

Multilayer perceptron as inverse model in a ground-based remote sensing temperature retrieval problem

Esteban García-Cuesta^{a,*}, Inés M. Galván^b, Antonio J. de Castro^a

^aPhysics Department, Universidad Carlos III Avda. Universidad 30, 28911 Leganés, Madrid, Spain

^bComputer Science Department, Universidad Carlos III Avda. Universidad 30, 28911 Leganés, Madrid, Spain

Abstract

In this paper, a combustion temperature retrieval approximation for high-resolution infrared ground-based measurements has been developed based on a multilayer perceptron (MLP) technique. The introduction of a selection subset of features is mandatory due to the problems related to the high dimensionality data and the worse performance of MLPs with this high input dimensionality. Principal component analysis is used to reduce the input data dimensionality, selecting the physically important features in order to improve MLP performance. The use of a priori physical information over other methods in the chosen feature's phase has been tested and has appeared jointly with the MLP technique as a good alternative for this problem.

Keywords: Neural networks; Principal component analysis; Dimensionality reduction; Retrieval temperature; Inverse models; Ground base remote sensing; Combustions

1. Introduction

This work concerns a method for the retrieval of physical properties (temperature) in a combustion process, which is very useful to monitor and control these kinds of processes in order to minimise pollutant emissions and to optimise energy losses also, for example in the industrial fuel-fired furnace. The flame temperature appears, among others, as a very important parameter to be monitored (Romero et al., 2005; Thakur et al., 2001; Liu and Jiang, 2001; Lu et al., 2004). Usually thermocouples have been used but are intrusive and disturb the measurement, and they must undergo the harsh furnace environment also. Remote optical measurements are more suitable because they are non-intrusive, and the interaction of electromagnetic radiation with matter modifies to some extent the incident wave. The infrared region is important because the majority of the emission of hot gases in the flame are due to CO₂ and H₂O, which exhibit important emission in this

region. These kinds of emissions are governed by the radiative transfer equation (Goody and Yung, 1989)

$$R_i = (I_0)_i \tau_i(z_0) + \int_{z_0}^{\text{inf}} B_i\{T(z)\} K_i(z) dz, \quad (1)$$

where i 's are the different channels.

This equation expresses the forward problem for the channel. However, when we receive the measurement from a spectroradiometer, we are faced with the inverse problem: given the measurements of energy, guess what is the state of the flame or the combustion in terms of spatial temperature and constituents. This inversion problem of Eq. (1) is ill-posed or under-constrained since we are trying to retrieve $T(z)$, a continuous function from a finite measurements (McCormick, 1992). Progress in optoelectronics has helped to enlarge the amount of data obtained in these kinds of measurements allowing a better approximation solving the inverse problem. However, this amount of data increases in a way that makes difficult the use of conventional data regression techniques and also many problems could be found in these kinds of retrievals, related to data dimension or complexity of regression

*Corresponding author. Tel.: +3491 6249962; fax: +3491 6248749.
E-mail addresses: egc@fis.uc3m.es (E. García-Cuesta), igalvan@inf.uc3m.es (I.M. Galván), decastro@fis.uc3m.es (A.J. de Castro).

models, which implies a large number of operations to solve it. Also this data dimensionality will lead to a problem called the curse of dimensionality (Bellman, 1961), which says in practice that we could find a certain point beyond it; adding new features can actually lead to a reduction in the performance of the machine learning system. This encourages the use of new techniques for these kinds of problems.

The artificial neural networks seem to be an interesting alternative technique to be used in these retrieval problems. One important advantage of neural networks in this field is their speed. Once the neural network has been trained the inversion method is almost instantaneous in comparison with other models as physical iterative methods, which use the negative gradient descent to find the true profile (Eyre, 2004). Advantages over other physical–statistical techniques are that they do not need a good initial condition for the inversion and do not need a rapid direct model for iterative inversion algorithms.

The goal of this study is to present an inversion method that retrieves the temperature profile of a hot gas cloud composed by CO₂ and water vapour from spectroradiometric measurements. This problem is related to ill-posed problems and corresponds with inverse radiative problems (McCornick, 1992). To do it, a multilayer perceptron (MLP) approximation has been adopted as inverse model for the radiative transfer equation (1).

In previous works, neural networks have been used to retrieve the atmospheric temperature (Aires et al., 2001), although the problem here is different because the temperature ranges are bigger and hotter, and in atmospheric retrieval some parameters are known a priori, which simplifies the inverse model under some assumptions. In our case we have a complex relationship between optical path and temperature whose influence varies in the straightforward model following Beer’s law, and in the inverse model we do not know how this relationship is because both factors have a non-linear influence in all the spectrum measurements.

As it has been previously mentioned, in this context the data dimensionality is so high and it could influence the performance capacity of MLP. Principal component analysis (PCA) has been used to reduce the number of input neurons of MLP. PCA is a multivariate statistical analysis introduced by Pearson (1901), and developed independently by Hotelling (1933). PCA involves a mathematical procedure that transforms a number of correlated variables into a smaller number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible.

There are many different techniques to make the selection of those variables which best represent the variability of the data which is being studied. A generalised method called B4 is widely used in many scenarios (Jolliffe, 2002). Here the results obtained with this method are not

good enough and a pick peak selector has been developed to make the selection of variables.

The rest of the paper is organised as follows: In Section 2, a context description of retrieval of temperature profiles in flames is made. Section 3 describes the retrieval methodology with high-resolution spectra. This section includes the use of PCA to reduce the dimensionality of the input space for the MLP and an explanation of how the selection of variables has been done. Section 4 describes the physical characteristics of the data which fit the combustion processes we are working in, and the obtained results. Finally, conclusions are presented in Section 5.

2. Retrieval of temperature profiles in flames

In an industrial fuel-fired furnace, it is very important to have devices that monitor and control the combustion process in order to minimise pollutant emissions as well as to optimise energy losses. Flame temperature appears, among others, as a very important parameter to be monitored (Romero et al., 2005; Thakur et al., 2001; Lu et al., 2004; Liu and Jiang, 2001). Conventional temperature monitoring devices such as thermocouples are intrusive and disturb the measurement, and they must undergo the harsh furnace environment. Remote optical measurements are more suitable because they are non-intrusive. Ultraviolet, visible and infrared detectors have been used in flame monitoring systems (Romero et al., 2005). Infrared sensing appears to be very promising, because the hot gases in the flame, mainly carbon dioxide (CO₂) and water vapour (H₂O), exhibit important emission bands in the infrared region. A recent trend in flame thermometry is based on spectrometric measurements that discriminate the received energy as a function of the wavelength. An example is the application of the so-called emission–transmission method by using tunable infrared laser and optical fibre (Lu et al., 2004). This technique is an active technique, because it uses an infrared source in addition to the sensor system. These methods are very sensitive, but their high cost and complexity make them not very suitable for routine operations in industrial furnaces.

This paper presents some results within the framework of an author’s general proposal to use passive infrared spectroscopy to recover the temperature profile inside a hot gas cloud composed of CO₂ and water vapour, representative of a fossil fuel combustion. The experimental equipment to be used is a sensor (for instance, a commercial spectroradiometer) that measures directly the spectral distribution of radiated energy by the flame in the infrared spectral range. The selected spectral range is 3–5 μm, because carbon dioxide presents a strong emission band in this region (Briz et al., 2003). Moreover, this range is commonly implemented in infrared systems, because atmospheric absorption is not very important.¹

¹Note the magnitude at the X -axis: it is common for spectroscopists to use the wavenumber n instead of the wavelength l . Wavenumber is defined

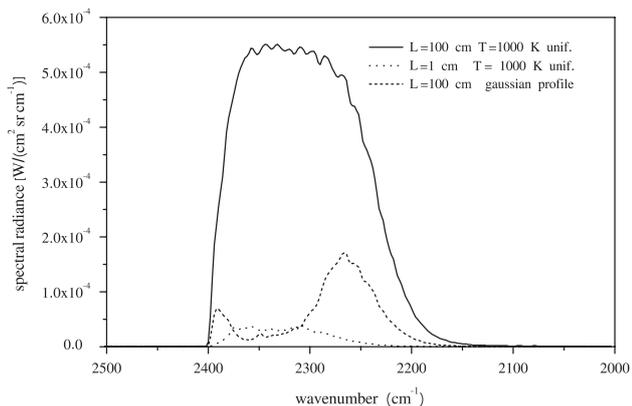


Fig. 1. CO₂ emission band calculated for: (a) a cell of width $L = 100$ cm and uniform temperature $T = 1000$ K; (b) a cell of width $L = 1$ cm and uniform temperature $T = 1000$ K; (c) a cell of width $L = 100$ cm and a temperature distribution (Gaussian profile).

The problem to obtain the temperature profile from such a spectrum is not straightforward. Energy emission at each wavenumber depends in a non-linear way on parameters like the spatial distribution of temperature and gas concentrations, or the gas cloud width. Moreover, each wavenumber emission depends in a different way on these parameters. To illustrate these difficulties, Fig. 1 shows a theoretical calculation of the CO₂ emission band for three different situations: a CO₂ cloud 100 cm wide at 1000 K, the same cloud but 1 cm wide and a CO₂ cloud 100 cm wide with a temperature exponentially decayed from 1000 to 400 K. It is very clear from the figure that the shape of the temperature profile and the width of the cloud are key parameters to understand the information involved in the energy emission spectrum.

3. Retrieval methodology

Due to the non-linearity of the inverse problem the use of MLP neural network as inverse model for the radiative transfer equation (1) has been proposed. With this approximation the physical properties in a parallel form can be obtained trying to make useful the non-linear correlations or dependencies among the channels and helping to solve the ambiguity due to the dependence of several constituents.

3.1. Multilayer perceptron

The MLP approximation constructs a unique model composed of three layers: input, hidden and output layer. Each layer consists of neurons or units which receive their input from units from a layer directly below and send their output to neurons in a layer directly above the unit. The

(footnote continued)

as the inverse of wavelength, and it is typically measured in cm^{-1} in the infrared spectral region.

input neurons are merely fan-out and any processing takes place in these units. The activation or output of the rest of the neurons in the network—hidden and output neurons—is a function f of the weighted inputs plus a bias value:

$$\text{out}_i = f\left(\sum_{j=1}^n w_{ij}x_j + b_i\right), \quad (2)$$

where x_j , w_{ij} and b_i are the inputs, weights and bias associated to the neuron i , respectively. The function f is called the activation function and the most often used are the sigmoidal functions, given by the following equations:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (3)$$

The output values, out_i , serve as input to the next layer to which the node is connected and the process is repeated until output values are obtained in the output layer. The activations of the units in the output layer are the outputs of the MLP.

The adjustable parameters—weights and bias—have to be determined to specify the network completely, which is known as the training process of the network. At this point, the learning rule comes into the operation using the samples about the problem. Most learning rules are formulated with a specific goal, e.g. move iteratively the parameter vector to a position which minimises or maximises some particular cost function. The parameters in MLP are determined using a supervised learning technique and they are generally adjusted to minimise the error between the output vector of the network y_k and the desired output vector y_k^{des} for the k th training sample:

$$\text{Error} = \frac{1}{N} \sum_{k=1}^N (y_k - y_k^{\text{des}}), \quad (4)$$

where N is the number of training samples.

As it was mentioned in the Introduction, neural networks have been used to retrieve the atmospheric temperature (Aires et al., 2001) although in combustion processes there are some differences as we do not have information about the opposite level temperature from where we take the measurement (the earth surface in atmospheric problem), or the temperatures are hotter and a priori know about the different levels.

3.2. High resolution spectral feature selection: PCA

It is known that high resolution leads a priori to a better accuracy in the retrieval problem. However, working with a huge dimensionality with mapping purposes conducts towards the so-called problem, the curse of dimensionality, as was mentioned above. For instance, spectral resolution used in this work is 0.075 cm^{-1} in the range of $2110\text{--}2410 \text{ cm}^{-1}$, which means a dimension of 4000.

Although a MLP could be trained with any input dimension, practically this dimensionality affects the results and the time processing either. The problem is to

find a transformation from the original high-dimensional space to a low-dimensional one, which will retain as much information as possible that is useful for our retrieval purposes. There are different ways to make this reduction of the dimensionality: feature extraction (linear or non-linear), transformation of data (PCA projections) and feature selection (Bishop, 1999). A PCA-based analysis has been adopted to make a selection of the channels which have the information about temperature without losing accuracy (even improving it). Thus it will make the method faster because of the significant input dimension reduction and will help to reach the best performance of MLP neural network. Input reduction also decreases the number of parameters in the regression model (i.e. weights in the neural network) and consequently decreases the number of degrees of freedom in the model.

In this work, we have used the feature selection based on PCA. This selection will choose those channels that contain most of the variance of the database which are also sensitive to the combustion parameters. In ad hoc studies the projections were used to make the reduction but because the results were not promising this alternative was discarded. The main idea of the statistical PCA is to reduce the dimensionality of data set in which there are a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This reduction is achieved by transforming to a new set of variables, the principal components, which are uncorrelated, and which are ordered so that the first few retain most of the variations present in all of the original variables (Jolliffe, 2002).

Let $C = \{e_1^M, \dots, e_n^M\}$ be a data set of n spectrum of dimension M variables. Let Σ be the covariance matrix of the data set C with dimension $M \times M$. Let V the $M \times M$ matrix with columns equal to eigenvectors of Σ and let L be the diagonal $M \times M$ matrix with the M associated eigenvalues (by definition $\Sigma \cdot V = V \cdot L$).

The selection of m specific channels from M variables, where $m \ll M$, allows to work with lower dimensionality. This m subset of variables contains virtually all the information available in M variables. The problem then is to find the value of m , and to decide the subset or subsets of m variables are the best. Here we want to find those variables which best represent the internal variation of C to find out which channels are significant (feature selection). In other cases the linear correlation between principal components (PCs) and channels are used to interpret the physical meaning (Huang and Antonelli, 2001), or to get a first retrieval approximation (Aires et al., 2002). To resolve the question about how many m variables we have to consider, we must check the number of PCs that account for most of the variations in a spectrum e_x^M of the data set C . This can also be interpreted as finding the effective dimensionality of e_x^M . If e_x^M can be successfully described by only m PCs, then it will often be true that M can be replaced by a subset m (or perhaps slightly more) variables, with a relative small loss of information (Jolliffe, 2002).

3.3. How to make a good selection of channels: pick peak algorithm

Regarding the choice of a subset of variables, a conservative procedure has been adopted in a sense of not trying to find the optimum number for the reduction but seeing how this reduction improves the results obtained over without using it. Most of the variations in data set (Section 4.1) can be described by the first PCs, and often will be true that this variation could be replaced by a subset of variables equal to those few PCs. The scree graph plotted in Fig. 2 gives an approximation about how many are they (Jolliffe, 2002). Looking at the figure, the “low elbow” shows that around the 5th PC the function reaches stability, so the size of the subset would be 5 or 6.

A second way to decide how many variables must be selected is to look at the accumulated variance (see Table 1) and determine what is the percentage of reconstruction we need. Then the number of variables will be the number of PCs needed to reach this percentage. In our case a pessimistic approach leads us to make a selection of 99.9% which is associated to 15 PCs and consequently to 15 variables.

In previous studies with medium spectra resolution (García-Cuesta et al., 2005) a reduction of a factor of 15 has been reached with successful results and rather more, the inputs reduction not only preserves the accuracy in retrieval but it could even improve it. However, in high resolution the results obtained with this number of variables and the method B4 (Jolliffe, 2002) were not as good as we got in previous works with medium resolution.

Afterwards, in order to introduce a priori knowledge of the problem to help the MLP to learn and improve results, a study of the different eigen-spectrums is done to see

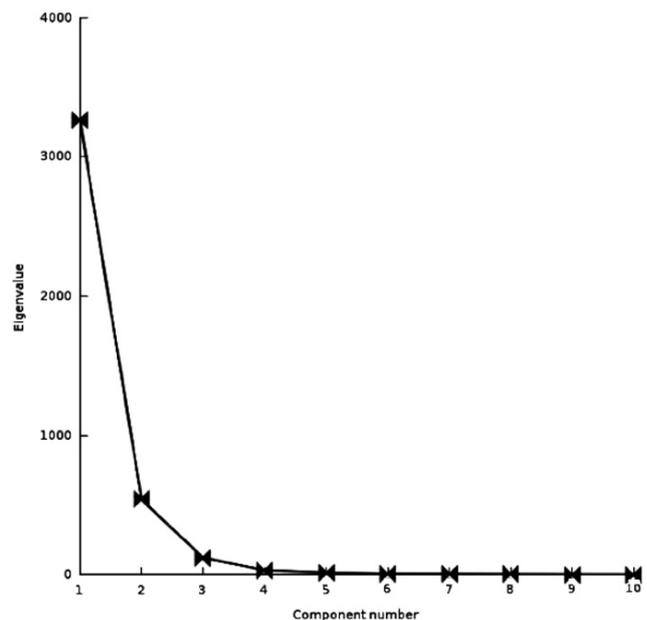


Fig. 2. Scree graph for the correlation matrix.

Table 1
Accumulated percentage of variance for spectrum data set generated

Number of PCA components	Accumulated variance (%)
1	95.15
2	98.20
3	99.04
4	99.42
5	99.58
6	99.70
7	99.77
8	99.81
9	99.83
10	99.85
11	99.87
12	99.88
13	99.89
14	99.90

where the most relevant information is. From the physical point of view, and closely related to the radiative transference equations solutions, appears the concept of weighting function (Rogers, 2005). Information measured in each channel of the spectrum is related to different spatial depths inside the flame. For a given depth, the weight functions show which channels are the most sensitive to the energy emitted from this depth. Our purpose is to use that physical information to figure out the most relevant channels. This idea is going to be exploited guiding the selection over all the possible variables M , obtaining a subset of selected variables $FS \in \mathcal{R}^m : 1 \leq m \leq M$, using a pick peak selector over the eigenspectrums $V = \{v_i : 1 \leq i \leq PC\}$ where PCs are the principal components. The pick peak selector will select channels searching for maximum absolute values in the eigenspectrums, but at the same time spreading the selection according with previous information coming from weighting function study. From this mixture of physical information and statistical information it is expected to improve the effectiveness of the selected channels.

The first six eigenspectrums are shown in Fig. 3. Each one gives specific information about different channels. The first eigenvector represents an average of the whole spectrum data set, so it is not casual that the most of the variations of PCs belong to this eigenvector. The shape of this first eigenvector is very similar to an average over all the spectrums of the data set. The following five eigenspectrums show the variations of the different channels over this first eigenvector. Physically each of these eigenspectrums do not contain only information about the temperature of one area of the gas, but also show information about the temperature and the spatial distribution over the different channels.

The following eigenspectrum shapes begin to lose the physical meaning and we cannot see a pattern related with any physical characteristics. The same idea could be inferred from the projections of each of our energy

spectrums of the database in this eigenspectrums. Observing the projections over these first PCs can be found out some patterns related with temperature and total depth, but after this it is very difficult to observe any pattern, it seems a randomness more than information. This is why the applicability of the peak selector will be these first five eigenspectrums.

In order to select the most important variables of this structure, we proposed here the use of the first five eigenspectrums $VS = \{V_i : 1 \leq i \leq 5\}$, and a pick peak selector method to get the subset of variables $FS \in \mathcal{R}^m$, where $1 \leq m \leq M$. This methodology agrees with the method called B4 in Jolliffe (2002), but instead of choosing only one variable per PC the pick peak selector is used. With this purpose, a window $win_thres \in \mathcal{R}^2$ is established, where first and second components are the threshold height and width, respectively. So, $win_thres = (win_thres_height, win_thres_width)$. Afterwards, every peak that has a cumulative value (since the last peak either positive or negative) greater than this win_thres_height will be chosen (see pick peak selector algorithm). This method is rough but it works quite well with sharp peaks, such as in this case, and will allow us to catch not only the most relevant information of each eigenspectrum but mainly second, third, ... and carry on most important variables either. To limit the number of peaks selected, the window thresholds also could be adjusted. **Pick peak selector algorithm.**

```

FS ← ∅
for each ( $v_i \in VS$ ) do
  for each ( $j : win\_thres\_width \leq j \leq M - win\_thres\_width$ )
    min_win ← MIN( $v_{i:w-j-win\_thres\_width \leq w \leq j+win\_thres\_width}$ )
    if (min_win < min)
      min = min_win
    end
    if ( $(v_{ij} - min) > win\_thres\_height$ )
      FS ← FS ∪ j
      min =  $v_{ij}$ 
    end
  end
end
end

```

The final selected channels FS are shown in Fig. 4 and are the ones we have used to make the test. The results are shown in Section 4.

4. Results for retrieval of temperature profiles

Using the methodology explained in Section 3, and the data set generated for typical flame combustions scenarios described in the next subsection 4.1, three different approaches have been adopted. The profiles of temperature are discretized to five cells, so the output layer will have six neurons, one for each temperature cell and another one for

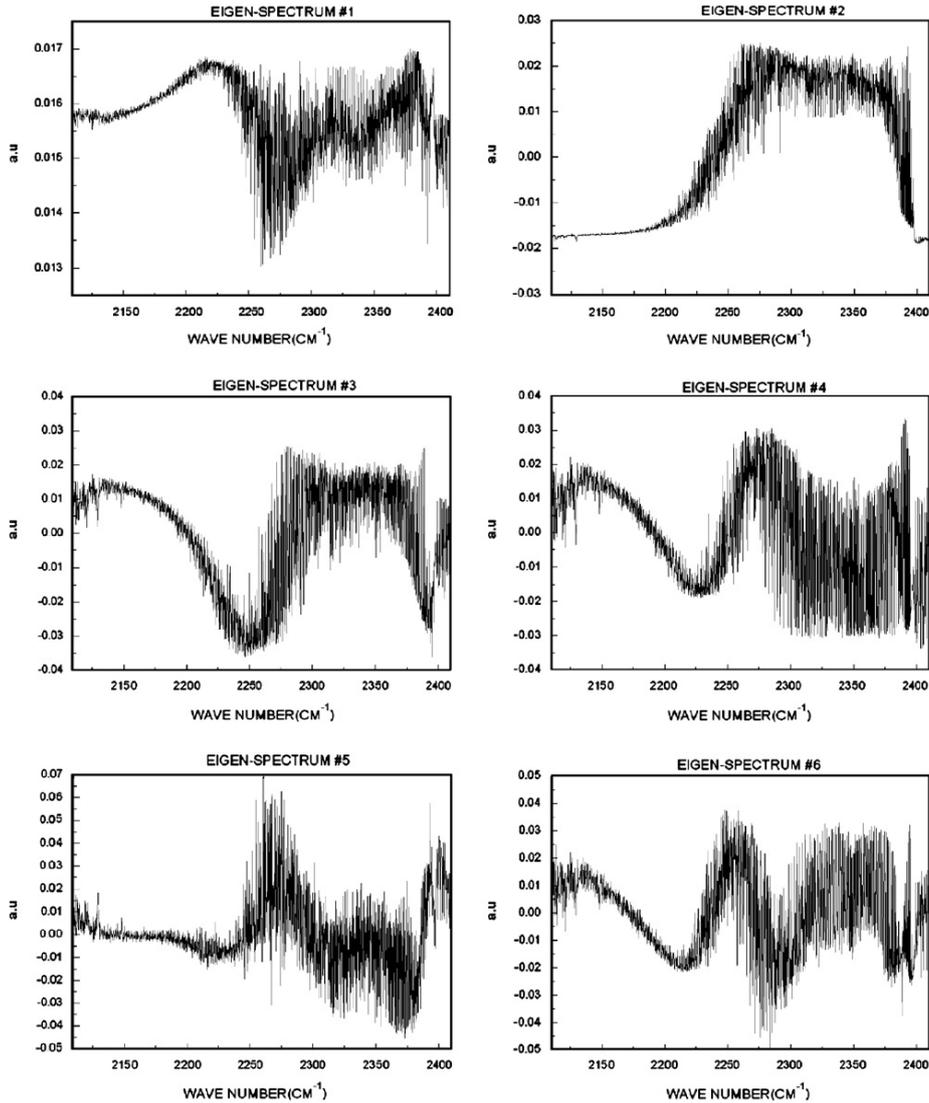


Fig. 3. First six eigenspectra of synthetic combustion simulations 2110–2410 cm^{-1} .

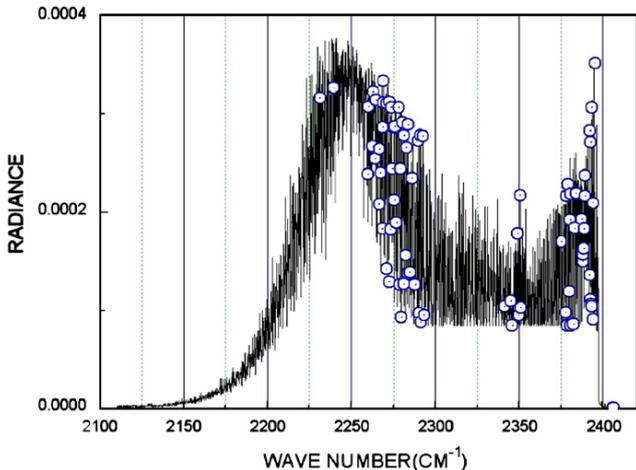


Fig. 4. Selected channels by the pick peak selector in the range 2110–2410 cm^{-1} .

the total length.² We must include the length because both parameters, temperature and optical depth or total length, have an influence on the composition of the spectrum in agreement with the Beer's law. Different architectures of MLP varying the number of hidden neurons have been trained until to reach the minimum value in validation error, not allowing overfitting. In these experiments, we have used a heuristic procedure to know what is the number of hidden neurons which best fits the generalisation errors.

4.1. Scenarios

The data set is composed of a large number of synthetic emission spectra generated with a computer code

²We are assuming in the discretization that all the cells have the same length so we do not include one per each cell.

developed at the University Carlos III called CASIMIR (Atmosphere’s Calculus and SIMulation in IR) (García-Cuesta, 2003) based on the well-known HITRAN/HITEMP spectral database (Rothman, 2003). This software has been developed under line-by-line concept and provides high spectral resolution taking into account the fine structure of vibrational–rotational spectra of individual molecules. The core of CASIMIR is based on the use of the well-known Beer’s law to calculate the spectral emissivity of the hot gases at each wavelength. Basic information is read from the HITEMP database. Afterwards, corrections to temperature, gas concentration and optical path length are performed. A line broadening mechanism is taken into account by using a Lorentzian broadening function. Finally, the computation of the radiative transfer equation is performed to obtain the spectrum of emitted energy by the flame.

The total number of cases simulated is 1040 covering many possible scenarios of a typical flame combustion. Data set generation has been performed under the following assumptions:

- Synthetic spectra will correspond to energy emission of hot gas cloud of width L . Temperature and gas concentrations present gradients inside the cloud.
- The spectral range selected for this data set is $2000\text{--}2500\text{ cm}^{-1}$. Most of the commercial infrared instruments have capabilities to measure in this range. In this spectral range, the CO_2 emission band is by far the most important emission feature, being the water emission nearly negligible. Due to this fact, only the emission associated to the CO_2 will be considered.
- For retrieval scheme, we have used a spatial discretization in a basic case with five cells of equal width ($L/5$). Each cell has an average value of temperature and gas concentration.
- The objective of this study is to focus on the dependence of the spectral energy distribution on temperature profile and length. For this reason, the concentration profiles for carbon dioxide and water vapour will be kept unchanged for the whole data set. Numerical values for these concentrations have been selected from typical combustion experiments.
- Four basic temperature profiles have been chosen to simulate different temperature gradient. The step between the temperature of two consecutive synthetic flames is $\Delta T = 50\text{ K}$, with a variation in the hottest cell among 540 and 1140 K. And for each of these variations of temperature, a variation of cell’s length is done. These variations have a step of $\Delta w = 0.02\text{ m}$ for each cell which means a total step variation $\Delta W = 0.1\text{ m}$, covering a range between 0.1 and 2 m. These profiles adjusted to a spatial discretization of five cells can be seen in Fig. 5. In these simulations the most unfavourable situation from energy detection point of view has been always chosen: the cell corresponding to the hottest part of the cloud is located opposite from the location of the

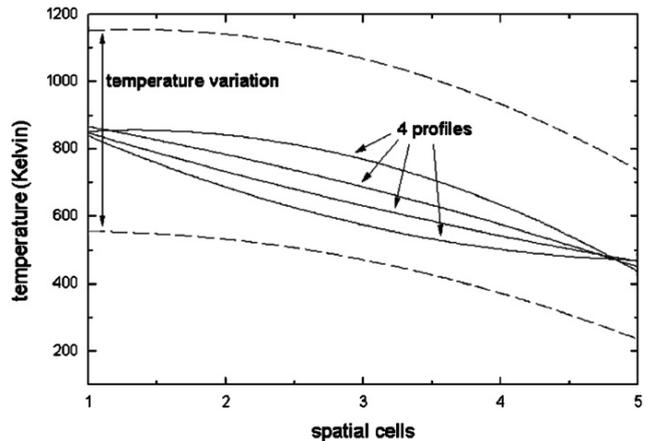


Fig. 5. Temperature profiles variations for retrieval scheme.

sensor system. All the value ranges for temperature and length have been chosen to be representative for hot gases clouds associated to fossil fuel combustion.

- Experimental noise of spectra has not been simulated in order to avoid the extraction of features associated with it when PCA is being used.

4.2. Initial approach: without feature selection

First, a MLP has been trained using the whole spectrum. In this case the experiments have been done with 4000 dimensions (high spectral resolution). In Table 2, the mean square error over the trained and test data for different architectures is shown. The table also included the mean error per profile, the mean error on the hottest cell and its standard deviation. In this case the MLP reach the minimum in a few cycles and then begins the overfitting.

As it was expected the results were not good and also the process was very high time consuming producing overfitting without reaching the results obtained with medium spectral resolution (García-Cuesta et al., 2005). The standard deviation for the best case (30 hidden neurons) in temperature is 13.22 K, which is also a bad result.

4.3. B4 Selection method

Our second approach was with a method called B4 in Jolliffe (1972, 1973). This method has been applied over the first 100 eigenspectrums taking the largest value of each one. With this criterion a selection of 59 channels (not a 100 because there are duplications) has been tested and the results obtained have improved the first one but it is not good enough to accept it. The results are shown in the Table 3.

Here, the best results also are obtained with 30 hidden neurons, with a mean error of 9.76 K over the profile and 8.31 K on the hottest cell with a standard deviation of 6.69 K, which although smaller than the previous one is still not acceptable.

Table 2

Errors for the hot gas temperature retrieval using a MLP with 4000 inputs and different architectures

Hidden neurons	MSE train	MSE test	Mean error per profile (K)	Mean error hottest cell (K)	Standard deviation
10	0.00418	0.02817	21.65	19.20	16.40
20	0.00178	0.03224	21.08	19.13	14.78
30	0.00099	0.02871	18.80	17.08	13.22
60	0.00989	0.03208	25.88	21.89	17.20

Table 3

Errors for the hot gas temperature retrieval using a MLP with B4 criteria

Hidden neurons	MSE train	MSE test	Mean error per profile (K)	Mean error hottest cell (K)	Standard deviation
10	0.00377	0.00436	16.42	11.79	11.70
20	0.00239	0.00255	11.49	8.84	6.63
30	0.00163	0.00228	9.76	8.31	6.69
40	0.00177	0.00236	10.27	9.43	7.67

Table 4

Errors for the hot gas temperature retrieval using a MLP with pick peak selector algorithm

Hidden neurons	MSE train	MSE test	Mean error per profile (K)	Mean error hottest cell (K)	Standard deviation
10	0.00058	0.00067	6.82	5.10	4.60
20	0.00021	0.00033	3.94	2.95	3.56
30	0.00024	0.00029	3.36	2.81	2.90
40	0.00022	0.00040	3.72	3.14	3.18

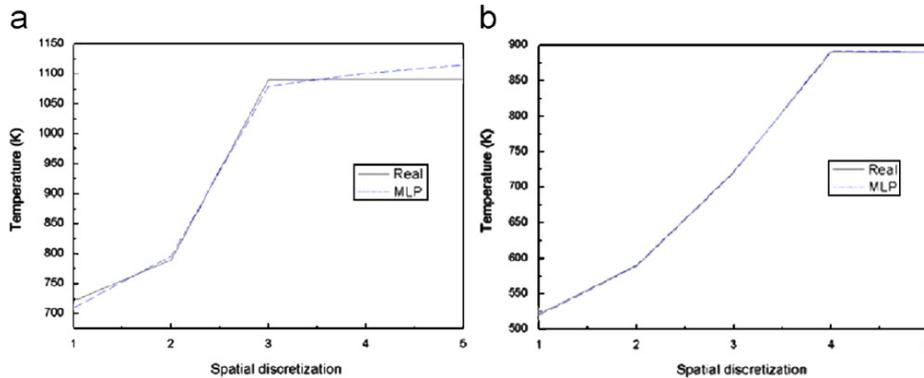


Fig. 6. The worst (a) and best (b) results obtained by the MLP with our method.

4.4. Refined approach: pick peak

The approach we suggest in Section 3.3 has obtained the best results. With the pick peak selector algorithm, 86 channels have been selected. In Table 4 we could see the results per different architectures. The mean errors per profile are 3.36 K and 2.81 K on the hottest cell with a relative error of 0.34%. Also the relative error of length of 2.1% for the best case is acceptable and these results improve quantitatively the ones obtained with 4000 variables and with B4 method.

There is also a significant difference in the standard deviation of these errors because versus 2.89 K deviation obtained with peak selector criteria, are the 6.69 K in the B4 criteria and 13.22 K with all the variables.

The representation of the results for the best and worst cases obtained with this last method could be seen in the Fig. 6.

5. Conclusions and perspectives

We have developed a method based on a reduction of dimensionality and on the MLP as inverse model for the retrieval of temperature profiles in combustion processes. This approach presents some advantages against the classical physical–statistical techniques as higher speed, autonomy of initial conditions and adaptability to other conditions. Physically, this remote sensing approach to control combustion processes has one main advantage, it is not intrusive and does not disturb the measurement. But

this is not the unique advantage. It allows to make measurements in difficult access areas and also undergo harsh environments.

As was expected the use of high spectral resolution measurements versus medium has lead to better results. We have improved the results over the hottest area in almost 3 K and the mean profile error has been reduced almost to a half. This has been possible due to the reduction done over the whole dimensionality of the problem. The reduction done is of factor 46 which helps in two different ways. Firstly, it allows the use of MLP as inverse model because with high dimensionality the time of calculus grows geometrically to the input dimension. And secondly it prevents the curse of dimensionality problem reducing the working space and thus enlarges the number of examples per dimension unit (the density), allowing better results in the retrieval of temperature as we have seen in Section 4.

We can conclude after these results that the mix between MLP and a guide high spectral feature selection is presented as a possible alternative on retrieval of temperature profile in hot gases combustion processes.

Until now we have been working with a low spatial discretization but in near future we expect to make retrievals with high spatial discretization which will suppose fortunately new problems and challenges. We think that in this new scenario the difference between using high spectral resolution and medium or low spectral resolution will be greater than in the results shown here, and high resolution measurements will take more importance.

Acknowledgements

This research is supported by the Spanish MEC projects–TRA2005-08892-C02-01 and OPLINK::UC3M, Ref: TIN2005-08818-C04-02.

References

- Aires, F., Chedin, A., Scott, N.A., Rossow, W.B., 2001. A regularized neural net approach for retrieval of atmospheric and surface temperatures with the iasi instrument. *Journal of Applied Meteorology* 41, 144–159.
- Aires, F., Rossow, W.B., Scott, N.A., Chedin, A., 2002. Remote sensing from the infrared atmospheric sounding interferometer instrument. *Journal of Geophysical Research* 107 (D22) ACH6–1–15.
- Bellman, R., 1961. *Adaptative Control Processes: A guide Tour*. Princeton University Press, NJ.
- Bishop, C.M., 1999. *Neural Networks for Pattern Recognition*. Oxford University Press.
- Briz, S., de Castro, A.J., Aranda, J.M., Meléndez, J., López, F., 2003. Reduction of false alarm rate in automatic forest fire infrared surveillance systems. *Remote Sensing of Environment* 86, 19–29.
- Eyre, J.R., 2004. Inversion methods for satellite sounding data. Lecture Notes NWP Course. European Centre for Medium-Range Weather Forecasts (ECMWF). ([http://www.ecmwf.int/newsevents/training/course_notes/DATA_ASSIMILATION/INVERSION_METHODS/Inversion_methods.html\(10/07/2006\)](http://www.ecmwf.int/newsevents/training/course_notes/DATA_ASSIMILATION/INVERSION_METHODS/Inversion_methods.html(10/07/2006))).
- García-Cuesta, E., 2003. CASIMIR: Cálculos Atmosféricos y Simulación de la Transmancia en el Infrarrojo. University Carlos III L/PFC 01781, Madrid (in Spanish).
- García-Cuesta, E., Galván, I.M., de Castro, A.J., 2005. Neural networks and spectral feature selection for retrieval of hot gases temperature profiles. In: *International Conference on Computational Intelligence for Modelling, Control and Automation, CIMCA*, vol. 2, 2005, pp. 81–86.
- Goody, R.M., Yung, Y.L., 1989. *Atmospheric Radiation. Theoretical basis*. Oxford University Press, New York (Chapter 2).
- Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. *Educational Psychology* 24, 417–441, 498–520.
- Huang, H.L., Antonelli, P., 2001. Application of principal component analysis to high-resolution infrared measurement compression an retrieval. *Journal of Climate and Applied Meteorology* 40, 365–388.
- Jolliffe, I.T., 1972. Discarding variables in principal component analysis i: artificial data. *Applied Statistics* 21.
- Jolliffe, I.T., 1973. Discarding variables in principal component analysis ii: real data. *Applied Statistics* 22.
- Jolliffe, I.T., 2002. *Principal Component Analysis*, second ed. Springer Series in Statistics. Springer, New York.
- Liu, L.H., Jiang, J., 2001. Inverse radiation problem for reconstruction of temperature profile in axisymmetric free flames. *Journal of Quantitative Spectroscopy & Radiative Transfer* 70, 207–215.
- Lu, G., Yan, Y., Colechin, M., 2004. A digital imaging based multi-funcional flame monitoring system. *IEEE Transactions on Instrumentation and Measurement* 53, 1152–1158.
- McCornick, N.J., 1992. Inverse radiative transfer problems: a review. *Nuclear Science and Engineering* 112 (3), 185–198.
- Pearson, K., 1901. On lines and planes of closet fit to systems of points in space. *Philosophical Magazine* 2 (6), 559–572.
- Rogers, C.D., 2005. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Review of Geophysics and Space Science*.
- Romero, C., Keyvan, S., Li, X., Rossow, R., 2005. Spectrometer-based combustion monitoring for flame stoichiometry and temperature control. *Applied Thermal Engineering* 25, 659–676.
- Rothman, L.S., 2003. The hitran molecular spectroscopic database: edition of 2000 including updates through 2001. *Journal of Quantitative Spectroscopy & Radiative Transfer*.
- Thakur, M., Vyas, A.L., Shakher, C., 2001. Measurement of temperature and temperature profile of an axisymmetric gaseous flames using lau phase interferometer with linear gratings. *Optics and Lasers in Engineering* 36, 373–380.