



Trabajo Fin de Máster

Desarrollo de un modelo basado en redes neuronales para la predicción de cargas aplicadas en la tibia humana. Aplicación al diseño de prótesis de rodilla de revisión

Autor/es

MªPaz Quílez Callén

Director/es

MªÁngeles Pérez Ansón

Escuela de Ingeniería y Arquitectura
2013

Agradecimientos

A mis padres, Andrés y M^a Jesús, y a mi tío Alberto por haberme proporcionado todo lo que he necesitado para llegar hasta aquí.

A mi novio Alex por su ayuda, comprensión y paciencia en los momentos más duros.

A M^a Ángeles por su gran trabajo y dedicación, y por haberme ayudado siempre que lo he necesitado a pesar de las cosas importantes que ha tenido entre manos.

A Noelia por echarme una mano cuando lo he necesitado y su paciencia para repetirme las cosas una y otra vez.

A Belén Seral, doctora en Traumatología en el Hospital Clínico Lozano Blesa, por darme la oportunidad de entrar en un quirófano y ver de cerca lo estudiado para este trabajo, así como lo que he aprendido al ver las cosas desde el punto de vista de un médico.

Y finalmente a todos esos compañeros que han hecho mi paso por este Máster una experiencia inolvidable, sobre todo a los buenos amigos que me llevo de él.

Resumen

Este trabajo se ha desarrollado en el Departamento de Ingeniería Mecánica, dentro del área de Mecánica de Medios Continuos y Teoría de Estructuras. Su principal objetivo es la estimación de las cargas aplicadas en la tibia de un paciente concreto mediante un método basado en Redes Neuronales Artificiales y, posteriormente, utilizar estas cargas para la evaluación del efecto de distintos tipos de prótesis de rodilla de revisión en la remodelación ósea.

El hueso está sometido a un proceso continuo de renovación que se conoce con el nombre de remodelación ósea. Este proceso adaptativo se produce debido a los importantes cambios morfológicos que sufre el hueso en respuesta a la situación mecánica a la que se encuentra solicitado. Al incorporar una prótesis al hueso, se modifica la distribución de tensiones a la que se encuentra sometido, produciendo una formación o reabsorción ósea dependiendo de la zona del hueso y del tipo de prótesis.

Un aspecto poco estudiado en este campo es la determinación de las cargas reales a las que es sometida la tibia de un paciente concreto. Este complejo problema se ha intentado abordar en este Trabajo Fin de Master mediante una combinación del método de los Elementos Finitos y las Redes Neuronales Artificiales. Para ello este trabajo se ha dividido en dos partes: la primera parte se centra en la determinación de las cargas que afectan a la tibia de un paciente concreto mediante un método basado en Redes Neuronales, y la segunda la utilización de dichas cargas para analizar la reabsorción ósea que producen cuatro tipos distintos de prótesis de rodilla de revisión.

Para la primera parte del trabajo se ha analizado el problema de remodelación ósea, aplicando un rango de condiciones de carga al modelo 3D de una tibia real. De este análisis se han obtenido unas densidades óseas, que sirven de entrada para la Red Neuronal Artificial. Después de entrenar la red con una función sigmoidea como función de entrada, distinto número de neuronas en la capa intermedia y una función lineal como salida, se han obtenido unas cargas, las cuales se han comparado con las condiciones de carga aplicadas inicialmente y se ha obtenido un error de cálculo. Tras determinar el número de entradas y neuronas de la capa intermedia más adecuado, se han introducido las cargas obtenidas en el modelo de la tibia con distintas prótesis, evaluando su efecto en la remodelación ósea.

Como parte clínica de este trabajo se asistió a una intervención quirúrgica en el Hospital Clínico Universitario Blesa, en la que se realizó un recambio de prótesis de rodilla. Además se han mantenido reuniones periódicas con médicos del Departamento de Traumatología de dicho hospital.

Índice

1. Introducción	6
1.1. La Biomecánica.....	6
1.2. La articulación de la rodilla y sus prótesis de revisión	7
1.2.1. Anatomía de la rodilla	7
1.2.2. Artroplastia total de rodilla	8
1.2.3. Prótesis de rodilla de revisión	8
1.3. Objetivos	10
1.4. Descripción del trabajo	10
 2. Redes neuronales aplicadas a la determinación de cargas en la tibia de un paciente concreto	12
2.1. Redes neuronales: Introducción	12
2.2. Planteamiento de la metodología desarrollada.....	13
2.3. Definición del problema de remodelación ósea	14
2.3.1. Modelo de remodelación ósea.....	14
2.3.2. Modelo 3D de elementos finitos de la tibia	15
2.3.3. Condiciones de carga del modelo.....	17
2.4. Definición del problema inverso de remodelación ósea.	18
2.4.1. Entradas y salidas del problema inverso de remodelación ósea.....	18
2.4.2. Características de la red neuronal utilizada.....	19
2.5. Resultados	20
2.5.1. Evaluación de la red neuronal	20
2.5.2. Modelo de entrenamiento.....	21
2.5.3. Reducción del número de entradas para la red neuronal.....	23
2.5.4. Validación	28
2.6. Conclusiones	29
 3. Diseño de una prótesis de rodilla de revisión.....	30
3.1. Introducción	30
3.2. Modelos de elementos finitos y remodelación ósea.....	30
3.3. Resultados	33
3.4. Conclusiones	41
 4. Conclusiones y líneas futuras.....	42
4.1. Resumen	42
4.2. Conclusiones	42
4.3. Futuras líneas de investigación	43
 Anexo I. Redes neuronales.....	44
I. El modelo biológico.....	44
II. Modelo de neurona artificial.....	45
i. Entradas y salidas	46
ii. Pesos sinápticos	46
iii. Regla de propagación.....	46

iv.	Función de activación	46
v.	Función de salida.....	47
III.	Arquitectura de una red neuronal	47
IV.	Funcionamiento de la red neuronal	48
i.	Fase de entrenamiento.....	48
ii.	Fase de operación o test	49
V.	Redes neuronales con conexión hacia delante	49
i.	Perceptrón simple.....	49
ii.	Perceptrón multicapa (MLP).....	49
VI.	Entrenamiento de la red neuronal	50
i.	Algoritmo Backpropagation.....	50
ii.	Generalización	51
	Bibliografía	53

1. Introducción

1.1. La Biomecánica

La biomecánica es una disciplina científica que tiene por objetivo el estudio de las estructuras de carácter mecánico que existen en los seres vivos, fundamentalmente del cuerpo humano. Esta área de conocimiento se apoya en diversas ciencias biomédicas, utilizando los conocimientos de la mecánica, la ingeniería, la anatomía, la fisiología y otras disciplinas, para estudiar el comportamiento del cuerpo humano y resolver los problemas derivados de las diversas condiciones a las que puede verse sometido. La biomecánica ha tenido un gran desarrollo en relación con aplicaciones de la ingeniería, informática y de modelos matemáticos, para el conocimiento de los sistemas biológicos y de partes del cuerpo humano, facilitando su aplicación práctica y su utilización como nuevos métodos de valoración y diagnóstico.

El objetivo de la biomecánica es resolver los problemas que surgen de las diversas condiciones a las que puede verse sometido nuestro cuerpo en distintas situaciones. Las posibilidades que la biomecánica ofrece al plantear y resolver problemas relacionados con la mejora de nuestra salud y calidad de vida la han consolidado como un campo de conocimientos en continua expansión, capaz de aportar soluciones científicas y tecnológicas muy beneficiosas para nuestro entorno más inmediato.

La proyección industrial de la biomecánica ha alcanzado a diversos sectores, sirviendo de base para la concepción y adaptación de numerosos productos: técnicas de diagnóstico, implantes e instrumental quirúrgico, prótesis, ayudas técnicas a personas con discapacidad, sistemas de evaluación de nuestras actividades, herramientas y sistemas de seguridad en automoción, entre otros muchos.

Uno de los desarrollos más importantes de la biomecánica está asociado con la ortopedia. Este campo se ha convertido en una disciplina utilizada tanto en la vertiente más relacionada con la ingeniería, para el desarrollo de nuevos implantes y materiales, como en el terreno de la biología, en investigaciones sobre los procesos experimentados por las células en relación a las tensiones y esfuerzos, así como en simulaciones de tejidos (cartílagos, tendones, huesos...).

Debido al envejecimiento de la población y al aumento de los accidentes de tráfico, uno de los grandes retos a los que se enfrentan tanto médicos como ingenieros es la mejora de los diseños de prótesis. La incorporación en un hueso fracturado de un implante, ya sea provisional o definitivo, produce una redistribución de tensiones que afectan a las propiedades del hueso modificándolas. La aparición de pérdidas óseas, el aflojamiento entre implante y hueso, y la falta de estabilización de la fractura son efectos importantes como consecuencia de la inclusión de un implante protésico.

El desarrollo de técnicas de simulación por ordenador del comportamiento de huesos y tendones hace posible conocer de forma más aproximada cuál será la respuesta del hueso a largo plazo. Es importante señalar que el uso de técnicas computacionales, aunque sean más rentables, no pueden sustituir completamente a las técnicas experimentales. Los resultados obtenidos mediante sistemas informáticos deben ser validados con ensayos en laboratorio.

1.2. La articulación de la rodilla y sus prótesis de revisión

1.2.1. Anatomía de la rodilla

La rodilla es la articulación más grande del cuerpo humano, así como una de las más complejas. Está compuesta por el juego de tres huesos, fémur, tibia y rótula. Los dos primeros conforman el cuerpo principal de la articulación, que soporta el peso corporal, y la rótula cumple una misión atípica, a modo de polea sobre la que se apoyan los tendones cuadripucital y rotuliano (*Figura 1.1*)

Es además una articulación bicondilea. Los dos cóndilos femorales ruedan sobre la superficie casi plana de los platillos tibiales. El apoyo de un hueso sobre otro es libre, sin topes óseos para mantenerla y necesita el amarre de los ligamentos.

Las superficies de contacto entre los huesos están cubiertas por una capa de cartílago. Todo el conjunto está envuelto por una cápsula articular, constituyendo un espacio cerrado. La cubierta íntima de la cápsula es la membrana sinovial, cuya misión principal es la secreción del líquido del mismo nombre, fundamental en la fisiología articular con misiones de lubricación y defensa. El exceso de secreción da lugar a un cúmulo sinovial que causa aumento de la presión intrarticular y genera el molesto y conocido derrame.

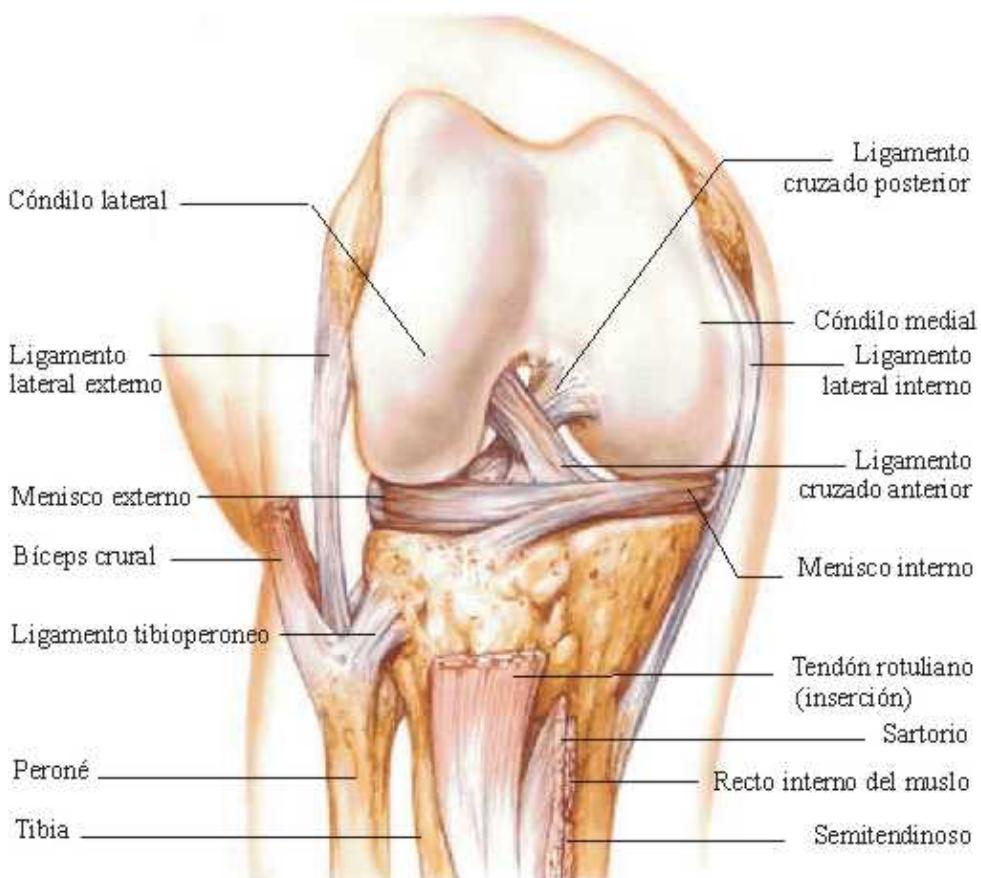


Figura 1.1. Articulación de la rodilla

1.2.2. Artroplastia total de rodilla

El reemplazo total de rodilla (RTR) es un procedimiento quirúrgico que se realiza para reemplazar la articulación de la rodilla por una prótesis artificial. Es también llamado artroplastia de rodilla, artroplastia total de rodilla, o artroplastia tricompartimental de rodilla.

Esta técnica fue realizada por primera vez en el año 1968 y es uno de los avances más importantes de la cirugía ortopédica. Cada año, se llevan a cabo aproximadamente 478 000 reemplazos totales de rodilla en los Estados Unidos.

El RTR es normalmente efectuado en pacientes que tienen desgastes o daños severos en la articulación, debido a la artritis o a una lesión. En estos casos las personas pueden presentar dificultad para realizar sus actividades de la vida diaria tales como caminar o subir y bajar escaleras e incluso pueden sentir molestias estando sentado o acostado.

Al restaurar las superficies dañadas y desgastadas de la rodilla, la cirugía de reemplazo total de rodilla puede aliviar el dolor, corregir la deformidad de la pierna y ayudarle a realizar las actividades cotidianas.

La operación proporciona un alivio del dolor, para que la persona pueda ser capaz de llevar a cabo las actividades del diario. La rodilla artificial podría permitirle retornar a sus actividades deportivas o laborales, pero las actividades que sobrecarguen la rodilla artificial, deben ser evitadas.

1.2.3. Prótesis de rodilla de revisión

Las prótesis primarias pueden sufrir un fallo por aflojamiento, metalosis o diversas infecciones. En ese caso se realiza una cirugía de revisión en la que se sustituyen la prótesis primaria por una prótesis con características diferentes y cuya diferencia fundamental es que los vástagos (femoral y tibial) tienen una mayor longitud. A estas prótesis se les llama prótesis de revisión.

Prótesis de vainas TC3 (Depuy, Johnson & Johnson)

El aflojamiento aséptico y el desgaste son las principales causas de las revisiones de rodilla en más del 40% de los pacientes. La bandeja de esta prótesis (*Figuras 1.2 y 1.3*) proporciona libertad de rotación para distribuir las fuerzas de aflojamiento, haciéndola adecuada para aumentar la restricción mecánica.



Figura 1.3. Prótesis de vainas TC3 (Depuy)



Figura 1.2. Prótesis de vainas TC3 (Depuy)

Su bandeja universal la hace compatible con todas las inserciones de plataforma giratoria de las prótesis constreñidas y con bisagra. Las vainas metafisarias cargan el hueso para formar una base sólida, evitando la resección ósea excesiva y ayudando a preservar la restauración de la articulación. A través de una serie de mejoras diseñadas para aumentar la artroplastia total de rodilla, esta solución quirúrgica puede resultar en soluciones más personalizadas y una mayor facilidad de uso.



Prótesis con offset LCCK (Zimmer)

El componente femoral NexGen Legacy LCCK (*Figura 1.4*) ha sido desarrollado para pacientes, que según el cirujano, requieran una estabilización protésica adicional debido a una función defectuosa de los ligamentos tanto laterales como cruzados y además requieran bloques de aumento y/o vástagos de extensión debido a stocks óseos inadecuados. Los componentes LCCK están disponibles únicamente en una versión no recubierta para fijación cementada.



Figura 1.4. Prótesis LCCK (Zimmer)

Para pacientes con unos ligamentos laterales no funcionales o para aquellas rodillas que no pueden ser estabilizadas mediante las usuales liberaciones de ligamentos, la rodilla LCCK presenta una eminencia tibial elevada y un cajón intercondíleo más profundo. Un ajuste adecuado entre la eminencia y el cajón proporciona estabilidad a medida que el rollback natural es inducido, inhibiendo la subluxación posterior, limitando el movimiento en varo/valgo a 1.25 grados y la rotación interna/externa a 2 grados. Este diseño acomoda un rango de movimiento teórico superior a los 120 grados (*Figura 1.5*).

Figura 1.5. Vástago en offset

1.3. Objetivos

El hueso está sometido a un proceso continuo de renovación que se conoce con el nombre de remodelación ósea. Este proceso adaptativo se produce debido a los importantes cambios morfológicos que sufre el hueso en respuesta a la situación mecánica a la que se encuentra solicitado. La incorporación de una prótesis acelera este proceso adaptativo.

Uno de los aspectos menos estudiado y más complejo en este campo es la determinación de las cargas que actúan en la tibia. De ahí que **uno de los principales objetivos de este TFM sea la estimación de las cargas que actúan sobre la tibia de un paciente concreto a partir de la distribución de su densidad ósea.**

Para alcanzar este objetivo se va a plantear una metodología que combina el **método de los elementos finitos** con las **redes neuronales**. Para extrapolar este problema a la vida real y obtener unos resultados más concretos se puede utilizar información de un paciente en particular como su geometría del hueso y la distribución de densidad ósea. En realidad, el problema que se plantea es de dinámica inversa. Partiendo de una distribución de densidades, se quiere saber qué cargas generan estas densidades, esto se debe a que la estructura morfológica del hueso varía en función de las condiciones de carga (remodelación). La aplicación de estas metodologías de dinámica inversa es muy compleja y costosa, sin embargo, existen otras técnicas que requieren un tiempo computacional menor. Por ello, en este trabajo se va a utilizar una metodología que consiste en la aplicación de redes neuronales, en concreto, MLP perceptrón multicapa y el análisis por elementos finitos con un modelo de remodelación ósea.

El **segundo objetivo** de este TFM es la **aplicación de las cargas anteriormente obtenidas para determinar la prótesis de rodilla de revisión más adecuada para un paciente concreto**. Cuando se incorpora un implante en el hueso, se produce una importante alteración del estado tensional sobre el tejido. Por ejemplo, el implante, que posee una mayor rigidez, absorbe las cargas a las que se somete el hueso, dejando el hueso descargado. Esta falta de estímulos mecánicos en el hueso provoca un proceso degenerativo conocido como reabsorción ósea. Para mantener sus propiedades mecánicas, el hueso necesita un cierto nivel de estímulo mecánico. Se pretende analizar entre varios diseños de prótesis de rodilla de revisión (prótesis de vainas, prótesis en offset sin suplemento, prótesis en offset con suplemento y prótesis con vástago recto) el más adecuado para el paciente.

1.4. Descripción del trabajo

Este Trabajo Fin de Máster está dividido en 4 capítulos, seguidamente se realizará una breve descripción de cada uno de ellos.

En el **Capítulo 1** hace una breve introducción a la biomecánica, la anatomía de la rodilla y los tipos de prótesis estudiados en este trabajo.

En el **Capítulo 2** se realiza una breve explicación de qué son las redes neuronales artificiales, el procedimiento llevado a cabo para la determinación de las cargas en la tibia de un paciente concreto y los resultados y conclusiones obtenidas del mismo.

En el **Capítulo 3** se explica la metodología seguida para el análisis de remodelación ósea tras la incorporación de una prótesis de revisión, utilizando las cargas obtenidas en el capítulo 2. De este procedimiento se analizará la influencia de cada tipo de prótesis en la reabsorción ósea, con su consiguiente aplicación a la mejora de los diseños aplicados a un paciente concreto.

En el **Capítulo 4** se detallan los resultados y conclusiones finales de los puntos anteriormente citados.

2. Redes neuronales aplicadas a la determinación de cargas en la tibia de un paciente concreto

2.1. Redes neuronales: Introducción

Las redes neuronales artificiales (RNA) son modelos matemáticos que intentan reproducir el funcionamiento del sistema nervioso. Como todo modelo realizan una simplificación del sistema real que simula y toma las características principales del mismo para la resolución de una tarea determinada.

La neurona artificial es un elemento de procesamiento simple que a partir de un vector de entradas produce múltiples salidas. En general podemos encontrar tres tipos de neuronas artificiales: de entrada, ocultas y de salida.

Una vez definida el tipo de neurona que se utilizará en un modelo de redes neuronales artificiales es necesario definir la topología de la misma. La organización y disposición de las neuronas dentro de una red neuronal se denomina topología, y viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad, y el tipo de conexión entre neuronas.

Una red puede estar formada por una única capa de neuronas. En este caso hablamos de redes monocapa, y las neuronas que conforman dicha capa cumplen la función de neuronas de entrada y salida simultáneamente. Cuando la red está compuesta por dos o más capas hablamos de redes multicapa. A su vez, hablamos de redes neuronales con conexión hacia delante (redes feedforward) cuando las conexiones entre las distintas neuronas de la red siguen un único sentido, desde la entrada de la red hacia la salida de la misma. Cuando las conexiones pueden ser tanto hacia delante como hacia atrás hablamos de redes recurrentes (redes feedback).

Durante la operatoria de una red neuronal podemos distinguir claramente dos fases o modos de operación: la fase de aprendizaje o entrenamiento, y la fase de operación o test. Durante la primera fase, la fase de aprendizaje, la red es entrenada para realizar un determinado tipo de procesamiento. Una vez alcanzado un nivel de entrenamiento adecuado, se pasa a la fase de operación, donde la red es utilizada para llevar a cabo la tarea para la cual fue entrenada.

El algoritmo de entrenamiento/aprendizaje supervisado más común utilizado para una MLP (multilayer perceptron), es el algoritmo llamado Backpropagation. Este algoritmo surgió de la búsqueda de un algoritmo para entrenar específicamente a la MLP. Es un tanto complejo ya que se toma la información del comportamiento de la red en el sentido directo de la red y en el sentido inverso, esto se realiza por la necesidad de modificar el comportamiento de las capas ocultas.

El algoritmo Backpropagation tiene como objetivo usar la diferencia entre las salidas deseadas y las salidas actuales en la capa de salida de la red para cambiar los pesos (indicados con valores aleatorios pequeños) con el fin de reducir al mínimo esta diferencia (error). Esto se logra mediante una serie de interacciones donde se modifica cada peso de derecha a izquierda (sentido inverso de la propagación de información en la red) hasta modificarse los pesos de la capa de entrada prosiguiendo nuevamente con la propagación de la información de entrada, esto hasta que la diferencia entre la salida deseada y la obtenida en cada neurona de salida sea mínima.

Más detalle de las redes neuronales se encuentran en el Anexo I.

2.2. Planteamiento de la metodología desarrollada

En este trabajo se ha desarrollado una metodología que consiste en resolver dos problemas y comparar los resultados obtenidos; estos dos problemas son el problema de remodelación ósea y el problema inverso de remodelación ósea (*Figura 2.1*).

El problema de remodelación ósea se ha resuelto utilizando un modelo de elementos finitos 3D de una tibia real de un paciente concreto. Este modelo de remodelación ósea se ha calculado varias veces con el objetivo de obtener un amplio rango de densidades óseas debidas a la acción de diversas cargas. Esta variación de carga se ha realizado con el objetivo de suplir la carencia de datos, en referencia a diferentes casos de pacientes específicos. Por ello, se ha cogido una geometría real de un único paciente y se ha sometido a diferentes condiciones de carga, obteniéndose diversas distribuciones óseas.

Por otro lado, la resolución del problema inverso de remodelación ósea nos dará como datos de salida el módulo de carga y el porcentaje de dicha carga repartido en cada cóndilo que genera una determinada distribución de densidad. Para obtener estos datos de salida se hace necesaria la utilización de un modelo matemático. En este trabajo se ha utilizado el modelo de red neuronal artificial (RNA). Una vez establecido el modelo matemático a utilizar debemos determinar los parámetros de entrada, en nuestro caso corresponden a los elementos seleccionados de la malla 3D de elementos finitos de la tibia de mayor desviación típica de densidad, los cuales se han obtenido del problema de remodelación ósea anterior (*Figura 2.1*).

Para finalizar, se validarán los datos de carga obtenidos mediante las redes neuronales con las condiciones de carga introducidas en el modelo de remodelación ósea, las cuales se corresponden con los datos de carga del paciente. Dicho desarrollo del trabajo se expone en el siguiente esquema, ver *figura 2.1*.

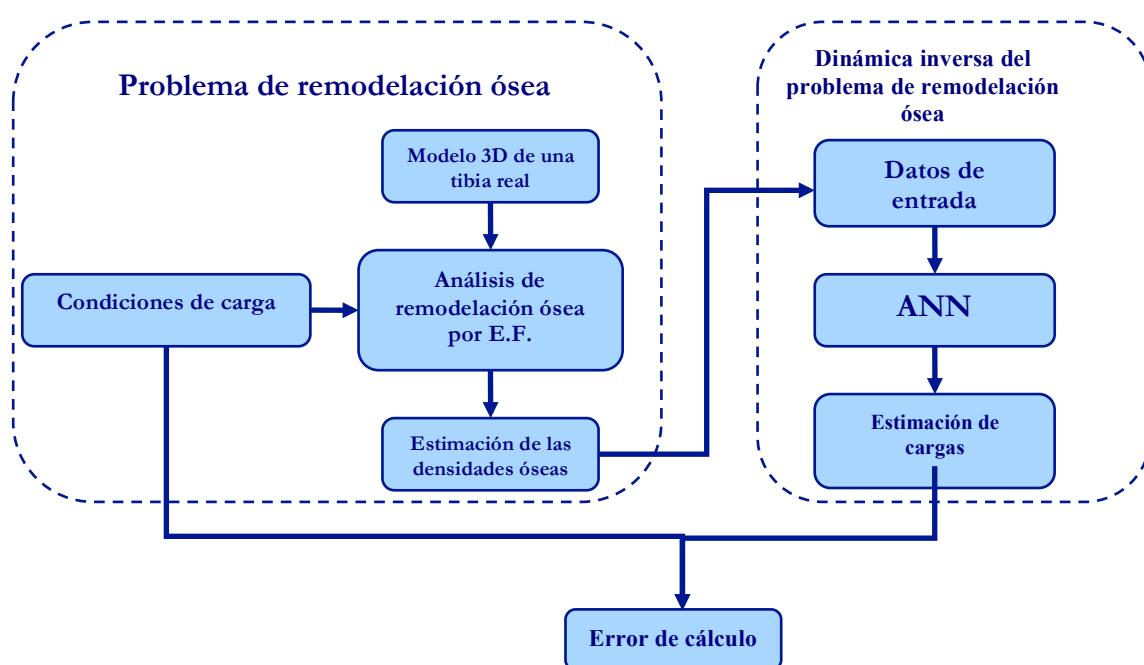


Figura 2.1. Esquema de la metodología desarrollada en el trabajo

2.3. Definición del problema de remodelación ósea

A continuación, se exponen en este apartado los pasos seguidos para la resolución del problema de remodelación ósea.

En primer lugar, es necesario tomar como referencia para la resolución del problema, un modelo de remodelación ósea que se aproxime al comportamiento real del hueso ante la acción de diversas cargas.

Una vez seleccionado el modelo de remodelación ósea más adecuado, se ha realizado un modelo 3D de elementos finitos de la tibia del paciente. Posteriormente, y a fin de simular diferentes casos del paciente, se ha sometido al modelo 3D de la tibia a diferentes condiciones de carga. Al final como resultado del problema de remodelación ósea obtendremos una serie de distribuciones de densidad ósea.

En los siguientes apartados se expone detalladamente los pasos seguidos en el problema de remodelación ósea, hasta la consecución de las citadas distribuciones de densidad ósea en la tibia del paciente.

2.3.1. Modelo de remodelación ósea

El hueso es un material poroso, heterogéneo y anisótropo. Además es un tejido vivo en constante cambio. Estos cambios que se producen en el hueso son importantes cambios morfológicos en respuesta a la situación mecánica a la que se encuentra solicitado, modificando su geometría externa y su estructura interna. Este proceso adaptativo se conoce con el nombre de remodelación ósea adaptativa. En concreto este proceso de remodelación puede ser externo, en que la forma geométrica externa cambia con el tiempo, o interno, en el que las propiedades del material cambian con el tiempo sin alterar su forma.

Estos procesos suceden simultáneamente. Para el desarrollo de este trabajo se ha utilizado un modelo anisótropo de remodelación ósea desarrollado por **Doblaré y García [2002]**.

Doblaré y García [2002] plantean la utilización de los conceptos de la Mecánica del Daño anisótropo como marco general para la formulación del problema de remodelación ósea. Es de destacar que en este modelo el estímulo no es una medida del microdaño, sino que el estímulo vuelve a ser la medida del estado de las deformaciones locales e instantáneas del hueso. Por tanto, cuando se habla de *variables de daño* habrá que entender mejor *variables internas microestructurales* sin identificación con ningún daño real.

Este modelo de remodelación se basa en los principios de la Mecánica del Daño Continuo utilizando la teoría de daño como herramienta matemática para la formulación de la capacidad de la remodelación ósea. Este modelo es capaz de predecir la distribución ósea anisotrópica y no homogénea en el hueso.

Las variables independientes internas son las que definen la microestructura del hueso, la densidad aparente y el “*fabric tensor*”. La densidad aparente cuantifica el grado de porosidad del hueso y el “*fabric tensor*” que mide el grado de anisotropía de la estructura ósea. El “*fabric tensor*” se demuestra que está alineado con el tensor de elasticidad, mientras que utilizando los principios de la teoría respecto a las variables internas queda demostrado que las direcciones de ortropía coinciden con los ejes principales de daño.

El modelo de remodelación ósea está implementado en una subrutina de Abaqus.

2.3.2. Modelo 3D de elementos finitos de la tibia

El método de elementos finitos (MEF) es un método numérico general para la aproximación de soluciones de ecuaciones diferenciales muy utilizado en diversos problemas de ingeniería y física.

El método de elementos finitos es muy usado debido a su generalidad y facilidad de introducir dominios de cálculo complejos (en dos o tres dimensiones). Dada la imposibilidad práctica de encontrar la solución analítica de estos problemas, con frecuencia en la práctica ingenieril los métodos numéricos, y en particular los elementos finitos se convierten en la única alternativa práctica de cálculo.

Para la obtención del modelo 3D de elementos finitos de la tibia se han seguido varios pasos, los cuales se exponen a continuación, *ver figura 2.2*:

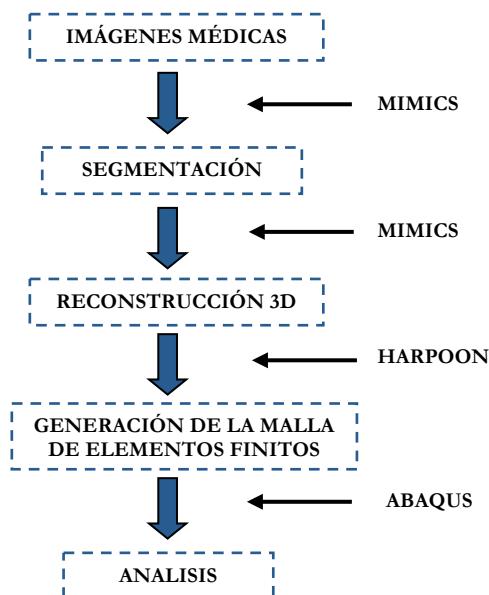


Figura 2.2. Proceso realizado para la obtención del modelo 3D

En primer paso del proceso realizado, consiste en la reconstrucción geométrica de la tibia, la cual se ha realizado mediante el programa comercial Mimics. Se ha partido de imágenes médicas de una tibia humana de un varón. Dichas imágenes médicas utilizadas en este trabajo para la reconstrucción de la tibia, fueron obtenidas mediante Tomografía Axial Computarizada (TAC).

Una vez con las imágenes en el programa Mimics, se inició la reconstrucción de la geometría. En primer lugar se realizó una segmentación automática a partir de la escala de grises. A partir de esta segmentación automática, las imágenes se fueron retocando manualmente. En función de la intensidad de dicha escala se ha podido diferenciar las diferentes partes de la tibia, esto quiere decir que, según aumenta el nivel de grises aumenta la densidad del material. A continuación en la *figura 2.3* se puede observar la imagen del proceso de reconstrucción con Mimics.

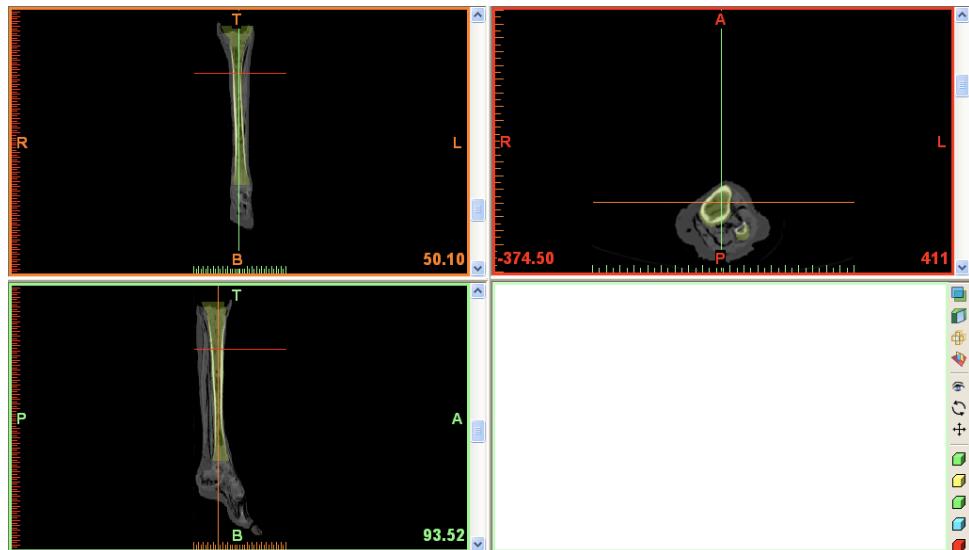


Figura 2.3. Segmentación y proceso de reconstrucción de la tibia mediante Mimics

Una vez realizada la reconstrucción de la tibia, se ha realizado una malla del volumen generado de la geometría. Para realizar la malla de la tibia, se ha utilizado el programa comercial Harpoon que nos ha generado dicha malla de forma automática.

Para la construcción del modelo tridimensional de la tibia generado en Mimics, se han exportado los ficheros en formato “stl”, para posteriormente ser importados al programa Harpoon. Cabe destacar, que este programa está especialmente indicado para mallar diferentes volúmenes de forma automática con diferentes niveles de exactitud.

Para el mallado de la tibia se utilizaron elementos tetraédricos C3D4 debido a su mejor adaptabilidad a la geometría, estos elementos se obtuvieron de la librería de elementos de Abaqus. Una vez obtenida la malla de toda la tibia se obtuvo en todo el modelo de elementos finitos completo, un total de 54047 elementos y 11074 nodos. Cabe destacar, que el tamaño medio de la malla es de 2 mm, previamente se realizó un estudio de sensibilidad de la malla para establecer el tamaño más adecuado de la misma. En la figura 2.4 se puede observar la tibia mallada.

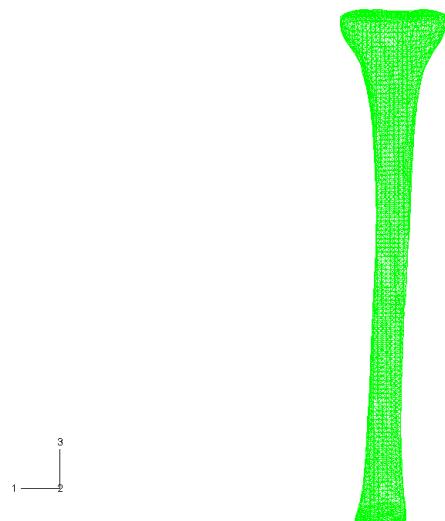


Figura 2.4. Tibia mallada

2.3.3. Condiciones de carga del modelo

Como condiciones de contorno se ha restringido el movimiento de la parte inferior de la diáfisis a lo largo de las direcciones vertical y horizontal en cuatro nodos. Se han considerado tres casos de carga correspondientes a tres etapas de la marcha. Las cargas consideradas son las fuerzas de reacción que se producen en las superficies condilares (*Figura 2.5 y Tabla 2.1*).

El primer caso es el momento en el que el pie se encuentra tocando el suelo y la carga se distribuye uniformemente entre los dos cóndilos y de forma normal a los mismos; el segundo y tercer caso son los momentos de abducción y aducción, en los cuales se distribuye la carga en un 70 y 30%, y un 30 y 70%, respectivamente, en ambos casos con un ángulo de inclinación de 5º [Pérez et al 2010, Nyman et al 2004].

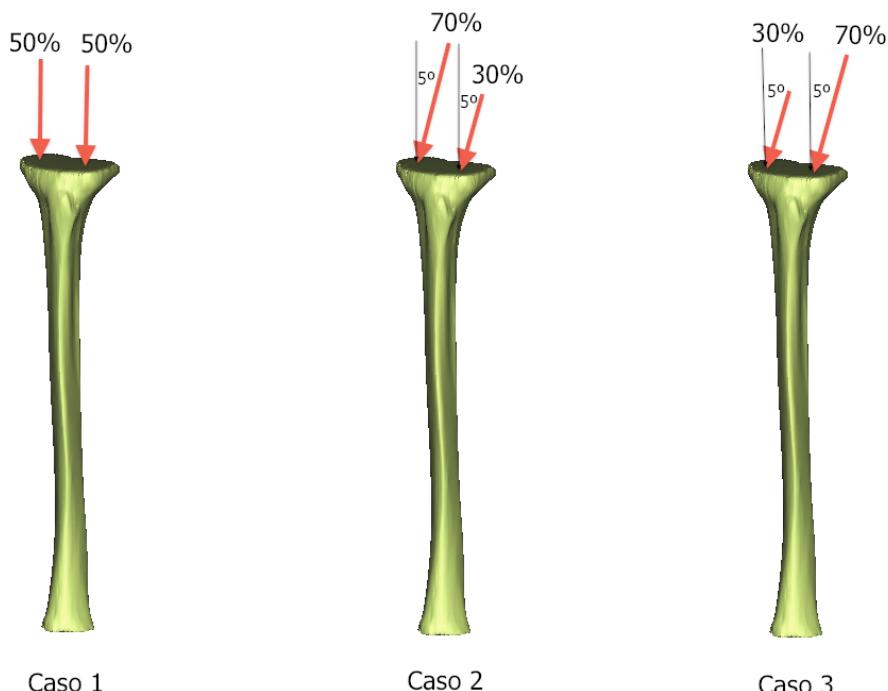


Figura 2.5. Modelo 3D de elementos finitos de la tibia con las condiciones de carga aplicadas. Caso 1: Pie entra en contacto con el suelo de forma completa. Caso 2: Momento de abducción. Caso 3: Momento de aducción

Tomamos como punto de partida para la simulación del problema de remodelación ósea una situación arbitraria (comportamiento isotrópico y una densidad uniforme de 0.5gr/cm³). Para esta situación arbitraria, aplicamos una secuencia de cargas anteriormente explicada. Los cambios producidos por esta secuencia de cargas nos producirán cambios en la distribución ósea real de la tibia. Estos cambios son debidos a que durante la simulación la estructura interna del hueso (densidad y propiedades mecánicas), se adapta a la mecánica del estímulo. Al final del proceso el hueso posee una distribución de la densidad y del módulo de Young heterogénea.

Debido al hecho de que las cargas presentan variaciones durante el desarrollo de las actividades más habituales y ante la imposibilidad de obtener estos datos reales, se han tomado un amplio rango de cargas. Esta variación se estima en un 20% para la magnitud de fuerza y de 70 y 30% hasta llegar a un 30 y 70% de la carga distribuida entre los cóndilos para el caso 2, y de 30 y 70% hasta llegar a 70 y 30% para el caso 3 [Bergmann et al 2001, Heller et al 2001]. Se han simulado 500 combinaciones de carga dentro de los rangos explicados anteriormente.

Case	Cycles/day	Medial condyle		Lateral condyle	
		X-axis	Z-axis	X-axis	Z-axis
Load values for the tibia: applied forces at the medial and lateral condyles					
1	3000	0.0	-1062.08	0.0	-1062.08
2	500	-129.6	-1353.28	-55.68	-634.88
3	500	55.68	-634.88	129.6	-1353.28

Tabla 2.1. Condiciones de carga de la tibia

2.4. Definición del problema inverso de remodelación ósea.

El problema inverso de remodelación ósea tiene como objetivo predecir las cargas musculoesqueléticas (módulo de carga y porcentaje de carga repartido a cada cóndilo) producidas en la tibia. Estas cargas representan las salidas del problema inverso de remodelación ósea.

Para obtener estos datos de salida es necesario la utilización de un modelo matemático. En este trabajo y como ya se explicó en capítulos anteriores se ha utilizado el modelo de red neuronal artificial (RNA). Una vez establecido el modelo matemático a utilizar debemos determinar los parámetros de entrada. Para garantizar el correcto funcionamiento de la red neuronal debemos someterla a un entrenamiento el cual nos permitirá seleccionar la opción más adecuada para nuestra red neuronal, de esta forma, nos garantiza el menor error posible. Para finalizar, una vez modelada de forma óptima nuestra red neuronal y predichas las cargas de salida, se procede a comparar los datos de salida obtenidos mediante la red neuronal con los datos del paciente.

A continuación se exponen más detalladamente las entradas y salidas del problema inverso de remodelación ósea, así como las características de la red neuronal utilizada en este trabajo.

2.4.1. Entradas y salidas del problema inverso de remodelación ósea

Las entradas del problema inverso corresponden a una selección de puntos de la malla 3D de elementos finitos con mayor desviación de densidad. Para esta selección de puntos de la malla se han utilizado técnicas estadísticas basadas en la media y desviación típica a fin de reducir el coste computacional de la red neuronal.

En un primer lugar se generó la media y la desviación típica de densidad de los 54047 elementos de la malla 3D de elementos finitos de la tibia a partir de los 500 casos de carga analizados (sección 2.3.3). Con todas estas densidades se seleccionaron los 40 elementos con mayor desviación típica, los cuales se han utilizado como entradas para nuestra red neuronal. Cabe destacar que los 40 elementos seleccionados no corresponden exactamente a los elementos de mayor desviación ya que muchos de ellos se localizaron en la zona cortical de la tibia. Debido a este hecho, se seleccionaron otros con menor desviación típica de densidad a fin de conseguir una distribución más armoniosa de las entradas de la red neuronal. En la figura 2.6 se observa como en la zona del cortical es donde se ha encontrado mayor concentración de densidades con mayor desviación típica, y de cómo se han seleccionado los puntos de la forma más homogénea posible a lo largo de la geometría de la tibia.

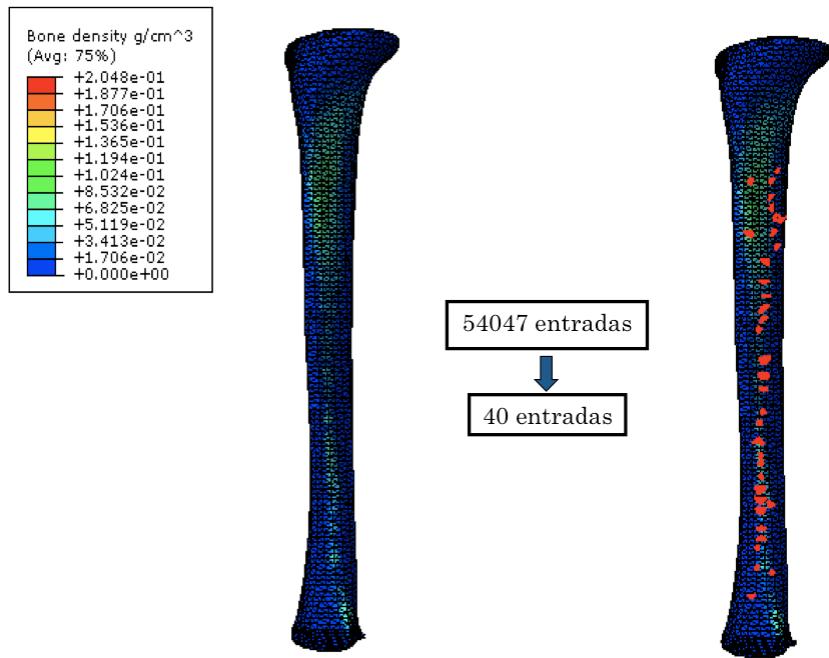


Figura 2.6. Distribución de desviaciones típicas obtenidas en la tibia para las condiciones de carga aplicadas:
a) Distribución de densidades de los 54047 elementos de la malla 3D de elementos finitos de la tibia. b) 40 elementos con mayor desviación típica, que representan las entradas para la red neuronal.

Las salidas del problema inverso, corresponden a las condiciones de carga (módulo de carga y porcentaje de carga repartido a cada cóndilo) proporcionadas por la red neuronal entrenada para una distribución de densidades de un paciente en concreto. El objetivo es que estos datos de salida proporcionados por la red neuronal sean similares a los proporcionados por el paciente concreto.

2.4.2. Características de la red neuronal utilizada

La red neuronal utilizada en este trabajo consta de 40 entradas, una capa oculta y dos salidas las cuales corresponden al módulo de la fuerza y el porcentaje de la misma repartido a cada cóndilo. Como función de entrada se ha utilizado una función sigmoidea y como función de salida se ha utilizado una función lineal. Para el entrenamiento de la red neuronal se ha utilizado el algoritmo Backpropagation (Anexo I).

Para la capa oculta se ha seleccionado entre diferentes opciones con neuronas intermedias (de 3 a 50 neuronas), la capa con un menor error relativo.

En próximos apartados observaremos que, a fin de obtener un entrenamiento más completo de la red neuronal, se obtuvieron varios modelos de entrenamiento, reduciéndose las entradas a 20.

2.5. Resultados

2.5.1. Evaluación de la red neuronal

Para evaluar correctamente nuestra red neuronal artificial hemos utilizado la técnica de validación cruzada. Consiste en dividir los datos de la muestra, en nuestro caso en dos subconjuntos: un subconjunto utilizado para entrenar al modelo (90% de los datos) y otro subconjunto utilizado para validar y testear el modelo (10% de los datos). En esta técnica aleatoria se han realizado 10 iteraciones y, para cada una de estas iteraciones, el subconjunto de validación ha ocupado una posición diferente a lo largo de los datos de la muestra.

La ventaja de la validación cruzada es que utiliza todos los datos disponibles de entrenamiento y el error final es independiente de los datos seleccionados en cada modelo. Con el fin de predecir la exactitud de nuestra red neuronal artificial se han utilizado los siguientes errores:

$$\text{Error relativo} = \text{abs} \frac{(x - y)}{y}$$

$$\text{Coeficiente de correlación}(r_{xy}) = \frac{\sigma_{xy}}{(\sigma_x \sigma_y)}$$

Donde x son los datos de predicción, y son los datos reales, σ_{xy} es la covarianza y σ_x y σ_y son las desviaciones típicas de las variables x e y .

Los resultados obtenidos para el error relativo se calculan en base a los subconjuntos de test y entrenamiento de la validación cruzada. Mediante el coeficiente de correlación medimos el grado de intensidad de la relación entre los datos de entrenamiento. Estos datos se utilizan para generar el modelo, de manera que el correspondiente error nos indicará si el modelo se ha entrenado correctamente. Por otro lado, los datos de test se utilizan para validar el modelo, de manera que el correspondiente error nos indica si el modelo es adecuado para simular el problema.

Antes de introducirnos en los diferentes modelos de entrenamiento utilizados para nuestra red neuronal y comentar los resultados obtenidos, es importante destacar los valores que aparecen en las gráficas. Estos valores corresponden a los errores de aprendizaje y a los errores de test.

Error de aprendizaje: corresponde a los errores obtenidos en función de los datos con que se entrena el modelo, es decir, los mismos valores utilizados para enseñar a la red neuronal.

Error de test: compara el error que se produce al introducir en nuestra red datos no utilizados anteriormente, comparándolo posteriormente con el obtenido para los datos reales.

2.5.2. Modelo de entrenamiento

Como modelo inicial de entrenamiento para nuestra red neuronal se han tomado 500 casos (para diferentes módulos de carga y porcentaje de la misma distribuida entre los cóndilos) y 40 elementos de entrada.

A continuación se exponen en la *Tabla 2.2* para diferentes números de neuronas en la capa intermedia, los errores obtenidos para el módulo de la carga, el porcentaje de la misma repartido entre los cóndilos y el error total obtenido, que es suma de los dos anteriores. También se muestra el coeficiente de correlación. Siempre que hablamos de error nos estaremos refiriendo a un error relativo, en este caso sería *error relativo de carga* y *error relativo del porcentaje de carga repartido entre los cóndilos*.

Error relativo (%) y coeficiente de correlación										
Neuronas capa intermedia										
		3	5	10	15	20	25	30	40	50
Error de carga (%)	Test	0,0291	0,0250	0,0354	0,0524	0,0530	0,0475	0,0418	0,0415	0,0404
	Aprendizaje	0,0210	0,0117	0,00339	3,34E-6	1,24E-5	4,13E-6	1,02E-6	3,78E-6	7,26E-6
	r _{xy}	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999
Error de porcentaje (%)	Test	2,6443	4,4203	2,5281	0,11777	0,10705	0,08199	0,05952	0,09258	0,08050
	Aprendizaje	4,3831	4,3831	2,3501	0,10589	0,10589	0,10589	0,10589	0,07243	0,07243
	r _{xy}	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999
Error total (%)	Test	2,67	4,45	2,56	0,17	0,16	0,13	0,101	0,134	0,121
	Aprendizaje	4,4	4,39	2,35	0,1059	0,1059	0,1059	0,1059	0,07244	0,07244
	r _{xy}	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999

Tabla 2.7. Error relativo y coeficiente de correlación para el modelo de entrenamiento de la red neuronal con 40 entradas

A continuación, en las *Figuras 2.7 y 2.8* se observa la evolución del error relativo obtenido respecto al número de neuronas de la capa intermedia, tanto para el módulo de la carga como para el porcentaje de carga repartido entre cóndilos, respectivamente.

En la *Figura 2.7* se muestran los errores de aprendizaje y de test obtenidos para el valor del módulo de la carga. Se observa que mientras el error de aprendizaje disminuye de forma continua según aumenta el número de neuronas, el error de test no sigue una trayectoria tan homogénea. El error disminuye hasta llegar al entrenamiento de capa intermedia con 5 neuronas, posteriormente el error va aumentando progresivamente hasta alcanzar las 20 neuronas y vuelve a disminuir hasta llegar al último punto de entrenamiento que corresponde a la capa intermedia con 50 neuronas. Cabe destacar que para 50 neuronas se alcanza un mínimo relativo, el cual será descartado debido a que es de valor superior al obtenido para una capa intermedia con 5 neuronas.

De esta forma y a la vista de los resultados del error de carga, se determina como solución óptima una capa intermedia con **5 neuronas**. Se ha considerado esta solución ya que

a partir de ella nos aparece el llamado sobreentrenamiento, el cual nos produce una pérdida de generalización por el uso excesivo de neuronas en la capa intermedia de la red neuronal. Es decir, al aumentar el número de neuronas del modelo el resultado no mejora.

Error relativo del módulo de la carga

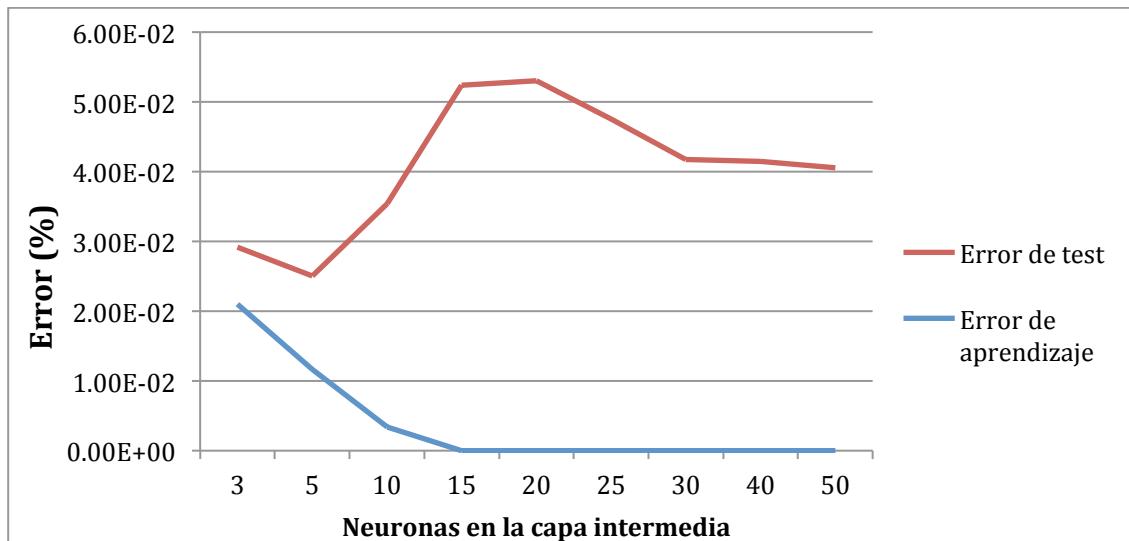


Figura 2.7. Error relativo del módulo de la carga obtenido para el modelo inicial de entrenamiento

En la *Figura 2.8* los valores representados corresponden a los errores obtenidos para el error relativo del porcentaje de carga repartido en cada cóndilo. Tanto el error de test como el error de aprendizaje tienen unos valores muy similares, los cuales disminuyen según incrementamos el número de neuronas. Estos errores son de un orden bastante superior a los obtenidos para el error relativo de carga. Gráficamente es difícil de determinar la solución que optimice el error, pero si nos dirigimos a la *Tabla 2.2* y a la *Figura 2.8* se observa que el menor error total (error de test) se consigue para una capa intermedia con 30 neuronas. De esta forma, seleccionaremos como solución óptima una capa intermedia con **30 neuronas**.

Error relativo del porcentaje de carga repartido a cada cóndilo

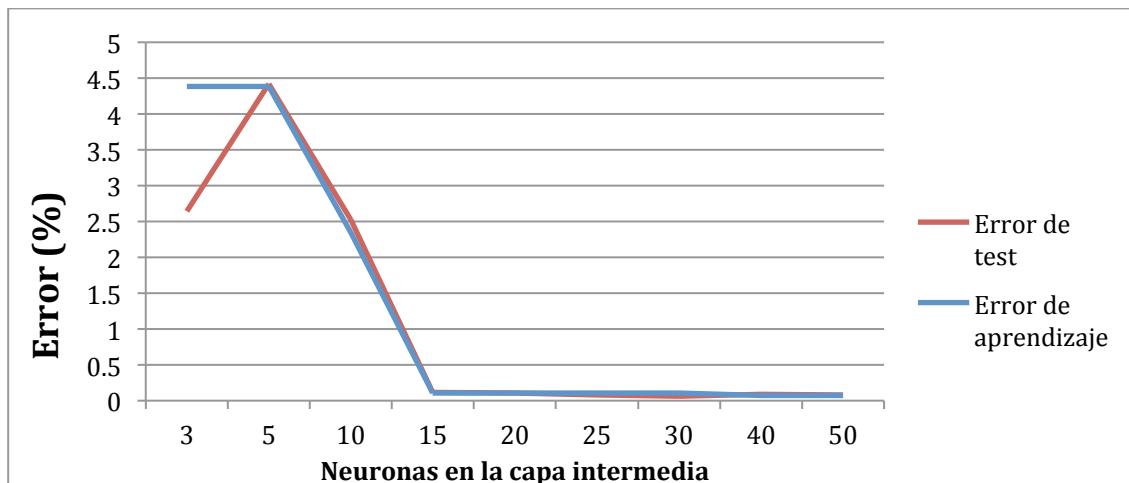


Figura 2.8. Error relativo del porcentaje de carga repartido entre cóndilos obtenido para el modelo inicial de entrenamiento

Para obtener una visión conjunta del modelo se ha englobado los errores obtenidos tanto para la carga como para el ángulo. En la *Figura 2.9* se muestra el error total producido en el modelo de entrenamiento para las diferentes capas intermedias. Los errores relativos obtenidos para el porcentaje de carga repartido entre cóndilos son de un orden mayor respecto a los errores relativos del módulo de la carga. De esta manera, será el error relativo obtenido para el porcentaje de carga repartido entre cóndilos el que tenga mayor peso a la hora de determinar el número óptimo de neuronas en la capa intermedia. Analizando los datos numérica y gráficamente, se ha determinado que el número de neuronas de la capa intermedia que optimizan el funcionamiento del modelo de entrenamiento es de **30 neuronas**.

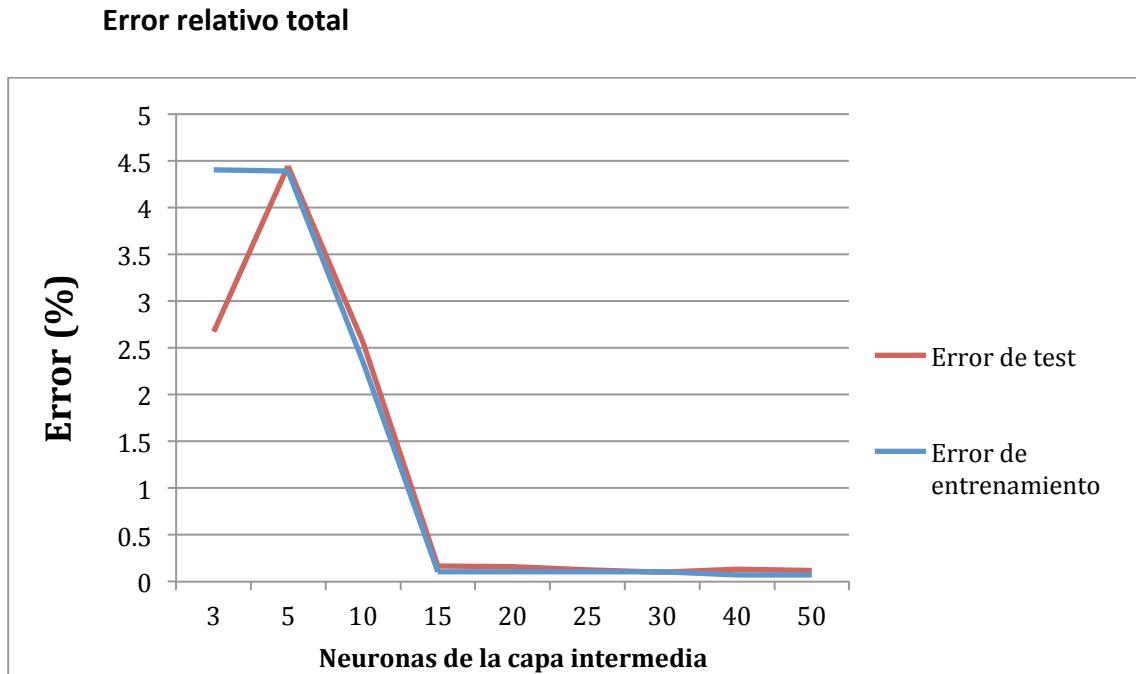


Figura 2.9. Error relativo total obtenido para el modelo inicial de entrenamiento

2.5.3. Reducción del número de entradas para la red neuronal

En un primer momento se utilizó una técnica estadística basada en la media y desviación típica para reducir el número de entradas (valores de densidad de los elementos de la tibia), ya que si tomamos todas las entradas el coste computacional que obtenemos es muy elevado (apartado 2.4.1). En un principio las entradas se redujeron de los 54047 puntos de integración conocidos del modelo de elementos finitos de la tibia a 40 entradas, se tomó como criterio de selección aquellos puntos de integración (entradas) con la desviación típica mayor.

Con el fin de reducir aún más el número de datos de entrada y por tanto reducir el coste computacional de entrenamiento de la red se realizó el siguiente estudio.

Para la reducción de los elementos de entrada a la red, se partió de los datos obtenidos para el modelo de entrenamiento anterior. Del error total se obtuvo como ya se explicó en apartado 2.5.2 una capa intermedia óptima para nuestra red neuronal con 30 neuronas. Para esta capa intermedia se evaluó la media de los pesos sinápticos de todas las entradas (30 neuronas, 40 elementos) y se seleccionaron los 20 elementos con mayor media en valor absoluto como nuevos elementos de entrada.

En la *Figura 2.10* podemos observar los valores medios en valor absoluto de los pesos sinápticos, obtenidos para los 40 elementos de entrada considerados inicialmente. Se seleccionaron los elementos con mayor media.

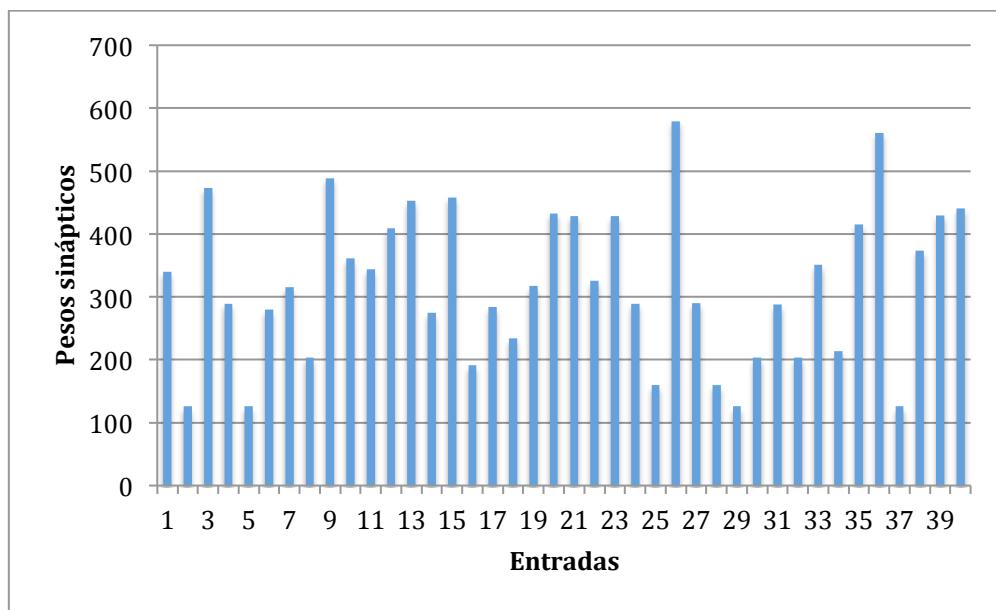


Figura 2.10. Valores medios en valor absoluto de los pesos sinápticos para las 40 entradas seleccionadas del modelo de elementos finitos de la tibia.

Cuando se obtuvieron los puntos con mayor desviación típica en la tibia la mayoría de ellos se localizaron en el borde de la zona cortical y del esponjoso, zona de mayor densidad ósea del modelo 3D (*Figura 2.11*). Con el fin de obtener un mejor estudio de la red neuronal se tomaron, a pesar de tener menor desviación típica, otros puntos de integración situados en zonas como las epífisis de la tibia. En las siguientes figuras (*Figura 2.11.a* y *Figura 2.11.b*) se observa la distribución de estos elementos de entrada a lo largo del modelo 3D de la tibia.

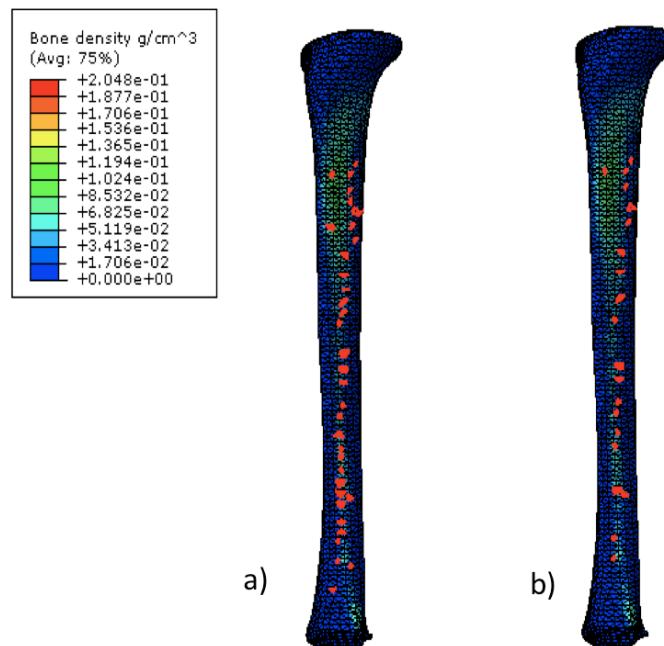


Figura 2.11. Distribución de desviaciones típicas obtenidas en la tibia para las condiciones de carga aplicadas:
a) 40 elementos con mayor desviación típica; b) 20 elementos con mayor desviación típica

A continuación se exponen los resultados obtenidos para el modelo de 500 casos con los 20 elementos de entrada seleccionados, ver *Figura 2.11.b*, con el fin de reducir el coste computacional de la red neuronal (*Tabla 2.3*). Podemos observar a simple vista que los errores obtenidos para el porcentaje de carga repartido a cada cóndilo (error más influyente a la hora de obtener el error total) al reducir el número de entradas, son mayores para los primeros casos de entrenamiento (hasta capa intermedia con 20 neuronas). Sin embargo, al llegar a los casos de 30, 40 y 50 neuronas en la capa intermedia se observa una disminución del error, llegando a ser prácticamente nulo (*Figuras 2.13, 2.14 y 2.15*).

Error relativo (%) y coeficiente de correlación										
Neuronas capa intermedia										
		3	5	10	15	20	25	30	40	50
Error de carga (%)	Test	0,0392	0,0313	0,0343	0,0466	0,0870	0,0632	0,0519	0,0535	0,0548
	Aprendizaje	0,0318	0,0205	0,0139	4,04E-3	3,52E-4	1,45E-4	4,60E-4	9,98E-7	1,49E-6
	r_{xy}	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999
Error de porcentaje (%)	Test	6,04	4,68	3,51	1,80	0,811	0,104	0,158	0,0454	0,0321
	Aprendizaje	4,38	4,38	3,47	3,47	0,792	0,792	0,792	0,792	0,792
	r_{xy}	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999
Error total (%)	Test	6,08	4,71	3,54	1,85	0,898	0,167	0,21	0,0988	0,0869
	Aprendizaje	4,41	4,40	3,48	3,47	0,793	0,793	0,793	0,792	0,792
	r_{xy}	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999	0,99999

Tabla 2.3. Error relativo y coeficiente de correlación para el modelo de entrenamiento de 20 entradas

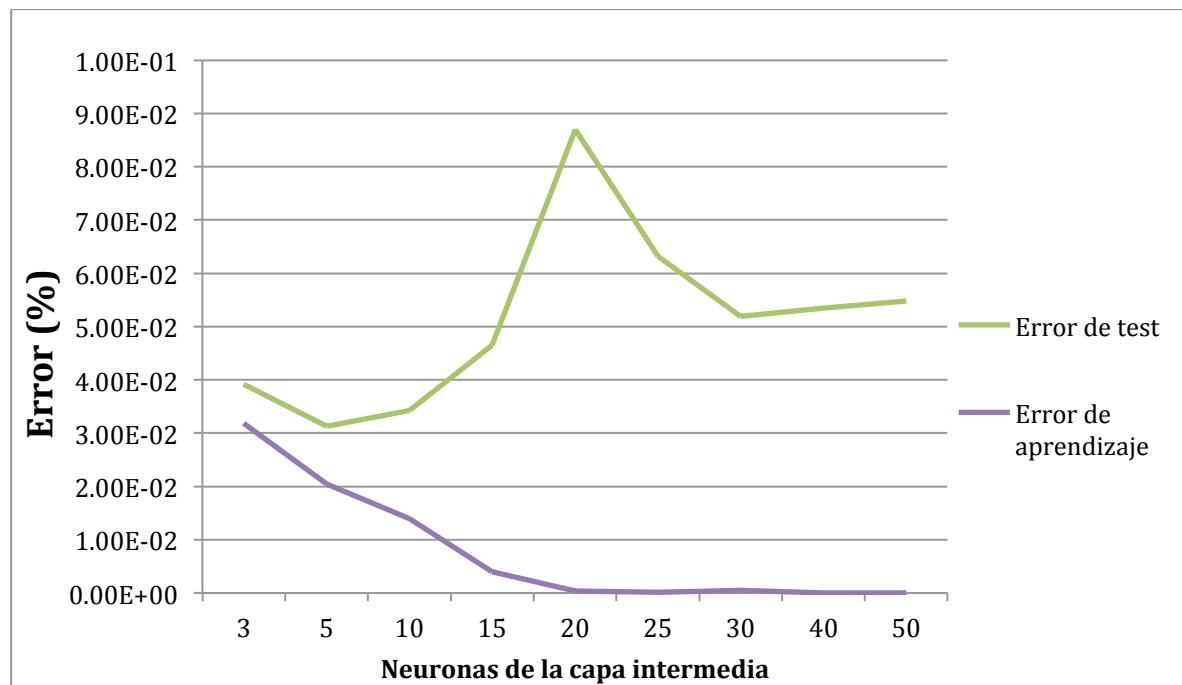


Figura 2.13. Error relativo del módulo de la carga obtenido para el modelo de entrenamiento con 20 entradas

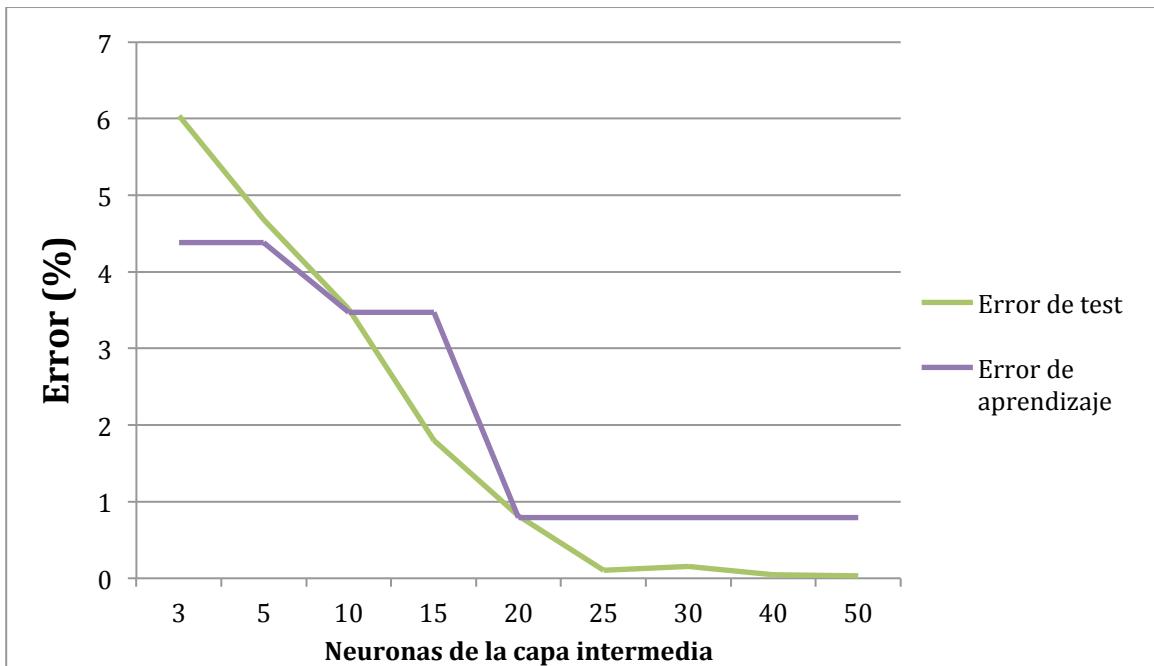


Figura 2.14. Error relativo del porcentaje de carga repartido entre cada cóndilo obtenido para el modelo de entrenamiento con 20 entradas

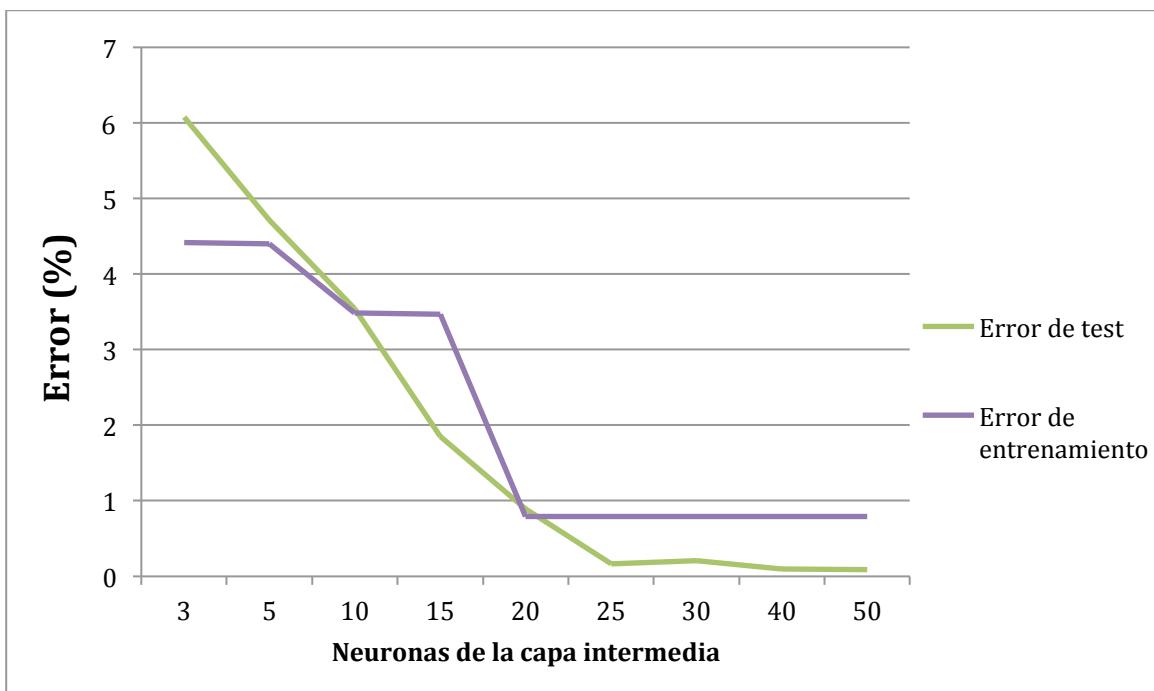


Figura 2.15. Error relativo total obtenido para el modelo de entrenamiento con 20 entradas

Al comparar los errores obtenidos con 40 y 20 entradas (*Figura 2.16*), observamos que al disminuir el número de entradas el error relativo total es ligeramente superior. En los modelos donde se ha utilizado el mayor número de neuronas esta diferencia se hace insignificante, ya que los errores de ambos modelos se aproximan a cero.

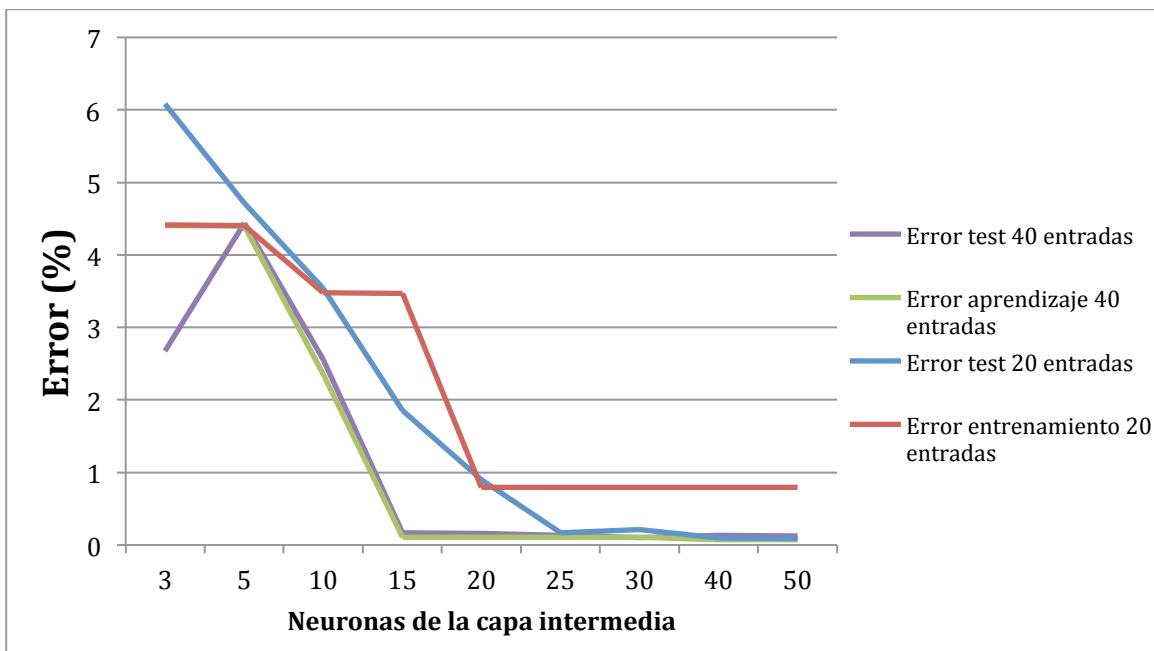


Figura 2.16. Error relativo total obtenido para los modelos de entrenamiento con 20 y 40 entradas

A continuación se exponen gráficamente los resultados obtenidos para el coste computacional de los diferentes modelos de entrenamiento, los cuales, nos ayudarán a determinar la solución óptima para nuestra red neuronal, ver *Figura 2.17*.

Se puede observar claramente que al aumentar el número de neuronas de la capa intermedia el tiempo computacional aumenta exponencialmente. Este aumento se produce de una forma más brusca para todos los modelos de entrenamiento, a partir de la capa intermedia con 15 neuronas. De esta forma, queda demostrado que el tiempo computacional aumenta al aumentar el número de neuronas de la capa intermedia.

Se determina claramente en la figura, cómo al reducir el número de entradas de la red neuronal, disminuye el tiempo computacional de entrenamiento de la red.

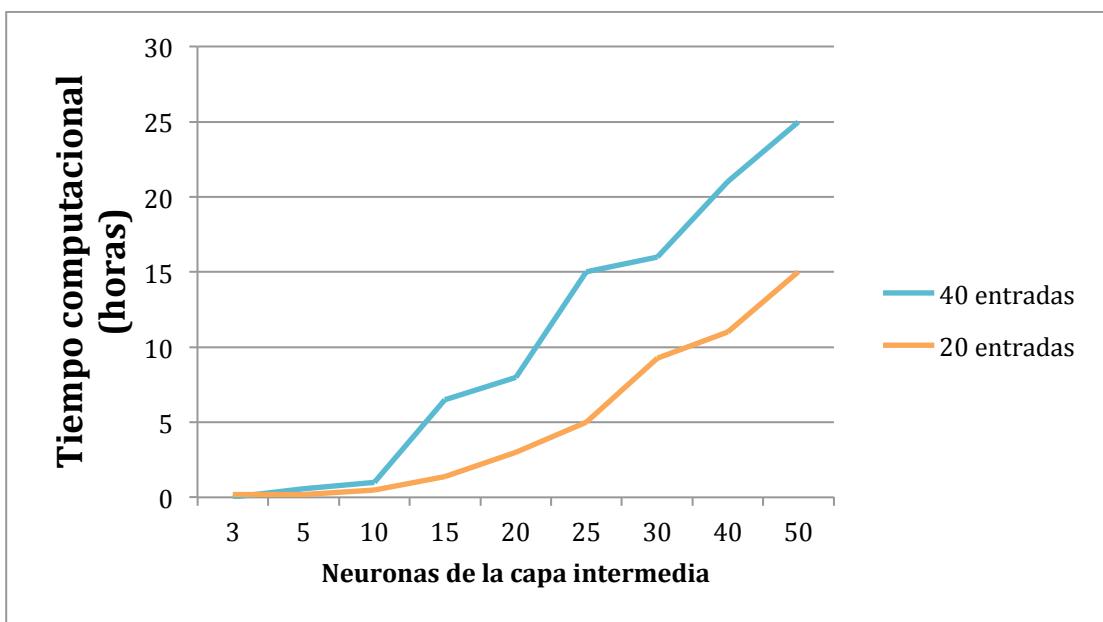


Figura 2.17. Tiempo computacional obtenido para los diferentes modelos de entrenamiento

2.5.4. Validación

En este apartado se va a proceder a realizar la validación de la metodología desarrollada utilizando la red con 20 entradas para predecir las cargas (módulo de la carga y porcentaje de la misma repartida entre los cóndilos) para un caso real de un paciente concreto.

El objetivo de esta validación es asegurarnos que la red neuronal seleccionada está correctamente entrenada. Con bien entrenada se quiere decir, que dicha red puede ser útil a la hora de predecir las cargas de determinados pacientes específicos, a partir de la distribución de densidades.

Para la realización de esta validación se han obtenido del TAC las densidades correspondientes a los 20 elementos de entrada de las redes a partir del nivel de grises de las imágenes (HU). Con los HU se ha obtenido el valor de la densidad en cada entrada de la red (punto de integración del elemento de la malla) utilizando la siguiente expresión propuesta por Peng et al (2008):

$$\rho = 1 + 7,185 \times 10^{-4} \text{ HU}$$

Una vez obtenidas se introducen en la red neuronal para obtener las cargas que han dado lugar a esas densidades. Por último se vuelven a introducir esas cargas en el modelo de la tibia, se realiza el análisis de remodelación ósea y comparamos las densidades resultantes con las obtenidas del TAC.

Al realizar una comparación cualitativa entre la distribución de densidades del TAC y obtenidos tras el análisis de remodelación ósea introduciendo las cargas de la red (*Figura 2.16*), observamos que obtenemos unos buenos resultados en la zona del cortical, pero obtenemos cierto error en las zonas de esponjoso, donde se aplicaron las cargas y las condiciones de apoyo.

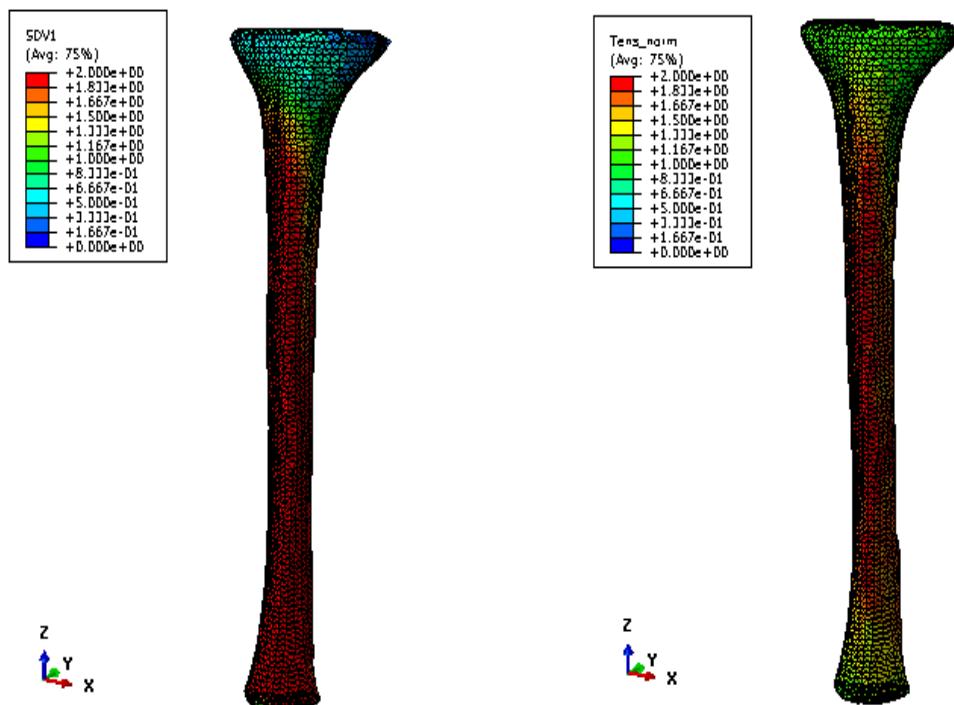


Figura 2.17. Comparación cualitativa de las densidades (gr/cm^3) obtenidas a partir de las HU del TAC y el modelo de redes neuronales

2.6. Conclusiones

Como se ha visto en apartados anteriores, cuantas más neuronas usemos para entrenar nuestra red mayor será el coste computacional y, a su vez, menor será el error, excepto si traspasamos el límite del sobreentrenamiento. Por otro lado, la disminución del número de entradas supone una disminución del coste computacional y un ligero aumento del error.

En función de los resultados obtenidos, tanto para el error relativo como para el tiempo computacional, podemos determinar como solución óptima para nuestra red neuronal, el modelo de entrenamiento con **20 entradas y una capa intermedia con 30 neuronas**.

El modelo de predicción de las cargas mediante el método de las redes neuronales es un método útil y preciso para las zonas con mayor densidad ósea. Sin embargo, cuando la densidad es baja o en las zonas donde se aplican condiciones de contorno el modelo presenta ciertos errores. Esta inexactitud puede ser debida a los fallos que presentan los modelos de remodelación ósea que, como se ha dicho anteriormente, producen errores en zonas trabeculares y donde se aplican las condiciones de contorno.

3. Diseño de una prótesis de rodilla de revisión

3.1. Introducción

Una vez obtenido un modelo capaz de predecir las cargas que dan lugar una determinada distribución de densidades, podemos aplicar este modelo al diseño de prótesis para un paciente concreto.

Partiendo de la tibia utilizada en el capítulo anterior y las cargas obtenidas en el apartado 2.5.4, se ha simulado la colocación de cuatro tipos de prótesis de rodilla de revisión con ayuda del software Mimics para estudiar su efecto sobre el hueso, en concreto sobre la reabsorción/formación ósea.

3.2. Modelos de elementos finitos y remodelación ósea

Para el estudio de la influencia de cada tipo de prótesis sobre la remodelación ósea de la tibia, se ha seguido el siguiente esquema (*Figura 3.1*):

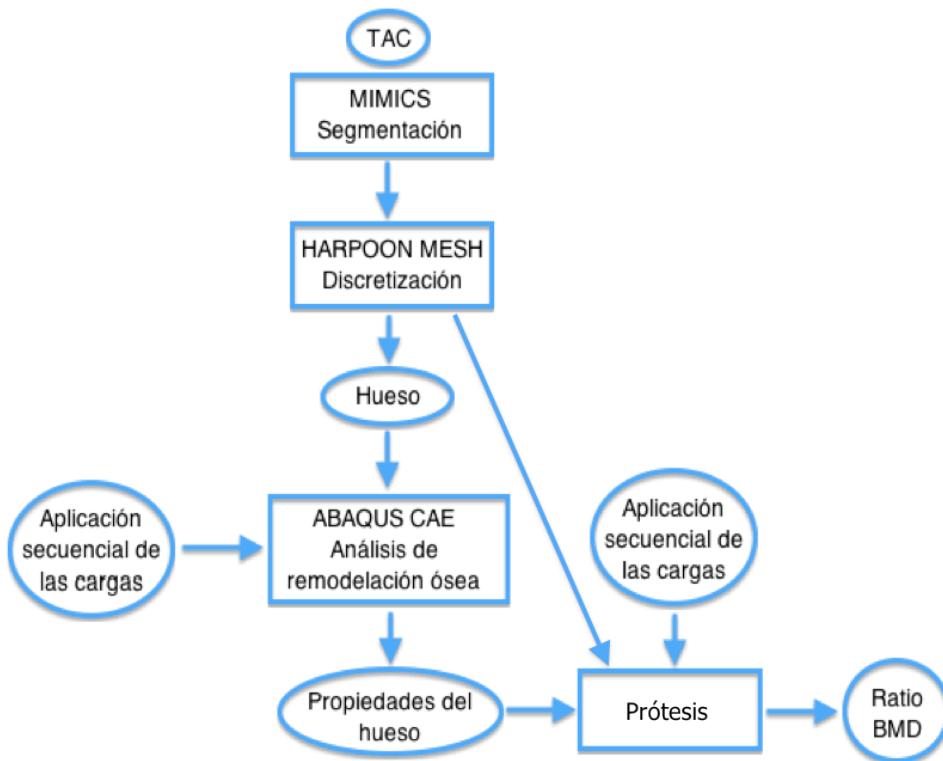


Figura 3.1. Esquema seguido para la realización del análisis de remodelación ósea

Se parte de las Tomografías Axiales Computerizadas (TAC) (*Figura 3.2*) tomadas a distintos pacientes a los que se le han colocado distintos tipos de prótesis de revisión. Se han importado las imágenes de cada una de ellas desde el programa Mimics y se ha desarrollado un modelo tridimensional de la componente tibial del implante, y que una vez suavizado (*Figura 3.3*), es introducido en la tibia sana utilizada en el capítulo anterior (*Figura 3.4*). Posteriormente es exportado al programa Harpoon Mesh para realizar un mallado (*Figura 3.5*). Finalmente se exporta al procesador de cálculo Abaqus CAE v. 6.11 (*Figura 3.6*) para realizar, mediante el método de los elementos finitos y el análisis de remodelación ósea.

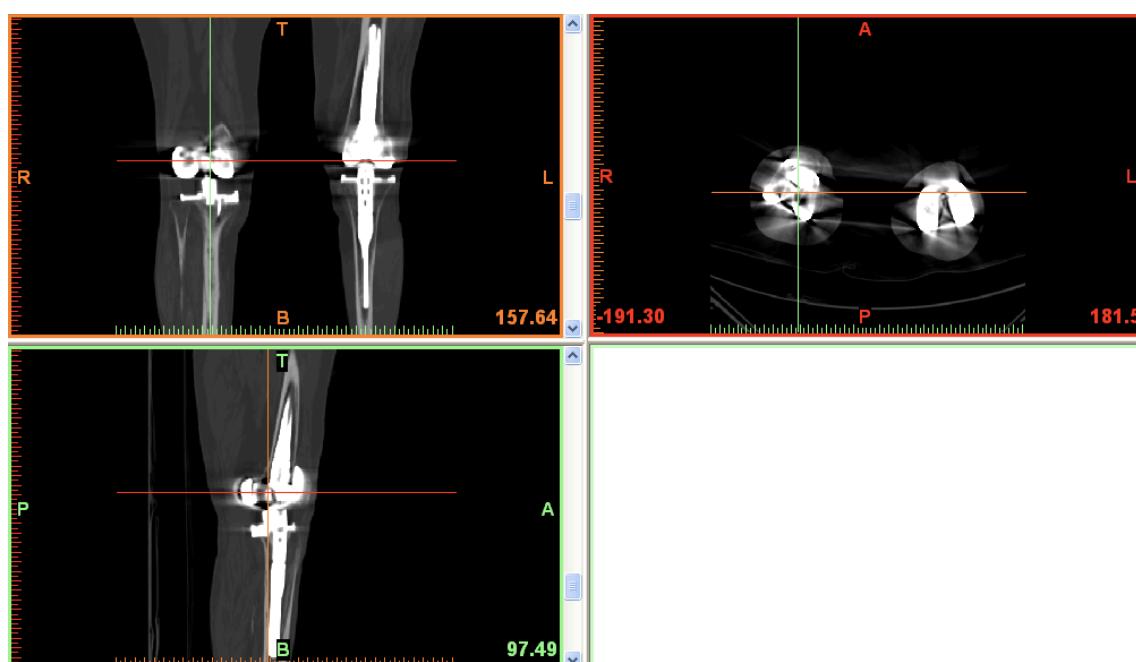


Figura 3.2. Tomografías Axiales Computarizadas en el software Mimics

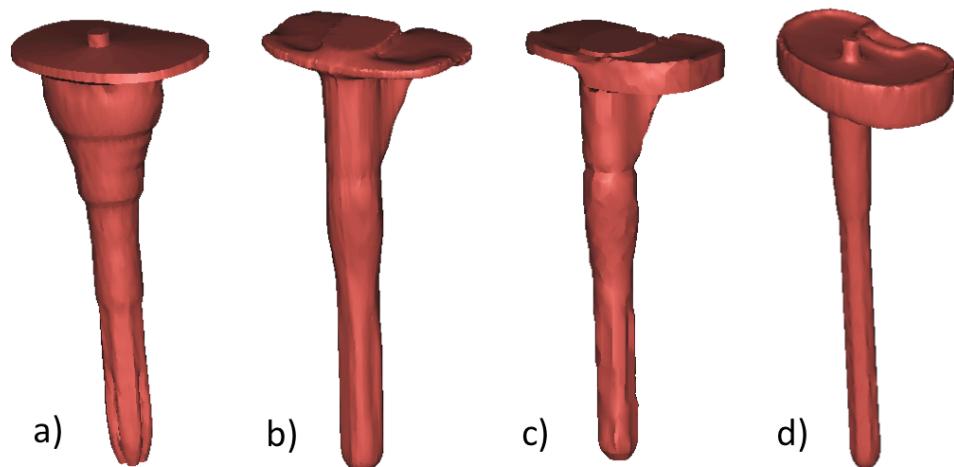


Figura 3.3. Modelos 3D de las componentes tibiales de las prótesis de rodilla de revisión: a) Prótesis de vainas. b) Prótesis en offset. c) Prótesis en offset con suplemento medial. d) Prótesis con vástago recto.

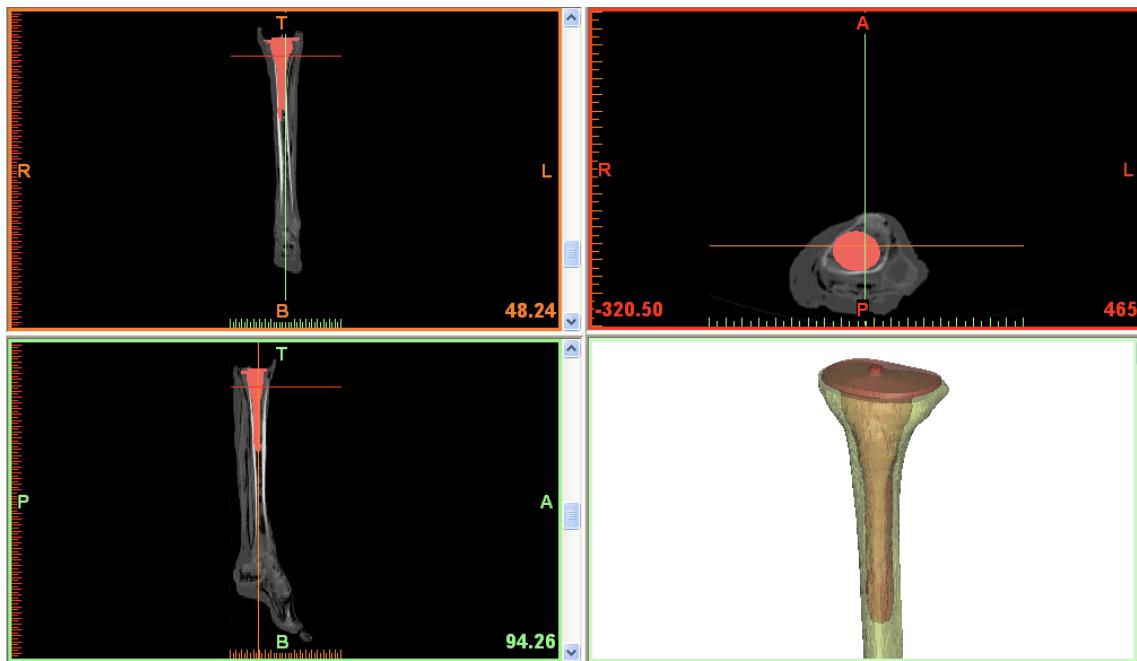


Figura 3.4. Prótesis introducida en la tibia sana

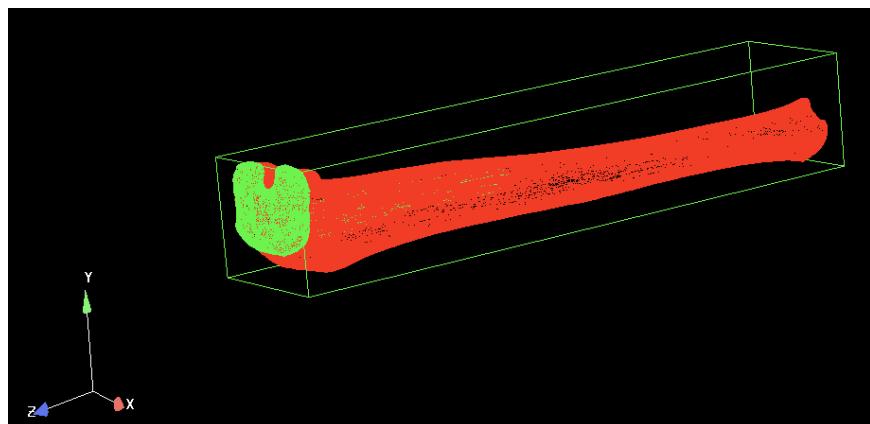


Figura 3.5. Modelos 3D de la tibia y la prótesis en offset malladas en Harpoon Mesh

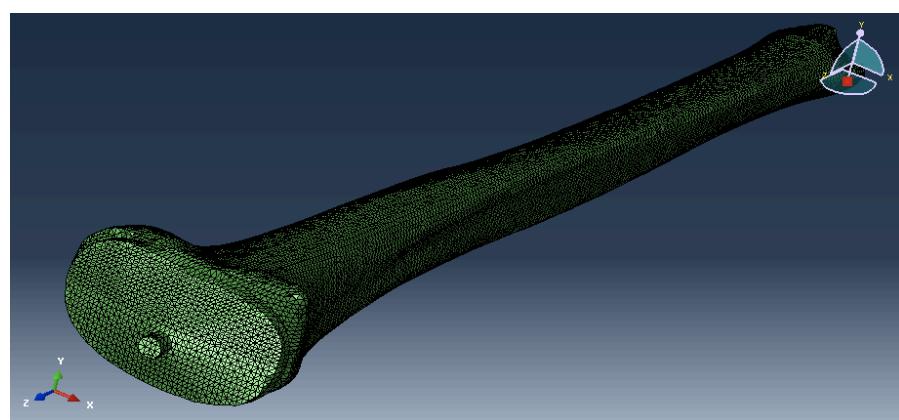


Figura 3.6. Modelos 3D de la tibia y la prótesis de vainas en Abaqus

La densidad ósea o BMD (bone mineral density) es un término médico que normalmente se refiere a la cantidad de materia mineral por centímetro cuadrado de los huesos. La densidad ósea se utiliza en la medicina clínica como un indicador indirecto de la osteoporosis y el riesgo de fractura. Se mide mediante un procedimiento llamado densitometría, el cual se realiza en el departamento de radiología o de medicina nuclear de los hospitales o clínicas. La medición es indolora y no invasiva e implica exposición a baja radiación.

El ratio de densidad ósea cuantifica cuánto cambia la densidad ósea a lo largo del tiempo; si desciende de forma considerable significa que la densidad ósea también ha descendido, lo que quiere decir que el hueso ha pasado de poseer mayor parte de hueso cortical (hueso más fuerte) a mayor parte de hueso trabecular (hueso esponjoso) debido al cambio en la transmisión de las cargas que supone la incorporación de la prótesis. Se puede determinar el ratio de densidad ósea mediante la siguiente expresión:

$$\text{BMD ratio} = \frac{\int \frac{\rho_i}{\rho_o} dV}{\sum dV} \times 100$$

Figura 3.7 Expresión utilizada para el cálculo del BMD ratio

El ratio de BMD se ha evaluado en las cuatro prótesis, así como la tensión principal máxima y la deformación equivalente de Von Mises.

3.3. Resultados

A continuación se muestran las distribuciones de densidades obtenidas al incorporar las distintas prótesis a la tibia (*Figuras 3.7 a 3.14*).

En todos los casos se observa un mismo patrón. La densidad en la epífisis va disminuyendo, es decir, se produce reabsorción ósea, al igual que en el canal medular en torno al vástago protésico. Por otro lado, se produce un aumento de la densidad, es decir, formación ósea, en la zona de la punta del vástago y en la diáfisis.

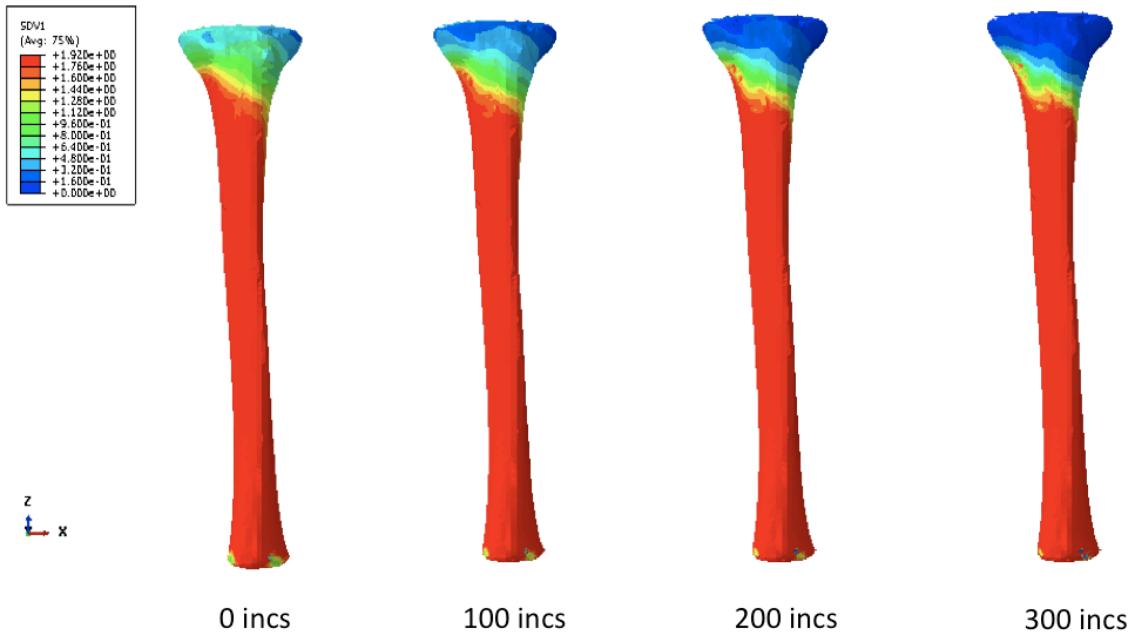


Figura 3.7. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis con vástago en offset sin suplemento

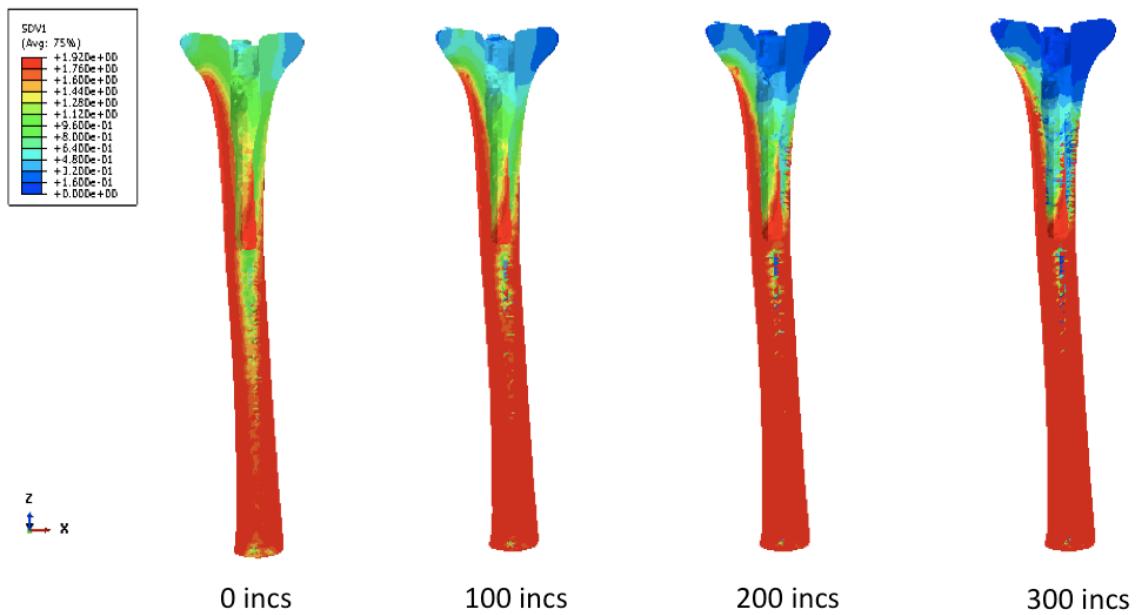


Figura 3.8. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis con vástago en offset sin suplemento. Vista corte transversal

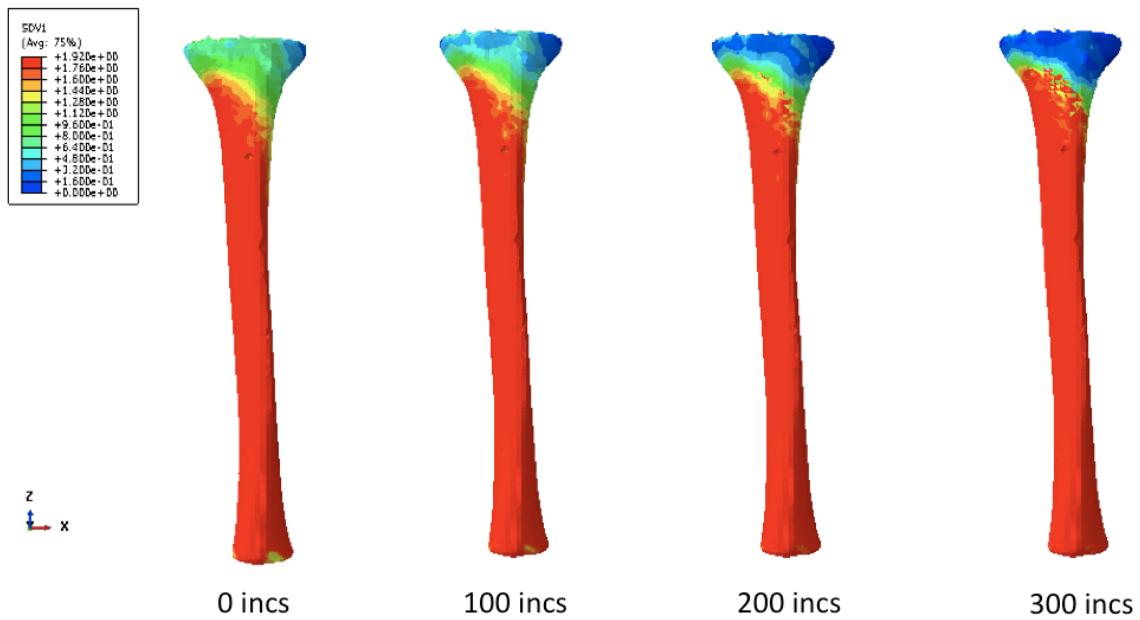
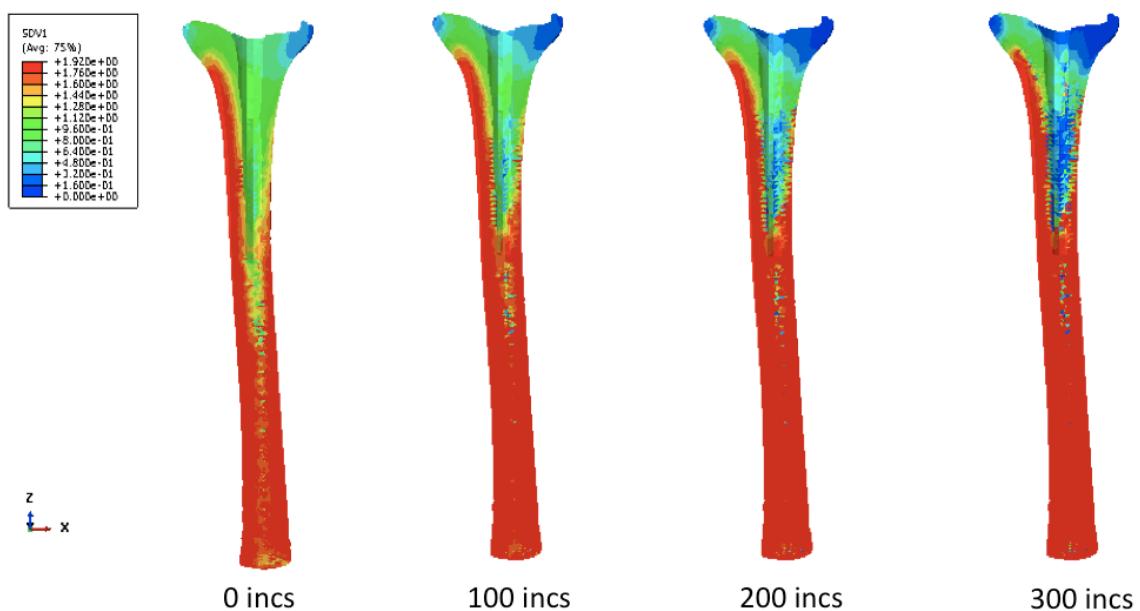


Figura 3.9. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis con vástago recto



**Figura 3.10. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis con vástago recto.
Vista corte transversal**

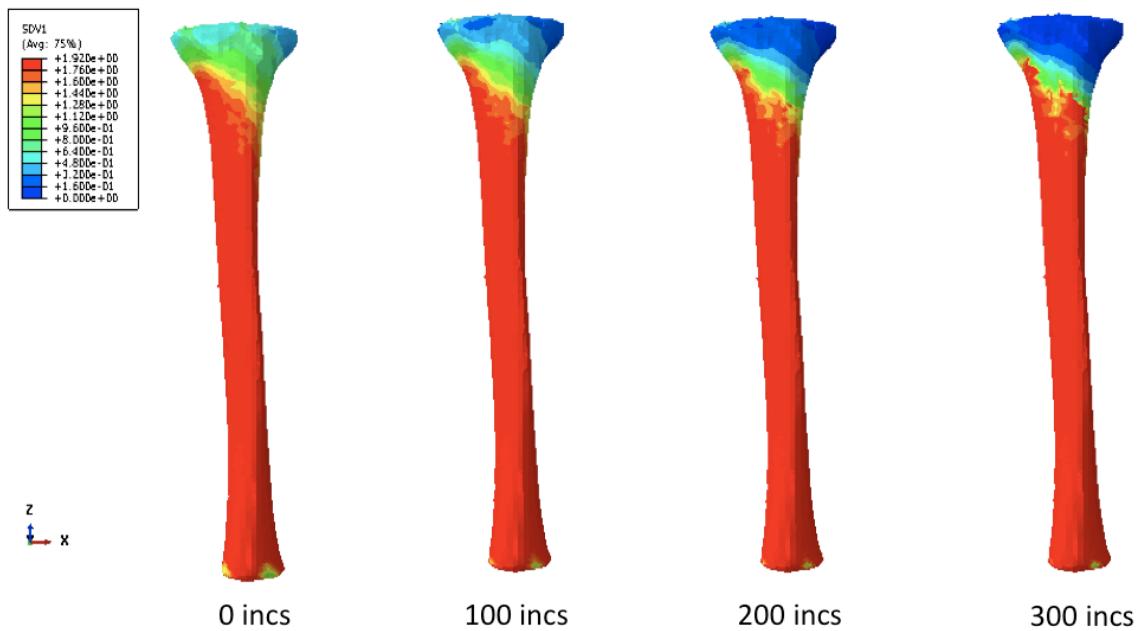


Figura 3.11. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis con vástago en offset con suplemento

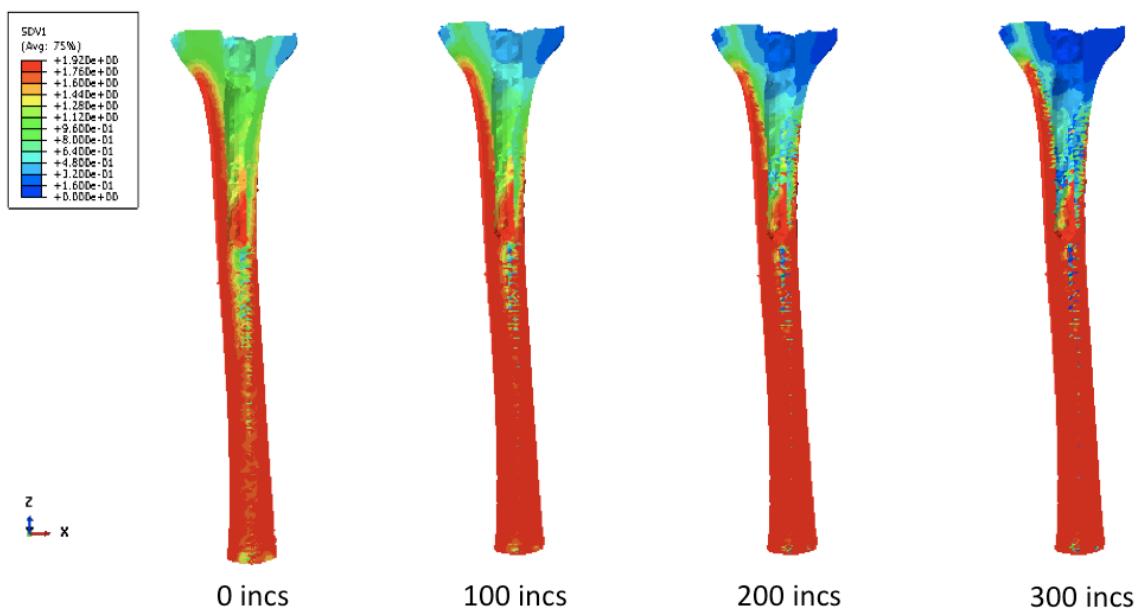


Figura 3.12. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis con vástago en offset con suplemento. Vista corte transversal

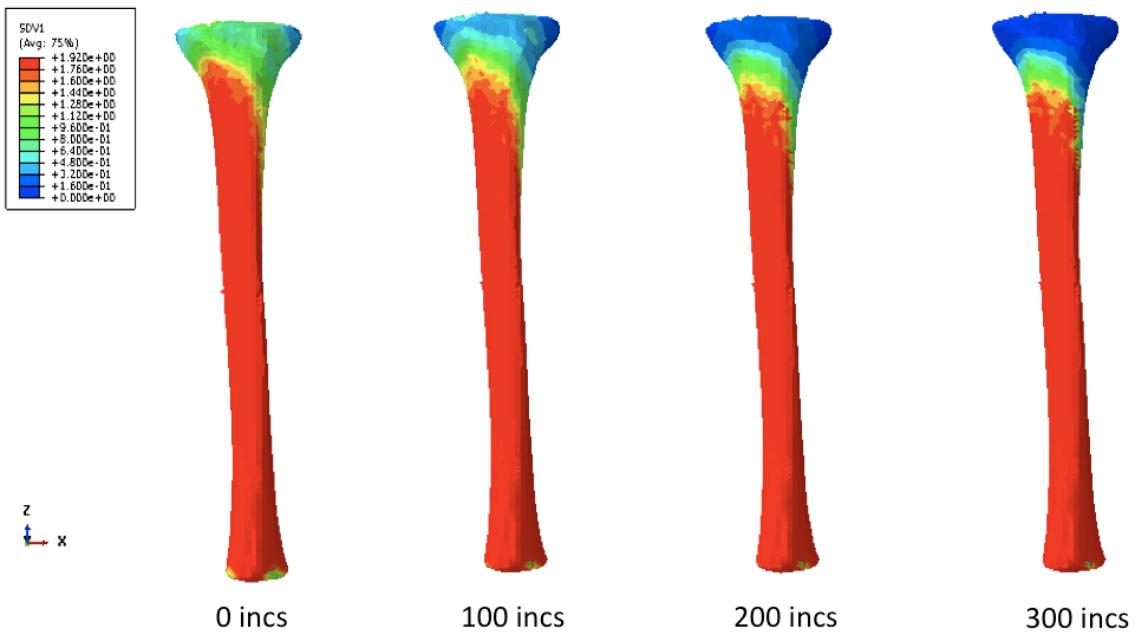


Figura 3.13. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis de vainas

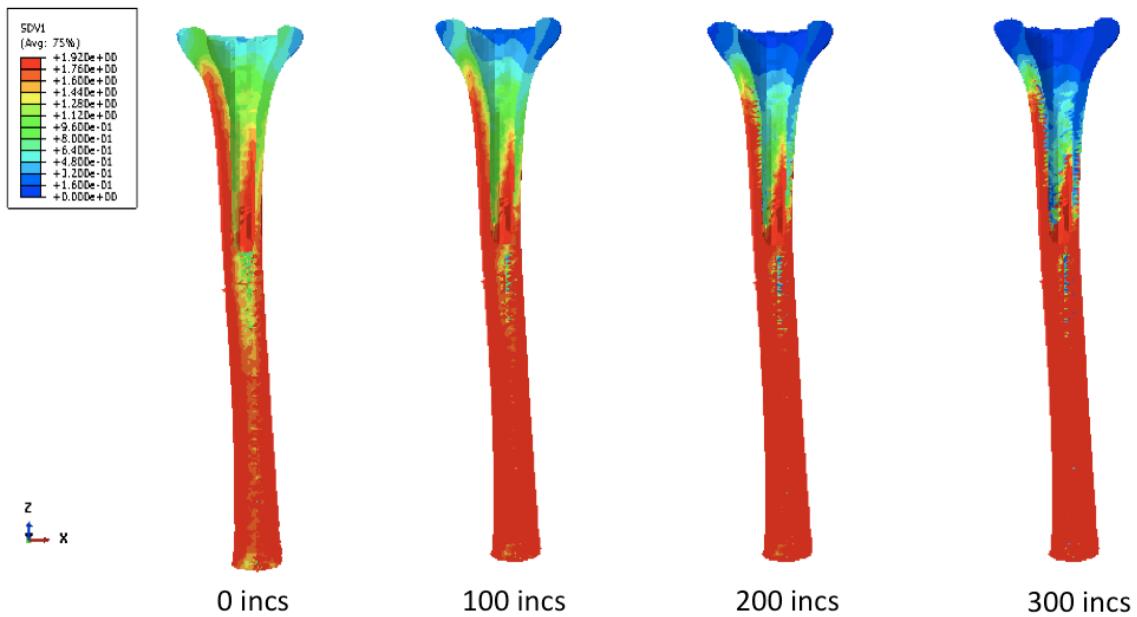


Figura 3.14. Evolución de la densidad ósea (gr/cm^3) a lo largo del tiempo para la prótesis de vainas. Vista corte transversal

En la *figura 3.15* se observan las diferencias entre las distribuciones de densidad que han dado lugar la implantación de las distintas prótesis. Se observa que la prótesis con vástago recto produce una gran reabsorción ósea en la zona del vástago tibial comparada con el resto, mientras que en la zona de la diáfisis produce menos reabsorción que las demás, seguida de la prótesis en offset con suplemento. La prótesis en offset sin suplemento y la prótesis de vainas dan lugar a densidades similares.

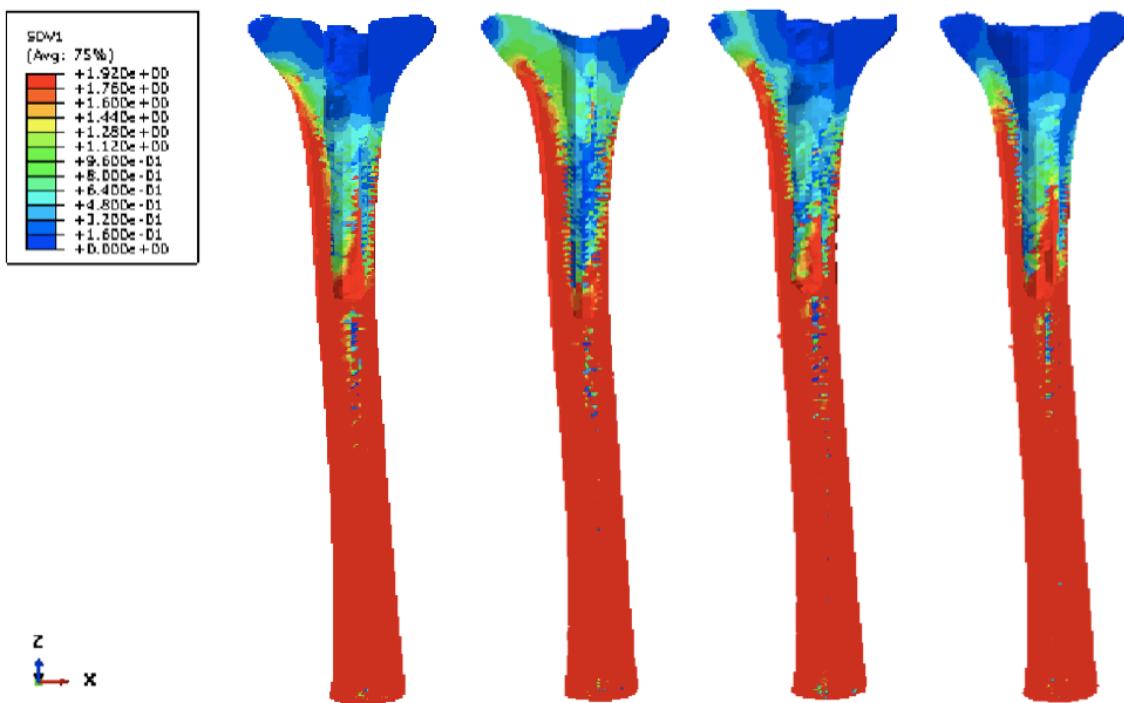


Figura 3.15. Comparación de la distribución de densidades (gr/cm^3) para las cuatro prótesis en el último incremento. a) Prótesis en offset sin suplemento. b) Prótesis con vástago recto. c) Prótesis en offset con suplemento. d) Prótesis con vainas

En la siguiente gráfica (Figura 3.16) se muestran los resultados del ratio de densidad ósea obtenidos para cada tipo de prótesis. La prótesis con vástago recto es, con diferencia, la que produce una menor reabsorción ósea, seguida de la prótesis con vástago en offset con suplemento y por último la prótesis en offset sin suplemento y la prótesis de vainas, que producen una reabsorción ósea similar.

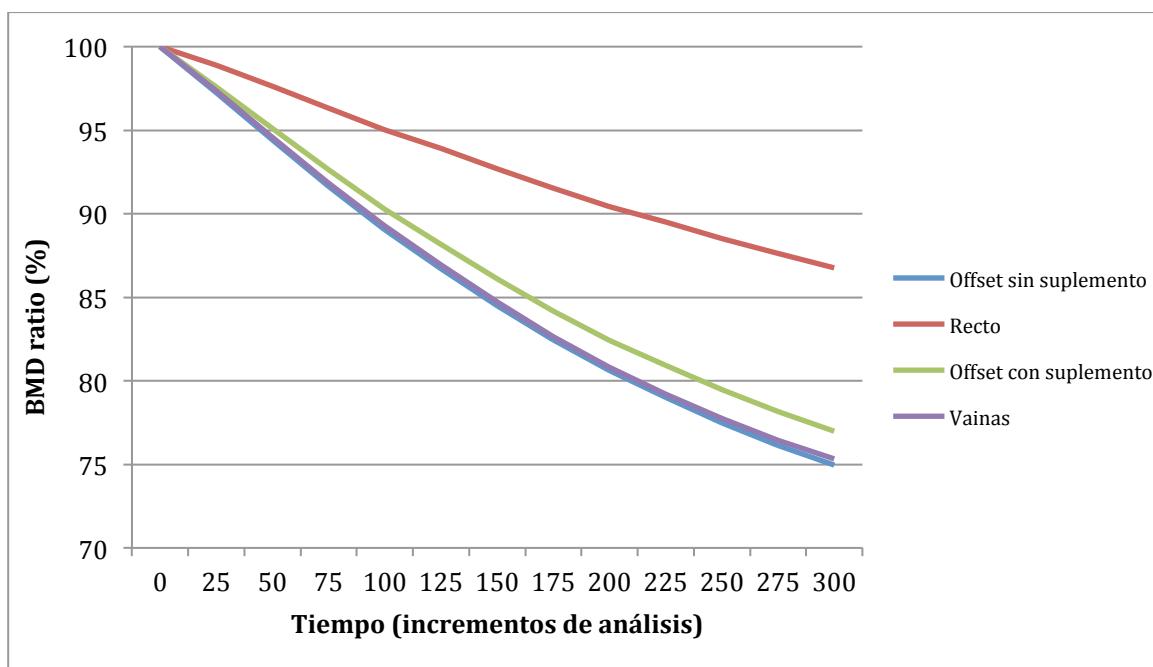


Figura 3.16. Variación del ratio de densidad ósea con el tiempo

También se ha evaluado la tensión máxima principal que produce cada prótesis en el hueso (*Figura 3.17 y 3.18*). La tensión principal máxima en el hueso no supera en ningún momento la resistencia a tracción del hueso (30 MPa). Todas las prótesis producen distribuciones de tensiones similares en el hueso. Las prótesis que producen una mayor zona de tensiones son la prótesis con vástago recto seguida de la prótesis en offset con suplemento. En cuanto a la zona donde aparece tensión puede ser debida al estrechamiento de la sección de la tibia en esa zona, que produce un efecto entalla y actúa como concentrador de tensiones. Los máximos locales que se producen en los nodos en los que se aplicaron las condiciones de apoyo.

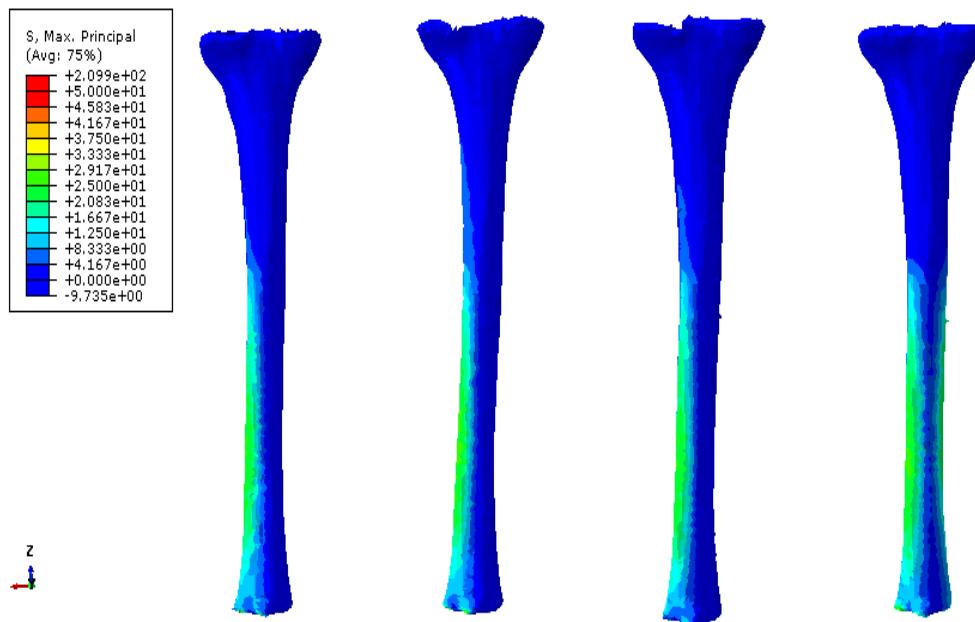


Figura 3.17. Tensión principal máxima en el hueso (MPa). a) Prótesis en offset sin suplemento. b) Prótesis con vástago recto. c) Prótesis en offset con suplemento. d) Prótesis con vainas

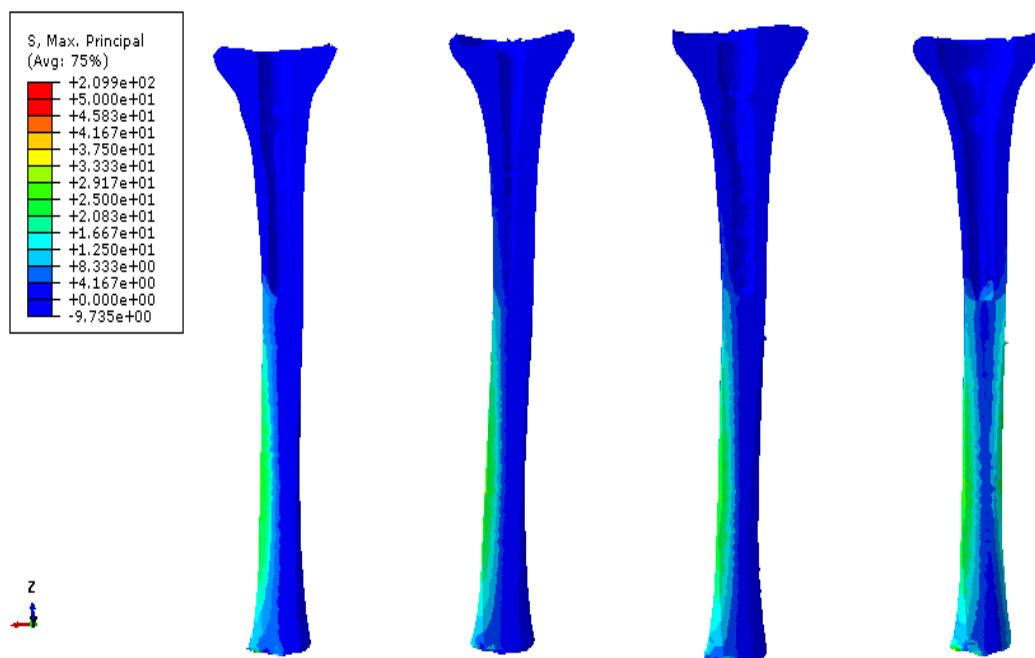


Figura 3.18. Tensión principal máxima en el hueso (MPa). Vista corte transversal. a) Prótesis en offset sin suplemento. b) Prótesis con vástago recto. c) Prótesis en offset con suplemento. d) Prótesis con vainas

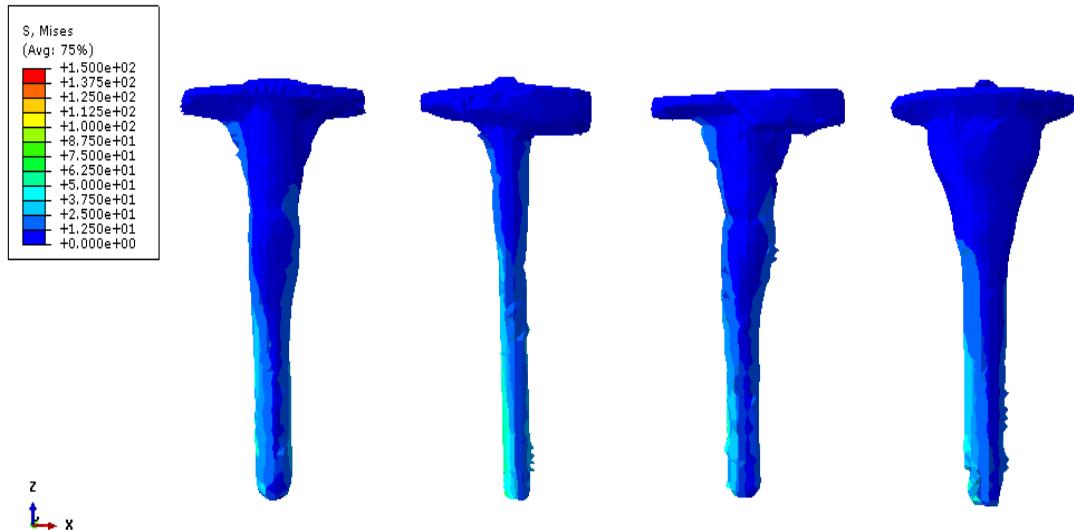


Figura 3.19. Tensiones de Von Mises en el implante (MPa). Vista delantera. a) Prótesis en offset sin suplemento. b) Prótesis con vástago recto. c) Prótesis en offset con suplemento. d) Prótesis con vainas



Figura 3.20. Tensiones de Von Mises en el implante (MPa). Vista trasera. a) Prótesis en offset sin suplemento. b) Prótesis con vástago recto. c) Prótesis en offset con suplemento. d) Prótesis con vainas

En las imágenes anteriores se observa la comparación entre las cuatro prótesis de las tensiones de Von Mises a las que se encuentran sometidos los vástagos (*Figuras 3.19 y 3.20*). La tensión de Von Mises en los implantes no supera en ningún momento la resistencia del mismo (450 MPa). Los mayores valores de tensión se localizan en la punta del vástago. La prótesis que está sometida a más tensiones es la de vástago recto, mientras que, la prótesis con offset sin suplemento es la que menos tensiones soporta.

3.4. Conclusiones

La prótesis que mayor reabsorción producen son la prótesis de vainas y la prótesis con vástago en offset sin suplemento, seguida muy de cerca por la prótesis con vástago en offset con suplemento. La prótesis que, con diferencia, produce una menor reabsorción ósea es la prótesis con vástago recto. Esto puede deberse a que no posee un offset ni unas vainas que puedan actuar como concentradores de tensiones, descargando el hueso y produciendo la reabsorción ósea. Esto también es debido a que la prótesis con vástago recto es la menos rígida, por lo que es capaz de transmitir de una manera más uniforme las tensiones, reduciendo la reabsorción ósea. La prótesis de vainas es la más rígida, de ahí que genere la mayor reabsorción ósea.

Las prótesis con vástago en offset son las que producen mayor tensión principal máxima en el hueso, aunque en ningún momento supera los máximos permitidos. Esto puede ser debido, como se ha comentado anteriormente, a que el offset actúa como un importante concentrador de tensiones localizado localmente. En cuanto a la tensión de Von Mises en las prótesis, los mayores valores aparecen localizados en la punta del vástago, destacando las tensiones que se encuentran en la prótesis de vástago recto.

4. Conclusiones y líneas futuras

4.1. Resumen

Este trabajo se ha dividido en dos partes: la primera la determinación de las cargas que afectan a la tibia de un paciente concreto mediante un método basado en Redes Neuronales, y la segunda la utilización de dichas cargas para analizar la reabsorción ósea que producen cuatro tipos distintos de prótesis de rodilla de revisión.

Para la primera parte del trabajo se ha analizado el problema de remodelación ósea, aplicando un rango de variación de un 20% de las condiciones de carga al modelo 3D de una tibia real. De este análisis se han obtenido unas densidades óseas, que sirven de entrada para la Red Neuronal Artificial. Después de entrenar la red con una función sigmoidea como entrada, distinto número de neuronas en la capa intermedia y una función lineal como salida, se han obtenido unas cargas, las cuales se han comparado con las condiciones de carga aplicadas inicialmente y se ha obtenido un error de cálculo. Tras determinar el número de entradas y neuronas de la capa intermedia más adecuado, se han introducido las cargas obtenidas en el modelo de la tibia con distintas prótesis, evaluando su efecto en la remodelación ósea.

4.2. Conclusiones

En cuanto a la primera parte de este trabajo, una de las principales conclusiones es que el modelado la red neuronal se puede llegar a predecir con exactitud las condiciones de carga de un paciente específico a partir de su distribución de densidades. Gracias a este método, podemos obtener las cargas que sufre la tibia y sus angulo de aplicación. De esta forma, el modelo matemático desarrollado en este proyecto puede ser utilizado en la vida real para predecir las cargas de un paciente, las cuales nos proporcionen la información necesaria para implantar de forma más eficiente una prótesis para dicho paciente si fuera necesario.

Como conclusiones secundarias se ha observado que cuantas más neuronas se utilicen para entrenar la Red Neuronal Artificial, mayor será el coste computacional y menor será el error, hasta llegar a un límite de sobreentrenamiento para el cual el error volverá a aumentar. La disminución del número de entradas para la Red Neuronal supone una disminución del coste computacional y un pequeño aumento del error. Con estas conclusiones, hemos obtenido como modelo óptimo para nuestro problema una red con 20 entradas y una capa intermedia con 30 neuronas. Los resultados obtenidos de la red han sido buenos para las zonas de cortical, pero en las zonas de baja densidad o donde se aplican las condiciones de contorno presenta errores de aproximación más altos.

En la segunda parte de este trabajo se han comparado las distintas prótesis. En el análisis de remodelación ósea realizado se ha visto que la prótesis con vástago recto es la que menos reabsorción ósea produce, lo que puede ser debido a que no tiene un offset ni unas vainas que actúen como concentradores de tensiones, descargando el hueso y produciendo una importante reabsorción ósea y por ser un vástago mucho menos rígido que el resto. Por este motivo, las prótesis con vástago en offset y con vainas han producido las mayores tensiones y deformaciones.

4.3. Futuras líneas de investigación

Las conclusiones obtenidas en este proyecto son importantes, pero siempre se puede mejorar el trabajo realizado y explorar otros aspectos que permitan avanzar en la investigación realizada. Por lo tanto las futuras líneas de acción que se proponen son las siguientes:

- Efecto de la incorporación del peroné en las simulaciones de la remodelación ósea
- Desarrollo de un modelo paramétrico de la tibia válido para cualquier paciente
- Nueva metodología basada en las redes neuronales donde las entradas sean los parámetros del modelo paramétrico anterior y que sea capaz de predecir las cargas de cualquiera paciente
- Comparación de distintos vástagos con un platillo tibial sin vástagos

Anexo I. Redes neuronales

Las redes neuronales artificiales (RNA) son modelos matemáticos que intentan reproducir el funcionamiento del sistema nervioso. Como todo modelo realizan una simplificación del sistema real que simula y toma las características principales del mismo para la resolución de una tarea determinada.

I. El modelo biológico

El cerebro es el elemento principal del sistema nervioso humano y está compuesto por un tipo especial de célula llamada neurona. Una neurona es una célula viva y como tal posee todos los elementos comunes de las células biológicas. A su vez, las neuronas tienen características propias que le permiten comunicarse entre ellas, lo que las diferencia del resto de las células biológicas. La figura I.1 muestra la estructura típica de una neurona biológica.

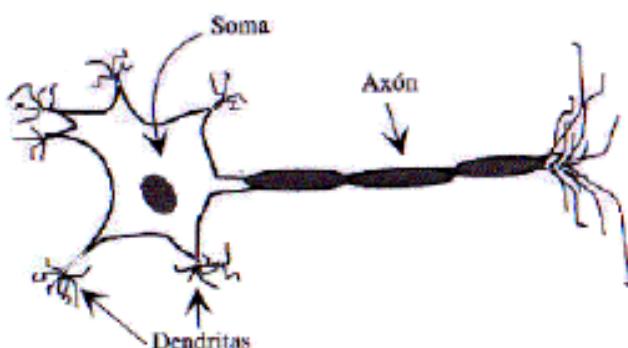


Figura I.1: Neurona biológica

Se observa que la neurona biológica está compuesta por un *cuerpo celular* o *soma*, del cual se desprende un árbol de ramificaciones llamado *árbol dendrítico*, compuesto por las *dendritas*. Del soma también parte una fibra tubular, llamada *axón*, el cual suele ramificarse cerca de su extremo. Las dendritas actúan como un canal de entrada de señales provenientes desde el exterior hacia la neurona, mientras que el axón actúa como un canal de salida. El espacio entre dos neuronas vecinas se denomina *sinapsis*.

Desde un punto de vista funcional, las neuronas conforman un procesador de información sencillo. Constan de un subsistema de entrada (dendritas), un subsistema de procesamiento (el soma) y un subsistema de salida (axón). Como característica principal destaca su capacidad de interacción con otras neuronas, las señales nerviosas pueden ser eléctricas o químicas. La transmisión química se da principalmente en la comunicación entre neuronas mientras que la eléctrica se produce dentro de una neurona [Martinez et al., 2003].

En general, una neurona recibe información de cientos de neuronas vecinas y la transmite a otras tantas neuronas. La comunicación entre neuronas se lleva a cabo de la siguiente manera: el soma de las neuronas transmisoras o presinápticas se genera un pulso eléctrico llamado potencial de acción. El pulso eléctrico se propaga a través del axón en dirección a las sinapsis. La información se transmite a las neuronas vecinas utilizando un proceso químico, mediante la liberación de neurotransmisores. Estos neurotransmisores se transmiten a través de la sinapsis hacia la neurona receptora. La neurona receptora o postsináptica toma la señal enviada por cientos de neuronas a través de las dendritas y la

transmite al cuerpo celular. Estas señales pueden ser excitadoras (positivas) o inhibidoras (negativas) [Gurmey, 1997]. El soma es el encargado de integrar la información proveniente de las distintas neuronas. Si la señal resultante supera un determinado umbral (umbral de disparo) el soma emite un pulso que se transmite a lo largo del axón dando lugar a la transmisión eléctrica a lo largo de neurona. Al llegar la señal al extremo del axón se liberan neurotransmisores que permiten transmitir la señal a las neuronas vecinas [Nascimiento, 1994]. Este modelo biológico es la base que siguen los distintos sistemas basados en neuronas artificiales y que brevemente se van a presentar en los siguientes apartados.

II. Modelo de neurona artificial

La neurona artificial es un elemento de procesamiento simple que a partir de un vector de entradas produce múltiples salidas. En general podemos encontrar tres tipos de neuronas artificiales y podemos clasificarlas según su función en:

Neuronas de entrada: Son aquellas que reciben directamente la información desde el exterior.

Neuronas ocultas: Llamadas de esta forma a las neuronas que reciben la información desde otras neuronas artificiales. Es en estas neuronas donde se realiza la representación de la información almacenada.

Neuronas de salida: Las cuales reciben la información ya procesada y la devuelven al exterior.

En la *figura I.2* se muestran todos los elementos que componen una neurona artificial:

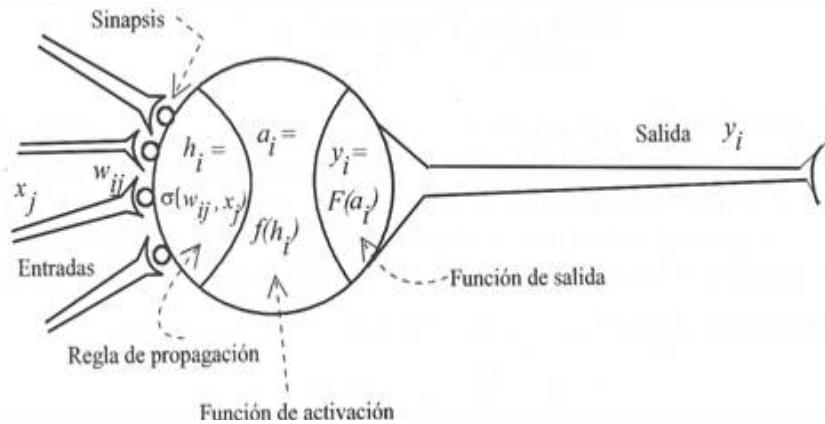


Figura I.2: Neurona artificial

Conjunto de entradas, $X_j(t)$: Reciben la información, esta puede ser proveniente del exterior o de otra neurona artificial.

Pesos sinápticos, W_{ij} : Representan el grado de comunicación entre dos neuronas artificiales. Estos pesos pueden ser excitadores o inhibidores.

Regla de propagación, σ : Integra la información proveniente de las distintas neuronas artificiales, entradas y pesos sinápticos, y proporcionan el valor potencial postsináptico de la neurona.

Función de activación, f_i : Provee el estado de activación de la neurona i.

Función de salida, F_i : Representa la salida actual de la neurona i.

A continuación se presentan los puntos expuestos anteriormente de una forma más detallada.

i. Entradas y salidas

Las entradas y salidas de una neurona pueden ser clasificadas en dos grandes grupos, binarias o continuas. Las neuronas binarias (digitales) sólo admiten dos valores posibles. En general en este tipo de neurona se utilizan los siguientes dos alfabetos {0,1} o {-1,1}. Por su parte, las neuronas continuas (analógicas) admiten valores dentro de un determinado rango, que en general suele definirse como [-1, 1]. La selección del tipo de neurona a utilizar depende de la aplicación y del modelo a construir.

ii. Pesos sinápticos

El peso sináptico W_{ij} define la fuerza de una conexión sináptica entre dos neuronas, la neurona presináptica i y la neurona postsináptica j . Los pesos sinápticos pueden tomar valores positivos, negativos o cero. En caso de una entrada positiva, un peso positivo actúa como excitador, mientras que un peso negativo actúa como inhibidor. En caso de que el peso sea cero, no existe comunicación entre el par de neuronas.

Mediante el ajuste de los pesos sinápticos la red es capaz de adaptarse a cualquier entorno y realizar una determinada tarea.

iii. Regla de propagación

La regla de propagación determina el potencial resultante de la interacción de la neurona i con las N neuronas vecinas. El potencial resultante h_i se puede expresar de la siguiente manera:

$$h_i(t) = \sigma_i(W_{ij}, X_j(t))$$

La regla de propagación más simple y utilizada consiste en realizar una suma de las entradas ponderadas con sus pesos sinápticos correspondientes:

$$h_i(t) = \sum W_{ij} * X_j(t)$$

iv. Función de activación

La función de activación determina el estado de activación actual de la neurona en base al potencial resultante h_i y al estado de activación anterior de la neurona $a_i(t-1)$. El estado de activación de la neurona para un determinado instante de tiempo t puede ser expresado de la siguiente manera:

$$a_i(t) = f_i(a_i(t-1), h_i(t))$$

Sin embargo, en la mayoría de los modelos se suele ignorar el estado anterior de la neurona, definiéndose el estado de activación en función del potencial resultante h_i :

$$a_i(t) = f_i(h_i(t))$$

La *Tabla I.1* muestra un listado de algunas de las funciones de activación más utilizadas en los distintos modelos de redes neuronales artificiales.

	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
Lineal a tramos	$y = \begin{cases} -1, & \text{si } x < -l \\ x, & \text{si } -l \leq x \leq l \\ +1, & \text{si } x > l \end{cases}$	$[-1, +1]$	
Sigmoidea	$y = \frac{1}{1+e^{-x}}$ $y = \text{tgh}(x)$	$[0, +1]$ $[-1, +1]$	
Gaussiana	$y = Ae^{-Bx^2}$	$[0, +1]$	
Sinusoidal	$y = A \sen(\omega x + \varphi)$	$[-1, +1]$	

Tabla I.1: Funciones de activación.

v. Función de salida

La función de salida proporciona el valor de salida de la neurona, en base al estado de activación de la neurona. En general se utiliza la función identidad, es decir:

$$y_i(t) = F_i(a_i(t)) = a_i(t)$$

III. Arquitectura de una red neuronal

Una vez definida el tipo de neurona que se utilizará en un modelo de redes neuronales artificiales es necesario definir la topología de la misma. La organización y disposición de las neuronas dentro de una red neuronal se denomina topología, y viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad, y el tipo de conexión entre neuronas.

Las neuronas suelen agruparse en unidades funcionales denominadas capas. Se denomina capa de entrada a aquella que está compuesta por neuronas de entradas y por lo tanto recibe información procedente desde el exterior. Análogamente, se denomina capa oculta y capa de salida a aquellas capas que están compuestas por neuronas ocultas y de salida

respectivamente. Una red neuronal artificial está compuesta por una o más capas, las cuales se encuentran interconectadas entre sí. Entre un par de neuronas de la red neuronal artificial pueden existir conexiones. Estas conexiones son las sinapsis, tienen asociadas un peso sináptico, y son direccionales. Cuando la conexión se establece entre dos neuronas de una misma capa hablamos de conexiones laterales o conexiones intra-capa. Por el contrario, si la conexión se establece entre neuronas de distintas capas se la denomina conexión inter-capa. Si la conexión se produce en el sentido inverso al de entrada-salida la conexión se llama recurrente o realimentada.

Una red puede estar formada por una única capa de neuronas. En este caso hablamos de redes monocapa, y las neuronas que conforman dicha capa cumplen la función de neuronas de entrada y salida simultáneamente. Cuando la red está compuesta por dos o más capas hablamos de redes multicapa. A su vez, hablamos de redes neuronales con conexión hacia delante (redes feedforward) cuando las conexiones entre las distintas neuronas de la red siguen un único sentido, desde la entrada de la red hacia la salida de la misma. Cuando las conexiones pueden ser tanto hacia delante como hacia atrás hablamos de redes recurrentes (redes feedback).

IV. Funcionamiento de la red neuronal

Durante la operación de una red neuronal podemos distinguir claramente dos fases o modos de operación: la fase de aprendizaje o entrenamiento, y la fase de operación o test. Durante la primera fase, la fase de aprendizaje, la red es entrenada para realizar un determinado tipo de procesamiento. Una vez alcanzado un nivel de entrenamiento adecuado, se pasa a la fase de operación, donde la red es utilizada para llevar a cabo la tarea para la cual fue entrenada.

i. Fase de entrenamiento

Una vez seleccionada el tipo de neurona artificial que se utilizará en una red neuronal y determinada su topología es necesario entrenarla para que la red pueda ser utilizada. Partiendo de un conjunto de pesos sinápticos aleatorio, el proceso de aprendizaje busca un conjunto de pesos que permitan a la red desarrollar correctamente una determinada tarea. Durante el proceso de aprendizaje se va refinando iterativamente la solución hasta alcanzar un nivel de operación suficientemente bueno.

El proceso de aprendizaje se puede dividir en tres grandes grupos de acuerdo a sus características en: [Isasi Viñuela y Galván León, 2004; Yao, 1999]

Aprendizaje supervisado. Se presenta a la red un conjunto de patrones de entrada junto con la salida esperada. Los pesos se van modificando de manera proporcional al error que se produce entre la salida real de la red y la salida esperada.

Aprendizaje no supervisado. Se presenta a la red un conjunto de patrones de entrada. No hay información disponible sobre la salida esperada. El proceso de entrenamiento en este caso deberá ajustar sus pesos en base a la correlación existente entre los datos de entrada.

Aprendizaje por refuerzo. Este tipo de aprendizaje se ubica entre medio de los dos anteriores. Se le presenta a la red un conjunto de patrones de entrada y se le indica a la red si la salida obtenida es o no correcta. Sin embargo, no se le proporciona el valor de la salida esperada. Este tipo de aprendizaje es muy útil en aquellos casos en que se desconoce cual es la salida exacta que debe proporcionar la red.

ii. Fase de operación o test

Una vez finalizada la fase de aprendizaje, la red puede ser utilizada para realizar la tarea para la que fue entrenada. Una de las principales ventajas que posee este modelo es que la red aprende la relación existente entre los datos, adquiriendo la capacidad de generalizar conceptos. De esta manera, una red neuronal puede tratar con información que no le fue presentada durante de la fase de entrenamiento.

V. Redes neuronales con conexión hacia delante

Este tipo de redes neuronales son las que mayor aplicación práctica tienen, se caracterizan por su organización en capas y conexiones estrictamente hacia delante, utilizando entrenamientos de tipo supervisado.

Dentro de este grupo destacaremos el perceptrón y el perceptrón multicapa (MLP) los cuales desarrollaremos a continuación.

i. Perceptrón simple

Este modelo tiene gran importancia histórica ya que fue el primer modelo en poseer un mecanismo de entrenamiento que permite determinar automáticamente los pesos sinápticos que clasifican correctamente a un conjunto de patrones a partir de un conjunto de ejemplos.

La arquitectura del perceptrón está compuesta por dos capas de neuronas, una de entrada y una de salida. La capa de entrada es la que recibe la información proveniente del exterior y la transmite a las neuronas sin realizar ningún tipo de operación sobre la señal de entrada. En general la información entrante es binaria. La función de activación de las neuronas de un perceptrón es del tipo escalón, dando de esta manera sólo salidas binarias. Cada neurona de salida del perceptrón representa a una clase. Una neurona de salida responde con 1 si el vector de entrada pertenece a la clase a la que representa y responde con 0 en caso contrario.

El algoritmo de entrenamiento del perceptrón se encuentra dentro de los denominados algoritmos por corrección de errores. Este tipo de algoritmos ajustan los pesos de manera proporcional a la diferencia entre la salida actual proporcionada por la red y la salida objetivo, con el fin de minimizar el error producido por la red.

ii. Perceptrón multicapa (MLP)

El perceptrón multicapa es una extensión del perceptrón simple. La topología de un perceptrón multicapa está definida por un conjunto de capas ocultas, una capa de entrada y una de salida. No existen restricciones sobre la función de activación aunque en general se suelen utilizar funciones sigmoideas (ver *tabla I.1*). A continuación se expone el esquema tipo para una MLP, el cual es el modelo más utilizado actualmente, ver *figura I.3*.

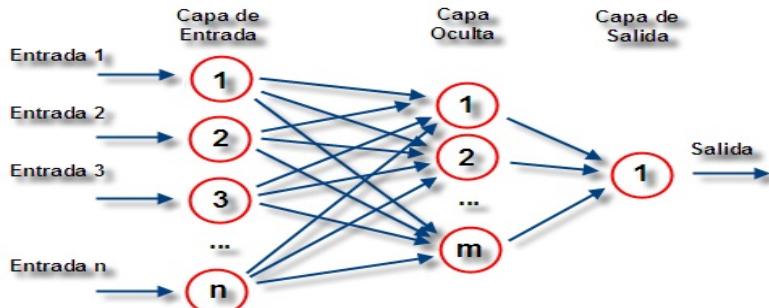


Figura I.3. Perceptrón multicapa (MLP)

VI. Entrenamiento de la red neuronal

En el contexto de las redes neuronales el entrenamiento o aprendizaje puede ser visto como el proceso de ajuste de los parámetros libres de la red [Yao, 1995]. Partiendo de un conjunto de pesos sinápticos aleatorios, el proceso de entrenamiento busca un conjunto de pesos que permitan a la red desarrollar correctamente una determinada tarea. Este proceso de entrenamiento, es un proceso iterativo, en el cual se va refinando la solución hasta alcanzar un nivel de operación lo suficientemente bueno.

La mayoría de los métodos de entrenamiento utilizados en las redes neuronales con conexión hacia delante consisten en proponer una función de error que mida el rendimiento actual de la red en función de los pesos sinápticos. El objetivo del método de entrenamiento es encontrar el conjunto de pesos sinápticos que minimizan (o maximizan) la función. El método de optimización proporciona una regla de actualización de los pesos que en función de los patrones de entrada modifica iterativamente los pesos hasta alcanzar el punto óptimo de la red neuronal.

i. Algoritmo Backpropagation

El algoritmo de entrenamiento/aprendizaje supervisado más común utilizado para una MLP, es el algoritmo llamado Backpropagation. Este algoritmo surgió de la búsqueda de un algoritmo para entrenar específicamente a la MLP. Es un tanto complejo ya que se toma la información del comportamiento de la red en el sentido directo de la red y en el sentido inverso, esto se realiza por la necesidad de modificar el comportamiento de las capas ocultas.

El algoritmo Backpropagation tiene como objetivo usar la diferencia entre las salidas deseadas y las salidas actuales en la capa de salida de la red para cambiar los pesos (indicados con valores aleatorios pequeños) con el fin de reducir al mínimo esta diferencia (error). Esto se logra mediante una serie de interacciones donde se modifica cada peso de derecha a izquierda (sentido inverso de la propagación de información en la red) hasta modificarse los pesos de la capa de entrada prosiguiendo nuevamente con la propagación de la información de entrada, esto hasta que la diferencia entre la salida deseada y la obtenida en cada neurona de salida sea mínima [Fristsch, 1996].

El algoritmo Backpropagation es el método que desde un principio se desarrollo con el fin de entrenar redes neuronales multicapa y se demostró su eficiencia suficiencia en el entrenamiento de redes para resolver diversos problemas, pero en muchos casos resultó ser muy lento. A través de los años han surgido algoritmos más poderosos, aunque más complejos, la mayoría partiendo de la base de este algoritmo, propagar el error hacia atrás. De cualquier forma, es recomendable el uso de este algoritmo cuando se trata el diseño de una MLP, ya

que no es demasiado complejo, se entiende fácilmente su finalidad, y sirve para comprender más rápido los algoritmos que se basan en él.

ii. Generalización

Una vez finalizada la fase de aprendizaje, la red puede ser utilizada para realizar la tarea para la que fue entrenada. Una de las principales ventajas que posee este modelo es que la red aprende la relación existente entre los datos, adquiriendo la capacidad de generalizar conceptos. De esta manera, una red neuronal puede tratar con información que no le fue presentada durante la fase de entrenamiento [Chirungrueng, 1988].

Cuando se evalúa una red neuronal no sólo es importante evaluar si la red ha sido capaz de aprender los patrones de entrenamiento. Es imprescindible también evaluar el comportamiento de la red ante patrones nunca antes vistos. Esta característica de las redes neuronales se la conoce como capacidad de generalización y es adquirida durante la fase de entrenamiento. Es necesario que durante el proceso de aprendizaje la red extraiga las características de las muestras, para poder luego responder correctamente a nuevos patrones.

De lo dicho anteriormente surge la necesidad de evaluar durante la fase de entrenamiento dos tipos de errores. El **error de aprendizaje**, que indica la calidad de la respuesta de la red a los patrones de entrenamiento y el **error de test**, que indica la calidad de la respuesta de la red a patrones nunca antes vistos. Para poder obtener una medida de ambos errores es necesario dividir el set de datos disponibles en dos, el set de datos de entrenamiento, y el set de datos de evaluación. El primero se utiliza durante la fase de entrenamiento para que la red pueda extraer las características de los mismos mediante el ajuste de sus pesos sinápticos, así la red logra una representación interna de la función. El set de evaluación se utiliza para evaluar la capacidad de generalización de la red.

La causa más común de la pérdida de capacidad de generalización es el **sobreaprendizaje**. Esto sucede cuando la cantidad de ciclos de entrenamientos tiende a ser muy alta. Se observa que la respuesta de la red a los patrones de entrenamiento es muy buena mientras que la respuesta a nuevos patrones tiende a ser muy pobre. Al aumentar el número de ciclos la red tiende a sobreajustar la respuesta a los patrones de entrenamiento, a expensas de una menor capacidad de generalización. La Figura I.4.a muestra una situación idealizada de lo dicho anteriormente. En la misma se observa que en un determinado punto se empieza a perder capacidad de generalización como consecuencia del sobreaprendizaje.

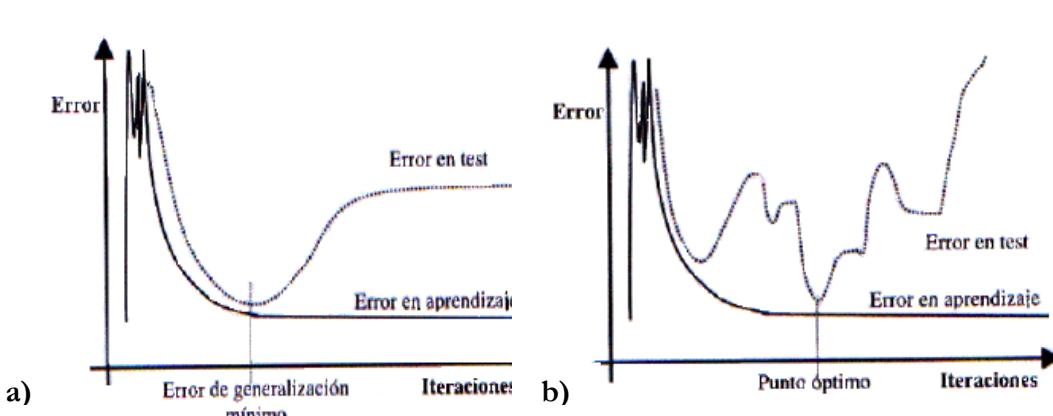


Figura I.4: Generalización. a) Situación idealizada. b) Situación real.

En la *Figura I.4.b* se muestra una situación más real del mismo caso. A medida que transcurre el proceso de aprendizaje se obtienen varios mínimos sobre el conjunto de evaluación. Existen diversas técnicas de parada temprana aunque en la mayoría de los casos se deja que el proceso de aprendizaje avance hasta alcanzar una cota de error razonable, guardando periódicamente las distintas configuraciones intermedias para luego seleccionar la de menor error de evaluación.

En ocasiones la pérdida de capacidad de generalización se produce por el uso excesivo de neuronas ocultas en la red neuronal. Esto hace que la red tienda a ajustar con mucha exactitud los patrones de entrenamiento, evitando que la red extraiga las características del conjunto.

Bibliografía

- [1] Doblaré M, García J. 2002. Anisotropic bone remodelling model based on a continuum damage-repair theory. *J Biomech.* 35(1):1 – 17.
- [2] M. A. Pérez, P. Fornells, M. Doblaré, J. García-Aznar. Comparative analysis of bone remodelling models with respect to computerised tomography-based finite element models of bone. *Computer Methods in Biomechanics and Biomedical Engineering*, 13 (2010), pp. 71-80
- [3] Nyman JS, Hazelwood SJ, Rodrigo JJ, Martin RB, Yeh OC. 2004. Long stemmed total knee arthroplasty with interlocking screws: a computational bone adaptation study. *J Orthop Res.* 22:51–57.
- [4] Bergmann G, Heller M, Duda GN. 2001. Preclinical testing of cemented hip replacement implants: pre-normative research for a European Standard, Final Report of Workpackage 5: Development of the Loading Configuration, EU Contract SMT4-CT96-2076, 1999, HIP98.
- [5] Heller MO, Bergmann G, Deuretzbacher G, Duurselen L, Pohl M, Claes L, Hass NP, Duda GN. 2001. Musculo-skeletal loading conditions at the hip during walking and stair climbing. *J Biomech.* 34:883–893.
- [6] Peng L, Bai J, Zeng X, Zhou Y. 2007. Comparison of isotropic and orthotropic material property assignments on femoral finite element models under two loading conditions. *Med Eng Phys.* 28:227–233
- [7] Martínez G, Pasquini R, ServenteD , Sistemas Inteligentes. Nueva Librería(2003).
- [8] Gurney K, Neural Nets. CRC Press. ISBN: 1857285034 (1997).
- [9] Nascimento C, Artificial Neural Networks in Control and Optimization. PhD Thesis. University of Manchester, Institute of Science and Technology (UMIST), Control Systems Centre, Manchester, United Kingdom (1995).
- [10] Isasi Viñuela P, Galván León I. M, Redes Neuronales Artificiales. Un enfoque Práctico, Prentice Hall (2004).
- [11] Yao X, Evolutionary Artificial Neural Networks, Encyclopedia of Computer Science and Technology, Kent, A., J. G. Williams(eds), Vol 33 (1995) 137-170.
- [12] Fritsch J, Modular Neural Networks for Speech Recognition. Master's Thesis. Carnegie Mellon University (1996).
- [13] Chinrungrueng C, Analysis of simple neural networks. Master's Report Computer Science Department, University of California at Berkeley (1988).