

AN ARTIFICIAL NEURAL NETWORKS APPROACH TO THE ESTIMATION OF PHYSICAL STELLAR PARAMETERS

Rodriguez, Alejandra

*Department of Information and Communication Technologies
University of A Coruña. A Coruña. 15071. Spain*

alejandra@udc.es

Carricajo, Iciar

*Department of Navigation and Earth Sciences
University of A Coruña. A Coruña. 15011. Spain*

iciar@udc.es

Dafonte, Carlos

*Department of Information and Communication Technologies
University of A Coruña. A Coruña. 15071. Spain*

dafonte@udc.es

Arcay, Bernardino

*Department of Information and Communication Technologies
University of A Coruña. A Coruña. 15071. Spain*

cibarcay@udc.es

Manteiga, Minia

*Department of Navigation and Earth Sciences
University of A Coruña. A Coruña. 15011. Spain*

manteiga@udc.es

Abstract

This paper presents an artificial neural networks approach to the estimation of effective stellar temperatures by means of optical spectroscopy.

The present work is included in a global project, whose final objective is the development of an automatic system for the determination of the physical and chemical parameters of stars. In previous works, we designed a hybrid system that integrated neural networks, fuzzy logic and expert systems to obtain the stellar spectral type and luminosity in the MK standard system. Considering those results, we now propose the design of several neural networks for the calculation of stellar temperatures.

The proposed networks have been trained with synthetic spectra that were previously contrasted with statistical clustering techniques. The final networks obtained a success rate of 88% for public catalogue spectra.

Our final objective is to calibrate the MK classification system, obtaining thus a new relation between the temperature and the spectral type.

Keywords: Neural Networks, Clustering Techniques, Stellar Physical Parameters, Optical Spectroscopy

1. Introduction

This work is part of a global project that studies the last phases of stellar evolution. Our main objective is to provide an automatic system, based on artificial intelligence techniques, that contributes to the determination of chemical and physical stellar parameters by means of optical spectroscopy.

Spectroscopy is among the most powerful currently available techniques for the study of stars and, in particular, their physical conditions (temperature, pressure, density, etc.) and chemical components (H, He, Ca, K, etc.). In general terms, a stellar spectrum consists of a black body continuum light distribution, distorted by the interstellar absorption and reemission of light, and by the presence of absorption lines, emission lines and molecular bands [Zombeck, 1990].

The stellar spectra are collected from telescopes with appropriate spectrographs and detectors. Observers collect the flux distribution of each object and reduce these data to obtain a one-dimensional spectrum calibrated in energy flux ($\text{erg}^{-1}\text{cm}^{-2}\text{s}^{-1}\text{\AA}^{-1}$) and wavelength (\AA).

As part of the above-mentioned global project, we have collected a sample of approximately 400 low-resolution stellar spectra from astronomical observations that were carried out at several telescopes. In order to extract useful information from the individual spectra and to study the evolution in the whole sample, we must determine the main stellar parameters, such as spectral type, luminosity, temperature, surface gravity, metallicity, etc.

The estimation of the stellar parameters is often carried out by human experts, who analyse the spectra by hand, with no more help than their own experience. These manual analyses usually lead to a MK classification of the spectra. The MK classification system was firstly proposed in 1943 by Morgan, Keenan & Kellman [Morgan et al., 1943] and has experienced many revisions ever since [Kurtz, 1984]. This bidimensional system is the only one that is widely used for stellar classification. One of its main advantages is that MK classifications are often static, because they are based on the visual study of the spectra and on a set of standard criteria. However, the same spectra can be classified differently by different experts and even differently by the same person at different times.

Any classification system should hold a compromise between maintaining the full information of the spectra and the need for a compact summary of this information. An optimal summary is obviously obtained by a study of the physical parameters.

The manual techniques that are currently used to estimate stellar parameters are very time-consuming and involve a great amount of human resources. It would be highly advisable to dispose of automatic, fast and efficient computational techniques that allow the experts to classify a large number of spectra according to their physical parameters.

In previous works, we developed an expert system for the classification of the stellar spectra of Post-AGB stars in the visible electromagnetic spectral range. The obtained results led us to extend this system to stars of different luminosities and to add new computational techniques, such as fuzzy logic, in order to refine the automatic processing of spectra. We have also tested feed-forward, radial-basis functions (RBF) and self-organising neural networks in order to obtain the spectral type and luminosity of stars through the analysis of their optical spectra. Combining both techniques, we formalised a hybrid system that obtains MK classifications and is able to determine the most appropriate method for each spectrum type. A complete description of these works can be found in [Rodríguez et al., 2003].

Our previous developments have proven that artificial techniques allow researchers to obtain the spectral type and luminosity of stars in a fast, easy and accurate way. Neural networks in particular have given excellent results, which is why this paper proposes the design and implementation of several artificial neural networks to estimate the effective stellar temperature. By carrying out a sensibility analysis of this technique in the estimation of the physical parameters of stars, we can determine the best learning algorithm and network structure for this specific problem.

2. Methodology

As mentioned in Sect. 1, we have implemented artificial techniques in order to obtain MK classifications. That first approach combined signal processing [Kalouptsidis, 1997], knowledge-based systems [Buchanan and Shortliffe, 1984], fuzzy logic techniques [Bouchon-Meunier et al., 1991] and artificial neural networks models [Arbib, 1995], integrated by a relational database that stores and structures all the stellar information. Table 1 contrasts the behaviour of the artificial techniques and that of two human experts who collaborated on this project, applied to 100 unclassified spectra belonging to the standard atlas of spectra from [Pickles, 1998].

Table 1. Artificial and Manual Techniques for MK Classifications

<i>Approach</i>	<i>Global</i>	<i>Spectral Type</i>	<i>Luminosity</i>
<i>Human Expert A</i>	99.0%	92.0%	81.0%
<i>Human Expert B</i>	95.0%	85.0%	70.0%
<i>Expert Systems</i>	95.6%	90.3%	78.2%
<i>Fuzzy Expert Systems</i>	98.6%	93.5%	79.0%
<i>Backpropagation</i>	97.0%	95.4%	81.0%
<i>RBF</i>	96.2%	94.6%	80.0%
<i>Self-Organizing Maps</i>	73.0%	68.0%	55.0%

These results led us to develop a final hybrid system based on an expert system that determines the global type of each star and, according to the type, sends the spectra to different neural networks in order to obtain their spectral type as well as their luminosity level. The success rate of that system was very similar to the agreement percentage between experts in the field (about 80%). This paper does not describe the development in detail: a more complete description of the models and stellar parameters used can be found in [Rodriguez et al., 2003]. We only include a brief explanation of the morphological algorithms so as to clarify how the spectral parameters, used as the input layer of most of the proposed neural networks models, are obtained and measured.

2.1 Morphological Analysis

Before presenting the spectra to the neural networks, we carry out a morphological analysis of all the spectra in order to obtain the values of the parameters that characterise each spectrum separately. These spectral parameters include the measurement of 25 spectral features that can be divided into three types:

- Absorption and emission lines: including hydrogen, helium and metallic lines (Ca, K, etc.).
- Molecular bands: hydrogen and carbon absorption bands.
- Rates between lines: CH-K rates, He-H rates.

From a morphological point of view, an absorption line is a descending (ascending for emission) deep peak that appears in an established wavelength zone [Zwicky, 1957]. To accurately calculate the intensity of each line, we carry out an estimation of the local spectral continuum. That is

$$C_j = \left(\frac{\sum_{j-n}^{j+n} E_i * X_i}{N} \right), \quad (1)$$

where C_j is the estimation of the continuum for sample j , E_i is the flux in sample i , N is the number of values used to calculate the local spectral continuum, and X is a binary vector that indicates the representative fluxes of the spectral continuum in the zone.

A molecular band is a spectral zone in which the flux suddenly decreases from the local continuum during a wide lambda interval [Zwicky, 1957]. For the molecular bands, this means that we only have to measure their energy to decide if they are significant enough. That is

$$B_{lr} = \int_l^r L(\lambda_i) - \int_l^r E(\lambda_i), \quad (2)$$

where B_{lr} is the flux of the band between the samples l and r , L is the projection line, E is the flux function, λ the wavelength, l the left limit of the band and r the right limit.

2.2 Training and Validation Spectra

In order to build the training set of the artificial neural networks that will be applied to the problem of estimation of stellar effective temperatures, we have chosen a complete and consistent set of synthetic optical spectra. The selected spectra were generated using the SPECTRUM v. 2.56. software written by Richard O. Gray. This software is a stellar spectral synthesis program that allows us to obtain prototype spectra of a chosen effective temperature and surface gravity. SPECTRUM is distributed with an atomic and molecular line list for the optical spectral region 3500 Å to 6800 Å, called luke.lst, suitable for computing synthetic spectra with temperatures between 4000K and 20000K. The details on the physics included in this software can be found in [Gray, 2003].

For the generation of synthetic spectra we also used the models calculated by Robert Kurucz [Kurucz, 1979]. The atmospheric model is a tabulation of temperature and pressure at range of mass depths in the stellar photosphere, calculated on the basis of a variety of opacity sources. Each model is characterised by four parameters: effective temperature (T_{eff}), metallicity ($[M/H]$), microturbulence velocity (V_{micro}) and surface gravity ($\log g$).

We generated a total of 350 solar metallicity spectra with effective temperatures between 4000K and 20000K and surface gravity from 0.5 to 5.0. This set of synthetic spectra covers the spectral range K-B with luminosity levels I to V (in the MK system). The spectra were generated in the 3500-6800 Å range, which is the wavelength interval where the software is capable of correctly obtaining synthetic spectra.

Before training the neural networks, the synthetic spectra were contrasted with statistical clustering techniques in order to verify their suitability and to avoid dispersions in the whole sample.

Grouping techniques are often used as a tool to verify and test the quality of an established clustering of data: the grouping algorithms make their own division into classes of the input data. Our purpose was to try to discard those spectra that were not similar to any others. We implemented the max-min and K-clustering algorithms [Kaufman and Rousseuw, 1990]. The data groups obtained by these two techniques prevent us from using some of the initial spectra that could not be included in any of the clusters. We also observed that the algorithms were able to accurately group the spectra of low and medium temperatures (4000K-8000K), although both techniques stored all the spectra with temperatures above 8000K in the same category. Obviously, they are not capable of separating spectra with temperatures higher than 8000K. Finally, we built the training set with 348 synthetic spectra, excluding those that were discarded by the clustering techniques.

As for the test set of spectra, we used the public catalogue of Pickles [Pickles, 1998], choosing only the 79 spectra that match our temperature range.

2.3 Artificial Neural Networks Implementation

The neural networks used in this experimentation are based on supervised learning models and backpropagation networks [Rojas, 1996] in particular. In our experiments, we have made use of three different learning algorithms:

- Standard backpropagation: as the most common learning algorithm, it updates the weights after each training pattern.

- Enhanced backpropagation: it uses a momentum term that introduces the old weight change as a parameter for the computation of the new weight change.
- Batch Backpropagation: in standard backpropagation, an update step is performed after each single pattern; in batch backpropagation, all the weight changes are summed over a full presentation of all the training patterns (one epoch). Only then, an update with the accumulated weight changes is performed.

As for the topology, the different networks that were implemented are shown in Table 2. These topologies have been tested for the three backpropagation learning algorithms.

We built the input patterns for the nets with 659 flux values (from 3510 Å to 6800 Å), as well as with the 25 spectral features obtained by means of the morphological algorithms described in Sect. 2.1. The output of all networks is a continuous function that obtains the effective temperature in the form $T_o = ((\log T_{eff} - 3)/2)$.

Table 2. Topologies tested for Backpropagation Networks

<i>Input Patterns</i>	<i>Hidden Layer</i>
659 flux values	100:50:10:3
659 flux values	10:5:3*
659 flux values	10:10
659 flux values	5:5
25 Spectral Parameters	10:5:3
25 Spectral Parameters	10:10
25 Spectral Parameters	5:5

* Best performance topology.

We used the Stuttgart Neural Network Simulator (SNNS v.4.1) [Zell, 2002] to implement the above described networks and to transform them into C code, so they can be integrated into the previously developed tool. In the training phase, the weights were updated in a topological order and initiated randomly with values in the [-1, 1] interval.

During our experiments, we found that the best topology corresponds to an enhanced backpropagation network 659:10:5:3:1; that is, a network with 659 flux values in the input layer, three hidden layers with 10, 5 and 3 neurons respectively, and 1 neuron for the effective temperature in the output layer.

2.4 Enhanced Techniques

While analysing the success rates of the implemented neural networks for each different temperature, we observed that all the networks have a higher performance for temperatures between 4000K-8000K (less than 200K mean error). Considering this fact, we separated the input spectra into two sets (4000K-8000K and 8000K-20000K) and trained the best performance network once again, but only with the patterns that correspond to spectra with temperatures in the interval 8000K-20000K. This “overtrained” network resulted in a better performance for all the test spectra.

An additional research consisted in analysing the weights of the output layer units of the best network so as to determine, for each temperature interval, which input parameters have more influence on the output. We were able to reduce the inputs of the network because we considered only the sufficiently significant flux values. This enhanced strategy allowed us to obtain networks that converge sooner, and has resulted in a significant improvement of the performance of the original artificial neural networks (around 2%).

3. Discussion and Results

The best implemented neural networks are able to determine the temperature of the 79 selected spectra from the test catalogue with an error rate below 15% and a mean deviation of 300 K, according to the estimation made by Pickles [Pickles, 1998].

The experts in this field usually work with a mean error of 10% of the correct temperature. Taking this error rate into account, our neural networks approach is able to correctly estimate 88% of the effective temperatures with a minimal mean deviation of 3% (for the coolest stars) and a maximum mean deviation of 7% (for the hottest stars). The final performance is shown in Table 3.

As shown in the results, the best temperature estimations are reached for cooler stars (4000K-8000K), whereas the stars with higher temperatures present a higher error rate; one explanation of this particularity could be that the training set includes few stars of this temperature range, because the models on which we based the generation of the synthetic spectra (Kurucz) can generate only 1 “hot spectrum” for every 4 “cool spectra”. Including more spectra for higher temperatures is not meaningful, because the error bar for this interval is established in 1000K. So, for higher temperatures, we can only include one sample spectrum for every 1000K, whereas for cool stars, where the error bar is 250K, we can include four samples. In future extensions, we hope

to increase the success rate for hot stars by designing a net of neural networks that includes different trained networks for hot and cool stars.

Table 3. ANN Performance for Effective Temperature

<i>T_{eff}</i>	<i>Spectra number</i>	<i>Success rate</i>	<i>Mean temperature error</i>
4000-6000	36	97.22%	±230 K
6000-8000	19	89.47%	±180 K
8000-10000	10	60.00%	±360 K
10000-20000	14	85.71%	±580 K
<i>Total</i>	<i>79</i>	<i>88.60%</i>	<i>± 300 K</i>

4. Conclusions

This paper has presented a computational approach to the estimation of the physical parameters of stars and, in particular, the calculation of effective temperature by means of artificial neural networks.

We have described several models of neural networks and analysed their performance and results to discover the best approach to the estimation of each temperature range. Backpropagation networks were trained with approximately 350 synthetic spectra that were previously contrasted with statistical clustering techniques. In order to obtain the input patterns of the neural networks and extract and measure spectral features, we used the morphological analysis algorithms that were developed in the expert systems approach. Several networks were trained with this parameterisation, and other networks with full flux values; the second strategy yielded the best performance.

The best networks have reached a success rate of approximately 88% for a sample of 79 testing spectra from public catalogues. For the development and evaluation of the system, we can count on the essential collaboration of experts from the area of Astronomy and Astrophysics of the University of A Coruña.

As an additional study, the best implemented networks were analysed to determine the flux values that are more influential in the output for each temperature interval; by training the networks with these small input patterns, we increased the performance of the original networks.

At present, we are analysing other types of neural structures and learning algorithms in order to improve the classification properties. We are also studying the influence of differences in signal to noise ratio in the resulting classification, since our final objective is to obtain a calibration that can determine our own relation between temperature

and MK classifications. We are also completing the development of our stellar database, STARMIND (<http://starmind.tic.udc.es>), to make it accessible through the Internet.

Acknowledgments

The authors acknowledge support from grants AYA2000-1691 and AYA2003-09499, financed by the Spanish Ministerio de Ciencia y Tecnología.

References

- [Arbib, 1995] Arbib, M. A. (1995). *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, Massachusetts.
- [Bouchon-Meunier et al., 1991] Bouchon-Meunier, R., Yager, R., and Zadeh, L. A. (1991). *Uncertainty in Knowledge Bases*. Springer-Verlag, Berlin.
- [Buchanan and Shortliffe, 1984] Buchanan, B. and Shortliffe, E. (1984). *Ruled-based Expert Systems*. Reading MA: Addison-Wesley.
- [Gray, 2003] Gray, R. (2003). Spectrum: A stellar spectra synthesis program. <http://www1.appstate.edu/dept/physics/spectrum>.
- [Kalouptsidis, 1997] Kalouptsidis, N. (1997). *Signal Processing Systems: Theory and Design*. Wiley, New York.
- [Kaufman and Rousseuw, 1990] Kaufman, L. and Rousseuw, P. J. (1990). *Finding Groups in Data*. Wiley, New York.
- [Kurtz, 1984] Kurtz, M. J. (1984). *The MK Process and Stellar Classification of Stars*. David Dunlap Observatory, Toronto.
- [Kurucz, 1979] Kurucz, R. L. (1979). Model atmospheres for G, F, A, B and O stars. *The Astrophysical Journal Suppl.*, 40.
- [Morgan et al., 1943] Morgan, W. W., Keenan, P. C., and Kellman, E. (1943). *An Atlas of Stellar Spectra with an Outline of Spectral Classification*. University of Chicago Press, Chicago.
- [Pickles, 1998] Pickles, A. J. (1998). A stellar spectral flux library. 1150-25000 Å. *Publications of the Astronomical Society of the Pacific*, 110:863-878.
- [Rodriguez et al., 2003] Rodriguez, A., Dafonte, C., Arcay, B. and Manteiga, M. (2003). An Artificial Neural Network Approach to Automatic Classification of Stellar Spectra. In *Proc. 7TH. Int. Work Conf. on Artificial and Natural Neural Networks (IWANN'03)*, 2: 639-646, Menorca, Spain.
- [Rojas, 1996] Rojas, R. (1996). *Neural Networks. A Systematic Introduction*. Springer-Verlag, Berlin.
- [Zell, 2002] Zell, A. (2002). Stuttgart neural network simulator. <http://www-ra.informatik.uni-tuebingen.de/SNNS/>.
- [Zombeck, 1990] Zombeck, M. V. (1990). *Handbook of Astronomy and Astrophysics*. Cambridge University Press, Cambridge, second edition.
- [Zwicky, 1957] Zwicky, F. (1957). *Morphological Astronomy*. Springer, Berlin.