

IDENTIFICACIÓN DE LOS PARÁMETROS DE LA ECUACIÓN DIFERENCIAL DE  
DIFUSIÓN DE CALOR, BASADOS EN LOS DATOS OBTENIDOS DE UNA  
SIMULACIÓN DE LA VARIACIÓN DE TEMPERATURA, RESPECTO AL TIEMPO  
Y AL ESPACIO PARA UNA GEOMETRÍA (PLACA PLANA)

JEFREY JOANY GUEVARA HUERTAS

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS  
ESCUELA DE INGENIERÍA MECÁNICA  
BUCARAMANGA

2023

IDENTIFICACIÓN DE LOS PARÁMETROS DE LA ECUACIÓN DIFERENCIAL DE  
DIFUSIÓN DE CALOR, BASADOS EN LOS DATOS OBTENIDOS DE UNA  
SIMULACIÓN DE LA VARIACIÓN DE TEMPERATURA, RESPECTO AL TIEMPO  
Y AL ESPACIO PARA UNA GEOMETRÍA (PLACA PLANA)

JEFREY JOANY GUEVARA HUERTAS

Trabajo de grado para optar al título de  
Ingeniero Mecánico

Director

JABID EDUARDO QUIROGA MENDEZ

Doctor

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FÍSICO-MECÁNICAS  
ESCUELA DE INGENIERÍA MECÁNICA  
BUCARAMANGA

2023



## DEDICATORIA

*A Dios*

*A mi madre maravillosa que la amo, con su amor y comprensión, siempre ha sido un apoyo incondicional, exigente muchas veces, esperando la superación de sus hijos, orgullosa de que sean buenas personas y segura que van a ser buenos profesionales.*

*A mi padre, "mi querido viejo" que lo amo, ha sido otro gran apoyo fundamental y siempre me recuerda con su genio que hay que terminar*

*A mis hermanos, los cuales quiero con todo mi corazón que siempre voy a estar agradecido con la vida de que compartan conmigo los buenos y malos momentos*

*A mis abuelos que siempre pusieron su granito de arena, algunos mientras estuvieron conmigo y especialmente para don Hernán, que siempre he me ha dado su cariño.*

*A mi toda mi familia*

*A Carlitos y a Fercho porque gracias a ellos sigo aquí.*

*A todos mis amigos, que los llevo en el corazón, los admiro demasiado con los cuales he compartido alegrías, tristezas y una que otra cerveza, también de una u otra manera me han apoyado*

*JEFREY J GUEVARA HUERTAS*

## **AGRADECIMIENTOS**

A la UIS, y a la escuela de ingeniería mecánica por brindarme todas las bases de conocimiento.

A Jabid E Quiroga Méndez, ingeniero mecánico, director de este proyecto de grado.

A cada uno de mis profesores, de los cuales aprendí.

A Juan F. Maradey, ingeniero mecánico, que me enseñó a ser más dedicado y a esforzarme más.

## TABLA DE CONTENIDO

|   | pág. |
|---|------|
| INTRODUCCIÓN  | 18   |
| 1. TÉCNICAS PARA LA IDENTIFICACIÓN DE MATERIALES          | 20   |
| 1.1 ALGUNAS PROPIEDADES TERMICAS Y FISICAS .....          | 22   |
| 1.1.1 CONDUCTIVIDAD TÉRMICA. ....                         | 23   |
| 1.1.2 CALOR ESPECIFICO.....                               | 23   |
| 1.1.3 DENSIDAD.....                                       | 24   |
| 1.1.4 DIFUSIVIDAD. ....                                   | 25   |
| 1.2 JUSTIFICACIÓN .....                                   | 25   |
| 1.3 OBJETIVOS.....  | 26   |
| 1.3.1 OBJETIVO GENERAL.....                               | 26   |
| 1.3.2 OBJETIVOS ESPECÍFICOS.....                          | 27   |
| <br>  |      |
| 2. DATOS HISTÓRICOS DE TEMPERATURA                        | 28   |
| 2.1 MEDICION EXPERIMENTAL.....                            | 28   |
| 2.2 SIMULACIÓN.....                                       | 29   |
| 2.2.1 PDE TOOLBOX .....                                   | 29   |
| 2.2.2 SIMULACIÓN USANDO <i>PDE TOOLBOX</i> DE MATLAB..... | 34   |
| <br>  |      |
| 3. MODEL DISCOVERY  | 45   |
| 3.1 DATOS PROVENIENTES DE (EDO). ....                     | 47   |
| 3.2 DATOS PROVENIENTES DE (PDE).....                      | 50   |
| 3.2.1 DERIVADAS DE LOS DATOS.....                         | 52   |
| 3.2.2 LIBRERÍA DE TÉRMINOS.....                           | 55   |
| 3.2.3 DETERMINACIÓN DE LOS PARÁMETROS. ....               | 56   |

|       |   |     |
|-------|---|-----|
| 3.3   | APLICACIÓN DEL MODEL DISCOVERY A LOS DATOS SIMULADOS .....  | 58  |
| 3.3.1 | CÓDIGO BASE PARA LA APLICACIÓN DE LA REGRESIÓN <i>MODEL DISCOVERY</i> . .....                       | 62  |
| 3.3.2 | OPTIMIZACIÓN DEL CÓDIGO Y NUEVO ALGORITMO. ....   | 67  |
| 4.    | PRUEBAS Y DISMINUCIÓN DEL ERROR   | 74  |
| 4.1   | AUMENTO DE LA CANTIDAD DE NODOS.....  | 75  |
| 4.1.1 | TAMAÑO MÁXIMO DEL ELEMENTO 0.01m. ....  | 75  |
| 4.1.2 | TAMAÑO MÁXIMO DEL ELEMENTO 0.007m. ....   | 76  |
| 4.1.3 | TAMAÑO MÁXIMO DEL ELEMENTO 0.004m. ....   | 78  |
| 4.1.4 | TAMAÑO MÁXIMO DEL ELEMENTO 0.002m. ....   | 79  |
| 4.1.5 | TAMAÑO MÁXIMO DEL ELEMENTO 0.001m. ....   | 80  |
| 4.1.6 | ANÁLISIS DE DATOS Y RESULTADOS. ....  | 81  |
| 4.2   | COMPARACIÓN DE LOS RESULTADOS USANDO DIFERENCIAS<br>CENTRALES EQUIDISTANTES Y NO EQUIDISTANTES..... | 85  |
| 4.2.1 | TAMAÑO MÁXIMO DEL ELEMENTO 0.007m. ....   | 85  |
| 4.2.2 | TAMAÑO MÁXIMO DEL ELEMENTO 0.004m. ....   | 86  |
| 4.2.3 | TAMAÑO MÁXIMO DEL ELEMENTO 0.002m. ....   | 86  |
| 4.2.4 | ANÁLISIS DE RESULTADOS. ....  | 87  |
| 4.3   | ANÁLISIS DE LOS DATOS .....   | 91  |
| 4.3.1 | GRAFICAS DE LOS VECTORES.....   | 91  |
| 4.3.2 | ¿DE DÓNDE PROVIENEN LOS OUTLIERS (VALORES ATÍPICOS)? .....  | 94  |
| 4.3.3 | ¿QUÉ HACER CON LOS VALORES ATÍPICOS? .....  | 95  |
| 4.3.4 | ELIMINACIÓN DE LOS DATOS ATÍPICOS ( <i>OUTLIERS</i> ) .....   | 97  |
| 4.3.5 | RESULTADOS.....   | 98  |
| 5.    | CONCLUSIONES  | 102 |

6. RECOMENDACIONES

104

BIBLIOGRAFÍA

105

## LISTA DE FIGURAS

|   | Pág. |
|---|------|
| Figura 1. Apariencia superficial de algunos metales                                 | 21   |
| Figura 2. Forma lineal y cuadrática del elemento de malla para 2D                   | 32   |
| Figura 3. Forma lineal y cuadrática del elemento de malla para 3D                   | 33   |
| Figura 4. Geometría para simular  | 34   |
| Figura 5. Diagrama de flujo para la simulación                                      | 35   |
| Figura 6. Geometría creada en PDE Toolbox   | 40   |
| Figura 7. Malla de la placa   | 42   |
| Figura 8. Solución bidimensional al cabo de 10 s                                    | 43   |
| Figura 9. Diagrama de flujo para la aplicación del Model Discovery                  | 46   |
| Figura 10. Sistema ODE  | 49   |
| Figura 11. Sistema PDE  | 51   |
| Figura 12. Model discovery para PDE   | 58   |
| Figura 13. Arreglo bidimensional de puntos con coordenadas (x, y)                   | 59   |
| Figura 14. Snapshots del fenómeno   | 60   |
| Figura 15. Diagrama de flujo del código base para la aplicación del Model Discovery | 62   |
| Figura 16. Modelo de rectángulos para encontrar nodos en una misma línea            | 63   |
| Figura 17. Líneas verticales de nodos   | 64   |
| Figura 18. Líneas horizontales de nodos   | 64   |
| Figura 19. Diagrama de barras del resultado para el código base                     | 66   |
| Figura 20. Diagrama de flujo para algoritmo Model Discovery para PDE                | 67   |
| Figura 21. Tamaño de la geometría y distancia entre nodos de los bordes             | 69   |
| Figura 22. Parámetros del rectángulo de búsqueda                                    | 70   |
| Figura 23. Malla de tamaño máximo de elemento 0.01m                                 | 75   |
| Figura 24. Malla de tamaño máximo de elemento 0.007m                                | 77   |
| Figura 25. Malla de tamaño máximo de elemento 0.004m                                | 78   |

|  |    |
|--|----|
| Figura 26. Malla de tamaño máximo de elemento 0.002m   | 79 |
| Figura 27. Malla de tamaño máximo de elemento 0.001m   | 80 |
| Figura 28. Grafica del error para la primera parte del vector de derivadas temporales                            | 83 |
| Figura 29. Grafica del error para la segunda parte del vector de derivadas temporales                            | 84 |
| Figura 30. Grafica del porcentaje de error para la primera parte del vector derivadas temporales para DCE y DNE. | 89 |
| Figura 31. Grafica del porcentaje de error para la segunda parte del vector derivadas temporales para DCE y DNE  | 90 |
| Figura 32. Grafica de los vectores resultantes   | 92 |
| Figura 33. Histograma de los vectores  | 93 |
| Figura 34. Nodos causantes de los valores atípicos   | 94 |
| Figura 35. Detección paramétrica de valores atípicos, basados en curva de distribución normal.                   | 97 |
| Figura 36. Distribución de temperatura de las placas   | 99 |

## LISTA DE TABLAS

|  | Pág. |
|--|------|
| Tabla 1. Clasificación de EDP de segundo orden   | 30   |
| Tabla 2. Creación de un círculo PDE Toolbox  | 37   |
| Tabla 3. Creación de un polígono PDE Toolbox   | 37   |
| Tabla 4. Creación de un rectángulo PDE Toolbox   | 38   |
| Tabla 5. Creación de una elipse PDE Toolbox  | 38   |
| Tabla 6. Fórmulas de diferencias centrales de segundo orden                                | 52   |
| Tabla 7. Fórmulas de diferencias centrales de cuarto orden                                 | 53   |
| Tabla 8. Fórmulas de diferencias hacia adelante y hacia atrás de segundo orden             | 53   |
| Tabla 9. Resultados Hmax 0.01m, Pseudoinversa y DCE  | 76   |
| Tabla 10. Resultados Hmax 0.007m, Pseudoinversa y DCE                                      | 77   |
| Tabla 11. Resultados Hmax 0.004m, Pseudoinversa y DCE                                      | 79   |
| Tabla 12. Resultados Hmax 0.002m, Pseudoinversa y DCE                                      | 80   |
| Tabla 13. Resultados Hmax 0.001m, Pseudoinversa y DCE                                      | 81   |
| Tabla 14. Resultados para la primera parte del vector de derivas temporales                | 82   |
| Tabla 15. Resultados para la segunda parte del vector de derivas temporales                | 82   |
| Tabla 16. Porcentaje de error para la primera parte del vector derivadas temporales        | 82   |
| Tabla 17. Porcentaje de error para la segunda parte del vector derivadas temporales        | 83   |
| Tabla 18. Resultados Hmax 0.007m, Pseudoinversa y DC no equidistantes                      | 86   |
| Tabla 19. Resultados Hmax 0.004m, Pseudoinversa y DC no equidistantes                      | 86   |
| Tabla 20. Resultados Hmax 0.002m, Pseudoinversa y DC no equidistantes                      | 87   |
| Tabla 21. Resultados para la primera parte del vector de derivas temporales para DCE y DNE | 87   |

|  |     |
|--|-----|
| Tabla 22. Resultados para la segunda parte del vector de derivas temporales para DCE y DNE         | 88  |
| Tabla 23. Porcentaje de error para la primera parte del vector derivadas temporales para DCE y DNE | 88  |
| Tabla 24. Porcentaje de error para la segunda parte del vector derivadas temporales para DCE y DNE | 88  |
| Tabla 25. Causas comunes de los valores atípicos   | 96  |
| Tabla 26. Resultados quitando los Outliers de la data  | 98  |
| Tabla 27. Resultado para los dos arreglos sin test de Tukey  | 100 |
| Tabla 28. Resultados para los dos arreglos aplicando test de Tukey                                 | 100 |
| Tabla 29. Porcentaje de error para los dos arreglos sin test de Tukey                              | 101 |
| Tabla 30. Porcentaje de error de los dos arreglos aplicando test de Tukey                          | 101 |

## LISTA DE ECUACIONES

|  |    |
|--|----|
| Ecuación 1. Definición de conductividad térmica                          | 23 |
| Ecuación 2. Definición de calor específico                               | 24 |
| Ecuación 3. Definición de densidad                                       | 24 |
| Ecuación 4. Definición de difusividad térmica                            | 25 |
| Ecuación 5. Ecuación escalar para resolver con PDE                       | 30 |
| Ecuación 6. Ecuación de valores propios a resolver con PDE               | 30 |
| Ecuación 7. Ecuación de valores propios a resolver con PDE               | 30 |
| Ecuación 8. Frontera tipo Dirichlet                                      | 31 |
| Ecuación 9. Frontera tipo Neumann  | 31 |
| Ecuación 10. Ecuación general de difusividad térmica                     | 35 |
| Ecuación 11. Ecuación de difusividad térmica para 2D                     | 36 |
| Ecuación 12. Ecuación general de una regresión lineal                    | 45 |
| Ecuación 13. Ecuación general de la regresión Model Discovery            | 46 |
| Ecuación 14. Ecuación ODE masa-resorte-amortiguador                      | 48 |
| Ecuación 15. Variables de estado   | 48 |
| Ecuación 16. Ecuación para $x_1$ en espacio de estados                   | 48 |
| Ecuación 17. Ecuación para $x_2$ en espacio de estados                   | 48 |
| Ecuación 18. Forma general del Model Discovery                           | 49 |
| Ecuación 19. Librería de términos para ODE                               | 50 |
| Ecuación 20. Forma de los datos para PDE                                 | 51 |
| Ecuación 21. Definición de derivada                                      | 52 |
| Ecuación 22. Ecuación general de un polinomio de Lagrange                | 54 |
| Ecuación 23. Factor L del polinomio de Lagrange                          | 54 |
| Ecuación 24. Polinomio de Lagrange de segundo orden                      | 54 |
| Ecuación 25. Polinomio de Lagrange de segundo orden variación            | 54 |
| Ecuación 26. Polinomio de Lagrange de segundo orden variación            | 55 |
| Ecuación 27. Ecuación de primeras derivadas para puntos no equidistantes | 55 |

|  |    |
|--|----|
| Ecuación 28. Ecuación de segundas derivadas para puntos no equidistantes | 55 |
| Ecuación 29. Librería de términos para PDE                               | 56 |
| Ecuación 30. Forma de la matriz de posiciones P                          | 59 |
| Ecuación 31. Forma de la matriz de temperatura T                         | 61 |
| Ecuación 32. Ecuación para el número de espacios entre nodos             | 69 |
| Ecuación 33. Ecuación para la distancia entre nodos                      | 69 |
| Ecuación 34. Ecuación para el número de búsquedas                        | 71 |
| Ecuación 35. Ecuación del error  | 74 |
| Ecuación 36. Ecuación de pérdidas de datos en las líneas de búsqueda     | 74 |

## **LISTA DE ANEXOS**

**(Ver anexos adjuntos y pueden visualizarlos en la Base de datos de la biblioteca UIS)**

Anexo A. Código para la simulación en un script live de Matlab.

Anexo B. Código base d la implementación del Model Discovery en un Script Live de Matlab.

Anexo C. Algoritmo Model Discovery para PDEs en un Script Live de Matlab.

Anexo D. Función (funordmatrix)

Anexo E. Datos de las simulaciones.

Anexo F. Script live del Model Discovery eliminando outliers de la data.

## RESUMEN

**TÍTULO:** IDENTIFICACIÓN DE LOS PARÁMETROS DE LA ECUACIÓN DIFERENCIAL DE DIFUSIÓN DE CALOR, BASADOS EN LOS DATOS OBTENIDOS DE UNA SIMULACIÓN DE LA VARIACIÓN DE TEMPERATURA, RESPECTO AL TIEMPO Y AL ESPACIO PARA UNA GEOMETRÍA (PLACA PLANA) \*

**AUTOR:** JEFREY JOANY GUEVARA HUERTAS\*\*

**PALABRAS CLAVE:** Descubrimiento de modelos, regresión lineal, ecuación de difusión de calor, Dirichlet.

### DESCRIPCIÓN:

En este trabajo de grado muestra el proceso seguido para obtener los parámetros de la ecuación de difusividad térmica, para dos dimensiones, mediante una regresión lineal llamada (Model Discovery) a partir de datos, estos se generan en una simulación por elementos finitos, usando el (PDE Toolbox), una herramienta de Matlab, la cual, permite resolver ecuaciones diferenciales parciales y sistemas de ecuaciones, se dibuja una placa cuadrada, la cual se discretiza en una malla con elementos triangulares, lineales o cuadráticos, los cuales contienen, tres, o seis nodos, respectivamente, se les puede variar el tamaño, como la cantidad, se define el material, las condiciones iniciales, y las condiciones de frontera ya sean tipo Dirichlet o Neuman, para este caso de estudio se usan condiciones de frontera tipo Dirichlet, se genera un vector de tiempo, que va desde cero en pequeños intervalos hasta el tiempo final, para el cual se va a hacer la simulación, al final la data se guarda con la información de posición y temperatura de cada nodo, además del vector tiempo.

La data obtenida carga en un script de Matlab, este ordena en líneas horizontales o verticales, calcula primeras y segundas derivadas respecto del espacio, luego respecto del tiempo, todas estas derivadas se acomodan en vectores columnas muy largos, deben ser del mismo tamaño, la variable dependiente es el vector de derivadas temporales y los otros vectores se acomodan en una matriz, se conoce como librería de términos, los parámetros se obtienen con una operación matricial, entre el vector de derivadas temporales y la librería de términos, esto se conoce como (Model Discovery). Las derivadas tienen ruido, hay que eliminarlo para disminuir el error, el cual, es la comparación de los coeficientes obtenidos con el que se usa en la simulación.

---

\*Trabajo de grado

\*\* Facultad de Físico-Mecánicas. Escuela de ingeniería mecánica. Ingeniería Mecánica.  
Director: Jabid Eduardo Quiroga Méndez. Doctor.

## ABSTRACT

**TITLE:** IDENTIFICATION OF THE PARAMETERS OF THE DIFFERENTIAL HEAT DIFFUSION EQUATION, BASED ON THE DATA OBTAINED FROM A SIMULATION OF THE VARIATION OF TEMPERATURE, WITH RESPECT TO TIME AND SPACE FOR A GEOMETRY (FLAT PLATE) \*

**AUTHOR:** JEFREY JOANY GUEVARA HUERTAS\*\*

**KEY WORDS:** model discovery, linear regression, heat diffusion equation, Dirichlet.

### DESCRIPTION:

In this degree work, he shows the process followed to obtain the parameters of the thermal diffusivity equation, for two dimensions, through a linear regression called (Model Discovery), from data, these are generated in a finite element simulation, using the (PDE Toolbox), a MATLAB tool, which allows solving partial differential equations and systems of equations, a square plate is drawn, which is discretized into a mesh with triangular, linear, or quadratic elements which contain, three, or six nodes, respectively, their size can be varied, such as the quantity, the material, the initial conditions, and the boundary conditions are defined, whether they are Dirichlet or Neuman type, for this case study, Dirichlet-type boundary conditions are used, a time vector is generated, which goes from zero in small intervals to the final time, for which the simulation is going to be done, at the end, the data is saved with the position and temperature information of each node, in addition to the time vector.

The data obtained is loaded into a MATLAB script, it orders in horizontal or vertical lines, calculates first and second derivatives with respect to space, then with respect to time, all these derivatives fit into very long column vectors, they must be the same size, the dependent variable is the vector of time derivatives and the other vectors are accommodated in a matrix, it is known as library of terms, the parameters are obtained with a matrix operation, between the vector of time derivatives and the library of terms, this is known as (Model Discovery). The derivatives have noise, it must be eliminated to reduce the error, which is the comparison of the coefficients obtained with the one used in the simulation.

---

\*Degree work

\*\*Physical-Mechanical Engineering Faculty. Mechanical Engineering School. Mechanical engineering. Director: Jabid Eduardo Quiroga Méndez. Doctor.

## INTRODUCCIÓN

En los tiempos actuales, que se vive tratando de anticiparse en los aspectos más cotidianos, por ejemplo, ir por la vía que contiene menos tráfico, comprar vuelos más baratos, vender productos por internet a personas que necesitan inmediatamente dicho producto, esto hace que las leyes físicas sean de vital importancia para la comprensión y predicción de comportamientos futuros, todo esto genera que últimamente se trabaje con mucha importancia en el descubrimiento de modelos basados en datos, lo que en inglés se conoce como (Data Driven-Discovery), consiste en que a partir de datos se conozcan las ecuaciones que gobiernan algún fenómeno.

El hecho de poder predecir comportamientos futuros permite anticiparse a inundaciones, sequías, fenómenos naturales, enfermedades, o quizá, anticiparse a hacer un producto estrella o cerrar un negocio importante, el hecho de que se puedan llegar a detectar enfermedades de la piel con cámaras, serían avances muy importantes en el ámbito de la salud.

En la ingeniería, es muy importante predecir comportamientos, ya sea con el fin de cambiar una pieza porque va a fallar, o de controlar algún sistema, como sistemas de refrigeración, de calefacción, controlar la calidad de aire en un recinto, la temperatura de un refrigerador o de un horno, anticiparse a factores que pueden influir en el diseño de cualquier cosa, como el diseño de un edificio, un puente, el bastidor de una máquina, o el hecho de identificar un material metálico por una propiedad.

La identificación de materiales metálicos siempre se ha hecho determinando propiedades físicas, para eso existen varios métodos convencionales, como pruebas de dureza, de textura, entre otros. Los métodos actuales involucran el uso de aparatos sofisticados, como ocurre cuando se hace una espectrometría de luz óptica, pero hasta ahora ningún material se ha podido identificar en base a sus propiedades térmicas, como la difusividad térmica propia de cada material.

Por esto se pretende determinar los parámetros de una ecuación de difusión de calor, primero desarrollando un algoritmo que sea capaz de aplicar una regresión lineal a unos datos provenientes de este fenómeno físico, en dos dimensiones, la forma correcta de iniciar las bases de este estudio, es generando datos a partir de simulaciones asistidas por programas computacionales, como Matlab, ya que, es económico y se puede obtener una gran cantidad de datos en tiempos muy cortos, para acelerar el proceso.

Una vez se tiene el programa piloto, se estudia, se optimiza y luego se puede aplicar a datos reales, por ejemplo, datos obtenidos en tiempo real a través de una cámara termográfica, para determinar materiales, o hacer control en base a una ecuación de gobierno que se genera a partir de datos.

## 1. TÉCNICAS PARA LA IDENTIFICACIÓN DE MATERIALES

A medida que transcurre el tiempo, la búsqueda de leyes físicas ha sido un objetivo fundamental de la ciencia. Las leyes físicas son primordiales para la comprensión de fenómenos naturales y la predicción de una dinámica futura, habitualmente se derivan de leyes físicas o se generalizan a partir de observaciones empíricas. Se ha trabajado últimamente en un tema llamado “*Data Driven-Discovery*” (descubrimiento basado en datos), dichos datos son experimentales y provienen de algún fenómeno que se basa en leyes físicas, donde el modelo físico que los produce no está claro, a partir de los datos se descubre un modelo fisicomatemático en ecuaciones diferenciales que gobiernan el comportamiento de los datos<sup>1</sup>.

Los modelos de “*Data Driven-Discovery*” se han venido usando en fenómenos donde se tenga una buena cantidad de datos, es decir, millones o miles de millones de estados, pero de pocas dimensiones con el fin de determinar comportamientos futuros<sup>2</sup>, por ejemplo:

- Convección de fluidos en la atmosfera.
- El clima.
- Epidemiología.
- Redes eléctricas.

En la industria existen varias técnicas tradicionales para determinar materiales metálicos, las cuales son; apariencia de la superficie, pruebas de chispa, la prueba de viruta, pruebas magnéticas, pruebas de dureza. Este tipo de pruebas que

---

<sup>1</sup> SHENG, Zhang; GUANG, Lin. [en línea]. revista journal of Computational Physics., marzo 1 de 2021: vol. 428, [consultado 20 marzo 2021]. ISSN 0021-9991. Disponible en: <https://bibliotecavirtual.uis.edu.co:2191/science/article/pii/S0021999120307361>.

<sup>2</sup> BRUNTON, Steven L.; PROCTOR, Joshua L.; y KUTZ, J. Nathan. [en línea]. Revista PNAS. Washington D.C, noviembre 5 de 2019: vol. 116, nro.45. [consultado 28 noviembre 2020]. ISSN 1091-6490. Disponible en: <https://www.pnas.org/content/pnas/116/45/22445.full.pdf>.

determina las propiedades físicas de los materiales, requieren de personas con mucha experiencia y perspicacia, capacitados, no son tan costosas en términos económicos, pero si requieren de bastante tiempo para llegar a la identificación del material, además que algunas de estas pruebas como el caso de la viruta hay que rayar el material y en el caso de la prueba de apariencia de la superficie requiere netamente el factor humano, que puede no ser muy preciso.

Figura 1. Apariencia superficial de algunos metales



*Fuente: Propiedades de los metales [en línea]. Andalucía España: Federación de Enseñanza de CC.OO. de Andalucía, 2011. [fecha de consulta: 2 diciembre 2022]. disponible en: <https://www.feandalucia.ccoo.es/docu/p5sd8631.pdf>*

Además de los métodos tradicionales, existen métodos modernos para la identificación de los metales los cuales son más rápidos y precisos, los cuales son: OES (*Optical Emission Spectrometry*), también conocida como espectrometría de emisión óptica, utiliza la intensidad de la luz emitida a una longitud de onda particular para determinar la composición elemental de una muestra. Al igual que las huellas dactilares, la emisión de rayos y luz es exclusiva del tipo de metal. XRF (*X-Ray*

*Fluorecence*), es una medida muy precisa y exacta de la composición elemental de los materiales, los espectrómetros XRF excitan una muestra con rayos X de alta energía, que obligan a la muestra a emitir ciertos rayos de luz característicos, posteriormente son leídos por el espectrómetro XRF. PMI (*Positive Metal Identification*), la cual es una técnica que usa tanto OES como XRF. XRD (*X-Ray diffraction*), se utiliza para identificar la información de composición química de los metales. XRD se puede usar de la mano con XRF, ya que XRD lleva las pruebas un paso más allá para brindar un contexto adicional. Estas pruebas suelen tener un costo monetario más elevado comparado con las pruebas tradicionales, ya que, requieren de una tecnología más avanzada.

No existe una técnica que involucre las propiedades térmicas de los materiales metálicos, sea rápida y precisa, además de ser más económica que las pruebas modernas. Por lo tanto, identificar las propiedades térmicas de los materiales como la difusividad, característica de cada material, mediante la toma de datos de temperatura, al calentarse o enfriarse por un periodo de tiempo corto, luego procesar estos datos, identificar las propiedades que hacen parte del fenómeno es el primer paso, para más adelante desarrollar un instrumento que pueda hacer todo el proceso, aplicándole el desarrollo matemático a un instrumento de recolección de datos como una cámara termográfica.

## **1.1 ALGUNAS PROPIEDADES TERMICAS Y FISICAS**

Las propiedades térmicas, so experimentadas por los materiales, como su nombre lo indica, bajo cargas térmicas, a diferencia de las propiedades físicas, que se obtienen de algunas mediciones físicas, como la masa o el volumen.

### 1.1.1 Conductividad térmica.

“Proveniente de las leyes de Fourier y asociada con la conducción de calor, la conductividad térmica está definida así, para un gradiente de temperatura prescrito, el flujo de calor por conducción aumenta al aumentar la conductividad térmica”<sup>3</sup>, está descrita por la siguiente ecuación.

$$k_x = -\frac{q_x''}{\left(\frac{\partial T}{\partial x}\right)} \quad (1)$$

Donde,  $k_x$  representa la conductividad térmica de un material,  $q_x''$  representa el flujo de calor a través de una superficie,  $\frac{\partial T}{\partial x}$  representa el gradiente de temperatura.

La definición se puede asociar para cualquier dimensión, a menos, de que el material sea isotrópico, es decir, sus propiedades no varíen con la dirección, es decir, que la conductividad térmica en cualquier dirección es la misma  $k_x = k_y = k_z = k$

### 1.1.2 Calor específico.

“Es la energía necesaria para aumentar la temperatura de una determinada sustancia en una unidad de temperatura, por unidad de masa”<sup>4</sup>, el calor específico está descrito por la siguiente ecuación.

---

<sup>3</sup> INCROPERA, Frank P, *et al.* Introduction to conduction. En: Fundamentals of heat and mass transfer. 6 ed. New York: John Wiley & Sons, 2007. p. 60

<sup>4</sup> ÇENGEL, Yunus A; BOLES, Michael A. Introducción y conceptos básicos. En: Termodinámica. 5 ed. México D.F: McGraw-Hill Interamericana editores,2006. p. 178

$$c_p = \frac{Q}{\Delta T * m} \quad (2)$$

Donde,  $c_p$  es el calor específico de una sustancia,  $Q$  es el calor o energía que absorbe la una sustancia,  $\Delta T$  es el cambio de temperatura que tiene la sustancia,  $m$  es la masa de la sustancia.

### 1.1.3 Densidad.

Es una propiedad física, “se define como la masa por unidad de volumen”<sup>5</sup>, la cual está representada por la siguiente ecuación.

$$\rho = \frac{m}{V} \quad (3)$$

Donde,  $\rho$  es la densidad de una sustancia,  $m$  la masa y  $V$  el volumen.

En general, la densidad de una sustancia depende de la temperatura y la presión. La densidad de la mayor parte de los gases es proporcional a la presión e inversamente proporcional a la temperatura. Por otro lado, los líquidos y sólidos son en esencia sustancias no compresibles y la variación de su densidad con la presión es por lo regular insignificante. A 20°C, por ejemplo, la densidad del agua cambia de 998 kg/m<sup>3</sup> a 1 atm, a 1 003 kg/m<sup>3</sup> a 100 atm, un cambio de sólo 0.5 por ciento. La densidad de líquidos y sólidos depende más de la temperatura que de la presión. A 1 atm, por ejemplo, la densidad del agua cambia de 998 kg/m<sup>3</sup> a 20°C a 975 kg/m<sup>3</sup> a 75°C, esto significa un cambio de 2.3 por ciento, lo cual puede ignorarse en muchos análisis de ingeniería.<sup>6</sup>

---

<sup>5</sup> ÇENGEL, Yunus A; BOLES, Michael A. Introducción y conceptos básicos. En: Termodinámica. 5 ed. México D.F: McGraw-Hill Interamericana editores,2006. p .13

<sup>6</sup> Ibid., p.13.

#### 1.1.4 Difusividad.

“Es la relación entre la conductividad térmica y la capacidad calorífica, es una propiedad muy importante, que tiene unidades de longitud al cuadrado por unidad de tiempo”<sup>7</sup>, la cual esta descrita por la siguiente ecuación.

$$\alpha = \frac{k}{\rho c_p} \quad (4)$$

Donde,  $\alpha$  es difusividad térmica de un material,  $k$  es la conductividad térmica,  $\rho$  la densidad y  $c_p$  el calor específico.

El producto de la densidad y el calor específico, conocido como capacidad volumétrica para almacenar calor, mide la capacidad de un material de almacenar energía térmica, por ende, la difusividad térmica mide la capacidad de un material para conducir energía térmica en relación con su capacidad para almacenar energía térmica. Los materiales con un gran coeficiente de difusividad térmica responderán rápidamente a los cambios en su ambiente térmico, mientras que los materiales de pequeño responderán más lentamente, tomando más tiempo para alcanzar una nueva condición de equilibrio<sup>8</sup>.

## 1.2 JUSTIFICACIÓN

Actualmente las empresas que usan materiales metálicos para procesarlos como materia prima, hacer piezas en masa o por encargo, necesitan saber el material de algún componente específico, que compran o proveen a un cliente, ya sea para cambiar, soldar, o reparar de forma adecuada alguna pieza, para esto se tiene que llevar a cabo una prueba, o técnica, para la determinación del metal que lo constituye. Las técnicas existentes en la industria para identificar materiales, tanto las tradicionales, como modernas, no satisfacen de manera íntegra los

---

<sup>7</sup> INCROPERA, Frank P, *et al.* Introduction to conduction. En: Fundamentals of heat and mass transfer. 6 ed. New York: John Wiley & Sons, 2007. p. 68

<sup>8</sup> Ibid., p. 67-68

requerimientos para que pueda ser accesible a todo el que las necesite, por tanto, las técnicas tradicionales son asequibles en términos económicos, pero son muy demoradas, además precisan del factor humano, el cual puede ser un causal de errores, donde depende de la perspicacia de la persona que efectúa la técnica de identificación, por otro lado, las técnicas modernas son monetariamente costosas. Se busca un método que sea rápido, eficaz, económico, y suficientemente preciso.

El proyecto es importante porque al tomar datos de temperatura de una geometría plana sometida a un proceso de difusión de calor, con un instrumento como una cámara termográfica, luego llevar estos datos a un programa o algoritmo, procesarlos e identificar los parámetros de la ecuación de difusión de calor de la geometría, los cuales solo dependen del tipo de material y por ende se puede identificar el constituyente de algún elemento que debe ser reemplazado, con esto se ahorra tiempo y dinero, puesto que, se hace la identificación sin necesidad de desarmar o dañar un pedazo, como se hace al aplicarle un ensayo tradicional, por otro lado, si se aplica una técnica moderna su costo es bastante alto, con la ventaja que se van a obtener resultados en un tiempo corto y bastante precisos.

### **1.3 OBJETIVOS.**

Los objetivos del proyecto de grado se fraccionan en dos categorías, objetivo general y objetivos específicos.

#### **1.3.1 Objetivo general.**

Contribuir a la misión del programa de ingeniería mecánica orientada hacia el desarrollo científico y tecnológico; al implementar una metodología para determinar los parámetros de la ecuación de difusión de calor, a base de datos de temperatura obtenidos con una simulación.

### 1.3.2 Objetivos específicos.

**1.3.2.1** Obtener los datos históricos de temperatura en cada nodo que resulta de discretizar una placa (2D) de cobre homogénea, sometida a un proceso de difusión de calor en estado transitorio, usando la *Toolbox* PDE de Matlab, con condiciones de frontera tipo "*Dirichlet*".

**1.3.2.2** Implementar una regresión lineal conocida como método de "*Model Discovery*" propuesto por J. Nathan Kutz en 2018<sup>9</sup>, para identificar los parámetros de la ecuación de difusión de calor a partir de los datos de temperatura obtenidos en el objetivo anterior.

**1.3.2.3** Validar el método implementado a través de la determinación del error, al comparar el coeficiente de difusión de calor usado en la simulación y el obtenido por el método "*Model Discovery*".

---

<sup>9</sup> UNIVERSITY OF WASHINGTON; Inferring structure of complex system; Model Discovery [sitio web]. Washington D. C.; Kutz, J. Nathan. [consulta 30 noviembre 2022]. Disponible en: <http://faculty.washington.edu/kutz/am563/page1/page5/am563.html>

## **2. DATOS HISTÓRICOS DE TEMPERATURA**

Los datos históricos de la temperatura son el elemento más importante, porque a partir de estos se determina el modelo fisicomatemático más cercano que puede representarlos, por lo general los datos se obtienen mediante medición experimental de un fenómeno físico o una simulación de este. Los datos experimentales son reales, por lo tanto, son más confiables, son almacenados y posteriormente se le hace el debido tratamiento, ya sea filtrado u cualquier otro tipo para eliminar ruidos o datos erróneos que no corresponden con la mayoría, mientras que los datos obtenidos mediante una simulación permiten un ambiente más controlado, por lo general, con bajo ruido, son almacenados y luego se hace con ellos el trabajo que se pretende.

### **2.1 MEDICION EXPERIMENTAL**

En la investigación es muy común hacer medición experimental, la cual consiste en tomar datos de cierto fenómeno físico, como por ejemplo los datos de temperatura para un fenómeno de difusividad térmica, la cual se desarrolla en un laboratorio con un ambiente controlado, ya sea para una geometría dada, como una placa plana, tanto el material del cual está hecha y las condiciones térmicas a las cuales está siendo sometida. Esto permite medir la temperatura de una esquina, un borde, de un punto central o de una malla de puntos de dicha placa plana en determinado instante el cual se desee estudiar.

## 2.2 SIMULACIÓN

“Puede definirse a la simulación como la experimentación con un modelo que imita ciertos aspectos de la realidad. Esto permite trabajar en condiciones similares a las reales, pero con variables controladas y en un entorno que se asemeja al real pero que está creado o acondicionado artificialmente”<sup>10</sup>. Para el caso de estudio los datos se obtienen mediante una simulación de elementos finitos, usando la herramienta *PDE Toolbox* que viene incorporada en el software MATLAB.

### 2.2.1 PDE TOOLBOX

Es una caja de herramientas la cual proporciona funciones para resolver problemas de mecánica estructural, transferencia de calor, y en general, ecuaciones diferenciales parciales, utilizando análisis de elementos finitos. Puede realizar un análisis estático lineal para calcular la deformación, el estrés y la tensión. Para modelar dinámica estructural y vibración, la caja de herramientas proporciona un solucionador de integración de tiempo directo. Se pueden estudiar las características estructurales de un componente realizando un análisis modal para encontrar frecuencias y formas modales. Puede modelar problemas de transferencia de calor, problemas dominados por conducción para calcular distribuciones de temperatura, flujos de calor y tasas de flujo de calor, a través de superficies. También se puede resolver problemas estándar como difusión, electrostática y magnetostática, así como personalizados. Permite importar geometrías 2D y 3D desde un formato STL o datos de malla. Puede generar automáticamente mallas con elementos triangulares y tetraédricos. Puede resolver PDE mediante el uso del método de elementos finitos y el pos-procesamiento de los resultados para explorarlos y analizarlos<sup>11</sup>.

---

<sup>10</sup> Pérez Porto, J, Merino, M. (15 de julio de 2011). Definición de simulación - Qué es, Significado y Concepto. Definicion.de. [en línea] [Consultado 10 diciembre 2022]. Disponible en <https://definicion.de/simulacion/>

<sup>11</sup> MathWorks. Partial Differential Equation Toolbox: User's Guide [en línea] r2021b. [Consultado 3 noviembre 2022]. Disponible en [https://la.mathworks.com/help/releases/R2021b/pdf\\_doc/pde/index.html](https://la.mathworks.com/help/releases/R2021b/pdf_doc/pde/index.html)

### 2.2.1.1 Ecuaciones que se pueden resolver con el *PDE Toolbox*.

Ecuaciones escalares de la forma:

$$m \frac{\partial^2 u}{\partial t^2} + d \frac{\partial u}{\partial t} - \nabla \cdot (c \nabla u) + au = f \quad (5)$$

Y ecuaciones de valores propios, de la forma:

$$-\nabla \cdot (c \nabla u) + au = \lambda du \quad (6)$$

$$-\nabla \cdot (c \nabla u) + au = \lambda^2 mu \quad (7)$$

Es decir que se pueden resolver ecuaciones de segundo orden, elípticas, parabólicas e hiperbólicas.

Tabla 1. Clasificación de EDP de segundo orden

| <b>Ecuación</b>                                      | <b>Nombre</b> | <b>Tipo</b> |
|--|---------------|-------------|
| $\nabla^2 u = 0$                                     | Laplace       | Elíptica    |
| $\nabla^2 u = f$                                     | Poisson       | Elíptica    |
| $\nabla^2 u = ku$                                    | Helmholtz     | Elíptica    |
| $\frac{\partial u}{\partial t} = k \nabla^2 u$       | Difusión      | Parabólica  |
| $\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u$ | Onda          | Hiperbólica |

Fuente: WIKIPEDIA. Ecuación en derivadas parciales [Sitio Web]. Clasificación de las EDP de segundo orden. [Consulta 30 de noviembre de 2022]. Disponible en: [https://es.wikipedia.org/wiki/Ecuaci%C3%B3n\\_en\\_derivadas\\_parciales](https://es.wikipedia.org/wiki/Ecuaci%C3%B3n_en_derivadas_parciales)

**2.2.1.2 Tipos de frontera.** Para las ecuaciones diferenciales parciales se pueden elegir dos tipos de condiciones frontera, ya sea para un borde o una cara.<sup>12</sup>

**2.2.1.2.1 Dirichlet.** En un borde o en una cara, la solución  $u$  satisface la ecuación.

$$hu = r \quad (8)$$

Donde  $h$  y  $r$  pueden ser funciones dependientes del espacio ( $x, y$ , y  $z$  para el caso de 3D), de la solución  $u$ , y del tiempo. A menudo se toma el valor de  $h = 1$  y se inserta en  $r$  un valor apropiado<sup>13</sup>.

**2.2.1.2.2 Condiciones de frontera generalizadas tipo Neuman.** En un borde o en una cara, la solución  $u$  satisface la ecuación.

$$\vec{n} \cdot (c\nabla u) + qu = g \quad (9)$$

Donde  $\vec{n}$  es un vector normal exterior unitario.  $q$  y  $g$  son funciones definidas en  $\partial\Omega$  y pueden ser funciones dependientes del espacio ( $x, y$ , y  $z$  para el caso de 3D), de la solución  $u$ , de ecuaciones dependientes del tiempo y del tiempo.<sup>14</sup>

---

<sup>12</sup> MathWorks. Partial Differential Equation Toolbox: User's Guide [en línea] r2021b. [Consultado 3 noviembre 2022]. Disponible en [https://la.mathworks.com/help/releases/R2021b/pdf\\_doc/pde/index.html](https://la.mathworks.com/help/releases/R2021b/pdf_doc/pde/index.html)

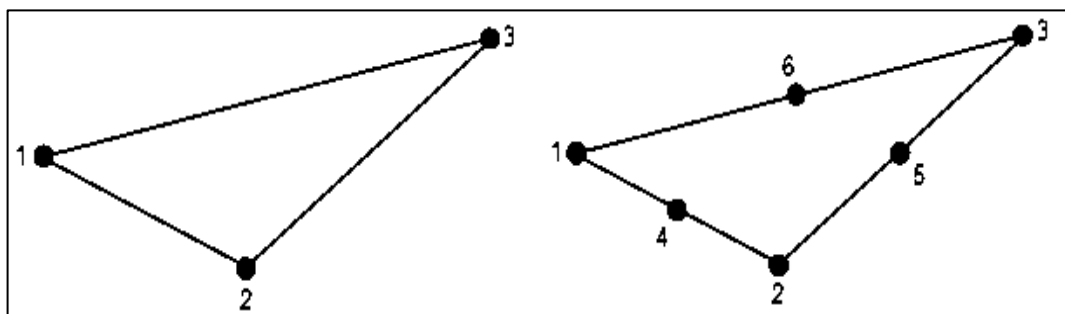
<sup>13</sup> Ibid.,

<sup>14</sup> Ibid.,

**2.2.1.3 Geometrías que se pueden dibujar en *PDE Toolbox*.** Se pueden importar geometrías ya hechas tanto en 2D como en 3D en formato SLT, también se pueden trazar dibujos a partir de geometrías básicas como, rectángulos, círculos, polígonos y elipses, las cuales se unen a partir de operaciones básicas, suma (+) para añadir una figura a otra, resta (-) para quitarle una figura a otra y multiplicación (\*) para marcar la intersección de dos figuras. Se pueden trabajar coordenadas cartesianas como polares<sup>15</sup>.

**2.2.1.4 Arreglo de la malla.** *PDE Toolbox* utiliza mallas con elementos triangulares para geometrías 2D y mallas con elementos tetraédricos para geometrías 3D. En ambos casos utiliza la geometría cuadrática por defecto, y ofrece la opción de cambiar al orden geométrico lineal. La malla siempre está formada por elementos del mismo orden, no permite generar mallas mixtas. Los triángulos en mallas 2D están especificados por tres nodos para elementos lineales o seis nodos para elementos cuadráticos. Un triángulo que representa un elemento lineal tiene nodos en las esquinas. Un triángulo que representa un elemento cuadrático tiene nodos en sus esquinas y centros de cada borde del triángulo.<sup>16</sup>

Figura 2. Forma lineal y cuadrática del elemento de malla para 2D



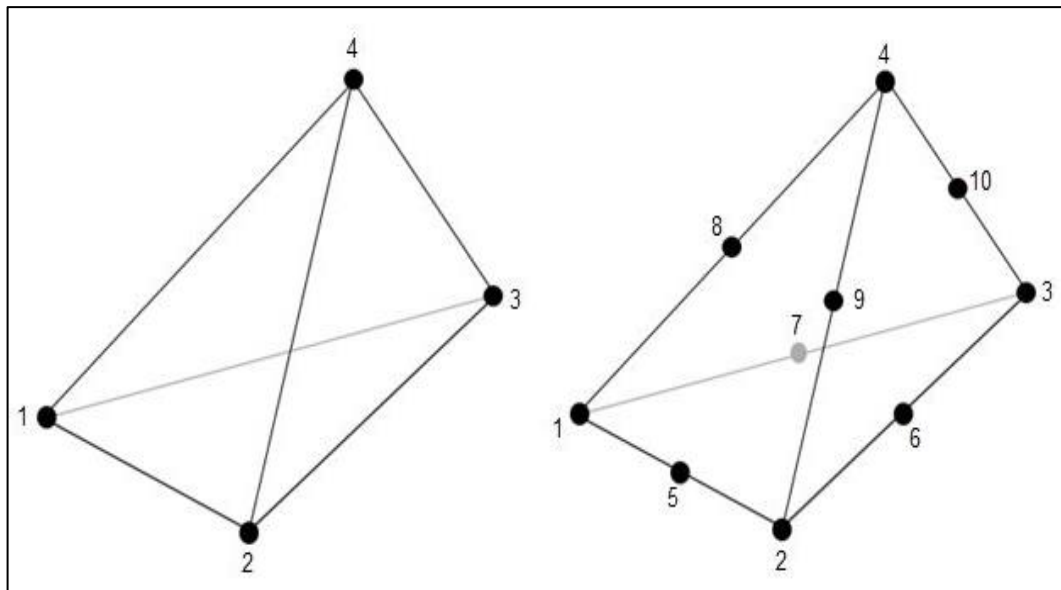
Fuente: MathWorks. *Partial Differential Equation Toolbox: User's Guide*

<sup>15</sup> MathWorks. *Partial Differential Equation Toolbox: User's Guide* [en línea] r2021b. [Consultado 3 noviembre 2022]. Disponible en [https://la.mathworks.com/help/releases/R2021b/pdf\\_doc/pde/index.html](https://la.mathworks.com/help/releases/R2021b/pdf_doc/pde/index.html)

<sup>16</sup> *Ibid.*,

Los tetraedros en mallas 3D están especificados por cuatro nodos para elementos lineales o diez nodos para elementos cuadráticos. Un tetraedro que representa un elemento lineal tiene nodos en las esquinas. El tetraedro que representa un elemento cuadrático tiene nodos en sus esquinas y centros de cada borde del tetraedro. Los nodos centrales en mallas cuadráticas siempre se agregan a la mitad de la distancia entre las esquinas. Para geometrías con superficies y bordes curvos, es posible que los nodos centrales no aparezcan en el borde o la superficie de sí mismo. El objeto contenedor del modelo almacena los parámetros del modelo PDE. La caja de herramientas ofrece varios tipos de objetos contenedor modelo, cada uno para un área de aplicación particular. Por ejemplo, para la elasticidad lineal problemas, el contenedor modelo es un objeto *StructuralModel*, y para problemas de transferencia de calor, el contenedor modelo es un objeto *ThermalModel*. Para problemas generales de PDE, la caja de herramientas utiliza el objeto *PDEModel*.<sup>17</sup>

Figura 3. Forma lineal y cuadrática del elemento de malla para 3D



Fuente: MathWorks. *Partial Differential Equation Toolbox: User's Guide*.

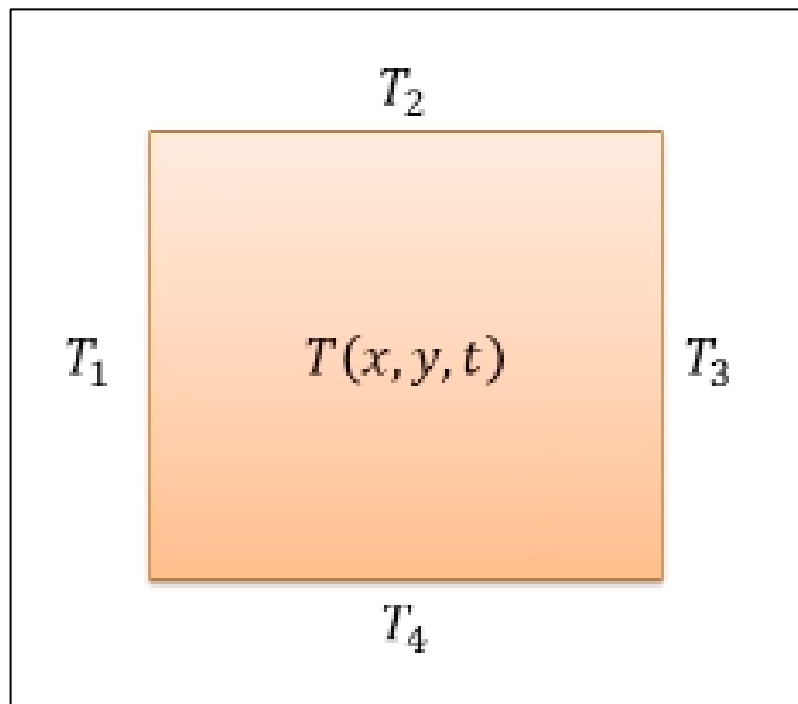
<sup>17</sup>MathWorks. *Partial Differential Equation Toolbox: User's Guide* [en línea] r2021b. [Consultado 3 noviembre 2022]. Disponible en [https://la.mathworks.com/help/releases/R2021b/pdf\\_doc/pde/index.html](https://la.mathworks.com/help/releases/R2021b/pdf_doc/pde/index.html)

## 2.2.2 SIMULACIÓN USANDO PDE TOOLBOX DE MATLAB.

En esta sección se va a mostrar la forma en la que se utiliza la herramienta *PDE Toolbox* de Matlab, usando un *Script* o un *Live script* de Matlab, para simular un caso de transferencia de calor, en base a la ecuación de difusividad térmica aplicada a una placa de cobre cuadrada.

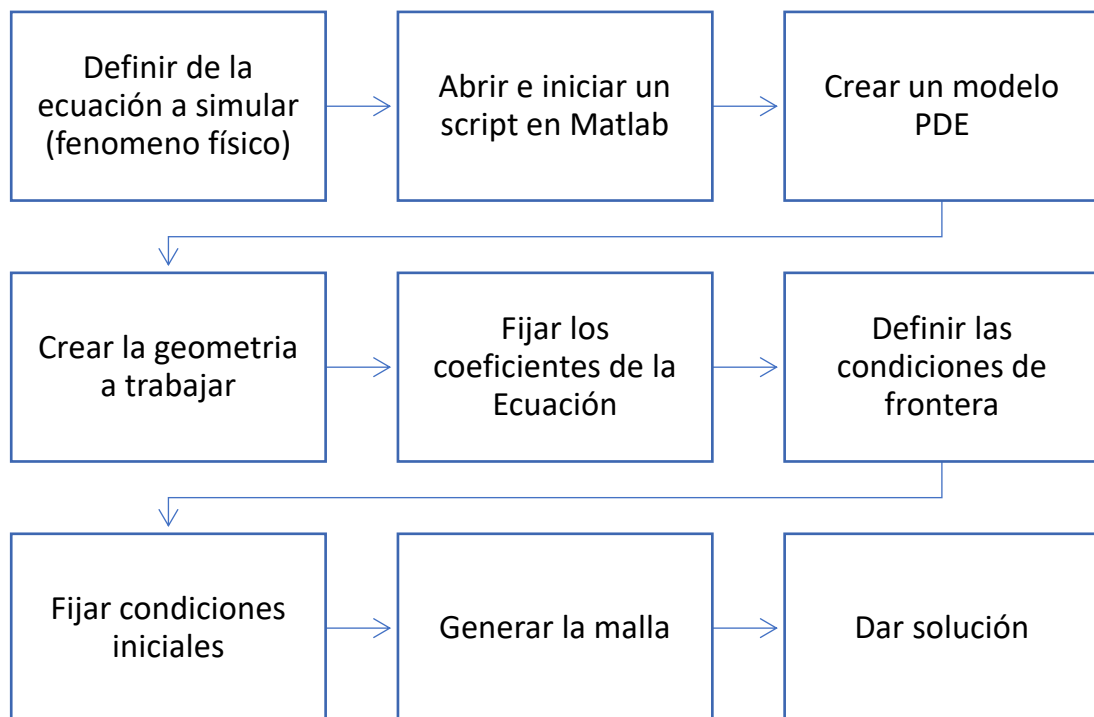
Se va a tomar como ejemplo una placa de cobre cuadrada de  $5 \times 5 \text{ cm}$  de ancho por alto, muy delgada, con coeficiente de difusividad térmica  $\alpha = 1,17 \times 10^{-4} \text{ m/s}^2$  con las siguientes condiciones de frontera en el borde izquierdo  $T_1 = 100 \text{ °C}$ , el borde superior  $T_2 = 0 \text{ °C}$ , el borde derecho  $T_3 = 0 \text{ °C}$ , y el borde inferior  $T_4 = 100 \text{ °C}$ , con temperatura inicial  $T_0 = 10 \text{ °C}$ . Los valores de temperatura de  $0$  y  $100 \text{ °C}$ , en las fronteras de la placa, están relacionadas con la temperatura de fusión y ebullición del agua a condiciones estándar.

Figura 4. Geometría para simular



El procedimiento que se usa para la simulación se muestra en el siguiente diagrama de flujo.

Figura 5. Diagrama de flujo para la simulación



**2.2.2.1 Se define la ecuación.** Esta describe el proceso físico que se desea simular, es decir, la forma y los coeficientes que la acompañan. En este caso es una ecuación de segundo orden, parabólica, llamada “Ecuación de difusión de calor” para dos dimensiones, la cual está definida por las ecuaciones (10) y (11).

$$\rho C_p \frac{\partial T}{\partial t} - \nabla(k\nabla T) = 0 \quad (10)$$

$$\frac{\partial T}{\partial t} = \alpha \left[ \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right] \quad (11)$$

Donde,  $T$  es la temperatura,  $\rho$  es la densidad del material,  $c_p$  es el calor específico,  $k$  es coeficiente de conducción,  $\frac{\partial T}{\partial t}$  es la razón de cambio de la temperatura respecto del tiempo,  $\nabla$  representa el gradiente, es decir, la razón de cambio de una magnitud respecto una dirección determinada, esto para la ecuación (10). Para la ecuación (11),  $\frac{\partial T}{\partial t}$  es la razón de cambio de la temperatura respecto del tiempo,  $\alpha$  el coeficiente de difusión térmica de un material,  $\frac{\partial^2 T}{\partial x^2}$  y  $\frac{\partial^2 T}{\partial y^2}$  representan la segunda derivada de la temperatura respecto la dirección  $x$  y  $y$  respectivamente.

**2.2.2.2 Se inicia el Script en Matlab.** Se inicia limpiando todo usando las funciones (*close all*, *clear* y *clc*), luego se definen las constantes del problema, difusividad térmica del material, en este caso cobre y temperatura inicial.

**2.2.2.3 Se crea el modelo PDE.** Se crea el modelo con la función (*createpde*), la cual se almacena en una variable.

**2.2.2.4 Se crea la geometría que se va a trabajar.** Para crear la geometría que se desea trabajar, se hace a partir de la suma o resta de figuras geométricas básicas, como rectángulos, círculos, polígonos y elipses.

**2.2.2.4.1 Círculo.** Se genera mediante un vector que tiene las siguientes especificaciones.

Tabla 2. Creación de un círculo *PDE Toolbox*

| Row | Value                         |
|-----|-------------------------------|
| 1   | 1 (indicate a circle)         |
| 2   | x-coordinate of circle center |
| 3   | y-coordinate of circle center |
| 4   | Radius (strictly positive)    |

*Fuente: MathWorks. Partial Differential Equation Toolbox: User's Guide*

**2.2.2.4.2 Polígono.** Se genera mediante un vector que tiene las siguientes especificaciones.

Tabla 3. Creación de un polígono *PDE Toolbox*

| Row               | Value                                |
|-------------------|--------------------------------------|
| 1                 | 2 (indicate a polygon)               |
| 2                 | Number of line segments n            |
| 3 through 3+n-1   | x-coordinate of edge starting points |
| 3+n through 2*n+2 | y-coordinate of edge starting points |

*Fuente: MathWorks. Partial Differential Equation Toolbox: User's Guide*

**2.2.2.4.3 Rectángulo.** Se genera mediante un vector que tiene las siguientes especificaciones.

Tabla 4. Creación de un rectángulo *PDE Toolbox*

| Row          | Value                                |
|--------------|--------------------------------------|
| 1            | 3 (indicate a rectangle)             |
| 2            | 4 (number of line segments)          |
| 3 through 6  | x-coordinate of edge starting points |
| 7 through 10 | y-coordinate of edge starting points |

*Fuente: MathWorks. Partial Differential Equation Toolbox: User's Guide*

**2.2.2.4.4 Elipse.** Se genera mediante un vector que tiene las siguientes especificaciones.

Tabla 5. Creación de una elipse *PDE Toolbox*

| Row | Value  |
|-----|--|
| 1   | 4 (indicate an ellipse)                        |
| 2   | x-coordinate of ellipse center                 |
| 3   | y-coordinate of ellipse center                 |
| 4   | First semiaxis length (strictly positive)      |
| 5   | Second semiaxis length (strictly positive)     |
| 6   | Angle in radians from x axis to first semiaxis |

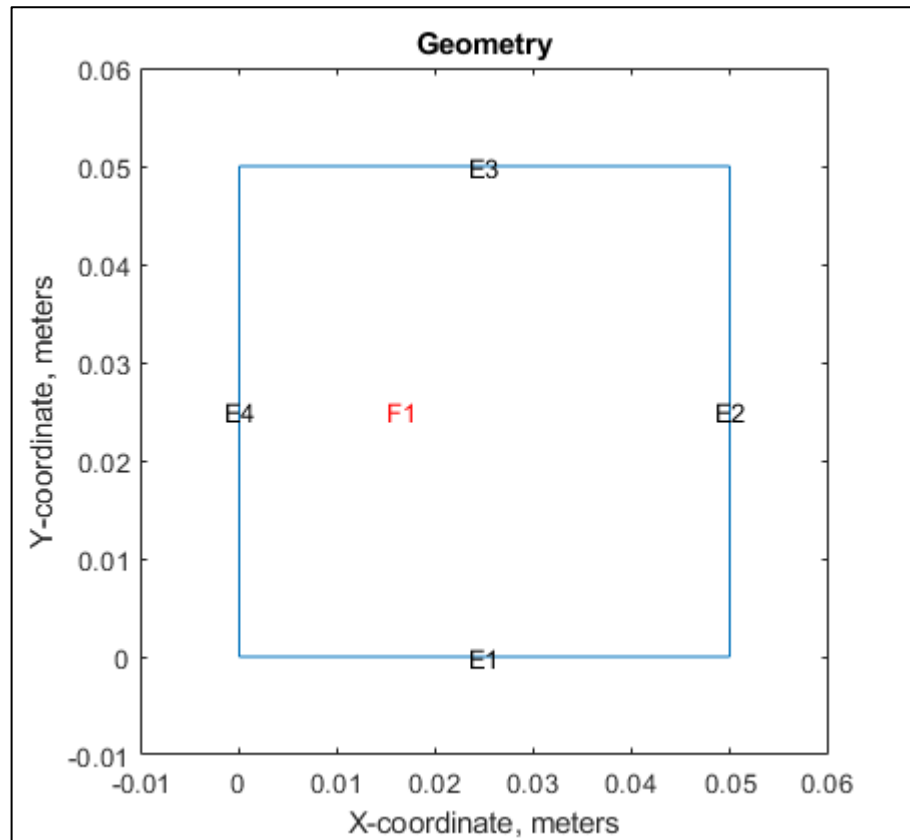
*Fuente: MathWorks. Partial Differential Equation Toolbox: User's Guide*

**2.2.2.4.5 Crear una superficie a partir de geometrías ecuaciones.** Se crea la geometría con una función predispuesta del PDE Toolbox de Matlab, cuya función es descomponer todas las formas en una matriz de mínimas regiones ( $dl = dectsg(gd, sf, ns)$ ), para este caso es una geometría simple, hecha a partir de una sola figura, se ingresa como valor de entrada en la función, la variable donde se creó el vector rectángulo. Cuando se crea una superficie más compleja, es decir a partir de muchas formas básicas, se ingresan otros valores de entrada a la función, se recomienda emplear la guía de usuario (*PDE Toolbox User Guide*), donde se especifica de manera correcta el uso de la función para geometrías más complejas.

**2.2.2.4.6 Creación de una figura a partir de matriz de geometría descompuesta.** Se crea la figura que se desea trabajar, ingresando en la función (*geometryFromEdges*) como variables de entrada, la variable donde se creó el modelo PDE y la matriz de mínimas regiones.

**2.2.2.5 Se grafica la geometría del modelo.** Se hace una gráfica de la geometría creada con el objetivo de verificar que la figura que se creó ya sea sencilla, o a partir de figuras básicas, tiene las dimensiones correctas y está bien hecha.

Figura 6. Geometría creada en PDE Toolbox



**2.2.2.6 Se fijan los coeficientes.** Se especifican los coeficientes de la ecuación (11), la difusividad térmica que se pretende evaluar, en la forma general que trabaja PDE Toolbox, es decir la ecuación (5), donde aparecen los coeficientes  $m$ ,  $d$ ,  $c$ ,  $a$  y  $f$ . para esto se usa la función (*specifyCoefficients*) la cual tiene como entrada la variable donde se creó el modelo PDE y los coeficientes.

**2.2.2.7 Se definen las condiciones de frontera.** Para el ejemplo son condiciones de frontera tipo Dirichlet, es decir temperatura constante. Si se tiene una frontera de calor constante, se usa fronteras tipo Neumann, es ese caso se

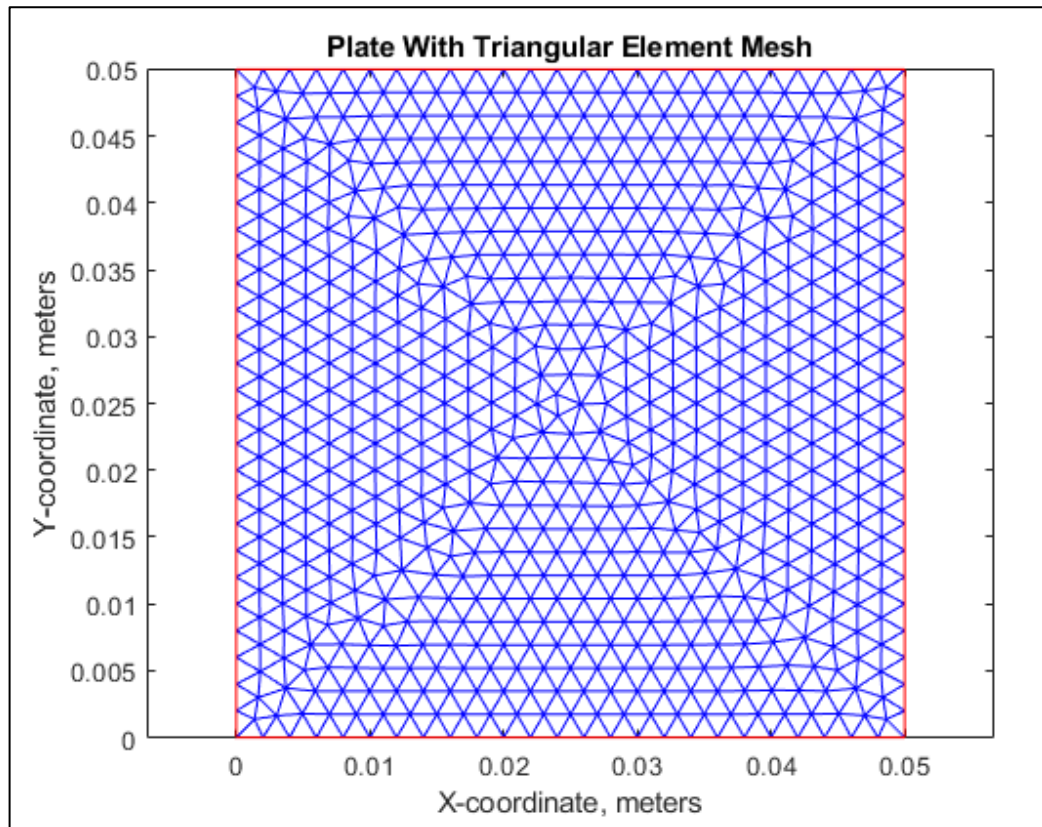
recomienda consultar la guía de usuario (*PDE Toolbox User Guide*), donde especifican el uso correcto de las condiciones de frontera. Para la aplicación de las condiciones de frontera se usa la función (*applyBoundaryCondition*) la cual tiene como entrada la variable donde se creó el PDE, el tipo de condición de frontera, la cara o el borde y el valor de la condición de frontera.

**2.2.2.8 Se fijan las condiciones iniciales.** En este caso solamente es la temperatura inicial a la que se encuentra la placa con la función (*setInitialConditions*) la cual tiene como entrada la variable donde se creó el modelo PDE y las condiciones iniciales. Para casos más específicos se recomienda consultar la guía de usuario (*PDE Toolbox User Guide*), donde especifican el ingreso de las condiciones iniciales.

**2.2.2.9 Se genera la malla.** Para generar la malla hay que ingresar el tamaño del arreglo, mediante el número de elementos que va a contener la geometría, o el tamaño máximo u mínimo de los elementos de la malla y si es un arreglo lineal o cuadrático, esto se hace con el uso de la función (*generateMesh*) la cual tiene como entrada la variable donde se creó el modelo PDE, el tamaño máximo del elemento y el orden geométrico, si es lineal o cuadrático.

Se recomienda generar una gráfica donde se muestre la malla, para verificar si se concibió correctamente el arreglo.

Figura 7. Malla de la placa

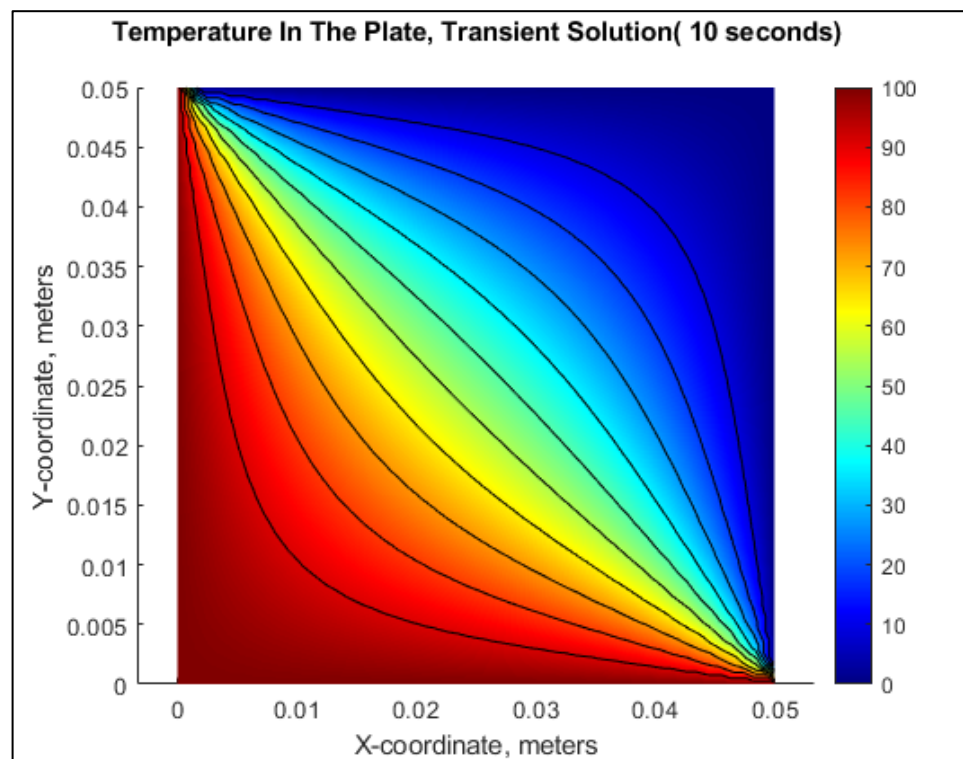


**2.2.2.10 Se soluciona el problema.** Para la solución del problema se crea una matriz ( $P$ ) que guarde los datos de cada nodo, es decir, las coordenadas  $(x, y)$  de cada uno de los nodos, un vector tiempo que vaya desde cero hasta un valor final ( $t$ ) en pequeños intervalos de tiempo ( $dt$ ), la solución se almacena en una variable ( $R$ ) con la función (*Solvepde*) que depende del modelo creado y al vector tiempo, la temperatura de cada nodo para cada instante se almacena en un matriz ( $T$ ).

**2.2.2.11 Disposición de la solución.** Una vez se tiene la solución, se puede saber el comportamiento de un nodo ubicado en cualquier parte del arreglo, u en general, de toda la malla de nodos, con esta información se puede hacer una gráfica en 2D, para un instante de tiempo determinado, o un recorrido por el tiempo, a modo de animación, de igual manera, se puede aplicar la animación para una gráfica en 3D, o simplemente almacenar algunos datos obtenidos de la simulación, para darles algún tratamiento, como en este caso. Para guardar se usa el comando (*save*) donde se necesita un nombre para lo que se va a guardar y los elementos que se guardan, para este caso, la matriz de Temperaturas, la matriz de posiciones de los nodos, el vector de tiempo, el intervalo de tiempo y la malla.

Se puede generar una gráfica de 2D que muestre el comportamiento de la placa para cualquier instante de tiempo.

Figura 8. Solución bidimensional al cabo de 10 s



**2.2.2.12 Error de la simulación.** Se reduce el error al refinar la malla, al ser un modelado por elemento finitos entre más pequeña sea la discretización del elemento a estudiar se reduce el error.

Es de vital importancia hacer un mallado correcto, si es bastante fino va a gastar más recursos computacionales, pero en caso contrario, si el mallado no es lo suficientemente fino se van a dar grandes errores, ya que, el fenómeno simulado se aleja del real.

Para ver todo el proceso que se siguió en la simulación, hay que ir al Anexo A. Código para la simulación en un script live de Matlab.

### 3. MODEL DISCOVERY

Es un modelo matemático en el cual a unos datos se le aplica una regresión lineal.

La regresión lineal es una técnica de modelado estadístico que se emplea para describir una variable de respuesta continua como una función de una o varias variables predictoras. Puede ayudar a comprender y predecir el comportamiento de sistemas complejos o a analizar datos experimentales, financieros y biológicos. Las técnicas de regresión lineal permiten crear un modelo lineal. Este modelo describe la relación entre una variable dependiente  $Y$  (también conocida como la respuesta) como una función de una o varias variables independientes  $X_i$  (denominadas predictores). La ecuación general correspondiente a un modelo de regresión lineal es la siguiente.<sup>18</sup>

$$Y = \beta_0 + \sum \beta_i X_i + \epsilon_i \quad (12)$$

Donde  $\beta$  representa las estimaciones de parámetros lineales que se deben calcular y,  $\epsilon$  representa los términos de error. Existen varios tipos de regresiones lineales, simple, múltiple, multivariante, y múltiple-multivariante.<sup>19</sup>

El modelo de regresión que se usa con *Model Discovery* difiere un poco del modelo general de una regresión lineal, es más simple, no cuenta con el término que representa el error, la ecuación que lo representa es la siguiente.

---

<sup>18</sup> MathWorks. [sitio web]. Descripción de relaciones matemáticas y realización de predicciones a partir de datos experimentales; ¿Qué es la regresión lineal? [Consultado 14 de diciembre 2022]. disponible en: <https://la.mathworks.com/discovery/linear-regression.html>

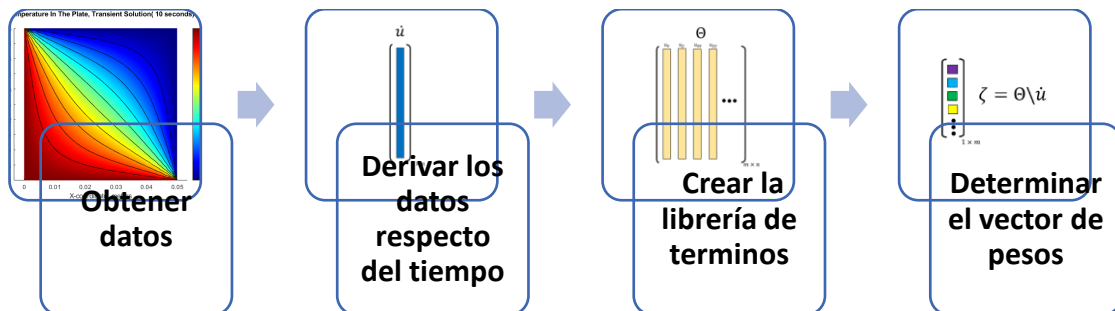
<sup>19</sup> Ibid.,

$$Y = \beta X \quad (13)$$

Donde  $Y$  es un vector de datos variantes en el tiempo, en otras palabras, la variable dependiente,  $\beta$  es una matriz llamada librería de términos, donde se contienen vectores de las posibles variables independientes, y, por último,  $X$  que es un vector de coeficientes a determinar, este vector es el peso asociado a cada una de las posibles variables independientes, es decir, muestra las variables de las cuales depende el fenómeno del que se adquirieron los datos y en qué proporción.

El funcionamiento de la regresión *Model Discovery* es muy sencillo, basta con tener unos datos, estos datos se derivan respecto al tiempo, el resultado se almacena en un vector columna, siguiente a esto se arma una librería de términos, esta es una matriz que contiene vectores columna del mismo tamaño del de derivadas temporales, estos vectores son posibles términos que acompañan la ecuación que genera los datos, por último se calcula el vector de peso, asociado a cada uno de los vectores columna que hacen parte de la matriz de términos, es decir, cada valor del vector de peso, corresponde al coeficiente que acompaña a cada uno de los posibles términos de la ecuación.

Figura 9. Diagrama de flujo para la aplicación del *Model Discovery*



El *Model Discovery* es un algoritmo de regresión lineal, usado inicialmente para fenómenos que provienen de ecuaciones diferenciales ordinarias (ODE del inglés *ordinary differential equations*), que posteriormente se aplicó a fenómenos provenientes de ecuaciones diferenciales parciales (PDE del inglés *partial differential equations*) en una sola dirección, funciona de la misma manera para (ODE) como para (PDE) aplicando el paso a paso que muestra la Figura 9. Diagrama de flujo para la aplicación del *Model Discovery*. La diferencia radica en la creación de la librería de términos, luego “los fenómenos descritos por (ODE) son representados por ecuaciones que contienen solo derivadas de una o más variables dependientes respecto a una sola variable independiente, en su gran mayoría la variable independiente es el tiempo, mientras que los fenómenos descritos por (PDE) son representados por una ecuación que involucra derivadas parciales de una o más variables dependientes de dos o más variables independientes”<sup>20</sup>.

### 3.1 DATOS PROVENIENTES DE (EDO).

Por lo general los datos provenientes de ecuaciones diferenciales ordinarias, son datos dependientes del tiempo, pues, es la variable independiente ( $t$ ), las variables dependientes son por lo general las dimensiones espaciales ( $x, y, z$ ), ya sea para fenómenos unidimensionales, bidimensionales o tridimensionales, pero que solo son dependientes del tiempo.

Un ejemplo claro, además de ser muy común, es el sistema que compone una masa, un resorte y un amortiguador, es un fenómeno que es descrito por una ecuación diferencial ordinaria, el cual busca manifestar el comportamiento de la masa, al ser

---

<sup>20</sup> ZILL, Dennis G; WRIGHT, Warren S. Definiciones y terminología En: Ecuaciones diferenciales con problemas con valores en la frontera. 8 ed. México D.F: Cengage Learning Editores S.A., 2015, p.2-12

excitada por un impulso  $P$ , el desplazamiento que esta realiza a lo largo de un determinado tiempo, este fenómeno es representado por la siguiente ecuación.

$$m \frac{d^2x}{dt^2} + b \frac{dx}{dt} + kx = P \quad (14)$$

Donde,  $m$  es la masa,  $b$  el coeficiente de amortiguamiento y  $k$  la constante elástica de un resorte.

Para poder representar el fenómeno en espacio de estados, se debe hacer los siguientes cambios de variables.

$$x_1 = x ; x_2 = \dot{x}_1 = \frac{dx}{dt} ; \dot{x}_2 = \ddot{x}_1 = \frac{d^2x}{dt^2} \quad (15)$$

En el análisis en el espacio de estados se centra la atención en los tres tipos de variables que aparecen en el modelado de los sistemas dinámicos; las variables de entrada, las variables de salida y las variables de estado, las variables de un sistema dinámico son las variables que constituyen el menor conjunto de variables que determinan el estado del sistema dinámico. La representación en el espacio de estados de un sistema dado no es única, salvo que el número de variables de estado es el mismo para cualquiera que sea la representación en variables de estado de un mismo sistema.<sup>21</sup>

Por tanto, el en fenómeno representado en espacio de estados queda definido por las siguientes ecuaciones.

$$\dot{x}_1 = x_2 \quad (16)$$

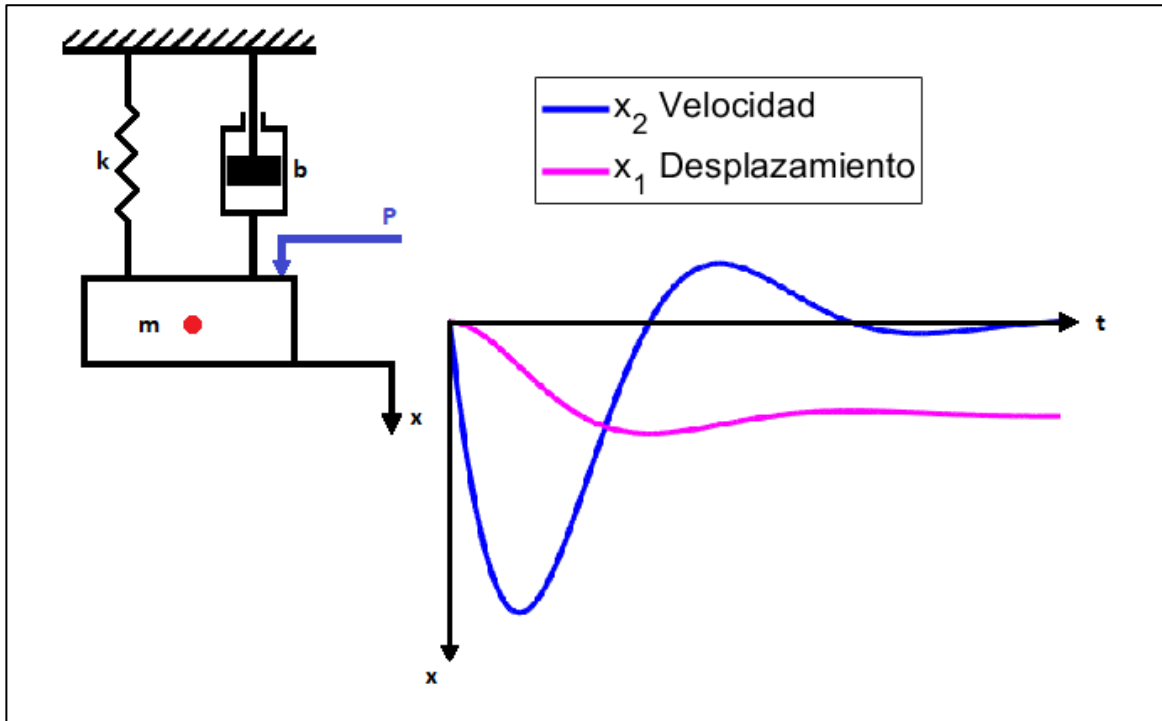
$$\dot{x}_2 = \frac{P}{m} - \frac{b}{m}x_2 - \frac{k}{m}x_1 \quad (17)$$

El desplazamiento y la velocidad del centroide de la masa, hasta que alcanza una posición de equilibrio se muestra en la siguiente figura.

---

<sup>21</sup> OGATA, Katsuhiko. Modelado matemático de sistemas de control En: Ingeniería de control moderna. 5 ed. Madrid: Pearson Educación S.A., 2010, p. 13-60

Figura 10. Sistema ODE



La matemática del modelo es bastante sencilla, pues se basa en la solucionar un arreglo algebraico de matrices y vectores, el cual está representado por la siguiente ecuación.

$$\dot{u} = \Theta \zeta \quad (18)$$

Donde,  $\dot{u}$  es un vector columna de las derivadas temporales de unos datos provenientes de un fenómeno dependiente de una variable, en la mayoría de los casos, el tiempo,  $\Theta$  representa la matriz conocida como librería de términos y  $\zeta$  es un vector columna conocido como el vector de pesos.

Se siguen los pasos descritos por la Figura 9. Diagrama de flujo para la aplicación del *Model Discovery*.

Se debe tener cuidado en el conformado de la librería de términos, porque cada vector que haga parte de esta es un posible termino de los que depende la ecuación que describe el fenómeno, para el caso de una ecuación diferencial ordinaria que representa el sistema masa resorte amortiguador, la librería de términos queda conformada de la siguiente forma.

$$\Theta = [A, x_1, x_2, x_1 * x_2, x_1^2, x_2^2 \dots] \quad (19)$$

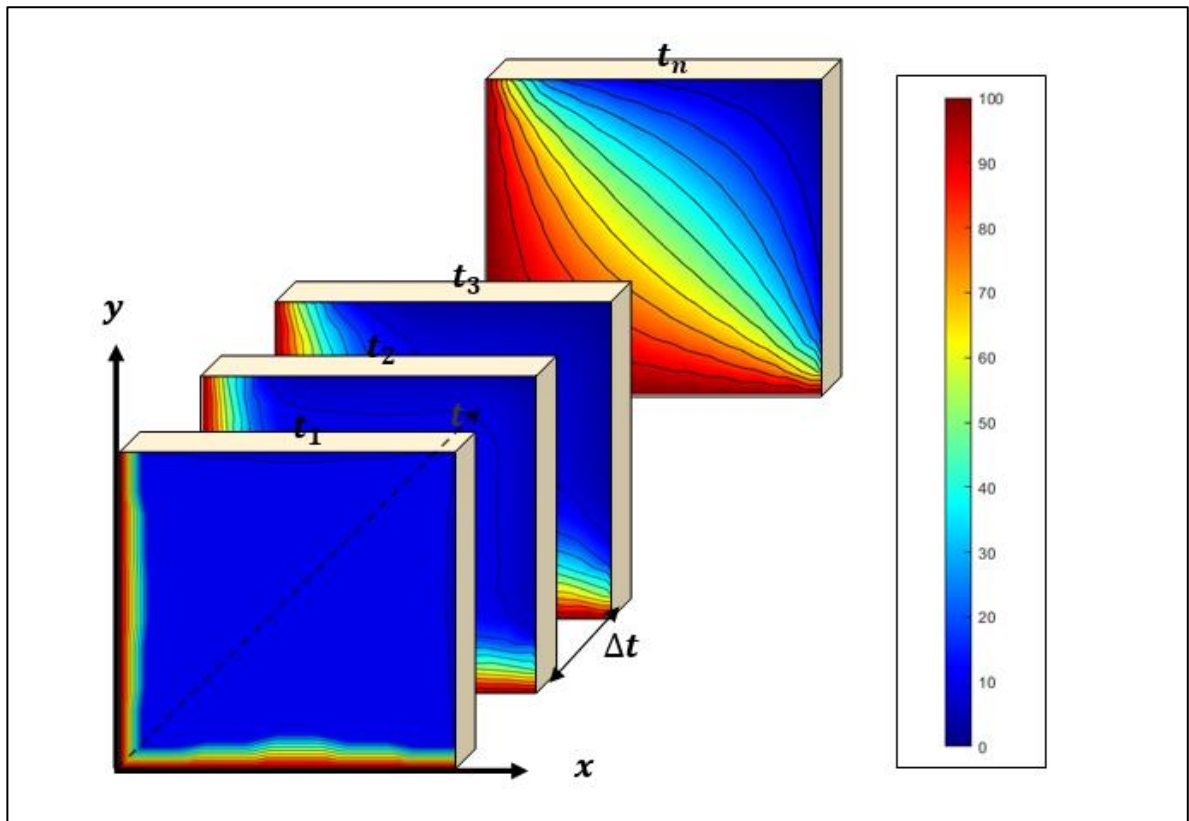
Donde,  $A$  representa un vector columna de unos, una contante,  $x_1$  representa el vector columna de datos de posición,  $x_2$  representa el vector columna de datos de velocidad, y una combinación de posibles términos de los cuales depende la misma ecuación, estos vectores columna son del mismo tamaño.

### 3.2 DATOS PROVENIENTES DE (PDE).

Por lo general los datos provenientes de ecuaciones diferenciales parciales, son datos dependientes del tiempo, y del espacio, pues, hay variables independientes  $(t, x, y, z)$ , las variables dependientes son por lo general propiedades como la temperatura, esfuerzos, entre otras, ya sea para fenómenos unidimensionales, bidimensionales o tridimensionales, pero que son dependientes del tiempo y del espacio.

Un ejemplo claro es el calentamiento de una placa, es un fenómeno representado por una ecuación diferencial parcial, llamada ecuación de difusión de calor para dos dimensiones, la ecuación (11). El de calentamiento de una placa se puede percibir como un cumulo de imágenes de la placa, cada imagen corresponde a un tiempo  $(t)$ , y contiene la información de temperatura de cada punto de la geometría para dicho tiempo. Se vería de la siguiente manera.

Figura 11. Sistema PDE



Los datos de temperatura de la placa son dependientes tanto del espacio como del tiempo, por lo tanto, vienen dados de la siguiente forma.

$$u = f(t, x, y, \dots) \quad (20)$$

Donde,  $u$  representa la cualquier propiedad, para este caso la temperatura y es una variable dependiente del espacio  $(x, y)$  y del tiempo  $(t)$ .

Una vez se tienen los datos del fenómeno se procede a aplicar la regresión, con los pasos descritos por Figura 9. Diagrama de flujo para la aplicación del *Model Discovery*.

### 3.2.1 Derivadas de los datos.

Dado un conjunto de datos o una función, puede ser útil diferenciar la cantidad considerada para determinar una propiedad físicamente relevante. Por ejemplo, dado un conjunto de datos que representa la posición de una partícula en función del tiempo, entonces la derivada y la segunda derivada dar la velocidad y la aceleración, respectivamente. Del cálculo, la definición de la derivada es dada por<sup>22</sup>.

$$\frac{df(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{f(t + \Delta t) - f(t)}{\Delta t} \quad (21)$$

La ecuación anterior es la definición de derivada, la idea es hacer que el  $\Delta t$  sea muy pequeño para reducir el error de cálculo, el cual se puede cuantificar y calcular usando la serie de expansión de *Taylor*<sup>23</sup>. Existen unas formulas ya desarrolladas para calcular las derivadas numéricamente.

Tabla 6. Fórmulas de diferencias centrales de segundo orden

| <b><math>O(\Delta t^2)</math> center-differences schemes</b>  |
|---|
| $f'(t) = [f(t + \Delta t) - f(t - \Delta t)]/2\Delta t$   |
| $f''(t) = [f(t + \Delta t) - 2f(t) + f(t - \Delta t)]/\Delta t^2$   |
| $f'''(t) = [f(t + 2\Delta t) - 2f(t + \Delta t) + 2f(t - \Delta t) - f(t - 2\Delta t)]/2\Delta t^3$         |
| $f''''(t) = [f(t + 2\Delta t) - 4f(t + \Delta t) + 6f(t) - 4f(t - \Delta t) + f(t - 2\Delta t)]/\Delta t^4$ |

*Fuente: KUTZ, J. Nathan. Data-driven modeling & scientific computation. Oxford: Oxford university press,2013. P.80*

<sup>22</sup> KUTZ, J. Nathan. Data-driven modeling & scientific computation. Oxford: Oxford university press,2013. P.77

<sup>23</sup> Ibid,.

Tabla 7. Fórmulas de diferencias centrales de cuarto orden

**$O(\Delta t^4)$  center-differences schemes**

|  |
|--|
| $f'(t) = [-f(t + 2\Delta t) + 8f(t + \Delta t) - 8f(t - \Delta t) + f(t - 2\Delta t)]/12\Delta t$  |
| $f''(t) = [-f(t + 2\Delta t) + 16f(t + \Delta t) - 30f(t) + 16f(t - \Delta t) - f(t - 2\Delta t)]/12\Delta t^2$  |
| $f'''(t) = [-f(t + 3\Delta t) + 8f(t + 2\Delta t) - 13f(t + \Delta t) + 13f(t - \Delta t) - 8f(t - 2\Delta t) + f(t - 3\Delta t)]/8\Delta t^3$             |
| $f''''(t) = [-f(t + 3\Delta t) + 12f(t + 2\Delta t) - 39f(t + \Delta t) + 56f(t) - 39f(t - \Delta t) + 12f(t - 2\Delta t) - f(t - 3\Delta t)]/6\Delta t^4$ |

Fuente: KUTZ, J. Nathan. *Data-driven modeling & scientific computation*. Oxford: Oxford university press,2013. P.80

Tabla 8. Fórmulas de diferencias hacia adelante y hacia a tras de segundo orden

**$O(\Delta t^2)$  forward- and backward-difference schemes**

|   |
|---|
| $f'(t) = [-3f(t) + 4f(t + \Delta t) - f(t + 2\Delta t)]/2\Delta t$                      |
| $f'(t) = [3f(t) - 4f(t + \Delta t) + f(t + 2\Delta t)]/2\Delta t$                       |
| $f''(t) = [2f(t) - 5f(t + \Delta t) + 4f(t + 2\Delta t) - f(t + 3\Delta t)]/\Delta t^2$ |
| $f''(t) = [2f(t) - 5f(t - \Delta t) + 4f(t - 2\Delta t) - f(t - 3\Delta t)]/\Delta t^2$ |

Fuente: KUTZ, J. Nathan. *Data-driven modeling & scientific computation*. Oxford: Oxford university press,2013. P.80

Una vez se elige la ecuación correcta para derivar los datos dependientes del tiempo, se hace el proceso de derivación, al terminar, se almacenan el producto de la derivación en un vector columna muy largo  $\dot{u}$ . Cabe aclarar que todas las tablas anteriores son para datos equidistantes.

**3.2.1.1 Derivación de datos no equidistantes.** Una manera de emplear datos irregularmente espaciados consiste en ajustar un polinomio de interpolación de Lagrange de segundo grado<sup>24</sup>.

$$f_n(x) = \sum_{i=0}^n L_i(x)f(x_i) \quad (22)$$

Donde

$$L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} \quad (23)$$

Por lo tanto, el polinomio de Lagrange se segundo grado  $n = 2$  queda de la siguiente manera.

$$f_2(x) = \left(\frac{x - x_1}{x_0 - x_1}\right)\left(\frac{x - x_2}{x_0 - x_2}\right)f(x_0) + \left(\frac{x - x_0}{x_1 - x_0}\right)\left(\frac{x - x_2}{x_1 - x_2}\right)f(x_1) + \left(\frac{x - x_0}{x_2 - x_0}\right)\left(\frac{x - x_1}{x_2 - x_1}\right)f(x_2) \quad (24)$$

$$f_2(x) = \left(\frac{x^2 - x(x_1 + x_2) + x_1x_2}{(x_0 - x_1)(x_0 - x_2)}\right)f(x_0) + \left(\frac{x^2 - x(x_0 + x_2) + x_0x_2}{(x_1 - x_0)(x_1 - x_2)}\right)f(x_1) + \left(\frac{x^2 - x(x_0 + x_1) + x_0x_1}{(x_2 - x_0)(x_2 - x_1)}\right)f(x_2) \quad (25)$$

Con el polinomio de la ecuación (25), se aplica la primera y segunda derivada

---

<sup>24</sup> CHAPRA, Steven C; CANALE, Raymond P. Diferenciación e integración numéricas En: Métodos numéricos para ingenieros. México D. F: McGraw Hill interamericana editores S.A. DE C.V. Séptima edición. 2015. p.515

$$f_2'(x) = \left( \frac{2x - x_1 - x_2}{(x_0 - x_1)(x_0 - x_2)} \right) f(x_0) + \left( \frac{2x - x_0 - x_2}{(x_1 - x_0)(x_1 - x_2)} \right) f(x_1) + \left( \frac{2x - x_0 - x_1}{(x_2 - x_0)(x_2 - x_1)} \right) f(x_2) \quad (26)$$

Si se hace símil con las diferencias centrales, el punto dos sería el paso adelante, el punto cero el paso atrás y el punto uno el punto donde se está parado o se quiere hallar la derivada. Por tanto, la primera derivada quedaría de la siguiente forma.

$$f'(x) = \left( \frac{2x - x_i - x_{i+1}}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})} \right) f(x_{i-1}) + \left( \frac{2x - x_{i-1} - x_{i+1}}{(x_i - x_{i-1})(x_i - x_{i+1})} \right) f(x_i) + \left( \frac{2x - x_{i-1} - x_i}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)} \right) f(x_{i+1}) \quad (27)$$

Por tanto, la segunda derivada quedaría de la siguiente manera.

$$f''(x) = \left( \frac{2}{(x_{i-1} - x_i)(x_{i-1} - x_{i+1})} \right) f(x_{i-1}) + \left( \frac{2}{(x_i - x_{i-1})(x_i - x_{i+1})} \right) f(x_i) + \left( \frac{2}{(x_{i+1} - x_{i-1})(x_{i+1} - x_i)} \right) f(x_{i+1}) \quad (28)$$

### 3.2.2 Librería de términos

Es una matriz de vectores, donde cada vector es un posible término, del cual dependen los datos obtenidos de fenómeno físico. Los vectores de los posibles términos se determinan a partir de los datos que se tienen, los cuales son

dependientes del espacio, además del tiempo, por ejemplo, si los datos recolectados tienen un valor cualquiera asignado a una posición en específico  $(x, y)$  es muy posible que estos dependan de estas coordenadas, el primer vector de la librería sería la derivada de los datos respecto de la coordenada  $x$ , el segundo vector sería la derivada de los datos respecto de la coordenada  $y$ , luego la segunda derivada y así se construye la librería de términos, con los posibles vectores de los cuales dependen los datos. La librería de términos se vería así.

$$\Theta = \left[ u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial^2 u}{\partial x^2}, \frac{\partial^2 u}{\partial x \partial y}, \frac{\partial^2 u}{\partial y^2}, \dots \right] \quad (29)$$

Donde  $u$  es el conjunto de datos que se tienen almacenado en un vector columna muy largo y los demás son también vectores columna muy largos que son calculados a partir de los datos obtenidos inicialmente, a partir de derivadas parciales.

### 3.2.3 Determinación de los parámetros.

Para la determinar los parámetros y por tanto la ecuación que gobierna los datos los cuales están dados por la ecuación (20), se siguen los siguientes pasos.

- Se calcula  $\dot{u}$  y se mete en un vector columna de tamaño  $1 \times n$
- Se arma la librería  $\Theta = [u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}, \dots]$  se debe tener en cuenta que la librería es una matriz  $m \times n$  de vectores columna  $1 \times n$  donde  $m$  corresponde a cantidad de vectores columna. Donde  $u$  son los datos  $u_x$  las derivadas de los datos respecto al eje  $x$ ,  $u_y$  la primera derivada de los datos respecto al eje  $y$ ,  $u_{xx}$  y  $u_{yy}$  las segundas derivadas.

- Se calcula el vector de peso asociado a cada uno de los términos  $\zeta$ , esto se puede hacer con una pseudoinversa  $\zeta = \Theta \backslash \hat{u}$ , o mediante una técnica de regularización como “LASSO”. El vector resultante es de tamaño  $1 \times m$ .
- Para validar el método se halla el porcentaje de error existente entre el coeficiente que se determina por la regresión lineal (*Model Discovery*) y el coeficiente de difusividad térmica que se usó inicialmente para la simulación y obtención de los datos.

**3.2.3.1 Pseudoinversa.** Resuelve el sistema de ecuaciones lineales como la ecuación (18) para  $\zeta$ . Las matrices  $\Theta$  y  $\hat{u}$  deben contener el mismo número de columnas.

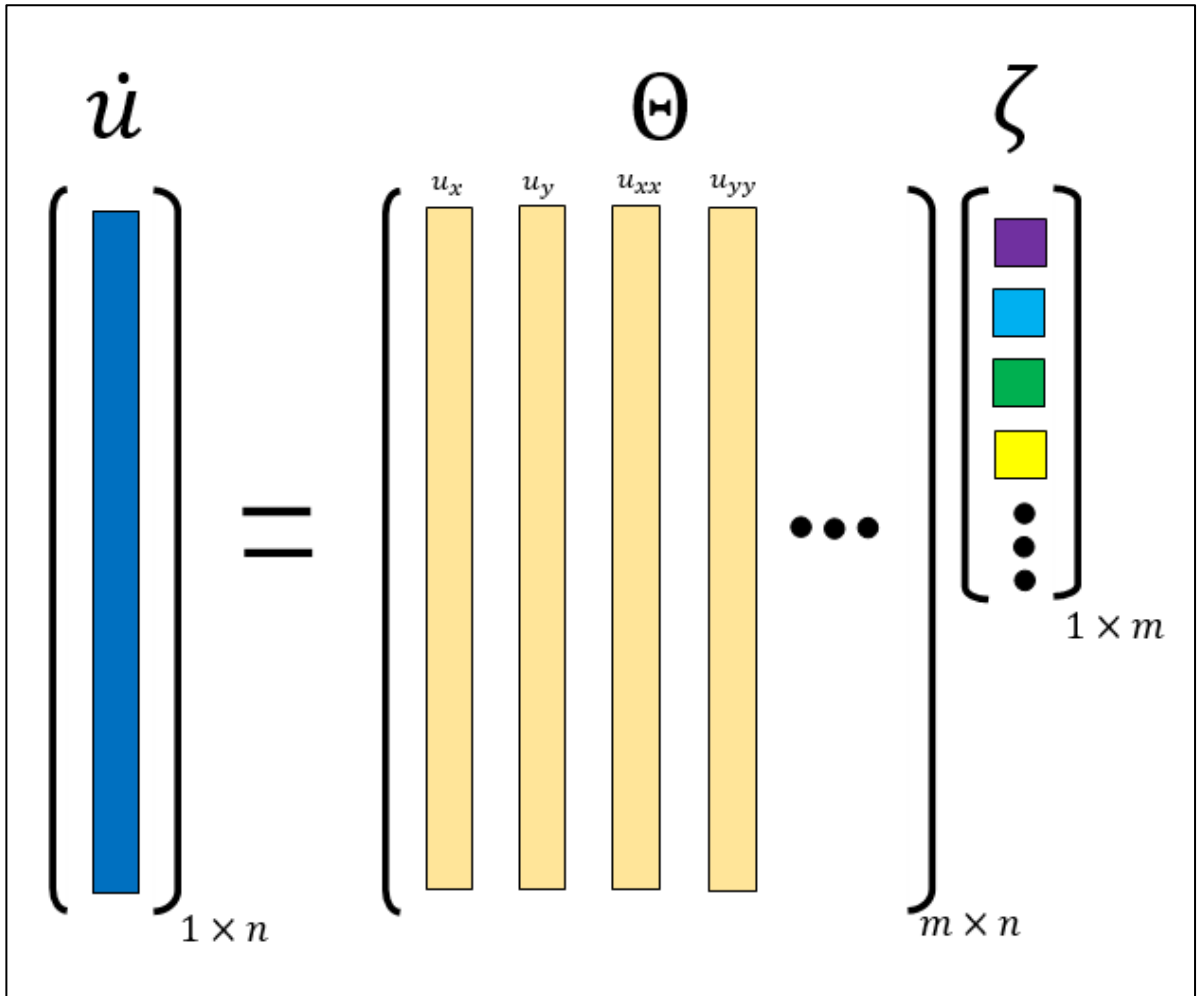
### 3.2.3.2 Lasso.

Es una técnica de regularización para realizar regresiones lineales, usada para reducir el número de predictores en un modelo de regresión, identificar predictores importantes, seleccionar entre predictores redundantes. Incluye un término de penalización  $\alpha$  que restringe el tamaño de los coeficientes estimados, por lo tanto, se asemeja a la regresión de cresta, es un estimador de contracción, genera estimaciones de coeficientes que están sesgadas para ser pequeñas, sin embargo, un estimador de Lasso puede tener un error cuadrático medio más pequeño que un estimador de mínimos cuadrados ordinario cuando lo aplica a datos nuevos. La red elástica es lo mismo que el lazo cuando  $\alpha = 1$ , a medida que  $\alpha$  se encoge hacia 0, la red elástica se aproxima a la regresión de la cresta, a diferencia de la regresión de cresta, a medida que aumenta el término de penalización, establece más coeficientes en cero, esto significa que el estimador de Lasso es un modelo más pequeño, con menos predictores, como tal, Lasso es una alternativa a la regresión por pasos y otras técnicas de reducción de dimensionalidad y selección de modelos.<sup>25</sup>

---

<sup>25</sup> MathWorks. [sitio web]. Lasso and Elastic Net; ¿what are Lasso an Elastic Net? [Consultado 16 de diciembre 2022]. [https://la.mathworks.com/help/stats/lasso-and-elastic-net.html#bs9ub\\_i](https://la.mathworks.com/help/stats/lasso-and-elastic-net.html#bs9ub_i)

Figura 12. *Model Discovery* para PDE



### 3.3 APLICACIÓN DEL MODEL DISCOVERY A LOS DATOS SIMULADOS

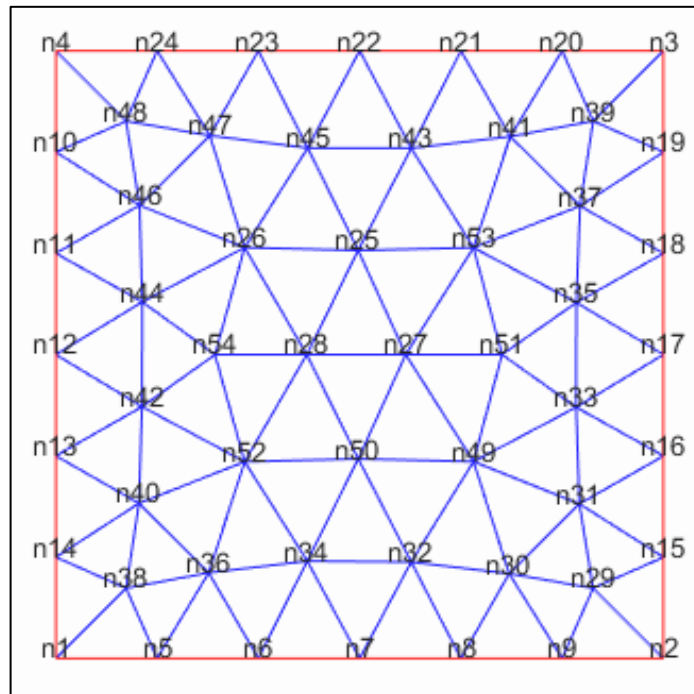
Una vez se tienen los datos que resultan de la simulación, se procede a cargarlos en el *script* de Matlab y conocer la forma en que están guardados. Se cargan mediante el comando (*load*).

Como son datos dependientes del tiempo, además vienen de una geometría bidimensional, también son dependientes del espacio, por ende, deben estar guardados de esa manera.

Si se hace una imagen de un arreglo bidimensional, el cual tiene cincuenta y cuatro puntos (nodos) de estudio en orden aleatorio, estos puntos  $P$  tienen una información de posición que no cambia en el tiempo, como se muestra en la siguiente imagen.

$$P = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ y_1 & y_2 & y_3 & \dots & y_n \end{bmatrix} \quad (30)$$

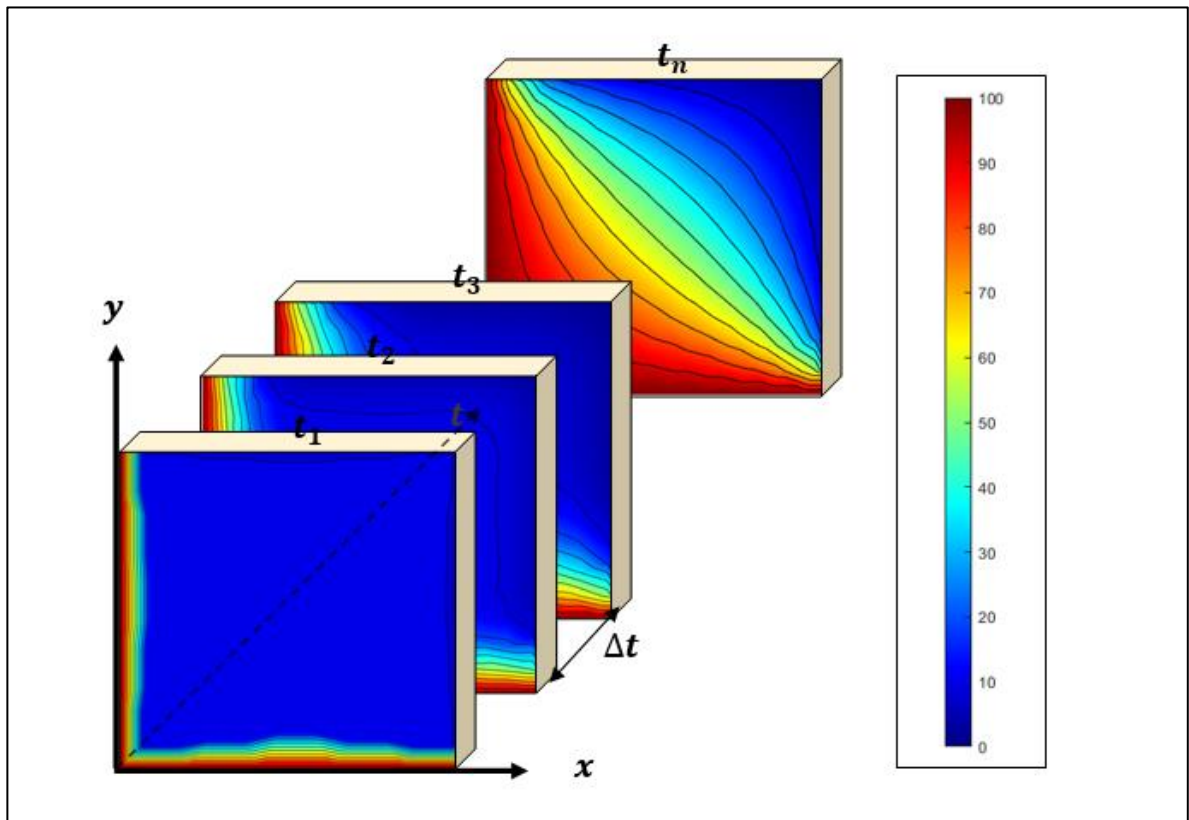
Figura 13. Arreglo bidimensional de puntos con coordenadas (x, y)



Si bien los puntos de estudio de la imagen anterior no cambian de posición, la conservan a lo largo del tiempo, estos si varían en una propiedad dependiente del tiempo, para este caso la temperatura, a menos de que por alguna condición esta

permanezca constante a lo largo del tiempo. Si se estudian todos los nodos a medida que transcurre el tiempo se ve que algunos cambian el valor de temperatura, pero ninguno cambia de posición, es decir que para un instante de tiempo dado la temperatura varía con la posición, la temperatura de un nodo  $n$  puede ser diferente a la de un nodo vecino para ese instante de tiempo  $t$ , como la temperatura de un mismo nodo  $n$  cambia de un instante  $t$  con respecto a un instante  $t + \Delta t$ , por eso el fenómeno se vería como una acumulación de imágenes de información, una a continuación de otras, esto recibe el nombre de snapshots.

Figura 14. Snapshots del fenómeno



Los datos se ven como imágenes instantáneas de información, una imagen a continuación de otra, pero esta información en el tiempo, al trabajar con imágenes no es cómodo ni sencillo, debe descomponerse en una matriz de datos, la cual es más fácil tratarla matemáticamente hablando, por eso la temperatura se almacena en una matriz como la que se muestra en la siguiente ecuación.

$$T = \begin{bmatrix} T_{1,1} & T_{1,2} & T_{1,3} & \dots & T_{1,n} \\ T_{2,1} & T_{2,2} & T_{2,3} & \dots & T_{2,n} \\ T_{3,1} & T_{3,2} & T_{3,3} & \dots & T_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T_{m,1} & T_{m,2} & T_{m,3} & \dots & T_{m,n} \end{bmatrix} \quad (31)$$

Donde el subíndice  $m$  de  $T$  es el nombre de cada nodo, y el subíndice  $n$  es el instante de tiempo para el cual ha sido medida la temperatura de cada nodo.

Con la data cargada, se procede a aplicar la regresión lineal *Model Discovery*, para ello se necesitan dos aspectos importantes, la derivada de los datos respecto del tiempo, y la construcción de la librería de términos, donde cada término de la librería, no, es más, que un vector de primeras, segundas, u, otras derivadas de los datos respecto al espacio, u alguna combinación de estas. La librería se basa en los posibles términos de los que depende la ecuación.

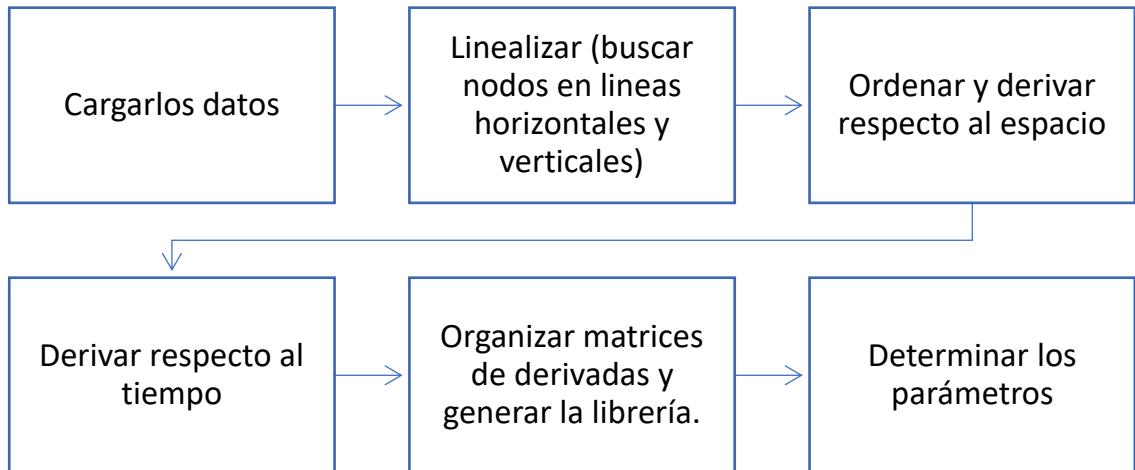
Al conocer la data se ve que la derivación respecto al tiempo no causa ningún problema, por otro lado, la derivación respecto al espacio de cada nodo sí, ya que el arreglo de nodos es aleatorio el primer nodo puede estar muy distante del segundo y así sucesivamente, uno que otro puede coincidir de simple casualidad. Se tienen que ordenar los datos para poder derivarlos, como los datos están en dos dimensiones, se deben ordenar respecto a una dimensión  $x$  y luego respecto a la dimensión sobrante  $y$ , con esto se puede armar la librería de términos y determinar los parámetros.

### 3.3.1 Código base para la aplicación de la regresión *Model Discovery*.

Para el código base usan los datos de una simulación, donde se va a tomar como ejemplo una placa de cobre cuadrada de  $5 \times 5 \text{ cm}$  de ancho por alto, muy delgada, con coeficiente de difusividad térmica  $\alpha = 1,17 \times 10^{-4} \text{ m/s}^2$  con las siguientes condiciones de frontera en el borde izquierdo  $T_1 = 25 \text{ }^\circ\text{C}$ , el borde superior  $T_2 = 0 \text{ }^\circ\text{C}$ , el borde derecho  $T_3 = 25 \text{ }^\circ\text{C}$ , y el borde inferior  $T_4 = 50 \text{ }^\circ\text{C}$ , con temperatura inicial  $T_0 = 10 \text{ }^\circ\text{C}$  para un tiempo de diez segundos. Ver Figura 4. Geometría para simular.

Una vez simulado, se guardan los datos de temperatura. Por tanto, la matriz de posiciones  $P_{[2 \times 761]}$  y  $T_{[761 \times 1001]}$ , es decir que se van a estudiar 761 nodos.

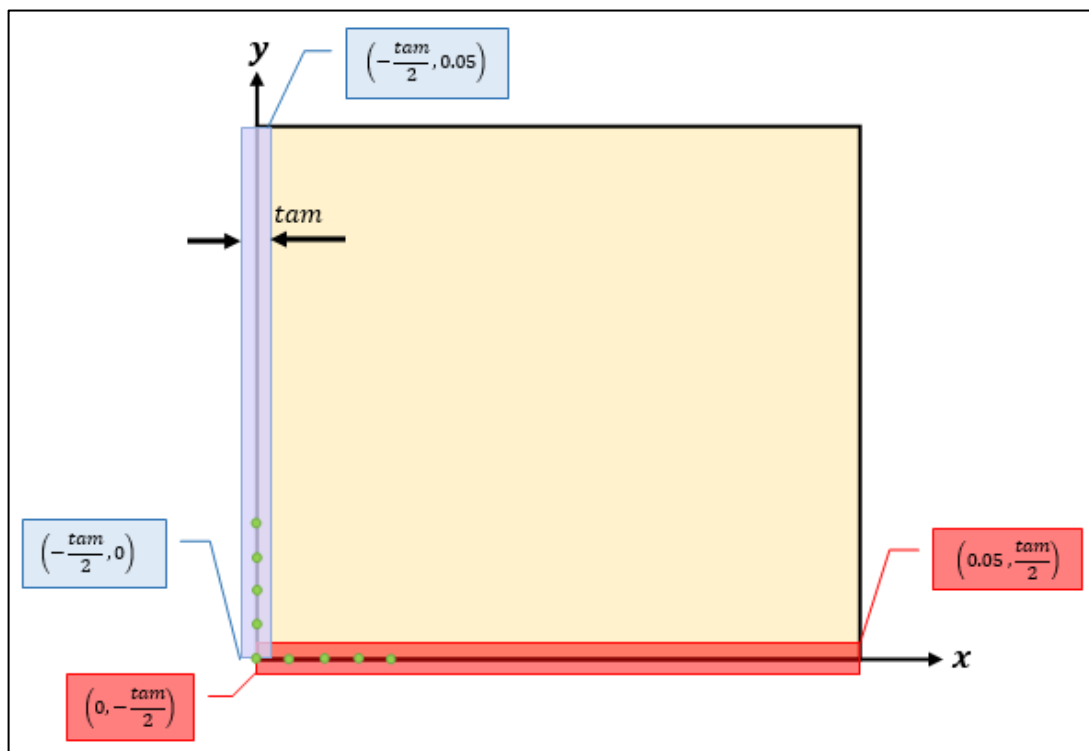
Figura 15. Diagrama de flujo del código base para la aplicación del *Model Discovery*



**3.3.1.1 Se cargan los datos.** El primer paso es limpiar todo, seguido de cargar los datos en un *script live* de Matlab, se carga con el comando (*load*), este arreglo tiene la particularidad que tiene 26 nodos en cada borde, se determina la relación de la longitud del lado de la placa y el número de nodos, esto se va a usar más adelante.

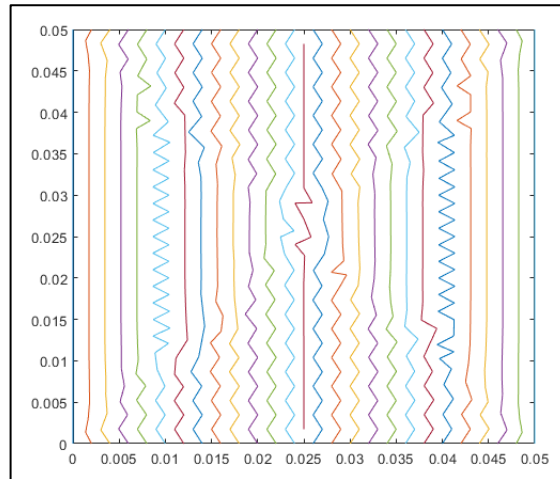
**3.3.1.2 Buscar los nodos correspondientes a una línea.** Para buscar los nodos se trazan pequeños rectángulos que barren una zona, como ejemplo, si se quiere buscar una línea de nodos paralela al eje  $y$ , se buscan nodos que estén en la dentro rectángulo que tiene como vértices los puntos  $p_1 = (-tam/2, 0)$  y  $p_2 = (tam/2, 0.05)$  estos puntos tienen coordenadas  $(x,y)$  respectivamente, luego el rectángulo se va corriendo para encontrar otros nodos, es decir que el rectángulo se corre una distancia de  $(tam)$ , lo mismo se hace para encontrar nodos paralelos al eje  $x$ . Gráficamente se vería de la siguiente manera.

Figura 16. Modelo de rectángulos para encontrar nodos en una misma línea



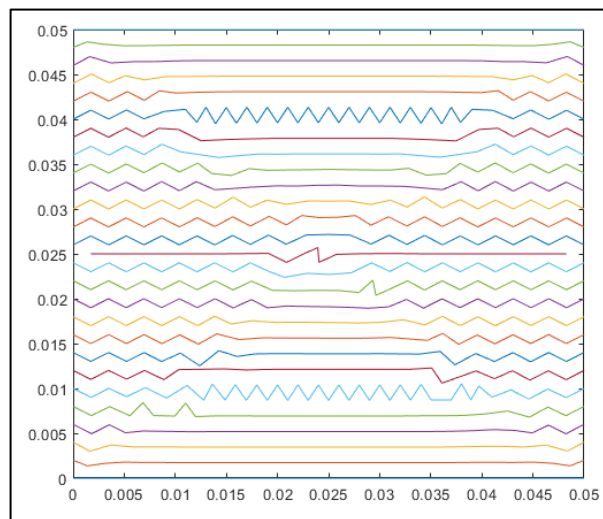
La búsqueda de las líneas de nodos se hace con el comando (*findNodes*) de Matlab, que busca en la malla, nodos dentro de un rectángulo con coordenadas específicas. Si existen dudas de cómo se utiliza el comando se puede recurrir al ayuda de la función.

Figura 17. Líneas verticales de nodos



Para buscar los nodos que se encuentren en líneas horizontales se hace de manera muy parecida a la forma como se buscan nodos en líneas verticales, se usa el mismo comando de Matlab, solo hoy que variar las coordenadas del rectángulo de búsqueda, se muestran solo dos líneas.

Figura 18. Líneas horizontales de nodos



**3.3.1.3 Ordenar y derivar espacialmente.** Una vez se tiene las líneas de nodos, hay que ordenarlas, porque, el comando de Matlab (*findNodes*), encuentra los nodos dentro de una caja de tamaño específico, pero estos no quedan ordenados. Por ejemplo, la primera línea vertical, se encuentra desde  $x = -tam/2$  hasta  $x = tam/2$  y desde  $y = 0$  hasta  $y = 0.05$  los nodos que están contenidos ahí no van a variar mucho en la coordenada  $x$ , pero se deben ordenar los nodos en la coordenada  $y$  ya sea de menor a mayor o viceversa, esto para poder derivar los nodos de la línea vertical respecto de  $y$ .

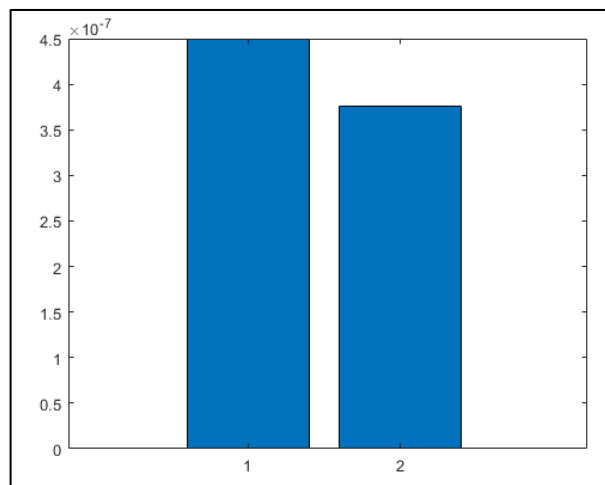
Una vez los nodos están ordenados se procede a derivar y guardar las derivadas en una variable.

**3.3.1.4 Derivadas respecto al tiempo.** Las derivadas respecto al tiempo son sencillas de calcular por la forma de los datos, en este caso no hay necesidad de ordenar, se calculan utilizando la primera fórmula que está en Tabla 6. Fórmulas de diferencias centrales de segundo orden.

**3.3.1.5 Organización de las matrices y construido de la librería de términos.** Existe una matriz de primeras derivadas y segundas derivadas para cada una de las líneas de nodos encontradas, ya sea vertical u horizontal, estas matrices hay que unificarlas, en una matriz que contenga todas las primeras derivadas de las líneas verticales, otra matriz que contenga las primeras derivadas de las líneas horizontales, de igual manera se hace con las matrices de segundas derivadas tanto para líneas verticales como horizontales. Una vez estén unificadas las matrices se convierten en un vector columna muy largo con el comando (*reshape*) de Matlab, también se hace con la matriz de derivadas respecto al tiempo, todos estos vectores columna deben tener el mismo tamaño. Al montar la librería solo se incluyen las segundas derivadas respecto de  $x$  y  $y$ , con el fin de obligar el sistema, pues es la primera prueba del algoritmo.

**3.3.1.6 Determinación de los parámetros.** Al tener el vector columna de las derivadas de la temperatura respecto del tiempo y la librería de términos, a partir de vectores columnas de los posibles términos, solo queda resolver el sistema lineal, ya sea mediante la pseudoinversa o con Lasso, en este caso se usó la pseudoinversa, al final se grafican los resultados en un diagrama de barras.

Figura 19. Diagrama de barras del resultado para el código base

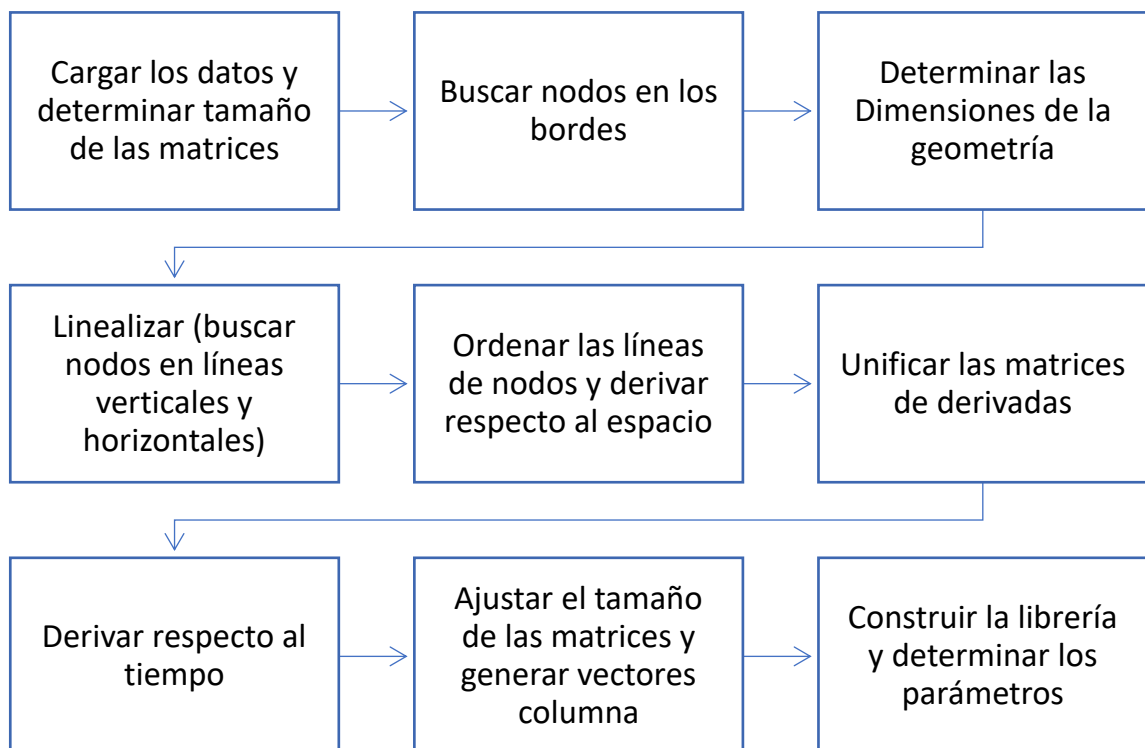


Existen varias limitaciones con el código base, la primera, el resultado da unos coeficientes muy alejados del esperado, la segunda, es muy largo y solo sirve para esos datos de temperatura iniciales, si se quieren cambiar los datos tendría que volverse a hacer todo el código, esto es tedioso, porque hay que hacer muchas pruebas para ver si se logra bajar el error, por eso, en base a este, hay que crear uno nuevo, donde se puedan hacer inmensas pruebas para encontrar la posible causa de error y tratar a reducirlo. El código se puede encontrar en el Anexo B. Código base d la implementación del Model Discovery en un Script Live de Matlab.

### 3.3.2 Optimización del código y nuevo algoritmo.

La optimización del código base generando un nuevo algoritmo, es una necesidad, ya que, con el código base no se pueden hacer variaciones en la data, ni usar una nueva data, donde se contengan mayor o menor cantidad de nodos, una geometría más grande o chica, un intervalo de tiempo más largo u más corto. Se quiere generar un algoritmo capaz de aplicar el método de regresión *Model Discovery* a cualquier data, proveniente de un fenómeno de difusividad térmica para dos dimensiones. El nuevo código (*script*) sigue los siguientes pasos.

Figura 20. Diagrama de flujo para algoritmo *Model Discovery* para PDE

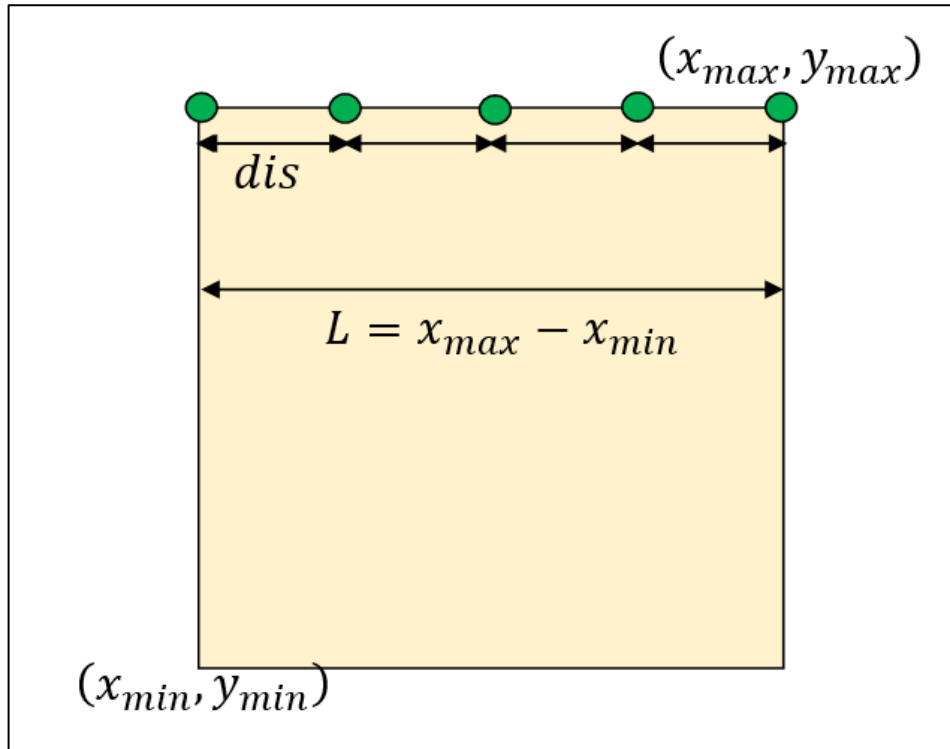


**3.3.2.1 Cargar los datos.** Al igual que para el código base se hace el proceso de limpiado, se cargan los datos y se determina el tamaño de las matrices que contienen la información, el tamaño se guarda en unas variables y se usa más adelante.

**3.3.2.2 Buscar la cantidad de nodos en los bordes.** Los datos provienen de una placa plana cuadrada, la cual tiene cuatro bordes, además cuenta con la ventaja de que todos los bordes tienen la misma cantidad de nodos, esto hace que se pueda usar un ciclo *for* y almacenar la data dentro de una matriz, el tamaño de dicha matriz está dado por el número de bordes y la cantidad de nodos, esto último es lo que se necesita. Dentro del ciclo se utiliza el comando *findNodes*, el cual, ya se había usado antes, con la diferencia que buscaba una caja, un rectángulo, esta vez busca los nodos en una región, es decir, en los bordes.

**3.3.2.3 Determinar las dimensiones de la geometría.** La matriz de posiciones guarda las coordenadas  $(x, y)$  de la geometría, es una matriz de tamaño  $(2 \times n)$  donde  $n$  es el número total de nodos, la primera fila de la matriz corresponde a las coordenadas  $x$  de cada nodo y la segunda fila a las coordenadas en  $y$ , al tener como precedente solo es buscar el valor máximo y mínimo de cada una de las filas, y con la resta de estos valores ya se tiene la longitud de la geometría, con esta longitud se puede determinar la distancia que hay entre cada nodo que está en los bordes, esta distancia es de vital importancia para la construcción de los rectángulos de búsqueda de nodos en líneas horizontales y verticales.

Figura 21. Tamaño de la geometría y distancia entre nodos de los bordes



De la Figura 21 se puede ver que si en un borde hay una cantidad específica de nodos, el número de espacios que hay entre los nodos va a ser igual al número de nodos disminuido en uno. Como lo muestra la siguiente ecuación.

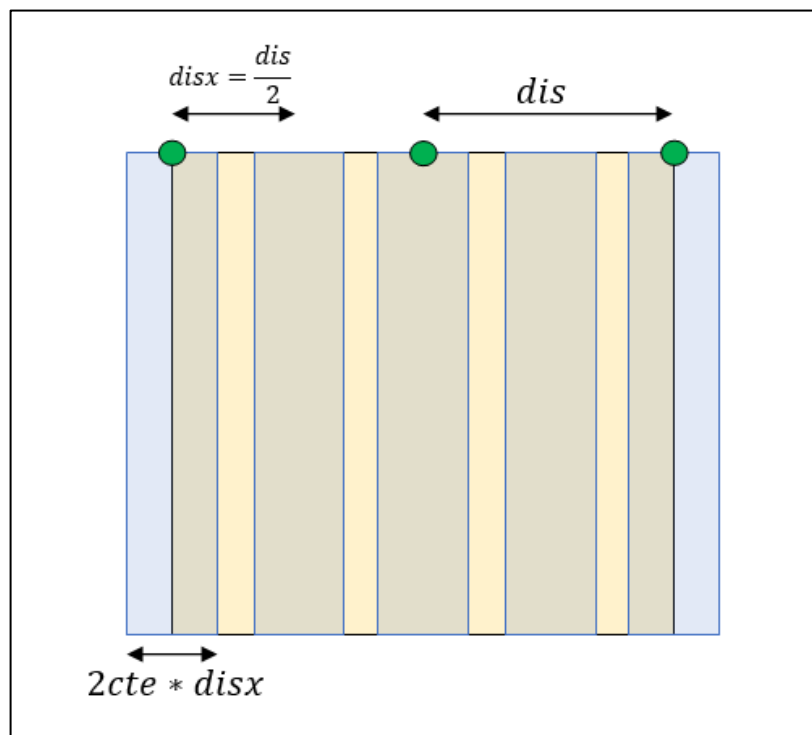
$$\#E_{nodos} = (n - 1) \quad (32)$$

$$dis = \frac{L}{\#E_{nodos}} = \frac{x_{max} - x_{min}}{n - 1} \quad (33)$$

Donde,  $\#E_{nodos}$  es el número de espacios entre nodos,  $n$  el número de nodos en un borde,  $L$  la distancia en entre el nodo más lejano y el más cercano y  $dis$  es la distancia entre los nodos del borde. Se hace de la misma manera para determinar las dimensiones en el eje  $y$ .

**3.3.2.4 Buscar los nodos de líneas verticales y horizontales.** Al igual que con el código base la búsqueda de nodos que estén en líneas horizontales y verticales se hace mediante rectángulos con el comando *findNodes*, existe un inconveniente, que todas las líneas verticales u horizontales no tienen la misma cantidad de nodos, esto hace que correr un ciclo *for* y almacenar en una matriz sea imposible, se le da solución almacenando en celdas (*cell*) el nombre de los vectores. Los parámetros de el rectángulo de búsqueda se ven en la siguiente figura.

Figura 22. Parámetros del rectángulo de búsqueda



De la figura anterior se tienen tres nodos, por tanto, el número de espacios entre nodos es dos, se deduce de la ecuación (32), al dividir la longitud en el número de espacios, da la distancia entre nodos, se deduce de la ecuación (33), si esta

distancia entre nodos la dividimos en dos tenemos la distancia  $disx$ , esta distancia cabe cuatro veces en la longitud total del borde, para barrer toda la longitud se necesitan cinco rectángulos, es decir, el doble del número de nodos disminuido en uno, de ahí sale el número de búsquedas a lo largo de la geometría, además el ancho máximo del rectángulo de búsqueda puede ser igual al  $disx$  es decir que el máximo valor que puede tomar la  $cte = 1/2$ . El número de búsquedas se representa en la siguiente ecuación.

$$\#_{búsquedas} = 2n - 1 \quad (34)$$

Donde,  $\#_{búsquedas}$  es el número de búsquedas de líneas de nodos que va a hacer el algoritmo,  $n$  es el número de nodos en un borde.

**3.3.2.5 Ordenar y derivar respecto al espacio.** Cada línea de nodos está almacenada en un arreglo de celdas, ya sean líneas verticales (columnas) o líneas horizontales (filas), el primer paso es sacar la primera línea de nodos del arreglo de celdas, la cual es un vector con el nombre de los nodos correspondientes a esa línea, una vez se saca el vector que contiene los nombres de los nodos, se halla el tamaño de ese vector, es decir la cantidad de nodos de esa línea, para en un ciclo *for* crear una nueva matriz que contenga los datos de posición de cada nodo correspondiente a la línea, y de la misma forma una nueva matriz de temperatura, luego se ordenan las matrices, para esto se creó una función llamada (*funordmatrix*), la cual ordena una matriz de acuerdo a filas u columnas en orden descendente o ascendente, con esta función se ordena el nuevo vector de posición, si se quiere ordenar una columna, entonces se ordena respecto a  $y$ , la función guarda el orden en un vector, lo que se hace es crear las matrices ordenadas tanto de posición como de temperatura en función del vector ordenado y de las matrices que corresponden a la línea, las cuales se crearon anteriormente, una vez se ordena se pasa a derivar esa línea ordenada, se usan las fórmulas de la Tabla 6. Fórmulas de diferencias

centrales de segundo orden, o también se usan las ecuaciones (27) y (28) provenientes de un polinomio de segundo orden de Lagrange, con las que se pueden derivar puntos no equidistantes.

Para ordenar y derivar filas es muy similar, hay que hacer algunos cambios en el nombre de las variables, y en la función (*funordmatrix*) que cambia un parámetro, si se quiere obtener más información sobre esta función ver el Anexo D. Función (*funordmatrix*).

**3.3.2.6 Unificar las matrices.** Para unificar las matrices de derivadas espaciales se usa un ciclo *for*, en el cual se crea una matriz a la que se va agregando matriz por matriz de cada línea de nodos, ya sean líneas horizontales o verticales.

**3.3.2.7 Derivar respecto al tiempo.** No cambian respecto al código base, en este caso no hay necesidad de ordenar, se calculan utilizando la primera formula que está en Tabla 6. Fórmulas de diferencias centrales de segundo orden.

**3.3.2.8 Ajustar el tamaño de las matrices y generar vectores columna.** Esta parte se hace manual, dependiendo de la data que se maneje, se cambia el valor dos variables  $r$  y  $s$ , se hace con el fin de dejar todas las matrices del mismo tamaño y poder reorganizarlas en vectores columna muy largos, los vectores columna se redondean con cinco cifras significativas.

**3.3.2.9 Construir la librería de términos y determinación de los parámetros.** Se crea la librería de términos ordenando, los vectores columna de las segundas derivadas espaciales al igual que en el código base y metiéndolos en una matriz, se hace con el fin de obligar el algoritmo a dar solo estos dos coeficientes,

se resuelve el sistema lineal, ya sea mediante la pseudoinversa o con Lasso, en este caso se usó la psudoinversa, al final se grafican los resultados en un diagrama de barras.

Para ver el algoritmo completo ir al Anexo C. Algoritmo Model Discovery para PDEs en un Script Live de Matlab.

#### 4. PRUEBAS Y DISMINUCIÓN DEL ERROR

En esta sección se van a hacer pruebas del algoritmo variando algunos parámetros de la data, se calcula el error comparando el coeficiente de la difusividad térmica usado para la simulación y los coeficientes que da como resultado el algoritmo, ya sea usando la pseudoinversa o Lasso, en el caso de Lasso hay que variar el termino de penalización  $\alpha$ .

$$e = \frac{teórico - real}{teórico} * 100 \quad (35)$$

Donde,  $e$  representa el porcentaje de error,  $teórico$  el valor teórico del coeficiente de difusividad térmica, el que se usó para la simulación y  $real$  el valor real del coeficiente de difusividad térmica, el que se halla con el algoritmo *Model Discovery*.

Al usar derivadas de centrales de segundo orden por cada línea de nodos que se deriven se pierden dos datos, es decir, se pierden datos igual al doble del número de búsquedas, por esta razón los vectores columna de las derivadas espaciales van a quedar de menor tamaño que el vector de derivadas respecto del tiempo, como este vector queda de mayor tamaño, se va a tomar la data de este vector en dos muestras, al inicio del vector y al final del vector. La ecuación de perdida de datos viene dada de la siguiente manera.

$$P_{datos} = 4n - 2 \quad (36)$$

Donde,  $P_{datos}$  representa la perdida de datos por la simulación,  $n$  es el número de nodos en los bordes.

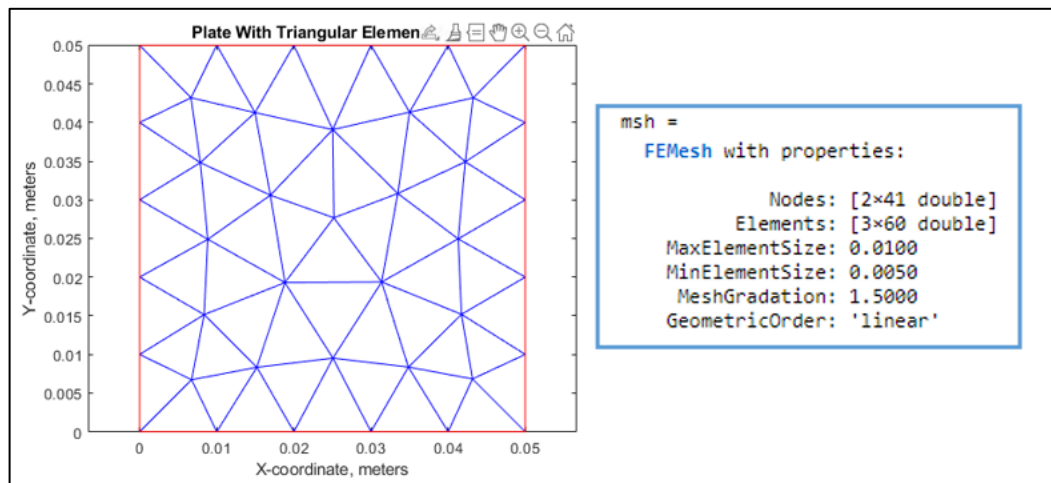
## 4.1 AUMENTO DE LA CANTIDAD DE NODOS

Para esta sección se va a hacer pruebas, para los datos obtenidos de una simulación de una placa de cobre cuadrada de  $5 \times 5 \text{ cm}$  de ancho por alto, muy delgada, con coeficiente de difusividad térmica  $\alpha = 1,17 \times 10^{-4} \text{ m/s}^2$  con las siguientes condiciones de frontera en el borde izquierdo  $T_1 = 100 \text{ }^\circ\text{C}$ , el borde superior  $T_2 = 0 \text{ }^\circ\text{C}$ , el borde derecho  $T_3 = 0 \text{ }^\circ\text{C}$ , y el borde inferior  $T_4 = 100 \text{ }^\circ\text{C}$ , con temperatura inicial  $T_0 = 10 \text{ }^\circ\text{C}$ , para un vector de tiempo  $t = 0:0.01:10\text{s}$ . Para un tamaño máximo del elemento de  $0.01\text{m}$ ,  $0.007\text{m}$ ,  $0.004\text{m}$ ,  $0.002\text{m}$  y  $0.001\text{m}$ , lineal y solo se va a tener en cuenta el resultado que da la pseudoinversa y las derivadas espaciales se resuelven por diferencias centrales para puntos equidistantes.

### 4.1.1 Tamaño máximo del elemento 0.01m.

La data de la simulación queda guardada como *Thmax1.mat*, en el Anexo E. Datos de las simulaciones.

Figura 23. Malla de tamaño máximo de elemento 0.01m



El arreglo de la figura anterior da una matriz  $p = [2 \times 41]$  y  $T = [41 \times 1001]$ , al aplicar el algoritmo *Model Discovery* a esta data, las matrices unificadas de las derivadas respecto del espacio  $T_x = [18 \times 1001]$ ,  $T_{xx} = [18 \times 1001]$ ,  $T_y = [17 \times 1001]$ ,  $T_{yy} = [17 \times 1001]$ , y la derivada de los datos respecto de la temperatura  $T_{dot} = [41 \times 999]$ , por tanto  $r = 17$  toma el valor de la fila de la matriz con el menor número de filas y  $s = 999$  toma el valor de la columna de la matriz con el menor de columnas. Por tanto, el código para tomar la muestra en la parte inicial y final del vector  $T_{dot}$  queda de la siguiente manera.

```
r=17;
s=999;
T_dot=Tdot(1,r,1:s);% para la parte inicial del vector
T_dot=Tdot(end-r+1:end,1:s); % para la parte final del vector
```

Se hace ese cambio y luego se determina el resultado

Tabla 9. Resultados Hmax 0.01m, Pseudoinversa y DCE

| <b>Coficiente</b> | <b>Inicio <math>T_{dot}</math></b> | <b>Final <math>T_{dot}</math></b> |
|-------------------|------------------------------------|-----------------------------------|
| $U_{xx}$          | 0                                  | $7.8724 \times 10^{-6}$           |
| $U_{yy}$          | 0                                  | $7.6209 \times 10^{-6}$           |

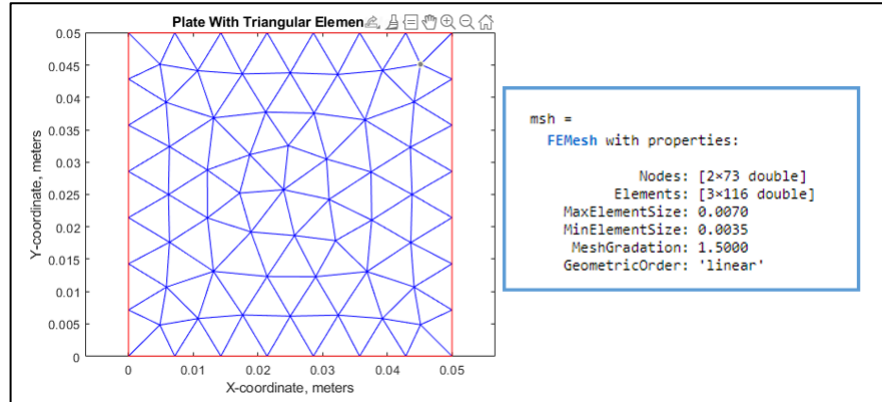
#### 4.1.2 Tamaño máximo del elemento 0.007m.

La data de la simulación queda guardada como *Thmax2.mat*, en el Anexo E.

Datos de las simulaciones.

Para este caso el tamaño máximo del arreglo triangular de la malla es de 0.007m, además son triángulos lineales, es decir, cuenta con tres nodos.

Figura 24. Malla de tamaño máximo de elemento 0.007m



El arreglo de la figura anterior da una matriz  $p = [2 \times 73]$  y  $T = [73 \times 1001]$ , al aplicar el algoritmo Model Discovery a esta data, las matrices unificadas de las derivadas respecto del espacio  $T_x = [38 \times 1001]$ ,  $T_{xx} = [38 \times 1001]$ ,  $T_y = [38 \times 1001]$ ,  $T_{yy} = [38 \times 1001]$ , y la derivada de los datos respecto de la temperatura  $T_{dot} = [73 \times 999]$ , por tanto  $r = 38$  toma el valor de la fila de la matriz con el menor número de filas y  $r = 999$  toma el valor de la columna de la matriz con el menor de columnas. Por tanto, el código para tomar la muestra en la parte inicial y final del vector  $T_{dot}$  queda de la siguiente manera.

```
r=38;
s=999;
T_dot=Tdot(1,r,1:s);% para la parte inicial del vector
T_dot=Tdot(end-r+1:end,1:s); % para la parte final del vector
```

Se hace ese cambio y luego se determina el resultado

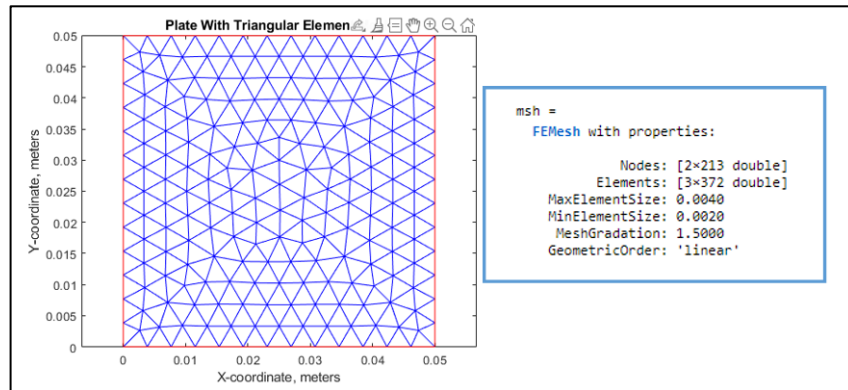
Tabla 10. Resultados Hmax 0.007m, Pseudoinversa y DCE

| <b>Coficiente</b> | <b>Inicio <math>T_{dot}</math></b> | <b>Final <math>T_{dot}</math></b> |
|-------------------|------------------------------------|-----------------------------------|
| $U_{xx}$          | $1.7329 \times 10^{-6}$            | $1.4746 \times 10^{-5}$           |
| $U_{yy}$          | $6.4566 \times 10^{-8}$            | $-1.2276 \times 10^{-5}$          |

### 4.1.3 Tamaño máximo del elemento 0.004m.

La data de la simulación queda guardada como *Thmax3.mat*, en el Anexo E. Datos de las simulaciones.

Figura 25. Malla de tamaño máximo de elemento 0.004m



El arreglo de la figura anterior da una matriz  $p = [2 \times 213]$  y  $T = [213 \times 1001]$ , al aplicar el algoritmo Model Discovery a esta data, las matrices unificadas de las derivadas respecto del espacio  $T_x = [125 \times 1001]$ ,  $T_{xx} = [125 \times 1001]$ ,  $T_y = [135 \times 1001]$ ,  $T_{yy} = [135 \times 1001]$ , y la derivada de los datos respecto de la temperatura  $T_{dot} = [213 \times 999]$ , por tanto  $r = 125$  toma el valor de la fila de la matriz con el menor número de filas y  $r = 999$  toma el valor de la columna de la matriz con el menor de columnas. Por tanto, el código para tomar la muestra en la parte inicial y final del vector  $T_{dot}$  queda de la siguiente manera.

```
r=125;  
s=999;  
T_dot=Tdot(1,r,1:s);% para la parte inicial del vector  
T_dot=Tdot(end-r+1:end,1:s); % para la parte final del vector
```

Se hace ese cambio y luego se determina el resultado.

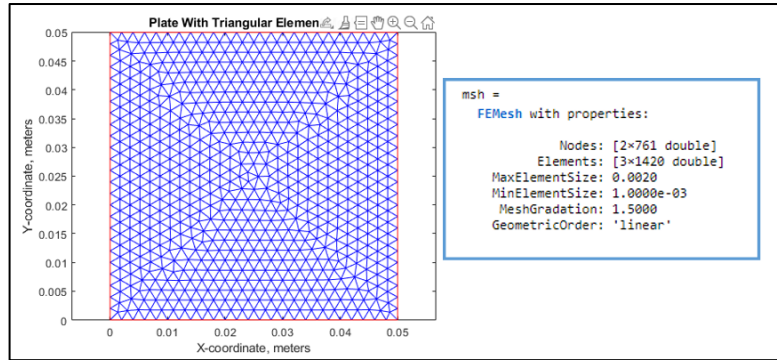
Tabla 11. Resultados Hmax 0.004m, Pseudoinversa y DCE

| <b>Coefficiente</b> | <b>Inicio <i>Tdot</i></b> | <b>Final <i>Tdot</i></b> |
|---------------------|---------------------------|--------------------------|
| $U_{xx}$            | $8.7199 \times 10^{-7}$   | $1.0423 \times 10^{-6}$  |
| $U_{yy}$            | $1.2194 \times 10^{-6}$   | $9.5858 \times 10^{-7}$  |

#### 4.1.4 Tamaño máximo del elemento 0.002m.

La data de la simulación queda guardada como *Thmax4.mat*, en el Anexo E. Datos de las simulaciones.

Figura 26. Malla de tamaño máximo de elemento 0.002m



El arreglo de la figura anterior da una matriz  $p = [2 \times 761]$  y  $T = [761 \times 1001]$ , al aplicar el algoritmo Model Discovery a esta data, las matrices unificadas de las derivadas respecto del espacio  $T_x = [584 \times 1001]$ ,  $T_{xx} = [584 \times 1001]$ ,  $T_y = [586 \times 1001]$ ,  $T_{yy} = [586 \times 1001]$ , y la derivada de los datos respecto de la temperatura  $Tdot = [761 \times 999]$ , por tanto  $r = 584$  toma el valor de la fila de la matriz con el menor número de filas y  $r = 999$  toma el valor de la columna de la

matriz con el menor de columnas. Por tanto, el código para tomar la muestra en la parte inicial y final del vector  $T\_dot$  queda de la siguiente manera.

```
r=584;
s=999;
T_dot=Tdot(1,r,1:s);% para la parte inicial del vector
T_dot=Tdot(end-r+1:end,1:s); % para la parte final del vector
```

Se hace ese cambio y luego se determina el resultado.

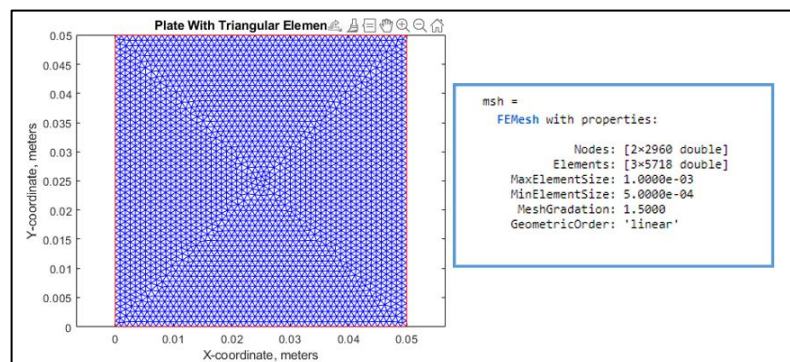
Tabla 12. Resultados Hmax 0.002m, Pseudoinversa y DCE

| <b>Coficiente</b> | <b>Inicio <math>T\_dot</math></b> | <b>Final <math>T\_dot</math></b> |
|-------------------|-----------------------------------|----------------------------------|
| $U_{xx}$          | $6.0150 \times 10^{-7}$           | $4.8380 \times 10^{-7}$          |
| $U_{yy}$          | $6.1072 \times 10^{-7}$           | $5.0416 \times 10^{-7}$          |

#### 4.1.5 Tamaño máximo del elemento 0.001m.

La data de la simulación queda guardada como *Thmax5.mat*, en el Anexo E. Datos de las simulaciones.

Figura 27. Malla de tamaño máximo de elemento 0.001m



El arreglo de la figura anterior da una matriz  $p = [2 \times 2960]$  y  $T = [2960 \times 1001]$ , al aplicar el algoritmo Model Discovery a esta data, las matrices unificadas de las derivadas respecto del espacio  $T_x = [2457 \times 1001]$ ,  $T_{xx} = [2457 \times 1001]$ ,  $T_y = [2474 \times 1001]$ ,  $T_{yy} = [2474 \times 1001]$ , y la derivada de los datos respecto de la temperatura  $T_{dot} = [2960 \times 999]$ , por tanto  $r = 2457$  toma el valor de la fila de la matriz con el menor número de filas y  $s = 999$  toma el valor de la columna de la matriz con el menor de columnas. Por tanto, el código para tomar la muestra en la parte inicial y final del vector  $T_{dot}$  queda de la siguiente manera.

```
r=2457;
s=999;
T_dot=Tdot(1,r,1:s);% para la parte inicial del vector
T_dot=Tdot(end-r+1:end,1:s); % para la parte final del vector
```

Se hace ese cambio y luego se determina el resultado

Tabla 13. Resultados Hmax 0.001m, Pseudoinversa y DCE

| <b>Coeficiente</b> | <b>Inicio <math>T_{dot}</math></b> | <b>Final <math>T_{dot}</math></b> |
|--------------------|------------------------------------|-----------------------------------|
| $U_{xx}$           | $1.9094 \times 10^{-7}$            | $1.1938 \times 10^{-7}$           |
| $U_{yy}$           | $1.3664 \times 10^{-6}$            | $1.3836 \times 10^{-7}$           |

#### 4.1.6 Análisis de datos y resultados.

En este apartado se reúne la data recolectada en los experimentos anteriores, donde se fue aumentando el número de nodos, en otras palabras, la cantidad de puntos de estudio, a través, de la disminución del tamaño de los triángulos que componen la malla, los resultados se tabularon, teniendo como base el hecho que las segundas derivadas espaciales son las que componen la librería de términos. Como se evidencia en la recolección de la data, el vector de derivadas espaciales es de mayor tamaño, ya que, se pierde data en las derivadas espaciales, por eso

los datos tabulados se toman en referencia a dividir el vector de derivadas en el tiempo en dos. Los resultados del error se calculan con la ecuación (35), el resultado teórico es el coeficiente de difusividad térmica del cobre  $\alpha = 1,17 \times 10^{-4}$ ,

Tabla 14. Resultados para la primera parte del vector de derivas temporales

| <b>Coeficiente</b> | <b><math>h_{max}</math><br/>= 0,01</b> | <b><math>h_{max}</math><br/>= 0,007</b> | <b><math>h_{max}</math><br/>= 0,004</b> | <b><math>h_{max}</math><br/>= 0,002</b> | <b><math>h_{max}</math><br/>= 0,001</b> |
|--------------------|--|---|---|---|---|
| U <sub>xx</sub>    | 0                                      | 1.7329E <sup>-6</sup>                   | 8.7199E <sup>-7</sup>                   | 6.015E <sup>-7</sup>                    | 1.9094E <sup>-7</sup>                   |
| U <sub>yy</sub>    | 0                                      | 6.4566E <sup>-8</sup>                   | 1.2194E <sup>-6</sup>                   | 6.1072E <sup>-7</sup>                   | 1.3664E <sup>-7</sup>                   |

Tabla 15. Resultados para la segunda parte del vector de derivas temporales

| <b>Coeficiente</b> | <b><math>h_{max}</math><br/>= 0,01</b> | <b><math>h_{max}</math><br/>= 0,007</b> | <b><math>h_{max}</math><br/>= 0,005</b> | <b><math>h_{max}</math><br/>= 0,002</b> | <b><math>h_{max}</math><br/>= 0,001</b> |
|--------------------|--|---|---|---|---|
| U <sub>xx</sub>    | 7.8724E <sup>-6</sup>                  | 1.4746E <sup>-5</sup>                   | 1,0423E <sup>-6</sup>                   | 4.8380E <sup>-7</sup>                   | 1.9038E <sup>-7</sup>                   |
| U <sub>yy</sub>    | 7.6209E <sup>-6</sup>                  | -1.2276E <sup>-5</sup>                  | 9.5858E <sup>-7</sup>                   | 5.0416E <sup>-7</sup>                   | 1.3836E <sup>-7</sup>                   |

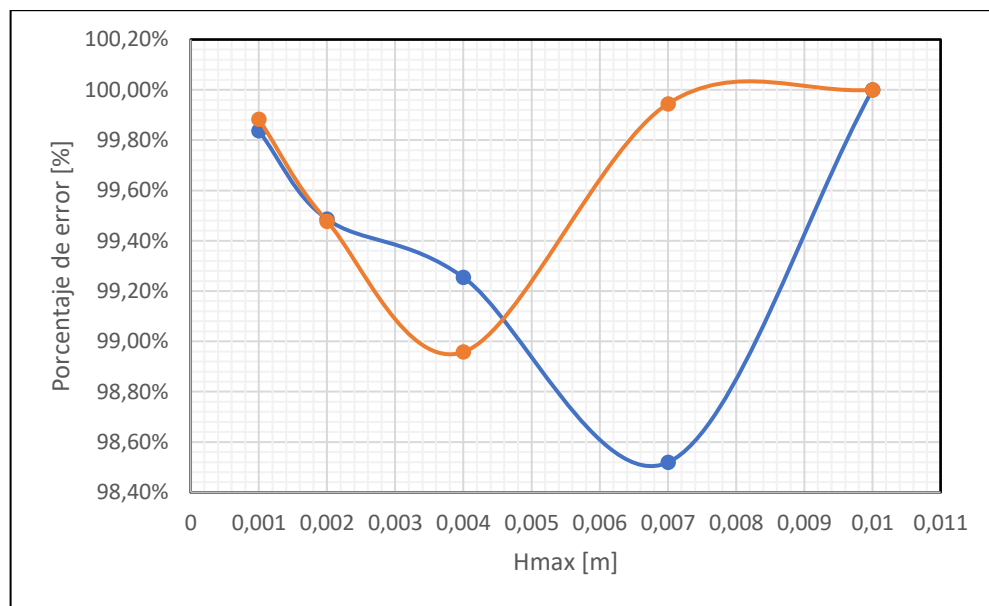
Tabla 16. Porcentaje de error para la primera parte del vector derivadas temporales

| <b>error</b>    | <b><math>h_{max}</math><br/>= 0,01</b> | <b><math>h_{max}</math><br/>= 0,007</b> | <b><math>h_{max}</math><br/>= 0,005</b> | <b><math>h_{max}</math><br/>= 0,002</b> | <b><math>h_{max}</math><br/>= 0,001</b> |
|-----------------|--|---|---|---|---|
| U <sub>xx</sub> | 100.00%                                | 98,52%                                  | 99,25%                                  | 99,49%                                  | 99,84%                                  |
| U <sub>yy</sub> | 100.00%                                | 99,94%                                  | 98,96%                                  | 99,48%                                  | 99,88%                                  |

Tabla 17. Porcentaje de error para la segunda parte del vector derivadas temporales

| <i>error</i> | $h_{max}$<br>= 0,01 | $h_{max}$<br>= 0,007 | $h_{max}$<br>= 0,005 | $h_{max}$<br>= 0,002 | $h_{max}$<br>= 0,001 |
|--------------|---------------------|----------------------|----------------------|----------------------|----------------------|
| $U_{xx}$     | 93.27%              | 87,40%               | 99,11%               | 99,59%               | 99,84%               |
| $U_{yy}$     | 93.49%              | 110,49%              | 99,18%               | 99,57%               | 99,88%               |

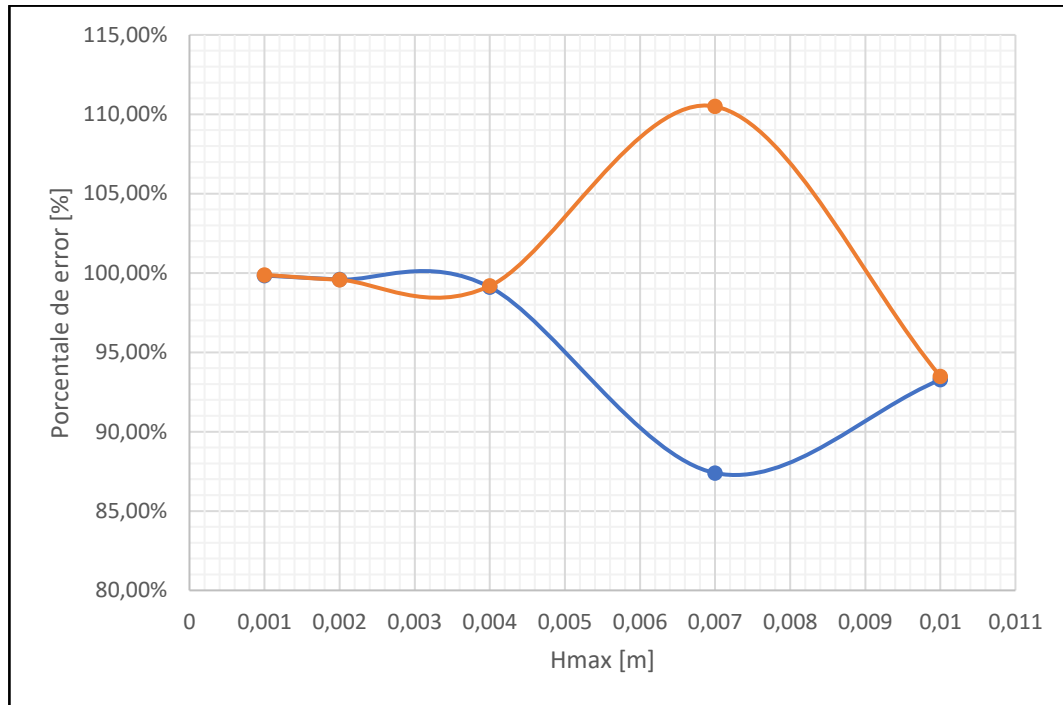
Figura 28. Grafica del error para la primera parte del vector de derivadas temporales



La Figura 28 muestra el comportamiento del porcentaje de error a medida que se aumenta el número de nodos, disminuye el tamaño de los triangulo elementos de la malla, para la primera parte del vector  $\dot{T}$ , en color azul el coeficiente que pertenece a  $\partial^2 T / \partial x^2$  y en color naranja el coeficiente que pertenece a  $\partial^2 T / \partial y^2$ , se evidencia que el error aumenta con el aumento de nodos, los valores de error son muy cercanos, a excepción, de los datos correspondientes al tamaño de malla  $h_{max} = 0.007m$ . donde se encuentra el mínimo valor de error para el coeficiente correspondiente a  $\partial^2 T / \partial x^2$  y el máximo valor de error para el coeficiente

correspondiente a  $\partial^2 T / \partial y^2$ , con una separación cercana a los dos puntos porcentuales.

Figura 29. Grafica del error para la segunda parte del vector de derivadas temporales



La Figura 29 muestra el comportamiento del porcentaje de error a medida que se aumenta el número de nodos, disminuye el tamaño de los triangulo elementos de la malla, para la primera parte del vector  $\dot{T}$ , en color azul el coeficiente que pertenece a  $\partial^2 T / \partial x^2$  y en color naranja el coeficiente que pertenece a  $\partial^2 T / \partial y^2$ , se evidencia un comportamiento similar a la Figura 28, donde de los datos correspondientes al tamaño de malla  $h_{max} = 0.007m$ . se encuentra el mínimo valor de error para el coeficiente correspondiente a  $\partial^2 T / \partial x^2$  y el máximo valor de error para el coeficiente correspondiente a  $\partial^2 T / \partial y^2$ , con una separación de más de 20 puntos porcentuales.

Tanto para la Figura 28 como para la Figura 29 los errores son muy altos, un posible factor de error puede ser el uso de diferencias centrales, a pesar de que se modificó un poco la formula, no se tiene en cuenta  $dx$  o un  $dy$  constantes, sino los valores reales de separación de los puntos, en teoría al discretizar más el arreglo de la malla, es decir aumentar el número de nodos para estudiar, se reduce el error, ya que el concepto de derivada funciona mejor al hacer más pequeño el espacio entre datos.

## **4.2 COMPARACIÓN DE LOS RESULTADOS USANDO DIFERENCIAS CENTRALES EQUIDISTANTES Y NO EQUIDISTANTES.**

Para este apartado, se van a tomar usar los datos simulados a partir de los arreglos de la Figura 24. Malla de tamaño máximo de elemento 0.007m y la Figura 26. Malla de tamaño máximo de elemento 0.002m, comparando los resultados de los parámetros, al usar el modelo de derivadas para puntos no equidistantes, para las derivadas espaciales, con los resultados dados con el método de diferencias centrales de segundo orden.

### **4.2.1 Tamaño máximo del elemento 0.007m.**

Se ejecuta el algoritmo de *Model Discovery*, con la data de la simulación guardada como *Thmax2.mat*, se usa el método de diferenciación para puntos no equidistantes, para obtener las derivadas espaciales. Se toma la muestra en la parte inicial y final del vector  $T\_dot$ , ya que, este queda más largo que los otros vectores. Los resultados se muestran en la siguiente tabla.

Tabla 18. Resultados Hmax 0.007m, Pseudoinversa y DC no equidistantes

| <b>Coefficiente</b> | <b>Inicio <math>T_{dot}</math></b> | <b>Final <math>T_{dot}</math></b> |
|---------------------|------------------------------------|-----------------------------------|
| $U_{xx}$            | $1.5380 \times 10^{-6}$            | $1.0740 \times 10^{-5}$           |
| $U_{yy}$            | $5.3661 \times 10^{-7}$            | $-7,9375 \times 10^{-6}$          |

#### 4.2.2 Tamaño máximo del elemento 0.004m.

Se ejecuta el algoritmo de *Model Discovery*, con la data de la simulación guardada como *Thmax3.mat*, se usa el método de diferenciación para puntos no equidistantes, para obtener las derivadas espaciales. Se toma la muestra en la parte inicial y final del vector  $T_{dot}$ , ya que, este queda más largo que los otros vectores. Los resultados se muestran en la siguiente tabla.

Tabla 19. Resultados Hmax 0.004m, Pseudoinversa y DC no equidistantes

| <b>Coefficiente</b> | <b>Inicio <math>T_{dot}</math></b> | <b>Final <math>T_{dot}</math></b> |
|---------------------|------------------------------------|-----------------------------------|
| $U_{xx}$            | $9.8283 \times 10^{-7}$            | $1.4021 \times 10^{-6}$           |
| $U_{yy}$            | $1.4939 \times 10^{-6}$            | $1.0517 \times 10^{-6}$           |

#### 4.2.3 Tamaño máximo del elemento 0.002m.

Se ejecuta el algoritmo de *Model Discovery*, con la data de la simulación guardada como *Thmax4.mat*, se usa el método de diferenciación para puntos no equidistantes, para obtener las derivadas espaciales. Se toma la muestra en la parte inicial y final del vector  $T_{dot}$ , ya que, este queda más largo que los otros vectores. Los resultados se muestran en la siguiente tabla.

Tabla 20. Resultados Hmax 0.002m, Pseudoinversa y DC no equidistantes

| <b>Coefficiente</b> | <b>Inicio <i>Tdot</i></b> | <b>Final <i>Tdot</i></b> |
|---------------------|---------------------------|--------------------------|
| U <sub>xx</sub>     | $6.8774 \times 10^{-7}$   | $5.3264 \times 10^{-7}$  |
| U <sub>yy</sub>     | $6.8932 \times 10^{-7}$   | $5.7533 \times 10^{-7}$  |

#### 4.2.4 Análisis de resultados.

Se analizan los resultados de correr el algoritmo *Model Discovery*, centrándose en la diferencia de efectos, al aplicar derivadas espaciales, a partir de fórmulas para puntos equidistantes y no equidistantes representadas con las siglas DCE y DNE respectivamente, como el vector de derivadas temporales es de mayor tamaño, este se divide en dos partes, una parte inicial y otra parte final, para aumentos en la cantidad de nodos a partir de la disminución del tamaño de los triángulos de la malla, los resultados se ven reflejados en las siguientes tablas y graficas.

Tabla 21. Resultados para la primera parte del vector de derivas temporales para DCE y DNE

|            | <b>coeficiente</b> | <b><math>h_{max} = 0.007</math></b> | <b><math>h_{max} = 0.004</math></b> | <b><math>h_{max} = 0.002</math></b> |
|------------|--------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <b>DCE</b> | U <sub>xx</sub>    | $1.7329E^{-6}$                      | $8.7199E^{-7}$                      | $6.0150E^{-7}$                      |
|            | U <sub>yy</sub>    | $6.4566E^{-8}$                      | $1.2194E^{-6}$                      | $6.1072E^{-7}$                      |
| <b>DNE</b> | U <sub>xx</sub>    | $1.5380E^{-6}$                      | $9.8283E^{-7}$                      | $6.8774E^{-7}$                      |
|            | U <sub>yy</sub>    | $5.3361E^{-7}$                      | $1.4939E^{-6}$                      | $6.8932E^{-7}$                      |

Tabla 22. Resultados para la segunda parte del vector de derivas temporales para DCE y DNE

|            | <b>coeficiente</b> | <b><math>h_{max} = 0.007</math></b> | <b><math>h_{max} = 0.004</math></b> | <b><math>h_{max} = 0.002</math></b> |
|------------|--------------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <b>DCE</b> | $U_{xx}$           | $1.4746E^{-5}$                      | $1.0423E^{-6}$                      | $4.8380E^{-7}$                      |
|            | $U_{yy}$           | $-1.2276E^{-5}$                     | $9.5858E^{-7}$                      | $5.0416E^{-7}$                      |
| <b>DNE</b> | $U_{xx}$           | $1.0740E^{-5}$                      | $1.4021E^{-6}$                      | $5.3264E^{-7}$                      |
|            | $U_{yy}$           | $-7.9375E^{-6}$                     | $1.0517E^{-6}$                      | $5.7533E^{-7}$                      |

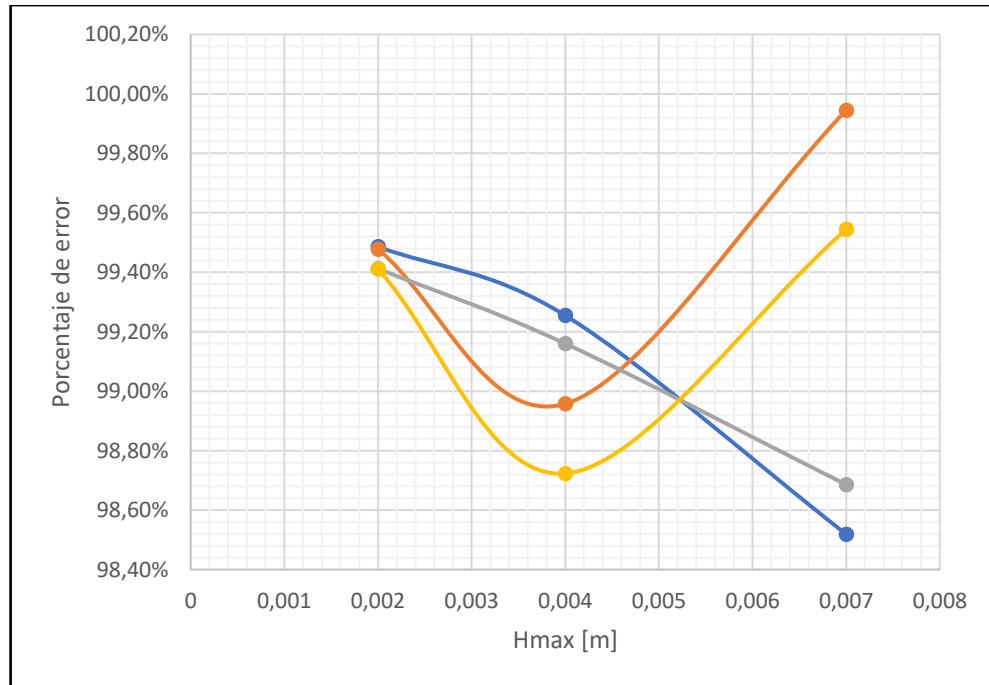
Tabla 23. Porcentaje de error para la primera parte del vector derivadas temporales para DCE y DNE

|            | <b>Error</b> | <b><math>h_{max} = 0.007</math></b> | <b><math>h_{max} = 0.004</math></b> | <b><math>h_{max} = 0.002</math></b> |
|------------|--------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <b>DCE</b> | $U_{xx}$     | 98.52%                              | 99.25%                              | 99.49%                              |
|            | $U_{yy}$     | 99.94%                              | 98.96%                              | 99.48%                              |
| <b>DNE</b> | $U_{xx}$     | 98.69%                              | 99,16%                              | 99,41%                              |
|            | $U_{yy}$     | 99.54%                              | 98,72%                              | 99,41%                              |

Tabla 24. Porcentaje de error para la segunda parte del vector derivadas temporales para DCE y DNE

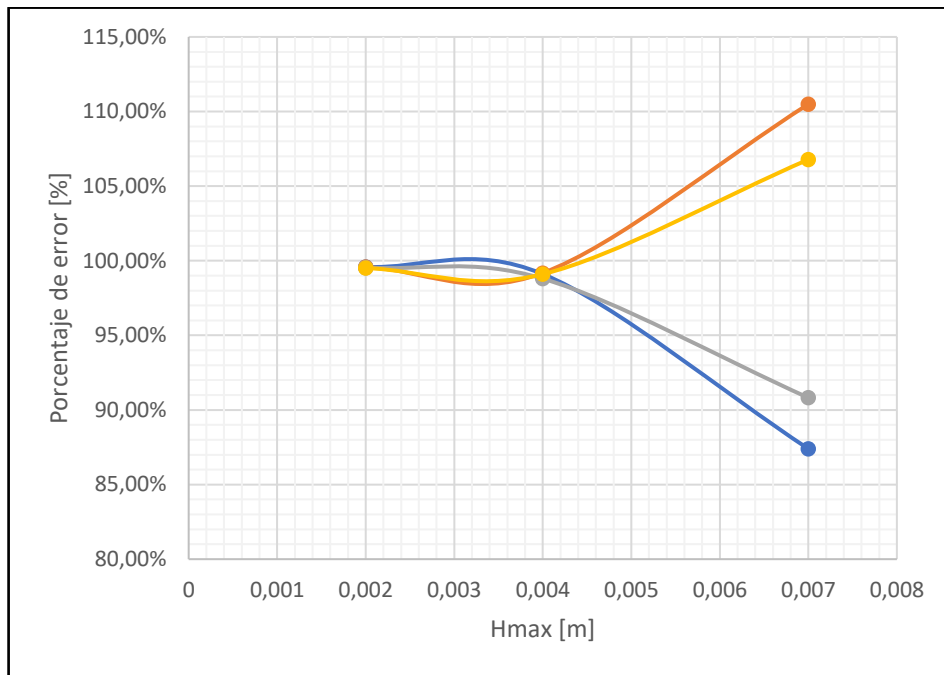
|            | <b>Error</b> | <b><math>h_{max} = 0.007</math></b> | <b><math>h_{max} = 0.004</math></b> | <b><math>h_{max} = 0.002</math></b> |
|------------|--------------|-------------------------------------|-------------------------------------|-------------------------------------|
| <b>DCE</b> | $U_{xx}$     | 87.40%                              | 99.11%                              | 99.59%                              |
|            | $U_{yy}$     | 110.49%                             | 99.18%                              | 99.57%                              |
| <b>DNE</b> | $U_{xx}$     | 90.82%                              | 98.80%                              | 99,54%                              |
|            | $U_{yy}$     | 106.78%                             | 99,10%                              | 99,51%                              |

Figura 30. Grafica del porcentaje de error para la primera parte del vector derivadas temporales para DCE y DNE.



La Figura 30 muestra el comportamiento del porcentaje de error a medida que se aumenta el número de nodos, disminuye el tamaño de los triangulo elementos de la malla, para la primera parte del vector  $\dot{T}$ , en color azul el coeficiente que pertenece a  $\partial^2 T / \partial x^2$  y en color naranja el coeficiente que pertenece a  $\partial^2 T / \partial y^2$ , al aplicar diferencias centrales para puntos equidistantes, en color gris el coeficiente que pertenece a  $\partial^2 T / \partial x^2$  y en color amarillo el coeficiente que pertenece a  $\partial^2 T / \partial y^2$ , al aplicar diferencias centrales para puntos no equidistantes se evidencia que el error aumenta, con la disminución del tamaño de los elementos triangulares, tanto los coeficientes que corresponden a  $\partial^2 T / \partial x^2$  como  $\partial^2 T / \partial y^2$  para las diferencias entre puntos equidistantes como no equidistantes, tienen el mismo comportamiento respecto al error, aunque con el método de diferencias para puntos no equidistantes baja el error, pero no de manera significativa.

Figura 31. Grafica del porcentaje de error para la segunda parte del vector derivadas temporales para DCE y DNE



La Figura 31 muestra el comportamiento del porcentaje de error a medida que se aumenta el número de nodos, disminuye el tamaño de los triangulo elementos de la malla, para la segunda parte del vector  $\vec{T}$ , en color azul el coeficiente que pertenece a  $\partial^2 T / \partial x^2$  y en color naranja el coeficiente que pertenece a  $\partial^2 T / \partial y^2$ , al aplicar diferencias centrales para puntos equidistantes, en color gris el coeficiente que pertenece a  $\partial^2 T / \partial x^2$  y en color amarillo el coeficiente que pertenece a  $\partial^2 T / \partial y^2$ , al aplicar diferencias centrales para puntos no equidistantes se evidencia que el error aumenta, con la disminución del tamaño de los elementos triangulares, tanto los coeficientes que corresponden a  $\partial^2 T / \partial x^2$  como  $\partial^2 T / \partial y^2$  para las diferencias entre puntos equidistantes como no equidistantes, tienen el mismo comportamiento respecto al error, aunque con el método de diferencias para puntos no equidistantes disminuye el error, pero no de manera significativa.

El porcentaje de error sigue siendo muy alto y persiste el aumento del error al aumentar el número de datos, esto no concuerda con que, al hacer la diferenciación entre elementos muy cercanos, se mejora el resultado.

### 4.3 ANALISIS DE LOS DATOS

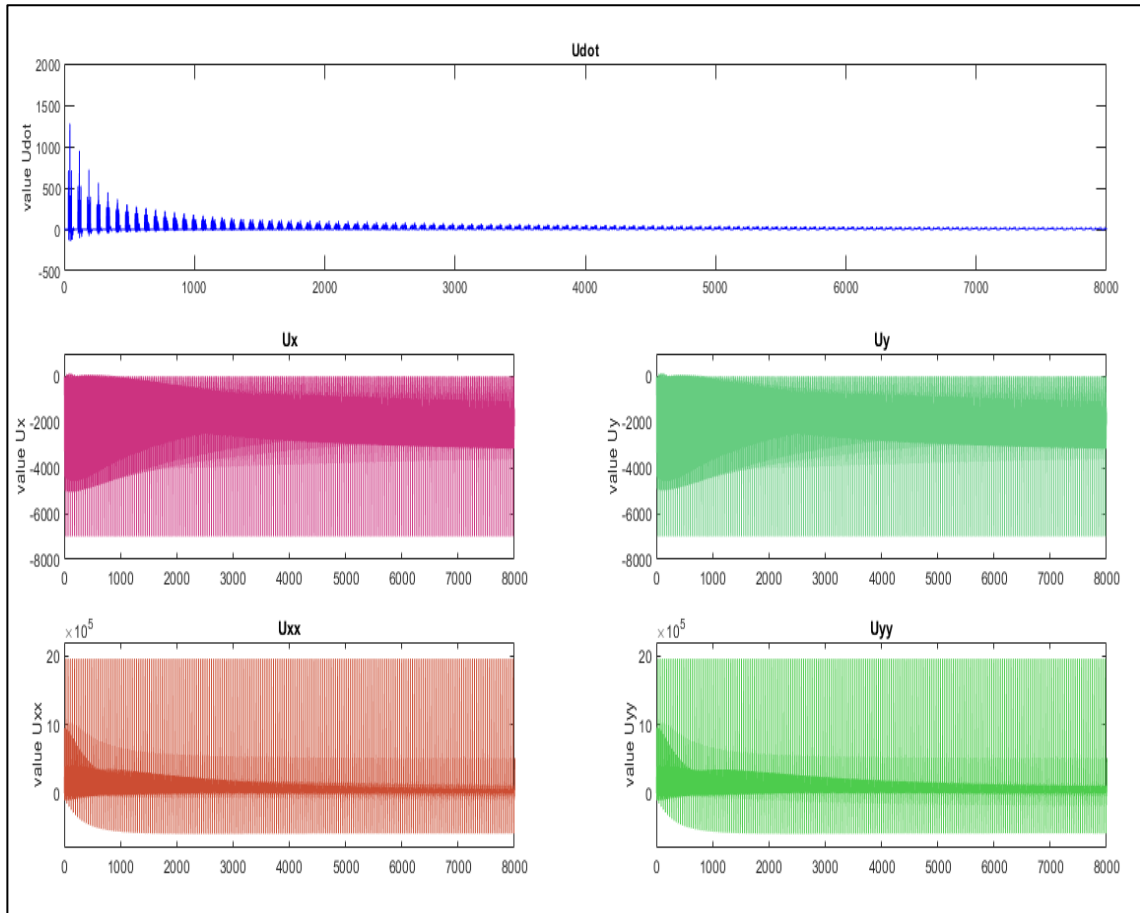
Se va a analizar la data generada por el algoritmo Model Discovery, para ver el comportamiento de esta, porque hasta ahora el porcentaje de error es muy alto, para identificar las posibles causas y tratar de bajar el error considerablemente.

Para esto se va a tomar una data no tan extensa, la cual permite darle un mejor manejo, la data que se va a analizar es la correspondiente a la Figura 24. Malla de tamaño máximo de elemento 0.007m, la cual tiene las siguientes características, una matriz de posiciones  $p = [2 \times 41]$ , una matriz de temperatura  $T = [41 \times 1001]$ , guardada como *Thmax2.mat*, al aplicar el algoritmo *Model Discovery* a esta data, las matrices unificadas de las derivadas respecto del espacio  $T_X = [18 \times 1001]$ ,  $T_{XX} = [18 \times 1001]$ ,  $T_Y = [17 \times 1001]$ ,  $T_{YY} = [17 \times 1001]$ , las cuales se calculan con el método de diferenciación para puntos no equidistantes a partir del polinomio de Lagrange, visto en las ecuaciones (27) y (28), la derivada de los datos respecto del tiempo  $Tdot = [41 \times 999]$ , que se calcula con primera ecuación de la Tabla 6. Fórmulas de diferencias centrales de segundo orden.

#### 4.3.1 Graficas de los vectores

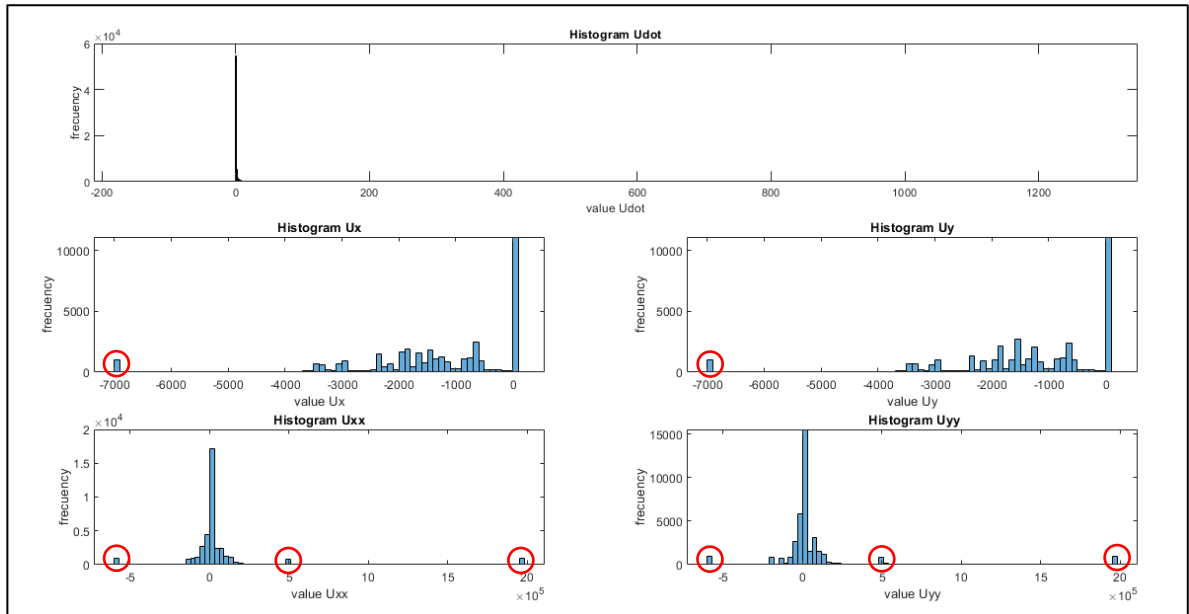
Después de obtener todas las derivadas de la data, tanto espaciales como la derivada dependiente del tiempo, toda esta data se almacena en vectores muy largos, se va a graficar estos vectores sin deshacerse de ningún dato, para ver el comportamiento de cada uno.

Figura 32. Grafica de los vectores resultantes



La Figura 32 muestra el comportamiento de los vectores, como se puede observar, la gráfica que tiene un comportamiento más claro es la de las derivadas temporales, las otras gráficas, las de las derivadas espaciales, parecen tener mucho ruido, porque en un color más claro hay datos en un rango de valores, comportándose de manera cíclica, y de un tono más oscuro, hay un área donde parecen estar concentrados la mayoría de los datos, para saber esto hay que recurrir a una gráfica de frecuencias, como un histograma.

Figura 33. Histograma de los vectores



La Figura 33 muestra la frecuencia con la que aparece un valor en la data, la gráfica superior es el histograma de el vector de derivadas espaciales, presenta valores por fuera de la mayor concentración de los datos por eso aparece en la gráfica valores cercanos a 1500, el factor común de todos los histogramas, es que el dato que más veces se repite, esto se debe a que tomamos para el ejemplo un arreglo que no tiene muchos nodos y por ende gran parte de estos están concentrados en los bordes de la geometría donde están las condiciones de frontera que no cambian en el tiempo, y espacialmente tampoco, para este caso de estudio, en todos los bordes, la temperatura va a ser constante. Para las primeras derivadas espaciales, como para las segundas, aparecen datos que están por fuera de la mayor concentración, lo curioso de estos datos, se repiten considerablemente, ya que, se alcanzan a ver en la gráfica los cuales están señalados con un círculo rojo, se pueden considerar como *outliers* (valores atípicos), hay que determinar porque aparecen, esto porque pueden estar generando el error y tal vez es correcto eliminarlos.

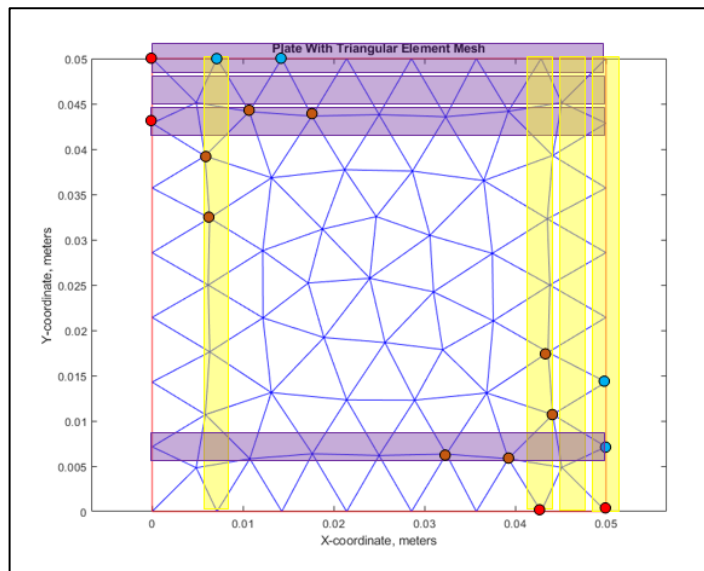
### 4.3.2 ¿De dónde provienen los outliers (valores atípicos)?

Esta es una pregunta la cual se debe responder, para saber cómo se procede con esos datos atípicos, si se pueden dejar en el estudio o si es necesario sacarlos. Para esto se va a ir a cada uno de los vectores, se busca el valor del dato señalado en la Figura 33 y se determina la procedencia, si puede ser un error de cálculo u cualquier otro tipo de error.

Como se necesita saber el porqué del valor de esos datos, la búsqueda se hace de forma manual, ya que no son muchos datos, la forma más sencilla de hacerlo es yendo hasta las celdas donde se almacenaron las derivadas espaciales fila por fila y columna por columna, buscar en cada una de las ellas el valor, se va a buscar el dato de las primeras y segundas derivadas.

El dato busca en las celdas de almacenamiento llamadas  $TTX$ ,  $TTY$ ,  $TTXX$  y  $TTY Y$ , se encuentra en la posición quince de las celdas y para las celdas,  $TTXX$  y  $TTY Y$  las cuales son el objeto de estudio, aparecen las celas de posición tres y trece, es decir, que son las mismas posiciones de búsqueda, se revisa que nodos de la malla son los que pueden generar este inconveniente y si es apropiado eliminar estos datos. Se mira que pasa en la siguiente imagen.

Figura 34. Nodos causantes de los valores atípicos



De la Figura 34 se puede concluir que los valores atípicos son generados por la diferenciación de tres nodos, donde uno de estos corresponde a un valor frontera, especialmente para la fila y columna de nodos, donde se encuentra la intersección de las temperaturas frontera de  $100^{\circ}\text{C}$  y  $0^{\circ}\text{C}$ , como ejemplo, los nodos de color rojo corresponden a temperaturas de  $100^{\circ}\text{C}$  y los de color azul a temperaturas de  $0^{\circ}\text{C}$ , al derivar, un nodo va a tener el máximo de temperatura y los otros dos el mínimo, por eso se generan esos valores tan grandes y lejanos de la mayoría. Los otros datos que generan error son los nodos de color café, de la zona cercana a las intersecciones de las fronteras, se comportan muy parecido a los empalmes de las temperaturas frontera, ya que son los primeros nodos en llegar al estado estable y tienden a generar datos constantes alejados de la mayoría, al ser derivados, esto sucede porque se toma un nodo que se encuentra en la condición de frontera. El error se puede deber a que las condiciones de frontera restringen el fenómeno de difusividad y se tienen muchas isotermas, al aumentar el número de nodos es posible que se aumente el número de isotermas.

### 4.3.3 ¿Qué hacer con los valores atípicos?

La opción de ignorarlos a veces no es la mejor de las soluciones posibles si pretendemos conocer qué ha pasado con estos valores. Lo bien cierto es que distorsionan los resultados del análisis, por lo que hay que identificarlos y tratarlos de forma adecuada. A veces se excluyen si son resultado de un error, pero otras veces son datos potencialmente interesantes en la detección de anomalías. Los valores atípicos pueden deberse a errores en la recolección de datos válidos que muestran un comportamiento diferente, pero reflejan la aleatoriedad de la variable en estudio. Es decir, valores que pueden haber aparecido como parte del proceso, aunque parezcan extraños. Si los valores atípicos son parte del proceso, deben conservarse. En cambio, si ocurren por algún tipo de error (medida, codificación...), lo adecuado es su eliminación. En la siguiente tabla se recogen algunas de las causas comunes de los valores atípicos y sus acciones posibles.<sup>26</sup>

---

<sup>26</sup> YEPES P, Víctor. ¿Qué hacemos con los valores atípicos? En: Universidad politécnica de Valencia. [sitio web]. Valencia: Poliblogs el blog de Víctor Yepes [consulta 30 de diciembre 2022]. Disponible en: <https://victoryepes.blogs.upv.es/2022/02/21/que-hacemos-con-los-valores-atipicos-outliers>

Tabla 25. Causas comunes de los valores atípicos

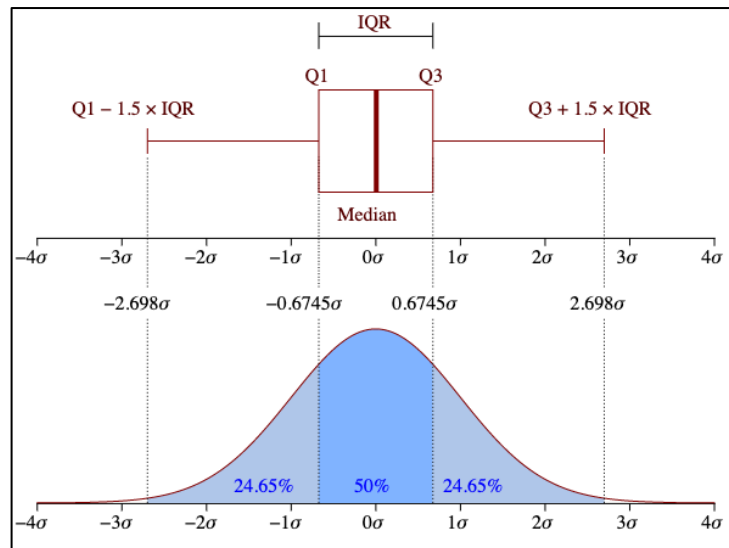
| <b>Causa</b>              | <b>Acciones posibles</b>  |
|---------------------------|---|
| Error de entrada de datos | Corregir el error y volver a analizar los datos.  |
| Problema de proceso       | Investigar el proceso para determinar la causa del valor atípico.   |
| Factor faltante           | Determinar si no se consideró un factor que afecta el proceso.  |
| Probabilidad aleatoria    | Investigar el proceso y el valor atípico para determinar si se produjo en virtud de las probabilidades; realice el análisis con y sin el valor atípico para ver su impacto en los resultados. |

Fuente: soporte de Minitab(R) 18

El método más habitual por su sencillez y resultados es la prueba de Tukey, que toma como referencia la diferencia entre el primer cuartil (Q1) y el tercer cuartil (Q3), o rango intercuartílico. En un diagrama de caja se considera un valor atípico el que se encuentra 1,5 veces esa distancia de uno de esos cuartiles (atípico leve) o a tres veces esa distancia (atípico extremo). Se trata de un método paramétrico que supone que la población es normal. No obstante, también existen métodos no paramétricos cuando la muestra no supere la prueba de normalidad correspondiente.<sup>27</sup>

<sup>27</sup> YEPES P, Víctor. ¿Qué hacemos con los valores atípicos? En: Universidad politécnica de Valencia. [sitio web]. Valencia: Poliblogs el blog de Víctor Yepes [consulta 30 de diciembre 2022]. Disponible en: <https://victoryepes.blogs.upv.es/2022/02/21/que-hacemos-con-los-valores-atipicos-outliers>

Figura 35. Detección paramétrica de valores atípicos, basados en curva de distribución normal.



Fuente: El blog de Víctor Yepes

#### 4.3.4 Eliminación de los datos atípicos (*ouliers*)

Los datos atípicos se eliminan porque son generados por un error en cálculo de las derivadas espaciales, los nodos que se encuentran condiciones de frontera del arreglo influyen mucho, en otras palabras, impiden el buen desarrollo del fenómeno, por ende, los resultados en las derivadas que sean iguales a cero se eliminan, esto con el fin de quitar la influencia de las condiciones de frontera en el fenómeno. La eliminación de los datos se hace mediante el siguiente código.

```
nuxx=uxx(uxx>-4*10^5 & uxx<4*10^5 & uxx~=0);
nyyy=uyy(uyy>-4*10^5 & uyy<4*10^5 & uyy~=0);
nudot=udot(udot~=0);
```

Todo el código para la eliminación de los valores atípicos se puede ver en el **¡Error! No se encuentra el origen de la referencia..**

### 4.3.5 Resultados.

Los resultados del estudio eliminando los datos atípicos de la data generada por el algoritmo *Model Discovery*, usando pseudoinversa y Lasso para determinar los parámetros de  $u_{xx}$  y de  $U_{yy}$  la data correspondiente. Los resultados se muestran en la siguiente tabla

Tabla 26. Resultados quitando los *Outliers* de la data

|                    | <b>Parámetro</b> | <b>Coficiente</b>       | <b>Error</b> |
|--------------------|------------------|-------------------------|--------------|
| Pesudoinversa      | $U_{xx}$         | $8.8706 \times 10^{-5}$ | 24.18%       |
|                    | $U_{yy}$         | $5.8672 \times 10^{-5}$ | 49.85%       |
| Lasso $\alpha = 0$ | $U_{xx}$         | $6.5001 \times 10^{-5}$ | 44.44%       |
|                    | $U_{yy}$         | $4.2665 \times 10^{-5}$ | 63.53%       |
| Lasso $\alpha = 1$ | $U_{xx}$         | $5.2359 \times 10^{-5}$ | 55.25%       |
|                    | $U_{yy}$         | $3.2427 \times 10^{-5}$ | 55.25%       |

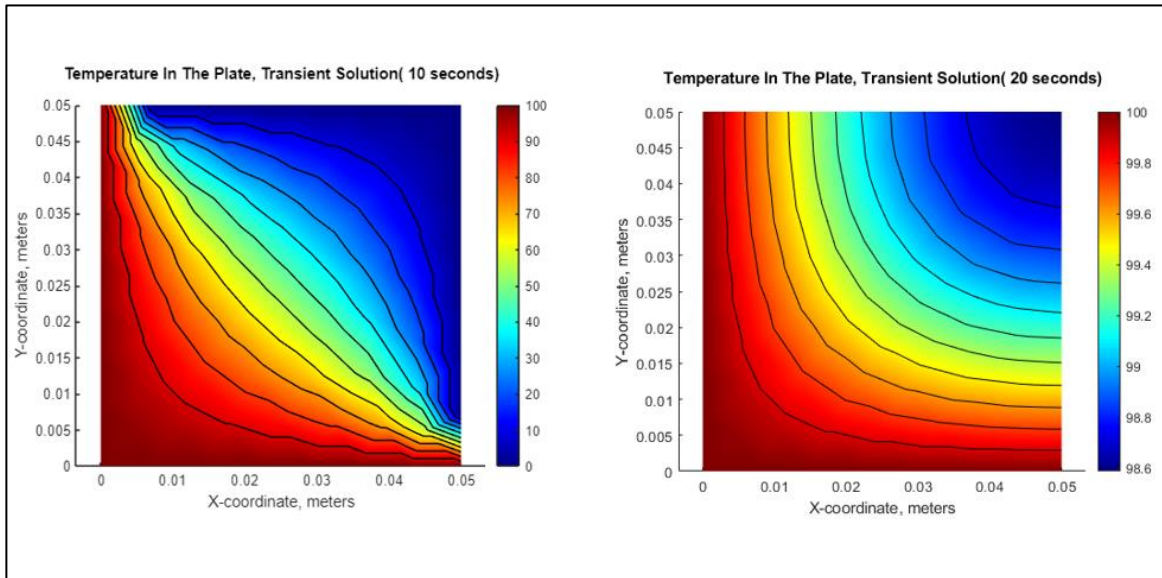
En la tabla anterior se evidencia que, al quitar los datos atípicos, de los vectores de derivadas, ya sea, respecto al tiempo o al espacio, usando la Pesudoinversa ( $\backslash$ ) para obtener los parámetros, o con el uso de la función Lasso, la reducción del error es muy notoria. Se puede observar que el error con la Pesudoinversa es menor que con la función Lasso.

Al ver la disminución del porcentaje de error de una manera tan drástica, paso de estar por encima del 90 % ,a estar en valores por debajo del 50 % con el uso de la pseudoinversa, que hasta ahora es más efectiva, solo con quitar unos datos atípicos, en su mayoría afectados por las condiciones de frontera.

#### 4.4 Pruebas finales.

Para las pruebas finales se va a comparar los resultados de los parámetros de aplicar el *Model Discovery* a los datos llamados *Thmax21.mat* en el Anexo E. Datos de las simulaciones, cuyas características son, una placa plana de cobre cuadrada de  $5 \times 5 \text{ cm}$  de ancho por alto, muy delgada, con coeficiente de difusividad térmica  $\alpha = 1,17 \times 10^{-4} \text{ m/s}^2$  con las siguientes condiciones de frontera en el borde izquierdo  $T_1 = 100 \text{ }^\circ\text{C}$ , y el borde inferior  $T_4 = 100 \text{ }^\circ\text{C}$ , con temperatura inicial  $T_0 = 10 \text{ }^\circ\text{C}$ , para un vector de tiempo  $t = 0:0.02:20\text{s}$ , con los resultados obtenidos en la Tabla 18. Resultados Hmax 0.007m, Pseudoinversa y DC no equidistantes y Tabla 26. Resultados quitando los *Outliers* de la data. Se va a evaluar el resultado de la Pseudoinversa, aplicando diferencias para puntos no equidistantes, sin y con la aplicación de la prueba de Tukey.

Figura 36. Distribución de temperatura de las placas



El código para aplicar la prueba de Tukey es el siguiente.

```

IQRudot=iqr(udot)
IQRux=iqr(ux)
IQRuy=iqr(uy)
IQRuxx=iqr(uxx)
IQRuyy=iqr(uyy)

lowrudot=prctile(udot,25)-1.5*IQRudot
lowrux=prctile(ux,25)-1.5*IQRux
lowruy=prctile(uy,25)-1.5*IQRuy
lowruxx=prctile(uxx,25)-1.5*IQRuxx
lowruyy=prctile(uyy,25)-1.5*IQRuyy

highrudot=prctile(udot,75)+1.5*IQRudot
highrux=prctile(ux,75)+1.5*IQRux
highruy=prctile(uy,75)+1.5*IQRuy
highruxx=prctile(uxx,75)+1.5*IQRuxx
highruyy=prctile(uyy,75)+1.5*IQRuyy

```

Los resultados se muestran en las siguientes tablas

Tabla 27. Resultado para los dos arreglos sin test de Tukey

| <b>Parámetro</b> | <b>Cuatro condiciones F</b> | <b>Dos condiciones F</b> |
|------------------|-----------------------------|--------------------------|
| $U_{xx}$         | $1.5380E^{-6}$              | $8.4356E^{-5}$           |
| $U_{yy}$         | $5.3361E^{-7}$              | $3.7966E^{-5}$           |

Tabla 28. Resultados para los dos arreglos aplicando test de Tukey

| <b>Parámetro</b> | <b>Cuatro condiciones F</b> | <b>Dos condiciones F</b> |
|------------------|-----------------------------|--------------------------|
| $U_{xx}$         | $4.2029E^{-5}$              | $9.3454E^{-5}$           |
| $U_{yy}$         | $3.1614E^{-5}$              | $7.3478E^{-5}$           |

Tabla 29. Porcentaje de error para los dos arreglos sin test de Tukey

| <b>Error</b>    | <b>Cuatro condiciones F</b> | <b>Dos condiciones F</b> |
|-----------------|-----------------------------|--------------------------|
| U <sub>xx</sub> | 98,69 %                     | 27,90%                   |
| U <sub>yy</sub> | 99.54%                      | 67,55%                   |

Tabla 30. Porcentaje de error de los dos arreglos aplicando test de Tukey

| <b>Error</b>    | <b>Cuatro condiciones F</b> | <b>Dos condiciones F</b> |
|-----------------|-----------------------------|--------------------------|
| U <sub>xx</sub> | 64,08 %                     | 20,12%                   |
| U <sub>yy</sub> | 72,98%                      | 37,20%                   |

Las tablas anteriores muestran que al quitar condiciones de frontera se baja el error, sin necesidad de eliminar datos atípicos, al aplicar la prueba de Tukey también se baja el error, quita datos dentro de una distribución normal, los cuales están por fuera de una porción de la distancia entre el primer cuartil y el segundo cuartil, pero no iguala cuando se buscan los datos que son causados por un error y se quitan.

Lo que está claro, es que las condiciones de frontera influyen bastante en la causa de error al aplicar el *Model Discovery*, de pasar de un modelo muy restringido, a un modelo con menos condiciones de frontera, se nota el cambio.

Para ver el código completo donde se quitan los outliers, porque son errores y no se consideran o usando la prueba de Tukey se va al Anexo F. Script live del Model Discovery eliminando outliers de la data.

## 5. CONCLUSIONES

- Se obtuvo la data a partir de una simulación por elementos finitos, de una placa plana de cobre, con condiciones de frontera tipo Dirichlet, en base a la utilización de la herramienta *PDE Toolbox* de Matlab.
- Se desarrolló un algoritmo base para la implementación de la regresión lineal *Model Discovery* para ecuaciones diferenciales en derivadas parciales.
- Se mejoró el algoritmo y se optimizó para que al cambiar la data no se necesite volver a generar un código.
- Se creó una función para ordenar matrices, ya que, una que provee Matlab, generó errores al almacenar datos en arreglos tipo celda.
- Se calculó el porcentaje de error comparando el dato de difusividad térmica que se usó para la simulación y el obtenido por el algoritmo de regresión lineal *Model Discovery*.
- Se determinó como el error aumentó, al incrementar el número de nodos a estudiar, para una data con las mismas que viene de una placa cuadrada plana, con condiciones de frontera en todos sus bordes.
- Se concluyó que el uso de diferencias centrales para datos equidistantes genera más error, que las diferencias para datos no equidistantes, en los resultados, porque la data no es equidistante entre sí.
- Se revisó la data, a través de gráficas temporales y de frecuencia, para determinar valores atípicos que generaban error.

- Se eliminaron los valores atípicos que se generaron por los cálculos de las derivadas, en puntos que estaban asociados a las condiciones de frontera y disminuyo de manera significativa el error.
- Se disminuyeron las condiciones de frontera del arreglo restringido y se comparó con el mismo arreglo, pero con menos condiciones de frontera lo cual bajo el error.
- Se eliminaron valores atípicos usando la prueba de Tukey, la cual se basa en una distribución normal, esta prueba bajo el error, pero no alcanzó a bajar como el hecho buscar el porqué del error de los datos atípicos y luego eliminarlos.
- Se concluyó que el error está muy asociado a las condiciones de frontera si están son muy restrictivas y no dejan desarrollar el fenómeno de difusividad térmica el error va a ser muy alto.
- Se determinó que la diferencia de error entre los coeficientes que se determinan se debe a que los datos son no equidistantes y un vector contiene mejor información que el otro.

## 6. RECOMENDACIONES

Para la aplicación adecuada del modelo de regresión lineal (*Model Discovery*) y poder implementarlo más adelante en un aparato real, como lo es una cámara termográfica, se debe estudiar en base a la data que se genera a partir de las derivadas espaciales, ya que esta genera valores atípicos, por las condiciones de frontera que se usaron para la obtención de la data, se debe revisar a fondo, si es necesario revisar cada caso o si con la implementación de un buen análisis estadístico se pueden resolver todos los inconvenientes que generan los datos atípicos.

Aplicar el método a otro tipo de condiciones de frontera para ver cómo se comporta a estas, aunque las condiciones de frontera son el medio para que el fenómeno ocurra y se pueda obtener la data.

## BIBLIOGRAFÍA

BRUNTON, Steven L.; PROCTOR, Joshua L.; y KUTZ, J. Nathan. [en línea]. Revista PNAS. Washington D.C, noviembre 5 de 2019: vol. 116, nro.45. [consultado 28 noviembre 2020]. ISSN 1091-6490. Disponible en: <https://www.pnas.org/content/pnas/116/45/22445.full.pdf>.

ÇENGEL, Yunus A; BOLES, Michael A. Introducción y conceptos básicos. En: Termodinámica. 5 ed. México D.F: McGraw-Hill Interamericana editores,2006. p. 178

CHAPRA, Steven C; CANALE, Raymond P. Diferenciación e integración numéricas En: Métodos numéricos para ingenieros. México D. F: McGraw Hill interamericana editores S.A. DE C.V. Séptima edición. 2015. p.515

INCROPERA, Frank P, et al. Introduction to conduction. En: Fundamentals of heat and mass transfer. 6 ed. New York: John Wiley & Sons, 2007. p. 60

KUTZ, J. Nathan. Data-driven modeling & scientific computation. Oxford: Oxford university press,2013. P.77

MathWorks. Partial Differential Equation Toolbox: User's Guide [en línea] r2021b. [Consultado 3 noviembre 2022]. Disponible en [https://la.mathworks.com/help/releases/R2021b/pdf\\_doc/pde/index.html](https://la.mathworks.com/help/releases/R2021b/pdf_doc/pde/index.html)

MathWorks. [sitio web]. Descripción de relaciones matemáticas y realización de predicciones a partir de datos experimentales; ¿Qué es la regresión lineal? [Consultado 14 de diciembre 2022]. disponible en: <https://la.mathworks.com/discovery/linear-regression.html>

MathWorks. [sitio web]. Lasso and Elastic Net; ¿what are Lasso an Elastic Net? [Consultado 16 de diciembre 2022]. [https://la.mathworks.com/help/stats/lasso-and-elastic-net.html#bs9ub\\_i](https://la.mathworks.com/help/stats/lasso-and-elastic-net.html#bs9ub_i)

OGATA, Katsuhiko. Modelado matemático de sistemas de control En: Ingeniería de control moderna. 5 ed. Madrid: Pearson Educación S.A., 2010, p. 13-60

Pérez Porto, J, Merino, M. (15 de julio de 2011). Definición de simulación - Qué es, Significado y Concepto. Definición de. [en línea] [Consultado 10 diciembre 2022]. Disponible en <https://definicion.de/simulacion/>

SHENG, Zhang; GUANG, Lin. [en línea]. revista journal of Computational Physics., marzo 1 de 2021: vol. 428, [consultado 20 marzo 2021]. ISSN 0021-9991. Disponible en: <https://bibliotecavirtual.uis.edu.co:2191/science/article/pii/S0021999120307361>

UNIVERSITY OF WASHINGTON; Inferring structure of complex system; Model Discovery [sitio web]. Washington D. C.; Kutz, J. Nathan. [consulta 30 noviembre

2022]. Disponible en:  
<http://faculty.washington.edu/kutz/am563/page1/page5/am563.html>

YEPES P, Víctor. ¿Qué hacemos con los valores atípicos? En: Universidad  
politécnica de Valencia. [sitio web]. Valencia: Poliblogs el blog de Víctor Yepes  
[consulta 30 de diciembre 2022]. Disponible en:  
[https://victoryepes.blogs.upv.es/2022/02/21/que-hacemos-con-los-valores-atipicos-  
outliers](https://victoryepes.blogs.upv.es/2022/02/21/que-hacemos-con-los-valores-atipicos-outliers)

ZILL, Dennis G; WRIGHT, Warren S. Definiciones y terminología En: Ecuaciones  
diferenciales con problemas con valores en la frontera. 8 ed. México D.F: Cengage  
Learning Editores S.A., 2015, p.2-12