

DESARROLLO DE UNA RED NEURONAL PARA ESTIMAR EL OXÍGENO DISUELTO EN EL AGUA A PARTIR DE INSTRUMENTACIÓN DE E.D.A.R.

Antonio Iriondo Saint-Gerons
Fundación CETENASA. Polígono Mocholí, Plaza Cein, nº 4, 31110 Noáin (Navarra)
E-mail: airiondo@cetenasa.es

Jorge Mota Adrados
Fundación CETENASA. Polígono Mocholí, Plaza Cein, nº 4, 31110 Noáin (Navarra)
E-mail: jmota@cetenasa.es

Resumen

El oxígeno disuelto constituye una variable fundamental en la evaluación de la calidad del agua de los ríos, ya que determina la diversidad de organismos presentes en su seno. Para monitorizar su valor, se ha desarrollado un sensor software haciendo uso de la técnica de redes neuronales. La red neuronal elegida es de tipo backpropagation, con una arquitectura sigmoideal/lineal de dos capas, que utiliza como entradas tres variables medibles: temperatura, turbidez y potencial redox. La solución propuesta se ha mostrado idónea en términos de error de estimación, complejidad y entradas necesarias. La red permite, además, reducir costes en instrumentación, al ser los datos requeridos como entradas deducibles, en una E.D.A.R., a partir de los sensores para el control de los procesos de desnitrificación y decantación. Pruebas de validación, realizadas sobre datos tomados de las estaciones medioambientales de Navarra, prueban la efectividad de la solución propuesta.

Palabras Clave: Redes neuronales; Aguas residuales; Monitorización.

1 INTRODUCCIÓN

El aumento en los últimos años de los vertidos procedentes de entornos urbanos e industriales ha convertido a nuestros ríos en uno de los problemas más serios de contaminación de nuestro planeta. La aparición de nuevas y más restrictivas directivas anticontaminación ha provocado la construcción de numerosas estaciones depuradoras de aguas residuales (E.D.A.R.), y con ello la necesidad de ahondar en la sensorización necesaria para controlar la calidad del agua.

Los sensores software, basados en estimar el valor de una o más variables a partir de otras medibles vía hardware, han experimentado en los últimos años un gran desarrollo debido a las ventajas de implementación y coste que ofrecen en relación con los sensores hardware. De entre las diferentes técnicas utilizables para el desarrollo de estos sensores, las redes neuronales han tomado gran auge en el entorno de las aguas residuales, debido a su habilidad para aprender complejas interrelaciones entre variables, como las que caracterizan a las variables biológicas que gobiernan la calidad de las aguas de los ríos [5]. Así por ejemplo, se han desarrollado redes neuronales para la monitorización de nutrientes derivados del nitrógeno y el fósforo [3], la decantación de sólidos suspendidos en el agua [1], o, con propósitos de control, el establecimiento del dosado de coagulante preciso para favorecer la precipitación de partículas sólidas [6].

En este trabajo, se aprovecha la técnica de redes neuronales para el desarrollo de un sensor software que permita monitorizar el nivel de oxígeno disuelto en aguas fluviales. Se trata de una de las variables más representativas de la calidad del agua, al ser el componente que da soporte a la vida vegetal y animal en el agua. En general, niveles altos de oxígeno disuelto están asociados a mayor calidad de agua, mientras que los niveles bajos imposibilitan la supervivencia de muchos peces y otros organismos. Su importancia ya se ha traducido en algunas propuestas para controlar su nivel, haciendo uso de técnicas clásicas de control [4]. El sensor desarrollado en este trabajo permite aprovechar instrumentación típica utilizada en control del proceso de depuración del agua en una E.D.A.R., con el fin de reducir costes de implantación. La exposición se ha organizado como sigue: en el apartado 2 se describe la instrumentación de una E.D.A.R. utilizable por la red neuronal. En el apartado 3 se describe la fase de entrenamiento de la red, evaluándose la solución propuesta en términos de precisión, complejidad y entradas requeridas por

la red. El apartado 4 muestra los resultados alcanzados con la red neuronal para estimar el oxígeno disuelto en el agua. Finalmente, en el apartado 5 se enumeran las conclusiones relativas a la solución propuesta.

2 DEFINICIÓN DE LA INSTRUMENTACIÓN

El nivel de oxígeno disuelto en aguas fluviales está relacionado, principalmente, con la temperatura del agua y los sólidos en suspensión. La temperatura determina la solubilidad del oxígeno en el agua, aumentando ésta al disminuir la temperatura. En cuanto a los sólidos en suspensión, son partículas de tamaño hasta 1 micra que bloquean la luz solar, y evitan que las plantas acuáticas obtengan la luz solar que necesitan para la fotosíntesis. En consecuencia, las plantas producen menos oxígeno, y se reduce el nivel de oxígeno disuelto. Además, las plantas morirán más fácilmente y serán descompuestas por las bacterias en el agua, lo cual reducirá aún más el nivel de oxígeno. Los sólidos en suspensión también absorben calor adicional de la luz solar, lo cual ocasiona que el agua sea más caliente, disminuyendo la solubilidad del oxígeno en el agua y con ello el nivel de oxígeno disuelto.

Estas complejas relaciones obligan al empleo de técnicas de redes neuronales para la estimación del nivel de oxígeno disuelto en el agua. La red se desarrollará a partir de datos de temperatura, turbidez y potencial redox, medidos en las estaciones medioambientales de Navarra. La utilización de estos datos presenta la ventaja de aprovechar instrumentación disponible en una E.D.A.R con tecnología de fangos activados, lo cual reduciría costes de implantación. En efecto, el potencial redox, que mide la energía química de oxidación-reducción y la concentración de compuestos oxidados, es el parámetro de control de los procesos de nitrificación-desnitrificación en la E.D.A.R. Su medida ya ha sido aprovechada, en propuestas similares, para la monitorización de nutrientes [3]. En cuanto a la turbidez, su medida está estrechamente ligada al de los sólidos en suspensión, ya que refleja el grado de penetración de la luz solar. Sin embargo, la turbidez refleja también pérdida de transparencia del agua ocasionada por los sólidos en suspensión, y en este sentido está, asimismo, estrechamente ligada con el color aparente del agua. Esta característica invita a la utilización de técnicas de procesamiento de imagen para la detección del color del agua. Esta técnica ya ha sido propuesta en E.D.A.R. como medida cualitativa, a nivel microscópico, del grado de decantación de la biomasa suspendida en el agua. En este contexto, en un trabajo anterior ya se ha presentado un algoritmo para la caracterización de

fenómenos de bulking y foaming en la decantación, a partir de la identificación de la concentración de flóculos y filamentos [2]. La solución desarrollada se puede aprovechar para identificar la turbidez del agua. Dependiendo de la imagen, bastará aplicar la primera etapa del algoritmo (segmentación), o extender la aplicación a las otras dos (definición y clasificación de objetos).

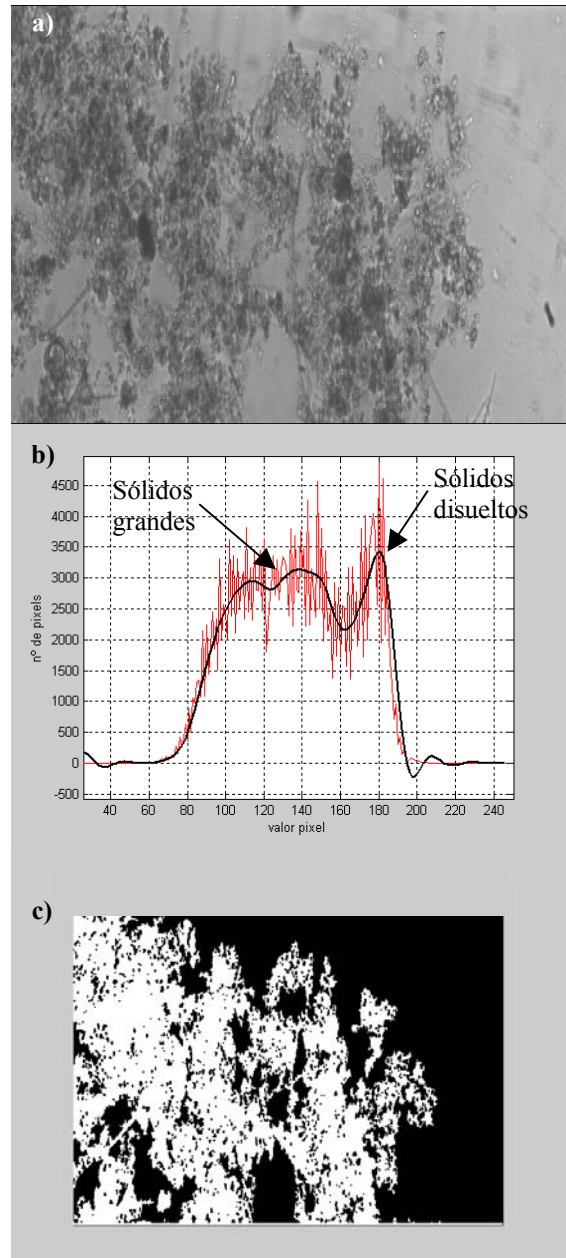


Figura 1. Detección de turbidez a partir de procesamiento de imágenes: a) imagen del agua; b) histograma; c) imagen discriminando los sólidos de gran tamaño.

La situación más desfavorable corresponde a una imagen como la que muestra la figura 1a, en la que, además del color del agua, se observan restos de

sólidos grandes, en tono más oscuro, cubriendo buena parte de la imagen. La figura 1b muestra el histograma filtrado resultante en la etapa de segmentación. Sin la presencia de los sólidos grandes, el histograma se caracterizaría por un único pico, cuyo valor de píxel, en abscisas del histograma, sería el correspondiente a la tonalidad del agua, y por consiguiente una medida de la turbidez. Sin embargo, la presencia de los sólidos grandes genera en el histograma un pico adicional de altura similar, por ser similares el número de pixels asociados a ambos, que dificulta la identificación del pico característico de la turbidez. No obstante, la discriminación automática de este pico se puede realizar a partir de la identificación de la anchura de los dos picos, ya que como se observa en la figura 1b, es mayor la del pico correspondiente a los sólidos grandes debido a la mayor dispersión de esta parte de la imagen. La anchura se puede identificar, indirectamente, a partir de la concentración de sólidos asociados a cada pico, aplicando a cada uno de estos picos las otras dos etapas del algoritmo de tratamiento de imágenes: 1)definición de objetos, mediante la aplicación de una erosión y una secuencia de filtros de la forma:

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}; \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}; \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}; \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

para separar formas sólidas en la imagen, y 2)clasificación de éstas, adecuando el ratio área/perímetro, que caracteriza cada forma, a objetos de gran tamaño. El resultado es una imagen para cada pico conteniendo los sólidos identificados, como la figura 1c obtenida para el pico de sólidos grandes, de la que se calcula la concentración. El pico de turbidez tendrá una imagen de concentración casi nula.

3 ENTRENAMIENTO DE LA RED

El desarrollo de la red neuronal contempla dos fases: 1) entrenamiento de la red, a partir de datos representativos de temperatura, turbidez, potencial redox y oxígeno disuelto, y 2) validación de la red a partir de datos distintos a los del entrenamiento. Para desarrollar la red se ha utilizado la Toolbox de redes neuronales de MATLAB (The MathWorks, Inc.). En este apartado se describe la primera fase, dejándose la segunda para un apartado posterior.

Mediante el entrenamiento, se ajustan los pesos de la red a partir de unos datos de entrada y de salida, de modo que la salida de la red neuronal reproduzca con suficiente exactitud la medida del oxígeno disuelto. El éxito del entrenamiento depende de la selección adecuada de los datos de entrada y salida, del tipo de red, y del algoritmo de entrenamiento.

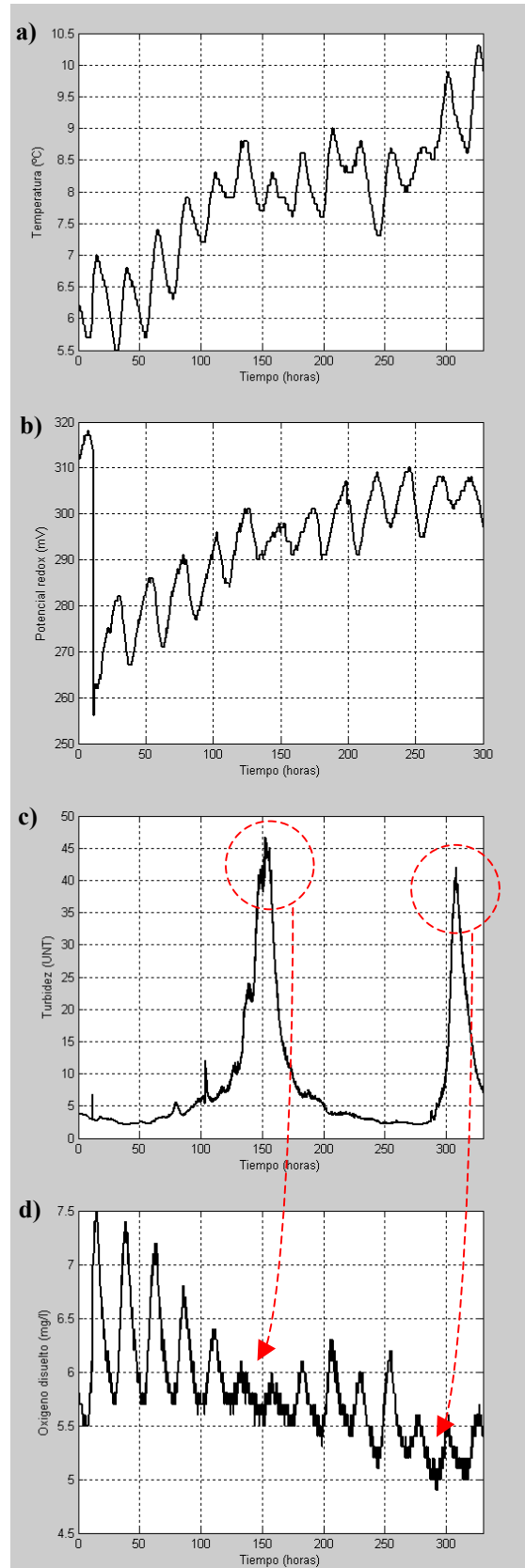


Figura 2. Datos para el entrenamiento de la red: a) temperatura; b) potencial redox; c) turbidez; d) oxígeno disuelto.

En la figura 2 se muestran los datos de temperatura, potencial redox, turbidez en Unidades de Turbidez Nefelométrica, y oxígeno disuelto seleccionados para el entrenamiento de la red. Se trata de datos diezminutales, correspondientes a la estación situada en Marcilla, tomados por un periodo de 14 días. Este periodo es suficiente para que los datos reflejen tanto las variaciones suaves producidas en el oxígeno disuelto, la temperatura y el potencial redox con el paso de los días, como las variaciones dentro del mismo día, a modo de dientes de sierra, debidas a los cambios de temperatura y actividad fotosintética de las plantas. Asimismo, en las figuras 2c y 2d se indica el efecto reductor producido sobre el oxígeno disuelto por los picos de turbidez.

Debido a su aptitud para representar relaciones no lineales, la red seleccionada para esta aplicación es de tipo backpropagation. La red sigmoideal/lineal de dos capas suele resultar suficiente para representar cualquier relación entre entradas y salidas, si la capa sigmoideal tiene suficientes neuronas. Esta arquitectura es la que se ha utilizado para el desarrollo de la red neuronal para la estimación del oxígeno disuelto. La figura 3 esquematiza la red neuronal propuesta. Los datos de temperatura, turbidez y potencial redox se aplican en una matriz de datos de entrada, en un número total de 2000 por variable para cubrir el periodo de 14 días. La salida es el vector de datos de oxígeno disuelto. Se ha establecido en 20 el número de neuronas sigmoideales de la capa de entrada que proporciona una convergencia suficiente, mientras que la capa de salida está integrada por una única neurona lineal.

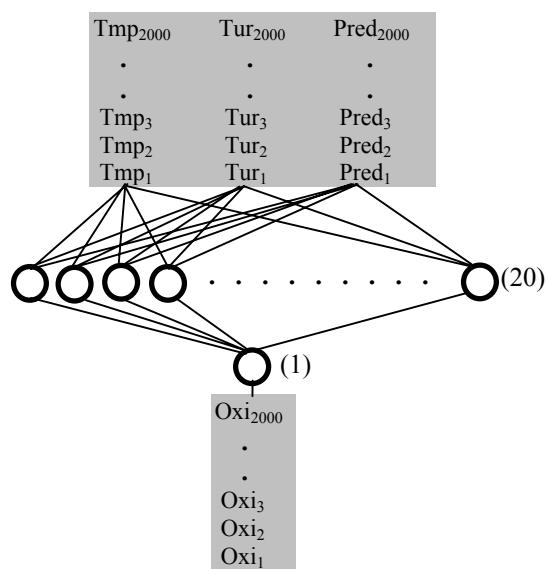


Figura 3. Esquema de la red neuronal para la monitorización del oxígeno disuelto.

Para entrenar a la red, se ha sometido a los datos a una normalización a partir del valor máximo, de

modo que su valor varíe entre 0 y 1. De este modo, las neuronas sigmoideales de la capa de entrada se han tomado de tipo logarítmico, aunque las de tipo tangencial también ofrecen un resultado similar. Como algoritmo de entrenamiento, se ha elegido el de Levenberg-Marquardt, ya que permite un entrenamiento rápido para redes de tamaño moderado. El algoritmo se ha aplicado con éxito sin modificar los parámetros de entrenamiento establecidos por defecto en la Toolbox de MATLAB.

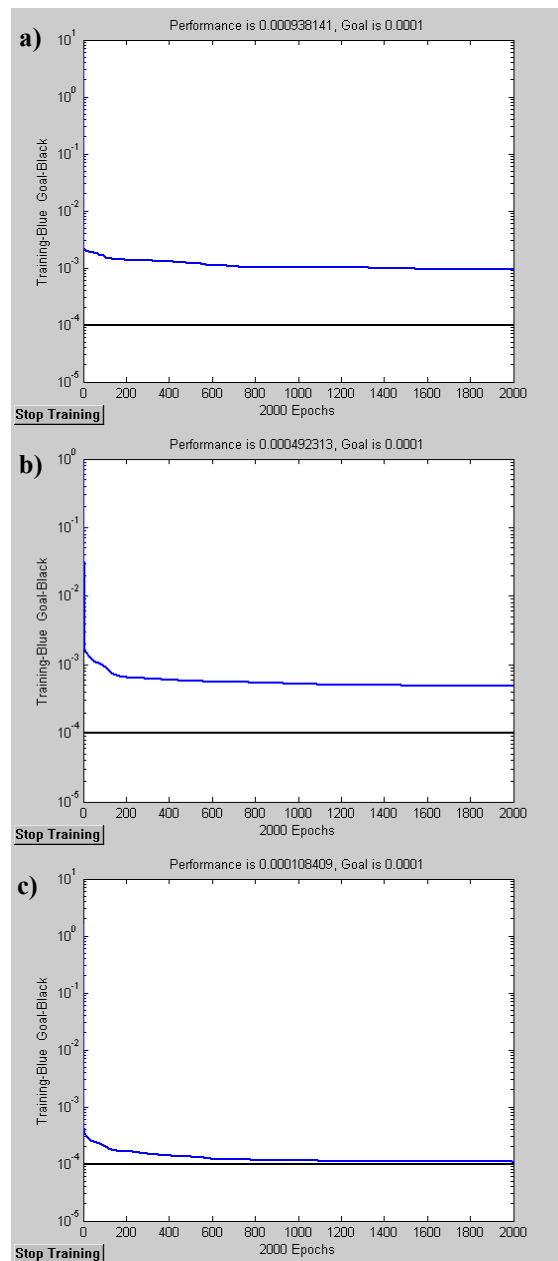


Figura 4. Evolución del error cuadrático medio: a) red de dos capas con entradas de temperatura y turbidez; b) red de tres capas con entradas de temperatura y turbidez; c) red de dos capas con entradas de temperatura, turbidez y potencial redox.

La bondad del entrenamiento de la red se suele evaluar en términos de error cuadrático medio entre la medida real del oxígeno disuelto y el resultado de la estimación por medio de la red neuronal. No obstante, para una valoración más completa de la solución, se ha querido evaluar la red, además, en términos de número de entradas necesarias y complejidad de la red. Para ello, se ha comparado la solución propuesta con otras dos: la misma red de dos capas, entrenada sólo con datos de temperatura y turbidez, y una red más compleja, integrada por tres capas, para ser utilizada también sólo con datos de temperatura y turbidez. Esta última red está integrada por las mismas capas de entrada y salida, y una capa intermedia entre ambas de 10 neuronas sigmoidales de tipo tangencial.

La figura 4 se muestra la evolución del error cuadrático medio a medida que avanzan los pasos del entrenamiento, para las tres redes consideradas. También se indica la tolerancia de error admitida para el ajuste. Como se puede observar, haciendo uso solamente de los datos de temperatura y turbidez, la red de tres capas permite reducir a la mitad el error cuadrático medio resultante con la red de dos capas, pero sin alcanzar el objetivo establecido. Sin embargo, la red de dos capas propuesta, entrenada con las tres entradas de temperatura, turbidez y potencial redox, permite alcanzar el objetivo deseado.

El resultado del entrenamiento se puede apreciar en la figura 5, donde se compara la medida del oxígeno disuelto con la estimación de la red neuronal, en los tres casos analizados. Como se puede observar, haciendo uso solamente de datos de temperatura y turbidez, la red de tres capas presenta, respecto de la de dos capas, solamente ligeras mejoras de estimación en algunos picos del oxígeno disuelto, como se puede apreciar en el pico señalado por la flecha. El mejor resultado corresponde a la solución propuesta con las tres entradas, que reproduce, con total fidelidad, los cambios de oxígeno disuelto producidos tanto en el día como a lo largo de los días. Se demuestra, en consecuencia, la mayor influencia de la inclusión de una tercera entrada sobre el incremento de la complejidad de la red, para mejorar la precisión de la red neuronal.

4 VALIDACIÓN DE LA RED

Para que la red entrenada resulte válida, ha de representar también la evolución del oxígeno disuelto al someterla a datos de entrada diferentes a los utilizados para su entrenamiento. Los nuevos datos se aplican a la red normalizados del mismo modo que en el entrenamiento. A modo de ejemplo, en la figura 6 se muestran nuevos datos de temperatura, potencial

redox, turbidez y oxígeno disuelto, que corresponden a un periodo de tiempo posterior a los datos utilizados para el entrenamiento de la red. Como se puede apreciar, los nuevos datos siguen una evolución similar a la reflejada en los datos de entrenamiento: variación suave de oxígeno disuelto a lo largo de los días en sentido contrario a la temperatura y el potencial redox, y variación diaria a modo de dientes de sierra, que se reduce con los picos de turbidez.

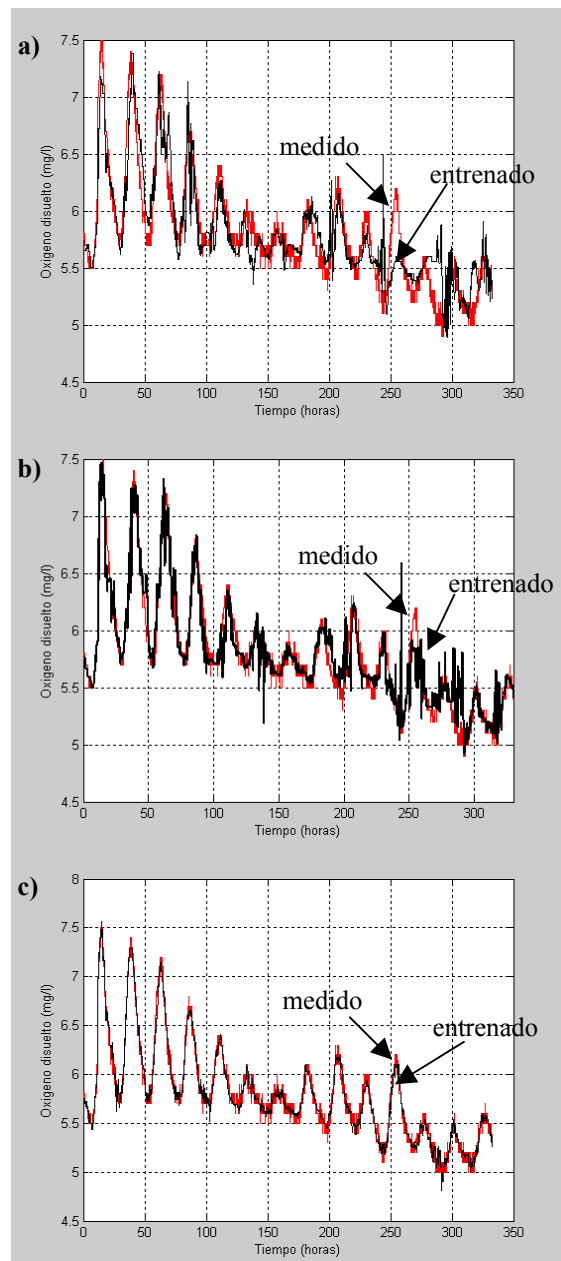


Figura 5. Simulación del entrenamiento de la red: a) red de dos capas con entradas de temperatura y turbidez; b) red de tres capas con entradas de temperatura y turbidez; c) red de dos capas con entradas de temperatura, turbidez y potencial redox.

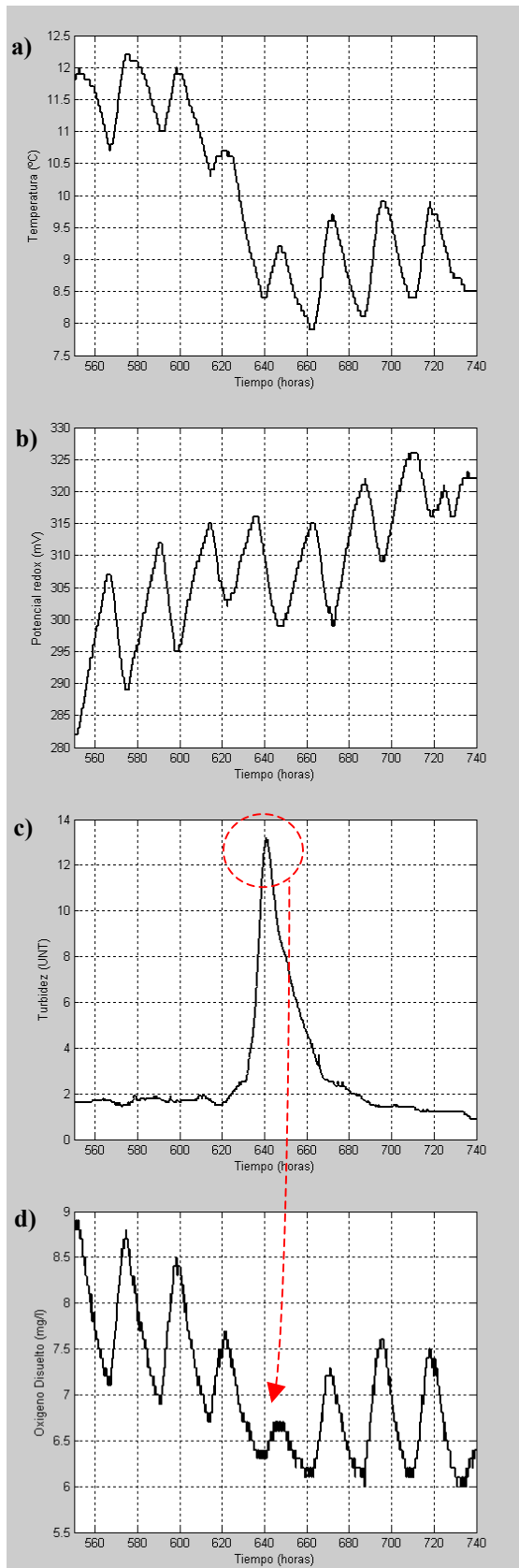


Figura 6. Datos para la validación de la red: a) temperatura; b) potencial redox; c) turbidez; d) oxígeno disuelto.

Sin embargo, esta tendencia es más difícil de mantener en los datos de temperatura, pudiéndose observar algunas desviaciones puntuales, como el descenso que se observa en la figura 6a en los datos comprendidos entre las 600 y 660 horas. En las redes neuronales entrenadas con un menor número de entradas, desviaciones puntuales pueden conducir a errores de estimación adicionales al valor del error cuadrático medio. Tal es el caso de la red neuronal entrenada solamente con los datos de temperatura y turbidez. En la figura 7a se muestra la estimación obtenida en este caso, con la red de dos capas (la estimación con la red de tres capas es similar). En la figura 7b se muestra la estimación obtenida con la red de dos capas con entradas de temperatura, turbidez y potencial redox, propuesta como solución. Sólo con datos de temperatura y turbidez, la red estima en el día variaciones de oxígeno disuelto mayores que las reales, y no sigue la evolución a lo largo de los días. Sin embargo, la solución propuesta, basada en la utilización, además, de los datos de potencial redox, es más robusta a variaciones imprevistas de la temperatura, resultando una estimación que reproduce con suficiente fidelidad las variaciones de oxígeno disuelto tanto durante el día, como a lo largo de los días.

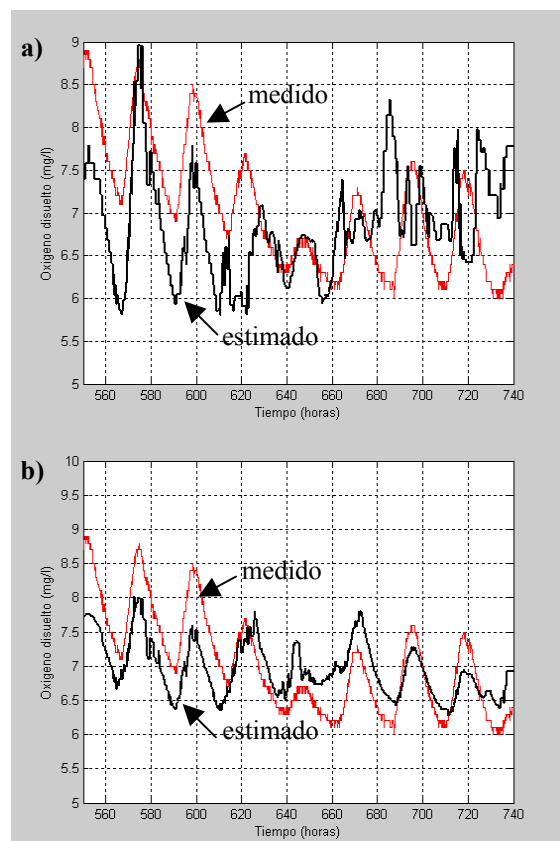


Figura 7. Validación de la red neuronal: a) red de dos capas con entradas de temperatura y turbidez; b) red de dos capas con entradas de temperatura, turbidez y potencial redox.

5 CONCLUSIONES

A partir de datos de calidad de agua tomados de las estaciones medioambientales de Navarra, se ha desarrollado una red neuronal para la estimación del oxígeno disuelto en aguas fluviales. La red desarrollada presenta las siguientes características:

- utiliza tres entradas de datos: temperatura, turbidez y potencial redox;
- se basa en una arquitectura simple de dos capas: la capa de entrada con neuronas no lineales, y la capa de salida con una neurona lineal;
- se ha demostrado que la solución propuesta resulta ideal en términos de error de estimación, complejidad de la red, y número de entradas precisas;
- se ha propuesto la técnica de procesamiento de imágenes, aplicada al control de sedimentación de fangos activos, como alternativa para estimar la turbidez. La aplicación de esta técnica, junto con el sensor de potencial redox de los procesos de desnitrificación, permitiría aprovechar la instrumentación típica en una E.D.A.R. para reducir los costes necesarios para implementar la red.

Agradecimientos

Los autores agradecen a la Comunidad de Trabajo de los Pirineos (CTP) la financiación recibida a través del proyecto "Monitorización y Control Centralizado de una Cuenca de Saneamiento" (MC³S), y al Departamento de Medio Ambiente del Gobierno de Navarra la cesión de los datos de las estaciones medioambientales utilizados para la realización de este trabajo.

Referencias

- [1] Capodaglio, A.G., Jones, H.V., Novotny, V., Feng, X., (1991) "Sludge bulking analysis and forecasting: Application of system identification and artificial neural computing technologies", *Water Research* 25 (10), pp. 1217-1224.
- [2] Iriondo, A., Goya, R., (2003) "Automatic Measurement of Floc and Filament Concentrations from Activated Sludge Images", *Proceedings del ITEE'2003 (Information Technologies in Environmental Engineering)*, Gdansk, Polonia.
- [3] Lee, L.S., Park, J.M., (1999) "Neural network modeling for on-line estimation of nutrient dynamics in a sequentially-operated batch reactor", *Journal of Biotechnology* 75, pp. 229-239.
- [4] Moreno, R., de Prada, C., Lafuente, J., Poch, M., Montague, G., (1992) "Non-linear predictive control of dissolved oxygen in the activated sludge process", *ICCAFT 5/IFAC-BIO 2*, Keystone, pp. 289-293.
- [5] Schlaeger, F., Schonlau, H., Köngelter, J., (2003) "An integrated water resources management approach for the River Spree and its catchment", *Water Science and Technology* 47 (7-8), pp. 191-199.
- [6] Valentin, N., Denoeux, T., (2001) "A neural network-based software sensor for coagulation control in water treatment plant", *Intelligent Data Analysis* 5, pp. 23-39.