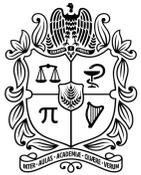


MODELOS GAMLSS PARA ANALIZAR EL GRADO DE SECADO

DE ACETATO DE CALCIO DIHIDRATADO



UNIVERSIDAD
NACIONAL
DE COLOMBIA

Brahian Cano Urrego
Edwin Caicedo Chamorro
Santiago Toro Zuluaga

Universidad Nacional de Colombia sede Medellín

1. Resumen

Se empleo metodología GAMLSS para reanalizar el problema propuesto por los autores Triana et al (2009); planteando una opción alternativa que explica el grado de secado (GS) del acetato de calcio dihidratado mediante un secador de tipo transportador. Encontrando con esta metodología una mejor distribución y adecuación para la variable respuesta, llegando así, a un modelo con distribución Weibull con parametrización 3 (WEI3) que cumple con las cualidades de interpolación y modelamiento de los datos, los resultados fueron obtenidos empleando el programa estadístico R, los cuales arrojaron una correlación de 0.95 y un AIC 42.93.

2. Modelos GAMLSS

Los modelos GAMLSS (Generalized Additive Model for Scale and Shape) propuestos por Rigby y Stasinopoulos (2005) son una opción que brinda al investigador la posibilidad de modelar la variable respuesta y sus parámetros en función de las covariables.

En la mayoría de los estudios para modelar una variable no siempre se cumple el supuesto de normalidad o varianza constante; lo cual brinda una ventaja significativa a la metodología GAMLSS, dado su flexibilidad, ya que ninguno de estos es un requerimiento indispensable para el modelamiento y este tipo de método permite seleccionar entre más de 80 distribuciones posibles. También permite que la distribución de la variable de interés sea continua, discreta o mixta, estos modelos se pueden aplicar fácilmente por medio del paquete gamlss disponible en R core Team (2017).

3. Introducción

Con base al artículo Triana et al. (2009) el cual trata de modelar con una distribución normal el conjunto de datos que explica el grado de secado (GS) del acetato de calcio dihidratado mediante un secador de tipo transportador en función de las covariables: velocidad rotacional del tornillo (N) en dos niveles 3 y 7 rpm, la temperatura (T) del sólido alimentado en tres niveles 45, 53 y 61 °C y el flujo de alimentación (C) en tres niveles 6, 8 y 10 L/min.

El secador de tornillo transportador (SCD) es un secador continuo de calentamiento indirecto, debido a la versatilidad que tiene este tipo de secador recientemente se han realizado estudios relacionados con el desempeño del secador y el comportamiento de las variables hidrodinámicas y térmicas. Nuestro objetivo principal es proponer modelos alternativos que podrían explicar de mejor manera la variable respuesta, utilizando metodología GAMLSS.

4. Análisis descriptivo

GS	Te	N	C
14.13	61	3	6
12.08	45	3	8
7.49	61	3	10
5.94	61	3	8
1.17	53	3	10
0.25	45	3	6
7.23	45	3	10
5.80	53	3	6
0.59	53	3	8
3.61	53	7	6
9.58	61	7	10
0.22	53	7	8
1.31	53	7	10
2.09	45	7	6
1.39	45	7	10
3.84	45	7	8
5.35	61	7	8
5.75	61	7	6

Tabla 1: Datos del artículo

Se presenta la base de datos del experimento del artículo, en la cual se aprecian los niveles de cada variable y además las observaciones que se consideran como outliers por los autores del artículo, se resaltan en negrilla.

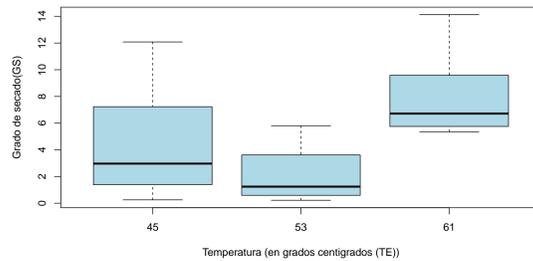


Figure 1: Grado de secado discriminado por temperatura

Se puede apreciar una diferencia con respecto a la media de grado de secado (GS) entre los niveles de temperatura de 53 y 61 siendo este último el mayor. Lo que nos da un indicio hacia qué nivel de temperatura podría optimizar el grado de secado.

5. Metodología

En primera instancia se buscaron las mejores distribuciones paramétricas que pudieran modelar la variable respuesta, en nuestro caso el grado de secado (GS); usando la función fitDist() del paquete GAMLSS, encontramos las mejores cuatro distribuciones para modelar el grado de secado (GS); de esto pudimos obtener las siguiente gráfica:

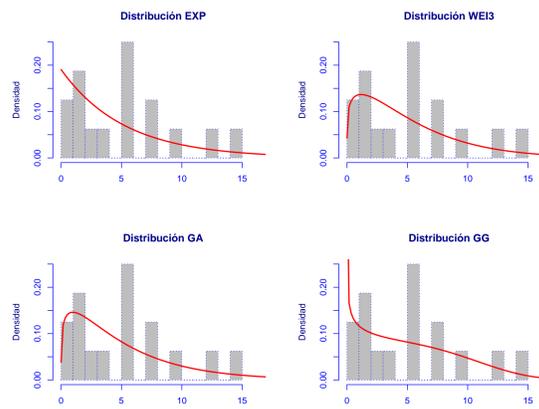


Figure 2: Posibles distribuciones del grado de secado

Dentro del desarrollo del trabajo, se propusieron diferentes tipos de modelos, entre ellos lm, no paramétricos, transformaciones y GAMLSS. Para la selección de los mejores modelos se utilizaron los criterios: Normalidad de los errores, AIC, Correlación con la variable respuesta (capacidad de interpolar). En busca del mejor modelo posible para cada distribución, se realizó un proceso de selección de variables con la función StepAICAll.A() para los mejores modelos considerados, arrojando como resultado final los siguientes modelos:

Modelos	Distribución	V.Respuesta	Correlación	AIC
modgs	Normal	GS	0.95	79.02
mod1	Normal	\sqrt{GS}	0.91	28.39
mod2	Exponencial	GS	0.76	86.45
mod3	Gamma	GS	0.76	51.58
mod4	Gamma G	GS	0.45	33.07
mod5	Weibull 3	GS	0.95	42.93

Tabla 2: Modelos propuestos

6. Resultados

Después de analizar los modelos con los criterios AIC, correlación y normalidad de los errores, descartamos los modelos que tienen violaciones graves de los supuestos normalidad, en segunda instancia filtramos por los de menor AIC y por ultimo elegimos el modelo con la mejor capacidad de explicar la variable respuesta en términos de correlación.

De este proceso se obtuvo como mejor modelo el WEI3, con $E(GS) = \mu$ y $\widehat{Var}(GS) = \hat{\mu}^2 \frac{\Gamma(\frac{1}{\sigma}+1)}{\Gamma(\frac{1}{\sigma}+1)^2-1}$ las características del modelo obtenido se muestran a continuación:

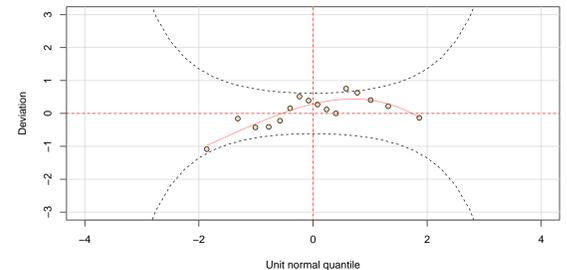


Figure 3: Wormplot del modelo WEI3

No podemos observar violaciones graves del supuesto de normalidad de los errores.

Este modelo tiene un AIC= 42.93, una Correlación con la variable respuesta de 0.954; las siguientes ecuaciones describen el modelo:

$$GS \sim WEI3(\mu, \sigma)$$

con parámetros

$$\hat{E}(GS) = \exp(109.58 - 2.05N - 3.30T - 4C + 0.03T^2 + 0.14C^2 + 0.02N \cdot T + 0.07N \cdot C + 0.02T \cdot C)$$

$$\widehat{Var}(GS) = \exp(152.47 + 0.05T^2 - 5.07T - 0.92N)$$

Con el modelo resultante calculamos los valores para los cuales se optimiza el grado de secado, y además generamos una superficie respuesta, fijando la temperatura en 43°, la cual es la óptima, para poder obtener una idea visual del fenómeno.

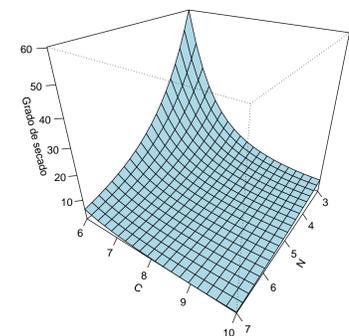


Figure 4: Superficie de respuesta de grado de secado(GS) a una temperatura de 43°

7. Conclusiones

Se obtiene como resultado un mejor modelo con respecto al propuesto en el artículo en términos de interpolación y adecuación del fenómeno evidenciado por la correlación y el AIC, con este resultado se evidencia la utilidad que tiene dar un uso adecuado de la metodología GAMLSS y no limitarse exclusivamente a modelar la variable respuesta con una distribución normal.

Se observó que el punto óptimo del grado de secado (GS) es 60.28, el cual se obtiene con unos valores de TE=45°, C=6, N=3, esto se contradice la creencia inicial de la parte descriptiva que nos indicaba que la temperatura que podría optimizar el grado de secado era de 61, esto puede deberse a que no se tomaba en cuenta la interacción entre las covariables o que solo tomaba en cuenta el termino lineal.

Referencias

- Triana et al (2009). Grado de secado y eficiencia térmica de un secador de tornillo transportador drying *DYNA*, **78(165)**, 196–206.
- Rigby, B.; Stasinopoulos, E (2005). Generalized additive models for location scale and shape. *Applied Statistics*, **54**, 507–554.
- R Core Team (2017). R: A Language and Environmental for Statistical Computing, R Foundation for Statistical Computing. Vienna, Austria. <http://www.R-project.org/>.