A&A 438, 1021–1028 (2005)
DOI: 10.1051/0004-6361:20052865
© ESO 2005

Astronomy & Astrophysics

An artificial neural network approach to the solution of molecular chemical equilibrium

A. Asensio Ramos[1] and H. Socas-Navarro[2]

[1] Istituto Nazionale di Astrofisica (INAF), Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Florence, Italy
e-mail: aasensio@arcetri.astro.it
[2] High Altitude Observatory, NCAR*, 3450 Mitchell Ln, Boulder CO 80307-3000, USA

Received 11 February 2005 / Accepted 19 April 2005

Abstract. A novel approach is presented for the solution of instantaneous chemical equilibrium problems. The chemical equilibrium can be considered, due to its intrinsically local character, as a mapping of the three-dimensional parameter space spanned by the temperature, hydrogen density and electron density into many one-dimensional spaces representing the number density of each species. We take advantage of the ability of artificial neural networks to approximate non-linear functions and construct neural networks for the fast and efficient solution of the chemical equilibrium problem in typical stellar atmosphere physical conditions. The neural network approach has the advantage of providing an analytic function, which can be rapidly evaluated. The networks are trained with a learning set (that covers the entire parameter space) until a relative error below 1% is reached. It has been verified that the networks are not overtrained by using an additional verification set. The networks are then applied to a snapshot of realistic three-dimensional convection simulations of the solar atmosphere showing good generalization properties.

Key words. molecular processes – astrochemistry – methods: numerical

1. Introduction

Molecules are usually found in highly dynamic systems (e.g., the solar atmosphere, winds of AGB stars, ...) and their formation is influenced by the time variation of the physical conditions in the medium. In many situations the dynamical timescales are much slower than the timescales of molecular formation, and the approximation of Instantaneous Chemical Equilibrium (ICE) can be used with extremely good results (Russell 1934; Tsuji 1964, 1973; McCabe et al. 1979, etc.). Under the ICE approximation, molecules and ions are assumed to form instantaneously and their abundances depend only upon the local temperature and density. Another consequence of this assumption is that the specific reaction mechanisms that create and destroy a given molecule are irrelevant and only the molecule and its constituents are important. Concerning the solar atmosphere, it has been recently shown with the comparison between ICE calculations and detailed chemical evolution calculations that the CO abundance can be well described using the ICE approximation for heights below ∼700 km (Asensio Ramos et al. 2003). On the other hand, chemical evolution effects in regions of the atmosphere above this height are of importance in setting the CO abundance.

In order to calculate the atomic and molecular number densities using the ICE approximation one needs to solve a non-linear system of equations (see below). Although efficient numerical methods exist for solving this kind of systems, the computing time when the ICE problem has to be solved in a large amount of points (e.g., dense grids, multi-dimensional geometries, iterative inversions, etc) becomes prohibitive. It is then very important to develop a numerical method for the rapid solution of chemical equilibrium problems.

Artificial Neural Networks (ANNs) have proven to be a powerful approach to a broad variety of problems (see, e.g., Bishop 1996). In the solar community, they have been recently applied to the problem of inferring the magnetic field from observations of the polarization profiles of selected atomic lines (Carroll & Staude 2001; Socas-Navarro 2003, 2005). In this paper, we make use of the ability of ANNs with one hidden layer to approximate any non-linear continuous function (e.g., Jones 1990; Blum & Li 1991) to solve the ICE problem.

The structure of the paper is as follows. Section 2 describes the ICE approximation and its standard solution. Section 3 describes our approach to the ICE problem using ANNs, discussing how it can be trained to associate a combination of physical parameters with the number density of each species included in the problem. Section 4 details how the trained network can be applied to solve the ICE problem in realistic convection simulations of the solar atmosphere. In Sect. 5 we show

* The National Center for Atmospheric Research (NCAR) is sponsored by the National Science Foundation.

the dependence of the outputs of the ANN on the physical parameters, which can be easily done because of the intrinsic analytic character of the mapping generated by the ANN. Finally, the most relevant conclusions are summarized in Sect. 6.

2. Instantaneous chemical equilibrium

2.1. Basic equations

Consider three elements A, B and C that stick together to form a molecule $A_aB_bC_c$, in which a, b and c are the stoichiometric coefficients indicating the number of times an element is present in the molecule. Because the explicit reactions which form the molecule are irrelevant under the ICE approximation, one needs only to consider the following dissociative process for the formation of the molecule $A_aB_bC_c$:

$$A_aB_bC_c \Leftrightarrow aA + bB + cC. \tag{1}$$

This reaction is characterized by its dissociative reaction constant or equilibrium constant. It is given by the ratio between the product of the partial pressures of the individual elements and the partial pressure of the molecule:

$$K_\mathrm{p}(T) = \frac{P_A^a P_B^b P_C^c}{P_{ABC}}, \tag{2}$$

where $K_\mathrm{p}(T)$ is a function of the temperature and P_i the partial pressure of the species i. For example, the equilibrium constants for the dissociation of CO and H_2 are given by

$$K_\mathrm{p}^\mathrm{CO}(T) = \frac{P_\mathrm{C} P_\mathrm{O}}{P_\mathrm{CO}}, \qquad K_\mathrm{p}^{H_2}(T) = \frac{P_\mathrm{H}^2}{P_{\mathrm{H}_2}}\cdot \tag{3}$$

The partial pressure is usually related to the number density n_i of a given species by the ideal gas equation:

$$P_i = n_i kT, \tag{4}$$

where k is the Boltzmann constant and T the local temperature. This partial pressure is the pressure of the gas if no other species were present. The sum of all the partial pressures is the total pressure in the medium.

Following the same line of reasoning, one can consider the ionization equilibria for the atomic species:

$$\begin{aligned} &A \Leftrightarrow A^+ + \mathrm{e}^- \\ &A + \mathrm{e}^- \Leftrightarrow A^-, \end{aligned} \tag{5}$$

with their corresponding equilibrium constants:

$$K_{A^+}(T) = \frac{P_{A^+} P_{\mathrm{e}^-}}{P_A} \qquad K_{A^-}(T) = \frac{P_{A^-}}{P_A P_{\mathrm{e}^-}}\cdot \tag{6}$$

We now define the *fictitious pressure* $P(i)$ as the pressure exerted by element i if all the gas were in the neutral form of the atomic species i. It is customary to specify the partial pressures of all the atomic species in terms of the fictitious pressure of hydrogen $P(H)$. The coefficient relating one and the other is the abundance of each element with respect to hydrogen, $A(i)$, so $P(i) = A(i)P(H)$. The abundance of each atomic species depends on the metallicity of the star we are considering. For the solar case, we have used the standard solar abundances given by Grevesse (1984).

Under ICE conditions, the number of atoms and molecules are obtained by solving the conservation of mass and the chemical equilibrium conditions given by Eq. (2). The conservation of mass establishes that the sum of the partial pressures of all the species containing a given atomic element (taking into account the stoichiometry) equals the fictitious pressure of the given element:

$$P(i) = P_i + P_{i^+} + P_{i^-} + \sum_k \omega_k^i P_k, \tag{7}$$

where P_i, P_{i^+} and P_{i^-} are the partial pressure of the neutral element i, and the ionized elements i^+ and i^-, respectively, while P_k is the partial pressure of molecule k which has species i in its composition. ω_k^i is the stoichiometric coefficient which indicates the number of times element i appears in molecule k. The sum is extended over all the molecules in which the element i takes part. The previous equation can be rewritten with the aid of Eqs. (2) and (6):

$$P(i) = P_i + K_{i^+}\frac{P_i}{P_{\mathrm{e}^-}} + K_{i^-}P_i P_{\mathrm{e}^-} + \sum_k \omega_k^i \frac{P_i^{\omega_k^i} P_j^{\omega_k^j} \cdots P_l^{\omega_k^l}}{K_\mathrm{p}^{(k)}(T)}, \tag{8}$$

where $i, j, \ldots, l$ are the atomic species composing molecule k, $\omega_k^i + \omega_k^j + \ldots + \omega_k^l = n_k$ is the number of atoms of molecule k and $K_\mathrm{p}^{(k)}(T)$ is its equilibrium constant. Charged molecules can be easily included by taking into account their dissociation and the ionization equilibrium, simultaneously. Consider the dissociation equilibrium of an ionized diatomic molecule. One only needs to take into account that when a charged molecule dissociates, the atomic element which remains ionized is the one with the lowest ionization potential:

$$AB^+ \Leftrightarrow A^+ + B, \tag{9}$$

with $D_0(A) < D_0(B)$. Therefore, using Eqs. (2) and (6) we can calculate the partial pressure of the ionized molecule:

$$K_\mathrm{p}(T) = \frac{P_{A^+} P_B}{P_{AB^+}} = \frac{K_{A^+}(T)}{P_{\mathrm{e}^-}} \frac{P_A P_B}{P_{AB^+}}\cdot \tag{10}$$

We can write an equation like Eq. (8) for each of the N_s atomic species included in the calculation. In our case we have selected the 21 species which are shown in Table 1. Therefore, once the molecular and ionization equilibrium constants are known, we have a set of N_s non-linear algebraic equations which depend on the local temperature and density. The ICE approximation is intrinsically local since it depends only on the local values of the temperature and density. Therefore, to calculate the molecular number densities in a model atmosphere, we have to solve the algebraic non-linear system of equations at each point in the atmosphere. This system is solved using a standard Newton-Raphson iterative method (Press et al. 1986). Once the partial pressures for all the atomic species are known, we can calculate the ensuing atomic number densities by using Eq. (4). The molecular number densities are calculated by solving for their partial pressures in Eq. (2), using the equilibrium constant appropriate for each molecule.

Table 1. Number of hidden neurons used for each ANN.

Species	Number of hidden neurons
He, Na, Mg, Al, P, K, Ca	20
Ti, Cr, Mn, Fe, Ni, Cu	
H, C, N, O, F, Si, S, Cl	30

2.2. Equilibrium constants

As we discussed above, the problem of obtaining the atomic and molecular number densities is completely defined once the temperature and the total density (or pressure) is known. However, we also need to know the value of the equilibrium constant for each temperature. It is known from considerations of statistical mechanics that the value of the equilibrium constant can be obtained with the aid of the partition function of the molecule and of the individual atomic components, their respective masses and the dissociation energy D_0 (see, e.g., Tejero Ordoñez & Cernicharo 1991, for detailed information on how the equilibrium constants can be calculated). The expression is given by (Tejero Ordoñez & Cernicharo 1991, and references therein):

$$K_\mathrm{p}(T) = (kT)^{NA-1} \left(\frac{2\pi kT}{h^2}\right)^{3(NA-1)/2} \left(\frac{m_A^a m_B^b m_C^c}{m_{ABC}}\right)^{3/2} \times \left(\frac{\phi_A^a \phi_B^b \phi_C^c}{\phi_{ABC}}\right) \mathrm{e}^{-D_0/kT}, \qquad (11)$$

where m_i is the mass of atomic or molecular species i, $NA = a + b + c$ is the number of atoms present in the molecule, ϕ_i is the partition function of species i. h and k stand for the Planck and Boltzmann constants, respectively. The partition functions can be obtained by performing the summation

$$\phi = \sum_j g_j \mathrm{e}^{-E_j/kT}, \qquad (12)$$

where g_j is the degeneracy of level j and E_j its energy. Ideally, the sum has to be extended over all the energy levels of the species. However, due to the difficulty of accounting for all the energy levels, the most complete available set of energy levels must be used. For molecules, the partition function has to include the contribution from the electronic, vibrational and rotational levels. It is not practical to compute this summation for every ICE calculations. As a workaround, some authors have tabulated polynomial fits to the partition functions (Russell 1934; Tsuji 1971; Sauval & Tatum 1984; Tejero Ordoñez & Cernicharo 1991). In this work, we make use of the results obtained by Tejero Ordoñez & Cernicharo (1991). It represents one of the most complete compilations of molecular equilibrium constants for more than 240 diatomic and polyatomic molecules. These equilibrium constants were obtained by calculating the partition function with the most up-to-date molecular constants at the time of the work and their dependence with the temperature are given, as in previous works, as polynomial fits. One of the advantages of selecting this database is its self-consistency.

3. Neural network

3.1. Description of the network

The ICE problem consists on obtaining the partial pressures of the atomic species (or the atomic species number densities) which are consistent with Eq. (8) once the local temperature, local hydrogen density and local electron density are given. One could also take a different point of view and see the ICE problem as a mapping between the three-dimensional space given by $(T, n(\mathrm{H}), n(\mathrm{e}))$ onto several one-dimensional spaces, one for each atomic species included in the problem. The functional form of the functions $f\colon \boldsymbol{R}^3 \to \boldsymbol{R}$ is not known.

The Artificial Neural Network (ANN) with one hidden layer is a universal approximant to any non-linear continuous function (e.g., Jones 1990; Blum & Li 1991). The schematic structure of such a network is shown in Fig. 1. We have constructed an ANN with an input layer of three neurons where the temperature, hydrogen density and electron density are introduced. The neurons of the input layer have a linear activation function. The hidden layer consists on N_h neurons with a non-linear activation function $\sigma(x)$. Each hidden neuron is connected to all the neurons of the input layer by a certain weight (indicated by arrows in the figure). The value obtained at each hidden neuron is a linear combination of the values at the neurons of the input layer multiplied by the weights. These values are then applied the non-linear function $\sigma(x)$, multiplied by another set of weights and summed to give the final output of the neural network. The output of the network can be written as:

$$N(T, n(\mathrm{H}), n(\mathrm{e}), \boldsymbol{m}) = \sum_j^{N_\mathrm{h}} v_j \sigma\Big[w_j^\mathrm{T} T + w_j^\mathrm{H} n(\mathrm{H}) + w_j^\mathrm{e} n(\mathrm{e}) + u_j\Big], \qquad (13)$$

where the activation function is usually given by the sigmoid or the hyperbolic-tangent function. In this case, we have selected $\sigma(x) = \tanh(x)$. In the previous expression, $\boldsymbol{m}$ formally represents the whole set of $5 \times N_\mathrm{h}$ weights which define the neural network.

The use of a neural network approach for the solution of the ICE problem is very appealing for several reasons. Firstly, we obtain an analytic function which is infinitely differentiable and which could be easily used in any subsequent calculation. This way, we can investigate with great detail the dependence of the atomic and/or molecular abundances with temperature, hydrogen density and electron density. Secondly, the powerful approximation capabilities of ANNs make them very suitable for multi-dimensional interpolation problems. Even more striking is the ability of ANNs to extrapolate data outside the range of parameters used in the learning process. A direct consequence of these properties is that the number of points that need to be used for the learning process can be very small compared to the case of standard interpolation techniques. Furthermore, the number of parameters needed for performing the interpolation is also small. Essentially, one transforms the problem from storing in memory the whole set of N points which relates T, $n(H)$ and $n(\mathrm{e})$ with the partial pressures P_i for every species to storing only the $5 \times N_\mathrm{h}$ $\boldsymbol{m}$ weights present in Eq. (13) for each

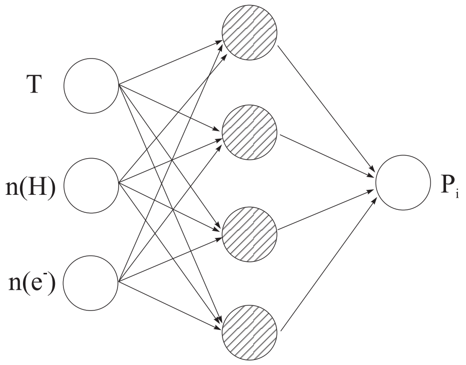

Fig. 1. Schematic structure of the neural network model used for approximating the ICE results. The input layer and the output layer are linear, while the hidden layer is assumed to be non-linear, with an hyperbolic-tangent activation function. The input layer consists of three neurons for the temperature, hydrogen number density and electronic number density. The output neuron gives the partial pressure of element i for the physical conditions in the input layer.

species. This reduction is proportional to the ratio N/N_{h}, which can be very large for large N, as we will see below. Finally, it is also important to note that the neural network can be straightforwardly implemented in parallel architectures.

3.2. Learning process

In order to obtain a neural network which is able to approximate the abundance of each species included in the chemical equilibrium calculation, we randomly select N_l sets of temperature, hydrogen density and electronic density. For each one of these combinations, we solve the ICE equations and calculate the abundances of all the atomic species. We have verified that taking $N_l = 1000$ gives a good coverage of the parameter space, considering that they are randomly selected. The temperature is varied in the range from 3500 K and 14 000 K, which is representative of the physical conditions in typical stellar atmospheres. Concerning the hydrogen and electronic abundances, we decided to use a logarithmic scaling. The hydrogen density is varied from 10^{12} cm^{-3} to 10^{18} cm^{-3} while the electronic density is varied from 10^{8} cm^{-3} to 10^{17} cm^{-3}. Again, these values are a reasonable representation of realistic conditions.

Once the learning set is selected, the training of the network for each species i reduces to the minimization of the following error function:

$$E_i = \sum_{l=1}^{N_l} \left[P_i^l - N(T_l, n_l(\mathrm{H}), n_l(\mathrm{e}), \boldsymbol{m}) \right]^2, \tag{14}$$

so that the optimum values of the parameters $\boldsymbol{m}$ are obtained. P_i^l is the partial pressure (or number density) of species i obtained for the combination of physical conditions l. Since almost every minimization algorithm makes use of the derivatives of the error function with respect to the parameters, it is of interest to notice that they can be obtained analytically:

$$\frac{\partial E}{\partial m_k} = -2 \sum_{l=1}^{N_l} \left[P_i^l - N(T_l, n_l(\mathrm{H}), n_l(\mathrm{e}), \boldsymbol{m}) \right] \frac{\partial N}{\partial m_k}, \tag{15}$$

where the partial derivatives of the neural network with respect to the parameters are obtained in a straightforward manner by using Eq. (13). In order to obtain the set of parameters $\boldsymbol{m}$ which minimize the error function, we have used a quasi-Newton BFGS method (Fletcher 1987) implemented in the Merlin package (Papageorgiou et al. 1998) which presents a very good convergence rate.

It is interesting to note that the usual approach of training the neural network with a learning set obtained from the original problem can be circumvented here. We do not need to rely on the precision given by the Newton-Raphson solution and the neural network can be made as precise as desired. This is a consequence of the fact that the equations describing the ICE approximations are known. One can see that the equations of ICE can be formally written as $f(T, n(\mathrm{H}), n(\mathrm{e}), \{P_i\}) = 0$, where $\{P_i\}$ is the set of N_{s} partial pressures (or atomic number densities) which satisfy the equations. Therefore, we can build as many neural networks as species included in the ICE problem and train the network by minimizing the following error function:

$$E = \sum_{l=1}^{N_l} \left[f(T_l, n_l(\mathrm{H}), n_l(\mathrm{e}), \{N_i(T_l, n_l(\mathrm{H}), n_l(\mathrm{e}), \boldsymbol{m})\}) \right]^2. \tag{16}$$

The solution to the problem is transformed into finding the set of parameters $\boldsymbol{m}$ for all the networks which minimizes the previous error function. The summation is extended over the N_l elements of the learning set. Several issues are noteworthy in this case. On the one hand, the derivatives of the error function with respect to the parameters $\boldsymbol{m}$ are not as straightforward as those found in Eq. (16) since the function $f(T, n(\mathrm{H}), n(\mathrm{e}), \{P_i\})$ is non-linear. On the other hand, the minimization process has to be carried out simultaneously for the N_{s} neural networks, so that the problem of finding the minimum might be more difficult to solve. The dimension of the hypersurface in which we have to find the minimum is $5 \times N_{\mathrm{h}} \times N_{\mathrm{s}}$, instead of finding N_{s} minima in hypersurfaces of dimension $5 \times N_{\mathrm{h}}$.

A critical parameter in the training ability of a neural network is the number of hidden neurons N_{h}. It is important to build networks with sufficient number of hidden networks to accurately approximate the non-linear function we are interested in. However, since the number of learning points is not infinite, the number of hidden neurons one can use is practically limited by the overtraining phenomenon (in principle, the number of training points should be much larger than the number of hidden neurons). If N_{h} is very large, the network is capable of correctly approximating all the points in the learning set but its interpolation capacity is lost, producing large oscillations between the points of the learning set. This is a direct consequence of the increase in the degrees of freedom of the network when the number of hidden neurons is increased. One way to avoid this overtraining is to use two different data sets: one for learning purposes and the other for testing. When the training process takes places, the error given by expression (14) is evaluated for the training set and for the test set. When the network is training correctly, both errors are reduced. When the network starts to be overtrained, the error of the test set starts to increase. We have followed this scheme in order to stop the

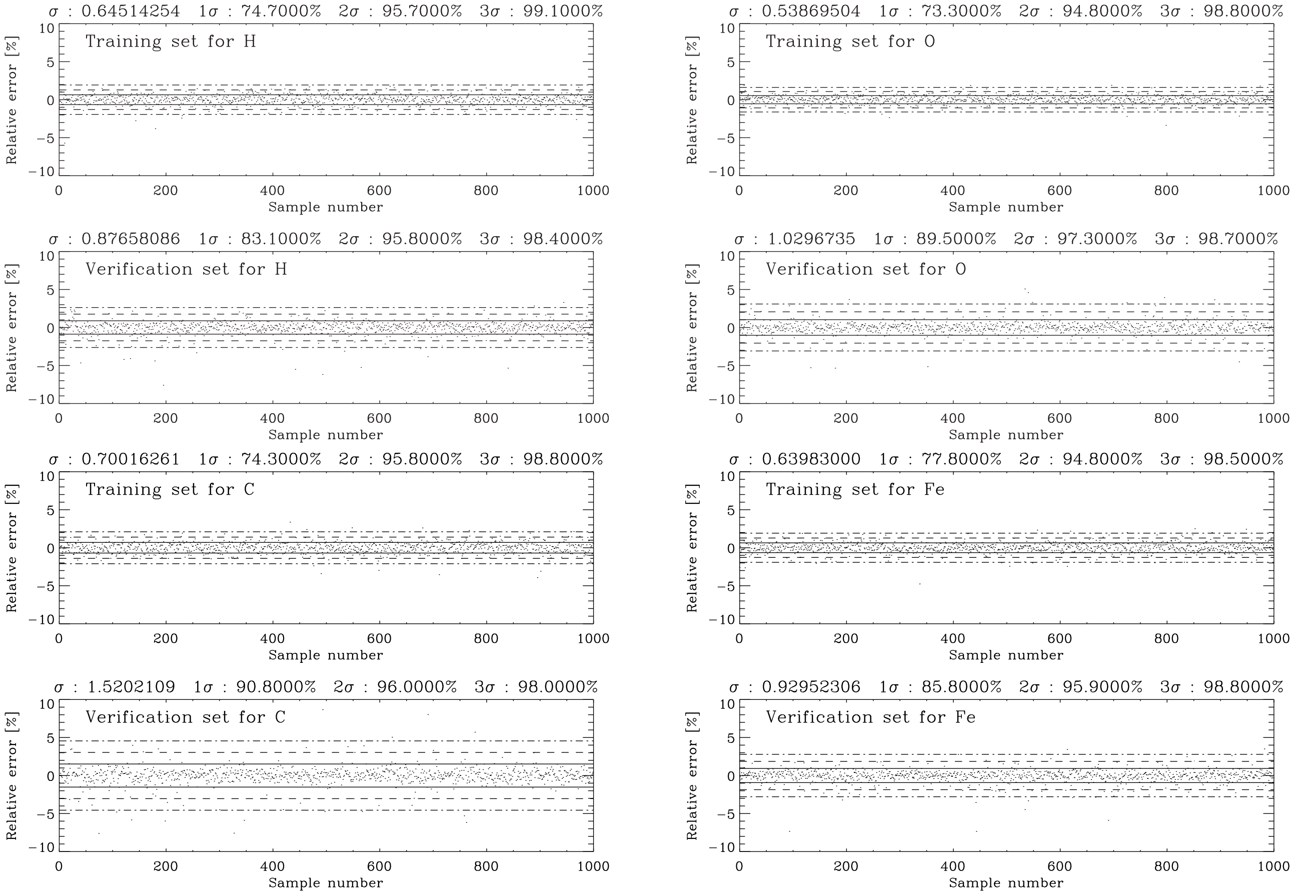

Fig. 2. Relative error obtained after the training of the neural network for hydrogen, oxygen, carbon and iron. The distribution of relative errors are shown for the learning set and for a verification set, which aids at detecting overtraining. We indicate the value of the standard deviation σ of the distribution and the number of points below 1σ, 2σ and 3σ. We also indicate these values in the plots as horizontal lines.

training process before overtraining. In our case, we have used two different values for the number of hidden neurons. We have verified that $N_\mathrm{h} = 20$ gives sufficiently accurate results for 13 of the 21 atomic species included in the ICE calculations, while we were forced to increase this number to $N_\mathrm{h} = 30$ for the rest of species. Table 1 lists the number of hidden neurons for each species. We did not investigate the dependence of the minimum error on the number of hidden neurons so that the increase to $N_\mathrm{h} = 30$ was somewhat arbitrary.

The training of the neural networks was stopped when the standard deviation of the relative error between the output of the neural network and the values obtained from the solution of the ICE equations for both the learning and the test sets was 1% or below. Obviously, the relative error was always larger for the test set than for the learning set, but the training always reached the limit of 1% and was stopped before any indication of overtraining was observed. In Fig. 2 we present the relative error obtained for the 1000 training points for the abundances of H, C, O and Fe. The behavior of the relative error for the rest of species follows the same behavior, except for He, which presents a much smaller dispersion. Since He cannot form molecules, it is always in its atomic form and the non-linear mapping that the neural network has to learn is strongly simplified.

We have indicated in each plot the value of the standard deviation and the percentage of points which are contained within 1σ, 2σ and 3σ. This is an indication of the precision obtained with the neural network. It is common in all the networks that more than ~95% of the points are inside 3σ, which translates into a relative error of ~3% (since σ is always of the order or below 1%). However, we note that it is common to all the networks that more than 75% of the points are below 1σ. Along with the plots of the results obtained for the learning set, we have included the results for the verification set. It is interesting to note that, although the standard deviation of the distribution of relative errors is larger than that obtained for the learning set, typically more than ~80% of the points are below this limit. In the case of the network for the carbon abundance, we have a relatively high value for σ, but more than 90% of the points are below this limit.

4. Comparison of results

Once the neural networks were trained to the desired precision, as shown in Fig. 2, we have applied them to realistic situations in order to compare their performance with the complete ICE calculations. To this end, we make use of a three-dimensional snapshot of the convection simulation of the solar atmosphere

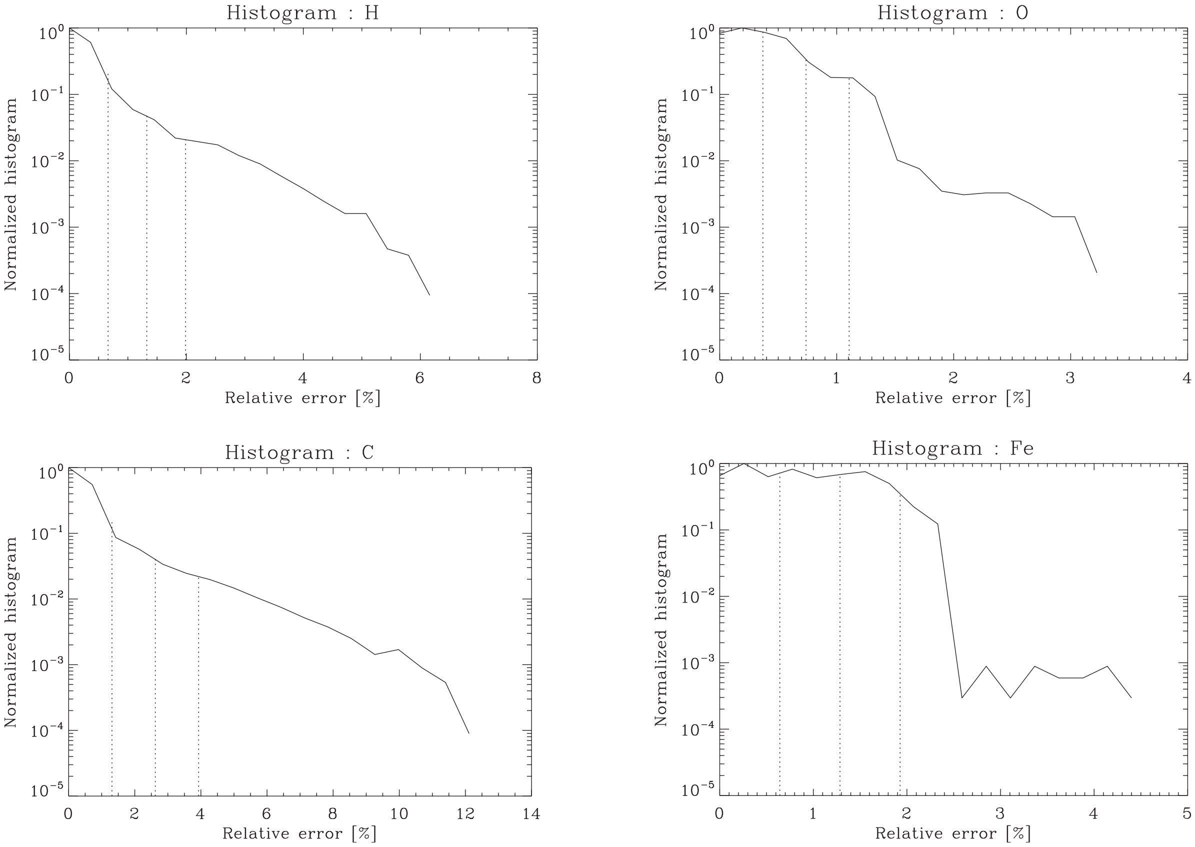

Fig. 3. Histograms of relative errors for the $2.55{\times}10^5$ points of the three-dimensional convection simulation of Asplund et al. (2000). We indicate the position of 1σ, 2σ and 3σ of the distribution. Note that they are similar to those obtained for the verification set during the training of the networks.

of Asplund et al. (2000). We consider this simulation to be an adequate representation of the physical conditions in the solar atmosphere since several investigations are reporting good agreements between the synthetic line profiles and the observed ones (e.g., Shchukina & Trujillo Bueno 2001; Asplund et al. 2004, 2003). The snapshot box size is $50 \times 50 \times 102$, so that the number of points is 2.55×10^5. We have solved the ICE problem in each point and confronted the results with those obtained from the neural networks.

The normalized error histograms are shown in Fig. 3 for H, C, O and Fe, with vertical dashed lines indicating the position of 1σ, 2σ and 3σ of the distribution. Note that the histograms are decreasing functions of the relative error, so that the number of points with large relative error is much smaller than the number of points with small relative errors. It is verified that the value of the standard deviation of the distribution of relative errors is quite similar to those obtained for the verification set. This reinforces our confidence that the learning set and the verification set provide sufficient coverage of the three-dimensional parameter space $(T, n(\mathrm{H}), n(\mathrm{e}))$. The shapes of the histograms turn out to be fairly linear in this log-linear scale, so that we can consider, as a first approximation, that the number of points N_p with a certain relative error r is given by $N_\mathrm{p} \propto r^{-\alpha}$. For example, we find that $\alpha \simeq 5$ for hydrogen, $\alpha \simeq 3.7$ for carbon and $\alpha \simeq 7$ for oxygen. There is an exception for iron, for which the distribution of points with relative errors below ~2.3% seems to be constant, while the number of points with errors larger than ~2.3% is three order of magnitude smaller. The behavior of the histograms for the rest of species is very similar to those shown in Fig. 3 for the selected ones, maintaining the properties obtained for the verification sets.

5. *T*, *n*(H) and *n*(e) dependence

Since the neural network approach leads to analytic formulae for the solution of ICE problems, we now take advantage of them and compare the results for abundant species with simplified models. Our aim is to investigate the variation with temperature, hydrogen density and electronic density of the number density of carbon and oxygen. We have selected these two species because they are the constituents of the highly abundant CO diatomic molecule. This way, we can safely obtain an approximation to the carbon and oxygen abundances by taking only into account the formation of CO in the ICE equations and neglecting the rest of molecular species. Note that this simplification is only valid for species like C and O and not for species which form less abundant molecules and/or form several high abundance molecules. Specializing Eq. (8) to oxygen and carbon, we obtain:

$$\begin{aligned} P(\mathrm{C}) &= P_\mathrm{C} + \frac{P_\mathrm{C} P_\mathrm{O}}{K_\mathrm{CO}(T)} + K_{\mathrm{C}^+} \frac{P_\mathrm{C}}{P_{\mathrm{e}^-}} + K_{\mathrm{C}^-} P_\mathrm{C} P_{\mathrm{e}^-} \\ P(\mathrm{O}) &= P_\mathrm{O} + \frac{P_\mathrm{C} P_\mathrm{O}}{K_\mathrm{CO}(T)} + K_{\mathrm{O}^+} \frac{P_\mathrm{O}}{P_{\mathrm{e}^-}} + K_{\mathrm{O}^-} P_\mathrm{O} P_{\mathrm{e}^-}. \end{aligned} \qquad (17)$$

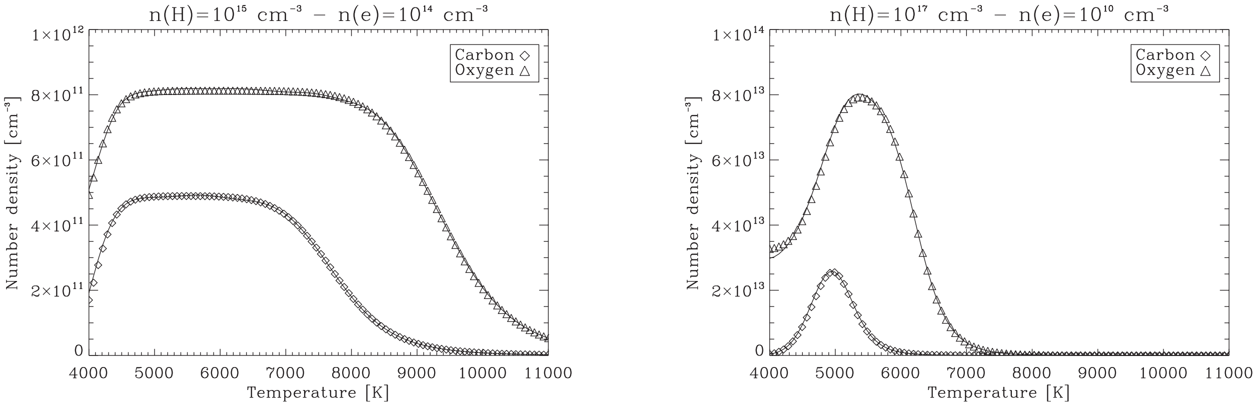

Fig. 4. Number density of carbon and oxygen obtained with the simplified model and with the neural network for two different combinations of hydrogen and electron densities and for different values of the temperature. Note that this kind of investigation is greatly simplified due to the analytical character of the neural network.

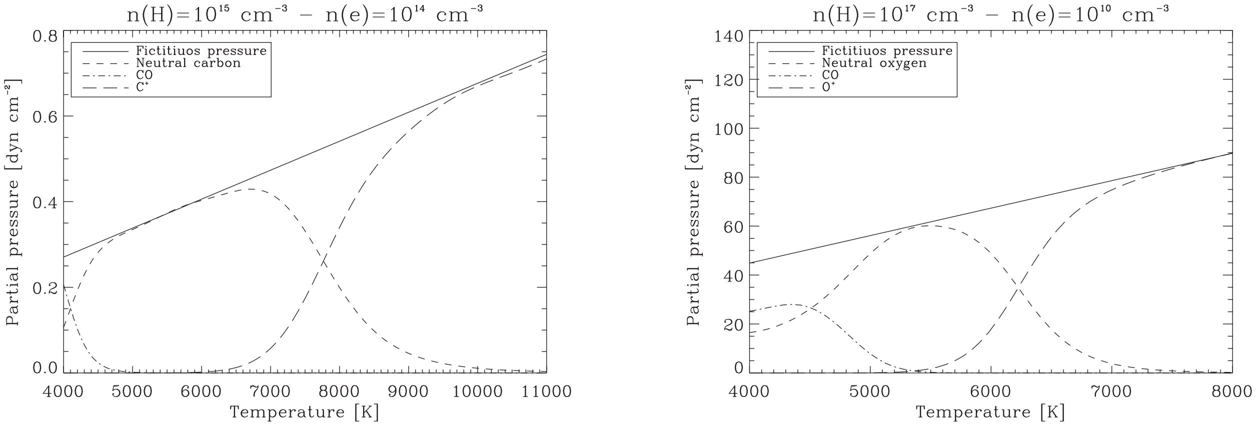

Fig. 5. Contribution to the total pressure of carbon and oxygen in two different cases. The curves have been obtained using the neural networks. The addition of all the contribution closely matches the fictitious pressure of carbon and oxygen. This is the reason why the very simplified model for C and O works very well.

This system of equations can be solved analytically for P_{C} and P_{O} and transformed into number densities by using the ideal gas equation Eq. (4). Their dependence with temperature for two different combinations of hydrogen and electronic densities are shown in Fig. 4. The output of the neural network is also shown in the plots with solid lines. Note the accurate results obtained with this simplified model in this high abundance case. This behavior can be better understood if we plot separately the contribution to the total pressure of each element of each term in Eq. (17). The results are shown in Fig. 5 for the two sets of physical conditions. The fictitious pressure of carbon and oxygen (essentially the abundance of carbon and/or oxygen times the hydrogen density) is shown by solid line, the partial pressure of the neutral element by dashed line, the partial pressure of the carbon monoxide molecule by dash-dotted line and the partial pressure of the ionized element by long dashes. All the quantities shown in the plots have been obtained using the neural network approach. Although the neural network takes into account the formation of other molecules which contain carbon and oxygen (i.e., it gives the solution to the complete ICE problem), we see that accounting only for CO appears to be enough since the addition of all the contribution closely approximates the value of the fictitious pressure.

Note that the CO molecule is efficiently formed when the temperature is below ~5000 K in both situations. The neutral species are mainly situated at intermediate temperatures, typical from photospheric regions, while the ionized species tend to dominate for high temperatures, above 6000−7000 K.

6. Conclusions

We have successfully trained 21 neural networks which approximate the solution to the Instantaneous Chemical Equilibrium problem. We have generated a learning set which is representative of the physical conditions in stellar atmospheres and have verified that the coverage of the three-dimensional parameter space is sufficiently good with only 1000 points. We have trained the neural networks so that the standard deviation of the relative error were below or of the order of 1%. In order to avoid overtraining, we have employed an independent verification set. The standard deviation of the relative error in this verification set is larger than that for the learning set but also of the order of 1%. We applied the neural networks obtained to the ICE problem in a three dimensional convection simulation, representative of the physical conditions in the solar atmosphere. The histograms of relative errors built with all the points in the simulation show that the neural networks are capable of solving the ICE problem with relative errors similar to those obtained for the verification set. The advantage of the neural network approach is their intrinsic analytical character, the possibility of parallelization and the very fast evaluation. The development of a fast approach to ICE is of importance for, among others, recent multi-dimensional simulations of stellar atmospheres which include very dense grids and iterative inversions of observed spectra.

Acknowledgements. This research has been partly funded by the Ministerio de Educación y Ciencia through project AYA2004-05792 and by the European Solar Magnetism Network (contract HPRN-CT-2002-00313).

References

Asensio Ramos, A., Trujillo Bueno, J., Carlsson, M., & Cernicharo, J. 2003, ApJ, 588, L61

Asplund, M., Ludwig, H. G., Nordlund, A., & Stein, R. F. 2000, A&A, 359, 669

Asplund, M., Carlsson, M., & Botnen, A. V. 2003, A&A, 399, 31

Asplund, M., Grevesse, N., Sauval, A. J., Allende Prieto, C., & Kiselman, D. 2004, A&A, 417, 751

Bishop, C. M. 1996, Neural networks for pattern recognition (Oxford University Press)

Blum, E. K., & Li, L. K. 1991, Neural Networks, 4, 511

Carroll, T. A., & Staude, J. 2001, A&A, 378, 316

Fletcher, R. 1987, Practical Methods of Optimization, 2nd ed. (New York: Wiley)

Grevesse, N. 1984, Phys. Scr, 8, 49

Jones, L. K. 1990, in IEEE Proc., 78, 1585

McCabe, E. M., Connon Smith, R., & Clegg, R. E. S. 1979, Nature, 281, 263

Papageorgiou, D. G., Demetropoulos, I. N., & Lagaris, I. R. 1998, Comput. Phys. Commun., 109, 227

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1986, Numerical Recipes (Cambridge: Cambridge University Press)

Russell, H. N. 1934, ApJ, 79, 281

Sauval, A. J., & Tatum, J. B. 1984, ApJS, 56, 193

Shchukina, N., & Trujillo Bueno, J. 2001, ApJ, 550, 970

Socas-Navarro, H. 2003, Neural Networks, 16, 355

Socas-Navarro, H. 2005, ApJ, 620, 517

Tejero Ordoñez, J., & Cernicharo, J. 1991, Modelos de Equilibrio Termodinámico Aplicados a Envolturas Circunestelares de Estrellas Evolucionadas (Madrid: IGN)

Tsuji, T. 1964, Ann. Tokio. Astron. Obs., 2nd ser., 9, 1

Tsuji, T. 1971, Ann. Tokio Astron. Obs. 2nd ser., 9, 1

Tsuji, T. 1973, A&A, 23, 411