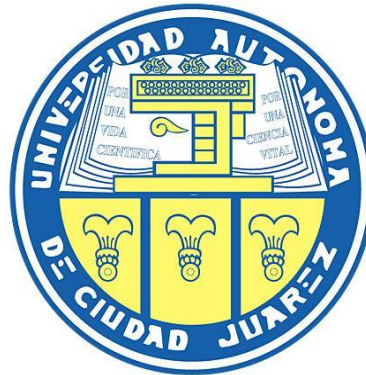


UNIVERSIDAD AUTÓNOMA DE CIUDAD JUÁREZ
Instituto De Ingeniería Y Tecnología
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL Y MANUFACTURA
MAESTRÍA EN INGENIERÍA INDUSTRIAL



**Modelado predictivo para robustecimiento del proceso de
impresión digital en prensas HP Indigo mediante modelos de
regresión y planta virtual**

PROYECTO DE INGENIERÍA QUE PRESENTA:

Ing. Juan Pablo Puentes Castro

**COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE:
MAESTRO(A) EN INGENIERÍA INDUSTRIAL**

DIRECTOR(A): Dr. Rey David Molina Arredondo
CODIRECTOR(A): Dr. Manuel Roman Piña Monarrez

CD. JUÁREZ, CHIH.

MAYO DE 2026

Dedicatoria

A los míos, a mi familia, que me sostiene y me impulsa a superar cada escalón de mi vida profesional y desarrollo académico. A mis seres queridos que partieron de este mundo sin verme alcanzar algunos de mis sueños, pero de quienes no tengo duda que guían y cuidan cada uno de ellos en silencio y con amor.

A mi madre Martha, cuyo esfuerzo ha sido mi primer ejemplo de trabajo y amor incondicional. A mi hermano Juan Manuel, quien me ha enseñado el respeto por la vocación y el valor del trabajo al salvar vidas día a día. A mi hermana Karol, que con su amor y determinación me enseña la pasión con la que se debe vivir la vida. A mi padre Juan Manuel, quien me enseñó el valor del sacrificio por quienes amas.

A mi amor Salma que me dio ejemplo de que la experiencia y la juventud pueden coexistir, y que es posible alcanzar grandes metas a corta edad, motivando profundamente el presente.

A mis maestros a lo largo de mi vida, quienes forjaron la identidad de mi trabajo y de quienes aprendí lecciones valiosas de humanidad, respeto y compromiso.

Al deporte, que me dio disciplina, carácter y me abrió las puertas para cumplir objetivos.

Y finalmente, me lo dedico a mí: a mi voluntad, a mi determinación y a mi trabajo duro.

Resumen

El presente estudio analizó el robustecimiento del proceso de impresión digital en una prensa HP Indigo serie 4 de hoja, centrándose en la evaluación de factores técnicos y ambientales que influyen en el desempeño y comportamiento del sustrato durante la operación. La problemática abordada se relacionó con la alta incidencia de atascos de papel y tiempos muertos, asociados a variaciones en las propiedades del sustrato, condiciones ambientales y falta de control en variables críticas del proceso.

El desarrollo de la investigación se fundamentó en información técnica proveniente de la guía de sustratos de HP, así como en los principios del diseño de experimentos y el diseño robusto. Se empleó análisis estadístico para evaluar la influencia de variables como la temperatura ambiente, la humedad relativa, la precisión del corte del papel y el gramaje del sustrato. Los datos utilizados fueron recopilados durante un periodo aproximado de 90 días de operación real, complementados mediante simulación del factor de corte dentro de un rango de 748 mm a 752 mm, con el objetivo de modelar su impacto sobre la tasa de atascos sin alterar la estructura observada del sistema.

Dentro del análisis, la regresión lineal constituyó la base inicial para identificar efectos principales; posteriormente, se desarrollaron modelos con interacciones y modelos de segundo orden para capturar efectos no lineales y relaciones complejas entre variables. El análisis de varianza (ANOVA) permitió validar la significancia estadística de los factores y comparar la calidad de ajuste entre modelos. Como resultado, se seleccionó un modelo cuadrático optimizado, el cual presentó el mejor desempeño en términos de coeficiente de determinación ajustado, criterio de información de Akaike (AIC) y capacidad predictiva (PRESS y R^2 predictivo).

Los resultados indicaron que las variables con mayor influencia en la generación de atascos fueron la humedad relativa y la precisión del corte del sustrato, así como su interacción. Se observó que pequeñas desviaciones respecto al valor nominal de 750 mm incrementan significativamente la probabilidad de fallas, mientras que la humedad actúa

como un factor de ruido que amplifica la variabilidad del sistema. La temperatura mostró efectos secundarios, principalmente en interacción con otras variables.

A partir del modelo seleccionado, se desarrolló una planta virtual predictiva en entorno R, incorporando simulación tipo Monte Carlo para evaluar la variabilidad del proceso bajo diferentes condiciones operativas. Esta herramienta permite estimar la distribución de producción, pérdidas asociadas a atascos y comportamiento del sistema ante cambios en los factores controlables, facilitando la toma de decisiones sin necesidad de experimentación directa en planta.

Con base en los hallazgos, se propone robustecer el proceso mediante el control estricto de la precisión de corte del sustrato, la implementación de estrategias de control ambiental, y la estandarización de condiciones operativas. Estas acciones permiten reducir la variabilidad del proceso, minimizar tiempos muertos y mejorar la eficiencia operativa del sistema de impresión.

El estudio contribuye a la comprensión integral del comportamiento del sustrato en procesos de impresión digital y demuestra la efectividad de la integración de diseño de experimentos, modelado estadístico y simulación como herramienta para la optimización de procesos industriales en entornos con restricciones operativas.

Palabras clave— Impresión digital, HP Indigo, diseño de experimento, modelo de regresión, simulación Monte Carlo, sustrato, procesos.

Índice Temático

CAPÍTULO I. INTRODUCCIÓN	1
1.1 Antecedentes	3
1.2 Descripción del Problema.....	4
1.3 Objetivos.....	6
1.4. <i>Objetivo General</i>	6
1.4.1 <i>Objetivos Específicos</i>	7
1.4 Hipótesis	7
1.5 Justificación.....	7
1.6 Alcances y limitaciones.....	9
CAPÍTULO II. MARCO TEÓRICO.....	11
2.1 Modelos de regresión	11
2.2 ANOVA.....	14
2.2.1 <i>Estadístico AIC y PRESS</i>	15
2.3 Diseño robusto.	17
2.4 Método Montecarlo.	19
2.5 Plantas virtuales y gemelos digitales	22
2.6 Tecnología de impresión HP Indigo	23
2.7 Comportamiento del sustrato en impresión digital.....	25
CAPÍTULO III. METODOLOGÍA.....	28
3.1 Revisión de la bibliografía.	28
3.2 Estudio del proceso actual.	29
3.3 Selección de sustrato a estudiar	30

3.4	Recolección de datos históricos del sustrato seleccionado.....	31
3.5	Selección y ajuste de modelo predictivo con herramientas de análisis estadístico	31
3.6	Creación de planta virtual para simulación	32
CAPÍTULO IV. RESULTADOS		33
4.1	El proceso actual.	33
4.2	Datos para el experimento.....	37
4.3	Análisis de regresión y selección del modelo predictivo	40
4.3.1	<i>Los supuestos.</i>	45
4.4	Planta Virtual	48
4.4.1	<i>Validación de la planta virtual.</i>	54
CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES.....		60
CAPÍTULO VI. REFERENCIAS.....		63
ANEXOS		65
Tabla 5.1. Datos históricos observados.....		65
Tabla 5.2 Funciones RStudio utilizadas		67
Código 5.1 Comparación de modelos Rstudio.....		69
Código 5.2 Planta virtual Rstudio		73

Índice de Tablas

Tabla 1.1 Mediciones tomadas en el periodo Diciembre 2025- Febrero 2026.	9
Tabla 2.1 Condiciones recomendadas para sustratos y ambiente de impresión	
.....	27
Tabla 3.1 Propiedades físicas del sustrato	31
Tabla 4.2 Anova del modelo seleccionado	44
Tabla 4.3 Comparación final de modelos	45
Tabla 4.4 Coeficientes de los términos del modelo	48

Índice de Figuras

Figura 1.1 Hojas impresas vs atascos.....	6
Figura 2.1 Prensa digital HP Indigo serie 4.....	24
Figura 3.1 Descripción general metodología del proyecto.	28
Figura 3.2 Grafica de pastel de sustratos utilizados en las prensas ...	30
Figura 4.1 Diagrama de Ishikawa del proceso.....	33
Figura 4.2. Sistema de corte CHM 1400.....	34
Figura 4.3. Llegada del material a sitio.....	35
Figura 4.4. Almacenamiento actual.....	36
Figura 4.5. Humedad relativa del sitio.....	39
Figura 4.6. Temperatura del sitio.....	40
Figura 4.7 Residuals vs Fitted.....	46
Figura 4.8 Q-Q Residuals.....	46
Figura 4.9 Scale-Location.....	47
Figura 4.10 Residuals vs Leverage.....	48
Figura 4.11 Interfaz de usuario de la planta virtual.....	51
Figura 4.12 Ejemplo de resultados de simulación en planta virtual.....	53
Figura 4.13 Grafica de barras de ejemplo de simulación.....	54
Figura 4.14 Histograma de ejemplo de simulación.....	54
Figura 4.15 Atascos registrados en Abril-Mayo 2026.....	55
Figura 4.16 Entradas de la simulación.....	56
Figura 4.17 Resultados de la simulación.....	57

CAPÍTULO I. INTRODUCCIÓN

En años recientes la impresión digital se ha posicionado como una tecnología clave en la industria del arte gráfico, no solo por su capacidad de personalizar y producir rápidamente, sino también por su notable evolución en la calidad de impresión y compatibilidad con diversos sustratos. Dentro de los sistemas más avanzados del mercado se encuentra la tecnología de impresión digital desarrollada por la empresa HP. En particular en el presente estudio abordaremos el modelo HP Indigo serie 4, esta prensa propone una solución que combina precisión, velocidad y versatilidad a través de un complejo sistema de impresión basado en la tecnología de electrofotografía líquida (LEP) de HP, la cual utiliza tinta ElectroInk con partículas de apenas 1 a 2 micras (HP Development Company, 2018). Este sistema permite una transferencia por contacto directo que garantiza la calidad de impresión. No obstante, se subraya que, para alcanzar este nivel de excelencia operativa, es imperativo contar con un entorno de trabajo optimizado y una gestión estricta de los sustratos.

Según las recomendaciones del fabricante explicadas en el documento de preparación de sitio para instalación, para alcanzar su máximo potencial se exige un entorno de operación optimizado, así como estrictas políticas de insumos compatibles especialmente en lo que se refiere al material o sustrato en el que se realizara la impresión. (HP,2014)

El papel o sustrato (como lo llamaremos en este estudio) es la base que recibe la impresión y representa mucho más que un soporte visual. Su comportamiento durante el proceso afecta directamente la calidad del producto final, la eficiencia de la prensa y la frecuencia de tiempos muertos debidos a fallas mecánicas. Las propiedades físicas de los sustratos que definiremos como variables son el gramaje, el tipo de acabado, la dirección de las fibras y el contenido de humedad. La combinación de estas variables define si el pliego (hoja) avanzará sin obstáculos a través de la prensa o si causará fallas frecuentes durante la impresión.

De acuerdo con la Guía práctica de sustratos HP Indigo (HP, 2024), la calidad de impresión en prensas de bobina (rollo) o de pliego (hoja) depende directamente de la manipulación correcta y almacenamiento de los sustratos, los cuales deben mantenerse bajo condiciones ambientales estrictamente controladas. Es entonces donde el desafío técnico está relacionado con la naturaleza higroscópica del papel, ya que este busca mantener un equilibrio con las condiciones de su entorno. La guía práctica de sustratos de HP Indigo indica que la mayoría de los papeles son estables en entornos óptimos de una humedad relativa entre 50% y 70%, y una temperatura ambiente de 20°C a 25°C, parámetros que garantizan la estabilidad dimensional y la conservación del tratamiento superficial aplicado en los materiales. Desviaciones en estas condiciones generan fenómenos de curvatura (curl), acumulación de electricidad estática y deformaciones que elevan la frecuencia de atascos de papel, traducándose en tiempos muertos y desperdicio de material.

El presente estudio tiene como objetivo analizar el cómo las variables técnicas y ambientales en los sustratos influyen en el desempeño de la prensa HP Indigo serie 4 de pliego. Se busca, además, fortalecer el proceso para la selección, preparación y control de los sustratos que se emplean en este tipo de impresión, a fin de reducir fallos operativos y optimizar los resultados. Para ello, se propone el análisis de regresión específicamente un diseño factorial, como metodología estadística para identificar interacciones críticas entre factores. El modelo de Taguchi destaca la importancia de evaluar el desempeño bajo condiciones de campo como parte del proceso de diseño y el hecho de que la variación funcional en el desempeño está influenciada por los factores de ruido los cuales varían en el ambiente en el que los procesos o productos están funcionando (Yáñez Hernández et al., 2023). Investigaciones industriales demuestran que el uso de DoE en prensas gráficas puede reducir los tiempos de preparación hasta en un 40% (Stat-Ease, s.f.).

1.1 Antecedentes

Sun Digital es una empresa especializada en la comercialización, instalación y servicio técnico de soluciones de impresión digital basadas en la tecnología de HP Indigo, posicionándose como el canal exclusivo (“channel partner”) de esta tecnología en México. Este rol implica no solo la distribución de equipos, sino también la integración de un ecosistema completo de soluciones que incluye soporte técnico, asesoría operativa y optimización de procesos productivos en la industria gráfica.

De acuerdo con información institucional, Sun Digital ha sido reconocida como Channel Partner Platinum en Latinoamérica (Sun Digital, s.f.), lo que refleja un alto nivel de especialización técnica y una relación estratégica directa con el fabricante original. Su portafolio incluye prensas digitales de hoja y bobina, así como soluciones complementarias de software y automatización orientadas a mejorar la eficiencia y productividad de las operaciones de impresión digital.

Dentro de sus principales actividades, destaca el asesoramiento técnico en el manejo de sustratos, un aspecto crítico en los procesos de impresión digital con tecnología HP Indigo. Esta tecnología se caracteriza por su capacidad de imprimir sobre una amplia variedad de materiales como papel, cartón, películas plásticas y etiquetas mediante el uso de tintas líquidas electrofotográficas (ElectroInk), lo que requiere un control preciso de variables como la energía superficial del sustrato, la humedad ambiental y la preparación previa del material.

En este contexto, Sun Digital cuenta con personal altamente capacitado que brinda soporte en la selección, acondicionamiento y tratamiento de sustratos, tanto para prensas de alimentación por bobina como de hoja. Este acompañamiento técnico permite asegurar la compatibilidad entre el sustrato y el proceso de impresión, minimizando defectos como problemas de adhesión de tinta, variabilidad de color o atascos de papel, los cuales impactan directamente en la eficiencia operativa y en los costos de producción.

El desarrollo de esta investigación se basa en una situación real identificada en un cliente atendido por Sun Digital, ubicado en el norte de México, donde se han reportado de manera recurrente quejas relacionadas con problemas de atascos de papel en equipos de impresión digital. Esta problemática ha generado afectaciones significativas en la productividad, tiempos de entrega y los costos operativos, evidenciando la necesidad de un análisis más profundo de las condiciones bajo las cuales opera el sistema.

Ante esta situación, el personal técnico especializado de Sun Digital trabaja de manera directa con el cliente, proporcionando soporte en sitio y asesoría técnica continua para la identificación de causas raíz. Este acompañamiento incluye la evaluación de factores críticos como el manejo y almacenamiento del sustrato, la aplicación de tratamientos superficiales, así como el control de variables ambientales dentro del área de producción. La colaboración entre el equipo técnico y el cliente ha permitido recopilar información relevante del proceso, la cual constituye la base empírica para el desarrollo del presente estudio.

De esta manera, el proyecto no solo responde a una necesidad académica, sino que se encuentra directamente vinculado con una problemática industrial real, lo que le confiere un alto valor aplicado y potencial impacto en la mejora de procesos dentro del entorno de impresión digital.

1.2 Descripción del Problema

En la práctica, se ha observado que las fallas no responden exclusivamente a errores de máquina, sino que tienen su origen en múltiples factores interrelacionados, entre ellos la calidad del corte del papel, la orientación del grano, el contenido de humedad, la correcta aplicación del primer, o las condiciones ambientales del área de impresión. Cada uno de estos aspectos puede alterar el comportamiento del pliego dentro de la prensa y, en conjunto, comprometer el rendimiento del proceso. Como resultado, la ocurrencia de atascos produce tiempos muertos que interrumpen el flujo de trabajo,

generan desperdicio de material, disminuyen la productividad y elevan los costos operativos (HP, 2014).

Aunque las prensas de la tecnología HP Indigo cuentan con herramientas de ajuste automático, no siempre es posible minimizar el impacto de un cuidado y manejo deficiente del sustrato y del ambiente. Más aún si la empresa en cuestión carece de registros sobre la frecuencia de fallos y su posible motivo, lo que impide cuantificar su impacto real y dificulta la toma de decisiones informadas para su corrección.

El problema se agrava cuando el mismo papel es procesado internamente, cortado desde bobinas y almacenado sin controles estrictos de humedad o temperatura. Bajo estas condiciones, es frecuente que los pliegos presenten deformaciones, memoria estructural, cargas electrostáticas o bordes imprecisos, lo cual aumenta la probabilidad de interferencias en las zonas críticas del equipo: alimentador, puente de entrada, cilindros de impresión o apilador de salida.

Desde una perspectiva económica, el impacto de esta variabilidad resulta significativo. Considerando un costo unitario de 0.26 USD por hoja impresa y estimando que durante el período analizado (aproximadamente 90 días) se imprimieron cerca de 503,339 hojas, el valor total de producción asciende aproximadamente a 130,868.14 USD.

Durante este mismo periodo se registraron 548 atascos, lo que implica, considerando una pérdida promedio de 4 hojas por evento, un desperdicio total de 2,192 pliegos, equivalente a 569.92 USD asociados únicamente al costo del sustrato.

Adicionalmente, de acuerdo con los datos recuperados en HP PrintOS, plataforma de monitoreo y recolección de datos de las prensas HP Indigo, el tiempo promedio estimado de recuperación para una prensa Serie 4 (S4) posterior a un atasco es cercano a 3 minutos por evento. Bajo esta consideración, los 548 atascos registrados representan aproximadamente 1,644 minutos de tiempo perdido, equivalentes a 27.4 horas de indisponibilidad operacional durante el periodo analizado.

Lo anterior evidencia que el impacto económico real del fenómeno es considerablemente mayor al estimado únicamente por desperdicio de material, ya que no contempla otros costos asociados como tiempo máquina, tiempo operador, pérdida de capacidad productiva, tiempos muertos y afectaciones a la eficiencia global del proceso.

Este escenario evidencia la urgencia de identificar claramente cuales son los factores significativos en la incidencia de los tiempos muertos por atascos, así como de establecer procedimientos técnicos para su control. De esta manera, se podría mejorar la confiabilidad operativa de la maquina y asegurar mayor eficiencia en la producción.

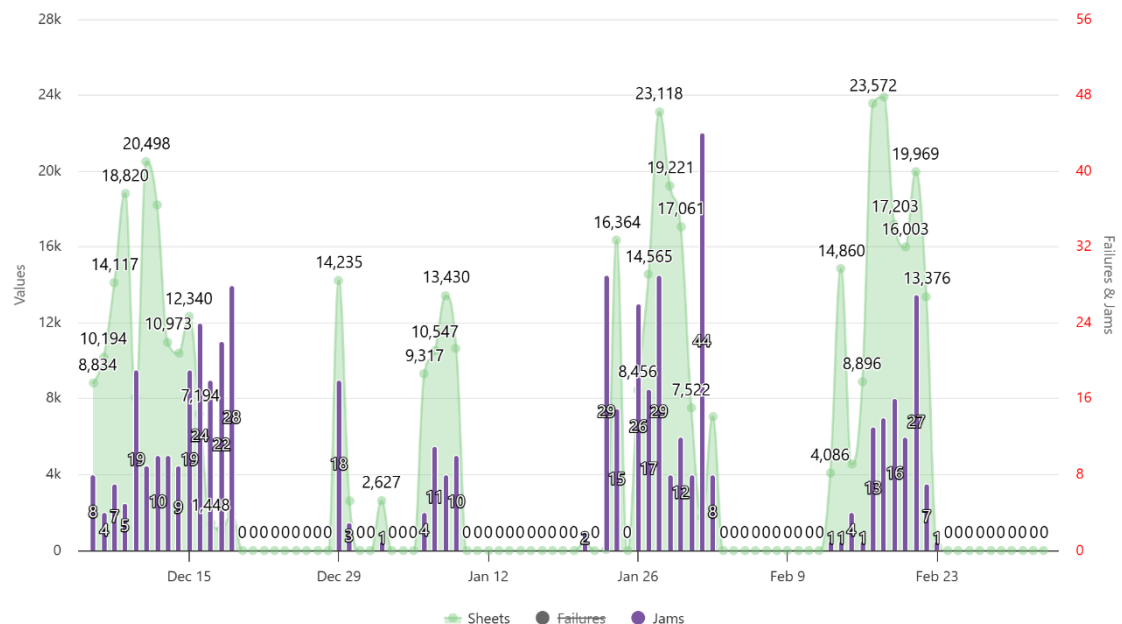


Figura 1.1 Hojas impresas vs atascos.

1.3 Objetivos

1.4. *Objetivo General*

Analizar de manera integral las variables técnicas y ambientales que influyen en el desempeño de la prensa digital HP Indigo serie 4 de hoja, con el propósito de identificar las causas frecuentes de fallas operativas particularmente los atascos de papel y proponer un modelo matemático de predicción para el sistema.

1.4.1 *Objetivos Específicos*

- Evaluar el impacto de las condiciones ambientales (temperatura, humedad, almacenamiento) sobre la estabilidad dimensional y mecánica del sustrato.
- Utilizar herramientas de análisis estadístico para analizar datos históricos recopilados y estudio de resultados.
- Selección del mejor modelo predictivo del proceso.
- Creación de planta virtual con base en el modelo seleccionado para simulación de proceso y obtención de respuesta predictiva.

1.4 Hipótesis

La implementación de un modelo predictivo que permita estimar la cantidad de atascos esperados y controlar los factores críticos del proceso, junto con la conversión de factores de ruido en variables controlables en el manejo y cuidado del sustrato, puede reducir significativamente los tiempos muertos ocasionados por fallas mecánicas (atascos de papel).

1.5 Justificación

La justificación del presente estudio se fundamenta en el análisis de los datos operativos mostrados en la figura 1.1, donde se observa el comportamiento de la producción (hojas impresas), así como la cantidad y tasa aproximada de atascos a lo largo del tiempo. A partir de esta información, es evidente la existencia de una alta variabilidad en el desempeño del proceso de impresión digital, reflejada en fluctuaciones significativas tanto en la cantidad de hojas procesadas como en la frecuencia de atascos. En particular, se identifican periodos con picos elevados de producción que coinciden con incrementos importantes en el número de atascos, alcanzando valores superiores a 40 eventos en ciertos días. Asimismo, se observan intervalos de baja producción asociados a la presencia de fallas, lo que indica interrupciones en la operación y tiempos muertos considerables. Esta relación sugiere la existencia de factores subyacentes no controlados que afectan directamente el rendimiento y la estabilidad del proceso.

La identificación y reducción de las causas raíz de estos eventos se convierte, por tanto, en un factor crítico para mejorar la rentabilidad y eficiencia operativa del sistema de impresión.

La dispersión de los datos evidencia que el sistema no opera bajo condiciones robustas, ya que pequeñas variaciones en condiciones de operación generan impactos significativos en los indicadores clave de desempeño y en el costo de la operación. Entre estos factores potenciales destacan las condiciones ambientales (temperatura y humedad), la precisión dimensional del corte del sustrato y su preparación previa, los cuales influyen directamente en la estabilidad del proceso.

Desde una perspectiva de ingeniería industrial, esta inestabilidad se traduce en pérdidas de productividad, incremento en costos operativos y disminución en la calidad del producto final. Por lo tanto, resulta indispensable implementar un enfoque sistemático que permita identificar, cuantificar y controlar las variables críticas del proceso.

En este contexto, el presente proyecto se justifica al proponer el uso de análisis de regresión como herramienta metodológica para analizar de manera sistemática las interacciones entre las variables del proceso, tales como la temperatura, la humedad relativa, la precisión del corte del sustrato y determinar las condiciones óptimas de operación. A diferencia de enfoques tradicionales basados en prueba y error, el análisis de regresión permite identificar no solo los efectos principales de cada factor, sino también sus interacciones y posibles efectos no lineales, lo cual resulta fundamental en procesos complejos como la impresión digital en prensas HP Indigo, donde múltiples variables influyen simultáneamente en la calidad y estabilidad del proceso.

Asimismo, este enfoque metodológico se alinea con los principios del diseño robusto, al permitir la evaluación y eventual transformación de factores de ruido como la variabilidad ambiental o las condiciones de almacenamiento del sustrato en variables controlables. Esto resulta particularmente relevante en el manejo de sustratos, donde condiciones como la humedad relativa pueden afectar directamente la rigidez, estabilidad

dimensional y comportamiento en el transporte del papel, incrementando la probabilidad de atascos. Al establecer rangos operativos óptimos y estrategias de control, se busca reducir la sensibilidad del proceso ante dichas variaciones externas.

De esta manera, la aplicación del DOE no solo contribuye a la reducción de la variabilidad del proceso, sino que también permite minimizar la incidencia de fallas mecánicas, específicamente los atascos de papel, que representan una de las principales fuentes de tiempo muerto y pérdidas económicas en la operación. Adicionalmente, los resultados obtenidos facilitan la construcción de modelos predictivos capaces de estimar el comportamiento del sistema bajo diferentes condiciones, lo que abre la posibilidad de desarrollar una “planta virtual” para la simulación y toma de decisiones.

En conjunto, este proyecto busca mejorar la eficiencia operativa de la prensa, optimizar el uso de recursos y fortalecer los indicadores clave de desempeño. Finalmente, la implementación de estas metodologías contribuye al fortalecimiento de una cultura de mejora continua y toma de decisiones basada en datos, alineada con las mejores prácticas de manufactura utilizadas en entornos de impresión digital avanzada.

Tabla 1.1 Mediciones tomadas en el periodo Diciembre 2025- Febrero 2026

Total hojas impresas	Total de atascos	Jam rate del periodo
503,339	548	0.11%

1.6 Alcances y limitaciones

El presente proyecto tiene como alcance encontrar el modelo que predice de mejor manera el sistema para simular condiciones reales del proceso y predecir con exactitud la respuesta (tasa de atasco) que serviría como base para generar un plan de producción más acertado, En función de variables críticas del proceso, tales como las condiciones ambientales, corte de papel y funciones operativas de la prensa.

A partir de este modelo, se busca simular distintos escenarios operativos que permitan reproducir condiciones cercanas a la realidad del proceso sin la necesidad de intervención directa en planta. Esto permitirá analizar la sensibilidad del sistema ante variaciones, identificar combinaciones óptimas de operación y estimar su impacto en la producción.

Asimismo, el modelo desarrollado servirá como una herramienta de apoyo para la toma de decisiones en la planeación de la producción, permitiendo anticipar comportamientos del sistema bajo diferentes condiciones y contribuyendo a la definición de estrategias orientadas a la mejora de la eficiencia operativa, reducción de desperdicio y optimización de recursos.

Una de las principales limitantes del estudio es la imposibilidad de realizar experimentación directa en planta bajo condiciones controladas, debido al alto costo asociado a la ejecución de pruebas industriales, el consumo de materiales (sustrato, cuchillas de corte), así como de tiempo máquina y tiempo hombre.

Adicionalmente, la operación de la prensa en un entorno productivo restringe la posibilidad de alterar parámetros de manera sistemática, ya que esto podría comprometer la calidad del producto final, tiempos de entrega y los indicadores clave de desempeño (KPIs) del área.

Por lo tanto, el desarrollo del modelo se basa en datos históricos, datos simulados y supuestos fundamentados en el comportamiento físico del proceso, lo cual introduce cierto nivel de incertidumbre inherente al modelo. En este sentido la precisión de las predicciones estará sujeta a la calidad y variabilidad de los datos disponibles.

Finalmente, el modelo no contempla la totalidad de variables que pueden influir en la generación de atascos y tiempos muertos, como el estado de la prensa, la experiencia del operador o variaciones en el lote del material, por lo que los resultados deben

interpretarse como una aproximación del sistema y no como una representación exacta del mismo.

CAPÍTULO II. MARCO TEÓRICO

2.1 Modelos de regresión

En el análisis estadístico aplicado a procesos industriales, la regresión se suele definir como una “técnica estadística para investigar y modelar la relación entre variables”. Su aplicación permite describir datos, estimar relaciones funcionales y desarrollar modelos firmes para la predicción en sistemas experimentales. (Montgomery et al., 2012).

En muchos sistemas industriales reales, un modelo estrictamente lineal no suele ser suficiente para la representación e interpretación adecuada del proceso, esto se debe a la presencia de interacciones y curvaturas. En tales casos, es necesario recurrir a modelos polinomiales de segundo orden. Según explican Montgomery et al. (2012), estos esquemas son esenciales cuando la respuesta es curvilínea, ya que “incluso las relaciones no lineales complejas pueden modelarse adecuadamente mediante polinomios en rangos razonablemente pequeños”.

Montero Granados (2016) señala que la regresión lineal permite expresar la relación entre variables mediante una función matemática, aportando no solo una medida de asociación, sino también información sobre el efecto marginal de cada factor, el valor esperado de la respuesta y la significancia estadística del modelo. En este sentido, la regresión supera a la simple correlación, ya que no se limita a mostrar si dos variables cambian conjuntamente, sino que permite cuantificar cuánto cambia la variable dependiente cuando se modifica una variable independiente, manteniendo constantes las demás.

En su forma general, el modelo de regresión lineal múltiple puede expresarse como la Ecuación (2.1):

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \cdots + \beta_k x_{kj} + u_j \quad (2.1)$$

donde y_j representa la variable dependiente o respuesta, x_{ij} son las variables independientes o explicativas, β_0 es la ordenada al origen, β_i son los coeficientes de regresión asociados a cada factor y u_j es el término de error, que representa la variabilidad no explicada por el modelo.

Teóricamente, este modelo supone que el valor esperado de la variable respuesta puede describirse como una combinación lineal de los factores de entrada, lo cual constituye la base del modelado estadístico clásico en ingeniería y en análisis experimental. El modelo lineal resulta especialmente útil como una primera aproximación para estudiar el efecto de variables como la temperatura, la humedad relativa, la precisión del corte del sustrato y otras condiciones operativas sobre la tasa de atascos. Una de sus principales ventajas es que permite interpretar de manera directa el efecto marginal de cada variable. Es decir, cada coeficiente estimado indica la variación esperada en la tasa de atascos ante un cambio unitario de un factor, mientras los demás permanecen constantes. Esto simplifica el proceso de identificar qué variables ejercen mayor influencia sobre la respuesta del proceso y, por lo tanto, cuáles deben priorizarse en actividades de control, mejora o simulación.

Sin embargo, en muchos sistemas industriales reales, la relación entre las variables de entrada y la respuesta no es estrictamente lineal. Existen procesos en los que la respuesta presenta curvatura, puntos óptimos o interacciones entre factores, de manera que el efecto de una variable depende del nivel en que se encuentre otra. Cuando esto ocurre, el modelo lineal simple o múltiple puede resultar insuficiente para representar adecuadamente el comportamiento del sistema. En estos casos, es necesario recurrir a modelos polinomiales de segundo orden, también conocidos como modelos cuadráticos, ampliamente utilizados en metodología de superficie de respuesta. Montgomery y sus coautores explican que estos modelos son apropiados cuando se requiere aproximar una

superficie de respuesta en una región experimental determinada y capturar tanto efectos lineales como curvatura e interacciones entre factores. La literatura de superficie de respuesta también establece que el modelo de segundo orden es la base de los diseños experimentales usados para buscar máximos, mínimos o condiciones óptimas de operación, porque ofrece una buena aproximación local del sistema cerca de una región de interés

El modelo de segundo orden puede expresarse como la Ecuación (2.2):

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j}^k \beta_{ij} x_i x_j + \varepsilon \quad (2.2)$$

donde y es la variable respuesta, β_0 es la constante, β_i representan los efectos lineales, β_{ii} los efectos cuadráticos, β_{ij} los efectos de interacción entre factores y ε es el término aleatorio de error. La inclusión de términos cuadrados x_i^2 permite modelar curvatura, mientras que los productos cruzados $x_i x_j$ permiten representar interacciones, es decir, situaciones en las que el efecto de un factor cambia según el valor de otro factor.

En la presente tesis, el uso de modelos de segundo orden resulta relevante debido a la posible interacción entre variables como la humedad y el corte del sustrato, las cuales pueden generar efectos no lineales en la tasa de atascos. De esta manera, el modelo cuadrático permite describir mejor el comportamiento del proceso y mejorar la capacidad predictiva del sistema. Este modelo resulta relevante para el estudio, ya que la tasa de atascos en un proceso de impresión digital puede depender de relaciones no lineales entre las condiciones ambientales y las características físicas del sustrato. Por ejemplo, una ligera desviación en el corte del papel podría tener un efecto pequeño bajo condiciones estables de humedad, pero ese mismo desvío podría amplificar su impacto cuando la humedad relativa se eleva o cuando la temperatura se aparta del rango óptimo.

Este tipo de comportamiento no puede describirse con suficiente precisión mediante un modelo puramente lineal, pero sí mediante un modelo de segundo orden que

incorpore efectos cuadráticos e interacciones. Por ello, desde el punto de vista teórico, el modelo cuadrático no solo representa una extensión matemática del modelo lineal, sino una herramienta de mayor capacidad para aproximar fenómenos reales, construir superficies de respuesta y apoyar decisiones relacionadas con la optimización del proceso.

2.2 ANOVA

La obra de Montgomery sobre diseño y análisis de experimentos presenta el ANOVA como una herramienta fundamental para analizar efectos de factores, contrastar hipótesis y validar modelos en estudios experimentales. En el modelado estadístico es fundamental interpretar indicadores como el coeficiente de determinación (R^2), el coeficiente de determinación ajustado (R^2 ajustado) y los valores p asociados a cada término del modelo. En este sentido, el análisis estadístico moderno no se limita únicamente a evaluar el ajuste, sino también a comparar múltiples estructuras de modelos mediante criterios adicionales como AIC y PRESS, los cuales permiten valorar el equilibrio entre precisión y complejidad del modelo (Keka & Cxixco, 2025).

El coeficiente R^2 mide la proporción de la variabilidad de la variable respuesta que es explicada por el modelo. Un valor cercano a 1 indica que el modelo tiene una alta capacidad explicativa, mientras que valores bajos indican que gran parte de la variabilidad no está siendo capturada. No obstante, Montgomery (2017) menciona que este indicador puede resultar engañoso cuando se incorporan variables adicionales, ya que “el valor de R^2 tiende a aumentar aun cuando las variables adicionales no aportan información significativa al modelo”. Adicionalmente, estudios comparativos recientes indican que el ajuste del modelo debe analizarse junto con otras métricas, ya que “el desempeño del modelo no solo depende del ajuste, sino también de su capacidad predictiva y de su parsimonia” (Keka & Cxixco, 2025).

Por esta razón, se utiliza el coeficiente de determinación ajustado (R^2 ajustado), el cual penaliza la inclusión de variables innecesarias y proporciona una medida más confiable del ajuste del modelo. Este indicador permite comparar modelos con diferente

número de variables y seleccionar aquel que mantiene un equilibrio entre capacidad explicativa y complejidad del modelo. Autores destacan que “un mayor valor de R^2 ajustado indica un mejor ajuste del modelo, ya que explica una mayor proporción de la variabilidad de la variable dependiente” (Keka & Cxixco, 2025). Esto resulta especialmente relevante en problemas reales, como el analizado en esta investigación, donde la inclusión de términos cuadráticos e interacciones puede mejorar significativamente la capacidad predictiva del modelo.

Adicionalmente, los valores p asociados a cada coeficiente permiten evaluar la significancia estadística individual de los factores. Un valor p bajo (generalmente menor a 0.05) indica que existe evidencia suficiente para afirmar que el factor es significativo sobre la variable respuesta. Por el contrario, valores p elevados sugieren que el factor no contribuye de manera relevante al modelo y podría ser eliminado para simplificarlo sin perder capacidad predictiva. Según Keka y Cxixco (2025), “los valores p permiten determinar si las variables independientes están significativamente relacionadas con la variable dependiente”, lo cual nos permite establecer un criterio al momento de seleccionar y mejorar modelos.

Asimismo, el análisis de varianza (ANOVA) permite evaluar la significancia global del modelo en base a la estadística F , en donde se contrasta la hipótesis nula la cual afirma que todos los coeficientes son iguales a cero, frente a la hipótesis alternativa, en la que al menos uno de ellos es significativo. Para complementar esta idea, Keka y Cxixco (2025) mencionan que “un valor alto de la estadística F indica que el modelo es significativo y que las variables independientes están relacionadas con la variable dependiente”. Este análisis es importante y esencial para validar la estructura del modelo antes de su implementación en entornos reales.

2.2.1 Estadístico AIC y PRESS

El AIC por sus siglas en inglés (Akaike Information Criterion), es una herramienta estadística utilizada para comparar y seleccionar modelos matemáticos. Este criterio fue

desarrollado por Hirotugu Akaike y tiene como objetivo identificar el modelo que mantiene un equilibrio entre calidad de ajuste y complejidad estadística, esto se debe a que un modelo con demasiados términos puede ajustarse de manera muy precisa a los datos históricos, pero perder capacidad de generalización al aplicarse en nuevos escenarios. Este fenómeno es conocido como sobreajuste (overfitting) y representa uno de los principales problemas en el modelado estadístico. Por esta razón, el criterio AIC favorece modelos más simples siempre que mantengan un desempeño predictivo adecuado.

En términos prácticos, el modelo con el valor más bajo de AIC es considerado el más adecuado dentro del conjunto de modelos evaluados, ya que representa el mejor compromiso entre precisión y simplicidad matemática. Este criterio es ampliamente utilizado en regresión lineal, análisis multivariable y selección de modelos predictivos dentro de aplicaciones industriales y científicas.

El estadístico PRESS (Prediction Error Sum of Squares) es una medida estadística utilizada para evaluar la capacidad predictiva de un modelo de regresión. Su objetivo principal consiste en cuantificar el error de predicción del modelo cuando se estiman observaciones que no participaron directamente en el proceso de ajuste estadístico. A diferencia de indicadores tradicionales como el coeficiente de determinación (R^2), el PRESS se enfoca específicamente en la capacidad del modelo para generalizar y predecir nuevos datos.

El cálculo del PRESS se basa en un procedimiento de validación cruzada interna. Durante este proceso, cada observación experimental es eliminada temporalmente del conjunto de datos y posteriormente estimada utilizando el modelo ajustado con el resto de las observaciones. El error obtenido entre el valor real y el valor predicho es elevado al cuadrado y acumulado para todas las observaciones, generando así la suma total de errores de predicción.

Un valor bajo de PRESS indica que el modelo presenta una alta capacidad predictiva y que los errores de estimación son reducidos. Por el contrario, valores elevados indican que el modelo posee dificultades para representar correctamente nuevos escenarios operativos, aun cuando el ajuste a los datos originales pueda parecer adecuado.

2.3 Diseño robusto.

Los métodos del diseño robusto desarrollados por Taguchi son particularmente útiles por que permiten incorporar factores de ruido tal como la humedad ambiental dentro del experimento, con el objetivo de identificar configuraciones que reduzcan la sensibilidad del proceso ante las variaciones no controlables (Gutiérrez, 2007). Por el contrario, los enfoques tradicionales, como los diseños factoriales clásicos, poseen mayor capacidad para estimar interacciones complejas entre factores de control, lo cual complementa el análisis cuando se requiere un entendimiento más profundo del proceso (Arias-Nava et al., 2015).

El diseño de experimentos robustos, también conocido como diseño de parámetros, es considerablemente utilizado en la optimización de procesos y productos, especialmente en áreas de manufactura y mejora continua. Su principio central es seleccionar adecuadamente los niveles de los factores de control para lograr que el desempeño del sistema sea estable y consistente, aun cuando existan variaciones en los factores de ruido que no pueden controlarse directamente (Gutiérrez, 2007; Arias-Nava et al., 2015). Por ende, un proceso robusto es aquel que mantiene su funcionamiento cerca del valor objetivo de la característica de calidad, con mínima variabilidad posible, incluso cuando se presenta en situaciones que no son controlables y pueden llegar a ser cambiantes.

La robustez, por lo tanto, no se obtiene mediante el control directo del ruido, sino mediante la elección inteligente de los niveles de los factores de control que compensan o reducen el impacto del ruido.

La metodología reconoce que pueden existir interacciones relevantes entre los factores de control y los factores de ruido, y que estas interacciones son esenciales para que un diseño pueda considerarse efectivamente robusto. Identificar dichas interacciones permite seleccionar la combinación de niveles de los factores de control que produzca la menor variabilidad posible frente a perturbaciones externas. De igual manera, el diseño robusto puede incorporar interacciones entre factores de control, lo que ofrece un análisis más detallado y preciso de las condiciones óptimas del sistema (Arias-Nava et al., 2015).

En conjunto, esta aproximación metodológica permite no solo mejorar la calidad del proceso al reducir la variabilidad, sino también establecer condiciones de operación que mantengan el desempeño del sistema lo más cercano posible al valor objetivo, aun ante fluctuaciones externas. Por ello, el diseño robusto constituye una herramienta clave para la optimización estadística de procesos donde la estabilidad y la consistencia del desempeño son elementos críticos.

La Guía de Sustratos de HP (HP, 2014) identifica numerosas condiciones que influyen en la operabilidad y la calidad de impresión, las cuales, bajo la metodología de diseño robusto se clasifican en factores de ruido (no controlables directamente) y factores de control (ajustables para optimizar el desempeño).

Factores de Ruido: Son aquellas fuentes de variabilidad que no pueden ser controladas directamente por el fabricante o el operador en el punto de uso, pero cuyo efecto debe ser minimizado mediante la elección correcta de los factores de control. Taguchi clasifica las variables ambientales, la temperatura y la humedad, como ruido externo. Según la guía de HP, los principales factores que el producto o proceso deben ser insensibles o resistentes a son:

Humedad relativa del ambiente y Temperatura ambiente: Variables ambientales que pueden causar que las fibras del papel se hinchen y encojan, lo que origina problemas como la curvatura, bordes ondulados o apretados. Aunque la guía establece rangos

óptimos de almacenamiento y operación (50% – 70% de humedad relativa y 20°C – 25°C de temperatura), estos factores causan disturbios en el diseño robusto porque el proceso debe funcionar bien a pesar de las fluctuaciones dentro de la sala de impresión.

Factores de Control: Los factores de control son las variables o parámetros del proceso que pueden ser ajustados por el usuario, el diseñador o el personal de preparación del material. La selección correcta de los niveles de estos factores es el objetivo del diseño robusto.

Según la Guía de Sustratos de HP, los factores de control clave relacionados con el manejo y la preparación del sustrato incluyen:

- **Dirección del grano:** La guía de HP hace una recomendación específica basada en el peso del papel (grano corto para >170 gsm y grano largo para <170 gsm). Esta es una decisión de selección y preparación del material que afecta la rigidez y la operabilidad
- **Precisión del corte del papel:** Se requiere un corte perfectamente cuadrado para la operabilidad y el registro, tolerancia ± 0.2
- **Tiempo de aclimatación del papel:** El papel debe aclimatarse en la sala de impresión por un mínimo de 24 horas para alcanzar el equilibrio con el ambiente y la prensa. Este es un parámetro de manejo controlable.

2.4 Método Montecarlo.

El método de Monte Carlo establece una técnica de simulación estocástica utilizada para el análisis de sistemas complejos, está basada en el uso de muestreo aleatorio y modelado estadístico con el objetivo de estimar funciones matemáticas y reproducir el comportamiento de procesos reales. En términos generales, esta metodología permite aproximar soluciones a problemas cuya formulación analítica resulta difícil o incluso imposible de resolver de manera exacta, mediante la generación de múltiples escenarios simulados a partir de variables aleatorias (Harrison, 2010).

En la ingeniería y la manufactura, particularmente dentro del paradigma de la Industria 4.0, esta técnica permite evaluar el desempeño de procesos mediante indicadores clave (KPI), analizar riesgos operativos y generar predicciones basadas en datos históricos, contribuyendo a la optimización de la toma de decisiones (Kostka et al., 2022). Por ende, el hecho de que la simulación de Monte Carlo utilice muestreo aleatorio y un modelado estadístico para poder estimar funciones matemáticas y simular el comportamiento de los sistemas, pone de manifiesto su capacidad para representar fenómenos cuya solución analítica resulta limitada o inviable. (Harrison, 2010, p. 17).

El fundamento del método radica en la generación de números pseudoaleatorios los cuales permiten realizar muestreos independientes sobre distribuciones de probabilidad previamente definidas. A partir de estos muestreos, se construyen estimaciones estadísticas de los parámetros de interés del sistema bajo estudio. La lógica central del método establece que, mediante la repetición masiva de experimentos aleatorios, es posible aproximar el valor esperado de una variable o caracterizar su distribución de probabilidad. Este principio se encuentra sustentado en la Ley de los Grandes Números, la cual establece la convergencia de los promedios muestrales hacia el valor esperado conforme aumenta el número de iteraciones.

En cuanto a su origen, aunque existen antecedentes históricos como el experimento de la aguja de Buffon en el siglo XVIII, el desarrollo formal del método se consolidó durante la Segunda Guerra Mundial. Fue en este contexto donde Stanislaw Ulam y John von Neumann desarrollaron esta técnica dentro del Proyecto Manhattan para analizar fenómenos asociados a la difusión de neutrones. En palabras de Harrison (2010), “esta forma de simulación fue desarrollada y utilizada de manera sistemática por primera vez durante el Proyecto Manhattan” (p. 18), lo que evidencia su origen en problemas físicos de alta complejidad. El término “Monte Carlo” hace referencia al casino de Mónaco, aludiendo al carácter probabilístico del método.

A lo largo de su evolución, el método ha dado lugar a diversas variantes que responden a diferentes necesidades de modelado. El método de Monte Carlo clásico se basa en la generación de muestras independientes a partir de distribuciones de probabilidad, siendo ampliamente utilizado por su simplicidad. Se reconoce que “el método estocástico conocido como simulación de Monte Carlo utiliza generadores de números pseudoaleatorios para el muestreo independiente y el análisis estadístico” (Kostka et al., 2022, p. 2). Por otra parte, el método de Monte Carlo por Cadenas de Markov (MCMC) introduce dependencia entre las muestras, lo que permite abordar problemas donde las distribuciones no pueden obtenerse de forma analítica. Asimismo, el método Quasi-Monte Carlo (QMC) sustituye el muestreo puramente aleatorio por secuencias deterministas de baja discrepancia, lo que mejora la eficiencia del muestreo y la velocidad de convergencia en problemas de alta dimensionalidad (Di, 2022).

En la actualidad, el método de Monte Carlo se ha consolidado como una herramienta fundamental en múltiples disciplinas debido a su capacidad para modelar incertidumbre y analizar sistemas complejos. En la ingeniería y la manufactura avanzada, particularmente dentro del paradigma de la Industria 4.0, su aplicación permite evaluar el desempeño de procesos mediante indicadores clave (KPI), analizar riesgos operativos y generar predicciones basadas en datos históricos. De acuerdo con Kostka et al. (2022), “las simulaciones de Monte Carlo pueden desempeñar un papel importante en la mejora de las estimaciones” (p. 1), destacando su relevancia en entornos industriales digitalizados. De manera complementaria, en el campo del aprendizaje automático, se utiliza para aproximar funciones complejas y reducir la carga computacional en modelos de gran escala, mientras que en las ciencias de la salud resulta esencial para la simulación de fenómenos físicos, como la propagación de partículas en tejidos biológicos (Harrison, 2010; Di, 2022).

En síntesis, el método de Monte Carlo representa una herramienta robusta para el análisis probabilístico de sistemas complejos, cuya principal fortaleza radica en su capacidad para transformar problemas deterministas o de alta incertidumbre en modelos

simulados que permiten obtener estimaciones confiables a partir de la repetición de experimentos aleatorios.

2.5 Plantas virtuales y gemelos digitales

El concepto de planta virtual, también conocido como Digital Twin o gemelo digital, es la representación digital de un sistema físico que permite modelar, simular y analizar su comportamiento a lo largo del tiempo. Este enfoque constituye una de las bases fundamentales de la manufactura inteligente, ya que integra datos reales, simulación y análisis para mejorar el desempeño de los sistemas productivos.

Desde una perspectiva conceptual, el gemelo digital puede definirse como “una representación virtual de un producto o proceso físico” (Soori et al., 2023, p. 1). Esta definición implica que el modelo digital no solo replica la geometría del sistema, sino también su comportamiento, desempeño e interacción con el entorno. Adicionalmente, Grieves (2014) plantea que el gemelo digital permite establecer una relación directa entre el producto diseñado y el producto fabricado, cerrando el ciclo entre diseño y ejecución. El modelo de planta virtual se compone de tres elementos principales: el sistema físico, su representación digital y el flujo bidireccional de datos que conecta ambos espacios. Esta integración permite construir modelos capaces de reflejar el comportamiento real del sistema productivo, incluso en condiciones dinámicas, facilitando su análisis y optimización continua.

Uno de los principales beneficios de las plantas virtuales es la mejora en la eficiencia operativa y la reducción de costos en los sistemas de manufactura. Esto se logra mediante la simulación de procesos en entornos virtuales, lo que permite evaluar diferentes configuraciones del sistema sin afectar la operación real. Por ello, el uso de gemelos digitales permite analizar múltiples escenarios de producción y optimizar el desempeño del sistema antes de implementar cambios físicos. Otro de los beneficios relevantes es la mejora en la calidad del producto. Al integrar datos provenientes del sistema físico, las plantas virtuales permiten identificar patrones, detectar desviaciones y

reducir la generación de productos defectuosos. Esta capacidad de análisis contribuye a la optimización del proceso productivo y al aseguramiento de la calidad.

Como se mencionó anteriormente, las plantas virtuales permiten validar procesos antes de su implementación. Por medio de la simulación en entornos digitales, es posible evaluar el comportamiento del sistema bajo diferentes condiciones operativas, lo que reduce los riesgos asociados a pruebas en el entorno real. Esto resulta particularmente importante en sistemas de manufactura complejos, donde los costos de error pueden ser elevados.

Otro aspecto clave es el monitoreo en tiempo real del sistema productivo. La integración de sensores y datos permite que el modelo virtual refleje continuamente el estado del sistema físico, facilitando la detección temprana de fallas y la toma de decisiones basada en datos. De esta manera, el gemelo digital permite no solo analizar el sistema, sino también intervenir de manera oportuna para mejorar su desempeño.

Finalmente, también es importante mencionar la ventaja de disminuir el tiempo de desarrollo de los productos. Al permitir la simulación y validación de diseños en entornos virtuales, se reduce la necesidad de prototipos físicos, lo que acelera el proceso de desarrollo y disminución de los costos asociados.

2.6 Tecnología de impresión HP Indigo

En la impresión digital comercial, la tecnología HP Indigo representa una de las soluciones más avanzadas en la actualidad, particularmente en sistemas de impresión en formato pliego. La prensa HP Indigo serie 4, se fundamenta en el principio de electrofotografía líquida (Liquid Electrophotography, LEP), un sistema propietario de HP que combina la precisión de la formación de imagen láser con las ventajas operativas de las tintas líquidas.

El proceso de impresión inicia con la generación de una imagen latente sobre un tambor fotoconductor conocido como PIP (Photo Imaging Plate), el cual es posteriormente revelado mediante la aplicación de tinta líquida ElectroInk. Esta tinta es transferida a una manta intermedia calentada, donde se forma una película delgada y uniforme que finalmente es depositada sobre el sustrato mediante la acción combinada de presión y temperatura controladas. Este mecanismo permite obtener una calidad de impresión superior, con resoluciones visuales equivalentes a 4876×4876 dpi, así como una excelente uniformidad de color y detalle fino (HP, 2023).

Adicionalmente, la prensa HP Indigo serie 4 de hoja ofrece una alta versatilidad en términos de compatibilidad de materiales, siendo capaz de imprimir sobre más de 2,900 tipos de sustratos certificados, incluyendo papeles estucados y no estucados, cartulinas, materiales sintéticos y superficies metalizadas. Su arquitectura modular, que puede incluir hasta siete estaciones de tinta, permite la incorporación de colores especiales como blanco opaco, tintas fluorescentes o tintas de seguridad, lo que amplía significativamente su rango de aplicaciones industriales.

Desde una perspectiva operativa, la integración del sistema PrintOSx proporciona capacidades avanzadas de monitoreo y control en tiempo real, permitiendo la optimización continua del proceso mediante análisis predictivo y ajustes automáticos. Esto convierte a la prensa no solo en un sistema de impresión de alta calidad, sino en una plataforma inteligente capaz de adaptarse dinámicamente a las condiciones del proceso.



Figura 2.1 Prensa digital HP Indigo serie 4.

2.7 Comportamiento del sustrato en impresión digital

El sustrato es uno de los elementos más críticos dentro del proceso de impresión digital, particularmente en tecnologías como HP Indigo, donde la tinta no es absorbida por el material, sino que permanece como una capa superficial adherida. Esta característica implica que propiedades como la energía superficial, la planicidad, la rigidez, la rugosidad, la dirección del grano y el contenido de humedad del papel adquieran una relevancia determinante en la calidad del producto final.

De acuerdo con la Guía de Sustratos de HP (HP, 2014), la selección del gramaje del papel debe considerar tanto la aplicación final como la capacidad mecánica de la prensa para manipular el material sin generar deformaciones. Por ejemplo, se recomienda evitar gramajes inferiores a 90 g/m² en trabajos dúplex, debido al riesgo de translucidez y pérdida de calidad visual. Asimismo, la dirección del grano influye directamente en la estabilidad dimensional del papel durante su paso por la prensa, siendo preferible el grano corto en papeles de mayor gramaje para mejorar su rigidez y comportamiento mecánico.

El contenido de humedad del sustrato representa otro factor crítico. Se recomienda mantener valores entre 4.5 % y 5.5 %, ya que niveles bajos pueden generar acumulación de carga electrostática, mientras que niveles elevados pueden provocar expansión del papel, curvatura o deformaciones conocidas como “memoria”. Estas condiciones afectan directamente la alimentación del material y aumentan la probabilidad de fallos operativos.

De igual manera, la precisión del corte del papel debe mantenerse dentro de tolerancias estrictas, típicamente del orden de ± 0.2 mm, ya que desviaciones acumuladas pueden generar problemas de registro e incluso atascos dentro del sistema. Por ello, la verificación dimensional del sustrato previo a su uso en máquina constituye una práctica fundamental para garantizar la estabilidad del proceso.

Más allá de sus propiedades físicas visibles, el comportamiento del papel en impresión digital también está determinado por características estructurales y químicas que influyen en su interacción con el entorno y con los procesos de impresión. Es por ello que la durabilidad del papel adquiere relevancia, especialmente en aplicaciones donde la permanencia del producto impreso es un requisito.

Estudios como el de Černic Letnar y Muck (2003) señalan que incluso papeles fabricados con fibras de alta pureza pueden degradarse rápidamente si no cuentan con una reserva alcalina adecuada. La presencia de carbonato de calcio, por ejemplo, actúa como un agente amortiguador frente a procesos de acidificación derivados de la humedad, el calor o contaminantes metálicos. La ausencia de esta protección puede traducirse en pérdida de resistencia mecánica, desprendimiento de fibras y generación de partículas, lo cual incrementa la probabilidad de fallos dentro del sistema de impresión.

En paralelo, investigaciones sobre sustratos recubiertos con materiales poliméricos han demostrado que tratamientos superficiales como descarga corona, plasma atmosférico o tratamiento por llama pueden modificar significativamente la energía superficial del material. Estos procesos inducen la formación de grupos funcionales polares, mejorando la adherencia de tintas líquidas. Este principio es consistente con el uso de primers como DigiPrime8020 en sistemas HP Indigo, los cuales buscan optimizar la compatibilidad entre la tinta y el sustrato.

El tratamiento por llama, en particular, ha mostrado ventajas en términos de uniformidad y calidad de impresión en comparación con otros métodos, lo que abre la posibilidad de implementar estrategias combinadas de tratamiento físico y químico para mejorar el desempeño de materiales complejos.

Sin embargo, incluso con un sustrato adecuadamente preparado, el desempeño del sistema depende también de la capacidad de control de la prensa. Por ello, la arquitectura de control de la HP Indigo permite ajustar automáticamente parámetros como

temperatura, conductividad, densidad óptica y carga electrostática mediante sistemas de retroalimentación en tiempo real (Tzori, 2004). Sin embargo, estos sistemas operan dentro de un rango de tolerancia, por lo que condiciones extremas del sustrato o del ambiente pueden superar su capacidad de compensación, generando defectos de impresión o fallos operativos.

En consecuencia, el desempeño óptimo del sistema no depende únicamente de la tecnología de la prensa, sino de la interacción entre el sustrato, las condiciones ambientales, los tratamientos aplicados y los parámetros de operación.

Tabla 2.1 Condiciones recomendadas para sustratos y ambiente de impresión

Variable	Valor recomendado
Humedad relativa del ambiente	50 % – 70 %
Temperatura ambiente	20 °C – 25 °C
Contenido de humedad del papel	4.5 % – 5.5 %
Dirección del grano	Grano corto (>170 g/m ²), largo (<170 g/m ²)
Corte del papel	Tolerancia ±0.2 mm, sin rebabas
Tiempo de aclimatación del papel	24 horas en sala de impresión

CAPÍTULO III. METODOLOGÍA

La metodología propuesta para la realización del presente estudio esta propuesta en la figura 3.1:

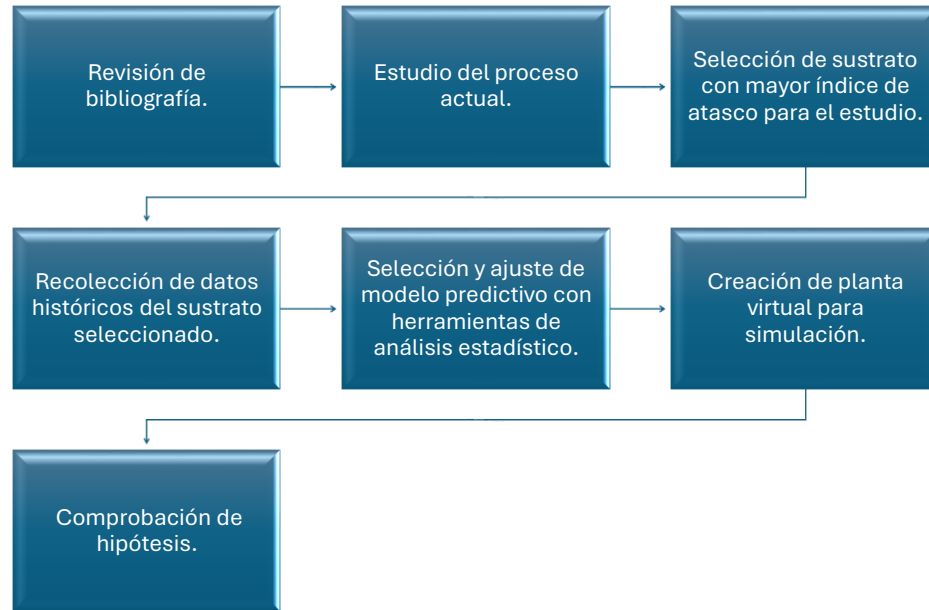


Figura 3.1 Descripción general metodología del proyecto.

3.1 Revisión de la bibliografía.

La revisión de la bibliografía fue la base fundamental de la investigación desarrollada en el presente marco teórico, permitiendo comprender y contextualizar los principales conceptos relacionados con los modelos de regresión, la simulación mediante plantas virtuales y la tecnología de impresión digital HP Indigo. A partir del análisis de libros, artículos científicos, documentación técnica y manuales especializados, fue posible establecer un sustento metodológico sólido para el desarrollo de la investigación y para la posterior construcción del modelo predictivo del sistema de impresión.

El estudio bibliográfico mostro la relevancia e importancia de los modelos estadísticos en los procesos industriales modernos, particularmente en sistemas donde existe una alta variabilidad operativa, donde múltiples factores interactúan

simultáneamente. En este contexto, los modelos de regresión lineal y cuadrática fueron analizados como herramientas capaces de describir matemáticamente el comportamiento de variables críticas del proceso, tales como la temperatura, humedad relativa, dimensiones reales del sustrato y frecuencia de atascos. También, se revisaron conceptos relacionados con análisis de varianza (ANOVA), capacidad predictiva, criterios de selección de modelos y validación estadística, los cuales sirvieron como fundamento para seleccionar el modelo con mejor desempeño predictivo.

De igual manera, la investigación documental permitió comprender el funcionamiento y la relevancia de las plantas virtuales dentro de la industria moderna. La simulación virtual de procesos industriales representa actualmente una herramienta estratégica para predecir escenarios operativos, reducir riesgos de producción y optimizar la toma de decisiones antes de implementar cambios directamente en la planta. Por medio del estudio de metodologías de simulación y técnicas de Monte Carlo, se identificó la posibilidad de incorporar variabilidad aleatoria y comportamiento probabilístico al sistema de impresión, permitiendo representar de forma más realista las condiciones reales de operación.

Por último, el análisis de la tecnología HP Indigo permitió comprender el comportamiento físico y operativo del sistema de impresión digital, particularmente la interacción entre la ElectroInk y las propiedades del sustrato. Se estudiaron aspectos relacionados con humedad ambiental, tolerancias dimensionales del papel, preparación del sustrato, aplicación de primer y parámetros de alimentación del sistema, identificándose como variables potencialmente influyentes en la generación de atascos y variaciones de calidad. Esta revisión fue esencial para definir las variables de entrada consideradas posteriormente en el modelo experimental y en la planta virtual predictiva.

3.2 Estudio del proceso actual.

En una primera etapa, se estudió el proceso actual con el objetivo de comprender las condiciones operativas de la prensa, identificar patrones de comportamiento y detectar

las condiciones operativas de las principales fuentes de variabilidad, tales como la ocurrencia de atascos, tiempos muertos y fluctuaciones en la producción. Para ello se utiliza la herramienta del diagrama de Ishikawa.

El diagrama de Ishikawa, también denominado como diagrama de causa-efecto, es una de las herramientas de calidad que facilita la identificación de las causas que generan problemas en los procesos productivos. La relevancia de esta herramienta radica en la capacidad de estructurar gráficamente posibles factores claves y facilita la visualización general del problema.

3.3 Selección de sustrato a estudiar

Posteriormente, se lleva a cabo la selección del sustrato a estudiar considerando el sustrato más utilizado en la prensa y que tiene un número alto de atascos como se muestra en la figura 3.2 para garantizar la consistencia de los datos recolectados. Para el estudio se utilizó el sustrato denominado Multicolor 17 el cual cada pliego es de dimensiones físicas:

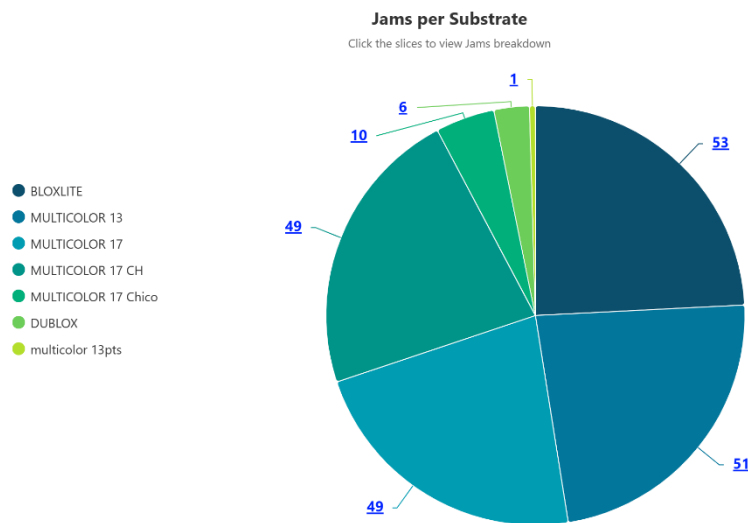


Figura 3.2. Grafica de pastel de sustratos utilizados en las prensas (datos obtenidos de PrintOS).

Tabla 3.1 Propiedades físicas del sustrato

Propiedad	Valor
Ancho	750 mm
Largo	530 mm
Peso	260 GSM
Grosor	350 μm
Revestimiento	Revestido
Grano	Largo

3.4 Recolección de datos históricos del sustrato seleccionado

Los datos analizados en este estudio serán recopilados durante un periodo aproximado de 90 días de operación de la prensa digital, durante el cual se registraron diariamente variables operativas tales como la temperatura promedio del área de impresión, la humedad relativa del ambiente, el número de hojas impresas y la cantidad de atascos de papel observados durante la producción.

3.5 Selección y ajuste de modelo predictivo con herramientas de análisis estadístico

Se analizarán los datos obtenidos en el periodo especificado para hacer el análisis estadístico en herramientas de software con el objetivo de identificar los factores significantes y obtener la ecuación del modelo que mejor represente el comportamiento del proceso. Esto se llevará a cabo mediante la recolección de datos operativos y la simulación de variables críticas, particularmente la longitud de corte del sustrato.

Posteriormente, se hará un análisis mediante ANOVA, con el fin de identificar la significancia estadística de las variables estudiadas y sus interacciones.

3.6 Creación de planta virtual para simulación

Con base en los resultados obtenidos, se llevó a cabo la elaboración de una planta virtual orientada a simular la optimización del proceso. Esta etapa tiene como finalidad comprobar la hipótesis del proyecto, mediante una simulación en la cual convertiremos el factor de ruido más significativo o el factor más alcanzable en un factor de control para comprobar que la tasa de atascos disminuye.

CAPÍTULO IV. RESULTADOS

4.1 El proceso actual.

Como beneficios del uso del diagrama de Ishikawa, se pueden mencionar los siguientes: una mejor comprensión de un proceso o equipo, la iniciación de un proceso de innovación, la facilitación del aprendizaje y el intercambio de ideas, una mejor gestión de los factores capaces de generar efectos menos favorables, y el eventual establecimiento de la necesidad de elaborar normas técnicas, entre otros (Botezatu et al., 2019).

Los autores concluyen que el uso del diagrama de Ishikawa permite identificar de manera estructurada los factores que inciden en la calidad de los procesos estudiados, clasificándolos, lo cual contribuye a esclarecer las acciones correctivas necesarias.

Este antecedente muestra que dicha herramienta no solo es aplicable en entornos de manufactura tradicional, sino que su aplicación abarca también procesos tecnológicos avanzados tales como la impresión digital en este caso. Confirma que la herramienta resulta eficaz para comprender la complejidad de las 6M (materiales, métodos, maquinaria, mano de obra, medición y medio ambiente). Por lo tanto, el uso de dicha herramienta en este estudio se justifica plenamente, ver figura 4.1.

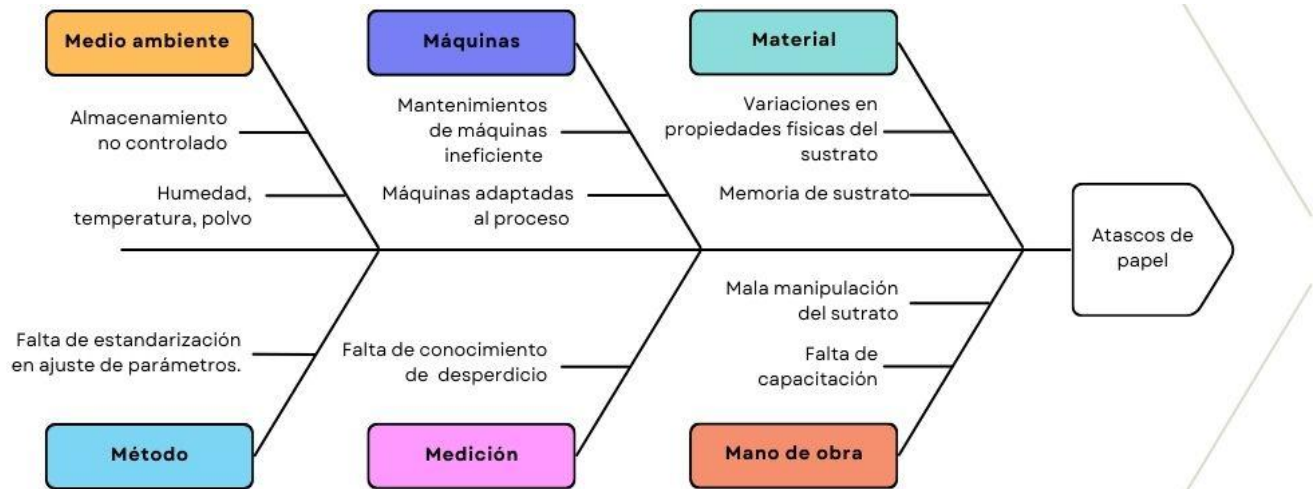


Figura 4.1 Diagrama de Ishikawa del proceso.

Dentro del estudio del proceso actual, se identificó que la empresa donde se desarrolla el análisis emplea un sistema de corte de sustrato mediante una CHM 1400 Sheeter Cutter (figura 4.2), el cual forma parte crítica en la preparación del papel previo a su ingreso a la prensa HP indigo. Este equipo realiza la conversión de bobinas en hojas mediante un sistema de cuchillas rotativas sincronizadas con la velocidad del material, con el objetivo de obtener una longitud nominal de corte de 740 mm; sin embargo, durante el levantamiento de información en planta se detecta que dicho sistema presenta condiciones de operación no óptimas asociadas principalmente a una falta de mantenimiento preventivo y control operativo. En particular, se observaron indicios de desgaste en cuchillas, posibles desajustes en los sistemas de tracción y control de tensión del papel, así como variabilidad en la sincronización del corte, lo que en conjunto genera desviaciones en la longitud real del sustrato (dentro de rangos aproximados de 748 a 752 mm).



Figura 4.2 Sistema de corte de papel tipo sheeter CHM 1400.

Adicionalmente, la ausencia de controles estrictos ya definidos en este análisis sobre las condiciones ambientales, especialmente la humedad relativa, contribuye a que el papel experimente fenómeno de expansión o contracción antes del corte, amplificando aún más dicha variabilidad dimensional. Esta condición es particularmente relevante, ya que, el corte real del papel se comporta como variable dentro del sistema que impacta directamente el desempeño de alimentación de la prensa, provocando desalineaciones,

incrementos en la fricción durante el transporte y en consecuencia, una mayor probabilidad de atascos. Por lo tanto, se establece que la operación de la cortadora no solo influye en la calidad del producto intermedio, sino que constituye una fuente significativa de variabilidad dentro del sistema global de impresión, justificando su inclusión en el análisis del proceso y en el modelado estadístico como un factor clave en la explicación de la tasa de atascos observada.



Figura 4.3 Llegada del material a sitio.

En el área de almacenamiento observada, los sustratos ya cortados se encuentran apilados directamente sobre tarimas de madera y cubiertos con película plástica, sin ningún tipo de control ambiental. Esta práctica contradice las recomendaciones del fabricante y representa un riesgo crítico para el desempeño del material durante la impresión. Las fluctuaciones de temperatura y humedad pueden generar deformaciones, curvaturas o variaciones de espesor, así como disminuir la energía superficial necesaria para la adherencia del primer y de la tinta ElectroInk (HP, 2024).

Por tanto, la falta de control ambiental en el área de almacenaje (figura 4.4) constituye un problema grave que afecta la eficiencia operativa, eleva el desperdicio de sustratos y reduce la vida útil de consumibles críticos.



Figura 4.4 Almacenamiento actual.

Para el desarrollo de la presente investigación se emplearon tanto recursos físicos como herramientas computacionales especializadas, que permitieron llevar a cabo el análisis y la experimentación del proceso de impresión digital.

Como equipo principal, se utilizó una prensa digital HP Indigo serie 4, la cual constituye el sistema de estudio donde se evaluó el comportamiento del sustrato bajo diferentes condiciones operativas y ambientales. Esta prensa permitió generar datos reales relacionados con la producción, la ocurrencia de atascos y el desempeño general del proceso. Los datos fueron recopilados en la aplicación HP Remote@Glance la cual permite observar en tiempo real los datos recopilados por máquina, así como descargar base de datos.

En cuanto a las herramientas de análisis, se utilizó el software estadístico Rstudio el cual permitió analizar los datos observacionales, así como en ANOVA y modelado de regresión, con el fin de identificar variables.

Adicionalmente, se empleó el software R como herramienta complementaria para la simulación del experimento, particularmente en la generación de datos sintéticos asociados a la variabilidad del corte del sustrato. Esto permitió modelar escenarios no observados directamente y fortalecer la robustez del análisis experimental.

El uso conjunto de estos recursos permitió integrar datos reales y simulados dentro de un marco metodológico estructurado, facilitando el análisis de las variables críticas del proceso y la validación de los resultados obtenidos.

4.2 Datos para el experimento.

Con el propósito de incorporar al análisis una variable dimensional relevante asociada al desempeño del transporte del sustrato, se consideró la longitud de corte del papel, cuyo valor nominal corresponde a 750 mm de acuerdo con las especificaciones del proceso de conversión del material. Debido a que no se contaba con registros históricos exactos de la longitud real de corte para cada jornada de producción, esta variable fue estimada mediante una simulación con fines académicos, asignando valores dentro de un rango comprendido entre 748 mm y 752 mm, rango que representa una variación dimensional razonable considerando tolerancias típicas de corte en procesos industriales de papel.

La generación de estos valores se realizó bajo la hipótesis operativa de que la probabilidad de ocurrencia de atascos aumenta conforme la longitud real del sustrato se aleja del valor nominal de 750 mm, mientras que valores cercanos al objetivo tienden a generar un comportamiento más estable del proceso. Esta suposición se fundamenta en el principio de que desviaciones dimensionales del sustrato pueden afectar la sincronización del transporte de hojas, el registro longitudinal y la interacción con los mecanismos de alimentación de la prensa.

Para representar estas variaciones, los valores de longitud fueron asignados utilizando incrementos discretos de 0.5 mm dentro del intervalo establecido. El comportamiento del proceso se representó mediante la relación entre el error dimensional del corte y la tasa de atascos, considerando el error absoluto respecto al valor nominal, el cual puede expresarse mediante la siguiente Ecuación (4.1):

$$Error_{corte} = | L_{real} - L_{nominal} | \quad (4.1)$$

donde:

- L_{real} = longitud real del sustrato
- $L_{nominal}$ = longitud objetivo del proceso (750 mm)

Este valor permite cuantificar la desviación del corte respecto al objetivo y relacionarla con el comportamiento del sistema de transporte de papel.

Adicionalmente, considerando que el sustrato utilizado posee un gramaje nominal de 260 g/m², también se evaluó la variación equivalente de masa asociada a la variación de longitud del papel. Dado que el gramaje se define como la masa por unidad de área, la relación entre la longitud del sustrato y el gramaje equivalente puede expresarse mediante la siguiente Ecuación (4.2):

$$GSM_{equivalente} = GSM_{nominal} \times \frac{L_{real}}{L_{nominal}} \quad (4.2)$$

donde:

- $GSM_{nominal}$ = 260 g/m²
- L_{real} = longitud simulada del sustrato
- $L_{nominal}$ = 750 mm

Esta relación permite estimar la variación proporcional de masa asociada a los cambios dimensionales del corte.

Como ejemplo ilustrativo, si la longitud real del sustrato es 748 mm, el gramaje equivalente se calcula de la siguiente manera:

$$GSM_{equivalente} = 260 \times \frac{748}{750}$$

$$GSM_{equivalente} = 259.31 \text{ g/m}^2$$

Esto representa una variación aproximada de -0.69 g/m^2 , equivalente a una diferencia relativa de aproximadamente 0.27 % respecto al valor nominal. Este cambio resulta prácticamente despreciable en términos de masa del material, por lo que se considera que el incremento en la tasa de atascos se debe principalmente a la desviación dimensional del corte y su impacto en la dinámica de transporte del papel dentro de la prensa, más que a una variación significativa en el peso del sustrato.

En consecuencia, los valores de longitud de corte utilizados en el conjunto de datos fueron calculados con fines académicos para construir un modelo experimental que permitiera evaluar la influencia conjunta de las condiciones ambientales y de la precisión dimensional del sustrato sobre la tasa de atascos del sistema, facilitando así el desarrollo del análisis aplicado en el presente estudio.

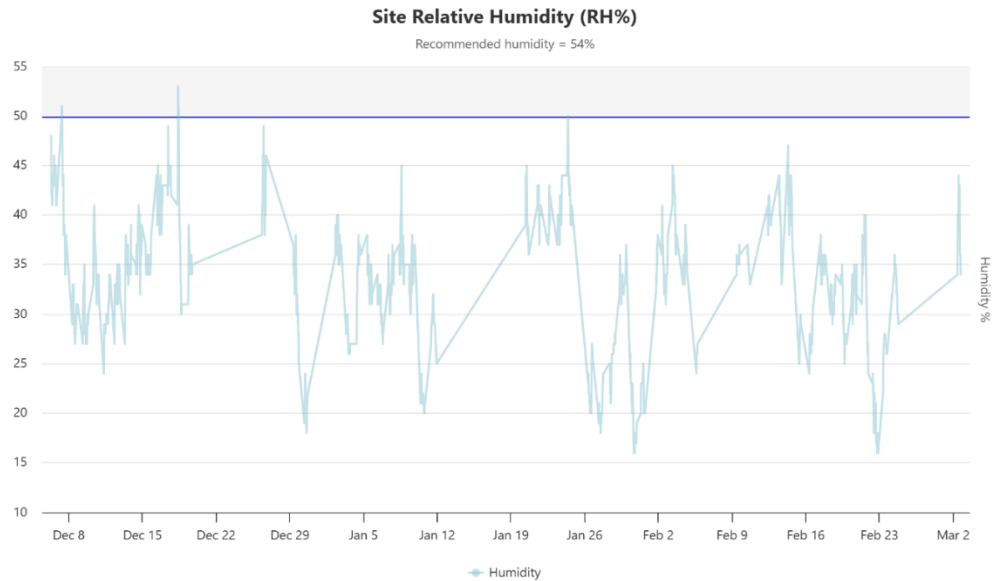


Figura 4.5 Humedad relativa del sitio (datos obtenidos de PrintOS).

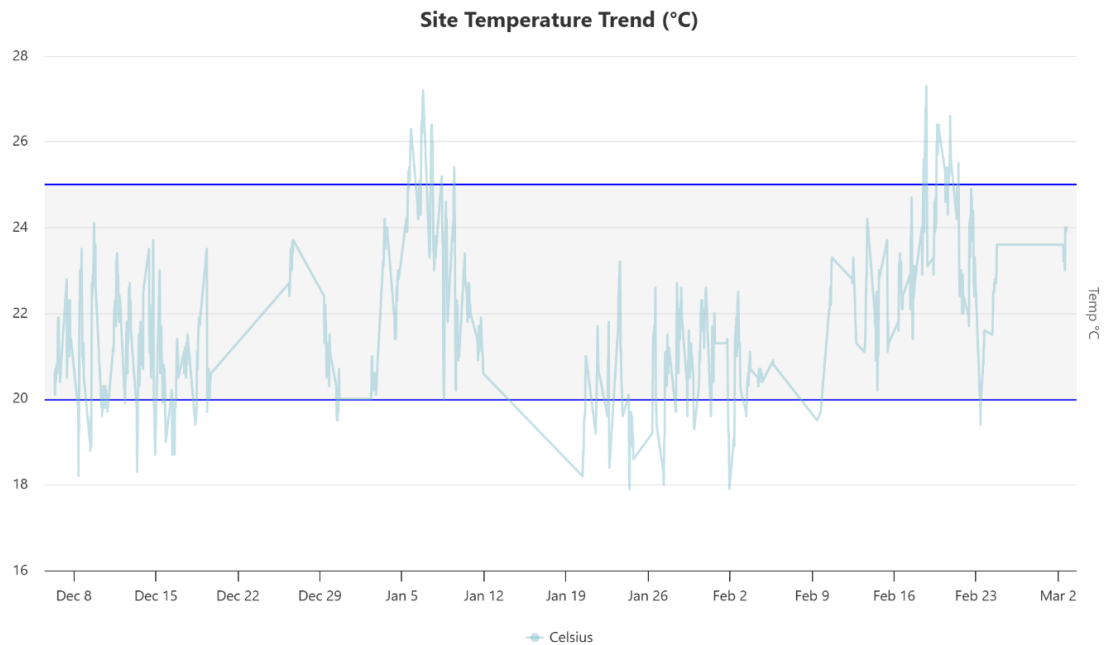


Figura 4.6 Temperatura del sitio (datos obtenidos de PrintOS).

4.3 Análisis de regresión y selección del modelo predictivo

Para analizar el comportamiento de la tasa de atascos en el proceso de impresión digital, se desarrollaron diferentes modelos de regresión en el software R. La variable respuesta del análisis fue definida como la tasa de atascos del sistema (jam rate). Como variables independientes se consideraron la temperatura ambiente, la humedad relativa y el corte real del sustrato.

La codificación de los factores quedó definida de la siguiente manera:

$$x_1 = \text{Temperatura } (^{\circ}\text{C})$$

$$x_2 = \text{Humedad relativa } (\%)$$

$$x_3 = \text{Corte real del sustrato } (mm)$$

El análisis inicio con la depuración de la base de datos, asegurando que las variables fueran numéricas y eliminando registros incompletos. Esto permitió trabajar con una base limpia y adecuada para el ajuste de modelos predictivos.

El primer modelo evaluado fue el modelo lineal:

$$m1 = \text{lm}(y \sim x1 + x2 + x3, \text{data} = \text{data})$$

Este modelo considera únicamente los efectos principales de cada factor sobre la tasa de atascos. Su forma general puede expresarse como la Ecuación (2.1):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon \quad (2.1)$$

El modelo lineal obtuvo un valor de $R^2 = 0.5651$, lo cual indica que explica aproximadamente el 56.51 % de la variabilidad observada en la tasa de atascos. Sin embargo, su R^2 predictivo fue de 0.4254, lo que muestra una capacidad limitada para predecir nuevos escenarios. Además, presentó el PRESS más alto, con un valor de 0.0006665, lo cual confirma que fue el modelo con menor capacidad predictiva dentro de los modelos comparados.

Posteriormente, se ajustó un modelo con interacciones:

$$m2 = \text{lm}(y \sim (x1 + x2 + x3)^2, \text{data} = \text{data})$$

Este modelo incluyo los efectos principales y las interacciones dobles entre los factores. Su inclusión es importante porque en el proceso real los atascos no dependen necesariamente de una sola variable aislada. Por ejemplo, el efecto del corte real del sustrato puede cambiar dependiendo de la humedad relativa del ambiente.

El modelo con interacciones mejoró el ajuste respecto al modelo lineal, alcanzando un $R^2 = 0.7421$ y un R^2 ajustado = 0.7013. Esto significa que el modelo explicó aproximadamente el 74.21 % de la variabilidad total, aun considerando el incremento en

el número de términos. También redujo el PRESS a 0.00050899 y aumentó el R^2 predictivo a 0.5612. Este resultado confirma que incluir interacciones mejora la representación del sistema.

Después se evaluó un modelo cuadrático o de segundo orden:

$$m3 = \text{lm} (y \sim x1 + x2 + x3 + x1:x2 + x1:x3 + x2:x3 + I(x1^2) + I(x2^2) + I(x3^2), \text{data} = \text{data})$$

Este modelo permitió analizar no solo los efectos lineales e interacciones, sino también posibles curvas en el comportamiento del sistema. Ver ecuación (2.2):

$$y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ij} x_i x_j + \sum \beta_{ii} x_i^2 + \varepsilon$$

El modelo cuadrático presentó un $R^2 = 0.9094$, lo que indica que logró explicar aproximadamente el 90.94 % de la variabilidad de la tasa de atascos. También obtuvo un R^2 ajustado = 0.8860, un PRESS = 0.00026488 y un R^2 predictivo = 0.7716. Estos resultados muestran una mejora importante respecto a los modelos anteriores, lo que sugiere que el comportamiento del sistema no es completamente lineal.

Finalmente, se aplicó selección automática mediante el criterio AIC al modelo cuadrático:

$$m3_step = \text{stepAIC} (m3, \text{scope} = \text{list(lower} = m0, \text{upper} = m3), \text{direction} = \text{"both"}, \text{trace} = \text{TRUE})$$

El modelo cuadrático optimizado fue el mejor modelo seleccionado. Su fórmula final fue la Ecuación (4.3):

$$y = \beta_0 + \beta_2 x_2 + \beta_3 x_3 + \beta_{33} x_3^2 + \beta_{23} x_2 x_3 + \varepsilon \quad (4.3)$$

El Anova del modelo seleccionado confirma que todos los términos conservados tienen un efecto estadísticamente significativo sobre la tasa de atascos y . En primer lugar, la humedad relativa x_2 presenta un valor F de 18.187 y un valor-p de 0.0001188, por lo

que su efecto es significativo. Esto indica que los cambios en la humedad del ambiente influyen en el comportamiento del sustrato y, por consecuencia, en la probabilidad de generar atascos durante la impresión.

El factor con mayor influencia dentro del modelo es el corte real del sustrato x_3 , ya que presenta la mayor suma de cuadrados, con 0.00058012, y el valor F más alto, con 195.062. Además, su valor-p es menor a $2.2e-16$, lo que confirma una significancia estadística muy alta. Este resultado demuestra que la dimensión real del corte es el factor principal para explicar la variación en la tasa de atascos. En términos prácticos, pequeñas desviaciones en el tamaño del sustrato pueden afectar su alineación, alimentación y transporte dentro de la prensa, incrementando el riesgo de fallas.

El término cuadrático del corte real también resulta altamente significativo, con un valor F de 123.954 y un valor-p de $7.988e-14$. Esto representa que el efecto del corte real no es únicamente lineal, sino que presenta curvatura. Por lo tanto, la tasa de atascos no cambia de forma constante conforme varía el corte, sino que puede aumentar o disminuir de manera más marcada dependiendo de que tan alejado se encuentra el sustrato de una zona dimensional óptima.

Finalmente, la interacción entre humedad relativa y corte real $x_2:x_3$ presenta un valor F de 12.808 y un valor-p de 0.0009228, por lo que también es significativa. Esto indica que el efecto del corte real sobre la tasa de atascos depende de la humedad relativa presente en el ambiente. Esto quiere decir que un mismo valor de corte puede comportarse de manera diferente bajo distintas condiciones de humedad, por lo que refuerza la necesidad de controlar ambas variables de forma conjunta.

En conjunto, el Anova permite concluir que el modelo no solo identifica factores significativos, sino que también evidencia que el comportamiento de los atascos está dominado principalmente por el corte real del sustrato, seguido por su efecto cuadrático y por la influencia de la humedad relativa. Esto valida el uso del modelo cuadrático

optimizado como herramienta predictiva para la planta virtual, ya que representa adecuadamente la relación entre las condiciones del proceso y la tasa esperada de atascos.

Tabla 4.2 Anova del modelo seleccionado.

Factor	GL (Df)	Suma de cuadrados (Sum Sq)	Cuadrado medio (Mean Sq)	Valor F	Valor-p
Humedad (x_2)	1	5.41E-05	5.41E-05	18.187	0.000119
Corte (x_3)	1	0.00058	0.00058	195.062	< 2.2e-16
Corte ² (x_3^2)	1	0.000369	0.000369	123.954	7.99e-14
$x_2:x_3$	1	3.81E-05	3.81E-05	12.808	0.000923
Error	40	0.000119	2.97E-06	—	—

Sustituyendo los factores reales:

$$\text{Tasa de atasco} = \beta_0 + \beta_2(\text{RH}) + \beta_3(\text{Corte real}) + \beta_{33}(\text{Corte real})^2 + \beta_{23}(\text{RH})(\text{Corte real}) + \varepsilon$$

Este resultado es relevante porque el modelo final ya no conserva la temperatura como término principal. Esto no significa necesariamente que la temperatura no tenga ningún efecto físico sobre el sustrato, sino que, dentro de los datos analizados, su contribución estadística fue menor en comparación con la humedad relativa, el corte real y la interacción entre ambos factores. Desde el punto de vista del proceso, la selección de este modelo indica que la tasa de atascos está principalmente relacionada con la humedad relativa, el corte real del sustrato, la curvatura del corte real y la interacción entre humedad y corte. Esto sugiere que controlar únicamente el valor promedio del corte no es suficiente; también es necesario considerar las condiciones ambientales bajo las cuales el papel es almacenado, preparado y alimentado a la prensa.

En términos prácticos, el modelo permite representar el sistema como una planta virtual, ya que puede recibir nuevas combinaciones de humedad relativa y corte real para estimar la tasa esperada de atascos. De esta forma, el modelo se convierte en una

herramienta predictiva para evaluar escenarios de operación, identificar condiciones de mayor riesgo y establecer recomendaciones para reducir la variabilidad del proceso y mejorar la estabilidad productiva de la prensa HP Indigo.

La comparación final de modelos fue la siguiente:

Tabla 4.3 Comparación final de modelos.

Modelo	R²	R² ajustado	AIC	PRESS	R² predictivo
Lineal	0.5650807	0.5332574	-375.2359	0.0006665	0.4253868
Interacciones	0.7420745	0.7013494	-392.7479	0.00050899	0.5611775
Interacciones optimizado	0.7420745	0.7013494	-392.7479	0.00050899	0.5611775
Cuadrático	0.9093533	0.8860441	-433.8045	0.00026488	0.7716377
Cuadrático optimizado	0.8974387	0.8871825	-438.2474	0.00020484	0.8234028

4.3.1 Los supuestos.

Las gráficas de diagnóstico del modelo cuadrático optimizado permiten evaluar el cumplimiento de los supuestos estadísticos fundamentales y validar la calidad del ajuste obtenido para la predicción de la tasa de atascos del sistema de impresión digital. La figura 4.7 Residuals vs Fitted muestra que los residuos se distribuyen de manera relativamente aleatoria alrededor del cero, indicando que el modelo representa adecuadamente la relación existente entre las variables independientes y la respuesta del sistema. La ausencia de patrones claramente definidos o tendencias pronunciadas sugiere que el modelo no presenta problemas severos de no linealidad y que la estructura cuadrática seleccionada logra capturar el comportamiento principal del proceso. Aunque existen algunas observaciones con residuos ligeramente mayores, estas pueden asociarse a la variabilidad natural presente en ambientes industriales reales, donde factores operacionales y ambientales generan fluctuaciones difíciles de controlar completamente.

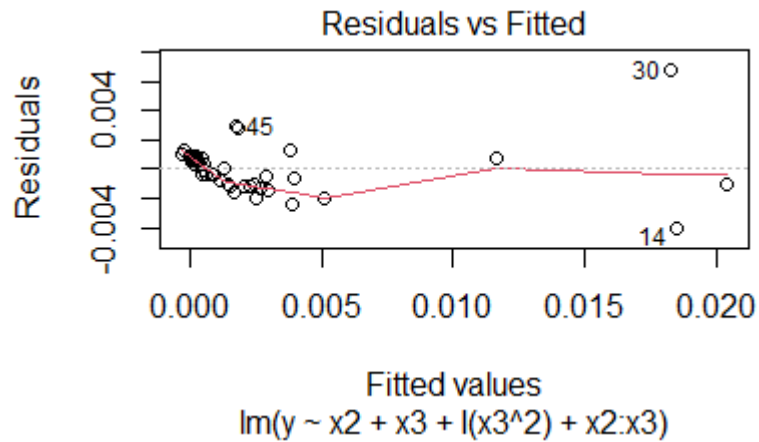


Figura 4.7 Residuals vs Fitted.

Por otra parte, la figura 4.8 Q-Q Residuals evidencia que la mayoría de los residuos siguen una distribución aproximadamente normal, ya que los puntos se alinean de manera cercana a la línea de referencia teórica. Este resultado es importante debido a que confirma que el modelo cumple razonablemente con el supuesto de normalidad requerido para la validez de pruebas estadísticas como ANOVA y el análisis de significancia de los coeficientes. Las ligeras desviaciones observadas en los extremos de la gráfica indican la presencia de algunos escenarios atípicos o extremos dentro del proceso productivo, lo cual resulta coherente con el comportamiento dinámico y variable característico de los sistemas de impresión digital.

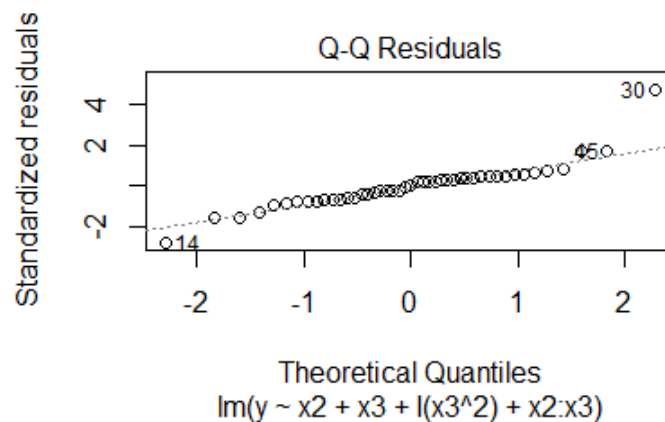


Figura 4.8 Q-Q Residuals.

Asimismo, la figura 4.9 Scale-Location indica que la dispersión de los residuos permanece relativamente estable a lo largo de los valores ajustados, aunque se aprecia un ligero incremento de variabilidad conforme aumentan las predicciones del modelo. Este comportamiento sugiere la presencia de una leve heterocedasticidad, lo cual implica que bajo ciertas condiciones operacionales el error del modelo puede incrementarse ligeramente. Sin embargo, esta variación no resulta suficientemente severa para comprometer la capacidad predictiva del modelo y puede considerarse aceptable dentro de un entorno industrial donde las condiciones de operación presentan incertidumbre y variabilidad inherente.

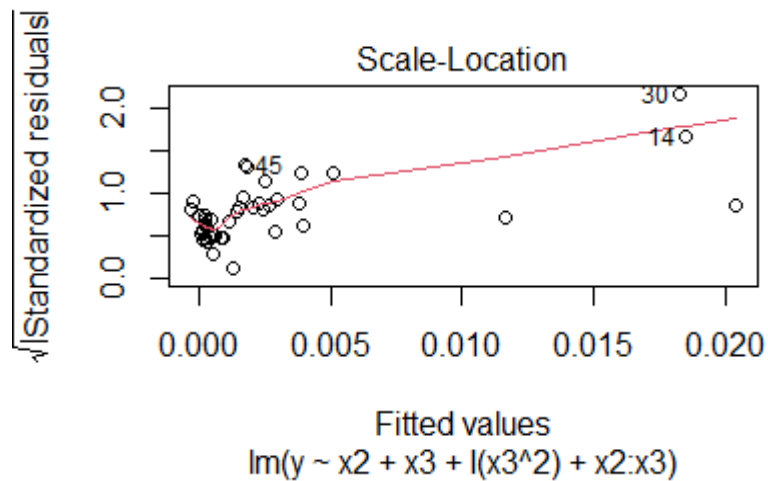


Figura 4.9 Scale-Location

Finalmente, la figura 4.10 Residuals vs Leverage muestra que la mayoría de las observaciones poseen una influencia moderada sobre el modelo y permanecen dentro de límites aceptables de distancia de Cook. Aunque algunas observaciones presentan valores de leverage relativamente altos, ninguna de ellas demuestra una influencia crítica capaz de alterar significativamente la estabilidad del ajuste estadístico. Esto indica que el modelo no depende excesivamente de una sola observación experimental y que los coeficientes obtenidos presentan robustez frente a la variabilidad de los datos analizados.

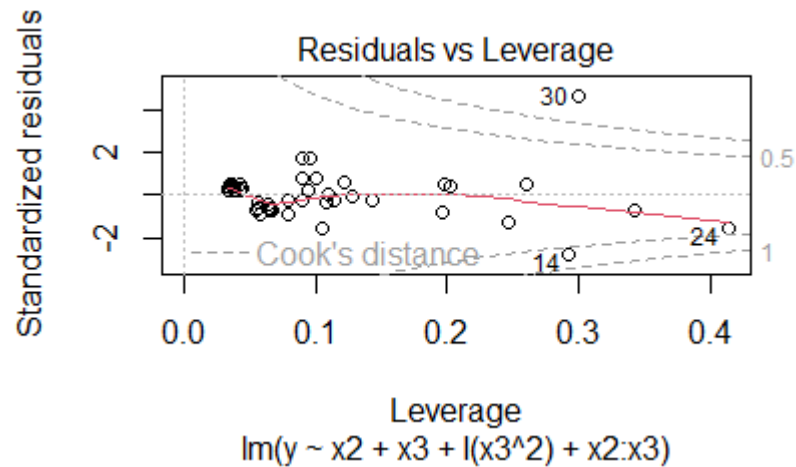


Figura 4.10 Residuals vs Leverage

4.4 Planta Virtual

Como resultado del análisis estadístico realizado, se seleccionó un modelo cuadrático optimizado como la mejor representación del comportamiento del sistema de impresión digital. Este modelo fue elegido con base en su alto poder explicativo y predictivo, considerando métricas como el R^2 ajustado, el criterio AIC y el R^2 predictivo, lo que permitió asegurar un adecuado balance entre ajuste y capacidad de generalización.

Los coeficientes del modelo fueron obtenidos a partir del análisis estadístico desarrollado previamente en la etapa de modelado del proceso, utilizando técnicas de regresión lineal y de segundo orden en R.

Tabla 4. 4 coeficientes de los términos

Término	Estimate
(Intercept)	3807
X ₂	0.1949
X ₃	-10.17
X ₃ ²	0.006786
X ₂ :X ₃	-0.0002599

La Ecuación (4.4) obtenida para predecir la tasa de atascos (y) es la siguiente:

$$y = 3807 + 0.1949x_2 - 10.17x_3 + 0.006786x_3^2 - 0.0002599x_2x_3 \quad (4.4)$$

donde x_2 corresponde a la humedad relativa y x_3 al corte real del sustrato.

Este modelo evidencia que la tasa de atascos depende principalmente de la interacción entre las condiciones ambientales y la precisión del corte del papel. En primer lugar, el término asociado a la humedad relativa presenta un coeficiente positivo, lo que indica que incrementos en la humedad tienden a aumentar la probabilidad de atascos. Este comportamiento es consistente con la naturaleza del sustrato, ya que la humedad influye directamente en propiedades físicas del papel como su rigidez, absorción y fricción, afectando su desempeño durante el transporte dentro de la prensa.

Por otra parte, el corte real del sustrato muestra un comportamiento no lineal, representado por la presencia de un término cuadrático. Aunque el coeficiente lineal sugiere una disminución inicial en la tasa de atascos al incrementar el valor del corte, el término cuadrático positivo evidencia la existencia de una curvatura en forma de “U”. Esto implica que existe una región óptima de operación en la cual los atascos se minimizan, mientras que desviaciones por encima o por debajo de dicho punto incrementan significativamente la probabilidad de falla. Este resultado es coherente con el comportamiento físico del proceso, donde variaciones en el corte afectan la alineación, alimentación y registro del papel.

Adicionalmente, el modelo incorpora un término de interacción entre la humedad y el corte, lo cual confirma que ambos factores no actúan de manera independiente. El coeficiente negativo asociado a esta interacción indica que el efecto de una variable depende del nivel de la otra. En términos operativos, esto significa que un mismo valor de corte puede ser aceptable en condiciones de baja humedad, pero puede generar atascos bajo condiciones de alta humedad, evidenciando un acoplamiento significativo entre ambas variables.

En conjunto, los resultados del modelo permiten concluir que el corte real del sustrato es el factor dominante en el comportamiento del sistema, mientras que la humedad actúa como un factor que amplifica o modifica dicho efecto. Este hallazgo tiene implicaciones directas en la operación del proceso, ya que demuestra que no es suficiente controlar las variables de manera individual. En cambio, es necesario establecer condiciones de operación conjuntas que permitan mantener el sistema dentro de una región óptima, donde la ocurrencia de atascos sea mínima.

Estos resultados constituyen la base para el desarrollo de la planta virtual predictiva, ya que la ecuación obtenida permite simular el comportamiento del sistema bajo diferentes escenarios operativos, facilitando la identificación de condiciones robustas sin necesidad de realizar pruebas físicas en planta.

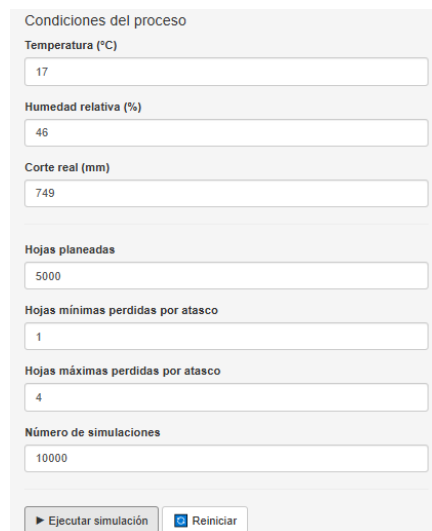
La planta virtual predictiva funciona como una herramienta de simulación desarrollada en R Shiny, cuyo objetivo es representar de manera simplificada el comportamiento del proceso de impresión bajo diferentes condiciones operativas. Esta aplicación permite estimar cómo variables críticas del proceso, como la temperatura, la humedad relativa y el corte real del sustrato, pueden influir en la tasa de atascos o jam rate, y posteriormente en la producción real esperada.

El funcionamiento inicia con la carga de la base de datos original y del modelo predictivo previamente seleccionado. Este modelo representa la relación estadística entre las condiciones del proceso y la variable de respuesta, en este caso el jam rate. Además, se extrae el error residual del modelo mediante, el cual es importante porque indica la variabilidad que el modelo no logra explicar completamente. Dicho error se utiliza después para incorporar incertidumbre en la simulación.

La interfaz de usuario permite ingresar manualmente las condiciones del proceso. El usuario puede modificar la temperatura, la humedad relativa y el corte real del papel, así como definir el volumen de hojas planeadas para producción. También se agregan dos parámetros importantes: la cantidad mínima y máxima de hojas perdidas por cada atasco.

Esto permite representar una situación más realista, ya que no todos los atascos generan la misma pérdida; en algunos casos puede perderse una sola hoja, mientras que en otros pueden perderse hasta cuatro hojas o más, dependiendo del tipo de atasco, el momento en que ocurre y la intervención necesaria para corregirlo.

La simulación se ejecuta cuando el usuario presiona el botón “Ejecutar simulación”. Primero, la aplicación toma los valores ingresados y construye un nuevo conjunto de datos con las variables x_1 , x_2 y x_3 . Posteriormente, el modelo predictivo calcula una predicción base del jam rate. Sin embargo, en lugar de utilizar únicamente ese valor como un resultado fijo, la planta virtual genera miles de escenarios posibles mediante simulación Monte Carlo.



The image shows a web-based user interface for a virtual plant simulation. It is titled "Condiciones del proceso" (Process Conditions). The interface contains several input fields with the following labels and values:

- Temperatura (°C): 17
- Humedad relativa (%): 46
- Corte real (mm): 749
- Hojas planeadas: 5000
- Hojas mínimas perdidas por atasco: 1
- Hojas máximas perdidas por atasco: 4
- Número de simulaciones: 10000

At the bottom of the form, there are two buttons: "▶ Ejecutar simulación" (Execute simulation) and "🔄 Reiniciar" (Reset).

Figura 4.11 Interfaz de usuario de la planta virtual.

La implementación de Monte Carlo fue necesaria porque el proceso real de impresión no se comporta de forma completamente determinista. Aunque el modelo estadístico entregue una predicción puntual, en la práctica existen fuentes de variabilidad que no pueden controlarse totalmente, como pequeñas diferencias entre lotes de papel, variaciones ambientales, condiciones de almacenamiento, estado mecánico de la máquina, experiencia del operador o cambios menores en el manejo del sustrato. Por esta razón, usar un solo valor de predicción podría dar una visión limitada del proceso. Mediante

Monte Carlo, el sistema genera una distribución de posibles valores de jam rate alrededor de la predicción base, usando una distribución normal con media igual al valor predicho por el modelo y desviación estándar igual al error residual. Después, los valores negativos se corrigen a cero ya que físicamente no tendría sentido tener una tasa de atascos negativa.

Después de simular el jam rate, la planta calcula la cantidad estimada de atascos multiplicando el volumen planeado de hojas por cada valor simulado de la tasa de atasco. Luego se incorpora la variabilidad en la cantidad de hojas perdidas por atasco, usando una distribución uniforme entre el mínimo y máximo definido por el usuario. Esta parte es relevante porque convierte el modelo en una representación más cercana a la operación real: no solo estima cuántos atascos podrían ocurrir, sino también cuánto impacto podría tener sobre la producción final.

Con estos valores, el sistema calcula las hojas perdidas, la producción real simulada y la eficiencia del proceso. La producción real se obtiene restando las hojas perdidas al volumen planeado, mientras que la eficiencia se calcula como el porcentaje de producción efectiva respecto a la producción planeada.

Los resultados mostrados por la aplicación incluyen el jam rate promedio, así como los percentiles P5 y P95. Estos percentiles permiten interpretar escenarios pesimistas y optimistas tanto de los atascos esperados como la cantidad de hojas esperadas. Por ejemplo, el P5 de producción representa un escenario desfavorable, donde solo el 5% de las simulaciones presentó una producción menor a ese valor. En cambio, el P95 representa un escenario favorable, donde la mayoría de las simulaciones queda por debajo de ese nivel de producción.

```
Jam Rate promedio: 0.004697
Jam Rate P5: 0.001897
Jam Rate P95: 0.007532

Atascos esperados (promedio): 23.5
Atascos pesimista (P95): 37.7
Atascos optimista (P5): 9.5

Producción promedio: 4941 hojas
Producción pesimista (P5): 4883
Producción optimista (P95): 4982

Eficiencia promedio: 98.83 %
```

Figura 4.12 Ejemplo de resultados de simulación en planta virtual.

La figura 4.13 gráfica de barras, compara la producción planeada contra tres escenarios simulados: promedio, pesimista y optimista. Esto facilita visualizar la diferencia entre lo que se espera producir idealmente y lo que realmente podría obtenerse considerando atascos y pérdidas. Por otro lado, la figura 4.10 histograma, muestra la distribución de la producción real simulada, permitiendo observar si los resultados están concentrados cerca del promedio o si existe una alta dispersión.

En conjunto, esta planta virtual permite evaluar el comportamiento probable del sistema antes de realizar cambios físicos en la producción. Su principal ventaja es que transforma el modelo predictivo en una herramienta práctica de análisis y toma de decisiones, ya que permite probar diferentes condiciones de temperatura, humedad, corte y volumen de producción sin necesidad de detener la máquina o generar desperdicio real. La simulación Monte Carlo fortalece el análisis porque no entrega un único resultado, sino un rango de posibles comportamientos del proceso, lo cual es más adecuado para representar la incertidumbre natural de un ambiente industrial.

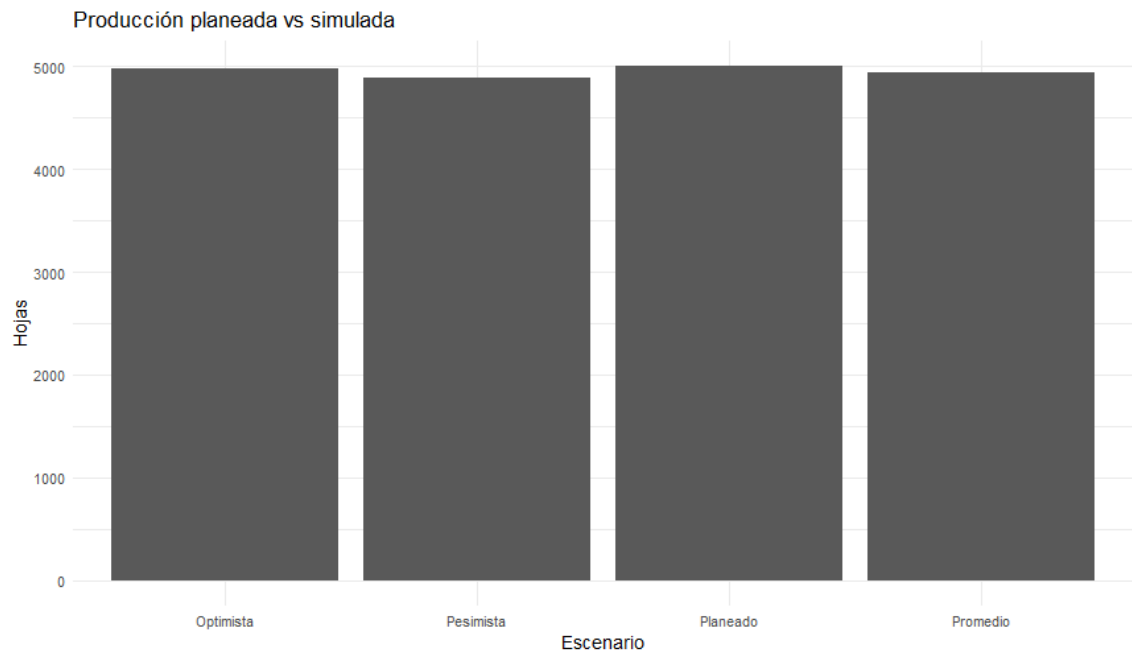


Figura 4.13 Grafica de barras de ejemplo de simulación.

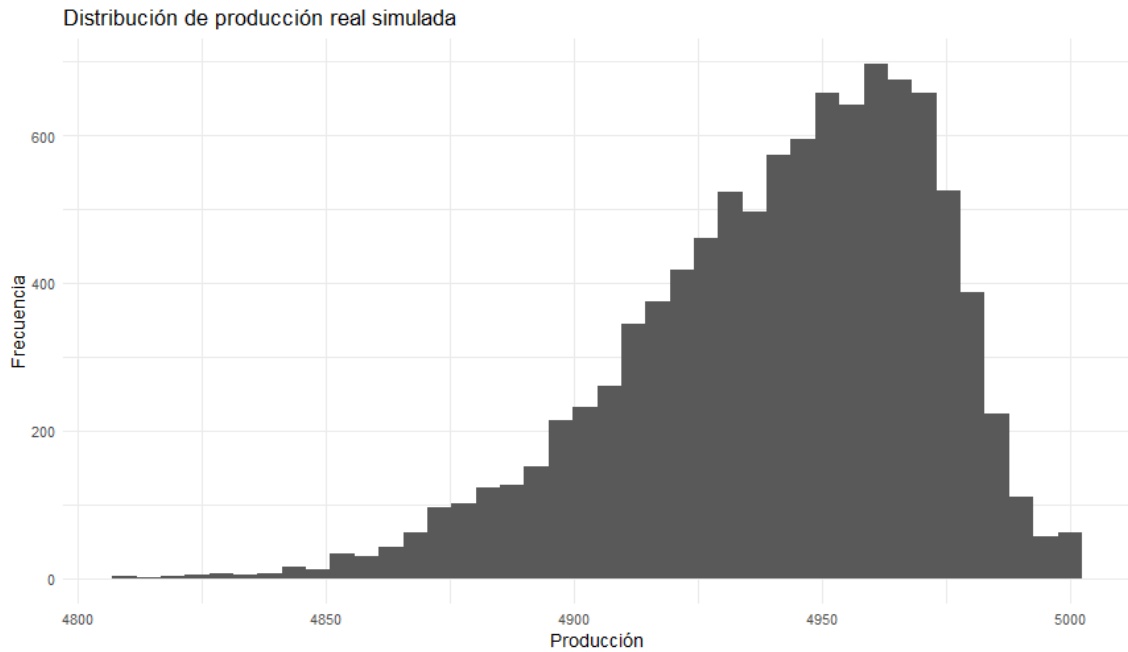


Figura 4.14 Histograma de ejemplo de simulación.

4.4.1 Validación de la planta virtual.

Para validar el proyecto desarrollado, se realizó un ejercicio de comparación

utilizando un día real de producción correspondiente a la misma máquina donde fue efectuado el estudio experimental. El propósito de esta validación consistió en contrastar la cantidad de atascos registrados durante la operación real contra los resultados obtenidos mediante la simulación de la planta virtual predictiva. A partir de este procedimiento fue posible evaluar la capacidad del modelo cuadrático optimizado y de la simulación basada en el método de Monte Carlo para representar el comportamiento real del sistema y generar predicciones confiables bajo condiciones futuras de operación.

El día seleccionado para la validación correspondió al 3 de mayo, fecha en la cual se registró la mayor cantidad de atascos dentro del periodo de 90 días posteriores a la última observación utilizada en la construcción del modelo histórico, ver figura 4.11. Durante dicha jornada, las condiciones ambientales registradas fueron de 19.9 °C de temperatura y 40.4 % de humedad relativa. Estas variables fueron utilizadas como parámetros de entrada en la planta virtual para reproducir las condiciones reales del proceso y analizar el comportamiento esperado del sistema de impresión.

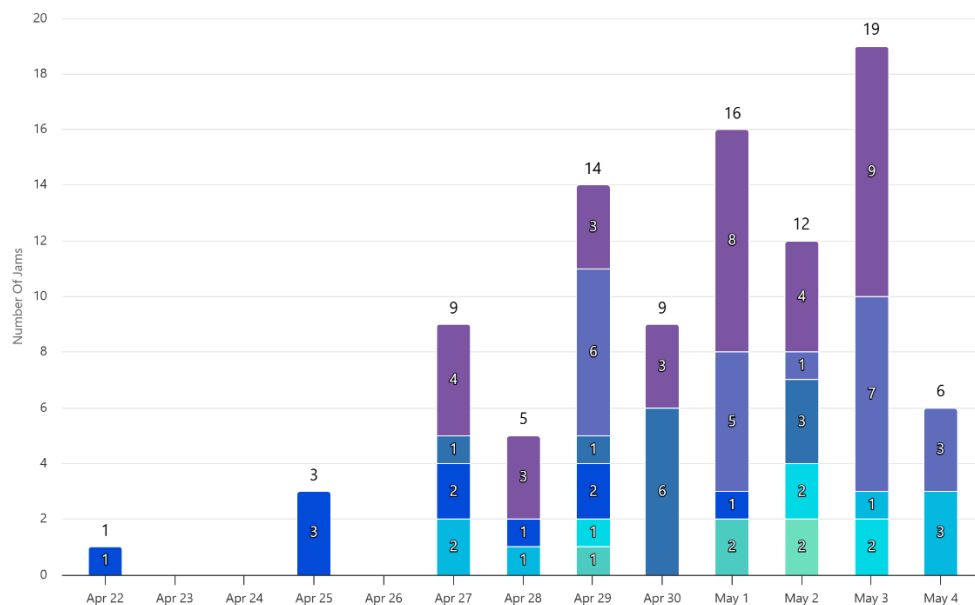
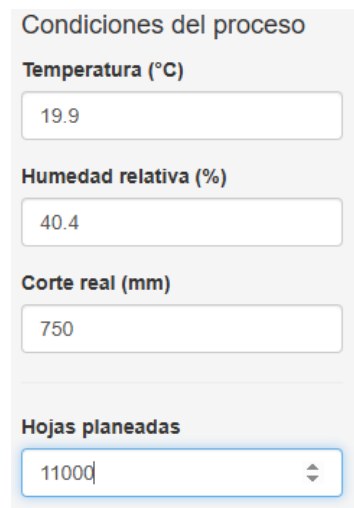


Figura 4.15 Atascos registrados en Abril-Mayo 2026 (datos obtenidos de PrintOS).

Como parte fundamental de la validación de la hipótesis y de los objetivos planteados en la investigación, el factor x_3 correspondiente al corte real del sustrato fue tratado como una variable de control y no como un factor de ruido. Esta decisión se fundamenta en los resultados obtenidos mediante el análisis ANOVA del modelo cuadrático optimizado, donde dicho factor presentó la mayor significancia estadística sobre la tasa de atascos del sistema. Lo anterior demuestra que pequeñas variaciones dimensionales en el corte del sustrato tienen un impacto directo sobre la estabilidad del proceso de alimentación y transporte del papel dentro de la prensa digital.

Este comportamiento coincide con lo descrito en la documentación técnica de sustratos de HP, donde se establece que el papel debe mantener un corte nominal de 750 mm y presentar bordes uniformes y libres de rebabas, debido a que el sistema mecánico de alimentación de la prensa se encuentra sincronizado con base en dichas dimensiones. Aunque el equipo posee cierta capacidad de compensación ante pequeñas desviaciones dimensionales, cuando la variación excede la capacidad de ajuste dinámico del sistema pueden generarse errores de alimentación y, en consecuencia, nuevos atascos durante la producción.



The image shows a software interface for setting simulation parameters. It is titled 'Condiciones del proceso' and contains four input fields with the following values: 'Temperatura (°C)' is 19.9, 'Humedad relativa (%)' is 40.4, 'Corte real (mm)' is 750, and 'Hojas planeadas' is 11000. The last field is a dropdown menu.

Condiciones del proceso
Temperatura (°C)
19.9
Humedad relativa (%)
40.4
Corte real (mm)
750
Hojas planeadas
11000

Figura 4.16 Entradas de la simulación.

A partir de esta validación se comprobó que transformar el factor más significativo de ruido en un factor de control permite reducir considerablemente la ocurrencia de atascos y maximizar la producción efectiva del sistema. De igual forma, los resultados obtenidos demostraron que la integración del modelo cuadrático optimizado con la simulación Monte Carlo constituye una herramienta viable para la predicción del desempeño operativo de la prensa digital bajo distintas condiciones de trabajo, permitiendo anticipar escenarios críticos y establecer estrategias de control orientadas a mejorar la estabilidad y eficiencia del proceso de impresión.

Los datos históricos correspondientes al día 3 de mayo indican que durante la operación real de la prensa se registraron 17 atascos bajo las condiciones ambientales previamente descritas, correspondientes a 19.9 °C de temperatura y 40.4 % de humedad relativa. Estos resultados representan uno de los escenarios de mayor inestabilidad observados dentro del periodo de validación posterior a la construcción del modelo estadístico.

```
=== RESULTADOS ===  
  
Jam Rate promedio: 0.000945  
Jam Rate P5: 0  
Jam Rate P95: 0.003312  
  
Atascos esperados (promedio): 10.4  
Atascos pesimista (P95): 36.4  
Atascos optimista (P5): 0  
  
Producción promedio: 10974 hojas  
Producción pesimista (P5): 10900  
Producción optimista (P95): 11000  
  
Eficiencia promedio: 99.76 %
```

Figura 4.17 Resultados de la simulación.

Una vez ejecutada la simulación en la planta virtual predictiva ver figura 4.13 utilizando las mismas condiciones operativas y estableciendo el corte real del sustrato en 750 mm como factor de control, se obtuvo una predicción promedio de 10.4 atascos para un volumen planeado de 11,000 hojas. Este resultado representa una reducción aproximada del 38.8 % respecto al comportamiento histórico observado en la operación real, lo cual evidencia el impacto significativo que tiene el control dimensional del sustrato sobre la estabilidad del proceso de impresión.

Adicionalmente, la simulación Monte Carlo permitió generar distintos escenarios probabilísticos de comportamiento del sistema, proporcionando una visión más realista de la variabilidad operacional existente dentro del proceso. El escenario optimista, correspondiente al percentil P5, estimó la posibilidad de operar sin presencia de atascos, mientras que el escenario pesimista, asociado al percentil P95, proyectó hasta 36.4 atascos bajo condiciones desfavorables del sistema. La importancia de estos escenarios radica en que permiten comprender que el proceso productivo no opera bajo condiciones completamente determinísticas, sino que existe una variabilidad inherente asociada a factores ambientales, mecánicos y operacionales que pueden modificar significativamente el desempeño diario de la prensa.

La incorporación de escenarios optimistas y pesimistas dentro de la simulación resultó fundamental para validar la robustez del modelo predictivo, ya que permitió evaluar no solamente un valor promedio esperado, sino también los posibles límites operacionales del sistema bajo diferentes niveles de incertidumbre. Desde un enfoque de ingeniería industrial, este análisis probabilístico proporciona una herramienta de gran valor para la toma de decisiones, debido a que permite anticipar riesgos operativos, estimar pérdidas potenciales y establecer estrategias preventivas orientadas a minimizar tiempos muertos y desperdicio de material.

De igual forma, la simulación predijo una producción promedio de 10,974 hojas útiles de un total planeado de 11,000 hojas, obteniendo una eficiencia operacional estimada de 99.76 %. En el escenario optimista, la producción alcanzó prácticamente la totalidad de las hojas planeadas, mientras que en el escenario pesimista se observó una reducción asociada al incremento de atascos y pérdidas derivadas de interrupciones operacionales. Esto demuestra que la ocurrencia de atascos impacta directamente sobre la productividad y disponibilidad del sistema de impresión digital.

A partir de estos resultados puede concluirse que el modelo cuadrático optimizado, combinado con la simulación Monte Carlo, posee una capacidad adecuada para representar el comportamiento real del sistema de impresión y evaluar escenarios futuros de operación. Asimismo, la validación confirma la hipótesis principal de la investigación: el control del corte real del sustrato, identificado como el factor más significativo mediante el análisis ANOVA, permite reducir considerablemente la ocurrencia de atascos y mejorar la estabilidad del proceso productivo.

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

A partir del análisis realizado, se evidenció que las variables críticas del proceso, particularmente la humedad relativa y la precisión de corte del sustrato, tienen un efecto significativo sobre la tasa de atascos y, en consecuencia, sobre la producción efectiva del sistema. Esto permitió transformar un problema tradicionalmente abordado de forma reactiva en un problema cuantificable y predecible, lo cual representó un cambio relevante en la forma de analizar el desempeño del proceso.

En términos de aportaciones, el desarrollo del modelo predictivo constituyó una herramienta clave para estimar la producción esperada bajo distintas condiciones operativas. Sin embargo, uno de los hallazgos más importantes fue que dicho modelo no solo funcionó como un mecanismo de predicción, sino también como una herramienta de diagnóstico. En particular, al comparar la producción real contra la producción estimada, se identificaron escenarios donde la producción observada resultó considerablemente menor a la esperada. Estas discrepancias se interpretaron como indicadores de desviaciones en el proceso, lo cual permitió establecer un criterio objetivo para detonar investigaciones técnicas orientadas a la identificación de causas raíz.

En este sentido, el estudio sirvió como punto de partida para un enfoque más amplio de monitoreo del proceso, en el cual la planta virtual puede utilizarse como un sistema de referencia para evaluar el desempeño real. Cuando el sistema no cumple con el volumen de producción esperado, el modelo puede generar una alerta que indique la necesidad de analizar condiciones no consideradas o fallas operativas, fortaleciendo así la toma de decisiones basada en datos.

Adicionalmente, se observó que la implementación de esta metodología permitió visibilizar de manera clara la magnitud del problema de atascos y su impacto en la producción. Este aspecto resulta particularmente relevante, ya que evidenció que es posible identificar oportunidades de mejora sin necesidad de incrementar costos operativos ni el consumo de materiales, sino mediante una mejor comprensión y control de las variables existentes. En este contexto, el uso de plantas virtuales facilitó la

evaluación de escenarios y la identificación de condiciones subóptimas del proceso sin intervenir directamente en la operación real, lo que representó una ventaja significativa en términos de reducción de riesgos y costos.

En relación con la hipótesis planteada, se comprobó que es posible desarrollar un modelo predictivo capaz de estimar la tasa de atascos y la producción del sistema en función de variables críticas del proceso. Asimismo, se confirmó que la utilización de este modelo contribuye a reducir la incertidumbre en la toma de decisiones y proporciona una base cuantitativa para la mejora del proceso.

No obstante, se identificaron algunas etapas que quedaron inconclusas al momento de la entrega de este trabajo, particularmente aquellas relacionadas con la ampliación del modelo hacia un mayor número de variables y su validación en diferentes periodos de operación. Estas limitaciones estuvieron asociadas principalmente a los alcances académicos del estudio, los cuales requirieron delimitar el análisis a un conjunto específico de variables para garantizar rigor metodológico.

Derivado de lo anterior, se plantearon diversas líneas de trabajo a futuro. En primer lugar, se propone la replicación del estudio utilizando nuevos conjuntos de datos históricos, con el objetivo de validar la robustez del modelo y fortalecer su capacidad predictiva. Esta replicabilidad permitiría consolidar el uso de la planta virtual como una herramienta de monitoreo continuo del proceso.

Asimismo, se recomienda ampliar el modelo mediante la incorporación de variables adicionales que potencialmente influyen en el desempeño del sistema, tales como el tiempo de reposo del sustrato, el uso y condiciones de aplicación de primer, los diferentes tipos y gramajes de sustrato, así como las condiciones de almacenamiento, manejo del material y mantenimiento del equipo. La inclusión de estas variables permitiría desarrollar modelos más completos y representativos del sistema real.

Finalmente, se propone la evolución de la metodología hacia sistemas más avanzados que integren el uso de datos en tiempo real y análisis automatizado, permitiendo que la planta virtual funcione como un sistema activo de apoyo a la toma de decisiones. De esta manera, el presente estudio no solo cumple con los objetivos planteados, sino que también sentó las bases para futuras investigaciones orientadas a la optimización, control y monitoreo inteligente de procesos industriales mediante enfoques predictivos.

CAPÍTULO VI. REFERENCIAS

- Taguchi, G. (1986). *Introduction to quality engineering: Designing quality into products and processes*. Asian Productivity Organization.
- Montgomery, D. C. (2017). *Design and analysis of experiments* (9th ed.). John Wiley & Sons.
- Franceschini, F., & Lazzarotti, V. (2008). *Design of experiments for industrial applications*.
- Gutiérrez, H. (2007). *Análisis y diseño de experimentos*. McGraw-Hill.
- HP. (2014). *Substrate guide for HP Indigo digital presses*. Hewlett-Packard.
- HP. (2020). *The HP Indigo LEP / LEPX technology & the future of digital printing* (White paper). Hewlett-Packard.
- HP. (2023). *HP Indigo 15K digital press: Technical overview*. Hewlett-Packard. <https://www.hp.com/mx-es/industrial-printers/indigo-digital-presses/commercial-digital-presses/15k-printing-press.html>
- HP. (2024). *HP Indigo solutions and PrintOS documentation*. Hewlett-Packard.
- Keka, I., & Cxixco, B. (2025). *Comparing linear and nonlinear models for load profile data using ANOVA, AIC, and BIC*. SAGE Open, 1–22. <https://doi.org/10.1177/21582440251326389>
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (5th ed.). John Wiley & Sons.
- Montero Granados, R. (2016). *Modelos de regresión lineal múltiple*. Universidad de Granada.
- Srinivasan, R. (2022). Application of DoE in additive manufacturing.
- Stat-Ease, Inc. (s.f.). *Designed experiments contribute to 5% printing cost and 40% setup time savings*.
- Arias-Nava, D., et al. (2015). *Diseño de experimentos y optimización de procesos*.
- Botezatu, C., Condrea, I., Oroian, B., Hrițuc, A., Ețcu, M., & Slătineanu, L. (2019). Use of the Ishikawa diagram in the investigation of some industrial processes. *IOP Conference*

Series: Materials Science and Engineering, 682, 012012. <https://doi.org/10.1088/1757-899X/682/1/012012>

Di, Q. L. (2022). Quasi-Monte Carlo Approximations for Exponentiated Quadratic Kernel in Latent Force Models. *Open Journal of Modelling and Simulation*, 10(4), 349-390. <https://doi.org/10.4236/ojmsi.2022.104021>

Harrison, R. L. (2010). Introduction to Monte Carlo Simulation. *AIP Conference Proceedings*, 1204(1), 17-21. <https://doi.org/10.1063/1.3295638>

Kostka, P., Rossi, B., & Ge, M. (2022). Monte Carlo Methods for Industry 4.0 Applications. *arXiv*. <https://doi.org/10.48550/arXiv.2206.11021>

Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication* (White paper). Michael W. Grieves, LLC.

Soori, M., Arezoo, B., & Dastres, R. (2023). Digital twin for smart manufacturing: A review. *Sustainable Manufacturing and Service Economics*, 2, 100017. <https://doi.org/10.1016/j.smse.2023.100017>

Černic Letnar, M., & Muck, T. (2003). The quality of graphic art paper: The influence of paper permanence and durability on graphic art. *Restaurator*, 24(3), 240–252.

Tagansky, B. (2012). HP-Indigo technology and its application to photo printing. In *International Symposium on Technologies for Digital Photo Fulfillment*.

Tuominen, J. (2010). *Surface treatment of polymer-coated paper*.

Tzori, G. (2004). Control mechanisms for print quality assurance in HP Indigo presses. In *IS&T's NIP20: International Conference on Digital Printing Technologies*.

R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.r-project.org/>

Posit Software, PBC. (2024). Shiny: Web application framework for R. <https://shiny.posit.co/>

ANEXOS

Tabla 5.1. Datos históricos observados

Date	Temp (°C)	RH (%)	Printed sheets	Jams	Corte real (mm)	GSM equivalente	Jam Rate
06/12/2025	20.9	43.13	8834	8	750	260	0.09%
07/12/2025	21.57	39.56	10194	4	750	260	0.04%
08/12/2025	21.54	29.79	14117	7	749.5	259.83	0.05%
09/12/2025	21.91	29.68	18820	5	750	260	0.03%
10/12/2025	19.98	35.04	8067	19	749	259.65	0.24%
11/12/2025	22.3	30.14	20498	9	750	260	0.04%
12/12/2025	21.53	32.84	18217	10	749.5	259.83	0.05%
13/12/2025	20.71	35.29	10973	10	749.5	259.83	0.09%
14/12/2025	21.21	36.73	10409	9	749.5	259.83	0.09%
15/12/2025	20.73	35.31	12340	19	749	259.65	0.15%
16/12/2025	19.95	41.25	7194	24	749	259.65	0.33%
17/12/2025	20.94	44.18	1448	18	748.5	259.48	1.24%
18/12/2025	20.76	39.83	1135	22	748	259.31	1.94%
19/12/2025	20.59	35.32	1929	28	748	259.31	1.45%
29/12/2025	21.21	32.08	14235	18	749	259.65	0.13%
30/12/2025	19.94	20.88	2613	3	750	260	0.11%
02/01/2026	20.42	36.91	2627	1	750	260	0.04%
06/01/2026	25.51	30.42	9317	4	750	260	0.04%
07/01/2026	24.4	33.5	10547	11	749.5	259.83	0.10%
08/01/2026	22.93	36.48	13430	8	749.5	259.83	0.06%
09/01/2026	22.5	34.35	10653	10	749.5	259.83	0.09%
21/01/2026	20.55	40.67	395	2	749	259.65	0.51%
24/01/2026	19.11	42.95	16364	15	749.5	259.83	0.09%
26/01/2026	20.84	22.11	8456	26	748.5	259.48	0.31%

27/01/2026	20.28	20.42	14565	17	749	259.65	0.12%
28/01/2026	21.62	25.33	23118	29	749	259.65	0.13%
29/01/2026	20.66	33.3	19221	8	750	260	0.04%
30/01/2026	21.69	19.75	17061	12	749.5	259.83	0.07%
31/01/2026	20.94	22.37	7522	8	749.5	259.83	0.11%
01/02/2026	19.72	34.7	1763	44	748	259.31	2.50%
02/02/2026	21.26	34.64	7058	8	749.5	259.83	0.11%
03/02/2026	20.43	41	19969	27	749	259.65	0.14%
04/02/2026	20.55	35.8	13376	7	750	260	0.05%
05/02/2026	20.83	26.33	219	1	750	260	0.46%
13/02/2026	22.34	38.69	4086	1	750	260	0.02%
14/02/2026	21.72	40.42	14860	1	750	260	0.01%
15/02/2026	21.83	28.57	4566	4	749.5	259.83	0.09%
16/02/2026	22.63	28	8896	1	750	260	0.01%
17/02/2026	22.93	35	23572	13	749.5	259.83	0.06%
18/02/2026	25.01	31.67	23905	14	749.5	259.83	0.06%
19/02/2026	25.03	29.81	17203	16	749	259.65	0.09%
20/02/2026	25.24	32.48	16003	12	749.5	259.83	0.07%
21/02/2026	23.22	33.62	19969	27	749	259.65	0.14%
22/02/2026	23.66	18.96	13376	7	750	260	0.05%
23/02/2026	20.52	26.89	219	1	750	260	0.46%

Tabla 5.2 Funciones RStudio utilizadas

Función	Paquete	Descripción
<code>readWorkbook()</code>	openxlsx	Permite importar datos desde archivos de Excel hacia R, facilitando la lectura de la base de datos experimental utilizada en el análisis.
<code>lm()</code>	base R	Ajusta modelos de regresión lineal, incluyendo modelos lineales, con interacciones y cuadráticos, para explicar la relación entre variables del proceso y la tasa de atascos.
<code>anova()</code>	base R	Realiza el análisis de varianza del modelo, permitiendo evaluar la significancia estadística de los factores y sus interacciones.
<code>summary()</code>	base R	Proporciona un resumen estadístico del modelo, incluyendo coeficientes, valores p y métricas de ajuste como R^2 .
<code>PRESS()</code>	MPV	Calcula el estadístico PRESS (Prediction Error Sum of Squares), utilizado para evaluar la capacidad predictiva del modelo.
<code>stepAIC()</code>	MASS	Implementa selección automática de variables basada en el criterio AIC, optimizando el balance entre ajuste y complejidad del modelo.
<code>compareLM()</code>	rcompanion	Permite comparar múltiples modelos de regresión mediante métricas como R^2 , facilitando la selección del mejor modelo.
<code>predict()</code>	base R	Genera predicciones del modelo para nuevos escenarios de operación, base fundamental de la planta virtual.
<code>saveRDS()</code> / <code>readRDS()</code>	base R	Permiten guardar y cargar modelos entrenados, facilitando su reutilización en aplicaciones como la interfaz Shiny.
<code>fluidPage()</code>	Shiny	Define la estructura general de la interfaz gráfica de la planta virtual.
<code>numericInput()</code>	Shiny	Crea campos de entrada numéricos para que el usuario defina condiciones del proceso (temperatura, humedad, corte, etc.).

<code>actionButton()</code>	Shiny	Permite ejecutar eventos dentro de la interfaz, como iniciar la simulación o reiniciar valores.
<code>eventReactive()</code>	Shiny	Controla la ejecución de la simulación de Monte Carlo únicamente cuando el usuario lo solicita.
<code>renderPrint()</code>	Shiny	Muestra resultados numéricos en la interfaz, como jam rate, producción y eficiencia.
<code>renderPlot()</code>	Shiny	Permite visualizar gráficamente los resultados de la simulación dentro de la aplicación.
<code>rnorm()</code>	base R	Genera números aleatorios con distribución normal, utilizado para simular la variabilidad del jam rate en el método de Monte Carlo.
<code>runif()</code>	base R	Genera valores aleatorios uniformes, utilizado para modelar la variabilidad en hojas perdidas por atasco.
<code>pmax()</code>	base R	Restringe valores a un mínimo (por ejemplo, evitar jam rates negativos o producción negativa).
<code>ggplot()</code>	ggplot2	Permite construir visualizaciones como histogramas y gráficos de barras para analizar los resultados de la simulación.
<code>geom_col()</code>	ggplot2	Genera gráficos de barras para comparar producción planeada vs simulada.
<code>geom_histogram()</code>	ggplot2	Permite visualizar la distribución de la producción real simulada, mostrando la variabilidad del sistema.
<code>updateNumericInput()</code>	Shiny	Permite restablecer los valores de entrada en la interfaz al usar el botón de reinicio.
<code>shinyApp()</code>	Shiny	Ejecuta la aplicación interactiva de la planta virtual en R.

Código 5.1 Comparación de modelos Rstudio

```
#Librerias
library(openxlsx)
library(MPV)
library(MASS)
library(rcompanion)

# 1. Revisar estructura de los datos
str(data)
summary(data)

# 2. DEFINICIÓN DE VARIABLES

# y = variable respuesta, por ejemplo Jam Rate
# x1 = Temperatura
# x2 = Humedad relativa
# x3 = Corte real
# x4 = GSM equivalente, si existe en tu base

# Asegurar que las variables sean numéricas
data$y <- as.numeric(data$y)
data$x1 <- as.numeric(data$x1)
data$x2 <- as.numeric(data$x2)
data$x3 <- as.numeric(data$x3)

# Eliminar filas con datos faltantes
data <- na.omit(data)

# 3. MODELO NULO
# Modelo base sin factores.

m0 <- lm(y ~ 1, data = data)

# 4. MODELO LINEAL

m1 <- lm(y ~ x1 + x2 + x3, data = data)

anova(m1)
summary(m1)

PRESS_m1 <- MPV::PRESS(m1)
R2pred_m1 <- 1 - (PRESS_m1 / sum(anova(m1)$"Sum Sq"))

# 5. MODELO CON INTERACCIONES
```

```

m2 <- lm(y ~ (x1 + x2 + x3)^2, data = data)

anova(m2)
summary(m2)

PRESS_m2 <- MPV::PRESS(m2)
R2pred_m2 <- 1 - (PRESS_m2 / sum(anova(m2)$"Sum Sq"))

# Selección automática del mejor modelo con interacciones
m2_step <- stepAIC(
  m2,
  scope = list(lower = m0, upper = m2),
  direction = "both",
  trace = TRUE
)

summary(m2_step)

# 6. MODELO CUADRÁTICO / SEGUNDO ORDEN

m3 <- lm(
  y ~ x1 + x2 + x3 +
    x1:x2 + x1:x3 + x2:x3 +
    I(x1^2) + I(x2^2) + I(x3^2),
  data = data
)

anova(m3)
summary(m3)

PRESS_m3 <- MPV::PRESS(m3)
R2pred_m3 <- 1 - (PRESS_m3 / sum(anova(m3)$"Sum Sq"))

# Selección automática del mejor modelo cuadrático
m3_step <- stepAIC(
  m3,
  scope = list(lower = m0, upper = m3),
  direction = "both",
  trace = TRUE
)

summary(m3_step)

PRESS_m3_step <- MPV::PRESS(m3_step)
R2pred_m3_step <- 1 - (PRESS_m3_step / sum(anova(m3_step)$"Sum Sq"))

```

7. COMPARACIÓN DE MODELOS

```
compareLM(m1, m2, m2_step, m3, m3_step)

comparacion <- data.frame(
  Modelo = c(
    "Lineal",
    "Interacciones",
    "Interacciones optimizado",
    "Cuadrático",
    "Cuadrático optimizado"
  ),
  R2 = c(
    summary(m1)$r.squared,
    summary(m2)$r.squared,
    summary(m2_step)$r.squared,
    summary(m3)$r.squared,
    summary(m3_step)$r.squared
  ),
  R2_Ajustado = c(
    summary(m1)$adj.r.squared,
    summary(m2)$adj.r.squared,
    summary(m2_step)$adj.r.squared,
    summary(m3)$adj.r.squared,
    summary(m3_step)$adj.r.squared
  ),
  AIC = c(
    AIC(m1),
    AIC(m2),
    AIC(m2_step),
    AIC(m3),
    AIC(m3_step)
  ),
  PRESS = c(
    PRESS_m1,
    PRESS_m2,
    MPV::PRESS(m2_step),
    PRESS_m3,
    PRESS_m3_step
  ),
  R2_Predictivo = c(
    R2pred_m1,
    R2pred_m2,
    1 - (MPV::PRESS(m2_step) / sum(anova(m2_step)$"Sum Sq")),
    R2pred_m3,
    R2pred_m3_step
  )
)
```

```

print(comparacion)

# 8. SELECCIÓN DEL MEJOR MODELO
# Criterio recomendado:
# 1. Mayor R2 predictivo
# 2. Menor PRESS
# 3. Menor AIC
# 4. Modelo más simple posible

mejor_modelo <- m3_step

summary(mejor_modelo)
anova(mejor_modelo)

# 9. DIAGNÓSTICO DEL MEJOR MODELO

par(mfrow = c(1, 1))
plot(mejor_modelo)
par(mfrow = c(1, 1))

# 10. PREDICCIÓN DEL SISTEMA

# planta virtual.

nuevo_escenario <- data.frame(
  x1 = 23,    # Temperatura
  x2 = 60,    # Humedad relativa
  x3 = 750.1  # Corte real
)

prediccion <- predict(
  mejor_modelo,
  newdata = nuevo_escenario,
  interval = "prediction"
)

print(prediccion)

# 11. GUARDAR EL MODELO PARA USARLO EN LA PLANTA VIRTUAL

saveRDS(mejor_modelo, "mejor_modelo_planta_virtual.rds")

```

Código 5.2 Planta virtual Rstudio

```
# PLANTA VIRTUAL PREDICTIVA

library(shiny)
library(ggplot2)

modelo <- readRDS("mejor_modelo_planta_virtual.rds")

# Error del modelo
error_residual <- summary(modelo)$sigma

# Interfaz de usuario

ui <- fluidPage(
  titlePanel("Planta Virtual Predictiva"),
  sidebarLayout(sidebarPanel(h4("Condiciones del proceso"),

    numericInput("x1", "Temperatura (°C)", 23, min = 15, max = 35),
    numericInput("x2", "Humedad relativa (%)", 55, min = 30, max = 80),
    numericInput("x3", "Corte real (mm)", 750, min = 748, max = 752, step = 0.05),

    hr(),

    numericInput("volumen", "Hojas planeadas", 10000),

    numericInput("perdida_min", "Hojas mínimas perdidas por atasco", 1),
    numericInput("perdida_max", "Hojas máximas perdidas por atasco", 4),

    numericInput("n_sim", "Número de simulaciones", 10000),

    hr(),

    actionButton("simular", "▶ Ejecutar simulación"),
    actionButton("reset", "🔄 Reiniciar")
  ),
  mainPanel(
    h3("Resultados"),
    verbatimTextOutput("salida"),

    plotOutput("plot"),
    plotOutput("histograma")
  )
)
```

```
# SERVER
```

```
server <- function(input, output, session) {
```

```
  # RESET
```

```
  observeEvent(input$reset, {  
    updateNumericInput(session, "x1", value = 23)  
    updateNumericInput(session, "x2", value = 55)  
    updateNumericInput(session, "x3", value = 750)  
  
    updateNumericInput(session, "volumen", value = 10000)  
    updateNumericInput(session, "perdida_min", value = 1)  
    updateNumericInput(session, "perdida_max", value = 4)  
    updateNumericInput(session, "n_sim", value = 10000)  
  })
```

```
  # MONTE CARLO
```

```
  simulacion <- eventReactive(input$simular, {
```

```
    nuevos_datos <- data.frame(  
      x1 = input$x1,  
      x2 = input$x2,  
      x3 = input$x3  
    )
```

```
    # Predicción base
```

```
    y_pred <- as.numeric(predict(modelo, newdata = nuevos_datos))
```

```
    # Simulación del jam rate
```

```
    y_sim <- rnorm(input$n_sim, mean = y_pred, sd = error_residual)
```

```
    y_sim <- pmax(y_sim, 0)
```

```
    # Simulación de atascos
```

```
    atascos_sim <- input$volumen * y_sim
```

```
    # Variabilidad en hojas perdidas
```

```
    hojas_por_atasco_sim <- runif(  
      input$n_sim,  
      min = input$perdida_min,  
      max = input$perdida_max  
    )
```

```

hojas_perdidas_sim <- atascos_sim * hojas_por_atasco_sim

produccion_real_sim <- input$volumen - hojas_perdidas_sim
produccion_real_sim <- pmax(produccion_real_sim, 0)

eficiencia_sim <- (produccion_real_sim / input$volumen) * 100

data.frame(
  jam_rate = y_sim,
  atascos = atascos_sim,
  hojas_perdidas = hojas_perdidas_sim,
  produccion_real = produccion_real_sim,
  eficiencia = eficiencia_sim
)
})

# OUTPUT TEXTO

output$salida <- renderPrint({

  sim <- simulacion()

  cat("=== RESULTADOS ===\n\n")

  cat("Jam Rate promedio:", round(mean(sim$jam_rate), 6), "\n")
  cat("Jam Rate P5:", round(quantile(sim$jam_rate, 0.05), 6), "\n")
  cat("Jam Rate P95:", round(quantile(sim$jam_rate, 0.95), 6), "\n\n")

  cat("Atascos esperados (promedio):", round(mean(sim$atascos), 1), "\n")
  cat("Atascos pesimista (P95):", round(quantile(sim$atascos, 0.95), 1), "\n")
  cat("Atascos optimista (P5):", round(quantile(sim$atascos, 0.05), 1), "\n\n")

  cat("Producción promedio:", round(mean(sim$produccion_real), 0), "hojas\n")
  cat("Producción pesimista (P5):", round(quantile(sim$produccion_real, 0.05), 0), "\n")
  cat("Producción optimista (P95):", round(quantile(sim$produccion_real, 0.95), 0), "\n\n")

  cat("Eficiencia promedio:", round(mean(sim$eficiencia), 2), "%\n")
})

# GRÁFICA COMPARACIÓN

output$plot <- renderPlot({

  sim <- simulacion()

  df <- data.frame(

```

```

Escenario = c("Planeado", "Promedio", "Pesimista", "Optimista"),
Hojas = c(
  input$volumen,
  mean(sim$produccion_real),
  quantile(sim$produccion_real, 0.05),
  quantile(sim$produccion_real, 0.95)
)
)

ggplot(df, aes(x = Escenario, y = Hojas)) +
  geom_col() +
  labs(title = "Producción planeada vs simulada") +
  theme_minimal()
})

# HISTOGRAMA

output$histograma <- renderPlot({

  sim <- simulacion()

  ggplot(sim, aes(x = produccion_real)) +
    geom_histogram(bins = 40) +
    labs(
      title = "Distribución de producción real simulada",
      x = "Producción",
      y = "Frecuencia"
    ) +
    theme_minimal()
})
}

# RUN

shinyApp(ui, server)

```