Available online at www.sciencedirect.com
SCIENCE DIRECT®
Chemical Physics 297 (2004) 153–161
Chemical Physics
www.elsevier.com/locate/chemphys

The fitting of potential energy and transition moment functions using neural networks: transition probabilities in OH ($A^2\Sigma^+ \rightarrow X^2\Pi$)

Ana Carla P. Bittencourt [a], Frederico V. Prudente [a,*], José David M. Vianna [a,b]

[a] *Instituto de Física, Universidade Federal da Bahia, Campus de Ondina, 40210-340 Salvador, Bahia, Brazil*
[b] *Instituto de Física, Universidade de Brasília, 70919-970 Brasília, DF, Brazil*

Received 19 September 2003; accepted 15 October 2003

Abstract

We have studied the performance of the back-propagation neural network with different architectures and activation functions to fit potential energy curves and dipolar transition moment functions of the OH molecule from the ab initio data points of Bauschlicher and Langhoff [J. Chem. Phys. 87 (1987) 4665]. The neural network fittings are tested in different moments of the training process by computing the vibrational levels, the transition probabilities between $A^2\Sigma^+$ and $X^2\Pi$ electronic states, and the radiative lifetimes. The results from the neural network fittings are then compared with experimental values, previous results calculated by Bauschlicher and Langhoff and the ones obtained by using of extended Rydberg function fitting.

PACS: 31.50.−x; 33.70.Ca; 33.50.Hv

Keywords: Neural networks; Back-propagation; Discrete variable representation; Potential energy surfaces; Transition probabilities; OH molecule

1. Introduction

The study of nuclear motions in molecules is relevant in such problems as the determination of rovibrational states, photodissociation process and molecular reactive scattering, where the Born–Oppenheimer approximation is usually considered. This approximation implies that the molecular dynamics is governed by a potential energy surface (PES) obtained from the electronic energy and the nuclear repulsion term. There are two important steps to obtain the PES: (i) the calculation of the electronic energy for a set of nuclear configuration using ab initio and/or semiempirical methods and (ii) the fitting of the potential energy surface from these points. The fitting process is usually done by power series in an appropriate coordinate system [1–3], using local functions such as the cubic spline [4–6], or semiempirical potentials with adjustable parameters to reproduce experimental and theoretical results [7,8]. However, alternative procedures to those methods have always been sought [9–15]. One of them is the multilayer neural network (NN) method to fit the PES [16–22].

Artificial neural network is a highly nonlinear computational device based on the organizational structure of the brain and in its learning process [23,24]. It consists in an interconnected processing units called "neurons" typically arranged in layers. Each neuron receives a set of input signals and transforms them, emitting a output signal, as a synapse, by using a function (activation or transfer function). The output signal is directed to other neurons through links affected by a connection weights. These weights are determined by an adaptive method known as learning or training process, where the neural network self-organizes to reproduce some desired behaviour. The learning process is, to modelling problems, fulfilled by control of the error performed by NN in the reproduction of known examples. The most simple and widely used algorithm to train the multilayer neural network is the back-propagation (BP) procedure [24,25]. The BP algorithm is an interactive gradient technique that minimizes the global error between the exact and the neural network output. However, there are many

* Corresponding author. Tel.: +55-71-2472033; fax: +55-71-2355592.
E-mail addresses: acpb@ufba.br (A.C.P. Bittencourt), prudente@ufba.br (F.V. Prudente), david@ufba.br (J.D.M. Vianna).

0301-0104/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemphys.2003.10.015

other more sophisticated minimization algorithms such as second order optimization methods [25–29]. Because its versatility, NNs have been used in a variety of applications in chemistry and physics [24,30–32]. Particularly in both theoretical chemistry and atomic and molecular physics NNs have been recently applied with great success in many subject areas [33–41].

In present work, the central aspect is the fitting of the fundamental ($X^2\Pi$) and excited ($A^2\Sigma^+$) electronic potential energy curves (PECs) and the corresponding dipolar transition moment function (DTMF) to the OH molecule from ab initio points by using multilayer neural networks. The choice of the OH system is motivated by the existence of several theoretical studies based on ab initio [42–46] and semiempirical calculations [47–49]. Our major goal is to discuss the implications of using different NN architectures (types of activation functions, number of neurons and number of layers) in the quality and accuracy of the fitted PEC and DTMF. For this, we employ them to determine vibrational frequencies, probabilities in the $A^2\Sigma^+ \rightarrow X^2\Pi$ transitions and the radiative lifetime of the $A^2\Sigma^+$ state of the title system. Note that the transition probabilities are very sensitive to the PECs and DTMF involved in the process. The study of an one-dimensional system enables us to realize an exhaustive and comparative analysis between different possible choices for NN architectures concerning the quality of the PEC and DTFM; this because it is easier to compute such properties as vibrational levels and transition probabilities. Similar studies for multidimensional problems should be prohibitive because the computational cost to calculate, for example, the vibrational energy levels should be very high. Such an analysis is useful in the attempt to establish the NNs as a technique to fit molecular potential energy surfaces.

In particular, we employ the BP procedure to train neural networks with different types of activation functions (sigmoidal, hyperbolic tangent and linear functions), number of neurons (3, 4, and 6), and number of layers (one or two hidden layers). For these learning processes, we consider the ab initio data points of $X^2\Pi$ and $A^2\Sigma^+$ electronic states and the corresponding dipolar transition moment function obtained by Bauschlicher and Langhoff [45]. It is interesting to point out that the stop condition for training process of the neural network is not based solely on the error function but also on the accuracy of the calculated vibrational levels obtained from the fitted NN potential energy curves. The vibrational levels are calculated using the discrete variable representation (DVR) method (see [50–52] and references therein). Transition probabilities in $A^2\Sigma^+ \rightarrow X^2\Pi$ and the radiative lifetimes are then obtained for the best neural network fittings. In order to test the accuracy of our results, we compare them with the experimental values [53–55], with the results calculated by Bauschlicher and Langhoff in [45], and with the values obtained using the extended Rydberg function [3] fitting applied to the same ab initio data.

The structure of the paper is the following. Section 2 presents the basic theory of neural network, while the theoretical framework and the main aspects of the DVR method are shown in Section 3. Section 4 presents the different neural networks fittings of the potential energy curves and dipolar transition moment function, and the results to vibrational levels, transition probabilities and lifetimes calculations. The paper is closed with some concluding remarks in Section 5.

2. Neural network

A neural network performs an input–output mapping based on the associations of basic units called neurons. A single neuron receives an input, a vector of elements $\{x_i\}$, and the output y_j is computed by

$$y_j = f\left(\sum_{i=1}^{n} w_{ji}x_i + w_{j0}\right), \tag{1}$$

where f is the activation function and w_{ji} are the synaptic weights associated to the synaptic connections (w_{j0} is a bias). The activation function is responsible for the non-linear behavior of the neural networks [56]. There are many types of them. But, as we already said in Section 1, here the following ones are considered:

$$f(z) = \frac{1}{1+\exp(-z)}, \tag{2}$$

$$f(z) = \tanh(z), \tag{3}$$

$$f(z) = z, \tag{4}$$

that are named as sigmoidal (s), hyperbolic-tangent (t) and linear (l) activation functions, respectively.

A neural network structured as a set of layers of neurons, where the response signal of each neuron in a layer is an input signal for the neurons in the next layer, is known as multi-layer perception. The multi-layer perception has a simple form and is widely used to treat interpolation problems [57]. In that case, the output of the ith neuron on the kth layer is

$$y_i^{(k)} = f^{(k)}(z_i^{(k)}), \tag{5}$$

with $z_i^{(k)}$ given by

$$z_i^{(k)} = \sum_{j=1}^{n_k} w_{ij}^{(k)} y_j^{(k-1)} + w_{i0}^{(k)}, \tag{6}$$

where $w_{ij}^{(k)}$ is the synaptic weight of the connection from the jth neuron of the $(k-1)$th layer to the ith neuron of the kth layer, $w_{i0}^{(k)}$ is the bias, and $y_j^{(k-1)}$ is the output of the jth neuron of the $(k-1)$th layer. An appropriated architecture for the present applications is a fully con-

nected NN with one input neuron without activation function (the nuclei distance R), some hidden layers, and an output layer with one neuron (the potential energy $[V(R)]$ or the dipolar transition moment function $[D(R)]$ values for the particular geometry). Usually, such an architecture is described by listing the number of neurons in each layer, followed by letters denoting the activation function used in the layer. For instance, Fig. 1 displays a 1-3s-1l structure which represents an NN with three layers, where the hidden layer has three neurons with the sigmoidal activation function and in the output layer is used the linear one. This NN corresponds to the following functional form:

$$y_1^{(2)} = f^{(2)}\left(\sum_{j=1}^{3} w_{1j}^{(2)} f^{(1)}(w_{j1}^{(1)} x_1 + w_{j0}^{(1)}) + w_{10}^{(2)}\right) \tag{7}$$

with $x_1 = R$ and $y_1 = V(R)$ or $D(R)$.

The next step is to obtain the best set of synaptic weights that fits a known set of input–output values $\{x_i, E_i\}$, $i = 1, \ldots, m$; in the present case we have, respectively, the nuclear configurations and the ab initio results of potential energy or the dipolar transition moment. By the training process of our NN we minimize the following error function:

$$\epsilon^2 = \frac{1}{2}\sum_{i=1}^{m}(E_i - y_i^{(N)})^2, \tag{8}$$

where $y_i^{(N)}$ is the output of the NN for the ith input (x_i). The BP algorithm that we have utilized here is composed by following steps: (i) the initial weights are obtained randomly and (ii) the weights are corrected according to

$$\Delta w_{ji}^{(l)} = w_{ji}^{(l)\text{new}} - w_{ji}^{(l)\text{old}} = -\eta \frac{\partial \epsilon}{\partial w_{ji}^{(l)}}, \tag{9}$$

where $\Delta w_{ji}^{(l)}$ is the correction for the weight $w_{ji}^{(l)}$ and η is a positive scale factor. Using the chain rule, we obtain

$$\Delta w_{ji}^{(l)} = \eta \delta_j^{(l)} y_i^{(l-1)}, \tag{10}$$

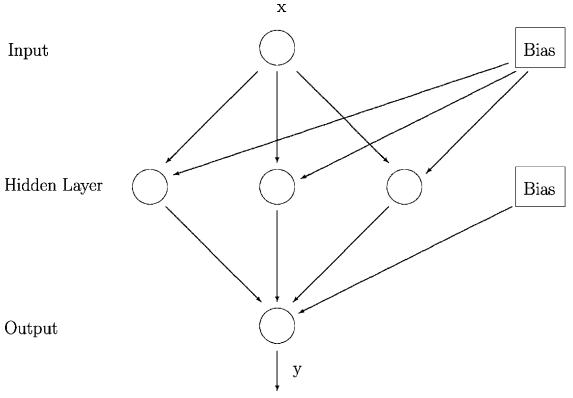

Fig. 1. The neural network with 1-3s-1l architecture.

where $y_i^{(l-1)}$ is the output of the ith neuron of the $(l-1)$th layer and the generalized delta function $\delta_j^{(l)}$ has the following definition:

$$\delta_i^{(l)} = \left(\sum_{k=1}^{r} \delta_k^{(l+1)} w_{ki}^{(l+1)}\right) \frac{\partial f^{(l)}(z)}{\partial z} \tag{11}$$

for $l \neq N$, and

$$\delta_i^{(N)} = (E_i - y_i) \frac{\partial f^{(N)}(z)}{\partial z} \tag{12}$$

for the output layer. In order to avoid that the numerical process is stopped at local minima during the minimization progress, we use the following generalization of Eq. (10):

$$\Delta w_{ji}^{(l)} = \eta \delta_j^{(l)} y_i^{(l-1)} + \mu \Delta w_{ji}^{(l)\text{previous}}, \tag{13}$$

where μ is the momentum constant. Both η and μ are taken between zero and one. The weights are corrected using Eq. (13) until the fitting (or the root-mean-square error) attains the desired accuracy. The term "desired accuracy" is defined in each context of use.

3. Transition probabilities

The rovibrational energy levels associated with a particular electronic state of OH system are characterized by the potential energy curve that describes such an electronic state. These rovibrational levels are then obtained by solving the radial equation for the nuclear motion [43]

$$-\frac{\hbar^2}{2\mu}\frac{\mathrm{d}^2}{\mathrm{d}R^2} X_{v,l}(R) + \left[V(R) + \frac{l(l+1)}{2\mu R^2} - E_{v,l}\right] X_{v,l}(R) = 0, \tag{14}$$

where $V(R)$ is the PEC, μ is the reduced mass, l and v are the rotational and the vibrational quantum numbers, and $X_{v,l}(R)$ is the rovibrational wave function with eigenenergy $E_{v,l}$. We point out that we have considered in present work the case $l = 0$ only, and we do not maintain this index in the notation. We emphasize that to solve Eq. (14) it is generally necessary to have an analytical function for the PEC and that the accuracy of the vibrational eigenfunctions and eigenvalues depends directly of the PEC quality.

Moreover, the knowledge of the DTMF, associated with accurated vibrational wave functions, allows the evaluation of the radiative transition probabilities $(A_{v'v''})$ for arbitrary transition between the vibrational levels of two electronic states and the respective lifetime of the excited level $(\tau_{v'})$. The transition dipole matrix elements and the Franck–Condon factor are obtained from

$$D_{v'v''} = \langle X_{v'}(R)|D(R)|X_{v''}(R)\rangle, \tag{15}$$

$$q_{v'v''} = \langle X_{v'}(R)|X_{v''}(R)\rangle, \tag{16}$$

where $D(R)$ is the electronic DTMF; $X_{v'}(R)$ and $X_{v''}(R)$ are the vibrational wave functions of the upper and lower states. In terms of $D_{v'v''}$, the spontaneous transition probability (or Einstein A-coefficient) from a bound vibrational level v' of an upper electronic state into a vibrational level v'' of a lower electronic state is given by [43]

$$A_{v'v''} = 2.03 \times 10^{-6} g \nu_{v'v''}^3 |D_{v'v''}|^2 s^{-1}, \tag{17}$$

where g is a statistical weighting factor equal to 2 for $\Sigma \rightarrow \Pi$ transitions; $D_{v'v''}$ is given in atomic unit and $\nu_{v'v''}$ is the transition frequency measured in cm^{-1}. The radiative lifetime of a vibrational energy level of the excited electronic state is the reciprocal of the sum of all Einstein coefficients for that level

$$\tau_{v'} = \left(\sum_{v''} A_{v'v''} \right)^{-1}. \tag{18}$$

In this work we use the discrete variable representation method [58–61] to calculate the vibrational eigenenergies and eigenfunctions. The DVR method is described with enough detail in many other papers (e.g., see [50–52] and references therein), but here we give a brief introduction.

The DVR procedure consists: (i) to build a set of basis functions $\Phi_\alpha(R)$ with the following property:

$$\Phi_\alpha(R_\beta) = \frac{\delta_{\alpha\beta}}{\sqrt{\omega_\alpha}}, \quad \alpha, \beta = 1, \ldots, k, \tag{19}$$

where $\{R_\alpha\}$ and $\{\omega_\alpha\}$, $\alpha = 1, \ldots, k$ are the pivots and weights of a Gaussian quadrature; (ii) to expand the trial wave function with the basis (19); and (iii) to solve the associated eigenvalue–eigenvector problem. In this method the matrix elements of the potential energy using the basis (19) are given by

$$V_{\alpha\beta} \simeq V(R_\alpha)\delta_{\alpha\beta}, \tag{20}$$

while the kinetic energy matrix elements can be calculated analytically. A characteristic of the DVR method is that the value of an eigenfunction in a quadrature point is simply the coefficient of the DVR function associated with this point divided by the root of the related weight ($X_v(R_\alpha) = c_\alpha^v/\sqrt{\omega_\alpha}$). Thus, integrals (15) and (16) are calculated by

$$\begin{aligned} D_{v'v''} &\simeq \sum_\alpha \omega_\alpha X_{v'}(R_\alpha) D(R_\alpha) X_{v''}(R_\alpha) \\ &= \sum_\alpha c_\alpha^{v'} D(R_\alpha) c_\alpha^{v''}, \end{aligned} \tag{21}$$

$$q_{v'v''} \simeq \sum_\alpha \omega_\alpha X_{v'}(R_\alpha) X_{v''}(R_\alpha) = \sum_\alpha c_\alpha^{v'} c_\alpha^{v''}. \tag{22}$$

In particular, here we have utilized the DVR method generated for an equally spaced set of points and weights [60], and Eq. (14) has been solved for different potential energy curves fitted by using neural network method and extended Rydberg function.

4. Results

The neural network method described in Section 2 is used to fit the potential energy curves of the $X^2\Pi$ and $A^2\Sigma^+$ electronic states and the associated dipolar transition moment function for the OH molecule. The ab initio points we have utilized as the input–output values of the neural network training process were the ones calculated by Bauschlicher and Langhoff [45]. They computed the potential energies and the dipolar transition moment for 19 nuclear distances between 1.3 and 4.4 bohr by using a multireference CI (MRCI), and employed them to calculate the differences between vibrational energy levels ($\Delta G_{v=1/2} = E_{v+1} - E_v$), the transition probabilities ($A_{v'v''}$) and radiative lifetimes ($\tau_{v'}$). We compare our results obtained from different architectures of neural networks, where the activation functions were varied, the number of neurons and the number of layers, with the Bauschlicher and Langhoff (BL) ones [45], the experimental values of Coxon et al. [53–55], and the ones calculated by using the extended Ryderberg function (ERF) [3] fitting. The ERF is a good representation of a typical diatomic potential given by

$$V_{\mathrm{ER}} = D_e \left[1 + \sum_k a_k (R - R_e)^k \right] \exp(-\gamma(R - R_e)), \tag{23}$$

which has a minimum of depth D_e at $R = R_e$. We point out that we have utilized $k \leqslant 8$ for the ERF fitting.

Initially, we fit $X^2\Pi$ and $A^2\Sigma^+$ PECs and the $A^2\Sigma^+ \rightarrow X^2\Pi$ DTMF for the OH system employing various fully connected NNs with one input neuron without activation function, one output neuron with the linear function and one or two hidden layers with the sigmoidal (s) and hyperbolic tangent (t) activation functions. To NNs with one hidden layer we varied the number of neurons of such a layer in 3, 4 and 6, while to NNs with two hidden layers we used two neurons at each one, totalizing eight different NNs to fit the PECs and the DTMF. The training process was stopped when the root-mean-square error of each neural network fitting reached the plateau of 10^{-4}–10^{-5} a.u. In order to eliminate the overfitting problem, we also supervised this process analysing the accuracy of the calculated vibrational levels when compared to experimental and BL values. Examples of these fittings for the $X^2\Pi$ and $A^2\Sigma^+$ electronic states, and the $A^2\Sigma^+ \rightarrow X^2\Pi$ DTM function are shown in Figs. 2 and 3, respectively.

To determine the quality of the NN fittings to potential energy curves we have calculated the vibrational energy levels using the equally spaced DVR method [60]. Specifically, the vibrational energy separations $\Delta G_{v-1/2}$

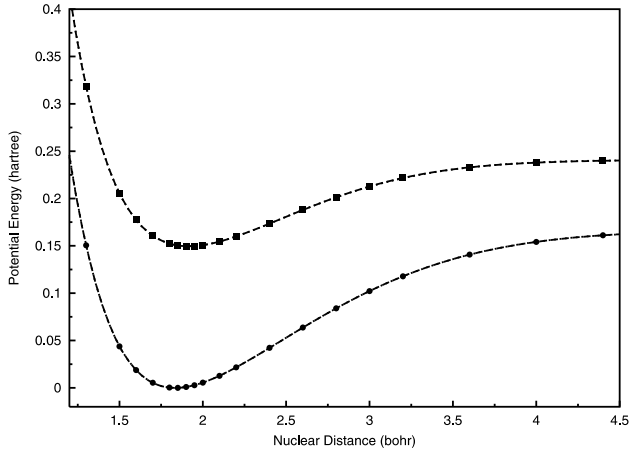

Fig. 2. The electronic ground state and $A^2\Sigma^+$ electronic excited state of the OH using 1-3*t*-1*l* (dashed line) and 1-3*s*-1*l* (dotted line) neural network fittings (they are indistinguishable here because both fittings are essentially identical). The ab initio data points [45] are denoted by circles and squares.

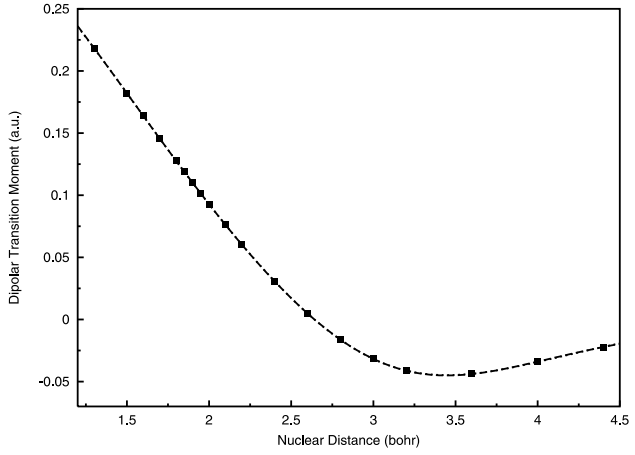

Fig. 3. The dipolar transition moment for $A^2\Sigma^+ - X^2\Pi$ of OH using neural network fitting: 1-3*t*-1*l* (dashed line) and 1-3*s*-1*l* (dotted line); they are indistinguishable here because both fittings are essentially identical. The ab initio data points [45] are denoted by squares.

for the $X^2\Pi$ and $A^2\Sigma^+$ electronic states of OH have been calculated from the potential energy curves we have fitted using the various neural network architectures and the extended Rydberg function. These results are then presented in Tables 1 and 2, respectively, where the BL and experimental ones are also displayed for illustration. Assuming as reference the $\Delta G_{v-1/2}$ obtained from ERF fitting, the best fits are obtained with the NN architecture 1-3*s*-1*l* for the $X^2\Pi$ state (root-mean-square error between ERF and NNs values of 2.34 cm^{-1}) and 1-6*t*-1*l* structure for the $A^2\Sigma^+$ state (root-mean-square error of 1.17 cm^{-1}). Moreover, we point out that all results from NN fitting present a good agreement with the ERF ones, even in the case where the neural networks have the smallest number of neurons.

Table 1
Experimental and calculated vibrational energy separations ($\Delta G_{v+1/2}$) for the $X^2\Pi$ state[a]

Levels	Neural network							ERF	Ref. [45]	Exp. [53–55]
	1-3*t*-1*l*	1-4*t*-1*l*	1-6*t*-1*l*	1-2*t*-2*t*-1*l*	1-3*s*-1*l*	1-4*s*-1*l*	1-6*s*-1*l*	1-2*s*-2*s*-1*l*			
0–1	3557.82	3555.21	3563.94	3570.48	3555.83	3550.12	3554.62	3557.23	3558.72	3560.36	3569.64
1–2	3392.84	3392.66	3395.56	3395.27	3392.18	3388.71	3391.67	3391.86	3392.94	3396.56	3404.03
2–3	3230.94	3231.63	3232.04	3228.59	3230.63	3229.28	3230.65	3231.45	3229.26	3238.05	3240.37
3–4	3069.95	3070.49	3070.61	3066.92	3069.44	3069.85	3069.81	3072.79	3066.64	3084.07	3077.76
4–5	2907.95	2907.79	2908.83	2906.75	2907.04	2908.68	2907.58	2912.94	2903.85	2926.75	2915.33
5–6	2743.24	2742.20	2744.49	2744.72	2741.95	2744.22	2742.50	2749.33	2739.47	2760.97	2752.17
6–7	2574.15	2572.42	2575.59	2577.74	2572.74	2574.96	2573.12	2579.76	2571.81	2585.49	2587.41
7–8	2399.01	2397.02	2400.18	2403.05	2397.83	2399.35	2397.91	2402.29	2398.87	2402.85	2420.15
8–9	2215.92	2214.34	2216.30	2218.12	2215.43	2215.67	2215.11	2215.13	2218.15	2212.93	2249.51
9–10	2022.57	2022.20	2021.77	2020.68	2023.23	2021.80	2022.48	2016.59	2026.45	2013.99	2073.85
rms error[b]	2.67	3.03	3.86	5.16	2.34	4.31	3.37	6.38			

[a] Results in cm^{-1}.
[b] Root-mean-square error between EFR and NN values.

Table 2
Experimental and calculated vibrational energy separations ($\Delta G_{v+1/2}$) for $A^2\Sigma^+$ state[a]

Levels	Neural network								ERF	Ref. [45]	Exp. [53–55]
	1-3*t*-1*l*	1-4*t*-1*l*	1-6*t*-1*l*	1-2*t*-2*t*-1*l*	1-3*s*-1*l*	1-4*s*-1*l*	1-6*s*-1*l*	1-2*s*-2*s*-1*l*			
0–1	2983.71	2989.81	2985.44	2989.58	2986.28	2987.49	2985.32	2984.81	2986.27	2987.25	2988.35
1–2	2791.37	2791.56	2792.32	2789.62	2794.77	2796.51	2795.29	2789.08	2791.24	2791.89	2793.05
2–3	2591.89	2588.18	2592.59	2586.95	2592.37	2594.71	2593.64	2590.69	2591.09	2591.63	2593.52
rms error[b]	1.55	2.65	1.17	3.20	2.17	3.76	2.82	1.52			

[a] Results in cm^{-1}.
[b] Root-mean-square error between EFR and NN values.

Another criteria to judge the accuracy of the neural network potential energy curver (NNPECs) is the calculation of Franck–Condon (FC) factors ($q_{v'v''}$) between vibrational wave functions of the ground and excited electronic states. The FC factor explicits the quality of the PEC by the analysis of the wave function of the nuclear problem. The differences between $q_{v'v''}$ calculated using NNPECs and using ERF fitting are shown in Table 3. The ERF results are shown also as reference data. We can see in Table 3 that the best results are obtained with the 1-4*t*-1*l* structure, while the neural networks with two hidden layers presented the worst results. Similar results have been verified also for the vibrational levels of the $X^2\Pi$ and $A^2\Sigma^+$ electronic states (see Tables 1 and 2).

Now, to verify the quality of the DTMFs we have fitted with various neural network structures, we have calculated the transition probabilities in $A^2\Sigma^+ \rightarrow X^2\Pi$. In particular, in Table 4 all calculations were carried out using the vibrational states obtained with the ERF fitting, being varied only the dipolar transition moment functions. This arbitrary choice was done to eliminate any influence of the PECs in the analysis of the DTMF fitting. The transition probabilities calculated by Bauschlicher and Langhoff [45] are shown also for comparison. Assuming as reference the ERF results, we can verify that the best results are those obtained with fitted DTMFs using the following NN structures: 1-3*t*-1*l* and 1-4*s*-1*l*. In spite of the worst results are calculated with the fits that used more neurons at the hidden layer, all values obtained from neural network fitting present excellent agreement with ERF ones.

The next step is then to determine the neural network transition probabilities (NNTP), i.e., the $A_{v'v''}$ calculated using the neural network dipolar transition moment functions and the vibrational states obtained from the neural network potential energy curves. In particular, we restricted the study for the case that all fittings have been carried out employing neural networks of same structure. In Table 5 we compare the calculated NNTP for the 1-3*t*-1*l*, 1-3*s*-1*l* and 1-4*s*-1*l* structures with the respective ones shown in Table 4. Note that the values of NNTP for the three structures are in excellent agreement with the transition probabilities calculated with the ERF procedure, being the NNTP obtained using the 1-3*t*-1*l* the best ones.

Finally, the radiative lifetimes Eq. (18) are determined from the transition probabilities presented in Table 5. They are then compared with the ones calculated by Bauschlicher and Langhoff [45] in Table 6. We note that the radiative lifetimes which have been obtained from neural network transition probabilities with 1-3*t*-1*l* structure present the best agreement with the ones calculated using the ERF transition probabilities.

Table 3
Calculated Franck–Condon factors ($q_{v'v''}$) between vibrational levels of $A^2\Sigma^+ \rightarrow X^2\Pi$ transition

$v'v''$	ERF	Neural network							
		1-3*t*-1*l*	1-4*t*-1*l*	1-6*t*-1*l*	1-2*t*-2*t*-1*l*	1-3*s*-1*l*	1-4*s*-1*l*	1-6*s*-1*l*	1-2*s*-2*s*-1*l*
0, 0	0.9057	−0.0016	−0.0011	−0.0018	−0.0060	−0.0004	−0.0016	−0.0009	0.0000
0, 1	0.0911	0.0016	0.0012	0.0019	0.0060	0.0005	0.0017	0.0012	0.0001
1, 0	0.0863	0.0013	0.0008	0.0015	0.0050	0.0004	0.0012	0.0005	0.0001
1, 1	0.7110	−0.0016	−0.0008	−0.0021	−0.0081	0.0013	−0.0009	0.0004	0.0024
1, 2	0.1903	0.0005	0.0002	0.0008	0.0034	−0.0013	0.0002	−0.0003	−0.0020
2, 1	0.1691	−0.0002	−0.0005	0.0001	0.0009	−0.0014	−0.0010	−0.0017	−0.0018
2, 2	0.5040	0.0011	0.0008	0.0012	−0.0015	0.0047	0.0030	0.0040	0.0062
2, 3	0.2875	−0.0006	−0.0004	−0.0011	0.0007	−0.0024	−0.0011	−0.0015	−0.0033
3, 1	0.0246	0.0001	0.0001	0.0002	0.0010	−0.0003	0.0001	0.0000	−0.0005
3, 2	0.2348	−0.0011	−0.0006	−0.0012	−0.0020	−0.0022	−0.0024	−0.0023	−0.0026
3, 3	0.3023	0.0027	0.0007	0.0040	0.0031	0.0053	0.0044	0.0055	0.0069

Table 4
Calculated vibrational transition probabilities for OH ($A^2\Sigma^+ \rightarrow X^2\Pi$) relative to $A_{00} = 1000$

$v'v''$	Ref. [45]	ERF	Neural network					
			1-3*t*-1*l*	1-4*t*-1*l*	1-6*t*-1*l*	1-3*s*-1*l*	1-4*s*-1*l*	1-6*s*-1*l*
0, 0	1000	1000	1000	1000	1000	1000	1000	1000
0, 1	4.5	4.5	4.5	4.6	4.6	4.5	4.5	4.6
1, 0	328.8	328.4	328.5	327.7	327.7	328.5	328.4	327.6
1, 1	597.4	597.5	597.4	596.9	596.5	598.3	597.7	598.5
1, 2	5.4	5.5	5.5	5.5	5.5	5.5	5.5	5.6
2, 0	68.6	68.5	68.6	68.1	68.0	68.7	68.5	68.4
2, 1	482.8	483.2	483.2	482.6	482.6	483.3	483.1	483.0
2, 2	314.6	314.7	314.7	313.9	313.5	315.5	314.9	315.2
2, 3	3.8	3.9	3.9	3.9	3.9	4.0	3.9	3.9
3, 0	12.9	12.9	12.9	12.8	12.7	13.0	12.9	13.0
3, 1	173.1	173.4	173.5	172.8	172.7	173.6	173.5	173.1
3, 2	486.6	488.6	488.5	488.1	487.9	489.1	488.6	489.4
3, 3	138.1	138.0	138.0	137.4	137.1	138.3	138.1	137.9
3, 4		1.6	1.6	1.5	1.5	1.6	1.6	1.5

All results utilize the vibrational states obtained from the ERF PEC fitting being varied only the DTMFs.

Table 5
Calculated vibrational transition probabilities ($A_{v'v''}$) for OH ($A^2\Sigma^+ \rightarrow X^2\Pi$) relative to $A_{00} = 1000$

$v'v''$	ERF	ERF + 1-3*t*-1*l*	1-3*t*-1*l*	ERF + 1-3*s*-1*l*	1-3*s*-1*l*	ERF + 1-4*s*-1*l*	1-4*s*-1*l*
0, 0	1000	1000	1000	1000	1000	1000	1000
0, 1	4.5	4.5	4.8	4.5	4.6	4.5	4.8
1, 0	328.4	328.5	331.7	328.5	329.7	328.4	332.9
1, 1	597.5	597.4	595.7	598.3	599.6	597.7	598.9
1, 2	5.5	5.5	5.5	5.5	5.4	5.5	5.6
2, 0	68.5	68.6	69.9	68.7	68.9	68.5	69.7
2, 1	483.2	483.2	482.8	483.3	481.9	483.1	485.7
2, 2	314.7	314.7	314.6	315.5	318.3	314.9	317.7
2, 3	3.9	3.9	3.8	4.0	3.7	3.9	3.9
3, 0	12.9	12.9	13.2	13.0	13	12.9	13.1
3, 1	173.4	173.5	173.7	173.6	172.0	173.5	174.0
3, 2	488.6	488.5	487.0	489.1	487.5	488.6	490.3
3, 3	138.0	138.0	138.7	138.3	140.5	138.1	140.2
3, 4	1.6	1.6	1.5	1.6	1.4	1.6	1.5

Table 6
Calculated radiative lifetimes $\tau_{v'}$ (ns) for $A^2\Sigma^+$ state

v'	Ref. [45]	ERF	1-3*t*-1*l*	1-3*s*-1*l*	1-4*s*-1*l*
0	672.2	661.3	660.8	661.5	663.1
1	724.5	712.9	711.4	710.8	710.4
2	774.9	761.9	761.0	760.1	758.4
3		811.8	811.9	812.2	809.7

5. Concluding remarks

The study we have carried out in the present work has the following motivation: to discuss the implications in the quality and accuracy of the potential energy surfaces and dipolar transition moment function when we employ different architectures of the neural network method in the fitting process using ab initio data points. For this, the vibrational eigenenergies and eigenfunctions of the $X^2\Pi$ and $A^2\Sigma^+$ electronic states, the $A^2\Sigma^+ \rightarrow X^2\Pi$ transition probabilities and the radiative lifetime of the $A^2\Sigma^+$ state of the OH molecule have been calculated from the fitted PECs and DTMFs. The results obtained using the fittings done with extended Rydberg functions were utilized as reference ones to judge the efficiency of neural network method to fit the curves of interest. From this comparison, we could confirm that the neural network method is able to fit the potential energy curves and dipolar transition moment functions of the diatomic molecules. We also could demonstrate that a neural network with one hidden layer and a reduced number of neurons is sufficient to fit PECs and DMTF with great accuracy. This result is particularly useful for multi-dimensional problems where it is of great importance the employment of interpolation functions with a small number of free parameters. Moreover, the neural network method does not require previous knowledge of the shape of the surface, uses a small set of ab initio points for the fit, and can be easily generalized to larger molecules. Therefore, the neural network approach can be considered a good choice to fit multi-dimensional PESs and DTMFs.

Acknowledgements

This work has been supported by CAPES and CNPq (Brazilian National Research Councils) through grants to the ACPB and FVP, respectively.

References

[1] G. Simons, R.G. Parr, J.M. Finlan, J. Chem. Phys. 59 (1973) 3229.
[2] V. Spirko, P. Jensen, P.R. Bunker, A. Cejchan, J. Mol. Spectrosc. 112 (1985) 183.

[3] J.N. Murrell, S. Carter, S.C. Farantos, P. Huxley, A.J.C. Varandas, Molecular Potential Energy Functions, Wiley, Chichester, 1984.
[4] N. Sathyamurthy, L.M. Raff, J. Chem. Phys. 63 (1975) 464.
[5] J.M. Bowman, J.S. Bittmann, L.B. Harding, J. Chem. Phys. 83 (1985) 660.
[6] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes, Cambridge University Press, London, 1986.
[7] P.G. Mezey, Potential Energy Hypersurfaces, Elsevier, Amsterdam, 1987.
[8] J.C. Tully, Adv. Chem. Phys. 42 (1980) 63.
[9] H.G. Yu, S. Andersson, G. Nyman, Chem. Phys. Lett. 321 (2000) 275.
[10] T.S. Ho, H. Rabitz, J. Chem. Phys. 104 (1996) 2584.
[11] G.S. Kedziora, I. Shavitt, J. Chem. Phys. 106 (1997) 8733.
[12] D.E. Makarov, H. Metiu, J. Chem. Phys. 108 (1998) 590.
[13] M.A. Collins, Theor. Chem. Acc. 108 (2002) 313.
[14] A. Frishman, D.K. Hoffman, D.J. Kouri, J. Chem. Phys. 107 (1997) 804.
[15] V. Aquilanti, G. Capecchi, S. Cavalli, C. Adamo, V. Barone, Phys. Chem. Chem. Phys. 2 (2000) 4095.
[16] B.G. Sumpter, D.G. Noid, Chem. Phys. Lett. 192 (1992) 455.
[17] T.D. Blank, S.D. Brown, A.W. Calhoun, D.J. Doren, J. Chem. Phys. 103 (1995) 4129.
[18] D.F.R. Brown, M.N. Gibbs, D.C. Clary, J. Chem. Phys. 105 (1996) 7597.
[19] K.T. No, B.H. Chang, S.Y. Kim, M.S. Jhon, H.A. Scheraga, Chem. Phys. Lett. 271 (1997) 152.
[20] F.V. Prudente, J.J. Soares Neto, Chem. Phys. Lett. 287 (1998) 585.
[21] F.V. Prudente, P.H. Acioli, J.J. Soares Neto, J. Chem. Phys. 109 (1998) 8801.
[22] C. Muñoz-Caro, A. Niño, Comput. Chem. 22 (1998) 355.
[23] S. Haykin, Neural Networks: A Comprehensive Foundation, second ed., Prentice-Hall, Englewood Cliffs, NJ, 1999.
[24] J. Zupan, J. Gasteiger, Neural Network for Chemists, VCH, Weinheim, 1993.
[25] R. Rojas, Neural Networks: A Systematic Introduction, Springer, Berlin, 1996.
[26] R.P. Brent, IEEE Trans. Neural Network 2 (1991) 346.
[27] M.F. Møller, Neural Networks 6 (1993) 525.
[28] T.B. Blank, S.D. Brown, J. Chemometrics 8 (1994) 391.
[29] D.J.C. MacKay, Network: Comput. Neural 6 (1995) 469.
[30] W. Duch, G.H.F. Diercksen, Comput. Phys. Commun. 82 (1994) 91.
[31] B.G. Sumpter, C. Getino, D.W. Noid, Annu. Rev. Phys. Chem. 45 (1994) 439.
[32] K.L. Peterson, Rev. Comput. Chem. 16 (2000) 53.
[33] K.H. Cho, K.T. No, H.A. Scheraga, J. Mol. Struct. 641 (2002) 77.
[34] M. Sugawara, Comput. Phys. Commun. 140 (2001) 366.
[35] L.E.S. de Souza, S. Canuto, Phys. Chem. Chem. Phys. 3 (2001) 4762.
[36] H.G. Bohr, K. Frimand, K.J. Jalkanen, R.M. Nieminen, S. Suhai, Phys. Rev. E 64 (2001) 021905.
[37] I.Z. Kiss, G. Mandi, M.T. Beck, J. Phys. Chem. A 104 (2000) 8081.
[38] A.P. Braga, J.P. Braga, J.C. Belchior, J. Chem. Phys. 107 (1997) 9954.
[39] G.M. Silva, P.H. Acioli, A.C. Pedroza, J. Comput. Chem. 18 (1997) 1407.
[40] I.E. Lagris, A. Likas, D.I. Fotiadis, Comput. Phys. Commun. 104 (1997) 1.
[41] J. Androsiuk, L. Kulak, K. Sienicki, Chem. Phys. 173 (1993) 377.
[42] S. Chu, M. Yoshimine, B. Liu, J. Chem. Phys. 61 (1974) 5389.
[43] S.R. Langhoff, E.F. van Dishoeck, R. Wetmore, A. Dalgarno, J. Chem. Phys. 77 (1982) 1379.
[44] E.F. van Dishoeck, A. Dalgarno, J. Chem. Phys. 79 (1983) 873.
[45] C.W. Bauschlicher, S.R. Langhoff, J. Chem. Phys. 87 (1987) 4665.
[46] D.R. Yarkony, J. Chem. Phys. 97 (1992) 1838.
[47] I.L. Chidsey, D.R. Crosley, J. Quantum Spectrosc. Radiat. Transf. 23 (1980) 187.
[48] A.J.C. Varandas, A.I. Voronin, Chem. Phys. 194 (1995) 91.
[49] J. Luque, D.R. Crosley, J. Chem. Phys. 109 (1998) 439.
[50] J.C. Light, T. Carrington Jr., Adv. Chem. Phys. 114 (2000) 263.
[51] F.V. Prudente, A. Riganelli, A.J.C. Varandas, Rev. Mex. Fis. 47 (2001) 568.
[52] R.G. Littlejohn, M. Cargo, T. Carrington Jr., K.A. Mitchell, B. Poirier, J. Chem. Phys. 116 (2002) 8691.
[53] J.A. Coxon, J. Mol. Spectrosc. 58 (1975) 1.
[54] J.A. Coxon, K.V.L.N. Sastry, J.A. Austin, D.H. Levy, Can. J. Phys. 57 (1979) 619.
[55] J.A. Coxon, Can. J. Phys. 58 (1980) 933.
[56] E.D. Sontag, Neural Comput. 1 (1989) 470.
[57] C.M. Bishop, C.M. Roach, Rev. Sci. Instrum. 63 (1982) 4450.
[58] J.V. Lill, G.A. Parker, J.C. Light, J. Chem. Phys. 89 (1982) 483.
[59] J.T. Muckerman, Chem. Phys. Lett. 173 (1990) 200.
[60] D.T. Colbert, W.H. Miller, J. Chem. Phys. 96 (1992) 1982.
[61] F.V. Prudente, L.S. Costa, J.J. Soares Neto, J. Mol. Struct. (Theochem) 394 (1997) 169.