UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INDUSTRIAS
VALPARAÍSO – CHILE

MODELOS DE ESTIMACIÓN DE LA DEMANDA PARA EMBONOR COCA-COLA S.A

SALAH KARAKI AHMAD

MEMORIA DE TITULACIÓN PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

PROFESOR GUÍA:	**SR. PABLO ESCALONA**
PROFESOR CORREFERENTE:	**SR. RAÚL STEGMAIER**

OCTUBRE-2017

I. RESUMEN

En los últimos años Embonor S.A ha aumentado el nivel de demanda insatisfecha. Una de las causas atribuibles es que la programación de la producción no se está llevando a cabo óptimamente. La programación de la producción en cada línea productiva se realiza mediante el criterio de cobertura de stock de cada producto. El ratio de cobertura es determinado a través de una estimación de la demanda. Actualmente, se estima la demanda para todos los productos (84 'SKU's') con el mismo modelo de predicción. En el presente estudio se evalúa 53 modelos de predicción de demanda para cada 'SKU'. Para cada producto se determina el modelo que minimiza la cuasivarianza muestral de los errores de pronóstico, determinándose de esta forma el mejor método de predicción para cada producto.

En este estudio se determinó que el modelo de predicción actual de la empresa es válido solo en seis de los 84 'SKU's', i.e., en este estudio se identificaron mejores métodos de pronóstico para 78 'SKU's'.

II. ABSTRACT

In recent years Embonor S.A has increased the level of unsatisfied demand. One of the attributable causes is that the scheduling of production is not being carried out optimally. The scheduling of production in each production line is done using the criterion of stock coverage of each product. The coverage ratio is determined through an estimate of demand. Currently, the demand for all products (84 'SKU's') is estimated using the same prediction model. The present study evaluates 53 demand prediction models for each 'SKU'. For each product, the model that minimizes the quasivariance of the forecast errors is determined, thus determining the best prediction method for each product.

In this study it was determined that the company's current prediction model is valid only in six of the 84 'SKU's', i.e., in this study better forecasting methods were identified for 78 'SKU's'.

III. Tabla de Contenido

1 Problemática 7
1.1 Variedad de Productos en Embonor 8
1.2 Programación de la Producción en Embonor 9
1.3 Definición de la Problemática 10
2 Objetivo 11
3 Marco Teórico 12
3.1 Modelos de Estimación 12
3.1.1 Promedio Móvil 12
3.1.2 Suavizamiento Exponencial 14
3.1.3 Regresiones Lineales 17
3.1.4 Modelos Autoregresivos 19
3.2 Indicadores de Desempeño 20
3.2.1 Cuasi-Varianza Muestral 20
3.3 Inteligencia de Negocios 21
3.3.1 Definición 21
3.3.2 Apoyo en el Análisis de Datos 21
4 Desarrollo del Problema 22
4.1 Definición Base de Datos 22
4.2 Definición de Conjuntos y Variables 24
4.3 Confección de la Base de Datos 25
4.3.1 Base de Niveles de Inventario 25
4.3.2 Base de Demanda 26
4.4 Análisis de la Data 27
4.4.1 Características Generales de los Datos 27
4.5 Modelo Actual 34
4.5.1 Deficiencias del Modelo 34
4.5.2 Efectos en la Programación de la Producción 35
4.6 Modelos Propuestos 36
4.6.1 Suavizamiento Exponencial 38
4.6.2 Regresión Lineal 38

4.6.3 Regresiones con Variables Exógenas Binarias 39
4.6.4 Regresiones con Variables Exógenas Autoregresivas 40
4.6.5 Regresiones con Variables Exógenas Suavizadas 42
4.7 Procedimiento y Metodología 44
4.7.1 Ejecución de los Modelos Propuestos 45
4.7.2 Confección de Matriz para el Ratio de Cobertura Real 46
4.7.3 Confección de Matriz para el Ratio de Cobertura por Modelo 46
4.7.4 Indicador Comparativo de Selección de Modelo Óptimo por 'SKU' 47
4.7.5 Ejecución del Algoritmo 48
4.8 Resultados Generales 49
4.8.1 Modelos Óptimos por 'SKU' 49
4.8.2 Método Seleccionado v/s Modelo Actual 54
5 Conclusiones y Comentarios 58
6 Referencias 60
7 Anexos 61
7.1 Anexo 1 62
7.2 Anexo 2: Código algoritmo generador de resultados. 63

IV. Índice de Ilustraciones

Ilustración 4-1: Demanda Acumulado 2016 v/s Demanda Acumulada 2015 Total Productos 28

Ilustración 4-2: Demanda Acumulada 2016 v/s Demanda Acumulada 2015 para Productos Retornables 29

Ilustración 4-3: Demanda por Día Para Coca-Cola Ref. Pet 2,0 L 29

Ilustración 4-4: Demanda por Mes para Coca-Cola Ref. Pet. 2,0 L. 30

Ilustración 4-5: Demanda por Trimestre para Coca-Cola Ref. Pet 2,0 L. 31

Ilustración 4-6: Diagrama del Procedimiento para la Programación de la Producción en Embonor. 35

Ilustración 4-7: Diagrama Descriptivo de la Metodología Aplicada. 44

Ilustración 4-8: Gráfica de Pareto, muestra la frecuencia de los modelos óptimos seleccionados, el porcentaje respecto al espectro total de productos evaluados, y el porcentaje acumulado asociado. 51

Ilustración 4-9: Gráfica de Pareto, muestra la frecuencia por categoría de modelos seleccionados. 52

Ilustración 4-10: Gráfica que relaciona el [%] de mejora de los modelos propuestos respecto al modelo actual 55

Ilustración 4-11: Media porcentual del indicador de mejora de modelo seleccionado respecto al modelo actual 56

Ilustración 4-12: Frecuencia de selección, respecto al porcentaje de mejora, por categoría de modelo. 57

V. Índice de Tablas

Tabla 1-2: Líneas Productivas Planta Viña del Mar *9*

Tabla 4-1: Horizonte de Evaluación para las Bases de Datos *23*

Tabla 4-2: Regresión Lineal con 3 componentes autoregresivos para el SKU 49 ..... *32*

Tabla 4-3: Regresión Lineal con 3 componentes autoregresivos para el SKU 43 ..... *32*

Tabla 4-4: Componentes de los Modelos Propuestos *36*

Tabla 4-5: Numeración de los Modelos Propuestos *37*

Tabla 4-6: Modelo Óptimo por SKU *49*

Tabla 4-7: Frecuencia de Selección de Modelos *50*

Tabla 4-8: Categoría de Modelos de Predicción y Su Frecuencia. *52*

Tabla 4-9: Indicador de Mejora Respecto al Modelo Actual *54*

Tabla 6-1: 84 SKU's A Evaluar. *62*

1 Problemática

Embonor Coca-Cola S.A. es una empresa de carácter comercial, dónde toda la cadena de valor gira en torno a la comercialización de sus productos. Por lo tanto, es imprescindible que la cadena de suministro de sus productos sea lo más eficiente. Para esto, es necesario definir una política de programación de la producción que asegure un alto nivel de servicio en términos de disponibilidad de producto. Los parámetros óptimos de toda política de programación de la producción están basados en estimaciones de la demanda de corto-mediano plazo. En la empresa no existen mecanismos ni procedimientos de control de pronósticos. Se observa un elevado error en las estimaciones de demanda que producen una mala programación de la producción y finalmente rechazo de ventas, producto de la no disponibilidad de producto en los centros de distribución de la empresa.

Según lo anterior, se identifican dos problemáticas concatenadas, a saber, (i) estimación de la demanda y (ii) programación de la producción. El alcance de esta memoria solo se refiere al punto (i).

1.1 Variedad de Productos en Embonor

Embonor S.A comercializa un total de 308 ‘SKU’s’. Estos productos tienen dos tipos de procedencias, a saber, (i) producción propia, y (ii) adquiridos. Aquellos productos referidos al punto (i) son producidos en alguna de las plantas de Embonor en Chile, mientras que los productos referidos al punto (ii) tienen distintas procedencias (importados, Embotelladora Andina, Vital Aguas, etc.). El alcance de este estudio evalúa aquellos productos producidos en la planta embotelladora de Viña del Mar, la cual produce 108 ‘SKU’s’.

De aquellos productos que la planta de Viña del Mar produce, 24 son productos bajo pedido que no requieren un pronóstico de demanda. Por lo tanto, la variedad de productos que requiere una estimación se reduce a 84 ‘SKU’s’, detallados en la tabla del anexo 1.

La cantidad de ‘SKU´s’ que requieren ser pronosticados han aumentado en los últimos años. Esto responde a una política de diversificación de los productos de la empresa para aumentar su competitividad y participación de mercado. Como la variedad de productos es creciente en el tiempo, el requerimiento de pronósticos es cada vez mayor. Dada la política actual de la empresa de utilizar un solo método para todo ‘SKU’ los errores inducidos en los pronósticos pueden ser mayores, y tener repercusión en los niveles de servicio al cliente.

La planta de Viña del Mar posee tres centros de distribución. (i) Sucursal Con Cón, (ii) sucursal San Felipe, y (iii) sucursal Valparaíso. Cada uno de ellos es abastecido por un gran almacén denominado bodega de productos terminados. La bodega de productos terminados, es el almacén que tiene como función recepcionar tanto los productos de origen propio (los producidos en la planta de Viña del Mar) como los de origen adquirido.

1.2 Programación de la Producción en Embonor

La planta de Embonor en Viña del Mar cuenta con cinco líneas productivas para la producción de bebidas retornables, desechables y vidrio. La categoría y cantidad de productos por línea se describen en la tabla 1-1,

Tabla 1-1: Líneas Productivas Planta Viña del Mar

Línea Productiva	# SKU's	Categoría
L1	20	Retornable Plástico
L2	17	Retornable Vidrio
L3	16	Desechable
L4	2	Desechable
L5	29	Desechable
Total	84	

La programación de la producción en cada linea productiva se realiza mediante prioridad. Esta prioridad es determinada en base al *ratio de cobertura* que consiste en un indicador de "¿Cuánto demanda nos alcanza a cubrir el inventario actual antes de que se agote el stock (en días)?", lo que deja en evidencia la importancia de contar con un buen pronóstico de demanda. Es decir, un buen pronóstico de demanda asegura una buena programación de la producción.

El ratio de cobertura estimado denominado como "días pisos" se utiliza como criterio principal debido a que entrega una referencia de que producto comenzará primero a generar demanda insatisfecha. La idea general es que en base a este ratio, y en orden de prioridad comienzan a entrar los productos a las líneas de producción. Si el ratio de cobertura de demanda no está bien estimado, no es posible realizar una programación certera, por lo que es probable que se generen quiebres de stock en los productos con nivel de inventario crítico, generando así demanda insatisfecha en el mercado. La demanda insatisfecha en Embonor se utiliza como un 'KPI' clave de gestión denominado "rechazo". Su minimización es considerada como un factor de éxito para la compañía.

1.3 Definición de la Problemática

El programador de la producción en la planta de Embonor Viña del Mar tiene por objetivo minimizar los rechazos. Como no tiene injerencia sobre las fallas de cada línea productiva, ni tampoco sobre la variabilidad de la demanda del mercado, entonces se hace imprescindible una programación robusta de la producción basada en una buena estimación del ratio de cobertura. Lamentablemente, la data analizada evidencia que el ratio de cobertura no está siendo estimado correctamente.

La problemática central se reduce entonces a que actualmente se utiliza como política para programar la producción el ratio de cobertura estimado de demanda. Como el cálculo del ratio no es eficaz, entonces, la programación de la producción por cada línea productiva obtiene una secuencia errada, resultado de esta secuencia se genera un aumento de la demanda insatisfecha.

2 Objetivo

El objetivo principal de esta memoria es determinar el método de pronóstico más adecuado para cada ‘SKU’ que se fabrica en las líneas de producción.

Los objetivos secundarios de esta memoria son los siguientes:

- Pre-procesar la base de datos histórica de tal forma que se pueda evaluar diferentes modelos de predicción.

- Utilizar herramientas de inteligencia de negocios sobre la base de datos confeccionada de tal forma que permita una interpretación eficaz de la data.

- Minimizar el error de estimación entre el ratio de cobertura real respecto al ratio de cobertura estimado por cada producto evaluado.

- Detectar oportunidades de mejora para los procedimientos y criterios que determinan la programación de la producción.

3 Marco Teórico

En este capítulo se describen tanto cualitativa como cuantitativamente los métodos de series de tiempo para realizar pronósticos de demanda. Además, se describen las metodologías utilizadas para determinar el método de pronóstico más adecuado. Finalmente, se hace una breve descripción acerca de las herramientas de inteligencia de negocios existentes en la actualidad.

El alcance del marco teórico se limita tanto a series de tiempos como a métodos de pronósticos debido a que se dispone de una base de datos histórica de demanda como herramienta principal para realizar este estudio. Note que la base de datos de Embonor contiene las ventas y los rechazos, es decir, la demanda.

3.1 Modelos de Estimación

El principal objetivo para realizar un pronóstico es elaborar un esquema que sirva como base para poder tomar una buena decisión basada en informaciones obtenidas previamente. De estas informaciones se decide cuál método de pronóstico utilizar, a continuación se describen en detalle los modelos de predicción utilizados actualmente según (Axsater, 2015)

3.1.1 Promedio Móvil

Suponiendo que la demanda de un periodo determinado varía lentamente, entonces el modelo más simple posible a utilizar es el modelo constante. Este modelo supone que las demandas en diferentes periodos están representadas por desviaciones estables que tienden a

un promedio y que son completamente independientes de las desviaciones aleatorias en el tiempo.

Introduciendo la notación del modelo:

$X_t : Demanda\ en\ el\ período\ t$

$A : Demanda\ media\ por\ período$

$\varepsilon_t : Desviación\ aleatoria\ independiente\ con\ media\ cero$

Un modelo constante asume que la demanda en el periodo t puede ser representada como,

$$X_t = A + \varepsilon_t \qquad (3\text{-}1)$$

Sin embargo, bajo el supuesto que la demanda en el tiempo varía lentamente, razonablemente es mejor utilizar un modelo que considere solo el período de las últimas demandas. A este modelo se le denomina modelo de promedio móvil y la idea es tomar el promedio sobre los N valores más recientes, como sigue,

$\hat{a}_t : Estimación\ de\ A\ después\ de\ observar\ la\ demanda\ en\ el\ período\ t$

$\hat{x}_{t,\tau} : Estimación\ del\ período\ \tau > t\ al\ observar\ la\ demanda\ en\ el\ período\ t$

Se obtiene,

$$\hat{x}_{t,\tau} = \hat{a}_t = \frac{(x_t + x_{t-1} + x_{t-2} + \cdots + x_{t-N+1})}{N} \qquad (3\text{-}2)$$

El valor de N depende de que tan lento varía la demanda en un período, razonablemente para una demanda que varía muy lentamente puede ser mejor un N grande, sin embargo, si esta demanda varía no tan lentamente, entonces se vuelve más ideal un N pequeño.

3.1.2 Suavizamiento Exponencial

El modelo de suavizamiento exponencial es en muchos aspectos similar al modelo de promedio móvil, sin embargo, la diferencia radica en que se utiliza una combinación lineal entre el pronóstico del período anterior y la demanda más reciente. El modelo entonces es,

$$\hat{x}_{t,\tau} = \hat{a}_t = (1-\alpha)\hat{a}_{t-1} + \alpha x_t \qquad (3\text{-}3)$$

Donde, $\alpha : Constante\ de\ suavizamiento$ ($0 < \alpha < 1$)

Si elegimos un $\alpha = 0$ significa que no se está modificando el pronóstico, si elegimos un $\alpha = 1$ significa que se está eligiendo la demanda más reciente como nuestro pronóstico.

3.1.2.1 *Suavizamiento Exponencial Simple*

Este método es adecuado para series que se mueven aleatoriamente por sobre y por debajo de una demanda media constante, y hace referencia al modelo descrito en la ecuación (3-3).

3.1.2.2 *Suavizamiento Exponencial Doble*

Este método es una extensión al modelo descrito en la ecuación (3-3) que es utilizado para series que presenten características de tendencia lineal, el cual se representa como,

$$S_t = \alpha x_t + (1-\alpha)S_{t-1} \qquad (3\text{-}4)$$

$$D_t = \alpha S_t + (1-\alpha)D_{t-1} \qquad (3\text{-}5)$$

Donde S_t es un suavizamiento simple y D_t un suavizamiento doble.

La estimación para el suavizamiento exponencial doble es entonces como sigue,

$$\hat{x}_{t+\tau} = \left(2 + \frac{\alpha\tau}{1-\alpha}\right) S_t - \left(1 - \frac{\alpha\tau}{1-\alpha}\right) D_t \qquad (3\text{-}6)$$

3.1.2.3 *Modelo Holt-Winters Multiplicativo*

Este método es apropiado para series que presenten tendencia lineal y una variación estacional multiplicativa. (i.e desviación estándar creciente en el tiempo). La estimación de este modelo se encuentra definida por,

$$\hat{x}_{t+\tau} = (a + b\tau)C_{t+\tau} \qquad (3\text{-}7)$$

Donde los coeficientes son,

$a = Componente\ constante\ (intercepto)$

$b = Tendencia$

$C_t = Factor\ de\ estacionalidad\ multiplicativa$

Los cuales están definidos de la siguiente forma recursiva,

$$a_t = \alpha \frac{x_t}{C_{t,t-s}} + (1-\alpha)(a_{t-1} + b_{t-1}) \tag{3-8}$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \tag{3-9}$$

$$C_{t,t} = \gamma \frac{x_t}{a_t} + (1-\gamma)C_{t,t-s} \tag{3-10}$$

Los parámetros están dentro del intervalo $0 < \alpha,\ \beta,\ \gamma < 1$ y el factor s se define como la frecuencia del período estacional.

3.1.2.4 *Modelo Holt-Winters Aditivo*

Este método es apropiado para series que presenten tendencia lineal y una variación estacional aditiva. (i.e desviación estándar constante en el tiempo). La estimación de este modelo se encuentra definida por,

$$\hat{x}_{t+\tau} = a + b\tau + C_{t+\tau} \tag{3-11}$$

Donde los coeficientes se definieron para la ecuación (3-7), solo que ahora el factor de estacionalidad es aditivo. Los coeficientes para este modelo están definidos por las siguientes recursiones,

$$a_t = \alpha\left(x_t - C_{t,t-s}\right) + (1-\alpha)(a_{t-1} + b_{t-1}) \tag{3-12}$$

$$b_t = \beta(a_t - a_{t-1}) + 1 - \beta b_{t-1} \tag{3-13}$$

$$C_{t,t} = \gamma(x_t - a_{t+1}) - \gamma C_{t,t-s} \tag{3-14}$$

Los parámetros están dentro del intervalo $0 < \alpha, \beta, \gamma < 1$ y el factor s se define como la frecuencia del período estacional.

3.1.3 Regresiones Lineales

Cuando se tiene una serie que presenta una tendencia se puede utilizar el modelo descrito anteriormente en la ecuación (3-3). Sin embargo, (Jenkins, 2013) plantea un método denominado regresión de mínimos cuadrados, este tiene como idea minimizar los errores asociados a las demandas que se encuentren por sobre o por debajo de una recta de tendencia asociada a un período. Supónganse entonces la siguiente recta de tendencia para un período t basada en la estimación de las N demandas más recientes (x_t , x_{t-1} , ... , x_{t-N+1})

$$y_{t+k} = \hat{\alpha}_t + \hat{\beta}_t k \qquad (3\text{-}15)$$

Donde $k > 0$ representa la estimación del periodo $t + k$ por lo que se define de la siguiente manera (i.e $\hat{x}_{t,t+k} = y_{t+k}$).

Para determinar los parámetros óptimos de $\hat{\alpha}_t$ y $\hat{\beta}_t$ se debe minimizar la siguiente función convexa,

$$\sum_{k=-N+1}^{0} \left(x_{t+k} - \hat{\alpha}_t - \hat{\beta}_t k\right)^2$$

Derivando ambos parámetros e igualando a cero se obtienen las siguientes ecuaciones,

$$-2\sum_{k=-N+1}^{0} x_{t+k} - \hat{\alpha}_t - \hat{\beta}_t k = 0$$

$$-2\sum_{k=-N+1}^{0} k\left(x_{t+k} - \hat{\alpha}_t - \hat{\beta}_t k\right) = 0$$

Usando la notación,

$$\bar{k} = \frac{1}{N}\sum_{k=-N+1}^{0} k = \frac{-N+1}{2}$$

$$\bar{x} = \frac{1}{N}\sum_{k=-N+1}^{0} x_{t+k}$$

Por lo tanto, se obtiene los parámetros que minimizan los errores en la recta de tendencia de la ecuación (3-15),

$$\hat{\beta}_t = \frac{\sum_{k=-N+1}^{0}(k - \bar{k})(x_{t+k} - \bar{x})}{\sum_{k=-N+1}^{0}(k - \bar{k})^2} \qquad (3\text{-}16)$$

$$\hat{\alpha}_t = \bar{x} - \hat{\beta}_t \bar{k} \qquad (3\text{-}17)$$

El modelo anteriormente busca minimizar los errores sobre la recta de tendencia (debido a que se está analizando una predicción en base exclusiva al comportamiento de tendencia de una serie). Sin embargo, este modelo puede extenderse a la utilización de múltiples variables denominadas exógenas, las cuales pueden ser o no significativas dentro de un modelo de regresión lineal. La matemática asociada a la resolución para la minimización de los cuadrados en un modelo de regresión múltiple es la misma del modelo de regresión lineal simple, pero extendida a sus variables.

3.1.4 Modelos Autoregresivos

Las variaciones estocásticas en los modelos de demanda que se han considerado en las ecuaciones (3-3) y (3-15) se asumen que son independientes. Sin embargo, es evidente que en situaciones esto no es cierto, ya que por ejemplo, a veces se puede esperar que la demanda en períodos consecutivos este negativa o positivamente correlacionada. Una de las técnicas de pronóstico que pueda manejar variaciones estocásticas correlacionadas es la desarrollada por (Jenkins, 2013). La técnica desarrollada es conocida como modelo 'AutoRegressive Integrated Moving Average' (ARIMA). Existe una gran variedad de estos modelos, y es común utilizar la notación ARIMA (p, d, q) donde,

$$AR: p = Orden\ de\ la\ parte\ autoregresiva\ del\ modelo.$$

$$I:\ d = Grado\ de\ primera\ diferencia$$

$$MA: p = Orden\ del\ promedio\ móvil$$

Si se considera el caso donde el orden integrado (o de primera diferencia) es cero, denotando ARIMA (p, 0, q) se tiene el siguiente modelo general,

$$x_t = a + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_2 x_{t-p} \ldots + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_1 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (3\text{-}18)$$

En la ecuación (3-18), ε_t son errores que son considerados como variables exógenas para el modelo que tienen por objetivo predecir el comportamiento estocástico de la demanda a través de la correlación entre los errores de periodos consecutivos.

3.2 Indicadores de Desempeño

Los indicadores de desempeño para los modelos de predicción de la demanda permiten comparar cuantitativamente el error de estimación. Y por lo tanto, demuestran que modelo es mejor sobre otro. En el libro "Inventory Control" (Axsater, 2015) se definen algunos indicadores utilizados, sin embargo, para este estudio se utiliza la cuasivarianza muestral.

3.2.1 Cuasi-Varianza Muestral

La cuasivarianza muestral se define como,

$$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \qquad (3\text{-}19)$$

Y caracteriza la dispersión o volatilidad de una muestra o población, la cuasivarianza muestral es el mejor estimador centrado (no sesgado) de la varianza poblacional. Si se evalúan distintos modelos de estimación con la misma muestra, el modelo óptimo en relación a que obtiene una menor dispersión es el que minimiza el indicador de cuasivarianza., i.e., sea k el conjunto de modelos de pronósticos y s_k^2 la cuasivarianza muestral del método. Luego, $k^* = Argmin_{R \in k} \{s_R^2\}$.

3.3 Inteligencia de Negocios

La Inteligencia de Negocios BI (‘Business Intelligence’) es una herramienta bajo la cual diferentes tipos de organizaciones, pueden soportar la toma de decisiones basadas en información precisa y oportuna; garantizando la generación del conocimiento necesario que permita escoger la alternativa que sea más conveniente para el éxito de la empresa.

3.3.1 Definición

La inteligencia de negocios se define como la habilidad corporativa para tomar decisiones. Esto se logra mediante el uso de metodologías, aplicaciones y tecnologías que permiten reunir, depurar, transformar datos, y aplicar en ellos técnicas analíticas de extracción de conocimiento (Parr, 2000) los datos pueden ser estructurados para que indiquen las características de un área de interés generando el conocimiento sobre los problemas y oportunidades del negocio para que puedan ser corregidos y aprovechados respectivamente. (Sobrosa, 2013)

3.3.2 Apoyo en el Análisis de Datos

Uno de los aspectos más interesantes de las herramientas actuales para desarrollar inteligencia de negocios es la fácil manipulación e interpretación de gran cantidad de datos. Al utilizar una herramienta ‘BI’ es posible dinamizar totalmente la data que se desea trabajar, y por lo tanto, la visualización de cualquier factor que se desee analizar es más eficaz que los métodos tradicionales.

Uno de los Software más reconocidos en el mercado es el desarrollado por Microsoft llamado ‘POWER BI’ el cual posee una interfaz amigable para el análisis estadístico.

4 Desarrollo del Problema

El desarrollo de este estudio consta de siete puntos de análisis (i) definición, componentes, y horizontes de evaluación para la base de datos (ii) análisis de los datos con herramientas de inteligencia de negocios (iii) definición del modelo actual en Embonor (iv) propuestas de modelos para la estimación de la demanda (v) metodología y procedimiento del estudio (vi) generación de un algoritmo para seleccionar los métodos de pronósticos más adecuados y (vii) resultados generales.

4.1 Definición Base de Datos

La realización de este estudio considera la obtención de una base de datos estructurada que permita el análisis de los productos. Para ello la compañía facilita su base histórica de ventas junto con la base de las demandas que no pudieron ser satisfechas (rechazos). Adicionalmente, se elabora una base de datos de los niveles de inventarios de cada producto por el horizonte de tiempo a evaluar.

Finalmente, se obtiene (i) base de datos para las ventas (ii) base de datos para los rechazos y (iii) base de datos para los niveles de inventario. Las tres bases de datos descritas para cada uno de los 84 productos a evaluar.

El horizonte de evaluación para las bases de datos definidas en los puntos (i)-(ii) y (iii) es el descrito en la siguiente tabla,

Tabla 4-1: Horizonte de Evaluación para las Bases de Datos

Base de Datos	Fecha Inicio	Fecha Fin	Periocidad
Ventas	01-01-2014	20-05-2017	Diaria
Rechazos	01-01-2014	20-05-2017	Diaria
Inventario	01-01-2017	20-05-2017	Diaria

Una consideración particular es que el horizonte de evaluación no incluye días en que no hay ventas. Para el caso de Embonor, domingos y festivos.

Otra consideración del horizonte de evaluación para la base de datos de inventario, es que es un subconjunto del horizonte de evaluación para la base de datos de demanda. Esto debido a que se utiliza un período conocido de la demanda.

4.2 Definición de Conjuntos y Variables

- **Conjuntos**

$I =$ Conjunto de Productos Evaluados

$\mathrm{T} =$ Conjunto de Período para la Demanda (Horizonte de Evaluación)

$\mathrm{T}' =$ Conjunto de Período para el Stock (Horizonte de Evaluación)

$\mathrm{M} = Conjunto\ de\ Meses$

$N = Conjunto\ de\ Días\ de\ la\ Semana$

- **Variables**

$\boldsymbol{V_{it}} = Ventas\ del\ producto\ i\ en\ el\ período\ t.$

$\boldsymbol{V'_{it}} = Rechazos\ del\ producto\ i\ en\ el\ período\ t.$

$\boldsymbol{D_{it}} = Demanda\ para\ el\ producto\ i\ en\ el\ período\ t.$

$\boldsymbol{S_{it}} = Nivel\ de\ inventario\ para\ el\ producto\ i\ en\ el\ período\ t$

$$\boldsymbol{M_t} = \begin{cases} 1, si\ mes\ está\ \ en\ curso\ en\ el\ período\ t \\ 0, si\ mes\ no\ está\ en\ curso\ en\ el\ período\ t \end{cases}$$

$$\boldsymbol{N_t} = \begin{cases} 1, si\ día\ de\ la\ semana\ está\ \ en\ curso\ en\ el\ período\ t \\ 0, si\ día\ de\ la\ semana\ no\ está\ en\ curso\ en\ el\ período\ t \end{cases}$$

$\boldsymbol{Corr_t} = Serie\ correlativa\ para\ el\ perído\ t$

$\boldsymbol{SE_{it}} = Suavización\ Exponencial\ Mult. para\ el\ periodo\ t\ del\ producto\ i$

4.3 Confección de la Base de Datos

Para el desarrollo de este estudio se estructuran dos bases de datos, a saber (i) base de datos para la demanda y (ii) base de datos para los niveles de inventarios. La primera sobre el horizonte de evaluación T y la segunda sobre el horizonte de evaluación T'. Como se menciona anteriormente, T' es un subconjunto de T. La relación entre ambos conjuntos (T' ⊆ T) es la siguiente,

- T = (1,2,…, 1024, 1025)
- T'= (912, 913,…, 1024,1025)

Donde, el subconjunto T' solo considera el horizonte de tiempo que va desde la fecha número 912 hasta la fecha 1025.

4.3.1 Base de Niveles de Inventario

Primero, es de importancia mencionar que se trabaja con los niveles de inventarios del total de la planta, esto es, la suma de los inventarios de la bodega de productos terminados más los niveles de inventario de cada sucursal. Se define entonces el stock total planta de la siguiente manera,

$$\mathbf{S}_{\text{Total Planta}} = \mathbf{S}_{\text{BPT}} + \mathbf{S}_{\text{Concón}} + \mathbf{S}_{\text{Valparaíso}} + \mathbf{S}_{\text{San Felipe,}} \qquad (4\text{-}1)$$

Donde $S_{\text{Total Planta}}$ corresponde al stock del total de planta, S_{BPT} corresponde al stock en la bodega de productos terminados, $S_{\text{Concón}}$ corresponde al stock en la sucursal de Concón,

$S_{Valparaíso}$ corresponde al stock en la sucursal de Valparaíso y, $S_{San\ Felipe}$ corresponde al stock en la sucursal de San Felipe.

$$\mathbf{S}_{Total\ Planta} = S_{it}, \qquad \forall \ \ t \ \in T'; \ \ i \in I$$

Se confecciona una matriz de tal forma que en sus columnas se contiene el elemento de cada producto, y en sus filas se contiene el elemento de cada fecha en que se efectúa demanda. La matriz generada tiene dimensión [114 x 84], es decir, 114 fechas en que se efectúa demanda para 84 productos.

4.3.2 Base de Demanda

A saber se tiene una base de datos para las ventas y otra para los rechazos. Se define como demanda la suma entre las ventas efectuadas y la demanda insatisfecha de la siguiente manera:

$$D_{it} = V_{it} + V'_{it}, \qquad \forall \ \ t \ \in T; \ \ i \in I \qquad (4\text{-}2)$$

Por lo tanto, se obtiene una matriz definida como D_{it} de dimensión [1025 x 84], es decir, 1025 fechas en que se efectúa demanda para 84 productos.

Es importante mencionar dos aspectos de la base de datos (i) que está estructurada en base a las unidades de ventas de cada producto, estás unidades de ventas son representadas a través del concepto de “cajas físicas” las cuales varían para cada producto. Por ejemplo, un retornable de 2000 cc posee una caja física que contiene 4 bebidas, sin embargo, un desechable de 3000 cc posee una caja física de 6 bebidas. A causa de lo anterior, la matriz D_{it} representa el volumen de demanda en unidades de venta de cada producto y no en unidades de bebidas, (ii) que la base de rechazos representa la demanda insatisfecha como la diferencia entre todos los pedidos de venta del período t cotejados con los niveles de inventario del mismo periodo t. Cómo los clientes realizan sus pedidos repetitivamente hasta

que son abastecidos, entonces se genera una base con demandas insatisfechas repetitivas. Por lo tanto, la base de rechazos se encuentra sobreestimada.

4.4 Análisis de la Data

Antes de comenzar a proponer modelos de predicción es necesario realizar un análisis previo de la data. Para ello, se hace uso de herramientas de inteligencia de negocios generando cuatro reportes descritos a continuación,

- Reporte 1: Tendencia de la demanda con filtros dinámicos para (i) el horizonte de evaluación, (ii) línea de producción asociada a cada 'SKU' y (iii) descripción de cada producto.

- Reporte 2: Proporción de demanda diaria (día de la semana) con filtros dinámicos para el (i) horizonte de evaluación, (ii) línea de producción asociada a cada 'SKU' y (iii) descripción de cada producto.

- Reporte 3: Proporción de demanda mensual y trimestral con filtros dinámicos para (i) el horizonte de evaluación, (ii) línea de producción asociada a cada 'SKU' y (iii) descripción de cada producto.

- Reporte 4: Demanda real respecto a la estimación con el modelo actual (promedio móvil de orden 6) y también promedios móviles de cualquier orden.

4.4.1 Características Generales de los Datos

A través del análisis de los reportes generados se pueden observar tres características esenciales presentes en la data de demanda de los productos. (i) Factor de tendencia (ii) Factor de Estacionalidad y (iii) Factor de autocorrelación. Por el momento no es relevante cuantificar el factor asociado a cada fenómeno, bastando solo con su detección.

Para el punto (i) se observa tendencia positiva en la demanda agregada de los productos. Es decir, la demanda total aumenta con el paso del tiempo. Se ejemplifica a través de la siguiente gráfica que muestra en las barras azules la demanda acumulada del año 2016 respecto a la línea amarilla que representa la demanda acumulada del año 2015. El aumento fue de un 8,17 [%] respecto al año anterior.

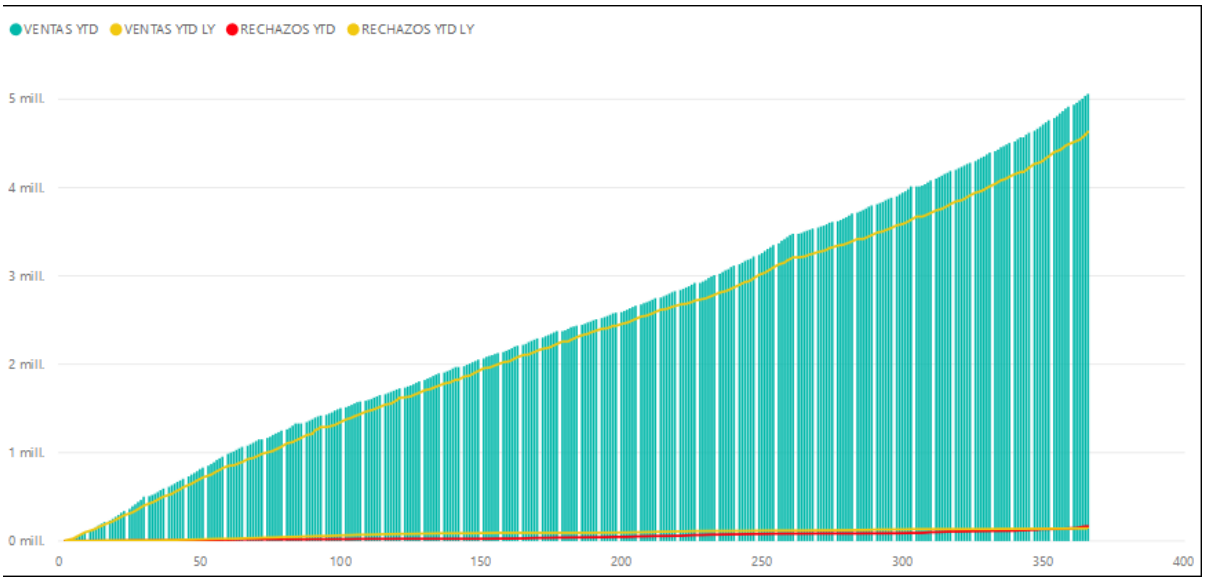

Ilustración 4-1: Demanda Acumulado 2016 v/s Demanda Acumulada 2015 Total Productos

Sin embargo, es posible visualizar demandas decrecientes para los productos retornables. En términos generales, los productos desechables aumentan su demanda mientras que los productos retornables disminuyen su demanda constantemente. La siguiente gráfica muestra el comportamiento de la demanda para los productos retornables, se visualiza que la línea amarilla está por sobre las barras azules. Es decir, que para el año 2016 se demandó menos que durante el año 2015. La disminución de la demanda fue de un 2,87 [%] respecto al año anterior.

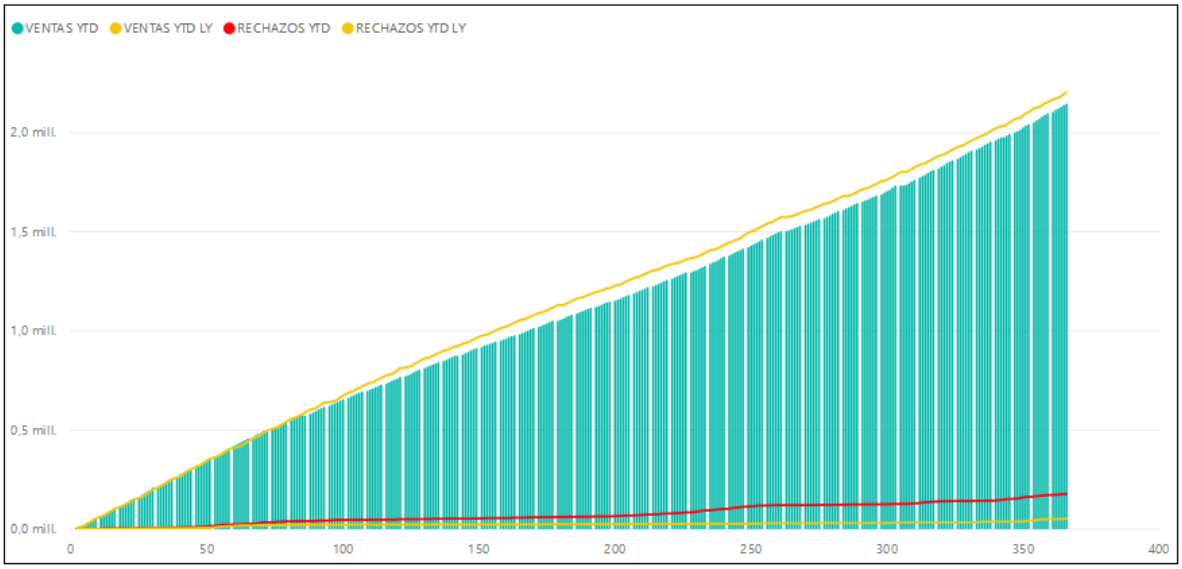

Ilustración 4-2: Demanda Acumulada 2016 v/s Demanda Acumulada 2015 para Productos Retornables

Para el punto (ii) se observa estacionalidad de tres tipos, la primera corresponde a una estacionalidad temporal asociada a las estaciones del año. La segunda, corresponde a una estacionalidad mensual, y la tercera corresponde a una estacionalidad diaria. Por ejemplo, para el producto más vendido de la compañía, la Coca-Cola 2,0 [l] retornable de plástico se tiene el siguiente comportamiento de estacionalidad,

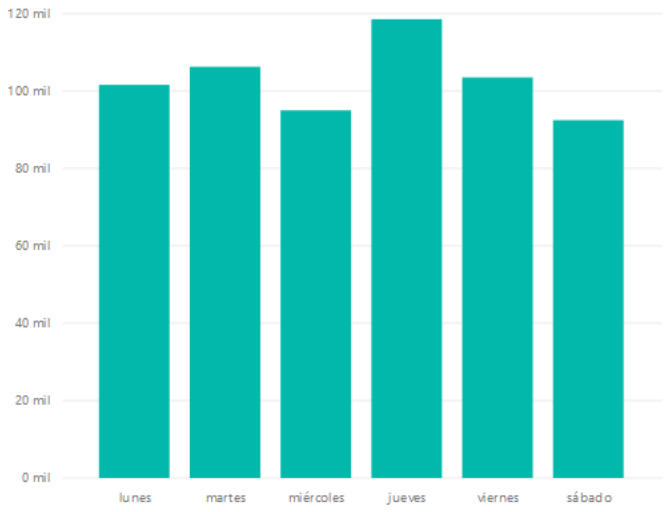

Ilustración 4-3: Demanda por Día Para Coca-Cola Ref. Pet 2,0 L

La Ilustración 4-3 muestra el comportamiento estacional diario de las demandas entre el 01/01/2017 al 20/05/2017. Por ejemplo, se observa que los días jueves se vende en promedio un 25 [%] más del producto respecto al día sábado. La proporción se hace más intuitiva conociendo que las ventas del día jueves son entregadas el día viernes a los clientes, y las ventas del día sábado son entregadas los días lunes. Es evidente que el factor estacional diario captura el comportamiento de este producto en particular.

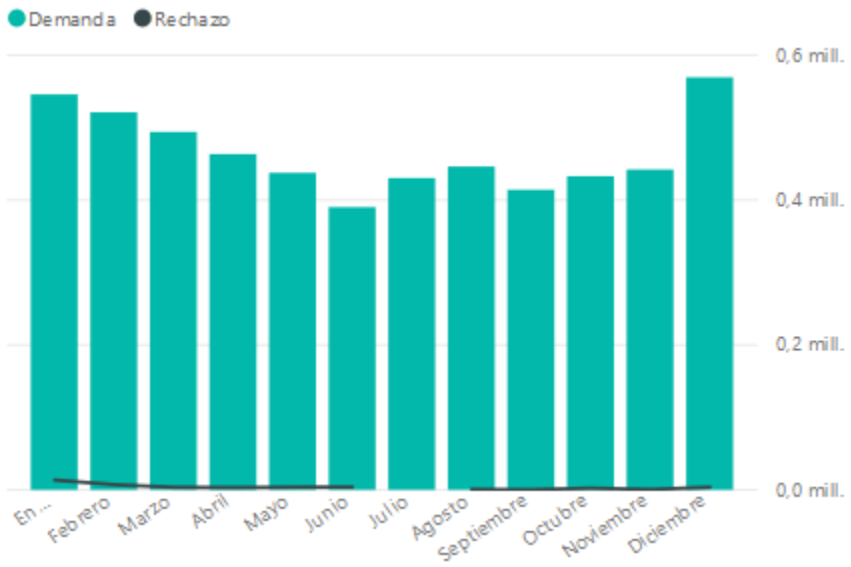

Ilustración 4-4: Demanda por Mes para Coca-Cola Ref. Pet. 2,0 L

La Ilustración 4-4 muestra el comportamiento estacional mensual de las demandas entre el 01/01/2014 al 31/12/2016 (Una muestra de tres años). Por ejemplo, se observa que la demanda en diciembre es aproximadamente superior en un 1/3 de la demanda de junio. Es evidente también el comportamiento estacional mensual de este producto en particular.

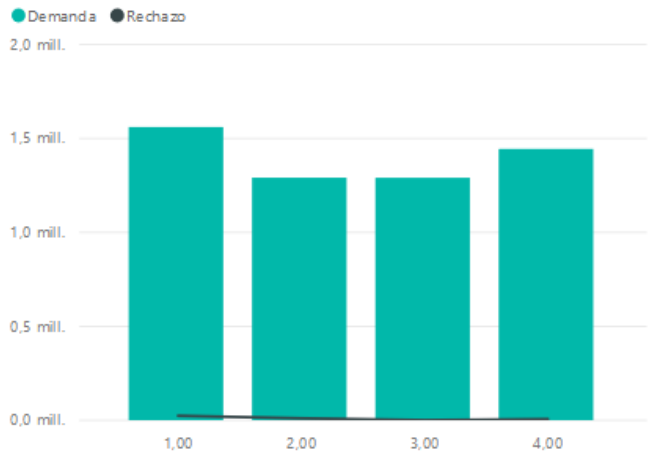

Ilustración 4-5: Demanda por Trimestre para Coca-Cola Ref. Pet 2,0 L.

1 representa el trimestre Ene-Feb-Mar; 2 representa el trimestre Abr-May-Jun; 3 representa el trimestre Jul-Ago-Sep; 4 representa el trimestre Oct-Nov-Dic.

La Ilustración 4-5 muestra el comportamiento estacional trimestral de las demandas entre el 01/01/2014 al 31/12/2016 (una muestra de tres años). Por ejemplo, se observa que las demandas para el primer y último trimestre son más altas que las demandas para el segundo y tercer trimestre. Debido principalmente a que el clima es un factor predominante en el consumo de este tipo de productos.

Para el punto (iii) se observa un factor de correlación entre las demandas y las demandas rezagadas. Es decir, la demanda del periodo anterior, se encuentra correlacionada con la demanda del periodo siguiente. Se observa que al menos debería ser significativa la correlación hasta tres períodos de rezago. Para ejemplificar lo anterior se muestra en las siguientes tabla información de un modelo autoregresivo de orden tres para dos ‘SKU’s’.

Tabla 4-2: Regresión Lineal con 3 componentes autoregresivos para el SKU 49

Dependent Variable: SKU49
Method: Least Squares
Date: 09/05/17 Time: 01:16
Sample (adjusted): 4 1025
Included observations: 1022 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
SKU49(-1)	0.404869	0.030816	13.13842	0.0000
SKU49(-2)	0.316943	0.031781	9.972669	0.0000
SKU49(-3)	0.180492	0.030784	5.863106	0.0000

Tabla 4-3: Regresión Lineal con 3 componentes autoregresivos para el SKU 43

Dependent Variable: SKU43
Method: Least Squares
Date: 09/05/17 Time: 01:16
Sample (adjusted): 4 1025
Included observations: 1022 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
SKU43(-1)	0.579010	0.030743	18.83401	0.0000
SKU43(-2)	0.184374	0.035273	5.227039	0.0000
SKU43(-3)	0.194307	0.030825	6.303585	0.0000

Las tablas descritas contienen la información de dos modelos autoregresivos, uno para el producto 43 y otro para el producto 49. Las variables regresoras para cada modelo corresponden a los rezagos de los periodos 1, 2 y 3. La segunda columna de la tabla contiene el coeficiente que minimiza la recta de regresión a través del método de mínimos cuadrados ordinarios y representa la ponderación asociada al número de rezago. Ambos modelos son significativos (la probabilidad de cometer un error tipo I [1]es de un 0,00 [%]). Es evidente, por ejemplo, que existe un alto factor de correlación en el SKU43 respecto al rezago de orden 1 ya que su coeficiente es de un 57,9 [%]. Es decir, la demanda en el periodo inmediatamente

[1] Un error Tipo I, se refiere a la probabilidad de rechazar una hipótesis nula cuando esta en realidad es verdadera.

anterior, representa un 57,9 [%] de la demanda si se evalúa un horizonte para tres rezagos. Es evidente entonces, la utilización de un factor de correlación en los modelos de predicción.

4.5 Modelo Actual

El modelo actual para el cálculo del ratio de cobertura de demanda ($\hat{R}_{it}$) en la compañía es el definido a continuación:

$$\hat{R}_{it} = \frac{S_{it}}{\frac{\sum_1^6 d_{t-n,i}}{6}} \qquad (4\text{-}3)$$

De la ecuación (4-3) se observa que el modelo actual en Embonor es simplemente un promedio móvil de orden $n = 6$

4.5.1 Deficiencias del Modelo

Es irracional para una empresa del tamaño de Embonor seguir estimando la demanda mediante el modelo actual. Esto debido a que un promedio móvil de orden $n = 6$ no es capaz de capturar efectos estacionales ni de tendencia. La idea general para mejorar estas deficiencias es considerar un modelo para cada producto que sea capaz de comprender la dinámica propia del comportamiento de su demanda.

Al estimar con un promedio móvil de orden $n = 6$ estamos dejando de considerar el comportamiento histórico de la demanda. Y por lo tanto, excluyendo efectos inherentes al comportamiento característico de cada producto.

4.5.2 Efectos en la Programación de la Producción

La programación de la producción en Embonor está enfocada en la minimización de los rechazos. Para ello, es adecuado programar mediante un orden de prioridad la fabricación de productos debido a que existen diversos factores que afectan la utilización de las líneas productivas. Cómo existen estos factores ajenos a lo programación de la producción que afectan la utilización de las líneas, es necesario replantear el escenario de producción todos los días. La forma en que se aborda esta problemática es la descrita en el siguiente diagrama,

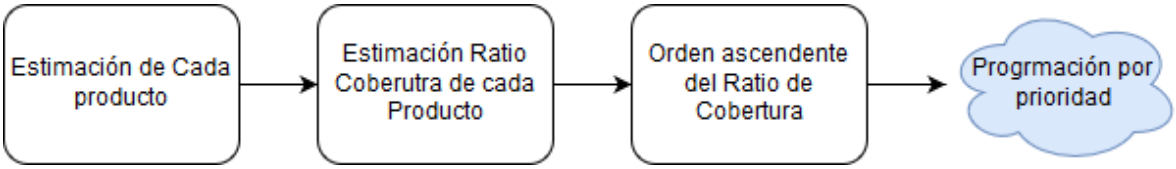

Ilustración 4-6: Diagrama del Procedimiento para la Programación de la Producción en Embonor.

La Ilustración 4-6 describe el procedimiento diario para determinar la programación de la producción. En primer lugar, se estima una demanda para cada producto de acuerdo al modelo actual, luego se estima el ratio de cobertura asociado. Una vez que se tiene el ratio de cobertura para todos los productos estos se separan por cada línea productiva y son ordenados de forma ascendente. Resultado de esto, es una prioridad de entrada de cada producto a cada línea de producción. Esta prioridad de cobertura es mejor cuando el modelo de predicción es más asertivo, por lo tanto, un modelo más asertivo para la estimación de la demanda incide directamente en una mejor programación.

La realización de este estudio nace cuándo el efecto de la mala programación de la producción mantiene un alto nivel de rechazo. Pero, ¿Qué provoca específicamente una mala estimación de la demanda en el ratio de cobertura estimado? La respuesta hace referencia a que es completamente irracional estimar todos los productos indiferentemente con un promedio móvil de orden $n = 6$ debido a que se está suponiendo un comportamiento similar en la demanda de cada uno de los productos. El aumento del rechazo nace justamente cuándo

se visualiza que la programación del día anterior estaba errada debido a la variabilidad del ratio de cobertura.

4.6 Modelos Propuestos

En esta sección se desarrollan las propuestas para los modelos de predicción. Estas propuestas responden al análisis previo de los datos.

La idea general, es estimar cada producto con el modelo que mejor capture sus características. Para ello, es necesario tener una gran variedad de modelos. Cada uno de estos modelos tiene características especiales de acuerdo a alguna de las siguientes variables (i) tendencia, (ii) estacionalidad diaria, (iii) estacionalidad mensual, (iv) correlación entre períodos rezagados, (v) suavizamiento. Es importante mencionar que la elección de estas variables en los modelos no es azarosa, debido a que por ejemplo, se iteraron modelos de predicción con variables como estacionalidad semanal, autocorrelación en los errores, y variables regresoras por categoría de productos. Todas ellas resultaron ser no significativas en los modelos.

Finalmente, mediante las variables seleccionadas se confeccionan 52 modelos de predicción para la demanda más el modelo actual. El conjunto de modelos es 53, utilizados para predecir la demanda en 84 productos distintos. La siguiente tabla muestra las características que pueden asociarse a cada modelo,

Tabla 4-4: Componentes de los Modelos Propuestos

Sigla	Característica
T1	Tendencia orden n=1
T2	Tendencia orden n=2
T3	Tendencia orden n=3
Día	Efecto Día de la Semana
Mes	Efecto Mes del año
AR(1)	Autoregresivo orden a=1
AR(2)	Autoregresivo orden a=2
AR(3)	Autoregresivo orden a=3
SE	Suavizado

Los modelos desarrollados son descritos en detalle dentro de esta sección, sin embargo, a modo de referencia se resumen las características asociadas a cada modelo en la siguiente tabla,

Tabla 4-5: Numeración de los Modelos Propuestos

#	Descripción Modelo
1	Modelo Actual
2	Suavizamiento E. Simple
3	Suavizamiento E. Doble
4	Suavizamiento H.W Multiplicativo
5	Suavizamiento H.W Aditivo
6	T1
7	T2
8	T3
9	T1 Mes
10	T1 Día
11	T1 Mes Día
12	T2 Mes
13	T2 Día
14	T2 Mes Día
15	T3 Mes
16	T3 Día
17	T3 Mes Día
18	T1 Mes AR(1)
19	T1 Día AR(1)
20	T1 Mes Día AR(1)
21	T2 Mes AR(1)
22	T2 Día AR(1)
23	T2 Mes Día AR(1)
24	T3 Mes AR(1)
25	T3 Día AR(1)
26	T3 Mes Día AR(1)
27	T1 Mes AR(1) AR(2)
28	T1 Día AR(1) AR(2)
29	T1 Mes Día AR(1) AR(2)
30	T2 Mes AR(1) AR(2)
31	T2 Día AR(1) AR(2)
32	T2 Mes Día AR(1) AR(2)
33	T3 Mes AR(1) AR(2)
34	T3 Día AR(1) AR(2)
35	T3 Mes Día AR(1) AR(2)
36	T1 Mes AR(1) AR(2) AR(3)
37	T1 Día AR(1) AR(2) AR(3)
38	T1 Mes Día AR(1) AR(2) AR(3)
39	T2 Mes AR(1) AR(2) AR(3)
40	T2 Día AR(1) AR(2) AR(3)
41	T2 Mes Día AR(1) AR(2) AR(3)
42	T3 Mes AR(1) AR(2) AR(3)
43	T3 Día AR(1) AR(2) AR(3)
44	T3 Mes Día AR(1) AR(2) AR(3)
45	T1 Mes SE
46	T1 Día SE
47	T1 Mes Día SE
48	T2 Mes SE
49	T2 Día SE
50	T2 Mes Día SE
51	T3 Mes SE
52	T3 Día SE
53	T3 Mes Día SE

4.6.1 Suavizamiento Exponencial

Se ejecutan los modelos descritos en las ecuaciones (3-3)-(3-6)-(3-7)-(3-11), correspondientes al modelo de suavización exponencial simple, suavizamiento exponencial doble, modelo de Holt-Winters multiplicativo y modelo de Holt-Winters aditivo respectivamente, los parámetros $\alpha,\ \beta,\ \gamma$ dependientes del modelo de estimación a evaluar corresponden a los parámetros óptimos que minimizan el error entre la demanda real respecto a la estimación de cada modelo evaluado. El factor de frecuencia estacional 's' definido en la sección 3.1.2 para ambos modelos de Holt-Winters se establece de acuerdo al análisis desarrollado en la sección 4.4.1 en el cuál se detecta la existencia de una estacionalidad de acuerdo a las temporadas del año (verano con mayor demanda, invierno con menor demanda). El factor de frecuencia estacional es entonces 's' = 4.

4.6.2 Regresión Lineal

Se desarrollan los modelos descritos mediante la ecuación (3-15), donde la variable endógena corresponde a la demanda, y la variable exógena corresponde a la variable correlativa definida en la sección 4.2. Estos modelos consideran la variable tendencia como relevante para la predicción de la demanda. La descripción de los modelos es definida de la siguiente manera,

- Regresión Lineal Simple

$$\hat{d}_{it} = \hat{\alpha}_0 + \hat{\beta}_1 corr_{it} \qquad (4\text{-}4)$$

- Regresión Lineal Doble

$$\hat{d}_{it} = \hat{\alpha}_0 + \hat{\beta}_1 corr_{it} + \hat{\beta}_2 {corr_{it}}^2 \qquad (4\text{-}5)$$

- Regresión Lineal Triple

$$\hat{d}_{it} = \hat{\alpha}_0 + \hat{\beta}_1 corr_{it} + \hat{\beta}_2 {corr_{it}}^2 + \hat{\beta}_3 {corr_{it}}^3 \qquad (4\text{-}6)$$

4.6.3 Regresiones con Variables Exógenas Binarias

La utilización de variables 'Dummy' o binarias en los modelos extendidos de regresiones lineales definidos en la ecuación (3-15) entrega una herramienta para poder detectar alzas y bajas en períodos específicos del horizonte de evaluación, de forma que permite identificar las estacionalidades presentes para cada tipo producto demandado. Según las variables definidas en la sección 4.2 y considerando la variable de tendencia se proponen los siguientes modelos,

- Regresión Lineal de nivel (n) con Efecto Mes

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n {corr_{it}}^n + \hat{\beta}_e M_{e(1x11)} \qquad (4\text{-}7)$$

$$Donde,\ se\ generan\ los\ modelos\ para\ n = 1;\ n = 2;\ n = 3$$

- Regresión Lineal de nivel (n) con Efecto Día de la Semana

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^{n} + \hat{\beta}_l N_{l(1x5)} \qquad (4\text{-}8)$$

$$Donde, se\ generan\ los\ modelos\ para\ n = 1;\ n = 2;\ n = 3$$

- Regresión Lineal de nivel (n) con Efecto Mes y Día de la Semana

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^{n} + \hat{\beta}_e M_{e(1x11)} + \hat{\beta}_l N_{l(1x5)} \qquad (4\text{-}9)$$

$$Donde,\ se\ generan\ los\ modelos\ para\ n = 1;\ n = 2;\ n = 3$$

4.6.4 Regresiones con Variables Exógenas Autoregresivas

Una alternativa para considerar las demandas más recientes dentro de un modelo de regresión es agregar variables exógenas autoregresivas como se encuentra definido en la ecuación (3-18). El modelo ARIMA se compone de variables autoregresivas (AR), un factor de integración (I), y una media móvil (MA). Un trabajo previo de distintas iteraciones considerando las tres variables de los modelos ARIMA entrega como resultado que sólo es significativa la variable autoregresiva, con orden de integración igual a cero y que además el orden autoregresivo significativo es hasta los últimos tres periodos rezagados, es decir, las últimas tres demandas más recientes (o menos). Se detalla los modelos evaluados, que consideran además de las demandas recientes, el factor estacional, y el factor de tendencia.

- Regresión Lineal de nivel (n) con Efecto Mes Autoregresiva de orden (a)

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^n + \hat{\beta}_e M_{e(1x11)} + \sum_{a=1}^{a} \hat{\beta}_a d_{i,t-a} \qquad (4\text{-}10)$$

$Donde, se\ generan\ los\ modelos\ para\ el\ nivel\ n = 1;\ n = 2;\ n = 3,$
$Y\ \ para\ el\ orden\ autoregresivo\ a = 1;\ a = 2;\ a = 3.$

- Regresión Lineal de nivel (n) con Efecto Día de la Semana Autoregresiva de orden (a)

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^n + \hat{\beta}_l N_{l(1x5)} + \sum_{a=1}^{a} \hat{\beta}_a d_{i,t-a} \qquad (4\text{-}11)$$

$Donde, se\ generan\ los\ modelos\ para\ el\ nivel\ n = 1;\ n = 2;\ n = 3,$
$Y\ para\ el\ orden\ autoregresivo\ a = 1;\ a = 2;\ a = 3$

- Regresión Lineal de nivel (n) con Efecto Mes y Día de la Semana Autoregresiva de orden (a)

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^n + \hat{\beta}_e M_{e(1x11)} + \hat{\beta}_l N_{l(1x5)} + \sum_{a=1}^{a} \hat{\beta}_a d_{i,t-a} \qquad (4\text{-}12)$$

$Donde, se\ generan\ los\ modelos\ para\ el\ nivel\ n = 1;\ n = 2;\ n = 3,$
$Y\ para\ el\ orde\ autoregresivo\ a = 1;\ a = 2;\ a = 3$

4.6.5 Regresiones con Variables Exógenas Suavizadas

Otra alternativa para considerar las demandas más recientes dentro de un modelo de regresión lineal, es implementar un modelo que contenga dentro de sus variables exógenas una estimación a través de cualquier otro modelo, en este caso particular, un modelo determinado por la suavización de sus demandas más recientes. Mediante iteraciones realizadas detallas más adelante en este estudio, se determina que la mejor variable a utilizar es la estimación a través del modelo de Holt-Winters multiplicativo. Por lo tanto, se agrega a los modelos que consideran un factor estacional y un factor de tendencia una variable suavizada que representa el factor de ponderación de las demandas más recientes. Se definen los modelos como,

- Regresión Lineal de nivel (n) con Efecto Mes con Suavizamiento

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^{n} + \hat{\beta}_e M_{e(1x11)} + SE_{it} \qquad (4\text{-}13)$$

$$Donde, se\ generan\ los\ modelos\ para\ n = 1;\ n = 2;\ n = 3$$

- Regresión Lineal de nivel (n) con Efecto Día de la Semana con Suavizamiento

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^{n} + \hat{\beta}_l N_{l(1x5)} + SE_{it} \qquad (4\text{-}14)$$

$$Donde, se\ generan\ los\ modelos\ para\ n = 1;\ n = 2;\ n = 3$$

- Regresión Lineal de nivel (n) con Efecto Mes y Día de la Semana con Suavizamiento

$$\hat{d}_{it} = \hat{\alpha}_0 + \sum_{n=1}^{n} \hat{\beta}_n corr_{it}{}^{n} + \hat{\beta}_e M_{e(1x11)} + \hat{\beta}_l N_{l(1x5)} + SE_{it} \qquad (4\text{-}15)$$

$$Donde, se\ generan\ los\ modelos\ para\ n = 1;\ n = 2;\ n = 3$$

4.7 Procedimiento y Metodología

El procedimiento general para obtener los resultados de este estudio sigue el comportamiento descrito en el siguiente diagrama,

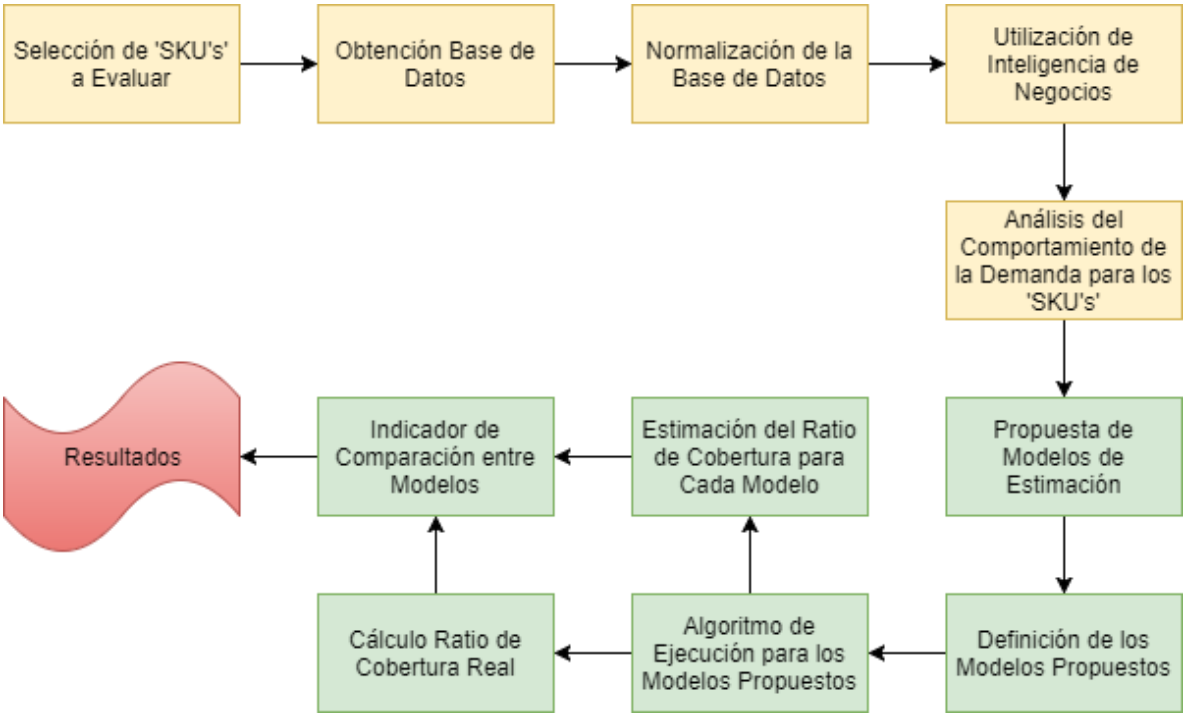

Ilustración 4-7: Diagrama Descriptivo de la Metodología Aplicada.

Los recuadros anaranjados representan el análisis a priori de este estudio, mientras que los recuadros en verde representan el análisis a posteriori.

El análisis a posteriori tiene por objetivo el desarrollo de un algoritmo que permita identificar el modelo óptimo de predicción para cada producto dentro del espectro total de modelos propuestos. A grandes rasgos el algoritmo realiza una estimación por cada modelo propuesto para cada producto. Luego estima un ratio de cobertura que es comparado respecto al ratio de cobertura real. Finalmente, establece en base a un criterio de selección el modelo óptimo por 'SKU'.

4.7.1 Ejecución de los Modelos Propuestos

La ejecución se desarrolla en el Software estadístico Eviews 9 ('Student Versión'), los comandos de ejecución para la codificación se encuentran en el siguiente enlace: http://www.eviews.com/help/helpintro.html

A modo general, se confecciona un programa que genere como resultado 53 matrices. Una para cada modelo. Se definen estas como,

$$M_{m,(t,i)} = Estimación\ del\ modelo\ m\ para\ el\ producto\ i\ en\ el\ período\ t.$$

$$M_{m,(t,i)} = \begin{bmatrix} \widehat{e_{1;1}} & \cdots & \widehat{e_{1;84}} \\ \vdots & \ddots & \vdots \\ \widehat{e_{1025;1}} & \cdots & \widehat{e_{1025;84}} \end{bmatrix}_m$$

Cómo cada modelo es diferente se desarrolla una codificación para cada uno de ellos de acuerdo a las ecuaciones definidas en la sección 4.6, el código para cada modelo se describe en el anexo 2

4.7.2 Confección de Matriz para el Ratio de Cobertura Real

El ratio de cobertura real se define como el alcance que tiene el inventario de un producto en una fecha determinada ‘t’ hasta ‘n’ períodos futuros, es decir, para cuántos días futuros alcanza el inventario de hoy hasta que se genere el punto de quiebre de stock. Este ratio se confecciona dentro del horizonte T’. Definido de la siguiente manera,

$$R_{it} = [n], \qquad tal\ que\ \frac{S_{it}}{\sum_{t}^{t+n} D_{it}} = 1 \tag{4-16}$$

Mediante la codificación descrita en el anexo 1 se obtiene la matriz del ratio de cobertura real para cada producto, la cual se define de la siguiente manera,

$$R_{real,(t,i)} = \begin{bmatrix} R_{1;1} & \cdots & R_{1;84} \\ \vdots & \ddots & \vdots \\ R_{97;1} & \cdots & R_{97;84} \end{bmatrix}$$

4.7.3 Confección de Matriz para el Ratio de Cobertura por Modelo

El ratio de cobertura por modelo se define como el alcance que tiene el inventario de un producto en una fecha determinada ‘t’ hasta ‘n’ períodos futuros de acuerdo a la demanda estimada del modelo ‘m’, es decir, para cuántos días futuros alcanza el inventario de hoy hasta que la demanda estimada de los ‘n’ períodos futuros del modelo ‘m’ genera el punto de quiebre de stock. Este ratio de cobertura por modelo se confecciona dentro del horizonte T’. Definido de la siguiente manera,

$$\hat{R}_{itm} = [n], \qquad tal\ que\ \frac{S_{it}}{\sum_{t}^{t+n} M_{m(t,i)}} = 1 \tag{4-17}$$

Mediante la codificación descrita en el anexo 2 se obtienen las 53 matrices para el ratio de cobertura estimado de cada modelo. Definida de la siguiente manera,

$$\hat{R}_{m,(t,i)} = \begin{bmatrix} \hat{R}_{1;1} & \cdots & \hat{R}_{1;84} \\ \vdots & \ddots & \vdots \\ \hat{R}_{97;1} & \cdots & \hat{R}_{97;84} \end{bmatrix}_m$$

4.7.4 Indicador Comparativo de Selección de Modelo Óptimo por 'SKU'

Una vez desarrollado tanto el cálculo para el ratio de cobertura real como para el ratio de cobertura estimado, es posible determinar una medida de dispersión en relación al indicador de cobertura. Se desea conocer qué tan alejado está el ratio de cobertura estimado respecto al ratio de cobertura real para cada período 't' del horizonte de evaluación T'. Para ello se define s_{itm} como la diferencia entre ellos de la siguiente manera,

$$s_{itm} = R_{it} - \hat{R}_{itm} \qquad \forall \ t \in T' \qquad (4\text{-}18)$$

Estas diferencias son almacenadas en 53 matrices. Una para cada modelo, y de la siguiente manera,

$$s_{m,(t,i)} = \begin{bmatrix} s_{1;1} & \cdots & s_{1;84} \\ \vdots & \ddots & \vdots \\ s_{97;1} & \cdots & s_{97;84} \end{bmatrix}_m$$

Luego, se determina la cuasivarianza muestral para el horizonte de evaluación T' de la siguiente manera,

$$I_{i,m} = \frac{\sum_{t=1}^{n} (s_{itm} - \overline{s_{itm}})^2}{n-1}, \qquad \forall n \in T' \qquad (4\text{-}19)$$

Donde, $I_{i,m}$ representa la cuasivarianza muestral para el producto ‘i’ en el modelo ‘m’. Estas son almacenas de la siguiente manera,

$$I_{(m,i)} = \begin{bmatrix} I_{1;1} & \cdots & I_{84;1} \\ \vdots & \ddots & \vdots \\ I_{1;53} & \cdots & I_{84;53} \end{bmatrix}$$

4.7.5 Ejecución del Algoritmo

La ejecución del algoritmo se desarrolla mediante el código descrito en el anexo 7.2. Y se desarrolla mediante la ejecución de los siguientes pasos,

- Paso 1. Se calcula $R_{real,(t,i)}$
- Paso 2. Se define ‘m’ = 0
- Paso 3. Se genera la estimación $M_{m,(t,i)}$ para ‘m’ = ‘m’ + 1
- Paso 4. Se calcula $\hat{R}_{m,(t,i)}$ para ‘m’ = ‘m’ + 1
- Paso 5. Se calcula $s_{m,(t,i)}$ para ‘m’ = ‘m’ + 1
- Paso 6. Se almacena $I_{(m,i)}$ para ‘m’ = ‘m’ + 1
- Paso 7. Se define ‘m’ = ‘m’ + 1
- Paso 8. Si $I_{(m,i)} < I_{(m+1,i)} \Rightarrow$ Modelo ‘m’ es óptimo.
  $\sim$ Modelo ‘m’ + 1 es óptimo. $(\forall i \in I)$
- Paso 9. Si ‘m’ < 53 volver al paso 2.

4.8 Resultados Generales

Esta sección muestra (i) el mejor modelo para cada uno de los 84 productos y (ii) las mejoras respecto al modelo actual. El primer punto, representa el resultado final de este estudio el cual es aplicable a la programación en Embonor. El segundo punto, entrega una medida de qué tan mejor es la solución aplicada en el estudio.

4.8.1 Modelos Óptimos por 'SKU'

Al ejecutar el algoritmo se obtiene como resultado la matriz $I_{(m,i)}$. Esta contiene la cuasivarianza del modelo 'm' en el producto 'i'. Por lo tanto, es posible identificar para cada producto 'i' el modelo que tenga la mínima cuasivarianza. Aquellos modelos que tengan la cuasivarianza más baja están pronosticando de mejor forma el ratio de cobertura debido a una menor variabilidad en el error de estimación. La siguiente tabla muestra el número de modelo óptimo para cada producto,

Tabla 4-6: Modelo Óptimo por SKU

SKU	#
SKU1	18
SKU2	14
SKU3	17
SKU4	6
SKU5	36
SKU6	42
SKU7	42
SKU8	24
SKU9	21
SKU10	24
SKU11	12
SKU12	36
SKU13	42
SKU14	8
SKU15	36
SKU16	12
SKU17	15
SKU18	7
SKU19	45
SKU20	9
SKU21	39
SKU22	36
SKU23	42
SKU24	39
SKU25	45
SKU26	9
SKU27	51
SKU28	48
SKU29	51
SKU30	7
SKU31	45
SKU32	53
SKU33	36
SKU34	51
SKU35	3
SKU36	42
SKU37	8
SKU38	36
SKU39	14
SKU40	36
SKU41	8
SKU42	36
SKU43	32
SKU44	42
SKU45	42
SKU46	45
SKU47	53
SKU48	36
SKU49	15
SKU50	42
SKU51	36
SKU52	51
SKU53	51
SKU54	45
SKU55	48
SKU56	48
SKU57	36
SKU58	36
SKU59	51
SKU60	53
SKU61	39
SKU62	16
SKU63	15
SKU64	3
SKU65	24
SKU66	1
SKU67	36
SKU68	1
SKU69	1
SKU70	12
SKU71	9
SKU72	39
SKU73	48
SKU74	6
SKU75	44
SKU76	39
SKU77	1
SKU78	44
SKU79	8
SKU80	15
SKU81	32
SKU82	1
SKU83	52
SKU84	6

Del espectro de 53 modelos posibles, existen 24 que al menos minimizan la cuasivarianza para algún producto. Por lo tanto, existen 29 modelos que no son utilizados. La siguiente tabla muestra en su primera columna el número del modelo propuesto, en su segunda columna las características asociadas al modelo, y en su tercera columna la frecuencia de productos que seleccionan el modelo como óptimo.

Tabla 4-7: Frecuencia de Selección de Modelos

#	Descripción Modelo	# SKU's
1	Modelo Actual	5
3	Suavizamiento E. Doble	2
6	T1	3
7	T2	2
8	T3	4
9	T1 Mes	3
12	T2 Mes	3
14	T2 Mes Día	2
15	T3 Mes	4
16	T3 Día	1
17	T3 Mes Día	1
18	T1 Mes AR(1)	1
21	T2 Mes AR(1)	1
24	T3 Mes AR(1)	3
32	T2 Mes Día AR(1) AR(2)	2
36	T1 Mes AR(1) AR(2) AR(3)	13
39	T2 Mes AR(1) AR(2) AR(3)	5
42	T3 Mes AR(1) AR(2) AR(3)	8
44	T3 Mes Día AR(1) AR(2) AR(3)	2
45	T1 Mes SE	5
48	T2 Mes SE	4
51	T3 Mes SE	6
52	T3 Día SE	1
53	T3 Mes Día SE	3

Se observa de la tabla 4-7 los siguientes resultados. (i) El modelo que captura mejor la demanda para la mayor cantidad de productos, es aquel que presenta tendencia, estacionalidad mensual, y correlación de tres rezagos. (ii) El modelo actual captura mejor el comportamiento de la demanda para el 6 [%] de los productos. (iii) 94 [%] de los productos pueden ser mejor estimados mediante los modelos propuestos.

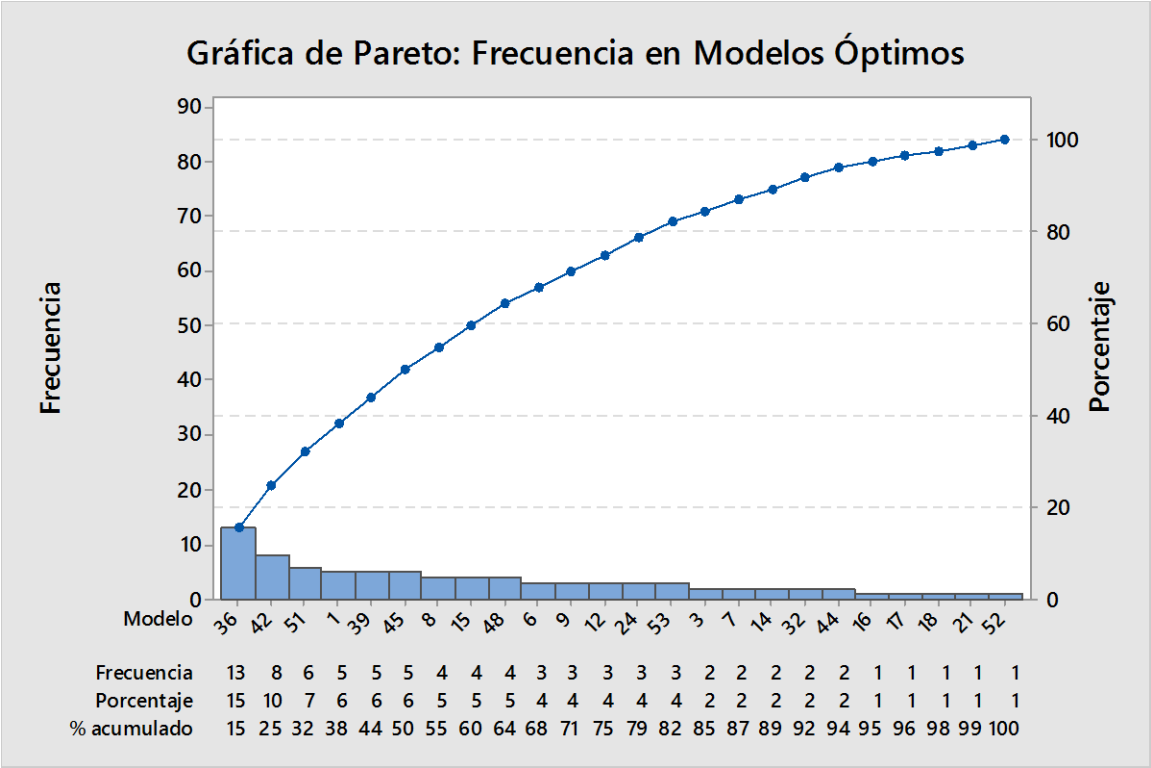

Ilustración 4-8: Gráfica de Pareto, muestra la frecuencia de los modelos óptimos seleccionados, el porcentaje respecto al espectro total de productos evaluados, y el porcentaje acumulado asociado.

La Ilustración 4-10 muestra de izquierda a derecha los modelos que más veces fueron seleccionados como óptimos. Por ejemplo, muestra que el modelo número 36, tiene una frecuencia de 13, el cual representa el 15 [%] del espectro total de productos evaluados. También es observable que el 50[%] de los productos seleccionan entre los primeros 6 modelos. Y que el 82[%] de los productos seleccionan entre 14 modelos.

Los 24 modelos de predicción seleccionados, están compuesto por al menos una de las siguientes características (i) componente de tendencia, (ii) componente de estacionalidad, (iii) componente autoregresiva (iv) componente suavizada. Por lo tanto, es posible realizar una agrupación según las componentes que contenga cada modelo. Ver Tabla 4-8.

Tabla 4-8: Categoría de Modelos de Predicción y Su Frecuencia.

Categoría de Modelo	Descripción de Categoría	Frecuencia
Modelo Actual	Promedio Móvil	5
S	Suavizamiento	2
T	Componente Tendencia	9
T;E	Componente Tendencia y Estacionalidad	14
T;E;A	Componente Tendencia, Estacionalidad y Autoregresiva	35
T;E;S	Componente Tendencia, Estacionalidad y Suavizada	19

Mediante la tabla anterior, se confecciona un diagrama de Pareto (ver Ilustración 4-9) para visualizar el comportamiento de las componentes de predicción.

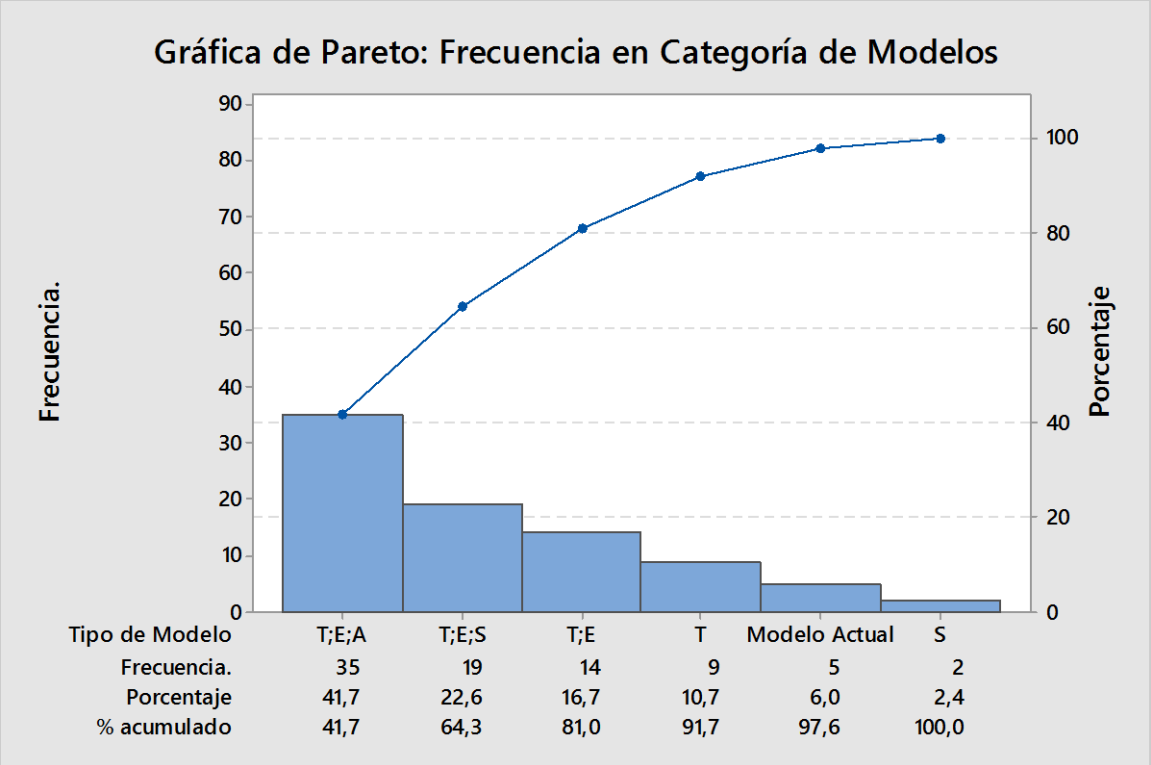

Ilustración 4-9: Gráfica de Pareto, muestra la frecuencia por categoría de modelos seleccionados.

De la gráfica anterior se observa que, el 41,7 [%] corresponden a modelos que contienen las componentes de tendencia, estacionalidad, y autoregresivas dentro de sus variables de predicción. El 22,6 [%] corresponden a modelos que contienen las componentes de tendencia, estacionalidad, y suavizamiento dentro de sus variables de predicción.

También es observable que, el 91,7 [%] corresponden a modelos que contienen la componente de tendencia, y el 81,0 [%] corresponden a modelos que contienen al menos una componente de estacionalidad.

4.8.2 Método Seleccionado v/s Modelo Actual

Para cuantificar qué tan mejor son los modelos seleccionados respecto al modelo actual se adjunta la siguiente tabla,

Tabla 4-9: Indicador de Mejora Respecto al Modelo Actual

#	CuasiVarianza		[%] Mejora
	Óptimo	Actual	
1	15.80	64.80	76%
2	25.01	51.81	52%
3	17.17	55.28	69%
4	4.14	12.70	67%
5	1.80	6.79	74%
6	140.50	896.81	84%
7	25.11	59.02	57%
8	15.70	34.58	55%
9	9.43	48.16	80%
10	3.84	8.51	55%
11	1.61	2.21	27%
12	1.22	6.22	80%
13	2.24	8.53	74%
14	84.99	1233521	100%
15	13.65	49.17	72%
16	19.93	46.07	57%
17	7.31	73.10	90%
18	12.28	61.84	80%
19	1.01	3.58	72%
20	0.17	0.20	16%
21	17.83	48.94	64%
22	1.38	4.25	67%
23	37.23	1166038	100%
24	5.64	17.88	68%
25	12.91	14.03	8%
26	0.89	1.42	37%
27	1.73	1.96	12%
28	1.01	1.73	41%
29	0.35	0.45	22%
30	0.70	0.92	24%
31	0.18	0.20	8%
32	0.13	0.14	5%
33	0.83	1.10	25%
34	0.23	0.30	23%
35	5.13	7.84	35%
36	2.73	7.72	65%
37	41.96	111099	100%
38	3.62	13.10	72%
39	5.44	10.50	48%
40	1.15	2.06	44%
41	26.09	68.44	62%
42	4.35	6.47	33%
43	4.73	831.12	99%
44	0.38	0.77	51%
45	0.72	1.06	32%
46	0.53	0.58	10%
47	0.32	0.37	13%
48	1.47	4.60	68%
49	138.85	210.46	34%
50	58.39	393.33	85%
51	2.22	5.71	61%
52	6.90	7.55	9%
53	1.12	1.22	8%
54	0.56	0.62	9%
55	4.28	5.25	18%
56	0.83	1.01	18%
57	0.21	0.30	30%
58	1.32	4.04	67%
59	18.08	25.60	29%
60	14.18	17.31	18%
61	4.78	17.21	72%
62	137.38	208.60	34%
63	27.32	68.92	60%
64	6.48	10.27	37%
65	1.41	4.88	71%
66	35.78	35.78	0%
67	4.03	6.53	38%
68	0.28	0.28	0%
69	15.07	15.07	0%
70	1.60	4.04	60%
71	2.68	34.61	92%
72	0.68	1.11	39%
73	3.30	3.71	11%
74	3.81	3816	100%
75	0.80	0.93	14%
76	4.26	11.21	62%
77	0.59	0.59	0%
78	1.31	1.94	32%
79	5.64	6.71	16%
80	292.03	3909	93%
81	86.18	2307	96%
82	9.44	9.44	0%
83	22.29	25.45	12%
84	14.81	178.41	92%

Esta tabla muestra las cuasivarianzas del modelo óptimo respecto a las del modelo actual. La última columna de la tabla muestra el porcentaje de mejora en la estimación calculado como,

$$[\%]\ de\ Mejora = 1 - \left[\frac{CuasiVarianza\ Modelo\ Óptimo}{CuasiVarianza\ Modelo\ Actual}\right] \qquad (4\text{-}20)$$

La gráfica muestra el porcentaje de mejora definido para los 84 productos ordenados ascendentemente. Es evidente entonces, que el modelo actual en Embonor no es el óptimo para programar la producción en base al criterio de cobertura, debido a que tan solo 5 de los 84 productos continúan eligiendo el modelo actual, pero que sin embargo el resto ha disminuido considerablemente la variabilidad en el error de estimación.

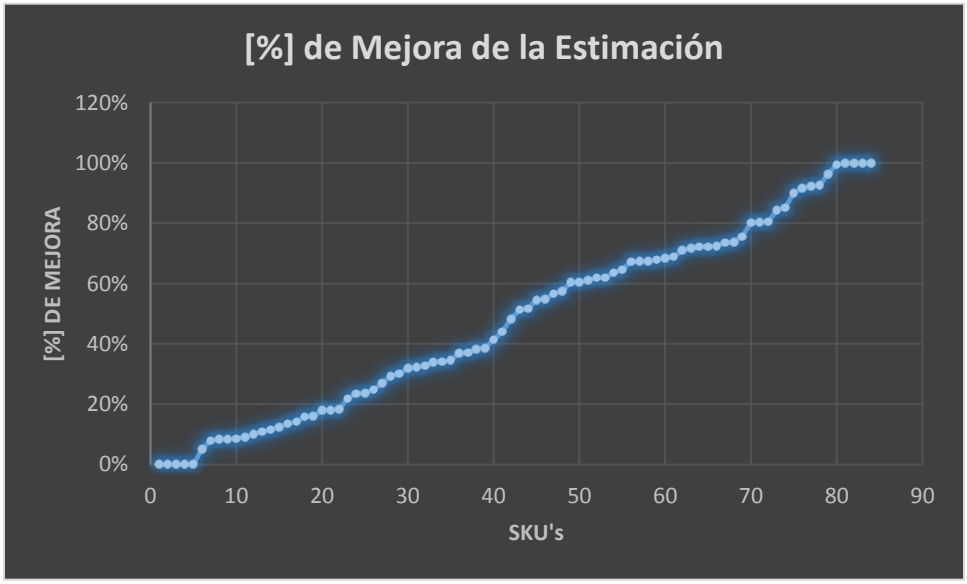

Ilustración 4-10: Gráfica que relaciona el [%] de mejora de los modelos propuestos respecto al modelo actual

Los resultados generales indican que el ratio de cobertura estimado mejoraría en un promedio de un 48 [%] con un 95 [%] de confiablidad (Ver Ilustración 4-11).

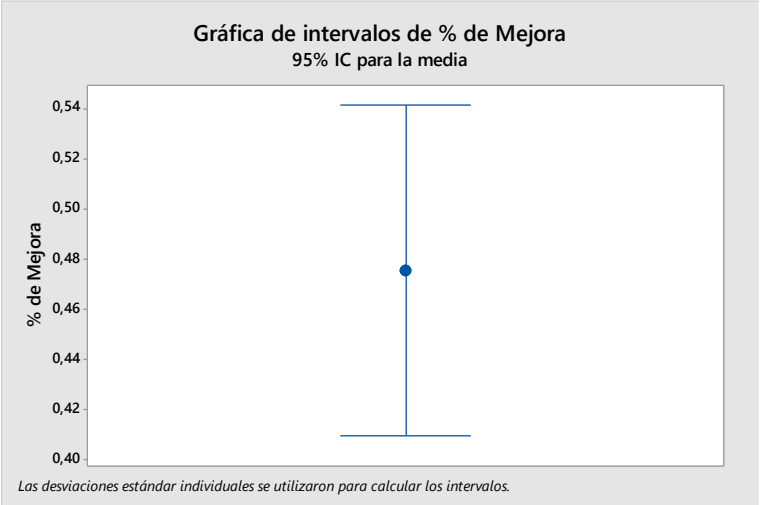

Ilustración 4-11: Media porcentual del indicador de mejora de modelo seleccionado respecto al modelo actual

Finalmente, se observa que las categorías de modelos difieren en sus comportamientos asociados al porcentaje de mejora (Ver Ilustración 4-12). Es posible visualizar, por ejemplo, que aquellos modelos seleccionados que únicamente tienen una componente de tendencia, mejoran respecto al modelo actual por sobre un 50 [%]. Luego, también es observable lo siguiente,

- Aquellos modelos que se encuentran dentro de la categoría que contiene exclusivamente las componentes de tendencia y estacionalidad, poseen un porcentaje de mejora respecto al modelo actual distribuido uniformemente.

- Aquellos modelos que se encuentran dentro de la categoría que contiene exclusivamente las componentes de tendencia, estacionalidad, y autoregresiva, poseen un porcentaje de mejora por sobre el promedio del resto de las categorías de modelos.

- Aquellos modelos que se encuentran dentro de la categoría que contiene exclusivamente las componentes de tendencia, estacionalidad, y suavizamiento, poseen un porcentaje de mejora por bajo el promedio del resto de las categorías de modelos.

Por lo tanto, es posible determinar que aquellos modelos que contienen una componente autoregresiva, representan una mejora de predicción mucho mejor respecto a aquellos modelos que contienen una componente suavizada.

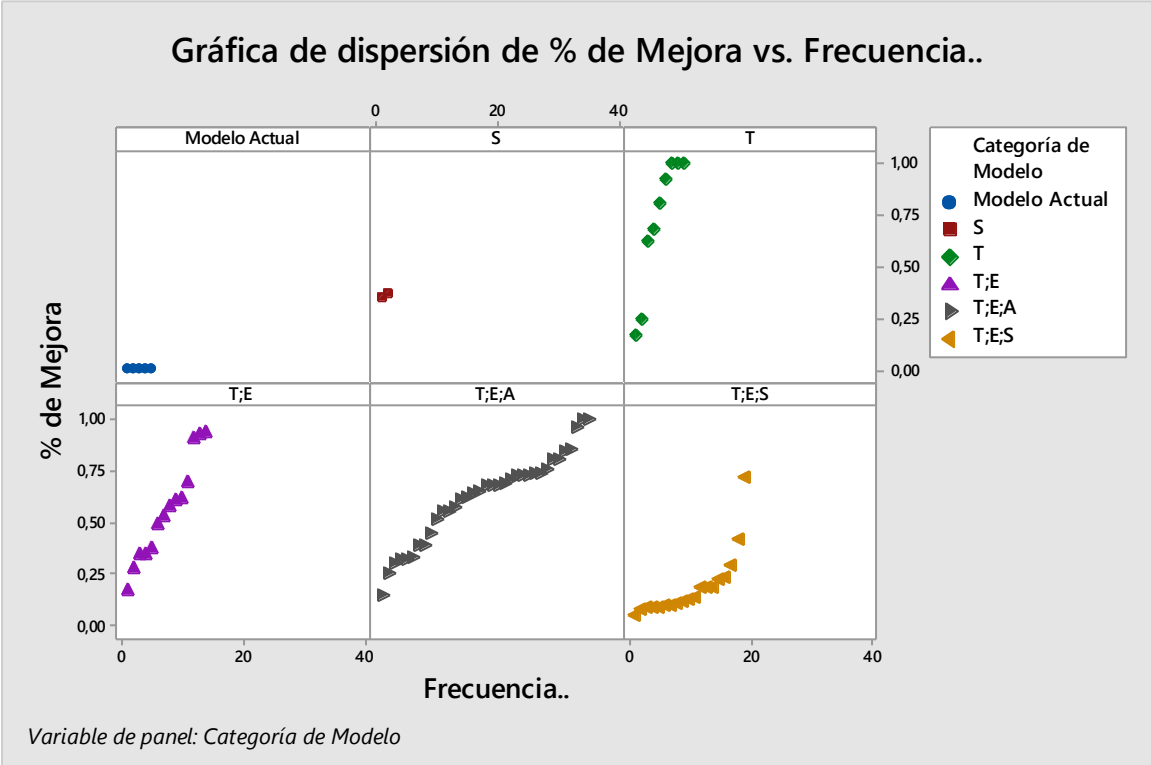

Ilustración 4-12: Frecuencia de selección, respecto al porcentaje de mejora, por categoría de modelo.

5 Conclusiones y Comentarios

El algoritmo entrega como resultados generales una mejora promedio del 48 [%]. Por lo tanto, este estudio ha mejorado la estimación del ratio de cobertura respecto a la estimación del modelo actual. Se espera como resultado directo de esto, una mejor programación de la producción, y una disminución de la demanda insatisfecha.

Del total de productos, un 94 [%] ha seleccionado un mejor modelo de estimación respecto al actual, es decir, solo el 6 [%] de los productos continúa capturando el comportamiento de la demanda a través de un promedio móvil orden seis. El 94 [%] representa 78 productos, de los cuales, 5 han mejorado su estimación en un 100 [%].

Del conjunto de 53 modelos inicialmente propuestos, solo 24 modelos han sido seleccionados por al menos un producto como óptimo. Estos modelos poseen características de tendencia (T), estacionalidad (E), suavizamiento (S) y autoregresión (A). La componente de tendencia está presente en los modelos para 76 productos, mientras que la componente de estacionalidad se encuentra presente para 68 productos.

El modelo que captura el comportamiento de la demanda para más 'SKU's' es el modelo que contiene tendencia de primer orden, con estacionalidad mensual y sus tres componentes autoregresivas. Este modelo es seleccionado como óptimo por 13 de los 84 productos. El 50 [%] de los productos seleccionan como óptimo entre tan solo 6 modelos diferentes.

Existen categorías de modelos que tienen mejores resultados respecto a otras. Un 42 [%] de los productos seleccionan como categoría óptima los modelos T; E; A, mientras que un 23 [%] la categoría T; E; S, y un 17 [%] la categoría T; E. Además, la categoría que captura la mayor cantidad de productos (T, E; A) es también la categoría que tiene un mayor porcentaje de mejora respecto al modelo actual.

El método actual de estimación de la demanda utilizado en Embonor S.A no es capaz de capturar el comportamiento característico de cada producto. Este estudio entrega una metodología, para determinar dentro de un gran espectro de modelos aquel que es más característico a las propiedades intrínsecas de la demanda de cada producto.

Como comentarios generales en relación al estudio, Embonor S.A debe implementar constantemente nuevas metodologías para realizar pronósticos de demanda a la vez que se van agregando nuevos productos al portafolio de ventas. Es también imprescindible, monitorear la asertividad de cada modelo constantemente, implementando una herramienta de control y seguimiento.

Finalmente, es interesante contestar a la siguiente pregunta: ¿Es el criterio de ratio de cobertura el mejor método para programar la producción del corto-mediano plazo? Este método (calculado mediante un buen pronóstico) ayuda a detectar la criticidad del nivel de inventario de los productos asociados a cada línea, sin embargo, no considera algunos factores relevantes para la programación, como por ejemplo, el tiempo de producción de cada formato, los SetUp de cada producto, los tiempos promedios de falla, la utilización y eficiencia de cada línea productiva, variables que sin duda inciden constantemente en la toma de decisiones. La mejor forma de reducir la demanda insatisfecha para Embonor, es un modelo de optimización que considere todas las variables de producción, cuya función objetivo sea la de minimizar los rechazos.

6 Referencias

Axsater, S. (2015). *International Series in Operation Research & Management science* (Vol. third Edition). CA, USA: Springer.

Jenkins, B. a. (2013). The Box-Jenkins Methodology for Time Series Model.

Parr, O. (2000). *Data Mining Cookbook: Modeling Data for Marketing, Risk and Costomer Relationship Management.* TX, USA: Wiley.

Sobrosa, F. (2013). Information architecture analysis using business intelligence tools based on the information needs of executives. *JISTEM - Journal of Information Systems and Technology Management*.

7 Anexos

7.1 Anexo 1

Tabla 7-1: 84 SKU's A Evaluar.

SKU	DESCRIPCIÓN	LÍNEA
SKU1	BENE_CG_X06_PE1500C	L5
SKU2	BENE_CG_X06_PE2000C	L5
SKU3	BENE_CG_X06_PET500C	L5
SKU4	BENE_SG_20_LTS_BIB	L4
SKU5	BENE_SG_X02_PE6000C	L4
SKU6	BENE_SG_X06_PE1500C	L5
SKU7	BENE_SG_X06_PE2000C	L5
SKU8	BENE_SG_X06_PET500C	L5
SKU9	COKE_LIGHX06_PE1000	L3
SKU10	COKE_LIGHX06_PET591	L5
SKU11	COKE_LIGHX24_VRE237	L2
SKU12	COKE_LIGX04_RP2000C	L1
SKU13	COKE_LIGX06_PE1500C	L5
SKU14	COKE_LIGX06_PE1750C	L3
SKU15	COKE_LIGX06_PE2500C	L5
SKU16	COKE_LIGX06_PE3000C	L5
SKU17	COKE_LIGX06_RP2500C	L1
SKU18	COKE_LIGX06_RP3000C	L1
SKU19	COKE_LIGX24_VRE350C	L2
SKU20	COKE_X04_RP2000CC	L1
SKU21	COKE_X06_PET1000CC	L3
SKU22	COKE_X06_PET1500CC	L5
SKU23	COKE_X06_PET1750CC	L3
SKU24	COKE_X06_PET2500CC	L5
SKU25	COKE_X06_PET250CC	L5
SKU26	COKE_X06_PET3000CC	L5
SKU27	COKE_X06_PET591CC	L5
SKU28	COKE_X06_RP2500CC	L1
SKU29	COKE_X06_RP3000CC	L1
SKU30	COKE_X06_RPE1500CC	L1
SKU31	COKE_X06_VR1000CC	L2
SKU32	COKE_X24_VR237CC	L2
SKU33	COKE_X24_VR350CC	L2
SKU34	COKE_ZERO_X04_RP2_0	L1
SKU35	COKE_ZERO_X06_P1000	L3
SKU36	COKE_ZERO_X06_P1500	L5
SKU37	COKE_ZERO_X06_P1750	L3
SKU38	COKE_ZERO_X06_P2500	L5
SKU39	COKE_ZERO_X06_P3000	L5
SKU40	COKE_ZERO_X06_PE591	L5
SKU41	COKE_ZERO_X06_RP1_5	L1
SKU42	COKE_ZERO_X06_RP2_5	L1
SKU43	COKE_ZERO_X06_VR1_0	L2
SKU44	COKE_ZERO_X24_VR237	L2
SKU45	COKE_ZERO_X24_VR350	L2
SKU46	COKE_ZEROX06_RP3000	L1
SKU47	F_NAR_X04_RP2000CC	L1
SKU48	F_NAR_X06_PET1500CC	L3
SKU49	F_NAR_X06_PET2500CC	L3
SKU50	F_NAR_X06_PET250CC	L5
SKU51	F_NAR_X06_PET3000CC	L5
SKU52	F_NAR_X06_PET591CC	L5
SKU53	F_NAR_X06_RP1500CC	L1
SKU54	F_NAR_X06_RP3000CC	L1
SKU55	F_NAR_X06_RPE2500CC	L1
SKU56	F_NAR_X06_VR1000CC	L2
SKU57	F_NAR_X24_VR237CC	L2
SKU58	F_NAR_X24_VR350CCPR	L2
SKU59	F_NAZ_CX06_PE1500CC	L3
SKU60	FANTA_X06_PET1100CC	L3
SKU61	INCA_KO_ZEX06_P1500	L3
SKU62	NOR_GA_LIX24_VRET350	L2
SKU63	NOR_TO_LIX24_VRET350	L2
SKU64	QUA_PO_ZEX06_PE1500	L3
SKU65	SPR_ZEROX04_RPE2000	L1
SKU66	SPR_ZEROX06_PET1500	L3
SKU67	SPR_ZEROX24_VR350CC	L2
SKU68	SPRITE_X04_RP2000CC	L1
SKU69	SPRITE_X06_PET1100C	L3
SKU70	SPRITE_X06_PET1500C	L3
SKU71	SPRITE_X06_PET2500C	L3
SKU72	SPRITE_X06_PET3000CC	L5
SKU73	SPRITE_X06_PET591CC	L5
SKU74	SPRITE_X06_RP1500CC	L1
SKU75	SPRITE_X06_RP3000CC	L1
SKU76	SPRITE_X06_RPE2500C	L1
SKU77	SPRITE_X06_VR1000CC	L2
SKU78	SPRITE_X24_VR350CC	L2
SKU79	SPRITE_X24_VRE237CC	L2
SKU80	TAI_LIL_X06_PET2000	L5
SKU81	TAI_MAN_X06_PET2000	L5
SKU82	TAI_NAR_X06_PET2000	L5
SKU83	TAI_PAP_X06_PET2000	L5
SKU84	TAI_PIN_X06_PET2000	L5

7.2 Anexo 2: Código algoritmo generador de resultados.

```
smpl @all


matrix(114, 84) _dias_piso

for !j = 1 to 84

    for !i = 1 to 114

        %demanda_acum = @str(!i) + " " + @str(114)
        series pisos = @cumsum(demandasku{!j}, %demanda_acum) <= @elem(stocksku{!j}, !i)
        _dias_piso(!i, !j) = @sum(pisos)

    next

next



for  !i =1 to 84

    group _SKU
    _SKU.add sku{!i}
    group _STOCKSKU
    _STOCKSKU.add STOCKSKU{!i}
    group _Modelo1
    series Moving_Average_6sku{!i} =( @movsum(sku{!i}, 6) ) / 6
    _Modelo1.add Moving_Average_6sku{!i}
    group _tendencia
    equation tendencia_sku{!i}.ls sku{!i} corr
    forecast tendenciasku{!i}
    _tendencia.add Tendenciasku{!i}


    for !l=2 to 5

        group _Modelo{!l}
        smooth(s)  sku{!i} SE2_sku{!i}
        smooth(d)  sku{!i} SE3_sku{!i}
        smooth(n)  sku{!i} SE4_sku{!i}
        smooth(a)  sku{!i} SE5_sku{!i}
        _Modelo{!l}.add SE{!l}_sku{!i}

    next

    for !k=6 to 53

        group _Modelo{!k}

        equation Reg1_sku{!i}.ls sku{!i} c corr
        forecast sku{!i}_6

        equation Reg2_sku{!i}.ls sku{!i} c corr corr^2
        forecast sku{!i}_7

        equation Reg3_sku{!i}.ls sku{!i} c corr corr^2 corr^3
        forecast sku{!i}_8
```

```
equation Reg4_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast)
forecast sku{!i}_9

equation Reg5_sku{!i}.ls sku{!i} c corr lunes martes miercoles jueves viernes
forecast sku{!i}_10

equation Reg6_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast)  lunes martes miercoles jueves viernes
forecast sku{!i}_11

equation Reg7_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast)
forecast sku{!i}_12

equation Reg8_sku{!i}.ls sku{!i} c corr corr^2 lunes martes miercoles jueves viernes
forecast sku{!i}_13

equation Reg9_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast) lunes martes miercoles jueves
viernes
forecast sku{!i}_14

equation Reg10_sku{!i}.ls sku{!i} c corr corr^2 corr^3 @expand(mes, @droplast)
forecast sku{!i}_15

equation Reg11_sku{!i}.ls sku{!i} c corr corr^2 corr^3 lunes martes miercoles jueves viernes
forecast sku{!i}_16

equation Reg12_sku{!i}.ls sku{!i} c corr corr^2 corr^3  @expand(mes, @droplast) lunes martes miercoles
jueves viernes
forecast sku{!i}_17

equation Reg13_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast) sku{!i}(-1)
forecast sku{!i}_18

equation Reg14_sku{!i}.ls sku{!i} c corr lunes martes miercoles jueves viernes sku{!i}(-1)
forecast sku{!i}_19

equation Reg15_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast)  lunes martes miercoles jueves
viernes sku{!i}(-1)
forecast sku{!i}_20

equation Reg16_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast)  sku{!i}(-1)
forecast sku{!i}_21

equation Reg17_sku{!i}.ls sku{!i} c corr corr^2 lunes martes miercoles jueves viernes sku{!i}(-1)
forecast sku{!i}_22

equation Reg18_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast) lunes martes miercoles jueves
viernes sku{!i}(-1)
forecast sku{!i}_23

equation Reg19_sku{!i}.ls sku{!i} c corr corr^2 corr^3 @expand(mes, @droplast) sku{!i}(-1)
forecast sku{!i}_24

equation Reg20_sku{!i}.ls sku{!i} c corr corr^2 corr^3 lunes martes miercoles jueves viernes sku{!i}(-1)
forecast sku{!i}_25

equation Reg21_sku{!i}.ls sku{!i} c corr corr^2 corr^3  @expand(mes, @droplast) lunes martes miercoles
jueves viernes sku{!i}(-1)
forecast sku{!i}_26

equation Reg22_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast) sku{!i}(-1) sku{!i}(-2)
forecast sku{!i}_27
```

```
equation Reg23_sku{!i}.ls sku{!i} c corr lunes martes miercoles jueves viernes sku{!i}(-1) sku{!i}(-2)
forecast sku{!i}_28

equation Reg24_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast)  lunes martes miercoles jueves
viernes sku{!i}(-1) sku{!i}(-2)
forecast sku{!i}_29

equation Reg25_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast)  sku{!i}(-1)  sku{!i}(-2)
forecast sku{!i}_30

equation Reg26_sku{!i}.ls sku{!i} c corr corr^2 lunes martes miercoles jueves viernes sku{!i}(-1)  sku{!i}(-
2)
forecast sku{!i}_31

equation Reg27_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast) lunes martes miercoles jueves
viernes sku{!i}(-1) sku{!i}(-2)
forecast sku{!i}_32

equation Reg28_sku{!i}.ls sku{!i} c corr corr^2 corr^3 @expand(mes, @droplast) sku{!i}(-1) sku{!i}(-2)
forecast sku{!i}_33

equation Reg29_sku{!i}.ls sku{!i} c corr corr^2 corr^3 lunes martes miercoles jueves viernes sku{!i}(-1)
sku{!i}(-2)
forecast sku{!i}_34

equation Reg30_sku{!i}.ls sku{!i} c corr corr^2 corr^3  @expand(mes, @droplast) lunes martes miercoles
jueves viernes sku{!i}(-1) sku{!i}(-2)
forecast sku{!i}_35

equation Reg31_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast) sku{!i}(-1) sku{!i}(-2) sku{!i}(-3)
forecast sku{!i}_36

equation Reg32_sku{!i}.ls sku{!i} c corr lunes martes miercoles jueves viernes sku{!i}(-1) sku{!i}(-2)
sku{!i}(-3)
forecast sku{!i}_37

equation Reg33_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast)  lunes martes miercoles jueves
viernes sku{!i}(-1) sku{!i}(-2) sku{!i}(-3)
forecast sku{!i}_38

equation Reg34_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast)  sku{!i}(-1)  sku{!i}(-2) sku{!i}(-
3)
forecast sku{!i}_39

equation Reg35_sku{!i}.ls sku{!i} c corr corr^2 lunes martes miercoles jueves viernes sku{!i}(-1)  sku{!i}(-
2) sku{!i}(-3)
forecast sku{!i}_40

equation Reg36_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast) lunes martes miercoles jueves
viernes sku{!i}(-1) sku{!i}(-2) sku{!i}(-3)
forecast sku{!i}_41

equation Reg37_sku{!i}.ls sku{!i} c corr corr^2 corr^3 @expand(mes, @droplast) sku{!i}(-1) sku{!i}(-2)
sku{!i}(-3)
forecast sku{!i}_42

equation Reg38_sku{!i}.ls sku{!i} c corr corr^2 corr^3 lunes martes miercoles jueves viernes sku{!i}(-1)
sku{!i}(-2) sku{!i}(-3)
forecast sku{!i}_43

equation Reg39_sku{!i}.ls sku{!i} c corr corr^2 corr^3  @expand(mes, @droplast) lunes martes miercoles
jueves viernes sku{!i}(-1) sku{!i}(-2) sku{!i}(-3)
```

```
        forecast sku{!i}_44

        equation Reg40_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast) SE4_sku{!i}
        forecast sku{!i}_45

        equation Reg41_sku{!i}.ls sku{!i} c corr lunes martes miercoles jueves viernes SE4_sku{!i}
        forecast sku{!i}_46

        equation Reg42_sku{!i}.ls sku{!i} c corr @expand(mes, @droplast)  lunes martes miercoles jueves
viernes SE4_sku{!i}
        forecast sku{!i}_47

        equation Reg43_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast) SE4_sku{!i}
        forecast sku{!i}_48

        equation Reg44_sku{!i}.ls sku{!i} c corr corr^2 lunes martes miercoles jueves viernes SE4_sku{!i}
        forecast sku{!i}_49

        equation Reg45_sku{!i}.ls sku{!i} c corr corr^2 @expand(mes, @droplast) lunes martes miercoles jueves
viernes SE4_sku{!i}
        forecast sku{!i}_50

        equation Reg46_sku{!i}.ls sku{!i} c corr corr^2 corr^3 @expand(mes, @droplast)  SE4_sku{!i}
        forecast sku{!i}_51

        equation Reg47_sku{!i}.ls sku{!i} c corr corr^2 corr^3 lunes martes miercoles jueves viernes SE4_sku{!i}
        forecast sku{!i}_52

        equation Reg48_sku{!i}.ls sku{!i} c corr corr^2 corr^3  @expand(mes, @droplast) lunes martes miercoles
jueves viernes SE4_sku{!i}
        forecast sku{!i}_53


        _Modelo{!k}.add sku{!i}_{!k}

    next
next

'-----------------------------------------------------------------------------------------------------------------------------------------------------------
--
for !m=1 to 53
    matrix(97, 84) _matrix!m
    matrix(97, 84) _diaspiso
    matrix(97,84) _error!m
    matrix(97, 84) _ep!m
    matrix(97, 84) _x!m
    matrix(97, 84) _z!m
    matrix(1, 84)  _w!m
    matrix(53, 84) _evaluador
    for !c =1 to 84
        for !r =1 to 97
            _matrix!m(!r, !c) = _stocksku (!c) (!r) / _modelo!m (!c) (!r + 911)
            _diaspiso(!r, !c)  = _dias_piso(!r, !c)
        next
    next
    _error!m = _diaspiso - _matrix!m
    matrix _t!m= @transpose(@cmean(_error!m))
    for !c=1 to 84
        for !r=1 to 97
            _ep!m(!r,!c) = _t!m(1, !c)
            _x!m =  _error!m - _ep!m
            matrix _z!m(!r, !c) = _x!m(!r, !c) * _x!m(!r, !c)
```

```
            matrix _w!m = @transpose(@csum(_z!m)/96)
        next
        _evaluador(!m, !c) = _w!m(1, !c)
    next
next

matrix __best_model = @cimin(_evaluador)

matrix __resultados = @sort(@uniquevals(__best_model),"a")
matrix(__resultados.@rows, 2) __resultados

for !i = 1 to __resultados.@rows
   !iden = 0
   for !c = 1 to 84
      if __best_model(!c, 1) = __resultados(!i, 1) then
         !iden = 1 + !iden
      endif
   next
   __resultados(!i, 2) = !iden
next
```