

Data Science Applied to a Market Study to Estimate the Demand for Bioplastics in the Peruvian Context

Jheyson Alexander Quiñones – Gabino¹; Maximiliano Arroyo - Ulloa²; Absalon Rivasplata – Sanchez³.

^{1, 2, 3} Universidad Católica Santo Toribio de Mogrovejo, Perú

¹72496938@usat.edu.pe, ²marroyo@usat.edu.pe, ³arivasplata@usat.edu.pe.

Abstract – The present study analyzes the unmet demand for bioplastics in Peru for the 2023–2027 period, considering the impact of policies such as Law No. 30884, which regulates the use of single-use plastics, and the growing interest in sustainability. Due to the lack of specific historical data for the bioplastics market, primary polypropylene (PP) destined for the packaging sector was used as a reference, given its relevance in the plastics industry and its potential to be replaced by biopolymers such as polyhydroxybutyrate (PHB). Information was gathered from official sources such as SUNAT and BCRP to estimate historical demand and supply. The analysis was conducted in three stages: demand estimation, supply estimation, and calculation of unmet demand. Tools such as Minitab and RStudio were employed to carry out linear regression, multiple stepwise regression, and advanced time series models (ARIMA and nonlinear regressions). The results reveal a significant unmet demand for bioplastics, representing an opportunity for investments in this emerging market. Additionally, a hypothetical investment scenario was developed to address part of this demand, projecting prices and market share. This study highlights the utility of data science in generating precise projections, optimizing strategic decisions, and fostering the transition toward a sustainable economy.

Keywords: bioplastics, unmet demand, regression analysis, time series, circular economy.

Data science aplicada a un estudio de mercado para estimar la demanda de bioplásticos en el contexto peruano

Jheyson Alexander Quiñones – Gabino¹; Maximiliano Arroyo - Ulloa²; Absalon Rivasplata – Sanchez³.

^{1, 2, 3} Universidad Católica Santo Toribio de Mogrovejo, Perú

¹72496938@usat.edu.pe, ²marroyo@usat.edu.pe, ³arivasplata@usat.edu.pe.

Resumen – El presente estudio analiza la demanda insatisfecha de bioplásticos en el Perú durante el horizonte 2023-2027, considerando el impacto de políticas como la Ley N° 30884, que regula el uso de plásticos de un solo uso, y el creciente interés en la sostenibilidad. Ante la falta de datos históricos específicos para el mercado de bioplásticos, se tomó como referencia el polipropileno (PP) en forma primaria destinada al sector de envases, dado su relevancia en la industria plástica y su potencial para ser sustituido por biopolímeros como el polihidroxibutirato (PHB). Se recopiló información de fuentes oficiales como SUNAT y BCRP para estimar la demanda y oferta históricas. El análisis se realizó en tres etapas: estimación de la demanda, estimación de la oferta y cálculo de la demanda insatisfecha. Se emplearon herramientas como Minitab y RStudio para llevar a cabo un análisis de regresión lineal, regresión múltiple escalonada y modelos avanzados de series temporales (ARIMA y regresiones no lineales). Los resultados muestran una significativa demanda insatisfecha de bioplásticos, que representa una oportunidad para inversiones en este mercado emergente. Además, se desarrolló un escenario de inversión hipotético que contempla cubrir parte de esta demanda, proyectando precios y participación de mercado. Este estudio resalta la utilidad de la ciencia de datos para generar proyecciones precisas, optimizar decisiones estratégicas y fomentar la transición hacia una economía sostenible.

Palabras clave: bioplásticos, demanda insatisfecha, análisis de regresión, series temporales y economía circular.

I. INTRODUCCIÓN

De acuerdo con European Bioplastics, en cooperación con el Nova – Institute [1], el mercado global de bioplásticos está experimentando un crecimiento exponencial. Se estima que la capacidad de producción aumentará de 2,8 millones de toneladas en 2024 a 6,3 millones de toneladas en 2026. Este notable incremento está impulsado por la creciente preocupación por el cambio climático, la necesidad de reducir la dependencia de plásticos convencionales derivados del petróleo y las crecientes exigencias del mercado para materiales más sostenibles y ambientalmente responsables.

En el Perú, la adopción de bioplásticos en el mercado es una realidad en pleno desarrollo. La Ley N° 30884, promulgada en 2018, marcó un punto de inflexión en la regulación de plásticos de un solo uso, fomentando un cambio

en la industria local y en las preferencias de los consumidores. Durante el primer año de vigencia de esta normativa, la participación de mercado de las bolsas biodegradables aumentó del 3,7% al 9,1% [2], mientras que las importaciones de envases biodegradables crecieron un 176% en 2022 con respecto al año anterior [3]. Estos datos reflejan no solo un interés creciente en materiales más sostenibles, sino también una oportunidad significativa para la producción y comercialización de bioplásticos a nivel nacional.

El Perú, además, cuenta con ventajas únicas para impulsar esta industria. Su vasta riqueza en recursos biológicos lo posiciona como un país con un potencial notable para la producción de bioplásticos de base biológica. Aunque estos recursos también son utilizados para generar productos como biogás, compost o biochar, investigaciones como las de Korkakaki *et al.* [4], Papa *et al.* [5] y Hereza [6] destacan que los bioplásticos, como el polihidroxialcanoato (PHA), generan ingresos hasta cinco veces mayores en comparación con dichos productos. Esta perspectiva hace que los bioplásticos sean una opción atractiva para los inversionistas interesados en alternativas sostenibles y rentables. Sin embargo, a pesar del creciente interés y las ventajas competitivas del país, una pregunta crítica permanece sin una respuesta clara: ¿qué tan grande es la demanda insatisfecha de bioplásticos en el Perú? Responder a esta interrogante es fundamental para comprender el verdadero potencial del mercado y tomar decisiones informadas sobre inversiones y estrategias de producción.

El presente estudio tiene como objetivo estimar la demanda insatisfecha de bioplásticos en el Perú. Al hacerlo, busca proporcionar una visión clara y detallada del mercado, identificando oportunidades estratégicas para consolidar la producción de estos materiales innovadores en el país. Este análisis, basado en un enfoque novedoso de ciencia de datos, es especialmente relevante, ya que permite realizar proyecciones más precisas al integrar un mayor volumen de información histórica. Este nivel de precisión contribuye no solo a la evaluación de viabilidad de proyectos, sino también a optimizar aspectos clave como costos de materia prima, planificación operativa y estrategias comerciales. En un mercado global competitivo y en un contexto local lleno de

potencial, contar con herramientas analíticas avanzadas se convierte en un recurso indispensable para posicionar a los bioplásticos como una alternativa viable y sostenible. Este estudio destaca cómo la aplicación de la ciencia de datos puede marcar la diferencia en la toma de decisiones estratégicas, reforzando su valor como herramienta clave en la transición hacia una economía más circular y sostenible.

II. MATERIALES Y MÉTODOS

El presente estudio adoptó un enfoque cuantitativo, predictivo y aplicado, orientado a estimar la evolución futura de la demanda y oferta del polipropileno (PP) en forma primaria destinado al sector de envases en el Perú. Dado que el mercado de bioplásticos no cuenta aún con registros históricos consolidados, se utilizó como variable proxy el comportamiento del PP, debido a su alta representatividad en la industria del plástico y su similitud funcional con biopolímeros como los PHA.

El análisis comenzó con una fase exploratoria de las series temporales correspondientes a la demanda y a la oferta. Para ello, se aplicaron el test de Dickey – Fuller aumentada (ADF) con el fin de evaluar la estacionariedad de las series, los correlogramas (ACF) para detectar autocorrelación, y el test de Durbin – Watson (DW) para examinar la independencia de los residuos. Esta etapa permitió identificar patrones estructurales en los datos y justificar el uso de regresión múltiple frente a modelos autoregresivos como ARIMA.

Seguidamente, se procedió a la recolección de variables macro y microeconómicas potencialmente explicativas, tales como el año (x1), el PBI de manufactura no primaria (x2), la inflación (x3), la población (x4), el precio del PP en forma primaria (x5) y la producción industrial de productos químicos, caucho y plástico (x6). Estas covariables fueron seleccionadas según su relevancia teórica y disponibilidad en fuentes oficiales.

Posteriormente, se llevó a cabo un análisis de regresión para determinar la existencia de relaciones funcionales estadísticamente significativas entre cada covariable y la variable dependiente (demanda u oferta). Aquellas que presentaron un valor p menor a 0,05 fueron consideradas candidatas para la modelación múltiple. Con este subconjunto de variables, se construyó un modelo de regresión múltiple escalonada, utilizando distintos métodos de selección (hacia adelante, hacia atrás, paso a paso, etc.) con el fin de obtener una estructura predictiva parsimoniosa y explicativamente robusta.

Definido el modelo final, se aplicó la prueba de causalidad de Granger a las covariables incluidas, con el propósito de determinar si sus valores pasados contribuían significativamente a predecir el comportamiento de la variable

dependiente. Esto permitió diferenciar relaciones contemporáneas de relaciones con poder predictivo temporal.

Para evaluar la precisión del modelo, se aplicó una validación retrospectiva mediante el método hold – out, empleando los años más recientes como conjunto de prueba. Las predicciones generadas fueron contrastadas con los valores reales y se calcularon métricas de error como el error absoluto y el error porcentual absoluto medio (MAPE), permitiendo evaluar la capacidad predictiva antes de su uso prospectivo.

Finalmente, se proyectaron los valores futuros de las covariables significativas utilizando técnicas de series temporales como modelos ARIMA y regresiones no lineales, empleando la plataforma RStudio y Minitab. Las estimaciones obtenidas se introdujeron en las ecuaciones de regresión para proyectar la demanda y la oferta en el horizonte 2023 – 2027. A partir de dichas proyecciones se estimó la demanda insatisfecha — definida como la diferencia entre demanda y oferta proyectadas — y se planteó un escenario de inversión enfocado en cubrir parcialmente esta brecha. Adicionalmente, se estimaron precios futuros mediante técnicas de descomposición de series y referencias internacionales, permitiendo construir un escenario económico integral para la toma de decisiones estratégicas en el sector de bioplásticos.

III. RESULTADOS

Si se requiere analizar el mercado de bioplásticos en el Perú, el análisis resulta complejo, ya que al ser un mercado que recientemente ha tomado impulso, no se cuenta con datos históricos para poder proyectar de manera eficaz y con poco margen de error la demanda y oferta que se espera en los próximos años. Es por tal motivo, que, para realizar el estudio de mercado se toma como cimiento el mercado de los plásticos en el Perú.

En el país, dicho mercado se divide en dos sectores, se tienen a los plásticos en forma primaria y a los productos de plástico. Este primer sector contribuye con un 36,9% a la oferta total de la industria de plástico, y el segundo aporta con un 63,1%, dando a notar, y tal y como muestran los registros, que en el Perú las industrias pertenecientes a este sector están más orientadas a la producción de artículos y materiales de plástico. Un dato curioso es que en el primer sector la industria peruana oferta el 12,8 % y en el segundo el 78,9%; lo restante en ambos casos les compete a las importaciones [7].

A grosso modo se puede notar que, en el primer sector se presenta una potencial oportunidad, puesto que, el Perú al carecer de industrias petroquímicas tiene que importar grandes volúmenes de plásticos en bruto. Por esta razón, el estudio de mercado se centró en este sector, en específico en el polipropileno en forma primaria destinado a la producción de envases. Se toma como referencia el polipropileno (PP),

porque los PHA, en especial el polihidroxibutirato (PHB), surge como una alternativa para sustituir este tipo de poliolefinas, además este biopolímero tiene un gran potencial de desarrollo en el país. Asimismo, el estudio se centra en el sector de envases, porque aparte de representar uno de los mayores destinos de este tipo de productos (10% de lo que se produce tiene como destino dicho sector [8]), es el que tiende acrecentarse mucho más en los próximos años, ya que está fuertemente impulsado por la ley N° 30884 [9].

Ya identificado el punto de partida, se procedió a analizar la variable demanda. Para ello, se tomaron datos de importación, exportación y producción nacional del polipropileno (PP) en bruto destinado a la fabricación de envases. La demanda se definió como la suma de las importaciones más la producción nacional, menos las exportaciones. La Figura 1 muestra la evolución de esta serie temporal entre los años 2010 y 2022.



Fig. 1 Serie temporal de la demanda (2010 – 2022)

Con el fin de comprender su comportamiento temporal, se aplicaron tres pruebas exploratorias: el test de Dickey-Fuller aumentada (ADF), el correlograma (ACF) y el test de Durbin-Watson (DW). El test ADF arrojó un p-valor de 0,705, lo que indica que la demanda no es estacionaria, es decir, presenta una tendencia creciente en el tiempo. Esta tendencia también se aprecia visualmente en la Figura 1 y en el correlograma (Figura 2), que muestra autocorrelación positiva en los primeros rezagos. Por su parte, el valor del test Durbin-Watson (1,30) evidenció autocorrelación positiva en los residuos de una regresión simple contra el año, lo que sugiere que la relación temporal no es capturada adecuadamente por un modelo lineal simple.

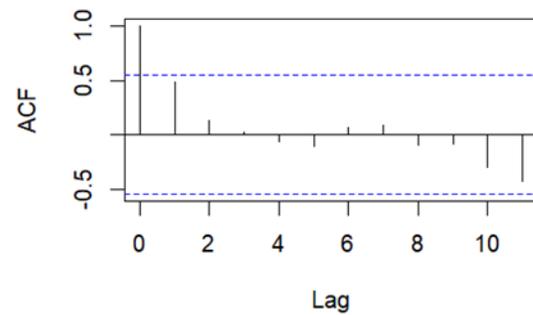


Fig. 2 Correlograma de la demanda histórica

Dado que la demanda presenta una estructura temporal influenciada por factores externos, se descartó el uso de modelos autoregresivos como ARIMA aplicados directamente sobre la variable dependiente. En su lugar, se optó por un enfoque de regresión múltiple escalonada, que permite incorporar tanto la tendencia temporal como covariables económicas que podrían explicar el comportamiento de la demanda.

Por ello, se recolectó información de variables macro y microeconómicas potencialmente influyentes, tales como: el PBI de manufactura no primaria (x2), la inflación (x3), la población (x4), el precio del PP en forma primaria (x5) y la producción industrial de productos químicos, caucho y plástico (x6). Al ser la variable dependiente y las covariables de tipo numérico, el análisis de regresión resultó adecuado como prueba de hipótesis para evaluar la existencia de una relación funcional significativa. En la Tabla I se muestran los datos históricos de cada una de las posibles variables predictoras.

Tabla I
DATA SET DE COVARIABLES Y VARIABLE DE INTERÉS

Año x_1	x_2 (millones de soles)	x_3	x_4 (mil. de personas)	x_5 (S/. por kg)	x_6 (mil. de soles)	Demanda histórica (t) y
2010	46 112	1,50	28,692	5,69	114,4	14 906,15
2011	49 087	3,40	28,905	6,66	122,2	17 300,78
2012	48 820	3,70	29,113	5,81	125,4	19 365,96
2013	50 562	2,80	29,341	6,13	135,7	19 904,46
2014	49 799	3,20	29,616	6,34	138,3	21 430,60
2015	48 553	3,50	29,964	4,97	134,4	22 074,50
2016	47 772	3,60	30,422	4,01	135,5	21 695,82
2017	47 363	2,80	30,973	4,49	131,2	19 862,77
2018	48 955	1,30	31,562	5,33	135,4	22 370,74
2019	49 457	2,10	32,131	4,60	136,7	19 713,83
2020	41 351	1,80	32,625	3,91	128,6	21 890,63
2021	51 772	4,00	33,035	6,14	146,0	25 294,10
2022	52 895	7,90	33,396	6,15	147,7	24 812,08

Se realizaron corridas individuales, es decir, se analizó cada “x” por separado (ej.: $y - x_1, y - x_2, \dots, y - x_6$), y aquellas que arrojaran un valor de p menor a 0,05 serían consideradas significativas, ya que hay suficiente prueba estadística para rechazar la H_0 y afirmar que sí existe una relación significativa entre las variables [10].

Realizado el análisis de regresión (véase la Tabla II) se puede aseverar que el año, la población y la producción industrial tienen influencia en la demanda histórica del PP en forma primaria, sin embargo, el estudio no concluye ahí, puesto que esta información solo es el input para un posterior análisis. Entonces, ya identificadas las variables significativas (x_1, x_4 y x_6), se procedió a elaborar una regresión múltiple escalonada, la cual es una técnica muy utilizada en data science para elaborar modelos predictivos eficientes, dado que se centra en seleccionar las variables más influyentes o relevantes, optimizando y mejorando así la precisión del modelo sin introducir ruido innecesario, además también toma en cuenta la interacción entre las covariables [11].

Tabla II
RESULTADOS DEL ANALISIS DE REGRESIÓN (P VALUE – R^2 – MODELO)

Covariables	P value	Coefficiente de determinación (R^2)	Ecuación o modelo
x_1	0,000	68,45%	$y = -1\ 182\ 110 + 597\ x_1$
x_2	0,214	13,65%	$y = 2\ 953 + 0,367\ x_2$
x_3	0,068	27,10%	$y = 18\ 007 + 878\ x_3$
x_4	0,002	58,89%	$y = -19\ 355 + 0,001306\ x_4$
x_5	0,816	0,51%	$y = 21\ 997 - 218\ x_5$
x_6	0,000	81,79%	$y = -16\ 462 + 279,9\ x_6$

En la Tabla III se presentan los resultados obtenidos a partir de los distintos métodos de regresión escalonada aplicados. Como se aprecia, todos los modelos presentan un coeficiente de determinación ajustado elevado (superior al 80 %), lo que indica un buen nivel de ajuste. Aunque podría optarse por el modelo con el mayor valor de este estadístico, dos de ellos presentan diferencias mínimas en dicho indicador (83,52 % y 83,85 %), las cuales no son significativas desde el punto de vista práctico. El primero corresponde al método de selección hacia adelante (que coincide con el procedimiento paso a paso), mientras que el segundo proviene del método de eliminación hacia atrás. En este estudio, se eligió el modelo del método de selección hacia adelante, en línea con el principio de parsimonia, el cual sugiere preferir modelos más simples cuando su desempeño es comparable al de alternativas más complejas [12].

Adicionalmente, se realizó una evaluación de los residuos del modelo seleccionado, verificándose el cumplimiento de los tres supuestos fundamentales: normalidad, estabilidad e independencia respecto a los valores ajustados. Estos

resultados brindan evidencia estadística suficiente para considerar que el modelo es adecuado para estimar los valores futuros de la demanda.

Tabla III
RESULTADOS DE CADA MÉTODO DE REGRESION ESCALONADA – I

Método	Ecuación de regresión	R^2 ajustado	N° de términos	Términos significativos
Criterios de información hacia adelante	$y = -16\ 462 + 279,9\ x_6$	80,33%	2	2
Paso a paso	$y = -475\ 864 + 233\ x_1 + 203,3\ x_6$	83,52%	3	1
Selección hacia adelante	$y = -475\ 864 + 233\ x_1 + 203,3\ x_6$	83,52%	3	1
Eliminación hacia atrás	$y = 9\ 858\ 730 - 4\ 755\ x_1 - 0,582\ x_4 + 0,000285\ x_1 * x_4$	83,85%	4	2

Una vez seleccionado el modelo de regresión múltiple escalonada, se procedió a evaluar si las covariables significativas que lo componen presentaban dependencia temporal respecto a la demanda. Para ello, se aplicó la prueba de causalidad de Granger, una técnica econométrica que permite determinar si los valores pasados de una variable ayudan a predecir el comportamiento de otra. En el caso específico de la producción industrial (x_6), los resultados indicaron que, si bien esta variable fue estadísticamente significativa dentro del modelo de regresión, no se encontró evidencia de causalidad de Granger respecto a la demanda. Esto sugiere que su influencia es de tipo contemporáneo y no predictiva en el tiempo. No obstante, debido a su alto poder explicativo y robustez dentro del modelo, se decidió mantenerla como parte de la ecuación predictiva.

Con el propósito de comprobar la precisión del modelo, se aplicó una validación retrospectiva mediante el método hold – out (Figura 3). Se utilizaron los datos históricos correspondientes al período 2019 – 2022, los cuales no fueron empleados en la calibración del modelo. La ecuación final seleccionada fue utilizada para estimar los valores de la demanda en dicho intervalo, empleando como entrada los valores reales de las covariables. La comparación entre los valores observados y los estimados permitió calcular el error porcentual absoluto medio (MAPE), obteniéndose un valor de 5,4 %, lo que evidencia que el modelo presenta una alta capacidad predictiva. Este resultado respalda la validez estadística del modelo y justifica su aplicación en la estimación futura de la demanda.

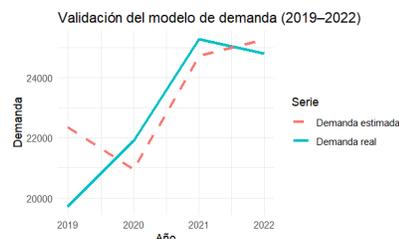


Fig. 3 Validación con el método hold – out

Validado el modelo predictivo, el siguiente paso fue estimar los valores futuros de las variables independientes x_1 y x_6 para el período 2023 – 2027, y posteriormente introducirlos en la fórmula obtenida. En el caso de x_1 , al tratarse del año, sus valores futuros se definieron directamente. Sin embargo, la variable x_6 requirió una estimación más elaborada al no contar con datos proyectados.

Por ello, se aplicaron técnicas avanzadas de series temporales. En primer lugar, se desarrolló un modelo ARIMA (0,1,1) que, aunque resultó estadísticamente viable, presentó un MAPE de 7,37. En búsqueda de mayor precisión, se evaluaron modelos de regresión no lineal, siendo el modelo cúbico el que ofreció mejor ajuste y menor error (MAPE = 2,58), como se muestra en la Figura 4. Debido a su desempeño superior, este último fue utilizado para proyectar los valores de x_6 .

Finalmente, una vez determinados los valores futuros de x_1 y x_6 , estos fueron introducidos en la ecuación de regresión múltiple escalonada, obteniéndose así los valores proyectados de la demanda. En la Tabla IV se presentan los resultados obtenidos, junto con sus respectivos intervalos de predicción al 95 % de confianza.

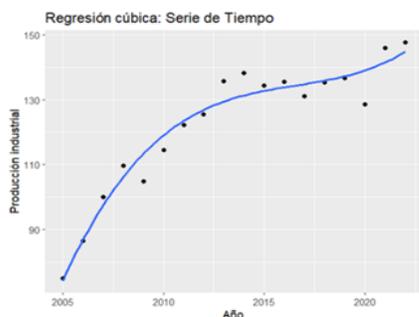


Fig. 4 Diagrama de dispersión – Regresión cúbica

TABLA IV
PRONÓSTICO DE LA DEMANDA

x_1	x_6	Modelo	Demanda proyectada y (t)	Intervalos de confianza al 95%
2023	142,30	$y = -475,864 + 233x_1 + 203,3x_6$	24 299,5	(22 744,4 – 25 854,6)
2024	145,16		25 113,9	(23 425,0 – 26 802,9)
2025	148,88		26 103,3	(24 267,2 – 27 939,4)
2026	153,62		27 300,0	(25 264,3 – 29 335,7)
2027	159,53		28 734,6	(26 391,9 – 31 077,2)

Obtenidas las predicciones de la demanda, se procedió a estimar la oferta, aplicando el mismo procedimiento metodológico. La diferencia radica únicamente en la forma en que se construyen los datos históricos. Para calcular los

valores de la oferta, se utilizaron datos de exportación y de la capacidad del oferente [13], esta última estimada a partir de dos variables clave: la producción nacional y la tasa de utilización de la capacidad instalada del sector plástico. La primera se obtuvo de la base de datos de la SUNAT [14] y la segunda del BCRP [15]. La oferta se definió como la diferencia entre la capacidad del oferente y las exportaciones. La Tabla V presenta los valores históricos de dicha variable y la Figura 5 muestra su serie temporal.

TABLA V
OFERTA HISTÓRICA DEL PP EN FORMA PRIMARIA DESTINADO AL SECTOR DE ENVASES

Año	Capacidad del oferente (t)	Exportaciones (t)	Oferta (t)
2010	3 689,31	381,42	3 307,88
2011	4 012,84	415,27	3 597,57
2012	4 528,15	536,67	3 991,48
2013	4 499,27	683,22	3 816,05
2014	4 435,81	599,43	3 836,38
2015	4 473,52	516,77	3 956,75
2016	4 912,48	621,42	4 291,06
2017	4 183,19	504,13	3 679,06
2018	4 972,72	565,93	4 406,79
2019	3 731,01	684,28	3 046,73
2020	5 865,38	145,78	5 719,60
2021	6 245,97	21,59	6 224,38



Fig. 5 Serie temporal de la oferta (2010 – 2022)

Al igual que en el caso de la demanda, se llevó a cabo un análisis exploratorio de la serie temporal de la oferta. El test de Dickey-Fuller arrojó un p-valor de 0,735, indicando que la serie no es estacionaria. Esta conclusión fue respaldada por el correlograma (Figura 6), que mostró autocorrelación positiva en los primeros rezagos, y por el test de Durbin-Watson, cuyo valor de 1,09 confirmó la presencia de autocorrelación positiva en los residuos de una regresión lineal simple.

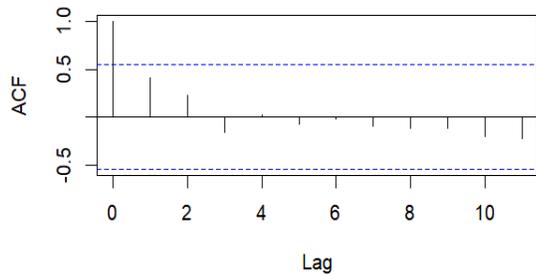


Fig. 6 Correlograma de la demanda histórica

Debido a estos resultados, y en coherencia con el enfoque aplicado a la demanda, se descartó el uso de modelos autoregresivos como ARIMA. En su lugar, se optó por un modelo de regresión múltiple escalonada, que permite capturar la evolución de la oferta en función de variables explicativas relevantes.

Se realizó un análisis de regresión para verificar si alguna covariable, de las vistas anteriormente, tenía relación o no con la oferta. Los resultados arrojaron que x_1 , x_3 , x_4 y x_6 resultaban ser variables significativas. Estas covariables fueron analizadas más a fondo en una regresión múltiple escalonada. De los cuatro métodos existentes, el modelo obtenido mediante el enfoque de eliminación hacia atrás fue el seleccionado para pronosticar los valores de la oferta (véase la Tabla VI). Dicho modelo solo toma en cuenta como variables verdaderamente influyentes a x_1 y x_4 . Se eligió debido al elevado coeficiente de determinación ajustado que presenta y, además, siguiendo el principio de Parsimonia, este modelo resulta más sencillo que el generado mediante el método paso a paso.

TABLA VI
RESULTADOS DE CADA MÉTODO DE REGRESION ESCALONADA – II

Método	Ecuación de regresión	R ² ajustado	Nº de términos	Términos significativos
Criterios de información hacia adelante	$y = 51,5 x_3 * x_3 + 0,000001 x_4 * x_6$	97,37%	2	2
Paso a paso	$y = -0,1389 x_4 - 66,0 x_6 + 0,1898 x_1 * x_1 + 0,00000 x_4 * x_4 + 0,000047 x_1 * x_4$	98,85%	5	4
Selección hacia adelante	$y = 53,7 x_3 * x_3 + 0,000000 x_4 * x_4 + 6,17 x_6 * x_6 - 0,0226 x_1 * x_6 - 0,000052 x_4 * x_6$	98,13%	5	1
Eliminación hacia atrás	$y = -618 x_1 - 0,03910 x_4 + 0,456 x_1 * x_1 + 0,00000 x_4 * x_4$	98,63%	4	3

Con el objetivo de evaluar si la población (x_4), una de las variables explicativas del modelo, presenta una relación temporal directa con la oferta, se aplicó la prueba de causalidad de Granger. El test F, aplicado con dos rezagos, arrojó un p-valor de 0,1305, lo que indica que no existe

evidencia estadísticamente significativa de causalidad temporal entre la población y la oferta. Este hallazgo sugiere que la influencia de la población es de tipo contemporáneo, es decir, explica parte de la variabilidad de la oferta en un mismo periodo, pero no anticipa su comportamiento futuro. Por esta razón, su inclusión en el modelo de regresión se mantiene, debido a su relevancia explicativa dentro del marco predictivo.

Para verificar la capacidad predictiva del modelo de oferta, se aplicó una validación retrospectiva utilizando como conjunto de prueba los años 2019 – 2022. Los valores reales de las variables independientes (x_1 y x_4) fueron introducidos en la ecuación final. La comparación entre los valores observados y los estimados arrojó MAPE de 8%, lo que indica una precisión adecuada para estudios de estimación de mercado. Este nivel de error confirma que el modelo es estadísticamente sólido y confiable para realizar proyecciones en el horizonte de análisis.

En consecuencia, se procedió a estimar los valores de ambas variables para posteriormente introducirlas en la ecuación predictiva. Para la primera (x_1), los valores se definieron directamente, mientras que en el caso de x_4 fue necesario aplicar un modelo ARIMA (2, 4, 2) para pronosticar su evolución futura. Dicho modelo presentó un nivel de precisión altamente satisfactorio, alcanzando un MAPE de 0,69%, lo cual valida su idoneidad para ser utilizado con fines proyectivos (la Figura 7 muestra las medidas de precisión del modelo seleccionado). Una vez obtenidas las estimaciones de x_1 y x_4 , se procedió al cálculo de la oferta proyectada. Los resultados obtenidos, junto con sus intervalos de predicción al 95% de confianza, se detallan en la Tabla VII.

```
> cat("RMSE:", rmse, "\n")
RMSE: 264611.5
> cat("MAE:", mae, "\n")
MAE: 221375.3
> cat("MAPE:", mape, "\n")
MAPE: 0.6881119
```

Fig. 7 Medidas de error del modelo Arima

TABLA VII
PRONÓSTICO DE LA OFERTA

x_1	x_4	Modelo	Oferta proyectada y (t)	Intervalos de confianza al 95%
2023	33 741 637	$y = -619 x_1 - 0,03910 x_4 + 0,456 x_1 * x_1 + 0,00000 x_4 * x_4$	9 279,4	(7 734,78 – 10 824,0)
2024	34 077 996		10 927,2	(8 739,03 – 13 115,3)
2025	34 403 514		12 961,3	(9 783,04 – 15 599,6)
2026	34 715 411		14 553,5	(10 863,6 – 18 243,4)
2027	35 012 495		16 495,6	(11 973,5 – 21 017,8)

Pronosticados los valores futuros de la demanda y oferta, se calculó la demanda insatisfecha proyectada como la diferencia entre ambas variables. Los resultados, presentados

en la Tabla VIII, evidencian altos valores de demanda insatisfecha para el horizonte temporal analizado (2023 – 2027), lo que representa una oportunidad atractiva para posibles inversiones.

Para enriquecer el análisis, se planteó un escenario de proyecto de inversión enfocado en cubrir parcialmente esta demanda insatisfecha. Según Quintero et al. [16] y Guerra [17], una estrategia conservadora es recomendada para empresas que se integran a un mercado emergente. En este caso, se propone una participación inicial del 1,85% en el primer año de operaciones, con un incremento progresivo hasta alcanzar el 3,35% en 2027.

Asimismo, se proyectaron precios de venta como base para estimar los ingresos del proyecto. Debido a la ausencia de datos históricos sobre precios de bioplásticos en Perú, se utilizaron referencias del mercado internacional [18], donde predominan los bioplásticos importados. La estimación de precios empleó el método de descomposición de series (modelo aditivo), que mostró mayor precisión en comparación con el modelo ARIMA y otros modelos de regresión. Se espera que los precios presenten un comportamiento cíclico, con un valor aproximado de S/20,09 por kilogramo en 2024, una disminución en 2025 y 2026, y un repunte en 2027, alcanzando S/22,10 por kilogramo.

como la Ley N° 30884 y la transición hacia una economía circular.

La integración de herramientas de ciencia de datos, como análisis de regresión y modelos avanzados de series temporales, permitió construir modelos predictivos robustos. Estos modelos no solo identifican las variables macro y microeconómicas más relevantes para la demanda de bioplásticos, sino que también ofrecieron proyecciones confiables que pueden orientar estrategias comerciales en mercados emergentes.

Asimismo, se destaca que los precios internacionales de bioplásticos son una referencia válida para el mercado peruano, dado que actualmente la mayoría de estos productos son importados. Este análisis permite plantear escenarios de inversión conservadores, con una participación inicial del 1,85% en 2023, aumentando progresivamente hasta el 3,35% en 2027, demostrando que incluso estrategias moderadas pueden generar retornos significativos.

Finalmente, se recomienda complementar este análisis con estudios técnicos y financieros que evalúen la viabilidad integral de proyectos de producción de bioplásticos en Perú. Este enfoque es esencial para capitalizar las oportunidades del mercado y contribuir al desarrollo sostenible del país.

TABLA VIII

DEMANDA INSATISFECHA Y ESCENARIO DE UN PLAN DE VENTAS

Año	Demanda insatisfecha (t)	Demanda del proyecto – Escenario planteado (t)	Precio proyectado por kg	Ingresos totales
2023	15 020,09	278	S/ 15,97	S/ 4 439 660
2024	14 186,70	305	S/ 20,09	S/ 6 127450
2025	13 412,00	329	S/18,18	S/ 5 981 220
2026	12 746,50	363	S/17,98	S/ 6 526 740
2027	12 239,00	410	S/22,10	S/ 9 061 000

Este escenario hipotético evidencia que el mercado de bioplásticos en Perú es altamente prometedor, impulsado por factores como la Ley N° 30884 y la creciente conciencia de la población hacia una economía sostenible. Una vez demostrada la relevancia del sector, los siguientes pasos serán abordar los aspectos de ingeniería y evaluar la viabilidad financiera del proyecto.

IV. CONCLUSIONES

Este estudio demuestra que el mercado de bioplásticos en Perú tiene un alto potencial de desarrollo, especialmente en el sector de envases, impulsado por la regulación ambiental y el creciente interés en la sostenibilidad. La demanda insatisfecha proyectada para el periodo 2023 – 2027 evidencia una clara oportunidad para nuevas inversiones, reforzada por factores

V. REFERENCIAS

[1] European Bioplastics, “Market – European Bioplastics e.V.” Accessed: Oct. 04, 2023. [Online]. Available: <https://www.european-bioplastics.org/market/>

[2] Christian Flores Acosta, “Plásticos biodegradables: un nuevo paso para su normalización.” Accessed: Oct. 04, 2023. [Online]. Available: <https://boletines.exportemos.pe/recursos/notas/plasticos-biodegradables:-un-nuevo-paso-para-su-normalizacion>

[3] La Cámara de Comercio de Lima, “Importación de envases y recipientes biodegradables creció 176% en 2022, impulsado por ley de plásticos - La Cámara.” Accessed: Sep. 18, 2023. [Online]. Available: <https://lacamara.pe/importacion-de-envases-y-recipientes-biodegradables-crecio-176-en-2022-impulsado-por-ley-de-plasticos/>

[4] E. Korkakaki, M. Mulders, A. Veeken, R. Rozendal, M. C. M. van Loosdrecht, and R. Kleerebezem, “PHA production from the organic fraction of municipal solid waste (OFMSW): Overcoming the inhibitory matrix,” *Water Res*, vol. 96, pp. 74–83, Jun. 2016, doi: 10.1016/J.WATRES.2016.03.033.

[5] G. Papa, T. Pepè Sciarria, A. Carrara, B. Scaglia, G. D’Imporzano, and F. Adani, “Implementing polyhydroxyalkanoates production to anaerobic digestion of organic fraction of municipal solid waste to diversify products and increase total energy recovery,” *Bioresour Technol*, vol. 318, p. 124270, Dec. 2020, doi: 10.1016/J.BIORTECH.2020.124270.

[6] Julia Hereza Atienza, “RES URBIS, generación de bioplásticos a partir de la fracción orgánica municipal,” Barcelona, 2020.

[7] A. Castillo, “Situación actual del sector plástico y perspectivas de la industria plástica,” Nov. 2021

- [8] L. De Cajamarca, A. Juliana, N. Correa, B. Asesor, D. Adolfo, and B. Sigueñas, “Propuesta de producción de bioplástico a partir del suero de leche generado por el clúster de derivados lácteos de Cajamarca,” 2021, Accessed: Oct. 04, 2023. [Online]. Available: <http://tesis.usat.edu.pe/handle/20.500.12423/4052>
- [9] D. Carhuavilca, “Industria del plástico en el Perú,” Nov. 2022.
- [10] L. Nel, *Estadística para ingenieros*, 1st ed. Lima: Editorial Macro, 2019.
- [11] K. Burnham and D. Anderson, “Multimodel inference: Understanding AIC and BIC in model selection,” *Sociol Methods Res*, vol. 33, no. 2, pp. 261–304, 2004.
- [12] J. Hanke and D. Wichern, *Pronósticos en los negocios*, Novena. México: Pearson Educación, 2010.
- [13] M. Casia, *Guía para la preparación y evaluación de proyectos con un enfoque administrativo*, Primera edición. Guatemala: Foundation Cholsamaj, 2018.
- [14] SUNAT, “SUNAT - Acumulado anual Subpartida Nacional/País.” Accessed: Oct. 04, 2023. [Online]. Available: <http://www.aduanet.gob.pe/cl-adi-itestadispartida/resumenPPaisS01Alias?accion=cargarFrmResumenPPais>
- [15] BCRP, “Banco Central de Reserva del Perú.” Accessed: Oct. 04, 2023. [Online]. Available: <https://www.bcrp.gob.pe/>
- [16] H. Nelson, Q. Bertel, G. Uriel, F. Morales, C. Camilo, and C. Duque, “Plan de negocios para la creación de la empresa Miscompetencias.com S.A.S.,” Jun. 2012, Accessed: Nov. 25, 2023. [Online]. Available: <https://repository.universidadean.edu.co/handle/10882/1961>
- [17] A. Belen, G. Salas, and J. E. Rojas Polo, “Estudio de prefactibilidad para la producción y comercialización de bolsas de almidón de papa peruana,” Nov. 2021, Accessed: Nov. 25, 2023. [Online]. Available: <https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/20816>
- [18] Consultoría de Exactitud, “Tamaño del mercado de biopolímeros bioplásticos, región, participación, pronóstico para 2029.” Accessed: Nov. 25, 2023. [Online]. Available: <https://exactitudeconsultancy.com/es/reports/11800/bioplasticos-biopolimeros-mercado/>