

Estimating Photometric Redshifts with Artificial Neural Networks and Multi-Parameters *

Li-Li Li^{1,2}, Yan-Xia Zhang¹, Yong-Heng Zhao¹ and Da-Wei Yang²

¹ National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012; lily@lamost.org

² College of Physics Science and Information Engineering, Hebei Normal University, Shijiazhuang 050016

Received 2006 October 15; accepted 2006 December 25

Abstract We calculate photometric redshifts from the Sloan Digital Sky Survey Data Release 2 (SDSS DR2) Galaxy Sample using artificial neural networks (ANNs). Different input sets based on various parameters (e.g. magnitude, color index, flux information) are explored. Mainly, parameters from broadband photometry are utilized and their performances in redshift prediction are compared. While any parameter may be easily incorporated in the input, our results indicate that using the dereddened magnitudes often produces more accurate photometric redshifts than using the Petrosian magnitudes or model magnitudes as input, but the model magnitudes are superior to the Petrosian magnitudes. Also, better performance results when more effective parameters are used in the training set. The method is tested on a sample of 79 346 galaxies from the SDSS DR2. When using 19 parameters based on the dereddened magnitudes, the rms error in redshift estimation is $\sigma_z = 0.020184$. The ANN is highly competitive tool compared to the traditional template-fitting methods when a large and representative training set is available.

Key words: galaxies: fundamental parameters — techniques: photometric — method: data analysis

1 INTRODUCTION

Photometric redshifts refer to galaxy redshifts estimated using only medium-band or broad-band photometry or imaging rather than spectroscopy. It is a fact that broad band photometry is less time consuming than spectroscopy by orders of magnitude. Furthermore, photometry is available for faint galaxies that may not be spectroscopically accessible at all, because of the limited telescope time. In addition, the greater area of the sky covered by imaging detectors usually means that more redshifts of more objects are measured at a time by photometry than by spectroscopy, the latter being limited to individual galaxies or those happening to lie on the slits or fibres. The importance of the technique is growing not only with the desire to gain a greater understanding of galaxy evolution (for example, the determination of luminosity function), but also in connection with weak gravitational lensing, where redshift estimates can reduce contamination from intrinsic alignments (Heymans & Heavens 2003; King & Schneider 2003). If a method can be found that can give accurate estimates of redshifts for the larger photometric catalogs, then we shall have much better constraints on the formation and evolution of large-scale structural elements such as galaxy cluster, filaments, and cosmological models (e.g. Blake & Bridle 2005) in general. True, photometric redshifts have a relatively lower precision, but for many applications such as determining the properties of large numbers of galaxies and the large-scale structure of the universe, it is quite tolerable and sometimes even more effective.

* Supported by the National Natural Science Foundation of China.

The concept of photometric redshift was first developed by Baum (1962). Since then, many new methods have been applied to calibrate the relevant relations. To date, these methods have typically been employed on multicolor deep-field and wide-field surveys, notably the Hubble Deep Field (e.g. Gwyn & Hartwick 1996; Sawicki et al. 1997; Connolly et al. 1998; Fernández-Soto et al. 1999; Fontana et al. 2000; Vanzella et al. 2004; Coe et al. 2006) and the Sloan Digital Sky Survey (SDSS, Sowards-Emmerd et al. 2000; Casbai et al. 2003; Weinstein et al. 2004). The commonest way of estimating photometric redshifts is template-matching. This requires converting the photometric data of galaxies into spectral energy distributions (SEDs) and compiling a library of template spectra covering all the galaxy types, luminosities and redshifts in the range of interest. For a particular target galaxy, the redshift of the most closely matching template spectrum is selected to be its photometric redshift. This is usually defined by minimizing the χ^2 between the template and actual magnitudes. The template-matching photometric redshift technique makes use of the available and reasonably detailed knowledge of galaxy SEDs, and in principle it may be used reliably even for populations of galaxies with few or no spectroscopically confirmed redshifts. However, its success strongly depends on the compilation of a library of accurate and representative template SEDs (see e.g. Hogg et al. 1998). In the situation where there is a large amount of prior information on the redshifts of the sample, the template-matching technique is not the best approach.

An alternative approach is fitting by a polynomial or some other function, mapping the photometric data to the known redshifts and using this to estimate the redshifts for the remainder of the sample with unknown redshifts (e.g. Sowards-Emmerd et al. 1999). In essence, its aim is to derive a parametrization of redshift in terms of photometric parameters. This requires a large and representative training set of galaxies with both photometry and precisely known redshifts. A simple example is to express the redshift as a polynomial in the galaxy colours (Connolly et al. 1995; Sowards-Emmerd et al. 2000). The coefficients in the polynomial are varied to optimize the fit between the predicted and measured redshifts. The photometric redshift for a galaxy with unknown spectroscopy can then be estimated by means of the optimized function to the colours of the target galaxy.

In recent years, a variety of techniques to estimate photometric redshifts have emerged based on machine learning. Artificial neural networks (ANNs) as a new possibility among the interpolative techniques have been used in astronomy. Popular applications include star/galaxy separation (e.g. Odewahn & Nielsen 1994; Bertin & Arnouts 1996), morphological classification of galaxies (Nieversity & Odewahn 1994; Lahav et al. 1996; Ball et al. 2004), spectral classification (Folkes et al. 1996; Weaver 2000) and astronomical objects classification (Zhang & Zhao 2003, 2007). Certainly, ANNs have also been applied in photometric redshift prediction (Tagliaferri et al. 2002; Firth et al. 2003; Ball et al. 2004; Vanzella et al. 2004; Collister & Lahav 2004).

ANNs are applicable to ‘mixed’ data sets in which a moderately large training set of objects for which both photometry in the survey filters and spectroscopic redshifts are available. The ANN method is general in the sense that any parameter can be used to train the network and then make predictions. In practice, one can measure and use an almost limitless number of parameters to describe the galaxy. However, it is desirable to have as much information as possible with the fewest possible parameters (continuous or discrete). The chosen parameters should be physically meaningful, i.e. they should be directly predicted by the theories of galaxy and large scale structure formation, or be related in a quantitative way. Hence, it is necessary to find out what parameters are the most helpful and useful for photometric redshift evaluation.

This paper explores the use of ANNs as a potential tool for photometric redshift determination. We mainly focus on establishing the best set of input parameters and compare the effect of different input parameter sets on the redshift estimation. The layout of this paper is as follows. In Section 2 the principle of ANN is introduced. Section 3 describes the data in detail and the parameters used in the experiments. The procedure to estimate redshifts with different input parameter sets is presented in Section 4. In Section 5 the performances of different parameter sets are discussed. The conclusions are given in Section 6.

2 ARTIFICIAL NEURAL NETWORKS

ANNs, being originally conceived as models of the brain, are collections of interconnected neurons each capable of carrying out simple processing. Thus, they are composed of massively parallel distributed processors that have an inherent property of storing experiential knowledge and making it available for use. The

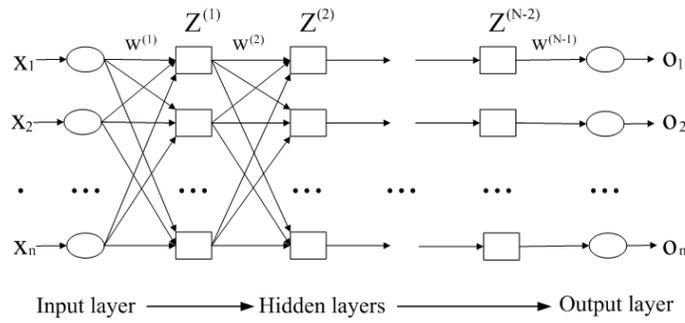


Fig. 1 Schematic diagram of artificial neural network, consisting of an input layer of input nodes taking, for example, the magnitudes in various bands, a middle hidden layer, and a single output node giving, e.g., a redshift z . Each connecting line has an assigned weight w_{ij} .

knowledge is acquired by the network through a learning process and is stored in interneuron connection strengths - known as synaptic weights (Haykin 1994).

Practical applications of ANNs most often employ supervised learning. For supervised learning, one must provide training data that includes both the input (a set of vectors of parameters, here each vector corresponds to a galaxy) and the desired result or the target value (the corresponding redshifts). After the network is trained successfully, one can present input data alone to the ANN (that is, input data without the desired result), and the ANN will compute an output value that approximates the desired result.

This is achieved by using a training algorithm to minimize the cost function which represents the difference (error) between the actual and desired output. The cost function E is commonly of the form

$$E = \frac{1}{p} \sum_{k=1}^p (o_k - t_k)^2, \quad (1)$$

where o_k and t_k are the output and target respectively for the objects, p is the sample size. Generally the topology of an ANN can be schematized as a set of N layers (see Fig. 1), with each layer composed of a number of neurons. The first layer ($i = 1$) is usually called the ‘input layer’, the intermediate ones, the ‘hidden layers’ and the last one ($i = N$) the ‘output layer’. Each neuron j in the s layer derives a weighted sum of the M output $z_i^{(s-1)}$ from the previous layer ($s - 1$) and, through either a linear or a non-linear function, produces an output,

$$z_j^{(s)} = f \left(\sum_{i=0}^M (w_{ji}^{(s)} z_i^{(s-1)}) \right). \quad (2)$$

Here w_{j0} denotes the bias for the hidden unit j , and f is an activation function such as the continuous sigmoid or, as used here, the tanh function, which has an output range of -1 to 1 :

$$f(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (3)$$

When the entire network has been executed, the output of the last layer is taken as the output of the entire network. The free parameters of ANNs are the weight vectors. During the training session, the weights of the connections are adjusted so as to minimize the total error function. The learning procedure is the so-called ‘back propagation’. The number of layers, the number of neurons in each layer, and the functions are chosen from the beginning and specify the so called ‘architecture’ of the ANN.

Neural networks learn by example. The neural network user gathers representative data into a training set and initiates the weight vector with a random seed, then invokes the training algorithms to automatically learn the structure of the data. Here, we use a method that is popular in neural network research: the

Levenberg-Marquardt method (Levenberg 1944; Marquardt 1963; also detailed in Bishop 1995). This has the advantage that it converges very quickly to a minimum of the error function. This error function may not have just a global minimum in the multidimensional weight space but could have a number of local minima instead. In general, network trained using exactly the same training set for the same given number of epochs but using different initial weights (different starting points in this space) will converge to slightly different final weights.

In order to avoid (possible) over-fitting during the training, another part of the data can be reserved as a validation set (independent both of the training and test sets, so not used in the updating of the weights), and used during the training to monitor the generalization error. After a chosen number of training iterations, the training terminates and the final weights chosen for the ANN are those from the iteration at which the cost function is minimal on the validation set. It is the so called ‘early stopping method’. This is useful to avoid over-fitting to the training set when the training set is small, but the disadvantage of this technique is that it reduces the amount of data available for both training and validation, which is particularly undesirable if the data set is small to begin with.

3 THE CHOSEN GALAXY SAMPLE AND PARAMETERS

The SDSS consortium has publicly released more than 10^5 spectroscopic galaxy redshifts in the Data Release 2 (DR2). In order to test the accuracy of the photometric redshifts derived from SDSS DR2, we selected all objects satisfying the following criteria (also see Vanzella et al. 2004): (1) r -band Petrosian magnitude $r < 17.77$; (2) the spectroscopic redshift confidence greater than 0.95 and no warning flags. This gave 159 346 galaxies, which we randomly partitioned into a training set of 60 000, a validation set of 20 000, and a test set of 79 346. We will explore different network complexities, the validation set is required to compare them, and the test set is used at the end to estimate the true error of the final network predictions.

Using different magnitude measurements given in SDSS, we compare the effect of input parameters for predicting the redshifts by considering many different input parameter sets, which mainly involve the Petrosian magnitudes, the model magnitudes and the dereddened magnitudes in five bands. The Petrosian (1976) magnitude is based on the flux (PSF) within an aperture defined by the ratio of the local surface brightness to the mean interior surface brightness. The model magnitude is used as a template to determine the PSF magnitude in each band. The galaxy images are fitted with a de Vaucouleurs profile and an exponential profile of arbitrary axis ratio and orientation. The total magnitude associated with the better fit of the two profiles is referred to as the ‘model’ magnitude. The magnitude after correcting for reddening is named the dereddened magnitude. One advantage of our ANN approach to photometric redshift estimation is that any additional parameters that can help in estimating the redshift can be easily incorporated as an additional input parameter. However, these parameters need to be chosen carefully: they must have a genuine dependence on the redshift. Here, we supplemented the 50% and 90% Petrosian flux levels of the SDSS training sample with the angular radii containing these fractions of the Petrosian flux. Each of these radii is a measure of the angular size of the galaxy, hence is a redshift-dependent property.

4 REDSHIFT PREDICTION WITH DIFFERENT INPUT PARAMETER SETS

The ANN experiments were performed using the Matlab nnet Toolboxes. The training and test samples are independent, but in fact it is required the former must be representative of the latter. A successfully trained neural network with the training set can be applied to the test sample. During the training, we alter the network architecture in many ways including the weights (which are initialized randomly) and final architecture is saved that corresponds to the smallest error in training sample (in almost all the cases this coincides with the last epoch).

To evaluate the accuracy of the prediction, we define the variance between the neural outputs (zNN) and the targets (spectral redshift z_{spec}) as

$$\sigma_z = \sqrt{\frac{1}{N} \sum_i (zNN_i - z_{\text{spec}_i})^2}, \quad (4)$$

where N is the number of galaxies, and $i = 1 \dots N$. This defines a statistical measure for the accuracy of the predicted redshifts by the ANN.

4.1 Petrosian Parameters

In this exploration, we select the Petrosian magnitudes in five different bands from SDSS DR2 as the root data. The Petrosian magnitude system which measures flux in apertures is determined by the shape of the surface brightness profile. By adding other parameters or changing to different parameter combination in the input set, the ANN was trained and the final test results were assessed with the Equation (4).

For the experiment, we directly use the Petrosian magnitudes in five bands (u, g, r, i, z) as the first set of input parameters. The number of hidden units was chosen by trial and error rather than some quantitative method, such as the Bayesian information criterion, for the lack a clearly better way. For our training set we tried one or two hidden layers and different numbers of nodes. The weights corresponding to the minimum training error were stored. The best error obtained was $\sigma_z = 0.027031$ for the architecture 5:10:10:1 (five input nodes, two hidden layers with ten nodes each and one output node).

To compare the effect of different input parameters, we used the Petrosian color index ($u - g, g - r, r - i, i - z$) and the r -band magnitude. For this, the final determined network architecture is 5:10:10:1 and the dispersion in test set is $\sigma_z = 0.026717$. We can see the result is slightly better than in the previous experiment. Unfortunately, the accuracy is not adequate for our photometric redshift estimation. So we need to consider increasing the amount of input information.

Certainly, there will be more information if more parameters are included, so we added the r -band 50 and 90 percent Petrosian flux radii (PetR50, PetR90) to the input. Now, there are seven input parameters. Again by trial and error, we found the ultimate structure is 7:16:16:1 and the rms scatter in the test set is $\sigma_z = 0.025048$. We can see the new information produced some improvement and the 7-parameter input is preferable to the 5-parameter input.

Indeed, increasing information in the training data is an obvious method to improve the generalization. Now let us include all the various parameters, the Petrosian 50 and 90 percent flux radii in all bands and the magnitudes in the five different bands so we now have 19 parameters (see Table 1) and the chosen architecture has just a single hidden layer with 20 neurons. The experiment shows that increasing the number of nodes in the architecture of the neural network does not change the results significantly. The final structure by trial and error is 19:20:1 and the scatter is reduced to 0.021596. This result shows that adding some new information gives a clearly better improvement. We compare the spectroscopic redshifts with the ANN photometric redshifts for our 19-parameter experiment in Figure 2. The resulting dispersions (σ_z) for the different sets of input parameters used are summarized in Table 1.

Table 1 Comparison of Different Sets of Input Parameters based on the Petrosian Magnitudes

Input	Parameters	σ_z (Train)	σ_z (Test)
5	Petrosian u, g, r, i, z	0.026939	0.027031
5	Petrosian $u - g, g - r, r - i, i - z, r$	0.026535	0.026717
7	Petrosian $u - g, g - r, r - i, i - z, r, \text{PetR50}, \text{PetR90}$	0.025002	0.025131
19	Petrosian $u - g, g - r, r - i, i - z, u, g, r, i, z$ PetU50, PetU90, PetG50, PetG90, PetR50 PetR90, PetI50, PetI90, PetZ50, PetZ90	0.021502	0.021596

Table 2 Comparison of Different Sets of Input Parameters based on the Model Magnitudes

Input	Parameters	σ_z (Train)	σ_z (Test)
5	model u, g, r, i, z	0.023354	0.023321
5	model $u - g, g - r, r - i, i - z, r$	0.022006	0.022097
7	model $u - g, g - r, r - i, i - z, r, \text{PetR50}, \text{PetR90}$	0.020765	0.02075
19	model $u - g, g - r, r - i, i - z, u, g, r, i, z$ PetU50, PetU90, PetG50, PetG90, PetR50 PetR90, PetI50, PetI90, PetZ50, PetZ90	0.02034	0.020465

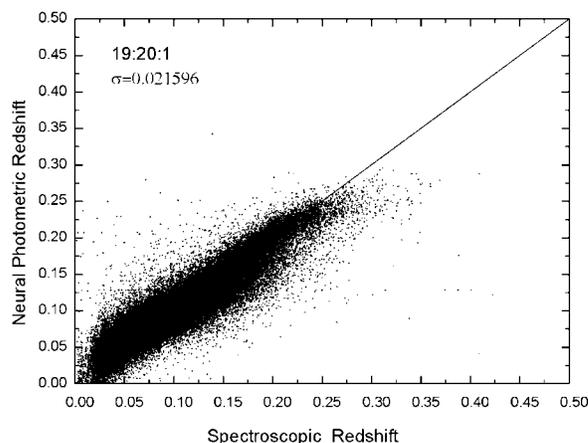


Fig. 2 Spectroscopic redshifts vs. the predicted photometric redshifts using 19 Petrosian magnitude parameters for 79 346 galaxies in the SDSS DR2 sample. The ANN architecture is 19:20:1 and the scatter is 0.021596.

4.2 The Model Parameters

We have attempted to find the optimal set of input parameters to use in a neural network for estimating photometric redshifts. A non-expert might ask, ‘how should I decide what parameters to use?’ The comparison process can be customized by specifying different input parameters, such as the model magnitudes. Thus, we now make other experiments, where we mainly focus on the model magnitudes in five different bands or these in combination with further parameters.

First, we use the model magnitudes in five bands (u, g, r, i, z) as the input parameters for the neural network. A 5:12:8:1 neural network was trained for 80 epochs with a validation set leading to an early termination (and again, a random initialization of the weights). This network produces a dispersion in the test sample of $\sigma_z = 0.0233$.

Instead of using only the magnitudes, we next took four model color index ($u - g, g - r, r - i, i - z$) and the model r -band magnitude as the input parameters. We used the same network architecture 5:12:8:1 and different distributions of weights. The ultimate prediction error at the output ($\sigma_z = 0.0221$) is relatively small.

For a further comparison, we added PetR50 and PetR90 to the above five input parameters. A 7:12:8:1 network were carried out with changes of the initial random distribution of weights and early stopping to avoid over-fitting during the training. The final error, $\sigma_z = 0.02075$, is a remarkable improvement. It is comparable to the other photometric redshifts in the literature based on neural networks, e.g., Tagliaferri et al. (2002), Firth, Lahav & Somerville (2003), Vanzella et al. (2004), Collister & Lahav (2004). So incorporating PetR50 and PetR90 in the input seems to be crucial in improving the agreement between photometric and spectroscopic redshifts.

Finally, we added some new information in the training set in order to reduce the systematic errors. Based on the above parameters, all the model magnitudes and the Petrosian 50 and 90 percent flux radii in the other bands were considered and 19 parameters in all (see Table 2) are input to the network. By trial and error, we came to a final network architecture of 19:12:8:1; its dispersion is $\sigma_z = 0.020465$ which is a slight improvement. Figure 3 compares the ANN redshifts with spectroscopic redshifts for the test set of 79 346 galaxies with the 19:12:8:1 network. In Table 2, we summarize some of the results obtained from the above experiments.

4.3 Dereddened Magnitude Parameters

Similar to the procedure to predict photometric redshifts based on Petrosian and model parameters, here we discuss parameter sets based on the dereddened magnitudes, as well as the dispersion of redshift estimation.

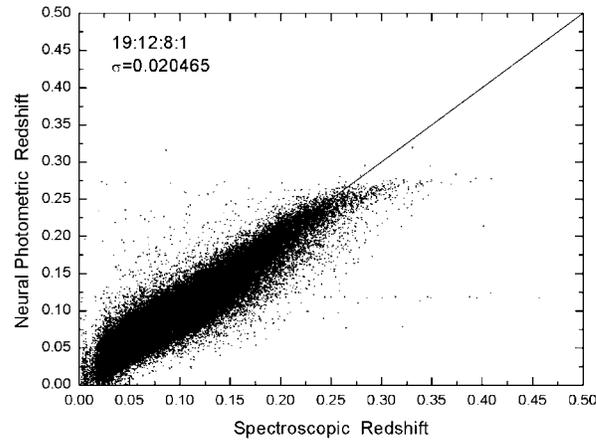


Fig. 3 A comparison of spectroscopic redshifts and predicted photometric redshifts on 19 model magnitude parameters. The ANN has a 19:12:8:1 architecture and was tested on a test set of size 79 346 (plotted) and gave $\sigma_z=0.020465$.

We adopt the same sample. The training was carried out to 3000 epochs. Different architectures with one or two hidden layers and different numbers of nodes were tried. For the different parameter sets, the ANN architectures are 5:5:5:1, 5:5:10:1, 7:10:1 and 19:12:1, respectively. These, and the corresponding RMS values are listed in Table 3. The best result is $\sigma_z = 0.020184$, with the 19-parameter input.

Moreover, we have studied the effect of adding the errors of the magnitudes (five parameters) to the input, resulting in a 24-parameter input set. The resulting network gave a scatter of $\sigma_z = 0.020053$.

Table 3 Comparison of Different Sets of Parameters Involving Dereddened Magnitudes

Input	Parameters	σ_z (Train)	σ_z (Test)
5	dereddened u, g, r, i, z	0.021371	0.02388
5	dereddened $u - g, g - r, r - i, i - z, r$	0.021081	0.021097
7	dereddened $u - g, g - r, r - i, i - z, r, \text{PetR50}, \text{PetR90}$	0.020821	0.020689
19	dereddened $u - g, g - r, r - i, i - z, u, g, r, i, z,$ $\text{PetU50}, \text{PetU90}, \text{PetG50}, \text{PetG90}, \text{PetR50},$ $\text{PetR90}, \text{PetI50}, \text{PetI90}, \text{PetZ50}, \text{PetZ90}$	0.020174	0.020184

5 DISCUSSION

We have presented extensive experiments with a variety of input parameters for the estimation of redshifts based on feed-forward neural networks. There are a few points to observe about the results. First, the experiments with the Petrosian magnitudes as the root data are listed in Table 1. The combination of four Petrosian color index ($u - g, g - r, r - i, i - z, r$) plus the r -band Petrosian magnitude performed better than just the five magnitudes (u, g, r, i, z). In order to improve the prediction and investigate the effect of other input parameters, we further added PetR50 and PetR90 and we got improved prediction results on the same data. Finally, as shown in Table 1 when 19 parameters are taken, the redshift prediction is markedly improved and the system error decreased.

Secondly, we used the model magnitudes as basic input. It is shown in Table 2 that the second set of parameters yielded a higher prediction accuracy than the first one, that the performance of the combination of color index with the r -band magnitude is better than using just the five magnitudes. Similarly, we added two other parameters (PetR50, PetR90) which will hopefully offer some new information for the training.

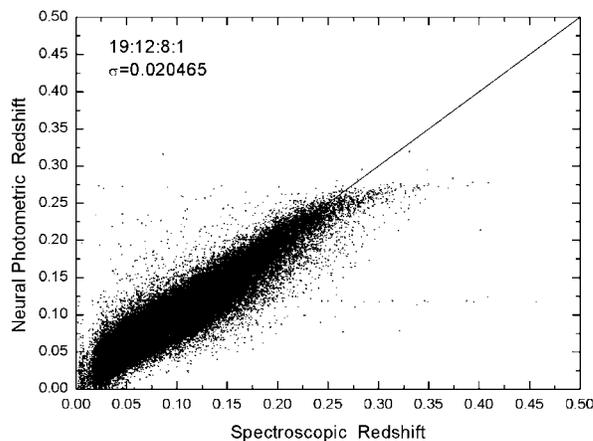


Fig. 4 Redshifts prediction using dereddened magnitude with 19 input parameters. The ANN architecture is 19:12:1 and the test sample size is 79 346 (plotted).

Generally, with increasing amount of information containing more features, the prediction should improve more and more. Indeed, the 7-parameter set gave a rather good performance and the 19-parameter set gave a further, though slight, improvement. See Table 2.

Thirdly, The results of adding the dereddened parameters are given in Table 3. We can see the combination of the dereddened color index with the r -band magnitude ($u - g, g - r, r - i, i - z, r$) is better than using just the five magnitudes (u, g, r, i, z). Moreover, PetR50 and PetR90 will further improve the performance of the neural network. With 19 parameters the result of redshift prediction is again increased.

Finally, as indicated in Tables 1–3, when similarly considering the magnitudes in five bands, the dereddened magnitudes as parameters gave the smallest dispersion σ_z^{test} among the three kinds of magnitudes and the model magnitudes are better than the Petrosian magnitude. In addition, all the combinations involving the dereddened magnitudes are superior to those involving the Petrosian magnitudes or model magnitudes. Furthermore, there is a slight improvement when we consider the error of model magnitudes.

6 CONCLUSIONS

In this paper, we have described experiments that compare the performance of a number of different sets of input parameters for estimating photometric redshifts. From the experimental results, we can easily see that, whether we use the Petrosian magnitudes, the model magnitudes or the dereddened magnitudes, we can conclude that the more parameters are considered, the higher the accuracy will be. More input parameters used in the training data provide more information for the network to improve its capability of prediction and generalization, so increasing the final accuracy correspondingly. Moreover, it is clear that the performance of the dereddened magnitudes is superior to that of the Petrosian magnitudes and model magnitudes for the same network structure and the same data set. Therefore, we can see the dereddened magnitudes offer some significant advantage over the Petrosian magnitudes and model magnitudes, and all the three sets of parameters are available for neural network estimation of photometric redshifts. Our best prediction accuracy of photometric redshifts is $\sigma_z = 0.020184$, which is a result from statistical computation of large-size samples and which, hopefully, will help large-scale structure researcher in their study of some cosmic issues.

With the advances in astronomical observation, more and more parameters are becoming available, it therefore becomes increasingly desirable to select the most suitable parameters among them. This is a major problem in the empirical photometric redshift estimation where inappropriate parameters that have no obvious redshift dependence will lead to larger scatter and error. Selecting appropriate and effective parameters is a challenging issue in future research. In order to improve the accuracy of photometric redshift estimation, we should consider more multiwavelength band parameters, such as J, H, K_s from 2MASS.

Moreover, we will further perform feature extraction (e.g. principal component analysis, PCA) to reveal the underlying factors or components in the multi-dimensional parameter space.

The above neural network applications were concerned with the photometric redshift, but neural networks have had wider applications in astronomy. The usefulness of neural networks derives from the fact that they are an efficient and effective means of tackling problems which are non-linear or concerned with multi-parameter problems. Neural network techniques for solving problem are designed primarily to give an accurate representation of the relationship between two sets of variables, and they are particularly successful when the relationship is highly complex. With their implementation in redshift estimation, it has become evident that neural networks are a very useful and adaptable addition to the tools available to astronomers in tackling a wide variety of problems (e.g., classification, regression and feature selection).

Acknowledgements This paper is funded by National Natural Science Foundation of China under grants Nos. 10473013, 90412016 and 60603057.

References

- Adams A., Woolley A., 1994, Hubble Classification of Galaxies Using Neural Networks., *Vistas in Astronomy*, 273, 280
- Ball N. M., Loveday J., Fukugita M. et al., 2004, *MNRAS*, 348, 1038
- Baum W. A., 1962, IAU Symp. No. 15, Problems of Extragalactic Research, p.390
- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
- Blake C., Bridle S., 2005, *MNRAS*, 363, 1329
- Bishop C. M., 1995, *Neural Networks for Pattern Recognition*, Oxford: Oxford University Press
- Csabai I., Budavári T., Connolly A. J. et al., 2003, *AJ*, 125, 580
- Coe D., Bentz N., Sánchez S. F. et al., 2006, *AJ*, 132, 926
- Collister A. A., Lahav O., 2004, *PASP*, 116, 345
- Connolly A. J., Csbai I., Szalay A. S., 1995, *AJ*, 110, 2655
- Connolly A. J., Szalay A. S., Brunner R. J., 1998, *ApJ*, 499, L125
- Fernández-Soto A., Lanzetta K. M., Yahil A., 1999, *ApJ*, 513, 34
- Firth A. E., Lahav O., Somerville R. S., 2003, *MNRAS*, 339, 1195
- Folkes S. R., Lahav O., Maddox S. J., 1996, *MNRAS*, 283, 651
- Fontana A., D'Odorico S., Poli F. et al., 2000, *AJ*, 120, 2206
- Gwyn S. D. J., Hartwick F. D. A., 1996, *ApJ*, 468, 77
- Haykin S., 1994, *Neural Networks: A Comprehensive Foundation*, New York: Macmillan College Publishing Company
- Heymans C., Heavens A., 2003, *MNRAS*, 339, 711
- Hogg D. W., Cohen J. G., Blandford R., 1998, *AJ*, 115, 1418
- King L. J., Schneider P., 2003, *A&A*, 398, 23
- Lahav O., Naim A., Sodr  L. J. et al., 1996, *MNRAS*, 283, 207
- Levenberg K., 1944, *Quarterly Journal of Applied MathematicII*, 2, 164
- Marquardt D. W., 1963, *Journal of the Society of Industrial and Applied Mathematics*, 11, 431
- Nielsen M. L., Odewahn S. C., 1994, 185th AAS Meeting, 26, 1498
- Odewahn S. C., Nielsen M. L., 1994, *Vistas in Astronomy*, 38, 281
- Sawicki M. J., Lin H., Yee H. K. C., 1997, *The Hubble Space Telescope and the High Redshift Universe*, Proceedings of the 37th Herstmonceux conference, Cambridge, UK., p.175
- Sowards-Emmerd D., McKay T. A., Sheldon E. et al., 1999, *American Astronomical Society*, 31, 827
- Sowards-Emmerd D., Smith J. A., McKay T. A. et al., 2000, *AJ*, 119, 2598
- Tagliaferri R., Longo G., Andreon S. et al., 2002, (astro-ph/0203445)
- Vanzella E., Cristiani S., Fontana A. et al., 2004, *A&A*, 423, 761
- Wadadekar Y., 2005, *PASP*, 117, 79
- Weaver W. B., 2000, *ApJ*, 541, 298
- Weinstein M. A., Richards G. T., Schneider D. P. et al., 2004, *ApJS*, 155, 243
- York D. G., Adelman J., Anderson J. E. et al., 2000, *AJ*, 120, 1579
- Zhang Y. X., Zhao Y. H., 2003, *Chin. J. Astron. Astrophys. (ChJAA)*, 3, 183
- Zhang Y. X., Zhao Y. H., 2007, *Chin. J. Astron. Astrophys. (ChJAA)*, 7, 289