



HAL
open science

SNEMO: Improved Empirical Models for Type Ia Supernovae

C. Saunders, G. Aldering, P. Antilogus, S. Bailey, C. Baltay, K. Barbary, D. Baugh,
K. Boone, S. Bongard, C. Buton, et al.

► **To cite this version:**

C. Saunders, G. Aldering, P. Antilogus, S. Bailey, C. Baltay, et al.. SNEMO: Improved Empirical Models for Type Ia Supernovae. *The Astrophysical Journal*, 2018, 869 (2), pp.167. <10.3847/1538-4357/aac7e>. <hal-01921643>

HAL Id: hal-01921643

<https://hal.science/hal-01921643v1>

Submitted on 19 Aug 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.





L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



SNEMO: Improved Empirical Models for Type Ia Supernovae

C. Saunders^{1,2,3}, G. Aldering³, P. Antilogus¹, S. Bailey³, C. Baltay⁴, K. Barbary⁵, D. Baugh⁶, K. Boone^{3,5} , S. Bongard¹, C. Buton^{7,8}, J. Chen⁶, N. Chotard^{7,8}, Y. Copin^{7,8}, S. Dixon^{3,5}, P. Fagrelus^{3,5}, H. K. Fakhouri^{3,5}, U. Feindt⁹, D. Fouchez¹⁰, E. Gangler¹¹, B. Hayden³ , W. Hillebrandt¹², A. G. Kim³, M. Kowalski^{13,14}, D. Küsters¹³, P.-F. Leget¹¹, S. Lombardo¹³, J. Nordin¹³, R. Pain¹, E. Pecontal¹⁵, R. Pereira^{7,8}, S. Perlmutter^{3,5}, D. Rabinowitz⁴, M. Rigault¹¹ , D. Rubin^{3,16} , K. Runge³, G. Smadja^{7,8}, C. Sofiatti^{3,5}, N. Suzuki³, C. Tao^{6,10}, S. Taubenberger¹², R. C. Thomas¹⁷, and M. Vincenzi³
(The Nearby Supernova Factory)

¹ Sorbonne Université, Université Paris Diderot, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Énergies, 4 Place Jussieu, Paris, France

² Sorbonne Université, Institut Lagrange de Paris (ILP), 98 bis Boulevard Arago, F-75014 Paris, France

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

⁴ Department of Physics, Yale University, New Haven, CT 06250-8121, USA

⁵ Department of Physics, University of California Berkeley, 366 LeConte Hall MC 7300, Berkeley, CA 94720-7300, USA

⁶ Tsinghua Center for Astrophysics, Tsinghua University, Beijing 100084, People's Republic of China

⁷ Université de Lyon, F-69622, Lyon, France

⁸ Université de Lyon 1, Villeurbanne; CNRS/IN2P3, Institut de Physique Nucléaire de Lyon, France

⁹ The Oskar Klein Centre, Department of Physics, AlbaNova, Stockholm University, SE-106 91 Stockholm, Sweden

¹⁰ Centre de Physique des Particules de Marseille, Aix-Marseille Université, CNRS/IN2P3, 163 Avenue de Luminy—Case 902-13288 Marseille Cedex 09, France

¹¹ Clermont Université, Université Blaise Pascal, CNRS/IN2P3, Laboratoire de Physique Corpusculaire, BP 10448, F-63000 Clermont-Ferrand, France

¹² Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748 Garching, Germany

¹³ Institut für Physik, Humboldt-Universität zu Berlin, Newtonstr. 15, D-12489 Berlin, Germany

¹⁴ DESY, D-15735 Zeuthen, Germany

¹⁵ Centre de Recherche Astronomique de Lyon, Université Lyon 1, 9 Avenue Charles André, F-69561 Saint Genis Laval Cedex, France

¹⁶ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

¹⁷ Computational Cosmology Center, Computational Research Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road MS 50B-4206, Berkeley, CA 94720, USA

Received 2018 July 11; revised 2018 October 5; accepted 2018 October 12; published 2018 December 21

Abstract

SN Ia cosmology depends on the ability to fit and standardize observations of supernova magnitudes with an empirical model. We present here a series of new models of SN Ia spectral time series that capture a greater amount of supernova diversity than is possible with the models that are currently customary. These are entitled SuperNova Empirical MOdels (SNEMO; <https://snfactory.lbl.gov/snemo>). The models are constructed using spectrophotometric time series from 172 individual supernovae from the Nearby Supernova Factory, comprising more than 2000 spectra. Using the available observations, Gaussian processes are used to predict a full spectral time series for each supernova. A matrix is constructed from the spectral time series of all the supernovae, and Expectation Maximization Factor Analysis is used to calculate the principal components of the data. K -fold cross-validation then determines the selection of model parameters and accounts for color variation in the data. Based on this process, the final models are trained on supernovae that have been dereddened using the Fitzpatrick and Massa extinction relation. Three final models are presented here: SNEMO2, a two-component model for comparison with current Type Ia models; SNEMO7, a seven-component model chosen for standardizing supernova magnitudes, which results in a total dispersion of 0.100 mag for a validation set of supernovae, of which 0.087 mag is unexplained (a total dispersion of 0.113 mag with an unexplained dispersion of 0.097 mag is found for the total set of training and validation supernovae); and SNEMO15, a comprehensive 15-component model that maximizes the amount of spectral time-series behavior captured.

Key words: cosmology: observations – supernovae: general

Supporting material: machine-readable table

1. Introduction

Supernovae (SNe) Type Ia have provided a powerful tool for measuring the accelerating expansion of the universe and constraining the properties of dark energy (Riess et al. 1998; Perlmutter et al. 1999). However, future progress will be limited by various uncertainties in their standardized magnitudes (Saunders et al. 2015, hereafter S15). While current and upcoming supernova surveys (DES, Bernstein et al. 2012; LSST, LSST Science Collaboration et al. 2009) will eliminate any significant source of statistical uncertainty, along with certain calibration errors, remaining intrinsic differences in the

supernovae still lead to a large amount of systematic uncertainty. Crucially, the presence of intrinsic dispersion in standardized magnitudes indicates the existence of latent unmodeled supernova processes, which can cause bias in standardized magnitudes if they are dependent on redshift or selection effects.

On the question of standardized magnitudes, progress should be possible. SN Ia standardization was originally made viable by the Phillips relation (Phillips 1993), which identified the relation between supernova magnitudes and the decay rate of the light curve. This was improved further by Riess et al. (1996) and Tripp (1998), who added a correction to the

supernova magnitude based on the difference between the B - and V -band magnitudes at maximum light, in other words the $B - V$ color. Data could be K -corrected with supernova templates, such as Nugent et al. (2002), and the B -band magnitudes could be corrected by the light-curve width and color relations in order to be used for cosmology. Current techniques for standardization still rely on simple empirical supernova models that have only one or two degrees of freedom (SALT2, Guy et al. 2007, MLCS2K2, Jha et al. 2007, SNOOPY, Burns et al. 2011). After standardization, there is a remaining dispersion in supernova magnitudes of about 0.15 mag (Betoule et al. 2014), which is a combination of scatter due to calibration errors and intrinsic and extrinsic sources (e.g., dust) of variation in supernova magnitudes. Alternative methods for using SN Ia as standard candles, such as finding “twin” supernovae that have identical spectra (Fakhouri et al. 2015, hereafter F15) or using the near-infrared maximum magnitude (Krisciunas et al. 2004; Mandel et al. 2011; Barone-Nugent et al. 2012), are restricted by the type of data required, which is not available for most existing supernova surveys.

Recent work has shown that SNe Ia exhibit more diversity than can be described by two components (Kim et al. 2013, S15, Sasdelli et al. 2015, etc.), suggesting that the supernova-dependent sources of dispersion in the standardized magnitudes can be reduced by using an improved empirical model for SNe Ia. Moreover, differences in the supernovae can be dependent on their environment, and the distribution of supernovae may shift with evolution in the environments with respect to redshift (Kelly et al. 2010; Sullivan et al. 2010; Gupta et al. 2011; Rigault et al. 2013, 2015, 2018). A model that can account for more of these differences among supernovae will reduce the potential for bias in the standardized magnitudes if the population changes with the redshift.

Since $\sim 85\%$ of a supernova’s bolometric luminosity is emitted in the optical range (Howell et al. 2009), a supernova model covering this range, as presented here, will cover the majority of the region where supernova variability is expressed. Further, the spectra in this optical range indicate which ions are present in the supernova ejecta, along with their velocities, as well as determining the broad-band shape of the supernova light curves. We are motivated by the idea that these indicators uniquely determine the luminosity of the supernova, meaning that two supernovae with the same spectral time series will have the same luminosity. This hypothesis is supported by F15, where a dispersion of 0.072 mag was attained by pairing supernovae with very similar spectral time series (see Foley et al. 2018 for an apparent exception). Thus, ideally, the better the supernova magnitudes can be modeled, the better they can be standardized. For these reasons, better spectrophotometric models are needed in order to extract the greatest possible amount of information from SNe Ia observations.

For the empirical models currently in use, the complexity that could be incorporated in their training is constrained by the limited availability of supernova observations, in particular the small numbers of homogeneous spectral time series uncontaminated by the host-galaxy light. With spectrophotometric time-series data from the Nearby Supernova Factory (SNfactory, Aldering et al. 2002), a much more complex empirical model can be built. The SNfactory data set includes hundreds of flux-calibrated, host-subtracted spectral time series that can

be used to build a more complete picture of SNe Ia because there is so much more information per supernova. To equivalently constrain a model with photometry, a photometric data set orders of magnitude larger and at a wide range of redshifts would be needed in order to deduce what spectral behavior drives variation in the observed broad-band measurements and to break degeneracies between different spectral behavior that result in identical broad-band behavior when integrated over the wavelength range of a filter. However, degeneracies between broad-band colors and dust extinction, along with redshift and cosmology, may mean that a photometric data set would not be able to replace spectral data regardless of its size.

Given the nature of variations among spectra, such as the strengths and velocities of emission and absorption features, it may be that supernova spectral time series can only be perfectly described by complex nonlinear models or by explosion and radiative transfer simulations, which are currently far from obtaining the accuracy needed for supernova cosmology. However, in this analysis, we aim to make the next logical extension beyond the current generation of empirical models. The SALT2 model is essentially performing Principal Component Analysis (PCA), finding the first two eigenvectors that describe the space of SN Ia light curves or spectral time series. These are combined with a color relation, which is either explicitly in the form of reddening due to dust, or else is intended to describe reddening due to a combination of dust and some type of intrinsic color variation. The simplest improvement to this approach is to find the remaining significant eigenvectors, and thus build a model that describes a larger fraction of the variation in Type Ia supernovae. The data from the SNfactory allows us to do this by performing PCA on a matrix made up of our collection of spectral time series. The resulting principal components, which are ranked eigenvectors of the matrix, become the linear components of a new spectral-temporal model. Using the model to fit new supernovae provides a method for standardizing their magnitudes, since correlations between the peak magnitude and fit coefficients can be calculated and corrected for.

Other methods for modeling SNe Ia have found alternative ways of reducing the dimensionality of the population space. The SUGAR model (P.-F. Léget et al. 2018, in preparation), also trained on the SNfactory supernova sample, is an example of one of these methods. This model focuses on the spectral features at maximum, which strongly impact the photometric magnitudes. SUGAR finds the principal components in the space of spectral features, and then relates those to the full spectral time series in order to build a model. Whether and in what cases this model may be preferable to the ones presented here is left to future testing.

We present the SNfactory data used here in Section 2. The method used to calculate the model components is discussed in Section 3. Three final Supernova Empirical Models (SNEMO) are presented in Section 4: SNEMO2, a simple model to compare with SALT2 and other current empirical supernova models, SNEMO7, a model chosen for standardizing supernova magnitudes, and SNEMO15, a model intended to capture the greatest possible amount of SN Ia behavior. In this section, empirical diagnostics of the model components are measured, and the models are compared with SALT2. Concluding remarks are given in Section 5.

2. The SN Ia Data Set

The supernova data used in this analysis are spectrophotometric time series observed by the SNfactory between 2004 and 2014 with the SuperNova Integral Field Spectrograph (SNIFS, Lantz et al. 2004). Many of these supernovae have been presented previously, for example in Thomas et al. (2007, 2011), Bailey et al. (2009), Chotard et al. (2011), Scalzo et al. (2012, 2014), Sasdelli et al. (2015), and Nordin et al. (2018). A publication containing the spectral time series used in this analysis is in preparation. Additionally, the supernovae in this analysis discovered by iPTF (Kulkarni 2013) are also presented in Papadogiannakis et al. (2018). SNIFS is a fully integrated instrument optimized for automated observation of point sources on a structured background, i.e., the supernova host galaxy, over the full ground-based optical window at moderate spectral resolution ($R \sim 600\text{--}1300$). It consists of a high-throughput wide-band pure-lenslet integral field spectrograph (IFS), an imaging channel to monitor the field in the vicinity of the IFS for atmospheric transmission variations simultaneous with spectroscopy, and for acquisition and guiding. The IFS possesses a fully filled $6''.4 \times 6''.4$ spectroscopic field of view subdivided into a grid of 15×15 spatial elements, a dual-channel spectrograph covering $3200\text{--}5200 \text{ \AA}$ and $5100\text{--}10000 \text{ \AA}$ simultaneously, and an internal calibration unit (continuum and arc lamps). SNIFS is continuously mounted on the south bent Cassegrain port of the University of Hawaii 2.2 m telescope on Maunakea and is operated remotely. A description of host-galaxy subtraction is given in Bongard et al. (2011).

The supernovae are flux-calibrated following the procedure described in Buton et al. (2013), using a system of both CALSPEC (Bohlin 2014 and references within) and Hamuy (Hamuy et al. 1992, 1994) standard stars, which are calibrated on the Hayes & Latham (1975) system and listed in Buton et al. (2013). Calibration of the data on non-photometric nights is described in Pereira et al. (2013). Note that this hybrid calibration system is mildly in tension with data calibrated purely with the CALSPEC standards (an effect seen earlier in S15). This calibration difference primarily affects the U band; the effect on the $B - V$ color is minimal but the $U - B$ color can be approximately 0.02 mag lower in the system used here. The redshifts of the supernovae are given in Childress et al. (2013) and Rigault et al. (Rigault et al. 2018). The supernovae have been corrected for Milky Way dust extinction (Schlegel et al. 1998; Cardelli et al. 1989). The wavelengths have been shifted to a common rest frame at $z = 0$. All references hereafter to the phase or wavelength of the supernova data will refer to the values in this rest frame.

The supernovae used in this analysis range in redshift from $z = 0.01$ to 0.08 . They are required to have observations on at least 5 epochs, with at least one epoch between 10 days before and 7 days after maximum light, as determined by fitting the data with SALT2, and 4 epochs between 10 days before and 35 days after maximum light. Stricter phase-based cuts on the data set are also discussed below. In this analysis, we include peculiar supernovae, such as SN 1991T and SN 1991bg-like objects, plus other supernovae that are not well fit by the SALT2 model. Whether and how to exclude such supernovae from a cosmology analysis are discussed in Section 3.4.2 and in the Appendix. The SNfactory data set used here has 223 supernovae that meet these data cuts. Future versions of the

SNfactory data production are expected to increase this number.

The date of maximum fit by SALT2 is used to align the supernovae by phase in the analysis that follows. This phase is well constrained because of the other constraints put on the data set, and its uncertainty is generally well under half a day even for supernovae that are not very well matched by the SALT2 model. This small level of uncertainty is confirmed in the final models, which give a small scatter ($\text{RMS} = 0.3$ days) in the date of maximum of the reconstructed supernovae. Additionally, in the final data set used in this analysis only supernovae with data before maximum are used, further constraining the date of maximum.

3. Method for Building the Supernova Models

The goal of the analysis is to build a model for the time-evolving spectral energy distribution of SNe Ia that will consist of spectral time-series components. There are a number of reasons for using a model for SNe Ia, and we present separate analysis products to address these: a simple model with two components (the first of which is essentially the average spectral time series) plus a color relation to compare our analysis method with SALT2; a model optimized to work best for standardizing supernova absolute magnitudes; and a model intended to best match the spectral time-series diversity of all SNe Ia.

Following the precedent of current empirical supernovae models, each of the produced models will consist of a number of spectral time-series components. For each model, the components are linearly combined to match the spectral time series of a given supernova. The spectral-temporal model components are produced by first regressing the supernova data onto a regular grid in phase and wavelength, discussed in Section 3.1, and then calculating the principal components of this regularly sampled data, discussed in Section 3.3.

In training these empirical models, a number of choices must be made. First, we consider the selection criteria for the model training supernova set, namely whether supernova used in training the model must have epochs observed at early or late times. Also considered is whether the supernova spectra should be dereddened to account for dust between the supernova and the observer as a separate step and what dust reddening relation to use, discussed in Section 3.2. Lastly, the number of model components must be determined for the model optimized for magnitude standardization and the comprehensive spectral time-series model. To avoid overfitting, these decisions are made using K -fold cross-validation, described in Section 3.4.

3.1. Regression of Supernova Observations onto a Regularly Sampled Grid in Wavelength and Phase

The supernova data from the SNfactory are by necessity not uniformly sampled in phase and have varying degrees of noise due to the redshift range, phase, and observing conditions of the supernovae. To facilitate the calculation of the principal components, Gaussian processes are used to regress the irregular data onto a uniform grid in wavelength and phase. Gaussian processes (Rasmussen & Williams 2006) assume that an observation is a sample drawn from some underlying function with an error model:

$$y = f(x) + \epsilon, \quad (1)$$

where ϵ is a normal distribution that accounts for the error in the data. A Gaussian process models the underlying function as

$$f \sim \mathcal{N}(a m(X), K(X, X)), \quad (2)$$

where X are the input dependent variables (here two-dimensional elements corresponding to the wavelength and phase for each flux measurement, $m(X_s)$ is the model for the mean function, a is a fit normalization parameter, and K is a covariance matrix made up of Matérn kernel (Matérn 1986) elements

$$k(x, x') = \sigma_f^2 \left[1 + \frac{\sqrt{5}(x-x')}{l} + \frac{5(x-x')^2}{3l^2} \right] \times \exp\left(-\frac{\sqrt{5}(x-x')}{l}\right) + \sigma_n^2 \delta(x, x'), \quad (3)$$

where l , σ_f , and σ_n describe the correlation scale between variables, the amount of variation allowed from the mean, and the independent variation in the data).

In this analysis, the mean function $m(X)$ is constructed by taking the smoothed and normalized mean of all the supernovae, with the Hsiao et al. (2007) template as a prior at epochs where there is little data. The correlation scale l is fit globally for all supernovae and is only applied to data points with different phases but the same wavelength. No correlation is enforced between elements at different wavelengths since the data has already been binned in wavelength and enforcing a correlation was not found to improve results. The other parameters (a , σ_f , and σ_n) are fit for each supernova by maximizing the log likelihood of the model. These parameters are fit for each individual supernova because their values relate to the observing conditions and the amount the supernova differs from the average.

Given input data at X with variance V , the prediction for the underlying supernova spectrum at point X' is then the mean of the conditional distribution of function values drawn from the normal distribution:

$$f' = a m(X') + K(X', X)[K(X, X) + V]^{-1}(y - a m(X)) \quad (4)$$

with covariance

$$\text{cov}(f') = \Sigma' = K(X', X') - K(X', X) [K(X, X) + V]^{-1} K(X, X'). \quad (5)$$

For each supernova in our data set, a prediction is made in this manner for the true spectral time series in flux of the object. For this analysis, we generate spectral time series spanning 10 days before maximum and 46 days after maximum at two day intervals, with 288 wavelength elements between 3300 and 8600 Å at a constant velocity resolution of 1000 km s⁻¹. These phase and wavelength ranges are determined by the range in which the SNfactory supernovae are well observed, which correspond to the phase and wavelength range in which the majority of supernova luminosity is emitted (Howell et al. 2009). Fluxes here and after are f_λ (erg s⁻¹ cm⁻² Å⁻¹). Figure 1 shows an example of the Gaussian process prediction for a moderately well-sampled supernova with some scatter in the data. Figure 2 shows the performance of the predictions over the whole supernova set, demonstrating that the predictions are unbiased and have residuals consistent with the estimated

uncertainty on the data around maximum light. The uncertainty at later times is somewhat overestimated, which can be attributed to the fact that the matrix K is fixed for all epochs of the supernova and over-estimates the variability at later times. A more advanced epoch-dependent model might be able to account for this, but this is left for future work.

3.2. Dereddening the Spectra to Account for Interstitial Dust

Supernova color differences have been shown to be a major contributor to inhomogeneity in supernova spectra and magnitudes (Riess et al. 1996; Tripp 1998). This color is some combination of intrinsic color and reddening due to dust, but a generally accepted method for separating these components has not yet been discovered (Mandel et al. 2014, Kim et al., submitted). The portion due to dust reddening is expected to obey an extinction relation, such as in Cardelli et al. (1989, hereafter C89), or another similar model such as Fitzpatrick & Massa (2007, hereafter FM07). Dust reddening is well described by an exponential multiplier to the flux and thus cannot be described by a single linear principal component. Dereddening the supernovae using the C89 relation before calculating the spectral time-series model components may lead to a model that better matches supernova behavior than a model without dereddening and means that a smaller basis of linear components should be needed to describe supernova variability. Thus, in addition to calculating a purely linear supernova model, we also calculate models using supernova data that are first dereddened with the C89 or the FM07 dust reddening models. Supernova dereddening was seen to be dependent on the phases of the spectra used, for example when only early phases, which are seen to be highly variable in time, or only more uniform later phases, are available to calculate the dereddening. Because of this, the dereddening process is done using the Gaussian process prediction for the supernova, instead of the raw spectra, so that all phases are included.

The color relation is applied by fitting the color difference between each supernova and a fiducial supernova and dereddening the supernova by that amount. This is done by minimizing the quantity

$$\sum_{i=1}^{N_{\text{SNe}}} (\mathbf{f}_{\text{SN}_i}(p, \lambda) - a_i 10^{-0.4 A_{s,i}} C(\lambda) \mathbf{f}_{\text{fid}}(p, \lambda))^T \times \sum_{\text{SN}_i}^{-1} (\mathbf{f}_{\text{SN}_i}(p, \lambda) - a_i 10^{-0.4 A_{s,i}} C(\lambda) \mathbf{f}_{\text{fid}}(p, \lambda)) \quad (6)$$

as a function of a_i , $A_{s,i}$, and $\mathbf{f}_{\text{fid}}(p, \lambda)$, where $A_{s,i}$ is the extinction difference between SN_i and the fiducial supernova, a is a normalization factor, $\mathbf{f}_{\text{fid}}(p, \lambda)$ is a freely fit spectral time series varying as a function of the phase p and wavelength λ , and \mathbf{f}_{SN_i} and Σ_{SN_i} are the predicted flux and covariance for each supernova from the Gaussian processes step (Equations (4) and (5)). $C(\lambda)$ is either the C89 or FM07 dereddening model, both using $R_V = 3.1$. (The degeneracy between a_i , $A_{s,i}$, and $\mathbf{f}_{\text{fid}}(p, \lambda)$ is removed by fixing the value for a_i for an arbitrarily chosen supernova and fixing the average of $A_{s,i}$ to zero.) A fit fiducial supernova is used, rather than a template such as Hsiao et al. (2007), in order to ensure that the maximum amount of variation due to dust-like reddening allowed by the data is removed.

The supernovae dereddened using the fit A_s values (but not including the normalization factor a) are then used as input to the calculation of the components that will form the final

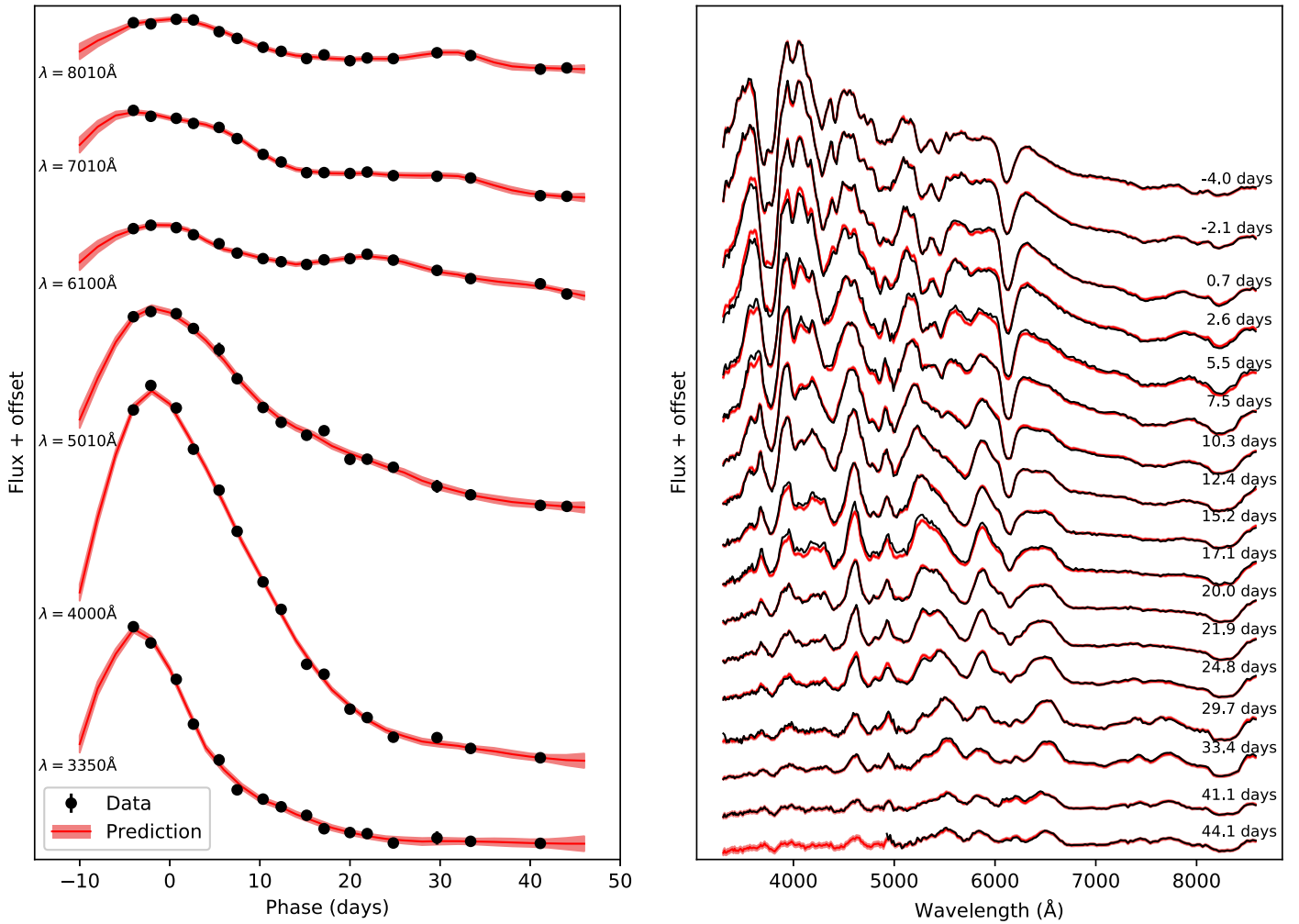


Figure 1. A demonstration of the supernova model predicted by Gaussian processes compared with the supernova data. The left panel shows monochromatic light curves while the right panel shows the spectra predicted by the Gaussian processes at the phases where there is supernova data.

spectral-temporal model. For brevity, $F_{\text{SN}}(p, \lambda)$ will refer hereafter to the dereddened spectral time series of the supernova, i.e.,

$$F_{\text{SN}}(p, \lambda) = f_{\text{SN}}(p, \lambda) 10^{0.4A_r C(\lambda)}. \quad (7)$$

3.3. Extracting Principal Components from the Spectral Time Series

In order to calculate linear components from the supernova spectral time-series predictions produced using Gaussian processes, we use a variation of PCA that is capable of incorporating data with uncertainties. This variation, Expectation Maximization Factor Analysis (EMFA) is a method for performing dimensionality reduction presented in Ghahramani & Hinton (1996). In addition to utilizing data uncertainty, EMFA has the added advantage of allowing the user to choose the number of model components that will be used to describe the data set, in other words, the basis of N vectors that best fits a given space. Note that this means that a model with N components is not a subset of a model with $N + 1$ components trained on the same data set.

EMFA models the given data as

$$F = e c + u, \quad (8)$$

where F is the $N_{\text{variables}} \times N_{\text{SNe}}$ dimensional data matrix (here the Gaussian process predictions for each phase and wavelength combination) and e is a $N_{\text{variables}} \times N_{\text{vectors}}$ dimensional matrix of model components. The factors c (forming a $N_{\text{vectors}} \times N_{\text{SNe}}$ dimensional matrix) correspond to the coefficients of the model. N_{vectors} is generally much smaller than $N_{\text{variables}}$. Here we have 8352 variables corresponding to 288 wavelength elements times 29 phase elements, while we predict something on the order of 10 vectors. Finally, u is a $N_{\text{variables}} \times N_{\text{SNe}}$ dimensional random variable with distribution $\mathcal{N}(0, \Sigma)$, corresponding to the noise Σ in F , and is here taken from the covariance predicted in the Gaussian processes step (see Section 3.1). While the method is geared toward data centered around zero, here we do not subtract the mean so as to prevent introducing assumptions about the absolute flux of the supernovae (in terms of calibration and cosmology) into the model. Testing has shown that the method is stable without subtracting the mean and is also insensitive to moderate changes in the normalization of the supernovae, though it starts to degrade if order-of-magnitude variations are introduced in the brightness scale of the data. The method of Ghahramani & Hinton (1996) has been adapted here to accommodate correlated uncertainties.

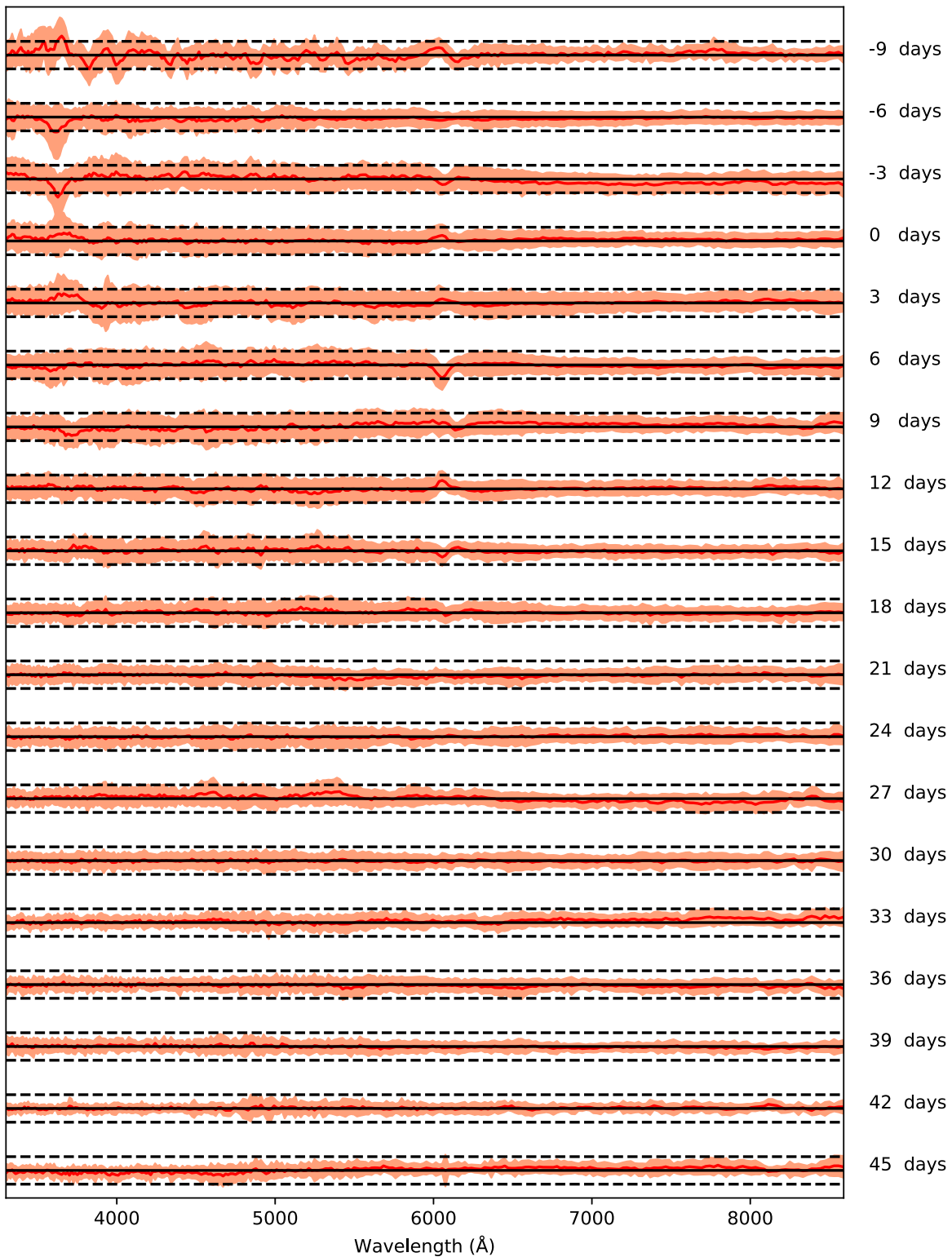


Figure 2. Average and standard deviation of the pulls between Gaussian process predictions and the data vs. the wavelength, in 3-day phase bins. The red lines show the average pull, which is defined as the residual divided by the error, of the Gaussian process prediction to the data at the phase indicated to the right of the figure, while the shaded region shows the standard deviation in the pulls. The solid black lines are the zeropoint for the averages, where a level of zero indicates that there is no bias. The dashed lines show the level of plus or minus one in the pulls.

The method consists of an iterative process in which the expected value of the factor c is calculated, given e , and then new values for e are calculated given the data and the new values for c . This is repeated until the log likelihood converges, here

determined by the point where the percentage change in the log likelihood is less than 10^{-7} . Once convergence is reached, the components in the matrix e can be orthogonalized and ranked using a standard Singular Value Decomposition algorithm.

Table 1
Numbers of Supernovae in Each of the Levels of Data Selection

Data Selection	Each Training Set	Each Test Set	Final Hold-out Set	Total
All Supernovae	120	39	64	223
Supernovae with pre-max data	92	30	49	171
Supernovae with pre-max and post-30 days data	77	25	38	140

The resulting model for a supernova takes the form

$$\mathbf{F}_{\text{SN}}(p, \lambda) = \sum_{i=0}^{N-1} c_{\text{SN},i} e_i(p, \lambda), \quad (9)$$

where \mathbf{F}_{SN} is the flux of the supernova at the phase p and wavelength λ , N is the number of model components, e_i are the model components (the eigenvectors), and $c_{\text{SN},i}$ are the model coefficients fit to the individual supernova. Since the model is expected to be used for supernovae at a range of redshifts, we rewrite the model in the form

$$\mathbf{F}_{\text{SN}}(p, \lambda) = c_{\text{SN},0} \left(e_0(p, \lambda) + \sum_{i=1}^N c'_{\text{SN},i} e_i(p, \lambda) \right), \quad (10)$$

where $c'_{\text{SN},i} = \frac{c_{\text{SN},i}}{c_{\text{SN},0}}$, so that $c_{\text{SN},0}$ is the only model coefficient that is dependent on the flux normalization and is thus the only coefficient that is a function of the luminosity distance. Because of the leading multiplicative coefficient, the model is not affected by the flux normalization of the training data. Hereafter, the prime is dropped and $c'_{\text{SN},i}$ is referred to as simply $c_{\text{SN},i}$. Note that this step creates correlations between the model coefficients, even though the eigenvectors are orthogonal.

3.4. Making Decisions on the Model Training Using Cross-Validation

As described above, there are a number of decisions to make regarding the training of the models. The first question is whether to make any cuts on the supernova set (in addition to those described in Section 2) based on the phase coverage of the observations. For this, we compare the model components calculated using the full data set to a set only including supernovae that have observations before maximum, which brings the total number of supernovae to 171, and a set only including supernovae that have observations before maximum and at least 30 days after maximum, comprising 140 supernovae. The number of supernovae in each level of cuts is given in Table 1. Second, as described in Section 3.2, we consider whether to deredden the supernovae prior to calculating the model components and whether to use the reddening relation of FM07 or C89. Last, we must determine how many components to use in the model optimized for standardizing supernova magnitudes and the model optimized to fully capture the spectral time series for all SNe Ia.

Cross-validation is used to determine the best option for each of these issues. After withholding $\sim 30\%$ of the total number of supernovae randomly selected to be used as a final demonstrative set, the remaining supernovae are divided into 4 training and test set pairs, following a K -fold cross-validation procedure where the model is trained on $3/4$ of the data and tested on the last $1/4$. The options that best match the supernovae are chosen by looking at the average of the given

metrics (described below) over the 4 sets. Once the options have been chosen in this step, the final models are trained accordingly on all of the supernovae (with the exception of the supernovae withheld as a final demonstration set).

3.4.1. Selecting the Phase Coverage Criterion and Dereddening Procedure

To select the training set phase coverage criterion and method of accounting for reddening, we first consider how well the models are able to fit the test set supernovae. To do this, we calculate the value

$$\chi^2(n) = \sum_{i=0}^{\text{SNe}} (\mathbf{F}_i^n(p, \lambda) - \mathbf{F}_{\text{SN}_i}(p, \lambda))^T \times \Sigma_{\text{SN}_i}^{-1} (\mathbf{F}_i^n(p, \lambda) - \mathbf{F}_{\text{SN}_i}(p, \lambda)), \quad (11)$$

where \mathbf{F}_i^n is the n -component model fit to supernova i , and \mathbf{F}_{SN_i} and Σ_{SN_i} are the predicted flux and covariance from the Gaussian processes step. The average of the reduced χ^2 , in other words the χ^2 per degree of freedom, over each of the 4 test sets is compared for each of the combinations of training sets and dereddening options. Reduced χ^2 values below 1 are driven by the overestimated uncertainties at later epochs from the Gaussian processes step, discussed in Section 3.1. Since we are concerned with the relative value of the χ^2 for different models using the same data, the absolute value of χ^2 is not a concern here. We find that the models using the FM07 reddening relation and trained with supernovae that have pre-maximum observations are equivalent or slightly preferred to the others, while also performing best on the metrics considered in the following sections. As a preference for the Fitzpatrick reddening relation was also seen in recent studies, such as Schlafly et al. (2010) and Huang et al. (2017), we select this dereddening procedure for the models.

The χ^2 metric, shown in Figure 3, does not provide a way to choose the optimal number of components to use for the model since the reduced χ^2 value follows roughly asymptotic behavior for the test sets and does not have a minimum indicating the number of components preferred by the data.

3.4.2. Selecting the Number of Components for the Magnitude Standardization Model

In terms of its utility for cosmology, the primary metric for a supernova model is how well it can standardize supernova magnitudes and minimize their unexplained dispersion, that is, the combination of effects leading to dispersion in the supernova magnitudes that are not captured by our model. To select the model that leads to the lowest unexplained dispersion, we perform a magnitude standardization procedure on each model following the method of the ‘‘Union’’ supernova compilation (Amanullah et al. 2010). The model chosen is the

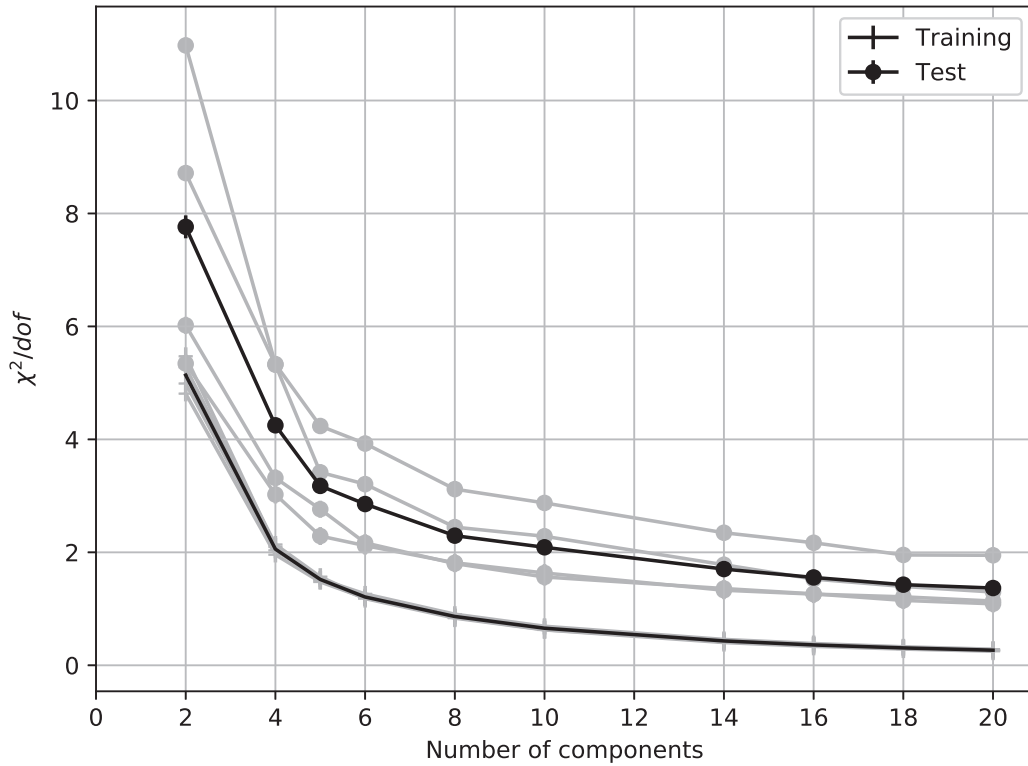


Figure 3. Reduced χ^2 values for supernovae with prior dereddening using the FM07 extinction curve vs. the number of model components. Gray lines show the results for each of the cross-validation training and test sets, while black lines show the median and error on the median for the training and test sets.

one for which the unexplained dispersion is lowest. This procedure performs a global fit for the level of unexplained dispersion simultaneously with the parameters α , α_c , and M_g used to calculate each supernova’s distance modulus:

$$\begin{aligned} \mu_g &= m_g^{\text{std}} - M_g = m_g + \sum_i \alpha_i c_i \\ &+ \alpha_c \times A_S - M_g, \end{aligned} \quad (12)$$

where c_i and A_S are the model parameters, determined by the previously obtained fits for the supernovae from Equation (10), m_g is the SDSS g -band magnitude at maximum obtained by performing synthetic photometry on the model formed by those fit parameters, M_g is the standardized absolute magnitude of the supernova, and μ is the distance modulus. In the case where no prior dereddening step is performed, the $\alpha_c \times A_S$ term is omitted. In this procedure we also reject supernovae that are outliers in the model coefficient space, as described in the Appendix, because in this section only a linear standardization is assumed. The supernovae identified in this way were already known to be peculiar (Scalzo et al. 2012), though this knowledge was not used to select them as outliers in this analysis.

The standardization parameters are fit using an iterative procedure in which the log likelihood function

$$\begin{aligned} \mathcal{L} &= \sum_i^{N_{\text{SNr}}} \left\{ -\ln[2\pi (\sigma_{pv,i}^2 + \sigma_{m,i}^2 + \sigma_u^2)] \right. \\ &\quad \left. - \frac{[\mu_g(\alpha, \alpha_c, M_g) - \mu(z_i, \theta)]^2}{\sigma_{pv,i}^2 + \sigma_{m,i}^2 + \sigma_u^2} \right\} \end{aligned} \quad (13)$$

is maximized with a fixed value of the unexplained dispersion σ_u , and then the value of the unexplained dispersion is fit such that the reduced χ^2 is equal to unity. The distance modulus $\mu(z, \theta) \approx 5 \log(1 + z) + \text{constant}$, where θ represents the parameters of the cosmological model, is insensitive to the cosmological model since the supernova redshifts are very low and the constant is degenerate with the supernova absolute magnitude M_g ; σ_{pv} accounts for uncertainty due to host-galaxy peculiar velocities of 300 km s^{-1} ; and σ_m is the uncertainty in the fit of the supernova model,

$$\sigma_m = \mathbf{V}^T \mathbf{C} \mathbf{V} \quad \text{where} \quad \mathbf{V} = \begin{pmatrix} 1 \\ \alpha_c \end{pmatrix}, \quad (14)$$

and \mathbf{C} is the covariance in the model fits to the supernova data.

It is important to note that this standardization method is purely empirical and not an analytic consequence of the model. It is based on finding empirical correlations between the $c_{i>0}$ components, which are independent of the luminosity distance (see Equation (10)), and the magnitudes of the supernovae in the data set used here, thus separating the intrinsic and distance-dependent parts of the supernova magnitude. More advanced techniques, potentially requiring more data, could be used to improve upon the linear standardization performed here (discussed further in Section 4.3).

The standardization parameters are calculated from the training set, then applied to the test set. To select the number of components to use for the model optimized to minimize dispersion in supernova magnitudes, we choose the point where the unexplained dispersion stops improving for the test set, shown in Figure 4. Fitting a curve to the modeled points gives a

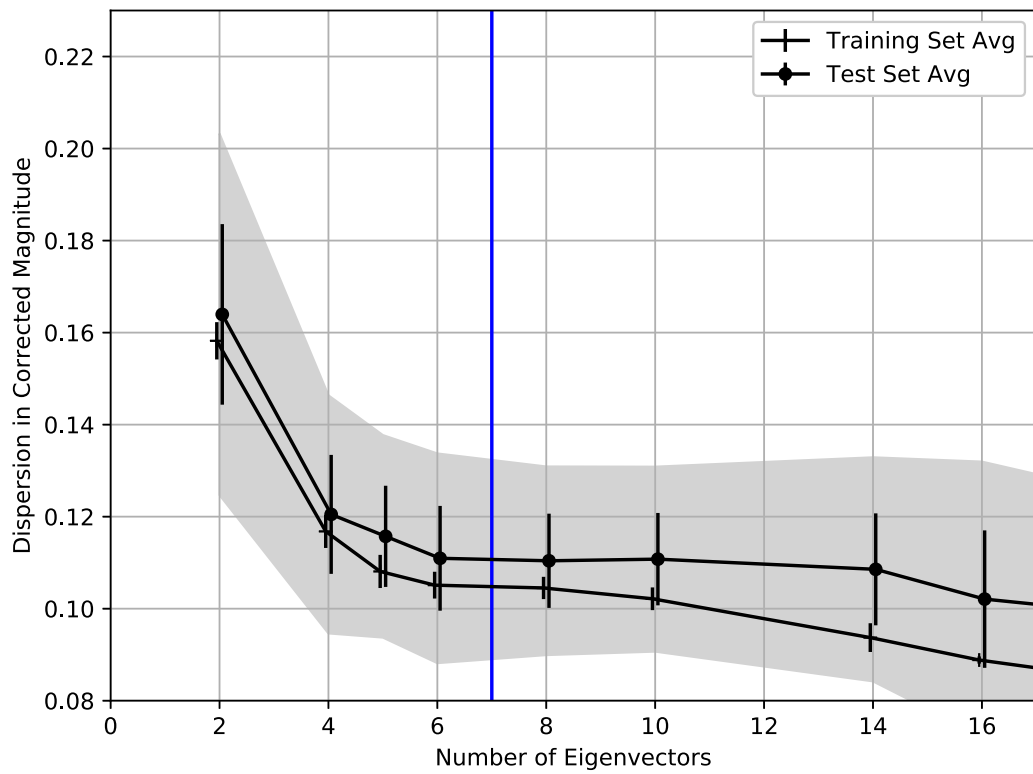


Figure 4. Unexplained dispersion in standardized magnitudes for supernovae with *FM07* dereddening vs. the number of model components. The solid lines show the result averaged over all of the K -fold sets and the error bars shows the error on the mean, while the shaded region shows the standard deviation among the test sets. Points are offset slightly on the x -axis to facilitate interpretation. The vertical blue line shows the local minimum at 7 vectors.

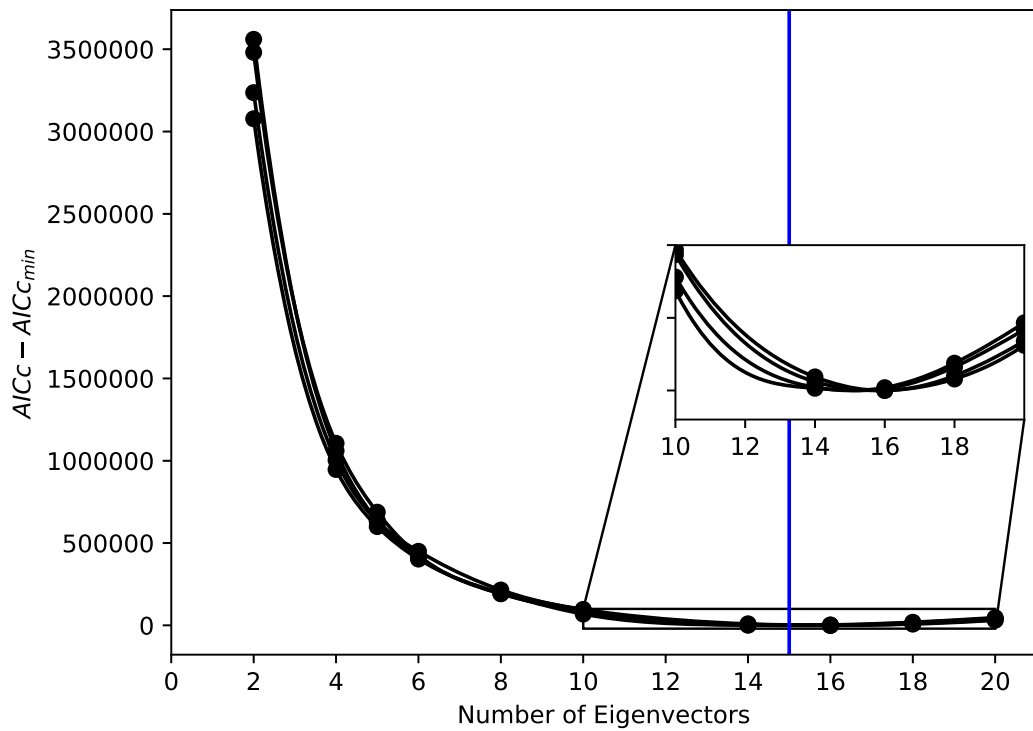


Figure 5. AIC metric applied to the models trained on the supernovae with data before maximum, corrected with the *FM07* reddening relation, vs. the number of model components. The four lines correspond to the different cross-validation training sets. The vertical blue line shows the minimum at 15 vectors.

Table 2
Components for the SNEMO Models

Phase (days)	Wavelength (Å)	SNEMO2 e_0	SNEMO2 e_1	SNEMO7 e_0	SNEMO7 e_1	SNEMO7 e_2
-10.00	3305	1.201	-0.2772	1.365	0.7203	0.03686
-10.00	3317	1.235	-0.2931	1.404	0.7471	-0.01845
-10.00	3328	1.271	-0.2961	1.448	0.7512	-0.04672
-10.00	3339	1.313	-0.2654	1.485	0.7356	-0.09023
-10.00	3350	1.306	-0.2396	1.482	0.7120	-0.05760
-10.00	3361	1.294	-0.2407	1.489	0.7122	-0.04199

Table 3
Training and Validation Set Model Parameters

SN Name	SNEMO2 $c_0 - \bar{c}_0$	SNEMO2 c_1	SNEMO7 $c_0 - \bar{c}_0$	SNEMO7 c_1	SNEMO7 c_2
Train_SN ₀	0.003742	-0.504857	0.003014	1.016900	-0.185422
Train_SN ₁	0.012838	0.135103	0.013745	0.572341	0.584960
Train_SN ₂	-0.024537	0.123520	-0.024508	0.652294	0.252721
Train_SN ₃	0.004242	0.563471	0.004089	-0.021505	0.326422

(This table is available in its entirety in machine-readable form.)

local minimum at 7 components. While the unexplained dispersion decreases again at 16 components, we choose the minimum at 7 components as this is more likely to be constrainable with photometric data, which is the target for this model.

3.4.3. Selecting the Number of Components for the Comprehensive Supernova Model

To select the model with the number of components that best describes the full set of supernovae we compare the Akaike Information Criterion (AIC, Burnham & Anderson 2003) for each number of model components. This method measures the relative quality of models for a given set of data and puts a stronger penalty on adding complexity to the model than the reduced χ^2 metric. Because of the fact that a very large number of variables (corresponding to the number of wavelength and phase elements of the spectral time series) are added for each additional model component, the models are not well suited to calculating the relative probability between different numbers of components, but a minimum AIC value can be found. The AIC metric versus the number of model components is shown in Figure 5, where the minimum at 15 components indicates that this model best describes the supernova population.

4. The Final Spectral-temporal Models

With all training decisions made as described in the previous section, the 2-, 7-, and 15-component models, called SNEMO2, SNEMO7, SNEMO15, are calculated with the full training set of 122 supernovae with observations before maximum (still reserving 49 for use in the hold-out final demonstration set). The resulting model components are given in Table 2 and the corresponding coefficients for the training and hold-out sets of supernovae are given in Table 3.

The components of the models are shown in Figures 6, 7, and 8 (Figure 9 shows the SALT2 model for comparison). For demonstration purposes, we only show spectra at a few intervals in phase and monochromatic light curves at illustrative wavelengths. Another version of the models is shown in the Appendix, in which no outlier supernovae (which here were all identified after the fact as known SN 1991T or SN

1991bg-like objects) are included in the analysis. In these no-outlier training versions, about 1 less component is needed for the SNEMO7 and SNEMO15-analog models.

The distributions of coefficients for SNEMO2 and SNEMO7 are shown in Figure 10. As expected, the A_g parameter correlates strongly with m_g . The distributions provide hints of a nonlinear structure in the supernova population, but do not give any immediate indications of subpopulations.

It would be expected that the more complex models contain all of the information in the less complex models. The following matrix equations demonstrate how the SNEMO2 and SNEMO7 models are reconstructed by the more complex models, with the R^2 values denoting coefficients of determination, i.e., the fraction of the variance explained in the model reconstruction of the other model (with 1 as a perfect reconstruction).

For SNEMO2,

$$e^{(2)} \approx e^{(7)} \cdot \begin{pmatrix} 0.996 & 0.001 \\ 0.515 & -0.612 \\ 0.153 & 0.450 \\ 0.080 & 0.067 \\ 0.118 & 0.282 \\ 0.254 & 0.089 \\ 0.195 & -0.588 \end{pmatrix} \approx e^{(15)} \cdot \begin{pmatrix} 0.990 & 0.001 \\ 0.546 & -0.492 \\ 0.119 & 0.479 \\ 0.103 & 0.210 \\ -0.173 & -0.253 \\ 0.119 & -0.348 \\ -0.122 & 0.591 \\ -0.038 & 0.168 \\ -0.354 & 0.456 \\ 0.204 & -0.285 \\ 0.129 & -0.074 \\ -0.087 & 0.053 \\ 0.160 & 0.047 \\ -0.322 & 0.157 \\ -0.178 & -0.240 \end{pmatrix} \quad (15)$$

with corresponding R^2 values

$$\begin{aligned} R^2_{(\text{SNEMO7})} &= (1.000, 0.857) \\ R^2_{(\text{SNEMO15})} &= (1.000, 0.871). \end{aligned} \quad (16)$$

While for SNEMO7,

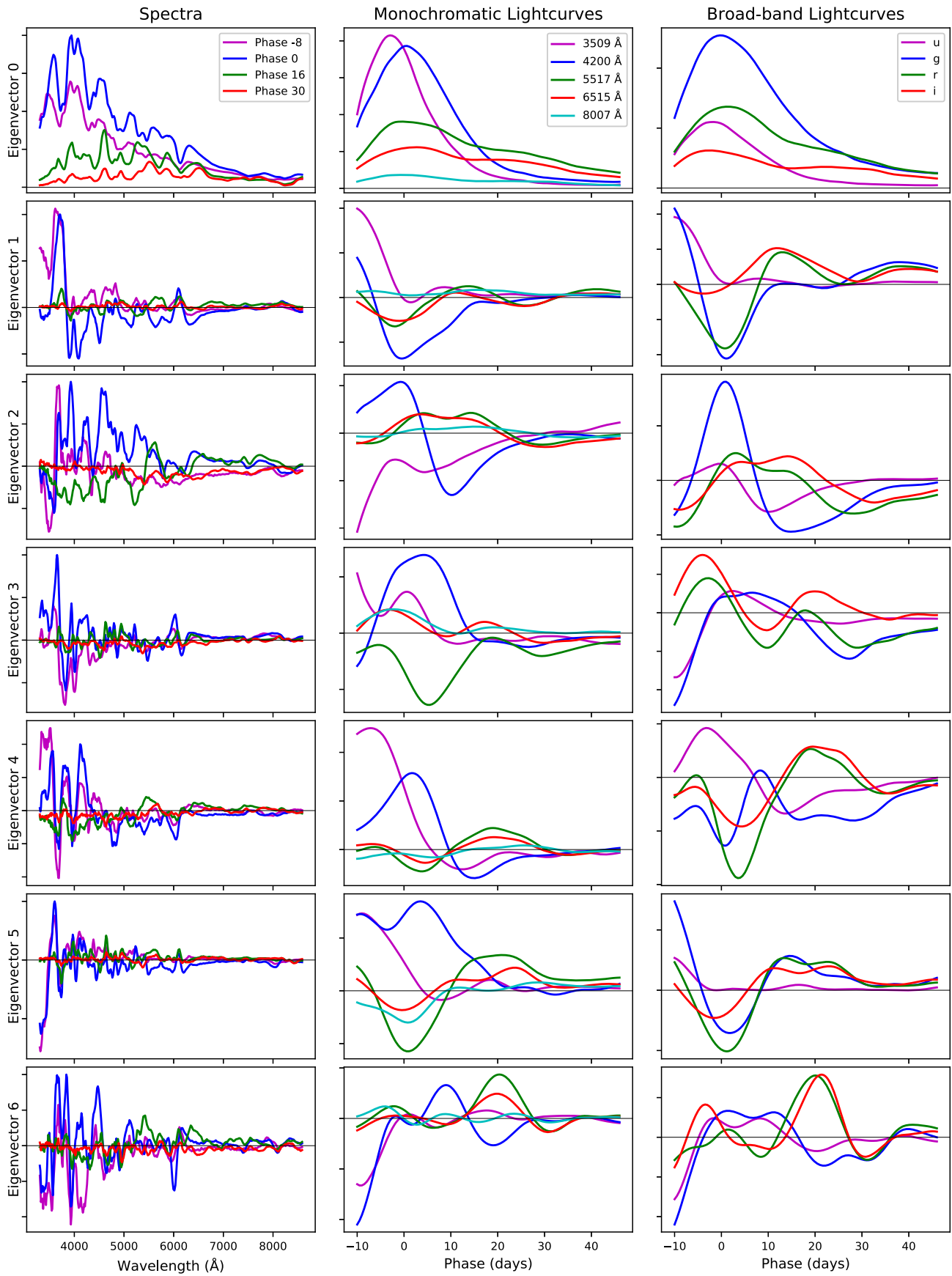


Figure 6. Demonstrative spectra and light curves produced using SNEMO7. See Figure 8 for descriptions of each panel.

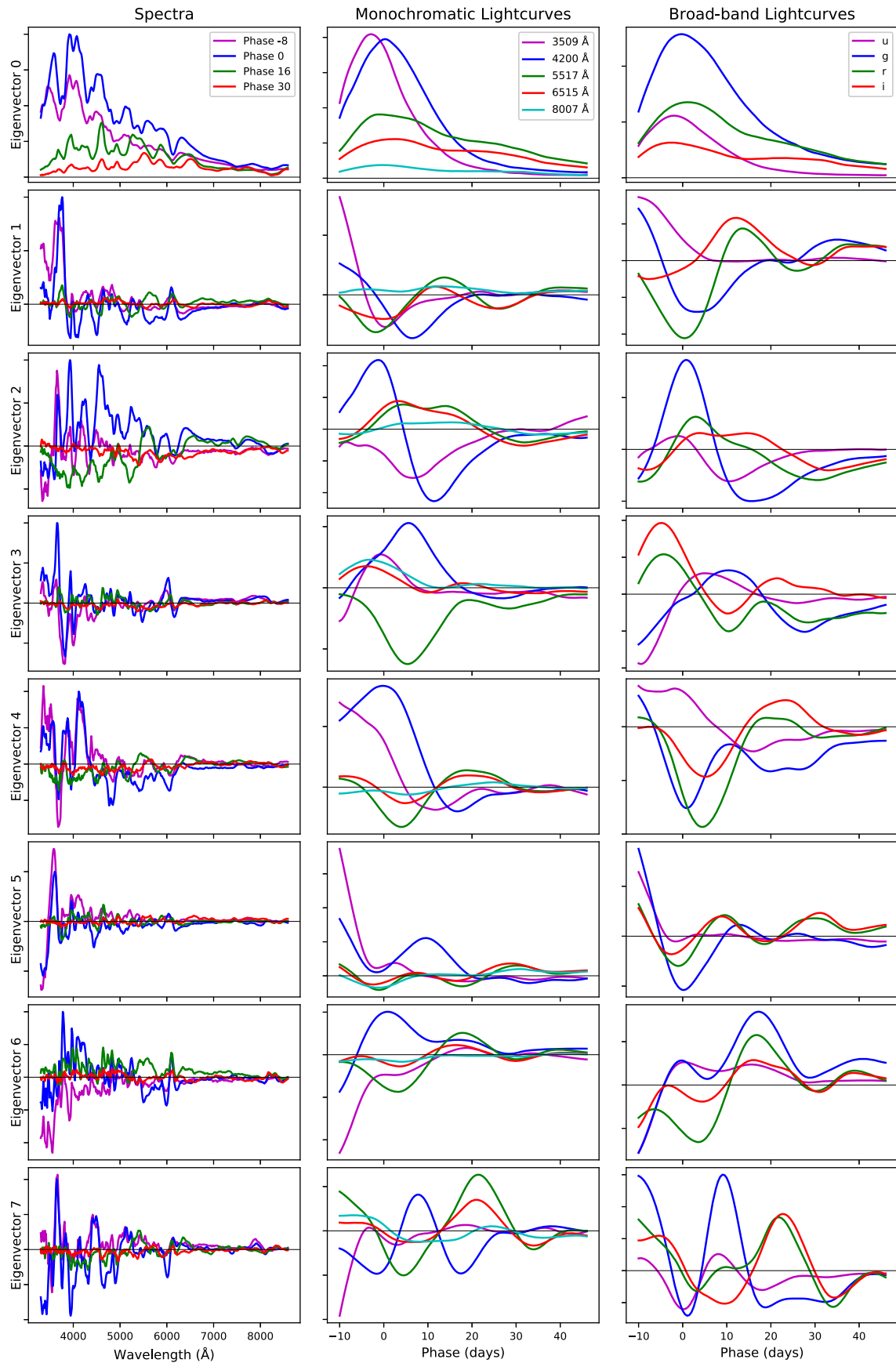


Figure 7. Demonstrative spectra and light curves produced using the first eight components of SNEMO15. See Figure 8 for descriptions of each panel.

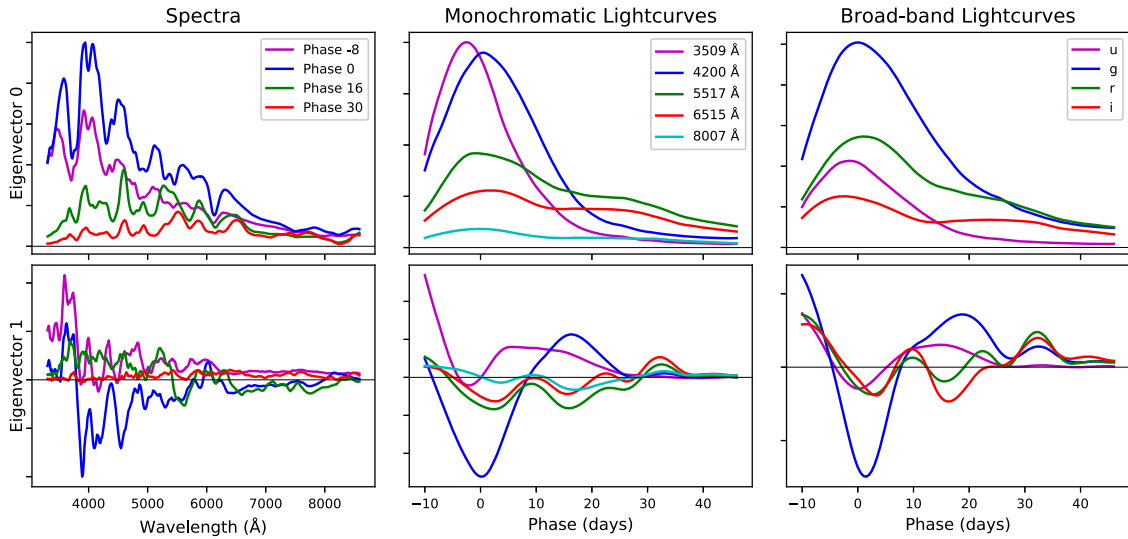


Figure 8. Demonstrative spectra and light curves produced using SNEMO2. The left column shows slices in the phase dimension of the spectral time series (i.e., spectra) at four demonstrative phases, while the middle column shows slices in the wavelength dimension (monochromatic light curves). The right column shows broad-band light curves in the SDSS u , g , r , and i filters.

$$e^{(7)} \approx e^{(15)} \times \begin{pmatrix} 0.994 & 0.000 & -0.000 & -0.000 & -0.000 & 0.000 & -0.000 \\ 0.039 & 0.952 & 0.158 & 0.008 & 0.060 & -0.055 & 0.012 \\ 0.003 & -0.077 & 0.951 & -0.064 & 0.021 & 0.037 & -0.009 \\ 0.008 & -0.038 & 0.108 & 0.999 & 0.132 & 0.055 & -0.025 \\ -0.093 & 0.108 & 0.023 & 0.088 & -0.844 & -0.082 & -0.045 \\ -0.203 & 0.255 & -0.269 & -0.092 & -0.211 & 0.858 & 0.124 \\ -0.018 & -0.257 & 0.317 & -0.090 & -0.047 & 0.496 & -0.262 \\ 0.189 & -0.432 & 0.190 & -0.051 & -0.088 & -0.232 & 0.678 \\ -0.157 & 0.088 & -0.159 & -0.308 & 0.401 & 0.008 & -0.422 \\ 0.019 & 0.249 & -0.155 & -0.060 & 0.608 & 0.204 & 0.295 \\ -0.059 & 0.119 & 0.135 & 0.077 & -0.158 & 0.129 & 0.128 \\ -0.064 & -0.089 & -0.227 & -0.245 & 0.226 & -0.181 & -0.226 \\ 0.143 & 0.145 & -0.235 & 0.084 & 0.163 & 0.238 & -0.081 \\ -0.386 & 0.251 & 0.183 & -0.100 & 0.541 & -0.261 & -0.167 \\ -0.461 & 0.045 & -0.278 & 0.172 & 0.410 & -0.214 & -0.018 \end{pmatrix} \quad (17)$$

with corresponding R^2

$$R^2 = (1.000, 0.988, 0.968, 0.984, 0.916, 0.986, 0.900). \quad (18)$$

The somewhat lower R^2 values for the second component of SNEMO2 and for the seventh component of SNEMO7 indicate small differences between the subspaces of the models.

4.1. Overview of the Three Models

Figure 11 shows an example of fitting a supernova with each of the three models. Here and after, the SNEMO models have been fit to the Gaussian processes' models of the hold-out supernovae, but the models can be equally applied to raw spectra or photometry by interpolating or integrating over the spectral time-series templates. It can be seen from the residuals between the supernova spectra and the models that major improvements have been made moving from the two-component model, SNEMO2, to the seven-component model, SNEMO7. In the example shown, SNEMO2 misses both broad-band behavior and spectral features of the supernova. Most of these are matched by SNEMO7. SNEMO15 makes some

additional improvements, particularly with adjustments to detailed shapes of the spectral feature Ca II H&K and the early spectra.

A comprehensive view is found looking at the standard deviation of the residuals for all of the supernovae. This is shown at intervals in phase in Figure 12. Areas with large standard deviations in the residuals indicate spectral diversity that is not captured by the model. For SNEMO2, there are large residuals at the Si II 6150 feature, with smaller residuals at the Si II features near 4100 Å. Large residuals are also apparent in the near-UV and Ca H&K region, particularly around maximum light, and in the Fe II blend around 5000 Å at and after maximum light. Aside from the spectral features in the residuals, broad-band residuals are also apparent at and after maximum light.

For SNEMO7, the spectral and broad-band residuals are greatly reduced. Some unmodeled supernova diversity is still apparent in the S II features, at Si II 6150 and in the near-UV near maximum light. Moving further to SNEMO15, almost all of these are gone and the residuals are at or below the scale of the error in the data.

Table 4 shows the effect of adding or subtracting the one-sigma contribution (as determined by the distribution of model coefficients) of a given model component from the mean spectrum,

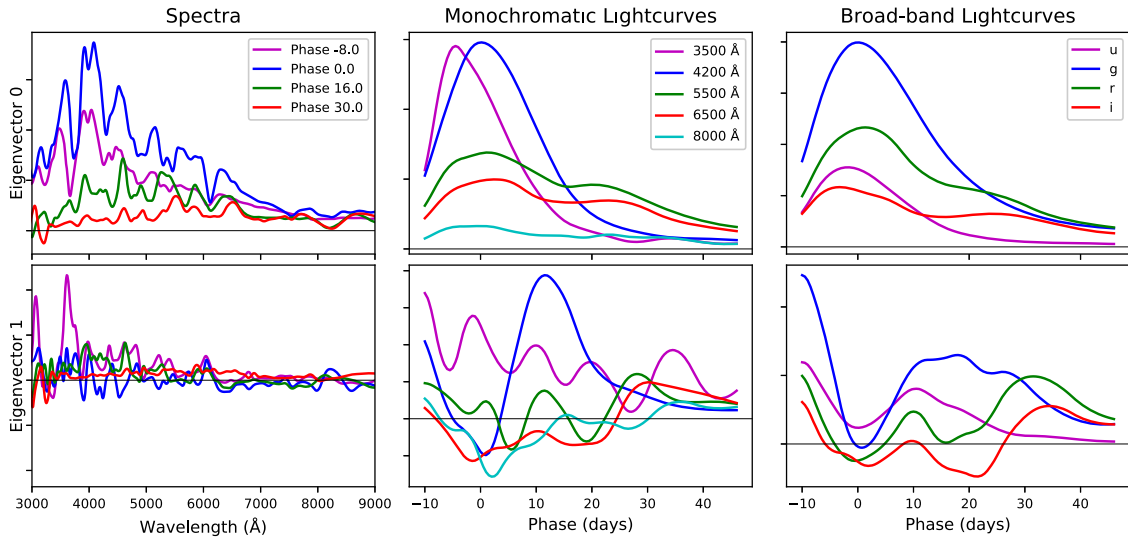


Figure 9. Demonstrative spectra and light curves produced using the SALT2 model. See Figure 8 for descriptions of each panel.

i.e., the difference between certain characteristics of e_0 versus $e_0 \pm \text{std}(c_i) e_i$. We consider the $g - r$ color at maximum and the change in the g -band magnitude between 0 and 15 days after maximum light, $\Delta m15(g)$ (thus the term here refers to the quantity itself, not a model fit parameter as, for example, in SNooPy, Burns et al. 2011). These quantities are calculated by evaluating the broad-band synthetic photometry of the model at 0 and 15 days. In terms of broad-band behavior, SNEMO2 is shown to be analogous to the SALT2 model, with Eigenvector 1 causing effects of a similar magnitude in $\Delta m15(g)$, though the effect on the $g - r$ color at maximum is larger for SNEMO2. The components can also have an effect on the date of maximum. Here we show the epoch of the g -band maximum, based on an interpolation of the components. The g -band maximum is slightly earlier than the phase zeropoint given by SALT2 for all of the models (as well as for SALT2). For the reconstructed supernova models, there is a small scatter of 0.3 days in the date of maximum, which is in agreement with the uncertainty in the SALT2 phases used as input.

In accordance with FM07, a constant value of $R_V = 3.1$ has been assumed in this analysis. However, it can be seen in Table 4 that the components affect the $g - r$ color at maximum and, if the training supernovae have been affected by variable R_V , the model components might have some capacity for fitting other values of R_V . This possibility was tested by fitting the models to the supernova spectral time series that were simulated with higher and lower values of R_V , with a range of values for A_V . It was found that the SNEMO7 and SNEMO15 can account for a portion of the variation due to a change in R_V (a roughly 25% improvement in the χ^2 values of fits to supernova spectra simulated using values of R_V higher or lower than 3.1), but do not have the power to fully account for variable R_V .

4.2. Comparison of SNEMO2 with SALT2

Comparing the two-component plus color model presented here with the SALT2 model, Figures 8 and 9, can give some context for understanding the new model components. The SALT2 auxiliary template, corresponding to the x_1 parameter, is known to be highly correlated with the width of the B-band supernova light curve, so comparing the new model

components with this template gives us a proxy for how correlated our model components are with light-curve width.

The supernova model parameters are compared in Figure 13. While the second components of each (Eigenvector 1) do not appear to be very similar, it is clear that the coefficients are highly correlated (bottom right panel of the figure). Thus, while they may be capturing the behavior differently, the two models find roughly the same way to discriminate among the supernovae. Some difference may be attributable to the fact that the SALT2 x_1 and c components are constructed to be uncorrelated, while this is not the case with the models shown here. A comparison of SNEMO7 and SNEMO15 (not shown) shows that each also has one or two components that have strong correlations with the SALT2 x_1 coefficient, though not so clearly as in SNEMO2.

As in Equations (15) and (17), the SALT2 model can be reconstructed with the SNEMO models. In this case, it is also necessary to fit the color difference between the models. The best-fit reconstructions are given by

$$\begin{aligned} \text{SALT2} &\approx \begin{pmatrix} e^{(2)} \\ \text{FM07} \end{pmatrix} \cdot \begin{pmatrix} -0.440 & -0.012 \\ 0.082 & -0.245 \\ 0.817 & 0.388 \end{pmatrix} \\ &\approx \begin{pmatrix} e^{(7)} \\ \text{FM07} \end{pmatrix} \begin{pmatrix} -0.440 & -0.013 \\ -0.197 & 0.106 \\ -0.070 & -0.143 \\ -0.049 & 0.038 \\ 0.020 & -0.230 \\ -0.014 & 0.218 \\ -0.238 & 0.166 \\ 0.821 & 0.413 \end{pmatrix} \approx \begin{pmatrix} e^{(15)} \\ \text{FM07} \end{pmatrix} \begin{pmatrix} -0.441 & -0.014 \\ -0.216 & 0.097 \\ -0.048 & -0.178 \\ -0.040 & -0.004 \\ 0.011 & 0.208 \\ -0.057 & 0.374 \\ 0.199 & 0.019 \\ -0.046 & -0.026 \\ 0.339 & -0.216 \\ -0.150 & -0.078 \\ 0.193 & 0.117 \\ 0.140 & -0.222 \\ -0.035 & -0.049 \\ 0.311 & -0.058 \\ -0.063 & 0.154 \\ 0.828 & 0.515 \end{pmatrix}, \end{aligned} \quad (19)$$

where the first column refers to the SALT2 mean template and the second column refers to the SALT2 auxiliary template

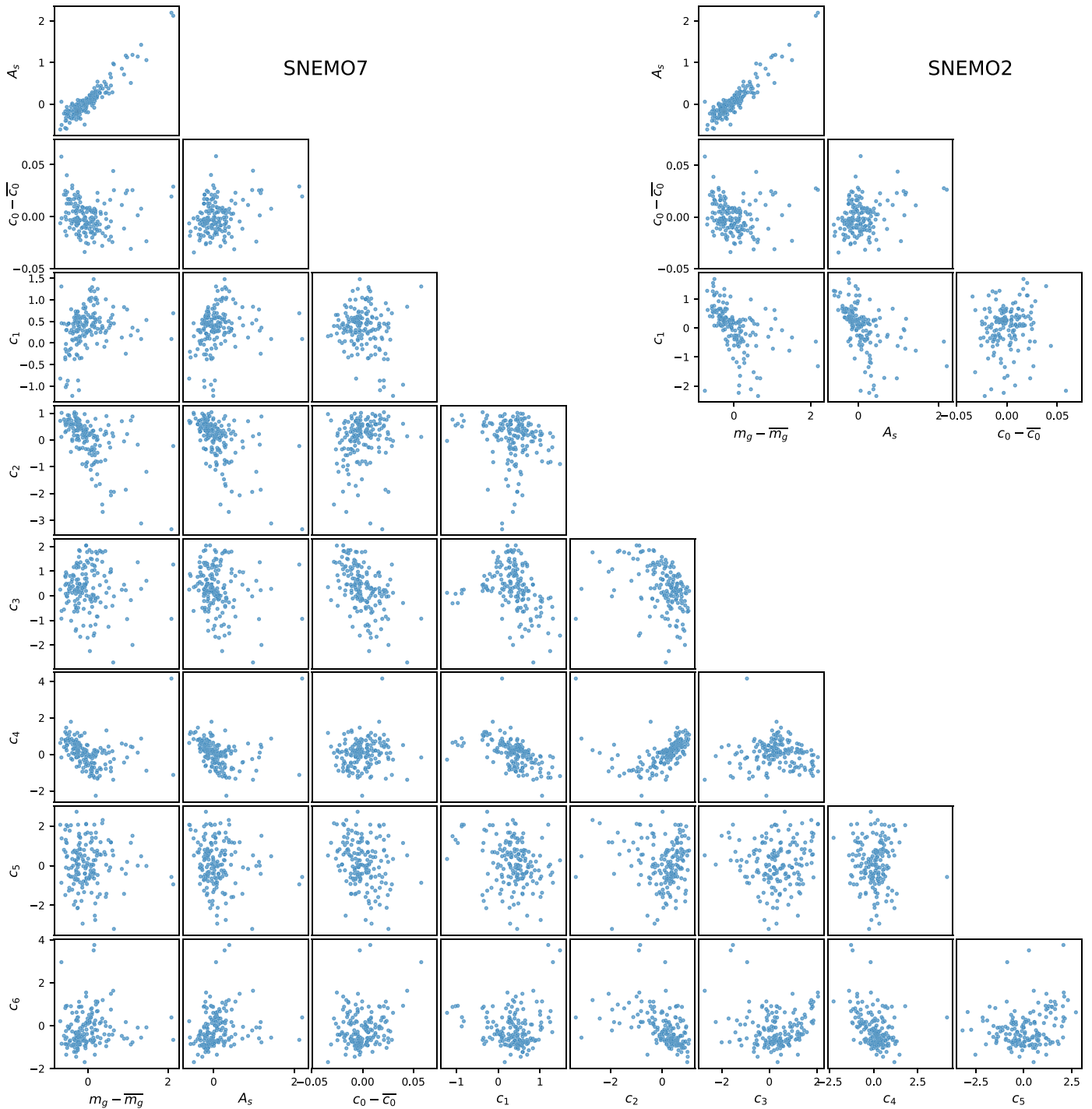


Figure 10. Coefficient distributions for SNEMO2 (small triangle in the top right) and SNEMO7 (large lower triangle).

(which corresponds with the x_1 parameter). The corresponding R^2 values are

$$\begin{aligned}
 R^2_{(\text{SNEMO2})} &= (0.994, 0.468) \\
 R^2_{(\text{SNEMO7})} &= (0.995, 0.673) \\
 R^2_{(\text{SNEMO15})} &= (0.996, 0.694).
 \end{aligned}
 \tag{20}$$

The low value of R^2 for the second component of SALT2 is further indication of the differences between the SNEMO

models and SALT2 on a spectral level, despite the correlations in the coefficient distributions in the supernova population.

4.3. Minimizing Unexplained Magnitude Dispersion with SNEMO7

The seven-component model, SNEMO7, demonstrates the improvements made in the linear standardization of magnitudes. As described in Section 3.4.2, we find a standardized magnitude by fitting a linear relation between the model coefficients and the

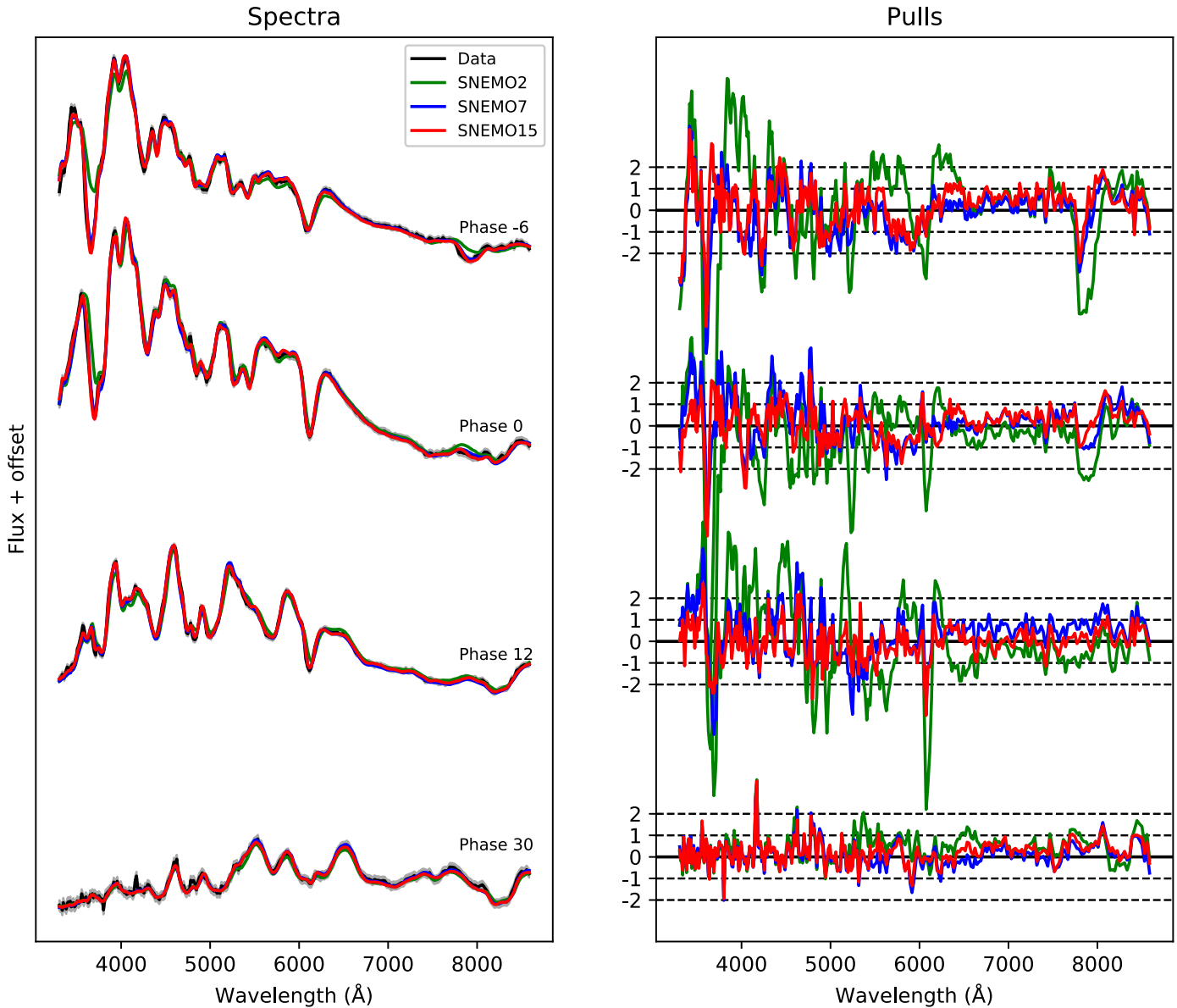


Figure 11. Model reconstructions of an out-of-sample supernova. The supernova spectra and models for four demonstrative phases are shown in the left panel, while the right panel shows the pulls (residual divided by the uncertainty) of the models.

g -band magnitude. The results are shown in Table 5 using the terminology defined in Section 3.4.2. The parameters of the linear magnitude correction are calculated using the final hold-out set, since in a cosmological analysis the corrections would be recalculated for that given set of supernovae. For comparison and as a check for consistency, we also show the dispersion found when calculating the corrections using the training set plus the hold-out set. Note that while the Gaussian processes step accounted for epoch-to-epoch flux calibration for an individual supernova, any calibration uncertainty between supernovae, aside from that due to peculiar velocity, is still part of the dispersion values found here.

Fitting the standardized magnitudes and the α and β standardization parameters for the same supernovae using the SALT2 model results in unexplained dispersions of $0.141^{+0.016}_{-0.021}$ mag (0.149 ± 0.019 mag total dispersion) when fitting the hold-out set and $0.140^{+0.009}_{-0.010}$ mag (0.150 ± 0.009 mag

total dispersion) when fitting the training set plus the hold-out set. For the SALT2 model, $\sigma_{lc} \sim 0.002$ mag, while σ_{ext} is the same as for SNEMO7.

The value for the unexplained dispersion is dependent on the values for σ_{lc} and σ_{ext} , but is not highly sensitive to changes in their values. For example, if the uncertainties in σ_{lc} were actually half of the value used here and reported in Table 5, the unexplained dispersion fit for all supernovae with 3σ cuts would increase from 0.097 to 0.103 mag. Changing the assumed value of the host-galaxy peculiar velocities used to calculate σ_{ext} by ± 100 km s $^{-1}$ changes the resulting unexplained dispersion by less than ± 0.01 mag.

As in the “Union” compilation (Amanullah et al. 2010), we remove supernovae with standardized magnitude pulls (the residual divided by the error) of more than 3. This removes one supernova from the training set and one supernova from the hold-out set, which are not otherwise outliers in the coefficient

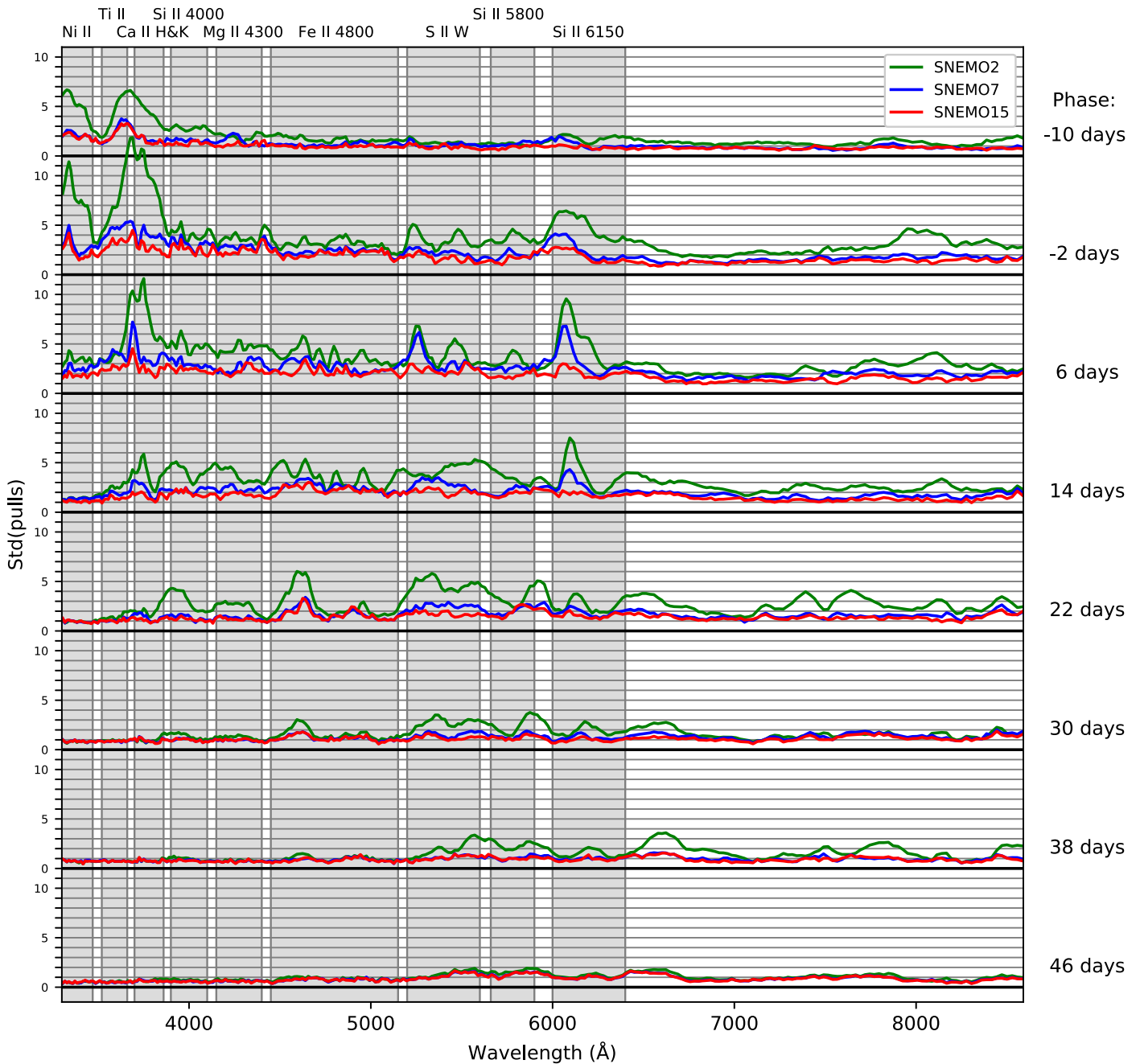


Figure 12. Standard deviation of the pulls between the three SNEMO models and the supernovae from the hold-out set at demonstrative phases, indicated at the right of the figure. The $\text{std}(\text{pulls}) = 1$ level indicates the level at which the residuals are on the same scale as the uncertainty in the data. The approximate regions of spectral features that strongly affect supernova spectra around maximum light are indicated by the vertical gray bands.

space. Similarly, one of these supernovae was also determined to be an outlier and was removed for calculating the SALT2 dispersion. The resulting dispersion is lower than the SALT2 values but not quite at a level equal to that of “twin” supernovae (F15, discussed further below), which found an unexplained dispersion of 0.072 ± 0.010 mag. (Note that the twinning technique makes no assumption about the linearity of the relation between the spectra or light curves and magnitudes.) This difference indicates that moving from linear standardization to a more sophisticated method may be necessary to maximize the possible improvement in standardized magnitudes. The standardized magnitude residuals are shown in Figure 14, compared to the residuals for the same supernovae using SALT2.

Table 6 shows the best-fit parameters for the linear correlations between the model coefficients and the g -band magnitude, corresponding to the α_i parameters in Equation (13). Those shown are the result of fitting to all of the data (the training plus hold-out sets) to give the most significant result. It can be seen that some of the components have a stronger significance, while the others are within $1-2\sigma$ of zero. This indicates that some components do not have a strong effect on the g -band magnitudes at maximum, though they may still affect other features of the light curve. Whether more data would enable significant correlations to be fit between all of the components and the g -band magnitude remains to be seen.

Table 4
Broad-band Characteristics of the Eigenvectors

Eigenvector	$g - r$ (mag)	Δm_{15} (mag)	Δt_{\max} (days)
SALT2			
Mean spectrum	-0.208	-0.918	-0.190
Eigenvector 1	-0.008 +0.008	+0.074 -0.079	-0.100 +0.080
SNEMO2			
Mean spectrum	-0.299	-0.921	-0.26
Eigenvector 1	+0.042 -0.039	+0.147 -0.150	-0.77 +0.35
SNEMO7			
Mean spectrum	-0.318	-0.927	-0.44
Eigenvector 1	-0.029 +0.025	+0.053 -0.050	-0.50 +0.29
Eigenvector 2	+0.051 -0.046	+0.189 -0.197	-0.48 +0.22
Eigenvector 3	-0.007 +0.007	+0.003 -0.003	-0.19 +0.17
Eigenvector 4	+0.021 -0.021	+0.028 -0.030	+0.06 -0.07
Eigenvector 5	-0.018 +0.018	+0.020 -0.020	-0.07 +0.07
Eigenvector 6	+0.004 -0.004	+0.009 -0.009	-0.05 +0.05
SNEMO15			
Mean spectrum	-0.330	-0.933	-0.53
Eigenvector 1	-0.065 +0.058	+0.012 -0.011	-0.35 +0.28
Eigenvector 2	+0.036 -0.033	+0.194 -0.204	-0.40 +0.20
Eigenvector 3	+0.012 -0.012	+0.016 -0.017	+0.17 -0.18
Eigenvector 4	+0.022 -0.023	+0.018 -0.018	+0.09 -0.10
Eigenvector 5	-0.001 +0.001	+0.023 -0.022	-0.07 +0.06
Eigenvector 6	-0.032 +0.031	+0.020 -0.020	+0.05 -0.05
Eigenvector 7	-0.005 +0.005	+0.002 -0.002	+0.08 -0.09
Eigenvector 8	-0.007 +0.007	+0.015 -0.016	-0.09 +0.08
Eigenvector 9	-0.014 +0.014	+0.000 -0.000	+0.13 -0.13
Eigenvector 10	+0.001 -0.001	+0.003 -0.003	-0.06 +0.07
Eigenvector 11	-0.005 +0.005	+0.006 -0.006	+0.00 +0.01
Eigenvector 12	-0.011 +0.011	+0.009 -0.009	+0.01 +0.00
Eigenvector 13	+0.002 -0.002	+0.000 -0.000	-0.10 +0.11
Eigenvector 14	+0.006 -0.006	+0.002 -0.002	+0.02 -0.02

Table 5
Final Results for the Dispersion in g -band Standardized Magnitudes Using SNEMO7

Data Set	Median σ_{ext} (mag)	Median σ_{lc} (mag)	σ_{μ} (mag)	σ_{tot} (mag)
Hold-out set	0.041	0.021	$0.121^{+0.023}_{-0.017}$	0.128 ± 0.020
Hold-out set with 3σ cuts	0.041	0.020	$0.087^{+0.021}_{-0.015}$	0.100 ± 0.014
All SNe	0.045	0.021	$0.110^{+0.010}_{-0.008}$	0.125 ± 0.010
All SNe with 3σ cuts	0.045	0.021	$0.097^{+0.009}_{-0.008}$	0.113 ± 0.007

Note. See Equation (13) for definition of terms.

4.4. Measures of the Reconstruction of the Spectral Time Series

Figure 15 shows the median and 95% confidence interval on the reduced χ^2 values for SNEMO2, SNEMO7, and SNEMO15 based on a bootstrap analysis of the training and hold-out set χ^2 distributions. As discussed in Section 3.4, we are more concerned here with the relative values of the reduced χ^2 , rather than the absolute value, which is known to be underestimated. The differences in the sizes of the uncertainties are partially driven by the different sample sizes and also by outliers, which can be seen in the histograms in Figure 16. The

Table 6
Corrected Magnitude Best-fit Parameters

Component	Parameter
α_c	-1.081 ± 0.040
α_1	-0.074 ± 0.014
α_2	-0.018 ± 0.021
α_3	-0.093 ± 0.015
α_4	-0.010 ± 0.017
α_5	-0.051 ± 0.010
α_6	0.034 ± 0.015

values for the training and hold-out sets are consistent with the results found in the cross-validation process. Reduced χ^2 values calculated in the same way for the SALT2 model are $7.68^{+1.72}_{-1.90}$ for the training set and $6.59^{+2.23}_{-1.36}$ for the hold-out set.

Examination of the supernovae from the hold-out set that still have somewhat large reduced χ^2 values when fit by SNEMO15 (see Figure 16) shows that most of these are peculiar supernovae. It may be that from the small number of SN 1991T-like objects in the training set, we do not have enough information to fully capture the behavior for this type of supernova. One outlier is PTF09dnp, Test_SN44 in Table 3, an otherwise normal object that is a rare case of a supernova with very narrow absorption lines where the Ca II H&K and Si lines can be separately resolved. It can be seen from the residual between the model and supernova spectra that this feature is driving the relatively large χ^2 , which is not unreasonable since this is the only supernova in our data set to have this feature. The quantity of such outlier χ^2 values gives an indication of the scale at which we have captured the full diversity of SN Ia spectral time series. It would be expected that an even larger sample of spectrophotometric time series would better incorporate this behavior, possibly requiring additional model components.

4.4.1. Comparison with “twin” Supernovae

The reduced χ^2 values for the SNEMO models can be compared with the analysis of “twin” supernovae from F15. In that analysis, pairs of supernovae were identified in the SNfactory data set that, after normalizing and dereddening, had very similar spectral time series. The concept underpinning that analysis was that such twin supernovae would have the same intrinsic luminosities and thus would be better standard candles, requiring little or no correction to relative distance measurements. These supernovae were called twins, and it was found that there was a dispersion of 0.072 mag among the magnitudes of twins, which is the best that has been found for a large sample of SNe Ia.

Here, we took Gaussian process predictions for the best matches from the supernova twins analysis and calculated the χ^2 values, showing how closely the spectral time series for a pair of twins match each other (after normalizing and dereddening the spectra, as in the twins analysis). The median value for these is shown as the red line in Figure 15, where it is at about the level of the median reduced χ^2 for SNEMO15 fit to the hold-out set.

This shows that at fifteen components, our model matches the supernova spectra at the same level that good twins match each other. The quality of the spectral matching implies that with SNEMO15 it should be possible to achieve the same level of magnitude standardization as with twins. The fact that the dispersion in the standardized magnitudes is poorer than for

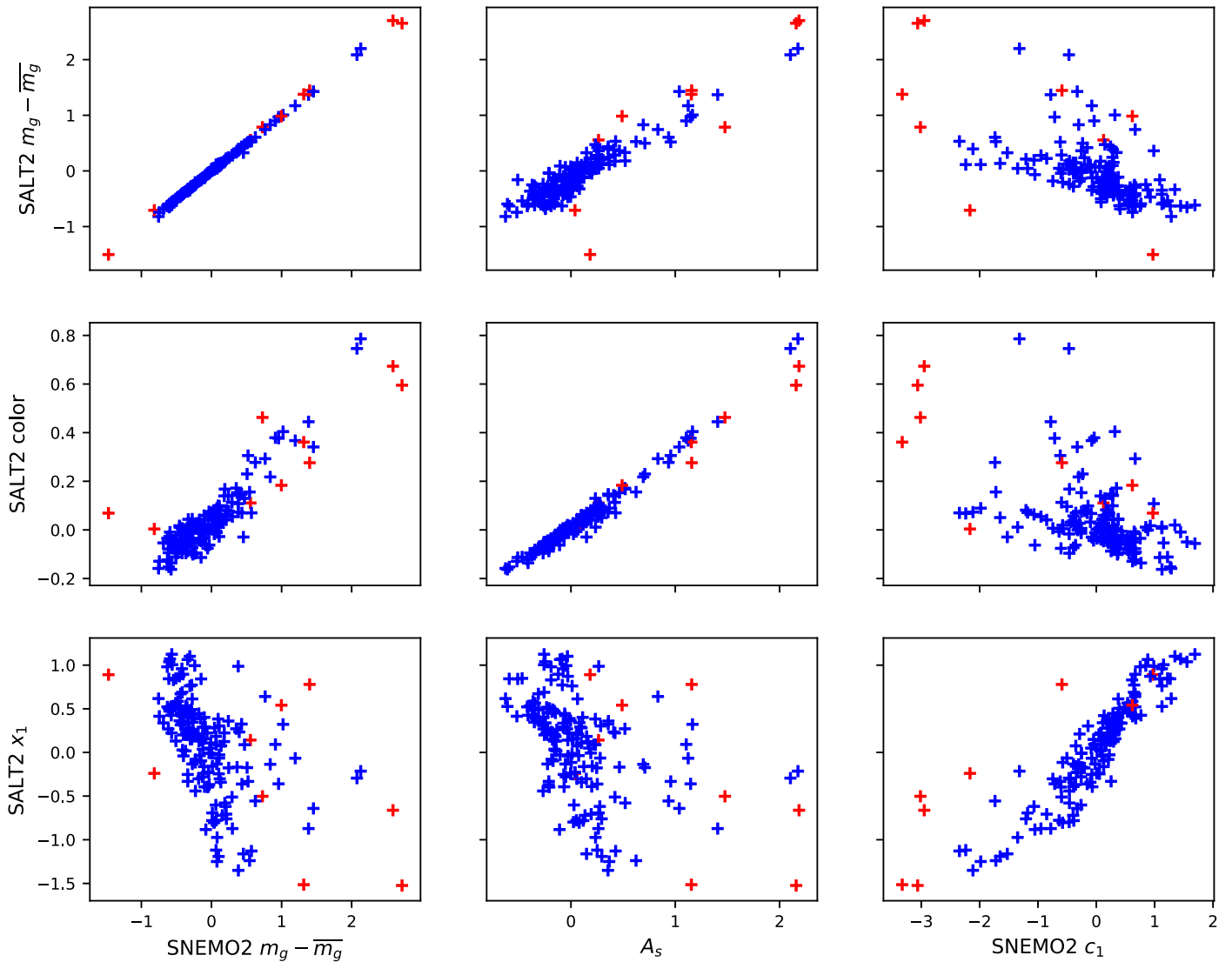


Figure 13. Coefficients for SNEMO2 vs. the model coefficients for the SALT2 model. Red points indicate supernovae that were identified as outliers.

twins, even though the reduced χ^2 is similar, points to the linear standardization method as the limit in improving magnitude dispersion with SNEMO. Improvements could be made in two ways: one way would be to fit supernovae with the model and identify cases that are very close to one another in the space of model coefficients, effectively finding twins through the coefficients of the model instead of directly fitting pairs of spectra. As with twins, one would then use the “model twins” as standard candles. The other possibility is to take a more holistic approach and find a mapping from the space of model coefficients onto the possible range of g -band magnitudes. If the coefficients are a unique identifier of the supernova spectral time series, then there must be some holomorphic map from them to the magnitudes, either by means of an analytic function (for example the linear method used in the standardization here) or something like a machine learning technique. This hypothetical mapping would need to be able to follow nonlinear and non-orthogonal relationships between the model coefficients and the supernova magnitudes. In principle, this should achieve a level of magnitude dispersion comparable with the twins method, to the extent that it describes an equivalent amount of the spectral time

series. However, such a mapping might not be discoverable without a larger data set than is available here.

5. Discussion and Conclusions

In summary, we have calculated three models for SNe Ia intended for a range of applications. The two-component model, SNEMO2, serves mainly as a tool for comparison with other two-component empirical models for supernova spectral time series currently in use. Though the components of the model appear to be quite different from the components in the SALT2 model, they describe similar effects in supernova behavior, which can be seen from the correlation in the model coefficients.

The seven-component model, SNEMO7, describes much more of supernova diversity and is also the most efficient among the models presented here for making linear corrections to the g -band magnitude and minimizing dispersion among supernovae, assuming that the components can be accurately fit by the data. The number of model components is small enough that it should be possible to meaningfully constrain this model to supernova data with only photometric light curves, though testing of the necessary temporal sampling and signal-to-noise is left for future study.

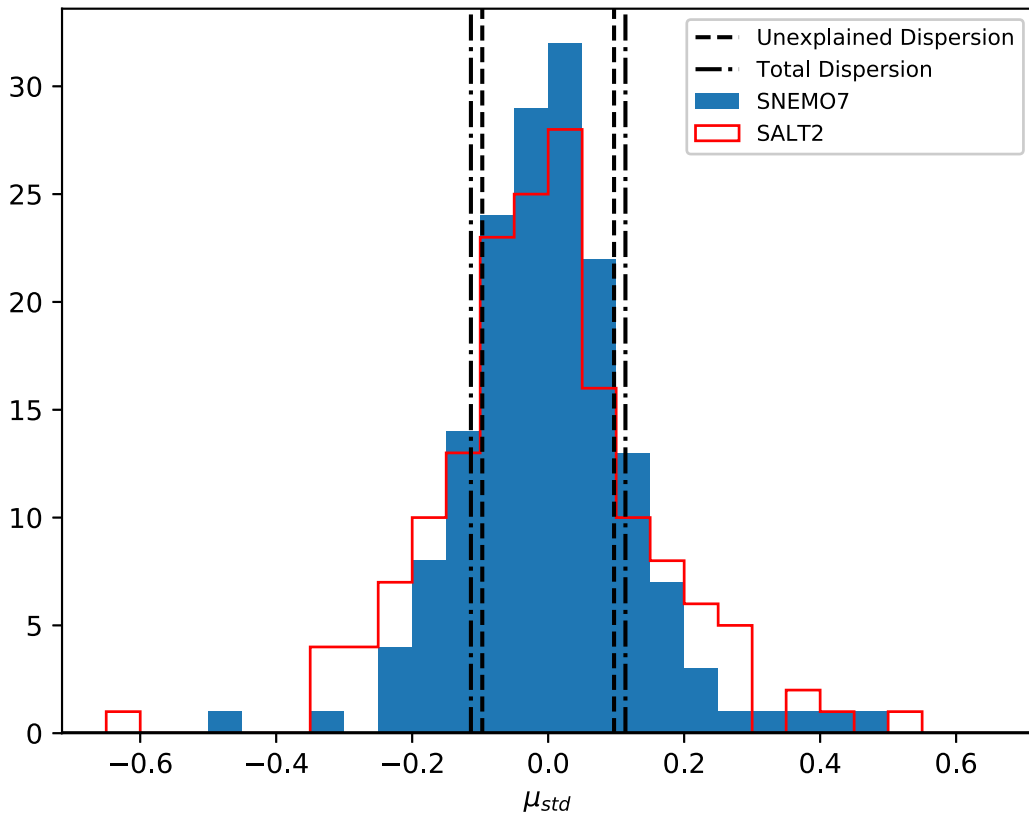


Figure 14. Resulting μ_{std} for SNEMO7 after minimizing Equation (13). The SALT2 μ_{std} for the same supernovae are shown as a comparison in red. The dashed lines show the level of the unexplained dispersion, while the dotted–dashed lines show the total dispersion.

Lastly, the fifteen-component model, SNEMO15, matches the spectral time series for all the supernovae in our data set on the same level that the best “twin” supernovae match each other. This model is particularly suited to producing realistic supernova simulations. Given the number of components, it will probably be difficult to fit the parameters of this model without at least some knowledge of the spectral behavior of the supernova being fit. If the data can sufficiently constrain the model parameters, it should also be possible to achieve magnitude standardization at the level of supernova twins, around 0.07 mag. A light-curve fitter that could match supernova spectra at this level would also remove certain systematic errors, such as those due to K -corrections discussed in S15, which may not be the case with SNEMO7. This improvement will be possible if the model parameters can be fit accurately enough to minimize the differences between the spectral-temporal model and the true supernova spectrum. If this accurate fitting could be done, the same standardized magnitude would be fit independent of the filters used, unlike what was found using current two-parameter light-curve fitters in S15.

Future work will determine how the models presented here can be best applied to observational data—whether external supernovae can be constrained by photometry alone or if spectral data will be required. Preliminary testing shows that the Si II velocity is recovered to within 700 km s^{-1} for SNEMO7 and within 630 km s^{-1} for SNEMO15 when the models are fit to a spectral time series, and simulations will show how well these numbers hold up with photometric data. We can also begin to test more advanced nonlinear standardization techniques, moving beyond fitting linear correlations between the magnitudes and the

model parameters, assuming that the number of supernovae available is sufficient to constrain such methods.

The models presented here will be incorporated into the supernova light-curve fitting Python package `sncosmo`¹⁸ (Barbary 2014), so that anyone can apply one of these models to their data. The coefficient distributions of the supernovae are published as part of the current paper so that the model can be used to create simulated data sets.

As described in Section 2, the hybrid standard star system used to calibrate the SNfactory data is in tension with data calibrated purely with the CALSPEC standards, an effect that must be accounted for in applying the current version of these models to new data. Since this calibration difference primarily affects the $U - B$ color, which is estimated to be ~ 0.02 mag lower in the system used here, there may be small systematic biases if the models are applied to data calibrated to CALSPEC or other calibration systems if the rest-frame U -band observations have a high weight in proportion to the rest of the data. However, work is in progress to allow the SNfactory data set to be recalibrated on a pure CALSPEC system, and when this is done the models presented here will be retrained to facilitate application to outside data sets (such as Betoule et al. 2014) that are calibrated to the CALSPEC system.

Supernova cosmology has relied on two-component models for over twenty years. For most of this time, models were limited by the lack of sufficient data to make any improvements upon this. We now have the potential to move past this barrier, with advances possible both for our understanding of SNe Ia and our capacity to measure dark energy.

¹⁸ <https://sncosmo.readthedocs.io>

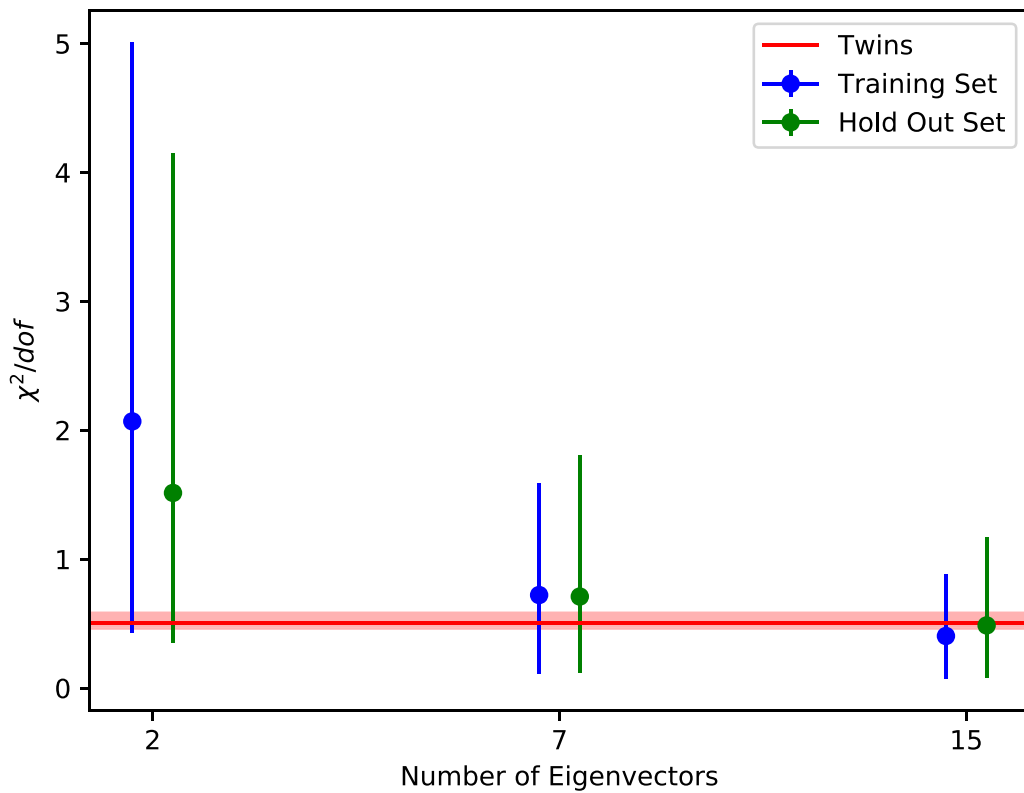


Figure 15. Median and 95% confidence interval for the reduced χ^2 of the three models in comparison to the reduced χ^2 of the best supernova twins. The blue points correspond to the result for the training set and the green points correspond to the smaller hold-out set supernovae. The red line is the median and 95% confidence interval of the reduced χ^2 between supernova twins. The corresponding reduced χ^2 values for the SALT2 model are $7.68^{+1.72}_{-1.90}$ for the training set and $6.59^{+2.23}_{-1.36}$ for the hold-out set.

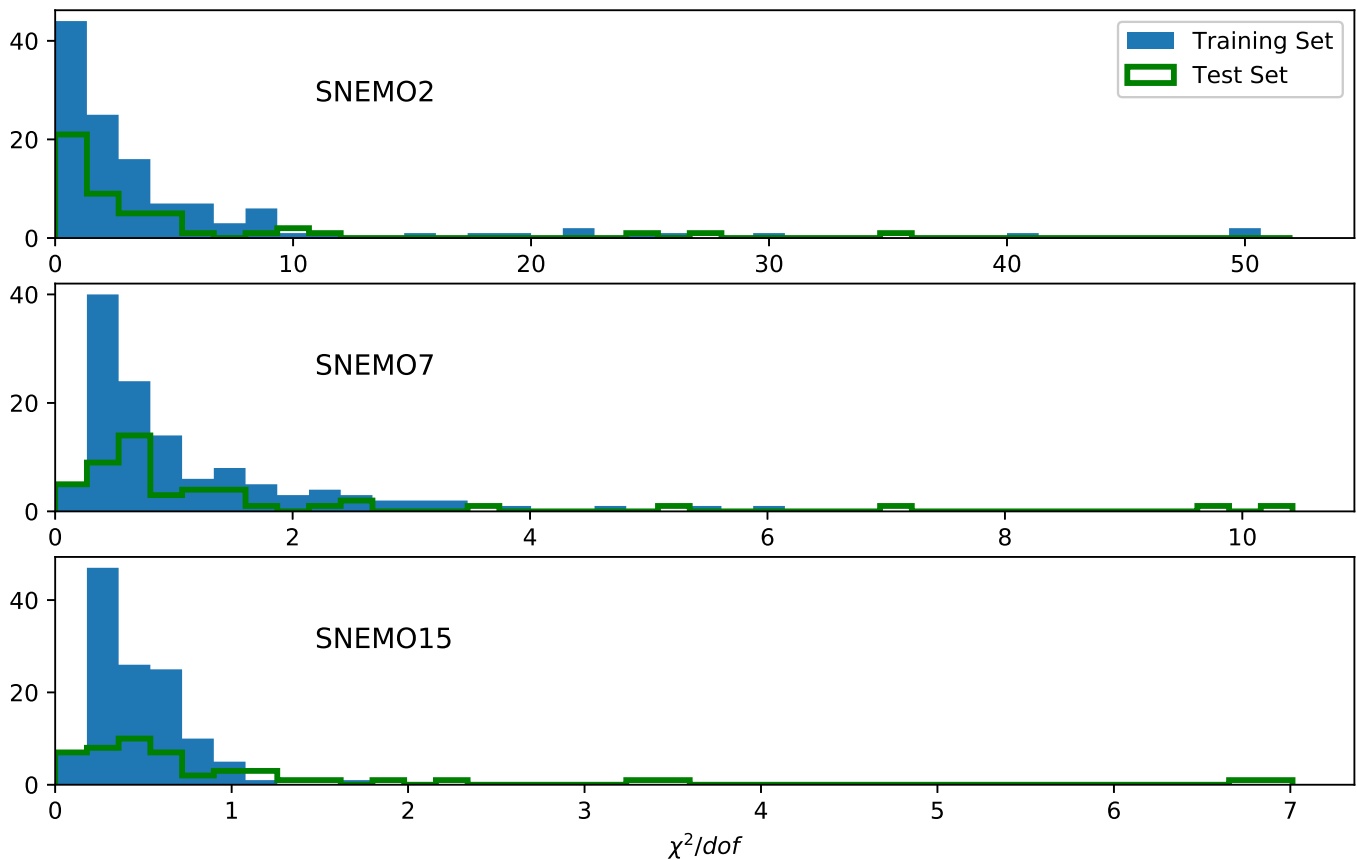


Figure 16. Three panels show normalized histograms of the reduced χ^2 of the supernovae for the three models in comparison. The histograms are normalized to more easily compare the training and hold-out set results. Note that the x-axis scales are different in each of the plots.

The authors are grateful to the technical and scientific staff of the University of Hawaii 2.2 m telescope, the Palomar Observatory, and the High Performance Wireless Radio Network (HPWREN). We thank Dan Birchall for his assistance in collecting data with SNIFS. We wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. This work was supported by the Director, Office of Science, Office of High Energy Physics, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; by a grant from the Gordon & Betty Moore Foundation; in France by support from CNRS/IN2P3, CNRS/INSU, and PNC; and in Germany by the DFG through TRR33 “The Dark Universe.” National Science Foundation grant No. ANI-0087344 and the University of California, San Diego provided funding for HPWREN. In China support was provided from Tsinghua University 985 grant and NSFC grant No. 11173017. C.S. was also supported by the Labex ILP (reference ANR-10-LABX-63) part of the Idex SUPER, and received financial state aid managed by the Agence Nationale de la Recherche, as part of the program “Investissements d’avenir” under the reference ANR-11-IDEX-0004-02. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement n759194—USNAC). Based in part on observations obtained with the Samuel Oschin Telescope and the 60 inch Telescope at the Palomar Observatory as part of the Palomar Transient Factory project, a scientific collaboration between the California Institute of Technology, Columbia University, Las Cumbres Observatory, the Lawrence Berkeley National Laboratory, the National Energy Research Scientific Computing Center, the University of Oxford, and the Weizmann Institute of Science.

Appendix

For calculating the correction function and the resulting dispersion in the standardized magnitudes we exclude the extreme outlier supernovae. To do this we calculate the Mahalanobis distance (Mahalanobis 1936), or normalized

metric distance, of each of the supernovae in the space of the model coefficients. The coefficient corresponding to the normalization of the model, c_0 , which is degenerate with the supernova magnitude, is not included in this calculation. Supernovae are excluded that have a probability of less than 0.001% according to the expected χ^2 distribution of the Mahalanobis distances. This procedure is then repeated once, since outliers affect the calculation of the Mahalanobis distances. This removed six supernovae from the standardization analysis. Three are 1991T-like supernovae, identified by the lack of Si II and Ca II H&K absorption before maximum and strong absorption from Fe II and Fe III (Filippenko et al. 1992a) and three are 1991bg-like supernovae, mainly identified by broad Ti II absorption around 4200 Å (Filippenko et al. 1992b). One of the SN 1991T-like supernovae is the super-Chandrasekhar mass supernova SN 2007if (here called Train_SN46), the SNfactory observations of which have been previously discussed in Scalzo et al. (2010).

Such supernovae, assuming they could be correctly identified, would generally be excluded in current cosmological analyses because they are not expected to be well fit by current light-curve fitters like SALT2 (Amanullah et al. 2010), though future nonlinear methods for magnitude standardization may allow these types of supernovae to be brought back into the set used for cosmology. Since these are not currently useful for cosmology, there is some motivation to exclude them from the entire analysis: the model might be able to describe “normal” supernovae better or with fewer components if it was not required to also fit these peculiar supernovae. Testing showed that removing these supernovae does change the individual components of the models, but does not lead to a significant improvement in the χ^2 of the model or the resulting dispersion in the standardized magnitudes, except in the two-component model, where it did improve the χ^2 metric. K -fold cross-validation showed that the favored model for standardizing magnitudes and the best spectral model were at six and fourteen components respectively, as opposed to seven and fifteen when the peculiar supernovae are included. Thus, there is the benefit of needing one less component to describe the more normal supernovae. We present the resulting models using two, six, and fourteen components in Figures 17 through 19.

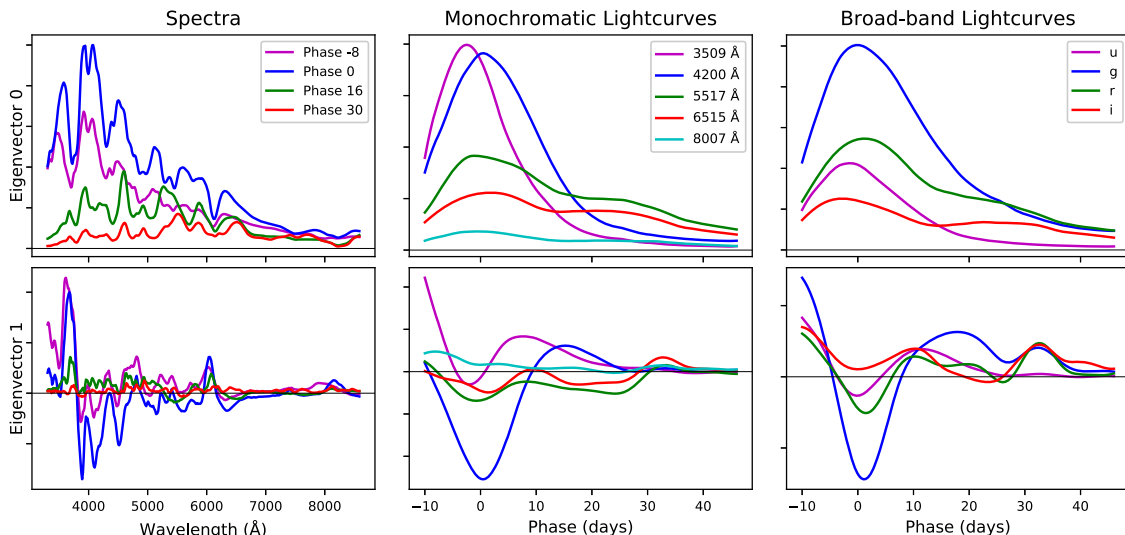


Figure 17. Demonstrative spectra and light curves for the eigenvectors of the two-component model trained with no outlier supernovae. See Figure 8 for comparison and for a description of the subplots.

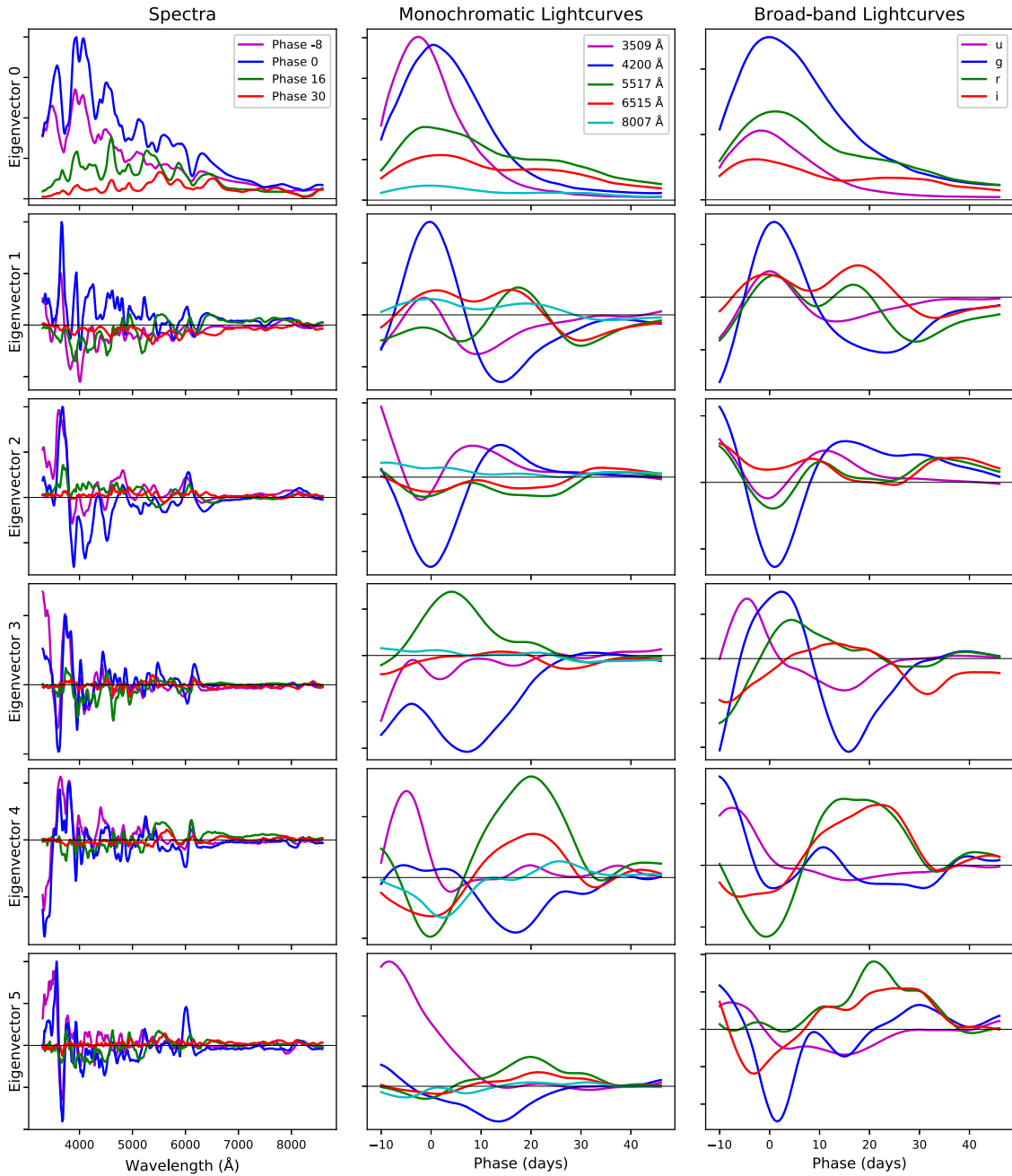


Figure 18. Demonstrative spectra and light curves for the eigenvectors of the six-component model trained with no outlier supernovae. See Figure 8 for a description of the columns.

Ultimately it was decided that excluding the peculiar supernovae did not provide enough benefits to outweigh the somewhat subjective decision to exclude them from the model training set. There is a continuum in the behavior that allows these supernovae to be identified as SN 1991T or SN 1991bg-like, and supernovae that share some of these features are seen in our data set that are not outliers in the model coefficient space. This means that to have a model trained only on “normal” supernovae, all of the SN 1991T or SN 1991bg-like supernovae

would have to be identified by eye in order to remove them, which would be dependent on some chosen limits and would probably be biased by the phase coverage of the supernovae, since SN 1991T-like supernova spectra appear normal at later phases. On balance, it is preferable to avoid subjective choices in training the model, with the benefit that since the model has been trained on peculiar supernovae it will be easier to identify these supernovae when the model is applied to a new supernova set, and to remove them if necessary for cosmological analyses.

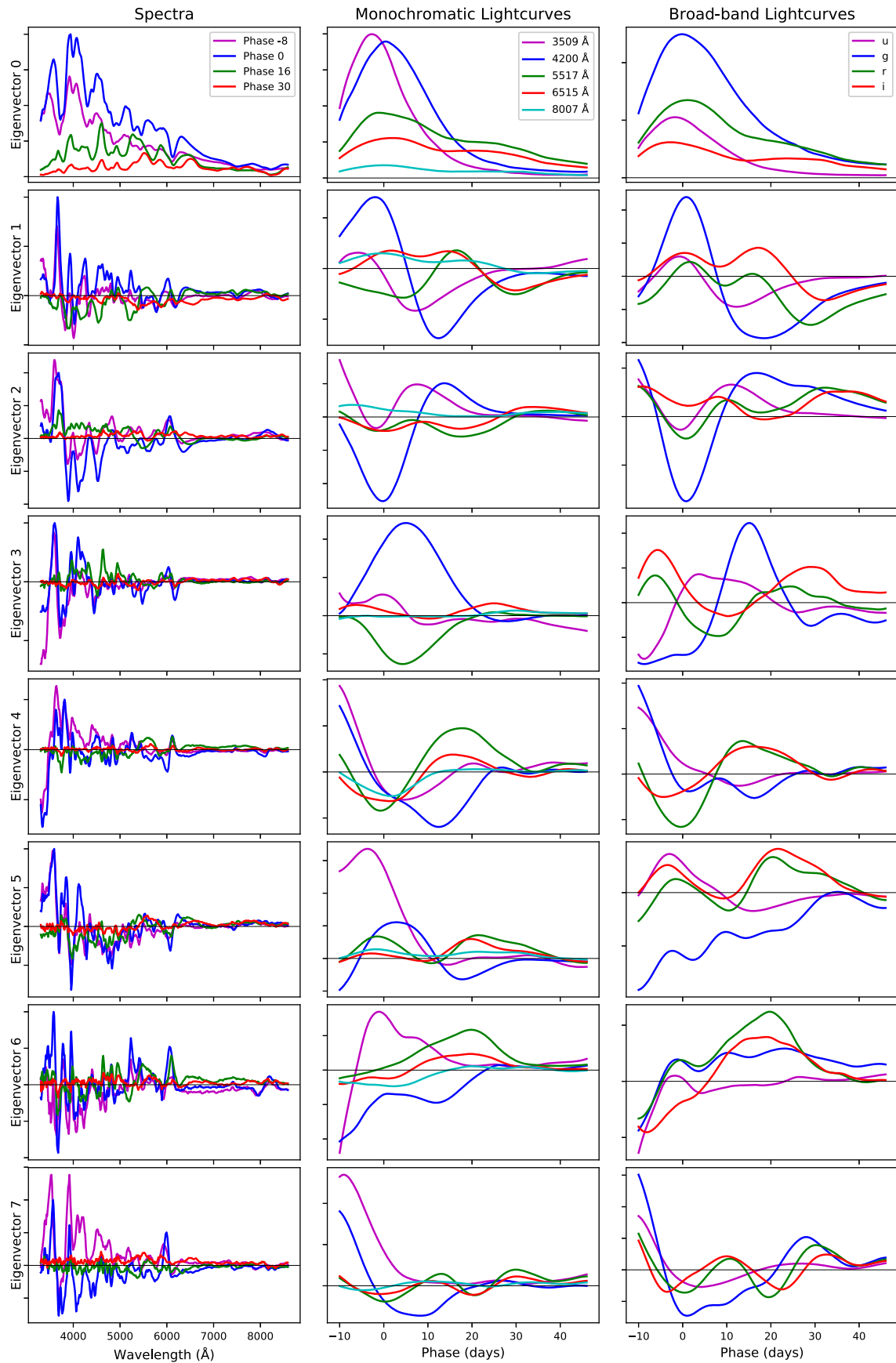




Figure 19. Demonstrative spectra and light curves for the first eight eigenvectors of the fourteen-component model trained with no outlier supernovae. See Figure 8 for a description of the columns.

ORCID iDs

K. Boone  <https://orcid.org/0000-0002-5828-6211>
 B. Hayden  <https://orcid.org/0000-0001-9200-8699>
 M. Rigault  <https://orcid.org/0000-0002-8121-2560>
 D. Rubin  <https://orcid.org/0000-0001-5402-4647>

References

- Aldering, G., Adam, G., Antilogus, P., et al. 2002, *Proc. SPIE*, 4836, 61
 Amanullah, R., Lidman, C., Rubin, D., et al. 2010, *ApJ*, 716, 712
 Bailey, S., Aldering, G., Antilogus, P., et al. 2009, *A&A*, 500, L17
 Barbary, K. 2014, sncosmo v0.4.2, Zenodo, doi:10.5281/zenodo.11938
 Barone-Nugent, R. L., Lidman, C., Wyithe, J. S. B., et al. 2012, *MNRAS*, 425, 1007
 Bernstein, J. P., Kessler, R., Kuhlmann, S., et al. 2012, *ApJ*, 753, 152
 Betoule, M., Kessler, R., Guy, J., et al. 2014, *A&A*, 568, A22
 Bohlin, R. C. 2014, *AJ*, 147, 127
 Bongard, S., Soulez, F., Thiébaud, É., & Pecontal, É. 2011, *MNRAS*, 418, 258
 Burnham, K., & Anderson, D. 2003, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (New York: Springer)
 Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2011, *AJ*, 141, 19
 Buton, C., Copin, Y., Aldering, G., et al. 2013, *A&A*, 549, A8
 Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245
 Childress, M., Aldering, G., Antilogus, P., et al. 2013, *ApJ*, 770, 107
 Chotard, N., Gangler, E., Aldering, G., et al. 2011, *A&A*, 529, L4
 Fakhouri, H. K., Boone, K., Aldering, G., et al. 2015, *ApJ*, 815, 58
 Filippenko, A. V., Richmond, M. W., Branch, D., et al. 1992b, *AJ*, 104, 1543
 Filippenko, A. V., Richmond, M. W., Matheson, T., et al. 1992a, *ApJL*, 384, L15
 Fitzpatrick, E. L., & Massa, D. 2007, *ApJ*, 663, 320
 Foley, R. J., Hoffmann, S. L., Macri, L. M., et al. 2018, *MNRAS*, submitted (arXiv:1806.08359)
 Ghahramani, Z., & Hinton, G. E. 1996, The EM algorithm for mixtures of factor analyzers, Tech. Rep. CRG-TR-96-1(Toronto: Univ. Toronto)
 Gupta, R. R., D'Andrea, C. B., Sako, M., et al. 2011, *ApJ*, 740, 92
 Guy, J., Astier, P., Baumont, S., et al. 2007, *A&A*, 466, 11
 Hamuy, M., Suntzeff, N. B., Heathcote, S. R., et al. 1994, *PASP*, 106, 566
 Hamuy, M., Walker, A. R., Suntzeff, N. B., et al. 1992, *PASP*, 104, 533
 Hayes, D. S., & Latham, D. W. 1975, *ApJ*, 197, 593
 Howell, D. A., Sullivan, M., Brown, E. F., et al. 2009, *ApJ*, 691, 661
 Hsiao, E. Y., Conley, A., Howell, D. A., et al. 2007, *ApJ*, 663, 1187
 Huang, X., Raha, Z., Aldering, G., et al. 2017, *ApJ*, 836, 157
 Jha, S., Riess, A. G., & Kirshner, R. P. 2007, *ApJ*, 659, 122
 Kelly, P. L., Hicken, M., Burke, D. L., Mandel, K. S., & Kirshner, R. P. 2010, *ApJ*, 715, 743
 Kim, A. G., Thomas, R. C., Aldering, G., et al. 2013, *ApJ*, 766, 84
 Krisciunas, K., Phillips, M. M., & Suntzeff, N. B. 2004, *ApJL*, 602, L81
 Kulkarni, S. R. 2013, *ATel*, 4807, 1
 Lantz, B., Aldering, G., Antilogus, P., et al. 2004, *Proc. SPIE*, 5249, 146
 LSST Science Collaboration, Abell, P. A., Allison, J., et al. 2009, arXiv:0912.0201
 Mahalanobis, P. C. 1936, *Proc. National Inst. Sci. India*, 2, 49
 Mandel, K. S., Foley, R. J., & Kirshner, R. P. 2014, *ApJ*, 797, 75
 Mandel, K. S., Narayan, G., & Kirshner, R. P. 2011, *ApJ*, 731, 120
 Matérn, B. 1986, *Lecture Notes Stat.*, vol. 36 (Berlin: Springer)
 Nordin, J., Aldering, G., Antilogus, P., et al. 2018, *A&A*, 614, A71
 Nugent, P., Kim, A., & Perlmutter, S. 2002, *PASP*, 114, 803
 Papadogiannakis, S., Goobar, A., Amanullah, R., et al. 2018, *MNRAS*, in press (doi:10.1093/mnras/sty3301)
 Pereira, R., Thomas, R. C., Aldering, G., et al. 2013, *A&A*, 554, A27
 Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *ApJ*, 517, 565
 Phillips, M. M. 1993, *ApJL*, 413, L105
 Rasmussen, C. E., & Williams, C. K. I. 2006, *Gaussian Processes for Machine Learning* (Princeton, NJ: Univ. Press)
 Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *AJ*, 116, 1009
 Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, *ApJ*, 473, 88
 Rigault, M., Aldering, G., Kowalski, M., et al. 2015, *ApJ*, 802, 20
 Rigault, M., Brinnel, V., Aldering, G., et al. 2018, *A&A*, submitted (arXiv:1806.03849)
 Rigault, M., Copin, Y., Aldering, G., et al. 2013, *A&A*, 560, A66
 Sasdelli, M., Hillebrandt, W., Aldering, G., et al. 2015, *MNRAS*, 447, 1247
 Saunders, C., Aldering, G., Antilogus, P., et al. 2015, *ApJ*, 800, 57
 Scalzo, R., Aldering, G., Antilogus, P., et al. 2012, *ApJ*, 757, 12
 Scalzo, R., Aldering, G., Antilogus, P., et al. 2014, *MNRAS*, 440, 1498
 Scalzo, R. A., Aldering, G., Antilogus, P., et al. 2010, *ApJ*, 713, 1073
 Schlafly, E. F., Finkbeiner, D. P., Schlegel, D. J., et al. 2010, *ApJ*, 725, 1175
 Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, 500, 525
 Sullivan, M., Conley, A., Howell, D. A., et al. 2010, *MNRAS*, 406, 782
 Thomas, R. C., Aldering, G., Antilogus, P., et al. 2007, *ApJL*, 654, L53
 Thomas, R. C., Aldering, G., Antilogus, P., et al. 2011, *ApJ*, 743, 27
 Tripp, R. 1998, *A&A*, 331, 815