

<https://doi.org/10.69639/arandu.v13i2.2187>

Análisis Situacional de la actividad ganadera bovina y su incidencia en el desarrollo económico

Situational Analysis of Bovine Livestock Activity and Its Impact on Economic Development

Nelson Germán Heredia Enríquez

nelsonheredia7302@gmail.com

<https://orcid.org/0000-0001-5219-794X>

Universidad Politécnica Estatal del Carchi
Tulcán – Ecuador

Andrés Alejandro Galvis Correa

a.galvis8307@gmail.com

<https://orcid.org/0000-0001-7762-2893>

Universidad Politécnica Estatal del Carchi
Tulcán – Ecuador

Artículo recibido: 18 marzo 2026- Aceptado para publicación: 20 abril 2026
Conflictos de intereses: Ninguno que declarar.

RESUMEN


El estudio analiza la actividad ganadera bovina y su incidencia en el desarrollo económico de la parroquia Julio Andrade, en la provincia del Carchi. El objetivo es identificar los factores que influyen en la producción de leche y su relación con el desempeño económico local. La investigación adopta un enfoque cuantitativo de tipo correlacional y utiliza un modelo de regresión lineal múltiple con transformación logarítmica, estimado mediante Mínimos Cuadrados Ordinarios (MCO), a partir de una base de datos de 162 unidades productivas. Los resultados muestran que variables como el acceso a crédito, el número de vacas en producción, la prevención sanitaria y el tipo de alimentación se asocian positivamente con la producción lechera. El modelo explica aproximadamente el 90 % de la variabilidad observada y cumple con los supuestos estadísticos de normalidad, homocedasticidad y ausencia de multicolinealidad significativa. La estimación en términos de elasticidades indica que el crédito, la sanidad y la alimentación influyen de manera relevante en el incremento de la producción. El análisis con variables instrumentales no evidencia problemas de endogeneidad y presenta bajo poder explicativo, por lo que no mejora la estimación respecto al modelo MCO. En consecuencia, se establece que el modelo MCO describe adecuadamente la relación entre las variables estudiadas. El estudio concluye que la producción lechera depende de factores productivos, financieros y sanitarios, y que su fortalecimiento puede contribuir al desarrollo económico local.

Palabras clave: ganadería bovina, producción de leche, regresión lineal múltiple, desarrollo económico

ABSTRACT

This study analyzes cattle farming activity and its impact on economic development in the parish of Julio Andrade, located in Carchi province. The objective is to identify the factors that influence milk production and their relationship with local economic performance. The research adopts a quantitative, correlational approach and applies a multiple linear regression model with logarithmic transformation, estimated using Ordinary Least Squares (OLS), based on a dataset of 162 productive units. The results show that variables such as access to credit, number of productive cows, sanitary prevention, and type of feeding are positively associated with milk production. The model explains approximately 90% of the observed variability and satisfies key statistical assumptions, including normality, homoscedasticity, and absence of significant multicollinearity. Elasticity estimates indicate that credit, sanitary practices, and feeding significantly contribute to increased production. The instrumental variables analysis does not reveal endogeneity problems and shows low explanatory power, thus not improving upon the OLS estimates. Therefore, the OLS model adequately describes the relationship among the studied variables. The study concludes that milk production depends on productive, financial, and sanitary factors, and that strengthening these aspects can contribute to local economic development.

Keywords: cattle farming, milk production, multiple linear regression, economic development

Todo el contenido de la Revista Científica Internacional *Arandu UTIC* publicado en este sitio está disponible bajo licencia Creative Commons Attribution 4.0 International. 

INTRODUCCIÓN

Según la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, 2025) la actividad ganadera bovina desempeña un papel crucial en la seguridad alimentaria, contribuye a casi el 40% de la producción agrícola total en los países desarrollados y el 20% en los países en desarrollo, apoyando la nutrición y los medios de vida de al menos 1300 millones de personas en todo el mundo.

La provincia del Carchi es una región eminentemente agrícola y pecuaria, donde la producción ganadera bovina destaca como una de las principales actividades económicas. Esta actividad, de gran impacto en la dinamización económica del país, se extiende por todos los cantones de la provincia. Gracias a sus favorables condiciones climáticas y de suelo, Carchi es una de las mejores zonas para la producción de pastos de alta calidad, que sostienen el desarrollo ganadero. La ganadería, a su vez, abastece constantemente de leche a gran parte del mercado nacional, siendo este uno de los alimentos básicos en la dieta de la población.

La provincia del Carchi cuenta con 190.381 hectáreas dedicadas a la ganadería bovina, con una producción diaria de leche de 380.000 litros, de los cuales se venden 353.000 litros que representa el 92,9% y el 7,1% se destina al procesamiento en unidades de producción agropecuaria (UPA). (MAG, 2021)

La parroquia Julio Andrade, la más extensa de la provincia del Carchi, sustenta su actividad económica en la agricultura y la ganadería. La superficie destinada a la ganadería comprende 4.546,81 hectáreas de pasto cultivado, equivalentes al 40,07% del suelo disponible, y 2.702 hectáreas de pasto natural, que representan el 0,24% de la superficie total de 11.346,25 hectáreas. La producción principal corresponde a leche cruda, cuyo precio varía entre 35 y 45 centavos de dólar por litro. A pesar de su relevancia económica, el sector enfrenta limitaciones estructurales, como la insuficiencia de sistemas de riego, la carencia de asistencia técnica y las restricciones de acceso a financiamiento para la modernización de procesos productivos. Asimismo, la comercialización de ganado constituye una fuente complementaria de ingresos, apoyada por un mercado local activo. No obstante, el crecimiento de la actividad ganadera ha generado impactos ambientales adversos, derivados del manejo inadecuado de residuos sólidos, el uso de antibióticos y fertilizantes químicos en los pastizales, lo que ha provocado la contaminación de las fuentes hídricas locales, comprometiendo la disponibilidad de agua limpia para la población y los ecosistemas circundantes. (Acosta, Gómez, & Rivera, 2020)

MATERIALES Y MÉTODOS

El estudio se desarrolló en la parroquia Julio Andrade, del cantón Tulcán, provincia del Carchi, al norte del Ecuador. La investigación adoptó un enfoque cuantitativo de tipo correlacional, con el objetivo de analizar la relación entre la producción de leche como variable dependiente y 39 variables explicativas. Se planteó la hipótesis de que dichas variables presentan

una asociación estadísticamente significativa con los niveles de producción lechera en la parroquia.

Para estimar dicha relación, se formuló un modelo de regresión lineal múltiple con transformación logarítmica en base natural aplicada tanto a la variable dependiente como a las explicativas, con el fin de mitigar posibles problemas de heterocedasticidad y mejorar la distribución de los residuos, para la selección de variables se utilizó el método stepwise. La estimación de los parámetros se realizó mediante el método de Mínimos Cuadrados Ordinarios (MCO), la validación de los supuestos de multicolinealidad, homocedasticidad, normalidad de los errores y funcionalidad del modelo, se realizó mediante el análisis de los factores de inflación de varianza (VIF) y la aplicación de los test de Breusch- Pagan, Shapiro-Wilk, Jarque-Bera y RESET de Ramsey respectivamente.

Con el objeto de abordar una posible endogeneidad en las variables explicativas del modelo, se empleó una estimación mediante el método de Variables Instrumentales (IV). La validez de estos instrumentos fue evaluada en términos de relevancia mediante la significancia estadística en la primera etapa del modelo, y de exogeneidad a través del test de Wu-Hausman.

La base de datos utilizada fue generada por el Gobierno Autónomo Descentralizado parroquial de Julio Andrade en el año 2021, en el marco de su programa de fomento a la actividad agropecuaria. Esta base está compuesta por 162 registros correspondientes a unidades productivas dedicadas a la ganadería lechera, y fue obtenida mediante un muestreo no probabilístico por conveniencia, dado el carácter operativo del levantamiento.

Los análisis estadísticos se llevaron a cabo en el entorno R, versión 4.2.1. Para la estimación por MCO se utilizó la función `lm()`, mientras que para el modelo con variables instrumentales se empleó la función `ivreg()` del paquete AER. Además, se usaron los paquetes complementarios `readxl`, `dplyr`, `car`, `forcats`, `lmtest`, `sándwich`, `glmnet`, `ggplot2`, `ggpubr`, `ggthemes`, `broom`, `stargazer`, `metrics`, `nortest`, `leaps`, `tseries`, para la determinación del modelo de regresión lineal múltiple.

Modelo de Regresión Lineal Múltiple

(Lind, Marchal, & Wathen, 2015), la regresión lineal múltiple es una técnica estadística utilizada para modelar la relación existente entre diversas variables independientes ($X_1, X_2, X_3 \dots, X_n$) y una variable dependiente (y), las variables independientes ayudan a explicar o predecir mejor a la variable dependiente (Y).

El objetivo es encontrar una ecuación que permita predecir la variable dependiente (y) en función de las variables independientes (x).

La ecuación general es:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Dónde:

β_0 Es el intercepto (Valor de (y) cuando todas las x son cero)

$\beta_1, \beta_2, \dots, \beta_3$ = Son los coeficientes que indican cuando cambia (y) por cada unidad de (x)

ε Es el término de error (Diferencia entre el valor real y el predicho)

Transformación logarítmica

Según (Peña, 2002), la transformación logarítmica es útil cuando los datos presentan una asimetría positiva pronunciada, permitiendo una mejor aproximación a la normalidad. (p. 134).

Por otro lado, (Montgomery & Runger, 2010), explican que, en muchos casos, la variabilidad de una variable depende de su magnitud, por lo que la transformación logarítmica ayuda a reducir la heterocedasticidad y mejorar el ajuste de modelos estadísticos. (p. 245).

Especificación del modelo

Para analizar la relación entre la variable dependiente y un conjunto de variables explicativas, se estimó un modelo de regresión lineal múltiple con transformación logarítmica tanto de la variable dependiente como de las variables independientes continuas y estrictamente positivas. La especificación del modelo log-log es la siguiente:

$$\ln(Y_i) = \beta_0 + \sum_{j=1}^n \beta_j \ln(X_{ij}) + \varepsilon_i$$

Donde Y_i es la producción diaria de leche en litros.

Variables independientes:

$$\ln(X_{ij})$$

Para cada predictor:

$\ln X_1 \dots \ln X_2 \dots \ln X_3$ Variables independientes

ε_i Término de error aleatorio

Este modelo permite interpretar los coeficientes asociados a las variables logarítmicas como elasticidades: es decir, la variación porcentual en la variable dependiente ante un cambio porcentual en las variables explicativas.

Supuestos del modelo

Exogeneidad:

$$E[\varepsilon_i | X] = 0$$

Varianza constante:

$$Var(\varepsilon_i) = \sigma^2$$

No multicolinealidad perfecta:

$$Rank(X) = n + 1$$

Normalidad de errores:

$$\varepsilon_i \sim N(0, \sigma^2)$$

Estimación por Mínimos Cuadrados (MCO)

Es una técnica estadística que se utiliza para realizar la estimación de los coeficientes de un modelo de regresión lineal, con la finalidad de encontrar la línea de mejor ajuste que minimice la suma de cuadrados de los errores entre los valores observados y los predichos por el modelo. (Wooldridge, 2020).

El modelo de estimación por Mínimos Cuadrados Ordinarios (MCO) permitió estimar el impacto de cada variable explicativa con respecto a la producción diaria de leche.

Derivación Matricial:

$$\hat{\beta} = (X'X)^{-1}X'y, \quad \text{Var}(\hat{\beta}) = \sigma^2(X'X)^{-1}$$

Donde:

$\hat{\beta}$ Vector de coeficientes estimados del modelo de regresión

$X'X$ Producto matricial de X' por X

$(X'X)^{-1}$ Inversa de la matriz $X'X$

y Vector de la variable dependiente (dimensión $n \times 1$)

$\text{Var}(\hat{\beta})$ Matriz de varianzas y covarianzas de los estimadores de β (dimensión $k \times k$)

σ^2 Varianza del término de error (asumida constante: homocedasticidad)

Validación y diagnóstico

A continuación, se presenta las fórmulas y pruebas aplicadas:

Multicolinealidad

Factor de inflación de varianza (VIF) X_j :

$$VIF_j = \frac{1}{1 - R_j^2}$$

Dónde R_j^2 es el coeficiente de determinación de la regresión de X_j contra las demás regresoras.

Heterocedasticidad

Test de Breusch–Pagan

Ajustar

$$\varepsilon_i^2 = \alpha_0 + \alpha'X_i + u_i$$

Estadístico Breusch–Pagan:

$$BP = NxR_{aux}^2 \sim X_{k-1}^2$$

Normalidad de residuos

Jarque–Bera:

$$JB = \frac{n}{6} \left(s^2 \frac{(K-3)^2}{4} \right)$$

Dónde S es la asimetría y K la curtosis de los residuos

Especificación funcional

Test RESET de Ramsey

Ajustar:

$$Y_i = X_i' \hat{\beta} + \gamma_1 \hat{Y}_i^2 + \gamma_2 \hat{Y}_i^3 + u_i$$

Estadístico F:

$$F = \frac{(SSR_r - SSR_u/m)}{SSR_u/(n - k)}$$

Predicción y Variables Instrumentales (IV)

Intervalos de confianza y predicción

Para un nuevo vector x_0 :

$$\hat{Y}_0 = x_0' \hat{\beta}, \quad \text{Var}(\hat{Y}_0) = \sigma^2 x_0' (X'X)^{-1} x_0$$

Intervalo de confianza $1 - \alpha$:

$$\hat{Y}_0 \pm t_{n-p, \alpha/2} \sqrt{\text{Var}(\hat{Y}_0)}$$

Variables instrumentales (IV)

Es una técnica econométrica utilizada cuando una o más variables explicativas en un modelo de regresión están correlacionadas con el término de error, generando endogeneidad y sesgo en las estimaciones de mínimos cuadrados ordinarios. (Wooldridge, 2020)

Se emplea el estimador de dos etapas:

Primera etapa:

$$X = Z_{\pi} + u$$

Segunda etapa:

$$Y = \beta_0 + \beta_1 \hat{X} + \varepsilon$$

Se emplea el test de Wu-Hausman para comparar estimadores de MCO e IV para detectar endogeneidad.

Se utiliza el test de Sargan para validar los instrumentos

RESULTADOS Y DISCUSIÓN

Especificación del modelo

$$\ln(Y_i) = \beta_0 + \sum_{j=1}^n \beta_j (X_{ij}) + \varepsilon_i$$

Tabla 1

Variables independientes

Variable	Notación	Variable	Notación
X ₁	Vacas	X ₂₁	Vacas producción raza Normando
X ₂	Hectáreas	X ₂₂	Vacas producción raza Pizan
X ₃	Experiencia	X ₂₃	Prevención
X ₄	Genera ganancias	X ₂₄	Cantidad pasto
X ₅	Vende animales	X ₂₅	Leche líquida cruda
X ₆	Vende tierras	X ₂₆	Queso amasado

X ₇	Vende mercancías	X ₂₇	Queso pasteurizado
X ₈	Solicita crédito	X ₂₈	industria
X ₉	No variación	X ₂₉	Piquero
X ₁₀	Disminuye producción	X ₃₀	quesera
X ₁₁	Incrementa ganado	X ₃₁	Capacitación empleados
X ₁₂	Mejor genética	X ₃₂	Tic
X ₁₃	Escaso alimento	X ₃₃	Acceso tierras
X ₁₄	Raza ineficiente	X ₃₄	Acceso riego
X ₁₅	Falta veterinario	X ₃₅	asesoramiento
X ₁₆	Fertilidad suelo	X ₃₆	Miembro gremio
X ₁₇	Vacas producción	X ₃₇	Recibe crédito
X ₁₈	Vacas producción raza Criolla	X ₃₈	Monto crédito
X ₁₉	Vacas producción raza Holstein	X ₃₉	Forrajes silo concentrado
X ₂₀	Vacas producción raza Jersey		

Selección de variables con el método Stepwise

En la tabla 2 se presentan los factores que tienen un efecto estadístico positivo significativo con un $p < 0.05$.

Tabla 2

Selección de variables con el método Stepwise

Variable	Estimate	Std. Error	t value	Pr(> t)	Signif.
(Intercept)	2,120205	0,188185	11,267	< 2e-16	***
Solicita crédito	0,24447	0,052468	4,659	7.26e-06	***
Vacas en producción	0,086778	0,006469	13,415	< 2e-16	***
Prevención sanitaria	0,3704	0,059773	6,197	5.97e-09	***
Tipo de alimentación	0,105691	0,017273	6,119	8.81e-09	***

Supuestos

Exogeneidad

$$E[\varepsilon_i|X] = 0$$

La exogeneidad de las variables explicativas, tales como vacas en producción, composición genética del hato, prácticas de alimentación, prevención sanitaria, acceso a financiamiento, no están relacionadas con el término de error. Se asume que estas variables, que representan decisiones de manejo y características estructurales de las unidades ganaderas inciden en la producción lechera, al cumplirse el supuesto de exogeneidad los estimadores son insesgados y confiables.

Varianza constante

$$Var(\varepsilon_i) = \sigma^2$$

La transformación logarítmica de las variables en el modelo log- X se justifica como un mecanismo que inherentemente tiende a estabilizar la varianza de los residuos. Al modelar relaciones en términos de elasticidades, se asume que las desviaciones del modelo son proporcionales, lo que usualmente conduce a una varianza del error más constante a lo largo de

los diferentes niveles de las variables explicativas. La confirmación empírica de este supuesto se llevó a cabo mediante pruebas formales de heterocedasticidad; en caso de identificar varianza no constante residual, se implementaron errores estándar robustos para asegurar la validez de la inferencia.

No multicolinealidad perfecta

$$\text{Rank}(X) = n + 1$$

La ausencia de multicolinealidad perfecta se evidencia por la estimación exitosa de todos los coeficientes del modelo. Las variables explicativas, de mayor influencia, representan factores importantes de la actividad ganadera, si bien puede existir cierto grado de correlación entre los regresores, esta no es lo suficientemente elevada como para comprometer la estabilidad o la significancia estadística de los coeficientes estimados.

Normalidad de errores

$$\varepsilon_i \sim N(0, \sigma^2)$$

El Teorema del Límite Central sustenta la aproximación a la normalidad de la distribución de los estimadores de Mínimos Cuadrados Ordinarios (MCO), independientemente de la distribución exacta de los errores subyacentes. Esto asegura la validez de las pruebas de hipótesis y la construcción de intervalos de confianza.

Estimación por mínimos cuadrados ordinarios (mco)

Se estimó un modelo de regresión lineal múltiple bajo especificación $\log(Y) \sim X$ con el objetivo de analizar los factores que inciden en la producción de leche. Los resultados indican las variables estadísticamente significativas y que se asocian positivamente con la productividad, confirmando su relevancia estratégica dentro del sistema productivo (Ver tabla 3).

Tabla 3
Resultados Modelo Mínimos Cuadrados Ordinarios

Variable	B	EE	t	p	Significancia
(Intercepto)	1.611	0.319	5.045	< .001	***
Solicita crédito	0.235	0.071	3.311	.001	**
Vacas producción	0.088	0.011	7.908	< .001	***
Vacas producción raza Normando	0.651	0.176	3.699	< .001	***
Prevención sanitaria	0.364	0.066	5.501	< .001	***
Tipo de alimentación	0.1062	0.019	5.542	< .001	***

Residual standard error: 0.2965 on 122 degrees of freedom
 Multiple R-squared: 0.8993, Adjusted R-squared: 0.8671
 F-statistic: 27.94 on 39 and 122 DF, p-value: < 2.2e-16

El error estándar residual (RSE) fue de 0.2965 en escala logarítmica, al transformarlo a la escala original mediante $(e^{0.2965} - 1) * 100$, se obtiene un error relativo promedio aproximado del 34.5%. Esta magnitud indica una precisión aceptable del modelo, ya que las observaciones individuales no presentan desviaciones significativas respecto a los valores estimados.

$R^2 = 0.8993$ Indica que aproximadamente el 90% de la variabilidad observada en la producción de leche es explicada por las variables independientes del modelo. Se trata de un ajuste muy alto que evidencia la pertinencia del conjunto de regresores seleccionados.

En términos de elasticidad, el acceso a crédito está relacionado con un incremento del 26.49% en la producción diaria de leche, un aumento de una unidad en vacas en producción se asocia con un incremento del 9.2%, de igual forma, el tipo de alimentación y la prevención sanitaria genera aumentos del 11.2% y 43.9% respectivamente (Ver tabla 4).

Tabla 4
Transformación de coeficientes a escala original

Variable	Estimación	Cálculo	Cambio (%) aproximado
Solicita crédito	0.2352	$(e^{0.2352-1}) \times 100$	26.5%
Vacas en producción	0.08817	$(e^{0.08817-1}) \times 100$	9.2%
Prevención sanitaria	0.3640	$(e^{0.3640-1}) \times 100$	43.9%
Tipo de alimentación	0.1062	$(e^{0.1062-1}) \times 100$	11.2%

Comparación de errores estándar clásicos y robustos

Se compararon los errores estándar clásicos y robustos, para evaluar posibles problemas de heterocedasticidad, el modelo es razonablemente estable, ya que los errores estándar clásicos y robustos son similares en casi todas las variables lo que fortalece la confiabilidad estadística del modelo (Ver tabla 5).

Tabla 5
Comparación de errores estándar clásicos y robustos

Variable	Estimate	Std. Error (Clásicos)	Std. Error (Robustos)
(Intercept)	1.6111	0.3193	0.4002
Solicita crédito	0.2352	0.0710	0.0607
Vacas producción	0.08817	0.0111	0.0112
Prevención sanitaria	0.3640	0.0662	0.0670
Tipo de alimentación	0.1062	0.0192	0.0188

Validación y diagnóstico

Factor de Inflación de varianza (VIF)

En la Tabla 6, se observa que todos los valores de VIF se encuentran por debajo del umbral crítico de 5. El valor más alto corresponde a la variable vacas en producción con un VIF de 2.77, seguido de número de hectáreas con 2.43, estos niveles indican una correlación moderada con otras variables explicativas, pero no comprometen la estabilidad ni la interpretabilidad del modelo estimado.

En consecuencia, se concluye que no existen indicios de multicolinealidad severa entre las variables incluidas en el modelo, lo que respalda la validez de los coeficientes estimados y fortalece la confiabilidad de los resultados.

Tabla 6*Factores de Inflación de Varianza*

Variables	VIF
Vacas en producción	2.773765
Número de hectáreas	2.431144
Prevención sanitaria	1.513019
Tipo de alimentación	1.184201
Genera ganancias	1.063487
Vende tierras	1.023418

Test de Breusch–Pagan

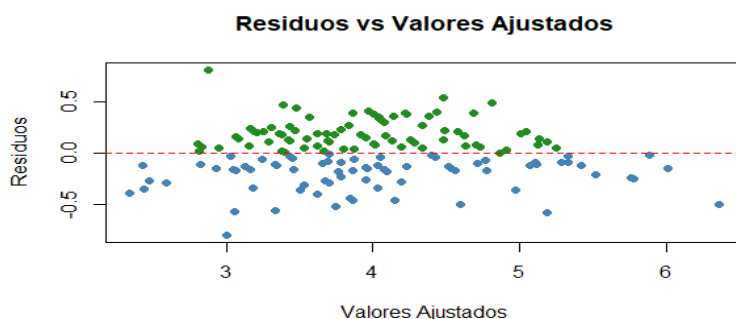
Dado que el p-valor 0.3388 es mayor al umbral común de significancia ($\alpha= 0.05$), no se rechaza la hipótesis nula de homocedasticidad. Por lo tanto, no se encuentran evidencias estadísticas que sugieran la presencia de heterocedasticidad en los residuos del modelo.

Este resultado respalda la validez del estimador de Mínimos Cuadrados Ordinarios (MCO), en cuanto a la consistencia de sus errores estándar bajo el supuesto de varianza constante (Ver tabla 7).

Tabla 7*Prueba studentized Breush-Pagan test*

Estadístico	Valor
BP	6.81
df	6
p-value	0.3388

En la Figura 1, se observa que los residuos representados por puntos de diferentes colores se distribuyen aproximadamente de forma simétrica y sin una estructura evidente alrededor de la línea roja punteada que representa el valor cero. Aunque existen ligeras fluctuaciones, no se identifica un patrón en forma de abanico, curva o embudo, lo cual indica que no hay evidencia visual de heterocedasticidad ni de una mala especificación funcional.

Figura 1*Homocedasticidad*

Para evaluar la normalidad de los residuos del modelo de regresión, se aplicó la prueba de Shapiro-Wilk y la prueba de Jarque – Bera.

Prueba de Shapiro-Wilk

Los resultados del test indicaron un valor de $W = 0.99153$ y un valor $p = 0.4528$, dado que es mayor al nivel de significancia de 0.05, no se rechaza la hipótesis nula de normalidad, esto sugiere que los residuos del modelo no presentan desviaciones significativas respecto a una distribución normal, cumpliendo así con uno de los supuestos fundamentales de la regresión lineal. Esto implica que las predicciones del modelo sobre la producción diaria de leche son consistentes y no están sesgadas por errores sistemáticos (Ver tabla 8).

Tabla 8

Prueba de Shapiro-Wilk

Estadístico	Valor
W	0.99153
p	0.4528

Test de Jarque-Bera

Dado que el valor-p (0.4799) es mayor que $\alpha = 0.05$, no se rechaza la hipótesis nula de normalidad de los residuos.

Los residuos del modelo de regresión lineal múltiple aplicado al análisis de producción de leche no presentan desviaciones significativas respecto a la normalidad. Esto respalda la validez estadística del modelo, permitiendo realizar inferencias fiables sobre las variables que explican la producción lechera (Ver tabla 9).

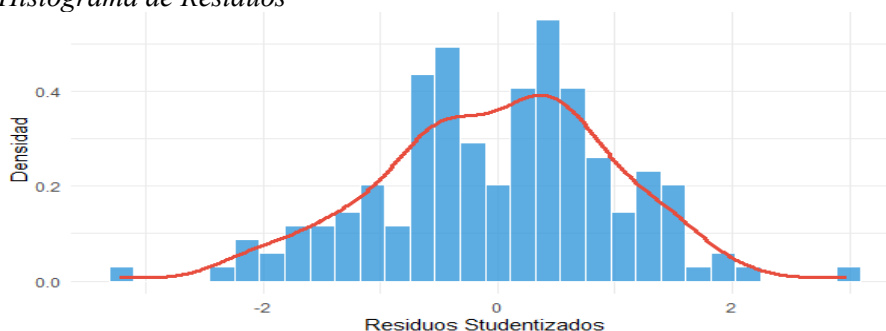
Tabla 9

Prueba Jarque Bera test

Estadístico	Valor
x-squared	1.4682
df	2
p-value	0.4799

El histograma presentado en la Figura 2 muestra una concentración central de los residuos en torno a cero y una forma simétrica moderadamente cercana a la campana de Gauss. La curva de densidad superpuesta en rojo permite observar que los residuos se distribuyen de manera aproximadamente normal, sin evidencia de colas largas ni de valores atípicos extremos.

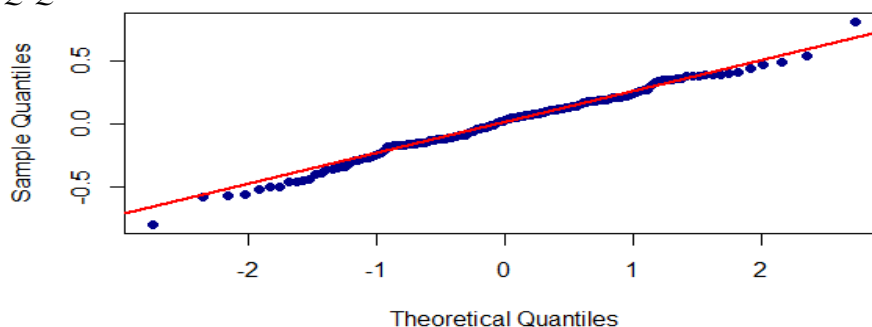
Figura 2
Histograma de Residuos



El resultado de la figura 3 es coherente con los valores obtenidos en las pruebas de normalidad de Shapiro–Wilk y Jarque–Bera (p -valores > 0.45), reforzando la conclusión de que los residuos no presentan desviaciones significativas respecto a la normalidad.

La adecuada alineación de los residuos en el Q-Q plot implica que las predicciones del modelo sobre la producción diaria de leche no están distorsionadas por errores no aleatorios. Esto refuerza la fiabilidad de las estimaciones obtenidas y apoya el uso del modelo para la toma de decisiones en unidades productoras de leche en la parroquia Julio Andrade.

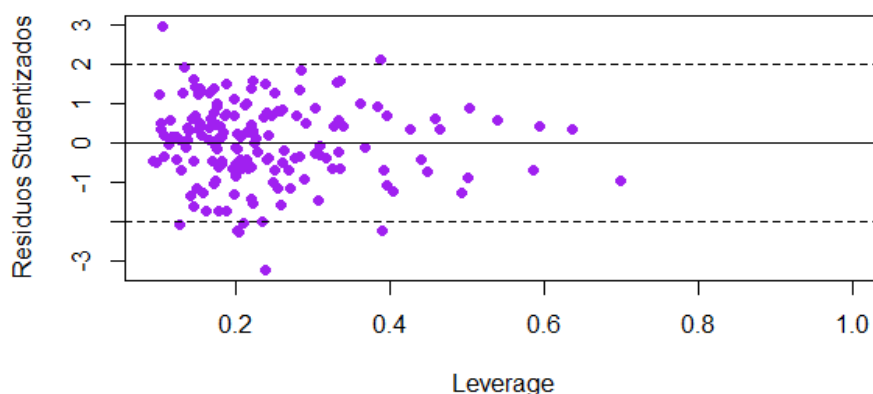
Figura 3
Q-Q Plot de Residuos



En la Figura 4 de residuos studentizados versus leverage permite detectar observaciones influyentes. En el modelo estimado, los puntos se concentran dentro del rango de ± 2 y presentan leverage menor a 0.5, sin evidencias de observaciones con alta influencia ni residuos anómalos. Esta distribución indica que no existen valores atípicos extremos que comprometan la estabilidad del modelo, lo cual respalda la validez de los resultados obtenidos.

La ausencia de valores influyentes extremos implica que las relaciones estimadas entre las prácticas ganaderas como la cantidad de vacas en producción, prevención sanitaria, tipo de alimentación no están sesgadas por unidades productivas atípicas o altamente distintas al resto. Esto sugiere que el modelo es representativo de la realidad ganadera local y proporciona resultados robustos para orientar políticas o decisiones técnicas en el sector.

Figura 4
Observaciones Influyentes



Test RESET de Ramsey

Se aplicó la prueba RESET de Ramsey para evaluar la especificación funcional del modelo de Mínimos Cuadrados Ordinarios. El resultado del test da un valor p de 0.4999, lo cual indica que al nivel de significancia del 5%, no se rechaza la hipótesis nula que afirma la correcta especificación del modelo. Esta evidencia sugiere que la forma funcional del modelo es adecuada para explicar el comportamiento de la variable dependiente en el contexto de la producción lechera (Ver tabla 10).

Tabla 10
Test RESET de Ramsey

Estadístico	Valor
RESET	3.0569
Df1	2
Df2	149
p-value	0.4999

Predicción y variables instrumentales

Intervalos de confianza y predicción

A partir de los resultados de predicción en escala logarítmica, se estima que la producción diaria de leche para una nueva unidad ganadera con características similares a la muestra es de aproximadamente 55.4 litros. Con un intervalo de predicción al 95 %, se espera que dicha producción se ubique entre 33.6 y 91.0 litros. Esta amplitud refleja la incertidumbre inherente a variaciones individuales, y el intervalo captura tanto la variabilidad del modelo como la del nuevo dato (Ver tabla 11).

Tabla 11*Intervalo de predicción para un nuevo dato*

Límites	Valores	
fit	4.013211	$e^{4.013211} \cong 55.36$ litros
lwr	3.515856	$e^{3.515856} \cong 33.65$ litros
upr	4.510566	$e^{4.510566} \cong 90.97$ litros

Variables Instrumentales

El modelo de variables instrumentales fue estimado para corregir posibles problemas de endogeneidad, utilizando instrumentos previamente identificados. A continuación, se presentan los coeficientes estimados y su significancia estadística (Ver tabla 12).

Tabla 12*Coefficientes del modelo IV*

Variable	Estimador	Error estándar	Valor t	p-valor
(Intercepto)	5.1502	8.6229	0.597	0.551
Hectáreas	-0.1360	0.2994	-0.454	0.650
Genera ganancias	-0.7080	2.5627	-0.276	0.783
Vende tierras	-0.4229	14.3442	-0.029	0.977
Vacas producción	0.1982	0.2151	0.921	0.358
Prevención Sanitaria	0.4657	0.3887	1.198	0.233
Tipo de alimentación	0.1807	0.2855	0.633	0.528

Residual standard error: 0.9135 on 155 degrees of freedom Multiple R-Squared: -0.2145, Adjusted R-squared: -0.2615 Wald test: 17.77 on 6 and 155 DF, p-value: 1.291e-15

Se evaluó la solidez de los instrumentos utilizados hectáreas, genera ganancias y vende tierras mediante pruebas específicas. Los resultados se presentan en la (Ver Tabla 13).

Tabla 13*Diagnóstico de instrumentos*

Instrumento	gl1	gl2	Estadístico F	p-valor
Hectáreas	3	155	1.129	0.339
Genera ganancias	3	155	0.175	0.913
Vende tierras	3	155	0.026	0.994

Los valores de p mayores a 0.05 indican que los instrumentos no están suficientemente correlacionados con las variables endógenas, por lo tanto, se consideran débiles. Esto compromete la fiabilidad de las estimaciones obtenidas mediante el modelo IV.

Test de endogeneidad (Wu-Hausman)

El p-valor de 0.459 es mayor a 0.05, sugiere que no hay evidencia estadísticamente significativa de endogeneidad en el modelo. Por tanto, el uso de variables instrumentales no mejora sustancialmente la estimación respecto al modelo de Mínimos Cuadrados Ordinarios (MCO), lo cual respalda el uso del modelo MCO original (Ver Tabla 14).

Tabla 14*Test de endogeneidad (Wu–Hausman)*

Estadístico	Valor
Wu–Hausman	0.868
p-valor	0.459

Error estándar residual: 0.9135 en escala logarítmica, al transformarlo a la escala original mediante $(e^{0.9135} - 1) * 100$, implica que las predicciones del modelo presentan un error típico relativo de aproximadamente 149.3 % respecto a los valores observados en la escala original. Es decir, en promedio, las estimaciones difieren de los valores reales en un 149.3 %, lo que sugiere una variabilidad considerable no explicada por el modelo.

R²: -0.2145 Un valor negativo implica que el modelo con instrumentos explica menos variabilidad que una media constante; es decir, el modelo es inapropiado.

A pesar de los intentos por abordar una posible endogeneidad mediante el uso de variables instrumentales, los resultados muestran que:

- No hay evidencia robusta de endogeneidad demostrada por el test de Wu–Hausman,
- Los instrumentos utilizados son estadísticamente débiles,
- El modelo IV tiene débil capacidad explicativa (R² negativo),
- No se puede validar los instrumentos por Sargan.

En consecuencia, el modelo de MCO sigue siendo preferible para analizar los determinantes de la producción lechera en las unidades ganaderas de la parroquia Julio Andrade.

DISCUSIÓN

Según (Wooldridge, 2020), valores de VIF menores a 10 descartan multicolinealidad severa, y valores inferiores a 5 pueden considerarse como evidencia de colinealidad mínima. En línea con esta postura, (Greene, 2018) resalta que la presencia de colinealidad solo debe considerarse problemática si compromete la precisión de las estimaciones o la estabilidad de los coeficientes. En el presente estudio, los valores más altos corresponden a las variables vacas en producción 2.77 y hectáreas 2.43, sin exceder niveles críticos. Esto respalda la fiabilidad de las estimaciones obtenidas mediante el modelo de Mínimos Cuadrados Ordinarios (MCO).

Por tanto, con base en los criterios establecidos por autores como (Wooldridge, 2020) y (Greene, 2018), puede afirmarse que el modelo estimado no presenta problemas relevantes de multicolinealidad, permitiendo interpretar los coeficientes de manera confiable y sin distorsiones atribuibles a redundancia lineal entre los predictores.

La prueba de Breusch–Pagan con un p-valor de 0.3388 que es mayor al nivel de significancia de $\alpha = 0.05$, tal como indican (Wooldridge, 2020), la presencia de homocedasticidad es deseable, ya que garantiza la eficiencia de los estimadores bajo el supuesto

clásico. En este caso, el modelo supera satisfactoriamente esta prueba diagnóstica, fortaleciendo la validez estadística de las inferencias realizadas.

El test de Shapiro–Wilk con un p-valor de 0.4528, es considerablemente mayor al umbral de significancia estadística $\alpha = 0.05$. Esto indica que los residuos del modelo no presentan desviaciones significativas respecto a una distribución normal, lo cual respalda la validez de las inferencias basadas en los estimadores de Mínimos Cuadrados Ordinarios (MCO).

Según (Lind, Marchal, & Wathen, 2015) recomiendan verificar este supuesto especialmente cuando se busca interpretar valores extremos o aplicar intervalos y predicciones en escala transformada.

El test de Jarque–Bera con un p-valor = 0.4799, mayor a 0.05, no se rechaza la hipótesis nula de normalidad, en consecuencia, se concluye que los residuos del modelo no presentan evidencia estadística de asimetría o curtosis anómalas.

Según (Greene, 2018), la normalidad de los errores es un supuesto esencial en muestras pequeñas para garantizar la precisión de los intervalos y la validez exacta de los tests inferenciales. (Lind, Marchal, & Wathen, 2015) destacan la utilidad de esta prueba en la validación de modelos econométricos en análisis predictivos.

El test RESET con un resultado de un p-valor de 0.4999, considerablemente mayor al nivel de significancia del 5 %, esto indica que no hay evidencia estadística de que la forma funcional del modelo esté mal especificada, lo que respalda la validez estructural del modelo log-X estimado.

De acuerdo con (Guajarati & Porter, 2010), la prueba RESET de Ramsey es una herramienta útil para evaluar errores de especificación no observables, especialmente en modelos no lineales transformados, como el logarítmico.

La ausencia de problemas funcionales sustenta que las variables seleccionadas y su transformación logarítmica capturan de forma adecuada las relaciones entre los predictores y la producción lechera, reforzando la robustez del modelo estimado.

Con base en el modelo de regresión log-log estimado mediante Mínimos Cuadrados Ordinarios (MCO), se generó un intervalo de predicción al 95 % para una unidad ganadera con características específicas. En la escala logarítmica, el valor ajustado fue de 4.013, con un límite inferior de 3.516 y superior de 4.511. Al aplicar la transformación exponencial, los valores se interpretan como una producción diaria estimada de 55.36 litros, con un intervalo que oscila entre 33.65 y 90.97 litros.

Este intervalo refleja tanto la incertidumbre asociada al valor esperado de la respuesta como la variabilidad inherente en futuras observaciones (Montgomery & Runger, 2010). La amplitud del rango es coherente con el error estándar residual del modelo, equivalente a un error relativo promedio del 34.5 %, lo que indica una precisión aceptable en la predicción para el contexto ganadero analizado.

El modelo de regresión lineal múltiple estimado mediante Mínimos Cuadrados Ordinarios (MCO) presenta un coeficiente de determinación

$R^2=0.899$, lo que indica que aproximadamente el 89.9% de la variabilidad observada en la producción de leche es explicada por las variables independientes incluidas en el modelo, de acuerdo con (Guajarati & Porter, 2010), un resultado alto sugiere un buen poder explicativo y mejor ajuste del modelo a los datos observados.

Los resultados del modelo de Variables Instrumentales (IV) revelan una capacidad explicativa deficiente, evidenciada por un R^2 negativo de -0.2145 y un error estándar residual elevado de 0.9135 , lo que implica una alta imprecisión en las predicciones. Ninguna variable resultó estadísticamente significativa, y los coeficientes muestran gran inestabilidad.

Según (Wooldridge, 2020), estos síntomas son característicos cuando los instrumentos son débiles o inválidos, lo cual compromete la consistencia del estimador IV. (Greene, 2018) añade que, en ausencia de evidencia clara de endogeneidad, el uso de IV puede inducir más error que el modelo MCO. Por tanto, los hallazgos sugieren que el modelo IV no es adecuado para este caso.

El modelo de regresión lineal múltiple con transformación logarítmica muestra cómo distintos factores influyen en la producción de leche, hallazgos que concuerdan con investigaciones previas sobre la eficiencia productiva en sistemas lecheros.

Los productores que acceden al crédito presentan un aumento estimado del 26.49% en su producción diaria de leche, en comparación con aquellos que no lo hacen. Esta relación sugiere un fuerte vínculo entre el financiamiento y la capacidad productiva, estos hallazgos coinciden con lo reportado por (Mishra, El-Osta, & Sandretto, 2009), quienes señalan que el acceso al crédito les permite una mayor probabilidad de innovar, diversificar e incrementar su producción mejorando la inversión en insumos de calidad, alimentación, infraestructura tecnología y genética animal.

El resultado de que cada vaca adicional aumenta la producción en un 9.2% con rendimientos decrecientes es consistente con estudios que destacan la importancia del manejo del hato para evitar sobrepoblación y mantener la productividad individual. Investigaciones como (Méndez, Caunedo, & Fernández, 2004) señalan que el crecimiento del número de vacas mejora la producción agregada, pero la eficiencia por animal disminuye debido a la etapa de lactancia.

El impacto positivo del 11.2% en la variable tipos de alimentación como, silos o concentrados es respaldado por (Delgado, Olivos, & Muroya, 2020), quienes destacan que una nutrición equilibrada con suplementos proteicos y energéticos optimiza la conversión alimenticia, incrementa la producción y mejora la composición de la leche, justificando el efecto positivo observado en el modelo.

Un efecto significativo proviene de la inversión en sanidad animal, con un incremento del 43.9%, lo que respalda estudios como el informe de (HealthforAnimal, 2022), que destacan

cómo las enfermedades reducen la producción y elevan los costos de tratamiento, el 60% de la implementación de estrategias preventivas, como vacunaciones y programas de control en ganado vacuno se correlaciona con un aumento de la producción de más del 50%.

CONCLUSIONES

El análisis realizado mediante un modelo de regresión lineal múltiple permitió identificar y cuantificar los principales determinantes de la producción diaria de leche en unidades ganaderas de la parroquia Julio Andrade. A partir de los resultados obtenidos en la estimación, validación y predicción del modelo, se concluye que:

Las pruebas de normalidad Shapiro–Wilk y Jarque–Bera no revelaron desviaciones significativas respecto a la distribución normal de los residuos. La prueba de Breusch–Pagan confirmó la homocedasticidad, y el test RESET de Ramsey no evidenció errores funcionales, lo cual respalda la validez de los supuestos clásicos.

Se descarta la presencia de multicolinealidad severa, los factores de inflación de la varianza (VIF) fueron inferiores al umbral crítico de 5, lo que garantiza que las variables explicativas no presentan redundancia significativa entre sí.

La estimación por Mínimos Cuadrados Ordinarios (MCO) presenta un ajuste sólido, con un $R^2=0.899$, lo que indica que el modelo explica el 89.9 % de la variabilidad en la producción diaria de leche en escala logarítmica. El error estándar residual de 34.5% refleja una dispersión baja de los residuos, varias variables resultan estadísticamente significativas y presentan coeficientes con errores estándar relativamente bajos, lo que indica estimaciones precisas y robustas. En particular, variables como vacas en producción, prevención sanitaria y tipo de alimentación muestran una relación positiva y significativa con la producción diaria de leche.

El modelo de Variables Instrumentales (IV) muestra serias deficiencias, un R^2 negativo (-0.2145), alta dispersión de residuos ($RSE = 0.913$), presenta coeficientes inestables con errores estándar elevados, lo que sugiere una alta varianza y menor precisión en las estimaciones. Además, ninguna variable en el modelo IV alcanza significancia estadística, la debilidad de los instrumentos y la ausencia de endogeneidad respaldan la validez y superioridad del modelo de Mínimos Cuadrados Ordinarios (MCO) para el análisis.

En conjunto, los resultados permiten afirmar que el modelo MCO es adecuado, parsimonioso y confiable para explicar la variación en la producción lechera, siendo preferible frente a modelos alternativos con variables instrumentales.

El modelo confirma que una gestión eficiente de los sistemas de producción depende de múltiples factores. Se encontró que el aumento en el número de vacas mejora la producción agregada. La alimentación basada en silos y concentrados tiene un impacto positivo significativo, lo que refuerza la importancia de una nutrición balanceada para maximizar la producción.

Además, la inversión en prevención sanitaria demuestra ser un factor determinante, con un alto impacto en la producción al reducir enfermedades y mejorar la eficiencia del sistema. La genética también juega un papel fundamental, siendo la raza Normando la más eficiente en términos de producción lechera.

En general, estos hallazgos subrayan la importancia de generar una estrategia integral que combine un manejo adecuado del hato, una nutrición optimizada, la implementación de medidas preventivas de salud y la selección genética adecuada para maximizar la eficiencia y sostenibilidad de la producción lechera.

La incorporación de variables como la calidad del pasto, la frecuencia del servicio veterinario, el nivel educativo del productor, el acceso a infraestructura para el ordeño, y el precio de venta de la leche podría enriquecer sustancialmente el análisis de los determinantes de la producción lechera, estas variables capturan dimensiones actualmente ausentes en el modelo. Su inclusión permitiría identificar con mayor precisión las restricciones y oportunidades que enfrentan los productores, mejorando la capacidad predictiva del modelo y fortaleciendo la formulación de estrategias diferenciadas para el desarrollo del sector lechero.

REFERENCIAS

- Acosta, D., Gómez, D., & Rivera, C. (2020). *Plan de Desarrollo y Ordenamiento Territorial*. Obtenido de <https://gpjulioandrade.gob.ec/carchi/wp-content/uploads/2023/12/PDOT-JULIO-ANDRADE-2019-2023.pdf>
- Delgado, A., Olivos, W., & Muroya, C. (2020). *La Alimentación de Las Vacas Lecheras*. Amazon Digital Services LLC - KDP Print US.
- FAO. (2025). *Organización de las Naciones Unidas para la Alimentación y Agricultura*. Obtenido de <https://www.fao.org/animal-production/es?>
- Greene, W. (2018). *Econometric Analysis*. Pearson.
- Guajarati, D., & Porter, D. (2010). *Econometría*. México: McGraw-Hill.
- HealthforAnimal. (2022). *Sanidad animal y Sostenibilidad - Resumen de un análisis global de datos*. Obtenido de <https://healthforanimals.org/wp-content/uploads/2023/06/Sanidad-animal-y-sostenibilidad.pdf>
- Lind, D., Marchal, W., & Wathen, S. (2015). *Estadística Aplicada a los Negocios y Economía*. México: McGraw-Hill.
- MAG. (23 de 05 de 2021). *Ministerio de Agricultura y Ganadería*. Obtenido de https://www.agricultura.gob.ec/wp-content/uploads/downloads/2023/07/PEI-MAG-2021-2025-22_04_2022-signed-signed-1_compressed.pdf?
- Méndez, A., Caunedo, J., & Fernández, M. (2004). Relación entre el porcentaje de vacas en ordeño y la producción láctea total del rebaño. *Revista Cubana de Ciencia Agrícola*.
- Mishra, A., El-Osta, H., & Sandretto, C. (2009). *Factors affecting farm profitability*.
- Montgomery, D., & Runger, G. (2010). *Estadística aplicada y probabilidad para ingenieros*. Wiley.
- Peña, D. (2002). *Análisis de datos multivariantes*. Madrid: McGraw-Hill Interamericana de España.
- Wooldridge, J. M. (2020). *Introducción a la econometría. Un enfoque moderno*. CENGAGE Learning.