

# Análisis y modelación estadística del proceso de añejamiento de rones en una ronera cubana

## Analysis and statistical modelling in the aging process of rums in a Cuban rum factory

Beatriz García-Castellanos<sup>1</sup>, Osney Pérez-Ones<sup>2</sup>, Lourdes Zumalacárregui de Cárdenas<sup>3</sup>,  
Idania Blanco-Carvajal<sup>4</sup>, Luis Eduardo López de la Maza<sup>5</sup>

<sup>1</sup> Instituto Cubano de Investigaciones de los Derivados de la Caña de Azúcar, La Habana, Cuba. beatriz.garcia@icidca.azcuba.cu,

<sup>2</sup> Universidad Tecnológica de La Habana "José Antonio Echeverría, La Habana, Cuba. osney@quimica.cujae.edu.cu,

<sup>3</sup> Universidad Tecnológica de La Habana "José Antonio Echeverría, La Habana, Cuba. lourdes@quimica.cujae.edu.cu,

<sup>4</sup> Instituto Cubano de Investigaciones de los Derivados de la Caña de Azúcar, La Habana, Cuba. idania.blanco@icidca.azcuba.cu.

<sup>5</sup> Universidad Tecnológica de La Habana "José Antonio Echeverría", La Habana, Cuba. llopezm@quimica.cujae.edu.cu

Recibido: 28 de mayo de 2020 Aceptado: 23 de junio de 2020

**Resumen**– El añejamiento de ron experimenta pérdidas de volumen, denominadas mermas. Las variables de operación: producto, rambra, posiciones horizontal y vertical, fecha, volumen, grado alcohólico, temperatura, humedad y tiempo de añejamiento, registradas en bases de datos, contienen información valiosa para estudiar el proceso. Las variables cualitativas se procesaron utilizando el *software Weka 3.8.0* mientras que las cuantitativas se sometieron a un análisis estadístico empleando *Statgraphics Centurion XVII.2*. Las mayores mermas corresponden a toneles ubicados en áreas donde irradia el sol, favoreciendo la evaporación del producto. Las variables temperatura y humedad presentan coeficientes de variación muy elevados; debido a la no existencia de control sobre estos factores y sugiriendo la necesidad de regulación. Se obtuvo un modelo de regresión que logra predecir las mermas en función de las variables: mes numérico, volumen y tiempo de añejamiento con valores de error cuadrático medio (ECM) y  $R^2$  de 0,115 y 95,88 % respectivamente.

**Palabras clave**– Añejamiento, ron, mermas, *software Weka*, *Statgraphics Centurion*, modelo de regresión.

**Abstract**– The rum aging process shows volume losses, called wastage. The operation variables: product,

boardwalk, horizontal and vertical positions, date, volume, alcoholic degree, temperature, humidity and aging time, recorded in databases, contain valuable information to study the process. The qualitative variables were processed using *Weka 3.8.0* software while the quantitative variables underwent a statistical analysis using *Statgraphics Centurion XVII.2*. The biggest reductions correspond to barrels located in areas which solar irradiation, favoring the evaporation of the product. The variable temperature and humidity present very high variation coefficients; these factors are uncontrolled so a regulation process is suggested. A regression model was obtained that predicts the losses based on the variables: numerical month volume and aging time with mean square error values (ECM) and  $R^2$  of 0.115 and 95.88 % respectively.

**Keywords**– Aging, rum, wastages, *Weka software*, *Statgraphics Centurion software*, regression model.

## 1. INTRODUCCIÓN

En el ron fresco, como en la mayoría de las bebidas alcohólicas destiladas, el aroma recuerda a la materia prima utilizada. Este varía cuando se deja en reposo en recipientes de roble por un tiempo determinado; es decir, durante

**Citar este artículo como:** García-Castellanos, B., Pérez-Ones, O., Zumalacárregui de Cárdenas, L., Blanco-Carvajal, I., & López de la Maza, L. E. Análisis y modelación estadística del proceso de añejamiento de rones en una ronera cubana. ITECKNE, 17(2), 2020 pp. 91 - 109. doi: <https://doi.org/10.15332/iteckne.v17i2.2468>

el proceso de permanencia de los aguardientes en barriles de roble y al cual se le conoce comúnmente como “envejecimiento” o tiempo de añejamiento [1]. Este proceso se ve afectado por gran cantidad de factores que, a medida que pasa el tiempo, le otorgan al producto una mayor calidad, consiguiéndose un mejor balance entre el sabor a madera y el impacto etanólico [2].

Durante el proceso de añejamiento del ron se producen pérdidas del producto. El añejamiento del ron no cambia ni transforma la bebida, sino que desarrolla y sublima sus cualidades latentes [1]. Es por lo que, en el contexto de excelencia en que compiten estas bebidas, surge el interés por estudiar la merma en el volumen de ron durante su añejamiento en relación con las condiciones ambientales existentes en la solera.

En el mundo el valor medio de pérdidas anuales que se considera para toneles de 180 litros es de 10% del volumen almacenado. Sin embargo, la literatura señala para los países del Caribe 13% de pérdidas anuales [3]. En un trabajo realizado en las soleras de la ronera de Cárdenas, en los años 1985-1986, se corroboró una pérdida de 10,5% anual. No se han encontrado reportes recientes del valor de estas pérdidas en Cuba. En Jamaica se reporta un 6% de pérdidas [4].

Martí et al. [5] analizaron el comportamiento de parámetros ambientales en las naves de añejamiento en la ronera central “Agustín Rodríguez Mena”, concluyendo que la temperatura dentro del barril se comporta uniformemente, debido a las propiedades térmicas de la madera de roble, que actúa como aislante moderado, lo que permite amortiguar las diferencias térmicas durante el día. Igualmente, no encontraron diferencias significativas en la temperatura y la humedad relativa tomadas en diferentes puntos de la nave.

Se han realizado estudios quimiométricos en los rones en Cuba, pero han estado dirigidos a poder identificar la autenticidad de los rones por comercializar [6].

Durante el proceso productivo se determinan numerosos parámetros, que pueden variar de una ronera a otra, que se registran sistemáticamente. Toda esta memoria almacenada constituye una valiosa fuente de información que puede ser útil en la comprensión del presente y en la predicción del futuro. Los modelos predictivos obtenidos por técnicas de minería de datos constituyen una alternativa a los modelos matemáticos y a la vez una herramienta para el análisis de la información almacenada [7] de los procesos de añejamiento de rones.

Por consiguiente, dado que se estima que hay altos porcentajes de mermas en volumen de las bases en solera y no están esclarecidos los factores que influyen en estas pérdidas, se realiza el trabajo con el objetivo de proponer un modelo de regresión que logre predecir las mermas en función de las variables medidas más influyentes en el añejamiento del ron Añejo.

## 2. METODOLOGÍA

### 2.1 Determinación de mermas en toneles de añejamiento de ron

El presente estudio se realizó entre los años 2018 y 2019 en las bodegas de ron Añejo, ubicadas en el Centro de Referencia de Alcoholes y Bebidas (CERALBE), pertenecientes al ICIDCA, bajo las condiciones normales de producción de temperatura y humedad ambiental, utilizando los datos históricos del periodo de tiempo entre octubre de 2014 y diciembre de 2015.

Para la determinación de las variables se contó con un dispositivo de medición de temperatura y humedad relativa ambiente marca HOB0, modelo H08-003-02 con registrador de datos en una solera. Estas variables físicas se midieron ininterrumpidamente durante 13 meses.

Los toneles se hallan almacenados en la ronera en cinco ramblas que poseen una distribución vertical de cinco niveles cada una; en ellas se observa la identificación numérica de cada tonel. Se escogieron tres puntos de la solera (centro y extremos) y se midieron los toneles del primer, tercer y quinto nivel, para evaluar las pérdidas en relación con su posición y su altura. Se escogieron 45 toneles, los cuales tienen almacenados diferentes bases de ron.

El dispositivo de medición se instaló sobre las duelas de uno de los barriles del nivel más elevado de la rambla, en las cercanías del tapón utilizado para el llenado o vaciado de la barrica y del techo de la solera.

Los toneles contienen cinco productos con diferentes concentraciones de etanol (formulaciones) que se encuentran distribuidos aleatoriamente en la solera. Los productos se identifican como: BVi, BVa, A, Ag, VV [8].

Esta distribución y selección de toneles y variedad de productos para añejar da lugar a cinco variables cualitativas: producto, rambla, fecha, posición horizontal y posición vertical.

### 2.2 Análisis de las variables cualitativas

El *software* Waikato Environment for Knowledge Analysis (Weka 3.8.0), creado por la Universidad de Waikato, Hamilton, Nueva Zelanda presenta en su interfaz de trabajo aplicaciones muy atractivas para el campo del análisis de procesos, entre las que se pueden citar: exploración, experimentación, ambiente de flujo de trabajo y banco de trabajo [9]. Dentro de la plataforma de exploración se puede acceder a las técnicas de preprocesamiento de datos, clasificación, agrupamiento, selección de atributos, visualización de los resultados y asociación.

Esta última técnica ofrece el conjunto de “mejores reglas encontradas” de acuerdo con el nivel de confianza establecido de 0,6 entre las cinco variables cualitativas de entrada y la de salida en cuestión (mermas). Se implemen-

tó el algoritmo Apriori debido a que constituye una opción para extraer reglas de asociación entre clases y es adaptable para explicar hasta segundas referencias. La variable merma inicialmente presentaba carácter numérico y para lograr que el *software* la reconociera como al resto de las variables cualitativas se transformó y categorizó, de manera que una vez calculada su media, se estableció que las mediciones que superaran ese valor no cumplieran y las menores a este eran admisibles.

### 2.3 Análisis de las variables cuantitativas

Las variables cuantitativas estudiadas fueron: volumen, grado alcohólico, temperatura, humedad y tiempo de añejamiento.

Las variables cuantitativas que regulan un proceso de producción determinado, medidas de manera continua durante un tiempo, pueden presentar incertidumbre desde el propio momento de la medición, lo cual unido al error intrínseco de los instrumentos de medición y a la propia distribución espacial que sigan, introducen ruido o lo que se conoce comúnmente como puntos discrepantes. Estos últimos pueden afectar la confianza en los resultados del proceso de modelado y conlleva la necesidad de realizar un análisis estadístico previo, del que se pueda arribar a información relevante referida a: la media, desviación estándar, coeficiente de variación, mínimo, máximo, rango, sesgo y kurtosis de los datos.

Se utilizó el *software* Statgraphics Centurion XVII.2 para el análisis estadístico de datos cuantitativos obteniendo los estadígrafos anteriormente mencionados de manera independiente, tanto para cada variable de operación como para las mermas [10].

#### 2.3.1 Limpieza de datos

Para la identificación preliminar de puntos discrepantes se utilizó la coincidencia de dos criterios: análisis de las variables cuantitativas con un coeficiente de variación > 10% y el análisis del gráfico de cajas y bigotes para los subconjuntos correspondientes a cada mes numérico.

#### 2.3.2 Pruebas estadísticas de KMO y de esfericidad de Bartlett

El *software* Statgraphics Centurion XVII.2 se utilizó para realizar las pruebas estadísticas de KMO (Kaiser-Meyer-Olkin) y la prueba de esfericidad de Bartlett. La prueba de KMO indica el grado en que cada variable de un conjunto se predice sin error por las restantes variables. Se sugiere un valor de aceptación de 0,5 o más para que se aconseje realizar el análisis de factores. Valores entre 0,5 y 0,7 son regulares y entre 0,7 y 0,8 son buenos. La prueba de esfericidad de Bartlett se aplica para encontrar la presencia de correlaciones entre las variables. Para el análisis de los factores hace falta que haya relaciones entre las variables. Por tanto, se requiere un  $p < 0,05$  para que exista significación [11].

### 2.3.3 Modelo de regresión

El software Statgraphics Centurion XVII.2 permite estudiar las posibilidades de explicar el proceso a través de un modelo de regresión entre la variable dependiente: mermas y las variables independientes que sigan una distribución normal. Si estas variables no presentan dicho comportamiento deben ser sometidas a un procesamiento previo para normalizarlas.

El ajuste del modelo de regresión permitió seleccionar cuáles son las variables que mayor variabilidad del proceso explican y se deben incluir en el modelo, atendiendo a los valores de error cuadrático medio (ECM), error estándar de los estimados (EEE), coeficiente de determinación ( $R^2$ ) y la prueba de Durbin-Watson.

## 3. RESULTADOS

### 3.1 Determinación de ruido, limpieza y selección de los datos que se utilizarán

Se trabajó con un conjunto inicial de datos de 900 instancias y diez variables, de ellas cinco cualitativas: producto, rambla, fecha y posiciones horizontal y vertical; mientras que las restantes: tiempo de añejamiento, volumen, temperatura, humedad y grado alcohólico inicial tienen carácter cuantitativo, como se declaró en el acápite Metodología. La matriz inicial de datos se redujo a 546 valores, ya que presentaba mediciones incompletas con respecto a tres toneles (19, 21 y 159) y al período de tiempo que comprende de febrero a septiembre de 2014.

### 3.2 Reglas obtenidas del procesamiento de las variables cualitativas

El empleo de la técnica de asociación utilizando el software Weka 3.8 arrojó la siguiente lista de resultados en que se reportan las reglas que relacionan las variables de operación con la variable de salida, mermas y que se presentan en la Tabla I.

Tabla I.

LISTA DE RESULTADOS REPORTADOS POR WEKA 3.8 PARA MOSTRAR LAS REGLAS QUE RELACIONAN LAS VARIABLES CUALITATIVAS

Regla No.	Texto de la regla obtenida
1	61 muestras de mermas no cumplen en la rambla uno; corresponden al producto BVi.
2	56 muestras de mermas admisibles en la rambla uno; corresponden al producto BVi.
3	62 muestras de mermas admisibles en la rambla dos; corresponden al producto BVi.
4	58 muestras de mermas no cumplen en la posición vertical derecha; corresponden al producto BVi.
5	67 muestras de mermas del total de muestras no cumplen para el mes de octubre.
6	59 muestras de mermas admisibles en la posición vertical derecha; corresponden al producto BVi.
7	62 muestras de mermas no cumplen en la posición vertical centro; corresponden al producto BVi.

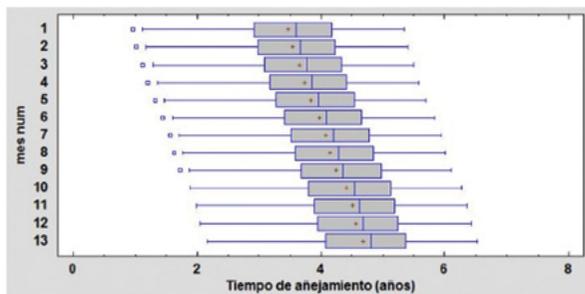
### 3.3 Resultados del análisis estadístico de las variables cuantitativas

En la Tabla II se muestran los estadígrafos resultantes del análisis descriptivo de las cinco variables de operación, así como de la variable de salida.

#### 3.3.1 Gráfico de cajas y bigotes

De estas variables independientes estudiadas, solo el tiempo de añejamiento presenta valores fuera de los bigotes del gráfico de cajas y bigotes, como se aprecia en la Fig. 1. Por no presentar comportamientos anómalos no se presentan los gráficos de las restantes variables.

Fig. 1 GRÁFICO DE CAJAS Y BIGOTES PARA LA VARIABLE TIEMPO DE AÑEJAMIENTO



Fuente: Los autores.

#### 3.3.2 Pruebas estadísticas KMO y esfericidad de Bartlett

La prueba estadística KMO arrojó un valor de 0,5219 que, aunque está muy cercana al límite inferior que se recomienda para la aplicación del análisis de componentes principales cumple con la premisa de ser mayor de 0,5. Esta es un indicio de la cantidad de varianzas común que está presente en la tabla de valores y la necesidad de extraer factores del conjunto de variables.

La prueba de esfericidad detectó un valor - P de 0,0 inferior a 0,05 y por tanto existe correlación significativa al menos entre algunas variables. De aquí que se recomiende aplicar el análisis de componentes principales para eliminar estas correlaciones y facilitar la interpretación de los resultados [12], [13].

#### 3.3.3 Modelo de regresión

Para encontrar un modelo de regresión que permitiera relacionar las mermas con algunas de las variables que se miden en el proceso se requirió realizar un tratamiento previo de las variables originales, ya que, como se discutirá posteriormente, presentaban alejamiento de la distribución normal. Se agruparon las mediciones correspondientes a cada mes estudiado (trece en total) y se determinó la media de dichas variables, de manera que se contaba con trece instancias para cada una. Además, se decidió incorpo-

rar al análisis la variable mes numérico. Posteriormente se comprobó que este subgrupo de muestras, representativo del conjunto original, sí presenta un comportamiento normal, como se muestra en los estadígrafos de la Tabla III, ya que los estadígrafos sesgo y kurtosis se encuentran en el intervalo (-2; 2) [14], [15].

El mejor modelo obtenido se muestra en la Ecuación (1); contiene tres variables: mes numérico, volumen y tiempo de añejamiento y presenta valores de error cuadrático medio (ECM), error estándar de los estimados (EEE) y coeficiente de determinación (R<sup>2</sup>) de 0,194; 0,339 y 95,88%, respectivamente. La prueba estadística de Durbin-Watson (DW) de residuales, con un valor de 1,97, cercano del valor esperado 2, se usa para determinar si existe alguna correlación significativa basada en el orden en que aparecen estos datos en la data.

$$\text{Mermas} = 1,27762 * \text{Mes numérico} - 0,17843 * \text{volumen} + 9,75862 * \text{Tiempo de añejamiento (1)}$$

## 4. DISCUSIÓN

De acuerdo con los resultados de la Tabla I, la interpretación de las reglas reportadas por el software Weka 3.8 permite concluir que las muestras con un valor de mermas mayor a su media están relacionadas con el producto BVi, el cual se encontraba en la rambla uno predominantemente, ubicada en el final de la solera, muy cerca de la pared que limita el lugar con el exterior. En dicha pared irradia el sol de manera ininterrumpida durante el horario de la mañana y presenta un vidrio que permite el paso de los rayos solares en mayor medida que el resto de las paredes circundantes. Esto provoca un calentamiento de esta área y el aumento de la temperatura en los toneles más cercanos, favoreciendo la evaporación del producto.

De la técnica aplicada se deduce que las muestras de BVi ubicadas en la rambla dos, en su mayoría, presentan valores de mermas admisibles. Este comportamiento se corresponde con su ubicación, situada en el centro de la bodega de añejamiento, alejada de la periferia y protegida entre las ramblas uno y tres, por lo que no se ve afectada de manera directa por las variables ambientales y donde existe un mayor flujo de aire.

El análisis además evidencia el elevado número de muestras, correspondientes al mes octubre, que presentan valores de mermas mayores a la media; debido al aumento brusco que experimentó la variable temperatura.

Las muestras ubicadas en las posiciones: vertical derecha y centro se encuentran exactamente debajo de la escalera que comunica con el piso superior; donde se realiza el tratamiento con viruta. Durante este procedimiento la energía calorífica cedida es alta, lo cual provoca el lógico aumento de la temperatura en esa zona, mayor evaporación y mermas. Se aprecia que el número de muestras con valores superiores a la media en cada una es similar, al estar afectada toda la zona por el efecto de calentamiento en el piso superior.

Tabla II.  
RESUMEN DE ESTADÍGRAFOS DESCRIPTIVOS DE LAS VARIABLES CUANTITATIVAS

Variable/ Estadígrafo	Volumen (L)	Grado alcohólico (°GL)	Temperatura (°C)	Humedad (%)	Tiempo de añejamiento (años)	Mermas (L)
Número de mediciones	546	546	546	546	546	546
Media o promedio	164,15	56,16	29,80	52,86	4,06	1,56
Desviación estándar	12,52	7,11	4,71	8,33	1,04	0,94
Coefficiente de variación (%)	7,63	12,66	15,80	15,77	25,48	60,55
Mínimo	120,76	48,7	23,82	43,75	0,96	0
Máximo	190,21	75,57	37,0	74,0	6,53	5,24
Rango	69,45	26,87	13,18	30,25	5,57	5,24
Sesgo estandarizado	-6,64	11,01	3,36	10,68	-3,91	9,57
Kurtosis estandarizada	1,91	2,06	-5,60	3,75	0,73	7,84

Tabla III.  
RESUMEN DE ESTADÍGRAFOS NORMALIZADOS PARA CONSTRUIR EL MODELO DE REGRESIÓN

Variable/Estadígrafo	Volumen (L)	Grado alcohólico (°GL)	Temperatura (°C)	Humedad (%)	Tiempo de añejamiento (años)	Mermas (L)
Cantidad	13	13	13	13	13	13
Media o promedio	164,14	56,16	29,80	52,86	4,06	1,39
Desviación estándar	5,93	0,36	4,90	8,67	0,40	0,41
Coefficiente de variación (%)	3,61	0,64	16,43	16,40	9,96	29,85
Mínimo	154,33	55,72	23,82	43,75	3,46	0,59
Máximo	172,51	56,74	37,0	74,0	4,69	2,11
Rango	18,18	1,03	13,18	30,25	1,22	1,52
Sesgo estandarizado.	-0,55	0,85	0,59	1,87	0,034	-0,40
Kurtosis estandarizada	-0,82	-0,96	-0,84	1,34	-0,94	0,19

Del análisis de la Tabla II, en la que se presenta el análisis descriptivo de cada variable cuantitativa independiente se puede plantear que, en términos estadísticos, cuatro de las variables cuantitativas independientes, exceptuando el volumen, presentan coeficientes de variación superiores a 10%.

Las mermas y el tiempo de añejamiento presentan altos coeficientes de variación en relación con el resto. Para la primera variable esta variación ha motivado la realización del trabajo por la necesidad de conocer los principales factores que condicionan el cambio. El tiempo de añejamiento debe su variabilidad a que existen toneles en la base de datos muy jóvenes, de menos de un año de añejamiento y otros de casi siete años, lo que representa un espectro amplio de condiciones.

Las variables ambientales temperatura y humedad relativa presentan valores similares de este estadígrafo y superiores a 10%, indicando la no existencia de control sobre estos factores durante el proceso de añejamiento y por tanto la posible necesidad de regulación. Estos valores son superiores a los reportados por Martí et al. [5].

Los estadígrafos sesgo y kurtosis, en su mayoría, presentan valores distantes del intervalo  $-2; 2$ , lo cual muestra un alejamiento del comportamiento normal para todas las variables, que puede invalidar cualquier prueba estadística en relación con la desviación estándar. Esto limitó el uso de herramientas para pronosticar valores de variables de datos en bruto.

Al observar el diagrama de cajas y bigotes de la Fig. 1 se aprecia que los puntos que se encuentran fuera de los bigotes corresponden a las mediciones 40, 41, 42, 43, 44, 45, 46, 47 y 48 de los primeros nueve meses de estudio. Estos constituyen los menores valores de la variable analizada y están directamente relacionados con el volumen. En estos casos se presentan los mayores valores de volúmenes registrados, pertenecientes a toneles que se encontraban prácticamente llenos cuando comenzó la investigación y, por lo tanto, los menos añejados.

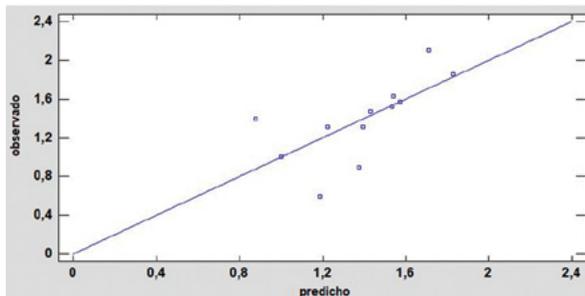
De la Tabla III se concluye que las normalizaciones realizadas a las variables transforman estas en conjuntos de datos con distribuciones normales. El modelo obtenido tiene una buena calidad de ajuste de acuerdo con los parámetros ECM,  $R^2$  y a la prueba de Durbin-Watson, la que con su valor

cercano a 2 indica que no hay correlación significativa basada en el orden en que aparecen los datos.

Las variables temperatura, humedad y grado alcohólico no resultaron estadísticamente significativas para la predicción de las mermas luego de ajustar el modelo de regresión. Como se aprecia a mayor tiempo de añejamiento mayores serán las mermas, hecho que viene dado por la mayor evaporación a que se somete el producto en los barriles con el cursar de los meses, resultados encontrados por otros autores [2]. Además, en el caso de que el barril disminuya su nivel de líquido, o sea el volumen, y la tapa se seque, se forman grietas por las cuales se facilita la pérdida de vapores y el sistema se comporta más como un sistema abierto en ciertos instantes del día. Con estas fisuras, es mucho más fácil este mecanismo de salida de vapores que el proceso de difusión a través de la madera, el cual es mucho más lento, propiciándose un aumento en las mermas. También el modelo manifiesta que con el trascurso de la investigación (el avance en meses) disminuyen las mermas en proceso, lo cual puede estar condicionado por el elevado aumento de la temperatura que se manifestó en el mes de octubre, al comienzo de las mediciones, conllevando un aumento excesivo de las mermas en un principio.

En la Fig. 2 se muestra el gráfico que relaciona datos observados vs predichos para la variable mermas.

Fig. 2 COMPARACIÓN ENTRE LOS VALORES REALES Y PREDICHOS PARA EL MODELO DE REGRESIÓN DE LA VARIABLE MERMAS



Fuente: Los autores.

## 5. CONCLUSIONES

El análisis de las variables cualitativas arrojó que los mayores valores de mermas se encuentran en los toneles ubicados cerca de las paredes que limitan con el exterior, ya que el incremento de temperatura en la zona en que la radiación solar es mayor, favorecerá la evaporación del producto.

Las variables ambientales temperatura y humedad relativa tienen altos coeficientes de variación, lo que se corresponde con el hecho de que en la solera no se controlan estos parámetros durante el proceso de añejamiento.

Se obtuvo un modelo de regresión lineal múltiple que logra predecir las mermas en función de las variables: mes numérico, volumen y tiempo de añejamiento con valores de error cuadrático medio de 0,194 y coeficiente de determinación de 95,88 %. La nave de añejamiento contiene toneles con tiempos de añejamiento muy diferentes, de ahí que se aprecie la alta influencia de esta variable en el modelo predictivo de las mermas. Durante todo el año no existe estabilidad en las condiciones ambientales presentes en la solera, las variables temperatura y humedad varían constantemente y están condicionadas por el mes en cuestión, lo cual explica la necesidad de identificar numérica y secuencialmente cada uno de los meses de investigación, para así definir cuáles son más propicios para desencadenar elevados volúmenes de mermas.

Se evidenció que son de utilidad las herramientas de análisis de procesos para extraer información útil de los datos almacenados durante años en la planta y modelar el proceso de añejamiento de ron.

## REFERENCIAS

- [1] J. Pérez and R. Delgado, "Bebidas alcohólicas destiladas y añejada," [CD-ROM], La Habana: IIIA., 2007.
- [2] P.E. Rico. "Reducción del porcentaje de merma de etanol en el proceso de añejamiento para la elaboración de ron", Tesis presentada en opción del título de Ingeniero Químico, Universidad Simón Bolívar, Venezuela, 2008.
- [3] M. Boscolo, B. S. Lima-Neto, and D. W. Franco, "Envelhecimento de aguardente de cana-de-açúcar em tonéis de madeira", *Engarrafador Moderno*, vol. 6 (41): 30-33, Set. 1995.
- [4] J. Marcano, (2019, Septiembre 15). "El ron. Tradición en Las Antillas" [Online]. Available: <https://mipais.jmarcano.com/economia/ron.html>
- [5] C. A. Martí, J. A. Fabelo, E. González, A. Lago y Y. Rodríguez, "Comportamiento de parámetros ambientales en las naves de añejamiento en la ronera central "Agustín Rodríguez Mena" *Centro Azúcar*, vol 45 (4): 41-50, Oct. 2018.
- [6] E. Zayas, M. Lorenzo, D. Hernández, I. Blanco. "Clasificación multivariante de rones añejos cubanos". *Revista Cubana de Ingeniería*, vol. 5 (2): 62-67, May. 2014.
- [7] L. E. López, L. Zumalacárregui, y O. Pérez, "Análisis de componentes principales aplicado a la fermentación alcohólica", *Revista Científica de la UCSA*, vol. 6 (2): 11-19, Ago. 2019.
- [8] M. D. González, M. A. Vázquez, y D. Redondo, "Estudio comparativo de rones y aguardientes añejados en barriles de roble y por el proceso acelerado", *ICIDCA*, vol. XL (3): 12-19, Sept. 2006.
- [9] M. A Espinoza. "Weka, áreas de aplicación y sus algoritmos: una revisión sistemática de literatura", *Revista Científica Ecociencia*, vol. 5, pp. 1-26. Dic. 2018.
- [10] StatPoint Technologies, Inc. *Statgraphics Centurion XVII. User's Manual*. 2014.
- [11] CAMO Software products. *The Unscrambler X v10.4. User's Manual*. 2016.

- [12] I. T. Jolliffe, and J. Cadima, "Principal component analysis a review and recent developments". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374, Apr. 2016. DOI: <https://doi.org/10.1098/rsta.2015.0202>.
- [13] A. Tharwat, "Principal component analysis - A tutorial", *International Journal of Applied Pattern Recognition*, vol. 3 (3): 197- 238. 2016. DOI: <https://doi.org/10.1504/ijapr.2016.079733>.
- [14] A. Dean, D. Voss, and D. Draguljić, *Design and analysis of experiments*. Berlin: Springer International Publishing, 2017. DOI: <https://doi.org/10.1007/978-3-319-52250-0>.
- [15] W. K. Hardle, and L. Simar, *Applied multivariate statistical analysis*. Berlin: Springer-Verlag, 2015. DOI: <https://doi.org/10.1007/978-3-662-45171-7>.