JAAS

PAPER

View Article Online
View Journal

Cite this: DOI: 10.1039/c6ja00226a

Corona discharge induced plasma spectroscopy (CDIPS) for quantitative analysis of gas mixtures

J. Vorobioff,[a] N. Boggio,[abc] S. Moncayo,[d] J. O. Caceres[d] and C. A. Rinaldi*[abc]

Received 24th June 2016
Accepted 23rd August 2016

DOI: 10.1039/c6ja00226a

www.rsc.org/jaas

The capability of the Corona Discharge Induced Plasma Spectroscopy (CDIPS) technique has been demonstrated for quantitative analysis of nitrogen–oxygen mixtures. Optical emission of the induced plasma in combination with multivariate chemometric methods of calibration such as Partial Least Squares (PLS) regression and Neural Networks (NNs) was able to produce a highly dynamic calibration curve (from 0% to 100% of O_2/N_2 mixtures) with a high sensitivity, for the measurement of medical grade O_2 and N_2. The predictive ability of NNs, supported by nonlinear data modelling, can be useful in the quantitative analysis, overcoming experimental effects and producing a smaller relative error in comparison to PLS. For validation purposes, all samples were also analyzed by gas chromatography showing an excellent agreement in the O_2/N_2 compositions measured within the limit of relative error. The methodology presented shows high potential as an alternative method for the monitoring and analysis of oxygen and nitrogen mixtures in production systems, especially for medical applications.

1. Introduction

The Corona Discharge Induced Plasma Spectroscopy (CDIPS) technique is based on generating plasma from an electrical corona discharge across a constant gas flow. The luminous plasma is composed of a set of partially ionized atoms, free electrons, excited atoms, molecules and a large number of neutral atoms,[1–3] and is characterized by the law of Peek–White.[4] The radiation from the plasma provides information about its physical properties such as composition, electron density and temperature. Thus, the behavior of the emitter species can be correlated with changes in the density of ions and the electron temperature allowing the quantification of such species. The emission from the plasma is collected with a UV-visible spectrometer obtaining a spectrum consisting of discrete and narrow atomic emission lines and wider bands of molecular species such as N_2 or O_2 characteristic of the elements present in the plasma. The band emission results from the decay of excited molecules into different rotational and vibrational levels. In the case of nitrogen the processes that lead to band emission are direct excitation, direct excitation and ionization and dissociative ionization and excitation.[5] Direct excitation $e^- + N_2 \rightarrow N_2^* + e^-$ corresponds to the First Positive Band (FPS) and Second Positive Band (SPS) of N_2 ($\lambda = 300$ and 400 nm). Direct excitation and ionization $e^- + N_2 \rightarrow (N_2^+)^* + e^-$ correspond to the First Negative System (FNS) Band of N_2 with the dominant emission line at $\lambda = 391.44$ nm. Dissociative ionization and excitation generate excited singly charged atoms, $e^- + N_2 \rightarrow (N^+)^* + N$ with main transitions at $\lambda = 399.5$ nm, 648 nm, 661 nm and 885 nm. These emissions are weak compared to those of excited molecular ions. The most intense emission of oxygen results from excited neutral atoms, originating from dissociative excitation: $e^- + O_2^+ \rightarrow O^* + O$, and is mainly associated with transitions between $3p^5P–3s^5S^0$ ($\lambda = 777.19$ nm) and $3p^3P–3s^3S^0$ ($\lambda = 844.67$ nm).

These characteristic signals can be verified by their correspondence with the data obtained from the simulated spectra, which provide the intensities at a given wavelength of different species and the theoretical estimation of parameters as the rotational and vibrational temperatures as well as their concentration.[5,6] There are many scientific articles concerning fundamental spectroscopy studies such as E. H. Lock *et al.*[5] have been working on a theoretical and experimental estimation of excited species generated by the electron beam. Staack *et al.*[7] characterized the spectra of a DC glow microplasma discharge at atmospheric pressure. Machala *et al.*[8] demonstrated the use of UV-visible spectroscopy for the characterization of atmospheric pressure air and nitrogen plasmas in environmental applications. Shin *et al.*[9] described a pulsed corona discharge system for the removal of air pollutants, in which emission spectroscopy is employed to characterize the generation of excited species by the corona discharge process. Previously, Hrachová *et al.*[10] studied the influence of impurities on the

[a]Departamento de Micro y Nanotecnología, Comisión Nacional de Energía Atómica – CAC, Av.Gral.Paz 1499, (1650) San Martin, Bs. As., Argentina. E-mail: rinaldi@cnea.gov.ar

[b]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), CABA, Godoy Cruz 2290, 1425, Argentina

[c]Escuela de Ciencia y Tecnología, Universidad Nacional de San Martin, Av.25 de Mayo y Francia, (1650) San Martin, Bs. As., Argentina

[d]Laser-Chemistry Lab, Faculty of Chemistry, Complutense University, 28040 Madrid, Spain

emission spectra of oxygen and nitrogen under different vacuum conditions and the influence of the presence of small quantities of rare gases or nitrogen on the oxygen spectra. However, the quantification of gas mixtures by CDIPS using chemometric methods, to the best of our knowledge, has not been reported before. Therefore, the goal of this study is to develop a new methodology to quantify the components of oxygen–nitrogen mixtures through the corona discharge emission spectra using multivariate regression methods and assess the performance/capabilities of partial least-squares (PLS) and artificial neural networks (NNs) for quantitative analysis. This method is being proposed as a cost effective and simple alternative to the already commercially implemented standard methods.

2. Experimental

The elements that compose the experimental setup, called hereinafter Corona Discharge Induced Plasma Spectroscopy (CDIPS), are given as follows.

2.1 Instrumentation

The custom made CDIPS equipment is shown in Fig. 1. The setup of the system has three main parts: (1) a discharge chamber, (2) a gas supply system that controls the injection, composition, working pressure and flow and (3) a spectrometer for spectra acquisition.

2.1.1 CDIPS system. The discharge chamber (1), shown in Fig. 1, was made of Teflon and had a volume of 0.3 cm^3. Two electrodes were inserted into the chamber forming a coaxial needle–cylinder geometry, where the central electrode (anode) was made up of a stainless steel needle and the cathode consisted of a copper cylinder with an internal diameter of 2 mm. The radius of the central electrode was 0.15 mm and the distance between the electrodes was fixed at 0.85 mm. Gases were introduced into the discharge chamber with a micro-pump at a flow rate of 100 mL min^{-1} and the working pressure was fixed at 0.8 atm.

To optimize the performance of this system the parameters that govern the corona discharge were carefully studied.

2.1.2 Corona discharge parameters. The main characteristics of plasma discharge, breakdown voltage, voltage–current relationship and discharge structure depend on the following parameters: cylinder geometry, electrode material, used gas, the pressure inside the chamber and the external circuit.[1,2,7] The calculation of the inception voltage V_i was obtained from the expression of Peek–White's law:[4]

$$V_i = 3 \times 10^3 a \left(\frac{T_0 P}{T P_0} + 0.3 \sqrt{\frac{10 T_0 P}{T P_0 a}} \right) \ln\left(\frac{4d}{\pi a}\right)$$

where d is the cylinder–needle distance (0.85 mm), a is the radius from the central electrode (0.15 mm), P_0 is the atmospheric pressure (101 325 Pa), P is the experimental pressure (81 060 Pa), T_0 is the reference temperature (298 K), T is the experimental temperature (298 K) and the working pressure is 0.8 atm. Under these conditions, in accordance with the dimensions described above, the inception voltage (V_i) required to start the discharge was estimated and set at 2.66 kV while the experimental voltage applied to maintain a stable discharge was set at 2.70 kV by maintaining a constant current of 0.5 mA.

2.1.3 Gas supply system. In order to establish the dynamic range of the CDIPS, two sets of gaseous mixture samples were analyzed (Table 1). The samples were prepared by mixing oxygen and nitrogen (both 99.999% from Indura S.A, Argentine with a content of H_2O < 8 ppm) in a 2 L cylinder controlling the pressure with a digital manometer. The minor gas fraction was first introduced and then the mixture was completed by the major gas fraction until the pressure reached the value corresponding to the desired composition. The relative error in the preparation of the samples was 1%. Sixteen different concentration samples were prepared as described in Table 1 and each

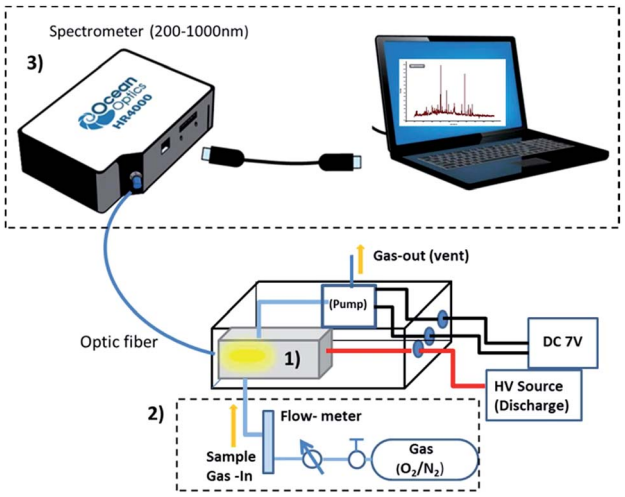

Fig. 1 Scheme of the experimental setup of the plasma generator system, (1) discharge chamber, (2) gas supply system and (3) spectrometer.

Table 1 Composition of the samples of N_2/O_2 mixtures

	Sample ID	O_2% (v/v)	N_2% (v/v)
SET 1	1	0	100
	2	15	85
	3	22	78
	4	30	70
	5	50	50
	6	80	20
	7	90	10
	8	93	7
	9	100	0
SET 2	10	80	20
	11	84	16
	12	85	15
	13	93	7
	14	97.7	2.3
	15	99	1
	16	100	0
	Unknown sample	?	?

sample was analyzed in triplicate. An evacuated cylinder was filled with a medical oxygen sample produced by a Pressure Swing Adsorption (PSA) Oxygen Generator (model AS-12 from AirSep, USA), named as "unknown sample" (Table 1). The composition of the samples was checked by gas chromatography.

2.1.4 Spectrometer for the spectra acquisition. The emission from the plasma was collected with a 4 mm aperture, and 7 mm focus fused silica collimator placed at 4 cm from the plasma and then focused into an optical fiber (1000 μm core diameter, 0.22 numerical aperture), coupled to a spectrometer. The spectrometer system was an Ocean Optics HR4000, with a CCD detector. A grating of 300 lines per mm was selected; a spectral resolution of 0.5 nm was achieved using a 7 μm entrance slit. The wavelength range used was 200 to 1000 nm. To optimize the signal-to-noise ratio, spectra were acquired by establishing an integration time of 2 s and averaging 5 spectra for each measurement. The plasma spectrum was acquired and stored as a column on a dataset. The dataset contains the intensity at different wavelengths in rows and the spectra in columns. Thus, our dataset has 2048 rows (one for each wavelength) and 15 columns (spectrum) for each sample. Each dataset (spectral library) contains the spectra at a given concentration. Ten spectra were used to create the training model considering each spectra library as a fingerprint of the sample, and the remaining five spectra were used to test the chemometric models.

2.1.5 Chromatographic set-up. Chromatographic analysis was performed using a standard gas chromatographic system, Buck Scientific model 910. A molecular sieve type column, helium as carrier gas at a flow of 85 mL min^{-1}, at a temperature of 50 °C and a TCD detector were used. A calibrated volume of 1 mL of the gas sample was injected.

3. Chemometric models

For quantitative elemental analysis of gases by the CDIPS technique, one possibility is to relate directly the emission line intensity of different elements in the plasma to the concentration of those elements in the gas (conventional calibration curve). However, the signal is easily affected by several experimental parameters that are difficult to control. The application of chemometric methods is necessary to take into account all these parameters providing a more reliable calibration.

Artificial Neural Networks (NNs) and Partial Least Squares (PLS) have been chosen as quantitative calibration methods and Principal Component Analysis (PCA) has been used for variable reduction for the NN model. Chemometric methods evaluated in this work have been widely described in the literature and only a brief discussion is presented here.

3.1 Principal component analysis

Using a complete spectrum to feed the neural network as input increases the model complexity and the presence of noise and redundant data makes the model inefficient. As an alternative, a reduction in the number of input data could be done in order to increase the prediction ability of the NN. In other induced plasma spectroscopy techniques this data reduction is done based on different approaches. One of them is to use the line intensity of the characteristic emission lines of the element of interest. However, it does not take into account important plasma information related to the properties of the line such as its width and shape. Another approach is to consider different wavelength ranges comprising the main emission lines and reproduce the spectrum without the loss of meaningful information.[11] In this work, a new approach to reduce the number of input variables has been considered which is based on a previous Principal Component Analysis (PCA). PCA involves a mathematical transformation of the data represented by eqn (1):

$$\boldsymbol{X} = \boldsymbol{TP} + \boldsymbol{E} \tag{1}$$

where $\boldsymbol{X}$ is the input data matrix of dimension $\boldsymbol{I} \times \boldsymbol{J}$. $\boldsymbol{I}$ is the number of reference spectra and $\boldsymbol{J}$ is variables (value at each wavelength). The individual variables (columns) of $\boldsymbol{X}$ are denoted as $\boldsymbol{x_j}$ and are all vectors in the $\boldsymbol{I}$-dimensional space. A linear combination of those $\boldsymbol{x}$ variables can be written as $\boldsymbol{t} = \omega_1\boldsymbol{x}_1 + \ldots \omega_J\boldsymbol{x_j}$, where ω is the weight of variables and $\boldsymbol{t}$ is a new vector called score, in the same space as the $\boldsymbol{x}$ variables. $\boldsymbol{T}$ is the score matrix with dimension $\boldsymbol{I} \times \boldsymbol{A}$, where $\boldsymbol{A}$ is the number of principal components (PCs) considered. $\boldsymbol{P}$ is the loading matrix with the $\boldsymbol{A} \times \boldsymbol{J}$ dimension, where each vector $\boldsymbol{p_A}$ contains the regression coefficient. The principal component is defined for the pair of eigenvector $\boldsymbol{t}$ and $\boldsymbol{p}$.

3.2 Neural networks

The NN topology consists of several neurons arranged in different layers (input, hidden and output).[12,13] The input data matrix is composed of the intensity emission values at each wavelength for all the reference samples and the target vector is generated with the actual concentration of such samples. The NN processing consists of two steps: first, the input dataset is successively presented to the neural networks in the training process and a comparison of the prediction output concentration with the actual concentration is carried out. The mean square error (MSE) shown in eqn (2) between the output of the NN and the target is calculated until a fixed value is reached improving the predictive ability of the network.

$$\mathrm{MSE} = \frac{1}{N}\sum_{k}^{N}(r_k - y_k)^2 \tag{2}$$

where N, y_k, and r_k are the number of input data, the response from each output neuron, and the observed output response, respectively.

Once the NN is trained, the weights are fixed and new spectra from unknown samples are tested obtaining the predicted concentration. As in other calibration procedures based on NNs,[14–17] a supervised multilayer feed forward perceptron model has been used due to its ability to model systems with a similar level of complexity.[18–20] A detailed description of the calculation process is provided in the literature.[21,22] The main advantage of

the method is the capacity to model non-linear effects, which are usually present in plasma spectra. Although the data matrix could be considerably large, the computation time in the training of the NN was always below 10 s. In this case several PCs have been used as numerical inputs to the NN model. An important decision is the number of components that should be used to retain the meaningful information for the entire reference dataset. Herein, the number of PCs was established by performing a cross-validation procedure and finally 6 PCs were used to estimate the NN model.

In order to ensure a high prediction ability of the NN model, the number of hidden layers has to be kept as low as possible. A large number of hidden layers produce a NN unable to generalize. It recalls the training set to perfection, but does not predict unknown samples for the model. Therefore, the number of hidden layers was set to one. Finally, NN architecture consists of 6 input neurons, one hidden layer and six output neurons (reference concentrations).

3.3 Partial least squares (PLS)

Partial Least Squares (PLS) is a statistical method that finds the minimum variance between the response and independent variables, projecting the predicted variables and the observable variables into a new space in a linear regression model.[23] The basic assumption of this model is that the system under study depends on a small number of latent variables (LVs). The latent variables are estimated as linear combinations of the observed variables. When a process model is used for the prediction of process variables, the quality of prediction must be evaluated. The cross-validation is a convenient and reliable method for assessing the predictive ability of the model and was therefore adopted as the standard in the analysis by PLS. The performance parameter for the models was calculated by using the root mean square error of prediction (RMSEP), and the goodness of the model was evaluated by using the coefficient of determination (R^2). Through this analysis, it is possible to determine the concentration of the sample analyzed. Partial least squares was performed using a free computational tool "Chemoface Software".[24]

4. Results and discussion

The typical spectrum acquired by the CDIPS system is shown in Fig. 2. As can be seen, in the case of pure N_2, the set of lines from the second positive system with the most intense line at 337.00 nm stands out between 300 nm and 400 nm and at a smaller intensity the emission lines from the first positive system lie between 600 and 800 nm. For pure O_2, emission lines at 777.40 nm and 844.00 nm lines predominate. In the case of mixtures, the intensities of the mentioned peaks decrease according to the amounts of both components. In order to study the linearity response of the system, the integral signal of the oxygen peak at 777.40 nm as a function of the O_2 concentration was plotted in Fig. 3. The non-linearity of the response can be observed, mainly in the high oxygen concentrations. Subsequently the results obtained using the CDIPS were analyzed using two chemometric methods, Neural Networks (NNs)[12] and Partial Least Squares (PLS).[11] These methods allowed us to obtain quantitative information on the quality of the gas mixture and the subsequent analysis of a real sample, coming from the medical oxygen generator.

4.1 NN analysis

In this work, an NN-UV/VIS approach was used to determine the oxygen and nitrogen concentrations in standard mixture

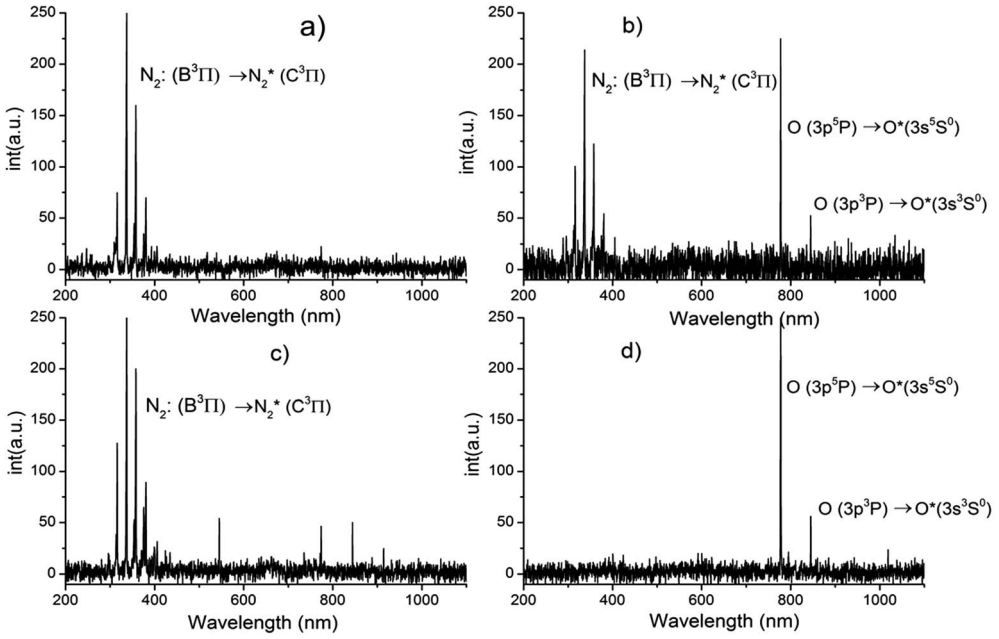

Fig. 2 Emission spectra corresponding to the samples analyzed: (a) pure nitrogen, (b) mixture N_2/O_2 10/90%, (c) mixture N_2/O_2 50/50% and (d) pure oxygen. The main transitions are indicated.

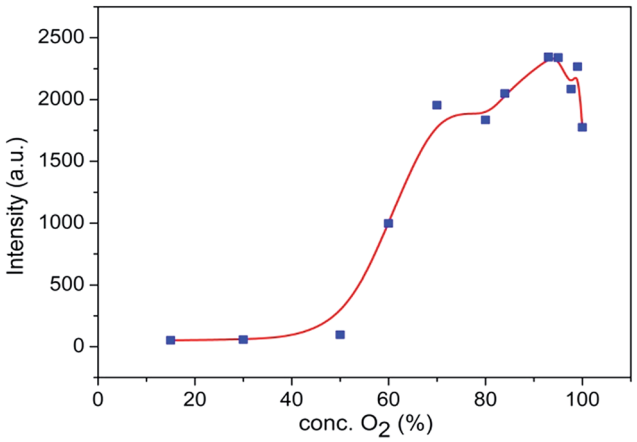

Fig. 3 Intensity of the oxygen peak at 777.40 nm *vs.* O_2 concentration. The continuous line in the graph has been shown only to help the viewer see the trend and does not indicate the fit line.

samples. The spectral information obtained using the spectrometer (measurement of the emission intensity) was previously analyzed by PCA and the scores of the first 6 PCs were used as the input for the NN model, and the actual concentration of reference samples (10 to 16 from Table 1) was used as the target vector. A total of 180 spectra were used for the NN training, considering thirty spectra for each reference sample. Fig. 4 shows the calibration curve of the predicted *versus* reference values of the quantitative analysis by NNs applying a previous PCA analysis for dimensionality reduction. The error bars were calculated taking into account the standard deviation of the output values for each reference concentration.

After the training of the NN, the obtained calibration curve was evaluated introducing into the model the spectra of the PSA sample generating a prediction value. Forty-five spectra of the PSA sample were estimated obtaining an average value of 99.0% of O_2 with a standard deviation of 0.1%. This value indicates that the NN can predict elemental concentrations in the range of 80–100%, for an unknown sample. As can be seen in this case the relative error is less than the relative error of the sample preparation.

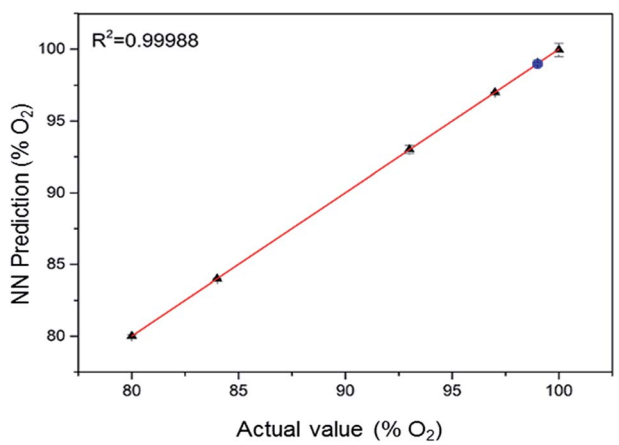

Fig. 4 NN correlation plot considering one hidden neuron.

4.2 PLS analysis

Table 2 shows the performance parameters R^2cal, R^2cv, RMSEC and RMSECV for the calibration done. According to these values in the case of a long dynamic range (0–100% O_2) a number of 4 LV is enough to obtain a precision of 6% and in the case of a short dynamic range (80–100% O_2) a number of 3 LV is enough to obtain a precision of 2.9%.

Then, the optimal number of LVs corresponded to a compromise that allowed obtaining a model presenting for a hand the relative lowest RMSECV and for other hand the relative highest R^2cv. This criterion is commonly used in these types of PLS models.[25]

Fig. 5 shows the correlation plot obtained for PLS analysis.

After the optimization of the calibration curve parameters, the spectra were used for quantification. The method[24] normally uses some of the samples as a test in order to determine the quality of the calibration. Thus, Fig. 5 presents the correlation plot in which samples containing 93% O_2 were used as a test dataset.

Table 2 PLS values of parameters R^2cal (determination coefficient for the calibration of the PLS model), R^2cv (determination coefficient for the cross-validation of the PLS model), RMSEC (root mean square error of calibration) and RMSECV (root mean square error of cross-validation) as a function of the number of LVs (latent variables)

Range of O_2 conc. (%)	Number of LV	R^2cal	R^2cv	RMSEC (% O_2)	RMSECV (% O_2)
LDR (0–100%)	2	0.90	0.90	10.87	11.14
	3	0.97	0.96	5.80	6.80
	4	0.99	0.96	3.48	6.42
	6	0.99	0.96	2.08	6.50
SDR (80–100%)	2	0.88	0.74	4.94	5.41
	3	0.90	0.84	2.78	3.35
	4	0.92	0.87	2.09	2.90
	6	0.98	0.89	1.03	2.51

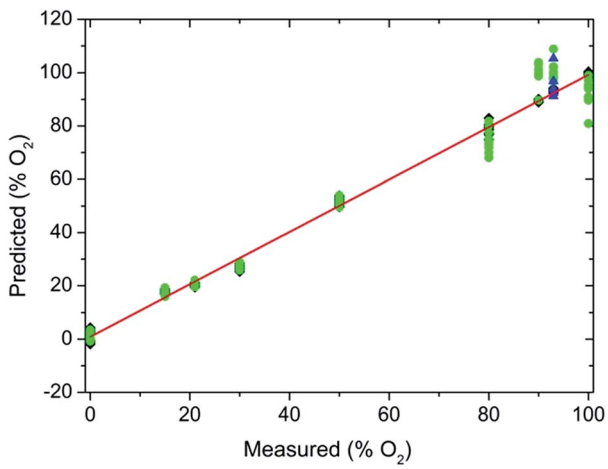

Fig. 5 PLS correlation plots of predicted *vs.* certified values, which include spectra of samples containing 93% O_2/7% N_2 included as a test (blue triangles). Black diamonds: calibration points. Green dots: cross-validation.

According to the results obtained using this method, it is possible to obtain a high linearity over the entire range of concentrations studied. However, an increase in the dispersion of spectral data is observed in Fig. 5 when the oxygen content in the sample increases from 80 to 100%. This situation is associated with discharge instabilities as it changes the composition of the injected sample. It is possible that the voltage/current ratio needs to be kept constant in smaller ranges of composition in order to perform several calibrations at different voltages, obtaining greater stability and therefore less spectral dispersion in the calibration range. Subsequent to the analysis of the samples in all ranges of compositions (samples 1–9, Table 1) the same methodology was applied to the set of samples with high concentrations of O_2 (samples 10–16, Table 1). The correlation plot was obtained applying normalization of spectral data and then PLS regression as in the previous case. The determination coefficients R^2, RMSEC and RMSECV values as a function of the latent variables (LVs) are shown in Table 2 and the correlation plot obtained is shown in Fig. 6. As can be seen, in both cases (LDR and SDR) the number of LVs is practically the same indicating that it is possible to perform a calibration in the entire range of concentrations, since the amount of variables that affect the calibration is the same.

The following step was performed to obtain the quantification prediction of the unknown sample. The results are shown in Fig. 6. The predicted value for the unknown sample (medical oxygen PSA) was (99 ± 3)% O_2, except for three samples in the lower part of the regression line that are considered outliers.

In order to compare the percentage of O_2 present in the oxygen PSA sample with the value obtained for spectrometry, its composition was determined by gas chromatography. As can be observed in Fig. 7 the ratio of areas under the curves indicates that the composition of the oxygen PSA sample is (98 ± 1)% O_2.

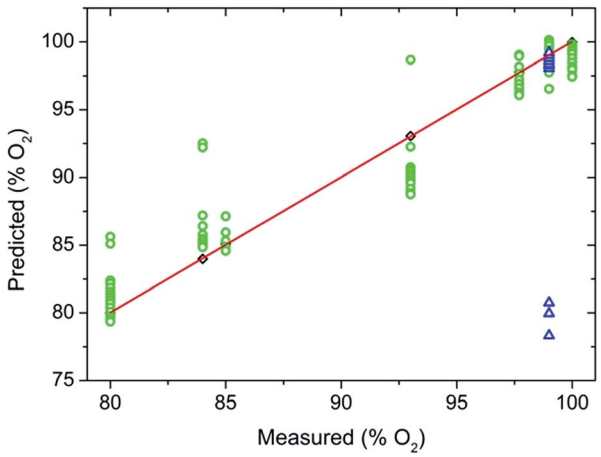

Fig. 6 PLS correlation plots of predicted *vs.* certified values with the results of the "Unknown sample" included as a test sample (blue triangles). Black diamonds: calibration points. Green circles: cross-validation points.

Table 3 Figures of merit for comparison of the different PLS methods and the GC analysis

	CDIPS data analysis methods			
Figure of merit	NN	LDR-PLS	SDR-PLS	Chromatography
Accuracy (%)	99.0	93	99	98
Precision (%)	±0.1	±6	±3	±1
Dynamic range (%)	80–100	0–100	80–100	0–100

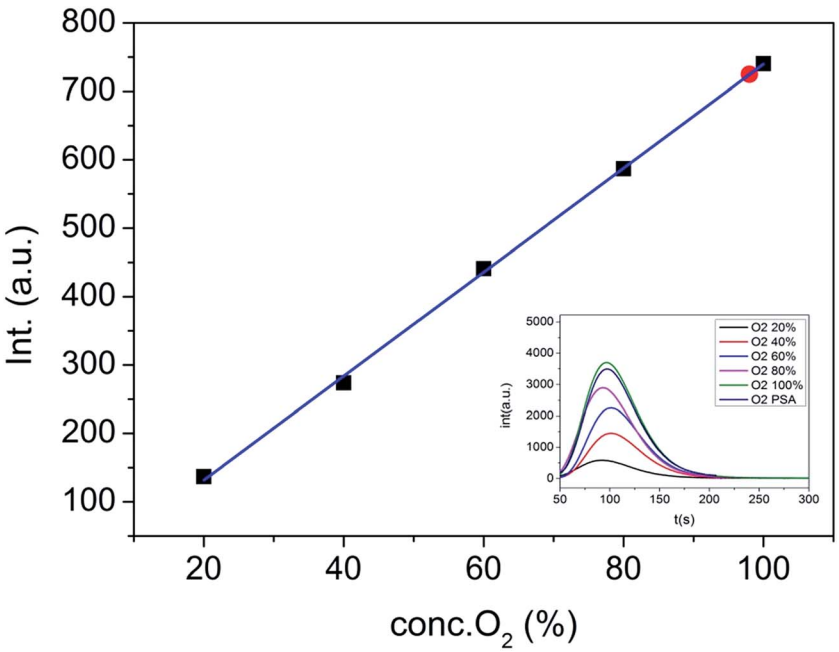

Fig. 7 Chromatographic calibration curve of samples of pure oxygen. The red dot represents the sample of medical oxygen PSA. In the inset, the chromatograms of the samples are shown.

Finally, Table 3 presents a set of figures of merit for the purpose of comparison between the performance of the developed system and analysis by gas chromatography.

5. Conclusions

In the present study, a new methodology based on corona discharge induced plasma emission spectrometry has been developed and applied to quantify the composition of binary mixtures of O_2/N_2 over the entire range of compositions in combination with chemometric methods. The methodology was validated with gas chromatography measurements achieving a precision lower than 1%. The predicted concentration values are closer to those determined through standard chromatographic analysis. The use of non-linear chemometric methods such as NNs provided an analysis with high generalization ability able to deal with experimental variables such as fluctuations in measurement conditions. This characteristic produces better prediction values and avoids the need to generate a NN model for each experimental condition. It is also clear that the reduction and selection of the inputs for the NN model clearly improve the quantitative prediction capacity of the NN. The application of PLS to spectroscopy data obtained by CDIPS has been demonstrated to be a useful tool as a predictive model in the analysis of binary gas mixtures in a wide range of compositions.

Acknowledgements

The authors gratefully acknowledge financial support from the Complutense University of Madrid, the National Council Of Scientific and Technical Research (CONICET), the University of San Martín, Buenos Aires, and Bell Export enterprise from Argentina.

References

1 Y. P. Raizer, *Gas Discharge Physics*, Springer, 1991.
2 B. M. Smirnov, *Theory of Gas Discharge Plasma*, Springer, 2015.
3 Y. W. Yoon, S. Y. Chae, M. Lim and S. K. Lee, *Chem. Phys. Lett.*, 2015, **634**, 118–121.
4 Z. Bo, G. Lu, P. Wang and J. Chen, *Ozone: Sci. Eng.*, 2013, **35**(1), 31–37.
5 E. H. Lock, R. F. Fernsler, S. Slinker, and S. G. Walton, *Experimental and Theoretical Estimation of Excited Species Generation in Pulsed Electron Beam-Generated Plasmas Produced in Pure Argon, Nitrogen, Oxygen, and Their Mixtures*, Naval Research Lab., U.S. Navy, 2011.
6 NIST Atomic Spectra Database, USA, http://www.nist.gov/pml/data/asd.cfm, accessed Dec. 2015.
7 D. Staack, F. Bakhtier, G. Alexander and F. Alexander, *Plasma Sources Sci. Technol.*, 2005, **14**(4), 700.
8 Z. Machala, M. Janda, K. Hensel, I. Jedlovský, L. Leštinská, V. Foltin, V. Martišovitš and M. Morvová, *J. Mol. Spectrosc.*, 2007, **243**(2), 194–201.
9 D. N. Shin, C. W. Park and J. W. Hahn, *Bull. Korean Chem. Soc.*, 2000, **21**(2), 228–232.
10 V. Hrachová, O. Kylián and A. Kanka, *Vacuum*, 2004, **76**, 433–436.
11 R. G. Brereton, *Applied Chemometrics for Scientists*, John Wiley & Sons Ltd., Chichester, UK, 2007.
12 A. J. Maren and C. T. Harston, *Handbook of Neural Computing Applications*, Academic Press, San Diego, USA, 1990.
13 S. Curteanu and H. Cartwright, *J. Chemom.*, 2011, **25**(10), 527–549.
14 C. M. Bishop, *Rev. Sci. Instrum.*, 1994, **65**(6), 1803–1832.
15 F. Despagne and D. L. Massart, *Analyst*, 1998, **123**(11), 157R–178R.
16 J. S. Torrecilla, M. Cámara, V. Fernández-Ruiz, G. Piera and J. O. Caceres, *J. Agric. Food Chem.*, 2008, **56**(15), 6261–6266.
17 Q. Shen, W. Zhou and K. Li, *Proc. SPIE 7854, Infrared, Millimeter Wave, and Terahertz Technologies*, Beijing, China, 2010, 78543Q, DOI: 10.1117/12.873406.
18 J. B. Sirven, B. Bousquet, L. Canioni and L. Sarger, *Anal. Chem.*, 2006, **78**(5), 1462–1469.
19 M. Cámara, J. S. Torrecilla, J. O. Caceres, M. C. Sánchez Mata and V. Fernández-Ruiz, *J. Agric. Food Chem.*, 2010, **58**(1), 72–75.
20 S. Moncayo, S. Manzoor, F. Navarro-Villoslada and J. O. Caceres, *Chemom. Intell. Lab. Syst.*, 2015, **146**, 354–364.
21 C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, USA, 1996.
22 H. B. Demuth, M. H. Beale and M. T. Hagan, *Neural network toolbox for use with MATLAB, User's guide 9th for version 6.0* (Release 2008a), https://filer.case.edu/pjt9/b378s10/nnet.pdf.
23 D. L. Massart, *Handbook of Chemometrics and Qualimetrics, Volume Part B*, Elsevier Science, 1998.
24 C. A. Nunes, M. P. Freitas, A. C. Pinheiro and S. C. Bastos, *J. Braz. Chem. Soc.*, 2012, **23**(11), 2003–2010.
25 C. Camps, R. Robic, M. Bruneau and F. Laurens, *LWT–Food Sci. Technol.*, 2010, **43**, 1164–1167.