

The inversion of Stokes profiles with artificial neural networks

T. A. Carroll and J. Staude

Astrophysikalisches Institut Potsdam, Telegrafenberg, Sonnenobservatorium Einsteinurm, 14473 Potsdam, Germany

Received 2 July 2001 / Accepted 10 August 2001

Abstract. We investigate the application of artificial neural networks (ANNs) for the interpretation of Stokes profiles. We have employed ANNs to approximate the nonlinear inverse mapping between the Stokes profiles and some of the underlying atmospheric parameters. This approximate model is used in the following to carry out a fast non-iterative inversion of synthetic Stokes profiles. We have used synthetic Stokes profiles of the photospheric infrared line Fe I $\lambda 15648$ to demonstrate that the ANNs are capable to yield accurate single valued estimates of the complete magnetic field vector, line-of-sight (LOS) velocity, microturbulence, macroturbulence and the filling factor with exceptional speed. For a stratified atmosphere we also demonstrate that these single valued parameters do represent very good averaged values of the input stratification. To retrieve some of the temperature information encoded in the Stokes profiles we modeled a neural network classifier on the basis of several semi-empirical model atmospheres (i.e. temperature and pressure stratification). With this classifier we are able to determine the probability that a given Stokes profile has its origin from a particular temperature stratification of a semi-empirical model.

Key words. line: formation – line: profiles – Sun: photosphere – Sun: magnetic fields – method: data analysis

1. Introduction

The reliable determination of magnetic fields and the thermodynamic structures of the solar atmosphere is still crucial for our understanding of many complex processes in and outside magnetic active regions. The inference of atmospheric parameters from Stokes profiles is a nonlinear inverse problem. Typically, those problems are solved by linearizing an appropriate forward model, computing the sensitivities and then iteratively solving a regularized optimization problem. In this sense most of the inversion procedures for Stokes profiles are based on a non-linear least-square minimization (i.e. Levenberg-Marquardt algorithm) (Auer et al. 1977; Landolfi et al. 1984; Skumanich & Lites 1987). All the latter are using a simplified forward model (Milne-Eddington model) in order to yield an analytical solution for the fast evaluation of the required derivatives. The most advanced Stokes profile inversion so far was introduced by Ruiz Cobo & del Toro Iniesta (1992). This inversion code is based on response functions and does not rely on the analytical Milne-Eddington solution and therefore it is able to retrieve height dependent information within a reasonable time.

But the improving techniques in spectro-polarimetry and the improving spatial resolution of especially future projects like GREGOR (von der Luehe et al. 1999), SOLIS (Keller 1998) or SolarB (Lites 2000) provide us with huge amounts of polarimetric data. This emphasizes the need for a fast and stable automated analysis and interpretation of Stokes profiles. In this sense the new inversion method described by Socas-Navarro et al. (2001) which operates in the low-dimensional eigenfeature space, produced by a principle component analysis (PCA) of Stokes profiles, bypass the iterative behaviour and seems to be a very promising alternative and supplement to existing inversion methods. In this paper we introduce another alternative to an iterative solution, a direct inversion, based on the approximation of the nonlinear inverse mapping between the Stokes parameter and some of the underlying atmospheric parameters. The inversion based on this approximate inverse mapping proved to be stable, accurate and extremely fast.

Artificial neural networks were already successfully applied in many different astrophysical fields such as the classification of stellar spectra (Bailor-Jones et al. 1998; Weaver & Torres-Dodgen 1997; Gulati et al. 1994), the retrieval of stellar parameters from low resolution spectra (Bailor-Jones 2000), the time series analysis of solar active regions (Calvo et al. 1995) or the sunspot index prediction (Fessant et al. 1996).

Send offprint requests to: T. A. Carroll,
e-mail: tcarroll@aip.de

We have used one of the most popular types of ANNs, the Multi-Layer Perceptron (MLP), which is a general model for approximating nonlinear multivariate functions. To adapt (i.e. train) our MLP to the inverse problem we take advantage of the fact that we have a good knowledge of the forward problem (i.e. the polarized radiative transfer) in the case of magnetic sensitive absorption lines formed under the conditions of local thermodynamical equilibrium (LTE). Thus we are able to generate a sufficient large database of synthetic Stokes profiles for the training process of the MLP. We have applied this model in the following work to simulated observation of the infrared line (IR) Fe I $\lambda 15648$. Once the MLP is trained and has found a good approximation of the inverse mapping it can recover the complete magnetic field vector, the line of sight velocity, the microturbulence, the macroturbulence and the filling factor as well as an estimate of the temperature stratification with exceptional speed.

This paper is organized as follows: in Sect. 2 we describe the general structure and properties of a Multi-Layer Perceptron. In Sect. 3 we investigate the capabilities of the MLP to retrieve some of the temperature information encoded in the Stokes profiles. We consider the problem if an MLP can recognize a specific signature in a given Stokes profile to distinguish between the temperature structure of different semi-empirical model atmospheres. Extensive statistical tests are made to evaluate the performance of the MLP. Section 4 deals with the inference of various atmospheric parameters from the Stokes parameters. The training of the MLPs to approximate the inverse relation and the statistical tests of the trained MLPs are described. Because our estimates are single valued we consider, in Sect. 5, the influence of magnetic field strength and LOS velocity gradients on the MLP calculations. To assess the results we use the concept of heights of formation (HOFs) defined by Sanchez Almeida et al. (1996), to demonstrate that the calculations of the MLP do represent very good averaged values of the stratified atmosphere. In Sect. 6 we employed the MLP model to retrieve estimates for the filling or stray light factor. We simulated a simple two component atmosphere to demonstrate that a trained MLP can yield reliable estimates about the fraction of the magnetic component. In Sect. 7 we give a short comment on the remarkable speed of the inversion. Finally, in Sect. 8, we summarize the main conclusions of this work and give an outlook to some future applications.

2. The neural network model

The adaptive model for approximating the inverse relationship between the Stokes profiles and the physical parameters is a fully connected two-layer MLP which is a practical framework for approximating arbitrary nonlinear multivariate functions. The general adaptive function,

which is represented by a fully connected two-layer MLP, is given by

$$y_k(\mathbf{x}; \mathbf{w}) = g_k \left(\sum_{j=0}^M w_{kj}^{(2)} g_j \left(\sum_{i=0}^L w_{ji}^{(1)} x_i \right) \right), \quad (1)$$

where y_k is the k th element of the output vector of the MLP. The adaptive parameters are the so called weights $w_{ji}^{(1)}$ in the first layer and $w_{kj}^{(2)}$ in the second layer. The input from the i th input value is given by x_i . The processing units g are also called the activation functions and are given by the nonlinear logistic sigmoid function, whose output lies in the range $(0, 1)$

$$g(a) = \frac{1}{1 + \exp(-a)}. \quad (2)$$

The process of determining the values of the weight parameters of the network function (1) on the basis of a given data set is called (supervised) learning or training. The data set should contain a representative and sufficient number of input-target relations (i.e. the Stokes profiles and the atmospheric parameters which give rise to that profiles) to allow an appropriate representation of the complex functional mapping of the problem. The training process which tries to find an approximation of the intrinsic functional mapping of the training data, is formulated as a nonlinear optimization problem. The common type of error function for that optimization problem is given by the sum of square error function

$$E = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^c \{y_k(\mathbf{x}^n; \mathbf{w}) - t_k^n\}^2, \quad (3)$$

where N is the number of training patterns (i.e. pairs of input-target vectors) in the training set, $y_k(\mathbf{x}^n; \mathbf{w})$ is the calculated output of the network for the n th input vector and t_k^n is the desired k th target value.

Once the network is successfully trained in terms of minimizing the error function (3) the network weights are frozen and the network, which represents a nonlinear function, can be applied to unknown data (application mode).

In the following work we make use of two theoretical concepts for the Multi-Layer Perceptron. It can be shown that the MLP network function (1) can in principle approximate any continuous functional mapping to arbitrary accuracy (Funahashi 1989; Hecht-Nielsen 1989; Hornik et al. 1989). In practice an arbitrary good approximation of the intrinsic mapping is never reached due to the limiting size (i.e. degree of freedom) of the network function in terms of weights, the limiting size of our training database and the possible inability of the optimization algorithm to find the global minimum within the high-dimensional model parameter space of the network function (1). Nevertheless we can get in many cases a satisfying result and a good approximation of the intrinsic mapping for a wide range of possible input parameters. The other remarkable theoretical result is that an MLP

trained as classifier is able to approximate the Bayesian posterior probability (Richard & Lippmann 1991) which make them a powerful tool in classification problem.

Both parameter estimation (regression) and classification can be seen as a particular case of function approximation (Richard & Lippmann 1991). In a regression problem (i.e. parameter estimation) the task is to find an approximation of an unknown mapping on the basis of a continuous input and target relation. In a classification problem the task is to assign inputs to a number of discrete and possible mutually exclusive classes or categories, to yield an approximation of the underlying discriminant function.

3. The inference of temperature information from Stokes profiles

The temperature is the dominating factor in the radiative transfer process, it affects the appearance of the resulting Stokes spectra in various ways. The paramount role of the temperature can be demonstrated in terms of a first order perturbation of the radiative transfer equation (e.g. response functions). The temperature response of the Stokes parameters, Fig. 1, is at least an order higher than the responses from other parameters. A resulting variation of the Stokes vector \mathbf{I} , for each wavelength, from a perturbation of different parameters $x_i(\tau)$, in optical depth, can be expressed in first order as,

$$\delta\mathbf{I}(0) = \sum_{i=1}^l \int_0^{\infty} R_i(\tau) \delta x_i(\tau) d\tau. \quad (4)$$

A general expression for the response functions $R_i(\tau)$ of the Stokes parameters for a perturbation of the physical parameters $x_i(\tau)$ is given by Sanchez Almeida (1992),

$$R_i(\tau) = \mathbf{O}(0, \tau) \left[\mathbf{K}(\tau) \frac{\partial \mathbf{S}}{\partial x_i} - \frac{\partial \mathbf{K}}{\partial x_i} (\mathbf{I} - \mathbf{S}) \right], \quad (5)$$

where \mathbf{O} stands for the evolution operator defined in the formal solution of the RTE (Landi Degl'Innocenti & Landi Degl'Innocenti 1985). \mathbf{K} represents the absorption matrix and \mathbf{S} is the source function vector. One of the tasks of the MLP is to identify these variations and to disentangle the effects produced by each parameter perturbation δx_i .

To avoid possible problems, due to the dominating effect of the temperature, for the convergence and accuracy of the MLP during the training phase we decided to separate the problem of estimating the temperature information from the inference of the other atmospheric parameters. To retrieve some of the information from the underlying temperature we analyse, in a first step, the Stokes profiles with a neural network classifier. The classifier determines possible similarities to a set of semi-empirical model atmospheres.

The main source of information for a classification of the underlying temperature structure is given by the variation of the complete intensity spectrum, (Fig. 1) shows the

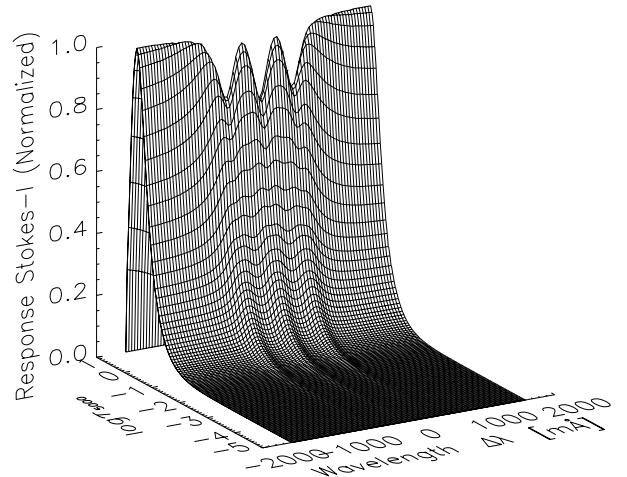


Fig. 1. The temperature response function of Stokes I for the Fe I $\lambda 15648$ line, calculated with the umbral model of Maltby et al. (1986).

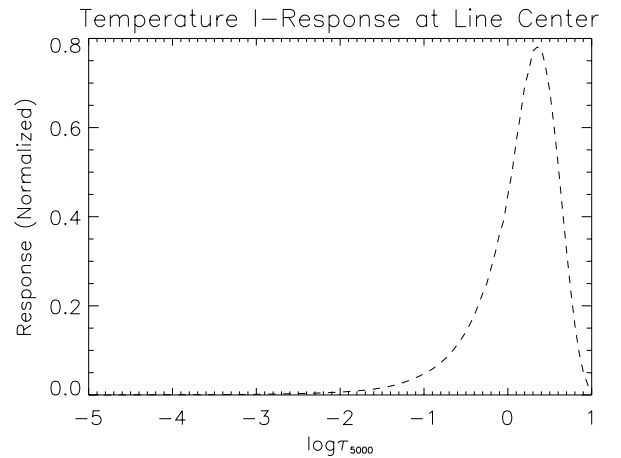


Fig. 2. The line center temperature response function of Stokes I for the Fe I $\lambda 15648$ line, calculated with the umbral model of Maltby et al. (1986).

temperature response for the Stokes I parameter. Since the response functions for the line cores have the slowest trend to zero with height, the line depth variation relative to the continuum, does also indicate a relative variation of the temperature gradient for the classifier (Figs. 1 and 2). To retain most of the information of the intensity variation we did not normalize our spectra by the local continuum, instead we have used the HSRA quiet-sun atmosphere from Gingerich et al. (1971) as an independent reference model.

3.1. The classification network

In this section we describe the basics of the training process for the classifier. To yield an estimate of the temperature structure from a given Stokes profile we train our MLP model on the basis of nine semi-empirical model atmospheres (temperature and pressure stratification). We have used three basic semi-empirical models, the umbra

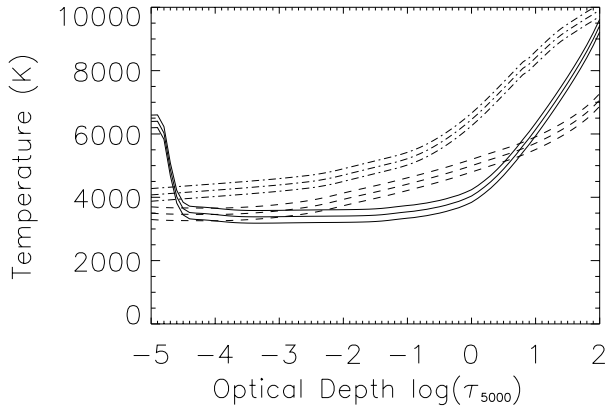


Fig. 3. Temperature stratification of the semi-empirical models: Maltby, Maltby+200, Maltby-200 (solid line), DingFang, DingFang+200, DingFang-200 (dotted line), Holweger, Holweger+200, Holweger-200 (dashed line).

model M from Maltby et al. (1986), the penumbral model from Ding & Fang (1989) and the quiet-sun model from Holweger & Mueller (1974). From each basic model we have calculated two variations in hydrostatic equilibrium with a uniform temperature offset of plus 200 and minus 200 Kelvin respectively. The temperature stratifications of the resulting nine model atmospheres are shown in Fig. 3. With this set of atmospheres we get a coarse grid of possible temperature structures. Our classifier should be able to find similarities of a given profile to a set of trained profile which were calculated on the basis of one of the nine semi-empirical model atmospheres. For each model the similarities are expressed in terms of a likelihood that this particular model has been responsible for producing the observed profile.

To create the database for the classification network we have used the DIAMAG synthesis code (i.e. forward model) from Grossmann-Doerth (1994) which implements a numerical integration of the radiative transfer equation (RTE) by means of the diagonal lambda iteration method (DELO) (Rees et al. 1989). For the photospheric Fe I $\lambda 15648.5$ we calculated the Stokes parameters in a wavelength range of -160 pm to $+160$ pm around the line center, with a sample of 4 pm, such that for each discrete Stokes parameter there are 81 wavelength points available. For the classification of the temperature model we have used the Stokes I and V parameters only because we did not expect any further information about the temperature variation from Q and U . That means, we have used 162 data points as an input vector to our classification network. The simulated profiles are assumed to be normalized to the HSRA quiet-sun atmosphere from Gingerich et al. (1971). On the basis of the nine semi-empirical model atmospheres we produced a data set of 9000 Stokes profiles for the training set (1000 for each model to account for an equal prior probability) and 9000 Stokes profiles for a validation set to check the performance of the network with an independent data set during the training process.

Table 1. Intervals for the atmospheric parameters.

Atmospheric Parameter	Range
Mag. Field Strength	[500 G–3500 G]
Inclination	[0°–180°]
Azimuth	[0°–175°]
LOS Velocity	[−2000 m s ^{−1} –+ 2000 m s ^{−1}]
Microturbulence	[0 m s ^{−1} –1800 m s ^{−1}]
Macroturbulence	[0 m s ^{−1} –1800 m s ^{−1}]

To take into account the full variability of all the other free atmospheric parameters, a random generator first determined the atmospheric values of a single-valued atmosphere for the magnetic field strength, inclination, azimuth, LOS velocity, microturbulence and macroturbulence. The values are taken from the intervals given in Table 1. In a second step another random generator picks out the temperature and pressure stratifications from one of the nine semi-empirical models. With this procedure we calculated the entire database for the training and the validation samples. To meet the demands of a more realistic inversion we added white noise to the synthetic Stokes profiles with a signal-to-noise ratio of 1000 (relative to the continuum). We applied a linear transformation (zero mean and unit standard deviation for each wavelength point) with respect to the training database. For the output variables of the MLP we adopt a typical 1-of- c coding scheme for the classification networks. The number c of output elements of the MLP corresponds to the number of categories we want to distinguish. We have used nine semi-empirical temperature models which means that each of our MLPs has nine output elements and the training and validation samples are also associated with nine target values. If a given pair of I and V profiles is formed under a specific temperature model, the corresponding target value in the training set is set to 1 (i.e. 100% probability) and the other eight are set to zero.

The training data set is then used to train different network architectures (with different numbers of hidden layers and elements in these layers) with a standard error Back-propagation algorithm (Rumelhart et al. 1986), together with a regularized conjugate gradient algorithm.

3.2. Results of the temperature classification

If the MLP has found a good approximation of the underlying discriminant function of the classification problem the activations of the output units should provide a direct estimate of the posterior probabilities $P(C_k|\mathbf{X})$ for a particular class (i.e. model atmosphere) C_k given the Stokes profile \mathbf{X} (Richard & Lippmann 1991; Bishop 1995). Since our classes are not mutually exclusive, which means that an observed profile may carry attributes from temperature stratifications of different semi-empirical models, the MLP classifies multiple independent attributes. Each single output activation (i.e. class activation) represents the

probability that the attributes of a particular temperature model are present.

The network which shows the best performance in the training session in terms of the error function (3) was taken to statistically determine the performance of the network. The free model parameters (i.e. weights) of the MLP are frozen and the MLP is now in application mode. In order to test the performance of the MLP in identifying the most probable model atmosphere for a given Stokes profile we have calculated a test database of 18 000 Stokes profiles (2000 for each semi-empirical model). The interpretation of the network results follow the Bayesian approach for classification, where minimizing the probability of misclassification is achieved by the selection of the largest posterior probability (Bishop 1995). In this way we interpret the largest output activation as a decision for the model which is represented by that particular output unit.

By interpreting the network results in this way the MLP could identify, in all 18000 cases, the correct semi-empirical model!!! For every single pair of Stokes I and V in the test database the neural network classifier was able to determine the correct underlying temperature structure. Figure 4 shows the distribution of the probabilities (output activation) for the corresponding correct model, the summed probability of all other output elements (classes) is in all cases less than 5%, this indicates the high certainty of the MLP decisions. The MLP has found distinct characteristics for each temperature model in the Stokes profiles, and this although the other free parameters like magnetic field strength, inclination, azimuth, LOS velocity, microturbulence and macroturbulence are randomly changed.

In a realistic application, none of the nine temperature models will match exactly the original temperature stratification of a given observed Stokes profile. As an example how the MLP classifier reacts in the presence of a profile which has its origin from an unknown temperature stratification we consider a set of Stokes profiles which were calculated on the basis of the umbral model of Kollatschny et al. (1980). The temperature stratification of the Kollatschny model is shown in Fig. 5 (solid line) together with the three trained umbral models Maltby, Maltby+200, Maltby-200 (dashed lines). Since the temperature stratification of the Kollatschny model shows different good agreements with these three umbral models, depending on the optical depth, we expect that profiles calculated for the Kollatschny model comprises attributes from all three umbral models. We calculated a set of 2000 Stokes I and V profiles for the Kollatschny model, the atmospheric parameters for each calculation are again chosen by a random generator from the parameter intervals given in Table 1.

Again we get a very impressive result, in all 2000 cases only the output units which represents the three umbral models showed activation, all the others remained silent. In all 2000 test cases the MLP classifier clearly favours the original Maltby atmosphere. In Table 2 we

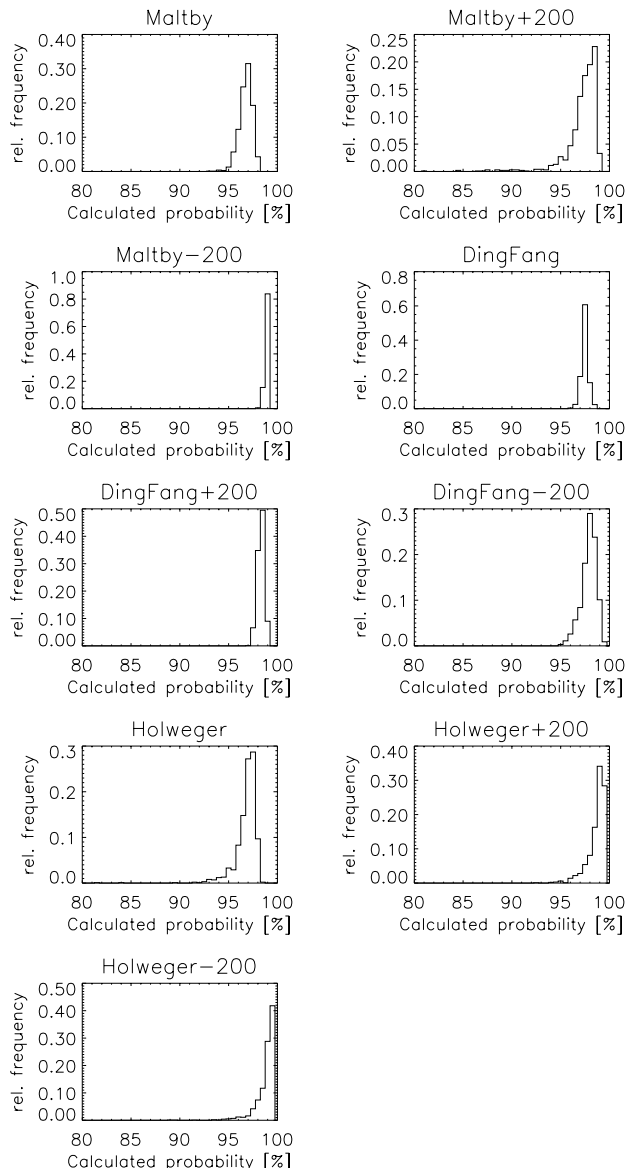


Fig. 4. Calculated distribution of the network output activity for the true semi-empirical model.

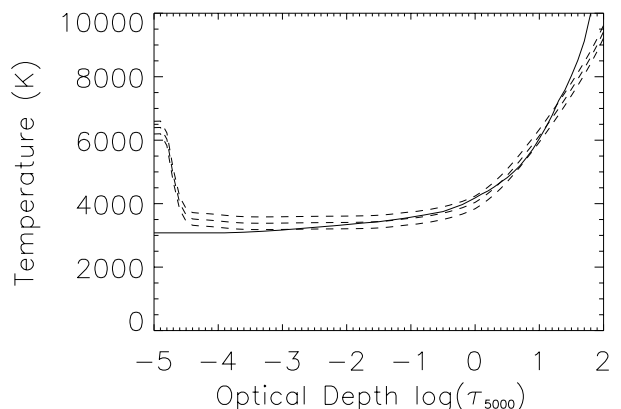


Fig. 5. Temperature stratification of the Kollatschny model (solid line) and the three trained umbral models, Maltby-200, Maltby, Maltby+200 (all dashed lines).

Table 2. Average output activation for each semi-empirical model.

Model	Average Output
Maltby-200	0.011
Maltby	0.868
Maltby+200	0.123
DingFang-200	0.000
DingFang	0.000
DingFang+200	0.000
Holweger-200	0.000
Holweger	0.000
Holweger+200	0.000

listed the average activation of each output unit for all 2000 test profiles. The original Maltby model has an average activation of 0.868 (or 86.8% probability that attributes from that temperature model are present), the Maltby+200 shows an average activation of 0.123 (12.3%) and the Maltby-200 has an activation of 0.011 (1.1%). This clearly shows that the classifier has found most similarities with Stokes profiles that were calculated under the original Maltby model. From Fig. 2, the line center response of Stokes I , we can see that the classifier does indeed provide reasonable results, the Stokes I line center response has its highest values in a depth range of about $\log(\tau) = +0.7$ to $\log(\tau) = -0.3$ right where the temperature stratification of the Kollatschny model runs just slightly above the temperature stratification of the original Maltby model (Fig. 5). Therefore the original Maltby temperature model can be used as an appropriate approximation of the Kollatschny model. Note that this is just valid for the considered Fe I $\lambda 15648$ line, a different line, with a different temperature sensitivity in optical depth, might be better approximated by a different temperature model.

Although the grid of available semi-empirical models is very coarse and should be increased in a realistic application to better constrain the real temperature stratification, the simple neural classifier has demonstrated that it can yield fast and reliable classifications of the underlying temperature model.

4. The inversion of Stokes parameter

One of the major goals of the Stokes profile inversion is the inference of the magnetic field and dynamic structure of the solar atmosphere. The Stokes parameters formalism provides a description of the Zeeman induced splitting and polarization properties of magnetic sensitive absorption lines. A formal solution of the polarized RTE given by Landi Degl'Innocenti & Landi Degl'Innocenti (1985) represents the forward model from which we want to infer some of the underlying atmospheric parameters. The emergent (observed) Stokes vector \mathbf{I} for a particular wavelength at, $\tau = 0$ is given by

$$\mathbf{I}(0) = \int_0^\infty \mathbf{O}(0, \tau) \mathbf{K}(\tau) \mathbf{S}(\tau) d\tau, \quad (6)$$

where \mathbf{O} stands for the evolution operator \mathbf{K} for the absorption matrix and \mathbf{S} for the source function vector. All of the considered atmospheric parameters are constituents of the absorption matrix \mathbf{K} which is also part of the evolution operator. There is no general analytical solution for the inverse problem of retrieving the atmospheric parameters from the observed Stokes profiles. Therefore the inverse problem is solved by transforming it to a nonlinear optimization problem. To circumvent the conventional iterative approach of the optimization process, we approximate the inverse relation with the MLP model and use the approximate inverse mapping (i.e. trained MLP) for a direct inversion of observed Stokes profiles.

4.1. Training of the regression networks

For each semi-empirical model atmosphere we have trained an MLP as a regression network. We again calculated a training database and a validation set for each network with the DIAMAG synthesis code for the Fe I $\lambda 15648$ line. The training set and validation set consist of 8000 patterns (Stokes profiles to physical parameters). The Stokes profiles are calculated in a wavelength range of -160 to $+160$ pm around the line center, with a sample of 1 pm, to make 81 wavelength points available for each discrete Stokes parameter. In order to constrain the numbers of free parameter of the MLPs we only make use of Stokes I and V for the determination of the magnetic field strength, inclination, LOS velocity, microturbulence and macroturbulence and have trained an extra network for the determination of the field azimuth which uses the Stokes Q and U parameters. Thus each regression network has a input vector of 162 elements. The networks for the determination of the azimuth angle have one single output while the others have five, one for each parameter. To account for an appropriate representation of the inverse mapping, for each calculated Stokes profile a random generator determines the values for the different free atmospheric parameters from the intervals given in Table 1. Once more noise with a signal-to-noise ratio of 1000 was added and a linear transformation of the input values (zero mean and unit standard deviation for each wavelength point) with respect to the training database was applied. The training data set is used to train different network architectures with a standard error Back-propagation algorithm together with a regularized conjugate gradient algorithm.

4.2. Results of the inversion

In order to limit the description of the results we demonstrate the good performance by presenting the results of the two MLPs which were trained on the basis of the original Maltby atmosphere. The other MLPs trained on the other semi-empirical temperature models are showing the same good results in terms of root-mean-square (rms) error and error distribution. Therefore we consider

in the following two MLPs, one for the estimation of the magnetic field strength, inclination, LOS velocity, microturbulence and macroturbulence (MLP-M-5 hereafter) and another one which only retrieves an estimate of the azimuthal angle of the magnetic field (MLP-M-1 hereafter). As for the classification network we have chosen the network which shows the best performance in the training session in terms of the error function (3), to statistically determine the performance of the inversion with an independent test data set.

For the independent test database we have calculated 10000 Stokes profiles under the same conditions as the training data set. Since we are now in the application mode all the free parameters of the network (i.e. weights) are frozen and the MLP represents a simple and fast to evaluate nonlinear function. The presentation of a Stokes I and V profile to the MLP-M-5 network yields a simultaneous calculation of all 5 atmospheric parameters within a fraction of a second.

For the complete test dataset of 10000 Stokes profiles we yield an rms error for the magnetic field strength of 24.02 Gauss, the distribution of the error and a plot of the calculated field strength values versus the true (known) values, are shown in Fig. 6. These very good results demonstrate that the MLP could recognize the characteristics of the magnetic field strength within the Stokes I and V profiles. The mapping of the profiles to the field strength values was performed with high accuracy (see Fig. 6 left). From the right plot in Fig. 6 we see that for the whole trained field strength range (500 G to 3500 G) the estimated values show a very dense alignment along the ideal line which indicates that the approximate mapping for the field strength has a very low variance over the entire trained parameter range.

For the inclination angle we yield an rms error of 2.01° . The distribution of the error for all 10000 test profiles are shown Fig. 7. The variation of Stokes I and V due to a change of the inclination angle is remarkably good recognized and quantified by the MLP. Over the entire parameter range the inclination angles are very accurately estimated by the MLP.

The LOS velocity was also very accurately estimated by the MLP, the relative line shifts of the Stokes I and V profile are easily detected by the MLP (Fig. 8). The determination of the LOS velocity for all 10000 Stokes I and V profiles yields an rms error of only 28.81 ms^{-1} .

Another remarkable result is the fact that the MLP is able to disentangle the non-thermal broadening effects of both microturbulence and macroturbulence. Although a small perturbation of both effects shows a rather small variation in the observed Stokes profiles (i.e. small response functions) the MLP has learned to identify even fairly small variations. This is expressed in the fact that the rms error of the retrieved microturbulence, for all 10000 test profiles, has a value of only 103.74 ms^{-1} . The macroturbulence could even be determined with an rms error of 57.62 ms^{-1} . Figures 9 and 10 show again the distributions of the error.

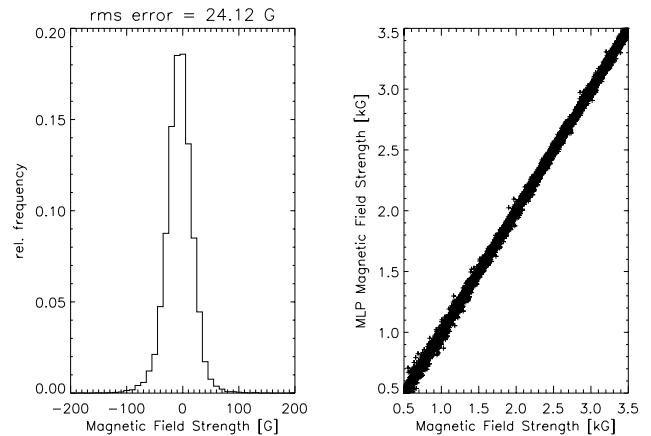


Fig. 6. Distribution of the deviation from the true field strength (left) and a plot of MLP calculated versus the true field strength (right). The calculated values are marked by a “+” sign.

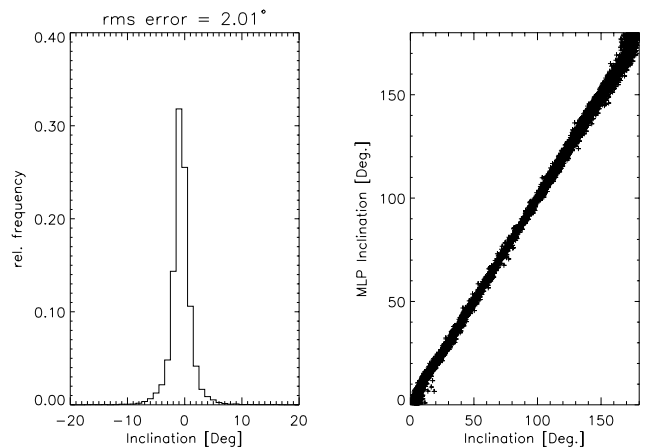


Fig. 7. Distribution of the deviation from the true inclination (left) and a plot of MLP calculated versus the true inclination angle (right).

For the test of the MLP-M-1 network which determines the azimuthal angle we have generated another test database of 10000 Stokes Q and U profiles. The 6 free atmospheric parameters are again chosen from the intervals shown in Table 1, with the exception that we have shortened the range of possible inclination angles. The inclination angles are now chosen from 15° to 165° to provide a sufficient signal in the Stokes Q and U components. In Fig. 11 the error distributions for the complete test data set are shown. The determination of the azimuth could not reach the accuracy of the inclination, but taking into account the presence of the added artificial noise and the relative weakness of the Stokes Q and U signals compared to Stokes I and V , the retrieval of the azimuthal angle for all test profiles is still remarkable. The rms error of the azimuth for all 10000 test profiles is 6.45° .

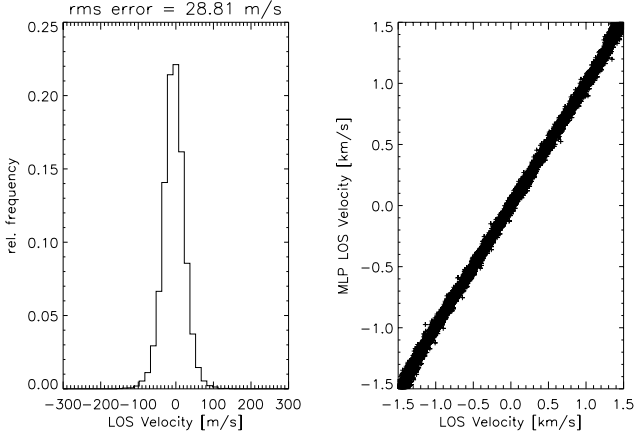


Fig. 8. Distribution of the deviation from the true velocities (left) and a plot of MLP calculated versus the true LOS velocities (right).

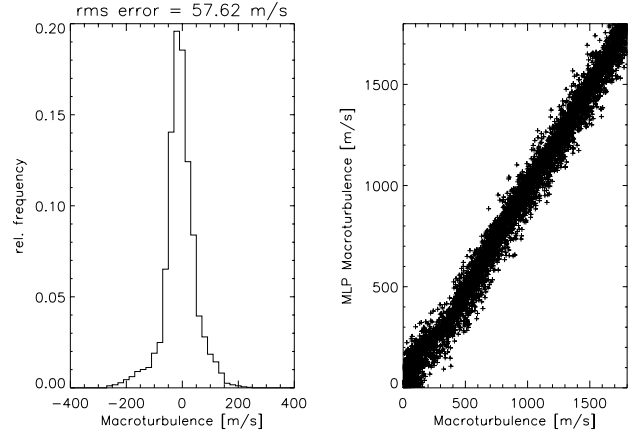


Fig. 10. Distribution of the deviation from the true macroturbulences (left) and a plot of MLP calculated versus the true macroturbulences (right).

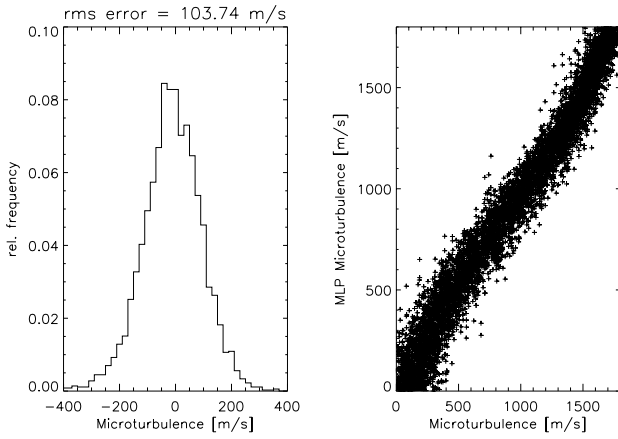


Fig. 9. Distribution of the deviation from the true microturbulences (left) and a plot of MLP calculated versus the true microturbulences (right).

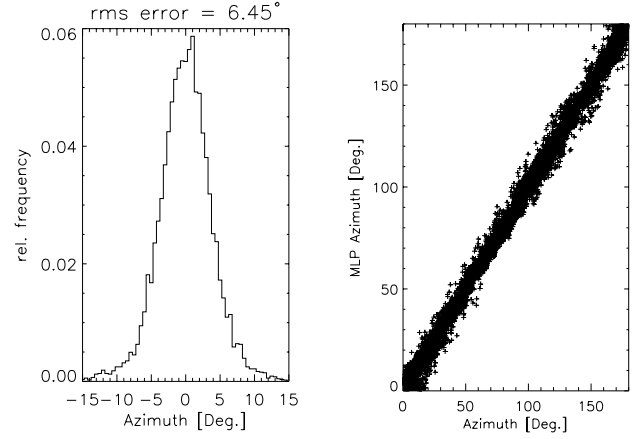


Fig. 11. Distribution of the deviation from the true azimuth (left) and a plot of MLP calculated versus the true azimuthal angle (right).

5. The presence of gradients

Since our MLP models for the inversion of Stokes parameters are trained on the basis of homogeneous single-valued atmospheres for the magnetic field vector, LOS velocity, microturbulence and macroturbulence we are interested to see how the nonlinear approximation of the MLP react in the presence of a profile which has originated from a stratified atmosphere. To test this behaviour we calculated a set of 1000 Stokes profiles with different gradients in the magnetic field strength and LOS velocity. We have adopted various linear dependencies of the field strength and the velocities with $\log(\tau)$. The gradients for the magnetic field strength are changed in a range from 50 G/ $\log(\tau)$ to 500 G/ $\log(\tau)$. For the line forming region of the Fe I $\lambda 15648$ line in the Maltby atmosphere this corresponds to a gradient of about 0.5 G/km to 5 G/km. The gradients for the LOS velocity are changed in a range from 20 m/s/ $\log(\tau)$ to 200 m/s/ $\log(\tau)$. Apart from any physical meaning we adopted a positive and a negative gradient for both parameters. A variable offset in $\log(\tau) = 0$ was also provided.

Since our expectation is that the network, trained on the basis of single-valued parameter, should provide us with a kind of averaged values for the stratified atmosphere, we need a method which yields an appropriate estimation of these averaged values. For this reason, we follow the definition of heights of formation (HOFs) for a measurement from Sanchez Almeida et al. (1996). They consequently related the determination of the HOFs with the physical parameter under consideration, the Stokes parameter and the underlying atmosphere. This relation is expressed in terms of the response functions R_{ij} for the corresponding Stokes parameter i and the considered physical parameter j . To take into account both Stokes I and V and the whole wavelength range, $+p$ to $-p$, we calculated the mean depth given in $\log(\tau_m)$ with:

$$\log(\tau_{mj}) = \sum_{i=1}^2 \frac{\int_{-p}^{+p} \int_0^{\infty} R_{ij}(\lambda, \tau) \log(\tau) d\tau d\lambda}{\int_{-p}^{+p} \int_0^{\infty} R_{ij}(\lambda, \tau) d\tau d\lambda}. \quad (7)$$

For all 1000 test profiles we have calculated the HOFs for the magnetic field strength and the LOS velocity with (6) from the stratified atmosphere. To yield an averaged value

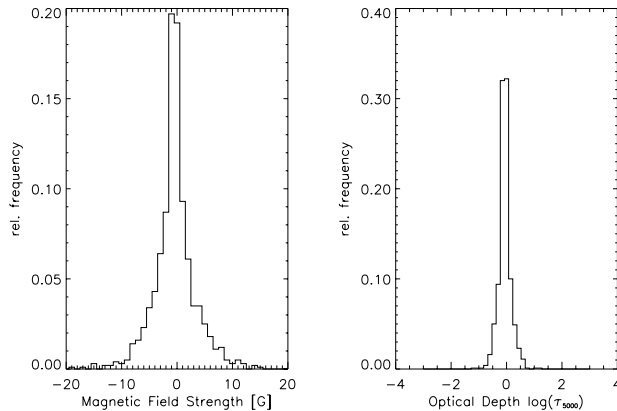


Fig. 12. Distribution of the deviation of the MLP-M-5 results from the averaged magnetic field strength (left). Distribution of deviations from the HOFs (right).

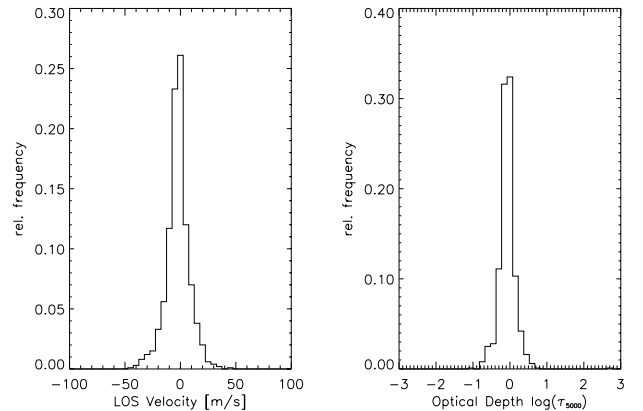


Fig. 13. Distribution of the deviation of the MLP-M-5 results from the averaged LOS velocity (left). Distribution of deviations from the HOFs (right).

for the field strength or velocity for the given stratified atmosphere we interpret the field strength or velocity at the corresponding HOFs, as an estimate for that value. The retrieved parameters of the MLP are compared to these averaged values.

We found extremely good agreement between the MLP calculations and the determined averaged field strength and velocity parameter. To illustrate how good these agreements are, in Fig. 12 (left) we have depicted the deviations of the retrieved magnetic field strength by the MLP-M-5 from the calculated averaged field strength. Figure 13 (left) shows these deviations for the retrieved LOS velocities. For a better comparison of how these deviations are related to a difference in optical depth, for each retrieved field strength and velocity of the MLP-M-5, we determined the position in optical depth, obtained by an interpolation from the original field strength and velocity stratification. With this procedure we have calculated the deviation, in optical depth, for each retrieved field strength and velocity, from the related HOFs. Figure 12 (right) and Fig. 13 (right) illustrates these deviations.

6. The determination of the filling factor

Most magnetic structures outside of sunspots are represented by small unresolved magnetic elements (i.e. magnetic flux tubes) (e.g. Stenflo 1989). Therefore the estimation of the fraction of the magnetic structure provides a valuable parameter and puts additional physical constraints on the inference of the other atmospheric parameters. We have adopted a simple two-component atmosphere which has a magnetic and a non-magnetic component. This simple two-component model is also a valid representation for the contribution of unpolarized stray light. The observed Stokes vector is then given by

$$\mathbf{I}^{\text{Observ}} = (1 - \alpha)\mathbf{I} + \alpha\mathbf{I}^{\text{Mag}}, \quad (8)$$

where $(1 - \alpha)$ does represent the factor of the non-magnetic or stray light component. For the following application of

the MLP, our focus is to demonstrate that an MLP is in principle able to recognize the presence of an additional non-magnetic component and to estimate the filling factor.

For the calculation of the training and validation data set we assumed the temperature structure of the quiet-sun model of Holweger & Mueller (1974) for the non-magnetic or stray light component and for the magnetic component we assumed the temperature structure from the Maltby et al. (1986) atmosphere. The Stokes parameters for the different components are calculated on the basis of the intervals given in Table 1. This means we have 6 free parameters for the magnetic component and 3 free parameters for the non-magnetic component. For the resulting Stokes vector we have one additional free parameter, the filling factor for the magnetic component, which is adjusted between 0.05 (5%) and 1.0 (100%). Since we are now only interested in the fraction of the magnetic or non-magnetic component our MLP has just one output element which represents the filling factor α . The free parameters are again determined by a random generator to produce a training and a validation set of 5000 synthetic Stokes profiles. As input we have used the Stokes I and V parameter. Therefore we have excluded inclination values within a range of 75° to 105° to the LOS to retain a sufficient Stokes V signal.

The training data set is then used to train different network architectures (with different numbers of hidden layers and elements in these layers) with a standard error Back-propagation algorithm (Rumelhart et al. 1986), together with a conjugate gradient algorithm. After a successful training the network which has the best performance in the training session is used to statistically determine the behaviour and accuracy of the estimations. Therefore we calculated another test set of 5000 synthetic Stokes profiles under the same conditions like the training data set.

The whole test set of 5000 profiles was statistically investigated and does once more demonstrate the remarkable abilities of the MLP. The filling factor could be

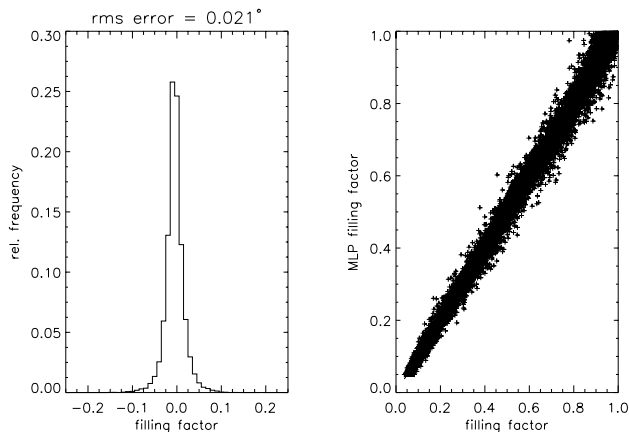


Fig. 14. Distribution of the deviation from the true filling factor (left). A plot of MLP calculated versus the true filling factors (right).

determined by the trained MLP with an rms error of only 0.021 (2.1%). Figure 14 shows again the distribution of the deviations from the true values (left) and a plot of the MLP estimates over the true filling factors (right) for all 5000 test profiles. Despite the variation of all the other free atmospheric parameters and the superimposed noise the trained MLP is able to determine the fraction of the magnetic component or non-magnetic component, respectively, with a very good accuracy.

7. The speed of the inversion

One of the striking features of the MLP is, that once it is adapted to the problem, the network parameters are frozen (application mode) and the network simply represents a mathematical function, see (1), which allows a fast evaluation either as a software program or as a hardware (VLSI) implementation. We implemented a software simulation of our neural networks (written as C++ code) which runs on a fairly ordinary Pentium III processor under the Linux operating system. The CPU time needed for the software simulated MLPs to invert 10 000 Stokes profiles (I and V or Q and U) is less than 10 s, which is many orders of magnitudes faster than any conventional inversion routine based on an iterative optimization and still significantly faster than the PCA based inversion of Stokes profiles proposed by Socas-Navarro et al. (2001).

8. Conclusions

We have investigated the ability of Multi-Layer Perceptrons for the inversion of Stokes profiles. We have used synthetic observations of the Fe I $\lambda 15648$ line. The application of IR line offers valuable advantages for the diagnostic of magnetic fields (Solanki 1992) but the application of ANNs are not limited to the IR lines, the application to other visible lines like the Fe I $\lambda 6302.5$ and Fe I $\lambda 5250.2$ lines (Carroll et al. 2000) revealed very similar results as presented in this paper.

The classification of Stokes profiles for the estimation of the underlying temperature structure on the basis of nine semi-empirical model atmospheres has demonstrated the remarkable classification properties of the MLP. Although the IR line has a rather small temperature sensitivity compared to other visible lines, the neural classifier was able to find distinct features within the profiles and to determine the correct underlying atmosphere or the most probable respectively. For a real application of that classifier the number of available semi-empirical models should be increased, but this is straightforward and no principle problem for the MLP classifier.

For the inference of atmospheric parameters from the Stokes profiles we have trained an MLP on the basis of an appropriate numerical forward model. Despite the presence of an artificial signal-to-noise ratio of 1000 the trained MLPs found a good approximation of the inverse mapping of the problem. For all free parameters, magnetic field vector, LOS velocity, microturbulence and macroturbulence, the retrieved estimates from the Stokes profiles are very close to the original true values.

Since the MLPs are trained on a single-valued atmosphere for the considered free parameters we investigated the results of the inversion in the presence of moderate gradients for the magnetic field strength and LOS velocity. These investigations reveal that the estimated single-valued parameters of the MLP do represent very accurate averaged values for the stratified atmosphere.

For the inference of an unresolved magnetic component or the presence of unpolarized stray light, a MLP could be trained on a two-component model and has retrieved remarkable good estimates of the underlying fraction of the magnetic or non-magnetic component respectively.

Although our investigation are based on synthetic observations it was demonstrated that the Multi-Layer Perceptron as one well known type of artificial neural networks is a suitable model for the inversion of Stokes profiles. Because of their speed, noise tolerance and nonlinear adaption abilities MLPs and other supervised ANN models provide us with a powerful statistical tool to tackle the problem of an automated and stable inversion of large amounts of polarimetric data.

ANNs offer a variety of possible outlooks for the analysis and interpretation of Stokes profiles, in forthcoming work we will extend the application to a multi-line inversion to improve the diagnostic capabilities and a possible inference of gradients for a stratified atmosphere. The coupling of ANNs with a conventional inversion procedure based on response functions is already in progress. For more complicated atmospheric structures (e.g. multi-component or interlaced atmospheres) which might contain inherent ambiguities, an ANN based modeling of mixture of experts (Jacobs et al. 1991) or mixture of models (Bishop 1995) offer a promising approach to deal in a probabilistic way with multi-valued inverse problems. Further outlooks are the forward modeling of the polarized RTE with ANNs and a fast approximation of response functions. Another interesting application might be the

auto-associative MLP (Kramer 1991) to perform a non-linear dimensionality reduction, this offers a nonlinear extension to the linear decomposition of Stokes profiles (Rees et al. 2000) which is performed by a principle component analysis.

Acknowledgements. The authors gratefully acknowledge financial support of the present work by the German Federal Ministry of Education and Research through the German Space Research Center (DLR) under grant No. 50 QL 0003.

References

- Auer, L. H., Heasley, J. M., & House, L. L. 1977, *Solar Phys.*, 55, 47
- Bailer-Jones, C. A. L., Irwin, M., & von Hippel, T. 1998, *MNRAS*, 298, 361
- Bailer-Jones, C. A. L. 2000, *A&A*, 357, 197
- Bishop, C. M. 1995, *Neural Networks for Pattern Recognition* (Oxford University Press)
- Calvo, R. A., Ceccato, H. A., & Piacentini, R. D. 1995, *ApJ*, 444, 916
- Carroll, T. A., Balthasar, H., Muglach, K., & Nickelt, I. 2000, *Advanced Solar Polarimetry - Theory, Observations and Instrumentation*, ed. M. Sigwarth (San Francisco: ASP), in press
- Ding, M. D., & Fang, C. 1989, *A&A*, 225, 204
- Fessant, F., Pienret, C., & Lantos, P. 1996, *Solar Phys.*, 168, 423
- Funahashi, K. 1989, *Neural Networks*, 2, 183
- Gingerich, O., Noyes, R. W., Kalkofen, W., & Cuny, Y. 1971, *Solar Phys.*, 18, 347
- Grossmann-Doerth, U. 1994, *A&A*, 285, 1012
- Gulati, R. K., Gupta, R., Gothoskar, P., & Khobragade, S. 1994, *ApJ*, 426, 340
- Hecht-Nielsen, R. 1989, *Proceedings of the International Joint Conf. on Neural Networks*, 1, 593
- Holweger, M., & Mueller, E. A. 1974, *Solar Phys.*, 39, 19
- Hornik, K., Stinchcombe, M., & White, H. 1989, *Neural Networks*, 2, 359
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. 1991, *Neural Comp.*, 3, 79
- Keller, C. U. 1998, *Proc. SPIE*, 3352, 732
- Kjeldseth-Moe, O. 1969, *Solar Phys.*, 8, 23
- Kollatschny, W., Wiehr, E., Stellmacher, G., & Falipou, M. A. 1980, *A&A*, 86, 245
- Kramer, M. A. 1991, *AI Journal*, 37(2), 233
- Landi Degl'Innocenti, E., & Landi Degl'Innocenti, M. 1985, *Solar Phys.*, 97, 239
- Landolfi, M., Landi Degl'Innocenti, E., & Arena, P. 1984, *Solar Phys.*, 93, 269
- Lites, B. W. 2001, *Advanced Solar Polarimetry - Theory, Observations and Instrumentation*, ed. M. Sigwarth (San Francisco: ASP), in press
- Maltby, P., Avrett, E. H., Carlson, M., et al. 1986, *ApJ*, 306, 284
- Rees, D. E., Murphy, G. A., & Durrant, C. J. 1989, *ApJ*, 339, 1093
- Rees, D. E., Lopez Ariste, A., Thatcher, J., & Semel, M. 2000, *A&A*, 355, 759
- Richard, M. D., & Lippmann, R. P. 1991, *Neural Comp.*, 3, 461
- Ruiz Cobo, B., & Del Toro Iniesta, J. C. 1992, *ApJ*, 398, 375
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986, *Parallel Distributed Processing*, 1, 318
- Sanchez Almeida, J. 1992, *Solar Phys.*, 137, 1
- Sanchez Almeida, J., Ruiz Cobo, B., & del Toro Iniesta, J. C. 1996, *A&A*, 314, 295
- Skumanich, A., & Lites, B. W. 1987, *ApJ*, 322, 473
- Socas-Navarro, H., Lopez Ariste, A., & Lites, B. W. 2001, *ApJ*, 553, 949
- Solanki, S. K., Ruedi, I., & Livingston, W. 1992, *A&A*, 263, 312
- Stenflo, J. O. 1989, *A&AR*, 1, 3
- del Toro Iniesta, J. C., & Ruiz Cobo, B. 1996, *Solar Phys.*, 164, 169
- von der Luehe, O., Schmidt, Soltau, D. W., Kneer, F., & Staude, J. 1999, in *ESA SP-463, Euroconf. Solar Cycle and Terrestrial Climate*, ed. A. Wilson
- Weaver, B. W., & Torres-Dodgen, A. V. 1997, *ApJ*, 487, 847