

# Pronóstico de precios de la carne de cerdo mexicana, con base en el crudo WTI, Maíz y Soya

## Mexican pork price forecast, based on WTI crude oil and Corn and Soybean grains

Sergio Orozco-Cirilo<sup>1</sup>, Juan Manuel Vargas-Canales<sup>1</sup>, Sergio Ernesto Medina-Cuéllar<sup>2</sup>, Juan Antonio Bautista<sup>3\*</sup>

<sup>1</sup>Universidad de Guanajuato, Campus Celaya Salvatierra, División de Ciencias Sociales y Administrativas, Departamento de Estudios Sociales. Guanajuato, México.

<sup>2</sup>Universidad de Guanajuato, Campus Irapuato-Salamanca, Departamento de Arte y Empresa. Guanajuato, México.

<sup>3</sup>Universidad de Guanajuato, Campus Irapuato-Salamanca, Departamento de Estudios Multidisciplinarios, Guanajuato, México.

\*Autor para correspondencia: [ja.bautista@ugto.mx](mailto:ja.bautista@ugto.mx)

### RESUMEN

La investigación realizada es de tipo correlacional y estudió la influencia del precio del barril del crudo WTI, el bushel de Maíz y la tonelada de harina de Soya como variables independientes, sobre el precio de la libra de carne de cerdo mexicana, con el objetivo de hallar una función que explique dichas variaciones. Se utilizaron datos promedios mensuales de los precios, recolectados en un periodo de 10 años desde el 01/10/2012 hasta el 01/09/2022. En los datos hay una clara tendencia lineal entre el precio de la carne de cerdo con el precio del barril de crudo WTI, y el precio de la tonelada de harina de soya, más no tan clara o evidente con el precio del bushel de maíz, siendo esta variable excluida del modelo final por ser estadísticamente no significativa (Sig. 0,184). Las variables independientes del modelo final son estadísticamente significativas (Sig. 0,000), con valores de *t*-student de 4,999 para el crudo WTI y 3,697 para la harina de soya y no existen problemas de colinealidad entre ellas. El modelo obtenido es de regresión lineal múltiple, y tiene como predictores del precio de la carne de cerdo : el precio del barril de crudo WTI y el precio de la tonelada de harina de soya . Pronostica que el precio de la carne de cerdo no puede descender de 15,50 centavos de USD por libra, y puede explicar las variaciones de esta en un 61,4 %. Los residuos estandarizados del modelo presentan una distribución normal, corroborado mediante una prueba de Kolmogorov-Smirnov de 0,071, existiendo un par valores extremos positivos, que pueden informar sobre las circunstancias de las variables para el interés del investigador en los meses de mayo y junio del año 2021.

**Palabras clave:** Correlación lineal; regresión múltiple; carne de cerdo; crudo WTI; cereales; pronóstico

### ABSTRACT

The research conducted is correlational and studied the influence of the price of a barrel of WTI crude oil, a bushel of Corn and a ton of Soybean meal as independent variables, on the price of a pound of Mexican pork, with the objective of finding a function that explains these variations. Monthly average price data, collected over a 10-year period, from 01/10/2012 to 01/09/2022, were used. In the data there is a clear linear trend between the price of pork with the price of a barrel of WTI crude oil, and the price of a ton of soybean meal, but not so clear or evident with the price of a bushel of corn, being this variable excluded from the final model for being statistically not significant (Sig. 0.184). The independent variables of the final model are statistically significant (Sig. 0.000), with *t*-student values of 4.999 for WTI crude oil and 3.697 for soybean meal, and there are no collinearity problems between them. The model obtained is multiple linear regression, and has as predictors of pork price: the price of a barrel of WTI crude oil and the price of a ton of soybean meal . It predicts that the price of pork cannot fall below 15.50 US cents per pound, and can explain the variations of pork by 61.4%. The standardized residuals of the model present a normal distribution, corroborated by a Kolmogorov-Smirnov test of 0.071, there being a pair of positive extreme values, which can inform about the circumstances of the variables for the researcher's interest in the months of May and June 2021.

**Key words:** Linear correlation; multiple regression; pork meat; WTI crude; cereals; forecast

## INTRODUCCIÓN

De acuerdo a Chuluunsaikan y col. [1], la ganadería es uno de los principales subsectores de la agricultura. La Organización para la Agricultura y la Alimentación de las Naciones Unidas (FAO) afirma que la ganadería juega un papel económico importante para alrededor del 60 % de los hogares rurales en los países en desarrollo. Además de los agricultores y los gobiernos, los consumidores también prestan atención al mercado de productos básicos de ganado, ya que estos productos están entre los más consumidos. Sin duda, la ganadería constituye una pieza relevante de la economía.

El objetivo de los sistemas de producción de ganado porcino (*Sus scrofa domestica*) es la necesidad que se tiene de abastecer la demanda de alimentos (carne) y de nutrientes de alto valor (proteína, energía, vitaminas, entre otros) para muchos habitantes en el mundo; en este sentido, las estadísticas de la Organización de las Naciones Unidas (ONU) ponen de manifiesto la evolución creciente de la población humana mundial que se espera llegue a superar los 9 mil millones en el año 2037 [2].

El costo de producción de carne de cerdo se encuentra íntimamente relacionado con algunos parámetros productivos que determinan su eficiencia, el porcentaje de los parámetros son aproximadamente de 59,4 % en la alimentación, 2,4 % en la mano de obra, 1,16 en la genética, 0,82 % de sanidad, 1,62 % en amortizaciones, 0,94 % en energía, 0,84 % en fletes y guías, y otros varios en 0,25 % [3]. Siendo el de mayor porcentaje y más significativo el de la alimentación.

En la alimentación de los cerdos existe una gran variedad de ingredientes que pueden utilizarse en la formulación de una dieta. El nivel de uso de estos ingredientes en la ración, estará determinado por la composición nutricional del producto, de las restricciones nutricionales que tenga para las diferentes etapas productivas y del requerimiento de nutrimentos que se quiera satisfacer. El maíz (*Zea mays*), es la principal fuente de energía utilizada en la alimentación porcina y la fuente de proteína de origen vegetal incluye principalmente a la harina de soya (*Glycine max*) [4]. La harina de soya es la única fuente disponible de proteína sin problemas para utilizarse en la alimentación de los cerdos, excepto en la alimentación de lechones recién destetados, donde ocurre una reacción antígeno - anticuerpo producida por las proteínas de origen vegetal [4]. Por estas razones, el maíz y la harina fueron tomados en cuenta para el análisis de este modelo de predicción, sin embargo, cabe resaltar la fuerte influencia de los crudos por su dependencia en el mundo y que afectan las economías en todos los rubros.

En cuanto a los precios que adquiere el petróleo están íntimamente influenciados por la demanda y la oferta que se tiene mayormente de economías que se encuentran en crecimiento. Según Henke [5], encontró que el precio del petróleo puede verse afectado por eventos que generen la interrupción o disminución de suministro de petróleo y sus derivados en el mercado, incluyendo los eventos geopolíticos y relacionados con el clima, los cuales pueden afectar también la demanda, creando incertidumbre sobre el futuro de los precios.

De acuerdo a Zhang y col. [6,7], los pronósticos confiables y precisos del precio del petróleo crudo son de gran interés para una amplia gama de aplicaciones, sus cambios de precios y los shocks tienen impactos importantes en los mercados financieros y la economía real. Según Boubaker y col. [8], el petróleo crudo es el producto básico más comercializado del mundo por su importancia estratégica para los mercados de economía nacional. Se ha

demostrado que las fluctuaciones en los precios del crudo tienen un impacto significativo en la economía mundial, igual que otros productos básicos. Además, la volatilidad del precio del petróleo es fundamental para fijación de precios de activos, asignación de activos y gestión de riesgos [9].

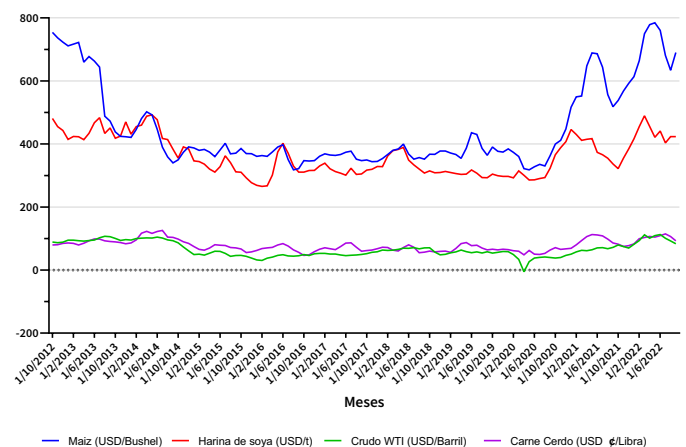
La influencia del precio de los crudos en sectores como el transporte de materias primas hacen que indirectamente impacte sobre las variaciones en los precios de la libra de carne de cerdo. El objetivo final de esta investigación fue determinar la función o modelo que sea capaz de explicar las variaciones en los precios de la carne de cerdo, en relación con las variaciones en los precios del crudo WTI y el precio de los cereales maíz y soya.

Para conocer la relación entre ambas partes se procesaron los datos obtenidos en la web *Investing.com*, con una frecuencia de precios promedios mensuales, un análisis de regresión [10], que por tratarse de dos tipos de crudos distintos es necesario observar los valores estadísticos de regresiones lineales y la existencia o no de colinealidad [11].

## MATERIALES Y MÉTODOS

Para el desarrollo de esta investigación se tomó la información disponible y descargable en la página web de *Investing.com* en su portal de México, donde los [precios](#) del crudo West Texas Intermediate (WTI: petróleo de referencia en los Estados Unidos de Norteamérica), están expresados en USD-Barril<sup>-1</sup>, los [cereales](#) seleccionados son el precio del maíz que esta expresado en USD/Bushel, y el precio de la harina de soya expresada en USD-t<sup>-1</sup> y los datos de precios de la [carne de cerdo](#) expresada cts USD-Libra<sup>-1</sup>. Se descargaron los precios máximos y mínimos con una frecuencia mensual y de un lapso de 10 años, específicamente desde el 01/10/2012 hasta el 01/09/2022, con los datos descargados se calculó un promedio de cada mes, para un total de 120 meses de estudio.

Con los datos aportados en *Investing.com* se elaboró un gráfico de secuencia (FIG. 1), donde se muestran los valores del precio del crudo WTI, los cereales y la carne de cerdo, evidenciándose la tendencia de correlación que puede existir entre las variables, es por esta razón que se analizarán más detalladamente las tres variables y su relación [12].



**FIGURA 1. Precio del crudo WTI, los cereales (maíz y harina de soya) y la carne de cerdo en el periodo de los 10 años de estudio. Fuente: <https://mx.investing.com/commodities/meats>**

### Metodología aplicada

Se utiliza como modelo cuantitativo de pronósticos un análisis mediante el método de los mínimos cuadrados para identificar la relación entre las variables. Para el procesamiento de todos los datos y la generación de los diferentes informes estadísticos y figuras se utilizó el software IBM SPSS Stactics en su versión 25.0.0. Esta investigación fue de tipo correlacional, dado que se buscaban explicaciones mediante el estudio de la relación que pueda existir entre dos o más variables que deben medirse cuantitativamente y brinda información sobre en qué medida un cambio en una variable dependiente es debido a la modificación experimentada en otra u otras variables independientes [13].

### Regresión múltiple

De acuerdo con Huang y Aldeeb [14], el análisis de regresión es esencialmente un proceso matemático. Describe cuantitativamente la correlación entre variables a través de expresiones matemáticas. Para describir el modelo de regresión múltiple se debe plantear que en una muestra de  $n$  elementos cada elemento se ha medido y se ha evaluado las  $k$  variables independientes ( $x_i, x_{1i}, \dots, x_{ki}$ ) a una variable dependiente  $y$ . Por lo que se puede expresar de la siguiente forma:  $y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki} + \epsilon_i$ . La cantidad  $\epsilon_i$  en la ecuación es una variable aleatoria, que se supone está normalmente distribuida con  $E(\epsilon_i) = 0$  y  $V(\epsilon_i) = \sigma^2$ .

Para el caso de una regresión múltiple de dos variables predictoras, se estableció la siguiente ecuación:  $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$ . Donde  $x_{1i}$  y  $x_{2i}$  son los datos en la ubicación  $i$  (en la base de datos recolectada) de las variables independientes de regresión. La  $y_i$  representa el dato de la variable dependiente del modelo que se genera al sustituir los valores de  $x_{1i}$  y  $x_{2i}$ . Los parámetros  $\beta_0$ ,  $\beta_1$  y  $\beta_2$  son desconocidos, pero mantienen una función lineal y deben ser estimados en el análisis. Y  $\epsilon_i$  es un error aleatorio que se puede generar en la determinación de  $y_i$ . La generalización apropiada de esta ecuación a un modelo probabilístico supone que el valor esperado de  $Y_i$  es una función lineal de  $x_{1i}$ , entonces, se puede representar el modelo probabilístico de  $Y_i$  de la siguiente forma para las dos variables predictoras:  $Y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + e_i$ . Siendo los coeficientes  $b_0$ ,  $b_1$  y  $b_2$  los parámetros estimados de  $\beta_0$ ,  $\beta_1$  y  $\beta_2$  respectivamente y  $e_i$  el error residual que se produce en la estimación. El análisis de regresión tiene ciertos requisitos previos. Se requiere que las variables explicativas no deben tener relaciones de multicolinealidad [11]. Por lo tanto, se debe diagnosticar la multicolinealidad del modelo. Se debe asegurar que la variable independiente tenga un impacto en la variable dependiente y tener una estrecha correlación para que el modelo logre el efecto deseado [15].

### Pruebas de hipótesis y medidas de adecuación en la regresión

**Prueba de la significancia de la regresión:** ésta es una prueba que determina si existe una relación lineal entre la variable respuesta  $Y_i$  y un subconjunto de las variables de regresión  $x_{1i}, x_{2i}, \dots, x_{ki}$ . Las hipótesis apropiadas son:  $H_0: b_1 = b_2 = \dots = b_k = 0$  y  $H_1: b_j \neq 0$  al menos para una  $j$ . El rechazo de:  $H_0: b_1 = b_2 = \dots = b_k = 0$  implica que al menos una de las variables de regresión  $x_{1i}, x_{2i}, \dots, x_{ki}$  tiene una contribución significativa en el modelo. Este rechazo de  $H_0$  se da si el valor calculado del estadístico de prueba  $f_0 > f_{\alpha, k, n-p}$ . El estadístico de prueba se calcula mediante la expresión:  $F = \frac{\frac{SSR}{k}}{\frac{SSE}{n-p}} = \frac{MSR}{MSE}$ .

Este valor se ve reflejado en las tablas de análisis de la varianza [16].

**Pruebas sobre los coeficientes individuales de regresión y sobre el subconjunto de coeficientes:** esta prueba es útil para determinar el valor potencial de cada una de las variables de regresión del modelo de regresión que se está obteniendo. Las hipótesis para la prueba de la significancia de cualquier coeficiente de regresión individual,  $b_j$ , son:  $H_0: b_j = 0$  y  $H_1: b_j \neq 0$ . Si no se rechaza  $H_0: b_j = 0$ , entonces esto indica que el regresor  $x_{kj}$  puede eliminarse del modelo. El estadístico de prueba para esta hipótesis es:  $T = \frac{b_j}{\sqrt{\sigma^2 C_{jj}}}$

La hipótesis nula  $H_0: b_j = 0$ , se rechaza si  $|t| > t_{\alpha/2}$ . A esto se le conoce como prueba parcial o marginal, porque el coeficiente de regresión  $b_j$ , depende de todas las demás variables de regresión ( $i \neq j$ ) que están en el modelo.

**Coefficiente de determinación múltiple:** este coeficiente está definido por:  $R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$ . Es una medida de la magnitud de la reducción en la variabilidad de  $y_i$  obtenida mediante el empleo de las variables de regresión  $x_{1i}, x_{2i}, \dots, x_{ki}$ . Este coeficiente adquiere valores entre 0 y 1 ( $0 \leq R^2 \leq 1$ ).

**Coefficiente de correlación múltiple de Pearson:** es una medida de la asociación lineal que existe entre  $y_i$  y las variables de regresión  $x_{1i}, x_{2i}, \dots, x_{ki}$ . Su cálculo proviene de la raíz cuadrada positiva de  $R^2$  [17].

### Pruebas de multicolinealidad

**Coefficiente de inflación de la varianza (VIF):** cuantifica la severidad de la multicolinealidad en un análisis de regresión. Si el  $VIF = 1$ , las variables no están correlacionadas, si  $VIF$  tiene valor entre 1 y 5, las variables están moderadamente correlacionadas y si el  $VIF > 5$ , las variables están altamente correlacionadas.

**Índice de condición de colinealidad:** se calculan como las raíces cuadradas de las razones del mayor autovalor con respecto a cada autovalor sucesivo. Los valores  $>15$  indican un posible problema con la colinealidad;  $>30$ , un problema grave [17].

## RESULTADOS Y DISCUSIÓN

**Prueba de normalidad:** Se realizó la prueba de Kolmogorov-Smirnov para conocer si los datos evaluados se rigen por una distribución normal, los resultados se muestran en la TABLA I, quedando demostrado que, tanto los datos dependientes como independientes tienen un comportamiento normal, por lo que se pueden aplicar pruebas estadísticas de tipo paramétrico.

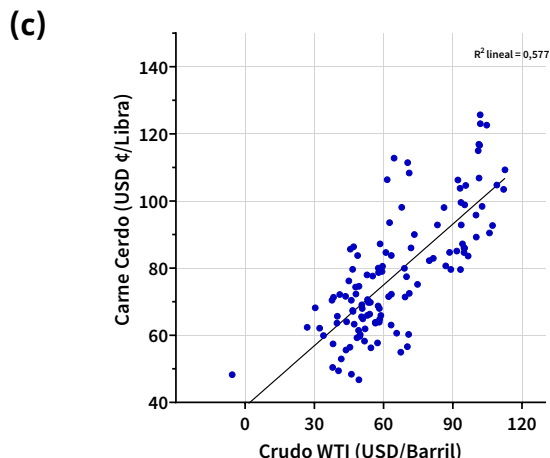
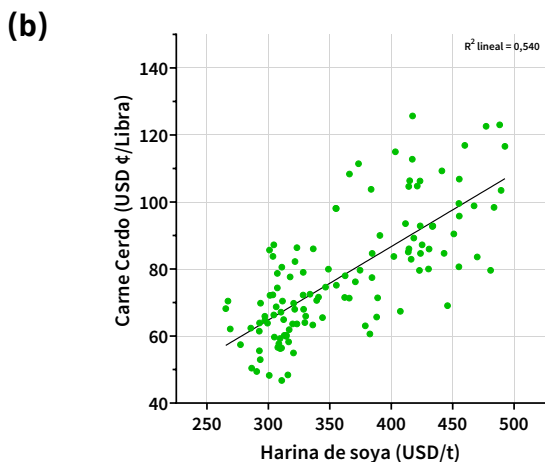
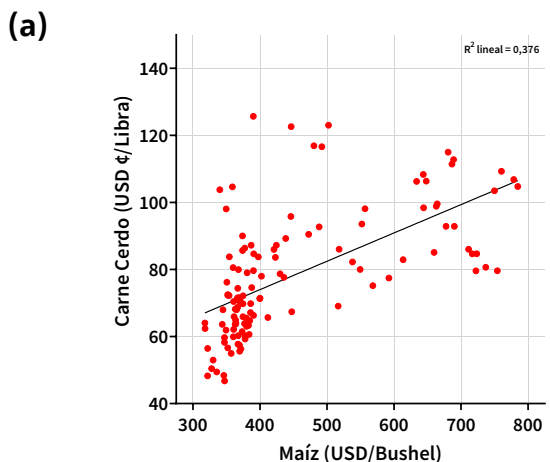
**TABLA I**  
Prueba de Kolmogórov-Smirnov para los datos de la investigación

	Carne Cerdo (USD c/Libra)	Maíz (USD/Bushel)	Harina de soya (USD/t)	Crudo WTI (USD/Barril)
N	120	120	120	120
Estadístico de prueba	0,109	0,261	0,154	0,128
Sig. asintótica (bilateral)	0,001 <sup>c</sup>	0,000 <sup>c</sup>	0,000 <sup>c</sup>	0,000 <sup>c</sup>

- La distribución de prueba es normal.
- Se calcula a partir de datos.
- Corrección de significación de Lilliefors

**CORRELACIÓN LINEAL ENTRE VARIABLES**

Los datos del crudo WTI y de los cereales, con la carne de cerdo se pueden representar en figuras de dispersión de puntos, observando así alguna tendencia lineal entre ellos, esto se representa en la FIG. 2, donde en el apartado (a) muestra una correlación de los datos con entre la carne de cerdo y el Maíz.



**FIGURA 2.** Diagramas de dispersión de la carne de cerdo con las variables independientes. Fuente: propia

Mientras que en el apartado (b) se visualiza una correlación entre la carne de cerdo y la harina de soya con y en el apartado (c) una correlación entre la carne de cerdo y el crudo WTI con . En estos diagramas se evidencia que existe una baja tendencia lineal entre la carne de cerdo y el precio del maíz para el periodo analizado, lo que contradice hallazgos de otros autores, como Ahumada y Cornejo [18] quienes aseguran que la carne muestra una fuerte correlación tanto con la soya como con el maíz. Esta discrepancia puede deberse a las diferencias en la extensión del periodo analizado y a la frecuencia utilizada, siendo trimestral y para 6 años en el caso de los autores, y mensual durante 10 años en el caso de esta investigación. Para conocer la correlación de las variables entre sí, se tiene una matriz de correlación de Pearson, como se presenta en la TABLA II.

La TABLA II muestra la correlación de Pearson de la variable dependiente (carne de cerdo) con cada variable independiente, siendo de 0,613 con el maíz, de 0,735 con la harina de soya y 0,759 con respecto al crudo WTI. Pero también es de notar el alto coeficiente de correlación de Pearson que existe entre las variables independientes de harina de soya y de crudo WTI de 0,802; lo cual indica que puede existir un problema de colinealidad para un modelo de regresión múltiple entre las variables [17]. Esto sería respaldado por los resultados de [12] que han demostrado la fuerte relación que existe entre el precio del petróleo y algunos productos agrícolas, incluyendo la soya y el maíz, lo que se reforzaría en el caso de la predicción de precios de la carne de cerdo ya que dependería tanto del precio del petróleo como del precio de la soya.

**TABLA II**  
**Matriz de correlación de Pearson entre las variables de estudio**

	Maíz (USD/Bushel)	Harina de soya (USD/t)	Carne Cerdo (USD c/Libra)	Crudo WTI (USD/Barril)
Maíz (USD/Bushel)	1	0,705	0,613	0,655
Harina de soya (USD/t)	0,705	1	0,735	0,802
Carne Cerdo (USD c/Libra)	0,613	0,735	1	0,759
Crudo WTI (USD/Barril)	0,655	0,802	0,759	1

**REGRESIÓN LINEAL MÚLTIPLE**

Para estudiar el comportamiento de las variables independientes con respecto a sus aportes en el modelo de regresión, se utilizó la herramienta informática IBM SPSS en su opción de “Regresión por pasos”, la cual arroja una comparación entre un modelo con todas las variables involucradas y otros modelos donde el software toma decisiones de exclusión. Al realizar estas corridas de regresión se obtuvieron los resultados expresados en la TABLA III.

El software elaboró dos modelos de regresión: el modelo 1 donde toma como variable predictora solo al precio del crudo WTI, y un modelo 2 donde las variables predictoras son: el crudo WTI y la harina de soya. Los modelos de regresión son una sólida herramienta para el pronóstico de precios en diferentes contextos [14], entre los que se destacan las proyecciones económicas [12] e incluso geográficas [10], lo que refuerza los resultados obtenidos en el presente análisis. En ambos modelos es excluida la variable sobre el precio del maíz. En la TABLA IV se puede observar la razón de la exclusión del precio del Bushel de maíz como variable significativa

**TABLA III**  
Resumen de los modelos de regresión lineal por pasos

Modelo <sup>c</sup>	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Estadísticos de cambio				Durbin-Watson
					Cambio en R cuadrado	Cambio en F	gl1	gl2	
1	0,759 <sup>a</sup>	0,577	0,573	11,99446	0,577	160,698	1	118	0,000
2	0,788 <sup>b</sup>	0,621	0,614	11,39812	0,044	13,670	1	117	0,000

- a. Predictores: (Constante), Crudo WTI (USD/Barril)
- b. Predictores: (Constante), Crudo WTI (USD/Barril), Harina de soya (USD/t)
- c. Variable dependiente: Carne Cerdo (USD €/Libra)

del modelo definitivo, donde en ambos modelos la significancia de esta variable es mayor a 0,05. Es decir, su aporte a los modelos no era significativo estadísticamente, por lo que nuevamente el modelo difiere de lo expuesto por otros autores [18] que consideran relevante la relación entre precio del maíz y de la carne de cerdo, discrepancia que nuevamente se considera debida a las diferencias en cantidad y frecuencia de los datos utilizados entre ambos estudios.

**TABLA IV**  
Variables excluidas<sup>a</sup> en la regresión lineal por pasos

Modelo	En beta	t	Sig.	Correlación parcial	Estadísticas de colinealidad Tolerancia
1 Maíz (USD/Bushel)	0,203 <sup>b</sup>	2,623	0,010	0,236	0,570
1 Harina de soya (USD/t)	0,352 <sup>b</sup>	3,697	0,000	0,323	0,357
2 Maíz (USD/Bushel)	0,109 <sup>c</sup>	1,337	0,184	0,123	0,480

- a. Variable dependiente: Carne Cerdo (USD €/Libra)
- b. Predictores en el modelo: (Constante), Crudo WTI (USD/Barril)
- c. Predictores en el modelo: (Constante), Crudo WTI (USD/Barril), Harina de soya (USD/t)

En este resumen, el valor de  $R^2_{ajustado}$  es ligeramente mayor para el modelo 2 con 0,614 en comparación al modelo 1 con 0,573. Los resultados del análisis de la varianza (ANOVA) para cada modelo está indicado en la TABLA V, donde se puede observar que ambos modelos se adecuan bien para ser utilizados en la relación de estas variables, ambos modelos tienen una significancia 0,000 y valores del estadístico es alto; siendo para el modelo 1 de mayor adecuación que el modelo 2.

**TABLA V**  
Análisis de varianza para cada modelo (ANOVA)<sup>a</sup>

Modelo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
1 Regresión	23119,126	1	23119,126	160,698	0,000 <sup>b</sup>
1 Residuo	16976,324	118	143,867		
1 Total	40095,450	119			
2 Regresión	24895,144	2	12447,572	95,812	0,000 <sup>c</sup>
2 Residuo	15200,305	117	129,917		
2 Total	40095,450	119			

- a. Variable dependiente: Carne Cerdo (USD €/Libra)
- b. Predictores: (Constante), Crudo WTI (USD/Barril)
- c. Predictores: (Constante), Crudo WTI (USD/Barril), Harina de soya (USD/t)

La TABLA VI señala el valor de los coeficientes para las variables de cada modelo y los coeficientes *Beta* y *t* que indican las significancias de cada variable, así como, el estadístico *VIF* (Variance Inflation Factor) o factor de inflación de la varianza para detectar los problemas de colinealidad entre las variables predictoras.

En la TABLA VI se muestran como en cada modelo, las variables aportan al modelo con valores de sig. 0,000 en casi todos los casos. El Crudo WTI en el modelo 1 es el mayor aporte estadístico da al modelo con un coeficiente estandarizado Beta de 0,759, sin embargo, los valores presentados en el modelo 2 también son significativos y no poseen ningún problema de colinealidad detectado por su *VIF* de 2,805 un valor aceptable de inflación de la varianza ya que se encuentra por debajo de 10 [19].

**TABLA VI**  
Coeficientes<sup>a</sup> de las variables de cada modelo

Modelo	Coeficientes no estandarizados		Coeficientes estandarizados	T	Sig.	Estadísticas de colinealidad	
	B	Desv. Error	Beta			Tolerancia	VIF
1 (Constante)	38,762	3,310		11,712	0,000		
1 Crudo WTI (USD/Barril)	0,604	0,048	0,759	12,677	0,000	1,000	1,000
2 (Constante)	15,479	7,039		2,199	0,030		
2 Crudo WTI (USD/Barril)	0,379	0,076	0,477	4,999	0,000	0,357	2,805
2 Harina de soya (USD/t)	0,105	0,028	0,352	3,697	0,000	0,357	2,805

- a. Variable dependiente: Carne Cerdo (USD €/Libra)

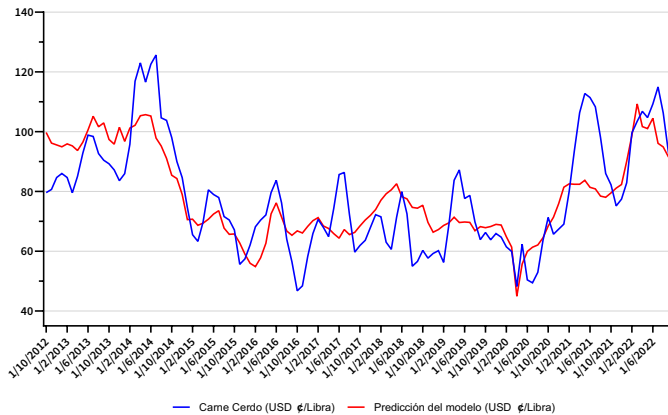
**Modelo de regresión lineal**

Para el modelo de regresión final se tiene como variables predictoras el precio del crudo WTI y el precio de la tonelada de harina de soya, y como variable dependiente, el precio de la carne de cerdo. Por lo tanto, la ecuación del modelo viene dada por la expresión:  $Y_i = b_0 + b_1x_{1i} + b_2x_{2i}$ . Donde  $Y_i$  es el valor estimado del precio de la carne de cerdo en la posición *i* de la base de datos. El coeficiente  $b_0$  es la constante del modelo,  $b_1$  es el coeficiente estimado de la variable predictor "Crudo WTI" y  $x_{1i}$  es el valor del precio del Crudo WTI para la posición *i*, mientras que  $b_2$  es el coeficiente estimado de la variable predictor "Harina de soya" y  $x_{2i}$  es el valor del precio de la tonelada de harina de soya para la posición *i* de la base de datos. Si se hace que  $y = y_{cerdo}$ ;  $x_1 = x_{WTI}$  (Precio del crudo);  $x_2 = x_{HS}$  (Precio de la harina de soya); quedando la ecuación del modelo como:  $y_{cerdo} = b_0 + b_1x_{WTI} + b_2x_{HS}$ . Los coeficientes  $b_0$ ,  $b_1$  y  $b_2$  se obtienen de la TABLA VI y son  $b_0 = 15,479$ ,



$b_1 = 0,379$  y  $b_2 = 0,105$ , entonces la ecuación del modelo de regresión finalmente es:  $y_{cerdo} = 15,479 + 0,379x_{WTI} + 0,105x_{HS}$ .

Con el modelo de regresión  $y_{cerdo}$  se pueden figurar los datos reales históricos recolectados al inicio de la investigación. Los resultados de los valores pronosticados debido a la serie inicial de precios del crudo WTI se presenta en la FIG. 3.



**FIGURA 3.** Distribución de los precios de la carne de cerdo registrada en los 10 años de muestra y la predicción de los precios en el mismo periodo. Fuente: propia

En la FIG. 3 se presenta en color rojo los resultados pronósticos del modelo  $y_{cerdo}$  y en línea azul los valores reales históricos de los precios de la libra de carne de cerdo. Se puede ver reflejado en la Figura que ambos valores mantienen una tendencia general, aunque los datos reales de la carne de cerdo tienen varios picos y valles en su curva. El software también proporciona información, en relación con los límites de confianza de la media y de los datos individuales, que se presentan en la TABLA VII

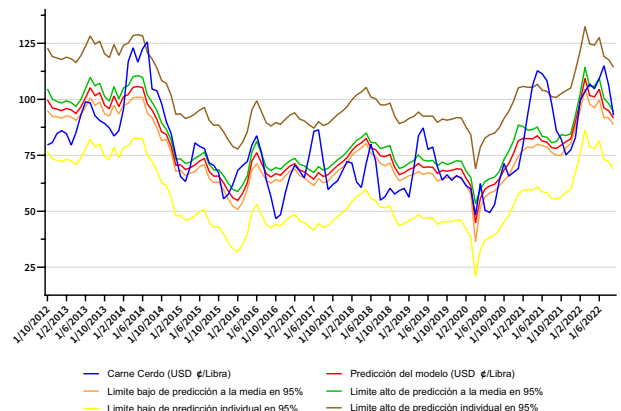
**TABLA VII**  
**Coefficientes<sup>a</sup> del modelo de regresión con límites de gráfica confianza**

Modelo	Coefficients no estandarizados		Coefficients estandarizados	T	Sig.	95% Intervalo de confianza para B	
	B	Desv. Error	Beta			Límite inferior	Límite superior
(Constante)	15,479	7,039		2,199	0,030	1,539	29,419
2 Crudo WTI (USD/Barril)	0,379	0,076	0,477	4,999	0,000	0,229	0,529
Harina de soya (USD/t)	0,105	0,028	0,352	3,697	0,000	0,049	0,161

a. Variable dependiente: Carne Cerdo (USD ¢/Libra)

En base a estos valores también se puede representar la secuencia de los valores pronósticos y sus límites de predicción para la media de los datos y para cada dato individual, como se representa en la FIG. 4.

En esta figura se observa como los límites de predicción de los datos generan una franja de probabilidad de la predicción con un nivel de confianza de 95 %, y los datos históricos fueron en su mayoría



**FIGURA 4.** Distribución de los valores de predicción del modelo, mostrando los límites de predicción para los datos en los 10 años de muestra. Fuente: propia

bien pronosticados, a excepción de un pico en el mes de mayo de 2021, donde el precio de la carne de cerdo subió por encima de lo esperado por este modelo [16]. Es de hacer notar en este caso que la predicción de precios del petróleo es un tema complejo [20, 21], que está lejos de considerarse resuelto y por tanto la proyección de precios de artículos que dependen de su precio también lo es. Sin embargo, diferentes metodologías se utilizan para aproximar el precio del petróleo, con diversos niveles de éxito. Esto mismo motiva el interés de aplicar las técnicas para la proyección del precio de otros productos, como en el caso de esta investigación. Ya que el modelo se ha ajustado en gran medida, salvo el pico extremo ya comentado, se cree que el mismo es una alternativa viable a lo presentado por otras metodologías como la de Wang y Sun [22] o Chuluunsaikan y col. [1].

**Análisis de los residuos del modelo**

Los residuos del modelo planteado sean estandarizados o no, fueron analizados para conocer su distribución, si cumple con las condiciones de normalidad y precisar cuáles son los valores extremos que se alejan un poco de los límites de confianza en el pronóstico. En la prueba de Kolmogorov-Smirnov para los residuos estandarizados, se pudo apreciar la distribución normal de los mismos, como se muestra en la TABLA VIII.

Con valor de 0,071 de significancia en la prueba Kolmogorov-Smirnov se puede concluir que, la distribución que tienen los residuos es Normal. La forma (a) de la FIG. 5 muestra el histograma de frecuencia de los residuos estandarizados, donde se observan las barras formando una distribución normal y en la parte (b), el diagrama de caja indica el valor más extremo que se tiene en los residuos estandarizados. En base a los resultados de la TABLA IX, los puntos que presentan valores extremos altos son cinco, todos por encima

**TABLA VIII**  
**Prueba de normalidad residuos del modelo**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estadístico	gl	Sig.	Estadístico	gl	Sig.
Residuos estandarizados	0,078	120	0,071	0,973	120	0,016

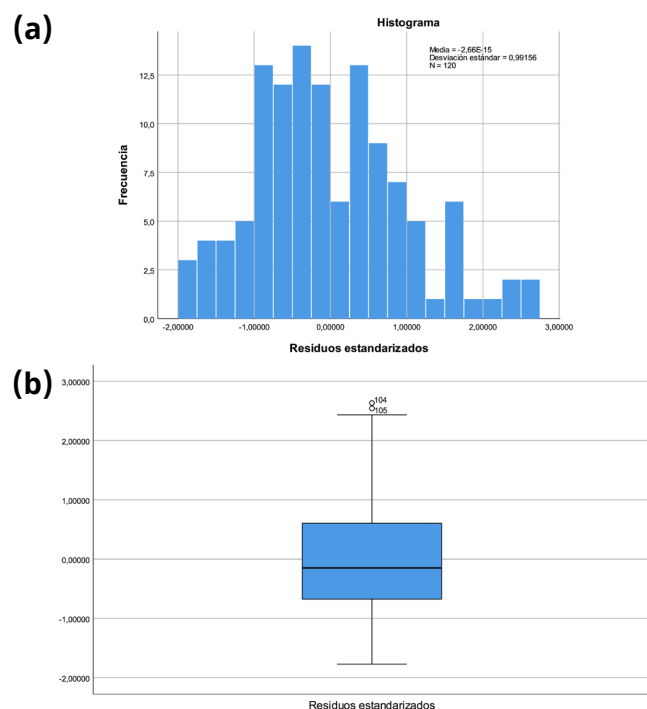
a. Corrección de significación de Lilliefors

de 2, y los valores extremos bajos son 5 todos por debajo de -1. Los puntos más extremos se ubican en las posiciones 104 y 105 de la base de datos, con un valor de 2,54307 y 2,63376 respectivamente, valores que coinciden con el diagrama de caja de la FIG. 5.

### CONCLUSIONES

Los resultados obtenidos en esta investigación demuestran que existe una clara tendencia lineal entre el precio de la carne de cerdo con el precio del crudo WTI y el precio de la harina de soja (materia prima importante en la preparación del alimento del cerdo). Mientras que esta tendencia lineal no está clara con el precio del bushel de maíz, quedando esta variable excluida del modelo final de regresión, dado que no tiene un aporte estadísticamente significativo en el modelo. El modelo obtenido según los datos analizados es de regresión lineal múltiple, y toma en cuenta como predictores del precio de la carne de cerdo: el precio del barril de crudo WTI y el precio de la tonelada de harina de soja en México. Pronostica que el precio de la carne de cerdo no puede descender de 15,50 centavos de USD por libra, y sostiene una relación más marcada con el barril del crudo WTI que con la harina de soja. Este modelo puede explicar los cambios del precio en la carne de cerdo en un 61,4 % dado su . Los residuos estandarizados del modelo presentan una distribución normal, existiendo un par valores extremos positivos, que pueden informar sobre las circunstancias de las variables para el interés del investigador en los meses de mayo y junio del año 2021.

Finalmente, esta información puede permitir una proyección del comportamiento del precio del cerdo a partir de datos del petróleo y la soja, sobre los que es posible tener mejores datos históricos y una previsión más certera. De esta manera se podrían tomar mejores decisiones a futuro sobre la producción.



**FIGURA 5. Histograma de frecuencia y diagrama de caja de los residuos estandarizados. Fuente: propia.**

Para la identificación de los datos extremos en el modelo se contó con los valores extremos de los residuos estandarizados que se presentan en la TABLA IX.

**TABLA IX**  
**Valores extremos en los residuos estandarizados**

		Número del caso	Valor	
Residuos estandarizados	Mayor	1	105	2,63376
		2	104	2,54307
		3	22	2,43354
		4	106	2,40875
		5	103	2,09950
Residuos estandarizados	Menor	1	1	-1,77221
		2	49	-1,76262
		3	67	-1,75188
		4	71	-1,73341
		5	15	-1,56967

### REFERENCIAS BIBLIOGRÁFICAS

- [1] Chuluunsaikan T, Ryu GA, Yoo KH, Rah H, Nasridinov A. Incorporating Deep Learning and News Topic Modeling for Forecasting Pork Prices: The Case of South Korea. *Agricult. [Internet]*. 2020; 10:513. doi: <https://doi.org/kmzp>
- [2] Parant, A. World population: Trends and prospects. *Futuribles. [Internet]*. 2023; 452:63-85. doi: <https://doi.org/ktx9>
- [3] El Sitio Porcino. Sector porcino: ¿Qué pasa con los costos de producción? [Internet] Charlbury (GBR): Global Ag Media. 2018 [consultado 20 Ene 2023]; Disponible en: <https://bit.ly/30FN3wv>.
- [4] Campabadal C. Guía técnica para la alimentación de cerdos. Costa Rica: Imprenta Nacional. 2009; p. 7-23.
- [5] Henke A. La relación entre el precio del petróleo WTI y los índices bursátiles S&P 500 y el S&P 500 del sector energético. [tesis de maestría en Internet]. Argentina: Universidad de San Andrés; 2017 [consultado 18 Feb 2023]. 32 p. Disponible en: <https://bit.ly/45dCpke>.
- [6] Zhang Y, Ma F, Wang Y. Forecasting crude oil prices with a large set of predictors: Can LASSO select powerful predictors? *J. Empir. Finance. [Internet]*. 2019; 54:97-117. doi: <https://doi.org/gmmqqk>
- [7] Zhang D, Li Q, Mugeraw AW, Ling L. A hybrid model considering cointegration for interval-valued pork price forecasting in China. *J. Forecasting. [Internet]*. 2020; 39:1324-1341. doi: <https://doi.org/ghhm6n>
- [8] Boubaker S, Liu Z, Zhang Y. Forecasting oil commodity spot price in a data-rich environment. *Ann. Oper. Res. [Internet]*. 2022; 1-10. doi: <https://doi.org/km2g>

- [9] Zhang Y, Wei Y, Zhang Y, Jin D. Forecasting oil price volatility: Forecast combination versus shrinkage method. *Energy Econom.* [Internet]. 2019; 80:423–433. doi: <https://doi.org/gjgdww>
- [10] Evelina D, Barca E, Mikosz R, de Araujo H, Berardi L, Passarella G. Linear and evolutionary polynomial regression models to forecast coastal dynamics: Comparison and reliability assessment. *Geomorph.* [Internet]. 2018; 300:128–140. doi: <https://doi.org/gctbdn>
- [11] Goyal L, Mittal M, Sethi J, Hemanth DJ. Monitoring the Impact of Economic Crisis on Crime in India Using Machine Learning. *Comput. Econom.* [Internet]. 2019; 53:1467–1485. doi: <https://doi.org/gmq32s>
- [12] Hung NT. Oil prices and agricultural commodity markets: Evidence from pre and during COVID-19 outbreak. *Resour. Policy.* [Internet]. 2021; 73(102236):1–10. doi: <https://doi.org/gk7t43>
- [13] Bianco V, Manca O, Nardini S. Linear Regression Models to Forecast Electricity Consumption in Italy. *Energy Sourc.* [Internet]. 2012; 8(Pt B):86–93. doi: <https://doi.org/gf8ntm>
- [14] Huang J, Aldeeb H. Mathematical Modeling and Forecasting of Economic Variables Based on Linear Regression Statistics. *Appl. Mathemat. Nonlinear.* [Internet]. 2022; 1–7. doi: <https://doi.org/km2k>
- [15] Silva D, de Souza Y, Antonini T, Valadares H, de Nobrega M. Unorganized machines and linear multivariate regression model applied to atmospheric pollutant forecasting. *Acta Scientif. Technol.* [Internet]. 2020; 42(48203):1–11. doi: <https://doi.org/km2m>
- [16] Walpole RE, Myers RH, Myers SL, Ye K. *Probabilidad y estadística para ingeniería y ciencias.* México. Pearson. 2012; p. 319–506.
- [17] Navidi W. *Estadística para ingenieros y científicos.* México: McGraw Hill. 2006; p. 368–527.
- [18] Ahumada H, Cornejo M. Forecasting food prices: The case of corn, soybeans and wheat. *International Journal of Forecasting.* 2016; 32(3):838–848. doi: <https://doi.org/gn8635>
- [19] Hayrapetyan HN, Hambaryan MA. Empirical Specification of Factors Affecting Per Capita Pork Consumption in Armenia. *AgriScience and Technology.* 2020; 70(2):35–42.
- [20] Wen F, Zhao Y, Zhang M, Hu C. Forecasting realized volatility of crude oil futures with equity market uncertainty. *Appl. Econ.* [Internet]. 2019; 51(59):6411–6427. doi: <https://doi.org/gpcm2p>
- [21] Li X, Shang W, Wang S. Text-based crude oil price forecasting: A deep learning approach. *Int. J. Forecast.* [Internet]. 2019; 35:1548–1560. doi: <https://doi.org/gmwq4k>
- [22] Wang, C. and Sun, Z. Monthly pork price forecasting method based on Census X12-GM(1,1) combination model. *PLoS ONE.* [Internet] 2021; 16(5):e0251436. doi: <https://doi.org/kt6d>