

**UNIVERSIDAD DE CONCEPCIÓN - CHILE  
FACULTAD DE INGENIERÍA  
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL**

**Metodologías de arribos y su impacto en  
las métricas de un estudio de simulación  
sobre capacidad de almacenamiento en  
un centro de distribución de consumo  
masivo**



Por  
**Juan Emilio Gatica Vásquez**  
Profesor Guía  
**Rodrigo De la Fuente Gallegos, Ph.D.**

Concepción, Abril de 2020

Tesis presentada a la

**DIRECCIÓN DE POSTGRADO  
UNIVERSIDAD DE CONCEPCIÓN**



Para optar al grado de

**MAGÍSTER EN INGENIERÍA INDUSTRIAL**

©

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.



## RESUMEN

En todo modelo de simulación existe incertidumbre en los parámetros de entrada, sobre todo en el generador de llegadas. Este trabajo compara diferentes métodos de arribos a un centro de distribución, para luego analizar los efectos en métricas de interés. Lo anterior, asistido por un modelo de simulación elaborado mediante software Simio LLC, con el cual se obtuvo la capacidad de almacenamiento y tiempo de elaboración de cargas.

La comparación de métodos de arribos se basa en cinco criterios; (1) validación estadística entre la capacidad promedio obtenida por los modelos de arribos y la capacidad promedio real, (2) validación entre la capacidad promedio obtenida por los modelos de arribos y la *capacidad promedio modelada*, (3) La variabilidad de las llegadas y resultados, (4) dependencia o correlación entre los arribos, y finalmente, (5) la exactitud de los resultados respecto a los valores reales.

Los resultados muestran que el modelo Markov-MECO (Mixture of Erlangs Common Order) presentó el mejor desempeño en base a los diferentes criterios de evaluación, mientras que el Modelo Exponencial Autorregresivo y Poisson no Homogéneo presentaron una volatilidad mayor a la presente en el proceso de arribo de camiones al centro de distribución.

Complementariamente, se determinó una estimación del impacto económico que puede causar la elección de un modelo de arribos inapropiado, al momento de hacer una recomendación basada en una simulación, se alcanza un monto de CLP 14,298,514 para el caso estudio de esta investigación.

**Palabras Claves:** Procesos de Arribos, Simulación, Markov-MECO.

## TABLA DE CONTENIDOS

<b>RESUMEN</b>	<b>II</b>
<b>LISTA DE FIGURAS</b>	<b>V</b>
<b>LISTA DE TABLAS</b>	<b>VII</b>
<b>1. INTRODUCCIÓN Y OBJETIVOS</b>	<b>1</b>
1.1. INTRODUCCIÓN . . . . .	1
1.2. OBJETIVOS . . . . .	5
<b>2. PROCESO Y MODELO DE SIMULACIÓN</b>	<b>7</b>
2.1. DESCRIPCIÓN PROCESO . . . . .	7
2.2. MODELO DE SIMULACIÓN . . . . .	9
2.2.1. Estructura . . . . .	9
<b>3. PROCESO DE DATOS</b>	<b>11</b>
3.1. ANÁLISIS PRELIMINAR DE DATOS . . . . .	11
3.2. AJUSTE DE LOS MODELOS DE ARRIBOS . . . . .	15
3.2.1. Modelo 1: Heurística. . . . .	16
3.2.2. Modelo 2: Proceso de Poisson no Homogéneo (NHPP). . . . .	17
3.2.3. Modelo 3: Exponencial autorregresivo EAR(p). . . . .	19
3.2.4. Modelo 4: Proceso Markov MECO (no estacionario y no renovable). . . . .	22
<b>4. RESULTADOS Y DISCUSIÓN</b>	<b>28</b>
<b>5. CONCLUSIÓN</b>	<b>36</b>
<b>ANEXOS</b>	<b>42</b>
<b>A. MODELO DE SIMULACIÓN</b>	<b>43</b>

**B. AUSENCIA DE DATOS**

**46**

**C. VALIDACIÓN MODELO DE SIMULACIÓN**

**51**



## LISTA DE FIGURAS

1.1. Taxonomía de modelos de entrada para procesos dependientes del tiempo. . . . .	2
2.1. Flujo de los productos desde su creación hasta la entrega al cliente. . . . .	7
2.2. Participación de las operaciones en los cambios stock. . . . .	8
2.3. Esquema modelo de simulación. . . . .	9
3.1. Número de llegadas diarias. . . . .	12
3.2. Número promedio de llegadas por hora durante una semana. . . . .	12
3.3. Relación entre día de la semana y hora del día. . . . .	13
3.4. Gráfico de cajas del número de llegadas por mes. . . . .	14
3.5. Gráfico de cajas del número de llegadas por día de la semana. . . . .	14
3.6. Gráfico de control con límites $3\sigma$ . . . . .	15
3.7. Ajuste Heurística. . . . .	16
3.8. Comparación semanal entre los datos reales y el ajuste de la Heurística. . . . .	17
3.9. Ajuste NHPP. . . . .	18
3.10. Comparación entre las llegadas generadas por NHPP y los datos reales. . . . .	19
3.11. Función Autocorrelación. . . . .	20
3.12. Autocorrelaciones parciales. . . . .	20
3.13. Iteraciones obtenidas en Octubre mediante el modelo EAR(2). . . . .	22
3.14. Comparación semanal entre las 20 iteraciones y los datos reales. . . . .	23
3.15. Proceso Markov MECO. . . . .	24
3.16. Llegadas reales y réplicas obtenidas mediante el método de Nelson and Gerhardt (2011). . . . .	26
3.17. Comparación llegadas reales con Modelo Markov MECO. . . . .	27
4.1. Comparación semanal entre las capacidades promedio obtenidas por cada modelo de arribos. . . . .	29
4.2. Diagrama de cajas de la capacidad del promedio del almacén. . . . .	32
4.3. Diagrama de cajas de la capacidad máxima del almacén. . . . .	34

A.1. Comparación resultados de todos los Experimentos. . . . .	45
B.1. Ajuste del error a una distribución normal. . . . .	50
B.2. Ajuste de una distribución normal al error. . . . .	50
C.1. Comparación entre las capacidades obtenidas en base a datos reales y en base al modelo de simulación. . . . .	51



## LISTA DE TABLAS

1.1. Comparación de investigaciones respecto a las temáticas que tratan. . . . .	6
2.1. Principales estructuras definidas en el modelo de simulación. . . . .	10
3.1. Correlaciones entre días de la semana. . . . .	15
3.2. Parámetros del modelo EAR(2). . . . .	21
3.3. Prueba de Box Pierce. . . . .	22
3.4. Ejemplo del cálculo de parámetros para días martes en temporada alta. . .	25
4.1. Pruebas de medias entre la capacidad de la bodega obtenida mediante el modelo de simulación y los datos reales. . . . .	29
4.2. Pruebas de medias entre la capacidad promedio modelada y las capacidades promedio obtenidas por los modelos de arribos. . . . .	30
4.3. Variabilidad presente en las llegadas generadas por cada modelo de arribos.	31
4.4. Resumen elección de modelo de arribos. . . . .	33
4.5. Resultados del tiempo de elaboración de cargas en Picking. . . . .	35
A.1. Resumen experimentos. . . . .	43
A.2. Tabla resumen de resultados. . . . .	44
A.3. Tabla resumen de resultados capacidad de almacenamiento. . . . .	44
B.1. Parámetros y estadísticas regresión lineal. . . . .	47
B.2. Parámetros $\beta_j$ . . . . .	47
B.3. Parámetros $\gamma_j$ . . . . .	48
B.4. Parámetros $\delta_{j,k}$ . . . . .	48
B.5. Corridas arriba y abajo. . . . .	49
B.6. Valores estimados para completar la serie de tiempo. . . . .	50
C.1. Pruebas de medias entre la capacidad de la bodega obtenida mediante el modelo de simulación y los datos reales. . . . .	52

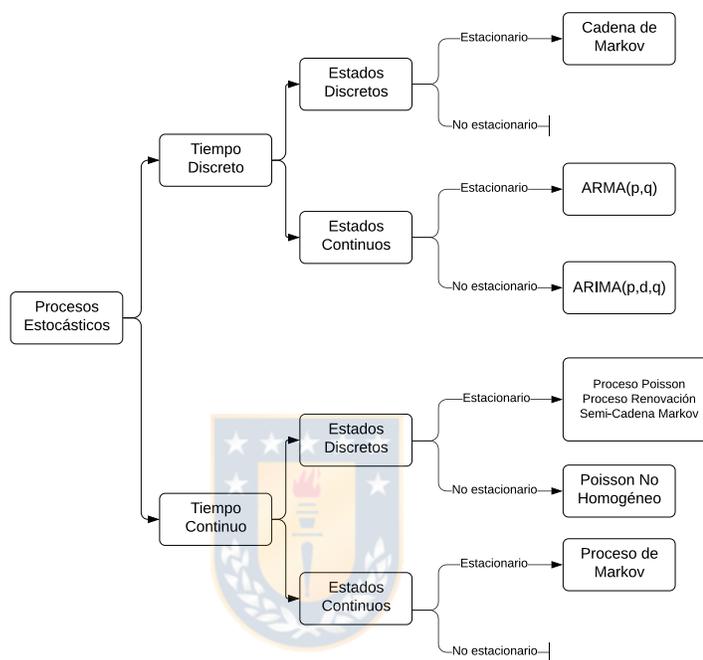
## 1. INTRODUCCIÓN Y OBJETIVOS

### 1.1. INTRODUCCIÓN

En los sistemas de servicio generalmente los arribos son una fuente primaria de incertidumbre, en donde es necesario acudir a procesos estocásticos complicados y difíciles de modelar según Ibrahim et al. (2012). De igual forma, Rabe and Scheidler (2014) son contundentes al comentar que una de las temáticas más relevantes en simulaciones de eventos discretos sobre cadenas de suministros es la recopilación y el análisis de las entradas del modelo, comentando que la precisión de la salida depende fuertemente de la calidad de los modelos de entrada. Esta investigación, tiene como objetivo evaluar diferentes metodologías que buscan parametrizar los arribos de un sistema de servicio, para luego analizar los efectos en métricas de interés. Dicha evaluación será asistida por un modelo de simulación elaborado mediante software Simio LLC, considerando un caso real basado en la llegada de camiones a un centro de distribución local. A continuación, presentaremos una revisión de los modelos utilizados para parametrizar un proceso de llegadas, indicando en qué instancias es pertinente aplicar un modelo determinado, en conjunto con indicar los métodos existentes. Luego, comentaremos de qué forma podemos escoger un modelo adecuado basados en como su ajuste se compara con los datos reales.

Ahora, es importante mencionar que esta investigación busca comparar distintos procesos de arribos aplicándolos en un estudio de caso. Cheng (2017) realizó una revisión sobre el desarrollo y uso de modelos de entrada con el enfoque principal en la simulación de eventos discretos (DES), señalando diferentes metodologías dependientes del proceso y de la disponibilidad de datos existente. De igual forma, Leemis (2004) menciona un esquema que permite seleccionar un modelo de entrada de acuerdo a las características del proceso, esto es, dependencias respecto al tiempo, estados discretos o continuos, y la presencia de estacionalidad (Ver Figura 1.1). Para procesos de entradas independientes del tiempo, Leemis (2004) aconseja ajuste de distribuciones. Adicionalmente, Law (2013) detalla dos métodos para modelar una fuente de aleatoriedad; ajustando una distribución de probabilidad teórica a los datos y el uso de una distribución empírica. Respecto a dicha temática existen adaptaciones de distribuciones más avanzadas, por ejemplo, Kuhl et al.

(2006) considera distribuciones Johnson, Beta Generalizada, Bézier y distribuciones bivariadas. Sin embargo, no profundizaremos mayormente en ajuste de distribuciones como método de entrada al modelo de simulación, dado que el enfoque de esta investigación no son modelos independientes del tiempo.



**Figura 1.1.** Taxonomía de modelos de entrada para procesos dependientes del tiempo.

Fuente: Leemis (2004)

Para procesos dependientes del tiempo, Leemis (2004) aconseja modelos de entrada basados en métodos autoregresivos, procesos de Markov y procesos de Poisson no homogéneos. Los dos primeros además, son recomendados para llegadas correlacionadas. Según Nelson (2013) los procesos Poisson estacionarios y no estacionarios son ampliamente utilizados en la práctica de la simulación, siendo incluso por muchos autores el proceso de arribos por defecto, sin embargo, se debe cumplir que los tiempos entre arribos deben ser distribuidos de forma exponencial y las llegadas deben ser independientes lo que en la práctica no siempre es cierto. Por lo anterior, existe un amplio número de investigaciones que han estudiado procesos de arribos a centros de llamados, dado que estos cumplen con las restricciones teóricas mencionadas, por ejemplo, Kim and Whitt (2014), Brown et al. (2005), He et al. (2016), Ibrahim et al. (2016), entre otros.

Respecto a procesos Poisson, Ibrahim et al. (2012) mencionan que en casos reales la dispersión es considerablemente mayor respecto a la media; por lo tanto, es recomenda-

ble considerar una tasa estocástica para dichos procesos. Dicho esto, si consideramos la variación en el tiempo y bajo ciertas suposiciones, los procesos Poisson no homogéneos (NHPP) son una alternativa válida para ciertos estudios de casos. Sin embargo, cuando los procesos de arribos presentan llegadas correlacionadas, se deben tratar con especial cuidado. Uhlig et al. (2016) proporcionan una revisión de las herramientas disponibles para ajustar y modelar los procesos de llegadas con dicha característica. Los autores evalúan las herramientas basándose en su capacidad para modelar procesos de entrada que se ajustan a una traza o definidos explícitamente por su distribución marginal y los coeficientes de autocorrelación, encontrando que las herramientas que se basan en modelos autorregresivos son las que mejor se desempeñan. Complementariamente, Bause et al. (2009) compararon dos metodologías; los procesos de llegada de Markov (MAP), y los modelos basados en ARMA/ARTA, ambos recomendados por Leemis (2004) para procesos estocásticos con tiempos discretos. Bause encontró, que efectivamente las dos opciones permiten modelar entradas correlacionadas pero indicando que los modelos ARMA/ARTA se desempeñan de mejor forma; sin embargo, requieren de un ajuste más elaborado y están restringidos a trabajar bajo supuestos de normalidad.

Otra metodología utilizada para entradas correlacionadas y sobre todo para el sector financiero, es el análisis de copulas. Pasupathy and Nagaraj (2015) mencionan su definición como una función de los componentes de una variable multivariada  $d$ -dimensional, comentando que son un modelo flexible y conveniente para incorporar dependencia. El requisito es que cada uno de los componentes transformados tiene una distribución marginal que es uniforme  $U(0, 1)$ . Esta estructura retiene la interdependencia entre los componentes de modo que una vez transformados a sus marginales originales, preservan la dependencia multivariada. Existen múltiples trabajos que abordan dicha metodología, por ejemplo, Kurowicka and Cooke (2002), Channouf and L'Ecuyer (2010) y Channouf and L'Ecuyer (2012). Estos últimos consideran un modelo de cópula normal para el proceso de llegadas en centros de llamadas, temática relacionada directamente con esta investigación.

Los modelos de efectos fijos y mixtos enfocado en procesos de arribos son otra herramienta que captura la dependencia de datos. Para estos modelos es necesario determinar factores en base a los datos históricos que representan, por ejemplo, los efectos del día de la semana, del período del día, de la interacción día-período, efecto de día especial del tipo  $k$  y el efecto de interacción del período con el día especial. Por ejemplo, Aldor et al. (2010) desarrollan un modelo mixto que incluye dichos factores, en donde tratan un caso de estudio basado en un centro de llamadas. En este utilizan el modelo para pronosticar el número de llamadas para determinar la carga laboral del recinto.

Una temática importante a considerar y que está relacionada directamente con esta in-

vestigación, es cuando hay diferentes tipos de llegadas (en nuestro caso distintos orígenes dados por las plantas productoras). Generalmente, hay dependencia entre tasas de llegada en los diferentes procesos de arribo. Respecto a esto, Jaoua et al. (2013) realizaron un análisis en los ingresos de llamadas de un call center, utilizando modelos de copulas para definir los arribos de los diferentes tipos de llamados. En su análisis concluyen que omitir la correlación entre llegadas tiene un impacto significativo en el desempeño de la simulación, por lo tanto, ignorando esta dependencia se puede conducir a una subestimación o sobreestimación sustancial del rendimiento del centro de llamadas.

Un enfoque no tan común para modelar las entradas de un modelo de simulación, es la formulación bayesiana. Esta requiere que se supongan en un inicio distribuciones para cada modelo de entrada y para los parámetros correspondientes a cada uno. Luego se obtienen posteriores parametrizaciones con el método de Cadena de Markov Monte Carlo, tal como se explica en Xie et al. (2014). Según Cheng (2017), un aspecto interesante del enfoque bayesiano es fácil aplicación cuando no hay datos existentes, dado que se pueden estimar parámetros desconocidos de los modelos de entrada, sólo utilizando suposiciones.

Un aspecto importante se refiere a la cantidad de datos que son necesarios para obtener un modelo con un buen desempeño, Freimer and Schruben (2002) discuten cuántos datos recolectar para ajustar parámetros importantes utilizando experimentos diseñados, ya sea con un análisis de modelo de efectos fijos de varianza, o un modelo de efectos aleatorios.

Los modelos de entrada paramétricos también figuran como opción. Estos a menudo se seleccionan en función de qué tan bien se ajustan a una muestra dada de datos reales. Sin embargo, Cheng (2017) menciona que los modelos de entrada paramétricos disminuyen la posibilidad de que las elecciones sean adecuadas, debido a las incertidumbres en conjunto tanto de la simulación y la estimación de los parámetros. Los sistemas Bézier son interesantes dado que son paramétricos y muy flexibles; sin embargo, esto significa que su ajuste es bastante subjetivo, haciendo difícil evaluar formalmente la calidad del ajuste logrado.

Al intentar parametrizar las llegadas a un sistema, existen ciertos comportamientos que es importante identificar para obtener estimaciones confiables. Ibrahim et al. (2012) revisaron algunos problemas claves al modelar la tasa de llegada estocástica, estos son; (1) La tasa de llegada varía con la hora del día y exhibe estacionalidades diarias, semanales o anuales, (2) La demanda total (número de llamadas entrantes) durante un período de tiempo determinado tiene una dispersión excesiva comparada con una distribución de Poisson, y (3) en ciertos tipos de centros de llamadas, la tasa de llegada a veces tiene picos altos inesperados en cortos períodos de tiempo.

Para procesos de llegadas que consideran varios días (los cuales pueden ser tratados

como series de tiempo), se pueden adaptar los modelos de entradas que analizan los arribos dentro del día iterándolos con los parámetros correspondientes al periodo en curso, considerando fin de semana, feriados, etc. Channouf et al. (2007) desarrollaron modelos aditivos simples para el número (pequeño) de llamadas de ambulancia en cada hora. Sus modelos capturan estacionalidades diarias, semanales y anuales con comportamientos similares al caso de estudio abordado en esta investigación.

Hasta el momento hemos mencionado diferentes opciones capaces de parametrizar las entradas a un modelo de simulación, en donde estos son utilizados de acuerdo al comportamiento de los datos. Pero ¿Cómo definir qué modelo es el más adecuado? Una vez elegida una metodología de ajuste a las entradas de un modelo de simulación es importante evaluar dicha estimación. Law (2015) da una revisión detallada a pruebas que evalúan los ajustes de datos reales. Comúnmente, la formulación de hipótesis y elección mediante *valor - p* suele ser muy utilizado, en donde una pequeña magnitud de dicha estadística indica un ajuste insatisfactorio, adicionalmente, menciona que evaluar una estimación mediante solo el *valor - p* no es aconsejable. Una metodología complementaria es el análisis de gráficos probabilidad-probabilidad (P-P) y cuantil-cuantil (Q-Q). Gupta and Parzen (2005) sugieren un examen visual con el uso de gráficos Q-Q para identificar distribuciones adecuadas ajustando los datos y posteriormente comparándolos con ajustes por gráficos P-P para comprobarlos.

Ahora bien, existen múltiples investigaciones que comparan metodologías de arribos. Como se aprecia en la Tabla 1.1, las temáticas donde existen más diferencias se refieren a la evaluación de salidas (efectos en los resultados), el tratamiento de la estacionalidad de los datos y cuando existen diferentes orígenes de arribos. Otra tendencia presente en la literatura es que generalmente las evaluaciones son en estudios de casos basados en call centers y centros médicos. Por su parte, esta investigación busca diferenciarse no solo abordando temáticas de tendencia y estacionalidad, sino que también correlación. Además, el estudio de caso se basa en un rubro no explorado para comparar metodologías de arribos, en donde se busca responder ¿Qué impactos tiene una metodología de arribos adecuada en las conclusiones de dotación de personal y capacidad de almacenamiento que el centro posee?

## 1.2. OBJETIVOS

### Objetivo General

Evaluar diferentes metodologías de arribos en base al impacto que estas generan en los resultados de una simulación de eventos discretos.

Tabla 1.1: Comparación de investigaciones respecto a las temáticas que tratan.

Autor	Efectos en Resultados	Metodología Arribos	Estacionalidad	Tipos de Arribos	Simulación	Caso Práctico
Ibrahim et al. (2012)	✓	MEM, MEF, SWH, ETD	✓		✓	Call center
Xie et al. (2014)	✓	EB			✓	Centro Médico
Channouf and L'Ecuyer (2012)		Copulas			✓	Call center
Freimer and Schruben (2002)		MEA, MEF	✓		✓	
Jaoua et al. (2013)	✓	Copulas		✓	✓	Call center
Channouf et al. (2007)		ARIMA	✓		✓	Centro Médico
Aldor et al. (2010)	✓	MEM, MEF				Call center
Bause et al. (2009)		ARMA/ARTA y MAP	✓		✓	Tráfico internet
Ibrahim et al. (2016)		MEM, MMU	✓	✓		Call center
He et al. (2016)	✓	NSNP			✓	Call center
Uhlig et al. (2016)		ARTA y MAP				
Esta Investigación	✓	HEU, NHPP, EAR, MECO	✓		✓	CD Bebestible

MEF: Modelo efectos fijos, MEM: Modelo efectos mixtos, ETD: Enfoque Top-Down, SWH: Suavización Winter-Holt, EB: Enfoque Bayesiano, MEA: Modelo efectos aleatorios, MMU: Modelo multiplicativo univariante, HEU: Heurística, NHPP: Proceso Poisson no Homogéneo, EAR: Exponencial autoregresivo y MECO: Markov MECO.

### Objetivos Específicos

- Identificar metodologías de arribos relevantes en la literatura actual y aplicarlas a un caso de estudio real.
- Proponer métricas que permitan escoger la metodología de arribos con mejor desempeño.
- Evaluar los beneficios y riesgos esperados de la implementación de la metodología con mejor desempeño.

Lo que queda de este trabajo está organizado de la siguiente manera. Para comenzar, el Capítulo 2 detalla el proceso y el modelo de simulación utilizado en esta investigación. En seguida, el Capítulo 3 presenta las diferentes metodologías de modelos de llegadas utilizados para parametrizar los arribos al sistema. Luego, el Capítulo 4 presenta los resultados y discusión obtenidos al aplicar dichas metodologías en el modelo de simulación. Finalmente, en el Capítulos 5 se entregan las principales conclusiones.

## 2. PROCESO Y MODELO DE SIMULACIÓN

En el siguiente capítulo se detallarán las características del proceso y el modelo de simulación desarrollado en De la Fuente et al. (2019) utilizado como medio para obtener los resultados de esta investigación. En primera instancia se describirá el proceso modelado y luego los detalles del modelo implementado que permitirá evaluar las distintas metodologías de entrada.

### 2.1. DESCRIPCIÓN PROCESO

Cómo se mencionó en el capítulo anterior, el caso real utilizado en esta investigación es sobre un centro de distribución local de consumo masivo. En primera instancia describiremos generalidades para entender la lógica del proceso modelado y luego las principales decisiones que se tomaron para realizar la simulación.

Para comenzar, los productos siguen un proceso estándar desde su elaboración hasta que son entregados a los clientes finales, como se presenta en la Figura 2.1. En primera instancia plantas productoras realizan despachos de acuerdo a un requerimiento mensual, luego es deber del centro de distribución almacenar y confeccionar las cargas para su posterior despacho.

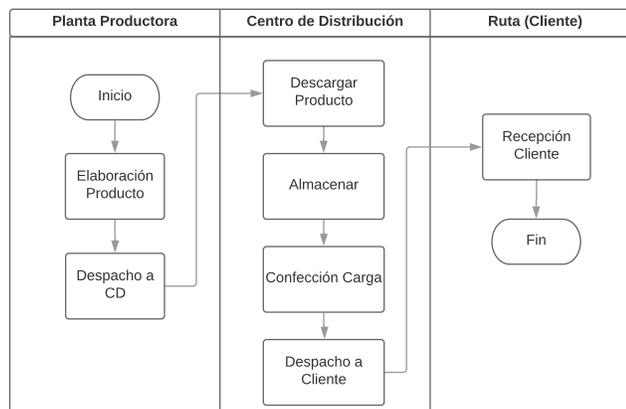


Figura 2.1: Flujo de los productos desde su creación hasta la entrega al cliente.

Actualmente, el centro de distribución administra productos con envases y líquidos de distintas características, lo que ocasiona una variedad de 324 SKU. Dicho volumen presenta una complejidad en la operación diaria del almacén dado que los productos se deben gestionar considerando variables como; el material del envase y su tiempo de vida. El material del envase limita el máximo almacenamiento vertical asegurando la integridad del producto, siendo así, la principal restricción en términos de capacidad de almacenaje. Por otro lado, por políticas de la empresa no se pueden distribuir productos cuyo período de vida útil remanente no supere los treinta días, obligando así, a mantener un nivel de stock suficiente para cubrir la demanda y a la vez evitar un sobre stock que aumente el riesgo de pérdidas de productos por edad avanzada o aumente los costos de almacenamiento.

Los niveles de stock de los productos varían por diversos motivos. Como es de esperar la operación que genera un aumento de stock es el abastecimiento, mientras que los principales motivos de salidas de stock son; demanda, despacho a sucursales y la utilización de productos para elaborar multi-pack. La Figura 2.2 presenta la participación en los cambios de stock considerando los datos de un año.

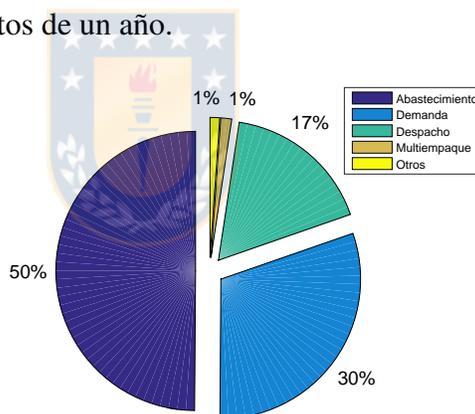


Figura 2.2: Participación de las operaciones en los cambios stock.

A continuación, se indica el procedimiento utilizado para parametrizar cada una de las operaciones que fueron incluidas en el modelo de simulación. En cuanto al personal del área de operaciones en la zona de picking, existen tres tipos de contratos: Planta, Temporada y Contratista. Considerando los costos previsionales, salud, el salario y pago de horas extras el contrato para personal de planta es el más costoso, seguido por el de temporada y finalmente el tipo contratista. Estos dos últimos además están restringidos a que no pueden exceder los once meses de duración. Un factor importante respecto al personal, es el cambio de productividad que presentan los diferentes trabajadores. Esto debido a la antigüedad y a la carga laboral que tenga cada uno. Además, la duración del

contrato afecta derechamente los resultados del modelo de simulación, ya que a medida que existan contratos con baja duración, aumenta la rotación y por ende los niveles de producción disminuyen.

## 2.2. MODELO DE SIMULACIÓN

### 2.2. Estructura

El modelo de simulación fue elaborado mediante software Simio LLC considerando los pasos enunciados por Law (2008). Se utilizaron las bases de datos entre el 3 de Octubre de 2016 y el 30 de Septiembre de 2017. En términos generales se modelaron los flujos de productos diarios de las operaciones del almacén, para así obtener un seguimiento del nivel de stock de cada familia (agrupación de productos con mismas características), permitiendo determinar la dotación de personal necesaria y el tamaño de las instalaciones que aseguren un buen funcionamiento. Para ello, fue necesario caracterizar la demanda y los arribos de camiones, en conjunto con la dinámica de movimientos de productos dentro del almacén y turnos laborales del personal. En la Figura 2.3 se presenta un esquema general del modelo de simulación y en la Tabla 2.1 se detalla la descripción de cada elemento.

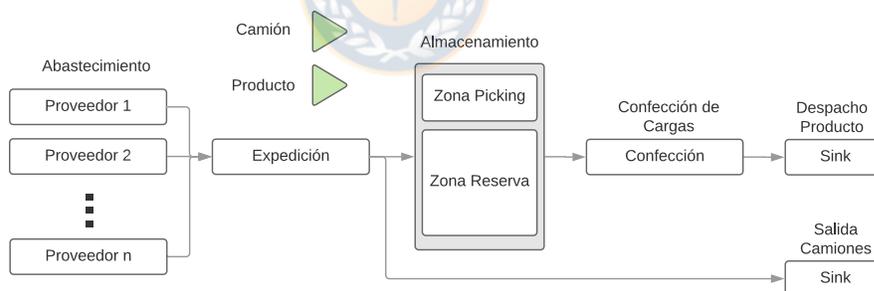


Figura 2.3: Esquema modelo de simulación.

El modelo obtiene dos medidas de desempeño que son de especial interés para nuestros objetivos, capacidad de almacenamiento y el costo total de la dotación de personal, este último relacionado directamente con el tiempo de elaboración de cargas (TEC). Estos indicadores son claves para evaluar las distintas metodologías de arribos. En el Anexo A, se presentan los escenarios y resultados del modelo de simulación mencionado en este capítulo, para más detalles ver De la Fuente et al. (2019). Es importante mencionar, que en esta investigación sólo se considerará un proveedor como input al modelo de simulación, el cual presenta el mayor flujo de ingreso al almacén del centro de distribución.

Tabla 2.1: Principales estructuras definidas en el modelo de simulación.

Tipo	Nombre	Descripción
Entidad	Camión	Representa el arribo generado desde las plantas productoras a centro de distribución. Contiene la información de los productos que contiene.
	Producto	Representa un lote de producto de mismo SKU y arribo. Contiene información de su cantidad y vida del producto.
Source	Proveedor	Figura a las plantas productoras. Genera arribos de camiones.
	Expedición	Punto de ingreso al centro de distribución.
Workstation	Confección	Zona utilizada para la confección de cargas. Determina el tiempo necesario para terminar el proceso y recolecta estadísticas
	Zona Picking	Este elemento almacena las entidades tipo producto, representando el actual stock en la zona de picking.
Storage	Zona Reserva	Este elemento almacena las entidades tipo producto, representando el actual stock en la zona de reserva.
	Salida	Salida de productos desde el almacén.
Sink	Salida Camion	Salida de camiones desde el almacén.

### 3. PROCESO DE DATOS

En este capítulo detallaremos las llegadas al centro de distribución. En primera instancia, se presenta el comportamiento que tienen actualmente, para contar con una vista preliminar de los datos que serán posteriormente parametrizados. Luego en la siguiente sección, detallaremos los diferentes modelos que fueron ajustados a los arribos del sistema.

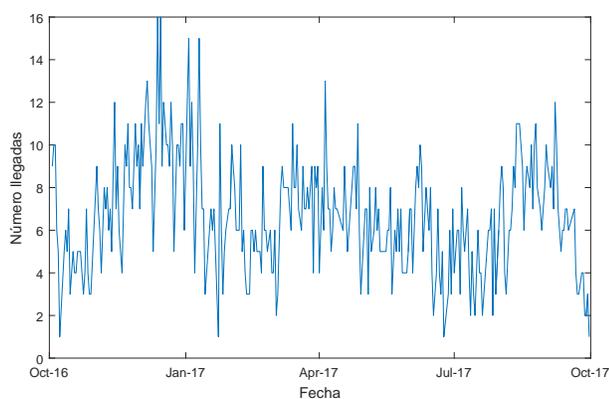
#### 3.1. ANÁLISIS PRELIMINAR DE DATOS

A continuación, presentaremos como introducción los arribos al sistema que ocurrieron entre Octubre del año 2016 y Septiembre del año 2017, con tal de visualizar el comportamiento presente en el proceso. Para nuestra investigación consideramos las llegadas provenientes de la ciudad de Talca, Chile. Estos son los que mueven el mayor volumen del almacén (57,5 %), y abarcan 13 de las 61 familias de productos.

En la Figura 3.1 se muestra el número de llegadas diarias durante el periodo mencionado. Las cifras sugieren una tendencia negativa, probablemente explicado por la inclinación de la población hacia estilos de vida más saludables, lo que gatilla una disminución en el consumo de bebidas gaseosas. De igual forma se observan mayores valores en Diciembre seguido por Abril, Agosto y Noviembre.

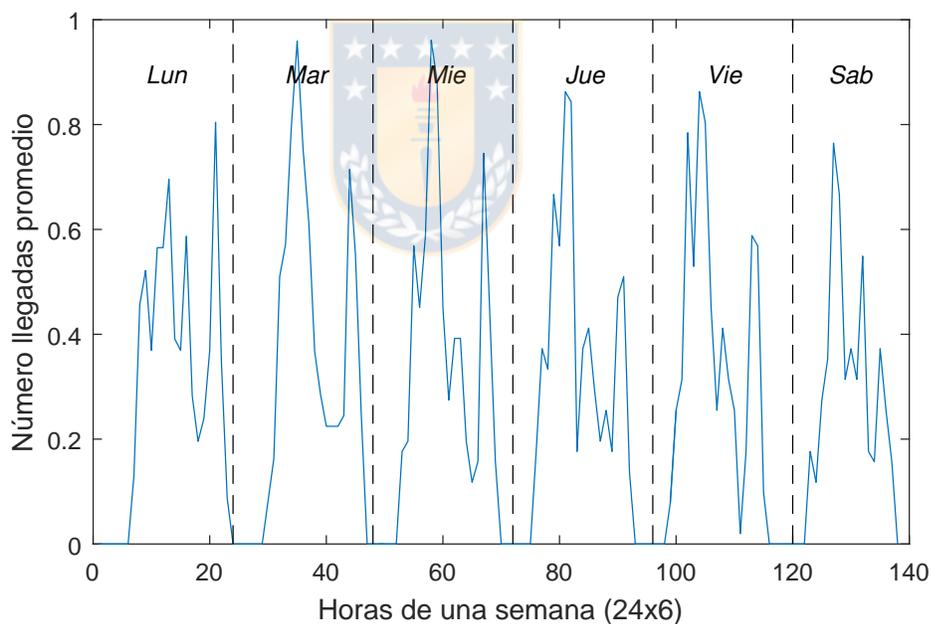
La Figura 3.2 muestra las llegadas promedio por hora durante el ciclo semanal. Se observa una clara estacionalidad de la hora del día (con una cierta variación los días lunes), generalmente se observan mayores volúmenes entre las 12:00 pm y las 8:00-10:00 p.m. Lo anterior, respaldado con la Figura 3.3 en donde se aprecia claramente las horas con mayor frecuencia de llegadas durante la semana. Con respecto al volumen diario, se observa un menor flujo durante el sábado respecto a otros días de la semana (Ver Figura 3.5). Estas observaciones deberían tenerse en cuenta al diseñar horarios de turnos para personal del almacén.

Las Figuras 3.4 y 3.5 coinciden con lo mencionado anteriormente, estas presentan diagramas de caja del volumen diario vistos de forma mensual y para cada día de la semana, respectivamente. Estos concuerdan con que los sábados tienen menor volumen y que



**Figura 3.1.** Número de llegadas diarias.

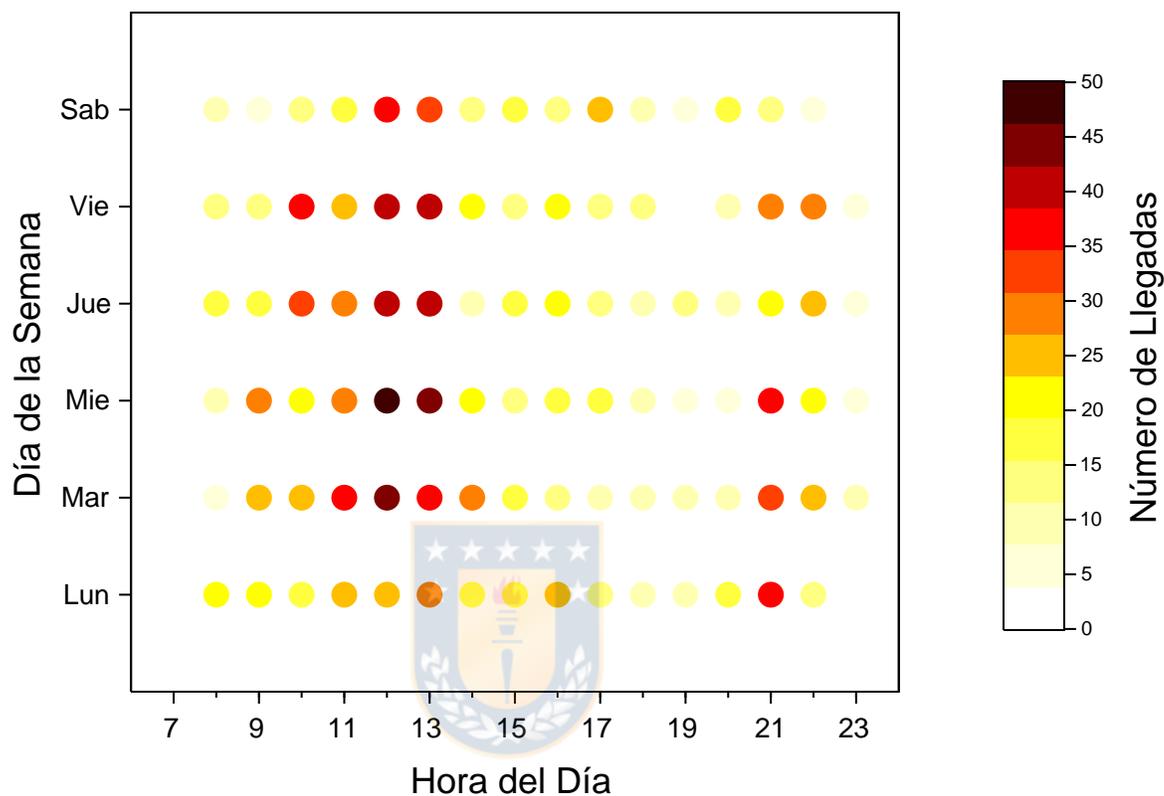
Fuente: Elaboración Propia



**Figura 3.2.** Número promedio de llegadas por hora durante una semana.

Fuente: Elaboración Propia

los días martes poseen una mayor variabilidad respecto a otros días (tamaño de la caja). Además, el mes que presenta mayor volumen es el mes de Diciembre (esperable por el comportamiento del rubro) y que en menor medida los meses de Noviembre, Enero (mayor variabilidad), Marzo, Abril y Agosto. Mientras que Febrero es el mes más tranquilo y



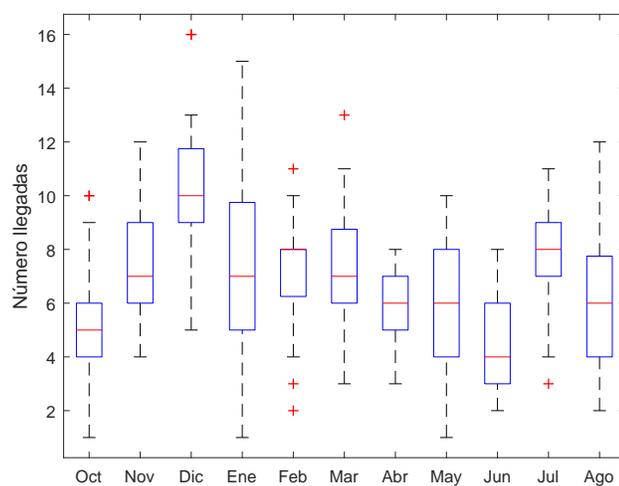
**Figura 3.3.** Relación entre día de la semana y hora del día.

Fuente: Elaboración Propia

estable (menor variabilidad).

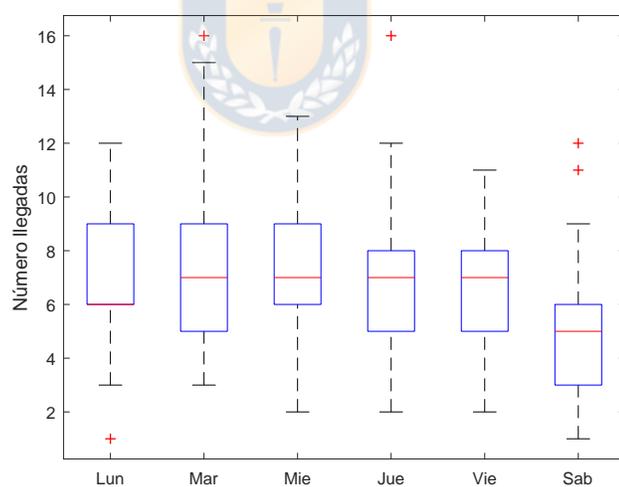
Al examinar las correlaciones entre días (Ver Tabla 3.1), notamos que generalmente son más pequeñas con retrasos mayores. Dicha cualidad es compartida con las llegadas de call center (Ibrahim et al. (2016)). Sin embargo existen diferencias importantes, por ejemplo, los días lunes en los centros de llamados se presentan menores correlaciones con los días restantes y existen correlaciones fuertes y positivas entre períodos sucesivos dentro de los días, características que no son apreciadas de forma clara en los arribos del centro de distribución en estudio.

Para continuar, eliminaremos de nuestra muestra los datos atípicos que solo generan complicaciones a la hora de ajustar modelos de llegadas. En la Figura 3.6 podemos visualizar las llegadas diarias al centro de distribución, y podemos identificar los datos que violan los límites de control ( $\pm 3\sigma$ ), estos son omitidos a la hora de ajustar los modelos y



**Figura 3.4.** Gráfico de cajas del número de llegadas por mes.

Fuente: Elaboración Propia



**Figura 3.5.** Gráfico de cajas del número de llegadas por día de la semana.

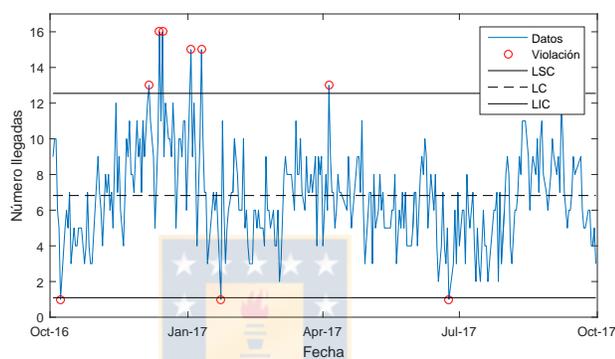
Fuente: Elaboración Propia

solo consideramos los 292 datos restantes.

Para concluir esta sección podemos aseverar que el tiempo tiene incidencia en el número de arribos al centro de distribución. Dado que vemos diferencias dependiendo del día de la semana o el mes en que un camión llega desde la planta productora. Por ejemplo, los

Tabla 3.1: Correlaciones entre días de la semana.

	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado
Lunes	1,00	0,46	0,60	0,59	0,45	0,24
Martes		1,00	0,50	0,55	0,43	0,36
Miércoles			1,00	0,52	0,53	0,43
Jueves				1,00	0,63	0,58
Viernes					1,00	0,67
Sábado						1,00

Figura 3.6. Gráfico de control con límites  $3\sigma$ .

Fuente: Elaboración Propia

días sábados generalmente tienen un menor número de llegadas, y a su vez, Septiembre presenta un aumento en el número de arribos al centro de distribución.

### 3.2. AJUSTE DE LOS MODELOS DE ARRIBOS

En la sección anterior mostramos que existen efectos entre el número de llegadas al almacén y el tiempo de los arribos. Se observa una tendencia negativa, confirmando la presencia de efectos correspondientes al mes del año y día de la semana. Ahora consideramos cuatro modelos para ajustar nuestros datos, (1) el primero se refiere a una heurística que agrega aleatoriedad local respecto a los datos reales, (2) Un proceso Poisson no homogéneo, (3) Modelo EAR(p) introducido por Lawrance and Lewis (1980), y por último, (4) Un proceso de Markov MECO para llegadas no estacionarias y no renovables propuesto por Nelson and Gerhardt (2011). Estos modelos son definidos y validados en las siguientes subsecciones, comparándolos en términos de calidad de ajuste y validación del modelo de simulación.

Para estandarizar la explicación en el capítulo, definiremos la siguiente notación. Sea

$t$  el índice de tiempo en días, por lo tanto, el número de llegadas en el día  $t$  se denota como  $Y_t$ , mientras que  $Y_t^*$  representa el número de llegadas estimadas en el día  $t$ , para  $t = 1, 2, \dots, n$ . Los modelos son ajustados con 1.488 observaciones que representan las llegadas por hora desde el 3 de octubre de 2016 hasta el 30 de septiembre de 2017.

### 3.2. Modelo 1: Heurística.

El primer modelo es una heurística utilizada en De la Fuente et al. (2019), que tiene como lógica dar una pequeña variación respecto al promedio de un rango de longitud  $T$ . Para este caso,  $T$  fue considerado como una semana, es decir, 6 días.

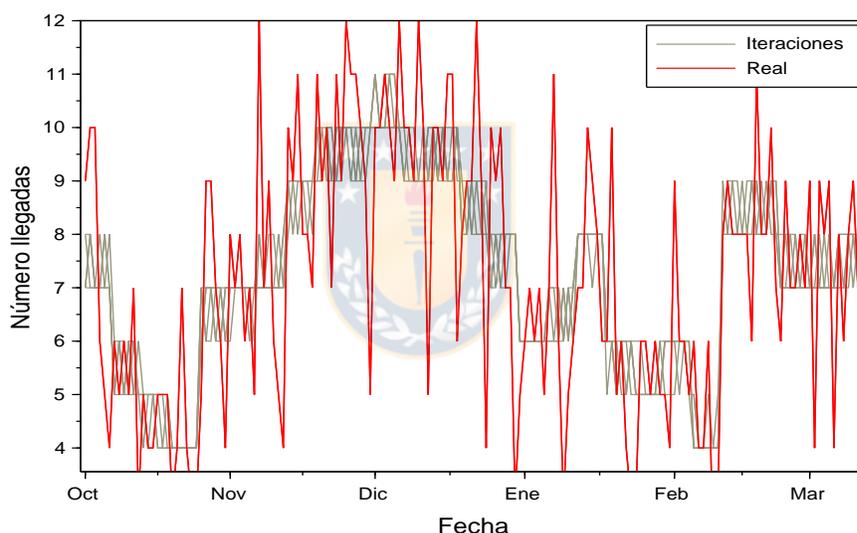


Figura 3.7: Ajuste Heurística.

Fuente: Elaboración Propia

El procedimiento consiste en determinar el promedio de llegadas diaria en cada semana  $j$ . En caso de que el promedio sea un número con decimales, la parte no entera de la tasa de arribos se representa con distribuciones discretas, de tal forma, que el número esperado de cada distribución semanal sea igual a la tasa promedio calculada. Por ejemplo, si la tasa es igual a 5,3 llegadas semanales (6 días hábiles) la distribución discreta asociada es *Random.Discrete*(5; 0,7; 6; 1) ya que  $5 \cdot 0,7 + 6 \cdot 0,3 = 5,3$ .

En la Figura 3.7 podemos ver como el comportamiento de dicha metodología no genera una gran variabilidad respecto al promedio. Lo anterior, es complementado con la

Figura 3.8, en donde el promedio de las iteraciones es muy cercano a los datos reales, generando una pequeña variación entre iteraciones.

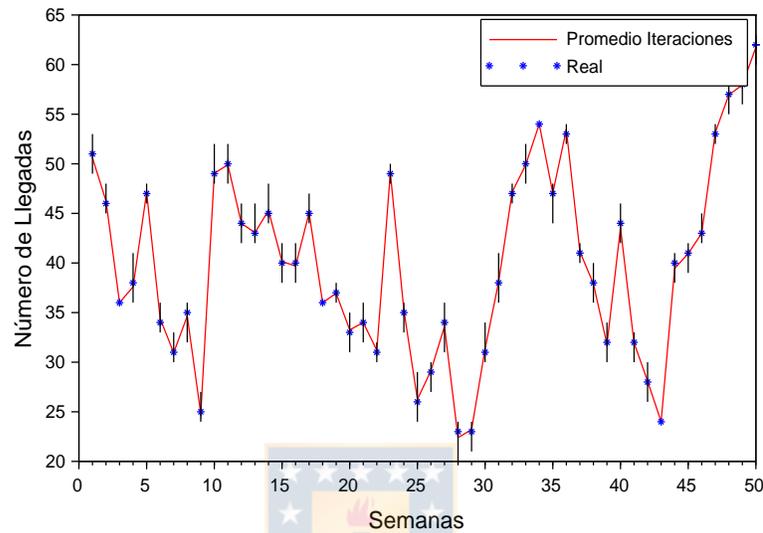


Figura 3.8: Comparación semanal entre los datos reales y el ajuste de la Heurística.

Fuente: Elaboración Propia

### 3.2. Modelo 2: Proceso de Poisson no Homogéneo (NHPP).

Para continuar con la explicación de los modelos ajustados, la metodología más utilizada en la literatura cuando un proceso presenta llegadas con una tasa que varía en el tiempo, es el proceso de Poisson no homogéneo.

En primer lugar, definiremos el modelo ajustado. Para que una colección de variables aleatorias  $N(t)$ ,  $t \geq 0$  se llame proceso de Poisson no homogéneo se deben cumplir las siguientes condiciones;

- $P(N(0) = 0) = 1$
- $N(t) \in \mathbb{N}, \forall t.$
- $N(t_1) \leq N(t_2)$ , si  $t_1 \leq t_2.$
- $[N(t_1) - N(0)], [N(t_2) - N(t_1)], [N(t_3) - N(t_2)],$  y así sucesivamente, para  $t_{n-1} < t_n$  (incrementos independientes).

- $\forall t > s \geq 0$  y  $c > 0$ , las v.a.  $[N(t) - N(s)]$  y  $[N(t+c) - N(s+c)]$  tienen la misma distribución de probabilidad (condición de estacionariedad).

Donde  $t_i$  es el índice de tiempo para el instante  $i$ . Podemos notar que la primera condición indica que la colección de variables aleatorias debe partir con un contador de ocurrencia igual a cero, es decir, para nuestro caso no debe haber ninguna llegada previa al tiempo 0. Las condiciones 3 y 4, se refieren a incrementos positivos e independientes. Mientras que la última condición expresa la estacionariedad en distintos tiempos  $t$  con un desfase de longitud  $c$ . En nuestro caso, definimos arbitrariamente una longitud de 3 días para cada intervalo (es decir se tienen dos tasas de llegadas por semana). Luego de eso obtuvimos las tasas de llegadas presentadas en la Figura 3.9.

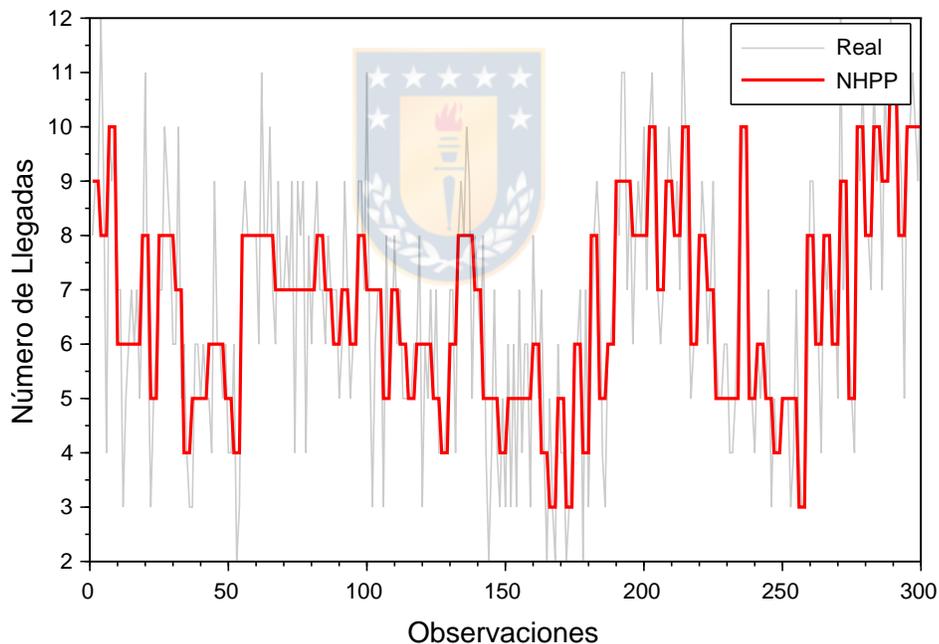


Figura 3.9: Ajuste NHPP.

Fuente: Elaboración Propia

Para obtener las diferentes iteraciones, creamos un modelo pequeño en software Simio LCC que solo genera llegadas de acuerdo al proceso de arribos mencionado en esta sección. Para esto ocupamos una tabla de tasas en donde ingresamos la tasa de arribos por hora durante un año obtenida de los datos históricos. Internamente, el software utiliza el procedimiento descrito en Joines and Roberts (2015, p.44). Luego, en la Figura 3.10

podemos observar la clara cercanía entre el promedio de las iteraciones respecto a los datos reales. Y adicionalmente, se observa una variabilidad visiblemente mayor respecto al proceso anterior.

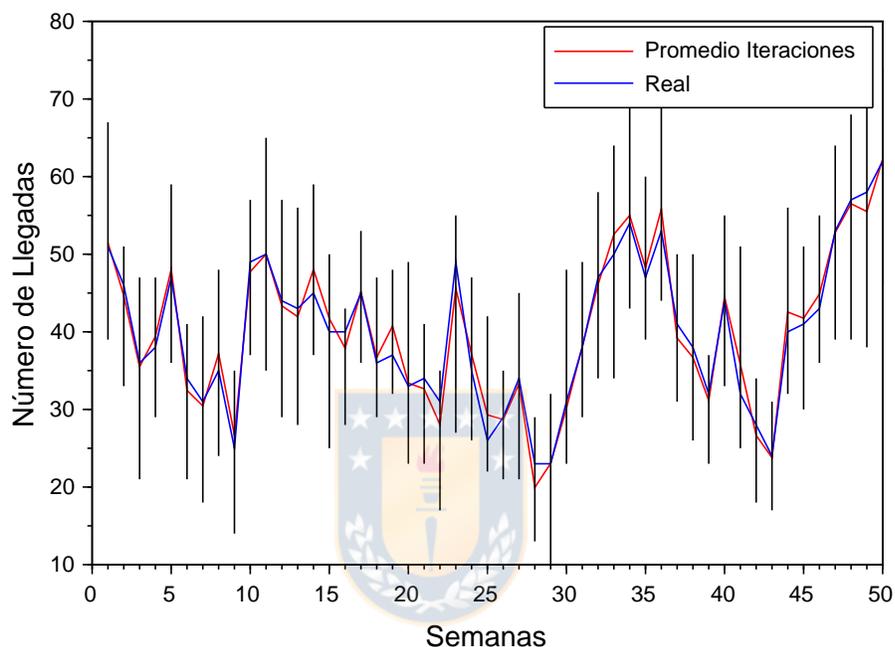


Figura 3.10: Comparación entre las llegadas generadas por NHPP y los datos reales.

Fuente: Elaboración Propia

### 3.2. Modelo 3: Exponencial autorregresivo EAR(p).

Los modelos exponenciales autoregresivos EAR(p) son introducidos por Lawrance and Lewis (1980), y luego extendidos en Lawrance and Lewis (1985) con una aplicación para modelar una serie de datos de velocidad del viento. En primera instancia veremos que comportamiento tienen los datos previos al ajuste para proponer un modelo adecuado, y luego obtendremos diferentes iteraciones que nos permitirán evaluar la estabilidad del modelo de simulación y el efecto en los resultados que provoca la variación del número de llegadas al almacén.

En primer lugar, notamos que por días feriados o festivos nuestra serie de datos no cuenta con todas las observaciones, esto puede afectar la identificación del modelo adecuado y por tanto su capacidad de ajuste. Por lo tanto, fue necesario aplicar un modelo

de regresión para obtener los valores faltantes. Se ajustó un modelo de efectos fijos desarrollado en el Anexo B el cual cuantifica los efectos del día de la semana, del mes del año y la interacción de ambos. Una vez completa la serie de tiempo, es necesario realizar una transformación de los datos debido a que esta metodología utiliza distribuciones exponenciales marginales estándar. De manera similar al procedimiento aplicado por Perera (2001), primero aplicamos una desestacionalización y luego una transformación que iguale tanto la media como la varianza. Para la primera transformación dividimos cada dato por el factor estacional correspondiente, es decir,  $Z_t = Y_t/\beta_k$ , con  $\beta_k$  como el factor estacional del día de la semana  $k$ . Luego aplicamos una segunda transformación a los datos,  $W_t = Z_t^{1,422}/3,461$ , obtenida en base al comportamiento de los datos. Finalmente la serie de tiempo está procesada para ser ajustada por un modelo EAR(p), cumpliendo las características de una serie estacionaria.

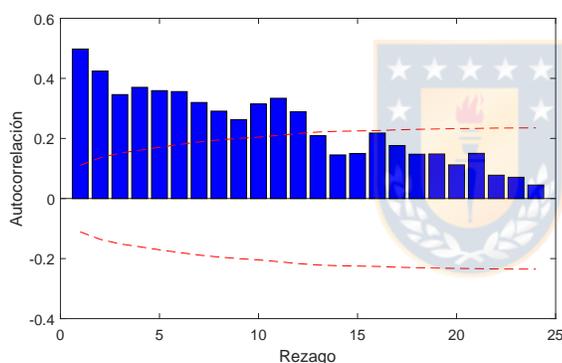


Figura 3.11: Función Autocorrelación.

Fuente: Elaboración Propia

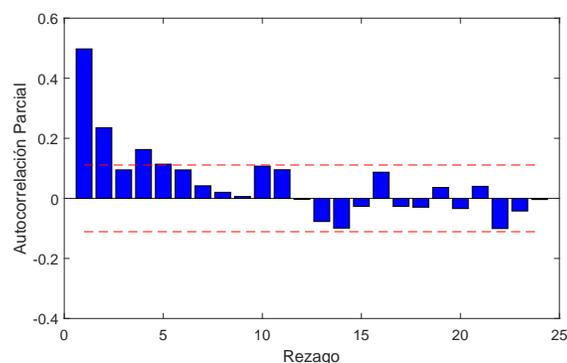


Figura 3.12: Autocorrelaciones parciales.

Fuente: Elaboración Propia

En primer lugar, analizaremos las funciones de autocorrelaciones de la serie  $W_t$ . Al ver las Figuras 3.11 y 3.12, podemos notar que el comportamiento de las funciones de autocorrelación (FAC) y la función de autocorrelaciones parciales (FAP) respectivamente, son similares a un comportamiento autorregresivo de orden 2, en donde tenemos una disminución aproximadamente gradual de la FAC y los dos primeros rezagos fuera de los límites permitidos en la FAC.

A continuación, veremos como se define un modelo EAR(2), el cual es de interés para esta investigación.

$$W_t = \left\{ \begin{array}{ll} \phi_1 W_{t-1} & w.p. \quad 1 - \phi_2 \\ \phi_2 W_{t-2} & w.p. \quad \phi_2 \end{array} \right\} + u_t \quad t = 0, 1, 2, \dots, T,$$

donde  $\phi_1$  y  $\phi_2$  se definen como medidas de dispersión con un ratio entre la covarianza y la varianza con 1 y 2 rezagos respectivamente. Mientras que  $u_t$  se define de la siguiente forma.

$$u_t = \begin{cases} 0 & w.p. & \frac{\phi_1}{1 + \phi_1 - \phi_2}, \\ v_t & w.p. & \frac{(1 - \phi_1)(1 - \phi_2)}{1 - \delta}, \\ \delta v_t & w.p. & \frac{(1 - \phi_2)(\phi_1 - \phi_2)^2}{(1 + \phi_1 - \phi_2)(1 - \delta)}, \end{cases}$$

donde  $\delta = (1 + \phi_1 - \phi_2)\phi_2$  y  $v_t$  son variables i.i.d. exponencialmente distribuidas con media  $\lambda$ . Ahora bien, la Tabla 3.2 muestra los parámetros obtenidos al ajustar el modelo a los datos.

Tabla 3.2: Parámetros del modelo EAR(2).

Item	Valor
$\lambda$	4,534
$Cov(W_t, W_{t-1})$	2,266
$Cov(W_t, W_{t-2})$	0,221
$\sigma$	2,139
$\phi_1$	0,4951
$\phi_2$	0,4237

Para continuar, al igual que en el ajuste del modelo de efectos fijos, debemos validar si es que los errores se distribuyen normal i.d.d. con media 0 y varianza igual a  $\sigma_{E,0}^2$ . Para lo anterior, realizaremos una prueba para analizar la independencia de los residuos conjuntamente con asumir que los errores están normalmente distribuidos con media 0.

Respecto a la aleatoriedad, aplicamos la prueba Box-Pierce la cual está basada en la suma de los cuadrados de los primeros 24 coeficientes de autocorrelación, y los resultados se muestran en la Tabla 3.3. Puesto que el *valor - p* es mayor o igual a 0,05, no se puede rechazar que la serie es aleatoria, con un nivel de confianza del 95 %. Lo anterior indica que no existe evidencia suficiente para aseverar que los residuos no son independientes. Concluyendo así el ajuste del modelo exponencial autorregresivo.

Finalmente, validamos el modelo ajustado y podemos obtener las iteraciones deseadas. En la Figura 3.13 vemos el ajuste presentado con datos diarios comparado con dichas réplicas. Mientras que en la Figura 3.14, mostramos el promedio y los rangos de las 20 iteraciones obtenidas por el modelo ajustado, comparado con los datos reales. Como podemos apreciar, el modelo captura la variabilidad de los datos.

Tabla 3.3: Prueba de Box Pierce.

	Valor
Estadístico de prueba	27,6739
valor - $p$	0,274036
Rechazar	No
Nivel de Confianza	95 %

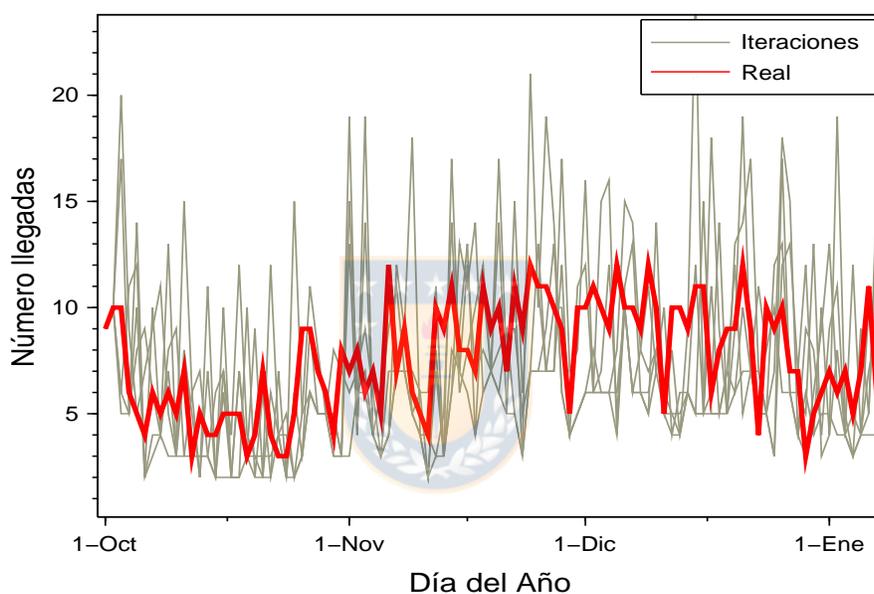


Figura 3.13. Iteraciones obtenidas en Octubre mediante el modelo EAR(2).

Fuente: Elaboración Propia

### 3.2. Modelo 4: Proceso Markov MECO (no estacionario y no renovable).

Como mencionamos en el Capítulo 1, los procesos de llegadas Poisson son ampliamente utilizados en la literatura actual, sin embargo; los arribos a sistemas generalmente presentan una dispersión considerablemente mayor que su media, por lo que no siempre se cumplen las características teóricas, por ejemplo, la independencia de llegadas y tiempo entre arribos distribuidos de forma exponencial.

Sriram and Whitt (1986) definen un ratio que permite medir dispersión e identificar si el proceso puede ser considerado como Poisson, el índice de dispersión para contadores

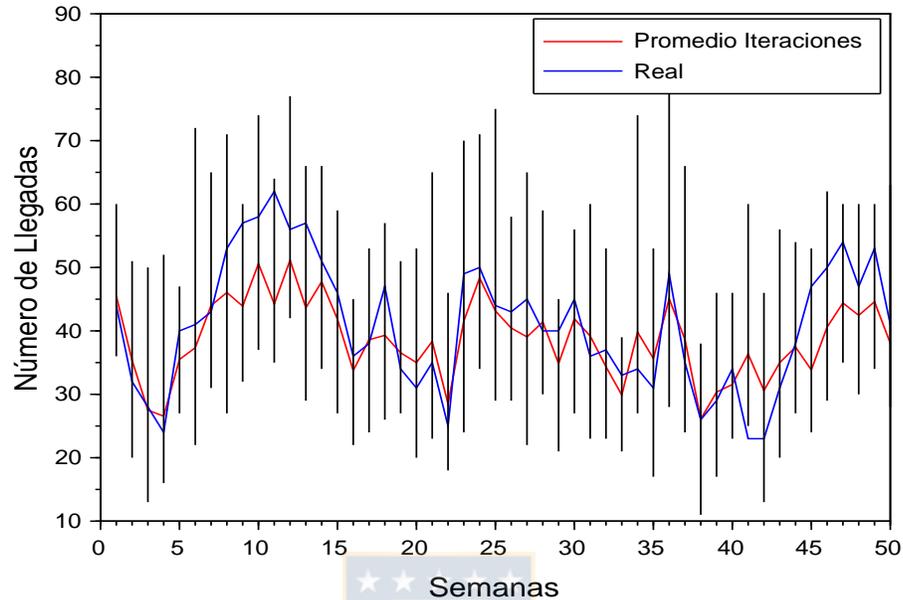


Figura 3.14: Comparación semanal entre las 20 iteraciones y los datos reales.

(IDC):

$$IDC = \lim_{t \rightarrow \infty} \frac{Var(N(t))}{E(N(t))} \quad (3.1)$$

Si  $IDC = 1$  el proceso es considerado como Poisson. Análogamente, para un proceso de renovación en equilibrio se cumple que  $IDC = cv^2$ . Ahora bien, al calcular dicho índice en nuestro proceso, obtenemos un valor mayor a 1 (para más detalles, ver De la Fuente et al. (2019, p.1746)), lo que indica mayor variabilidad que un proceso Poisson. Sumado a que las llegadas no son independientes, como lo hemos mostrado al inicio de este capítulo, parece ser inadecuado considerar el proceso de arribos como uno del tipo Poisson no homogéneo.

Así, Nelson and Gerhardt (2011) propusieron un método para simular una secuencia de tiempos entre llegadas tal que el proceso de conteo de llegadas es no estacionario no renovable (NSNR) asistidos por un proceso de arribos de Markov. El método consiste en generar tiempos entre arribos con un proceso de llegadas Markov Mix de Erlangs de Orden Común (Markov MECO) presentado por Johnson (1998). Dicho proceso es un caso especial de los procesos de arribos de Markov que incluye un proceso de renovación compuesto por un mix de dos distribuciones Erlangs de orden común (MECO), para procesos de no renovación presentado por Johnson and Taaffe (1989).

Primero, MECO es una distribución de tipo "Fase" que se definen como la distribución de probabilidad del tiempo hasta la absorción de un proceso de Markov con un número finito de estados transitorios y un estado de absorción. La ventaja de usar esta distribución radica en su conveniencia para momentos coincidentes, lo que permite replicar la mayoría de los procesos de renovación especificando sus primeros tres momentos. Sobre la base de esta distribución, el Markov-MECO extiende la no renovación de llegadas controlando la dependencia de tiempos entre llegadas usando un proceso de Markov como se muestra en la Figura 3.15. Aquí, la probabilidad de transición determina la distribución Erlang que generará el próximo tiempo entre llegadas, por lo tanto, la generación de un tiempo entre llegadas depende del estado desde el cual se generó el tiempo previo a la llegada.

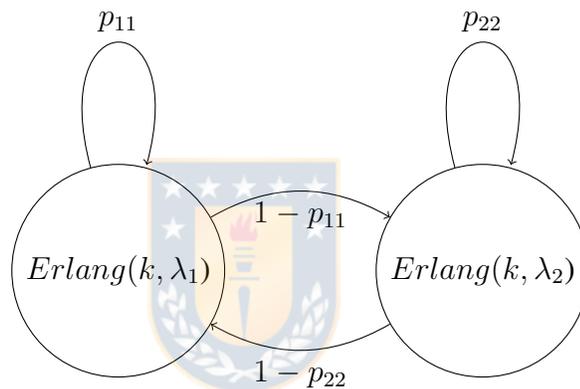


Figura 3.15: Proceso Markov MECO.

Para entender el método, realizaremos algunas definiciones. El conjunto de tiempos entre arribos no negativos es definido como  $\{X_n, n \geq 1\}$ , siendo  $S_n$  el tiempo de el  $n$ -ésimo arribo (considerando  $S_0 = 0$  y  $S_n = \sum_{i=1}^n X_i$ , para  $n = 1, 2, \dots$ ), y  $N(t)$  como el número de arribos que han ocurrido antes del tiempo  $t$ , con  $N(t) = \max\{n \geq 0 : S_n \leq t\}$ . Adicionalmente, asumimos  $N(t)$  inicializado en equilibrio. A continuación definiremos el algoritmo mencionado y luego discutiremos los detalles del proceso base utilizado para generar las llegadas (Markov MECO).

Primero para el algoritmo, los autores lo generalizan en base a uno previamente propuesto por Gerhardt and Nelson (2009) para generar procesos NSNR. Se define  $r(t), t \geq 0$  como la tasa de arribos no negativa e integrable, luego  $R(t) = \int_0^t r(s) ds$ , donde  $s \in \mathbb{R}^+$  y  $R^{-1} \equiv \inf\{t : R(t) \geq s\}$ .

Una vez definidas las variables, el algoritmo es el siguiente:

- Paso 1: Definir  $V_0 = 0$  y índice contador  $n=1$ . Generar  $S_1$  y fijar  $V_1 = R^{-1}(S_1)$ .
- Paso 2: Obtener tiempo entre arribos  $W_n = V_n - V_{n-1}$ .

- Paso 3: Definir  $n = n + 1$ . Generar  $X_n$ . Definir  $S_n = S_{n-1} + X_n$  y  $V_n = R^{-1}(S_n)$ .
- Paso 4: Ir a Paso 2.

En segundo lugar respecto al proceso base, este es utilizado para generar  $X_n$  en el algoritmo anterior. Dado que la tasa de arribos para el proceso base Markov MECO debe ser 1 (por lo tanto la media debe ser 1), solo se requieren tres entradas al modelo generador de tiempos entre arribos, estas son, (1) especificar una tasa de arribos variable en el tiempo  $r(t)$ , (2) un coeficiente de variación estimado al cuadrado ( $cv^2$ ) y (3) la autocorrelación del primer retraso de nuestros datos ( $\rho_1$ ).

Ahora bien, una vez definida nuestra metodología de generación de llegadas procedemos a calcular las entradas necesarias para aplicarla. En primer lugar, en la Tabla 3.4 presentamos el cálculo de los parámetros de interés, donde la media de las llegadas acumuladas por periodo  $t$  está dada por  $\hat{\Lambda}(t)$ , mientras que la varianza esta dada por la Ecuación (3.2), siendo  $k$ , el número de observaciones por cada realización.

$$V(t) = \frac{1}{k-1} \sum_{j=1}^k (C_j(t) - \hat{\Lambda}(t))^2 \quad (3.2)$$

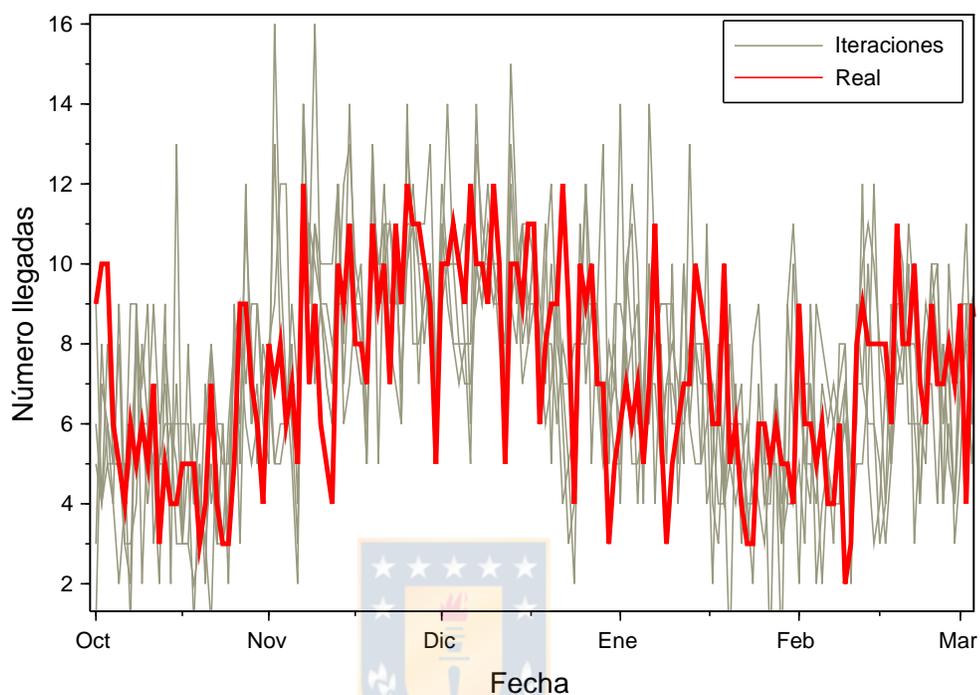
Tabla 3.4: Ejemplo del cálculo de parámetros para días martes en temporada alta.

$t$	#	$C_1(t)$	#	$C_2(t)$	...	#	$C_{14}(t)$	$\hat{\Lambda}(t)$	$V(t)$	$V(t)/\hat{\Lambda}(t)$	$r(t)$
1	8	8	10	5	...	8	8	7.059	2.184	0.309	7.059
2	9	17	9	14	...	7	15	15.059	5.184	0.344	8
3	9	26	10	24	...	8	23	22.882	8.86	0.387	7.824
4	12	38	7	31	...	6	29	30.118	20.485	0.68	7.235
5	9	47	7	38	...	7	36	37.176	29.779	0.801	7.059
6	4	51	3	41	...	5	41	42.647	26.993	0.633	5.471

Finalmente, para obtener el coeficiente de variación ( $cv$ ) lo estimamos a partir de la Ecuación (3.3), donde  $m$  es el número de periodos  $t$  que comprenden cada realización.

$$\hat{\sigma} = \frac{1}{m} \sum_{i=1}^m \frac{V(t_i)}{\hat{\Lambda}(t_i)} = cv \quad (3.3)$$

Ahora bien, dado que tenemos las entradas necesarias para el modelo tratado en esta sección, procedemos a aplicar la metodología mediante la herramienta disponible en Nelson and Gerhardt (2011). En la Figura 3.16 vemos el ajuste presentado con datos diarios comparado con dichas réplicas. Mientras tanto, en la Figura 3.17, mostramos el promedio

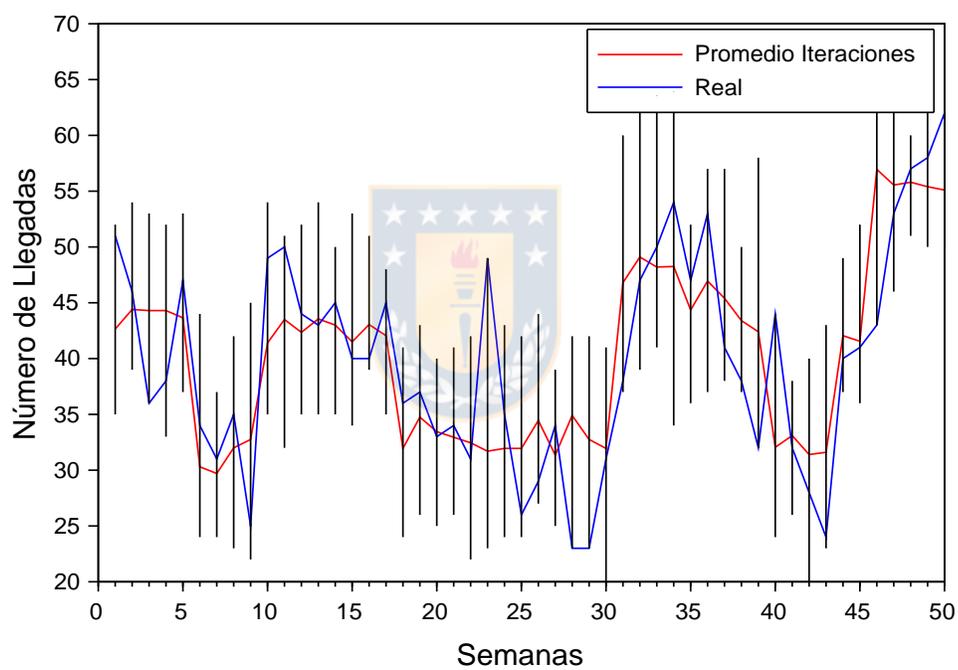


**Figura 3.16.** Llegadas reales y réplicas obtenidas mediante el método de Nelson and Gerhardt (2011).

Fuente: Elaboración Propia

y los rangos de las 20 iteraciones obtenidas por el modelo ajustado en comparación con los datos reales. Como podemos apreciar el modelo captura la variabilidad de los datos.

Finalmente, en este capítulo hemos desarrollamos las metodologías de ajuste de datos o generación de réplicas utilizando los datos de las llegadas al centro de distribución. Se obtuvo 20 réplicas para cada metodología las cuales serán ingresadas como entrada al modelo de simulación descrito en el Capítulo 2. De esta forma, obtendremos los resultados de esta investigación para así poder discutir el efecto en estos y la estabilidad del modelo de simulación.



**Figura 3.17.** Comparación llegadas reales con Modelo Markov MECO.

Fuente: Elaboración Propia

## 4. RESULTADOS Y DISCUSIÓN

En este capítulo, presentamos los resultados de esta investigación junto con la propuesta de un modelo de arribos en base a criterios de selección. Primero, realizaremos pruebas de medias entre los resultados obtenidos por el modelo de simulación y los datos reales, para luego discutir y comparar el comportamiento de los resultados tanto individuales como en conjunto de los modelos de llegadas abordados en capítulos anteriores. Los cinco criterios de selección utilizados para comparar los procesos de arribos son; (1) validación estadística entre la capacidad promedio obtenida por los modelos de arribos y la capacidad promedio real, (2) validación entre la capacidad promedio obtenida por los modelos de arribos y la *capacidad promedio modelada*<sup>1</sup>, (3) La variabilidad de las llegadas y resultados, (4) dependencia o correlación entre los arribos, y finalmente, (5) la exactitud de los resultados respecto a los valores reales.

Para comenzar, es necesario mencionar que el modelo es válido y representa de forma adecuada los procesos operativos del centro de distribución. Para esto, en el Anexo C se desarrolla la validación del modelo de simulación realizada por De la Fuente et al. (2019), en donde se indica que el modelo es apto para obtener resultados mediante su utilización.

Para continuar, dado que el modelo de simulación está validado se utiliza como medio para obtener resultados variando el número de llegadas de acuerdo a los diferentes modelos de arribos desarrollados en la Sección 3.2; Heurística (HEU), Proceso Poisson no Homogéneo (NHPP), Exponencial Autorregresivo (EAR) y Proceso de Markov-MECO (MECO).

El primer criterio de selección se refiere a la validación estadística entre la capacidad promedio obtenida por los modelos de arribos y la capacidad promedio real. Para esto, se realizó una prueba de medias bajo el mismo procedimiento presentado en el Anexo C, considerando la capacidad del almacén obtenida entre el promedio de las 20 iteraciones de cada modelo de arribos, y los datos semanales reales (Ver Figura 4.1). Los resultados son mostrados en la Tabla 4.1. Puesto que los intervalos de las diferencias de las medias del modelo NHPP no contienen el valor de 0, podemos rechazar la hipótesis nula indicando

---

<sup>1</sup>La capacidad de almacenamiento modelada, se obtiene mediante el modelo de simulación considerando los arribos históricos (reales) como proceso de entrada.

Tabla 4.1: Pruebas de medias entre la capacidad de la bodega obtenida mediante el modelo de simulación y los datos reales.

	HEU	NHPP	EAR	MECO
LSC	233,68	-8,47	51,09	48,23
LIC	-139,58	-172,59	-126,96	-139,74
Nivel de Confianza	95 %	95 %	95 %	95 %
¿Rechazar?	No	Sí	No	No

LSC - LIC: Límite superior e inferior del intervalo de confianza de la diferencia de medias.

que existe evidencia suficiente para aseverar una diferencia estadísticamente significativa entre las medias de las dos muestras de datos, con un nivel de confianza del 95,0 %.

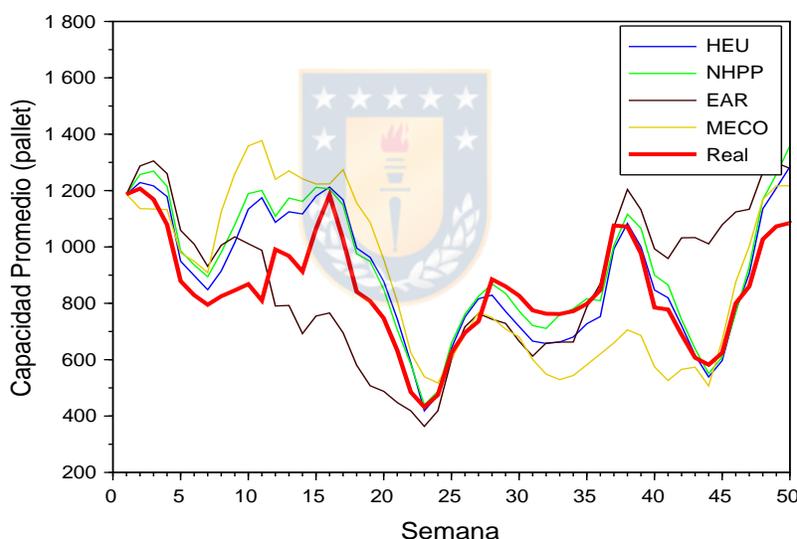


Figura 4.1: Comparación semanal entre las capacidades promedio obtenidas por cada modelo de arribos.

Los resultados del primer criterio de selección coinciden con lo mencionado por Nelson (2013), en donde a pesar de que los procesos Poisson estacionarios y no estacionarios son ampliamente utilizados en la práctica de la simulación, deben cumplir que los tiempos entre arribos deben ser distribuidos de forma exponencial y las llegadas deben ser independientes, y cómo hemos mencionado en el Capítulo 3 estas características no se cumplen, perjudicando el ajuste NHPP.

El segundo criterio de selección es una validación entre la capacidad promedio obtenida por los modelos de arribos y la *capacidad promedio modelada*. Al observar la

Figura 4.1, podemos notar un comportamiento similar entre los resultados de cada modelo de arribos y la *capacidad promedio modelada*, incluso presentando una mayor similitud que con la capacidad real (efecto más notorio entre las semanas 5 y 15). Es lógico pensar que existe una incidencia de la estructura del modelo sobre los resultados creando un sesgo para nuestros objetivos de investigación, en donde buscamos evaluar los impactos de cada metodología de arribos. Para capturar dicha problemática, realizamos una comparación respecto a la *capacidad promedio modelada* y no respecto a la capacidad de almacenamiento real, de esta forma, logramos obtener el impacto más explícito de cada modelo de arribos en los resultados de la simulación.

La Tabla 4.2, muestran la validación estadística de las metodologías de arribos. Las mayores diferencias respecto a la previa validación se refieren a que el modelo NHPP en esta ocasión no rechaza la hipótesis nula, indicando que la variabilidad propia del modelo puede afectar el rechazo o aprobación de un modelo de arribos.

Tabla 4.2: Pruebas de medias entre la capacidad promedio modelada y las capacidades promedio obtenidas por los modelos de arribos.

	HEU	NHPP	EAR	MECO
LSC	101,46	69,56	67,35	125,16
LIC	-80,71	-103,66	-128,45	-79,71
Nivel de Confianza	95 %	95 %	95 %	95 %
¿Rechazar?	No	No	No	No

LSC - LIC: Límite superior e inferior del intervalo de confianza de la diferencia de medias.

El tercer criterio de selección se refiere a la variabilidad de las llegadas ajustadas y de los resultados obtenidos. Es importante notar el comportamiento diferente entre las distintas metodologías de arribos, y su capacidad de representar la variabilidad en las llegadas de camiones al centro de distribución. Para comenzar, definimos  $A$  como la amplitud promedio mediante la Ecuación (4.1), donde  $H_j^{max}$  es el número máximo de llegadas en la semana  $j$  considerando las 20 iteraciones de un modelo de arribos, y de forma análoga  $H_j^{min}$  como el mínimo.

$$A_{max} = \frac{H_j^{max} - H_j^{min}}{20} \quad (4.1)$$

La amplitud máxima en conjunto con la desviación estándar de cada modelo de arribos (Ver Tabla 4.3) buscan capturar la variabilidad promedio entre las llegadas generadas por cada metodología de arribos. Como es de esperar, la menor variación se presenta en el modelo HEU, debido a su metodología de cálculo basada en el nivel promedio de llegadas por semana, que genera una variabilidad local. Por su parte, podemos observar que

Tabla 4.3: Variabilidad presente en las llegadas generadas por cada modelo de arribos.

	Amplitud Promedio	Desviación Estándar
HEU	3.23	0.94
NHPP	22.75	5.95
EAR	33.10	8.72
MECO	17.87	4.63

la variabilidad más amplia se ve en los modelos NHPP y EAR. Lo anterior, debido a que ambos representan dicho concepto mediante muestreos de funciones exponenciales, generando una variabilidad mayor a la inherente en el proceso de arribos real. Esto verifica lo mencionado por Ibrahim et al. (2012) sobre procesos Poisson, indicando que los casos reales generalmente tienen una dispersión considerablemente mayor respecto a la media; por lo tanto, no son recomendables dicho tipo de procesos.

Por lo tanto, respecto a variabilidad el modelo Markov-MECO resulta recomendable, debido a que captura la volatilidad de los arribos mediante el índice de dispersión para intervalos (*IDI*) (Ver Ecuación (4.2)), propuesto por Gusella (1991) y luego utilizado por Nelson and Gerhardt (2011). Donde  $\rho_j$  es la autocorrelación con rezago  $j$  de los tiempos entre arribos y  $cv$  el coeficiente de variación estimado mediante la Ecuación 3.3. Así, el  $IDC = IDI$  capturan tanto la variabilidad ( $cv^2$ ) como la dependencia ( $1 + 2 \sum_{j=1}^{\infty} \rho_j$ ) del proceso de arribos en análisis.

$$IDC = IDI \equiv \lim_{x \rightarrow \infty} \frac{Var\{S_n\}}{n \cdot E^2\{X_2\}} = cv^2(1 + 2 \sum_{j=1}^{\infty} \rho_j) \quad (4.2)$$

Es evidente, que cuando existe una mayor estocasticidad en las entradas al modelo, obtendremos una mayor variabilidad en las salidas de este. Lo anterior, se ve respaldado al observar la amplitud de los resultados de la capacidad promedio del almacén en la Figura 4.2. Además, podemos notar una vez más la menor variabilidad del modelo Markov-MECO respecto a los modelos NHPP y EAR.

Un tema importante que permite discriminar entre un ajuste adecuado, es la capacidad de parametrizar las dependencias de los arribos. Ibrahim et al. (2016) menciona que se debe tener especial cuidado en los procesos de llegadas correlacionados. En nuestro caso, por definición los modelos EAR representan procesos correlacionados mientras que el Markov-MECO lo realiza mediante el factor ( $1 + 2 \sum_{j=1}^{\infty} \rho_j$ ) mencionado recientemente. Por otro lado, los modelos HEU y NHPP solo generaron variabilidad en base a tasas arbitrarias obtenidas desde los datos históricos, es decir, no consideran correlación en sus iteraciones haciendo menos confiables sus ajustes.

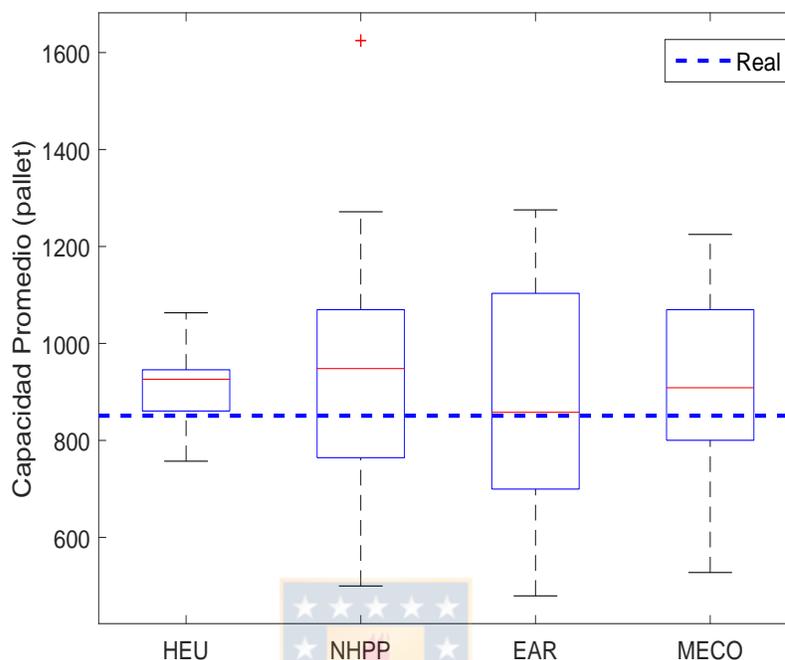


Figura 4.2: Diagrama de cajas de la capacidad del promedio del almacén.

Como mencionamos en el primer capítulo, Bause et al. (2009) menciona que los modelos ARMA/ARTA tienen mejor desempeño que los procesos de arribos de Markov; sin embargo, requieren de un ajuste más elaborado y están restringidos a trabajar bajo supuestos de normalidad. La recomendación propuesta en esta investigación es complementaria a lo indicado por Bause, debido a que las dos metodologías abordadas en esta investigación similares a las mencionadas (EAR y MECO) son recomendadas para llegadas correlacionadas, sin embargo, los modelos EAR presentaron problemas al representar la variabilidad del proceso de arribos.

Finalmente, el último criterio de elección se refiere a la exactitud de los resultados respecto a los valores reales. Se espera que un modelo de arribos ajuste y represente las características del proceso de llegadas, esto es, que sea capaz de generar arribos con la dependencia y variabilidad inherente del proceso respecto a valores reales. En base a lo anterior, se observa en el diagrama de cajas y bigotes, en la Figura 4.2, qué modelos tienen una variabilidad cercana a la capacidad promedio real. Se aprecia que los modelos NHPP, EAR y MECO tienden a distribuirse correctamente respecto a la capacidad promedio real, esto es, que dicho valor esté comprendido entre el segundo y tercer cuantil de las capacidades promedio obtenidas por las 20 iteraciones de cada modelo.

Recapitulando, en la Tabla 4.4 podemos observar un resumen de los criterios de elección. Los resultados sugieren que el modelo de Markov-MECO es la metodología que presentó el mejor desempeño en base a los diferentes criterios definidos en esta investigación.

Tabla 4.4: Resumen elección de modelo de arribos.

Modelo	Validación 1	Validación 2	Amplitud	Dependencia	Exactitud
HEU	✓	✓			
NHPP		✓			✓
EAR	✓	✓		✓	✓
MECO	✓	✓	✓	✓	✓

Para continuar, mencionamos el impacto económico que puede causar escoger un modelo de arribos inapropiado o no recomendado. Para esto primero debemos definir la capacidad de almacenaje de la bodega del centro de distribución, la cual consideramos como la máxima capacidad observada durante el año en análisis. Esto debido a que por políticas corporativas de la empresa no se pueden almacenar productos fuera del almacén. La Figura 4.3, muestra como los máximos de cada modelo son mayores a la capacidad máxima real, esto es esperable dado que la lógica de cada modelo es agregar una dispersión o estocasticidad respecto a los valores reales, por ende, se espera obtener dichos resultados.

Al cuantificar el impacto económico que significa elegir una metodología de arribos para nuestro caso estudio, comparamos las diferencias máximas de la capacidad entre los resultados de cada modelo, observamos que la mayor desviación está comprendida entre los modelos de arribos EAR y HEU. De estas diferencias obtenemos un promedio de 190,01 pallets piso<sup>2</sup>(cada pallet piso es igual a un 1  $m^2$  de bodega) considerando los cuatro cuantiles de cada metodología. Ahora bien, si el costo por metro cuadrado es de CLP 75.253<sup>3</sup> para construcciones de almacenes de hormigón con techumbre de acero, y asumimos que la construcción extra del almacén no incluye pasillos, obtenemos un impacto económico de CLP  $75,253 \cdot 190,01 = \text{CLP } 14,298,514$ .

La segunda medida de desempeño del modelo de simulación, es el tiempo de elaboración de cargas (TEC), en la Tabla 4.5 podemos observar los resultados obtenidos en base a cada modelo de arribos. Existe una variación despreciable debido a que el tiempo de elaboración de cargas depende directamente de la demanda diaria de cada familia de productos, por lo tanto, mientras no existan quiebres de stock ocasionados por desabas-

<sup>2</sup>El concepto de pallet piso se refiere a la capacidad efectiva que utilizan los productos dentro del almacén.

<sup>3</sup>Tabla de costos unitarios, categoría otras construcciones. [www.minvu.cl](http://www.minvu.cl), Enero 2020.

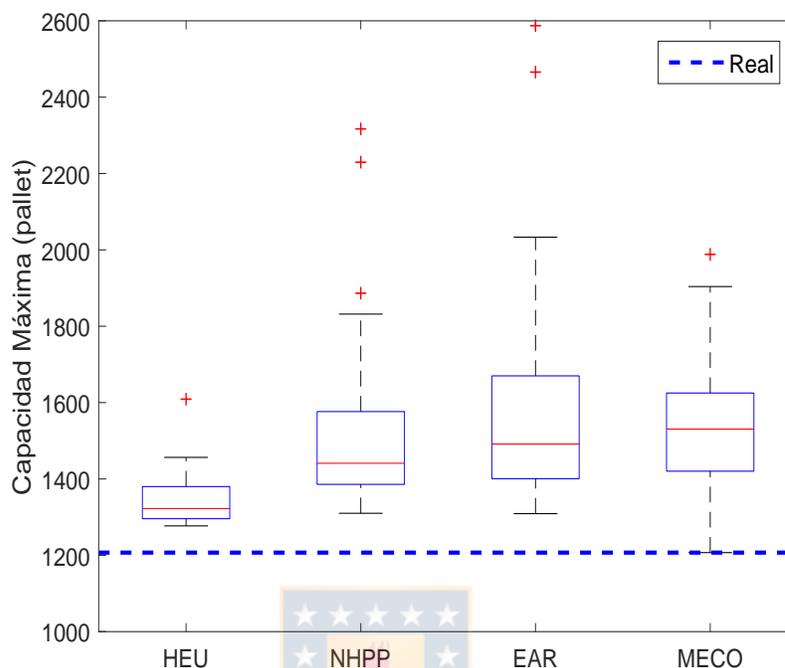


Figura 4.3: Diagrama de cajas de la capacidad máxima del almacén.

tecimiento, las llegadas de camiones al centro de distribución no tendrán mayor impacto en el TEC. Lo anterior ocurre, debido a que el modelo de simulación tiene programado controlar y evitar los quiebres de stock mediante una lógica de requerimientos o pedidos desde el centro de distribución a los proveedores. Dicha programación se realiza mediante un proceso de Simio LLC en donde se crea la relación entre la estocasticidad de las llegadas y los productos que estas contienen de tal forma de evitar quiebres de stock dentro de lo posible.

Los resultados del TEC, nos permiten aseverar que es relevante escoger indicadores claves para evaluar distintas metodologías de arribos, debido a que se corre el riesgo de obtener impactos implícitos propios del proceso e independientes de la entrada en análisis.

En cuanto a las suposiciones o debilidades presentes en esta investigación. Mencionamos la incertidumbre debido a la falta de datos históricos. Al momento de realizar los ajustes de los modelos de arribo nos encontramos con la falta de realizaciones (sólo contamos con una correspondiente a un periodo de un año), por lo mismo, fue necesario crear iteraciones de cada metodología de arribos y así obtener llegadas factibles. Lo anterior implica una mayor incertidumbre de los ajustes aplicados. Con más realizaciones, los ajustes de las metodologías de arribos serían aún más confiables.

Tabla 4.5: Resultados del tiempo de elaboración de cargas en Picking.

Modelo	Iteraciones	
	Promedio	Desviación Estándar
HEU	6,6288	0,0028
NHPP	6,6250	0,0120
EAR	6,5824	0,0316
MECO	6,6265	0,0091

En síntesis, este capítulo presentó un argumento para proponer una metodología de arribos, en donde al comparar los resultados obtenidos en esta investigación estos sugieren que el modelo de Markov MECO presentó el mejor desempeño en base a diferentes criterios. Además, comentamos el impacto económico que puede causar la elección de un modelo de arribos en base a nuestro caso estudios. Y finalmente, mencionamos la pertinencia de escoger indicadores adecuados para la evaluación de metodologías de arribos en conjunto con explicar la mayor debilidad o incertidumbre de las metodologías utilizadas en esta investigación.



## 5. CONCLUSIÓN

Este trabajo evalúa diferentes metodologías que buscan parametrizar los arribos de un centro de distribución, para luego analizar los efectos en métricas de interés (capacidad de almacenamiento y tiempo de elaboración de cargas). Dicha evaluación es asistida por un modelo de simulación elaborado mediante software Simio LLC, considerando un caso real basado en la llegada de camiones a un centro de distribución de consumo masivo.

La comparación de métodos de arribos se basa en cinco criterios; (1) validación estadística entre la capacidad promedio obtenida por los modelos de arribos y la capacidad promedio real, (2) validación entre la capacidad promedio obtenida por los modelos de arribos y la *capacidad promedio modelada*, (3) La variabilidad de las llegadas y resultados, (4) dependencia o correlación entre los arribos, y finalmente, (5) la exactitud de los resultados respecto a los valores reales.

Los resultados muestran dos puntos importantes a considerar. Primero, referido a la elección de un modelo de arribos en conjunto con el comportamiento de las diferentes metodologías utilizadas, y segundo, las consecuencias e impactos que puede tener escoger un modelo inadecuado para generar el proceso de arribos. Respecto al primer punto se presentaron dos metodologías, Heurística y Poisson no Homogénea, que no capturan la dependencia de los arribos ya que simplemente generan volatilidad respecto parámetros deterministas obtenidos desde los datos históricos. Lo preocupante, es que los casos reales generalmente las llegadas son correlacionadas y los modelos NHPP son ampliamente utilizados en la literatura, lo que representa un riesgo en las conclusiones que se obtienen. Adicionalmente, los modelos NHPP y EAR generan la volatilidad de sus ajustes en distribuciones exponenciales, sin embargo, a pesar de que se ajustaron correctamente la variabilidad es mayor a la inherente en el proceso de arribos en análisis. Respecto al modelo de arribos, Markov MECO presentó el mejor desempeño en base a los diferentes criterios de evaluación demostrando que ajusta correctamente la variabilidad y dependencia de los datos.

Posteriormente, se realizó una estimación del impacto económico que pueden causar la utilización de diferentes métodos de arribos, alcanzando un total de CLP 14,298,514, demostrando numéricamente la importancia que representan las entradas a los modelos

de simulación y la variación en la práctica que puede provocar en sus conclusiones. Conjuntamente, debido a los resultados obtenidos para el tiempo de elaboración de cargas, se sugiere identificar los indicadores apropiados para realizar evaluaciones en base a modelos de simulación, dado que deben estar directamente relacionados con las medidas que deseamos evaluar, de esta forma, capturamos correctamente la volatilidad de las entradas.

Finalmente, este trabajo logró comparar los efectos de las llegadas en los resultados de un sistema de servicio mediante el uso de la simulación. Dado lo anterior, se pudo estimar un impacto económico ante una eventual elección de un modelo inapropiado, revelando la importancia que significa la temática abordada por esta investigación.



## REFERENCIAS

- Aldor, S., Feigin, P., and Mandelbaum, A. (2010). Workload forecasting for a call center: Methodology and a case study. *Annals of Applied Statistics*, 3.
- Bause, F., Buchholz, P., and Krieger, J. (2009). A comparison of markovian arrival and arma/arta processes for the modeling of correlated input processes. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, pages 634–645.
- Brown, L., Gans, N., Mandelbaum, A., Sakov, A., Shen, H., Zeltyn, S., and Zhao, L. (2005). Statistical analysis of a telephone call center: A queueing-science perspective. *Journal of the American statistical association*, 100(469):36–50.
- Caridad, J. M. et al. (1998). *Econometría. Modelos econométricos y series temporales con los paquetes uTSP y TSP*, volume 1. Reverté.
- Channouf, N. and L'Ecuyer, P. (2010). Fitting a normal copula for a multivariate distribution with both discrete and continuous marginals. pages 352 – 358.
- Channouf, N. and L'Ecuyer, P. (2012). A normal copula model for the arrival process in a call center. *International Transactions in Operational Research*, 19.
- Channouf, N., L'Ecuyer, P., Ingolfsson, A., and Avramidis, A. N. (2007). The application of forecasting techniques to modeling emergency medical system calls in calgary, alberta. *Health Care Management Science*, 10(1):25–45.
- Cheng, R. (2017). History of input modeling. In *2017 Winter Simulation Conference (WSC)*, pages 181–201.
- De la Fuente, R., Gatica, J., and Smith, R. L. (2019). A simulation model to determine staffing strategy and warehouse capacity for a local distribution center. In *2019 Winter Simulation Conference (WSC)*, pages 1743–1754.
- Freimer, M. and Schruben, L. (2002). Collecting data and estimating parameters for input distributions. *Proceedings of the Winter Simulation Conference*, 1:392–399 vol.1.

- Gerhardt, I. and Nelson, B. (2009). Transforming renewal processes for simulation of nonstationary arrival processes. *INFORMS Journal on Computing*, 21:630–640.
- Gupta, A. and Parzen, E. (2005). Input modeling using quantile statistical methods. pages – 736.
- Gusella, R. (1991). Characterizing the variability of arrival processes with indexes of dispersion. *IEEE Journal on Selected Areas in Communications*, 9(2):203–211.
- He, B., Liu, Y., and Whitt, W. (2016). Staffing a service system with non-poisson non-stationary arrivals. *Probability in the Engineering and Informational Sciences*, 30(4):593–621.
- Ibrahim, R., Regnard, N., L’Ecuyer, P., and Shen, H. (2012). On the modeling and forecasting of call center arrivals. In *Proceedings of the 2012 Winter Simulation Conference (WSC)*, pages 1–12.
- Ibrahim, R., Ye, H., L’Ecuyer, P., and Shen, H. (2016). Modeling and forecasting call center arrivals: A literature survey and a case study. *International Journal of Forecasting*, 32(3):865 – 874.
- Jaoua, A., L’Ecuyer, P., and Delorme, L. (2013). Call-type dependence in multiskill call centers. *Simulation*, 89(6):722–734.
- Johnson, M. A. (1998). Markov meco: a simple markovian model for approximating nonrenewal arrival processes. *Communications in Statistics. Stochastic Models*, 14(1-2):419–442.
- Johnson, M. A. and Taaffe, M. R. (1989). Matching moments to phase distributions: Mixtures of erlang distributions of common order. *Communications in Statistics. Stochastic Models*, 5(4):711–743.
- Joines, J. A. and Roberts, S. D. (2015). *Simulation modeling with SIMIO: a workbook*. Simio LLC Pittsburgh.
- Kim, S.-H. and Whitt, W. (2014). Choosing arrival process models for service systems: Tests of a nonhomogeneous poisson process. *Naval Research Logistics (NRL)*, 61(1):66–90.
- Kuhl, M. E., Steiger, N. M., Lada, E. K., Wagner, M. A., and Wilson, J. R. (2006). Introduction to modeling and generating probabilistic input processes for simulation. In *Proceedings of the 2006 Winter Simulation Conference*, pages 19–35.

- Kurowicka, D. and Cooke, R. M. (2002). The vine copula method for representing high dimensional dependent distributions: application to continuous belief nets. In *Proceedings of the Winter Simulation Conference*, volume 1, pages 270–278 vol.1.
- Law, A. M. (2008). How to build valid and credible simulation models. In *Simulation Conference, 2008. WSC 2008. Winter*, pages 39–47. IEEE.
- Law, A. M. (2013). A tutorial on how to select simulation input probability distributions. In *2013 Winter Simulations Conference (WSC)*, pages 306–320.
- Law, A. M. (2015). *Simulation Modeling & Analysis*. McGraw-Hill Education.
- Lawrance, A. and Lewis, P. (1980). The exponential autoregressive-moving average earma( p,q ) process. *Journal of the Royal Statistical Society Series B-Methodological*, 42:150–161.
- Lawrance, A. J. and Lewis, P. A. W. (1985). Modelling and residual analysis of nonlinear autoregressive time series in exponential variables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 47(2):165–202.
- Leemis, L. M. (2004). Building credible input models. In *Proceedings of the 2004 Winter Simulation Conference, 2004.*, volume 1, page 40.
- Nelson, B. (2013). *Foundations and methods of stochastic simulation: a first course*. Springer Science & Business Media.
- Nelson, B. L. and Gerhardt, I. (2011). Modelling and simulating non-stationary arrival processes to facilitate analysis. *Journal of Simulation*, 5(1):3–8.
- Pasupathy, R. and Nagaraj, K. (2015). Modeling dependence in simulation input: The case for copulas. In *2015 Winter Simulation Conference (WSC)*, pages 1850–1864.
- Perera, S. (2001). An analysis of daily peak stream discharge using a non-gaussian time series model. *Conference on Applied Statistics in Agriculture*.
- Rabe, M. and Scheidler, A. A. (2014). An approach for increasing the level of accuracy in supply chain simulation by using patterns on input data. In *Proceedings of the Winter Simulation Conference 2014*, pages 1897–1906.
- Sriram, K. and Whitt, W. (1986). Characterizing superposition arrival processes in packet multiplexers for voice and data. *IEEE Journal on Selected Areas in Communications*, 4(6):833–846.

Uhlig, T., Rose, O., and Rank, S. (2016). Evaluation of modeling tools for autocorrelated input processes. In *Proceedings of the 2016 Winter Simulation Conference*, pages 1048–1059. IEEE Press.

Xie, W., Nelson, B., and Barton, R. (2014). A bayesian framework for quantifying uncertainty in stochastic simulation. *Operations Research*, 62(6):1439–1452.



**ANEXOS**



## A. MODELO DE SIMULACIÓN

En esta sección detallaremos los resultados obtenidos por el modelo de simulación utilizado como base en esta investigación, respecto a los valores obtenidos tanto en tiempo de elaboración de cargas y capacidad de almacenamiento se podrá medir el impacto que genera el proceso de arribos en las salidas del modelo de simulación.

Tabla A.1: Resumen experimentos.

Experimento	Descripción
Experimento I	Determina la capacidad máxima que requiere el almacén dentro del período de simulación bajo los 3 escenarios de demanda (pesimista, esperado y optimista).
Experimento II	Se busca disminuir la rotación de personal valorando el tiempo de servicio, mediante escenarios con aumentos en el número de trabajadores con contratos tipo <i>planta</i> , en base a niveles esperados de demanda.
Experimento III	Se basa en el supuesto de aumentar la rotación de personal, es decir, tener trabajadores con contratos laborales de tipo <i>temporada</i> y <i>contratista</i> . Los escenarios evaluados buscan un aumento gradual de trabajadores con contratos tipo <i>contratista</i> , en un ambiente de producción y demanda esperada.
Experimento IV	Analiza la situación de despidos y contrataciones de personal, considerando distintas variaciones entre trabajadores de <i>planta</i> , <i>temporada</i> o <i>contratista</i> .
Experimento V	Reúne los escenarios de los experimentos anteriores cuyos resultados de tiempos de confección de cargas están por sobre la meta de 15 horas. El análisis consiste en aumentar el nivel de producción en 300 cajas por trabajador lo que se puede lograr por medio de capacitaciones, bonificaciones por producción o cambios en los procedimientos de las operaciones que permitan un aumento en la productividad.

En la Tabla A.1, se presenta una descripción de los experimentos evaluados en temáticas de dotación de personal. Luego, en la Tabla A.2 mostramos los resultados numéricos en conjunto con su representación gráfica (Ver Figura A.1). Respecto a dotación de personal no se esperan grandes impactos al variar el proceso de llegadas al modelo de simulación, dado que estos resultados están directamente relacionados con la demanda por

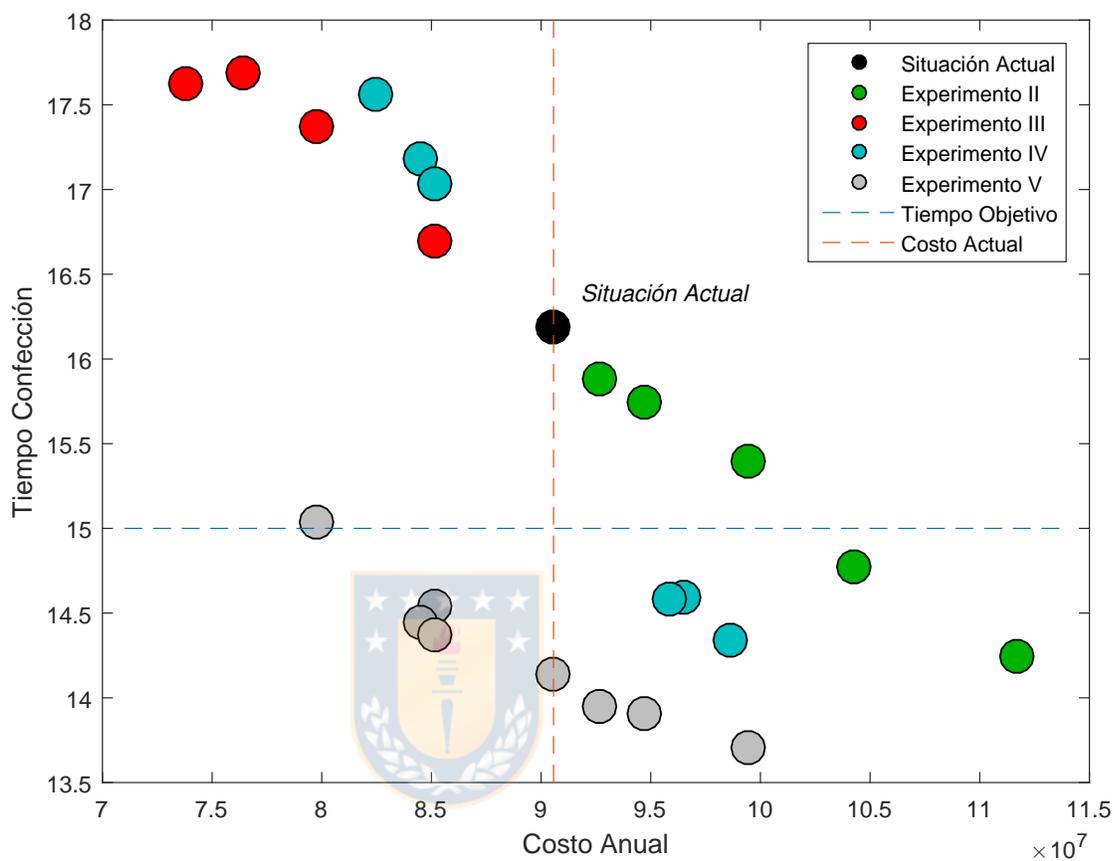
sobre el stock de entrada. Sin embargo, en caso de producirse quiebres de stock e impactar en los niveles disponibles para satisfacer la demanda diaria, efectivamente existiría un movimiento en los resultados de dotación de personal.

Tabla A.2: Tabla resumen de resultados.

Experimento	Escenario	Métricas		
		Tiempo Confección Promedio	Costo Personal	Variación Costo
Original	Original	16,184 (0,034)	90.567.000	0
	Escenario 1	15,881 (0,033)	92.635.200	2.068.200
Dos	Escenario 2	15,744 (0,031)	94.699.200	4.132.200
	Escenario 3	15,399 (0,028)	99.478.200	8.911.200
	Escenario 4	14,777 (0,028)	104.252.400	13.685.400
	Escenario 5	14,241 (0,026)	111.672.000	21.105.000
	Escenario 1	16,697 (0,036)	85.151.400	-5.415.600
Tres	Escenario 2	17,368 (0,037)	79.732.200	-10.834.800
	Escenario 3	17,692 (0,040)	76.382.400	-14.184.600
	Escenario 4	17,625 (0,038)	73.816.800	-16.750.200
	Escenario 1	14,344 (0,027)	98.593.800	8.026.800
Cuatro	Escenario 2	17,562 (0,053)	82.429.800	-8.137.20
	Escenario 3	14,594 (0,027)	96.525.600	5.958.600
	Escenario 4	17,183 (0,051)	84.495.000	-6.072.000
	Escenario 5	14,586 (0,027)	95.884.200	5.317.200
	Escenario 6	17,037 (0,051)	85.135.800	-5.431.200
	Capacitado	14,135 (0,026)	90.567.000	0
	Escenario 1	13,944 (0,025)	92.635.200	2.068.200
Cinco	Escenario 2	13,907 (0,027)	94.699.200	4.132.200
	Escenario 3	13,706 (0,053)	99.478.200	8.911.200
	Escenario 4	14,541 (0,027)	85.151.400	-5.415.600
	Escenario 5	15,033 (0,051)	79.732.200	-10.834.800
	Escenario 6	14,447 (0,027)	84.495.000	-6.072.000
	Escenario 7	14,372 (0,051)	85.135.800	-5.431.200

Tabla A.3. Tabla resumen de resultados capacidad de almacenamiento.

Escenario	Capacidad Máxima Bodega (pallets)
Original	2242,611 (58,975)
Optimista	2136,389 (56,719)
Esperado	2119,722 (52,718)
Pesimista	2073,361 (49,206)



**Figura A.1.** Comparación resultados de todos los Experimentos.

Fuente: Elaboración Propia

Respecto a capacidad de almacenamiento se esperan impactos relevantes en los resultados (salidas modelo de simulación). Dado a su relación directa con el abastecimiento. La Tabla A.3 presenta los resultados obtenidos.

## B. AUSENCIA DE DATOS

Generalmente, la identificación y estimación de modelos parten del supuesto de que las series de tiempo que se van a analizar no contienen datos faltantes. Sin embargo, en la práctica, esto no siempre ocurre, afectando la identificación del modelo adecuado y por tanto su capacidad de ajuste. En esta sección presentamos un ajuste de un modelo de efectos fijos que permite estimar los datos faltantes.

La metodología que utilizaremos es un modelo de efectos fijos que capturará los efectos del mes del año, del día de la semana y de la interacción de estos factores. De acuerdo a Channouf et al. (2007), mencionan el siguiente modelo lineal de efectos fijos como primera aproximación para capturar los diferentes comportamientos que presentan los datos (Ver Ecuación (B.1)).

$$Y_t = a + bt + \sum_{j=1}^7 \beta_j C_{t,j} + \sum_{k=1}^{12} \gamma_k S_{t,k} + \sum_{j=1}^7 \sum_{k=1}^{12} \delta_{j,k} \cdot M_{t,j,k} + E_t \quad (\text{B.1})$$

Donde  $a, b, \beta, \gamma$  y  $\delta$  son constantes.  $C_{t,j}$  es una variable binaria que tiene valor 1 si la observación  $t$  corresponde al día  $j$  de la semana. De forma similar  $S_{t,k}$  es una variable binaria que tiene valor 1 si la observación  $t$  corresponde al mes  $k$  del año. Luego asumimos que los residuos  $E_t$  son independientes e idénticamente distribuidos (i.i.d.) de forma normal con media 0 y varianza  $\sigma_{E,0}^2$  (suposición validada más adelante).

Para evitar la redundancia en los parámetros se agregan las restricciones de identificabilidad mencionadas en la Ecuación (B.2), de esta forma no existe suma y resta de las mismas magnitudes en conjunto con asegurar que se pueden estimar todos los parámetros, para más detalles ver Channouf et al. (2007) y Caridad et al. (1998).

$$\sum_{j=1}^7 \beta_j = \sum_{k=1}^{12} \gamma_k = \sum_{j=1}^7 \sum_{k=1}^{12} \delta_{j,k} = 0 \quad (\text{B.2})$$

Una vez definido el modelo de predicción estimamos todos los parámetros por medio de mínimos cuadrados simples. Esto es minimizar la suma del error cuadrático sujeto a las restricciones mencionadas en la Ecuación (B.2). Por lo tanto, el modelo generado es el

siguiente.

$$\text{Minimizar } \sum_{t=1}^n [Y_t - (a + bt + \sum_{j=1}^7 \beta_j C_{t,j} + \sum_{k=1}^{12} \gamma_k S_{t,k})]^2 \quad (\text{B.3})$$

sujeta a :

$$\sum_{j=1}^7 \beta_j = 0 \quad (\text{B.4})$$

$$\sum_{k=1}^{12} \gamma_k = 0 \quad (\text{B.5})$$

$$\sum_{j=1}^7 \sum_{k=1}^{12} \delta_{j,k} = 0 \quad (\text{B.6})$$

En las Tablas B.1, B.2, B.3 y B.4 presentamos los parámetros mencionados luego de resolver el sistema de ecuaciones respectivo.

Tabla B.1: Parámetros y estadísticas regresión lineal.

Item	Valor
Constante $a$	13.5823
Constante $b$	-0.0439
Error Cuadrático	834.8236
Desv. Estándar Error	1.6996
Media Error	$-1,315 \cdot 10^{-8}$
Varianza Estimada Error Cuadrático	4.5126
Desv. Estándar Error Cuadrático	2.1243

Tabla B.2: Parámetros  $\beta_j$ .

Día Semana	$\beta_j$
Lunes	0.3448
Martes	0.4090
Miércoles	0.4141
Jueves	0.0607
Viernes	0.0879
Sábado	-1.3564
Total	0

Tabla B.3: Parámetros  $\gamma_j$ .

Mes	$\gamma_j$
Ene	-2.4887
Feb	-2.1197
Mar	-0.0168
Abr	0.7134
May	0.5312
Jun	1.7129
Jul	1.639
Ago	5.3796
Sept	5.6507
Oct	-6.5697
Nov	-3.5876
Dic	-0.8442
Total	0

Tabla B.4: Parámetros  $\delta_{j,k}$ .

	Lun	Mar	Mie	Jue	Vie	Sab
Ene	-0.0601	0.8020	0.4455	-0.1572	-0.1404	-1.1522
Feb	-1.6643	-0.6445	0.6307	0.0280	-0.4552	-0.2170
Mar	-0.4630	0.5568	-0.5862	0.0111	0.0279	0.2343
Abr	-0.1390	0.6308	-0.1987	1.0669	-0.5391	-0.3099
May	-0.6391	-0.2011	0.2377	-0.4468	1.3200	0.0582
Jun	0.5183	0.5031	0.0419	0.5075	0.7242	-0.7845
Jul	1.4433	0.1313	-0.8300	-0.1827	0.3341	0.5406
Ago	0.6886	0.0767	0.4154	1.2128	0.6477	2.1359
Sept	-0.2389	0.4476	0.0303	1.1776	1.6126	2.4190
Oct	-1.2475	-1.1884	-0.1496	-1.5023	-1.4855	-1.1989
Nov	1.6275	-0.6028	0.7043	-1.6802	-1.9134	-1.9252
Dic	0.1883	-0.0284	-0.2532	0.1002	0.0291	-1.0827

Para continuar debemos validar el modelo, verificando si es que los errores son i.i.d. normalmente con media 0 y varianza igual a  $\sigma_{E,0}^2$ . Para lo anterior, analizaremos dos puntos;

1. Realizar una prueba para analizar la independencia de los residuos.
2. Evaluar un ajuste de una distribución normal con media 0.

Antes de validar dichos puntos es necesario calcular los errores con la Ecuación (B.7)

y estimar la varianza mediante la Ecuación (B.8).

$$E_t = Y_t - Y_t^* \quad (\text{B.7})$$

$$\hat{\sigma}_{E,0}^2 = \frac{1}{n-s} \sum_{t=1}^n \hat{E}_t^2 \quad (\text{B.8})$$

Respecto al primer punto, aplicamos la prueba de corridas arriba y abajo mediante el software Statgraphics XVI, la cual cuenta el número de veces que la secuencia ascendió o descendió. La Tabla B.5 se muestran los resultados obtenidos. Puesto que el p-valor es mayor o igual a 0,05, no se puede rechazar que la serie es aleatoria, con un nivel de confianza del 95 %. Lo anterior indica que no existe evidencia suficiente para aseverar que los errores están correlacionados.

Tabla B.5: Corridas arriba y abajo.

	Valor
Número de corridas arriba y abajo	183
Número esperado de corridas	192,333
Estadístico Z	1,2362
P-Valor	0,2164
Rechazar	No
Nivel de Confianza	95 %

Para continuar con segundo punto, las Figuras B.1 y B.2 se presentan los gráficos pdf y Q-Q plot de los errores ajustados a una distribución normal, en donde visualmente podemos suponer un ajuste de dicha distribución. Ahora bien, al calcular las estadísticas de dicha distribución obtenemos que la media del error es aproximadamente cero (Ver Tabla B.5). Con lo anterior, y en conjunto con la prueba de aleatoriedad respaldamos el ajuste propuesto.

Por lo tanto, podemos concluir que el modelo esta ajustado correctamente, en donde hemos obtenido los parámetros correspondientes mediante mínimos cuadrados y nos aseguramos que los residuos son i.i.d. de forma normal con media cero y con varianza  $\sigma_{e,0}^2$ .

Una vez ajustado el modelo procedemos a calcular los valores faltantes en la serie de tiempo, los cuales son presentados en la Tabla B.6.

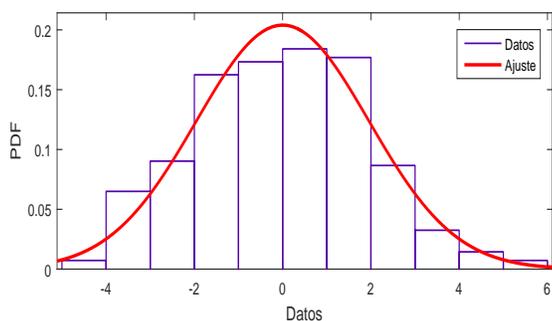


Figura B.1: Ajuste del error a una distribución normal.

Fuente: Elaboración Propia

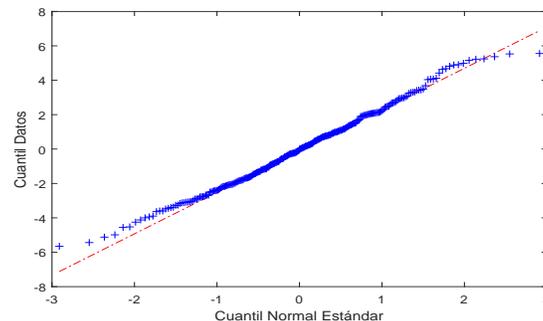


Figura B.2: Ajuste de una distribución normal al error.

Fuente: Elaboración Propia

Tabla B.6: Valores estimados para completar la serie de tiempo.

Periodo	Fecha	Estimación
6	08-10-2016	4.1937
7	10-10-2016	5.8424
25	31-10-2016	5.0517
26	01-11-2016	8.6589
56	06-12-2016	10.6588
58	08-12-2016	10.3514
62	13-12-2016	10.3953
64	15-12-2016	10.0879
79	02-01-2017	7.9483
80	03-01-2017	8.7906
86	10-01-2017	8.5271
97	23-01-2017	7.1576
159	05-04-2017	7.5271
167	14-04-2017	6.5090
168	15-04-2017	5.2500
171	19-04-2017	7.0000
181	01-05-2017	5.9089
228	24-06-2017	3.1396
229	26-06-2017	6.1396
272	15-08-2017	7.5000
301	18-09-2017	6.1576
302	19-09-2017	6.8243
312	30-09-2017	6.5911

### C. VALIDACIÓN MODELO DE SIMULACIÓN

Este capítulo tiene como objetivo detallar la validación del modelo de simulación, con objetivo de justificar que este representa de forma adecuada los procesos operativos del centro de distribución. Es necesario tener presente que existen dos salidas o variables respuesta que entrega el modelo; la capacidad de almacenamiento y el tiempo de elaboración de cargas.

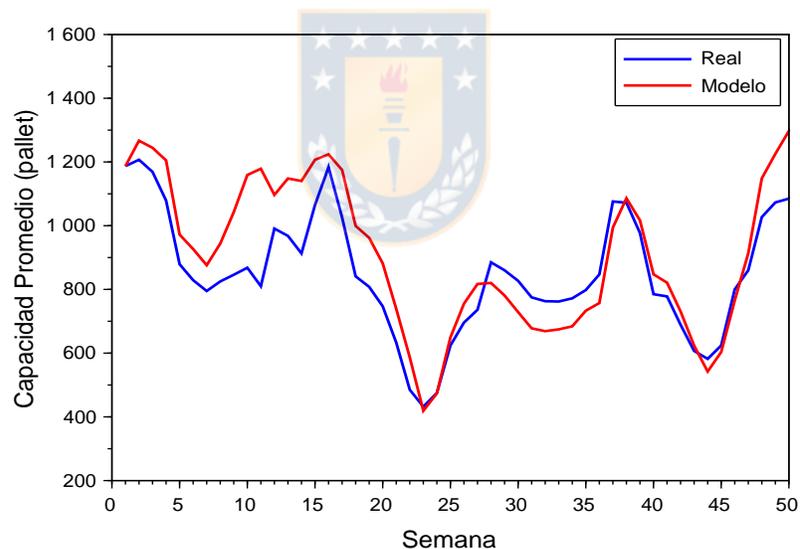


Figura C.1: Comparación entre las capacidades obtenidas en base a datos reales y en base al modelo de simulación.

Para comenzar, se corrió el modelo de simulación con un número de 36 réplicas (número justificado en De la Fuente et al. (2019)) para luego comparar los resultados obtenidos con los datos reales. Para el caso de la capacidad, se realizó una prueba de medias considerando una muestra de resultados semanales de la capacidad del almacén, obtenidos desde la simulación, y los datos de la capacidad semanal real (Ver Figura C.1). Para este análisis, se utilizó software Statgraphics Centurion considerando una hipótesis nula  $h_0 : \mu_2 - \mu_1 = 0$  y una hipótesis alternativa  $h_0 : \mu_2 - \mu_1 \neq 0$ . El software ejecuta una

prueba-t para comparar las medias de dos muestras no pareadas. También construye los intervalos de confianza para la diferencia entre las medias, los que se muestran en la Tabla C.1. Puesto que los intervalos de las diferencias de las medias contiene el valor de 0, no se puede rechazar la hipótesis nula indicando que no existe evidencia suficiente para aseverar una diferencia estadísticamente significativa entre las medias de las dos muestras de datos, con un nivel de confianza del 95,0 %.

Tabla C.1: Pruebas de medias entre la capacidad de la bodega obtenida mediante el modelo de simulación y los datos reales.

	Modelo
LSC	14,09
LIC	-151,06
Nivel de Confianza	95 %
¿Rechazar?	No

LSC - LIC: Límite superior e inferior del intervalo de confianza de la diferencia de medias.

En segundo lugar, para el tiempo de elaboración de cargas se validó en base a opinión de expertos debido a la ausencia de datos históricos, en donde mencionan que el tiempo varía entre 6,5 y 7 horas para las 10 familias de productos involucradas en este análisis. La validación es exitosa dado a que el modelo de simulación brinda un resultado de 6,229(0,010) horas para el tiempo de elaboración de cargas.

Finalmente, en base a lo expuesto el modelo representa de forma confiable el comportamiento del centro de distribución, y es apto para obtener resultados mediante su utilización.