



Redes neuronales aplicadas al control estadístico de procesos con cartas de control EWMA

Neural networks applied to statistical process control with EWMA control charts

Ruth Milena Suárez Castro ¹, Iván Darío Ladino Vega ²

Fecha de Recepción: 23 de mayo de 2022

Fecha de Aceptación: 26 de septiembre de 2022

Cómo citar: Suárez-Castro., R.M. y Ladino-Vega., I.D. (2023). Redes neuronales aplicadas al control estadístico de procesos con cartas de control EWMA. *Tecnura*, 27(75), 72-88. <https://doi.org/10.14483/22487638.18623>

Resumen

Contexto: Existe una creciente necesidad de monitorear y predecir variables críticas en procesos productivos; por tanto, a partir del enfoque de control estadístico, se ha asumido desde el uso de cartas de control para mediciones individuales. Así, en este artículo se presentan los resultados del diseño de una red neuronal recurrente *long short term memory* (LSTM) para predecir el valor promedio de la variable temperatura en mediciones individuales, y así evaluar la capacidad de la red para obtener valores similares a los cálculos del promedio móvil ponderado EWMA para mediciones individuales.

Método: Se obtuvieron 1768 registros de mediciones individuales de temperatura realizadas por un sensor, en el conjunto de datos denominado *gas sensors for home activity monitoring data set*. Los datos de temperatura se representaron en una carta de control de promedios móviles ponderados exponenciales EWMA, con el fin de obtener los valores de la media del proceso y de identificar que este estuviera dentro del control estadístico. Posteriormente, se entrenó una red neuronal LSTM a una muestra de entrenamiento de 1184 datos con algoritmo *backpropagation* que permitiera obtener valores similares a EWMA, los cuales se validaron en una muestra de prueba de 584 datos de temperatura.

Resultados: Se obtuvo el diseño de una red neuronal con una unidad en la puerta de entrada, cuatro en la puerta de olvido y una en la puerta de salida entrenada con el algoritmo *Backpropagation*, la cual permitió calcular valores muy cercanos a los representados en la carta de control EWMA, con un MSE de 1.1405e-04.

Conclusión: Las redes neuronales LSTM son una buena alternativa para el cálculo de valores EWMA, cuando se requiera hacer control estadístico de un proceso que genera gran cantidad de datos obtenidos de mediciones y no se cuente con un *software* para procesarlos.

Financiación: Fundación Universitaria Los Libertadores

Palabras clave: redes neuronales LSTM, carta control, EWMA, temperatura.

¹Ingeniera industrial, magíster en Ingeniería Industrial. Docente de la Fundación Universitaria Los Libertadores. Bogotá (Colombia). Email: rmsuarezc@libertadores.edu.co

²Ingeniero electrónico, magíster en Ingeniería Electrónica, magíster en Ingeniería en Teleinformática. Docente de la Fundación Universitaria Los Libertadores. Bogotá (Colombia). Email: idladinov@libertadores.edu.co

Abstract

Context: There is a growing need to monitor and predict critical variables in production processes, from the statistical process control approach it has been assumed from the use of control charts for individual measurements, for that reason this article presents the results of the design of a *long short term memory* (LSTM) recurrent neural network to predict the average value of the variable temperature in individual measurements and thus evaluate the ability of the network to obtain values similar to the EWMA weighted moving average calculations for individual measurements. Being this.

Methodology: 1768 records of individual temperature measurements made by a sensor were obtained, in the data set called: Gas sensors for home activity monitoring data set. Temperature data was plotted on an EWMA exponential weighted moving average control chart to obtain process mean values and identify that the process was within statistical control. Subsequently, an LSTM neural network was trained on a training sample of 1184 data with a Backpropagation algorithm that allowed obtaining values similar to EWMA, which were validated in a test sample of 584 temperature data.

Results: The design of a neural network with a unit in the input gate, 4 units in the forgetting gate and 1 unit in the output gate was obtained, trained with the Backpropagation algorithm, it allowed to calculate values very close to those represented in the control chart. EWMA, with an MSE of $1.1405e-04$.

Conclusions: LSTM neural networks are a good alternative for calculating EWMA values, when statistical control of a process that generates a large amount of data obtained from measurements is required and there is no software to process them.

Financing: Fundación universitaria Los Libertadores

Keywords: LSTM neural networks, control chart, EWMA, temperature.

Tabla de contenidos

	Pág
Introducción	73
Metodología	80
Resultados	81
Conclusiones	83
Financiamiento	84
Referencias	84

INTRODUCCIÓN

El continuo avance de la ciencia y la tecnología ha permitido el desarrollo dinámico de sistemas de control a partir de la automatización industrial. El uso de dispositivos que capturan y registran información de variables químicas o físicas, tiene la particularidad de revelar lo que ocurre en un proceso en tiempo real y genera el registro de una gran cantidad de datos de manera continua. La información obtenida a través de los dispositivos que monitorean el proceso favorece su análisis, así como el diseño y la aplicación de controles, además de la toma de decisiones.

La integración de dispositivos de control automáticos con la informática hace posible la detección inmediata de desviaciones en los valores obtenidos por variables que se suelen monitorear en los procesos industriales, que pueden afectar los resultados esperados y poner en riesgo la seguridad. Cuando las variaciones que se dan afectan la obtención de productos dentro de las especificaciones de calidad, es recomendable el uso de herramientas de control estadístico, especialmente de herramientas que tienen la capacidad de analizar una gran cantidad de información.

El control estadístico tiene como propósito la identificación de causas de variabilidad asignables y no asignables en el desarrollo de procesos industriales. Con el fin de realizar su control, es necesario monitorear cada actividad para verificar el cumplimiento de parámetros establecidos de funcionamiento, de tal forma que se haga evidente la existencia de eventos asignables como fallos en máquinas, mezcla de materias primas en proporciones inadecuadas (Flores Sánchez, 2018), que deben ser controlados y eliminados, para así asegurar el desarrollo de productos que cumplan con las especificaciones requeridas por clientes y consumidores finales. Por esa razón, en la industria se han venido implementando técnicas estadísticas que permiten monitorear y vigilar cada una de las etapas, con el propósito de identificar y reducir las causas de variabilidad, de tal forma que no afecten la conformidad de los productos.

Según Montgomery, 2013, en el control de procesos, la variabilidad solo puede describirse en términos estadísticos. Razón por la cual se puede afirmar que el análisis de datos derivado de tareas de medición y monitoreo forma parte del control estadístico de procesos (SPC, por su sigla en inglés), y tienen una elevada influencia no solo en la conformidad de productos, sino también en los costos y la productividad de la organización. Desde 1950, el SPC ha tomado gran importancia y actualmente forma parte de metodologías de gestión como *lean manufacturing* (Govindarajana et al., 2010), que aplican técnicas de estadísticas, no solo como mecanismo para obtener beneficios para la organización fruto de la competitividad, sino también para minimizar los despilfarros representados en reprocesos y mayor uso de recursos.

La carta de control es un recurso gráfico utilizado en actividades de monitoreo y por eso se considera una de las principales técnicas de SPC (Montgomery, 2013). Consta de dos límites de control superior e inferior, y un límite central, este último asociado al comportamiento ideal del proceso, si no estuviera afectado por fuentes de variabilidad. El monitoreo a través de esta herramienta facilita la detección de causas de variabilidad asignables a fallas en el proceso a partir de señales que se encuentran fuera de los límites de control, y también favorecen el seguimiento y monitoreo de manufactura, a fin de identificar patrones normales de desempeño, así como patrones anormales (Xu et al., 2019). Se clasifican en dos tipos: cartas de control de variables y cartas de control de atributos.

Cuando las características de interés en el proceso son medibles, se hace referencia a cartas de control por variables, donde es posible estimar la media, el rango, la mediana o la desviación estándar de las características en estudio. Las cartas de control más conocidas están basadas en la estimación de media y desviación estándar de un número n de muestras organizadas en subgrupos, sin embargo los actuales sistemas de control automatizado permiten la obtención de muestras individuales del

proceso y no en subgrupos, razón por la cual, también se han desarrollado cartas de control por variables, que tienen en cuenta la influencia de datos pasados en el comportamiento de la media, y que se conocen como *cartas de control de sumas acumuladas* (Cusum) y *de promedio móviles ponderados exponencialmente* (EWMA), que tienen la facilidad para detectar pequeños corrimientos de la media en el proceso.

La EWMA es conocida como carta de control para promedios móviles ponderados exponencialmente, y fue creada por Roberts en 1959 (Herrera Acosta *et al.*, 2018). La ventaja principal de esta consiste en su elevada sensibilidad para detectar cambios muy pequeños en la media de la variable del proceso a controlar (Vergara Benavides, 2012); además, la detección se logra a gran velocidad, característica que comparte con las Cusum (Quintana *et al.*, 2015). La EWMA se interpreta como un caso particular de series de tiempo (Rius *et al.*, 1998). La elaboración de la EWMA utiliza el valor (L) recomendado entre 2,6 y 2,8 como múltiplo de sigma, y un valor lambda (λ) con valores más utilizados de 0,05; 0,1; o 0,20 (Montgomery, 2013), donde el promedio móvil ponderado exponencial se define con la ecuación (1):

$$z_i = \lambda x_i + (1 - \lambda)z_{i-1} \quad (1)$$

Con $z_0 = \mu_0$ con límites de control definidos por las ecuaciones (2), (3) y (4):

$$\text{Límite central} = \mu_0 \quad (2)$$

$$\text{Límite superior de control} = \text{UCL} = \mu_0 + L\sigma\sqrt{\frac{\lambda}{(2 - \lambda)}} \quad (3)$$

$$\text{Límite inferior de control} = \text{LCL} = \mu_0 - L\sigma\sqrt{\frac{\lambda}{(2 - \lambda)}} \quad (4)$$

Ahora bien, el auge que han venido tomando las nuevas tendencias tecnológicas que promueven la automatización de procesos ha generado el surgimiento de un control de procesos más detallado en frecuencia y en tiempos cortos de medición, situación que permite el monitoreo del proceso en tiempo real y que genera abundancia de datos para ser procesados y para identificar cambios en las variables de funcionamiento del proceso, que muestren cuando el proceso se encuentra fuera del control. Por consiguiente, es imprescindible el uso de técnicas de manejo de altos volúmenes de información a fin de realizar el seguimiento con mayor facilidad.

Las redes neuronales artificiales forman parte de los algoritmos sofisticados de optimización de *machine learning*, entre los que se encuentra un amplio rango de técnicas como algoritmos de clasificación y *clustering*, modelos de regresión, árboles de decisión, etc. Un aspecto en común de las técnicas de *machine learning*, como las redes neuronales, es que incluyen tres tipos de aprendizaje: supervisado, no supervisado (Peláez Chávez, 2012) y de refuerzo (Yang, 2019, Maisueche Cuadrado, 2019).

Las redes neuronales artificiales son modelos de cálculo operados por algoritmos sofisticados (Flórez López y Fernández, 2008), técnicas de procesamiento de información programables (Rivas

Asanza y Mazón Olivo, 2018), máquinas adaptativas que están hechas de la interconexión de neuronas artificiales (Rios *et al.*, 2020), conformadas por una gran cantidad de neuronas conectadas en capas (Theodorids, 2020); las cuales adquieren conocimiento a través de un aprendizaje tomado del entorno y lo almacenan en los pesos sinápticos de la red; donde el comportamiento de un neurona está definido por un modelo neuronal (Yang, 2019).

Cuando se utiliza lenguaje supervisado (*supervised metric learning*) (Pérez Verona y Arco García, 2016) para el entrenamiento de la red, este se realiza proporcionando un conjunto de ejemplos de los valores de salida que debe asumir la red. En el caso del aprendizaje no supervisado (*unsupervised metric learning*), se hace referencia a la inexistencia de conocimiento *a priori* de la salida que se puede obtener cuando se ingresa una entrada a la red y lo que produce es la categorización de las entradas a través del *clustering*. Por su parte, el aprendizaje de refuerzo o semisupervisado se compone de un ejemplo parcial del comportamiento deseado donde se debe comparar continuamente las entradas y salidas al interactuar con el medio ambiente.

Las redes neuronales están conformadas por varios elementos: enlaces de conexión o sinapsis, uniones sumatorias, funciones de activación y sesgos (*bias*) (Rios *et al.*, 2020). Los enlaces de conexión están dispuestos en los elementos receptores de la neurona artificial, donde sucede una parametrización de los pesos sinápticos. Cuando el valor del peso sináptico es mayor que cero se considera una conexión excitatoria, con pesos positivos; de lo contrario la conexión es inhibitoria, con pesos negativos (Berzal, 2018).

Las uniones sumatorias, por su parte, ponderan los factores y los pesos sinápticos para obtener la intensidad total de la señal recibida; la neurona se activa si la intensidad de la señal alcanza un valor umbral definido (Yang, 2019); mientras que la función de activación recibe el resultado de la unión sumatoria y, a partir de allí, define la salida de la neurona en un rango de valores. Por último, la *bias*, o sesgo, actúa como un parámetro adicional de la neurona vinculada a una entrada fija con valor igual a 1, tal como se aprecia en la figura 1.

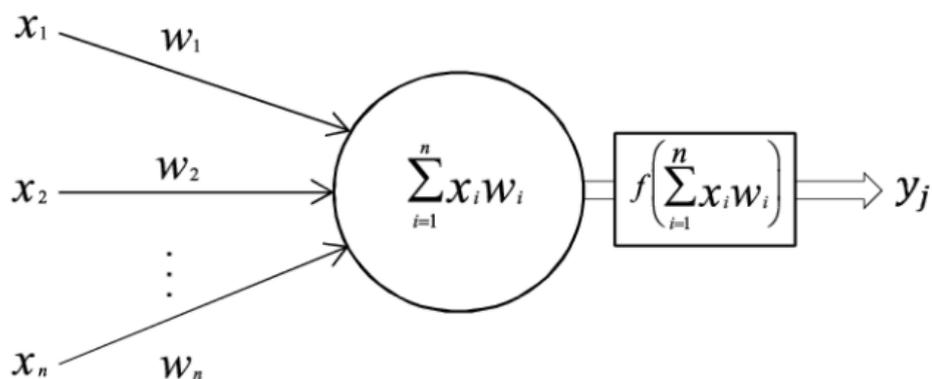


Figura 1. Estructura de la neurona artificial

Fuente: tomada de Vieira *et al.*, 2020.

La representación de la etapa de integración con uniones sumatorias y *bias* para la neurona artificial se presenta en la ecuación (5):

$$Z_j = b_j + \sum_{i=1}^n X_i W_i \quad (5)$$

Donde: Z es la suma ponderada de los factores x de entrada con los pesos sinápticos w ; b_j hace referencia al sesgo o bias; X_i representa a cada una de las n entradas o factores que recibe una capa de neuronas; W_i representan a los pesos sinápticos.

La representación de la etapa de activación con funciones de activación para la neurona artificial se presenta en la ecuación (6). Donde, y_j es el valor de salida de la neurona, $f(z_j)$ hace referencia a la función de activación.

$$y_j = f(z_j) = f\left(\sum_i w_{ij} x_i\right) \quad (6)$$

La estructura computacional de las redes neuronales artificiales se asemeja a la estructura del sistema nervioso humano de quien son análogas, por esa razón logran aprender con entrenamiento previo (Rivas Asanza y Mazón Olivo, 2018). Generalmente, las estructuras de redes neuronales artificiales se forman por conjuntos de neuronas que se agrupan en capas (de entrada, ocultas, de salida), donde las neuronas de cada capa comparten características comunes (Yang, 2019).

Es posible determinar dos etapas en el procesamiento de redes neuronales, la primera etapa llamada de integración ocurre cuando se combinan los factores de entrada de las neuronas con los pesos sinápticos, con el fin de establecer el tipo de conexión (excitatoria o inhibitoria), la segunda etapa es la activación, donde se determina el valor de salida de la neurona (Berzal, 2018).

La estructura de redes neuronales se forma cuando se conectan diversas neuronas según una arquitectura dada, es decir, cada neurona estará conectada a neuronas de capas adyacentes y conformarán una red, en la que la neurona transforma los datos de entrada por medio de la suma ponderada de la salida de las neuronas de la anterior capa y luego la pasa por una función no lineal que genera la salida de esa neurona; esta salida se convierte en la entrada de la siguiente capa hasta llegar al a capa final (Vieira et al., 2020). Tal como se observa en la figura 2.

Una manera de determinar el mejor desempeño de las redes neuronales está dada por la capacidad de minimizar los errores que se generen en el aprendizaje y la predicción (Yang, 2019). Es posible hacer una diferenciación entre las arquitecturas de redes neuronales artificiales, desde diversos criterios. El primero está dado por el número de capas, donde se encuentran las redes neuronales monocapa (de entrada y de salida) y redes multicapa (de entrada, intermedias y de salida), donde las redes multicapa pueden tener conexiones totales o parciales (Torres Álvarez et al., 2011).

El segundo criterio se da por el tipo de conexiones, entre las que se encuentran las redes no recurrentes (no retroalimentadas y unidireccionales) y redes neuronales recurrentes (permiten retroalimentación). Como último criterio está el grado de conexión, en el que se encuentran redes neuronales

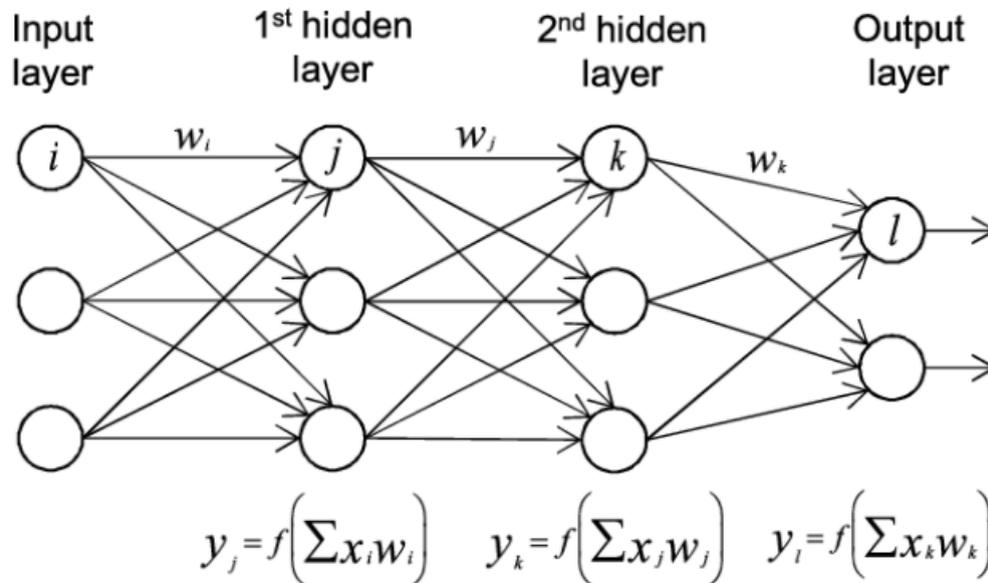


Figura 2. Estructura de red neuronal

Fuente: tomada de [Vieira et al., 2020](#).

con conexiones completas entre capas de neuronas, redes parcialmente conectadas, redes de propagación hacia adelante (*feed forward*) y redes de propagación hacia atrás (*feedback*) ([Rivas Asanza y Mazón Olivo, 2018](#)), redes de función de base radial (RBF, por su sigla en inglés) (activación según vecino más cercano), redes Hopfield, máquinas Boltzman ([Yang, 2019](#)).

En términos del mecanismo de aprendizaje que presentan las redes neuronales, se tiene el aprendizaje supervisado y el no supervisado. Entre las principales redes que utilizan aprendizaje supervisado están la red Adaline, el perceptrón y la red de retropropagación; este tipo de redes son utilizadas principalmente para clasificación, series de tiempo y eliminar ruido de señales de entrada. Para el caso del aprendizaje no supervisado, se encuentran las redes Hopfield, mapas de autoorganización de Kohonen, algoritmo *wake sleep* y la *learning vector quantization*; estas pueden solucionar problemas de análisis de componentes principales, *clustering*, prototipado, codificación, extracción y relación de patrones.

Son diversas las aplicaciones que se obtienen del uso de redes neuronales, van desde procesos de modelado, predicción, control, clasificación, reconocimiento de patrones, asociación de datos, agrupación, procesamiento, hasta optimización de señales ([Truong Pham et al., 2007](#), [Rivas Asanza y Mazón Olivo, 2018](#), [Montiel Ariza, 2015](#)). Entre las principales características de funcionamiento de las redes neuronales artificiales se encuentran la capacidad de resolver funciones no lineales, la transformación de entradas en salidas, la adaptabilidad, la tolerancia a los fallos, y la uniformidad en análisis y diseño ([Rios et al., 2020](#)).

Las redes neuronales recurrentes (RNN, por su sigla en inglés) fueron creadas en la década de 1980; son dinámicas, y en estas las entradas son realimentadas con muestras anteriores de las salidas; además, permiten conexiones hacia atrás en las capas que las conforman, por lo que se denominan *feedback networks* (Serrano et al., 2010). Esta característica hace que sean utilizadas para simular sistemas no lineales y para predecir series de tiempo, por su capacidad para seguir el tratamiento de secuencias, aun cuando existen otras herramientas para obtener análisis similares (Villarreal y Arango, 2013).

Para evitar el problema generado por el descenso del gradiente que se causa en aquellas tipo *feed forward networks*, se generaron dos tipos de redes recurrentes denominadas *long short term memory* (LSTM) y *gated recurrent neural networks* (GRU) las cuales están compuestas por tres capas: una de entrada, una oculta y una de salida, denominadas en ocasiones *puertas* (Núñez Castro, 2017). La figura 3 ilustra un modelo de red neuronal recurrente, donde A es una red neuronal, X_t es la entrada de la red y h_t es la salida de la red.

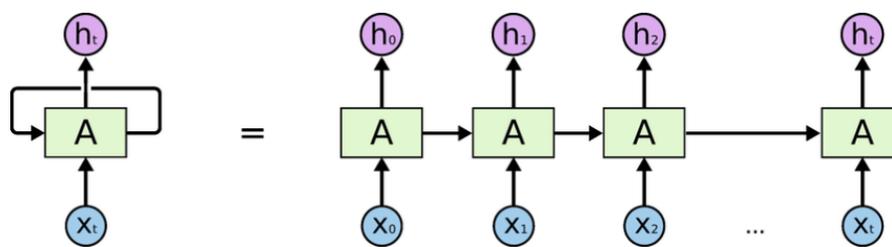


Figura 3. Estructura de red neuronal recurrente

Fuente: tomada de (Olah, (s. f.)).

La red neuronal artificial LSTM, creada por Hochreiter y Schmidhuber en 1997, es considerada una de las RNN de mayor aplicación. Es un tipo de modelo recurrente que implica tres unidades de control, una puerta de entrada (*input gate*), una de olvido (*forget gate*) y una de salida (*output gate*) (Wang et al., 2020), situación diferente de las capas utilizadas en otros modelos de red.

La característica más importante de este tipo de red neuronal es su capacidad para recordar patrones en una memoria de largo y corto plazo que captura información sobre lo que se ha calculado previamente a través del entrenamiento con el algoritmo de *backpropagation*; razón por la cual es muy utilizada para tratar problemas de predicción y previsión de secuencias (Elsheikh et al., 2021). Este tipo de redes solucionan el problema de dependencia entre los datos a largo plazo, situación que no es posible con el uso de redes *feed forward* (Núñez Castro, 2017).

La red neuronal LSTM utiliza funciones de activación sigmoide con valores entre 0 y 1, por lo que pueden incorporarse al algoritmo *backpropagation*, y solucionar el problema del descenso del gradiente, debido a que la red LSTM mantiene el valor del gradiente elevado (pendiente en un punto elevado). Esto favorece el aprendizaje del modelo, y así se evita que el valor de la pendiente sea cero, lo cual impide la continuación del proceso de aprendizaje.

En múltiples estudios se ha utilizado el control estadístico de procesos (SPC, por su sigla en inglés) en conjunto con redes neuronales. Los estudios más recientes se han encaminado al reconocimiento de patrones que indican fallas en los procesos, tal es el caso de las redes neuronales de base radial RBF que se han utilizado con el algoritmo de entrenamiento Bee-RBF (Addeha *et al.*, 2018), con el algoritmo *backpropagation*; igualmente, se han utilizado redes convolucionales (Fuquaa y Razzaghi, 2020) y redes neuronales de impulsos *spikinf* SNN (Awadallaa y Abdellatif Sadekb, 2012). Otros estudios han sido aplicados en la construcción de gráficos de control para pequeños corrimientos de la media del proceso con métodos Cusum y EWMA (Rius *et al.*, 1998). Para el monitoreo de la media en procesos multivariados se han utilizado cartas de control MWEMA (Aparisi y Carrión-García, 2010), incluyendo T^2 y cartas MEWMA en conjunto con redes neuronales del tipo perceptrón, con redes RBF (Chiñas-Sánchez y Vázquez-López, 2014) y redes que trabajan el algoritmo *backpropagation* (Beshah y Muluneh, 2017); por último, se han estudiado patrones de cambio en la varianza del proceso bivariado con redes neuronales de retropropagación (Cheng y Cheng, 2011).

En menor medida, se han estudiado redes neuronales que buscan pronosticar los valores medios de las variables que son monitoreadas con las cartas de control. En el caso de la media, se desarrollaron redes *feed-forward* con entrenamiento *backpropagation* (Cheng y Cheng, 2001), con perceptrón multicapa y algoritmo de entrenamiento Broyden-Fletcher-Goldfarb-Shanno (BFGS) (Cheng y Cheng, 2011).

Igualmente, las redes neuronales permiten hacer monitoreo de la media con cartas de control EWMA, para procesos univariados, por ejemplo, en el seguimiento de la temperatura de la cadena de frío con redes de retropropagación (Chen y Shaw, 2010) o del consumo de energía en sistemas de generación de aire comprimido (Santolamazza, 2018) por medio de redes neuronales *feed-forward* y regresión lineal. Las RNN se han utilizado como complemento de las EWMA, en un proceso de detección de anomalías en la transmisión de datos por cuenta de nodos de sensores (Zhang *et al.*, 2020).

METODOLOGÍA

El propósito de este estudio fue medir la capacidad de las redes neuronales para tratar datos obtenidos por sensores y así predecir valores que permitan el control estadístico de procesos, a partir de la construcción de cartas de control de medias móviles ponderadas exponenciales EWMA para mediciones individuales. En el desarrollo del estudio la serie de datos *gas sensors for home activity monitoring data set* utilizada fue obtenida del UCI Machine Learning Repository. Este conjunto de datos contiene grabaciones de un grupo de sensores de gas compuesto por 8 sensores de gas metal-óxido-semiconductor (MOX) y un sensor de temperatura y de humedad. El interés se centró en el análisis de los datos de 1768 registros de temperatura medidos por uno de los sensores anteriormente mencionados.

Los datos se procesaron inicialmente con el lenguaje de programación R para la obtención de la carta de control EWMA. A partir de allí se utilizó un tipo de red neuronal: *red long short term memory* (LSTM), *feedback forward* con el uso del lenguaje de programación Python; posteriormente, se compararon los resultados.

RESULTADOS

El objetivo del estudio se centró en pronosticar los valores de la media móvil exponencial con el uso de redes neuronales, con el fin de determinar el ajuste de la red neuronal creada con la carta de control EWMA. Para ello se tomó la serie de 1768 datos obtenidos por un sensor de temperatura y presentados en el *data set* denominado *sensores de gas para conjunto de datos de monitoreo de actividad en el hogar*. Luego, se procedió a construir la carta de control EWMA con parámetros iniciales $\mu_0 = 26,11$, parámetro de suavización $\lambda = 0,1$ y $L = 2,7$; se encontraron todos los datos de temperatura dentro de los límites de control.

De acuerdo con la figura 4, el valor promedio de las mediciones individuales corresponde a 26,1138, con una desviación estándar de 1. Cada uno de los valores EWMA calculados se representan en la carta de control a partir de puntos que indican cada medición realizada. En la carta de control se observa que, para las 1768 muestras, las mediciones se mantuvieron dentro de los límites de control variables.

La EWMA para mediciones individuales evidenció un corrimiento de la media a partir de la muestra 820, donde la temperatura descendió abruptamente de una temperatura de 26,11 °C hasta un 25,98 °C que permaneció así hasta la muestra 1112, donde vuelve a incrementarse; sin embargo, se sigue manteniendo dentro de los límites de control estadístico.

Se procedió a desarrollar una red neuronal LSTM con la finalidad de pronosticar los valores de la media móvil exponencial y de realizar una nueva carta de control; las entradas de la red corresponden al registro de 1768 datos de temperatura captados por un sensor. Se utilizó el lenguaje de programación Python con el paquete Keras, con el objetivo de correr una red neuronal que tuviera la capacidad de calcular valores semejantes al obtenido en la carta de control EWMA; teniendo en cuenta que cada dato proyectado depende de los datos anteriores se considera que los datos están correlacionados.

La red neuronal LSTM utilizada consta de una puerta de entrada con 1 unidad, 1 puerta de olvido con 4 unidades y 1 puerta de salida con 1 unidad y funciones de activación sigmoidea. Para el entrenamiento de la red se tomó una muestra de 1184 datos de temperatura, correspondientes al 67 % del total medido por el sensor, así como un algoritmo de entrenamiento *backpropagation* que al tener en cuenta determinados instantes de tiempo se ha denominado *backpropagation through time* (BPTT). La ventaja de la aplicación de las redes neuronales LSTM es que permiten registrar los datos de entrenamiento en orden cronológico y no de manera aleatoria, situación que mejora la estimación de los valores de salida de la red, dado que aumenta la calidad de las predicciones.

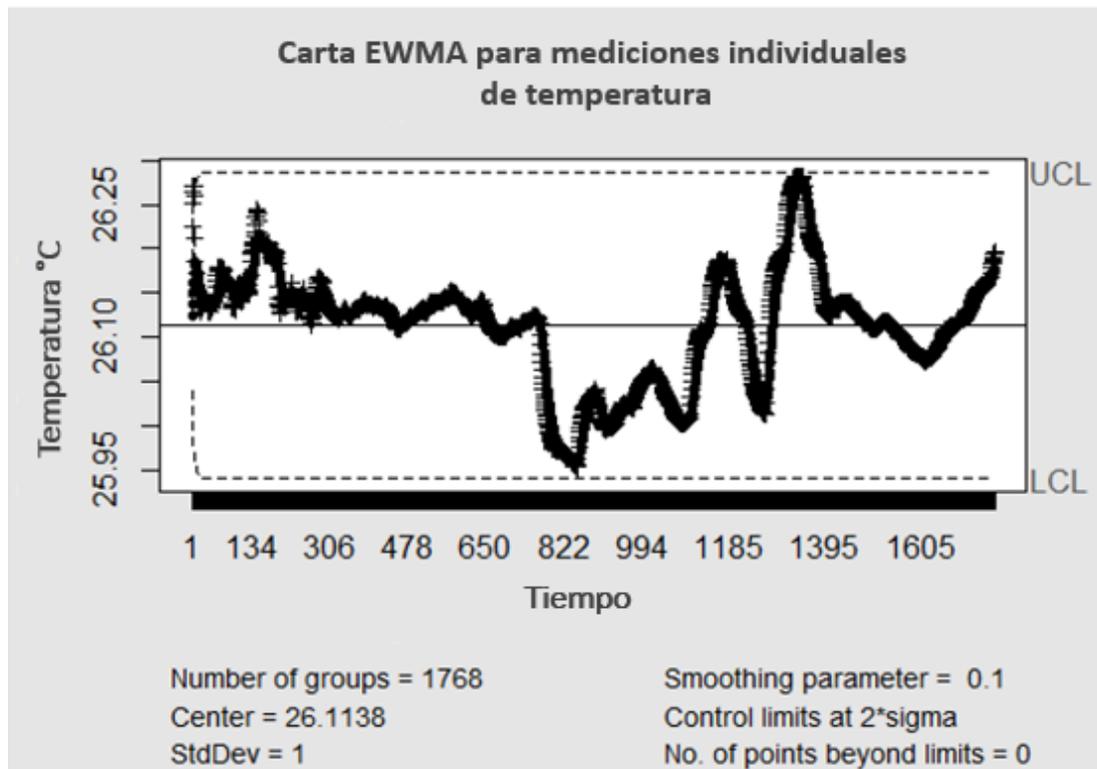


Figura 4. Carta de control EWMA para mediciones individuales de temperatura

Fuente: Elaboración propia.

El modelo de la red neuronal se configuró con lote (*batch*) de 1 dato y con 100 épocas, o que indica que el modelo procesa los datos de uno en uno y se realizarán 100 corridas del modelo hasta obtener el menor error cuadrado medio (MSE, por su sigla en inglés). Para el modelo en estudio se encontró que el MSE en la época 1 fue de 0,0444 y en la época 100 el MSE había disminuido hasta 1,1405e-04.

La red LSTM cuenta con 96 parámetros en la puerta de olvido (*forget gate*) y 5 en la puerta de salida (*output gate*), para un total de 101 parámetros, tal como lo muestra la tabla 1. Las RNN utilizan un tamaño de lote *batch* fijo = 1, para asegurar que pasan de un instante de tiempo al siguiente. Para realizar las pruebas sobre la red entrenada se utilizó el 33 % del restante de la muestra de temperatura, correspondiente a 584 datos y del cual se obtuvo la figura 5, en la cual se identifican los valores reales de la medición realizada por el sensor de temperatura (azul), la predicción realizada por la red con los datos de entrenamiento (naranja) y la predicción realizada con los datos de prueba (verde).

El resultado de la aplicación de la red neuronal LSTM arrojó valores pronosticados con un elevado grado de semejanza con los valores EWMA representados en la carta de control, lo que indica la capacidad de la red neuronal LSTM para procesar datos de mediciones individuales, tal como se obtienen de las mediciones a partir de sensores. Igualmente, el algoritmo de protección el descenso del gradiente Adam, utilizado en esta red, asegura que los valores obtenidos de la predicción tengan

Tabla 1. Configuración de la red neuronal LSTM

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 4)	96
dense (Dense)	(None, 1)	5
Total Paramas: 101		
Trainable params: 101		
Non-trainable params: 0		

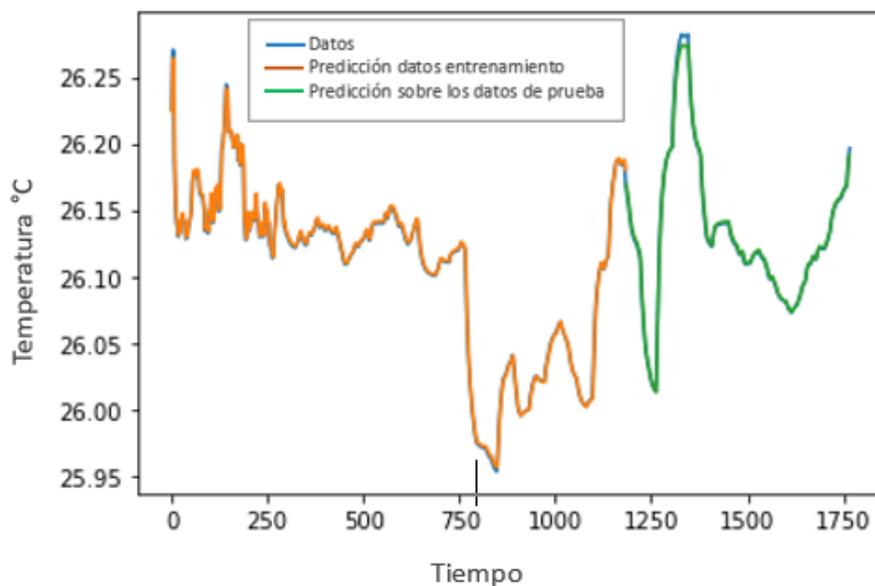


Figura 5. Comparación valores EWMA de temperatura vs. Predicción de red LSTM

Fuente: Elaboración propia.

en cuenta los datos correlacionados, como se requiere para obtener el cálculo de valores EWMA. De esta manera es posible identificar los corrimientos de la media de temperatura, de la misma manera que lo hace una carta de control EWMA.

CONCLUSIONES

Las cartas de control de la media ponderada exponencial EWMA son herramientas de control estadístico que permiten monitorear con facilidad aquellos procesos de los que se obtienen mediciones individuales; estas no solamente indican si se encuentran dentro de los límites de control, sino que además evidencian pequeños cambios que se van dando en las variables de estudio, como en este caso es la temperatura.

Las redes neuronales recurrentes (RNN) del tipo *long short term memory* (LSTM) son una herramienta de gran utilidad para el análisis de series de tiempo, debido a que poseen algoritmos que evitan que el gradiente descienda hasta tal punto que frene la capacidad de aprendizaje de la red; esto la hace adecuada para el análisis de datos obtenidos de mediciones individuales de variables importantes de monitorear.

Las redes neuronales LSTM son una buena alternativa para el cálculo de valores EWMA, cuando se requiera hacer control estadístico de un proceso que genera gran cantidad de datos obtenidos de mediciones y no se cuente con un *software* para procesarlos.

FINANCIAMIENTO

Artículo de investigación científica derivado del proyecto de investigación “Análisis de datos sobre una red neuronal embebida en un microcontrolador ARM7”, financiado por la Fundación Universitaria Los Libertadores (año de inicio: 2021, año de finalización: 2021).

REFERENCIAS

- [Addeha *et al.*, 2018] Addeha, A., Khormalib, A. y Golilarz, N. A. (2018). Control chart pattern recognition using RBF neural network with new training algorithm and practical features. *ISA Transactions*, 79, 202-212. 10.1016/j.isatra.2018.04.020 <https://doi.org/10.1016/j.isatra.2018.04.020> ↑Ver página 80
- [Aparisi y Carrión-García, 2010] Aparisi, F. y Carrión-García, A. (2010). Artificial neural networks for identifying the signals of multivariate EWMA control charts. En *10th International Conference on Intelligent Systems Design and Applications* (pp. 427-431). ISDA. <https://doi.org/10.1109/ISDA.2010.5687226> ↑Ver página 80
- [Awadallaa y Abdellatif Sadekb, 2012] Awadallaa, M. y Abdellatif Sadekb, M. (2012). Spiking neural network-based control chart pattern. *Alexandria Engineering Journal*, 51(1), 27-35. 10.1016/j.aej.2012.07.004 <https://doi.org/10.1016/j.aej.2012.07.004> ↑Ver página 80
- [Berzal, 2018] Berzal, F. (2018). *Redes neuronales y deep learning*. Edición independiente. ↑Ver página 76, 77
- [Beshah y Muluneh, 2017] Beshah, B. y Muluneh, A. (2017). Control chart pattern recognition of multivariate auto-correlated processes using artificial neural network. *Journal of EEA*, 35, 47-57. ↑Ver página 80

- [Cheng y Cheng, 2001] Cheng, C.-S. y Cheng, S.-S. (2001). A neural network-based procedure for the monitoring of exponential mean. *Computers % Industrial Engineering*, 40, 309-321. [https://doi.org/10.1016/S0360-8352\(01\)00031-6](https://doi.org/10.1016/S0360-8352(01)00031-6) ↑Ver página 80
- [Cheng y Cheng, 2011] Cheng, C.-S. y Cheng, H.-P. (2011). Using neural networks to detect the bivariate process variance shifts pattern. *Computers % Industrial Engineering*, 60(2), 269-278. [10.1016/j.cie.2010.11.009](https://doi.org/10.1016/j.cie.2010.11.009) <https://doi.org/10.1016/j.cie.2010.11.009> ↑Ver página 80
- [Chen y Shaw, 2010] Chen, K.-Y. y Shaw, Y.-C. (2010). Applying back propagation network to cold chain temperature monitoring. *Advanced Engineering Informatics*, 25(1), 11-22. [10.1016/j.aei.2010.05.003](https://doi.org/10.1016/j.aei.2010.05.003) <https://doi.org/10.1016/j.aei.2010.05.003> ↑Ver página 80
- [Cheng *et al.*, 2011] Cheng, C.-S., Chen, P.-W. y Huang, K.-K. (2011). Estimating the shift size in the process mean with support vector regression. *Expert Systems with Applications*, 38, 10624-10630. <https://doi.org/10.1016/j.eswa.2011.02.121> ↑Ver página
- [Chiñas-Sánchez y Vázquez-López, 2014] Chiñas-Sánchez, P. y Vázquez-López, J. A. (2014). Multivariate variables recognition using Hotelling's T2 and MEWMA via ANN's. *Ingeniería, Investigación y Tecnología*, 15(1), 125-138. [https://doi.org/10.1016/S1405-7743\(15\)30011-1](https://doi.org/10.1016/S1405-7743(15)30011-1) ↑Ver página 80
- [Elsheikh *et al.*, 2021] Elsheikh, A., Katekar, V., Muskens, O., Deshmukh, S., Elaziz, M. y Dabour, S. (2021). Utilization of LSTM neural network for water production forecasting of a stepped solar still with a corrugated absorber plate. *Process Safety and Environmental Protection*, 148, 273-282. <https://doi.org/10.1016/j.psep.2020.09.068> ↑Ver página 79
- [Flores Sánchez, 2018] Flores Sánchez, M. (2018). *Nuevas aportaciones del análisis de datos funcionales en el control estadístico de procesos* [Tesis de doctorado]. Universidad de Coruña. ↑Ver página 74
- [Flórez López y Fernández, 2008] Flórez López, R. y Fernández, J. M. (2008). *Las redes neuronales artificiales. Fundamentos teóricos y aplicaciones prácticas*. Netbiblo. ↑Ver página 75
- [Fuquaa y Razzaghi, 2020] Fuquaa, D. y Razzaghi, T. (2020). A cost-sensitive convolution neural network learning for control chart pattern recognition. *Expert Systems with Applications*, 150, 113275. <https://doi.org/10.1016/j.eswa.2020.113275> ↑Ver página 80
- [Govindarajana *et al.*, 2010] Govindarajana, R., Lluguerab, E., Melero, A., Molero, J. y Soler, N. (2010). El control estadístico de proceso puede ayudar a prevenir los errores de tratamiento sin aumentar los costes en radioterapia. *Revista de Calidad Asistencial*, 25(5), 281-290. <https://doi.org/10.1016/j.cali.2010.04.004> ↑Ver página 74

- [Herrera Acosta *et al.*, 2018] Herrera Acosta, R., Romero Cabrera, I. y Wasinki-Zúñiga, R. (2018). Contraste entre las cartas de control MR Shewart y Cusum varianza en el monitoreo del potencial de hidrógeno en protectores de planta. *Iteckne, Innovación e Investigación en Ingeniería* 15(2), 88-98. 10.15332/iteckne.v15i2.2070. <https://doi.org/10.15332/iteckne.v15i2.2070> ↑Ver página 75
- [Maisueche Cuadrado, 2019] Maisueche Cuadrado, A. (2019). *Utilización del machine learning en la industria 4.0* [Tesis de maestría]. Repositorio institucional de la Universidad de Valladolid. <https://uvadoc.uva.es/handle/10324/37908> ↑Ver página 75
- [Montgomery, 2013] Montgomery, D. (2013). *Control estadístico de la calidad*. Limusa Willey. ↑Ver página 74, 75
- [Montiel Ariza, 2015] Montiel Ariza, H. M. (2015). Using neural networks for face recognition in controlled environments. *Tecnura*, 19, 67-77. <https://doi.org/10.14483/udistrital.jour.tecnura.2015.SE1.a05> ↑Ver página 78
- [Núñez Castro, 2017] Núñez Castro, J. F. (2017). *Aprendizaje automático en fusión nuclear con deep learning* [Tesis de pregrado]. Pontificia Universidad Católica de Valparaíso. ↑Ver página 79
- [Olah, (s. f.)] Olah, Ch. (s. f.). Recurrent neural networks. *Colah's Blog*. <http://colah.github.io/> ↑Ver página 79
- [Peláez Chávez, 2012] Peláez Chávez, N. (2012). *Aprendizaje no supervisado y el algoritmo wake sleep en redes neuronales* [Tesis de pregrado]. Universidad Tecnológica de la Mixteca. ↑Ver página 75
- [Pérez Verona y Arco García, 2016] Pérez Verona, I. C. y Arco García, L. (2016). Una revisión sobre aprendizaje no supervisado de métricas de distancia. *Revista Cubana de Ciencias Informáticas*, 10(4), 43-67. ↑Ver página 76
- [Quintana *et al.*, 2015] Quintana, A. E., Pisani, M. V. y Casal, R. N. (2015). Desempeño de cartas de control estadístico con límites bilaterales de probabilidad para monitorear procesos Weibull en mantenimiento. *Ingeniería, Investigación y Tecnología*, 16(1), 143-156. [https://doi.org/10.1016/S1405-7743\(15\)72115-3](https://doi.org/10.1016/S1405-7743(15)72115-3) ↑Ver página 75
- [Rios *et al.*, 2020] Rios, J., Alanis, A., Arana-Daniel, N. y López-Franco, C. (2020). Appendix A - Artificial neural networks. En *Neural networks modeling and control. Applications for unknown nonlinear delayed systems in discrete time* (pp. 117-124). Academic Press ↑Ver página 76, 78
- [Rius *et al.*, 1998] Rius, A., Ruisanchez, I., Callao, M. y Rius, F. (1998). Reliability of analytical systems: Use of control charts, time series models and recurrent neural networks RNN. *Chemometrics and Intelligent Laboratory Systems*, 40, 1-18. [https://doi.org/10.1016/S0169-7439\(97\)00085-3](https://doi.org/10.1016/S0169-7439(97)00085-3) ↑Ver página 75, 80

- [Rivas Asanza y Mazón Olivo, 2018] Rivas Asanza, W. y Mazón Olivo, B. (2018). *Redes neuronales artificiales aplicadas al reconocimiento de patrones*. Editorial UTMACH. ↑Ver página 75, 77, 78
- [Santolamazza, 2018] Santolamazza, A. C. (2018). Anomaly detection in energy consumption for condition-based maintenance of compressed air generation systems: An approach based on artificial neural networks. *IFAC-PapersOnLine*, 51(11), 1131-1136. 10.1016/j.ifacol.2018.08.439 <https://doi.org/10.1016/j.ifacol.2018.08.439> ↑Ver página 80
- [Serrano et al., 2010] Serrano, A., Soria, E. y Martín, J. D. (2010). *Redes neuronales artificiales*. Universidad de Valencia. ↑Ver página 79
- [Theodoridis, 2020] Theodoridis, S. (2020). Chapter 18 - Neural networks and deep learning. En *Machine learning: A Bayesian and optimization perspective* (pp. 901-1038). Academic Press. <https://doi.org/10.1016/B978-0-12-818803-3.00030-1> ↑Ver página 76
- [Torres Álvarez et al., 2011] Torres Álvarez, N., Hernández, C. y Pedraza, L. (2011). Redes neuronales y predicción de tráfico. *Tecnura*, 15, 90-97. ↑Ver página 77
- [Truong Pham et al., 2007] Truong Pham, D., Packianather, M. y Afify, A. (2007). Artificial neural networks. En D. Andina y D. Truong Pham, *Computacional intelligence: For engineering and manufacturing* (pp. 67-92). Springer. https://doi.org/10.1007/0-387-37452-3_3 ↑Ver página 78
- [Vergara Benavides, 2012] Vergara Benavides, M. C. (2012). Aplicación de la carta de control EWMA-CV para la optimización del monitoreo del peso en la leche pasteurizada en bolsa. *Revista Científica Technos*, 8(1), 7-16. ↑Ver página 75
- [Vieira et al., 2020] Vieira, S., López Pinaya, W. y Garcia-Dias, R. M. (2020). Chapter 9 - Deep neural networks. En S. Vieira y A. Mechelli, *Machine learning methods and applications to brain disorders* (pp. 157-172). Academic Press. <https://doi.org/10.1016/B978-0-12-815739-8.00009-2> ↑Ver página 76, 77, 78
- [Villarreal y Arango, 2013] Villarreal, E. y Arango, D. (2013). Estrategias para el entrenamiento de redes neuronales de números difusos. *Tecnura*, 18, 36-47. <https://doi.org/10.14483/udistrital.jour.tecnura.2014.2.a03> ↑Ver página 79
- [Wang et al., 2020] Wang, P., Zheng, X., Ai, G., Liu, D. y Zhu, B. (2020). Time series prediction for the epidemic trends of COVID-19 using the improved LSTM deep learning method: Case studies in Russia, Peru and Iran. *Chaos, Solitons & Fractals*, 140, 110240. 10.1016/j.chaos.2020.110214 <https://doi.org/10.1016/j.chaos.2020.110214> ↑Ver página 79
- [Xu et al., 2019] Xu, J., Lv, H., Zhuan, Z., Lu, Z., Zou, D. y Qin, W. (2019). Control chart pattern recognition method based on improved one-dimensional convolutional neural network. *IFAC-*

PapersOnLine, 52(13), 1537-1542. <https://doi.org/10.1016/j.ifacol.2019.11.418> ↑Ver página 74

[Yang, 2019] Yang, X.-S. (2019). 8-Neural networks and deep learning. En X.-S. Yang, *Introduction to algorithms for data mining and machine learning* (pp. 139-161). Academic Press. <https://doi.org/10.1016/B978-0-12-817216-2.00015-6> ↑Ver página 75, 76, 77, 78

[Zhang *et al.*, 2020] Zhang, M., Guo, J., Li, X. y Jin, R. (2020). Data-driven anomaly detection approach for time series streaming data. *Sensors*, 20(19), 5646. 10.3390/s20195646 <https://doi.org/10.3390/s20195646> ↑Ver página 80

