TESIS DOCTORAL

2021

Predicción de propiedades de materiales para la fabricación de componentes destinados a aplicaciones de alta exigencia mediante simulación numérica, análisis avanzado de datos e inteligencia artificial

David Merayo Fernández

Programa de doctorado en tecnologías industriales

Directores:
Dr. D. Álvaro Rodríguez Prieto
Dra. Dª. Ana María Camacho López


UNED
Escuela
Internacional
de Doctorado
EIDUNED

Agradecimientos

Quiero mostrar mi agradecimiento a mis directores de tesis D. Álvaro Rodríguez Prieto y Dª. Ana María Camacho López ya que ellos me han acompañado durante estos años aportándome todo lo necesario para llevar a cabo mi trabajo: ideas, correcciones y comentarios. Así mismo, quiero hacer extensible este agradecimiento al Grupo de Investigación Producción Industrial e Ingeniería de Fabricación (*Industrial Production and Manufacturing Engineering - IPME*) de la UNED.

Así mismo, quiero hacer llegar mi agradecimiento a mi familia y amigos por todo su apoyo y confianza.

Este trabajo se ha desarrollado en el marco del Programa de Doctorado en Tecnologías Industriales de la UNED y ha sido financiado parcialmente por las Ayudas de la Escuela Técnica Superior de Ingenieros Industriales de la UNED en diferentes convocatorias anuales.

Índice

Listado de tablas

Listado de figuras

Tabla de símbolos y abreviaturas

Símbolo	Significado
A	Elongación a rotura
ADAM	*Adaptive Moment Estimation*
AI	Inteligencia artificial
ANN	Red neuronal artificial
α	Parámetro Ramberg-Osgood
β_n	Parámetro ADAM
E	Módulo de Young
E_1	Módulo de Young en la dirección principal
E_2	Módulo de Young en la dirección secundaria
E_T	Módulo de endurecimiento por deformación
ϵ	Factor de estabilidad ADAM
ε	Error predictivo
ε	Deformación
η	Tamaño del paso ADAM
f	Función de error
FEM	Análisis de elementos finitos
g	Gradiente de la función de error
HB	Dureza Brinell
m	Estimador del primer momento ADAM
n	Parámetro Ramberg-Osgood
v	Estimador del Segundo momento ADAM
v	Coeficiente de Poisson
v_{12}	Coeficiente de Poisson en la dirección principal
v_{21}	Coeficiente de Poisson en la dirección secundaria
R^m	Ponderación de las propiedades mecánicas
R^ρ	Ponderación de la densidad
R^ϵ	Ponderación del coste
$R_i^{\sigma_m}$	Ponderación de *UTS*
$R_i^{\sigma_s}$	Ponderación de *YS*
$R_i^{\varepsilon_R}$	Ponderación de *A*
RNA	Red neuronal artificial
S_x	Desviación típica
SL_g	Nivel de severidad global
SL_m	Nivel de severidad mecánico
SL_ρ	Nivel de severidad de la densidad
SL_{σ_m}	Nivel de severidad de *UTS*
SL_{σ_s}	Nivel de severidad de *YS*
SL_{ε_R}	Nivel de severidad de *A*
σ	Tensión
t	Espesor de la tubería
τ	Tensión cortante
UTS	Resistencia máxima a tracción
w	Vector de pesos de ponderación
$\bar{x}$	Media
$\bar{x}_{90\%}$	Media recortada al 90%
YS	Tensión de fluencia

Resumen

Las propiedades del material con que se fabrica un componente industrial determinan, en gran medida, su comportamiento y su capacidad para soportar las condiciones de trabajo a las que se ve sometido durante su operación; ya que, un rendimiento insuficiente puede provocar un fallo prematuro, especialmente en entornos de alta exigencia donde las condiciones de servicio pueden ser particularmente críticas. Por ello, es fundamental disponer de herramientas que permitan seleccionar el material más adecuado para cada aplicación, siendo la simulación numérica, la analítica avanzada de datos y la inteligencia artificial algunas de las técnicas que presentan un mayor potencial gracias al desarrollo tecnológico y al acceso a la información. Concretamente, los materiales metálicos y algunos materiales compuestos juegan un papel fundamental en el campo industrial ya que pertenecen a dos de los grupos de materiales estructurales más destacados debido a sus cualidades; pero puede resultar muy costoso, en tiempo y recursos, caracterizar su comportamiento y sus propiedades de manera experimental mediante las técnicas que tradicionalmente se emplean para este fin. En particular, resulta de interés conocer el comportamiento de un material metálico para poder determinar su respuesta frente a diversas solicitaciones. Adicionalmente, se debe considerar que ciertos materiales compuestos pueden ser excelentes alternativas en determinadas aplicaciones de alta exigencia. En esta Tesis Doctoral se examina la capacidad de las simulaciones mediante elementos finitos, de las redes neuronales artificiales y de las técnicas basadas en aprendizaje automático e inteligencia artificial para predecir el comportamiento y las propiedades de materiales metálicos y compuestos de matriz polimérica. Estas herramientas pueden ser empleadas para extraer tendencias en los datos que permiten generar nueva información acerca de otros materiales y, por tanto, sirven de apoyo a la toma de decisiones. Se ha demostrado que existen aplicaciones de muy alta exigencia donde los materiales compuestos pueden sustituir a los materiales metálicos y se ha logrado desarrollar un modelo predictivo, basado en inteligencia artificial, para predecir algunas propiedades fundamentales de aleaciones metálicas.

Abstract

The properties of the material with which an industrial component is manufactured determine, to a large extent, its behavior and its ability to withstand the working conditions to which it is subjected during its operation; as insufficient performance can lead to premature failure, especially in demanding environments where service conditions can be particularly critical. For this reason, it is essential to have tools that allow selecting the most appropriate material for each application, being numerical simulation, advanced data analytics and artificial intelligence some of the techniques that present the greatest potential thanks to technological development and access to information. Specifically, metallic materials and some composite materials play a fundamental role in the industrial field since they belong to two of the most prominent groups of structural materials due to their qualities; but it can be very costly, in time and resources, to characterize its behavior and its properties experimentally using the techniques traditionally employed for this purpose. In particular, it is of interest to know the behavior of a metallic material in order to determine its response to various stresses. Additionally, it should be considered that certain composite materials can be excellent alternatives in certain demanding applications. This Doctoral Thesis examines the ability of finite element simulations, artificial neural networks and techniques based on machine learning and artificial intelligence to predict the behavior and properties of metallic materials and polymeric matrix composites. These tools can be used to extract trends in the data that allow generating new information about other materials and, therefore, serve as support for decision-making. It has been shown that there are very demanding applications where composite materials can replace metallic materials and a predictive model, based on artificial intelligence, has been developed to predict some fundamental properties of metallic alloys.

Capítulo 1. Introducción

La industria y los usuarios demandan, cada día, materiales más avanzados con los que sea posible cumplir con la creciente demanda de recursos y nuevas aplicaciones, al tiempo que se busca un compromiso entre comportamiento, durabilidad, coste y sostenibilidad [1]. Muchos de los grandes avances que están teniendo lugar en los últimos tiempos están íntimamente ligados a nuevos materiales con propiedades avanzadas [2]: materiales más ligeros, más resistentes y capaces de soportar entornos de operación más hostiles, entre otras características.

Los metales son una parte indispensable de nuestra vida y, cada día, surgen nuevos usos industriales de estos materiales debido a que presentan un amplio espectro de características. Por ello, es necesario disponer de herramientas que permitan investigar y desarrollar nuevas aleaciones con propiedades optimizadas para cada necesidad [3, 4]. Las propiedades mecánicas de un material juegan un papel fundamental en el desempeño de los componentes industriales.

Por otra parte, los materiales compuestos son ya una parte muy destacada en algunas industrias de alta tecnología como la aeronáutica [5], la aeroespacial o la alta competición [6]. Debido a la forma en que se diseñan este tipo de materiales, es posible fabricarlos con propiedades optimizadas para cumplir con la misión y los requisitos que sean precisos [7]. Los materiales compuestos también se están convirtiendo en una opción muy atractiva en algunos sectores industriales de alta exigencia [8].

Los diseñadores, científicos e ingenieros deben hacer frente a importantes retos en la selección de materiales, especialmente para aplicaciones de alta exigencia, donde se exploran los límites de los materiales de los componentes sometidos a condiciones críticas [9]. En estas circunstancias, es fundamental disponer de herramientas confiables que nos permitan hacer frente a estos retos [10]. Los materiales con que se fabrican algunos componentes de sectores tan exigentes como el sector nuclear, de generación de energía eléctrica, están sometidos a unas condiciones de funcionamiento que exigen emplear el material más adecuado para cada aplicación, cumpliendo con estrictos requisitos funcionales [11].

La creciente demanda de energía y recursos hace que los científicos se vean obligados a buscar nuevas soluciones que permitan hacer frente a los retos futuros y posibiliten que se puedan construir equipos industriales capaces de funcionar, de forma continua, en condiciones más adversas sin que sus componentes fallen [11]. En estas circunstancias, es fundamental contar con buenas herramientas de selección de materiales y apoyo a la decisión [12].

1.1 Necesidad de desarrollo de nuevas técnicas para la selección de materiales

Un correcto comportamiento en servicio depende, en gran medida, de las características de los materiales que lo constituyen, ya que unas propiedades inadecuadas o insuficientes pueden provocar fallos prematuros [12, 13]. Por tanto, la decisión de elegir un material específico para fabricar un componente industrial afecta en gran medida a su capacidad para soportar los esfuerzos y desempeñar la función para la que fue diseñado [14, 15].

Existen miles de materiales de interés industrial, aunque solo algunas de ellas se usan profusamente [16, 17], en algunos casos, debido a las dificultades para encontrar nuevas soluciones y, en otros, porque son materiales específicos con características perseguidas *ex profeso* en las etapas de diseño de acuerdo con la aplicación en la que desempeñarán su función.

Conocer las propiedades de los materiales empleados en los diseños industriales es fundamental; sin embargo, la obtención de estos datos, a menudo, requiere de grandes cantidades de recursos que, normalmente, no están disponibles. Se necesitan multitud de ensayos para obtener información significativa, lo cual implica que se debe disponer de tiempo, personal e instalaciones suficientes con los costes asociados que ello conlleva [12]. El proceso de caracterización de un material puede implicar numerosos ensayos y pruebas, que demandan una cantidad considerable de tiempo y la inversión de muchos recursos [18].

A pesar de que existen múltiples sistemas de apoyo a la toma de decisiones y metodologías de selección aplicadas a la ciencia, tecnología y procesado de los materiales [19], aún no están tan extendidas las técnicas basadas en el uso de inteligencia artificial en estos campos [3, 12, 20, 21]; ya que son más habituales otros métodos, tanto analíticos [9] como experimentales. Por otra parte, no existen trabajos donde se combinen estos estudios con el análisis particular de alternativas construidas a partir de materiales compuestos para una aplicación tan exigente como es el entorno de un reactor nuclear [9].

1.2 Materiales metálicos y propiedades para su empleo en aplicaciones de alta exigencia

Las aleaciones metálicas constituyen uno de los grupos de materiales más ampliamente empleados en la actualidad y juegan un papel fundamental en el mundo industrial debido a sus excelentes propiedades [18]. En general, se trata de materiales que presentan una buena conductividad térmica y eléctrica, poseen una alta densidad y una alta resistencia mecánica [22]. No obstante, se debe tener en cuenta que sus propiedades son muy variadas debido a que existen miles de aleaciones metálicas [12].

Conocer las propiedades de los materiales empleados en los diseños industriales es fundamental; sin embargo, la obtención de estos datos a menudo implica acceder a grandes cantidades de recursos, que normalmente no están disponibles. Se necesitan multitud de pruebas para obtener información significativa, lo que implica que se debe disponer de suficiente tiempo, personal e instalaciones a un coste razonable [3, 12]. El proceso de caracterización de un material puede implicar una gran cantidad de pruebas que requieren tiempo y la inversión de muchos recursos [18].

A pesar de que existen múltiples sistemas de apoyo a la toma de decisiones y metodologías de selección de materiales [19] aplicados a la ciencia de los materiales, las referencias que emplean tecnologías basadas en

inteligencia artificial en el campo del procesamiento e ingeniería de metales son más escasas [12, 20, 21]. Aunque existen muchos estudios que utilizan técnicas de aprendizaje automático para investigar la microestructura de los metales [23, 24], apenas existen referencias con un enfoque industrial en las que se trate la prognosis de las propiedades mecánicas de las aleaciones metálicas [3, 12, 25, 26].

En esta Tesis Doctoral se profundiza en la capacidad de las redes neuronales artificiales (RNA) para predecir, gracias al aprendizaje automático, algunas de las propiedades más importantes que caracterizan el comportamiento mecánico de un material [1]: la densidad (ρ), el módulo Young (E), la tensión de fluencia (YS), la resistencia a tracción (*Ultimate Tensile Strength - UTS*) y la elongación máxima (A) [3, 12, 25] . De esta forma, es posible realizar una prognosis del comportamiento de un elemento industrial ya existente o se puede estudiar la viabilidad de un nuevo material sin ser necesaria la realización de un extenso protocolo de ensayos, como suele ser habitual en la industria [21, 26].

La densidad es una propiedad física que determina en gran medida el comportamiento de un material y está íntimamente ligada con su química fundamental [11]. Por otra parte, las otras cuatro propiedades pueden ser empleadas para definir el modelo bilineal de la curva tensión-deformación de un material y, por tanto, definen el comportamiento elasto-plástico de la aleación [3].

La curva tensión-deformación contiene una gran cantidad de información sobre el material ya que muestra, de forma sencilla, las tensiones necesarias para deformar un material cuando se somete a cargas bajo diferentes estados tensionales. En este tipo de gráfico, la tensión se representa en el eje de ordenadas y su deformación correspondiente en el eje de abscisas [27]. El ensayo de tracción es el ensayo de caracterización mecánica por excelencia y proporciona información sobre la resistencia y ductilidad de los materiales bajo tensiones de tracción uniaxiales [11, 27, 28] . Esta información puede ser útil en comparaciones de materiales, desarrollo de aleaciones, control de calidad, simulación numérica mediante modelado con elementos finitos y diseño en determinadas circunstancias [28].

La curva tensión-deformación es una herramienta crucial y existen varios métodos de ensayo estándar para caracterizarla, como el ensayo de tracción, de compresión y de torsión [27, 28].

La curva de tensión-deformación también indica la cantidad de energía que un material puede almacenar antes de fracturarse [29], ya que el área encerrada debajo de la curva es la energía que el material absorbe durante su deformación [11, 30]. La energía que absorbe un material se llama resiliencia si la deformación es elástica y tenacidad si la deformación es plástica. Esta energía se puede calcular mediante la ecuación (1):

$$U = U_r + T = \int_0^A \sigma \cdot d\varepsilon \tag{1}$$

, donde U es la energía total de deformación (energía absorbida), U_r es la resiliencia, T es la tenacidad, A es la elongación a rotura, σ es la tensión y ε es la deformación.

La transición del comportamiento elástico al plástico se especifica a través de lo que se conoce como límite elástico o tensión de fluencia inicial, siendo complicado establecer su ubicación de manera exacta [31]. Por ello, en ingeniería se ha normalizado un criterio que garantiza la reproducibilidad de las pruebas: la definición del límite elástico convencional, considerado como la tensión a la cual el material tiene una deformación plástica del 0.2% [10].

Por otro lado, la curva tensión-deformación de las aleaciones metálicas es de especial relevancia, no sólo para estimar las propiedades óptimas para las aplicaciones en que suelen emplearse estos materiales, sino también

de cara a estudiar su procesado mediante técnicas como el conformado plástico. El comportamiento plástico de las aleaciones metálicas determina en gran medida su capacidad para experimentar las deformaciones permanentes que se producen durante las operaciones de conformado. De hecho, para el estudio de este tipo de procesos (como son la forja, laminación, extrusión, estirado), se emplean, típicamente, la simulación por elementos finitos [32, 33] y otros métodos analíticos [34], siendo en ambos casos de gran importancia tener caracterizadas correctamente lo que se conoce como curvas de fluencia del material, que pueden obtenerse a partir de las curvas tensión-deformación. En aplicaciones de ingeniería, el comportamiento plástico se ha caracterizado, tradicionalmente, mediante modelos simplificados, para facilitar el manejo de los datos y la intercomparación entre materiales. Además de modelos bien conocidos como la ley potencial de *Hollomon* y el modelo *Swift* [22], una aproximación lineal empleando el límite elástico (YS) y la resistencia máxima a la tracción (UTS) es la forma más sencilla de modelar el comportamiento plástico, especialmente cuando el endurecimiento por deformación tiene un papel importante, como en las condiciones de conformado en frío.

La forma de la curva tensión-deformación (real o aproximada) de una aleación metálica y sus valores dependen de [3, 27]:

- Composición química de la aleación.
- Tratamiento térmico de partida.
- Historia previa de deformación plástica.
- La tasa de deformación de la prueba.
- Temperatura.
- Orientación de la tensión aplicada en relación con la estructura de las probetas.
- Tamaño y forma.

Los últimos cuatro parámetros se describen en las normas pertinentes [28, 35]. Los tres primeros parámetros son los que se consideran en esta Tesis Doctoral, por ser los más relevantes y, por tanto, de los que más información se dispone en abierto para el empleo de técnicas de inteligencia artificial.

1.3 Materiales compuestos de matriz polimérica y propiedades para su empleo en aplicaciones de alta exigencia

Un material compuesto está formado por, al menos, dos componentes no miscibles que presentan una alta capacidad de penetración o unión entre ellos y cuyas propiedades son complementarias [7]. De esta forma, el nuevo material heterogéneo posee unas características intermedias entre las de sus componentes que dependen de la fracción volumétrica de cada fase y de su geometría [36].

En general, en los materiales compuestos es posible identificar dos fases o componentes principales: una fase discontinua y más rígida llamada refuerzo; y otra fase continua y, normalmente, menos rígida, que se denomina matriz [36].

Por una parte, la principal función del refuerzo consiste en absorber las tensiones e incrementar la rigidez y resistencia del conjunto, por tanto, se trata del elemento resistente. Por otra parte, la matriz sirve de transmisor de tensiones entre los elementos de refuerzo, funciona como ligante manteniendo el refuerzo en una posición fija y lo protege del medio exterior [7].

En general, los materiales compuestos son muy anisótropos, debido, fundamentalmente, a la falta de homogeneidad en la distribución y colocación de los elementos de refuerzo dentro de la matriz [37]. La resistencia de los materiales compuestos depende, en gran medida, de la naturaleza de las fibras utilizadas

como elementos de refuerzo, de su disposición y orientación en la matriz, y de la efectividad de la interfaz fibra/matriz [7].

Los materiales compuestos ya se aplican en casi todos los sectores industriales debido a su amplia gama de propiedades [37] y son de especial interés en aplicaciones de alta exigencia. Generalmente, los materiales compuestos tienen relaciones de resistencia y módulo a densidad más altas que los materiales de ingeniería tradicionales. Estas características pueden reducir el peso de un sistema entre un 20 y un 30% [30, 31]. La reducción del peso se traduce en ahorro de energía y mayor rendimiento. Los compuestos avanzados exhiben propiedades dinámicas deseables y tienen una alta resistencia a la fluencia y buenas características de amortiguación [37]. De hecho, el rendimiento superior a la fatiga de los materiales compuestos les permite ser utilizados para reparar fuselajes metálicos dañados por este mecanismo de degradación [7].

La ingeniería de materiales compuestos necesita enfoques sistemáticos e interactivos, que permiten lograr las características óptimas para los materiales con que se diseña cada componente industrial [5, 9]. Este proceso requiere la aplicación de varios métodos y tecnologías dirigidas a la investigación de las propiedades físicas y mecánicas de cada componente, así como del material compuesto; la optimización de las propiedades del material compuesto según las condiciones específicas de trabajo; la comprensión de los efectos generados por el procesado y la composición sobre las propiedades del material compuesto; y el desarrollo de métodos computacionales para la caracterización, análisis y predicción del desempeño de materiales en diferentes condiciones de trabajo [7, 36]. Esto es especialmente importante en el caso de materiales compuestos ya que sus propiedades son muy heterogéneas.

Los materiales metálicos y compuestos pueden cohabitar para aumentar la eficiencia general de un sistema. Todavía hay espacio para la evolución de los materiales conocidos, pero los nuevos desarrollos se unen al panel de opciones de los diseñadores y hacen más complicado el problema de decisión [10]. No obstante, estas nuevas opciones permiten generar diseños más eficientes y optimizados capaces de cumplir mejor con su cometido [30].

1.4 Inteligencia artificial y redes neuronales artificiales

La inteligencia artificial (IA) es la simulación de procesos de inteligencia humana por máquinas, especialmente por sistemas informáticos [1]. Estos procesos comprenden la autocorrección (detectar errores y resolverlos), el razonamiento (utilizar reglas para llegar a nuevas conclusiones y conocimiento) y el aprendizaje (adquirir procedimientos para emplear la información) [12, 38]. En la actualidad, IA es un término de amplio alcance que, recientemente, ha ganado gran importancia debido al aumento de velocidad, tamaño y variedad de los datos recopilados por las empresas [38]. La IA puede realizar tareas, como el reconocimiento de patrones en los datos de manera más eficiente que los humanos, lo que permite a los usuarios extraer más información de sus conjuntos de datos [20].

La IA es un término que engloba una multitud de técnicas y tecnologías destinadas a dotar a una máquina de la capacidad de presentar un comportamiento inteligente[1] [39]. Dentro de estas técnicas, podemos encontrar modelos matemáticos simples (aunque potentes) como árboles de decisión, capaces de categorizar datos [1]; y

[1] Se entiende por comportamiento inteligente la capacidad que exhiben estos sistemas para simular el comportamiento humano ante la toma de decisiones.

otras herramientas mucho más complejas y avanzadas como las redes neuronales convolucionales profundas, capaces de identificar imágenes y patrones [40].

La IA ha demostrado que se puede aplicar a una multitud de disciplinas que no están únicamente relacionadas con la informática o la robótica. Entre los nuevos usos más relevantes, se pueden destacar la medicina [41, 42], los procesos bélicos [43], la ecología [44], la seguridad [45], la educación [46], la exploración petrolera [47] o la ciencia de los materiales [48]. La IA se puede aplicar a casi todas las ramas de la ciencia y la ingeniería, y cada día surgen nuevos usos y aplicaciones [25, 39].

La Industria 4.0 conlleva la digitalización de las cadenas de valor gracias a la integración de nuevas tecnologías de procesamiento de datos, software basado en agentes inteligentes y sensores [39]. La inteligencia artificial ya juega un papel fundamental en la concepción de las fábricas y las cadenas de producción del futuro, cuya optimización está basada en el uso y gestión inmediato de la información [1]. Entre todas las herramientas incluidas dentro del campo de la inteligencia artificial, se deben destacar las redes neuronales artificiales multicapa (RNAM) debido a su actual relevancia y capacidades [49]. Una red multicapa es un algoritmo de aprendizaje supervisado que, mediante el entrenamiento en un conjunto de datos etiquetado, es capaz de aprender una función no lineal que se puede utilizar para realizar clasificaciones y regresiones [50]. Las redes neuronales multicapa están formadas por perceptrones (neuronas artificiales) que se organizan creando capas que se comunican entre sí (en general, los perceptrones no se conectan con sus propios compañeros de capa) [12].

Teniendo en cuenta la topología de conexión de los perceptrones, se pueden definir tres tipos de capas: capa de entrada, que incluye todos los perceptrones que reciben datos de una fuente externa; capa de salida, que incluye todos los perceptrones que devuelven resultados; y capa oculta, que incluye todos los demás perceptrones, que no se comunica con el exterior de la red [51].

Se puede entrenar una red neuronal multicapa para aprender una función no lineal [22] de la forma de la ecuación (2):

$$F(X): \mathbb{R}^m \rightarrow \mathbb{R}^o \tag{2}$$

, donde $X = \{x_i \ / \ i \in 1 \dots m\}$ es el vector de entrada, m es el tamaño del vector de entrada y o es el tamaño del vector de salida [1].

El procedimiento de aprendizaje de la red neuronal se conoce como entrenamiento y, matemáticamente, se basa en el problema del descenso del gradiente para minimizar la función de error asociada [39]. Esa función de error depende de los pesos relativos de cada uno de los perceptrones. Este vector de pesos (cuyo tamaño es igual al número de neuronas en la red) se representa como w, de tal forma que se puede indicar que $f(w)$ es la función de error cuando se asignan los pesos w a cada uno de los perceptrones de la red. Con esta formalización, el objetivo del entrenamiento es encontrar el vector w^* para el cual se obtiene un mínimo global de la función f, lo que convierte el problema de aprendizaje en un problema de optimización [12].

De esta forma, en la primera iteración, se asignan valores, generalmente, aleatorios al vector de pesos de la red neuronal y, luego, se calcula un nuevo vector para reducir la función de error [39]. Este proceso se repite hasta que el error se reduzca por debajo de cierto límite o hasta que se cumpla una condición de parada específica. Dado que la función de error es diferenciable, el gradiente de esta función se puede definir para cada uno de los pasos de optimización (ver ecuación (3)) [12]:

$$g_i = \nabla f_i = \nabla f(w_i) \tag{3}$$

, donde g_i es el valor de gradiente de la función de error en el i-ésimo paso de la iteración, f_i es el valor de la función de error en el i-ésimo paso y w_i es el vector de pesos en la i-ésima iteración.

ADAM (*Adaptive Moment Estimation*) es una metodología con tasa de aprendizaje adaptativa que utiliza las estimaciones del primer y segundo momento del gradiente para adaptar la tasa de aprendizaje para cada peso de la red neuronal [52]. Usando este método, en cada iteración, el nuevo vector de peso se calcula según la ecuación (4) [52]:

$$w_{i+1} = w_i - \eta \frac{\widehat{m}_{i+1}}{\sqrt{\hat{v}_{i+1}} + \epsilon} \tag{4}$$

, donde η es el tamaño del paso (un valor que gradúa la relevancia del factor de gradiente), ϵ es el factor de estabilidad del algoritmo (constante) y $\widehat{m}_{i+1}$ y $\hat{v}_{i+1}$ son las estimaciones con sesgo corregido del primer y segundo momento, que se calculan de la siguiente manera (ver ecuaciones (5) y (6)) [52]:

$$\widehat{m}_{i+1} = \frac{m_{i+1}}{1 - \beta_1^{i+1}} \tag{5}$$

$$\hat{v}_{i+1} = \frac{v_{i+1}}{1 - \beta_2^{i+1}} \tag{6}$$

, donde β_1 y β_2 son parámetros del algoritmo con valor cercano a 1 [52]; m_{i+1} y v_{i+1} se calculan de la siguiente manera (ver ecuaciones (7) y (8)) [52]:

$$m_{i+1} = \beta_1 m_i + (1 - \beta_1) g_{i+1} \tag{7}$$

$$v_{i+1} = \beta_2 v_i + (1 - \beta_2) {g_{i+1}}^2 \tag{8}$$

, donde m_{i+1} y v_{i+1} son, respectivamente, las tasas de decrecimiento del gradiente y del gradiente al cuadrado, los cuales son estimadores del primer momento (media) y del segundo momento (varianza no centrada) de los gradientes [52].

Con todo lo indicado hasta este punto, el proceso de optimización y el método de entrenamiento de la red se han definido matemáticamente.

Una vez que la red ha sido entrenada convenientemente, se pueden obtener predicciones basadas en la función de aproximación aprendida por la red neuronal [12]. La desviación de la predicción se calcula como el valor absoluto del error relativo del valor resultante (ver ecuación (9)).

$$\varepsilon = \left| \frac{v_{predicción} - v_{real}}{v_{real}} \right| \tag{9}$$

, donde ε es el error predictivo relativo (en valor absoluto), $v_{predicción}$ es el valor estimado (valor resultante de la red) y v_{real} es el valor real.

Los nodos de una red neuronal artificial se pueden conectar de muchas maneras, formando diferentes topologías de red. El comportamiento del sistema, su capacidad de aprendizaje y la cantidad de recursos que necesitará durante las fases de entrenamiento y predicción depende en gran medida de la topología elegida [50]. Una red neuronal artificial completamente conectada consiste en un conjunto de capas plenamente conexas, es decir, aquellas en que todos los nodos están conectados a todos los nodos de la siguiente capa [39].

Para una red neuronal multicapa completamente conectada, la complejidad del proceso de entrenamiento por retro-propagación viene dada por la ecuación (10). Por lo tanto, es muy recomendable minimizar el número de nodos ocultos para reducir el tiempo de entrenamiento [50].

$$\mathcal{O}\left(n \cdot m \cdot o \cdot N \cdot \prod_{i=1}^{k} h_i\right) \tag{10}$$

, donde n es el tamaño del conjunto de datos de entrenamiento, m es el número de características, o es el número de perceptrones de salida, N es el número de iteraciones y k es el número de capas ocultas (cada una de ellas contiene h_i nodos).

1.5 Big data en tecnología de materiales y su procesado

La ciencia e ingeniería de los materiales depende de experimentos y modelos basados en simulaciones para comprender la física de los materiales con el fin de conocer mejor sus características y descubrir otros nuevos con propiedades mejoradas [53]. Todos estos experimentos y simulaciones generan una gran cantidad de datos, que son cada vez más difíciles de manejar utilizando técnicas tradicionales de procesamiento de datos [1, 12]. Debido al enorme volumen de datos que se producen, a una velocidad sin precedentes, estos datos no se procesan de manera eficaz para crear información, lo cual retrasa la producción de nuevo conocimiento [54].

Tradicionalmente, el conocimiento se ha organizado a través de la denominada pirámide del conocimiento o jerarquía de la información [12]. Este modelo consta de cuatro pasos, cada uno de los cuales se deriva del anterior: datos, información, conocimiento y sabiduría (DIKW) [55]. De esta manera, los datos procesados constituyen información, que se organiza para generar conocimiento, que, finalmente, se resume en sabiduría [56].

Nuestra tecnología actual ha alcanzado un nivel nunca antes visto en términos de generación de datos [53]; sin embargo, las técnicas destinadas a su procesamiento aún no están tan avanzadas y su uso no está muy extendido [1, 12]. Por lo tanto, tenemos que hacer frente al reto de transformar datos en información y conocimiento. Extraer valor de los datos brutos requiere un enfoque sistemático y bien definido para resolver estos problemas emergentes del mundo real y, por lo tanto, se necesita un nuevo enfoque multidisciplinar [25] [54].

Dos de los principales factores que contribuyen al auge de la era del *big data* incluyen los rápidos avances en la tecnología de la información y la consiguiente explosión de datos: el primero se deriva de la mejora de la tecnología de procesamiento, que incluye CPU más rápida, mayor ancho de banda de red y la evolución de la tecnología del software; el último se deriva del uso generalizado de sensores [57]. Estos factores han traído una repentina explosión de la cantidad de datos disponible y han contribuido al auge de la tecnología de *big data* [53]. En cualquier campo de conocimiento, los conjuntos de datos se consideran "*big*" cuando se presentan en numerosos, son complejos y son difíciles de procesar y analizar. Los datos en la ciencia de materiales tienden a ser particularmente heterogéneos en términos de tipo y fuente [12].

Uno de los primeros pasos en el procesamiento de grandes conjuntos de datos es la reducción [2]. Por ejemplo, los experimentos en el Gran Colisionador de Hadrones (*Large Hadron Collider-LHC*) retienen solo una pequeña fracción del 1% de los datos que producen porque se vuelve impracticable con la tecnología actual almacenar y analizar más de los cientos de megabytes por segundo que se consideran más valiosos: depende de un software sofisticado determinar qué datos son los más relevantes [58].

Aunque el término "*big data*" es relativamente nuevo, la acción de recopilar y almacenar grandes cantidades de información para su posterior análisis se ha realizado durante muchos años. La definición actual de macrodatos se basa en los tres pilares siguientes [12, 59]:

- Volumen: se procesan grandes cantidades de datos no estructurados de baja densidad. Los datos pueden ser de valor desconocido, como las condiciones de mecanizado, las propiedades del material o las medidas de control de fabricación.
- Velocidad: se reciben datos a gran velocidad y, posiblemente, se aplica alguna acción.
- Variedad: los tipos de datos convencionales están estructurados y pueden organizarse claramente en una base de datos relacional; sin embargo, los macrodatos se presentan como registros dispersos no estructurados.

Los macrodatos permiten obtener respuestas más completas al disponer de más información [60]. Disponer de soluciones más completas a un problema permite abordarlo con mayores garantías de éxito y permite gestionar los recursos de forma más eficiente [61]. La disponibilidad de respuestas más completas también significa una mayor confiabilidad de la información, lo que implica un enfoque completamente diferente para abordar los problemas.

Los macrodatos han demostrado ser una herramienta muy útil en la investigación en ciencia y tecnología de materiales. En los últimos años se han desarrollado muchos trabajos sobre este tema, han surgido numerosas aplicaciones y han aparecido importantes iniciativas destinadas a servir de base a los científicos en este campo. Se han desarrollado modelos predictivos basados en *big data* y aprendizaje automático [62], se han desarrollado herramientas de diseño asistido por ordenador soportadas por grandes bases de datos que transforman el software en un sistema experto [63] y han aparecido entidades como la *Material Genome Initiative* [64]. La misión de esta iniciativa es reducir el costo y el tiempo de desarrollo de los descubrimientos, la optimización y el despliegue de materiales ofreciendo acceso a grandes bases de datos de materiales y proporcionando las herramientas necesarias para facilitar la investigación.

Las bibliotecas de materiales disponibles en Internet contienen grandes colecciones de registros que, en general, están incompletas; es decir, algunas propiedades no están disponibles para un material [65], lo que dificulta la posibilidad de obtener conclusiones válidas utilizando herramientas convencionales [66]. Aunque algunas bibliotecas de materiales son de libre acceso [65], en general, descargar grandes cantidades de datos es difícil ya que las empresas guardan esta información tan importante con gran celo. Algunas bibliotecas de materiales disponibles son: *Matmatch, MatWeb, MatDAT* o *Material Project*.

En parte de esta Tesis Doctoral, los datos contenidos en estas bibliotecas se han utilizado para entrenar un sistema basado en técnicas de inteligencia artificial (IA) para extraer nueva información que inicialmente no estaba disponible. Téngase en cuenta que la calidad de las conclusiones de este tipo de metodología es tan buena como la calidad de los datos de los que se han extraído [13].

Matmatch [67] es una conocida y consolidada biblioteca de materiales de acceso abierto que contiene información sobre miles de materiales comerciales y estándar [12, 25]. Los usuarios registrados pueden acceder libremente a la información almacenada en las bases de datos y pueden descargar, para cada material, una ficha técnica que contiene todos los datos disponibles [12]. En esta biblioteca, están archivados datos muy heterogéneos acerca de más de 100k materiales [66]. Estos datos son proporcionados por los fabricantes y proveedores de los materiales y, aunque son muy precisos, deben procesarse, filtrarse y analizarse para generar

un corpus de información útil y significativa [56]. Para la realización de esta Tesis Doctoral se han empleado los datos de *Matmatch* para llevar a cabo los estudios descritos en [3, 12, 25].

1.6 Justificación de la unidad temática de la Tesis

Esta Tesis Doctoral comienza abordando el problema de encontrar nuevas soluciones para una aplicación de alta exigencia como es un componente presurizado de una central nuclear [9, 68], en un entorno que expone al material a moderadas presiones y temperaturas y a radiación (gamma y neutrónica). Se aborda mediante simulación numérica, por el método de los elementos finitos, la posibilidad de fabricar este tipo de tuberías empleando materiales compuestos de matriz polimérica (frente al empleo de acero), aplicando una metodología de selección de materiales basada en niveles de severidad [69]. Tal y como se demuestra en este primer trabajo, una cuidada selección de materiales basada en sus propiedades y comportamiento permite que los materiales compuestos puedan ser una opción válida para sustituir a los elementos metálicos en determinadas aplicaciones.

Una vez empleada la simulación numérica para la selección de materiales compuestos de matriz polimérica en aplicaciones de alta exigencia, esta Tesis examina las herramientas disponibles para aplicar las tecnologías basadas en inteligencia artificial y el aprendizaje automático a la selección de materiales [1]. Aunque existen muchos métodos muy potentes, las redes neuronales basadas en aprendizaje supervisado son especialmente atractivas y han demostrado su potencial como elementos disruptivos en otros muchos campos [51, 53].

En un primer paso, se examina la capacidad de las redes neuronales multicapa para realizar predicciones sobre la densidad y el módulo de Young de una gran colección de materiales metálicos [12]. Esta red neuronal es entrenada para predecir el valor de estas propiedades físicas y mecánicas tomando como dato de partida la composición química de la aleación metálica en cuestión.

A continuación, se demuestra que estas técnicas avanzadas de análisis de datos también permiten entrenar una RNAM para aproximar la curva tensión-deformación de las aleaciones de aluminio a partir de la composición química y el tratamiento (térmico o mecánico) a que ha sido sometido el metal en cuestión [3]. Se realiza una aproximación bilineal a la curva basándose en el valor de cuatro propiedades mecánicas muy importantes: el módulo de Young, la tensión de fluencia, la tensión última y la elongación a rotura [22, 62].

Finalmente, se aborda la posibilidad de refinar los modelos predictivos mediante la incorporación de nueva información a la metodología de análisis de datos: se añade, como dato de entrada adicional, la dureza del material [25]. Esta información suplementaria permite llevar a cabo predicciones más precisas (y, por tanto, más útiles) acerca de la tensión de fluencia y la tensión última de las aleaciones de aluminio. Los datos que requiere la metodología para realizar predicciones son el resultado de ensayos no destructivos que se pueden realizar con facilidad, incluso en componentes en servicio [14, 21].

Capítulo 2. Hipótesis y objetivos

2.1 Hipótesis

Esta Tesis Doctoral se cimienta sobre la gran importancia que tienen las propiedades de los materiales con que se fabrican los componentes industriales para que puedan desempeñar adecuadamente su cometido; esto es especialmente importante en entornos de alta exigencia, donde los materiales están sometidos a solicitaciones más severas y, por tanto, conocer con precisión sus características y comportamiento es más crítico. Este estudio, se aborda mediante diferentes técnicas de análisis avanzado de datos tales como el método de los elementos finitos o las metodologías basadas en inteligencia artificial y aprendizaje automático.

Las principales hipótesis de este trabajo son:

- Los entornos de alta exigencia requieren nuevas soluciones en el campo de la ciencia y tecnología de materiales que pueden estar basadas en el uso de materiales compuestos para aplicaciones donde, tradicionalmente, se emplean metales [9].
- La simulación mediante el método de los elementos finitos es una herramienta de gran potencial para la evaluación de materiales alternativos en el diseño de componentes industriales
- Las técnicas basadas en inteligencia artificial y aprendizaje automático pueden ser empleadas como apoyo a la decisión en problemas de selección de materiales [1].
- Las técnicas de inteligencia artificial basadas en aprendizaje supervisado (y, especialmente, las redes neuronales artificiales multicapa) son capaces, mediante entrenamiento, de aprender una función de regresión de datos que relaciona las propiedades químicas, mecánicas y físicas de una aleación metálica.
- La composición química y el tratamiento (térmico o mecánico) al que se ha sometido una aleación metálica están relacionados con sus propiedades físicas y mecánicas y con su comportamiento a tracción.

2.2 Objetivos

El objetivo fundamental de esta Tesis Doctoral consiste en estudiar las metodologías basadas en análisis avanzado de datos más prometedoras para predecir las propiedades y el comportamiento de materiales candidatos para la fabricación de componentes destinados a aplicaciones de alta exigencia. Los principales objetivos de este trabajo y las publicaciones donde se han alcanzado dichos objetivos son:

- Estudiar metodologías avanzadas basadas en análisis de datos que pueden ser empleadas para realizar predicciones sobre las propiedades de un material [1, 3, 12, 25].
- Estudiar la posibilidad de sustituir componentes metálicos por otros fabricados con materiales compuestos para aplicarlos en entornos de alta exigencia [9].
- Establecer un marco metodológico que permita comparar el desempeño, en entornos de alta exigencia, de un acero y de un material compuesto [9].
- Examinar las distintas herramientas que se pueden englobar dentro de la rama de la inteligencia artificial y entender cuáles de ellas pueden ser más prometedoras y cuales se podrían emplear para obtener resultados de utilidad en la selección de materiales [1].
- Establecer una metodología de trabajo que permita gestionar y tratar grandes bases de datos y colecciones *big data* acerca de propiedad de materiales para generar información útil [3, 12, 25].
- Analizar la capacidad de las redes neuronales para predecir la densidad de una aleación metálica tomando como dato de entrada, únicamente, su composición química [12].
- Analizar la capacidad de las redes neuronales artificiales para predecir, a través de aprendizaje supervisado, las propiedades que definen el comportamiento elasto-plástico de un material sometido a tracción: módulo de Young (*E*) [3], tensión de fluencia (*YS*) [3, 25], resistencia a tracción (*UTS*) [3, 25] y elongación máxima (*A*) [3].
- Examinar la bondad de los modelos predictivos basados en redes neuronales artificiales comparándolos con otras metodologías clásicas descritas en las referencias bibliográficas [3, 25].
- Estudiar el conjunto de datos de partida mínimo necesario para que el proceso de entrenamiento de la red neuronal artificial converja y, además, las posteriores predicciones que se lleven a cabo usando dicha red tengan una tasa de error baja [3, 12, 25].
- Evaluar la capacidad de las redes neuronales para aprender las relaciones existentes entre las diferentes propiedades físicas, mecánicas y químicas que se emplean como datos de partida [3, 12, 25].
- Examinar el potencial de las redes neuronales artificiales para ser usadas como herramientas de apoyo a la decisión en la selección de materiales mediante la construcción de diagramas de *Ashby* [12].

Capítulo 3. Metodología

Cada una de las publicaciones describe, en detalle, la metodología y el marco teórico sobre el cual se asienta cada uno de los trabajos. La propuesta metodológica de cada uno de los artículos trata de responder al interrogante y los objetivos que ha motivado cada uno de los artículos. En cualquier caso, se debe resaltar, como hilo conductor entre los diferentes métodos propuestos, la gran importancia del análisis de los datos obtenidos. La analítica avanzada de datos permite sustentar ideas y rechazar aquellos conceptos que no sean pertinentes.

A continuación, se describen detalladamente las bases metodológicas del trabajo de investigación desarrollado a través de esta Tesis Doctoral.

3.1 Método basado en simulación numérica

El método de los elementos finitos (MEF) es uno de los métodos numéricos más utilizados para resolver problemas de ingeniería. Una metodología basada en el método de los elementos finitos para la simulación del comportamiento de un componente industrial conlleva múltiples pasos (Figura 1). La robustez del modelo numérico desarrollado determina, en gran medida, la calidad del resultado final [70].

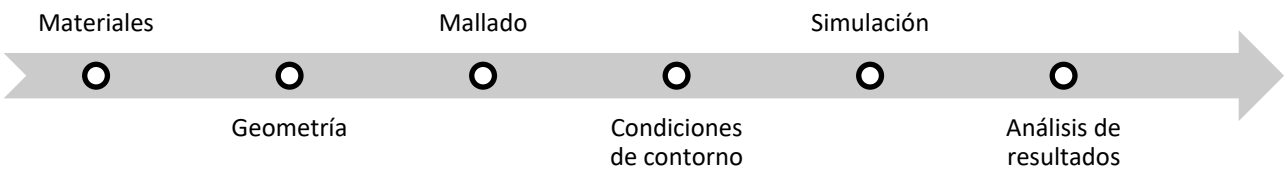


Figura 1. Pasos en el análisis por el método de los elementos finitos para simulación del comportamiento de un componente industrial

El proceso comienza con la definición de las propiedades de los materiales que se van a emplear para simular el comportamiento del componente. En este paso es especialmente importante la correcta caracterización del

comportamiento elasto-plástico de los materiales empleados y se definen los límites de fallo del mismo [9, 70]. La precisión y exactitud de esta información determina, en gran medida, el resultado final [71].

La definición geométrica del componente a simular se basa en un modelo teórico que, en general, prescinde de las imperfecciones inherentes a la fabricación [30]. Por tanto, si bien se intenta crear modelos geométricos de la calidad suficiente, éstos nunca pueden abarcar toda la casuística, fallos y defectos existentes en el mundo real. En cualquier caso, un modelo geométrico realista mejora notablemente la precisión de los resultados de la simulación [70].

Para resolver un problema mediante el MEF, se subdivide un continuo en un conjunto de elementos interconectados por medio de puntos denominados nodos; esta discretización permite pasar de un sistema continuo de infinitos grados de libertad a un sistema discreto con un número de grados de libertad finito, donde las ecuaciones del continuo rigen el comportamiento del elemento [70, 72]. La discretización del problema se implementa matemáticamente mediante la construcción de una malla. De esta forma, el dominio numérico de la solución tiene un número finito de puntos [73]. El mallado de un objeto que debe simularse numéricamente es un proceso que requiere adoptar un compromiso entre la precisión de la solución y los recursos necesarios: una malla muy fina y precisa requerirá muchos recursos para llevar a cabo la simulación, mientras que una malla más gruesa podría no representar correctamente la realidad del proceso que se está simulando [11].

Sobre este modelo discretizado se aplican las condiciones de contorno, las cuales incluyen: condiciones ambientales como la temperatura, las cargas a que está sometido y los apoyos sobre los que se asienta el sistema [70]. Las condiciones de contorno están directamente relacionadas con el resultado de la simulación ya que, en la medida en que estas se ajusten a la realidad del funcionamiento del sistema, la simulación será más precisa y útil.

Una vez que el modelo numérico está correctamente definido, es posible llevar a cabo la simulación numérica mediante el código apropiado. Los resultados de esta simulación permiten refinar y mejorar el modelo o extraer conclusiones válidas que sirvan para los objetivos del estudio [9]. En la medida en que el proceso completo de simulación se ajuste a la realidad, las predicciones que se alcancen serán más precisas [73].

Se ha utilizado *ANSYS Workbench* 14.5 para realizar la simulación de los componentes diseñados [9]. Se trata de un paquete comercial de herramientas de diseño y simulación que permite realizar estudios de fenómenos físicos en varias áreas: dinámica de fluidos, mecánica, electromagnetismo y otras. Cuenta con unos algoritmos de cálculo muy utilizados en el ámbito industrial y destaca por su intuitiva interfaz. Además, cuenta con algunos módulos específicos para el análisis y simulación de materiales compuestos.

3.2 Métodos basados en inteligencia artificial

Para desarrollar una metodología basada en inteligencia artificial, se requieren una gran cantidad de pasos y se involucran diferentes disciplinas: se ha desarrollado un extenso software [74] capaz de trabajar sin la intervención del usuario para descargar datos de una biblioteca de materiales en línea [67], filtrar y organizar datos, definir y entrenar las RNA [39] y, finalmente, hacer predicciones utilizando esas redes. Por otro lado, se ha necesitado mucho trabajo para analizar datos y definir criterios basados en la ciencia e ingeniería de los materiales [71]. Desarrollar el software capaz de obtener los datos es una de las fases más delicadas y a la que se le ha prestado mayor atención, ya que esos datos son la base sobre la que se elabora todo el estudio. A continuación, se describen las etapas principales.

3.2.1 Obtención de los datos de entrada y filtrado

Como ya se indicó, el conjunto de datos de entrada utilizado en este trabajo se obtuvo de una biblioteca de materiales de acceso abierto en línea (*Matmatch* [67]). En este portal web, es posible acceder a la información proporcionada por miles de fabricantes y proveedores de materiales de diferente índole, incluidas aleaciones metálicas [12].

Para cada material, los datos registrados pueden ser muy diversos y, en cualquier caso, cabe señalar que estos datos no son exhaustivos: no toda la información está disponible para todos los materiales ya que la tarea de registrar los datos de cada material depende de los propios comercializadores. En el campo del *big data*, es muy habitual tratar con información heterogénea y dispersa [59].

Esta biblioteca ofrece información acerca de más de 100k materiales diferentes [67]; incluidos varios miles de registros de aleaciones metálicas. Es posible acceder a una hoja de datos específica para cada material y descargarla; sin embargo, no es posible obtener un paquete completo con la información de múltiples materiales, sino que es necesario descargar los datos de cada material uno por uno [12].

Para llevar a cabo la tarea de descargar los datos brutos de los materiales empleados, se ha desarrollado una aplicación *Python* que es capaz de descargar secuencialmente las hojas de especificaciones [74, 75]. Cada registro se descarga como un documento de Excel que contiene toda la información disponible sobre el material.

Una vez descargadas las fichas técnicas de todas las aleaciones relevantes, la información contenida en estos archivos se lee e interpreta secuencialmente. Como ya se ha indicado, se dispone de muchos más datos de los necesarios para realizar este estudio [67] y, en cualquier caso, cada una de las publicaciones pone el acento, únicamente, en las propiedades que se van a emplear.

3.2.2 Definición de la red neuronal

Una vez filtrados los datos y garantizada su consistencia, se define la RNA que se encargará de realizar las predicciones: se ha elegido una arquitectura multicapa y una topología totalmente conectada [50]. Esta estructura consta de una capa de entrada, 3 capas ocultas (que contienen 100, 100 y 10 perceptrones respectivamente) y una capa de salida.

Esta topología es el resultado de sucesivos pasos de optimización para equilibrar su capacidad predictiva y los recursos necesarios para su entrenamiento [76]. Se debe tener en cuenta que una topología compleja es capaz de aprender funciones más complejas que una topología simple, pero requiere de más recursos durante su entrenamiento: tiempo, capacidad de cálculo y datos de entrada [12]. Con esta topología, se obtuvo un equilibrio entre la profundidad y el ancho de la red.

3.2.3 Entrenamiento y predicción

Una vez que los datos de entrada ya están disponibles y se define la topología de la red neuronal, comienza la fase de entrenamiento y predicción. Durante esta fase, cada una de las publicaciones tiene en cuenta únicamente las propiedades que se están estudiando. Para cada una de las propiedades, se realizan varias iteraciones de aprendizaje y predicción. Cada una de estas iteraciones (independientes entre sí) se subdivide en cuatro pasos:

- División del conjunto de datos de entrada: se divide, aleatoriamente, en dos subconjuntos disjuntos que contienen, respectivamente, 80% (subconjunto de entrenamiento) y 20% (subconjunto de prueba)

de los registros. Para evitar sesgos, no se deben utilizar los mismos datos durante el entrenamiento y durante las pruebas y el cálculo del rendimiento predictivo ya que podría producirse un sobreajuste y se obtendrían métricas incorrectas (resultados demasiado buenos) [50].

- Entrenamiento de las redes neuronales con el subconjunto de entrenamiento.
- Predicción sobre el subconjunto de prueba.
- Almacenamiento de datos para su posterior análisis.

Todo este proceso de entrenamiento y predicción genera una gran cantidad de información que proporciona evidencia muy significativa sobre el rendimiento y las capacidades de la red neuronal.

3.2.4 Análisis de datos

Una vez realizadas todas las iteraciones y estando disponible toda la información resultante, se lleva a cabo la fase de análisis. Se calculan una gran batería de métricas estadísticas y se realizan varios diagramas y gráficas para resumir tanto el entrenamiento como la fase de predicción. Esta información permite la discusión de los resultados obtenidos con la metodología descrita en cada trabajo.

La información más notable que se puede obtener del entrenamiento es la evolución de la función de error a lo largo del aprendizaje. Aunque el número de pasos no es significativo, es muy importante comprobar que la función de error converge asintóticamente a un valor relativamente bajo [50].

Por otro lado, el rendimiento del proceso de predicción se estima utilizando la desviación relativa (en valor absoluto) para cada muestra del subconjunto de prueba. Con esta información, es posible calcular diversos estimadores estadísticos y métricas que permiten conocer la bondad y corrección de la metodología completa. Además, es posible trazar gráficas que representen esta información.

Capítulo 4. Publicaciones

Para la realización de esta tesis por compendio de publicaciones, se ha elegido la primera opción descrita en el documento regulador de Tesis por compendio de publicaciones aprobado por el Comité de Dirección de la EIDUNED, en su reunión de 16 de enero de 2017, y por la Comisión de Investigación y Doctorado de la UNED, con fecha 21 de febrero de 2017:

> *«Un mínimo de 3 artículos (al menos, dos ya publicados y el tercero aceptado) en revistas de índices de impacto en los dos primeros cuartiles de la relación de revistas del ámbito de la especialidad del Programa en el que está inscrita dicha tesis y referenciadas en la última relación publicada por el Journal Citation Reports (SCI y/o SSCI) y de SCOPUS. Todos los artículos deben estar publicados con fecha posterior a la primera matrícula de tutela académica en la EIDUNED. El doctorando o doctoranda debe ser primer firmante o segundo, en este último caso, el primero debe ser el director o directora de la tesis.»*

A continuación, se resume la información principal de cada una de las publicaciones, incluyendo el factor de impacto, así como una copia de cada una de las publicaciones que comprenden el compendio.

4.1 *Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications* [9]

Los indicios de calidad de este artículo pueden encontrarse en el Apéndice A.

4.1.1 Datos de la publicación y factor de impacto

Tabla 1. Factor de impacto de Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications

Título	Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications
Autores	David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Revista	Proceedings of the Institution of Mechanical Engineers Part L-Journal of Materials-Design and Applications
ISSN	1464-4207
Editorial	SAGE Publications
País	Inglaterra
Volumen	233 (10)
Páginas	2072-2083
Fecha	2019
doi	10.1177/1464420718817334
Factor de impacto	2.104 (2019 Journal Citation Reports) 193/314 (Materials Science, Multidisciplinary, Q3)

4.1.2 Resumen y copia de la publicación

Este estudio describe una propuesta metodológica para seleccionar materiales compuestos adecuados para la fabricación de tuberías que puedan resistir apropiadamente entornos sometidos a radiación gamma y neutrónica. La metodología se utiliza para seleccionar el material compuesto óptimo, entre un conjunto de opciones, cuyas propiedades se utilizan posteriormente para simular varias secciones de tubería mediante análisis por el método de los Elementos Finitos. Los resultados se comparan con un acero habitualmente empleado en estos componentes en la industria nuclear, WWER 15Kh2MFAA. El material compuesto más adecuado según los criterios definidos está compuesto por una matriz de resina fenólica reforzada con fibras largas de boro y presenta interesantes propiedades para ser utilizado en el entorno de un reactor nuclear: buena resistencia a la radiación, buenas propiedades mecánicas con una densidad muy baja y a bajo costo. Se puede concluir que, en algunos casos, las tuberías de material compuesto pueden ser una mejor opción que las de acero.

This study describes a methodological proposal to select composite materials which are suitable to be employed to manufacture pipes that can properly withstand environments subjected to gamma and neutronic radiation. The methodology is used to select, among many others, the optimal composite material whose properties are used afterwards to simulate several pipe sections by Finite Element Analysis, comparing the results with a well-known nuclear-grade steel, WWER 15Kh2MFAA. The most suitable composite material according to the defined criteria is composed of a phenolic resin matrix reinforced with long boron fibers and exhibit great properties to be used in a nuclear reactor environment: good radiation resistance and mechanical properties with a very low density at low cost. It can be concluded that, in some cases, composite material pipes can be a better option than steel ones.

Original article

Institution of MECHANICAL ENGINEERS
Journal of MATERIALS: DESIGN AND APPLICATIONS

Proc IMechE Part L:
J Materials: Design and Applications
2019, Vol. 233(10) 2072–2083

DOI: 10.1177/1464420718817334
journals.sagepub.com/home/pil
SAGE

Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications

David Merayo[1], Álvaro Rodríguez-Prieto[1,2] and Ana María Camacho[1]

Abstract

This study describes a methodological proposal to select composite materials which are suitable to be employed to manufacture pipes that can properly withstand environments subjected to gamma and neutronic radiation. The methodology is used to select, among many others, the optimal composite material whose properties are used afterwards to simulate several pipe sections by finite element analysis, comparing the results with a well-known nuclear-grade steel, WWER 15Kh2MFAA. The most suitable composite material according to the defined criteria is composed of a phenolic resin matrix reinforced with long boron fibres and exhibit great properties to be used in a nuclear reactor environment: good radiation resistance and mechanical properties with a very low density at low cost. It can be concluded that, in some cases, composite material pipes can be a better option than steel ones. Extending the method to be employed in other industries or with other components could be seen as future works.

Keywords

Material selection, nuclear industry, steel, polymeric matrix composite, stringency level, finite element analysis

Date received: 13 August 2018; accepted: 14 November 2018

Introduction

In the nuclear industry, as in many others, the steel has been widely employed because, during centuries XIX and XX, it has become the most relevant construction material (with the exception of the concrete). Although there is a huge amount of relevant metallic materials of great structural importance, its use could be seen as residual compared to steel and iron.

People working in the nuclear industry know perfectly that steel plays a very critical role in the operation safety of an electric power plant.[1] There are many components made of steel in the zones subjected to radiation, such as pipes or high-pressure vessels.

It is worth noting that steels employed to build components for the primary circuit of a nuclear power plant should be able to withstand, in the case of pressurized water reactors (PWR), temperatures of 300 °C and an internal pressure of around 18 MPa, in addition to corrosion and high doses of radiation.[2]

According to Murty and Charit,[3] even nickel-based superalloys could show problems when they are subjected to radiation conditions because helium bubbles can appear and the internal material structure could weaken. Nonetheless, new materials are currently being developed to solve all those issues.[4]

Most of the scientists agree that the clean energy demand will keep growing for the next decades. As society requests more environmentally friendly energy generation technologies, new materials will play a fundamental role in the energy industry: low-density metal alloys, high-performing ceramics, composite materials, new polymers, etc.[5]

Nowadays, a large amount of research on the use of composite materials in the nuclear industry is being done because new materials able to cope with more extreme conditions are requested: greater temperatures, greater pressures and higher radiation levels. Traditionally, metals have been excellent materials to build structural components; however, some physical limits are being reached and, so, it is compulsory to find new solutions. Even though, metallurgy investigation will keep being very important.

According to Ashby and Smidman,[4] before 2050, China and India will build about 700 new nuclear

[1]Department of Manufacturing Engineering, UNED, Madrid, Spain
[2]Applied Materials Division, Argonne National Laboratory, Lemont, IL, USA

Corresponding author:
David Merayo, Department of Manufacturing Engineering, UNED, Juan del Rosal 12, Madrid 28040, Spain.
Email: dmerayo1@estudiante.uned.es

installations to deal with the currently increasing energy demand. So, more research and new solutions are required to take advantage of these power plants in a more efficient and safe way.

Most of the current research related to composite materials in the nuclear industry is due to their expected future relevance because these materials could be potentially used in all areas of the power plant, including those considered critical.[6,7] In addition to their use in conventional nuclear facilities, in the next decades, the International Thermonuclear Experimental Reactor (ITER), which takes advantage of very advanced composite materials, will be inaugurated.

The current development of advanced composite materials to be employed in the next generation of reactors has been boosted by the need of resisting, for longer periods, extreme environments subjected to high radiation fluxes, high temperatures and high stresses that, together, could exceed the traditional materials capabilities. Materials able to face reactor conditions require low activation levels, high structural integrity, high dimensional stability and high thermal resistance.[8] Even if these materials look very promising, according to Memory et al.,[9] there is still a lot of research to be done to understand their behaviour under radiation conditions.

However, future designs will employ a great variety of materials, such as composites and metals. Designers, scientists and engineers will be able to use the best option of material to take advantage of the better available properties, creating thus more efficient power plants.

Structural materials should be able to face extreme environments. A suitable tool for selecting materials in high demanding applications is the stringency level methodology[1] which takes into account a large number of materials requirements to approach the decision problem in a multi-criteria way. This methodology could be spread or modified to consider other materials besides steel.

New materials and new design approaches, along with new manufacturing techniques, are necessary to minimize all those problems and issues and to achieve the optimal performance of systems at more extreme conditions.

Soneda[10] indicates that, currently, designers concentrate on using the more convenient available material to build each component to obtain the best performance and to take advantage of its properties. Metallic and composite materials can cohabit to increase the overall efficiency of the system. There is still room for well-known materials like steel, but new choices will join the designers' panel of options and will make more complicated the decision problem.

In this paper, polymeric matrix composites are proposed to be used in manufacturing of pipes subjected to gamma and neutronic radiation, as an alternative to consolidated ferritic steels such as DIN 20MnMoNi55, ASME SA 533 Grade B Cl.1, and WWER 15Kh2MFAA, typically used in harsh environment areas in a nuclear power plant.[11]

Methodology

The composite materials have been designed considering that these materials are intended to build a pipe section used in the containment of a nuclear power plant. Thus, mechanical stresses, temperature, radiation and corrosion resistance are critical aspects to consider. Composite materials tend to be fragile, chemically stable and, in general, bear very well with radiation but not with temperature.[12,13]

The methodology followed in this work is presented in Figure 1. Several reinforcements and matrices have been considered and their properties have been compared to create the most performant fibre-reinforced materials for this application.

After defining the candidate materials, a decision methodology based on the algorithm described by Rodríguez-Prieto et al.[1] is used to select the optimal material.

Finally, the composite material behaviour has been analysed by the finite element method and simulation results are compared with those of the well-known nuclear-grade steel WWER 15Kh2MFAA to show its strengths and weaknesses.

Matrix pre-selection

Composite material matrices subjected to radiation could decompose or deteriorate because most of the

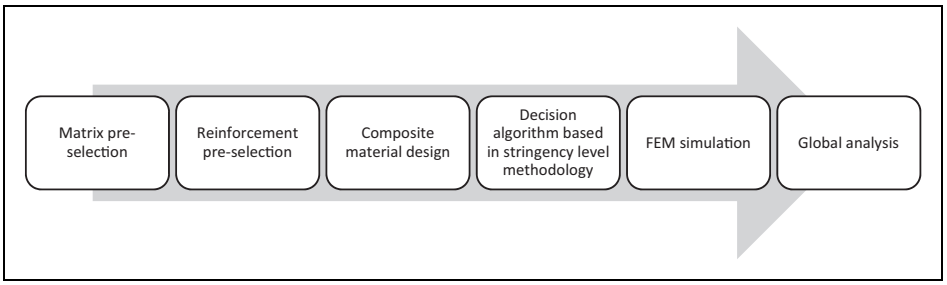


Figure 1. Methodology overview.

regular matrices are polymeric and, so, they can be strongly impacted by neutronic irradiation. In the last decades, a lot of research has been conducted about the effects of radiation on polymers degradation.[14] Carfagno[15] reviewed the ageing theory of polymers subjected to radiation and addressed the neutronic irradiation problem in a very detailed way with great amount of information related to typical composite matrices, that has been employed in this work.

Epoxy, phenolic and polyester resins have been chosen as the main candidates because they can withstand the typical radiation doses found in a nuclear reactor. Table 1 contains some relevant information related to these resins.[15,16]

Reinforcement pre-selection

The selection of reinforcement fibres requires taking into consideration the same principles that have already been advised to select the matrices. Moreover, physical and chemical compatibility between both phases should be ensured.

This study contemplates seven options of fibres to design the composite materials (among the most common ones in the field of composite materials), whose properties can be seen in Table 2.

Anyway, all these reinforcements show a much better stability against ionic radiation than the pre-selected matrices. Table 3 contains some relevant information related to these reinforcements.

Table 1. Properties of the analysed resins.[15,16]

Material	Epoxy	Phenolic	Polyester
Young modulus, E [MPa]	3500	3000	2500
Poisson's ratio, ν	0.33	0.34	0.37
Shear modulus, G [MPa]	1316	1125	1314
Density [kg/m^3]	1540	1320	1380
Gamma dose [Gigarads]	6.00	9.00	1.00
Max. working temp. [°C]	300	340	320
UTS [MPa]	73	100	78
Yield point (0.2%) [MPa]	23	27	18
Max. deformation [%]	7.10	6.80	6.30

Composite material design

Several composite materials have been designed by combining the candidate reinforcements and matrices as shown in Table 4.

The resulting composite materials properties have been estimated by using the mixing rule.[12,13,20] Mechanical properties for composite materials could largely vary for different manufacturing processes and technologies. The composite material properties have been calculated for uniaxial non-knit long fibres. Moreover, the mixing rule has been applied using typical values employed in the aeronautic and automotive industry: 50% volumetric content of fibre, 49.5% volumetric content of resin and 0.5% volumetric empty space.[12]

Tables 5 to 7 contain the 21 designed composite materials properties, which have been considered in this study.

Decision algorithm

The decision methodology based on stringency levels[1] has been adapted to compare the designed composite materials. To do so, a value (in the range of 1–5) is assigned to each relevant property using the three cases formulae described in this work.

The mechanical stringency levels have been calculated using the procedure described by Rodríguez-Prieto et al.;[1] a good behaviour against ionic radiation has been required. Nevertheless, no chemical requirements have been considered.

In addition to consider the mechanical properties of each material, its density and price are also taken into account because there is a greater diversity than among steels.[21] These two additional factors are very relevant since composite materials are usually lighter than steels but, in general, are very expensive. In one hand, lighter materials allow larger pipe sections to be built and handled.[12] In the other hand, little mechanical properties improvements could not always justify large expenses and, so, a balance is needed.[22]

The density and price stringency levels have been calculated using case 1 formulae[1] to ensure that

Table 2. Properties of the pre-selected mechanical reinforcements.[17–19]

Fibre material	Carbon	Glass	Aramid	Boron	SiC	Alumina	Aluminium
Young modulus, E [GPa]	385	71	154	420	406	385	69
Poisson's ratio, ν	0.2	0.22	0.35	0.2	0.2	0.19	0.35
Shear modulus, G [GPa]	7.7	30	2.9	170	169	154	26
Density [kg/m^3]	1940	2450	1470	2450	3200	3900	2700
Max. working temp. [°C]	1500	800	400	1600	2100	1900	400
UTS [MPa]	1632	1100	1200	1853	706	1029	420
Yield point (0.2%) [MPa]	1387	920	970	1571	657	875	354
Max. deformation [%]	0.96	2.05	4.10	3.36	1.37	2.04	6.84

Table 3. Features and applications in the nuclear industry of the reinforcement fibres.

Type of fibre	Features	Nuclear industry applications
Carbon fibre	Expensive, resistant, excellent behaviour against radiation	Structural primary structures, light structures
Glass fibre	Dense, bad behaviour against radiation, very cheap	Structural secondary structures
Aramid fibre	Excellent impact resistance, polymeric	Structural primary structures, structures subjected to impacts
Boron fibre	Low-density, reduce neutron speed, high resistance, cheap	Structural primary structures, light structures
Silicon carbide fibre	Very expensive, excellent resistance against temperature and radiation	Structural primary structures, high-temperature applications
Alumina fibre	Excellent resistance/cost ratio	Structural secondary structures
Aluminium wire	Metallic, low-density, high deformation before failing	Structural primary structures, structures subject to deformation

Table 4. Designed composite materials.

ID	Composite material
A	Carbon + epoxy
A′	Carbon + phenolic
A″	Carbon + polyester
B	Glass + epoxy
B′	Glass + phenolic
B″	Glass + polyester
C	Aramid + epoxy
C′	Aramid + phenolic
C″	Aramid + polyester
D	Boron + epoxy
D′	Boron + phenolic
D″	Boron + polyester
E	SiC + epoxy
E′	SiC + phenolic
E″	SiC + polyester
F	Alumina + epoxy
F′	Alumina + phenolic
F″	Alumina + polyester
G	Aluminium + epoxy
G′	Aluminium + phenolic
G″	Aluminium + polyester

lighter and cheaper materials would obtain better results.

For the calculation, the relative weightings shown in Table 8 are used.[1]

Finally, the global stringency level has been calculated (equations (1) and (2)) as a weighted average:

$$SL_g = R_i^m \cdot SL_m + R_i^\rho \cdot SL_\rho + R_i^\epsilon \cdot SL_\epsilon \quad (1)$$

where

$$SL_m = R_i^{\sigma_m} \cdot SL_{\sigma_m} + R_i^{\sigma_s} \cdot SL_{\sigma_s} + R_i^{\varepsilon_R} \cdot SL_{\varepsilon_R} \quad (2)$$

Finite element modelling

ANSYS Workbench 14.5 has been used to carry out the simulation of the designed components. It is a commercial package of design and simulation tools that allows carrying out studies of physical phenomena in several areas: fluid dynamics, mechanics, electromagnetism and others. It has some calculation algorithms widely used in the industrial field and stands out for its intuitive interface. In addition, it has some specific modules for the analysis and simulation of composite materials. The simulation stages are presented in Figure 2. ANSYS Workbench 14.5 allows to generate parametric geometries, mesh the parts that the model contains, define materials, apply boundary conditions, perform the calculation and post-process the results.

A parametrized geometry (pipe diameter and length) based on a cylindrical surface (membrane stress theory hypothesis has been considered) has been built, which has been meshed by using a curvature approximation with quadratic two-dimensional elements of type Quad. Figure 3 shows a mesh example.

During the pipes meshing, it has been demanded that the quality of the elements was, at least, 0.95. However, the simplicity of the geometry has allowed to achieve better levels of quality. Table 9 shows the meshing data related to each pipe.

Number of elements gives an idea of the complexity of the problem to solve. Note that, for larger diameters, fewer elements are required because the structure will have a less pronounced curvature. Average quality is an indication of the accuracy of the results associated with that element; it is a value in the interval 0 to 1, where values close to 1 indicate the maximum quality. And quality standard deviation represents the dispersion of the values relative to the quality of the elements.

Table 5. Designed composite materials properties 1/3.

Material	A	A′	A″	B	B′	B″	C	C′	C″
Young modulus, E_1 [GPa]	194.23	194.28	193.74	37.73	37.78	37.24	78.73	78.78	78.24
Young modulus, E_2 [GPa]	4.74	4.80	4.02	11.54	11.81	8.67	3.84	3.89	3.26
Poisson's ratio, ν_{12} [mm/mm]	0.26	0.28	0.26	0.27	0.29	0.27	0.34	0.36	0.34
Poisson's ratio, ν_{21} [mm/mm]	0.01	0.01	0.01	0.08	0.09	0.06	0.02	0.02	0.01
Shear modulus, G_{12} [MPa]	2758	2755	2148	3539	3535	2604	1925	1923	1584
Density [kg/m^3]	1730	1614	1653	2062	1944	1983	1497	1379	1418
Max. working temp. [°C]	300	340	320	300	340	320	300	340	320
Cost [€/kg]	100.00	94.00	96.00	24.00	18.00	20.00	43.00	37.00	39.00
UTS [MPa]	1632	1587	1576	1100	1082	1108	1116	1131	1103
Yield point (0.2%) [MPa]	1387	1349	1340	935	920	942	949	961	938
Max. deformation [%]	1.26	1.32	1.25	2.37	2.39	2.37	5.67	5.74	5.64

Table 6. Designed composite materials properties 2/3.

Material	D	D′	D″	E	E′	E″	F	F′	F″
Young modulus, E_1 [GPa]	211.73	211.78	211.24	204.73	204.78	204.24	194.23	194.28	193.74
Young modulus, E_2 [GPa]	13.50	13.87	9.74	13.48	13.85	9.73	13.45	13.82	9.72
Poisson's ratio, ν_{12} [mm/mm]	0.26	0.28	0.26	0.26	0.28	0.26	0.26	0.28	0.26
Poisson's ratio, ν_{21} [mm/mm]	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.02	0.01
Shear modulus, G_{12} [MPa]	3868	3862	2779	3867	3862	2778	3860	3854	2774
Density (kg/m^3)	1987	1869	1908	2362	2244	2283	2712	2594	2633
Max. working temp. [°C]	300	340	320	300	340	320	300	340	320
Price [€/kg]	35.00	29.00	31.00	410.00	404.00	406.00	14.00	8.00	10.00
UTS [MPa]	1853	1853	1848	706	773	735	1029	1064	1036
Yield point (0.2%) [MPa]	1575	1575	1571	600	657	625	875	904	881
Max. deformation [%]	3.50	3.50	3.50	1.38	1.51	1.44	2.12	2.19	2.14

Table 7. Designed composite materials properties 3/3.

Material	G	G′	G″
Young modulus, E_1 [GPa]	36.23	36.28	35.74
Young modulus, E_2 [GPa]	11.45	11.72	8.62
Poisson coeff., ν_{12} [mm/mm]	0.31	0.33	0.31
Poisson coeff., ν_{21} [mm/mm]	0.10	0.11	0.08
Shear modulus, G_{12} [MPa]	3493	3489	2579
Density (kg/m^3)	2112	1994	2033
Max. working temp. [°C]	300	340	320
Price [€/kg]	13.00	7.00	9.00
UTS [MPa]	417	420	417
Yield point (0.2%) [MPa]	354	357	354
Max. deformation [%]	7.00	6.97	6.84

The thickness adjustment, for metals, is directly found by mass-optimization methods because there is no limitation about which pipe-wall thicknesses can be manufactured; however, for composite materials, there is a restriction imposed by the fabrics (in this case, 0.6 mm thick) and, therefore, it is compulsory that the pipe thickness is the result of conveniently stacking these layers.

Table 8. Selected relative weightings.

Coefficient	Weight value
Mechanical properties (R_i^m)	0.6
Density (R_i^ρ)	0.2
Cost ($R_i^€$)	0.2

Figure 4 shows the simulated deformation of the composite material and steel pipes of the well-known nuclear-grade steel, WWER 15Kh2MFAA, for all the considered diameters. Note that this deformation is always lower than 2 mm.

Reserve factor criteria

Several cylindrical pipe sections with different sizes have been simulated by using the finite element method (FEM) to determine the optimum

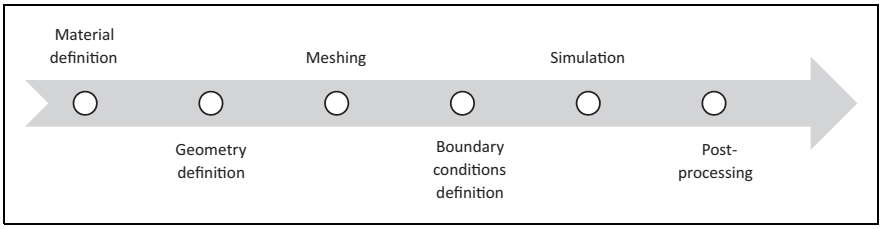


Figure 2. Figure element method simulation tasks.

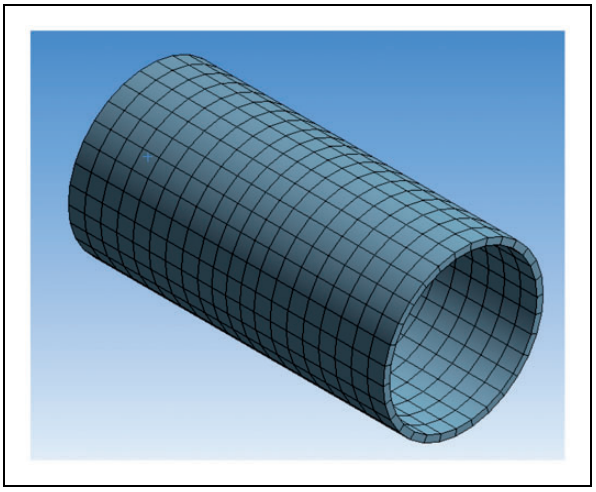


Figure 3. Example of pipe mesh ($D = 508$ mm).

Table 9. Dimensions and mesh quality for each pipe diameter.

Diameter [mm]	76.2	152.4	304.8	508.0	914.4
Number of elements	1760	1708	990	600	330
Average quality	0.997	0.996	0.994	0.996	0.997
Quality standard deviation	0.0086	0.0111	0.0156	0.0057	0.0029

thickness-diameter ratio. The components have been designed to be able to overcome a hydrostatic pressure of 18 MPa with a minimum reserve factor (RF) of 2.

The composite materials RFs have been calculated as the inverse of the tridimensional Tsai–Wu failure criteria,[23] whereas the steel factors have been computed with the Von Mises criteria referred to the maximum allowable stress before plastic deformation takes place.

The Tsai–Wu failure criterion is a theoretical formulation generally applied to anisotropic materials to predict material failure. This failure occurs when the failure Tsai–Wu index reaches the value of 1. The formulation of that index is given by the expression (equation (3))

$$F_1\sigma_1 + F_2\sigma_2 + F_3\sigma_3 + F_4\sigma_4 + F_5\sigma_5 + F_6\sigma_6 + F_{11}\sigma_1^2 + F_{22}\sigma_2^2 + F_{33}\sigma_3^2 + F_{44}\sigma_4^2 + F_{55}\sigma_5^2 + F_{66}\sigma_6^2 + 2F_{12}\sigma_1\sigma_2 + 2F_{13}\sigma_1\sigma_3 + 2F_{23}\sigma_2\sigma_3 \leqslant 1 \quad (3)$$

where

$$F_1 = \frac{1}{\sigma_{1t}} - \frac{1}{\sigma_{1c}} \quad (4)$$

$$F_2 = \frac{1}{\sigma_{2t}} - \frac{1}{\sigma_{2c}} \quad (5)$$

$$F_3 = \frac{1}{\sigma_{3t}} - \frac{1}{\sigma_{3c}} \quad (6)$$

$$F_4 = F_5 = F_6 = 0 \quad (7)$$

with

$$F_{11} = \frac{1}{\sigma_{1c}\sigma_{1t}} \quad (8)$$

$$F_{22} = \frac{1}{\sigma_{2c}\sigma_{2t}} \quad (9)$$

$$F_{33} = \frac{1}{\sigma_{3c}\sigma_{3t}} \quad (10)$$

$$F_{44} = \frac{1}{\tau_{23}^2} \quad (11)$$

$$F_{55} = \frac{1}{\tau_{31}^2} \quad (12)$$

$$F_{66} = \frac{1}{\tau_{12}^2} \quad (13)$$

where the uniaxial and compression stresses are σ_{1t}, σ_{1c}, σ_{2t}, σ_{1c}, σ_{3t}, σ_{1c}; and the shear stresses are τ_{23}, τ_{12}, τ_{31}. The coefficients F_{12}, F_{13} and F_{23} can be determined by equibiaxial stress testing and, therefore (knowing that $\sigma_1 = \sigma_2 = \sigma_{b12}$, $\sigma_1 = \sigma_3 = \sigma_{b13}$, $\sigma_2 = \sigma_3 = \sigma_{b23}$), these coefficients are given by

$$F_{12} = \frac{1}{2\sigma_{b12}^2}\left[1 - \sigma_{b12}(F_1 + F_2) - \sigma_{b12}^2(F_{11} + F_{22})\right] \quad (14)$$

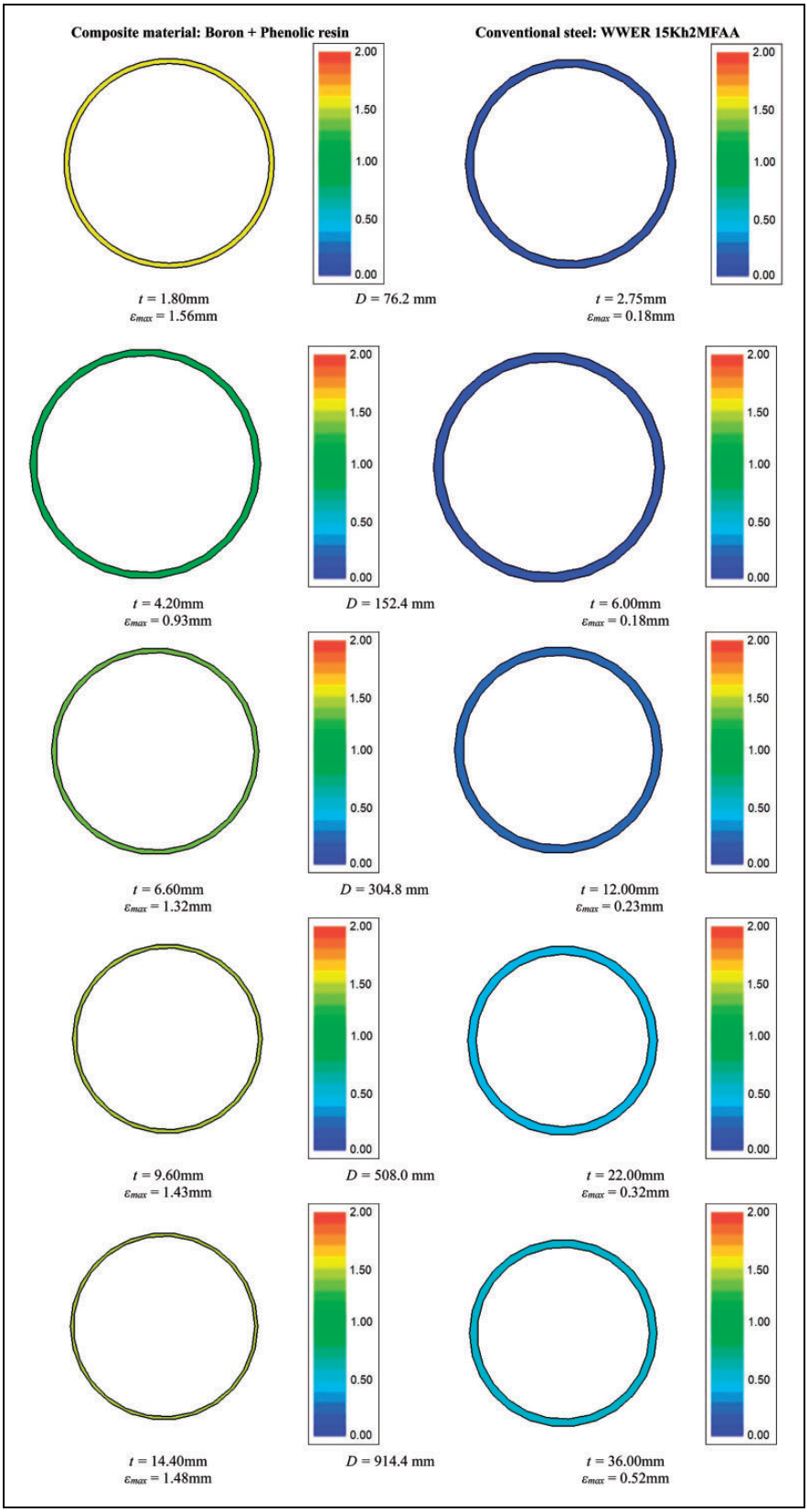


Figure 4. Values of deformation reached under 18 MPa hydrostatic pressure.

$$F_{12} = \frac{1}{2\sigma_{b13}^2}\left[1 - \sigma_{b13}(F_1 + F_3) - \sigma_{b13}^2(F_{11} + F_{33})\right] \quad (15)$$

$$F_{12} = \frac{1}{2\sigma_{b23}^2}\left[1 - \sigma_{b23}(F_2 + F_3) - \sigma_{b23}^2(F_{22} + F_{33})\right] \quad (16)$$

Table 10. Mechanical stringency levels.

	$SL_{\sigma m}$	$SI_{\sigma s}$	$SL_{\varepsilon R}$	**SL_m**
R_i	0.40	0.20	0.40	
A	4.40	4.40	1.00	**3.04**
A′	4.28	4.28	1.00	**2.97**
A″	4.25	4.25	1.00	**2.95**
B	2.97	2.97	1.69	**2.46**
B′	2.92	2.92	1.71	**2.43**
B″	2.99	2.99	1.69	**2.47**
C	3.01	3.01	4.05	**3.43**
C′	3.05	3.05	4.10	**3.47**
C″	2.98	2.98	4.03	**3.40**
D	5.00	5.00	2.50	**4.00**
D′	5.00	5.00	2.50	**4.00**
D″	4.99	4.99	2.50	**3.99**
E	1.91	1.91	1.00	**1.54**
E′	2.09	2.09	1.08	**1.68**
E″	1.98	1.98	1.03	**1.60**
F	2.78	2.78	1.51	**2.27**
F′	2.87	2.87	1.56	**2.35**
F″	2.80	2.80	1.53	**2.29**
G	1.13	1.13	5.00	**2.68**
G′	1.13	1.13	4.98	**2.67**
G″	1.13	1.13	4.89	**2.63**

On the other hand, the equivalent Von Mises stress, σ_v, predicts yielding of isotropic materials subjected to multiaxial loading. Plastic deformation occurs when the equivalent tensile stress in a given point of the material is greater than the equivalent Von Mises yield stress, which is estimated as (equation (17)

$$\sigma_v = \sqrt{\frac{(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_1 - \sigma_3)^2}{2}} \quad (17)$$

Results

Determination of the best composite material based in SL methodology

As indicated, Tables 5 to 7 contain the 21 designed composite materials properties. From these data, following the adapted methodology, the stringency values for each material can be estimated. Table 10 contains the stringency levels related to mechanical properties, such as the ultimate tensile stress, the yield stress (0.2%), maximum deformation and the weighted mechanical stringency level. The information in the latter column is shown in Figure 5 in a more intuitive way.

Due to the large price variability of the composite materials, the standard estimative assumptions employed in the aeronautical industry have been considered: a fixed manufacturing cost of 8 €/kg and a 10% increase are added to the manufacturer prices for raw materials.[16–19] Table 11 contains the mechanical, density, cost and global weighted stringency levels for each composite material considered in this study. Figure 6 shows these global values in a graphic way.

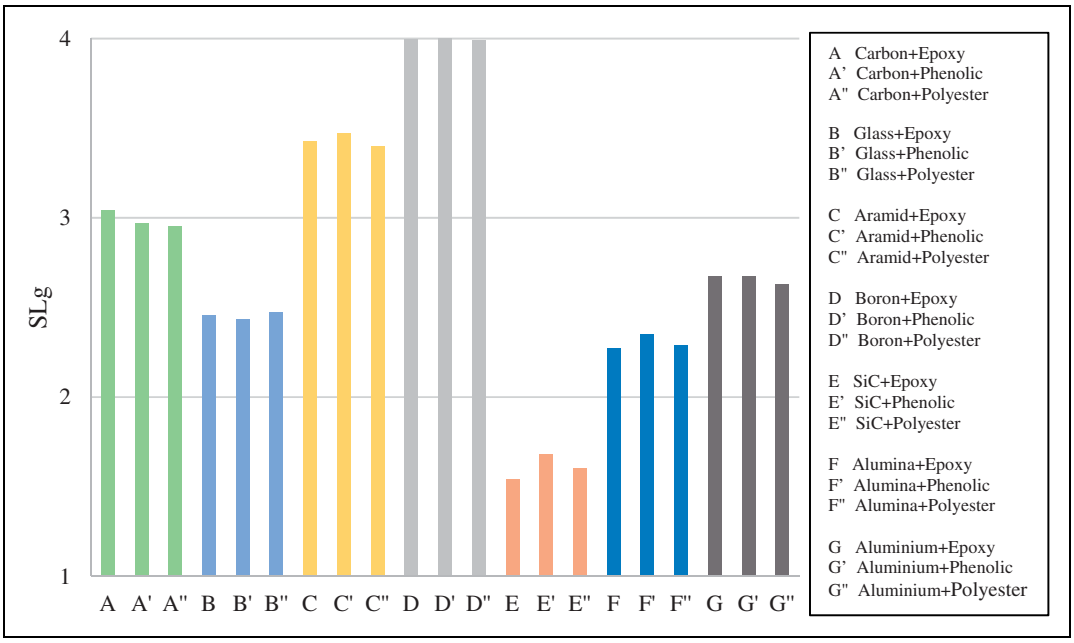


Figure 5. Mechanical stringency level for each composite material.

Given these values, the following comments can be made:

- The results obtained using the methodology allow a positive evaluation of the aramid and boron fibres.

Table 11. Global stringency levels.

		SL_m	SL_ρ	SL_ϵ	**SL_g**
	Ri	0.60	0.20	0.20	
Carbon + epoxy	A	3.04	3.98	1.00	**2.82**
Carbon + phenolic	A′	2.97	4.27	1.00	**2.84**
Carbon + polyester	A″	2.95	4.17	1.00	**2.80**
Glass + epoxy	B	2.46	3.34	1.46	**2.43**
Glass + phenolic	B′	2.43	3.55	1.94	**2.56**
Glass + polyester	B″	2.47	3.48	1.75	**2.53**
Aramid + epoxy	C	3.43	4.6	1.00	**3.18**
Aramid + phenolic	C′	3.47	5.00	1.00	**3.28**
Aramid + polyester	C″	3.4	4.86	1.00	**3.21**
Boron + epoxy	D	4.00	3.47	1.00	**3.29**
Boron + phenolic	D′	4.00	3.69	1.21	**3.38**
Boron + polyester	D″	3.99	3.61	1.13	**3.34**
SiC + epoxy	E	1.54	2.92	1.00	**1.71**
SiC + phenolic	E′	1.68	3.07	1.00	**1.82**
SiC + polyester	E″	1.6	3.02	1.00	**1.76**
Alumina + epoxy	F	2.27	2.54	2.50	**2.37**
Alumina + phenolic	F′	2.35	2.66	4.38	**2.82**
Alumina + polyester	F″	2.29	2.62	3.50	**2.60**
Aluminium + epoxy	G	2.68	3.26	2.69	**2.80**
Aluminium + phenolic	G′	2.67	3.46	5.00	**3.29**
Aluminium + polyester	G″	2.63	3.39	3.89	**3.03**

- The best material according to the methodology is a phenolic matrix reinforced with boron fibre, *D′* ($SL_g = 3.38$).
- For each reinforcing fibre, the best matrix according to the methodology is the phenolic resin.

Therefore, the most suitable composite material according to the defined criteria is composed of a phenolic resin matrix reinforced with long boron fibres. This selected composite exhibits another additional interesting feature, since boron has the ability to reduce the speed of neutrons[24–26] and, so, this material could have a secondary function as protection shield against neutronic radiation.

Element design

The pipe diameter selection is not exhaustive but allows us envisage the tendency. The considered sizes have been:

- 76.2 mm (3″)
- 152.4 mm (6″)
- 304.8 mm (12″)
- 508.0 mm (20″)
- 914.4 mm (36″)

Table 12 contains the results obtained after performing the simulations with steel and composite material for each pipe diameter. This table shows the linear density (mass per length), the minimum thickness to be able to withstand the constraints (note that the composite material thickness is restricted by the manufacturing process: automated filament winding), RF, maximum deformation and

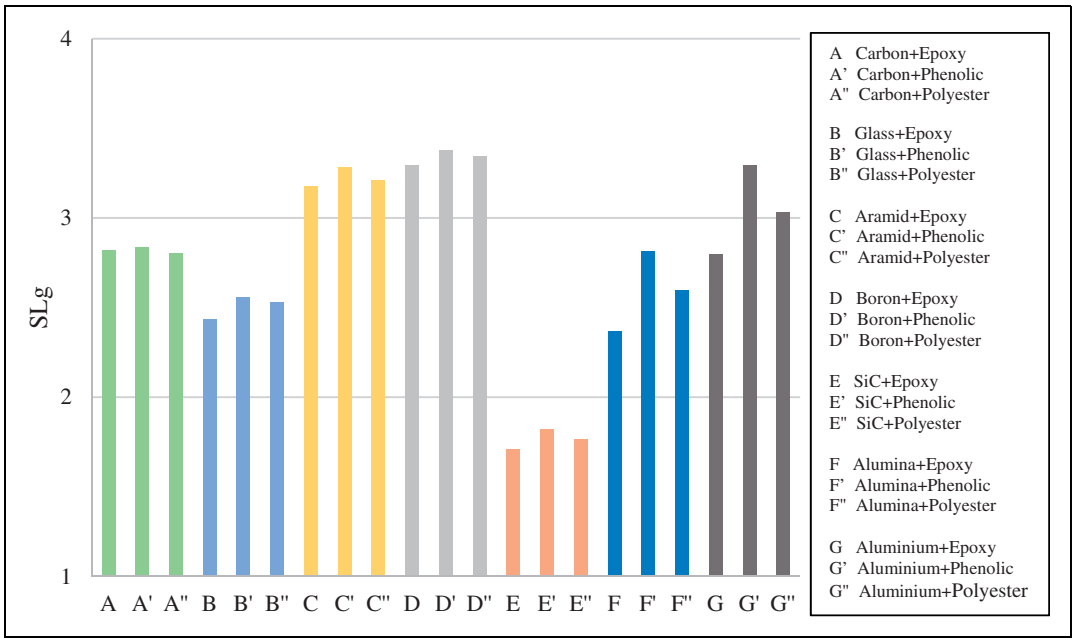


Figure 6. Global stringency level for each composite material.

cost (only the bulk material and the manufacturing expenses are considered). WWER 15Kh2MFAA steel has been selected to compare with the composite material according to Rodríguez-Prieto et al.,[11,27,28] who selected it as one of the most promising for nuclear applications. Figures 7–9 show the comparison information obtained by simulation in a graphic way.

Figure 7 shows the linear density curves for the considered pipe diameters for both materials. Not only the composite material pipes are lighter but also their growing curve stays always below the steel one, which allows the designers to think about larger pipe sections and slender supporting structures.

It must be noted that, as shown in Figure 8, pipe thickness grows almost linearly with the diameter for both materials; however, the curve slopes are very different and the steel conductions become thicker than the composite material ones. That can also explain the differences found when their linear densities have been compared.

Figure 9 shows the expected manufacturing cost for the considered pipes. It can be seen that both options are almost equal and so, both options can be considered as potential candidates to be explored during the design phases. If we look the figure closer, it can be seen that there is a diameter around 508 mm for which the steel pipes are no more economically better because the composite material becomes cheaper.

Table 12. Finite element simulations results.

Diameter [mm]	76.2	152.4	304.8	508.0	914.4
Boron + phenolic resin					
RF	2.12	2.92	2.36	2.20	2.23
Deformation [mm]	1.56	0.93	1.32	1.43	1.48
Density [kg/m]	1.60	7.49	23.55	57.10	154.17
Minimum thickness [mm]	1.80	4.20	6.60	9.60	14.40
Cost [€/m]	46.40	217.21	682.95	1655.90	4470.93
WWER 15Kh2MFAA steel					
RF	2.05	2.19	2.12	2.09	2.10
Deformation [mm]	0.18	0.18	0.23	0.32	0.52
Density [kg/m]	5.17	22.55	90.20	282.00	811.82
Minimum thickness [mm]	2.75	6.00	12.00	22.00	36.00
Cost [€/m]	31.00	135.30	541.21	1692.00	4870.92

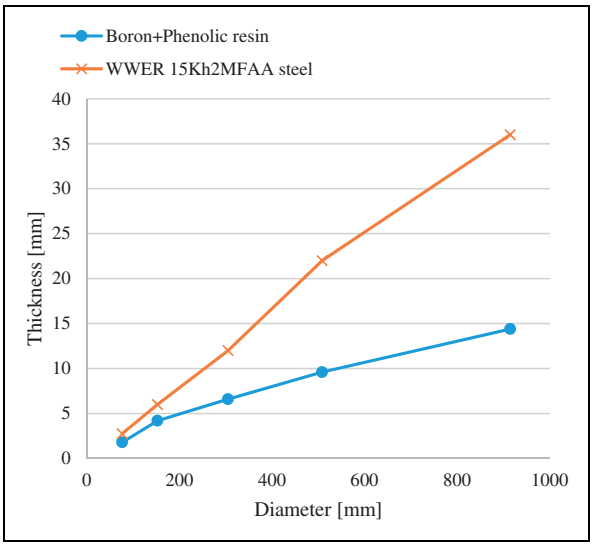


Figure 8. Minimum thickness versus diameter for WWER 15Kh2MFAA steel and boron + phenolic resin.

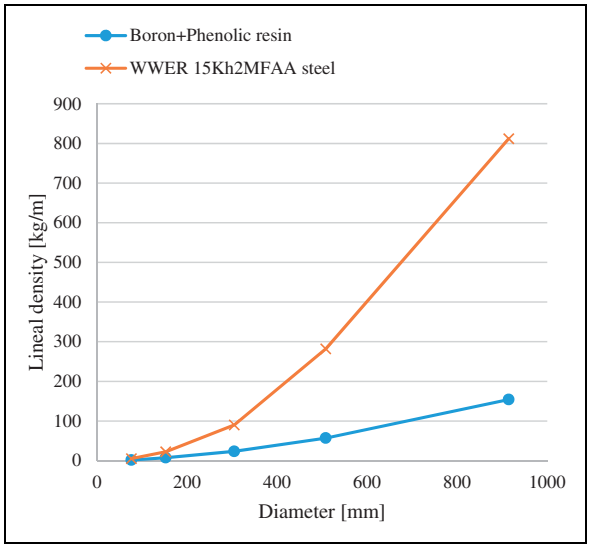


Figure 7. Linear density versus diameter for WWER 15Kh2MFAA steel and boron + phenolic resin.

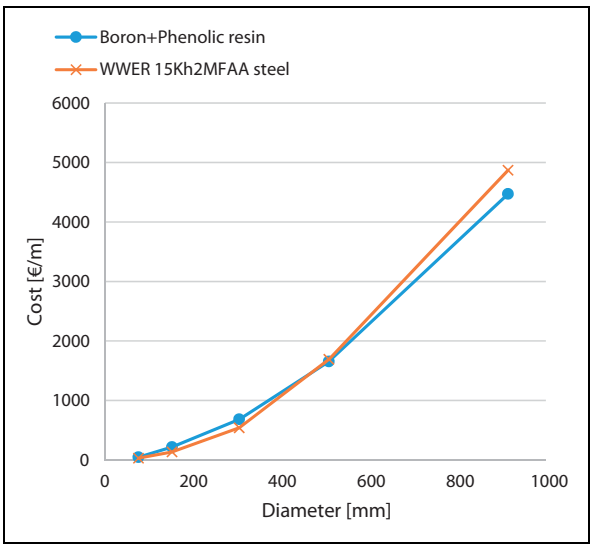


Figure 9. Cost versus diameter for WWER 15Kh2MFAA steel and boron + phenolic resin.

Conclusions

Here are some of the conclusions that can be drawn from this study:

- It has been explored the possibility of using composite materials that are able to meet, at least, the same requirements that have been imposed on steels and, thus, can be considered as potential valid substitutes for metallic components in components that can properly withstand environments subjected to gamma and neutronic radiation.
- In the case of the selection of the elements that constitute a composite material, it is necessary to consider, not only the particular properties of each component, but also the interactions that occur between them.
- After the study carried out on the preselected composites, a good behaviour and characteristics have been observed that allow them to be considered as excellent candidates for the manufacture of structural elements subjected to radiation conditions.
- Composite materials reinforced with boron fibre have the ability to meet the requirements of structural steels, adding, also, the ability of boron to reduce the speed of the neutrons that make up the neutron radiation.
- One of the main advantages arising from the use of composite structures lies in their lower density, which makes the final structures lighter and, therefore, it is possible to manufacture larger elements or handle larger structures.
- Manufacturing pipes made of phenolic resin reinforced with long boron fibres could be economically better than manufacturing them with conventional nuclear-grade steel, WWER 15Kh2MFAA.
- In the selection of composite materials, a large part of the cost is related to the price of the selected reinforcing fibre. Anyway, the composite material pipes have a similar price to the steel ones.
- There is a wide range of matrices for composite material capable of withstanding the radiation and heat conditions that have been imposed as working conditions. It must be taken into account that one of the main missions of the resin (although not the only one) is to serve as protective wrapping of the reinforcing fibres.
- It can be considered that the stringency level methodology[1] adapted to the composite materials is a very useful tool that has allowed, in a very simple way, to compare the properties of the different selected materials.
- The combination of the stringency level methodology with numerical simulation techniques such as the finite element method is a powerful tool for designers in demanding engineering applications.

Future works

- The possibility of extending or developing specific methodologies for other industries, such as automotive or aeronautics is proposed.
- Based on the conclusions drawn in this study, the redesign of other components is proposed to be adapted to be manufactured with composite material and to use the full potential that this methodology can provide.

Acknowledgements

This work has been developed within the framework of the doctorate program in Industrial Technologies of the UNED.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was financially supported by the funds provided through the Annual Grant Call of the E.T.S. Industriales of UNED of reference 2018-ICF04.

ORCID iD

David Merayo http://orcid.org/0000-0002-2018-3241
Álvaro Rodríguez-Prieto http://orcid.org/0000-0002-0712-7472

References

1. Rodríguez-Prieto A, Camacho AM and Sebastián MA. Materials selection criteria for nuclear power applications: a decision algorithm. *JOM* 2016; 2: 496–506.
2. Rodríguez-Prieto A. *Análisis de requisitos tecnológicos de materiales especificados en normativas reguladas y su repercusión sobre la fabricación de recipientes especiales para la industria nuclear*. PhD Dissertation, Departamento de Ingeniería de Construcción y Fabricación, Universidad Nacional de Educación a Distancia, Spain, 2014.
3. Murty K and Charit I. Structural materials for Gen-IV nuclear reactors: challenges and opportunities. *J Nucl Mater* 2008; 383: 189–195.
4. Ashby M and Smidman M. *Materials for nuclear power systems*. Cambridge: Granta Design, 2010.
5. Exxon Mobil. Worldwide energy demand to rise 35% by 2030. *Pipeline Gas J* 2010; 1: 237.
6. Katoh Y, Ozawa K, Hinoki T, et al. Mechanical properties of advanced SiC fiber composites irradiated at very high temperatures. *J Nucl Mater* 2011; 1: 416–420.
7. Rodríguez-Prieto A, Camacho AM and Sebastián MA. Evaluation method for pressure vessel manufacturing codes: the influence of ASME unit conversion. *Int J Mater Prod Technol* 2017; 4: 259–274.
8. Simos N. *Composite materials under extreme radiation and temperature environments of the next generation nuclear reactors*. Brookhaven: Department of Energy, Office of Science, 2011.

9. Memory JD, Fornes RE and Gilbert RD. Radiation effects on graphite fiber reinforced composites. *J Reinf Plast Compos* 1988; 1: 33–65.
10. Soneda N. *Irradiation embrittlement of reactor pressure vessels (RPVs) in nuclear power plants*. Cambridge, UK: Elsevier, 2014.
11. Rodríguez-Prieto A, Camacho AM and Sebastián MA. Selection of candidate materials for reactor pressure vessels using irradiation embrittlement prediction models and a stringency level methodology. *Proc IMechE, Part L: Materials: Design and Applications*. Epub ahead of print 30 August 2017. DOI: 10.1177/1464420717727769.
12. Baker A, Stuart D and Kelly D. *Composite materials for aircraft structures*. Bozeman: AIAA Education Series, 2004.
13. Chung DD. *Carbon fiber composites*. New York: Butterworth-Heinemann, 1994.
14. Clough RL. *High-energy radiation and polymers: A review of commercial processes and emerging applications*. Albuquerque: Organic Materials Department, Sandia National Laboratorie, 2001.
15. Carfagno S. *A review of equipment aging theory and technology*. Philadelphia: EPRI Report NP-1558, 1980.
16. Huntsman. *Huntsman datasheet*. The Woodlands: Huntsman, 2016.
17. Specialty Materials. *Data sheets*. Tulsa: Specialty Materials, 2016.
18. Deltatech. *Data sheets catalog*. Altopascio, Italy: Deltatech, 2016.
19. Zircar Ceramics. *Data sheets*. Florida: Zircar Ceramics, 2016.
20. Boisse P. *Composite reinforcements for optimum performance*. Cambridge, UK: Woodhead Publishing, 2011.
21. Rodríguez-Prieto A, Camacho AM and Sebastián MA. Quantitative analysis of prediction models of hot cracking in stainless steels using standardized requirements. *Sadh Acad Proc Eng Sci* 2017; 12: 2147–2155.
22. Genovese A, Russo M and Strano S. Mechanical characterization and modeling of an innovative composite material for railway applications. *Proc IMechE, Part L: J Materials: Design and Applications* 2017; 231: 122–130.
23. Tsai SW and Wu EM. A general theory of strength for anisotropic materials. *J Compos Mater* 1971; 1: 58–80.
24. Shin JW, Lee JW, Yu S, et al. Polyethylene/boron-containing composites for radiation shielding. *Thermochim Acta* 2014; 585: 5–9.
25. Abdel-Aziz MM, Gwaily SE, Makarious AS, et al. Ethylene-propylene diene rubber/low density polyethylene/boron carbide composites as neutron shields. *Polym Degrad Stab* 1995; 2: 235–240.
26. Harrison C, Weaver S, Bertelsen C, et al. Polyethylene/boron nitride composites for space radiation shielding. *J Appl Polym Sci* 2008; 4: 2529–2538.
27. Rodríguez-Prieto A, Camacho AM, Merayo D, et al. An educational software to reinforce the comprehensive learning of materials selection. *Comput Appl Eng Educ* 2018; 26: 125–140.
28. Rodríguez-Prieto A, Camacho AM and Sebastián MA. Multicriteria materials selection for extreme operating conditions based on a multiobjective analysis of irradiation embrittlement and hot cracking prediction models. *Int J Mech Mater Des* 2018; 14: 617–634.

Appendix

Notation

D	pipe diameter
E	Young modulus
E_1	Young modulus in the main direction
E_2	Young modulus in the secondary direction
FEM	finite element analysis
G	shear modulus
G_{12}	in-plane shear modulus
ITER	International Thermonuclear Experimental Reactor
R^m	mechanical properties average weight
R^ρ	density average weight
$R^\in$	cost average weight
$R_i^{\sigma_m}$	UTS average weight
$R_i^{\sigma_s}$	yield stress average weight
$R_i^{\varepsilon_R}$	ultimate elongation average weight
SL_g	global stringency level
SL_m	mechanical stringency level
SL_ρ	density stringency level
SL_{σ_m}	UTS stringency level
SL_{σ_s}	yield stress stringency level
SL_{ε_R}	ultimate elongation stringency level
t	pipe thickness
UTS	ultimate tensile strength
ϵ	deformation
ν	Poisson's ratio
ν_{12}	in-plane Poisson's ratio (tension in main direction)
ν_{21}	in-plane Poisson's ratio (tension in secondary direction)
σ	tensile/compression stress
τ	shear stress

4.2 *Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0* [1]

Los indicios de calidad de este artículo pueden encontrarse en el Apéndice B.

4.2.1 Datos de la publicación y factor de impacto

Tabla 2. Factor de impacto de Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0

Título	Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0
Autores	David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Revista	Procedia Manufacturing
ISSN	23519789
Editorial	Elsevier BV
País	Países Bajos
Volumen	41
Páginas	42-49
Fecha	2019
doi	10.1016/j.promfg.2019.07.027
Factor de impacto	0.516 (SJR – 2019; SCImago Journal Rank) (Industrial and Manufacturing Engineering, Q2)

4.2.2 Resumen y copia de la publicación

Este artículo tiene como objetivo examinar las técnicas computacionales más innovadoras para seleccionar el material óptimo para fabricar un producto en la Industria 4.0. Se examina cómo las técnicas de inteligencia artificial (IA) más relevantes pueden ayudar al proceso de fabricación eligiendo el material más conveniente para las aplicaciones previstas de acuerdo con sus propiedades y comportamiento en servicio. El artículo presenta una descripción general, que incluye una lista de ventajas y desventajas, una explicación matemática y su aplicación a la selección de materiales para la fabricación de las siguientes técnicas: arboles de decisión, redes neuronales multicapa, algoritmos K-medias y análisis de componente principal.

This paper is intended to examine the most innovative computing techniques to select the optimal material to manufacture a product in Industry 4.0. It examines how the most relevant artificial intelligence (AI) techniques can assist the manufacturing process by choosing the most convenient material for the envisaged applications according to their properties and in-service behavior. The paper includes an overview which includes a list of advantages and disadvantages, a mathematical explanation and its application to the selection of materials for manufacturing for each of the following techniques: decision trees, multilayer neural networks, K-means algorithms and principal component analysis.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 41 (2019) 42–49

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

8th Manufacturing Engineering Society International Conference

Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0

D. Merayo[a], A. Rodríguez-Prieto[a,b], A. M. Camacho[a]*

[a] *Department of Manufacturing Engineering, UNED, Juan del Rosal 12, Madrid 28040, Spain*
[b] *Applied Materials Division, Argonne National Laboratory, 9700 Cass Ave, Lemont, IL 60439, USA.*

Abstract

This paper is intended to examine the most innovative computing techniques to select the optimal material to manufacture a product in Industry 4.0. It examines how the most relevant artificial intelligence (AI) techniques can assist the manufacturing process by choosing the most convenient material for the envisaged applications according to their properties and in-service behaviour. The paper includes an overview of each pertinent technique, which includes a list of advantages and disadvantages, mathematical explanation and its application to the selection of materials for manufacturing.

Keywords: Artificial intelligence; material selection; industry 4.0; supervised learning; unsupervised learning

1. Introduction

Selecting the optimal material in manufacturing for an industrial application is becoming more and more complex and time-consuming due to the large amount of options which currently exist. Many decision techniques have been developed but most of them are only intended to be used in certain fields or contexts and, in general, there is not a standardized way to proceed [1].

The objective of this field is to explore the way in which some of the intelligent functions of the human brain can be imitated and executed, so that people can concentrate on carrying out tasks that have a higher added value [2].

* Corresponding author: D. Merayo.
E-mail address: dmerayo1@alumno.uned.es

Currently, we are immersed in a thriving scene of research and application of artificial intelligence methods, especially relevant in the application of the multilayer neural network model presented by Hopfield [3].

The application of artificial intelligence (AI) has proven to be a key tool dealing with complex problems [4] involved in demanding applications. However, the possibilities of these techniques have not been studied yet in detail within the world of manufacturing. At present, AI can only be seen as a powerful assistant, since these systems still require the supervision of human agents who validate their decisions [5]. In any case, these are systems capable of greatly reducing the complexity of the tasks entrusted.

Since the birth of computing, a large number of methods and techniques have been developed that can be framed within the field of AI and, of course, some of them will be more useful than others [6]. Throughout this research, it is intended, among other things, to study their performance and capabilities. Nowadays, technology allows us to produce materials with a great variety of properties; moreover, these materials can perform in a very different way depending on the environment and the working conditions [7, 8]. All of that implies that a huge amount of data is available and so, the situation is about to become unmanageable by only using customary calculation methods. We require tools which were able to dive into these data and propose the optimal solution among the options [9].

Our current technological development enables to produce massive amounts of material behaviour data related to the industrial manufacturing process, the conditions of the machines and their components, the environment and the performance. Nevertheless, these data should be processed to obtain reliable and useful information which allows to take better decisions [9]. AI can help designers to select the appropriate material for each task, which can potentially increase the over-all performance and safety, reducing the maintenance costs and allowing the system to operate in more extreme conditions [5].

2. Main objectives

Currently, designers face such a complex task when selecting the material that will be used for each application that, in many cases, a suboptimal one is eventually used. Industry 4.0 is committed to optimize the material selection process and, at this point, AI techniques can be of great help [10].

The aim of this work is to determine which of the main AI techniques currently available can help more efficiently in the task of selecting materials for the manufacture of demanding applications components. This paper deals with some of the most important techniques including them in two categories: supervised learning and unsupervised learning [11]. The present study also aims to determine which category can be more useful for materials selection and if there is a certain consistency within each group, that is, if it can be said that certain types of techniques yield better results or perform better than the others [6, 11].

To make the selection of materials more refined and less tedious, AI techniques could be integrated within the tools that the designers have at their disposal [6, 12]. It should not be forgotten that, currently, there are a large number of very diverse materials with very heterogeneous properties that, in many applications, must be considered [7]. In addition, AI techniques can consider other factors than purely technical or economic ones, such as environmental or performance factors among others [13].

AI has been a great aid allowing to obtain excellent results as an engineering assistance tool [14, 15]. Since new uses for these methods arise day-to-day, AI techniques will have an increasing importance in the future industrial designs because they allow to optimize the available resources associated to each project [16, 17].

3. Methodology

The studied methods are framed into two main categories: Supervised Learning (SL) techniques and Unsupervised Learning (UL) techniques [11]. This pooling allows us to globally compare the main two machine learning working philosophies to determine what category could deliver better results in applications related to the material selection for the manufacturing industry [11].

Establishing a comparison among methods based on supervised and unsupervised learning has a great relevance during the training phase since, in general, much more starting data is required for the second ones to obtain good results [17, 18].

For each technique, a brief development of the principles of operation and of its mathematical model is introduced. This explanation serves as a basis for addressing, from a formal and rigorous perspective, the main strengths and weaknesses, which could easily give an idea of how powerful it is. A good understanding of the operating principles and mathematical foundations of each method allows to mitigate the errors that may appear during the learning and exploitation phases.

4. Results

4.1. Decision trees

Decision trees are a set of methods categorized as supervised learning techniques that are generally used to perform classifications. Its objective is to create a model capable of predicting the value of an objective variable based on a series of decision rules inferred from the training data features. Decision trees learn from the input data to model the boundaries that allow describing the different categories [19, 20]. One of the input variables is selected in each internal node of the tree according to a method that depends on the selected algorithm. Each branch to a child node corresponds to a set of values of an input variable, so that all branches cover all possible values of the input variable [20].

Each leaf (or terminal node of the tree) represents a value (or range of values) of the target variable. The combination of the values of the input variables is represented by the path from the root to the leaf [19]. The tree is usually constructed by separating all the data into subsets according to the value of an input characteristic. This process is repeated in each subset recursively, so it is a recursive partition [20, 21].

The recursion is completed in a node, either when the defined recursion depth is achieved or when the separation no longer improves the prediction [22]. This process is called top-down induction of decision trees (TDIDT) and it is a greedy algorithm, since we are looking at each node of the tree for the optimal distribution, in order to obtain the best possible participation in the whole decision tree. In decision trees, this is the most common strategy of learning from data [19, 21, 22]. Figure 1 shows the boundary learned by training a decision tree with a 2-dimensional labelled dataset.

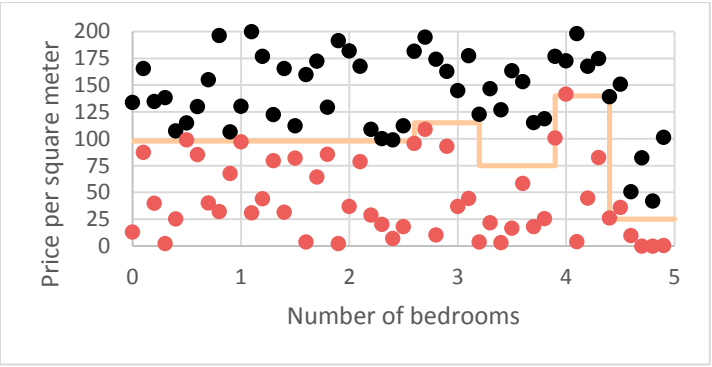


Fig. 1. Decision tree example in a 2-dimensional space.

Some strengths of these methods are:

- The computational requirements are low.
- Decision trees don't require data pre-processing and the learning process is usually fast.

However, they also have some weaknesses:

- Trees are very given to overfitting and tend to have bias if a class dominate.
- Learning the optimal decision tree is known to be an NP-complete problem (nondeterministic polynomial time).

The computational construction complexity of decision trees is $\mathcal{O}(n_{features} \cdot n_{samples} \cdot \log n_{samples})$ and the querying complexity is $\mathcal{O}(\log n_{samples})$ [21].

The simplest case consists on a 2-classes categorizer. Given a training vector set $x_i \in \mathbb{R}^n, i = 1, \dots, N$ and a training label vector $l \in \{-1,1\}^N$, the algorithm builds the decision tree by recursively splitting the space so that, in each new division, the number of elements with the same label that are grouped is maximized. Given that the data at node m is referred as Q. For each possible split $\theta = (j, t_m)$, where j is the feature and t_m is the threshold, the data is parted into the subsets $Q_{right}(\theta)$ and $Q_{left}(\theta)$ as shown in Eq. 1:

$$Q_{right}(\theta) = (x, l)/x_j \geq t_m$$
$$Q_{left}(\theta) = Q \backslash Q_{right}(\theta) \quad (1)$$

Being H the impurity function that calculates the number of wrong classed elements in a subset, the optimal split θ^* is the one that minimizes the function G (Eq. 2):

$$G(Q, \theta) = \frac{n_{right}}{N_m} H\left(Q_{right}(\theta)\right) + \frac{n_{left}}{N_m} H\left(Q_{left}(\theta)\right) \quad (2)$$

The decision tree construction continues by recursively split the subsets $Q_{right}(\theta^*)$ and $Q_{left}(\theta^*)$ until the maximum depth is attained, the minimum subset size is reached $N_m < min_{samples}$ or until the subset contains only one element $N_m = 1$ [21].

These methods can be used to help in the decision of the material that should be used to manufacture an industrial element. For example, if we have the working boundary conditions of an industrial fluids ducting facility and its construction material, we can train a decision tree algorithm so that it can propose the most suitable material that should be used to build a new pipeline.

4.2. Neural network - Multi-layer perceptron

A multi-layer perceptron (MLP) is a supervised learning algorithm able to learn a non-linear function by training on a labelled dataset which can be employed to perform classifications and regressions [23, 24]. Considering the connection topology of the perceptrons, 3 types of layers can be defined: input layer, which includes all perceptrons that receive data from an external source; output layer, which includes all perceptrons that return results (information); and hidden layer, which includes all other perceptrons which doesn't communicate with the network exterior [24] [25].

An MLP can be trained to learn a function (Eq. 3):

$$f(X): \mathbb{R}^m \rightarrow \mathbb{R}^o \quad (3)$$

where $X = \{x_i \ / \ i \in 1 \dots m\}$ is the input vector, m is the size of the input vector and o is the size of the output vector [26].

Input layer perceptrons only receive data from the outside, output layer perceptrons only return the function result and, finally, hidden layer perceptrons are involved in learning and performing all the required calculations [23].

Given a set of input data vectors X and its related set of target vectors $Y = \{y_i \ / \ i \in 1 \dots o\}$, the MLP can be trained to learn the function that approximates the function $f(\cdot)$ that relies them [23]. Each node (except the ones in the input layer) receives data from the previous perceptrons, calculates a weighted linear summation and returns the result of its non-linear activation function as shown in Eq. 4:

$$g\left(\sum_{i=1}^{n} w_i p_i\right): \mathbb{R}^n \rightarrow \mathbb{R} \quad (4)$$

Where $g(\cdot)$ is the activation function, n is the input vector size for this node, $W = \{w_i \ / \ i \in 1 \dots n\}$ is the weights vector and $P = \{p_i \ / \ i \in 1 \dots n\}$ is the input vector for this node. The first hidden layer receives the network input data from the input layer and the output layer receives the results from the last hidden layer. The weights vector values W of all perceptrons are learned by training by using the backpropagation algorithm [26]. Figure 2 shows a 3-layer artificial neural network able to learn $\mathbb{R}^3 \rightarrow \mathbb{R}$ functions.

The advantages of MLP are:

- Capable of learning non-linear models.

- Capable of reaching great complexity.

The weaknesses of MLP are:

- Neural networks are prone to finding local minima and its training is a slow and complex task.
- Requires tuning very relevant parameters before starting the training.

Given that n is the size of the training dataset, m is the number of features, k is the number of hidden layers (containing h nodes each, for simplicity), o is the number of output perceptrons and i is the number of iterations, then, the time complexity of the backpropagation training is $\mathcal{O}(n \cdot m \cdot h^k \cdot o \cdot i)$. So, it is highly recommended to minimize the number of hidden nodes to reduce the training time [26].

This algorithm can be employed to recognize all sort of patterns among a dataset. For example, if we have a broad and concise database of materials, we can train an MLP algorithm to recommend the best material to manufacture an industrial component based on all known features and all the indicated boundary conditions. The solution of the algorithm will condense all the knowledge contained in the starting data.

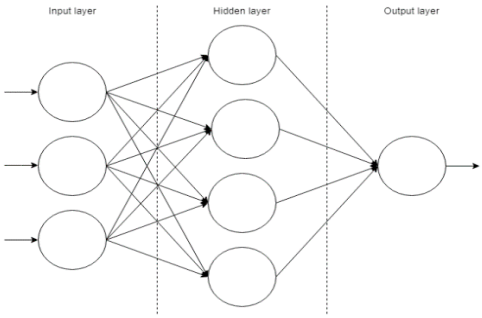


Fig. 2. 3 layers neural network example.

4.3. K-means clustering algorithms

Data clustering is a group of data analysis methods based on unsupervised learning. Its objective is to divide a set of data vectors into different homogeneous packages (clusters) to create data subsets that share common characteristics, which often correspond to criteria of proximity (information similarity) that are defined as common values or classes of distance between objects [27, 28].

The clustering strategy is very important when it comes to obtain good results and depends, to a great extent, on the inertia of the data set and subsets. In this case, inertia is understood as a measure of the distance (in terms of information) that exists between two data vectors or two data groups [28]. In order to obtain good partitions, it is required to minimize intraclass inertia to obtain groups as homogeneous as possible and to maximize the inertia between classes to obtain well-differentiated subsets.

These methods are commonly used in classification (which may eventually allow the data to be condensed), data segmentation and search of outliers. There are many clustering strategies and inertial management methods that give rise to different algorithms, which, in general, reach distinct results. The most commonly used clustering algorithm is the K-means [29]. Figure 3 shows an example of a dataset partitioned by using K-means.

The K-means method requires that the number of clusters to be built was specified and is often referred as the Lloyd's algorithm. It groups the data through a strategy that separates the samples into subsets of the same variance, minimizing the intra-cluster inertia. Given a set of N samples X to be divided into K disjoint clusters C of size n_j, each described by the mean $\mu_j, j \in 1, \dots, K$ (centroid), the algorithm chooses the centroids that minimize the inertia given by Eq. 5 [30]:

$$\sum_{i=0}^{n_j} \min_{\mu_j \in C} \left(\left\| x_i - \mu_j \right\|^2 \right) \tag{5}$$

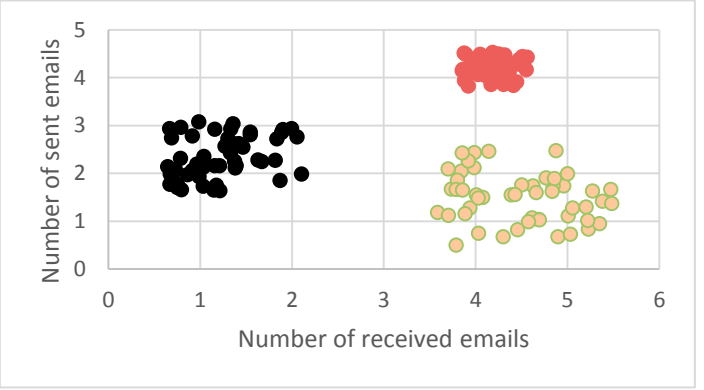


Fig. 3. K-means clustering algorithm example in a 2-dimensional space.

The average complexity is given by $\mathcal{O}(k \cdot N \cdot T)$, where N is the number of samples and T is the number of iterations even though the worst-case complexity is given by $\mathcal{O}\left(N^{\frac{k+2}{f}}\right)$ where N is the number of samples and f is the number of features [30].

The advantages of the clustering algorithms are:

- There are efficient polynomial complexity heuristics that can solve these problems.
- K-means is generally very fast but there are bad cases where convergence is delayed.

The disadvantages of the clustering algorithms are:

- K-means is proud to sub-optimal results and is known to be a NP-complex problem.
- K-means is proud to find clusters of equal size and the number of clusters is an input.

These methods can be used to carry out groupings of components that share similar functionalities or are subjected to similar working conditions. In general, the mechanisms that work under the same boundary conditions may use the same manufacturing materials. For example, if we have the data about the boundary conditions of all existing pipelines in a large industrial plant, this type of algorithm can be used to find patterns of use or to make clusters that can be useful when making new designs.

4.4. Principal Component Analysis (PCA)

Principal component analysis (PCA) refers to a group of methods which are employed to accomplish classifications by trying to perform dimensional reduction on the data. It can be also used to speed up other algorithms by decreasing the sample vectors size [31]. The main components of a data set are the directions in which the variance is greater and represent, in some way, the underlying structure [31, 32].

The way in which the variance is distributed in the data cloud gives an idea of the directions in which there is more information [31]. Those in which the samples are more spread out will have more relevance than those that are constant. The main components allow to summarize the data eliminating the redundant or little differentiating information and highlighting the one that is likely to be more important [32].

These algorithms have a limitation that makes them work better when the data are organized forming a linear manifold, that is, when the data is organized so that it can be projected on a hyperplane without producing relevant distortions. The PCA algorithm minimizes the error that occurs when projecting the data on a dimension $(n-1)$ hyperplane. At the same time, the algorithm will look for the orientation of the hyperplane that maximizes the variance [31, 32]. Figure 4 displays the 2 principal components of a dataset which shows that there is a high-variance direction for the data.

Some advantages of PCA are:

- PCA allows to reduce data dimensionality.
- PCA is a polynomial algorithm.

Some weaknesses of PCA are:

- PCA assumes data linearity and that the observed data are a linear combination of a certain base.
- PCS completely relies on the average and variance, which can produce statistical problems.

Given that $n_{max} = \max(n_{samples}, n_{features})$ and $n_{max} = \min(n_{samples}, n_{features})$, the complexity of PCA is $\mathcal{O}({n_{max}}^2 \cdot n_{min})$ and the memory complexity is $\mathcal{O}(n_{samples} \cdot n_{features})$ [31].

Given that the data matrix X is dimesion $n \times p$, where n is the number of elements in the dataset and p is the vector size of each element, the PCA algorithm will calculate an $n \rightarrow k$ reduction by following several steps [31]. Firstly, the covariance matrix is calculated (Eq. 6):

$$\sigma = \frac{1}{n}(X^T \times X) \tag{6}$$

Then, the eigenvalues of the covariance matrix are calculated (Eq. 7):

$$U = svd(\sigma) \tag{7}$$

where U is an $n \times n$ matrix whose columns will be the principal components. The reduction from matrix $X \in \mathbb{R}^{n \times p}$ to matrix $Z \in \mathbb{R}^{k \times p}$ is accomplished by truncating matrix U to obtain $\tilde{U}$, which implies taking only the first k columns to create the truncated matrix $\tilde{U}$ defined by Eq. 8:

$$\tilde{U} = U_{1:k} \in \mathbb{R}^{n \times k} \tag{8}$$

Finally, the reduced data matrix $Z \in \mathbb{R}^{k \times p}$ is calculated as shown in Eq. 9:

$$Z = \tilde{U} \times X \in \mathbb{R}^{k \times p} \tag{9}$$

These algorithms can be used to determine which features are most important when materials selection is performed and, therefore, reduce the range of possibilities.

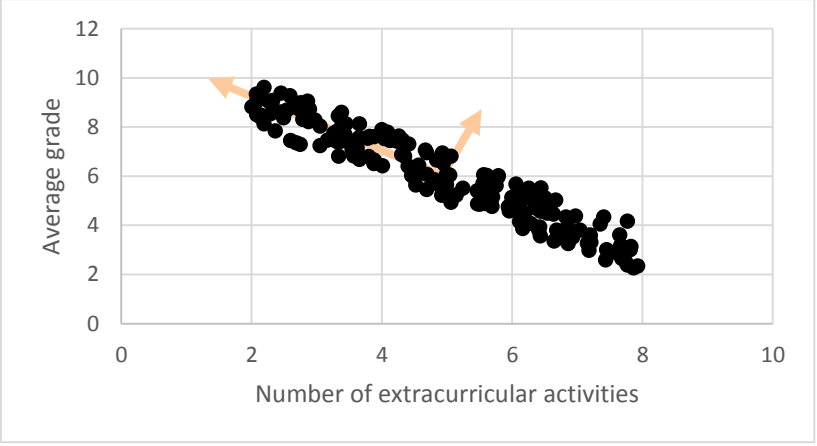


Fig. 4. PCA algorithm example in a 2-dimensional space.

5. Conclusions

Here are some of the conclusions that can be drawn from this study:

- AI techniques can be used to help designers to select the most suitable materials for each industrial application.
- Although the methodologies based on Supervised Learning require preliminary training and, in general, are more computationally demanding, more complex results can be obtained.
- AI-based algorithms can be used to extract information from large datasets, which are difficult to manage using traditional techniques.

- Artificial neural networks are a set of techniques able to learn nonlinear functions.
- The methodologies based on Unsupervised Learning allow to reach less ambitious results but can serve as a tool to simplify a dataset before carrying out a more complex treatment.

6. Acknowledgements

This work has been developed within the framework of the doctorate program in Industrial Technologies of the UNED. This work has been financially supported by the funds provided through the Annual Grant Call of the E.T.S.I. Industriales of UNED of reference 2019-ICF04.

References

[1] L. Pengzhen, C. Shengyong, Z. Yujun. Artificial Intelligence in Civil Engineering. Mathematical Problems in Engineering, 2012.
[2] S. Helal. The Expanding Frontier of Artificial Intelligence. Computer, 51, 9 (2018) 14-17.
[3] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79, 8 (1982) 2554-2558.
[4] D. Dimiduk, E. Holm, S. Niezgoda. Perspectives on the Impact of Machine Learning, Deep Learning, and Artificial Intelligence on Materials, Processes, and Structures Engineering. Integrating Materials and Manufacturing Innovation, 7, 3 (2018) 157-172.
[5] Z. Chang-Chun, Y. Guo-Fu, H. Xiao-Bing. Multi-objective optimization of material selection for sustainable products: Artificial neural networks and genetic algorithm approach. Materials & Design, 30, 4 (2009) 1209-1215.
[6] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, C. Kim. Machine learning in materials informatics: recent applications and prospects. NPJ Computational Materials, 3 (2017) 54.
[7] D. Merayo, A. Rodríguez-Prieto, A. M. Camacho. Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications. Proc IMechE, Part L: Materials Design and Applications. 2018.
[8] A. Rodríguez-Prieto, A. M. Camacho, M. A. Sebastián. Materials selection criteria for nuclear power applications: a decision algorithm. JOM 2, 68 (2016) 496-506.
[9] S. K. Pal, S. K. Meher, A. Skowron. Data science, big data and granular mining. Pattern Recognition Letters, 67 (2015) 109-112.
[10] D. Kim et al. Smart Machining Process Using Machine Learning: A Review and Perspective on Machining Industry. International Journal of Precision Engineering and Manufacturing-Green Technology, 5, 4 (2018) 555-568.
[11] D. Akdemir, J. Jannink. Ensemble learning with trees and rules: Supervised, semi-supervised, unsupervised. IDA, 18, 5 (2014) 857-872.
[12] A. Hafezalkotob, A. Hafezalkotob. Risk-based material selection process supported on information theory: A case study on industrial gas turbine. Applied Soft Computing, 52 (2017) 1116-1129.
[13] T. Tambouratzis, D. Karalekas, N. Moustakas. A Methodological Study for Optimizing Material Selection in Sustainable Product Design. Journal of Industrial Ecology, 18, 4 (2014) 508-516.
[14] C. Urrea, G. Henríquez, M. Jamett. Development of an expert system to select materials for the main structure of a transfer crane designed for disabled people. Expert Systems with Applications, 42, 1 (2015) 691-697.
[15] C. Lo et al. Editorial: Scientific advances in product experience engineering. Journal of Intelligent Manufacturing. 28, 7, 1581-1584, 2017.
[16] M. Przybyła-Kasperek, A. Wakulicz-Deja. Global decision-making in multi-agent decision-making system with dynamically generated disjoint clusters. Applied Soft Computing, 40 (2016) 603-615.
[17] V. S. Costa et al. Combining multiple algorithms in classifier ensembles using generalized mixture functions. Neurocomputing, 313 (2018) 402-414.
[18] W. Fan et al. Proportional data modelling via entropy-based variational Bayes learning of mixture models. Applied Intelligence, 47, 2 (2017) 473-487.
[19] J.R. Quinlan. C4. 5: programs for machine learning. Morgan Kaufmann, 1993.
[20] T. Hastie, R. Tibshirani, J. Friedman. Elements of Statistical Learning. Springer series in statistics, 2009.
[21] L. Breiman, J. Friedman, R. Olshen, C. Stone. Classification and Regression Trees. Wadsworth, 1984.
[22] K. Kim. A hybrid classification algorithm by subspace partitioning through semi-supervised D-tree. Pattern Recognition, 60 (2016) 157-163.
[23] J. Schmidhuber. Deep learning in neural networks: An overview. Neural Networks, 61 (2015) 85-117.
[24] A. Hyvärinen, E. Oja. Independent component analysis: algorithms and applications. Neural Networks, 13, 4–5 (2000) 411-430.
[25] G. Huang, G. B. Huang, S. Song, K. You. Trends in extreme learning machines: A review. Neural Networks, 61 (2015) 32-48.
[26] A. Ng, J. Ngiam, C. Y. Foo, Y. Mai. Backpropagation, Caroline Suen - Website, 2011.
[27] A. Vattani. k-medias requires exponentially many iterations even in the plane. Discrete and Computational Geometry, 45, 4 (2011) 596–616.
[28] M. Mahajan et al. The Planar k-medias Problem is NP-Hard. Lecture Notes in Computer Science, 5431 (2009) 274–285.
[29] J.A. Hartigan. Clustering algorithms. John Wiley & Sons, Inc. 1975.
[30] D. Arthur, S. Vassilvitskii. How slow is the k-means method?. Proceedings of the 22nd symposium on Computational geometry. ACM. 2006.
[31] M. Hoffman, D. Blei, F. Bach. Online Learning for Latent Dirichlet Allocation. NIPS, (2010) 856-864
[32] M. D. Hoffman, D. M. Blei, C. Wang, J. Paisley. Stochastic variational inference. JMLR, 14, 1 (2013) 1303-1347.

4.3 *Prediction of physical and mechanical properties for metallic materials selection using big data and artificial neural networks* [12]

Los indicios de calidad de este artículo pueden encontrarse en el Apéndice C.

4.3.1 Datos de la publicación y factor de impacto

Tabla 3. Factor de impacto de Prediction of physical and mechanical properties for metallic materials selection using big data and artificial neural networks

Título	Prediction of physical and mechanical properties for metallic materials selection using big data and artificial neural networks
Autores	David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Revista	IEEE Access
ISSN	2169-3536
Editorial	IEEE-INST ELECTRICAL ELECTRONICS ENGINEERS INC
País	Estados Unidos
Volumen	8
Páginas	13444-13456
Fecha	2020
doi	10.1109/ACCESS.2020.2965769
Factor de impacto	3.745 (2019 Journal Citation Reports) 61/266 (Engineering, Electrical & Electronic, Q1)

4.3.2 Resumen y copia de la publicación

En este trabajo se desarrolla una herramienta asistida por ordenador para predecir propiedades físicas y mecánicas destacadas que intervienen en las tareas de selección de materiales metálicos. El sistema se basa en el uso de redes neuronales artificiales apoyadas en una colección *big data* de información sobre las características tecnológicas de miles de materiales. Por tanto, el volumen de datos supera los 43k. El sistema puede acceder a una biblioteca de materiales gratuita en línea (un sitio web donde se registran los datos de los materiales), descargar la información requerida, leerla, filtrarla, organizarla y pasar a la fase basada en inteligencia artificial. Se construye una red neuronal artificial (RNA) con miles de perceptrones, cuya topología y conexiones se han optimizado para acelerar el entrenamiento y la capacidad predictiva de la RNA. Después del entrenamiento correspondiente, el sistema es capaz de realizar predicciones sobre la densidad del material y el módulo de Young con una precisión promedio superiores al 99% y 98%, respectivamente.

In this work, a computer-aided tool is developed to predict relevant physical and mechanical properties that are involved in the selection tasks of metallic materials. The system is based on the use of artificial neural networks supported by big data collection of information about the technological characteristics of thousands of materials. Thus, the volume of data exceeds 43k. The system can access an open online material library (a website where material data are recorded), download the required information, read it, filter it, organize it and move on to the step based on artificial intelligence. An artificial neural network (ANN) is built with thousands of perceptrons, whose topology and connections have been optimized to accelerate the training and predictive capacity of the ANN. After the corresponding training, the system is able to make predictions about the material density and Young's modulus with average confidences greater than 99% and 98%, respectively.

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 18, 2019, accepted December 31, 2019, date of publication January 10, 2020, date of current version January 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2965769

Prediction of Physical and Mechanical Properties for Metallic Materials Selection Using Big Data and Artificial Neural Networks

D. MERAYO, A. RODRÍGUEZ-PRIETO, AND A. M. CAMACHO
Department of Manufacturing Engineering, National Distance Education University (UNED), 28040 Madrid, Spain

Corresponding author: D. Merayo (dmerayo1@alumno.uned.es)

This work was supported by the Annual Grants Call of the Escuela Técnica Superior de Ingenieros Industriales (Higher Technical School of Industrial Engineers) (ETSII), of the Universidad Nacional de Educación a Distancia (National University of Distance Education) (UNED), through a project under Grant 2020-ICF04.

ABSTRACT In this work, a computer-aided tool is developed to predict relevant physical and mechanical properties that are involved in the selection tasks of metallic materials. The system is based on the use of artificial neural networks supported by big data collection of information about the technological characteristics of thousands of materials. Thus, the volume of data exceeds 43k. The system can access an open online material library (a website where material data are recorded), download the required information, read it, filter it, organise it and move on to the step based on artificial intelligence. An artificial neural network (ANN) is built with thousands of perceptrons, whose topology and connections have been optimised to accelerate the training and predictive capacity of the ANN. After the corresponding training, the system is able to make predictions about the material density and Young's modulus with average confidences greater than 99% and 98%, respectively.

INDEX TERMS Artificial intelligence, big data, material selection, multilayer feedforward networks, neural network, property prediction, software-based web browser control.

LIST OF SYMBOLS AND ABBREVIATIONS

ADAM	Adaptive Moment Estimation
AI	Artificial intelligence
ANN	Artificial Neural Networks
β_n	ADAM algorithm parameter
ϵ	ADAM stability factor
ε	Prediction error of a neural network
η	ADAM step size
f	Error function
g	Gradient of the error function
HTML	HyperText Markup Language
m	ADAM first moment estimate
v	ADAM second moment estimate
w	Weights vector

I. INTRODUCTION

The physical and mechanical properties of a material have a fundamental role in the performance of industrial components. A correct operation depends, to a large extent, on the characteristics of the materials that constitute it as insufficient material properties can cause premature component failure.

In this work, the density and Young's modulus have been chosen as study variables due to their relevance in engineering materials selection tasks. These two characteristics have been profusely applied and investigated, which enables the construction of Ashby diagrams [37], which is an efficient tool in the selection of materials.

Knowing the properties of the materials involved in industrial designs has importance; however, obtaining these data often requires access to a large amount of resources, which are generally not available. A multitude of tests are needed to obtain really significant information, which implies that sufficient time, personnel and facilities must be available at a cost [1]. The process of characterisation of a material may require a battery of tests that requires a considerable amount of time and the investment of vast amounts of resources.

The use of a methodology this is based on trial and error should not be considered since each test implies, as previously indicated, a considerable consumption of resources that are generally scarce. To develop a material from scratch, and

The associate editor coordinating the review of this manuscript and approving it for publication was Huanqing Wang.

after a long testing period, realise that it is not valid or does not satisfy the initial specifications, is not possible.

Although extensive libraries of materials with abundant and very detailed information exist, some data needed to perform the work is not always available [2]. Even the most basic information may not be accessible or available, especially when dealing with infrequent, recently developed or the latest technology materials. However, the industry demands that our designs should be taken to their limits, which implies increased use of technological materials [3].

Material libraries can be used to extract trends or general characteristics; however, as the amount of data to be considered increases, the task becomes painful or even impossible [4]. To solve this type of situation, assistance tools can be developed to help the designer [5]. In this way, design loops are reduced and people can concentrate on more creative tasks. The availability of open data becomes increasingly important, which justifies the development of tools to treat information, extract trends or generate new knowledge.

Thousands of material data with adequate tools to extract trends enable the anticipation of the expected properties of a new material without the need for any type of test and limits errors [6]. In this way, research routes can be eliminated before they consume valuable resources that can be exploited in other more productive studies [7].

In this work, the ability of neural networks to extract trends is employed to obtain new information from a material database. The heterogeneity of these data makes traditional database management techniques insufficient, and therefore, the problem must be addressed using newer techniques based on big data.

Big data consists of large and complex data sets, especially from new data sources. Due to the bulkiness of these data sets, they cannot be managed by conventional data processing software [8]. However, these massive volumes of data can be used to address problems that, previously, would not have been possible to solve.

Although the term ''big data'' is relatively new, the action of collecting and storing large amounts of information for further analysis has been performed for many years [9]. The concept gained momentum when the industry analyst Doug Laney articulated the current definition of big data as the three Vs [10]:

- Volume: the amount of data matters. With big data, large volumes of unstructured low-density data are processed. The data can be data of unknown value, such as machining conditions, material properties or manufacturing control measures. For some organisations, tens of terabytes of data are employed; for other companies, even hundreds of petabytes of data are utilised.
- Velocity: the rate at which the data are received, and possibly, to which some action is applied. The higher data-transmission rate is usually conducted directly to memory rather than written to a disc. Some intelligent products work in real time (or practically in real time) and require instantaneous evaluation and action.
- Variety: the various types of data that are available. Conventional data types are structured and can be clearly organised in a relational database. With the rise of big data, data are presented in new types of unstructured data. Unstructured and semi-structured data types, such as text, audio or video, require additional pre-processing to obtain meaning and enable metadata. In addition, big data applications usually cope with sparse information.

Big data enables answers that are more complete to be obtained since more information is available [10]. Having more complete solutions to a problem enables the problem to be addressed with greater guarantees of success and enables resources to be managed a more efficient way [9]. The availability of more complete answers also means greater reliability of information, which implies a completely different approach to addressing problems.

Big data have proven to be a very useful tool in research in materials science and technology. In recent years, many works on this subject have been developed, numerous applications have emerged and important initiatives aimed at serving as a basis for scientists in this field have appeared. Predictive models have been developed based on big data and machine learning [11], and assisted design tools supported by large databases that transform software into an expert system [12] and entities such as the Material Genome Initiative have appeared. The mission of this initiative is to reduce the cost and development time of material discoveries, optimisation and deployment by offering access to large material databases and providing the tools required to ease investigation.

The libraries of materials available on the Internet contain large collections of records that, in general, are incomplete, that is, some properties are not available for a material [13]. The information is actually very sparse, which hinders the ability to obtain valid conclusions using conventional tools [14].

Even though some material libraries are freely accessible [13], in general, downloading large amounts of data is difficult as companies regain their most important information with great zeal.

In this work, the data contained in these libraries have been used to train a system based on artificial intelligence (AI) techniques to extract new information that was not initially available. Note that the quality of the conclusions of this type of methodology is as good as the quality of the data from which the conclusions have been drawn [13]. In this case, the available data have been provided by the material manufacturers themselves and have been certified by standard tests.

Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems [6]. These processes include learning (acquisition of information and rules for the use of information), reasoning (using the rules to reach approximate or definitive conclusions) and self-correction [15].

Often classified in the cognitive science group, AI uses computational neurobiology (particularly neural networks),

mathematical logic (sub-discipline of mathematics and philosophy) and computer science [16]. AI looks for problem-solving methods with high logic or algorithmic complexity. By extension, AI designates devices or tools by imitating humans in certain implementations of its cognitive functions [15].

The term ''artificial intelligence'' was coined in 1956 during the Dartmouth Conference, where the discipline emerged [17]. Currently, AI is a general term that addresses automation of robotic processes to current robotics. AI has recently gained prominence due to the large volumes of data or to the increase in speed, size and variety of information collected by companies [6]. AI can perform tasks, such as identifying patterns in data more efficiently than humans, which enables users to obtain more information about their data [16].

Since the birth of computer science, a large number of methods and techniques that can be framed within the field of AI have been developed, of which some are more useful than others [18].

Technology enables us to produce materials with an extensive variety of properties; these materials can work very differently depending on the environment and working conditions [6]. This finding implies that a large amount of data are available, and therefore, the situation will become unmanageable using only usual calculation methods. We need tools that can immerse themselves in these data and propose the optimal solution among the options.

A large variety of methods can be framed within the field of artificial intelligence. Among them, artificial neural networks (ANNs) stand out as they have become paradigmatic techniques due to their incredible achievements and unstoppable progress within the field of computing [19].

Artificial neural networks comprise a computational model that is slightly inspired by the behaviour observed in the brain. An artificial neural network consists of a set of units that are referred to as artificial neurons, which are connected to each other to transmit signals. The input information transmits across a neural network, undergoes various operations and produces output values [19].

Frank Rosenblatt defined the perceptron model, which is a supervised learning algorithm of binary classification [20]. This model is a formal neuron with a learning rule that automatically determines synaptic weights to decide whether an input belongs or does not belong to a class. If the problem is linearly separable, a theorem ensures that the rule of the perceptron enables the identification of a separator between the two classes [19].

A set of perceptrons that is organised in several layers has the ability to correctly manage nonlinearly separable problems. These systems rely on error gradient backpropagation models and enable the construction of neuronal models that are as complex as necessary [21].

As systems that are capable of learning, neural networks implement the principle of induction, which is learning by experience [6]. By confronting specific situations, neural networks infer an integrated decision system whose generic character is a function of the number of learning cases that are encountered and their complexity in relation to the complexity of the problem to be solved [39], [41]. Therefore, complex problems require more training.

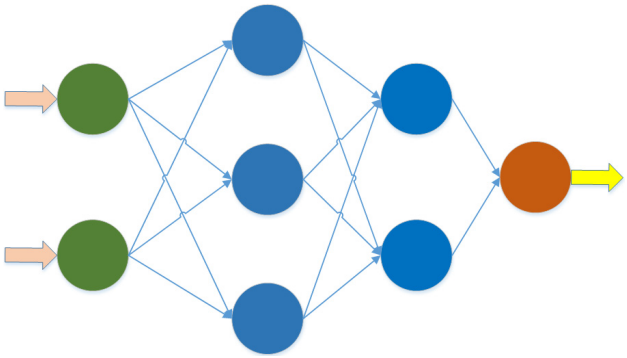


FIGURE 1. Simplified model of a multi-layer artificial neural network.

Many topological models and neural network architectures exist, among which multi-layer neural networks account for the most prominent networks [21]. Each network attempts to solve some of the problems posed by this type of algorithm: learning process optimisation, reduction of resources for training, and ability to quickly learn certain functions. These systems generally have great plasticity and can be adapted to most types of problems. However, they require large amounts of resources and input data for their training [22].

Fig. 1 shows a simplified model of a multi-layer neural network that is composed by 8 perceptrons, which are organised in 4 layers: 2 perceptrons in the input layer (the layer that receives the data from the exterior), 5 perceptrons in the two hidden layers and 1 perceptron in the output layer (the layer that outputs the processed information).

Table 1 contains some of the most relevant neural network topologies, including a small explanation about its model.

Currently, neural networks are applied in many fields of science and engineering, from control systems [44], [48] to business. Neural networks are a suitable alternative to a large number of methods that are applied in numerous fields. In most cases, even for problems that have been solved by other means and due to different theories, neural networks have identified other more efficient forms of resolution.

Neural networks have a large number of real uses in the industry; they have already identified many commercial applications, as they show better results in the recognition of data patterns or trends than other techniques based on mathematical analysis. Artificial neural networks have proven to be excellent tools for analysing large data sets as, after a training process, they can extract trends that may remain hidden for other conventional systems of information analysis [21].

Our current technological development enables us to produce large amounts of data. However, these data must be processed to obtain reliable and useful information that enables better decisions [6]. Artificial intelligence can help

TABLE 1. Classical neural network topologies [23], [39]–[42].

Topology	Description
Perceptron (P)	A simple mathematical model of a neuron, which simulates the behaviour of a single biological neuron.
Multi-layer feed forward (MLFF)	A model that contains 3 types of layers (input, hidden and output), which propagates the error during the training. These networks can learn any nonlinear function.
Recurrent neural network (RNN)	A MLFF model, whose neurons are trained, not only use information from other neurons but also information from themselves from the previous iteration.
Long/short term memory (LSTM)	A model that attempts to avoid the vanishing/exploding gradient problem of the RNN by introducing gates and an explicitly defined memory cell.
Gated recurrent unit (GRU)	A model similar to LSTM that contains update gates that determines the quantity of information to retain from the last state and the quantity of information to receive from the previous layer.
Auto encoder (AE)	Similar to MLFF. A model that intended to automatically encode information in a compressed way.
Variational AE (VAE)	A model with the same architecture as AE, including some probabilistic cells that avoid non-desired information propagation.
Denoising AE (DAE)	An AE model that is intended to avoid the network to learn details by introducing more noise as input.
Sparse AE (SAE)	An AE model whose aim is to encode information using more space.
Hopfield network (HN)	A fully connected model that always converges to a local minima and whose nodes simultaneously act as input, hidden and output nodes.
Boltzmann machine (BM)	Similar to HN although some neurons are considered input neurons and other neurons remain hidden.
Restricted BM (RBM)	A model similar to BM where some nodes are linked to other nodes; they are mostly grouped.
Deep belief network (DBN)	A model that stacks several RBM or VAE. These networks are able to generate new data.
Deep convolutional network (DCN)	A model that can easily process images or audio and tag them.
Deconvolutional network (DN)	A model that, after training, can produce pictures that are related to a given concept.

designers identify the optimal solution for each task, which can potentially increase the total performance [18].

One of the main tools for the search of the optimum material for an industrial application are Ashby diagrams [37]. This type of scattered plot enables two or more properties of many materials (or classes of materials) to be simultaneously and intuitively visualised. To generate this type of graphic, however, the required data are necessary.

As previously indicated, the libraries of existing free access materials on the internet enable a large amount of data to be available [13]. However, they generally do not authorise users to download complete libraries in a simple way and only allow individual downloading of the data of a single material. In this context, the development of tools that capable of iteratively carrying out this painful task is necessary [14].

These types of websites are protected against the action of bots that can carry out attacks that are aimed at causing the system downfall. Thus, the data download procedure that is developed acts in a similar way to the procedure that would follow a human being, and therefore, building software tools based on web browser control that imitates a user's behaviour is essential [24].

Simplifying, this technique consists of developing code that instructs the web browser that it should behave as if the user had performed daily actions, such as pressing a button, entering a text or downloading a file.

Software-based web browser control consists of the development of code that is aimed at employing browser capabilities: HTML code interpretation, file downloading, permission management, and cookies administration. A web browser provides a wrapper that frees the developed code from the complexity of the operations that are needed to surf the internet [24].

The objective of this work is to determine how artificial neural network technology can be used to exploit a large set of metallic material data [40] to predict its physical and mechanical properties. This work explores the feasibility of training a neural network to ensure that it can predict some physical properties based on the chemical composition of the new materials presented to it.

The purpose of this paper is to establish a framework that establishes the basis for developing the required software and algorithms to determine a broad spectrum of material properties based on their chemical composition. The framework is envisioned as a support tool for the study of materials science.

The main goal of this study is to build an artificial neural network that can receive the material information as input data and make valid predictions. The network topology is optimised to perform this task and enables the collection of output data that is subsequently analysed to calculate the performance of an entire system.

The objectives include the development of software tools that facilitate this work to be carried out (including tools with which the obtained results are analysed) and the procurement of a large set of input data about metallic materials, including their physical properties and chemical composition, and the organisation of big data to ensure that the data were exploitable via the developed algorithms and tools.

The innovation of this work is the use of neural network techniques to build a complete Ashby diagram based on the chemical composition of metals. In addition, the use of artificial neural networks that are trained with a large set of metallic material data to predict their mechanical and physical properties is remarkable.

Neural networks have proven to be very useful in multiple fields to obtain very significant results when combined with reinforced learning techniques using large data sets. However, a limited amount of work applies ANNs to the field of metal science, and most studies that develop procedures aimed at predicting their physical and mechanical properties are based on mathematical or statisticians models instead of artificial intelligence.

The procedure proposed in this work is perfectly scalable and applicable to different types of metals; however, most models can only be applied to a restricted group of

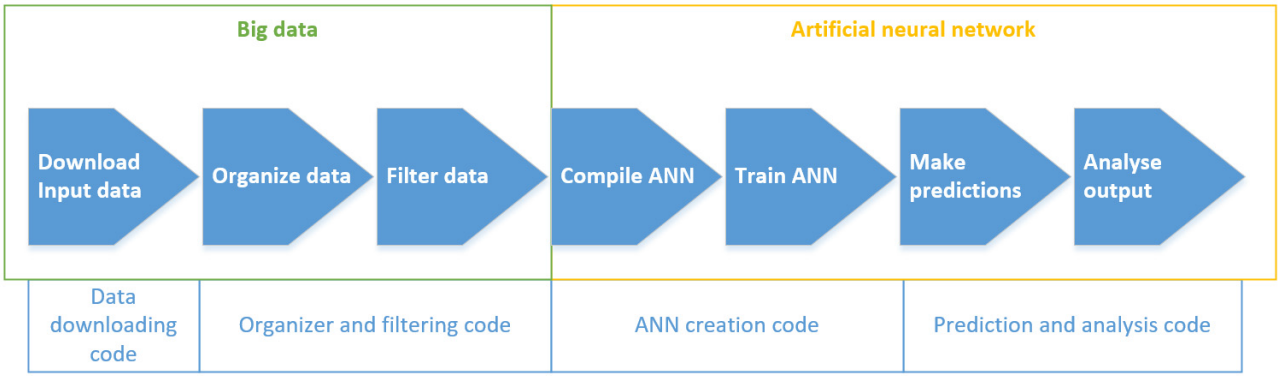


FIGURE 2. Methodology scheme.

materials [49]. The novelty of this methodology is the application of artificial intelligence to the prediction of physical and mechanical properties of metal alloys.

The added value of this methodology consists of the use of artificial intelligence techniques to predict some important properties for the selection of materials in engineering applications. Numerical methodologies only focus on a limited and closed number of material characteristics, while systems based on neural networks can potentially take advantage of all available information to make better decisions.

II. METHODOLOGY

The development of the work has focused on obtaining an artificial neural network that is capable of making adequate predictions about material properties while maintaining a limited average error. Subsequently, the output data, data about the network training process and data about the prediction step are conveniently analysed [16].

A. PRINCIPLES OF THE METHODOLOGY: GENERAL OVERVIEW

Fig. 2 schematically shows the different steps for developing the system. The system has two main stages: the stage related to big data and the collection of input data; and the stage related to the development and use of the artificial neural network (ANN), which also comprises the data analysis. The schematic shows which programming code manages each step.

Matmatch [25] is a well-known open-access materials library that contains information about thousands of different commercial and standard materials. All registered users can freely access the information stored in their databases, whose data are provided by suppliers and manufacturers. A specification sheet that contains all available data can be downloaded for each material. However, the full bulk cannot be downloaded at once, that is, the entire library of materials cannot be downloaded. A software bot that is capable of iteratively downloading (separately) all materials available on the website has been developed.

Once the bot completes its task, therefore, we obtain a collection of thousands of specification sheets about materials that generally contain information about chemical composition, mechanical properties and physical properties. This set of datasheets covers all kinds of materials: metals, ceramics, composites, and polymers.

A second software application runs through all available specification sheets, extracts all contained information (even information that is not relevant for the development of this work), and carries out a small processing and organises it into an easy access matrix. The previously mentioned initial processing consists of the homogenisation of units of measurement and the elimination of non-numerical information (with the exception of the name and the identification, a unique code for each material), which is not susceptible of being exploited.

At this point, we obtain an immense matrix of very sparse data, where each row corresponds to a material and each column represents a property [8]. As previously indicated, some data are not available for certain properties of materials and since the database is very heterogeneous, some properties do not make sense for certain materials [1]. Therefore, this matrix must be filtered to obtain valid input information.

A third code is executed to carry out the filtering of data. Only metallic materials whose chemical composition is defined in more than 90% are considered. During the development of this work, a small amount of the specification sheets did not include the chemical composition of a material or did not correctly detail it. To avoid bias during the training phase, incomplete materials are eliminated [26].

The following chemical elements have been investigated: Al, Fe, Mn, Si, Cu, C, Cr, P, S, Ni, Zn, Mg, Ti, Mo, Pb, Sn, V, N, Nb, W, Sb, As, Bi and Co [27]. The presence of other elements in small quantities does not affect the learning process of the neural network [26]. Note that only metallic materials whose main components are included in this list are considered (i.e., no information related to gold or silver has been considered even if their datasheets have been downloaded).

Fig. 3 graphically shows the results of filtering all available downloaded materials (43575). Materials that are not considered are split into two categories: non-metallic materials (16473) and materials whose chemical definition is

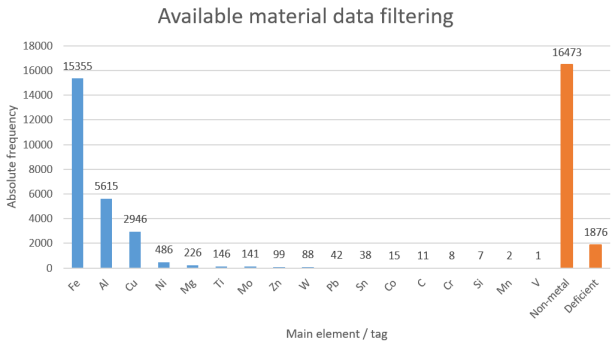


FIGURE 3. Available material data filtering detail.

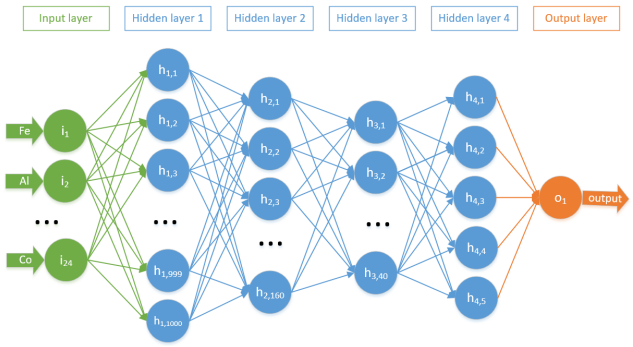


FIGURE 4. Simplified artificial neural network topology.

deficient (1876). The materials that have been effectively employed (25226) are divided according to their main chemical element (17 categories). Materials with a deficient chemical definition include those whose main element is not included in the previously mentioned elements list.

Once the material data are filtered and all information is guaranteed useful and relevant to develop this study, an adequate big data structure is already available. Therefore, the fourth software code is launched to initiate the step based on artificial intelligence [28].

A multilayer feedforward architecture has been chosen and an artificial neural network with a fully connected topology, which consists of an input layer, 4 hidden layers and an output layer, has been defined [22]. The input layer consists of 24 nodes, which correspond to each of the considered chemical elements; the hidden layers are formed by 1000, 160, 40 and 5 perceptrons; and the output layer is formed by a single node. Fig. 4 shows a simplified representation of the neural network topology in this work. The input layer, the hidden layers and the output layer have been differentiated using colours.

This topology is the result of successive optimisation steps to balance its learning capacity and the necessary resources for its training [29]. Note that a complex topology is capable of learning more complex functions than a simple topology but requires additional resources during its training: additional time, calculation capacity and input data [16]. A balance between the network depth and the network width was obtained.

A fully-connected artificial neural network consists of a set of fully connected layers; a fully-connected layer is a layer in which all nodes are connected to all nodes of the next layer [21]. This network is able to learn the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, where n is the input vector size (in this case, the amount of considered chemical elements is 24) [42].

Hornik [30] showed that the multilayer feedforward architecture provides neural networks with the potential of being universal approximators. Even if a fully connected ANN can represent any function, it may not be able to learn some functions as backpropagation convergence is not guaranteed [22].

Since the network architecture and topology were previously defined, the training phase of the network begins. The available data are randomly divided into two disjoint subsets: training subset and test subset. The first subset comprises 80% of the data, while the second subset contains the remaining 20% of the data. For a data to be used during the learning phase, all input information and expected results are necessary since neural networks are a supervised learning technique.

The network training is subject to the following conditions:

- Calculation of the learning rate for each parameter using Adaptive Moment Estimation (ADAM) where $\beta_1 = 0.9$, $\beta_2 = 0.999$ (algorithm parameters), $\eta = 0.001$ (step size) and $\epsilon = 10^{-8}$ (stability factor) [31].
- Early stopping after 10 iterations without significant changes to avoid overfitting.
- Training stops when a training error of less than 0.001 is reached as it is considered negligible [32].
- Maximum of 100000 training epochs to avoid infinite loops.
- Sigmoid activation function.
- Verbose output to analyse the training process.

Once the neural network has been trained using the corresponding subset of data, predictions can be made using the information from the test subset. In this way, metrics that enable us to measure the performance of the system can be extracted. Several statistical parameters are calculated by comparing the output estimation of the neural network with the real value; these parameters are employed to describe the performance of the network.

B. SOFTWARE AND TOOLS

As previously indicated, multiple software codes have been developed to carry out each of the tasks of this work (refer to Fig. 2). These codes have been developed entirely in Python 3.7 language, although multiple libraries and non-standard modules have been used to simplify some stages of the project. This language has been chosen as Python is a high-level, multi-platform and multi-paradigm programming language with dynamic typing that has great popularity among developers [33], especially, among those who develop software related to artificial intelligence [34].

Some external modules have been employed, such as Selenium (library that enables control of the web browser by using code), BeatifulSoup (library that facilitates working

with HTML files), SciPy (library that contains numerous scientific, mathematical and statistics functions), NumPy (library that enables easy management of large amounts of data and large matrices and numbers), MatPlot (library that eases the production of plots, figures and graphics) [33] and TensorFlow with Keras (high-level library that contains a vast amount of functions and procedures related to artificial intelligence, especially artificial neural networks) [34].

This system comprises more than 9000 lines of code distributed in 11 files that manage different operations.

C. MATERIAL INPUT DATA ACQUISITION FOR FURTHER EVALUATION

As previously indicated, information about materials has been downloaded from an online library via a code that mimics the behaviour of a human user. This task has been achieved by using Gecko Driver 0.24 for Firefox 67.0 and the Selenium module for Python [33]. Gecko Driver and Selenium enables a programmer to ask the web browser to perform actions such as going to a given web page, clicking on an element (i.e., a button or hyperlink) or writing some text in a text-box.

The data downloading code works by the following several steps: identify location of the names of the materials to download, filter to eliminate repeated data, access web record of each material and download the summary data file. This part of the work is extremely slow since it is substantially affected by the speed of the servers where the online library is hosted and the bandwidth of the local network.

After downloading the summary record of each material, we obtain more than 43000 individual files that contain a vast amount of data. These files must be read again to extract their information and organise it in a way that is easily accessible. Once the information is organised in a vast sparse matrix, filtering should be carried out to eliminate data that are not useful or that will not be employed.

These organisation and filtering processes are also carried out via code that was specifically developed for this work.

D. TRAINING AND PREDICTION USING ARTIFICIAL NEURAL NETWORKS

The procedure by which a neural network learns is referred to as training and is mathematically based on the problem of gradient descent [32]. The learning problem of an artificial neural network is the minimisation of the associated error function. This function usually consists of 2 terms: the first term evaluates the adjustment of the network output with respect to the available data (error term) and the second term (regularisation term) is used to avoid overfitting (overfitting is avoided by allowing the training process to eventually stop before overfitting occurs) [35].

The error function entirely depends on the weights associated with each of the perceptrons. This vector of weights with a size equal to the number of neurons (1230) is represented as w and enables us to indicate that $f(w)$ is the error made by the neural network when assigning the weights w to each of the perceptrons that compose it [36]. With this formalisation, the objective of the training is to find the vector $w*$ for which a global minimum of the function f is obtained, which converts the learning problem into an optimisation problem [35].

In this way, a neural network is initialised with some weight vectors (in general, randomly chosen). A new parameter vector is calculated to reduce the error function. This process is repeated until the error has been reduced under a tolerable threshold or when a specific stop condition is satisfied [34].

Since the error function is derivable, the gradient of this function can be defined for each of the optimisation steps (refer to Eq. 1):

$$g_i = \nabla f_i = \nabla f\,(w_i) \tag{1}$$

where w_i is the weight vector for the *ith* optimisation step; f_i is the value of the error in the *ith* step of the iteration; and g_i is the gradient value of the error function in the *ith* step of the iteration [35].

Adaptive moment estimation (ADAM) is an adaptive learning rate method that computes individual learning rates for different parameters. ADAM uses estimations of first and second moments of a gradient to adapt the learning rate for each weight of the neural network [31]. Using this method, in each iteration, the new weights vector is calculated as (refer to Eq. 2):

$$w_{i+1} = w_i - \eta \frac{\hat{m}_{i+1}}{\sqrt{\hat{v}_{i+1}} + \in} \tag{2}$$

where η is the step size (a value that graduates the relevance of the gradient factor), $\in$ is the stability factor of the algorithm (constant) and the bias-corrected first moment estimate $\hat{m}_{i+1}$ and second moment estimate $\hat{v}_{i+1}$ are calculated as follows (refer to Eq. 3 and 4) [31]:

$$\hat{m}_{i+1} = \frac{m_{i+1}}{1 - \beta_1^{i+1}} \tag{3}$$

$$\hat{v}_{i+1} = \frac{v_{i+1}}{1 - \beta_2^{i+1}} \tag{4}$$

where β_1 and β_2 are the algorithm parameters that are set to a value near 1 [31]. m_{i+1} and v_{i+1} are calculated as follows (refer to Eq. 5 and 6):

$$m_{i+1} = \beta_1 m_i + (1 - \beta_1)\, g_{i+1} \tag{5}$$

$$v_{i+1} = \beta_2 v_i + (1 - \beta_2)\, g_{i+1}^2 \tag{6}$$

where m_{i+1} and $v_{i+1}$are the decaying averages of past gradients and past squared gradients, respectively, and are estimates of the first moment (mean) and the second moment (uncentred variance), respectively, of the gradients [31].

The optimisation process and the network training method has been mathematically defined. Once the network has been conveniently trained, predictions can be obtained based on the approximation function learned by the neural network [34].

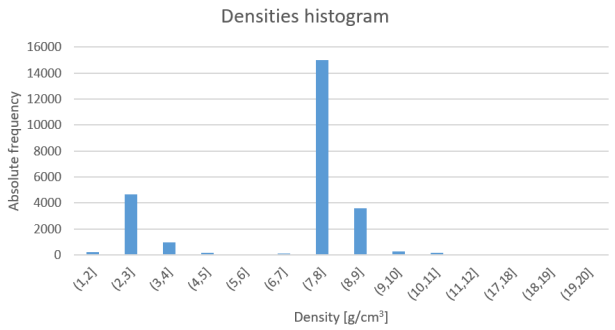


FIGURE 5. Densities histogram.

The prediction deviation is calculated as the absolute value of the relative error of the resulting value (refer to Eq. 7):

$$\varepsilon = \left| \frac{v_{prediction} - v_{real}}{v_{real}} \right| \tag{7}$$

III. RESULTS AND DISCUSSION

Although the training algorithms are randomly initialized, the results (both during training and during prediction) are very stable and converge to the same values with differences below 1% [36]. Only the results of the network that have obtained the best predictive metrics are shown.

Once the neural network is conveniently trained with 80% of the records (training subset), the ANN is requested to make predictions with the remaining 20% of the data (testing subset). In the second step, the network is not given any information about the expected results since this information is information that it must return.

Training and testing using the same dataset is not recommended as bias and overfitting can occur; thus, the obtained results can be fabricated since the performance of the network is not realistic [34].

A. DENSITY TRAINING AND PREDICTION

The neural network is trained with 20180 records of randomly chosen materials, that is, 80% of available materials. For each sample, the ANN is given the percentages of each chemical element (of the 24 elements that are considered) and the density of each material. In this way, the neural network is asked to identify a function that serves as an approximator for the calculation of the density.

Fig. 5 shows a histogram of all densities of the considered materials. The distribution is very irregular as the number of samples exceeds the number of elements. Three remarkable ranges appear: (7, 8], which includes ferrous materials; (2, 3], which includes aluminium alloys; and (8, 9], which includes, among others, brass (Cu + Zn). For these elements, the dataset contains a larger number of samples as they are broadly used materials.

The density of a metal alloy can be well approximated by calculating a weighted average, in which the total density is equal to the sum of the products of the proportions of each element multiplied by its individual density.

TABLE 2. Density training details.

Training samples	20180
Iterations	159
Min. error function value	0.002
Initial error function value	20.976

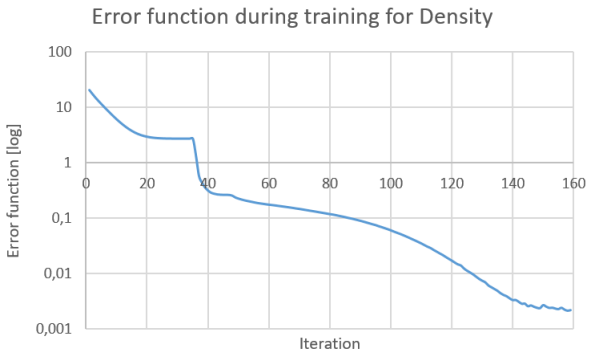


FIGURE 6. Error function during training for the density (log scale).

Note that the ANN is not given the density of each of the individual chemical elements, which is a concept that the neural network must learn supported by the supervised training.

The density-related training process that required 159 iterations ended as no significant improvements occurred during 10 iterations (early stop condition), and a minimum error function value of 0.002 was attained. These data are listed in Table 2. Note that this very low final error function value indicates that the approximation that the network has learned matches the provided data [34].

Fig. 6 shows the evolution (on logarithmic scale) of the error function during the training. Near iteration 35, a very important change in trend occurs; in this step, the ANN learned a significant concept that enabled the error function to be considerably reduced. After iteration 140, the slope of the error function becomes flatter, which indicates that the training process is almost finished (no additional significant improvement).

The neural network is asked to make predictions about the 5046 records contained in the testing subset. For these records, the real density values are known but are not communicated to the ANN as they are retained to calculate some performance metrics.

Table 3 shows some statistical metrics that enable evaluation of the average prediction performance of the network. As shown, the average prediction deviation is 0.448%, and half of the samples are below 0.237% (median). The difference among the average, trimmed mean and median indicates that outliers exist and should be carefully investigated.

Fig. 7 shows the histogram of the errors made in the prediction of the densities of the materials of the testing subset. The samples are accumulated in the first two error ranges, that is, for deviations less than 0.5%. The error is distributed and forms a large peak near 0%, and a long tail of low height appears.

TABLE 3. Density prediction error testing details.

Test samples	5046
Average	0.448%
Std. Dev.	0.855%
Median	0.237%
Maximum	15.157%
Minimum	0.000%
Trimmed mean (±5%)	0.352%
Quartile 1	0.171%
Quartile 2	0.237%
Quartile 3	0.341%

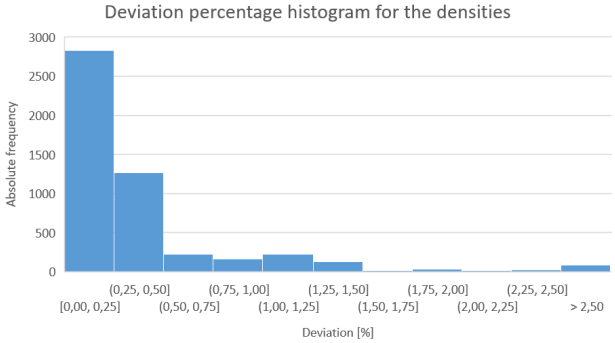


FIGURE 7. Prediction deviation histogram for the density.

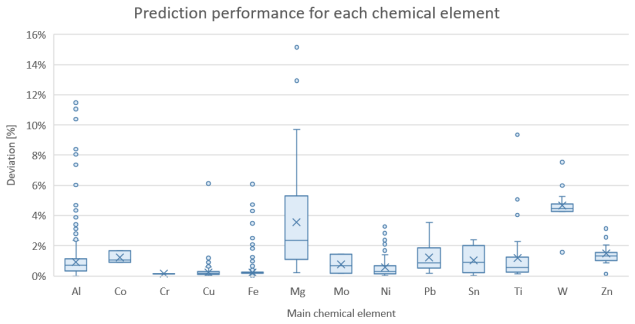


FIGURE 8. Density prediction performance for each chemical element.

Fig. 8 shows the performance of the predictions for each of the main chemical elements of the alloys in the testing subset. Note that the ANN offers substantially worse results for magnesium and tungsten (note that a few samples can be used during the training, which impairs the learning performance). The big box and long whisker related to the magnesium alloys indicates a large predictive variability.

As foreseen by the statistical metrics, Fig. 8 shows some outliers that have been associated with poorly defined materials (note that the quality of the input data is not perfect, although the ANN can address these imperfections). In addition, the neural network encounters problems with some duralumin alloys (they appear as outliers in the chart).

For each main chemical element, Table 4 shows the amount of samples that have been employed during the tests (materials in the testing subset) and the average deviation of the artificial neuronal network when it is making predictions. Significant differences are observed among the elements, although the average error remains bounded. For elements

TABLE 4. Main element related average error.

Main element	Average error [%]	Samples
Al	0.91%	1113
Co	1.22%	3
Cr	0.15%	1
Cu	0.22%	584
Fe	0.23%	3084
Mg	3.55%	48
Mo	0.78%	30
Ni	0.58%	100
Pb	1.23%	9
Sn	1.05%	4
Ti	1.20%	32
W	4.67%	15
Zn	1.48%	23
TOTAL	0.45%	5046

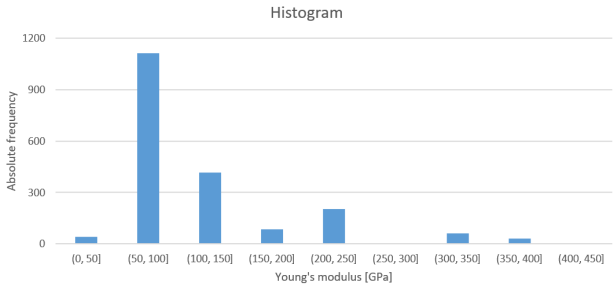


FIGURE 9. Young's modulus histogram.

with a large number of samples (i.e.: aluminium, iron or copper), note that acceptable results are obtained as the ANN successfully learns the density calculation function.

B. YOUNG'S MODULUS TRAINING AND PREDICTION

In this case, the Young's modulus (E) is not available for all materials. Therefore, only materials for which this information is accessible will be considered as these data are necessary in the supervised training phase and the performance estimation phase. Fig. 9 shows the distribution of the Young's modulus among the considered materials. Note that the ranges that contain aluminium alloys (50-100GPa), copper alloys (100-150GPa) and iron alloys (200-250GPa) are much more prominent than the other materials as the dataset contains more samples.

The ANN was trained with 1571 randomly chosen material samples obtained from the samples that can be considered (only 1962 registers contain the Young's modulus value), which is 80% of the data (training subset). For each record, the ANN is given the percentage of each chemical element (among the 24 elements considered) and its corresponding Young's modulus. The neural network is asked to find a function that serves as an approximator for the calculation of the Young's modulus.

The Young's modulus training process required 6693 iterations and ended as no significant improvements were observed during 10 iterations (early stop condition) and a minimum error function value of 548.94 was reached. These data are listed in Table 5.

TABLE 5. Young's modulus training details.

Training samples	1571
Iterations	6693
Min. error function value	548.94
Initial error function value	9123.58

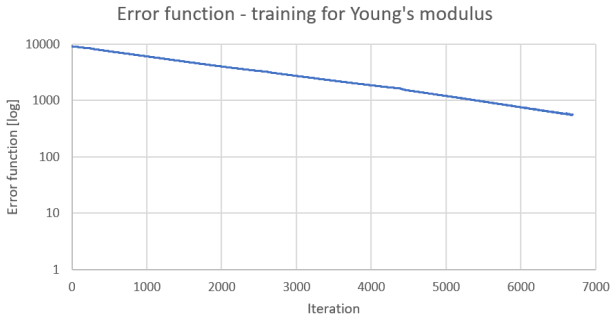


FIGURE 10. Error function during training for the Young's modulus (log scale).

TABLE 6. Young's modulus prediction error testing details.

Testing samples	390
Average	1.561%
Std. Dev.	2.233%
Median	0.667%
Maximum	12.446%
Minimum	0.003%
Trimmed mean (±5%)	1.394%
Quartile 1	0.219%
Quartile 2	0.667%
Quartile 3	1.843%

The final error function value is substantially larger than the density value, which indicates that the training process was not equally performant, and therefore, the predictive capacity of the ANN will be inferior to the previous capacity.

Fig. 10 shows the evolution (on logarithmic scale) of the error function during the training of the neural network for the approximation of the Young's modulus function. A continuous and regular descent that progressively slows until an early stop condition is attained. Although they cannot be observed in the chart, oscillations of some importance occur near the end of the training process, which indicates that the network is not able to continue learning from the available data. Three steps (that can barely be observed due to the chart scale) near iterations 300, 2600 and 4400 are related to significant approximation improvements.

Subsequently, the network is asked to make predictions about the data contained in the testing subset (390 registers). Although we know the real value of the Young's modulus for these samples, this information is not communicated to the network since it will be used to calculate several statistical metrics that can be used to measure the network's performance.

Table 6 contains some statistical metrics that enable us to measure the predictive performance of the neural network. As shown, the average deviation is 1.561% and the error of

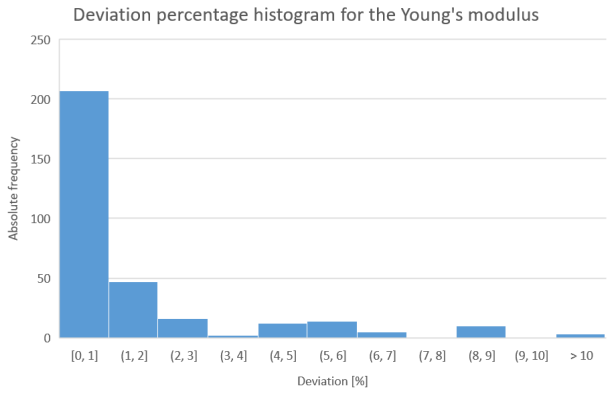


FIGURE 11. Prediction deviation histogram for the Young's modulus.

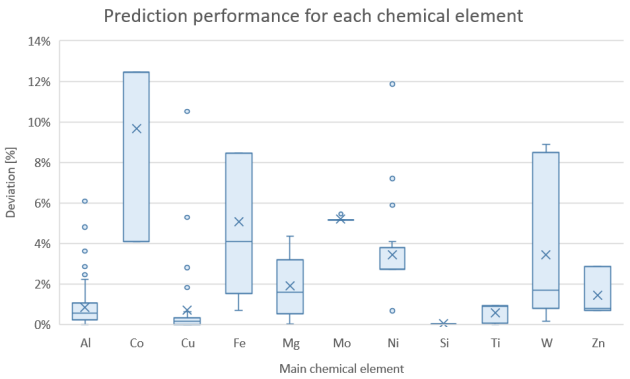


FIGURE 12. Young's modulus prediction performance for each chemical element.

half of the obtained results is less than 0.667%. Note that the average (1.561%), median (0.667%) and trimmed mean (1.394%) are not coincident, which can indicate that some outliers should be carefully investigated.

Fig. 11 shows a histogram in which each sample is related to the range of error in which it is observed. Note the high peak for the range [0, 1] and the long and flat tail. This distribution indicates that the network can make adequate predictions and makes mistakes with some samples.

Fig. 12 shows the performance of the predictions for each of the main chemical elements of the materials in the testing subset. Note that the ANN offers significantly worse results for cobalt and tungsten (note that a few samples are available to use during the training, and therefore, the learning performance is impaired). Fig. 12 also shows some outliers. The big boxes for cobalt, iron and tungsten indicate a large predictive variance.

For each main chemical element, Table 7 shows the number of samples that have been used during the tests (materials in the testing subset) and the average deviation. Significant differences are observed among the groups, although the average error remains bounded. However, for the elements for which a few samples exist, note that the artificial neural network returns worse results. The prediction performance for cobalt and molybdenum alloys is very poor due to the lack of training samples.

TABLE 7. Main element related average error.

Main element	Average error [%]	Samples
Al	0.86%	200
Co	9.67%	3
Cu	0.71%	67
Fe	4.99%	24
Mg	1.92%	10
Mo	5.22%	13
Ni	3.46%	27
Si	0.05%	2
Ti	0.58%	31
W	3.44%	10
Zn	1.46%	3
TOTAL	1.56%	390

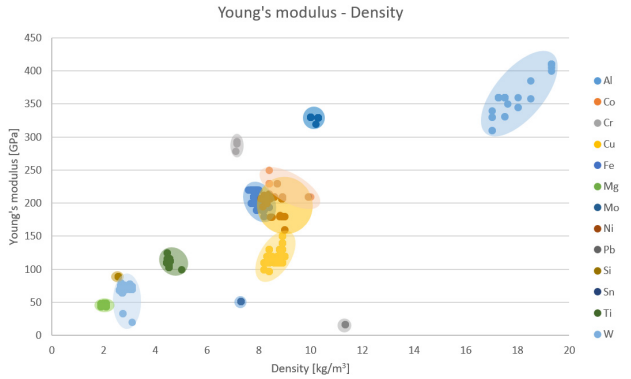


FIGURE 13. Young's modulus vs. density Ashby chart.

C. YOUNG'S MODULUS AND DENSITY ASHBY CHART

An Ashby diagram is a type of scatter plot that enables the relationship between two properties to be established by grouping the points according to a certain criterion [37]. The diagram is a methodological approach to the selection of materials that applies objective principles for the evaluation of the considered properties [38].

Fig. 13 shows an Ashby diagram that relates the Young's moduli of the materials to their respective densities [37]. This Ashby materials selection chart shows clusters of materials that share their main alloy element, which tend to have similar properties.

The diagram has been elaborated using the material data obtained from the predictions made by the artificial neural network. Some elements show substantially greater variability than other elements, which causes a more extensive field. This variability has a negative impact on the performance of the network training since it implies that the function to be learned is more complex.

D. LIMITS OF THE STUDY

The main limitation of this work is the size of the starting data set and the ability of the neural network to learn from this information. As previously indicated, the results of this system improve when the learning process is carried out using a larger input set. However, obtaining large amounts of material data is difficult since the input set it is a very important asset that is not published in an open and easily accessible way. Therefore, a larger initial information corpus can improve the obtained results.

As previously explained, the artificial neural network topology model in this study has some disadvantages (e.g., the results are considerably affected by a reduced initial data set and the local minimums generate substantial attraction) that constitute a limitation of the procedure. Other neural networks architectures can improve the results or reduce the required resources to carry out the training phase.

This study is based on the assumption that the data obtained from the materials library are correct and reliable. Otherwise, the methodology would not change but the obtained results could be affected.

IV. CONCLUSIONS AND FUTURE WORKS

This paper has investigated the feasibility of using artificial neural networks and big data for the prediction of properties of metallic materials whose chemical composition is known. The possibilities of artificial intelligence techniques have been explored based on large data sets. Thus, the major conclusions from this work are presented as follows:

- The artificial neural network technology, as a supervised learning technique, can be employed to exploit large datasets of material information to predict physical and mechanical properties from the chemical composition of different alloys. An ANN can learn to approximate the value of a material property as a function of its chemical constituents.
- An artificial neural network can be trained to predict the density and Young's modulus of a metallic material if its chemical composition is correctly defined. The error of the prediction remains bounded and the average deviation in this work is 0.448% and 1.561%, respectively.
- Supervised-learning methodologies require large training datasets to attain satisfactory predictive performance. The predictive capacity of a neural network improves as the dataset increases as it has more samples to learn from, and therefore, the network is able to learn more precise functions (functions that better approximate the reality of the problem).
- The compiled neural network can adequately approximate the density of a material but requires additional training to attain satisfactory results when it attempts to predict the Young's modulus of an alloy. Note that approximating the density function is theoretically easier with a larger number of samples for training.
- An artificial neural network with multilayer topology can be trained to approximate nonlinear functions related to materials science. Theoretically, a multilayer neural network can learn to approximate any nonlinear function if a sufficient number of samples are provided during the training process and if it has a sufficient number of perceptrons [30].

This work contributes new forms of research for the prediction of other physical or mechanical properties of different types of materials (metals, polymers, ceramics or composites) and enables us to focus future research on materials that are more prominent or for which better results are obtained. In the same way that satisfactory results have been obtained in the prediction of the Young's modulus and the density, the possibility of obtaining useful results for the ultimate stress or the yield stress can be explored.

Artificial neural networks have proven to be a suitable ally for the prediction of physical properties; therefore, they can be employed to describe the elastoplastic behaviour of industrial materials of great relevance without the need to perform expensive and complicated tests of stress-strain. The design of a system that is based on artificial intelligence and capable of completely predicting the stress-strain curve is feasible.

The possibilities of a system whose operation proceeds in the opposite direction to that in this work can also be investigated, that is, we can indicate to the system the properties that we want to achieve, and a system that is based on artificial intelligence can chemically define a material that complies with the indicated requirements.

ACKNOWLEDGMENT

The authors would like to thank the Research Group of the UNED Industrial Production and Manufacturing Engineering (IPME). This work has been developed within the framework of the doctorate program in Industrial Technologies of the UNED.

REFERENCES

[1] J. R. Davis, *Tensile Testing*, 2nd ed. Cleveland, OH, USA: ASM International, 2014.

[2] H. Talebi, M. Silani, S. P. A. Bordas, P. Kerfriden, and T. Rabczuk, "A computational library for multiscale modeling of material failure," *Comput. Mech.*, vol. 53, no. 5, pp. 1047–1071, May 2014.

[3] D. Merayo, Á. Rodríguez-Prieto, and A. M. Camacho, "Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications," *Proc. Inst. Mech. Eng. L, J. Mater., Des. Appl.*, vol. 233, no. 10, pp. 2072–2083, Dec. 2018.

[4] S. K. Pal, S. K. Meher, and A. Skowron, "Data science, big data and granular mining," *Pattern Recognit. Lett.*, vol. 67, pp. 109–112, Dec. 2015.

[5] A. Rodriguez-Prieto, A. M. Camacho, and M. Á. Sebastián, "Materials selection criteria for nuclear power applications: A decision algorithm," *JOM*, vol. 68, no. 2, pp. 496–506, Nov. 2015.

[6] D. Merayo, A. Rodriguez-Prieto, and A. M. Camacho, "Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0," *Procedia Manuf.*, to be published.

[7] T. Tambouratzis, D. Karalekas, and N. Moustakas, "A methodological study for optimizing material selection in sustainable product design," *J. Ind. Ecol.*, vol. 18, no. 4, pp. 508–516, Aug. 2014.

[8] I. Yaqoob, I. Hashem, A. Gani, S. Mokhtar, E. Ahmed, N. Anuar, and A. Vasilakos, "Big data: From beginning to future," *Int. J. Inf. Manage.*, vol. 36, no. 6, pp. 1231–1247, Dec. 2016.

[9] C. Dobre and F. Xhafa, "Intelligent services for big data science," *Future Gener. Comput. Syst.*, vol. 37, pp. 267–281, Jul. 2014.

[10] D. Laney, "3D data management: Controlling data volume, velocity, and variety," META Group Res. Note 6, 2001.

[11] S. Guo, J. Yu, X. Liu, C. Wang, and Q. Jiang, "A predicting model for properties of steel using the industrial big data based on machine learning," *Comput. Mater. Sci.*, vol. 160, pp. 95–104, Apr. 2019.

[12] P. V. Balachandran,"Machine learning guided design of functional materials with targeted properties," *Comput. Mater. Sci.*, vol. 164, pp. 82–90, Jun. 2019.

[13] M. Danylenko, "Aluminium alloys in aerospace," *Aluminium Int. Today*, vol. 31, no. 4, p. 35, 2018.

[14] J. Weinbub, M. Wastl, K. Rupp, F. Rudolf, and S. Selberherr, "ViennaMaterials—A dedicated material library for computational science and engineering," *Appl. Math. Comput.*, vol. 267, pp. 282–293, Sep. 2015.

[15] S. Helal, "The expanding frontier of artificial intelligence," *Computer*, vol. 51, no. 9, pp. 14–17, Sep. 2018.

[16] D. M. Dimiduk, E. A. Holm, and S. R. Niezgoda, "Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering," *Integr. Mater. Manuf. Innov.*, vol. 7, no. 3, pp. 157–172, Sep. 2018.

[17] J. McCarthy, M. Minsky, N. Rochester, and C. Shannon, "A proposal for the dartmouth summer research project on artificial intelligence: August 31, 1955," *AI Mag.*, vol. 27, no. 4, pp. 12–14, Aug. 1955.

[18] C.-C. Zhou, G.-F. Yin, and X.-B. Hu, "Multi-objective optimization of material selection for sustainable products: Artificial neural networks and genetic algorithm approach," *Mater. Des.*, vol. 30, no. 4, pp. 1209–1215, Apr. 2009.

[19] G. Huang, G.-B. Huang, S. Song, and K. You, "Trends in extreme learning machines: A review," *Neural Netw.*, vol. 61, pp. 32–48, Jan. 2015.

[20] F. Rosenblatt, *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. New York, NY, USA: Spartan, 1962.

[21] I. J. Betere, H. Kinjo, K. Nakazono, and N. Oshiro, "Investigation of multilayer neural network performance evolved by genetic algorithms," *Artif Life Robot.*, vol. 24, no. 2, pp. 183–188, Jun. 2019.

[22] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[23] F. Van Veel. (Mar. 31, 2017). Neural Network Zoo Prequel: Cells and Layers. Isaac Asimov Institut. [Online]. Available: https://www.asimovinstitute.org/neural-network-zoo-prequel-cells-layers/

[24] P. Marks and M. Robins, "Who's flying this thing? Web browser control brings hijacking threats to spacecraft," *New Scientist*, vol. 167, no. 2246, p. 5, 2000.

[25] *Matmatch GmbH (N.D.) Matmatch*. Accessed: Mar. 2019. [Online]. Available: https://matmatch.com/

[26] T. Thankachan, K. S. Prakash, C. David Pleass, D. Rammasamy, B. Prabakaran, and S. Jothi, "Artificial neural network to predict the degraded mechanical properties of metallic materials due to the presence of hydrogen," *Int. J. Hydrogen Energy*, vol. 42, no. 47, pp. 28612–28621, Nov. 2017.

[27] M. Bauccio, *ASM Metals Reference Book*, 3rd ed. Cleveland, OH, USA: ASM International, 1993.

[28] A. Deshpande, M. Kumar, *Artificial Intelligence for Big Data: Complete Guide to Automating Big Data Solutions Using Artificial Intelligence Techniques*. Birmingham, U.K.: Packt Publishing, 2018.

[29] M. B. Takahashi, J. C. Rocha, and E. G. F. Nuñez, "Optimization of artificial neural network by genetic algorithm for describing viral production from uniform design data," *Process Biochem.*, vol. 51, no. 3, pp. 422–430, Mar. 2016.

[30] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Netw.*, vol. 4, no. 2, pp. 251–257, 1991.

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, Sacramento, CA, USA, Dec. 2014.

[32] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*.

[33] M. Lutz, *Programming Python*. Sebastopol, CA, USA: O'Reilly Media, 2010.

[34] P. Joshi, *Artificial intelligence With Python: Build Real-World Artificial Intelligence Applications With Python to Intelligently Interact With the World Around You*. Birmingham, U.K.: Packt Publishing, 2017.

[35] K. Fukushima, "Training multi-layered neural network neocognitron," *Neural Netw.*, vol. 40, pp. 18–31, Apr. 2013.

[36] T. Takase, S. Oyama, and M. Kurihara, "Effective neural network training with adaptive learning rate based on training loss," *Neural Netw.*, vol. 101, pp. 68–78, May 2018.

[37] M. Ashby, and K. Johnson, *Materials and Design: The Art and Science of Material Selection in Product Design*, 3rd ed. Oxford, U.K.: Butterworth-Heinemann, 2014.
[38] A. Rodriguez-Prieto, A. M. Camacho, A. M. Aragon, M. A. Sebastian, and A. Yanguas-Gil, "Polymers selection for harsh environments to be processed using additive manufacturing techniques," *IEEE Access*, vol. 6, pp. 29899–29911, 2018.
[39] H. Castro and M. Sweet, "Radiation exposure effects on the performance of an electrically trainable artificial neural network (ETANN)," *IEEE Trans. Nucl. Sci.*, vol. 40, no. 6, pp. 1575–1583, Dec. 1993.
[40] N. N. Qaddoumi, A. H. El-Hag, and Y. Saker, "Outdoor insulators testing using artificial neural network-based near-field microwave technique," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 2, pp. 260–266, Feb. 2014.
[41] H. Zhang, C. Cao, L. Xu, and T. A. Gulliver, "A UAV detection algorithm based on an artificial neural network," *IEEE Access*, vol. 6, pp. 24720–24728, 2018.
[42] J. J. Wang, S. G. Hu, X. T. Zhan, Q. Luo, Q. Yu, Z. Liu, T. P. Chen, Y. Yin, S. Hosaka, and Y. Liu, "Predicting house price with a memristor-based artificial neural network," *IEEE Access*, vol. 6, pp. 16523–16528, 2018.
[43] L. Ma, X. Huo, X. Zhao, B. Niu, and G. Zong, "Adaptive neural control for switched nonlinear systems with unknown backlash-like hysteresis and output dead-zone," *Neurocomputing*, vol. 357, pp. 203–214, Sep. 2019.
[44] Y. Wang, H. Shen, and D. Duan, "On stabilization of quantized sampled-data neural-network-based control systems," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3124–3135, Oct. 2017.
[45] Y. Wang, W. Zhou, J. Luo, H. Yan, H. Pu, and Y. Peng, "Reliable intelligent path following control for a robotic airship against sensor faults," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 6, pp. 2572–2582, Dec. 2019, doi: 10.1109/tmech.2019.2929224.
[46] H. Wang, P. X. Liu, X. Zhao, and X. Liu, "Adaptive fuzzy finite-time control of nonlinear systems with actuator faults," *IEEE Trans. Cybern.*, to be published, doi: 10.1109/tcyb.2019.2902868.
[47] L. Ma, G. Zong, X. Zhao, and X. Huo, "Observed-based adaptive finite-time tracking control for a class of nonstrict-feedback nonlinear systems with input saturation," *J. Franklin Inst.*, to be published, doi: 10.1016/j.jfranklin.2019.07.021.
[48] H. Wang, P. X. Liu, J. Bao, X. Xie, and S. Li, "Adaptive neural output-feedback decentralized control for large-scale nonlinear systems with stochastic disturbances," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: 10.1109/TNNLS.2019.2912082.
[49] E. Real, I. Arrayago, E. Mirambell, and R. Westeel, "Comparative study of analytical expressions for the modelling of stainless steel behaviour," *Thin-Walled Struct.*, vol. 83, pp. 2–11, Oct. 2014.

D. MERAYO received the B.S. degree in computer science from the University of Leon, Spain, in 2007, the B.S. degree in aeronautical engineering from the Polytechnic University of Valencia, Spain, in 2012, and the M.S. degree in investigation in industrial technologies from National Distance Education University, Madrid, Spain, in 2016, where he is currently pursuing the Ph.D. degree in mechanical engineering. From 2009 to 2017, he was a Structural Engineer with Airbus Spain. Since 2017, he has been a Project Engineer with Renault Seville, Spain. He is the author of several research and technical articles. His research interests include material science and technology, artificial intelligence, and computer science.

A. RODRÍGUEZ-PRIETO received the M.Sc. degree in materials engineering from the Complutense University of Madrid (UCM), in 2007, and the M.Sc. and Ph.D. degrees in advanced manufacturing engineering from the National Distance Education University (UNED), Spain, in 2011 and 2014, respectively. He is currently an Assistant Professor with the Department of Manufacturing Engineering, UNED. He is also a Senior Engineer and Project Manager of SGS Tecnos. His major current research interests include innovation in materials selection methodologies and the analysis of advanced manufacturing processes for demanding applications.

A. M. CAMACHO received the M.Sc. degree in industrial engineering from the University of Castilla–La Mancha (UCLM), in 2001, and the Ph.D. degree in industrial engineering from National Distance Education University (UNED), Spain, in 2005. She is currently an Associate Professor with the Department of Manufacturing Engineering, UNED. Her main research interests include innovation in manufacturing engineering and materials technology, with a focus on the analysis of metal forming and additive manufacturing techniques using computer aided engineering tools and experimental testing, and the development of methodologies for materials selection in demanding applications.

4.4 *Prediction of the Bilinear Stress-Strain Curve of Aluminum Alloys Using Artificial Intelligence and Big Data* [3]

Los indicios de calidad de este artículo pueden encontrarse en el Apéndice D.

4.4.1 Datos de la publicación y factor de impacto

Tabla 4. Factor de impacto de Prediction of the bilinear stress-strain curve of aluminum alloys using artificial intelligence and Big Data

Título	Prediction of the bilinear stress-strain curve of aluminum alloys using artificial intelligence and Big Data
Autores	David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Revista	Metals
ISSN	2075-4701
Editorial	MDPI AG
País	Suiza
Volumen	10 (7-904)
Páginas	1-29
Fecha	2020
doi	10.3390/met10070904
Factor de impacto	2.117 (2019 Journal Citation Reports) 18/79 (Metallurgy & Metallurgical Engineering, Q1)

4.4.2 Resumen y copia de la publicación

Las aleaciones de aluminio se encuentran entre los materiales más utilizados en industrias exigentes como la aeroespacial, la automoción o el envasado de alimentos y, por tanto, es fundamental predecir el comportamiento y las propiedades de cada componente. Se pueden utilizar herramientas basadas en inteligencia artificial para afrontar este complejo problema. En este trabajo, se desarrolla una herramienta asistida por ordenador para predecir las propiedades mecánicas de las aleaciones de aluminio: módulo de Young, límite elástico, resistencia máxima a la tracción y elongación a rotura. Estas predicciones se basan en la composición química y los tratamientos de la aleación y se emplean para estimar la aproximación bilineal de la curva tensión-deformación, muy útil como herramienta de decisión que ayuda en la selección de materiales. El sistema se basa en el uso de redes neuronales artificiales respaldadas por una gran recopilación de datos sobre las características tecnológicas de miles de materiales comerciales. Por tanto, el volumen de datos supera las 5000 entradas. Una vez recuperados, filtrados y organizados los datos relevantes, se define una red neuronal artificial y, tras el entrenamiento, el sistema es capaz de realizar predicciones sobre las propiedades del material con una confianza media superior al 95%. Finalmente, la red ya entrenada se emplea para mostrar cómo se puede utilizar para respaldar decisiones sobre aplicaciones de ingeniería.

Aluminum alloys are among the most widely used materials in demanding industries such as aerospace, automotive or food packaging and, therefore, it is essential to predict the behavior and properties of each component. Tools based on artificial intelligence can be used to face this complex problem. In this work, a computer-aided tool is developed to predict relevant mechanical properties of aluminum alloys—Young's modulus, yield stress, ultimate tensile strength and elongation at break. These predictions are based on the alloy chemical composition and tempers, and are employed to estimate the bilinear approximation of the stress-strain curve, very useful as a decision tool that helps in the selection of materials. The system is based on the use of artificial neural networks supported by a big data collection about technological characteristics of thousands of commercial materials. Thus, the volume of data exceeds 5k entries. Once the relevant data have been retrieved,

filtered and organized, an artificial neural network is defined and, after the training, the system is able to make predictions about the material properties with an average confidence greater than 95%. Finally, the trained network is employed to show how it can be used to support decisions about engineering applications.

metals

MDPI

Article

Prediction of the Bilinear Stress-Strain Curve of Aluminum Alloys Using Artificial Intelligence and Big Data

David Merayo Fernández *, Alvaro Rodríguez-Prieto and Ana María Camacho

Department of Manufacturing Engineering, Universidad Nacional de Educación a Distancia (UNED), Juan del Rosal 12, 28040 Madrid, Spain; alvaro.rodriguez@ind.uned.es (A.R.-P.); amcamacho@ind.uned.es (A.M.C.)

* Correspondence: dmerayo1@alumno.uned.es

Received: 9 June 2020; Accepted: 2 July 2020 ; Published: 6 July 2020

check for updates

Abstract: Aluminum alloys are among the most widely used materials in demanding industries such as aerospace, automotive or food packaging and, therefore, it is essential to predict the behavior and properties of each component. Tools based on artificial intelligence can be used to face this complex problem. In this work, a computer-aided tool is developed to predict relevant mechanical properties of aluminum alloys—Young's modulus, yield stress, ultimate tensile strength and elongation at break. These predictions are based on the alloy chemical composition and tempers, and are employed to estimate the bilinear approximation of the stress-strain curve, very useful as a decision tool that helps in the selection of materials. The system is based on the use of artificial neural networks supported by a big data collection about technological characteristics of thousands of commercial materials. Thus, the volume of data exceeds 5*k* entries. Once the relevant data have been retrieved, filtered and organized, an artificial neural network is defined and, after the training, the system is able to make predictions about the material properties with an average confidence greater than 95%. Finally, the trained network is employed to show how it can be used to support decisions about engineering applications.

Keywords: aluminum alloy; artificial intelligence; multi-layer artificial neural network; big data; python; stress-strain curve; material selection; decision support system; material characterization

1. Introduction

Aluminum alloys are some of the most relevant metallic materials of the industry and they play a very important role in some high-technology fields such as aerospace and in everyday industries such as food packaging [1], among other reasons, due to its high strength-to-weight ratio. Aluminum production and consumption has grown by approximately 50% in the last decade and this rate is estimated to accelerate over the next few years [2,3]. In addition, it is expected to play a fundamental role at the ecological and environmental level because it is a relatively easy material to recycle [3].

Besides, aluminum alloys are the most frequent type of non-ferrous material employed for an extensive range of applications, namely in the automotive, aerospace, and structural industries, among others [4]. Widespread use of these alloys in the modern world is due to an exceptional blend of material properties, combining low density, excellent strength, corrosion resistance, toughness, electrical and thermal conductivity, recyclability and manufacturability. Another key factor is the relatively low cost of aluminum machining, making its alloys very attractive for applications in different sectors [4].

Aluminum was only discovered in the early 19th century, however, despite its short history, it has become an essential material. Every day, new uses for aluminum alloys are emerging in various

industrial sectors due to its excellent properties [5] and the fact that the price of the raw material has been decreasing since then [2]. Therefore, it is necessary to provide material scientists with tools that can be used to develop new alloys with properties optimized for each need. The mechanical properties of a material play an important role in the performance of industrial components. A correct in-service behavior depends largely on the characteristics of the materials that constitute it, as inadequate material properties can cause premature failure [6,7]. Therefore, the decision to choose a specific material to manufacture an industrial component greatly affects its ability to carry out the work for which it was designed [8–10].

There are thousands of aluminum alloys although only a few of them are commonly used in the industry [11], in some cases due to the difficulty of finding new solutions and, in others, because they are specific materials with optimized characteristics for the mission they fulfill.

Knowing the properties of the materials employed in industrial designs is crucial; however, obtaining these data often involves accessing large amounts of resources, which are commonly not available. Multitude of tests are needed to obtain significant information, which entails that enough time, personnel and facilities must be available at a given cost [6]. The process of characterizing a material may involve a bulk of tests that requires a substantial amount of time and the investment of vast quantities of resources [12].

Despite the fact that there are multiple decision support systems and materials selection methodologies [13] applied to materials science, there are few references which mention the use of artificial intelligence-based technologies in the field of metal processing and engineering [6,14–18]. Although there are many studies that use machine learning to investigate the microstructure of metals and their properties [19–21], there are hardly any references with an industrial approach that take into account the tempers of aluminum alloys [22].

Nevertheless, it is possible to find a greater number of references that develop techniques based on artificial intelligence applied to other industrial materials, mainly steel [23]. These studies take advantage of the ability of these tools to obtain predictions about the behavior or properties of a certain material or industrial component [24–28].

In this work, a decision support system is developed which is capable of predicting some of the most important properties that define the stress-strain curve of aluminum alloys whose chemical composition and treatments (thermal and mechanical) are known. This system is capable of predicting the Young's modulus (E), the yield stress (YS), the ultimate tensile strength (UTS) and the elongation at break (A). These four properties define the elastic and plastic behavior of a material under tension [29].

The difficulty of developing this study lies in the large number of steps and disciplines involved in carrying it out: an extensive software has been developed in Python 3.7 [30] capable of working without user intervention to download data from an on-line material library [31], filter and organize data, define and train an artificial neural network [32], and, finally, make predictions using that network. On the other hand, a great deal of work has been required to analyze data and define criteria based on materials science [33]. Developing the software to obtain and download the data to carry out this study has been one of the most delicate and time consuming steps.

1.1. Designation and Main Characteristics of Aluminum Alloys

Aluminum alloys are light materials with a high strength-to-weight ratio combined with excellent thermal conductivity and good corrosion resistance [5]. Aluminum has a density of about 2700 kg/m^3, approximately one-third as much as steel (7830 kg/m^3) [34]. Such lightweight, along with the high strength of some aluminum alloys (higher than some structural steels), allow designing and manufacturing of strong, lightweight structures that are particularly beneficial for vehicles [1,35] and for the environment.

Aluminum alloys are able to withstand the progressive oxidization that causes steel to rust away. The bare surface of aluminum reacts with oxygen to procedure an inert aluminum oxide film, that blocks further oxidation [35]. In addition, unlike iron rust, the aluminum oxide film does not

flake off to expose a fresh surface that could be further oxidized. If that protective layer is scratched, it will immediately reseal itself. The thin oxide layer sticks tightly to the metal and is colorless and transparent [36,37].

Aluminum alloys and their tempers comprise a wide and adaptable assortment of manufacturing materials. For optimum product design and effective development, it is important to understand the differences between the available alloys, their performance and characteristics [34].

Aluminum is an example of a ductile material because it can withstand significant plastic deformation so they are very used in metal forming operations; such materials can be compressed to form thin plates and sheets or pulled to form wires [11]. Typical ductile materials show a stress-strain curve that is very steep at the beginning (elastic zone, where the stress-strain curve is almost a straight line) and, after the yield point, the curve slope decreases (plastic zone). At one point, the slope of the curve becomes zero at the ultimate tensile strength. The strain difference between the yield point and the ultimate point is relatively large for aluminum [38], due to their excellent ductility. Ductile materials have generally high toughness and are able to absorb a large amount of energy before breaking [12].

Appendix A contains a brief introduction to the nomenclature and standardization of aluminum alloys.

1.2. Modelization of Stress-Strain Curve

The stress-strain curve shows, in a simple way, the deformation of a material when it is subjected to mechanical load. In this diagram, the stress is plotted on y-axis and its corresponding strain on the x-axis [39]. Tension tests provide information on the strength and ductility of materials under uniaxial tensile stresses. This information may be useful in comparisons of materials, alloy development, quality control, numerical simulation such as finite element modeling, and design under certain circumstances [40].

The stress-strain curve is a crucial material asset and there are several standard testing methods to measure this curve, such as the tensile test [40], the compression test and the torsion test [38]. Although several studies have reported extension of the strain range [41], achieving a large strain with those methods sometimes can be difficult because the specimen tends to break at relative small strain points.

The simplest loading to visualize is a one-dimensional tensile test, in which a slender test specimen is stretched along its axis [42]. The stress-strain curve is a representation of the deformation of the specimen as the applied load is increased monotonically, usually to fracture [39]. Stress-strain curves are usually presented as:

- Engineering stress-strain curves, in which the initial dimensions of the specimens are used in the calculations.
- True stress-strain curves, where the instantaneous dimensions of the specimen at each point during the test are used in the calculations. The true curves are always above the engineering curves, notably in the higher strain portion of the curves [40].

A stress-strain curve combines a lot of information about the material and its behavior [43]. In this work, four of its properties will be studied:

- Young's modulus (E)—it is a mechanical property that measures the stiffness of a material and characterizes its behavior in the elastic zone according to the Hooke's law. It defines the ratio between uniaxial applied force and deformation of a material in the linear elastic regime (see Equation (1)) [44]:

$$E = \frac{\sigma}{\varepsilon}, \tag{1}$$

where E is the Young's modulus, σ is the stress and ε is the strain.

- Yield strength (YS): it is a property of the material that indicates the point at which the material begins to deform plastically. Stresses lower than the YS do not produce permanent deformations, whereas higher ones produce deformations that will remain even when the applied forces are eliminated [45].
- Ultimate tensile strength (UTS): the maximum stress that the material can withstand without area reduction [43].
- Elongation at break (A): the maximum strain that the material can withstand before failure [43].

These four properties completely define the bilinear approximation of the stress-strain curve of a material and allow summarizing the elasto-plastic stress behavior of a material to four values.

The stress-strain curve also indicates the amount of energy a material can store before fracture since the area enclosed below the curve is the energy that the material absorbs during its deformation [43,46]. The energy that a material absorbs is called resilience if the deformation is elastic and toughness if the deformation is plastic. This energy can be calculated using the Equation (2):

$$U = U_r + T = \int_0^A \sigma \cdot d\varepsilon, \tag{2}$$

where U is the total deformation energy (absorbed energy), U_r is the resilience, T is the toughness, A is the elongation at break, σ is the stress and ε is the strain.

Since the transition from elastic to plastic behavior is continuous, for aluminum alloys (and for many other materials), there is no singular point that delimits them [47]. Therefore, standardization organizations have selected a criterion that guarantees the reproducibility of the tests—the yield point is defined as that in which there is a deviation of 0.2% of strain with respect to the elastic linear behavior [48,49].

Figure 1 shows the true stress-strain curves of some relevant aluminum alloys. It is easy to distinguish the elastic regime (linear and very steep) and the plastic regime, where the curve slope decreases and becomes flatter. Thus, there is an obvious rapid change near the yield point.

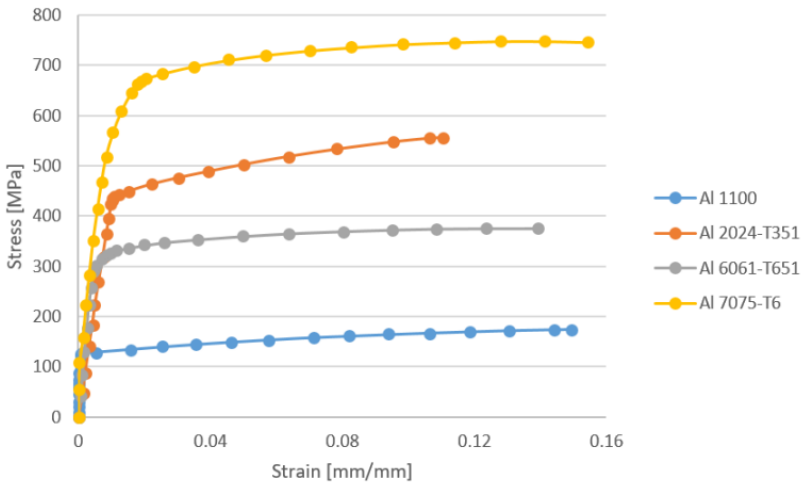


Figure 1. Example of true stress-strain curves of some aluminum alloys (data from Reference [50]).

To carry out some industrial design tasks, it is very common to use analytical models that allow the real curve of a material to be approximated using mathematical functions [51]. The behavior of aluminum alloys can be approximated very well by the expression of the Ramberg-Osgood stress-strain law [52] or by a bilinear stress-strain diagram, which is an accurate approximation away from the yield point [46,51,53–55].

The Ramberg-Osgood expression represents the elastoplastic behavior of the material throughout all its admissible strain values (see Equation (3)) [52,56].

$$\varepsilon = \frac{\varepsilon}{E} + \alpha \frac{\varepsilon}{E} \left(\frac{\sigma}{\sigma_{YS}} \right)^{n-1}, \tag{3}$$

where ε is the strain, σ is the applied stress, E is the Young's modulus, σ_{YS} is the yield strength, and α and n are two parameters that depend on the material.

On the other hand, the bilinear approximation of the stress-strain curve consists of two lines that represent, respectively, the linear behavior (whose slope is the Young's modulus, E) and the plastic behavior (whose slope is the strain hardening modulus, E_T) [43,56]. These two lines intersect at the yield point (see Equations (4) and (5)).

$$\varepsilon = \frac{\sigma}{E}, \sigma \leq \sigma_{YS} \tag{4}$$

$$\varepsilon = \frac{\sigma}{E_T}, \sigma > \sigma_{YS}, \tag{5}$$

where ε is the strain, σ is the applied stress, E is the Young's modulus, σ_{YS} is the yield strength and E_T is strain hardening modulus (the slope of the line that defines the plastic behavior).

Figure 2 shows a comparison between the actual stress-strain curve of a generic aluminum alloy and its bilinear approximation [55]. As can be seen, the fit of the simplified model is good away from the yield point, where the discrepancies are significant [56].

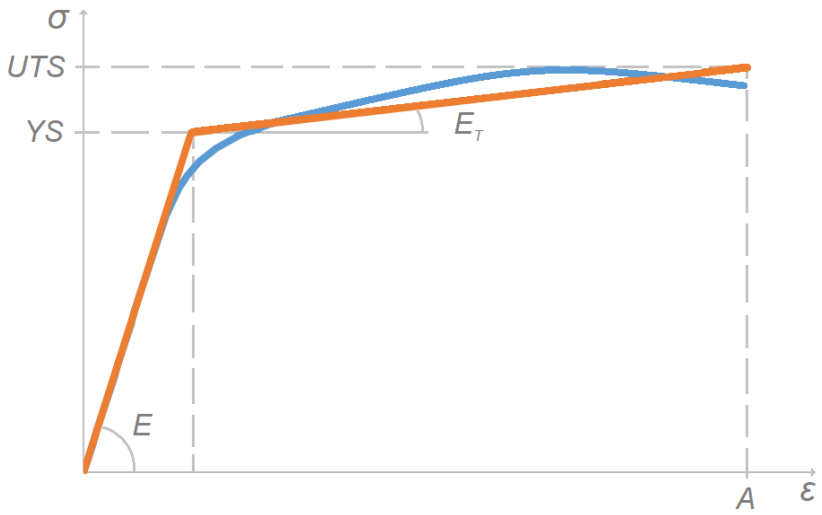


Figure 2. Actual stress-strain curve and bilineal approximation for an aluminum alloy, data from [55].

The shape of the stress-strain curve (real and approximate) and its values depend on [39]:

- Alloy chemical composition.
- Heat treatment and conditioning.
- Prior history of plastic deformation.
- The strain rate of the test.
- Temperature.
- Orientation of applied stress relative to the structure of the test specimens.
- Size and shape.

The latter four parameters are described in the pertinent standards, including the case of aluminum testing specimens [40,57]. The former three parameters are the ones that are considered in this study.

1.3. Sources of Big Data

Materials science depends on experiments and simulation-based models to understand the physics of materials in order to better know their characteristics and discover new materials with enhanced properties [58]. All these experiments and simulations generate a huge amount of data, which is becoming increasingly difficult to handle using traditional data processing techniques [6]. Due to the massive volume of data being produced at unprecedented speed, these data are not effectively processed to create information, delaying the production of new knowledge [59].

Traditionally, knowledge has been organized through the so-called "knowledge pyramid" or "information hierarchy". This model is made up of four steps, each of which derives from the previous one—data, information, knowledge and wisdom (DIKW) [60]. In this way, the processed data constitutes information, which is organized to generate knowledge, which, finally, is summarized as wisdom [61].

Our current technology has reached a level never seen before in terms of generating data [58]; however, the techniques aimed at their processing are not yet as advanced and their use is not widespread [6]. Therefore, our society faces challenging problems to transform data into information and knowledge. Extracting value from raw data requires a systematic and well-defined approach to solve these emerging real-world problems and so, a new multidisciplinary approach is needed [59].

In any field, datasets are considered "big" when they are large, complex, and difficult to process and analyze. Materials science data tend to be particularly heterogeneous in terms of type and source. One of the first steps in processing large datasets is data reduction [62]. Experiments on the Large Hadron Collider, for example, retain only a small fraction of 1% of the data they produce because it becomes impractical with the current technologies to store and analyze more than the hundreds of megabytes per second that are considered more valuable: it is up to sophisticated software to determine which data are more relevant [63].

Although the term "big data" is relatively new, the action of collecting and storing large amounts of information for further analysis has been performed for many years. The current definition of big data is based on the three Vs [6,64]:

- Volume: large volumes of unstructured low-density data are processed. The data can be of unknown value, such as machining conditions, material properties or manufacturing control measures.
- Velocity: the rate at which the data are received, and possibly, to which some action is applied.
- Variety: conventional data types are structured and can be clearly organized in a relational database; nevertheless, big data is presented as unstructured sparse registers.

Matmatch® Munich, Germany [31] is a well-known open-access materials library that contains information about thousands of different commercial and standard materials. Registered users can freely access the information stored in the databases. A description sheet, which contains all available data, can be downloaded for each material [6].

Matmatch® [31] offers widely sparse and heterogeneous data about more than 70,000 materials [65]. These data is provided by the manufacturers and suppliers of the materials. Although the data is believed to be accurate, it must be processed, filtered and parsed to generate a corpus of useful and meaningful information [61].

1.4. Artificial Intelligence and Artificial Neural Networks

Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems [66]. These processes comprise self-correction (spotting errors and solving them), reasoning (using rules to reach new conclusions and knowledge) and learning (acquiring procedures to employ the information) [6,67]. The term "artificial intelligence" was created in 1956 during the Dartmouth Conference, where the discipline arose [68]. At present, AI is a wide-ranging

term that has lately gained importance due to the increase in speed, size and variety of the data collected by companies [67]. AI can perform tasks, such as recognizing patterns in data, more proficiently than humans do, which enables users to extract more information from their datasets [14].

AI is a term that encompasses a multitude of techniques and technologies aimed at endowing a machine with the ability to exhibit "intelligent behavior" [32]. Within these techniques, we can find simple (although powerful) mathematical models such as decision trees, capable of categorizing data [66]; and other much more complex and advanced technologies such as deep convolutional neural networks, able to identify images and patterns [69].

AI has shown that it can be applied to a multitude of disciplines not directly related to computing or robotics. Among the most relevant new uses, it can be highlighted medicine [70], warfare [71], ecology [72], security [73], education [74], oil exploration [75] or material science [76]. AI can be applied to almost all branches of science and engineering, and new uses and applications emerge every day [16,32].

Among all the tools included within the artificial intelligence field, multi-layer artificial neural networks (ANN) can be highlighted due to their current relevance and proven capabilities [77]. A multi-layer network is a supervised learning algorithm able to learn a non-linear function by training on a labelled dataset that can be used to perform classifications and regressions [78]. Multi-layer neural networks are made up of perceptrons that organize themselves forming layers (groups of neurons) that communicate with each other (in general, perceptrons do not communicate with their own layer companions) [6,17].

Bearing in mind the connection topology of the perceptrons, three types of layers can be defined: input layer, which includes all perceptrons that receive data from an external source; output layer, which includes all perceptrons that return results; and hidden layer, which includes all other perceptrons, which does not communicate with the network exterior [79].

Appendix B contains a mathematical explanation of the fundamentals of neural network technology.

2. Methodology

This work is focused on obtaining an artificial neural network capable of making acceptable predictions of the main parameters that define the stress-strain curve of aluminum alloys while maintaining a limited average error. Subsequently, the output data, the data about the network training process and the data about the prediction step are conveniently analyzed.

Figure 3 schematically shows an overview of the methodology of this work. It consists of two main phases: phase of dataset creation and phase of prediction and analysis. Each of these phases is made up of several stages that are based on the results of the previous one. This work scheme has already demonstrated its ability to obtain adequate results predicting material properties [6].

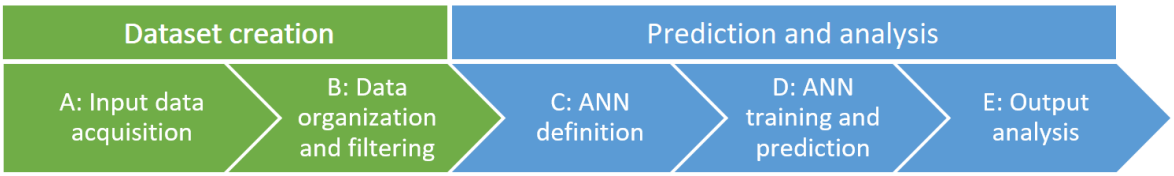


Figure 3. Methodology scheme.

2.1. Stage A—Input Data Acquisition

As already indicated, the input dataset used in this work has been obtained from an online open access material library (Matmatch® [31]). In this web portal, it is possible to access the information provided by thousands of suppliers of materials of different kinds, including aluminum alloys [6].

For each material, the registered data can be very diverse and, in any case, it should be noted that these data are not at all exhaustive, but quite sparse: not all information is available for all materials

since the task of recording the data of each material depends on the marketers themselves. In the field of big data, it is very common to deal with sparse, heterogeneous and disperse information [64].

This material library offers information about more than 70,000 different materials [31]; including several thousand aluminum alloys registries. It is possible to access a specific datasheet for each material and download it; however, it is not possible to obtain a complete package with the information of multiple materials; instead, it is necessary to download the data of each material one by one [6].

To carry out the task of downloading the raw data of the relevant materials, a Python application has been developed which is capable of sequentially downloading the datasheets [30,80]. In this way, the raw data about 5341 aluminum alloys have been obtained. This bulk of registries contains data about 351 material properties, including chemical composition and mechanical, physical, electrical or acoustic properties. Each record is downloaded as an Excel document which contains all available information about the material; the datasheet format is not uniform neither the data is shown homogeneously [6].

2.2. Stage B—Data Organization and Filtering

Once the datasheets of all the materials have been downloaded, the information contained in all these files is sequentially read and interpreted. As already indicated, much more data is available than is necessary to carry out this study [31]. The following considerations have been taken when filtering and organizing the available data:

- The average value is taken for those properties that are registered as ranges in the datasheets. Some properties (especially chemical properties and some mechanical ones) are shown by specifying the maximum and minimum values because the standards and norms are written in this way [81,82].
- Only materials whose chemical composition is defined at more than 95% are considered. For some alloys, the chemical composition is not specified or is poorly done [6].
- Only the four properties that define the bilinear approximation of the stress-strain curve are taken into account [39,56]: Young's modulus (E), yield stress (YS), ultimate tensile strength (UTS) and elongation at break (A).
- Only records in which these four properties of the stress-strain curve are specified are considered [6]. Although the methodology is capable of inferring the missing information, it is necessary to know the real data in order to carry out the training or to calculate the precision of the prediction.
- Only eleven chemical elements (the main ones) are taken into account when defining the chemical composition of the alloys [35]: Al, Zn, Cu, Si, Fe, Mn, Mg, Ti Cr, Ni and Zr. All other chemical elements are considered non-relevant and their mass contribution is regrouped as "Other". The presence of the discarded elements in the considered alloys is, in all cases, lower than 0.4% (by mass) [81,82].
- The methodology only considers 35 different treatments: F (as fabricated, single type), O (annealed, single type), H (strain hardening, 19 types of treatment) and T (thermally treated, 14 types of treatments) [81,82]. Despite the fact that there are data about alloys with other treatments, it has been considered that the sample is so scarce that it causes bias [78] in the training process of the neural network and, so, these other treatments and their related registries have been discarded.

Approximately 84% of discarded records are due to not indicating the four properties that define the bilinear stress-strain curve or because they do not specify any treatment. Note that an alloy whose manufacturing process does not involve treatments (therefore, F, as fabricated) is different from a material that does not specify any treatment (lack of data).

After conveniently filtering and organizing the 5341 datasheets, 2101 aluminum alloys records are kept. Only the following data are considered from now on:

- Young's modulus (E), [GPa].
- Yield strength (YS), [MPa].
- Ultimate tensile strength (UTS), [MPa].
- Elongation at break (A), [mm/mm].
- Chemical composition (11 elements are considered), [% mass].
- Temper (35 treatments are considered).

2.3. Stage C—Artificial Neural Network Definition

Once the data has been filtered and has been guaranteed to be relevant, the artificial neural network that will be in charge of carrying out the predictions is defined: a multilayer feedforward architecture and a fully connected topology have been chosen [78]. This structure consists on one input layer, 3 hidden layers (which contain 100, 100 and 10 perceptrons respectively) and one output layer.

The multilayer feedforward architecture provides neural networks with the potential of being universal approximators [83]. Even though a fully connected ANN can represent any function, it may not be able to learn some functions because backpropagation convergence is not guaranteed [78].

This topology is the result of successive optimization steps to balance its learning capacity and the necessary resources for its training [84]. Note that a complex topology is capable of learning more complex functions than a simple topology but requires additional resources during its training: additional time, calculation capacity and input data [6]. A balance between the network depth and the network width was obtained.

2.4. Stage D—Artificial Neural Network Training and Prediction

Once the input data are already available and the neural network topology is defined, the training and prediction phase begins. During this phase, each of the four properties that define the bilinear approximation of the stress-strain curve is taken into account: Young's modulus, yield strength, ultimate tensile strength and elongation at break.

For each of the four properties, 10 learning and prediction iterations are performed. Each of these iterations (independent from each other) is subdivided into four steps:

- Division of the input dataset: it is randomly divided into two disjoint subsets containing, respectively, 80% (training subset) and 20% (testing subset) of the records. To avoid bias, the same data should not be used to train and to make predictions since overfitting could occur and incorrect metrics (too good results) would be obtained [78].
- Neural network training with the training subset.
- Prediction of the properties of the test subset.
- Data storage for further analysis.

Figure 4 shows an overview of the iterative steps of the training and prediction phase. Repeating each iteration 10 times allows for a clearer view of the network performance metrics since better statistical analyzes can be carried out. The network training is subject to the following conditions:

- Calculation of the learning rate for each parameter using Adaptive Moment Estimation (ADAM) with $\beta_1 = 0.9$, $\beta_2 = 0.999$ (algorithm parameters), $\eta = 0.001$ (step size) and $\epsilon = 10^{-8}$ (stability factor) [85].
- Early stopping after 100 iterations without significant changes to avoid overfitting.
- Training stops when a training error of less than 0.001 is reached as it is considered negligible [6].
- Maximum of 100,000 training epochs to avoid infinite loops (this condition was never reached during this study).
- Sigmoid activation function.

This entire training and prediction process generates a large amount of information that provides very significant evidence about the performance and capabilities of the neural network.

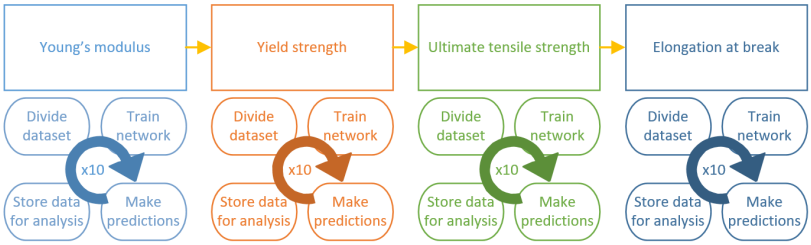


Figure 4. Training and prediction phases overview.

2.5. Stage E—Output Analysis

Once all the training and prediction iterations have already been carried out and all the resulting information is available, the analysis phase is carried out. A complete battery of statistical metrics are calculated and several figures are plotted to summarize both the training and the prediction steps. This information allows the discussion of the results obtained with the methodology described in this paper.

The most remarkable information that can be obtained from the training is the evolution of the error function throughout the learning epochs. Although the number of epochs is not relevant, it is very important to check that the error function converges asymptotically to a relatively low value [78].

On the other hand, the performance of the prediction process is estimated using the absolute relative deviation for each sample of the test subsets. With this information, it is possible to calculate various statistical estimators and metrics that allow knowing the goodness and correctness of the complete methodology. In addition, it is possible to plot figures that represent this information.

Results out of 4.4 sigma interval (98% confidence) are considered abnormal and are marked as outliers.

2.6. Software and Tools

The decision support system has been developed in Python 3.7 (Python Software Foundation: Beaverton, OR, USA) using an object-oriented paradigm [86] and the code architecture consists of more than 25 classes that interact and handle the different phases of the methodology. Multiple standard libraries and modules have been used to simplify the development, promote code reuse and take advantage of the latest technology [87]. Python has been chosen because it is a high-level, cross-platform, multi-paradigm programming language that is very popular among developers [30], especially those who develop artificial intelligence related software [88].

The most relevant external modules that have been used are:

- Selenium (version 3.141, Software Freedom Conservancy: New York, NY, USA): library that enables control of the web browser by using code [89].
- BeautifulSoup (v4.7.1, Free Software Foundation: Boston, MA, USA): library that facilitates working with HTML files and parsing them [90].
- SciPy (v1.2.3, Python Software Foundation: Beaverton, OR, USA): library that contains numerous scientific, mathematical and statistics functions [91].
- NumPy (v1.18.0, Python Software Foundation: Beaverton, OR, USA): library that enables easy management of large amounts of data and large matrices and numbers [91].
- MatPlot (v3.2.2, Python Software Foundation: Beaverton, OR, USA): library that eases the production of plots, figures and graphics [30].
- TensorFlow with Keras (v2.2, Google: Mountain View, CA, USA): high-level library that contains a vast amount of functions and procedures related to artificial intelligence, especially artificial neural networks [43,92].

The complete project includes more than 10,000 lines of code and works, mainly, on command line through batch processing. Only data analysis has really required active user intervention.

3. Results and Discussion

Even if the training algorithms are randomly initialized, the outcomes (during both training and prediction) are very stable and converge to similar results. Once the neural network is appropriately trained with the training subset, it is requested to make predictions. In this second step, the network is not given any clue about the expected results because this is the information that should be returned.

For each of the four properties, the network is trained with 1681 randomly chosen registers and the remaining 420 are employed to test the prediction performance of the network. Note that both subsets (training and testing) are randomly created for each of the 10 iterations; therefore, each iteration is fully independent from the others.

3.1. Young'S Modulus

Figure 5 shows the Young's modulus histogram of the input dataset. It can be seen that the registers are grouped around the range $(69, 71]$. This is an expected behavior since $E = 70$ GPa is the most common value for aluminum alloys. It can also be seen that the range of values is quite small with very few records out of the range $(67, 73]$.

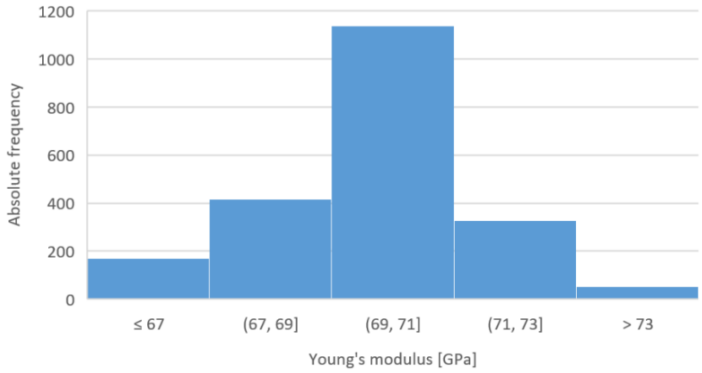


Figure 5. Young's modulus histogram of the input dataset.

Appendix C contains some notions about the neural network training process for predicting this property.

After the training, the neural network is asked to make predictions about the remaining records contained in the testing subset. For these records, the real values of the Young's modulus are known but are not communicated to the ANN as they are retained to calculate some performance metrics afterwards.

The values contained in Table 1 are the relative errors of the prediction of the Young's modulus (calculated using Equation (A8)). It shows several statistical indicators related to the deviation (as percentage) in the Young's modulus prediction: average deviation (Avg. Dev.), statistical standard deviation (Std. Dev.), median and trimmed average deviation at 90% of the interval (Avg. Dev. 90%). The same information can be seen on Figure 6 as a box and whisker plot.

The overall average error is 3.07%, the median is 2.35% and the trimmed average deviation at 90% is 2.87%. These three statistical values are quite close to each other, which means that the results are grouped around the mean value and few abnormal values appear.

Figure 6 shows the combined results of all 10 iterations. It is relevant to highlight the presence of some sparse outliers. These anomalous values are easily identifiable and, in general, are linked to very specific alloys that exhibit unusual properties. Although these outliers reduce the overall performance of the system, they allow knowing the capacity of the methodology in the worst conditions.

Table 1. Average deviation (as %) of the prediction of the Young's modulus.

Avg. Dev.	Std. Dev.	Median	Avg. Dev. 90%
3.07	2.24	2.35	2.87

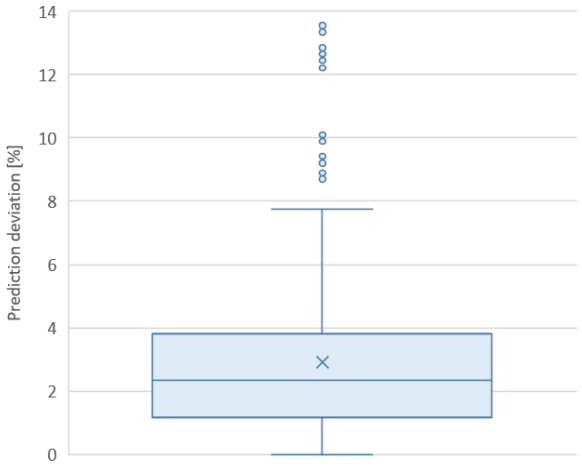


Figure 6. Prediction deviation of the Young's modulus.

Figure 7 shows the histogram of the deviations in the Young's modulus prediction for all iterations (it displays the 4200 predictions that are carried out in the 10 iterations). This plot shows that most of the errors are lower than 4%, however, some high values appear for alloys with unusual properties. The neural network has trouble learning the properties of these alloys because the sample in the input dataset is small and they diverge with respect to the behavior of the other alloys (this issue would be solved with a more complete input dataset).

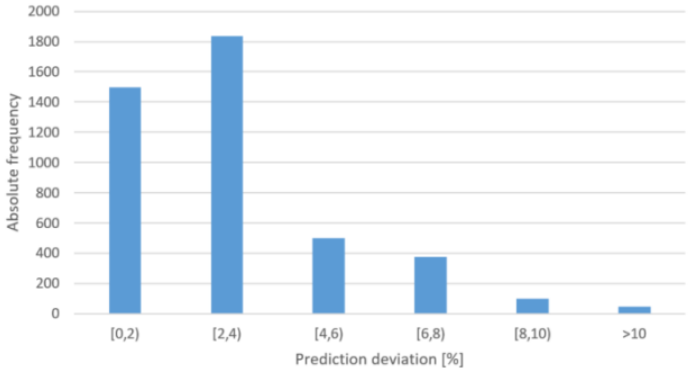


Figure 7. Histogram of the prediction error of the Young's modulus for all iterations.

Since the overall average deviation is 3.07%, it can be said that the system makes very low errors when predicting the value of the Young's modulus. Furthermore, the median and the trimmed average are very close to the average, so it can be confirmed that hardly any bad results appear.

3.2. Yield Strength

Figure 8 shows the histogram of the yield strength values of the input dataset. These data are quite disperse and do not exhibit any predominant value. The yield strength strongly depends on the chemical composition of the alloy and the treatment applied to it. For instance, Al 7075-O (no treatment) alloy has a $YS = 140$ MPa, while Al 7075-T6 (heat treatment) shows a $YS = 455$ MPa [93]. As shown in Figure 8, the yield strength of aluminum alloys is a property that exhibits a wide range of values.

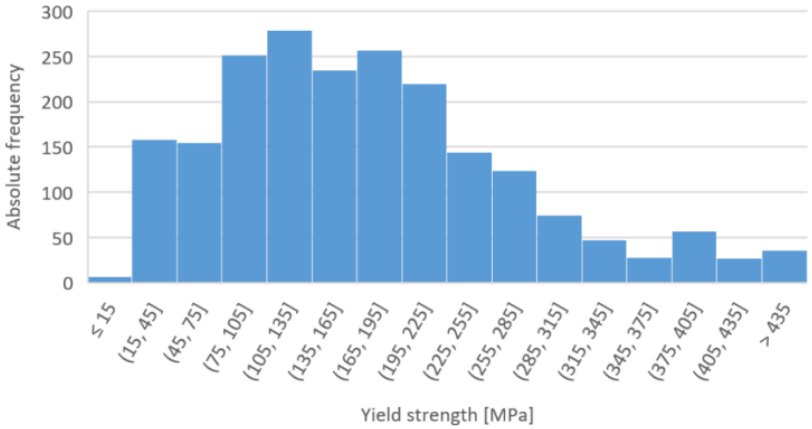


Figure 8. Yield strength histogram of the input dataset.

As already indicated, the position of the yield point is based on conventions (usually a deviation of 0.2% from the linear behavior) since, in fact, no significant physical phenomenon occurs in it [48]. Therefore, it is a property for which there is usually considerable uncertainty even in the reference bibliography (this data is usually given in the form of a range of values) [93].

Appendix C contains some notions about the neural network training process for predicting this property.

Once the training has been successfully completed, the neural network is asked to make predictions about the data from the testing subset. The averaged statistical metrics of Table 2 are obtained after performing the 10 training-prediction iterations.

Table 2. Average deviation (as %) of the prediction of the yield strength.

Avg. Dev.	Std. Dev.	Median	Avg. Dev. 90%
4.58	3.40	3.78	4.33

Table 2 shows the averaged information regarding the relative errors (according to Equation (A8)) of the 10 iterations: relative average deviation (Avg. Dev.), statistical standard deviation (Std. Dev.), median and trimmed average deviation at 90% (Avg. Dev. 90%).

The average precision of the prediction (average relative error) is 4.58% with a standard deviation of 3.40%. It is noteworthy that the average deviation, the median and the trimmed average deviation at 90% show very similar values (4.58%, 3.78% and 4.33% respectively), which indicates that the results are concentrated and few anomalous values appear. Figure 9 shows this same information in the form of a box and whisker diagram. In addition, this figure shows the outliers that have appeared during the process.

Figure 10 shows the histogram of the deviations of the yield strength prediction for all iterations (it shows the 4200 predictions that are made in the 10 iterations). This plot shows that most of the errors are lower than 6%. The error of the yield strength estimation is low (average 4.58%) but the results are more dispersed than in the case of Young's modulus because, as already indicated, it is a property that has an inherent uncertainty.

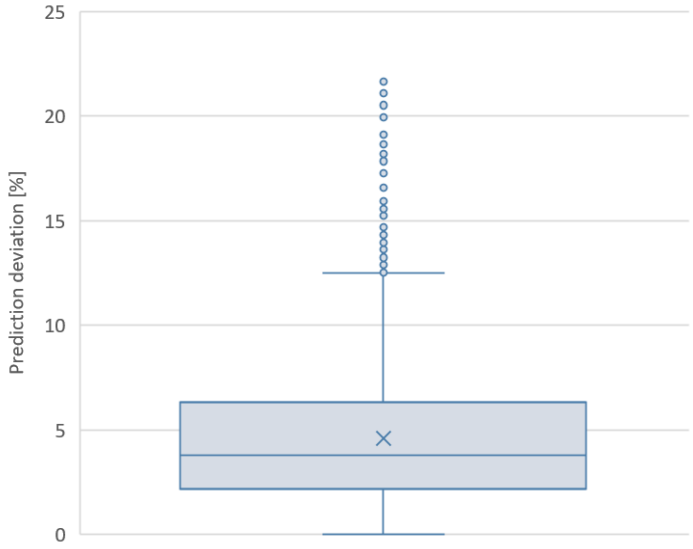


Figure 9. Prediction deviation of the yield strength.

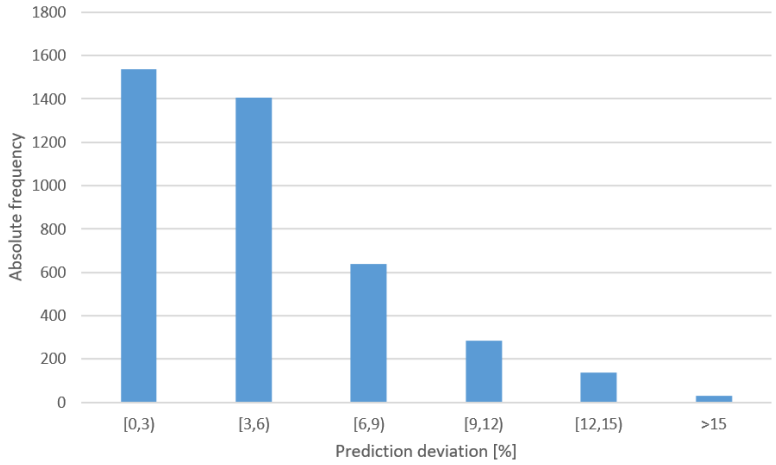


Figure 10. Histogram of the prediction error of the yield strength for all iterations.

3.3. Ultimate Tensile Strength

Figure 11 shows the histogram of the ultimate tensile strength values for all records in the input dataset. These data are heterogeneously distributed over a very wide range of values and, although a maximum appears in the range $(175, 205]$, it cannot be considered as truly remarkable. This figure highlights the great diversity of values that this property takes.

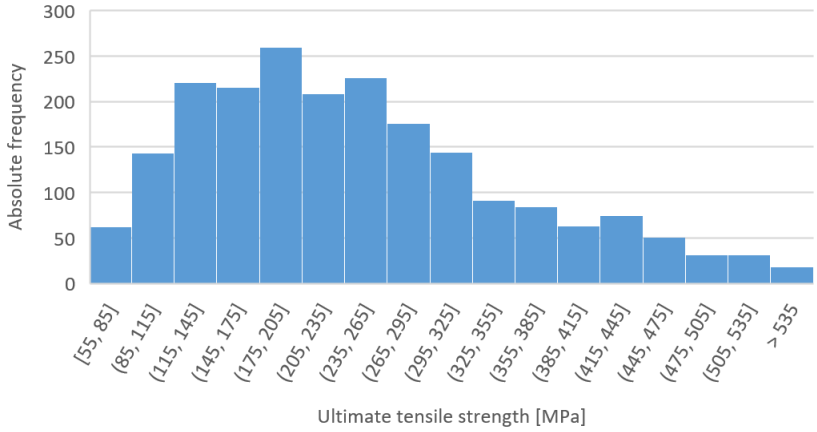


Figure 11. Ultimate tensile strength histogram of the input dataset.

Appendix C contains some notions about the neural network training process for predicting this property.

Table 3 shows various averaged statistical metrics about the relative error of the predictions (see Equation (A8)) that have been carried out: average deviation (Avg. Dev.), statistical standard deviation (Std. Dev.), median and trimmed average deviation at 90% (Avg. Dev. 90%). The average relative error of the system is 3.30%, being 2.55% and 3.08% the median and the trimmed average deviation at 90% respectively. These very low values account for the performance of the methodology.

Table 3. Average deviation (as %) of the prediction of the ultimate tensile strength

Avg. Dev.	Std. Dev.	Median	Avg. Dev. 90%
3.30	2.82	2.55	3.08

Figure 12 shows the result of the averaged prediction precision in the form of a box and whisker diagram. The presence of some abnormal values that have been marked as outliers should be highlighted. In this case, those anomalous results are related to alloys that exhibit unusually low ultimate tensile strength values and for which there are few samples in the input dataset.

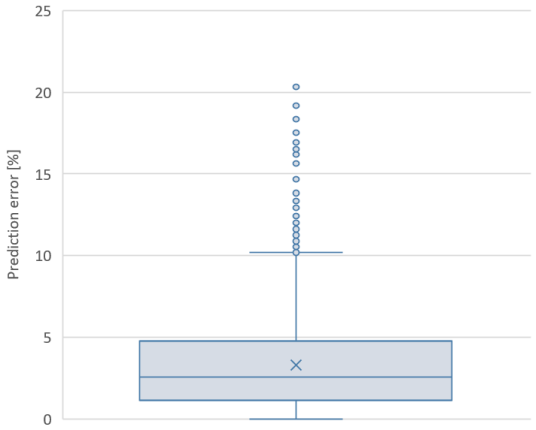


Figure 12. Prediction deviation of the ultiamte tensile strength.

Figure 13 shows a histogram of the errors made by the system throughout the 10 iterations that have been carried out. It is remarkable that there are few results greater than 6% and, in any case, most of the values are included in the range $[0, 3)$.

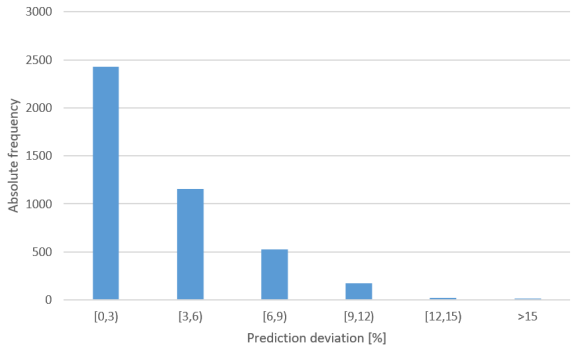


Figure 13. Histogram of the prediction error of the ultimate tensile strength for all iterations.

The system is more performant predicting the ultimate tensile strength than the yield strength because the former has a physical meaning and, therefore, the data in the input dataset is more precise.

3.4. Elongation at Break

Figure 14 shows the histogram of elongation at break for the entire input dataset. In this case, the data exhibits a wide range of values although they are concentrated around low values. Aluminum alloys, in general, are more ductile than steel and therefore easier to work with [34].

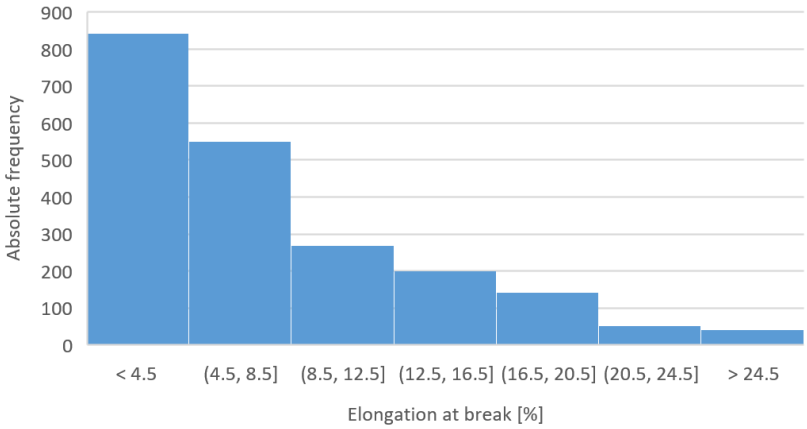


Figure 14. Elongation at break histogram of the input dataset.

Elongation at break is a very difficult property to determine since it requires an exhaustive test campaign that involves working with very high deformations, which implies very low straining rates [50]. Moreover, the behavior of the testing probes greatly depends on the metallurgic microstructure, the exact chemical composition and the treatments [34,50]. Therefore, the available data for this property is not very precise and is usually shown in the form of ranges, for example, the elongation at break of the Al 7075-T6 is 5–11% [39,93].

Appendix C contains some notions about the neural network training process for predicting this property.

Table 4 shows various statistical metrics related to the performance of the predictions. In the table, each column contains, respectively, the average deviation (Avg. Dev.), the statistical standard deviation (Std. Dev.), the median and the trimmed average deviation at 90% (Avg. Dev. 90%).

These results show a lower predictive performance than in the case of the other three considered properties: the mean deviations are higher (5.90%, 5.33% and 5.73% for the average, the median and the trimmed average). It is also noteworthy that the statistical standard deviation (Std. Dev.) is also greater (4.05%), which indicates that the results of these predictions are more scattered.

Table 4. Average deviation (as %) of the prediction of the elongation at break.

Avg. Dev.	Std. Dev.	Median	Avg. Dev. 90%
5.90	4.05	5.33	5.73

Figure 15 shows the averaged result of the predictive performance of the 10 iterations in the form of a box and whisker diagram. The results are more dispersed than in the other three cases and a few outliers with very high values appear. The network has been trained with data that, by its own nature, are imprecise (ranges) and it causes the results to be more heterogeneous. Figure 16 shows the histogram of the relative errors obtained in the prediction of the elongation at break for all the iterations. This plot shows that the deviations are concentrated on low values, with few abnormally high results.

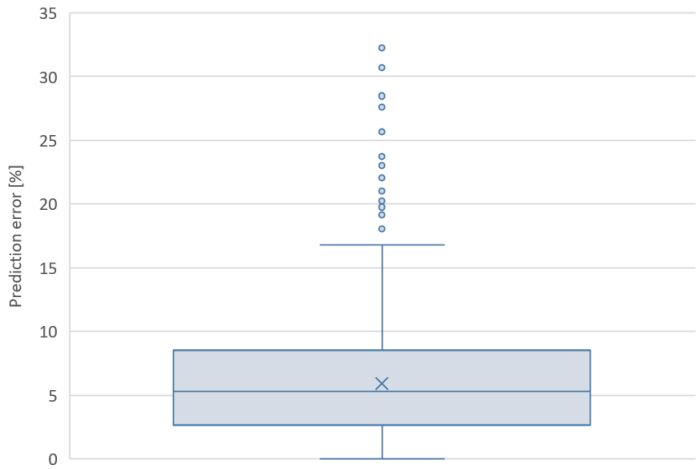


Figure 15. Prediction deviation of the elongation at break.

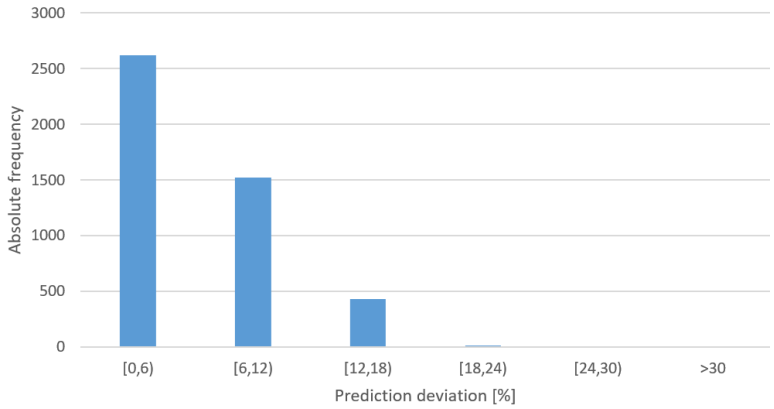


Figure 16. Histogram of the prediction error of the elongation at break for all iterations.

3.5. Limitations of the Methodology

The main limitation of this study is the size of the input dataset and the ability of the neural network to learn from it [32]. As already indicated, the outcomes of this methodology improve when the training process is carried out using a larger input set. However, obtaining large amounts of material data is difficult because it consumes a huge amount of resources (time, money, people...). Therefore, a larger initial information corpus can improve the results.

As previously described, the topology model that has been employed in this study has some disadvantages (e.g., the results are affected by a limited initial dataset and the local minimums

generate substantial attraction) that constitute a drawback of the procedure [6]. Other neural networks architectures can improve the results or reduce the required resources to carry out the training phase.

This study is founded on the assumption that the data obtained from the material library are correct and reliable [31]. The correctness of the input data do not modify the methodology but it can affect the results because the neural network would learn incorrect information.

4. Example of Application

The Al 2024-T4 alloy has been selected to develop this example because there is extensive information about it, it is easily comparable with data from leading sources and it is a widely used industrial material. Al 2024-T4 is a copper-based aluminum alloy (Al 2xxx) that has been treated with the T4 temper (solution heat-treated and natural aged) [81]. It has the highest ductility compared to the other variants of 2024 aluminum [1].

This is one of the best-known aluminum alloys due to its high strength and excellent fatigue behavior; it is widely used in structures and parts where a good strength-to-weight ratio is required [34]. Al 2024 alloy is easily machined to a high quality surface finish; moreover, it is easily plastically formed in the annealed condition (Al 2024-O) and, then, can be heat-treated to become Al 2024-T4. Since its resistance to corrosion is relatively low, this aluminum is commonly used with some type of coating or surface treatment [37].

Table 5 shows the chemical composition of Al 2024-T4 and Table 6 shows the mechanical properties that are relevant to this study [93].

Table 5. Al 2024-T4 chemical composition [81].

Element	Weight %
Al	90.7–94.7
Cr	Max. 0.1
Cu	3.8–4.9
Fe	Max. 0.5
Mg	1.2–1.8
Mn	0.3–0.9
Other	Max. 0.15
Si	Max. 0.5
Ti	Max. 0.15
Zn	Max 0.25

Table 6. Actual mechanical properties of the Al 2024-T4 [93].

Property	Value
Young's modulus [GPa]	73
Yield strength [MPa]	395
Ultimate tensile strength [MPa]	470
Elongation at break [%]	19

Before launching the software that carries out the training and prediction, to avoid overfitting, all references to Al 2024-T4 and -T351 (this is an identical standard regarding mechanical properties [93]) have been removed from the input dataset. In the same way as previously explained, 10 training-prediction iterations have been executed.

Table 7 shows the actual values (Actual val.) and the results of the prediction of the mechanical properties of Al 2024-T4, as well as some other statistical metrics that allow quantifying the error and the performance of the methodology for this particular case: average predicted value (Avg. val.), statistical standard deviation of the predictions (Std. Dev.), median, maximum (Max.) and minimum (Min.).

Table 7. Properties prediction for Al 2024-T4.

Property	Actual val.	Avg. val.	Std. Dev.	Median	Max.	Min.
E [GPa]	73	73.3	0.7	73.4	74.3	71.9
YS [MPa]	395	395.1	9.3	395.9	409.4	376.5
UTS [MPa]	470	471.5	8.0	470.8	483.2	460.1
A [%]	19	19.0	0.8	18.9	20.1	17.7

Table 8 shows various statistical results that summarize the predictive error for this alloy (the results are shown as a percentage). Note that the average errors do not exceed, in any case, 3.5%. The same information can be seen in Figure 17. With this information, it can be assured that the results adjust very well to the actual values.

Table 8. Prediction error for Al 2024-T4 (as %).

Property	Avg. error	Std. Dev.	Median	Max.	Min.
E [%]	0.84	0.55	0.81	1.77	0.08
YS [%]	1.63	1.62	0.95	4.68	0.20
UTS [%]	1.35	0.98	1.44	2.81	0.13
A [%]	3.21	2.34	3.08	6.89	0.05

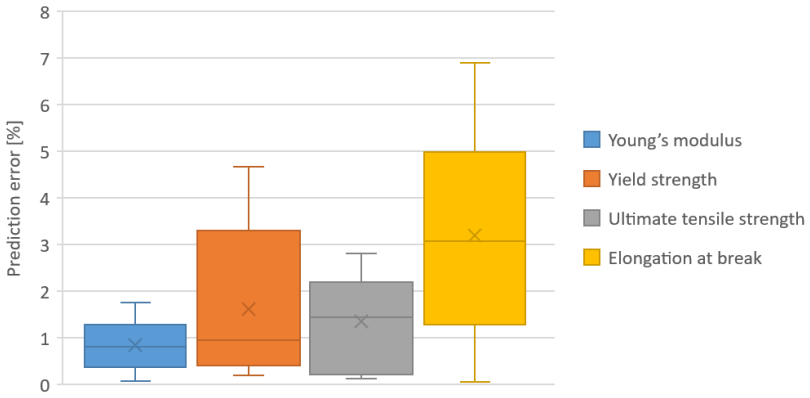


Figure 17. Prediction error for Al 2024-T4.

Note that the distribution of average errors is consistent with what was said previously: the better predictive performances have been obtained for the Young's modulus and the ultimate tensile strength, and the worse results for the yield strength and the elongation at break. This is also true for the statistical standard deviation values.

Figure 18 shows the actual stress-strain curve for Al 2024-T6 [50] and its bilinear approximation using the average values resulting from the prediction using the methodology described in this work (see Table 7). Note that the predicted curve fits the actual one (especially in the elastic region); however, discrepancies appear near the yield point and in the plastic zone.

The discrepancy between the two curves can be quantified by calculating the difference in deformation energies (see Equation (2)) [43,46]. This is equivalent to calculating the area enclosed between both curves. The deformation energy difference between the two curves is 2.74 MJ, so, 3.3% of the actual energy (83.8 MJ). This deviation is also an indication of the error made when using the approximation instead of the real curve.

Keeping the methodology error below 5% implies a similar performance as the typical artificial intelligence-based methodologies applied to materials science [24,25]. On the other hand, a similar error rate would be comparable, according to the Lean Manufacturing framework, to that of an industrial

system working at a four-sigma level, which has traditionally been associated with the average industry in developed countries [94,95].

As already indicated, obtaining the stress-strain curve of a material is a slow, expensive and resource-intensive process. However, based on this example, it can be said that using the methodology described in this paper allows shortening deadlines and having an estimate of the expected results.

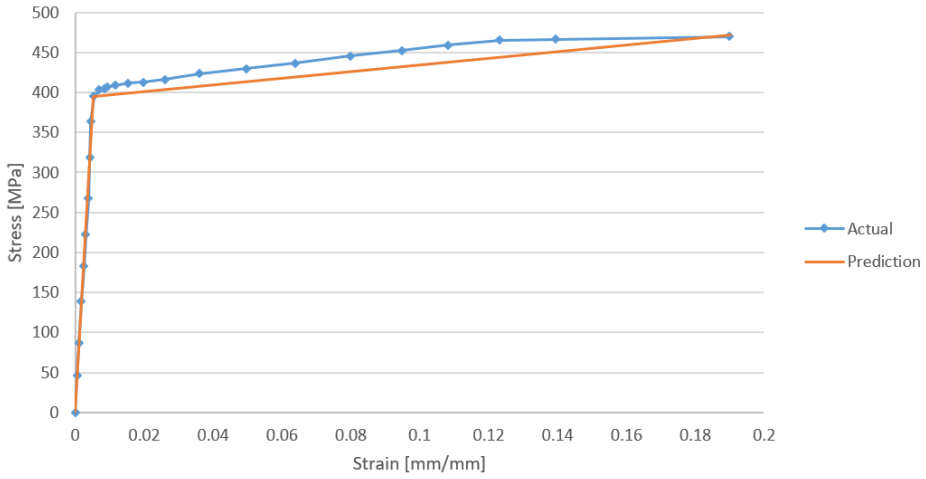


Figure 18. Actual stress-strain curve and its bilinear predicted approximation for Al 2024-T4 (actual curve from Reference [50]).

5. Conclusions and Future Work

This article has investigated the feasibility of using artificial neural networks and big data to predict the stress-strain curve of aluminum alloys whose chemical composition and previous treatments are known. The possibilities of artificial intelligence techniques have been explored based on large datasets. Therefore, the main conclusions of this work are presented as follows:

- Artificial neural network technology can be employed to exploit large material datasets to predict the mechanical properties of aluminum alloys. An ANN can learn to estimate the value of a material property based on its chemical composition and temper.
- An artificial neural network can be trained to predict the bilinear approximation of the stress-strain curve of an aluminum alloy if its chemical composition and tempers are well defined. The prediction error remains limited and the average deviations in this work for the Young's module, the yield strength, the ultimate tensile strength and the elongation at break are, respectively, 3.07%, 4.58%, 3.30% and 5.90%.
- Supervised learning methodologies require large training datasets to achieve satisfactory predictive performance. The predictive ability of a neural network improves as the dataset grows because it has more samples to learn from, and therefore, the network can approximate better the reality of the problem.
- A multilayer artificial neural network can be trained to approximate nonlinear functions related to materials science. Theoretically, a multilayer neural network can learn to approximate any nonlinear function if the training dataset is large enough and if it has a sufficient number of perceptrons [83].

This work contributes to applying innovative techniques such as those based in artificial intelligent techniques in materials science and technology research as it provides a new development tool to consider new aluminum alloys. It allows obtaining a first approximation and, therefore, focusing resources on the most promising materials. In addition, it opens the door to investigate similar solutions applied to other metals.

Artificial neural networks have proven to be a suitable ally to describe the elastoplastic behavior of highly relevant industrial materials without the need of expensive and complicated stress-strain tests. It can be studied whether it is possible to design a system based on artificial intelligence capable of predicting the stress-strain curve more accurately or using other better approaches such as the Ramberg-Osgood one [52].

Other more performant network architectures can be explored since this work scheme has shown that it is possible to use them to make these predictions. There is a wide spectrum of network topologies that cover different needs [78], which suggests that other solutions can be investigated.

Author Contributions: Conceptualization, D.M.F., A.R.-P. and A.M.C.; methodology, D.M.F.; software, D.M.F.; validation, D.M.F., A.R.-P. and A.M.C.; formal analysis, D.M.F.; investigation, D.M.F.; resources, A.R.-P. and A.M.C.; data curation, D.M.F.; writing–original draft preparation, D.M.; writing–review and editing, D.M., A.R.-P. and A.M.C.; visualization, D.M.; supervision, A.R.-P. and A.M.C.; project administration, A.M.C.; funding acquisition, A.M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been developed within the framework of the Doctorate Program in Industrial Technologies of the UNED and has been funded by the Annual Grants Call of the E.T.S.I.I. of the UNED via the projects of reference 2020-ICF04/B and 2020-ICF04/D.

Acknowledgments: We extend our acknowledgments to the Research Group of the UNED "Industrial Production and Manufacturing Engineering (IPME)". We also thank Matmatch GmbH for freely supplying all the material data employed to accomplish this study.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations and symbols are used in this manuscript:

A	Elongation at break
α	Ramberg-Osgood parameter
ADAM	Adaptive Moment Estimation
AI	Artificial intelligence
ANN	Artificial Neural Networks
β_n	ADAM algorithm parameter
E	Young's modulus
E_T	Strain hardening modulus
ϵ	ADAM stability factor
ε	Prediction error of a neural network
ε	Strain
η	ADAM step size
f	Error function
g	Gradient of the error function
m	ADAM first moment estimate
n	Ramberg-Osgood parameter
σ	Stress
σ_{YS}	Yield stress
UTS	Ultimate tensile strength
v	ADAM second moment estimate
w	Weights vector
YS	Yield stress

Appendix A. Aluminum Designation

The Aluminum Association Inc. is the main entity (among others) in charge of the regulation and standardization of all matters related to aluminum alloys. Although it is possible to subdivide these materials according to multiple criteria, this association distinguishes two basic categories: casting alloys and wrought alloys [81], being the latter the most widely produced and consumed [2]. The nomenclature of the different aluminum alloys is based on a 4-digit system that determines

the limits of the material composition and uniquely identifies it [81]. However, the meaning of each of the digits of this identification system varies between the casting and wrought alloys.

In the case of wrought alloys, the first digit (Xxxx) designates what is the main alloying element (see Table A1), the second digit (xXxx) indicates a modification or evolution of the original alloy (if it is different from 0) and the last two digits (xxXX) are simply arbitrary numbers that identify a specific alloy [81]. For example, in Al-2014, number 2 refers to an alloy whose main alloying agent is copper, number 0 indicates that there have been no modifications and 14 identifies this particular alloy.

Table A1. Principal alloying element for wrought aluminum alloys.

Alloy	Principal Alloying Element
1xxx	99% minimum aluminum
2xxx	Copper
3xxx	Manganese
4xxx	Silicon
5xxx	Magnesium
6xxx	Magnesium and silicon
7xxx	Zinc
8xxx	Others

On the other hand, for casting alloys, the first digit (Xxx.x) also identifies the main alloying element (see Table A2); the second and third digits (xXX.x) identify a particular alloy; and the fourth digit (decimal) indicates whether it is a final shape casting (.0) or an ingot (.1 or .2). Moreover, a capital letter prefix indicates a modification to a specific alloy [82]. For example, A256.0 indicates that this material is a modification (A) of an alloy whose main alloying element is copper (2) and which is offered in its final form (.0) and not as an ingot.

Table A2. Principal alloying element for cast aluminum alloys.

Alloy	Principal Alloying Element
1xx.x	99% minimum aluminum
2xx.x	Copper
3xx.x	Silicon plus copper and/or magnesium
4xx.x	Silicon
5xx.x	Magnesium
6xx.x	Unused series
7xx.x	Zinc
8xx.x	Tin
9xx.x	Other elements

Each of these alloys can be subjected to different heat and mechanical treatments (not all alloys are capable of undergoing all treatments) to modify their properties. To differentiate the treatment, there is a nomenclature (standardized by the Aluminum Association) whose identification is based on a letter, which indicates the type of process that the material has undergone (see Table A3), and numbers that identify the specific treatment [34]. For example, the 6012-H18 alloy has been strain hardened.

Table A3. Basic temper designation for aluminum alloys [34].

Letter	Meaning
F	As fabricated, it applies to products of a forming process in which no special control over thermal or strain hardening conditions is employed
O	Annealed, it applies to products which have been heated to produce the lowest strength condition to improve ductility and dimensional stability
H	Strain hardened, it applies to products that are strengthened through cold-working
W	Solution heat-treated, an unstable temper applicable only to alloys that age spontaneously at room temperature
T	Thermally treated, it applies to products that have been heat-treated

Appendix B. Neural Network Mathematical Explanation

A multi-layer neural network can be trained to learn a non-linear function [32] of the form (see Equation (A1)):

$$F(X) : \mathbb{R}^m \rightarrow \mathbb{R}^o, \tag{A1}$$

where $X = \{x_i / i \in 1...m\}$ is the input vector, m is the size of the input vector and o is the size of the output vector [66].

The neural network learning procedure is known as training, which is mathematically based on the gradient descent problem that tries to minimize the associated error function [32]. That error function depends on the weights related with each of the perceptrons. This vector of weights (whose size is equal to the number of neurons in the network) is represented as w and allows indicating that $f(w)$ is the error function when the weights w are assigned to each of the perceptrons of the network. With this formalization, the objective of the training is to find the vector w^* for which a global minimum of the function f is obtained, which turns the learning problem into an optimization problem [6].

In this way, a neural network is initialized with a vector of weights (in general, random) and, then, a new vector is calculated to reduce the error function [32]. This process is iterated until the error has been limited or until a specific stopping condition is satisfied. Since the error function is differentiable, the gradient of this function can be defined for each of the optimization steps (see Equation (A2)) [6]:

$$g_i = \nabla f_i = \nabla f(w_i), \tag{A2}$$

where g_i is the gradient value of the error function in the *i-th* step of the iteration, f_i is the value of the error function in the *i-th* step and w_i is the vector of weights in the *i-th* iteration.

Adaptive Moment Estimation (ADAM) is an adaptive learning rate methodology that calculates individual learning rates for different parameters. ADAM uses estimates of the first and second moment of a gradient to adapt the learning rate for each weight of the neural network [85]. Using this method, in each iteration, the new weight vector is calculated as (see Equation (A3)) [85]:

$$w_{i+1} = w_i - \eta \frac{\widehat{m}_{i+1}}{\sqrt{\widehat{v}_{i+1}} + \epsilon}, \tag{A3}$$

where η is the step size (a value that graduates the relevance of the gradient factor), ϵ is the stability factor of the algorithm (constant) and $\widehat{m}_{i+1}$ and $\widehat{v}_{i+1}$ are the bias-corrected first and second moment estimate, which are calculated as follows (refer to Equations (A4) and (A5)) [85]:

$$\widehat{m}_{i+1} = \frac{m_{i+1}}{1 - \beta_1^{i+1}} \tag{A4}$$

$$\widehat{v}_{i+1} = \frac{v_{i+1}}{1 - \beta_2^{i+1}}, \tag{A5}$$

where β_1 and β_2 are the algorithm parameters that are set to a value near 1 [72]; m_{i+1} and v_{i+1} are calculated as follows (refer to Equations (A6) and (A7)) [85]:

$$m_{i+1} = \beta_1 m_i + (1 - \beta_1)\, g_{i+1} \tag{A6}$$

$$v_{i+1} = \beta_2 v_i + (1 - \beta_2)\, {g_{i+1}}^2, \tag{A7}$$

where m_i and v_i are the decaying averages of past gradients and past squared gradients, respectively, and are estimates of the first moment (mean) and the second moment (non-centered variance) of the gradients [85].

Therefore, the optimization process and the network training method have been mathematically defined.

Once the network has been conveniently trained, predictions can be obtained based on the approximation function learned by the neural network [6]. The prediction deviation is calculated as the absolute value of the relative error of the resulting value (refer to Equation (A8)):

$$\varepsilon = \left| \frac{v_{prediction} - v_{real}}{v_{real}} \right|, \tag{A8}$$

where ε is the relative predictive error (in absolute value), $v_{prediction}$ is the predicted value (resulting value of the network) and v_{real} is the actual value.

The nodes of an artificial neural network can be connected in many ways, forming different network topologies. The behavior of the system, its learning capacity and the amount of resources it will need during the training and prediction phases depends greatly on the chosen topology [78]. A fully connected artificial neural network consists of a set of fully connected layers and a fully connected layer is a layer in which all nodes are connected to all nodes of the next layer [32].

For a fully connected multilayer neural network, the time complexity of the backpropagation training is given by Equation (A9). So, it is highly recommended to minimize the number of hidden nodes to reduce the training time [78].

$$\mathcal{O}\left(n \cdot m \cdot o \cdot N \cdot \prod_{i=1}^{k} h_i \right), \tag{A9}$$

where n is the size of the training dataset, m is the number of features, o is the number of output perceptrons, N is the number of iterations and k is the number of hidden layers (each of them containing h_i nodes).

Appendix C. Learning Curves

Figure A1 shows the averaged evolution of the error functions (on a logarithmic scale) relative to the training phase of each of the properties. Each curve is the result of averaging those obtained from each of the ten iterations.

In the case of the Young's modulus (E), the error function started at a value close to 2400 and evolved to converge asymptotically to about 30. It took around 1700 training epochs to reach the end of the process due to non-improvement conditions. Reaching a non-improvement condition, in general, indicates that the network is no longer capable of learning more from the provided data and, therefore, continuing the training could produce overfitting or some type of bias [78].

In the case of the Yield strength (YS), the curve evolved from approximately 20,000 to converge asymptotically to a value close to 300. It has taken around 8500 training epochs for the process to finish.

In the case of the Ultimate tensile strength (UTS), the curve started from a value close to 36,000 to descend until reaching a value close to 50, where it stabilizes. Training has required almost 12,000 epochs to stop.

In the case of the Elongation at break (A), the curve started from a value of approximately 50 and descend until reaching a value close to 3. Due to the scale, in this plot, it is possible to observe

the small oscillations that occur in the curves, which create small irregularities and hops. It is also very interesting to highlight the big steps that the curves create; these are usually related to instants in which the neural network learned an important rule [32].

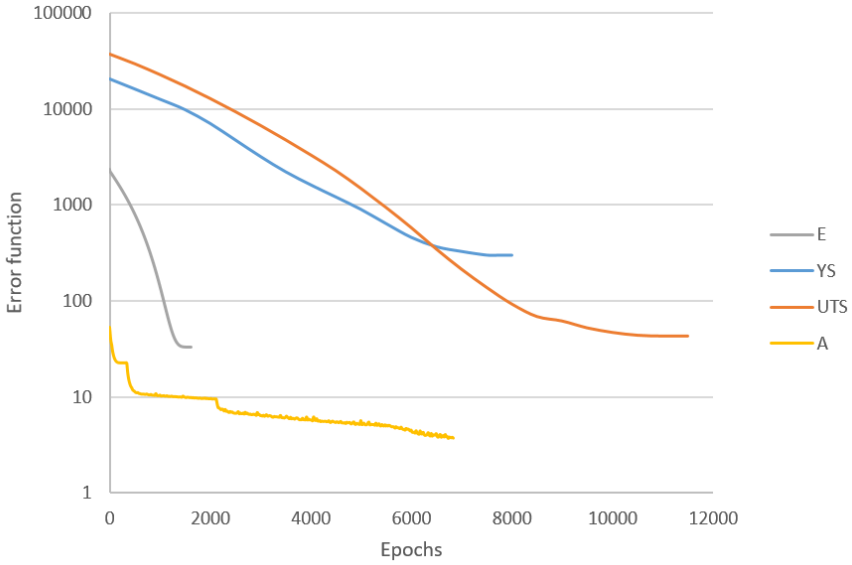


Figure A1. Averaged evolution of the error function during the training for each property.

References

1. Danylenko, M. Aluminium alloys in aerospace. *Alum. Int. Today* **2018**, *31*, 35–35.
2. Galevsky, G.; Rudneva, V.; Aleksandrov, V. *Current State of the World and Domestic Aluminium Production and Consumption*; IOP Conference Series: Materials Science and Engineering; IOP Publishing: Novokuznetsk, Russia, 2018; Volume 411, p. 012017.
3. Soo, V.K.; Peeters, J.; Paraskevas, D.; Compston, P.; Doolan, M.; Duflou, J.R. Sustainable aluminium recycling of end-of-life products: A joining techniques perspective. *J. Clean. Prod.* **2018**, *178*, 119–132.
4. Branco, R.; Berto, F.; Kotousov, A. *Mechanical Behaviour of Aluminium Alloys*; MDPI Applied Sciences: Basel, Switzerland, 2018.
5. Ashkenazi, D. How aluminum changed the world: A metallurgical revolution through technological and cultural perspectives. *Technol. Forecast. Soc. Chang.* **2019**, *143*, 101–113.
6. Merayo, D.; Rodríguez-Prieto, A.; Camacho, A. Prediction of Physical and Mechanical Properties for Metallic Materials Selection Using Big Data and Artificial Neural Networks. *IEEE Access* **2020**, *8*, 13444–13456.
7. Morini, A.A.; Ribeiro, M.J.; Hotza, D. Early-stage materials selection based on embodied energy and carbon footprint. *Mater. Des.* **2019**, *178*, 107861.
8. Piselli, A.; Baxter, W.; Simonato, M.; Del Curto, B.; Aurisicchio, M. Development and evaluation of a methodology to integrate technical and sensorial properties in materials selection. *Mater. Des.* **2018**, *153*, 259–272.
9. Mousavi-Nasab, S.H.; Sotoudeh-Anvari, A. A comprehensive MCDM-based approach using TOPSIS, COPRAS and DEA as an auxiliary tool for material selection problems. *Mater. Des.* **2017**, *121*, 237–253.
10. Das, D.; Bhattacharya, S.; Sarkar, B. Decision-based design-driven material selection: A normative-prescriptive approach for simultaneous selection of material and geometric variables in gear design. *Mater. Des.* **2016**, *92*, 787–793.
11. Alam, T.; Ansari, A.H. Review on Aluminium and its alloys for automotive applications. *Int. J. Adv. Technol. Eng. Sci.* **2017**, *5*, 278–294.
12. Kamaya, M.; Kawakubo, M. A procedure for determining the true stress–strain curve over a large range of strains using digital image correlation and finite element analysis. *Mech. Mater.* **2011**, *43*, 243–253.
13. Rodríguez-Prieto, Á.; Camacho, A.M.; Sebastián, M.Á. Materials selection criteria for nuclear power applications: A decision algorithm. *JOM* **2016**, *68*, 496–506.

14. Dimiduk, D.M.; Holm, E.A.; Niezgoda, S.R. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. *Integr. Mater. Manuf. Innov.* **2018**, *7*, 157–172.
15. Liu, Y.; Zhao, T.; Ju, W.; Shi, S. Materials discovery and design using machine learning. *J. Mater.* **2017**, *3*, 159–177.
16. Wang, Y.; Wu, X.; Li, X.; Xie, Z.; Liu, R.; Liu, W.; Zhang, Y.; Xu, Y.; Liu, C. Prediction and Analysis of Tensile Properties of Austenitic Stainless Steel Using Artificial Neural Network. *Metals* **2020**, *10*, 234.
17. Javaheri, E.; Kumala, V.; Javaheri, A.; Rawassizadeh, R.; Lubritz, J.; Graf, B.; Rethmeier, M. Quantifying Mechanical Properties of Automotive Steels with Deep Learning Based Computer Vision Algorithms. *Metals* **2020**, *10*, 163.
18. Abbas, A.T.; Pimenov, D.Y.; Erdakov, I.N.; Taha, M.A.; El Rayes, M.M.; Soliman, M.S. Artificial intelligence monitoring of hardening methods and cutting conditions and their effects on surface roughness, performance, and finish turning costs of solid-state recycled Aluminum alloy 6061 chips. *Metals* **2018**, *8*, 394.
19. Zhou, T.; Song, Z.; Sundmacher, K. Big Data Creates New Opportunities for Materials Research: A Review on Methods and Applications of Machine Learning for Materials Design. *Engineering* **2019**, *5*, 1017–1026.
20. Schmidt, J.; Marques, M.R.; Botti, S.; Marques, M.A. Recent advances and applications of machine learning in solid-state materials science. *Npj Comput. Mater.* **2019**, *5*, 1–36.
21. Ly, H.B.; Le, L.M.; Duong, H.T.; Nguyen, T.C.; Pham, T.A.; Le, T.T.; Le, V.M.; Nguyen-Ngoc, L.; Pham, B.T. Hybrid artificial intelligence approaches for predicting critical buckling load of structural members under compression considering the influence of initial geometric imperfections. *Appl. Sci.* **2019**, *9*, 2258.
22. Ling, J.; Antono, E.; Bajaj, S.; Paradiso, S.; Hutchinson, M.; Meredig, B.; Gibbons, B.M. Machine Learning for Alloy Composition and Process Optimization. In Proceedings of the ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition, Oslo, Norway, 11–15 June 2018; American Society of Mechanical Engineers Digital Collection: New York, NY, USA, 2018.
23. Twardowski, P.; Wiciak-Pikuła, M. Prediction of Tool Wear Using Artificial Neural Networks during Turning of Hardened Steel. *Materials* **2019**, *12*, 3091.
24. Asteris, P.G.; Roussis, P.C.; Douvika, M.G. Feed-forward neural network prediction of the mechanical properties of sandcrete materials. *Sensors* **2017**, *17*, 1344.
25. De Filippis, L.A.C.; Serio, L.M.; Facchini, F.; Mummolo, G.; Ludovico, A.D. Prediction of the vickers microhardness and ultimate tensile strength of AA5754 H111 friction stir welding butt joints using artificial neural network. *Materials* **2016**, *9*, 915.
26. Moayedi, H.; Kalantar, B.; Abdullahi, M.M.; Rashid, A.S.A.; Nazir, R.; Nguyen, H.; others. Determination of Young Elasticity Modulus in Bored Piles Through the Global Strain Extensometer Sensors and Real-Time Monitoring Data. *Appl. Sci.* **2019**, *9*, 3060.
27. Sun, D.; Lonbani, M.; Askarian, B.; Armaghani, D.J.; Tarinejad, R.; Pham, B.T.; Huynh, V.V. Investigating the Applications of Machine Learning Techniques to Predict the Rock Brittleness Index. *Appl. Sci.* **2020**, *10*, 1691.
28. Abambres, M.; Rajana, K.; Tsavdaridis, K.D.; Ribeiro, T.P. Neural Network-based formula for the buckling load prediction of I-section cellular steel beams. *Computers* **2019**, *8*, 2.
29. Szumigała, M.; Polus, Ł. An numerical simulation of an aluminium-concrete beam. *Procedia Eng.* **2017**, *172*, 1086–1092.
30. Lutz, M. *Programming Python: Powerful Object-Oriented Programming*; O'Reilly Media, Inc.: Newton, MA, USA, 2010.
31. GmbH, Matmatch Matmatch. Available online: https://matmatch.com/ (accessed on 15 April 2020).
32. Jackson, P.C. *Introduction to Artificial Intelligence*; Courier Dover Publications: Mineola, NY, USA, 2019.
33. Callister, W.D.; Rethwisch, D.G. *Materials Science and Engineering*; John wiley & sons: New York, NY, USA, 2011; Volume 5.
34. Kaufman, J.G. *Introduction to Aluminum Alloys and Tempers*; ASM international: Almere, The Netherlands, 2000.
35. Davis, J.R. *Alloying: Understanding the Basics*; ASM international: Almere, The Netherlands, 2001.
36. Scamans, G.; Butler, E. In situ observations of crystalline oxide formation during aluminum and aluminum alloy oxidation. *Metall. Trans. A* **1975**, *6*, 2055–2063.
37. Gui, F. Novel corrosion schemes for the aerospace industry. In *Corrosion Control in the Aerospace Industry*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 248–265.

38. Yogo, Y.; Sawamura, M.; Iwata, N.; Yukawa, N. Stress-strain curve measurements of aluminum alloy and carbon steel by unconstrained-type high-pressure torsion testing. *Mater. Des.* **2017**, *122*, 226–235.
39. ASM. *Atlas of Stress-Strain Curves*; ASM: Almere, The Netherlands, 2002.
40. ASTM, E8–99. *Standard Test Methods for Tension Testing of Metallic Materials (ASTM E8/E8M–16AE1)*; ASTM: West Conshohocken, PA, USA, 2001.
41. Bacha, A.; Maurice, C.; Klocker, H.; Driver, J.H. The large strain flow stress behaviour of aluminium alloys as measured by channel-die compression (20-500 C). *Mater. Sci. Forum* **2006**, *519*, 783–788.
42. Huang, C.; Jia, X.; Zhang, Z. A modified back propagation artificial neural network model based on genetic algorithm to predict the flow behavior of 5754 aluminum alloy. *Materials* **2018**, *11*, 855.
43. Nageim, H.; Durka, F.; Morgan, W.; Williams, D. Structural Mechanics–Loads, Analysis. In *Materials and Design of Structural Elements*, 7th ed.; Pearson International: England, UK, 2010.
44. ASTM, Committee E-28 on Mechanical Testing. *Standard Test Method for Young's Modulus, Tangent Modulus, and Chord Modulus*; ASTM International: West Conshohocken, PA, USA, 2004.
45. Hahn, G.; Rosenfield, A. Metallurgical factors affecting fracture toughness of aluminum alloys. *Metall. Trans. A* **1975**, *6*, 653–668.
46. Fertis, D.G. *Infrastructure Systems: Mechanics, Design, and Analysis of Components*; John Wiley & Sons: Hoboken, NJ, USA,1997; Volume 3.
47. Christensen, R.M. Observations on the definition of yield stress. *Acta Mech.* **2008**, *196*, 239–244.
48. Christensen, R.M. *The Theory of Materials Failure*; Oxford University Press: Oxford, UK, 2013.
49. Johnson, G.; Holmquist, T. *Test Data and Computational Strength and Fracture Model Constants for 23 Materials Subjected to Large Strains, High Strain Rates, and High Temperatures*; Los Alamos National Laboratory, LA-11463-MS: Los Alamos, NM, USA, 1989; Volume 198.
50. Nicholas, T. Material behavior at high strain rates. *Impact Dyn.* **1982**, *1*, 277–332.
51. Gere, J.; Goodno, B. Deflections of Beams. In *Mechanics of Materials*, 8th ed.; Cengage Learning: Boston, MA, USA, 2012.
52. Ramberg, W.; Osgood, W.R. *Description of Stress-Strain Curves by Three Parameters*; NASA: Washington, DC, USA, 1943.
53. Rasmussen, K.J.; Rondal, J. Strength curves for metal columns. *J. Struct. Eng.* **1997**, *123*, 721–728.
54. Pelletier, H.; Krier, J.; Cornet, A.; Mille, P. Limits of using bilinear stress–strain curve for finite element modeling of nanoindentation response on bulk materials. *Thin Solid Films* **2000**, *379*, 147–155.
55. *Eurocode 9—Design of Aluminium Structures*; BSI: London, UK, 2007.
56. Mazzolani, F. EN1999 Eurocode 9— Design of aluminium structures. In *Proceedings of the Institution of Civil Engineers-Civil Engineering*; Thomas Telford Ltd.: London, UK, 2001; Volume 144, pp. 61–64.
57. *ISO-EN. 6892-1. Metallic Materials-Tensile Testing—Part 1: Method of Test at Room Temperature*; International Organization for Standardization: Geneva, Switzerland, 2009.
58. Agrawal, A.; Choudhary, A. Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science. *APL Mater.* **2016**, *4*, 053208.
59. Song, I.Y.; Zhu, Y. Big data and data science: What should we teach? *Expert Syst.* **2016**, *33*, 364–373.
60. Rowley, J. The wisdom hierarchy: Representations of the DIKW hierarchy. *J. Inf. Sci.* **2007**, *33*, 163–180.
61. Batra, S. Big data analytics and its reflections on DIKW hierarchy. *Rev. Manag.* **2014**, *4*, 5.
62. White, A.A. Big data are shaping the future of materials science. *MRS Bull.* **2013**, *38*, 594–595.
63. García-Gil, D.; Ramírez-Gallego, S.; García, S.; Herrera, F. Principal components analysis random discretization ensemble for big data. *Knowl. Based Syst.* **2018**, *150*, 166–174.
64. Erl, T.; Khattak, W.; Buhler, P. *Big Data Fundamentals: Concepts, Drivers & Techniques*; Prentice Hall Press: Upper Saddle River, NJ, USA, 2016.
65. Weinbub, J.; Wastl, M.; Rupp, K.; Rudolf, F.; Selberherr, S. ViennaMaterials–A dedicated material library for computational science and engineering. *Appl. Math. Comput.* **2015**, *267*, 282–293.
66. Merayo, D.; Rodriguez-Prieto, A.; Camacho, A. Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0. *Procedia Manuf.* **2019**, *41*, 42–49.
67. Helal, S. The expanding frontier of artificial intelligence. *Computer* **2018**, *51*, 14–17.
68. McCarthy, J.; Minsky, M.L.; Rochester, N.; Shannon, C.E. A proposal for the dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Mag.* **2006**, *27*, 12.

69. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
70. Johnson, K.W.; Soto, J.T.; Glicksberg, B.S.; Shameer, K.; Miotto, R.; Ali, M.; Ashley, E.; Dudley, J.T. Artificial intelligence in cardiology. *J. Am. Coll. Cardiol.* **2018**, *71*, 2668–2679.
71. Cummings, M. *Artificial Intelligence and the Future of Warfare*; Chatham House for the Royal Institute of International Affairs London: London, UK, 2017.
72. Villa, F.; Ceroni, M.; Bagstad, K.; Johnson, G.; Krivov, S. ARIES (Artificial Intelligence for Ecosystem Services): A new tool for ecosystem services assessment, planning, and valuation. In Proceedings of the 11th Annual BIOECON Conference on Economic Instruments to Enhance the Conservation and Sustainable Use of Biodiversity, Venice, Italy, 21–22 September 2009; pp. 21–22.
73. Allen, G.; Chan, T. *Artificial Intelligence and National Security*; Belfer Center for Science and International Affairs: Cambridge, MA, USA, 2017.
74. Ee, J.H.; Huh, N. A study on the relationship between artificial intelligence and change in mathematics education. *Commun. Math. Educ.* **2018**, *32*, 23–36.
75. Kolesov, V. Cognitive Modelling in Oil & Gas Exploration and Reservoir Prediction. In Proceedings of the 80th EAGE Conference and Exhibition 2018, Copenhaga, Dinamarca, 11–14 Novermber 2018; European Association of Geoscientists & Engineers: Houten, The Netherlands, 2018; Volume 2018, pp. 1–5.
76. Thankachan, T.; Prakash, K.S.; Pleass, C.D.; Rammasamy, D.; Prabakaran, B.; Jothi, S. Artificial neural network to predict the degraded mechanical properties of metallic materials due to the presence of hydrogen. *Int. J. Hydrog. Energy* **2017**, *42*, 28612–28621.
77. Qian, L.; Winfree, E.; Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **2011**, *475*, 368–372.
78. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117.
79. Huang, G.; Huang, G.B.; Song, S.; You, K. Trends in extreme learning machines: A review. *Neural Netw.* **2015**, *61*, 32–48.
80. Joshi, P. *Artificial Intelligence with Python*; Packt Publishing Ltd.: Birmingham, UK, 2017.
81. The Aluminum Association. *International Alloy Designations and Chemical Composition Limits for Wrought Aluminum and Wrought Aluminum Alloys*; The Aluminum Association: Arlington, VA, USA, 2015.
82. The Aluminum Association. *Designations and Chemical Composition Limits for Aluminum Alloys in the Form of Castings and Ingot*; The Aluminum Association: Arlington, VA, USA, 2006.
83. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251–257.
84. Deshpande, A.; Kumar, M. *Artificial Intelligence for Big Data: Complete Guide to Automating Big Data Solutions Using Artificial Intelligence Techniques*; Packt Publishing Ltd.: Birmingham, UK, 2018.
85. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference for Learning Representations (ICLR 15). San Diego, CA, USA, 7–9 May 2015.
86. Elmishali, A.; Stern, R.; Kalech, M. An artificial intelligence paradigm for troubleshooting software bugs. *Eng. Appl. Artif. Intell.* **2018**, *69*, 147–156.
87. Bouanan, Y.; Zacharewicz, G.; Vallespir, B. DEVS modelling and simulation of human social interaction and influence. *Eng. Appl. Artif. Intell.* **2016**, *50*, 83–92.
88. Perikos, I.; Hatzilygeroudis, I. Recognizing emotions in text using ensemble of classifiers. *Eng. Appl. Artif. Intell.* **2016**, *51*, 191–201.
89. Li, W.; Le Gall, F.; Spaseski, N. A survey on model-based testing tools for test case generation. In Proceedings of the International Conference on Tools and Methods for Program Analysis, Moscow, Russia, 3–4 March 2017; Springer: Berlin, Germany, 2017; pp. 77–89.
90. Lal, A.; others. SANE 2.0: System for fine grained named entity typing on textual data. *Eng. Appl. Artif. Intell.* **2019**, *84*, 11–17.
91. Siegel, J.E.; Pratt, S.; Sun, Y.; Sarma, S.E. Real-time deep neural networks for internet-enabled arc-fault detection. *Eng. Appl. Artif. Intell.* **2018**, *74*, 35–42.
92. Martín, A.; Rodríguez-Fernández, V.; Camacho, D. CANDYMAN: Classifying Android malware families by modelling dynamic traces with Markov chains. *Eng. Appl. Artif. Intell.* **2018**, *74*, 121–133.
93. ASM International Handbook Committee. *Properties and Selection: Nonferrous Alloys and Special-Purpose Materials Volume 2*; ASM Handbook; ASM International: Novelty, OH, USA, 2010.

94. Socconini, L.V.; Reato, C. *Lean Six Sigma*; Marge Books: Barcelona, Spain, 2019.
95. Furterer, S.L. *Lean Six Sigma in Service: Applications and Case Studies*; CRC press: Boca Raton, FL, USA, 2016.

4.5 *Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys* [25]

Los indicios de calidad de este artículo pueden encontrarse en el Apéndice E.

4.5.1 Datos de la publicación y factor de impacto

Tabla 5. Factor de impacto de Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys

Título	Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys
Autores	David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Revista	Materials
ISSN	1996-1944
Editorial	MDPI
País	Suiza
Volumen	13 (5227)
Páginas	1-22
Fecha	2020
doi	10.3390/ma13225227
Factor de impacto	3.057 (2019 Journal Citation Reports) 132/314 (Materials Science, Multidisciplinary, Q2)

4.5.2 Resumen y copia de la publicación

En el conformado de metales, el comportamiento plástico de las aleaciones metálicas está directamente relacionado con su conformabilidad, y tradicionalmente se ha caracterizado por modelos simplificados de las curvas de fluencia, especialmente en el análisis por simulación de elementos finitos y métodos analíticos. Las herramientas basadas en redes neuronales artificiales han mostrado un alto potencial para predecir el comportamiento y las propiedades de los componentes industriales. Las aleaciones de aluminio se encuentran entre los materiales más utilizados en industrias desafiantes como la aeroespacial, la automoción o el envasado de alimentos. En este estudio, se desarrolla una herramienta asistida por computadora para predecir dos de las propiedades mecánicas más útiles de los materiales metálicos para caracterizar el comportamiento plástico, el límite elástico y la resistencia máxima a la tracción. Estos pronósticos se basan en la composición química de la aleación, el temple y la dureza Brinell. En este estudio, se emplea una base de datos de materiales para entrenar una red neuronal artificial que es capaz de hacer predicciones con una confianza superior al 95%. También se muestra que esta metodología logra un desempeño similar al de las ecuaciones empíricas desarrolladas expresamente para un material específico, pero brinda mayor generalidad ya que puede aproximar las propiedades de cualquier aleación de aluminio. La metodología se basa en el uso de redes neuronales artificiales respaldadas por una gran colección de datos sobre las propiedades de miles de materiales comerciales. Por lo tanto, los datos de entrada superan las 2000 entradas. Cuando se ha recopilado y organizado la información relevante, se define una red neuronal artificial, y después del entrenamiento, la inteligencia artificial es capaz de hacer predicciones sobre las propiedades del material con una confianza promedio superior al 95%.

In metal forming, the plastic behavior of metallic alloys is directly related to their formability, and it has been traditionally characterized by simplified models of the flow curves, especially in the analysis by finite element simulation and analytical methods. Tools based on artificial neural networks have shown high potential for predicting the behavior and properties of industrial components. Aluminum alloys are among the most broadly used materials in challenging industries such as aerospace, automotive, or food packaging. In this study, a

computer-aided tool is developed to predict two of the most useful mechanical properties of metallic materials to characterize the plastic behavior, yield strength and ultimate tensile strength. These prognostics are based on the alloy chemical composition, tempers, and Brinell hardness. In this study, a material database is employed to train an artificial neural network that is able to make predictions with a confidence greater than 95%. It is also shown that this methodology achieves a performance similar to that of empirical equations developed expressly for a specific material, but it provides greater generality since it can approximate the properties of any aluminum alloy. The methodology is based on the usage of artificial neural networks supported by a big data collection about the properties of thousands of commercial materials. Thus, the input data go above 2000 entries. When the relevant information has been collected and organized, an artificial neural network is defined, and after the training, the artificial intelligence is able to make predictions about the material properties with an average confidence greater than 95%.

materials

MDPI

Article

Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys

David Merayo *, Alvaro Rodríguez-Prieto and Ana María Camacho

Department of Manufacturing Engineering, Universidad Nacional de Educación a Distancia (UNED), Juan del Rosal 12, 28040 Madrid, Spain; alvaro.rodriguez@ind.uned.es (A.R.-P.); amcamacho@ind.uned.es (A.M.C.)

* Correspondence: dmerayo1@alumno.uned.es

Received: 2 October 2020; Accepted: 16 November 2020 ; Published: 19 November 2020

Abstract: In metal forming, the plastic behavior of metallic alloys is directly related to their formability, and it has been traditionally characterized by simplified models of the flow curves, especially in the analysis by finite element simulation and analytical methods. Tools based on artificial neural networks have shown high potential for predicting the behavior and properties of industrial components. Aluminum alloys are among the most broadly used materials in challenging industries such as aerospace, automotive, or food packaging. In this study, a computer-aided tool is developed to predict two of the most useful mechanical properties of metallic materials to characterize the plastic behavior, yield strength and ultimate tensile strength. These prognostics are based on the alloy chemical composition, tempers, and Brinell hardness. In this study, a material database is employed to train an artificial neural network that is able to make predictions with a confidence greater than 95%. It is also shown that this methodology achieves a performance similar to that of empirical equations developed expressly for a specific material, but it provides greater generality since it can approximate the properties of any aluminum alloy. The methodology is based on the usage of artificial neural networks supported by a big data collection about the properties of thousands of commercial materials. Thus, the input data go above 2000 entries. When the relevant information has been collected and organized, an artificial neural network is defined, and after the training, the artificial intelligence is able to make predictions about the material properties with an average confidence greater than 95%.

Keywords: aluminum; artificial neural network; chemical composition; heat treatment; Brinell hardness; material properties' prognosis; yield strength; UTS

1. Introduction

In metal forming, the plastic behavior of metallic alloys is directly related to their formability, that is the material's ability to undergo plastic deformation. In fact, for the study of forming processes, the plastic behavior of metals is typically implemented in finite element simulation [1,2] and other analytical methods [3] by the so-called flow curves. In engineering applications, the plastic behavior has been traditionally characterized by simplified models; apart from well-known models such as Hollomon's power law and the Swift model [4], a linear approximation is the easiest way to model the plastic behavior through the mechanical properties' yield strength (YS) and ultimate tensile strength (UTS), especially when strain hardening has an important role such as in cold forming conditions.

1.1. Aluminum Alloys

Aluminum alloys have one of the lowest densities among structural metals, high resistance to corrosion, high strength (even higher than some steels), and good electrical and thermal conductivity [5]. On the other hand, they have an excellent workability and machinability and allow a large number of surface finishes [5]. Aluminum is a ductile and malleable material that can be shaped using a wide variety of techniques; however, its alloys have very different properties that significantly affect its forming behavior [6,7].

Pure aluminum (1100-O) is a relatively soft material [5]; however, its alloys exhibit an enormous variety of resistance and ductility values, which make it an exceptional material. The main alloying elements of aluminum are copper, magnesium, silicon, manganese, nickel, and zinc [8]. The relatively low hardness of aluminum can sometimes cause abrasive wear of the material [9]. The low hardness of pure aluminum makes it a non-ideal material for building structures. For this application, alloying elements must be added to make it stronger [10]. These supplementary elements do not simply improve the hardness of the metal, but also modify other properties [11,12]. Besides, heat-treated aluminum alloys can withstand more load due to the precipitation hardening process of aluminum, even though the hardness is different due to the addition of various alloying agents [9].

Nondestructive approaches to approximate the yield strength (YS) and the ultimate tensile strength (UTS) have been often of interest to engineers [5] because mechanical data can be collected quickly without requiring samples for testing [13]. One very relevant technique to estimate these tensile properties has been the hardness test because it is almost nondestructive (leaving behind only a small indentation) [14]. Moreover, there are techniques that, in the industrial environment, are considered as totally non-destructive, such as the ultrasonic contact impedance hardness test [15,16].

The methodological approach presented in this paper has the advantage of allowing the yield stress and the ultimate tensile strength to be estimated based on known values (chemical requirements specified by standards [17] and heat treatments, both typically accessible in databases) and an almost non-destructive test (hardness test). Using traditional techniques, these two properties can only be obtained by conducting tensile tests (destructive) that require specimens and access to facilities and other resources such as equipment. On the other hand, characterizing the mechanical resistance of the material together with a good design improve the safety level.

There is a wide variety of decision support systems focused on materials engineering [18]; however, very few really take advantage of technologies based on artificial intelligence [9,10,19–22]. Although several studies that use machine learning to address metallotechnics and the properties of metals have been published [23,24], aluminum alloys have hardly been investigated considering their tempers from an industrial perspective using these tools [25].

In this study, a computer-aided methodology is developed to predict some fundamental properties of aluminum alloys whose chemical composition and treatments (thermal and mechanical) are known [10]. The system presented in this work is able to predict accurately the yield stress (YS) and the ultimate tensile strength (UTS) of aluminum alloys based on Brinell hardness data [9,26]. This methodology uses artificial neural networks (ANNs) and technology based on machine learning to carry out the predictions.

1.2. Brinell Hardness and Material Strength

The Brinell hardness scale, proposed in 1900, is an indentation hardness measurement scale in which the penetration of a spherical ball into the test material is measured [27]. This indenter, when subjected to a standard load, deforms the material, creating a spherical cap, whose diameter is used to calculate the Brinell hardness [28,29]. The Brinell (HB) and Vickers (HV) hardness scales are among the most widely used, and in the case of aluminum alloys, their values are equivalent [30,31]. Aluminum alloys have hardness values ranging from HB∼20 to HB∼200 [32].

These two scales are based on the measurement of the plastic deformation that occurs on a material when applying a standard load through a standard penetrator [28,29]. Therefore, the mechanical

process is closely related to the properties that define the plastic deformation of a material [31], namely the YS and the UTS. In the development of the indentation that occurs during the Brinell hardness test, the main mechanical process that takes place is the plastic flow of the metal around the indenter [11].

The relationship between the hardness of a material and their tensile properties is so intimate that mathematical expressions have been proposed; however, these equations are not universal [33]. Both the YS and the UTS of steels and aluminums exhibit a correlation with the hardness over the entire range of strength values, and so, empirical relationships can be provided that enable the estimation of strength from a bulk hardness measurement if a certain error is allowed [34–37].

A good understanding of the correlation between the tensile properties and the hardness of materials is very noteworthy [38]:

- Reliable hardness strength relationships allow for quick mechanical property evaluations by means of fast and low-cost hardness testing instead of complicated tensile testing.
- Contrary to tensile tests, the hardness can be measured non-destructively in situ on fully assembled devices, therefore allowing for structural integrity tests in service.
- Often, new materials could only be manufactured at a small scale; these materials are not sufficient to accomplish extensive tensile testing, and so, hardness testing is frequently the only option.

There has been a great interest in trying to find approximations to the tensile properties from the results of the hardness tests. These efforts can be grouped as follow:

- Estimation of the tensile properties directly from the results of hardness tests [13,14,35,38–40].
- Estimation of the tensile properties indirectly from a material constant obtained from the results of hardness tests [13,14,38].

In this study, a third new method is going to be developed, similar to the first one, but with the particularity that the function that relates the tensile properties with Brinell hardness will be obtained using artificial intelligence and machine learning [41]. This new methodology takes, as input, the chemical composition, the temper, and the Brinell hardness and, then, approximates the yield strength and the ultimate tensile strength [9,10].

This new approach will be compared with the equations that directly estimate the tensile properties from the Brinell hardness test values. The literature is consistent in reflecting a linear relationship between the yield strength and the result of the Brinell hardness test (see Equation (1)) [13,14,38,40].

$$YS = \beta_1 \cdot HB - \beta_0 \tag{1}$$

where YS is the yield strength, HB is the Brinell hardness, and β_0 and β_1 are coefficients related to the material that are strongly affected by the tempers. For aluminum alloys, $\beta_1 \approx 3$ [13,14,38].

Estimating the UTS from hardness data has been mostly experimental because the phenomenon of tensile instability after which engineering strain increases while engineering stress falls does not take place during indentation [14]. However, it is common to find linear approximations with slopes close to three (see Equation (2)) [38,42,43].

$$UTS \approx 3 \cdot HB \tag{2}$$

where UTS is the ultimate tensile strength and HB is the Brinell hardness.

In the case of UTS, a linear expression similar to the yield stress approximation is proposed. In this case, the y-intersection coefficient γ_0 is positive (see Equation (3)) [14]:

$$UTS = \gamma_1 \cdot HB - \gamma_0 \tag{3}$$

where UTS is the ultimate tensile strength, HB is the Brinell hardness, and γ_0 and γ_1 are coefficients related to the material.

Although estimating tensile properties from hardness data has been seen as a very practical method, some persistent discrepancies between theoretical models and best-fit equations resulting from empirical data have to be resolved [40]. Only the equations that have been developed for a precise alloy are able to make good predictions at the cost of being too specific. In general, the literature focuses its efforts on specific materials or on certain types of alloys [44]. It is not common to find studies that contain general expressions that are valid for all aluminum alloys, except those that require obtaining empirical coefficients that are employed to fit the equations [35,39].

1.3. Artificial Neural Networks

Since its inception [45], artificial intelligence (AI) has shown that it can be applied to a wide spectrum of disciplines not directly related to computing. Among the most significant new uses, medicine [46,47], warfare [48], ecology [49], security [50], education [51], oil exploration [52], or material science [44,53] can be emphasized. Today, ANNs have become one of the most notable AI methods due to their incredible accomplishments and unstoppable advancement [10,54].

An ANN is a mathematical model inspired by the biological behavior of neurons and the structure of the brain [41,55]. In a multi-layer neural network, perceptrons, the basic units that form an ANN, are hierarchically organized into layers [54]. A layer is a set of neurons not connected to themselves nor to each other that receive their input data from the same source (the outside or another layer) and that send their information to the same destination (another layer or the outside) [41].

Therefore, three types of layers can be distinguished: the input layer, which receives information from the outside; the hidden layers are those whose inputs and outputs are within the system and, therefore, have no contact with the outside; and finally, the output layer, which sends the response from the network to the outside [9,10]. Many different topological models can be defined depending on the organization of the perceptrons and the type of connections they establish between them [41,54].

A multilayer ANN is a supervised learning algorithm capable of learning a nonlinear function by training on a labeled input dataset that can be used to perform classifications and regressions [41,54,56] and that can overcome traditional programming in some tasks. However, neural systems are not without certain drawbacks. One of the most important is that they usually carry out such complex processing that involves thousands of operations, so it is not possible to follow, step by step, the reasoning that led them to draw their conclusions [41,57]. Nevertheless, in small networks, by simulation or by studying synaptic weights, it is possible to know, at least, which input variables have been relevant in decision-making [58–60].

Various artificial intelligence techniques have been implemented in the material science field to carry out different types of analysis or predictions about the properties and behavior of industrial components and materials [20,22]: prediction of elastic properties of metals [10,61], prediction of metallic components behavior [19,62–64], optimization of alloy composition [25,65], or early prediction of the degradation of metallic materials [53,66]. As already stated, artificial intelligence and neural networks can be applied to almost all science fields [57,67].

The main objective of this work is to develop an ANN capable of making precise predictions about two of the main tensile properties of commercial aluminum alloys. Therefore, it must be ensured that the predictive error is small and that all the information obtained throughout the training and prediction phases is analyzed to obtain statistical metrics about the performance of the methodology.

2. Methodology

Figure 1 schematically shows the different stages of the methodology of this work. It consists of three main stages: the stage of input dataset creation; the stage of ANN definition; and the stage of training, prediction, and analysis. With the aim of increasing reliability, this work scheme guarantees that the data that reach each phase have been correctly prepared and processed in the previous one and are ready to be employed [9,10].

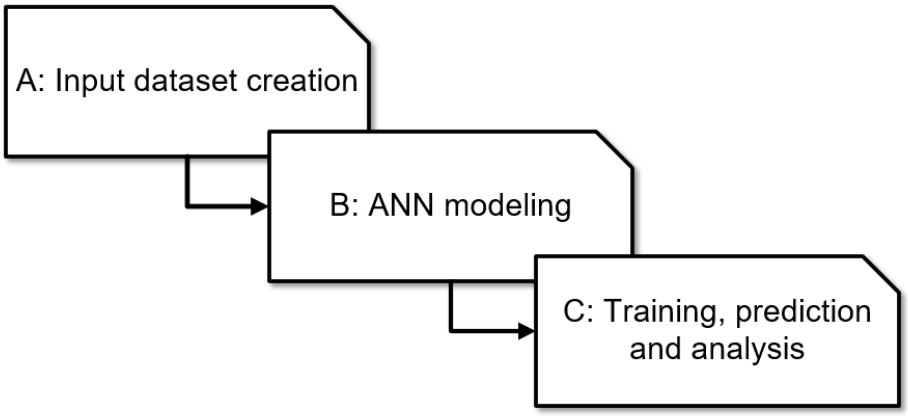


Figure 1. Overview of the methodology.

2.1. Stage A: Input Dataset Creation

The data employed to carry out this study comprised both wrought and casting alloys [9,10] and were obtained from Matmatch (Munich, Germany) [32], which is an open-access online materials library, which is comprised of thousands of entries [10,68]. For each record, it is possible to obtain a datasheet (in general, heterogeneous and non-exhaustive) that must be filtered, organized and processed to obtain a corpus of accurate and useful information [69].

The construction of the input dataset requires several processes aimed at guaranteeing the quality of the final corpus of information: reading the datasheets and organizing them as a table; unifying units and eliminating ranges (the conservative criterion of maintaining the maximum of the range is adopted); filtering the data according to the established criteria and eliminating duplicates. Figure 2 shows an overview of this process.

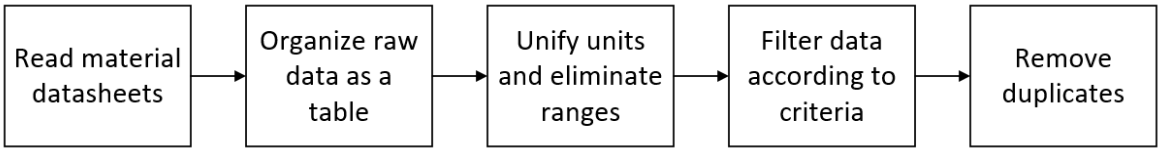


Figure 2. Overview of the data filtering process.

The following considerations and criteria were taken when filtering and organizing the available data [32]:

- Only records containing information on its Brinell hardness (HB), its yield stress (YS), and its ultimate stress (UTS) were considered [70,71]. Even if the methodology is able to infer the missing information, it is necessary to have all the data to carry out the training or to calculate the precision of the prediction [10].
- Only alloys whose chemical composition is defined at more than 95% are taken into account [10].
- Only 11 chemical elements are considered to define the chemical composition of the alloys [72]: Al, Zn, Cu, Si, Fe, Mn, Mg, Ti Cr, Ni, and Zr. All other chemical elements are tagged as non-relevant, and their mass contribution is regrouped as "other" [17,73].
- The methodology only considers 35 different treatments: F (as fabricated, single type), O (annealed, single type), H (strain hardening, 19 types of treatment), and T (thermally treated, 14 types of treatments) [17,73]. All records indicating other treatments were eliminated because the sample was so small that it could cause errors during training and prediction [56].

Most of the discarded records were eliminated because the information they contained was imprecise (the definition of the chemical composition was poor) or incomplete (some of the relevant

properties were missing, especially the Brinell hardness). After conveniently filtering and organizing the 5341 datasheets, seven-hundred thirteen aluminum alloys records were kept.

2.2. Stage B: ANN Modeling

Once the input dataset is prepared (filtered and organized), the neural network is defined: a fully connected multilayer feedforward topology [56], which comprises one input layer, three hidden layers, and one output layer. This network topology is made up of these five layers, and all the perceptrons of each layer are only connected to all the perceptrons in the next layer so that the information only travels in a single direction, from the input layer to the output layer (see Figure 3) [74].

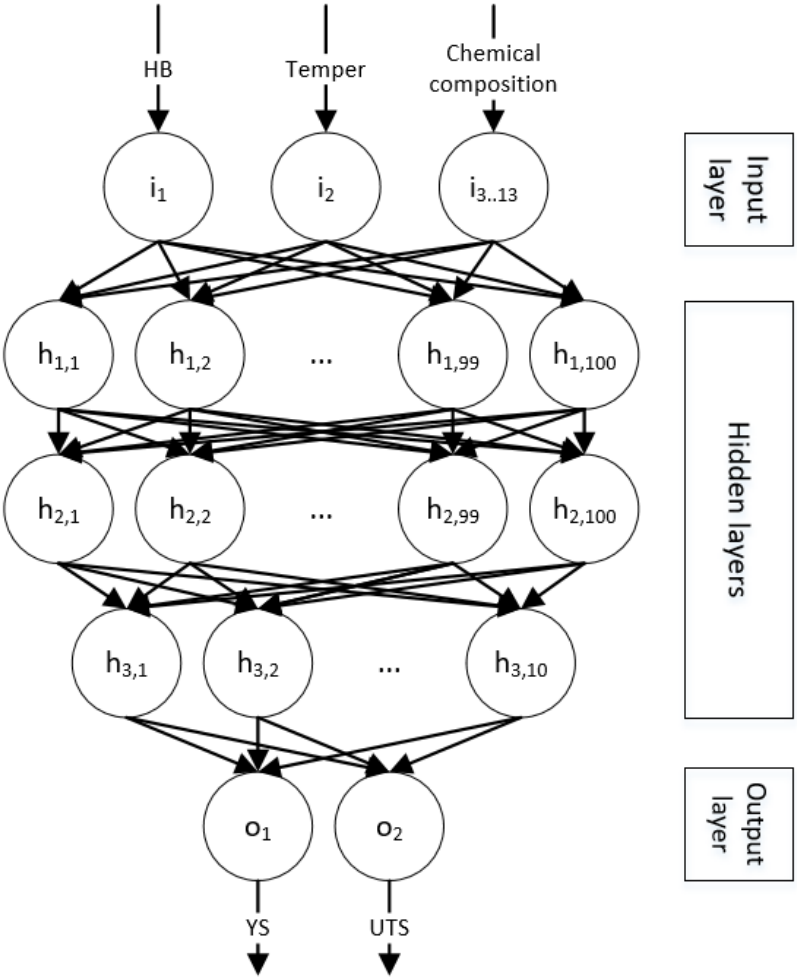


Figure 3. Artificial neural network model.

The topology of a neural network refers to the way in which perceptrons are connected and is a fundamental characteristic in the performance and learning of the network [41]. Each layer is shapeless in the sense that all of its perceptrons are equally important, have the same connections, and lack differentiators [74]. Only the network initialization process and the subsequent training will make its relevance change [75].

This topology is one of the most common because multilayer feedforward neural networks are, potentially, universal approximators [58], but even if a fully connected network can represent any function (including non-linear ones), backpropagation convergence is not guaranteed; therefore, it may not be able to learn some functions [56].

The number of hidden layers, the activation functions, and the number of neurons per layer are the main parameters of a neural network and must be defined before starting the training [54], but despite the fact that there have been great advances in this field [76,77], there is no formal method to optimize them [9].

The chosen network topology has more than 200 perceptrons linked by more than 20,000 connections and uses sigmoid-type activation function [56]. This topology is the outcome of continuous optimization steps intended to balance its learning ability and the needed resources for its training [78]. Note that a complex topology is able to learn more complex functions than a simple one, but it needs extra resources throughout its training: higher calculation capacity, additional time, and more input data [9,10].

2.3. Stage C: Training, Prediction, and Analysis

When the input dataset is ready and the neural network model has been correctly defined, it is time to start the training and prediction phase. Each of the two tensile properties that were taken into account in this study will be treated independently: yield strength (YS) and ultimate tensile strength (UTS).

The ANN executes 10 iterations of training and prediction for each of the two properties. Each of these iterations is fully independent of the others and is subdivided into four main steps:

- The input dataset is randomly split into two subsets, which comprise, respectively, 80% (training subset) and 20% (testing subset) of the records. Using disjoint groups to carry out these two steps (training and prediction) ensures that unwanted effects such as bias or overfitting will not occur [56].
- The ANN training with the training subset.
- The prediction of the properties of the testing subset.
- Results and data storage for further analysis.

Figure 4 shows a scheme of the steps of the training and prediction phase. Iterating the process 10 times allows better measuring the network performance because more precise statistical analyzes can be carried out.

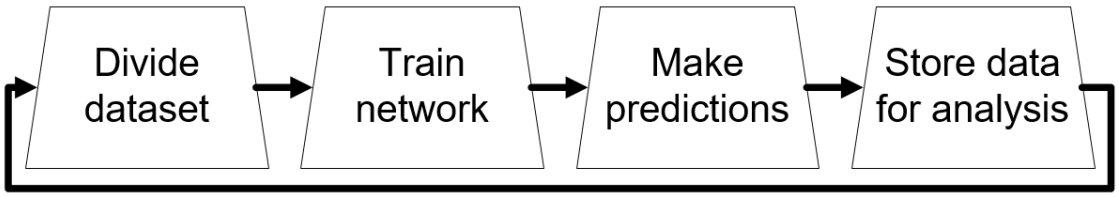


Figure 4. Training and prediction scheme.

The training is configured as follows [9]:

- Calculation of the learning rate for each parameter using adaptive moment estimation (ADAM) with $\beta_1 = 0.9$, $\beta_2 = 0.999$ (algorithm parameters), $\eta = 0.001$ (step size), and $\epsilon = 10^{-8}$ (stability factor) [79].
- Early stopping after 20 iterations without significant changes.
- Training stops when a training error of less than 0.1 is reached.
- Maximum of 100,000 training epochs to avoid infinite loops.

The prediction and training process generates a large amount of information that, after a detailed analysis, makes it possible to estimate the performance and capabilities of the methodology.

3. Results and Discussion

The outcomes of the training and prediction procedures are very stable and converge to similar results although they are randomly initialized. Once the ANN is trained with 80% of the registries (training subset), it is requested to make predictions about the remaining 20% of the data (testing subset). During the predictive step, no clues about the expected results are given to the network [56].

In this study, the ANN is trained with 570 randomly selected materials from the input dataset, and the remaining 143 are used to test the predictive capabilities of the ANN. Both subsets (training and testing) are randomly built for each of the 10 iterations, and so, they are fully independent.

Figure 5 shows the input data in a tridimensional scatter plot (blue dots are the real data and cyan ones the projections on the HB-YS and HB-UTS planes). The graph shows that, for the input dataset, the correlation between the three variables is mostly linear (which was already discussed in the Introduction); however, the fit of the data to the corresponding regression line is $R^2 \sim 0.8$, which is acceptable for a first approximation, but in most cases, causes excessive uncertainty about the predictions [9,35].

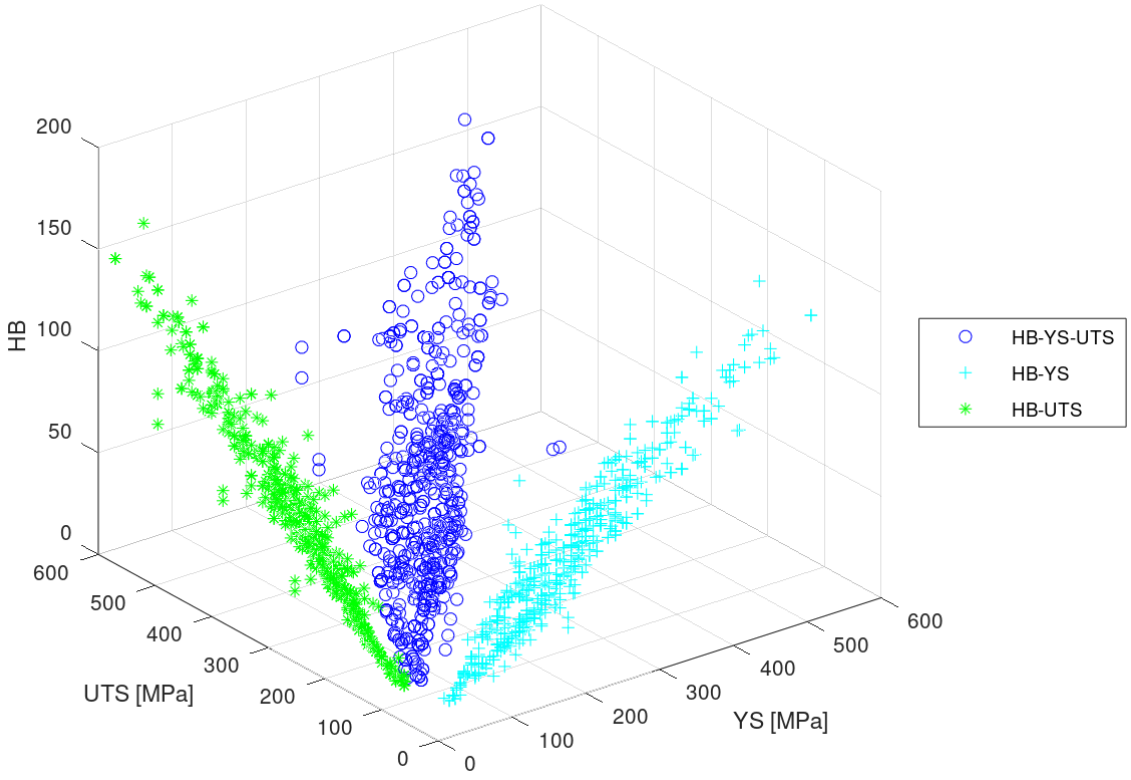


Figure 5. Scatter plot of the input dataset showing the relation HB-YS-UTS.

3.1. Yield Strength

Figure 6 shows a scatter plot that relates the yield stress and the Brinell hardness of the input registries. As already indicated, both properties show an approximately linear correlation although, in most situations, insufficiently precise to carry out approximations that may be useful [14]. The Pearson correlation coefficient of the related regression curve is $R^2 = 0.86$ [80], and its slope is $m = 2.96$.

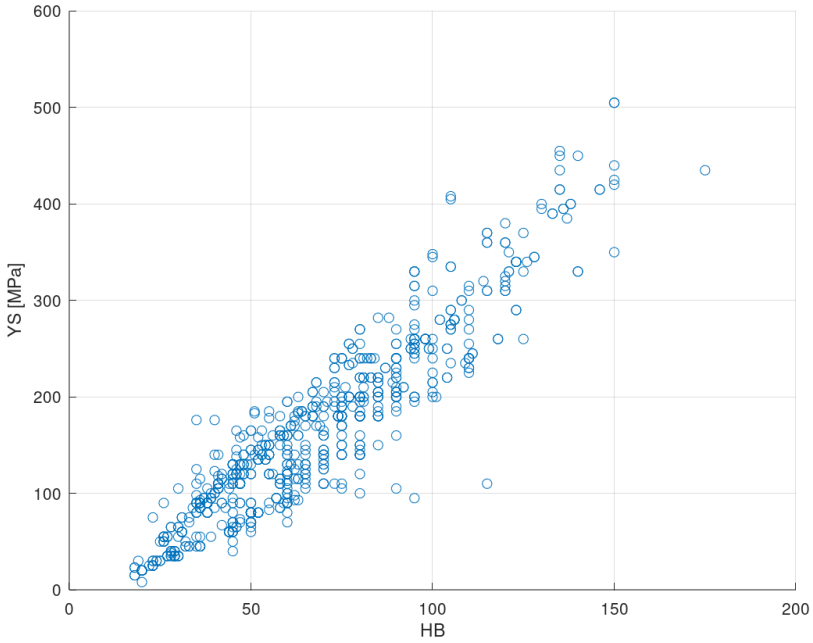


Figure 6. Yield strength-Brinell hardness scatter plot of the input dataset.

It is possible to observe that some scattered values appear that are notably far from the general trend. These data cause defects in the results since there is not enough sample for the ANN to correctly learn to predict them. However, it has been decided not to delete any records in order to obtain real metrics about the performance of the method.

Table 1 shows the results of the yield strength prediction phase for the 10 iterations that were carried out. The table shows, for each iteration, the averaged deviation ($\overline{x}$), the statistical standard deviation (S_x), the maximum deviation (Max), the median deviation (Median), and the trimmed averaged deviation at 90% ($\overline{x}_{90\%}$). Furthermore, the average of all iterations is displayed (Avg.).

Table 1. Prediction deviation (as %) of the yield strength.

Iteration	$\overline{x}$	S_x	Max	Median	$\overline{x}_{90\%}$
1	3.65%	2.48%	9.94%	3.00%	3.56%
2	3.29%	2.41%	13.51%	3.04%	3.12%
3	3.60%	2.78%	13.14%	3.11%	3.42%
4	3.67%	2.69%	14.39%	3.28%	3.51%
5	3.82%	2.45%	10.80%	3.17%	3.70%
6	3.81%	2.99%	12.80%	2.98%	3.66%
7	3.63%	2.48%	9.86%	3.23%	3.52%
8	3.81%	2.82%	12.59%	3.38%	3.68%
9	3.88%	2.58%	9.99%	3.45%	3.77%
10	3.53%	2.36%	9.90%	3.27%	3.42%
Avg.	3.67%	2.61%	11.69%	3.19%	3.54%

As can be seen, the averaged error of the yield strength prediction is 3.67%, with a standard deviation of 2.61% and a median of 3.19%. It can be observed that the results of the 10 iterations are very homogeneous, except in the case of the maximum deviations because these extreme values are usually associated with alloys that exhibit characteristics for which there is a scarce sample.

Figure 7 shows the results of the 10 iterations (1430 predictions) in the form of a box-and-whiskers plot. It is interesting to note that some outliers (out of 4.4σ interval, 98% confidence) appear that relate to the aforementioned extreme values. Even though these outliers reduce the global performance of the network, they allow knowing the capability of the methodology in the worst circumstances [10].

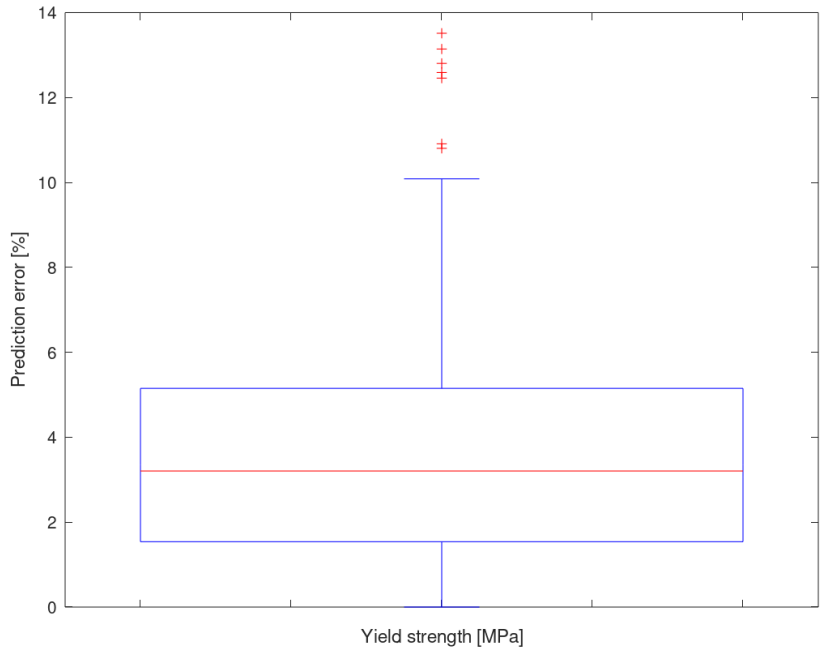


Figure 7. Prediction deviation of the yield strength for all iterations.

Figure 8 shows a histogram of the predictive deviations resulting from the 10 iterations. It can be seen that most of the yield strength predictions show an error of less than 4%, and few of them are greater than 10%.

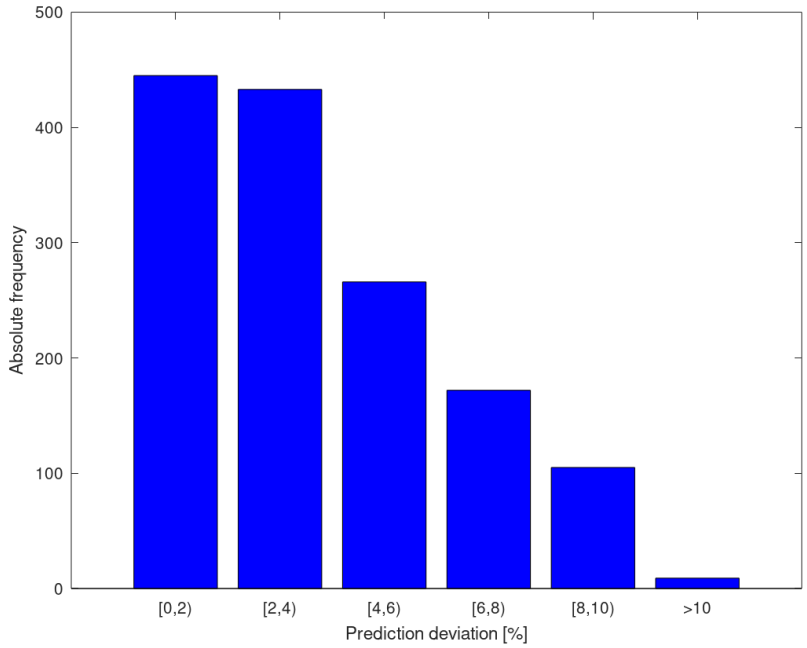


Figure 8. Histogram of the prediction error of the yield strength for all iterations.

Appendix A includes a deeper analysis of the input dataset and contains more information about the training process.

3.2. Ultimate Tensile Strength

Figure 9 shows the relationship between the ultimate tensile strength and the Brinell hardness of the input dataset. As already indicated, both properties show an approximately linear correlation [14].

The Pearson correlation coefficient of the related regression curve is $R^2 = 0.88$ [80], and its slope is $m = 3.28$.

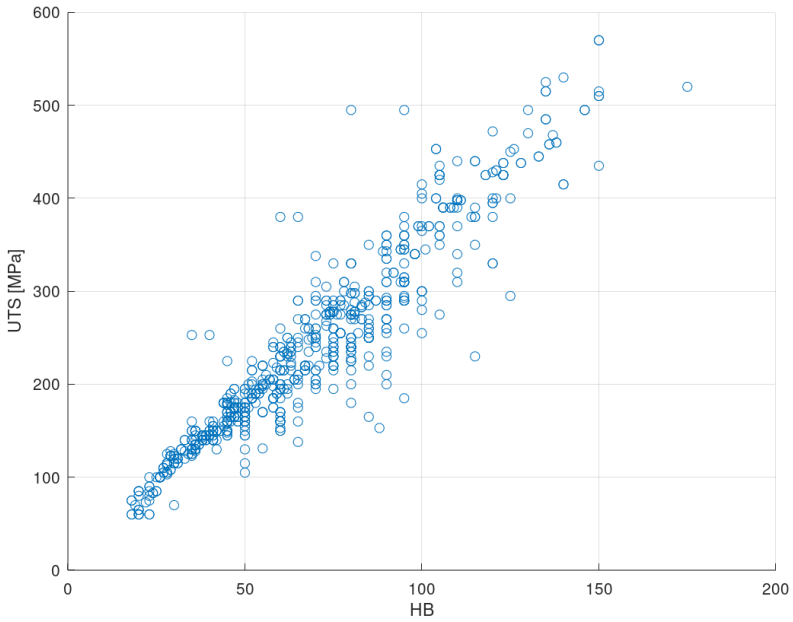


Figure 9. Ultimate tensile strength-Brinell hardness scatter plot of the input dataset.

There is a multitude of registries that deviate from the general trend. These materials constitute a challenge for the ANN and can cause the performance of the methodology to significantly drop. However, the possibility of eliminating these alloys from the input dataset should not be considered since these properties are not sufficiently aberrant to judge them as outliers.

Table 2 shows the results of the ultimate tensile strength prediction phase. The table shows, for each iteration, the averaged deviation ($\overline{x}$), the statistical standard deviation (S_x), the maximum deviation (Max), the median deviation (Median), and the trimmed averaged deviation at 90% ($\overline{x}_{90\%}$). Furthermore, the average of all iterations is displayed (Avg.).

Table 2. Prediction deviation (as %) of the ultimate tensile strength.

Iteration	$\overline{x}$	S_x	Max	Median	$\overline{x}_{90\%}$
1	4.26%	2.55%	10.97%	3.18%	4.11%
2	4.38%	2.54%	11.32%	4.03%	4.22%
3	3.98%	2.62%	14.24%	2.75%	3.76%
4	4.04%	2.64%	10.60%	2.92%	3.86%
5	4.38%	2.46%	11.30%	3.81%	4.23%
6	4.41%	2.87%	14.81%	3.40%	4.19%
7	4.32%	2.46%	11.32%	3.38%	4.13%
8	4.33%	2.95%	15.44%	3.10%	4.10%
9	4.22%	2.52%	11.29%	3.35%	4.02%
10	4.27%	2.39%	11.27%	3.62%	4.12%
Avg.	4.26%	2.60%	12.26%	3.35%	4.07%

As can be seen, the averaged error of the UTS prediction is 4.26%, with a standard deviation of 2.60% and a median of 3.35%. The predictive performance of the UTS is slightly lower than that of the YS, and its maximum error results are higher; however, the values of the standard deviation are very similar between both properties. The results of the 10 UTS prediction iterations are also very homogeneous.

Figure 10 shows a box-and-whiskers plot showing, together, the results of the 10 iterations (1430 predictions). The median is located at 3.35%, and there are some outliers associated with alloys with properties that are far from the global trend and with which the ANN has trouble making predictions.

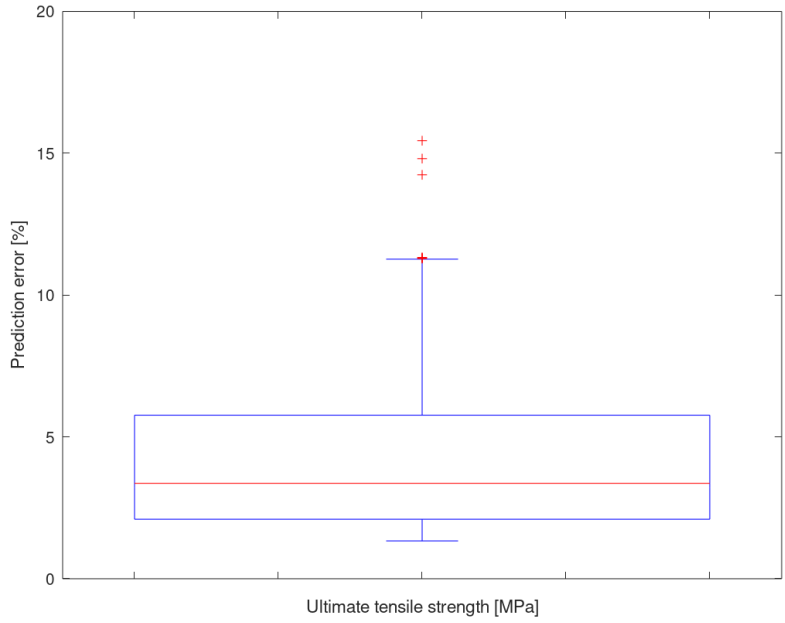


Figure 10. Prediction deviation of the ultimate tensile strength for all iterations.

Figure 11 shows a histogram with the predictive error of the 10 iterations. As can be seen, most of the predictions have an error of less than 4%, and few of them have values above 10%.

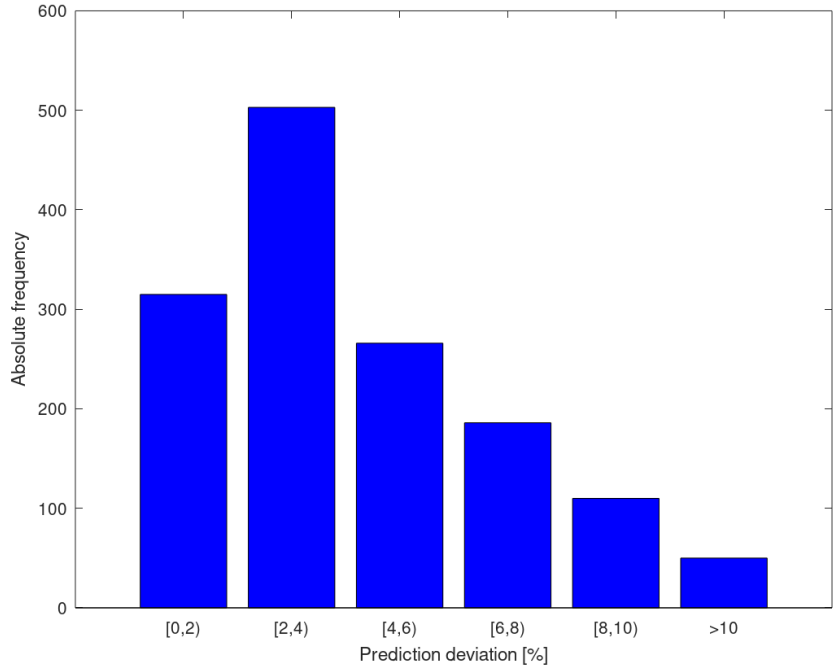


Figure 11. Histogram of the prediction error of the ultimate tensile strength for all iterations.

Appendix B includes a deeper analysis of the input dataset and contains more information about the training process.

3.3. Case Study: Application to Al 7010 Alloy

The material Al 7010 was selected because the performance of the ANN will be compared with the empirical formulas specifically developed for this alloy in [14]. In that study, the authors developed two equations that estimate the YS and the UTS as a function of the result of the hardness test, obtaining very good approximations [14].

Al 7010 is a 7000-series aluminum alloy whose main alloying elements are zinc, magnesium, and copper. Table 3 shows its chemical composition [12,73].

Table 3. Al 7010-T6's chemical composition.

Element	Content
Al	Remainder
Zn	5.7–6.7%
Mg	2.1–2.6%
Cu	1.5–2.0%
Zr	0.10–0.16%
Fe	≤0.15%
Si	≤0.12%
Mn	≤0.10%
Ti	≤0.06%
Cr	≤0.05%
Ni	≤0.05%
Other	≤0.15%

Table 4 contains the values of the relevant properties to carry out this discussion [32]. Only the information about one temple (T6) is available [12].

Table 4. Al 7010-T6's aluminum alloy mechanical properties.

Property	Value
YS (MPa)	530
UTS (MPa)	590
HB (kp/mm^2)	190

Tiryakioğlu et al. [14] developed two equations to approximate the values of YS and UTS (see Equations (4) and (5)) (the expressions require that the value of HB was expressed in units of MPa instead of the more frequently used kp/mm^2).

$$YS = 0.383 \cdot HB - 182.3 \tag{4}$$

$$UTS = 0.247 \cdot HB + 113.1 \tag{5}$$

Table 5 shows the predicted values of both tensile properties and the deviation with respect to the real values. As can be seen, the predictive error of this approximation is very small (0.16% for the YS and −2.88% for the UTS) due to the fact that they are empirical formulas developed specifically for this alloy. It is important to point out that these results are better than the ones reported by the authors (∼0.95) [40].

Table 5. Predicted values of the YS and UTS for Al 7010 using the empirical formulas.

Property	Actual Value (MPa)	Predicted Value (MPa)	Deviation
YS	530	530.85	0.16%
UTS	590	573.01	−2.88%

Table 6 contains the results of the prediction of the tensile properties using the ANN: averaged predicted value and deviation. As can be seen, these results show a greater error than those of the empirical equations. However, the predictive error is limited and is consistent with the previous analysis of the methodology.

Table 6. Predicted values of the YS and UTS for Al-7010 using the ANN.

Property	Actual Value (MPa)	Predicted Value (MPa)	Deviation
YS	530	519.12	−2.05%
UTS	590	564.69	−4.29%

Table 7 shows some statistical metrics about the predictive process of the ANN. To obtain these results, ten training-prediction iterations (fully independent) were launched, and to avoid bias and overfitting, the references to Al 7010 that appeared in the input dataset were eliminated. As can be seen in the table, the ANN tends to underestimate the value of both properties. This phenomenon may be due to the fact that the alloys of the 7000 series have particularly high tensile properties.

Table 7. Statistical metrics about the predictive iterations of the ANN.

Property	Actual Value	$\bar{x}$	S_x	Median	Max	Min
YS (MPa)	530	519.12	4.18	519.00	525.30	514.10
UTS (MPa)	590	564.69	4.95	562.95	572.55	559.55

Figure 12 shows a box-and-whiskers plot showing the results of the 10 predictive iterations. It can be seen that the results are very homogeneous and show few variations (the value of the standard deviation is low).

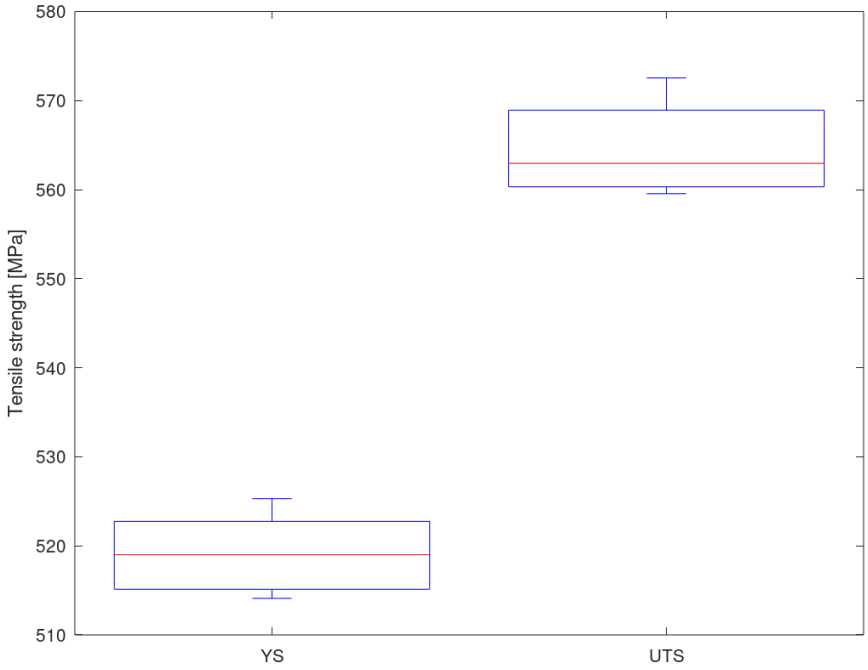


Figure 12. Yield strength and ultimate tensile strength prediction for Al 7010-T6 using the ANN.

Table 8 summarizes the predictive deviations of both methods to ease comparison. As can be seen in the table, the performance of the empirical equations is greater than that of the ANN. It should be taken into account that the equations have been developed expressly for this particular alloy; however, the methodology depicted in this paper is able to make predictions about any aluminum

alloy (it is more general, but less precise). Anyway, the ANN error is only ∼2% greater than that of the empirical equations.

Table 8. Comparison between the predictive precision of the empirical equations and the ANN.

Property	Equations Deviation	ANN Deviation
YS	0.16%	−2.05%
UTS	−2.88%	−4.29%

4. Conclusions and Future Work

This paper predicts, by means of an ANN, the YS and the UTS of aluminum alloys taking as the input their chemical composition, their tempers, and the results of a hardness test. It is a contribution of great industrial interest as it allows knowing the mechanical behavior based on an almost non-destructive test (hardness test) and parameters specified in standards and scientific databases. In addition to its multiple applications in the selection of the suitability of materials for a specific application and evaluation of their ability to be manufactured by forming processes, this methodology can be an interesting tool to help in the prediction of the possible degradation of properties or the loss of mechanical integrity that could cause failure or breakage. Therefore, the main conclusions of this work are presented as follows:

- ANN can be employed to predict the YS and the UTS of aluminum alloys based on its hardness, chemical composition, and temper. In this study, the average deviations for both properties are, respectively, 3.67% and 4.26%.
- The results of a methodology based on artificial intelligence can achieve a similar performance to that obtained through empirical equations. In the example shown in this study, for Al 7010-T6, the resulting difference between both methods was ∼2%.
- Although the predictive performance of this methodology is slightly lower than that of empirical equations, it provides greater generality since it can make predictions about any aluminum alloy.
- A multilayer ANN can be trained to make predictions about the mechanical behavior of an in-service industrial component on which a hardness test can be performed.

This study shows that it is possible to predict the tensile properties of aluminum alloys using AI based techniques. In the same way, it opens the door to investigate similar solutions applied to other metallic alloys such as steel or, even, to try to apply the same methodology to ceramics.

ANNs have proven to be a powerful tool for predicting the tensile properties of highly relevant industrial materials without the need for costly and complicated stress-strain tests. It can be of interest to study whether it is possible to design an artificial intelligence based system capable of predicting these properties using the results of other widespread hardness tests such as Rockwell.

Once this working scheme has shown that it can be used to make adequate predictions, other network architectures can be tested to improve the overall performance of the methodology. There is a wide variety of network topologies that meet diverse needs and solve different problems [56].

Author Contributions: Conceptualization, D.M., A.R.-P., and A.M.C.; data curation, D.M. and A.R.-P.; formal analysis, D.M.; funding acquisition, A.M.C.; investigation, D.M.; methodology, D.M.; project administration, A.M.C.; resources, A.R.-P. and A.M.C.; software, D.M.; supervision, A.R.-P. and A.M.C.; validation, D.M., A.R.-P., and A.M.C.; visualization, D.M.; writing, original draft, D.M.; writing, review & editing, D.M., A.R.-P., and A.M.C. All authors read and agreed to the published version of the manuscript.

Funding: This work was developed within the framework of the Doctorate Program in Industrial Technologies of the UNED and was funded by the Annual Grants Call of the E.T.S.I.I. of the UNED via the project References 2020-ICF04/B and 2020-ICF04/D.

Acknowledgments: We extend our acknowledgments to the Research Group of the UNED Industrial Production and Manufacturing Engineering (IPME) and the Industrial Research Group "Advanced Failure Prognosis for Engineering Applications". We also thank Matmatch GmbH for freely supplying all the material data employed to accomplish this study.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations and symbols are used in this manuscript:

ADAM	Adaptive moment estimation
AI	Artificial intelligence
ANN	Artificial neural network
β_n	ADAM algorithm parameter
ϵ	ADAM stability factor
η	ADAM step size
HB	Hardness Brinell
m	ADAM first moment estimate
UTS	Ultimate tensile strength
S_x	Standard deviation
v	ADAM second moment estimate
$\overline{x}$	Average
$\overline{x}_{90\%}$	Trimmed mean at 90%
YS	Yield stress

Appendix A. Yield Strength

Figure A1 shows a histogram of the yield strength values of the input dataset, which are quite disperse even if it exhibits a concentration of registries around 125 MPa. The yield strength strongly depends on the chemical composition of the alloy and the treatment applied to it [9].

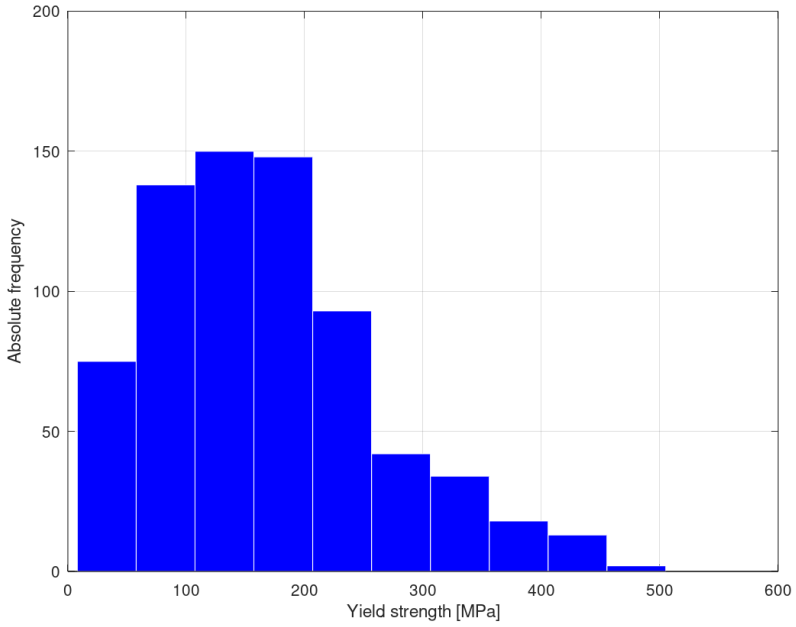


Figure A1. Yield strength histogram of the input dataset.

The location of the yield point is based on conventions (typically a deviance of 0.2% from the elastic linear behavior) because, in fact, it is not related to any substantial physical phenomenon [81]. Consequently, it is a property for which there is frequently significant uncertainty even in the bibliography (these data are commonly given as a range of values) [82].

Figure A2 shows the averaged curve of the evolution of the error function for the ANN training with the yield strength data. This is the mean curve resulting from the 10 training iterations that were carried out. As can be seen, the error function starts at values close to 18,000 and, after 23,000 epochs, descends until it reaches a value of about 100, at which point, the training process stops due to a no-improvement condition. It is interesting to note that, around the epoch of 18,000, a step appears; these are usually linked to moments in which the ANN has learned a relevant concept [9,54].

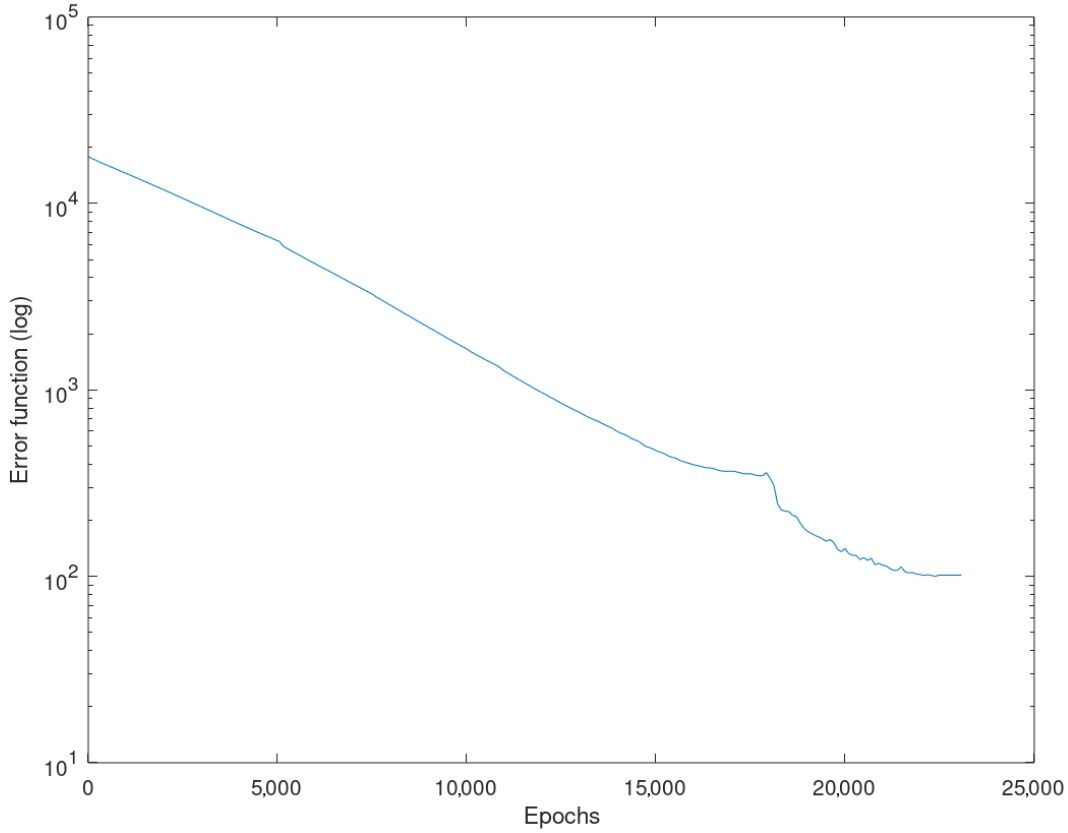


Figure A2. Error function averaged evolution curve for the training on the yield strength data (logarithmic scale).

During training, the error function descends until it asymptotically stabilizes. At that moment, the network has learned everything that is possible from the input dataset, and it is not advisable to persist in training since bias can be generated [41,54].

Appendix B. Ultimate Tensile Strength

Aluminum is an example of a ductile material because it can withstand significant plastic deformation before breaking [83,84]. After the yield point, the slope of the aluminum stress-strain curve falls to become zero at the UTS point [9,70].

Figure A3 shows a histogram representing the ultimate tensile strengths of the records contained in the input dataset. As can be seen, these alloys have a wide range of values, although most of the materials accumulate around 175 MPa. This property is highly dependent on the chemical composition of the alloy and the treatments to which it has been subjected [9].

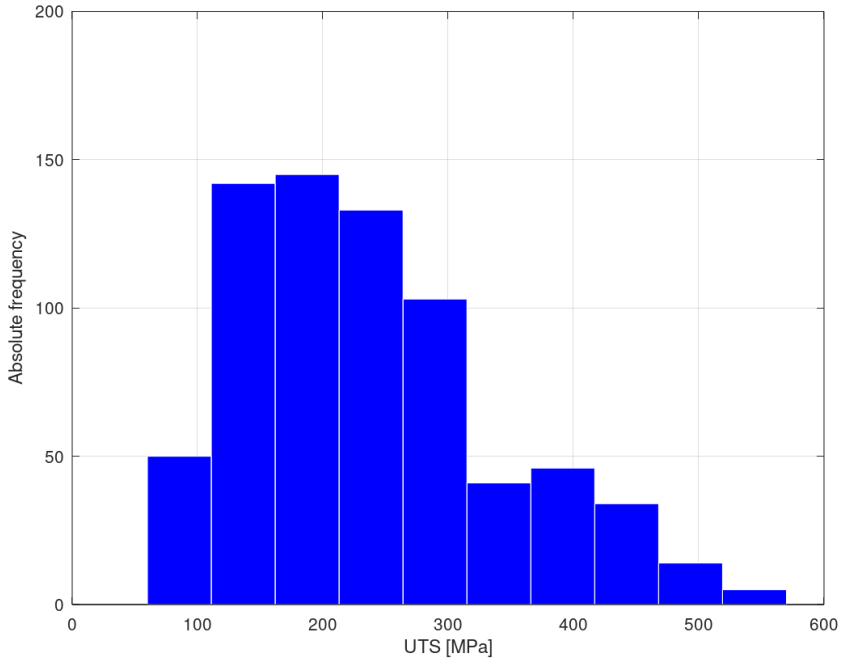


Figure A3. Ultimate tensile strength histogram of the input dataset.

Figure A4 shows the averaged curve resulting from the 10 training iterations that were accomplished with the ultimate tensile strength data. This is the averaged curve of the evolution of the error function, which starts at values close to 33,000 and, after 24,000 epochs, reaches a value of approximately 50, at which point, the training process is stopped due to a non-improvement condition.

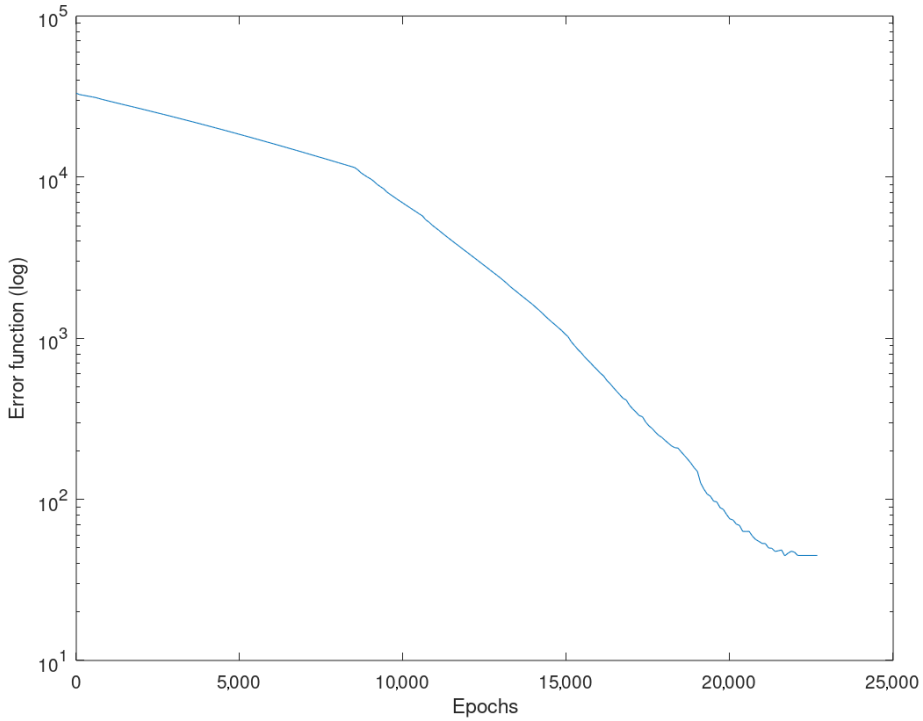


Figure A4. Error function averaged evolution curve for the training on the ultimate tensile strength data (logarithmic scale).

References

1. Camacho, A.M.; Rodríguez-Prieto, A.; Herrero, J.M.; Aragón, A.M.; Bernal, C.; Lorenzo-Martin, C.; Yanguas-Gil, Á.; Martins, P.A. An Experimental and Numerical Analysis of the Compression of Bimetallic Cylinders. *Materials* **2019**, *12*, 4094. [CrossRef] [PubMed]
2. Regueras, J.M.G.; Camacho, A.M. Investigations on the influence of blank thickness (t) and length/wide punch ratio (LD) in rectangular deep drawing of dual-phase steels. *Comput. Mater. Sci.* **2014**, *91*, 134–145. [CrossRef]
3. Rubio, E.M.; Camacho, A.M.; Pérez, R.; Marín, M.M. Guidelines for selecting plugs used in thin-walled tube drawing processes of metallic alloys. *Metals* **2017**, *7*, 572. [CrossRef]
4. Tu, S.; Ren, X.; He, J.; Zhang, Z. Stress–strain curves of metallic materials and post-necking strain hardening characterization: A review. *Fatigue Fract. Eng. Mater. Struct.* **2020**, *43*, 3–19. [CrossRef]
5. Danylenko, M. Aluminium alloys in aerospace. *Alum. Int. Today* **2018**, *31*, 35.
6. Rojas, J.I.; Crespo, D. Onset frequency of fatigue effects in pure aluminum and 7075 (AlZnMg) and 2024 (AlCuMg) alloys. *Metals* **2016**, *6*, 50. [CrossRef]
7. Lin, Y.; Dong, W.Y.; Zhu, X.H.; Wu, Q.; He, Y.J. Deformation Behavior and Precipitation Features in a Stretched Al–Cu Alloy at Intermediate Temperatures. *Materials* **2020**, *13*, 2495. [CrossRef]
8. Ashkenazi, D. How aluminum changed the world: A metallurgical revolution through technological and cultural perspectives. *Technol. Forecast. Soc. Chang.* **2019**, *143*, 101–113. [CrossRef]
9. Merayo Fernández, D.; Rodríguez-Prieto, A.; Camacho, A.M. Prediction of the Bilinear Stress-Strain Curve of Aluminum Alloys Using Artificial Intelligence and Big Data. *Metals* **2020**, *10*, 904. [CrossRef]
10. Merayo, D.; Rodríguez-Prieto, A.; Camacho, A. Prediction of Physical and Mechanical Properties for Metallic Materials Selection Using Big Data and Artificial Neural Networks. *IEEE Access* **2020**, *8*, 13444–13456. [CrossRef]
11. Tabor, D. *The Hardness of Metals*; Oxford University Press: Oxford, UK, 2000.
12. Kaufman, J.G. *Introduction to Aluminum Alloys and Tempers*; ASM International: Almere, The Netherlands, 2000.
13. Tiryakioğlu, M.; Campbell, J.; Staley, J.T. On macrohardness testing of Al–7 wt.% Si–Mg alloys: II. An evaluation of models for hardness–yield strength relationships. *Mater. Sci. Eng. A* **2003**, *361*, 240–248. [CrossRef]
14. Tiryakioğlu, M.; Robinson, J.; Salazar-Guapuriche, M.; Zhao, Y.; Eason, P. Hardness–strength relationships in the aluminum alloy 7010. *Mater. Sci. Eng. A* **2015**, *631*, 196–200. [CrossRef]
15. Li, Y.; Yang, Z. Evaluation of Measurement Uncertainty for Ultrasonic Contact Impedance Hardness Testing Results. In *Physical Testing and Chemical Analysis (Part A: Physical Testing)*; Shanghai Research Institute of Materials and Mechanical Engineering: Shanghai, China, 2017; p. 7.
16. DIN. *DIN 50159-1—Metallic Materials—Hardness Testing with the UCI Method—Part 1: Test Method*; Deutsches Institut für Normung: Berlin, Germany, 2015.
17. The Aluminum Association. *Designations and Chemical Composition Limits for Aluminum Alloys in the Form of Castings and Ingot*; The Aluminum Association: Arlington, VA, USA, 2006.
18. Rodríguez-Prieto, A.; Camacho, A.M.; Sebastián, M.A. Materials selection criteria for nuclear power applications: A decision algorithm. *JOM* **2016**, *68*, 496–506. [CrossRef]
19. Dimiduk, D.M.; Holm, E.A.; Niezgoda, S.R. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. *Integr. Mater. Manuf. Innov.* **2018**, *7*, 157–172. [CrossRef]
20. Liu, Y.; Zhao, T.; Ju, W.; Shi, S. Materials discovery and design using machine learning. *J. Mater.* **2017**, *3*, 159–177. [CrossRef]
21. Ponnuvel, S.; Senthilkumar, N. A study on machinability evaluation of Al-Gr-B4C MMC using response surface methodology-based desirability analysis and artificial neural network technique. *Int. J. Rapid Manuf.* **2019**, *8*, 95–122. [CrossRef]
22. Shi, S.; Gao, J.; Liu, Y.; Zhao, Y.; Wu, Q.; Ju, W.; Ouyang, C.; Xiao, R. Multi-scale computation methods: Their applications in lithium-ion battery research and development. *Chin. Phys. B* **2015**, *25*, 018212. [CrossRef]
23. Zhou, T.; Song, Z.; Sundmacher, K. Big Data Creates New Opportunities for Materials Research: A Review on Methods and Applications of Machine Learning for Materials Design. *Engineering* **2019**, *5*, 1017–1026. [CrossRef]

24. Schmidt, J.; Marques, M.R.; Botti, S.; Marques, M.A. Recent advances and applications of machine learning in solid-state materials science. *NPJ Comput. Mater.* **2019**, *5*, 1–36.
25. Ling, J.; Antono, E.; Bajaj, S.; Paradiso, S.; Hutchinson, M.; Meredig, B.; Gibbons, B.M. Machine learning for alloy composition and process optimization. In Proceedings of the ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition, Oslo, Norway, 11–15 June 2018; American Society of Mechanical Engineers Digital Collection: New York City, NY, USA, 2018.
26. Mrówka-Nowotnik, G.; Sieniawski, J. Influence of heat treatment on the microstructure and mechanical properties of 6005 and 6082 aluminium alloys. *J. Mater. Process. Technol.* **2005**, *162*, 367–372. [CrossRef]
27. Barajas, C.; de Vicente, J.; Caja, J.; Maresca, P.; Gómez, E. Considerations to the hardness Brinell measurement using optical equipment. *Procedia Manuf.* **2017**, *13*, 550–557. [CrossRef]
28. ISO. *ISO 6507-1 Metallic Materials—Vickers Hardness Test—Part 1 Test Method*; International Organization for Standardization: Geneva, Switzerland, 2005.
29. ASTM. *ASTM E10-18, Standard Test Method for Brinell Hardness of Metallic Materials*; ASTM International: West Conshohocken, PA, USA, 2018.
30. ASTM. *ASTM E140-12B(2019)e1—Standard Hardness Conversion Tables for Metals Relationship among Brinell Hardness, Vickers Hardness, Rockwell Hardness, Superficial Hardness, Knoop Hardness, Scleroscope Hardness, and Leeb Hardness*; ASTM International: West Conshohocken, PA, USA, 2019.
31. Li, H.; Hu, Z.; Chen, Y.; Sun, Q.; Zhou, X. Modeling mechanical properties and plastic strain for hot forming-quenching AA6061 aluminum alloy parts. *Int. J. Lightweight Mater. Manuf.* **2020**, *3*, 66–72. [CrossRef]
32. Matmatch GmbH. Matmatch. 2020. Available online: https://www.matmatch.com/ (accessed on 15 April 2020).
33. Hill, R.; Storåkers, B.; Zdunek, A. A theoretical study of the Brinell hardness test. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **1989**, *423*, 301–330.
34. Chang, S.; Jahn, M.; Wan, C.; Lee, J.; Hsu, T. The determination of tensile properties from hardness measurements for Al-Zn-Mg alloys. *J. Mater. Sci.* **1976**, *11*, 623–630. [CrossRef]
35. Pavlina, E.; Van Tyne, C. Correlation of yield strength and tensile strength with hardness for steels. *J. Mater. Eng. Perform.* **2008**, *17*, 888–893. [CrossRef]
36. Cahoon, J.; Broughton, W.; Kutzak, A. The determination of yield strength from hardness measurements. *Metall. Trans.* **1971**, *2*, 1979–1983.
37. Cahoon, J. An improved equation relating hardness to ultimate strength. *Metall. Mater. Trans. B* **1972**, *3*, 3040. [CrossRef]
38. Zhang, P.; Li, S.; Zhang, Z. General relationship between strength and hardness. *Mater. Sci. Eng. A* **2011**, *529*, 62–73. [CrossRef]
39. Salazar-Guapuriche, M.A.; Zhao, Y.; Pitman, A.; Greene, A. *Correlation of Strength with Hardness and Electrical Conductivity for Aluminium Alloy 7010*; Materials Science Forum; Trans Tech Publ: Stafa-Zurich, Switzerland, 2006; Volume 519, pp. 853–858.
40. Tiryakioğlu, M.; Robinson, J.S. On the representative strain in Vickers hardness testing of 7010 aluminum alloy. *Mater. Sci. Eng. A* **2015**, *641*, 231–236. [CrossRef]
41. Merayo, D.; Rodriguez-Prieto, A.; Camacho, A. Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0. *Procedia Manuf.* **2019**, *41*, 42–49. [CrossRef]
42. Callister, W.D.; Rethwisch, D.G. *Materials Science and Engineering: An Introduction*; Wiley: New York, NY, USA, 2018; Volume 9.
43. Davis, J.; Committee, A.I.H. *Metals Handbook Desk Edition*, 2nd ed.; 75th Anniversary ASM Handbooks; Taylor & Francis: Geauga County, OH, USA, 1998.
44. Jenab, A.; Sarraf, I.S.; Green, D.E.; Rahmaan, T.; Worswick, M.J. The use of genetic algorithm and neural network to predict rate-dependent tensile flow behaviour of AA5182-O sheets. *Mater. Des.* **2016**, *94*, 262–273. [CrossRef]
45. McCarthy, J.; Minsky, M.L.; Rochester, N.; Shannon, C.E. A proposal for the Dartmouth summer research project on artificial intelligence, august 31, 1955. *AI Mag.* **2006**, *27*, 12.
46. Liu, Y.; Gopalakrishnan, V. An overview and evaluation of recent machine learning imputation methods using cardiac imaging data. *Data* **2017**, *2*, 8. [CrossRef] [PubMed]

47. Rahmanpanah, H.; Mouloodi, S.; Burvill, C.; Gohari, S.; Davies, H.M. Prediction of load-displacement curve in a complex structure using artificial neural networks: A study on a long bone. *Int. J. Eng. Sci.* **2020**, *154*, 103319. [CrossRef]
48. Cummings, M. *Artificial Intelligence and the Future of Warfare*; Chatham House for the Royal Institute of International Affairs London: London, UK, 2017.
49. Villa, F.; Ceroni, M.; Bagstad, K.; Johnson, G.; Krivov, S. ARIES (Artificial Intelligence for Ecosystem Services): A new tool for ecosystem services assessment, planning, and valuation. In Proceedings of the 11th Annual BIOECON Conference on Economic Instruments to Enhance the Conservation and Sustainable Use of Biodiversity, Veneto, Italy, 21–22 September 2009; pp. 21–22.
50. Allen, G.; Chan, T. *Artificial Intelligence and National Security*; Belfer Center for Science and International Affairs: Cambridge, MA, USA, 2017.
51. Ee, J.H.; Huh, N. A study on the relationship between artificial intelligence and change in mathematics education. *Commun. Math. Educ.* **2018**, *32*, 23–36.
52. Galán-Freyle, N.J.; Ospina-Castro, M.L.; Medina-González, A.R.; Villarreal-González, R.; Hernández-Rivera, S.P.; Pacheco-Londoño, L.C. Artificial intelligence assisted Mid-infrared laser spectroscopy in situ detection of petroleum in soils. *Appl. Sci.* **2020**, *10*, 1319. [CrossRef]
53. Thankachan, T.; Prakash, K.S.; Pleass, C.D.; Rammasamy, D.; Prabakaran, B.; Jothi, S. Artificial neural network to predict the degraded mechanical properties of metallic materials due to the presence of hydrogen. *Int. J. Hydrog. Energy* **2017**, *42*, 28612–28621. [CrossRef]
54. Jackson, P.C. *Introduction to Artificial Intelligence*; Courier Dover Publications: Mineola, NY, USA, 2019.
55. Yamanaka, A.; Kamijyo, R.; Koenuma, K.; Watanabe, I.; Kuwabara, T. Deep neural network approach to estimate biaxial stress-strain curves of sheet metals. *Mater. Des.* **2020**, *195*, 108970. [CrossRef]
56. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2015**, *61*, 85–117. [CrossRef]
57. Helal, S. The expanding frontier of artificial intelligence. *Computer* **2018**, *51*, 14–17. [CrossRef]
58. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251–257. [CrossRef]
59. Huang, G.; Huang, G.B.; Song, S.; You, K. Trends in extreme learning machines: A review. *Neural Netw.* **2015**, *61*, 32–48. [CrossRef] [PubMed]
60. Anysz, H.; Brzozowski, Ł.; Kretowicz, W.; Narloch, P. Feature Importance of Stabilised Rammed Earth Components Affecting the Compressive Strength Calculated with Explainable Artificial Intelligence Tools. *Materials* **2020**, *13*, 2317. [CrossRef] [PubMed]
61. Merayo, D.; Rodríguez-Prieto, A.; Camacho, A.M. Prediction of the Yield Stress of Aluminum Alloys using Big Data and Artificial Neural Networks. In Proceedings of the 3rd International Conference on Material Design and Applications, MDA2020, Porto, Portugal, 5–6 November 2020.
62. De Albuquerque, V.H.C.; de Alexandria, A.R.; Cortez, P.C.; Tavares, J.M.R. Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images. *NDT E Int.* **2009**, *42*, 644–651. [CrossRef]
63. Baldo, N.; Manthos, E.; Miani, M. Stiffness modulus and marshall parameters of hot mix asphalts: Laboratory data modeling by artificial neural networks characterized by cross-validation. *Appl. Sci.* **2019**, *9*, 3502. [CrossRef]
64. Wei, R.; Bi, Y. Research on Recognition Technology of Aluminum Profile Surface Defects Based on Deep Learning. *Materials* **2019**, *12*, 1681. [CrossRef]
65. Chokshi, P.; Dashwood, R.; Hughes, D.J. Artificial Neural Network (ANN) based microstructural prediction model for 22MnB5 boron steel during tailored hot stamping. *Comput. Struct.* **2017**, *190*, 162–172. [CrossRef]
66. Liu, G.; Jia, L.; Kong, B.; Guan, K.; Zhang, H. Artificial neural network application to study quantitative relationship between silicide and fracture toughness of Nb-Si alloys. *Mater. Des.* **2017**, *129*, 210–218. [CrossRef]
67. Li, Y.; Yu, B.; Wang, B.; Lee, T.H.; Banu, M. Online quality inspection of ultrasonic composite welding by combining artificial intelligence technologies with welding process signatures. *Mater. Des.* **2020**, *194*, 108912. [CrossRef]
68. Weinbub, J.; Wastl, M.; Rupp, K.; Rudolf, F.; Selberherr, S. ViennaMaterials–A dedicated material library for computational science and engineering. *Appl. Math. Comput.* **2015**, *267*, 282–293. [CrossRef]
69. Batra, S. Big data analytics and its reflections on DIKW hierarchy. *Rev. Manag.* **2014**, *4*, 5.

70. ASM. *Atlas of Stress-Strain Curves*; ASM: Almere, The Netherlands, 2002.
71. Mazzolani, F. EN1999 Eurocode 9: Design of aluminium structures. In *Proceedings of the Institution of Civil Engineers-Civil Engineering*; Thomas Telford Ltd.: Brussels, Belgium, 2001; Volume 144, pp. 61–64.
72. Davis, J.R. *Alloying: Understanding the Basics*; ASM International: Almere, The Netherlands, 2001.
73. The Aluminum Association. *International Alloy Designations and Chemical Composition Limits for Wrought Aluminum and Wrought Aluminum Alloys*; The Aluminum Association: Arlington, VA, USA, 2015.
74. Joshi, P. *Artificial Intelligence with Python*; Packt Publishing Ltd.: Birmingham, UK, 2017.
75. Teng, S.; Chen, G.; Liu, G.; Lv, J.; Cui, F. Modal strain energy-based structural damage detection using convolutional neural networks. *Appl. Sci.* **2019**, *9*, 3376. [CrossRef]
76. Walczak, S.; Cerpa, N. Heuristic principles for the design of artificial neural networks. *Inf. Softw. Technol.* **1999**, *41*, 107–117. [CrossRef]
77. Stier, J.; Gianini, G.; Granitzer, M.; Ziegler, K. Analysing neural network topologies: A game theoretic approach. *Procedia Comput. Sci.* **2018**, *126*, 234–243. [CrossRef]
78. Deshpande, A.; Kumar, M. *Artificial Intelligence for Big Data: Complete Guide to Automating Big Data Solutions Using Artificial Intelligence Techniques*; Packt Publishing Ltd.: Birmingham, UK, 2018.
79. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference for Learning Representations (ICLR 15), San Diego, CA, USA, 7–9 May 2015.
80. Benesty, J.; Chen, J.; Huang, Y. On the importance of the Pearson correlation coefficient in noise reduction. *IEEE Trans. Audio Speech Lang. Process.* **2008**, *16*, 757–765. [CrossRef]
81. Christensen, R.M. *The Theory of Materials Failure*; Oxford University Press: Oxford, UK, 2013.
82. ASM International Handbook Committee. *Properties and Selection: Nonferrous Alloys and Special-Purpose Materials Volume 2*; ASM International: Almere, The Netherlands, 2010.
83. Kamaya, M.; Kawakubo, M. A procedure for determining the true stress–strain curve over a large range of strains using digital image correlation and finite element analysis. *Mech. Mater.* **2011**, *43*, 243–253. [CrossRef]
84. Alam, T.; Ansari, A.H. Review on Aluminium and Its Alloys for automotive applications. *Int. J. Adv. Technol. Eng. Sci.* **2017**, *5*, 278–294.

Capítulo 5. Otras aportaciones científicas derivadas de la Tesis Doctoral

5.1 Contribuciones en congresos internacionales

Congreso: **2nd International Conference on Materials Design and Applications (MDA2018)**
Entidad organizadora: Faculdade de Engenharia da Universidade do Porto (FEUP)
Carácter: Internacional
Lugar celebración: Oporto (Portugal)
Fecha: 5-6 de julio de 2018
Título: Analytical and Numerical Study for Selecting Polymeric Matrix Composites Intended to Demanding Nuclear Applications
Autores: David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Tipo de participación: comunicación oral
Publicación: Book of abstracts, ISBN 978-989-8927-09-5, 57, 2018 (ver Apéndice F)

Congreso: **The 8th Manufacturing Engineering Society International Conference (MESIC'19)**
Entidad organizadora: Sociedad de Ingeniería de Fabricación (SIF) y Universidad Politécnica de Madrid
Carácter: Internacional
Lugar celebración: Madrid
Fecha: 19-21 de junio de 2019
Título: Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0
Autores: David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Tipo de participación: póster
Publicación: Procedia Manufacturing, ISSN 2351-9789 (ver Apéndice G)
doi: 10.1016/j.promfg.2019.07.027

Congreso: **3rd International Conference on Materials Design and Applications (MDA2020)**
Entidad organizadora: Faculdade de Engenharia da Universidade do Porto (FEUP)
Carácter: Internacional
Lugar celebración: Oporto (Portugal)
Fecha: 5-6 de noviembre de 2020
Título: Prediction of the yield stress of aluminum alloys using Big Data and Artificial Neural Networks
Tipo de participación: comunicación oral
Autores: David Merayo; Álvaro Rodríguez-Prieto; Ana María Camacho
Publicación: Book of abstracts, ISBN 978-989-9017-41-2, 103, 2020 (ver Apéndice H)

Capítulo 6. Conclusiones y desarrollos futuros

En esta Tesis Doctoral se ha examinado la capacidad de las técnicas de simulación numérica y basadas en inteligencia artificial y redes neuronales artificiales para servir como herramientas de apoyo a la decisión en el ámbito de la tecnología de materiales y su procesado. Este trabajo aborda la problemática y la factibilidad de encontrar nuevas soluciones para aplicaciones de alta exigencia, desarrollando metodologías para llevar a cabo la prognosis de algunas de las propiedades mecánicas más importantes de las aleaciones metálicas.

6.1 Conclusiones generales

Se pueden extraer las siguientes conclusiones generales de las publicaciones que conforman esta Tesis Doctoral:

- La combinación de la metodología de niveles de severidad junto con técnicas de simulación numérica (método de los elementos finitos) constituye una herramienta de gran potencial para los ingenieros que trabajan en aplicaciones de alta exigencia [9].
- Se ha explorado la posibilidad de utilizar materiales compuestos en aplicaciones nucleares de alta exigencia que sean capaces de cumplir, al menos, los mismos requisitos que se han impuesto a los aceros y, por tanto, puedan ser considerados como potenciales sustitutos válidos de componentes sometidos a radiación gamma y neutrónica [9].
- Las técnicas basadas en IA se pueden utilizar para extraer información de grandes conjuntos de datos, los cuales pueden ser difíciles de explotar de forma tradicional. Además, estas técnicas pueden utilizarse para ayudar a los diseñadores a seleccionar los materiales más adecuados para cada aplicación industrial [1, 3, 9, 12, 25].
- Aunque las metodologías basadas en aprendizaje supervisado requieren un entrenamiento previo y, en general, son computacionalmente más exigentes, se pueden obtener resultados muy prometedores. Por otra parte, aunque las metodologías basadas en aprendizaje no supervisado logran resultados menos ambiciosos, pueden servir como una herramienta para tratar de simplificar un conjunto de datos de entrada [1].

- La tecnología de redes neuronales artificiales puede emplearse para explotar grandes conjuntos de datos sobre materiales para predecir propiedades físicas y mecánicas de aleaciones metálicas a partir de datos conocidos o de fácil obtención como su composición química, temple o dureza [3, 12, 25].
- Se puede entrenar una red neuronal artificial para predecir la densidad y el módulo de Young de un material metálico tomando como dato de partida, únicamente, su composición química [12]. Además, se puede predecir la aproximación bilineal de la curva tensión-deformación de una aleación de aluminio si su composición química y tratamientos están bien definidos [3]. Por otra parte, se puede predecir el límite elástico y la resistencia máxima a la tracción de las aleaciones de aluminio en función de su dureza, composición química y temple [25].
- Las metodologías basadas en aprendizaje supervisado requieren grandes conjuntos de datos de entrenamiento para lograr un rendimiento satisfactorio puesto que su capacidad predictiva mejora a medida que la red aprende funciones más precisas (funciones que se aproximan mejor a la realidad del problema) [3, 12, 25].
- Se puede entrenar una red neuronal artificial con topología multicapa para aproximar funciones no lineales relacionadas con la ciencia de los materiales [3, 12, 25].
- Los resultados de una metodología basada en inteligencia artificial pueden alcanzar un rendimiento similar al obtenido mediante ecuaciones empíricas y, aunque el rendimiento predictivo de esta metodología pueda ser ligeramente inferior al de las ecuaciones, proporciona una mayor generalidad [3, 12, 25].

6.2 Conclusiones particulares

De cada uno de los trabajos publicados, se pueden extraer las siguientes conclusiones particulares:

- Los materiales compuestos reforzados con fibra de boro pueden cumplir con los requisitos de los aceros estructurales, agregando, además, la capacidad del boro para reducir la velocidad de los neutrones que componen la radiación neutrónica [9].
- Fabricar tubos de resina fenólica reforzada con fibras largas de boro puede ser una alternativa económicamente aceptable con respecto a fabricarlos con un tipo de acero típicamente empleado en la industria nuclear, WWER 15Kh2MFAA [9].
- Se puede entrenar una red neuronal artificial para predecir la densidad y el módulo de Young de una aleación metálica si su composición química está definida correctamente. El error predictivo permanece acotado y las desviaciones promedio, en este trabajo, son, respectivamente, 0.5% y 1.5% [12].
- Se puede entrenar una red neuronal artificial para predecir la aproximación bilineal de la curva tensión-deformación de una aleación de aluminio si su composición química y tratamiento (térmico y mecánico) están bien definidos. El error de predicción se mantiene acotado y las desviaciones medias, en este trabajo, para el módulo de Young, el límite elástico, la resistencia máxima a la tracción y el alargamiento a la rotura son, respectivamente, 3,07%, 4,58%, 3,30% y 5,90% [3].
- Se pueden emplear redes neuronales artificiales para predecir el límite elástico y la resistencia máxima a la tracción de las aleaciones de aluminio en función de su dureza, composición química y temple. En este estudio, el desempeño predictivo es estable y las desviaciones promedio para ambas propiedades son, respectivamente, 3.67% y 4.26% [25].
- Los resultados de una metodología basada en inteligencia artificial pueden alcanzar un rendimiento similar al obtenido mediante ecuaciones empíricas. En este estudio, para el Al 7010, la diferencia resultante entre ambos métodos fue ~2%. Aunque el rendimiento predictivo de la metodología basada

en IA es ligeramente inferior al de las ecuaciones empíricas, proporciona una mayor generalidad ya que puede realizar predicciones sobre cualquier aleación de aluminio [25].

6.3 Desarrollos futuros

Se ha demostrado que una metodología de apoyo a la selección de materiales basada en simulación numérica, analítica de datos e inteligencia artificial puede ser una herramienta de gran utilidad tareas de predicción de propiedades de materiales que operan en aplicaciones de alta exigencia. Como trabajos futuros, se podría abordar el empleo de dicha metodología en otros sectores industriales, implicando, asimismo, el análisis de otros tipos de materiales como, por ejemplo, los materiales cerámicos y sus compuestos. De forma más específica:

- Este estudio ha demostrado que es posible predecir las propiedades mecánicas más importantes de aleaciones metálicas utilizando técnicas basadas en IA. De la misma forma, se abre la puerta a investigar soluciones similares aplicadas a otros materiales.
- Las redes neuronales han demostrado ser una herramienta de gran potencial para predecir propiedades de materiales industriales de gran relevancia sin la necesidad de costosos ensayos experimentales. Puede ser interesante estudiar si es posible diseñar un sistema basado en inteligencia artificial capaz de predecir otras propiedades.
- Una vez que este esquema de trabajo ha demostrado que se puede utilizar para realizar predicciones adecuadas, se pueden probar otras arquitecturas de red para mejorar el rendimiento general de la metodología. Existe una amplia variedad de topologías de red que satisfacen diversas necesidades y resuelven diferentes problemas.

Referencias

[1] D. Merayo, A. Rodriguez-Prieto y A. M. Camacho, «Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0,» *Procedia manufacturing,* vol. 41, pp. 42-49, 2019.

[2] A. A. White, «Big data are shaping the future of materials science,» *MRS Bulletin,* vol. 38, nº 8, pp. 594-595, 2013.

[3] D. Merayo, A. Rodríguez-Prieto y A. M. Camacho, «Prediction of the Bilinear Stress-Strain Curve of Aluminum Alloys Using Artificial Intelligence and Big Data,» *Metals,* vol. 10, p. 904, 2020.

[4] D. Ashkenazi, «How Aluminum Changed the World: A Metallurgical Revolution through Technological and Cultural Perspectives,» *Technological Forecasting & Social Change,* vol. 143, pp. 101-113, 2019.

[5] T. W. Clyne y D. Hull, An introduction to composite materials., Cambridge, UK: Cambridge university press, 2019.

[6] R. F. Gibson, Principles of composite material mechanics., Boca Raton, Florida, USA: CRC Press, 2016.

[7] E. J. Barbero, Introduction to composite materials design, Boca Raton, Florida, USA: CRC press, 2017.

[8] Y. Lee, J. J. Dang, J. Jo, K. J. Chung, Y. S. Hwang, M. S. Cheon y L. Bertalot, «Preliminary study on capsule material for ITER neutron activation system,» *Fusion Engineering and Design,* vol. 89, nº 9-10, pp. 1894-1898., 2014.

[9] D. Merayo, A. Rodríguez-Prieto y A. M. Camacho, «Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications,» *Journal of materials: design and applications,* vol. 233, nº 10, pp. 2072-2083, 2019.

[10] R. Christensen, The theory of materials failure, Stanford: Oxford University Press, 2013.

[11] F. Durka, W. Morgan y D. Williams, Structural mechanics: loads, analysis, design, and materials, Liverpool: Prentice Hall, 2010.

[12] D. Merayo, A. Rodriguez-Prieto y A. M. Camacho, «Prediction of Physical and Mechanical Properties for Metallic Materials Selection Using Big Data and Artificial Neural Networks,» *IEEE Access,* vol. 8, pp. 13444-13456, 2020.

[13] A. A. Morini, M. J. Ribeiro y D. Hotza, «Early-stage materials selection based on embodied energy and carbon footprint,» *Materials & Design,* vol. 178, p. 107861, 2019.

[14] A. Piselli, W. Baxter, M. Simonato, B. Del Curto y M. Aurisicchio, «Development and evaluation of a methodology to integrate technical and sensorial properties in materials selection,» *Materials & Design,* vol. 153, pp. 259-272, 2018.

[15] S. H. Mousavi-Nasab y A. Sotoudeh-Anvari, «A comprehensive MCDM-based approach using TOPSIS, COPRAS and DEA as an auxiliary tool for material selection problems,» *Materials & Design,* vol. 121, pp. 237-253, 2017.

[16] D. Das, S. Bhattacharya y B. Sarkar, «Decision-based design-driven material selection: a normative-prescriptive approach for simultaneous selection of material and geometric variables in gear design,» *Materials & Design,* vol. 92, pp. 787-793, 2016.

[17] T. Alam y A. H. Ansari, «Review of aluminium and its alloys for automotive applications,» *International journal of advanced technology in engineering and science,* vol. 5, nº 5, pp. 278-294, 2017.

[18] M. Kamaya y M. Kawakubo, «A procedure for determining the true stress–strain curve over a large range of strains using digital image correlation and finite element analysis,» *Mechanics of Materials,* vol. 43, nº 5, pp. 243-253, 2011.

[19] Á. Rodriguez-Prieto, A. M. Camacho y M. A. Sebastian, «Materials selection criteria for nuclear power applications: a decision algorithm,» *JOM,* vol. 2, nº 68, pp. 496-506, 2015.

[20] D. Dimiduk, M. Holm y E. Niezgoda, «Perspectives on the Impact of Machine Learning, Deep Learning, and Artificial Intelligence on Materials, Processes, and Structures Engineering,» *Integrating Materials and Manufacturing Innovation,* vol. 7, nº 3, pp. 157-172, 2018.

[21] Y. Liu, T. Zhao, W. Ju y S. Shi, «Materials discovery and design using machine learning,» *Journal of Materiomics,* vol. 3, nº 3, pp. 159-177, 2017.

[22] S. Tu, X. Ren, J. He y Z. Zhang, «Stress–strain curves of metallic materials and post-necking strain hardening characterization: A review,» *Fatigue Fract. Eng. Mater. Struct.,* vol. 43, pp. 3-19, 2020.

[23] T. Zhou, Z. Song y K. Sundmacher, «Big Data Creates New Opportunities for Materials Research: A Review on Methods and Applications of Machine Learning for Materials Design,» *Engineering,* vol. 5, nº 6, pp. 1017-1026, 2019.

[24] J. Schmidt, M. Marques, S. Botti y M. Marques, «Recent advances and applications of machine learning in solid-state materials science,» *npj Computational Materials,* vol. 5, nº 1, pp. 1-36, 2019.

[25] D. Merayo, A. Rodríguez-Prieto y A. M. Camacho, «Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys,» *Materials,* vol. 13, p. 5227, 2020.

[26] J. Ling, E. Antono, S. Bajaj, S. Paradiso, M. Hutchinson, B. Mereding y B. Gibbons, «Machine learning for alloy composition and process optimization,» de *Proceedings of the ASME Turbo Expo 2018: Turbomachinery Technical Conference and Exposition*, Oslo, Norway, Jun. 2018.

[27] ASM International, Atlas of stress-strain curves, Ohio: ASM International, 2002.

[28] ASTM International, Standard test methods for tension testing of metallic materials (ASTM E8/E8M - 16AE1), West Conshohocken: ASTM International, 2016.

[29] Y. Yogo, M. Sawamura, N. Iwata y N. Yukawa, «Stress-strain curve measurements of aluminum alloy and carbon steel by unconstrained-type high-pressure torsion testing,» *Materials and design,* vol. 122, pp. 226-235, 2017.

[30] D. Fertis, Infrastructure Systems: Mechanics, Design, and Analysis of Components, New York : Wiley-Interscience, 1997.

[31] R. Christensen, «Observations on the definition of yield stress,» *Acta Mechanica 196,* vol. 196, nº 3-4, pp. 239-244, 2008.

[32] A. M. Camacho, A. Rodríguez-Prieto, J. M. Herrero, A. M. Aragón, C. Bernal, C. Lorenzo-Martín, A. Yanguas-Gil y P. A. F. Martins, «An Experimental and Numerical Analysis of the Compression of Bimetallic Cylinders,» *Materials,* vol. 12, p. 4094, 2019.

[33] J. M. Gutiérrez y A. M. Camacho, «Investigations on the influence of blank thickness (t) and length/wide punch ratio (LD) in rectangular deep drawing of dual-phase steels,» *Comput. Mater. Sci.,* vol. 91, pp. 134-145, 2014.

[34] E. M. Rubio, A. M. Camacho, R. Pérez y M. M. Marín, «Guidelines for selecting plugs used in thin-walled tube drawing processes of metallic alloys,» *Metals,* vol. 7, p. 572, 2017.

[35] International Organization for Standardization (ISO), Metallic materials — Tensile testing — Part 1: Method of test at room temperature (DIN EN ISO 6892-1:2019), 2019.

[36] F. C. Campbell, Structural composite materials, Russell Township, Ohio, USA: ASM international, 2010.

[37] G. P. Sendeckyj, Mechanics of Composite Materials: Composite Materials, Amsterdam, Netherlands: Elsevier, 2016.

[38] S. Helal, «The Expanding Frontier of Artificial Intelligence,» *IEEE,* vol. 51, nº 9, pp. 14-17, 2018.

[39] P. Jackson, Introduction to artificial intelligence, New York: Dover Publications Inc., 2019.

[40] A. Krizhevsky, I. Sutskever y G. E. Hinton, «Imagenet classification with deep convolutional neural networks,» *Advances in neural information processing systems,* vol. 1, pp. 1097-1105, 2012.

[41] K. Johnson, J. Torres Soto, B. Glicksberg, K. Shameer, R. Miotto, M. Ali, E. Ashley y J. Dudley, «Artificial Intelligence in Cardiology,» *Journal of the American College of Cardiology,* vol. 71, nº 23, pp. 2668-2679, 2018.

[42] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre y T. Green, «Improved protein structure prediction using potentials from deep learning,» *Nature,* vol. 577, nº 7792, pp. 706-710, 2020.

[43] M. Cummings, Artificial intelligence and the future of warfare, London: Chatham House for the Royal Institute of International Affairs, 2017.

[44] F. Villa, M. Ceroni, K. Bagstad, G. Johnson y S. Krivov, «ARIES (ARtificial Intelligence for Ecosystem Services): a new tool for ecosystem services assessment, planning, and valuation,» de *Proceedings of the 11th Annual BIOECON Conference on Economic Instruments to Enhance the Conservation and Sustainable Use of Biodiversity*, Venice, Italy, 2009.

[45] G. Allen y T. Chan, Artificial intelligence and national security, Cambridge: Belfer Center for Science and International Affairs, 2017.

[46] J. H. Ee y N. Huh, «A study on the relationship between artificial intelligence and change in mathematics education,» *Communications of Mathematical Education,* vol. 32, nº 1, pp. 23-36, 2018.

[47] V. Kolesov, «Cognitive Modelling in Oil & Gas Exploration and Reservoir Prediction,» *80th EAGE Conference and Exhibition 2018,* vol. 2018, nº 1, pp. 1-5, 2018.

[48] T. Thankachan, K. Prakash, C. Pleass, D. Rammasamy, B. Prabakaran y S. Jothi, «Artificial neural network to predict the degraded mechanical properties of metallic materials due to the presence of hydrogen,» *International Journal of Hydrogen Energy,* vol. 42, nº 47, pp. 28612-28621, 2017.

[49] L. Qian, E. Winfree y J. Bruck, «Neural Network Computation with DNA Strand Displacement Cascades,» *Nature,* vol. 475, nº 7356, pp. 368-372, 2011.

[50] J. Schmidhuber, «Deep learning in neural networks: an overview,» *Neural networks,* vol. 61, pp. 85-117, 2015.

[51] G. Huang, G. Huang, S. Song y K. You, «Trends in extreme learning machines: a review,» *Neural Networks,* vol. 61, pp. 32-48, 2015.

[52] D. Kingma y J. Ba, «ADAM: a Method for Stochastic Optimization,» de *3rd International Conference for Learning Representations*, San Diego, 2015.

[53] A. Agrawal y A. Choudhary, «Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science,» *Apl. Materials,* vol. 4, nº 5, p. 053208, 2016.

[54] I.-Y. Song y Y. Zhu, «Big data and data science: what should we teach?,» *Expert Systems,* vol. 33, pp. 364-373, 2016.

[55] J. Rowley, «The wisdom hierarchy: representations of the DIKW hierarchy,» *Journal of Information Science,* vol. 33, nº 2, pp. 163-180, 2007.

[56] S. Batra, «Big Data Analytics and Its Reflections on DIKW Hierarchy,» *Review of knowledge management,* vol. 4, nº 1-2, pp. 5-17, 2014.

[57] R. Akerkar, Big data computing, Sogndal: Crc Press, 2013.

[58] D. García-Gil, S. Ramírez-Gallego, S. García y F. Herrera, «Principal components analysis random discretization ensemble for big data. Knowledge-Based Systems, 150, 166-174.,» *Knowledge-Based Systems,* vol. 150, pp. 166-174, 2018.

[59] T. Erl, W. Khattak y P. Buhler, Big data fundamentals: concepts, drivers & techniques, New Jersey: Prentice Hall Press, 2016.

[60] D. Laney, «3D Data Management: Controlling Data Volume, Velocity, and Variety,» *META Group Research Note,* vol. 6, nº 949, p. 1, 2001.

[61] C. Dobre y F. Xhafa, «Intelligent services for Big data science,» *Future Generation Computer Systems,* vol. 37, pp. 267-281, 2014.

[62] S. Guo, J. Yu, X. Liu, C. Wang y Q. Jiang, «A predicting model for properties of steel using the industrial big data based on machine learning,» *Computational Materials Science,* vol. 160, pp. 95-104, 2019.

[63] P. Balachandran, «Machine learning guided design of functional materials with targeted properties,» *Computational Materials Science,* vol. 164, pp. 82-90, 2019.

[64] U.S. Federal government, «Material Genome Initiative,» MaterialsGenome Inc., 2021. [En línea]. Available: https://www.mgi.gov/. [Último acceso: enero 2021].

[65] M. Danylenko, «Aluminium alloys in aerospace,» *Aluminium International Today,* vol. 31, nº 4, p. 35, 2018.

[66] J. Weinbub, M. Wastl, K. Rupp, F. Rudolf y S. Selberherr, «ViennaMaterials – A dedicated material library for computational science and engineering,» *Applied Mathematics and Computation,* vol. 267, pp. 282-293, 2015.

[67] Matmatch GmbH, «Matmatch,» [En línea]. Available: https://matmatch.com/. [Último acceso: April 2019].

[68] Á. Rodríguez-Prieto, Análisis de requisitos tecnológicos de materiales especificados en normativas reguladas y su repercusión sobre la fabricación de recipientes especiales para la industria nuclear, Tesis Doctoral, UNED, 2014.

[69] Á. Rodríguez-Prieto, A. M. Camacho and M. Á. Sebastián, "Materials selection criteria for nuclear power applications: a decision algorithm," *JOM,* vol. 2, no. 68, pp. 496-506, 2016.

[70] J. N. Reddy, Introduction to the finite element method, New York, USA: McGraw-Hill Education, 2019.

[71] W. D. Callister y D. G. Rethwisch, Materials science and engineering: an introduction, New York: Wiley, 2018.

[72] A. M. Camacho, Análisis por el método de los elementos finitos de procesos estacionarios de conformado por deformación plástica, Tesis Doctoral, UNED, 2005.

[73] G. R. Liu y S. S. Quek, The finite element method: a practical course, Oxford, UK: Butterworth-Heinemann, 2013.

[74] M. Lutz, Programming Python, Sebastopol, CA, USA: O'Reilly Media, 2010.

[75] P. Joshi, Artificial intelligence with Python: Build real-world artificial intelligence applications with Python to intelligently interact with the world around you, Birmingham, UK: Packt Publishing, 2017.

[76] A. Deshpande y M. Kumar, Artificial Intelligence for Big data: Complete guide to automating Big data solutions using Artificial Intelligence techniques, Birmingham, UK: Packt Publishing, 2018.

[77] J. McCarthy, M. Minsky, N. Rochester y C. Shannon, «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence: August 31, 1955,» *AI magazine,* vol. 27, nº 4, pp. 12-14, 1955.

[78] A. Bacha, C. Maurice, H. Klocker y J. Driver, «The large strain flow stress behaviour of aluminium alloys as measured by channel-die compression (20–500 C),» *Material science forum,* vol. 519, pp. 783-788, 2006.

[79] K. Hornik, «Approximation capabilities of multilayer feedforward networks,» *Neural Networks,* vol. 4, nº 2, pp. 251-257, 1991.

[80] ASM international, ASM Handbook volume 2 - Properties and selection: Nonferrous alloys and special-purpose materials, Ohio: ASM International, 1990.

[81] ASTM International, ASTM E10-18, Standard Test Method for Brinell Hardness of Metallic Materials, West Conshohocken, PA: ASTM International, 2018.

[82] ASTM, ASTM E140-12B(2019)e1 - Standard Hardness Conversion Tables for Metals Relationship Among Brinell Hardness, Vickers Hardness, Rockwell Hardness, Superficial Hardness, Knoop Hardness, Scleroscope Hardness, and Leeb Hardness, West Conshohocken, PA: ASTM International, 2019.

[83] ASTM International, ASTM E8 / E8M-16ae1, Standard Test Methods for Tension Testing of Metallic Materials, West Conshohocken, PA, 2016.

[84] The aluminum association, Designations and Chemical Composition Limits for Aluminum Alloys in the Form of Castings and Ingot, vol. 1, The aluminum association, 2008, pp. 1-20.

[85] DIN, DIN 50159-1 - Metallic materials - Hardness testing with the UCI method - Part 1: Test method, Berlin, Germany: Deutsches Institut für Normung, 2015.

[86] Eurocode, Eurocode 9 - Design of aluminium structures. Part 1-1, Brussels: European Commitee for standardization, 2009.

[87] The aluminum association, International Alloy Designations and Chemical Composition Limits for Wrought Aluminum and Wrought Aluminum Alloys, vol. 1, The aluminum association, 2018, pp. 1-45.

[88] ISO, ISO 6507-1:2005 Metallic materials — Vickers hardness test — Part 1: Test method, Geneva, Switzerland: International Organization for Standardization, 2005.

[89] C. Barajas, J. de Vicente, J. Caja, P. Maresca y E. Gómez, «Considerations to the hardness Brinell measurement using optical equipment,» *Procedia Manufacturing,* vol. 13, pp. 550-557, 2017.

[90] J. Benesty, J. Chen y Y. Huang, «On the importance of the Pearson correlation coefficient in noise reduction,» *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 16, nº 4, pp. 757-765, 2008.

[91] Y. Bouanan, G. Zacharewicz y B. Vallespir, «DEVS modelling and simulation of human social interaction and influence,» *Engineering Applications of Artificial Intelligence,* vol. 50, pp. 83-92, 2016.

[92] H. Boyer y T. Gall, Metals Handbook, Desk Edition, Metals Park, Ohio, USA: ASM International, 1985.

[93] R. Branco, F. Berto y A. Kotousov, «Editorial of the special issue on Mechanical Behaviour of Aluminium Alloys,» *Mechanical Behaviour of Aluminium Alloys,* vol. 1, pp. 1-3, 2018.

[94] J. R. Cahoon, «An improved equation relating hardness to ultimate strength,» *Metallurgical and Materials Transactions B,* vol. 3, nº 11, pp. 3040-3040, 1972.

[95] J. R. Cahoon, W. H. Broughton y A. R. Kutzak, «The determination of yield strength from hardness measurements,» *Metallurgical transactions,* vol. 2, nº 7, pp. 1979-1983, 1971.

[96] A. M. Camacho, Á. Rodríguez-Prieto, J. M. Herrero, A. M. Aragón, C. Bernal, C. Lorenzo-Martin, Á. Yanguas-Gil y P. A. F. Martins, «An Experimental and Numerical Analysis of the Compression of Bimetallic Cylinders,» *Materials,* vol. 12, nº 24, p. 4094, 2019.

[97] P. Chokshi, R. Dashwood y D. J. Hughes, «Artificial Neural Network (ANN) based microstructural prediction model for 22MnB5 boron steel during tailored hot stamping,» *Computers & Structures,* vol. 190, pp. 162-172, 2017.

[98] J. Davis, Alloying: Understanding the Basics, Ohio: ASM International, 2001.

[99] V. H. C. de Albuquerque, A. R. de Alexandria, P. C. Cortez y J. M. R. Tavares, «Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images,» *NDT & E International,* vol. 42, nº 7, pp. 644-651, 2009.

[100] S. Dey, N. Sultana, M. S. Kaiser, P. Dey y S. Datta, «Computational intelligence based design of age-hardenable aluminium alloys for different temperature regimes,» *Materials & Design,* vol. 92, pp. 522-534, 2016.

[101] O. J. Dunn, «Multiple Comparisons among Means,» *Journal of the American Statistical Association,* vol. 56, nº 293, pp. 52-64, 1961.

[102] C. W. Dunnett, «A multiple comparison procedure for comparing several treatments with a control,» *Journal of the American Statistical Association,* vol. 50, pp. 1096-1121, 1955.

[103] A. Elmishali, R. Stern y M. Kalech, «An artificial intelligence paradigm for troubleshooting software bugs,» *Engineering Applications of Artificial Intelligence,* vol. 69, pp. 147-156, 2018.

[104] S. Feng, H. Zhou y H. Dong, «Using deep neural network with small dataset to predict material defects.,» *Materials & Design,* vol. 162, pp. 300-310, 2019.

[105] G. Galevsky, V. Rudneva y V. Aleksandrov, «Current state of the world and domestic aluminium production and consumption,» *IOP Conference Series: Materials Science and Engineering,* vol. 411, nº 1, pp. 1-6, 2018.

[106] J. Gere y B. Goodno, Mechanics of materials, Stanford: Cengage learning, 2012.

[107] A. E. Giannakopoulos, P. L. Larsson y R. Vestergaard, «Analysis of Vickers indentation,» *International journal of solids and structures,* vol. 31, nº 19, pp. 2679-2708, 1994.

[108] F. Gui, «Novel corrosion schemes for the aerospace industry,» *Corrosion Control in the Aerospace Industry,* vol. 1, pp. 248-265, 2009.

[109] J. M. Gutiérrez Regueras y A. M. Camacho López, «Investigations on the influence of blank thickness (t) and length/wide punch ratio (LD) in rectangular deep drawing of dual-phase steels,» *Computational materials science,* vol. 91, pp. 134-145, 2014.

[110] G. Hahn y A. Rosenfield, «Metallurgical factors affecting fracture toughness of aluminum alloys,» *Metallurgical Transactions A,* vol. 6, nº 4, pp. 653-668, 1975.

[111] R. Hill, B. Storåkers y A. B. Zdunek, «A theoretical study of the Brinell hardness test,» *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences,* vol. 423, nº 1865, pp. 301-330, 1989.

[112] R. Hill, B. Storåkers y A. B. Zdunek, «A theoretical study of the Brinell hardness test,» *Proc. R. Soc. Lond. A,* vol. 423, pp. 301-330, 1989.

[113] D. C. Howell, Statistical Methods for Psychology, 8th Ed., Boston, Massachusetts, USA: Wadsworth Cengage Learning, 2012.

[114] A. Jenab, I. S. Sarraf, D. E. Green, T. Rahmaan y M. J. Worswick, «The use of genetic algorithm and neural network to predict rate-dependent tensile flow behaviour of AA5182-O sheets.,» *Materials & Design,* vol. 94, pp. 262-273, 2016.

[115] G. Johnson y T. Holmquist, Test data and computational strength and fracture model constants for 23 materials subjected to large strains, high strain rates and high temperatures, Los Alamos: Los Alamos National Laboratory, 1988.

[116] J. Kaufman, Introduction to aluminum alloys and tempers, Ohio: ASM international, 2000.

[117] A. Lal, «SANE 2.0: System for fine grained named entity typing on textual data,» *Engineering Applications of Artificial Intelligence,* vol. 84, pp. 11-17, 2019.

[118] H. Li, Z. Hu, Y. Chen, Q. Sun y X. Zhou, «Modeling mechanical properties and plastic strain for hot forming-quenching AA6061 aluminum alloy parts,» *International Journal of Lightweight Materials and Manufacture,* vol. 3, nº 1, pp. 66-72, 2020.

[119] W. Li, F. Le Gall y N. Spaseski, «A survey on model-based testing tools for test case generation,» de *International Conference on Tools and Methods for Program Analysis*, Springer, Cham, 2017.

[120] Y. Li y Z. Yang, «Evaluation of Measurement Uncertainty for Ultrasonic Contact Impedance Hardness Testing Results,» *Physical Testing and Chemical Analysis (Part A: Physical Testing),* vol. 6, p. 7, 2017.

[121] Y. Li, B. Yu, B. Wang, T. H. Lee y M. Banu, «Online quality inspection of ultrasonic composite welding by combining artificial intelligence technologies with welding process signatures,» *Materials & Design,* vol. 192, p. 108912, 2020.

[122] G. Liu, L. Jia, B. Kong, K. Guan y H. Zhang, «Artificial neural network application to study quantitative relationship between silicide and fracture toughness of Nb-Si alloys,» *Materials & Design,* vol. 129, pp. 210-218, 2017.

[123] A. MARTÍN, V. RODRÍGUEZ-FERNÁNDEZ y D. Camacho, «CANDYMAN: Classifying Android malware families by modelling dynamic traces with Markov chains,» *Engineering Applications of Artificial Intelligence,* vol. 74, pp. 121-133, 2018.

[124] F. Mazzolani, «EN1999 Eurocode 9: Design of aluminium structures,» *Proceedings of the Institution of Civil Engineers-Civil Engineering,* vol. 144, nº 6, pp. 61-64, 2001.

[125] S. D. Mesarovic y N. A. Fleck, «Spherical indentation of elastic–plastic solids,» *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences,* vol. 455, nº 1987, pp. 2707-2728, 1999.

[126] C. Nappi , «The Global Aluminium Industry 40 years from 1972,» *World aluminium,* vol. 1, pp. 1-27, 2013.

[127] T. Nicholas, Material behavior at high strain rates, Ohio: Wright-Patterson AFB, 1982.

[128] E. Pavlina y C. Van Tyne, «Correlation of Yield Strength and Tensile Strength with Hardness for Steels,» *Journal of Materials Engineering and Performance ,* vol. 17, pp. 888-893, 2008.

[129] H. Pelletier, J. Krier, A. Cornet y P. Mille, «Limits of using bilinear stress–strain curve for finite element modeling of nanoindentation response on bulk materials,» *Thin Solid Films,* vol. 379, nº 1-2, pp. 147-155, 2000.

[130] I. Perikos y I. Hatzilygeroudis, «Recognizing emotions in text using ensemble of classifiers,» *Engineering Applications of Artificial Intelligence,* vol. 51, pp. 191-201, 2016.

[131] W. Ramberg y W. Osgood, «Description of stress-strain curves by three parameters,» NACA, Washington DC, 1943.

[132] K. Rasmussen y J. Rondal, «Strength curves for metal columns,» *Journal of Structural Engineering,* vol. 123, nº 6, pp. 721-728, 1997.

[133] N. M. Razali y Y. B. Wah, «Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests,» *Journal of statistical modeling and analytics,* vol. 2, nº 1, pp. 21-33, 2011.

[134] E. M. Rubio, A. M. Camacho, R. Pérez y M. M. Marín, «Guidelines for selecting plugs used in thin-walled tube drawing processes of metallic alloys,» *Metals,* p. 7, 2017.

[135] M. A. Salazar-Guapuriche, Y. Y. Zhao, A. Pitman y A. Greene, «Correlation of strength with hardness and electrical conductivity for aluminium alloy 7010,» *Materials science forum,* vol. 519, pp. 853-858, 2006.

[136] G. Scamans y E. Butler, «In situ observations of crystalline oxide formation during aluminum and aluminum alloy oxidation,» *Metallurgical transactions A,* vol. 6, nº 11, pp. 2055-2063, 1975.

[137] S. S. Shapiro y M. B. Wilk, «An analysis of variance test for normality (complete samples),» *Biometrika,* vol. 52, nº 3-4, pp. 591-611, 1965.

[138] J. E. Siegel, S. Pratt, Y. Sun y S. Sarma, «Real-time deep neural networks for internet-enabled arc-fault detection,» *Engineering Applications of Artificial Intelligence,* vol. 74, pp. 35-42, 2018.

[139] V. K. Soo, J. Peeters, D. Paraskevas, P. Compston, M. Doolan y J. Duflou, «Sustainable aluminium recycling of end-of-life products: A joining techniques perspective,» *Journal of cleaner production,* vol. 178, pp. 119-132, 2018.

[140] J. Stier, G. Gianini, M. Granitzer y K. Ziegler, «Analysing neural network topologies: a game theoretic approach,» *Procedia Computer Science,* vol. 126, pp. 234-243, 2018.

[141] M. Szumigała y Ł. Polus, «An numerical simulation of an aluminium-concrete beam,» *Procedia engineering,* vol. 172, pp. 1086-1092, 2016.

[142] D. Tabor, The hardness of metals, Oxford, UK: Oxford university press, 2000.

[143] B. Taljat, T. Zacharia y F. Kosel, «New analytical procedure to determine stress-strain curve from spherical indentation data,» *International Journal of Solids and Structures,* vol. 35, nº 33, pp. 4411-4426, 1998.

[144] M. Tiryakioglu y J. Campbell, «On macrohardness testing of Al-7 wt.% Si–Mg alloysI. Geometrical and mechanical aspects,» *Materials Science and Engineering,* vol. 361, pp. 232-239, 2003.

[145] M. Tiryakioğlu, J. Robinson, M. Salazar-Guapuriche, Y. Zhao y P. Eason, «Hardness–strength relationships in the aluminum alloy 7010,» *Materials Science and Engineering: A,* vol. 631, pp. 196-200, 2015.

[146] M. Tiryakioğlu, J. Campbell y J. Staley, «On macrohardness testing of Al–7 wt.% Si–Mg alloys: II. An evaluation of models for hardness–yield strength relationships,» *Materials Science and Engineering: A,* vol. 361, nº 1-2, pp. 240-248, 2003.

[147] M. Tiryakioğlu y J. Robinson, «On the representative strain in Vickers hardness testing of 7010 aluminum alloy,» *Materials Science and Engineering: A,* vol. 641, pp. 231-236, 2015.

[148] S. Tu, X. Ren, J. He y Z. Zhang, «Stress–strain curves of metallic materials and post-necking strain hardening characterization: A review,» *Fatigue Fract. Eng. Mater. Struct.,* vol. 43, pp. 3-19, 2020.

[149] S. Walczak y N. Cerpa, «Heuristic principles for the design of artificial neural networks,» *Information and software technology,* vol. 41, nº 2, pp. 107-117, 1999.

[150] S. Walker, «Bauxite and Alumina: Growth Maintained,» *Engineering and Mining Journal,* vol. 216, nº 3, pp. 42-47, 2015.

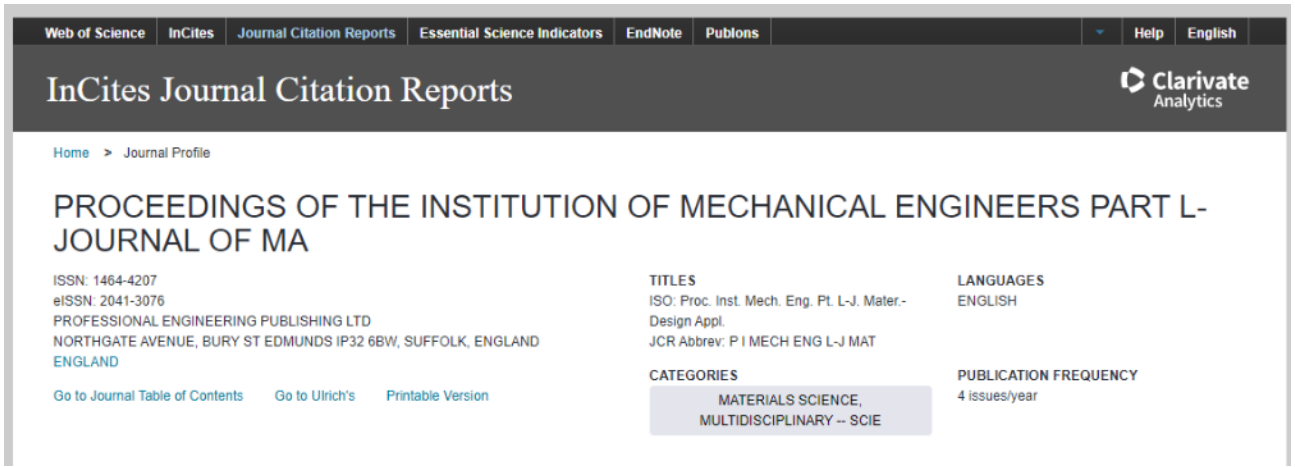
[151] A. Yamanaka, R. Kamijyo, K. Koenuma, I. Watanabe y T. Kuwabara, «Deep neural network approach to estimate biaxial stress-strain curves of sheet metals,» *Materials & Design,* p. 108970, 2020.

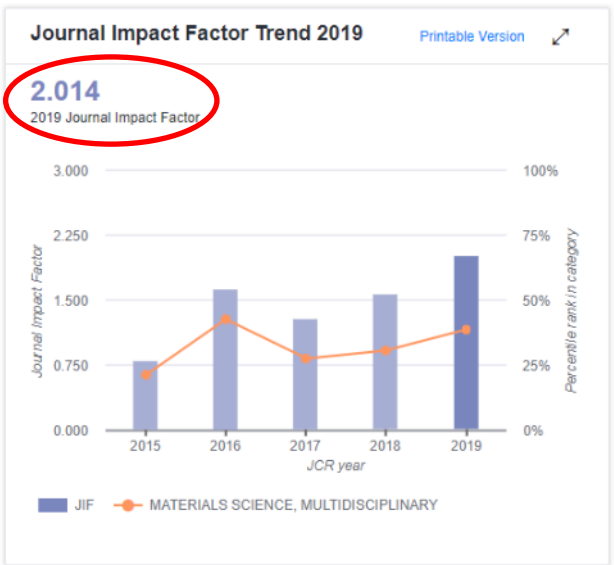
[152] J. Yoon, J. Kim, J. Jung, D. Lee, H. Jeong, M. Shahbaz, S. Lee y H. Kim, «Obtaining reliable true plastic stress-strain curves in a wide range of strains using digital image correlation in tensile testing,» *Korean J Met Mater,* vol. 54, pp. 231-236, 2016.

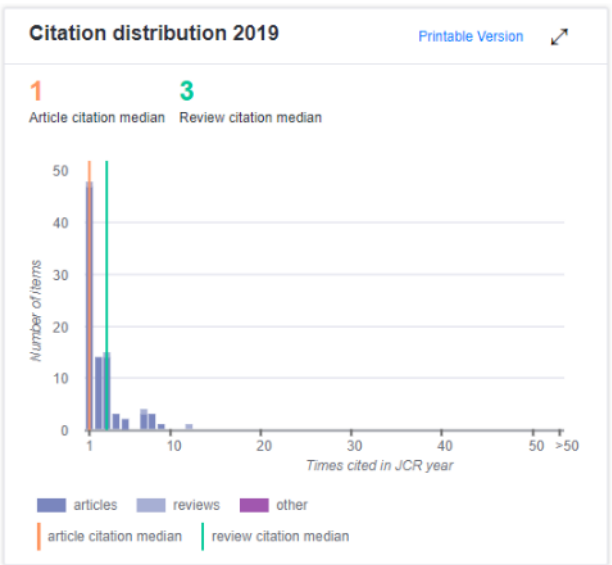
[153] P. Zhang, S. X. Li y Z. F. Zhang, «General relationship between strength and hardness,» *Materials Science and Engineering: A,* vol. 529, pp. 62-73, 2011.

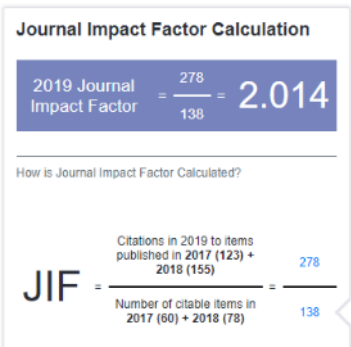
Apéndices

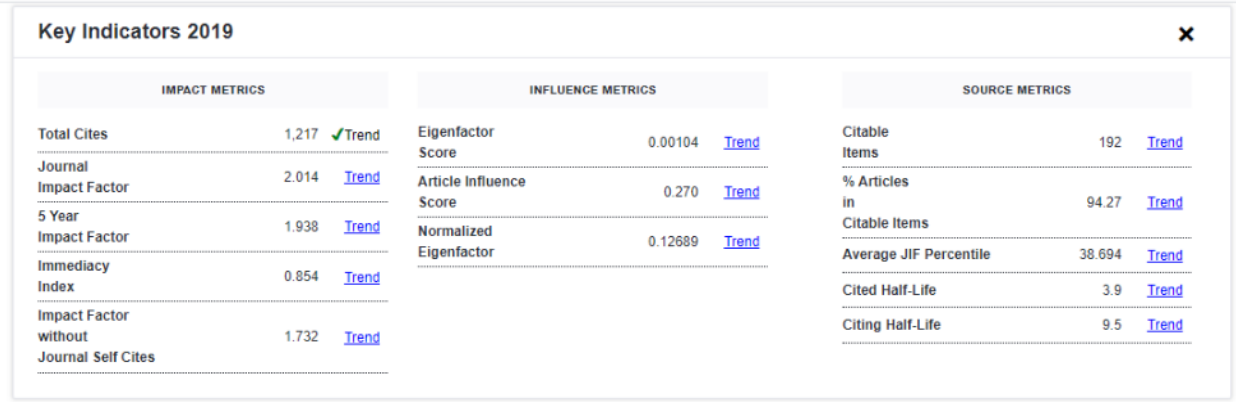
Apéndice A. Indicios de calidad del artículo "*Analytical and numerical study for selecting polymeric matrix composites intended to nuclear applications*" [9]









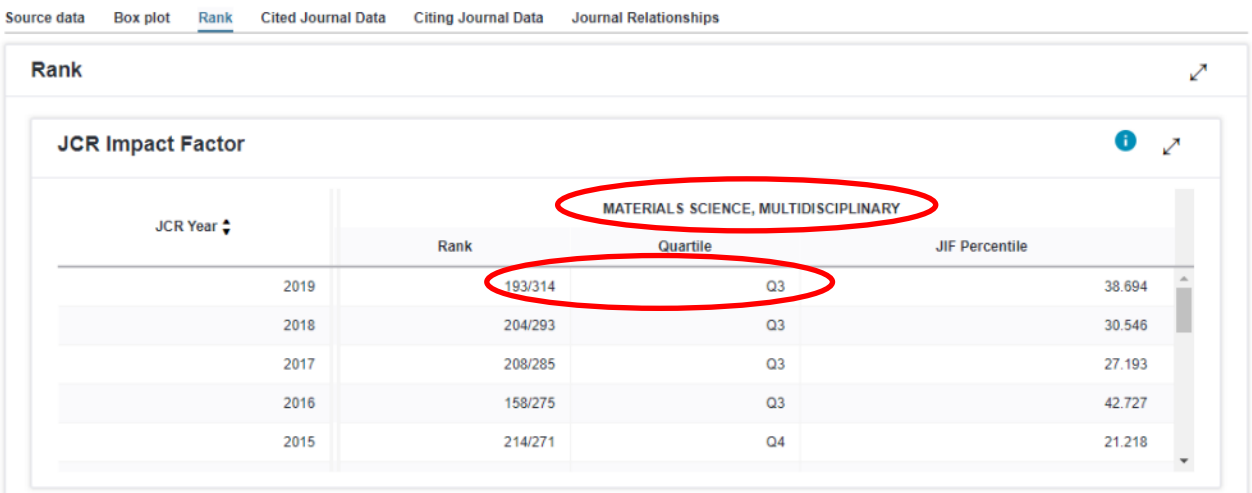


Key Indicators 2019

IMPACT METRICS		
Total Cites	1,217	✓Trend
Journal Impact Factor	2.014	Trend
5 Year Impact Factor	1.938	Trend
Immediacy Index	0.854	Trend
Impact Factor without Journal Self Cites	1.732	Trend

INFLUENCE METRICS		
Eigenfactor Score	0.00104	Trend
Article Influence Score	0.270	Trend
Normalized Eigenfactor	0.12689	Trend

SOURCE METRICS		
Citable Items	192	Trend
% Articles in Citable Items	94.27	Trend
Average JIF Percentile	38.694	Trend
Cited Half-Life	3.9	Trend
Citing Half-Life	9.5	Trend



Source data | Box plot | Rank | Cited Journal Data | Citing Journal Data | Journal Relationships

Rank

JCR Impact Factor

JCR Year	MATERIALS SCIENCE, MULTIDISCIPLINARY		
	Rank	Quartile	JIF Percentile
2019	193/314	Q3	38.694
2018	204/293	Q3	30.546
2017	208/285	Q3	27.193
2016	158/275	Q3	42.727
2015	214/271	Q4	21.218

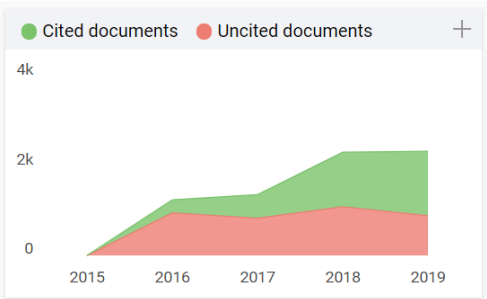
Apéndice B. Indicios de calidad del artículo "*Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0*" [1]

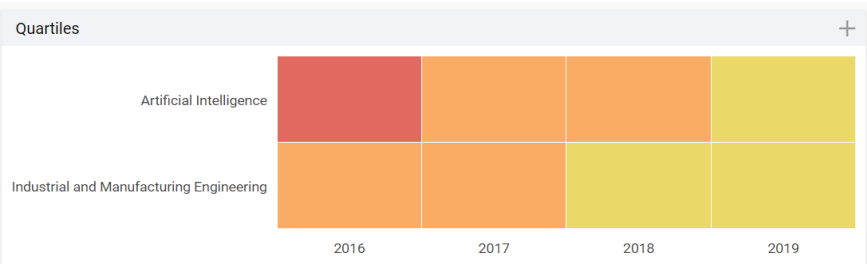
Procedia Manufacturing

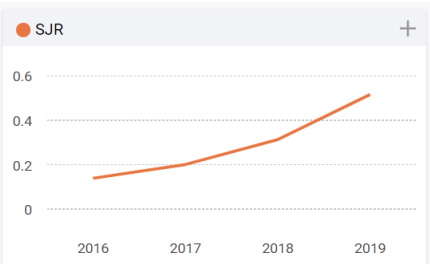
Country	Netherlands - SCIMAGO INSTITUTIONS RANKINGS
Subject Area and Category	Computer Science Artificial Intelligence Engineering Industrial and Manufacturing Engineering
Publisher	Elsevier BV
Publication type	Journals
ISSN	23519789
Coverage	2015-2019
Scope	Procedia Manufacturing is an open access journal focusing entirely on publishing high quality conference proceedings. Procedia Manufacturing enables fast dissemination of its content so that conference delegates can publish their papers in a dedicated online issue on Sciverse ScienceDirect under the Creative Commons license BY-NC-ND. (For further details see our open access license policy.) Papers published on Procedia Manufacturing will be available online on ScienceDirect within 8 weeks of acceptance of the final manuscripts. Procedia Manufacturing will publish papers from conferences on all important topics in the field of manufacturing engineering, including but not limited to manufacturing processes, systems and emerging topics in manufacturing, such as: -Manufacturing process modeling -Automation and robotics in manufacturing -Sensing and control -Manufacturing system design and operations -Biomanufacturing -Micro/nano manufacturing

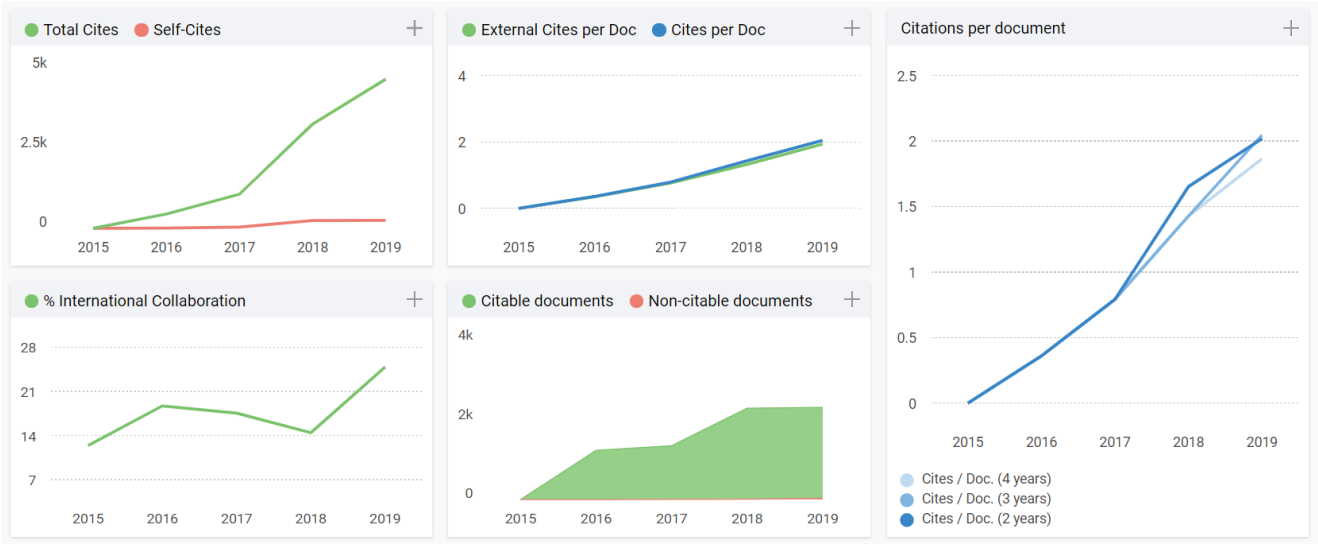
29

H Index



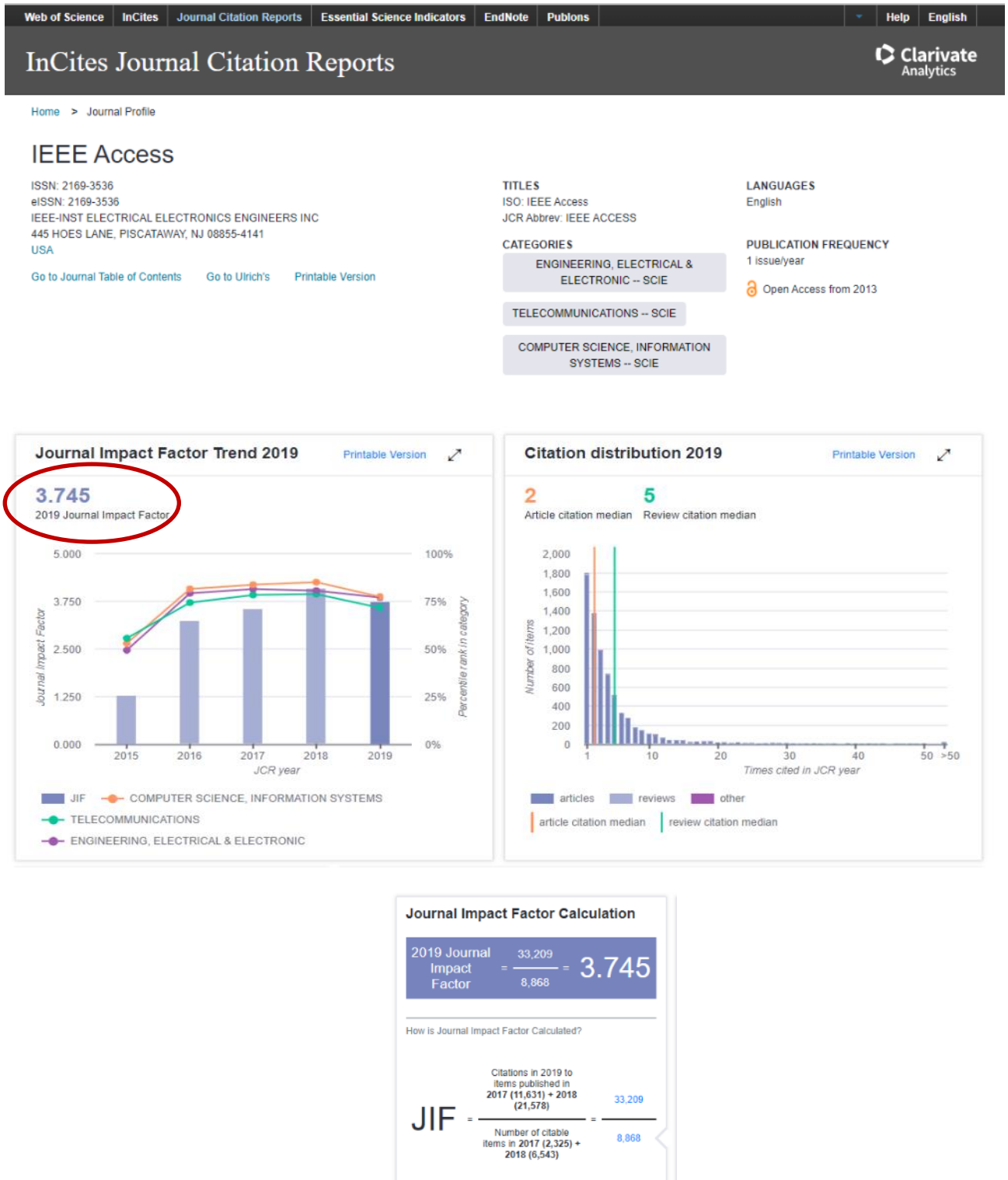






Category	Year	Quartile
Artificial Intelligence	2016	Q4
Artificial Intelligence	2017	Q3
Artificial Intelligence	2018	Q3
Artificial Intelligence	2019	Q2
Industrial and Manufacturing Engineering	2016	Q3
Industrial and Manufacturing Engineering	2017	Q3
Industrial and Manufacturing Engineering	2018	Q2
Industrial and Manufacturing Engineering	2019	Q2

Apéndice C. Indicios de calidad del artículo "*Prediction of physical and mechanical properties for metallic materials selection using big data and artificial neural networks*" [12]



Source data | Box plot | Rank | Cited Journal Data | Citing Journal Data | Journal Relationships

Journal source data 2019

	Articles	Reviews	Combined(C)	Other(O)	Percentage(C/(C+O))
Number in JCR Year 2019 (A)	14,811	174	14,985	39	100%
Number of References (B)	582,720	20,078	602,798	59	100%
Ratio (B/A)	39.3	115.4	40.2	1.5	

Key Indicators 2019

IMPACT METRICS		
Total Cites	51,038	✓Trend
Journal Impact Factor	3.745	Trend
5 Year Impact Factor	4.076	Trend
Immediacy Index	0.615	Trend
Impact Factor without Journal Self Cites	2.429	Trend

INFLUENCE METRICS		
Eigenfactor Score	0.08168	Trend
Article Influence Score	0.643	Trend
Normalized Eigenfactor	9.95471	Trend

SOURCE METRICS		
Citable Items	14,985	Trend
% Articles in Citable Items	98.84	Trend
Average JIF Percentile	75.602	Trend
Cited Half-Life	1.8	Trend
Citing Half-Life	5.1	Trend

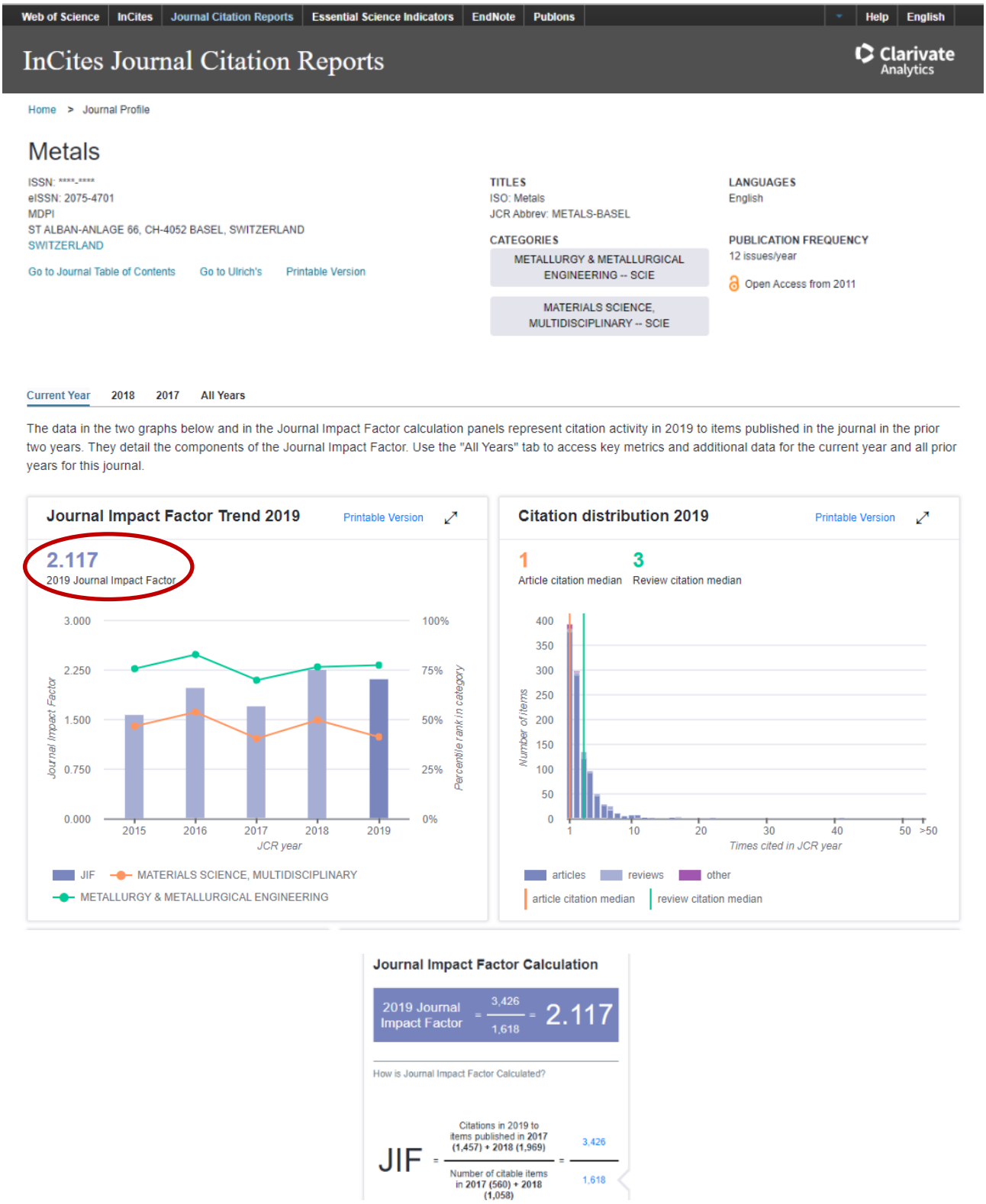
Source data | Box plot | Rank | Cited Journal Data | Citing Journal Data | Journal Relationships

Rank

JCR Impact Factor

JCR Year	COMPUTER SCIENCE, INFORMATION SYSTEMS			ENGINEERING, ELECTRICAL & ELECTRONIC			TELECOMMUNICATIONS		
	Rank	Quartile	JIF Percentile	Rank	Quartile	JIF Percentile	Rank	Quartile	JIF Percentile
2019	35/156	Q1	77.885	61/266	Q1	77.256	26/90	Q2	71.667
2018	23/155	Q1	85.484	52/266	Q1	80.639	19/88	Q1	78.977
2017	24/148	Q1	84.122	48/260	Q1	81.731	19/87	Q1	78.736
2016	27/146	Q1	81.849	54/262	Q1	79.580	23/89	Q2	74.719
2015	68/144	Q2	53.125	131/257	Q3	49.222	37/82	Q2	55.488

Apéndice D. Indicios de calidad del artículo "*Prediction of the Bilinear Stress-Strain Curve of Aluminum Alloys Using Artificial Intelligence and Big Data*" [3]



Source data | Box plot | Rank | Cited Journal Data | Citing Journal Data | Journal Relationships

Journal source data 2019

	Articles	Reviews	Combined(C)	Other(O)	Percentage(C/(C+O))
Number in JCR Year 2019 (A)	1,287	45	1,332	26	98%
Number of References (B)	43,012	4,219	47,231	319	99%
Ratio (B/A)	33.4	93.8	35.5	12.3	

Key Indicators 2019

IMPACT METRICS		
Total Cites	5,708	✓Trend
Journal Impact Factor	2.117	Trend
5 Year Impact Factor	2.244	Trend
Immediacy Index	0.594	Trend
Impact Factor without Journal Self Cites	1.633	Trend

INFLUENCE METRICS		
Eigenfactor Score	0.00925	Trend
Article Influence Score	0.343	Trend
Normalized Eigenfactor	1.12793	Trend

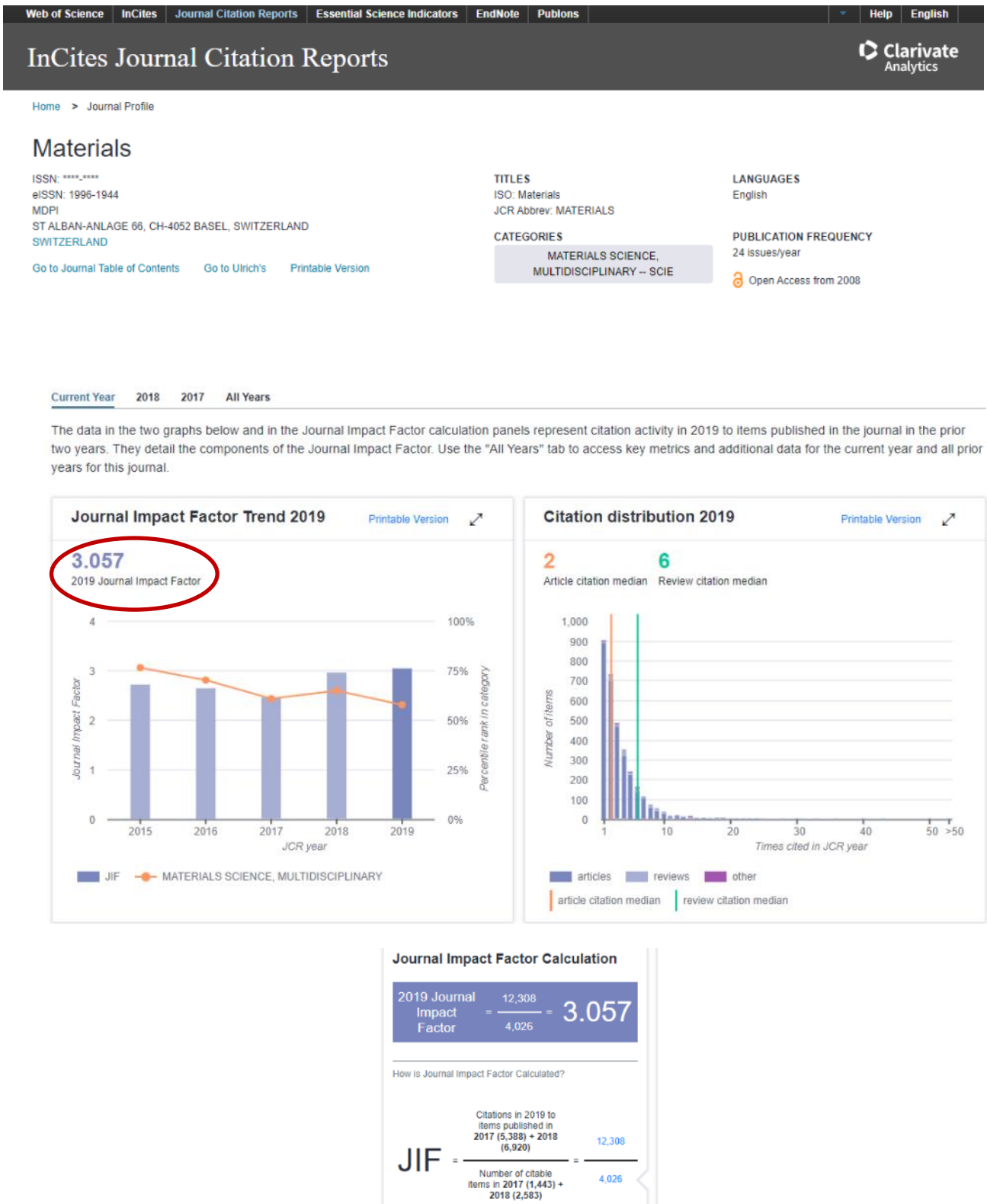
SOURCE METRICS		
Citable Items	1,332	Trend
% Articles in Citable Items	96.62	Trend
Average JIF Percentile	59.545	Trend
Cited Half-Life	2.1	Trend
Citing Half-Life	7.8	Trend

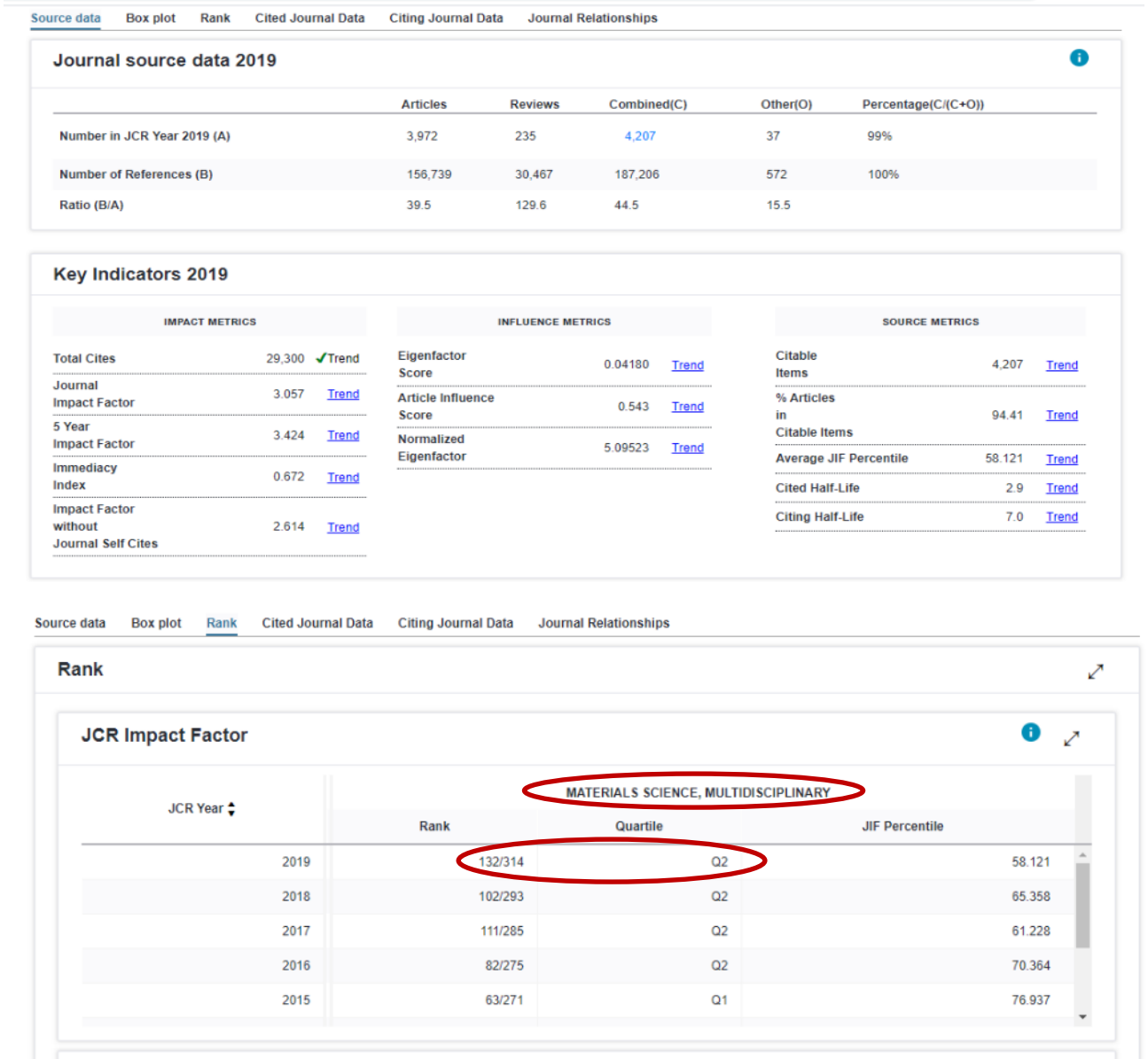
Rank

JCR Impact Factor

JCR Year	MATERIALS SCIENCE, MULTIDISCIPLINARY			METALLURGY & METALLURGICAL ENGINEERING		
	Rank	Quartile	JIF Percentile	Rank	Quartile	JIF Percentile
2019	185/314	Q3	41.242	18/79	Q1	77.848
2018	148/293	Q3	49.659	18/76	Q1	76.974
2017	170/285	Q3	40.526	23/75	Q2	70.000
2016	127/275	Q2	54.000	13/74	Q1	83.108
2015	145/271	Q3	46.679	18/73	Q1	76.027

Apéndice E. Indicios de calidad del artículo "*Prediction of Mechanical Properties by Artificial Neural Networks to Characterize the Plastic Behavior of Aluminum Alloys*" [25]





Source data | Box plot | Rank | Cited Journal Data | Citing Journal Data | Journal Relationships

Journal source data 2019

	Articles	Reviews	Combined(C)	Other(O)	Percentage(C/(C+O))
Number in JCR Year 2019 (A)	3,972	235	4,207	37	99%
Number of References (B)	156,739	30,467	187,206	572	100%
Ratio (B/A)	39.5	129.6	44.5	15.5	

Key Indicators 2019

IMPACT METRICS		
Total Cites	29,300	✓Trend
Journal Impact Factor	3.057	Trend
5 Year Impact Factor	3.424	Trend
Immediacy Index	0.672	Trend
Impact Factor without Journal Self Cites	2.614	Trend

INFLUENCE METRICS		
Eigenfactor Score	0.04180	Trend
Article Influence Score	0.543	Trend
Normalized Eigenfactor	5.09523	Trend

SOURCE METRICS		
Citable Items	4,207	Trend
% Articles in Citable Items	94.41	Trend
Average JIF Percentile	58.121	Trend
Cited Half-Life	2.9	Trend
Citing Half-Life	7.0	Trend

Source data | Box plot | Rank | Cited Journal Data | Citing Journal Data | Journal Relationships

Rank

JCR Impact Factor

JCR Year	MATERIALS SCIENCE, MULTIDISCIPLINARY		
	Rank	Quartile	JIF Percentile
2019	132/314	Q2	58.121
2018	102/293	Q2	65.358
2017	111/285	Q2	61.228
2016	82/275	Q2	70.364
2015	63/271	Q1	76.937

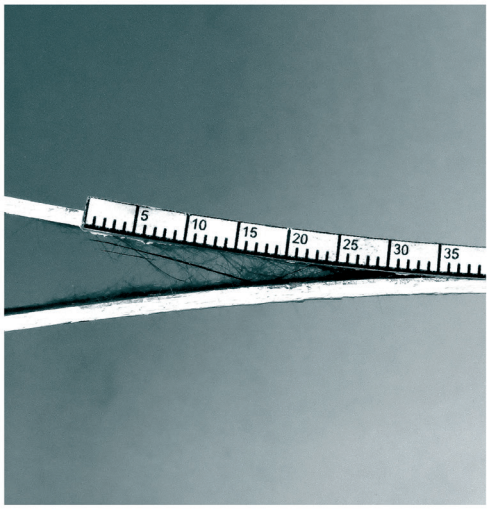
Apéndice F. Extracto del libro de resúmenes del "2nd International Conference on Materials Design and Applications (MDA2018)"

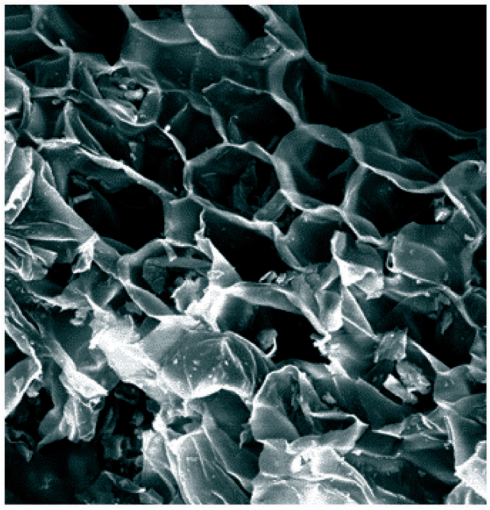
Book of abstracts of the

2nd International Conference on Materials Design and Applications 2018

5 - 6 July 2018

FACULTY OF ENGINEERING
UNIVERSITY OF PORTO
PORTO - PORTUGAL





Book of abstracts

MDA2018

2nd International Conference
on Materials Design and Applications

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Book of abstracts of the

2nd International Conference on Materials Design and Applications 2018

5 - 6 July 2018

FACULTY OF ENGINEERING
UNIVERSITY OF PORTO
PORTO - PORTUGAL

EDIÇÃO

Publindústria, Produção de Comunicação, Lda.

www.booki.pt

ISBN

978-989-8927-09-5

www.booki.pt

Utilize o seu SmartPhone para aceder automaticamente ao link através deste QR code.

Session 3A – Composites (Chair: R Campilho, V Tita)

53 Disastrous Technical Condition of the Material of New Cement Floors in Public Utility Building - Analysis of Causes **(MDA18_6)**

J Hoła, Ł Sadowski (Wrocław University of Science and Technology, Poland), A Hoła

54 Free Vibration and Torsional Buckling Analysis of Composite Drive Shaft Interleaved by Polyamide-6,6 (PA66) Nanfibers **(MDA18_8)**

B Beylergil (Sabanci University, Turkey)

55 Development of a Hybrid Aluminum Metal Matrix Composite Through Squeeze Casting Process **(MDA18_26)**

R Arunachalam (Sultan Qaboos University, Sultanate of Oman), I Al-Fori, R Muraliraja, M Al-Maharbi, S Piya

56 Effects of Fiber Treatment on the Properties of Epoxy Curaua-Reinforced Composites **(MDA18_31)**

FC Amorim, JFB Souza, JML Reis (Universidade Federal Fluminense, Brazil)

57 Analytical and Numerical Study for Selecting Polymeric Matrix Composites Intended to Demanding Nuclear Applications **(MDA18_50)**

D Merayo, Á Rodríguez-Prieto, AM Camacho (UNED, Spain)

58 Mechanical, Fire and Smoke Behaviour of Novel Hybrid Composites Based on Polyamide 6 with Carbon/Basalt Fibres **(MDA18_97)**

S Kuciel (Cracow University of Technology, Poland), K Mazur, K Sałasinska

Session 3B – Machining (Chair: JF Chatelain, A Jesus)

59 Effect of Additives on the Machinability of Glass Fiber Reinforced Polymer **(MDA18_4)**

B Lasseur, C Ouellet-Plamondon, J-F Chatelain (École de Technologie Supérieure, Canada)

60 Surface Integrity of Gamma Titanium Aluminides Milled with WC Tools **(MDA18_15)**

SD Castellanos (INEGI, Portugal), J Lino Alves, R Neto

61 Evaluation of Machining Defects in a Composite Laminate by Combining Non-Destructive and Tensile Testing **(MDA18_5)**L-A Généreux, G Lebrun, M Viens, J-F Chatelain (École de Technologie Supérieure, Canada)

62 Machinability of Titanium Aluminides: A Review **(MDA18_16)**

SD Castellanos (INEGI, Portugal), J Lino Alves, R Neto

63 Machinability Analysis of Leaded and Minimally-Leaded Brass Alloys **(MDA18_112)**

L Amaral, TE Silva (FEUP, Portugal), R Quinta, RMB Soares, SDC Villa, AMP de Jesus

64 Fracture Characterization of a Cast Aluminium Alloy Aiming Machining Simulation **(MDA18_138)**

TEF Silva (University of Porto, Portugal), S Gain, D Pinto, AMP de Jesus, J Xavier, PAR Rosa

Session 3C – Metals (Chair: SB Leen, AD Santos)

65 Inverse Methodology for Heat Transfer Coefficient Estimation in Cast Super Duplex Stainless Steel Parts of Complex Shape **(MDA18_84)**

RO Sousa (INEGI, Portugal), LMM Ribeiro, PJ Ferreira, AM Deus, I Felde

66 Crystal Plasticity Modelling of Ferrite Phases Effect on the Cyclic Response of IC-HAZ in a Welded 9Cr Martensitic Steel at High Temperature **(MDA18_87)**

M Li, PE O'Donoghue, SB Leen (National University of Ireland Galway, Ireland)

67 Effect of Ultrasonic Treatment on the Deformation Behaviour of AZ91D Magnesium Alloy at Room Temperature **(MDA18_90)**

IV Gomes, VE Lopes (CMEMS, Portugal), VH Carneiro, H Puga

Analytical and Numerical Study for Selecting Polymeric Matrix Composites Intended to Demanding Nuclear Applications

D. Merayo, Á. Rodríguez-Prieto, A.M. Camacho

Department of Manufacturing Engineering, UNED, Juan del Rosal 12, Madrid, Spain

This study describes a methodological proposal to select composite materials which are suitable to be employed to manufacture pipes that can properly withstand environments subjected to gamma and neutronic radiation. The methodology is used to select (see Figure 1) the optimal composite material whose properties are afterwards used to simulate several pipe sections by Finite Element Analysis, comparing the results with a nuclear ferritic steel used in the manufacture of primary loop components [1].

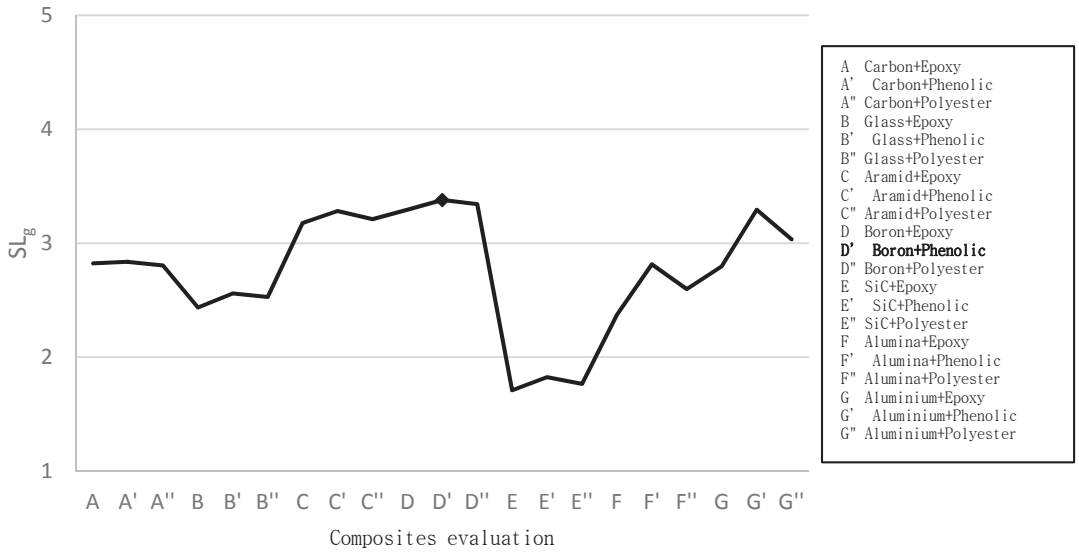


Figure 1. Global stringency level for each composite material

The most suitable composite material according to the defined criteria is formed of a phenolic resin matrix reinforced with long boron fibres and exhibit great properties to be used in a nuclear reactor environment: good radiation resistance and mechanical properties with a very low density and low cost. It can be concluded that, in some cases, composite material pipes could be an interesting alternative compared with usual ferritic steel pipes [2].

[1] Á. Rodríguez-Prieto, A. Camacho and M. Sebastián. Selection of candidate materials for reactor pressure vessels using irradiation embrittlement prediction models and a stringency level methodology, *P. I. Mech. Eng. L-J. Mat., 1-4,* 2017. DOI: 10.1177/1464420717727769

[2] K. Murty and I. Charit. Structural materials for Gen-IV nuclear reactors: Challenges and opportunities, *J. Nucl. Mater.,* 383, 1-2, 189-195, 2008.

MDA2018

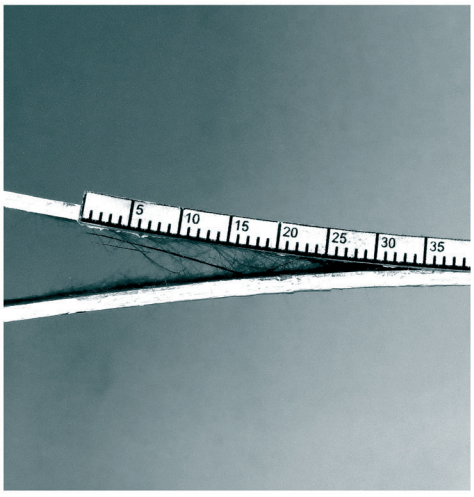
2nd International Conference on Materials Design and Applications

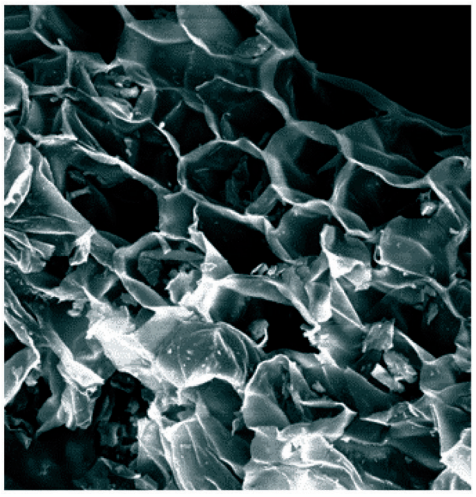
Book of abstracts

5 - 6 July 2018

FACULTY OF ENGINEERING
UNIVERSITY OF PORTO
PORTO - PORTUGAL

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO






CONTACTS
Lucas F. M. da Silva
Faculdade de Engenharia
da Universidade do Porto
Departamento de Engenharia Mecânica
Rua Dr. Roberto Frias
4200-465 Porto, Portugal
tel.: +351 225 081 716
fax: +351 225 081 445
email: lucas@fe.up.pt

Apéndice G. Extracto del libro de resúmenes del "The 8th Manufacturing Engineering Society International Conference (MESIC'19)"

ABSTRACT BOOK

8th Manufacturing Engineering Society International Conference| June 2019

ISBN 978-84-09-10387-4

Edited by: Cintia Barajas, Jesús Caja, Roque Calvo, Piera Maresca, Teresa Palacios

This Book contains the Abstracts as they were sent by the authors in text format to 8th MESIC Conference.

Book printed at the Publishing Services of ETSI Aeronáutica y del Espacio – UPM

Cover design: Laura Reyes

ISBN 978-84-09-10387-4

Editors: Cintia Barajas, Jesús Caja, Roque Calvo, Piera Maresca, Teresa Palacios

Comparative analysis of artificial intelligence techniques for material selection applied to manufacturing in Industry 4.0

Authors:
David Merayo Fernández [1], Álvaro Rodríguez Prieto [1], Ana María Camacho López [1]
[1]UNED

GOAL(S): This paper is intended to examine the most innovative techniques to select the optimal material to manufacture a product in Industry 4.0. The paper examines how the most relevant artificial intelligence techniques can assist the manufacturing process by selecting the optimal material.

INNOVATIVE ASPECTS: This paper introduces artificial intelligence (AI) assisting the material selection task, which has not been profusely examined as a tool to help manufacturing yet.
This paper examines the most relevant AI techniques to settle what are the most helpful in the manufacturing context of Industry 4.0.
This research tries to find synergies between the examined AI techniques.

BRIEF INTRODUCTION: Selecting the optimal material in manufacturing for an industrial application is becoming more and more complex and time-consuming due to the large amount of options which exist today. Many decision techniques have been developed but most of them are only intended to be used in certain fields or contexts and, in general, there is not a general way to proceed.
The application of artificial intelligence has proven to be a key tool when dealing with complex problems and, every day, new uses for this technology arise. However, the possibilities of these techniques have not yet been studied in detail within the world of manufacturing. At present, artificial intelligence can only be seen as a powerful assistant since, still, these systems require the supervision of human agents who validate their decisions. In any case, these are systems capable of greatly reducing the complexity of the tasks entrusted.
Since the birth of computing, a large number of methods and techniques have been developed that can be framed within the field of artificial intelligence and, of course, some of them will be more useful than others. Throughout this research, it is intended, among other things, to study their performance and capabilities.

METODOLOGHY: The paper includes a detailed overview of each relevant technique, which includes a list of advantages and disadvantages, mathematical explanation and its application to the selection of materials for manufacturing.
After that, combined techniques are introduced and, finally, the most performant techniques are compared.
Performance data are obtained by using relevant data and coding programs which allow us to study each technique and obtain metrics.

PERFORMANCES AND RESULTS: Techniques that could be framed within two categories are studied: Supervised Learning and Unsupervised Learning. Within the first group, the decision trees, the regressions in hyperplanes, the Support Vector Machines (SVM) and the neural networks are studied; In the second group, the Clustering Algorithms, the Principal Component Analysis (PCA), the Singular Value Decomposition (SVD) and the Independent Component Analysis (ICA) are analysed.
It is hypothesized that unsupervised techniques will present a higher performance in the selection of materials for manufacturing. However, composite techniques should yield even better results.

REFERENCES: Spector L. Evolution of artificial intelligence. Artificial Intelligence. 2006; 170 (18): 1251-1253.
Rodríguez-Prieto A., Camacho AM, Sebastián MA. Materials selection criteria for nuclear power applications: a decision algorithm. JOM. 2016; 2 (68): 496-506.
Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review. 1958; 65 (6): 386.
Dagli CH. Artificial Neural Networks for Intelligent Manufacturing. Springer Science & Business Media. 2012.
Spelt PF, Knee HE, Glover CW. Hybrid artificial intelligence architecture for diagnosis and decision-making in manufacturing. Journal of Intelligent Manufacturing. 1991; 2 (5): 261-268.

S

T

U

V

W

X

Y

Z

Apéndice H. Extracto del libro de resúmenes del "3rd International Conference on Materials Design and Applications (MDA2020)"

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

BOOK OF ABSTRACTS

MDA2020

3RD INTERNATIONAL CONFERENCE ON MATERIALS DESIGN AND APPLICATIONS

5-6 NOVEMBER 2020

FACULTY OF ENGINEERING - UNIVERSITY OF PORTO
PORTO - PORTUGAL

SPONSOR
INSTRON

BOOK OF ABSTRACTS OF THE

3RD INTERNATIONAL CONFERENCE ON MATERIALS DESIGN AND APPLICATIONS 2020 MDA 2020

PORTO - PORTUGAL
5-6 NOVEMBER 2020

EDIÇÃO

Quântica Editora, Lda.

www.quanticaeditora.pt

ISBN

978-989-9017-41-2

www.quanticaeditora.pt

Utilize o seu *SmartPhone* para aceder automaticamente ao *link* através deste *QR code*.

93 Use of wood in the mobility sector – Opportunities, benefits and challenges **(MDA20_144)**
D Berthold (Fraunhofer Institute for Wood Research, Germany), C Burgold, N Ritter

94 Development and data analysis of a process-integrated pretension monitoring system for thermoset fiber composites in robotic coreless winding**(MDA20_71)**
P Mindermann (University of Stuttgart, Germany), SI Bodea, D Gubetini, B Rongen, A Menges, J Knippers, GT Gresser

95 Adhesive strength of gypsum composites with lightweight fillers **(MDA20_179)**
M Doleželová (Czech Technical University in Prague, Czech Republic), J Krejsová, A Vimmrová

96 **Session 4C – Metals IV**
(Chair: M Vieira, J Belinha)

97 Analysis and evaluation of mechanical descriptors from micro compression tests on spherical samples **(MDA20_105)**
H Sonnenberg (University of Bremen, Germany), B Clausen

98 Phase-field ductile fracture diffusive approach using residual control staggered solution method **(MDA20_160)**
E Azinpour (University of Porto, Portugal), AD Santos, JMA Cesar de Sá

99 Dilatometric study and microstructure evolution during one-step and double-step isothermal treatment of 4Mn multiphase automotive sheet steel **(MDA20_180)**
A Skowronek (Silesian University of Technology, Poland), A Grajcar, M Morawiec, C Garcia-Mateo, V Ruiz-Jimenez, K. Radwański

100 Influence of strain rate effects on temperature field distribution and mechanical behavior of dual-phase steel sheets **(MDA20_123)**
RL Amaral (INEGI, Portugal), AD Santos, SS Miranda, J César de Sá

101 Microstructure and mechanical performance of cast and heat treated Al-12%Si-piston alloy with rare earth metals (Sm, Tb and Ce) **(MDA20_134)**
MM Tash (Cairo University, Egypt), WM Khalifa, IS El-Mahallawi

Poster session

102 **Metals**

103 Poster 1
Prediction of the yield stress of aluminium alloys using Big Data and Artificial Neural Networks **(MDA20_59)**
D Merayo (UNED, Spain), A Rodriguez-Prieto, AM Camacho

104 Poster 2
Microstructure of Ti50Ni25Cu25 rapidly quenched ribbons crystallized by electric pulse treatment **(MDA20_116)**
NN Sitnikov (National Research Nuclear University MEPhI, Russia), AV Shelyakov, IA Khabibullina, KA Borodako, AA Diadechko

105 Poster 3
Seedless copper electroplating on CoW thin films in low pH electrolyte - early stages of formation **(MDA20_162)**
BMC Oliveira (INEGI, Portugal), RF Santos, F Viana, S Simões, P Alpuim, PJ Ferreira, M Vieira

106 Poster 4
Finite element implementation of the Cazacu ductile damage law for porous orthotropic hcp crystalline structure materials exhibiting strength differential effect **(MDA20_169)**
C Rojas-Ulloa (Universidad de La Frontera, Chile), V Tuninetti, AM Habraken

107 Poster 5
Effect of solution heat treatment conditions on the microstructure of a Ni-Si-B alloy **(MDA20_176)**
GM Gorito (University of Porto, Portugal), P Lacerda, M Vieira, LMM Ribeiro

108 Poster 6
Minisample tensile-compression testing for sheet metal mechanical characterization **(MDA20_117)**
DJ Cruz (INEGI, Portugal), AD Santos, RL Amaral, JC Xavier, JG Mendes, SS Miranda

109 Poster 7
Temperature-dependent mechanical stability of retained austenite and strain hardening behavior of thermomechanically-processed medium-Mn TRIP-aided steel **(MDA20_182)**
A Kozłowska (Silesian University of Technology, Poland), A Grajcar

110 Poster 8
Theoretical and experimental studies of bainite formation in lean medium-Mn steel **(MDA20_183)**
M Morawiec (Silesian University of Technology, Poland), A Grajcar, C Garcia-Mateo, V Ruiz-Jimenez

Prediction of the yield stress of aluminium alloys using Big Data and Artificial Neural Networks

D. Merayo, A. Rodriguez-Prieto, A.M. Camacho

Department of Manufacturing Engineering, UNED, Juan del Rosal 12, 28040, Madrid, Spain

In this work, a computer-aided tool is developed to predict the yield stress of aluminium alloys. The system is based on the use of artificial neural networks supported by a large dataset containing the most relevant chemical and mechanical properties of thousands of aluminium alloys. The volume of considered data exceeds 2k materials manufactured by more than 100 companies. The information is retrieved from an open online material library, filtered, organized and prepared to train the artificial neural network (ANN). An ANN has been defined to optimize the predictive capacity of the network. After the training, the ANN is able to predict the yield stress of aluminium alloys with an average confidence greater than 95%.

Aluminum alloys are some of the most relevant metallic materials of the industry and they play a very important role in some high-technology fields such as aerospace [1]. The elastic behaviour of these metals can be described by using two mechanical properties: Young's modulus and yield stress. The yield stress is also very useful for the analysis of metal forming processes. Artificial neural networks fed by an appropriate dataset have proven to be able to predict the Young's modulus of metals in a very performant way [2]. Material libraries can be used to extract trends or general characteristics; however, as the amount of data to be considered increases, the task becomes painful or even impossible [3]. The artificial neural network developed in this work is able to predict the yield stress of an aluminium alloy with errors less than 5%.

[1] M. Danylenko, Aluminium Int. Today, 31, 4, 35 (2018).
[2] D. Merayo, A. Rodriguez-Prieto and A.M. Camacho, IEEE Access, 8, 13444-13456 (2020).
[3] S.K. Pal, S.K. Meher, and A. Skowron, Pattern Recognit. Lett., 67, 109–112 (2015).

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

BOOK OF ABSTRACTS

MDA2020

3RD INTERNATIONAL CONFERENCE ON MATERIALS DESIGN AND APPLICATIONS

5-6 NOVEMBER 2020

FACULTY OF ENGINEERING - UNIVERSITY OF PORTO
PORTO - PORTUGAL

SPONSOR
INSTRON