



TecnoLógicas

ISSN: 0123-7799

ISSN: 2256-5337

Instituto Tecnológico Metropolitano - ITM

Pinto-Altamiranda, Sania; Gómez R., Sara Manuela;
González, María Eugenia; Barrera-Causil, Carlos
Modelación estadística para analizar el rendimiento y contenido de carbono de biomasas agroindustriales
TecnoLógicas, vol. 26, núm. 57, e207, 2023, Mayo-Agosto
Instituto Tecnológico Metropolitano - ITM

DOI: <https://doi.org/10.22430/22565337.2677>

Disponible en: <https://www.redalyc.org/articulo.oa?id=344276660009>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](https://www.redalyc.org)



Sistema de Información Científica Redalyc
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso
abierto

Recibido: 23 marzo 2022
Aceptado: 15 agosto 2023
Disponible: 24 agosto 2023

©Instituto Tecnológico Metropolitano
Este trabajo está licenciado bajo
una Licencia Internacional
Creative Commons Atribución
(CC BY-NC-SA)



Modelación estadística para analizar el rendimiento y contenido de carbono de biomasas agroindustriales

Statistical Modeling to Analyze the Performance and Carbon Content of Agro-industrial Biomasses

 Sania Pinto-Altamiranda ¹;
 Sara Manuela Gómez R ²;
 María Eugenia González ³;
  Carlos Barrera-Causil ⁴

¹ Institución Universitaria Pascual Bravo, Medellín-Colombia,
saniapintoa@gmail.com

² Instituto Tecnológico Metropolitano, Medellín-Colombia,
saragomez266593@correo.itm.edu.co

³ Universidad de la Frontera, Temuco-Chile,
mariaeugenia.gonzalez@ufrontera.cl

⁴ Instituto Tecnológico Metropolitano, Medellín-Colombia,
carlosbarrera@itm.edu.co

Cómo citar / How to cite

S. Pinto-Altamiranda, S. M. Gómez, M. E. González, C. Barrera-Causil,
“Modelación estadística para analizar el rendimiento y contenido de
carbono de biomasas agroindustriales,” *Tecnológicas*, vol. 26, nro. 57,
e2677, 2023. <https://doi.org/10.22430/22565337.2677>

Resumen

En la agroindustria se genera una considerable cantidad de residuos, los cuales pueden ser tratados usando diversas tecnologías termoquímicas como la carbonización hidrotermal. El rendimiento y contenido de carbono de biomásas son dos de las características más comunes que se estudian dentro del proceso generado en estas tecnologías termoquímicas, y usualmente se aplican análisis químicos y técnicas estadísticas, tales como pruebas t-student, análisis de varianza o modelos de superficies de respuestas para optimizar estas respuestas o estimar el efecto que ciertos factores puedan tener sobre estas. A diferencia de las investigaciones abordadas en esta área de la química, este estudio tuvo como propósito introducir diferentes técnicas alternativas de la estadística para la modelación de este tipo de datos con el fin de proponer diferentes estrategias de análisis que permitan ampliar el conocimiento de los fenómenos estudiados en esta área. Para ello, se presentó la modelación estadística de dos bases de datos provenientes de bagazo de manzana y de arándanos que contienen un total de cuatro factores (tiempo, humedad, potencia, temperatura) y dos respuestas a analizar por separado (contenido de carbono y rendimiento del proceso). En este estudio se observa que el tiempo, la temperatura y la humedad tienen un efecto conjunto sobre el rendimiento del proceso y el contenido de carbono de la biomasa proveniente de la manzana. Se concluye que, técnicas como modelos lineales generalizados con respuesta beta y los modelos aditivos generalizados de posición, escala y forma, proporcionan un mayor conocimiento del fenómeno de interés y la capacidad de estimar el efecto de los factores estudiados sobre respuestas que naturalmente no poseen un comportamiento distribucional como el modelo normal.

Palabras clave

Carbonización hidrotermal, contenido de carbono, modelación estadística, rendimiento del hidrocarbón, residuos agroindustriales.

Abstract

In agroindustry, a significant amount of waste is generated, which can be treated using various thermochemical technologies such as hydrothermal carbonization. Biomass yield and carbon content are two of the most common characteristics studied within the processes generated by these thermochemical technologies, and chemical analyses and statistical techniques are usually employed. These techniques include t-student tests, analysis of variance, or response surface models to optimize or estimate the effects of certain factors. Unlike research conducted in this field of chemistry, this study aimed to introduce alternative statistical techniques for modeling such data, proposing diverse analysis strategies to enhance understanding of the studied phenomena. To achieve this, the statistical modeling of two datasets derived from apple pomace and blueberries was presented, encompassing four factors (time, humidity, power, temperature) and two separate responses (carbon content and process yield). This study reveals that time, temperature, and humidity collectively affect process yield and carbon content in apple biomass. It is concluded that techniques like generalized linear models with beta response and generalized additive models for location, scale, and shape provide a deeper understanding of the phenomenon of interest and the ability to estimate the effects of studied factors on responses that do not naturally follow a normal distribution.

Keywords

Hydrothermal carbonization, carbon content, statistical modeling, hydrocarbon yield, agro-industrial residues.

1. INTRODUCCIÓN

Actualmente en el mundo existe una problemática frente a la generación de residuos orgánicos, entre ellos los provenientes de la industria agrícola. Si estos residuos se liberan al medio ambiente sin el tratamiento de depuración adecuado, pueden causar efectos nocivos para la salud humana y el medio ambiente [1].

Los residuos agroindustriales de distintos sectores fomentan la producción de biomásas, de los cuales la gran mayoría se eliminan mediante vertimientos, rellenos sanitarios o quema [2]. La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO por sus siglas en inglés) estima que aproximadamente 1600 millones de toneladas de alimentos se desperdicia por año. Las frutas, verduras, tubérculos representan entre el 40 % y el 50 % de los alimentos desperdiciados [3].

Bajo este contexto, esto se convierte en una gran problemática para los gobiernos debido a las posibles afectaciones medioambientales y de salud pública. Este fenómeno ha llevado a la comunidad científica mundial a buscar una solución para la cual, actualmente, se han escalado procesos de transformación de biomásas agroindustriales para diferentes aplicaciones, tales como las biorrefinerías. En este proceso, es necesario utilizar técnicas de modelación estadística para la toma de decisiones, especialmente cuando se pretende explicar el comportamiento de variables derivadas de experimentos controlados mediante tecnologías tecnoquímicas, como la carbonización hidrotermal.

En la literatura científica, asociada a tecnologías termoquímicas, es posible encontrar una amplia variedad de estudios en los cuales se ha implementado una metodología específica para el análisis estadístico de datos. Cuando se desea optimizar el valor de una respuesta en función de un conjunto de factores, la metodología de superficie de respuesta es ampliamente utilizada para este fin. En [4] hacen uso de esta a través de un diseño central compuesto (CCD por sus siglas en inglés) con puntos axiales. Su interés es analizar los efectos entre el rendimiento y el contenido de carbono de la carbonización de ramas de manzano podadas para obtener biocarbón. El diseño experimental usado es multifactorial con dos respuestas y una de ellas fija. Otra técnica de superficie de respuesta muy usada es el diseño Box-Behnken, por ejemplo, [5] lo implementan en un problema con tres factores a dos niveles cada uno con el fin de optimizar compuestos fenólicos mediante pirolisis de residuos de espuma floral. En la actualidad, el hidrocarbón es considerado como absorbente carbonoso renovable emergente y en conjunto con los diseños experimentales son de interés [6]. Por otro lado, los efectos de los parámetros de activación de un hidrocarbón de bagazo de caña de azúcar, mediante un medio álcali frío fueron estudiados mediante un CCD en complemento con un análisis de varianza (ANOVA) para optimizar el rendimiento y la capacidad de adsorción [7].

Por su parte, [8] estudian el rendimiento del bioaceite y biocarbón de la cáscara de lentejas por medio de un CCD con tres factores, ocho puntos factoriales, seis puntos axiales y puntos centrales y un ANOVA. [9] usan el CCD para evaluar tres factores con dos variables de respuesta: rendimiento y propiedades del biocarbón obtenido mediante una pirolisis de la cascara de arroz y el alquitrán de residuos sólidos urbanos. Asimismo, [10] estudian la valorización del racimo de fruta vacío y la cáscara de palmiste mediante una pirolisis lenta optimizando el biocarbón obtenido de cada biomasa. Aquí, la metodología estadística aplicada es superficie de respuestas con un CCD y un ANOVA para determinar las interacciones entre los factores estudiados. Aplicando el diseño Box-Behnken, diseño de superficie de respuestas y el ANOVA, [11] implementaron un modelo de diseño de experimentos para optimizar el rendimiento del bioaceite proveniente de la cáscara de nuez de macadamia.

Es relevante señalar que existe un patrón recurrente en la literatura estudiada, donde ninguno de los análisis estadísticos implementados tiene en cuenta que la variable de

respuesta no sigue una distribución normal. En el modelo ANOVA, el supuesto de normalidad de los errores debe ser chequeado para garantizar su usabilidad, y el cumplimiento de este supuesto, lleva a que la respuesta también se comporte como una distribución normal. Cuando se analizan las variables rendimiento y contenido de carbono, similar a los anteriores estudios, se observa que estas son medidas en porcentaje, por lo que su rango de valores natural está entre cero (0) y 100, es así como asumir un modelo distribucional como una normal puede no ser adecuado.

En este artículo se propone una metodología para el análisis de datos estadísticos para observaciones obtenidas a través de procesos de carbonización hidrotermal asistida por microondas. Esta metodología se enfoca en respuestas no normales o en modelos mucho más generales que los modelos de análisis de varianza. La metodología se ilustra usando dos bases de datos reales correspondiente al bagazo de manzana y de arándanos en dos diferentes estudios, que son analizadas usando el *software* R (los datos y el código de R usado se encuentran disponibles en [Material suplementario](#)). El interés en las ilustraciones es modelar el contenido de carbono y el rendimiento de este como respuesta, teniendo en cuenta cuatro factores: humedad, tiempo, potencia y temperatura a la que son expuestas las biomásas en el microondas.

El contenido del artículo se describe como sigue: inicialmente se exponen las definiciones correspondientes a las técnicas estadísticas aplicadas en esta propuesta metodológica; la sección 2 presenta la metodología; la sección 3, de resultados, expone como eje central la propuesta metodológica para la modelación estadística del rendimiento y contenido de carbono de biomásas agroindustriales, e ilustra dicha propuesta a través del análisis de dos bases de datos reales; la sección 4 presenta las conclusiones; las secciones 5 y 6 presentan agradecimientos y referencias, respectivamente; por último, se presentan anexos.

2. ASPECTOS METODOLÓGICOS

El desarrollo metodológico tiene como punto de partida una revisión literaria con la cual se identifican los tipos de estudios y las técnicas de análisis de datos implementadas para investigar el rendimiento y contenido de carbono de biomásas. Se propone, además, una metodología experimental alternativa para la modelación y análisis de datos provenientes de biomásas agroindustriales. Si bien aquí la propuesta metodológica muestra dos componentes, el análisis químico y el estadístico, se hace mayor énfasis en el análisis estadístico y se presentan dos ilustraciones con datos reales, los cuales son descritos a continuación, al igual que la metodología experimental con la cual se recolectan los datos.

La primera ilustración consiste en el análisis de datos provenientes de bagazo de arándanos, donde se observan sus características fisicoquímicas y se explica, a través de modelos estadísticos, el comportamiento del rendimiento del hidrocarbón y del contenido de carbono encontrado. De igual manera, la segunda ilustración se lleva a cabo con datos provenientes del bagazo de manzana, donde se presenta el análisis elemental y proximal de la muestra, y diferentes modelos estadísticos para el rendimiento del hidrocarbón como respuesta. En ambas ilustraciones se hace chequeo de normalidad de las respuestas y se evidencia que los errores de los modelos no siguen el comportamiento de una distribución normal, algo que, como se menciona en la introducción de este artículo, es común que ocurra para este tipo de estudios, pero que los investigadores omiten.

La metodología experimental implementada en ambas ilustraciones corresponde a un diseño experimental de dos niveles tipo factorial 2^4 con una sola réplica, cuya matriz de diseño se obtiene a través del software Design-Expert 11. Para el caso del bagazo de manzana

se usan cuatro puntos centrales, mientras que para los datos de bagazo de arándanos se toman tres puntos centrales. Los factores estudiados son tiempo (30, 60 y 90 minutos), temperatura (160 °C, 180 °C y 200 °C), contenido de humedad (70 %, 80 % y 90 %) y potencia del microondas (50, 150 y 250 vatios). Es importante mencionar que para disminuir el margen de error en las muestras de bagazo de manzana se decide trabajar con cuatro puntos centrales debido a la dificultad para que esta materia prima se mantuviera húmeda; para el bagazo de arándanos no se presenta ese inconveniente, por ello se trabaja con solo tres puntos centrales (ver Tablas A1 y A2 en anexos).

Para el tratamiento químico se realiza el acondicionamiento de las muestras, reduciéndolas de tamaño a través de trituración con minipimer durante cinco minutos, con el propósito de homogeneizar las materias primas antes de la caracterización. Posteriormente, como parte de la caracterización fisicoquímica, se realiza un análisis elemental usando un analizador elemental CHNS, Eurovector EA 3000 y un análisis proximal o termogravimétrico para ambas biomásas. En este estudio no se mide el tamaño de la partícula debido a que las muestras, luego de ser trituradas, no eran sólidas, sino líquidas; esta característica favorece el proceso de carbonización hidrotermal.

La muestra de bagazo de arándanos se obtuvo en el campo experimental Maquehue de la Facultad de Ciencias Agropecuarias y Forestales de la Universidad de la Frontera, predio ubicado a diecisiete kilómetros al sur de Temuco, en la Comuna de Freire, en la Región de Araucanía (Chile) y la muestra de bagazo de manzana se obtuvo del Fundo Loncovaca, ubicado en el km 4 Camino Botrolhue, Temuco, en la Región de Araucanía (Chile), que cuenta con veintiocho hectáreas y una producción de 5 ton/año. En el Fundo Loncovaca se cosechan cinco variedades de manzana, las cuales son: Elstar, Royal Gala, Braeburn, Red Chief y Jonagold.

Las dos ilustraciones se centran en mostrar dos modelos estadísticos: el modelo lineal generalizado mixto tipo beta y el modelo GAMLSS, los cuales se implementan con el software estadístico R. Note que han sido propuestos para datos que no siguen una distribución normal, por lo que se realiza el chequeo del supuesto de normalidad para ambos casos, con la prueba Shapiro Wilk.

Con el fin de conocer aquellos factores que tienen influencia sobre las respuestas, los modelos explorados se evalúan con y sin interacciones. Dada la discusión que se ha generado en años recientes en torno al nivel de significancia estadístico para las pruebas (ver [12]), en este artículo se recomienda y adopta la propuesta de concluir con evidencia estadísticamente significativa cuando p-valor sea inferior a 0.005. Además, cuando p-valor se encuentre entre 0.005 y 0.05, se concluye en contra de H_0 con una evidencia prometedora [13]. La selección de los modelos se hace a través de los criterios de información de Akaike (AIC por sus siglas en inglés) y de Bayes (BIC por sus siglas en inglés), donde el AIC está basado en la función de verosimilitud e incluye una penalización que aumenta con el número de parámetros estimados en el modelo y, además, premia los modelos que dan un buen ajuste en términos de verosimilitud, pero, a la vez, tienen pocos parámetros. Por su parte, el BIC comprende una modificación al AIC que tiene en cuenta también el número de datos utilizados en el ajuste del modelo [14].

En el estudio, los datos generados por medio de tecnologías tecnoquímicas, en particular carbonización hidrotermal asistida por microondas, y el rendimiento y contenido de carbono de biomásas, son dos de las características más comunes que se estudian dentro del proceso generado en estas tecnologías tecnoquímicas, y usualmente se aplican análisis químicos y técnicas estadísticas, tales como pruebas t-student, análisis de varianza o modelos de superficies de respuestas para optimizar estas respuestas o estimar el efecto que ciertos factores puedan tener sobre las mismas. Los modelos de regresión lineal múltiple, lineales

generalizados y los modelos GAMLSS (modelos aditivos generalizados para posición, escala y forma) son alternativas adecuadas cuando las respuestas no siguen una distribución de probabilidad como la distribución normal, la cual es la más común en la estadística y la que se asume usualmente en estudios para esta área de la química.

Cuando se analizan el rendimiento y el contenido de carbono como resultados en un estudio químico, estos valores se expresan en forma de porcentajes que varían de 0 % a 100 %. Debido a su naturaleza, no siguen una distribución normal, por lo que asumir la normalidad en el análisis podría ser incorrecto. Por lo tanto, basados en datos obtenidos bajo metodología experimental, tales como son estos tipos de estudio, se puede implementar modelación estadística para respuestas no normales como alternativa para los análisis con el fin de estimar el efecto de los factores sobre las respuestas y poder así optimizar la misma.

2.1 Modelo de regresión lineal múltiple

Un modelo de regresión lineal múltiple es una función lineal que intenta modelar una respuesta Y a través de una combinación lineal de predictores o covariables x [15]. En general se tienen n observaciones de la variable Y . $y_1, y_2, \dots, y_n$ son observaciones independientes de Y , y siguen un modelo lineal si se puede expresar como (1) [16].

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{p-1} x_{i,p-1} + \varepsilon_i, i = 1, 2, 3, \dots, n \quad (1)$$

La forma matricial del modelo se puede escribir como se muestra en (2):

$$Y = X\beta + \epsilon \quad (2)$$

Este modelo se basa en supuestos, los cuales incluyen la linealidad de la relación entre las variables, la independencia de los errores, la homocedasticidad de la varianza de los errores, la ausencia de multicolinealidad entre las variables independientes y la normalidad de los errores. Estos supuestos garantizan la validez y confiabilidad del modelo, por lo que es necesario chequear el cumplimiento de estos.

2.2 Modelos lineales generalizados

Los modelos lineales generalizados (GML por sus siglas en inglés) son una alternativa de transformaciones de la respuesta, justificada por una falta de linealidad y homocedasticidad de la varianza. En todos los GML, las propiedades básicas de los estimadores son sensibles a la distribución asumida [17]. Los GML pueden incluir como distribuciones Poisson, binomial, gamma, beta.

En los modelos lineales generalizados se introduce una función invertible g , denominada función enlace (3) [18].

$$g(E(Y|X)) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (3)$$

2.2.1 Modelos generalizados mixtos con respuesta beta

La regresión de tipo beta fue introducida por Ferrari y Cribari mediante una reparametrización en función de la media y la varianza. El modelo propuesto es útil para situaciones en las que la variable de interés es continua y restringida al intervalo de 0 % a

100 % (0, 1) y se relaciona con otras variables a través de una estructura de regresión. Los parámetros del modelo son interpretables en términos de la media de la respuesta (4) [19].

$$g(E(Y_i|u)) = X_1\beta + Z_iu \quad (4)$$

Siendo $Y_i|u$, $X_1\beta$, Z_iu tal y como se definió en el modelo lineal mixto. La estimación se realiza por máxima verosimilitud, que es un método para obtener estimaciones de parámetros desconocidos mediante la optimización de una función de verosimilitud [20].

2.3 Modelos aditivos generalizados para posición, escala y forma (GAMLSS)

Los modelos GAMLSS son modelos de regresión simple semiparamétricos. Paramétricos porque requieren que la variable respuesta siga una distribución paramétrica (beta, gamma, entre otros), y semi, porque los parámetros de la distribución son modelados de forma independiente [21].

Los términos empleados dentro del marco de los GAMLSS para referirse a los parámetros de localización, escala y forma son μ, σ, ν, τ (5).

$$Y \sim D(\mu, \sigma, \nu, \tau) \quad (5)$$

Los modelos GAMLSS y los modelos lineales generalizados aditivos (GAM), asumen que la variable respuesta sigue distribuciones de la familia exponencial, donde la varianza se calcula mediante una función $\nu(\mu)$ y una constante de dispersión, y la media es modelada en función de predictores, por lo que la media calculada en los modelos es buena, sin embargo, la incertidumbre no es acorde, ya que la varianza se modela a través de su relación con la media [22].

3. RESULTADOS Y DISCUSIÓN

3.1 Propuesta metodológica

El tratamiento de las biomásas sugiere varias tecnologías termoquímicas que se pueden usar, como pirólisis y gasificación, entre otras; el uso de estos procedimientos químicos está sujeto a la investigación realizada y al objetivo de estudio. Inicialmente, definir este aspecto es importante, ya que hace parte del desarrollo metodológico. Posteriormente, para realizar un análisis de una biomasa se realiza acondicionamiento de muestras y caracterización fisicoquímica según sea el residuo para tratar. Antes de realizar las medidas en un laboratorio, es recomendable plantear el diseño experimental de acuerdo con los niveles y factores definidos por los investigadores, luego se construye la matriz de diseño, la cual, por medio de una aleatorización, proporciona el orden en el que se deben tomar las medidas en el laboratorio.

La descripción de la propuesta metodológica para el análisis de datos experimentales provenientes de la aplicación de tecnologías termoquímicas planteada en este estudio es descrita en la Figura 1. Aquí, se considera como esencial, luego de planear y diseñar el experimento, proceder a la selección de la tecnología termoquímica adecuada para la experimentación, esto teniendo en cuenta los equipos con los que se cuenta para realizar el estudio.

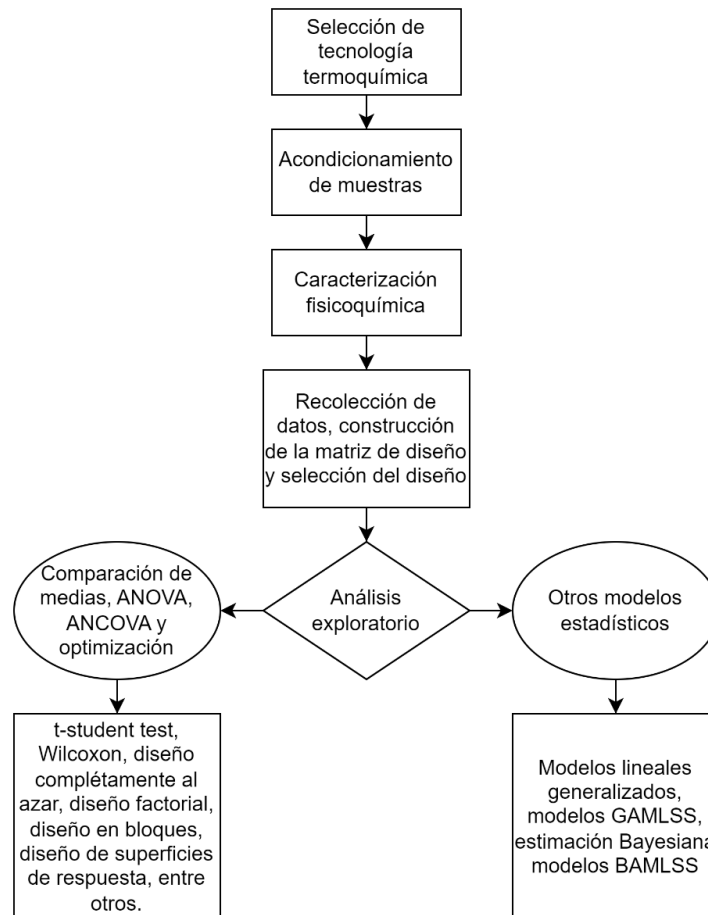


Figura 1. Propuesta metodológica de análisis de datos provenientes de biomásas agroindustriales
Fuente: elaboración propia.

Después, se debe proceder al acondicionamiento de las muestras con el fin de llevar a cabo su caracterización fisicoquímica y poder así obtener información detallada sobre las propiedades físicas y químicas de la muestra. Ahora bien, con miras a la toma de decisiones respecto a las hipótesis iniciales del problema, se propone la recolección y organización de datos, la construcción de la matriz de diseño y selección del diseño de experimento considerando posibles cambios respecto a la planeación inicial para dar solución al problema.

Continuando con la propuesta metodológica presentada en la Figura 1, se observa la necesidad de incorporar en esta el análisis exploratorio de datos. A pesar de que en muchos estudios no se presentan los resultados del análisis exploratorio, este es sumamente importante y necesario, ya que ayuda a comprender la naturaleza de los datos y las relaciones entre las variables o factores involucrados en el estudio. Además, esta herramienta permite al investigador identificar patrones o tendencias en los datos que pueden ser útiles para entender el efecto de los factores sobre la variable respuesta. De aquí, y basados en lo observado con el análisis exploratorio de datos, se debe proceder a la construcción de modelos estadísticos para identificar los posibles efectos y el tamaño de estos sobre la respuesta. Este paso se puede ejecutar aplicando técnicas como las expuestas en la parte izquierda de la Figura 1 (*test* t-student, *test* de Wilcoxon, diseño completamente al azar, diseño factorial, diseño en bloques, diseño de superficie de respuestas, entre otros.), las cuales son las que comúnmente se encuentran en la literatura y/o implementando técnicas de modelación

estadística bayesiana, o asumiendo no normalidad de la respuesta (modelos lineales generalizados), algo que es lógico asumir dados los tipos de respuesta que se estudian en esta área de las ciencias, o modelación bayesiana (ver parte derecha de la Figura 1).

Es de resaltar en la propuesta metodológica de análisis de datos provenientes de experimentos basados en tecnologías termoquímicas, que la modelación bayesiana y los modelos lineales generalizados (MLG) son más flexibles que los modelos clásicos de diseño de experimentos, como ANOVA, y pueden manejar diferentes tipos de variables predictoras y distribuciones no normales. Asimismo, pueden proporcionar predicciones precisas para nuevos conjuntos de datos y calcular intervalos de credibilidad en lugar de intervalos de confianza para una mejor interpretación de la incertidumbre en las predicciones. Estos modelos también son efectivos en el manejo de datos faltantes y pueden imputar valores faltantes basados en la información disponible. Por último, la modelación bayesiana permite la incorporación de información previa en la construcción del modelo, lo que puede ser útil cuando se dispone de información adicional sobre los efectos de los predictores o cuando se desea utilizar la información de estudios previos para informar la construcción del modelo. En fin, esta propuesta metodológica abre la posibilidad de no solo implementar metodologías estadísticas comunes para este tipo de datos, sino complementar los análisis con otras herramientas mucho más flexibles que permitan dar mayor claridad en las interpretaciones y aportar a la toma de decisiones.

Esta propuesta metodológica es ilustrada con dos aplicaciones a datos reales y se evidencian las ventajas de la implementación de modelación estadística diferente a la que comúnmente se viene implementando. A continuación, se presentan los análisis y resultados obtenidos en cada una de las aplicaciones.

3.2 Bagazo de arándanos

3.2.1 Caracterización fisicoquímica

Como parte de la caracterización fisicoquímica de la biomasa bagazo de arándanos, se recomienda realizar un análisis elemental de las muestras, y luego un análisis termogravimétrico (TGA) o proximal. El análisis elemental es una técnica que permite determinar el contenido total de carbono (C), hidrógeno (H), nitrógeno (N) y azufre (S) y oxígeno (O) presentes en las muestras (ver Tabla 1).

De acuerdo con la Tabla 1, se observa que el contenido de carbono (C) es más significativo con respecto a los demás; sin embargo, el contenido de oxígeno (O) es considerable en la muestra, mientras que el contenido de azufre (S) se considera despreciable.

El análisis proximal comprende, entre otros aspectos, la determinación de los porcentajes de humedad y cenizas (las siglas “ND” significan “No definido”) (ver Tabla 2).

Tabla 1. Análisis elemental de bagazo de arándanos (%m/m). Fuente: elaboración propia.

Elemento	Experimental
C	57.13 %
H	7.07 %
O	34.96 %
N	0.83 %
S	ND

Tabla 2. Análisis proximal de bagazo de arándanos (%m/m). Fuente: elaboración propia.

Elemento	Experimental	Teórico [23]
Humedad	90 %	83.0 %
Cenizas	0.15 %	ND
Proteína	ND	0.7 %
Fibra	ND	1.5 %
Carbohidratos	ND	15.3 %

En la Tabla 2 se evidencia que el bagazo de arándano presenta en su estado inicial un alto contenido de humedad, lo cual facilita el proceso de carbonización hidrotermal, muestra un bajo contenido de cenizas, permitiendo con ello un mayor porcentaje útil en la reacción y un bajo nivel de impurezas; la fibra y carbohidratos no fueron analizados experimentalmente.

3.2.2 Análisis de varianza bagazo de arándanos

Comúnmente, en este tipo de datos se realizan análisis de varianzas o técnicas de modelación que buscan identificar la presencia de interacción entre los factores estudiados, pues puede existir una influencia conjunta de dos o más factores en la variable respuesta. Para eso, el modelo lineal con interacciones es una buena alternativa. En la Tabla 3, por ejemplo, se observa la influencia de los factores sobre la respuesta rendimiento (para la biomasa bagazo de arándanos). Además, se nota que existe interacción triple, excepto entre el tiempo, temperatura y potencia, las demás interacciones triples poseen un nivel de significancia estadística valor-p mayor a 0.005. De igual manera, las interacciones dobles también poseen un nivel de significancia prometedor. Solamente la humedad y la temperatura, de forma independiente, presentan un nivel de significancia concluyente, con un valor-p menor a 0.005 SE es el error estándar, R es el rendimiento del hidrocarbón, T es la temperatura, Ti es el tiempo, P es la potencia y H es la humedad.

Tabla 3. Resumen estadístico del bagazo de arándanos para la respuesta rendimiento
Fuente: elaboración propia.

Modelo	Factores	Estimativo	SE	Valor p	Otros estimativos
Modelo lineal múltiple $lm(R \sim (Ti + H + T + P)^3)$	<i>Intercepto</i>	-1.87E+04	2.81E+03	0.00266	R-cuadrado ajustado: 0.9358 Valor p: 0.005387 AIC: -53.10785 BIC: -37.99683
	Ti	1.23E+02	3.52E+01	0.02504	
	H	2.49E+02	3.45E+01	0.00195	
	T	1.00E+02	1.54E+01	0.00285	
	P	5.47E+01	1.08E+01	0.00714	
	Ti:H	-1.65E+00	4.30E-01	0.01838	
	Ti:T	-5.21E-01	1.92E-01	0.05316	
	Ti:P	-1.77E-01	5.66E-02	0.03544	
	H:T	-1.29E+00	1.89E-01	0.00241	
	H:P	-6.83E-01	1.30E-01	0.00631	
	T:P	-3.11E-01	5.83E-02	0.00593	
	Ti:H:T	6.71E-03	2.34E-03	0.04562	
	Ti:H:P	1.55E-03	4.68E-04	0.02960	
	Ti:T:P	4.35E-04	2.34E-04	0.13628	
	H:T:P	3.84E-03	7.02E-04	0.00545	

Además, el símbolo “:” en la columna 2, significa interacción entre los factores indicados. Por ejemplo, H:T significa interacción entre humedad y temperatura. Esta nomenclatura es igual para las siguientes tablas, pero adicionando que C representa el contenido de carbono del hidrocarbón.

Se puede analizar gráficamente la interacción triple entre la temperatura, humedad y potencia en la Figura 2, observando que principalmente la interacción se da cuando la potencia es baja (50 vatios); aquí, cuando pasa la humedad de 70 % al 90 %, se nota un incremento en el rendimiento del proceso con temperaturas de 160 °C, pero para este mismo cambio de humedad, el rendimiento decrece considerablemente cuando la temperatura se incrementa a 200 °C, (ver Figura 2).

Note que este tipo de modelo asume que la respuesta sigue una distribución normal, pero como es natural para estas variables, la normalidad no se puede garantizar siempre. Ahora bien, respecto al componente estadístico de la propuesta metodológica, se recomienda explorar el uso de modelos con respuesta no normales. Para esta ilustración, un modelo con respuesta tipo beta es adecuado debido a que considera variables con soporte entre 0 y 1. Es así como se ilustra el modelo beta para la respuesta *Contenido de Carbono* transformando la respuesta, que naturalmente está entre 0 % y 100 %, a 0 y 1.

En la Tabla 4 se presentan las estimativas de los parámetros del modelo lineal generalizado beta para la variable *Contenido de Carbono*. Se evidencia que este modelo, al igual que en el anterior, la interacción entre tiempo, temperatura y potencia no es significativa, las demás interacciones triples presentan significancia estadística valor-p mayor a 0.005, excepto la interacción triple de humedad, tiempo y potencia que presenta significancia concluyente valor-p menor a 0.005. Este modelo presenta menor BIC en comparación con el modelo anterior, por tanto, este modelo tiene un mejor ajuste, lo que, en términos de calidad de ajuste a los datos, el modelo lineal generalizado con respuesta tipo beta es una mejor alternativa que el tradicional modelo lineal general.

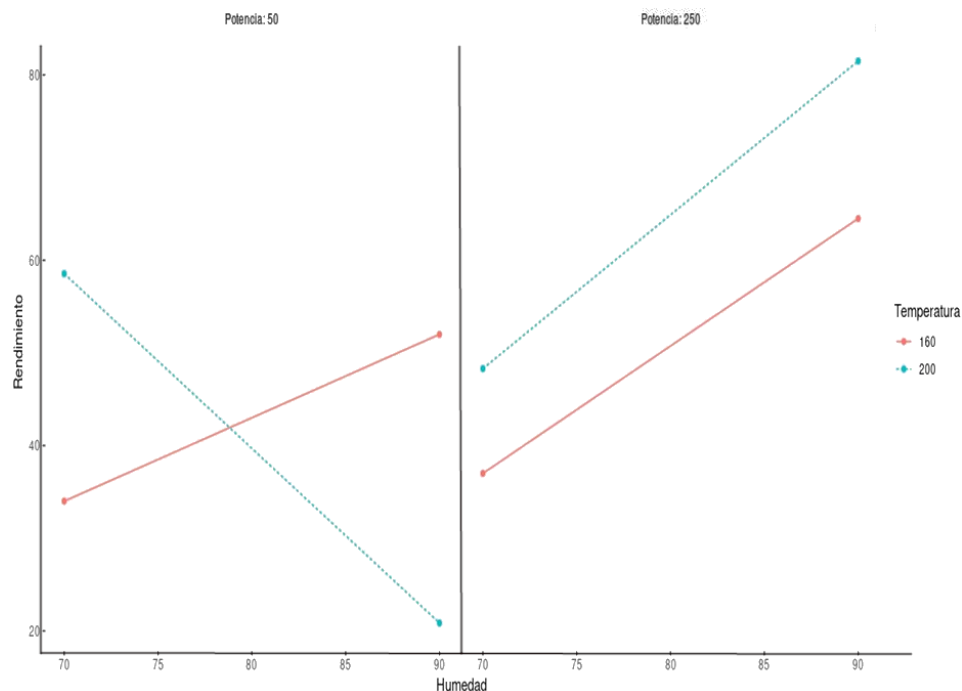


Figura 2. Interacción triple entre humedad (%), temperatura (°C) y potencia (W) para la respuesta rendimiento de bagazo de arándanos. Fuente: elaboración propia.

Tabla 4. Resumen estadístico del bagazo de arándanos para la respuesta contenido de carbono
Fuente: elaboración propia.

Modelo	Factores	Estimativo	SE	Valor p	Otros estimativos
Modelo lineal generalizado tipo beta Betareg ($C \sim (Ti + H + T + P)^3$)	<i>Intercepto</i>	1.39E+04	2.99E+03	3.13e-06	Pseudo R-cuadrado: 0.907 AIC: -74.37286 BIC: -59.26184
	Ti	-7.43E+01	3.76E+01	0.048161	
	H	-1.79E+02	3.68E+01	1.06e-06	
	T	-8.73E+01	1.64E+01	9.83e-08	
	P	-3.96E+01	1.15E+01	0.000599	
	Ti:H	8.27E-01	4.59E-01	0.071240	
	Ti:T	6.14E-01	2.05E-01	0.002777	
	Ti:P	-1.19E-01	6.04E-02	0.048286	
	H:T	1.12E+00	2.02E-01	2.76e-08	
	H:P	4.68E-01	1.39E-01	0.000762	
	T:P	2.72E-01	6.22E-02	1.25e-05	
	Ti:H:T	-6.90E-03	2.50E-03	0.005789	
	Ti:H:P	1.73E-03	5.00E-04	0.000540	
	Ti:T:P	-1.20E-04	2.50E-04	0.631218	
	H:T:P	-3.23E-03	7.51E-04	1.67e-05	
	<i>Phi coeficiente</i>	1131.0	366.8	0.00205	

3.3 Bagazo de manzanas

3.3.1 Caracterización

Al igual que en el caso anterior (bagazo de arándanos), es recomendable, desde la metodología planteada, realizar la caracterización fisicoquímica mediante una análisis elemental y proximal de la biomasa (ver Tablas A3 y A4 en anexos). Sin embargo, el resto del análisis correspondiente a esta ilustración se centra en la modelación estadística.

3.3.2 Análisis estadístico de varianza bagazo de manzanas

La propuesta metodológica incorpora inicialmente un cambio en el paradigma de la normalidad, por lo que para esta biomasa se inicia el análisis con un modelo lineal generalizado beta (ver Tabla 5). Se observa también interacción entre todos los factores (Ti:H:T:P), con un nivel de significancia menor a 0.005; incluso, en este segundo modelo, todos los factores, incluido el tiempo, son significativos de manera independiente.

Si se analiza la significancia estadística de todos los parámetros de cada modelo, se puede ver según el índice AIC y BIC de -91.25805 y -74.3306, respectivamente, que este segundo modelo lineal generalizado tipo beta tiene un mejor ajuste, y en términos predictivos, puede ser más útil que el modelo anterior.

Una nueva alternativa de análisis, desde lo metodológico, es considerar respuesta tipo beta, pero determinar cómo la varianza de la respuesta puede ser explicada por los factores que se estudian. Es así como un modelo GAMLSS comprende una alternativa adecuada para este objetivo.

Tabla 5. Resumen estadístico del bagazo de manzana para la respuesta rendimiento
Fuente: elaboración propia.

Modelo	Factores	Estimativo	SE	Valor p	Otros estimativos
Modelo lineal generalizado tipo beta Betareg ($R \sim (Ti$ $H * T * P)$)	<i>Intercepto</i>	4.830e+01	3.689e+00	< 2e-16	Pseudo R-cuadrado: 0.9952 AIC: -91.25805 BIC: -74.3306
	Ti	-1.343e-01	5.410e-02	0.0131	
	H	-5.227e-01	4.524e-02	< 2e-16	
	T	-2.692e-01	2.007e-02	< 2e-16	
	P	-3.526e-01	1.960e-02	< 2e-16	
	Ti:H	3.651e-04	6.710e-04	0.5864	
	Ti:T	1.154e-03	2.947e-04	9.06e-05	
	H:T	3.017e-03	2.468e-04	< 2e-16	
	Ti:P	3.144e-03	2.923e-04	< 2e-16	
	H:P	4.070e-03	2.418e-04	< 2e-16	
	T:P	1.978e-03	1.073e-04	< 2e-16	
	Ti:H:T	-8.461e-06	3.658e-06	0.0207	
	Ti:H:P	-3.086e-05	3.621e-06	< 2e-16	
	Ti:T:P	-1.980e-05	1.607e-06	< 2e-16	
	H:T:P	-2.313e-05	1.326e-06	< 2e-16	
	Ti:H:T:P	2.036e-07	1.990e-08	< 2e-16	
	<i>Phi coeficiente</i>	2020.6	638.8	0.00156	

Para la variable respuesta rendimiento, y aplicando un modelo GAMLSS (tal como se propone en este artículo), se aprecian los parámetros estimados cuando se modela la media y la varianza. Se observa que los parámetros, cuando se modela la media, cambian con respecto al modelo beta, debido a que aquí se incorpora la modelación de la varianza. Cabe centrarse aquí en el componente de la varianza, debido a que esto sería lo diferencial respecto a los análisis propuestos anteriormente. En este sentido, se observan dos interacciones significativas, una de ellas con significancia estadística prometedora con un valor-p mayor a 0.005. En los factores independientes se presenta significancia estadística en la temperatura y potencia con un valor-p menor a 0.005 (ver Tabla 6). Nótese que los modelos GAMLSS pueden ser implementados en este tipo de análisis y esta metodología propuesta abre la puerta a la exploración de nuevas herramientas estadísticas en el estudio de datos provenientes de tecnologías termoquímicas.

4. CONCLUSIONES

Este estudio proporciona una propuesta metodológica para el análisis estadístico de datos provenientes de biomásas que no siguen una distribución de probabilidad como una normal y son tratadas con tecnologías termoquímicas similares a la carbonización hidrotermal. Se observa que, técnicas como los modelos lineales generalizados y modelos GAMLSS, son alternativas adecuadas de análisis para el investigador cuando se enfrenta a este tipo de datos. Se propone un esquema metodológico con un componente químico y estadístico con el fin de exponer una estructura que aborde un análisis estadístico para analizar el rendimiento y contenido de carbono de biomásas agroindustriales tratadas mediante carbonización hidrotermal.

Tabla 6. Resumen estadístico del bagazo de manzana para la respuesta rendimiento modelo GAMLSS
Fuente: elaboración propia.

Modelo	Factores	Estimativo	SE	Valor p	Otros estimativos
Modelo GAMLSS Media	<i>Intercepto</i>	2.684e+01	7.954e+01	0.753	
	Ti	-4.332e-01	1.187e+00	0.734	
	H	-3.886e-01	9.881e-01	0.714	
	T	-1.619e-01	3.977e-01	0.705	
	P	-2.668e-01	3.182e-01	0.449	
	Ti:H	5.894e-03	1.473e-02	0.709	
	Ti:T	2.648e-03	5.937e-03	0.679	
	H:T	2.347e-03	4.941e-03	0.659	AIC: -97.32256
	Ti:P	4.335e-03	4.750e-03	0.413	SBC: -70.43779
	H:P	3.535e-03	3.952e-03	0.422	
	T:P	1.549e-03	1.591e-03	0.385	
	Ti:H:T	-3.611e-05	7.364e-05	0.650	
	Ti:H:P	-5.293e-05	5.891e-05	0.420	
	Ti:T:P	-2.574e-05	2.375e-05	0.339	
	H:T:P	-2.046e-05	1.976e-05	0.359	
	Ti:H:T:P	3.139e-07	2.946e-07	0.347	
Modelo GAMLSS Sigma	<i>Intercepto</i>	5.433e+01	1.601e+01	0.0079	
	Ti	1.405e-02	7.978e-02	0.86409	
	H	1.421e-01	1.895e-01	0.47244	
	T	-2.991e-01	8.205e-02	0.00535	
	P	-2.975e-01	2.416e-02	6.18e-07	
	Ti:H	-3.625e-05	6.299e-04	0.95536	
	Ti:T	-2.292e-04	3.225e-04	0.49516	
	Ti:P	1.899e-04	6.449e-05	0.01636	
	H:T	-6.926e-04	9.674e-04	0.49220	
	H:P	-4.011e-05	1.935e-04	0.84037	
	T:P	1.447e-03	9.821e-05	1.32e-07	

En el modelo lineal múltiple sobre la respuesta rendimiento (para la biomasa bagazo de arándanos), se evidencia que existe interacción triple con un nivel de significancia estadística valor-p mayor a 0.005, sin embargo, la interacción triple entre tiempo, temperatura y potencia no es significativo. Gráficamente, la interacción entre humedad, potencia y tiempo se da cuando la potencia es baja (50 vatios), en este caso, cuando la humedad pasa de 70 % al 90 %, se observa un incremento en el rendimiento del proceso con una temperatura de 160 °C, pero para este mismo intervalo de humedad, el rendimiento decrece cuando la temperatura se incrementa a 200 °C.

El modelo lineal generalizado con respuesta tipo beta presenta menor BIC en comparación con el modelo lineal múltiple para el bagazo de arándanos, por lo tanto, en términos de calidad de ajuste a los datos, el modelo lineal generalizado con respuesta tipo beta es una mejor alternativa que el modelo lineal múltiple.

Con el propósito de determinar cómo la varianza de la respuesta puede ser explicada por los factores que se estudian, una alternativa adecuada para este objetivo es un modelo GAMLSS. Para el bagazo de manzana y la variable respuesta rendimiento, se estiman los

parámetros al moldearse la media y la varianza. Los modelos GAMLSS pueden ser implementados en este tipo de análisis de datos provenientes de tecnologías termoquímicas. Esta propuesta metodológica abre la puerta a la exploración de nuevas herramientas estadísticas.

5. AGRADECIMIENTOS

Este estudio es un producto financiado por el Instituto Tecnológico Metropolitano (ITM) bajo el programa de “Jóvenes Investigadores” y por FONDECYT 11180752.

CONFLICTOS DE INTERÉS

Se declara que no existe ningún conflicto de intereses económicos, profesionales o personales que puedan influir de forma inapropiada en los resultados obtenidos en este artículo.

CONTRIBUCIONES DE LOS AUTORES

El reconocimiento de la autoría de este artículo se basa en la contribución de cada uno de los autores, cuyos aportes fueron importantes para el desarrollo del documento. En este orden de ideas, todos los autores aportaron en la construcción del artículo. María Eugenia González aportó las bases de datos estudiadas y revisiones al documento en sus diferentes fases de construcción; Sania Pinto-Altamiranda, Sara Gómez y Carlos Barrera-Causil llevaron a cabo el análisis de los datos, la propuesta metodológica y la redacción del manuscrito. El documento final fue completamente revisado, ajustado y aprobado por todos los autores.

6. REFERENCIAS

- [1] P. R. Yaashikaa, P. Senthil Kumar, and S. Varjani, “Valorization of agro-industrial wastes for biorefinery process and circular bioeconomy: A critical review,” *Bioresour Technol*, vol. 343, p. 126126, Jan. 2022. <https://doi.org/10.1016/j.biortech.2021.126126>
- [2] P. K. Sath, S. Duhan, and J. S. Duhan, “Agro-industrial wastes and their utilization using solid state fermentation: a review,” *Bioresour and Bioprocess*, vol. 5, no. 1, pp. 1–15, Jan. 2018. <https://doi.org/10.1186/s40643-017-0187-z>
- [3] L. C. Freitas, J. R. Barbosa, A. L. C. da Costa, F. W. F. Bezerra, R. H. H. Pinto, and R. N. de Carvalho Junior, “From waste to sustainable industry: How can agro-industrial wastes help in the development of new products?,” *Resour Conserv Recycl*, vol. 169, p. 105466, Jun. 2021. <https://doi.org/10.1016/j.resconrec.2021.105466>
- [4] X. Yang, K. Kang, L. Qiu, L. Zhao, and R. Sun, “Effects of carbonization conditions on the yield and fixed carbon content of biochar from pruned apple tree branches,” *Renew Energy*, vol. 146, pp. 1691–1699, Feb. 2020. <https://doi.org/10.1016/j.renene.2019.07.148>
- [5] R. Ebaid, Q. Wang, S. Faisal, L. Li, and A. Abomohra, “Valorization of floral foam waste via pyrolysis optimization for enhanced phenols recovery,” *Chemosphere*, vol. 310, p. 136758, Jan. 2023. <https://doi.org/10.1016/j.chemosphere.2022.136758>
- [6] X. Cao, S. Sun, and R. Sun, “Application of biochar-based catalysts in biomass upgrading: a review,” *RSC Adv*, vol. 7, pp. 48793–48805, Oct. 2017. <https://doi.org/10.1039/C7RA09307A>

- [7] F. M. Jais, C. Y. Chee, Z. Ismail, and S. Ibrahim, “Experimental design via NaOH activation process and statistical analysis for activated sugarcane bagasse hydrochar for removal of dye and antibiotic,” *J Environ Chem Eng*, vol. 9, no. 1, p. 104829, Feb. 2021. <https://doi.org/10.1016/j.jece.2020.104829>
- [8] N. Kazemi Targhi, O. Tavakoli, and A. H. Nazemi, “Co-pyrolysis of lentil husk wastes and *Chlorella vulgaris*: Bio-oil and biochar yields optimization,” *J Anal Appl Pyrolysis*, vol. 165, p. 105548, Aug. 2022. <https://doi.org/10.1016/j.jaap.2022.105548>
- [9] Z. Gan, X. Zhuang, K. Cen, Y. Ba, J. Zhou, and D. Chen, “Co-pyrolysis of municipal solid waste and rice husk gasification tar to prepare biochar: An optimization study using response surface methodology,” *Fuel*, vol. 330, p. 125574, Dec. 2022. <https://doi.org/10.1016/j.fuel.2022.125574>
- [10] X. J. Lee, L. Y. Lee, B. Y. Z. Hiew, S. Gan, S. Thangalazhy-Gopakumar, and H. K. Ng, “Valorisation of oil palm wastes into high yield and energy content biochars via slow pyrolysis: Multivariate process optimisation and combustion kinetic studies,” *Mater Sci Energy Technol*, vol. 3, pp. 601–610, Jan. 2020. <https://doi.org/10.1016/j.mset.2020.06.006>
- [11] M. M. Hasan, M. G. Rasul, M. I. Jahirul, and M. M. K. Khan, “Fast pyrolysis of macadamia nutshell in an auger reactor: Process optimization using response surface methodology (RSM) and oil characterization,” *Fuel*, vol. 333, Part. 2, p. 126490, Feb. 2023. <https://doi.org/10.1016/j.fuel.2022.126490>
- [12] V. Amrhein, S. Greenland, and B. McShane, “Scientists rise up against statistical significance,” *Nature*, vol. 567, no. 7748, pp. 305–307, Mar. 2019. <https://doi.org/10.1038/d41586-019-00857-9>
- [13] D. J. Benjamin *et al.*, “Redefine statistical significance,” *Nature Human Behaviour*, vol. 2, no. 1, pp. 6–10, Sep. 2017. <https://doi.org/10.1038/s41562-017-0189-z>
- [14] K. P. Burnham, and D. R. Anderson, “Multimodel inference: Understanding AIC and BIC in model selection,” *Sociological Methods and Research*, vol. 33, no. 2, pp. 261–304, Nov. 2004. <https://doi.org/10.1177/0049124104268644>
- [15] A. Galecki, and T. Burzykowski. “Linear Mixed-Effects Model.” In: *Linear Mixed-Effects Models Using R. A Step-by-Step Approach. Springer Texts in Statistics*. New York, NY, Springer. 2013, pp 245–273. https://doi.org/10.1007/978-1-4614-3900-4_13
- [16] A. Zamora Saiz, C. Quesada González, L. Hurtado Gil, and D. Mondéjar Ruiz, “An Introduction to Data Analysis in R. Hands-on Coding, Data Mining, Visualization and Statistics from Scratch,” 2020. <https://doi.org/10.1007/978-3-030-48997-7>
- [17] W. W. Stroup, *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*, Boca Ratón, CRC Press, 2016. <https://doi.org/10.1201/B13151>
- [18] H. Madsen, and P. Thyregod, *Introduction to general and generalized linear models*, Boca Ratón, CRC Press 2010. <https://doi.org/10.1201/9781439891148>
- [19] C. Armero, “Data Analysis Using Hierarchical Generalized Linear Models with R,” *J Stat Softw*, vol. 88, no. 1, pp. 1–3, Jan. 2019. <https://doi.org/10.18637/jss.v088.b01>
- [20] J. J. Faraway, *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*, Second Edition, New York, Chapman and Hall/CRC, 2016. <https://doi.org/10.1201/9781315382722>
- [21] M. D. Stasinopoulos, R. A. Rigby, G. Z. Heller, V. Voudouris, and F. De Bastiani, *Flexible regression and smoothing: Using GAMLSS in R*, New York, Chapman and Hall/CRC, 2017. <https://doi.org/10.1201/b21973>
- [22] D. M. Stasinopoulos, and R. A. Rigby, “Generalized Additive Models for Location Scale and Shape (GAMLSS) in R,” *J Stat Softw*, vol. 23, no. 7, pp. 1–46, Dec. 2007. <https://doi.org/10.18637/JSS.V023.I07>
- [23] B. Zhang *et al.*, “Hydrothermal Carbonization of Fruit Wastes: A Promising Technique for Generating Hydrochar,” *Energies*, vol. 11, no. 8, p. 2022, Aug. 2018. <https://doi.org/10.3390/en11082022>

ANEXO**Tabla A1.** Matriz experimental para la carbonización hidrotermal asistida por microondas de bagazo de manzana. Fuente: elaboración propia.

Tratamiento	Valores codificados				Valores reales				Comentario
	X1	X2	X3	X4	Temperatura (°C)	Tiempo (min)	Humedad (%)	Potencia (W)	
1	-1	-1	1	1	160	30	90	250	Ninguno
2	1	-1	-1	1	200	30	70	250	Ninguno
3	-1	-1	-1	1	160	30	70	250	Ninguno
4	1	1	1	-1	200	90	90	50	Ninguno
5	-1	-1	1	-1	160	30	90	50	Ninguno
6	-1	1	1	1	160	90	90	250	Ninguno
7	0	0	0	0	180	60	80	150	Punto central
8	1	1	-1	1	200	90	70	250	Ninguno
9	0	0	0	0	180	60	80	150	Punto central
10	1	-1	1	-1	200	30	90	50	Ninguno
11	1	-1	1	1	200	30	90	250	Ninguno
12	-1	1	-1	1	160	90	70	250	Ninguno
13	0	0	0	0	180	60	80	150	Punto central
14	1	-1	-1	-1	200	30	70	50	Ninguno
15	-1	1	1	-1	160	90	90	50	Ninguno
16	0	0	0	0	180	60	80	150	Punto central
17	1	1	1	1	200	90	90	250	Ninguno
18	1	1	-1	-1	200	90	70	50	Ninguno
19	-1	1	-1	-1	160	90	70	50	Ninguno
20	-1	-1	-1	-1	160	30	70	50	Ninguno

Tabla A2. Matriz experimental para la carbonización hidrotermal asistida por microondas de bagazo de arándanos. Fuente: elaboración propia.

Tratamiento	Valores codificados				Valores reales				Comentario
	X1	X2	X3	X4	Temperatura (°C)	Tiempo (min)	Humedad (%)	Potencia (W)	
1	1	-1	-1	1	200	30	70	250	Ninguno
2	-1	-1	1	-1	160	30	90	50	Ninguno
3	1	1	-1	-1	200	90	70	50	Ninguno
4	1	1	-1	1	200	90	70	250	Ninguno
5	-1	-1	1	1	160	30	90	250	Ninguno
6	1	1	1	1	200	90	90	250	Ninguno
7	1	1	1	-1	200	90	90	50	Ninguno
8	1	-1	-1	-1	200	30	70	50	Ninguno
9	0	0	0	0	180	60	80	150	Punto central
10	-1	1	-1	-1	160	90	70	50	Ninguno
11	-1	1	-1	1	160	90	70	250	Ninguno
12	-1	-1	-1	1	160	30	70	250	Ninguno
13	0	0	0	0	180	60	80	150	Punto central
14	-1	1	1	-1	160	90	90	50	Ninguno
15	-1	-1	-1	-1	160	30	70	50	Ninguno
16	1	-1	1	-1	200	30	90	50	Ninguno
17	1	-1	1	1	200	30	90	250	Ninguno
18	0	0	0	0	180	60	80	150	Punto central
19	-1	1	1	1	160	90	90	250	Ninguno

Análisis elemental

Es una técnica realizada con un analizador elemental que permite determinar el contenido total de carbono (C), hidrógeno (H), nitrógeno (N) y azufre (S) y oxígeno (O) presentes en las muestras (ver Tabla A3).

Tabla A3. Análisis elemental de bagazo de manzana (%m/m). Fuente: elaboración propia.

Elemento	Experimental
C	37.68 %
H	6.15 %
O	55.35 %
N	0.34 %
S	No definido

La caracterización elemental de bagazo de manzana, a diferencia del análisis del bagazo de arándano (Tabla 1), presenta menor contenido de carbono (C) y mayor contenido de oxígeno (O); de igual forma, el azufre (S) se considera despreciable.

Análisis proximal

El análisis proximal comprende la determinación de los porcentajes de humedad y cenizas (las siglas “ND” significan “No definido”) (ver Tabla A4).

Tabla A4. Análisis proximal de bagazo de manzana (%m/m). Fuente: elaboración propia.

Elemento	Experimental	Teórico [23]
Humedad	5.97 %	ND
Cenizas	0.48 %	2.47 %
Proteína	ND	0.19 %
Fibra	ND	2.70 %
Carbohidratos	ND	15.25 %

En este análisis proximal se observa que el contenido de ceniza presente en la muestra disminuye el nivel de impurezas, observación similar a la del bagazo de arándano, la humedad inicial es pequeña con respecto a la biomasa de arándano (Tabla 2), lo que no favorecer la carbonización hidrotermal.