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 CO_2 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 (Jollife, 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 (CO_2) and water vapour (H_2O), 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 CO_2 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]

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