

RELEVANCIA DE LAS VARIABLES METEOROLÓGICAS EN EL DISEÑO DE UN MODELO DE PREDICCIÓN DE LOS NIVELES DE OZONO, EN TIEMPO REAL, BASADO EN EL USO DE REDES NEURONALES

Elena Agirre Basurko⁽¹⁾, Alvaro Anta Sanz⁽²⁾, Luis Javier R. Barrón⁽³⁾, Marivi Albizu Etxeberria⁽⁴⁾.

⁽¹⁾ Universidad del País Vasco, Escuela Universitaria de Ingeniería Técnica Industrial, La Casilla, 3 – 48012, Bilbao, elena.agirre@ehu.es

⁽²⁾ Egailan, Nieves Cano, 29 D – 01006, Vitoria-Gasteiz, alvaro.anta@lanbide.net

⁽³⁾ Universidad del País Vasco, Facultad de Farmacia, Paseo de la Universidad, 7 – 01006, Vitoria-Gasteiz, luisjavier.rbarron@ehu.es

⁽⁴⁾ Gobierno Vasco, Departamento de Medio Ambiente y Ordenación del Territorio, Gran Vía, 85-8ª planta – 48013, Bilbao, marivi-albizu@ej-gv.es

1. INTRODUCCION

La calidad del aire, el estudio de los principales contaminantes atmosféricos, el comportamiento de éstos, los niveles y focos de emisión, las variables que toman parte en la formación de estos contaminantes y aspectos similares han sido tema de investigación en las últimas décadas (Finlayson y Pitts, 1986). La vigilancia de los niveles registrados, la determinación de la evolución de los contaminantes atmosféricos así como la elaboración de modelos de predicción de los niveles de estos contaminantes son temas fundamentales en el diseño de estrategias de control y vigilancia de la contaminación atmosférica, que permitirán mejorar la calidad del aire.

Los contaminantes atmosféricos son compuestos que repercuten de forma nociva en el entorno (Williamson, 1973). Los contaminantes primarios como el monóxido de carbono, dióxido de azufre, partículas y compuestos orgánicos se emiten directamente desde el foco de emisión (tráfico rodado, actividad industrial,...) a la atmósfera. Por otro lado, los contaminantes secundarios se originan como consecuencia de las reacciones que sufren los contaminantes primarios en la atmósfera. El ozono troposférico es un contaminante secundario que se produce debido a la interacción de los óxidos de nitrógeno con los componentes orgánicos volátiles bajo la influencia de la radiación solar. En la actualidad el ozono troposférico (O₃) es uno de los contaminantes más relevantes de la atmósfera de la CAPV (Gobierno Vasco, 2001; Gobierno Vasco, 2004). Diversos modelos de predicción han sido

desarrollados con el propósito de obtener las predicciones de las concentraciones de los contaminantes atmosféricos. Estos modelos admiten la siguiente clasificación (Zanetti, 1990):

- (i) Los modelos causa/efecto son modelos basados en relaciones físico-químicas que relacionan las emisiones de los contaminantes, el transporte de éstos en el medio circundante y las reacciones que dichas emisiones sufren en la atmósfera. Un ejemplo de modelo causa/efecto es el modelo euleriano tridimensional UAM (Urban Airshed Model) (Scheffe y Morris, 1993).
- (ii) Los modelos estadísticos se basan en la utilización de técnicas estadísticas para establecer relaciones funcionales entre las variables de entrada y las variables de salida, y no requieren conocer previamente cuáles son exhaustivamente las relaciones matemáticas entre las variables que integran el modelo. Así, en el caso de la predicción de los contaminantes atmosféricos se han desarrollado modelos de predicción basados en el análisis de series temporales (Simpson et al., 1983)(Kuang-Jung, 1992) o la regresión lineal múltiple (Cassmassi, 1998; Cardelino et al., 2001; Agirre, 2003).
- (iii) Los modelos basados en el uso de redes neuronales artificiales (RNAs) son modelos matemático-computacionales de tratamiento de la información, cuyo origen parte de la simulación del cerebro humano. Estos modelos se han utilizado con fines de predicción, de

aproximación de funciones y de clasificación de patrones. La ventaja de los modelos basados en la utilización de redes neuronales artificiales reside en su elevada capacidad de generalización y de tratamiento de problemas no lineales. Además, no requieren del conocimiento riguroso de la distribución de la variable de estudio. En el área de la calidad del aire, durante los últimos años se han elaborado diversos modelos de predicción basados en el uso de las redes neuronales artificiales (Gardner et al., 1999; Perez et al., 2000; Elkamel et al., 2001; Agirre et al. 2006).

Estudios comparativos entre las diferentes técnicas de predicción de los niveles de contaminantes atmosféricos como el ozono han probado la eficiencia de los modelos basados en el uso de las redes neuronales (Comrie, 1997; Gardner et al., 2000; Agirre, 2003). Por ello, el objetivo fundamental de este equipo de investigación hasta la fecha ha sido la elaboración y evaluación de un modelo de predicción basado en el uso de las redes neuronales (concretamente, en el uso del perceptrón multicapa), en tiempo real, de los niveles de ozono a corto plazo en diferentes estaciones de la Red de Control y Vigilancia de la Calidad del Aire de la CAPV que gestiona el departamento de Medio Ambiente y Ordenación del Territorio del Gobierno Vasco.

En este trabajo se van a mostrar los resultados de las predicciones de ozono obtenidas en los estudios realizados en dos zonas de la Red de Control y Vigilancia de la Calidad del Aire de la CAPV. En primer lugar se presentarán los resultados de las predicciones de ozono obtenidas en las estaciones de Elorrieta, Deusto, Mazarredo y Txurdinaga de la ciudad de Bilbao y en segundo lugar se tratarán las predicciones de ozono obtenidas en las estaciones de Muskiz, Zierbena, Mundaka, Algorta y Pagoeta en la zona denominada Kostaldea. En ambos estudios se obtuvieron predicciones, en tiempo real, de los niveles de ozono hasta ocho horas vista. Finalmente, mediante el método de Garson se observará de forma cuantitativa la importancia relativa de cada variable de entrada en el modelo de predicción.

2. EL PERCEPTRON MULTICAPA

Las redes neuronales artificiales se han ido difundiendo a través de las distintas áreas de la ciencia y de la tecnología. Las redes neuronales artificiales son unas estructuras que originariamente se crearon para simular el sistema nervioso humano, donde el elemento fundamental es la neurona. En la figura 1 se muestra de forma general una neurona o perceptrón simple. Los valores x_j representan las entradas de la

red, los valores w_{kj} son los pesos sinápticos que conectan la entrada j -ésima con la neurona k -ésima, el valor b representa el umbral que puede ser considerado como otro peso, f representa la función de activación o función de transferencia que se aplicará al valor procesado en la neurona, y finalmente el valor y es la salida de la neurona.

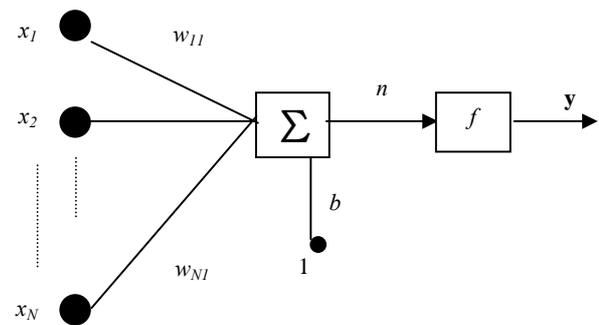


Fig. 1.- Neurona simple o perceptrón.

La expresión analítica es la siguiente:

$$y = f(n) = f\left(\sum_{j=1}^N w_{kj}x_j + b\right) \quad (1)$$

Las neuronas se conectan entre sí de forma que dan lugar a las redes neuronales artificiales. La estructura y el tipo de conexiones de la red, las características de las neuronas y la regla de aprendizaje son aspectos fundamentales en la determinación del tipo de red neuronal artificial a diseñar en un modelo. Así, como consecuencia de sus diversas aplicaciones (Rumelhart, 1986), en la actualidad el perceptrón multicapa (MLP, Multilayer Perceptron) es el tipo de red neuronal con mayor número de aplicaciones en problemas que involucran relaciones no lineales, como son las relaciones entre el ozono y las variables meteorológicas (Gardner y Dorling, 1998).

El perceptrón multicapa es una red neuronal artificial en la que la información que se introduce se propaga hacia adelante. Consta al menos de tres capas: la capa de entrada, la(s) capa(s) intermedia(s) o capa(s) oculta(s) y la capa de salida. La capa de entrada recibe la información o entradas a introducir en la red. Las entradas se multiplican por los pesos sinápticos y la suma total de estos productos se propaga hacia las neuronas de la capa intermedia. Al sumatorio correspondiente a cada neurona de la capa intermedia se le aplica una función de transferencia, y el resultado de dicha función es la entrada de la siguiente capa. Puede haber una o varias capas intermedias. Finalmente, como se puede observar en la figura 2 la

capa de salida produce la salida del perceptrón multicapa. Este perceptrón multicapa consta de 9 neuronas en la capa de entrada, 4 neuronas en la capa intermedia y 2 neuronas en la capa de salida.

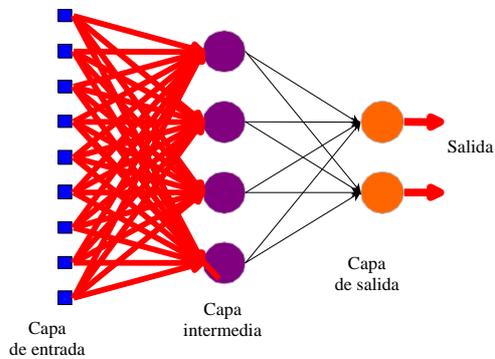


Fig. 2.- Perceptrón multicapa 9-4-2.

La función de transferencia determinará las relaciones entre las neuronas de las distintas capas. Las funciones de transferencia no lineales más habituales son la función logarítmico sigmoideal de la ecuación 2 y la función tangente sigmoideal de la ecuación 3.

$$\text{logsig}(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

$$\text{tansig}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3)$$

Las redes neuronales artificiales, y en concreto el perceptrón multicapa, poseen la capacidad de aprender de los patrones que les son introducidos y de los errores que cometen en el proceso de aprendizaje de patrones, de forma que finalmente son capaces de identificar un patrón no visto anteriormente. Por ello se dice que una red neuronal artificial adecuadamente entrenada ha de tener capacidad de generalización. El proceso de aprendizaje es equivalente al proceso de minimización del error E entre los valores objetivo $(t_1, t_2, \dots, t_L)^t$ y las salidas del perceptrón multicapa $(y_1, y_2, \dots, y_L)^t$ definido como sigue:

$$E = \frac{1}{L} \sum_{k=1}^L (t_k - y_k)^2 \quad (4)$$

donde L es el número de neuronas de la capa de salida. La salida del perceptrón multicapa se compara con la salida objetivo o salida deseada, y se retrocede hacia atrás en la red, produciéndose un ajuste en los pesos de la red para que la diferencia entre la salida de la red y la salida objetivo sea menor. Este método de aprendizaje es conocido como *backpropagation* o

método de aprendizaje con retropropagación. Una vez se ha alcanzado el mínimo de la función error E , finaliza el aprendizaje. Para evitar situaciones de sobreaprendizaje y pérdida de generalización se pueden aplicar técnicas como la parada temprana o *early stopping* (Sarle, 1995). Mediante esta técnica la base de datos total se subdivide en tres conjuntos: el conjunto de entrenamiento, el conjunto de validación y el conjunto de test. En el periodo de entrenamiento la red aprende con los datos del conjunto de entrenamiento, calcula los valores de los pesos sinápticos y los ajusta. Mediante los datos del conjunto de validación se determina el número de iteraciones del aprendizaje, de modo que el aprendizaje finaliza cuando se incrementa el error en el conjunto de validación. Finalmente, para medir la bondad del ajuste de los resultados obtenidos se utiliza una base de datos que el perceptrón multicapa desconoce, denominado conjunto de test.

3. BASE DE DATOS

En los dos estudios que se presentan se utilizaron los valores horarios de los contaminantes atmosféricos y variables meteorológicas registrados en la Red de Control y Vigilancia de la Calidad del Aire de la CAPV. En la tabla 1 se muestran las variables de contaminación, meteorología y tráfico utilizadas en los estudios de Bilbao y Kostaldea.

Notación	Variables
TE	Temperatura ($^{\circ}C$)
HU	Humedad relativa (%)
PR	Presión (kpa)
RA	Radiación ($\text{cal cm}^{-2} \text{h}^{-1}$)
GR	Gradiente térmico ($^{\circ}C$)
VV	Velocidad del viento (m/s)
DV	Dirección del viento (N°)
$O3$	Ozono ($\mu\text{g}/\text{m}^3$)
$NO2$	Dióxido de nitrógeno ($\mu\text{g}/\text{m}^3$)
NV	Nº de vehículos (vehic./10 min.)
OP	Porcentaje de ocupación (%)
KH	Velocidad ($\text{km h}^{-1} 100^{-1}$)
	$\text{sen}(2\pi h/24), \text{cos}(2\pi h/24),$
$SENCOS$	$\text{sen}(2\pi d/7), \text{cos}(2\pi d/7)$

Tabla 1.- Variables de los modelos de predicción.

3.1 Base de datos del estudio de Bilbao

En este primer estudio se recogieron los valores horarios de las variables de ozono, dióxido de nitrógeno, temperatura, presión, humedad relativa, gradiente térmico (diferencia entre la temperatura de la estación de Feria y la temperatura de la estación de Banderas), radiación solar, velocidad y dirección del viento. Asimismo se tomaron las variables registradas en la Red de Tráfico del Ayuntamiento de Bilbao, que miden en cada estación el número de vehículos cada diez minutos y el porcentaje de ocupación de la calzada. Para indicar una idea de la velocidad del tráfico, junto a estas dos variables se introdujo una nueva variable como cociente de las dos variables de tráfico indicadas, $KH = NV/OP$. El estudio se realizó en Bilbao en las estaciones de Elorrieta, Deusto, Mazarredo y Txurdinaga. El modelo fue construido mediante la base de datos correspondiente al año 1993 y mediante los datos del año 1994 se midió la bondad del modelo diseñado para obtener predicciones de ozono, en tiempo real, en cada estación hasta ocho horas vista. Debido al comportamiento cíclico del ozono, a estas variables de contaminación atmosférica, meteorología y tráfico se les añadieron las variables $sen(2\pi h/24)$, $cos(2\pi h/24)$, $sen(2\pi d/7)$ y $cos(2\pi d/7)$ a utilizar como variables de entrada del modelo, donde $h = 1, \dots, 24$ es la hora del día y $d = 1, 2, \dots, 7$ es el día de la semana.

3.2 Base de datos del estudio de Kostaldea

En este segundo trabajo se obtuvieron las predicciones de ozono hasta ocho horas vista en las estaciones de Muskiz, Zierbena, Mundaka, Algorta y Pagoeta en la zona denominada Kostaldea, utilizando como variables de entrada del modelo los valores horarios de las variables de ozono, dióxido de nitrógeno, temperatura, humedad relativa, presión, radiación solar, velocidad y dirección del viento. Para la construcción del modelo se utilizaron los datos correspondientes al periodo 2001-2003. El modelo se validó sobre los datos registrados en el año 2004. Al igual que en el estudio anterior se sumaron las variables $sen(2\pi h/24)$, $cos(2\pi h/24)$, $sen(2\pi d/7)$ y $cos(2\pi d/7)$ a utilizar como variables de entrada del modelo, siendo $h = 1, 2, \dots, 24$ la hora del día y $d = 1, 2, \dots, 7$ el día de la semana.

4. METODOLOGIA

4.1 Construcción del modelo

En ambos estudios se diseñaron modelos basados en el uso del perceptrón multicapa con el fin de obtener predicciones, en tiempo real, de los niveles de ozono

hasta ocho horas vista en cada una de las estaciones. Un perceptrón multicapa con una única capa intermedia es un aproximador universal de funciones (Hornik et al., 1989). Por ello, los modelos diseñados contienen una sola capa intermedia en la que mediante un procedimiento de prueba-error se determinará el número de neuronas de ésta. A su vez, la capa de salida contiene una neurona, que es la predicción de ozono o variable de salida del modelo, $O3(t+k)$ para el instante $t+k$, $k = 1, \dots, 8$. El procedimiento de prueba-error seguido se rige por la siguiente regla: "el ratio entre el número de patrones y el número de parámetros de la red a ser al menos igual a 30" (Amari et al., 1997). Para determinar el número de neuronas de la capa de entrada se tomaron los valores históricos de las variables de contaminación, meteorología y variables sinusoidales en ambos estudios (más los valores históricos de las variables de tráfico en el estudio de Bilbao) y, en general, mediante el método de regresión por pasos sucesivos con filtro de tolerancia se fijaron las entradas de cada modelo de predicción para cada instante $t+k$ ($k = 1, 2, \dots, 8$) en cada una de las estaciones. De este modo se redujo el número de variables de entrada notablemente. Así, por ejemplo en el caso del estudio de Kostaldea el número inicial de variables de entrada era de 204 y se redujo a un número de variables de entrada comprendido entre 4 y 9. En la determinación de las variables de entrada en el estudio realizado en Bilbao se utilizó la base de datos del año 1993, mientras que en el estudio de Kostaldea se utilizó la base de datos del periodo 2001-2003. Mientras en el estudio de Bilbao se utilizaron como variables de entrada las cinco variables determinadas mediante el método de regresión por pasos sucesivos con filtro de tolerancia más las cinco componentes sinusoidales dependientes del momento en el que se realice la predicción, en el estudio de Kostaldea se utilizaron menos variables de entrada, habiéndose determinado éstas mediante la misma metodología aplicada también sobre las variables sinusoidales.

La tabla 2 muestra las variables de contaminación, meteorología y tráfico que actuaron como variables de entrada en el modelo de predicción de ozono hasta ocho horas vista en la estación de Mazarredo.

Salidas	Entradas en Mazarredo				
$O3(t+1)$	$O3(t)$	$RA(t)$	$RA(t-3)$	$NO2(t-1)$	$O3(t-15)$
$O3(t+2)$	$TE(t-10)$	$O3(t-8)$	$PR(t-6)$	$NO2(t-8)$	$O3(t)$
$O3(t+3)$	$O3(t)$	$TE(t-12)$	$NO2(t-3)$	$NO2(t-7)$	$O3(t-15)$
$O3(t+4)$	$RA(t)$	$O3(t)$	$RA(t-3)$	$PR(t-2)$	$RA(t-12)$
$O3(t+5)$	$O3(t)$	$O3(t-15)$	$RA(t)$	$RA(t-3)$	$KH(t)$
$O3(t+6)$	$O3(t-15)$	$RA(t)$	$O3(t)$	$NV(t-6)$	$KH(t)$
$O3(t+7)$	$O3(t-15)$	$O3(t)$	$NV(t-6)$	$KH(t)$	$RA(t-5)$
$O3(t+8)$	$O3(t-15)$	$NV(t-6)$	$O3(t)$	$RA(t-15)$	$RA(t-11)$

Tabla 2. Variables de entrada en el modelo de predicción de ozono hasta ocho horas vista en Mazarredo.

Según se observa, entre las variables utilizadas como entradas de los modelos diseñados para obtener las predicciones $O3(t+k)$ ($k = 1, \dots, 8$) en la estación de Mazarredo destacan los valores de la variable de ozono en el instante en el que se realiza la predicción, $O3(t)$, y el valor pasado 15 horas atrás, $O3(t-15)$, los valores históricos de la variable radiación, $RA(t-z)$, $z = 0, 3, 5, 11, 12, 15$ y los valores pasados de las variables de temperatura, dióxido de nitrógeno, número de vehículos y velocidad del tráfico. Además de estas cinco variables, las cuatro variables sinusoidales actúan como variables de entrada de los modelos de predicción.

Análogamente, la tabla 3 muestra las variables de contaminación, meteorología y variables sinusoidales que actuaron como variables de entrada en el modelo de predicción de ozono hasta ocho horas vista en la estación de Pagoeta.

Salidas	Entradas en Pagoeta				
$O3(t+1)$	$O3(t)$	$NO2(t-1)$	$\cos(2\pi h/24)$	$DV(t)$	$DV(t-21)$
$O3(t+2)$	$O3(t)$	$NO2(t-1)$	$\cos(2\pi h/24)$	$DV(t-21)$	$HU(t-21)$
$O3(t+3)$	$O3(t)$	$\cos(2\pi h/24)$	$NO2(t-1)$	$RA(t-19)$	$DV(t-21)$
$O3(t+4)$	$O3(t)$	$\cos(2\pi h/24)$	$NO2(t-1)$	$RA(t-19)$	$DV(t-20)$
$O3(t+5)$	$O3(t)$	$\cos(2\pi h/24)$	$NO2(t-1)$	$RA(t-19)$	$DV(t-19)$
$O3(t+6)$	$O3(t)$	$\cos(2\pi h/24)$	$RA(t-19)$	$NO2(t-1)$	$DV(t-18)$
$O3(t+7)$	$O3(t)$	$RA(t-21)$	$NO2(t-1)$	$RA(t-11)$	$DV(t-17)$
$O3(t+8)$	$O3(t)$	$RA(t-20)$	$NO2(t-1)$	$RA(t-10)$	$DV(t-6)$

Tabla 3. Variables de entrada en el modelo de predicción de ozono hasta ocho horas vista en Pagoeta.

En este caso hay menos variación entre las variables de entrada de los modelos de predicción. Así, a la hora de explicar los ocho modelos de las ocho predicciones de ozono $O3(t+k)$ ($k = 1, \dots, 8$) en la estación de Pagoeta destacan la variable de ozono en el instante en el que se realiza la predicción, $O3(t)$, la variable dióxido de nitrógeno una hora antes de que se realice la predicción, $NO2(t-1)$, la variable $\cos(2\pi h/24)$, $h = 1, 2, \dots, 24$, los valores históricos de la variable radiación, $RA(t-z)$, con $z = 10, 11, 19, 20, 21$, y los valores pasados de la variable dirección del viento $DV(t-z)$, siendo $z = 0, 6, 17, 18, 19, 20, 21$. Aunque en algunos casos la metodología utilizada indicara que un valor histórico de ozono fuera variable de entrada, las predicciones de ozono mostraron un ajuste más adecuado cuando el valor histórico la variable de ozono que actuaba como variable de entrada era la variable $O3(t)$.

En ambos estudios las funciones de transferencia que se utilizaron fueron la función tangente sigmoideal y la función lineal. La función tangente sigmoideal estableció las conexiones entre la capa de entrada y la capa intermedia, y la función lineal conectó las neuronas capa de la capa intermedia con la neurona de la capa de salida.

Finalmente, se aplicó el algoritmo del gradiente conjugado escalado (SCG, *scaled conjugate gradient*), derivado del algoritmo de retropropagación y con mayor velocidad de convergencia que otros algoritmos de aprendizaje (Moller, 1993). Asimismo, para evitar problemas de sobreentrenamiento se aplicó el método de parada temprana. En el estudio desarrollado en las estaciones de Bilbao, el 85% de los datos de la primera parte del año 1993 formaron el conjunto de entrenamiento, el 15% de datos restantes del año 1993 formó el conjunto de validación y el modelo se probó sobre los datos del año 1994 o conjunto de test. En el estudio elaborado en Kostaldea el conjunto de test lo formaron los datos correspondientes al año 2004, y en función del número de casos válidos de cada estación, en general, el conjunto de entrenamiento lo formaron los datos correspondientes al periodo 2001-2002 y el conjunto de validación lo formaron los datos del año 2003.

4.2 Validación del modelo

Tras la construcción del modelo, mediante el cálculo de los estadísticos del Kit de Validación del Modelo se estudió cuantitativamente la eficacia de éste (Hanna et al., 1991; European Commission, 1994). La notación que se utiliza en los estadísticos del kit es la siguiente: C_p es la predicción, C_o el valor real, \bar{C}_p y \bar{C}_o son las medias aritméticas de C_p y C_o y SC_p y SC_o son las desviaciones típicas de estos valores.

Los estadísticos del kit de validación del modelo son los siguientes:

1. Coeficiente de correlación, R . Mide la bondad de ajuste del modelo, cuantificando de forma global la relación entre la predicción y el valor observado.

$$R = \frac{(\overline{C_o - C_p}) - (\overline{C_p} - \overline{C_o})}{(SC_p)(SC_o)} \quad (5)$$

2. Error cuadrático medio normalizado (NMSE, *Normalized Mean Square Error*). Es una variedad del error cuadrático medio, normalizado por el producto de la media de la predicción y la media del valor observado.

$$NMSE = \frac{(\overline{C_o - C_p})^2}{(\overline{C_o})(\overline{C_p})} \quad (6)$$

3. Factor de dos (FA2, *factor of two*). Muestra en qué proporción están los valores del ratio (valor predicho/valor real) en el intervalo [0,5, 2]. Es decir,

explica qué proporción de los datos verifica la siguiente expresión:

$$0,5 \leq C_o/C_p \leq 2 \quad (7)$$

4. Sesgo fraccional (*FB, Fractional Bias*). Esta medida es válida para comparar de forma global la media de la predicción y la media del valor real general.

$$FB = 2 \frac{\overline{C_o} - \overline{C_p}}{\overline{C_o} + \overline{C_p}} \quad (8)$$

5. Varianza fraccional (*FV, Fractional Variance*). Esta medida normalizada permite comparar globalmente la desviación típica del modelo y la desviación típica de los valores reales.

$$FV = 2 \frac{SC_o - SC_p}{SC_o + SC_p} \quad (9)$$

En ambos estudios el cálculo de los valores de los estadísticos del kit sobre el conjunto de test determinó cuantitativamente la bondad del ajuste de cada modelo.

5. RESULTADOS

Tras la construcción de los modelos de predicción de ozono en cada estación, en cada estudio se determinó la bondad del ajuste de cada modelo. La tabla 4 muestra como ejemplo los resultados obtenidos en el estudio de Kostaldea, donde se indican los valores de los estadísticos del kit de validación del modelo para cada una de las predicciones de ozono en la estación de Pagoeta. Dichos estadísticos se calcularon sobre la base de datos del conjunto de test y se compararon con los valores correspondientes a la observación o valor real de ozono en el instante $t+k$ ($k = 1, 2, \dots, 8$) y con los valores correspondientes a la persistencia (según la cual el valor de ozono en el instante $t+k$ es el mismo que el valor de ozono en el instante t).

Tal como indican los valores de los estadísticos de la tabla 4, las predicciones de ozono obtenidas mediante el modelo basado en la utilización del perceptrón multicapa (MLP) se ajustan al valor real mejor que la persistencia. Además, los valores de los estadísticos del kit demuestran la calidad de los resultados obtenidos mediante el modelo en cuestión.

Pagoeta	NMSE	R	FA2	FB	FV
O3(t+1)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0007	0,9969	0,9978	-0,0023	0,0468
Persistencia	0,0071	0,9568	0,9962	-0,0001	0,0001
O3(t+2)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0032	0,9840	0,9964	-0,0007	0,1061
Persistencia	0,0164	0,9000	0,9915	-0,0001	0,0001
O3(t+3)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0046	0,9792	0,9932	-0,0036	0,1556
Persistencia	0,0250	0,8474	0,9880	-0,0001	0,0001
O3(t+4)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0077	0,9690	0,9926	-0,0024	0,2376
Persistencia	0,0330	0,7992	0,9837	-0,0001	0,0001
O3(t+5)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0090	0,9611	0,9905	-0,0003	0,2487
Persistencia	0,0400	0,7562	0,9781	-0,0001	0,0001
O3(t+6)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0104	0,9599	0,9912	0,0049	0,2953
Persistencia	0,0459	0,7201	0,9747	-0,0001	0,0001
O3(t+7)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0117	0,9543	0,9895	0,0006	0,3194
Persistencia	0,0508	0,6905	0,9728	-0,0001	0,0001
O3(t+8)					
Observación	0,0000	1,0000	1,0000	0,0000	0,0000
MLP	0,0138	0,9360	0,9901	0,0030	0,3166
Persistencia	0,0550	0,6652	0,9695	-0,0001	0,0001

Tabla 4.- Estadísticos del kit de validación del modelo sobre el conjunto de test en Pagoeta.

El valor del error cuadrático medio normalizado está comprendido entre 0,0007 y 0,0138; los valores del coeficiente de correlación son positivos y varían en el intervalo [0,9360, 0,9969], siendo el valor inferior el correspondiente a la predicción de ozono a ocho horas vista y el valor superior el correspondiente a la predicción de ozono a una hora vista; análogamente, los valores del factor de dos son elevados, e indican en el peor de los casos que el 99,01% de los valores de la predicción de ozono a ocho horas vista están comprendidos en el intervalo [0,5, 2], y en el mejor de los casos que el 99,78% de los valores de la predicción de ozono a una hora vista están en el intervalo [0,5, 2]. Estos tres valores de los estadísticos ya indican un ajuste de calidad de las predicciones de ozono hasta ocho horas vista. En cuanto a los valores del sesgo fraccional y la varianza fraccional se puede observar que son relativamente pequeños, y que se incrementan a partir de la predicción a una hora vista, alcanzándose los valores mayores para la predicción de ozono a ocho horas vista.

6. RELEVANCIA DE LAS VARIABLES DE ENTRADA

En el caso de modelos elaborados mediante la utilización de redes neuronales artificiales, el método de Garson (Garson, 1991) establece una forma de medición de la importancia relativa de las variables de entrada en el modelo diseñado. Así, por ejemplo, en la tabla 5 se presenta la importancia relativa de cada una de las variables de entrada en el diseño del modelo para predecir los valores de la variable $O3(t+k)$ en la estación de Mazarredo, con $k=1, 4, 6$.

Predicción	Entradas	Importancia relativa (%)
$O3(t+1)$	$O3(t)$	27,05
	$NO2(t-1)$	11,90
	$sen(2\pi h/24)$	11,37
	$sen(2\pi d/7)$	10,99
	$O3(t-15)$	10,47
	$RA(t-3)$	10,26
	$RA(t)$	9,07
	$cos(2\pi h/24)$	8,34
$O3(t+4)$	$cos(2\pi d/7)$	0,55
	$RA(t-3)$	16,01
	$RA(t-12)$	13,38
	$O3(t)$	12,12
	$cos(2\pi h/24)$	11,77
	$RA(t)$	11,73
	$PR(t-2)$	11,30
	$sen(2\pi h/24)$	8,66
$O3(t+6)$	$cos(2\pi d/7)$	7,72
	$sen(2\pi d/7)$	7,30
	$O3(t)$	16,33
	$O3(t-15)$	14,29
	$KH(t)$	12,24
	$RA(t)$	11,45
	$cos(2\pi h/24)$	10,97
	$sin(2\pi h/24)$	10,70
$NV(t-6)$	10,03	
$sin(2\pi d/7)$	7,43	
$cos(2\pi d/7)$	6,57	

Tabla 5.- Importancia relativa de las variables de entrada en las predicciones de ozono a 1, 4 y 6 horas vista en Mazarredo.

Según indican los valores de dicha tabla la variable $O3(t)$ o concentración de ozono en el momento de la predicción y la variable $O3(t-15)$ o variable que mide la concentración de ozono 15 horas antes del momento en el que se calcula la predicción juegan un papel a destacar en las predicciones a una y seis horas vista. La suma de la importancias relativas de las variables $O3(t)$ y $O3(t-15)$ es del 37,52% en la predicción a una hora vista y supera el 30% en la predicción de ozono a

seis horas vista. Respecto de las variables meteorológicas, en la predicción a una hora vista la importancia relativa suma el 19,33%, en la predicción a cuatro horas vista supone el 52,42% del total, mientras en la predicción de ozono a seis horas vista la importancia relativa de la radiación solar en el instante en el que se efectúa la predicción es del 11,45%. La importancia restante se atribuye a las variables de tráfico y a las variables sinusoidales, siendo la suma de las importancias relativas de estas últimas como media del 30%.

Resultados similares se obtuvieron en el resto de las estaciones del estudio.

7. CONCLUSIONES

En ambos trabajos se destaca el uso del perceptrón multicapa como base del diseño de modelos de predicción de ozono hasta ocho horas vista. Un estudio conjunto de los estadísticos del kit de validación del modelo muestra que las predicciones de ozono obtenidas en el estudio de Kostaldea son más eficientes aún que las predicciones de ozono obtenidas en el estudio de las estaciones de Bilbao. Este hecho puede estar directamente influenciado por el número de parámetros utilizado en el entrenamiento del perceptrón multicapa, que en el caso del estudio de Kostaldea está formado por un periodo de cuatro años (2001-2004) mientras que en el estudio de Bilbao la base de datos está formada por los datos del bienio 1993-1994. Asimismo, la calidad del ajuste de las predicciones de ozono obtenidas es independiente de la utilización de variables de tráfico como variables de entrada del modelo, ya que el modelo de predicción elaborado en las estaciones de Kostaldea no dispone de variables de tráfico donde las variables de meteorología como variables de entrada desempeñan un papel a destacar. Las variables meteorológicas como la radiación solar o la dirección del viento, tanto en las predicciones de ozono del estudio realizado en Bilbao como en el estudio realizado en Kostaldea son variables determinantes de la eficacia del modelo de predicción. Habría que subrayar también el papel de la variable de ozono en el instante en el que se realiza la predicción como variable de entrada. Adicionalmente, las variables de entrada $sen(2\pi h/24)$, $cos(2\pi h/24)$, $sen(2\pi d/7)$ y $cos(2\pi d/7)$, $h = 1, 2, \dots, 24$, $d = 1, 2, \dots, 7$ muestran el carácter cíclico del ozono. El método de Garson permite medir de forma cuantitativa la participación de cada una de estas variables de entrada en el modelo de predicción diseñado.

La ventaja principal del desarrollo de modelos de predicción basados en el uso del perceptrón multicapa es que estos modelos no requieren información

exhaustiva de los contaminantes atmosféricos, las reacciones que sufren éstos o las variables que toman parte en su formación. Estos modelos han probado ser de gran eficacia en problemas no lineales, y además los modelos basados en la utilización de redes neuronales artificiales, en general, proporcionan mejores resultados. Consecuentemente, la técnica utilizada para elaborar modelos de predicción de ozono basados en la utilización del perceptrón multicapa, es una técnica de gran utilidad en el área de la contaminación atmosférica.

Agradecimientos

Los autores de este trabajo quisiéramos agradecer al departamento de Medio Ambiente y Ordenación del Territorio del Gobierno Vasco y al Ayuntamiento de Bilbao por habernos proporcionado los datos necesarios para la consecución de los estudios que se presentan en este artículo.

Referencias

Agirre, E., 2003. *Elaboration of a statistical prognostic model for short-term real-time prediction of O₃ and NO₂ levels in the city of Bilbao*. Ph. D. Universidad del País Vasco, Bilbao.

Agirre, E., Ibarra, G., Madariaga, I., 2006. *Regresión and multilayer perceptron based models to forecast hourly O₃ and NO₂ levels in the Bilbao area*. Environmental Modelling and Software 21, 430-446.

Amari, S.I., Murata, N., Müller, K.R., Finke, M., Yang, H., 1997. *Asymptotic statistical theory of overtraining and cross-validation*. IEEE Transactions on Neural Networks 8, 985-996.

Cardelino, C., Chang, M., St. John, J., Murphey, B., Cordle, J., Ballagas, R., Patterson, K, Stogner, J., Zimmer-Dauphinee, S., 2001. *Ozone predictions in Atlanta, Georgia: Analysis of the 1999 Ozone Season*. Journal of the Air and Waste Management Association 51, 1227-1236.

Cassmassi J.C., 1998. *Objective ozone forecasting in the South Coast Air Basin: Updating the objective prediction models for the late 1990s and southern California ozone study (SCOS97-NARSTO) application*. Proceedings of the 12th Conference on Numerical Weather Prediction, Phoenix, AZ. American Meteorology Society, Boston, MA. Pp 54-58.

Comrie A.C., 1997. *Comparing neural networks and regression model for ozone forecasting*. Journal of the Air & Waste Management Association 47, 653-663.

Elkamel, A., Abdul-Wahab, S., Bouhamra, W., Alper, E., 2001. *Measurement and prediction of ozone levels around a heavily industrialized area: a neural network approach*. Advances in Environmental Research 5, 47-59.

European Commission, 1994. *The Evaluation of Models of Heavy Gas Dispersion*. Model Evaluation

Finlayson-Pitts, B.J., Pitts, J.N., 1986. *Atmospheric chemistry: fundamentals and experimental techniques*. Wiley, New York.

Gardner, M.W., Dorling, S.R., 1998 *Artificial neural networks (the multilayer perceptron) – a review of applications in the atmospheric sciences*. Atmospheric Environment 32 (14/15), 2627-2636.

Gardner, M.W., Dorling, S.R., 1999. *Neural network modelling and prediction of hourly NO_x and NO₂ concentrations in urban air in London*. Atmospheric Environment 33, 2627-2636.

Gardner, M.W., Dorling, S.R., 2000. *Statistical surface ozone models: an improved methodology to account for non-linear behaviour*. Atmospheric Environment 34, 21-34.

Garson G.D., 1991. *Interpreting neural-network connection weights*. AI Expert 6(7), 47-51.

Gobierno Vasco, 2001. *Calidad del aire en la Comunidad Autónoma Vasca en el periodo 1996-2000*. Servicio de Publicaciones del Gobierno Vasco, Vitoria-Gasteiz.

Gobierno Vasco, 2004. *Aire-Ruido*. Servicio de Publicaciones del Gobierno Vasco, Vitoria-Gasteiz.

Hanna, S.R., Strimaitis, D.G., Chang, J.C., 1991. *Hazard response modeling uncertainty (a quantitative method). User's guide for software for evaluating hazardous gas dispersion models*. American Petroleum Institute, Washington.

Hornik K., Stinchcombe M., White H. 1989. *Multilayer feedforward networks are universal approximators*. Neural Networks 2, 359-366.

Kuang-Jung, H., 1992. *Time series analysis of the interdependences among air pollutant*. Atmospheric Environment 26B, 491-503.

Moller M.F., 1993. *A scaled conjugate gradient algorithm for fast supervised learning*. Neural Networks 6, 525-533.

Perez, P., Trier, A., Reyes, J., 2000. *Prediction of PM_{2.5} concentrations several hours in advance using neural networks in Santiago, Chile*. Atmospheric Environment 34, 1189-1196.

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. *Learning internal representations by error propagation*. MIT Press, Cambridge.

Sarle W.S., 1995. *Stopped training and other remedies for overfitting*. Proceedings on the 27th Symposium on The Interface.

Scheffe, R.D., Morris, R.E., 1993. *A review of the development and application of the urban airshed model*. Atmospheric Environment 27B, 23-39.

Simpson, R.W., Layton, A.P., 1983. *Forecasting peak ozone levels*. Atmospheric Environment 17, 1649-1654.

Zanetti, P., 1990. *Air pollution modelling: Theories, Computational Methods and Available Software*. Computational Mechanics Publications, Southampton, Boston.