

SPECTRAL CLASSIFICATION OF GALAXIES AT $0.5 \leq z \leq 1$ IN THE CDFS: THE ARTIFICIAL NEURAL NETWORK APPROACH

H. TEIMOORINIA

Department of Physics and Astronomy, University of Victoria, Victoria, British Columbia, V8P 1A1, Canada; hteimoo@uvic.ca

Received 2012 March 28; accepted 2012 September 26; published 2012 November 8

ABSTRACT

The aim of this work is to combine spectral energy distribution (SED) fitting with artificial neural network techniques to assign spectral characteristics to a sample of galaxies at $0.5 < z < 1$. The sample is selected from the spectroscopic campaign of the ESO/GOODS-South field, with 142 sources having photometric data from the GOODS-MUSIC catalog covering bands between ~ 0.4 and $24 \mu\text{m}$ in 10–13 filters. We use the CIGALE code to fit photometric data to Maraston’s synthesis spectra to derive mass, specific star formation rate, and age, as well as the best SED of the galaxies. We use the spectral models presented by Kinney et al. as targets in the wavelength interval $\sim 1200\text{--}7500 \text{ \AA}$. Then a series of neural networks are trained, with average performance $\sim 90\%$, to classify the best SED in a supervised manner. We consider the effects of the prominent features of the best SED on the performance of the trained networks and also test networks on the galaxy spectra of Coleman et al., which have a lower resolution than the target models. In this way, we conclude that the trained networks take into account all the features of the spectra simultaneously. Using the method, 105 out of 142 galaxies of the sample are classified with high significance. The locus of the classified galaxies in the three graphs of the physical parameters of mass, age, and specific star formation rate appears consistent with the morphological characteristics of the galaxies.

Key words: galaxies: fundamental parameters – galaxies: high-redshift – methods: data analysis – methods: statistical

Online-only material: color figure

1. INTRODUCTION

The physical properties of galaxies can be determined by fitting model spectral energy distributions (SEDs) to photometric data in different spectral passbands. In this way, it is possible to obtain properties such as stellar mass, age, and star formation rate (SFR). The models may be obtained from different observations, e.g., Coleman et al. (1980, hereafter CWW), Kinney et al. (1996, hereafter K96), and Gregg et al. (2003), or estimated by population synthesis codes based on theoretical methods, e.g., Leitherer et al. (1999), Bruzual & Charlot (2003), Maraston (2005), Lancon et al. (2008), and Molla et al. (2009). The question of which of them is preferable is the subject of ongoing research (Coelho et al. 2009).

On the empirical side, the content of the libraries is limited by the available observations. Limiting the range from the ultraviolet to the near-infrared proves rather restrictive. In these libraries the morphological nature of a galaxy is usually clear. On the theoretical side, the parameter space of the libraries can be tuned in order to generate the desired spectra in a wide range of wavelengths. They can also be subjected to environmental effects such as dust attenuation in a physically realistic manner. The aim is that the models should cover a realistic range of observational properties. They are not directly presented in terms of morphological properties and so it is useful to consider the advantages of applying both theoretical and empirical analyses.

Good fitting of galaxy data requires a realistic treatment of the effects of gas and dust on a galaxy’s SED. To this end, Noll et al. (2009, hereafter N09) developed a new code, called CIGALE. This focuses on the models of Maraston (2005) that also consider the thermally pulsating asymptotic giant branch and provide theoretical spectra of stars at a resolution of $10\text{--}20 \text{ \AA}$ FWHM in the wavelength range from 91 \AA to $160 \mu\text{m}$.

CIGALE offers the use of attenuation laws differing from and more comprehensive than the Calzetti law (Calzetti et al. 2000). It can also combine two composite stellar populations (CSP). According to their mass fraction, the two CS can be linked and can result in a more complex SED. The code can consider UV/optical interstellar emission and absorption lines, using empirical templates based on the spectra presented by K96 and Noll et al. (2004). To test the code, N09 use a well-known classified sample, SINGS (Kennicutt et al. 2003), which contains different morphological galaxy types.

From a variety of observations, K96 introduce 12 spectroscopic patterns of the galaxies from different morphological types. These models may be used to classify a new set of galaxies. There are several features in the spectra that can be detected by suitable pattern recognition tools, namely the break at 4000 \AA , strong lines in emission (Em) or absorption (Abs), and the overall pattern of the models.

Artificial neural networks (ANNs) are the most powerful topics in pattern-recognition problems. They have been the subjects of an explosion of interest in many different scientific areas. They provide us with the statistical information about a data set in order to categorize the data in a more quantitative way than a simple χ^2 minimization problem. ANNs can learn to detect regularities, correlations, and specific patterns in a certain set of data and so they can be considered to be a powerful tool. The applications of ANNs in astronomy include star–galaxy discrimination and galaxy classification (Cortiglioni et al. 2001; Andreon et al. 2001; Ball et al. 2004; Bazarghan et al. 2008), which also show that their importance has been increasing since last decade.

The morphological properties of nearby galaxies are well correlated with stellar mass, age, and specific SFR (N09). Here we use the CIGALE code in order to obtain these parameters for a sample of galaxies at $z = 0.5\text{--}1$, using photometric data in

the wavelength interval $\sim 0.4\text{--}24\ \mu\text{m}$. We wish to derive the best SEDs of the sample galaxies obtained from the fitting procedure and then classify them based on K96 models, using ANN in a supervised manner. In Section 2, we present the sample. In Section 3 we describe the K96 models and some points about the CIGALE code are presented in Section 4. In Section 5 we describe our method and in Section 6 we test the method in different ways. In Section 7 we classify the best SED of the sample galaxies. We investigate the correlation between the morphology of the classified SED and the physical parameters of mass, age, and SFR in Section 8. A short summary is presented in Section 9.

Throughout the paper we assume a cosmology with Ω_{tot} , Ω_M , $\Omega_\Lambda = 1.0, 0.3, 0.7$ and $H_0 = 70\ \text{km s}^{-1}\ \text{Mpc}^{-1}$.

2. SAMPLE

Our sample is selected from the spectroscopic campaign of the ESO GOODS-South field at $0.5 \leq z \leq 1$. The spectroscopic data are obtained with the FORS2 spectrograph at the ESO Very Large Telescope (VLT; Vanzella et al. 2005, 2006, 2008). The spectral coverage of the spectrograph is 6000–10000 Å. Redshift estimates carry one of four quality flags, “A”, “B”, “C”, and “X”, indicating secure, probable, tentative, and no measurement possible, respectively. For our sample, we choose galaxies with quality flags “A” and “B” only. The rest-frame spectra of the sample all include spectral information $\sim 3500\text{--}5100\ \text{\AA}$. This interval brackets the important features such as the 4000 Å break and also the strong emission lines such as $[\text{O II}]\lambda 3727$, $\lambda H_\beta 4861$, and $[\text{O III}]\lambda\lambda 4959, 5007$. The spectroscopic data for the sample galaxies contain different spectral types from starbursts (with a flatter continuum and stronger emission lines) to early-type galaxies (with a redder continuum, strong absorption lines, and large 4000 Å break). There are 190 spectra of 176 galaxies in the spectroscopic sample.

We use the point-spread-function matched photometric data from the most recently updated GOODS-MUSIC catalog provided by Santini et al. (2009). Measurements in filters in the wavelength range $\sim [0.4\text{--}24]\ \mu\text{m}$ of the catalog are used in this paper. They are U38 (VLT/VIMOS); B , V , i , z (*Hubble Space Telescope*/ACS); J , H , K_s (VLT/ISAAC), channels 1–4 from *Spitzer*/IRAC, and $24\ \mu\text{m}$ from *Spitzer*/MIPS.

At the redshift considered here, the Advanced Camera for Surveys (ACS) data bracket the 4000 Å break that is an important factor for deriving the effective age of the stellar population. The data from the ISAAC, IRAC, and MIPS instruments are also needed for obtaining robust results for the stellar mass and SFR. Incomplete spectral coverage for some galaxies forces us to limit our sample to those that have flux information in the ACS, ISAAC, and IRAC instruments at the same time. A total of 142 galaxies of the spectroscopic sample satisfy this condition and 71 of these also have secure flux information from the MIPS filter.

3. K96 MODELS

We use the K96 models as the target sample for the network training. The sample galaxies are therefore classified in terms of these models. They are the co-added spectra from 70 galaxies ranging from ellipticals to starbursts. They consist of Elliptical (E), Bulge (B), S0, Sa, Sb, Sc, and six starburst galaxies (SB1–SB6). The starburst patterns in K96 are labeled from 1 to 6 on the basis of their low to high internal extinction with $(B - V)$ ranging from 0.10 to 0.70 in steps of 0.1. The spectra

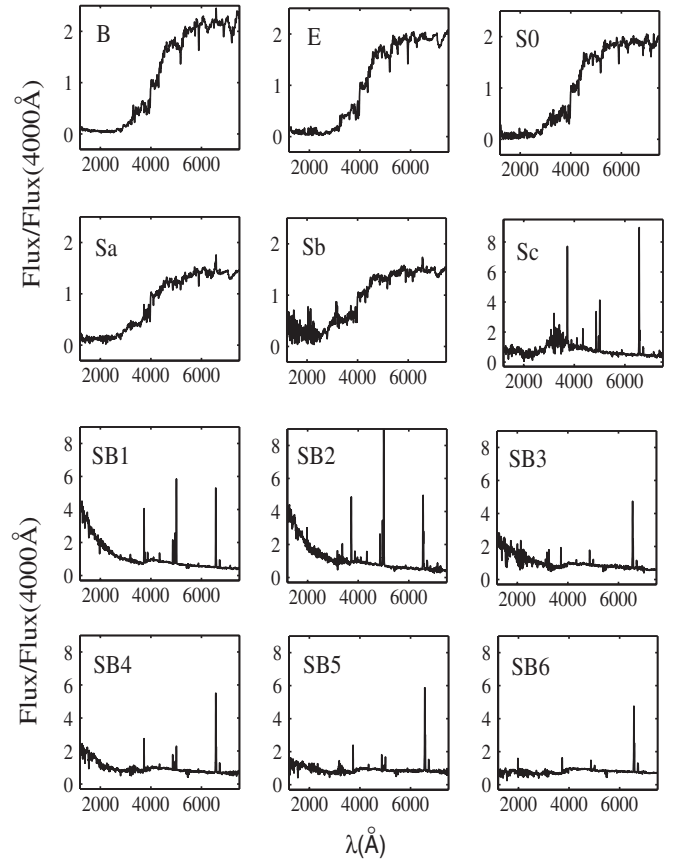


Figure 1. K96 models ranging from elliptical to starburst galaxies. The early-type galaxies (E, B, S0, Sa, Sb) show redder spectra with respect to the starburst ones. The starburst galaxies and the Sc galaxy show strong emission lines.

cover the interval $\sim 1200\text{--}10000\ \text{\AA}$ and have a resolution of $\sim 10\ \text{\AA}$. This interval samples a wide range of stellar populations from star-forming to early-type galaxies. All the spectra have complete flux information between ~ 1200 and $7500\ \text{\AA}$ and are distinguishable in the wavelength interval. Figure 1 shows these models. The early-type galaxies (E, B, S0, Sa, Sb) show redder spectra with respect to the starburst ones and the Sc galaxy also shows strong emission lines (see K96 and references therein for details).

4. THE CIGALE CODE

Briefly, CIGALE uses two combined CSPs to generate different models. In other words, it can consider a fraction, f_{burst} , of young star-forming populations on top of an older passive one. The code uses a modified version of Calzetti’s law (Calzetti et al. 2000) in a complex way to describe the attenuation in the model galaxies. CIGALE is also tested by selecting a series of input parameters to investigate the physical parameters of 39 nearby galaxies selected from SINGS. It is the first application of the code performed by N09. This sample contains galaxies with different morphological characteristics, ranging from young irregular star-forming systems to old and quiescent early-type galaxies. N09 evaluate the reliability of the selected parameters by studying the resulting expectation values and their errors in relation to the input model grid. In order to derive the physical properties, such as mass, age, and SFR of our sample galaxies, we use an exponentially decreasing SFR, τ model. The similar parameters to those listed in Table 3 of N09 are selected except for the age of the old population. N09 fix the

age of the old stellar population to 10 Gyr. Here, this parameter is selected so that it does not exceed the age of the universe at the selected redshift. We use a Salpeter initial mass function (Salpeter 1955) with solar metallicity (see N09 for details).

CIGALE uses two template files (emlin1 and emlin2) of the starburst galaxies to generate the emission lines in the SEDs. In the code, all the starburst models in the K96 library are used. Spectra 1–4 are taken for deriving the optical line template emlin2. Spectra 5 and 6, with relatively red UV SED, are used for the emlin1 template. The main criterion for the grouping is the SED slope in the UV and not the intensity of the optical emission lines.

5. METHOD

5.1. Artificial Neural Networks

ANNs are widely used in a large variety of applications in various areas. They are suitable tools for different purposes. For example, a fairly simple neural network can fit any practical function. In addition, clustering data and recognizing patterns are excellent applications for neural networks. They can be trained in supervised or unsupervised methods (e.g., Hagan et al. 1994; Kohonen 1997). In a supervised manner, a network can be trained to assign an unknown case to one desired class in the following steps.

Briefly, after collecting data (here, the sample and K96 models) a network is created and then configured. Configuration is the step in which the network is arranged to be compatible with the problem under study (i.e., classification problem). Initialization of the ANN parameters (weights and biases) and training the network (finding suitable weights and biases) are the next steps. The latter can be done by several algorithms. For example, in a gradient descent algorithm, and in an iterative way, the aim is to upgrade weights and biases in the direction where a performance function (e.g., a mean-squared error function) decreases most rapidly (i.e., in the direction of negative gradient). If the W_k are the current weights and biases and G is the current gradient, then the algorithm may be written as

$$W_{k+1} = W_k - \alpha_k G_k, \quad (1)$$

where α_k is the learning rate which is an adjustable parameter. It controls the sizes of weight and bias changes during learning. Finally, after validation of the network, it will be ready to use (or test) for new data.

There are several kinds of neural networks which can be configured and trained for solving different problems. Multi-layer feed-forward neural networks are among the most popular and are used in both forecasting and classification applications. For example, these networks can classify input data into C desired classes. In the simplest case, any problem requiring 1/0 (or yes/no) classification is a binary classification application where $C = 2$. In this case, one can associate an output value of zero to one of the two desired classes, and a value of one to the other. The network output will be the estimated probability that the input pattern belongs to one of the two categories (Hopfield 1987; Bishop 1995). This approach can be extended to $C > 2$.

5.2. Defining the Problem

In this work $C = 12$, which is the number of K96 models. In the training step, the normalized flux of the spectral models becomes the network input. The target is made as follows.

We consider the K96 models from bulge to starburst-6 and associate an output value of one with a selected K96 model and

Table 1
The Network Parameters

Parameter	Description
Number of layers	2
Activation function	tansig
Performance function	mse
Normalization method	mapminmax [−1 1]
Training algorithm	Levenberg–Marquardt

a value of zero with the rest. For example, we associate the output (1 0 0 0 0 0 0 0 0 0 0 0) with the bulge galaxy and we associate the output (0 0 0 0 0 0 0 0 0 0 1) with starburst-6. As a result, the network's target will be a 12×12 diagonal matrix (target matrix). The diagonal members of this matrix are one and the non-diagonal members are zero. With these input and output data, one can train a network. This trained network can classify a new data set into 12 classes with a certain probability.

If we have N spectra of galaxies for classification, the output of a trained network will be a matrix probability, P_{ij} , where $i = 1$ to N and $j = 1$ to C . This is the estimated probability that the i th galaxy belongs to the j th class. In this paper, the aim is to determine P_{ij} , where $i = 1$ to $N = 142$ (the number of sample galaxies) and $j = 1$ to $C = 12$.

To find the P_{ij} , we use the MATLAB network toolbox and create a two-layer feed-forward neural network. Table 1 shows the parameters used in the network. These are the default parameters in MATLAB that are used for pattern recognition problems (see the ANN user's guide of MATLAB and references therein).

6. TESTING THE TRAINED NETWORKS

Once a network is trained, we can apply it to a new sample with N members. A well-trained network should be able to distinguish all of the K96 models. So, if we select one member of the K96 models as input data (for example, the E model), then we should find a high value of P for this galaxy. This can be done for all members to test the trained network in terms of performance. To avoid any bias and to obtain a statistical result, we repeat the training procedure many times and save all the outputs. Then we compute the average value of these outputs, $\hat{P}_{E,j}$, and the associated errors.

To do this, we train 100 networks and find the $\hat{P}_{E,j}$ that is shown as the vertical axis of the top panel of Figure 2 (hereafter the probability diagram). The horizontal axis is the target class ($j = 1$ –12) from B (Bulge) to SB6 (starburst-6). The top graph shows that the input galaxy is an E-type galaxy ($j = 2$). The probability of it being an E galaxy is high ($\hat{P}_{E,2} \sim 90\%$), whereas all other probabilities are negligible. The bottom panel shows the same situation for another K96 model, i.e., the SB1. All K96 models show similar behavior.

There are many different features in the K96 models, making them distinguishable from one another. By visual inspection, one can find from Figure 1 that the most prominent features are the overall pattern (blue and red stellar continuum of the starburst and early-type galaxies, respectively), the 4000 Å break, the effect of dust on the spectra of starburst galaxies specially at $\lambda < 3000$ Å, and the presence of strong emission lines in the spectra of the starburst galaxies.

The strong spectral lines can change the flux average in a filter and this may present an alternative physical interpretation for the SED of high-redshift galaxies (Raiter et al. 2010). This effect can

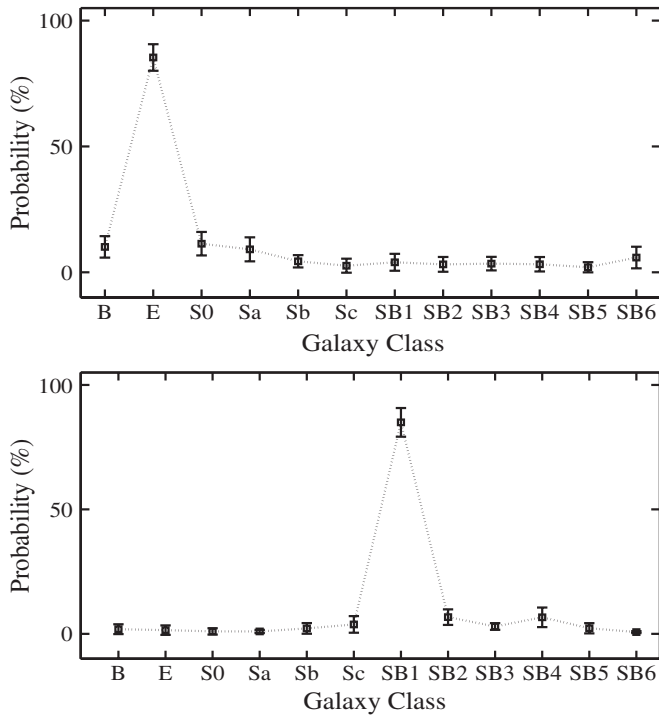


Figure 2. Probability diagrams, \hat{P}_{ij} , for two **K96** models. The horizontal axis is the target class ($j = 1-12$) from B (Bulge) to SB6 (starburst-6). These plots show that the two input data to the networks are correctly recognized as the E (top panel) and SB1 (bottom panel) galaxies.

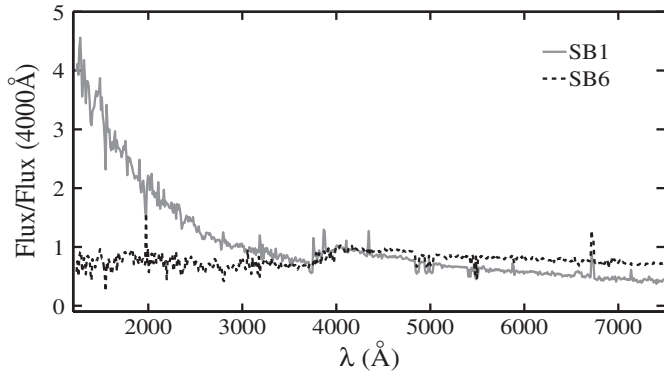


Figure 3. Two examples of the **K96** models, i.e., SB1 (solid gray line) and SB6 (black dashed line). Their strong emission lines ([O II], [O III], H_α , and H_β) are masked.

be more noticeable for the broadband filters located in a suitable wavelength interval. For example, this effect can be significant for a filter with $\Delta\lambda \sim 500 \text{ Å}$ located in the wavelength range 4959–5050 Å, where the strong emission lines [O III] $\lambda\lambda 4959, 5007$ are present.

Here, we show the effect of emission lines on the output of trained networks. To this end, we mask the prominent emission lines of [O II], [O III], H_α , and H_β in the **K96** models. Figure 3 shows two examples of the masked models, i.e., the SB1 (solid gray line) and the SB6 (dashed black line).

In this case, the probability diagrams show that starburst models with higher dust content have the potential to be confused with the Sb model. Figure 4 shows the probability diagrams for the two models, SB1 (top) and SB6 (bottom). Even after masking the lines, the SB1 model is significantly distinguishable. This is also clear from Figure 3. In this model,

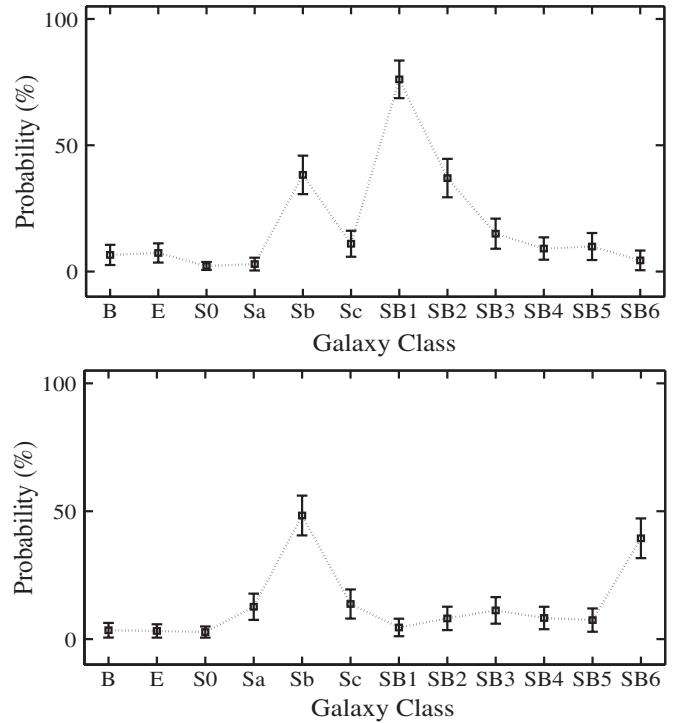


Figure 4. Probability diagram for two models of **K96** when the strong emission lines are masked: the SB1 model (top) and the SB6 model (bottom).

the bluer continuum at $\lambda < 3000 \text{ Å}$ makes it distinguishable from SB6. This shows that the trained networks can combine different features at the same time. By masking the lines in this model, the associated probability drops from $\sim 90\%$ (see Figure 2) to $\sim 76\%$. This drop appears gradually as we mask the emission lines one by one. In the bottom panel and after the masking, the SB6 model is not significantly distinguishable from the Sb model.

6.1. Testing the Networks with **CWW** Spectra

During the training step, one important problem that can occur is the overfitting issue. In this problem, the error on the training set is small but the error becomes large when the network is applied to a new data set. In other words, the network can recognize the training examples, i.e., the **K96** models, but it has not learned to generalize to new situations. To avoid this problem, MATLAB uses a method which is called the early stopping technique. To apply this, we separate the sample into three data sets: training (70%), validating (15%), and testing (15%). The training set is used for computing the gradient and updating the network weights and biases. A well-trained network should be able to validate the class membership of the known objects in the validation set. The validating set is also used to stop the training procedure before overfitting. During the initial phase of the training step, both the errors of the validation and the training sets decrease normally. When the data are being overfitted, the error on the validation set typically begins to rise. In this step, the training procedure stops updating the weights and biases. The test set is not used during the training procedure. By comparing minimum errors of the validation and testing sets, an unsuitable data division can be recognized.

The overfitting problem is checked in the training procedure. Nevertheless, we also check the trained networks using the well-known spectra of **CWW**. These are the four spectra of the

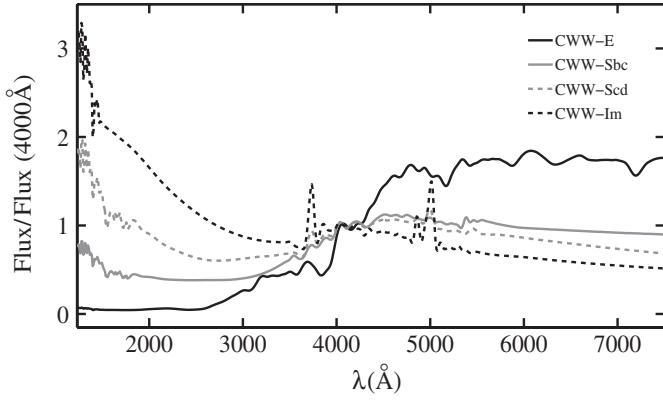


Figure 5. Four normalized spectra of *CWW* galaxies.

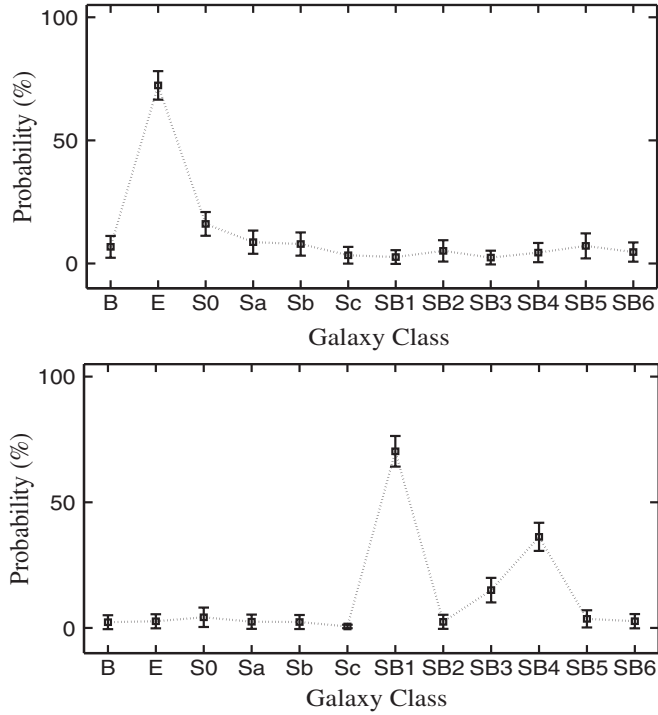


Figure 6. Probability diagrams for the E-type (top) and Im (bottom) of *CWW* spectra. They are recognized as the E and SB1 galaxies of the *K96* models.

E-type, Sbc, Scd, and Im galaxies that are shown in Figure 5. We present *CWW* spectra to the networks and estimate the associated \hat{P}_{ij} , ($i = 1-4$). Figure 6 shows the probability diagrams for the two E-type and Im galaxies in the top and bottom panels, respectively. The E-type and Im of *CWW* spectra are recognized as the E-type and SB1 of the *K96* models. They are significantly distinguishable but with lower associated probabilities when compared to *K96* models (see Figure 2) because the spectra have a lower resolution with respect to the *K96* models and the Im galaxy shows (only) two strong emission lines at $\lambda = 3727, 5007 \text{ \AA}$. Figure 7 shows the associated spectra in which the dashed black and solid gray lines represent the *K96* models and *CWW* spectra, respectively.

The probability diagram for the Sbc of *CWW* spectra is illustrated in the top panel of Figure 8. This shows that this galaxy can be the Sa of *K96* models. Another high-probability candidate is an SB6 model that is a dusty starburst candidate. They are not significantly different. The bottom panel of this figure shows that the Scd galaxy of *CWW* is marginally

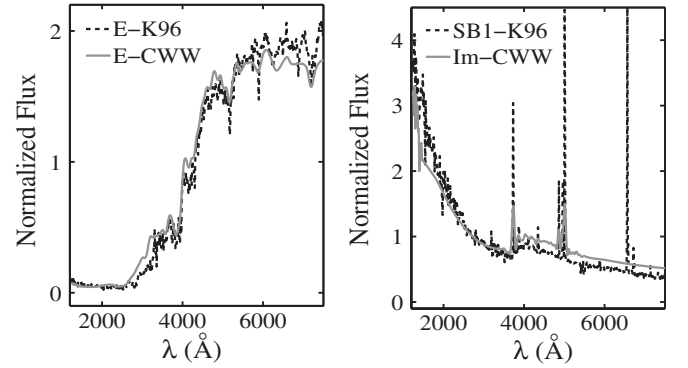


Figure 7. Spectra of the E-type and Im galaxies of *CWW* (solid gray lines) and E-type and SB1 of the *K96* models (dashed black lines).

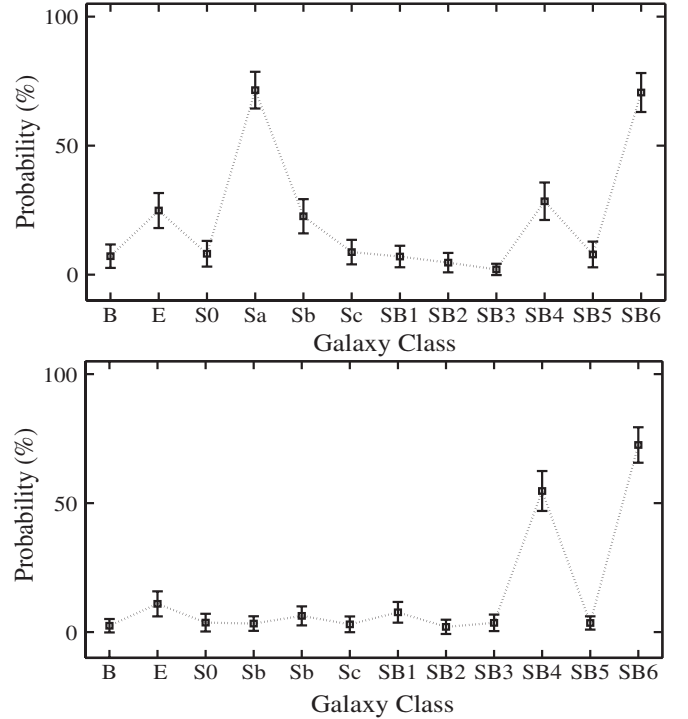


Figure 8. Probability diagrams for Sbc (top) and Scd (bottom) galaxies of *CWW* templates.

recognized as an SB6 galaxy. The second candidate is the SB4 model that is also a star-forming galaxy. The blue continuum of this Scd galaxy in Figure 5 at $\lambda < 3000 \text{ \AA}$ also suggests star formation activity for this galaxy. These spectra, along with the mentioned *K96* models, are drawn in Figure 9. The solid gray lines show *CWW* spectra, which are located between the two spectra of the *K96* models. For example, the left panel shows two uncertain spectra of *K96* models, i.e., the Sb and SB6 models. The Sbc spectrum of *CWW* is located between these two spectra. In summary, according to the probability diagrams obtained for the four *CWW* spectra, two of them can be classified in terms of *K96* models with high significance. The Sbc and Scd galaxies of *CWW* are not significantly recognized as single *K96* models. We will return to the uncertain cases in the following section.

7. CLASSIFICATION OF THE BEST SEDs

Finally, we apply the trained networks on the best SED obtained from the SED fitting procedure, in the wavelength

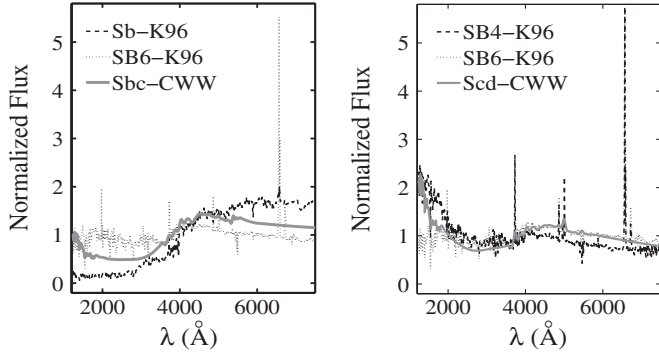


Figure 9. Spectra of the Sbc and Scd of CWW galaxies (the solid gray line). None of them is significantly recognized as a K96 model. They are located between two K96 models (dotted and dashed black lines).

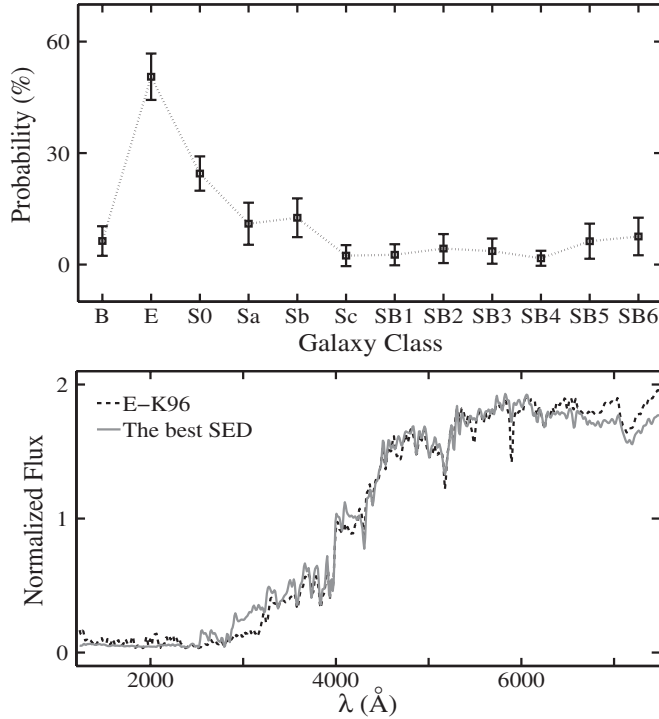


Figure 10. Top panel shows the probability diagram of the best SED for a galaxy recognized as an E-type galaxy. The bottom panel shows the best SED of this galaxy and the K96 model. The dashed black lines represent the best K96 model matched to the best SED of the sample galaxy (the solid gray lines).

interval 1200–7500 Å. In this way, one can find the associated probability diagram for each SED. For example, the top plot of Figure 10 shows the probability diagram for the best SED of a galaxy in the sample. Within this plot, it can be seen that networks return the highest probability for the E-type galaxy. The spectra of the early-type galaxies are more similar to each other. There is another candidate S0 with a lower probability. Therefore, we choose the E as a classified galaxy and draw the best SED of this galaxy (the solid gray line) in the bottom plot of Figure 10. The dashed black line represents the E-type model of the K96. The correlation between the two spectra is ~ 0.98 . Figure 11 shows the situation for another galaxy of the sample: the galaxy with a ~ 0.99 correlation with a starburst-1 galaxy of the K96 models.

Not all of the best SEDs of the 142 galaxies are significantly recognized as a K96 single model. As an example, we show

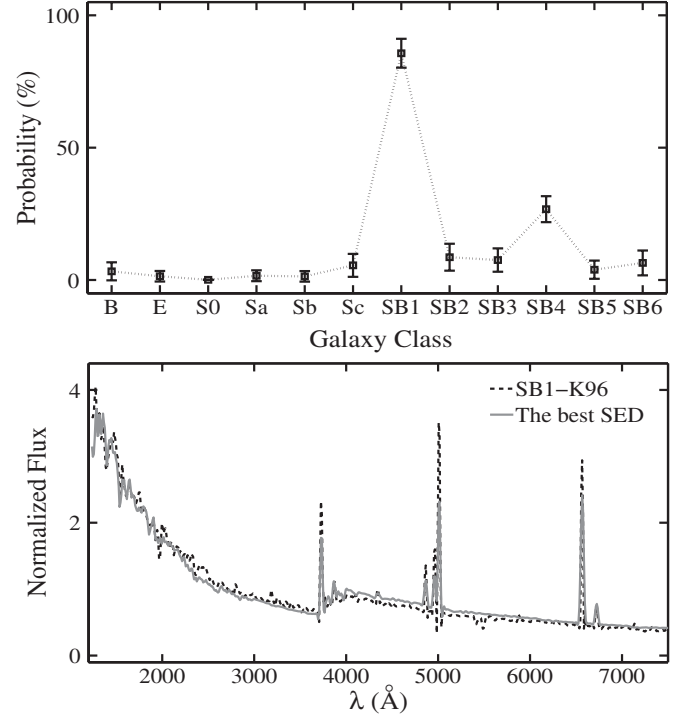


Figure 11. Same as Figure 10 for an SB1 galaxy.

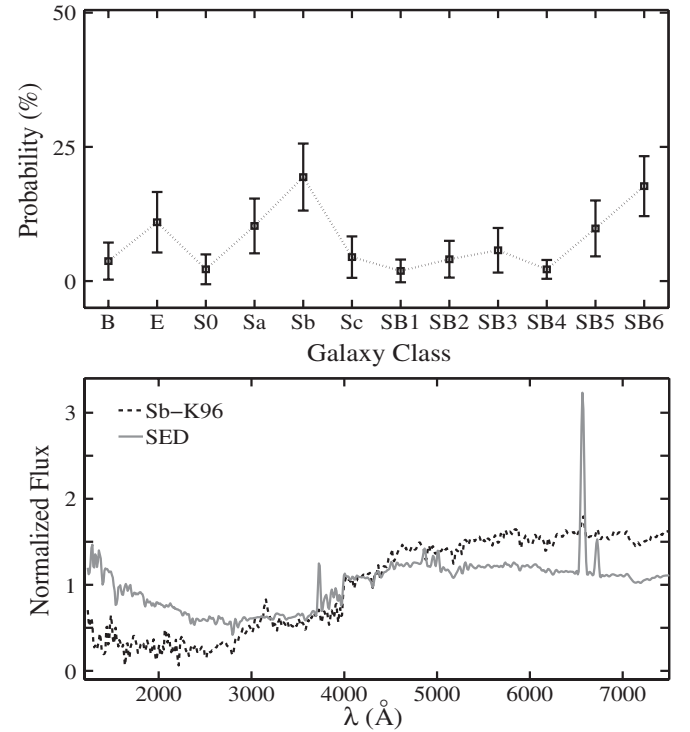


Figure 12. Top plot: the probability diagram of the best SED of a sample galaxy. This does not present a significant classification for the galaxy. Bottom plot: the best K96 model (an Sb galaxy, in dashed black line) and the best SED (the solid gray line). They do not show a well-correlated case.

the probability diagram of a galaxy in the top plot of Figure 12. This is the worst case in the sample in which the probabilities are distributed almost equally between all the models. The highest probability belongs to the Sb model, but it is not significantly different, for example, from an SB6. We show the highest probability candidate, the Sb model, as the dashed black line

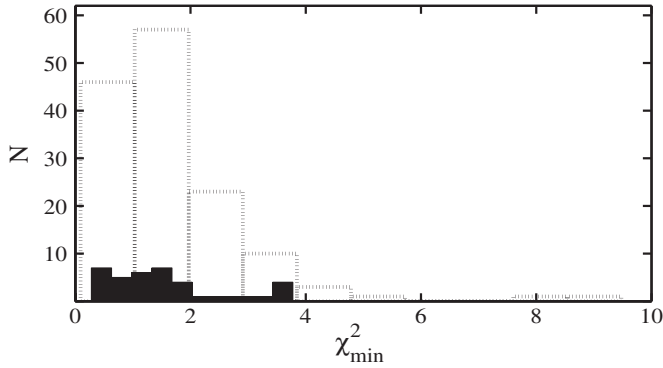


Figure 13. Distribution of χ_{\min}^2 . The dotted line is related to all galaxies. The filled histogram is the value of χ_{\min}^2 for the unclassified galaxies.

in the bottom plot. The solid gray line is the best SED. They do not show a well-correlated case. In this work, we choose the cases in which the best candidates are significantly different from the rest. In this way, the uncertain models do not appear in the final classified result. Out of 142 galaxies, 37 have SEDs that cannot be classified based on a specific K96 model.

Searching for the reasons for the unclassified cases, we compare the distribution of χ_{\min}^2 of the unclassified galaxies with the total distribution. They do not show a different χ_{\min}^2 distribution for the SED fitting. The dotted histogram of Figure 13 shows the distribution of χ_{\min}^2 of all galaxies. The filled black histogram shows this value for the unclassified SEDs. The mean value of χ_{\min}^2 of all galaxies is ~ 1.7 . This value is smaller for those galaxies that do not have photometric data in the MIPS filter. These galaxies have a mean value of $\chi_{\min}^2 \sim 1.14$. The MIPS data suffer a greater uncertainty when compared, for example, with ACS data. Nevertheless, they can reduce the degeneracy problem in the SED fitting and consequently can provide more reliable physical parameters such as mass. The mean value for the 71 galaxies detected in MIPS is ~ 2.15 . Despite these differences in χ_{\min}^2 , a substantial fraction of MIPS detections have been classified well.

The K96 models are only 12 co-added spectra of 70 galaxies and it is possible that some of the best SEDs cannot be significantly matched to a single model. As an example, we find that the Sc model, which is the most distinguishable one among the K96 models ($P_{\text{Sc}} \sim 96\%$), has not been significantly matched to any of the best SEDs. This model at $\lambda > 4000 \text{ \AA}$ is similar to the starburst models with the strongest emission lines and at $\lambda < 4000 \text{ \AA}$ the spectrum of this galaxy has a unique arc-shaped feature. The best SED may not be matched to these features at the same time. The best SEDs obtained from CIGALE are the combined spectra of the young and old populations. They are also modified by different effects, such as dust and the emission lines. For Figure 12, one scenario might be that this best SED can be a combination of models. To test this, we make a combined weighted spectrum of the K96 models based on the probability diagram of Figure 12. The \hat{P}_{ij} , ($j = 1-12$) are the weights. In this way, each K96 model contributes to the combined spectrum according to its weight. We construct the composite spectrum and show the result in Figure 14. The dashed black line is the composite spectrum of the K96 models and the solid gray line is the best SED. This shows that the values within the probability diagram in Figure 12 generate a suitable combined spectrum of K96 models, matched to the best SED.

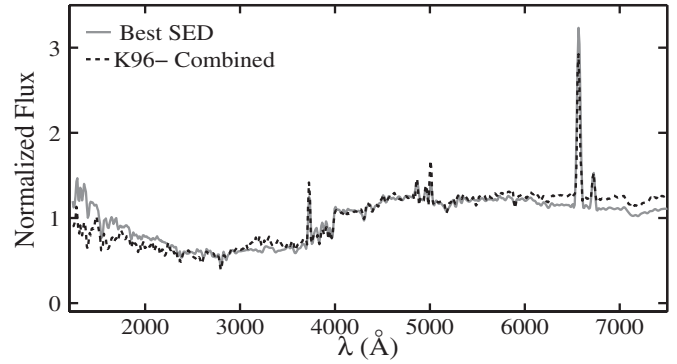


Figure 14. Dashed black line shows a composite spectrum of the K96 models making from the probability values of an unclassified galaxy. The solid gray line represents the best SED of the galaxy.

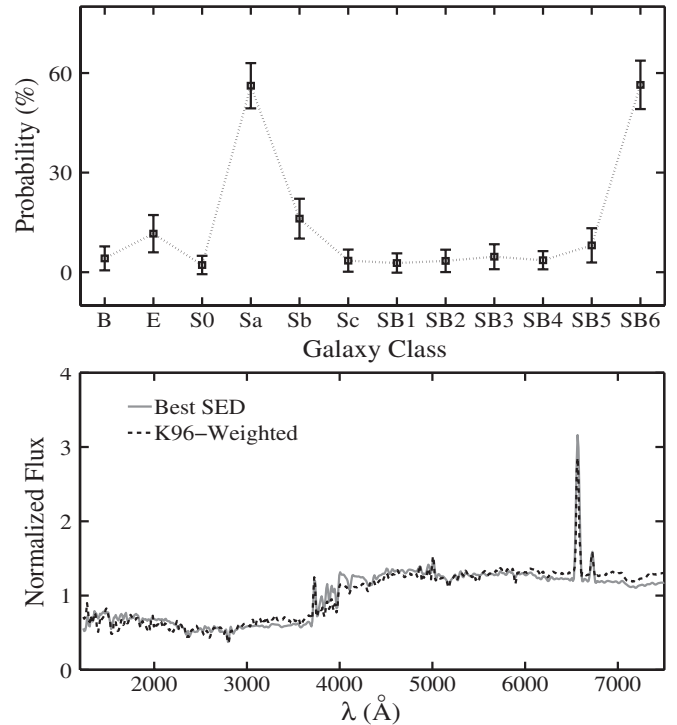


Figure 15. Another example of an unclassified galaxy. The top graph shows the probability diagram for the best SED of the galaxy. It does not introduce a significant, single candidate model for this galaxy. In the bottom graph, the dashed black line shows the combined weighted spectrum from the K96 models. The solid gray line represents the best SED of the unclassified galaxy.

The top graph of Figure 15 shows another example of an unclassified galaxy. This is similar to the probability diagram obtained for the Sbc of the CWW galaxies in Figure 8. The bottom graph of Figure 15 shows the combined weighted spectrum of the K96 models, matched to the SED. This shows that all of the features of the best SED are considered by the networks at the same time.

To summarize of our approach, the K96 models may be considered to be the components of a 12 dimensional space. For the best SED of a sample galaxy, the \hat{P}_{ij} , estimated by the trained networks, can produce a vector in this space. This vector can be expanded in terms of the components, i.e., the K96 models. The study of these vectors as functions of redshift or environment is left as an open question. In this work, the most significant component of a vector is considered to be a classified galaxy. In the following section we investigate the

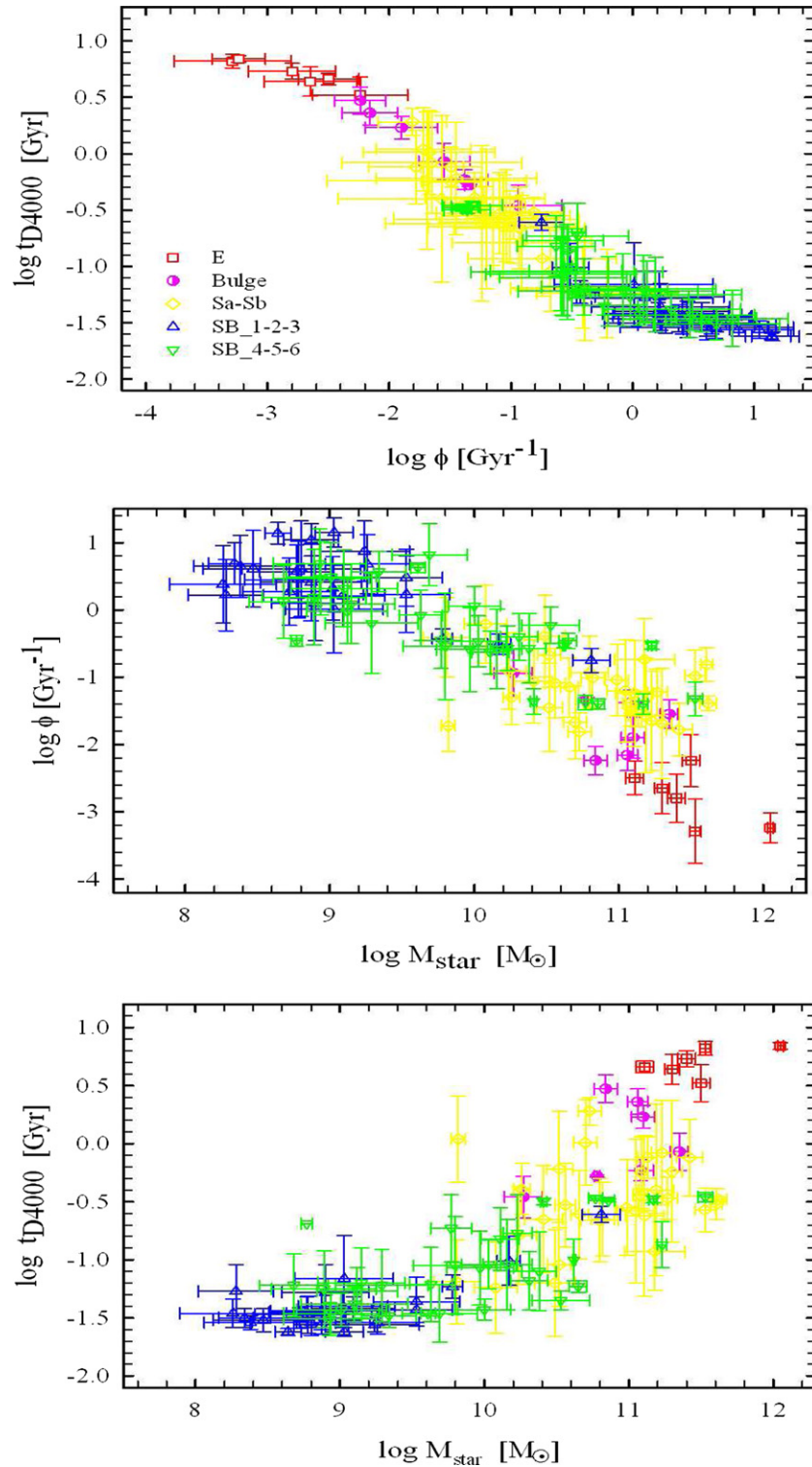


Figure 16. Mutual correlation between the three parameters ϕ (specific SFR), t_{D4000} , and the stellar mass of the galaxies. The different symbols show the different classified galaxies.

(A color version of this figure is available in the online journal.)

relation between the morphology of the classified galaxies and three physical parameters obtained from the SED fitting.

8. RESULTS

Not all the parameters extracted from CIGALE are independent of each other, for example, mass and SFR or f_{burst} and t_{D4000} .

For the three parameters of specific SFR, ϕ , the stellar mass, and t_{D4000} , the models do not show a direct relation (N09). The t_{D4000} is the age derived from the D4000 break (Balogh et al. 1999) and ϕ is the SFR of a galaxy divided by the total stellar mass.

There is a strong correlation between mass and the total SFR. It is much easier to use ϕ to study the relationship between star formation activity and the physical parameters of the galaxies.

The variation of ϕ with stellar mass is a consequence of the so-called downsizing scenario. In this scenario the early universe was dominated by a relatively small number of giant galaxies. The activity in the current universe is distributed among a large number of smaller galaxies. In this respect, Brinchmann et al. (2004) present a comprehensive study of the physical properties of $\sim 10^5$ galaxies and build a picture of the nature of star-forming galaxies at $z < 0.2$. Considering ϕ , they show the correlation between the present to past average star formation rates, the morphological type, and the 4000 Å break. One conclusion of their work is that a considerable amount of star formation takes place in low-mass galaxies.

Following N09, we draw three diagrams of the mutual correlation between the mentioned parameters. Figure 16 shows this situation. In comparison with the results obtained by N09 there is also a tighter correlation between the quantities ϕ and t_{D4000} with respect to the mass-related graphs. According to the top plot of this figure, the younger stellar populations show a higher star formation activity. The mass- ϕ graph shows that the lower mass galaxies are more active ones and also shows that the specific SFR decreases with increasing mass. Karim et al. (2011) show that the specific SFR increases with redshift at a rate independent of mass. This increment is also seen in our mass- ϕ graph when it is compared with the results obtained by Brinchmann et al. (2004). Bauer et al. (2005) and Damen et al. (2009) also investigate the relationships between ϕ and galaxy stellar mass at $0 < z < 1.5$ and $0 < z < 1.8$, respectively. They find that ϕ decreases with increasing galaxy stellar mass and that massive galaxies have substantially lower ϕ compared to lower mass galaxies. The bottom graph shows that the age derived from the D4000 break increases with increasing mass and that this trend is more prominent for stellar mass $> 10^{10} M_{\odot}$. Buat et al. (2011), using CIGALE, study a sample of 363 galaxies in the nearby universe. Their galaxy sample is mostly constituted of massive, actively star-forming galaxies. Nevertheless, this trend can also be seen in their D4000 mass diagram.

In the three plots of Figure 16, the loci of the classified galaxies also correlate well with the morphological characteristics. This is in agreement with the result obtained by N09. For example, E-type and starburst galaxies are the most and the least massive galaxies, respectively. Sa and Sb galaxies occupy the intermediate locations and we put them into a group. We roughly divide the starburst into two groups, SB1–3 and SB4–6, according to their dust content. There is the least amount of overlap between the Sa–Sb and the starburst 1–3 galaxies because the latter galaxies contain less dust and their overall patterns are distinguishably different from the former ones (see Section 6). The dusty starbursts (4–6) have the potential to be confused with the Sa–Sb galaxies. This potential is also seen in their probability diagrams. The bulge models in K96 are two Sb galaxies that are dominated by the stellar population in the bulge (K96). This group is approximately located between the E and Sa–Sb groups.

The co-added spectra of a certain class can show the average characteristic of the classified data (Shapley et al. 2003; Balestra et al. 2010). To make use of available observational data, we construct the co-added spectra of the FORS2 spectroscopic data of the groups shown by different symbols in Figure 16. In this way, one can compare them with the associated spectra of the K96 models. All of the FORS2 spectra, at $0.5 < z < 1$, have the flux information in the wavelength interval ~ 3500 – 5100 Å and can contribute to the co-added spectra. They are plotted in this region.

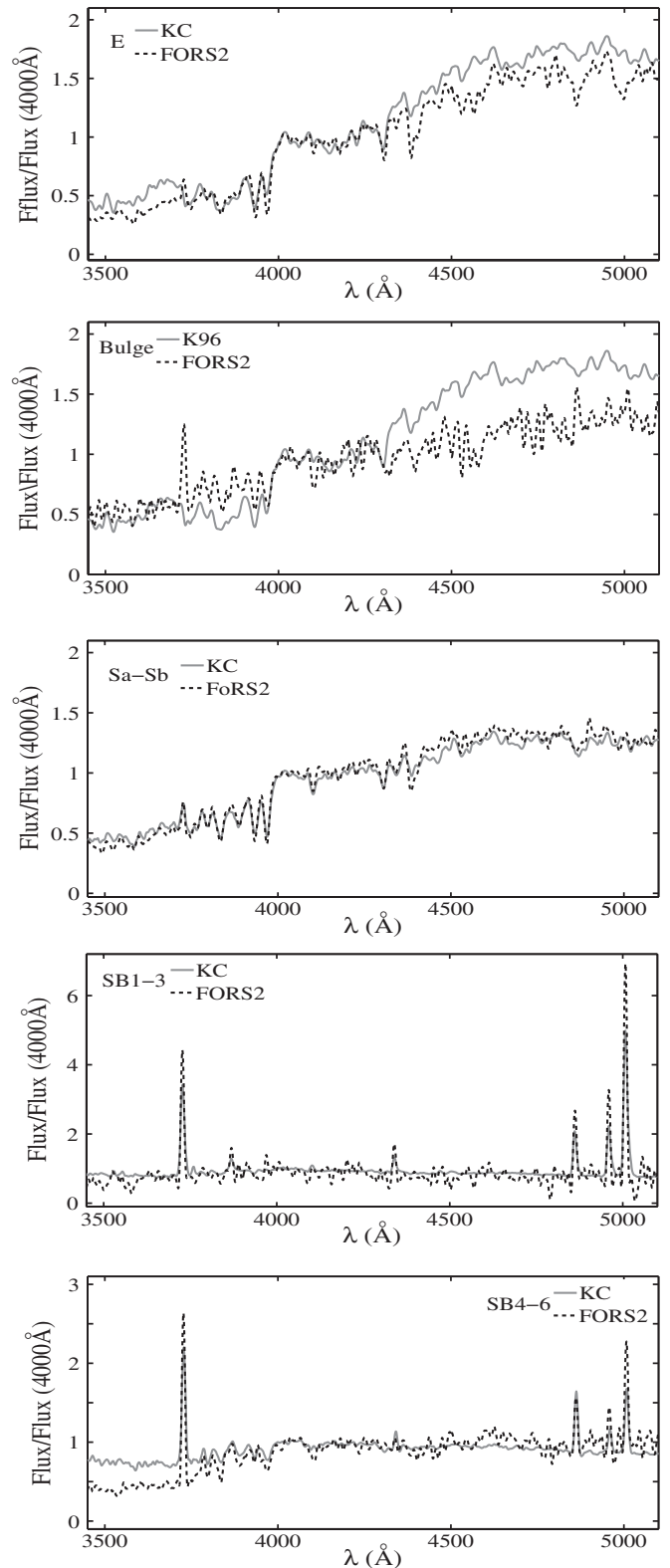


Figure 17. Co-added spectra of the FORS2 spectroscopic data (the dashed black line) and the associated K96 models (the solid gray line). Within the graphs, the groups are denoted as E-type, Bulge, Sa–Sb, SB1–3, and SB4–6.

In Figure 17, we draw the co-added spectra of the FORS2 spectroscopic data (the dashed black line) and the associated K96 models (the solid gray line). Within the graphs, the groups are denoted by the E-type, Bulge, Sa–Sb, SB1–3, and SB4–6.

The number of the galaxies contributing to the co-added spectra of the Bulge and E galaxies is small. This affects the S/N of the co-added spectra. A redder continuum of the co-added spectrum of E-type galaxy is distinguishable when it is compared to the rest of the co-added spectra. In E-type galaxies, the 4000 Å break and the absorption lines in this region are very compatible. The co-added spectrum of the Bulge candidates also shows a red continuum, but it does not show a very compatible case with the Bulge galaxy of K96. This spectrum has the lowest S/N among the co-added spectra.

The Sa–Sb group is in excellent agreement with that of the K9 models. The situation is the same for the starburst galaxies (1–3). The FORS2 data in this group show the stronger emission lines ([O II] ($\lambda 3727$), H β ($\lambda 4861$), and [O III] ($\lambda \lambda 4959, 5007$)) in contrast to the K96 models. In the SB4–6, the FORS2 data and the models show a well-matched case. The FORS2 spectra of this group show a redder continuum at $\lambda < 3850$ Å in contrast to the models.

9. SUMMARY

Fitting the SED of galaxies is a technique that is widely used to estimate the physical parameters of galaxies in the local and the high-redshift universe. We use the CIGALE code to find parameters such as mass, specific SFR, and age of a sample of galaxies at $0.5 < z < 1$ in CDFS. We also extract the best SED obtained from χ^2_{\min} of the SED fitting. We classify the best SED using a neural network approach in a supervised manner in which K96 models are used as target classes. These models contain 12 spectra ranging from starburst to early-type galaxies. When using galaxy spectra as input data, the output of the network presents 12 probabilities, which can determine the class of the galaxy.

The performance of the trained networks is found to be $\sim 90\%$. We check the effect of the prominent features such as the strong emission lines, the effect of dust on spectra, and the overall pattern of spectra on the performance of the networks. We also apply the trained networks to CWW spectra of E-, Sbc-, Scd-, and Im-type galaxies. They have lower resolution compared to K96 models and consequently their probability diagrams estimated by networks show lower values when compared with K96 models. In this way, we show that the trained networks can combine different features in the galaxy spectra at the same time.

After testing the trained networks, we apply them to the best SED of the sample galaxies and find that 105 out of 142 galaxies are significantly classified according to the models. The unclassified galaxies cannot be matched to a single spectrum of K96 models. We find that if we construct a combined weighted spectrum of the K96 models, then a best SED of an unclassified galaxy is matched to the composite model. In this approach, the K96 models may be considered to be the components of a 12 dimensional space and the best SED of a sample galaxy can be considered to be a vector in this space. This vector can be expanded in terms of the components. The study of these vectors as functions of redshift or environment is an open question. In this work, the most significant component of a vector is considered to be a classified galaxy.

The loci of the classified galaxies in the three graphs of the physical parameters of mass, age, and specific SFR show proper connections with the morphological characteristics of the galaxies. It should be noted that these morphological characteristics are indirectly extracted from the models and are not classified from their shapes in the images. Finally, to make use of available observational data, the co-added spectra of these data are compared with those of K96 models, which shows that they are in good agreement.

We limit our work to the K96 models, which are limited in turn to the wavelength interval of 1200–7500 Å. Today, far-IR, submillimeter, and radio data are available and one may find many full empirical SEDs. This may help to improve the desired classifications.

H.T. sincerely thanks R. A. E. Fosbury for reading the manuscript and for suggesting valuable comments. H.T. also thanks S. Noll for giving the valuable information about the CIGALE code. Observations have been carried out using the Very Large Telescope at the ESO Paranal Observatory under program IDs 170.A-0788, 074.A-0709, and 275.A-5060.

REFERENCES

- Andreon, S., Gargiulo, G., Longo, G., et al. 2000, *MNRAS*, **319**, 700
 Balestra, I., Mainieri, V., Popesso, P., et al. 2010, *A&A*, **512**, A12
 Ball, N. M., Loveday, J., Fukugita, M., et al. 2004, *MNRAS*, **348**, 1038
 Balogh, M. L., Morris, S. L., Yee, H. K. C., et al. 1999, *ApJ*, **527**, 54
 Bauer, A. E., Drory, N., Hill, G. J., et al. 2005, *ApJ*, **621**, 89
 Bazarghan, M., Safari, H., Innes, D. E., et al. 2008, *A&A*, **492**, 13
 Bishop, C. M. 1995, *Neural Networks for Pattern Recognition* (Oxford: Oxford University Press)
 Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, *MNRAS*, **351**, 1151
 Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
 Buat, V., Giovannoli, E., & Takeuchi, T. T. 2011, *A&A*, **529**, A22
 Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
 Coelho, P. 2009, in *AIP Conf. Proc.* 1111, *Probing Stellar Populations Out to the Distant Universe: CEFALU 2008*, ed. G. Giobbi et al. (Melville, NY: AIP), 67
 Coleman, G. D., Wu, C. C., & Weedman, D. W. 1980, *ApJS*, **43**, 393
 Cortiglioni, F., Mähönen, P., Hakala, P., et al. 2001, *ApJ*, **556**, 937
 Damen, M., Labbè, I., Franx, M., et al. 2009, *ApJ*, **690**, 937
 Gregg, M. D., Silva, D., Rayner, J., et al. 2004, *BAAS*, **36**, 1496
 Hagan, M. T., & Menhaj, M. 1994, *IEEE Trans. Neural Netw.*, **5**, 989
 Hopfield, J. J. 1987, *Proc. Natl Acad. Sci. USA*, **84**, 8429
 Karim, A., Schinnerer, E., & Martinez-Sansigre, A. 2011, *ApJ*, **730**, 61
 Kennicutt, R. C., Jr., Armus, L., Bendo, G., et al. 2003, *PASP*, **115**, 928
 Kinney, A. L., Calzetti, D., Bohlin, R. C., et al. 1996, *ApJ*, **467**, 38
 Kohonen, T. 1997, in *The Self-Organizing Map (SOM)*, ed. T. S. Huang & M. R. Schroeder (Springer Series in Information Sciences; New York: Springer)
 Lancon, A., Gallagher, J. S., III, Mouhcine, M., et al. 2008, *A&A*, **486**, 165
 Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *ApJS*, **123**, 3
 Maraston, C. 2005, *MNRAS*, **362**, 799
 Molla, M., Garcya-Vargas, M. L., Bressan, A., et al. 2009, *MNRAS*, **398**, 451
 Noll, S., Burgarella, D., & Giovannoli, E. 2009, *A&A*, **507**, 1793
 Noll, S., Mehlert, D., Appenzeller, I., et al. 2004, *A&A*, **418**, 885
 Raiter, A., Fosbury, R. A. E., & Teimoorinia, H. 2010, *A&A*, **510**, 109
 Salpeter, E. E. 1955, *ApJ*, **121**, 161
 Santini, P., Fontana, A., & Grazian, A. 2009, *A&A*, **504**, 751
 Shapley, A. E., Steidel, C. C., Pettini, M., et al. 2003, *ApJ*, **588**, 65
 Vanzella, E., Cristiani, S., Dickinson, M., et al. 2005, *A&A*, **434**, 53
 Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, *A&A*, **454**, 423
 Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, **478**, 83