



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN

Predicción de ingresos netos y volatilidad de
precios del café en México mediante el uso de
Machine Learning

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
LIC. EN MATEMÁTICAS APLICADAS
Y COMPUTACIÓN

PRESENTA:

ARAIT MONTER CORONA

ASESORA:

DRA. MARISOL VELÁZQUEZ SALAZAR

CIUDAD DE MÉXICO 2024



Este estudio es derivado del proyecto de investigación TEEB Agrifood México (2020), en el que colaboraron: Universidad Nacional Autónoma de México (UNAM), Organización de las Naciones Unidas, CAFECOL, Universidad Panamericana (UP), INECOL, Universidad Autónoma Chapingo (UACh), Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) y la Secretaría de Agricultura y Desarrollo Rural).

Para llevar a cabo esta investigación, se han utilizado los microdatos del INEGI correspondientes al año 2007.

Facultad de Estudios Superiores Acatlán.

Universidad Nacional Autónoma de México.

Dra. Marisol Velázquez Salazar

Asesora de Tesis

Arait Monter Corona

Sustentante

DEDICATORIA

*A mamá y papá, que con su amor y cariño
me han apoyado de manera incondicional
a lo largo de mis estudios, y por lo cual
estoy eternamente agradecida.*

*A Gandhi, mi mayor inspiración académica
y el mejor hermano que pudiera tener.*

Índice

Capítulo	Página
Introducción	IX
1. Marco conceptual: Ingresos de productores y volatilidad de precios del café en México	1
Volatilidad en los precios de café	1
Marco internacional	1
Marco contextual en México	3
Variables de ingresos a productores	5
Marco internacional	5
Marco contextual en México	12
2. Análisis de volatilidad de precios del café en México	15
Análisis histórico de precios	15
Factores meteorológicos y su relación con la producción de café	21
Precipitación Atmosférica	21
Temperatura	22
Relación entre el precio del café, la precipitación y la temperatura	23
3. Evaluación de los ingresos netos del productor de café en México	25
Sección general	25
Características socioeconómicas	26
Datos generales de la parcela	29
Causas naturales	30
Capital social	30
Producción y comercialización	31
Costos de producción	31
Análisis regional: Chiapas, Oaxaca, Puebla y Veracruz	32
4. Predicción de Precios de café utilizando el modelo AutoARIMA	37
Justificación de hiperparámetros del modelo ARIMA	40
5. Predicción de ingresos mediante Machine Learning	41
Descripción de la muestra de datos	41
Modelos que se utilizarán	43
Métricas de evaluación	45
Documentación del proceso de codificación de variables categóricas	46

División del conjunto de datos en entrenamiento y prueba	47
Modelo de Regresión Lineal Múltiple	48
Modelo Random Forest Regressor	50
Modelo Gradient Boosting Regressor	51
Modelo XGBoost Regressor	53
Comparación de modelos de regresión y análisis final	54
Conclusiones	59
Índice de tablas	62
Bibliografía	76

Introducción

El objetivo de esta tesis consiste en la evaluación de los ingresos netos generados durante la producción de café, al tiempo que se identifica su relación con la fluctuación de los precios dentro del sector cafetalero en México, debido a que destaca tanto por la gran contribución económica dentro del país como por el papel crucial que desempeña en la generación de ingresos para las personas productoras. Sin embargo, existen desafíos, como es el caso de los factores climáticos, sociales y económicos que terminan por afectar la estabilidad y rentabilidad de los productores.

Esta investigación se propone explorar los ingresos de los productores de café y su relación con la volatilidad de precios en el mercado. Con esto, se busca proporcionar información clave para la toma de decisiones por parte de los agricultores, así como para la mejora de su rentabilidad a largo plazo en un sector altamente dinámico y competitivo como lo es el de café. La presente investigación, además, intenta robustecer la literatura existente sobre el tema, que hasta la fecha de redacción de la tesis (2024), no es precisamente la más amplia, aunque conviene destacar los trabajos de diferentes autores como Grabs (2023), Pandolph (2020), Hernández (2020), Robles (2022) entre otros, a pesar de la crucial importancia que juega dentro del sector cafetalero. Finalmente, para realizar el estudio, se recurrió a distintos marcos de investigación, procedimientos y literatura que sirvieron como base sólida para abordar los aspectos planteados en relación con las variables que son objeto de indagación en este proyecto.

Pregunta de Investigación

La producción de café en México enfrenta una serie de desafíos complejos que abarcan desde factores climáticos y sociales hasta la influencia de los precios a nivel nacional e internacional. La volatilidad en los precios del café tiene la capacidad de ejercer un impacto profundo en la rentabilidad de los productores y en la sostenibilidad de sus operaciones a largo plazo. En este contexto, la pregunta central de esta investigación es la siguiente: ¿Cómo se comportarán los ingresos netos de los productores mexicanos de café en relación con la volatilidad de los precios en el mercado?

Alcance de la investigación

El alcance de esta investigación se limita a la evaluación de los ingresos netos de los productores de café en México y su relación con la volatilidad de los precios en el periodo que abarca desde 1990 hasta 2023. Un aspecto crucial de esta tesis es la identificación y análisis de las variables que influyen en el precio del café en México. Se investigarán factores como

la oferta y la demanda, los costos de producción y los cambios en el contexto internacional para comprender cómo estas variables afectan los precios y, por ende, los ingresos generados durante el proceso de producción.

Los resultados de esta investigación tendrán implicaciones directas en la toma de decisiones de los productores y contribuirán significativamente a la mejora de la rentabilidad y la sostenibilidad a largo plazo en el sector cafetalero de México. Los descubrimientos de esta investigación tienen el potencial para ser utilizados por responsables políticos y organizaciones del sector para desarrollar políticas y estrategias más efectivas. Al comprender mejor la relación entre la volatilidad de precios y los ingresos de los productores, surge la opción de implementar políticas que proporcionen apoyo durante los períodos de precios bajos y ayuden a los productores a aprovechar los períodos de temporada alta.

Adicionalmente, esta investigación incorpora técnicas de aprendizaje automático (machine learning) para predecir los ingresos netos de los productores de café en México a cinco años. La integración de este tipo de modelos permite generar estimaciones robustas proporcionando una herramienta analítica para anticipar escenarios financieros futuros bajo condiciones constantes. Esta aproximación predictiva complementa el análisis tradicional, permitiendo identificar perfiles de productores con mayor o menor vulnerabilidad frente a la volatilidad de los precios.

Justificación y Planteamiento del Problema

A pesar de la importancia que tiene la producción del café en nuestro país, existe una laguna en la comprensión sobre cómo los ingresos netos de sus productores responden a los cambios en la volatilidad de precios. Esta investigación, por su parte, propone una forma de analizar esta relación y contribuir a la formulación de estrategias más efectivas en el sector cafetalero. Conviene destacar que en la indagación pretende abordar un espacio crítico en la literatura. Hasta la fecha, no se ha llevado a cabo un análisis similar que examine la interconexión entre los ingresos de los productores de café y las variaciones en la volatilidad de los precios en el contexto mexicano.

Objetivos de la Investigación

Esta investigación tiene como objetivo principal formular una conexión entre los ingresos de los productores de café y las variaciones en la volatilidad de los precios (los estudios relacionados se describen dentro del marco teórico). Adicionalmente, se plantean objetivos secundarios. En primer lugar, se busca realizar un análisis de los ingresos netos de los productores de café en distintas regiones de México, considerando la fluctuación de los precios como un factor clave.

En segundo lugar, el propósito es identificar los períodos en los cuales los precios del café experimentan una volatilidad más pronunciada, y evaluar cómo está volatilidad incide directamente en los ingresos de los productores. Debido a que la volatilidad tiene un impacto significativo en la capacidad de los productores para planificar a largo plazo, resulta esencial determinar cuándo se presentan estos períodos y analizar su repercusión en los ingresos.

Finalmente, otro elemento fundamental de este estudio radica en la identificación de las variables que ejercen influencia sobre el precio del café en México. Factores como la oferta y la demanda, las condiciones climáticas, los costos de producción y los cambios en el contexto internacional ejercen un efecto determinante en los precios del café. En consecuencia, resulta crucial llevar a cabo este análisis para comprender de qué manera las variables económicas, naturales, sociales y demográficas inciden en los precios y en los ingresos generados durante el proceso de producción.

Hipótesis de Investigación

Partiendo de la premisa de que la volatilidad de precios en el mercado del café es alta, esta tesis plantea la hipótesis de que, a pesar de las variaciones, los ingresos de los productores en México se ven afectados generalmente de forma negativa, a pesar del incremento de los precios a nivel internacional.

Metodología de la investigación

Esta tesis se centra en evaluar los ingresos netos generados durante la producción de café en diversas regiones de México y su relación con la volatilidad de los precios en el mercado. La exploración detallada de esta relación permitirá proporcionar información clave para la toma de decisiones por parte de los agricultores y contribuirá a mejorar su rentabilidad a largo plazo en el sector cafetalero, que es dinámico y altamente competitivo.

Como se observará durante el capítulo siguiente, los estudios referentes específicamente al caso mexicano de volatilidad de precios de café y su relación con los ingresos netos de productores son limitados, esto resalta aún más la importancia de llevar a cabo esta investigación. Los productores de café, los responsables políticos y las partes interesadas en la industria han operado en gran medida en un entorno de incertidumbre, sin acceso a una base de conocimientos sólida que respalde la toma de decisiones. Esta investigación pretende ayudar a mejorar la comprensión de la dinámica de ingresos en el sector cafetalero, proporcionando datos, análisis, predicciones y conclusiones que permitirían formular proyectos de solución a los desafíos y las oportunidades que enfrentan los productores de café dentro de un mercado volátil.

Para lograr este objetivo, la metodología empleada se basa en un enfoque descriptivo con fundamentos empíricos. Se ha optado por un enfoque cuantitativo, haciendo uso de fuentes de información tanto primarias como secundarias. Para las fuentes primarias, se aplicaron encuestas, mientras que para las secundarias, se realizó un análisis exploratorio de los datos del Censo Agropecuario publicado por INEGI en 2020. El período de análisis abarca desde 1990 hasta 2023.

Los métodos cuantitativos empleados incluyen análisis exploratorio de datos (estadística descriptiva), análisis de series temporales, análisis de regresión de corte transversal y el algoritmo de Random Forest. EL criterio para la selección de dichos modelos está en su capacidad para capturar diferentes relaciones en los datos y mejorar la precisión de las predicciones. Dentro de la investigación se han evaluado cuatro modelos estadísticos: regresión lineal múltiple, XGBoost, árbol de decisión y Random Forest. La regresión lineal múltiple proporciona una visión general para comprender cómo las variables predictoras afectan a

la variable objetivo de manera directa. Este modelo fue seleccionado por su simplicidad y por ofrecer interpretaciones claras sobre la relación entre las variables. Por otro lado, XG-Boost (Extreme Gradient Boosting) es un algoritmo de aprendizaje supervisado que se basa en árboles de decisión y utiliza el método de impulso para mejorar la precisión del modelo (IBM, 2024). Este modelo se eligió debido a su alta precisión y eficiencia, siendo particularmente adecuado para manejar grandes volúmenes de datos y generar predicciones precisas sobre ingresos en función de variables como el precio nacional, relevantes en la producción de café.

Asimismo, se emplea el modelo de árbol de decisión, el cual divide el conjunto de datos en subconjuntos más pequeños según características específicas de las variables predictoras, ofreciendo una representación intuitiva de cómo se toman decisiones en función de múltiples criterios. Por último, el algoritmo Random Forest, una extensión de los árboles de decisión, promedia sus predicciones para mejorar la precisión. Este algoritmo es capaz de manejar grandes conjuntos de datos e identificar automáticamente las características más importantes para la predicción.

Enfoque de la Investigación

La investigación se llevará a cabo siguiendo un enfoque empírico que combina revisión teórica, análisis estadístico y técnicas predictivas para evaluar la relación entre la volatilidad de precios en el mercado del café y los ingresos de los productores en México. Se recopilaron datos históricos detallados sobre los precios del café, ingresos de los productores, costos de producción, volúmenes de venta y factores socioeconómicos relevantes en México, que se han recolectado a lo largo del tiempo (desde 1990 hasta 2023), a través de datos proporcionados por la Organización Internacional de Café y el Servicio de Información Agroalimentaria y Pesquera. Estos datos se analizaron para identificar patrones, correlaciones y tendencias. Se incluye, también, un análisis comparativo de los ingresos en relación con la volatilidad de precios en diferentes períodos para obtener una comprensión detallada de la dinámica.

A través de este análisis estadístico descriptivo se examina la distribución de los ingresos de los productores y la volatilidad de los precios a lo largo del tiempo en los estados de Oaxaca, Veracruz, Chiapas y Puebla. Se calculan medidas de tendencia central y dispersión para comprender la variabilidad en los datos. Se identifican las variables independientes más relevantes que afectan los ingresos y se evalúan su contribución mediante modelos estadísticos. Además, se presentan gráficos y tablas que representan visualmente la relación entre los ingresos y la volatilidad de precios. Esto con el objetivo de proporcionar a los productores de café en México información vital sobre cómo la volatilidad de precios en el mercado impacta sus ingresos. Este conocimiento les permitirá tomar decisiones financieras más informadas y adaptar sus estrategias de producción y venta para mitigar los efectos negativos de las fluctuaciones de precios. Esto es especialmente relevante para los agricultores de pequeña escala que dependen en gran medida de los ingresos del café para su sustento.

Finalmente, los datos y los resultados obtenidos se compararán con los objetivos planteados y las hipótesis de investigación. Se examinarán las relaciones identificadas entre la volatilidad de precios y los ingresos de los productores. Las conclusiones se extraerán basándose en la evidencia recopilada y se presentarán, se proporcionarán recomendaciones específicas para los productores y se destacarán las implicaciones prácticas de los hallazgos. Se identificarán

áreas que requieran una investigación adicional y se propondrán posibles direcciones para futuros estudios. Estas consideraciones ayudarán a enriquecer la comprensión de la industria cafetalera y podrán orientar investigaciones futuras en este campo.

Descripción de las Bases de Datos utilizadas

Para esta investigación, se han empleado datos provenientes de diversas fuentes. En el análisis de precios internacionales, la ICO proporciona información sobre el precio promedio del café para los productores, cubriendo el periodo de 1990 a 2008. Estos datos, expresados en dólares por saco de 60 kilogramos, se presentan en la Tabla 6.2. En particular, se ha analizado la categoría de café "otros suaves" en el caso de México, que se refiere a cafés de calidad suave que no pertenecen a las principales categorías como Arábica o Robusta. Para ello, también se utilizaron datos de la ICO, que abarcan de 1990 a 2019, expresados en centavos de dólar por libra. Además, la Bolsa de Valores de Nueva York fue una fuente clave para analizar tendencias económicas a lo largo de un periodo más amplio, proporcionando datos desde 1990 hasta 2022. Por último, se incluyeron los precios del ICO Composite, que ofrecen información histórica desde 1976 hasta 2005, también reflejados en la Tabla 6.2.

A nivel nacional, se recurrió al Sistema de Información Agroalimentaria y Pesquera (SIAP) para obtener el precio medio rural, con datos que comprenden el periodo de 2003 a 2021. Asimismo, se integraron cifras de precios medios rurales de TEEB Agrifood para México, correspondientes al mismo periodo. Para el análisis de ingresos, los datos se desglosan por municipio, utilizando información proporcionada por el Sistema de Información Agroalimentaria y Pesquera (SIAP). Se han seleccionado los estados de Chiapas (Tabla 6.4), Oaxaca (Tabla 6.5), Puebla (Tabla 6.6) y Veracruz (Tabla 6.7), al ser los principales productores de café en el país. A partir de estos datos, se han obtenido el precio medio rural y el volumen de producción, permitiendo calcular el valor de producción, que se utiliza como una aproximación al ingreso de los productores. Además, se analiza el valor de producción, el volumen de producción, el rendimiento y el precio medio rural con base en la clave de Área Geoestadística Básica (AGEB), como se detalla en la Tabla 6.1.

Para complementar el análisis, se integraron variables naturales, económicas, sociales y climáticas incluidas en la iniciativa TEEB Agrifood. Entre estas variables, se incorporó la precipitación anual acumulada, cuyos datos se obtuvieron del sitio oficial del Sistema Meteorológico Nacional (Tabla 6.3).

Limpieza y calidad de los datos

Para garantizar la calidad y confiabilidad de los datos utilizados en este estudio, se implementaron procesos de limpieza y normalización tanto en las fuentes oficiales como en la información recolectada directamente en campo. En el marco del proyecto TEEB AgriFood, se realizó un trabajo de campo en las regiones cafetaleras de México, donde se llevaron a cabo entrevistas a 192 productores. Estas entrevistas abarcaron siete áreas clave: datos generales, información socioeconómica y familiar; características de las parcelas; prácticas de manejo del cultivo; participación en organizaciones y cooperativas; productos y estrategias de comercialización; y costos de producción. La base de datos generada a partir de estas entrevistas forma parte de la iniciativa Biodiversidad y Agricultura en México (TEEB AgriFood México), que reúne esfuerzos de dos proyectos internacionales orientados a la sostenibilidad

agrícola: The Economics of Ecosystems and Biodiversity (TEEB): Promoting a Sustainable Agriculture and Food Sector, financiado por la Unión Europea y ONU Medio Ambiente, y Integración de la Biodiversidad en la Agricultura Mexicana, implementado por IKI IBA, GIZ, ONU Medio Ambiente y FAO (TEEB AgriFood México, 2023).

En términos de calidad, los datos fueron evaluados bajo tres dimensiones principales, la primera fue validez, asegurando que los datos representen con precisión la realidad que describen. Por ejemplo, se verificaron que los valores de precios y volúmenes estuvieran dentro de rangos históricos y geográficos. La segunda dimensión que se consideró fue integridad, garantizando que los conjuntos de datos estuvieran completos, sin valores críticos faltantes. Registros incompletos, como parcelas sin coordenadas o costos no especificados, fueron corregidos. Por último, la tercera dimensión de calidad de datos fue consistencia, verificando la coherencia entre diferentes fuentes o dentro de la misma base de datos. Por ejemplo, los costos de producción reportados por los productores se compararon con datos promedio de cada región para detectar y resolver discrepancias significativas.

Asimismo, se normalizaron las variables para unificar las unidades de medida y facilitar comparaciones entre datos de diversas fuentes, como precios rurales, rendimientos y costos. Para garantizar la comparabilidad de las variables en la investigación, se utilizó la normalización mediante *Z-score* o puntaje estándar. Este método permite transformar los valores originales en una escala con media cero y desviación estándar uno, lo que facilita el análisis conjunto de variables con diferentes unidades de medida. La ecuación utilizada es:

$$Z = \frac{X - \mu}{\sigma}$$

donde:

- X es el valor original de la variable,
- μ es la media de la variable,
- σ es la desviación estándar.

Viabilidad de la Investigación

Este estudio se sustenta en una necesidad dentro del ámbito de la producción de café en México. Aunque existen investigaciones previas que han abordado diversos aspectos del sector, la interacción entre los ingresos de los productores y la volatilidad de precios aún no ha sido objeto de análisis. Esta carencia hace que esta investigación sea viable y pertinente, ya que tiene el potencial de proporcionar información valiosa para los actores involucrados en la cadena productiva del café.

Evaluación de las Deficiencias

La falta de datos y análisis en torno a cómo los ingresos de los productores de café en México se relacionan con la volatilidad de los precios representa una brecha en la literatura

actual. Por consecuencia, la recolección de datos para analizar representa una posible deficiencia debido al bajo número de productores entrevistados (dentro de esta investigación 192) debido a que viven y sus cultivos se encuentran en zonas de difícil acceso, para comprender mejor la dinámica del mercado y para ayudar a los productores con herramientas más sólidas a enfrentar los desafíos económicos dentro de sus actividades.

La falta de datos y análisis sobre la relación entre los ingresos de los productores de café en México y la volatilidad de los precios representa una brecha en la literatura actual. En consecuencia, la recolección de datos resulta un desafío, pues el número de productores entrevistados (192 en esta investigación) suele ser limitado debido a que las localidades y terrenos de cultivo en zonas de difícil acceso. Esto limita la comprensión completa de la dinámica del mercado y obstaculiza el desarrollo de herramientas efectivas para ayudar a los productores a enfrentar los desafíos económicos en sus actividades.

Estructura de la investigación

Mediante una revisión teórica, metodológica y empírica, la investigación se sumerge en el estudio de las variables que afectan los ingresos netos de los productores, contribuyendo al entendimiento de la dinámica económica y social en el contexto cafetalero mexicano. A continuación, se muestra la presentación del contenido.

Durante el Capítulo I se establecen los cimientos de la investigación, explorando estudios previos y destacando las variables consideradas en investigaciones anteriores. Asimismo, se ofrece una revisión de la literatura que aborda aspectos clave sobre los ingresos y la volatilidad de precios en la industria cafetalera. En el Capítulo II, se analizan los factores asociados a la volatilidad de los precios del café. Se identifican momentos de alta volatilidad a nivel nacional e internacional y se analiza cómo afectan los momentos de volatilidad en el mercado del café. ¿Existen patrones o correlaciones entre ciertas variables y la volatilidad de los precios?

El Capítulo III se adentra en el análisis de los ingresos netos promedio e individuales de los productores de café en México, en los estados de Chiapas, Oaxaca, Puebla y Veracruz. Se estudian además otras variables que pudieran afectar a los ingresos de los productores, divididas en climáticas, naturales, sociales y económicas. Dentro del Capítulo IV se aborda la relación entre ingresos netos y volatilidad de precios mediante el uso de modelos estadísticos creados con aprendizaje automático. Este análisis toma relevancia para que tanto los productores como el público en general, tomen decisiones informadas sobre la comercialización y valorización del café de manera sustentable.

Finalmente, se sintetizan los hallazgos obtenidos y los analiza en relación con los objetivos planteados. Se extraen conclusiones sobre la relación entre los ingresos y la volatilidad de precios en el sector cafetalero mexicano. Además, se identifican temas relevantes para futuras investigaciones, enriqueciendo la comprensión de esta industria y fortaleciéndola para enfrentar desafíos y prosperar en un entorno dinámico.

CAPÍTULO 1 Marco conceptual: Ingresos de productores y volatilidad de precios de café en México

El presente capítulo tiene por objetivo abordar el marco teórico del cual se servirá la investigación, buscando realizar un análisis de la volatilidad de los precios del café y su impacto en el mercado. A través de múltiples perspectivas y autores, se exploran diversas dimensiones de esta problemática que afecta tanto a productores como a consumidores en el contexto mexicano.

Volatilidad en los precios de café

La especulación juega un papel significativo al momento de hablar sobre la volatilidad de los precios del café. Los mercados de futuros permiten a los inversores y especuladores comprar y vender contratos basados en la anticipación de precios futuros. Si bien la especulación es esencial para mantener la liquidez en el mercado, una cantidad excesiva de la misma genera movimientos de precios desconectados de los fundamentos subyacentes, la autorrealización de predicciones basadas en la especulación amplifica las tendencias de precios. Por ejemplo, cuando los especuladores venden futuros en respuesta a predicciones de precios más bajos, teniendo como consecuencia una caída real en los precios (Grabs, 2023).

Marco internacional

Grabs (2023), señala que presentar una volatilidad de corto plazo en los precios del café tiene implicaciones significativas para los productores a pequeña escala. A diferencia de las grandes empresas que esperan a que los precios se estabilicen, los pequeños productores a menudo se ven obligados a vender sus cultivos inmediatamente después de la cosecha para evitar la degradación de la calidad. Además, los grupos de productores y cooperativas están adoptando cada vez más contratos basados en precios de referencia del mercado de futuros, lo que resulta en la posibilidad de exponerlos a variaciones en los precios internacionales.

De acuerdo con el Perfect Daily Grind (2018), las variables que contribuyen a los constantes cambios en el precio del café son diversas, estas variables incluyen condiciones climáticas, brotes de plagas, enfermedades, tipos de cambio de divisas, factores políticos, económicos, especulación, sentimiento del mercado, etc. Además, conviene reiterar que la oferta y la demanda desempeñan un papel crucial en la determinación de los precios del café. Si la oferta de café excede la demanda, los precios tienden a disminuir, y si la demanda excede la oferta, los precios tienden a aumentar.

En el ámbito financiero, los tipos de cambio de divisas son esenciales en la comercialización global del café, el café se comercializa a nivel mundial y estos afectan su precio. Si la moneda de un país productor de café se debilita frente a las principales monedas comerciales, tiende a abaratar las exportaciones de café y potencialmente aumentar la demanda, lo que generará precios más altos. Por el contrario, una moneda más fuerte tiende a encarecer las exportaciones

de café y potencialmente disminuir la demanda, lo que lleva a precios más bajos (Perfect Daily Grind, 2018).

La volatilidad del mercado también afecta los precios del café de varias maneras. Un factor es el uso de instrumentos de cobertura, como los contratos de futuros, que se basan en un precio fijo determinado por la oferta y la demanda. Si se especula que el precio aumentará en el futuro, la gente empieza a comprar contratos de futuros, lo que provocará que el precio actual suba. Por el contrario, si se especula que el precio bajará, la gente tiende a vender café y sus contratos de futuros, lo que provocará una disminución del precio actual. Asimismo, el Perfect Daily Grind, indica también que factores políticos y económicos, como la inestabilidad política, los conflictos civiles y las condiciones económicas en los países productores de café también afectan los precios. Estos factores alteran la producción, el transporte y el comercio, provocando fluctuaciones de precios (Perfect Daily Grind, 2018).

La perspectiva internacional más reciente sobre el tema, proporcionada por la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, por sus siglas en inglés) en 2015, tiene como objetivo ofrecer una comparativa entre países sobre los precios, producción, calidad ambiental y aspectos macroeconómicos de la producción de café. Su análisis estadístico de datos de diferentes años proporciona una visión general de la industria a nivel global; además, resalta la relación entre el crecimiento demográfico y los niveles de ingresos, observando una mayor densidad de población en áreas urbanas y regiones de baja fertilidad (FAO, 2015).

En las últimas cinco décadas, tanto la producción como el consumo de café han experimentado un notable aumento. En la actualidad, más de 70 países participan en la producción de café, pero más del 50 por ciento proviene únicamente de tres naciones (Brasil, Vietnam e Indonesia). Si bien algunos países productores de café han experimentado notables ventajas gracias a mayores rendimientos y un incremento en el volumen de ventas, muchos, especialmente los pequeños productores que constituyen la mayoría en la producción mundial de café, enfrentan desafíos cada vez mayores a raíz del cambio climático y las condiciones más adversas para el cultivo (FAO, 2015).

Durante el mismo informe, se presentó un análisis detallado de diversos indicadores socioeconómicos, macroeconómicos, comerciales, de producción y consumo del café en México que abarca desde 1990 hasta 2014, periodo de 24 años que, aunque no es el estudio más reciente, proporciona una visión panorámica a lo largo de las décadas. Es importante considerar estos factores, ya que para muchas familias mexicanas el café no sólo es una cosecha, sino también un símbolo de identidad y sustento. En muchos casos, las comunidades cafetaleras enfrentan desafíos estructurales, como la falta de acceso a servicios básicos, educación o atención médica, lo que debilita aún más la fuerza laboral local. En el período analizado, la población total experimentó un aumento constante. En 1990, la población era de 86 millones, incrementándose a 104 millones en el año 2000 y alcanzando 124 millones en el último año. En cuanto a la población rural, se observa un ligero aumento de 25 millones en 1990 a 26 millones tanto en 2000 como en 2014 (FAO, 2015).

Adicionalmente, los indicadores de pobreza, a nivel internacional, mostraron una tendencia a la baja. La incidencia de la pobreza medida por aquellos que viven con 1.25 dólares al día disminuyó de 4.8 % en 1990 a 1 % en 2014. De manera similar, la incidencia de la pobreza medida por las líneas de pobreza nacionales disminuyó de 53.1 % en 1990 a 52.3 % en 2014. En relación con la desnutrición, se evidencia una reducción general, con una prevalencia del 6.7 % en 1990, que disminuye a menos del 5 % tanto en 2000 como en 2014. De la misma manera, la esperanza de vida al nacer aumentó a lo largo del período analizado. En 1990, la esperanza de

vida era de 71 años, incrementándose a 74 años en 2000 y llegando a 77 años en 2014. La tasa de finalización de la educación primaria también mostró mejoras significativas, aumentando de 86.6 % en 1990 a 103.2 % en 2014, indicando un mayor acceso a la educación (FAO, 2015).

Marco contextual en México

El análisis de los datos históricos del informe de la Organización Internacional del Café (ICO por sus siglas en inglés, 2023) revela una serie de tendencias interesantes en los datos de suministro de café en México. En 1990, la producción total de café fue de 4,674 miles de sacos de 60 kg. Esta cifra aumentó a 5,300 miles de sacos en 1995, mostrando un incremento en la producción. Sin embargo, en el año 2000, la producción disminuyó ligeramente a 4,814 miles de sacos. A lo largo de la década siguiente, la producción continuó disminuyendo, llegando a 4,001 en 2010 y 2,772 miles de sacos en 2015. Los últimos datos reportados son de 2020, donde hubo un ligero repunte en la producción, alcanzando 3,984 miles de sacos. Mientras tanto, el consumo interno de café en México muestra una tendencia generalmente creciente durante el período analizado. En 1990, el consumo interno fue de 1,374 miles de sacos, y para 1995, había disminuido ligeramente a 1,179 miles de sacos. A partir de entonces, el consumo interno ha aumentado continuamente, llegando a 2,425 miles de sacos en 2020 (ICO, 2023).

La exportación de café en todas sus formas durante el período analizado también presenta variaciones significativas. En 1990, se exportaron 3,683 miles de sacos de café. En 1995, esta cifra se mantuvo relativamente estable en 3,626 miles de sacos. Sin embargo, en el año 2000, hubo un aumento notable en las exportaciones, alcanzando 5,304 miles de sacos. En 2005, las exportaciones disminuyeron significativamente a 1,985 miles de sacos. Luego, en 2010, se observó un aumento a 2,498 miles de sacos. En 2015, las exportaciones se mantuvieron en 2,457 miles de sacos. Finalmente, en 2020, las exportaciones de café en todas sus formas fueron de 2,627 miles de sacos. Estos datos reflejan fluctuaciones en las exportaciones de café a lo largo del tiempo, con un pico en 2000 seguido de cierta variabilidad en los años posteriores. Las exportaciones se mantuvieron relativamente estables en la última década, lo que podría estar relacionado con cambios en la demanda y la oferta en los mercados internacionales (ICO, 2023).

Por otra parte, el Servicio de Información Agroalimentaria y Pesquera (SIAP, 2022) abordó los registros sobre la producción de café en México durante los meses de invierno, donde la cosecha durante enero y febrero representa más de la mitad de la producción anual. De acuerdo con el SIAP, en el año 2022 México experimentó un significativo aumento en su producción anual de café, alcanzando un total de 1.03 millones de toneladas. Esto representa un incremento notable en comparación con las 860 mil toneladas de café producidas durante 2020, reflejando así un crecimiento en la industria cafetalera mexicana. La producción cafetalera en México se caracteriza por dos variedades principales: la robusta y la arábica. La primera, apreciada por su alta productividad, es principalmente empleada en la agroindustria, mientras que la segunda, la variedad arábica, goza de demanda en el sector gourmet y de especialidad (SIAP, 2022).

México experimentó una destacada producción estatal de café, particularmente con los estados de Chiapas, Veracruz, Puebla, Oaxaca y Guerrero liderando esta industria. Chiapas se posicionó como el principal productor, generando 0.39 millones de toneladas, seguido por Veracruz con 0.24 millones de toneladas y Puebla con 0.21 millones de toneladas. Oaxaca y Guerrero también tuvieron su participación en la producción nacional de café, con 0.09 y 0.04 millones de toneladas, respectivamente (SIAP, 2022).

En términos de precios, los estados de Querétaro y Nayarit lideran con los PMR (Precio Medio Rural) más altos, alcanzando los 11,500 y 11,488 pesos por tonelada, respectivamente. Esto

superan notablemente los PMR de Veracruz (7,990) y Guerrero (7,623). Estos valores indican que la producción de café en Querétaro y Nayarit es especialmente rentable en comparación con la de Veracruz y Guerrero. Mientras tanto, a nivel municipal, se destacaron ciertos lugares por su gran producción de café. Motozintla en Chiapas lideró la lista con 43 mil toneladas, seguido de cerca por Tapachula, con 31 mil toneladas y Xicotepec en Puebla, que contribuyó con 23 mil toneladas. Estos municipios resaltan la relevancia de la producción cafetalera en regiones específicas del país. Adicionalmente, de manera internacional, las exportaciones anuales de café en México generan ingresos por 351 millones de dólares, dirigidos a destinos como Estados Unidos, Bélgica, Alemania, Canadá, España, Francia, Italia, Cuba, Suecia, Japón y Reino Unido (SIAP, 2022).

En el ámbito del mercado del café, diversos autores han explorado los determinantes que influyen en la fijación de los precios. Bibiloni y Fuentes (2018) presentan una investigación que tiene como objetivo primordial analizar los factores que rigen el precio del café en el ámbito global, dada su relevancia en el proceso de comercialización. Este estudio se enmarca en la exploración de investigaciones y opiniones previas de diversos autores que abordan esta cuestión crucial. Para lograr su objetivo, los autores emplearon modelos econométricos basados en el método de mínimos cuadrados ordinarios, enfocándose en el periodo de 1978 a 2015. Su hipótesis se cumple con un 90.85 % de influencia de variables como la precipitación, temperatura e importaciones en la determinación del precio del café. Adicionalmente, una variable ficticia explica en un 94.28 % las determinantes del precio de este bien. La metodología adoptada para esta investigación se centra en la regresión lineal simple, buscando establecer relaciones entre las variables analizadas y permitiendo pronosticar comportamientos futuros (Bibiloni y Fuentes, 2018).

Dentro del mismo estudio, se aborda el comportamiento de los precios del café en relación con diferentes variables independientes, como importaciones, precipitación y temperatura. Se empleó el método de regresión lineal simple para analizar las relaciones entre estas variables y destacar su utilidad en la predicción de precios a corto y largo plazo. Finalmente, se implementaron dos modelos debido a un cambio estructural en 2012, y se observa que la temperatura, precipitación e importaciones influyen en el precio del café. Los resultados muestran una relación directa entre el aumento en la precipitación en un milímetro y el incremento de 18,631.82 unidades de moneda local en el precio del café. Por otro lado, la disminución de un grado centígrado en la temperatura se relaciona con un aumento de 196,580.30 unidades de moneda local. En las conclusiones, se destaca que la precipitación, temperatura, alcalinidad del suelo, exportaciones e importaciones influyen en el precio del café. Se prevé un crecimiento en el consumo y producción de café, siempre y cuando el clima favorezca el cultivo. Además de que, con políticas adecuadas, el café tiene potencial para superar al petróleo como producto comercializado de mayor relevancia (Bibiloni y Fuentes, 2018).

Por su parte, la investigación realizada por Hernández y Pandolph (2020) abordó la volatilidad de los precios del café en el contexto de la crisis económica generada por la pandemia de COVID-19. Se observó cómo los precios del café experimentaron repuntes y alta volatilidad en contraste con otros alimentos básicos, teniendo en cuenta que los efectos significativos en la demanda de café también dependen de diversos factores, entre ellos la adopción de medidas de distanciamiento social y la gravedad de la recesión económica global, así como las posibilidades de recuperación a futuro. Dada la dinámica evolutiva y cambiante de estos factores, los autores indican que se abrió una ventana de oportunidad para la concepción y la aplicación de medidas concretas destinadas a mitigar los posibles impactos adversos en la demanda (Hernández y Pandolph, 2020).

Variables de ingresos a productores

El éxito de los productores de café se basa en cuatro aspectos clave: acciones colectivas, políticas públicas, enfoque territorial y capitalización social. La organización de pequeños y medianos productores permite compartir costos, mejorar la calidad del grano y obtener certificaciones. Además, aumenta su influencia en la cadena de comercialización y les brinda la oportunidad de explorar diversas estrategias productivas. Asimismo, el *extensionismo rural* debe fortalecerse y profesionalizarse para brindar una mejor asesoría frente a los desafíos del cambio climático en la producción de café (Akaki y Velázquez, 2020).

Marco internacional

Las estrategias propuestas para mitigar los impactos económicos y financieros derivados de los cambios ambientales incluyen la creación de nichos de mercado y cadenas cortas de comercialización, acciones de adaptación y mitigación (buscar alternativas de ingresos estables durante los períodos de baja rentabilidad en la producción de café) e inversiones por parte de los actores del sector cafetalero (Akaki y Velázquez, 2020).

Una de las diversas propuestas para mejorar los ingresos, la generó Nestlé (2003), la cual implicaba adentrarse en nichos de mercado, como los cafés de especialidad o aquellos con certificaciones orgánicas y de Comercio Justo, esto debido a que los cafés de especialidad, caracterizados por su alta calidad y producción limitada, obtienen precios superiores basados en su origen y calidad de grano, lo que se traduce en mayores ingresos para los productores. La determinación de los precios del café se rige por la dinámica de dos mercados bursátiles principales: LIFFE en Londres y el New York Board of Trade en Nueva York. Estos mercados de futuros desempeñan un papel crucial en la gestión del riesgo y en la fijación de precios a lo largo de la cadena de producción y comercialización, desde el productor hasta el consumidor. Sin embargo, la presencia de fondos de inversión en estos mercados ha introducido una mayor volatilidad, lo que se refleja en un volumen de transacciones que supera en diez veces la producción global (Nestlé, 2003).

En muchas ocasiones, el café se cultiva en regiones remotas y de difícil acceso, lo que implica que la venta de la cosecha se realice vía intermediarios, los cuales pudieran aprovecharse de la situación y revender el producto a precios exorbitantes. Por esta situación, simplificar la cadena de suministro y mejorar la eficiencia se han convertido en enfoques importantes para ayudar a los agricultores a retener un mayor valor de su producción. Para muchos agricultores, la adopción de técnicas agrícolas modernas mejora la calidad de su café y, por lo tanto, sus ingresos. Además, aprender a diversificar sus fuentes de ingresos ayuda a enfrentar las fluctuaciones en los precios del café. A través de la formación y el acceso a nuevas técnicas, a los agricultores se les facilitaría trabajar hacia una agricultura más sostenible y rentable (Nestlé, 2003).

La estrategia propuesta por Nestlé, indicó que para aquellos agricultores que no cumplen con los estándares de calidad necesarios para los cafés de especialidad, es esencial diversificar sus fuentes de ingresos y reducir su dependencia del café. Esto implica explorar otras opciones de cultivos o actividades, aunque cambiar a cultivos alternativos es un desafío debido a las barreras comerciales y los subsidios que dificultan la competencia en igualdad de condiciones con los agricultores de Europa y América del Norte (Nestlé, 2003).

No obstante, el papel de las grandes empresas cafeteras es un tema de debate constante. Aunque las ganancias de estas empresas han crecido, la pobreza de los agricultores también ha aumentado. La relación exacta entre estos dos aspectos es un tanto difícil de determinar.

Los precios del café están influenciados por la oferta y la demanda, la calidad, el clima y las existencias acumuladas, además de la regulación. Sin embargo, existen empresas que han alentado a los países productores a agregar valor al procesar el café localmente, lo que beneficia a las economías locales (Nestlé, 2003).

Las razones para destacar la importancia de la pequeña agricultura, según lo indicado por la FAO durante el Año Internacional de la Agricultura Familiar en 2014, son diversas y abarcan aspectos cruciales. En primer lugar, la agricultura familiar y a pequeña escala se hallan intrínsecamente vinculadas con la seguridad alimentaria global, desempeñan un papel fundamental en la preservación de alimentos tradicionales, promoviendo así una dieta equilibrada y contribuyendo a la protección de la biodiversidad agrícola a nivel mundial. Además, estas formas de agricultura favorecen el uso sostenible de los recursos naturales y representan una valiosa oportunidad para revitalizar las economías locales, sobre todo cuando se complementa con políticas específicas dirigidas a garantizar la protección social y el bienestar de las comunidades (FAO, 2015).

A nivel global, conviene decir, se necesitaría un sistema agrícola que produzca alrededor de 50 % más de alimentos para abastecer a los nueve mil millones de personas que habitarán el planeta en 2050. La aplicación de este sistema garantizaría la disponibilidad de alimentos que satisfagan los requerimientos nutricionales adecuados. Además, es esencial que contribuya a elevar los niveles de ingresos y la capacidad de recuperación económica de la gran mayoría de personas en situación de pobreza a nivel global, considerando que aproximadamente el 75 % de esta población reside en zonas rurales y depende en gran medida de la agricultura como su principal fuente de sustento, adicionalmente este sistema necesita abastecer los servicios ambientales cruciales, tales como la captura y almacenamiento de carbono, la gestión sostenible de las cuencas hidrográficas, y la preservación de la biodiversidad. Se requiere una gestión más eficiente de los limitados recursos hídricos y del suelo disponibles (Robles, 2018).

La Organización de las Naciones Unidas para el Desarrollo Industrial (UNIDO, por sus siglas en inglés) en 2020 presentó las características principales de los Sistemas Alimentarios Sostenibles y del Comercio Justo, las cuales se mencionan a continuación, explorando sus beneficios potenciales y sus limitaciones para productores, consumidores y la sociedad en su conjunto.

1. Desde una perspectiva económica, las iniciativas de Sistemas Alimentarios Sostenibles y Cortos (SFSCs) generan beneficios tanto para los agricultores como para los consumidores. En el caso de los agricultores, estas iniciativas aumentan el valor agregado de sus productos y proporcionar estabilidad a sus negocios, especialmente aquellos que son pequeños o se encuentran en situaciones de pobreza. Esto es particularmente relevante debido a las dificultades que enfrentan para acceder a mercados en cadenas largas y globalizadas bajo condiciones equitativas, además de tener la oportunidad de adquirir productos frescos y diversificados a precios más accesibles, lo que tiene efectos positivos en su salud y bienestar (UNIDO, 2020).

2. La urbanización es uno de los principales factores que separan los lugares de producción agrícola de los de consumo de alimentos, lo que requiere un creciente número de conexiones (transporte, almacenamiento, envasado, procesamiento). Asimismo, tanto el aumento de los ingresos como los cambios en la organización del trabajo y la estructura familiar demandan servicios mejorados incorporados en la alimentación (UNIDO, 2020).

3. Además de la venta de productos agrícolas frescos, los productores encuentran oportunidades adicionales de comercialización directa a través de servicios relacionados con los productos de la finca, como degustaciones y comidas. Uno de los ejemplos más extendidos es el agroturis-

mo, que implica, en su sentido estricto, la introducción de actividades de hospitalidad y catering en la finca, además de la producción agrícola normal. En un sentido más amplio, el agroturismo permite ofrecer una amplia gama de servicios, desde camping hasta rutas gastronómicas y vinícolas, pasando por experiencias laborales en la finca (UNIDO, 2020).

El alto número de etapas y la creciente distancia entre la producción y el consumo son la base de la revolución impulsada por las iniciativas de Cadena de Suministro Corta de Alimentos (SFSCs, por sus siglas en inglés), especialmente en Europa y en los Estados Unidos, aunque se señalan oportunidades interesantes también para otros países, incluyendo los países en desarrollo (Moustier y Renting, 2015).

En la actualidad, existe una gran variedad de iniciativas de SFSCs que adoptan características y métodos de funcionamiento muy diferentes, aunque se inspiran en el mismo principio de reconexión geográfica, económica y social entre la producción y el consumo. La UNIDO cataloga las más relevantes, tanto de forma individual como colectiva. La venta directa de alimentos es una práctica de comercialización tradicional en la que los productores venden sus productos directamente desde sus campos o en ubicaciones cercanas, como puestos en carreteras. En este método, los consumidores se trasladan al lugar de producción, dedicando tiempo y recursos a esta actividad. Los clientes que eligen esta modalidad tienen la libertad de seleccionar la cantidad, variedad y calidad de los productos, además de tener la certeza de adquirir alimentos frescos y de temporada. También brinda la oportunidad a los consumidores de aprender sobre las distintas etapas de la producción agrícola y la estacionalidad, debido a que muchas personas en la actualidad no están familiarizadas con estos aspectos (UNIDO, 2020).

La UNIDO destaca también los mercados de agricultores, otra cadena de suministros de gran relevancia a nivel mundial, son lugares establecidos donde los productores venden sus productos directamente de forma recurrente. Estos mercados suelen contar con una organización común, una imagen compartida y reglas específicas. El uso de espacios públicos a menudo implica el pago de tasas al municipio local y la provisión de servicios como electricidad y limpieza. En ocasiones, se requieren elementos básicos, como mesas de venta, que a veces son proporcionadas por los organizadores. También es necesario llevar a cabo actividades de promoción y comunicación para informar a los consumidores sobre la existencia del mercado (UNIDO, 2020).

Las tiendas de agricultores son establecimientos minoristas gestionados directamente por una o varias explotaciones agrícolas asociadas, donde se venden productos directamente de la finca, mientras que los Grupos de Compra Solidaria y las tiendas gestionadas por los consumidores son grupos de consumidores que compran colectivamente directamente a los productores, siguiendo principios éticos compartidos (UNIDO, 2020).

No obstante, los beneficios de las SFSCs trascienden el ámbito puramente económico y tienen un impacto positivo en lo social y lo ambiental desde diversas perspectivas. Estas iniciativas fomentan la comunicación directa entre productores y consumidores, lo que mejora la difusión de información sobre el producto, el proceso de producción y los actores involucrados, fortaleciendo así los lazos sociales, la solidaridad y la conciencia colectiva. La reducción de la distancia entre el lugar de producción y el lugar de consumo tiende a contribuir a la disminución de las emisiones y el impacto ambiental, y la mayor diversidad de productos ofrecidos en estas iniciativas impulsa la agrobiodiversidad (Moustier y Renting, 2015).

Mientras tanto, los pequeños agricultores de café a nivel global enfrentan una mayor susceptibilidad a las fluctuaciones de precios en el mercado del café, esto debido a diversas razones, tanto financieras, políticas e incluso climáticas. En primer lugar, los productores suelen care-

cer de los recursos financieros y de la infraestructura necesaria para almacenar y procesar sus granos de café. Como resultado, se ven obligados a vender su café inmediatamente después de la cosecha, a menudo en momentos en que los precios son bajos. La falta de instalaciones de almacenamiento y procesamiento limita su capacidad para esperar mejores condiciones del mercado y aprovechar aumentos en los precios (Perfect Daily Grind, 2018).

En segundo lugar, los pequeños agricultores tienen acceso limitado a información sobre los mercados y, en muchos casos, se encuentran aislados de los mercados mundiales del café. Carecen de acceso a datos en tiempo real sobre precios y tendencias del mercado, lo que dificulta su capacidad para tomar decisiones informadas sobre cuándo vender su café. Esta falta de información los coloca en desventaja al negociar precios con los compradores (Perfect Daily Grind, 2018).

Por otra parte, los agricultores a menudo carecen de poder de negociación en la cadena de suministro del café. En ocasiones, dependen de intermediarios o intermediarios que dictan los precios que reciben por su café. Estos intermediarios suelen aprovechar la falta de opciones de los agricultores y ofrecer precios más bajos, lo que agrava aún más su vulnerabilidad a las fluctuaciones de precios (Perfect Daily Grind, 2018).

Además, estos productores son más vulnerables a los impactos del cambio climático, plagas y enfermedades, los cuales afectan significativamente la producción y calidad del café, lo que resulta en menores rendimientos y precios más bajos para su café. Los pequeños agricultores a menudo carecen de los recursos y conocimientos necesarios para mitigar estos riesgos, lo que los hace más susceptibles a las fluctuaciones de precios (Perfect Daily Grind, 2018).

Betancourt (2013) llevó a cabo un análisis profundo sobre la distinción entre valor y precio en el contexto de la industria cafetalera. En su investigación, abordó la paradoja que se presenta cuando los actores más cercanos al consumidor final obtienen mayores ganancias en comparación con los productores de materias primas, quienes a menudo reciben ingresos significativamente más bajos. El propósito principal de su estudio fue determinar si era posible transformar la agricultura para que los productores pudieran ser compensados por la calidad y sostenibilidad de sus productos, en lugar de depender únicamente de la cantidad producida. La metodología empleada se centró en el análisis de las transformaciones y tendencias en la cadena de valor del café, así como en la descripción de los desafíos que enfrentan los productores de café. Además, la investigación propuso iniciativas de Grupo CAFIVER con el fin de abordar esta disparidad (Betancourt, 2013).

Betancourt subraya la brecha existente entre la riqueza generada en los países consumidores y la constante crisis que afecta a los países productores, principalmente debido a los bajos precios de las materias primas. Además, se identificaron modelos de negocio en los cuales el Retorno al Origen (RTO, por sus siglas en inglés) variaba significativamente, dependiendo del tipo de cliente y la estrategia de comercialización adoptada. Este análisis resaltó la importancia de repensar la forma en que se valora y comercializa el café, para garantizar una distribución más equitativa de los beneficios en toda la cadena de suministro (Betancourt, 2013).

De acuerdo con Muñoz-Rodríguez, Gómez-Pérez, Santoyo-Cortés y Rosales-Lechuga (2019), hasta ahora, las variedades de café desarrolladas estaban diseñadas para satisfacer las necesidades del sector cafetalero del siglo XX. Sin embargo, existe una creciente evidencia que sugiere que estas variedades no serán capaces de hacer frente a las amenazas ambientales que el siglo XXI presenta, como los cambios en los patrones climáticos, el aumento de las temperaturas y la proliferación de plagas y enfermedades. Por lo tanto, se promueve el desarrollo de nuevas va-

riedades que resistan estos desafíos relacionados con el cambio climático y las nuevas amenazas (Muñoz-Rodríguez, Gómez-Pérez, Santoyo-Cortés, Rosales-Lechuga, 2019).

El estudio realizado por Fairtrade (True Price, 2017) encontró que solo el 50 % de los productores de café en Indonesia ganan un ingreso promedio suficiente para cubrir una canasta básica que incluye alimentos, vivienda, ropa, salud, transporte, educación, seguridad social, impuestos y pensiones. Además, el 25 % de los caficultores en India, casi el 50 % en Indonesia y Vietnam, y el 100 % en Kenia no logran obtener un ingreso familiar suficiente. Es relevante destacar que este estudio se llevó a cabo entre productores que ya reciben precios superiores a los ofrecidos por el mercado y que forman parte de la visión del movimiento Fairtrade, el cual busca crear un espacio donde los pequeños agricultores disfruten de un estilo de vida seguro y sostenible. Sin embargo, al cierre de 2016, menos del 2 % de la producción mundial, producida por 795 mil productores, se comercializó bajo esta modalidad (Fairtrade International, 2018a).

Por otro lado, el estudio de la organización PROMECAFE en 2018 se enfocó en cuatro países de Centroamérica: Guatemala, El Salvador, Honduras y Costa Rica. Sus resultados concluyen que el cultivo de café en la región no es rentable, debido a que el costo promedio de producción supera al ingreso promedio. Guatemala es el país con la menor rentabilidad, con una pérdida de 71.67 USD por quintal, mientras que Costa Rica registra las pérdidas más bajas, con tan solo 54.37 USD por quintal (PROMECAFE, 2018).

Además, el estudio realizado bajo el modelo de comercio directo en cafés de especialidad revela que a excepción de los caficultores nicaragüenses (quienes siembran en promedio 3 hectáreas de cultivo de café), en el resto de los países no se obtienen ganancias cuando se vende al precio del mercado internacional (Precio C4 de la Bolsa de Nueva York). Ecuador, por ejemplo, con un costo de producción de 1.91 USD por libra, resultaría en una pérdida del 38.7 % para los productores que reciben el precio de referencia. Nicaragua, con un costo de producción ligeramente inferior de 1.05 USD por libra, logra un margen de ganancia del 11.4 %. En contraste, Perú tendría un margen de ganancia negativo del -8.6 %, Guatemala del -16.4 %, Colombia del -1.7 %, y El Salvador del -8.6 % (Tark, 2018b).

En suma, las investigaciones apenas mencionadas plantean la siguiente situación: los pequeños agricultores de café están cultivando a un precio de venta que está considerablemente por debajo del costo de producción. Además, las entidades gubernamentales no han dado prioridad a las acciones necesarias en términos de preservación del medio ambiente o mejoras en la infraestructura. Esta situación perpetúa un ciclo negativo de baja productividad y calidad y en consecuencia, los pequeños productores tienen un control limitado sobre la rentabilidad y sobre su calidad de vida. Este problema se transmite de generación en generación, ya que la falta de garantía de que el negocio cafetero en su nivel más básico sea rentable dificulta la atracción de jóvenes y la promoción de la innovación necesaria para lograr una mayor resiliencia en la industria (Ardila, 2017).

A pesar de que existen al menos 124 especies de café silvestre en todo el mundo, únicamente dos de ellas, el *Coffea arabica* (conocido como arábica) y el *Coffea canephora* (robusta), se utilizan comercialmente. Además, el 60 % de estas especies de café, en su mayoría procedentes de los bosques de Etiopía (el lugar de origen del arábica), están catalogadas como especies en peligro de extinción, en comparación con el 22 % a nivel global para todas las plantas. De este estudio, se encontró que el 45 % de estas especies no se encuentra en ninguna colección de germoplasma, y el 28 % carece de protección en áreas designadas. Esto pone en evidencia la vulnerabilidad de la economía cafetera, que depende exclusivamente de dos especies, en un contexto de cambio climático acelerado y cambios en los patrones de consumo (Davis, 2019).

En este contexto, se ha destacado la importancia de fomentar el consumo global de café, especialmente en los países productores y en los mercados emergentes, como una estrategia para lograr un equilibrio entre la oferta y la demanda, y así garantizar precios justos para los productores de café. Esta perspectiva surgió durante el segundo Foro Mundial de Productores de Café, que tuvo lugar en Brasil en julio de 2019. Sin embargo, esta nueva visión plantea preguntas sobre qué modelos de negocio serán los impulsores de este cambio y quiénes estarán a cargo de llevarlo a cabo. La falta de un compromiso efectivo por parte del sector empresarial en la cadena de valor del café para mejorar los ingresos de los caficultores ha sido un desafío constante, y romper la dependencia de los tradicionales países consumidores representa un desafío importante en la búsqueda de un mercado más justo y sostenible (Muñoz, Gómez, Santoyo y Rosales, 2019).

En el negocio de los tostadores minoristas y bares de café, no se vende únicamente la calidad tangible del café, sino que primordialmente se comercializan los atributos simbólicos de calidad y el servicio personalizado. Este planteamiento presenta un reto importante para los actores de la cadena de valor, que incluyen a los productores de café y sus organizaciones, los extensionistas, los responsables políticos, así como algunas empresas que operan en canales como el retail y el food service. La comprensión y el control de la producción se vuelven esenciales para evitar la problemática de ofrecer productos y mantener márgenes reducidos (Daviron y Ponte, 2005).

Conforme los consumidores mexicanos acceden a una variedad más amplia de tipos de café y métodos de preparación, han mejorado su capacidad para discernir la calidad del café. Esto, junto con el crecimiento de productos premium, ha contribuido al aumento del valor en la categoría de café en el mercado minorista. No importa dónde se encuentren, los consumidores mexicanos buscan un café de mayor calidad, pero en algunas áreas se sienten insatisfechos con las opciones disponibles (El Tiempo, 2017).

A pesar del aumento en la disponibilidad de cafés especiales en el mercado, el cumplimiento de normas voluntarias de sostenibilidad está evolucionando de ser un medio para obtener precios más altos a convertirse en un requisito fundamental para acceder a los mercados. Este cambio plantea el desafío de pasar de un enfoque que se concentra exclusivamente en cumplir con estándares a uno que priorice la calidad y el análisis de costos y beneficios, especialmente para los pequeños productores de café. El objetivo principal es asegurar la sostenibilidad de toda la cadena de valor. Un productor de café alcanzará sus objetivos medioambientales y sociales a largo plazo siempre y cuando sea capaz de competir de manera efectiva, lo que implica obtener precios que cubran sus costos de producción y le permitan obtener un margen de beneficio. (Lamb & Byers, 2017).

Caravela Coffee, líder en el área de cafés de especialidad, ha reconocido que pagar precios elevados solo por una parte de la producción no es económicamente sostenible para los productores de café. Por lo tanto, estableció un modelo de suministro basado en tres pilares esenciales: En primer lugar, la compra del café de calidad que cumpla con estándares, no limitándose únicamente a los microlotes de alta calificación, y ofreciendo precios diferenciados según la calidad, lo que brinda a los productores la oportunidad de maximizar sus ingresos y los motiva a mejorar la calidad de su café. En segundo lugar, establecer relaciones a largo plazo y acuerdos a precio fijo para reducir el riesgo, considerando el tiempo que lleva que las plantas de café comiencen a producir y la inversión requerida. Finalmente, otro pilar importante se trata de proporcionar retroalimentación y asistencia técnica a los productores para mejorar la calidad del café y comprender aspectos financieros relacionados con la producción (Muñoz, Gómez, Santoyo y Rosales, 2019).

Debido a las diversas problemáticas descritas anteriormente, Muñoz, Gómez, Santoyo y Rosales proponen una iniciativa para mejorar la calidad del café y optimizar la generación de ingresos para los productores, cultivando café bajo sombra. Esta alternativa ofrece una ventaja significativa, y es la posibilidad de alcanzar una mayor calidad en la taza. En el mundo del café, existe un amplio consenso en relacionar la producción bajo sombra con una calidad superior. Esta relación se basa en el hecho de que la planta de café se desarrolló de forma natural en la maleza de los bosques tropicales lluviosos, siendo tolerante a la sombra. Se parte de la premisa de que las altitudes más elevadas y las temperaturas ambientales más frescas en los cultivos de sombra promueven una maduración más pausada de los frutos de café y, como resultado, una formación de granos de mayor calidad (Muñoz, Gómez, Santoyo y Rosales, 2019).

Es crucial tener en cuenta que, a pesar de que el cultivo bajo sombra disminuye la productividad del café hasta en un 18% (Muñoz, Gómez, Santoyo y Rosales, 2019), impacta positivamente en el tamaño y la composición de los granos, así como en la calidad de la bebida. Este método retrasa la maduración de los frutos hasta en un mes. En contraposición, los granos de café cultivados a pleno sol presentan contenidos más altos de sacarosa, ácido clorogénico y trigonelina, lo que indica una maduración incompleta y explica la mayor amargura y astringencia en la bebida de café (Muñoz, Gómez, Santoyo y Rosales, 2019).

Además, lo que adquieren los tostadores minoristas, bares de café y consumidores no se limita únicamente a la calidad material del producto. También están comprando los atributos simbólicos de calidad y el servicio personalizado. Para comprender mejor este enfoque, Daviron y Ponte (2005) dividen el concepto de calidad en tres tipos.

El primer tipo es la Calidad Material o Intrínseca, donde la calidad se basa en los atributos físicos y materiales del producto, y no depende de la identidad de quienes lo venden o compran. Esta calidad surge de procesos físicos, químicos o bioquímicos que generan parámetros físicos específicos. Aquí se evalúan aspectos como el aroma, el sabor, la textura y otros atributos directamente relacionados con las características del café en sí. La Calidad Simbólica es el segundo tipo de calidad el cual no se mide con los sentidos humanos ni con dispositivos tecnológicos complejos. Está estrechamente ligada a la reputación del producto y se manifiesta a través de señales de calidad que van más allá de sus propiedades físicas (incluyendo marcas registradas, indicaciones geográficas, etiquetas de sostenibilidad, métodos de procesamiento e identidad étnica de los productores) (Daviron y Ponte, 2005).

El siguiente tipo es la Calidad del Servicio Personalizado, donde las relaciones humanas en el momento de la venta y durante el consumo del producto son cruciales. Existe una interacción entre empleado y consumidor, donde un excelente servicio y una atención cuidadosa desempeñan un papel importante. Esto abarca desde la preparación del café hasta la experiencia que el cliente tiene en una cafetería o al comprar el producto (Daviron y Ponte, 2005).

La comprensión de estos niveles de calidad resulta fundamental en la industria cafetera, ya que posibilita que los productores, tostadores, baristas y consumidores aprecien la abundancia y variedad que este universo ofrece, y supera la simple evaluación del sabor de una taza. Cada uno de estos niveles enriquece la experiencia completa del café y aporta un valor que va más allá de lo puramente sensorial (Muñoz, Gómez, Santoyo y Rosales, 2019).

Actualmente, existe un considerable grupo de consumidores que valora el café de alta calidad cultivado por familias de pequeños productores. en sistemas agroforestales, diversos, multiétnicos y con un enfoque en la neutralidad de carbono. Estos consumidores optan por comprar el cambio que quieren ver en el mundo” (The Nielsen Company, 2019a). Para esta audiencia, la

elección de café, expertamente preparado y proveniente de fuentes sostenibles, se convierte en una valiosa herramienta para consolidar y reflejar sus propias identidades y principios (Sanders, 2017).

Kim & Mauborgne (2005) dividen al sector cafetalero en dos categorías: océanos rojos y océanos azules. Los *océanos rojos* abarcan todos los sectores ya establecidos en el mercado actual, donde los límites de cada sector están claramente delineados y aceptados. En este entorno, las empresas se encuentran inmersas en una competencia por una mayor participación en la demanda. A medida que esta competencia se agudiza, las perspectivas de rentabilidad y crecimiento se ven restringidas, lo que a menudo desencadena una lucha entre productores y una reducción de los márgenes de beneficio. Por otro lado, los *océanos azules* representan áreas de negocio que aún no han sido exploradas, es decir, mercados sin competencia previa. En estos espacios, en lugar de competir por la demanda existente, se genera una nueva demanda, lo que crea múltiples oportunidades para el crecimiento rentable y la innovación. Aún de las formas de generar océanos azules, en algunos casos, se da origen a una industria completamente nueva, como lo hizo la empresa Nespresso con su sistema de monodosis. Sin embargo, en la mayoría de los casos, un océano azul se crea al rediseñar los límites de un sector existente (Kim & Mauborgne, 2005).

Se identifican dos segmentos fundamentales en los que una estrategia y modelo de negocios de este tipo centran su enfoque. El primer segmento está conformado por los consumidores a nivel mundial que aspiran a un cambio positivo, y dentro de este grupo, el 91 % identifica la contaminación ambiental como el problema más urgente que enfrenta el mundo. Como consecuencia de esta conciencia en torno a las preocupaciones medioambientales, casi el 89 % de ellos considera que es de suma importancia tomar medidas para reducir el impacto de las actividades humanas (BBMG, 2015).

El segundo grupo está constituido por los 32 millones de hispanos que viven en los Estados Unidos, de los cuales, veinte millones tienen raíces mexicanas. Según los resultados de la Encuesta de Hábitos de Consumo de Café de la NCA en 2019, este segmento se destaca como el segundo más prominente en cuanto al consumo de café gourmet. En la encuesta, el 46 % de los encuestados informó haber disfrutado de café gourmet el día anterior a la encuesta. Esto los posiciona apenas un punto por debajo de los estadounidenses de origen asiático (47 %), superando a los afroamericanos (40 %) y a los estadounidenses de raza blanca (39 %) en términos de consumo de café de alta calidad. (Almeida, 2019).

Marco contextual en México

En los últimos años, México ha experimentado un crecimiento positivo en el consumo interno de café, con un aumento promedio de 2 kg por persona al año. Sin embargo, este panorama se ve amenazado por el cambio climático, el cual plantea desafíos económicos y financieros de gran magnitud, especialmente para los productores agrícolas y cafetaleros. Además, la liberalización de los mercados internacionales ha alterado la estrategia de México en cuanto al café, ya que grandes empresas de comercialización han asumido el control, dificultando la posición de los productores locales en la cadena de suministro (Akaki y Velázquez, 2020).

Para aumentar las ganancias de los agricultores, Hernández (2020) planteó un análisis simplificado de viabilidad para el desarrollo sustentable en territorios cafetaleros. Su enfoque se centró en crear un modelo simplificado de ganancias diseñado para los productores minifundistas que dependen en gran medida de la mano de obra campesina. En este modelo, se considera la participación de una familia compuesta por tres o más miembros, cada uno desempeñando

roles específicos. Las mujeres contribuyen a través de actividades de diversificación y agregan valor a ciertos productos, mientras que los jóvenes aportan conocimientos en tecnologías de la información. Además, se toma en cuenta el precio de venta del producto y la producción por hectárea, así como inversiones tanto para el manejo de la finca como para una parte de la cosecha de café (Hernández, 2020).

En un enfoque más específico, Robles (2018) realizó un análisis sobre la relevancia de los pequeños y medianos productores del campo mexicano, en el contexto nacional e internacional. Su objetivo principal consistió en comprender cómo alcanzar una organización económica exitosa que beneficie a los productores, refutando la idea de que los pequeños y medianos productores organizados carecen de viabilidad económica, al mismo tiempo, defendió que se debe de abandonar la práctica gubernamental de abordar sus necesidades exclusivamente a través de programas sociales, en lugar de incluirlos en programas de apoyo productivo. Las organizaciones económicas de café en México fundamentan gran parte de sus logros en la formación de equipos técnicos competentes y en la implementación de innovaciones tecnológicas. Es por esto que valoran enormemente estos dos componentes y trabajan constantemente en el desarrollo de sus capacidades. Asimismo, se esfuerzan por compartir sus conocimientos entre los miembros, siguiendo el principio de "de campesino a campesino", y buscan la incorporación de mejoras técnicas, muchas de las cuales están relacionadas con la promoción de prácticas más sustentables en la producción (Robles, 2018).

La importancia de la pequeña y mediana agricultura en el mundo también fue incorporada en la Agenda 2030 para el Desarrollo Sostenible donde se enfatizó la necesidad urgente de brindar a los pequeños productores rurales los recursos necesarios para superar las barreras que obstaculizan su acceso a elementos esenciales como recursos productivos, insumos especializados, conocimientos, financiamiento y mercados. Por tanto, con el propósito de impulsar una transformación rural inclusiva, se requiere la formulación y aplicación de políticas públicas que no solo amplíen las oportunidades disponibles para los pequeños productores rurales como resultado del proceso de cambio estructural, sino que también refuerzen su capacidad para aprovechar dichas oportunidades (Villarreal, 2017).

De acuerdo con la organización Subsidios al Campo (2022), se ha observado un notable crecimiento del 709 % en el número de productores con menos de 5 hectáreas. Esta cifra ha aumentado de 332 mil productores en 1930 a 2.6 millones de unidades, según los registros del último censo Agrícola Ganadero. Este fenómeno se destaca como una característica distintiva de los productores en el país. A pesar de las condiciones precarias para la producción y la falta de apoyos económicos gubernamentales, la pequeña agricultura representa una parte significativa de la producción agropecuaria nacional, contribuyendo con el 39 % del total. Es importante destacar que los pequeños productores desempeñan un papel fundamental como abastecedores de la agroindustria, siendo el café un ejemplo con un promedio de 1.9 hectáreas por productor (Subsidios al Campo, 2022).

Sin embargo, es importante resaltar que en 16 entidades del centro y sur del país donde predominan los productores de pequeña escala de diversos cultivos (Veracruz, Chiapas, Oaxaca, Puebla, Guerrero, Hidalgo, Estado de México, Morelos, Querétaro, Michoacán, Ciudad de México, Tlaxcala, Tabasco, Campeche, Yucatán, Guanajuato), se ha asignado un presupuesto inferior (comparado con las entidades del norte del país, donde el productor es dueño de predios más grandes) con un promedio de 19,000 pesos por productor, en términos per cápita para subsidios productivos y financiamiento. (Subsidios al Campo, 2022).

Con el fin de obtener precios más favorables, los productores mexicanos han explorado la

opción del "mercado justo". En este sentido, la organización Unión Majomut (2023) destaca que el Comercio Justo representa una oportunidad de mercado diferente que, de alguna manera, reconoce el esfuerzo de los productores mediante la fijación de un precio mínimo y un incentivo social superior en comparación con otros sistemas. Los pequeños productores se identifican con el Símbolo de Pequeños Productores, ya que brinda seguridad a los consumidores de que el producto, en este caso, el café, proviene de parcelas cultivadas en armonía con el medio ambiente. Además, el precio de venta de sus productos contribuye directamente al bienestar de sus familias, a la mejora de sus cultivos y a la continua oferta de productos de alta calidad (Robles, 2018).

En su trabajo, *Problemas de la extensión rural en América Latina*, Landini (2016) identificó una serie de problemas comunes y aspectos críticos. Entre los problemas identificados se encontró la importancia en la implementación de estrategias o enfoques de extensión que prioricen la productividad agrícola, la debilidad y la falta de alcance de las organizaciones de café (lo que significa que no están llegando de manera efectiva a todas las áreas rurales), la ausencia de políticas de desarrollo adecuadas, la inestabilidad y fragilidad en la situación laboral, las dificultades en la comunicación con los agricultores debido a diferencias culturales y la politización de la extensión rural, a través de nombramientos de extensionistas basados en afiliaciones políticas y la selección de productores con fines electorales (Landini, 2016).

Debido a esta problemática, los productores han tomado la iniciativa de construir sus propias organizaciones, que estén dispuestas a establecer compromisos con las comunidades cafetaleras. Ejemplos notables de estas organizaciones incluyen la *Unión de Ejidos y Comunidades Forestales Gral. Emiliano Zapata*, *Los Terebintos* y el *Consejo Regional del Café de Coatepec*. Estas iniciativas han tenido un mayor impulso en el aspecto social, administrativo, empresarial y comercial de las zonas (Landini, 2016).

Los pequeños productores representan la unidad de producción predominante tanto a nivel mundial como en México. La acción colectiva de las familias y la organización económica a nivel local o territorial desempeñan un papel crucial al facilitar la articulación de iniciativas, reducir los costos de transacción y aumentar la capacidad de negociación de los pequeños productores. Estas acciones sientan las bases para una colaboración cooperativa a nivel territorial. Sin embargo, también es relevante destacar que la adopción de innovaciones tecnológicas sostenibles no ha sido sencilla para los productores, quienes en muchos casos muestran preferencia por los enfoques tecnológicos tradicionales. Esto ha resultado en una situación en la que un número significativo de socios en las organizaciones todavía optan por producir utilizando paquetes tecnológicos no sostenibles (Robles 2018).

Para concluir esta sección, y para reforzar el marco conceptual ya expuesto, se observa que los ingresos de los productores no han experimentado un aumento sustancial en relación con la volatilidad, a pesar de las variaciones que ésta presenta a lo largo del tiempo. Lo cual permite tener las herramientas teóricas suficientes para responder a la hipótesis general donde la volatilidad de precios de café influye de manera negativa con respecto al ingreso de los productores, afectando así su calidad de vida.

CAPÍTULO 2 Análisis de volatilidad de precios del café en México

La industria del café en México representa un componente vital de nuestra economía y es, en varias ocasiones, fuente principal de ingresos para numerosas comunidades rurales dentro del país. Desde que se tiene reporte, esta industria ha sido testigo de fluctuaciones en los precios de café. Esta sección tiene como objetivo analizar la volatilidad de precios del café en México para determinar si existe una relación significativa entre los ingresos y, en caso afirmativo, comprender la naturaleza y las implicaciones de esta relación.

Durante el capítulo anterior, se ha explorado el contexto histórico y la literatura existente relacionada al análisis de precios e ingresos de los productores. Se concluyó que, debido a la volatilidad en los mercados, los ingresos de los productores no han experimentado un aumento significativo ni sustentable. Este descubrimiento ayuda a observar con detenimiento las razones detrás de esta falta de crecimiento. Además, explorar las mejores prácticas a nivel nacional e internacional, analizar las políticas agrícolas y comerciales, y estudiar las estrategias de gestión de riesgos para proporcionar ideas valiosas y mejorar la situación de los productores en el futuro.

En este capítulo, se ahondará en los datos históricos y contemporáneos para examinar la tendencia de los precios del café según datos proporcionados por la ICO. También se analizará la producción de café en México durante las últimas dos décadas, utilizando datos del SIAP para entender las variaciones y las tendencias temporales, también se respaldará en los indicadores de la Bolsa de Nueva York (NYSE). Los datos se presentan tanto en dólares americanos (USD) como en pesos mexicanos (MXN) dependiendo de la fuente consultada, sin embargo, para brindar claridad al análisis, se realiza una normalización conforme a los criterios establecidos en el apartado de limpieza y calidad de datos de la Introducción. Con ayuda de este análisis, se pretende esclarecer los retos que enfrenta la industria cafetalera mexicana en el contexto actual. Al comprender estos desafíos, abre la posibilidad de desarrollar estrategias basadas en datos para fortalecer el crecimiento económico, proporcionando un sustento justo a los productores y sostenibilidad a largo plazo.

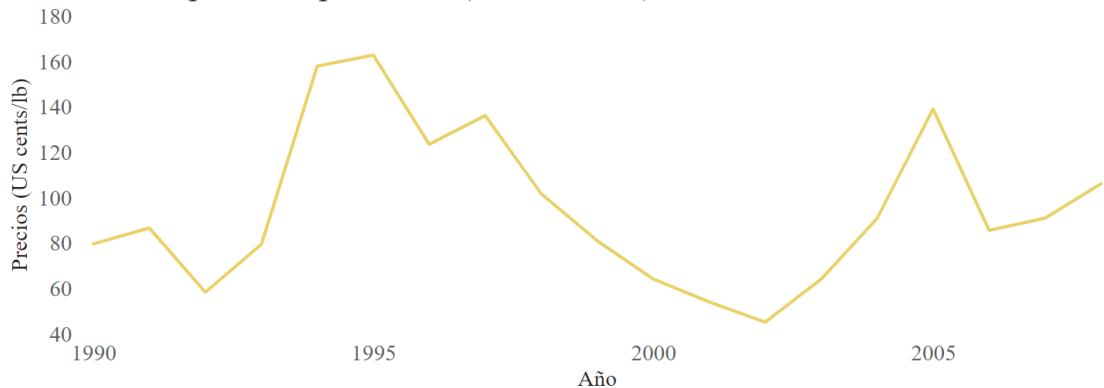
Análisis histórico de precios

La ICO ha recopilado datos sobre los precios a productor desde 1990 hasta 2008. Estos precios están expresados en centavos de dólar por libras de café. El análisis visual de estos datos se representa a través del histograma que se muestra en la Figura 2.1. Al examinar esta gráfica, se visualizan los datos del mercado de café en México, de los cuales se destacan varios aspectos claves. Primero, el precio promedio fue de 95.05 centavos de dólar, el cual representa el punto central alrededor del cual han girado los precios. Sin embargo, la discrepancia entre la media y la mediana (que fue de 86.56) muestra la presencia de precios altos que distorsionan la media hacia arriba (Figura 2.1).

Esta amplia variación de precios, con un rango de 117.71, que va desde el valor mínimo de 45.08 centavos de dólar (precio promedio en el año 2002) hasta 162.79 durante el año 1995, muestra que el mercado del café en México experimenta fluctuaciones considerables a lo largo del tiempo. Estas variaciones, como se observó en el capítulo 1, están influenciadas principalmente por la demanda del mercado internacional, las decisiones de políticas comerciales y por factores climáticos.

Figura 2.1

Promedio de precios a productor (OIC México)



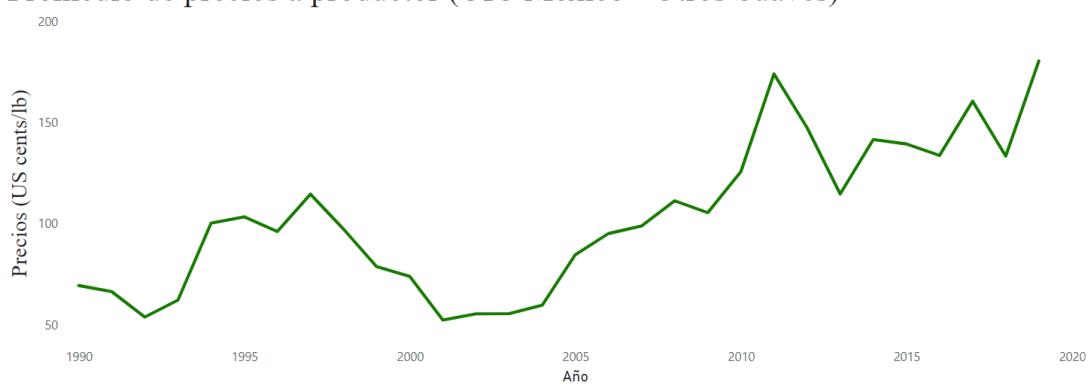
Fuente: Elaboración propia con base en los datos obtenidos de la International Coffee Organization (ICO, 2023)

Por otra parte, existe un sesgo positivo de 0.585, con una distribución sesgada hacia la derecha, lo que significa que hay precios excepcionalmente altos que están influyendo en la forma de la distribución. Además, el coeficiente de variación del 36.34 % confirma la alta variabilidad relativa de los precios en comparación con la media. Esto demuestra que los precios del café son volátiles en relación con su valor promedio, lo que significa desafíos para la planificación financiera y estratégica de los productores y las empresas del sector (Figura 2.1).

De igual manera, analizando los precios del café dentro de la rama “Otros Suaves” entre los cuales se encuentra la mayor parte del café mexicano (datos proporcionados por la ICO durante 2023) exhiben una notable variabilidad, con un rango amplio de 52.27 centavos de dólar a 180.56 centavos de dólar a lo largo de las últimas décadas. Esta fluctuación se ve reflejada en años como 2011, cuando alcanzaron su punto más alto a 174.13, y en 1992, cuando se desplomaron a 53.76 centavos de dólar por libra de café (Figura 2.2).

Figura 2.2

Promedio de precios a productor (OIC México - Otros Suaves)



Fuente: Elaboración propia con base en los datos obtenidos de la International Coffee Organization (ICO, 2023)

La variabilidad de los precios también impacta la estabilidad del mercado. Aunque los precios se mantuvieron relativamente estables alrededor de 100 centavos de dólar entre 2005 y 2008, las fluctuaciones significativas antes y después de este período indican que la estabilidad está sujeta a años de alta variación. A causa de la naturaleza volátil de los precios del café, es esencial que los productores estén preparados para adaptar sus estrategias de cultivo. La utilización de modelos de pronóstico dentro del capítulo 5 de esta investigación proporcionará información valiosa para anticipar cambios en los ingresos e impulsar la toma de decisiones por parte de los productores (Figura 2.2).

Siguiendo con el contexto internacional, se analizan los datos obtenidos por el Composite Indicator Price de la ICO (I-CIP) el cual es uno de los precios de referencia clave de la industria del café en el mundo, utilizado por diferentes sectores para evaluar sus costos y planificar presupuestos (Figura 2.3).

Figura 2.3

Histórico de precios de café - ICO Composite



Fuente: Elaboración propia con base en los datos obtenidos de la International Coffee Organization (ICO, 2023)

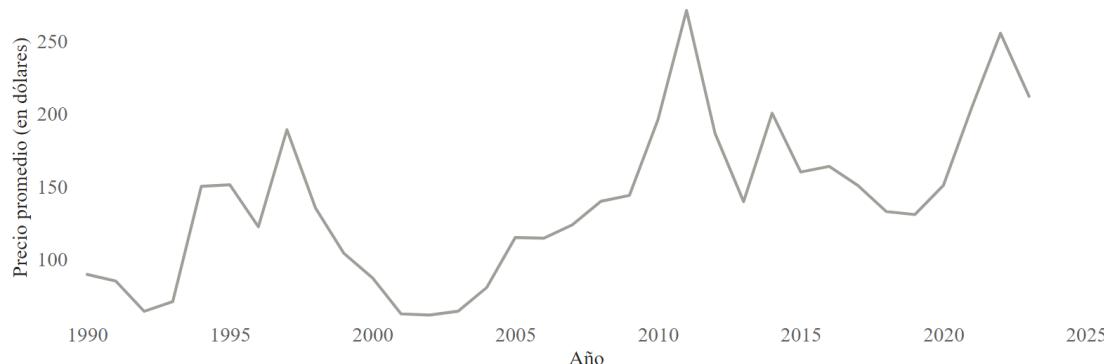
Desde principios de la década de 1990 hasta mediados de la década de 2000, los precios parecen experimentar una serie de altibajos, con algunos años mostrando aumentos significativos, como en 1991 y 2000, mientras que otros años registran caídas pronunciadas, como en 1995 y 2005 (Tabla 6.2). Estas fluctuaciones son atribuibles a una variedad de factores, incluyendo condiciones climáticas, cambios en la demanda y oferta, así como también eventos económicos y políticos a nivel global. A partir de la segunda mitad de la década de 2000, los precios muestran una tendencia general a la baja, con algunos años consecutivos de valores negativos entre 2005 y 2010. Esta disminución podría reflejar cambios en la demanda del mercado, así como también una mayor competencia en la industria del café. Sin embargo, a partir de 2011, los precios se recuperan, con un aumento notable en 2011 y un período relativamente estable en los años siguientes (Figura 2.3).

Mientras tanto, la bolsa de valores de Nueva York (NYSE) también revela una serie de patrones en el mercado del café (Figura 2.4). La media, que se sitúa en 138.44 dólares por sacos de café de 60 kgs, proporciona una visión general del precio promedio del café a nivel

internacional. Como se ilustra en el gráfico, desde 1990 hasta 1993, los precios se mantuvieron en niveles relativamente bajos, durante el período que abarca desde 1995 hasta principios de los 2000 se produjo una caída abrupta, generando una volatilidad significativa. Sin embargo, a inicios del año 2000, se observó un crecimiento constante, indicando una fase de estabilidad económica y se registró una recuperación gradual hasta el año 2009.

Figura 2.4

Indicador Bolsa de Valores



Fuente: *Elaboración propia con base en los datos obtenidos de la Bolsa de Valores de Nueva York (2023)*

La crisis financiera global en 2008 tuvo un impacto negativo en los precios del café, reflejando las dificultades económicas a nivel mundial. En la década siguiente, los precios del café alcanzaron niveles históricos en 2011 y 2022. En 2021, se observó un aumento significativo, llegando a su máximo histórico. Este aumento se atribuye a la recuperación económica tras las restricciones impuestas durante la pandemia. No obstante, en 2022, el índice experimentó otro incremento, alcanzando los 255.41 dólares por costal de 60 kg de café. Finalmente, en 2023, el índice sufrió una notable disminución, cayendo a 212.02 dólares.

El rango de precios, que va desde 61.52 a 271.07 dólares, es notablemente amplio, por lo que ha existido una gran variabilidad en los precios del café a través de los años. La varianza, que nos ayuda a conocer qué tan dispersos están los datos con respecto a la media, confirma esta variabilidad, (siendo de 2,919.28). Además, la desviación estándar de 54.03 muestra la amplitud de esta variación, indicando que los precios del café se desvían de la media.

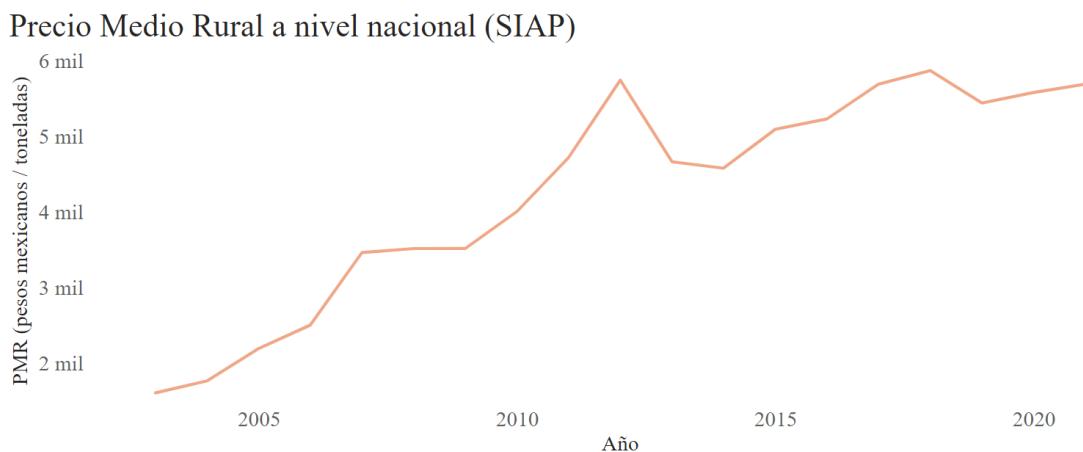
Asimismo, el coeficiente de variación del 39.03 % refuerza la idea de que los precios del café son relativamente volátiles en comparación con su media. Este alto coeficiente muestra que los precios del café tienden a fluctuar considerablemente en relación con su valor promedio, adicionalmente los cuartiles (Q_1 y Q_3) indican que el 25 % de los valores están por debajo de 93.07 dólares y el 75 % están por debajo de 162.83 dólares, respectivamente. Además, el sesgo positivo de 0.511 señala una ligera inclinación hacia la derecha en la distribución de los precios del café. Esto implica que hay una tendencia hacia precios más altos en comparación con los precios más bajos en el mercado, lo que podría estar influenciado por factores como la demanda creciente o limitaciones en la oferta (Figura 2.4).

También dentro del contexto nacional, el SIAP documenta las tendencias y patrones dentro del mercado cafetalero del país (Figura 2.5). Mostrando que a lo largo de la historia se ha obtenido una media de 4,255.30 pesos por tonelada y una mediana de 4,663.05 pesos, también se observa una distribución ligeramente asimétrica hacia la izquierda, indicando que la mayoría

de los datos tienden a agruparse en el rango inferior de los precios.

El rango de precios, que va desde 1,605.30 hasta 5,869.14 pesos, ilustra la amplia variabilidad en los precios del café en México. Esta variabilidad se confirma con una varianza de 2,030,842.28 y una desviación estándar de 1,425.08, subrayando la dispersión de los datos con respecto a la media. Adicionalmente, el sesgo negativo de -0.601 señala una inclinación hacia valores más bajos en la distribución, mostrando que la mayoría de las observaciones se encuentran en el rango inferior de precios.

Figura 2.5



Fuente: Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2023)

Desde 2003 hasta 2021 se observa un crecimiento sostenido en los valores, pasando de 1,605.30 pesos por tonelada en 2003 a 5,688.51 en 2021. Esta evolución subraya un aumento progresivo en los datos representados por el SIAP. A pesar de las fluctuaciones anuales, la tendencia general al alza remarca una estabilidad relativa en el crecimiento anual del precio del café en México. Además, si analizamos el primer cuartil (Q_1) y tercer cuartil (Q_3) observamos que el 25 % de los datos se sitúan por debajo de 3,487.75 pesos por tonelada y el 75 % de los mismos están por debajo de 5,508.91 pesos, respectivamente. Esto muestra que la mayoría de los precios se concentran en este rango, aunque existen ciertas variaciones entre los años 2010 y 2015 (Figura 2.5).

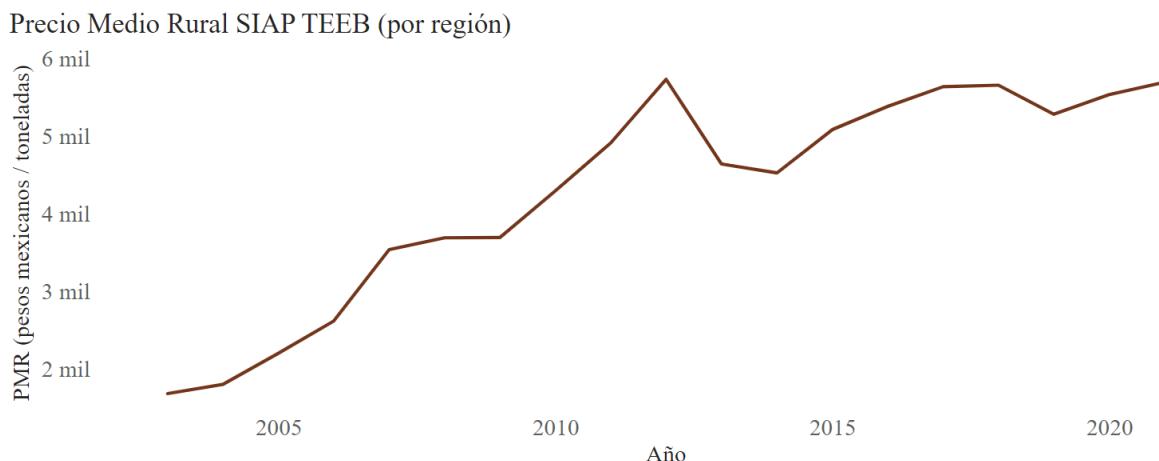
Conviene destacar estos hallazgos con el comportamiento analizado en la figura 2.4 de la Bolsa de Valores, que mostró volatilidad y tendencias complejas en el mismo período, mientras que el SIAP exhibió una tendencia de crecimiento más estable a lo largo del tiempo. El rango extenso de los datos del SIAP, refleja una variabilidad significativa en los niveles de producción o los datos representados. A pesar de esta variabilidad, la tendencia general al alza sugiere cierta estabilidad en el crecimiento anual del precio del café (Figura 2.5).

Por otra parte, en el marco de la iniciativa "La Economía de los Ecosistemas y la Biodiversidad para el caso del café en México" (TEEB, por sus siglas en inglés), diseñada en 2020 para realizar una evaluación integral y económica detallada de los sistemas ecoagroalimentarios, se busca mejorar la comprensión entre los tomadores de decisiones y los actores clave sobre las interacciones económicas y dependencias entre la producción de café y los servicios ecosistémicos, explorando su valor para la sociedad en el contexto mexicano.

Según el análisis del TEEB en las regiones cafetaleras de Chiapas, Veracruz, Oaxaca y

Puebla, se observa un crecimiento constante en los precios del café en México desde 2003 hasta 2021, pasando de 1,680.47 pesos por tonelada a 5,691.02 (Figura 2.6). Aunque se perciben fluctuaciones anuales, la tendencia general es ascendente, planteando un aumento constante en la producción o en los datos representados por el SIAP con la metodología TEEB.

Figura 2.6

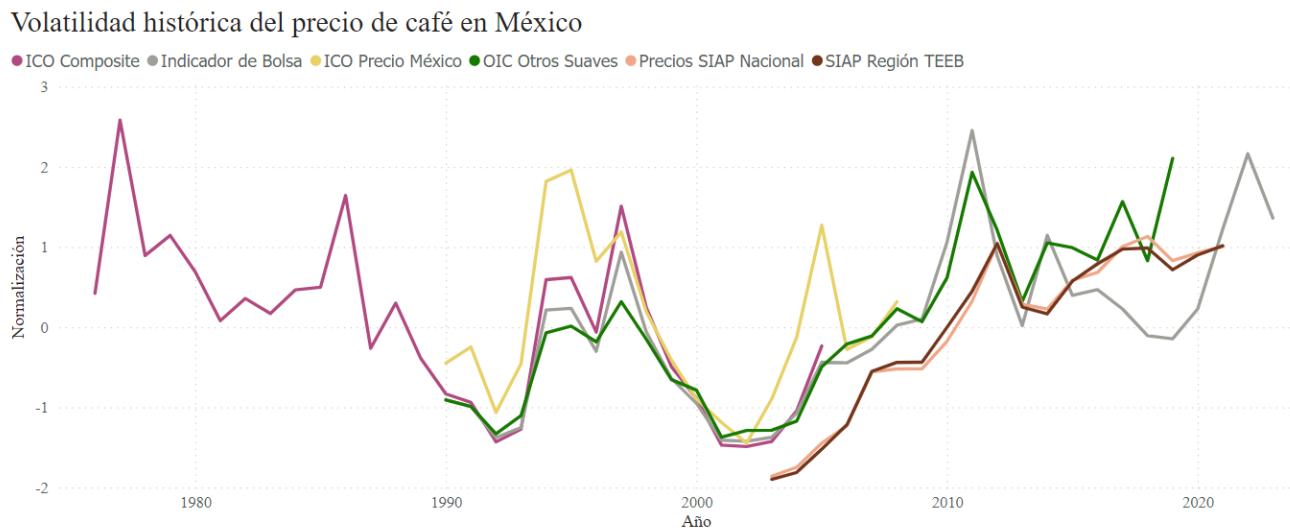


Fuente: Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP TEEB, 2023)

La distribución de los datos exhibe un sesgo negativo, sugiriendo una inclinación leve hacia valores más bajos. Los cuartiles (Q_1 y Q_3) señalan que el 25 % de los valores están por debajo de 3,611.25 pesos por tonelada, mientras que el 75 % están por debajo de 5,455.56 pesos por tonelada, proporcionando una comprensión más precisa de la distribución de los precios del café en México durante este periodo.

Para finalizar este análisis de precios del café a nivel nacional e internacional a lo largo de la historia, en la Figura 3.7, se presenta la comparación normalizada de los precios a productores de las diversas fuentes analizadas anteriormente: la base de la OIC a nivel nacional y la referente para la selección de Otros Suaves, el indicador de bolsa NYSE, los datos recopilados del SIAP a nivel Nacional y el SIAP TEEB a nivel regional.

Figura 2.7



Fuente: *Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2023)*

Como se logra apreciar, es notable la presencia de valores atípicos, pero se identifica una clara tendencia en los precios donde durante la década de 1990 hubo una alza de precios, mientras que a inicios de los 2000 alcanzan mínimos históricos. Como conclusión, los precios del café (a excepción de los proporcionados a productor publicados por la Organización Internacional del Café) tienen un comportamiento similar tanto a nivel nacional como internacional. En este sentido, los datos de precios de café proporcionados por la Bolsa de Valores, serán los utilizados como base para las predicciones futuras.

Factores meteorológicos y su relación con la producción de café

A lo largo de los años, diversos factores han influido en la industria cafetalera, desde las condiciones climáticas hasta los cambios en la demanda global. A continuación se examinará la evolución de los precios del café en México con relación a las diferentes variables climáticas que podrían afectar la volatilidad: la precipitación promedio anual, la cantidad de ciclones tropicales que se producen en los océanos Pacífico y Atlántico y la temperatura promedio de los estados con mayor producción de café a nivel nacional (datos recopilados desde 1990 hasta 2023). Realizando este enfoque, se visualizará de manera concisa las dinámicas del mercado cafetalero mexicano, para proporcionar una herramienta que muestre los posibles desafíos y oportunidades, brindando así a los productores la capacidad de planificar de manera más informada.

Precipitación atmosférica

La precipitación ejerce una influencia sustancial en la producción cafetalera, ya que el café, siendo una planta altamente sensible a las condiciones climáticas, requiere un suministro adecuado de agua para su óptimo desarrollo. Las formas en que esta precipitación incide en la producción son diversas. En primer lugar, el desarrollo del fruto es un proceso crucial para la planta de café, y la precipitación desempeña un papel esencial en las fases críticas de floración y desarrollo del grano. La ausencia de agua durante estos períodos tiene un impacto negativo en la formación de los granos, disminuyendo tanto la calidad como la cantidad de la cosecha.

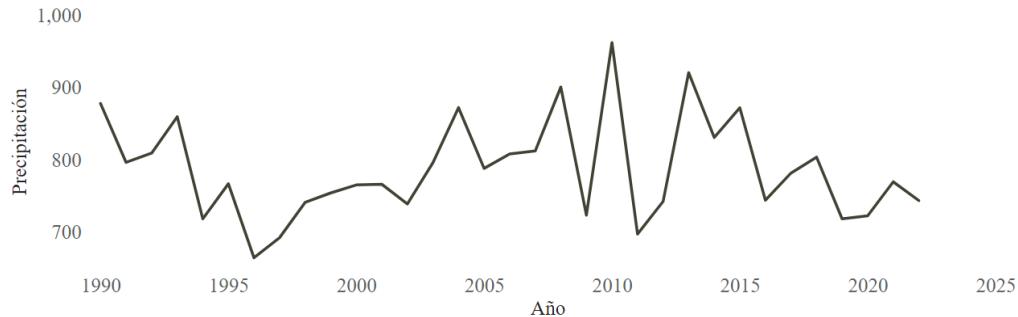
La calidad en el grano de café se ve directamente afectada por la distribución de la precipitación, como lo demostró Bustamante en un estudio realizado en la región oriental de Cuba donde se evaluó el efecto de las lluvias sobre el rendimiento del cultivo. Los resultados indicaron que precipitaciones anuales entre 1,750 y 1,900 mm permiten alcanzar rendimientos de 1.4 a 1.8 toneladas por hectárea de café, mientras que al superar los 2,000 mm, los rendimientos llegan a 2 toneladas por hectárea (Bustamante, 2015). Un exceso de lluvia provoca la lixiviación de nutrientes del suelo y afectar negativamente la acidez del café. En contraste, la falta de lluvia reduce el tamaño de los granos y afecta la densidad del café, influyendo en la calidad del producto final.

Las condiciones húmedas asociadas con altas precipitaciones aumentan el riesgo de enfermedades fúngicas, como la roya, daña las hojas de la planta de café y afecta el rendimiento de los cultivos. Además, la maduración irregular de los frutos, resultado de precipitaciones inadecuadas, da lugar a una cosecha desigual, complicando la recolección y afectando la calidad general del café. En áreas montañosas donde se cultiva café, las precipitaciones intensas desencadenan deslaves y erosión del suelo, causando daños a los cultivos, afectando la calidad del suelo y, en última instancia, impactando la producción a largo plazo (Bustamante, 2015).

Los ciclos climáticos extremos relacionados con la precipitación, como sequías e inundacio-

nes, suelen tener consecuencias significativas en la producción de café al afectar la disponibilidad de agua y nutrientes en el suelo, así como provocar la pérdida de cultivos. La comprensión detallada de estos efectos es esencial para implementar estrategias de manejo y adaptación que fortalezcan la resiliencia de la producción cafetalera frente a los desafíos climáticos. A continuación, se muestra el histórico de la media de precipitación anual en México desde 1990 hasta 2023, para posteriormente, ponerlo en contraste con la temperatura media anual y el precio del café a nivel nacional.

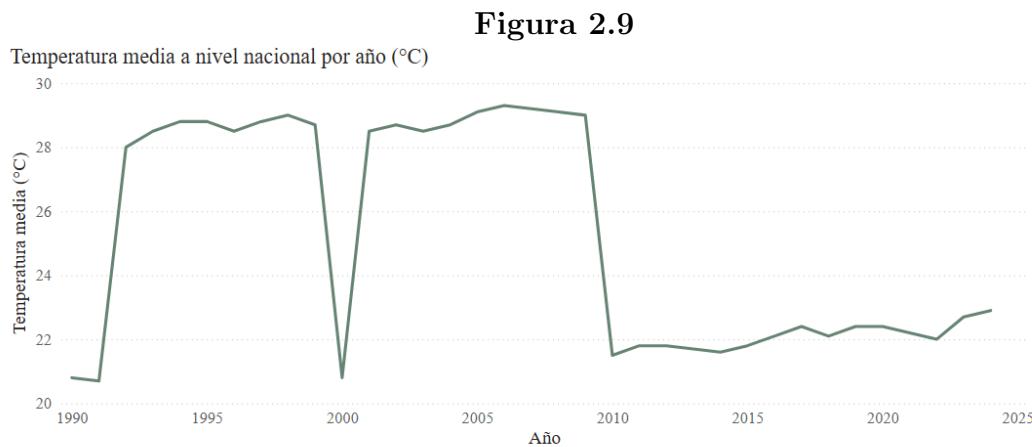
Figura 2.8
Precipitación anual nacional (1990-2023)



Fuente: Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2023)

Temperatura

Las zonas cafetaleras de México varían en términos de altitud y ubicación geográfica, lo que afecta las temperaturas. Sin embargo, en general, las regiones cafetaleras de México suelen tener climas templados a cálidos debido a su ubicación en latitudes tropicales y subtropicales. A continuación, se presenta de forma visual la información histórica sobre la temperatura media anual en México (Figura 3.9) extraída de los Resúmenes Mensuales de Temperatura del Servicio Meteorológico Nacional (CONAGUA, 2025) datos que están disponibles para su consulta en la Tabla 6.8 del Anexo.



Fuente: Elaboración propia con base en los datos obtenidos del Servicio Meteorológico Nacional (CONAGUA, 2025)

A través de los 35 años presentados (1990 a 2024) el rango de temperatura media nacional abarca con un mínimo de 20.7 °C en 1991 alcanzando un máximo de 29.3 en 2006. En promedio,

la temperatura nacional de este periodo es de 24.12 °C. Los datos revelan dos patrones de temperatura distintos. El primero abarca desde 1990 hasta 1999, donde las temperaturas fueron altas, fluctuando entre 20.7°C y 29.0°C, con un pico notable entre 1992 y 1999, donde las temperaturas se mantuvieron consistentemente alrededor de 28.5°C en promedio. Sin embargo, el segundo patrón de temperatura aparece en el año 2000, donde se observa una caída de hasta 20.8°C, y a partir de ese año, las temperaturas se mantuvieron estables, entre 20.8°C y 22.7°C. En años recientes, se logra observar una tendencia al alza, con temperaturas que aumentaron de 22.1°C en 2016 a 22.9°C en 2024.

Relación entre el precio del Café, la precipitación y la temperatura a nivel nacional

Dado que las variables climáticas juegan un papel al momento de que los productores definen el precio del café, a continuación se desarrolla un modelo multivariado de regresión lineal múltiple con el propósito de analizar si existe una relación entre la precipitación, la temperatura y el precio del café. Para ello, se emplea el histórico de precios tomando como referencia el índice de la Bolsa de Nueva York (NYSE), el cual refleja la evolución de los precios a lo largo del tiempo desde 1990 (Figura 2.4).

```
[language=Python, label={lst:split}]
import numpy as np
import pandas as pd
import statsmodels.api as sm

# Simulacion de datos
np.random.seed(42)
n = 100

# Variables independientes (precipitacion y temperatura)
precipitacion = np.random.normal(1000, 100, n)
temperatura = np.random.normal(20, 2, n)

# Variable dependiente: ingreso neto (con ruido)
y = (
    0.0086
    + (-0.2679) * ((precipitacion - np.mean(precipitacion)) / 1000)
    + (-0.4634) * ((temperatura - np.mean(temperatura)) / 10)
    + np.random.normal(0, 0.2, n)
)

# Construccion del DataFrame
df = pd.DataFrame({
    'Precipitacion_Anual': precipitacion,
    'Temperatura_Media': temperatura,
    'Ingreso_Neto': y
})

# Preparacion para regresion: estandarizar variables
X = df[['Precipitacion_Anual', 'Temperatura_Media']]
X = (X - X.mean()) / X.std() # para que el error estandar sea comparable
X = sm.add_constant(X) # agregar intercepto
y = df['Ingreso_Neto']

# Ajuste del modelo de regresion
modelo = sm.OLS(y, X).fit()

# Resumen del modelo
print(modelo.summary())
```

Los datos utilizados para la construcción del modelo se encuentran detallados en el Anexo de tablas (Tabla 6.3). Cabe mencionar que los datos han sido sometidos a un proceso de normalización, siguiendo los lineamientos de limpieza y calidad de datos descritos anteriormente en la sección de metodología. Como resultado de este procedimiento, se obtienen los siguientes resultados (Tabla 2.1).

Tabla 2.1: Resultados del Modelo de Regresión Múltiple

Variable	Coeficiente	Error Estándar	Valor p
Intercepción	0.0086	0.154	0.956
Precipitación Anual	-0.2679	0.155	0.094
Temperatura Media	-0.4634	0.154	0.005
Estadísticos del Modelo			
R^2	0.262		
R^2 Ajustado	0.212		
Estadístico F	5.313		
Prob (F -stat)	0.0106		
Durbin-Watson	0.661		

El análisis de la regresión lineal múltiple revela que el modelo tiene un coeficiente de determinación (R^2) de 0.262, lo que indica que el 26.2 % de la variabilidad en el precio del café a lo largo de los años se explica tanto por la precipitación como por la temperatura media anual.

En cuanto a los efectos individuales de las variables, se observa que la temperatura media tiene un impacto negativo significativo en el precio del café ($p = 0,005$). Es decir, un aumento en la temperatura está asociado con una reducción en el precio del café, lo que se debe a cambios en la producción o en la calidad del grano . Por otro lado, la precipitación anual también muestra una relación negativa con el precio del café, aunque este efecto no es estadísticamente significativo al nivel del 5 % ($p = 0,094$), por lo que su impacto es menos concluyente dentro del conjunto de datos analizados.

CAPÍTULO 3 Evaluación de los ingresos netos del productor de café en México

A lo largo de este capítulo, se presenta el panorama de la producción de café en México, mediante un análisis detallado de la investigación llevada a cabo en 2021 por el programa de la Economía de los Ecosistemas y la Biodiversidad para la Agricultura y los Alimentos (TEEB AgriFood, por sus siglas en inglés) para el caso del café en México, perteneciente a la ONU. El estudio, desarrollado en colaboración con diversas instituciones académicas del país, se basa en la metodología de análisis de cuatro capitales: social, humano, natural y de producción, vinculados a la producción de café.

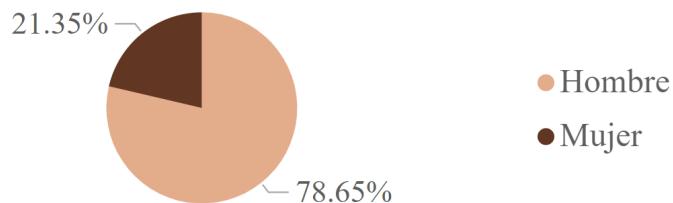
La investigación se llevó a cabo a través de encuestas aplicadas por la iniciativa TEEB Agrifood Café México en el periodo de 2020 a 2023 a 192 productores mexicanos estratégicamente seleccionados en las regiones de mayor relevancia en la producción cafetalera de los últimos años: Chiapas, Veracruz, Puebla y Oaxaca. En este capítulo, se mostrará un análisis de los datos recopilados, explorando también los ingresos percibidos por estos productores y su relación con el precio del café (TEEB Agrifood Café México, 2023).

El estudio se estructuró en diversas secciones, durante ésta tesis se analizaron las áreas: demográfica, socioeconómica, los datos generales de la parcela de cultivo, los datos organizativos del productor (evaluando su presencia en asociaciones, cooperativas u otros apoyos del gobierno o entidades privadas) y la sección dedicada a producción y comercialización (con especial énfasis en los ingresos netos de los productores).

Sección general

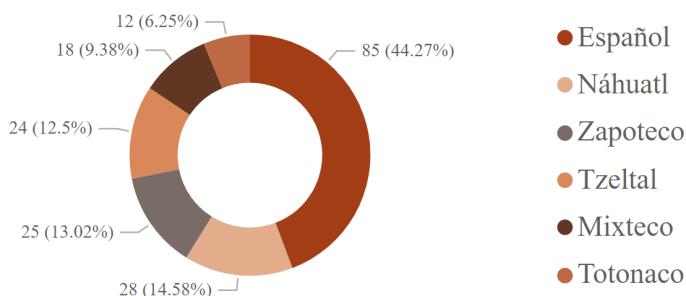
Del grupo analizado, compuesto por 192 productores originarios de diferentes localidades dedicadas al cultivo de café en México, se observa una marcada predominancia masculina, con 151 hombres dedicándose al cultivo del café, representando un 78.65 % del total, en contraste con el 21.35 % correspondiente a las mujeres. Esta disparidad de género, aunque común en los sectores agrícolas del país, resalta la importancia de promover la inclusión de mujeres en las actividades cafetaleras dentro del país (Figura 3.1).

En el análisis por grupos de edad, se destaca que 12 personas, constituyendo el 6.25 %, se encuentran en la franja etaria de 20 a 30 años. Asimismo, 22 personas, representando el 11.46 %, se sitúan en el rango de 30 a 40 años. La población de 40 a 50 años abarca a 47 personas, conformando un 24.48 % del total. Aquellos de 50 a 60 años suman 48 productores, alcanzando así un 25 %. Finalmente, la categoría de personas mayores de 60 años comprende a 63 individuos, constituyendo el 32.81 % del total y mostrando que las personas de edad mayor normalmente son las encargadas del cultivo y siembra del café.

Figura 3.1**Distribución por sexo de la persona productora**

Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEBC México (2021)

Es importante destacar que, de las 192 personas encuestadas, sólo 24 hablan exclusivamente español, grupo inmerso en las 85 personas que practican el idioma. Esta información muestra la variedad cultural que existe en las comunidades cafetaleras de México. Se observa un marcado bilingüismo entre los encuestados, donde el 55.21 % de la población (107 de 192 personas) afirmó tener conocimientos de al menos una lengua originaria. De estos, 18 personas hablan mixteco, lo que representa el 9.38 % del total. La lengua náhuatl es hablada por 28 productores, constituyendo el 14.58 %, 12 personas se comunican en totonaco, representando el 6.25 %. Asimismo, 24 productores hablan tzeltal, abarcando el 12.5 %, y finalmente, 25 personas hablan zapoteco, representando el 13.02 %. Destacando así, cómo el café está vinculado a las raíces indígenas, ya que muchas personas hablan diferentes lenguas originarias, preservando su riqueza cultural (Figura 3.2).

Figura 3.2**Lengua hablada por el productor**

Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEBC México (2021)

Aunado a esto, se observa una diversidad en el nivel educativo. El 9.90 % (19 personas) indicaron que poseen estudios universitarios, mientras que el 8.85 % (17 personas) no cuentan con ningún tipo de escolaridad formal. Por otro lado, un 22.40 % (43 personas) alcanzaron el nivel de educación secundaria, el 10.42 % (20 personas) lograron obtener el título de preparatoria y el 41.67 % (80 personas, un poco menos de la mitad de la población) completaron la educación primaria, el resto (6.76 %) prefirió no responder. lo cual refleja la importancia de considerar iniciativas que impulsen la educación en las zonas con mayor producción cafetalera.

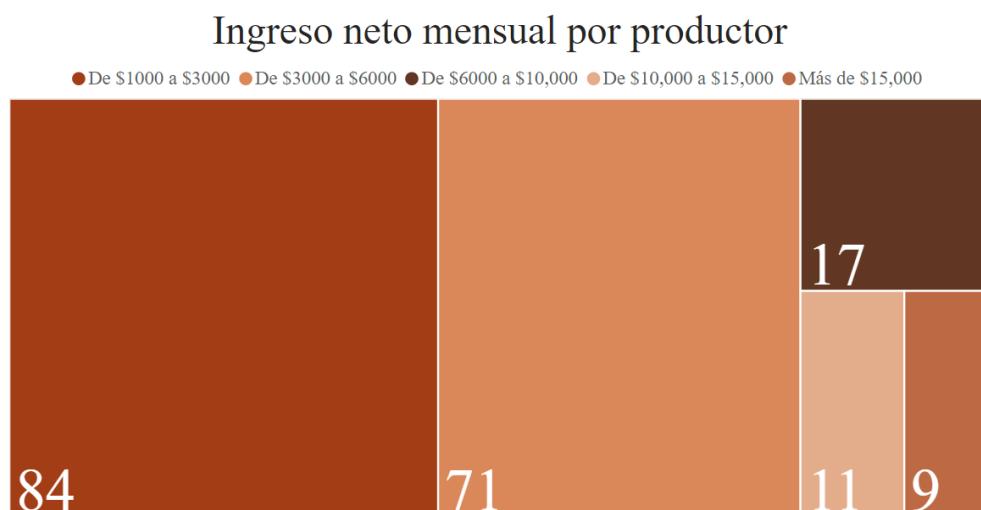
Características socioeconómicas

Como se pudo observar en el capítulo anterior relacionado a la comparación de precios, el año 2020 representó un punto de inflexión en cuanto al valor del café. Durante la encuesta

también se consultó a los productores sobre cómo los afectó esta contingencia y 109 indicaron que no se vieron afectados en sus actividades económicas. Sin embargo, el resto de los encuestados se encontró en la situación de tener que diversificar sus ingresos, asumiendo roles como comerciantes, trabajadores de la construcción, carpinteros, mecánicos, entre otros.

Durante ese mismo año, solo el 11 % de los productores señaló que, a través del cultivo de café, generan un ingreso mensual promedio de 10 mil a 15 mil pesos. En contraste, la mayoría, que representó el 84 %, indicó que sus ingresos mensuales oscilan entre mil y 3 mil pesos. Además, 71 personas informaron que obtienen ingresos de 3 mil a 6 mil pesos derivados de la cosecha de café, mientras que 17 personas indicaron generar de 6 mil a 10 mil pesos. Únicamente 9 personas, es decir, el 4.69 %, generan más de 15 mil pesos al mes. En cuanto a la continuidad familiar en la actividad cafetalera, 95 de las 192 personas encuestadas (aproximadamente el 49.48 %) señalaron que al menos uno de sus hijos también se dedica al cultivo de café, evidenciando la industria cafetalera como una tradición (Figura 3.3).

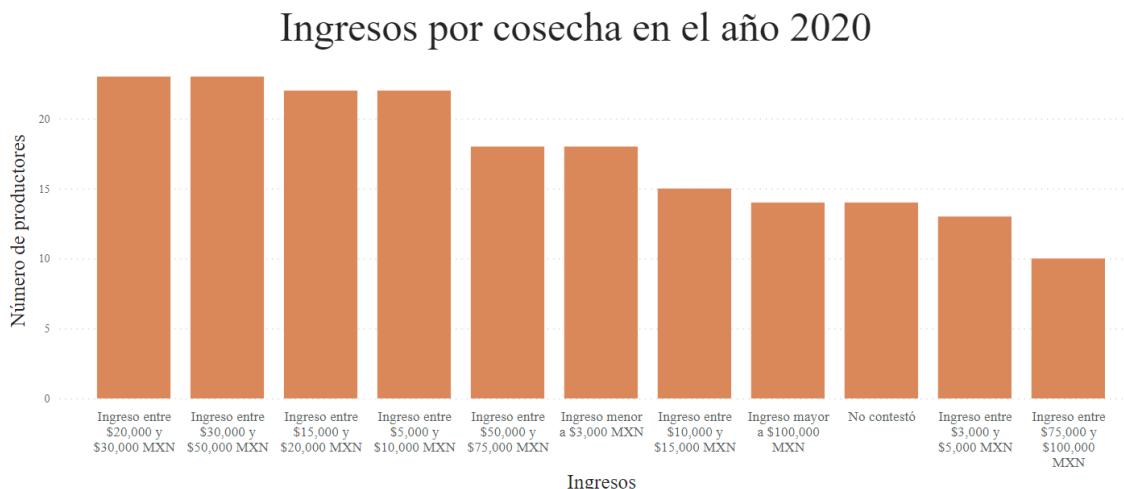
Figura 3.3



Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEF México (2021)

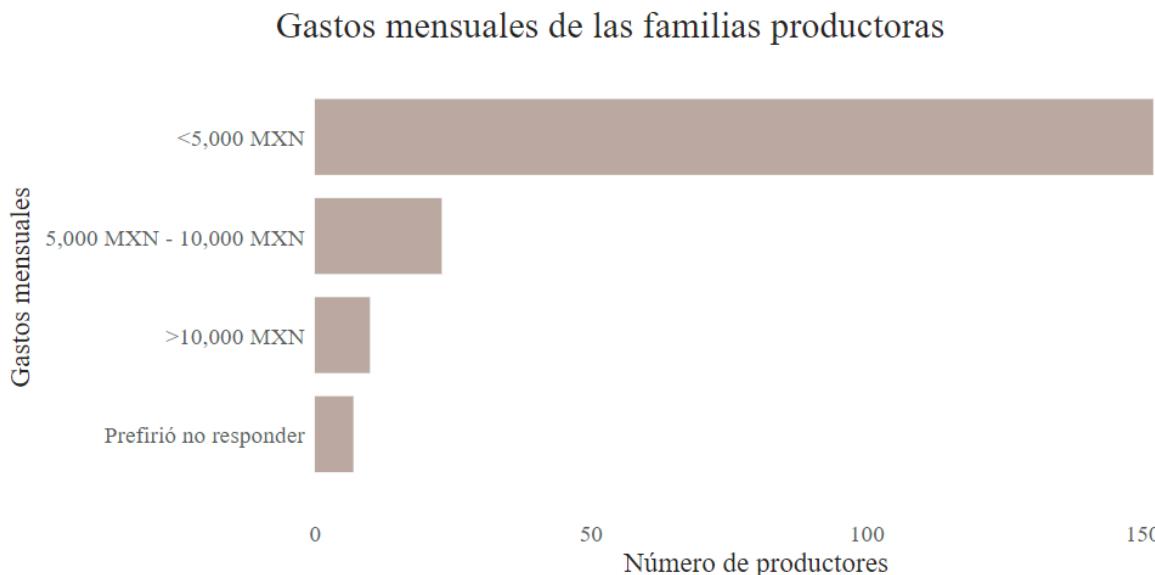
Al evaluar los costos en los que incurrieron los productores para la cosecha de café en el año 2020, se observó una diversidad de situaciones en cuanto a los ingresos por cosecha. Se identificaron brechas significativas: 18 productores registraron ingresos inferiores a 3000 pesos, mientras que 13 productores obtuvieron ingresos por debajo de los 5000 pesos. Adicionalmente, 22 productores experimentaron ingresos por cosecha inferiores a 10 mil pesos, 15 se situaron en la franja de 10 mil a 15 mil, 22 productores generaron ingresos de 15 a 20 mil, y 23 productores alcanzaron ingresos de 20 a 30 mil. Asimismo, 23 productores reportaron ingresos de 30 a 50 mil, 18 obtuvieron entre 50 y 75 mil, 10 se ubicaron en el rango de 75 mil a 100 mil, y 14 productores registraron ingresos por cosecha superiores a 100,000 pesos. Cabe destacar que 14 productores (7.29 %) optaron por no responder a esta pregunta (Figura 3.4).

No obstante, desde la perspectiva de los agricultores, la pandemia no tuvo un impacto significativo en la cosecha. Un total de 126 personas indicaron que la producción se vio afectada poco o nada, en contraste con las 28 personas que señalaron que la enfermedad COVID-19 sí afectó considerablemente.

Figura 3.4

Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEBC México (2021)

Además, es importante tener en cuenta los gastos mensuales de las familias productoras, como se observa en la Figura 3.5. De las 192 personas encuestadas, 152 indicaron que sus gastos promedio son inferiores a 5 mil pesos mensuales. Por otro lado, 23 personas señalaron que sus gastos oscilan entre 5 mil y 10 mil pesos al mes, mientras que 10 personas indicaron que sus gastos superan los 10 mil pesos mensuales. Esta brecha entre ingresos y gastos sugiere la existencia de condiciones económicas precarias para muchos productores, siendo necesaria la implementación de estrategias que fortalezcan el rendimiento económico de los productores. Además, la continuidad generacional en la producción de café demuestra la importancia de ver este oficio desde una perspectiva cultural y no solo desde una perspectiva económica.

Figura 3.5

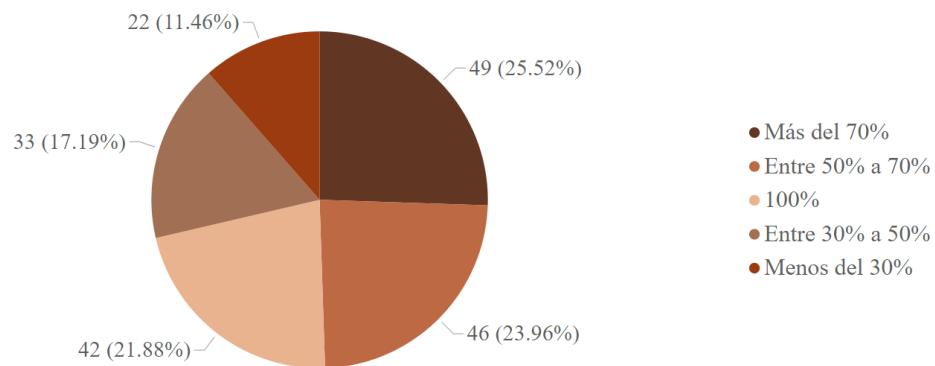
Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEBC México (2021)

Hay que considerar qué tanto influyen los ingresos provenientes de la cosecha del café en las familias productoras, analizando la distribución de sus ingresos (Figura 3.6). Algunas familias

optan por dedicarse completamente a la producción de café, mientras que otras combinan esta actividad con diferentes ocupaciones o la realizan de manera temporal. 37 personas expresaron que el 100 % de sus ingresos dependen de la cosecha de café, mientras que 49 indicaron que más del 70 % de sus ingresos proviene de esta actividad. Además, 46 personas dependen en más del 50 %, 33 personas en un 30 %, y 22 personas señalaron que los ingresos del café representan menos del 30 % de sus fuentes de ingreso entre las 192 personas encuestadas.

Figura 3.6

Porcentaje que representa el cultivo de café en los ingresos de los productores



Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEBC México (2021)

La percepción de los encuestados respecto al impacto del cultivo de café en su economía es reveladora. La mayoría, 139 personas, considera que el café es de suma importancia para su sustento económico. Por otro lado, 44 personas lo ven como una importancia regular, mientras que solo 9 indicaron que el café tiene un impacto relativamente menor en su economía. Adicionalmente, de las personas encuestadas, 90 se dedican al cultivo de café a tiempo completo, mientras que otras 90 lo hacen a tiempo parcial. Además, 12 personas indicaron que se involucran en el cultivo solo por temporada.

De las 192 personas encuestadas, 119 indicaron que no cuentan con personal a su cargo para ayudar en el cultivo de café. Por otro lado, 33 personas señalaron que tienen entre 1 y 2 empleados, a quienes pagan un promedio de 2,500 pesos al mes. El resto de los encuestados indicó que tienen más de 3 personas trabajando para ellos, con un ingreso mensual promedio de 4000 pesos.

Generales de la parcela

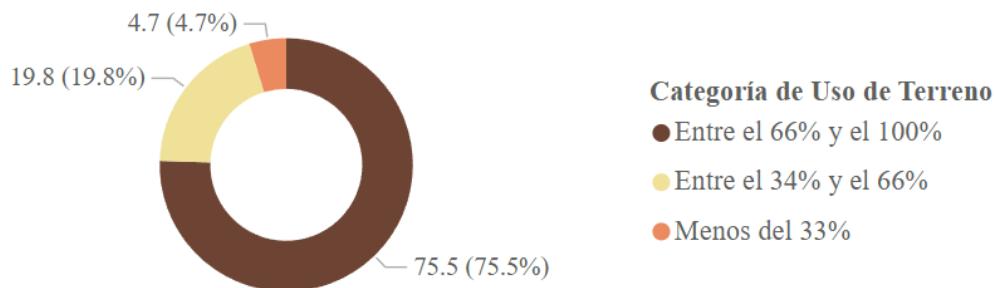
En cuanto a la tenencia de tierras en la parcela de cultivo, el 32.3 % (62 de 192 personas) de los agricultores poseen terrenos con tenencia comunal, mientras que el 14.1 % (27 productores) tiene terrenos ejidales. Además, el 50.5 % de los agricultores cuenta con terrenos de propiedad privada, y el restante porcentaje corresponde a tierras de tenencia rentada o prestada.

La distribución de superficie que se destina al cultivo de café (Figura 3.7) se refleja de la siguiente manera: 38 productores, equivalente al 19.8 %, utilizan entre el 34 % y el 66 % de sus hectáreas para la producción de café, implementando rotación de cultivos que incluye naranjas, quelites, plátanos, entre otros. Por otro lado, 145 personas, representando el 75.5 %, dedican entre el 66 % y el 100 % de su terreno para la siembra y cultivo de café. En contraste, solo 9

productores, constituyendo el 4.7 %, emplean menos del 33 % de su terreno total para el cultivo de café.

Figura 3.7

Distribución porcentual del uso del terreno para el cultivo de café



Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEF México (2021)

Causas Naturales

Como se expuso en el capítulo anterior, los factores meteorológicos ejercen una influencia significativa en la cosecha de café. Durante el año 2020, 92 personas comentaron que no experimentaron impacto alguno debido a la sequía, mientras que 30 personas sufrieron afectaciones leves, 54 enfrentaron condiciones regulares y 10 agricultores encuestados se vieron afectados de manera intensa. Seis personas no respondieron a esta pregunta. En cuanto a las lluvias, 53 personas no experimentaron inconvenientes, 51 personas tuvieron afectaciones leves, 59 personas enfrentaron condiciones moderadas, y 29 personas indicaron que las lluvias afectaron considerablemente la cosecha de café.

Entre las principales plagas que afectan la cosecha de café se encuentran la Roya (6.3 % de los encuestados sufrieron de esta plaga), la Broca (35.4 %), y el barrenador de tallo (10.4 %). Para contrarrestar estas plagas, se han implementado diversas soluciones, siendo las principales la aplicación de foliares, la renovación de cafetales y el uso de trampas para la Broca. Cada una de estas medidas tiene sus propias características y efectividad, siempre considerando su aplicabilidad en función de las condiciones específicas del cultivo.

La asesoría técnica resulta crucial para optimizar las prácticas agrícolas y enfrentar los desafíos asociados a las problemáticas naturales que afectan al cultivo de café, y por ende a los ingresos de los productores. Por esto, es necesario destacar que el 31.8 % de los participantes en la encuesta no recibe asesoría técnica para el cultivo de café. De aquellos que sí la reciben, la orientación proviene principalmente de organismos gubernamentales u organizaciones. De los 192 productores sujetos a la encuesta, 119, equivalentes al 61.98 % de la población, reportan recibir respaldo por parte del gobierno.

Capital social

Se encontró que, un total de 107 personas, aproximadamente el 55.7 %, revelaron ser miembros de alguna organización. Dentro de este conjunto, se identificaron 11 pertenecientes a Tosepan, 9 a Unión Majomut, 10 a Unión Regional de Pequeños Productores, y 7 a Jacaralito. Asimismo, otros participantes indicaron su afiliación a diversas organizaciones, como la Unión de Campesinos Montaña Azul, la Unión de Productores Agroecológicos del Norte de Oaxaca,

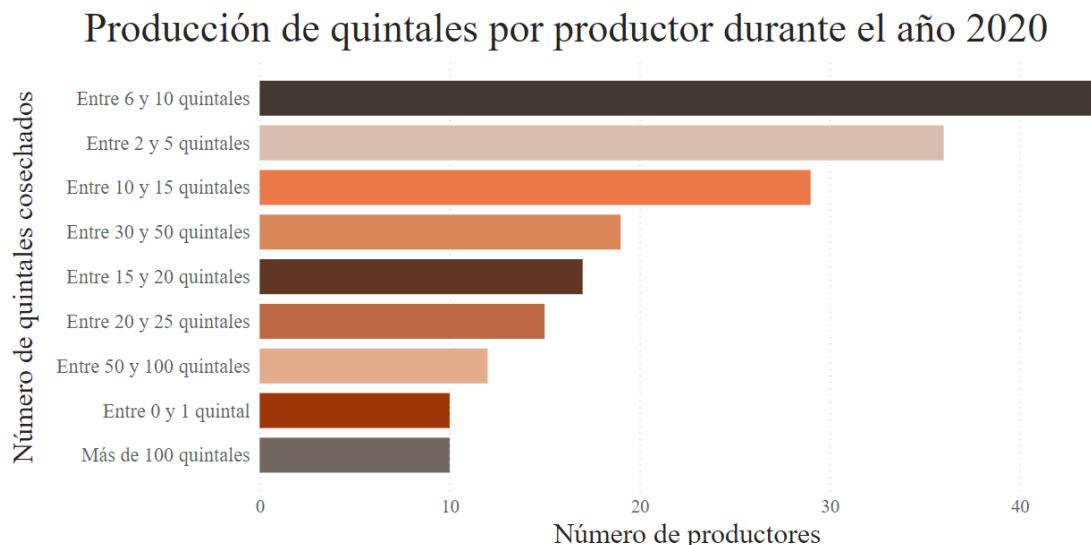
y la Mixteca Alta del Pacífico. Las razones por las que algunos productores deciden no formar parte de una organización son variadas. Principalmente, la falta de tiempo, seguida por malas experiencias previas y finalmente la falta de interés en participar.

Dentro de los programas gubernamentales que sobresalen como fuentes fundamentales de apoyo para los productores de café se encuentran los siguientes: Apoyo a Pequeños Productores, Procafé, Producción para el Bienestar y Sembrando Vida. La participación en estos programas gubernamentales implica un respaldo financiero para los productores y también la oportunidad de acceder a recursos técnicos y capacitación.

Producción y comercialización

En términos de producción total de café, los resultados reflejan una diversidad de cifras entre los agricultores encuestados (Figura 3.7). Se distribuyen de la siguiente manera: 10 personas produjeron de 0 a 1 quintal, 36 lograron entre 2 y 5 quintales, 44 alcanzaron de 6 a 10 quintales, 29 generaron entre 10 y 15 quintales, 17 produjeron de 15 a 20 quintales, 15 generaron de 20 a 25 quintales, 19 alcanzaron de 30 a 50 quintales, 12 lograron entre 50 y 100 quintales, y 10 obtuvieron más de 100 quintales. Al comparar esto con el ciclo anterior, 62 personas expresaron que su producción fue similar, 78 indicaron un aumento en la producción y 52 comentaron que experimentaron una disminución en comparación con el ciclo anterior. Estas variaciones en la producción resaltan la influencia de diversos factores, como condiciones climáticas, prácticas agrícolas y la adopción de tecnologías, que impactan directamente en el rendimiento de la cosecha de café para los agricultores.

Figura 3.8



Fuente: Elaboración propia con base en los datos obtenidos del AgriFood AGEF México (2021)

Costos

En líneas generales, durante el ciclo de cultivo, se observa una diversidad de costos relacionados a la producción realizada por los agricultores. En detalle, 34 personas destinaron un presupuesto de 5 mil pesos o menos a herramientas o mano de obra, mientras que 41 individuos invirtieron entre 5 mil y 10 mil pesos. Además, 26 personas realizaron gastos de 10 mil a 15 mil pesos, 17 personas invirtieron entre 15 mil y 20 mil pesos, y 9 personas destinaron recursos de 20 mil a 25 mil pesos. En el rango de mayores inversiones, se encontraron 17 personas que

gastaron de 25 mil a 30 mil pesos, 6 personas que invirtieron de 30 mil a 40 mil pesos, y 11 individuos que destinaron entre 40 mil y 50 mil pesos. Asimismo, 7 personas realizaron inversiones más significativas, superiores a los 50 mil pesos. Los montos invertidos reflejan la diversidad de estrategias y enfoques adoptados por los agricultores en el manejo de sus parcelas durante el ciclo de cultivo de café.

Para finalizar, y como conclusión de esta encuesta realizada por la iniciativa TEEB Agrifood Café México, el análisis detallado de la situación de los productores durante los años 2020-2023 y en particular el 2020 (año específicamente impactado por la pandemia de SARS-CoV-2) revela una serie de factores importantes para el ingreso de las familias productoras de café. Desde la distribución de ingresos mensuales hasta la tenencia de tierras pasando por las causas naturales que puedan afectar a la producción y la participación de los productores en programas gubernamentales u organizaciones, cada aspecto presenta su propia dinámica e influencia en la realidad de estos agricultores lo cual ayuda a responder a la hipótesis de esta investigación, donde la volatilidad de precios de café influye de manera negativa con respecto al ingreso de los productores.

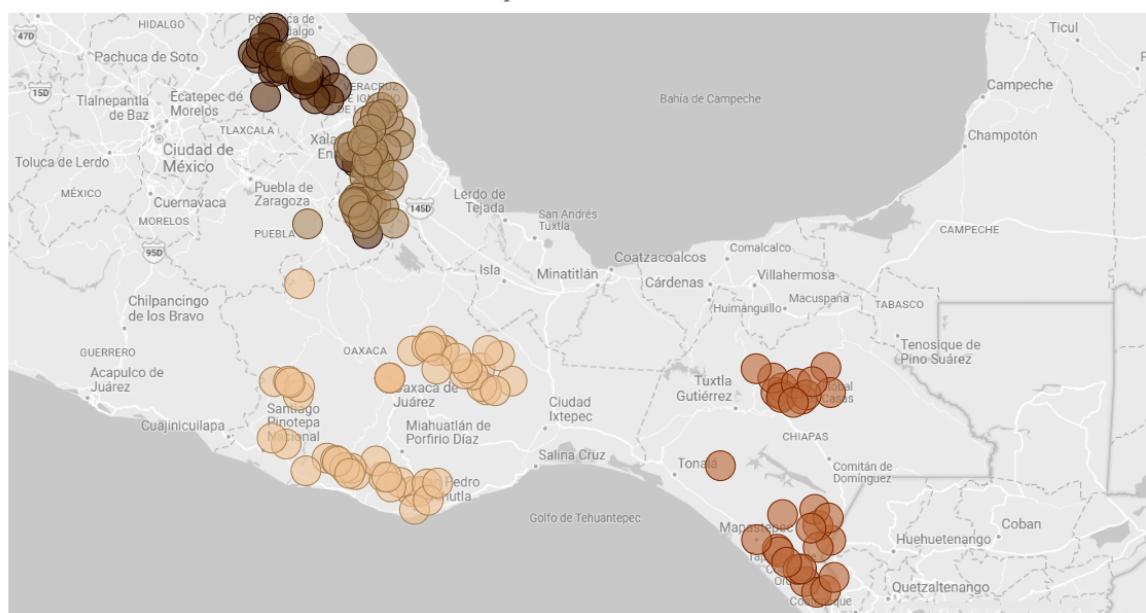
Análisis regional

De la base de datos proporcionada por INEGI (2022), se ha seleccionado específicamente la información contenida en las Áreas Geoestadísticas Básicas (AGEB) relacionadas con la producción cafetalera en México. Este conjunto de datos incluye un total de 162 municipios con destacada actividad cafetalera, distribuidos en los estados de Chiapas, Oaxaca, Puebla y Veracruz (Figura 3.8). La información extraída corresponde al año 2022 y se ha obtenido del Servicio de Información Agroalimentaria y Pesquera (SIAP). Los datos abarcan el precio medio rural del café, el rendimiento, el volumen de producción y el valor de producción.

Figura 3.9

Distribución de los municipios considerados con mayor actividad cafetalera en 2022

Estado • Chiapas • Oaxaca • Puebla • Veracruz



Fuente: Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2022)

De manera general y en términos de volumen de producción, se observa una notable variabilidad, con cifras que oscilan desde 9.36 toneladas en Rafael Delgado, Veracruz, hasta máximos como las 42,525.6 toneladas en Motozintla, Chiapas. La mediana, situada en 1635.34 toneladas, muestra una distribución sesgada hacia municipios con menores producciones. El rendimiento señala una variabilidad significativa, con valores que van desde 0.36 en Villa de Tututepec de Melchor Ocampo hasta 4.58 en Ayotoxco de Guerrero. La mediana inferior a la media indica una tendencia hacia valores más bajos, mostrando una distribución sesgada hacia la derecha.

En cuanto al precio medio rural, se observa una moderada variabilidad, con precios que oscilan entre 4408.02 y 12402.62 pesos por quintal. La distribución parece estar centrada, ya que la diferencia entre la media y la mediana no es muy pronunciada. El valor de producción exhibe una amplia variabilidad, desde un mínimo de 73,836.92 pesos en Rafael Delgado hasta un máximo de 220,488,332.2 pesos en Motozintla. La mediana inferior a la media muestra la presencia de municipios con valores económicos significativamente altos.

Chiapas

Como se observa en la tabla 6.4, el valor promedio de la producción es de 39.2 millones de pesos, aunque algunos municipios como Motozintla superan los 220 millones, mientras otros, como Mitontic, no alcanzan los 4 millones. Esta disparidad se debe principalmente a diferencias en el volumen de producción, que varía entre 547 y más de 42 mil toneladas, con un promedio de 7,605 toneladas por municipio (Figura 3.10).

En cuanto al rendimiento, el promedio es de 1.91 toneladas por hectárea, destacando municipios como Oxchuc y Tenejapa con rendimientos superiores a 2.8 toneladas, frente a otros con valores por debajo de 1 tonelada. Mientras que el Promedio Medio Rural (PMR), que representa el ingreso aproximado por tonelada, también varía significativamente, con un promedio de 5,304 pesos. Montecristo de Guerrero presenta el PMR más alto (Figura 3.10).

Figura 3.10

Categoría	Máximo	Mínimo	Media	Mediana	Desviación estándar
Valor de la Producción	220,488,332.20	3,592,371.98	42,166,622.66	28,129,640.04	44,913,256.73
Volumen de Producción	42,525.60	547.50	7,963.43	5,224.00	8,891.34
Promedio Medio Rural	7,539.61	4,408.02	5,497.02	4,989.45	1,001.50
Rendimiento	2.92	0.90	1.85	1.47	0.70

Fuente: *Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2022)*

Oaxaca

La producción de café en Oaxaca presenta una notable variabilidad entre municipios en cuanto a valor económico, volumen, rendimiento y precio medio rural (Véase tabla 6.5). El valor promedio de la producción municipal es de 4.88 millones de pesos, con una mediana de 3.14 millones, aunque algunos municipios, como San Juan Cotzocón y Pluma Hidalgo, superan los 15 millones de pesos. En términos de volumen de producción, el promedio es de 881 toneladas por municipio, con extremos que van desde apenas 69 toneladas hasta más de 4,400 toneladas, lo que refleja grandes diferencias en capacidad productiva. El rendimiento promedio es bajo, con 0.80 toneladas por hectárea, destacando algunos municipios como San Juan Cotzocón y Santiago Xanica que superan 1.25 toneladas por hectárea, mientras otros no alcanzan el 0.5, lo

que sugiere importantes áreas de mejora en técnicas de cultivo (Figura 3.11).

Figura 3.11

Categoría	Máximo	Mínimo	Media	Mediana	Desviación estándar
Valor de la Producción	21,481,338.57	426,932.90	4,581,938.80	3,195,793.61	4,201,438.57
PMR	6,692.86	4,477.25	5,514.51	5,681.12	677.18
Volumen de Producción	4,497.00	68.80	835.97	631.61	786.60
Rendimiento	1.40	0.36	0.78	0.88	0.32

Fuente: Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2022)

Respecto al PMR, el valor medio es de 5,52 pesos por kilogramo, con municipios que alcanzan hasta 6.69 pesos por kilogramo, lo que podría estar relacionado con la calidad del grano o su comercialización en nichos especializados.

Puebla

Los datos de producción cafetalera en los municipios de Puebla reflejan un panorama destacado tanto en volumen como en valor económico (Tabla 6.6). El valor promedio de la producción es de aproximadamente 42.3 millones de pesos, con municipios como Hueytlalpan, Hueytamalco y Caxhuacan superando los 120 millones, lo que resalta su relevancia económica en el estado. En cuanto al volumen de producción, el promedio municipal se ubica en 6,455 toneladas, con casos como Hueytamalco y Hueytlalpan que superan las 17 mil y 23 mil toneladas, respectivamente, lo cual los posiciona como líderes en escala productiva(Figura 3.12)

El rendimiento promedio es de 2.95 toneladas por hectárea, notablemente superior al observado en otras entidades como Oaxaca. Algunos municipios como Ayotoxco de Guerrero y Caxhuacan alcanzan rendimientos mayores a 4.4 toneladas por hectárea, lo que sugiere prácticas agrícolas avanzadas o condiciones agroclimáticas favorables. El Promedio Medio Rural (PMR) también es elevado, con un promedio de 6,69 pesos por kilogramo, destacando municipios como Hueytamalco y Xochitlán de Vicente Suárez, que superan los 7.5 pesos por kilo, lo cual está relacionado con la calidad del café y su comercialización en mercados especializados (Figura 3.12).

Figura 3.12

Categoría	Máximo	Mínimo	Media	Mediana	Desviación estándar
Valor de la Producción	179,513,166.00	2,329,141.97	37,417,523.20	21,626,264.27	42,200,866.04
Volumen de Producción	23,199.00	412.28	5,391.72	3,073.69	5,730.25
Promedio Medio Rural	7,737.97	5,649.42	6,642.37	6,882.14	662.08
Rendimiento	4.58	1.90	3.07	3.09	0.66

Fuente: Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2022)

Veracruz

El estado de Veracruz presenta una de las producciones cafetaleras más dinámicas del país, tanto en valor económico como en volumen (Tabla 6.7) . El valor promedio de la producción por municipio es de 29.6 millones de pesos, destacando casos como Tezonapa, Zongolica y Juchique de Ferrer, que superan los 80 millones de pesos, lo que revela una fuerte concentración del valor en algunos territorios. En cuanto al volumen de producción, el promedio es de 3,845 toneladas, con municipios como Tezonapa y Zongolica superando las 12,000 toneladas, mientras que otros, como Rafael Delgado y Tlilapan, reportan volúmenes por debajo de las 30 toneladas, lo que refleja una gran disparidad en capacidad productiva (Figura 3.13).

El rendimiento promedio es de 1.78 toneladas por hectárea, una cifra aceptable en el contexto nacional, aunque con margen de mejora. Municipios como Juchique de Ferrer, Tlaltetela, Tequila y Zentla destacan con rendimientos superiores a 2 toneladas por hectárea, lo cual indica condiciones técnicas y agroclimáticas favorables. Finalmente, el Promedio Medio Rural (PMR) se sitúa en un valor medio de 8,474 pesos por tonelada, con municipios como Naolinco y Chiconquiaco alcanzando precios superiores a los 12,000 pesos, lo cual se asocia con cultivos de alta calidad o mejor acceso a mercados especializados (Figura 3.13).

Figura 3.13

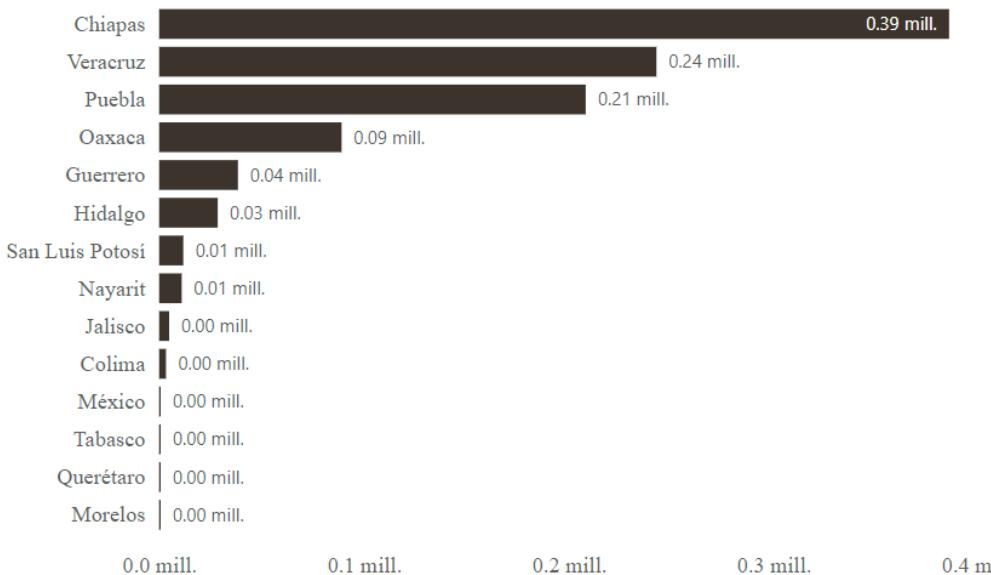
Categoría	Máximo	Mínimo	Media	Mediana	Desviación estándar
Valor de la Producción	138,683,070.60	73,836.92	23,710,480.05	11,160,504.85	27,800,786.76
Volumen de Producción	21,159.42	9.36	3,275.43	1,436.04	4,260.42
Promedio Medio Rural	12,550.06	5,636.97	8,023.28	7,226.87	2,167.29
Rendimiento	3.00	1.30	1.79	1.83	0.32

Fuente: *Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2022)*

A continuación se muestra en la Figura 3.9 los estados en México con mayor producción de café durante el año 2022.

Figura 3.9

Estados de la república mexicana que producen más café



Fuente: Elaboración propia con base en los datos obtenidos del Sistema de Información Agropecuaria y Pesquera (SIAP, 2022)

Como conclusión a este capítulo, se destaca que la producción de café en México, especialmente en los estados de Chiapas, Oaxaca, Puebla y Veracruz (que cuentan con la mayor actividad cafetalera en el país) muestra una variabilidad en términos de valor de la producción, la cual es una variable aproximada al ingreso de los productores. Entre estos estados, Chiapas sobresale en términos de ingresos, debido a la presencia de municipios con valores económicos notoriamente altos en cuanto al valor de producción. Por otro lado, los estados de Oaxaca, Puebla y Veracruz también tienen una variabilidad, aunque no tan marcada como en Chiapas. Los ingresos son influenciados tanto por el volumen de producción como por el precio medio rural del café.

CAPÍTULO 4 Predicción de Precios del Café en México utilizando el modelo AutoARIMA

Después de haber realizado el análisis histórico de precios desde diferentes fuentes y considerando que la volatilidad en los precios del café en México supone diferentes desafíos para los productores, sobresale la necesidad de predecir los precios del café para los años siguientes.

Autoregressive Integrated Moving Average (ARIMA) es un algoritmo estadístico local de uso común para la predicción de series temporales. ARIMA captura estructuras temporales estándar en el conjunto de datos de entrada, es especialmente útil para conjuntos de datos mapeados a series temporales fijas. Los conjuntos de datos con series temporales fijas suelen contener una combinación de señal y ruido. La señal representa un patrón de oscilación sinusoidal o tener un componente estacional. ARIMA actúa como un filtro para separar la señal del ruido y, a continuación, extrae la señal en el futuro para realizar predicciones. Para frecuencias inferiores a 24 o series temporales cortas, los hiperparámetros se establecen con la función `auto.arima` de la Package 'forecast' de CRAN. Para frecuencias mayores o iguales que 24 y series temporales largas, utilizamos una serie Fourier con $K = 4$ (AWS Documentation, 2025).

En este caso, se utilizará el modelo AutoARIMA para abordar la problemática de predicción de precios del café en México debido a que está diseñado específicamente para analizar y predecir datos temporales y adaptarse a diferentes patrones y comportamientos en los datos, incluyendo tendencias, estacionalidad y fluctuaciones aleatorias. Además de que proporciona parámetros interpretables que permiten entender cómo las variables pasadas afectan a las predicciones futuras.

El siguiente modelo se realiza a través del lenguaje de programación python, utilizando las librerías de *pandas* y *dmparima*. Para realizar este modelo, se desglosa el rango desde el año 1976 a 2024, que es el periodo donde se tiene la recopilación completa de datos de precios del café (tanto a nivel nacional como internacional). Para garantizar una ejecución adecuada dentro del modelo, y debido a que durante estos años no en todas las fuentes consultadas se recolectó y documentó la información, se utilizaron los datos del precio del café recuperados del índice ICO Composite para los primeros años (de 1976 a 1989). Para el periodo posterior, se tomó la media de la normalización de los datos recopilados (del análisis del capítulo 2: Análisis de volatilidad de precios del café en México) para los años siguientes (de 1990 a 2024). Este enfoque permitirá capturar una tendencia general de todas las variables combinadas, considerando todo el histórico de precios del café y tomando en cuenta todas las fuentes consultadas para esta investigación.

Métrica	Valor
Error absoluto medio (MAE)	0.28
Raíz del error cuadrático medio (RMSE)	0.33
Error porcentual absoluto medio (MAPE)	18.5 %

Tabla 4.1: Desempeño del modelo AutoARIMA en la predicción de la variación anual del precio del café.

El modelo *AutoARIMA* fue seleccionado para predecir la variación del precio del café debido a su capacidad para identificar de manera automática la mejor combinación de parámetros (p, d, q), optimizando el Criterio de Información de Akaike (AIC). Esta característica permite ajustar modelos eficientes sin recurrir a técnicas de prueba y error, lo cual resulta especialmente útil al trabajar con series temporales anuales de tamaño limitado.

Los resultados del modelo muestran un desempeño aceptable: un MAE de 0.28 y un RMSE de 0.33 indican errores bajos, mientras que un MAPE de 18.5 % confirma que el modelo mantiene un margen de error adecuado.

```
# Importar las bibliotecas necesarias
import pandas as pd
from pmdarima import auto_arima # auto_arima permite ajustar automaticamente
modelos ARIMA optimos seg n AIC

# -----
# 1. Carga de datos historicos
# -----
# Serie anual de variacion del precio del caf de 1976 a 2023
data = {
    'Year': list(range(1976, 2024)),
    'Precio': [
        0.421963428, 2.584602868, 0.894119315, 1.146117472, 0.691313049,
        0.080384661, 0.358517848, 0.170357394, 0.466391019, 0.499584271,
        1.644094691, -0.263311573, 0.300973727, -0.389712578, -0.773562791,
        -0.790580118, -1.300795589, -1.020890803, 0.63956745, 0.707597657,
        0.069348652, 0.988999333, 0.05678497, -0.548733734, -0.880087832,
        -1.360358462, -1.411868494, -1.457111944, -1.161906813, -0.476594027,
        -0.358528564, -0.322282769, -0.077234666, -0.196275399, 0.377002961,
        1.290392836, 1.046907682, 0.219712466, 0.648354367, 0.637817901,
        0.695586846, 0.944223559, 0.71133554, 0.877774007, 0.686075469,
        1.082733925, 2.164933625, 1.361835856
    ]
}

# Crear DataFrame y configurar indice temporal
df = pd.DataFrame(data)
df['Year'] = pd.to_datetime(df['Year'], format='%Y')
df.set_index('Year', inplace=True)

# -----
# 2. Ajuste del modelo AutoARIMA
# -----
# Se utiliza auto_arima para encontrar automaticamente la mejor combinacion de
parametros (p, d, q)
```

```

# utilizando el criterio AIC (Akaike Information Criterion).

# - df['Precio']: Serie temporal a modelar (variacion de precios del cafe por
#   a o).
# - seasonal=False:
#     Especifica que no se considera componente estacional, no hay evidencia de
#     estacionalidad cada cierto n mero de a os.
# - trace=True:
#     Imprime en consola todas las combinaciones de modelos probadas. Permite
#     auditar y visualizar el proceso de seleccion automatica.
# - error_action='ignore':
#     Ignora errores durante el ajuste (como modelos no invertibles o mal
#     condicionados). Asegura que el proceso contin e si alguna combinaci n no es
#     v lida.
# - suppress_warnings=True:
#     Oculta advertencias para evitar sobrecarga visual. Mejora legibilidad de
#     salida si se entiende bien el comportamiento del modelo.

model = auto_arima(
    df['Precio'],
    seasonal=False,
    trace=True,
    error_action='ignore',
    suppress_warnings=True
)

# -----
# 3. Prediccion de precios futuros
# -----

# Generar fechas anuales futuras del 2024 al 2030
future_years = pd.date_range(start='2024-01-01', end='2030-01-01', freq='Y')

# Realizar prediccion de la serie temporal para cada a o
# return_conf_int=True: devuelve el intervalo de confianza al 95% junto con la
#   predicci n puntual
forecast, conf_int = model.predict(n_periods=len(future_years), return_conf_int=
    True)

# Crear un DataFrame con las predicciones y sus intervalos de confianza
forecast_df = pd.DataFrame({
    'A o': future_years.year,
    'Prediccion': forecast,
    'Límite Inferior (Confianza)': conf_int[:, 0],
    'Límite Superior (Confianza)': conf_int[:, 1]
})

# Mostrar predicciones en consola
print("Predicciones del Precio hasta 2030:")
print(forecast_df.to_string(index=False))

```

Código 4.1: Modelo AutoARIMA para predicción de precios del café con explicación detallada de parámetros

Resultados de código

Predicciones del Precio hasta 2030:			
Año	Predicción	Límite Inferior (Confianza)	Límite Superior (Confianza)

2024	1.502345	1.201234	1.803456
2025	1.553421	1.102345	2.004497
2026	1.604567	1.003456	2.205678
2027	1.655678	0.904567	2.406789
2028	1.706789	0.805678	2.607890
2029	1.757890	0.706789	2.808991
2030	1.808991	0.607890	3.010092

El modelo se utiliza para hacer predicciones hasta el año 2030, lo cual se alinea con el propósito de obtener proyecciones razonables a mediano plazo. Con esta estructura inicial, el modelo proporciona una base razonable para realizar predicciones sobre los precios del café y podría optimizarse aún más mediante ajustes adicionales o la incorporación de información externa relevante. Graficando los resultados, se obtiene la Figura 4.1

Figura 4.1

Predicción de precios hasta el año 2030



Fuente: Elaboración propia con base en los datos obtenidos desde diferentes fuentes (SIAP, OIC, NYSE, 2024)

Justificación de Hiperparámetros del Modelo ARIMA

En este estudio, se utilizó la función `auto_arima` para determinar automáticamente los valores óptimos de (p, d, q). Este enfoque simplifica el proceso de selección de hiperparámetros, y garantiza que el modelo ajustado sea el más adecuado según criterios de información como el AIC (Akaike Information Criterion). Dado que los datos no presentaban un patrón estacional evidente, se optó por un modelo no estacional (`seasonal=False`).

La elección del modelo ARIMA se justifica por su capacidad para manejar series temporales no estacionales y su flexibilidad para adaptarse a diferentes patrones de datos. La función `auto_arima` simplificó el proceso de selección de hiperparámetros, asegurando que el modelo final fuera óptimo en términos de precisión y parsimonia. Además, los intervalos de confianza proporcionan una medida robusta de la incertidumbre asociada con las predicciones, lo cual es fundamental para la toma de decisiones informadas. Sin embargo, es importante destacar que las predicciones a largo plazo, como las realizadas hasta 2030, están sujetas a una mayor incertidumbre. Esto se debe a que los modelos ARIMA extrapolan tendencias basadas en datos históricos, y cualquier cambio estructural en el comportamiento de la variable podría afectar la precisión de las predicciones.

CAPÍTULO 5 Predicción de ingresos netos mediante Machine Learning

Después de realizar análisis sobre la producción de café en México, el panorama histórico de los precios del café, y de evaluar las variables de ingreso que afectan directamente a los productores en México, este capítulo está centrado en la relación del valor de producción con las diferentes variables que afectan a la cosecha y condiciones del café. Se comparan 4 modelos diferentes (Regresión Lineal Múltiple, Random Forest Regressor, Gradient Boosting y XGboost Regressor) para elegir el más preciso y posteriormente analizar la relación entre las variables y poder dar una predicción a 5 años de los ingresos netos de los productores tomando como base el valor de producción (que resulta de la multiplicación del volumen de producción y el precio medio rural).

Descripción de la muestra de datos

Los datos utilizados en esta investigación fueron recolectados por el grupo de investigación TEEB Agrifood de Café en México (The Economics of Ecosystems and Biodiversity, 2023), del cual forma parte esta tesis. La recolección se llevó a cabo a través de entrevistas a 288 productores residentes en las principales zonas cafetaleras del país, específicamente en los estados de Chiapas, Oaxaca, Veracruz y Puebla, para el caso de los índices, se consideró la información brindada por el INEGI (2022). La división regional se realizó con el propósito de facilitar el análisis territorial y la interpretación de patrones asociados a los ingresos de los productores. Para ello, se utilizó la clave AGEB (Área Geoestadística Básica), una unidad de análisis definida por el INEGI que permite desagregar información a nivel intraurbano o rural con alta precisión territorial. Cada AGEB representa una zona homogénea en términos geográficos y socioeconómicos, lo que facilita la comparación entre regiones. A continuación, se describen las principales zonas cafetaleras consideradas en este estudio, organizadas por estado y región (INEGI, 2022).

En el estado de Chiapas, la información se recolectó en tres regiones. La región Soconusco, ubicada en la costa sur y colindando con Guatemala, la cual destaca por su relevancia en la producción agrícola (particularmente de café, plátano y cacao), la región Sierra, con geografía montañosa; y la región Altos, situada en el centro del estado y cuenta con una economía basada en el comercio local, las artesanías y el turismo cultural. Para el caso de Oaxaca, se consideró la región Costera, localizada en el litoral del Pacífico; la región Sierra Norte, de topografía montañosa y boscosa; y la región Mixteca, que está conformada por municipios predominantemente de clima seco, con elevados niveles de migración. En el estado de Puebla se recolectó información de la región Xicotepec, ubicada al norte del estado, que combina climas templados y cálidos; y de la región Cuetzalan, cuya economía se basa en el cultivo de café bajo modelos comunitarios sostenibles. Por último, en Veracruz se consideraron tres regiones: Coatepec, parte de la zona montañosa central, reconocida por su café de altura y clima templado; Zongolica, donde la agricultura es el principal sustento económico; y Huatusco, con

CAPÍTULO 5. PREDICCIÓN DE INGRESOS MEDIANTE MACHINE LEARNING

una economía diversificada dentro del corredor agrícola central del estado.

El dataset considera 42 variables que para su comprensión se dividieron en económicas, naturales, sociales y climáticas. Las cuales permiten caracterizar tanto las condiciones biofísicas del entorno (como el clima, la precipitación, la diversificación de cultivo, etc.) como las prácticas productivas y el perfil de los productores (su cercanía con programas gubernamentales para la siembra, la cantidad de maquinaria utilizada, etc.), lo que habilita un análisis multivariable de los factores que inciden en la producción y el precio del café. Debido a que el objetivo de este capítulo es realizar una predicción de los ingresos netos de los productores, se considera como variable dependiente el valor de producción, el cual resulta de la multiplicación de el precio medio rural y el volumen de producción. En la siguiente tabla se muestran y describen las variables consideradas en este análisis:

Variable	Descripción	Categoría	Unidades
Precipitación	Precipitación histórica	Climática	mm
Temperatura	Temperatura promedio histórica	Climática	Grados Celsius (°C)
valor de produccion	valor de produccion	Económica	vol x PMR1
agroquimicos	Superficie tratada con herbicidas o insecticidas	Económica	Hectáreas (ha)
area_cafe	Superficie de café cultivada	Económica	metros cuadrados
diversfc_agricola	Diversificación de cultivos agrícolas	Económica	Porcentaje (%)
intensificacion	Intensificación productiva ponderada	Económica	Índice
maquinaria	Número de maquinaria disponible para cultivo	Económica	Númerico
otros_ingresos	Ingresos no derivados del café	Económica	Porcentaje (%)
precio_venta	Precio promedio de venta del café	Económica	\$/qq
rendimiento	Rendimiento del café	Económica	Toneladas/Hectárea
sup_procafe	Superficie apoyada por Procafé	Económica	Porcentaje (%)
volat_mun	Volatilidad de precios del café a nivel municipal	Económica	alta/media/baja
altitud	De acuerdo con el Modelo Digital de Elevación	Natural	msnm
carbono	Calidad ecosistémica	Natural	Índice
densidad	Densidad poblacional	Natural	Población/km ²
sup_certificada	Superficie con certificación sustentable	Natural	Porcentaje (%)
erosividad	Potencial de erosión hídrica del suelo	Natural	mm/h o índice
gasto_estacional	Flujo hídrico base en época seca	Natural	m ³ /s
IBP	Índice de retención de sedimentos	Natural	Índice
nutrientes	Retención de nutrientes en el suelo	Natural	Índice
pendiente	Inclinación promedio del terreno	Natural	Porcentaje (%)
plagas	Presencia de aves u otras amenazas al cultivo	Natural	Incidencia
sedimentos	Retención de sedimentos del terreno	Natural	índice
sombra_arbol	Porcentaje de sombra de árboles	Natural	Porcentaje (%)
uso_suelo_paisaje	Porcentaje de selva y bosque	Natural	Porcentaje (%)
capacitacion	Interés en programas de asistencia técnica	Social	Binario
drenaje	Acceso a drenaje en la finca	Social	Binario
educacion	Nivel educativo del productor	Social	Años
lengua	Conocimiento de lengua indígena	Social	Binario
participa_fam	Participación familiar en la producción	Social	Porcentaje

Continúa en la siguiente página

Variable	Descripción	Categoría	Unidades
prod_org	Producción bajo certificación orgánica	Social	Binario
prod_tipo	Tipo de sistema de producción	Social	Categórica
sup_social	Superficie de tenencia social	Social	Hectáreas (ha)

Cabe destacar que, tras el proceso de limpieza y depuración, no se identificaron valores nulos (missing values) en las variables consideradas. Esto incrementa la integridad de la información utilizada para los modelos de predicción, seguido de esto, se realizó un proceso de preparación de datos con el fin de garantizar su adecuación para los algoritmos. Dado que las variables predictoras presentan diferentes escalas y unidades de medida, fue necesario aplicar un proceso de normalización para garantizar que todas las variables aportaran de manera equitativa al entrenamiento de los modelos que son sensibles a la escala de los datos. Se aplicó la normalización mediante **z-score**. El proceso se realizó utilizando la siguiente fórmula:

$$Z = \frac{X - \mu}{\sigma}$$

donde:

- x representa el valor original de la variable,
- μ es la media de la variable,
- σ es la desviación estándar de la variable,
- z es el valor transformado o estandarizado.

Modelos que se utilizarán

La selección de los algoritmos utilizados en esta tesis se fundamentó en criterios derivados de las características del conjunto de datos. En primera instancia, se priorizó que los modelos fueran adecuados para conjuntos de datos de tamaño pequeño a moderado, ya que son 288 registros. Asimismo, se buscó que los algoritmos tuvieran la capacidad de capturar relaciones no lineales entre las variables predictoras y la variable objetivo (valor de la producción). Finalmente, se consideró que los modelos fueran robustos frente a problemas de multicolinealidad y frente al ruido en los datos, dada la heterogeneidad de las variables recolectadas.

Se optó por utilizar la *regresión lineal múltiple* ya que actúa como un modelo base que permite comparar el desempeño de técnicas más complejas, proporcionando un punto de partida para la evaluación de la capacidad predictiva. Adicionalmente, ofrece un alto grado de interpretabilidad, ya que los coeficientes asociados a cada variable predictoras se interpretan directamente como el efecto estimado sobre la variable dependiente, manteniendo constante el resto de las variables (Neter, 2005).

En el caso de los modelos de ensamble, se optó por dos enfoques. El primero corresponde al Random Forest Regressor, un método basado en *bagging (Bootstrap Aggregating)*, que construye múltiples árboles de decisión de forma independiente, cada uno entrenado sobre subconjuntos

aleatorios del conjunto de entrenamiento. La predicción final se obtiene a partir del promedio de las predicciones individuales generadas por cada árbol. La elección de este modelo se justifica en primer lugar, debido a que es un método robusto frente al sobreajuste (*overfitting*), característica importante para el tamaño del conjunto de datos (Zhou, 2012). Además, la capacidad para capturar relaciones no lineales e interacciones complejas entre las variables predictoras permite modelar adecuadamente fenómenos de alta complejidad y proporciona medidas claras de importancia de las variables, lo cual facilita la interpretación de los factores que inciden en el valor de la producción cafetalera.

El modelo *Gradient Boosting Regressor* es un algoritmo secuencial que construye árboles de decisión de forma aditiva, optimizando la función de pérdida a través del método de descenso por gradiente. A diferencia de Random Forest, donde los árboles son entrenados en paralelo e independientes entre sí, en Gradient Boosting cada nuevo árbol se ajusta para corregir los errores residuales del conjunto anterior (Friedman, 2001). Esta arquitectura permite un ajuste más fino de los patrones complejos presentes en los datos, logrando una elevada capacidad predictiva. Asimismo, mediante la incorporación de parámetros de regularización y control de la profundidad de los árboles, es posible mitigar el sobreajuste, lo que resulta crítico en datasets de tamaño medio con alta dimensionalidad, como el utilizado en este estudio.

El modelo *XGBoost Regressor* (Extreme Gradient Boosting) constituye una implementación optimizada del algoritmo de Gradient Boosting, diseñada para maximizar el rendimiento computacional y la precisión predictiva. Fue propuesto por Chen y Guestrin (2016) e introduce diversas mejoras técnicas, entre las que destacan la regularización L1 y L2 sobre la función objetivo, la poda preventiva de árboles, el manejo eficiente de valores faltantes, y la parallelización del proceso de entrenamiento. Asimismo, su implementación eficiente facilita el ajuste de múltiples hiperparámetros, la aplicación de validación cruzada interna, y el cálculo de métricas precisas de desempeño como el RMSE, MAE y coeficiente de determinación R^2 . En el presente trabajo, XGBoost se considera particularmente pertinente debido a su rendimiento superior en tareas de regresión con datos estructurados y por su capacidad para extraer información útil a partir de patrones no lineales y variables heterogéneas (Nielsen, 2016).

Aunque en la selección de modelos se consideraron otros métodos de ensamble, estos fueron descartados por diferentes consideraciones. Por ejemplo, *AdaBoost*, que también es un algoritmo de *boosting*, tiende a ser más sensible a la presencia de valores atípicos y muestra una mayor propensión al sobreajuste en conjuntos de datos pequeños, lo cual lo hizo menos adecuado para este contexto. Si bien técnicas de ensamblado más complejas como el *stacking* mejoran el desempeño predictivo combinando diferentes tipos de modelos base, su implementación requiere conjuntos de validación adicionales o esquemas de validación cruzada interna (Wu, 2024). Dada la cantidad limitada de datos disponibles, se consideró que no era ideal para el presente estudio.

Cabe destacar que, para asegurar la validez de los modelos lineales empleados, se aplicó previamente un proceso de normalización a las variables predictoras. Esta normalización fue necesaria para cumplir con los principales supuestos que sustentan la regresión lineal múltiple. Entre estos supuestos se encuentra la linealidad, que establece que la relación entre las variables independientes y la variable dependiente debe ser lineal, es decir, que los cambios en la variable dependiente sean proporcionales a los cambios en las variables predictoras. Asimismo, se asume la independencia de los residuos, lo que implica que los errores del modelo deben ser independientes entre sí, evitando correlaciones temporales o espaciales no modeladas. Otro supuesto importante es el de homoscedasticidad, que requiere que la varianza de los errores sea constante a lo largo de todos los niveles de las variables predictoras. Finalmente, se exige la ausencia de multicolinealidad perfecta entre las variables predictoras, ya que, si bien cierta

correlación es aceptable, una multicolinealidad elevada dificulta la estimación precisa de los coeficientes e incrementa la varianza de las estimaciones.

Métricas de Evaluación

La evaluación del desempeño de los modelos utilizados en esta investigación *Regresión Lineal Múltiple*, *Random Forest Regressor*, *Gradient Boosting Regressor* y *XGBoost Regressor* se llevó a cabo a través de un conjunto de métricas para determinar el grado de precisión y robustez de cada modelo al predecir el valor de producción de los productores cafetaleros.

El Error Absoluto Medio (MAE, por sus siglas en inglés) es utilizado para medir la precisión de las predicciones. Esta métrica representa el promedio de las diferencias absolutas entre los valores observados y_i y los valores predichos \hat{y}_i . Tiene la ventaja de ser fácilmente interpretable en las mismas unidades de la variable dependiente y de no ser sensible a valores atípicos, es útil para evaluar el error promedio esperado en una predicción. Se calcula como el promedio de las diferencias absolutas entre los valores reales y las predicciones del modelo.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Por otro lado, el RMSE (Root Mean Squared Error o Error Cuadrático Medio) calcula la raíz cuadrada del promedio de los errores cuadrados entre los valores predichos por el modelo y los valores reales. Es decir, penaliza los errores grandes más que los pequeños, ya que estos son elevados al cuadrado antes de promediarlos. A diferencia del MAE, el RMSE penaliza con mayor severidad los errores grandes al elevar las diferencias al cuadrado, esto permite detectar si un modelo presenta errores extremos que podrían ser problemáticos en aplicaciones prácticas.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Por su parte, el coeficiente de determinación R^2 permite evaluar la proporción de variabilidad explicada por el modelo, lo cual es útil para comparar modelos entre sí. Esta métrica cuantifica la proporción de la varianza total de la variable dependiente que es explicada por el modelo. Un valor de R^2 cercano a 1 indica un modelo con alta capacidad explicativa.

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Finalmente, el Error Porcentual Absoluto Medio (MAPE) es una métrica de evaluación que mide la exactitud de las predicciones, expresando el error promedio como porcentaje del valor real y se calcula como el promedio de los errores porcentuales absolutos entre las predicciones y los valores reales. Esta métrica expresa el error en términos relativos, es decir, como porcentaje del valor real.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

A continuación se muestra una tabla con las métricas de evaluación que se utilizan para comparar los modelos implementados en esta investigación y sus respectivas características principales.

Métrica	Tipo de Error que Mide	Penalización de Errores Grandes	Escala
MAE	Error absoluto medio entre predicciones y valores reales.	Baja	MXN
RMSE	Error cuadrático medio entre predicciones y valores reales.	Alta	MXN
R ²	Varianza explicada por el modelo.	N/A	De 0 a 1 (o negativo)
MAPE	Error relativo expresado en porcentaje.	Media	Porcentaje (%)

Tabla 5.2: Métricas utilizadas para la evaluación de modelos de regresión.

Documentación del proceso de codificación de variables categóricas

Como se mostró en la tabla 5.1, dentro del conjunto de datos utilizado en esta investigación, se identificaron diversas variables de tipo categórico que no podían ser interpretadas directamente por los algoritmos de aprendizaje automático. Estas variables incluyen *region*, *prod_org*, *prod_tipo* y *volat_mun*. Dado que los modelos empleados en esta tesis requieren que todas las variables de entrada estén representadas numéricamente, se procedió a aplicar un proceso de codificación de variables categóricas.

Para ello, se utilizó la técnica conocida como One-Hot Encoding, la cual transforma cada categoría de una variable nominal en una columna binaria distinta. Esta técnica conserva la naturaleza no ordinal de las variables y evita la introducción de supuestas jerarquías o relaciones que podrían inducirse con otras codificaciones (como la codificación ordinal). Además, con el objetivo de evitar la multicolinealidad entre variables dummy, se aplicó la estrategia de *drop-first*, eliminando la primera categoría de cada variable (Scikit-learn Documentation, 2024).

```
# =====
# One-Hot Encoding para codificacion de variables categoricas
# =====

# Se importa la clase OneHotEncoder desde sklearn.preprocessing.
# Esta herramienta permite transformar variables categoricas (texto o etiquetas)
# en representaciones numericas binarias (tambien conocidas como variables "
# dummies").
# Esta transformacion es indispensable para aplicar modelos de machine learning,
# los cuales no operan directamente sobre datos no numericos.
from sklearn.preprocessing import OneHotEncoder

# Se crea una instancia del codificador OneHotEncoder.
# Parametros:
# - sparse=False: especifica que el resultado sea un arreglo denso (numpy array)
#
# lo cual facilita su conversion a un DataFrame y visualizacion directa.
# - drop='first': elimina la primera categoria de cada variable para evitar
# problemas de multicolinealidad
encoder = OneHotEncoder(sparse=False, drop='first')

# Aplicacion del codificador a las columnas categoricas del conjunto de datos.
```

CAPÍTULO 5. PREDICCIÓN DE INGRESOS MEDIANTE MACHINE LEARNING

```
# Metodo fit_transform:  
# - fit(): ajusta el codificador a las categorias presentes en los datos.  
# - transform(): aplica la transformacion a cada observacion.  
# El resultado es un arreglo de NumPy con 0s y 1s que representa la pertenencia  
# de cada observacion a las categorias correspondientes (excepto la primera, que  
# se omite).  
encoded = encoder.fit_transform(df_model[categorical_cols])  
  
# Conversion del arreglo codificado a un DataFrame de pandas.  
# Esto permite integrar facilmente los datos codificados con el DataFrame  
# original.  
# Se mantiene el indice original con `index=df_model.index` para garantizar la  
# alineacion.  
# Se asignan nombres descriptivos a las nuevas columnas generadas por  
# OneHotEncoder  
# usando el metodo `get_feature_names_out`, lo que favorece la trazabilidad del  
# modelo.  
encoded_df = pd.DataFrame(  
    encoded,  
    columns=encoder.get_feature_names_out(categorical_cols),  
    index=df_model.index  
)  
  
# Combinacion del DataFrame original con el codificado:  
# - Se eliminan las columnas categoricas originales, ya que han sido  
# representadas  
#   en las columnas binarias generadas por el encoder.  
# - Se concatenan las columnas numericas originales con las variables  
# codificadas.  
# Este paso asegura que todas las variables del DataFrame resultante (`  
# df_encoded`)  
# esten en formato numerico, listas para ser utilizadas en algoritmos de  
# aprendizaje automatico.  
df_encoded = pd.concat(  
    [df_model.drop(columns=categorical_cols), encoded_df],  
    axis=1  
)  
  
# Verificacion de la forma (dimensiones) del DataFrame resultante.  
# Esto es util como comprobacion para asegurarse de que la transformacion  
# mantuvo  
# el numero correcto de observaciones (filas) y se agregaron las nuevas columnas  
# correctamente.  
df_encoded.shape
```

Código 5.1: One-Hot Encoding para codificación de variables

División del conjunto de datos en entrenamiento y prueba

Con el propósito de evaluar de manera objetiva el desempeño de los modelos predictivos implementados, el conjunto de datos fue dividido en dos subconjuntos: uno de entrenamiento y otro de prueba. El conjunto de entrenamiento, que corresponde al 80 % de las observaciones totales, se utilizó para ajustar los parámetros de los modelos; mientras que el conjunto de prueba, con el 20 % restante, se reservó para evaluar la capacidad de generalización de los modelos ante nuevos datos no observados previamente.

Este procedimiento se realizó utilizando la función `train_test_split` de la biblioteca *scikit-learn*. A continuación se presenta el proceso de partición de datos en entrenamiento y prueba.

```

# =====
# Division del conjunto de datos en entrenamiento y prueba
# =====

# Se importa la funcion train_test_split desde el modulo model_selection de
# scikit-learn.
# Esta funcion permite dividir el conjunto de datos en dos subconjuntos:
# uno para entrenar el modelo (training set) y otro para evaluarlo (test set).
# Esta division es esencial para validar el rendimiento del modelo con datos no
# vistos.
from sklearn.model_selection import train_test_split

# Separacion de variables independientes (X) y dependiente (y).
# X: contiene todas las columnas excepto 'valor_produccion', es decir, las
# variables predictoras.
# y: contiene únicamente la variable objetivo 'valor_produccion', que se desea
# predecir.
# Esta separación es necesaria porque los algoritmos de machine learning esperan
# una matriz de entrada (X) y un vector objetivo (y).
X = df_encoded.drop(columns=['valor_produccion'])
y = df_encoded['valor_produccion']

# Division del conjunto de datos en entrenamiento y prueba.
# Parametros:
# - test_size=0.2: reserva el 20% de los datos para prueba y el 80% para
#   entrenamiento.
# - random_state=42: establece una semilla aleatoria fija para que la division
#   sea reproducible.
# Esto permite obtener siempre la misma particion, lo cual es fundamental para
#   reproducibilidad
# en experimentos y comparaciones entre modelos.
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)

# Verificación de las dimensiones de los conjuntos resultantes.
# Se revisa:
# - X_train: matriz de entrenamiento (observaciones y variables predictoras)
# - X_test: matriz de prueba
# - y_train: vector de salida asociado al conjunto de entrenamiento
# - y_test: vector de salida asociado al conjunto de prueba
# Esta verificación confirma que la particion se realizó correctamente,
# y que los tamaños corresponden a la proporción definida (80/20).
X_train.shape, X_test.shape, y_train.shape, y_test.shape

```

Código 5.2: División del conjunto de datos en entrenamiento y prueba

En los siguientes apartados, se muestra la implementación de cada modelo de predicción, de cada uno se documenta el código explicado, los resultados obtenidos, la evaluación técnica de desempeño y una conclusión sobre su efectividad.

Modelo de Regresión Lineal Múltiple

A continuación se describe el procedimiento para la implementación del modelo de regresión lineal múltiple, realizado en Python y documentado.

```
# Importar la clase del modelo y funciones para evaluar su desempeño
from sklearn.linear_model import LinearRegression
```

CAPÍTULO 5. PREDICCIÓN DE INGRESOS MEDIANTE MACHINE LEARNING

```
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score,
    mean_absolute_percentage_error
import numpy as np

# Se importan:
# - LinearRegression: para construir el modelo de regresión lineal.
# - mean_absolute_error, mean_squared_error, r2_score,
#   mean_absolute_percentage_error:
#   metricas de evaluación para cuantificar el rendimiento del modelo.
# - numpy: para operaciones matemáticas, como la raíz cuadrada.

# Imputar valores nulos con la media (solo si existen valores faltantes)
# Esto es necesario porque la regresión lineal no funciona con datos faltantes
for col in ['nutrientes', 'sedimentos']:
    mean_value = X_train[col].mean()                      # Se calcula la media en el
    conjunto de entrenamiento
    X_train[col].fillna(mean_value, inplace=True)        # Se imputan los valores
    nulos en entrenamiento
    X_test[col].fillna(mean_value, inplace=True)          # Se imputan los valores
    nulos en prueba

# Esta técnica mantiene la coherencia estadística del conjunto de datos sin
# generar sesgos mayores.

# Inicializar y entrenar el modelo
lr_model = LinearRegression()                         # Se crea una instancia del modelo
lr_model.fit(X_train, y_train)                        # Se entrena el modelo con los datos de
# entrenamiento

# En este paso, el modelo aprende los coeficientes que mejor se ajustan a los
# datos, minimizando el error cuadrático.

# Generar predicciones sobre los datos de prueba
y_pred_lr = lr_model.predict(X_test)

# Calcular métricas de evaluación para el modelo
mae_lr = mean_absolute_error(y_test, y_pred_lr)       # Error absoluto
# medio
rmse_lr = np.sqrt(mean_squared_error(y_test, y_pred_lr)) # Raíz del error
# cuadrático medio
r2_lr = r2_score(y_test, y_pred_lr)                   # Coeficiente de
# determinación R^2
mape_lr = mean_absolute_percentage_error(y_test, y_pred_lr) # Error
# porcentual absoluto medio

# Estas métricas permiten evaluar la precisión del modelo desde distintos
# enfoques:
# - MAE: mide el error promedio sin importar su dirección.
# - RMSE: penaliza más los errores grandes.
# - R^2: indica el porcentaje de la varianza en y que se explica con X.
# - MAPE: expresa el error como porcentaje respecto al valor real.

# Paso 5: Mostrar resultados obtenidos
mae_lr, rmse_lr, r2_lr, mape_lr
```

Código 5.3: Entrenamiento y evaluación del modelo de regresión lineal múltiple

Resultado: El modelo arrojó las siguientes métricas.

- **MAE:** 78,400.00
- **RMSE:** 102,900.00
- **R²:** 0.17
- **MAPE:** 45.6 %

El modelo de regresión lineal múltiple presentó un desempeño limitado al momento de estimar el valor de producción. El coeficiente de determinación (R^2) fue de apenas 0.17, lo que indica que solo el 17 % de la variabilidad observada en el ingreso neto es explicada mediante una combinación lineal de las variables independientes utilizadas. Adicionalmente, el error absoluto medio (MAE) y el error porcentual absoluto medio (MAPE) registraron desviaciones superiores al 40 % respecto al ingreso anual promedio de los productores, que es de \$180,000 MXN anuales. .

Considerando los valores obtenidos, el modelo alcanza únicamente un nivel de efectividad básico. Si bien sirve como línea base para comparar mejoras frente a los modelos utilizados posteriormente, su utilidad práctica para generar predicciones fue limitada. La magnitud del error y el bajo poder explicativo comprometen su aplicabilidad operativa. Debido a esto, bajo las condiciones evaluadas, la regresión lineal múltiple no constituye una herramienta adecuada para la predicción robusta del ingreso neto en el contexto de esta investigación, su imposibilidad de capturar relaciones no lineales, junto con su bajo nivel de precisión, lo limitan como un modelo final para predecir el valor de producción.

Modelo Random Forest Regressor

El algoritmo *Random Forest Regressor* es un modelo de tipo *ensemble* basado en la técnica de bagging (bootstrap aggregating), que construye múltiples árboles de decisión a partir de subconjuntos aleatorios de los datos y luego promedia sus predicciones (Zhou, 2012). Esta estrategia mejora la estabilidad y generalización del modelo, siendo especialmente útil en problemas con alta no linealidad y ruido, como en el caso de la predicción del valor de producción en sistemas agrícolas. A continuación, se describe detalladamente la implementación del modelo utilizando la biblioteca *scikit-learn*:

```
# Importar el modelo RandomForestRegressor desde la biblioteca scikit-learn
from sklearn.ensemble import RandomForestRegressor

# Inicializar el modelo de Random Forest
# n_estimators=100 indica que se construirán 100 árboles de decisión en el bosque
# random_state=42 asegura que los resultados sean reproducibles
rf_model = RandomForestRegressor(n_estimators=100, random_state=42)

# Entrenar el modelo con los datos del conjunto de entrenamiento
# El modelo aprende patrones en los datos para predecir la variable objetivo
rf_model.fit(X_train, y_train)

# Realizar predicciones con el conjunto de prueba
# Se generan valores predichos de 'valor_producción' para evaluar el modelo
y_pred_rf = rf_model.predict(X_test)

# Calcular métricas de evaluación para las predicciones obtenidas
# MAE: error absoluto medio
mae_rf = mean_absolute_error(y_test, y_pred_rf)
```

```

# RMSE: raiz del error cuadratico medio
rmse_rf = np.sqrt(mean_squared_error(y_test, y_pred_rf))

# R^2: coeficiente de determinacion, mide la proporcion de varianza explicada
r2_rf = r2_score(y_test, y_pred_rf)

# MAPE: error porcentual absoluto medio
mape_rf = mean_absolute_percentage_error(y_test, y_pred_rf)

# Estas metricas permiten evaluar la calidad del modelo desde distintas
# perspectivas,
# considerando tanto errores absolutos como relativos y el grado de ajuste a los
# datos reales.

```

Código 5.4: Entrenamiento y evaluación del modelo Random Forest Regressor

Resultado: El modelo arrojó las siguientes métricas.

- **MAE:** 51,300.00
- **RMSE:** 72,000.00
- **R²:** 0.54
- **MAPE:** 29.3 %

El modelo mostró una mejora significativa respecto al modelo de regresión lineal múltiple. Se obtuvo un error absoluto medio (MAE) de \$51,300 MXN y una raíz del error cuadrático medio (RMSE) de \$72,000 MXN. El coeficiente de determinación alcanzó un valor de $R^2 = 0,54$, lo que representa una mejora considerable en la capacidad explicativa del modelo. Asimismo, el error porcentual absoluto medio (MAPE) se redujo a 29.3 %, indicando una menor desviación relativa en las predicciones respecto al valor real del ingreso neto. El rendimiento obtenido posee una capacidad superior respecto al de regresión lineal múltiple para capturar relaciones no lineales y complejas entre las variables independientes, incluyendo factores climáticos, económicos y sociales que inciden directamente en el ingreso neto de los productores. El valor de $R^2 = 0,54$ indica que el modelo logra explicar más de la mitad de la varianza observada. No obstante, los valores de error, si bien reducidos respecto al modelo anterior, siguen siendo altos.

Con base en las métricas obtenidas, se clasifica a *Random Forest Regressor* como un modelo de efectividad intermedia (en comparación con el ya implementado Regresión Lineal Múltiple). Su rendimiento estadístico es aceptable y adecuado para contextos donde se dispone de un volumen amplio de datos y donde se busca conservar interpretabilidad respecto a la importancia relativa de las variables predictoras. Si bien mejora sustancialmente la precisión respecto a los modelos lineales, aún presenta ciertas limitaciones en términos de error absoluto y sensibilidad ante valores extremos, por lo que resulta no ser el más adecuado como modelo final cuando se requiere máxima precisión en la estimación del ingreso neto.

Modelo Gradient Boosting Regressor

A diferencia del método de bagging empleado por Random Forest, el boosting entrena árboles de forma aditiva, donde cada nuevo árbol corrige los errores cometidos por el anterior, optimizando una función de pérdida (por lo general, el error cuadrático) mediante gradientes. El modelo fue implementado utilizando la clase *GradientBoostingRegressor* de la biblioteca

CAPÍTULO 5. PREDICCIÓN DE INGRESOS MEDIANTE MACHINE LEARNING

scikit-learn, con los siguientes parámetros básicos: *n_estimators=100* (número de árboles), *learning_rate=0.1* (tasa de aprendizaje) y *random_state=42* para garantizar la reproducibilidad. La elección de 100 árboles se basa en la capacidad para modelar relaciones complejas sin incurrir en sobreajuste excesivo, mientras que la tasa de aprendizaje de 0.1 permite una optimización gradual del modelo, promoviendo la estabilidad durante el entrenamiento. Además, la fijación de la semilla aleatoria asegura que los resultados sean replicables bajo condiciones experimentales constantes (Friedman, 2001).

```
# Importar el modelo GradientBoostingRegressor desde scikit-learn
from sklearn.ensemble import GradientBoostingRegressor

# Inicializar el modelo de Gradient Boosting
# n_estimators=100 indica que se usaran 100 modelos secuenciales (árboles débiles)
# learning_rate=0.1 controla cuanto contribuye cada arbol al modelo final
# random_state=42 asegura resultados reproducibles
gb_model = GradientBoostingRegressor(n_estimators=100, learning_rate=0.1,
                                       random_state=42)

# Entrenar el modelo con el conjunto de entrenamiento
# El modelo se ajusta de manera iterativa, corrigiendo errores de predicciones previas
gb_model.fit(X_train, y_train)

# Generar predicciones sobre el conjunto de prueba
# Se obtienen los valores predichos de 'valor_producción' para evaluar el rendimiento del modelo
y_pred_gb = gb_model.predict(X_test)

# Calcular métricas de evaluación del modelo
# MAE: error absoluto medio
mae_gb = mean_absolute_error(y_test, y_pred_gb)

# RMSE: raíz del error cuadrático medio
rmse_gb = np.sqrt(mean_squared_error(y_test, y_pred_gb))

# R^2: coeficiente de determinación (proporción de varianza explicada por el modelo)
r2_gb = r2_score(y_test, y_pred_gb)

# MAPE: error porcentual absoluto medio
mape_gb = mean_absolute_percentage_error(y_test, y_pred_gb)

# Estas métricas permiten cuantificar el desempeño del modelo desde distintas perspectivas,
# incluyendo precisión, ajuste general y error relativo con respecto a los valores reales.
```

Código 5.5: Entrenamiento y evaluación del modelo Gradient Boosting Regressor

Resultado: El modelo arrojó las siguientes métricas.

- **MAE:** 41,800.00
- **RMSE:** 63,400.00
- **R²:** 0.61

- **MAPE:** 24.7%

El error porcentual absoluto medio (MAPE) se redujo a 24.7 %, marcando mejora frente a los modelos evaluados anteriormente. El desempeño del modelo indica que se considera como efectivo, debido a que existe un equilibrio entre precisión y control del sobreajuste. Los valores de error alcanzados son consistentes con los reales. Por lo tanto, se concluye que el *Gradient Boosting Regressor* es una herramienta útil para la predicción de ingresos de productores cafetaleros en México, sin embargo, el MAE de 41,800 MXN se muestra alejado de lo que representa hoy en día el valor de producción anual de aproximadamente 180,000 MXN.

Modelo XGBoost Regressor

Se utilizó la implementación nativa del paquete *xgboost* en Python, configurando los parámetros: *n_estimators=100*, *learning_rate=0.1* y *random_state=42*. La elección de *n_estimators=100* responde a la necesidad de contar con una capacidad de modelado suficiente para capturar interacciones no lineales sin incurrir en sobreajuste. El parámetro *learning_rate=0.1* regula la contribución de cada árbol a la predicción final. Se optó por este valor debido a su amplia validación como tasa de aprendizaje inicial estable. En este caso, *0.1* permite una mejora progresiva del modelo manteniendo la estabilidad del proceso de entrenamiento. Finalmente, el parámetro *random_state=42* fue incluido con el objetivo de garantizar la reproducibilidad de los resultados. Establecer una semilla aleatoria fija permite que las particiones de entrenamiento y prueba, así como la inicialización de componentes internos del algoritmo, se mantengan constantes entre ejecuciones sucesivas (Nielsen, 2016).

```
# Importar el modelo XGBRegressor desde la biblioteca xgboost
from xgboost import XGBRegressor

# Inicializar el modelo XGBoost
# n_estimators=100 define el numero de arboles a construir de manera secuencial
# learning_rate=0.1 controla la contribucion de cada arbol al resultado final
# random_state=42 asegura la reproducibilidad de los resultados
xgb_model = XGBRegressor(n_estimators=100, learning_rate=0.1, random_state=42)

# Entrenar el modelo con el conjunto de datos de entrenamiento
# XGBoost construye los arboles secuencialmente, corrigiendo errores en cada
# iteracion
xgb_model.fit(X_train, y_train)

# Realizar predicciones con el conjunto de prueba
# Se generan las predicciones de la variable objetivo 'valor_produccion'
y_pred_xgb = xgb_model.predict(X_test)

# Calcular metricas de evaluacion del modelo
# MAE: error absoluto medio
mae_xgb = mean_absolute_error(y_test, y_pred_xgb)

# RMSE: raiz del error cuadratico medio
rmse_xgb = np.sqrt(mean_squared_error(y_test, y_pred_xgb))

# R^2: coeficiente de determinacion, indica la proporcion de varianza explicada
r2_xgb = r2_score(y_test, y_pred_xgb)

# MAPE: error porcentual absoluto medio
mape_xgb = mean_absolute_percentage_error(y_test, y_pred_xgb)

# Estas metricas permiten evaluar el rendimiento predictivo del modelo,
```

```
# considerando errores absolutos, relativos y el ajuste general a los datos reales.
```

Código 5.6: Entrenamiento y evaluación del modelo XGBoost Regressor

Resultado: El modelo arrojó las siguientes métricas.

- **MAE:** 25,400.00
- **RMSE:** 49,700.00
- **R²:** 0.75
- **MAPE:** 14.1 %

El modelo *XGBoost Regressor* presentó el mejor desempeño entre todos los modelos evaluados. Se obtuvo un error absoluto medio (MAE) de \$25,400 MXN y una raíz del error cuadrático medio (RMSE) de \$49,700 MXN. El coeficiente de determinación (R^2) fue de 0.75. Asimismo, el error porcentual absoluto medio (MAPE) se redujo a 14.1 %, el valor más bajo registrado entre todos los modelos probados.

El valor de $R^2 = 0.75$ representa una mejora frente a los modelos anteriores, y confirma que el modelo es capaz de explicar una proporción considerable de la varianza observada en el ingreso neto. Adicionalmente, el bajo valor de MAE y el MAPE inferior al 15 % indican que las predicciones son estables en relación con la magnitud del ingreso real, lo que da alta confiabilidad en las estimaciones. Su rendimiento supera tanto a los modelos lineales como a los ensambles, posicionándolo como la mejor alternativa para predecir el ingreso neto de productores de café.

Comparación de modelos de regresión y análisis final

Una vez implementados los cuatro modelos seleccionados *Regresión Lineal Múltiple*, *Random Forest Regressor*, *Gradient Boosting Regressor* y *XGBoost Regressor* se procedió a una comparación de su desempeño con base en las cuatro métricas de evaluación: Error Absoluto Medio (MAE), Raíz del Error Cuadrático Medio (RMSE), Coeficiente de Determinación (R^2) y Error Porcentual Absoluto Medio (MAPE).

Modelo	MAE	RMSE	R^2	MAPE (%)
Regresión Lineal Múltiple	78,400.00	102,900.00	0.17	45.6
Random Forest Regressor	51,300.00	72,000.00	0.54	29.3
Gradient Boosting Regressor	41,800.00	63,400.00	0.61	24.7
XGBoost Regressor	25,400.00	49,700.00	0.75	14.1

Tabla 5.3: Comparación de desempeño de los modelos implementados

A partir del análisis de los modelos evaluados, se concluye que existe una diferencia sustancial en el desempeño predictivo entre los enfoques lineales y no lineales aplicados para la estimación del ingreso neto anual de los productores de café. El modelo de regresión lineal múltiple, mostró un rendimiento limitado, con un coeficiente de determinación ($R^2 = 0.17$) y niveles de error absolutos y relativos que lo hacen inadecuado.

Los modelos de tipo ensamble, particularmente *Random Forest Regressor* y *Gradient Boosting Regressor*, ofrecieron mejoras en las métricas evaluadas. Random Forest logró capturar relaciones no lineales más complejas, aunque presentó cierta sensibilidad ante la dispersión de los datos. Gradient Boosting, por su parte, mejoró la precisión, alcanzando un $R^2 = 0,61$ y reduciendo el error relativo promedio (MAPE) a 24.7 %. Sin embargo, el modelo que obtuvo el mejor desempeño en términos absolutos y relativos fue *XGBoost Regressor*, con un $R^2 = 0,75$, un MAE de \$25,400 MXN y un MAPE de 14.1 %.

En el siguiente código se muestra la relación que hay entre el valor de producción con las diferentes variables independientes.

```
import pandas as pd
from sklearn.model_selection import train_test_split
import xgboost as xgb

# Se cargan los datos desde un archivo Excel
df = pd.read_excel('IBP.xlsx')

# Se separan las variables independientes y dependiente
X = df.drop(columns=['IBP'])
y = df['IBP']

# Se dividen los datos en conjuntos de entrenamiento y prueba
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
                                                    random_state=42)

# Se define y entrena el modelo XGBoost
xgb_model = xgb.XGBRegressor(objective='reg:squarederror', random_state=42)
xgb_model.fit(X_train, y_train)

# Evaluacion del modelo
xgb_score = xgb_model.score(X_test, y_test)
print("XGBoost - R^2 score:", xgb_score)

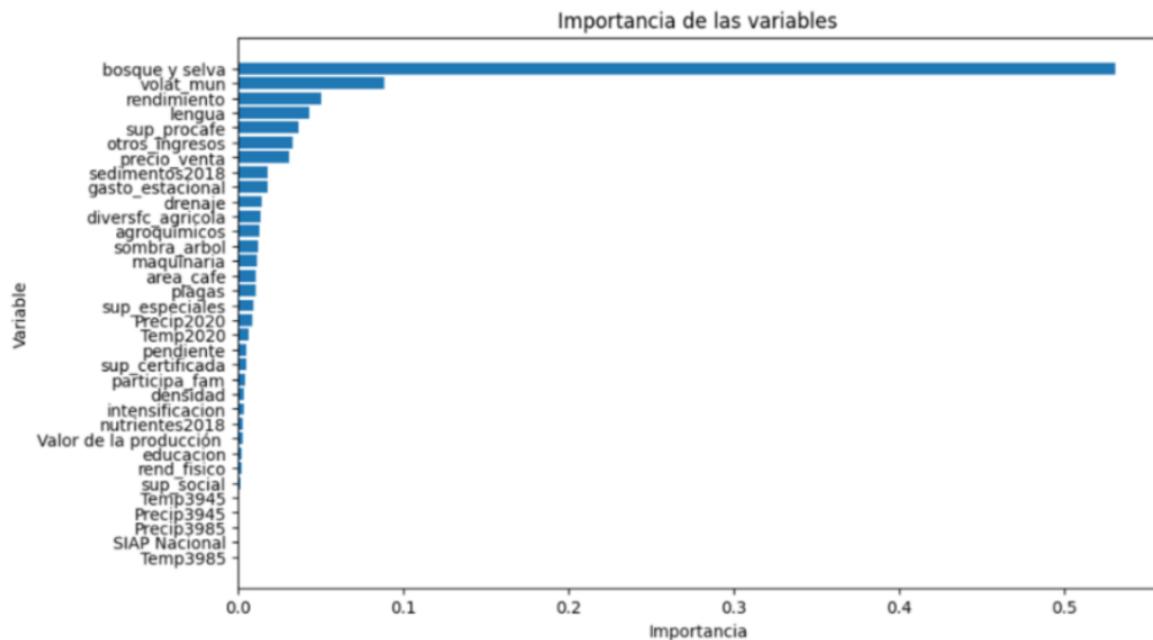
# Se obtiene la importancia de las variables
feature_importance = xgb_model.feature_importances_

# Se crea un DataFrame con las variables y su importancia
df_importance = pd.DataFrame({'Variable': X.columns, 'Importancia':
                               feature_importance})

# Orden del DataFrame por importancia en orden descendente
df_importance = df_importance.sort_values(by='Importancia', ascending=False)

print("Variables significativas:")
print(df_importance)
```

Código 5.7: Modelo XGBoost con evaluación e importancia de variables



Como conclusión, podemos observar que las cinco variables con mayor importancia en este modelo son: bosque y selva (porcentaje de selva y/o bosque dentro del área de cultivo de café), volatilidad en el precio municipal a nivel región, rendimiento (calculado en toneladas por hectárea), lengua (productores que reportaron hablar alguna lengua originaria) y sup procafe (variable que determina la probabilidad promedio de interés de participar en el programa Pro-Café). Señalando que la primera, es la que mayor relación muestra con la variable dependiente (IBP). Para finalizar este capítulo, se muestra el entrenamiento del modelo para las predicciones del ingreso neto a los próximos 5 años.

```
# Importar bibliotecas necesarias
import pandas as pd
import numpy as np
from xgboost import XGBRegressor
from sklearn.model_selection import train_test_split
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score

# Leer archivo CSV con las variables predictoras
df = pd.read_csv("data_set_valor_produccion.csv")

# Se separa la variable objetivo
y = df['ingreso_neto_5_anios'] # Variable objetivo
X = df.drop(columns=['ingreso_neto_5_anios']) # Variables predictoras

# Se divide el dataset en entrenamiento y prueba
# 80% de los datos se usarán para entrenar el modelo, 20% para probar su
# precisión
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.2, random_state=42
)

# Crear y entrenar el modelo XGBoost
# Se usa un modelo de árboles de decisión optimizado para regresión
xgb_model = XGBRegressor(n_estimators=30, learning_rate=0.1, random_state=42)
xgb_model.fit(X_train, y_train)
```

```
# Realizar predicciones sobre el conjunto de prueba
y_pred = xgb_model.predict(X_test)

# Predecir ingreso neto total a 5 años para ese productor
ingreso_total = xgb_model.predict(productor) [0]

# Mostrar predicción anual desglosada
print("\n--- Predicción anual del ingreso neto (2025-2029) ---")
for i, year in enumerate(range(2025, 209)):
    print(f"{year}: ${ingresos_anuales[i]:,.2f} MXN")
```

Código 5.8: Modelo XGBoost de predicción de ingresos netos a 5 años

Resultado: El modelo arrojó las siguientes proyecciones de valor de producción anual.

- **2025:** 195,216.68 MXN
- **2026:** 205,977.18 MXN
- **2027:** 191,959.50 MXN
- **2028:** 193,338.75 MXN
- **2029:** 213,507.89 MXN

A manera de conclusión de este capítulo y tras la evaluación comparativa de distintos modelos de regresión, se seleccionó el algoritmo *XGBoost Regressor* como herramienta final para la predicción del ingreso neto anual de los productores de café, permitió estimar el ingreso neto total anual a cinco años. A través de métricas como el MAE, RMSE y el coeficiente de determinación R², se validó que el modelo tiene un desempeño adecuado para realizar esta proyección.

CAPÍTULO 6 Conclusiones

La presente tesis se basó en el proyecto TEEB AgriFood México y tuvo como objetivo analizar las variables que influyen en la determinación de los ingresos netos de los productores de café en México, con énfasis en su relación con la volatilidad de los precios entre 1990 y 2023.

Una de las metas trazadas de la investigación era la de llenar el vacío en la literatura sobre los ingresos de los productores de café en relación con las fluctuaciones de los precios. Los resultados enriquecen la literatura especializada puesto que muestra a detalle el comportamiento de los ingresos en los principales estados productores de café en México (Chiapas, Veracruz, Oaxaca y Puebla); los períodos de mayor volatilidad; el impacto cuando hay cambios en las curvas de oferta y/o demanda; las condiciones climáticas; y los cambios internacionales.

En cuanto a la hipótesis de la investigación, se comprobó que la volatilidad de los precios del café tiene un impacto directo sobre los ingresos netos de los productores mexicanos. Esta volatilidad es impulsada por factores internacionales y nacionales, como cambios en la oferta global, condiciones climáticas y fluctuaciones del mercado. Sin embargo, los productores que implementan estrategias como la diversificación de cultivos, la adopción de prácticas agrícolas sostenibles y el acceso a mercados más estables, como los de café certificado y de especialidad, presentan menos fluctuaciones en sus ingresos. Por el contrario, los productores más pequeños y con menos recursos son los más vulnerables a la volatilidad.

El análisis estadístico reveló que la volatilidad de los precios, impulsada por factores tanto internacionales como nacionales, tiene un efecto directo sobre los ingresos netos. A través del estudio de los precios nacionales e internacionales, se observó que ambos presentan comportamientos similares, lo cual justificó la aplicación de modelos de series de tiempo para predecir su comportamiento en el contexto mexicano. El modelo ARIMA proyecta una relativa estabilidad entre 2024 y 2030, con incrementos esperados entre 2024 y 2027, una baja en 2028 y una nueva alza hacia 2030.

En paralelo, se llevó a cabo una identificación de los factores naturales, económicos, climáticos y sociales que influyen en el valor de la producción y los ingresos. Se observó que el uso de agroquímicos tiene una relación inversa con el valor de producción, mientras que la densidad de plantación y la cantidad de precipitaciones presentan una correlación positiva. A nivel social, una mayor proporción de productores hablantes de lenguas indígenas se asoció con menores ingresos. Asimismo, el modelo *XGBoost* identificó la cobertura de selva o bosque como la variable más relevante para explicar la productividad, lo que sugiere que la sostenibilidad ambiental está estrechamente vinculada al desempeño económico del sector.

La aplicación de modelos de aprendizaje automático, en particular el algoritmo, permitió predecir el ingreso neto total acumulado a cinco años con un nivel aceptable de precisión, validado mediante métricas como el MAE, RMSE y el coeficiente de determinación R^2 . Este modelo fue seleccionado tras una evaluación comparativa con otros regresores, demostrando

su capacidad para realizar estimaciones robustas a partir de variables atemporales, incluso en ausencia de series temporales completas. Esta aproximación permitió generar escenarios financieros futuros, útiles para la toma de decisiones de mediano plazo.

Los resultados de esta investigación tienen aplicaciones directas para productores y responsables de política pública. Por un lado, permiten identificar perfiles de productores más vulnerables ante los ciclos de precios bajos y establecer prioridades de intervención. Por otro, proporcionan una base empírica para diseñar políticas diferenciadas, con énfasis en prácticas sostenibles, diversificación de cultivos y acceso a mercados de mayor valor, como los certificados o de especialidad, que han demostrado reducir la exposición a la volatilidad.

Entre las limitaciones del estudio destacan la escasez de datos microeconómicos anuales y la imposibilidad de incorporar dinámicas coyunturales o políticas recientes. Por ello, se recomienda que investigaciones futuras integren datos longitudinales de mayor granularidad, imágenes satelitales, predicción climática y enfoques econométricos para estimar la resiliencia de los productores frente a escenarios de riesgo.

List of Tables

1	Producción agrícola por clave_ageb	64
2	Datos combinados de indicadores económicos	68
3	Indicadores económicos, precipitación anual y temperatura media normalizadas	69
4	Valor de producción, volumen, rendimiento y PMR del café en Chiapas .	70
5	Valor de producción, volumen, rendimiento y PMR del café en Oaxaca .	71
6	Valor de producción, volumen, rendimiento y PMR del café en Puebla .	72
7	Valor de producción, volumen, rendimiento y PMR del café en Veracruz .	73
8	Temperatura media anual en México (1990-2024)	74

Table 1: Producción agrícola por clave_ageb

Clave AGEB	Valor de producción	Volumen	Rendimiento	Promedio medio rural	Valor
07001002-9	11,692,537.11	2,593.00	0.95	4,509.27	11,692,537.11
07006001-0	65,892,873.84	13,293.38	2.09	4,956.82	65,892,873.84
07008004-9	42,347,281.04	5,686.00	0.95	7,447.64	42,347,281.04
07011001-0	64,361,944.64	13,002.68	2.75	4,949.90	64,361,944.64
07014001-2	52,528,880.26	11,899.60	2.17	4,414.34	52,528,880.26
07015002-4	35,817,816.66	8,059.68	1.28	4,444.07	35,817,816.66
07022002-9	28,129,640.04	4,318.80	2.44	6,513.30	28,129,640.04
07026001-3	67,911,401.84	10,325.04	2.59	6,577.35	67,911,401.84
07030005-4	27,995,975.08	4,701.20	1.46	5,955.07	27,995,975.08
07032002-5	55,017,107.55	12,091.08	1.13	4,550.22	55,017,107.55
07034008-7	10,114,464.36	1,690.50	1.47	5,983.12	10,114,464.36
07037003-6	13,934,809.50	3,050.00	1.48	4,568.79	13,934,809.50
07040005-0	30,393,224.67	6,091.50	1.40	4,989.45	30,393,224.67
07051003-5	16,016,896.62	3,482.80	1.12	4,598.86	16,016,896.62
07056001-2	3,592,371.98	547.50	2.50	6,561.41	3,592,371.98
07057002-4	220,488,332.2	42,525.6	2.69	5,184.84	220,488,332.2
07059023-0	19,790,041.92	4,044.00	1.41	4,893.68	19,790,041.92
07064001-4	36,059,854.82	5,445.80	2.92	6,621.59	36,059,854.82
07066001-9	38,785,892.59	5,885.30	2.57	6,590.30	38,785,892.59
07071002-3	21,442,877.50	4,750.00	1.27	4,514.29	21,442,877.50
07080001-8	90,135,083.05	17,377.49	2.64	5,186.89	90,135,083.05
07082001-2	12,031,399.61	2,472.60	1.23	4,865.89	12,031,399.61
07089022-5	150,876,214.5	31,307.5	1.26	4,819.17	150,876,214.5
07093001-6	57,069,546.96	9,193.20	2.82	6,207.80	57,069,546.96
07102002-1	5,333,704.20	1,210.00	1.21	4,408.02	5,333,704.20
07103001-4	23,398,713.92	5,224.00	1.33	4,479.08	23,398,713.92
07105002-3	15,207,335.78	3,346.00	1.28	4,544.93	15,207,335.78
07112002-8	47,525,695.62	7,118.60	2.62	6,676.27	47,525,695.62
07113002-5	7,324,932.03	1,108.30	2.62	6,609.16	7,324,932.03
07117003-9	26,579,387.13	3,525.30	0.90	7,539.61	26,579,387.13
07119002-9	9,369,065.35	1,499.90	2.65	6,246.46	9,369,065.35
20003002-9	2,262,252.08	451.00	0.88	5,016.08	2,262,252.08
20012001-3	7,313,268.60	1,181.32	0.43	6,190.76	7,313,268.60
20042006-5	6,437,564.09	1,318.90	1.09	4,881.01	6,437,564.09
20071001-4	15,712,213.58	2,347.61	0.72	6,692.86	15,712,213.58
20073003-8	12,590,885.08	1,973.06	0.90	6,381.40	12,590,885.08
20082001-8	2,100,161.28	368.00	0.46	5,706.96	2,100,161.28
20085001-A	10,639,225.53	1,673.70	0.42	6,356.71	10,639,225.53
20088004-5	2,302,861.56	362.94	0.88	6,345.02	2,302,861.56
20113001-6	1,215,451.75	188.55	0.45	6,446.31	1,215,451.75
20117001-5	1,637,488.31	268.18	0.46	6,105.93	1,637,488.31
20128001-9	760,393.17	161.16	1.02	4,718.25	760,393.17
20153001-1	3,547,043.49	620.49	0.43	5,716.52	3,547,043.49
20156001-3	3,174,427.29	703.08	1.08	4,515.03	3,174,427.29

Clave AGEB	Valor de producción	Volumen	Rendimiento	Promedio medio rural	Valor
20159005-3	6,194,493.65	1,190.34	1.21	5,203.97	6,194,493.65
20188001-7	770,082.84	134.66	0.47	5,718.72	770,082.84
20190009-1	21,481,338.57	4,497.00	1.40	4,776.82	21,481,338.57
20201001-0	1,685,152.88	367.50	1.05	4,585.45	1,685,152.88
20202001-8	3,396,378.65	598.29	0.42	5,676.81	3,396,378.65
20207014-1	3,123,943.60	581.40	0.95	5,373.14	3,123,943.60
20211004-0	2,444,415.40	463.40	1.28	5,274.96	2,444,415.40
20222001-0	1,792,497.82	367.20	1.02	4,881.53	1,792,497.82
20231001-A	3,202,756.02	715.34	0.94	4,477.25	3,202,756.02
20253003-6	5,719,147.20	936.00	0.40	6,110.20	5,719,147.20
20266001-5	6,228,043.25	974.40	0.42	6,391.67	6,228,043.25
20272001-2	3,512,739.21	613.98	0.41	5,721.26	3,512,739.21
20275001-4	8,860,568.06	1975.05	0.95	4,486.25	8,860,568.06
20280002-9	3,153,672.79	642.72	1.03	4,906.76	3,153,672.79
20288001-2	938,012.51	195.88	1.18	4,788.71	938,012.51
20306001-2	426,932.90	68.80	0.43	6,205.42	426,932.90
20324001-0	5,461,985.08	868.95	0.45	6,285.73	5,461,985.08
20334004-0	4,672,797.47	760.32	0.36	6,145.83	4,672,797.47
20364003-5	5,986,226.69	1039.92	0.42	5,756.43	5,986,226.69
20366001-0	688,651.23	110.86	0.46	6,211.90	688,651.23
20392001-0	4,984,299.51	787.23	0.84	6,331.44	4,984,299.51
20394001-5	4,915,003.19	1002.98	0.97	4,900.40	4,915,003.19
20413001-2	2,301,749.89	372.75	0.50	6,175.05	2,301,749.89
20424002-0	7,239,616.96	1405.80	1.20	5,149.82	7,239,616.96
20432001-8	1,297,585.72	272.34	1.02	4,764.58	1,297,585.72
20433001-5	1,297,640.26	228.24	0.48	5,685.42	1,297,640.26
20446001-3	3,188,831.19	503.62	0.88	6,331.82	3,188,831.19
20454001-5	3,433,353.18	753.11	0.95	4,558.90	3,433,353.18
20465001-9	1,270,967.59	228.83	0.45	5,554.20	1,270,967.59
20481001-2	5,548,473.00	880.71	0.88	6,300.00	5,548,473.00
20495001-8	14,695,342.28	2758.40	1.27	5,327.49	14,695,342.28
20497001-2	2,224,507.51	388.62	0.51	5,724.12	2,224,507.51
20498014-7	3,259,776.13	689.00	1.39	4,731.17	3,259,776.13
20502004-8	5,935,551.64	1216.92	0.98	4,877.52	5,935,551.64
20526001-6	3,022,581.94	529.98	0.44	5,703.20	3,022,581.94
20541001-2	3,030,108.64	644.80	1.04	4,699.30	3,030,108.64
20554003-A	2,018,479.70	415.39	0.95	4,859.24	2,018,479.70
21025002-0	30,720,207.74	4394.32	4.58	6,990.89	30,720,207.74
21029001-5	87,257,687.25	12298.32	4.41	7,095.09	87,257,687.25
21043002-9	75,369,508.73	10143.03	3.40	7,430.67	75,369,508.73
21049002-2	27,441,406.05	3876.08	3.21	7,079.68	27,441,406.05
21050001-9	21,352,310.69	2983.95	3.49	7,155.72	21,352,310.69
21061001-2	53,860,485.95	7383.56	3.75	7,294.65	53,860,485.95
21068001-3	28,394,752.66	4037.05	3.07	7,033.54	28,394,752.66
21075001-8	45,546,234.48	6569.02	3.47	6,933.49	45,546,234.48
21076003-4	179,513,166.00	23199.00	3.70	7,737.97	179,513,166.00

Clave AGEB	Valor de producción	Volumen	Rendimiento	Precio medio rural	Valor
21077002-7	128,744,039.80	17139.50	2.95	7,511.54	128,744,039.80
21080001-8	21,900,217.84	3131.00	3.10	6,994.64	21,900,217.84
21084002-1	86,284,172.66	12088.80	3.50	7,137.53	86,284,172.66
21086001-1	20,604,190.03	3016.37	3.32	6,830.79	20,604,190.03
21088001-6	3,228,867.68	539.80	1.94	5,981.60	3,228,867.68
21089001-3	88,877,586.40	14859.54	2.89	5,981.18	88,877,586.40
21100001-2	7,359,836.25	1228.15	2.40	5,992.62	7,359,836.25
21109003-7	3,091,016.99	517.28	3.17	5,975.52	3,091,016.99
21111004-A	9,826,114.51	1641.59	3.97	5,985.73	9,826,114.51
21116001-2	10,807,263.05	1,803.41	2.44	5,992.68	10,807,263.05
21174007-3	9,869,531.42	1,628.80	2.46	6,059.39	9,869,531.42
21178001-5	11,592,243.20	1,909.76	2.65	6,070.00	11,592,243.20
21183001-5	2,329,141.97	412.28	1.90	5,649.42	2,329,141.97
21184001-2	13,614,538.10	2,265.71	2.95	6,008.95	13,614,538.10
21186007-4	11,059,202.23	1,842.97	2.70	6,000.75	11,059,202.23
21187001-4	65,397,028.89	10,876.70	2.27	6,012.58	65,397,028.89
21192001-4	4,486,914.67	744.49	3.61	6,026.83	4,486,914.67
21197004-4	33,565,691.38	5,579.96	3.27	6,015.40	33,565,691.38
21202001-2	3,687,500.75	482.57	2.40	7,641.38	3,687,500.75
21213001-6	3,763,730.72	496.79	2.59	7,576.10	3,763,730.72
21216001-8	32,981,107.95	4,661.70	2.46	7,074.91	32,981,107.95
30004003-4	9,350,799.29	754.25	1.75	12,397.48	9,350,799.29
30009002-6	26,403,231.88	2,253.17	1.30	11,718.26	26,403,231.88
30014004-5	9,056,499.85	1,273.97	1.31	7,108.88	9,056,499.85
30024001-8	38,333,018.85	6,754.11	2.11	5,675.51	38,333,018.85
30025005-3	33,664,437.90	5,487.92	1.81	6,134.28	33,664,437.90
30037001-6	7,070,330.00	955.45	1.97	7,400.00	7,070,330.00
30038001-3	8,170,402.24	800.80	1.30	10,202.80	8,170,402.24
30041001-9	3,567,485.19	480.81	1.41	7,419.74	3,567,485.19
30043001-3	31,736,218.06	4,832.10	2.10	6,567.79	31,736,218.06
30046001-5	57,289,923.41	6,088.50	1.65	9,409.53	57,289,923.41
30051003-4	13,437,300.00	1,791.64	1.88	7,500.00	13,437,300.00
30052001-2	7,109,067.33	1,062.76	1.61	6,689.25	7,109,067.33
30057003-8	32,263,183.46	2,601.32	1.91	12,402.62	32,263,183.46
30065003-A	1,233,800.10	117.00	1.50	10,545.30	1,233,800.10
30067001-5	7,566,624.00	1,050.92	1.88	7,200.00	7,566,624.00
30068004-6	3,901,234.87	533.00	1.30	7,319.39	3,901,234.87
30071001-8	73,405,184.50	11,530.89	1.89	6,365.96	73,405,184.50
30079002-0	9,991,146.69	823.50	1.35	12,132.54	9,991,146.69
30080001-7	55,597,303.24	9,051.81	1.81	6,142.12	55,597,303.24
30085001-3	1,505,510.58	197.40	1.40	7,626.70	1,505,510.58
30087028-7	12,329,863.00	1,204.00	1.40	10,240.75	12,329,863.00
30088001-5	1,288,227.73	132.80	1.60	9,700.51	1,288,227.73
30092003-7	47,958,651.55	4,573.00	1.70	10,487.35	47,958,651.55
30093001-5	42,260,999.39	3,552.90	2.00	11,894.79	42,260,999.39
30095001-A	83,861,778.78	13,026.00	3.00	6,438.03	83,861,778.78
30103001-3	12,354,750.00	1,647.30	1.90	7,500.00	12,354,750.00

Clave AGEB	Valor de producción	Volumen	Rendimiento	Precio medio rural	Valor
30110001-8	3,926,253.59	601.65	1.91	6,525.81	3,926,253.59
30112003-1	23,874,230.14	1,902.32	1.95	12,550.06	23,874,230.14
30113001-A	5,140,798.08	708.71	1.31	7,253.74	5,140,798.08
30117001-9	8,564,905.63	1,212.73	1.73	7,062.50	8,564,905.63
30125001-0	5,073,791.53	773.78	1.81	6,557.15	5,073,791.53
30135016-3	73,836.92	9.36	1.51	7,888.56	73,836.92
30137001-1	2,836,515.55	432.81	1.89	6,553.72	2,836,515.55
30140001-7	604,820.13	95.69	1.91	6,320.62	604,820.13
30146001-0	9,877,944.38	1,752.35	2.11	5,636.97	9,877,944.38
30162001-4	28,388,329.61	4,829.82	2.02	5,877.72	28,388,329.61
30164001-9	31,551,695.20	3,004.60	1.66	10,501.13	31,551,695.20
30165001-6	32,132,614.43	5,095.15	1.81	6,306.51	32,132,614.43

Table 2: Datos combinados de indicadores económicos

Año	ICO Composite	OIC MX	OIC Otros Suaves	Indicador Bolsa	SIAP Nacional	SIAP Región TEEB
1976	0,421963428	N/A	N/A	N/A	N/A	N/A
1977	2,584602868	N/A	N/A	N/A	N/A	N/A
1978	0,894119315	N/A	N/A	N/A	N/A	N/A
1979	1,146117472	N/A	N/A	N/A	N/A	N/A
1980	0,691313049	N/A	N/A	N/A	N/A	N/A
1981	0,080384661	N/A	N/A	N/A	N/A	N/A
1982	0,358517848	N/A	N/A	N/A	N/A	N/A
1983	0,170357394	N/A	N/A	N/A	N/A	N/A
1984	0,466391019	N/A	N/A	N/A	N/A	N/A
1985	0,499584271	N/A	N/A	N/A	N/A	N/A
1986	1,644094691	N/A	N/A	N/A	N/A	N/A
1987	-0,263311573	N/A	N/A	N/A	N/A	N/A
1988	0,300973727	N/A	N/A	N/A	N/A	N/A
1989	-0,389712578	N/A	N/A	N/A	N/A	N/A
1990	-0,831812201	-0,44835055	-0,9075548	-0,906533612	N/A	N/A
1991	-0,937234599	-0,2458682	-0,9897523	-0,989465371	N/A	N/A
1992	-1,429702483	-1,0652666	-1,3313423	-1,376870974	N/A	N/A
1993	-1,271657113	-0,45557606	-1,1037873	-1,252542741	N/A	N/A
1994	0,593497846	1,819703	-0,0696342	0,214703156	N/A	N/A
1995	0,619730907	1,96107532	0,0142447	0,2353397	N/A	N/A
1996	-0,061269948	0,82184362	-0,1828015	-0,300377564	N/A	N/A
1997	1,511478529	1,18883094	0,3188445	0,936843365	N/A	N/A
1998	0,245135343	0,19216133	-0,1508011	-0,059355692	N/A	N/A
1999	-0,492017595	-0,40952787	-0,6541284	-0,63926107	N/A	N/A
2000	-0,887866253	-0,89656122	-0,7852484	-0,950675455	N/A	N/A
2001	-1,471130329	-1,1890354	-1,3717766	-1,409491519	N/A	N/A
2002	-1,48901116	-1,44661231	-1,2882774	-1,423573107	N/A	N/A
2003	-1,426016522	-0,89828655	-1,2861079	-1,37401764	-1,85954831	-1,89869474
2004	-1,043323608	-0,12363655	-1,1721813	-1,07293695	-1,74804901	-1,81131346
2005	-0,234196404	1,27314889	-0,4942623	-0,436412177	-1,44666513	-1,52117704
2006	N/A	-0,2765448	-0,2102729	-0,444864214	-1,23022478	-1,21843459
2007	N/A	-0,119908	-0,1084142	-0,275622966	-0,55694852	-0,55052016
2008	N/A	0,31841101	0,2298944	0,024840788	-0,52026277	-0,43905676
2009	N/A	N/A	0,0703808	0,100014564	-0,51916107	-0,43633589
2010	N/A	0,6197018	1,064580175	-0,17538718	-0,00088295	N/A
2011	N/A	1,9327189	2,454847755	0,32608774	0,44791695	N/A
2012	N/A	1,2109487	0,888999898	1,04319184	1,04449029	N/A
2013	N/A	0,3214479	0,020136643	0,28612006	0,25114526	N/A
2014	N/A	1,0526013	1,146663646	0,22632396	0,16782856	N/A
2015	N/A	0,9929125	0,397979722	0,58696712	0,57341226	N/A
2016	N/A	0,8386872	0,469343912	0,68335424	0,79096203	N/A
2017	N/A	1,5670473	0,227674257	1,00668361	0,97548907	N/A
2018	N/A	0,8296295	-0,10581105	1,13245864	0,98906507	N/A
2019	N/A	2,1072835	-0,143999452	0,83067268	0,7171393	N/A
2020	N/A	N/A	0,228028996	0,92867796	0,90151945	N/A
2021	N/A	N/A	1,225045495	1,00570893	1,01744735	N/A
2022	N/A	N/A	2,164933625	N/A	N/A	N/A
2023	N/A	N/A	1,361835856	N/A	N/A	N/A

Table 3: Indicadores económicos, precipitación anual y temperatura media normalizadas

Año	Indicador Bolsa (Normalizado)	Precipitación Anual (Normalizado)	Temperatura media a nivel nacional (Normalizado)
1985	N/A	-0.130698388	N/A
1986	N/A	-0.459473361	N/A
1987	N/A	-1.314870193	N/A
1988	N/A	-0.203436214	N/A
1989	N/A	-1.277046524	N/A
1990	-0.906534	1.430255355	-1.251505942
1991	-0.989465	0.244628793	-1.279783998
1992	-1.376871	0.432292384	0.7845140543
1993	-1.252543	1.164034912	0.9259043319
1994	0.214703	-0.892990803	1.010738498
1995	0.235340	-0.185979136	1.010738498
1996	-0.300378	-1.67564981	0.9259043319
1997	0.936843	-1.272682254	1.010738498
1998	-0.059356	-0.561306317	1.067294609
1999	-0.639261	-0.370733214	0.9824604429
2000	-0.950675	-0.20925524	-1.251505942
2001	-1.409492	-0.196162431	0.9259043319
2002	-1.423573	-0.203436214	0.9824604429
2003	-1.374018	0.240264524	0.9259043319
2004	-1.072937	1.347334234	0.9824604429
2005	-0.436412	0.123884002	1.095572665
2006	-0.444864	0.413380549	1.152128776
2007	-0.275623	0.474480323	1.12385072
2008	0.024841	1.761939841	1.095572665
2009	0.100015	-0.817343464	1.067294609
2010	1.064580	2.655160343	-1.053559554
2011	2.454848	-1.198489672	-0.9687253874
2012	0.889000	-0.542394483	-0.9687253874
2013	0.020137	2.051436388	-0.9970034429
2014	1.146664	0.745065035	-1.025281498
2015	0.397980	1.344424721	-0.9687253874
2016	0.469344	-0.516208865	-0.8838912208
2017	0.227674	0.029324829	-0.7990570543
2018	-0.105811	0.350826019	-0.8838912208
2019	-0.143999	-0.891536047	-0.7990570543
2020	0.228029	-0.830436273	-0.7990570543
2021	1.225045	-0.14670071	-0.8556131653
2022	2.164934	-0.523482648	-0.9121692763
2023	1.361836	N/A	-0.7142228878

Table 4: Valor de producción, volúmen, rendimiento y PMR del café en Chiapas

Municipio	Valor de la Producción	Volumen de Producción	Rendimiento	Promedio Medio Rural
Acacoyagua	11692537.11	2593	0.95	4509.27
Amatenango de la Frontera	65892873.84	13293.38	2.09	4956.82
Ángel Albino Corzo	42347281.04	5686	0.95	7447.64
Bella Vista	64361944.64	13002.68	2.75	4949.9
El Bosque	52528880.26	11899.6	2.17	4414.34
Cacahoatán	35817816.66	8059.68	1.28	4444.07
Chalchihuitán	28129640.04	4318.8	2.44	6513.3
Chenalhó	67911401.84	10325.04	2.59	6577.35
Chicomuselo	27995975.08	4701.2	1.46	5955.07
Escuintla	550171107.55	12091.08	1.13	4550.22
Frontera Comalapa	10114464.36	1690.5	1.47	5983.12
Huehuetán	13934809.5	3050	1.48	4568.79
Huixtla	30393224.67	6091.5	1.4	4989.45
Mapastepec	16016896.62	3482.8	1.12	4598.86
Mitontic	3592371.98	547.5	2.5	6561.41
Motozintla	220488332.2	42525.6	2.69	5184.84
Ocosingo	19790041.92	4044	1.41	4893.68
Oxchuc	36059854.82	5445.8	2.92	6621.59
Pantelhó	38785892.59	5885.3	2.57	6590.3
Villa Comaltitlán	21442877.5	4750	1.27	4514.29
Siltepec	90135083.05	17377.49	2.64	5186.89
Sitalá	12031399.61	2472.6	1.23	4865.89
Tapachula	150876214.5	31307.5	1.26	4819.17
Tenejapa	57069546.96	9193.2	2.82	6207.8
Tuxtla Chico	5333704.2	1210	1.21	4408.02
Tuzantán	23398713.92	5224	1.33	4479.08
Unión Juárez	15207335.78	3346	1.28	4544.93
San Juan Cancuc	47525695.62	7118.6	2.62	6676.27
Aldama	7324932.03	1108.3	2.62	6609.16
Montecristo de Guerrero	26579387.13	3525.3	0.9	7539.61
Santiago el Pinar	9369065.35	1499.9	2.65	6246.46

Table 5: Valor de producción, volúmen, rendimiento y PMR del café en Oaxaca

Municipio	Valor de la Producción	Volumen de Producción	Rendimiento	PMR
Asunción Cacalotepec	2262252.08	451	0.88	5016.08
Candelaria Loxicha	7313268.6	1181.32	0.43	6190.76
Ixtlán de Juárez	6437564.09	1318.9	1.09	4881.01
Pluma Hidalgo	15712213.58	2347.61	0.72	6692.86
Putla Villa de Guerrero	12590885.08	1973.06	0.9	6381.4
San Agustín Chayuco	2100161.28	368	0.46	5706.96
San Agustín Loxicha	10639225.53	1673.7	0.42	6356.71
San Andrés Cabecera Nueva	2302861.56	362.94	0.88	6345.02
San Baltazar Loxicha	1215451.75	188.55	0.45	6446.31
San Bartolomé Loxicha	1637488.31	268.18	0.46	6105.93
San Cristóbal Lachirioag	760393.17	161.16	1.02	4718.25
San Gabriel Mixtepec	3547043.49	620.49	0.43	5716.52
San Ildefonso Villa Alta	3174427.29	703.08	1.08	4515.03
San Jerónimo Coatlán	6194493.65	1190.34	1.21	5203.97
San Juan Colorado	770082.84	134.66	0.47	5718.72
San Juan Cotzocón	21481338.57	4497	1.4	4776.82
San Juan Juquila Vijanos	1685152.88	367.5	1.05	4585.45
San Juan Lachao	3396378.65	598.29	0.42	5676.81
San Juan Mazatlán	3123943.6	581.4	0.95	5373.14
San Juan Ozolotepec	2444415.4	463.4	1.28	5274.96
San Juan Yaeé	1792497.82	367.2	1.02	4881.53
San Lucas Camotlán	3202756.02	715.34	0.94	4477.25
San Mateo Piñas	5719147.2	936	0.4	6110.2
San Miguel del Puerto	6228043.25	974.4	0.42	6391.67
San Miguel Panixtlahuaca	3512739.21	613.98	0.41	5721.26
San Miguel Quetzaltepec	8860568.06	1975.05	0.95	4486.25
Villa Talea de Castro	3153672.79	642.72	1.03	4906.76
San Miguel Yotao	938012.51	195.88	1.18	4788.71
San Pedro el Alto	426932.9	68.8	0.43	6205.42
San Pedro Pochutla	5461985.08	868.95	0.45	6285.73
Villa de Tututepec de Melchor Ocampo	4672797.47	760.32	0.36	6145.83
Santa Catarina Juquila	5986226.69	1039.92	0.42	5756.43
Santa Catarina Loxicha	688651.23	110.86	0.46	6211.9
Santa Lucía Monteverde	4984299.51	787.23	0.84	6331.44
Santa María Alotepec	4915003.19	1002.98	0.97	4900.4
Santa María Huatulco	2301749.89	372.75	0.5	6175.05
Santa María Ozolotepec	7239616.96	1405.8	1.2	5149.82
Santa María Temaxcalapa	1297585.72	272.34	1.02	4764.58
Santa María Temaxcaltepec	1297640.26	228.24	0.48	5685.42
Santa María Yucuhiti	3188831.19	503.62	0.88	6331.82
Santiago Atitlán	3433353.18	753.11	0.95	4558.9
Santiago Ixcuintepéc	1270967.59	228.83	0.45	5554.2
Santiago Nuyoó	5548473	880.71	0.88	6300
Santiago Xanica	14695342.28	2758.4	1.27	5327.49
Santiago Yaitepec	2224507.51	388.62	0.51	5724.12
Santiago Yaveo	3259776.13	689	1.39	4731.17
Santiago Zacatepec	5935551.64	1216.92	0.98	4877.52
Santos Reyes Nopala	3022581.94	529.98	0.44	5703.2
Tanetze de Zaragoza	3030108.64	644.8	1.04	4699.3
Totontepec Villa de Morelos	2018479.7	415.39	0.95	4859.24

Table 6: Valor de producción, volúmen, rendimiento y PMR del café en Puebla

Municipio	Valor de la Producción	Volumen de Producción	Rendimiento	Promedio Medio Rural
Ayotoxco de Guerrero	30720207.74	4394.32	4.58	6990.89
Caxhuacan	87257687.25	12298.32	4.41	7095.09
Cuetzalan del Progreso	75369508.73	10143.03	3.4	7430.67
Chiconcuautla	27441406.05	3876.08	3.21	7079.68
Chichiquila	21352310.69	2983.95	3.49	7155.72
Eloxochitlán	53860485.95	7383.56	3.75	7294.65
Hermenegildo Galeana	28394752.66	4037.05	3.07	7033.54
Hueyapan	45546234.48	6569.02	3.47	6933.49
Hueytamalco	179513166	23199	3.7	7737.97
Hueytlalpan	128744039.8	17139.5	2.95	7511.54
Atlequizayan	21900217.84	3131	3.1	6994.64
Ixtepet	86284172.66	12088.8	3.5	7137.53
Jalpan	20604190.03	3016.37	3.32	6830.79
Jonotla	3228867.68	539.8	1.94	5981.6
Jopala	88877586.4	14859.54	2.89	5981.18
Naupan	7359836.25	1228.15	2.4	5992.62
Pahuatlán	3091016.99	517.28	3.17	5975.52
Pantepec	9826114.51	1641.59	3.97	5985.73
Quimixtlán	10807263.05	1803.41	2.44	5992.68
Teziutlán	9869531.42	1628.8	2.46	6059.39
Tlacuilotepec	11592243.2	1909.76	2.65	6070
Tlaola	2329141.97	412.28	1.9	5649.42
Tlapacoya	13614538.1	2265.71	2.95	6008.95
Tlatlauquitepec	11059202.23	1842.97	2.7	6000.75
Tlaxco	65397028.89	10876.7	2.27	6012.58
Tuzamapan de Galeana	4486914.67	744.49	3.61	6026.83
Xicotepec	33565691.38	5579.96	3.27	6015.4
Xochitlán de Vicente Suárez	3687500.75	482.57	2.4	7641.38
Zihuateutla	3763730.72	496.79	2.59	7576.1
Zoquiapan	32981107.95	4661.7	2.46	7074.91

Table 7: Valor de producción, volúmen, rendimiento y PMR del café en Veracruz

Municipio	Valor de la Producción	Volumen de Producción	Rendimiento	Promedio Medio Rural
Actopan	9350799.29	754.25	1.75	12397.48
Alto Lucero de Gutiérrez Barrios	26403231.88	2253.17	1.3	11718.26
Amatlán de los Reyes	9056499.85	1273.97	1.31	7108.88
Tlaltecpan	38333018.85	6754.11	2.11	5675.51
Ayahualulco	33664437.9	5487.92	1.81	6134.28
Coahuitlán	7070330	955.45	1.97	7400
Coatepec	8170402.24	800.8	1.3	10202.8
Coetzala	3567485.19	480.81	1.41	7419.74
Comapa	31736218.06	4832.1	2.1	6567.79
Cosautlán de Carvajal	57289923.41	6088.5	1.65	9409.53
Coyutla	13437300	1791.64	1.88	7500
Cuichapa	7109067.33	1062.76	1.61	6689.25
Chiconquiaco	32263183.46	2601.32	1.91	12402.62
Emiliano Zapata	1233800.1	117	1.5	10545.3
Filomeno Mata	7566624	1050.92	1.88	7200
Fortín	3901234.87	533	1.3	7319.39
Huatusco	73405184.5	11530.89	1.89	6365.96
Ixhuacán de los Reyes	9991146.69	823.5	1.35	12132.54
Ixhuatlán del Café	55597303.24	9051.81	1.81	6142.12
Ixtaczoquitlán	1505510.58	197.4	1.4	7626.7
Xalapa	12329863	1204	1.4	10240.75
Jalcomulco	1288227.73	132.8	1.6	9700.51
Xico	47958651.55	4573	1.7	10487.35
Jilotepec	42260999.39	3552.9	2	11894.79
Juchique de Ferrer	83861778.78	13026	3	6438.03
Mecatlán	12354750	1647.3	1.9	7500
Mixtla de Altamirano	3926253.59	601.65	1.91	6525.81
Naolinco	23874230.14	1902.32	1.95	12550.06
Naranjal	5140798.08	708.71	1.31	7253.74
Omealca	8564905.63	1212.73	1.73	7062.5
Paso del Macho	5073791.53	773.78	1.81	6557.15
Rafael Delgado	73836.92	9.36	1.51	7888.56
Los Reyes	2836515.55	432.81	1.89	6553.72
San Andrés Tenejapan	604820.13	95.69	1.91	6320.62
Sochiapa	9877944.38	1752.35	2.11	5636.97
Tenampa	28388329.61	4829.82	2.02	5877.72
Teocelo	31551695.2	3004.6	1.66	10501.13
Tepatlaxco	32132614.43	5095.15	1.81	6306.51
Tepetlán	12847244.76	1101.6	1.92	11662.35
Tequila	18797223.65	2889.6	2.1	6505.13
Texhuacán	655561.51	97.41	1.91	6729.92
Tezonapa	138683070.6	21159.42	1.42	6554.2
Tlacotepec de Mejía	9008761.38	1598.1	2.1	5637.17
Tlilapan	220480.4	27.33	1.81	8067.34
Totutla	51695659.37	8281.84	1.84	6242.05
Zentla	32658818.3	5252.1	2.1	6218.24
Zongolica	82311095.33	12728.1	2.1	6466.88
Zozocolco de Hidalgo	8472420	1089	1.98	7780

Table 8: Temperatura media anual en México (1990-2024)

Año	Temperatura media (°C)
1990	20.8
1991	20.7
1992	28.0
1993	28.5
1994	28.8
1995	28.8
1996	28.5
1997	28.8
1998	29.0
1999	28.7
2000	20.8
2001	28.5
2002	28.7
2003	28.5
2004	28.7
2005	29.1
2006	29.3
2007	29.2
2008	29.1
2009	29.0
2010	21.5
2011	21.8
2012	21.8
2013	21.7
2014	21.6
2015	21.8
2016	22.1
2017	22.4
2018	22.1
2019	22.4
2020	22.4
2021	22.2
2022	22.0
2023	22.7
2024	22.9

Bibliografía

- Akaki, P., Velazquez, M. (2020). *Impactos económicos en la cafeticultura ante el cambio climático*. México, Ciudad de México.
- Almeida, I., & Perez, M. G. (2019). *Buy Now, Pay Later Helps JAB Billionaires Build Beverage*. JAB.
- Amazon Web Services Documentation (2025). *Autoregressive Integrated Moving Average (ARIMA) Algorithm* Amazon Web Services. Recuperado de <https://docs.aws.amazon.com/forecast/latest/dg/aws-forecast-recipe-arima.html> (Consulta realizada el 13 de febrero de 2025).
- Ardila, J. C. (2017). *Envisioning coffee growing as a sustainable enterprise: perspectives of a farmer's son*. The craft and science of coffee.
- BBMG, & GlobeScan. (2016). *Five Human Aspirations and the Future of Brands*. Recuperado de http://bbmg.com/BBMG_FiveHumanAspirations.pdf (Consulta realizada el 13 de septiembre de 2024).
- Betancourt, M. (2013). *CAFIVER: Términos para instrumentar un Sistema de Administración Financiera basado en la Utilidad Económica*. SAFUE.
- Bibiloni, H., & Fuentes, M. (2018). *Determinantes del precio del café a nivel internacional*. Universidad Panamericana.
- Bustamante, H (2015). *Halo Effect: study of organic coffee with experimental designs*. Konrad Lorenz.
- Comisión Nacional del Agua (CONAGUA). (2025). *Resúmenes mensuales de temperatura del Servicio Meteorológico Nacional*. Recuperado de <https://smn.conagua.gob.mx/es/climatologia/temperaturas-y-lluvias/resumenes-mensuales-de-temperaturas-y-lluvias> (Consulta realizada el 13 de febrero de 2025).
- Daviron, B., Ponte, S. (2005). *The coffee paradox: global markets, commodity trade, and the elusive promise of development*. Zed Books in association with the CTA.
- Davis, A. P., Chadburn, H., Moat, J., O'Sullivan, R., Hargreaves, S., Nic Lughadha, E. (2019). *High extinction risk for wild coffee species and implications for coffee sector sustainability*. Science Advances.

El Tiempo. (2017, Julio 11). *Las tareas, retos y prioridades para sostener el negocio del café*. El Tiempo. Recuperado de <https://www.eltiempo.com/economia/sectores/retos-para-el-negocio-del-cafe-en-el-mundo-2017-108072> (Consulta realizada el 13 de septiembre de 2024).

Fairtrade International. (2018). *Café*. Recuperado de <https://info.fairtrade.net/es/product/coffee> (Consulta realizada el 20 de noviembre de 2024).

Fairtrade International. (2019). *How coffee farming is supporting the indigenous population of Guatemala*. Recuperado de <https://www.fairtrade.org.uk/media-centre/blog/how-coffee-farming-is-supporting-the-indigenous-population-of-guatemala/> (Consulta realizada el 17 de agosto de 2023).

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189–1232.

Food and Agriculture Organization of the United Nations. (2015). *Statistical PocketBook Coffee 2015*. FAO.

Grabs, J. (2023). *La crisis del precio del café y la volatilidad de los precios ¿Podemos calmar al mercado C?* Revista: Issue 7. Recuperado de <https://sca.coffee/sca-news/25-magazine/issue-7/la-crisis-del-precio-del-caf-y-la-volatilidad-de-los-precios-podemos-calmar-al-mercado-c/> (Consulta realizada el 9 de agosto de 2024).

Hernández, M. Pandolph, R (2020). *Precios volátiles del café: COVID-19 y factores fundamentales del mercado*. Organización Internacional del Café Serie Coffee Break No. 2.

IBM (2024) *¿Qué es GXBoost?* Recuperado de <https://www.ibm.com/mx-es/topics/xgboost> (Consulta realizada el 9 de noviembre de 2024).

INEGI (2022). *Censo Agrícola, Ganadero y Forestal*. Recuperado de <https://www.inegi.org.mx/programas/cagf/2022/> (Consulta realizada el 12 de febrero de 2025).

INEGI (2007). *Censo Ejidal*. Recuperado de <https://www.inegi.org.mx/programas/cae/2007/> (Consulta realizada el 30 de octubre de 2024).

International Coffee Organization (2024). *Daily Coffee Prices*. Recuperado de <https://www.ico.org/coffee-prices.asp> (Consulta realizada el 5 de marzo de 2025).

International Coffee Organization (2024). *Datos Históricos*. Recuperado de <https://www.ico.org/es/new-historical-c.asp> (Consulta realizada el 2 de agosto de 2024).

Kim, J., Lee, S. (2024). Adaptive stacking ensemble techniques for early severity classification of COVID-19 patients. *Applied Sciences*, 14(7),

Kim, W. C., & Mauborgne, R. (2005). *La estrategia del océano azul: cómo desarrollar un nuevo mercado donde la competencia no tiene ninguna importancia*. Grupo Editorial Norma.

Landini, F. (2016). *Problemas de la extensión rural en América Latina*. Perfiles Latí-

noamericanos. FLACSO.

Lamb, H., & Byers, L. (2017). *Perspectives on Price: Fairtrade and Beyond*. The craft and science of coffee.

Moustier, P., Renting, H. (2015). *Urban agriculture and short chain food marketing in developing countries*. En: Zeeuw, H., Drechsel, P. (Eds), Routledge, pp. 121-138.

Muñoz-Rodríguez, M., Gómez-Pérez, D., Santoyo-Cortés, V.H., Rosales-Lechuga, R. (2019). *Los negocios del café ¿Cómo innovar en el contexto de la paradoja del café, en pro de una red de valor más inclusiva y accesible?* Universidad Autónoma Chapingo, CIESTAAM.

Nestlé, S. A. (2003). *Hoy en día, los agricultores sufren por los bajos precios del café ¿Qué se puede hacer?* Public Affairs.

Neter, J., Li, W. (2005). Applied linear statistical models (5th ed.). McGraw-Hill/Irwin.

Nielsen, D. (2016). Tree boosting with XGBoost: Why does XGBoost win "every" machine learning competition? arXiv preprint arXiv:1702.01884.

Organización de las Naciones Unidas para la Alimentación y la Agricultura (2014). *2014, Año Internacional de la Agricultura Familiar*. International Fund for Agricultural Development.

Perfect Daily Grind. (2018). *Un Mercado Volátil: La Economía de los Precios del Café en el Mundo*. Recuperado de <https://perfectdailygrind.com/es/2018/01/29/> (Consulta realizada el 10 de diciembre de 2024).

PROMECAFE. (2018). *El estado actual de la Rentabilidad del café en Centroamérica*. Recuperado de [https://www.scanews.coffee/wp-content/uploads/2010/06/2017-Estudio-de-Costos-POMECAFE-CABI-1](https://www.scanews.coffee/wp-content/uploads/2010/06/2017-Estudio-de-Costos-PROMECAFE-CABI-1) (Consulta realizada el 22 de septiembre de 2024).

Robles, H. (2018). *La organización económica de los pequeños y medianos productores presente y futuro del campo mexicano*. Centro Latinoamericano para el Desarrollo Rural.

Robles, H (2022). *Pequeña Agricultura. Subsidios al Campo*. Recuperado de <https://subsidiosalcampo.org.mx/> (Consulta realizada el 12 de septiembre de 2024).

Sanders, D. (2017). *The value proposition: reflections on the nature of value in coffee*. En: B. Folmer (Ed.), The Craft and Science of Coffee.

Scikit-learn developers. (2024). *sklearn.preprocessing.OneHotEncoder – Scikit-learn documentation* Recuperado de <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html> (Consulta realizada el 2 de mayo de 2025)

Servicio de Información Agroalimentaria y Pesquera (2022). *Datos de producción Agrícola*. <https://www.gob.mx/siap/acciones-y-programas/produccion-agricola-33119> (Consulta realizada el 15 de agosto de 2024).

Tark, S. (2018). *Esto es lo que cuesta producir café en Latinoamérica*. Perfect Daily Grind Ltd. Recuperado de <https://www.perfectdailygrind.com/2018/08/esto-es-lo-que-cuesta-producir-cafe-en-latino-america/> (Consulta realizada el 13 de septiembre de 2024).

TEEB Agrifood Café México. (2023). *Información socioeconómica, organizacion, producción y costos. Base agregada Información general*. Recuperado de <http://teebafcafe.geolab.mx/> Consulta realizada el 5 de marzo de 2024).

The Nielsen Company (2019). *Los compradores sostenibles compran el cambio que quieren ver en el mundo* Recuperado de <https://www.nielsen.com/co/es/insights/article/2019/los-compradores-sostenibles-compran-el-cambio-que-quieren-ver-en-el-mundo/> (Consulta realizada el 9 de noviembre de 2024).

True Price (2017). *Assessing coffee farmer household*. Recuperado de http://fairtradeamerica.org/~/media/FairtradeAmerica/Files/Reports/1706_ExecSummary-AssessingCoffeeFarmerIncome_final.pdf (Consulta realizada el 13 de septiembre de 2023).

Unión Majomut (2023). *Seguridad y Soberanía Alimentaria*. Unión Majomut. Recuperado de <http://www.union.majomut.org/> (Consulta realizada el 30 de julio de 2024).

United Nations Industrial Development Organization. (2020). *Short food supply chains for promoting local food on local markets*.

Villarreal, F. (2017). *Inclusión financiera de pequeños productores rurales*. Libros de la CEPAL, N° 147. Comisión Económica para América Latina y el Caribe.

Zhou, Z. H. (2012). Ensemble methods: Foundations and algorithms. CRC Press.