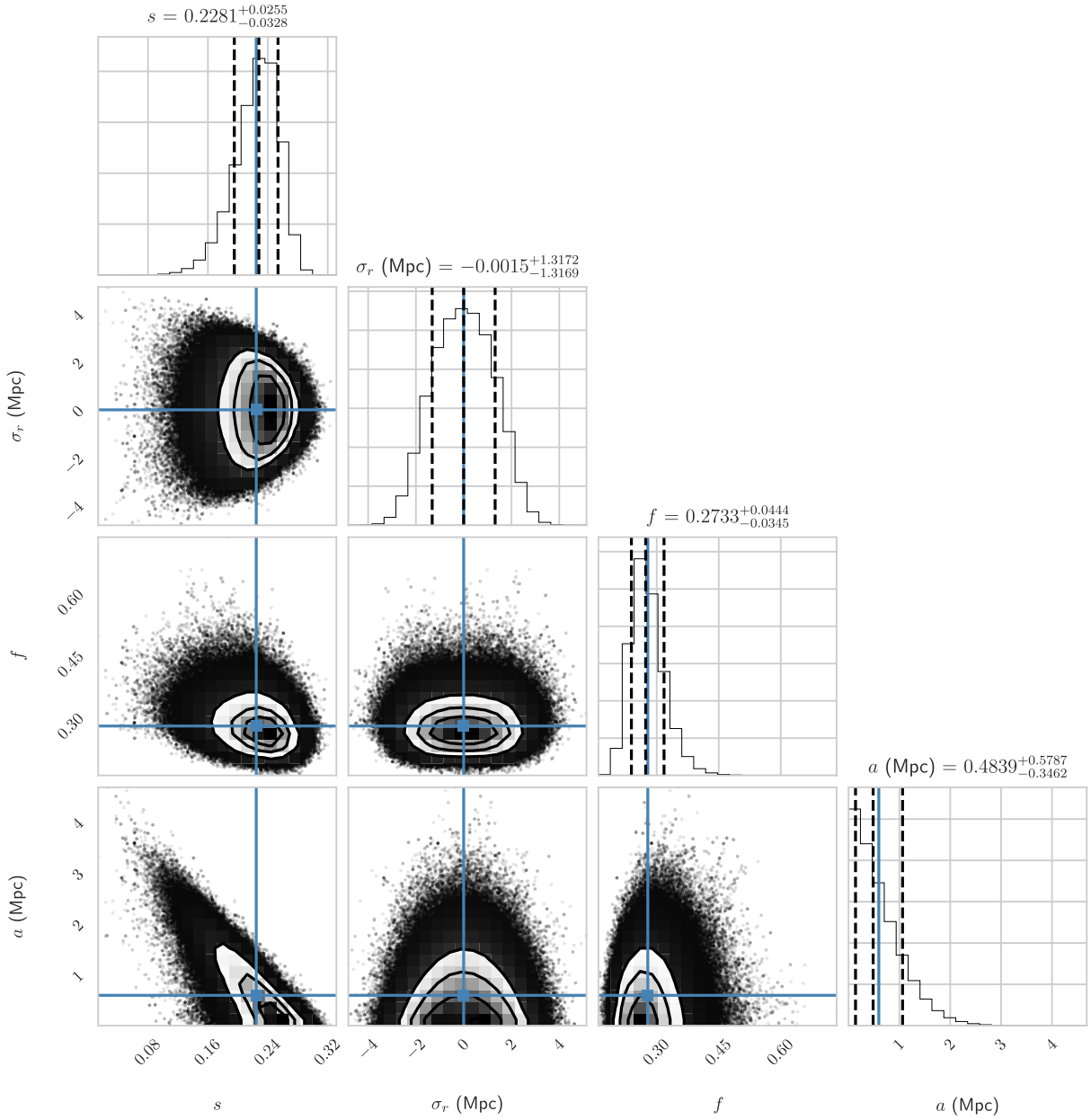

Figure 6. Corner plot showing the EMCEE sampling of the posterior probability distribution (equation 8) for the quadrature Bayesian model parameters $\boldsymbol{\theta} = (s, \sigma_r, f, a)$ based on errors estimated using method H for galaxies with more than 25 TF distance measurements. The dashed lines indicate the 16th, 50th, and 84th percentile of the marginalized distribution of each parameter (shown at the top of each column), and the blue solid lines indicate the mean. This plot was made using the CORNER PYTHON module.

Fig. 2 shows that for a small but representative sample of galaxies with more than seven distance measurements, the centre and variance of the posterior distribution of each extragalactic distance is best explained using the H method, whereas the less robust P and Q methods underpredict the variance. On the other hand, the M method also underpredicts the variance because it is a robust measure, and thus not as sensitive to outliers as methods P and Q, as seen in the case of NGC 1558 in Fig. 2. For the more symmetrical posterior distribution of UGC 12792, the M and Q methods predict the same centre and variance.

Fig. 3 shows that the Q and P methods underpredict distance errors for galaxies with more than five TF distance measurements. On the other hand, method Q underpredicts distance errors with respect to the M method, which again shows a tighter linear correlation due to the robustness of the M measure. However, the scale of H and M errors (relative errors) does not depend strongly on the limiting number of measurements for $N_{\rm TF} > 3$, as Fig. 4 shows.

The general correlation between distance and distance error (Figs 1 and 3) means that there is a strong systematic component in the variance of $P(D_{\rm G})$, which is expected from the conversion of distance modulus to metric distance. To improve visualization, only errors for galaxies with more than five TF distance measurements are shown in Figs 3, 13, and 14.

Given that each $\sigma_{\rm D}$ calculated using the H and M methods is obtained from many realizations from the distribution of extra-

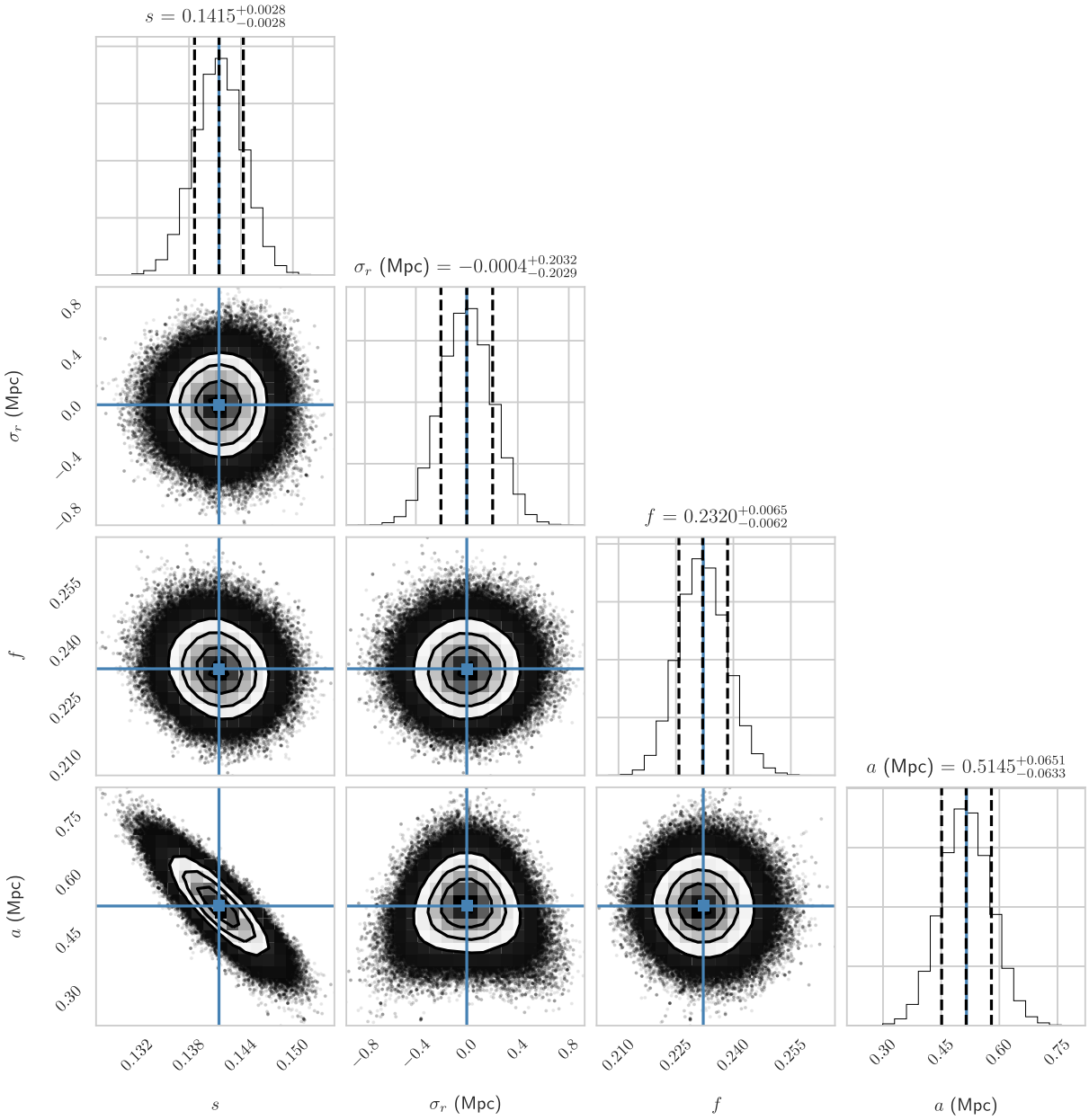

Figure 7. Corner plot showing the EMCEE sampling of the posterior probability distribution (equation 8) for the quadrature Bayesian model parameters $\boldsymbol{\theta} = (s, \sigma_r, f, a)$ based on errors estimated using method M for galaxies with more than 13 TF distance measurements. The dashed lines indicate the 16th, 50th, and 84th percentile of the marginalized distribution of each parameter (shown at the top of each column), and the blue solid lines indicate the mean. This plot was made using the CORNER PYTHON module.

galactic distances, it is also possible to calculate its variance as the half-distance between the 84th and 16th percentile of bootstrap σ_D realizations. Fig. 13 (left) shows that the variance of the estimated error is proportional to the error for the H and M methods. This will be relevant in Section 4 when we construct a predictive model for non-reported errors.

4 PREDICTIVE BAYESIAN MODELS FOR TF MISSING ERRORS

In the multimeasurement catalogues considered here, we observe that the scatter of reported distance measurements and reported individual measurement errors does not match in most cases. This situation happens because there are hidden systematic sources intrinsic to each method of distance estimation. These systematics cannot be removed, but they can be marginalized over in order to estimate the true variance of a distance estimation method based on multiple measurements. The central limit theorem indicates that as the number of measurements increases, the behaviour of the distance errors should settle toward being normally distributed. Thus, if a correlation trend between distance measurements and estimated errors can be explained from a Bayesian viewpoint, then it should be possible to use a Bayesian model to predict missing distance errors for a distance determination method, given enough data. Since more measurements can increase our knowledge of systematic uncertainties in distance measurements, the way we

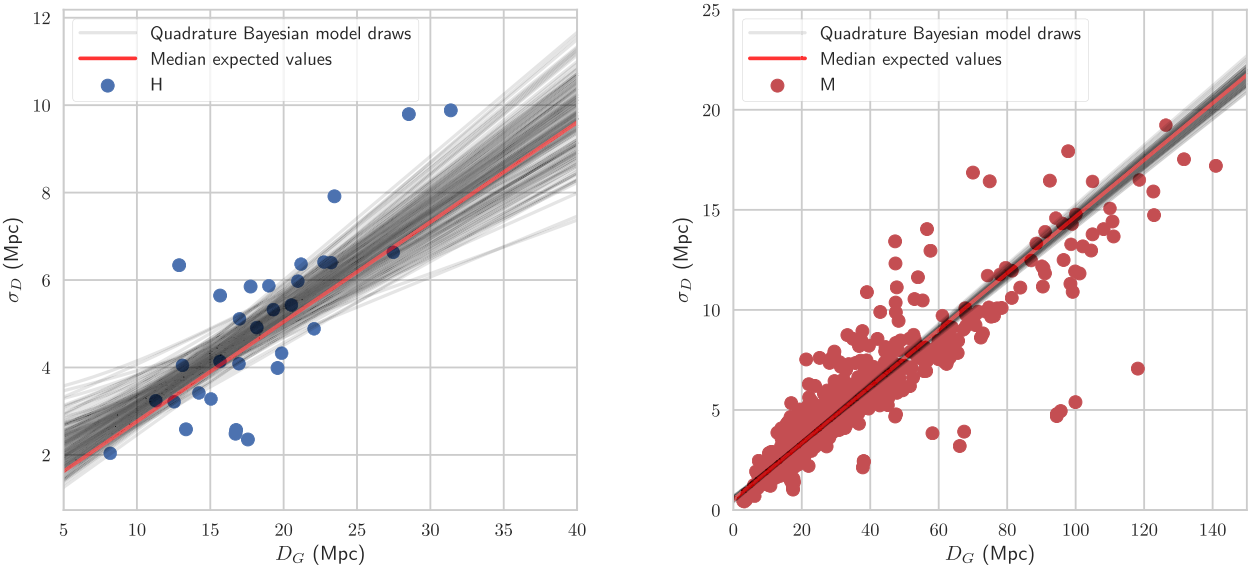

Figure 8. Projection of parameter set samples from the posterior probability distribution of the Bayesian quadrature model on to the σ_D versus D_G scatter plot for errors estimated using method H for galaxies with more than 25 TF distance measurements (left) and using method M for galaxies with more than 13 TF distance measurements (right).

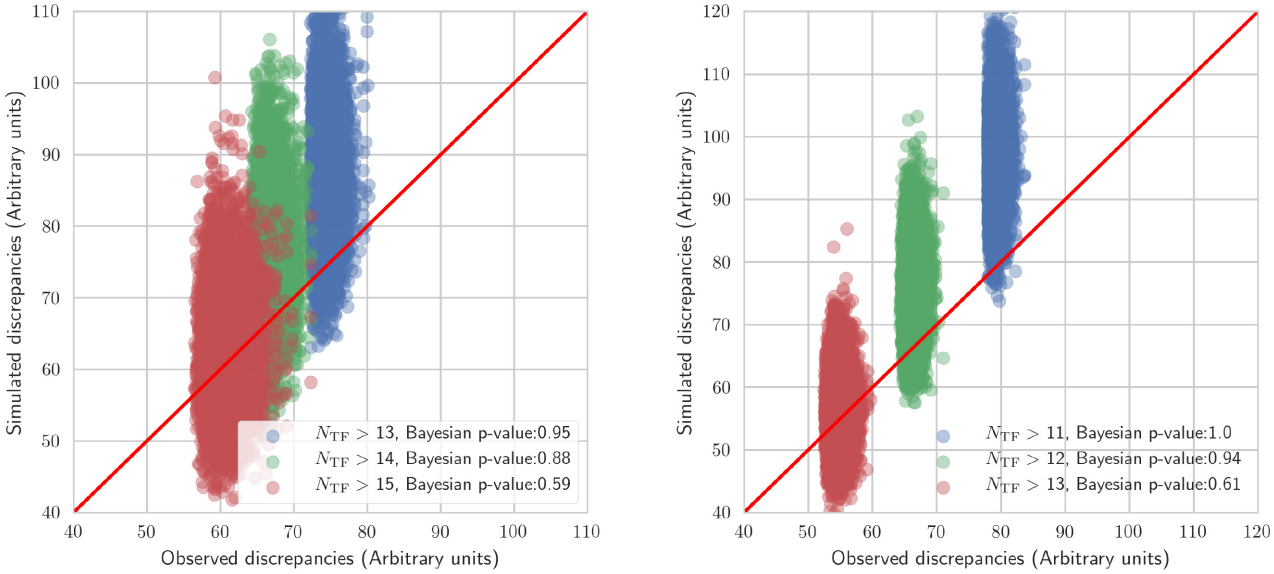

Figure 9. Discrepancy plot for the Bayesian linear model (equation 11) based on errors estimated using method H for $N_{TF} > 13$, 14, and 15 (left) and using method M for $N_{TF} > 11$, 12, and 13 (right).

explore and validate our Bayesian models is based on partitioning our data by choosing different lower thresholds for N, the number of measurements per galaxy.

As seen in Fig. 3, TF distance errors estimated using the robust methods H and M grow in a roughly linear fashion with distance and seem to be randomly distributed around this trend line. For this reason we try out simple linear and quadratic Bayesian models in order to be able to predict the value of missing distance errors. For this, we use the EMCEEaffine invariant Markov chain Monte Carlo (MCMC) ensemble sampler (Foreman-Mackey et al. 2013). Recently, `emcee` has been proved to be useful in obtaining probabilistic estimations for photometric redshifts (Speagle & Eisenstein 2017a,b). Since we want to be able to predict non-reported errors, our model selection is based on posterior predictive checks, i.e. we rely on models that can create synthetic data sets similar to the original data set (Gelman et al. 1996). This allows us to reproduce the original variance of the error (Fig. 13, left). Many Bayesian analyses often do not use posterior predictive checks, like in the work of Zhang & Shields (2018) and Jesus et al. (2018), where they used `emcee` for posterior sampling, and instead using Bayesian and Akaike information criteria along with Bayes factors for model assessment, but without attempting to reproduce the original variance of the data. This is also the case in other Bayesian tools like LINMIX (Kelly 2007), which is widely used in astronomy for approximating unobserved data.

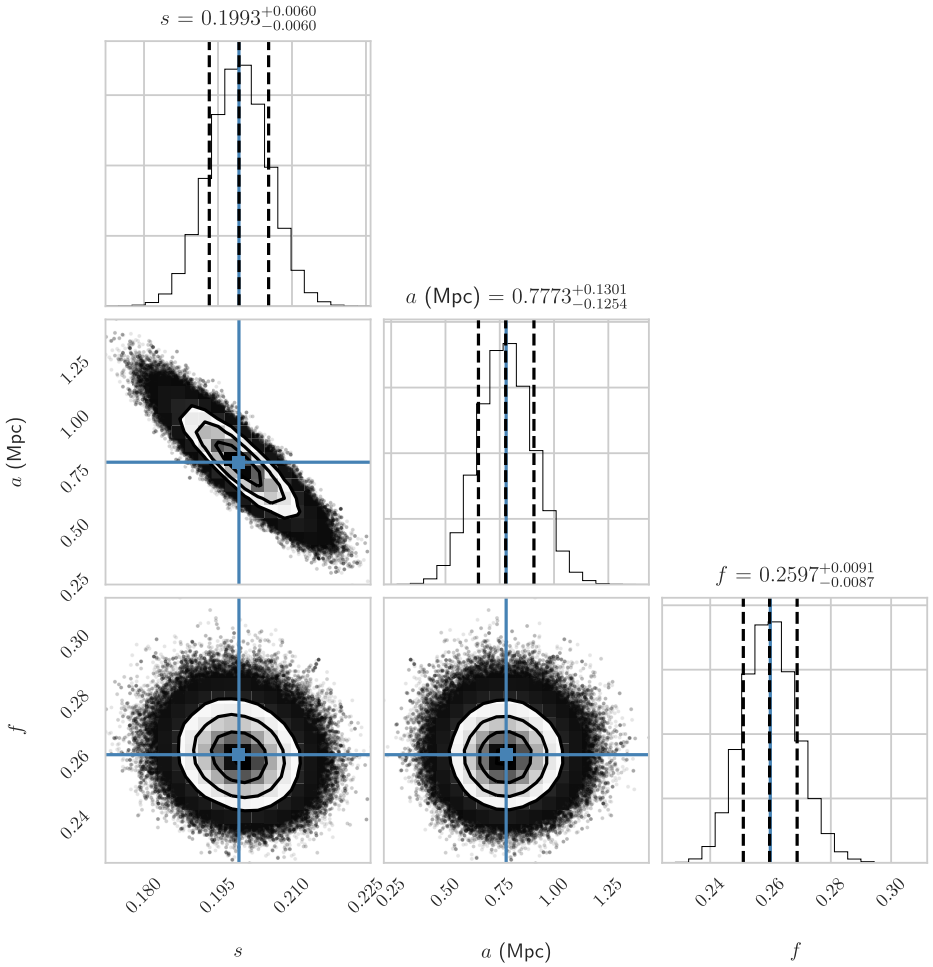

Figure 10. Corner plot showing the EMCEE sampling of the posterior probability distribution (equation 8) for the linear Bayesian model parameters $\boldsymbol{\theta} = (s, a, f)$ based on errors estimated using method H for galaxies with more than 15 TF distance measurements. The dashed lines indicate the 16th, 50th, and 84th percentile of the marginalized distribution of each parameter (shown at the top of each column), and the blue solid lines indicate the mean. This plot was made using the CORNER PYTHON module.

First, we assume that for any galaxy j the distance error $\sigma_{\mathrm{D}j}$ is a random normal variable, with variance $\sigma_{\sigma j}$ and mean $\hat{\sigma}_{\mathrm{D}j}$,

$$P(\sigma_{\mathrm{D}j}|\hat{\sigma}_{\mathrm{D}j}, \sigma_{\sigma j}) = \mathcal{N}(\hat{\sigma}_{\mathrm{D}j}, \sigma^2_{\sigma j}). \tag{6}$$

Our likelihood function is the joint probability that each of the $\sigma_{\mathrm{D}} = \{\sigma_{\mathrm{D}j}\}$ in the original data set of m galaxies is generated by the above probability,

$$P(\sigma_{\mathrm{D}}|\hat{\sigma}_{\mathrm{D}}, \sigma_{\sigma}) = \prod_{j}^{m} P(\sigma_{\mathrm{D}j}|\hat{\sigma}_{\mathrm{D}j}, \sigma_{\sigma j}). \tag{7}$$

We want to test the hypothesis mentioned above that all errors and their variances ($\hat{\sigma}_{\mathrm{D}} = \{\hat{\sigma}_{\mathrm{D}j}\}$, $\sigma_{\sigma} = \{\sigma_{\sigma j}\}$) can be estimated from a single model depending on the extragalactic distances $D_{\mathrm{G}} = \{D_{\mathrm{G}j}\}$ and a set of distance-independent parameters $\boldsymbol{\theta}$. Thus the likelihood can be expressed as

$$P(\sigma_{\mathrm{D}}|D_{\mathrm{G}}, \boldsymbol{\theta}) = \prod_{j}^{m} P(\sigma_{\mathrm{D}j}|D_{\mathrm{G}j}, \boldsymbol{\theta}).$$

Following Bayes' theorem we can compute the posterior probability up to a constant,

$$P(\boldsymbol{\theta}|D_{\mathrm{G}}, \sigma_{\mathrm{D}}) \propto P(\boldsymbol{\theta})P(\sigma_{\mathrm{D}}|D_{\mathrm{G}}, \boldsymbol{\theta}). \tag{8}$$

Because of the simplicity of the models used here, we will only use reasonably conservative (flat) priors on all model parameters, which are described in the next subsection.

A common feature across our models is that $\sigma_{\sigma} = f\hat{\sigma}_{\mathrm{D}}$, where the error variance scale factor f is one of the parameters in $\boldsymbol{\theta}$. This model choice is supported by Fig. 13 (left), which shows a roughly linear correlation between estimated errors and their variances. On the other hand, our models will differ by the proposed functional forms of $\hat{\sigma}_{\mathrm{D}}(D_{\mathrm{G}}, \boldsymbol{\theta})$.

We obtain computationally credible samplings of the posterior probability (equation 8) by removing the burn-in steps of the random walk according to the autocorrelation time. We can then create synthetic data sets by drawing a parameter sample $\boldsymbol{\theta}_k$ from the posterior and using it to draw from the likelihood to create a new data set, i.e. drawing new $\sigma_{\mathrm{D}j}$ from the probability distribution for all galaxies in the original data set using equation (6). We then assess the validity of the model by comparing synthetic data with the observed (i.e. original) data. This comparison is done by

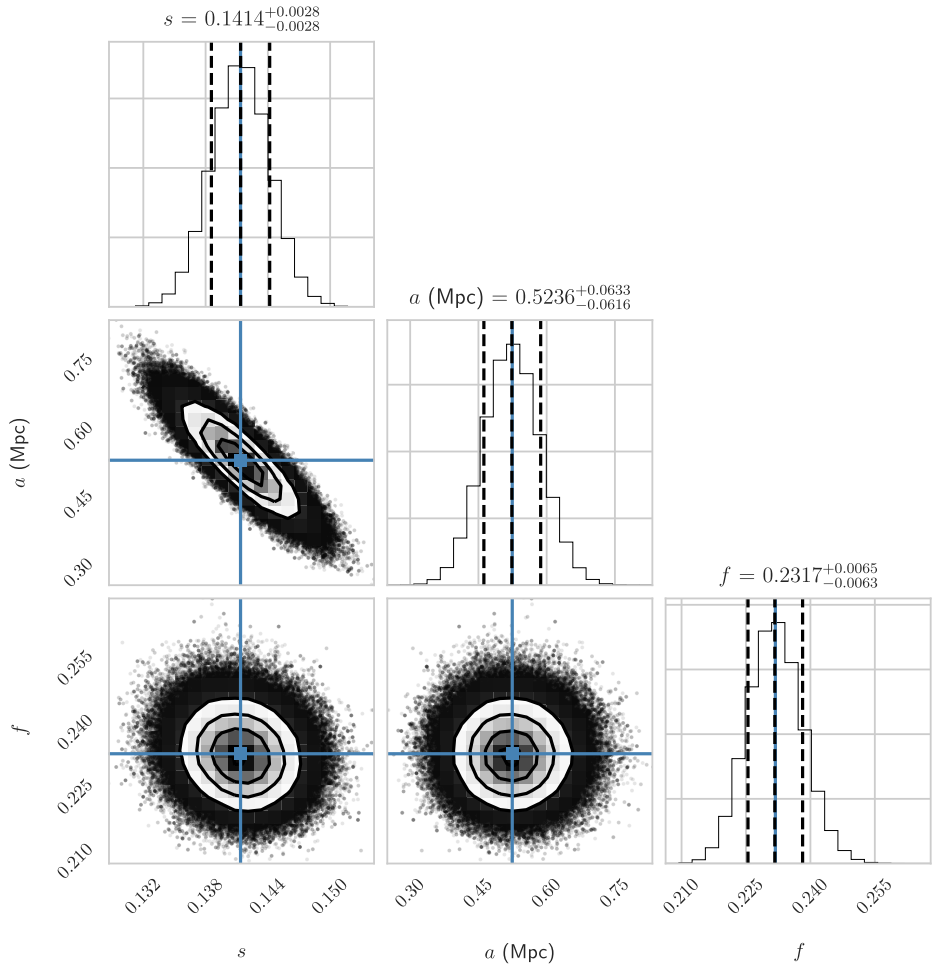

Figure 11. Corner plot showing the EMCEE sampling of the posterior probability distribution (equation 8) for the linear Bayesian model parameters $\boldsymbol{\theta} = (s, a, f)$ based on errors estimated using method M for galaxies with more than 13 TF distance measurements. The dashed lines indicate the 16th, 50th, and 84th percentile of the marginalized distribution of each parameter (shown at the top of each column), and the blue solid lines indicate the mean. This plot was made using the CORNER PYTHON module.

using a discrepancy measure $\mathcal{D}(\sigma_{\rm D}|\boldsymbol{\theta}_k)$ between data and model-derived expected values for the same data $e = \{e_j(\boldsymbol{\theta}_k)\}$, where $\boldsymbol{\theta}_k$ is drawn from the posterior distribution and $\sigma_{\rm D}$ can be the observed errors or the model-generated synthetic errors. The discrepancy can be calculated using a statistic like χ^2 (de la Horra 2008; de la Horra & Teresa Rodriguez-Bernal 2012), but here we will work with the Freeman–Tukey discrepancy since it is weight independent (Brooks, Catchpole & Morgan 2000; Bishop, Fienberg & Holland 2007),

$$\mathcal{D}(\sigma_{\rm D}|\boldsymbol{\theta}_k) = \sum_j^m \left(\sqrt{\sigma_{{\rm D}j}} - \sqrt{e_j(\boldsymbol{\theta}_k)}\right)^2.$$

For each parameter draw k, it is possible to compare the simulated discrepancy with the observed discrepancy. If the model is representative of the data, then for many parameter draws, the simulated and observed discrepancies should be similar. We can then calculate a Bayesian 'p-value' as the ratio of 'draws when the observed discrepancies are larger than the synthetic discrepancies' to 'total draws'. If this Bayesian p-value is too close to 0 or to 1 we can reject the model, otherwise it is generating synthetic data that are similar to the original data. This is better visualized using a discrepancy plot, where for each draw k, a synthetic discrepancy is paired with its corresponding observed discrepancy. If the discrepancy points are roughly equally distributed about the $\mathcal{D}_{\rm obs} = \mathcal{D}_{\rm sim}$ line, then we cannot reject the model. As mentioned above, we expect that galaxies with the largest number of measurements are sampling more completely the 'true' distribution of the distance. Therefore, we need to find the minimum number of measurements per galaxy for which the Bayesian p-value shows an agreement between on the partitioned data set and the model predictions.

4.1 Bayesian quadrature model

Our first model is based on the hypothesis that there are distinct systematic and random contributions to the distance measurement error, both of which are normally distributed. For this reason they are added in quadrature,

$$\sigma_{\rm D}^2 = \sigma_{\rm s}^2 + \sigma_{\rm r}^2. \tag{9}$$

Here $\sigma_{\rm r}$ is a random (constant) error and the systematic error is modelled allowing for scale factor (s) and zero setting (a) errors, i.e. $\sigma_{\rm s} = sD + a$, as Fig. 3 suggests. We set our prior to be symmetrical

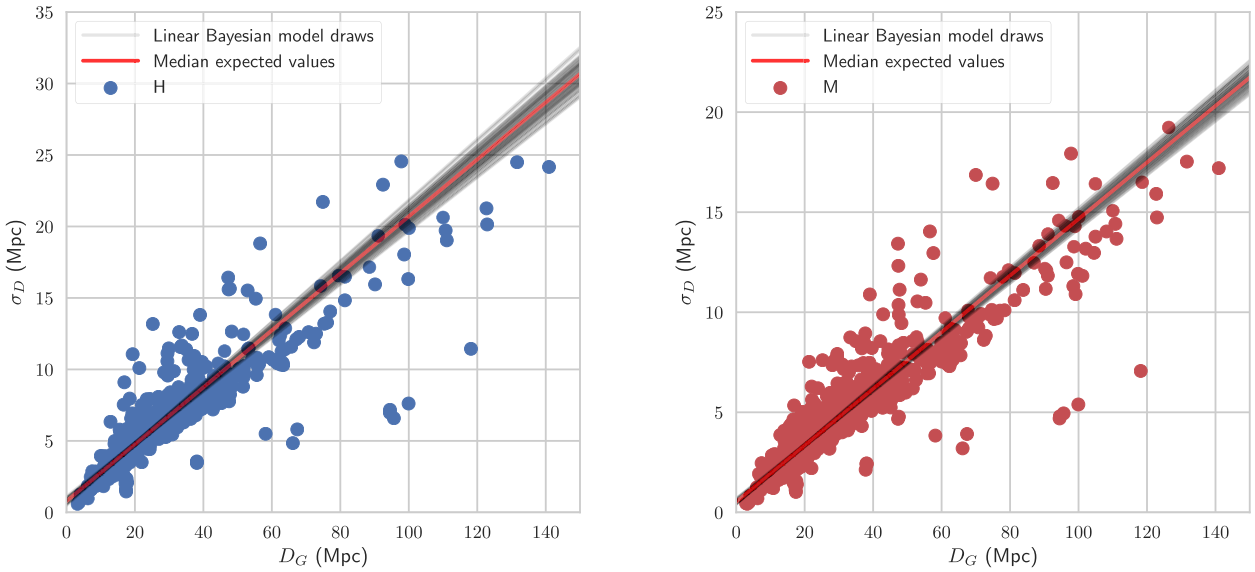

Figure 12. Projection of parameter set samples from the posterior probability distribution of the Bayesian linear model on to the σ_D versus D_G scatter plot for errors estimated using method H for galaxies with more than 15 TF distance measurements (left) and using method M for galaxies with more than 13 TF distance measurements (right).

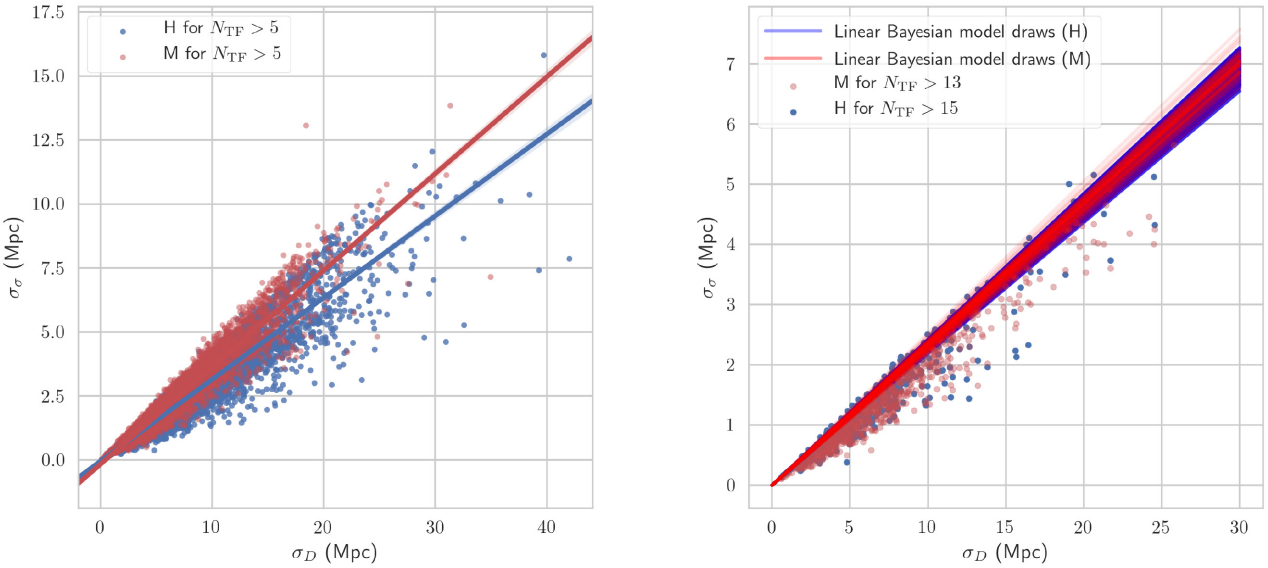

Figure 13. Variance of distance error estimates versus estimated extragalactic distance errors, showing linear regressions and confidence intervals computed using the SEABORN.REGPLOT PYTHON function (left), and showing a projection of parameter set samples from the posterior probability distribution of the Bayesian linear model (right) as determined by the H and M methods.

around $\sigma_r = 0$ in order to better visualize its behaviour near this point, so

$$P(s, a, \sigma_r, f) \propto \begin{cases} 1, & \text{if } 0 < s < 1 \text{ and} \\ & 0 < a < 10\,\text{Mpc and} \\ & -10 < \sigma_r < 10\,\text{Mpc and} \\ & 0 < f < 1, \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

We now use EMCEE to sample the posterior over the parameter set $\boldsymbol{\theta} = (s, \sigma_r, f, a)$ using 100 walkers and 20 000 steps ($\bar{t}_{\text{autocorr}} \lesssim 90$ steps). According to the discrepancy plot in Fig. 5 (left), this model is able to replicate method H errors for the 31 galaxies with $N > 25$ measurements (866 measurements in total). The corner plot showing the posterior sampling made by EMCEE is shown in Fig. 6, which shows the 16th, 50th, and 84th percentiles of the marginalized posterior distributions for the systematic scale factor s, the random error component σ_r, the error variance scale factor f, and the zero offset systematic error a. From the large variance in the marginalized posterior distribution for σ_r and a, we see that there is a significant degeneracy between those parameters. However, it should be noted that the marginalized posterior distribution of σ_r is symmetric around zero (because of its own degeneracy), while the distribution of a can only take positive values. The working

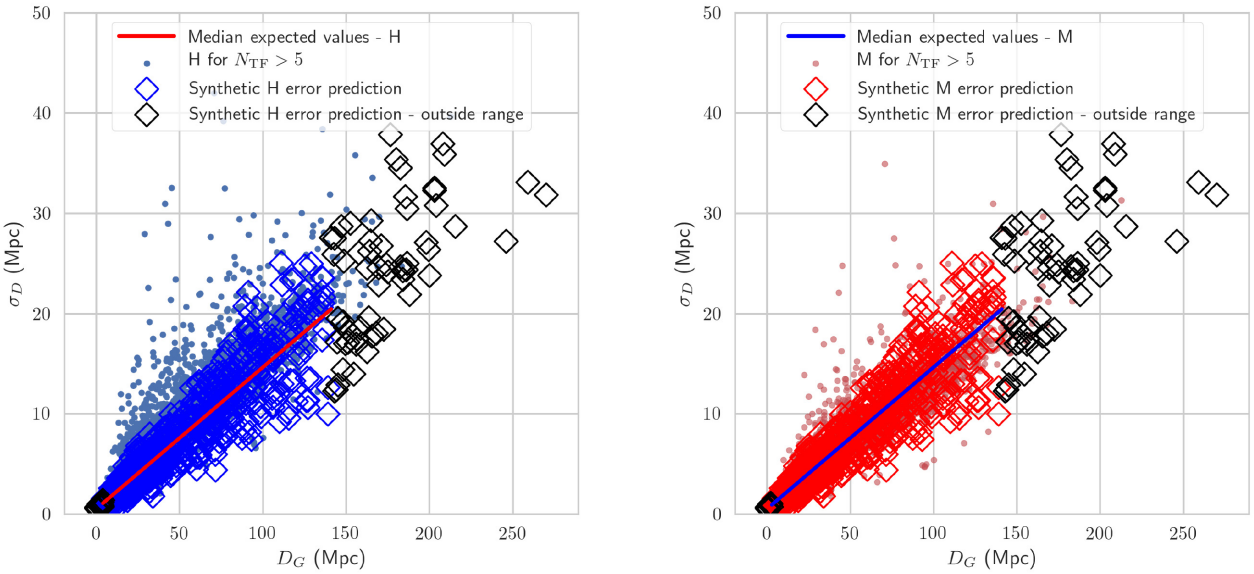

Figure 14. Synthetic H method (left) and M method (right) σ_D and their median expected values versus D_G for the 884 galaxies in NED-D for which no TF distance measurements report an error, generated using the corresponding Bayesian linear model. Predicted errors for galaxies outside of the working distance range of the model are plotted in black. H (left) and M (right) errors for galaxies with more than five TF measurements are also plotted for comparison.

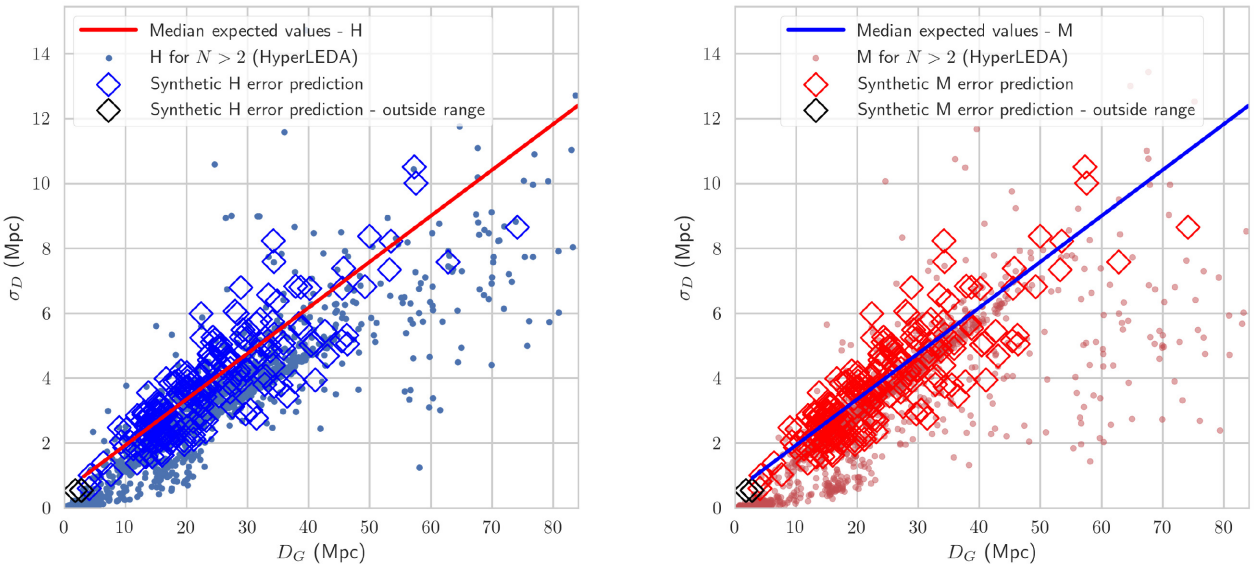

Figure 15. Synthetic H method (left) and M method (right) σ_D and their median expected values versus D_G for the 71 galaxies in HyperLEDA for which no TF distance measurements report an error, generated using the corresponding Bayesian linear model. Predicted errors for galaxies outside of the working distance range of the model are plotted in black. H errors for galaxies with more than two distance measurements are also plotted for comparison.

distance range and overall fitting of this model are shown in Fig. 8 (left), where method H errors corresponding to galaxies with more than 25 TF distance measurements are plotted along the expected values $e = \{e_j(\boldsymbol{\theta}_k)\}$ for parameter sets $\boldsymbol{\theta}_k$ drawn from the posterior probability distribution. Now we sample the posterior distribution for the Bayesian quadrature model with method M errors using EMCEE with 100 walkers and 20 000 steps s ($\bar{t}_{\rm autocorr} \lesssim 50$ steps). The discrepancy plot for method M errors in Fig. 5 (right) shows that the quadrature model also replicates method M errors, but for the 732 galaxies with more than 13 measurements (13 054 measurements in total). Fig. 7 shows that values for the random error component σ_r are so low that the model draws are almost indistinguishable from straight lines in Fig. 8 (right). Additionally, and just as for the quadrature model for H errors above, the symmetry of the marginalized posterior distribution of σ_r leads us to set this parameter to zero in our next (linear) model.

4.2 Bayesian linear model

In Section 4.1 above we conclude that we can ignore the random error component in equation (9) in order to work with a simpler, numerically stable, linear model that only considers a systematic

error with scale factor and zero setting error components,

$$\sigma_{\rm D} = \sigma_{\rm s} = sD + a. \tag{11}$$

We also update our prior considering that the quadratic model yielded lower values for the zero setting error a than previously considered in equation (10),

$$P(s, a, f) \propto \begin{cases} 1, & \text{if } 0 < s < 1 \text{ and} \\ & 0 < a < 2\,\text{Mpc and} \\ & 0 < f < 1, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

We use EMCEE to sample the posterior over $\boldsymbol{\theta} = (s, a, f)$ using 100 walkers and 10 000 steps ($\bar{t}_{\rm autocorr} < 50$ steps) for the linear Bayesian model applied to H errors. The discrepancy plot (Fig. 9, left) shows a significant improvement over the quadratic model, as it shows an acceptable Bayesian p-value for the 477 galaxies with $N > 15$ measurements (9361 in total), whereas the quadratic model replicated errors only for galaxies with $N > 25$ measurements. Fig. 10 shows the 16th, 50th, and 84th percentiles of the marginalized posterior distributions for the systematic scale factor s, the error variance scale factor f, and the zero offset systematic error a for the linear Bayesian model using H errors for galaxies with more than 15 measurements. We sample the posterior for the linear model applied to M errors using EMCEE with 100 walkers and 10 000 steps ($\bar{t}_{\rm autocorr} < 50$ steps). Fig. 9 (right) shows the corresponding discrepancy plot, which does not show a significant improvement of the linear over the quadratic model for M errors, as it also works for galaxies with $N > 13$ measurements. This happens because the sampling of the posterior for the quadratic model (Fig. 7) does not show a degeneracy between $\sigma_{\rm r}$ and a, and also because the marginalized posterior distribution for $\sigma_{\rm r}$ is a near-zero-centred distribution with a variance of ∼0.2 Mpc. The 16th, 50th, and 84th percentiles of the marginalized posterior distributions for the systematic scale factor s, the error variance scale factor f, and the zero offset systematic error a according the linear model for M errors are shown in Fig. 11.

4.3 Predictions for missing errors

Our linear Bayesian model is able to predict the intrinsic variance of TF H and M distance errors in NED-D by considering systematic zero setting and scale factor error components. The lower limit of distance measurements for which the model works for H and M errors is 15 and 13, respectively.[1] Fig. 12 shows the linear model draws for H and M errors, for which the working range is approximately $D_{\rm G} \in [3, 140]$ Mpc. We also show in Fig. 13 that the model draws for f the scale parameter for the variance of $\sigma_{\rm D}$ fit the bootstrap variance of H and M errors well, which means that our model choice for the variance of the error (σ_σ) was the right one.

Now, galaxies for which the models shown above work are not intrinsically different to other galaxies, as long as they are within the same distance range. Thus, we use the posterior predictive distribution of the linear Bayesian model for predicting H and M errors for the 884 galaxies in NED-D for which all TF measurements lack a reported error. Fig. 14 shows synthetic errors generated from the posterior predictive distribution for the $\sigma_{\rm D}$ linear model, along with the expected values of $\sigma_{\rm D}$ using the median of the posterior probability distribution in equation (8), and the $D_{\rm G}$ versus $\sigma_{\rm D}$ points for galaxies with more than five TF distance measurements (for contrast) for methods H and M, respectively. The median expected values are only drawn for points within the predictive range of each model, and synthetic predicted errors for galaxies outside of this range are plotted in black. The distance was calculated using the median of the reported distances whenever there was more than one TF distance measurement.

The HyperLEDA catalogue has distance measurements for 4224 galaxies, of which 1064 galaxies have reported measurements without errors. Of these galaxies with unreported distance errors, 203 report measurements obtained with the TF method. We create synthetic errors for these using our Bayesian predictive models for H and M TF errors. Fig. 15 (left) shows that predicted H errors are somewhat higher than those estimated for HyperLEDA, although acceptably within the range. Fig. 15 (right) shows that predicted M errors are even closer to the HyperLEDA M error estimates. This outstanding result is an independent validation of our linear Bayesian model for predicting TF distance errors, and its capacity to estimate systematic effects of the TF distance determination method.

This predictive model may work for other distance determination methods, but a cursory overview of methods that require error prediction due to missing errors [e.g. tip of the red giant branch (TRGB), colour–magnitude diagram (CMD), eclipsing binary, red clump, planetary nebulae luminosity function (PNLF), Sunyaev–Zeldovich (SZ) effect, brightest stars, horizontal branch in NED-D] suggests that such attempts need to be evaluated in a case-by-case basis. For instance, in NED-D, FP measurements are by far the most numerous (∼130 000 galaxies), but only 28 of those have more than three FP distance measurements. We attempted to create a model similar to what we did for TF, but we were only able to find a working predictive model (i.e. yielding a good Bayesian p-value) for the 16 galaxies with more than four distance measurements. The comparatively low number of galaxies for which this model works makes us wary of predicting FP errors, therefore we do not report these results.

5 CONCLUSIONS

We propose methods for robustly estimating the uncertainty in extragalactic distances in multimeasurement, multimethod catalogues. First we propose to report 16th, 50th, and 84th percentiles of the bootstrap-sampled distance distribution for each galaxy. We also propose the use of the half-distance between the 84th and 16th percentiles (method H), and the median absolute deviation (method M) if the bootstrap-sampled distance distribution for each galaxy as straightforward measures of the uncertainty in extragalactic distances. Method H gives errors that faithfully measure the variance of the distance probability distribution, whereas traditional frequentist propagation-of-error methods fail to match this variance measure. On the other hand, method M should be used whenever a specific application requires to ignore outdated or possibly wrong outliers.

We produce error data tables using the robust (H, M) and frequentist (P, Q) methods for NED-D, HyperLEDA, and Cosmicflows-3, along with the 16th, 50th, and 84th percentiles of the bootstrap-sampled distance distribution for each galaxy in those catalogues. These tables can be found in the repository for this paper, located at `http://github.com/saint-germain/errorprediction`. A description and analysis for each catalogue can be

[1] Our model validation has also worked with the two 2017 versions of the NED-D extragalactic distance catalogue, albeit with different thresholds for the number of measurements per galaxy.

found in the Appendix. We consider that these error tables should be a fundamental tool for future precision cosmology, catalogue-wide studies, as it should be possible to quote errors according to the method that the reader considers most relevant for specific applications.

We create a Bayesian predictive model for TF distance errors in the NED-D catalogue based on a Bayesian analysis of the systematic and random components in distance errors. We perform a posterior predictive check in the form of the computation of a Bayesian p-value based on simulated versus observed discrepancies measured with the Freeman–Tukey statistic. Thus we create models that can reproduce the intrinsic variance of distance errors along with systematic zero-setting and scale factor components from the posterior predictive distribution of the models, using NED-D estimated H and M errors.

We use these models to predict H and M errors for 884 galaxies in NED-D that report TF distance measurements but do not report measurement errors. Our predictive models are independently validated against the HyperLEDA catalogue by the agreement between our pre-computed H and M errors and our predictions for 203 galaxies in HyperLEDA with non-reported TF errors. Similar Bayesian predictive methods can be set up for other distance determination methods but with caveats, as model validation works better for methods for which there are many galaxies with a high number of distance measurements.

Finally, we want to advocate for the widespread use of discrepancy plots and their derived Bayesian p-values for Bayesian model checking in astronomy, as inference is based on the model's ability to reproduce the original distribution of the data and not only on a relative comparison to other models.

ACKNOWLEDGEMENTS

The authors would like to thank O. L. Ramírez-Suárez and J. E. Forero-Romero for their valuable input during the early stages of this work. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

Barris B., Tonry J., 2004, ApJ, 613, L21
Bishop Y. M. M., Fienberg S. E., Holland P. W., 2007, Discrete Multivariate Analysis: Theory and Practice. Springer-Verlag, New York
Brooks S. P., Catchpole E. A., Morgan B. J. T., 2000, Stat. Sci., 15, 357
Brugger R. M., 1969, Am. Stat., 23, 32
Chambert T., Rotella J. J., Higgs M. D., 2014, Ecol. Evolution, 4, 1389
Chaparro Molano G., Restrepo Gaitán O. A., Cuervo Marulanda J. C., Torres Arzayus S. A., 2018, Rev. Mex. Astron. Astrofis. Conf. Ser., 50, 63
Courtois H. M., Hoffman Y., Tully R. B., Gottloeber S., 2012, ApJ, 744, 43
de la Horra J., 2008, Commun. Stat. – Theory Methods, 37, 1412
de la Horra J., Teresa Rodriguez-Bernal M., 2012, SORT (Stat. Operations Res. Trans.), 36, 69
Dhawan S., Jha S. W., Leibundgut B., 2018, A&A, 609, A72
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Freedman W. L., Madore B. F., 2010, ARA&A, 48, 673
Freedman W. L. et al., 2001, ApJ, 553, 47
Gelman A., 2003, Int. Stat. Rev., 71, 369
Gelman A., Meng X.-L., Stern H., 1996, Stat. Sinica, 6, 733
Humphreys E. M. L., Reid M. J., Moran J. M., Greenhill L. J., Argon A. L., 2013, ApJ, 775, 13
Jarrett T. H., Chester T., Cutri R., Schneider S., Skrutskie M., Huchra J. P., 2000, AJ, 119, 2498
Javanmardi B., Kroupa P., 2017, A&A, 597, A120
Jesus J. F., Gregório T. M., Andrade-Oliveira F., Valentim R., Matos C. A. O., 2018, MNRAS, 477, 2867
Kelly B. C., 2007, ApJ, 665, 1489
Kourkchi E., Tully R. B., 2017, ApJ, 843, 16
Ma Y.-Z., Taylor J. E., Scott D., 2013, MNRAS, 436, 2029
McClure M. L., Dyer C. C., 2007, New Astron., 12, 533
Makarov D., Prugniel P., Terekhova N., Courtois H., Vauglin I., 2014, A&A, 570, A13
Mazzarella J. M., Team N., 2007, in Shaw R., Hill F., Bell D., eds, ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI. Astron. Soc. Pac., San Francisco, p. 153
Mould J., Sakai S., 2008, ApJ, 686, L75
Nasonova O. G., Karachentsev I. D., 2011, Astrophysics, 54, 1
Obreschkow D., Meyer M., 2013, ApJ, 777, 140
Riess A. G. et al., 2016, ApJ, 826, 56
Roman J., Trujillo I., 2017, MNRAS, 468, 703
Rubin D. et al., 2015, ApJ, 813, 137
Said K., Kraan-Korteweg R. C., Staveley-Smith L., Williams W. L., Jarrett T. H., Springob C. M., 2016, MNRAS, 457, 2366
Sorce J. G. et al., 2013, ApJ, 765, 94
Sorce J. G., Courtois H. M., Gottlober S., Hoffman Y., Tully R. B., 2014, MNRAS, 437, 3586
Speagle J. S., Eisenstein D. J., 2017a, MNRAS, 469, 1186
Speagle J. S., Eisenstein D. J., 2017b, MNRAS, 469, 1205
Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2007, ApJS, 172, 599
Springob C. M. et al., 2014, MNRAS, 445, 2677
Steer I. et al., 2017, AJ, 153, 37
Torres S., Cuervo J. C., 2018, Tecciencia, 24, 53
Tully R. B., Fisher J. R., 1977, A&A, 54, 661
Tully R. B., Pierce M. J., 2000, ApJ, 533, 744
Tully R. B., Courtois H. M., Sorce J. G., 2016, AJ, 152, 50
Watkins R., Feldman H. A., 2015, MNRAS, 450, 1868
White D. J., Daw E. J., Dhillon V. S., 2011, Classical Quantum Gravity, 28, 085016
Zhang J., Shields M. D., 2018, Mech. Syst. Signal Processing, 98, 465

APPENDIX A: PRE-COMPUTED DISTANCE ERROR DATA TABLES

We estimated errors for the HyperLEDA, Cosmicflows-3, and NED-D redshift-independent extragalactic distance data bases using the methods described in Section 3 across all distance determination methods, only considering measurements with more than two reported errors. Our pre-computed distance error tables can be found in the repository for this paper at `http://github.com/saint-germain/errorprediction`. The fields included for each catalogue are:

(i) `meas` – Number of distance measurements.
(ii) `D (Mpc)` – This is the median of the posterior distribution of the corresponding extragalactic distance.
(iii) `Dmin (Mpc)` – This is the 16th percentile of the posterior distribution of the corresponding extragalactic distance.
(iv) `Dmin (Mpc)` – This is the 84th percentile of the posterior distribution of the corresponding extragalactic distance.
(v) `H (Mpc)` – Error estimated using the H method (`Dmax-Dmin`)/2.
(vi) `M (Mpc)` – Error estimated using the M method.
(vii) `P (Mpc)` – Error estimated using the P method.
(viii) `Q (Mpc)` – Error estimated using the Q method.

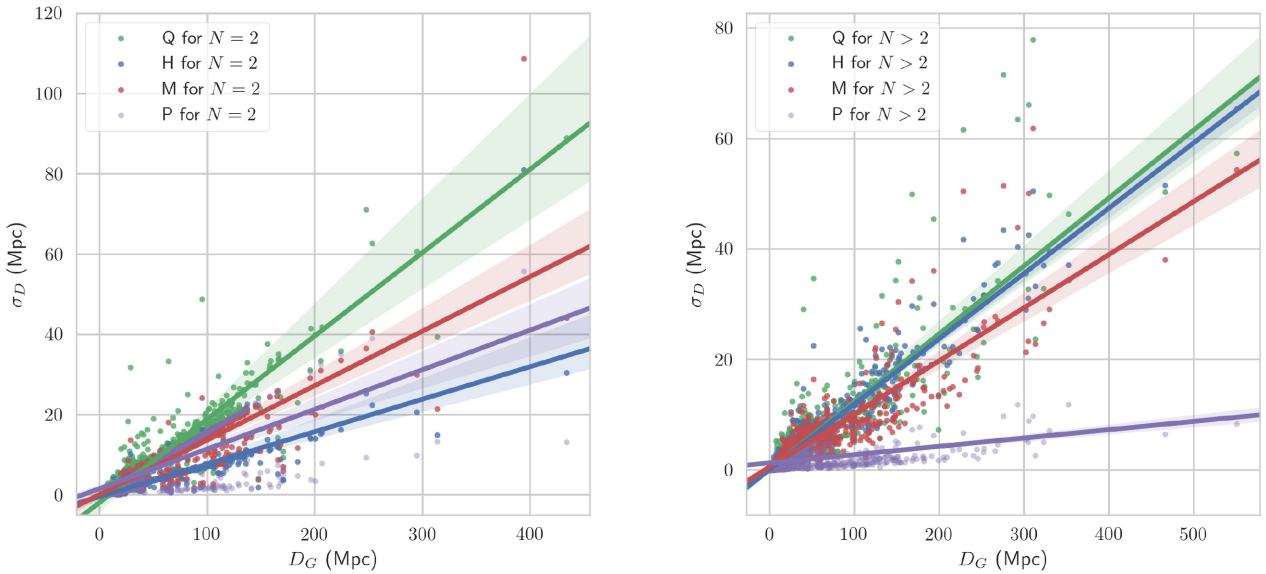

Figure A1. Estimated extragalactic distance errors versus median extragalactic distance for galaxies with $N = 2$ (left) and $N > 2$ (right) redshift-independent distance measurements in HyperLEDA according to the H, M, Q, and P error models, showing a linear regression and confidence intervals computed using the SEABORN.REGPLOT PYTHON function.

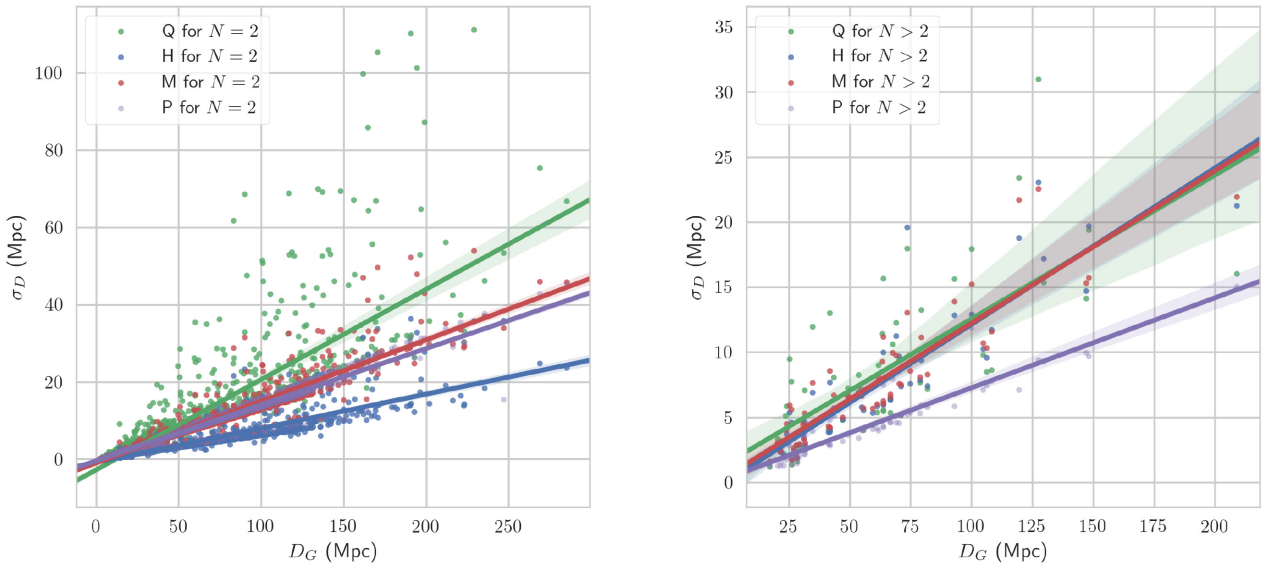

Figure A2. Estimated extragalactic distance errors versus median extragalactic distance for galaxies with $N = 2$ (left) and $N > 2$ (right) redshift-independent distance measurements in Cosmicflows-3 according to the H, M, Q, and P error models, showing a linear regression and confidence intervals computed using the SEABORN.REGPLOT PYTHON function.

A1 Estimation of errors for HyperLEDA

As expected from our analysis of TF errors in NED-D, errors calculated with methods P, Q, and M overpredict the error with respect to the H method for galaxies with a low number of distance measurements ($N = 2$), as shown in Fig. A1 (left). Fig. A1 (right) shows that for galaxies with a higher number of distance measurements, the P method significantly underpredicts the error with respect to the other methods. Even though the H and Q methods show a similar trend, the variance of Q errors around this trend is higher than for H methods. Errors obtained with method M are lower, due to the method's intrinsic robustness. These estimations are reported in the file called `hl_bootstrap_results.csv` in the repository. Special fields for this catalogue are:

(i) `objname` – Object name according to the HyperLEDA data base.
(ii) `j2000` – J2000 coordinates.

A2 Estimation of errors for Cosmicflows-3

The Extragalactic Distance Database (EDD) of Cosmicflows-3, which has the most up-to-date calibrated distance measurements using the TF, FP, and SNIa methods for more than 17 000 galaxies. We estimated errors for the approximately 10 per cent of them that have more than one reported distance, using the methods described

in Section 3. Fig. A2 shows that the P method, which is the suggested method in Tully et al. (2016), overpredicts errors with respect to the H method for galaxies with two distance measurements, as was the case for the errors of HyperLEDA. For galaxies with more than two distance measurements, Fig. A2 shows that the P method underpredicts the errors with respect to the H method. Even though the M, H, and Q methods show a similar trend with distance, the Q method has a significantly larger scatter around this trend. We compiled the estimated errors in a companion table to the Cosmicflows-3 EDD data base and in a similar format, in the file called `cf3_bootstrap_results.csv` in the repository for this work. Special fields for this catalogue are:

(i) `pgc` – Principal Galaxies Catalogue ID number.
(ii) `Name` – Object name according to the Cosmicflows-3 data base, where available.

A3 Estimation of errors for NED-D

The 2018 version of the NASA/IPAC extragalactic distance catalogue NED-D has ∼300 000 redshift-independent distance measurements with reported errors for ∼180 000 galaxies. We estimated the errors for the ∼16 000 galaxies with more than one distance measurement. The data base of errors for NED-D is in the file called `ned_bootstrap_results.csv`. The only special field for this catalogue is

(i) `Galaxy ID` – Object name according to the NED-D data base.

This paper has been typeset from a TEX/LATEX file prepared by the author.

List of astronomical key words (Updated on 2017 March)

This list is common to *Monthly Notices of the Royal Astronomical Society, Astronomy and Astrophysics*, and *The Astrophysical Journal*. In order to ease the search, the key words are subdivided into broad categories. No more than *six* subcategories altogether should be listed for a paper.

The subcategories in boldface containing the word 'individual' are intended for use with specific astronomical objects; these should never be used alone, but always in combination with the most common names for the astronomical objects in question. Note that each object counts as one subcategory within the allowed limit of six.

The parts of the key words in italics are for reference only and should be omitted when the keywords are entered on the manuscript.

General
editorials, notices
errata, addenda
extraterrestrial intelligence
history and philosophy of astronomy
miscellaneous
obituaries, biographies
publications, bibliography
sociology of astronomy
standards

Physical data and processes
acceleration of particles
accretion, accretion discs
asteroseismology
astrobiology
astrochemistry
astroparticle physics
atomic data
atomic processes
black hole physics
chaos
conduction
convection
dense matter
diffusion
dynamo
elementary particles
equation of state
gravitation
gravitational lensing: micro
gravitational lensing: strong
gravitational lensing: weak
gravitational waves
hydrodynamics
instabilities
line: formation
line: identification
line: profiles
magnetic fields
magnetic reconnection
(magnetohydrodynamics) MHD
masers
molecular data
molecular processes
neutrinos
nuclear reactions, nucleosynthesis, abundances
opacity
plasmas
polarization
radiation: dynamics
radiation mechanisms:general
radiation mechanisms: non-thermal
radiation mechanisms: thermal
radiative transfer
relativistic processes
scattering
shock waves
solid state: refractory
solid state: volatile
turbulence
waves

Astronomical instrumentation, methods and techniques
atmospheric effects
balloons
instrumentation: adaptive optics
instrumentation: detectors
instrumentation: high angular resolution
instrumentation: interferometers
instrumentation: miscellaneous
instrumentation: photometers
instrumentation: polarimeters
instrumentation: spectrographs
light pollution
methods: analytical
methods: data analysis
methods: laboratory: atomic
methods: laboratory: molecular
methods: laboratory: solid state
methods: miscellaneous
methods: numerical
methods: observational
methods: statistical
site testing
space vehicles
space vehicles: instruments
techniques: high angular resolution
techniques: image processing
techniques: imaging spectroscopy
techniques: interferometric
techniques: miscellaneous
techniques: photometric
techniques: polarimetric
techniques: radar astronomy
techniques: radial velocities
techniques: spectroscopic
telescopes

Astronomical data bases
astronomical data bases: miscellaneous
atlases
catalogues
surveys
virtual observatory tools

Astrometry and celestial mechanics
astrometry
celestial mechanics
eclipses
ephemerides
occultations
parallaxes
proper motions
reference systems
time

The Sun
Sun: abundances
Sun: activity
Sun: atmosphere
Sun: chromosphere
Sun: corona
Sun: coronal mass ejections (CMEs)
Sun: evolution
Sun: faculae, plages
Sun: filaments, prominences
Sun: flares
Sun: fundamental parameters
Sun: general
Sun: granulation
Sun: helioseismology
Sun: heliosphere
Sun: infrared
Sun: interior
Sun: magnetic fields
Sun: oscillations
Sun: particle emission
Sun: photosphere
Sun: radio radiation
Sun: rotation
(Sun:) solar–terrestrial relations
(Sun:) solar wind
(Sun:) sunspots
Sun: transition region
Sun: UV radiation
Sun: X-rays, gamma-rays

Planetary systems
comets: general

comets: individual: . . .
Earth
interplanetary medium
Kuiper belt: general

Kuiper belt objects: individual: . . .
meteorites, meteors, meteoroids
minor planets, asteroids: general

minor planets, asteroids: individual: . . .
Moon
Oort Cloud
planets and satellites: atmospheres
planets and satellites: aurorae
planets and satellites: composition
planets and satellites: detection
planets and satellites: dynamical evolution and stability
planets and satellites: formation
planets and satellites: fundamental parameters
planets and satellites: gaseous planets
planets and satellites: general

planets and satellites: individual: . . .
planets and satellites: interiors
planets and satellites: magnetic fields
planets and satellites: oceans
planets and satellites: physical evolution
planets and satellites: rings
planets and satellites: surfaces
planets and satellites: tectonics
planets and satellites: terrestrial planets
planet–disc interactions
planet–star interactions
protoplanetary discs
zodiacal dust

Stars
stars: abundances
stars: activity
stars: AGB and post-AGB
stars: atmospheres
(stars:) binaries *(including multiple)*: close
(stars:) binaries: eclipsing
(stars:) binaries: general
(stars:) binaries: spectroscopic
(stars:) binaries: symbiotic
(stars:) binaries: visual
stars: black holes
(stars:) blue stragglers
(stars:) brown dwarfs
stars: carbon
stars: chemically peculiar
stars: chromospheres
(stars:) circumstellar matter
stars: coronae
stars: distances
stars: dwarf novae
stars: early-type
stars: emission-line, Be
stars: evolution
stars: flare
stars: formation
stars: fundamental parameters
(stars:) gamma-ray burst: general
(stars:) **gamma-ray burst: individual: . . .**
stars: general
(stars:) Hertzsprung–Russell and colour–magnitude diagrams
stars: horizontal branch
stars: imaging
stars: individual: . . .
stars: interiors

stars: jets
stars: kinematics and dynamics
stars: late-type
stars: low-mass
stars: luminosity function, mass function
stars: magnetars
stars: magnetic field
stars: massive
stars: mass-loss
stars: neutron
(stars:) novae, cataclysmic variables
stars: oscillations *(including pulsations)*
stars: peculiar *(except chemically peculiar)*
(stars:) planetary systems
stars: Population II
stars: Population III
stars: pre-main-sequence
stars: protostars
(stars:) pulsars: general
(stars:) **pulsars: individual: . . .**
stars: rotation
stars: solar-type
(stars:) starspots
stars: statistics
(stars:) subdwarfs
(stars:) supergiants
(stars:) supernovae: general
(stars:) **supernovae: individual: . . .**
stars: variables: Cepheids
stars: variables: Scuti
stars: variables: general
stars: variables: RR Lyrae
stars: variables: S Doradus
stars: variables: T Tauri, Herbig Ae/Be
(stars:) white dwarfs
stars: winds, outflows
stars: Wolf–Rayet

Interstellar medium (ISM), nebulae
ISM: abundances
ISM: atoms
ISM: bubbles
ISM: clouds
(ISM:) cosmic rays
(ISM:) dust, extinction
ISM: evolution
ISM: general
(ISM:) HII regions
(ISM:) Herbig–Haro objects

ISM: individual objects: . . .
(except planetary nebulae)
ISM: jets and outflows
ISM: kinematics and dynamics
ISM: lines and bands
ISM: magnetic fields
ISM: molecules
(ISM:) photodissociation region (PDR)
(ISM:) planetary nebulae: general
(ISM:) **planetary nebulae: individual: . . .**
ISM: structure
ISM: supernova remnants

The Galaxy
Galaxy: abundances
Galaxy: bulge
Galaxy: centre
Galaxy: disc
Galaxy: evolution
Galaxy: formation
Galaxy: fundamental parameters
Galaxy: general
(Galaxy:) globular clusters: general
(Galaxy:) **globular clusters: individual: . . .**
Galaxy: halo
Galaxy: kinematics and dynamics
(Galaxy:) local interstellar matter
Galaxy: nucleus
(Galaxy:) open clusters and associations: general
(Galaxy:) **open clusters and associations: individual: . . .**
(Galaxy:) solar neighbourhood
Galaxy: stellar content
Galaxy: structure

Galaxies
galaxies: abundances
galaxies: active
(galaxies:) BL Lacertae objects: general
(galaxies:) **BL Lacertae objects: individual: . . .**
galaxies: bulges
galaxies: clusters: general

galaxies: clusters: individual: . . .
galaxies: clusters: intracluster medium
galaxies: distances and redshifts
galaxies: dwarf
galaxies: elliptical and lenticular, cD
galaxies: evolution
galaxies: formation
galaxies: fundamental parameters
galaxies: general
galaxies: groups: general

galaxies: groups: individual: . . .
galaxies: haloes
galaxies: high-redshift

galaxies: individual: . . .
galaxies: interactions
(galaxies:) intergalactic medium
galaxies: irregular
galaxies: ISM
galaxies: jets
galaxies: kinematics and dynamics
(galaxies:) Local Group
galaxies: luminosity function, mass function
(galaxies:) Magellanic Clouds
galaxies: magnetic fields
galaxies: nuclei
galaxies: peculiar
galaxies: photometry
(galaxies:) quasars: absorption lines
(galaxies:) quasars: emission lines
(galaxies:) quasars: general

(galaxies:) **quasars: individual: . . .**
(galaxies:) quasars: supermassive black holes
galaxies: Seyfert
galaxies: spiral
galaxies: starburst
galaxies: star clusters: general

galaxies: star clusters: individual: . . .
galaxies: star formation
galaxies: statistics
galaxies: stellar content
galaxies: structure

Cosmology
(cosmology:) cosmic background radiation
(cosmology:) cosmological parameters
(cosmology:) dark ages, reionization, first stars
(cosmology:) dark energy
(cosmology:) dark matter
(cosmology:) diffuse radiation
(cosmology:) distance scale
(cosmology:) early Universe
(cosmology:) inflation
(cosmology:) large-scale structure of Universe
cosmology: miscellaneous
cosmology: observations
(cosmology:) primordial nucleosynthesis
cosmology: theory

Resolved and unresolved sources as a function of wavelength
gamma-rays: diffuse background
gamma-rays: galaxies
gamma-rays: galaxies: clusters
gamma-rays: general
gamma-rays: ISM
gamma-rays: stars
infrared: diffuse background
infrared: galaxies
infrared: general
infrared: ISM
infrared: planetary systems
infrared: stars
radio continuum: galaxies
radio continuum: general
radio continuum: ISM
radio continuum: planetary systems
radio continuum: stars
radio continuum: transients
radio lines: galaxies
radio lines: general
radio lines: ISM
radio lines: planetary systems
radio lines: stars
submillimetre: diffuse background
submillimetre: galaxies
submillimetre: general
submillimetre: ISM
submillimetre: planetary systems
submillimetre: stars
ultraviolet: galaxies
ultraviolet: general
ultraviolet: ISM
ultraviolet: planetary systems
ultraviolet: stars
X-rays: binaries
X-rays: bursts
X-rays: diffuse background
X-rays: galaxies
X-rays: galaxies: clusters
X-rays: general
X-rays: individual: . . .
X-rays: ISM
X-rays: stars