RAA 2021 Vol. **21** No. **3**, 61(10pp) doi: 10.1088/1674-4527/21/3/61

http://www.raa-journal.org *http://iopscience.iop.org/raa*

***R**esearch in*
***A**stronomy and*
***A**strophysics*

Measurement of Hubble constant: were differences in secondary distance indicators apparent as early as the HST Key Project?

Rahul Kumar Thakur[1], Shashikant Gupta[2], Rahul Nigam[1] and PK Thiruvikraman[1]

[1] Birla Institute of Technology & Science, Pilani-Hyderabad Campus, India
[2] G D Goenka University, Gurgaon, Haryana, 122103, India; *shashikantgupta.astro@gmail.com*

Received 2020 May 1; accepted 2020 August 29

Abstract Different measurements of the Hubble constant (H_0) are not consistent, and a tension between the CMB based methods and cosmic distance ladder based methods has been observed. Measurements from various distance based methods are also inconsistent. To aggravate the problem, the same cosmological probe (Type Ia SNe for instance) calibrated through different methods also provides different values of H_0. We compare various distance ladder based methods through the already available unique data obtained from the Hubble Space Telescope (HST). Our analysis is based on parametric (t-test) as well as non-parametric statistical methods such as the Mann-Whitney U test and Kolmogorov-Smirnov test. Our results show that different methods provide different values of H_0 and the differences are statistically significant. The biases in the calibration would not account for these differences as the data have been taken from a single telescope with a common calibration scheme. The unknown physical effects or issues with the empirical relations of distance measurement from different probes could give rise to these differences.

Key words: Cosmology — Supernovae — Hubble constant

1 INTRODUCTION

Our understanding of the Universe increased exponentially during the last century. Initially, static models were suggested, but after Edwin Hubble's observations, it was established that the Universe is expanding (Hubble 1929; Freedman 2000; Nussbaumer & Bieri 2009). These observations can be summarised in the form of a graph between the distance to several galaxies and their recessional velocities. The graph, known as the Hubble diagram, results in a straight line whose slope is called the Hubble parameter (van den Bergh 1995). The Hubble parameter is dynamic in nature, and its present value is referred to as the Hubble Constant (H_0). It is among the most fundamental parameters of standard cosmology, as it is a measure of the expansion rate and age of the Universe. The Hubble constant also determines the critical density, $\rho_c = 3H_0/8\pi G$, essential for flat geometry of the Universe. Many other cosmological parameters such as physical properties of galaxies and quasars, growth of large scale structures, etc., depend on the numerical value of the Hubble constant. Thus, obtaining an accurate value of H_0 is of great importance to present day cosmology.

Various fundamentally different methods have been employed to determine an accurate value of the Hubble constant including Cosmic Microwave Background Radiation (CMBR), gravitational waves and the methods based on cosmic distance ladder such as type Ia supernovae (SNe Ia) (these methods are discussed in Sect. 2). The latest CMBR observations by the Planck satellite (Hinshaw et al. 2013; Planck Collaboration et al. 2018) along with the ΛCDM cosmological model predicted $H_0 = 66.93 \pm 0.62\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. It differs by around 3σ or more with that measured with the distance ladder based methods (Riess et al. 2016, 2019). Several solutions have been suggested, including a dark component in the early Universe (Slatyer & Wu 2018; Sakstein & Trodden 2020; Vagnozzi 2020), to resolve this tension. It is also a matter of concern that various methods based on the cosmic distance ladder also provide different values of H_0. Moreover, inconsistency has been observed in the data obtained from the same secondary distance indicator (SNe Ia) calibrated through different probes (Cepheids vs. Tip of Red Giant Branch (TRGB)) (Freedman et al. 2019, 2020). We wish to explore the statistical significance of the differences in these distance ladder based measurements. Significant differences would indicate either a lack of

understanding of physical concepts inherent in these probes, or presence of systematic effects, or issues related to the calibration of telescopes. This paper is organised as follows: we discuss various distance ladder based methods to measure the Hubble constant in Section 2. The data and statistical tools used for our analysis are addressed in Sections 3 and 4, respectively. The results and conclusions are presented in Sections 5 and 6, respectively.

2 MEASUREMENT OF HUBBLE CONSTANT AND COSMIC DISTANCE LADDER

Determination of the Hubble constant requires measurement of the distances to the galaxies and their recessional velocities up to sufficiently large scales. The progress in accurate determination of H_0 has been slow due to various issues in distance measurement methods. A distance indicator, in principle, should fulfill the following basic criteria: (i) it should be bright enough to be detectable at cosmological distances, (ii) the physics of the distance indicator should be clear, (iii) empirical relations among various quantities which are utilized for distance measurement should be free of systematic effects and (iv) statistically significant samples of such objects should be available.

The period luminosity (P-L) relation of young Cepheid stars found in spiral galaxies is a promising primary distance indicator as the physics of this correlation is well understood (Lanoix et al. 1999). However, they are found in dusty regions and hence the observed P-L relation often requires correction against scattering, absorption, reddening and extinction, leading to systematic errors. The P-L relation may also depend on the chemical composition (Freedman & Madore 1990), which is difficult to model. It is also difficult to resolve the Cepheid stars in distant galaxies; moreover, they are often observed mainly in spiral galaxies. Thus Cepheid variables alone cannot be relied on to estimate H_0, and we need to use secondary distance indicators such as SNe Ia and the Tully-Fisher (T-F) relation. However, accurate measurements of extragalactic distances have always been challenging; often the uncertainties in measurement are underestimated and systematic errors dominate. Even today, identifying and reducing the sources of systematic errors in distance measurement are challenging tasks. Measurement of H_0 up to 1% accuracy is a difficult goal. However, as per the results of the Hubble Space Telescope (HST) Key Project (Freedman et al. 2001), an accuracy of 10% was achieved in 2003; and 2.4% accuracy was claimed recently (Riess et al. 2016). The following different methods were applied as secondary indicators: (i) SNe Ia, (ii) T-F relation, (iii) surface brightness fluctuations, (iv) type II supernovae (SNe II) and (v) the fundamental plane relation of elliptical galaxies. We review these methods in the next section. The current status of these methods is available in Verde et al. (2019) and references therein.

2.1 Type Ia Supernovae

SNe Ia are believed to arise from the explosion of a carbon-oxygen white dwarf (Nomoto et al. 1997). Their peak luminosity can outshine the entire host galaxy, which makes them observable at cosmological distances (Kowal 1968; Sandage & Tammann 1982). The peak luminosities of all SNe Ia are found to be in a narrow range. Moreover, the decline rate of their brightness is strongly correlated with the peak luminosity (Phillips 1993). Using this correlation, the absolute magnitude of an individual SN and hence its distance can be measured with high precision (less than 10%).

Current Status: So far, SNe Ia are the most promising cosmological distance indicators. Unfortunately, the exact mechanism of an SN Ia explosion has not yet been well understood and recently, subclasses within the SN Ia family have also been explored (Umeda et al. 1999); some peculiar SNe Ia follow a special category recently termed SNe Iax (Jha 2017). Confidence in this empirical method will be strengthened once we gain a theoretical understanding of the explosion process.

2.2 The Tully-Fisher Relation

An empirical relation, known as the T-F relation, exists between the total luminosity and rotational speed of spiral galaxies (Tully & Fisher 1977; Tully & Pierce 2000). The correlation becomes tighter at longer wavelengths; especially at I band the rms scatter is only 0.4 mag (Willick et al. 1997). The T-F relation reflects the fact that massive and hence luminous galaxies rotate more rapidly and can be exploited to measure the distances to spiral galaxies (Giovanelli et al. 1997). Earlier measurements of H_0 based on the T-F relation provided a large range of values indicating the presence of systematic effects. Sakai et al. (2000) applied $BVRIH$ photometry calibrations to several galaxy clusters to measure H_0 and investigate the possible systematic uncertainties including the metallicity dependence. Using the new P-L calibration and adopting the metallicity corrections, Freedman et al. (2001) obtained the H_0 value, which is not different from that of Sakai et al. (2000).

Current Status: Recently, the relation has been measured for hundreds of galaxies and amounts to around 15% uncertainty in distance measurement (McGaugh 2005, 2012).

2.3 Fundamental Plane Relation

A correlation between luminosity and central velocity of elliptical galaxies was first discovered by Faber & Jackson (1976); Jorgensen et al. (1996). It is similar to the T-F relation of spiral galaxies; however, it has a lot of scatter. A similar noisy correlation also exists between the effective radius and mean surface brightness of elliptical galaxies (Djorgovski & Davis 1987). Both of these correlations are now understood as the projections of a plane, known as the fundamental plane of elliptical galaxies, defined as $r_e \propto \sigma^{\alpha} I_e^{\beta}$, where r_e is the effective radius, I_e is the average surface brightness within r_e and σ is the stellar velocity dispersion (Schaeffer et al. 1993; Kelson et al. 2000). The scatter in the fundamental plane relation is much smaller than its projections and thus can be utilized to measure the distance to elliptical galaxies in various clusters (Seljak 2002).

Current Status: Although the scatter in the fundamental plane relation is small, it suffers from systematic effects. For instance, early type brightest cluster galaxies exhibit a steeper size-luminosity relation which would indicate curvature in the plane (Bernardi et al. 2007). Many improvements have been suggested (see Joachimi et al. 2015; Saulder et al. 2019 for a recent review). Saulder et al. (2019) calculated reliable galaxy distances by removing the biases and selection effects.

2.4 Surface Brightness Fluctuations

Since each galaxy contains a finite number of stars, which are not distributed uniformly, the number of stars in a small patch of a galaxy varies from point to point (Lauer et al. 1998). This leads to fluctuations in the surface brightness of the galaxy. These fluctuations smooth out with distance since the resolution of stars within the galaxies depends on distance (Tonry & Schneider 1988; Tonry et al. 2000). This technique has been effectively applied as a distance indicator in the HST Key Project. However, HST Advanced Camera for Surveys exhibits geometrical distortion due to its off-axis position (Mei et al. 2005).

Current Status: This method is appropriate for elliptical galaxies because they have fairly consistent stellar populations or to spirals with prominent bulges (Jensen et al. 2001; Liu et al. 2002). Corrections due to variations in metallicity and age of galaxies are often required. Stellar population modelling and Cepheid variables are used for calibration of the method (Verde et al. 2019). Recently, the method has been employed to calibrate SNe Ia and hence measure the Hubble constant (Khetan et al. 2020).

2.5 Type II Supernovae

SNe II originate from core collapse of massive stars and are fainter than SNe Ia. Often they are observed in spiral arms of galaxies and HI clouds in the interstellar medium, but rarely in elliptical galaxies. Although they are not standard candles, their distance can be measured by combining the spectra of their expanding photosphere and photometric observations of angular size. The technique, known as the Baade-Wesselink method, has been utilized by Kirshner et al. (1994) to reliably determine the distances to various SNe II and hence H_0. Freedman et al. (2001) found the distances to SN 1970G, SN1987A and 1989L using Cepheids in their host galaxies. The distances measured by the above two methods agree quite well.

Current Status: Many new theoretical and empirical methods to measure distances to SNe II have been proposed in the last decade, e.g., the standard candle method (SCM) based on the correlation between the luminosity and the expansion velocities (Hamuy & Pinto 2002; de Jaeger et al. 2020b), the photospheric magnitude method (PMM) which is a generalisation of the SCM (Rodríguez et al. 2014, 2019), and the photometric colour method (PCM) which relies on the relation between the luminosity and the slope of the plateau (de Jaeger et al. 2015, 2017). New measurements of H_0 from these methods, which are 69 ± 16 km s^{-1} Mpc^{-1} for SCM (Olivares E. et al. 2010) and 71 ± 8 km s^{-1} Mpc^{-1} for PMM in the V band (Rodríguez et al. 2019), have been derived. However, their precision is not yet comparable to those of Planck or SNe Ia owing to a lack of large number of SNe II in the Hubble flow, as well as to a small number of Cepheids or resolved red giants in SN II host galaxies (de Jaeger et al. 2020a).

3 HST KEY PROJECT AND HST DATA

The data for our analysis have been taken from the Hubble Space Telescope Key (HST Key) Project (Freedman et al. 2001). The main goal of the HST Key Project was to determine H_0 to an accuracy of $< 10\%$ by calibrating the secondary distance indicators by examining Cepheid variables. Several new Cepheid stars were discovered by HST in various galaxies within a distance of 25 Mpc. The Cepheid P-L relation was also calibrated against the metallicity by HST. The better seeing conditions and the ability to schedule the observations independent of phases of the Moon or weather conditions were the biggest advantage of HST compared to ground based observatories. Due to this, the number of Cepheids available for calibration increased drastically which is responsible for reduced uncertainties in the distance

measurement. Table 1 of Freedman et al. (2001) compares the status of Cepheid calibrators pre and post HST.

3.1 The Revised P-L Relation of Cepheid Variables

The Cepheid P-L relation was first introduced by Leavitt & Pickering (1912). Various authors, e.g., Gieren (1993), used Cepheid surface brightness to estimate distances and absolute magnitudes. Currently, Cepheids are among the best stellar distance indicators and an important initial step on the cosmic distance ladder. However, the sensitivity of the zero point of the P-L relation to the chemical composition has always been a matter of concern.

Measuring an accurate value of H_0 was one of the motivating reasons for building the HST. Thus, in the mid 1980s, accurate measurement of H_0 with an accuracy of 10%, by observing several Cepheids and hence calibrating the secondary distance indicators, was designated as one of the key objectives of the HST. Before the launch of HST, most Cepheid searches were restricted to our own Local Group of galaxies and the very nearest surrounding groups (M101, Sculptor, and M81 groups) (Madore & Freedman 1991). By that time, only five galaxies with well-measured Cepheid distances were available for calibration of the T-F relation. The calibration of the surface brightness fluctuation method (Tonry 1991) was done by considering a single Cepheid distance, i.e., M31. Moreover, before HST no Cepheid calibrators were available for SNe Ia. A large number of Cepheid variables were required to be observed to improve the calibration of the P-L relation and hence the calibration of secondary distance indicators. Several improvements and refinements were made by the HST team, including installation of HST Wide Field and Planetary Camera 2 (WFPC2) for photometric calibration. Observations of several Cepheids in the Large Magellanic Cloud (LMC) were carried out to obtain the fiducial P-L relation and to study the dependence of the P-L relation on metallicity. The final results of the HST Key Project were based on a Cepheid calibration of several secondary distance methods applied up to a distance of 400 Mpc.

3.2 Data

Calibration is often a challenging issue in astronomical measurements. Different primary indicators (such as Cepheids, TRGB, etc.) and different instruments are used for photometric calibration, which lead to systematic biases. The HST Key Project data are unique as they have been obtained through calibration of several different methods by a single primary distance indicator, i.e., Cepheid variables. Use of a single instrument (WFPC2 of HST) for photometric calibration also makes it special.

Based on the revised Cepheid P-L relation, 36 SNe Ia were calibrated, which are available in table 6 of the key paper Freedman et al. (2001). The value of H_0 obtained from this sample, $71 \pm 2 \pm 6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, is slightly higher than the previous measurement considering SNe Ia ($68 \pm 2 \pm 5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) (Freedman et al. 2001). Twenty-one galaxies in the general field and in various clusters and groups were calibrated for the T-F relation incorporating the newly available Cepheids. The measurements are available in table 7 of the key paper and provide a value of $H_0 = 71 \pm 3 \pm 7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. These results have not changed much from the previous measurements available in the literature (Sakai et al. 2000), indicating the self-consistency of the T-F relation with respect to the Cepheid P-L relation. Distances to 11 elliptical galaxies in various clusters were measured through the fundamental plane relation with a calibration relying on the revised P-L relation. The new value of H_0 is $82 \pm 6 \pm 9\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, which is substantially different from the previous measurements, with the reason being the galaxies in the Key Project were quite distant, and their metallicities were quite high. Thus, the new calibration had a larger impact on this sample. The other two methods, surface brightness fluctuations and SNe II, were also calibrated and considered for H_0 measurement. However, the sample sizes of these methods are very small, and hence we do not include them in our analysis. Altogether, we have three samples for our analysis, 36 SNe Ia, 21 T-F galaxies and 11 FPR galaxies.

4 METHODOLOGY

Five different probes have been employed to measure the Hubble constant in the HST Key Project. All the methods have been calibrated utilizing a common mechanism, i.e., the P-L relation of Cepheid variables. However, empirical relations and the physical concepts involved in these methods are different. It would be interesting to ask the following question: Do these methods provide the same value of H_0? If the answer is no, it raises doubts about our understanding of the physical concepts and the empirical relations involved. Alternatively, it could also be due to the inconsistency of the systematics involved in the measurement process. Various statistical methods, related to hypothesis testing, can also be employed to answer the above question. If μ_1 and μ_2 are the averages of H_0 from two different samples, then the null hypothesis can be set as

$$\mu_1 = \mu_2. \tag{1}$$

The alternative hypothesis which does not emphasise any of the samples, the non-directional hypothesis, would be

$$\mu_1 \neq \mu_2\,. \tag{2}$$

Table 1 Hubble Constant Measurements Using SNe Ia, σ Represents the Uncertainty in H_0 (table 6 of Freedman et al. 2001)

SN	H_0	σ	SN	H_0	σ	SN	H_0	σ
SN1990O	67.3	2.3	SN1990T	75.6	3.1	SN1990af	75.8	2.8
SN1991S	69.8	2.8	SN1991U	83.7	3.4	SN1991ag	73.7	2.9
SN1992J	74.5	3.1	SN1992P	64.8	2.2	SN1992ae	81.6	3.4
SN1992ag	76.1	2.7	SN1992al	72.8	2.4	SN1992aq	64.7	2.4
SN1992au	69.4	2.9	SN1992bc	67.0	2.1	SN1992bg	70.6	2.4
SN1992bh	66.7	2.3	SN1992bk	73.6	2.6	SN1992bl	72.7	2.6
SN1992bo	69.7	2.4	SN1992bp	76.3	2.6	SN1992br	67.2	3.1
SN1992bs	67.8	2.8	SN1993B	69.8	2.4	SN1993O	65.9	2.1
SN1993ag	69.6	2.4	SN1993ah	71.9	2.9	SN1993ac	72.9	2.7
SN1993ae	75.6	3.1	SN1994M	74.9	2.6	SN1994Q	68.0	2.7
SN1994S	72.5	2.5	SN1994T	71.5	2.6	SN1995ao	78.8	2.7
SN1995ak	80.9	2.8	SN1996C	66.3	2.5	SN1996bl	78.7	2.7

Table 2 Hubble Constant Measurements Using T-F Relation, σ Represents the Uncertainty in H_0 (table 7 of Freedman et al. 2001)

Cluster/Group	H_0	σ	Cluster/Group	H_0	σ	Cluster/Group	H_0	σ
Abell 1367	75.2	12.8	Abell 0262	70.9	12.7	Abell 2634	77.7	12.7
Abell 3574	76.2	11.9	Abell 0400	79.3	12.6	Antlia	68.8	10.3
Cancer	67.1	10.9	Cen30	75.8	17.4	Cen45	70.7	10.9
Coma	83.5	13.9	Eridanus	77.6	13.1	ESO50	79.8	11.9
Fornax	92.2	15.2	Hydra	69.6	10.6	MDL59	73.6	13.7
NGC 3557	85.0	12.9	NGC 0383	73.9	12.9	NGC 0507	84.9	14.6
Pavo2	86.3	15.0	Pegasus	66.4	11.7	Ursa Major	54.8	8.6

Table 3 Hubble Constant Measurements Using Fundamental Plane Relation, σ Represents the Uncertainty in H_0 (table 9 of Freedman et al. 2001)

Cluster/Group	H_0	σ	Cluster/Group	H_0	σ	Cluster/Group	H_0	σ
Dorado	81.9	8.2	GRM15	95.6	10.7	Hydra I	82.8	8.0
Abell S753	87.5	7.2	Abell 3574	92.0	9.7	Abell 194	91.3	7.6
Abell S639	109.7	10.0	Coma	83.2	6.2	Abell 539	86.2	6.4
DC 2345-28	83.2	6.5	Abell 3381	88.9	8.2			

The alternative hypothesis would be directional if it emphasises a particular sample, i.e., $\mu_1 < \mu_2$ *or* $\mu_1 > \mu_2$.

Various parametric methods such as the t-test are often employed for hypothesis testing when the data samples are drawn from a Gaussian distribution. As the central limit theorem suggests that outcomes of a measurement process would follow a Gaussian distribution, we expect the same for our data sets as well. However, non-parametric methods provide more reliable results when the distribution of data values is far from a bell curve. We briefly discuss these techniques in the next section.

4.1 Parametric Methods: t-test

Average values of H_0 obtained from two different methods can be compared applying an unpaired t-test since the number of data points are different for different samples. If the mean values of H_0 for two different samples with n_1 and n_2 data points are M_1 and M_2 respectively then the t-score is defined as

$$T = \frac{M_1 - M_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}, \tag{3}$$

where s_1^2 and s_2^2 are the variances of the first and second samples respectively (Sheskin 1997). The uncertainties are an important part of the measurement process and contain vital information. We thus weigh the measurements, i.e., H_0 values, with the uncertainties, so that more precise values get more weight.

$$H_0^{'i} = \frac{H_0^i/\sigma_i^2}{\sum_j 1/\sigma_j^2}, \tag{4}$$

where H_0^i represents an individual measurement of Hubble constant (Crandall & Ratra 2015; Podariu et al. 2001). Since our alternative hypothesis in Equation (2) is non-directional, as it does not emphasise any particular method ($\mu_1 \neq \mu_2$), a "two-tailed test" would be required. One can compare the t-score calculated from Equation (3) with a critical value which is the probability of obtaining the data samples assuming that the null hypothesis is true. If the t-score is larger than the critical value, the null hypothesis is rejected but a smaller t-score supports the null hypothesis.

4.2 Non-parametric Methods

Often due to systematic effects, the data values are far from a normal distribution. In such cases, distribution free tests or non-parametric tests are applied. If the data values follow a normal distribution, the relative efficiency of parametric methods is higher and vice-versa. We now outline some non-parametric methods which can be applied to the data (Sheskin 1997).

4.2.1 Mann-Whitney U test

Due to its resistance to outliers, the median is a more robust estimate of central tendency than mean. The Mann-Whitney U test compares the median of two samples. The null hypothesis, in this case, would be $M_1 = M_2$, where M_1 and M_2 are the medians of the samples, respectively. The non-directional alternative hypothesis would be $M_1 \neq M_2$ and the directional hypothesis would be either $M_1 < M_2$ or $M_1 > M_2$. The data values of the two samples are arranged in ascending order, ranks are assigned to the combined data, and the sum of ranks of each sample ΣR_i is calculated. The U values are computed by relying on Equation (5)

$$U_i = n_i n_j + \frac{n_i(n_i+1)}{2} - \sum R_i, \tag{5}$$

where n_i and n_j are the sample sizes. One can easily verify that U_i and U_j are always positive and $U_i + U_j = n_i n_j$. The smaller value of U_i and U_j is designated as the U statistic. Statistical tables can be referenced to calculate the critical value of U for small samples. For large samples, the approximate normal deviate z is calculated (Sheskin 1997),

$$z = \frac{(|U - \frac{n_i n_j}{2}| - 0.5)}{A}, \tag{6}$$

where $A = \sqrt{\frac{n_i n_j (n_i + n_j + 1)}{12}}$.

4.2.2 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test (K-S test) is utilized to compare a data sample with a reference probability distribution (Sheskin 1997) or to compare two different samples by computing their cumulative distribution functions (CDFs). The distributions are calculated under the null hypothesis that both the samples are drawn from an identical distribution. The K-S test is quite useful since it is sensitive to the difference in location as well as shape of the underlying distributions. The two-sample K-S test estimates the difference between the CDFs of the two sample data vectors over a given range of x in each data set. The test statistic is the maximum distance, D, between the CDFs

$$D = \max|F_1(x) - F_2(x)|, \tag{7}$$

Table 4 Mean and standard deviation of H_0 values obtained from different methods: (a) SNe Ia, (b) T-F and (c) fundamental plane relation.

Method	Mean H_0	Std dev	No of data points
SNe	72.18	4.87	36
TF	75.68	8.34	21
FPR	89.3	8.07	11

where $F_1(x)$ and $F_2(x)$ are the proportions of x_1 and x_2 values that are less than or equal to x respectively. For our analysis, we applied the K-S test function available in Matlab, $h = kstest2(x_1, x_2)$, where x_1 and x_2 are the two samples. Based on the maximum distance D, it returns a value of zero or one. For large values of D, the function returns $h = 1$ which implies rejection of the null hypothesis. For small values of D, the function returns $h = 0$, which implies a failure to reject the null hypothesis at the given significance level, which is $\alpha = 0.01$ in our analysis.

5 RESULTS AND DISCUSSION

We first calculate the average value of the Hubble constant and its standard deviation obtained from each method and present it in Table 4. It is clear that although the average values of H_0 are close to each other for SNe and TF, their standard deviations are significantly different. The average in case of FPR is quite different from the other two methods.

Now, we sketch a histogram of the H_0 values in the three samples which are presented in Figure 1. A first glance at the figures indicates a deviation from the normal distribution. Although a small sample size in the case of T-F and fundamental plane relations could be one of the reasons for the deviation, the SN sample size is sufficient to display a bell shape. In any case, this motivates us to use non-parametric methods. Thus, along with the regular parametric tests, we apply non-parametric methods as well.

5.1 Results for the Parametric Test: t-test

As mentioned in Section 3.2, there are three samples of data and, hence, three pairs of data samples, namely, pair 1: SNe-TF, pair 2: TF-FPR and pair 3: SNe-FPR. Our first pair of samples consists of data from SNe Ia and T-F methods. A t-test is performed to check the null hypothesis (Eq. (1)) that the average value of H_0 is the same in SNe and T-F methods. The t-score has been calculated using Equation (3) and is presented in Table 5 (row 1). The critical value for the non-directional alternative hypothesis (Eq. (2)): "average values of H_0 for the two methods are not the same" taken from the table at 99% significance level is 2.7. Since the observed t value is smaller than the critical value, the null hypothesis is not rejected. However,

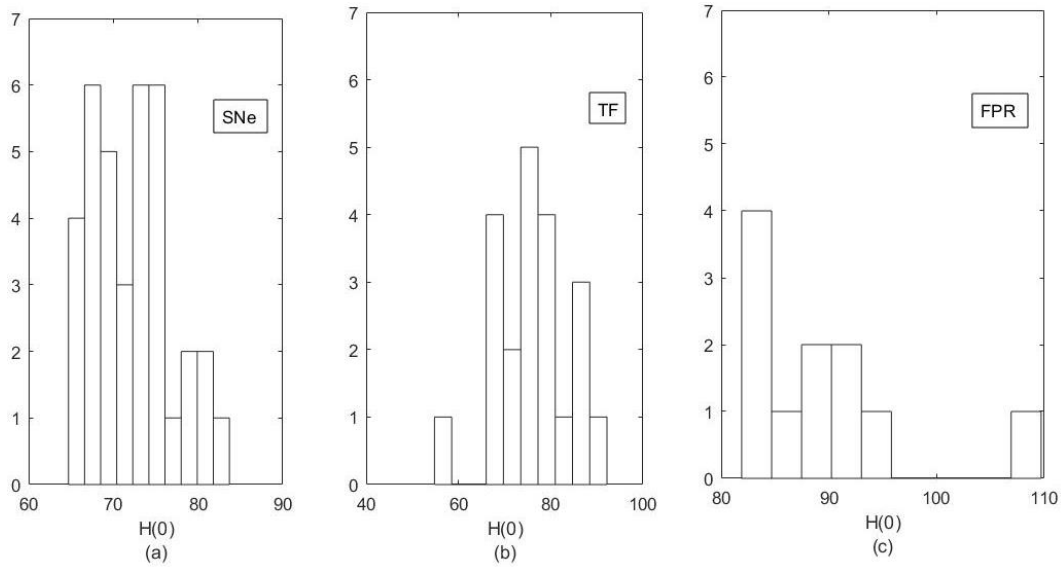

Fig. 1 Frequency distribution of H_0 values obtained from different methods: (a) SNe Ia, (b) T-F and (c) fundamental plane relation.

a one-tailed t-test is implemented, if the alternative hypothesis is directional, i.e., $H_0(\mathrm{SNe})$ is smaller than $H_0(\mathrm{TF})$. The critical value at 95% significance level is 1.7, which is smaller than the t-score, and the null hypothesis is rejected. A similar analysis for the other two pairs is also presented in Table 5. It is clear that for both pairs, the obtained value is greater than the critical value; hence, the null hypothesis is rejected at the 99% significance level.

We now include measurement uncertainties in the analysis by defining a new variable $H_0^{'}$ as the H_0 values weighted by the measurement uncertainties as defined in Equation (4) in each sample. The t-scores for various pairs of samples are again calculated and are presented in Table 6. Clearly, the t-score is larger compared to the critical value in all cases. Hence the null hypothesis "the measurement values are the same using different methods" is rejected in all cases.

5.2 Results for Non-parametric Tests

Since the histograms of data samples in Figure 1 do not show a clear bell shape, we apply the non-parametric tests as well. The tests are described in Section 4.2, and the results are discussed below.

5.2.1 U test

The null hypothesis for U test has been set up by replacing the mean with the median in Equation (1), and the numerical values of U_1 and U_2 have been calculated using Equation (5). Since the sample sizes are relatively large to obtain the critical U from the table, the normal approximate z has been calculated utilizing Equation (6).

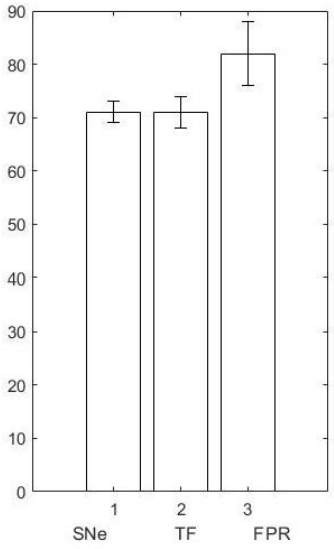

Fig. 2 Comparison of H_0 obtained from different methods. Errors are signified by vertical lines on top.

All these values are presented in Table 7. The critical value of z corresponding to the 99% significance level is 2.58, which is common in all cases. Since the z value in column 7, obtained using Equation (6), is smaller for the SN-TF pair, it supports the null hypothesis. On the other hand, it is greater than 2.58 for the TF-FPR and SN-FPR pairs and the null hypothesis is rejected for these sample pairs. However, the table value of z at 95% significance level is 1.96; thus, the null hypothesis is rejected for this case also at the 95% level.

In order to make use of the uncertainties in the measurement, we apply the U test on the $H_0^{'}$ values obtained utilizing Equation (4) for all the three pairs of

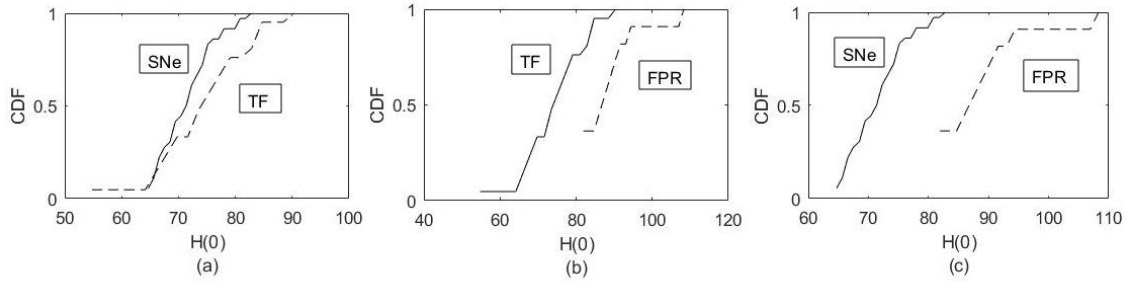

Fig. 3 K-S Test: CDF for original values of Hubble constant obtained from different methods. CDFs of different pairs have been plotted together for comparison: (a) SNe Ia and T-F Fig, (b) T-F and FPR and (c) SNe Ia and FPR.

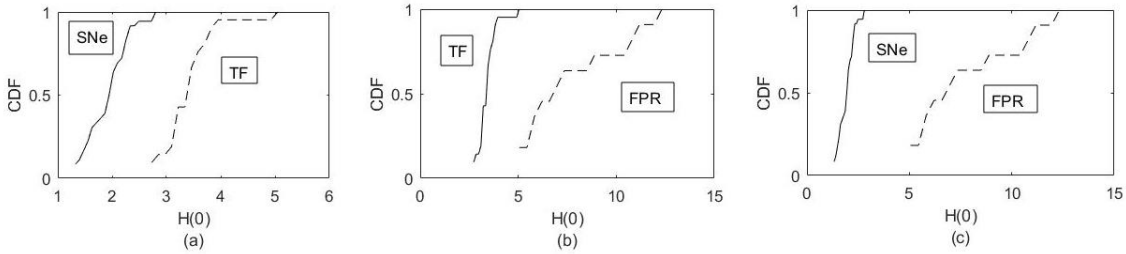

Fig. 4 K-S Test: CDF for Hubble constant values weighted by uncertainties. Comparison of CDF for different samples: (a) SNe Ia and T-F, (b) T-F and FPR and (c) SNe Ia and FPR.

Table 5 Unpaired t-test between various pairs of data samples without considering uncertainties. Null hypothesis (Eq. (1)) is rejected at the 99% significance level except for the SNe-TF pair.

Pair	μ_1	μ_2	S_1^2	S_2^2	$T_{\rm theory}$	$T_{\rm critical}$	DOF	Result
SN-TF	72.2	75.7	23.8	69.6	1.8	2.7	55	NR
TF-FPR	75.7	89.3	69.6	65.2	4.5	2.8	30	R
SN-FPR	72.2	89.3	23.8	65.2	6.7	2.7	45	R

samples. As in the previous case, the U_1, U_2 and z values have been calculated by relying on Equations (5) and (6) and are presented in Table 8. It is clear from the table that since in all cases, the tabulated z value is smaller than the calculated z value, the null hypothesis is rejected at the 99% significance level in all cases.

5.2.2 K-S test

Finally, we perform the K-S test for the same three pairs of samples to verify the equality of different methods (see Sect. 4.2.2). The original H_0 values were supplied to the Matlab function $kstest2$ to calculate the value of h at the 99% significance level. The results are presented in Table 9 (A). The value of h for the SN-TF pair is zero which supports the null hypothesis. However, in the remaining two cases, h is equal to one, indicating the rejection of the null hypothesis. The CDFs of these pairs have been plotted in Figure 3. The CDFs obtained for the TF-FPR pair and for the SN-FPR pair are quite far from each other. Since the distance between the CDFs is very large in both cases, the null hypothesis is rejected. Although the CDFs for SN and TF do not match well, they are not too different. This is the reason for the non-rejection of the null hypothesis.

In order to exploit the information available in the uncertainties, we apply the K-S test on the uncertainty weighted measurements, i.e., $H_0^{'}$. The results are presented in Table 9 (B). This time the null hypothesis is rejected in all three cases. Thus we conclude that the different methods of measuring H_0 provide different results. The CDFs of H_0 values from different methods have been compared in Figure 4. It is clear from the figure that in all the cases the distance between the CDFs is too large, hence the null hypothesis is rejected in all cases.

6 CONCLUSIONS

We have applied both the parametric and non-parametric methods to test if HST Key measurements of H_0 from different methods based on the cosmic distance ladder are statistically different. The use of HST Key data is favoured to avoid the instrumental biases arising from different telescopes. Our null hypothesis is that the average or median H_0 values obtained from different methods are

Table 6 Unpaired t-test between various pairs of samples (H_0 values weighted by uncertainties). Null hypothesis is rejected at the 99% significance level in all cases.

Pair	μ_1	μ_2	S_1^2	S_2^2	$T_{\rm theory}$	$T_{\rm critical}$	DOF	Result
SN-TF	2.0	3.5	0.2	0.3	11.9	2.7	55	R
TF-FPR	3.5	8.0	0.3	7.1	5.5	2.8	30	R
SN-FPR	2.0	8.0	0.2	7.1	7.5	2.7	45	R

Table 7 U test between various pairs of samples (using original H_0 values). z values in column 8 were referenced from a table of a standard normal distribution at the 99% significance level. The null hypothesis is rejected in two cases.

Pair	n_1	n_2	U_1	U_2	U	z from Eq. (6)	z from table	Result
SN-TF	36	21	501.5	254.5	254.5	2.03	2.58	NR
TF-FPR	21	11	205	26	26	3.53	2.58	R
SN-FPR	36	11	392	4	4	4.86	2.58	R

Table 8 U test between various pairs of samples (H_0 values weighted by uncertainties). z values in column 8 were referenced from a table of standard normal distribution for 99% significance level. The null hypothesis is rejected in all cases.

Pair	n_1	n_2	U_1	U_2	U	z from Eq. (6)	z from table	Result
SN-TF	36	21	3	753	3	6.19	2.58	R
TF-FPR	21	11	1	230	1	4.5	2.58	R
SN-FPR	36	11	0	396	0	4.96	2.58	R

Table 9 (A) indicates the K-S test results for sample pairs without including uncertainties. The null hypothesis is rejected in two cases. (B) signifies the K-S test results for samples including uncertainties. The null hypothesis is rejected in all cases.

Pair	h value	Result	Pair	h value	Result
SN-TF	0	NR	SN-TF	1	R
TF-FPR	1	R	TF-FPR	1	R
SN-FPR	1	R	SN-FPR	1	R
(A)			(B)		

the same. Based on the results presented in Section 5, we conclude that the null hypothesis is rejected at the 99% confidence level in most cases and at the 95% level in all cases. A possible reason for the failure could be the gaps in our understanding of the physical processes involved in these cosmological probes. As an example, SNe Ia were assumed to be among the most precise secondary distance indicators. However, the possibility of sub-classes within the SN Ia class (discovered after the HST Key era) could bias the correlation between the peak luminosity and decline rate which might lead to undesired systematic effects in distance measurement. The physics of the T-F relation was also poorly understood, for instance, the constant mass to light ratio assumption may not be reliable, and the role of dark matter in the galaxy rotation curve is also debatable (Gurovich et al. 2004), which could bias the distances measured through the T-F relation. The difference in the metallicity of various elliptical galaxies in FPR method could also affect the distance measurement. The SNe Ia measurements have improved over time, and precise measurement of H_0 relying on SNe Ia is available now (Riess et al. 2016). Although other methods have been improved since the HST Key Project era (Verde et al. 2019), it will take a long time to reach the required level of sensitivity.

Acknowledgements SG thanks SERB for financial assistance (EMR/2017/003714). RN, RKT and PKT extend their gratitude to BITS-Pilani Hyderabad Campus for providing all the required infrastructure.

References

Bernardi, M., Hyde, J. B., Sheth, R. K., et al. 2007, AJ, 133, 1741
Crandall, S., & Ratra, B. 2015, ApJ, 815, 87
de Jaeger, T., González-Gaitán, S., Anderson, J. P., et al. 2015, ApJ, 815, 121
de Jaeger, T., González-Gaitán, S., Hamuy, M., et al. 2017, ApJ, 835, 166
de Jaeger, T., Stahl, B. E., Zheng, W., et al. 2020a, MNRAS, 496, 3402
de Jaeger, T., Galbany, L., González-Gaitán, S., et al. 2020b, MNRAS, 495, 4860
Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
Freedman, W. L., & Madore, B. F. 1990, ApJ, 365, 186
Freedman, W. L. 2000, Phys. Rep., 333, 13
Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, ApJ, 553, 47
Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, ApJ, 882, 34
Freedman, W. L., Madore, B. F., Hoyt, T., et al. 2020, ApJ, 891, 57
Gieren, W. P. 1993, MNRAS, 265, 184
Giovanelli, R., Haynes, M. P., da Costa, L. N., et al. 1997, ApJL, 477, L1

Gurovich, S., McGaugh, S. S., Freeman, K. C., et al. 2004, PASA, 21, 412
Hamuy, M., & Pinto, P. A. 2002, ApJL, 566, L63
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168
Jensen, J. B., Tonry, J. L., Thompson, R. I., et al. 2001, ApJ, 550, 503
Jha, S. W. 2017, Type Iax Supernovae, eds. A. W. Alsabti, & P. Murdin, Handbook of Supernovae (Springer International Publishing AG), 375
Joachimi, B., Singh, S., & Mandelbaum, R. 2015, MNRAS, 454, 478
Jorgensen, I., Franx, M., & Kjaergaard, P. 1996, MNRAS, 280, 167
Kelson, D. D., Illingworth, G. D., Tonry, J. L., et al. 2000, ApJ, 529, 768
Khetan, N., Izzo, L., Branchesi, M., et al. 2020, arXiv e-prints, arXiv:2008.07754
Kirshner, B. P. S. R. P., Eastman, R. G., Phillips, M. M., et al. 1994, arXiv e-prints, astro-ph/9407098
Kowal, C. T. 1968, AJ, 73, 1021
Lanoix, P., Paturel, G., & Garnier, R. 1999, ApJ, 517, 188
Lauer, T. R., Tonry, J. L., Postman, M., et al. 1998, ApJ, 499, 577
Leavitt, H. S., & Pickering, E. C. 1912, Harvard College Observatory Circular, 173, 1
Liu, M. C., Graham, J. R., & Charlot, S. 2002, ApJ, 564, 216
Madore, B. F., & Freedman, W. L. 1991, PASP, 103, 933
McGaugh, S. S. 2005, ApJ, 632, 859
McGaugh, S. S. 2012, AJ, 143, 40
Mei, S., Blakeslee, J. P., Tonry, J. L., et al. 2005, ApJS, 156, 113
Nomoto, K., Iwamoto, K., & Kishimoto, N. 1997, Science, 276, 1378
Nussbaumer, H., & Bieri, L. 2009, Discovering the Expanding Universe (Cambridge, UK: Cambridge Univ. Press)
Olivares E., F., Hamuy, M., Pignata, G., et al. 2010, ApJ, 715, 833
Phillips, M. M. 1993, ApJL, 413, L105
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2018, arXiv e-prints, arXiv:1807.06209
Podariu, S., Souradeep, T., Gott, J. Richard, I., et al. 2001, ApJ, 559, 9
Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, ApJ, 826, 56
Riess, A. G., Casertano, S., Yuan, W., et al. 2019, ApJ, 876, 85
Rodríguez, Ó., Clocchiatti, A., & Hamuy, M. 2014, AJ, 148, 107
Rodríguez, Ó., Pignata, G., Hamuy, M., et al. 2019, MNRAS, 483, 5459
Sakai, S., Mould, J. R., Hughes, S. M. G., et al. 2000, ApJ, 529, 698
Sakstein, J., & Trodden, M. 2020, Phys. Rev. Lett., 124, 161301
Sandage, A., & Tammann, G. A. 1982, ApJ, 256, 339
Saulder, C., Steer, I., Snaith, O., & Park, C. 2019, arXiv e-prints, arXiv:1905.12970
Schaeffer, R., Maurogordato, S., Cappi, A., & Bernardeau, F. 1993, MNRAS, 263, L21
Seljak, U. 2002, MNRAS, 334, 797
Sheskin, D. J. 1997, Handbook of Parametric and Nonparametric Statistical Procedures (CRC Press)
Slatyer, T. R., & Wu, C.-L. 2018, Phys. Rev. D, 98, 023013
Tonry, J., & Schneider, D. P. 1988, AJ, 96, 807
Tonry, J. L. 1991, ApJL, 373, L1
Tonry, J. L., Blakeslee, J. P., Ajhar, E. A., & Dressler, A. 2000, ApJ, 530, 625
Tully, R. B., & Fisher, J. R. 1977, A&A, 500, 105
Tully, R. B., & Pierce, M. J. 2000, ApJ, 533, 744
Umeda, H., Nomoto, K., Kobayashi, C., et al. 1999, ApJL, 522, L43
Vagnozzi, S. 2020, Phys. Rev. D, 102, 023518
van den Bergh, S. 1995, ApJL, 453, L55
Verde, L., Treu, T., & Riess, A. G. 2019, Nature Astronomy, 3, 891
Willick, J. A., Courteau, S., Faber, S. M., et al. 1997, ApJS, 109, 333