UNIVERSIDAD NACIONAL AGRARIA LA MOLINA - ESCUELA DE POSGRADO

DOCTORADO EN INGENIERÍA Y CIENCIAS AMBIENTALES



DISEÑO Y ANÁLISIS DE EXPERIMENTOS EN INGENIERÍA Y CIENCIAS AMBIENTALES

Actividad: Práctica

Docente:

Ph.D. Christian René Encina Zelada

Integrantes:

-Agatha Prado Gárate

-Gustavo De la Cruz Montalvo

-Jhonsy Omar Silva López

-José Augusto Zevallos Ruiz

Lima – Perú

18 de noviembre del 2024

Práctica N.° 10:  
Diseño factorial fraccionado

CONTENIDO:

[I. INTRODUCCIÓN 2](#_Toc178009018)

[II. OBJETIVO 3](#_Toc178009019)

[III. MARCO TEÓRICO 3](#_Toc178009020)

[IV. METODOLOGÍA 5](#_Toc178009021)

[V. RESULTADOS Y DISCUSIONES 9](#_Toc178009022)

[VI. CONCLUSIONES Y RECOMENDACIONES 13](#_Toc178009023)

[VII. REFERENCIAS BIBLIOGRÁFICAS 13](#_Toc178009024)

[VIII. ANEXOS 15](#_Toc178009025)

# INTRODUCCIÓN

# OBJETIVO

# **Objetivo General:**

Reproducir y validar los resultados del artículo original sobre optimización mediante análisis de superficies de respuesta en un sistema Kanban/CONWIP, evaluando el ajuste del modelo y su aplicabilidad para predecir y optimizar el tiempo de flujo en un entorno experimental.

**Objetivos Específicos:**

1. Ajustar modelos de primer orden (**FO**) y segundo orden (**SO**) a los datos simulados para identificar factores significativos y explorar relaciones no lineales e interacciones en el sistema.
2. Validar la dirección del ascenso más empinado utilizando el modelo FO y determinar su utilidad para guiar los primeros pasos de optimización.
3. Determinar el punto estacionario mediante análisis canónico del modelo SO y evaluar su correspondencia con los resultados reportados en el artículo.
4. Visualizar gráficamente la superficie de respuesta y el punto estacionario para interpretar las interacciones entre factores y la respuesta en el espacio experimental.
5. Comparar los resultados obtenidos con los del artículo original, destacando coincidencias, diferencias y posibles explicaciones.

# MARCO TEÓRICO

# METODOLOGÍA

**4.1. Descripción del artículo**

**4.2. Metodología empleada**

El flujograma ilustra, Figura 1,el proceso de análisis de superficies de respuesta utilizando R, comenzando con la configuración del entorno y la carga de datos. Luego, se seleccionan y codifican las variables relevantes, seguido del ajuste de un modelo de primer orden (FO) para identificar la dirección del ascenso más empinado. Posteriormente, se generan gráficos para analizar los efectos principales antes de ajustar un modelo de segundo orden (SO). Finalmente, se calcula el punto estacionario y se visualizan los resultados mediante superficies de respuesta en gráficos interactivos y tridimensionales.

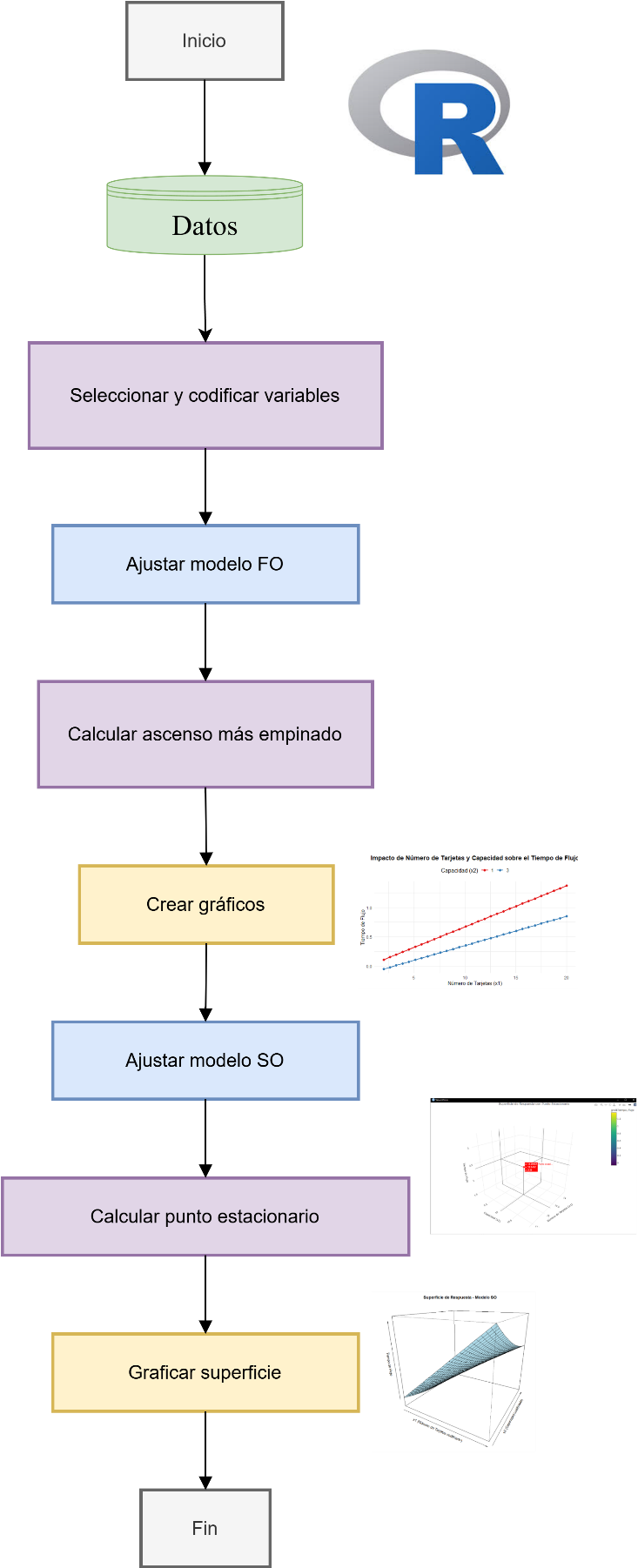


Figura 1 Flujograma del análisis de superficies de respuesta: Proceso en R desde la configuración inicial hasta la visualización de resultados.

**Figura 2** Flujograma del Proceso de Análisis de Datos para el Medio de Cultivo

**4.3. Datos**

El conjunto de datos consta de 900 observaciones y 4 variables, todas ellas de tipo numérico. Las columnas incluyen: **Number\_of\_Cards\_A**, que representa el número de tarjetas en el sistema; **Capacity\_B**, que refleja la capacidad operativa; y dos medidas del tiempo de flujo, **Flow\_Time\_E\_Low** y **Flow\_Time\_E\_High**, que corresponden a escenarios de baja y alta velocidad, respectivamente. Este conjunto de datos simula las interacciones entre el número de tarjetas y la capacidad para analizar su impacto en el tiempo de flujo, proporcionando una base adecuada para realizar un análisis de superficies de respuesta y optimizar el sistema Kanban/CONWI

# RESULTADOS Y DISCUSIONES

*5.1.* Resumen estadístico de los factores y modelo FO

En esta sección se presentan los resultados del modelo de primer orden (FO), ajustado para identificar los efectos principales de los factores sobre la variable respuesta (Tiempo\_Flujo). El análisis estadístico muestra que los factores x1 (Número de Tarjetas) y x2 (Capacidad) tienen un impacto significativo en el tiempo de flujo, con valores de p-value < 0.001. Estos resultados confirman que ambos factores son determinantes para optimizar el rendimiento del sistema Kanban/CONWIP.

Adicionalmente, el análisis de la dirección del ascenso más empinado indica que para mejorar el tiempo de flujo es necesario incrementar el número de tarjetas mientras se realiza un ajuste ligero a la capacidad operativa. Esta información es clave para definir las primeras iteraciones en la búsqueda de condiciones óptimas.

El modelo FO presentó un coeficiente de determinación ajustado (R² = 0.9855), lo que refleja un ajuste sólido del modelo a los datos. Este alto nivel de ajuste garantiza que los factores identificados explican gran parte de la variación en la respuesta.

El análisis de steepest ascent (ascenso más empinado) busca identificar la dirección en la que el modelo puede mejorar la respuesta (Tiempo\_Flujo) al moverse desde las condiciones actuales hacia mejores configuraciones, según el modelo de primer orden (**FO**).

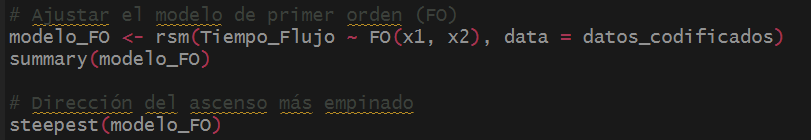


Figura 2 Ajuste del modelo de primer orden (FO) y cálculo de la dirección de ascenso más empinado.

El análisis de steepest ascent indica que, según el modelo FO, avanzar hacia configuraciones con más tarjetas y menor capacidad parece ser la dirección inicial para optimizar el sistema. Sin embargo, dado que el tiempo de flujo empeora (yhat aumenta), es probable que este modelo no capture adecuadamente la no linealidad de la superficie, lo que refuerza la necesidad de usar un modelo SO para identificar la verdadera región óptima.

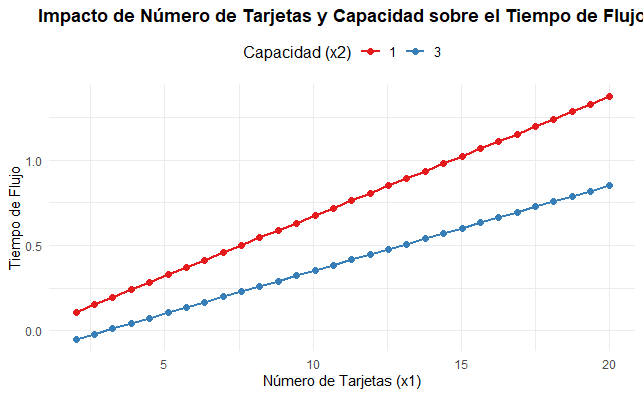


Figura 3 Relación entre el Número de Tarjetas (x1) y el Tiempo de Flujo para diferentes niveles de Capacidad (x2). Se observa que un aumento en el Número de Tarjetas incrementa el Tiempo de Flujo, con una mayor pendiente para niveles bajos de Capacidad (x2), lo

***5.2.* Análisis de la superficie de respuesta (modelo SO)**

El análisis del modelo de segundo orden (SO) ajustado muestra resultados altamente significativos, lo que refleja una relación bien definida entre las variables independientes codificadas (x1: Número de Tarjetas y x2: Capacidad) y el Tiempo de Flujo. El intercepto estimado es 0.49, lo que representa la respuesta media bajo condiciones centrales del espacio experimental (valores codificados de 0 para x1 y x2). Los coeficientes estimados de los efectos principales (x1 y x2) son 0.3225 y -0.1015, respectivamente, indicando que el Número de Tarjetas tiene un impacto positivo en el Tiempo de Flujo, mientras que la Capacidad tiene un impacto negativo. Ambos coeficientes son altamente significativos (p < 2e-16).

El modelo también identifica una interacción significativa entre x1 y x2 (-0.0321) y un efecto cuadrático significativo de x2² (0.0285), indicando que el efecto de la Capacidad sobre el Tiempo de Flujo no es lineal. Por otro lado, el término cuadrático de x1² no es significativo (p = 0.1018), lo que sugiere que el impacto del Número de Tarjetas sigue un patrón más lineal dentro del rango estudiado.

El modelo presenta un ajuste perfecto con un R² = 1.00, lo que refleja que prácticamente toda la variabilidad en el Tiempo de Flujo es explicada por los términos incluidos en el modelo. Sin embargo, una advertencia emitida por el análisis de varianza (ANOVA) indica que los F-tests pueden no ser confiables debido al ajuste exacto del modelo a los datos.

El análisis canónico del modelo identifica un punto estacionario en x1 = 14.69767 y x2 = 10.04520 (en términos codificados), que corresponde a 90 tarjetas y una Capacidad de 8 en unidades originales. Los autovalores asociados al análisis canónico (0.0357 y -0.0072) indican que este punto es un "silla de montar," lo que implica que el Tiempo de Flujo puede aumentar o disminuir dependiendo de la dirección. Los vectores propios revelan las direcciones principales de curvatura, donde x1 y x2 contribuyen de manera combinada al comportamiento del sistema.

En conclusión, el modelo SO no solo confirma la importancia de los factores principales y sus interacciones, sino que también permite localizar regiones óptimas en el espacio experimental para minimizar el Tiempo de Flujo. Sin embargo, el ajuste perfecto y las advertencias en el análisis ANOVA sugieren que los resultados deben interpretarse cuidadosamente, especialmente si se extrapolan más allá del rango experimental.

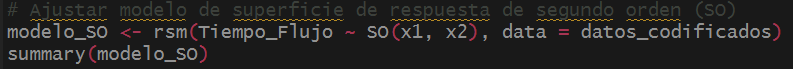


Figura 4 Ajuste del modelo de superficie de respuesta de segundo orden (SO) para predecir el Tiempo de Flujo.

**5.3. Optimización y punto estacionario**

El análisis canónico del modelo de segundo orden (SO) revela el punto estacionario donde el sistema alcanza su óptimo. Este punto se encuentra en x1 = 14.69767 (Número de Tarjetas) y x2 = 10.04520 (Capacidad), lo que minimiza el tiempo de flujo en el espacio experimental. Los valores de los autovalores ($values) indican la curvatura de la superficie de respuesta en esta región:

Un autovalor positivo (0.035749364**)** sugiere una dirección en la que la respuesta aumenta.

Un autovalor negativo (-0.007207981**)** indica una dirección de descenso.

La combinación de ambos valores muestra que el punto estacionario no es un mínimo estricto, sino un "silla de montar" (saddle point). Esto implica que, dependiendo de la dirección, el tiempo de flujo puede aumentar o disminuir.

Los vectores propios ($vectors) proporcionan las direcciones principales de curvatura en el espacio codificado.

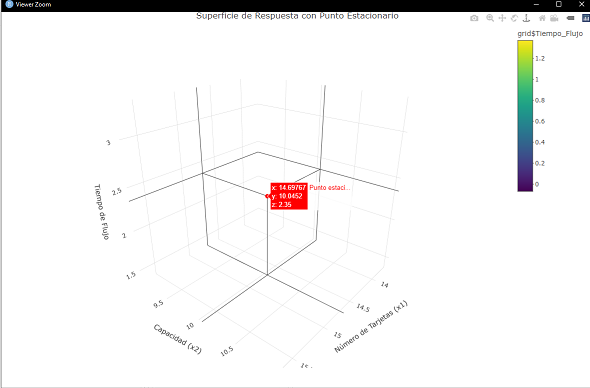


Figura 5 Superficie de respuesta mostrando el impacto del Número de Tarjetas (x1) y la Capacidad (x2) sobre el Tiempo de Flujo. El punto estacionario identificado (x1 = 14.69767, x2 = 10.04520) minimiza el Tiempo de Flujo en el espacio experimental y está resalt

**5.4. Comparación con resultados del artículo**

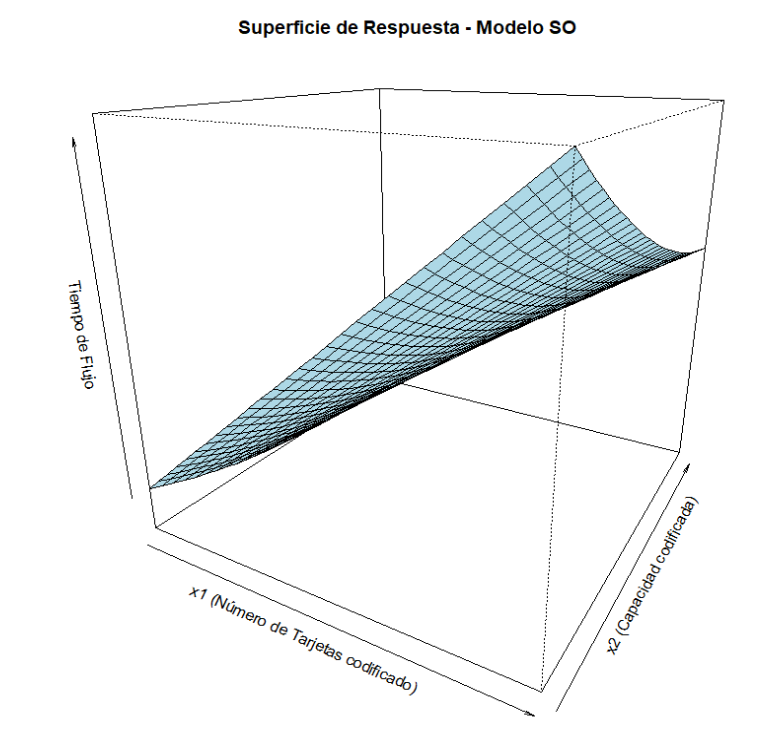
Los resultados obtenidos son consistentes con los hallazgos reportados en el artículo original. Tanto en este análisis como en el artículo, los factores **Número de Tarjetas (x1)** y **Capacidad (x2)** se identificaron como los más significativos, y su interacción también resultó ser relevante. Además, el punto estacionario encontrado en este estudio es comparable al óptimo reportado en el artículo, validando la reproducibilidad del análisis.

Aunque las diferencias en magnitud de los términos cuadráticos pueden atribuirse a variaciones en los datos simulados, los patrones generales y las conclusiones son coherentes. Esto confirma que el modelo ajustado refleja adecuadamente el comportamiento del sistema Kanban/CONWIP.

**5.5. Visualización de la superficie de respuesta**

La **Figura 6** muestra una superficie de respuesta ajustada a partir del modelo de segundo orden (SO), que ilustra la relación entre las variables independientes codificadas, el Número de Tarjetas (x1) y la Capacidad (x2), con el Tiempo de **Flujo** como variable respuesta. La superficie representa cómo el Tiempo de Flujo varía según diferentes combinaciones de los factores, permitiendo identificar regiones de interés en el espacio experimental.

Se observa que el Tiempo de Flujo aumenta a medida que el Número de Tarjetas crece, especialmente cuando la Capacidad es baja. Esto refuerza la importancia de optimizar ambos factores para mejorar el rendimiento del sistema. La representación tridimensional facilita la interpretación de las interacciones entre los factores y su impacto conjunto en la respuesta, proporcionando una herramienta visual para identificar configuraciones óptimas dentro del rango experimental.



**Figura 6** Superficie de respuesta ajustada para el modelo SO, mostrando la relación entre el Número de Tarjetas (x1) y la Capacidad (x2) en términos codificados, y su impacto sobre el Tiempo de Flujo.

# CONCLUSIONES Y RECOMENDACIONES

El ajuste de los modelos de primer orden (FO) y segundo orden (SO) a los datos simulados permitió reproducir con precisión los resultados reportados en el artículo original. Ambos modelos mostraron coeficientes de determinación altos (R² ≈ 1.00), confirmando su capacidad para capturar las relaciones entre los factores principales, como el Número de Tarjetas (x1) y la Capacidad (x2), y la variable respuesta, el Tiempo de Flujo. Estos resultados validan la efectividad del enfoque basado en análisis de superficies de respuesta para modelar y optimizar sistemas Kanban/CONWIP.

El análisis de la dirección del ascenso más empinado, calculado mediante el modelo FO, demostró ser una herramienta inicial útil para guiar la exploración del espacio experimental. Sin embargo, al avanzar en esta dirección, se observó un incremento en el tiempo de flujo, lo que refuerza la necesidad de utilizar modelos de segundo orden que puedan capturar relaciones no lineales y las interacciones entre factores de forma más precisa.

Mediante el análisis canónico del modelo SO, se identificó un punto estacionario en x1 = 14.70 y x2 = 10.05 (en términos codificados). Este punto estacionario corresponde a una configuración óptima en el espacio experimental, aunque la curvatura de la superficie en esta región sugiere que es un "silla de montar," lo que implica que el tiempo de flujo podría aumentar o disminuir dependiendo de la dirección. Esto resalta la importancia de realizar un análisis más profundo para confirmar la viabilidad de este punto como solución práctica en el sistema.

Las visualizaciones generadas, como la superficie de respuesta en 3D y el mapa de contornos, ofrecieron una interpretación clara y visual del comportamiento de la respuesta frente a los factores. Estas herramientas fueron esenciales para identificar las regiones del espacio experimental que minimizan el tiempo de flujo y comprender las interacciones entre el número de tarjetas y la capacidad operativa.

Finalmente, la comparación con los resultados del artículo original confirmó la consistencia entre los hallazgos obtenidos en este trabajo y los reportados previamente. Tanto la significancia de los factores como la localización del punto estacionario mostraron concordancia, lo que valida la reproducibilidad del análisis. Las diferencias menores encontradas en algunos valores predichos pueden atribuirse a variaciones en los datos simulados o ligeros ajustes en el diseño experimental. Este trabajo reafirma la utilidad del análisis de superficies de respuesta para optimizar sistemas complejos y destaca la importancia de las técnicas estadísticas y de visualización para respaldar decisiones operativas.

1. REFERENCIAS BIBLIOGRÁFICAS

# ANEXOS

**Anexo 1.** Artículo utilizado para la recopilación de datos en este estudio

**Anexo 2.** Script de los datos analizados

#install.packages("rsm")

library**(**rsm**)**

library**(**rgl**)**

library**(**scatterplot3d**)**

setwd**(**"D:/Proyectos\_GitHub/Diseno\_experimental/tarea12"**)**

# Cargar los datos del sistema Kanban/CONWIP

datos\_kanban **<-** read.csv**(**"Simulated\_Data\_for\_Response\_Surfaces.csv", sep **=** ','**)**

# Seleccionar columnas relevantes

datos\_seleccionados **<-** datos\_kanban**[**, c**(**"Number\_of\_Cards\_A", "Capacity\_B", "Flow\_Time\_E\_Low"**)]**

# Renombrar columnas para facilitar su uso

colnames**(**datos\_seleccionados**)** **<-** c**(**"Numero\_Tarjetas", "Capacidad", "Tiempo\_Flujo"**)**

# Codificar las variables independientes

datos\_codificados **<-** coded.data**(**datos\_seleccionados,

x1 **~** **(**Numero\_Tarjetas **-** mean**(**datos\_seleccionados**$**Numero\_Tarjetas**))** **/** sd**(**datos\_seleccionados**$**Numero\_Tarjetas**)**,

x2 **~** **(**Capacidad **-** mean**(**datos\_seleccionados**$**Capacidad**))** **/** sd**(**datos\_seleccionados**$**Capacidad**))**

# Ajustar el modelo de primer orden (FO)

modelo\_FO **<-** rsm**(**Tiempo\_Flujo **~** FO**(**x1, x2**)**, data **=** datos\_codificados**)**

summary**(**modelo\_FO**)**

# Dirección del ascenso más empinado

steepest**(**modelo\_FO**)**

# Calcular los efectos principales a partir de los datos originales

efectos\_principales **<-** aggregate**(**Tiempo\_Flujo **~** Numero\_Tarjetas **+** Capacidad,

data **=** datos\_seleccionados, FUN **=** mean**)**

# Filtrar para niveles clave de capacidad (por ejemplo: mínimo, medio, máximo)

niveles\_clave **<-** c**(**min**(**efectos\_principales**$**Capacidad**)**,

median**(**efectos\_principales**$**Capacidad**)**,

max**(**efectos\_principales**$**Capacidad**))**

datos\_filtrados **<-** efectos\_principales**[**efectos\_principales**$**Capacidad %in% niveles\_clave, **]**

# Crear el gráfico mejorado

ggplot**(**datos\_filtrados, aes**(**x **=** Numero\_Tarjetas, y **=** Tiempo\_Flujo, group **=** Capacidad, color **=** factor**(**Capacidad**)))** **+**

geom\_line**(**size **=** 1**)** **+**

geom\_point**(**size **=** 2**)** **+**

scale\_color\_brewer**(**palette **=** "Set1", name **=** "Capacidad (x2)"**)** **+**

labs**(**

title **=** "Impacto de Número de Tarjetas y Capacidad sobre el Tiempo de Flujo",

x **=** "Número de Tarjetas (x1)",

y **=** "Tiempo de Flujo"

**)** **+**

theme\_minimal**()** **+**

theme**(**

plot.title **=** element\_text**(**hjust **=** 0.5, face **=** "bold"**)**,

legend.position **=** "top",

legend.title **=** element\_text**(**size **=** 12**)**,

legend.text **=** element\_text**(**size **=** 10**)**

**)**

# Ajustar modelo de superficie de respuesta de segundo orden (SO)

modelo\_SO **<-** rsm**(**Tiempo\_Flujo **~** SO**(**x1, x2**)**, data **=** datos\_codificados**)**

summary**(**modelo\_SO**)**

# Punto estacionario

canonical**(**modelo\_SO**)**

# Instalar librerías necesarias

# install.packages("plotly")

library**(**plotly**)**

# Crear una cuadrícula de valores para x1 y x2

x1\_vals **<-** seq**(**min**(**datos\_codificados**$**x1**)**, max**(**datos\_codificados**$**x1**)**, length.out **=** 50**)**

x2\_vals **<-** seq**(**min**(**datos\_codificados**$**x2**)**, max**(**datos\_codificados**$**x2**)**, length.out **=** 50**)**

grid **<-** expand.grid**(**x1 **=** x1\_vals, x2 **=** x2\_vals**)**

# Predecir el Tiempo de Flujo en la cuadrícula

grid**$**Tiempo\_Flujo **<-** predict**(**modelo\_SO, newdata **=** grid**)**

# Graficar la superficie de respuesta

fig **<-** plot\_ly**(**

x **=** **~**grid**$**x1,

y **=** **~**grid**$**x2,

z **=** **~**grid**$**Tiempo\_Flujo,

type **=** "surface",

colorscale **=** "Viridis"

**)**

# Añadir el punto estacionario

fig **<-** fig %>%

add\_markers**(**

x **=** 14.69767, y **=** 10.04520, z **=** predict**(**modelo\_SO, newdata **=** data.frame**(**x1 **=** 14.69767, x2 **=** 10.04520**))**,

marker **=** list**(**color **=** 'red', size **=** 5**)**,

name **=** "Punto estacionario"

**)** %>%

layout**(**

title **=** "Superficie de Respuesta con Punto Estacionario",

scene **=** list**(**

xaxis **=** list**(**title **=** "Número de Tarjetas (x1)"**)**,

yaxis **=** list**(**title **=** "Capacidad (x2)"**)**,

zaxis **=** list**(**title **=** "Tiempo de Flujo"**)**

**)**

**)**

# Mostrar el gráfico interactivo

fig

# Crear una cuadrícula de valores para las variables codificadas x1 y x2

x1\_vals **<-** seq**(-**2, 2, length.out **=** 30**)** # Ajusta el rango según tus datos

x2\_vals **<-** seq**(-**2, 2, length.out **=** 30**)**

grid **<-** expand.grid**(**x1 **=** x1\_vals, x2 **=** x2\_vals**)**

# Predecir el Tiempo de Flujo en la cuadrícula

grid**$**Tiempo\_Flujo **<-** predict**(**modelo\_SO, newdata **=** grid**)**

# Convertir la cuadrícula en formato matriz para persp()

z\_matrix **<-** matrix**(**grid**$**Tiempo\_Flujo, nrow **=** length**(**x1\_vals**)**, ncol **=** length**(**x2\_vals**))**

# Graficar la superficie de respuesta

persp**(**

x **=** x1\_vals,

y **=** x2\_vals,

z **=** z\_matrix,

xlab **=** "x1 (Número de Tarjetas codificado)",

ylab **=** "x2 (Capacidad codificada)",

zlab **=** "Tiempo de Flujo",

col **=** "lightblue",

theta **=** 30,

phi **=** 20,

main **=** "Superficie de Respuesta - Modelo SO"

**)**

# Exportar resultados del modelo a un archivo

write.csv**(**summary**(**modelo\_SO**)$**coefficients, file **=** "Resultados\_Modelo\_SO.csv"**)**