

Estimación del nivel de deterioro en la madera de las haciendas de una compañía balsera en el Ecuador, mediante el diseño de Data Marts y modelamiento de Árboles de Decisiones.

Elizabeth Ponce Santana, Wendy Rodríguez Santiana, Fabricio Echeverría Briones
Instituto de Ciencias Matemáticas
Escuela Superior Politécnica del Litoral
Campus Gustavo Galindo Km 30.5 Vía Perimetral, Guayaquil, Ecuador
tponce@espol.edu.ec, wbrodrig@espol.edu.ec, pechever@espol.edu.ec

RESUMEN

En este estudio se analizaron los procedimientos utilizados en una empresa del Ecuador exportadora de madera de balsa. Los cuales van desde plantaciones (inicio del proceso), hasta la elaboración en las fábricas del producto final (bloques formados por piezas de balsa).

Luego de mejorar algunos de los procedimientos en la empresa, se estimó el nivel de deterioro de la madera en las haciendas ó plantaciones; con el objetivo de reducirlo para evitar pérdidas futuras. Se utilizó para este fin la metodología de Data Marts y Árboles de decisión.

Palabras claves: *Deterioro, satélites, haciendas, secadora, pies rechazados o muestreados.*

ABSTRACT

This study focused in a balsa wood exporter company in Ecuador which has procedures that go from the plantations (start of process) until the making of the final product (block formed by pieces of balsa). After having improved some of the procedures of the long process we estimated the balsa wood's level of deterioration in the plantations with the purpose of reducing it to avoid greater future lost.

1. Introducción

El fundamento de la presente investigación prevalece en establecer los factores o causas del deterioro de la madera, materia prima de una empresa balsera en el Ecuador cuya actividad es la exportación de la madera de balsa previamente procesada hacia diferentes países alrededor del mundo.

Las condiciones iniciales que condujeron a una pérdida en la calidad en el producto terminado, partieron de los informes desfavorables de los inspectores de calidad de los importadores o consumidores finales de la balsa procesada hecho que se dió a finales del año 2009, iniciándose controles exhaustivos en la calidad de la materia prima y el tratamiento de la balsa en las haciendas y en los centros de acopio, donde se considera radica el problema. A fin de confirmar esta hipótesis se utilizarán técnicas de Minerías de Datos como Árboles de decisión, partiendo de la elaboración de un Modelo Estrella (Data Marts) específico para

el Dpto. de Satélites, área del negocio inmerso en este estudio. El objetivo es reducir el deterioro encontrado en la balsa, y que éste no exista en el producto final.

2. Descripción del proceso

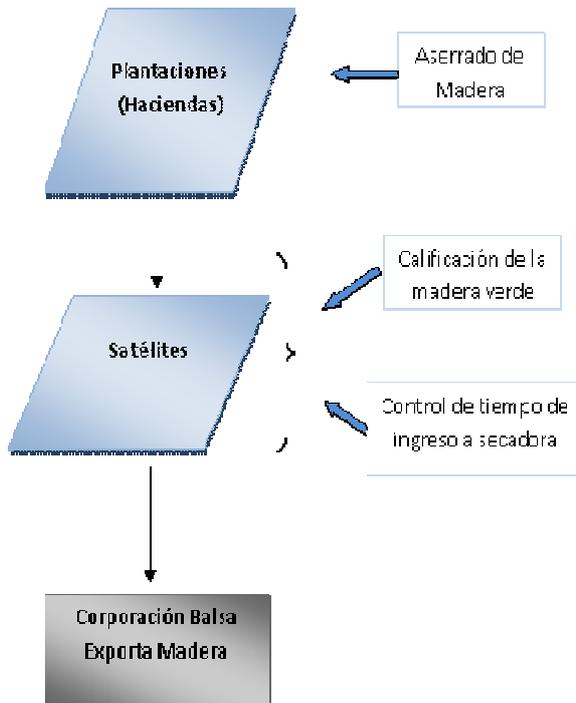
El proceso que sigue la materia prima (balsa), está directamente relacionado a exigencias preestablecidas por el comprador (importador), es así que la forma, tamaño y nivel de secado es totalmente conducido por el destinatario final del producto y por ende las condiciones del producto terminado esta basado en parámetros muy estrictos de calidad. En términos generales los pasos que debe seguir la madera (balsa) desde que es cortada hasta ser embarcada en los contenedores son los siguientes:

- Cuando están listos para la cosecha pasan a ser cortados, proceso que es denominado “*tumba*” del árbol.

- A cada árbol tumbado se le retiran las hojas quedando únicamente el tronco, el cual es llamado “troza”.
- Cada troza es pintada en sus extremos de un color de acuerdo al día de la semana.
- Cada árbol tumbado, pasa a la siguiente fase (proceso de *aserrado*).
- Dicho proceso consiste en cortar el árbol en pequeñas piezas que pueden tener tres medidas: 4, 5 ó 6 pies de largo; así mismo, pueden ser de diferente espesor o ancho.
- Un conjunto de 70 piezas aproximadamente es reunido para formar un “bulto”.
- Estos bultos son distribuidos a las fábricas de madera denominadas “satélites”, que prestan servicios de procesamiento de madera de balsa para la empresa en estudio.

2.1 Flujo del Proceso

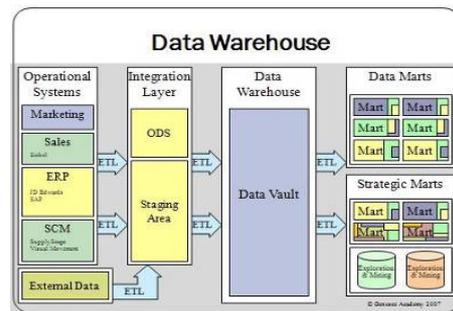
Gráfico 1.1
Flujo del proceso



3. Data Warehouse

Un Data Warehouse es una solución que permite centralizar en un solo punto, toda la información definida por la compañía como relevante para la gestión de su negocio y la toma de decisiones.

Figura 2.1

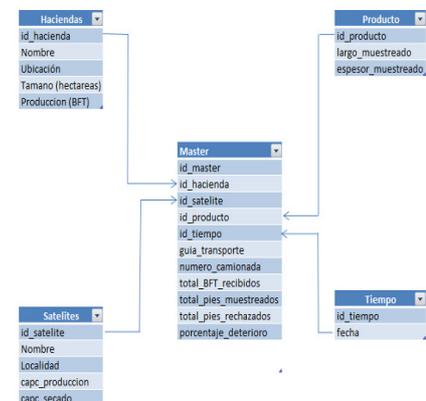


3.1 DataMart

Un DataMart es una solución, que compartiendo tecnología con el Data Warehouse (pero con contenidos específicos, volumen de datos más limitado y un alcance histórico menor), permite dar soporte a una empresa pequeña, o un departamento o área de negocio de una empresa grande.

3.2 Modelo Estrella

Figura 2.3
Data Mart (Modelo Estrella)



Este es el Data Mart de donde se obtendrá la información para aplicar la técnica de

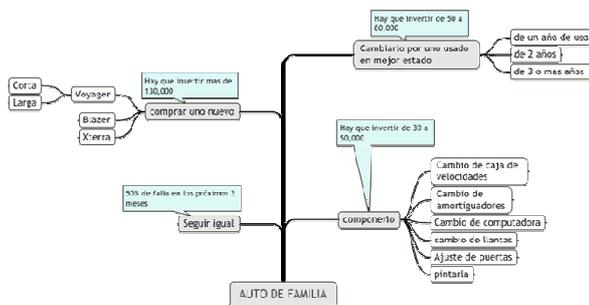
minería de datos, árboles de decisión, para la estimación del deterioro en las haciendas de la empresa balsaera.

4. Árbol de decisiones

Se define un árbol de decisión como una estructura en forma de árbol en la que las ramas representan conjuntos de decisiones. Estas decisiones generan sucesivas reglas para la clasificación de un conjunto de datos en subgrupos disjuntos y exhaustivos. Las ramificaciones se realizan de forma recursiva hasta que se cumplen ciertos criterios de parada.

El objetivo de estos métodos es obtener individuos más homogéneos con respecto a la variable que se desea discriminar dentro de cada subgrupo y heterogéneos entre los subgrupos. Para la construcción del árbol se requiere información de variables explicativas a partir de las cuales se va a realizar la discriminación de la población en subgrupos.

Figura 3.1
Ejemplo de árbol de decisiones



4.1 Prueba de Bondad de Ajuste

Existen dos procedimientos para realizar pruebas de bondad de ajuste que son los más conocidos. El primero se basa en una técnica gráfica muy útil llamada **gráfica de probabilidad** y el segundo procedimiento se basa en la **distribución Chi-cuadrada**. En este estudio se aplicará la distribución Chi-cuadrada la que describimos a continuación.

4.2 Prueba de bondad de ajuste de la Chi-cuadrada

El procedimiento de prueba de la Chi-cuadrada es un método analítico, requiere una muestra aleatoria de tamaño n de la variable aleatoria X . Estas n observaciones se arreglan en histogramas de frecuencias, teniendo k intervalos de clase (donde $k = \sqrt{n}$). Sea n_i la frecuencia observada en el i -ésimo intervalo de clase. De la distribución de probabilidad hipotética, se calcula la frecuencia esperada en el i -ésimo intervalo de clase, identificada como p_i .

H_0 : La distribución es $F_0(X)$
(Distribución Teórica dada)

H_a : La distribución no puede ser $F_0(X)$

$$EP: \chi^2 = \sum_{i=1}^k \frac{(n_i - np_i)^2}{np_i}$$

$$RR: \chi^2 > \chi_{\alpha, k-r-1}^2$$

4.3 Grados de Libertad

Es un estimador del número de categorías independientes en una prueba particular o experimento estadístico. Se encuentran mediante la fórmula $k - r$, donde k =número de grupos, cuando se realizan operaciones con grupos y r es el número de sujetos o grupos estadísticamente dependientes.

5 Conclusiones

- En el árbol elaborado con el algoritmo CHAID, con el total de pies rechazados y satélites como variables predictoras, muestra que la hacienda con mayor nivel de deterioro es la denominada MS09, siguiéndole la hacienda PB08 cuando el total de pies rechazados son mayores a 33,5 y menores o iguales a 35,5.
- Elaborado este árbol, se pudo constatar que independientemente de la forma como sea calificada la madera en las satélites, la hacienda

con mayor nivel de deterioro sigue siendo la MS09. Sin embargo se utilizó en los siguientes árboles otros algoritmos para confirmar los resultados antes mencionados o verificar si podían diferenciarse de alguna manera.

- Finalmente se puede concluir que para mejorar los índices de deterioro actuales, es importante realizar un análisis de las dos plantaciones mencionadas en el numeral 4, a fin de establecer las razones por las cuales estas dos haciendas están teniendo madera defectuosa, y poder determinar si de alguna manera la

forma de aserrar la balsa, esta influyendo en los resultados encontrados en las haciendas y en este estudio.

**Msc. Pedro Fabricio Echeverría
Briones**