

Modelo Matemático de Regresión Lineal Múltiple para Terapia de Hipertermia con Nanopartículas y Calentamiento Inducido por Láser

Mathematical model from Multiple Linear Regression for hyperthermia therapy with nanoparticles and heating induced by laser

Santiago Amado Ospina
santiago.amado00@usc.edu.co

Daniel Mauricio Valdes Cabrera
daniel.valdes00@usc.edu.co

Leonardo Antonio Bermeo Varón
leonardo.bermeo00@usc.edu.co

Universidad Santiago de Cali, Facultad de Ingeniería, Programa de Bioingeniería

Resumen

Actualmente, existe interés en el estudio de la terapia de hipertermia para el tratamiento del cáncer. Esta terapia consiste en aumentar la temperatura corporal entre 37 °C y 43 °C, la cual es posible obtenerlo mediante la inducción de ondas electromagnéticas como radiofrecuencia, microondas y/o láser, y ondas de ultrasonido. Adicionalmente, el uso de nanopartículas cargadas en el tumor permite un tratamiento focalizado reduciendo el daño a los tejidos sanos circundantes. Para la implementación de este tratamiento es necesario realizar experimentos *in vitro* y simulaciones de modelos matemáticos que permitan obtener un mejor conocimiento de la terapia. Existen diferentes modelos matemáticos los cuales representan la dinámica de este tratamiento. Sin embargo, estos modelos son complejos y, en muchos casos, requieren un alto costo computacional. Lo anterior limita los procesos de optimización debido a que tardarían mucho tiempo en ser resueltos. En este sentido, este trabajo propone un modelo con un tiempo de ejecución computacional reducido mediante la técnica de regresión lineal múltiple. El entrenamiento y validación del modelo de regresión se realizó sobre una base de datos de 200 simulaciones producidas por un modelo matemático desarrollado en Comsol Multiphysics® por medio de Python. Las medidas experimentales fueron tomadas en el proceso de hipertemia con inducción de calor por medio infrarrojo cercano durante 100s en agua destilada pura y dos proporciones diferentes de nanopartículas de óxido de hierro disueltas en agua destilada, i.e., 0,025% y 0,050% en peso. Una vez se obtiene el modelo, se realiza un proceso de estimación de parámetros por medio de los algoritmos de Gauss-Newton (GN) y Levenberg-Marquardt (LM). Los resultados indican que el campo de temperatura predicho por el modelo de regresión lineal múltiple es cercano a los valores de salida del modelo matemático completo y a las mediciones experimentales. La correlación del modelo propuesto fue de 0,99 con respecto al modelo realizado con Comsol Multiphysics® y de 0,95 con respecto al experimento. Además, el modelo presenta una reducción considerable del tiempo de ejecución computacional equivalente a 1,95 ms en comparación con el modelo de Comsol Multiphysics® que tardó 25,29 minutos. Esta reducción del tiempo de ejecución podría permitir planificar la terapia de hipertermia antes y durante su implementación. En cuanto al proceso de estimación, ambos algoritmos convergen a una solución estable, obteniendo así valores óptimos para los parámetros a estimar, derivados del promedio de las 50 ejecuciones de cada algoritmo.

Palabras Clave: hipertemia por láser, nanopartículas de óxido de hierro, tratamiento del cáncer, regresión lineal múltiple

Abstract

Nowadays, there is interest in the study of hyperthermia therapy for the treatment of cancer. This therapy consists on increasing body temperature to between 37 °C and 43 °C, which can be achieved by inducing electromagnetic waves like radiofrequency, microwave and/or laser and ultrasound waves. Additionally, the use of nanoparticles loaded in the tumor allows a focalized treatment reducing damage to surrounding healthy tissues. For the implementation of this treatment it is necessary to perform *in vitro* experiments and mathematical model simulations to obtain a better understanding of the therapy. There are different mathematical models which represent the dynamics of this therapy. However, these models are complex and, in many cases, require a high computational cost. This limits the optimization processes since they would take a long time to be solved. In this sense, this work proposes a model with a reduced computational execution time using the multiple linear regression technique. The training and validation of the regression model was performed on database of 200 simulations produced by a mathematical model developed in Comsol Multiphysics® by Python. Experimental measurements were taken in the hyperthermia process with heat induction by near infrared during 100s in pure distilled water and two different ratios of iron oxide nanoparticles dissolved in distilled water, i.e., 0.025% and 0.050% by weight. Once the model is obtained, a parameter estimation process is performed by means of the Gauss-Newton (GN) and Levenberg-Marquardt (LM) algorithms. The results indicate that the temperature field predicted by the multiple linear regression model is close to the output

values of the full mathematical model and the experimental measurements. The correlation of the model proposed was 0.99 with respect to the model performed with Comsol Multiphysics and 0.95 with respect to the experiment. Also, the model presents a considerable reduction of the computational execution time equivalent to 1.95 ms compared to the Comsol Multiphysics model that took 25.29 minutes. This reduction in run time could allow planning the hyperthermia therapy before and during its implementation. The finding for the estimation process was that both algorithms converge to a stable solution, thus obtaining estimated values for the estimable parameters derived from the mean of the 50 runs of each algorithm.

Keywords: cancer treatment, hyperthermia by laser, iron oxide nanoparticles, multiple linear regression

1. INTRODUCCIÓN

La hipertermia o terapia térmica, es el proceso donde el tejido corporal está expuesto a altas temperaturas (37 – 43°C) que buscan dañar o destruir tejidos cancerosos que se caracterizan por tener bajos niveles de oxígeno, nutrientes y pH (Bermeo, Orlande & Elicabe, 2015). Este tratamiento ha sido foco de investigación durante las últimas décadas por sus prometedores resultados como método alternativo contra el cáncer, siendo eficaz para destruir células tumorales y no inducir efectos secundarios de los tratamientos convencionales para combatir el cáncer. La hipertermia puede clasificarse como local, regional y de cuerpo entero. Por otro lado, el método de tratamiento hipertérmico se puede coadministrar con tratamientos convencionales, siendo en algunos casos mucho más efectivo. Se han identificado diferentes tratamientos de hipertermia, entre otros, por medio de ondas de radiofrecuencia, microondas e infrarrojo (Giombini, y otros, 2007; Lamien, Orlande & Elicabe, 2017; Singh, Das & Misra, 2014), que usan técnicas distintas para llegar al mismo fin que es la apoptosis o muerte celular.

En el caso de la hipertermia inducida por láser, es un procedimiento fototérmico (PTT) basado en inducción de radiación térmica provocada por un láser de infrarrojo cercano (NIR) como fuente de energía para tratar enfermedades (Singh, Das & Misra, 2014). El calentamiento del tejido con láser puede sufrir algunas deficiencias, como el sobrecalentamiento de tejidos sanos circundantes, que provocaría daños térmicos en lugares no deseados. Estudios recientes sugieren una posible solución utilizando la hipertermia asistida por nanopartículas, para el calentamiento selectivo y mejorado de los tejidos con un daño mínimo a los tejidos sanos. Los resultados numéricos y experimentales muestran la importancia y la eficacia de la terapia térmica asistida por nanopartículas (Lamien, Bermeo, Orlande & Elicabe, 2017; Larreur, Lamien & Orlande, 2019).

Gracias a los avances en nanotecnología, el uso de nanopartículas ha permitido una mayor absorción de calor en el tejido tumoral emitido por la radiación térmica generada en los procedimientos PTT, para inducir selectivamente hipertermia a células cancerosas con mínimo impacto en las células sanas circundantes (Sazgarnia, Naghavi, Mehdizadeh & Shahamat, 2015). Este tratamiento también puede actuar como un método de tratamiento independiente y alternativo para algunos tipos de cánceres. Además, puede apoyarse en conjunto con métodos convencionales como la radioterapia y la quimioterapia (Niemz, 2007; van der Zee, 2002; Wust, y otros, 2002). En particular, el uso de nanopartículas de óxido de hierro (Fe₂O₃) para las aplicaciones médicas son de gran interés porque tienen características de biocompatibilidad y no son tóxicos. Estas nanopartículas se usan como agentes fototérmicos que en combinación con agua destilada crean un nanofluido y permiten simular el tratamiento de hipertermia del cáncer (Bermeo, Loiola, da Silva, Lamien, Pereira, Orlande & Dos Santos, 2020).

Los modelos matemáticos basados en la transferencia de calor se han venido estudiando en las últimas cuatro décadas como el modelo de biocalentamiento de Pennes y los modelos de biocalentamiento de Weinbaum-Jiji son algunos ejemplos. Estos modelos se basan en varias suposiciones sobre la temperatura de la sangre que ingresa al tejido y la temperatura local del mismo. Dado que Pennes se basa en el concepto de tasa de perfusión sanguínea, ignora la presencia de vasos sanguíneos grandes y su efecto sobre la transferencia de calor general en el tejido (Okajima, Takeda, Komiya & Maruyama, 2009).

La transferencia de calor que implica la hipertermia combina varios fenómenos físicos que pueden hacer complejo su modelado. Por eso las simulaciones numéricas son importantes para obtener resultados precisos y aproximados sobre estos procedimientos (Larreur, Lamien & Orlande, 2019). Actualmente, la terapia por hipertermia por medio de láser está siendo estudiada correlacionando modelos matemáticos y experimentos con el objetivo de establecer protocolos para la implementación de la terapia fototérmica. Sin embargo, estos modelos requieren de un alto coste computacional para obtener una reproducción del sistema, debido a su alta complejidad matemática, en las que no solo se incluye el formalismo

matemático, sino también la geometría que en muchos casos no tiene una solución analítica, lo que tendría que ser modelado por métodos numéricos, que si bien permitirían obtener respuesta del problema en menor tiempo, no es suficiente si se quiere optimizar o estimar variables de estado y parámetros, importantes para establecer un protocolo de tratamiento.

El modelado numérico y la simulación de dicho procedimiento es importante para comprender el comportamiento termo físico de los sistemas biológicos y posibles efectos dañinos colaterales al tratamiento, dado que la predicción de la temperatura en el desarrollo de la hipertermia inducida por láser es sustancial para la eficiencia del mismo (Fanjul, Romanov & Arce, 2009; Singh, Das & Misrha, 2014). Para lograr un modelo predictivo de este procedimiento se hace uso de modelos matemáticos a partir de datos, esa definición se alinea muy bien a lo que se conoce como aprendizaje automático.

Bélangier, y otros (2005) estiman la temperatura del agua a partir de parámetros hidrometeorológicos a través de redes neuronales y regresión lineal múltiple y concluyen que ambos algoritmos tienen un desempeño satisfactorio en la predicción de la temperatura del agua para diversas condiciones meteorológicas. Bondarev (2019) compara las predicciones de los modelos de regresión lineal múltiple (LRM) y los modelos de una red neuronal artificial para modelar las constantes de estabilidad de los coronatos de sodio y potasio. Mokarram (2016) mediante el uso de regresiones lineales múltiples investiga la calidad del agua subterránea y determina la relación entre los parámetros efectivos en la calidad del agua. Delbari, Sharifazari & Mohammadi (2019) evalúan el rendimiento de un modelo basado en la regresión de vectores de soporte (SRV) para estimar la temperatura diaria del suelo a (10, 30 y 100) cm de profundidad en diferentes condiciones climáticas y se comparó con los resultados obtenidos a partir de un modelo de regresión lineal múltiple. El rendimiento de los modelos LRM y SVR fue bastante bueno en la capa superficial (10 cm de profundidad) y en condiciones de clima húmedo que en áreas áridas e hiperáridas.

Con respecto a procesos de optimización, se han realizado procesos de estimación de parámetros en hipertermia por radiofrecuencia (Bermeo, y otros, 2016; Bermeo, Orlande & Elicabe, 2015; Pacheco, y otros, 2020; López, Serna, Bermeo & Castillo, 2020), estimación de variables y parámetros usando nanopartículas, estimación de variables de estado por inducción de calor por medio de láser y usando nanopartículas (Lamien, Bermeo, Orlande & Elicabe, 2017; Lamien, Orlande & Elicabe, 2017; Lamien, y otros, 2017) y estimación de parámetros en hipertermia con ultrasonido (de los Ríos, Bermeo & Pereira, 2020; Huang & Huang, 2007). Los cuales tiene buenas aproximaciones al valor de parámetros y estimación de variables y contribuyen al conocimiento de las propiedades que participan en el proceso de hipertermia. Sin embargo, el proceso es realizado con medidas experimentales simuladas lo que genera incertidumbre. Lamien, Bermeo, Orlande, & Elicabe (2017) realizaron procesos de estimación usando medidas experimentales *in vitro* encontrando buenos resultados para el conocimiento de propiedades en el proceso de hipertermia con láser.

Así, en este trabajo realiza (i) el aprendizaje supervisado para entrenar un modelo de regresión a partir del algoritmo de regresión lineal múltiple entre las variables dependientes e independientes del modelo de hipertermia con láser en diferentes nanofluidos. El entrenamiento se hace con un conjunto de datos de valores aleatorios para cada parámetro con desviación estándar de $\pm 10\%$ extraídos de simulaciones numéricas realizadas en COMSOL Multiphysics®, con el fin de predecir la temperatura transitoria de salida del sistema, (ii) una vez obtenido el modelo de regresión múltiple, se implementan los algoritmos de Gauss – Newton y Levenberg – Marquardt para estimar los parámetros del sistema.

2. MARCO TEÓRICO

2.1 Hipertermia con ondas electromagnéticas

La hipertermia se conoce como el proceso de elevación de la temperatura dentro de un tumor, en un rango de 37 °C a 43 °C, sin sobrepasar los límites de tolerancia de los tejidos normales vecinos. Incluso, en ocasiones se utilizan temperaturas superiores a los 45°C para alcanzar la ablación térmica. La hipertermia puede ser superficial a través de microondas, regional por medio de ondas de radiofrecuencia o local utilizando infrarrojos. En todos estos sistemas las radiaciones son no ionizantes, donde la energía presenta una distribución heterogénea en el interior de los tejidos, en función de sus características térmicas y de la perfusión sanguínea. Esta terapia actualmente está haciendo uso de nanopartículas las cuales permiten la concentración de calor en una región determinada, por ejemplo, en el tumor. La elevación térmica persistente en la zona del tumor hace que los tejidos normales vecinos no se dañen, pero las células tumorales sí. Tras conseguir daño

en el núcleo celular, la membrana plasmática y el citoesqueleto, se llega a la apoptosis o muerte celular (Palazzi, Maluta, Dall'Oglio & Romano, 2010).

Las principales acciones de la hipertermia en los tejidos neoplásicos incluyen: (i) mayor sensibilidad al calor de los tejidos neoplásicos a la hipertermia, (ii) efecto letal de la temperatura de 42-43°C sobre las células tumorales, según el tiempo de aplicación, (iii) estabilización temporal del crecimiento de las células tumorales después de una hipertermia moderada (39-41 °C), y (iv) acción prolongada de la temperatura, debido a la menor disipación térmica (Palazzi, Maluta, Dall'Oglio & Romano, 2010).

La hipertermia como tratamiento del cáncer inducido por láser usando nanopartículas incluye varios fenómenos a tener en cuenta. Para empezar, las nanopartículas atraen una fuerte absorción de calor que combinado con la luz láser, exhibe selectivamente regiones del tejido a la radiación NIR generada por la fuente de energía, concentrando la transferencia de calor a dichas regiones inyectadas con nanopartículas, siendo una alternativa mínimamente invasiva y precisa (Lamien, Orlande & Eliçabe, 2017). En esta modalidad de tratamiento, las nanopartículas se sintetizan a partir de metales nobles o nanomateriales semiconductores, como nanoestructuras de carbono o nanopolímeros. Sin embargo, el uso de nanopartículas mayormente constituidas por metales nobles, como las de óxido de hierro (Fe₂O₃) han representado mayor aplicabilidad en estudios de investigación, por su alta absorción óptica, bajos niveles de toxicidad, bajo costo relativo y biocompatibilidad (Lamien, y otros, 2018; Sani, Abeywardena & Psomopoulou, 2020). No obstante, hay que considerar otros fenómenos físicos complejos derivados de la terapia fototérmica determinantes en el tratamiento de hipertermia, como el daño térmico causado en los tejidos, relacionado a su vez, con la necesidad de definir una dosis térmica determinada en función del tiempo de exposición a la radiación, la distribución de temperatura determinada por la absorción óptica, transferencia de biocalentamiento y perfusión sanguínea (Bruno, Maurente, Lamien & Orlande, 2017; Larreur, Lamien & Orlande, 2019).

Para modelar el problema físico de hipertermia se usa la ecuación de bio-transferencia de calor desarrollada por Pennes (1948), basada en el postulado de Fourier de conducción de calor y caracteriza el efecto de la perfusión sanguínea y la tasa de generación de calor metabólico sobre la transferencia de calor dentro de un tejido vivo basándose en un análisis experimental del antebrazo humano (Dutta & Kundu, 2020). La ecuación (1) describe la dinámica de transferencia de calor, y adicionalmente se relación la fuente de calor externa que en este caso es a la fuente generada por el láser.

$$\rho c_p \frac{\partial T}{\partial t} = \Delta \cdot (k \Delta T) - \rho_b c_b \omega_b (T_b - T) + Q_m + Q_{laser} \quad (1)$$

donde ρ es la densidad, c_p es la capacidad de calor, k es la conductividad térmica, ρ_b es la densidad de sangre, c_b es el calor específico de sangre y T_b es la temperatura de la sangre, ω_b es la tasa de perfusión sanguínea, Q_m es la fuente de calor metabólica y Q_{laser} es la fuente de calor resultante de la absorción del láser. La ecuación (2) define la propagación del láser dentro del fluido la cual es determinada por la ley de Beer-Lambert.

$$Q_{laser} = \kappa \Phi(t) \quad (2)$$

donde κ es el coeficiente de absorción y Φ es la tasa de densidad de energía (fluencia) total del láser.

2.2 Algoritmos de aprendizaje automático

El aprendizaje de máquinas mejor conocido como *Machine Learning* (ML), es una técnica en la cual modelos matemáticos crean instrucciones finitas para que un agente artificial aprenda a partir de la experiencia y los datos. El aprendizaje automático es el proceso mediante el cual los algoritmos aprenden a identificar patrones o tendencias en los datos para crear modelos predictivos, consiste en extraer conocimiento de los datos. Es un campo de investigación que integra la estadística, la inteligencia artificial y la informática, denominado como análisis predictivo o aprendizaje estadístico, estos métodos pueden resolver problemas complejos a partir de un conjunto datos, donde la cantidad es importante ya que se puede tener mayor variabilidad y en consecuencia podrían mejorar el rendimiento y precisión (Asgari, Scalzo & Kasproicz, 2019; Ram, y otros, 2019). La adaptabilidad de estos algoritmos permite un análisis estadístico y predictivo de los modelos, aprenden de manera autónoma mejorando su rendimiento progresivamente para comprender el comportamiento de un

sistema permitiendo realizar tareas por sí mismo. Entre las ventajas que tiene en muchas de sus aplicaciones está la de prever el comportamiento y descubrir información útil a través de la extracción de datos, mejorar la integridad de los mismos, reducir riesgos gracias a la estimación y verificación previa a concretar cualquier decisión, reducir costos por automatizar procesos, distribuir adecuadamente el tiempo y los recursos enfocándose en la tarea a realizar (Müller & Guido, 2017).

El aprendizaje automático está basado en un conjunto de algoritmos que pretenden modelar sistemas que manejan grandes cantidades de datos, analizan esos datos de forma distinta y se dividen en tres técnicas: (i) aprendizaje supervisado, (ii) aprendizaje no supervisado y (iii) aprendizaje por refuerzo. El aprendizaje supervisado genera datos de salida esperados para calcular el error entre la salida estimada y la salida real. Esta técnica es útil cuando se sabe cuál es el valor de salida del modelo. Por otro lado, el aprendizaje no supervisado ajusta el modelo predictivo únicamente tomando en cuenta parámetros de entrada. Por último, el aprendizaje por refuerzo dota la capacidad de tomar decisiones o acciones sin indicaciones por parte de los humanos debido a que aprenden de los resultados y es útil en sistemas automatizados (Das, Dey, Pal & Roy, 2015). Todos estos algoritmos utilizan la regresión o la clasificación, es ese sentido como se clasifica su categoría. En la clasificación, los algoritmos pueden ser de dos clases o multiclases, ambos dividen los datos y los asignan a categorías en función de valores de respuesta categóricos, es decir, se entrenan para reconocer datos de entrada. En la regresión, se crea un modelo a partir de los valores para llegar a la predicción, en otras palabras, predicen el valor de un punto específico en función de los parámetros de entrada dados.

Existe una secuencia de pasos que se debe seguir para dar solución a los problemas complejos que se abordan con ML. Lo primero es recopilar, preparar y compilar los datos de entrada que se le van a dar al modelo. El segundo paso es dividir los datos en dos conjuntos: el conjunto de entrenamiento y el conjunto de validación. En la fase de entrenamiento se ajusta el modelo siendo la mayor parte de los datos. En el conjunto de validación se prueba el modelo con datos que no conoce previamente, así evalúa el rendimiento y la precisión. Finalmente, se da una interpretabilidad a los resultados.

Las técnicas de ML dependen en gran medida de un conjunto amplio de datos necesarios para entrenar un modelo en su etapa de aprendizaje, es decir, basan las predicciones en función de los datos existentes, haciendo más flexible el aprendizaje autónomo. Por otro lado, es importante mencionar la capacidad que deben tener estos algoritmos para procesar un gran número de datos, siendo la diversidad de estos la complejidad del modelo.

2.2.1 Algoritmo de Regresión Lineal

La regresión lineal (RL) es uno de los algoritmos de aprendizaje automático bajo la técnica de aprendizaje supervisado para un análisis estadístico y predictivo del modelo. En ML, se denomina regresión lineal simple cuando sólo se busca encontrar la relación lineal entre una variable independiente X y una variable dependiente Y . En el caso de la regresión lineal múltiple, se busca la misma relación, pero con más de una variable independiente. Es un método estadístico que permite ajustar una ecuación lineal para modelar la relación entre sus variables independientes que se conocen como predictoras y la variable dependiente que se conoce como respuesta (Amat, 2016, 2020).

La ecuación (3) describe el comportamiento de un modelo de regresión lineal y la ecuación (4) representa modelos lineales múltiples.

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \quad (3)$$

$$Y_i = (\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_m X_{im}) + \varepsilon_i \quad (4)$$

donde β_0 es el intercepto en Y . i.e., el valor promedio de la variable de respuesta Y esperada cuando las variables independientes son igual a cero, β_1 es la pendiente de la línea de regresión estimada. Es el efecto promedio que tiene el incremento en una unidad de la variable predictora X sobre la variable dependiente Y , manteniéndose constantes el resto de las variables, esto se conocen como coeficientes parciales de regresión (Amat, 2016), y ε es el error aleatorio, diferencia entre el valor estimado y el valor real del modelo.

Al implementar un modelo de regresión lineal se intenta estimar el valor de los coeficientes β_0 y β_1 , estos se conocen

como coeficientes de regresión. Para esto, los métodos más usados son el de ajuste por mínimos cuadrados ordinarios (OLS) y el gradiente descendente (Amat, 2020). A pesar de la existencia de distintos algoritmos de regresión existentes i.e. k- vecino más cercano (kNN), *Random Forest* (RF) y *Gradient Boosting* (GB) y la regresión de soporte vectorial (SVR) (Kumar & Sahu, 2021), el algoritmo que se ajusta adecuadamente a los datos conocidos y extraíbles del modelo completo y la independencia lineal de los parámetros de entrada es la regresión lineal de entradas y salidas múltiples.

2.3 Algoritmos de optimización

Particularmente, el análisis de la sensibilidad es una consecuencia del análisis de la incertidumbre. El peso de cada parámetro en la sensibilidad se mide por la correlación entre el parámetro y su respuesta. Mayores cambios indican parámetros con mucha sensibilidad (García-Moreno, y otros, 2016; Kaipio & Somersalo, 2005).

El método de Gauss – Newton, el método iterativo más fundamental, que proporciona una solución numérica para encontrar el mínimo cuadrado de una función en problemas lineales (Gill & Murray, 1978). Es un método rápido debido a que localmente converge cuadráticamente y solo cuando los parámetros están cerca del valor ideal. La ecuación (5) determina el valor estimado de los parámetros.

$$\mathbf{P}^{k+1} = \mathbf{P}^k + [\mathbf{J}^T \mathbf{W} \mathbf{J}]^{-1} \mathbf{J}^T \mathbf{W} [\mathbf{Y} - \mathbf{T}(\mathbf{P}^k)] \quad (5)$$

donde \mathbf{P} representa los parámetros del sistema a ser estimados, k representa el instante temporal de los parámetros, \mathbf{J} es el coeficiente de sensibilidad de los parámetros, \mathbf{W} es la matriz de covarianza de errores de medición, \mathbf{Y} son las medidas experimentales y \mathbf{T} es vector de la variable de estado dependiente.

El algoritmo de Levenberg – Marquardt es utilizado para resolver el problema genérico de ajuste de curvas. Se interpola el método de Gradiente Descendente con el método de Gauss – Newton, permitiendo realizar de manera unificada los resultados de su respectiva convergencia (Oliveira, 2019; Pereyra, 2013). A medida que el valor del parámetro se aleja de su valor óptimo, opera el descenso del gradiente, y cuando está cerca del valor óptimo, opera el método de Gauss – Newton (Dokht, y otros, 2021). Este ha sido reconocido por su potencial uso en el ambiente de la optimización de parámetros, en particular, en la minimización de mínimos cuadrados en busca de un mínimo local (Fernández, 2010). Resolver problemas de regresión no lineal y encontrar el mínimo óptimo de funciones cuadradas, es una de las aplicaciones más importantes del método LM.

En la ecuación (6), se observa la descripción que determina el valor de los parámetros por medio del algoritmo de LM. donde μ es el factor de dumping y $\mathbf{\Omega}$ es la matriz diagonal de $\mathbf{J}^T \mathbf{J}$ que depende de la sensibilidad del sistema.

$$\mathbf{P}^{k+1} = \mathbf{P}^k + [\mathbf{J}^T \mathbf{W} \mathbf{J} + \mu \mathbf{\Omega}]^{-1} \mathbf{J}^T \mathbf{W} [\mathbf{Y} - \mathbf{T}(\mathbf{P}^k)] \quad (6)$$

3 MATERIALES Y MÉTODOS

3.1 Marco experimental

Los datos experimentales fueron tomados de (Bermeo, y otros, 2020), donde realiza un análisis comparativo de las medidas de temperatura experimentales y el modelo matemático desarrollado en COMSOL Multiphysics®. Los parámetros del modelo y las propiedades térmicas son tomados de la literatura, el coeficiente de absorción (κ) es determinado por medio de ensayo y error en el modelo matemático. El experimento es realizado sobre tres muestras una de agua destilada y dos nanofluidos.

Para los dos nanofluidos están compuestos por nanopartículas de óxido de hierro (Fe_2O_3) producidas en (Lamien, y otros, 2018) las cuales se dispersan en agua destilada con dos concentraciones diferentes. Las nanopartículas son fabricadas por medio de aleación mecánica y son dispersas con un procesador ultrasónico (SONICS -VC 505) en agua destilada durante 30 min a una potencia de 250 W. Los nanofluidos se produjeron disolviendo 10 mg y 20 mg de nanopartículas en 40 mL de agua destilada, esto equivale a concentraciones de 0,025% y 0,050% en peso, respectivamente. Se depositan 10 mL de cada muestra en placas de Petry de radio 28 mm y altura de 13,4 mm, se calentaron con un diodo láser (Newport, modelo 525B - $k = 829,1$ nm) a una corriente de 2200 mA, correspondiente a una potencia de calentamiento (P) de 0,327 W en el rango del infrarrojo cercano durante 100 s. La potencia del láser se calibró previamente y exhibió un perfil

Gaussiano con 11 mm de diámetro (tarjeta láser THORLABS VRC2). Las temperaturas de las superficies de las muestras se midieron con una cámara de infrarrojos (FLIR SC660) (Bermeo, y otros, 2020).

3.2 Simulación numérica

La simulación numérica fue implementada por el método de elementos finitos por medio del software COMSOL Multiphysics® con condiciones iguales al experimento. i.e., transferencia de calor por inducción de laser determinado por las ecuaciones (1-2) sobre una geometría en 3D como se presenta en la Figura 1. Las condiciones de contorno se describen en las ecuaciones (7-7a), donde $T_0 = 24,8 \text{ }^\circ\text{C}$, y $h_\infty=10 \text{ W/m}^2 \cdot \text{K}$. La temperatura se obtiene en un punto específico del modelo $\{10; 0; 4\} \text{ mm}$, con 165495 elementos disponibles cada segundo, durante 97 s. Los parámetros son considerados como exactos como se describen en la Tabla 1 considerando que no hay perfusión sanguínea ($\omega_b = 0$) y fuente de calor metabólico ($Q_m = 0$). Este modelo se considera como el modelo completo para describir el perfil de temperatura y base para la construcción de un modelo reducido.

$$T(x, z, t) = T_0 \quad (7)$$

$$k(x, z, t) \frac{\partial T(x, z, t)}{\partial x \partial z} + h_\infty T(x, z, t) = h_\infty T_0, \quad x = d, \quad z = d, \quad t > 0 \quad (7a)$$

Figura 1. Dominio del Sistema

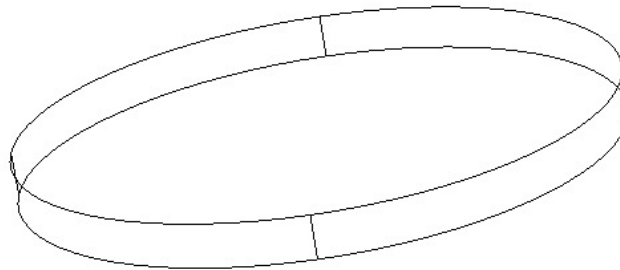


Tabla 1. Parámetros del Sistema

Parámetros	Valor exacto	Dimensión	Descripción
c_p	4184,10	J / (Kg. K)	Calor específico del agua
ρ	998,20	Kg/m ³	Densidad del agua
k	0,6034	W / (m.K)	Conductividad térmica del agua
κ	1,96	m ⁻¹	Coefficiente de absorción del agua
κ_1	13,72	m ⁻¹	Coefficiente de absorción del nanofluido al 0,025%
κ_2	45,08	m ⁻¹	Coefficiente de absorción del nanofluido al 0,050%
ν	1,0017x10 ⁻³	Pa.s	Viscosidad dinámica del agua

Fuente: (Bermeo et al, 2020)

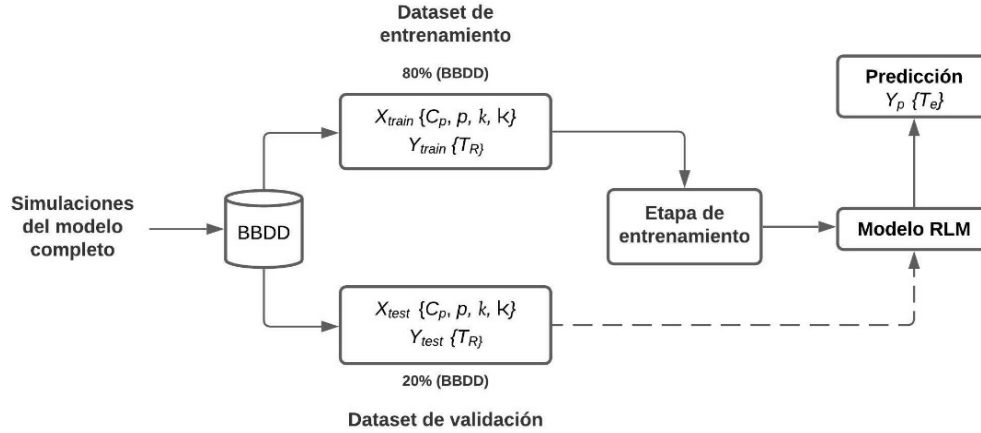
3.3 Modelo de Regresión Lineal Múltiple

Para el entrenamiento del modelo de Regresión Lineal Múltiple el conjunto de datos fue obtenido a través de simulación del modelo descrito en la sección anterior 200 veces con cambios en sus parámetros. Se consideró entonces una distribución Gaussiana con promedio centrado en los parámetros exactos para c_p , ρ y k con una desviación estándar de 10% y el parámetro κ se determina con una distribución normal entre $1.96 - 90.00 \text{ m}^{-1}$ para un total de 4 variables de entrada. Para la salida se consideró la temperatura disponible en un solo punto de medición durante 97s. Así, tenemos una base de datos correspondiente a 4 variables de entrada y una variable de salida disponible cada segundo durante 97s.

Para la construcción del modelo de Regresión Lineal Múltiple para predecir la temperatura transitoria de salida, se realizó

a través del lenguaje de programación Python utilizando la librería scikit-learn. En la Figura 2, se presenta el diagrama de bloques de la estructura del modelo de entrenamiento, donde T_R es la temperatura de referencia y T_e es la salida predicha. La etapa de entrenamiento del modelo se realizó utilizando 160 simulaciones correspondientes al 80% del conjunto de datos generados. Después de entrenar el modelo se evalúa la capacidad predictiva empleando la etapa de validación utilizando 40 simulaciones correspondientes al 20% restante del conjunto de datos. Después de la etapa de validación se procede a calcular el coeficiente de determinación R^2 y el margen de error asociado, lo cual refleja la diferencia existente entre los valores reales y estimados. Para el caso de este estudio, el coeficiente de determinación R^2 es una medida que representa la proporción de varianza en el campo de temperatura como variable dependiente predecible a partir de las variables de entrada o independientes.

Figura 2. Diagrama de bloques del modelo RLM



Fuente: Propia.

Una vez el modelo es validado, se realiza la simulación con los parámetros exactos y se compara su respuesta con la producida por COMSOL Multiphysics® y con las medidas experimentales, y se determina la correlación del modelo.

3.4 Algoritmos de optimización

3.4.1 Análisis de sensibilidad

Se realiza el análisis de sensibilidad permite entender cómo la temperatura de salida en el proceso de la hipertermia está influenciada por cambios pequeños y repentinos de los valores de los parámetros que se desean estimar o predecir. A través de la aplicación de la ecuación (8) se pueden obtener los valores de análisis de sensibilidad.

$$\mathbf{J}_{ij} = \left[\frac{\partial T_i(\mathbf{P})}{\partial P_j} \right] \cong \frac{T_i(P_1, P_2, \dots, P_i + \varepsilon P_i, \dots, P_n) - T_i(P_1, P_2, \dots, P_i - \varepsilon P_i, \dots, P_n)}{2\varepsilon P_i} \quad (8)$$

donde \mathbf{J} es la matriz de sensibilidad temporal en i y para cada parámetro j , \mathbf{P} es el vector de parámetros para el análisis de dependencia y ε es la perturbación de cada parámetro y se representa en el 10% de su valor nominal designado (Kaipio & Somersalo, 2005).

Con el objetivo de conocer la relación que existe entre la salida del modelo propuesto con cada uno de los parámetros de entrada los cuales corresponden a: calor específico (c_p), densidad (ρ), conductividad térmica (κ) y el coeficiente de absorción (κ), se calcula la matriz de sensibilidad \mathbf{J} para los parámetros y a partir del resultado se podrá identificar el nivel de sensibilidad a los cambios y/o dependencia lineal entre cada uno de los parámetros antes mencionados para su posterior interpretación.

3.4.2 Estimación de Parámetros

Se ejecuta la estimación de parámetros 50 veces por medio de los algoritmos de Gauss – Newton (GN) y Levenberg-Marquardt (LM) utilizando el modelo de Regresión Lineal Múltiple. Con los datos obtenidos se realizó un test de normalidad de los resultados por medio de Shapiro-Wilk (Shapiro & Wilk, 1965) para determinar si se cumple la hipótesis

nula que determina que los datos provienen de una distribución normal o la hipótesis alternativa que indica que no es una distribución normal, con un nivel de significancia de 0,05. Si los datos corresponden a una distribución normal, el resultado de los parámetros será el promedio de las ejecuciones, y se incluyen la desviación estándar y el intervalo de confianza del 95% mediante la prueba de t-student, de lo contrario se debe hacer un análisis sobre el resumen de los datos, estos son mediana, moda, y cuartiles.

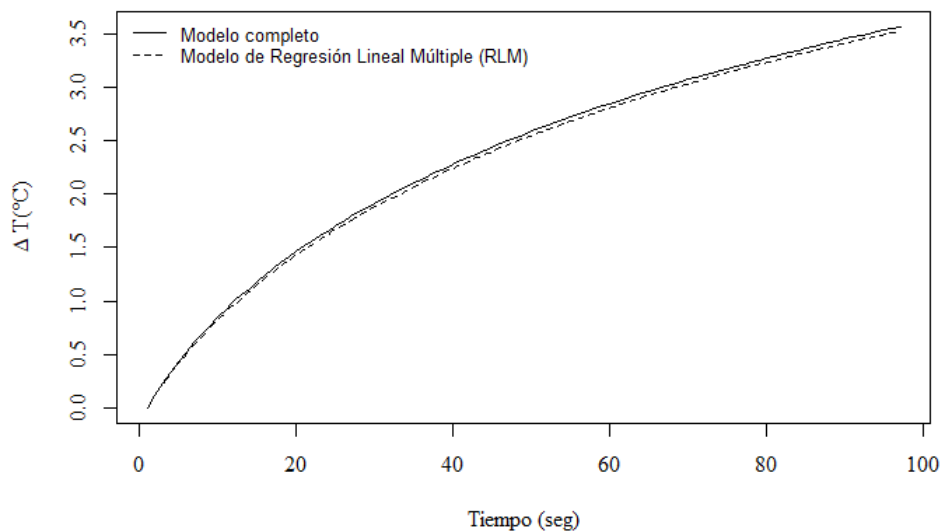
El desarrollo del modelo predictivo propuesto se llevó a cabo en un computador portátil Lenovo, con procesador Intel Core i5, RAM de 4 GB y sistema operativo Windows 10. Todos los resultados estadísticos y métodos aplicados se implementaron con el lenguaje de programación Python y el análisis estadístico se realizó en el lenguaje de programación R.

4 RESULTADOS Y DISCUSIÓN

4.1 Modelo de Regresión Lineal Múltiple

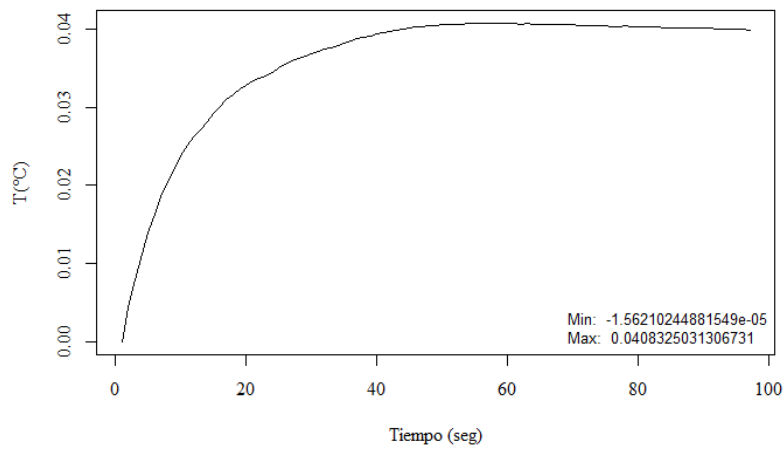
En la Figura 3 se presentan el delta de temperatura del modelo de regresión lineal múltiple (RLM) y la simulación numérica con el modelo completo. Nótese que las curvas de perfil de temperatura siguen la misma tendencia, lo que indica que el modelo RLM representa adecuadamente la dinámica del sistema. En la Figura 4 se observa el error absoluto entre el campo de temperatura de referencia y el estimado por RLM, obtenido por la diferencia entre la temperatura de salida del modelo completo y el modelo RLM durante 97 s en una sola muestra, observando un rango de $\{-1,56 \times 10^{-5}; 0,04\}$. En la Figura 5 se presenta el rango del error absoluto total entre las 200 muestras, se observa un rango de error de $\{-0,48; 0,29\}$ y con un error cuadrático medio (RMSE) = $7,09 \times 10^{-3}$, esos resultados indican que la respuesta del modelo completo y la predicción del modelo RLM son muy cercanos entre sí. En cuanto al coeficiente de determinación R^2 se obtiene un valor de 0,991 para la etapa de entrenamiento y de 0,987 en el proceso de validación. Se observa que el coeficiente de determinación R^2 es cercano a 1, lo que indica una buena predicción de la variable respuesta sin tanta variabilidad en la temperatura estimada. Los tiempos de cada etapa también son significativamente cortos en el proceso de entrenamiento fue de 78,67 ms y en el proceso de validación de 1,02 ms, esto es para un set de parámetros.

Figura 3. Temperatura de salida de una simulación aleatoria real vs. su estimación mediante el modelo RLM



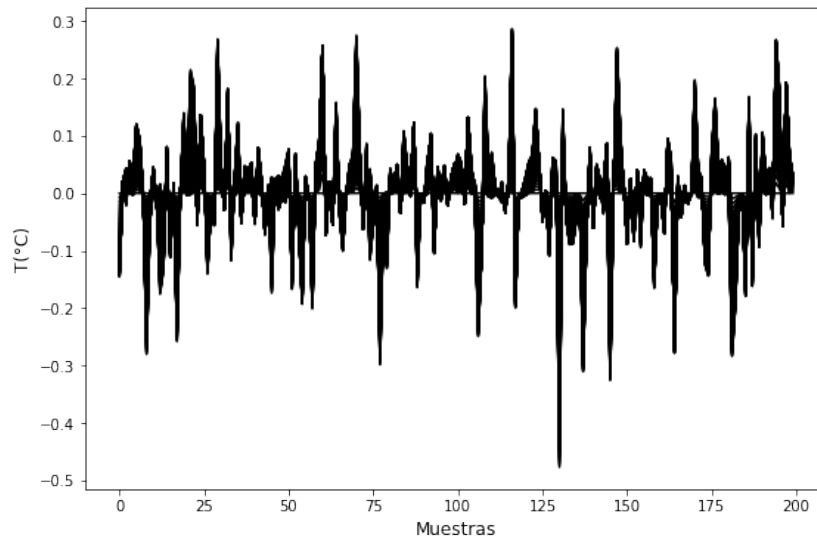
Fuente: Propia

Figura 4. Error absoluto. Diferencia entre el campo de temperatura de salida de la simulación numérica y el estimado mediante el modelo RLM



Fuente: Propia

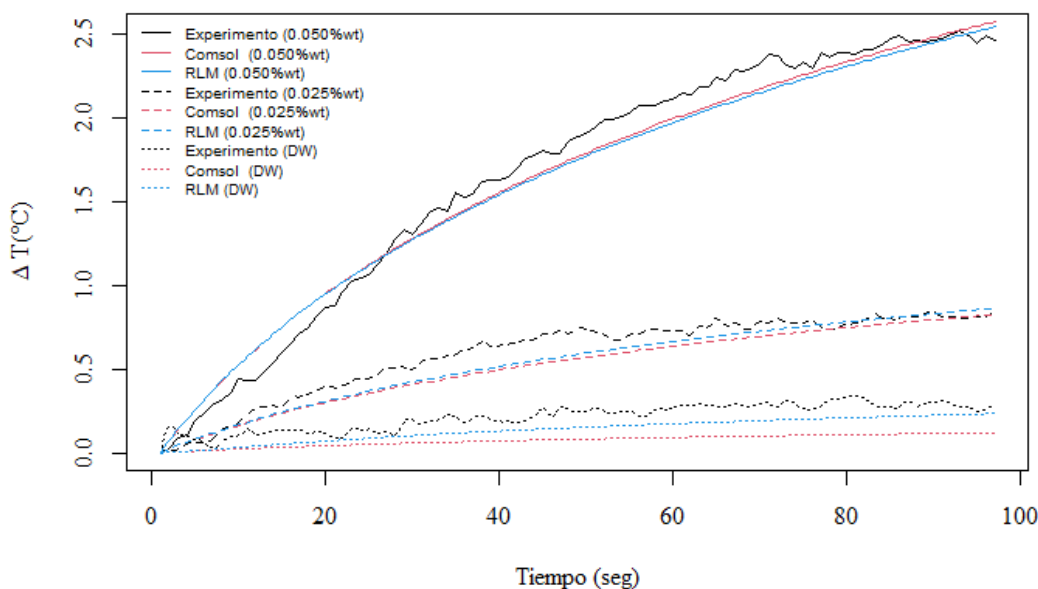
Figura 5. Error absoluto de las 200 muestras



Fuente: Propia

En la Figura 6, se representa la comparación entre las estimaciones del modelo de regresión lineal múltiple, la simulación numérica y los resultados experimentales de (Bermeo et al., 2020) con respecto al tiempo, para el agua destilada (DW) y los dos nanofluidos con concentraciones de 0,025% y 0,050% en proporción de peso (0,025%wt y 0,050%wt), variando únicamente el valor del coeficiente de absorción del modelo propuesto (κ) los cuales fueron 1,96, 13,72 y 45,08, respectivamente. Es de notarse, que los modelos matemáticos tienen una tendencia igual, lo que indica que el modelo RLM representa adecuadamente el problema físico. En cuanto a las medidas experimentales ambos modelos intentan representar la dinámica del sistema. Sin embargo, existen variaciones debido a que no se tiene un conocimiento apropiado de los parámetros del sistema, sobre todo del coeficiente de absorción, solo se tiene buen conocimiento del coeficiente de absorción del agua y no de los nanofluidos.

Figura 6. Temperatura de salida de los resultados experimentales, simulación numérica y predicción del modelo de Regresión Lineal Múltiple



Fuente: Propia

En la Tabla 2, se presentan los coeficientes de correlación entre el modelo RLM y los resultados experimentales, entre el modelo RLM y el modelo completo y entre los resultados experimentales y el modelo completo. Los resultados indican que el campo de temperatura predicho por el modelo RLM está muy cerca de los valores de salida del modelo matemático y las mediciones experimentales, i.e., una correlación de 0,99 con respecto al modelo realizado con COMSOL Multiphysics®. Respecto al experimento, la correlación fue de 0,93, 0,97 y 0,99 correspondientes al agua destilada y los dos nanofluidos en concentraciones de 0,025% y 0,050%, respectivamente. Estos valores muestran la alta eficiencia del modelo propuesto en este estudio. Adicionalmente, la correlación es similar a la determinada entre el modelo completo y las mediciones experimentales, esto significa que el modelo de RLM realiza una predicción del campo de temperatura con buena calidad. Con respecto al tiempo computacional, el modelo RLM presenta una reducción considerable del tiempo de ejecución computacional. La predicción del modelo de RLM en promedio fue de 1,95 ms en comparación con el modelo completo con COMSOL Multiphysics es de 25,29 minutos por una simulación, reduciendo notablemente el tiempo de ejecución.

Tabla 2. Coeficientes de correlación de las tres muestras de estudio

Ítem	Coeficiente de correlación		
	Agua destilada	Nanofluido al 0.025%	Nanofluido al 0.050%
RLM y los datos experimentales	0,927	0,971	0,994
RLM y el modelo completo (Comsol Multiphysics®)	0,999	0,999	0,999
Datos experimentales y el modelo completo	0,926	0,972	0,994

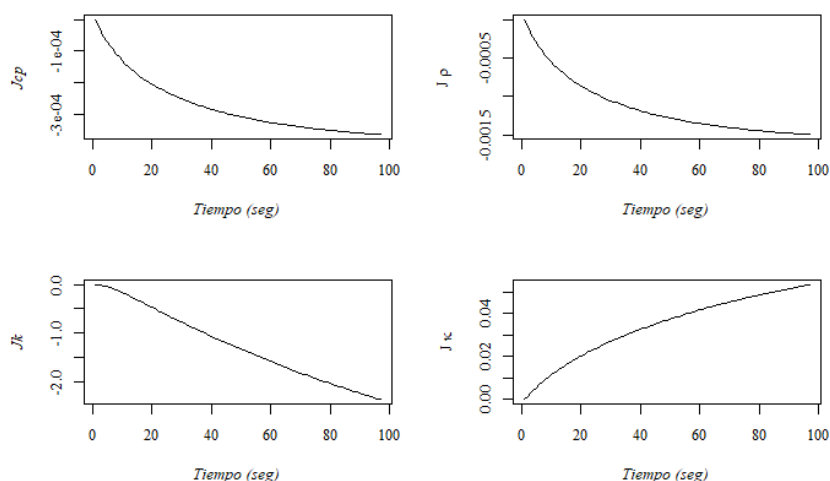
Fuente: Propia

4.2 Métodos de optimización

4.2.1 Análisis de sensibilidad

En la Figura 7, se muestran las funciones de los coeficientes de sensibilidad con respecto al calor específico del agua (c_p), densidad del agua (ρ), conductividad térmica del agua (k) y el coeficiente de absorción del agua (κ) y así comprender el efecto que tienen estos parámetros sobre los perfiles de temperatura. Estas funciones evidencian que el parámetro k es independiente a los demás parámetros presentes, ya que no sigue una tendencia similar con respecto al resto. Los parámetros c_p , ρ y κ son linealmente dependientes, lo que indica que podría ser estimado sólo uno de ellos. Nótese que el parámetro con mayor sensibilidad corresponde al coeficiente de absorción del agua (κ). Así, el análisis de coeficientes de sensibilidad sugiere que es posible estimar la conductividad térmica (k) y uno de los tres parámetros dependientes, en este caso el que presenta mayor sensibilidad y que corresponde al coeficiente de absorción (κ).

Figura 7. Coeficientes de sensibilidad de los parámetros del modelo de Regresión Lineal Múltiple.



Fuente: Propia

4.2.2 Algoritmos de Gauss – Newton y Levenberg – Marquardt

En la Tabla 3, se muestran los resultados de la prueba de normalidad de los datos implementando Shapiro-Wilk, donde se puede evidenciar que en todos los casos el valor de la probabilidad (*p – value*) es mayor al nivel de significancia de 0,05, lo que indica que no se rechaza la hipótesis nula que dice que la distribución de los datos es normal.

Tabla 3. Resultados de la prueba de normalidad por medio de Shapiro Wilk.

Concentración	Algoritmo	<i>p – value</i>	
		<i>k</i>	<i>κ</i>
Agua destilada	Levenberg - Marquardt	0,8089	0,8085
	Gauss - Newton	0,9191	0,9210
Nanofluido al 0,025%	Levenberg - Marquardt	0,4422	0,4419
	Gauss - Newton	0,9033	0,6449
Nanofluido al 0,050%	Levenberg - Marquardt	0,5631	0,5628
	Gauss - Newton	0,0672	0,0815

Fuente: Propia

En la Tabla 4, se presenta un resumen estadístico de las 50 ejecuciones realizadas para la estimación de *k* y *κ* para cada concentración con un criterio de parada de 10^{-5} del valor de verosimilitud. Para esto, se determina un valor inicial aleatorio para cada parámetro en cada ejecución entre un rango del $\pm 10\%$ del valor exacto (ver Tabla 1). Se presenta el promedio de la estimación de cada parámetro para cada algoritmo, la desviación estándar y el tiempo de ejecución del proceso de estimación. En la Tabla 5, se presentan los límites del rango del intervalo de confianza de cada parámetro estimado para el agua destilada y los dos nanofluidos con un nivel de confianza del 95%. Nótese que el valor estimado para la conductividad térmica (*k*) es muy cercano al valor de literatura, en el caso del coeficiente de absorción (*κ*) se encuentran valores que define bien la dinámica del perfil de temperatura y las medidas experimentales. Esto se confirma con el valor de la desviación estándar la cual está en el orden de 10^{-7} máximo, lo que indica que los resultados de cada proceso de estimación tienen un rango de dispersión bastante bajo, en cuanto a este factor el algoritmo de GN presenta una menor dispersión de los datos toda vez que la desviación estándar está en el rango de 10^{-12} máximo. En cuanto al intervalo de confianza se observa que los valores de los parámetros de todas estimaciones se encuentran en un rango con un margen pequeño. Sin embargo, el valor considerado como exacto en este trabajo referente al parámetro *k*, se encuentra por fuera de este rango. En ese sentido, los resultados de este trabajo contribuyen al conocimiento que se tiene de los parámetros, teniendo en cuenta que el valor considerado como exacto también es un valor estimado.

Tabla 4. Resumen estadístico entre algoritmos de optimización.

Concentración	Algoritmo	Número de ejecuciones	Promedio parámetro k	Promedio parámetro κ	Desviación estándar parámetro k	Desviación estándar parámetro κ	Tiempo promedio para cada ejecución (segundos)
Agua destilada	Levenberg - Marquardt	50	0,70320972785609769939	7,63155807835350152857	$1,48 \times 10^{-8}$	$5,60 \times 10^{-7}$	$5,59 \times 10^{-4}$
	Gauss – Newton	50	0,70320972786140789612	7,63155807855394119343	$4,07 \times 10^{-14}$	$1,55 \times 10^{-12}$	$1,80 \times 10^{-4}$
Nanofluido al 0.025%	Levenberg - Marquardt	50	0,83597121910393989985	23,65479895426072687314	$1,14 \times 10^{-12}$	$4,31 \times 10^{-11}$	$5,18 \times 10^{-4}$
	Gauss – Newton	50	0,83597121911432092922	23,65479895465262316634	$2,01 \times 10^{-14}$	$7,26 \times 10^{-13}$	$1,80 \times 10^{-4}$
Nanofluido al 0.050%	Levenberg - Marquardt	50	0,51803547225598123305	43,74309725011383420679	$2,09 \times 10^{-8}$	$7,90 \times 10^{-7}$	$5,05 \times 10^{-4}$
	Gauss – Newton	50	0,51803547225248391950	43,74309724998185089362	$5,77 \times 10^{-14}$	$2,17 \times 10^{-12}$	$1,59 \times 10^{-4}$

Fuente: Propia

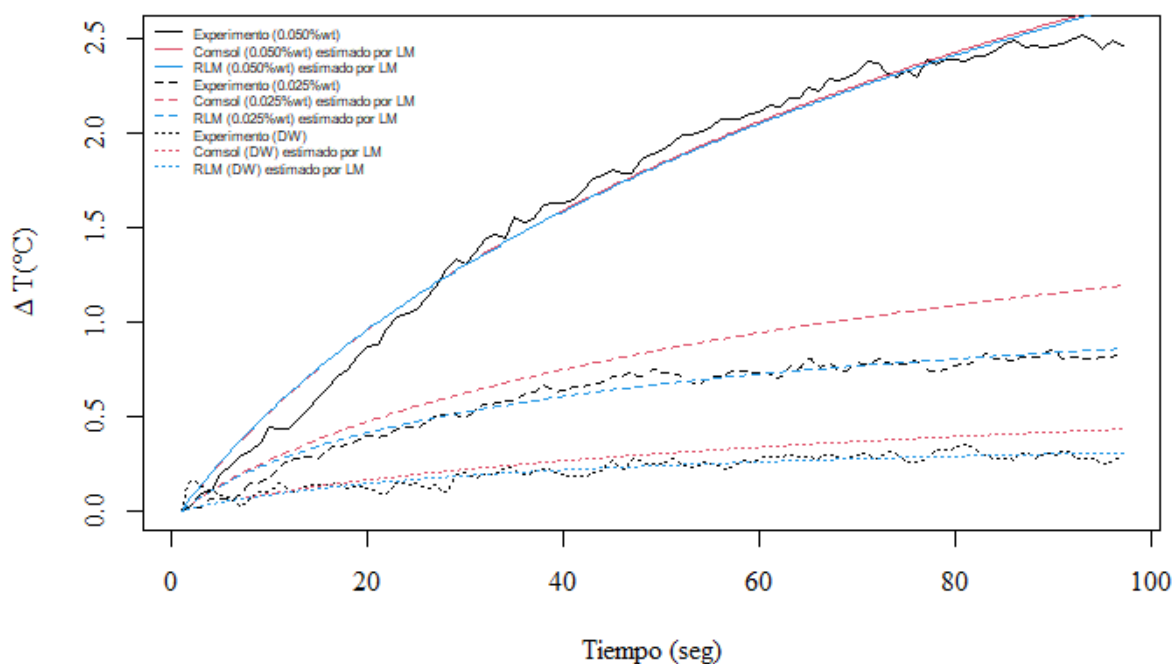
Tabla 5. Intervalo de confianza del 95%

Concentración	Algoritmo	Intervalo de confianza (k)	Intervalo de confianza (κ)
Agua destilada	Levenberg - Marquardt	0,70320972785573743202	7,63155807833990085243
		0,70320972785645796677	7,63155807836710220471
	Gauss – Newton	0,70320972786139568367	7,63155807855347667612
		0,70320972786142010857	7,63155807855440571075
Nanofluido al 0.025%	Levenberg - Marquardt	0,83597121910351568363	23,65479895424471123988
		0,83597121910436411607	23,65479895427674250641
	Gauss – Newton	0,83597121911431504504	23,65479895465239579266
		0,83597121911432681340	23,65479895465285054001
Nanofluido al 0.050%	Levenberg - Marquardt	0,51803547225508395080	43,74309725007995552915
		0,51803547225687851530	43,74309725014771288443
	Gauss – Newton	0,51803547225247381647	43,74309724998147430597
		0,51803547225249402253	43,74309724998222748127

Fuente: Propia

En la Figura 8, se presenta la comparación de la temperatura de salida entre el modelo de Regresión Lineal Múltiple y el modelo completo con los parámetros estimados por medio de los algoritmo de LM, y las mediciones experimentales de (Bermeo, y otros, 2020). Se observa que la respuesta del modelo lineal y las medidas experimentales están muy cercanos entre sí, en todos los casos, resultados coherentes debido al que el proceso de estimación se realizó precisamente con el modelo lineal. Sin embargo, si los parámetros estimados sirven de entrada en el modelo completo se observa que existe una diferencia, esto quiere decir que no hay un ajuste adecuado a los valores del modelo completo y medidas experimentales, a excepción del nanofluido al 0,050% de proporción de peso (0,050%wt). Para el caso de 0,025% de proporción de peso (0,025%wt) y el agua destilada pura (DW), se observa que los datos estimados para el parámetro k están por fuera de los datos considerados en el entrenamiento. Esto sugiere que el modelo puede mejorar si se amplía los datos de entrenamiento y paralelamente disminuir el rango del parámetro k . Es importante mencionar que en la Figura 8 solo se considera la estimación con LM debido a que el resultado obtenido con GN es considerado igual, por lo tanto, no se aprecia una diferencia notable en la predicción de la temperatura.

Figura 8. Temperatura de salida del modelo de Regresión Lineal Múltiple y modelo completo con parámetros estimados y las mediciones experimentales para cada concentración



En el proceso de estimación de parámetros, se encuentran valores adecuados que reproducen adecuadamente el perfil de temperatura sobre un modelo de transferencia de calor. de los Ríos, Bermeo & Pereira, (2020) y López, Serna, Bermeo & Castillo, (2020) realizaron procesos similares también con modelos de transferencia de calor, donde la diferencia radica en la fuente de calor. Si bien los datos encontrados en el proceso de estimación son adecuados, ellos utilizaron en estos procesos tiempo de ejecución bastante grandes debido a que el modelo directo es simulado en Comsol Multiphysics®. En el proceso propiamente dicho, los autores utilizan medidas simuladas, en este trabajo se utilizó medidas experimentales, lo que permite establecer parámetros más cercanos a la realidad. Así, el aporte principal en este trabajo es el modelo de regresión lineal múltiple para predecir la temperatura en el proceso de hipertermia con láser utilizando nanopartículas y la determinación de parámetros del sistema que permiten tener un mejor conocimiento del problema, para establecer protocolos para la implementación de este tipo de terapias.

5 CONCLUSIONES

En este artículo se realizó un estudio para la estimación del campo de temperatura de un modelo reducido propuesto implementando el algoritmo de Regresión Lineal Múltiple (RLM), el cual su respuesta se acerca notablemente a los resultados del modelo completo con un error cuadrático medio (RMSE) = $7,09 \times 10^{-3}$, esto indica que el modelo RLM puede ser usado sustituyendo el modelo completo con una diferencia aceptable. En cuanto al tiempo computacional se obtiene 1,95 ms por ejecución en comparación con el modelo completo con un tiempo por ejecución de 25,29 minutos, lo cual es muy beneficioso para problemas de optimización ya que se puede ejecutar en múltiples ocasiones, obteniendo buenos resultados por ejemplo en procesos de estimación de parámetros del modelo, lo que permite planificar la terapia de la hipertermia antes y durante su implementación.

En el proceso de estimación de parámetros se realizó un estudio de optimización del modelo, estimando los parámetros de entrada que mayor influencia tienen en el mismo, a través de los métodos de Gauss – Newton y Levenberg – Marquardt. Con el método de Gauss – Newton, se tuvo un mejor rendimiento con menos dispersión en los datos en 50 ejecuciones. En ambos métodos se obtienen valores estimados adecuados con un intervalo de confianza pequeño.

Se observa que la respuesta del modelo RLM para la concentración de nanopartículas de óxido de hierro correspondiente a 0,025% en peso, sugiere un conjunto más amplio de entrenamiento para estimar la temperatura de salida de esa concentración, dado que el intervalo de confianza y los valores propuestos de cada algoritmo de optimización para la

conductividad eléctrica (κ) están por fuera del rango de entrenamiento.

El análisis de sensibilidad estableció que se pueden estimar dos parámetros de entrada del modelo: conductividad térmica del agua (k) el cual es linealmente independiente y el coeficiente de absorción (κ) que presentan una mayor sensibilidad con respecto a la densidad (ρ) y el calor específico (c_p). El análisis estadístico de las 50 ejecuciones de los algoritmos LM y GN, indican que los resultados de las estimaciones corresponden a una distribución normal.

Finalmente, en estudios posteriores, se podrían implementar otros tipos de algoritmos de aprendizaje automático para evaluar la efectividad de estos y realizar la comparación con el modelo de Regresión Lineal Múltiple propuesto en este trabajo, que permitan establecer protocolos para la implementación de la terapia en un entorno clínico y otros procesos de estimación que considere el total de parámetros.

AGRADECIMIENTOS

Los autores expresan su agradecimiento a la Dirección General de Investigación (DGI) y al proyecto “Optimización en tratamientos para el cáncer por medio de hipertermia con inducción de ondas electromagnéticas” código DGI: 819-621120-1767 de la Universidad Santiago de Cali.

REFERENCIAS

- Amat, J. (2016). *Introducción a la Regresión Lineal Múltiple*. https://www.cienciadedatos.net/documentos/25_regresion_lineal_multiple.html
- Amat, J. (2020). *Regresión lineal con Python*. <https://www.cienciadedatos.net/documentos/py10-regresion-lineal-python.html>
- Asgari, S., Scalzo, F., & Kasprowicz, M. (2019). Pattern Recognition in Medical Decision Support. *BioMed Research International*, 2019. <https://doi.org/10.1155/2019/6048748>
- Bélanger, M., El-Jabi, N., Caissie, D., Ashkar, F., & Ribí, J. M. (2005). Water temperature prediction using neural networks and multiple linear regression. *Revue Des Sciences de l'Eau*, 18(3), 403–421. <https://doi.org/10.7202/705565ar>
- Bermeo, L. A., Orlande, H. R., & Eliçabe, G. E. (2015). Estimation of state variables in the hyperthermia therapy of cancer with heating imposed by radiofrequency electromagnetic waves. *International Journal of Thermal Sciences*, 98, 228–236. <https://doi.org/10.1016/j.ijthermalsci.2015.06.022>
- Bermeo, L. A., Loiola, B. R., Da Silva Abreu, L. A., Lamien, B., Pereira da Silva, N., Orlande, H. R. B., & Dos Santos, D. S. (2020). Thermal Effect by Applying Laser Heating in Iron Oxide Nanoparticles Dissolved in Distilled Water. *IFMBE Proceedings*, 76, 1239–1245. https://doi.org/10.1007/978-3-030-31635-8_151
- Bermeo, L. A., Orlande, H. R. B., & Eliçabe, G. E. (2016). Combined parameter and state estimation in the radio frequency hyperthermia treatment of cancer. *Numerical Heat Transfer, Part A: Applications*, 70(6), 581–594. <https://doi.org/10.1080/10407782.2016.1193342>
- Bondarev, N. V. (2019). Artificial Neural Network and Multiple Linear Regression for Prediction and Classification of Sustainability of Sodium and Potassium Coronates. *Russian Journal of General Chemistry*, 89(7), 1438–1446. <https://doi.org/10.1134/S1070363219070144>
- Bruno, A. B., Maurente, A., Lamien, B., & Orlande, H. R. B. (2017). Numerical simulation of nanoparticles assisted laser photothermal therapy: a comparison of the P1-approximation and discrete ordinate methods. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 39(2), 621–630. <https://doi.org/10.1007/s40430-016-0553-3>
- Cornejo Zúñiga, O., & Rebolledo Vega, R. (2016). Estimación De Parámetros En Modelos No Lineales: Algoritmos Y Aplicaciones. *Revista ELA*, 13(25), 81–98. <https://doi.org/10.24050/reia.v13i25.1019>
- Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of Artificial Intelligence in Machine Learning: Review and Prospect. *International Journal of Computer Applications*, 115(9), 31–41. <https://doi.org/10.5120/20182-2402>
- de los Rios, L., Bermeo Varon, L., & Pereira, A. (2020). Comparison of Attenuation Coefficient Estimation in High Intensity Focused Ultrasound Therapy for Cancer Treatment by Levenberg Marquardt and Gauss-Newton Methods. *Communications in Computer and Information Science*, 1194, 108–118. <https://doi.org/10.1007/978-3-030-42520-3>
- Delbari, M., Sharifazari, S., & Mohammadi, E. (2019). Modeling daily soil temperature over diverse climate conditions in Iran—a comparison of multiple linear regression and support vector regression techniques. *Theoretical and Applied Climatology*, 135(3–4), 991–1001. <https://doi.org/10.1007/s00704-018-2370-3>
- Dokht Shakibjoo, A., Moradzadeh, M., Moussavi, S. Z., Mohammadzadeh, A., & Vandeveld, L. (2021). Load frequency

- control for multi-area power systems: A new type-2 fuzzy approach based on Levenberg–Marquardt algorithm. *ISA Transactions*, xxx: <https://doi.org/10.1016/j.isatra.2021.03.044>
- Dutta, J., & Kundu, B. (2020). Exact Analytical Formulation of Three-Dimensional Pennes Bioheat Model in Regional Hyperthermia with Modified Initial Condition. *Journal of The Institution of Engineers (India): Series C*, 101(2), 271–279. <https://doi.org/10.1007/s40032-019-00534-w>
- Fanjul-Vélez, F., Romanov, O. G., & Arce-Diego, J. L. (2009). Efficient 3D numerical approach for temperature prediction in laser irradiated biological tissues. *Computers in Biology and Medicine*, 39(9), 810–817. <https://doi.org/10.1016/j.combiomed.2009.06.009>
- Fernández Slezak, D. (2010). *Estimación de parámetros en modelos biológicos complejos. Aplicación a modelos de crecimiento tumoral*. (Tesis de Doctorado). Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales. Buenos Aires, Argentina. http://hdl.handle.net/20.500.12110/tesis_n4766_FernandezSlezak
- García Moreno, A. I., González Barbosa, J. J., Hurtado Ramos, J. B., Ornelas Rodríguez, F. J., & Ramírez Pedraza, A. (2016). Análisis de la sensibilidad en un modelo de calibración cámara-LiDAR. *Revista Internacional de Metodos Numericos Para Calculo y Diseno En Ingenieria*, 32(4), 193–203. <https://doi.org/10.1016/j.rimni.2015.05.002>
- Gill, P. E., & Murray, W. (1978). Algorithms for the Solution of the Nonlinear Least-Squares Problem. *SIAM Journal on Numerical Analysis*, 15(5), 977–992. <http://www.jstor.org/stable/2156716>
- Giombini, A., Giovannini, V., Di Cesare, A., Pacetti, P., Ichinoseki-Sekine, N., Shiraishi, M., Naito, H., & Maffulli, N. (2007). Hyperthermia induced by microwave diathermy in the management of muscle and tendon injuries. *British Medical Bulletin*, 83(1), 379–396. <https://doi.org/10.1093/bmb/ldm020>
- Huang, C. H., & Huang, C. Y. (2007). An inverse problem in estimating simultaneously the effective thermal conductivity and volumetric heat capacity of biological tissue. *Applied Mathematical Modelling*, 31(9), 1785–1797. <https://doi.org/10.1016/j.apm.2006.06.002>
- Kaipio, J., & Somersalo, E. (2005). *Statistical and Computational Inverse Problems* (Vol. 160). Springer-Verlag. <https://doi.org/10.1007/b138659>
- Kumar, V., & Sahu, M. (2021). Evaluation of nine machine learning regression algorithms for calibration of low-cost PM2.5 sensor. *Journal of Aerosol Science*, 157(January), 105809. <https://doi.org/10.1016/j.jaerosci.2021.105809>
- Lamien, B., Barreto Orlande, H. R., & Enrique Eliçabe, G. (2017). Particle Filter and Approximation Error Model for State Estimation in Hyperthermia. *Journal of Heat Transfer*, 139(1). <https://doi.org/10.1115/1.4034064>
- Lamien, B., Bermeo, L. A., Orlande, H. R. B., & Eliçabe, G. E. (2017). State estimation in bioheat transfer: a comparison of particle filter algorithms. *International Journal of Numerical Methods for Heat & Fluid Flow*, 27(3), 615–638. <https://doi.org/10.1108/HFF-03-2016-0118>
- Lamien, B., Orlande, H., & Eliçabe, G. E. (2017). Inverse problem in the hyperthermia therapy of cancer with laser heating and plasmonic nanoparticles. *Inverse Problems in Science and Engineering*, 25(4), 608–631. <https://doi.org/10.1080/17415977.2016.1178260>
- Lamien, B., Orlande, H. R., Bermeo Varón, L. A., Queiroga Basto, R. L., Eliçabe, G. E., Silva dos Santos, D., & Machado Cotta, R. (2018). Estimation of the temperature field in laser-induced hyperthermia experiments with a phantom. *International Journal of Hyperthermia*, 35(1), 279–290. <https://doi.org/10.1080/02656736.2018.1496283>
- Lamien, B., Barreto, H.R., Bermeo Varón, L., Basto, R.L., Eliçabe, E., dos Santos, D.S., & Cotta, R. M. (2018). Estimation of the temperature field in laser-induced hyperthermia experiments with a phantom. *International Journal of Hyperthermia*, 35(1), 279–290. <https://doi.org/10.1080/02656736.2018.1496283>
- Larreur, M., Lamien, B., & Orlande, H. R. . (2019). Optimization of the hyperthermia treatment of a skin tumor containing nanoparticles. *VIII International Conference on Computational Methods for Coupled Problems in Science and Engineering*.
- López Perez, J. I., Serna Maldonado, R. D., Bermeo Varon, L. A., & Castillo García, J. F. (2020). Estimation of Electrical Conductivity from Radiofrequency Hyperthermia Therapy for Cancer Treatment by Levenberg Marquardt Method. *Communications in Computer and Information Science*, 1195 CCIS, 141–152. https://doi.org/10.1007/978-3-030-42531-9_12
- Mokarram, M. (2016). Modeling of multiple regression and multiple linear regressions for prediction of groundwater quality (case study: north of Shiraz). *Modeling Earth Systems and Environment*, 2(1), 1–7. <https://doi.org/10.1007/s40808-015-0059-5>
- Müller, A. C., & Guido, S. (2017). Introduction to with Python Learning Machine. In *Proceedings of the Speciality Conference on Infrastructure Condition Assessment: Art, Science, Practice*.

- Niemz, M. H. (2007). *Laser-Tissue Interactions*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-540-72192-5>
- Okajima, J., Maruyama, S., Takeda, H., & Komiya, A. (2009). Dimensionless solutions and general characteristics of bioheat transfer during thermal therapy. *Journal of Thermal Biology*, 34(8), 377–384. <https://doi.org/10.1016/j.jtherbio.2009.08.001>
- Pacheco, C. C., Orlande, H. R. B., Colaço, M. J., Dulikravich, G. S., Varón, L. A. B., & Lamien, B. (2020). Real-time temperature estimation with enhanced spatial resolution during MR-guided hyperthermia therapy. *Numerical Heat Transfer, Part A: Applications*, 77(8), 782–806. <https://doi.org/10.1080/10407782.2020.1720409>
- Palazzi, M., Maluta, S., Dall'Oglio, S., & Romano, M. (2010). The role of hyperthermia in the battle against cancer. *Tumori*, 96(6), 902–910. <https://doi.org/10.1177/548.6507>
- Paula, M., Díaz, R., David, J., Carmona, P., & Mir, J. (2021). *Estimación de parámetros para la quimioterapia en cáncer mediante el algoritmo de Levenberg-Marquardt*. (Tesis de Pregrado). Universidad Santiago de Cali, Facultad de Ingeniería, Bioingeniería. Cali, Colombia.
- Ram Kumar, R. P., Polepaka, S., Lazarus, S. F., & Krishna, D. V. (2019). An insight on machine learning algorithms and its applications. *International Journal of Innovative Technology and Exploring Engineering*, 8(11 Special issue 2), 432–436. <https://doi.org/10.35940/ijitee.K1069.09811S219>
- Sani, M. F., Abeywardena, S., & Psomopoulou, E. (2020). *Towards Finger Motion Tracking* (Vol. 1). Springer International Publishing. <https://doi.org/10.1007/978-3-030-31635-8>
- Sazgarnia, A., Naghavi, N., Mehdizadeh, H., & Shahamat, Z. (2015). Investigation of thermal distribution for pulsed laser radiation in cancer treatment with nanoparticle-mediated hyperthermia. *Journal of Thermal Biology*, 47, 32–41. <https://doi.org/10.1016/j.jtherbio.2014.10.011>
- Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4), 591. <https://doi.org/10.2307/2333709>
- Singh, R., Das, K., & Mishra, S. C. (2014a). Laser-induced hyperthermia of nanoshell mediated vascularized tissue - A numerical study. *Journal of Thermal Biology*, 44(1), 55–62. <https://doi.org/10.1016/j.jtherbio.2014.07.001>
- Singh, R., Das, K., & Mishra, S. C. (2014b). Laser-induced hyperthermia of nanoshell mediated vascularized tissue – A numerical study. *Journal of Thermal Biology*, 44, 55–62. <https://doi.org/10.1016/j.jtherbio.2014.07.001>
- van der Zee, J. (2002). Heating the patient: a promising approach? *Annals of Oncology*, 13(8), 1173–1184. <https://doi.org/10.1093/annonc/mdf280>
- Wust, P., Hildebrandt, B., Sreenivasa, G., Rau, B., Gellermann, J., Riess, H., Felix, R., & Schlag, P. (2002). Hyperthermia in combined treatment of cancer. *The Lancet Oncology*, 3(8), 487–497. [https://doi.org/10.1016/S1470-2045\(02\)00818-5](https://doi.org/10.1016/S1470-2045(02)00818-5)