

The intrinsic colour dispersion in Type Ia supernovae

S. Nobili¹, A. Goobar¹, R. Knop², and P. Nugent³

¹ Department of Physics, Stockholm University, SCFAB, 106 91 Stockholm, Sweden

² Department of Physics & Astronomy, Vanderbilt University, USA

³ Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

Received 14 February 2003 / Accepted 6 April 2003

Abstract. The properties of low-redshift Type Ia supernovae are investigated using published multi-band optical broadband data from the Calan/Tololo and CfA surveys. The average time evolution of $B - V$, $V - R$, $R - I$, $B - I$ and $V - I$, the intrinsic dispersion and time correlations are studied. This information is required to deduce the extinction of such explosions from the measured colours. We find that extinction corrections on individual SNe based on their colours up to 40 days past the B -band lightcurve maximum are generally limited to $\sigma_{A_V} \gtrsim 0.1$, due to intrinsic variations, as far as it can be conservatively deduced with the current sample of data. However, we find that the $V - R$ colour, especially at late times, is consistent with a negligible intrinsic spread, and may be the most accurate estimator for extinction.

Key words. supernovae: general – stars: statistics

1. Introduction

In the last few years Type Ia supernovae (SNe Ia) have proved to be excellent distance estimators and have been successfully used to investigate the fate of the universe (Perlmutter et al. 1999; Riess et al. 1998). Despite the broad use of these objects by cosmologists, the current knowledge of the nature of SNe Ia is rather limited. Thus, there is a strong demand for further understanding to assess important issues about these explosions. For cosmological implications, the main concerns are related to the possible evolution of the SN properties with redshift. Moreover, critical tests for extinction along the line of sight based on supernova colours require good knowledge of the intrinsic properties of these objects.

In this paper, a statistical study on 48 well observed nearby SNe Ia is carried out. In particular, the intrinsic dispersion in SN colours is investigated using published $BVRI$ data. We also focus on the time correlation of intrinsic optical colours. This information is needed to address the possible host galaxy or intergalactic extinction by dust of supernovae used for cosmological tests (see e.g. Riess et al. 2000; Nobili et al. 2003), and also to probe for other exotic sources of dimming at high- z with differential extinction (Mörtsell et al. 2002).

2. The data set

Published $BVRI$ lightcurves of well observed nearby SNe Ia were analysed. The considered sample consists of 48 SNe Ia from 2 different sets, the Calan/Tololo data published by Hamuy et al. (1996), and the set in Riess et al. (1999), usually

referred to as the CfA data. The list of SNe is given in Table 1, along with the observed filter data available for each of them, their redshift and the B -band lightcurve “stretch”, s , as defined in Perlmutter et al. (1997) and Goldhaber et al. (2001).

The selected samples include a broad variety of SNe, well distributed in stretch factor, s . This parameter, related to Δm_{15} (Phillips 1993), has been found to correlate with the supernova luminosity. Thus, a sample well distributed in stretch should imply a broad distribution in luminosity. Figure 1 shows the distribution of the B -band light curve stretch factor and SN redshifts ($0.003 \leq z \leq 0.12$), for both samples. The timescale stretch parameter was determined from the lightcurve fits as in Goldhaber et al. (2001) using their B -band lightcurve template with a parabolic behavior for the earliest epoch after explosion.

K-corrections were applied to account for the small cosmological redshift as described in e.g. Goobar et al. (2002), using the ($s = 1$) spectroscopic template of Nugent et al. (2002) as a starting point. The results of this analysis were used to improve the spectral template of SN Ia’s (Sect. 6) and we iterated the analysis once re-calculating the K-corrections with the improved template. Note that even though the K-corrections for the used data set are small, typically of the order of a few hundreds of a magnitude, for some of the more distant objects in the sample considered, they reach up to ~ 0.5 mag.

The SNe light curves were corrected for both Milky Way and host galaxy extinction as in Phillips et al. (1999) using the method first proposed by Lira (1995) for estimating host galaxy extinction using late epoch light curves. There is empirical evidence that the $B - V$ colors of SNe Ia show extremely small scatter for the period between 30 and 90 days post B -band maximum, despite any difference in the light curve shapes at earlier

Send offprint requests to: S. Nobili, e-mail: serena@physto.se

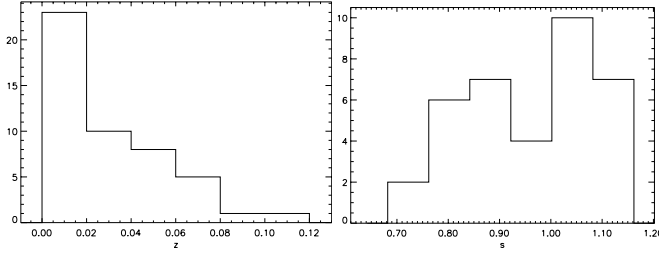


Fig. 1. Distribution of redshift, z , (binsize = 0.02), and stretch factor, s , (binsize = 0.08), for the analysed sample.

epochs. As the set of SNe used in this article is a subset of the one analysed in Phillips et al. (1999), the host galaxy extinctions listed in Table 2 of their paper were used. These were derived combining the late time $B - V$ colour with information on the $B - V$ and $V - I$ at maximum light.

The Cardelli et al. (1989) relation, modified by O’Donnell (1994), was used to compute the extinction in other colours given $E(B - V)$. Spectral templates of Type Ia SNe were used to compute the evolution of the extinction with the supernova epoch.

The extinction corrected lightcurves were further screened to exclude the most peculiar SNe, as the main emphasis of this work is to establish the properties of “normal” supernovae. Figure 2 shows the difference of the B -band lightcurve maximum, B_{\max} and the V -band light curve maximum V_{\max} plotted against the decline rate parameter, Δm_{15} , as reported in Phillips et al. (1999). A 3σ clipping rejection criteria was applied, iterating until no data points were further rejected. At least 5 SNe in the 2 data sets deviate significantly from the expected linear relation derived in Phillips et al. (1999) from an independent set of 20 non-reddened SNe (solid line in Fig. 2):

$$B_{\max} - V_{\max} = -0.07(\pm 0.012) + 0.114(\pm 0.037)(\Delta m_{15} - 1.1).$$

The outliers are SN 1993ae, SN 1995bd, SN 1996ai, and SN 1996bk for the CfA data, and SN 1992K for the Calan/Tololo set. Moreover SN 1995E and SN 1995ac have been excluded since they are similar to the peculiar SN 1991T (Branch et al. 1993)¹.

3. The average optical colours

Figures 3–8 show the time evolution of the extinction and K-corrected $B - V$, $V - R$, $R - I$, $B - I$ and $V - I$, colours. The plotted errors include the uncertainty on the host galaxy extinction correction. The time axis (t') has been corrected by the SN redshift and the B -band stretch factor s , to account for the dependence of colours on the stretch. This rescaling of the time axis was found very effective in reducing the measured intrinsic dispersion (Sect. 4)

$$t' = \frac{t - t_{B_{\max}}}{s \cdot (1 + z)}. \quad (1)$$

For each colour lightcurve we have spline-interpolated the weighted mean values computed in four days wide, non-overlapping bins. The aim of this procedure was to find a

¹ Including these 2 SNe in the analysis does not change the results significantly.

Table 1. List of SNe used for the analysis.

SN	Band	z	s	Ref.
1993ac	B, V, R, I	0.049	0.865	(1)
1993ae ^a	B, V, R, I	0.019	0.846	(1)
1994ae	B, V, R, I	0.004	1.033	(1)
1994M	B, V, R, I	0.023	0.865	(1)
1994Q	B, V, R, I	0.029	1.116	(1)
1994S	B, V, R, I	0.015	1.061	(1)
1994T	B, V, R, I	0.035	0.890	(1)
1995ac ^a	B, V, R, I	0.050	1.123	(1)
1995ak	B, V, R, I	0.023	0.857	(1)
1995al	B, V, R, I	0.005	1.044	(1)
1995bd ^a	B, V, R, I	0.016	1.131	(1)
1995D	B, V, R, I	0.007	1.081	(1)
1995E ^a	B, V, R, I	0.012	1.024	(1)
1996ai ^a	B, V, R, I	0.003	1.110	(1)
1996bk ^a	B, V, R, I	0.007	0.761	(1)
1996bl	B, V, R, I	0.036	1.030	(1)
1996bo	B, V, R, I	0.017	0.902	(1)
1996bv	B, V, R, I	0.007	1.106	(1)
1996C	B, V, R, I	0.030	1.102	(1)
1996X	B, V, R, I	0.007	0.889	(1)
1990af	B, V	0.051	0.792	(2)
1990O	B, V, R, I	0.030	1.116	(2)
1990T	B, V, R, I	0.040	0.998	(2)
1990Y	B, V, R, I	0.039	1.007	(2)
1991ag	B, V, R, I	0.014	1.084	(2)
1991S	B, V, R, I	0.055	1.114	(2)
1991U	B, V, R, I	0.032	1.068	(2)
1992ae	B, V	0.075	0.970	(2)
1992ag	B, V, I	0.025	0.951	(2)
1992al	B, V, R, I	0.015	0.963	(2)
1992aq	B, V, I	0.102	0.868	(2)
1992au	B, V, I	0.061	0.787	(2)
1992bc	B, V, R, I	0.020	1.039	(2)
1992bg	B, V, I	0.035	0.983	(2)
1992bh	B, V, I	0.045	1.048	(2)
1992bk	B, V, I	0.058	0.825	(2)
1992bl	B, V, I	0.044	0.812	(2)
1992bo	B, V, R, I	0.019	0.767	(2)
1992bp	B, V, I	0.079	0.907	(2)
1992br	B, V	0.088	0.682	(2)
1992bs	B, V	0.064	1.030	(2)
1992J	B, V, I	0.045	0.798	(2)
1992K ^a	B, V, I	0.010	0.787	(2)
1992P	B, V, I	0.037	0.952	(2)
1993ag	B, V, I	0.049	0.917	(2)
1993B	B, V, I	0.070	1.023	(2)
1993H	B, V, R, I	0.024	0.774	(2)
1993O	B, V, I	0.051	0.950	(2)

(1), Riess et al. (1998); (2), Hamuy et al. (1996).

^a Excluded from the analysis presented here.

curve that describes the *average* time evolution of the colours. This parameterization will be referred as a “*model*” in the

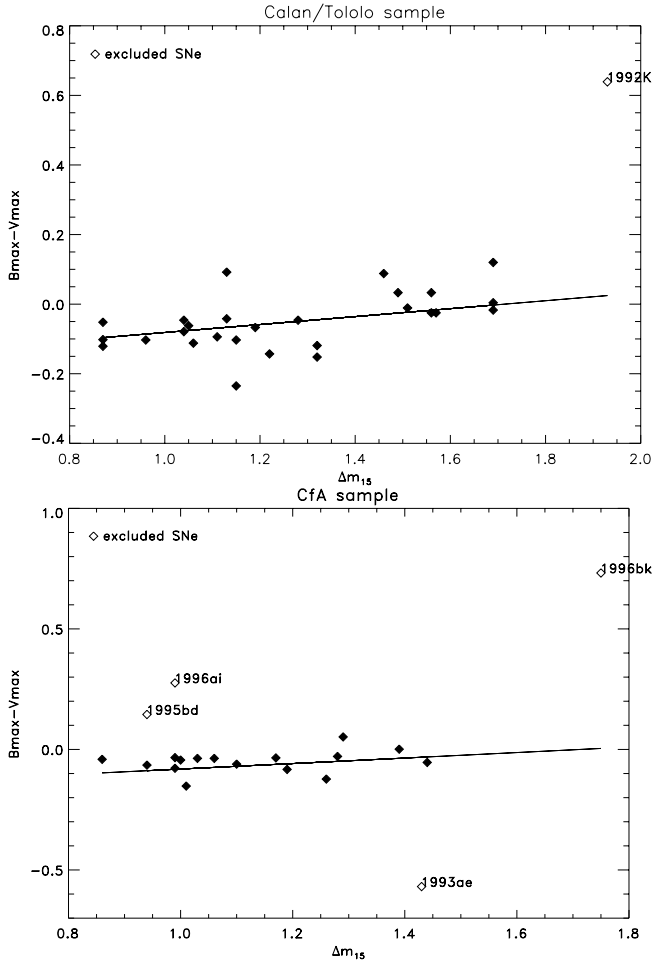


Fig. 2. $B_{\max} - V_{\max}$ vs. Δm_{15} for the 2 sets of data, k-corrected, corrected for Galactic extinction and Host Galaxy extinction, compared with the result of Phillips et al. (1999).

following discussion. Other methods were used to fit the data, as for example a least squares cubic spline fit. These result in curves that differ from the *model* typically by 0.01 mag. The *models* show some systematic discrepancies between the two data sets, especially in $R - I$ for which the C-T *model* is always redder than the CfA *model*. In Fig. 9 we investigate the differences of each of the *models* from the one built on both sets together. The differences are usually of the order of a few hundreds of a magnitude. The largest deviation was found for $R - I$ in the Calan/Tololo set, resulting in a difference of about 0.2 mag at maximum with respect to the *model* built up using data from both sets. Note, however, that the statistics in the Calan/Tololo set for $R - I$ colour, all along the evolution and in particular around the time of B_{\max} , is extremely poor, as shown in Fig. 5. Due to the smaller quoted observational error bars, the CfA set dominates the weighted average used to build the *model* out of both data sets.

A source of uncertainty in this analysis is the ability of the observers to convert the instrumental magnitudes from the used filter+CCD system transmission into the standard $BVRI$ system. The apparent systematic effects for $R - I$ in Fig. 9 may be indicative of this.

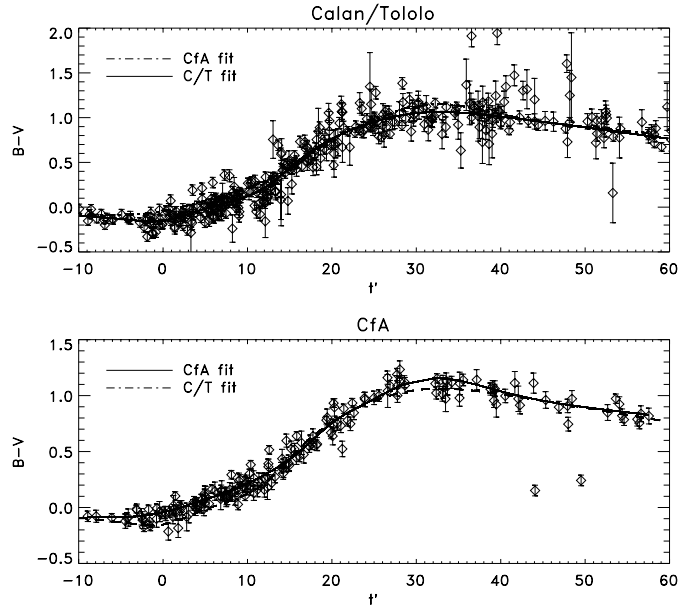


Fig. 3. $B - V$ for the two sets of data. The solid line represent always the curve found for the current set, while the dashed line is the curve found for the same colour of the other set.

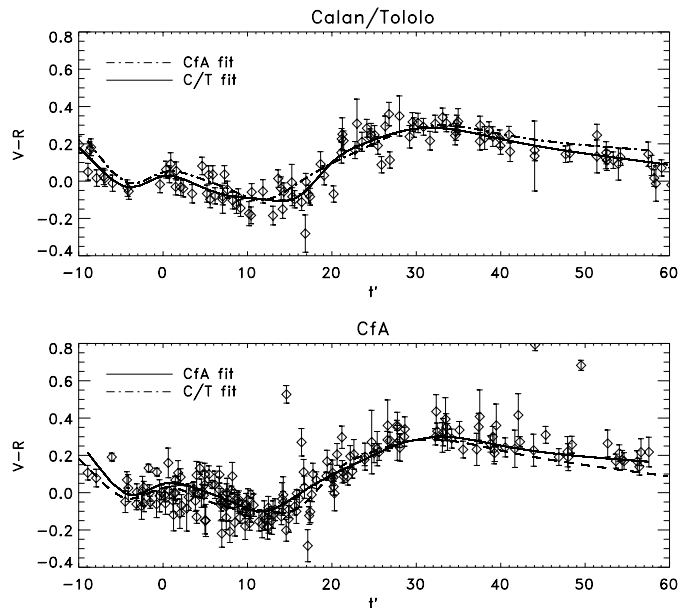


Fig. 4. $V - R$ for the two sets of data. The solid line represent always the curve found for the current set, while the dashed line is the curve found for the same colour of the other set.

In order to assess systematic effects in building the colour *models*, we checked whether the host galaxy extinction was over-corrected. Thus, we compared the *models* with the data of those SNe that, according to Phillips et al. (1999), suffered no extinction from the host galaxy. The comparison, shown in Fig. 7, exhibits no obvious deviations and we may conclude that the extinction corrected colours are consistent with the uncorrected sub-sample.

A possible remaining dependence of colours on the stretch factor s was investigated. For data points in a 5 days broad bin around time of B-maximum and around day 15, a linear

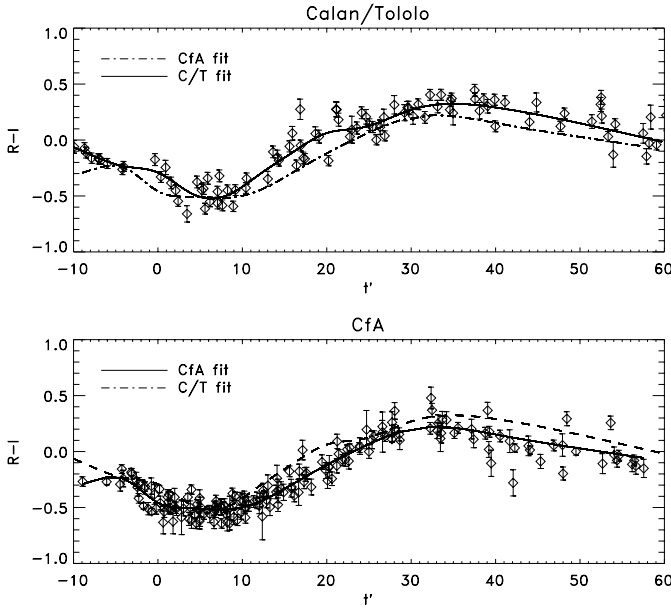


Fig. 5. $R - I$ for the two sets of data. The solid line represent always the curve found for the current set, while the dashed line is the curve found for the same colour of the other set.

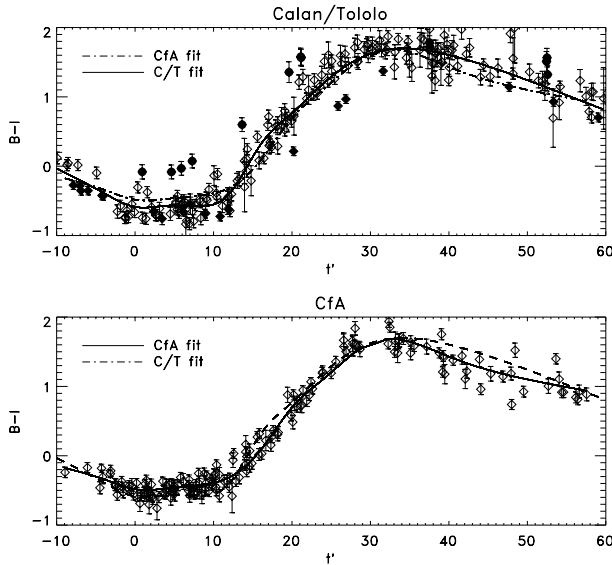


Fig. 6. $B - I$ for the two sets of data. The solid line represent always the curve found for the current set, while the dashed line is the curve found for the same colour of the other set. The filled circles represent SN 1993H, while the filled diamonds are data of SN 1992bc. Refer to Sect. 7 for a discussion about these 2 SNe.

weighted fit was done on the residuals from the models not showing any further significant dependence on Δm_{15} or equivalently on the stretch parameter s . Note that this is not in contradiction with Fig. 2, as we are considering $(B - V)_{t_{B_{\max}}}$ instead of $B_{\max} - V_{\max}$. Moreover, our analysis includes data points in a broad bin around day zero, and not a fitted lightcurve maximum as in Fig. 2.

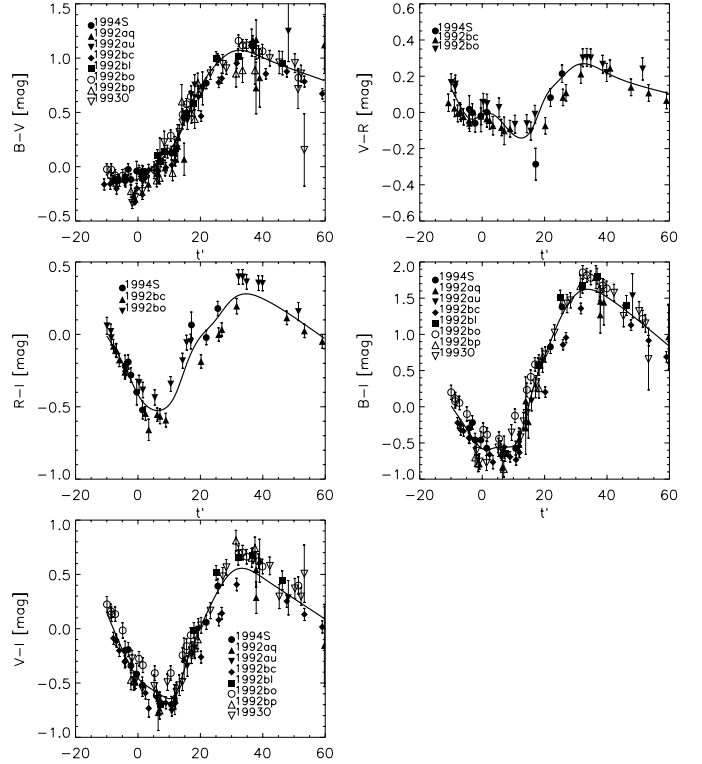


Fig. 7. Comparison of the models with data of unreddened SNe only.

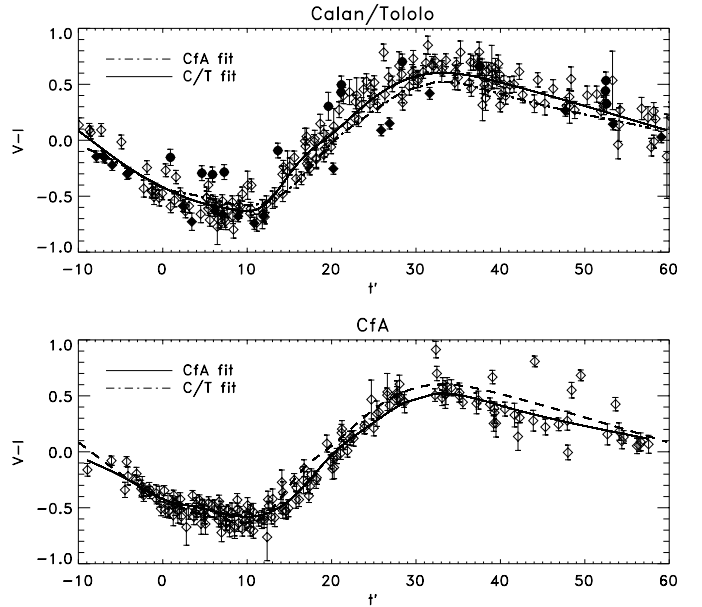


Fig. 8. $V - I$ for the two sets of data. The solid line represent always the curve found for the current set, while the dashed line is the curve found for the same colour of the other set. The filled circles represent SN 1993H, while the filled diamonds are data of SN 1992bc. Refer to Sect. 7 for a discussion about these 2 SNe.

4. Intrinsic colour dispersion

The observed colour dispersion around the *models* is overall inconsistent with a statistical random distribution of the data with the reported measurement errors. The χ^2 value per degree of freedom computed for each model is typically between 3

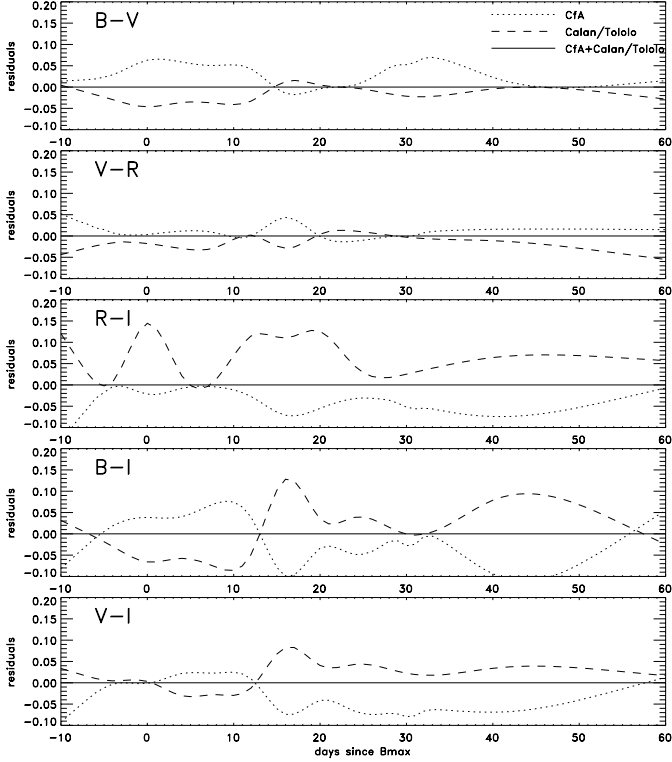


Fig. 9. Residuals of each of the models from the one built for all the data set.

and 6, giving indication of a very poor fit of the data. Thus, we conclude that the analysed data supports the existence of an *intrinsic colour dispersion* of Type Ia supernovae².

The intrinsic colour dispersion was computed on the residuals of each of the data sets from the corresponding *model*, built as explained above. Labeling XY_i any of the measured colours, $B - V$, $V - R$, $R - I$, $B - I$ and $V - I$, the residual with respect to the model expectation is referred as R_{XY_i} , i.e.,

$$R_{XY_i} = XY_i - XY_i^{\text{model}}. \quad (2)$$

Next, the time axis was binned into 5 days wide bins to compensate for the modest statistics. The bins are centered at days $d = 5 \cdot n$, with n ranging from 0 to 8. For each bin k , we compute the weighted average for the data XY , (see Appendix A)

$$XY^k = \frac{\sum_{i=1}^{N_k} w_i XY_i}{\sum_{i=1}^{N_k} w_i} \quad (3)$$

where N_k is the number of points in bin k and w_i is the inverse of the uncertainty on the i th measurement squared, $1/\sigma_i^2$, where σ_i include both measurement errors and uncertainty on the host galaxy extinction corrections. The weighted sample standard deviation was computed as the square root of the weighted second moment:

$$s_{XY}^k = \sqrt{m_{w_2}} = \sqrt{\frac{\sum_{i=1}^N w_i R_{XY_i}^2}{\sum_{i=1}^N w_i}}. \quad (4)$$

² This has been further investigated as described at the end of this section.

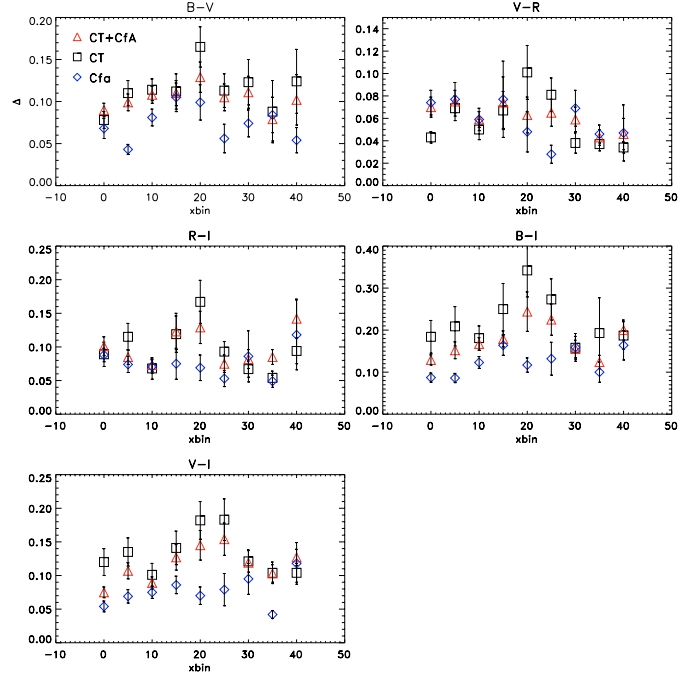


Fig. 10. Comparison of the results for the intrinsic dispersion in each colour given in Table 2 for both data sets (triangles) and the results obtained keeping the 2 sets of data separated (squares and diamonds for the C-T and CfA respectively).

The uncertainty on the expression in Eq. (4) is given by the square root of the variance, $V[s_{XY}^k]$, computed as:

$$V[s_{XY}^k] = \frac{V[m_{w_2}]}{4m_{w_2}} \quad (5)$$

where $V[m_{w_2}]$ is the variance of the weighted second moment. We can consider the result of Eq. (4) as an estimate of the intrinsic dispersion in each bin, for each XY colour: $\Delta_{XY} = s_{XY}^k$. This has been applied at each one of the sets separately and at the whole set of data. The results of the method for the whole set of data are given in Table 2.

Figure 10 shows a comparison of the intrinsic dispersion computed for each of the sets separately and their combination. In most of the cases the differences exceed the statistical uncertainties, seemingly arising from systematic differences between the two data sets. The computed intrinsic dispersion for the CfA sample was found to be smaller than for the CT data set and the whole (combined) sample in most cases, again pointing at systematic differences in the reported magnitudes.

Note that the adopted method leads to an overestimation of the intrinsic colour dispersion due to the contribution from the measurement errors. However, the weighting procedure ensures that the most accurate measurements dominate the result. To assess the impact of the measurement accuracy, we run a Monte-Carlo simulation to generate a synthetic colour data set with a dispersion given by the measurement uncertainties alone, i.e. no *intrinsic* dispersion. Three hundred data sets, with the same distributions in epochs and formal error bars as the CfA and Calan/Tololo were simulated and the weighted standard deviation (and its error) were computed, according to Eqs. (4) and (5). Note that the simulation, for simplicity,

Table 2. Results of the analysis of all SNe. First column: central value in days for each time bin; N_k is the number of points for each bin; XY^k is the weighted mean value and its 1-sigma uncertainty; Δ_{XY} is the intrinsic dispersion computed according to Eqs. (4) and (5); $\Delta_{XY}^{\text{corr}}$ is the corrected intrinsic dispersion, computed as in Eq. (6), and in the last column is the estimated lower limit at 99% C.L. (see text).

Day	N_k	BV^k	Δ_{BV}	$\Delta_{BV}^{\text{corr}}$	L.L.
0	57	-0.11 ± 0.01	0.09 ± 0.01	0.07	0.05
5	70	0.02 ± 0.01	0.10 ± 0.01	0.08	0.05
10	75	0.16 ± 0.01	0.11 ± 0.01	0.09	0.06
15	59	0.47 ± 0.02	0.11 ± 0.01	0.09	0.06
20	47	0.75 ± 0.03	0.13 ± 0.02	0.11	0.08
25	39	0.95 ± 0.02	0.10 ± 0.02	0.08	0.06
30	34	1.08 ± 0.02	0.11 ± 0.02	0.09	0.07
35 ^a	27	1.07 ± 0.02	0.08 ± 0.03	<0.05	
40	35	1.02 ± 0.02	0.10 ± 0.03	0.07	0.06
Day	N_k	VR^k	Δ_{VR}	$\Delta_{VR}^{\text{corr}}$	L.L.
0	31	0.04 ± 0.02	0.07 ± 0.01	0.06	0.05
5	36	-0.01 ± 0.02	0.08 ± 0.01	0.06	0.05
10	38	-0.08 ± 0.01	0.06 ± 0.01	0.04	0.04
15	34	-0.07 ± 0.02	0.07 ± 0.02	0.06	0.05
20	24	0.09 ± 0.02	0.06 ± 0.02	0.05	0.04
25 ^a	22	0.22 ± 0.02	0.06 ± 0.01	<0.04	
30 ^a	18	0.31 ± 0.01	0.06 ± 0.01	<0.04	
35 ^a	16	0.30 ± 0.01	0.04 ± 0.01	<0.03	
40 ^a	19	0.25 ± 0.01	0.05 ± 0.01	<0.03	
Day	N_k	RI^k	Δ_{RI}	$\Delta_{RI}^{\text{corr}}$	L.L.
0	30	-0.44 ± 0.02	0.10 ± 0.01	0.09	0.06
5	36	-0.50 ± 0.02	0.09 ± 0.01	0.07	0.05
10	32	-0.50 ± 0.02	0.07 ± 0.01	0.05	0.05
15	30	-0.25 ± 0.04	0.12 ± 0.03	0.11	0.08
20	20	-0.06 ± 0.04	0.13 ± 0.02	0.12	0.09
25	22	0.09 ± 0.02	0.08 ± 0.01	0.05	0.05
30	18	0.23 ± 0.03	0.08 ± 0.01	0.06	0.06
35 ^a	16	0.27 ± 0.03	0.09 ± 0.01	<0.06	
40	17	0.22 ± 0.04	0.14 ± 0.03	0.13	0.10
Day	N_k	BI^k	Δ_{BI}	$\Delta_{BI}^{\text{corr}}$	L.L.
0	39	-0.54 ± 0.02	0.13 ± 0.01	0.10	0.08
5	48	-0.52 ± 0.02	0.15 ± 0.02	0.13	0.09
10	48	-0.44 ± 0.03	0.17 ± 0.01	0.15	0.10
15	35	0.10 ± 0.05	0.18 ± 0.02	0.16	0.11
20	34	0.74 ± 0.06	0.24 ± 0.05	0.23	0.15
25	25	1.25 ± 0.05	0.22 ± 0.04	0.21	0.15
30	27	1.66 ± 0.03	0.16 ± 0.02	0.13	0.10
35 ^a	23	1.65 ± 0.03	0.12 ± 0.02	<0.08	
40	33	1.51 ± 0.04	0.20 ± 0.02	0.17	0.13
Day	N_k	VI^k	Δ_{VI}	$\Delta_{VI}^{\text{corr}}$	L.L.
0	37	-0.43 ± 0.02	0.08 ± 0.01	0.05	0.05
5	51	-0.54 ± 0.02	0.11 ± 0.01	0.09	0.06
10	48	-0.60 ± 0.01	0.09 ± 0.01	0.06	0.05
15	38	-0.34 ± 0.03	0.13 ± 0.02	0.11	0.08
20	36	0.01 ± 0.03	0.14 ± 0.02	0.13	0.09
25	28	0.33 ± 0.03	0.15 ± 0.02	0.14	0.10
30	28	0.57 ± 0.03	0.12 ± 0.02	0.10	0.08
35	25	0.58 ± 0.02	0.10 ± 0.01	0.08	0.07
40	33	0.48 ± 0.03	0.13 ± 0.01	0.11	0.08

^a Compatible with null intrinsic dispersion at 99% C.L.; an upper limit is given instead of the corrected intrinsic dispersion.

generates Gaussian distributed and completely uncorrelated data. The averages, δ , were used to disentangle the contribution of the intrinsic dispersion from the measurement errors. First, an hypothesis test was run to verify whether the simulated data and the measured data had the same dispersion; e.g. implying null intrinsic dispersion:

$$H_0 : \Delta = \delta$$

$$H_1 : \Delta \neq \delta.$$

A level of significance $\alpha = 0.01$ was set for rejecting the null hypothesis (Cowan 1998). Only 10 cases were not rejected, indicated by a ^(a) in Table 2. For all the other cases, for which the H_0 hypothesis was rejected, the intrinsic dispersion was computed as:

$$\Delta^{\text{corr}} = \sqrt{\Delta^2 - \delta^2} \quad (6)$$

and a lower limit on its value was set at a 99% confidence limit. The cases for which the null hypothesis was not rejected, were considered as compatible with *no intrinsic dispersion*, and an upper limit on its value was set at a 99% confidence level. The corrected intrinsic dispersions are listed in the 5th column of Table 2, together with upper and lower limits. We notice that the narrowest colour dispersion happens for $V - R$, especially at late times. At 25 days after B_{max} and later, this colour is compatible with no intrinsic spread at all. Further, it should be noted that at day 35, all the colours but $V - I$ are consistent with vanishing intrinsic dispersion.

5. Correlation

The correlation between optical colours at different epochs was also estimated. The property that was tested is whether a supernova that is blue at a certain epoch for example, say at maximum, stays blue at all epochs. In Riess et al. (2000), the authors argue that data measurements more than 3 days apart may be considered as uncorrelated estimators of colour. Our analysis does not support that assumption³. We find significant correlations for data points up to a month apart, as shown below. The method followed is essentially the one used to compute the intrinsic dispersion. One can summarize the following steps:

- Bin the data in time;
- Select only the SNe present in all the time bins;
- For each bin compute the weighted average of the measurements belonging to the same SN;
- Compute the linear correlation coefficient between bins as in Eq. (A.8);
- Test the correlation coefficient significance (Appendix A.2).

We refer to Appendix A.2 for what follows. The correlation coefficient between different epochs h and k is:

$$r_{hk} = \frac{\sum_{i=1}^n (R_{XY_i}^h - \bar{R}_{XY}^h) (R_{XY_i}^k - \bar{R}_{XY}^k)}{\sqrt{\sum_{i=1}^n (R_{XY_i}^h - \bar{R}_{XY}^h)^2 \sum_{j=1}^n (R_{XY_j}^k - \bar{R}_{XY}^k)^2}} \quad (7)$$

³ For high- z objects this is even more questionable when some of the main sources of uncertainty is the subtraction of a common image of the host galaxy and the K-corrections.

Table 3. Correlation coefficients between the different bins. The indicated $xbin$ is the central value of each bin. The bin size is 7 days for all the bins. The errors indicate the 1σ confidence level for the computed coefficients. See text for details.

$B - V/xbin$	0	7	15	22	30
0	1.00	$0.80^{+0.09}_{-0.14}$	$0.61^{+0.16}_{-0.24}$	$0.47^{+0.21}_{-0.28}$	$0.45^{+0.21}_{-0.28}$
7		1.00	$0.75^{+0.11}_{-0.17}$	$0.69^{+0.13}_{-0.20}$	$0.63^{+0.15}_{-0.23}$
15			1.00	$0.59^{+0.17}_{-0.24}$	$0.45^{+0.22}_{-0.28}$
22				1.00	$0.56^{+0.18}_{-0.25}$
30					1.00
$V - R/xbin$	0	7	15	22	30
0	1.00	$0.81^{+0.13}_{-0.31}$	$0.46^{+0.33}_{-0.54}$	$0.29^{+0.41}_{-0.56}$	$0.03^{+0.51}_{-0.53}$
7		1.00	$0.24^{+0.44}_{-0.56}$	$0.48^{+0.32}_{-0.53}$	$0.24^{+0.44}_{-0.56}$
15			1.00	$0.48^{+0.32}_{-0.53}$	$-0.15^{+0.55}_{-0.47}$
22				1.00	$-0.44^{+0.55}_{-0.34}$
30					1.00
$R - I/xbin$	0	7	15	22	30
0	1.00	$0.73^{+0.20}_{-0.51}$	$0.57^{+0.31}_{-0.63}$	$0.56^{+0.31}_{-0.63}$	$0.31^{+0.46}_{-0.68}$
7		1.00	$0.66^{+0.25}_{-0.57}$	$0.78^{+0.16}_{-0.45}$	$0.58^{+0.30}_{-0.63}$
15			1.00	$0.89^{+0.08}_{-0.28}$	$0.53^{+0.33}_{-0.65}$
22				1.00	$0.83^{+0.13}_{-0.38}$
30					1.00
$V - I/xbin$	0	7	15	22	30
0	1.00	$0.84^{+0.08}_{-0.17}$	$0.92^{+0.04}_{-0.09}$	$0.93^{+0.04}_{-0.08}$	$0.49^{+0.25}_{-0.36}$
7		1.00	$0.72^{+0.15}_{-0.26}$	$0.74^{+0.14}_{-0.25}$	$0.50^{+0.24}_{-0.36}$
15			1.00	$0.90^{+0.06}_{-0.12}$	$0.49^{+0.25}_{-0.36}$
22				1.00	$0.65^{+0.18}_{-0.30}$
30					1.00
$B - I/xbin$	0	7	15	22	30
0	1.00	$0.62^{+0.20}_{-0.35}$	$0.74^{+0.14}_{-0.27}$	$0.44^{+0.29}_{-0.42}$	$0.30^{+0.34}_{-0.44}$
7		1.00	$0.76^{+0.13}_{-0.26}$	$0.47^{+0.27}_{-0.41}$	$0.51^{+0.25}_{-0.39}$
15			1.00	$0.76^{+0.14}_{-0.26}$	$0.40^{+0.30}_{-0.42}$
22				1.00	$0.73^{+0.15}_{-0.28}$
30					1.00

where the summation is on the i th SN, which by construction is present in all the bins. The uncertainty on r_{hk} has been computed converting it into the normally distributed variable z , as described in Appendix A.2. The results, given in Table 3, show that the correlation is important and non-zero all along the time evolution. Figure 11 shows the weighted mean colours for the selected SNe in the bins centered at day 0, 15 and 30. Note that the supernovae selected are different in different colours, but, by construction, are the same in all the bins for each of the colour. It appears that supernovae that deviate from the average colour at a certain epoch are likely to keep their colour excess all along the 30 days evolution considered here.

6. Supernovae template

Nugent et al. (2002) analysed the relation between colours of SNe Ia and K-corrections. In particular they showed how K-corrections are mainly driven by the overall colour of the SN rather than by peculiarities of single features. Using spectra and

colour light curves they give a recipe to build a *SN Ia template*, to be used for computing K-corrections of $s \sim 1.0$ SNe. As the sample used in our work is larger than the one considered in their paper, we have used our results to improve the spectral template. Note that this affects only the *BVRI* magnitudes, which are the only bands treated here. Referring to Nugent et al. (2002), we proceed as follows:

- correct the *BVRI* of their list of *UBVRIJHK* magnitudes using our *B - V*, *V - R* and *R - I models*;
- correct the corresponding spectral templates with this colours⁴;
- iterate the whole analysis described in this paper, using the newly created templates to compute the K-corrections for both Calan/Tololo and CfA data set;
- correct again the templates using the most recent version of the *models*.

⁴ The same code used by Nugent et al. (2002) was used for this.

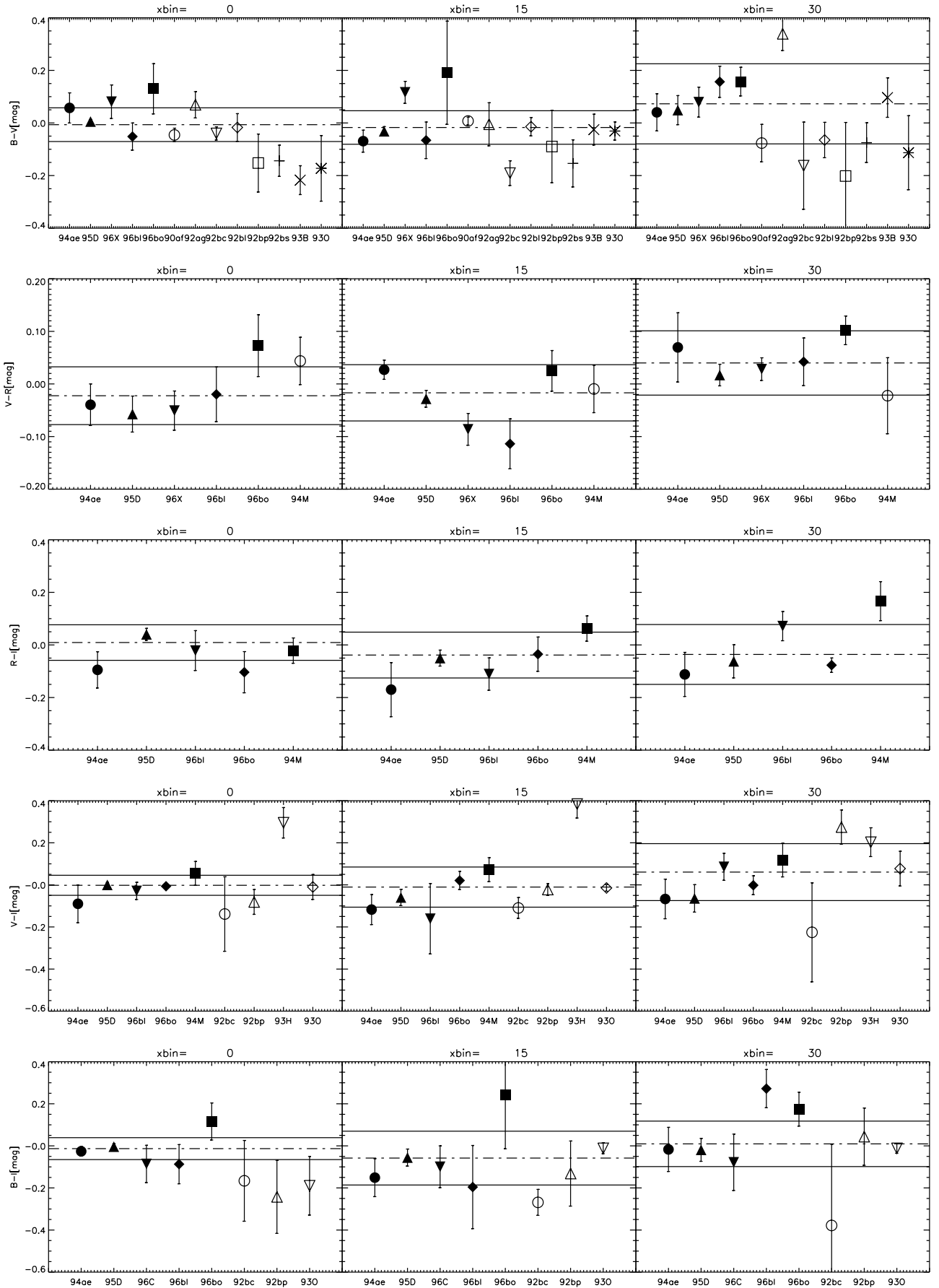


Fig. 11. From top to bottom and left to right. Residuals from the models for the SNe selected for each colour, in bin centered at day 0,15 and 30. The band defined by the solid lines correspond to the intrinsic dispersion found in at the same epoch.

Table 4. Results of the analysis done excluding SN 1992bc and SN 1993H (***) for $B - I$ and $V - I$. $xbin$ is the central value in days for each bin; XY^{k**} is the weighted mean value and its uncertainty; Δ_{BI}^{**} is computed according to Eqs. (4) and (5); Δ_{XY}^{corr**} is the corrected intrinsic dispersion and L.L. is the lower limit at 99% C.L.; in the last column is the difference between the results given in Table 2 and this analysis; the errors are computed as sum in quadrature.

day	BI^{k**}	Δ_{BI}^{**}	Δ_{BI}^{corr**}	L.L.	$\Delta_{BI} - \Delta_{BI}^{**}$
0	-0.53 ± 0.02	0.13 ± 0.01	0.10	0.08	0.000 ± 0.017
5	-0.53 ± 0.02	0.13 ± 0.01	0.10	0.08	0.024 ± 0.025
10	-0.41 ± 0.02	0.15 ± 0.02	0.13	0.09	0.018 ± 0.022
15	0.10 ± 0.05	0.18 ± 0.02	0.16	0.11	0.003 ± 0.026
20	0.74 ± 0.05	0.15 ± 0.02	0.12	0.09	0.098 ± 0.052
25	1.31 ± 0.05	0.12 ± 0.02	0.09	0.08	0.101 ± 0.044
30	1.68 ± 0.03	0.12 ± 0.02	0.08	0.08	0.038 ± 0.028
35	1.65 ± 0.03	0.13 ± 0.02	<0.09		-0.006 ± 0.024
40	1.49 ± 0.04	0.20 ± 0.03	0.18	0.13	-0.004 ± 0.035
day	VI^{k**}	Δ_{VI}^{**}	Δ_{VI}^{corr**}	L.L.	$\Delta_{VI} - \Delta_{VI}^{**}$
0	-0.43 ± 0.02	0.07 ± 0.01	0.05	0.05	0.002 ± 0.011
5	-0.55 ± 0.01	0.09 ± 0.01	0.06	0.05	0.019 ± 0.014
10	-0.58 ± 0.01	0.09 ± 0.01	0.06	0.05	0.002 ± 0.013
15	-0.34 ± 0.03	0.11 ± 0.01	0.09	0.07	0.016 ± 0.022
20	0.01 ± 0.03	0.11 ± 0.02	0.09	0.07	0.036 ± 0.027
25	0.35 ± 0.03	0.09 ± 0.01	0.07	0.06	0.060 ± 0.028
30	0.57 ± 0.03	0.11 ± 0.02	0.09	0.07	0.007 ± 0.029
35	0.58 ± 0.02	0.10 ± 0.01	0.08	0.07	0.000 ± 0.019
40	0.47 ± 0.03	0.13 ± 0.01	0.11	0.08	0.000 ± 0.018

Even though the template itself is modified considerably from the original one, iterating once results in small corrections (~ 0.001 – 0.005 mag) in the intrinsic dispersion in all colours at any epoch, and a few percent in the colour *models*. As the correction obtained are small and well within the given uncertainties, we do not iterate further. The results of this analysis should *not* be considered definitive and may change as SNe Ia and their colour evolutions are studied in even more detail. However, it can be considered a good estimate of the average SN Ia template (in BVI), as it is constructed from a quite broad sample of SNe. The final corrected template is available upon request.

7. Discussion

Throughout our analysis all data points were treated as independent measurements. This is particularly important for the cases with significant host galaxy light underneath the supernova where a single reference image was used, introducing a correlation between the data points not considered here.

A first estimate of the intrinsic dispersion was calculated for each data set separately and the two data sets together. A comparison of the results shows that the colours extracted from the CfA sample have smaller scatter indicating that the contribution from measurement errors is not negligible. Further, in the analysis the measurements are assumed to be Gaussian distributed and the weighted standard deviation has been taken as an estimate of the intrinsic dispersion. However, we noticed that 2 SNe, SN 1993H and SN 1992bc, seem to be rather deviant in $B - I$ and $V - I$ for the C-T set, as shown in Figs. 6

and 8. The effect of the “outliers” is particularly important around 20 days after B -maximum. For comparison, we recalculated the intrinsic spread in $B - I$ and $V - I$ excluding these 2 SNe. The results are shown on Table 4. The agreement between the intrinsic dispersion between the data sets improves when these 2 SNe are excluded. We emphasize that there are systematic differences between the Calan/Tololo and CfA data sets, and they might have been introduced while converting from the instrumental system used to the standard BVI system.

An attempt of disentangling the intrinsic dispersion from the contribution of the uncertainties has been done resulting in upper limit values for the intrinsic dispersion in some cases. The most relevant is the case of $V - R$, that seems compatible with null intrinsic dispersion for most of the epochs at 99% C.L. Moreover this analysis brings out an important feature at day 35, when all the colours but $V - I$ are consistent with zero intrinsic dispersion. This indicates that further studies will be needed to investigate this intriguing finding. When computing the correlation coefficients we considered data up to 30 days after B -band maximum, even though later epoch data are available for several supernovae. However, the necessary condition of each SN being observed in all the time bins reduces the statistics if later epochs are introduced. This limitation is specific to the sample used and can be overcome with the use of a more extensively observed sample, such as what will be provided by SNfactory (Aldering et al. 2002).

The intrinsic dispersion sets constraints on the ability to determine the host galaxy extinction, A_V . This will depend on

the colour, as the intrinsic dispersion is different for different colours. As an example we used the Cardelli et al. (1989) relation at the effective wavelength for each bandpass to compute the expected uncertainty in the extinction, neglecting any dependence on the supernova phase.

$$\begin{aligned}
 \sigma_{A_V}^{B-V} &= 3.1 \cdot \Delta_{BV}^{\text{corr}} \\
 \sigma_{A_V}^{V-R} &= 6.2 \cdot \Delta_{VR}^{\text{corr}} \\
 \sigma_{A_V}^{R-I} &= 4.1 \cdot \Delta_{RI}^{\text{corr}} \\
 \sigma_{A_V}^{B-I} &= 1.4 \cdot \Delta_{BI}^{\text{corr}} \\
 \sigma_{A_V}^{V-I} &= 2.5 \cdot \Delta_{VI}^{\text{corr}}
 \end{aligned} \tag{8}$$

where R_V was assumed equal to 3.1. Table 5 shows the results of the Eq. (8) for the epochs for which the intrinsic dispersion was calculated. The $\Delta_{XY}^{\text{corr}}$ was used for this. The results indicate that, with the present knowledge, extinction by dust with $R_V = 3.1$ may only be determined to $\sigma_{A_V} \gtrsim 0.10$ with Type Ia restframe optical data within the first 40 days after B -band lightcurve maximum, for the colours and epochs with non-zero intrinsic dispersion. However observations in $V - R$ are preferable to other colours to set limit on the extinction of an observed supernova. To account for possible different extinction parameters R_V in the different supernova host galaxies (as noticed in e.g. Riess et al. 1996 and Krisciunas et al. 2000), we repeated the analysis considering a Gaussian uncertainty $\sigma_{R_V} = 1$ on R_V , propagating this scatter on the host galaxy extinction corrections. This changes the values of the intrinsic dispersion, but typically within the quoted errors in Table 2.

8. Conclusion

A statistical analysis of colours of SNe Ia using a sample of 48 nearby SNe was performed. Of this sample 7 SNe have been excluded based on their large extinction, or peculiar behavior of their colours at maximum. With the present knowledge and data quality we computed the average colour evolution for $B - V$, $V - R$, $R - I$, $B - I$ and $V - I$ with time and the derived intrinsic scatter. We find that the correlation of colours during the first 30 days after restframe B -band maximum is not negligible, i.e. arguing against the assumptions made in the analysis of SN 1999Q ($z = 0.46$; Riess et al. 2000) where five measurements of restframe $B - I$ along its lightcurve were treated as *independent* estimates of extinction. According to our findings, their limit on the presence of intergalactic grey dust must be revised. A reanalysis of the data from SN 1999Q is in preparation (Nobili et al. 2003). With the data at hand, host galaxy extinction corrections from restframe optical colours within the first 30 days after maximum light are generally limited to $\sigma_{A_V} \gtrsim 0.1$ due to the intrinsic variation of Type Ia colours at those epochs, with the possible exception of extinction corrections derived from the rest-frame $V - R$ colour. The results of this analysis have been used to correct spectroscopic templates which are available upon request.

Acknowledgements. S. Nobili wishes to thank Christian Walck for his valuable help and constructive discussions on the statistical analysis presented in this paper and Don Groom, Chris Lidman and

Table 5. σ_{A_V} represents the constrain on the extinction A_V that it is possible to compute for a given intrinsic dispersion. The $\Delta_{XY}^{\text{corr}}$ and its L.L. given in Table 2 have been used for these estimates.

	xbin	σ_{A_V}	L.L.
$B - V$	0	0.23	0.16
$B - V$	5	0.25	0.17
$B - V$	10	0.28	0.18
$B - V$	15	0.28	0.20
$B - V$	20	0.34	0.24
$B - V$	25	0.25	0.20
$B - V$	30	0.27	0.22
$B - V$	35	<0.16	
$B - V$	40	0.21	0.20
$V - R$	0	0.36	0.28
$V - R$	5	0.38	0.29
$V - R$	10	0.27	0.22
$V - R$	15	0.36	0.29
$V - R$	20	0.29	0.26
$V - R$	25	<0.27	
$V - R$	30	<0.25	
$V - R$	35	<0.19	
$V - R$	40	<0.20	
$R - I$	0	0.37	0.27
$R - I$	5	0.28	0.22
$R - I$	10	0.22	0.19
$R - I$	15	0.46	0.32
$R - I$	20	0.48	0.36
$R - I$	25	0.22	0.21
$R - I$	30	0.25	0.23
$R - I$	35	<0.25	
$R - I$	40	0.54	0.41
$V - I$	0	0.12	0.12
$V - I$	5	0.22	0.16
$V - I$	10	0.16	0.13
$V - I$	15	0.27	0.20
$V - I$	20	0.32	0.22
$V - I$	25	0.35	0.25
$V - I$	30	0.25	0.19
$V - I$	35	0.19	0.17
$V - I$	40	0.26	0.20
$B - I$	0	0.14	0.11
$B - I$	5	0.18	0.13
$B - I$	10	0.20	0.14
$B - I$	15	0.22	0.16
$B - I$	20	0.32	0.21
$B - I$	25	0.29	0.21
$B - I$	30	0.18	0.14
$B - I$	35	<0.12	
$B - I$	40	0.24	0.18

Vallery Stanishev for their careful reading and useful suggestions. SN is supported by a graduate student grant from the Swedish Research Council. AG is a Royal Swedish Academy Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation. PEN acknowledges support from a NASA LTSA grant and by the Director, Office of Science under U.S. Department of Energy Contract No. DE-AC03-76SF00098.

Appendix A

A.1. Algebraic moments

In what follows we refer to Cowan (1998) and Kendall & Stuart (1958, see Vol. 1, Chap. 10), our aim here is to give uniformity to the notations used. Given n independent observations of a variable, x_1, x_2, \dots, x_n , the r th moment or algebraic moment is given by:

$$m_r = \frac{1}{n} \sum_{i=1}^n x_i^r. \quad (\text{A.1})$$

The expectation value, or mean value, and the variance of m_r are:

$$\begin{aligned} E[m_r] &= \mu_r \\ V[m_r] &= \frac{1}{n} (\mu_{2r} - \mu_r^2) \end{aligned} \quad (\text{A.2})$$

where μ_{2r} is the $2 \cdot r$ th moment. Equations (A.2) are exact formulae as long as one knows μ_{2r} and μ_r . However this is not always the case, and one has to use their estimators m_{2r} and m_r from the sample itself. The variance of the standard deviation can be computed using standard error propagation:

$$V[\sigma] = V[\sqrt{m_2}] = \frac{V[m_2]}{4m_2} = \frac{m_4 - m_2^2}{4nm_2}. \quad (\text{A.3})$$

In experimental situation each observation x_i is often attached to a certain weight w_i . Supposing that the weight themselves are known without errors, one can define the following formula for the r th weighted moments:

$$m_{w,r} = \frac{\sum_{i=1}^n w_i x_i^r}{\sum_{i=1}^n w_i}. \quad (\text{A.4})$$

In order to evaluate the variance on the weighted standard deviation we simply extended Eq. (A.3) for the case of weighted moments, obtaining:

$$V[\sigma_w] = V[\sqrt{m_{w_2}}] = \frac{V[m_{w_2}]}{4m_{w_2}} = \frac{m_{w_4} - m_{w_2}^2}{4n_{\text{eff}} m_{w_2}} \quad (\text{A.5})$$

where the m_{w_4} and m_{w_2} are the 4th and 2nd weighted moments respectively and $n_{\text{eff}} = (\sum_{i=1}^n w_i)^2 / \sum_{i=1}^n w_i^2$.

A simple Monte Carlo simulation has been run to verify the accuracy of the approximated formula (A.5).

A.2. Correlation coefficient

Given 2 random variables x and y the correlation coefficient is defined as

$$\rho_{xy} = \frac{V_{xy}}{\sigma_x \sigma_y}. \quad (\text{A.6})$$

The unbiased estimator of the covariance V_{xy} is:

$$\widehat{V}_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (\text{A.7})$$

so that the estimator of the correlation coefficient will be:

$$r_{xy} = \frac{\widehat{V}_{xy}}{s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{j=1}^n (y_j - \bar{y})^2}}. \quad (\text{A.8})$$

For Eq. (A.8) to be a good estimator of the correlation coefficient there are a few caveats to check. First condition is that the samples are randomly defined from the population, that is to make sure that the samples x and y are not selected in some ways that would operate as to increase or decrease the value of r . Even though one has random samples it is possible to compute the errors due to sampling. Commonly this is computed as $(1 - r^2) / \sqrt{n}$. Unfortunately this is just an approximation. Moreover r 's for successive samples are not distributed normally unless n is large and the true value ρ is near zero. This yields to a distinction:

- if $n > 30$:
in order to know whether the value calculated for r is significantly different from zero, one can compute its standard error as:

$$\sigma_r = 1 / \sqrt{n-1}. \quad (\text{A.9})$$

If r/σ is greater than 2.58, one can conclude that the universe value of r is likely to be greater than zero.

- if $n < 30$:
the variable

$$t = r \frac{\sqrt{n-2}}{\sqrt{1-r^2}} \quad (\text{A.10})$$

follows the t -distribution with $d.o.f. = n - 2$. This can be used only for testing the hypothesis of zero correlation.

R. A. Fisher developed a technique to overcome these difficulties. The variable r is transformed into another variable that is normally distributed. This is especially useful for high value of r , when none of the above test can be safely applied. The transformation to the variable z :

$$z = \frac{1}{2} \ln(1+r) - \frac{1}{2} \ln(1-r) \quad (\text{A.11})$$

allows some important simplifications. The distribution of z 's for successive samples does not depend on the universe value ρ and the distribution of z for successive samples is so near to normal that it can be treated as such without any loss of accuracy, see for example (Vol. 1, Chap. 16 and Vol. 2, Chap. 26) of Kendall & Stuart (1958). Moreover the standard error for z is independent on its σ :

$$\sigma_z = 1 / \sqrt{n-3}. \quad (\text{A.12})$$

The way to proceed is then very simple:

- Compute r according to (A.8);
- Transform r into z according to (A.11);
- Compute σ_z according to (A.12);
- In order to state a 1 sigma confidence limit for r , transform the 2 values $z \pm \sigma_z$ back to r , using the inverse of (A.11).

References

- Aguirre, A. 1999, *ApJ*, 512, L19
Aguirre, A. 1999, *ApJ*, 525, 583
Aguirre, A., & Haiman, Z. 2000, *ApJ*, 532, 28
Aldering, G., et al. 2002, Supernova Factory Webpage,
<http://snfactory.lbl.gov>
Branch, D., Fisher, A., & Nugent, P. 1993, *AJ*, 106, 2383
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245
Cowan, G. 1998, *Statistical data analysis* (Oxford University Press)
Cramer, H. 1957, *Mathematical methods of statistics* (Princeton University Press, Seventh Printing)
Goldhaber, G., Groom, D. E., Kim, A., et al. 2001, *ApJ*, 558, 359
Goobar, A., Mörtzell, E., Amanullah, R., et al. 2002, *A&A*, 392, 757
Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 1996, *AJ*, 112, 2408
Kendall, M. G., & Stuart, A. 1958, *The advanced theory of statistics* (London: Charles Griffin & Company Limited)
Krisciunas, K., Hastings, N. C., Loomis, K., et al. 2000, *ApJ*, 539, 658
Lira, P. 1995, Master Thesis, Univ. Chile
Mörtzell, E., Bergström, L., & Goobar, A. 2002, *Phys. Rev. D*, 66, 047702
Nobili, S., et al. 2003, in prep.
Nugent, P., Kim, A., & Perlmutter, S. 2002, *PASP*, 114, 803
O'Donnell, J. E. 1994, *ApJ*, 422, 1580
Perlmutter, S., Gabi, S., Goldhaber, G., et al. 1997, *ApJ*, 483, 565P
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, 517, 565P
Phillips, M. M. 1993, *ApJ*, 413, L105
Phillips, M. M., Lira, P., Suntzeff, N. B., et al. 1999, *ApJ*, 118, 1766
Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, *ApJ*, 473, 588R
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, 116, 1009
Riess, A. G., Kirshner, R. P., Schmidt, B. P., et al. 1999, *ApJ*, 117, 707-724
Riess, A. G., Filippenko, A. V., Liu, M. C., et al. 2000, *ApJ*, 536, 62