The Astrophysical Journal Supplement Series, 250:12 (39pp), 2020 September https://doi.org/10.3847/1538-4365/ab9a3b

Artificial Intelligence–Assisted Inversion (AIAI) of Synthetic Type Ia Supernova Spectra

Xingzhuo Chen[1,2,3], Lei Hu[3], and Lifan Wang[1]
[1] George P. and Cynthia Woods Mitchell Institute for Fundamental Physics & Astronomy, Texas A&M University, Department of Physics and Astronomy, 4242 TAMU, College Station, TX 77843, USA; lifan@tamu.edu
[2] College of Physical Science and Technology, Sichuan University, Chengdu, 610064, People's Republic of China
[3] Purple Mountain Observatory, Nanjing 210008, People's Republic of China
Received 2019 November 13; revised 2020 June 3; accepted 2020 June 5; published 2020 September 3

Abstract

We generate ∼100,000 model spectra of Type 1a supernovae (SNe Ia) to form a spectral library for the purpose of building an artificial intelligence–assisted inversion (AIAI) algorithm for theoretical models. As a first attempt, we restrict our studies to the time around *B*-band maximum and compute theoretical spectra with a broad spectral wavelength coverage from 2000 to 10000 Å using the code TARDIS. Based on the library of theoretically calculated spectra, we construct the AIAI algorithm with a multiresidual convolutional neural network to retrieve the contributions of different ionic species to the heavily blended spectral profiles of the theoretical spectra. The AIAI is found to be very powerful in distinguishing spectral patterns due to coupled atomic transitions and has the capacity to quantitatively measure the contributions from different ionic species. By applying the AIAI algorithm to a set of well-observed SN Ia spectra, we demonstrate that the model can yield powerful constraints on the chemical structures of these SNe Ia. Using the chemical structures deduced from AIAI, we successfully reconstructed the observed data, thus confirming the validity of the method. We show that the light-curve decline rate of SNe Ia is correlated with the amount of ^{56}Ni above the photosphere in the ejecta. We detect a clear decrease of ^{56}Ni mass with time that can be attributed to its radioactive decay. Our code and model spectra are available on the website https://github.com/GeronimoChen/AIAI-Supernova.

Unified Astronomy Thesaurus concepts: Type Ia supernovae (1728); Convolutional neural networks (1938)

1. Introduction

Type 1a supernovae (SNe Ia) have been used as standard candles for cosmological probes (Riess et al. 1998; Perlmutter et al. 1999; Wang et al. 2003, 2006b; Rubin et al. 2013). The tremendous success is built upon empirical methods (Pskovskii 1977; Phillips 1993; Riess et al. 1996; Perlmutter et al. 1999). In contrast to such success, theoretical models of SNe Ia have not yet produced satisfactory fits with a precision matching that of observational data. In particular, it has been extremely difficult to establish a quantitative approach to reliably model the spectral features of SNe Ia and determine their luminosities based on physical models.

The root of these difficulties resides in the complexity of radiation transfer through SN Ia ejecta. The process involves quantitative models of the chemical and kinematic structures of SN ejecta and detailed radiative transfer calculations. The spectral features of SNe Ia are normally broadened to about 10,000 km s^{-1} by ejecta motion. The atomic lines are heavily blended such that it is hard or impossible to separate spectral features arising from different atomic transitions based on conventional spectral feature measurement. A large number of input parameters and physical uncertainties affect the spectral feature formation of theoretical models. It is difficult to explore the relevant parameter space in great detail to search for the models that provide the best fits to observations. This is true even for the simple code SYNOW (Branch et al. 2009). For example, if we want to optimize over 10 free parameters with each parameter modeled in 10 different parameter values, a grid search for the optimal model would require 10 billion models to be calculated. This is formidably difficult.

Further complications arise from the unclear explosion physics of SNe Ia. The origin of SNe Ia is believed to involve one or two white dwarfs in a binary system, either the merger of two white dwarf stars (double-degenerate scenario) or the accretion of a white dwarf from a nondegenerate companion star (single-degenerate scenario). For some SNe Ia, the double-detonation scenario seems to provide excellent fits to observations (Shen et al. 2018). It is possible that the progenitors of SNe Ia form a diverse class of objects, and we do not yet know how to relate different physical systems to observations.

The SN Ia explosion occurs when the temperature and density inside the progenitor white dwarf become appropriate for carbon ignition. At least for some spectroscopically normal SNe Ia (Branch et al. 2009), the explosion is likely to involve an early deflagration phase followed by a phase of detonation (the delayed-detonation (DDT) model; Khokhlov 1991a). The physics of the transition from deflagration to detonation is still not clear. Recently, a general theory of a turbulence-induced deflagration-to-detonation transition (tDDT) in unconfined systems was published by Poludnenko et al. (2011, 2019). The tDDT has the potential of generating a class of models based on first-principle explosion mechanisms, but its application to SN Ia modeling has not yet been explored. A variety of chemical elements are synthesized during the explosions. The decay of radioactive material serves as the energy source to power the radiation of the SNe.

Nonetheless, first-principle calculations based on parameterized explosion models and detailed radiative transfer may prove to be useful (Khokhlov 1991b; Hoeflich et al. 1996a, 2017), although the density at which the transition to detonation occurs is set as a free parameter in these models. Such models usually take days to weeks to calculate, even with today's fastest computer. This makes a thorough exploration of model parameter space impossible. For example, the progenitor metallicity, especially the C/O ratio, plays an important role in the production of radioactive material (Timmes et al. 2003), as

well as the density at the time of detonation and the mass of the progenitor. Departure from spherical symmetry may also make the models viewing angle–dependent (Wang et al. 1996; Wang & Wheeler 2008; Cikota et al. 2019; Yang et al. 2019). For these reasons, it is difficult to perform thorough parameter space searches to optimize the model fits to observational data. Currently, the best fits to observations are usually derived from only a small number of model trials, and there is ample room for further improvement.

As an example, W7 (Nomoto et al. 1984) is an early model of SNe Ia that is still widely used in SN Ia spectral syntheses. In W7 and other deflagration models (e.g., WS15, WS30; Iwamoto et al. 1999), the flame speed of the deflagration is 0.01–0.3 of the acoustic speed in the exploding white dwarf (Iwamoto et al. 1999). The DDT models were introduced to produce enough intermediate-mass elements (IMEs) and radioactive material and have been validated by radiative transfer modeling (Hoeflich et al. 1996a; Blondin et al. 2013) and comparisons to observational data. These models have now been extended to 2D and 3D hydrodynamic simulations (e.g., Gamezo et al. 2003). Several variations of the DDT model, such as the gravitationally confined detonation (GCD) model (Plewa et al. 2004), are proposed to address the diversity and asymmetry of the spectral behavior of SNe Ia (Kasen & Plewa 2005). The complexities intrinsic to these physical processes explain the difficulties in identifying exactly identical spectral twins despite the fact that the broadband photometries of the majority of all SNe Ia can be precisely modeled empirically with a one- or two-parameter light-curve family.

In this study, we want to focus on spectral fitting around optical maximum. Around this phase, the ejecta of SNe Ia can be assumed to be expanding homologously (e.g., Hoeflich et al. 1996a; Kasen et al. 2002). For 1D ejecta structure, the SN is spherically symmetric, and the radiative process can be modeled by Monte Carlo algorithms (Lucy 1971, 1999; Abbott & Lucy 1985; Mazzali & Lucy 1993). Other codes with varying levels of physical detail and complexity that have been applied to SNe Ia include Hydra (Hoeflich et al. 1996a), SYNAPPS (Thomas et al. 2011), PHOENIX (Baron & Hauschildt 1998), and CMFGEN (Hillier & Miller 1998). Many spectral models have been computed for some well-observed SNe Ia, e.g., SN 1990N (Mazzali et al. 1993), SN 1992A (Mazzali 2000), SN 1991bg (Mazzali et al. 1993), SN 2005bl (Hachinger et al. 2009), SN 1984A (Lentz et al. 2001), SN 1999by (Blondin et al. 2017), and, more recently, SN 2011fe (Mazzali et al. 2014). Sometimes both deflagration and detonation models are explored. Within the context of these models, the abundance stratification can be studied as the photosphere recedes in mass coordinates while the ejecta expand and become progressively optically thin (Stehle et al. 2005). This "abundance tomography" was applied to some well-observed SNe Ia, such as SN 2011fe (Mazzali et al. 2014), SN 2002bo (Stehle et al. 2005), and SN 2011ay (Barna et al. 2017). Also, 3D time-dependent radiation transfer programs such as SEDONA (Kasen et al. 2006) and ARTIS (Kromer & Sim 2009) have been constructed, which also enable calculations of polarization spectra (Hoeflich et al. 1996b; Wang & Wheeler 2008; Bulla et al. 2015).

Recent advancement in computer sciences, especially in artificial intelligence (AI), opens a new possibility for theoretical modeling of SN spectra. In general, a large number of parameters are needed to properly describe an SN Ia. These parameters need to cover critical ejecta properties, such as the density structure, chemical abundances at different layers, expansion velocity, radioactive heating, geometric symmetry, and temperature structure of the ejecta. We present in this paper an attempt to construct an artificial intelligence–assisted inversion (AIAI) model to study a library of theoretical spectra of SNe Ia around optical maximum. The AIAI trains a deep-learning neural network to inverse the modeling procedure and deduce the correlations between the resulting spectra and input model parameters.

The spectral lines of SNe Ia are usually blended and form "pseudocontinua" that make it very difficult to isolate contributions from individual ions for even the most isolated spectral features. Most spectral lines have contributions from multiple ionic species and atomic transitions; their profiles are strongly affected by the density and kinematic structures of the ejecta. The variations of spectral features are governed by fundamental physics, although it is difficult to establish a one-to-one match between a given atomic transition and the associated spectral line. The spectral features and their variation with time can be studied by voluminous realization of the atomic processes in numerous theoretical models under different physical conditions. A statistical study of the theoretical model realizations may very well unveil the hidden correlations between the atomic processes and the heavily blended observable spectral features. A neural network is by design remarkably suitable for such a study.

In the study, we find that a multiresidual neural network (MRNN; Abdi & Nahavandi 2016) can be trained to provide the best correlation between spectral features and input physical parameters. This MRNN is further tested with a different set of theoretical models to verify model stability and reliability. We then apply the MRNN to a set of observed spectra of SNe Ia to derive the model parameters for the observational data.

By calculating a large number of simulated spectra and using them as input to train the deep-learning neural network, we demonstrate that AIAI can indeed reveal the underlying chemical structures of SN Ia ejecta. Many of the radiative transfer codes for SN atmospheres are technically expensive, which prohibits a large amount of models from being calculated. For our purpose, and as a first attempt at AIAI, we choose the code Temperature And Radiation Diffuse In Supernovae, also known as TARDIS (Kerzendorf & Sim 2014), to generate the theoretical models. TARDIS is a 1D radiation transfer code using a Monte Carlo algorithm. In previous studies, TARDIS has been applied to the modeling of normal SNe Ia (Ashall et al. 2018), SNe Iax (Barna et al. 2017, 2018), SNe II (Vogl et al. 2019), and kilonovae (Smartt et al. 2017). The agreements to observations in these models have been moderately successful considering the simplicity of TARDIS and the limited coverage of the model parameter space.

We restrict our studies to SNe Ia with UV coverage. The UV is important, as it is very sensitive to the density and chemical structure of the SNe. Optical data alone may provide tight constraints on some IMEs but are less efficient in constraining iron group elements.

For a further test and application of AIAI, we apply the MRNN-trained models to a set of observational data. We generate neural network–matched spectral models for six SNe Ia (SN 2011fe, SN 2011iv, SN 2015F, SN 2011by, SN 2013dy, and ASASSN-14lp) with well-observed UV and

optical (2000–10000 Å) spectra and 15 SNe Ia with wavelength coverage of 3000–5200 Å around their *B*-band maximum luminosity. Moreover, we applied MRNN to predict the *B*-band absolute luminosities of these SNe Ia based on the ejecta structure deduced from their spectral profiles. Such predictions are still rough, due largely to approximations inherit to TARDIS, but future refinement may prove to be useful if these approximations are systematic and can be calibrated by other models with more complete treatment of radiative transfer physics.

The paper is structured as follows. The TARDIS configuration and model SN spectral library are introduced in Section 2. The neural network structure and its performance on synthetic spectra are discussed in Section 3. In Section 4, we apply the results of the neural network and present the resulting abundance and ejecta structure of a sample of SNe Ia near *B*-band maximum luminosity. In Section 5 we present further applications of AIAI and show the correlations between the ejecta structure and the luminosity of the SNe Ia with UV/optical coverage, the spectral evolution of SN 2011fe and SN 2013dy near maximum, and the predicted luminosities of all selected SNe in comparison with the absolute luminosities derived from well-calibrated optical light curves. Section 6 gives the conclusions and discussions.

2. The Generation of the Model Spectral Library

2.1. TARDIS Spectral Syntheses

Our study needs a library of SN spectra. This library of spectra should capture the radiative processes involved in SNe Ia as much as possible and cover a broad range of the physical properties of the ejecta. The parameter space is large, and the number of models to be calculated can be huge. Codes that are very CPU-demanding are apparently inappropriate for this study, at least for the current attempt, which is still an exploratory first step. For the determination of chemical structures, the primary requirement is that the code should be approximately correct in producing physical models to the spectral profiles of the most important observable ionic species. Luckily, there are several codes that fit this criterion. For example, the code SYNOW (Branch et al. 2009; Parrent et al. 2010; Thomas 2013) can run at very high speed and generate spectral libraries of a broad range of ejecta parameters. However, the code only allows a very crude description of the ejecta geometry. The input parameters are given in terms of the optical depth of some reference lines of certain ionic species. Quantitative constraints to parameters related to ejecta structure and SN luminosity are difficult using the available versions of SYNOW.

With the computing power available to us, the Monte Carlo code TARDIS (Kerzendorf & Sim 2014) is a good compromise that matches the requirement of our study. The input parameters for TARDIS include the elemental abundances and density structures of the ejecta, which can be flexibly modified. A spectrum can be calculated for any given day during the photospheric phase, once the location of the photosphere and the luminosity at that day are provided. TARDIS calculates the electron density, level population of ions and atoms, and temperature of different velocity layers. The code offers several options for photon transfer through the ejecta. In this study, we use dilute local thermodynamic equilibrium, `dilute-lte`, to calculate the atomic level population and nebular local thermodynamic equilibrium, `nebular`, to calculate the ionization fraction; both equilibria are part of TARDIS (Kerzendorf & Sim 2014). The program then evaluates the transition probabilities of atomic energy levels. TARDIS generates energy packets (an ensemble of photons with the same energy) at an inner boundary that propagate through the SN ejecta, calculate the electron optical depth by path-integrating the electron density on the trajectory, and simulate the photon–atom interaction optical depth by randomly sampling the atomic transition probabilities. When the optical depth of an energy packet reaches 1 during the electron scattering process, its energy and direction are reassigned following the Compton scattering process. When the optical depth of an energy packet reaches 1 during line interaction, its energy and direction are reassigned following the transition probabilities of atomic lines. By collecting the emitted energy packets, the code then compares the resulting luminosity to the input luminosity and updates the photospheric temperature and the temperature throughout the ejecta. There are two line-interaction strategies, `downbranch` and `macroatom`, available in TARDIS. In `downbranch` (Lucy 1999), photoexcited atoms are allowed to reemit photons with the same excited energy or other de-excitation transition energies, and the de-excitation channel is selected according to transition probabilities. Based on `downbranch`, `macroatom` (Lucy 2002) is a more sophisticated line-interaction strategy that allows upward and downward internal transitions with different probabilities for a photoexcited particle. Considering `macroatom` is closer to the physical reality while the computational time differences between two line-interaction strategies are not large, we choose `macroatom` for all of our calculations.

2.2. The Initial Guess Model

Our goal is to generate a large number of models that cover the parameter space of observed SNe as much as possible. For this purpose, we need a fiducial ejecta model, and we will perturb the parameters of that model to account for ejecta diversity.

We restricted our studies to models at 16–23 days after explosion, which corresponds to the time around optical maximum. We used the DDT model `5p0z22d20_20_27g` (Khokhlov 1991b; Hoeflich et al. 1996a) and the W7 model (Nomoto et al. 1984) as the ejecta density profiles to derive an initial guess model (IGM) for the ejecta structure (see Figure 1). These initial models were chosen for their success in providing reasonable fits to a number of SNe in previous studies (e.g., Hoeflich et al. 1996a). The exact details of these models are not important, as they only serve as a starting point for the ejecta structures and will be heavily modified later to generate the spectral library for further analyses.

The IGM was derived by comparing model spectral shapes with the data of two well-observed SNe: SN 2011fe and SN 2005cf. The spectrum of SN 2011fe was acquired at UT 2011 September 10 09:22:00 (0.39 day after *B*-band maximum luminosity) by the Hubble Space Telescope (HST) Space Telescope Imaging Spectrograph (STIS; Mazzali et al. 2014). The SN 2005cf spectrum was acquired at UT 2005 June 12 00:00:00 (the *B*-band maximum luminosity) from the SSO 2.3 m DBS (Garavini et al. 2007). The original chemical profile of (DDT) model `5p0z22d20_20_27g` is not optimized for SN 2011fe and SN 2005cf. The density and chemical structure

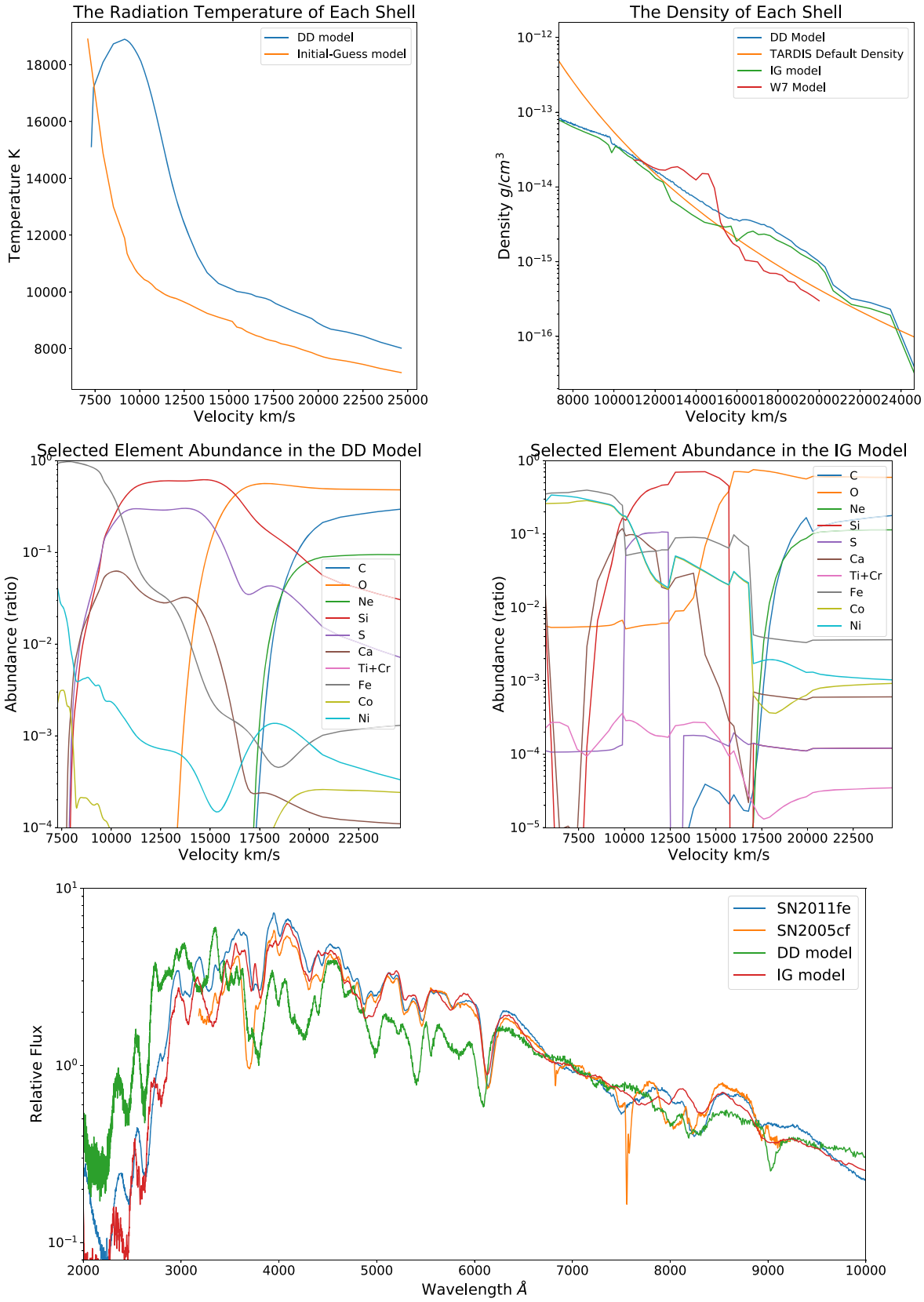

Figure 1. Upper left: temperature in each shell of the DDT model `5p0z22d20_20_27g` and the IGM calculated with TARDIS. Upper right: density profile of the DDT model, IGM, W7 model, and TARDIS's default seven-order approximated W7 model ($\rho = 3 \times 10^{29} \times v^{-7} \times (t/0.000231481)^{-3}$ g cm^{-3}, with t in days and v in km s^{-1}). All of the densities are normal SNe Ia at 19 days after explosion. Middle left: elemental abundances of the DDT model. Middle right: elemental abundances of the IGM. The elemental abundance ratios are normalized to 1 in each shell. Lower panel: TARDIS synthesized spectra of the DDT model and the IGM and the observed SN 2011fe/SN 2005cf spectra.

Table 1
Photospheric Velocity and Explosion Date

Days after Explosion	16	17	18	19	20	21	22	23
Photosphere velocity (km s^{-1})	8090	7850	7430	7100	6650	6290	6050	5690

of model `5p0z22d20_20_27g` was adjusted manually to match the strongest observed spectral features of SN 2011fe and SN 2005cf. Major changes to the elemental abundances were introduced when the profiles of strong lines such as Si II 6355 Å were examined closely. Increasing the iron group elements of DDT `5p0z22d20_20_27g` significantly improves the spectral match in the UV of these two SNe.

The observed spectra were found to be well fitted by a model with a photospheric velocity of $v_{\rm ph} = 7300\,{\rm km\,s^{-1}}$ and the integrated luminosity between 6500 and 7500 Å being $10^{8.52}$ times solar luminosity. Note that to ensure the convergence of the temperature structure, the temperature profiles were calculated with 40 iterations using the default temperature convergence parameters (`type:damped. damping_constant:1. threshold:0.05. fraction:0.8. hold_iterations:3. t_inner_damping_constant:1.`). The default iteration in TARDIS is 20, which is usually sufficient to reach temperature convergence.

The temperature and density profiles of the IGM model that fits the major spectral features of SN 2011ef and SN 2005cf are shown in the upper left and right panels of Figure 1, respectively. The chemical profiles of DDT `5p0z22d20_20_27g` and the IGM are shown in the middle left and right panels of Figure 1, respectively. The simulated spectra of the IGM are shown in Figure 1 (middle right), where we also show the spectra computed using DDT `5p0z22d20_20_27g` for comparison. Note that in Figure 1 (middle right), the flux levels of the models and data were arbitrarily scaled to match the spectral features. The fits to the observed spectra of SN 2011fe and SN 2005cf across most spectral lines and UV continua are apparently better for the IGM model than for the original DDT `5p0z22d20_20_27g` model. Note that we did not attempt to construct a quantitative model of the elemental structure at this stage. This manual step only modified the masses of a limited number of elements (listed in Section 3) with prominent spectral features to achieve a crude fit to the data.

2.3. The Model Spectral Library

For our deep-learning neural network, the IGM was used as a baseline model that was perturbed to generate the library of model spectra. It is impossible to build a complete model grid with varying elemental abundances at all velocity layers considering the large number of chemical elements involved, so we simplified the model into a limited number of velocity zones and used random sampling of parameter space to cover a broad range of physical possibilities. For the structure of the ejecta, we divided the ejecta into four distinct zones defined by different velocity boundaries; the velocity ranges of zones 1, 2, 3, and 4 are 5690–10,000, 10,000–13,200, 13,200–17,000, and 17,000–24,000 km s^{-1}, respectively. The ejecta structure includes all 23 elements with atomic numbers from 6 to 28. With four velocity zones, this leads to 92 free parameters on the masses of the chemical elements in each zone. The other parameters to run TARDIS that are not included in MRNN training are the luminosity, date of explosion, and photospheric velocity. The total number of parameter space dimensions is 95. It is impossible to construct a grid of models in such a vast dimension. The total number of models would reach a staggering value of $2^{95} \sim 4 \times 10^{28}$ even if only two grid values were sampled for each dimension.

The chemical compositions and densities were allowed to fluctuate independently within each velocity zone, but inside each zone, the velocity dependence of the density of each element is scaled to that of the same element of the IGM. For elements other than iron group, we used the following equation:

$$\rho_{ik} = \rho_{ik}^{\rm IGM} \times 3U_k, \quad (1)$$

where i denotes each different chemical element; k denotes the four different zones, where $k = 1$, 2, 3, and 4 stand for the velocity shells in the ranges 5690–10,000, 10,000–13,200, 13,200–17,000, and 17,000–24,000 km s^{-1}, respectively; and U_k is a random number drawn from a [0, 1) uniform distribution. The locations of the velocity boundaries were chosen to approximately match the major chemical layers in the IGM, and $\rho_i^{\rm IGM}$ is the density profile of element i in the IGM of shell i. The total density including all elements is calculated from the sum over all elements. With Equation (1), the masses of the elements in the four zones were artificially scaled from 0% to 300% relative to their respective values in the IGM.

Because zone 1 contains mostly Fe, Co, and Ni and is partially inside the location of the photosphere, we choose the variation of Fe, Co, and Ni in zone 1 to obey

$$\rho_{ik} = \rho_i^{\rm IGM} \times 3^{(1-2U_k)}, \quad (2)$$

where U_k is again a random number drawn from a uniform distribution between zero and 1. Equation (2) samples the chemical structures around $\rho_i^{\rm IGM}$ more frequently than large deviations from it. Hereafter, the elemental mass ratios between the library models and the IGM are delineated as "multiplication factors."

Another input parameter of TARDIS is the time after explosion. For the current study, we restricted the models to the time around optical maximum. This parameter was drawn from a uniform distribution between 16 and 23. The location of the photosphere was defined by interpolating the values given in Table 1, with an additional random number from a uniform distribution between −120 and 120 km s^{-1} to sample the range of possible variations of the location of the photosphere. The ranges of the location of the photosphere in Table 1 were derived using the IGM as input to TARDIS for which the corresponding photospheric velocities provide reasonable matches to the observed spectra of SN 2011fe between day 16 and day 23 after explosion. In this process, we found that the photosphere recesses approximately 300 km s^{-1} day^{-1}, which is consistent with the results of previous model fits by others (e.g., Mazzali et al. 2014).

TARDIS also requires the luminosity of the SN as an input parameter. For the spectral library, the luminosities in the range of 6500–7500 Å of the different models observe a uniform distribution in log space between $10^{8.64}$ and $10^{8.7}$ solar luminosity.

TARDIS assumes a predefined sharp inner boundary and does not calculate the gamma-ray transport in the ejecta. The radial temperature profile is only approximately correct, and in some regions may be severely incorrect. This caveat makes it difficult to constrain the physical quantities derived from TARDIS based on first-principle physics. The models are calculated independently at each epoch. This may introduce error when comparing the models with observed luminosities. However, we expect the spectral profiles to be governed by fundamental physics, and the model spectra bear the imprints of the atomic processes. We rely on AIAI to extract the fingerprints of these atomic processes through analyses of a large set of spectral models. Furthermore, some of the problems indigenous to TARDIS may be coped with by comparing a subset of TARDIS-generated models with more sophisticated radiative transfer codes such as PHOENIX (Baron & Hauschildt 1998), CMFGEN (Hillier & Miller 1998), and HYDRA (Hoeflich et al. 1996a) to identify systematic differences among these models. These systematic comparisons may lead to a large library of SN models with more physical processes being taken into account, although the calculations of a large library of PHOENIX, CMFGEN, or HYDRA models are too computationally expensive to be practical. We leave a comparative study of different models to future studies.

We used TARDIS's default temperature convergence strategy in calculating these models but with 15 temperature iterations (i.e., `type:damped. damping_constant:1. threshold:0.05. fraction:0.8. hold_iterations:3. t_inner_damping_constant:1.`). Additionally, in order to generate spectra with different signal-to-noise ratios, we set the number of Monte Carlo packets to vary from 1.5×10^5 to 2×10^6. The wavelength coverage of the model spectra was set to 2000–10000 Å.

2.4. Model Spectra Computation

The calculation of a single spectrum of the spectral library with about 10^7 energy packets takes approximately 0.5–2 CPU hr on our workstation, which is mounted with two Intel Xeon E5-2650 v4 CPUs. After about 1 month's calculation utilizing 80 cores (two chips of Intel Xeon E5-2650 v4 and one chip of Intel Xeon E5-2650 v1), a total of 99,510 spectra were generated with varying elemental abundances, photospheric velocities, explosion dates, and luminosities, as prescribed in the previous section.

2.5. Response of Spectral Profiles to the Variations of Input Abundances

The working hypothesis of this study is that the spectral profiles are sensitive to the input parameters and there is predictive power in the TARDIS models, albeit the spectral profiles may respond to the variations of chemical elements in a very complicated way. We demonstrate in Figure 2 that this is indeed the case. To show the effect of varying chemical elements, a single element at a given layer was artificially altered while all other parameters remained the same as in the reference model. Here the reference model was chosen from the spectral library that best matches SN 2011fe at day 0.4 (see Section 4). Figure 2 shows the results of artificially altering Fe and Ni by 0, 0.8, 1.5, and 3 times with respect to a reference spectrum.

We see that the spectral patterns of Fe and Co are very different. The spectra vary strongly with the UV wavelength and create spectral "wiggles" that are characteristic of the input chemical elements. The features generated by elements in different velocity zones are also distinctively different.

3. The Multiresidual Connected CNN Model

In this section, we apply the MRNN (Abdi & Nahavandi 2016) to the spectral library synthesized through the procedures outlined in Section 2 and train deep-learning models to infer the ejecta structure from the synthesized spectral library.

3.1. Model Data Preprocessing

The spectra generated by TARDIS are in units of erg s^{-1} $Å^{-1}$, which represents the luminosities of the SNe. They have a typical value of $\sim 10^{38}$ around optical maximum. This study explores the models in relative flux scale, as TARDIS is more reliable in modeling spectral features than absolute luminosity. The flux of each spectrum is normalized by dividing its average flux between 6500 and 7500 Å. Only the overall spectral shapes are of importance here. The absolute level of the flux is ignored throughout the neural network analyses. A discussion of the correlation between spectral features and luminosity will be given after the establishment of the neural network (see Section 5.3).

Moreover, to account for the problems related to TARDIS Monte Carlo noise when comparing to observational data, we further employed two methods for data augmentation: Gaussian noisification and Savitzky–Golay filtering (Savitzky & Golay 1964). Gaussian noisification is done by applying Gaussian noise for each wavelength bin following the formula $F_n(\lambda) = F_o(\lambda) \times (1 + \mathcal{N}(0, F_o(\lambda))/F_o^{1/2}(5500)/S)$, where $\mathcal{N}(0, F_o)$ is a normal distribution with $\mu = 0$ and $\sigma^2 = F_o$, S is a measure of the signal-to-noise ratio at the reference wavelength 5500 Å for the spectra with noise added, and F_n and F_o are the flux of the noise-added flux and the original TARDIS model flux, respectively. Savitzky–Golay filtering is achieved with smoothing windows randomly selected from (7, 9, 11, 13, 15, 17, 19, 21, 23, 25) and the order randomly chosen from (2, 3, 4, 5, 6). With this data augmentation strategy, we generated a new data set that is eight times larger than the original spectral library.

3.2. The Neural Network Structure

Deep-learning techniques have been developed for stellar spectroscopy (Fabbro et al. 2017; Bialek et al. 2019). For stellar spectroscopy, most of the spectral features are narrow and well separated, and the required network depth is minimal. Fabbro et al. (2017) applied a convolutional neural network (CNN) with two convolution layers and two fully connected layers to the infrared spectral data of the APOGEE (Nidever et al. 2015) sky survey program and, later, the Gaia-ESO database (Bialek et al. 2019). Liu et al. (2019) applied a CNN with eight convolution layers and two fully connected layers to mineral Raman spectrum classifications. Bu et al. (2019) utilized a five-layered CNN (combined with a support vector machine

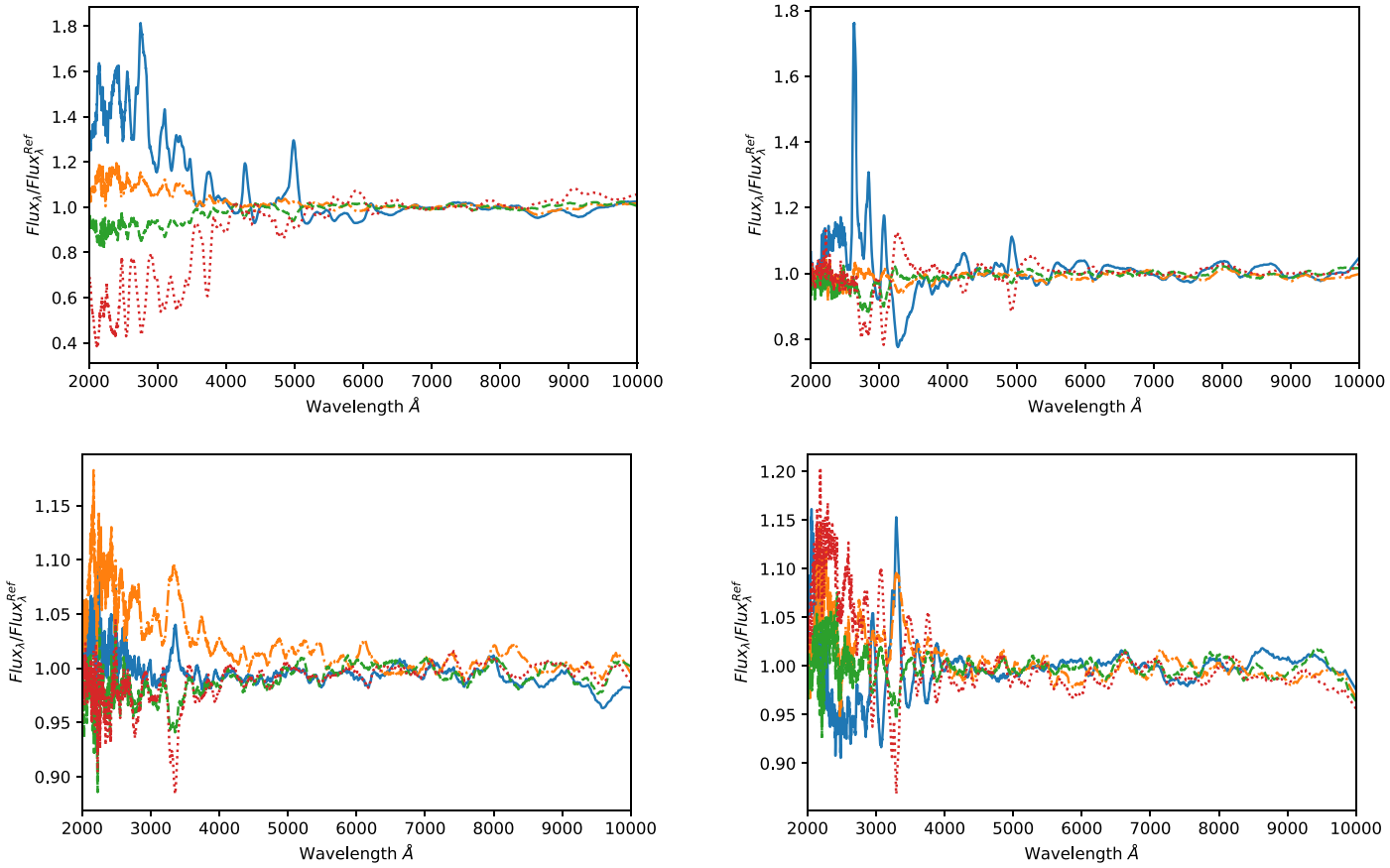

Figure 2. Effect on spectral profiles by only varying Fe in zone 2 (upper left), Fe in zone 3 (upper right), Ni in zone 2 (lower left), and Ni in zone 3 (lower right). The spectra were normalized by a reference spectrum that provides a good fit to the observed spectrum of SN 2011fe at day 0.4. Note that the solid blue lines with a scaling factor of zero are the model equivalent of removing that component from the ejecta. The dashed–dotted orange lines, dashed green lines, and dotted red lines are for models with that component enhanced by a factor of 0.8, 1.5, and 3, respectively.

(SVM)) to detect hot subdwarf stars from LAMOST DR4 spectra. In all of the above studies, the spectral features are much less blended than in the case of SNe. The SN spectra we aim to model are characterized by broad spectral lines due to the high velocity of the ejecta and the blending of atomic multiplets; models of SNe Ia require a more complicated neural network architecture to reach sufficient accuracy and sensitivity.

Between the input and output, a typical CNN contains stacked convolutional layers above one or two fully connected layers incorporated with pooling and activation layers. Starting from the input layer, the data propagate through the convolutional layers via convolution with convolution cores and pass through the activation layer with nonlinear functions (i.e., recitified linear activation $f(x) = 0$ when $x \leqslant 0$; x when $x > 0$). In the pooling layer, the dimensions of the data are reduced by binning adjacent data while preserving the maximum (MaxPooling method) or average (AveragePooling method) value. The fully connected layer shares the same structure as normal neural networks, which calculates linear combinations using the weights and biases between every neuron in the adjacent layers.

The trainable parameters (mainly the weights and biases in the fully connected layer and the convolution cores in the convolution layers) are randomly assigned at the outset and updated in the training process through forward and backward propagation. Forward propagation calculates the output using the current neural network parameters and input, then compares the network output (predictions) with the target values (the real values) under a loss function (i.e., mean squared error (MSE): $\text{Loss} = \text{Mean}(\hat{M}_{\text{scaled}} - \hat{M}_{\text{predict}})^2$). As all of the calculations in the neural network are analytical, the gradients of all trainable parameters can be deduced. In the back-propagation process, the trainable parameters are updated by multiplying a predefined learning rate by their gradients (for a review, see LeCun et al. 2015). Limited by computational efficiency, every forward–backward-propagation iteration only avails a subset of the training data set (batch), so a full review of the training data set consists of several batches.

It has been stated (Montúfar et al. 2014) and experimentally demonstrated (i.e., AlexNet, Krizhevsky et al. 2012; VGG16, Simonyan & Zisserman 2014) that additional layers endow the neural network with an exponential increase in agility. Nonetheless, a deeper network (a network with more layers) suffers from "gradient explosion" or "gradient diminish" problems: during the forward-propagation calculations, an input signal may be consecutively multiplied with <1 or >1 numbers, which results in the signal turn to zero or infinity in machine accuracy and affects the calculation on gradients. As a consequence, Ioffe & Szegedy (2015) introduced batch normalization layers, which normalize the input batch in every dimension to be average zero and standard derivative 1. Additionally, He et al. (2015) suggested adding the output from the previous convolution layers onto the current layer output and using this two-convolutional-layered structure as the

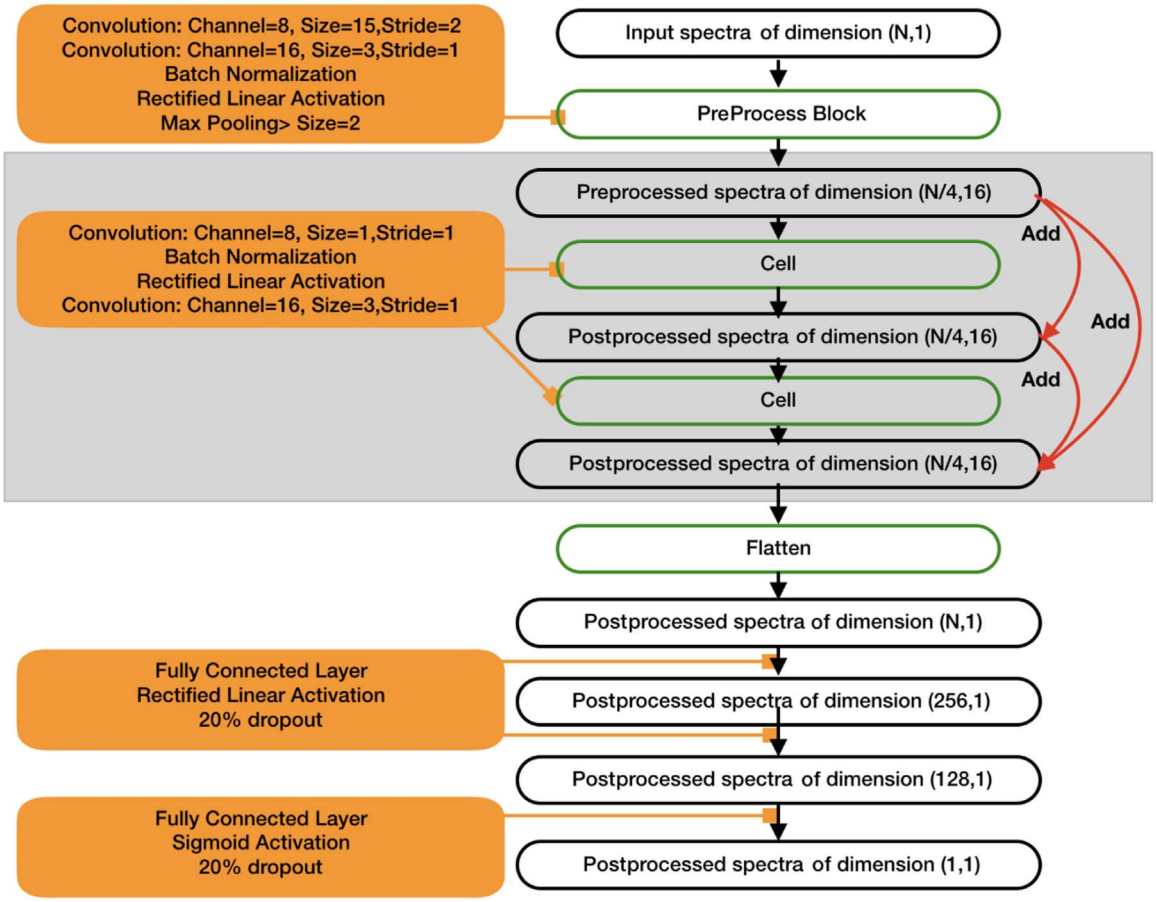

Figure 3. The MRNN structure. The black boxes are data tensors transferring in the network, and the green boxes are the "block" structures that contain multiple neural network layers whose details are shown in the yellow boxes. Note that the input and output of the "cell" structure have the same array size and can be repeated multiple times in MRNN, together with the boxes in the shaded area. The outputs of each cell are added with all of the inputs of previous cells to form the input of the next layer, as is marked by the red arrows. For the DNN, the "adding" processes are replaced with "concatenate" processes, so the data size of the intermediate, processed, and first fully connected layers is doubled.

building block for CNN structure. Such residual neural network (ResNet) structure shows great accuracy and is easy to optimize even in the 152-layer scenario (He et al. 2015). Based on ResNet, Huang et al. (2016) proposed a densely connected neural network (DNN) with another signal shortcut method: directly concatenate the output of all of the previous layers as the input of the current layer and insert low-dimension layers serving as "bottlenecks" to break the cumulative dimension increase caused by the consecutive concatenations.

Consequently, we adopt the MRNN structure (Abdi & Nahavandi 2016) for our studies. The structure of MRNN is relatively simple. Compared to ResNet, which allows a signal shortcut in two layers, MRNN introduces a signal shortcut in multiple layers by adding all of the previous layers' (or blocks') outputs together to be the next layer's input. Such an architecture is much simpler than DNN and requires less computation time to find a suitable network structure. Second, MRNN balances the training demands and accuracy. According to the test on the CIFAR-10 and CIFAR-100 image classification data sets (Krizhevsky 2009), MRNN shows an equivalent performance compared to a much deeper ResNet structure (Abdi & Nahavandi 2016) and better than all plain CNN structure available at that time. However, we notice that our CNN and MRNN with similar depth consume a comparable amount of CPU time to finish one epoch of training. After these preliminary assessments of the performance of CNN, ResNet, MRNN, and DNN, we chose MRNN to probe the best performance element prediction.

All of the models were trained using `keras`[4] `https://keras.io` with `tensorflow` (Abadi et al. 2016) as a back end. In the training step to build up the initial neural network structure, we selected the iron abundance in zone 3 and the TARDIS synthesized spectra in the wavelength range 2000–10000 Å. We chose MSE as the loss function and "adam" (Kingma & Ba 2014) as the optimizer. We adopted a two-step training scheme in order to achieve convergence while avoiding overfitting (the model performs exceptionally well on the training data set but fails on the testing data set). In the first step, the learning rate is 3×10^{-5} and decays 10^{-8} per time step, the batch size is 4000, and the training session jumps to the second step when there is no progress in the loss function of the testing data set after 10 epochs. In the second step, the learning rate is 3×10^{-7} and decays 10^{-8} per time step, the batch size is 10,000, and the training session stops when there is no progress in the loss function of the testing data set after five epochs.

The network architecture is shown in Figure 3. The neural network structure for the absolute luminosities is similar but with a different training schedule; the details are discussed in Section 5.3. We also tried to train the neural network to predict the photospheric velocities and the date of maximum; however,

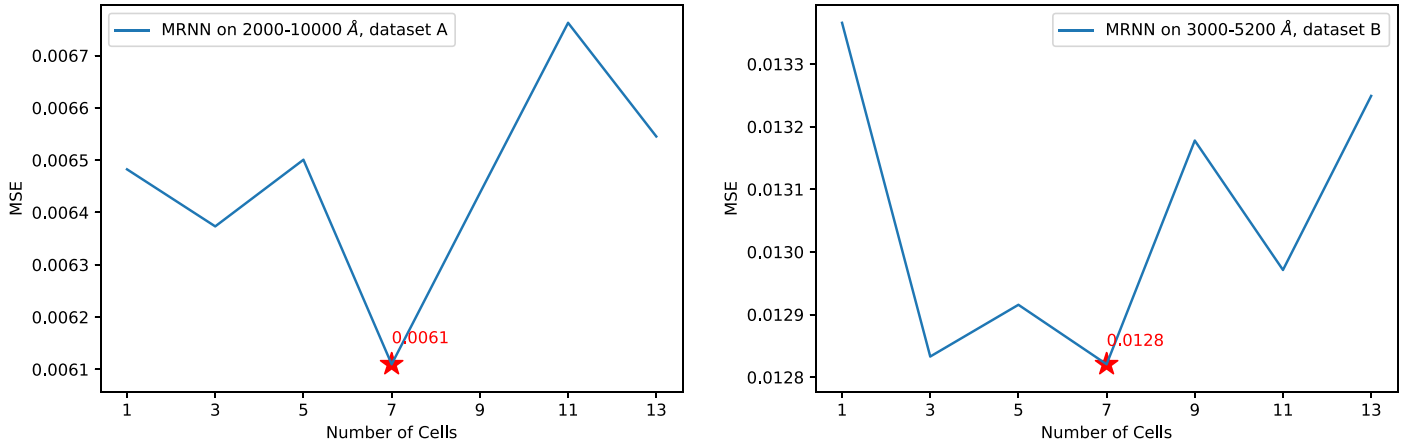

Figure 4. The MSE of Fe in zone 3 on the testing data set of the MRNN model. In this trial, 20,000 spectra were used in the training set, and 2666 spectra were used for testing. Left panel: MSE from the MRNN trained on 2000–10000 Å spectra (WR-Full). The MRNN with seven cells performs best, which shows a 0.0061 MSE. Right panel: MSE from the MRNN trained on 3000–5200 Å spectra (WR-Blue). The MRNN with seven cells performs best, with a 0.0128 MSE.

we failed to use these two predictions to refit the observed spectra. For the spectral fits in Section 5, we adopted the explosion date from the light curves and used a grid search to find the best photospheric velocities instead.

We trained the neural network using spectral data with two different wavelength ranges. One set used the entire spectral wavelength range from 2000 to 10000 Å (hereafter WR-Full), and the other only used the wavelength range from 3000 to 5200 Å (hereafter WR-Blue). These two sets of data are useful for applications to observational data with different spectral coverage.

In the next step, we chose the number of "cells" to be 1, 3, 5, 7, 9, 11, and 13 and explore the performances of MRNN. We randomly selected 20,000 spectra for training and reserved 2666 spectra for testing. We found that MRNN with seven cells performs the best among all models, as is shown in Figure 4. Increasing the number of cells actually deteriorates the fits.

We thus adopted the MRNN for the training on the chemical elements from 6 to 28 and the *B*-band absolute magnitudes. The chemical elements in the four different zones (plus the luminosity) were trained separately. In principle, the parameters can be trained jointly. However, the size of such a neural network could be too large for our computational capability, so we opted to choose a fast and robust neural network structure and applied it to all prediction tasks.

We have also tried the DNN (Huang et al. 2016) on the 2000–10000 Å models by simply changing the "adding" process in the MRNN by "concatenation." The MSE when using three cells for calculation is 0.0080, which is higher than MRNN with similar depth. However, due to our limited RAM capacity and the complexity in modifying DNN's cell structures, "bottleneck" (Huang et al. 2016) positions, and other hyperparameters, we did not explore more of this network structure.

3.3. Target Parameters for Machine Learning

Our goal is to find the ejecta structures that best fit the observed spectral features of an SN Ia at around optical maximum. We choose the "multiplication factor" (as mentioned in Section 2) in the zone of our concern as the neural network output. The "multiplication factor" was restricted to have a range from zero to 3. Our experiments with the models and data indicate that this range is sufficient to ensure coverage of a broad range of SNe Ia. Then, we applied the following normalization strategy in order to constrain the output of the neural network into (0, 1):

$$\hat{M}_{\rm scaled} = \tanh\left(\frac{(\hat{m} - \mu(\hat{m}))}{\sigma(\hat{m})}\right) \Big/ 2 + 0.5, \tag{3}$$

where $\hat{M}_{\rm scaled}$ is the scaled elemental abundances of the neural network output, and μ and σ are the average and standard derivative of the multiplication factor $\hat{m}$, respectively. This nonlinear normalization strategy allows the trained model to predict values outside the parameter space (extrapolation) within a small parameter range but suppresses erratic values derived by the neural network; elemental abundances less than zero or approaching infinity are remapped to values close to zero and 1, respectively.

3.4. Training Results

Not all the elements in the IGM are significantly influencing the spectra. As a first trial, we adopted a subset of model spectra and a simplistic neural network structure to probe the effect of various chemical elements. In this trial, 10,000 spectra were selected as the training set and 1829 spectra as the testing set, and we chose the MRNN with one cell structure. All elements from atomic number 6 (carbon) to 28 (nickel) in the four velocity zones are trained on the training set and verified on the testing set. By comparing the MSEs from the testing set, as well as the correlation between the neural network–predicted scaled elemental abundance (Equation (3)) and the corresponding value of the TARDIS input (Figures 5 and 6), we found models with MSEs larger than 0.1 to be poorly trained with little predictive power on elemental abundances. Consequently, we chose only the elements in zones with MSEs less than 0.1 for further training. There are 34 and 31 trainable chemical constituents located in the four velocity zones for training sets WR-Full and WR-Blue, respectively. The MSEs of these chemical constituents are shown in Tables 2 and 3 for training sets WR-Full and WR-Blue, respectively. In the tables, we also

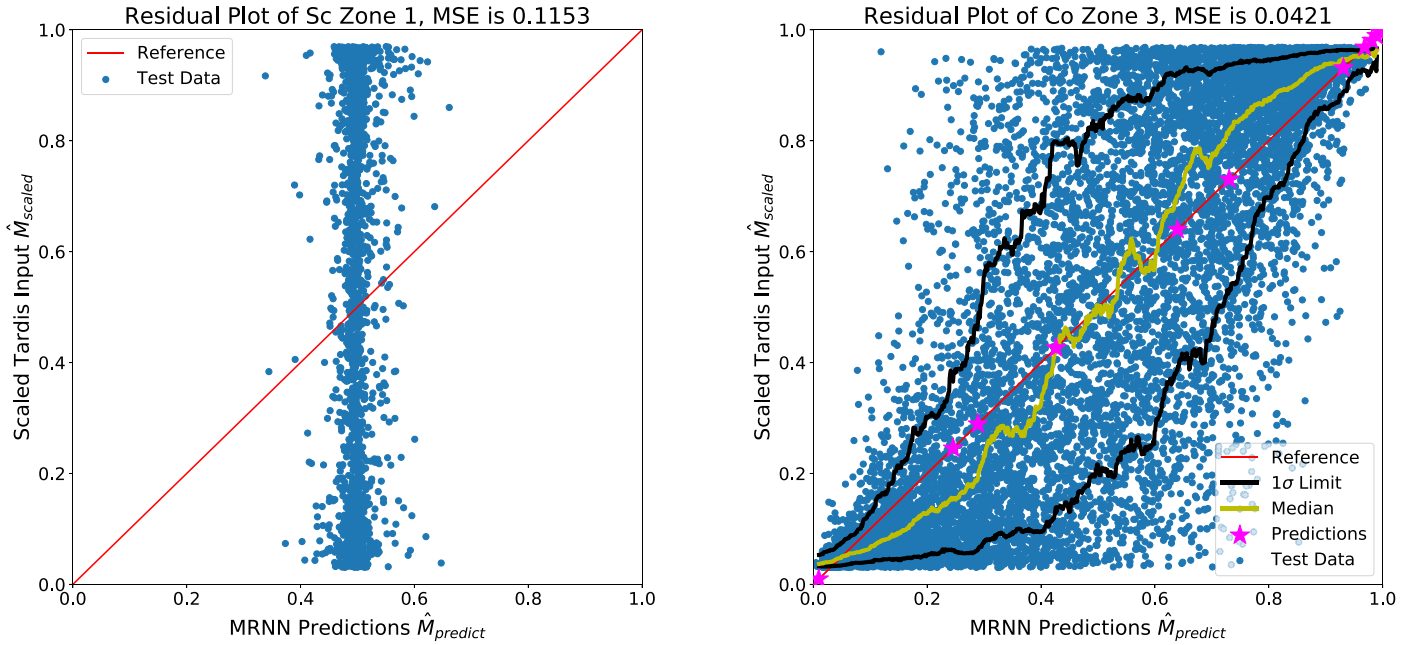

Figure 5. Correlation between the scaled elemental abundances of MRNN predictions and TARDIS inputs. The *X*-axis is the prediction from the MRNN model, and the *Y*-axis is the input into TARDIS for spectral calculations. In both panels, the diagonal red lines indicate the ideal model predictions, and the blue dots were derived from the test data set. The yellow and black lines show the median and $1-\sigma$ error limit of the given prediction in the testing set. Predictions from the selected 11 SN Ia spectra with wavelength coverage from 2000 to 10000 Å (which are analyzed in Section 4) are shown as magenta stars. Left panel: correlation plot of Sc in zone 1. The MSE is 0.1153. This model was trained on 10,000 spectra and tested on 1829 spectra. The predicted values are close to 0.5, which violates the "real" value for spectral synthesis, indicating the poor performance of the neural network on this element zone. Right panel: plot showing the correlation of the scaled elemental abundances of Co in zone 3, with an MSE value of 0.042. This model is trained on 89,559 spectra and tested on 9951 spectra.

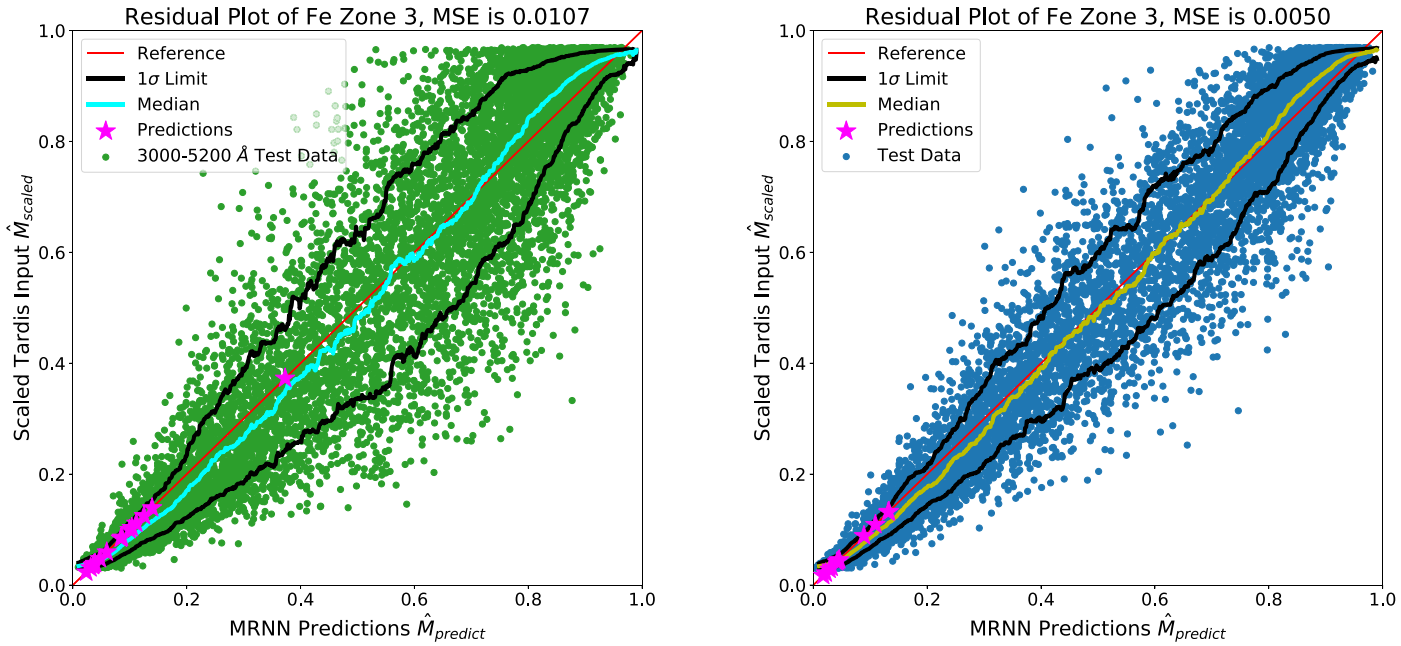

Figure 6. Same as Figure 5 but for Fe in zone 3. Left panel: correlation between scaled elemental abundances of MRNN predictions and TARDIS inputs for the neural network trained for the wavelength range 3000–5200 Å. The MSE is 0.0107. Right panel: same plot for the wavelength range 2000–10000 Å. The MSE is 0.0050.

show the correlation coefficients for each trainable element that can be used to gauge the reliability of the results.

As discussed in Section 3.2, we adopted the MRNN with seven cells for elemental abundance estimation. The training data set contains 89,559 spectra (90% of the total), and the testing data set contains 9951 spectra (10% of the total). For a typical neural network, it takes approximately 1 hr to finish 200 training epochs on two Tesla P100 GPUs.

In Tables 2 and 3, we list the MSE on the testing data set of the 34 and 31 selected elements and zones for two sets of neural networks built for WR-Full and WR-Blue, respectively. As an example, the scaled elemental abundances of the MRNN

Table 2
MSE and Correlation for Selected Elements and Zones of WR-Full (Wavelength Range 2000–10000 Å)

Element	Zone Number	MSE	Corr.	Element	Zone Number	MSE	Corr.
O	4	0.071	0.610	Mg	2	0.010	0.957
Mg	3	0.075	0.585	Mg	4	0.040	0.809
Si	1	0.051	0.747	Si	2	0.013	0.942
Si	3	0.019	0.915	S	2	0.067	0.641
Ar	2	0.092	0.425	Ca	1	0.054	0.726
Ca	2	0.018	0.920	Ca	3	0.048	0.762
Ca	4	0.019	0.913	Sc	4	0.058	0.698
Ti	1	0.086	0.498	Ti	2	0.055	0.718
Ti	3	0.046	0.772	Ti	4	0.044	0.787
V	1	0.050	0.748	V	2	0.019	0.909
Mn	1	0.026	0.875	Mn	2	0.017	0.926
Fe	1	0.004	0.984	Fe	2	0.005	0.981
Fe	3	0.050	0.978	Fe	4	0.025	0.885
Co	1	0.088	0.958	Co	2	0.029	0.864
Co	3	0.042	0.796	Co	4	0.078	0.567
Ni	1	0.009	0.960	Ni	2	0.028	0.869
Ni	3	0.050	0.754	Ni	4	0.075	0.586

Note. Neural networks are trained on the 89,559 spectra training set, and all MSEs are calculated on the 9951 spectra testing data set.

Table 3
MSE and Correlation for Selected Elements and Zones of WR-Blue (Wavelength Range 3000–5200 Å)

Element	Zone Number	MSE	Corr.	Element	Zone Number	MSE	Corr.
Mg	2	0.070	0.627	Mg	3	0.086	0.487
Mg	4	0.051	0.747	Si	1	0.085	0.511
Si	2	0.058	0.701	Si	3	0.053	0.728
S	2	0.071	0.607	Ar	2	0.093	0.420
Ca	1	0.062	0.480	Ca	2	0.072	0.605
Ca	3	0.073	0.605	Ca	4	0.052	0.739
Sc	4	0.062	0.680	Ti	2	0.066	0.653
Ti	3	0.061	0.685	Ti	4	0.046	0.773
V	1	0.087	0.499	V	2	0.046	0.773
Mn	1	0.069	0.640	Mn	2	0.058	0.702
Fe	1	0.037	0.808	Fe	2	0.024	0.893
Fe	3	0.011	0.952	Fe	4	0.036	0.829
Co	1	0.030	0.845	Co	2	0.045	0.783
Co	3	0.046	0.776	Ni	1	0.039	0.793
Ni	2	0.056	0.714	Ni	3	0.051	0.747
Ni	4	0.080	0.543				

Note. Neural networks are trained on the 89,559 spectra training set, and all MSEs are calculated on the 9951 spectra testing data set.

prediction and those of the TARDIS input of Fe in zone 3 are shown in Figure 6. The neural network–predicted chemical abundances clearly correlate with the values used to generate the model spectra. The neural network using the full 2000–10000 Å wavelength coverage outperforms the ones with partial coverage from 3000 to 5200 Å. The fact that the relation between MRNN prediction and the original TARDIS input elemental mass is approximately given by $\hat{M} \approx \hat{m}$ in all cases suggests that the training data set yields results that are consistent with the testing data set. This ensures that overfitting is not severely affecting the MRNN we have constructed. However, overfitting does appear to be an issue for elements that have weak spectral lines and where the correlation between the MRNN predicted and the original TARDIS input values are weak. This can be seen from Figure 5, where the predicted mass of Sc is biased toward the mean value of the true values of the TARDIS input (∼0.5). For Co in zone 3, the correlation can be detected, but the predicted values are biased toward values higher and lower than the true TARDIS inputs for input values close to zero and 1, respectively. Such bias is much weaker when the correlations are strong, as shown in Figure 6 for Fe in zone 3. This bias can, in principle, be corrected by using the median values of the scaled TARDIS input $\hat{M}$ to estimate the original model input.

In the current study, we did not correct this bias. Instead, we used the original MRNN predictions directly as the estimates of the mass of the input chemical elements but set the confidence levels of the estimates according to the correlations that can be derived from the test data set. We did not attempt to estimate the confidence intervals using Bayesian statistics based on a Markov Chain Monte Carlo algorithm (for example, see Foreman-Mackey et al. 2013), due primarily to the limited processing power of our computers. Instead, we estimated the $1-\sigma$ error by using the testing set itself. For each input model abundance, shown as examples in the Y-axes of Figure 6, we derive the $1-\sigma$ confidence levels based on the dispersion of

Table 4
Extinction and Stretch of SNe Ia with 2000–10000 Å Spectra

SN Name	Redshift	MW $E(B-V)$	Host $E(B-V)$	Stretch	References
SN 2011fe	0.000804	0	0	1.062 ± 0.005	Graur et al. (2018b), Munari et al. (2013)
SN 2011by	0.002843	0.013	0.039	0.93 ± 0.02	Maguire et al. (2012), Graham et al. (2015), Foley et al. (2018)
SN 2013dy	0.00389	0.135	0.206	1.098 ± 0.008	Zhai et al. (2016), Pan et al. (2015)
SN 2015F	0.0049	0.175	0.035	0.912 ± 0.005	Graur et al. (2018b)
SN 2011iv	0.006494	0	0	0.830 ± 0.007	Foley et al. (2012)
ASASSN-14lp	0.0051	0.33	0.021	1.101 ± 0.004	Shappee et al. (2016), Graur et al. (2018a)

the predicted values (the horizontal axis). To be more specific, we collected all of the data in the test data set that agree with the MRNN predictions to within the [−0.02, +0.02] range to build a histogram of the scaled TARDIS input mass and adopted the position of 15.8%, 50% (median), and 84.1% of the histogram as $1-\sigma$ error, median value, and 84.1% estimates. The $1-\sigma$ errors of Fe in zone 3 are overplotted in Figure 6 as an example. The errors for other elements in various zones are available in the online material.

Assuming that the TARDIS model spectra capture major spectral features of observed SNe Ia, we may apply the neural network trained by theoretical models straightforwardly to observational data. When an observed spectra is inserted into the well-trained MRNN, a set of chemical abundances and their associated uncertainties can be derived. Unfortunately, we do not have a large enough library of observed spectra to train the link between observational data and theoretical models. The consistency and reliability of the derived abundances can only be evaluated by comparing the results for different SNe and their spectral time sequences.

Alternatively, we may consider the chemical elements determined by feeding the observed spectra to the MRNN as empirical parameters that yield optimal TARDIS fits to the observed spectra. These parameters do not necessarily serve as true estimates of the elemental abundances of the SNe, but nonetheless, they can be used as model-based empirical parameters to analyze the properties of SNe Ia.

Note that the neural network approach is different from a simple observation-to-theory spectral match. The neural network model aims to match spectral features that are associated with differential chemical structure variations. Each chemical element in a certain zone is trained separately, and the neural networks thus trained are most sensitive to the changes of the spectral features related to this particular chemical element throughout the spectral range under study. Even though the spectral features of various elements are highly blended, the differential changes of spectral features due to varying chemical abundances are still detectable by the neural network we have constructed. Furthermore, the theoretical models may have intrinsic shortcomings and do not include all of the essential physics. The assumption of a sharp photosphere, for example, cannot be correct in a stricter sense. The lack of time dependence of the radiative transfer may also limit the precision of the theoretical models. However, our approach to elemental abundance may be less sensitive to these problems by construction, although the current study cannot establish a quantitative assessment of the uncertainties caused by approximations intrinsic to TARDIS models.

Moreover, the neural network allows for theoretical luminosity to be derived based on the theoretical model once the chemical structure of the ejecta is fully constrained. From the MRNN, the latter can be determined by spectral features without the need of knowing the absolute level of the spectral fluxes. When the global spectral profiles between observations and models are matched, the luminosity (or the temperature of the photosphere) is then completely constrained theoretically. This provides a theoretical luminosity of the SN understudy that is independent of the flux calibration of the spectra.

4. Applications of the Neural Network to Observational Data

Assuming the AIAI method constructed using MRNN in Section 3 to be correct, we may apply it to observed spectra as a sanity check of the method. We stress again that this step is detached from the deep-learning neural network and is not a validation of the MRNN.

The parameter space describing the ejecta structure is so large that our spectral library covers it only sparsely. The chance of having a perfect fit to any specific observation with spectra already in the library is low. The model spectra are thus recalculated using TARDIS with the ejecta structures determined from the MRNN. To derive the optimal TARDIS models, we also allowed the photospheric positions and the luminosities of the SNe to vary.

4.1. Applications to SNe with Wavelength Coverage 2000–10000 Å

There are six well-observed SNe with wavelength coverage from 2000 to 10000 Å. They are shown in Table 4. The UV data are all acquired by the HST. These data are rebinned to 1 Å pixel^{-1} and normalized by dividing their respective average flux between 6500 and 7500 Å, similar to what was done for TARDIS model spectra during neural network training.

4.1.1. SN 2011fe

On 2011 August 24, SN 2011fe was detected at the M101 galaxy at a distance of approximately 6.4 Mpc (Nugent et al. 2011). Its luminosity decline in the B band within 15 days after the B-band maximum is $\Delta M_{B,15} = 1.12 \pm 0.05$ mag (Munari et al. 2013). Based on optical and radio observations, SN 2011fe is not heavily affected by any interstellar (ISM)/circumstellar material (CSM) or Galactic dust extinction (Chomiuk et al. 2012; Patat et al. 2013). Consequently, we did not introduce any host galaxy extinction correction for SN 2011fe spectra.

We chose the HST spectra at −2.6, 0.4, and 3.7 days relative to the B-band maximum date for the elemental abundance calculations and spectral fittings. The results are shown in Figure 7.

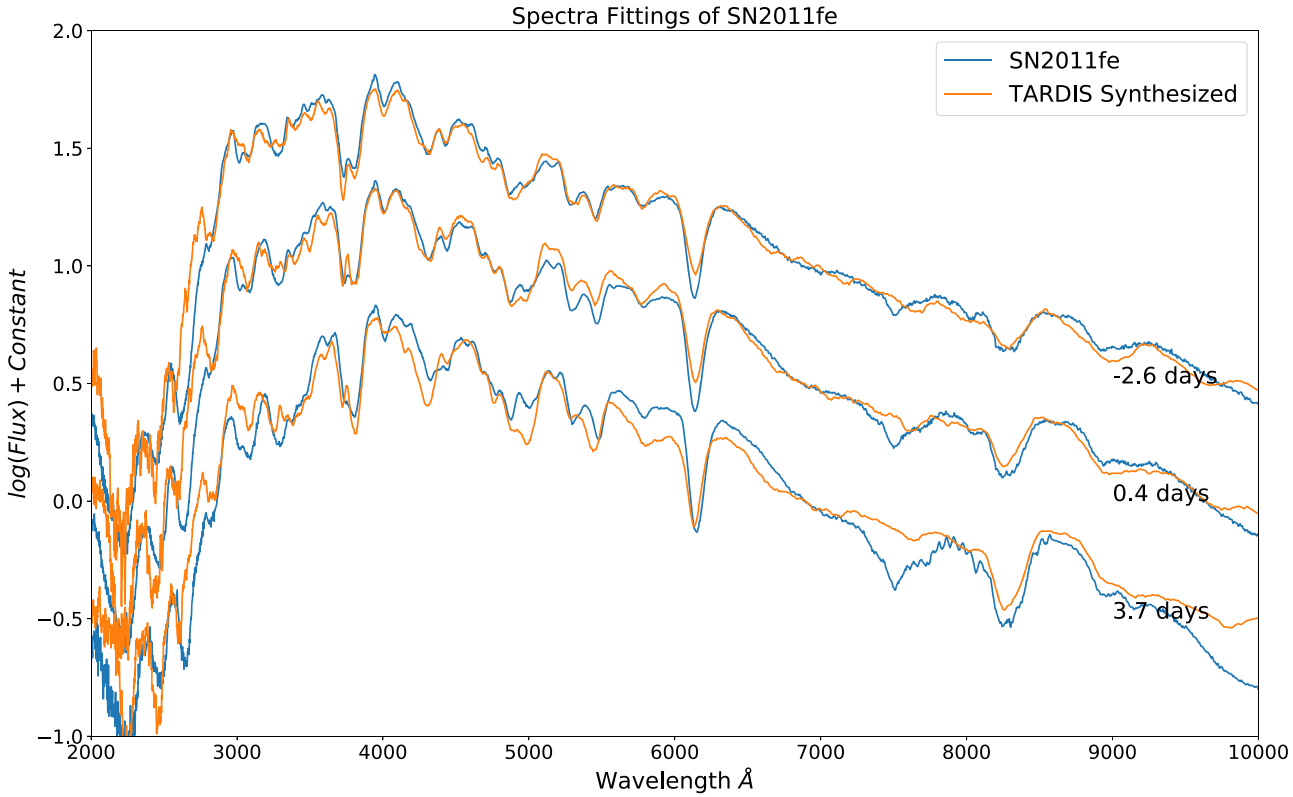

Figure 7. Observed spectra (blue lines) and TARDIS synthesized spectra (orange lines) of SN 2011fe at −2.6 (upper), 0.4 (middle), and 3.1 days (lower). The TARDIS synthetic spectra were filtered by a Savitzky–Golay filter with window = 9, order = 1 to remove Monte Carlo noise. The spectra are in logarithmic scale and arbitrarily shifted in the vertical axis for the different dates.

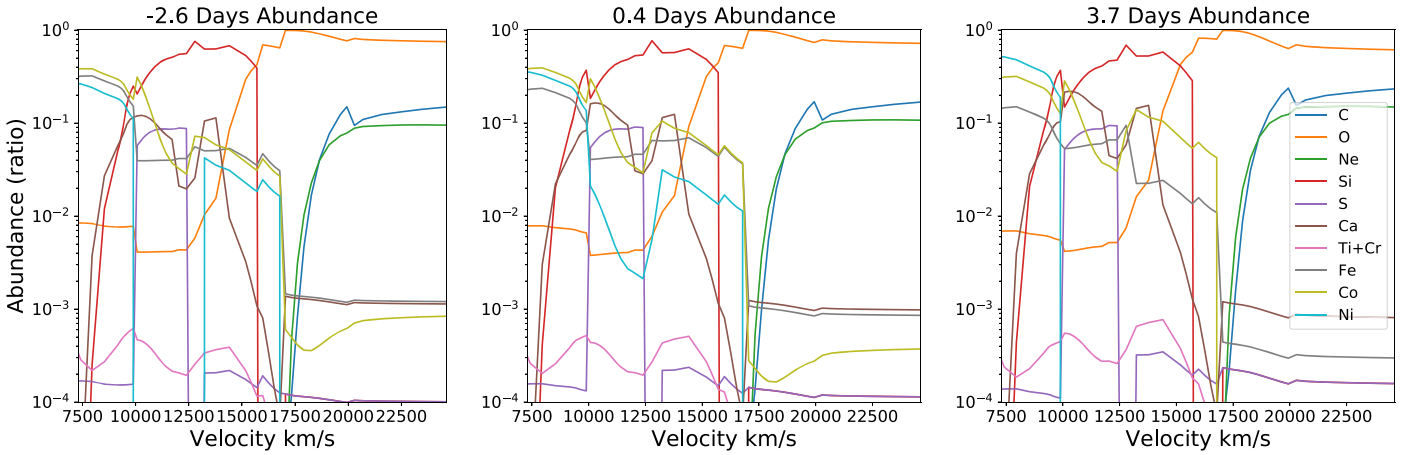

Figure 8. Elemental abundances predicted by the neural networks from SN 2011fe at −2.6 (left panel), 0.4 (middle panel), and 3.7 (right panel) days. These elemental abundances and the densities in Figure 9 were used for the synthetic spectra shown in Figure 7.

The temperature profiles of the TARDIS models of SN 2011fe in the three spectroscopically observed phases are shown in Figure 9. The three SN structures shown in Figure 8 are predicted by the neural network individually for each epoch, and the density structures may not strictly observe homologous expansion. However, it is encouraging to see that the density structures in Figure 9 (first panel) show adequate similarity with each other, and the corrections compared to the DDT model appear to be consistent at different velocity layers; this cross-validates the prediction of the neural networks.

4.1.2. SN 2011iv

The transitional SN (between Type Ia normal and Type Ia 91bg-like) SN 2011iv is located at NGC 1404 with a *B*-band decline rate $\Delta M_{B,15} = 1.69 \pm 0.05$ (Gall et al. 2018). According to Foley et al. (2012), this SN has a negligible dust extinction effect in the line of sight, so we did not apply extra extinction corrections to it (see Table 4).

In the elemental prediction and spectral fitting process, we adopted the combined spectra of HST and the Magellan

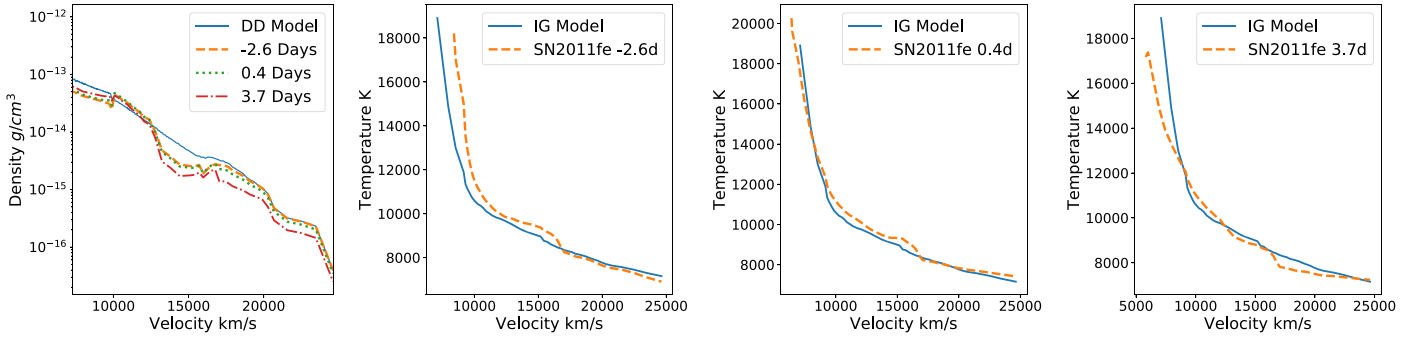

Figure 9. First panel: density structures of SN 2011fe at −2.6 (orange line), 0.4 (cyan line), and 3.7 (red line) days as predicted by the neural networks. The DDT model is shown with a blue line for comparison. Second panel: temperature structure for SN 2011fe at −2.6 days. Third panel: temperature structure for SN 2011fe at 0.4 day. Fourth panel: temperature structure for SN 2011fe at 3.7 days. Note that for convenience of comparison, we have converted all density profiles to 19 days after explosion, assuming homogeneous expansion, using the relation $\rho \propto t^{-3}$. The temperature profiles for the IGM are shown as blue lines for comparison, and MRNN fits are shown as orange lines.

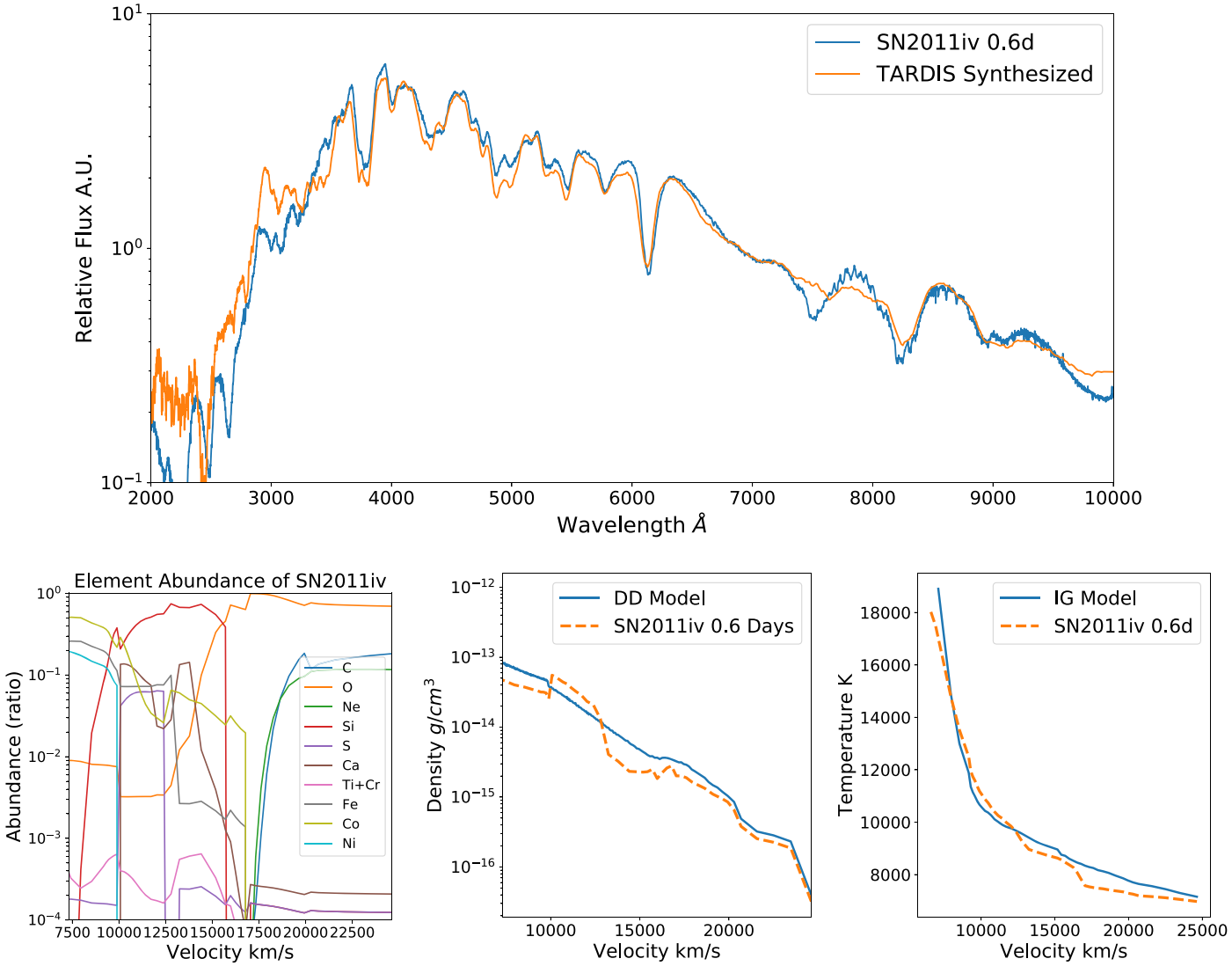

Figure 10. Upper panel: observed spectrum (blue line) and TARDIS synthetic spectrum (orange line) of SN 2011iv at 0.6 day after *B* maximum. Lower left: elemental abundances of SN 2011iv predicted from neural networks. Lower middle: density structure of SN 2011iv predicted from neural networks (orange line) and the DDT model density structure for comparison (blue line). Both densities are converted to that of day 19 using the $\rho \propto t^{-3}$ relation. Lower right: temperature structure for SN2011iv spectral fitting (orange line) and the IGM temperature structure for comparison (blue line).

telescope at 0.6 day after the *B*-band maximum date (Foley et al. 2012). The results are shown in Figure 10.

The model spectrum agrees reasonably well with the observed spectrum across major spectral features in the optical. The disagreement across the O I 7300 Å line is obvious. This is likely due to an insufficient amount of oxygen, and the feature is not well fit even when the oxygen abundance is enhanced to three times that of the IGM. A similar problem may also be

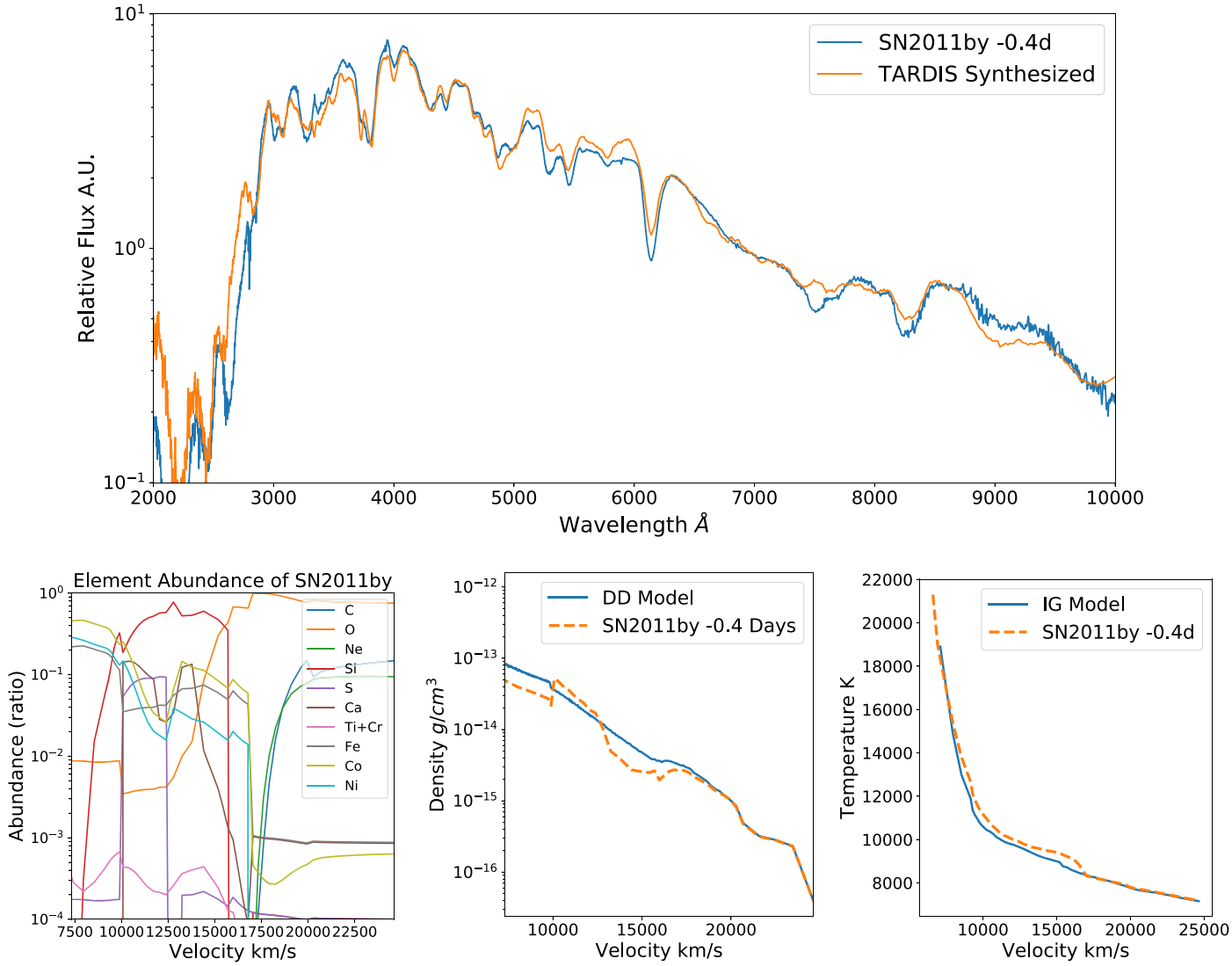

Figure 11. Upper panel: extinction-corrected observed spectra (blue line) and TARDIS synthesized spectra (orange line) of SN2011by at −0.4 day. Lower left: element abundance of SN2011by predicted from neural networks. Lower middle: density structure of SN2011by predicted from neural networks (orange line) and the DDT model density structure as a comparison. Both densities are converted to 19 days after explosion using the $\rho \propto t^{-3}$ relation. Lower right: comparison of the temperature structure of the SN2011by spectral model (orange line) and that of the IGM (blue line).

seen in Figure 7 for SN 2011fe. We see also that the spectral features below 3000 Å are poorly fit, suggesting again that the elemental structure may not have appropriately covered this transition SN Ia. We will improve these fits in future studies with the construction of a larger spectral library.

4.1.3. SN 2011by

The SN 2011by in NGC 3972 has a luminosity decline rate $\Delta M_{B,15} = 1.14 \pm 0.03$ (Silverman et al. 2013). It has remarkably similar optical spectra and light curves to those of SN 2011fe, and these two SNe are identified as optical "twins" (Graham et al. 2015). The only prominent difference is that SN 2011fe is significantly more luminous in the UV (1600 Å $< \lambda <$ 2500 Å) than SN 2011by before and around peak brightness (Foley et al. 2020). However, based on the distance deduced from Cepheid variables in NGC 3972, SN 2011by is about 0.335 ± 0.069 mag dimmer than SN 2011fe (Foley et al. 2020). This apparent magnitude difference can be a concern for SN cosmology, as its origins are unknown and thus difficult to correct.

As in Foley et al. (2020), we adopted the Fitzpatrick (1999) extinction model with $R_v = 3.1$ and $E(B-V) = 0.039$ to correct the host galaxy extinction and Milky Way extinction models of Gordon et al. (2009) with $E(B-V) = 0.013$ to correct for Milky Way extinction. The HST spectrum of SN 2011by at 0.4 day before the B-band maximum is employed to derive the elemental abundances using MRNN. The extinction-corrected observed spectrum and the corresponding synthetic spectrum are shown in Figure 11.

It is worth noticing that the absolute luminosity of SN 2011by found by the MRNN process is slightly more luminous than that of SN 2011fe. Our MRNN spectral fits thus do not provide a theoretical explanation of the apparent luminosity difference of the two SNe. This may be caused by model uncertainties and the fact that we adopted an extinction correction without further iterations to improve the overall

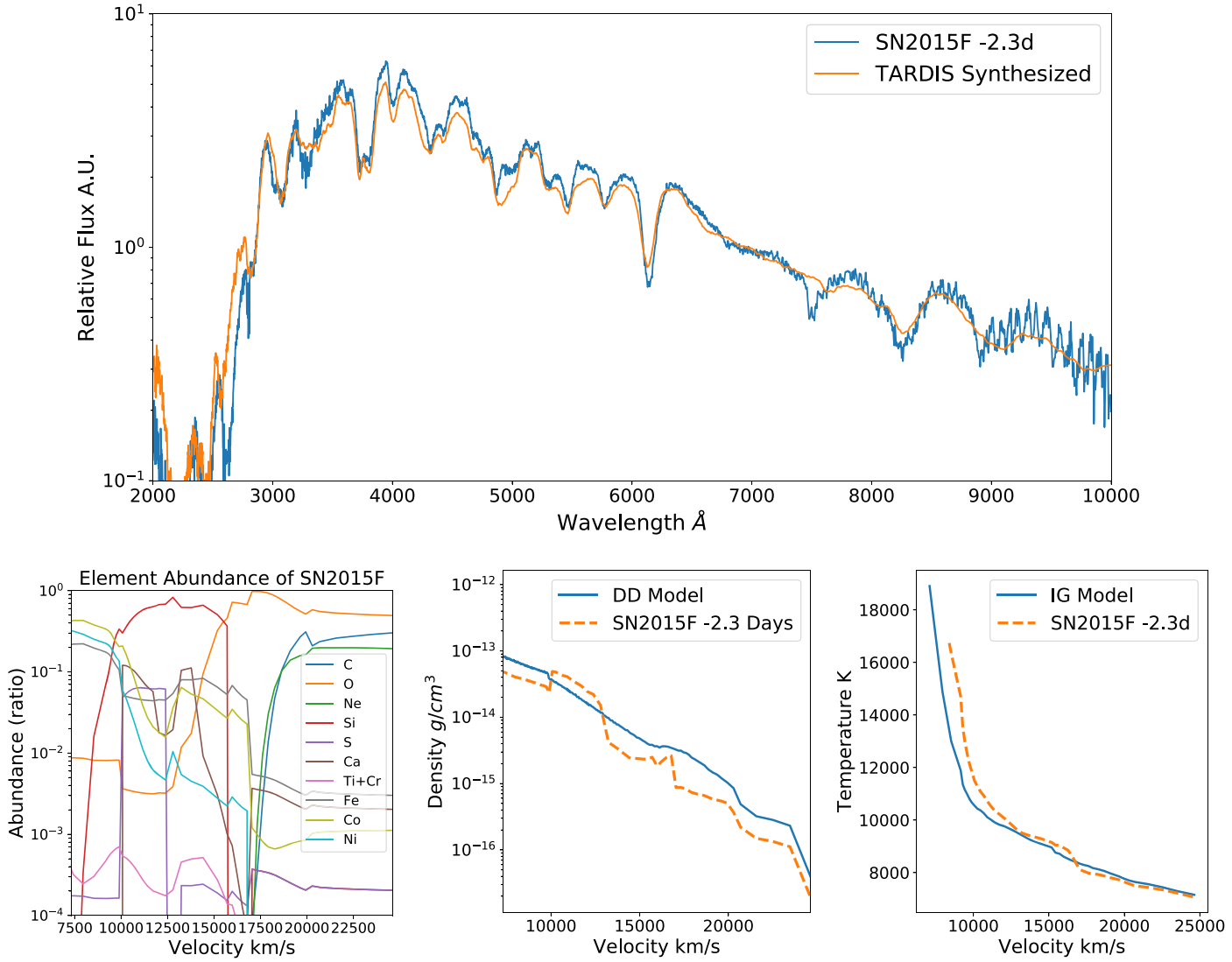

Figure 12. Upper panel: extinction-corrected observed spectra (blue line) and TARDIS synthesized spectra (orange line) of SN 2015F at −2.6 days. Lower left: elemental abundances of SN 2015F predicted by MRNN. Lower middle: density structure of SN 2015F predicted from MRNN (orange line) and the DDT model density structure for comparison. Lower right: temperature structures for the SN 2015F spectral model (orange line) and the IGM for comparison (blue line).

model fits in the UV (see Figure 11). Indeed, the model spectrum is too bright in the wavelength range shorter than 2800 Å, which may again suggest that there is room for improvement by enlarging the ranges of the elemental abundances for the spectral library to better sample the physical conditions of the observed spectrum.

4.1.4. SN 2015F

The SN 2015F in NGC 2442 is a slightly subluminous SN with a decline rate $\Delta M_{B,15} = 1.35 \pm 0.03$ (Cartier et al. 2016). We employed the HST spectrum at −2.3 days relative to B-band maximum for elemental abundance predictions and spectral fitting. The host galaxy extinction was corrected using the model from Cardelli et al. (1989) with $R_v = 3.1$ and $E(B-V) = 0.035$ mag. The Milky Way extinction was corrected with the Gordon et al. (2009) model with $E(B-V) = 0.175$ (Graur et al. 2018b). The fitting results are shown in Figure 12. Notice the absence of high-velocity Ca II and O I and the apparent lower density of the ejecta, as is obvious from the lower middle panel of Figure 12.

4.1.5. ASASSN-14lp

Located in NGC 4666, ASASSN-14lp is a bright SN Ia. Its luminosity decline is $\Delta M_{B,15} = 0.80 \pm 0.05$ (Shappee et al. 2016). In order to correct the host galaxy extinction, we adopted the Cardelli et al. (1989) extinction relation with $R_v = 3.1$ and $E(B-V) = 0.33$ mag (Shappee et al. 2016). For Milky Way extinction, we adopted the Gordon et al. (2009) extinction model with $E(B-V) = 0.021$ mag.

We used the HST spectrum at −4.4 days from B-band maximum for both the elemental abundance prediction and spectral fitting. The results are shown in Figure 13.

The ejecta show enhanced density at a velocity above 17,500 km s^{-1}. The Ca II H and K and IR triplet are clearly detected and highly blueshifted. The temperature profile shown in the lower right panel of Figure 13 is higher than that of the

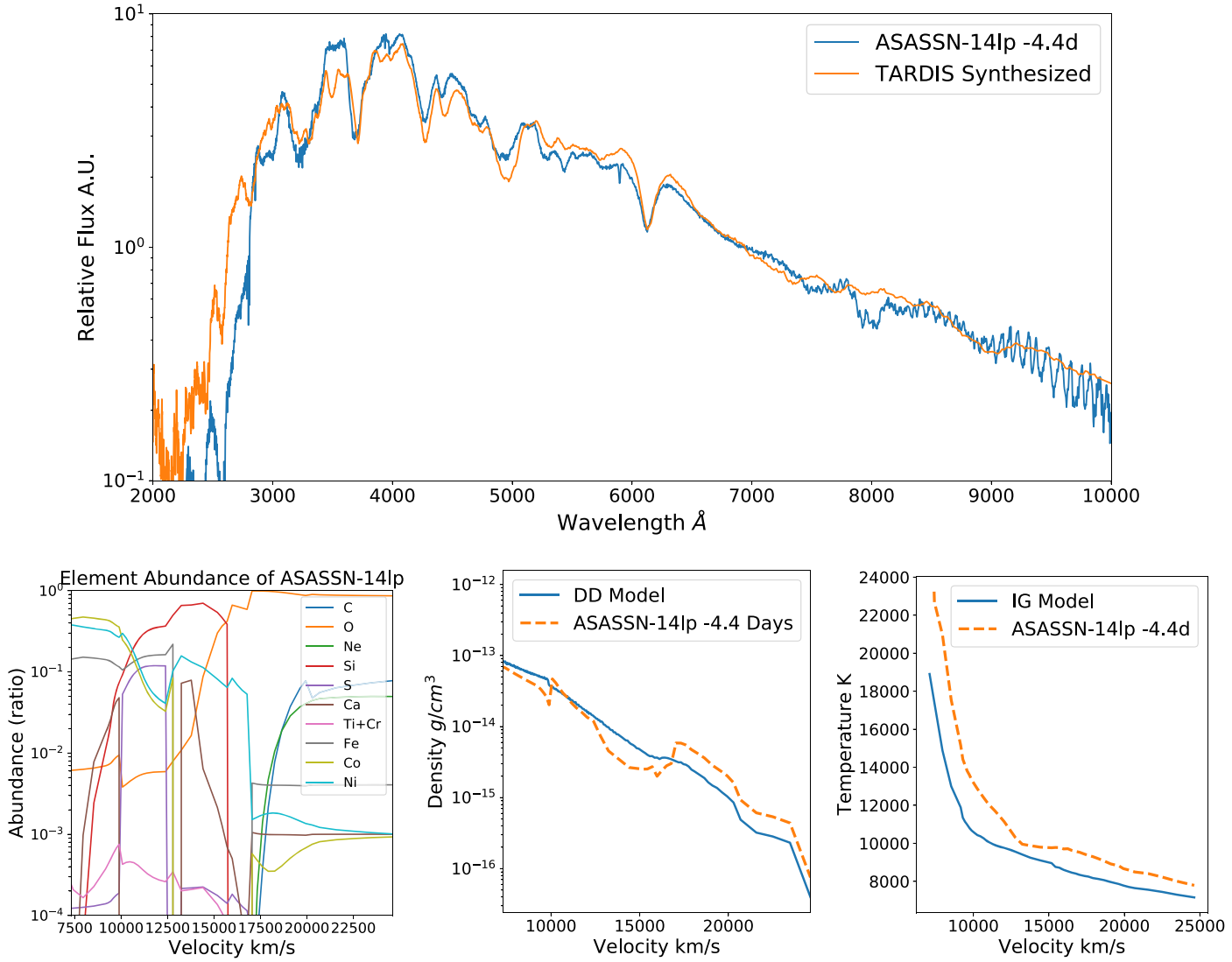

Figure 13. Upper panel: extinction-corrected observed spectrum (blue line) and TARDIS synthesized spectrum (orange line) of ASASSN-14lp at −4.4 days. Lower left: elemental abundances of ASASSN-14lp predicted by MRNN. Lower middle: density structure of ASASSN-14lp predicted by MRNN (orange line) and the DDT model density structure for comparison. Both densities are converted to 19 days from B maximum using the $\rho \propto t^{-3}$ relation. Lower right: temperature structure for ASASSN-14lp (orange line) and the IGM temperature structure for comparison (blue line).

IGM throughout the ejecta. This is consistent with what is typically expected for SNe with slow decline rates.

4.1.6. SN 2013dy

The SN 2013dy is located in NGC 7250 with a luminosity decline rate of $\Delta M_{B,15} = 0.90 \pm 0.03$ (Zhai et al. 2016). Like ASASSN-14lp, SN 2013dy is of the group with slow decline rates. We adopted the extinction model of Cardelli et al. (1989) with $R_v = 3.1$ and $E(B-V) = 0.206$ mag to correct the host galaxy extinction and the Gordon et al. (2009) model with $E(B-V) = 0.135$ mag for Milky Way reddening correction. The extinction parameters $E(B-V)$ were taken from Vinkó et al. (2018).

The HST spectra at −3.1, −1.1, 0.9, and 3.9 days were used for spectral modeling. The fitting spectra and the observed spectra are shown in Figure 14, and the ejecta structures and temperature profiles are shown in Figure 15.

The four ejecta structures for SN 2013dy show consistently higher densities at velocities above 17,500 km s^{-1} and a higher temperature before optical maximum than that of the IGM. The O I feature at 7500 Å is again poorly fit. It implies that the oxygen abundances are too low for the spectra in the spectral library.

4.2. Application to SNe with Wavelength Coverage 3000–5200 Å

4.2.1. HST Spectra of 15 PTF Targets

Maguire et al. (2012) presented 32 low-redshift ($0.001 < z < 0.08$) SN Ia spectra; the UV spectra were obtained with the HST using the STIS. The spectral wavelength coverage of these data is 3000–5200 Å. Meanwhile, the photometric data are obtained in the Palomar Transient Factory (PTF; Law et al. 2009; Rau et al. 2009). There are no published light curves on these targets yet. As our models are built only for

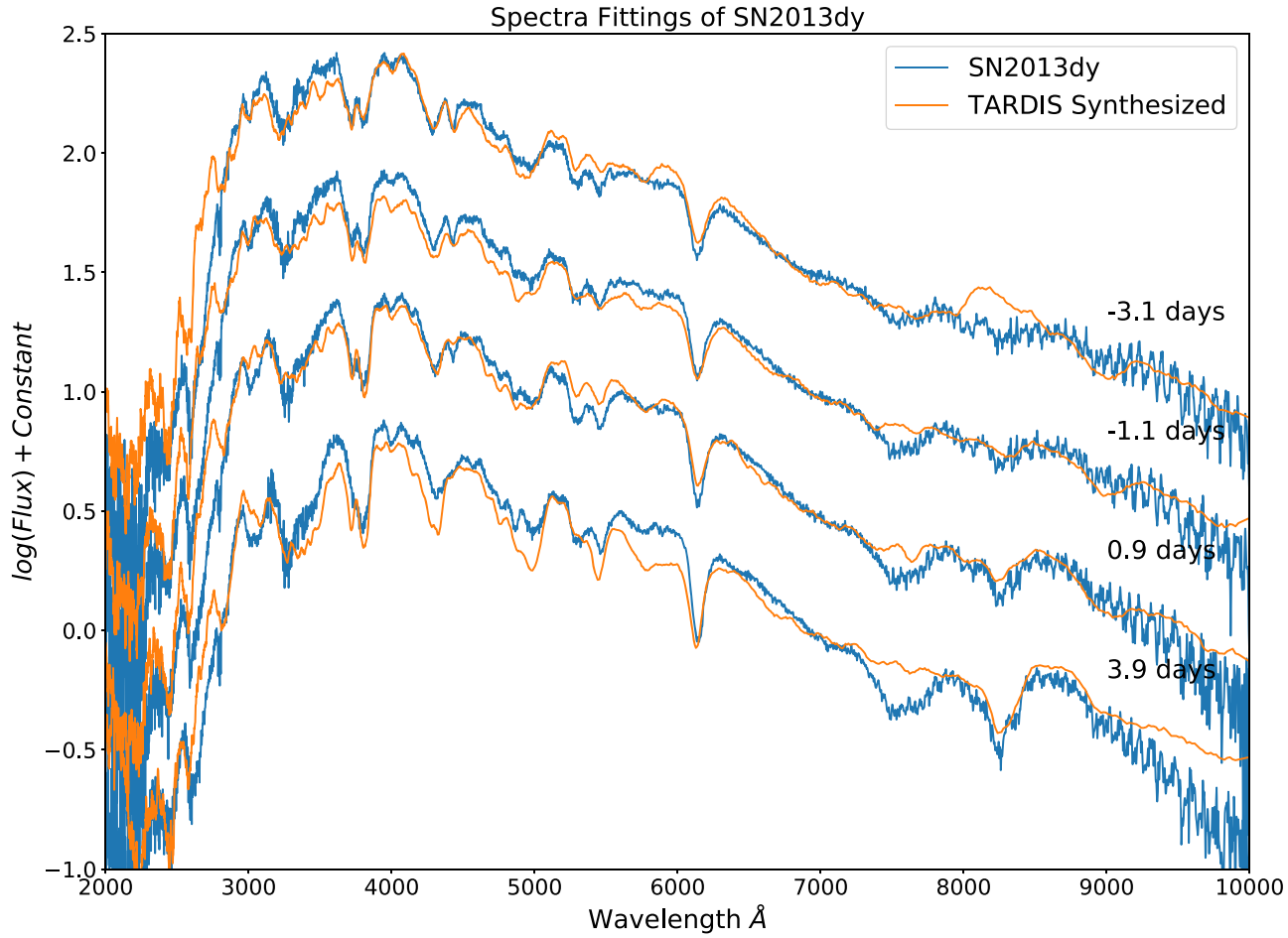

Figure 14. Extinction-corrected observed spectra (blue lines) and TARDIS synthetic spectra (orange lines) of SN 2013dy at −3.1, −1.1, 0.9, and 3.9 days. The fluxes are in logarithmic scale, and the spectra of different dates are arbitrarily offset by a constant.

data close to optical maximum, we selected only spectra taken between −3 and 4 days relative to the date of *B*-band maximum.

4.2.2. Dust Extinction Correction

To model the observed SN luminosity, we need to properly treat the dust reddening by the Milky Way and the host galaxies.

The extinction data of the 15 HST spectra in Maguire et al. (2012) are not available. In order to correct the reddening effect, we developed an iterative algorithm by searching for the best match of theoretical models and observations with reddening as one of the χ^2 minimization parameters, as is shown in the flowchart in Figure 16. The procedure starts with an initial guess of reddening index $E(B-V)$, which is applied to the observed spectrum; it then calculates the rms of the difference between the observed spectrum and all of the model spectra in the spectral library to find the best-match spectral model, a new reddening index $E(B-V)_{\rm newer}$ is calculated assuming the best-match spectral model is the true unreddened spectrum, the reddening index is updated using the formula $E(B-V) = 0.1 \times E(B-V) + E(B-V)_{\rm newer} \times 0.9$, and the procedure repeats for 20 iterations to ensure convergence of the algorithm.

The algorithm was found to be sensitive to the initial value, so we set the initial $E(B-V)$ to be 0, 0.2, 0.4, and 0.6, respectively. With different extinction values and the resulting best-fit spectra, we compare the spectral fitting fidelity and adopt the $E(B-V)$ value and spectral model with the least rms of the difference between the observed and model spectra reddened by the $E(B-V)$ value.

We also tested our extinction-correction algorithm on the 11 SN spectra with wavelength coverage of 2000–10000 Å and compared the derived extinction values with the extinction values in published literature. The results are shown in Table 5. We found that our algorithm can reproduce low and intermediate extinctions, while for the high-extinction case $E(B-V) \sim 0.3$, the algorithm seems to underestimate the extinction intensity. As it is difficult to assess the precision of the flux calibration of the observed spectra, we consider the values shown in Table 5 to be broadly in agreement.

In Table 6, we list the extinction values derived from this procedure for the 15 SNe under study.

The AIAI results, showing the spectral profiles, density structures, chemical structures, and temperature profiles, are shown in Appendix C.

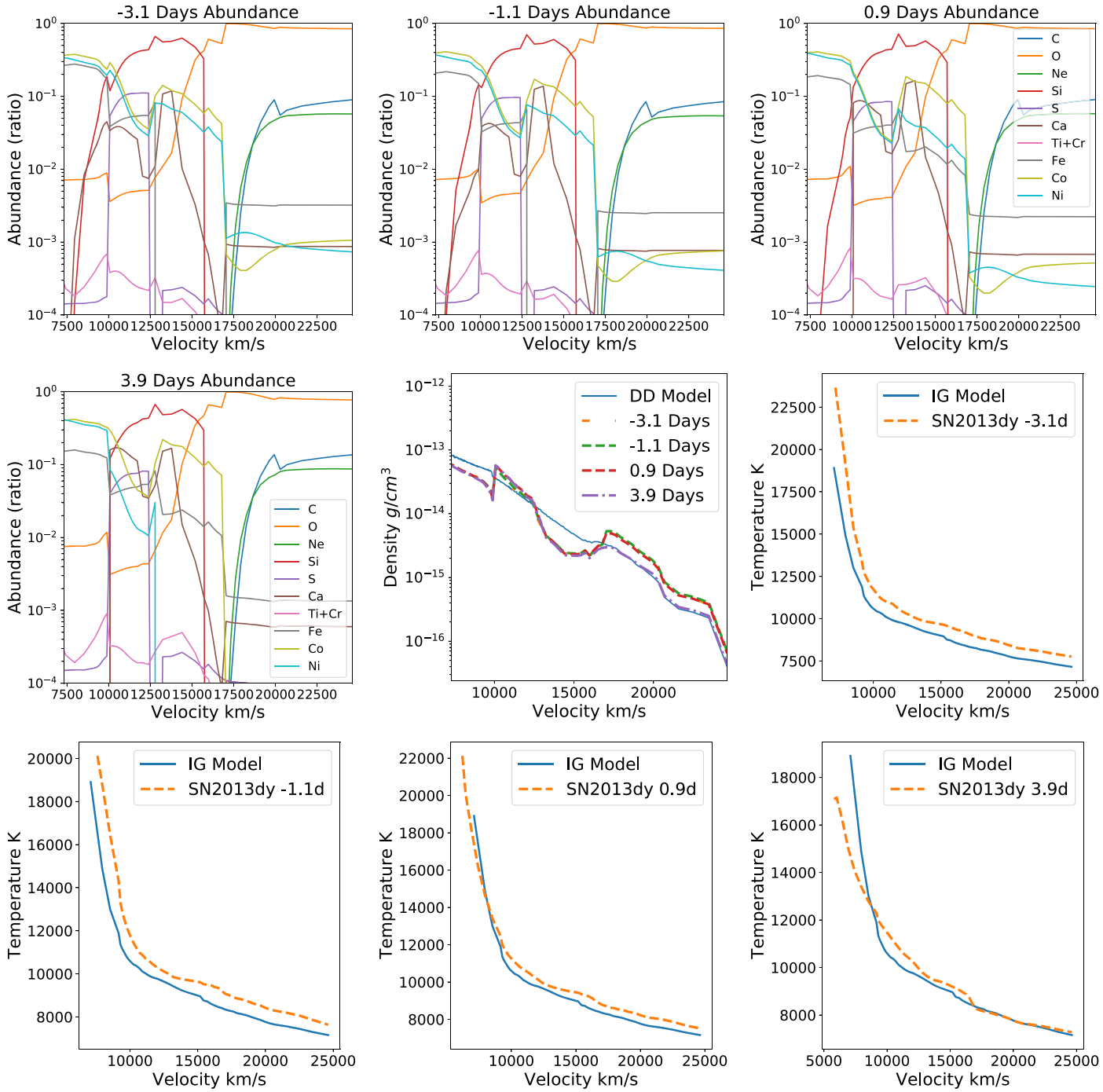

Figure 15. Top left, top center, top right, and middle left: elemental abundance structure of SN 2013dy predicted by neural networks from the spectra at −3.1, −1.1, 0.9, and 3.9 days after *B* maximum. Middle center: density structure of SN2013dy predicted from the spectra of −3.1, −1.1, 0.9, and 3.9 days. The density profiles are converted to the profile of day 19 using the $\rho \propto t^{-3}$ relation. Middle right, lower left, lower center, and lower right: temperature structure for SN2013dy spectra of day −3.1, −1.1, 0.9, and 3.9, respectively.

5. Discussions and Conclusions

5.1. Elemental Abundances and the Light-curve Stretch

The stretch values are known to be correlated to the SN luminosity (Phillips 1993; Perlmutter et al. 1999). To compare the elemental abundance derived from the best-match TARDIS spectra with light-curve stretch parameters, we employed the `SiFTO` (Conley et al. 2008) program to calculate the stretch factors. The resulting stretch values are listed in Table 4.

We may consider the theoretical models as a toolbox to construct empirical parameters to improve the precision of SNe Ia as standard candles, in a way similar to the light-curve shape parameters. This would be particularly interesting for projects based on spectrophotometry, such as has been planned for WFIRST. A few more clarifications of the uncertainties of the derived chemical abundances are necessary before carrying out such a study. Based on the testing data set discussed in Section 3, we can estimate the $1-\sigma$ limits of the elemental abundances. Not all of the predictions from the neural network

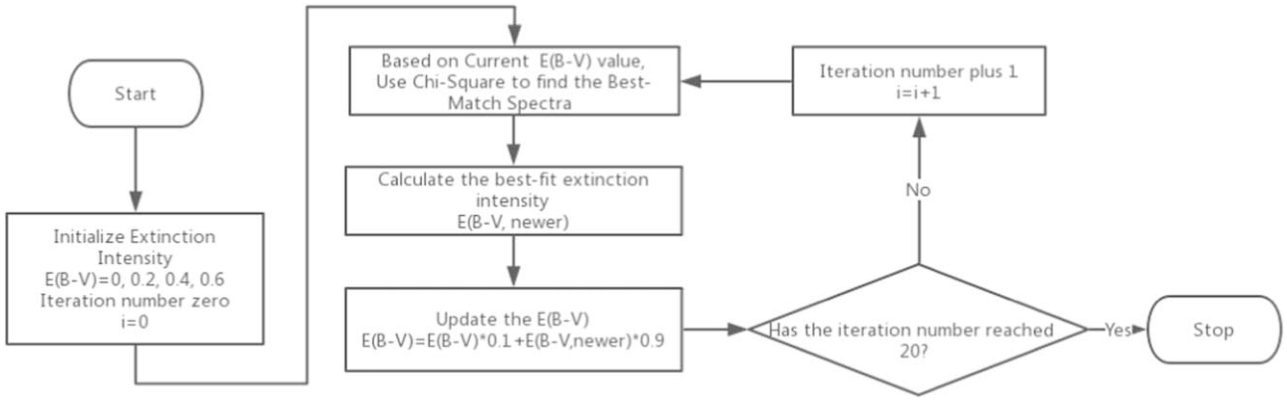

Figure 16. Illustrative flowchart of the algorithm for dust reddening estimates (see Section 4.2.2 for details).

Table 5
The Extinction Comparison

SN Name	Phase (days)	$E(B-V)$ from Papers	$E(B-V)$ from Algorithm
SN 2011fe	0.4	0.000	0.021
SN 2011fe	−2.6	0.000	0.000
SN 2011fe	3.7	0.000	0.000
SN 2013dy	−3.1	0.341	0.222
SN 2013dy	−1.1	0.341	0.253
SN 2013dy	0.9	0.341	0.231
SN 2013dy	3.9	0.341	0.253
SN 2011iv	0.6	0.000	0.092
SN 2015F	−2.3	0.210	0.129
ASASSN-14lp	−4.4	0.351	0.281
SN 2011by	−0.4	0.052	0.018

Table 6
The HST-observed SNe Ia from Maguire et al. (2012)

SN Name	Phase (days)	Redshift[a]	Stretch	$E(B-V)$
PTF 09dlc	2.8	0.068	1.05 ± 0.03	0.21
PTF 09dnl	1.3	0.019	1.05 ± 0.02	0.04
PTF 09fox	2.6	0.0718	0.92 ± 0.04	0.08
PTF 09foz	2.8	0.05	0.87 ± 0.06	0.15
PTF 10bjs	1.9	0.0296	1.08 ± 0.02	0.02
PTF 10hdv	3.3	0.054	1.05 ± 0.07	0.16
PTF 10hmv	2.5	0.032	1.09 ± 0.01	0.02
PTF 10icb	0.8	0.086	0.99 ± 0.03	0.13
PTF 10mwb	−0.4	0.03	0.94 ± 0.03	0.02
PTF 10pdf	2.2	0.0757	1.23 ± 0.03	0.23
PTF 10qjq	3.5	0.0289	0.96 ± 0.02	0.14
PTF 10tce	3.5	0.041	1.07 ± 0.02	0.23
PTF 10ufj	2.7	0.07	0.95 ± 0.02	0.15
PTF 10xyt	3.2	0.049	1.07 ± 0.04	0.24
SN 2009le	0.3	0.017786	1.08 ± 0.01	0.17

Note.
[a] In Maguire et al. (2012), only cosmological redshifts are given; the collected redshift data here are adopted from WISeREP.

are reliable due to the sparsity of our parameter space coverage in generating the data set and the limitations of the sensitivity of the neural network. For example, some of the predictions of Co in zone 3 are not in the region where the testing data set has sufficient coverage, as is shown in the right panel of Figure 5. When this situation occurs, we replace the predictions that go above or below the relevant $1-\sigma$ region with an upper or lower limit, respectively. Also, for simplicity, we have set the lower limits of the ejecta structure to be 7100 km s^{-1} when calculating the abundances of zone 1 to avoid the effect introduced by a varying photospheric velocity.

Taking the elemental masses derived from AIAI at their face values, we show in Figure 17 the correlations between the stretch parameters and Fe, Co, and Ni masses all corrected to the values at 19 days after explosion. The correlations for the remaining 22 elements in various zones are presented in Appendix B.

Although we intend to leave a thorough analysis of these correlations to an upcoming paper, we identify immediately from Figure 17 that the Fe, Co, and Ni masses in zone 4 are more strongly correlated with the stretch parameter than other parameters. The lower panel of Figure 17 also shows that the masses of Ni in zones 2, 3, and 4 are all correlated with the stretch parameter. The mass derived from zone 1 may not be reliable, as the photosphere is located in zone 1. The correlations that can be identified for Fe and Co in zone 4 and Ni in zones 2, 3, and 4 are suggestive of the presence of radioactive materials at the surface of the SN ejecta and are likely critical measures of the luminosities of SNe Ia.

It is noteworthy that the Ni mass in zones 2, 3, and 4 varies by more than 1 order of magnitude. The masses of Fe and Co do not appear to share such a behavior. No strong correlation is found between the mass of Fe and Co and the stretch factor either. This may be explained if a fraction of Fe and Co are nonradioactive and a dominant fraction of Ni in these zones is radioactive. Indeed, 1D models of thermonuclear explosions predict the existence of a high-density electron-capture burning region during the deflagration phase (Hoeflich et al. 1996a; Gerardy et al. 2007), which can lead to the production of a significant amount of nonradioactive Ni and Fe but little Co. These early deflagration products can be mixed out to higher-velocity layers, such as shown in some 3D models (e.g., Gamezo et al. 2004; Plewa et al. 2004; Röpke et al. 2006; Jordan et al. 2008); they are likely to distort and weaken the correlation between masses of Co and Fe and the light-curve shapes of the SNe.

5.2. Time Evolution of Elemental Abundances in SN 2011fe and SN 2013dy

Among all six SNe with spectroscopic observations near B maximum and wavelength coverage over 2000–10000 Å, there are three spectra of SN 2011fe and four spectra of SN 2013dy that are compatible with our neural network setups. In this section, we investigate the time evolution of the elemental abundances of these two SNe. The time evolution of the masses

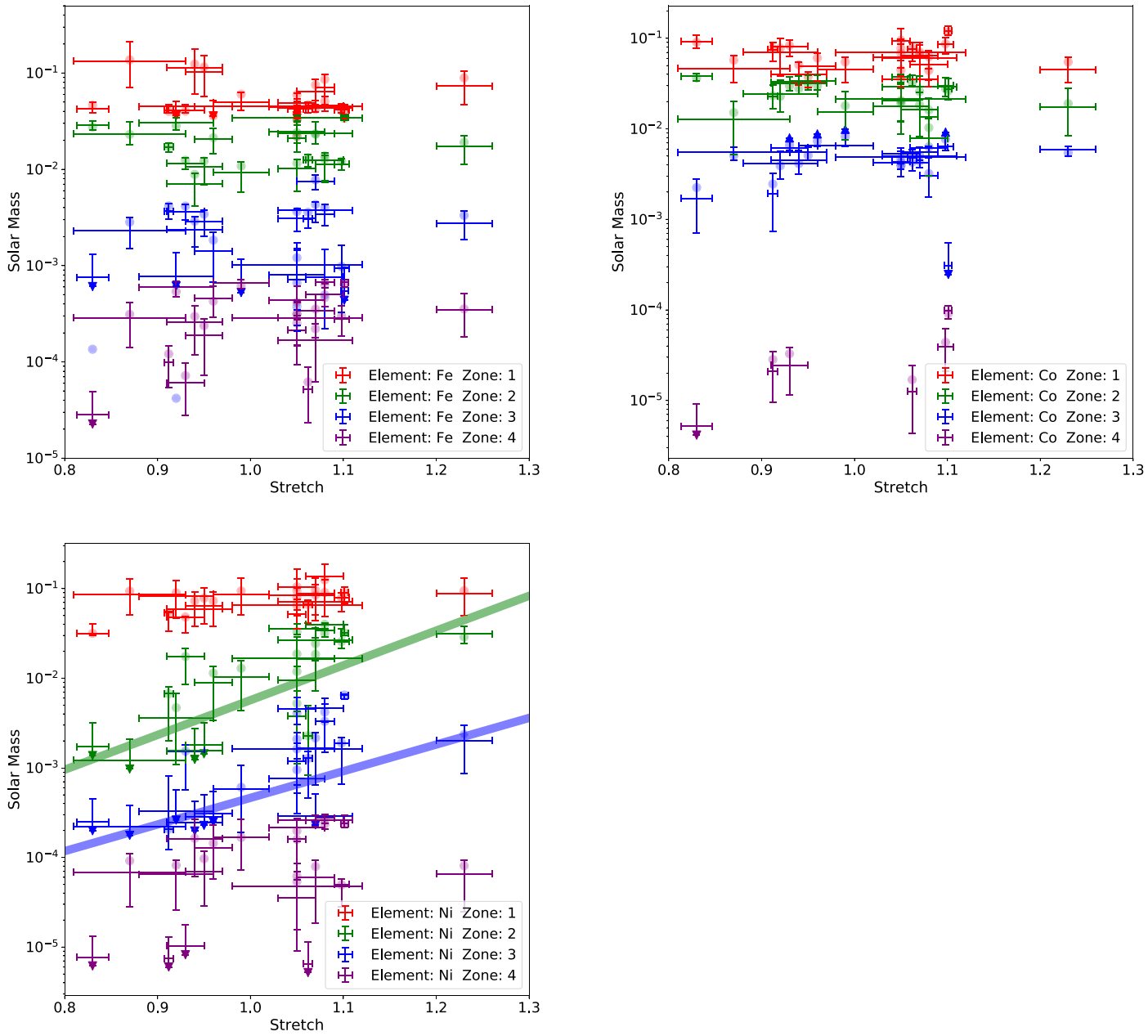

Figure 17. Correlation between the stretch parameter and the mass of various chemical elements. The lower and upper limits are marked with triangles, the median values are the crosses, and the predicted values are the circles. For simplicity, we choose the lower limit of the ejecta structures to be 7100 km s^{-1} when calculating the element masses. Upper left: correlation between the stretch parameter and Fe mass in zones 1, 2, 3, and 4. Upper right: correlation between the stretch parameter and Co mass in zones 1, 2, 3, and 4. Lower panel: correlation between stretch and Ni mass in zones 1, 2, 3, and 4. For the neural network WR-Blue (wavelength range 3000–5200 Å), the Co mass in zone 4 cannot be determined, so we did not predict the Co mass in zone 4 for the 15 spectra with such wavelength coverage. Notice that the Ni mass is corrected by its radioactive decay, and the predictions are for the Ni mass at the *B*-band maximum rather than the date of observations. We fitted the Ni mass in zones 2 and 3 with $M_{\rm Ni,2} = 7.511 \times 10^{-5} e^{8.933\times{\rm Stretch}}$ and $M_{\rm Ni,3} = 4.934 \times 10^{-7} e^{6.843\times{\rm Stretch}}$ $M_\odot$, respectively.

of radioactive materials can serve as an important check on the fidelity of the elemental abundances we deduce from the neural network whose performance is very difficult to track precisely from first-principle mathematical models.

In Figure 18 (top row), we notice a general agreement between the time evolution of the Ni mass in zones 2, 3, and 4 and the radioactive decay rate of ^{56}Ni. This agreement suggests that most of the ^{56}Ni is newly synthesized after the explosion. Here SN 2013dy shows a significantly larger nickel mass than SN 2011fe. A significant difference in Ni mass is found for the two SNe in zones 2, 3, and 4 or at a velocity above 10,000 km s^{-1}. If this difference is true, it may provide an explanation of the difference in the luminosity of the two SNe (Zhai et al. 2016).

Figure 18 (middle row) shows the mass of Co in zones 2, 3, and 4. Again, SN 2013by is more abundant in Co in zones 3 and 4 than SN 2011fe. Notice that the mass of Co peaks at around 20 days past explosion if its time variation is related to the radioactive decays of ^{56}Ni; Figure 18 (left and middle panels of middle row) is in agreement with this. Likewise, more iron is found at the highest velocity (zone 4) in SN 2013by than in SN 2011fe. Iron seems to be less abundant in SN 2013by than in SN 2011fe in zone 3.

Moreover, the predicted Ni masses from the spectra 3–4 days past maximum are significantly lower than earlier epochs. We

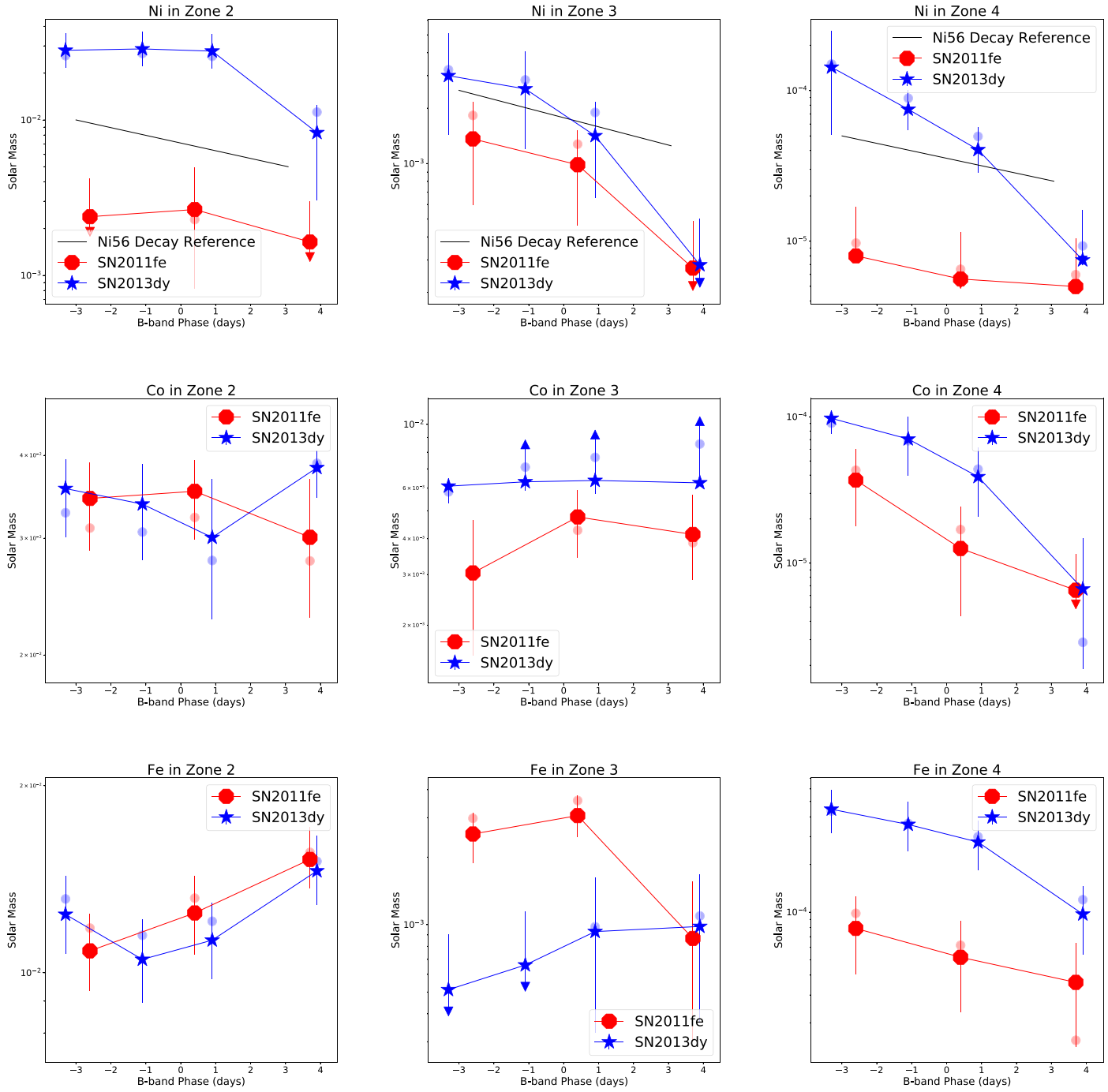

Figure 18. Time evolution of Ni mass in different zones of SN 2011fe and SN 2013dy near *B*-band maximum dates. Upper panels, from left to right: Ni mass in zones 2, 3, and 4. The decay rate of Ni_{56} is plotted with black lines for comparison. Middle panels, from left to right: Co mass in zones 2, 3, and 4. Lower panels, from left to right: Fe mass in zones 2, 3, and 4.

surmise that such an anomaly could be due to a temperature change that strongly affects the UV spectral features. When Fe group elements are in low-temperature regions, like zones 3 or 4, the relevant atomic levels are less activated, which results in weaker or absent spectral lines in the TARDIS model spectra due to Monte Carlo noise, which may misguide the neural network.

Curiously enough, the nebular spectra of SN 2013by show a smaller Ni II λ7378/Fe II λ7155 flux ratio than that of SN 2011fe (Pan et al. 2015), which may suggest that SN 2013dy produced a lower mass ratio of stable to radioactive iron group elements than SN 2011fe. The decay of ^{56}Ni leads to a comparative overabundance of Fe in SN 2013by as compared to that in SN 2011fe. This is in agreement with what we found in zones 2, 3, and 4, although the nebular lines measure ejecta at much lower velocities. The primary difference between SN 2011fe and SN 2013by that affects their early UV spectra and likely also their luminosities is the amount of radioactive Ni at velocities above $\sim$10,000 km s^{-1}.

Zhai et al. (2016) noticed that SN 2013dy is dimmer in the near-IR than SN 2011fe. This could also be suggestive of a difference in the chemical structures of the two. It may be related to the excessive amount of radioactive material at the outer layers of SN 2013dy. The difference in Fe abundance at the highest velocity may also be an indication of a genuine difference of the chemical

Table 7
B-band Absolute Luminosity Comparisons

SN Name	Phase	Obs. Mag.	$E(B-V)$	μ	MRNN Mag.	Abs. Mag.	Difference
SN 2011fe	−2.6	10.034	0	29.13	$-19.96^{+0.12}_{-0.12}$	−19.10	0.86
SN 2011fe	0.4	9.976	0	29.13	$-20.00^{+0.11}_{-0.11}$	−19.15	0.85
SN 2011fe	3.7	10.070	0	29.13	$-19.96^{+0.11}_{-0.12}$	−18.96	0.90
SN 2013dy	−3.1	13.353	0.341	31.50	$-20.06^{+0.08}_{-0.09}$	−19.54	0.52
SN 2013dy	−1.1	13.294	0.341	31.50	$-20.08^{+0.10}_{-0.09}$	−19.60	0.48
SN 2013dy	0.9	13.291	0.341	31.50	$-20.02^{+0.10}_{-0.09}$	−19.61	0.41
SN 2013dy	3.9	13.366	0.341	31.50	$-19.99^{+0.11}_{-0.11}$	−19.53	0.46
SN 2011by	−0.4	12.933	0.052	31.59	$-20.03^{+0.09}_{-0.09}$	−18.87	1.16
SN 2011iv	0.6	12.484	0	31.26	$-19.87^{+0.10}_{-0.09}$	−18.82	1.05
SN 2015F	−2.3	13.590	0.210	31.89	$-19.89^{+0.11}_{-0.10}$	−19.16	0.73
ASASSN-14lp	−4.4	12.496	0.351	30.84	$-20.04^{+0.09}_{-0.09}$	−19.74	0.30

Note. The column names are as follows. SN Name: name of the SN. Phase: days relative to *B*-band maximum time. Obs. Mag.: observed magnitude, interpolated from the photometry using `SNooPy`. $E(B-V)$: total extinction parameters, including both Milky Way and host galaxy extinction. μ: distance modulus. Data sources are listed in Table 4. MRNN Mag.: absolute magnitude predicted by our MRNN. We adopt the median values and $1-\sigma$ intervals in the testing data set. Abs. Mag.: absolute magnitude calculated from the observational magnitude. Difference: difference between MRNN predicted absolute magnitude and the absolute magnitude deduced from observed light curves.

Table 8
B-band Absolute Luminosity Comparisons of 15 SNe from Maguire et al. (2012)

SN Name	Phase	Stretch	MRNN Luminosity
PTF 09dlc	2.8	1.05 ± 0.03	$-19.80^{+0.12}_{-0.08}$
PTF 09dnl	1.3	1.05 ± 0.02	$-19.91^{+0.11}_{-0.1}$
PTF 09fox	2.6	0.92 ± 0.04	$-19.94^{+0.11}_{-0.11}$
PTF 09foz	2.8	0.87 ± 0.06	$-19.78^{+0.13}_{-0.08}$
PTF 10bjs	1.9	1.08 ± 0.02	$-19.89^{+0.11}_{-0.1}$
PTF 10hdv	3.3	1.05 ± 0.07	$-19.85^{+0.1}_{-0.09}$
PTF 10hmv	2.5	1.09 ± 0.01	$-19.80^{+0.12}_{-0.08}$
PTF 10icb	0.8	0.99 ± 0.03	$-19.93^{+0.12}_{-0.11}$
PTF 10mwb	−0.4	0.94 ± 0.03	$-19.87^{+0.11}_{-0.09}$
PTF 10pdf	2.2	1.23 ± 0.03	$-19.73^{+0.15}_{-0.08}$
PTF 10qjq	3.5	0.96 ± 0.02	$-19.92^{+0.11}_{-0.1}$
PTF 10tce	3.5	1.07 ± 0.02	$-19.67^{+0.13}_{-0.11}$
PTF 10ufj	2.7	0.95 ± 0.02	$-19.75^{+0.14}_{-0.08}$
PTF 10xyt	3.2	1.07 ± 0.04	$-19.83^{+0.11}_{-0.09}$
SN 2009le	0.3	1.08 ± 0.01	$-19.84^{+0.1}_{-0.09}$

Note. The observed luminosities of these SNe are not available due to the lack of published photometric data.

abundance of the progenitors of the two SNe if they are primordial to the progenitor. Alternatively, they may also suggest different explosion mechanisms if they are produced during the explosion. Note further that SN 2013dy was discovered to show very strong unburned C II lines before maximum, in contrast to the weak C II features of SN 2011fe. It sits on the border of the "normal velocity" SNe Ia and 91T/99aa-like events (Zhai et al. 2016), while SN 2011fe is no normal SN Ia of normal velocity (Wang et al. 2009, 2013). Mechanisms such as double detonation may enrich the high-velocity ejecta with Fe (Wang & Han 2012).

In Appendix D, we show the time evolution of other chemical elements derived from the MRNN.

5.3. The Absolute Luminosity

The neural network with the same structure as in Section 3 may be constructed to retrieve the luminosity assumed in the synthetic spectral models. This opens the possibility of predicting the absolute luminosity of an SN Ia based on its spectral shapes. This can be achieved by training the AIAI with predicting power on the absolute magnitude of any filter bands, here chosen to be the Bessel *B* band.

Unlike elemental abundance predictors, the training on *B*-band luminosities is prone to systematic errors due to TARDIS model limitations. The luminosity is sensitive to both the location of and the temperature at the photosphere. In reality, the location of the photosphere is sensitive to wavelength, whereas TARDIS treats the photosphere as wavelength-independent. The temperature is sensitive to the UV fluxes, and the location of the photosphere is sensitive to the absorption minimum of weak lines. At around optical maximum, the photosphere has receded into the iron-rich layers for normal SNe Ia, and the absorption minima of IMEs are no longer good indicators of the location of the photosphere. Such insensitivity may introduce degeneracy in the dependence of luminosity and spectral profiles. The luminosity of two SNe may be very different, whereas the spectral profiles in the optical may appear to be very similar.

Nonetheless, one may try to study the relations between the luminosities and spectral shapes in a similar way to what has been done for elemental abundances. However, when the neural networks are trained with multiple iterations, the MSE on the training set decreases while the MSE on the testing set increases. This indicates that the neural network overfits features in the training set, and the results cannot be used for model predictions. Consequently, we curtailed the performance tolerance to one iteration only and adopted a smaller learning rate, which is one-tenth of the value used for training the elemental abundances (3×10^{-6} in the first stage and 3×10^{-8} in the second stage).

Having done the training and testing, we insert the observed spectra into the trained neural network. The luminosities of the 11 spectra of the six SNe with HST spectra were predicted by the neural network; the results are listed in Table 7. We also use the neural network to predict the absolute magnitude of the other 15 SNe that are from Magurie et al. (2012); the results are listed in Table 8.

We have also estimated the *B*-band maximum luminosities of the six SNe using their light curves with SNooPy

(Burns et al. 2010). They are converted to absolute magnitude as shown in Table 7.

The *B*-band magnitudes from the neural network are systematically more luminous than the values deduced from well-calibrated light curves. This difference is perhaps the result of the simplified assumption of a sharp photosphere, as mentioned above. Despite this limitation, the neural network prediction is largely a measurement of the spectral properties and may be used as an empirical indicator of luminosity, which can still be useful in exploring the diversity of SN Ia luminosity based on spectroscopic data.

6. Discussion and Summary

We have developed an AIAI method using the MRNN for the reconstruction of the chemical and density structures of SN Ia models generated using the code TARDIS. With this MRNN architecture, we successfully trained and tested the predictive power of the neural network.

The relevance of this study to real observations is explored with the limited amount of observational data. With the elemental abundances predicted from the neural network, we successfully derived model fits to the spectra of SN 2011fe, SN 2011iv, SN 2015F, SN 2011by, SN 2013dy, ASASSN-14lp, and 15 other SNe Ia near their *B*-band maximum. The AIAI allows derivations of the chemical structures of these SNe.

From the AIAI-deduced elements, we found that SNe Ia with higher stretch factors contain larger Ni masses at velocities above 10,000 km s^{-1} (in zones 2, 3, and 4). We also observed the decline of the mass of ^{56}Ni due to radioactive decay in some well-observed SNe.

We attempted to predict the *B*-band luminosity from the spectral shapes using the AIAI network. The predicted *B*-band absolute magnitude is systematically higher than the luminosity derived from light-curve fits by 0.47–1.26 mag. We surmise that the discrepancy is due to approximations of physical processes made in TARDIS.

Despite these successes, we must stress that the present study is only a preliminary exploration of an exciting approach to SN modeling. The study proves that the combination of deep-learning techniques with physical models of complicated spectroscopic data may yield unique insights into the physical processes in SNe Ia. However, there are a few caveats that need to be kept in mind, and in a way, these caveats also point in the direction of further improvements.

1. In TARDIS, some major assumptions need to be kept in mind. The temperature profiles are calculated based on the assumption that the photospheric spectra follow that of blackbody radiation. There is radioactive material very close to or above the photosphere, so the energy deposition is significantly more complicated than the TARDIS assumptions.
2. The current implementation of the code only applies to data around optical maximum.
3. We subdivided the ejecta into some artificial grids. This introduces unphysical boundaries that are inconsistent with the physics of nuclear burning in the ejecta.
4. The dependence on the adopted baseline model has not been explored yet. Models with different density and chemical profiles need to be studied and built into the spectral library.
5. The spectral models are drawn from a uniform distribution of parameters around the baseline model, which serves as a plain Bayesian prior for the uncertainty estimation discussed in Section 3.4. Different Bayesian priors need to be explored.
6. The highest ejecta velocity explored in this study is 25,000 km s^{-1}, which may be too low for some high-velocity SNe, such as SN 2004dt (Wang et al. 2006a, 2009).
7. The size of the spectral library is still tiny.
8. Our MRNN architecture is constructed to be most sensitive to heavily blended spectral features produced by iron group elements. Other neural network architecture, such as LSTM (Hochreiter & Schmidhuber 1997), may be explored that can improve less blended features.

On a positive note, the current study is sensitive to spectral variations by construction, and the performance of the AIAI network confirms that. We expect the physical approximations made in TARDIS to have only a weak effect on the derivation of chemical elemental abundances.

In summary, we have developed a deep-learning technique to extract physical quantities of SN spectra. Preliminary application of the methods to a set of observational data proves the method to be powerful. More studies are needed to fully realize the potential of the techniques presented in this study.

We thank WISeREP https://wiserep.weizmann.ac.il for the observed SN spectra. We thank Prof. Peter Höflich (Florida State University) for the DDT models. We thank Prof. Ryan J. Foley (University of California Santa Cruz) for the computer-readable version of SN 2011by spectra. We thank Dr. Kate Maguire (Queen University at Belfast) for the light curves and spectra of the SNe in Maguire et al. (2012). X.Z. and L.W. acknowledge support by NSF grant AST-1817099. X.Z. thanks Dr. Wolfgang Kerzendorf (European Southern Observatory) for supportive discussions. Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing. This research made use of TARDIS, a community-developed software package for spectral synthesis in supernovae (Kerzendorf & Sim 2014; Kerzendorf et al. 2019). The development of TARDIS received support from the Google Summer of Code initiative and ESA's Summer of Code in Space program. TARDIS makes extensive use of Astropy and PyNE.

Software: python-keras (https://keras.io), python-dust_extinction https://github.com/karllark/dust_extinction, TARDIS (Kerzendorf & Sim 2014), SiFTO (Conley et al. 2008), SNooPy (Burns et al. 2010).

Appendix A
The MSE for Elements and Zones

In this section, we present two sets of MSEs in Tables A1 and A2. Table 2 is for the neural networks on 2000–10000 Å spectra, which contain 34 trainable element–zone combinations. Table 3 is for the neural networks on 3000–5200 Å spectra, which contain 31 trainable element–zone combinations.

Table A1
MSE for Elements 6 (C) to 28 (Ni) in Four Zones

Element	Zone 1	Zone 2	Zone 3	Zone 4
C	0.1144	0.1155	0.1171	0.1106
N	0.1159	0.1137	0.1165	0.1147
O	0.1147	0.1146	0.1155	0.0822
F	0.1128	0.1166	0.1159	0.1158
Ne	0.1147	0.1131	0.1174	0.1154
Na	0.1134	0.1127	0.1125	0.1121
Mg	0.1120	0.0135	0.0886	0.0509
Al	0.1143	0.1146	0.1149	0.1156
Si	0.0686	0.0177	0.0265	0.1173
P	0.1161	0.1142	0.1147	0.1145
S	0.1162	0.0678	0.1174	0.1134
Cl	0.1109	0.1136	0.1159	0.1153
Ar	0.1147	0.0964	0.1121	0.1127
K	0.1114	0.1122	0.1153	0.1155
Ca	0.0672	0.0285	0.0615	0.0248
Sc	0.1152	0.1145	0.1135	0.0670
Ti	0.0974	0.0673	0.0587	0.0546
V	0.0624	0.0286	0.1096	0.1149
Cr	0.1150	0.1119	0.1164	0.1164
Mn	0.0416	0.0241	0.1105	0.1081
Fe	0.0052	0.0078	0.0090	0.0347
Co	0.0136	0.0409	0.0567	0.0931
Ni	0.0143	0.0434	0.0585	0.0813

Note. All 23 × 4 neural networks involved in this table are trained on 10,000 spectra. The MSEs are tested on a 1829 spectra testing data set. Wavelengths between 2000 and 10000 Å are used as input.

Table A2
MSE for Elements 6 to 28 in Four Zones

Element	Zone 1	Zone 2	Zone 3	Zone 4
C	0.1158	0.1149	0.1160	0.1130
N	0.1156	0.1162	0.1146	0.1129
O	0.1158	0.1150	0.1147	0.1022
F	0.1137	0.1137	0.1153	0.1145
Ne	0.1135	0.1164	0.1154	0.1132
Na	0.1146	0.1144	0.1124	0.1160
Mg	0.1148	0.0733	0.0989	0.0602
Al	0.1148	0.1152	0.1147	0.1134
Si	0.0901	0.0652	0.0593	0.1140
P	0.1148	0.1151	0.1138	0.1152
S	0.1166	0.0750	0.1144	0.1146
Cl	0.1150	0.1145	0.1144	0.1164
Ar	0.1146	0.0992	0.1143	0.1152
K	0.1150	0.1157	0.1149	0.1136
Ca	0.0959	0.0775	0.0788	0.0600
Sc	0.1149	0.1151	0.1157	0.0677
Ti	0.1046	0.0738	0.0676	0.0521
V	0.0925	0.0524	0.1145	0.1150
Cr	0.1137	0.1154	0.1144	0.1123
Mn	0.0757	0.0644	0.1089	0.1153
Fe	0.0473	0.0311	0.0136	0.0465
Co	0.0372	0.0507	0.0568	0.1055
Ni	0.0462	0.0657	0.0583	0.0876

Note. All 23 × 4 neural networks involved in this table are trained on 10,000 spectra. The MSEs are tested on a 1829 spectra testing data set. Wavelengths between 3000 and 5200 Å are used as input.

Appendix B
The Relations between Stretch and Elemental Abundance

The relations between the stretch parameter and the chemical abundance are shown in Figure B1 for IMEs. No obvious correlation is identified.

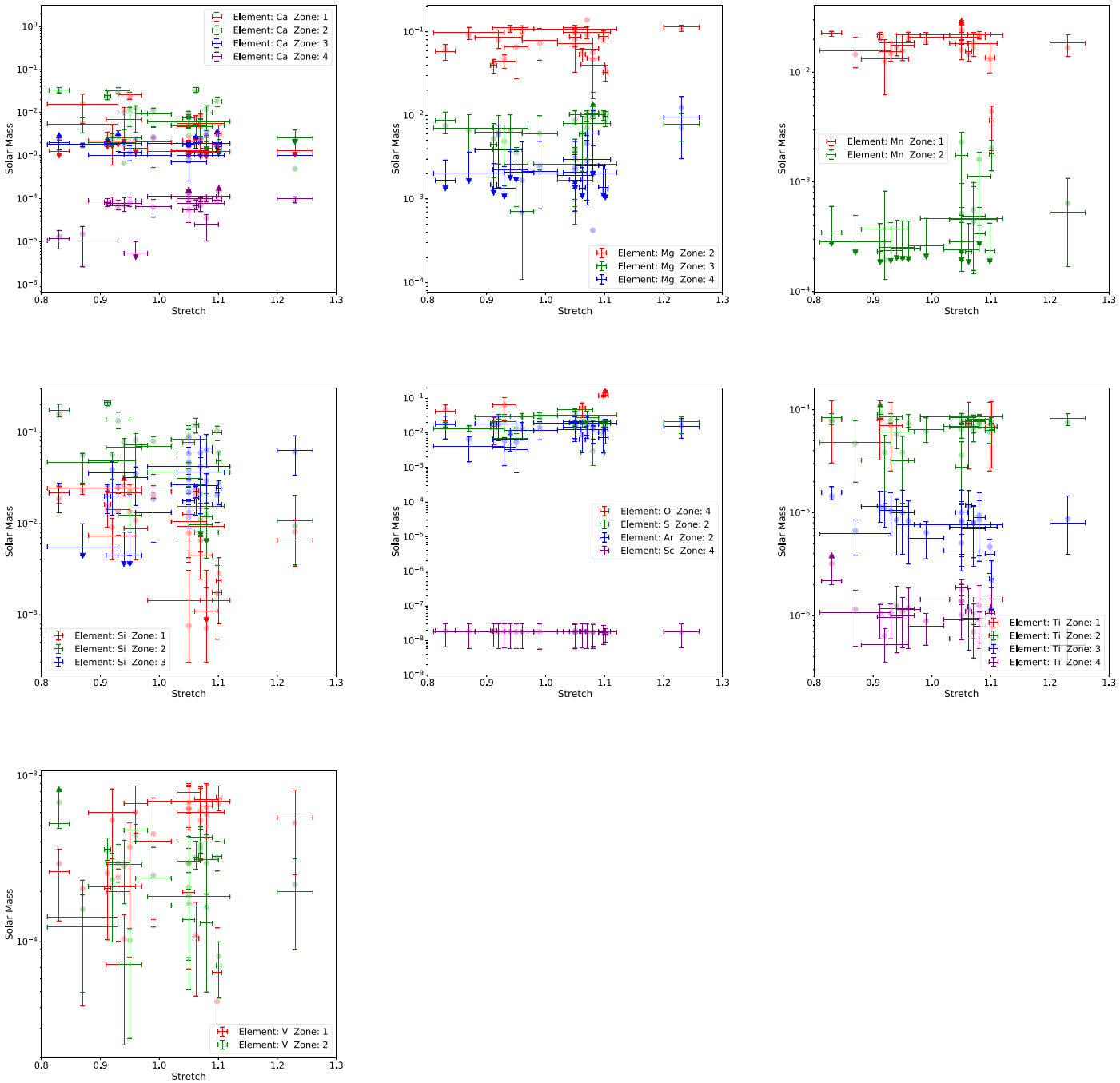

Figure B1. The masses of IMEs are compared with the stretch factors. No obvious correlation is found.

Appendix C
15 HST Spectral Fitting Results

In this section, we present the spectral fitting results of the 15 HST UV spectra, including the synthetic spectra, elemental abundances, density, and temperature profiles (Figures C1–C4).

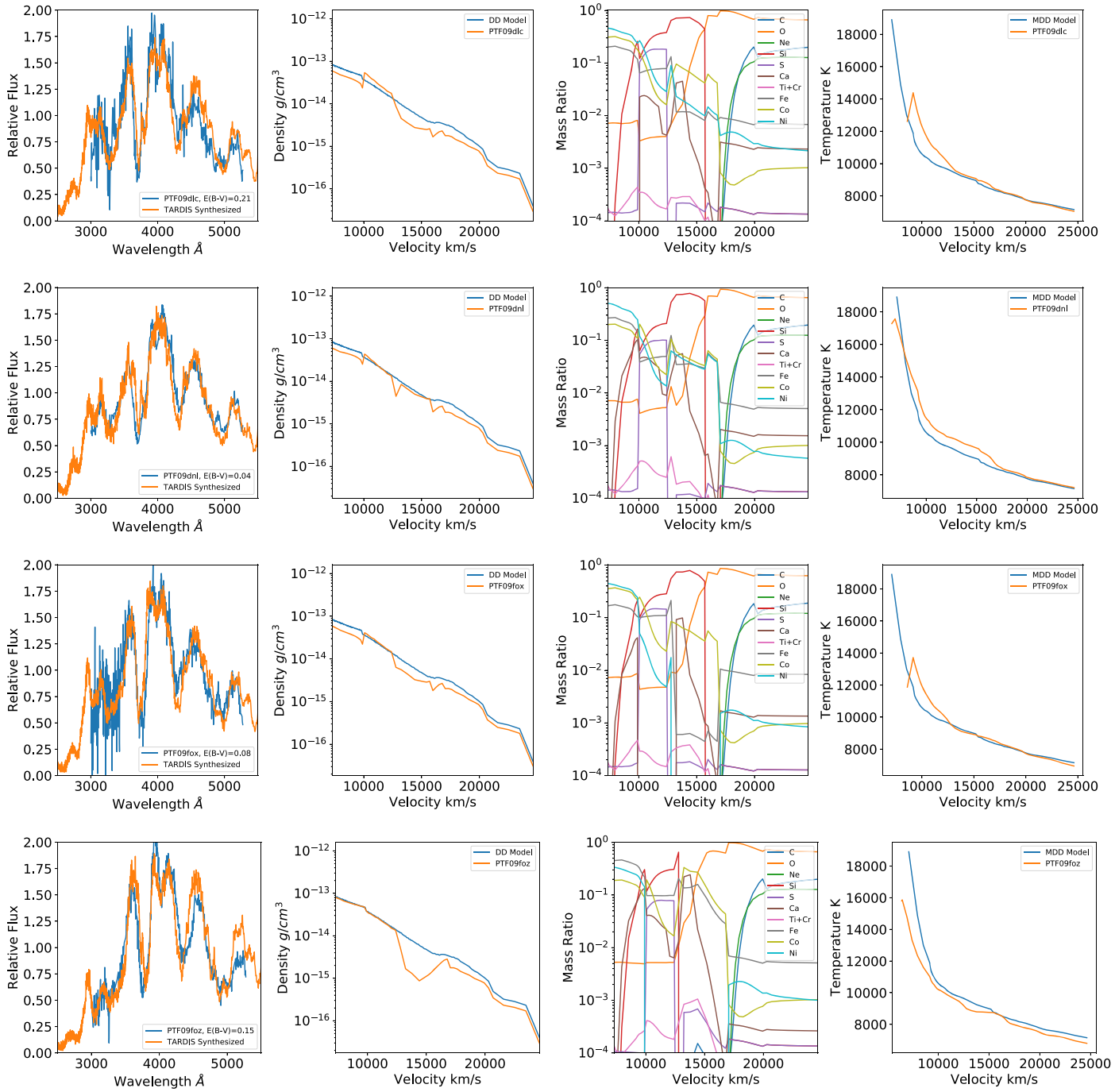

Figure C1. From upper to lower rows: models for PTF 09dlc, PTF 09dnl, PTF 09fox, and PTF 09foz. First column: observed spectra (blue line) after dust extinction correction using the CCM model, with $E(B-V)$ values given in the legend. The TARDIS synthesized spectra (orange line) are calculated with the MRNN derived chemical structure. Second column: density of the ejecta derived from the MRNN (orange line) and the IGM density as a comparison (blue line). All densities are converted to 19 days after explosion. Third column: elemental abundances derived from MRNN and used for TARDIS spectra calculation. Fourth column: temperature structure for the synthesized spectra (orange line) and the temperature for IGM as a comparison (blue line).

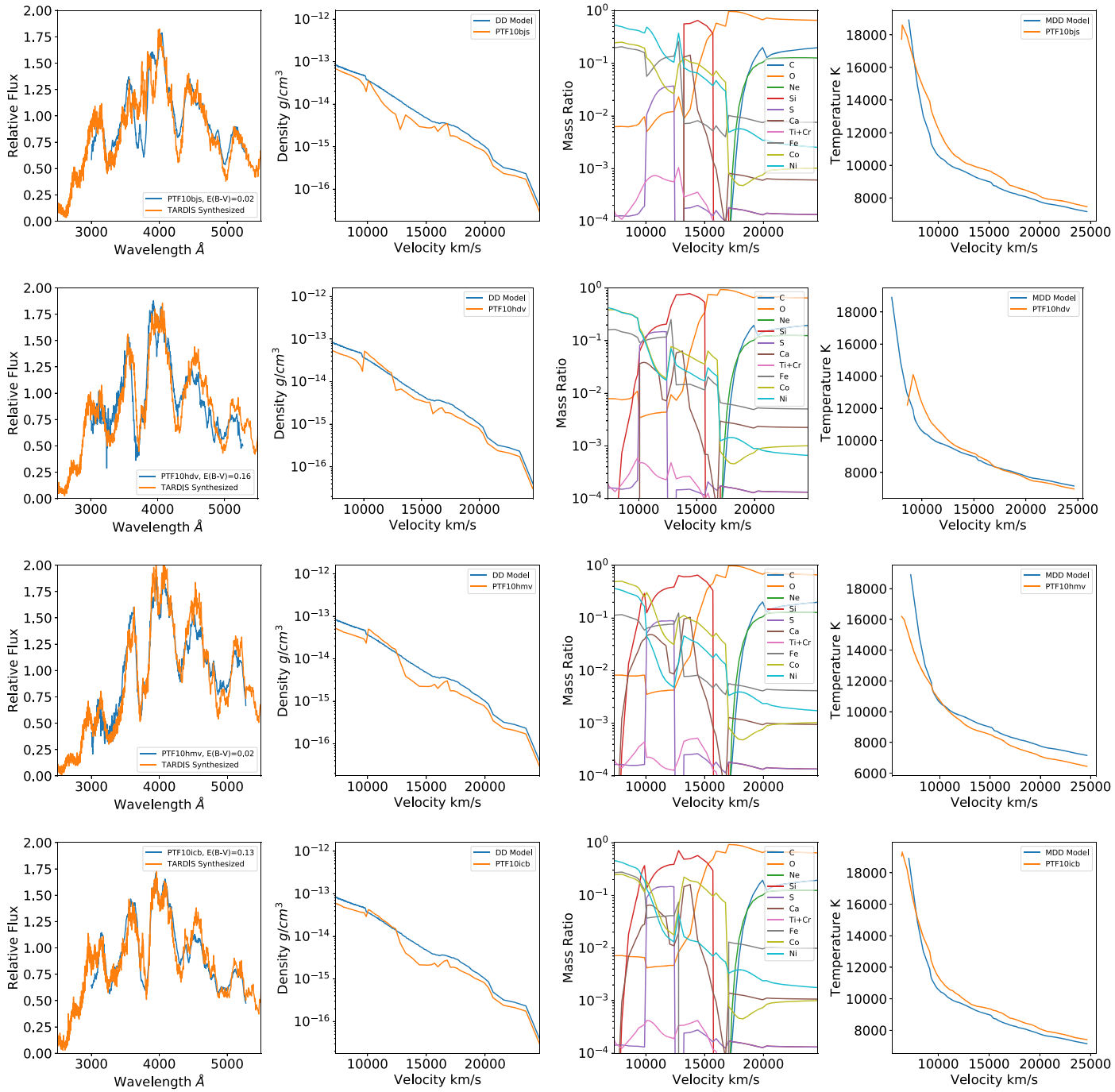

Figure C2. Same as Figure C1 but for PTF 10bjs, PTF 10hdv, PTF 10hmv, and PTF 10icb.

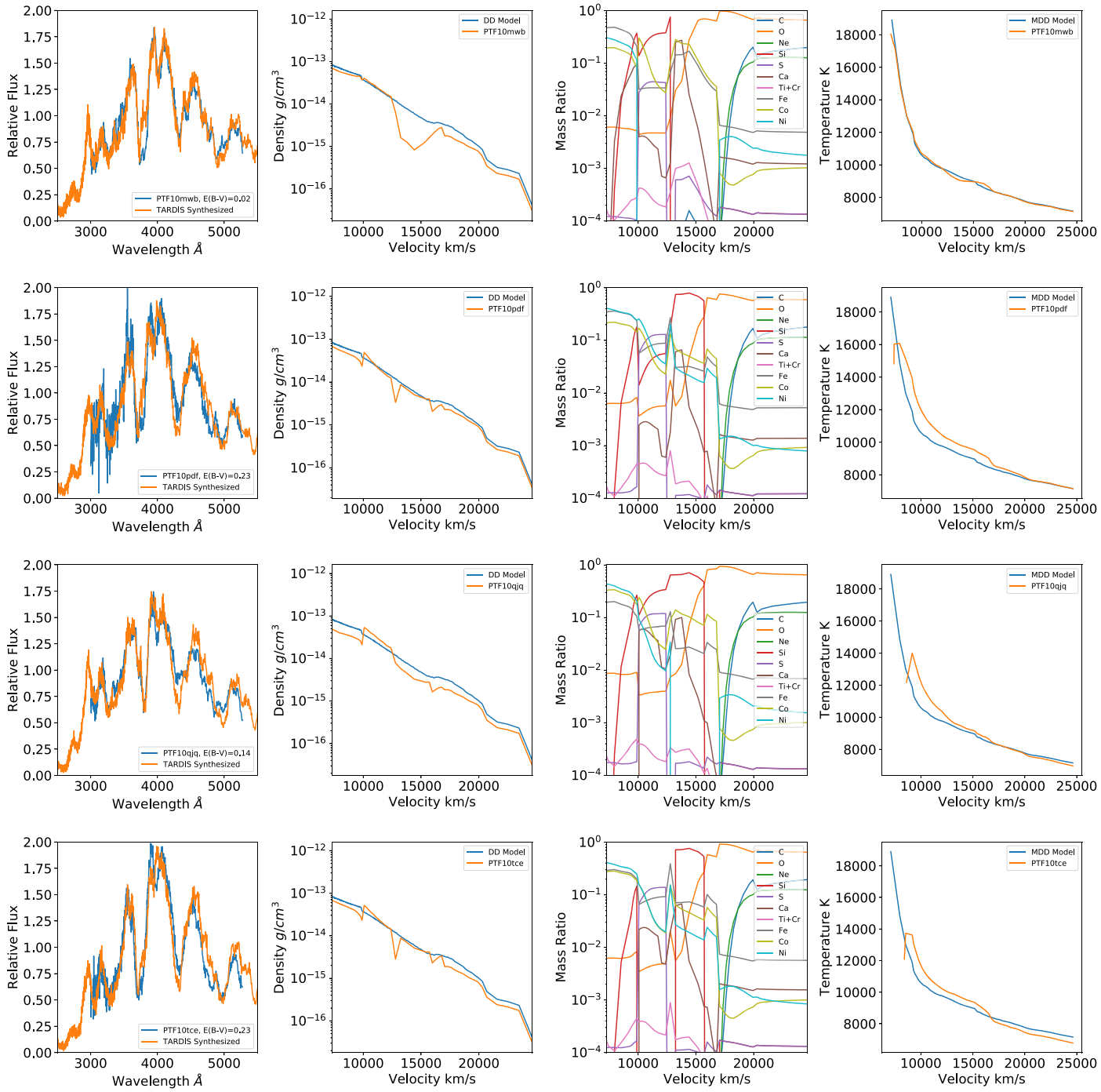

Figure C3. Same as Figure C1 but for PTF 10mwb, PTF 10pdf, and PTF 10qjq.

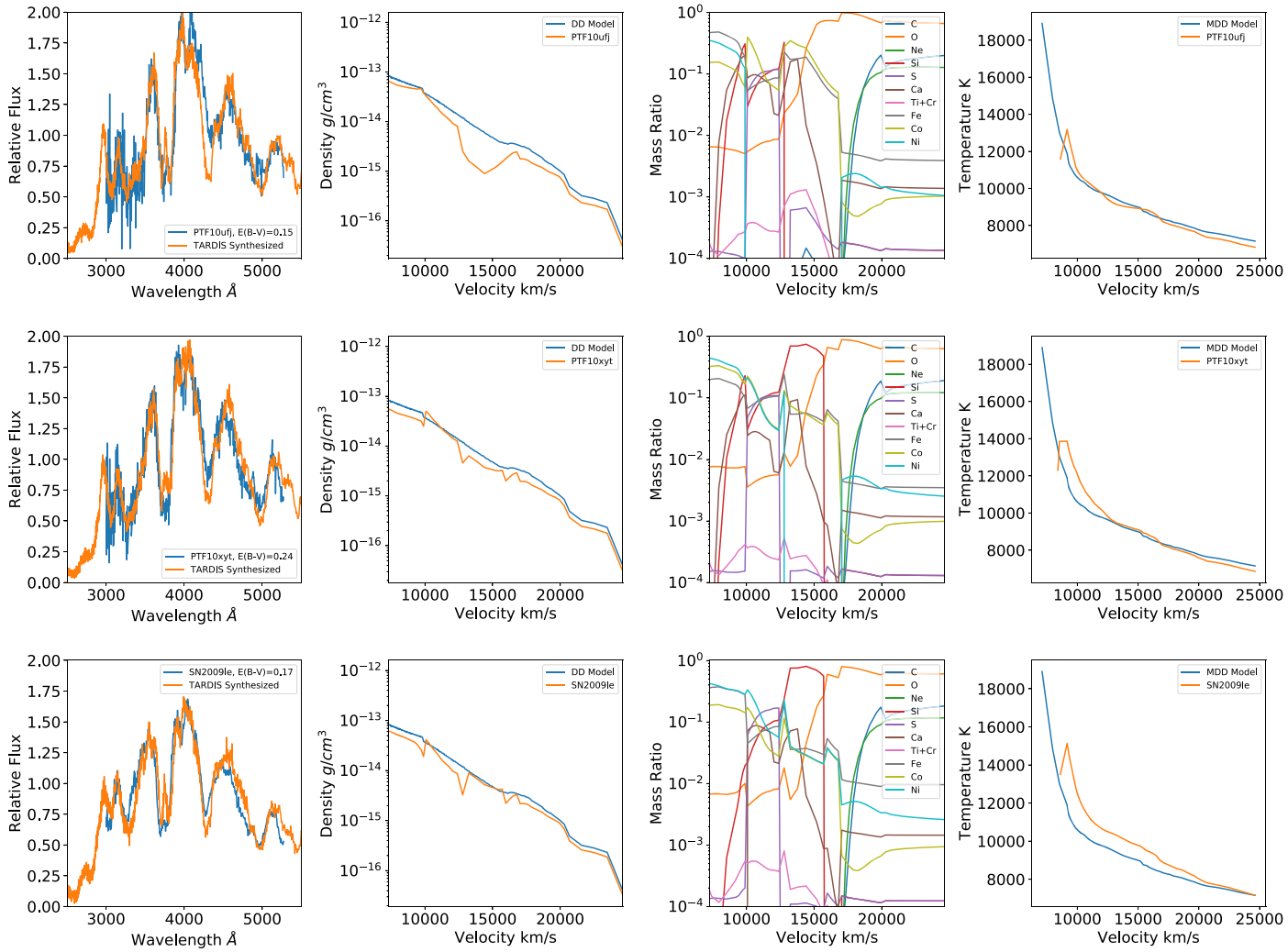

Figure C4. Same as Figure C1 but for PTF 10ufj, PTF 10xyt, and SN 2009le.

Appendix D
The Time Evolution of Elements

Here we show (Figures D1–D4) the time evolution of the masses of IMEs.

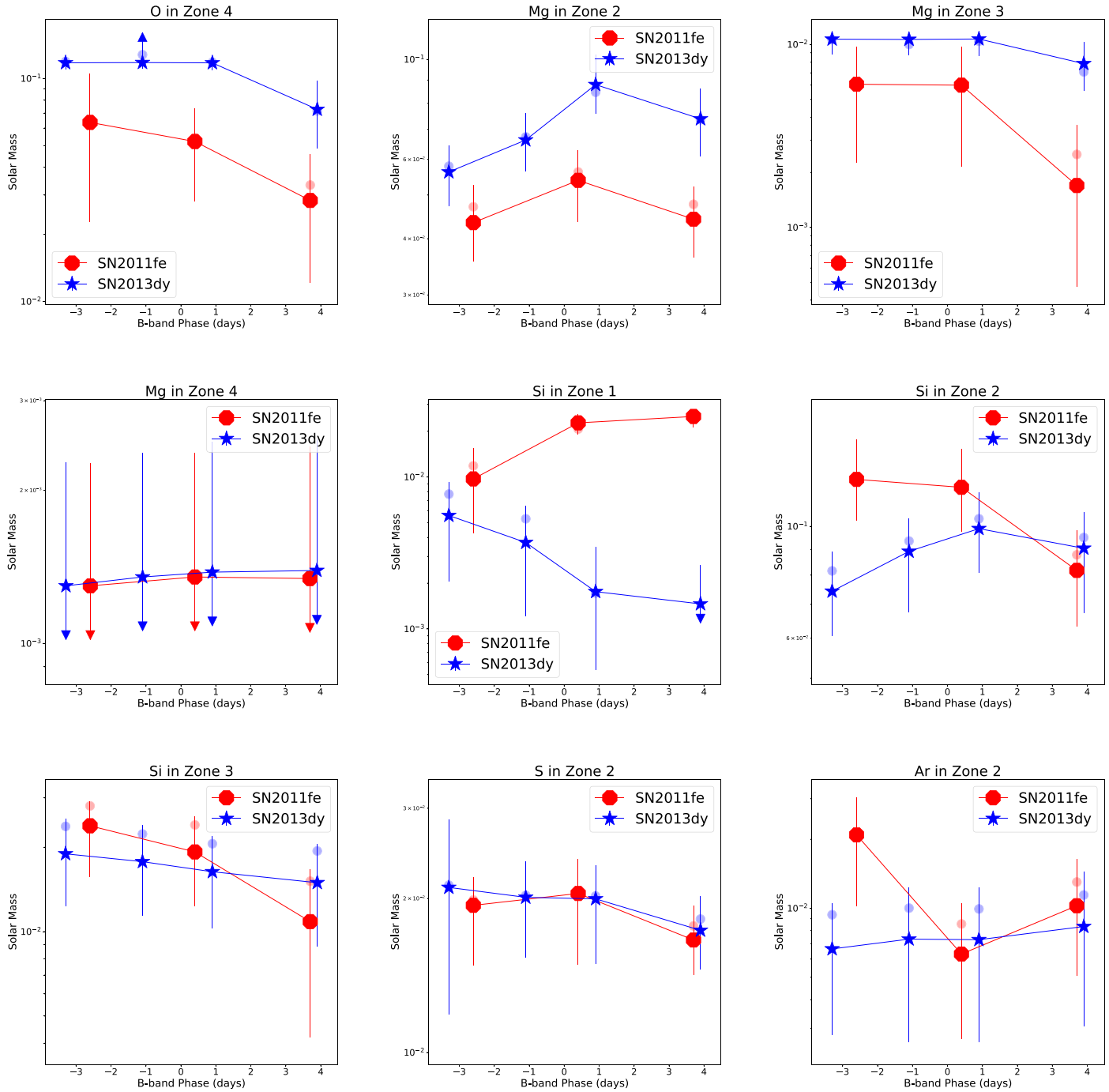

Figure D1. Masses of elements in different zones predicted by the neural network and their evolution with the time after explosion. Red lines are for SN 2011fe, and blue lines are for SN 2013dy. Transparent circles are the predictions from the neural network, crosses with error bars are the median values from the testing data set, error bars indicate the $1-\sigma$ limits, and upper and lower limits are marked with triangles. The elements and zones are labeled in the title of every panel.

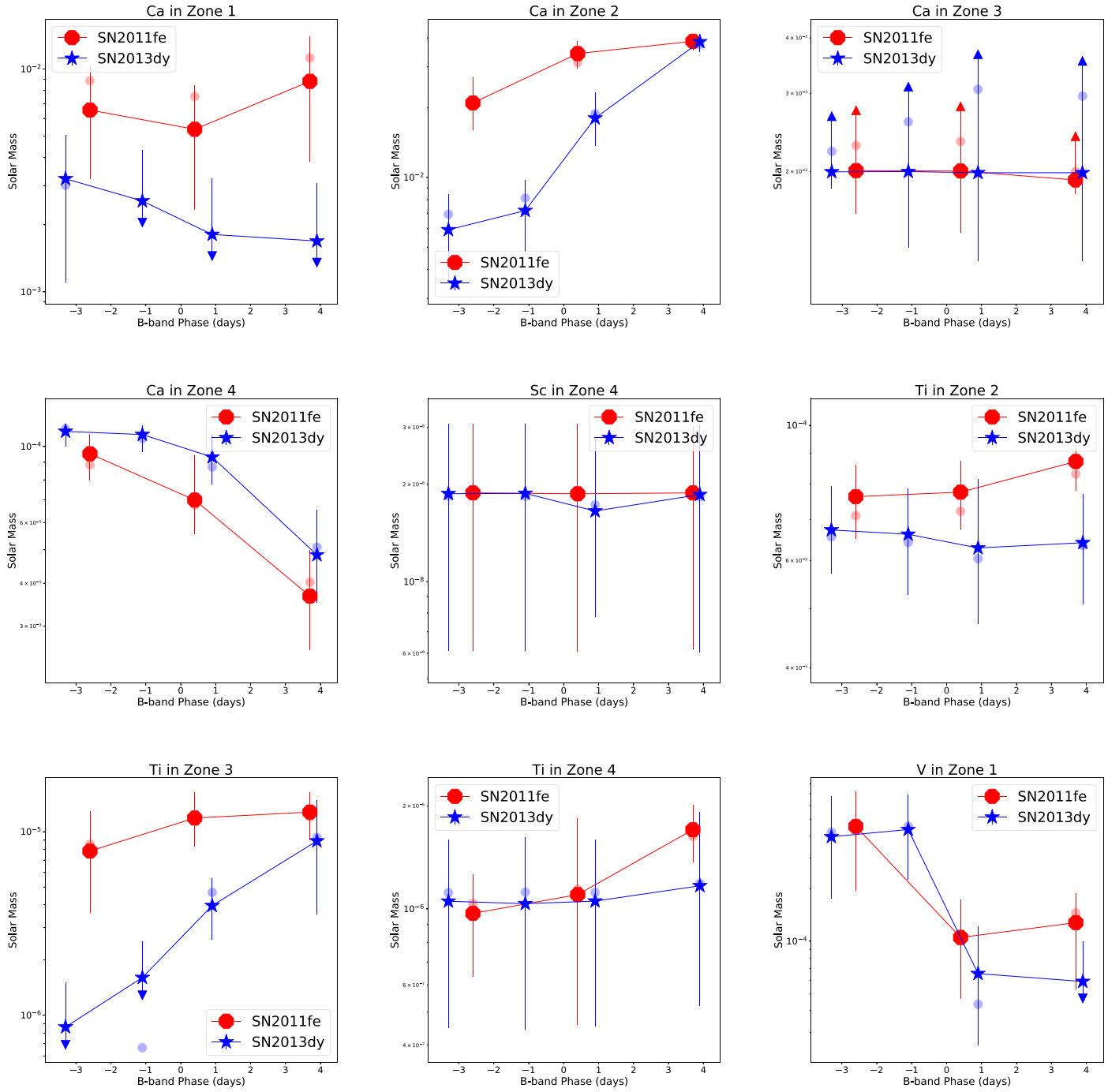

Figure D2. Same as Figure D1 but for different elements and zones.

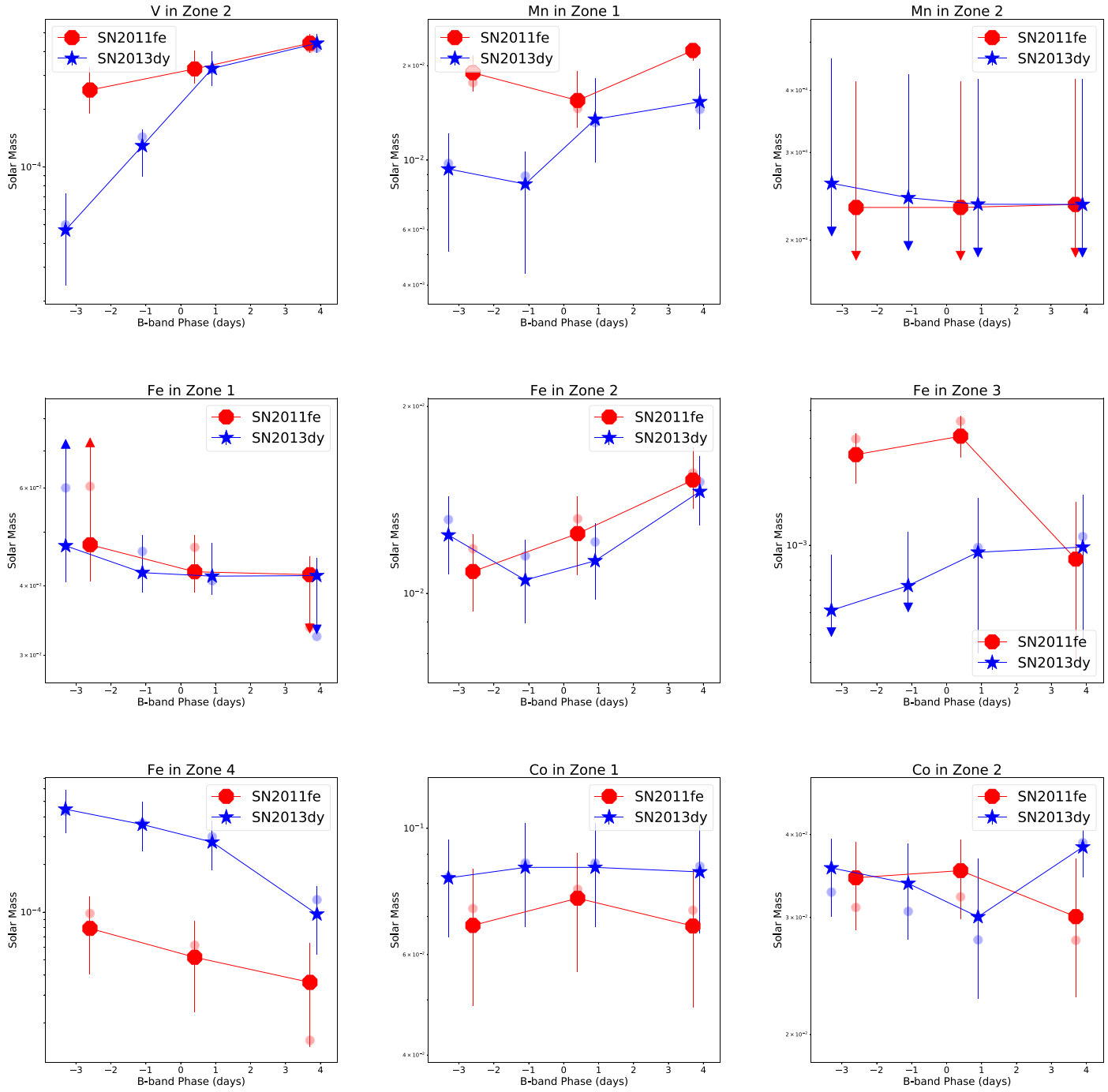

Figure D3. Same as Figure D1 but for different elements and zones.

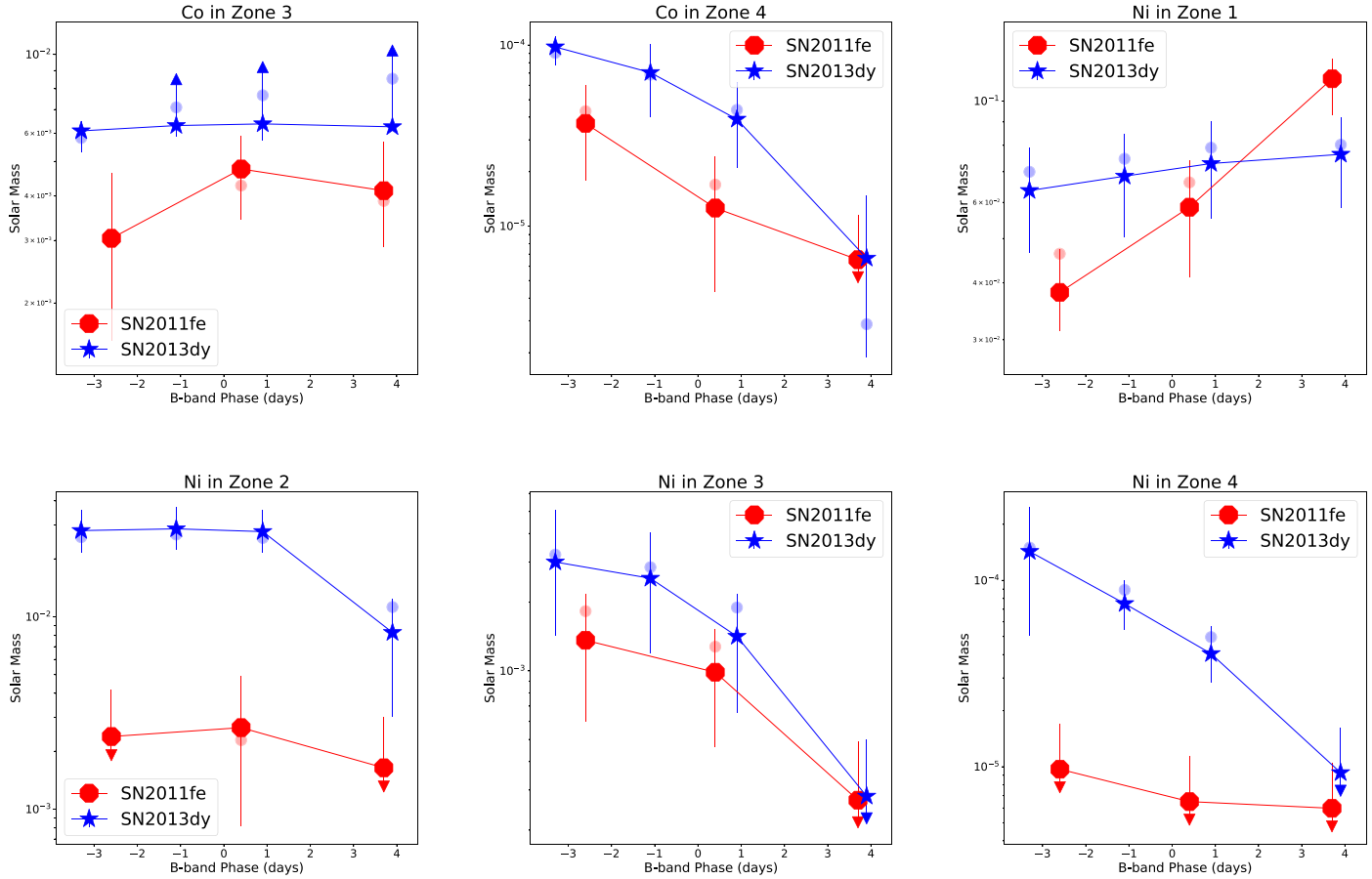

Figure D4. Same as Figure D1 but for different elements and zones.

Appendix E
The Fidelity of 1σ from the Testing Set

There are two caveats for our neural network. First, the MRNN using MSE as a loss function is not designed to mimic the input parameter distributions. Second, the parameter space used in Section 2.3 is a priori and not necessarily similar to the real SN elemental mass-zone distribution. Therefore, the results from the median value on the testing set may be biased, and we directly adopt the predictions from the neural network to calculate the spectra in Section 4. However, the $1-\sigma$ errors from the testing set can be indicative of the sensitivity of MRNNs on different elements and zones and allow us to assess the fitting fidelity.

We calculated the spectra using the median values as the mass estimates, as shown in Figure E1. Moreover, we modified the abundances of Fe in zone 2, Ni in zone 2, and Ni in zone 3 mass by $\pm 1\sigma$ based on neural network predictions of the 11 spectra with a wavelength range of 2000–10000 Å to evaluate their effect on the spectral profiles; these are shown in Figures E2, E3 and E4, respectively. These calculations prove that the results from both the median and direct estimates of TARDIS model parameters reproduce the observations well.

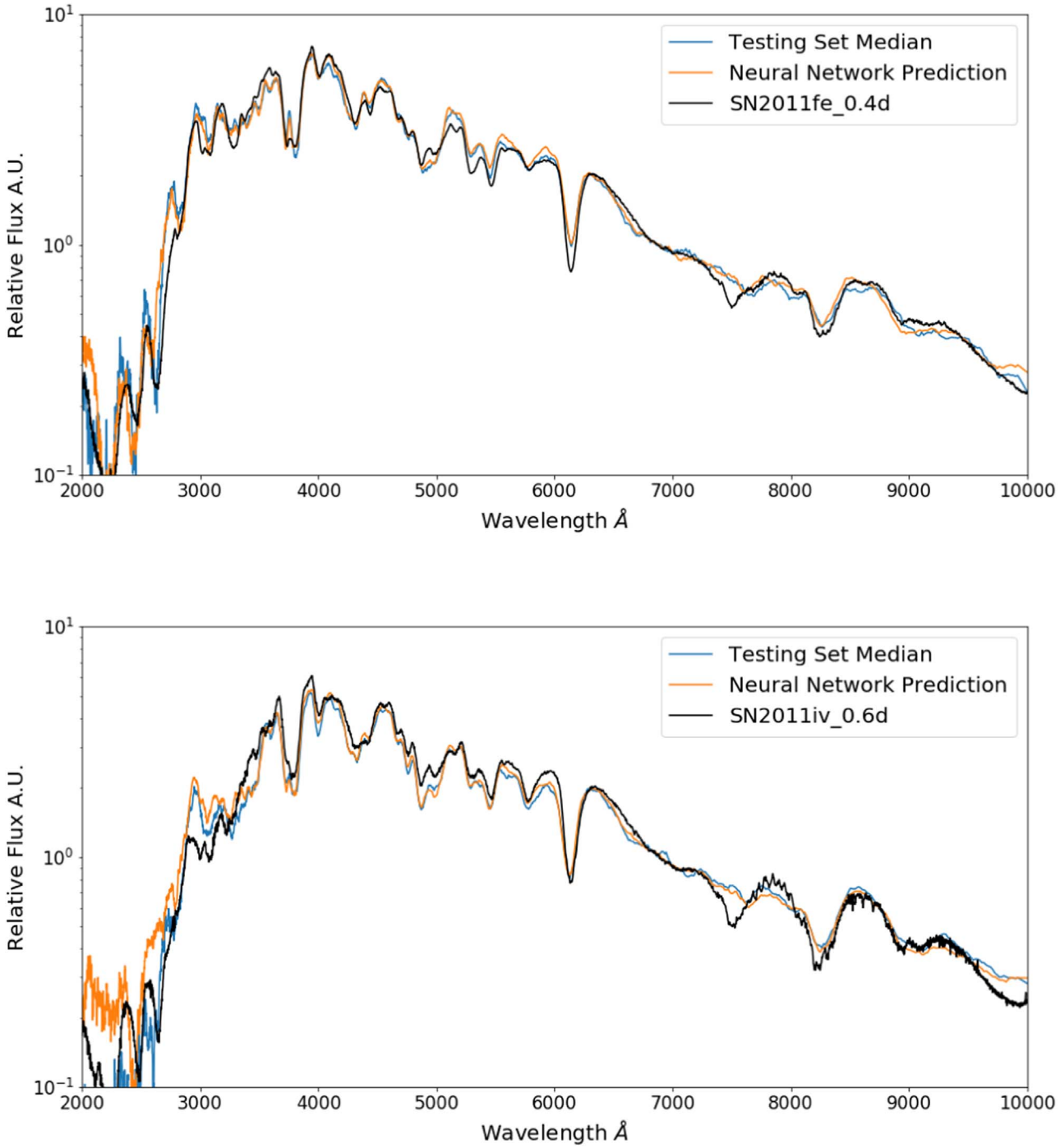

Figure E1. Top: observed spectrum of SN 2011fe at day 0.4 (black) compared with the TARDIS spectrum calculated using the MRNN predicted ejecta structure (orange) and the median values as the estimates of the ejecta structure (blue). This demonstrates that the predicted spectrum is robust to the methods of ejecta structure estimation. Bottom: same as the top panel but for SN 2011iv at 0.6 day.

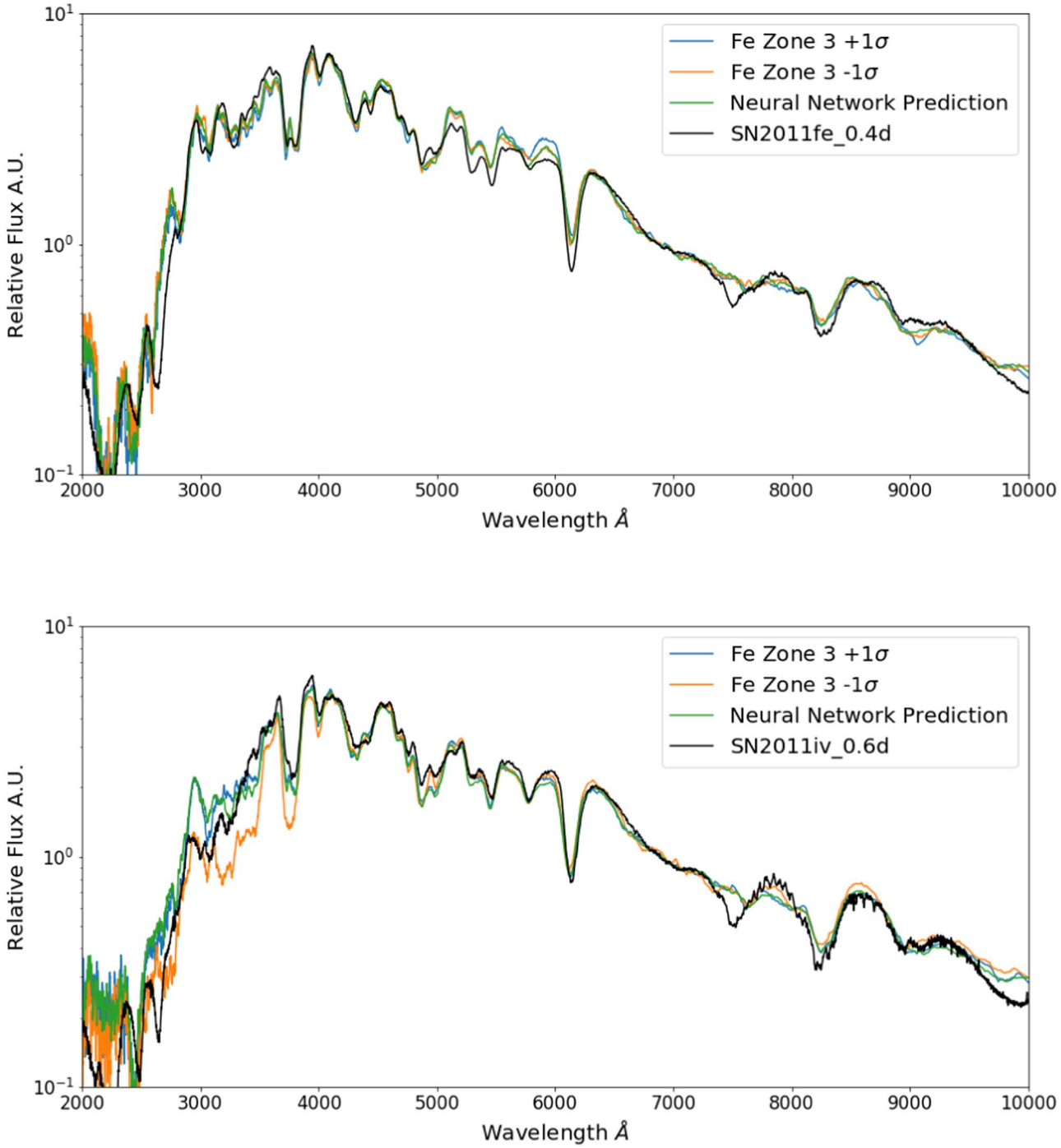

Figure E2. Top: SN 2011fe on day 0.4. The observed spectrum (black line) is compared with the TARIDS spectra calculated using the MRNN estimated chemical structure (green line) and the Fe abundance in zone 3 enhanced by 1σ (blue line) and reduced by 1σ (orange line). Bottom: same as the top panel but for SN2011iv at 0.6 day.

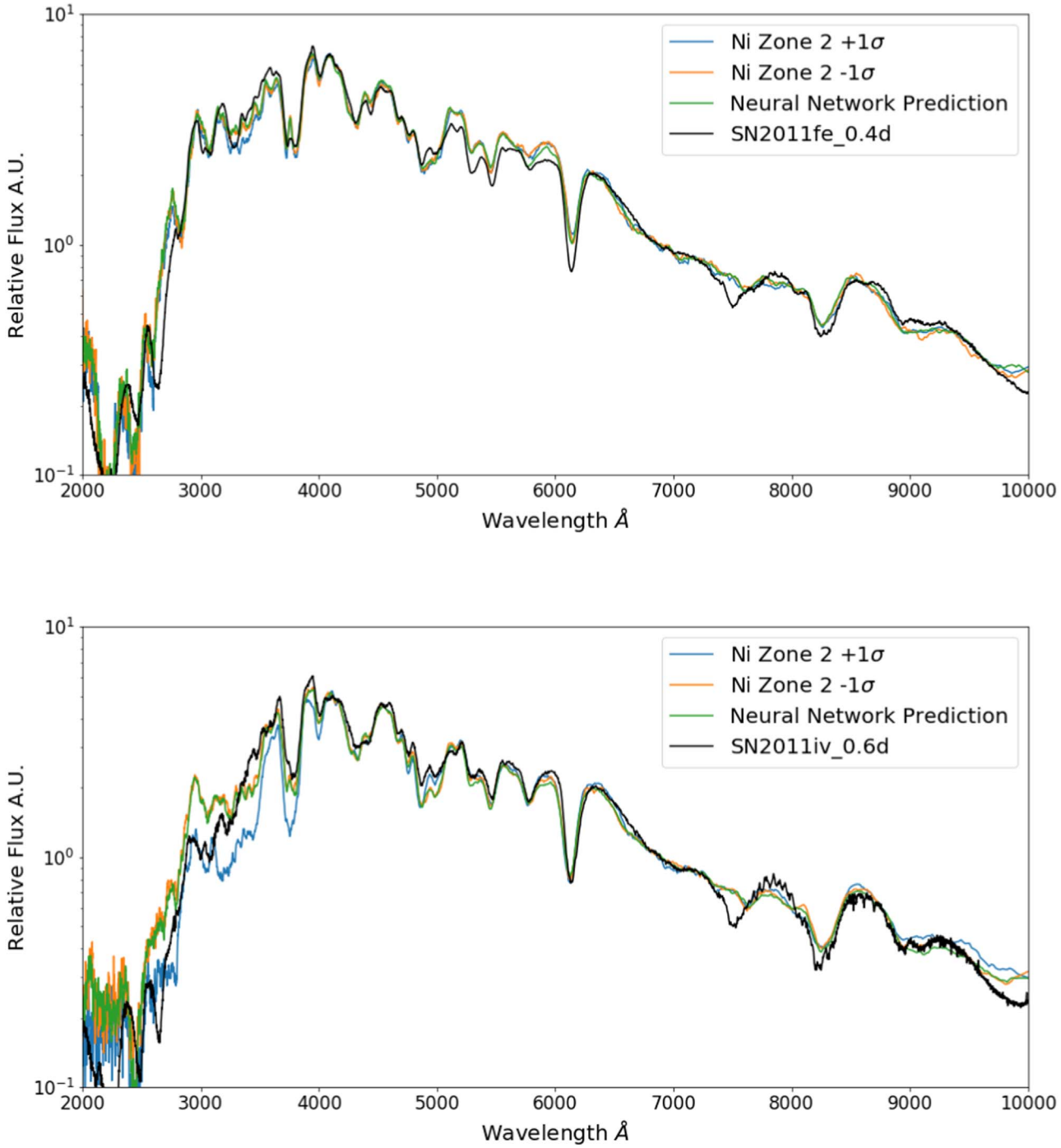

Figure E3. Same as Figure E2 but with Ni in zone 2 enhanced or reduced by 1σ.

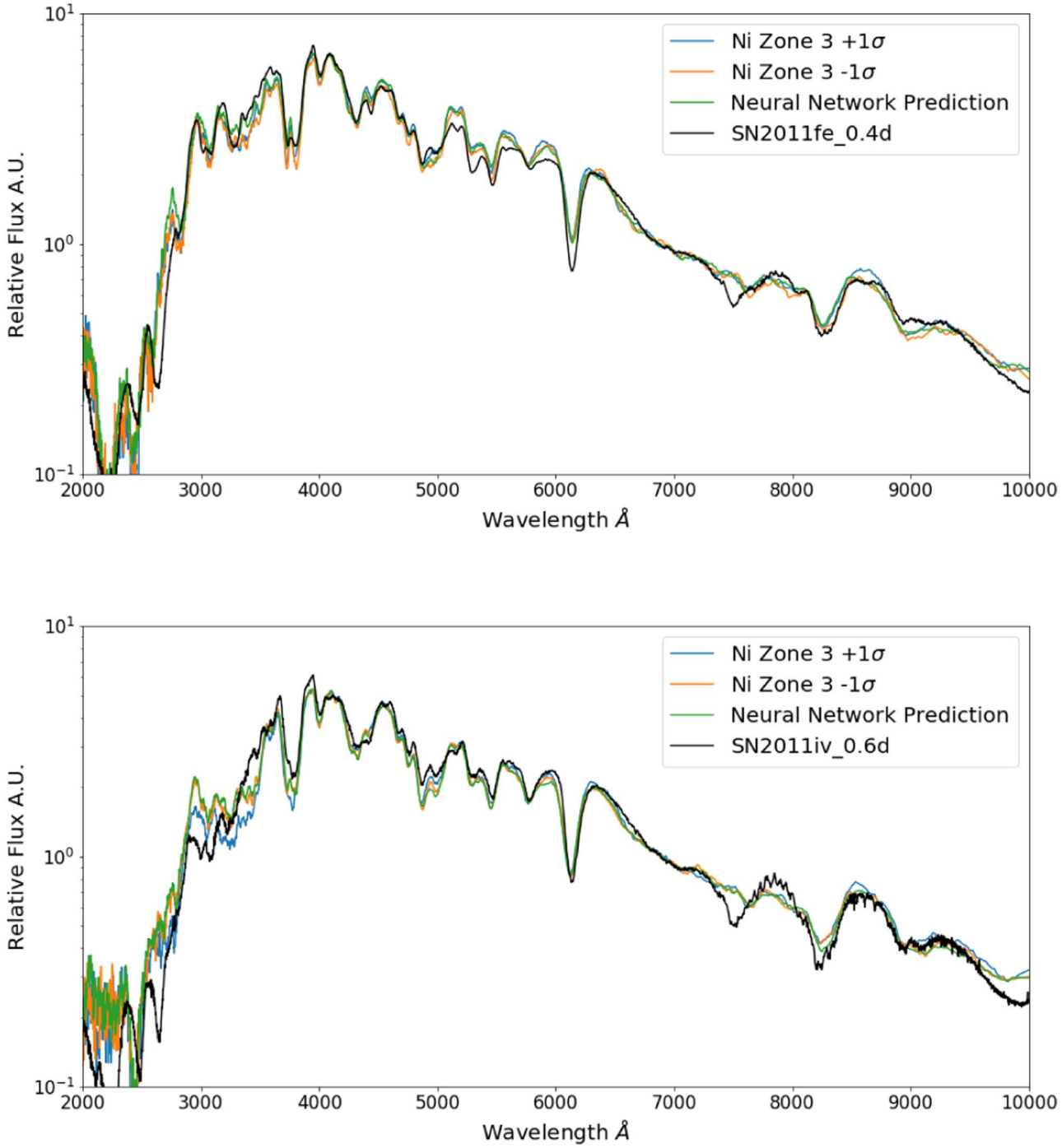

Figure E4. Same as Figure E2 but with Ni in zone 3 enhanced or reduced by 1σ.

References

Abadi, M., Barham, P., Chen, J., et al. 2016, in 12th USENIX Symp. on Operating Systems Design and Implementation, OSDI 16 (Berkeley, CA: USENIX), 265
Abbott, D. C., & Lucy, L. B. 1985, ApJ, 288, 679
Abdi, M., & Nahavandi, S. 2016, arXiv:1609.05672
Ashall, C., Mazzali, P. A., Stritzinger, M. D., et al. 2018, MNRAS, 477, 153
Barna, B., Szalai, T., Kerzendorf, W. E., et al. 2018, MNRAS, 480, 3609
Barna, B., Szalai, T., Kromer, M., et al. 2017, MNRAS, 471, 4865
Baron, E., & Hauschildt, P. H. 1998, ApJ, 495, 370
Bialek, S., Fabbro, S., Venn, K. A., et al. 2019, arXiv:1911.02602
Blondin, S., Dessart, L., & Hillier, D. J. 2017, MNRAS, 474, 3931
Blondin, S., Dessart, L., Hillier, D. J., & Khokhlov, A. M. 2013, MNRAS, 429, 2127
Branch, D., Dang, L. C., & Baron, E. 2009, PASP, 121, 238
Bu, Y., Zeng, J., & Lei, Z. 2019, ApJ, 886, 128
Bulla, M., Sim, S. A., & Kromer, M. 2015, MNRAS, 450, 967
Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2010, AJ, 141, 19
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Cartier, R., Sullivan, M., Firth, R. E., et al. 2016, MNRAS, 464, 4476
Chomiuk, L., Soderberg, A. M., Moe, M., et al. 2012, ApJ, 750, 164
Cikota, A., Patat, F., Wang, L., et al. 2019, MNRAS, 490, 578
Conley, A., Sullivan, M., Hsiao, E. Y., et al. 2008, ApJ, 681, 482
Fabbro, S., Bialek, S., O'Briain, T., et al. 2017, MNRAS, 475, 2978
Fitzpatrick, E. L. 1999, PASP, 111, 63
Foley, R. J., Hoffmann, S. L., Macri, L. M., et al. 2020, MNRAS, 491, 5991
Foley, R. J., Kromer, M., Howie Marion, G., et al. 2012, ApJL, 753, L5
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gall, C., Stritzinger, M. D., Ashall, C., et al. 2018, A&A, 611, A58

Gamezo, V. N., Khokhlov, A. M., & Oran, E. S. 2004, PhRvL, 92, 211102
Gamezo, V. N., Khokhlov, A. M., Oran, E. S., Chtchelkanova, A. Y., & Rosenberg, R. O. 2003, Sci, 299, 77
Garavini, G., Nobili, S., Taubenberger, S., et al. 2007, A&A, 471, 527
Gerardy, C. L., Meikle, W. P. S., Kotak, R., et al. 2007, ApJ, 661, 995
Gordon, K. D., Cartledge, S., & Clayton, G. C. 2009, ApJ, 705, 1320
Graham, M. L., Foley, R. J., Zheng, W., et al. 2015, MNRAS, 446, 2073
Graur, O., Zurek, D. R., Cara, M., et al. 2018a, ApJ, 866, 10
Graur, O., Zurek, D. R., Rest, A., et al. 2018b, ApJ, 859, 79
Hachinger, S., Mazzali, P. A., Taubenberger, S., Pakmor, R., & Hillebrandt, W. 2009, MNRAS, 399, 1238
He, K., Zhang, X., Ren, S., & Sun, J. 2015, arXiv:1512.03385
Hillier, D. J., & Miller, D. L. 1998, ApJ, 496, 407
Hochreiter, S., & Schmidhuber, J. 1997, Neural Computation, 9, 1735
Hoeflich, P., Hsiao, E. Y., Ashall, C., et al. 2017, ApJ, 846, 58
Hoeflich, P., Khokhlov, A., Wheeler, J. C., et al. 1996a, ApJL, 472, L81
Hoeflich, P., Wheeler, J. C., Hines, D. C., & Trammell, S. R. 1996b, ApJ, 459, 307
Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. 2016, arXiv:1608.06993
Ioffe, S., & Szegedy, C. 2015, arXiv:1502.03167
Iwamoto, K., Brachwitz, F., Nomoto, K., et al. 1999, ApJS, 125, 439
Jordan, G. C. I., Fisher, R. T., Townsley, D. M., et al. 2008, ApJ, 681, 1448
Kasen, D., Branch, D., Baron, E., & Jeffery, D. 2002, ApJ, 565, 380
Kasen, D., & Plewa, T. 2005, ApJL, 622, L41
Kasen, D., Thomas, R. C., & Nugent, P. 2006, ApJ, 651, 366
Kerzendorf, W., Nöbauer, U., Sim, S., et al. 2019, tardis-sn/tardis: TARDIS v3.0 alpha2, Zenodo, doi:10.5281/zenodo.2590539
Kerzendorf, W. E., & Sim, S. A. 2014, MNRAS, 440, 387
Khokhlov, A. M. 1991a, A&A, 246, 383
Khokhlov, A. M. 1991b, A&A, 245, 114
Kingma, D. P., & Ba, J. 2014, arXiv:1412.6980
Krizhevsky, A. 2009, The CIFAR-10 Dataset (Toronto: Univ. Toronto), https://www.cs.toronto.edu/~kriz/cifar.html
Krizhevsky, A., Sutskever, I., & Hinton, G. E. 2012, Adv. Neural Info. Process. Systems, 25, 1097, https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks
Kromer, M., & Sim, S. A. 2009, MNRAS, 398, 1809
Law, N. M., Kulkarni, S. R., Dekany, R. G., et al. 2009, PASP, 121, 1395
LeCun, Y., Bengio, Y., & Hinton, G. 2015, Natur, 521, 436
Lentz, E. J., Baron, E., Branch, D., & Hauschildt, P. H. 2001, ApJ, 547, 402
Liu, J., Gibson, S. J., Mills, J., & Osadchy, M. 2019, Chemometrics and Intelligent Laboratory Systems, 184, 175
Lucy, L. B. 1971, ApJ, 163, 95
Lucy, L. B. 1999, A&A, 345, 211
Lucy, L. B. 2002, A&A, 384, 725
Maguire, K., Sullivan, M., Ellis, R. S., et al. 2012, MNRAS, 426, 2359
Mazzali, P. A. 2000, A&A, 363, 705
Mazzali, P. A., & Lucy, L. B. 1993, A&A, 279, 447
Mazzali, P. A., Lucy, L. B., Danziger, I. J., et al. 1993, A&A, 269, 423
Mazzali, P. A., Sullivan, M., Hachinger, S., et al. 2014, MNRAS, 439, 1959
Montúfar, G., Pascanu, R., Cho, K., & Bengio, Y. 2014, arXiv:1402.1869
Munari, U., Henden, A., Belligoli, R., et al. 2013, NewA, 20, 30
Nidever, D. L., Holtzman, J. A., Prieto, C. A., et al. 2015, AJ, 150, 173
Nomoto, K., Thielemann, F.-K., & Yokoi, K. 1984, ApJ, 286, 644
Nugent, P. E., Sullivan, M., Cenko, S. B., et al. 2011, Natur, 480, 344
Pan, Y.-C., Foley, R. J., Kromer, M., et al. 2015, MNRAS, 452, 4307
Parrent, J., Branch, D., & Jeffery, D. 2010, SYNOW: A Highly Parameterized Spectrum Synthesis Code for Direct Analysis of SN Spectra, Astrophysics Source Code Library, ascl:1010.055
Patat, F., Cordiner, M. A., Cox, N. L. J., et al. 2013, A&A, 549, A62
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Phillips, M. M. 1993, ApJL, 413, L105
Plewa, T., Calder, A. C., & Lamb, D. Q. 2004, ApJL, 612, L37
Poludnenko, A. Y., Chambers, J., Ahmed, K., Gamezo, V. N., & Taylor, B. D. 2019, Sci, 366, aau7365
Poludnenko, A. Y., Gardiner, T. A., & Oran, E. S. 2011, PhRvL, 107, 054501
Pskovskii, I. P. 1977, SvA, 21, 675
Rau, A., Kulkarni, S. R., Law, N. M., et al. 2009, PASP, 121, 1334
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
Röpke, F. K., Gieseler, M., Reinecke, M., Travaglio, C., & Hillebrandt, W. 2006, A&A, 453, 203
Rubin, D., Knop, R. A., Rykoff, E., et al. 2013, ApJ, 763, 35
Savitzky, A., & Golay, M. J. E. 1964, AnaCh, 36, 1627
Shappee, B. J., Piro, A. L., Holoien, T. W.-S., et al. 2016, ApJ, 826, 144
Shen, K. J., Boubert, D., Gänsicke, B. T., et al. 2018, ApJ, 865, 15
Silverman, J. M., Ganeshalingam, M., & Filippenko, A. V. 2013, MNRAS, 430, 1030
Simonyan, K., & Zisserman, A. 2014, arXiv:1409.1556
Smartt, S. J., Chen, T.-W., Jerkstrand, A., et al. 2017, Natur, 551, 75
Stehle, M., Mazzali, P. A., Benetti, S., & Hillebrandt, W. 2005, MNRAS, 360, 1231
Thomas, R. C. 2013, SYN++: Standalone SN Spectrum Synthesis, Astrophysics Source Code Library, ascl:1308.008
Thomas, R. C., Nugent, P. E., & Meza, J. 2011, PASP, 123, 237
Timmes, F. X., Brown, E. F., & Truran, J. W. 2003, ApJL, 590, L83
Vinkó, J., Ordasi, A., Szalai, T., et al. 2018, PASP, 130, 064101
Vogl, C., Sim, S. A., Noebauer, U. M., Kerzendorf, W. E., & Hillebrandt, W. 2019, A&A, 621, A29
Wang, B., & Han, Z. 2012, NewAR, 56, 122
Wang, L., Baade, D., Hoflich, P., et al. 2006a, ApJ, 653, 490
Wang, L., Goldhaber, G., Aldering, G., & Perlmutter, S. 2003, ApJ, 590, 944
Wang, L., Strovink, M., Conley, A., et al. 2006b, ApJ, 641, 50
Wang, L., & Wheeler, J. C. 2008, ARA&A, 46, 433
Wang, L., Wheeler, J. C., Li, Z., & Clocchiatti, A. 1996, ApJ, 467, 435
Wang, X., Filippenko, A. V., Ganeshalingam, M., et al. 2009, ApJL, 699, L139
Wang, X., Wang, L., Filippenko, A. V., Zhang, T., & Zhao, X. 2013, Sci, 340, 170
Yang, Y., Hoeflich, P. A., Baade, D., et al. 2019, arXiv:1903.10820
Zhai, Q., Zhang, J.-J., Wang, X.-F., et al. 2016, AJ, 151, 125