

Influencia del rango de las variables de entrenamiento de RNA para la determinación de movimientos de presas arco

Francisco J. Riquelme, Jesús Fraile-Ardanuy, Miguel Á. Toledo

Universidad Politécnica de Madrid

friquelme@caminos.upm.es, jefar@caminos.upm.es, matoledo@caminos.upm.es

1 Introducción

La análisis de la auscultación de presas requiere de la interpretación de gran cantidad de variables de entrada y de salida. En este análisis hay involucrados parámetros que no se conocen con certeza y relaciones entre variables que no están claramente definidas. Frente a estas incertidumbres, las redes neuronales se presentan como una herramienta de gran utilidad [1] que permite resolver estos problemas altamente complejos en multitud de disciplinas de la ingeniería [2].

Sin embargo, la falta de una base teórica que permita definir a priori la red más adecuada a cada problema exige de un estudio detallado de todos los parámetros de la red [3].

En este trabajo se hace una revisión del estado del arte de las aplicaciones de redes neuronales. Se describe el caso piloto que servirá de referencia. Se define el modelo de red neuronal que se usa en este estudio y se muestran los resultados obtenidos en el estudio de la influencia del rango de entrenamiento de las variables de entrada y salida en la bondad de los resultados. Finalmente, se obtienen conclusiones y se presentan recomendaciones respecto al uso de redes neuronales para la interpretación de la auscultación de las presas.

2 Objetivo

El objetivo de esta investigación es valorar la importancia del rango de entrenamiento de las variables de entrada y control de las redes neuronales en la bondad de los resultados obtenidos en los modelos estudiados.

3 Estado del arte

Hay un gran número de publicaciones acerca de aplicaciones de redes neuronales a la ingeniería civil. La gran mayoría de estas aplicaciones están basadas en el algoritmo de aprendizaje back-propagation (BP) [4, 5, 6], aunque también hay aplicaciones de otros modelos de redes neuronales más poderosos y eficientes [2]. En la literatura hay trabajos que integran modelos de redes neuronales con otros paradigmas de cálculo como elementos finitos, lógica borrosa o algoritmos genéticos.

En [7] se utilizan redes neuronales como alternativa a los métodos estadísticos tradicionales como regresiones lineales múltiples para el análisis de variables de monitorización de presas como los desplazamientos, tensiones y flujos para detectar posibles anomalías. En particular, se estudia una presa de gravedad monitorizada mediante péndulos invertidos, medidores de flujo y juntas, así como termómetros midiendo la temperatura del agua y del hormigón. Utiliza redes neuronales con algoritmo de aprendizaje BP para predecir desplazamientos

de los péndulos en las tres direcciones, probando con un gran número de arquitecturas de redes. Un análisis preliminar mostró que la temperatura del hormigón es el mejor predictor de los desplazamientos de los péndulos. Mediante pruebas y errores construyó la red neuronal más apropiada, observando que al incrementar el número de capas ocultas los resultados del ajuste en los casos de entrenamiento mejoraba, sin embargo, la habilidad de

predicción en los casos de validación disminuía con la complejidad del modelo. El número óptimo de capas ocultas exige un compromiso entre el ajuste a los datos de entrenamiento y la capacidad de predicción.

La mayoría de los problemas requiere del uso de una única capa oculta o, a veces, dos capas. Generalmente es mejor comenzar con una capa e incrementar el número hasta obtener los resultados deseados, por ejemplo en [7] se usa una única capa en su análisis y, usa un número de nodos ocultos menor que el número de nodos de entrada, incrementando desde uno hasta seis (número de variables) el número de nodos ocultos. Finalmente, se concluye que redes de una capa oculta con dos o tres nodos aportan predicciones suficientemente buenas de los desplazamientos de los péndulos. Constató que el coste computacional de entrenar las redes es mayor que el de los modelos estadísticos, con la ventaja de que los modelos de redes neuronales no requieren de hipótesis que relaciones los datos de entrada y de salida, de modo que resultan ventajosos en aplicaciones donde no se conocen bien las relaciones entre variables. Además muestra que el grado de no linealidad de las aproximaciones puede ser fácilmente cambiado modificando el número de capas del modelo.

En [8] se desarrolla un modelo de redes neuronales y otro de elementos finitos para comparar los resultados con las mediciones de flujos a través de una presa de materiales sueltos. El modelo de redes neuronales emplea aprendizaje mediante BP, una capa oculta y función de activación sigmoidea. Ambos modelos fueron calibrados y verificados utilizando un piezómetro en una sección de la presa. Ambos modelos obtuvieron resultados satisfactorios comparados con las mediciones de niveles de agua medidas por los piezómetros. El estudio reveló que los modelos de redes neuronales pueden ser una buena herramienta para estudiar supresiones y, que en algunos casos, puede incluso superar a los modelos de elementos finitos. En sus experiencias [8] utilizó redes de tres neuronas en la capa de entrada, cinco neuronas en la capa oculta y una en la capa de salida. Las variables de entrada usadas fueron el nivel de aguas arriba, el nivel de aguas abajo y la identificación numérica de los piezómetros. El número óptimo de neuronas en la capa oculta fue encontrado por prueba y error. El período de entrenamiento abarcó aproximadamente dos años.

En [9] se destacan las ventajas de los sistemas de inteligencia artificial debido a la subjetividad involucrada en la interpretación de los datos de monitorizaciones de presas entre los cuales destaca el difícil manejo de grandes cantidades de información muchas veces incierta o incompleta, la necesidad de criterio ingenieril, conocimiento de la estructura particular, experiencia en el desarrollo de estructuras en general, y conocimientos generales de ingeniería para interpretar los datos. Se pretende minimizar la subjetividad de los expertos humanos en la monitorización y predicción del comportamiento de las estructuras hidráulicas.

4 Caso piloto

Se ha usado como caso piloto la presa de la Baells. Se trata de una estructura tipo bóveda de 3 centros y doble curvatura en hormigón en masa con 102,35 m sobre cimientos, una longitud de coronación de 403 m. La anchura en el pie de la presa es de 20,10 m y un espesor en el arco superior de 4 m (ver figura 1).

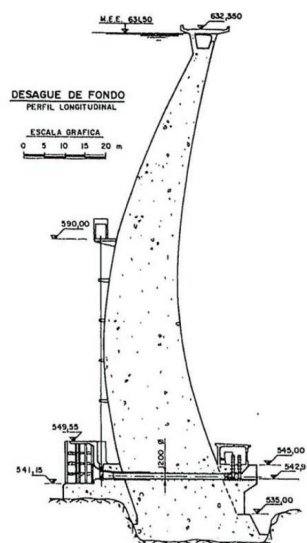


Figura 1 Presa de la Baells. Perfil.

Se estudia la relación entre las variables externas e internas de la presa, entendiendo como externas a todas aquellas medidas fuera del cuerpo de la presa, en este caso se restringe el estudio al nivel de embalse y la temperatura; y variables internas a aquellas medidas dentro del cuerpo de presa, que en este caso se corresponde con el desplazamiento radial de un péndulo central en su registro a mayor altura (ver figura 5).

4.1 Nivel de embalse

Se dispone de registros del nivel de embalse (H) entre el 25 de septiembre de 1980 y el 23 de octubre de 2008, fechas entre las cuales el nivel varió entre 585, 81 m.s.n.m. y 630, 18 m.s.n.m. En la figura 2 se muestra la variación del nivel de embalse a lo largo del período de estudio. El nivel de embalse depende, además de los condicionantes ambientales, de la explotación del embalse.

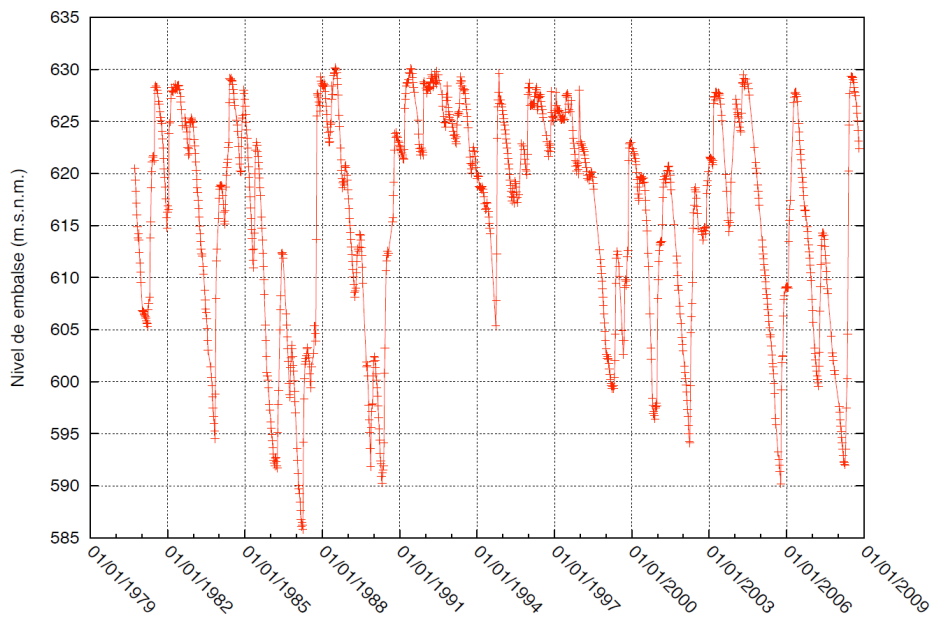


Figura 2 Nivel de embalse durante el período de estudio.

4.2 Temperatura

Se dispone de registros de temperatura ambiental (T) entre el 25 de septiembre de 1980 y el 23 de octubre de 2008, fechas entre las cuales la temperatura varió de forma cíclica con un período de un año. La temperatura informada corresponde al promedio simple entre la temperatura diaria máxima ($T_{\text{máx}}$) y la temperatura diaria mínima ($T_{\text{mín}}$).

$$T = \frac{T_{\text{máx}} + T_{\text{mín}}}{2}$$

En la figura 3 se muestra la variación de temperatura ambiental a lo largo del período de estudio.

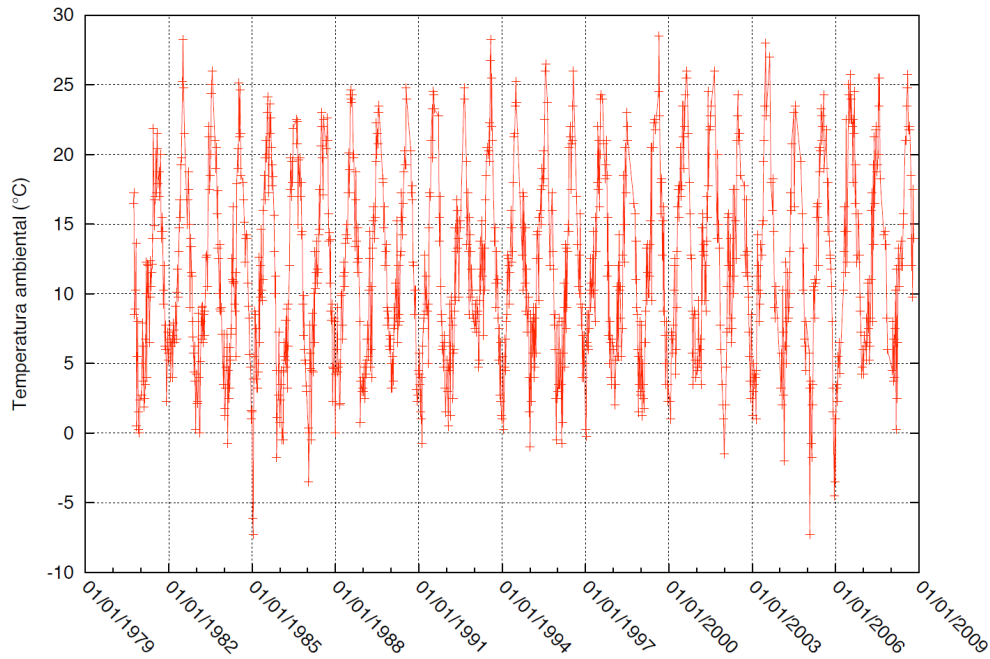


Figura 3 Temperatura ambiental durante el período de estudio.

Esta variable externa se puede caracterizar usando una curva periódica de ciclo anual (P) y función del tiempo cronológico (t), por ejemplo, los primeros términos de la serie de Fourier correspondiente y aproximada con los primeros términos de ésta.

$$T(t) \cong \frac{a_0}{2} + a_1 \cos\left(\frac{2\pi}{P} t\right) + a_2 \sin\left(\frac{2\pi}{P} t\right)$$

Para encontrar los coeficientes de la aproximación es necesario definir un origen del tiempo cronológico. En la figura 4 se muestra la aproximación obtenida junto a las mediciones de temperatura ambiental diaria usando un período de 365, 25 días.

Las medias móviles de temperatura para diferentes períodos, quedan definidas según la ecuación 3, en la que se puede reemplazar directamente la aproximación 2, obteniendo así medias móviles de temperatura consistentes con la aproximación de temperatura.

$$T_{MMP}(t) = \int_{t-P}^t \frac{T(t)}{P} dt$$

En este estudio se usa la media móvil de 90 días como medida indirecta de la temperatura, que ha mostrado, en estudios preliminares, ser la representación de la temperatura con la que se obtiene mejores resultados.

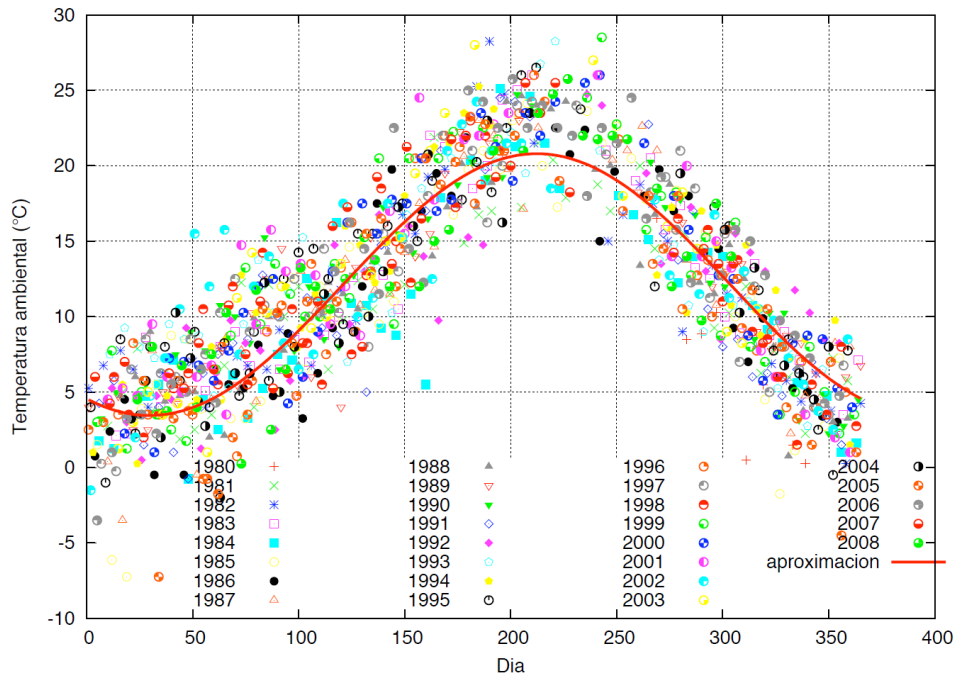


Figura 4 Aproximación de los ciclos anuales de temperatura ambiental.

4.3 Desplazamiento radial del péndulo

En la modelación del comportamiento de una presa tiene especial interés el estudio de los desplazamientos de los péndulos. El caso piloto tiene cuatro sistemas de péndulos, dos centrales y dos laterales dispuestos simétricamente como se muestra en la figura 5.

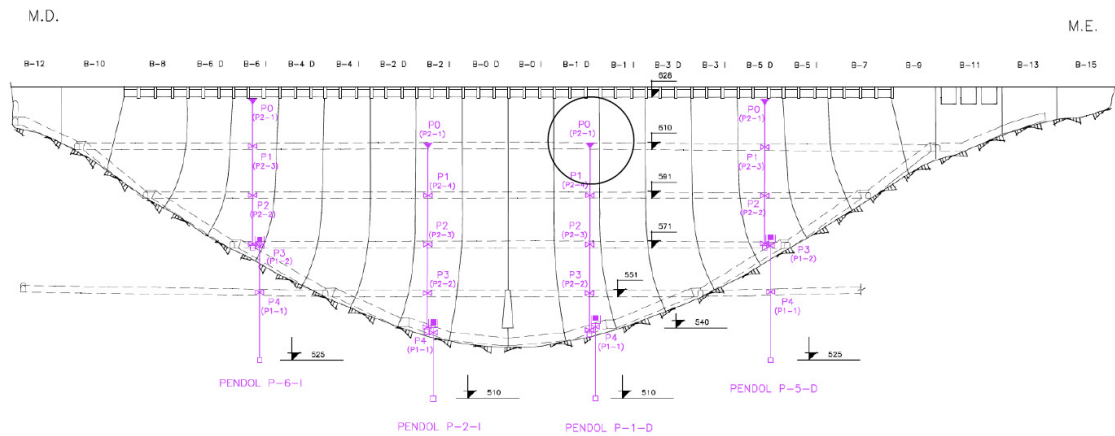


Figura 5 Esquema de los péndulos en la presa. Se señala con un círculo la cota más alta del péndulo estudiado.

De los desplazamientos radiales registrados en estos péndulos se ha escogido los correspondientes al péndulo P-1-D a la cota 610 m.s.n.m. que, por su posición, cabría esperarse que fueran los más amplios.

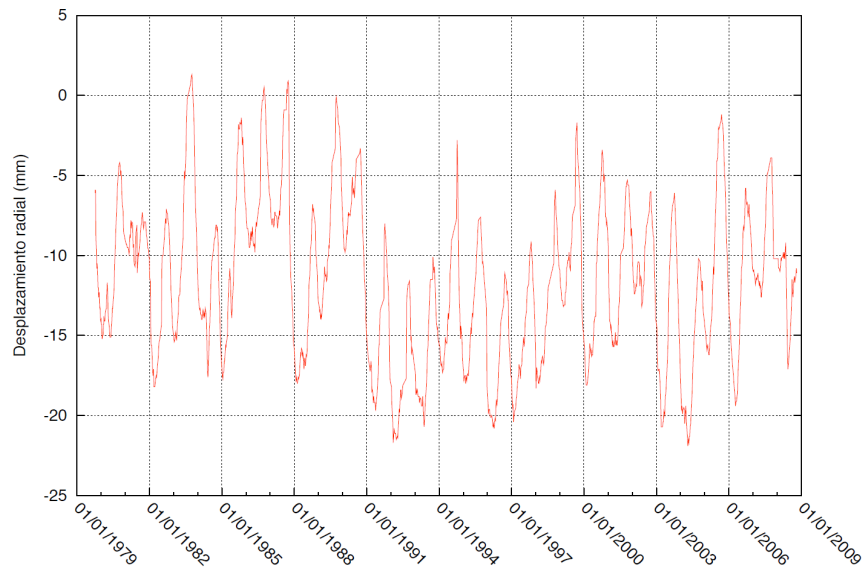


Figura 6 Desplazamiento radial del péndulo 1D a la cota 610 m.s.n.m.

5 Modelo de redes neuronales

Para estudiar la influencia del rango de datos de entrada y salida en la bondad de los resultados se usa un modelo simplificado de red neuronal que no incluye explícitamente al tiempo como variable de entrada. Las únicas dos variables de entrada son la temperatura, representada mediante la media móvil de 90 días (T_{MM90}) y, el nivel de embalse (H). La variable de salida es el desplazamiento radial del péndulo PID a la cota 610 m.s.n.m. ("). La arquitectura de la red se muestra en la figura 7, en la que los vínculos son estrictamente intercapa y el flujo de información ocurre en un único sentido, desde la capa de entrada a la de salida. El vínculo se produce entre todas las neuronas de una capa con todas las neuronas de la capa adyacente.

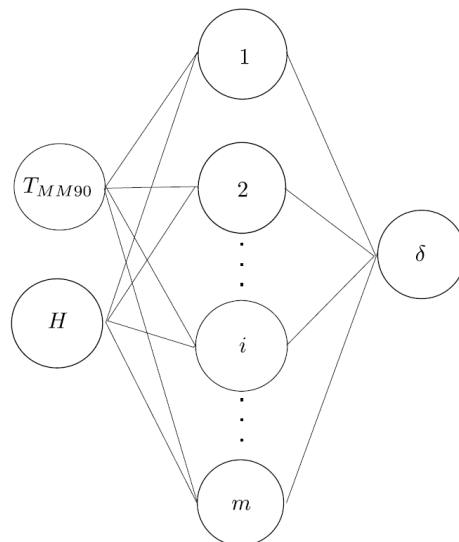


Figura 7 Arquitectura de red del modelo simplificado de 1 capa oculta y m neuronas ocultas.

El número de neuronas en los modelos de una capa oculta está directamente relacionado con la capacidad de aproximación y de predicción de estos. Pocas neuronas en la capa oculta podrían conducir a un error elevado en la aproximación y también en la predicción, esto es debido a que la red tendría una baja capacidad de aprendizaje. Por otro lado, un número muy elevado de neuronas permitiría memorizar los patrones de entrenamiento, disminuyendo el error en la aproximación, pero incrementándolo en la predicción. Esto es

equivalente a lo que ocurre en un sistema sobredeterminados en el que se aproximan exactamente los valores conocidos, pero se pierde bondad en la interpolación y extrapolación.

Se puede afirmar que si el número de neuronas en la capa oculta m es mayor que el número total de entradas, obtenido como el producto del número de variables de entrada y el número de patrones de entrenamiento, entonces existe un número de ciclos de entrenamiento que permitirá a cada neurona de la red especializarse en una entrada, reduciendo el error de aproximación a cero.

Las funciones de transferencia empleadas en este estudio son del tipo sigmoideo para la capa oculta e identidad para la capa de salida. El algoritmo de entrenamiento usado es BP, que es susceptible de bloquearse en mínimos locales, por lo que habrá que entrenar los modelos un número elevado de veces para evitar resultados sesgados por este fenómeno.

6 Rango de datos

Se estudia la influencia del rango de las variables de entrenamiento en la bondad de los resultados. Se define una red neuronal de 1 capa oculta con 4 neuronas. Se dividen las variables de entrada en dos grupos principales: el de entrenamiento y el de predicción. El grupo de entrenamiento se divide en un subgrupo correspondiente al primer 60% de los datos en que se aproximan los resultados y el restante 40% que se utiliza para validar la aproximación. El tamaño de los grupos de entrenamiento estudiados es de 12, 24 y 36 datos, correspondientes a 3, 6 y 9 meses aproximadamente. El resto de los datos se utilizan para estudiar la bondad de los resultados en predicción. Debe notarse que al tratarse de un modelo estático no tiene importancia el orden cronológico de los datos pues no se incluye el tiempo como variable de entrenamiento. Cada red es entrenada cien veces ($I = 100$) y se escoge la red con la que se obtiene el menor error relativo medio en la validación.

Se define el rango completo de cada variable según las siguientes ecuaciones:

$$R_{\delta} = \text{máx}\delta - \text{mín}\delta$$

$$R_{TMM90} = \text{máx}T_{MM90} - \text{mín}T_{MM90}$$

$$R_H = \text{máx}H - \text{mín}H$$

Análogamente, se define el rango parcial R^i de un subconjunto de entrenamiento i de cada variable de entrada y salida en las ecuaciones siguientes, en que el superíndice i indica que se trata del subconjunto de entrenamiento i y el subíndice indica la variable. En la figura 8 se explica esta nomenclatura.

$$R_{\delta}^i = \text{máx}\delta^i - \text{mín}\delta^i$$

$$R_{TMM90}^i = \text{máx}T_{MM90}^i - \text{mín}T_{MM90}^i$$

$$R_H^i = \text{máx}H^i - \text{mín}H^i$$

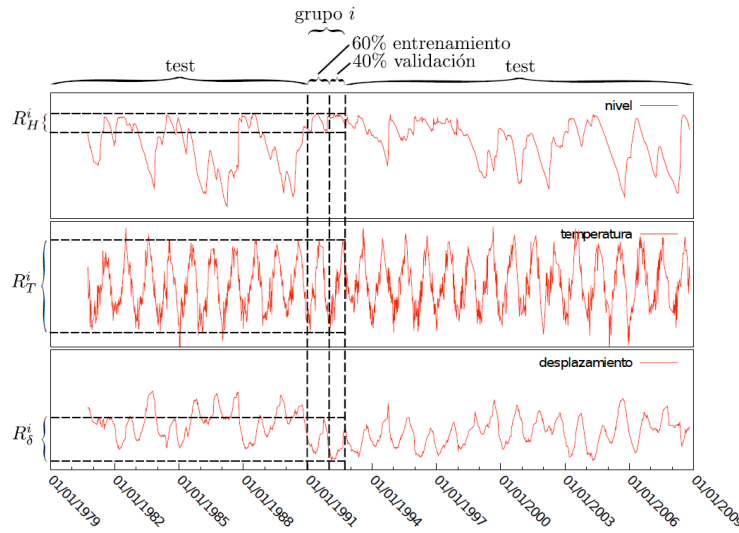


Figura 8 Ejemplo de análisis de influencia de rango de datos. El grupo de entrenamiento i tiene un rango de nivel de embalse R_H^i , de temperatura R_T^i y de desplazamiento R_δ^i . El grupo de test de la bondad del entrenamiento se compone por dos grupos, el que precede cronológicamente al grupo i y el que le sigue cronológicamente.

En las figuras 9, 10 y 11 se grafica el error relativo medio definido según la ecuación 10 para diversos tamaños del grupo de entrenamiento para las distintas variables de entrada y salida. Se diferencia entre el error relativo medio de predicción de desplazamientos al usar variables dentro y fuera del rango de entrenamiento.

Se observa que en todos los casos el error relativo medio de predicción es menor cuando las variables de entrada y salida están dentro del rango de entrenamiento; este fenómeno es más acusado en rangos estrechos de desplazamientos ($R_\delta^i / R_\delta < 0,2$) y nivel de embalse ($R_H^i / R_H < 0,2$). Por otro lado, la influencia del rango de temperaturas R_{TMM90}^i parece poco relevante en la magnitud de los errores.

$$e = \frac{1}{n} \sum_i^n \left| \frac{\hat{\delta}_i - \delta_i}{\delta_i} \right|$$

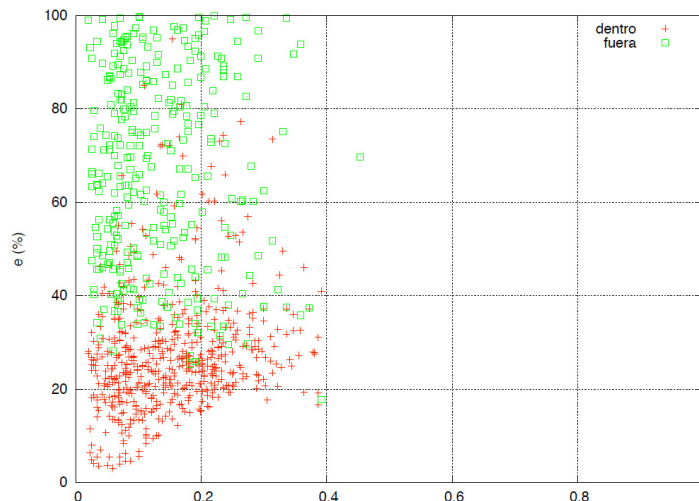


Figura 9 Variación del error según la razón del rango de desplazamientos R_δ^i / R_δ para grupos de tamaño de 3 meses aproximadamente.

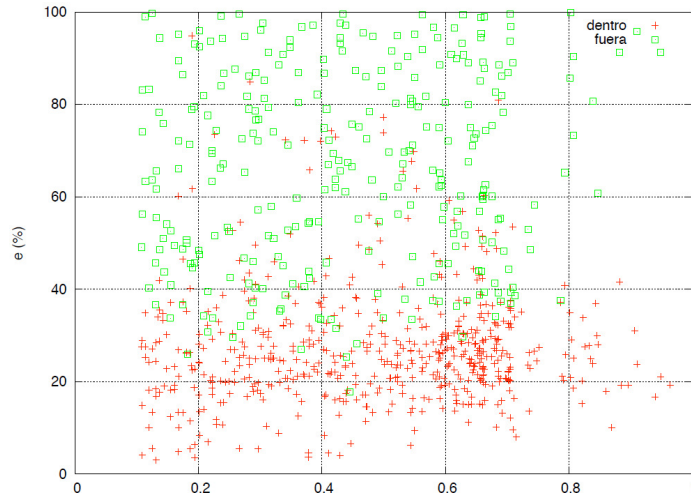


Figura 10 Variación del error según la razón del rango de temperaturas R_{TMM90}^i / R_{TMM90} para grupos de tamaño de 3 meses aproximadamente.

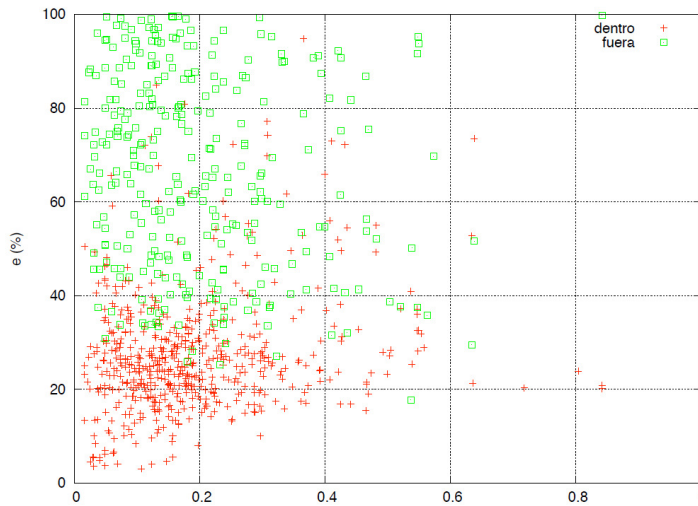


Figura 11 Variación del error según la razón del rango de niveles R_H^i / R_H para grupos de tamaño de 3 meses aproximadamente.

7 Conclusiones

Existe una clara dependencia entre la magnitud del error de predicción y el rango de los datos de entrada y de salida. Los errores de predicción son menores cuando se usa la red para predecir con valores de entrada y salida con los que ha sido entrenada, independientemente de la amplitud del rango. Los mejores resultados de predicción se obtienen en rangos estrechos de desplazamientos y de niveles de embalse.

Lo anterior sugiere que, los resultados obtenidos con redes neuronales son más fiables cuando los valores de las variables de entrada y de salida están dentro del rango de valores con los que han sido entrenadas.

8 Agradecimientos

Agradecemos la financiación del Ministerio de Medio Ambiente y Medio Rural y Marino del Gobierno de España a través del proyecto de investigación Estudio de la Seguridad de Presas e Identificación de Escenarios de Riesgo mediante Sistemas Inteligentes (SEPRISIS) con código 048/RN08/04.5.

Los autores también quieren expresar su agradecimiento a la empresa Ofiteco y a la Agencia Catalana del Agua por los datos proporcionados de la presa de La Baells.

9 Referencias

1. M.Y. Rafiq, G. Bugmann and D.J. Easterbrook. Neural network design for engineering applications. *Computers and Structures*, 79:1541-1552, 2001.
2. Hojjat Adeli. Neural networks in civil engineering: 1989-2000. *Computer-Aided Civil and Infrastructure Engineering*, 16:126-142, 2001.
3. Ian Flood and Nabil Kartam. Neural networks in civil engineering I: principles and understanding. *Journal of Computing in Civil Engineering*, 8(2):131-148, 1994.
4. G. Habibagahi. Reservoir induced earthquakes analyzed via radial basis function networks. *Soil Dynamics and earthquake Engineering*, 17:5356, 1998.
5. S. Barai, P. Pandey. Performance of the generalized delta rule in structural damage detection. *Engineering Applications of Artificial Intelligence*, 8:211221, 1995.
6. M. Meireles, P. Almeida, M. Godoy. A comprehensive review for industrial applicability of artificial neural networks. *IEEE Transactions on Industrial Electronics*, 50:585601, 2003.
7. B. Ahmadi, L. Chouinard. Use of artificial neural networks for real time analysis of dam monitoring data. *Annual Conference of the Canadian Society for Civil Engineering*, 2003.
8. G. Tayfur, D. Swiatek. Case study: Finite element method and artificial neural network model for flow through jeriorsko earthfill dam in poland. *Journal of Hydraulic Engineering*, 131:431440, 2005.
9. P. Salvaneschi, M. Cadei, M.Lazzari. Applying ai to structural safety monitoring and evaluation. *AI in Civil and Structural Engineering*, 2434, 1996.