

# Modelo basado en redes neuronales para predecir la estructura temporal y espectral de paisajes sonoros urbanos



PACS: 43.50.Rq

## Resumen

Los paisajes sonoros urbanos están dominados por el ruido generado por el tráfico rodado, que actúa como una fuente ruidosa que enmascara el resto de sonidos presentes, lo cual crea diversos problemas que van más allá de la molestia y desagrado en la población. Por otro lado, la percepción que la población tiene de los paisajes sonoros no se explica solamente por el nivel de presión sonora a que está expuesta, sino que entran en juego las características temporales y espectrales del paisaje sonoro, así como otras variables no acústicas como pueden ser, la actitud hacia la fuente, sensibilidad y habituación al ruido, significado para la población, etc. Todo ello supone que, tal y como establecen diferentes autores y, dentro del marco de la actual legislación Europea, el diseño, planificación y evaluación de los paisajes sonoros debería ser incluido en la futura planificación urbana. En este trabajo se propone un modelo, basado en redes neuronales, para la predicción de la estructura temporal y espectral de los paisajes sonoros, el cual, junto con la evaluación perceptual de los paisajes sonoros, nos permita la adaptación de estos a las necesidades y requerimientos de la población en los entornos urbanos.

## Abstract

The urban soundscapes are dominated by the noise generated by the road traffic, which acts as a noisy source that masks the rest of present sounds, which creates diverse problems that go beyond that the annoyance and displeasure in the population. On the other hand, the perception of the soundscapes by the population is not explained only for the

*Antonio J. Torija<sup>1</sup>; Diego P. Ruiz<sup>1</sup>; Ángel Ramos-Ridao<sup>2</sup>*

*<sup>1</sup>Dpto. Física Aplicada Facultad de Ciencias*

*<sup>2</sup>Dpto. Ingeniería Civil E.T.S. de I.C.C.P.*

*Universidad de Granada*

*ajtorija@ugr.es; druiz@ugr.es; ramosr@ugr.es*

sound pressure level to which it is exposed, but temporal and spectral characteristics of the soundscape must be taking into account, as well as other variables not acoustics, e.g., the attitude towards the source, sensitivity and habituation to the noise, meant to the population, etc. All this supposes that, as different authors establish and, within the framework of the current European legislation, the design, planning and assessment of the soundscapes should be included in the future town planning. In this work a model, based on neural networks, for the prediction of the temporal and spectral structure of the soundscapes is proposed, which, together with the perceptual assessment of the soundscapes, allows us the adjustment of these to the needs and requirements of the population in the urban environments.

## 1. Introducción

Tradicionalmente, la evaluación del ruido ambiental se ha centrado en la reducción del nivel de ruido de sonidos indeseados [1]. Más recientemente, con la introducción del concepto de paisajes sonoros [2], el problema se enfoca desde un punto de vista más holístico, de manera que el ambiente acústico urbano es observado como una multitud de sonidos los cuales pueden evocar emociones específicas, siendo considerado, por lo tanto, el paisaje sonoro urbano como una parte integral del ambiente vital urbano [3-5]. La evaluación y descripción de un paisaje sonoro debería, por tanto, tener en cuenta la relación entre la experiencia individual y subjetividad, con el contexto socio-cultural y físico en el que se desarrolla. Es en este último punto en el que se centra este trabajo.

La gran complejidad de las aglomeraciones urbanas supone un problema difícil de solventar desde el punto de vista de la predicción de las variables acústicas ambientales. En los entornos urbanos aparecen una gran variedad de fuentes sonoras y espacios, los cuales generan una gran heterogeneidad de situaciones acústicamente relevantes. Así, localizaciones con diferente composición de tráfico rodado, diferentes configuraciones urbanas, presencia de fuentes ruidosas difíciles de caracterizar (por ejemplo, actividades comerciales o de ocio, obras de construcción,...), existencia de espacios verdes, etc., genera la aparición de espacios sonoros con diferentes niveles de presión sonora, así como con grandes diferencias en cuanto a su composición temporal y espectral. La heterogénea fisonomía de los espacios urbanos junto con las características del ruido ambiental, con una gran variabilidad espacial, temporal y espectral, hace de su modelización y predicción un problema complejo y no-lineal, lo cual obliga a la utilización de una potente herramienta de minería de datos, las redes neuronales artificiales [6,7]. Las redes neuronales artificiales son un paradigma de aprendizaje y procesamiento automático que busca como último objetivo emular el comportamiento cerebral, o al menos algunas de sus funciones, tales como el aprendizaje. La utilización de este método proporciona flexibilidad, exactitud y una gran capacidad de adaptación [8], lo cual resulta muy interesante a la hora de abordar la modelización de las características acústicas de un entorno urbano.

Por esta razón, en esta investigación hemos desarrollado un modelo, basado en redes neuronales de retro-propagación, para obtener una adecuada predicción de las características temporales y espectrales, así como los niveles sonoros, para las diferentes localizaciones urbanas consideradas. El objetivo de este modelo es proporcionar una herramienta para el diseño, planificación y evaluación de los paisajes sonoros urbanos con la meta final de incorporar las necesidades de la población y obtener aglomeraciones urbanas acústicamente sostenibles.

## 2. Redes neuronales artificiales (ANN) como herramienta para la predicción sonora

Desde hace algunos años, las Redes Neuronales Artificiales (ANN) vienen recibiendo un fuerte interés como una metodología para la minería de datos, ya que ofrece los medios para modelar de manera eficiente y efectiva problemas complejos y no lineales [9-11]. Un modelo basado en ANN es capaz de encontrar relaciones (patrones) de forma inductiva por medio de varios algoritmos de aprendizaje basados en los datos existentes.

Se podría establecer que una Red Neuronal Artificial (ANN) es un paradigma de aprendizaje y procesamiento automático que busca como último objetivo emular el

comportamiento cerebral, o al menos algunas de sus funciones, tales como el aprendizaje [12].

Una red neuronal artificial (ANN) es un esquema de computación distribuida inspirada en la estructura del sistema nervioso humano. La arquitectura de una red neuronal está conformada de manera que se conectan múltiples procesadores elementales, siendo éste un sistema adaptativo que posee un algoritmo para ajustar sus pesos, y así alcanzar los requerimientos de desempeño del problema basado en muestras representativas.

Por lo tanto, se puede señalar que una ANN es un sistema de computación distribuida caracterizada por:

- Un conjunto de unidades elementales, cada una de las cuales posee bajas capacidades de procesamiento.
- Una densa estructura interconectada usando enlaces ponderados.
- Parámetros libres (pesos) que deben ser ajustados para satisfacer los requerimientos de desempeño.
- Un alto grado de paralelismo.

Las redes neuronales artificiales aprenden a solventar problemas desde la observación de una serie de ejemplos o muestran que representan el tipo de problema. De manera que, la ANN debe ser entrenada, usando parte de las muestras (conjunto de entrenamiento), para modificar los pesos de las conexiones, con la meta de dar una respuesta correcta para cada uno de los ejemplos del problema considerado. Para verificar que se produce un correcto aprendizaje, un grupo de muestras, las cuales no son utilizadas en el proceso de aprendizaje, deben ser usadas como conjuntos de verificación. Si la red entrenada ofrece una buena respuesta para el conjunto de verificación, entonces se puede establecer que la ANN generaliza bien y que el entrenamiento ha sido adecuado.

### 2.1. Redes neuronales unidireccionales (Feedforward)

Las redes neuronales unidireccionales (FANN) se corresponden con la clase de ANN más estudiada por el ámbito científico y, quizás la más utilizada en los diversos campos de aplicación. Este apartado se centra específicamente en las redes tipo Perceptron Multicapa (MLP). El MLP es una red neuronal unidireccional, es decir, con neuronas artificiales organizadas en capas y con ausencia de realimentación. El entrenamiento supervisado es realizado con el algoritmo de error de retro-propagación (o sus variantes) [13] basado en técnicas de gradiente, habitualmente utilizadas para la optimización matemática no-lineal [14]. El algoritmo minimiza una función, comúnmente cuadrática, de los errores entre las salidas esperadas (salidas del sistema) y las calculadas (por la ANN) en cada uno de los ejemplos de entrenamiento, modificando los parámetros del sistema (pesos de conexión de la ANN).

Dado un conjunto de observaciones, la tarea del aprendizaje neuronal es construir un estimador  $g_\lambda(\bar{x}, \bar{w})$  de la función desconocida  $h(\bar{x})$  del cual se conoce sólo un conjunto de datos.

$$g_\lambda(\bar{x}, \bar{w}) = \gamma_2 \left( \sum_{j=1}^k w_j^{[z]} \gamma_1 \left( \sum_{i=1}^m w_{ij}^{[1]} + w_{m+1,j}^{[1]} \right) + w_{\lambda+1}^{[z]} \right) \quad (1)$$

donde  $\bar{w} = (w_1, \dots, w_d)^T$  es el vector paramétrico a ser estimado y equivale a las ponderaciones de las conexiones entre las neuronas de la red,  $\gamma_1$  es una función no-lineal acotada y diferenciable con forma de función sigmoide o de radio basal,  $\gamma_2$  es una función que puede ser lineal o no lineal y  $\lambda$  es el parámetro de control que indica el número de neuronas en capa oculta.

La función  $\gamma_1$  es típicamente la función sigmoideal dada por

$$\gamma_1(z) = \frac{1}{1 + e^{-z}} \quad (2)$$

Si la función  $\gamma_2$  se elige no-lineal, deber ser estrictamente monótona, acotada y diferenciable. Estos requisitos son satisfechos por la función sigmoideal.

Una vez definida la arquitectura que se desea utilizar en un determinado problema, la red neuronal debe ajustarse a una muestra dada a través del proceso de aprendizaje. El algoritmo de aprendizaje consiste en un problema de inferencia de los parámetros libres de la red dado por las conexiones entre las neuronas. Como se ha establecido antes, el proceso de aprendizaje de una red neuronal unidireccional se caracteriza por ser supervisado, pues los parámetros de la red, conocidos como pesos, son estimados a partir de un conjunto de patrones de entrenamiento compuesto por patrones de entrada y salida [15].

Por tanto, este tipo de red neuronal está entrenado de acuerdo con un esquema de aprendizaje supervisado, mediante el cual, la red aprende del conjunto de datos de entrada y ofrece un conjunto de datos de salida. Es, por lo tanto, posible calcular un error  $E$  entre los valores calculados por la red neuronal ( $Y$ ) para cada entrada ( $X$ ) y la respuesta esperada u objetivo  $T$ . Durante el entrenamiento, para un conjunto de pares de datos ( $X, T$ ), el error entre la respuesta de la red y la respuesta deseada para el conjunto completo de pares deber ser minimizado:

$$\min E = \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^N (T_j - Y_j)^2 \quad (3)$$

donde  $M$  es el número de ejemplos de entrenamiento y  $N$  es el número de neuronas en la capa de salida.

Para estimar los parámetros de la red minimizando iterativamente la función de error, en [16] se reformuló el algoritmo de retropropagación como una generalización de la Regla Delta propuesta por Widrow-Hoff [17]. El término

retropropagación se refiere a la forma en que el gradiente de la función de error es calculado para redes neuronales unidireccionales. El aprendizaje involucra un ajuste de los pesos comparando la salida deseada con la respuesta de la red de manera que el error sea mínimo. La Regla Delta Generalizada expresa la manera en la que los pesos cambian para reducir el error entre la respuesta deseada y la respuesta de la ANN, para cada entrada  $X$ , y es expresada como:

$$w_{ji}(t+1) = w_{ji}(t) + \alpha(T_j - Y_j)X_i + \beta \Delta w_{ji}(t-1) \quad (4)$$

donde  $w_{ji}$  es el peso de conexión entre las neuronas  $i$  y  $j$ ;  $T_j$  e  $Y_j$  son los valores esperados y calculados respectivamente, en la neurona de salida  $j$ , para la entrada  $X$ ;  $X_i$  es el valor en la neurona  $i$  de la entrada  $X$ ;  $\beta$  es un coeficiente opcional de momento, el cual sirve para acelerar el entrenamiento de la red;  $\alpha$  es la tasa de aprendizaje de la red.

Cualquier valor introducido en la capa de entrada de la red neuronal es propagado hacia las subsiguientes capas. La respuesta de la red es comparada con los valores deseados y un valor de error es producido, el cual es propagado hacia las capas anteriores, de forma que los pesos son cambiados de acuerdo con la Regla Delta Generalizada. Con los nuevos pesos, la señal de entrada es de nuevo propagada hacia la capa de salida y un nuevo valor de error es calculado. Este proceso se repite hasta que el error llegue a ser mínimo o bien hasta que se alcanza un valor de error previamente especificado.

Una de las características más importantes y poderosas de las redes neuronales unidireccionales es su capacidad como aproximadores universales, esto es, dada cualquier función derivable y acotada, existe una red neuronal con un número suficiente de neuronas en la capa oculta que puede aproximar dicha función con el nivel de precisión deseada. La demostración de esta propiedad fue desarrollada por [18] a partir de una generalización del teorema de Stone-Weierstrass (ver [19]). En el trabajo [18] se demostró que las redes neuronales unidireccionales, en particular las redes neuronales unidireccionales con 3 capas, y con función de activación de tipo sigmoideal pueden aproximar cualquier función derivable y acotada con grado de precisión arbitrario para un número suficiente de neuronas en la capa oculta. Posteriormente, en [20] se extendieron los resultados obtenidos a las redes neuronales unidireccionales con funciones de activación no-sigmoideal en la capa oculta.

### 3. Metodología

#### 3.1. Proceso de obtención de datos

Para el estudio de los distintos paisajes sonoros de la ciudad de Granada, planteamos la idea de crear un sistema que, a través de redes neuronales, nos ayudara a la obtención de resultados precisos. Fueron seleccionadas 80 localizaciones de la ciudad de Granada, tratando incluir la gran heterogeneidad de situaciones

acústicamente relevantes característica de los entornos urbanos. Situaciones con diferentes condiciones de tráfico rodado, en distintos periodos del día, con diferente geometría y fisonomía de la localización, en parques y plazas urbanas (sin la presencia de tráfico rodado), con la presencia de fuentes ruidosas diferentes del tráfico rodado (obras de construcción, actividades comerciales,...), etc., fueron seleccionadas. Una vez realizadas todas las mediciones, se llevó a cabo el cálculo de los diferentes indicadores acústicos, de manera que, a partir de las diferentes variables de entrada se realizó la predicción de cada uno de los descriptores acústicos utilizados, con 5 minutos como periodo de integración.

### 3.2 Variables de entrada y salida del modelo

Para el desarrollo del modelo de predicción, como podemos observar en la tabla 1, han sido seleccionadas 33 variables de entrada, incluidas en 2 módulos, un módulo de propagación sonora y otro de emisión sonora, el cual, a su vez, está formado por 3 sub-módulos. Estos sub-módulos contienen variables para la caracterización del entorno, variables relacionadas con la evolución temporal del nivel sonoro y variables para la caracterización de la emisión

sonora generada por el tráfico rodado, respectivamente. Además, tenemos 36 variables de salida, el nivel de presión sonora ponderado A ( $L_{Aeq}$ ), el nivel de presión sonora no ponderado ( $L_{eq}$ ), la diferencia entre ambos descriptores, la variabilidad temporal y la impulsividad del nivel sonoro, así como el nivel sonoro para cada una de las bandas de 1/3 de octava entre 20 Hz – 20 KHz.

Con respecto al caudal de tráfico rodado, han sido considerados 5 tipos de vehículos: vehículos ligeros, vehículos pesados, autobuses, motocicletas-ciclomotores y vehículos de saneamiento urbano. Para la evaluación de la variabilidad temporal y de la impulsividad del nivel de presión sonora, en este trabajo han sido utilizados los descriptores Varianza Temporal del Nivel Sonoro (TSLV) y Factor Cresta (CF), respectivamente [21,22].

### 3.3. Estructura del modelo basado en redes neuronales de retro-propagación

Por sus propiedades para generalizar y su posibilidad de realizar un aprendizaje supervisado, hemos utilizado una red neuronal artificial, con la variante Levenberg-Marquardt con

Tabla 1. Variables de entrada y salida del modelo de predicción.

Variables de entrada			Variables de salida
Módulo	Sub-módulo	Variable	
Emisión sonora	Entorno	Tipo de día Periodo del día Actividad comercial/ocio Aparición de obras de construcción Tipo de localización Aparición de fuentes de agua Presencia de vegetación	$L_{Aeq}$ (0-100 dBA)  $L_{eq}$ (0-110 dB)  $L_{eq} - L_{Aeq}$ (0-30)  TSLV (0-60 dB <sup>2</sup> )
	Evolución temporal del nivel sonoro	Tiempo de estabilización Tipo de flujo de tráfico Eventos sonoros anómalos relacionados con el tráfico Eventos sonoros anómalos no relacionados con el tráfico	CF (0-10)
	Emisión sonora del tráfico rodado	Caudal de tráfico rodado ascendente/descendente Número de vehículos con sirena Velocidad media Pendiente de circulación Número de carriles ascendentes/descendentes Tipo de pavimento Estado de la superficie de rodadura	Bandas de 1/3 octava entre 20-20000 Hz (0-90 dB)
Propagación sonora		Geometría de la calle Anchura de la vía Altura de la vía Anchura de la calzada Distancia fuente-receptor	

\*Entre paréntesis se muestran los rangos de valores de las distintas variables de salida

# Ingeniería para el Control del Ruido

Servicios personalizados de Vibroacústica

**Ingeniería especializada**

Dominamos la ciencia del silencio

**Servicios únicos en acústica y vibraciones**

Tecnología líder

**Capacidad de adaptación**

Actuación en todos los sectores

[www.icrsl.com](http://www.icrsl.com)

E-mail: [icr@icrsl.com](mailto:icr@icrsl.com)

Tel/Fax. +34 93 428 63 39

**ICR**

Ingeniería para el Control del Ruido

retro-propagación con regulación Bayesiana (Función de entrenamiento Trainbr). La estructura de esta ANN (mostrada en la figura 1) es 33 variables de entrada, 36 neuronas en la capa oculta y 36 variables de salida.

Una vez seleccionada la estructura de la red neuronal, mostrada en la figura 1, construimos 5 conjuntos de datos diferentes. Los 5 conjuntos de datos fueron creados aleatoriamente a partir de los 533 registros de entrada. Por lo tanto, obtuvimos 5 conjuntos de datos, cada uno de los cuales contenía un conjunto de entrenamiento y un conjunto de validación, formados por registros diferentes. Los conjuntos de entrenamiento y validación contienen 350 y 183 registros, respectivamente. La ANN fue ejecutada 5 veces, cada una con diferentes pesos iniciales, con cada uno de los 5 conjuntos de datos. Esto nos ofrece 25 pruebas para evaluar la precisión de la predicción de los diferentes descriptores acústicos utilizados en este trabajo.

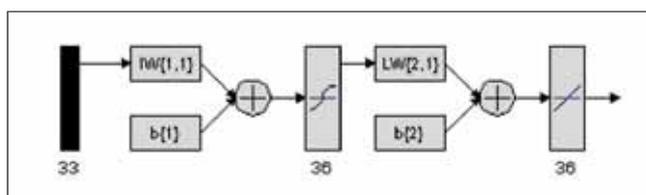


Figura 1. Estructura de la red neuronal propuesta

## 4. Resultados

### 4.1. Características acústicas de los paisajes sonoros de la ciudad de Granada

Antes de analizar los resultados obtenidos por el modelo desarrollado para la predicción del nivel sonoro y de la composición temporal y espectral de los paisajes sonoros se ha realizado un análisis de conglomerados jerárquicos para identificar los principales tipos de paisajes sonoros presentes en la ciudad y, observar las características acústicas de cada uno de ellos. Los diferentes paisajes sonoros observados en la ciudad han sido agrupados en 5 tipos.

Como se puede observar en la tabla 2, cuatro de los cinco tipos de paisajes sonoros urbanos están fuertemente dominados por el tráfico rodado. Los paisajes sonoros 2 y 5 están caracterizados por la presencia de un elevado caudal de vehículos ligeros (paisaje sonoro 2) y vehículos pesados, autobuses y motocicletas-ciclomotores. Los paisajes sonoros 1 y 3 tienen un caudal de tráfico rodado similar, diferenciándose en que el paisaje sonoro 1 es característico del casco antiguo de la ciudad, con un flujo intermitente de tráfico y la aparición de un gran número de eventos sonoros anómalos [23] y, el paisaje sonoro 3 es típico del borde

Tabla 2. Descripción de los paisajes sonoros analizados

Paisaje Sonoro	Descripción
1	Aparición, en calles del casco antiguo de la ciudad (calles muy estrechas tipo “U”), de un flujo intermitente de vehículos, con un muy alto tiempo de estabilización del nivel de presión sonora, principalmente, durante los periodos vespertinos-nocturnos y fines de semana. Aparición de un gran número de eventos sonoros anómalos, debido principalmente, a la gran pendiente de circulación o a la mala condición de la superficie de rodadura.
2	Vías principales de distribución de tráfico rodado, con un muy alto caudal de vehículos ligeros circulando a alta velocidad y, con la presencia de actividades comerciales o de ocio. Calles tipo “U” con un elevado número de carriles de tráfico (avenidas con gran anchura) y una altura de edificios circundantes muy elevada.
3	Aparición de un caudal de tráfico rodado no muy elevado, circulando a alta velocidad, en localizaciones situadas en el borde exterior de la ciudad. Calles cercanas al tipo “campo libre”, caracterizadas por ser avenidas muy anchas con una altura de edificios circundantes no muy elevada.
4	Parques, plazas y calles peatonales, con la presencia de vegetación (y fuentes de agua, especialmente en parques y plazas), con la ausencia de tráfico rodado en sus proximidades.
5	Vías principales de distribución de tráfico rodado, con un muy alto caudal de vehículos pesados, autobuses y motocicletas-ciclomotores circulando a alta velocidad y, con la presencia de actividades comerciales o de ocio. Calles tipo “U” con un elevado número de carriles de tráfico (avenidas con gran anchura) y una altura de edificios circundantes muy elevada.

Tabla 3. Valor de los descriptores acústicos en los diferentes paisajes sonoros analizados

Descriptor Acústico	Paisaje Sonoro 1	Paisaje Sonoro 2	Paisaje Sonoro 3	Paisaje Sonoro 4	Paisaje Sonoro 5
$L_{Aeq}$ [dB(A)]	66.50	71.10	66.35	63.59	74.11
$L_{eq}$ [dB]	76.15	80.61	74.91	72.61	83.21
$L_{eq}-L_{Aeq}$	9.65	9.51	8.57	9.02	9.10
TSLV [dB <sup>2</sup> ]	14.23	7.18	8.39	1.30	5.53
CF	1.07	0.95	0.83	0.49	1.03

exterior de la ciudad, con una alta velocidad de circulación en amplias avenidas (campo libre).

Los paisajes sonoros 2 y 5 tienen los valores más elevados para los descriptores acústicos  $L_{Aeq}$  y  $L_{eq}$  (tabla 3).

Además, estos dos paisajes sonoros tienen el mayor nivel sonoro para la totalidad de bandas de tercio de octava, siendo muy importante su composición de bajas-medias frecuencias (figura 2). Los paisajes sonoros 1 y 3 tienen valores similares para los descriptores  $L_{Aeq}$  y  $L_{eq}$  (tabla 3), sin embargo, debido a sus características (flujo de tráfico intermitente, eventos sonoros anómalos, etc.), el paisaje sonoro 1 posee un valor muy elevado de variabilidad temporal (TSLV) e impulsividad (CF) del nivel de presión sonora. Por lo que respecta a la composición espectral (figura 2), el paisaje sonoro 1 tiene una mayor composición en bajas frecuencias, debido a la gran pendiente de circulación, mala condición de la superficie de rodadura, geometría de la calle, etc., mientras que el paisaje sonoro 3 tiene una composición en medias frecuencias ligeramente superior.

Finalmente, el paisaje sonoro 4 está caracterizado por poseer los valores más reducidos en todos los descriptores acústicos utilizados (tabla 3). Tiene los valores más bajos de nivel de presión sonora, variabilidad temporal e impulsividad de todos los paisajes sonoros analizados. Además, observando su composición espectral (figura 2), se comprueba que el

contenido en bajas frecuencias es muy pequeño, debido a la ausencia de tráfico rodado en sus proximidades.

#### 4.2. Análisis de la relación entre las variables de entrada y los descriptores acústicos utilizados

En la tabla 4, podemos observar las variables con un mayor nivel de correlación con respecto a los diferentes indicadores acústicos utilizados en este trabajo.

En primer lugar, observamos que para los descriptores  $L_{Aeq}$  y  $L_{eq}$ , el caudal de motocicletas-ciclomotores ( $r = 0.645$ ) y el caudal de autobuses ( $r = 0.645$ ), son las variables de entrada con un mayor nivel de correlación respectivamente. El número de vehículos con sirena ( $r = -0.283$ ) es la variable de entrada con el descriptor  $L_{eq}-L_{Aeq}$ . Por su parte, las variables tiempo de estabilización ( $r = 0.840$ ) y presencia de eventos sonoros anómalos ( $r = 0.827$ ), son las variables con mayor grado de correlación con los descriptores para la caracterización temporal del nivel sonoro, TSLV y CF, respectivamente. Por lo que respecta a las bandas de 1/3 de octava (20 Hz – 20 KHz), se observan diferencias en función del rango de frecuencias en que nos situemos. En el rango 20-25 Hz, 31.5-125 Hz y 160-2500 Hz, el caudal de autobuses es la variable con un mayor nivel de correlación ( $r = 0.264, 0.534$  y  $0.549$ , respectivamente). Mientras, en el rango 3.15-20 KHz, el caudal de motocicletas-ciclomotores tiene el mayor grado de correlación ( $r = 0.449$ ).

#### 4.3. Resultados del modelo de predicción basado en redes neuronales

Una vez seleccionada la estructura de la red neuronal, ésta fue entrenada y validada 25 veces, 5 veces por cada uno de los 5 conjuntos de entrenamiento-validación previamente establecidos. En las tablas 5 y 6, podemos observar los resultados obtenidos a partir del modelo de predicción propuesto, basado en redes neuronales, para cada uno de los diferentes descriptores acústicos utilizados.

En función de los rangos de valores de los parámetros acústicos (mostrados en la tabla 1), comprobamos que la red neuronal propuesta alcanza una elevada precisión en

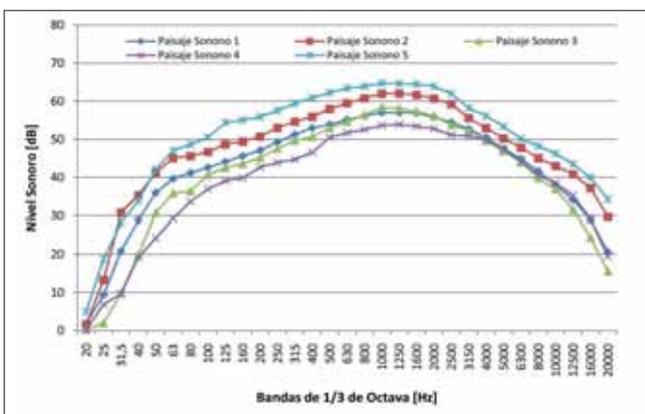


Figura 2. Composición espectral de los paisajes sonoros analizados

Tabla 4. Variables de entrada con mayor nivel de correlación con los descriptores acústicos utilizados

Descriptores Acústicos		Variables de Entrada
$L_{Aeq}$		Caudal de Motocicletas-Ciclomotores (0.645), Caudal de Vehículos Ligeros (0.514), Caudal de Autobuses (0.433), Caudal de Vehículos Pesados (0.425), Actividades Comerciales/Ocio (0.405)
$L_{eq}$		Caudal de Autobuses (0.645), Caudal de Vehículos Ligeros (0.508), Caudal de Motocicletas-Ciclomotores (0.504), Caudal de Vehículos Pesados (0.478), Tipo de Día (-0.417)
$L_{eq}-L_{Aeq}$		Número de Vehículos con Sirena (-0.283), Velocidad Media (-0.266), Pendiente de Circulación (0.238)
TSLV		Tiempo de Estabilización (0.840), Presencia de Eventos Sonoros Anómalos (0.582), Número de Vehículos con Sirena (0.478), Pendiente de Circulación (0.448)
CF		Presencia de Eventos Sonoros Anómalos (0.827), Número de Vehículos con Sirena (0.623), Tiempo de Estabilización (0.514), Caudal de Motocicletas-Ciclomotores (0.412)
Bandas de 1/3 de octava	20-25 Hz	Caudal de Autobuses (0.264), Caudal de Vehículos Pesados (0.227), Tipo de Pavimento (0.226), Pendiente de Circulación (0.207).
	31,5-125 Hz	Caudal de Autobuses (0.534), Caudal de Vehículos Pesados (0.486), Caudal de Vehículos Ligeros (0.447), Caudal de Motocicletas-Ciclomotores (0.431), Tipo de Pavimento (0.366).
	160-2500 Hz	Caudal de Autobuses (0.549), Caudal de Motocicletas-Ciclomotores (0.543), Caudal de Vehículos Ligeros (0.492), Caudal de Vehículos Pesados (0.482), Número de Vehículos con Sirena (0.426), Actividades Comerciales/Ocio (0.405)
	3150-20000 Hz	Caudal de Motocicletas-Ciclomotores (0.499), Actividades Comerciales/Ocio (0.426), Caudal de Vehículos Ligeros (0.408), Caudal de Autobuses (0.356), Caudal de Vehículos Pesados (0.319).

Nota: Entre paréntesis se especifica el nivel de correlación mediante el factor  $r$  de Pearson.

la predicción de los indicadores sonoros. Los descriptores que son estimados con mayor exactitud (tabla 4) son los parámetros  $L_{Aeq}$ ,  $L_{eq}$  y TSLV, mientras que los descriptores CF, bandas de 1/3 de octava (20 Hz – KHz) y  $L_{eq}-L_{Aeq}$  son estimados con un nivel de precisión ligeramente inferior.

Como observamos en la figura 3, el factor de correlación  $R^2$  varía dependiendo de la banda de 1/3 de octava considerada. Las bandas con un mayor valor del factor  $R^2$  están localizadas entre los 125-3150 Hz, mientras que las bandas con muy menor grado de correlación están

MURCIA : Plaza de la Cerámica, 6, 1º I. 30500 Molina de Segura (Murcia). ALICANTE : Avda. Juan

GANDIA (central) : Polígono Industrial Benieto. C/ Del Transporte, nº 12., Bloque 2, Nave 14. Apdo. correos 27. 46702 Gandia (Valencia)



 **Acusttel**<sup>®</sup>  
compromiso por la calidad acústica

empresa patrocinadora de Tecniacústica 06 en Gandía

## ingeniería para el control del ruido

instalaciones acústicas  
laboratorio de ensayos  
ingeniería y consultoría acústica  
I+D+i

# 902 36 02 52



Empresa con acreditación ENAC como laboratorio de ensayos acústicos

Sanctis Candela, 28. 03015 Alicante. MADRID : Cl. Isla Graciosa, 4/4ªA. 28034 Madrid

[www.acusttel.com](http://www.acusttel.com) | [info@acusttel.com](mailto:info@acusttel.com)

Tabla 5. Error medio de predicción (MPE) y valor  $R^2$  de los conjuntos de entrenamiento

Descriptores Acústicos	Conjunto 1		Conjunto 2		Conjunto 3		Conjunto 4		Conjunto 5	
	MPE	$R^2$								
$L_{Aeq}$ [dB(A)]	0.71±0.64	0.96	0.70±0.67	0.96	0.71±0.69	0.96	0.68±0.62	0.97	0.70±0.63	0.97
$L_{eq}$ [dB]	0.86±0.80	0.92	0.94±0.81	0.91	0.89±0.76	0.92	0.88±0.78	0.92	0.89±0.79	0.92
$L_{eq}-L_{Aeq}$	0.85±0.78	0.83	0.90±0.80	0.82	0.86±0.78	0.83	0.78±0.72	0.86	0.83±0.76	0.84
TSLV [dB <sup>2</sup> ]	0.91±0.84	0.98	0.92±0.87	0.98	0.89±0.88	0.98	0.92±0.90	0.98	0.91±0.92	0.98
CF	0.15±0.17	0.86	0.15±0.17	0.86	0.15±0.16	0.87	0.14±0.16	0.87	0.14±0.16	0.87
Bandas de 1/3 de octava [Hz]	1.54±1.46	0.89	1.59±1.46	0.89	1.57±1.48	0.89	1.51±1.44	0.90	1.52±1.45	0.90

Tabla 6. Error medio de predicción (MPE) y valor  $R^2$  de los conjuntos de validación.

Descriptores Acústicos	Conjunto 1		Conjunto 2		Conjunto 3		Conjunto 4		Conjunto 5	
	MPE	$R^2$								
$L_{Aeq}$ [dB(A)]	0.82±0.76	0.93	0.86±1.05	0.93	0.85±0.95	0.94	0.83±0.92	0.95	0.90±0.81	0.95
$L_{eq}$ [dB]	0.90±0.73	0.91	1.08±0.84	0.90	0.97±0.85	0.91	1.05±0.89	0.91	1.04±0.86	0.91
$L_{eq}-L_{Aeq}$	0.86±0.70	0.82	1.05±0.95	0.78	0.92±0.83	0.82	1.02±0.93	0.82	0.97±0.91	0.80
TSLV [dB <sup>2</sup> ]	0.96±1.05	0.98	1.14±1.25	0.97	1.10±1.15	0.97	1.18±1.12	0.97	1.07±1.18	0.96
CF	0.15±0.15	0.85	0.16±0.18	0.87	0.15±0.16	0.89	0.16±0.16	0.89	0.14±0.15	0.89
Bandas de 1/3 de octava [Hz]	1.70±1.69	0.83	1.79±1.86	0.85	1.76±1.78	0.86	1.72±1.73	0.87	1.76±1.72	0.86

situadas entre 20-31.5 Hz, 80-100 Hz y 16-20 KHz. Sin embargo, observando el valor medio del factor  $R^2$  de los 5 conjuntos de datos, verificamos que el valor mínimo es 0.77, alcanzando valores de 0.92.

En vista de los resultados obtenidos, podemos establecer que la red neuronal propuesta es capaz de predecir con una considerable precisión y exactitud el nivel de presión sonora (ponderado A y no ponderado), así como la composición temporal y espectral de los diferentes tipos de situaciones presentadas a la red, situaciones que incluyen una gran heterogeneidad y complejidad (características de las aglomeraciones urbanas).

#### 4.4. Comparación con los resultados obtenidos mediante regresión lineal múltiple

Una vez analizada la precisión y exactitud del modelo de predicción basado en redes neuronales, hemos comparado los resultados obtenidos con los resultados que obtendríamos si utilizáramos una regresión lineal múltiple (a partir de las mismas variables de entrada), en su lugar.

Observando los resultados mostrados en la tabla 7, podemos establecer que la utilización de una red neuronal supone una considerable mejora en cuanto a la consecución de un mayor grado de precisión y exactitud en la estimación

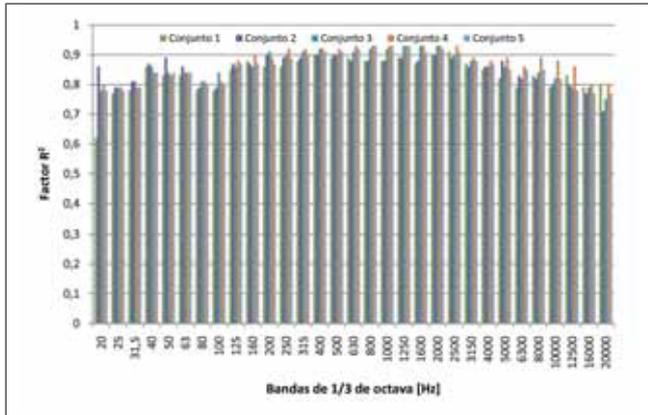


Figura 3. Valor  $R^2$  de las bandas de 1/3 de octava (20-20000 Hz) para los conjuntos de validación

de los parámetros acústicos utilizados, especialmente para los parámetros  $L_{eq}$ - $L_{Aeq}$  y bandas de 1/3 de octava.

Por lo que respecta al valor del factor  $R^2$  de las distintas bandas de 1/3 de octava, obtenido mediante la utilización de la red neuronal propuesta y una regresión lineal múltiple (figura 4), comprobamos que la red neuronal estima con mayor precisión todo el espectro sonoro.

Por otro lado, al igual que en el caso de la utilización de la red neuronal propuesta (figura 3), observamos que utilizando una regresión lineal múltiple, las bandas con mayor valor del factor  $R^2$  son las comprendidas entre 125-3150 Hz, sin embargo, en este caso la estimación de las muy bajas y las muy altas frecuencias se realiza con un nivel de precisión muy bajo.

#### 4.5. Análisis de los casos atípicos

Tras analizar la capacidad de predicción de la red neuronal para el conjunto de los 533 registros, los cuales incluyen, como

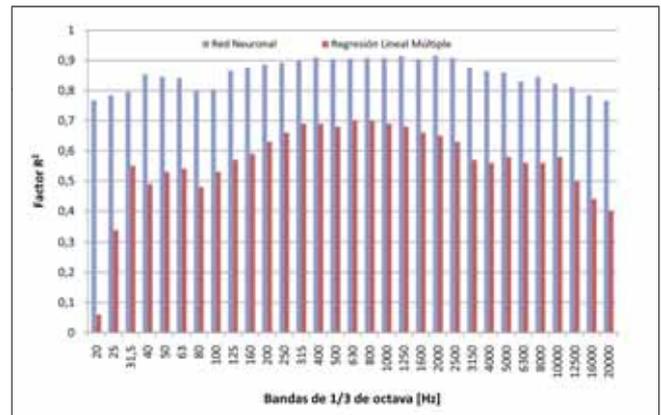


Figura 4. Comparación de los resultados obtenidos mediante la red neuronal propuesta y una regresión lineal múltiple en cuanto al valor  $R^2$  de las bandas de 1/3 de octava (20-20000 Hz) para los conjuntos de validación

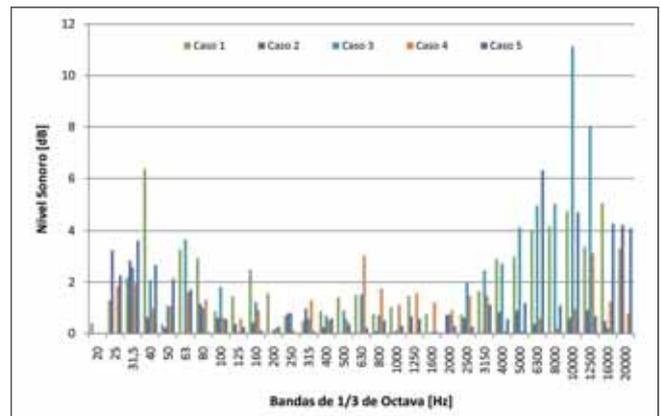


Figura 5. Valor del error medio de predicción (MPE) para las bandas de 1/3 de octava (20-20000 Hz) de los casos atípicos analizados

hemos mencionado anteriormente, una gran heterogeneidad de situaciones acústicamente relevantes, se ha analizado la respuesta del modelo a una serie de casos atípicos, situaciones

Tabla 7. Comparación de los resultados obtenidos mediante la red neuronal propuesta y una regresión lineal múltiple

Descriptor Acústico	Entrenamiento				Validación			
	Red Neuronal		Regresión Lineal Múltiple		Red Neuronal		Regresión Lineal Múltiple	
	MPE	$R^2$	MPE	$R^2$	MPE	$R^2$	MPE	$R^2$
$L_{Aeq}$ [dB(A)]	0.76±0.72	0.96	1.92±1.80	0.67	0.87±0.90	0.94	1.96±1.78	0.71
$L_{eq}$ [dB]	0.92±0.83	0.91	1.70±1.43	0.69	1.01±0.89	0.91	1.79±1.47	0.71
$L_{eq}$ - $L_{Aeq}$	0.85±0.78	0.84	1.50±1.17	0.51	1.03±0.99	0.81	1.96±1.56	0.29
TSLV [dB <sup>2</sup> ]	0.90±0.87	0.98	2.97±2.93	0.77	0.92±1.00	0.97	2.84±2.56	0.75
CF	0.14±0.16	0.87	0.16±0.21	0.75	0.16±0.16	0.88	0.19±0.30	0.75
Bandas de 1/3 de octava [Hz]	1.59±1.51	0.90	2.92±2.67	0.61	1.82±1.78	0.86	3.15±2.83	0.56

Tabla 8. Descripción de los casos atípicos analizados.

Caso	Descripción
1	Período nocturno con un muy alto tiempo de estabilización del nivel sonoro.
2	Plaza urbana con la presencia de una fuente de agua.
3	Localización con una muy reducida intensidad de tráfico rodado.
4	Localización con una gran cantidad de vehículos con sirena.
5	Congestión de tráfico.

Tabla 9. Valor del error medio de predicción (MPE) de los casos atípicos analizados

Descriptor Acústico	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Promedio
$L_{Aeq}$ [dB(A)]	0.27	0.42	0.27	0.30	0.29	0.47±0.26
$L_{eq}$ [dB]	0.19	0.25	0.71	0.65	1.15	0.66±0.32
$L_{eq}-L_{Aeq}$	0.23	0.24	1.03	0.47	0.93	0.55±0.52
TSLV [dB <sup>2</sup> ]	1.70	0.40	1.48	4.92	0.37	1.50±1.77
CF	0.32	0.05	0.02	0.62	0.07	0.16±0.18
Nivel Sonoro Bandas de 1/3 octava	2.10	0.72	1.95	1.07	1.32	1.52±1.34

donde aparecen valores extremos en algunas de las variables de entrada. Este tipo de situaciones son muy frecuentes en las aglomeraciones urbanas. Los casos atípicos seleccionados se pueden observar en la tabla 8.

En la tabla 9, se puede observar que para cada uno de los descriptores acústicos, entre el 60 y 100% de los casos atípicos estudiados tienen un error medio de predicción menor que el MPE promedio de los 5 conjuntos de validación para el mismo descriptor (tabla 6). Para los descriptores  $L_{eq}$  y  $L_{eq}-L_{Aeq}$  el porcentaje de casos atípicos con MPE menor que el MPE promedio de los 5 conjuntos de validación es del 80 %, mientras que para los descriptores TSLV, CF y composición espectral (bandas de 1/3 de octava) es del 60 %. Para el caso del descriptor  $L_{Aeq}$ , todos los casos estudiados tienen un MPE más reducido que el MPE de los 5 conjuntos de validación.

En la figura 5, se muestra el error medio de predicción de los casos atípicos analizados, para las bandas de tercio de octava entre 20 Hz y 20 KHz. En este caso se verifica que los mayores valores de MPE aparecen entre 40 Hz – 100 Hz, así como por encima de 6300 Hz..

## Conclusiones

Los paisajes sonoros urbanos están fuertemente afectados por el tráfico rodado, sin embargo, en un ambiente

urbano podemos encontrar situaciones en las que el tráfico rodado no tiene una incidencia directa (e.g. parques urbanos, paseos peatonales, ubicaciones comerciales, etc.). Además, el impacto del tráfico rodado será diferente dependiendo de multitud de factores, e.g. intensidad de tráfico, tipo de flujo de circulación, tipo de vehículo, velocidad de circulación, pendiente, etc., así como la propagación sonora será diferente dependiendo de la geometría de la calle, tipo de superficie de rodadura, etc. Esto genera la aparición de diferentes tipos de paisajes sonoros urbanos, los cuales, como hemos observado en este trabajo (sección 4.1), tienen características acústicas totalmente diferentes. Todo esto genera la aparición, en la ciudad, de una amplia variedad de situaciones, que deben ser modeladas con el fin de conseguir una predicción de las características temporales y espectrales así como del nivel de presión sonora, con la meta final de evaluar y gestionar los diferentes espacios sonoros urbanos, los cuales son percibidos e interpretados por la población, además de por factores no acústicos tales como información aportada, valor para el oyente, contexto social, etc., en función de su magnitud (nivel de presión sonora) y su composición temporal y espectral.

En vista de los resultados obtenidos, podemos concluir que la red neuronal propuesta logra una buena predicción de la composición temporal y espectral, así como del nivel de presión sonora (ponderado A y no ponderado) con una gran precisión y exactitud. Esta metodología y los resultados obtenidos, nos permiten ser optimistas con la posibilidad de

tener una herramienta para la predicción del nivel de presión sonora y de la composición temporal y espectral, la cual pueda ayudar a la integración de las variables acústicas en la planificación urbana, con el objetivo final de adaptar los paisajes sonoros urbanos a la población que interactúa con ellos.

## Agradecimientos

El autor principal de este trabajo quiere expresar su agradecimiento a la Sociedad Española de Acústica por la oportunidad concedida a través del premio “Andrés Lara” para continuar con su investigación.

## Referencias

- [1] De Coensel, B., 2007. Introducing the temporal aspect in environmental soundscape research. PhD dissertation. ISBN 978-90-8578-133-2. Dept. of Information Technology, Faculty of Engineering Sciences, Ghent University, Belgium.
- [2] Schafer, R.M., 1994. The soundscape: our sonic environment and the tuning of the world. Destiny Books, Rochester, Vermont, USA.
- [3] Dubois, D., Guastavino, C., Raimbault, M., 2006. A cognitive approach to urban soundscapes: using verbal data to access everyday life auditory categories. *Acta Acustica united with Acustica* 92(6), 865-874.
- [4] De Coensel, B., Botteldooren, D. 2006. The quiet rural soundscape and how to characterize it. *Acta Acustica united with Acustica* 92(6), 887-897.
- [5] Berglund, B., Nilsson, M.E. On a tool for measuring soundscape quality in urban residential areas. *Acta Acustica united with Acustica* 92(6), 938-944.
- [6] Avsar, Y., Saral, A., Gönüllü, M.T., Arslankaya, E., Kurt, U. 2004. Neural network modeling of outdoor noise levels in a pilot area. *Turkish Journal of Engineering & Environmental Sciences* 28, 149-155.
- [7] Cammarata, G., Cavalieri, S., Fichera, A. 1995. A neural network architecture for noise prediction. *Neural Networks* 8(6), 963-973.
- [8] Patra, J.C., Panda, G. 1998. ANN-based intelligent pressure sensor in noisy environment. *Measurement* 23, 229-238.
- [9] Akay, M.F., Inan, C., Bradshaw, D.I., George, J.D. 2009. Support vector regression and multilayer feed forward neural networks for non-exercise prediction of  $VO_{2max}$ . *Expert Systems with Applications* 36, 10112-10119.
- [10] Liu, M. 2009. Optimal exponential synchronization of general chaotic delayed neural networks: An LMI approach. *Neural Networks* 22, 949-957.
- [11] Wang, Z., Liu, Y., Liu, X. 2009. State estimation for jumping recurrent neural networks with discrete and distributed delays. *Neural Networks* 22, 41-48.
- [12] Haykin, S. 1998. *Neural networks. A comprehensive foundation*. Prentice Hall, New York, USA.
- [13] Werbos, P. 1994. The roots of backpropagation: From ordered derivatives to neural networks and political forecasting. Wiley, Hoboken, New Jersey, USA.
- [14] Kelley, C.T. 1999. *Iterative methods for optimization*. SIAM Press, Philadelphia, USA.
- [15] Allende, H., Maraga, C., Salas, R. 2002. Artificial neural networks in time series forecasting: A comparative analysis. *Kybernetika* 38, 685-707.
- [16] Rumelhart, D.E., Hinton, G.E., Williams, R.J. 1986. Learning representations by backpropagation errors. *Nature* 323, 533-536.
- [17] Widrow, B., Hoff, M.E. 1960. Adaptive switching circuits. *IRE WESCON Convention Record*, 96-104. New York, USA.
- [18] Hornik, K., Stinchcombe, M., White, F. 1989. Multilayer feedforward networks are universal approximators. *Neural Networks* 2, 359-366.
- [19] Rudin, W. 1964. *Principles of Mathematical Analysis*. McGraw-Hill, New York, USA.
- [20] White, H. 1992. *Artificial neural networks: Approximation and learning theory*. Basil Blackwell, Oxford.
- [21] Torija, A.J., Ruiz, D.P., Ramos, A. 2007. A method for prediction of the stabilization time in traffic noise measurements. In *Proceedings of the 19<sup>th</sup> International Congress on Acoustics, Madrid, Spain*.
- [22] Torija, A.J., Ruiz, D.P., Botteldooren, D., De Coensel, B. 2008. Taking into account the temporary and spectral structure of the sound energy for the characterization of the annoyance generated by the road and railway traffic. In *Proceedings of the 5<sup>th</sup> Iberian Congress of Acoustics (Acustica '08), Coimbra, Portugal*.
- [23] Torija, A.J., Ruiz, D.P., Ramos, A. 2007. Obtaining of a factor to describe the anomalous sound events in traffic noise measurements. In *Proceedings of the 19<sup>th</sup> International Congress on Acoustics, Madrid, Spain*.



Y tú, ¿qué valor le das al silencio?

# ChovACUSTIC®

Aislamiento acústico para la edificación

Disponible el libro "Aislamiento Acústico para la Edificación".  
Solicítelo gratuitamente en el 902 10 90 20 o en [www.chova.com](http://www.chova.com)



**ChovA**  
SISTEMAS DE IMPERMEABILIZACIÓN  
Y AISLAMIENTO

Obtenga la Propuesta Técnica Idónea para  
su proyecto y ejecución de obra.



LÍNEA ATENCIÓN AL CLIENTE: 902 10 90 20

[www.chova.com](http://www.chova.com)

e mail: [chova@chova.com](mailto:chova@chova.com)

