



Modelo matemático multivariante para estimar el valor CTOD en función de propiedades mecánicas, químicas y microestructurales aplicado a uniones soldadas para la industria Offshore

Álvaro Presno Vélez

Tesis Doctoral

Programa Doctoral: Ingeniería de producción y computación

Universidad de León

Director: Don Antonio Bernardo Sánchez

León, 2019



El Doctor D. Antonio Bernardo Sánchez, como Directores de la Tesis Doctoral terminada y titulada "Modelo matemático multivariante para estimar el valor CTOD en función de propiedades mecánicas, químicas y microestructurales aplicado a uniones soldadas para la industria Offshore" realizada por D. Álvaro Presno Vélez en el programa de *Doctorado en Ingeniería de producción y computación*, informa favorablemente ante el depósito de esta, reuniendo todas las condiciones necesarias para su defensa.

El firmante, en León a 25 de Noviembre de 2010

D. Antonio Bernardo Sánchez

### Resumen

Las solicitaciones mecánicas cada vez más grandes a las que son sometidas las estructuras en alta mar han puesto de relevancia la importancia del control de calidad, basado en la mecánica de fractura, de las uniones soldadas. Para este propósito, el CTOD (conocido como desplazamiento de apertura del frente de grieta) se ha considerado uno de los parámetros más adecuados para modelar y controlar el posible crecimiento de grietas y es usado ampliamente a nivel industrial. A través de técnicas de análisis multivariante, se ha modelado la relación entre el valor CTOD y otras propiedades del material (tales como dureza, composición química, tenacidad o morfología microestructural) aplicadas a uniones soldadas de acero de alta resistencia para estructuras en alta mar. Para crear este modelo, se realizaron cientos de pruebas en 72 muestras reales soldadas con una amplia gama de parámetros industriales reales. Los resultados obtenidos han sido procesados y evaluados con diferentes técnicas matemáticas, y se ha podido establecer la importancia de todas las variables explicativas elegidas y la buena capacidad predictiva del modelo dentro de los límites del conjunto experimental. Al establecer el uso de este modelo, se podrían lograr ahorros significativos en la fabricación de estas estructuras ya que las pruebas CTOD son más caras y complejas que las alternativas propuestas. Además, este modelo permite algunas conclusiones técnicas importantes para la industria fabricante de estructuras Offshore.

**Palabras clave:** CTOD; acero; soldadura o uniones soldadas; modelos de regresión multivariantes; redes neuronales; estructuras Offshore

### **Abstract**

The increasingly mechanical requirements of the offshore structures have established the relevance of fracture mechanics-based quality control in welded joints. For this purpose, Crack-Tip Opening Displacement (CTOD) at a given distance from Crack-tip has been considered one of the most suited parameters for modelling and control of crack growth and is it broadly used at the industrial level. We have modelled through multivariate analysis techniques the relationship among CTOD values and other material properties (as hardness, chemical composition, toughness or microstructural morphology) in high-thickness offshore steel welded joints. In order to create this model, hundreds of tests were done in 72 real samples welded with a wide range of real industrial parameters. The results obtained have been processed and evaluated with different multivariate techniques, and we have established the significance of all the chosen explanatory variables and the good predictive capability of the CTOD value within the limits of the experimental set variation. By establishing the use of this model, significant savings are achieved in the manufacturing of wind generator as CTOD tests are more expensive and complex than the proposed alternatives. Besides, this model allows for some technical conclusions.

**Keywords:** crack-tip opening displacement; steels; welded or bonded joints; multivariate regression model; marine structures

### **Agradecimientos**

-Ad maiorem *Dei* gloriam-

Me gustaría agradecer al director de este trabajo, Don Antonio Bernardo (Universidad de León) su continua tutela, ayuda y guía. También quisiera agradecer a mi querida familia, amigos y compañeros de trabajo su ayuda, paciencia y comprensión durante todos estos años.

Finalmente quisiera dedicar este trabajo a mi prometida, Ana, por navegar conmigo a través de aguas revueltas.

**Índice de contenidos**

1	Introducción .....	17
1.1	Presentación de las estructuras Offshore dentro de la industria eólica y su relevancia dentro del mercado energético global .....	17
1.1.1	El aerogenerador o OWT .....	20
1.1.2	Diferentes estructuras de soporte .....	21
1.2	Introducción a las principales sollicitaciones mecánicas consideradas durante la fase de diseño de estas estructuras .....	23
1.2.1	Cargas ambientales .....	24
1.2.2	Cargas por oleaje.....	24
1.2.3	Olas lineales regulares y Teoría de la onda de Airy .....	24
1.2.4	La ecuación de Morison.....	25
1.2.5	Carga de viento .....	27
1.2.6	Carga de hielo .....	28
1.2.7	Cargas sísmicas .....	28
1.2.8	Carga accidental.....	29
1.2.9	Carga muerta.....	29
1.2.10	Carga viva .....	29
1.3	Relevancia de la fatiga y el desgaste en la integridad de las estructuras Offshore. ...	29
1.4	Introducción al CTOD (dentro del contexto del presente trabajo) y su uso como parámetro de control del comportamiento a fatiga .....	31

## MODELO MATEMÁTICO: CTOD-PROPIEDADES DE LOS MATERIALES

1.5	Enunciado del problema, hipótesis y punto de partida para la realización del presente trabajo académico .....	32
1.5.1	Hipótesis de trabajo.....	33
1.5.2	Objetivos a alcanzar .....	34
2	Revisión bibliográfica.....	35
2.1	Ensayo CTOD, cálculo y modelos de fractura.....	35
2.2	Selección de las variables explicativas (input) .....	38
2.2.1	Microestructura (y su relación con otras variables).....	39
2.2.2	Composición química (y su relación con otras variables) .....	42
2.2.3	Resistencia mecánica (y su relación con otras variables) .....	44
2.2.4	Resiliencia (y su relación con otras variables).....	45
2.3	Análisis predictivo .....	46
2.3.1	Preselección de métodos apropiados .....	47
2.3.2	Método de regresión lineal.....	47
2.3.3	Árboles de regresión (aprendizaje computacional).....	48
2.3.4	Curvas de regresión adaptativa multivariable.....	48
2.3.5	Redes neuronales (aprendizaje computacional).....	48
2.3.6	Descartados.....	48
2.3.7	Modelos de regresión lineal multivariante.....	49
2.3.8	MARS (Multivariate Adaptive Regression Splines).....	53
2.3.9	Árboles de clasificación o decisión.....	55
2.3.10	Modelos lineales generalizados .....	58
2.3.11	Componente aleatoria .....	59

## MODELO MATEMÁTICO: CTOD-PROPIEDADES DE LOS MATERIALES

2.3.12	Componente sistemática .....	59
2.3.13	Función de enlace .....	59
2.4	Redes neuronales .....	60
2.4.1	El Perceptrón multicapa.....	62
2.4.2	Fases en la aplicación de un perceptrón multicapa.....	64
3	Metodología y fase experimental.....	65
3.1	Aproximación al diseño experimental .....	65
3.2	Muestreo y naturaleza de las muestras, obtención y controles de calidad preliminares (incluyendo las variables de soldadura utilizadas).....	69
3.3	Ensayos CTOD .....	73
3.4	Ensayos de resistencia mecánica .....	81
3.5	Ensayo de composición química.....	84
3.6	Ensayos para determinar la tenacidad.....	86
3.7	Ensayos para caracterizar la microestructura.....	88
3.8	Resumen de los datos obtenidos durante la fase experimental .....	92
4	Análisis .....	93
4.1	Análisis preliminar de los datos.....	93
4.2	Modelos de regresión lineal .....	96
4.2.1	Modelo de regresión lineal 1.....	96
4.2.2	Modelo de regresión lineal 2.....	100
4.2.3	Modelos de regresión lineales 3 y 4.....	101



## MODELO MATEMÁTICO: CTOD-PROPIEDADES DE LOS MATERIALES

4.3	Splines de regresión adaptativa multivariante (MARS) .....	105
4.4	Modelos de regresión lineal generalizada.....	108
4.5	Árboles de regresión .....	109
4.6	Modelo con Redes Neuronales .....	112
5	Discusión.....	113
6	Resumen.....	115
7	Conclusiones .....	117
8	Bibliografía .....	119
9	Anexos .....	127

**Lista de figuras**

*Figura 1.* (a) Uso global de electricidad por sectores económicos principales expresado en 1024 BTU y su evolución estimada hasta el 2050. (b) Estimación de la evolución de producción energética por tipo (renovables, combustibles fósiles líquidos, gas natural, carbón y energía nuclear) expresado en 1024 BTU hasta el año 2050..... 17

*Figura 2.* (a) Imagen del análisis geoespacial publicado por el Imperial College (Londres) y el IEA donde se marcan en color las zonas de alto potencial eólico (tonos naranjas para las aguas profundas y tonos azules para las aguas poco profundas). (b) Zonas de instalación de los últimos proyectos eólicos marinos..... 19

*Figura 3.* Estructura “Jacket” de los aerogeneradores del proyecto “Wikinger” antes de la fase de instalación. .... 23

*Figura 4.* Ejemplo de curva de diseño CTOD ..... 32

*Figura 5.* Arriba-Izquierda, cambio en la temperatura de transición dúctil-frágil para diferentes composiciones de carbono. Abajo-Izquierda, cambio en la temperatura de transición dúctil-frágil para diferentes composiciones de Manganeso. Derecha, cambios en el límite elástico con respecto a la temperatura y diferentes cantidades de Fósforo..... 43

*Figura 6.* Resistencia mecánica frente a dureza Vickers ..... 45

*Figura 7.* Ejemplo de árbol CART..... 56

*Figura 8.* Modelo de neurona para NN..... 61

*Figura 9.* Posición de extracción de las probetas y zonas de ensayo. Las zonas de color marcan áreas específicas para el análisis microestructural, dureza HV10 y químico. .... 68

## MODELO MATEMÁTICO: CTOD-PROPIEDADES DE LOS MATERIALES

<i>Figura 10.</i> Sección cilíndrica de una torre Offshore para un prototipo de aerogenerador de 8MW. Las probetas de homologación de la soldadura utilizada para unir la virola (soldadura longitudinal) y la virola a la brida (soldadura circular) fueron parte de set experimental.....	69
<i>Figura 11.</i> Detalle del tratamiento térmico de precalentamiento para la unión de una de las soldaduras de prueba.....	70
<i>Figura 12.</i> Macrografía de una de las uniones soldadas del set experimental que fue rechazada al no superar las pruebas no destructivas. En este caso el ensayo por ultrasonidos. En el corte macrográfico puede observarse una falta de fusión en la raíz de la soldadura .....	73
<i>Figura 13.</i> Diagrama de un proceso de pre-agrietado para el ensayo CTOD. El eje vertical se subdivide en; fuerza [KN], factor intensidad de tensiones [MPa.m <sup>1/2</sup> ], amplitud de movimiento [mm].....	75
<i>Figura 14.</i> Parte de la colección de probetas (72) utilizadas para la determinación de la variable CTOD (en la imagen las correspondientes a espesor 25 mm). .....	77
<i>Figura 15.</i> Detalle (1) de la superficie de fractura de una de las probetas CTOD. ....	78
<i>Figura 16.</i> Detalle (2) de la superficie de fractura de una de las probetas CTOD. ....	78
<i>Figura 17.</i> Detalle (3) de la superficie de fractura de una de las probetas CTOD. ....	79
<i>Figura 18.</i> (a), (b) y (c) Modelos tridimensionales sobre las superficies de fractura con un brazo de metrología láser. La reflectividad de la superficie de fractura resultó no ser adecuada para este sistema de medición y se decidió utilizar sistemas ópticos. ....	80
<i>Figura 19.</i> Imágenes secuenciales (de izquierda a derecha) mostrando el crecimiento de grieta en dos de las probetas CTOD. Para realizar estas fotografías se detuvo el proceso de “pre-cracking” en distintos momentos.....	81

## MODELO MATEMÁTICO: CTOD-PROPIEDADES DE LOS MATERIALES

<i>Figura 20.</i> En esta imagen se pueden observar una serie de indentaciones de dureza HV10 realizadas sobre una muestra transversal a una unión soldada..	83
<i>Figura 21.</i> Indentación real fotografiada a través del ocular del sistema semiautomático de medición de las dos diagonales principales.	84
<i>Figura 22.</i> Tabla comparativa para el porcentaje de fractura dúctil-frágil.	87
<i>Figura 23.</i> Micrografía a 100X (izquierda) y 500 X (derecha).	90
<i>Figura 24.</i> Micrografía a 100X de una muestra aceptada para el set experimental. Imágenes utilizadas para la determinación del tamaño y morfología de grano.	91
<i>Figura 25.</i> Micrografía a 500X. Detalle de la misma muestra que la anterior.	91
<i>Figura 26.</i> Correlación, KDE (estimación de densidad del Kernel) y diagramas de dispersión (también se representa la línea de tendencia que mejor se ajusta con una regresión lineal en color azul) entre las diferentes variables.	94
<i>Figura 27.</i> Cuantiles de la muestra con respecto a los cuantiles normales estandarizados	96
<i>Figura 28.</i> Distribución normal de los residuos (Arriba) Gráfica de probabilidad normal de los residuos (Abajo) Gráfica de residuos con respecto a los valores ajustados.	103
<i>Figura 29.</i> Efecto de cada variable explicativa sobre el CTOD: (a) Resistencia mecánica, (b) Tenacidad, (c) Microestructura y (d) Composición química. Se ha marcado un valor de cada variable explicativa de 1,0000 (estandarizado) y se representa como varía el CTOD (eje vertical).	105
<i>Figura 30.</i> (a) Diagrama del modelo de Splines de regresión adaptativa multivariante (MARS) para dos de las variables explicativas junto con sus ubicaciones de nudos (arriba) y (b) la función de análisis de varianza (ANOVA) para los pares CTOD-Resistencia mecánica y (c) CTOD-Microestructura.	107

## MODELO MATEMÁTICO: CTOD-PROPIEDADES DE LOS MATERIALES

<i>Figura 31.</i> Árbol de regresión para el conjunto de todos los datos disponibles (después de eliminar los Outliers). .....	110
<i>Figura 32.</i> Árbol de regresión con una muestra de prueba de 13 observaciones .....	111
<i>Figura 33.</i> Modelo de redes neuronales perceptrón con una capa oculta.....	112
<i>Figura 34.</i> Modificación de un diagrama de dispersión 3D .....	116

**Lista de tablas**

Tabla 1. *Selección no exhaustiva de variables con probado efecto sobre las propiedades de los materiales metálicos* ..... 39

Tabla 2. *Resultados preliminares de un diseño de experimentos factorial*..... 67

Tabla 3. Cuadro resumen de los principales rangos de parámetros de soldadura utilizados en la fabricación de los cupones de prueba ..... 71

Tabla 4. *Extracto de los datos recogidos (Muestra 01 y 72)* ..... 92

Tabla 5. *Modelo de regresión lineal 1  $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$* ..... 98

Tabla 6. *Resultados de la validación cruzada* ..... 99

Tabla 7. *Comparación de los valores de R-squared (RSQ) y el error cuadrático medio (RMSE)*  
..... 100

Tabla 8. *Modelo de regresión lineal 2  $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_4X_4$* ..... 101

Tabla 9. *Modelo de regresión lineal 3  $Y = \beta_0 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$* ..... 102

Tabla 10. *Modelo de regresión lineal 4  $Y = \beta_0 + \beta_1X_{12} + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$* . .... 104

Tabla 11. *Resultados del Modelo lineal generalizado 1* ..... 108

Tabla 12. *Resultados del Modelo lineal generalizado 2 sin la variable explicativa “Resistencia mecánica”* ..... 109

**Lista de anexos**

Anexo 1. Certificado de publicación del artículo .....	127
Anexo 2. Carta de aceptación para la presentación de resultados parciales en la IX Conferencia Internacional de Aprovechamiento de Recursos Minerales “CINAREM 2017” .....	127
Anexo 3 Referencia a las normas de ensayo.....	127
Anexo 4. Artículo publicado en la Revista ENERGIES en su número especial BIG DATA AND MACHINE LEARNING .....	127

**Lista de abreviaturas y acrónimos**

ASM	American Society for Metals
CE index	Índice del Carbono equivalente
CVN	Charpy con entalla en V
CTOD	Crack tip opening displacement
e	Error
FL	Línea de Fusión
PWHT	Post Weld Heat Treatment
RMSE	Root-Mean-Square Error
RSQ	R-squared
SEM	Microscopio Electrónico de barrido
UTS	Carga de rotura
YS	Límite elástico
$X_1, X_2, \dots, X_K$	Variables explicativas
$\beta_1, \beta_2, \dots, \beta_K$	Coefficientes de regresión

*(Otras abreviaturas o acrónimos están explicados dentro del propio texto)*



## 1 Introducción

### 1.1 Presentación de las estructuras Offshore dentro de la industria eólica y su relevancia dentro del mercado energético global

Las principales predicciones de la evolución de la demanda energética durante los próximos años calculan un aumento de esta de hasta el 50% entre el año 2018 y 2050 (International Energy Outlook 2019, IEO, 2019). La mayor parte de este crecimiento corresponde a países en los que se prevé una alta tasa de crecimiento demográfico y desarrollo industrial, principalmente en el sudeste asiático y que, en general, no son miembros de la OECD (Organization for Economic Cooperation and Development).

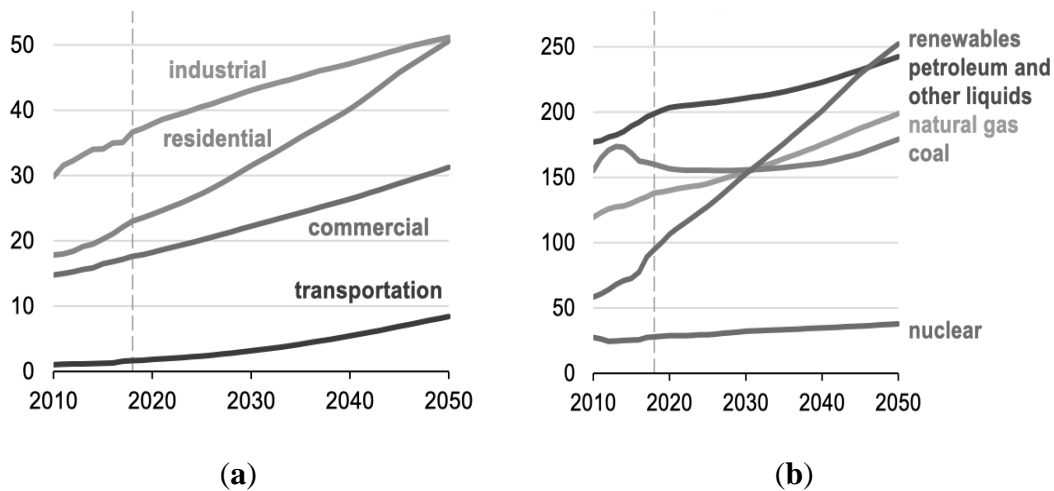


Figura 1. (a) Uso global de electricidad por sectores económicos principales expresado en 1024 BTU y su evolución estimada hasta el 2050. (b) Estimación de la evolución de producción energética por tipo (renovables, combustibles fósiles líquidos, gas natural, carbón y energía nuclear) expresado en 1024 BTU hasta el año 2050.

Fuente: (International Energy Outlook, 2019)

De estas predicciones se desprende la importancia, creciente, del desarrollo de las distintas tecnologías de generación de energía eléctrica. De entre ellas, las energías renovables se presentan con unas expectativas de crecimiento muy superiores a otras alternativas (Figura 1).

El número limitado de recursos renovables en tierra, los problemas de explotación de estas áreas y las posibles interferencias con otras actividades han desviado la atención hacia el aprovechamiento de la energía eólica desde instalaciones situadas de forma temporal o permanente en el mar. En el mar no existen limitaciones de impacto sonoro que limiten la velocidad del extremo de las palas, el perfil de velocidad del viento está menos distorsionado por la orografía o los obstáculos y el espacio disponible permite colocar los aerogeneradores en configuraciones más ventajosas (entre otras ventajas). El mercado global de energía eólica marina creció casi un 30% por año entre 2010 y 2018, beneficiándose de las rápidas mejoras tecnológicas, y cerca de 150 nuevos proyectos de energía eólica marina están en desarrollo activo en todo el planeta. Europa en particular ha fomentado el desarrollo de esta tecnología, liderado por el Reino Unido, Alemania y Dinamarca (*Offshore Wind Outlook, 2019*; World Energy Outlook Special Report) donde los recursos eólicos de alta calidad y las aguas relativamente poco profundas han proporcionado condiciones excepcionalmente buenas para desarrollar las tecnologías eólicas marinas y su traslado eficaz al mercado energético (Figura 2). Sin embargo, y con la intención de desbancar el Eurocentrismo de la industria Offshore, es importante señalar que ha sido China quien ha aumentado su capacidad más que cualquier otro país durante el último año (*Energy Information Administration, EIA, 2018*)

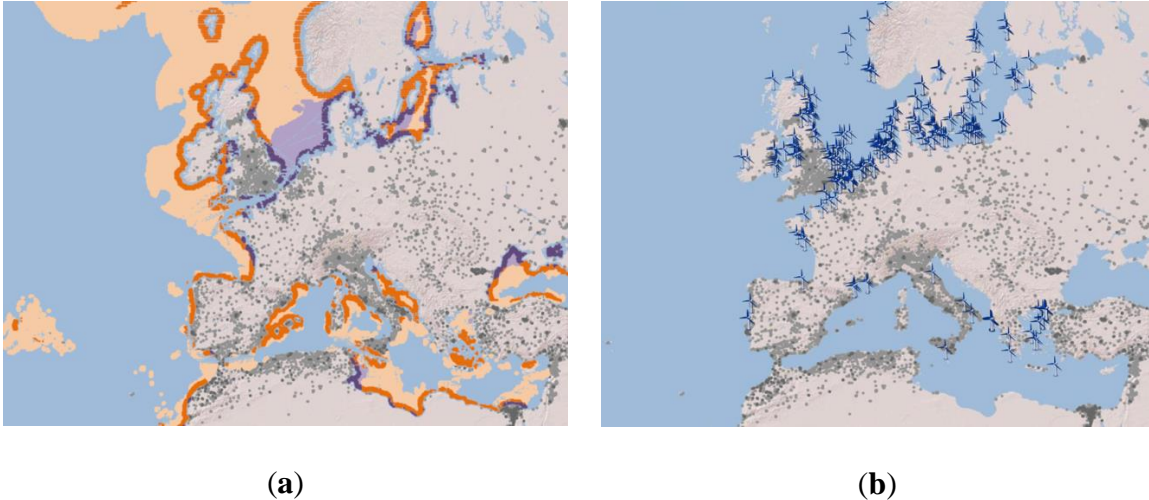


Figura 2. (a) Imagen del análisis geoespacial publicado por el Imperial College (Londres) y el IEA donde se marcan en color las zonas de alto potencial eólico (tonos naranjas para las aguas profundas y tonos azules para las aguas poco profundas). (b) Zonas de instalación de los últimos proyectos eólicos marinos.

Las predicciones sugieren que, en particular, este mercado se expandirá significativamente durante las próximas dos décadas, creciendo a una tasa del 13% anual en lo que el último informe de la Agencia Internacional para la Energía denomina escenario de políticas declaradas (es decir, aquel escenario de referencia en el que no se contemplan grandes cambios en las políticas energéticas actuales). Reforzada por los objetivos de la política y la caída de los costes tecnológicos, se prevé que solo la capacidad eólica mundial situada en aguas profundas (más de 60m hasta el lecho marino) aumente hasta en un 1500% antes del año 2040, convirtiéndose en una industria de aproximadamente  $10^9$ € en las próximas dos décadas. Este nivel de inversión implica que la industria eólica Offshore representa en la actualidad el 10% de la inversión en plantas de energía renovables a nivel mundial.

Como nota marginal cabe decir que el aprovechamiento de la energía eólica no es en absoluto un desarrollo reciente, ya a principios del pasado siglo numerosos proyectos de investigación

exploraron la posibilidad de generar electricidad a través del viento. La generación a gran escala de electricidad a partir del viento comenzó a principios de la década de 1980 y puede considerarse en la actualidad como una industria madura y sólida.

### **1.1.1 El aerogenerador o OWT**

A pesar de la complejidad actual de las plantas de generación eólica, el dispositivo principal utilizado en la conversión de la energía eólica en energía eléctrica es la turbina eólica, aerogenerador marino o OWT (Offshore Wind turbines). Estas OWT son artefactos de grandes dimensiones que se sitúan en zonas de alto potencial eólico a una altura suficiente para aprovechar los gradientes de potencial y mantener las distancias de seguridad razonables para el funcionamiento seguro de las enormes partes móviles (álabes o palas). Está fuera del objeto de este trabajo profundizar en el diseño o funcionamiento de estos dispositivos, pero si debe señalarse que, en líneas generales, se considera que, desde el punto de vista de eficiencia y productividad, a mayor tamaño del dispositivo mejores resultados y capacidades. Esto a propiciado una evolución constante en el potencial generador (y en el volumen, peso y tamaño) de las OWT durante los últimos años. Por lo tanto, los aerogeneradores están sometidos a condiciones de carga cada vez más severas; debido a las mayores velocidades del viento, las condiciones extremas de oleaje y corriente y las adversidades climáticas a las que deben enfrentarse. Esto tiene como consecuencia la importancia de la integridad estructural de las torres, estructuras y cimentaciones que sostienen, bajo las mencionadas sollicitaciones, a las OWT, y las tecnologías de diseño, inspección y fabricación que permiten a estos soportes hacerlo con fiabilidad.

### 1.1.2 Diferentes estructuras de soporte

Atendiendo con mayor detalle a estas estructuras, “Design of Offshore Wind Turbine Structures” (DNVGL-OS-J101, 2007) clasifica en cinco tipologías principales a las principales soluciones estructurales que sostienen estas OWT. Las estructuras “monopile” o monopilar, estructuras en forma de trípode (utilizando generalmente cimentaciones con pilotes), estructuras de celosía, estructuras “Jacket” (formadas por perfiles tubulares en estructuras tridimensionales) y estructuras de gravedad (formadas por grandes masas de cemento que utilizan su gran volumen y masa para garantizar la integridad estructural).

De entre ellas, las bases de monopilares son las estructuras de soporte más utilizadas. La mayoría de los parques eólicos marinos situados en aguas poco profundas utilizan esta tecnología ya que tienen la ventaja de un diseño simple para el proceso de fabricación. Sin embargo, el fracaso de las conexiones agrupadas entre los monopilares y la pieza de transición (la pieza responsable de conectar el monopilar a la torre de la turbina) ha demostrado ser una desventaja que imposibilita, en la práctica, el uso de esta solución para turbinas de más de 5 MW de capacidad.

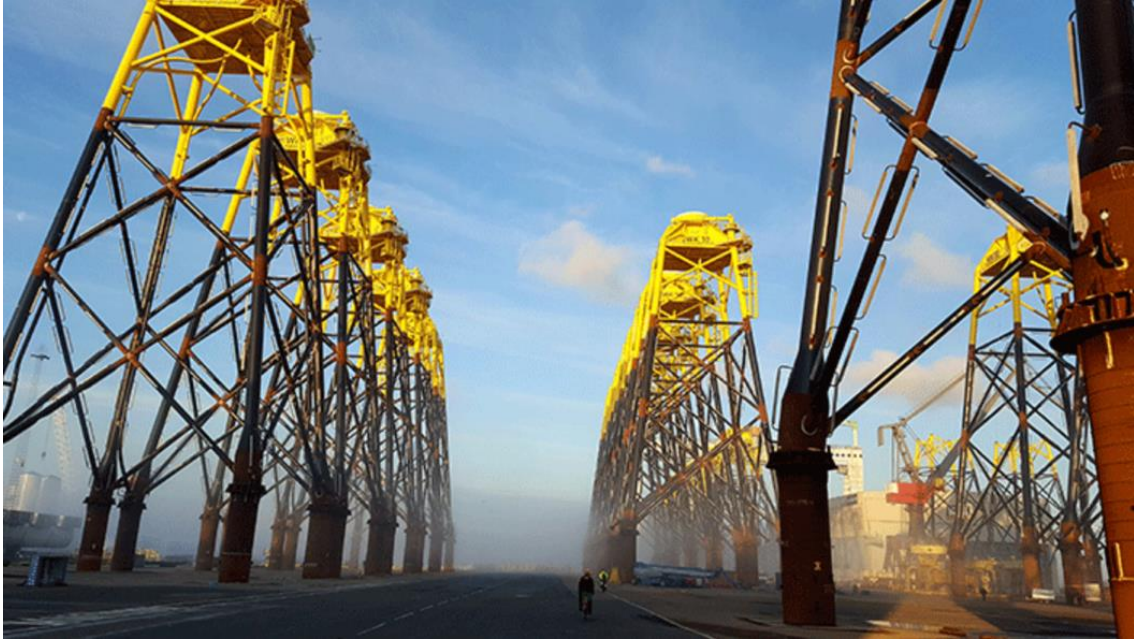
Las bases de gravedad son el segundo tipo de estructura de soporte más habitual. Este tipo de base logra su estabilidad únicamente al proporcionar suficientes cargas muertas por medio de su propia masa. En ocasiones se introducen cargas adicionales bombeando arena, cemento, roca o mineral de hierro para agregar peso extra. Las estructuras de gravedad son especialmente adecuadas para cargas ambientales modestas, como la carga producto del oleaje ligero o cuando es técnicamente factible añadir lastre adicional de forma sencilla.

Se han utilizado principalmente para soportar turbinas pequeñas, en aguas poco profundas o cerca de lugares costeros con un fondo marino rocoso donde la operación de “hincado” de los

monopilares es extremadamente complicada y costosa. Sin embargo, para profundidades de agua superiores a 35m es necesario buscar otro tipo de soluciones.

Una de las posibles opciones es el uso de estructuras flotantes. Estas estructuras son especialmente competitivas en aquellas localizaciones donde el lecho marino se encuentra a gran profundidad (y donde la profundidad hace que las estructuras convencionales con soporte en el fondo marino no sean económicamente competitivas) (DNV-OS-J103, "Design of Floating Wind Turbine Structures"). También las fundaciones y estructuras "Jacket" proporcionan una solución factible para aguas profundas (> 60 m).

En cualquier caso, hay una serie de características comunes a todos los diseños considerados; están contruidos utilizando acero, generalmente con perfiles tubulares en estructuras tridimensionales, simétricas con respecto a un eje de revolución vertical y sus componentes se unen entre si utilizando técnicas de soldadura por fusión.



*Figura 3.* Estructura “Jacket” de los aerogeneradores del proyecto “Wikinger” antes de la fase de instalación.

Fuente. Fotografía usada con el permiso del autor

## **1.2 Introducción a las principales solicitaciones mecánicas consideradas durante la fase de diseño de estas estructuras**

Se puede deducir de lo ya mencionado que las solicitaciones a las que están sometidas las estructuras Offshore son un factor decisivo para su diseño, especialmente en el caso de las torres, soportes y anclajes que cumplen funciones estructurales. De nuevo, un análisis exhaustivo está fuera del alcance de este trabajo sin embargo es importante considerar el origen principal de estas cargas y su naturaleza para comprender la especial relevancia que los criterios de mecánica de fractura tendrán sobre las etapas de diseño, fabricación y control de calidad.

En el diseño de una estructura Offshore, se deben considerar las siguientes cargas principales:

### 1.2.1 Cargas ambientales

Tales como la carga de las olas, las cargas del viento, las cargas de hielo, la carga del nivel del agua y las cargas por seísmos.

### 1.2.2 Cargas por oleaje

La carga producto de las fuerzas del oleaje son el tipo más importante de cargas ambientales en estructuras Offshore (DNVGL-OS-J101, 2007). Están diseñados considerando la máxima carga de ocurrencia-frecuencia que tenga en cuenta los escenarios de oleaje y clima para los últimos 50 y/o 100 años, de modo que se alcance un nivel de seguridad razonable. En general estas instalaciones son una inversión de capital considerable que solo puede resultar rentable con largos periodos de vida útil.

De acuerdo con *Design of Offshore Wind Turbine Structures* (DNVGL-OS-J101, 2007) debe utilizarse para el cálculo de esfuerzos por oleaje una teoría de olas reconocida, teniendo en cuenta la profundidad hasta el lecho marino y el rango de validez de dicha teoría. El tamaño, el perfil y el tipo de estructura y como interactúa con la superficie del mar también debe ser ponderado.

### 1.2.3 Olas lineales regulares y Teoría de la onda de Airy

La función que modela el oleaje lineal puede definirse como una función de onda sinusoidal variable en el tiempo y representarse de acuerdo con la siguiente ecuación (o alguna modificación de la siguiente ecuación):

$$\eta(x, t) = \frac{H}{2} \cos(kx - \omega t) \quad (1)$$



Donde  $H$  es la altura de la ola,  $t$  es el periodo y  $x$  la longitud de onda. Esta teoría, a menudo denominada teoría de ondas lineales, ofrece una descripción linealizada de la propagación del flujo de ondas de gravedad en la superficie del agua utilizando el enfoque de flujo potencial (Dean & Dalrymple, 1991).

La teoría de la onda de Airy calcula al perfil de presión dinámica, velocidades y aceleración de las partículas en las olas considerándolas como ondas armónicas, regulares y lineales. Se utiliza con frecuencia en lugares donde la profundidad del agua es mucho mayor que la longitud y altura de la onda que simula el oleaje (lo cual es el caso de la mayor parte de las estructuras Offshore). Para el uso de este modelo se supone que el nivel medio del agua es uniforme, con un flujo invisible, incompresible e irracional. La precisión de la teoría es suficiente para la aplicación de estructuras en alta mar en aguas poco profundas bajo el efecto de oleaje aleatorio, e incluso puede usarse para considerar algunas propiedades no lineales de ondas de segundo orden.

#### **1.2.4 La ecuación de Morison**

Para estructuras delgadas ( $\text{Diámetro/Longitud} < 0.2$ ) como los componentes de las estructuras “Jackets” y las estructuras monopilares, la ecuación de Morison se puede aplicar para calcular las cargas producto del oleaje. La fuerza en la dirección de propagación de las olas en un segmento de objeto cilíndrico sumergido en agua puede calcularse utilizando la ecuación semiempírica de Morison, donde la fuerza horizontal (por unidad de longitud en el miembro vertical) es la suma de dos componentes de fuerza; arrastre (proporcional al cuadrado de la velocidad de flujo) y de inercia (en fase con la aceleración local del fluido).

Se puede expresar como:

$$dF = dF_M + dF_D = C_M \rho \pi \frac{D^2}{2} \ddot{x} dz + C_D \rho \frac{D}{2} |\dot{x}| \dot{x} dz \quad (2)$$

Donde  $dF$  es la fuerza horizontal,  $dF_M$  es el componente por fuerza inercial,  $dF_D$  el componente de arrastre (drag),  $C_M$  y  $C_D$  coeficientes,  $D$  el diámetro del cilindro,  $\rho$  la densidad del fluido y  $z$  la altura con respecto a la superficie del mar.

Cabe señalar que la ecuación de Morison no tiene en cuenta el historial de flujo o que el vector de velocidad instantánea surge como una superposición de varios procesos de flujo, además, explica cierta no linealidad del flujo a través del término de arrastre (Morison et al., 1950). Las fuerzas de inercia que actúan sobre la estructura se deben al movimiento del cuerpo de agua sobre la estructura. Las fuerzas inerciales se pueden dividir en dos partes; fuerza de Froude-Krylov y la fuerza de masa añadida. Los coeficientes de arrastre e inercia son, en general, funciones del número de Reynolds, el número de Keulegan-Carpenter y la rugosidad relativa. El coeficiente también depende de la forma de la sección transversal de la estructura y de la orientación del cuerpo (Sarpkaya & Isaacson, 1981).

Para estructuras de gran volumen, para las cuales la cinemática de la onda se ve perturbada por la presencia de la propia estructura, se debe realizar un análisis de difracción de onda para determinar las cargas de onda locales y globales. Por otro lado para estructuras flotantes se deben incluir las fuerzas de radiación de onda. Tanto los efectos viscosos como los posibles efectos de flujo pueden ser importantes para determinar las cargas inducidas por las olas en una estructura de soporte.

### 1.2.5 Carga de viento

Naturalmente, los aerogeneradores se sitúan en zonas de alto potencial eólico. Por lo tanto, la carga del viento es otra carga importante en la consideración del diseño. La mayoría de los estándares de diseño (DNV-RP-C205, Condiciones ambientales y cargas ambientales, Design of Offshore Wind Turbine Structures DNVGL-OS-J101 (2007), BSI BS-6235: 1982 Code of Practice for Fixed Offshore Structures) distinguen entre condiciones de viento normales y extremas. Por término general las condiciones normales (que se obtienen de mapas y cartas con mediciones estadísticas de la velocidad del viento en la zona) se utilizan como parámetro para determinar el efecto de estas cargas dinámicas a fatiga. Las condiciones extremas presentan escenarios límite contemplados para asegurar la larga funcionalidad de los equipos ante situaciones imprevistas. DNVGL-OS-J101, 2004 proporciona metodologías para calcular la velocidad del viento, el parámetro fundamental de cálculo, y en donde se distingue entre velocidad media en un periodo  $T$  a una altura  $z$  del nivel del mar y los picos de los periodos medios de retorno de unos 50 años.

Este parámetro fundamental junto con su dirección, frecuencia, cortante y distribución permiten calcular las principales componentes de los esfuerzos aerodinámicos, como momentos flectores, esfuerzos cortantes en la base y efector turbulentos. Una descripción más detallada puede encontrarse en DNVGL-ST-0437 *Loads and site conditions for wind turbines* o IEC 61400-1 *Wind energy generation systems - Part 1: Design requirements*. Como conclusión llamar la atención al lector sobre la naturaleza dinámica de estas fuerzas, con una magnitud y frecuencia constantemente cambiante.

### 1.2.6 Carga de hielo

Algunas de las zonas con más capacidad instalada se encuentran en áreas especialmente frías, como el Mar del Norte o New-Found Bay (Canadá). Por lo tanto, la carga de hielo es fundamental en el dimensionamiento de la estructura. La carga es causada principalmente por el movimiento lateral del hielo contra la estructura. Evaluar el efecto de esta carga es bastante complejo para modelos teóricos, depende de la naturaleza y la calidad del hielo, incluido el tamaño, la edad, salinidad y temperatura. Por lo tanto, este tipo de carga se evalúa principalmente con la ayuda de modelos experimentales a escala que luego se transfieren a los equipos reales. Se puede encontrar información detallada sobre las cargas de hielo en la ISO 19906:2019 *Petroleum and natural gas industries, Arctic offshore structures*

### 1.2.7 Cargas sísmicas

Cuando una de estas estructuras se diseña para su instalación en un sitio en el que pueda estar sujeto a movimientos sísmicos (por ejemplo, China, Japón o EE. UU), la estructura se diseñará para soportar esas cargas. Los métodos de cálculo son similares a los utilizados en construcción civil y se basan en los espectros de respuesta que pueden calcularse a partir del historial de aceleración del terreno y la integral de Duhamel (IEC 61400-1 Ed.3: Wind Turbines—Part 1: Design Requirements)

Generalmente es suficiente reducir el análisis en dos direcciones horizontales a un análisis en una dirección horizontal, debido a la simetría del sistema dinámico. Como curiosidad, es necesario considerar el riesgo de pandeo vertical en la torre causado por la aceleración vertical para tipos específicos de terremotos.

### **1.2.8 Carga accidental**

Impactos con vehículos marinos o aéreos, explosiones, fuego y caída de objetos. Ver API RP 2A-WSD Recommended Practice for Planning, Designing, and Constructing Fixed Offshore Platforms, Working Stress Design: Section 18, Fire, Blast, and Accidental Loading, 2008

### **1.2.9 Carga muerta**

Formada por el peso de la plataforma, torre, soportes principales, “nacelle”, aerogenerador y palas. Es una de las cargas con mayor magnitud relativa y se modela generalmente como una carga por gravedad. Es de especial relevancia para dimensionar los soportes y “anclaje” (si aplica) al lecho marino. Es, además, esencial para el cálculo del centro de gravedad de la estructura y su equilibrio dinámico con respecto al centro de empuje (en el caso de las estructuras flotantes) o al eje central en el caso de las fijas.

### **1.2.10 Carga viva**

Se trata del efecto de aquellas cargas por gravedad que no son necesariamente propias de la estructura, como equipos no permanentes, personal temporal de mantenimiento o instalación, almacenes o depósitos, péndulos de aceite para contrapeso dinámico, etc.

## **1.3 Relevancia de la fatiga y el desgaste en la integridad de las estructuras Offshore.**

Dentro de este sector, el control de calidad de las uniones soldadas de estas estructuras es de suma importancia teniendo en cuenta que los defectos de soldadura son ampliamente considerados como puntos potenciales para el inicio de fallas estructural (Rodrigues, Wong, & Rogerson, 1980). El estudio basado en la mecánica de la fractura de parámetros como el CTOD,

en el contexto de la nucleación de grietas y el crecimiento de grietas por fatiga, se ha vuelto esencial para los fabricantes, diseñadores, sociedades de clasificación e inspectores. El cálculo de la vida a fatiga ocupa un lugar destacado en los códigos, estándares y reglas (ver por ejemplo ABS, 2011; Smith API 579, 2007 o DNV-OS-C101). Dicho análisis de fatiga se basa en métodos "rule-based" o en cálculos directos basados en modelos de datos S-N (Stress-No. of cycles), determinados por pruebas de fatiga de los detalles de soldadura considerados y la hipótesis de daño lineal.

Como este enfoque rara vez es posible (debido a la imposibilidad de realizar pruebas de fatiga completa de los detalles soldados), el análisis de fatiga puede basarse alternativamente en la mecánica de fractura. Los modelos de crecimiento de grietas de las sociedades de clasificación utilizan la formulación clásica de la ley de Paris y Erdogan con desarrollos de los modelos clásicos de bisagras de plástico (desarrollados por primera vez por la British Standards Institution y publicados en 1979). Según el vasto trabajo de Zhu y Joyce (2012), el factor de intensidad de estrés  $K$  (Irwin, 1954), el desplazamiento de apertura de frente de grieta (CTOD) (Wells, 1963), la integral  $J$  (Rice, 1968), y el ángulo de apertura de la punta de la grieta (CTOA) (desarrollado para materiales de paredes delgadas) son los parámetros más relevantes utilizados en la mecánica de fracturas. A partir de estos diversos parámetros que estudian la interacción de los materiales con la formación y propagación de grietas o defectos, el desplazamiento crítico de apertura de punta de grieta (CTOD) a una distancia dada de la punta de grieta es el más adecuado para modelar el crecimiento estable de grietas e inestabilidad durante el proceso de fractura (Newman, James, & Zerbst, 2003).

#### **1.4 Introducción al CTOD (dentro del contexto del presente trabajo) y su uso como parámetro de control del comportamiento a fatiga**

La prueba CTOD requiere la preparación de una entalla, con una geometría muy específica, que promueva la nucleación de una grieta estable y uniforme en un área delimitada (Akourri et al., 2000). El crecimiento de la grieta bajo la acción de fuerzas mecánicas dinámicas se transmite generalmente con enormes equipos oleo-hidráulicos y se controlan mediante extensómetros de precisión. La incertidumbre de los métodos de prueba, así como la sensibilidad a cualquier defecto interno, hacen que sea necesario realizar varias de estas pruebas para garantizar valores representativos. Por lo tanto, las pruebas CTOD son costosas, ya que requieren inversiones significativas en maquinaria de prueba, software, experiencia y subcontratación de servicios (Xiangqing et al., 2018). También se requiere la destrucción de grandes cantidades de material soldado de prueba ad-hoc (ASTM E1290-08e1c, 2008). Los resultados CTOD y las llamadas curva de diseño CTOD se puede usar para evaluar la resistencia a fractura para una amplia gama de estructuras, como tuberías, recipientes a presión, barcos, estructuras, edificios y puentes en alta mar. Una de las curvas de diseño CTOD más utilizadas es la desarrollada por el British Welding Institute (BWI) que relaciona el CTOD con algún evento crítico; el límite elástico  $\sigma_Y$ , la deformación nominal en una entalla  $\varepsilon$  o el tamaño del defecto  $a$  (Dawes, 1974). Se puede determinar una longitud de grieta crítica o un nivel de tensión utilizando la mencionada curva límite. La curva límite ha sido adoptada por estándares, por ejemplo, la British Standard BS-7910.

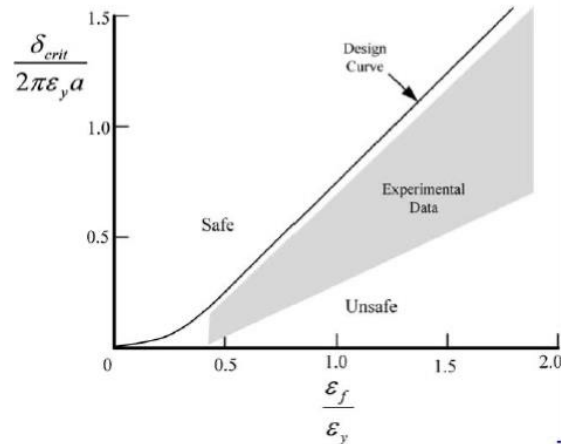


Figura 4. Ejemplo de curva de diseño CTOD

Fuente: (Anderson, 2005)

### 1.5 Enunciado del problema, hipótesis y punto de partida para la realización del presente trabajo académico

Se desprende de todo lo anterior que fabricantes, las sociedades de clasificación, los inspectores, las oficinas técnicas (y otros grupos involucrados) avanzan cada día en sus exámenes caracterizando propiedades cada vez más y más complejas de los materiales con los que trabajan e interrelacionando esos resultados para comprender el comportamiento de las uniones soldadas (y su criticidad) y aumentar de esa forma el control sobre las mismas.

Un buen ejemplo de esta dinámica es el estudio basado en mecánica de fractura de parámetros tales como el CTOD. En relación con estos parámetros la industria de fabricación se encuentra con dos escollos principales a la hora de controlar la calidad de estas uniones soldadas, a saber:

- 1) Los ensayos son costosos. Precisan el estudio con métodos relativamente avanzados, usando probetas de tipo y morfología muy estricto, que requieren grandes inversiones en maquinaria, know-how o en subcontratación de servicios. Así como la destrucción de grandes cantidades de material soldado ad-hoc para los ensayos.



- 2) Si los resultados no cumplen las altas expectativas es difícil para el fabricante interrelacionar los cambios en las variables de fabricación o diseño que permitan desplazar los resultados finales al valor adecuado.

El objetivo del presente trabajo es construir modelos matemáticos que relacionen el valor CTOD con el de otras variables más sencillas y económicas de medir. A partir de estos modelos se pretende analizar la influencia de cada variable en el CTOD y estimar este último a partir de los valores de esas variables “explicativas”. Para ello se han seleccionado variables representativas de la microestructura, resistencia mecánica y composición química del material de soldadura y las zonas afectadas por el proceso de soldeo por fusión.

### **1.5.1 Hipótesis de trabajo**

Se parte de las siguientes hipótesis de trabajo:

- a) Existe una relación entre propiedades mecánicas, microestructurales y químicas con la respuesta del material en un ensayo CTOD.
- b) Es posible establecer un modelo matemático que permita estimar un valor CTOD, en función de un conjunto de variables explicativas, con una exactitud coherente con la heterogeneidad de la muestra y la incertidumbre del ensayo.
- c) Del mismo modelo matemático se puede derivar un análisis de sensibilidad del resultado CTOD con respecto de las variables de entrada.
- d) Existe capacidad técnica y know-how a nivel industrial para modificar significativamente el valor de las variables explicativas.

### **1.5.2 Objetivos a alcanzar**

El presente trabajo pretende lograr la creación de un modelo matemático multivariante que relacione cuatro variables; microestructurales, mecánicas y químicas, con el valor CTOD que presenta una unión soldada de acero de gran espesor para la industria eólica Offshore. El modelo se debe sustentar en una extensa fase experimental que incluya ensayos reconocidos en la industria. Una vez determinado el modelo a través del estudio de los resultados experimentales se pretende determinar la sensibilidad del resultado final con respecto las variables de estudio, así como catalogar las formas de modificar esos valores de entrada. De igual forma se pretende poder predecir con solidez el rango de valores CTOD sin realizar el ensayo, solo a través del análisis, mucho más barato y habitual, de las variables de entrada.

## 2 Revisión bibliográfica

Los objetivos de esta revisión bibliográfica son de carácter descriptivo y panorámico, de forma que pueda presentarse un sustrato acumulado de conocimiento en las distintas áreas consideradas (como la modelización matemática, las propiedades de los materiales metálicos o la integridad estructural), de forma que sea suficiente como para impulsar un avance a través de refutación, hipótesis y confirmación. En todo caso, la revisión bibliográfica no tendrá un sentido crítico, comparativo o sistemático.

El uso de modelos matemáticos predictivos para el estudio de propiedades de los materiales está ampliamente tratado en la bibliografía (García et al., 2008, Rajakumar, 2012, Mohammed et al., 2016, Brown, 2018, solo por citar algunos de los ejemplos con métodos más novedosos como redes neuronales o redes bayesianas). Por lo tanto, la aplicabilidad “a priori” de esta metodología se considera suficientemente justificada

### 2.1 Ensayo CTOD, cálculo y modelos de fractura

Continuando con lo expuesto en el capítulo anterior sobre mecánica de fractura, sus parámetros principales y sus métodos de ensayo, se presenta en esta sección un análisis que justifica la elección del método de ensayo específico. Sobre esta cuestión se volverá en el apartado de metodología.

Se han publicado varios estándares para la estimación de los parámetros  $K$ ,  $J$ , CTOD y la curva  $R$ . Estos estándares son especialmente considerados con las condiciones de ensayo y se invierte un considerable esfuerzo para garantizar la representatividad y reproducibilidad de los resultados ya que los resultados son especialmente sensibles al efecto de la temperatura de ensayo (BSI 2014a; JWES 1995), espesor de la muestra (Hunt & McCartney 1979; Anderson,

1984), transición dúctil-frágil (Moskovic, 1993), endurecimiento por acritud, ..., etc. En esas pruebas, generalizando, se somete al material a solicitaciones de severidad creciente hasta que la fuerza impulsora de crecimiento de la grieta exceda la resistencia a la fractura del material.

Estos parámetros se usan habitualmente en Evaluaciones Críticas de Ingeniería (ECA) para evaluar la idoneidad de una estructura con respecto a la prevención de fracturas catastróficas (Shen et al. 2004; Gordon, Keith, & Gordon, 2013; Sarzosa, Souza y Ruggieri (2015); BSI 2014a; API (2007). Anderson y Osage (2000) también recuerdan que es importante que la resistencia a la fractura se evalúe adecuadamente, ya que es la variable principal para la evaluación (y aceptación) de defectos en servicio.

Los diferentes modelos para la estimación del CTOD se formularon en base a diferentes supuestos; el método de estimación geométrica para las BS/ISO, estimación geométrica con consideración de endurecimiento por deformación en JWES, y conversión basada en J en ASTM. También debe considerarse que la definición BS de CTOD es la apertura de la punta de grieta original, mientras que el CTOD se describe como la apertura en la intersección de 45 grados desde la punta de grieta para las ecuaciones basadas en J.

Se muestran a continuación algunos de los modelos de cálculo del CTOD utilizados en las principales normativas de ensayo:

$$\delta = K^2 \frac{(1 - \nu^2)}{2\sigma_{YS}E} + \frac{r_p B_0 V_p}{r_p B_0 + a_0 + z} \quad (3)$$

(BS7448-1, ASTM E1290, ASTM E1820)

$$\delta = \frac{J}{m\sigma_{YS}} \quad (4)$$

(ASTM E1820, ASTM E1290)

$$J = \frac{K^2(1 - \nu^2)}{E} + \frac{\eta_{pl}A_p}{B_n B_o} \quad (5)$$

(BSI EN ISO 15653)

$$\delta = K^2 \frac{(1 - \nu^2)}{2\sigma_{YS}E} + \frac{V_p(0.6\Delta_a + 0.4B_o)}{0.4W + 0.6(a_o + \Delta_a) + z} \quad (6)$$

(ISO 12135, BS 7448-1)

Todos los estándares actuales están de acuerdo en que el CTOD debe determinarse mediante la adición de dos componentes; el CTOD “elástico”,  $\delta_{el}$  y el CTOD “plástico”,  $\delta_{pl}$  (Wu, 1981). BS 7448-1 e ISO 12135 utilizan la misma ecuación para la determinación de CTOD. El componente elástico se determina a partir del factor de intensidad de tensión, K, mientras que el componente plástico asume un modelo de bisagra plástica fija en el ligamento de la muestra por delante de la grieta, y se calcula utilizando el método de triángulos similares. Basado en Lin et al. (1982) e Ingham et al. (1971), se supuso que el factor de rotación era 0.4.

La fórmula BSI / ISO no tiene en cuenta el endurecimiento por deformación del acero, y a pesar de haber sido correctamente validado para aceros de resistencia media y alta, la fórmula es menos precisa para otros aceros con una menor relación de rendimiento a tracción (Wei & Pisarski 2007; Khor et al.2016).

ASTM E1820 utiliza un enfoque diferente para la determinación de CTOD, donde se calcula primero la integral J (mediante la suma del componente elástico y plástico) y luego se convierte a CTOD usando un factor m, que incluye el rendimiento del material y las propiedades de tracción en su cálculo (ASTM 2014).

Sin embargo, se sabe que el método ASTM para determinar CTOD subestima CTOD significativamente para muchos aceros de mayor resistencia en comparación con el método BSI / ISO (Tagawa et al. 2010; Tagawa et al. 2014; Pisarski et al. 2010; Kayamori, Inoue, & Tagawa, 2008). Como se tratará más adelante en el apartado de metodología y desarrollo experimental se seleccionará la norma ASTM E1820 como la más idónea para la creación del modelo.

## **2.2 Selección de las variables explicativas (input)**

La selección de las llamadas variables explicativas se realizó considerando el enfoque industrial de este trabajo de investigación. Entre las numerosas variables con efectos probados en las propiedades del material (ver *Tabla 1* para una selección no exhaustiva propuesta por Dunne, Tsuei, & Sterjovski (2004); Haque y Sudhakar 2002, las siguientes fueron seleccionadas debido a su uso generalizado en la industria, la medición relativamente barata y la posibilidad de determinarse en laboratorios de control de calidad sin acceso a equipos de ensayo dinámicos. Además, las variables elegidas son parte de los ensayos requeridos por los códigos de diseño, reglas y estándares para el diseño, calificación y control de uniones soldadas. Por lo tanto, estos valores están generalmente disponibles (o son ciertamente fáciles de recopilar) y existen criterios de aceptación claros y estandarizados.

Tabla 1. *Selección no exhaustiva de variables con probado efecto sobre las propiedades de los materiales metálicos*

<b>Variables</b>	
<b>Carbono (wt %)</b>	Espesor (mm)
<b>Manganeso (wt %)</b>	Tratamiento térmico post-soldadura (PWHT)
<b>Silicio (wt %)</b>	PWH Ratio de enfriamiento
<b>Azufre (wt %)</b>	Orientación de las muestras de ensayo
<b>Fósforo (wt %)</b>	Temperatura de ensayo
<b>Aluminio (wt %)</b>	Límite elástico (MPa)
<b>Boro (wt %)</b>	Carga de rotura (MPa)
<b>Molibdeno (wt %)</b>	Tenacidad -Charpy- (J)
<b>Oxígeno (wt %)</b>	Juntas de grano y su orientación
<b>Nitrógeno (wt %)</b>	Dureza
<b>V% Volumen reaustenizado</b>	Juntas de grano ferríticas
<b>Ciclos térmicos</b>	Ferrita poligonal intergranular
<b>Subzonas de tamaño de grano afinado</b>	Subzonas de tamaño de grano engrosado
<b>Inclusiones no metálicas</b>	Tamaño tridimensional medio de las inclusiones

### 2.2.1 Microestructura (y su relación con otras variables)

La microestructura del material en el área en la que se determinará el valor CTOD se considerará una de las variables de entrada (o explicativas). Muchos autores han estudiado la relación entre las características de la microestructura, tales como el tamaño o multiaxialidad del grano (Chen et al., 2013 y 2018), el ángulo de las juntas que separan los granos cristalinos (Pegues, Roach, & Shamsaei, 2017), la orientación y tamaño de las inclusiones en la nucleación y propagación de grietas (Castelluccio & McDowell, 2016, Wei et al., 2014, Yun et al., 2017 y 2018) con respecto a la tenacidad a la fractura del material.

La relación de Hall-Petch, basada en los trabajos de Hall (1951) y Petch (1953) (quienes establecieron esta relación, por separado, entre el límite elástico y el tamaño de grano) es especialmente relevante, ya que conecta la microestructura con algunas de las demás variables explicativas, y se basa en un fenómeno, el efecto de los puntos de fijación (pinning-points) en las juntas de grano que es especialmente considerado en las tecnologías de fabricación.

La expresión de la relación es la siguiente:

$$\sigma_y = \sigma_0 + K \cdot D^{-\frac{1}{2}} \quad (7)$$

Donde  $\sigma_0$  se interpreta como el esfuerzo de fricción necesario para desplazar una dislocación (no bloqueada) en un plano de deslizamiento del sistema de apilamiento microestructural. Esto depende, fundamentalmente, de la temperatura, deformación y el nivel de aleación o impurezas del material (Rodríguez, Cabrera, & Benito, 2011).

La constante K que, si bien depende del material, es independiente del diámetro medio del grano (D) y de la temperatura (Dieter, 1988). Además, considerando estudios que relacionan la resistencia mecánica y la dureza;  $H = (2,5 \text{ a } 3) \sigma_0$  (Courtney, 2000). La relación Hall -Petch puede expresarse en términos de dureza como:

$$H_y = H_0 + K \cdot D^{-\frac{1}{2}} \quad (8)$$

Este efecto es aún más notable al evaluar el efecto del tamaño de grano en la transición del modo de fractura. Esto abre investigaciones futuras interesantes sobre el efecto del modelo considerando tamaños de grano críticos con la relación de Petch, la teoría de Cotrell de fractura por escisión (Cotrell, 1958), el modelo de Smith o el modelo de distancia característica de (Ritchie, Knott and Rice, 1973)

La mayoría de los modelos que relacionan microestructura y crecimiento o nucleación de grietas lo hacen a nivel macroscópico y considerando la variabilidad intrínseca de la



microestructura. Son pocos, sin embargo, los modelos que se han centrado en el crecimiento y nucleación, a nivel nanoscópico, de grietas a través de la microestructura, por ejemplo, estimando el número de ciclos necesarios para hacer crecer una grieta a través de los primeros granos metálicos. Algunas de las notables excepciones, como Krupp et al. (2007) o Castelluccio, Musinski y McDowell (2014) han relacionado ambas escalas de estudio.

El trabajo de este último ha esbozado un modelo elasto-plástico del propio grano cristalino (metálico),

$$\dot{\gamma}^{(\alpha)} = \dot{\gamma}_o \exp \left[ - \left( \frac{F_o}{k_b T} \right) \left\langle 1 - \left\langle \frac{|\tau^{(\alpha)} - B^{(\alpha)}| - S^{(\alpha)} \mu / \mu_0}{\tau_o \mu / \mu_0} \right\rangle^p \right\rangle^q \right] \operatorname{sgn} (\tau^{(\alpha)} - B^{(\alpha)}), \quad (9)$$

Donde  $\dot{\gamma}^{(\alpha)}$  es el ratio de cizalla del sistema de deslizamiento  $\alpha$ , T la temperatura absoluta y  $F_o$ , p, q,  $\dot{\gamma}_o$ ,  $\tau_o$ ,  $\mu$ , y  $\mu_0$  parámetros que dependen del propio material.

Podemos concluir, por lo tanto, que existe una sólida base en la bibliografía para relacionar la microestructura (a nivel microscópico, pero no necesariamente nanoscópico) con las propiedades de tenacidad a la fractura de los materiales metálicos y, especialmente con otras propiedades de los materiales como la resistencia mecánica o la dureza. Es por lo tanto una variable de especial interés para nuestro modelo.

### 2.2.2 Composición química (y su relación con otras variables)

La composición química del material es un factor, bien conocido, que ejerce influencia sobre las diferentes propiedades mecánicas (Talas, 2010; Lee et al., 2017; Yurioka, 2001).

Seok Gyu Lee et al. (2017, 2019) han estudiado el efecto de la adición de Ni, Mn o Mo en la composición química de los aceros (HSLA) sobre el CTOD en las zonas afectadas térmicamente de uniones soldadas. Los trabajos de Andrews (1965) ya reflejan el efecto de la composición química en las temperaturas de transformación microestructural de los aceros bajos en carbono. El efecto de la ferrita acicular inducida por Mn-Mo-Nb y su efecto sobre la tenacidad y la resistencia mecánica (Smith, Coldren y Cryderman, 1972). Efecto del Mn y del Ni en la variación de propiedades mecánicas y microestructura para soldaduras de acero bajo en carbono, Kang et al, 2000. Donde se identificó como mecanismo principal el cambio en el modo de fractura (intergranular) y el cambio de la ruta de avance de grieta en los granos de estructura columnar si no se formaba ferrita delta. Incluso el efecto de las inclusiones, como los óxidos de Ti-Ca estudiados por Wang et al. (2019), resulta determinante al propiciar la aparición de microestructuras aciculares (con un mayor efecto concentrador de tensiones).

La composición química siempre ha tenido un papel destacado en las tecnologías de fabricación, su conocimiento es avanzado y las medidas que se utilizan en su diseño y control son elevadas y es habitual un control de la composición química para los materiales base, aportes metálicos de soldadura, gases de protección, flux, etc. En la siguiente figura se ilustra a título de ejemplo algunos de estos efectos y la magnitud de su influencia.

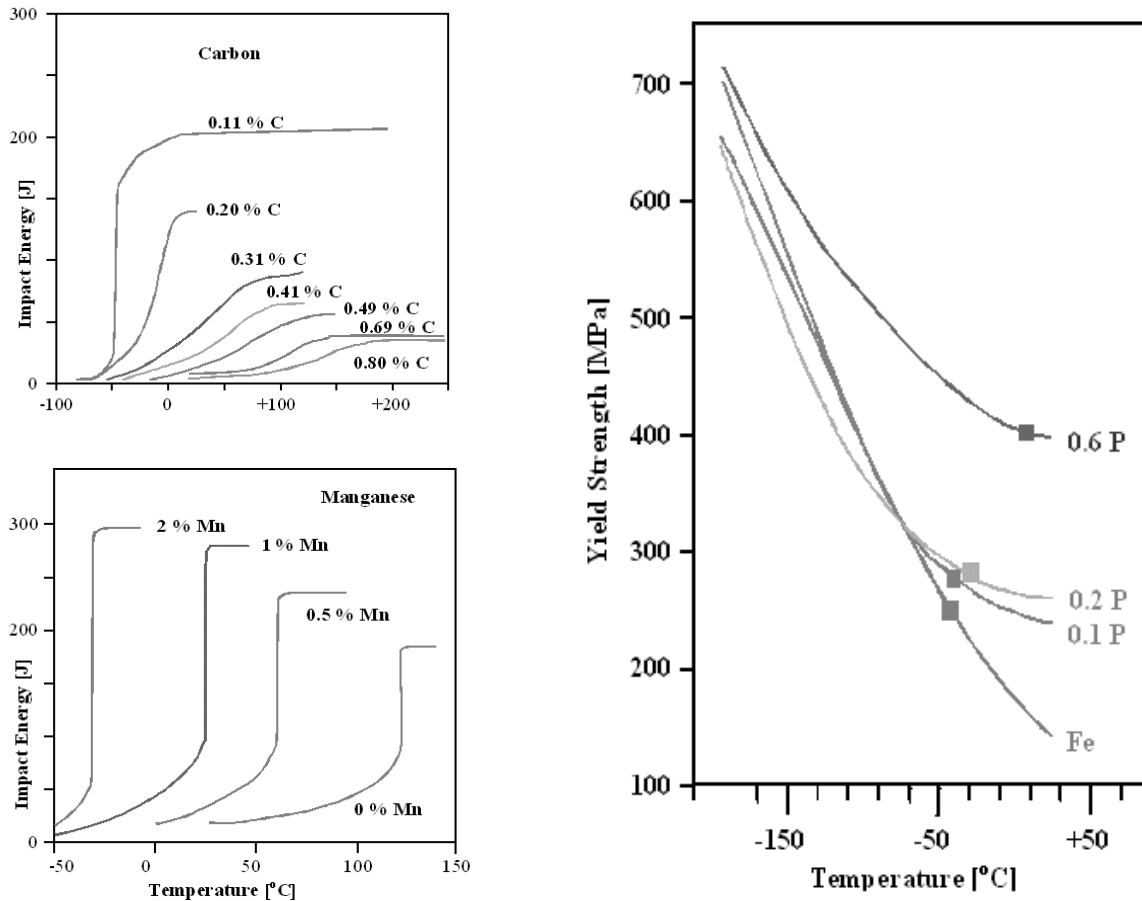


Figura 5. Arriba-Izquierda, cambio en la temperatura de transición dúctil-frágil para diferentes composiciones de carbono. Abajo-Izquierda, cambio en la temperatura de transición dúctil-frágil para diferentes composiciones de Manganeso. Derecha, cambios en el límite elástico con respecto a la temperatura y diferentes cantidades de Fósforo.

Fuente: (Petch, 1953)

Los principales mecanismos por los cuales la composición química afecta a los resultados de otras propiedades de los materiales metálicos es el efecto catalizador de ciertas microestructuras, el efecto de distorsión de la red cristalina de algunos elementos aleantes, el efecto de las inclusiones (óxidos y sulfuros generalmente) de algunos componentes (pinning-points) y, naturalmente, la solubilidad y soluciones sólidas entre los distintos componentes.

### 2.2.3 Resistencia mecánica (y su relación con otras variables)

La resistencia mecánica juega un papel fundamental y forma parte constitutiva en la mayor parte de método de ensayo en mecánica de la fractura (Anderson, 2005). Además, la determinación y el control de su valor es una parte fundamental del control de calidad de las propiedades del material (en el caso de materiales estructurales). El uso de ensayos de tracción se descartó debido a la imposibilidad de tomar medidas exclusivamente de la pequeña área de interés, ya que todas las probetas propuestas por las normas de ensayo exceden la capacidad de la máquina de prueba (por ser demasiado pequeña) o destruyen demasiado material de prueba valioso (demasiado grande). Sin embargo, según numerosas publicaciones (por ejemplo, “Handbook” Carbon Steel ASM -American Society for Metals- o Pavlina & Vantyne, 2008), existe una relación constante y casi lineal entre la resistencia a la tracción (UTS) y la dureza,

$$TS = \left(\frac{H}{2.9}\right) \left(\frac{n}{0.217}\right)^n, \quad (10)$$

$$YS = \left(\frac{H}{3}\right) (0.1)^n, \quad (11)$$

Donde TS es “Tensile Strenght” ( Carga de rotura) , YS es “Yield Strength” (límite elástico) y n es el exponente de endurecimiento “strain-hardening” (Cahoon, 1972).

Por lo tanto, las mediciones de dureza de acuerdo con ASTM E92 -Dureza Vickers 10- (2017) se tomaron de las muestras para estimar la resistencia mecánica del material. Esto se tratará de nuevo en la descripción de la metodología de ensayo.

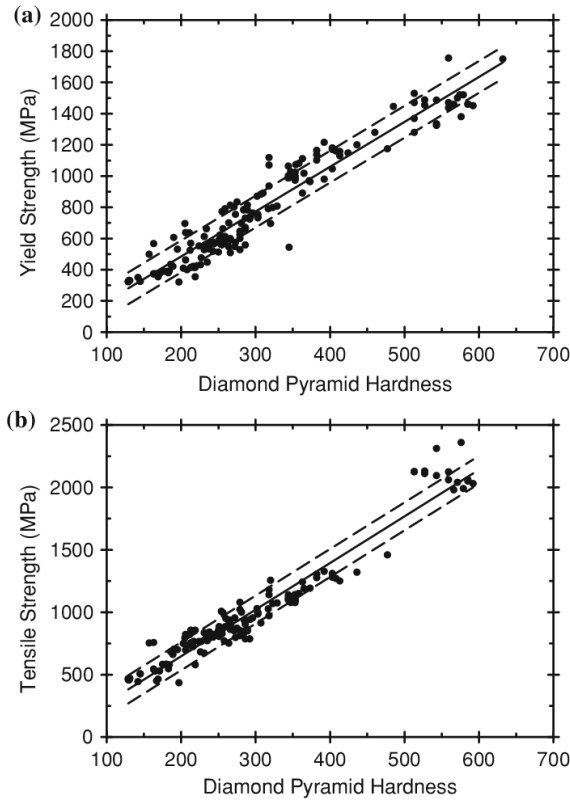


Figura 6. Resistencia mecánica frente a dureza Vickers

Fuente: (Pavlina & Vantyne, 2008)

### 2.2.4 Resiliencia (y su relación con otras variables)

Estudios previos, por ejemplo, Khor et al. (2016); Maropoulos et al. (2004), Takashima et al. 2006 respaldan la relación entre los resultados de las pruebas de impacto (medidos como valores de energía de Charpy con entalla en “V” (CVN)) y la resistencia a la fractura. Son numerosos los modelos que relacionan el resultado CVN y el parámetro  $K_{Ic}$ . Entre ellos, Barsom y Rolfe (1970) sugieren una relación expresada como (nótese la relevancia de  $K_{Ic}$  como parámetro de fractura):

$$\left(\frac{K_{Ic}}{\sigma_{ys}}\right)^2 = \frac{5}{\sigma_{ys}} \left(CVN - \frac{\sigma_{ys}}{20}\right), \quad (12)$$

Marendet y Sanz (1977) sugieren:

$$K_{Ic} = 19(CVN)^{1/2}, \quad (13)$$

Sailor y Corten (1972) :

$$\frac{K_{Ic}^2}{\sigma_{ys}} = 8(CVN) \quad (14)$$

A su vez las correlaciones entre  $K_{Ic}$  y CTOD son numerosas y siguen la forma:

$$CTOD = \frac{K_{Ic}}{\sigma_{ys}E} \quad (15)$$

Algunas correlaciones incluso han sido adoptadas por las normas ASME Boiler and Pressure Vessel Code (BPVC) XI (2017) y API 579.

Es por lo tanto razonable incluir esta variable como explicativa en el modelo matemático que se pretende determinar.

### 2.3 Análisis predictivo

El análisis predictivo es un área de las matemáticas que estudia la extracción de información (existente) en una colección de datos y su utilización para detectar tendencias y patrones de comportamiento de sistemas, pudiendo aplicarse sobre cualquier evento desconocido, ya sea en el pasado, presente o futuro. El análisis predictivo se fundamenta en la identificación de relaciones entre variables en eventos pasados, para luego interpolar-extrapolar dichas relaciones y predecir posibles resultados en el futuro.

### 2.3.1 Preselección de métodos apropiados

Los enfoques y técnicas utilizados para realizar el análisis predictivo pueden agruparse (de una manera muy general) en técnicas de regresión y técnicas de aprendizaje computacional. Los modelos de regresión son, probablemente, el pilar central de la analítica predictiva. Dependiendo de la situación, hay una gran variedad de modelos que se pueden aplicar durante la realización del análisis predictivo. (Espino & Martínez, 2017, pp.17-18)

El enfoque general se basa en el establecimiento de una ecuación matemática como modelo para representar las interacciones entre las diferentes variables en consideración (Espino & Martínez, 2017, p.18).

### 2.3.2 Método de regresión lineal

El análisis de la relación entre la variable dependiente o de respuesta y la/las variables independientes es el objeto del análisis de los modelos de regresión lineal. Normalmente esta expresión se plasma en una expresión matemática que tiene como resultado un estimador predictivo de la variable respuesta (en función lineal de los valores de entrada de la/las variables independientes). Estos parámetros se modulan para minimizar el error cometido en la predicción y ajustar el modelo de forma óptima para que este error se distribuya de forma aleatoria.

El objetivo de la regresión es seleccionar los parámetros del modelo que minimizan la suma de los errores al cuadrado. Conocido como estimación de mínimos cuadrados (ordinarios). Los resultados son las mejores estimaciones lineales no sesgadas de los parámetros si y solo si se satisfacen las suposiciones de Gauss-Markov. (Espino & Martínez, 2017, p.18).

### **2.3.3 Árboles de regresión (aprendizaje computacional)**

Los árboles de clasificación y/o regresión (Classification And Regression Trees, CART) son una técnica de aprendizaje no paramétrica que produce los llamados árboles de clasificación o regresión, dependiendo de si la variable dependiente es categórica o numérica, respectivamente y son usados para clasificar o predecir nuevos datos.

### **2.3.4 Curvas de regresión adaptativa multivariable**

Las curvas de regresión adaptativa multivariable (Multivariate Adaptive Regression Splines, MARS) son una técnica no paramétrica que construye modelos flexibles al ajustar regresiones lineales por piezas. Un “nudo” (donde un modelo de regresión local da paso a otro y por lo tanto es el punto de intersección entre dos curvas) es concepto importante asociado con el método MARS. (Espino & Martínez, 2017, p.20)

### **2.3.5 Redes neuronales (aprendizaje computacional)**

“Las redes neuronales son técnicas de modelización no-lineal que son capaces de modelar funciones complejas. Pueden aplicarse a problemas de predicción, clasificación o control en un amplio espectro de campos” (Araneda, 2019, párr.26).

### **2.3.6 Descartados**

Para el presente trabajo se han descartado otros métodos de aprendizaje computacional como Máquinas de vectores de soporte o Redes Bayesianas.



### 2.3.7 Modelos de regresión lineal multivariante

Si observamos un conjunto de variables  $K$ ,  $X_1, X_2, \dots, X_K$  en un conjunto de  $n$  elementos de una población de donde se quiere resumir los valores de las variables y describir su estructura de dependencia.

Cada una de estas  $K$  variables se denomina variable escalar o invariada y el conjunto de estas  $K$  variables forma un vector o variable multivariante. Todos estos valores se pueden representar en una matriz,  $X$ , de dimensiones  $n \times p$ , denominada matriz de datos, donde cada fila representa los valores de las variables  $K$  sobre el  $i$  individuo, y cada columna representa la variable escalar correspondiente medida en los  $n$  elementos de la población. En el elemento  $x_{ij}$ ,  $i$  denota al individuo y  $j$  es la variable.

A continuación, se procede al análisis multivariado de las observaciones. Para hacer esto, de calcula el vector de medias  $\bar{X} = [\bar{X}_1 \quad \bar{X}_2 \quad \dots \quad \bar{X}_k]$  de dimensión  $p$ , cuyos componentes son las medias de cada una de las variables  $p$  y la matriz de covarianza. Utilizando la matriz de datos centrados  $\tilde{X}$ ,

$$\tilde{X} = X - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \bar{X}, \quad (16)$$

Se calcula la matriz semidefinida simétrica y positiva de covarianza  $S = \frac{1}{n} \tilde{X}^T \tilde{X}$ .

El objetivo de describir datos multivariados es comprender la dependencia entre la variable objetivo y las variables explicativas. Para esto estudiamos:

- A. La relación entre pares de variables;
- B. Dependencia entre la variable objetivo y todas las variables explicativas;

C. Dependencia entre la variable objetiva y la explicativa, pero eliminando el efecto de algunas de ellas.

La dependencia por pares entre las variables se mide mediante la matriz de correlación semidefinida simétrica y positiva R

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1K} \\ r_{21} & 1 & \dots & r_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ r_{K1} & r_{K2} & \dots & 1 \end{bmatrix}, r_{jk} = \frac{S_{jk}}{S_j S_k} \quad (17)$$

Tal que existe una relación lineal exacta entre las variables  $X_j$  y  $X_k$  si  $|r_{jk}| = 1$ .

Puede suceder que haya variables que dependen mucho de otras, en cuyo caso es conveniente medir su grado de dependencia. Asumiendo que  $Y = X_j$  es la variable de interés, y llamando  $\hat{Y}$  a la variable utilizada para estimar  $Y$ , el mejor predictor lineal de las otras variables, llamadas variables explicativas, es:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K, \quad (18)$$

Donde el parámetro  $\beta_i$  se determina a través de los datos que tenemos a nuestra disposición.

El problema es encontrar el conjunto de parámetros que minimicen  $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ , llegando a,

$$\begin{aligned} y &= Y - \bar{Y} \\ x_j &= X_j - \bar{X}_j, \quad j = 1, \dots, K \end{aligned} \quad (19)$$

Y definiendo  $\hat{y} = \hat{Y} - \bar{Y}$ , tenemos  $Y - \hat{Y} = y - \hat{y}$ , y la ecuación (8) puede reescribirse como:

$$\hat{y} = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_K x_K, \quad (20)$$

Como minimizar  $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$  es equivalente a minimizar  $\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2$ , derivando esta suma con respecto a los  $\alpha_k$  parámetros, obtenemos un sistema de  $p - 1$  ecuaciones que puede escribirse,

$$\sum_{i=1}^n e_i x_{il} \quad l = 1, \dots, K, \quad l \neq j \quad (21)$$

La ecuación (19) indica que los errores de predicción no deben correlacionarse con las variables explicativas, de modo que la covarianza de ambos sea cero, o el vector residual debe ser ortogonal al espacio generado por las variables explicativas. Al definir la matriz  $X_R$ , de tamaño  $n \times (p - 1)$ , obtenida al eliminar la columna  $\bar{X}$  correspondiente a la variable que queremos predecir,  $y = x_j$ , los parámetros son calculados por el siguiente sistema de ecuaciones,

$$\boldsymbol{\alpha} = (X_R^T X_R)^{-1} X_R^T \mathbf{y} \quad (21)$$

Y la ecuación (20), con sus coeficientes, es la ecuación de regresión múltiple entre variable  $y = x_j$  y las variables restantes  $x_i$ ,  $i \neq j$ ,  $i = 1, \dots, K$ .

Para expresar este resultado basado en las  $X_1, \dots, X_K$  variables de la ecuación (18), debemos considerar:

$$\begin{aligned} \beta_i &= \alpha_i, \quad i = 1, \dots, K \\ \beta_0 &= \alpha_0 + \bar{Y} - \sum_{i=1}^K \alpha_i \bar{X}_i \end{aligned} \quad (22)$$

El cuadrado del coeficiente de correlación múltiple (que puede ser mayor, menor o igual que la suma de los cuadrados de las correlaciones simples entre la variable objetivo y cada una de las variables explicativas) (Cuadras, 1981) entre la variable  $y = x_j$  y el resto es,

$$R_j^2 = 1 - \frac{SS_{resid}}{SS_{total}} = 1 - \frac{1}{s_{jj}S^{jj}}, \quad (23)$$

Donde  $s_{jj} = s_j^2$  es el elemento diagonal j-ésimo de la matriz de covarianza S y  $S^{jj} = \frac{1}{s_r^2(j)}$  es el elemento diagonal j-ésimo de la matriz  $S^{-1}$ , que representa la varianza residual de una regresión entre la variable j-ésima y el resto. Como cada vez que se agrega una variable al modelo, se reduce el número de grados de libertad y se aumenta el ajuste, es necesario hacer una corrección de este coeficiente y calcular el ajuste  $\bar{R}_j^2$ ,

$$\bar{R}_j^2 = 1 - \frac{\frac{SS_{resid}}{(n-k)}}{\frac{SS_{total}}{(n-1)}}, \quad (25)$$

Donde  $n$  es el número total de observaciones y  $k$  es el número de variables del modelo; es decir, se realiza el mismo cálculo, pero ponderado por los grados de libertad de los residuos,  $n - k$ , y el modelo,  $n - 1$ .

El R-squared  $RSQ = \frac{\sum(\hat{y}_i - y_i)^2}{\sum(y_i - \bar{y})^2}$  es una medida descriptiva de la capacidad predictiva del modelo, y para una sola variable explicativa es el cuadrado del coeficiente de correlación simple entre las dos variables.

### 2.3.8 MARS (Multivariate Adaptive Regression Splines)

Vanegas y Vásquez (2017) define las MARS (Multivariate Adaptive Regression Splines) como una herramienta matemática flexible que a través de; seleccionar las variables relevantes, transformar las predictoras, establecer las interacciones de las variables predictoras, tratar los valores perdidos y realizando pruebas internas de autodiagnóstico (contra el sobreajuste), permite automatizar la construcción de modelos predictivos.

Este método puede revelar patrones y relaciones que es difícil, si no imposible, para otros métodos (Friedman, 1991; Salford, 2001).

Este método de modelización matemática (no paramétrico) extiende el modelo lineal al introducir no linealidades (e incluso interacciones entre variables). Es, por lo tanto, una generalización de las *Recursive Partitioning Regression* (RPR), que divide el espacio de las variables (predictoras) en diferentes subregiones.

El modelo puede escribirse como:

$$y_t = f(x_t) = \beta_0 + \sum_{i=1}^k \beta_i B(x_{it}) \quad (26)$$

Donde  $y_t$  es la variable respuesta en el instante  $t$  y  $\beta_i$  son los parámetros del modelo para las respectivas variables  $x_{it}$ , que van de  $i = 1, \dots, k$ . El valor  $\beta_0$  representa al intercepto, las funciones bases  $B(x_{it})$  son funciones que dependen de las respectivas variables  $x_{it}$ , en donde cada  $B(x_{it})$  puede escribirse como  $B(x_{it}) = \max(0, x_{it} - c)$  o  $B(x_{it}) = \max(0, c - x_{it})$ ,  $c$  es un valor umbral y  $k$  representa el número de explicativas, que incluye interacciones de las variables predictoras (Vanegas & Vásquez, 2017).

De alguna forma el número de funciones base resultantes da una idea de la complejidad del modelo (Salford, 2001).

MARS genera puntos de corte para las diferentes variables. Los puntos son identificados a través de las funciones basales, las que indican el inicio y el término de una región. En cada región en que se va dividiendo el espacio se ajusta una función base de una variable, la cual es lineal. El modelo final se constituye como una combinación de las funciones base generadas. Para determinar estos puntos de corte usa un algoritmo *forward-backward stepwise* por etapas. Primero, mediante el algoritmo *forward stepwise* se genera un modelo sobreestimado con un gran número de funciones base; posteriormente, mediante el algoritmo *backward stepwise*, se eliminan los nodos que menos contribuyen al ajuste global. El algoritmo se detiene cuando la aproximación construida incluye un número máximo de funciones fijadas por el investigador. (Vanegas & Vásquez, 2017, p.236)

Cuando se tienen varios modelos teóricos y se desea identificar el mejor, puede utilizar los siguiente (Lin et al., 2011; Lewis & Stevens, 1991):

El criterio de validación cruzada (GCV), Friedman (1991), que es una medida de ajuste a los datos y penalización, debido a la complejidad del modelo y el aumento de la varianza. La idea subyacente es que un modelo más sencillo debería ser preferido frente a otro más complejo.

El coeficiente de determinación ( $R^2$  ajustado) entre el valor observado y el predicho, el cual permite la adecuación del modelo utilizado para las predicciones.

La proporción media de error absoluto, que viene dada por el valor predicho y los valores observados y puede expresarse como:

$$\sum_{i=1}^n \left\{ \left| \frac{\text{valor observado}(i) - \text{valor predicho}(i)}{\text{valor observado}(i)} \right| \right\} \quad (27)$$

Esta ecuación muestra el porcentaje de error en el que se incurre con la predicción en comparación con los datos observados, considerándose, claro, que el mejor modelo es aquel que tiene el porcentaje de error más bajo.

### 2.3.9 Árboles de clasificación o decisión

El algoritmo conocido como particionamiento recursivo es el proceso paso a paso (Díaz, 2012) para construir un árbol de decisión y es la clave para el método estadístico no paramétrico “CART” (Classification And Regression Trees) (Izenman, 2008).

Si consideramos  $Y$  como una variable respuesta y  $p$  como variables predictoras  $x_1, x_2, \dots, x_p$ , donde las  $x$  son fijas e  $Y$  es aleatoria (como variable). El problema estadístico es establecer una relación entre  $Y$  y las  $x$  de tal forma que sea posible predecir  $Y$  basado en los valores de las  $x$  (la misma idea que hemos visto anteriormente). Traducido a lenguaje matemático, se quiere estimar la probabilidad condicional de la variable aleatoria  $Y$ , cuando la variable  $Y$  es discreta (o una función de su probabilidad tal como la esperanza condicional).

$$P\{Y = y | x_1, x_2, \dots, x_p\} \quad (28)$$

$$E\{Y | x_1, x_2, \dots, x_p\}$$

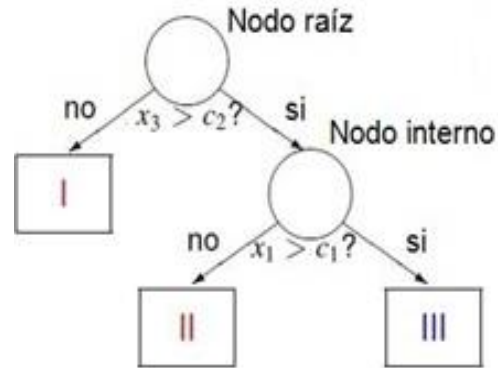


Figura 7. Ejemplo de árbol CART.

Fuente: (Zhang & Singer, 2010)

Según Zhang y Singer (2010) y para ilustrar las ideas básicas, considérese el diagrama de la figura anterior (Figura 7). El árbol presenta tres niveles diferenciados. El primer nivel, tiene un nodo único en la cima que podemos llamar nodo raíz. Un nodo interno en el segundo nivel (derecha), y tres nodos terminales (representados con cuadrángulos) que están respectivamente en el segundo y tercer nivel. El nodo raíz y el nodo interno son partidos cada uno en dos nodos en el nivel subsecuente, los cuales son llamados nodos hijos izquierdo y derecho.

El nodo raíz contiene una muestra de sujetos desde la cual se aumenta el árbol, es decir, desde donde se desprenden los demás nodos. Estos sujetos constituyen lo que se llama una “muestra de aprendizaje” del modelo, la cual puede ser la muestra total en estudio o solo una parte de ésta (Díaz, 2012).

El objetivo del particionamiento recursivo es acabar en nodos terminales que se puedan considerar homogéneos. La completa homogeneidad de los nodos terminales es un ideal raramente alcanzado en el análisis de datos real. Así, el objetivo de la partición recursiva es hacer las variables resultantes en los nodos terminales tan homogéneas como sea posible.



Una medida cuantitativa de esa homogeneidad es la noción de impureza. La idea es que está relacionada con la siguiente relación:

$$\text{Impureza del nodo} = \frac{\text{Num. de sujetos que cumplen la característica en el nodo}}{\text{Num. total de Sujetos}}$$

Esta metodología se puede subdividir en tres pasos fundamentales, la construcción del llamado árbol saturado, la selección del tamaño adecuado y la clasificación de nuevos datos usando el árbol.

La construcción del árbol saturado se hace con particionamiento recursivo. La diferencia en la construcción de árboles de clasificación o de regresión es el criterio de división de los nodos, es decir, la medida de impureza es diferente (Díaz, 2012).

La función impureza es:

$$i(\tau) = \phi(P\{Y = 1|\tau\}) \quad (29)$$

donde,

$$\phi \geq 0, \text{ Para cualquier } p \in (0,1), \phi(p) = \phi(1-p), \phi(0) = \phi(1) < \phi(p)$$

con,

$$\phi(p) = \min(p, 1-p) \text{ (Error de Bayes) ó}$$

$$\phi(p) = -p \log(p) - (1-p) \log(1-p) \text{ (Entropía)}$$

Pueden existir problemas de sobreajuste si un árbol tiene una gran cantidad de datos. He esto se desprende la importancia del tamaño del árbol a la hora de dar conclusiones sobre la población de estudio.

El proceso de particionamiento recursivo continua hasta que el árbol sea saturado en el sentido de que los sujetos en los nodos descendientes no se pueden partir en una división adicional. Esto sucede, por ejemplo, cuando queda solo un sujeto en un nodo. El número total de divisiones permitidas para un nodo disminuye cuando aumentan los niveles del árbol. Cualquier nodo que no pueda o no sea dividido es un nodo terminal. El árbol saturado generalmente es bastante grande para utilizarse porque los nodos terminales son tan pequeños que no se puede hacer inferencia estadística razonable debido a que los datos quedan "sobre-ajustados", es decir, el árbol alcanza un ajuste tan fiel a la muestra de aprendizaje que cuando en la práctica se aplique el modelo obtenido a nuevos datos los resultados pueden no tener valor. (Díaz, 2012, pp.9-10)

Breiman et al., (1984) afirma que, considerando el límite de parada, el proceso de partición tiende a terminar muy pronto o muy tarde. De forma que subsecuentes pasos consistentes en encontrar un subárbol del árbol saturado que sea el más "predictivo" de los resultados y menos vulnerable al ruido en los datos son habituales. La metodología para construir árboles de clasificación y árboles de regresión es la misma (o equivalente).

### **2.3.10 Modelos lineales generalizados**

Los modelos lineales generalizados tienen tres componentes básicos: Una componente aleatoria (cuyo objetivo es identificar la variable respuesta y su distribución de probabilidad), una componente sistemática (que pretende especificar las variables explicativas, independientes o predictoras, utilizadas en la función predictiva final) y una función de unión o enlace (una función del valor esperado de  $Y$ ,  $E(Y)$ , como una combinación lineal de las variables predictoras).

### 2.3.11 Componente aleatoria

La componente aleatoria de un modelo lineal generalizado consiste en una variable aleatoria  $Y$  con observaciones independientes  $(y_1, \dots, y_N)$ . En muchas aplicaciones, las observaciones de  $Y$  pueden ser binarias y se identifican como *éxito/fracaso*. Aunque de modo más general, cada  $Y_i$  indica el número de éxitos de entre un número fijo de ensayos, y se modeliza como una distribución binomial (Marín, 2014)

En otras ocasiones cada observación podría ser un recuento, con lo que se puede asignar a  $Y$  una distribución de binomial negativa o una de Poisson. Finalmente, si las observaciones son continuas se puede asumir para  $Y$  una distribución de tipo normal. Todos estos modelos se pueden incluir dentro de la llamada familia *exponencial* de distribuciones  $f(y_i|\theta_i) = a(\theta_i) \cdot b(y_i) \cdot \exp[y_i Q(\theta_i)]$ , por lo tanto,  $Q(\theta)$  recibe el nombre de *parámetro natural*.

### 2.3.12 Componente sistemática

La llamada componente sistemática de un modelo lineal generalizado especifica las variables explicativas que entran (en forma de efectos fijos) en un modelo lineal, es decir, las variables  $x_j$ . El denominado *predictor lineal* es  $\alpha + \beta_1 x_1 + \dots + \beta_k x_k$ , una combinación lineal de variables explicativas.

### 2.3.13 Función de enlace

Se denota el valor esperado de  $Y$  como  $\mu = E(Y)$ , entonces la función de enlace (*link*) especifica una función  $g(\cdot)$  que relaciona  $\mu$  con el predictor lineal como:

$$g(\mu) = \alpha + \beta_1 x_1 + \dots + \beta_k x_k \quad (30)$$

Así, la función enlace  $g(\cdot)$  relaciona las componentes aleatoria y sistemática.

Los modelos de regresión lineal típicos para respuestas continuas son un caso particular de los GLM. Estos modelos generalizan la regresión ordinaria de dos modos: permitiendo que  $Y$  tenga distribuciones diferentes a la normal y, por otro lado, incluyendo distintas funciones link de la media. Los modelos GLM permiten la unificación de una amplia variedad de métodos estadísticos como la regresión, los modelos *ANOVA* y los modelos de datos categóricos. En realidad, se usa el mismo algoritmo para obtener los estimadores de máxima verosimilitud en todos los casos. (Halweb, s.f., p.3)

## 2.4 Redes neuronales

Las redes neuronales (Neural Networks o NN) fueron originalmente una simulación teórica de los sistemas nerviosos orgánicos, constituidos por un conjunto de unidades, neuronas, o nodos conectados los unos con los otros.

Los primeros modelos de red neuronal fueron propuestos en 1943 por McCulloch y Pitts en términos de un modelo computacional de actividad nerviosa. Este modelo era un modelo de tipo binario, donde cada neurona tenía un escalón (o umbral) prefijado. Estos primeros modelos inspirados en las redes neuronales biológicas se han ido desarrollando. Las características principales de estas NN o redes neuronales son las siguientes:

- Procesado en paralelo. Normalmente se usa un gran número de nodos de procesado, con alto nivel de interconectividad.
- La autoorganización y la adaptabilidad. Se utilizan algoritmos de aprendizaje adaptativo y auto-organización, por lo que ofrecen mejores posibilidades de procesado robusto y adaptativo.

- Procesado no-lineal. Se aumenta la capacidad de la red para aproximar funciones, clasificar patrones y aumenta su inmunidad frente al ruido.

El elemento indivisible y básico de computación dentro de los modelos con redes neuronales es el nodo o unidad (a veces llamado modelo de neurona). Este elemento recibe un input desde otras unidades o de una fuente externa de datos. Cada uno de estos inputs de entrada tiene un valor asociado  $w$  (también llamado peso), que se modifica en el llamado proceso de aprendizaje. Por otro lado, cada una de esta unidad aplica una función dada  $f$  de la suma de los inputs (ponderadas mediante los pesos  $y_i = \sum_j w_{ij}y_j$ ) y donde el resultado puede servir como output (salida)de otras unidades.

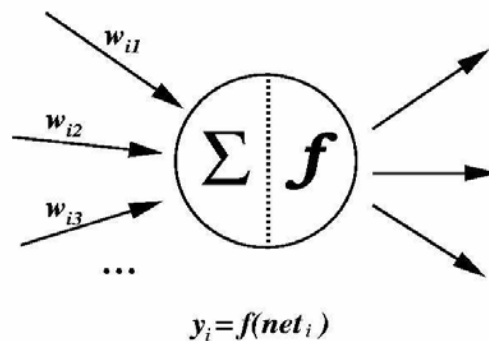


Figura 8. Modelo de neurona para NN

Las características de las NN juegan un importante papel, por ejemplo, en el procesado de señales e imágenes. Se usan arquitecturas que comprenden elementos de procesado adaptativo paralelo, combinados con estructuras de interconexiones jerárquicas. Hay dos fases principales en la modelización con redes neuronales: Fase de entrenamiento: se usa un conjunto de datos o patrones de entrenamiento para determinar los pesos (parámetros) que definen el modelo de red neuronal. Se calculan de manera iterativa, de acuerdo con los valores de los valores de entrenamiento, con el objeto de minimizar el error cometido entre la salida obtenida por la red neuronal y la salida deseada. Fase de Prueba: en la fase

anterior, el modelo puede que se ajuste demasiado a las particularidades presentes en los patrones de entrenamiento, perdiendo su habilidad de generalizar su aprendizaje a casos nuevos (sobreajuste). (Ingenieria Matematica, 2017, párr.5)

### 2.4.1 El Perceptrón multicapa

El perceptrón multicapa es un modelo que actúa como aproximador “universal” de funciones y esta formado por una red *back-propagation* que contiene al menos una capa oculta con suficientes unidades no-lineales. Son por lo tanto herramientas flexibles y de propósito general.

Rumelhart, Hinton y Williams (1986) proponen un método para que una red de este tipo “aprendiera” la asociación que existe entre un conjunto de patrones de entrada y sus salidas correspondientes con el método *back-propagation error* (propagación del error hacia atrás). Esta red tiene la capacidad de generalización, es decir, genera resultados satisfactorios a entradas que el sistema no ha visto nunca en su fase de entrenamiento.

Un perceptrón multicapa está formado por los siguientes elementos: una capa de entrada, una capa de salida y una o más capas ocultas; aunque se ha demostrado que para la mayoría de los problemas bastará con una sola capa oculta.

Si se considera  $w_{ji}$  como el peso de conexión entre la neurona de entrada  $i$  y la neurona oculta  $j$ , y  $v_{kj}$  como el peso de conexión entre la neurona oculta  $j$  y la neurona de salida  $k$ . Ante un patrón  $p$  de entrada  $X^p: x_1^p, x_i^p, \dots, x_N^p$ , este se transmite a través de los pesos  $w_{ji}$  desde la capa de entrada hacia la capa oculta.

Las neuronas de la capa intermedia procesan las señales recibidas transformándolas mediante la aplicación de una función de activación. Este se transmite a través de los pesos  $v_{kj}$  hacia la

capa de salida, donde operando de forma análoga al caso anterior, las neuronas de esta última capa proporcionan la salida final de la red.

Este proceso se resume en lo siguiente (J. M. Marín Diazarique, 2014):

La entrada total o neta que recibe una neurona oculta  $j$ ,  $net_j^p$ , es:

$$net_j^p = \sum_{i=1}^N w_{ji} x_i^p + \theta_j \quad (31)$$

donde  $\theta_j$  es el umbral de la neurona que se considera como un peso asociado a una neurona ficticia con valor de salida igual a 1. El output de la neurona oculta  $j$ ,  $y_j^p$ , se obtiene aplicando una función, que llamaremos  $f(\cdot)$ , sobre su entrada neta:

$$y_j^p = f(net_j^p) \quad (32)$$

De igual forma, la entrada neta que recibe una neurona de salida  $k$ ,  $net_k^p$ , es:

$$net_k^p = \sum_{i=1}^H v_{ki} x_i^p + \theta_k \quad (33)$$

Y el valor de salida de la neurona de salida  $k$ ,  $y_k^p$ , es:

$$y_k^p = f(net_k^p) \quad (34)$$

El objetivo, durante la etapa de aprendizaje, es hacer mínimo el error entre la salida obtenida por la red y la salida deseada por el usuario.

De esta forma, en las redes back-propagation el aprendizaje es de tipo supervisado. Si quisiésemos expresar el error que se trata de minimizar para cada  $p$ :

$$E^p = \frac{1}{2} \sum_{k=1}^M (d_k^p - y_k^p)^2 \quad (35)$$

donde  $d_k^p$  es la salida deseada para la neurona de salida k. Se puede obtener una medida general de error mediante:

$$E = \sum_{p=1}^P E^p \quad (35)$$

#### **2.4.2 Fases en la aplicación de un perceptrón multicapa**

Una red de este tipo (perceptrón multicapa) se orienta a tratar de enfrentarse a dos tipos principales de problemas: Problemas de predicción, que consisten en la estimación de una variable continua de salida, a partir de la presentación de un conjunto de variables predictoras de entrada (nuestro caso) o problemas de clasificación, que consisten en la asignación de la categoría de pertenencia.



### 3 Metodología y fase experimental

#### 3.1 Aproximación al diseño experimental

El objetivo de este proyecto es, en última instancia, construir un modelo que explique el comportamiento de una variable de interés, la variable respuesta, como resultado del efecto de un conjunto de variables explicativas y utilizar este (o estos) modelos para realizar predicciones o tomar decisiones.

Con la intención de cuantificar la influencia de cada una de esas variables y sus interrelaciones con respecto a la variable objetivo, debemos construir un diseño experimental que nos permita obtener resultados que se adapten a las necesidades del proyecto.

Estas necesidades (y sus limitaciones) son principalmente:

1. Para el estudio estadístico multivariante con una incertidumbre razonable se necesita alcanzar una determinada masa crítica de resultados. Este número es, en principio, indeterminado y solo tras el estudio final de los datos podemos verificar que el número de datos ha sido adecuado. A priori, solo podemos utilizar estimaciones de experiencias previas o resultados bibliográficos. Se estima que necesitaremos, como mínimo, uno 50 resultados de cada una de las variables.

2. Debería evaluarse el efecto que sobre la variable objetivo tiene cualquier cambio en los factores explicativos, es decir, estos últimos a distintos niveles (o intervalos) y estudiar el efecto que cada variación induce en la respuesta.

3. En relación con el punto anterior y con el objeto de evaluar de forma consistente el efecto de cada variable se debe procurar que cada una de ellas varíe en un intervalo suficientemente amplio para tener en cuenta los posibles efectos de todos los valores que podrían darse en la realidad.

4. Por el principio de aleatorización en el diseño experimental se espera que todas las variables que puedan influir en el resultado y estén fuera de control del experimentador, se repartan aleatoriamente entre las observaciones.

5. El posible efecto de estas variables no identificadas pretende ser absorbido a través de aumentar el número de experimentos, considerando la aproximación del punto 1 ( $>50$ ) como razonable.

6. Teniendo en cuenta las heterogeneidades en las propiedades físicas de nuestro objeto de estudio (uniones soldadas), es indispensable realizar las pruebas que conducen a la obtención de las distintas variables de la forma más compacta posible, es decir, todos los ensayos de variables explicativas deben realizarse en una zona lo más próxima posible a aquella donde se determine la variable objetivo (CTOD).

7. Por su extensión, sencillez y universalidad, se utiliza un sencillo diseño experimental factorial completo de cuatro variables en tres y dos niveles con una repetición.

Como no podemos asegurar la independencia de cada uno de los factores ni existen herramientas tecnológicas para fijar a voluntad los niveles de cada variable este sencillo diseño experimental esboza las líneas generales que después deberán adaptarse a los resultados reales.

Es decir, es imposible crear una unión soldada que presente niveles bajos en ciertas variables y altos en otras, pues todo el proceso es la interacción de un gran número de factores. Lo que sí es posible es modificar las condiciones y parámetros con las que se crean las unidades experimentales (la soldadura de los cupones soldados de prueba) para favorecer que determinadas variables presenten niveles más o menos altos.

Tabla 2. *Resultados preliminares de un diseño de experimentos factorial*

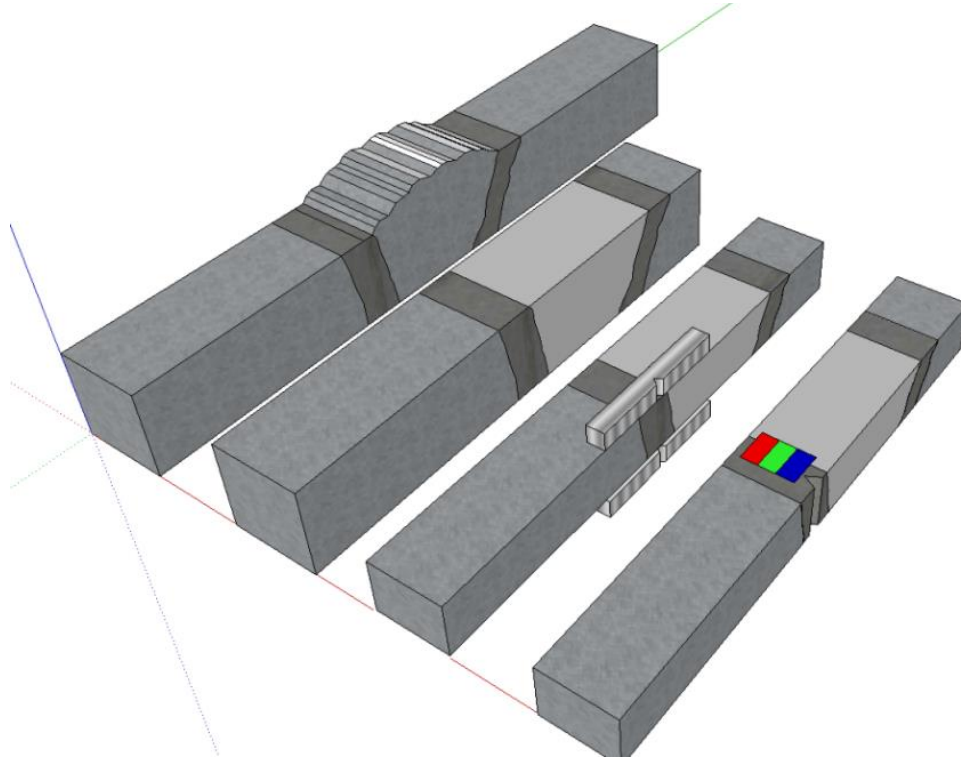
<b>N° Experimento</b>	<b>R. Mecánica</b>	<b>Tenacidad</b>	<b>Microestructura</b>	<b>C. Química</b>
<b>1</b>	1	2	1	2
<b>2</b>	2	2	2	2
<b>3</b>	2	1	2	2
<b>4</b>	2	1	1	2
<b>5</b>	2	2	1	1
...	...	...	...	...
<b>67</b>	2	3	2	1
<b>68</b>	1	3	1	2
<b>69</b>	1	1	1	2
<b>70</b>	3	1	1	1
<b>71</b>	1	1	1	1
<b>72</b>	1	1	1	2

En este diseño las variables Resistencia mecánica y Tenacidad tienen tres niveles, bajo, medio y alto (respectivamente 1, 2 y 3) para representar su mayor variabilidad. Se considera que las variables Microestructura y C. Química se valoran de forma adecuada con dos niveles, bajo y alto (1 y 2).

Para resumir, el estudio estadístico (con una incertidumbre razonable) se requiere alcanzar una masa crítica determinada de datos de entrada. Este número es indeterminado y se verificará después del modelado (Rencher, 2001). Además, se requiere un amplio rango que cubra el interés industrial para las variables explicativas.

Se espera que las heterogeneidades en las propiedades físicas de las uniones soldadas y las variables no controladas se distribuyan arbitrariamente entre las observaciones de acuerdo con el principio aleatorio (Yang, Mahfouf, & Panoutsos, 2012).

En todo caso, la posición de muestreo y ensayo se ha diseñado para minimizar el efecto de estas heterogeneidades siguiendo el diseño presentado a continuación, (Figura 9), donde se pretende reducir, tanto como sea posible, el área desde donde se obtienen los resultados. obtenido.



*Figura 9.* Posición de extracción de las probetas y zonas de ensayo. Las zonas de color marcan áreas específicas para el análisis microestructural, dureza HV10 y químico (rojo, verde y azul).

### 3.2 Muestreo y naturaleza de las muestras, obtención y controles de calidad preliminares (incluyendo las variables de soldadura utilizadas)

Cada una de las unidades de ensayo lo conforma un cupón soldado de al menos 350 mm de ancho y 300 mm de largo (en el sentido del eje de soldadura), lo cual puede no cumplir con alguna de las normativas referentes a cupones de prueba para cualificación de soldaduras. Sin embargo, dado que el número de pruebas es limitado, no es necesario tamaños mayores.

Es espesor de estos cupones de prueba es representativo de los utilizados en la fabricación de estructuras Offshore y varían entre 20-75mm.



*Figura 10.* Sección cilíndrica de una torre Offshore para un prototipo de aerogenerador de 8MW. Las probetas de homologación de la soldadura utilizada para unir la virola (soldadura longitudinal) y la virola a la brida (soldadura circular) fueron parte de set experimental.

Fuente: Fotografía del autor



*Figura 11.* Detalle del tratamiento térmico de precalentamiento para la unión de una de las soldaduras de prueba.

Fuente: Fotografía del autor

Se han utilizado como preparación de bordes biseles en K o biseles de tipo V (en ambos o un lado si el espesor es reducido).

Los materiales base y los consumibles de soldadura son los usuales dentro del sector y se contabiliza el uso de al menos tres consumibles diferentes y cuatro materiales base. El material base es acero o acero de alto límite elástico con propiedades mecánicas que los clasificarían como S275-S460 según la denominación de la UNE-EN 10225:2001.

Se presenta, a continuación, un cuadro resumen de los principales rangos de parámetros de soldadura utilizados en la fabricación de los cupones de prueba.

Tabla 3. Cuadro resumen de los principales rangos de parámetros de soldadura utilizados en la fabricación de los cupones de prueba

<b>Variable</b>	<b>Mínimo</b>	<b>Máximo</b>	<b>Variable</b>	<b>Mínimo</b>	<b>Máximo</b>
<b>Diámetro del hilo</b>	1.2 mm	2.8 mm	Material base	S275	S460
<b>Intensidad</b>	80 A	230 A	Nº Proc. Sold.	1	2
<b>Polaridad</b>	DC	AC	Voltaje	10 V	30 V
<b>Velocidad</b>	40 mm/min	240 mm/min	Tamaño del talón	No	Yes
<b>Heat Input</b>	0.6 KJ/mm	4.2 KJ/mm	Uso de respaldo	20 mm	75 mm
<b>Pre calentamiento</b>	No	150 °C	Espesor soldadura	PA	PF
<b>Flujo de gas</b>	10 l/min	25 l/min	Posición	35°	90°

Con objeto de minimizar el efecto de variables no controladas en los resultados de los experimentos se somete a todos los cupones de prueba a una batería de ensayos no destructivos. Se realiza sobre cada unidad de ensayo una inspección visual completa, tanto de aspecto superficial, defectos de soldadura o discontinuidades geométricas. La inspección visual se realiza por personal cualificado y según los criterios de la norma UNE-EN 5817:2013.

Asimismo, se realiza un examen por partículas magnéticas para identificar defectos superficiales o subsuperficiales. Este test se realiza con yugo portátil, partículas en base acuosa sobre laca de contraste blanca, en dos direcciones perpendiculares y con intensidades tangenciales dentro de las especificaciones de la norma UNE-EN ISO 17638:2010

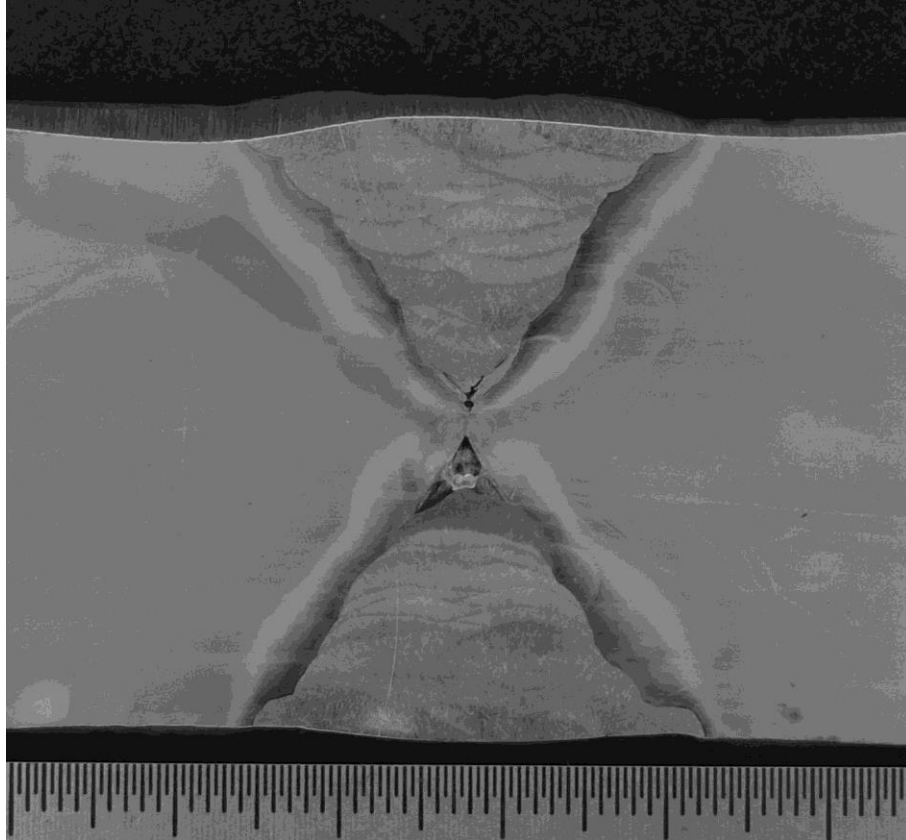
Para realizar una evaluación volumétrica se realiza un ensayo de ultrasonidos completo. Se utiliza un equipo convencional pulso-eco, al menos tres palpadores (dos de ellos angulares) y con frecuencias entre 2 y 5 MHz. Tanto el número de escaneos, el nivel de ensayo, los criterios de aceptación y la calibración del equipo se realizan según la norma UNE-EN ISO 17640:2011

Se identifican 14 zonas con imperfecciones de alguna clase. Estas zonas son marcadas y descartadas para su ensayo. Se trata de defectología leve como pequeñas mordeduras, sobreespesor del cordón, grietas o pequeñas faltas de fusión.

El resultado son 36 cupones soldados en condiciones representativas de las uniones soldadas de elevado espesor en industria Offshore. Estos cupones de prueba son evaluados para garantizar que ningún factor extraño (como una microfisura, un concentrador de tensiones o una bolsa de gas atrapada) afecte a los resultados.

Se utilizan sierras de cinta para desbastar cada uno de los cupones soldados. Se separan y extraen los bloques principales que después se mecanizarán con una máquina fresadora de control numérico.





*Figura 12.* Macrografía de una de las uniones soldadas del set experimental que fue rechazada al no superar las pruebas no destructivas. En este caso el ensayo por ultrasonidos. En el corte macrográfico puede observarse una falta de fusión en la raíz de la soldadura

### **3.3 Ensayos CTOD**

Como uno de los requisitos previos para la futura obtención del valor CTOD se extrae y ensaya una probeta de tracción cilíndrica en dirección transversal a la de laminación del material base. Esta probeta se somete a un ensayo de tracción a temperatura ambiente según la norma UNE-EN ISO 6892 donde se determina, principalmente, la carga máxima de rotura del material y el límite elástico, cuyo valor se necesitará posteriormente. Se utiliza extensometría de contacto, método B de control y velocidad en zona elástica de 15 MPa/s.

La variable objetivo CTOD es una variable continua, expresada en [mm] y con unos valores que suelen oscilar (para los materiales objeto de estudio) entre (0.1-2.5) mm. Esta magnitud (que físicamente mide la apertura del frente de grieta antes del colapso de esta) representa el comportamiento de un determinado tipo de materiales con respecto a su interacción con la dupla tensiones-grietas. Es usual determinar su valor para diferentes zonas de soldaduras críticas en grandes construcciones sometidas a esfuerzos de fatiga.

Para su determinación se han mecanizado 72 probetas de tamaño entre 200 y 500 mm de longitud, de ancho aproximadamente igual al espesor de los cupones de prueba y forma rectangular. Se han mecanizado las 6 caras hasta conseguir una rugosidad inferior a 2 Ra y se ha mecanizado por electroerosión una entalla artificial que sirve como iniciador de grieta. El pre-agrietado de las probetas se realiza con un equipo dinámico de 200 kN y tanto esta fase como la de ensayo se lleva a cabo según la norma ASTM E1290 (contemplada dentro del código de diseño DNV-GL que se utiliza como guía).

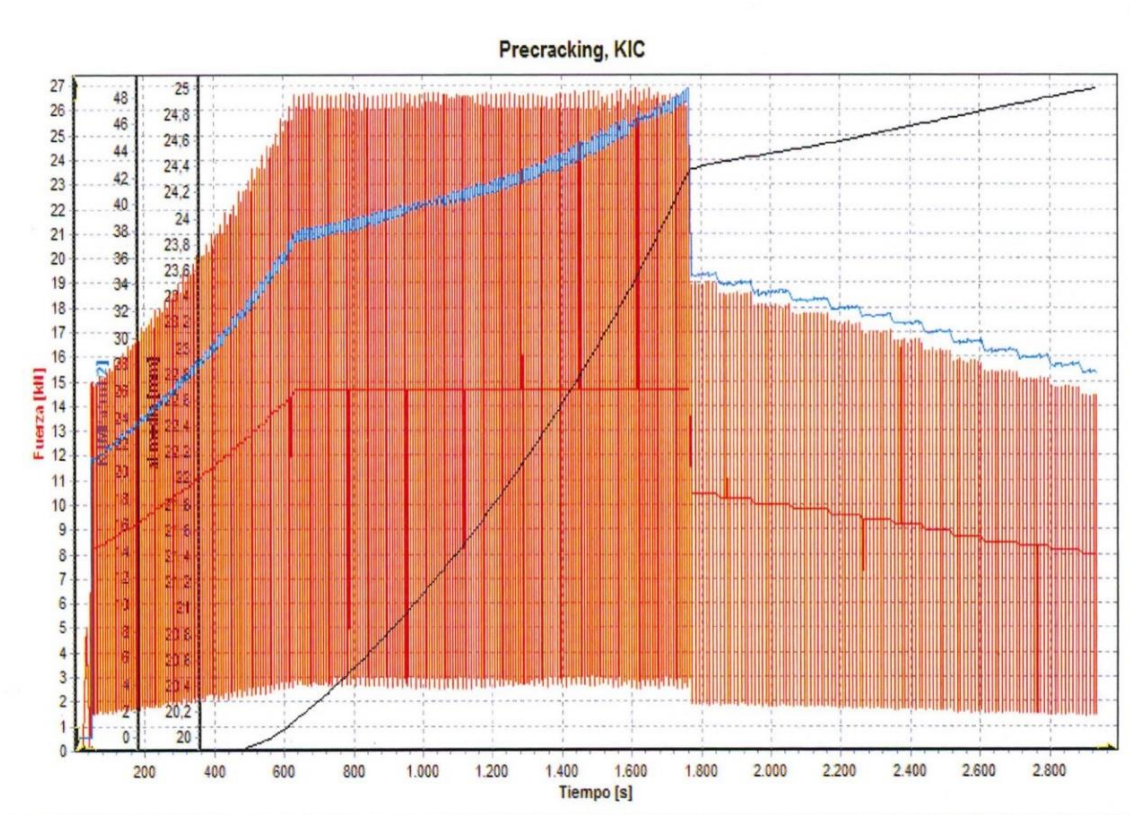


Figura 13. Diagrama de un proceso de pre-agrietado para el ensayo CTOD. El eje vertical se subdivide en; fuerza [kN], factor intensidad de tensiones [MPa.m<sup>1/2</sup>], amplitud de movimiento [mm]

Los 72 ensayos se realizarán a temperatura ambiente, pero en condiciones controladas (20-25 °C) y la entalla (zona efectiva de estudio) se situará en la zona afectada térmicamente de la probeta, a una distancia de 1 mm de la línea de fusión.

Todas las 72 pruebas CTOD se realizaron de acuerdo con ASTM E1290-08e1c (2008) con muestras estándar (muestra SE (B) de curvatura con muesca de un solo borde con sección transversal BxB cuadrada) y la muesca recomendada (Ávila et al., 2016). El aparato utilizado fue una máquina dinámica oleohidráulica (modelo UFIB-200E-MD5W) configurada para una configuración de flexión de 3 puntos y utilizando medidores de clip como sistema de medición

del crecimiento de grietas. Como muestra la Figura 1, la muesca se alineó a 1 mm de la línea de fusión.

El método de prueba elegido, ASTM E1290-08e1, calcula el valor CTOD con la siguiente expresión:

$$\delta = \frac{1}{m\sigma_Y} \left[ \frac{K^2(1-\nu^2)}{E} + \frac{\eta_{CMOD} A_{CMOD}^{pl}}{B(W-a_0) \left\{ 1 + Z / (0.8a_0 + 0.2W) \right\}} \right] \quad (36)$$

donde  $Z$  es la distancia de la cara frontal de las muestras SE (B) al punto de medición del filo de la cuchilla,  $A_{CMOD}^{pl}$  es el área plástica bajo carga desde la curva CMOD plástica, y la expresión de  $m$  es:

$$m = A_0 - A_1 \left( \frac{\sigma_{YS}}{\sigma_{ts}} \right) + A_2 \left( \frac{\sigma_{YS}}{\sigma_{ts}} \right)^2 - A_3 \left( \frac{\sigma_{YS}}{\sigma_{ts}} \right)^3 \quad (37)$$

donde,

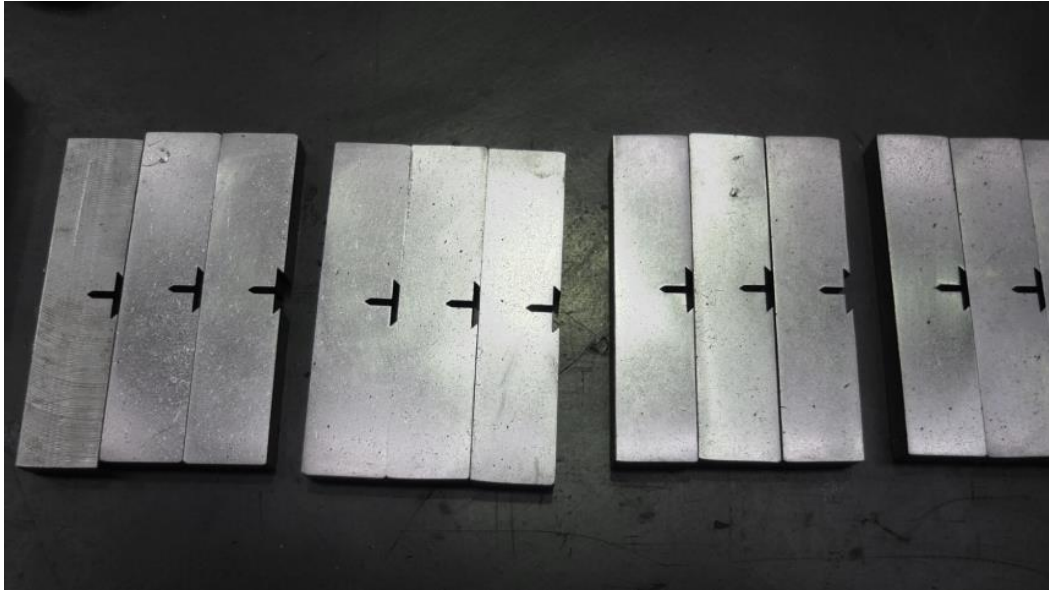
$$\begin{aligned} A_0 &= 3.18 - 0.22 \left( \frac{a_0}{W} \right), A_1 = 4.32 - 2.23 \left( \frac{a_0}{W} \right), A_2 \\ &= 4.44 - 2.29 \left( \frac{a_0}{W} \right), A_4 = 2.05 - 1.06 \left( \frac{a_0}{W} \right) \end{aligned} \quad (38)$$

y

$$\eta_{CMOD} = 3.667 - 2.199 \left( \frac{a_0}{W} \right) + 0.437 \left( \frac{a_0}{W} \right)^2, \quad (39)$$

Otros cálculos alternativos, fórmulas y predicciones fueron estudiados por (Khor et al., 2016-2017; Antunes et al., 2016-2917; Janssen, Zuidema y Wanhill (2004); Kawabata et al., 2016).

Todas las pruebas se realizaron en las instalaciones privadas de pruebas de laboratorio del grupo TAM (acreditación no. 808 / LE1532).



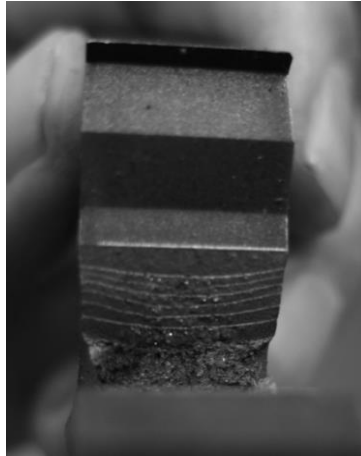
*Figura 14.* Parte de la colección de probetas (72) utilizadas para la determinación de la variable CTOD (en la imagen las correspondientes a espesor 25 mm).



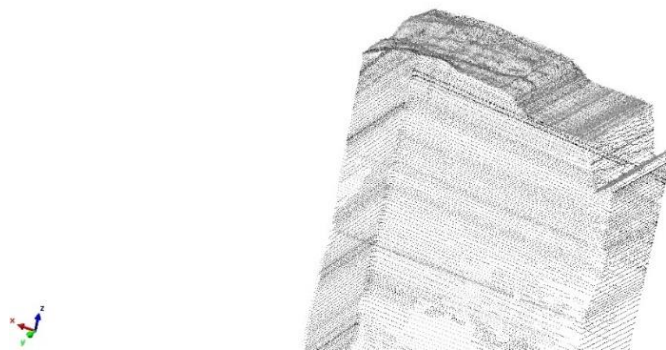
*Figura 15.* Detalle (1) de la superficie de fractura de una de las probetas CTOD.



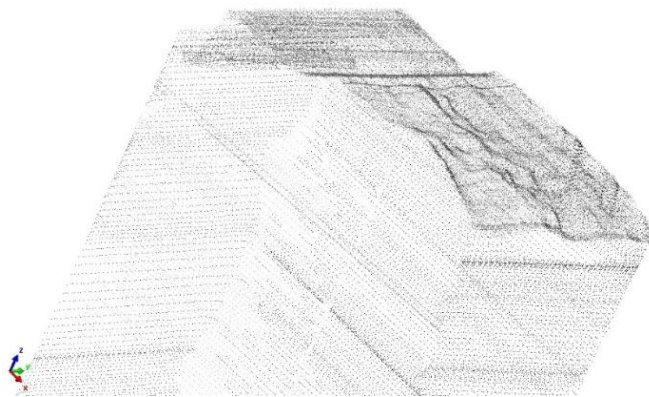
*Figura 16.* Detalle (2) de la superficie de fractura de una de las probetas CTOD.



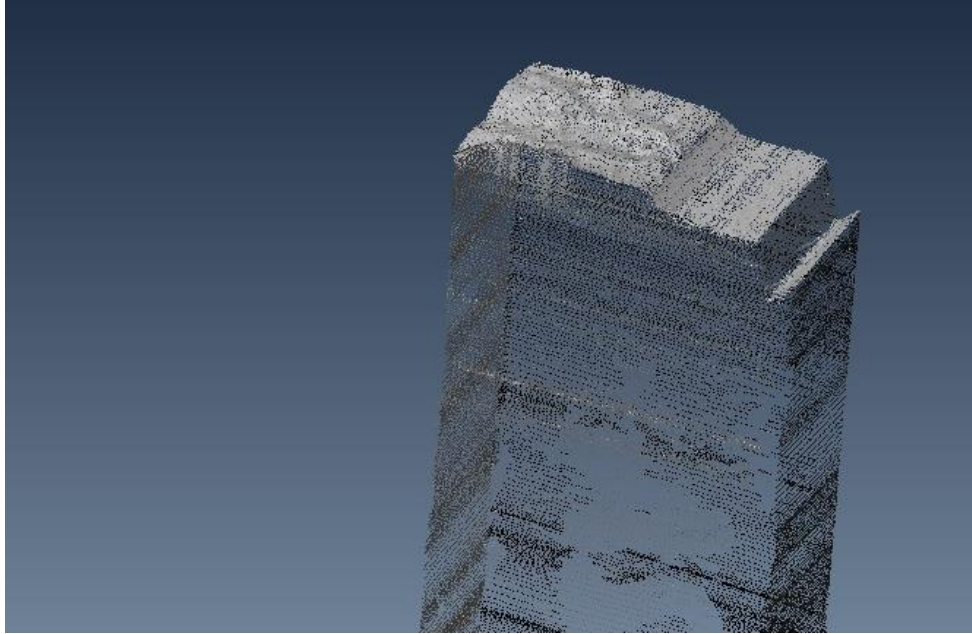
*Figura 17.* Detalle (3) de la superficie de fractura de una de las probetas CTOD.



(a)



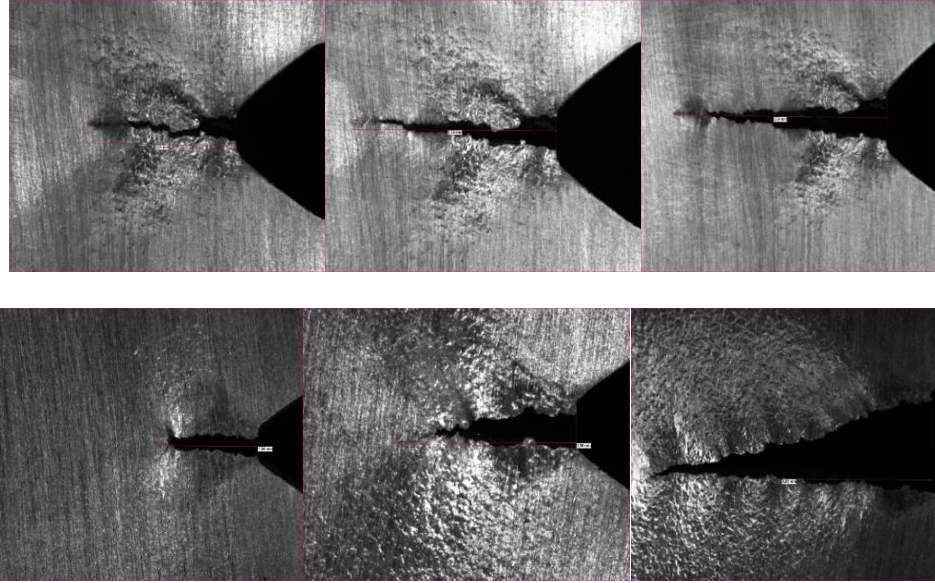
(b)



(c)

*Figura 18.* (a), (b) y (c) Modelos tridimensionales sobre las superficies de fractura con un brazo de metrología láser. La reflectividad de la superficie de fractura resultó no ser adecuada para este sistema de medición y se decidió utilizar sistemas ópticos.





*Figura 19.* Imágenes secuenciales (de izquierda a derecha) mostrando el crecimiento de grieta en dos de las probetas CTOD. Para realizar estas fotografías se detuvo el proceso de “pre-cracking” en distintos momentos. Después de cada fotografía se retomó el ensayo.

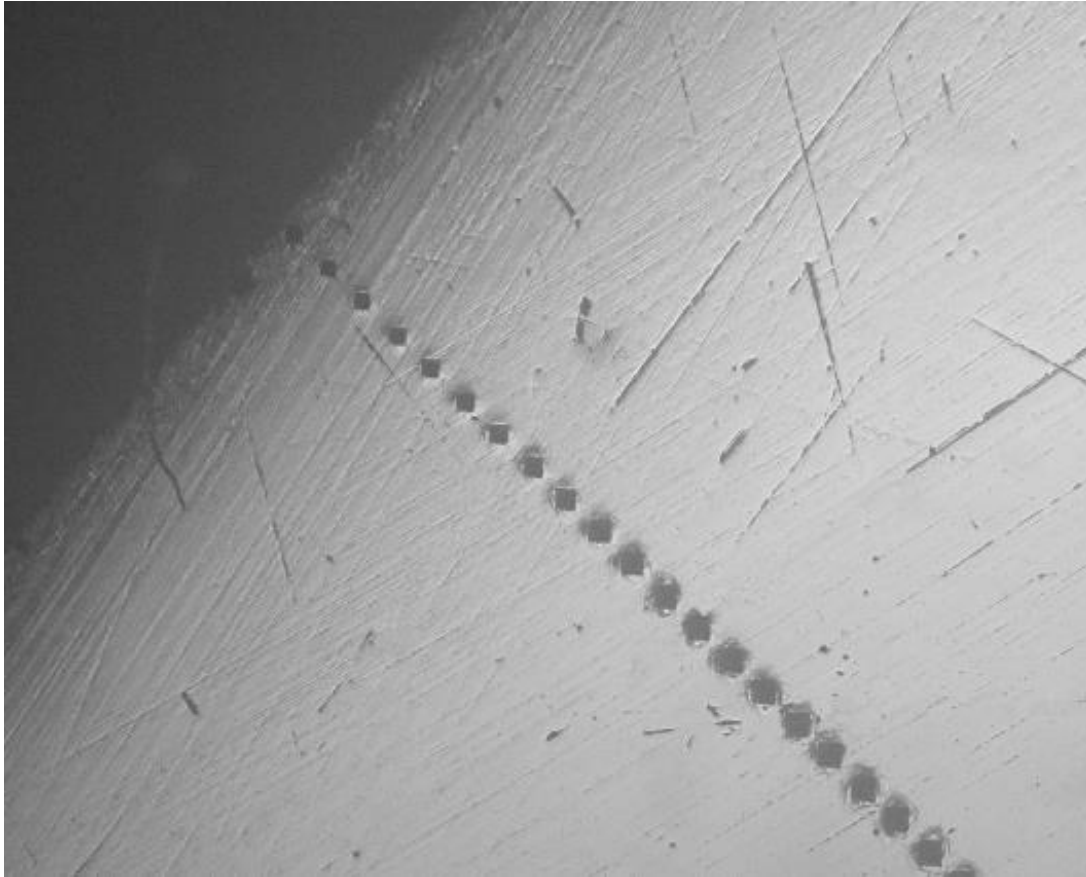
### 3.4 Ensayos de resistencia mecánica

Para la caracterización de la variable denominada “Resistencia mecánica” se utilizará la dureza Vickers del material en la zona de interés. La correlación entre dureza y resistencia a tracción del material es ampliamente utilizada a nivel industrial y responde a una relación lineal entre ambas. La principal diferencia es que un ensayo de tracción evalúa de forma global pues actúa sobre una zona relativamente grande, mientras que la dureza Vickers puede determinar el valor de la propiedad en una zona puntual de apenas unos  $\text{mm}^2$ .

Además, es una prueba barata, conocida a nivel industrial y que permite tener en cuenta la heterogeneidad de las propiedades mecánicas de las uniones soldadas (es decir, a lo largo de la sección de una soldadura vamos a encontrar valores muy diferentes separados apenas unos milímetros).

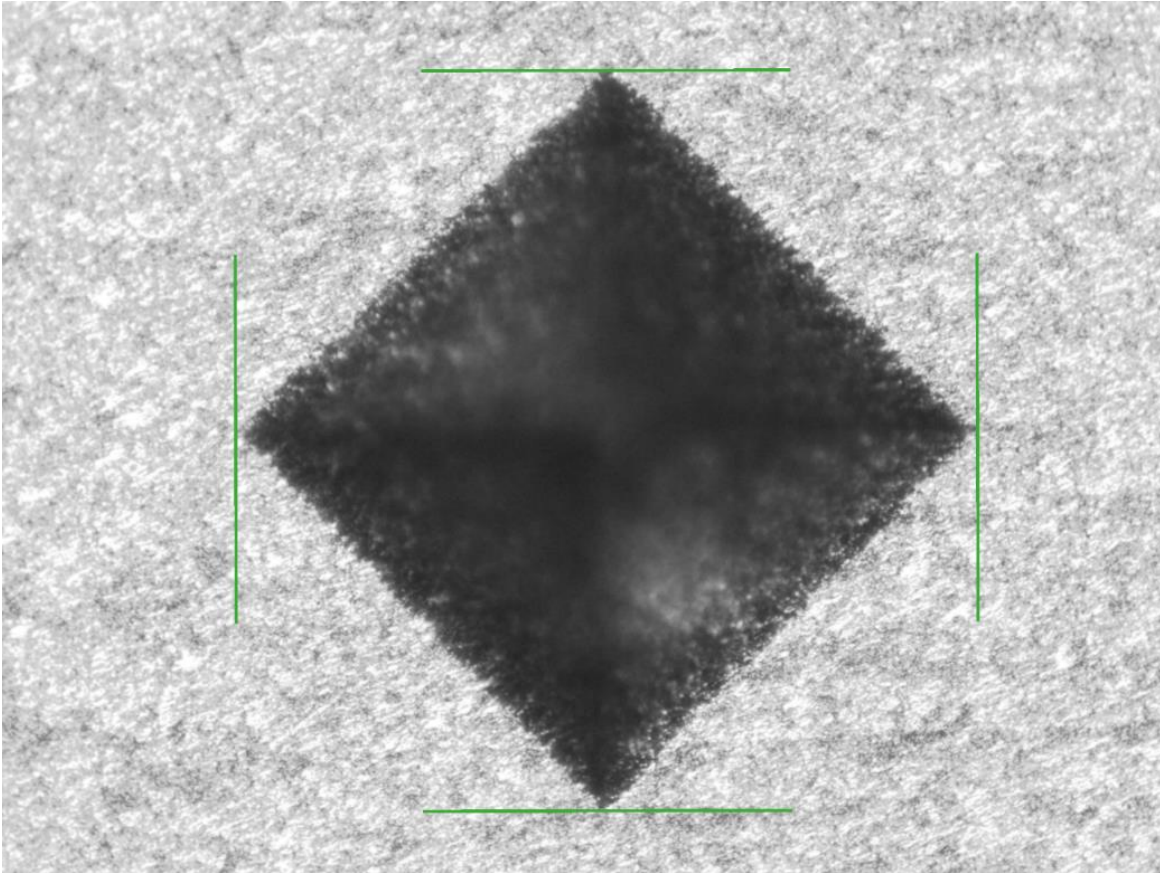
El durómetro es semiautomático y determina la dureza mediante la medición (en micras) de las diagonales de la huella cuadrangular de una indentación. Se utiliza como resultado la media de 3 mediciones. En general las mediciones no son especialmente sesgadas y no varían más de 5 unidades entre ellas. Todos los ensayos de dureza se realizan según la norma UNE-EN ISO 6507. Se realizan un total de 216 determinaciones de dureza, la incertidumbre media asociada a este ensayo es de 7 HV10.

Se usaron penetradores Vickers estandarizados (Clase B) con una carga de 98.7 N (HV 10) y una medición de sangría óptica. Se examinó el valor promedio de un conjunto de tres muescas (considerando 2 mm de espacio entre pruebas) para cada muestra.



*Figura 20.* En esta imagen se pueden observar una serie de indentaciones de dureza HV10 realizadas sobre una muestra transversal a una unión soldada. Para que sea fácilmente observable a bajos aumentos (10X) no se ha atacado la superficie con reactivos químicos.

Puede observarse la diferencia del tamaño de la huella mecánica sobre la superficie mientras las mediciones avanzan hacia la zona afectada térmicamente (arriba-izquierda). Este experimento ilustra la necesidad de determinar la dureza en una zona muy precisa para garantizar la representatividad considerando la heterogeneidad de la mismo.



*Figura 21.* Indentación real fotografiada a través del ocular del sistema semiautomático de medición de las dos diagonales principales.

### **3.5 Ensayo de composición química**

La composición química del material es un factor bien conocido que ejerce influencia sobre las propiedades mecánicas (Talas, 2010; Lee et al., 2017; Yurioka, 2001).

Las muestras se analizaron por espectrometría de emisión óptica y difracción de rayos X utilizando un analizador Niton® XL2 y un analizador de metales Spectromax. Los resultados se procesaron estadísticamente para ofrecer el estimador promedio mejor ponderado considerando las diferentes incertidumbres del método de prueba y para los siguientes elementos: C, Mn, Si, Cr, Ni, Mo y V. Tanto el procedimiento de prueba como el cálculo de incertidumbre utilizado

fueron aprobados por el laboratorio de pruebas. Para la implementación de la composición química en el modelo matemático, consideramos la influencia de los diferentes elementos utilizando el índice de carbono equivalente (CE), expresado en la ecuación (40). Entre las numerosas fórmulas de CE disponibles en la bibliografía, elegimos la American Welding Society (AWS) D1.1, que se citó en (Yurioka, 2001) y que también se conoce como el carbono equivalente del Instituto Internacional de Soldadura (IIW).

Esta expresión fue seleccionada teniendo en cuenta su precisión para las propiedades mecánicas y microestructurales (Talas, 2010):

$$CE_{index} = C + \frac{(Mn + Si)}{6} + \frac{(Cr + Mo + V)}{5} + \frac{(Ni + Cu)}{15} \quad (40)$$

donde todos los valores involucrados representan la composición porcentual de masa [w / w%]. Por lo tanto, el resultado es una variable continua no dimensional.

Por lo tanto, la variable “Composición química” está compuesta por entre 10 y 20 variables continuas correspondientes a la composición másica de cada elemento relevante del material. Para su determinación se utilizan dos tecnologías diferentes, espectrometría de emisión óptica y fluorescencia de rayos X (con la intención de cubrir posibles debilidades de una u otra).

La utilización de los mencionados índices pondera la tendencia del material a desarrollar microestructuras frágiles y está muy relacionado con el equilibrio alfa y gamma del acero. Teniendo en cuenta la conocida relación entre tenacidad y la tendencia a la fragilización parece especialmente adecuado utilizar este índice.

Se utiliza como resultado la media de 6 determinaciones para todos los elementos menos el carbono. De estas determinaciones 3 se realizarán con espectrometría óptica y 3 con

fluorescencia de rayos X. Para el caso del carbono solo podrán determinarse mediciones son espectrometría y se utilizará la media de tres mediciones.

Se realizan un total de 2808 mediciones que garantizan unos resultados fiables cuya incertidumbre es menor del 8%.

### **3.6 Ensayos para determinar la tenacidad**

La tenacidad es la capacidad del material para absorber energía dinámica antes de colapsar. Este valor se determina como una variable continua de energía expresada en [J]. Otra consideración relevante, que también se extrae del ensayo de tenacidad, es la propensión natural del material a fracturarse de modo frágil o dúctil.

Para su cálculo se realizarán tres ensayos de resiliencia con péndulo Charpy de 300 Julios y cuchilla R2 por unidad de ensayo, en los que se determinará el valor de tenacidad en J (media de los tres resultados) y se utilizará una tabla comparativa cualitativa para valorar, sobre la superficie fracturada de la probeta, el porcentaje aproximado de la misma que se ha roto de forma frágil o dúctil. La diferencia entre ambos modos de rotura es fundamental pues implica mecanismos diferentes y está fuertemente ligada a la microestructura y a la forma que tiene el material de reaccionar ante flujos tensionales. Lo que tendrá, presumiblemente, relevancia de cara a la variable objetivo CTOD.

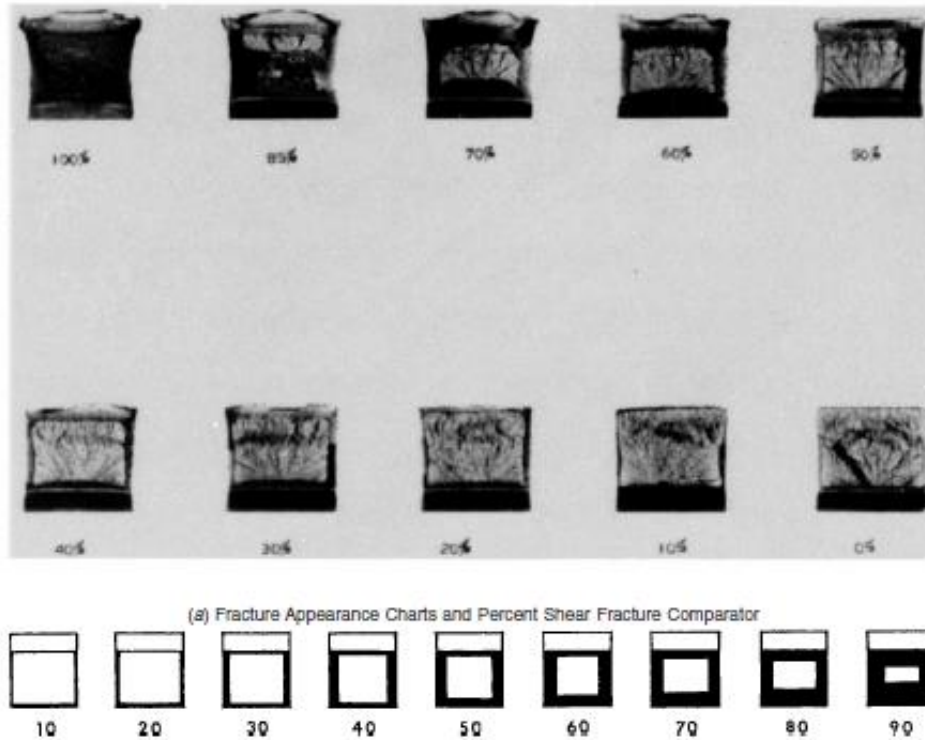


Figura 22. Tabla comparativa para el porcentaje de fractura dúctil-frágil.

Fuente: ASTM E23

Para tener en cuenta esta propiedad se consideró aplicar al valor de tenacidad un pequeño factor multiplicativo según sea frágil, frágil-dúctil o dúctil. En la práctica, la totalidad de los resultados fueron de rotura dúctil. Por lo tanto el uso de este factor se descarto en fases subsecuentes del proyecto. Los ensayos se llevarán a cabo según la norma UNE-EN ISO 148-1 y se obtienen un total de 216 determinaciones de energía y 216 determinaciones de porcentaje de rotura frágil-dúctil.

Las pruebas CVN, de acuerdo con ASTM E23 (2018), se realizaron en las muestras. Se utilizaron muestras de prueba de impacto de muesca en V de haz simple Charpy de tamaño medio (2,5 mm, de acuerdo con la Figura A3.1 de ASTM E-23), con la muesca alineada con la muesca futura de la muestra CTOD (Akourri et al., 2000). Todas las pruebas se realizaron a

temperatura ambiente (entre 20 y 25 ° C) con un dispositivo pendular de 300 J. Se utilizaron tres muestras (en lugar de dos) para cada caracterización de tenacidad para garantizar valores representativos (ver Figura 1), debido a la limitación del tamaño de la muestra. No se consideró la medición de la expansión lateral o el tamaño de la región de fractura.

### **3.7 Ensayos para caracterizar la microestructura**

Se considera relevante la microestructura del material en la zona en la que quiere determinarse el valor CTOD y por tanto es una de las variables explicativas. Se debe evaluar el tamaño medio de los granos en la zona de interés, la forma de estos en términos de acicularidad o esfericidad y la orientación con respecto a la dirección de propagación de la grieta en el futuro ensayo CTOD. Además, debe descartarse la presencia de martensita, ferrita widmanstatten, bainita, gran número de inclusiones de sulfuros y otras estructuras que puedan producir fragilización. De esta forma, la variable es continua y en [ $\mu\text{m}$ ] y producto de una variable continua (tamaño) y dos consideraciones cualitativas (forma y orientación). Teniendo en cuenta la zona de interés de este trabajo ni la forma ni la orientación deberían variar excesivamente.

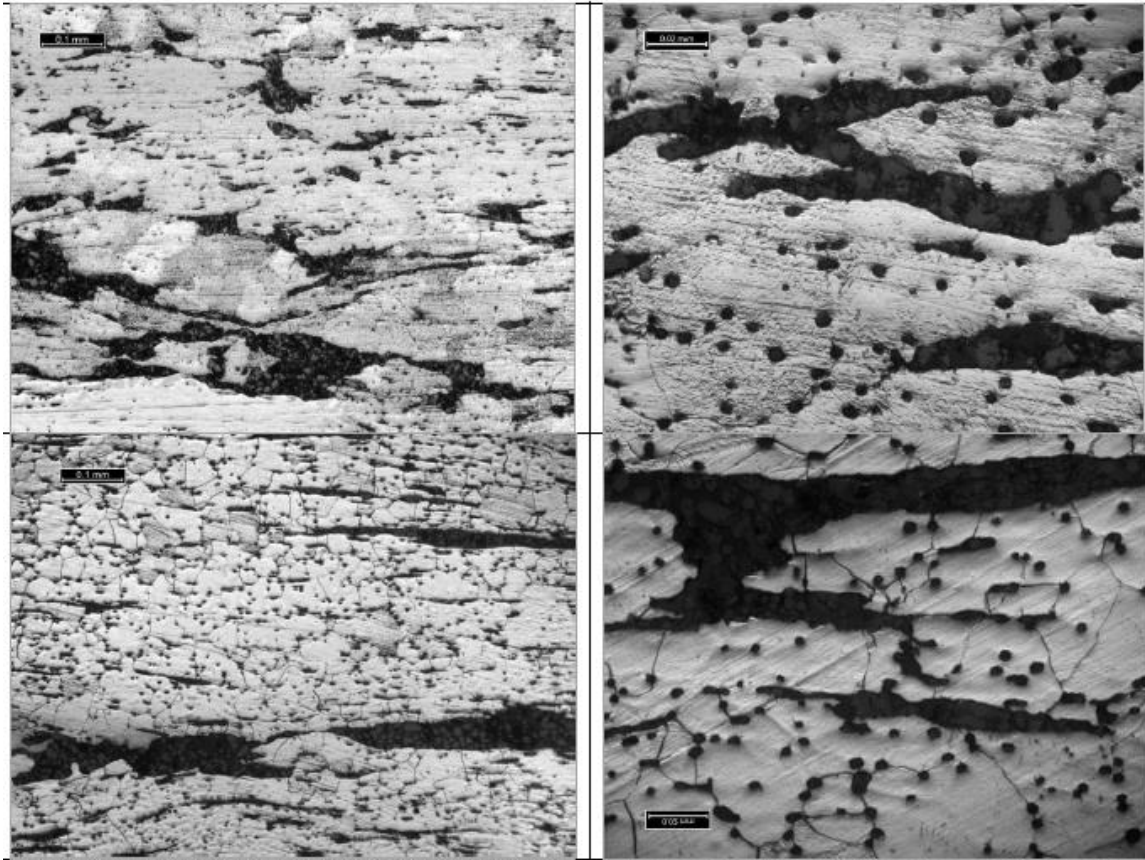
Para el análisis micrográfico se utiliza un microscopio óptico de 500X y un software de visión artificial. Se realizan 5 imágenes micrográficas por unidad de ensayo y el resultado es la media de cada una de las magnitudes determinadas. Un total de 360 imágenes tratadas.

La microestructura del material en el área de interés en la que se determinará el valor CTOD se considera una de las variables de entrada. Algunos autores [Castellucio, 2016; Cheng et al., 2018; Pegues et al., 2017; Wei et al., 2014; Yuan et al., 2017-2018] han estudiado la relación entre las características de la microestructura y las propiedades y resultados de los parámetros de mecánica de fractura, la influencia del tamaño del grano, el ángulo de los límites del grano, así como la influencia de la orientación y las inclusiones en la nucleación y la propagación de

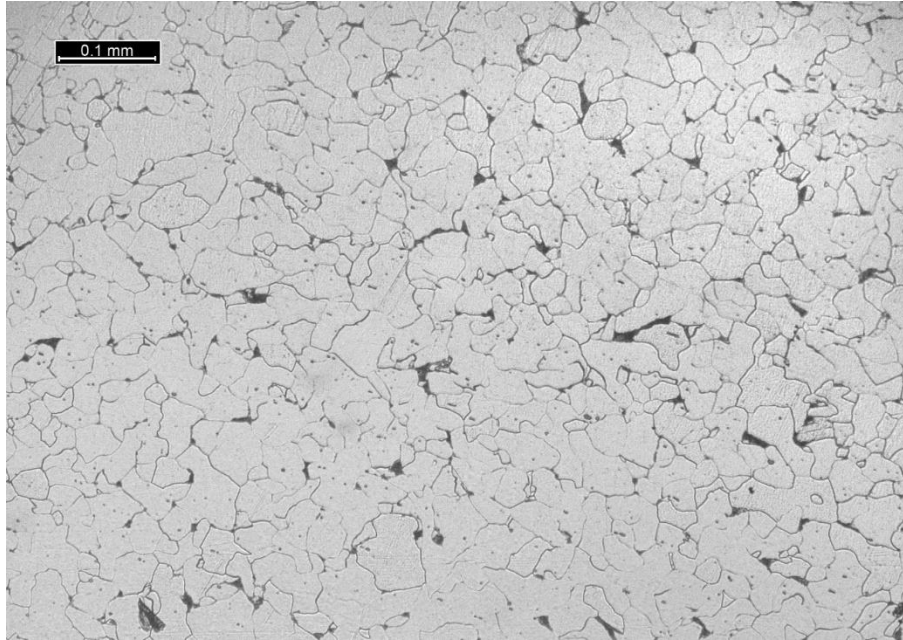


grietas. El tamaño promedio de los granos metálicos en el área de interés se determinó de acuerdo con ASTM E112 (2013) (determinado por microscopía óptica) para representar esta variable. Las muestras se pulieron y prepararon de acuerdo con la recomendación de E3-11 (Guía para la preparación de muestras metalográficas) con abrasivo de  $Al_2O_3$  (número de grano 1200 según la American National Standards Institute), rotación y reactivo de ataque químico número 77 (E407-07 Standard Practice for Microteaching Metals and Alloys).

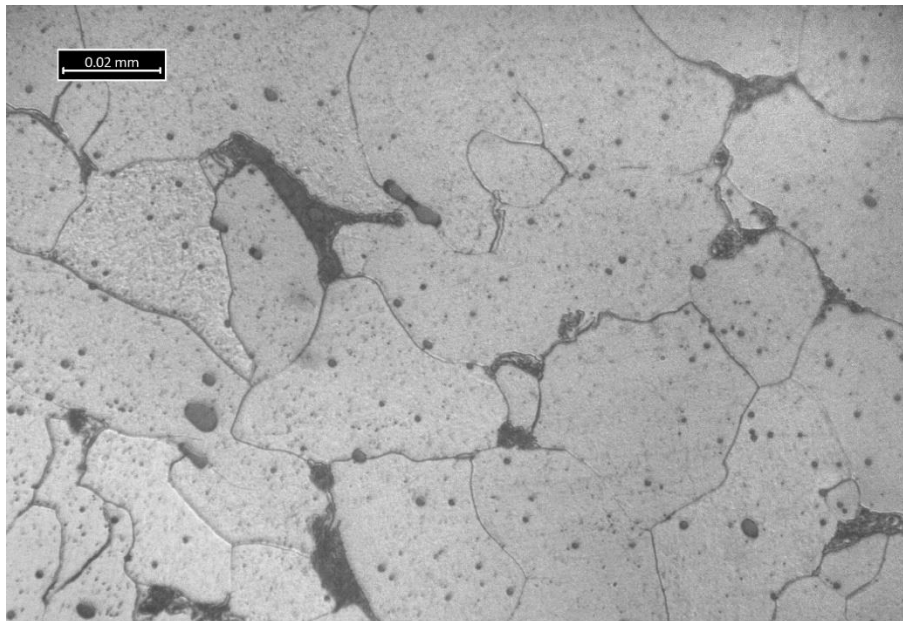
La limitación en la obtención de muestras con diferentes orientaciones de la superficie de análisis, que fuesen apropiadas para eventuales formas de grano no equiaxiales, se corrigió con la implementación de un factor de multiplicación arbitrario, dependiendo del contorno del grano. Se espera que cualquier posible heterogeneidad en el área de interés esté cubierta estadísticamente por el diseño experimental. Habiendo considerado el enfoque industrial, no se utilizaron otras técnicas bien conocidas que requerirían equipos específicos, como el microscopio electrónico de barrido (SEM). Es importante mencionar que el modo de fallo no se considera variables de estudio y, por lo tanto, no se consideró el tamaño de grano crítico para la fractura frágil en este análisis.



*Figura 23.* Micrografía a 100X (izquierda) y 500 X (derecha). Se puede apreciar un altísimo contenido de inclusiones metálicas esféricas (en este caso correspondientes a sulfuros metálicos). Esta muestra fue descartada debido a su alto contenido en azufre (por encima de los requisitos que estipula la norma de producto para el material base). Estas inclusiones metálicas tienen conocidos efectos perniciosos en la tenacidad a la fractura.



*Figura 24.* Micrografía a 100X de una muestra aceptada para el set experimental. Imágenes utilizadas para la determinación del tamaño y morfología de grano.



*Figura 25.* Micrografía a 500X. Detalle de la misma muestra que la anterior.

### 3.8 Resumen de los datos obtenidos durante la fase experimental

Se presenta a continuación a título de ejemplo un extracto de los datos recogidos para dos de las 72 unidades de ensayo.

Tabla 4. Extracto de los datos recogidos (Muestra 01 y 72)

ID:		01-CWS									
Dureza HV 10		Tenacidad		% Fractura D-F		Variable tenacidad		Tamaño de grano		Morfología	
1	197	1	129	1	D-F	<b>151 J</b>		1	147	1	Esférica
2	196	2	132	2	D-F			2	164	2	Esférica
3	189	3	100,5	3	D-F			3	163	3	Esférica
$\bar{x}$	<b>194 HV10</b>	$\bar{x}$	120,5	$\bar{x}$	D-F			$\bar{x}$	<b>158 <math>\mu\text{m}</math></b>	$\bar{x}$	Esférica
Análisis químico								Orientación		V. microestructura	
Elemento	1	2	3	4	5	6	$\bar{x}$	1	-	<b>158</b>	
Mn	1,12	1,11	1,12	1,08	1,07	1,07	1,1	2	-		
Cr	0,02	0,018	0,021	0,018	0,024	0,023	0,021	3	-		
Mo	0,003	0,003	0,002	0,003	0,002	0,003	0,003	$\bar{x}$	-		
V	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD	0,0				
Ni	0,02	0,03	0,02	0,02	0,02	0,003	0,019			CTOD	
Cu	0,06	0,06	0,05	0,04	0,04	0,04	0,05			1,14 mm	
C	0,018	0,017	0,017	-	-	-	0,017				
						Ceq	<b>0,31</b>				
...	...	...	...	...	...	...	...	...	...	...	...

ID:		72-CWS									
Dureza HV 10		Tenacidad		% Fractura D-F		Variable tenacidad		Tamaño de grano		Morfología	
1	215	1	115	1	D-F	<b>174 J</b>		1	174	1	Esférica
2	224	2	134	2	D-F			2	181	2	Esférica
3	211	3	168	3	D-F			3	172	3	Esférica
$\bar{x}$	<b>217 HV10</b>	$\bar{x}$	139	$\bar{x}$	D-F			$\bar{x}$	176 $\mu\text{m}$	$\bar{x}$	Esférica
Análisis químico								Orientación		V. microestructura	
Elemento	1	2	3	4	5	6	$\bar{x}$	1	-	<b>176</b>	
Mn	0,89	0,91	0,94	1,06	1,04	1,04	1,0	2	-		
Cr	0,03	0,03	0,04	0,04	0,05	0,05	0,040	3	-		
Mo	0,01	0,01	0,008	0,02	0,018	0,021	0,015	$\bar{x}$	-		
V	0,001	0,001	0,001	<LOD	<LOD	<LOD	0,0				
Ni	0,81	0,76	0,79	1,02	1,01	1,03	0,903			CTOD	
Cu	0,04	0,04	0,04	0,04	0,03	0,04	0,04			1,26 mm	
C	0,024	0,024	0,022	-	-	-	0,023				
						Ceq	<b>0,42</b>				

## **4 Análisis**

### **4.1 Análisis preliminar de los datos**

Los diferentes coeficientes de correlación se determinaron entre las distintas variables de estudio. Se puede observar un alto grado de correlación entre algunas de las variables, como la tenacidad (CVN) y la microestructura, lo que está fuertemente respaldado en la bibliografía (relación de Hall-Petch). Estas relaciones también dependen de otras variables que no se han considerado en este trabajo, como la temperatura, el estado de tensión residual o la geometría de la muestra. Por lo tanto, esta relación particular entre ambas variables es exclusiva de este experimento y no se puede considerar como una generalidad extensiva.

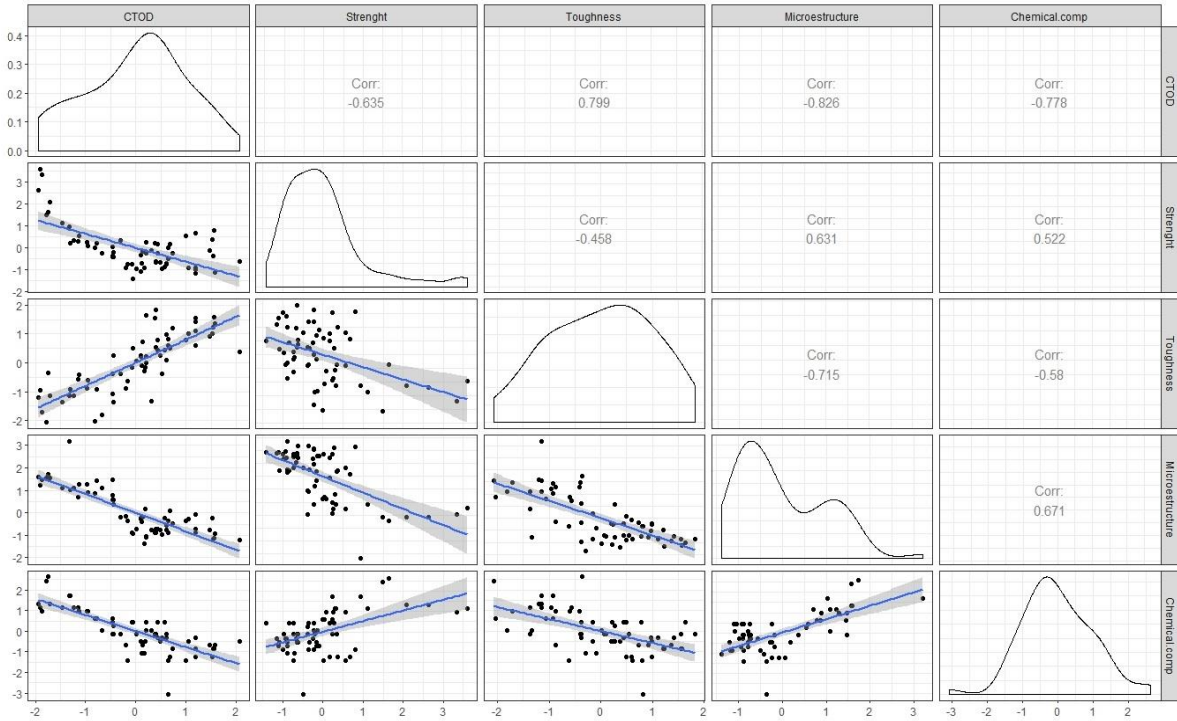


Figura 26. Correlación, KDE (estimación de densidad del Kernel) y diagramas de dispersión (también se representa la línea de tendencia que mejor se ajusta con una regresión lineal en color azul) entre las diferentes variables.

La Figura 26 muestra los diagramas de correlación y diagramas de dispersión entre todas las variables (objetivo y explicativas) tomadas de dos en dos. La representación de la estimación de densidad del núcleo (KDE) también es una forma de estimar la función de densidad de probabilidad de una variable aleatoria. Se puede observar una fuerte correlación entre el CTOD y las variables explicativas, particularmente la tenacidad, la microestructura y la composición química. Excluyendo la composición química, otras variables no parecen seguir una distribución normal.

La Figura 27 muestra los cuantiles de las muestras de entrada (variables explicativas) con respecto a los cuantiles normales estándar (cuantiles teóricos de una distribución normal). Si la distribución de la variable explicativa es normal, la gráfica estará cerca de aparecer como lineal.

Excepto por la composición química y la tenacidad, el resto de las variables independientes (*resistencia mecánica* y *microestructura*) no parecen seguir una distribución normal, por lo que sería aconsejable realizar una transformación previa (por ejemplo, tipo logarítmico) antes de realizar un análisis de regresión multivariante. Esto puede explicarse por la observación del KDE para la variable correspondiente en la *Figura 26*, donde la variable **resistencia mecánica** muestra un sesgo positivo hacia valores más bajos y la *microestructura* una distribución ligeramente bimodal. Este efecto se elimina mediante una transformación logarítmica después de la exclusión de valores atípicos. Esta transformación modera los valores extremadamente altos y magnifica los significativamente más bajos mediante la ecuación general  $b_{ij} = \log(x_{ij} + 1)$  (la inclusión del +1 evita cualquier indefinición de  $\log(0)$ ).

Con el objetivo de descartar los valores atípicos que pueden ser observaciones influyentes, se utilizó la distancia de Mahalanobis (Cuadras, 1981; Everitt, 1993) para su detección, y se excluyeron diez conjuntos de datos completos (14%).

Este parámetro, introducida por Mahalanobis (1936) es una forma de determinar la similitud entre dos variables aleatorias multidimensionales (Salas & Portillo, 2008). Como nota al margen se diferencia de la distancia euclídea en que tiene en cuenta la correlación entre las variables aleatorias lo que la hace más adecuada en nuestro caso).

La expresión general es:

$$d_m(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T \Sigma^{-1} (\vec{x} - \vec{y})}, \quad (41)$$

donde  $\vec{x}$  e  $\vec{y}$  son las dos variables con la misma distribución y  $\Sigma$  la matriz de covarianza entre ellas.

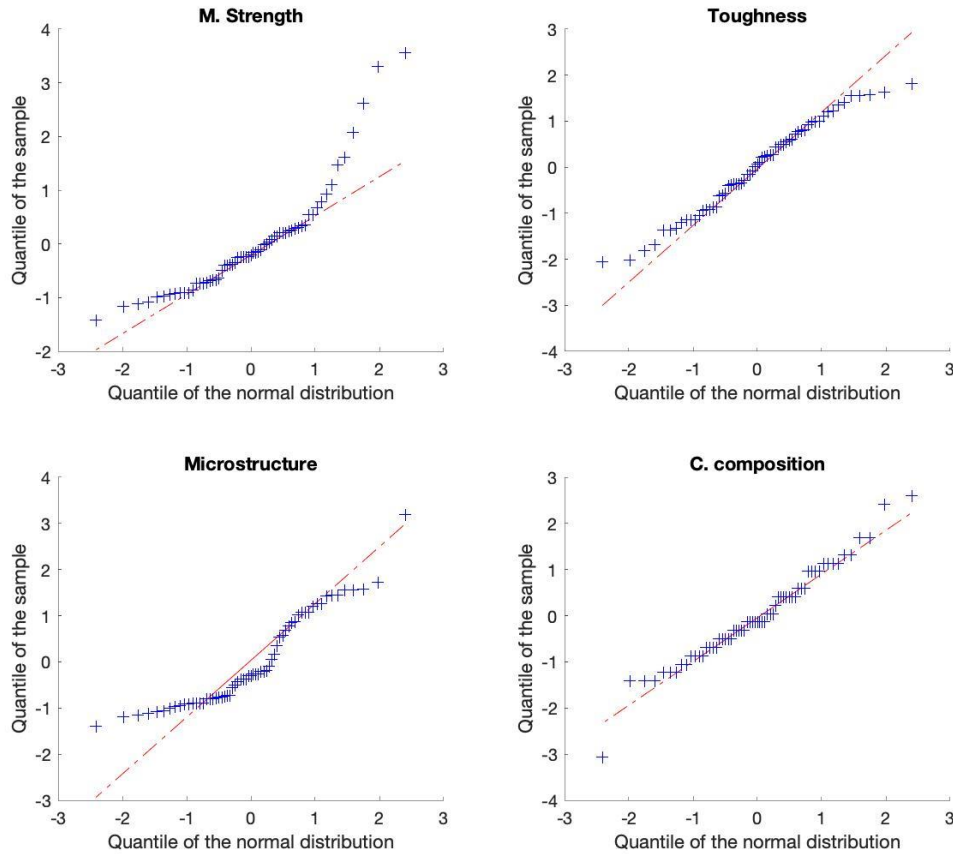


Figura 27. Cuantiles de la muestra con respecto a los cuantiles normales estandarizados

## 4.2 Modelos de regresión lineal

### 4.2.1 Modelo de regresión lineal 1



Consideraremos  $Y$  como la variable de estudio (CTOD) que puede estar relacionada linealmente con  $K$  variables explicativas  $X_1, X_2, \dots, X_K$  a través de  $X_1, X_2, \dots, X_K$  (coeficientes de regresión). Un modelo de regresión lineal múltiple se podría describir entonces como:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K + e \quad (42)$$

donde  $e$  es la diferencia entre el ajuste del modelo y las observaciones realizadas (Rao et al., 2008).

Usando las fórmulas (22) y (23), se calculan los valores de los distintos parámetros. En la Tabla X se encuentran los coeficientes para esta ecuación de regresión lineal múltiple (42). Podemos comprobar que todos los coeficientes son significativamente diferentes de cero. No obstante, la tenacidad es la variable con el valor absoluto más alto. En este caso, el número de observaciones es 63, siendo los grados de libertad 58.

El error cuadrático medio es  $RMSE = 0.216$ , el cual, en comparación con el rango de los valores de la variable de estudio  $Y$  da como resultado:

$$\frac{RMSE}{(Y_{MAX} - Y_{MIN})} = 0.1048 \approx 10\%, \quad (43)$$

Lo que proporciona una estimación del posible error obtenido en los valores reales de la variable CTOD. En la Figura X podemos observar que el coeficiente de correlación entre CTOD y la tenacidad es 0.799. Considerando todas las variables independientes, el RSQ es 0.866, y el valor ajustado de RSQ es 0.856, por lo que hay una mejora limitada al considerar la correlación de la tenacidad y el CTOD (o la microestructura y el CTOD).

Tabla 5. Modelo de regresión lineal  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$ 

Parámetros	Estimadores	SE	p-Value
$\beta_0$	1.2202	0.022731	$1.0629 \times 10^{-46}$
$\beta_1 - M.Str.$	-0.080323	0.035816	$2.8756 \times 10^{-2}$
$\beta_2 - Tough.$	0.22424	0.040094	$6.3259 \times 10^{-7}$
$\beta_3 - Micros.$	-0.12972	0.047887	$8.8595 \times 10^{-3}$
$\beta_4 - C.Comp.$	-0.19243	0.038258	$5.0415 \times 10^{-6}$
F-Statistic p-value	-	-	$1.4 \times 10^{-24}$

En adelante, para los modelos mostrados, se calculará e incluirá el estadístico  $t$  (tStat) y el estadístico  $F$ . El primero de ellos, tStat, calculado como Estimador / Standard-Error (SE), prueba la hipótesis nula de que el coeficiente correspondiente es cero frente a la alternativa de que es diferente de cero. Para evaluar este coeficiente, se calcula el valor  $p$  (p-valor) correspondiente asociado con una distribución  $t$  de Student (para  $n$  observaciones) y se compara con un intervalo de confianza del 95%. Si el valor  $p$  es menor que 0.05, podemos concluir que la variable es significativa para el modelo.

Análogamente, el estadístico  $F$ , calculado como:

$$F = \frac{\sum_{i=1}^n \frac{(\hat{y}_i - \bar{y})^2}{(p-1)}}{\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{(n-p)}} \quad (44)$$

Prueba la hipótesis nula de que uno o más de los coeficientes de regresión son significativamente diferentes de cero (por lo tanto, existe una relación de regresión lineal significativa para todo el modelo).

Este valor se compara con una distribución tipo  $F$  para un intervalo de confianza dado (95%) y se evalúa de la misma manera que el estadístico  $t$  (valor  $p$  asociado menor que 0.05). La distribución  $F$  es más apropiada que las pruebas “Chi-Cuadrado” (Chi-Square) para pequeños conjuntos de datos (Goldstein, 2014).

Se utilizaron dos métodos diferentes para verificar que el modelo obtenido es independiente de la población de datos elegida: muestras de validación cruzada y test de entrenamiento. La validación cruzada se calculó con el método LeaveMout (dentro del paquete de funciones de MATLAB “crossvalind”) con un valor  $M=1$  que selecciona aleatoriamente un valor y lo excluye de la evaluación. Este proceso repetido 50 veces ayuda a evaluar, llegado el caso, verificar que el análisis estadístico sea independiente del conjunto de datos. El número de observaciones fue de 62, con un RMSE = 0.218, RSQ = 0.866 y RSQ ajustado = 0.856. Los resultados se muestran en la Tabla 5.

Tabla 6. *Resultados de la validación cruzada*

Parámetros	Estimadores	SE	p-Value
$\beta_0$	1.2215	0.027668	$8.971 \times 10^{-46}$
$\beta_1 - M.Str.$	-0.08113	0.036144	$2.8692 \times 10^{-2}$
$\beta_2 - Tough.$	0.22244	0.040667	$1.0436 \times 10^{-6}$

$\beta_3 - \text{Micros.}$	-0.13334	0.049165	$8.8225 \times 10^{-3}$
$\beta_4 - \text{C. Comp.}$	-0.19006	0.039041	$9.3064 \times 10^{-6}$
F-Statistic p-value	-	-	$3.5 \times 10^{-24}$

---

La prueba de entrenamiento se realizó considerando un conjunto de 500 ejecuciones de muestras formadas por 50 observaciones (seleccionadas al azar de entre todo el conjunto de datos) y muestras de prueba de 13 conjuntos de datos. Los promedios de todos los resultados de RMSE y RSQ son  $RMSE = 0.2275$  y  $RSQ = 0.8284$ , respectivamente. Es decir, se seleccionan aleatoriamente conjuntos de 50 observaciones con las que se construye un modelo equivalente al *Modelo de regresión lineal 1*, se utilizan 13 conjuntos de datos para verificar el funcionamiento del modelo y se repite este proceso 500 veces.

La Tabla 6 contiene los valores de RSQ y RMSE obtenidos con el modelo de referencia (*Modelo de regresión lineal 1*), validación cruzada y prueba de entrenamiento. Como los valores son similares (menos del 5% de discrepancia), podemos concluir que la relación entre el CTOD y las variables explicativas es independiente del conjunto de datos utilizado.

Tabla 7. Comparación de los valores de R-squared (RSQ) y el error cuadrático medio (RMSE)

Parámetros	RSQ	RMSE
<b>Modelo de regresión lineal 1</b>	0.866	0.216
<b>Referencia cruzada</b>	0.866	0.218
<b>Test de entrenamiento</b>	0.828	0.227

#### 4.2.2 Modelo de regresión lineal 2

La relevancia de todas las variables consideradas (todas las variables explicativas) ha sido evaluada mediante la prueba del p-valor. Sin embargo, considerando que, a raíz de lo observado

previamente, la microestructura parecía estar altamente correlacionada con la tenacidad, se propone un nuevo modelo (*Modelo de regresión lineal 2*), donde la microestructura se eliminó del modelo original.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 \quad (45)$$

La *Tabla 7* muestra los valores de los parámetros calculados para el *Modelo de regresión lineal 2*, donde el ajuste obtenido (RMSE = 0.227, RSQ = 0.849 y RSQ ajustado = 0.841) fue similar al anterior (*I*).

Tabla 8. *Modelo de regresión lineal 2*  $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4$

Parámetros	Estimadores	SE	p-Value
$\beta_0$	1.2202	0.028657	$5.353 \times 10^{-46}$
$\beta_1 - M.Str.$	-0.11829	0.034686	$1.1768 \times 10^{-3}$
$\beta_2 - Tough.$	0.27944	0.036337	$1.8301 \times 10^{-10}$
$\beta_4 - C.Comp.$	-0.22775	0.03785	$1.2105 \times 10^{-7}$
F-Statistic p-value	-	-	$3.79 \times 10^{-24}$

#### 4.2.3 Modelos de regresión lineales 3 y 4

Teniendo en cuenta que el valor del parámetro  $\beta_1$  (coeficiente correspondiente a la variable explicativa *-Resistencia mecánica-*) en el *Modelo de regresión lineal 1* era significativamente más pequeño que el resto de los parámetros correspondientes al resto de variables explicativas se elimina esta variable para construir un nuevo modelo (*Modelo de regresión lineal 3*). De esta forma se considera que la información que esta variable aporta al modelo no es significativa. Los valores de los coeficientes de este nuevo modelo se presentan en la siguiente tabla (*Tabla 8*).

Tabla 9. *Modelo de regresión lineal 3*  $Y = \beta_0 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$ .

<b>Parámetros</b>	<b>Estimadores</b>	<b>SE</b>	<b>p-Value</b>
$\beta_0$	1.2202	0.028146	$1.9133 \times 10^{-46}$
$\beta_2 - Tough.$	0.22225	0.04143	$1.4246 \times 10^{-6}$
$\beta_3 - Micros.$	-0.17174	0.045549	$3.793 \times 10^{-4}$
$\beta_4 - C. Comp.$	-0.20715	0.038957	$1.961 \times 10^{-6}$
F-Statistic p-value	-	-	$1.32 \times 10^{-24}$

La calidad del ajuste,  $RMSE = 0.223$  y  $RSQ = 0.854$  es muy similar a la del modelo que utiliza las cuatro variables independientes. La Figura X muestra los residuos del *Modelo de regresión lineal 3*, que pueden considerarse distribuidos de forma normal.

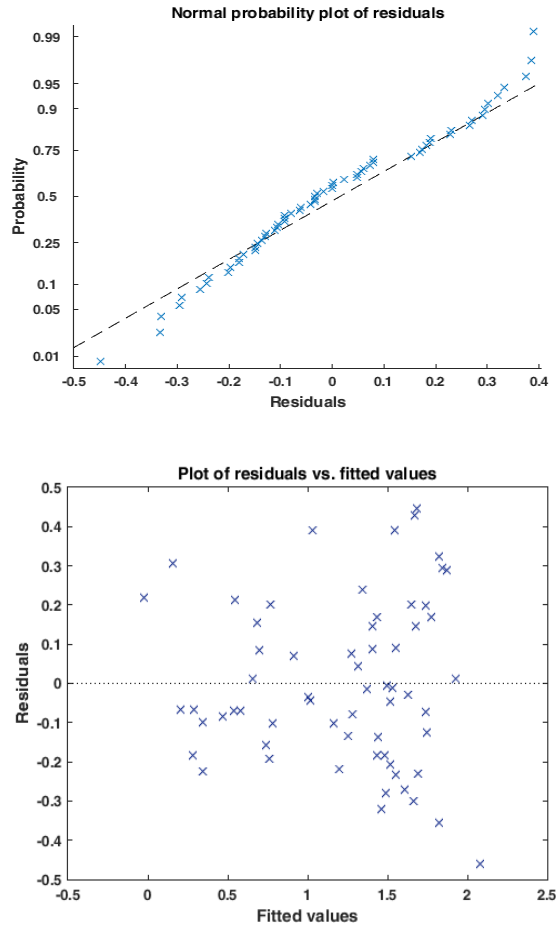


Figura 28. Distribución normal de los residuos (Arriba) Gráfica de probabilidad normal de los residuos (Abajo) Gráfica de residuos con respecto a los valores ajustados.

Finalmente, se adopta un nuevo modelo (*Modelo de regresión lineal 4*) considerando el cuadrado de la primera variable ( $\beta_1 X_1^2$ ), y se verifica la contribución de las variables independientes a la variable CTOD (Tabla 9). En este caso, el coeficiente de determinación RSQ = 0.874, mayor que en el modelo puramente lineal.

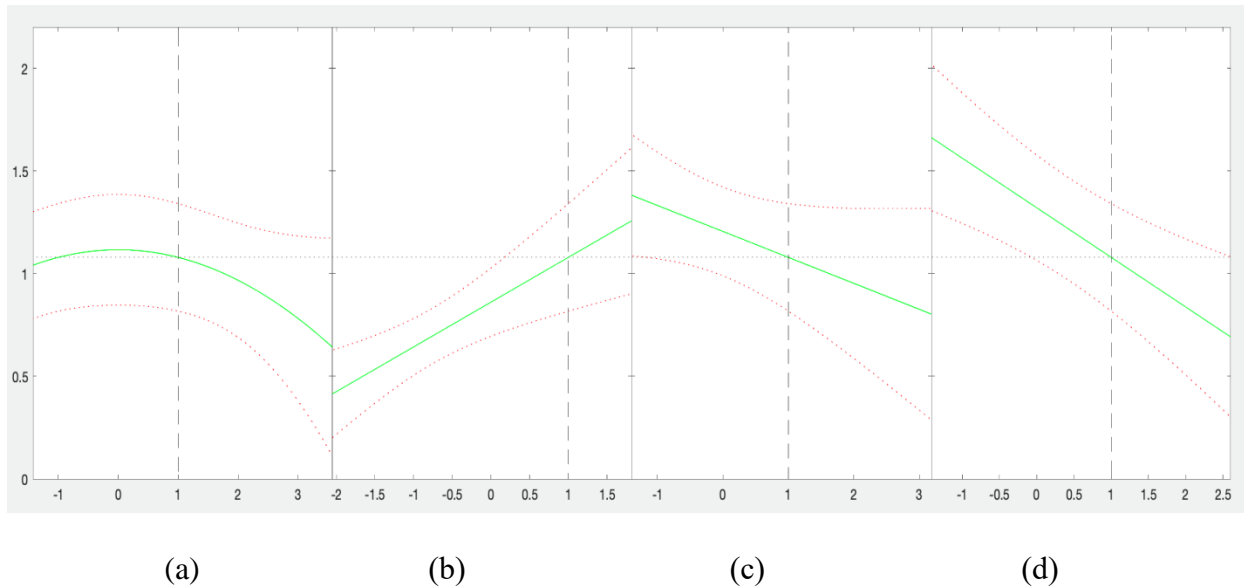
Tabla 10. *Modelo de regresión lineal 4*  $Y = \beta_0 + \beta_1 X_1^2 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$ .

Parámetros	Estimadores	SE	p-Value
$\beta_0$	1.2584	0.029113	$8.0522 \times 10^{-46}$
$\beta_1 - M.Str^2$	-0.038872	0.012639	$3.2021 \times 10^{-3}$
$\beta_2 - Tough$	0.21641	0.038792	$6.6704 \times 10^{-7}$
$\beta_3 - Micros$	-0.1562	0.042897	$5.7976 \times 10^{-4}$
$\beta_4 - C.Comp$	-0.19142	0.036789	$2.68 \times 10^{-6}$
F-Statistic p-value	-	-	$1.97 \times 10^{-25}$

Se han realizado otras pruebas con diferentes interacciones entre variables, pero ninguna de ellas mejora los resultados. La siguiente figura (*Figura 29*) muestra los efectos de cada variable predictiva sobre la variable dependiente para el *Modelo de regresión lineal 1*.

*Modelo de regresión lineal 4*  $Y = \beta_0 + \beta_1 X_1^2 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$ .





*Figura 29.* Efecto de cada variable explicativa sobre el CTOD: (a) Resistencia mecánica, (b) Tenacidad, (c) Microestructura y (d) Composición química. Se ha marcado un valor de cada variable explicativa de 1,0000 (estandarizado) y se representa como varía el CTOD (eje vertical).

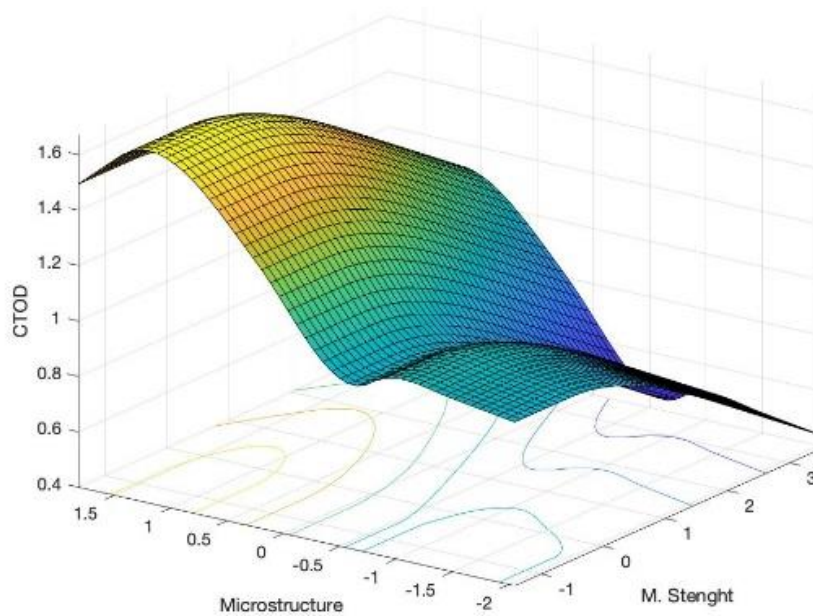
### 4.3 Splines de regresión adaptativa multivariante (MARS)

Las Splines de regresión adaptativa multivariante (MARS) es un método de modelización no paramétrico que extiende el modelo de regresión lineal incorporando interacciones y no linealidades. Es una generalización de la regresión de partición recursiva (RPR), que divide el espacio de las variables explicativas en diferentes subregiones. El método MARS genera puntos de corte para las variables.

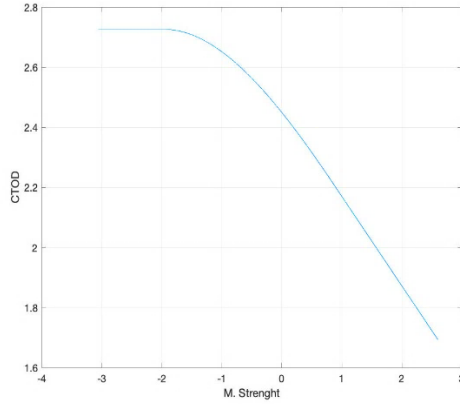
En cada región en la que se divide el espacio, se ajusta una función lineal base de una variable. El modelo final está constituido por una combinación de las funciones base generadas previamente.

Se aplicó un modelo MARS utilizando *Splines* cúbicos. Este método considera las relaciones no lineales entre la variable CTOD y las explicativas (utilizando los mencionados *Splines*

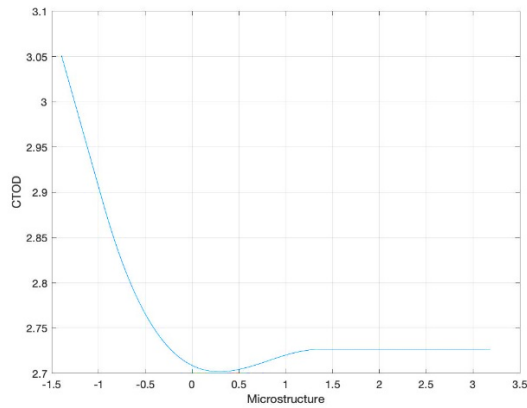
cúbicos) obteniendo un  $RSQ = 0.86$  y  $RMSE = 0.16$ . Con una muestra de entrenamiento de 50 conjuntos de datos y una muestra de prueba de 13, los resultados fueron  $RSQ = 0.84$  y  $RMSE = 0.26$ . Se puede encontrar información adicional en la *Figura 30*, donde el modelo MARS se representa para dos de las variables explicativas y dos funciones de análisis de varianza (ANOVA) (esto permite visualizar la contribución de las funciones ANOVA para los pares *CTOD-Resistencia mecánica* y *CTOD- Microestructura* en el modelo MARS).



(a)



(b)



(c)

Figura 30. (a) Diagrama del modelo de Splines de regresión adaptativa multivariante (MARS) para dos de las variables explicativas junto con sus ubicaciones de nudos (arriba) y (b) la función de análisis de varianza (ANOVA) para los pares CTOD-Resistencia mecánica y (c) CTOD-Microestructura (usando la herramienta ARESLab: Jakobsons G., ARESLab: Adaptive Regression Splines Toolbox para Matlab / Octave).

Los resultados obtenidos no mejoran los proporcionados por los modelos de regresión lineales.

#### 4.4 Modelos de regresión lineal generalizada

En la *Tabla 10* se representan los resultados de un modelo lineal generalizado (*Modelo lineal generalizado 1*) aplicado a los datos de entrenamiento. Se ha considerado una distribución gaussiana y una función de enlace identidad (=1).

Tabla 11. *Resultados del Modelo lineal generalizado 1*

Parámetros	Estimadores	SE	p-Value
$\beta_0$	1.239	0.023026	$1.7138 \times 10^{-42}$
$\beta_1 - M.Str^2.$	-0.077613	0.040691	$6.2871 \times 10^{-2}$
$\beta_2 - Tough.$	0.14999	0.035248	$1.0440 \times 10^{-4}$
$\beta_3 - Micros.$	-0.17391	0.043083	$2.0811 \times 10^{-4}$
$\beta_4 - C.Comp.$	-0.24981	0.036190	$1.4201 \times 10^{-8}$
F-Statistic p-value	-	-	$1.06 \times 10^{-23}$

Se observa que el p-valor de la variable *Resistencia mecánica* es mayor que 0,05 (el intervalo de confianza asociado a la prueba p-valor). Los intervalos de confianza de cada coeficiente son los siguientes:

$\beta_0$	(1.1927 ; 1.2854)
$\beta_1 - M.Str^2.$	(-0.1596 ; 0.0043)
$\beta_2 - Tough.$	(0.0790 ; 0.2210)
$\beta_3 - Micros.$	(-0.2607 ; -0.0871)

$$\beta_4 - C. Comp. \quad (-0.3227; -0.1769)$$


---

Parece que la variable *Resistencia mecánica* podría no ser significativa. Se construye un nuevo modelo sin esta variable (*Modelo lineal generalizado 2*). Los resultados del modelo están recogidos en la siguiente *Tabla 11*.

Tabla 12. *Resultados del Modelo lineal generalizado 2 sin la variable explicativa "Resistencia mecánica"*

Parámetros	Estimadores	SE	p-Value
$\beta_0$	1.2407	0.023659	$1.1069 \times 10^{-42}$
$\beta_2 - Tough.$	0.14034	0.035869	$2.9916 \times 10^{-4}$
$\beta_3 - Micros.$	-0.19866	0.042243	$2.3652 \times 10^{-5}$
$\beta_4 - C. Comp.$	-0.27453	0.034745	$4.1328 \times 10^{-10}$
F-Statistic p-value	-	-	$4.44 \times 10^{-24}$

Se determinan los siguientes coeficientes de ajuste:  $RSQ = 0,91$  y el  $RMSE = 0,2739$ .

#### 4.5 Árboles de regresión

Los árboles de regresión permiten construir modelos de regresión no-lineal. La metodología consiste en ir realizando una partición del espacio de las variables dependientes de manera secuencial. La siguiente figura (*Figura 31*) muestra un árbol de regresión aplicado a todos los datos de partida.

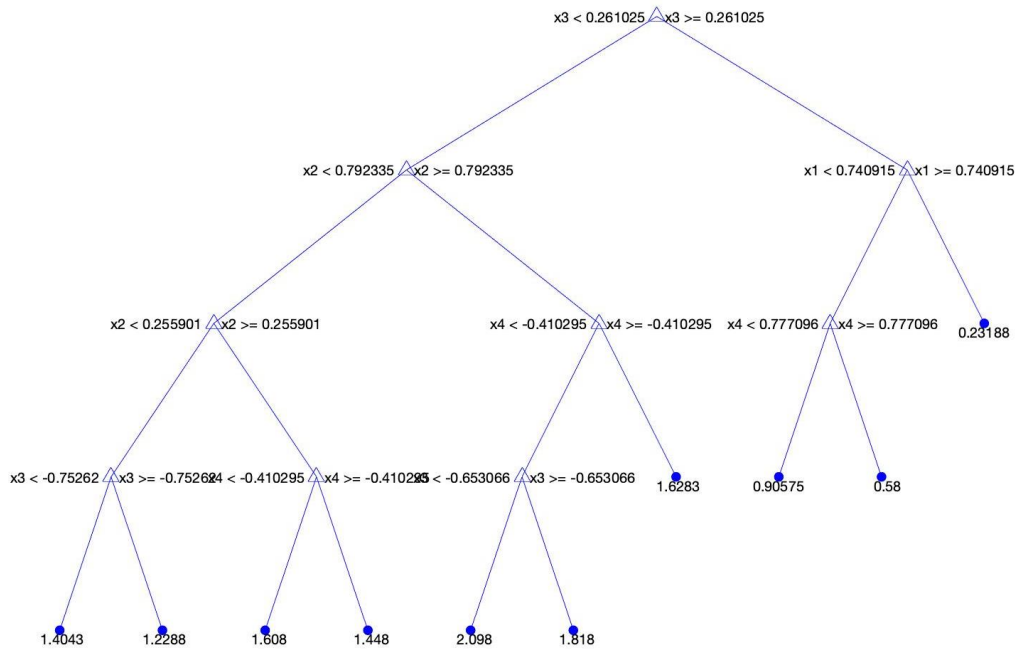


Figura 31. Árbol de regresión para el conjunto de todos los datos disponibles (después de eliminar los Outliers).

El *coste* del árbol completo, calculado mediante validación cruzada es de 0,0485. Este valor es la suma, sobre los nodos terminales, de la probabilidad estimada por el *coste* del nodo. Este último representa el error cuadrático medio de las observaciones en cada uno de los nodos.

El orden de importancia de las variables predictivas, de mayor a menor, es: *Microestructura*, *Tenacidad*, *Resistencia mecánica* y *Composición química*.

Se observa que la variable *Microestructura* es la que proporciona un valor tal que maximiza la información sobre la variable dependiente  $Y$  (CTOD). Si *Microestructura*  $< 0,26$  es *Tenacidad* la variable que proporciona más información de  $Y$  (CTOD), mientras que si *Microestructura*  $\geq 0,26$  entonces *Resistencia mecánica* es la variable con más información de la variable dependiente (es importante destacar que estos valores, por ejemplo 0.26, se refieren a la variable explicativa estandarizada y no a su valor absoluto).

Para una muestra de entrenamiento de 50 observaciones y una muestra de prueba de 13 datos, el valor del coste obtenido es 0,0629 (representación del error cuadrático medio). El orden de importancia de las variables, de mayor a menor, es: *Microestructura, Tenacidad, Composición química y Resistencia mecánica*.

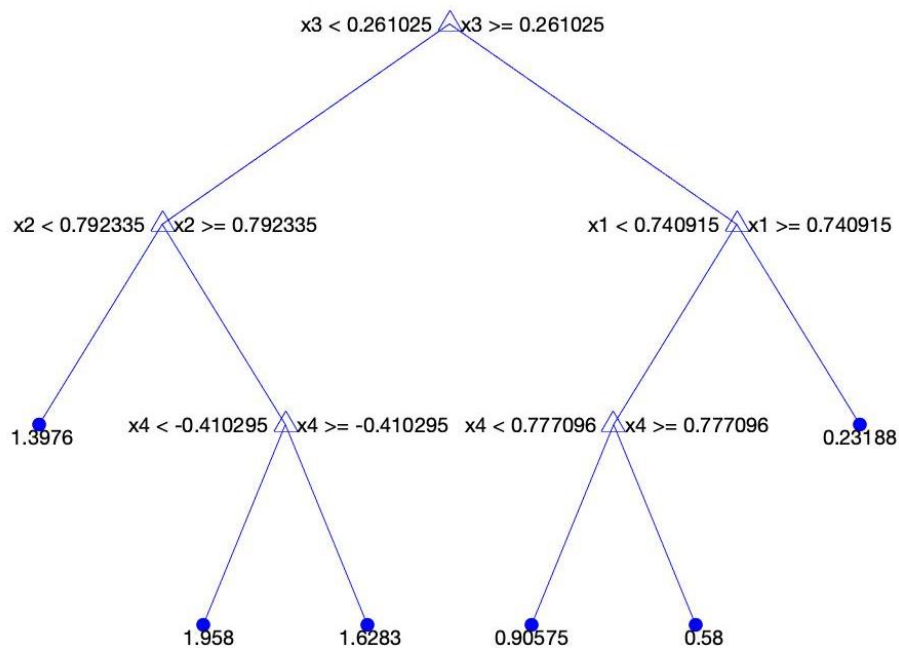
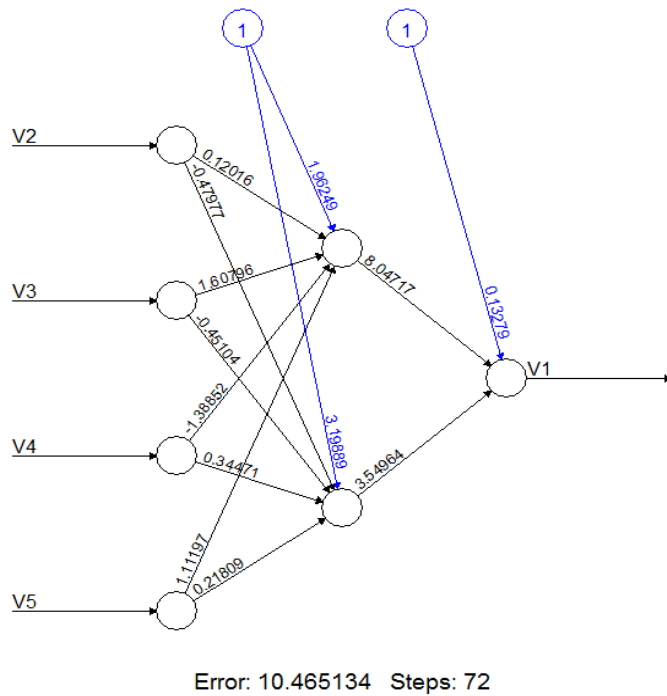


Figura 32. Árbol de regresión con una muestra de prueba de 13 observaciones

Los costes asociados a cada subárbol para la muestra de entrenamiento son entre 3 y 7 veces los de la muestra de prueba, lo que supone un mal comportamiento del modelo y por lo tanto de descarta su uso.

### 4.6 Modelo con Redes Neuronales

Se utiliza un modelo, simple, perceptrón con una capa oculta (hidden layer) denominado *Red Neuronal 1*. Se han utilizado todas las observaciones para calcular el modelo presentado en la siguiente figura (*Figura 33*). A la vista de los resultados, (error 10,465) no es posible considerar que este modelo mejore los resultados proporcionados por modelos previos de regresión lineal.



*Figura 33.* Modelo de redes neuronales perceptrón con una capa oculta



## 5 Discusión

Después de comparar los modelos previos y debido a la simplicidad y razonable precisión (y a pesar del peso no equilibrado de sus diferentes variables ( $\beta_i$ )); el modelo de regresión lineal 1 (RSQ=0.866) se propone como modelo predictivo de los valores CTOD.

El modelo estandarizado puede expresarse como:

$$CTOD = 1.2202 - 0.080323[M.Str] + 0.22424[Tough] - 0.12972[Micros] - 0.19243[C.Comp] \quad (46)$$

donde [...] representa los valores estandarizados de las correspondientes variables. El proceso de estandarización debe ser revertido para poder utilizar con comodidad el modelo con los datos directamente obtenidos de los ensayos de laboratorio:

$$\frac{(CTOD[mm]-1.0269)}{0.586} = 1.2202 - 0.080323 \frac{(M.Str[HV10]-221.09)}{41.87} + 0.22424 \frac{(Tough[J]-175.421)}{51.84} - 0.12972 \frac{(Micros[\mu m]-169.23)}{55.44} - 0.19243 \frac{(CE-0.3831)}{0.054} \quad (47)$$

donde  $CTOD[mm]$  es el valor del CTOD (Crack Tip Opening Displacement) expresado en [mm].  $M.Str[HV10]$  es el efecto de la resistencia mecánica del material representado como media aritmética de tres medidas de dureza superficial expresado en unidades HV10].

$Tough[J]$  es el valor medio de dos ensayos Charpy con probetas reducidas extraídas de la zona de interés expresado en [Julios].  $Micros[\mu m]$  es el tamaño medio de los granos metálicos cristalinos expresado en [ $\mu m$ ]. Y, finalmente,  $CE$  es el efecto de la composición química

expresada según el índice del carbono equivalente calculado según la ecuación propuesta en el AWS D1.1 (adimensional).

## 6 Resumen

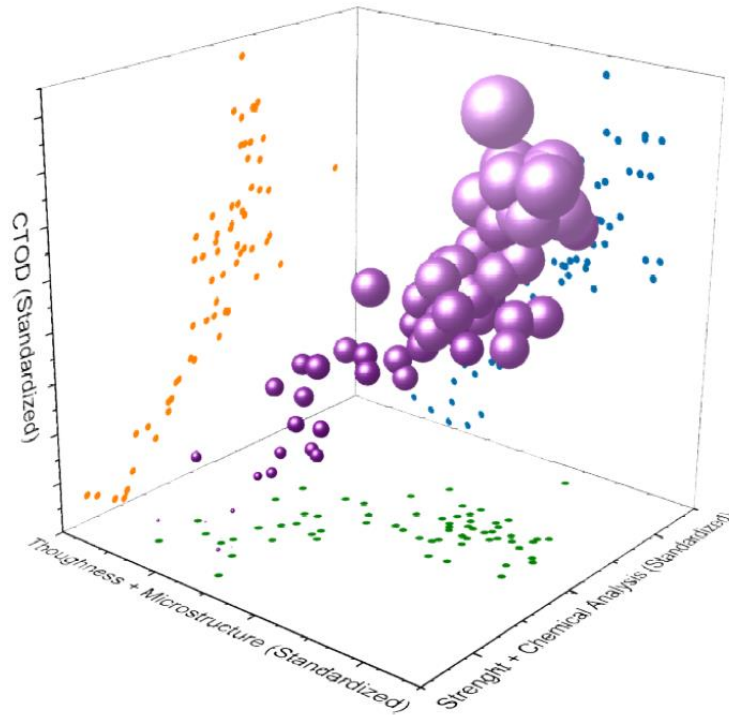
El objetivo de este estudio es construir un modelo matemático que reproduzca el efecto sobre el valor CTOD de cuatro variables características, todo ello en uniones soldadas de acero de alto límite elástico.

Con ello se podría comprender mejor la relación entre la compleja mecánica de fractura en uniones soldadas y variables más simples, de extendido uso industrial, como la resistencia mecánica, tenacidad, composición química o análisis microestructural.

Se realizan más de 3600 tomas de datos individuales en 72 probetas soldadas con parámetros de amplio uso industrial, los cuales, se procesan y evalúan con diversas técnicas matemáticas de regresión múltiple. Todas las variables parecen ser significativas y los grados de correlación son elevados.

Finalmente se selecciona un modelo que parece predecir con gran precisión el (complejo y costoso) valor CTOD a partir de las sencillas (y baratas) pruebas para determinar las cuatro variables de estudio. Se extraen diversas conclusiones técnicas al respecto.

A modo de resumen del modelo se representa en la siguiente figura una modificación de un diagrama de dispersión 3D.



*Figura 34.* Modificación de un diagrama de dispersión 3D

En el centro de la visualización y con un color llamativo, podemos encontrar los puntos que representan el valor CTOD en base a dos variables (una obtenida como una combinación de los datos de tenacidad y microestructura y otra de resistencia mecánica y composición química). El tamaño de estos puntos (representados como burbujas) es proporcional al valor de CTOD, por lo tanto, a medida que los puntos ascienden por el eje z, el tamaño aumenta. En los planos laterales se muestra la proyección de los datos. de esta forma se obtiene un diagrama de dispersión que relaciona las variables en pares. Se hace evidente la existencia de un patrón que relaciona las variables.

## 7 Conclusiones

El uso del análisis predictivo multivariante ha demostrado ser confiable. La idoneidad industrial de la metodología depende del conjunto experimental, específicamente, la disponibilidad de muestras, el número de pruebas y la elección de las variables explicativas. Estas variables elegidas están significativamente correlacionadas con el CTOD (ver, por ejemplo, el p-valor para el Modelo de regresión lineal 1). Además, existe una experiencia bien conocida dentro de la industria de fabricación Offshore que relaciona el cambio de estas variables con cambios reales durante el proceso de soldadura.

Por ejemplo, hay un amplio trasfondo sobre cómo el gas protector, la velocidad de soldadura o la secuencia de cordones de soldadura configura el tamaño de grano o la dureza de una unión soldada dada. Usando el modelo propuesto, es posible transferir este conocimiento al valor de la variable CTOD.

El modelo final es preciso y funcional, con un error estimado de  $\sim 10\%$  (dentro de los límites cubiertos por el conjunto experimental). Este error es compatible con la incertidumbre actual del ensayo CTOD. Además, el modelo muestra independencia con respecto a que subgrupo de datos se utiliza para la creación del modelo.

Se propone utilizar este modelo final de forma predictiva, utilizando el resultado de las pruebas para las variables explicativas (más barato, simple y habitual) para calcular el estimador del valor CTOD. Si el valor estimado para el CTOD es mayor que el especificado en el código de diseño, regla o estándar, se podría considerar no realizar la costosa prueba CTOD.

La utilidad del modelo ha sido probada dentro de los límites del conjunto experimental para uniones soldadas de acero de alta densidad para estructuras Offshore. Sin embargo, la influencia de otras variables no explícitamente consideradas en este trabajo no ha sido probadas y está fuera

del alcance del modelo presentado. Futuros desarrollos de este modelo podrían incluir, entre otros, la influencia de la temperatura de ensayo, las diferentes posiciones u orientaciones de la entalla o la inclusión de variables como los tratamientos térmicos posteriores a la soldadura o el tipo de fractura (frágil, dúctil-frágil o dúctil).

## 8 Bibliografía

- Akourri, O., Louah, M., Kifani, A., Gilgert, G., & Pluvinage, G. (2000). The effect of notch radius on fracture toughness  $J_{Ic}$ . *Eng. Fract. Mech.*, *65*, 491–505.
- Anderson, T. (2005). *Fracture Mechanics: Fundamentals and Applications*. FL, USA: Taylor & Francis: Boca Raton.
- Anderson, T. (2005). *Fracture Mechanics: Fundamentals and Applications, 3rd ed.*; . FL, USA: Taylor & Francis: Boca Raton.
- Anderson, T., & Osage, D. (2000). API 579: A comprehensive fitness-for-service guide. *International Journal of Pressure Vessels and Piping*, *77(14–15)*, 953–963.
- Andrews, K. (1965). Empirical formulae for calculation of some transformation temperature. *Iron Steel Inst*, 721-727.
- Antunes, F., Branco, R., Prates, P., & Borrego, L. (2017). Fatigue crack growth modeling based on CTOD for the 7050-T6 alloy. *Fatigue Fract. Eng. Mater. Struct.*, *40*, 11.
- Antunes, F., Rodrigues, S., Branco, R., & Camas, D. (2016). A numerical analysis of CTOD in constant amplitude fatigue crack growth . *Theor. Appl. Fract. Mech.*, *85*, 45–55.
- Araneda, P. (2019). *Manual de Análisis Predictivo*. Obtenido de <https://rpubs.com/paraneda/predictivo>
- ASTM E1290-08e1c. (2008). *Standard Test Method for Crack tip Opening Displacement (CTOD) Fracture Toughness Measurement; ASTM International*. PA, USA: West Conshohocken.
- Ávila, J., Lima, V., Ruchert, C., Mei, P., & Ramírez, A. (2016). Guide for Recommended Practices to Perform Crack Tip Opening Displacement Tests in High Strength Low Alloy Steels. *Soldag. Inspeção*, *21*, 290–302.

- Barsom, J., & Rolfe, M. (1970). ASTM STP 466 - Impact testing of metals: correlations between KIC and Charpy V-notch test results in the transition-temperature range. *West Conshohocken: ASTM*, 281-302.
- Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification And Regression Trees* . Boca Raton: CHAPMAN & HALL/CRC.
- Brown, J., & Hund, L. (2018). Estimating material properties under extreme conditions by using Bayesian model calibration with functional outputs. *United States*, doi:10.1111/rssc.12273.
- Cahoon, J. (1972). An improved equation relating hardness to ultimate strength. *Metallurgical Transactions*, 3, 3040.
- Castelluccio, G., & McDowell, D. (2016). Microstructure-sensitive small fatigue crack growth assessment: Effect of strain ratio, multiaxial strain state, and geometric discontinuities. *Int. J. Fatigue*, 82, 521–529.
- Castelluccio, G., Musinski, W., & McDowell. (2014). Recent developments in assessing microstructure-sensitive early stage fatigue of polycrystals. *Curr Opin Solid State Mater Sci*, 18, 180–187.
- Chen, Y., Pan, S., Zhou, M., Yi, D., Xu, D., & Xu, Y. (2013). Effects of inclusions, grain boundaries and grain orientations on the fatigue crack initiation and propagation behavior of 2524-T3 Al alloy. *Mater. Sci. Eng.*, 580, 150–158.
- Cottrell, A.H. (1958). Theory of brittle fracture in steel and similar metals. *Trans. Met. Soc. AIME*, 212, 192-203
- Courtney, T. (2000). *Mechanical behavior of materials*. Boston, U: McGraw-Hill.
- Cuadras, C. (1981). *Métodos de Análisis Multivariante*. Barcelona: Eunibar.



- Dawes, M. G. (1974). Fracture Control in High Yield Strength Weldments. *Welding Journal Research Supplement*, 53(9), pp. 396-379
- Dean, R., & Dalrymple, R. (1991). *Water wave mechanics for engineers and scientists. Advanced Series on Ocean Engineering*. Singapore: World Scientific.
- Díaz, J. (2012). *Comparación entre Árboles de Regresión CART y Regresión Lineal*. Obtenido de Universidad Nacional de Colombia:  
<http://www.bdigital.unal.edu.co/9474/1/71269839.2013.pdf>
- Dieter, G. (1988). *Mechanical metallurgy SI Metric., ed. ..*. McGraw-Hill Book Company cop: London, UK.
- Dunne, D., Tsuei, H., & Sterjovski, Z. (2004). Artificial neural networks for modeling of the impact toughness of steel. *ISIJ Int.*, 44, 1599–1607.
- Espino, C., & Martínez, X. (2017). *Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso*. Obtenido de UOC:  
<http://openaccess.uoc.edu/webapps/o2/bitstream/10609/59565/6/caresptimTFG0117mem%C3%B2ria.pdf>
- Everitt, B. (1993). *Cluster Analysis*. London, UK: Edward Arnold.
- Friedman, J. (1991). Multivariate Adaptive Regression Splines. *Ann Statist*, 19, 1–141.
- García, F., Esteban, L., Palacios, P., Cano, N., & Conde, M. (2008). Prediction of standard particleboard mechanical properties utilizing an artificial neural network and subsequent comparison with a multivariate regression model. *Investigación Agraria: Sistemas y Recursos Forestales*, 17, 10.5424/srf/2008172-01033.
- Goldstein, H. (2014). *Introduction to F-testing in Linear Regression Models; Lecture note of the Department of Statistics*. Oslo, Norway: University of Oslo.

Gordon, J., Keith, G., & Gordon, N. (2013). *Defect and Strain Tolerance of Girth Welds in High Strength Pipelines*. In *International Seminar on Welding High Strength Pipeline Steels*.

CBMM and TMS.

Hall, E. (1951). The deformation and Ageing of Mild Steel. *Proceeding Physical Society*, 64, 747-753.

Halweb. (s.f.). *Tema 3: Modelos lineales generalizados*. Obtenido de

<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/Categor/Tema3Cate.pdf>

Haque, M., & Sudhakar, K. (2002). ANN back-propagation prediction model for fracture toughness in microalloy steel. *Int. J. Fatigue*, 24, 1003–1010.

Hunt, R., & McCartney, L. (1979). A new approach to Weibull's statistical theory of brittle fracture. *International Journal of Fracture*, 15(4), 365–375.

Ingham, T. GR., Egan, D. Elliott & Harrison, TC. (1971). The effect of Geometry on the interpretation of COD test data, *Inst. Mech. Engng*. C54/71:200

Ingenieria Matematica. (2017). *Redes Neuronales Artificiales ANN*. Obtenido de

<https://mathengineering.blogspot.com/2017/01/redes-neuronales-artificiales-ann.html>

International Energy Outlook . (2019). *Energy Information Administration*. Obtenido de

<https://www.eia.gov/outlooks/ieo/pdf/ieo2019.pdf>

Izenman, A. (2008). *Modern Multivariate Statistical Techniques*. New York: Springer.

Janssen, M., Zuidema, J., & Wanhill, R. (2004). *Fracture Mechanics*. New York, NY, USA: Spon Press.

Kawabata, T., Tagawa, T., Sakimoto, T., Kayamori, Y., Ohata, M., Yamashita, Y., y otros. (2016).

Proposal for a new CTOD calculation formula . *Eng. Fract. Mech.*, 159, 16–34.

- Kayamori, Y., Inoue, T., & Tagawa, T. (2008). Transformation of BS7448-CTOD to ASTM E1290- CTOD. *ASME Pressure Vessels and Piping Division Conference, 1*, 1–8.
- Khor, W. e. (2016). Measurement and prediction of CTOD in austenitic stainless steel. *Fatigue & Fracture of Engineering Materials & Structures*, 39, 1433–1442.
- Khor, W., Moore, P. P., Haslett, M., & Brown, C. (2016). Measurement and prediction of CTOD in austenitic stainless steel. *Fatigue Fract. Eng. Mater. Struct.*, 39, 1433–1442.
- Krupp, U., Düber, O., Christ, H., Künkler, B., Köster, P., & Fritzen, C. (2007). Propagation mechanisms of microstructurally short cracks--Factors governing the transition from short- to long-crack behavior. *ater Sci Eng A*, 462, 174–177.
- Lewis, P., & Stevens, J. (1991). No linear modeling of time series using multivariate adaptive regression splines (MARS). *J Am Stat Assoc.*, 86, 1–36.
- Lin, I.H., T.L. Anderson, R. De Witt & M.G. Dawes (1982). Displacements and rotational factors in single edge notched bend specimens. *Int. J. Fracture*, 20, R1-7
- Marín, J. (2014). *Apuntes de: Análisis de Datos Categóricos* . Obtenido de <http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/Categor/Categ.html>
- Maropoulos, S., Ridley, N., Kechagias, J., & Karagiannis, S. (2004). Fracture toughness evaluation of a H.S.L.A. steel. *Eng. Fract. Mech.*, 71, 1695–1704.
- Morison, J., O'Brien, M., Johnson, J., & Schaaf, S. (1950). The force exerted by surface waves on piles, Petroleum Transactions. *American Institute of Mining Engineers*, 189, 149-154.
- Moskovic, R. (1993). Statistical analysis of censored fracture toughness data in the ductile to brittle transition temperature region. *Engineering Fracture Mechanics*, 44(1), 21–41.
- Newman, J., James, M., & Zerbst, U. (2003). A review of the CTOA/CTOD fracture criterion. *Eng. Fract. Mech.*, 70, 371–385.

- Offshore Standard Det Norske Veritas DNV-OS-J101. (2007). *Design of offshore wind turbine structures*. Obtenido de:  
[https://rules.dnvgl.com/servicedocuments/\\_ASPscripts/downloadPDF.asp?url=http://rules.dnvgl.com/docs/pdf/DNV/codes/docs/2009-11/Os-J101.pdf](https://rules.dnvgl.com/servicedocuments/_ASPscripts/downloadPDF.asp?url=http://rules.dnvgl.com/docs/pdf/DNV/codes/docs/2009-11/Os-J101.pdf)
- Pavlina, E., & Vantyne, C. (2008). Correlation of Yield Strength and Tensile Strength with Hardness for Steels. *Journal of Materials Engineering and Performance*, 17, 888-893.
- Pegues, J., Roach, M., & Shamsaei, N. (2017). Influence of microstructure on fatigue crack nucleation and microstructurally short crack growth of an austenitic stainless steel. *Mater. Sci. Eng.*, 707, 657–667.
- Petch, N. (1953). The cleavage strength of polycrystals. *Journal of the Iron and Steel Institute*, 25-30.
- Rao, C., Toutenburg, H., Shalabh, & Heumann, C. (2008). *Linear Models and Generalizations*. Germany : Springer: Berlin/Heidelberg.
- Rencher, A. (2001). *Methods of Multivariate Analysis*. NJ, USA: John Wiley & Sons: Hoboken.
- Rice, J. (1968). A path independent integral and the approximate analysis of strain concentration by notches and cracks. *J. Appl. Mech.*, 35, 379–386.
- Rodrigues, P., Wong, W., & Rogerson, J. (1980). *Weld Defect Distributions in Offshore Platforms and Their Relevance to Reliability Studies, Quality Control and In-Service Inspection*. TX, USA: In Proceedings of the Offshore Technology Conference, Houston.
- Rodríguez, R., Cabrera, J., & Benito, J. (2011). Estudio de la relación Hall-Petch en aceros (0,6% C) submicrométricos. *Ingeniería e investigación*, 31(3), 112-120.
- Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning representations by back-propagating errores. *Nature*, 323, 533–536 .

- Sailors, S., & Corten, H. (1972). Relations between material fracture toughness using fractures mechanics and transition temperature test. *Fracture toughness*. 164-191.
- Salas, J., & Portillo, M. (2008). Mahalanobis y las aplicaciones de su distancia estadística. *CULCyT: Cultura Científica y Tecnológica*, 27, 13-20.
- Salford. (2001). *Systems Multivariate Adaptive Regression Splines (MARS): user guide*. Chapter MARS Basics - Smoothing, splines and knot selection. 9-34.
- Sarpkaya, T., & Isaacson, M. (1981). *Mechanics of wave forces on offshore structures*. New York: Van Nostrand Reinhold.
- Sarzosa, D., Souza, R., & Ruggieri, C. (2015). Relaciones J- CTOD en muestras de fractura SE (T) fijadas incluyendo análisis estacionario y de crecimiento en 3-D. *Ingeniería Mecánica de Fractura*, 147, 331-354.
- Shen, G et al. (2004). *Fracture Toughness Testing of Pipeline Girth Welds*. In *International Pipeline Conference*. Canada: Minister of Natural Resources.
- Smith, Y., Coldren, A., & Cryderman, R. (1972). Mn-Mo-Nb acicular ferrite steels with high strength and toughness . *Toward Improved Ductility and Toughness*, 119-142.
- Tagawa, T. e. (2010). Difference between ASTM E1290 and BS 7448 CTOD Estimation Procedures. *Welding in the World*, 54(7-8), 182-188.
- Tagawa, T. e. (2014). Experimental measurements of deformed crack tips in different yield-to-tensile ratio steels. *Engineering Fracture Mechanics*, 128, 157-170.
- Vanegas, J., & Vásquez, F. (2017). Multivariate Adaptive Regression Splines (MARS), una alternativa para el análisis de series de tiempo. *Gac. Sanit*, 235-237.
- Vanegas, J., & Vásquez, F. (2017). Multivariate Adaptive Regression Splines (MARS), una alternativa para el análisis de series de tiempo. *Gac Sanit*, 31(3), 235-237.

- Wang, C.; Wang, X.; Kang, J.; Yuan, G.; Wang, G. (2019). Effect of Thermomechanical Treatment on Acicular Ferrite Formation in Ti–Ca Deoxidized Low Carbon Steel. *Metals*, 9, 296.
- Wei, L., & Pisarski, H. (2007). FEA investigations into the effects of geometry and tensile properties on J and CTOD in standard fracture specimens. En ESIA9, *International Conference on Engineering Structural Integrity Assessment*. Beijing, China: EMAS.
- Wei, L., Pan, Q., Huang, H., Feng, L., & Wang, Y. (2014). Influence of grain structure and crystallographic orientation on fatigue crack propagation behavior of 7050 alloy thick plate. *Int. J. Fatigue*, 66, 55–64.
- Wells, A. (1963). Application of fracture mechanics at and beyond general yielding. *Br. Weld. J.*, 10, 563–570.
- Yang, Y., Mahfouf, M., & Panoutsos, G. (2012). Probabilistic Characterization of Model Error Using Gaussian Mixture Model—with Application to Charpy Impact Energy Prediction for Alloy Steel. *Control Eng. Pract.*, 20, 82–92.
- Yuan, H., Zhang, W., Castelluccio, G., Kim, J., & Liu, Y. (2018). Microstructure-sensitive estimation of small fatigue crack growth in bridge steel welds. *Int. J. Fatigue*, 112, 183–197.
- Yurioka, N. (2001). Physical Metallurgy of Steel Weldability. *ISIJ Int*, 41, 566–570.
- Zhang, H., & Singer, B. (2010). *Recursive Partitioning and Applications*. Springer: New York.
- Zhu, X., & Joyce, A. (2012). Review of fracture toughness (G, K, J, CTOD, CTOA) testing and standardization. *Eng. Fract. Mech.*, 85, 1–46.

## 9 Anexos

**Anexo 1.** Certificado de publicación del artículo

**Anexo 2.** Carta de aceptación para la presentación de resultados parciales en la IX Conferencia Internacional de Aprovechamiento de Recursos Minerales “*CINAREM 2017*”

**Anexo 3.** Referencias a estándares de ensayo

**Anexo 4.** Artículo publicado en la Revista *ENERGIES* en su número especial *BIG DATA AND MACHINE LEARNING*

Anexo 1



**CERTIFICATE of PUBLICATION**



Certificate of publication for the article titled:

Multivariate Analysis to Relate CTOD Values with Material Properties in Steel Welded Joints for the Offshore Wind Power Industry

Authored by:

Álvaro Presno Vélez; Antonio Bernardo Sánchez; Marta Menéndez Fernández; Zulima Fernández Muñiz

Published in:

*Energies* 2019, Volume 12, Issue 20, 4001



## Anexo 2



IX Conferencia Internacional de Aprovechamiento de Recursos Minerales  
CINAREM'2017

Moa, 25 de septiembre de 2017  
"Año 59 de la Revolución"

Dr. Antonio Bernardo Sánchez  
Ing. Álvaro Presno Vélez  
Dra. Marta Menéndez Fernández  
Dra. Zulima Fernández Muñiz  
Dr. Roberto López González

Ref.- aceptación de trabajo para presentar en CINAREM'2017

En nombre del Comité Organizador de la IX Conferencia Internacional de Aprovechamiento de Recursos Minerales "CINAREM'2017" a celebrarse en el Instituto Superior Minero Metalúrgico de Moa durante los días del 14 al 16 de noviembre del presente año me dirijo a ustedes para comunicarles que el trabajo titulado **MODELO MATEMÁTICO MULTIVARIANTE QUE RELACIONA VALORES CTOD CON PROPIEDADES MECÁNICAS, QUÍMICAS Y MICROSTRUCTURALES EN SOLDADURAS DE ACERO PARA LA INDUSTRIA OFFSHORE** ha sido aprobado por el Comité Científico para su presentación en el referido evento. La casa de altos estudios de la comunidad minera de Moa se alista para acoger a delegados de diferentes nacionalidades y se sentirá orgullosa de contar con su presencia; la que sin dudas prestigiará la ocasión.

A la espera de su confirmación, les saluda cordialmente,

Josefa Mestre Lamoru  
Directora de Relaciones Internacionales  
Secretaria Ejecutiva Comité Organizador  
"CINAREM'2017"



**Anexo 3**

American Bureau of Shipping (2011) *Design Standards for Offshore Wind Farms*;  
American Bureau of Shipping: Huston Texas, USA

American Welding Society (2010) *D1.1/D1.1M. Structural Welding Code Steel*;  
American Welding Society: Miami, FL, USA

ASM International (1996) *ASM Handbook; ASM International*: Materials Park, OH,  
USA, Volume 19.

ASME BPVC IX (2017) *Boiler Pressure Vessel Code (BPVC) Welding, Brazing, and  
Fusing Qualifications*; The American Society of Mechanical Engineers (ASME): New  
York, NY, USA

ASTM E112 (2013) *Standard Test Methods for Determining Average Grain Size*; ASTM  
International: West Conshohocken, PA, USA

ASTM E1290-08e1c (2008) *Standard Test Method for Crack tip Opening Displacement  
(CTOD) Fracture Toughness Measurement*; ASTM International: West Conshohocken,  
PA, USA

ASTM E23 (2018) *Standard Test Methods for Notched Bar Impact Testing of Metallic  
Materials*; ASTM International; West Conshohocken, PA, USA

ASTM E3 11 (2018) *Standard Guide for Preparation of Metallographic Specimens*;  
ASTM International: West Conshohocken, PA, USA

ASTM E407 07 (2015) *Standard Practice for Micro-Etching Metals and Alloys*; ASTM International: West Conshohocken, PA, USA

ASTM E92 (2017) *Standard Test Methods for Vickers Hardness and Knoop Hardness of Metallic Materials*; ASTM International: West Conshohocken, PA, USA

EN 10225 (2009) *Weldable Structural Steels for Fixed Offshore Structures*; European Committee for Standardization: Brussels, Belgium

EN ISO 10025 (2006) *Hot Rolled Products of Structural Steels*; European Committee for Standardization: Brussels, Belgium

EN ISO 17637 (2017) *Non-Destructive Testing of Welds—Visual Testing of Fusion-Welded Joints*; European Committee for Standardization: Brussels, Belgium

EN ISO 17638 (2017) *Non-Destructive Testing of Welds—Magnetic Particle Testing*; European Committee for Standardization: Brussels, Belgium

EN ISO 17640 (2011) *Non-Destructive Testing of Welds—Ultrasonic Testing—Techniques, Testing Levels, and Assessment*; European Committee for Standardization: Brussels, Belgium.

## Anexo 4



Article

## Multivariate Analysis to Relate CTOD Values with Material Properties in Steel Welded Joints for the Offshore Wind Power Industry

Álvaro Presno Vélez <sup>1</sup>, Antonio Bernardo Sánchez <sup>1,\*</sup>, Marta Menéndez Fernández <sup>1</sup>  
and Zulima Fernández Muñiz <sup>2</sup>

<sup>1</sup> Department of Mining Technology, Topography and Structures, University of León, 24071 León, Spain; a.presno.velez@gmail.com (A.P.V.); marta.menendez@unileon.es (M.M.F.)

<sup>2</sup> Department of Mathematics, University of Oviedo, 33007 Oviedo, Spain; [zulima@uniovi.es](mailto:zulima@uniovi.es)

\* Correspondence: [abers@unileon.es](mailto:abers@unileon.es); Tel.: +34-987-293-554

Received: 28 September 2019; Accepted: 18 October 2019; Published: 21 October 2019

**Abstract:** The increasingly mechanical requirements of offshore structures have established the relevance of fracture mechanics-based quality control in welded joints. For this purpose, crack tip opening displacement (CTOD) at a given distance from the crack tip has been considered one of the most suited parameters for modeling and control of crack growth, and it is broadly used at the industrial level. We have modeled, through multivariate analysis techniques, the relationships among CTOD values and other material properties (such as hardness, chemical composition, toughness, and microstructural morphology) in high-thickness offshore steel welded joints. In order to create this model, hundreds of tests were done on 72 real samples, which were welded with a wide range of real industrial parameters. The obtained results were processed and evaluated with different multivariate techniques, and we established the significance of all the chosen explanatory variables and the good predictive capability of the CTOD tests within the limits of the experimental variation. By establishing the use of this model, significant savings can be achieved in the manufacturing of wind generators, as CTOD tests are more expensive and complex than the proposed alternatives. Additionally, this model allows for some technical conclusions.

**Keywords:** crack tip opening displacement; steels; welded or bonded joints; multivariate regression model; marine structures

---

## 1. Introduction

As the burgeoning offshore wind power industry grows, so too do the technical demands on the metal frames and primary structures that sustain them. These structures are under enormous dynamic stresses due to the effects of their moving parts, wind, currents, tides, and waves. Within this sector, the quality control of the welded joints of these structures is of the utmost importance, considering that welding defects are widely considered as potential spots for structural failure initiation [1]. The study based on fracture mechanics of parameters such as CTOD, in the context of crack nucleation and fatigue crack growth, has become essential for manufacturers, designers, classification societies, and inspectors. The fatigue life calculation occupies a prominent place in codes, standards, and rules [2–4]. Such fatigue analysis is based on “rule-based” methods or direct calculation based on Stress-Cycles data models, determined by fatigue testing of the considered welded details and linear damage hypothesis.

As this approach is rarely possible (due to the full fatigue test required for the welded details), the fatigue analysis may alternatively be based on fracture mechanics. The classification societies’ crack growth models use the classic formulation of the Paris–Erdogan law, with developments for the classical plastic hinge models (firstly developed by the British Standards Institution and published in 1979). According to the vast work of Zhu and Joyce [5], the stress intensity factor  $K$  [6], the crack tip opening displacement (CTOD) [7], the J-integral [8], and the crack tip opening angle (CTOA) (developed for thin-walled materials) are the most relevant parameters used in fracture mechanics. Out of these various parameters of the interaction of the materials with the formation and propagation of cracks or defects, the critical crack tip opening displacement (CTOD) at a given distance from the crack tip is the most suited for modeling stable crack growth and instability during the fracture process [9]. Currently, the tests are carried out by discarding the plastic hinge model and adopting the J-conversion, using recognized standards such as the (British Standard) BS-7910, (American Petroleum Institute) API-579, and (American Society for Testing and Materials) ASTM E1290.

CTOD testing requires the preparation of a notch with a specific geometry that promotes the nucleation of a stable and uniform crack in a delimited area [10]. The crack grows under the action of dynamic mechanical forces that are generally transmitted with huge oleo-hydraulic equipment and controlled by precision extensometers. The uncertainty of the test methods, as well as the sensitivity to any internal defect, make it necessary to carry out several of these tests to guarantee representative values.

The CTOD tests are expensive, as they require significant investments in testing machinery, software, expertise, and outsourcing of services [11]. The destruction of large quantities of ad hoc welded material is also required (ASTM E1290-08e1c (2008) [12]). Additionally, deadlines offered by the testing laboratories exceed the average for other quality control tests in welded unions. Considering the case of welded joints, in addition to the properties of the base material, dozens of other variables related to the welding process could affect the features of the final welded material. Therefore, if the CTOD test result does not fulfil the requirements, it is very difficult for technicians to infer which changes in the variables could lead to an improvement of the CTOD results.

The aim of the present work is to evaluate the possibility of using multivariate mathematical models to correlate the CTOD parameter with other test results that are simpler and cheaper to measure, and also well known by the parties involved.

## 2. Selection of Input Variables and Experimental Phase

The multivariate analysis consists of a series of appropriate statistical methods (such as multiple regression, logistic regression, analysis of variance (ANOVA), or cluster analysis, to name a few) used when numerous observations are performed on the same object in one or several samples. Those methods allow the creation of formal hypothesis tests when given a structure of input–output data. Expressing a variable as a function of a set of underlying intercorrelated variables is among the possible hypotheses [13].

### 2.1. Selection of Input Variables

The selection of these so-called explanatory variables was done considering the industrial approach of this research work. Among the numerous variables with proven effects on the material properties (see Table 1 for a non-exhaustive selection proposed by Dunne et al. [14] and Haque and Sudhakar [15]), the following ones were selected due to their widespread use in the industry, relatively cheap measurement, and possibility to be determined in modest-quality control laboratories. Also, the chosen variables are part of the testing process required by the design codes, rules, and standards for the design, qualification, and control of welded joints. Therefore, these values are usually available (or easy to gather), there are clear acceptance criteria, and their effects on the CTOD and on welded joints are widely recognized.

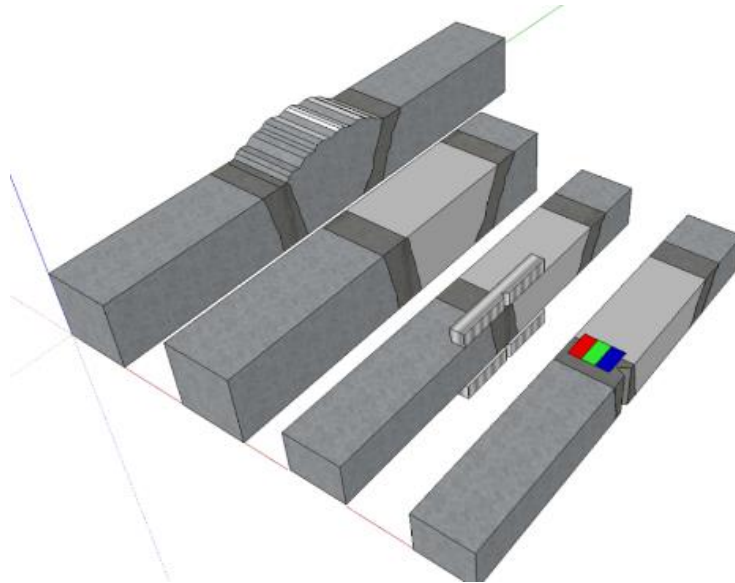
**Table 1.** Non-exhaustive selection of variables with proven influence on material properties.

Variables	
Carbon (wt %)	Plate thickness (mm)
Manganese (wt %)	Post Welding Heat Treatment (PWHT) holding time
Silicon (wt %)	PWH cooling rate/method
Sulphur (wt %)	Test piece orientation
Phosphorus (wt %)	Test temperature
Aluminum (wt %)	Yield Strength (MPa)
Boron (wt %)	Ultimate Tensile Strength (MPa)
Molybdenum (wt %)	Charpy toughness (J)
Oxygen (wt %)	Grain boundaries and orientation
Nitrogen (wt %)	Hardness
V% fraction of reaustenized region	Grain boundary ferrite
V% fraction of double-reheated zone	Intragranular polygonal ferrite
Grain refined subzone	Grain coarsened subzone
Non-metallic inclusions	Mean 3D diameter of inclusions

### 2.1.1. Microstructure

The microstructure of the material in the area in which the CTOD value is to be determined will be considered one of the input variables. Some authors [16–23] have studied the relation between microstructure characteristics and fracture mechanics properties and supports, and the influence of grain size, angle of grain boundaries, orientation, and inclusions on the nucleation and propagation of cracks. The average size of the metallic grains in the area of interest was determined according to ASTM E112 (2013) [24] (determined by optical microscopy) to represent this variable. The specimens were polished and prepared according to the recommendation of E3-11 (Guide for Preparation of Metallographic Specimens) [25] for Al<sub>2</sub>O<sub>3</sub> abrasive (1200 American National Standards Institute grit number), with rotation and etching reagent no. 77 (E407-07 Standard Practice for Microteaching Metals and Alloys) [26].

The limitation in obtaining samples with different surface orientations (see Figure 1) appropriate for eventual non-equiaxed grain shapes was corrected with the implementation of an arbitrary multiplication factor, depending on the grain contour. Any possible heterogeneity in the area of interest is expected to be statistically covered by the experimental design. Having considered the industrial approach, other well-known techniques that require specific equipment, such as scanning electron microscope (SEM), were not used. Also, as failure types are not considered as study variables, the critical grain size for brittle fracture was not considered.



**Figure 1.** Sampling position. Color zones mark targeted areas for microstructural, hardness, and chemical analysis (red, green and blue).

### 2.1.2. Chemical Composition

The chemical composition of the material is a well-known factor that exerts influence on the mechanical properties [27–29].

Samples were analyzed by optical emission spectrometry and X-ray diffraction using a Niton® XL2 analyzer and a Spectromax metal analyzer. Results were statistically processed to offer the



best-weighted average estimator considering the different uncertainties of the testing method and for the following elements: C, Mn, Si, Cr, Ni, Mo, and V. Both the test procedure and the uncertainty calculation used were approved by the testing laboratory. For the implementation of the chemical composition into the mathematical model, we considered the influence of the different elements using the carbon equivalent (CE) index, expressed in Equation (1). Among the numerous CE formulae available in the bibliography, we chose American Welding Society (AWS) D1.1 [30], which was cited in [29] and is also known as the International Institute of Welding (IIW) carbon equivalent.

This expression was selected considering its precision for mechanical and microstructural properties [27]:

$$CE_{\text{index}} = C + \frac{(Mn + Si)}{6} + \frac{(Cr + Mo + V)}{5} + \frac{(Ni + Cu)}{15} \quad (4)$$

where all values involved represent the mass percentage composition [w/w%]. Therefore, the result is a non-dimensional continuous variable.

### 2.1.3. Mechanical Strength

The mechanical resistance plays a fundamental role, and forms a constitutive part, in fracture mechanics [31]. Also, the determination and control of its value is a fundamental part of the quality control of the material properties (for structural materials). Tensile test results were discarded due to the impossibility to take measurements exclusively in the small area of interest, as all the subsized specimens proposed by the standards exceed the capability of the testing machine (too small) or destroy valuable testing material (too big). Nevertheless, according to numerous publications (e.g., ASM Handbook for carbon steel), there is a consistent and almost linear relation between ultimate tensile strength (UTS) and hardness. Therefore, hardness measurements according to ASTM E92 (Hardness Vickers 10) (2017) [32] were taken from the samples to estimate the mechanical resistance of the material. Standardized Vickers indenters (Class B) were used with a load of 98.7 N (HV 10) and an optical indentation measurement. The average value of a set of three indentations (considering 2 mm of space between tests) were examined for each sample.

### 2.1.4. Toughness

Previous studies [15,16,33,34] support the relation between impact testing results (measured as Charpy V-notch (CVN) energy values) and fracture toughness. Some correlations have been adopted by the standards ASME Boiler and Pressure Vessel Code (BPVC) XI (2017) [35] and API 579 [3].

CVN tests, according to ASTM E23 (2018) [36], were performed on the samples. Subsize Charpy simple-beam V-notch impact test specimens were used (2.5 mm, according to Figure A3.1 from ASTM E-23), with the notch aligned with the future CTOD sample notch [10]. All tests were performed at room temperature (between 20 and 25 °C) with a 300 J pendulum device. Three specimens (instead of two) were used for each toughness characterization to ensure

representative values (see Figure 1), due to the sample size limitation. Measurement of lateral expansion or the fracture region size was not considered.

## 2.2. *Experimental Design*

For a multivariate statistical study (with a suitable uncertainty), it is required to reach a determined critical mass of input data. This number is undetermined, and it will be verified after the modeling [17]. In addition, a wide range that covers the industrial interest is required for the explanatory variables.

It is expected that the heterogeneities on the physical properties of the welded joints and the uncontrolled variables were arbitrarily distributed among the observations, according to the random principle [37]. Nevertheless, the sampling and test position was designed to minimize the effect of these heterogeneities (see Figure 1) by reducing, as much as possible, the area from where the results are obtained.

The first approximation of the complete number of tests was estimated considering the guide of factorial experimental design, computed for four variables, each with two levels, and one replication. A total of 72 complete sets of data were obtained through testing campaigns.

The welded coupons (from where the specimens were extracted for testing) were kindly transferred from manufacturing companies. This guarantees the reproducibility with respect to real welding designs and manufacturing processes, and also the applicability of the ranges used. Nevertheless, it also limits the number of available samples, and as the authors cannot control the range of variability of the study variables, experimental designs with surface analysis or complex factorial designs cannot be used. These limitations of the used experimental set may affect the accuracy of future models (by not gathering the critical amount of data for multivariate models) and prevent the use of more explanatory variables.

## 2.3. *Samples*

All the samples were extracted from 36 welded coupons of at least 400 mm length in the direction of the weld. Those coupons were welded for real Welding Procedure Qualification records following real Welding Procedure Specifications, then kindly transferred from manufacturing companies for this project. The thickness of the coupons varied between 20–75 mm and were considered representative of offshore manufacturing. *K* and *V*-bevels were used, and the base materials and consumables were standard within the manufacturing sector. Several different consumables and four base materials classified according to EN 10225 [38] (low-alloy steels S355 G5 + M, S355 G10 + M, S420 G2 + M, and S460G2 + M together) were used, together with two structural steels from EN 10025:2010 [39] (low-carbon steels S275 J2 and S355 K2). The following table (Table 2) summarizes the range of different relevant variables during the welding process that may have an influence on the properties of the welding coupons. These ranges are considered as representative of the structural welding processes of the offshore wind power industry.

**Table 2.** The range for different variables of the test coupons.

Variables	Min.	Max.	Variables	Min.	Max.
Wire diameter	1.2 mm	2.8 mm	Material base	S275	S460
Intensity	80 A	230 A	No. of welding processes	1	2
Polarity	DC	AC	Voltage	10 V	30 V
Speed	40 mm/min	240 mm/min	Use of backup	No	Yes
Heat input	0.6 KJ/mm	4.2 KJ/mm	Welding thickness	20 mm	75 mm
Pre-heating	No	150 °C	W. position	PA	PF
Gas flux	10 l/min	25 l/min	Bevel angle	35°	90°

Abbreviations: minimum = Min.; maximum = Max.

All the welded coupons were subjected to extend non-destructive tests, according to EN 17637 [40], EN 17638 [41], and EN ISO 17640 [42]. A total of 14 small indications were found, and consequently the zone was marked and discarded for destructive tests.

#### 2.4. CTOD Test

All 72 CTOD tests were done according to ASTM E1290-08e1c (2008) [12] with standard specimens (single-edge notched bend SE(B) specimen with square  $B \times B$  cross-section) and the recommended notch [43]. The apparatus used was an oleo-hydraulic dynamic machine (model UFIB-200E-MD5W) configured for a 3-point bending setup and using clip-gauges as the crack growth measuring system. The testing temperature was in the range of 20–25 °C. As Figure 1 shows, the notch was aligned 1 mm from the fusion line.

The chosen testing method, ASTM E1290-08e1, calculate the CTOD value with the following expression:

$$\delta = \frac{1}{m\sigma_Y} \left[ \frac{K^2(1-\nu^2)}{E} + \frac{\eta_{CMOD} A_{CMOD}^{pl}}{B(W-a_0) \left\{ 1 + Z / (0.8a_0 + 0.2W) \right\}} \right] \quad (2)$$

where  $Z$  is the distance of the front face of the SE(B) specimens to the knife-edge measurement point,  $A_{CMOD}^{pl}$  is the plastic area under load from the plastic  $CMOD$  curve, and the expression of  $m$  is:

$$m = A_0 - A_1 \left( \frac{\sigma_{YS}}{\sigma_{ts}} \right) + A_2 \left( \frac{\sigma_{YS}}{\sigma_{ts}} \right)^2 - A_3 \left( \frac{\sigma_{YS}}{\sigma_{ts}} \right)^3 \quad (3)$$

where

$$A_0 = 3.18 - 0.22 \left(\frac{a_0}{W}\right), A_1 = 4.32 - 2.23 \left(\frac{a_0}{W}\right), A_2 = 4.44 - 2.29 \left(\frac{a_0}{W}\right), A_4 = 2.05 - 1.06 \left(\frac{a_0}{W}\right) \quad (4)$$

and

$$\eta_{CMOD} = 3.667 - 2.199 \left(\frac{a_0}{W}\right) + 0.437 \left(\frac{a_0}{W}\right)^2, \quad (5)$$

Alternatives calculations, formulas, and predictions were studied by [33,44–48].

All the tests were performed in the private laboratory testing facilities of the TAM group (accreditation no. 808/LE1532).

## 2.5. Results

The data obtained were processed according to the respective test procedures. Finally, for each of the 72 test samples, the results were collected for the explanatory and objective variables. In the Table 3 the results of the testing process are summarized and expressed as the minimum (Min.) and maximum (Max.), giving the range, the average value (Avg.), standard deviation (SD), and coefficient of variation (CV).

**Table 3.** Experimental phase results summary.

Para m.	CTOD [mm]	Mechanical Strength [HV10]	Toughness [J]	Microstructure [μm. Correction]	Chemical Composition [CE]
Min	0.10	165	76	101	0.22
.					
Max	2.45	375	278	354	0.53
.					
Avg	1.24	224.4	183.4	177.0	0.39
.					
SD	0.59	41.87	51.84	55.44	0.05
CV	47.3	18.7	28.3	31.3	14.1
[%]					

Abbreviations: parameters = Param.; minimum = Min.; maximum = Max.; average value = Avg.; standard deviation = SD; coefficient of variation = CV; crack tip opening displacement = CTOD.

### 3. Modeling

We observed a set of  $K$  variables  $X_1, X_2, \dots, X_K$  in a set of  $n$  elements of a population and wanted to summarize the values of the variables and describe their dependency structure. Each of these  $K$  variables is called a scalar or univariate variable and the set of these  $K$  variables form a vector or multivariate variable. All these values can be represented in a matrix,  $X$ , of dimensions  $n \times p$ , called a data matrix, where each row represents the values of the  $K$  variables over the individual  $i$ , and each column represents the corresponding scalar variable measured in the  $n$  elements of the population. In the element  $x_{ij}$ ,  $i$  denotes the individual and  $j$  is the variable.

Next, we proceed to the multivariate analysis of the observations. To do this, we calculate the vector of means  $\bar{x} = [\bar{x}_1 \quad \bar{x}_2 \quad \dots \quad \bar{x}_k]$  of dimension  $p$ , whose components are the means of each of the  $p$  variables and the covariance matrix. From the matrix of centered data  $\tilde{X}$ ,

$$\tilde{X} = X - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \bar{x}, \quad (6)$$

the symmetric and positive semidefinite matrix of covariance  $S = \frac{1}{n} \tilde{X}^T \tilde{X}$  is calculated.

The objective of describing multivariate data is to understand the dependence between the objective variable and the explanatory variables. For this we studied:

- 1 The relationship between pairs of variables;
- 2 Dependence between the objective variable and all the explanatory variables;
- 3 Dependence between the objective variable and the explanatory ones, but eliminating the effect of some of them.

The pairwise dependence between the variables is measured by the symmetric and positive semidefinite correlation matrix  $R$

$$R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1K} \\ r_{21} & 1 & \dots & r_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ r_{K1} & r_{K2} & \dots & 1 \end{bmatrix}, r_{jk} = \frac{S_{jk}}{S_j S_k} \quad (7)$$

so that there is an exact linear relationship between the variables  $X_j$  and  $X_k$  if  $|r_{jk}| = 1$ .

It may happen that there are variables that are very dependent on others, in which case it is convenient to measure their degree of dependence. Assuming that  $Y = X_j$  is the variable of interest, and calling  $\hat{Y}$  the variable used to estimate  $Y$ , the best linear predictor from the other variables, called the explanatory variables, is:

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \dots + \beta_K X_K, \quad (8)$$

where the parameter  $\beta_i$  is determined through the data that we have at our disposal. The problem is finding the set of parameters that minimizes  $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ , leading to

$$\begin{aligned} y &= Y - \hat{Y} \\ x_j &= X_j - \hat{X}_j, \quad j = 1, \dots, K \end{aligned} \quad (9)$$

and defining  $\hat{y} = \hat{Y} - \bar{Y}$ , we have  $Y - \hat{Y} = y - \hat{y}$ , and Equation (8) can be written as follows

$$\hat{y} = \alpha_0 + \alpha_1 x_1 + \dots + \alpha_K x_K, \quad (10)$$

Since minimizing  $\sum_{i=1}^n (Y_i - \hat{Y}_i)^2$  is equivalent to minimizing  $\sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2$ , by deriving this sum with respect to the  $\alpha_k$  parameters, we obtain a system of  $p - 1$  equations that can be written as follows:

$$\sum_{i=1}^n e_i x_{il} \quad l = 1, \dots, K, \quad l \neq j \quad (11)$$

Equation (9) indicates that the prediction errors must not be correlated with the explanatory variables, so that the covariance of both is zero, or else the residual vector must be orthogonal to the space generated by the explanatory variables. By defining the matrix  $X_R$ , of size  $n \times (p - 1)$ , obtained by eliminating the column in the matrix  $\bar{X}$  corresponding to the variable that we want to predict,  $y = x_j$ , the parameters are calculated by the normal equation system as follows

$$\alpha = (X_R^T X_R)^{-1} X_R^T y \quad (12)$$

and Equation (10), with these coefficients, is the multiple regression equation between variable  $y = x_j$  and the remaining variables  $x_i$ ,  $i \neq j$ ,  $i = 1, \dots, K$ .

To express this result based on the  $X_1, \dots, X_K$  variables of Equation (8), we must consider

$$\begin{aligned} \beta_i &= \alpha_i, i = 1, \dots, K \\ \beta_0 &= \alpha_0 + \bar{Y} - \sum_{i=1}^K \alpha_i \bar{X}_i \end{aligned} \quad (13)$$

The square of the multiple correlation coefficient (which can be greater than, less than, or equal to the sum of the squares of the simple correlations between variable  $y$  and each of the explanatory variables) [49] between the variable  $y = x_j$  and the rest is

$$R_j^2 = 1 - \frac{SS_{resid}}{SS_{total}} = 1 - \frac{1}{s_{jj} s^{jj}}, \quad (14)$$

where  $s_{jj} = s_j^2$  is the  $j$ -th diagonal element of the covariance matrix  $S$  and  $s^{jj} = \frac{1}{s_r^2(j)}$  is the  $j$ -th diagonal element of the  $S^{-1}$  matrix, which represents the residual variance of a regression between the  $j$ -th variable and the rest. As each time a variable is added to the model the number of degrees of freedom is reduced and the adjustment is increased, it is necessary to make a correction of this coefficient and calculate the adjusted  $R_j^2$ ,

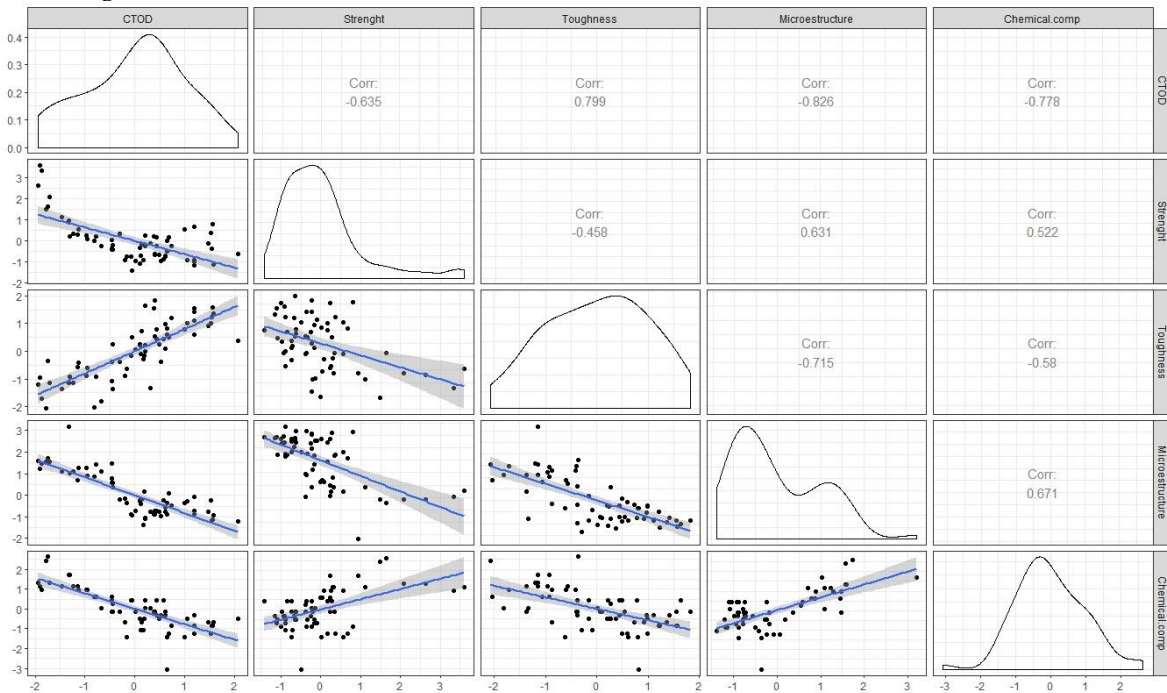
$$\bar{R}_j^2 = 1 - \frac{\frac{SS_{resid}}{(n - k)}}{\frac{SS_{total}}{(n - 1)}}, \quad (15)$$

where  $n$  is the total number of observations and  $k$  is the number of model variables; that is, the same calculation is made, but weighted by the degrees of freedom of the residuals,  $n - k$ , and the model,  $n - 1$ .

The R-squared  $RSQ = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$  is a descriptive measure of the predictive capacity of the model, and for a single explanatory variable is the square of the simple correlation coefficient between the two variables.

### 3.1. Previous Data Processing

Correlation coefficients were determined among the study variables. A high degree of correlation between toughness (CVN) and microstructure was observed, which was strongly supported in the bibliography. This relationship also depends on other variables that have not been considered in this experiment, such as temperature, tension state, or specimen geometry. Therefore, this particular relation between both variables is exclusive to this experiment and cannot be generalized.



**Figure 2.** Correlation, kernel density estimation (KDE), and scatterplots (the trendline that best fit linear relation is represented in blue) among the different variables.

Figure 2 shows the correlation and scatterplot diagrams between all the variables (objective and explanatory) taken two-by-two. The kernel density estimation (KDE) representation is also a way to estimate the probability density function of a random variable. A strong correlation can be observed among the CTOD and the explanatory variables, particularly toughness, microstructure, and chemical composition. Excluding the chemical composition, other variables do not seem to follow a normal distribution.

Figure 3 shows the quantiles of input samples (explanatory variables) versus standard normal quantiles (theoretical quantiles from a normal distribution). If the distribution of the explanatory variable is normal, the plot will be close to linear. Except for the chemical composition and toughness, the rest of the independent variables (the mechanical strength, called M. Strength onwards, and microstructure) do not seem to follow a normal distribution, so it would be advisable to make a transformation (for example, logarithmic type) before carrying out a multiple regression analysis. This can be explained by the observation of the KDE of the corresponding variable in Figure 2, where the M. Strength variable shows a positive skewness

towards lower values and the microstructure shows a slightly bimodal distribution (this effect is eliminated through a logarithmic transformation after the outlier exclusion).

With the aim of discarding the outliers that could influence observations, the Mahalanobis distance was used [49,50] for their detection and ten complete data sets were excluded (14%).

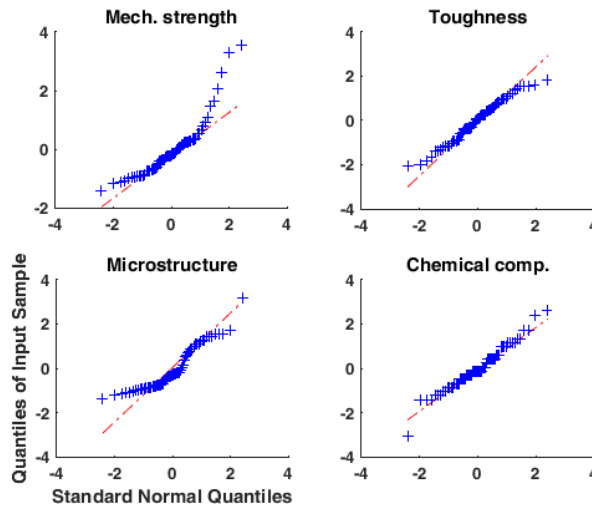


Figure 3. Quantiles of input sample versus standard normal quantiles.

### 3.2. Linear Regression Models

#### 3.2.1. Linear Model 1

Here,  $Y$  is considered as the study variable that may be linearly related with  $K$  explanatory variables  $X_1, X_2, \dots, X_K$  through  $\beta_0, \beta_1, \beta_2, \dots, \beta_K$  (regression coefficients). A multiple linear regression model can be written as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K + e \tag{16}$$

where  $e$  is the difference between the fitted relationship and the observations [51].

Using Equations (12) and (13), the values of the parameters are calculated. In Table 4, the coefficients for the multiple linear regression (Equation (16)) can be found. It can be seen that all coefficients are significantly different from zero, but toughness is the variable with the highest absolute value. In this case, the number of observations is 63, and the error degrees of freedom is 58.

The root mean square error (RMSE) is 0.216, which when compared to the range of the values of  $Y$  results in:

$$\frac{RMSE}{(Y_{MAX} - Y_{MIN})} = 0.1048 \approx 10\%, \tag{17}$$

Which provides an estimate of the possible error obtained from the real values of the CTOD variable. In Figure 2, it can be observed that the correlation coefficient between CTOD and toughness is 0.799. Considering all the independent variables the R-squared (RSQ) is 0.866, and



the adjusted RSQ value is 0.856, so there is a limited improvement from considering the CTOD toughness (or CTOD microstructure) correlation.

**Table 4.** Multiple linear model  $1 Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$ .

Parameters	Estimate	SE	p-Value
$\beta_0$	1.2202	0.022731	$1.0629 \times 10^{-46}$
$\beta_1 - M.Str.$	-0.080323	0.035816	$2.8756 \times 10^{-2}$
$\beta_2 - Tough.$	0.22424	0.040094	$6.3259 \times 10^{-7}$
$\beta_3 - Micros.$	-0.12972	0.047887	$8.8595 \times 10^{-3}$
$\beta_4 - C.Comp.$	-0.19243	0.038258	$5.0415 \times 10^{-6}$
F-Statistic p-value	-	-	$1.4 \times 10^{-24}$

Henceforth, for the models shown, the t-statistic (tStat) and F-statistic will be calculated and included. The first of them, tStat, calculated as estimated or standard error (SE), tests the null hypothesis that the corresponding coefficient is zero against the alternative that it is different from zero. To evaluate this coefficient, the corresponding p-value associated with a Student’s t distribution (for n observations) is calculated and compared with a confidence interval of 95%. If the p-value is less than 0.05, we can conclude that the variable is significant for the model.

Analogously, the F-statistic, calculated as:

$$F = \frac{\sum_{i=1}^n \frac{(\hat{y}_i - \bar{y})^2}{(p - 1)}}{\sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{(n - p)}} \tag{18}$$

tests the null hypothesis that one or more of the regression coefficients are significantly different from zero (meaning a significant linear regression relationship exists for the whole model). This value is compared with an F-distribution for a given confidence interval (95%) and is evaluated in the same way as the t-statistic (associated p-value less than 0.05). The F-distribution is more appropriate than Chi-square tests for small data sets [52].

Two different methods were used to verify that the obtained model was independent of the chosen data population: cross-validation and training-test samples.

The cross-validation was calculated with the *LeaveMout* method (see *crossvalind* Matlab function) with an M value of 1, which randomly selects one value and excludes it from the evaluation. This process is repeated 50 times and helps to verify that the statistical analysis is independent of the data set. The number of observations was 62, with  $RMSE = 0.218$ ,  $RSQ = 0.866$ , and adjusted  $RSQ = 0.856$ . The results are shown in Table 5.

**Table 5.** Cross-validation results.

Parameters	Estimate	SE	p-Value
$\beta_0$	1.2215	0.027668	$8.971 \times 10^{-46}$
$\beta_1 - M.Str.$	-0.08113	0.036144	$2.8692 \times 10^{-2}$
$\beta_2 - Tough.$	0.22244	0.040667	$1.0436 \times 10^{-6}$
$\beta_3 - Micros.$	-0.13334	0.049165	$8.8225 \times 10^{-3}$
$\beta_4 - C.Comp.$	-0.19006	0.039041	$9.3064 \times 10^{-6}$
F-Statistic p-value	-	-	$3.5 \times 10^{-24}$

The training test was done considering a set of 500 executions of samples from 50 observations (randomly selected from the whole data set) and test samples from 13 data sets. The averages of all RMSE and RSQ results are  $RMSE = 0.2275$  and  $RSQ = 0.8284$ , respectively.

Table 6 contains the values of RSQ and RMSE obtained with the reference model (linear model 1), cross-validation, and training test. As the values are similar (less than 5% discrepancy), we can conclude that the relation between the CTOD and the explanatory variables is independent of the data set.

**Table 6.** Comparison of the values of R-squared (RSQ) and root mean square error (RMSE).

Parameters	RSQ	RMSE
Linear Model 1	0.866	0.216
Cross reference	0.866	0.218
Training-Test	0.828	0.227

### 3.2.2. Linear Model 2

The significance of all variables was checked for all the explanatory variables, but it was observed that the microstructure was highly correlated with toughness. For that reason, a new model (linear model 2) was proposed, where the microstructure was eliminated from the original model.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4 \quad (19)$$

Table 7 shows the values of the parameters calculated for linear model 2, and the adjustment obtained ( $RMSE = 0.227$ ,  $RSQ = 0.849$ , and adjusted  $RSQ = 0.841$ ) was similar to the previous one (linear model 1).

**Table 7.** Multiple linear regression model 2  $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_4X_4$ .

Parameters	Estimate	SE	p-Value
$\beta_0$	1.2202	0.028657	$5.353 \times 10^{-46}$
$\beta_1 - M.Str.$	-0.11829	0.034686	$1.1768 \times 10^{-3}$
$\beta_2 - Tough.$	0.27944	0.036337	$1.8301 \times 10^{-10}$
$\beta_4 - C.Comp.$	-0.22775	0.03785	$1.2105 \times 10^{-7}$
F-Statistic p-value	-	-	$3.79 \times 10^{-24}$

### 3.2.3. Linear Models 3 and 4

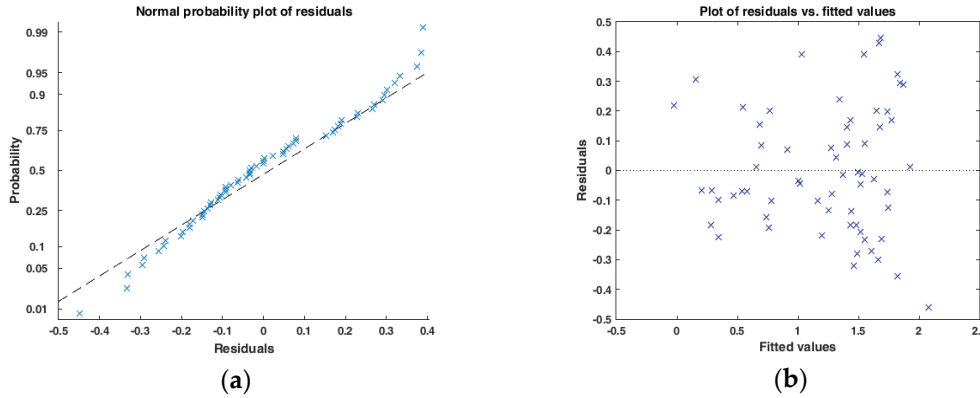
As the value of parameter  $\beta_1$  (coefficient of the mechanical strength) in linear model 1 was small compared to the values of the rest of the parameters, it was that the corresponding variable be eliminated to obtain a new model (linear model 3), considering that its contribution to the value of the CTOD variable was small. The values of the coefficients of linear model 3 are represented in Table 8.

**Table 8.** Linear regression model 3  $Y = \beta_0 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$ .

Parameters	Estimate	SE	p-Value
$\beta_0$	1.2202	0.028146	$1.9133 \times 10^{-46}$
$\beta_2 - Tough.$	0.22225	0.04143	$1.4246 \times 10^{-6}$
$\beta_3 - Micros.$	-0.17174	0.045549	$3.793 \times 10^{-4}$
$\beta_4 - C.Comp.$	-0.20715	0.038957	$1.961 \times 10^{-6}$
F-Statistic p-value	-	-	$1.32 \times 10^{-24}$

The quality of the adjustment is almost similar to that of the model with the four independent variables, with  $RMSE = 0.223$  and  $RSQ = 0.854$ .

Figure 4 shows the residuals of linear model 3, which can be considered as normally distributed.



**Figure 4.** Normal distribution of the residuals without variables. (a) Normal probability plot of residuals (b) Plot of residuals vs. fitted values.

Finally, a new model (linear model 4) is adopted considering the square of the first variable ( $M.Str^2$ ), and the contribution of the independent variables to the variable CTOD is checked (see Table 9). In this case, the coefficient of determination  $RSQ = 0.874$  is larger than in the purely linear model.

**Table 9.** Linear regression model 4  $Y = \beta_0 + \beta_1 X_1^2 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4$ .

Parameters	Estimate	SE	p-Value
$\beta_0$	1.2584	0.029113	$8.0522 \times 10^{-46}$
$\beta_1 - M.Str^2$	-0.038872	0.012639	$3.2021 \times 10^{-3}$
$\beta_2 - Tough$	0.21641	0.038792	$6.6704 \times 10^{-7}$
$\beta_3 - Micros$	-0.1562	0.042897	$5.7976 \times 10^{-4}$
$\beta_4 - C.Comp$	-0.19142	0.036789	$2.68 \times 10^{-6}$
F-Statistic p-value	-	-	$1.97 \times 10^{-25}$

Other tests have been done with different interactions between variables, but they do not improve the results.

### 3.3. Multivariate Adaptive Regression Splines (MARS)

Multivariate adaptive regression splines (MARS) is a non-parametric modeling method that extends the linear model (incorporating nonlinearities and interactions). It is a generalization of the recursive partitioning regression (RPR), which splits up the space of the explanatory variables

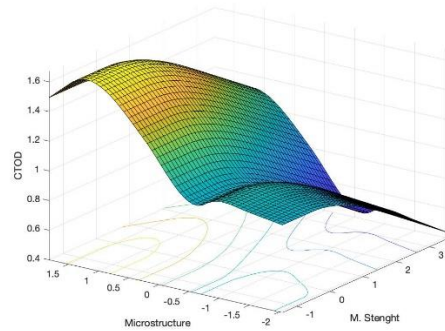
into different subregions. MARS generates cut points for the variables. These knots are identified through baseline functions, which indicates the beginning and end of a region.

In each region in which the space is divided, a base linear function of one variable is adjusted. The final model is constituted from a combination of the generated base functions [53]. The general expression of the model is:

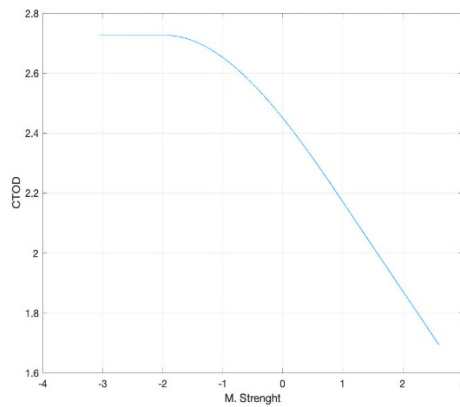
$$\hat{Y} = \sum_{i=1}^k c_i B_i(x), \tag{20}$$

where  $c_i$  is the constant coefficient and  $B_i$  is the base function.

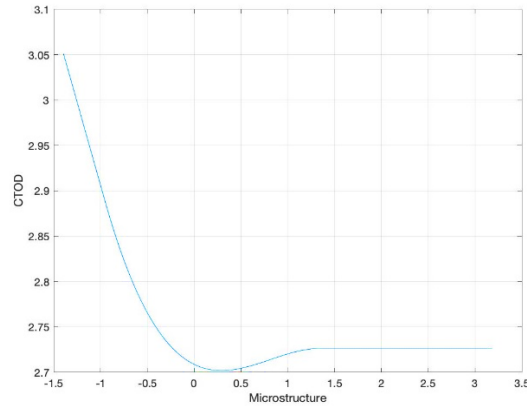
A MARS model was applied using cubic splines. This method considers nonlinear relationships among the CTOD variable and the explanatory ones using a spline adjustment, obtaining a  $RSQ = 0.86$  and  $RMSE = 0.16$ . With a training sample of 50 data sets and test sample of 13, the results were  $RSQ = 0.84$  and  $RMSE = 0.26$ . Additional information may be found in Figure 5, where the MARS model is plotted for two of the explanatory variables and two analysis of variance (ANOVA) functions (this visualizes the contribution of the ANOVA functions for the pairs CTOD-M. Strength and CTOD-microstructure in the MARS model).



(a)



(b)



(c)

**Figure 5.** (a) Multivariate adaptive regression splines (MARS) model plot for two of the explanatory variables together with its knot locations (up) and (b) the analysis of variance (ANOVA) function for the pairs CTOD-M. Strength (left) and (c) CTOD-microstructure (right) (using ARESLab toolbox: Jekabsons G., ARESLab: Adaptive Regression Splines Toolbox for Matlab/Octave, 2016, available at <http://www.cs.rtu.lv/jekabsons/>).

Again, these values do not improve on those obtained with previous models.

### 3.4. Other Models

Other models were studied in order to observe a possible improvement with respect to the initial model (linear model 1).

In the first place, we proposed a generalized linear model considering a Gaussian distribution and an identity linking function, the parameters for which are included in Table 10 (Generalized linear regression model 1—GLM1). It is noted that the p-value of the mechanical strength is greater than 0.05, therefore, the variable  $X_1$  (mechanical strength) may not be significant.

**Table 10.** Generalized linear regression model 1 (GLM1)  $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$ .

Parameters	Estimate	SE	p-Value
$\beta_0$	1.2202	0.027231	$1.0629 \times 10^{-46}$
$\beta_1 - M.Str.$	-0.08032	0.035816	$2.8756 \times 10^{-2}$
$\beta_2 - Tough.$	0.22424	0.040094	$6.3259 \times 10^{-7}$
$\beta_3 - Micros.$	-0.12972	0.047887	$8.859 \times 10^{-3}$
$\beta_4 - C.Comp.$	-0.19243	0.038259	$5.0415 \times 10^{-6}$
F-Statistic p-value	-	-	$1.4 \times 10^{-24}$

For this reason, a generalized linear model was calculated without the mechanical strength influence (GLM2), whose results are shown in Table 11, with  $RSQ = 0.675$  and  $RMSE = 0.3396$  obtained. These values do not improve on those obtained with previous models.

**Table 11.** Generalized linear regression model 2 (GLM2)  $Y = \beta_0 + \beta_2X_2 + \beta_3X_3 + \beta_4X_4$ .

Parameters	Estimate	SE	p-Value
$\beta_0$	1.0169	0.043522	$3.701 \times 10^{-27}$
$\beta_2 - Tough.$	-0.032814	0.037709	$3.8872 \times 10^{-2}$
$\beta_3 - Micros.$	0.3569	0.059269	$2.6872 \times 10^{-7}$
$\beta_4 - C.Comp.$	0.15866	0.031816	$9.1927 \times 10^{-6}$
F-Statistic p-value	-	-	$3.93 \times 10^{-13}$

In the second place, we considered a regression tree model [54]. To make a prediction for a given observation, we used the mean (or the mode) of the observations that were in the same region of the multidimensional space of predictors. The rules that were used to divide the predictor space can be represented as a tree [55].

The order of importance of the predictive variables, from highest to lowest, is microstructure, toughness, mechanical strength, and chemical composition. Therefore, the variable microstructure is the one that provides the value that maximizes the information about the

dependent variable (CTOD) if it is smaller than 0.26, otherwise it is the toughness that carries more information. Nevertheless, the values associated with each subtree for the training sample (13) are between 3 to 7 times bigger than those of the test sample (50), which indicates bad behavior of the model.



#### 4. Results and Discussion

After having compared the previous model, due to the simplicity and reasonable accuracy, and despite the unbalanced weigh of the different variable's parameters ( $\beta_i$ ), linear model 1 ( $RSQ = 0.866$ ) is proposed as a predictive model of the values of CTOD.

The standardized model can be expressed as:

$$CTOD = 1.2202 - 0.080323[M.Str] + 0.22424[Tough] - 0.12972[Micros] - 0.19243[C.Comp] \quad (21)$$

where [.] represents the standardized values of the corresponding variable. The standardization process should be reverted to allow the use of the testing data directly:

$$\begin{aligned} \frac{(CTOD[mm]-1.0269)}{0.586} = & 1.2202 - 0.080323 \frac{(M.Str[HV10]-221.09)}{41.87} \\ & + 0.22424 \frac{(Tough[J]-175.421)}{51.84} - 0.12972 \frac{(Micros[\mu m]-169.23)}{55.44} \\ & - 0.19243 \frac{(CE-0.3831)}{0.054} \end{aligned} \quad (22)$$

where  $CTOD[mm]$  is the value of the crack tip opening displacement, expressed in  $mm$ .  $M.Str[HV10]$  is the effect of the mechanical strength of the material as the average of three hardness measurements expressed in  $[HV10]$  units.  $Tough[J]$  is the average value of the two Charpy V-notch measurements, with subsized specimens extracted from the interest zone and expressed in Joules.  $Micros[\mu m]$  is the average size of the metallic grains expressed in  $\mu m$ . Finally,  $CE$  is the effect of the chemical composition as the carbon equivalent calculated with the equation proposed in AWS D1.1 (adimensional).

## 5. Conclusions

The use of multivariate analysis has been proven viable for relating complex fracture mechanics parameters to well-known material properties. The industrial suitability of the methodology depends on the experimental set, specifically the availability of samples, the number of tests, and the choice of variables.

These chosen variables are significantly related with the CTOD (see p-value for linear regression model 1). Also, there is well-known experience within the manufacturing industry relating these variables with actual changes during the welding process. As an example, there is a wide background of knowledge on how the shielding gas, the welding speed, or the bead scheme affect the grain size or the hardness of a given welded joint. Using the proposed model, it is possible for the industry to transfer this knowledge on how these variables may affect the CTOD value.

The final model is precise and functional, with an estimated error of ~10% (within the limits covered by the experimental set). This error is compatible with the current uncertainty of the CTOD testing process. Besides, the model is not dependent on which subgroup of data is used for the modeling process. It is proposed to use this final model predictively, using the results of the tests for the explanatory variables (it is cheaper, simpler, and more available than the CTOD) to compute the CTOD value estimator. If this value (considering the mentioned error) is greater than the critical value (acceptance criteria) specified in the design code, rule, or standard, the expensive CTOD test can be dispensed.

The usefulness of the model has been proven within the limits of the experimental set for offshore steel welded joints of high thickness. Nevertheless, the influences of other variables not explicitly considered in this work were not tested, even for the mentioned category, and are out of the scope of the presented model. Future developments of the model could include, among others, the influence of testing temperature, different positioning or shape of the notch, post-weld heat treatments, or type of failure category (brittle, ductile-brittle, and ductile).

**Author Contributions:** Conceptualization, A.P.V. and A.B.S.; methodology, A.P.V. and A.B.S.; software, A.P.V. and Z.F.M.; validation, Z.F.M. and M.M.F.; formal analysis, A.P.V.; investigation, A.P. V. and A.B.S.; resources, A.P.V. and A.B.S.; data curation, Z.F.M.; writing—original draft preparation, A.P.V.; writing—review and editing, A.P.V., Z.F.M., and M.M.F.; visualization, A.P.V.; supervision, A.B.S.

**Funding:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Acknowledgments:** The authors would like to thank the private laboratory TAM for access to testing equipment, software, and mediation for the transfer of testing samples.

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

## References

1. Rodrigues, P.E.; Wong, W.K.; Rogerson, J.H. Weld Defect Distributions in Offshore Platforms and Their Relevance to Reliability Studies, Quality Control and In-Service Inspection. In Proceedings of the Offshore Technology Conference, Houston, TX, USA, 5–8 May 1980.
2. American Bureau of Shipping. *Design Standards for Offshore Wind Farms*; American Bureau of Shipping: Houston Texas, USA, 2011.
3. Smith, M. *Fitness for Service*; API 579-1/ASME FFS-1; 2007.
4. Offshore Standard Det Norske Veritas DNV-OS-C101. Available online: <https://www.scribd.com/document/50782302/DNV-OS-C101-Design-of-offshore-steel-structures-general-LRFD-Method-October-2008> (accessed on 27 September 2019)
5. Zhu, X.; Joyce, A. Review of fracture toughness (G, K, J, CTOD, CTOA) testing and standardization. *Eng. Fract. Mech.* **2012**, *85*, 1–46.
6. Irwin, G.R.; Kies, J.A. Critical energy rate analysis for fracture strength. *Weld. J. Res. Suppl.* **1954**, *19*, 193–198.
7. Wells, A.A. Application of fracture mechanics at and beyond general yielding. *Br. Weld. J.* **1963**, *10*, 563–570.
8. Rice, J.R. A path independent integral and the approximate analysis of strain concentration by notches and cracks. *J. Appl. Mech.* **1968**, *35*, 379–386.
9. Newman, J.C.; James, M.A.; Zerbst, U. A review of the CTOA/CTOD fracture criterion. *Eng. Fract. Mech.* **2003**, *70*, 371–385.
10. Akourri, O.; Louah, M.; Kifani, A.; Gilgert, G.; Pluvinage, G. The effect of notch radius on fracture toughness J<sub>Ic</sub>. *Eng. Fract. Mech.* **2000**, *65*, 491–505.
11. Li, X.; Song, Y.; Ding, Z.; Bao, S.; Gao, Z. A modified correlation between K<sub>JIC</sub> and Charpy V-notch impact energy of Chinese SA508-III steel at the upper shelf. *J. Nucl. Mater.* **2018**, *505*, 22–29.
12. ASTM E1290-08e1c. *Standard Test Method for Crack tip Opening Displacement (CTOD) Fracture Toughness Measurement*; ASTM International: West Conshohocken, PA, USA, 2008.
13. Rencher, A.C. *Methods of Multivariate Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2002.
14. Dunne, D.; Tsuei, H.; Sterjovski, Z. Artificial neural networks for modeling of the impact toughness of steel. *ISIJ Int.* **2004**, *44*, 1599–1607.
15. Haque, M.E.; Sudhakar, K.V. ANN back-propagation prediction model for fracture toughness in microalloy steel. *Int. J. Fatigue* **2002**, *24*, 1003–1010.
16. ASM International. *ASM Handbook*; ASM International: Materials Park, OH, USA, 1996; Volume 19.
17. Castelluccio, G.M.; McDowell, D.L. Microstructure-sensitive small fatigue crack growth assessment: Effect of strain ratio, multiaxial strain state, and geometric discontinuities. *Int. J. Fatigue* **2016**, *82*, 521–529.
18. Chen, Y.Q.; Pan, S.P.; Zhou, M.Z.; Yi, D.Q.; Xu, D.Z.; Xu, Y.F. Effects of inclusions, grain boundaries and grain orientations on the fatigue crack initiation and propagation behavior of 2524-T3 Al alloy. *Mater. Sci. Eng.* **2013**, *580*, 150–158.
19. Cheng, T.C.; Yu, C.; Yang, T.C.; Huang, C.Y.; Lin, H.C.; Shiue, R.K. Microstructure and Impact Toughness of Offshore Steel. *Arch. Metall. Mater.* **2018**, *63*, 167–172.
20. Pegues, J.W.; Roach, M.D.; Shamsaei, N. Influence of microstructure on fatigue crack nucleation and microstructurally short crack growth of an austenitic stainless steel. *Mater. Sci. Eng.* **2017**, *707*, 657–667.
21. Wei, L.; Pan, Q.; Huang, H.; Feng, L.; Wang, Y. Influence of grain structure and crystallographic orientation on fatigue crack propagation behavior of 7050 alloy thick plate. *Int. J. Fatigue* **2014**, *66*, 55–64.

22. Yuan, H.; Zhang, W.; Castelluccio, G.M.; Kim, J.; Liu, Y. Microstructure-sensitive estimation of small fatigue crack growth in bridge steel welds. *Int. J. Fatigue* **2018**, *112*, 183–197.
23. Yuan, H.; Zhang, W.; Kim, J.; Liu, Y. A nonlinear grain-based fatigue damage model for civil infrastructure under variable amplitude loads. *Int. J. Fatigue* **2017**, *104*, 389–396.
24. ASTM E112. *Standard Test Methods for Determining Average Grain Size*; ASTM International: West Conshohocken, PA, USA, 2013.
25. ASTM E3–11. *Standard Guide for Preparation of Metallographic Specimens*; ASTM International: West Conshohocken, PA, USA, 2018.
26. ASTM E407–07. *Standard Practice for Micro-Etching Metals and Alloys*; ASTM International: West Conshohocken, PA, USA, 2015.
27. Talaş, Ş. The assessment of carbon equivalent formulas in predicting the properties of steel weld metals. *Mater. Des.* **2010**, *31*, 2649–2653.
28. Lee, S.G.; Lee, D.H.; Sohn, S.S.; Kim, W.G.; Um, K.K.; Kim, K.S.; Lee, S. Effects of Ni and Mn addition on critical crack tip opening displacement (CTOD) of weld-simulated heat-affected zones of three high-strength low-alloy (HSLA) steels. *Mater. Sci. Eng.* **2017**, *697*, 55–65.
29. Yurioka, N. Physical Metallurgy of Steel Weldability. *ISIJ Int.* **2001**, *41*, 566–570.
30. American Welding Society D1.1/D1.1M. *Structural Welding Code Steel*; American Welding Society: Miami, FL, USA, 2010.
31. Anderson, T.L. *Fracture Mechanics: Fundamentals and Applications*, 3rd ed.; Taylor & Francis: Boca Raton, FL, USA, 2005.
32. ASTM E92. *Standard Test Methods for Vickers Hardness and Knoop Hardness of Metallic Materials*; ASTM International: West Conshohocken, PA, USA, 2017.
33. Khor, W.; Moore, P.L.; Pisarski, H.G.; Haslett, M.; Brown, C.J. Measurement and prediction of CTOD in austenitic stainless steel. *Fatigue Fract. Eng. Mater. Struct.* **2016**, *39*, 1433–1442.
34. Maropoulos, S.; Ridley, N.; Kechagias, J.; Karagiannis, S. Fracture toughness evaluation of a H.S.L.A. steel. *Eng. Fract. Mech.* **2004**, *71*, 1695–1704.
35. ASME BPVC IX. *Boiler Pressure Vessel Code (BPVC) Welding, Brazing, and Fusing Qualifications*; The American Society of Mechanical Engineers (ASME): New York, NY, USA, 2017.
36. ASTM E23. *Standard Test Methods for Notched Bar Impact Testing of Metallic Materials*; ASTM International; West Conshohocken, PA, USA, 2018.
37. Yang, Y.Y.; Mahfouf, M.; Panoutsos, G. Probabilistic Characterization of Model Error Using Gaussian Mixture Model—with Application to Charpy Impact Energy Prediction for Alloy Steel. *Control Eng. Pract.* **2012**, *20*, 82–92.
38. EN 10225. *Weldable Structural Steels for Fixed Offshore Structures*; European Committee for Standardization: Brussels, Belgium, 2009.
39. EN ISO 10025. *Hot Rolled Products of Structural Steels*; European Committee for Standardization: Brussels, Belgium, 2006.
40. EN ISO 17637. *Non-Destructive Testing of Welds—Visual Testing of Fusion-Welded Joints*; European Committee for Standardization: Brussels, Belgium, 2017.
41. EN ISO 17638. *Non-Destructive Testing of Welds—Magnetic Particle Testing*; European Committee for Standardization: Brussels, Belgium, 2017.
42. EN ISO 17640. *Non-Destructive Testing of Welds—Ultrasonic Testing—Techniques, Testing Levels, and Assessment*; European Committee for Standardization: Brussels, Belgium, 2011.
43. Ávila, J.A.; Lima, V.; Ruchert, C.O.; Mei, P.R.; Ramírez, A.J. Guide for Recommended Practices to Perform Crack Tip Opening Displacement Tests in High Strength Low Alloy Steels. *Soldag. Inspeção* **2016**, *21*, 290–302.
44. Antunes, F.V.; Branco, R.; Prates, P.A.; Borrego, L. Fatigue crack growth modeling based on CTOD for the 7050-T6 alloy. *Fatigue Fract. Eng. Mater. Struct.* **2017**, *40*, 11.

45. Antunes, F.V.; Rodrigues, S.M.; Branco, R.; Camas, D. A numerical analysis of CTOD in constant amplitude fatigue crack growth. *Theor. Appl. Fract. Mech.* **2016**, *85*, 45–55.
46. Janssen, M.; Zuidema, J.; Wanhill, R.J.H. *Fracture Mechanics*, 2nd ed.; Spon Press: New York, NY, USA, 2004.
47. Kawabata, T.; Tagawa, T.; Sakimoto, T.; Kayamori, Y.; Ohata, M.; Yamashita, Y.; Tamura, E.I.; Yoshinari, H.; Aihara, S.; Minami, F.; et al. Proposal for a new CTOD calculation formula. *Eng. Fract. Mech.* **2016**, *159*, 16–34.
48. Khor, W.L.; Moore, P.; Pisarski, H.; Brown, C. Comparison of methods to determine CTOD for SENB specimens in different strain hardening steels. *Fatigue Fract. Eng. Mater. Struct.* **2017**, *41*, 551–564.
49. Cuadras, C.M. *Métodos de Análisis Multivariante*; Eunibar: Barça, Barcelona, 1981.
50. Everitt, B.S. *Cluster Analysis*; Edward Arnold: London, UK, 1993.
51. Rao, C.R.; Toutenburg, H.; Shalabh Heumann, C. *Linear Models and Generalizations*; Springer Series in Statistics; Springer: Berlin/Heidelberg, Germany, 2008.
52. Goldstein, H. *Introduction to F-testing in Linear Regression Models*; Lecture note of the Department of Statistics, University of Oslo: Oslo, Norway, 2014.
53. Vanegas, J.; Vásquez, F. Multivariate Adaptive Regression Splines (MARS), una alternativa para el análisis de series de tiempo. *Gac. Sanit.* **2017**, *31*, 235–237.
54. Berk, R.A. *Classification and Regression Trees (CART). Statistical Learning from a Regression Perspective*; Springer Series in Statistics; Springer: New York, NY, USA, 2008.
55. Seoane, J.; Carmona, C.P.; Tarjuelo, R.; Planillo, A. *Árboles de Regresión y Clasificación; Análisis Bioestadístico con Modelos de Regresión en R*, UAM: Mexico City, Mexico, 2014.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).