

## Aplicación de Redes Neuronales en la Detección de Regímenes Degradados en el Proceso Wedm

E. Portillo, I. Cabanes, M. Marcos, A. Zubizarreta

Departamento de Ingeniería de Sistemas y Automática,  
Universidad del País Vasco, C/ Alameda Urquijo s/n, 48013, Bilbao, España (e-mail: [marga.marcos@ehu.es](mailto:marga.marcos@ehu.es))

Resumen: Este artículo presenta los resultados de un estudio comparativo de distintas configuraciones de redes neuronales aplicadas al proceso de corte por electroerosión por hilo (WEDM). El objetivo perseguido es detectar con antelación comportamientos de corte degradado que alertan del riesgo creciente de la rotura de la herramienta empleada en este proceso de mecanizado: el hilo. Cuando esto sucede, disminuye la productividad de manera significativa. Así, partiendo de un trabajo previo en el que se identificaron diferentes tipos de comportamientos degradados, se ha realizado un estudio comparativo contemplando distintos criterios. Entre ellos, destaca la comparación de arquitecturas clásicas de red y, más concretamente, la arquitectura estática Perceptrón Multicapa, y la arquitectura recurrente Elman. La conclusión del trabajo ha sido que la arquitectura Elman constituye la alternativa más adecuada para la detección de la degradación del proceso. Copyright © 2009 CEA.

Palabras Clave: WEDM, electroerosión, RNA, redes neuronales artificiales, Perceptrón Multicapa, Elman.

### 1. INTRODUCCIÓN

El proceso de corte por electroerosión por hilo (WEDM) es uno de los procesos de mecanizado no convencional más utilizado en la actualidad. Este proceso se utiliza en la fabricación de utillajes para troquelaría que, por su dureza o por sus características mecánicas, resultan difíciles de mecanizar por medios convencionales.

El proceso de electroerosión consiste en el arranque de material originado por sucesivas descargas eléctricas entre la pieza a mecanizar y el hilo (ambos electrodos conductores) sin que se produzca contacto entre ellos. Este espacio de separación entre electrodos, en el cual se producen las sucesivas descargas, se denomina *gap*. Tras la ionización del *gap* se produce la descarga en presencia de fluido dieléctrico. Un generador mantiene la diferencia de potencial o tensión de vacío ( $U_i$ ) necesaria en el *gap* hasta que este es ionizado por completo, esto es, hasta que transcurre el tiempo de ionización  $t_d$ . Una vez transcurrido este tiempo, el generador se descarga dando lugar a una chispa capaz de erosionar material de la pieza. Antes de repetir este ciclo para generar la siguiente descarga, se deja transcurrir un tiempo que se emplea principalmente para refrigerar y limpiar el canal de descarga del material arrancado durante el ciclo anterior. Este tiempo se conoce como tiempo de pausa ( $t_{off}$ ). La duración de cada descarga es del orden de microsegundos (entre 3 y 8). En la Figura 1 se muestra un esquema del proceso WEDM.

Uno de los problemas fundamentales del corte por electroerosión por hilo consiste en la disminución del rendimiento del proceso debida a la aparición de perturbaciones que pueden acabar produciendo la rotura del hilo (Ho, *et al.*, 2004). La dificultad en el estudio y mejora del proceso WEDM radica en que se trata de un proceso estocástico y no lineal en el que entran en juego múltiples parámetros de mecanizado.

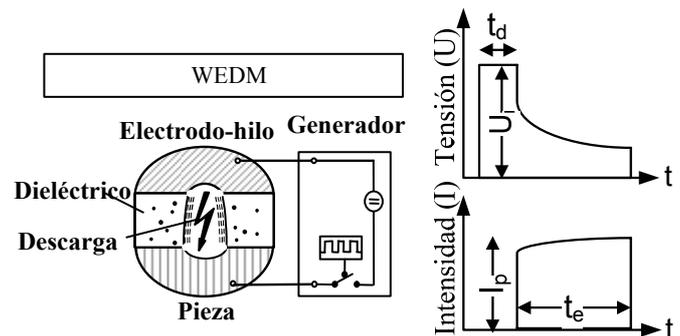


Figura 1. Esquema del proceso WEDM

Dada su característica fuertemente no lineal y estocástica, la descripción de modelos matemáticos que representen su comportamiento no es una tarea trivial. Es por ello que la alternativa de aplicación de técnicas de control inteligente, que generalmente se basan en utilizar el conocimiento empírico que el experto tiene sobre el proceso, constituye una opción claramente aplicable. Entre todas las técnicas que se engloban en el término *control inteligente*, en este trabajo se tratan las técnicas heurísticas y las técnicas basadas en redes neuronales.

Las primeras han sido tradicionalmente empleadas en el proceso WEDM. El denominador común de los trabajos realizados en este campo es la captura de las señales tensión e intensidad de las descargas, debido a que dichas señales aportan información de relevancia sobre la estabilidad del *gap*.

Una de las principales líneas de investigación se basa en la detección de la posición de cada descarga en el hilo, ya que la concentración de sucesivas descargas en las vecindades de un punto puede originar la rotura del hilo (Kunieda, *et al.*, 2001; Lauwers, *et al.*, 1998; Obara, *et al.*, 1999; Yan y Liao, 1995;

Shoda, *et al.*, 1995). Para ello los autores definen circuitos de detección que implementan en sistemas hardware propietarios. La mayoría de estos trabajos se basan en modelos matemáticos de la temperatura en el hilo. Sin embargo, el proceso WEDM se caracteriza por una naturaleza fuertemente estocástica cuyas características dependen considerablemente de los parámetros de corte. Es por ello que se considera que el modelado adecuado de este proceso es difícil si no imposible (Liao *et al.*, 2000).

Otra de las líneas considera la clasificación de las descargas en diferentes tipos (Liao, *et al.*, 1997a; Liao, *et al.*, 1997b; Wu y Li, 2001). La clasificación más comúnmente utilizada diferencia entre un mínimo de 2 y un máximo de 4 tipos de pulsos que se citan a continuación: descarga normal, arco, cortocircuito y descarga vacía. Cada tipo de descarga se establece definiendo umbrales en las señales tensión y/o intensidad. Para conocer el grado de estabilidad del proceso, los autores desarrollan sistemas hardware propietarios que permiten contabilizar cada tipo de descarga en un intervalo de registro predeterminado. Así, en los trabajos de Liao, *et al.* la clasificación se estableció únicamente a través de la tensión de las descargas. Dichos trabajos se basaron en la identificación de dos tipos de fenómeno de rotura: rotura por subida brusca de la frecuencia total de descarga, y rotura por subida paulatina de la frecuencia total de descarga e incremento del ratio de arcos y cortocircuitos.

La implementación hardware de las reglas heurísticas permite detectar comportamientos degradados de forma rápida, lo que se adapta muy bien a la elevada dinámica del proceso WEDM. Sin embargo, la definición de un conjunto de reglas acotado con una capacidad de generalización satisfactoria para todas las casuísticas es un objetivo difícil de alcanzar. Este hecho es más significativo si el estudio se realiza sobre máquinas con generador de relajación y si se contemplan distintos espesores de pieza.

Las técnicas basadas en redes neuronales han sido aplicadas en múltiples trabajos con distintos objetivos (Valverde, *et al.*, 2007). Así, también han sido aplicadas en el proceso WEDM (y en otros procesos de mecanizado no convencionales: Electroerosión por Penetración -SEDM, Mecanizado Electroquímico -ECM, etc). En concreto, en la bibliografía se pueden distinguir tres áreas principales de aplicación: determinación de los parámetros óptimos de mecanizado; predicción, mantenimiento y control de procesos de mecanizado no convencional así como la clasificación de descargas.

En relación con la primera de ellas, Tarng, *et al.* (1995) diseñaron una red de tipo *feedforward* (con conexiones hacia delante) para modelar las relaciones entre los distintos parámetros que intervienen en el proceso. Wang (2004) propuso una red neuronal *feedforward* híbrida de dos fases para determinar los parámetros óptimos de corte para espesores desconocidos en el proceso WEDM cuando se dispone de pocos ejemplos. Fenggou y Dayong (2004) han empleado técnicas de redes neuronales y algoritmos genéticos para la determinación de los parámetros óptimos de mecanizado en electroerosión por penetración (SEDM). Concretamente, se ha aplicado para obtener el parámetro de intensidad en régimen de desbaste.

La segunda área de aplicación de redes neuronales se encuentra estrechamente relacionada con la predicción, diagnóstico y control de procesos de mecanizado no convencional. Tsai y Wang (2001) diseñaron una red neuronal para predecir el

acabado superficial de las piezas en SEDM. Para ello, los autores compararon la eficacia de distintos modelos de redes neuronales estáticas utilizando el entorno Matlab<sup>TM</sup>. Además, también evaluaron distintos algoritmos de entrenamiento. Huang y Liao (2000) utilizaron técnicas de redes neuronales para desarrollar un sistema de detección de fallos fuera de línea para mantenimiento de máquinas WEDM. Concretamente, la red se trataba de una Perceptrón Multicapa entrenada mediante el algoritmo de retropropagación. Yan, *et al.* (2001) propusieron un sistema de control cuyo objetivo es seleccionar los parámetros de corte en función del espesor de la pieza. Concretamente, para la estimación de la altura de la pieza utilizaron una red tipo *feedforward*. Recientemente Behrens y Ginzl (2003) han desarrollado un sistema de control de la anchura del gap a través de una red neuronal en el proceso SEDM. De nuevo, la arquitectura de red seleccionada es la red Perceptrón Multicapa.

Por último, se han utilizado redes neuronales para clasificar los distintos tipos de descargas. Kao y Tarng (1997) propusieron un sistema de clasificación basado en redes neuronales que opera en línea. El tipo de red empleado es nuevamente Perceptrón Multicapa con algoritmo de aprendizaje de retropropagación. Mediliyegedara, *et al.* (2004) han desarrollado una red neuronal para la clasificación de pulsos en ECM.

El análisis del estado del arte revela que la aplicación de técnicas basadas en redes neuronales, concretamente, de redes estáticas, ha sido foco de interés en los procesos de mecanizado no convencional. En el caso del proceso WEDM, estas técnicas resultan apropiadas para dotar de mayor generalidad a la detección anticipada de la degradación del proceso de corte. Sin embargo, hasta donde los autores conocen, las redes neuronales no han sido utilizadas para prever la rotura del hilo, tal y como se plantea en el trabajo que aquí se presenta. Además, todos los trabajos referenciados utilizan, para lograr sus objetivos, redes estáticas y, en la mayoría de ellos, la arquitectura Perceptrón Multicapa. Sin embargo, no se han encontrado referencias al uso de redes recurrentes, que en principio parecen ser adecuadas para el proceso. En este trabajo se presenta un estudio comparativo entre diferentes configuraciones de redes estáticas y recurrentes cuyo objetivo es detectar la degradación del comportamiento del proceso que, finalmente, llevará a la rotura del hilo.

Tras este apartado de introducción al proceso WEDM y a la aplicación de las técnicas de redes neuronales en el mismo, en la sección 2 se describen brevemente trabajos previos de los autores sobre la detección de situaciones degradadas, de los que se parte en este artículo. En la sección 3 se muestra la estrategia seguida para el diseño de redes neuronales que permita detectar regímenes de corte degradado antes de que se produzca la rotura del hilo. En la sección 4 se analizan los resultados del estudio comparativo, seleccionando la configuración de red más apropiada para abordar los objetivos planteados. Por último, en la sección 5 se resumen las conclusiones del trabajo.

## 2. IDENTIFICACIÓN DE REGÍMENES DE CORTE DEGRADADO

Trabajos previos de los autores propusieron una metodología (resumida de forma gráfica en la Figura 2) para detectar regímenes de corte degradado y así anticipar la rotura del hilo (Cabanes, *et al.*, 2007). En primer lugar, se obtuvo una extensa base de datos experimental mediante un sistema de adquisición que permitía capturar las señales básicas del proceso: tensión e

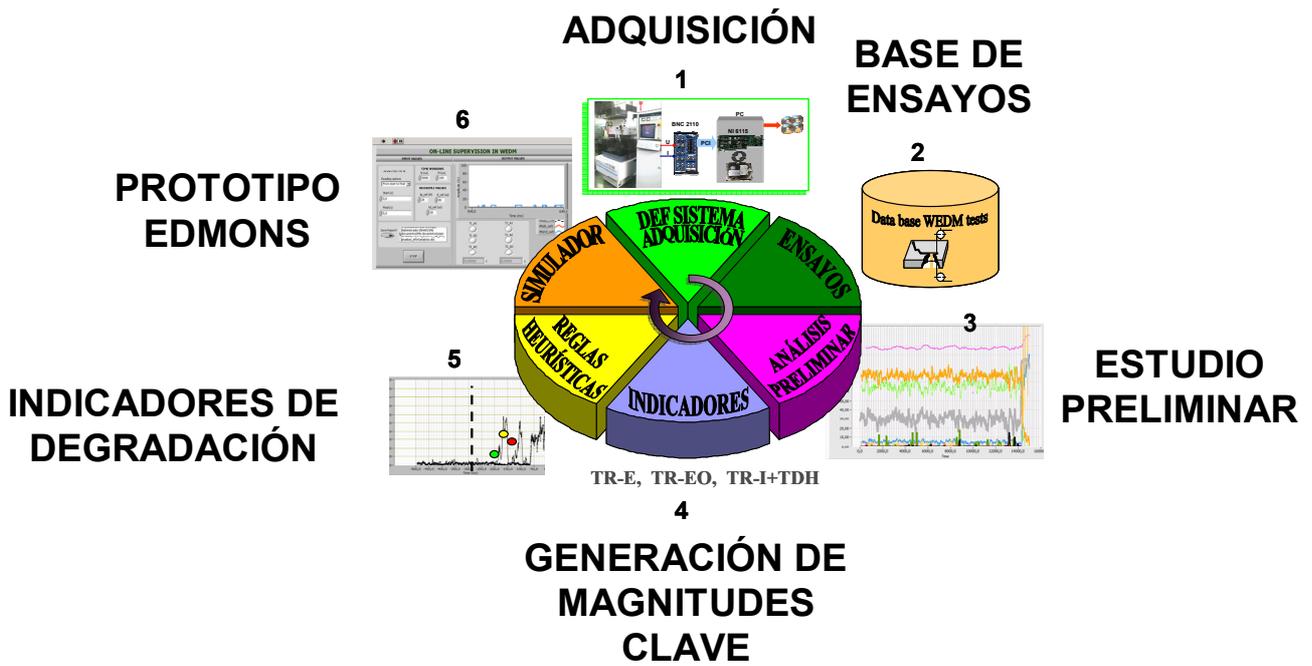


Figura 2. Metodología para la detección de la degradación del proceso WEDM

intensidad de las descargas. Las señales se adquieren a una frecuencia de 5 MHz y se almacenan en disco sin pérdida de información (Portillo, *et al.*, 2007). La batería de ensayos contempla tanto el corte en régimen estable como regímenes de corte degradado provocados por diferentes causas: disminución de la presión del dieléctrico, corte de geometrías complejas e incremento de la frecuencia de las descargas.

Los resultados del análisis preliminar de los experimentos mostraron tres variables clave para la detección de la rotura: la energía de las descargas  $E_i$ , la intensidad de pico  $I_{pi}$  y el tiempo de ionización  $t_{di}$  (ver Figura 3). La energía de una descarga  $E_i$  se define como:

$$E_i = \int_0^{t_{e_i}} I_{p_i} U_i dt \quad (1)$$

Durante el análisis se observó que una única descarga es insuficiente para conocer la calidad del régimen de corte. Por el contrario, un conjunto de descargas sucesivas con un comportamiento determinado sí puede informar sobre la degradación del proceso que, en última instancia, llevará a la rotura de la herramienta.

Así, mediante el preprocesamiento de las señales básicas del proceso se definieron una serie de indicadores o *medidas virtuales (MV)* relacionadas con estas tres variables clave: medida virtual de intensidad MV-I, medida virtual de energía MV-E y medida virtual de tiempos de ionización elevados MV-TDH (Portillo, *et al.*, 2007).

Las medidas virtuales se definen como sucesivos porcentajes de descargas que superan respectivamente una serie de valores de referencia en una ventana de tiempo deslizando representativa N. Los valores de referencia (VR) se obtienen a partir de los

ensayos en régimen de corte estable. Cada valor de referencia se ha definido como el valor de la clase cuando la frecuencia acumulada relativa alcanza el 99 % en corte estable.

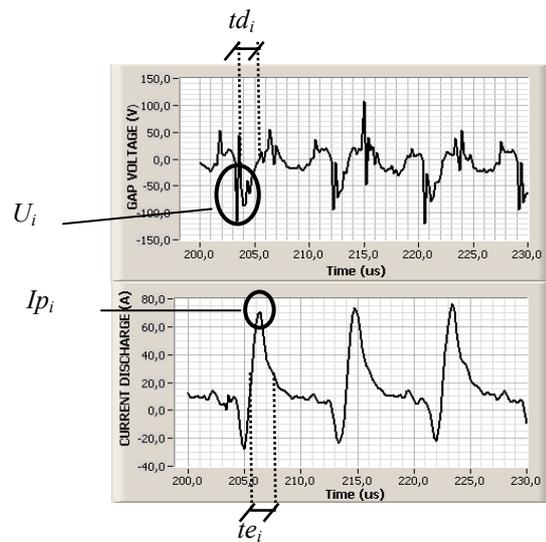


Figura 3. Variables clave de las descargas en el proceso WEDM

Así, el algoritmo genérico que define la evolución de cada medida virtual es el siguiente:

- 1) Por cada ventana básica de tamaño M, se realiza el cálculo del número total de descargas  $M_T$  y del número de descargas cuyas variables básicas BV ( $I_{pi}$ ,  $E_i$  or  $t_{di}$ ) exceden el respectivo valor de referencia RV: Así, para  $MV(I_p)$ ,  $MV(E)$ ,  $MV(t_{th})$ :

$$M(BV) = \sum_{j=1}^{M_T} p_j ; p = \begin{cases} 1 \rightarrow \text{if } BV \geq RV \\ 0 \rightarrow \text{if } BV < RV \end{cases} \quad (2)$$

2) En la primera ventana deslizante de tamaño  $N$ , se realiza el cálculo del número total de descargas  $N_T$  y del número de descargas cuyas variables básicas  $BV$  ( $I_{pi}$ ,  $E_i$  or  $t_{di}$ ) exceden (o son menores que) el respectivo valor de referencia  $RV$ :

$$N_{BV} = \sum_{j=1}^{N/M} M(BV)_j \quad (3)$$

$$N_T = \sum_{j=1}^{N/M} M_{T_j} \quad (4)$$

3) En las siguientes ventanas deslizantes de tamaño  $N$ , se realiza el cálculo deslizante:

$$N_{BV_{j+1}} = N_{BV_j} + M(BV)_{j+1} - M(BV)_{j+1-N} \quad (5)$$

$$N_{T_{j+1}} = N_{T_j} + M_{T_{j+1}} - M_{T_{j+1-N}} \quad (6)$$

$$MV_{j+1} = \frac{N_{BV_{j+1}}}{N_{T_{j+1}}} \quad (7)$$

Donde  $j=1, 2, \dots$

Mediante el análisis del comportamiento de las medidas virtuales, se identificaron tres tipos de comportamiento degradado (Cabanes, *et al.*, 2008):

*Incremento brusco de la energía (TR-E).* La Figura 4 muestra un ejemplo de comportamiento de la medida virtual de la energía  $MV-E$  un segundo antes de producirse la rotura de la herramienta.

*Oscilación de la energía (TR-EO).* La Figura 5 presenta un ejemplo de este tipo de rotura. Se observa un comportamiento oscilatorio de la medida virtual de la energía de aproximadamente 0,3 segundos antes de la rotura.

*Sucesivos picos de intensidad junto con un tiempo de ionización elevado (TR-I+TDH).* De nuevo, se muestra un ejemplo del comportamiento de las medidas virtuales relacionadas un segundo antes de la rotura (ver Figura 6).

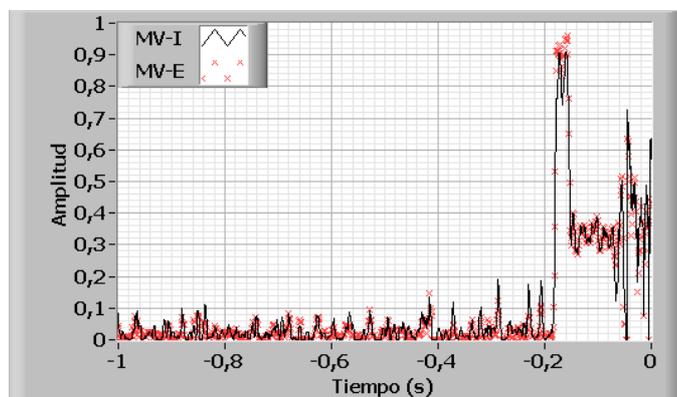


Figura 4. Ejemplo de rotura TR-E, 50 mm.

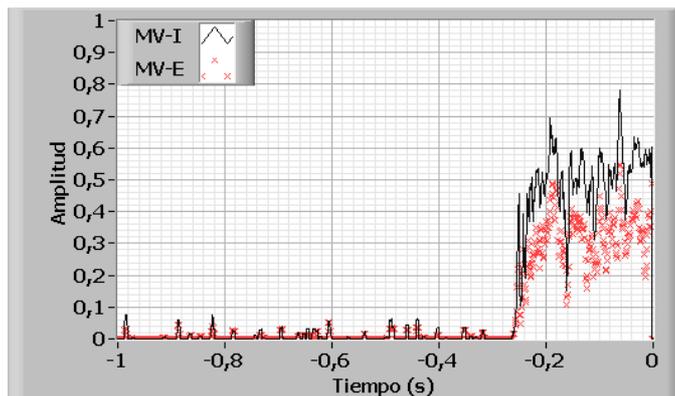


Figura 5. Ejemplo de rotura TR-EO, 50 mm.

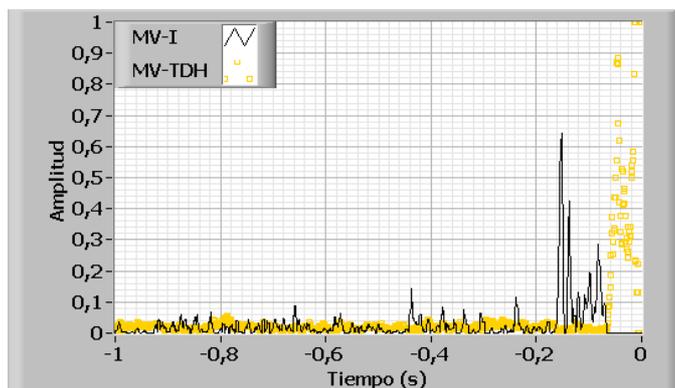


Figura 6. Ejemplo de rotura TR-I+TDH, 50 mm.

### 3. ESTRATEGIA BASADA EN REDES NEURONALES PARA LA DETECCIÓN DE LA DEGRADACIÓN DEL PROCESO WEDM

El objetivo de este trabajo es más ambicioso ya que, partiendo de la información de las medidas virtuales se pretende: detectar regímenes de corte degradado con antelación suficiente a la rotura del hilo como para aplicar estrategias que la eviten; además, se pretende estimar el grado de influencia de cada uno de los comportamientos degradados en la degradación del proceso. De esta forma, la estrategia de actuación puede ser diferente en función del tipo de situación degradada de mayor influencia. Así por ejemplo, dada la inminencia de la rotura en el tipo de comportamiento degradado TR-E, la estrategia que se presume como más adecuada estaría basada en la desconexión controlada del generador. Sin embargo, en el caso denominado TR-I+TDH, la estrategia aplicada estaría fundamentalmente basada en el control del gap.

#### 3.1 Entradas y salidas de la configuración de red

Las entradas son las medidas virtuales de intensidad  $MV-I$ , energía  $MV-E$  y de valores elevados de tiempo de ionización  $MV-TDH$ . Las salidas son neuronas destinadas a la detección del régimen de corte degradado, caracterizado en diferentes niveles de alarma: nivel bajo, A1, nivel medio, A2 y nivel alto, A3. Asimismo, se determina el grado de influencia de cada tipo de comportamiento degradado. Esto se ilustra en la Figura 7.

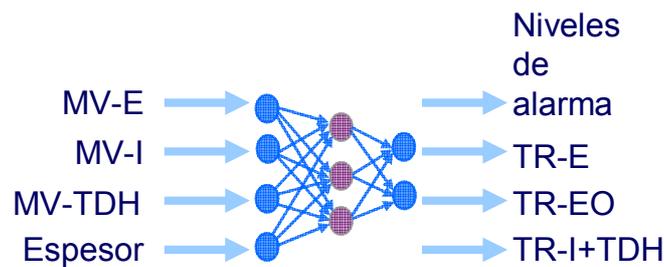


Figura 7. Entradas y salidas de las redes neuronales

### 3.2 Estudio comparativo

Para alcanzar los objetivos establecidos, el enfoque adoptado consiste en diseñar redes neuronales a partir del conocimiento extraído durante el análisis de los comportamientos degradados. Por tanto, se trata de aplicar aprendizaje supervisado.

Además, se apuesta por utilizar arquitecturas clásicas de redes neuronales que se encuentren disponibles en herramientas comerciales, como es el caso de la utilizada en este trabajo, Matlab<sup>TM</sup> (toolbox de redes neuronales). Así, el objetivo abordado no es proponer una nueva arquitectura de red neuronal, sino aplicar arquitecturas clásicas cuya validez haya sido ampliamente demostrada en aplicaciones industriales (siendo algunos ejemplos los mostrados en la introducción), y que además den solución a la problemática aquí planteada.

Así, para poder obtener aquella estructura neuronal que mejor responda a las necesidades planteadas, se ha realizado un estudio comparativo en función de distintos criterios que se presentan a continuación:

#### Arquitectura de red

Una de las decisiones más trascendentales es la selección de la arquitectura de red neuronal. Como se ha comentado en el apartado de la introducción, en la totalidad de los trabajos encontrados en el campo del mecanizado no convencional se han utilizado redes estáticas. Además, cabe destacar que la inmensa mayoría de los casos se han basado en el modelo Perceptron Multicapa. Así, esta arquitectura presenta la ventaja de que es la arquitectura de red en la que existe mayor experiencia en su aplicación industrial. Es por ello que una de las arquitecturas consideradas durante el estudio comparativo es la Perceptron Multicapa.

Además de la arquitectura estática, también se analiza uno de los modelos recurrentes más extendidos y que se encuentra disponible en Matlab<sup>TM</sup>: la red Elman (1990). Se trata de un tipo de red con conexiones hacia adelante similar a la Perceptron Multicapa pero que además presenta una conexión recurrente desde la salida de la capa oculta en el instante  $t$  a la entrada de la capa oculta en el instante  $t+1$ . Para ello se incorporan en la capa de entrada un tipo de neuronas especiales conocidas como neuronas de contexto que son receptoras de las conexiones recurrentes. La conexión recurrente dota a la red de una memoria exponencial (Elman, 1990).

En este tipo de redes existe claramente una menor experiencia, resultando más dificultoso su entrenamiento (definición de los valores deseados, etc). Sin embargo, es de prever que esta arquitectura responda mejor a la naturaleza dinámica de las

medidas virtuales. Es más, las tendencias hacia la rotura previamente identificadas ponen de manifiesto la necesidad de recordar estados anteriores para detectar regímenes de corte degradado.

Otra ventaja de las redes recurrentes es que permiten reducir drásticamente el tamaño de la red ya que las series temporales de las medidas virtuales se pueden introducir punto a punto a lo largo del tiempo, sin necesidad de introducir de manera simultánea todo el histórico de datos necesario para detectar la rotura del hilo. Por tanto, es de esperar que se pueda incorporar con mayor facilidad en un sistema empotrado tanto desde el punto de vista de requisitos temporales como de requisitos de memoria.

#### Escalado de las entradas y salidas de la red

El escalado o estandarización de las entradas y salidas resulta una decisión no trivial que puede condicionar de manera importante el éxito del entrenamiento (Sarle, 2002). En el caso de las entradas, su importancia depende en gran medida del algoritmo de entrenamiento utilizado. Así, el algoritmo de retropropagación es muy sensible al escalado de las entradas, mientras que algoritmos como Gauss-Newton son teóricamente independientes (siempre y cuando se utilice correctamente y no se aplique un escalado demasiado extremo).

En primer lugar, es necesaria la estandarización de las entradas para igualar su importancia relativa y evitar que los valores iniciales de los pesos dependan de la escala de las variables de entrada. Así se reduce la probabilidad de saturación que, dado el carácter asintótico de las funciones sigmoideas, aparece cuando la entrada total a una neurona toma valores absolutos muy elevados. Este fenómeno da lugar a que los parámetros de la red permanezcan invariables y puede producir la finalización prematura del entrenamiento. Por esta misma razón también es preferible utilizar rangos cuyos extremos sean inferiores a los de las funciones sigmoideas. A priori, la aplicación de un rango simétrico en la arquitectura de red Perceptron Multicapa suele dar lugar a buenos valores iniciales aleatorios de los pesos (Sarle, 2002).

Considerando todas estas observaciones, el estudio comparativo que se presenta, contempla las variables escaladas en distintos rangos [L-H]. En primer lugar, se comparan los resultados proporcionados por el rango simétrico [-1-1] y el rango asimétrico [0-1], seleccionándose uno de los enfoques (simétrico o asimétrico). Una vez elegido, se aplican rangos cuyos extremos sean inferiores a los de las funciones sigmoideas ([-0,95-0,95], [-0,9-0,9], [-0,85-0,85] en el caso simétrico, y [0,05-0,95], [0,1-0,9], [0,15-0,85] en el caso asimétrico).

Existen otros tipos de estandarizaciones, siendo una de las más extendidas la normalización estándar con media cero y desviación típica. Sin embargo, de cara a una implementación en tiempo real, el cálculo de la media y de la desviación típica de las entradas supone una mayor carga computacional. Como se ha incidido en la introducción, este aspecto es fundamental en el proceso WEDM dado que la duración de las descargas se encuentra entre 3 y 8 microsegundos.

#### Agrupación de las entradas

Este criterio hace referencia a si se define una única red neuronal

con todas las medidas virtuales como entradas (configuración única) o si por el contrario, se definen varias redes neuronales con una o varias medidas virtuales como entradas (configuración paralela). La configuración única solamente se ha valorado en el caso de las redes recurrentes ya que en el caso de las redes estáticas, el tamaño de la red sería excesivamente grande.

#### Codificación de los valores deseados (*targets*)

En este punto se trata la codificación de los valores deseados que permita expresar las salidas de la red, esto es, los niveles de alarma y los tipos de comportamiento degradado.

La codificación de las salidas depende tanto de la arquitectura de red (estática-PM-, o recurrente-Elman-) como de la configuración adoptada (única o paralela).

En el caso de las redes Perceptron Multicapa, existen trabajos que recomiendan la codificación *1 de (C-1)*, es decir, una codificación binaria en la que se activa simultáneamente una única neurona de salida por todas las categorías menos una (Sarle, 2002). Esta última categoría corresponde a la situación en la que no se produce ningún nivel de alarma y se codifica haciendo que todas las salidas adquieran un nivel bajo. En general, cada valor de la neurona de salida adquiere un *nivel bajo o alto* dependiendo del ejemplo de entrenamiento introducido. El nivel bajo o alto se aplica en función del rango de escalado.

En el caso de las redes recurrentes, la codificación de los valores deseados se ha definido mediante tres algoritmos: IAPA-1, IAPA-2 e IAPA-3 (IAPA: Integración Acumulativa con Pendientes Asimétricas). Estos algoritmos se han desarrollado con el objetivo de destacar distintas características de las medidas virtuales en función del tipo de comportamiento degradado. La aplicación de uno u otro algoritmo dependerá de la agrupación de las entradas. Así, los *targets* o valores deseados de la red se obtienen aplicando alguno de los algoritmos a la medida o medidas virtuales que corresponda.

El primero de estos algoritmos desarrollado, IAPA-1, consiste básicamente en amplificar el valor de la neurona de salida correspondiente cuando la pendiente de la medida virtual prefiltrada es creciente. Por el contrario, si la pendiente es decreciente, los valores de las neuronas de salida disminuyen, aunque en menor medida. Esto facilita que las neuronas de salida adquieran valores elevados, no solo cuando se producen incrementos “instantáneos” de las medidas virtuales, sino también cuando se produce una oscilación. En la Figura 8 se muestra un ejemplo ilustrativo en el que se aplica el algoritmo IAPA-1 a una hipotética medida virtual.

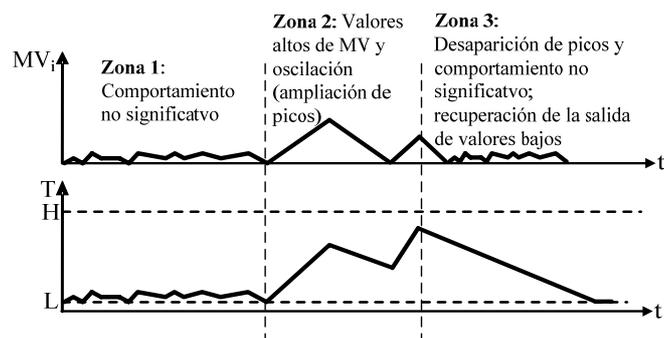


Figura 8. Aplicación de algoritmo IAPA-1

Zona 1: en la zona 1 no se observan variaciones significativas en la medida virtual, por lo que se concluye que no se ha producido la degradación del proceso y por ello únicamente se aplica el valor de escala de la Ecuación 8. Este es el único procesamiento realizado si no se detecta en la medida virtual un nivel mayor que el especificado mediante el parámetro *nivel*, que habitualmente toma el valor de 0,15 (15% del rango de valores de las medidas virtuales).

$$T_i = MV(BV)_i \cdot (H - L) + L \text{ desde } i = 1; 2 \dots \quad (8)$$

Zona 2 y zona 3: en estas zonas, se aplica el algoritmo de generación propiamente dicho. El algoritmo consta de los siguientes pasos:

- Filtrado de las medidas virtuales: el filtrado consiste en extraer las pendientes dominantes de las formas de onda de las medidas virtuales. Para considerar una pendiente significativa, y no filtrarla, la medida virtual debe permanecer con la misma tendencia (positiva, negativa o nula) durante al menos el número de puntos determinados con el parámetro *Filtro* (cuyo valor suele ser 3), o bien provocar una diferencia de valores entre un punto y el siguiente de un porcentaje de la amplitud máxima de la señal. Esto último se expresa mediante el parámetro *Límite*, cuyo valor se fija normalmente en un 10%. Si alguna de estas reglas no se cumple, se realiza la interpolación lineal entre los puntos extremos del intervalo a filtrar.
- Integración acumulativa con pendientes asimétricas: tras el filtrado de las medidas virtuales, se aplica un algoritmo que consiste en sumar o restar punto a punto los incrementos o decrementos de las pendientes de la medida virtual procesada. El resultado de esta operación es el valor del *target* en el punto correspondiente. Sin embargo, el incremento o decremento de este valor se aplica de forma asimétrica, es decir, si la medida virtual sube, el incremento depende de una ganancia  $K_s$ , mientras que si baja, depende de otra diferente  $K_b$ . Si el *target* alcanza alguno de los límites del rango  $[L-H]$ , se hace que su valor se sature a uno de dichos límites evitando así que los sobrepase. Con todo ello se consigue que el *target* refleje, no solo incrementos importantes de la medida virtual, sino también la existencia de picos sucesivos.

El algoritmo IAPA-1 se aplica a las medidas virtuales de la energía MV-E y de la intensidad MV-I<sub>p</sub> debido a la relevancia de detectar tanto incrementos bruscos como de picos sucesivos en ambas.

El algoritmo IAPA-2 parte del mismo enfoque que el IAPA-1 pero, en este caso, en vez de considerar los incrementos o decrementos punto a punto, se consideran los incrementos o decrementos del área de la medida virtual en una ventana de tiempo. Este algoritmo se aplica a las medidas virtuales asociadas al tiempo de ionización debido a su relevancia cuando permanecen elevadas a lo largo del tiempo. Así, en este algoritmo se integra MV-TDH en una ventana de tiempo y se escala el valor del área obtenida mediante un multiplicador  $K_{id}$ , cuyo cálculo se expresa en la Ecuación 9. Este se ha definido a partir de las reglas definidas en Cabanes (2008). Así, se definió la activación de la alarma A1 cuando en una ventana de tiempo, la suma de la MV-TDH supera un determinado umbral UB.

$$K_{td} = \frac{H-L}{UB} \quad (9)$$

La Figura 9 muestra un ejemplo hipotético de la aplicación el algoritmo IAPA-2.

El algoritmo IAPA-3 también parte del mismo enfoque que el IAPA-1. Se trata de un algoritmo que se adapta a los tipos de comportamiento degradado TR-I+TDH. Consiste en aplicar IAPA-1 a dos medidas virtuales distintas de manera secuencial. Así, el algoritmo IAPA-1 se aplica a la medida virtual asociada al tiempo de ionización hasta que se produce el incremento significativo de MV-I, momento en el que se pasa a aplicar el algoritmo sobre esta última. La aplicación de este algoritmo se ilustra en la Figura 10.

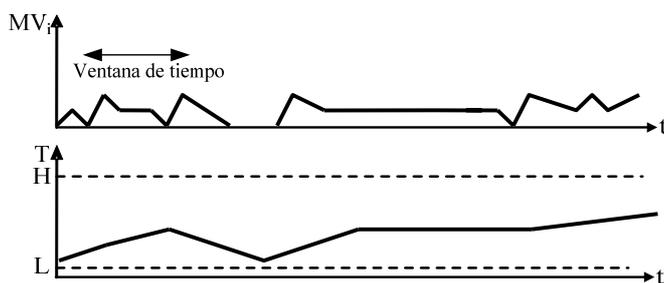


Figura 9. Aplicación de algoritmo IAPA-2

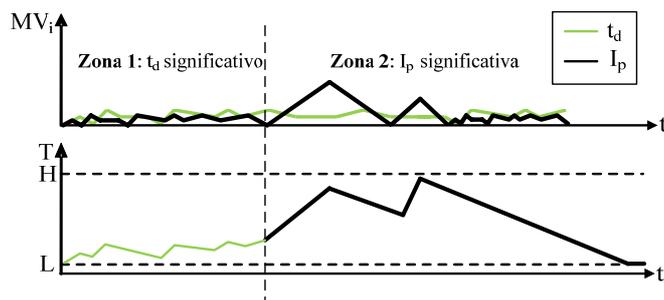


Figura 10. Aplicación de algoritmo IAPA-3

#### Entrenamiento

Uno de los aspectos más importantes es el método de generalización a utilizar en la etapa de entrenamiento. Dos métodos comúnmente conocidos son la regularización y la detención temprana. Para tomar esta decisión se han considerado las características propias en cuanto a la implementación de ambos métodos en la herramienta de redes neuronales de Matlab<sup>TM</sup>.

El algoritmo que incorpora la regularización bayesiana se denomina *trainbr*, el cual se utiliza en combinación con el algoritmo de optimización de Levenberg-Marquardt LM. El rendimiento de este algoritmo no es adecuado en el caso de reconocimiento de patrones debido a que el algoritmo LM no es tan exacto cuando se saturan las salidas de la red (Matlab, 2004). Por tanto, dado que en la configuración paralela Perceptron Multicapa se realiza un reconocimiento de patrones, se aboga por el método de detención temprana. En el caso de las configuraciones Elman, también se aplica la detención temprana, tal y como lo recomienda la herramienta (Matlab, 2004).

Así, durante el entrenamiento se ha empleado el algoritmo de aprendizaje convencional, válido tanto para el entrenamiento de la arquitectura de red Perceptrón Multicapa PM como para el de la arquitectura de red Elman. Se trata del algoritmo de retropropagación por descenso de gradiente con momento y tasa de aprendizaje adaptativa.

Este algoritmo se caracteriza por que el error es dependiente de los valores iniciales de los pesos. Por tanto, se han realizado 10 iteraciones por cada configuración de red. La función de activación elegida ha dependido del rango de las salidas: sigmoideal en el caso asimétrico, y tangente hiperbólica en el caso simétrico.

Durante el entrenamiento de las configuraciones Perceptron Multicapa se han empleado 533 ejemplos distribuidos entre las tres redes en paralelo. Respecto a las configuraciones Elman, se han utilizado 122 secuencias de 200 puntos cada una en el caso de la configuración única. En la configuración paralela, se han empleado 289 secuencias de 200 puntos distribuidas en las tres redes en paralelo. El 70% de los ejemplos se destinan al entrenamiento, y el 30% a validación. Los ejemplos se han obtenido a partir de la exhaustiva base de datos experimental obtenida en el trabajo previo. Así, las señales tensión e intensidad de las descargas, contenidas en la base de datos, son procesadas obteniendo como resultado las medidas virtuales.

Concretamente, los ejemplos de entrenamiento y validación contienen:

- Alarmas de tres niveles: nivel bajo A1, nivel medio A2 y nivel alto A3 de riesgo de rotura.
- Tipos de comportamiento degradado: incremento brusco de la energía TR-E, oscilación de la energía TR-EO, y sucesivos picos de intensidad junto con un tiempo de ionización elevado TR-I+TDH.

En (Sarle, 2002) se mantiene que un número de casos de entrenamiento igual al doble del número de pesos puede ser más que suficiente para evitar el sobreajuste. Por otro lado, en otros trabajos se concluye que, en el caso de parada del entrenamiento por detención temprana se debe utilizar una gran cantidad de neuronas ocultas (*sobredimensionamiento*) para evitar óptimos locales (Sarle, 1995), (Tetko, et al., 1995). Considerando estos aspectos, por cada arquitectura se ha valorado distinto número de neuronas y/o capas ocultas:

- Estática: 1 ó 2 capas ocultas y entre 1 y 30 neuronas por capa oculta.
- Recurrente: entre 5 y 30 neuronas en la capa oculta.

#### 4. RESULTADOS DEL ESTUDIO COMPARATIVO

En esta sección se muestran los resultados del estudio comparativo realizado. Así, durante la fase de análisis de resultados el objetivo ha sido identificar la estructura neuronal que resulte más adecuada en la detección tanto de los niveles de alarma como del tipo de comportamiento degradado. Por tanto, en este apartado, teniendo en cuenta los resultados del estudio comparativo, se aborda la preselección de la arquitectura, de la configuración, del tipo de codificación de las salidas y, por último, de algunos rangos de las variables.

Este análisis se ha realizado sobre los ensayos almacenados en la

base de datos experimental. Se trata de ensayos de corte estable y degradado realizados sobre piezas de acero de espesor 50 mm (AISI D2) en la máquina de electroerosión por hilo ONA E-250. Se ha utilizado hilo de latón de 0,25 mm de diámetro (DIN 160).

4.1 Redes estáticas: Perceptrón Multicapa (PM)

La Figura 11 muestra los resultados con menor error de validación por cada rango en el caso de una red paralela que procesa la medida virtual de energía, denominada a partir de ahora PM-E. En ella se muestra el error de validación por cada rango así como la configuración de las capas ocultas. Esta última aparece indicada sobre cada una de las barras. El error se refiere al error cuadrático medio mínimo (MSE).

Los rangos asimétricos han dado lugar a menores errores de validación (aproximadamente 4 veces inferior). La configuración que presenta el menor error de validación tiene dos capas con 20 neuronas cada una en el rango [0,15-0,85]. Sin embargo, la segunda red con menor error de validación presenta un error ligeramente superior y su tamaño es menor (dos capas con 5 neuronas respectivamente en el rango [0,1-0,9]).

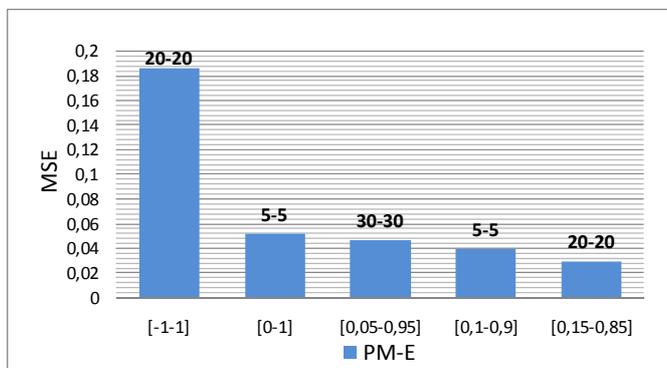


Figura 11. Menores errores de validación en PM Energía

La Figura 12 muestra los resultados con menor error de validación por cada rango en el caso de una red paralela que procesa la medida virtual de la intensidad PM-I. En este caso el menor error de validación corresponde a una configuración con dos capas ocultas de 5 neuronas cada una, en el rango [0,15-0,85]. También se observa que las redes más grandes no dan como resultado los errores de validación menores. Además, al igual que en la red PM Energía, el rango [-1-1] proporciona errores aproximadamente 4 veces mayores en comparación con el rango asimétrico [0-1].

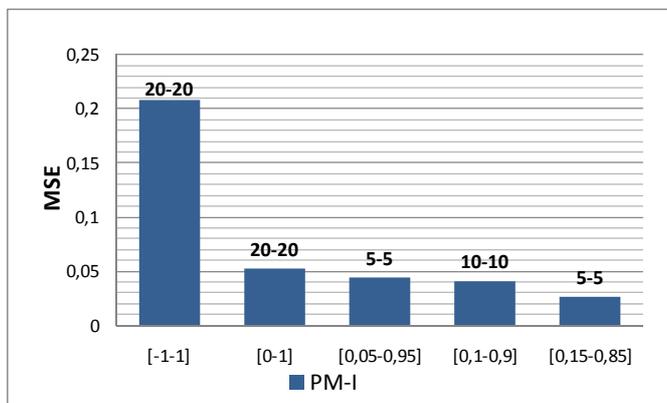


Figura 12. Menores errores de validación en PM Intensidad

En el caso de la red que procesa la medida virtual correspondiente al tiempo de ionización, se han obtenido resultados satisfactorios para varias configuraciones. Los casos con menor error de validación se han obtenido para redes de dos capas. Sin embargo, se consideran como resultados más satisfactorios los dados por otras configuraciones de red más pequeñas, que proporcionan una mejor relación tamaño de red-error de validación. Por ello son las únicas mostradas en la Figura 13. De nuevo el rango [-1-1] da errores de validación significativamente superiores a los proporcionados por el entrenamiento en el resto de rangos.

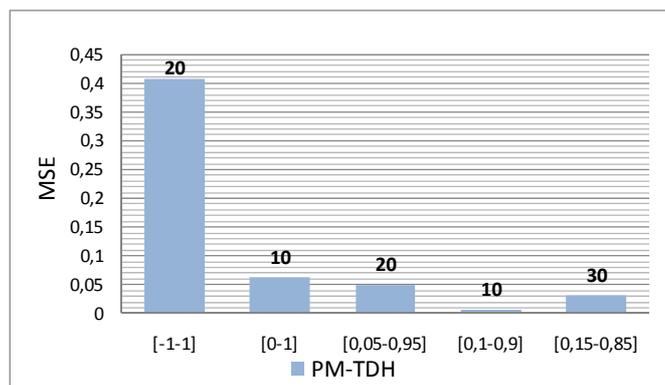


Figura 13. Menores errores de validación en PM TDH

A modo de ilustración, la Figura 14 muestra un ejemplo de validación cuyo funcionamiento es correcto. Se trata de una situación de alarma nivel alto provocada por sucesivos picos de energía. En la gráfica aparecen los patrones o ejemplos de validación. En la parte superior del marco de la figura se encuentran indicados los valores deseados o *targets* como, por ejemplo, 0,1 0,1.... Sobre la gráfica aparecen los valores de salida reales, 0,00040149 0,18802... Estos valores también se muestran en la tabla que se encuentra sobre la gráfica.

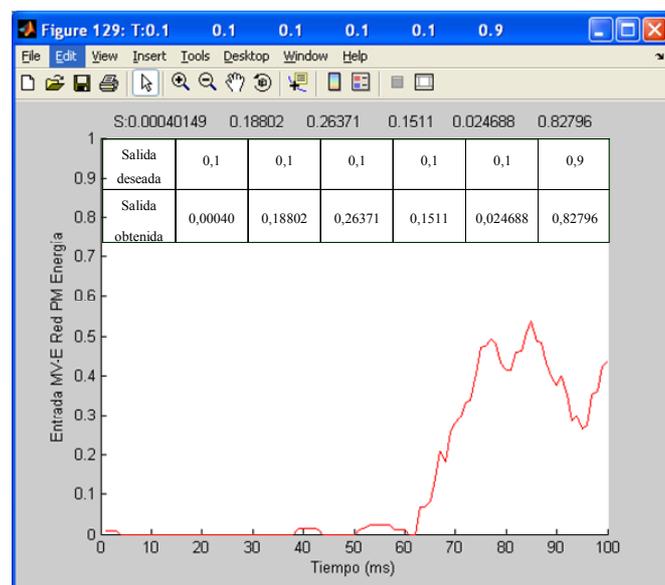


Figura 14. Ejemplo de validación con mejor caso en la alternativa PM Energía: TR-EO

4.2 Redes recurrentes: Elman (E)

La Figura 15 muestra las configuraciones de red que han dado lugar a los menores errores de validación por cada rango en la red paralela Elman de energía.

Si bien el rango [0,15-0,85] tiene el menor error de validación, el rango [0,05-0,95] presenta un error prácticamente similar y con una configuración de red más pequeña (5 neuronas en la capa oculta), por lo que esta última es la que se considera el mejor caso. Los resultados muestran que, de manera similar a los obtenidos en PM, las redes que presentan pocas neuronas ocultas (entre 5 y 10) proporcionan los menores errores de validación. De nuevo, el rango [-1-1] da lugar a errores de validación significativamente superiores.

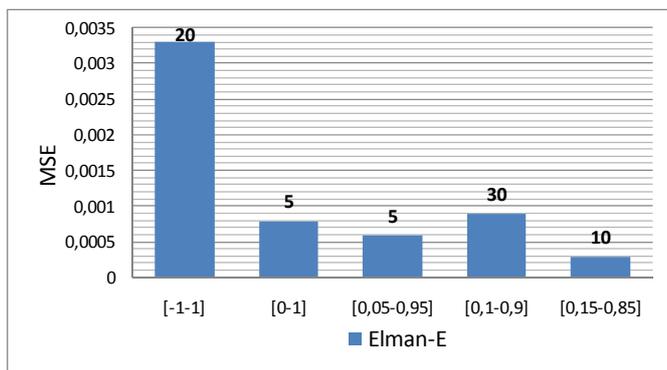


Figura 15. Menores errores de validación en Elman Energía

La Figura 16 muestra los resultados más significativos en la red paralela Elman de intensidad por cada rango evaluado. En este caso, el menor error de validación corresponde directamente a la red más pequeña (5 neuronas en la capa oculta en el rango [0,1-0,9]). De forma similar a las configuraciones anteriores, los errores de validación obtenidos con el rango simétrico son claramente superiores a los obtenidos con rangos asimétricos.

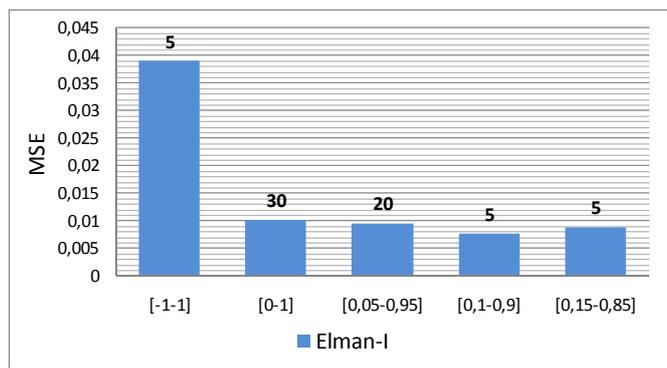


Figura 16. Menores errores de validación en Elman Intensidad

En la Figura 17 se muestran los resultados de las redes entrenadas de tipo paralelo Elman de tiempo de ionización elevado TDH. En este caso la configuración seleccionada es la correspondiente al rango [0,1-0,9] ya que es la que mejor relación tamaño/rendimiento presenta.

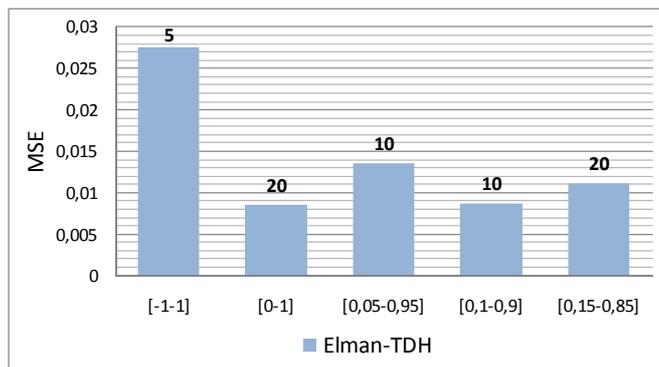


Figura 17. Menores errores de validación en Elman TDH

La Figura 18 muestra las configuraciones que presentan un menor error de validación por cada rango evaluado en la red única Elman. Esta red procesa todas las medidas virtuales simultáneamente. Debido a los malos resultados obtenidos con el rango simétrico en las configuraciones paralelas, este rango no se considera durante el entrenamiento de la configuración única Elman.

En este caso, se observa que, si bien los tipos de comportamiento degradado TR-E y TR-EO los distingue muy satisfactoriamente, el ratio de aciertos para el tipo de comportamiento degradado caracterizado por un incremento en la intensidad y en el tiempo de ionización (TR-I+TDH) es del 0%. Así, el buen comportamiento de los dos primeros tipos disminuye considerablemente la media del error, ocultando un funcionamiento inaceptable de la red en uno de los casos.

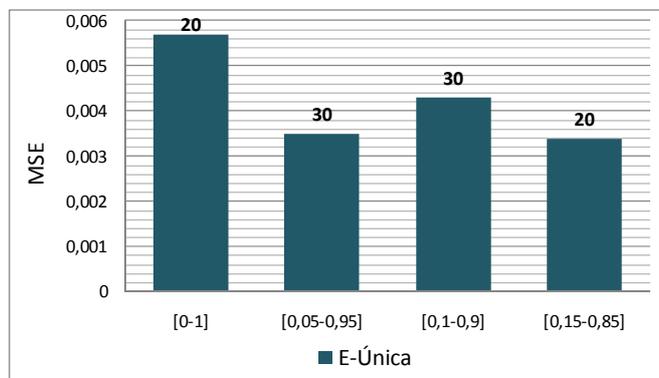


Figura 18. Menores errores de validación en Elman con configuración única

La Figura 19 ilustra este mal funcionamiento en un ejemplo de validación del tipo de comportamiento degradado TR-I+TDH. La gráfica superior muestra los patrones o medidas virtuales que constituyen las entradas de la red. La gráfica situada en el medio de la figura representa las salidas de la red, y la gráfica inferior los valores deseados o targets. Como se puede observar, las salidas muestran un comportamiento indeseado, por lo que esta configuración no resulta adecuada.

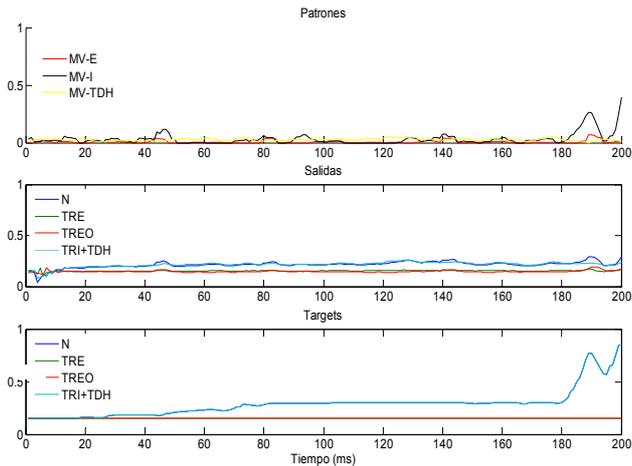


Figura 19. Ejemplo de TR-I+TDH con E-U en el rango [0,15-0,85]

En cuanto a la configuración paralela, la Figura 20 ilustra un ejemplo de validación correspondiente al tipo de comportamiento degradado TR-EO. Como se puede apreciar, la neurona de salida TR-EO sigue adecuadamente al target, mientras la neurona TR-E permanece en nivel bajo.

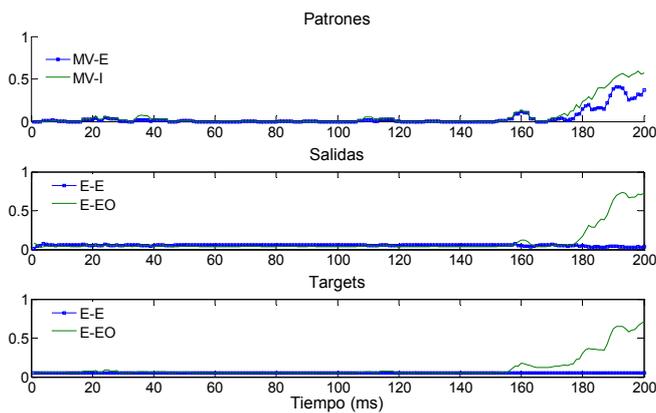


Figura 20. Ejemplo de validación con Elman Energía: TR-EO

4.3 Configuración de las redes neuronales seleccionadas

A la vista de los resultados obtenidos, cuyos resultados se han mostrado en los dos apartados anteriores, las configuraciones de red que mejor responden a los objetivos planteados son las basadas en la arquitectura Elman paralela que, entre otras razones, presentan menores errores de validación. La Tabla 1 muestra las configuraciones de las distintas redes seleccionadas así como el ratio de validación. Este último es el porcentaje de aciertos de la red en todos los ensayos de validación que incluyen comportamientos tanto en corte estable como en corte degradado.

La aplicación de esta arquitectura también se justifica desde el punto de vista de requisitos temporales y de memoria. Así, la arquitectura recurrente permite en este caso obtener tamaños de red significativamente menores. Además, evita tanto el coste computacional como de desarrollo que implicaría la gestión del buffer de las entradas en una arquitectura estática.

Tabla 1. Configuraciones de las redes neuronales E-E, E-I y E-TDH

RNA	E-E	E-I	E-TDH
Neuronas ocultas	10	10	10
Rango	[0.15-0.85]	[0.15-0.85]	[0.1-0.9]
Error V (MSE)	0,0005	0,0059	0,0088
Ratio de validación	91%	90,4%	100%

La Figura 21 muestra el esquema general del sistema de detección de comportamientos degradados en WEDM basado en RNA. Como se puede observar, está constituido por tres redes neuronales Elman en paralelo: red paralela de la energía E-E, red paralela de intensidad E-I y red paralela de tiempo de ionización elevado E-TDH.

El post-procesamiento de las salidas de la red consiste básicamente en lo siguiente:

- Post-procesamiento del nivel de alarma: consiste en la activación de las alarmas que superan un determinado umbral establecido durante la etapa de análisis de resultados. Este umbral se define como el 95% del peor caso correcto, es decir, del menor umbral alcanzado por las salidas en la fase de entrenamiento.
- Post-procesamiento del tipo de rotura: consiste en la cuantificación del grado de influencia de cada una de los tipos de rotura en la degradación el proceso. En concreto, se trata del porcentaje de contribución de cada salida.

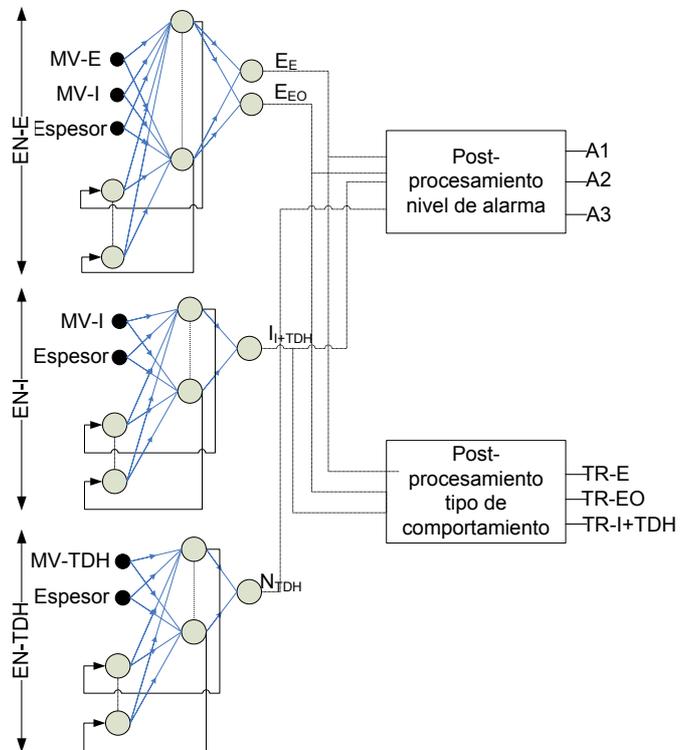


Figura 21. Esquema general del sistema de detección de inestabilidades en WEDM basado en ANN

## 5. EJEMPLOS DE APLICACIÓN

En este apartado se ilustra la operación de la red Elman que procesa los tipos de comportamientos degradados relacionados con la energía. En concreto, se presentan dos ejemplos (ver Figura 22 y Figura 23). En la parte superior de la figura de cada ejemplo se muestran las entradas de la red, y en la parte inferior las salidas.

En primer lugar, se muestra un ejemplo relacionado de rotura caracterizado por la sucesión de picos de elevada energía. (ver Figura 22). En la gráfica de las salidas, se puede apreciar que el tipo de comportamiento degradado TR-EO es correctamente estimado puesto que su salida (E-EO) presenta un comportamiento creciente. Mientras, la salida correspondiente a TR-E permanece en un nivel bajo.

El ejemplo presentado en la Figura 23, corresponde al tipo de comportamiento degradado caracterizado por un incremento brusco de la energía (TR-E). En este caso, la salida E-E, correspondiente al tipo de degradación de este ejemplo, presenta un comportamiento creciente, mientras la salida E-EO se mantiene en un nivel bajo.

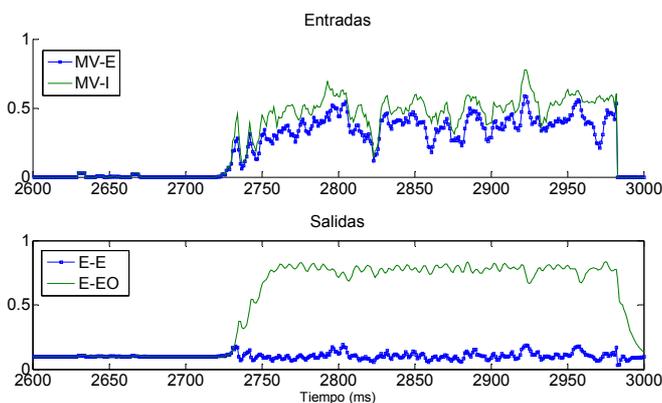


Figura 22. Ejemplo de aplicación: TR-EO

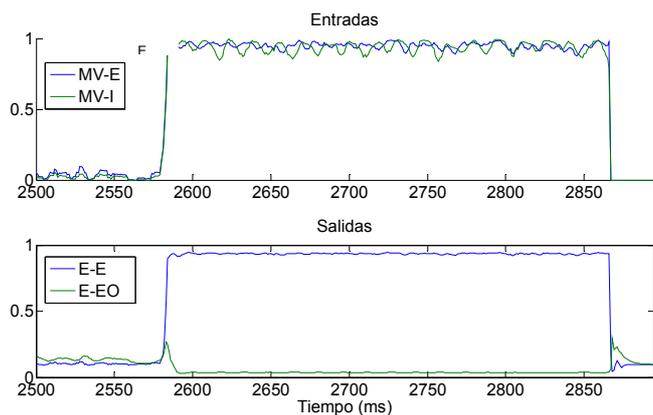


Figura 23. Ejemplo de aplicación: TR-E

## 6. CONCLUSIÓN

En este trabajo se muestra que la aplicación de redes neuronales en la detección de la degradación del proceso WEDM constituye una alternativa claramente viable. Esta alternativa, además, presenta la ventaja de que tiene visos de responder a dos objetivos simultáneamente: predecir la rotura del hilo en espesores intermedios y cuantificar el grado de influencia de

cada una de las posibles causas de rotura en la degradación el proceso. Otro aspecto fundamental a tener en cuenta es que el proceso WEDM se caracteriza por una dinámica muy rápida. Un sistema basado en redes neuronales, por tanto, se presenta como una técnica altamente apropiada para ser integrada en un sistema de control en tiempo real dada la rapidez computacional de las redes neuronales.

Tras la elaboración de un estudio comparativo de distintas configuraciones de red en arquitecturas estáticas (PM) y recurrentes (Elman), se ha observado que esta última es la más apropiada dada la naturaleza dinámica de los patrones y targets. Además, esta arquitectura permite obtener tamaños de red significativamente menores si se comparan con PM, facilitando así su aplicación en tiempo real.

## AGRADECIMIENTOS

Este trabajo ha sido subvencionado por la Universidad del País Vasco (Subvención a Grupo de Investigación, GIU07/36). Nos gustaría agradecer la colaboración del Departamento de Ingeniería Mecánica de la Universidad del País Vasco, y muy especialmente la de J.A. Sánchez.

## REFERENCIAS

- Behrens, A. y J. Ginzl (2003). Neuro-Fuzzy Process Control System for Sinking EDM. *Journal of Manufacturing processes*, **Vol. 5**, No. 1, pp. 33-39.
- Cabanes, I., E. Portillo, M. Marcos, y J.A. Sánchez (2008). On-line prevention of wire breakage in wire electro-discharge machining. *Robotics and Computer-Integrated Manufacturing*, **Vol. 24**, No. 2, pp. 287-298.
- Cabanes, I., E. Portillo, M. Marcos, y J.A. Sánchez (2008). Industrial application for on-line detection of instability and wire breakage in wire EDM). *Journal of Materials Processing Technology*, **Vol. 195**, No. 1-3, pp. 101-109.
- Elman, J. L. (1990). Finding Structure in Time. *Cognitive Science*, **Vol. 14**, pp. 179-211.
- Fenggou, C. y Y. Dayong (2004). The study of high efficiency and intelligent optimization system in EDM sinking process. *Journal of Materials Processing Technology*, **Vol. 149**, pp. 83-87.
- Ho K. H., S. T. Newman, S. Rahimifard y R. D. Allen (2004). State of the art in wire electrical discharge machining (WEDM). *International Journal of Machine Tools and Manufacture*, **Vol. 44**, pp. 1247-1259.
- Huang, J.T. y Y.S. Liao (2000). A wire edm maintenance and fault-diagnosis expert system integrated with an artificial neural network. *Int. J. Prod. Res.*, **Vol. 38**, No. 5, pp. 1071-1082.
- Isasi Viñuela, P. y I. M. Galván León (2004). *Redes de Neuronas Artificiales. Un enfoque práctico*, Ed. Pearson Educación, S.A., Madrid.
- Kao, J. Y. y Y. S. Tarnq (1997). A neural-network approach for the on-line monitoring of the electrical discharge machining process. *Journal of Materials Processing Technology*, **Vol. 69**, No. 1-3, pp. 112-119.
- Kunieda, M.; S. Saga, H. Yoshino, T. Ohta, y M. Kobayashi (2001). Control of discharge locations in EDM with locally imposed high electric field. *Proceedings of the XIII ISEM*, **Vol. 2**, pp. 485-495.
- Lauwers B., J. P. Kruth, P. H. Bleys, B. Van Coppenolle, L. Stevens, R. Derighetti (1999). Wire rupture prevention

- using on-line pulse localisation in WEDMB. *Vdi Berichte*, **Vol. 1405**, pp. 203-213.
- Liao, Y.S. y J.C. Woo (1997a). The effects of machining settings on the behaviour of pulse trains in the WEDM process. *Journal of Materials Processing Technology*, **Vol. 71**, pp. 433-439.
- Liao, Y.S., Y.Y. Chu and M.T. Yan (1997b). Study of wire breaking process and monitoring of WEDM. *International Journal of Machine Tools and Manufacture*, **Vol. 37 (4)**, pp. 555-567.
- Liao Y. S, Woo, J. C. Design of a fuzzy control system for the adaptive control of WEDM process. *International Journal of Machine Tools and Manufacture*. 2000; **Vol. 40 (15)**, p. 2293-2307.
- Mediliyegedara, T.K.K.R., A.K.M. De Silva,, D.K. Harrison, J.A. McGeough (2004). An intelligent pulse classification system for electro-chemical discharge, machining (ECDM)—a preliminary study- *Journal of Materials Processing Technology*, **Vol. 149**, pp. 499-503.
- Obara, H.; M. Abe y T. Ohsumi (1999). Control of Wire Breakage during Wire EDM. *International Journal of Electrical Machining*, No. 4.
- Portillo, E., I. Cabanes, M. Marcos, D. Orive, J. A. Sánchez (2007). Design of a virtual instrumentation system for a machining process. *IEEE Transactions on Instrumentation and Measurement*. **Vol. 56**, No. 6, 2007, pp. 2616-2622.
- Sarle W.S. Neural FAQ (2002) <ftp://ftp.sas.com/pub/neural/FAQ.html>.
- Sarle W.S. (1995) Stopped training and other remedies for overfitting. *Proceedings of the 27th Symposium on the Interface of Computing Science and Statistics*. 1995. - pp. 352-360.
- Shoda K., Y. Kaneko, H. Nishimura, M. Kunieda, M. X. Fan (1995). Development of adaptive control system to prevent EDM wire breakage. *EDM technology*, **Vol. 3**, pp. 17–22.
- Tarn, Y.S., S.C. Ma, L.K Chung (1995). Determination of optimal cutting parameters in WEDM. *International Journal of Machine Tools and Manufacture*, **Vol. 35**, No. 12, pp. 1693-1701.
- Tetko I.V., Livingstone D.J. y Luik A.I. (1995) Neural Network Studies. 1. Comparison of overfitting and overtraining. *Journal of Chemical Information and Computer Science*. **Vol. 35**. – pp. 826-833.
- Tiira, T. (1999). Detecting teleseismic events using artificial neural networks. *Computers & Geosciences*, **Vol. 25**, pp. 929-38.
- Tsai, K. y P. Wang (2001). Predictions on surface finish in electrical discharge machining based upon neural network models. *International Journal of Machine Tools & Manufacture*, **Vol. 41**, pp. 1385–1403.
- Valverde, R., Gachet, D. (2007) Identificación de sistemas dinámicos utilizando redes neuronales RBF. *RIAI*, **Vol. 4**, No. 2, pp. 32-42.
- Wang, G (2004). Two-Phase Reverse Neural Network Approach for Modeling a Complicate Manufacturing Process with Small Sample Size. *Neural Information Processing*, **Vol. 2**, No. 1.
- Wu, J y M.H. Li (2001). The identification of the servo control state in wire electrical discharge machining process. *ISEM XIII*, pp. 423-433.
- Yan, M.T. y Y.S. Liao (1995). Adaptive control of WEDM process using the fuzzy control strategy. *ISEM XI*, pp. 343-352.
- Yan, Liao y Chang (2001), On-line Estimation of Workpiece Height by using Neural Networks and Hierarchical adaptive Control of WEDM. *J. Advanced Manufacturing Technology*, **Vol.18**, No. 2, pp. 884-89.