



SYSTEMATIC REVIEW OF METHODS FOR SPECIALTY COFFEE PRICE ESTIMATION

Una revisión sistemática de métodos para la estimación del precio del café especial

Victor-Hugo Pinto-Rodriguez 

Universidad del Cauca (Popayán, Colombia). Centro de Desarrollo Tecnológico Creativo (Popayán, Colombia).
victorpinto@unicauca.edu.co

Carlos-Alberto Cobos-Lozada 

Universidad del Cauca (Popayán, Colombia). 
ccobos@unicauca.edu.co

Adriana-Marcela Nieto-Muñoz 

Corporación Universitaria Comfacaucá (Popayán, Colombia). 
anieto@unicomfacaucac.edu.co

Received: 30-08-2024

Accepted: 12-02-2025



ABSTRACT

Despite being the initial and fundamental actor in specialty coffee's global production and marketing chain, coffee growers receive a disproportionately small share of the product's final value. This situation, aggravated by the adverse effects of climate change, market volatility, and increased production costs, exposes them to a precarious economic situation. There are many ways to help these producers, such as estimating production costs and selling prices, estimating climate, improving infrastructure, supporting the creation of cooperatives, and promoting fair trade. This review aims to identify several methods employed across different countries for estimating the price of specialty coffee. For this purpose, a systematic review methodology is conducted. This methodology involves identifying the need for the review, selecting and analyzing primary studies, and disseminating findings. Several studies employ mathematical models based on regression analysis, time series models, and artificial neural network-based models for coffee price estimation. In addition, it was identified that the most used evaluation metrics for the models mentioned above were R², Akaike Information Criterion (AIC), and Average Mean Squared Error (MSE), respectively. Findings also reveal that these models often used data from online auctions, supermarkets, and the stock market, emphasizing the Cup of Excellence (CoE) auction contest. It was also identified that the main analyzed variables for the estimation were date or year, price of coffee, country of origin and destination, variety, ranking in the auction, and altitude. The study emphasizes the need for new methods and variables to estimate specialty coffee prices and their potential positive impact on the industry. These methods must adjust parameters flexibly as the models are prone to change over time.

Keywords: artificial intelligence; coffee price; sustainability.

RESUMEN

A pesar de ser el actor inicial y fundamental en la cadena global de producción y comercialización del café de especialidad, los caficultores reciben una parte desproporcionadamente pequeña del valor final del producto. Esta situación, agravada por los efectos adversos del cambio climático, la volatilidad del mercado y el aumento de los costos de producción los expone a una situación económica precaria. Hay muchas maneras de ayudar a estos productores, como estimar los costos de producción y los precios de venta, estimar el clima, mejorar la infraestructura, apoyar la creación de cooperativas y promover el comercio justo. Esta revisión tiene como objetivo identificar varios métodos empleados en diferentes países para estimar el precio del café de especialidad. Para ello, se emplea una metodología de revisión sistemática. Esta implica identificar la necesidad de la revisión, seleccionar y analizar estudios primarios y difundir los hallazgos. Varios estudios emplean modelos matemáticos basados en análisis de regresión, modelos de series de tiempo y modelos basados en redes neuronales artificiales para la estimación del precio del café. Además, se identificó que las métricas de evaluación más utilizadas para los modelos mencionados anteriormente fueron R^2 , Akaike Information Criterion (AIC) y Average Mean Squared Error (MSE), respectivamente. Los hallazgos también revelan que estos modelos a menudo usaban datos de subastas en línea, supermercados y el mercado de valores, enfatizando el concurso de subastas Cup of Excellence (COE). También se identificó que las principales variables analizadas para la estimación fueron fecha o año, precio del café, país de origen y destino, variedad, ranking en la subasta y altitud. El estudio enfatiza la necesidad de nuevos métodos y variables para estimar los precios del café de especialidad y su posible impacto positivo en la industria. Estos métodos deben ajustar los parámetros de manera flexible, ya que los modelos son propensos a cambiar con el tiempo.

Palabras clave: inteligencia artificial; precio del café; sostenibilidad.

REVISÃO SISTEMÁTICA DOS MÉTODOS PARA ESTIMAR O PREÇO DO CAFÉ ESPECIAL

RESUMO

Apesar de ser o ator inicial e fundamental na cadeia global de produção e comercialização do café especial, os cafeicultores recebem uma parte desproporcionalmente pequena do valor final do produto. Essa situação, agravada pelos efeitos adversos das mudanças climáticas, a volatilidade do mercado e o aumento dos custos de produção, expõe os produtores a uma situação econômica precária. Existem várias maneiras de ajudar esses produtores, como estimar os custos de produção e os preços de venda, estimar o clima, melhorar a infraestrutura, apoiar a criação de cooperativas e promover o comércio justo. Esta revisão tem como objetivo identificar diversos métodos empregados em diferentes países para estimar o preço do café especial. Para isso, é utilizada uma metodologia de revisão sistemática. Isso envolve identificar a necessidade da revisão, selecionar e analisar estudos primários e divulgar os resultados encontrados. Vários estudos utilizam modelos matemáticos baseados em análise de regressão, modelos de séries temporais e modelos baseados em redes neurais artificiais para estimar o preço do café. Além disso, identificou-se que as métricas de avaliação mais utilizadas para os modelos mencionados foram R^2 , Akaike Information Criterion (AIC) e o Erro Quadrático Médio (MSE), respectivamente. Os resultados também revelaram que esses modelos frequentemente utilizavam dados de leilões online, supermercados e do mercado de ações, com ênfase no concurso de leilões *Cup of Excellence* (COE). Também foi identificado que as principais variáveis analisadas para a estimativa foram data ou ano, preço do café, país de origem e destino, variedade, classificação no leilão e altitude. O estudo enfatiza a necessidade de novos métodos e variáveis para estimar os preços do café especial e seu possível impacto positivo na indústria. Esses métodos devem ajustar os parâmetros de maneira flexível, pois os modelos tendem a mudar ao longo do tempo.

Palavras-chave: inteligência artificial; preço do café; sustentabilidade.

1. INTRODUCTION

In several countries, most of the population dedicated to coffee production is highly impoverished. Factors such as ineffective business models, unreliable supply chains, and sudden climate changes contribute to the challenging economic conditions endured by this population sector [1].

In Colombia, coffee production supports over 541,000 families and generates more than 700,000 direct jobs; it is the primary agricultural export, representing 22% of the agricultural Gross Domestic Product (GDP) [2]. However, despite being the most invested actors along the value chain, coffee growers receive only 6% of the final product price due to the prevailing business model. For instance, a pound of roasted and ground 100% Colombian coffee costs approximately US\$15 in Europe. However, the breakdown of this value illustrates the disparities: local buyers retain USD\$0.13, transporters receive USD\$1.5, roasters earn USD\$1.2, maritime transport and customs expenses amount to USD\$0.3, exporters obtain USD\$3.5 and market intermediaries in the destination country keep USD\$7.45 [3]. Moreover, specialty coffee production incurs higher costs than traditional coffee. Shade-grown specialty coffee, for example, undergoes slower production cycles, yields lower quantities, and is often subject to fungal diseases [4].

The National Federation of Coffee Growers [*Federación Nacional de Cafeteros*] regulates the internal sales price of coffee, encompassing both commodity and specialty in Colombia. Specialty coffee refers to coffee that scores 80 points or higher in sensory evaluations [5]. This regulation estimates a base market price as a safeguard for coffee growers. Additionally, any applicable bonuses for specialty coffee programs are incorporated [6]. However, this pricing mechanism overlooks factors that could enhance the accuracy of estimations, such as sensory attributes, tree variety, lot size, and destination country. Considering that the Colombian coffee base price is determined based on the New York Stock Exchange, coffee growers are at constant economic vulnerability, as this price is subject to fluctuations due to the implicit market volatility [7]. This volatility is a consequence of the speculative nature of the business model implemented by investors within the New York Stock Exchange [3].

So far, the mathematical regression models used to estimate the price of specialty coffee are based on the analysis of outdated data sources, such as those proposed by Donnet et al. (2008) [8], Teuber & Hermann (2012) [9], Wilson & Wilson (2014) [10] and Traore et al. (2018) [11]. This reliance on outdated data results in inappropriate correlation coefficients to estimate coffee prices accurately today. Over time, the availability of information and variables that can contribute to a more accurate estimation of specialty coffee prices has increased significantly. These include data from various stock exchanges, geographic information systems providing price and supply data, supermarket sales, auction sites, climatic data sources, and national production data. By incorporating data from current sources into estimation models, it is possible to minimize the errors associated with changes in independent variables over time.

This systematic review summarizes the current methodologies for estimating the price of specialty coffees. It was based on the guidelines proposed by Kitchenham et al. [12]; it also discusses the main characteristics of the variables and models employed by different authors. Hence, potential future research is proposed to enhance the prediction and estimation of specialty coffee prices.

2. METHODOLOGY

The systematic review was conducted using the guidelines proposed by Kitchenham et al. [12], which outlines three main stages. In the first stage, planning, the need for the review must be clearly defined. A detailed review protocol is developed to reduce research bias and ensure a structured approach. In the second stage execution, documents are systematically searched and analyzed following the predefined review protocol. The review is compiled and formatted in the last stage, dissemination, to present the findings.

2.1 Review Protocol

As Kitchenham emphasizes, the precise formulation of research questions is essential for a systematic review since the process is fundamentally based on them. In this sense, the PICOC (Population, intervention, comparison, outcome and context) [13] method was employed to define the research questions and search queries. Table 1 shows the research questions and their justification.

Table 1. Research questions and their justification.

Research questions	Justification
What variables are considered in models for calculating specialty coffee prices?	Identify the factors influencing specialty coffee prices and explore how these factors can be measured and incorporated into pricing models.
What data sources are available to estimate specialty coffee prices?	A critical factor in specialty coffee price estimation is identifying publicly available data sources that provide relevant information.
What models, techniques, or methods have been used to estimate specialty coffee prices?	Investigate and review the different models and techniques that have been applied to estimate the prices of specialty coffee, as well as their results and benefits.
What metrics are used to evaluate the quality of specialty coffee price predictions?	It is essential to compare different techniques that researchers use to evaluate the metrics and assess the accuracy and effectiveness of specialty coffee price estimation models.

1) Search terms. A comprehensive set of search terms was used to gather relevant literature for this research, focusing on specialty coffee price estimation. To identify literature related to specialty coffee, we employed a variety of synonymous terms to cover the various aspects of the study. This search string included: ("specialty coffee" OR "premium coffee" OR "gourmet coffee" OR "high-quality coffee" OR "single-origin coffee" OR "certified coffee"). We also needed to capture different methodologies for estimating or calculating the price (intervention). Therefore, we included different terms related to various modeling and analytical techniques: (model OR technique OR algorithm OR function OR regression) and the word price with an AND operator. Finally, the resulting search string was constructed as follows to ensure comprehensive coverage of the topic: ("specialty coffee" OR "premium coffee" OR "gourmet coffee" OR "high-quality coffee" OR "single-origin coffee" OR "certified coffee") AND (model OR technique OR algorithm OR function OR regression) AND (price*). Moreover, the time period from 2008 to 2023 was considered.

To broaden the scope and include studies on general coffee price forecasting models, we used an additional search string focused specifically on price prediction and estimation for coffee. This string

targeted documents that may provide insights into broader pricing models applicable to coffee: TITLE-ABS-KEY ((pric* AND predict*) OR (pric* AND estima*) OR (pric* AND forecast*) OR “Consumer Price Index”) AND TITLE (“coffee”). Here, we searched for titles containing the word to focus on the product. We then explored various combinations involving the word “price” to find documents discussing prediction, estimation, or forecasting methods.

Applying the initial search string focused on specialty coffee pricing and related methodologies, we conducted searches on Scopus and Web of Science. This search yielded a total of 42 articles. After applying our defined inclusion and exclusion criteria based on language, research topic, and year (as detailed in [Table 2](#)), we narrowed the selection to six articles most relevant to our research questions.

The second search string, designed to capture broader coffee price forecasting and estimation models, was also applied to both databases. This search resulted in 184 articles, 23 of which were selected. This set includes all relevant articles from the first search, except for one study by Traore et al. (2018) [11], which was not duplicated in the second search results. Therefore, 18 new articles were found with this search string. The total number of studies found with both search strings was 24. Finally, the same inclusion and exclusion criteria of the previous search were used ([Table 2](#)).

2) Quality assessment. After selecting articles for the systematic review, a comprehensive quality assessment was performed to ensure that only high-quality studies were included; relevance to the research questions and the journal’s quartile in which the study was published were assessed, as shown in [Table 3](#).

Table 2. Inclusion and exclusion criteria.

Type of criterion	Criteria	Type
Language	Documents must be published in either English or Spanish.	Inclusion
Alignment with the research topic	The abstract suggests that the paper contributes to solving at least one research question.	Inclusion
Year	Documents must be published between 2008 and 2023.	Inclusion

Table 3. Criteria used to evaluate the quality of the selected studies.

Criterion	Description	Score		
		-1	0	1
C1	The study presents a detailed description of the variables of the models.	No	Partially	Yes
C2	The study presents how the data were obtained in detail.	No	Partially	Yes
C3	The study presents a detailed description of the models.	No	Partially	Yes
C4	The study presents clearly and in detail the results obtained.	No	Partially	Yes
C5	The study has been published in a relevant journal.	No	Q4, Q3	Q2, Q1

3. RESULTS AND DISCUSSION

[Table 4](#) presents the results of the quality criteria evaluation for the studies selected in the systematic review. It is noted that all the studies (100%) offer a thorough description of the variables used in the models; 96% of the selected studies describe the data acquisition process; moreover, 92% of the evaluated studies describe the used model; and finally, 83% of the studies present the obtained results,

which facilitates the evaluation of the model's effectiveness. Regarding the quality of the publication, most articles were published in journals categorized as Q1 and Q2, reflecting that the reviewed studies have been published in high-quality journals.

Table 4. Quality assessment of selected studies

Reference	C1	C2	C3	C4	C5	Score	Reference	C1	C2	C3	C4	C5	Score
[8]	1	1	1	0	1	4	[14]	1	1	1	1	0	4
[9]	1	1	1	1	1	5	[15]	1	1	1	1	-1	3
[10]	1	1	1	1	1	5	[16]	1	1	1	1	1	5
[11]	1	1	1	1	-1	3	[17]	1	1	1	1	0	4
[18]	1	1	1	1	1	5	[19]	1	1	1	1	1	5
[20]	1	1	1	1	1	5	[21]	1	1	1	1	1	5
[22]	1	1	1	1	0	4	[23]	1	1	1	1	1	5
[24]	1	1	1	-1	0	2	[25]	1	1	1	1	-1	3
[26]	1	1	1	1	-1	3	[27]	1	1	1	0	0	3
[28]	1	1	1	1	0	4	[29]	1	1	1	1	1	5
[30]	1	1	1	1	0	4	[31]	1	1	0	1	1	4
[32]	1	1	0	-1	1	2	[33]	1	0	1	1	0	3

3.1 Article Summary

For estimating the price of specialty coffee, Donnet et al. (2008) [8] leveraged a dataset comprising 541 entries from 21 coffee auctions sourced from CoE [34]. Their approach focused primarily on evaluating sensory attributes and the reputation of the coffee. They employed the Chow test to analyze these variables [35] one may wish to test whether m additional observations belong to the same regression. This paper presents systematically the tests involved, relates the prediction interval (for $m = 1$; a summary of variables is shown in Table 5. They created a linear regression model to estimate the hedonic price of specialty coffee presented in Equation (1). Concerning metrics, they focused on the variable's significance and the precision of the results using the R^2 coefficient; the authors obtained an R^2 of 0.67.

Table 5. Variables identified by Donnet et al.

Variable	Description
Quality	This variable represents the coffee's sensory attributes, including aroma, flavor, body, and aftertaste. Higher quality scores typically correlate with higher auction processes.
E-auction ranking	The auction ranking indicates a coffee lot's position within the auction results. Donnet et al. observed that coffee prices significantly increase for the coffee lot that secures the first four places.
Country of origin	The origin often influences the flavor profile and market perception, thus impacting the price. The study used data from auctions held in several countries known for producing high-quality coffee: Bolivia, Colombia, El Salvador, Honduras, and Nicaragua.
Tree Variety	The research included coffee tree varieties: Catuai, Caturra, Typica, and Pacamara. Additionally, they introduced a category label called 'others' to account for other varieties not explicitly named.
Lot size	The size of the coffee lot is an essential factor in pricing. Due to their limited availability, smaller lot sizes are often seen as more exclusive and can command higher processes.
Year of auction	The study considered 2004, 2005, and 2006 auctions. The year of the auction can reflect market trends or consumer preferences over time, which can influence the price dynamics for specialty coffee.

$$\ln(P_i) = \beta_0 + \sum_j f(Z_{ij})\beta_{ij} + \varepsilon_i \quad (1)$$

Where:

$i=1, \dots, n$ are the observed specialty coffees.

$j=1, \dots, m$ are the attributes considered in the model.

ε_i is the error term, assumed to be an independently distributed error term with mean 0 and variance σ^2 .

$\ln(P_i)$ is the natural logarithm of the individual price function.

$f(Z_{ij})$ is the function applied to the attribute j for coffee i . This function can be an identity function, a logarithm function, or a dummy variable (1 if the attribute j is present for the i -th observation, or 0 otherwise).

β_{ij} are the implicit prices or coefficients of the attributes.

In 2012, Teuber and Hermman [9] used CoE to estimate the hedonic price of specialty coffee and obtained 1260 observations from 46 electronic auctions; this approach was aligned with Donnet et al.'s methodology by incorporating similar variables. However, they expanded the scope by considering auction data from 2003 to 2009, adding bourbon variety. Additionally, they included Brazil, Costa Rica, and Guatemala as new countries of origin in their analysis. They compared five models derived from a base equation (see Equation (2)), adding or removing variables. Among the models, they highlighted variety and origin effects: Model 1 includes variety and origin effects; interaction effects; Model 2 model incorporates interactions between cup score (SQS) and country of origin (CO), and Model 3, only country of origin. Model 2 presented 356 points in the AIC. However, this model does not differ significantly considering Model 1, which presents 368 points in the AIC. According to the Bayesian Information Criterion (BIC), the best fitting equation is Model 3, with 471 points, followed by the so-called Model 1, which includes both variety and country of origin. Regarding the adjusted R^2 , no significant differences were observed between the models in the same year.

$$\log(P_i) = \beta_0 + \beta_1 \log(q_i) + \beta_2 \text{SQS}_i + \beta_3 \text{Organic}_i + \beta_4 \text{ranking}_i + \beta_5 \text{variety}_i + \beta_6 \text{CO}_i + \beta_7 \text{CO}_i \text{SQS}_i + \beta_8 \text{auctionyear}_i + u_i \quad (2)$$

Where:

i : represents an individual coffee auctioned.

$\log(P_i)$: Natural logarithm of the price of coffee auctioned in US dollars per pound.

q_i : quantity sold expressed in pounds.

SQS_i : sensory quality score assigned to the coffee.

organic_i : Presence of organic certification.

ranking_i : Position held by the coffee at auction.

$variety_i$: Variety of specialty coffee.

CO_i : Country of origin of the coffee.

$auctionyear_i$: Year of the coffee auction.

u_i : stochastic error term.

Schollenberg [31] used sales in a representative sample of supermarkets (101 – 400 m²) and megastores all over Sweden collected and provided by Nielsen from March 2005 to March 2008 to investigate the impact of the Fair-Trade label on the coffee market in Sweden. The variables were Fair Trade characteristics, Organic characteristics, Intrinsic characteristics (roast quality/ degree: mellan, mörk, and extra mörk), coffee characteristics (brygg, and kok), production region label, decaffeinated region label, flavoring, brands (41 dummies), and week (157 dummies). They found that the Fair-Trade label is essential to Swedish consumers due to their high sustainability awareness. The author used Rosen's method [36] to create a regression model for the supermarket sales dataset. She got 0.64 of R².

In 2014, Wilson and Wilson [10] expanded upon the research conducted by Teuber and Donnet; they used CoE data from 2005 to 2010. The examined variables included sensory test scores, auction ranking, lot size, the altitude at which the coffee was grown, certifications, country of origin, and varieties. An additional innovation included the market in which the coffee was sold. This variable categorizes sales into regions such as Asia, Nordic countries, Europe, cooperative buyer groups, and other markets. Regarding models, the authors start with Donnet's model and create six others combining different variables; Model 1 reproduces Donnet; the favorite model of the authors was Model 5 (See equation (3)). They compared their model against Donnet's model, and with a variation of Teuber & Herrmann using the AIC method, the author mentions that Model 5 proposed by him improves Donnet's model. However, he does not mention the values found.

$$\ln(P_i) = \beta_0 + \beta_1 \text{Quality } y_i + \sum j \beta_j \text{Reputation}_{ij} + \sum k \beta_k \text{MarcoCorrection}_{ik} + \sum m \beta_m \text{BuyerLocation}_{im} + \sum r \beta_r \text{CountryOfOrigin}_{ir} \text{TreeVariety}_{ir} + \varepsilon_i \quad (3)$$

In 2018, Traore et al. [11] got their data from CoE, analyzing 1260 observations from 2005 to 2010. They included several variables used in previous research and added the sensory profile and processing type to their analysis. They divided the dataset attributes into two categories to evaluate the influence between symbolic attributes such as country of origin, certification, lot size, and crop height and material attributes such as sensory profile. Hence, the model proposed by the authors incorporates symbolic and material variables mentioned and an error term (See equation 4); it is essential to consider that this equation is truncated because the minimum score to participate in the CoE is 84. They compared their model with that of Wilson and Wilson (2014) using AIC and BIC, finding better results by including more attributes, as can be seen below: Traore obtained 9789 and 9661 from BIC and AIC, respectively; while Wilson and Wilson obtained 8576 and 8264 from BIC and AIC, respectively.

$$\ln(\text{price}_i) = M_i' \sigma + R_i' \zeta + \psi_i \quad (4)$$

Where:

M_i' : vector of material attributes.

R_i' : vector of symbolic attributes.

ψ_i : residual value.

Novanda et al. [17] took secondary data from Index Mundi for World for international coffee prices and the Central Bureau of Statistics and Ministry of Agriculture of Indonesia for domestic price data. The data collected were monthly from 2008 to 2016, with 96 observations. The variables used were coffee prices in global and domestic markets and the corresponding dates. Regarding models, Novanda et al. tried different models to predict the volatility of coffee commodity prices in Indonesia, and the best model was the Autoregressive Integrated Moving Average (ARIMA). They compared the models they used with Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), and Mean Square Deviation (MSD); the model selected for the two datasets was ARIMA since it had a MAPE of 3.76, MAD of 0.074, and an MSD of 0.010 for World Price; and had a MAPE of 0.9, MAD of 141.2, MSD of 43455 in Domestic Price.

Also in 2018, Crespo et al. [29] analyzed the variables that determine the price of Arabica coffee from a data set of 363 records made up of monthly values from 1985 to 2016; among the variables contemplated in the study are coffee production in Brazil and the world; macroeconomic variables; different stock markets in the world; and climatic variables such as temperature and precipitation, in coffee-growing areas in Brazil. They evaluated the historical explanation of coffee prices using the Bayesian Model Averaging (BMA) method and the Auto-Regressive (AR) model with different lags. From the above, they found that variables related to macroeconomics and coffee prices in the stock market explain the historical behavior of coffee prices.

In 2019, Figueroa-Henandez et al. [27] compiled information from various sources to get the futures of coffee. They included total coffee world production and total coffee exports, the price of International Coffee Organization (ICO), the price of different coffee types, and monthly average future prices from London and New York stock exchanges measured in dollar cents by the pound. They used the Principal Components Analysis Technique to estimate coffee production and exportation. They didn't report an evaluation for their model.

In 2020, Rodríguez and Melgarejo [23] got 5478 samples released by the Colombian National Federation of Coffee Growers (FNC) for its acronym in Spanish; the data analyzed was the internal Colombian coffee price from 2003 to 2017, they analyzed the price variable as a signal, identifying its properties. Moreover, they selected a nonlinear dynamical system from a set of candidates by comparing the phase plane of the possible solution with that. A qualitative analysis based on phase diagrams selected the chaotic multi-scroll system; this model was calibrated using the Artificial Bee Colony optimization algorithm. They used the Nash Sutcliffe Efficiency, MSE, and Predictions of Change in Direction to evaluate models like Multiscroll Chua. Finally, they discovered that the Multiscroll Chua System, a three-state variable model, could be the best.

Holmes and Otero [32] extracted spot and future prices of Arabica and robusta coffees from the ICO; the future prices of other mild varieties from the settlement price traded on the New York Coffee, Sugar, and Cocoa Exchange for delivery standard contracts of 37500 pounds. Robusta's futures prices were obtained from the London International Financial Futures and Options Exchange for standard contracts of approximately 22046 pounds. This spans the period between 1993 and 2017. They tested

the co-integration between spot and futures prices to find the validity of futures market efficiency with the Dickey-Fuller test (DF). They found that only at the end of the study period could the other mild varieties (Robusta) futures price be regarded as an unbiased predictor of the Robusta spot price. They did not build a model. Therefore, they did not need a test to evaluate any model.

In 2020, Acevedo et al. [22] got their data from the FNC database on the price of coffee, and they gathered 3619 observations from 2010 to 2019. They applied the DF test and determined that the data wasn't stationary; they finally found that the ARIMA (1,1,0) model was optimal; however, the results of ARIMA presented some volatility clusters; then found a Generalized Autoregressive Conditional Heteroscedasticity model (GARCH) (1,2). They evaluated the fit of the ARIMA model with the Chi2 test; they got a Chi2 of 39.159 and a p-value of 0.39. Regarding the GARCH, they applied the Ljung-Box test and were found to have a 0.34 statistical and p-value of 0.55; since p is more significant than this significance level, it can be said that GARCH represents an excellent fit to the residual series of the ARIMA model.

In 2021, Ababu and Getahun [28] employed monthly secondary data collected over a decade (2000 – 2010) from the Ethiopian Commodity Exchange (ECX). The coffee price data were not stationary; for this reason, the Augmented Dickey-Fuller Test (ADF) found stationarity in the second difference. After that, they created some ARIMA models; the preferred was ARIMA (1,2,1), selected with the AIC test. They used MAPE, MSD, and MAD to measure their models' accuracy; the best model had a MAPE of 11.74, a MAD of 2.55, and an MSD of 9.23.

Wang et al. [21] collected historical coffee bean price data from the Giacaphe website [37] from 2014 to 2019, they aimed at predicting the variable price coffee beans from Lam Dong province. Then, they built a forecasting nonlinear grey Bernoulli model (NGBM) model; it was integrated with the Fourier residual modification model to forecast coffee bean prices. Finally, they evaluated their models with the mean absolute percentage error, calculated to find the forecasting error; the best model succeeded in a MAPE of 2.077.

Herrera et al. [26] extracted 6347 registers between 2010 and 2020 from the FNC database; the data was used to train different Long-Short Term Memory (LSTM) model configurations to forecast the price of coffee; the best model was one with one hidden layer, 20 LSTM units in each layer, and 1781 hyperparameters. Finally, they used MSE and the Mean Absolute Error (MAE) to compare the model's performance; the best model achieved an MSE of 0.000125 and a MAE of 0.007260 when trained with 1000 epochs and 0.0001 of learning rate.

Nugroho et al. [30] used daily data on the spot price of Robusta Coffee, futures prices of Robusta Coffee, exchange rates, and daily case-positive COVID-19 in Indonesia since 2020, with 213 observations, created a Vector Error Correction Model (VECM) to forecast the price of coffee. ADF test found that the data was not stationary, then compared the optimal lag in the VECM model to forecast the spot price of Robusta coffee; the best lag for the VECM model was two with AIC of 13.91 and Schwarz Information Criterion of 14.01. Finally, they found that the futures price of Robusta Coffee significantly influenced the spot price of Robusta Coffee, while the COVID-19 pandemic did not affect it.

Gálvez and Cortés [16] used the price of coffee in the Mexican market concerning the production of coffee and the international price of coffee to get data over 16 years from 2004 to 2019 (192 observations); they used a single equation conditional Error Correction Model (ECM) and found that all series were non-stationary. Finally, they evaluated their model using a CUSUM test. CUSUM values are between 95 percent confidence intervals. This means that the fitted ECM is parsimonious and stable. They concluded that a 1% increase in the international price of coffee implied an increment of 0.9% in

the Mexican coffee price and not a long-run relationship between Mexican price and production.

In 2022, Benavidez and Xia [24] collected data on quantity, organic, and conventional coffee prices from the Global Agricultural Trade System. The distance between the exporting countries and the United States [38] and other variables to analyze the effects of economic and regulatory factors on the trade volume of organic Arabica green coffee from Central America to the United States. They found that the green organic coffee import prices hurt the import volume by a 1% increase in the import price, and the import quantity would be reduced by 1.899%. In addition, the Gross Domestic Product (GDP) per capita negatively impacted the import volume of organic green coffee from Central America due to higher U.S. incomes, leading to more imports of mild coffee from Kenya, Tanzania, and Colombia [39]. They didn't report test methods used to compare or evaluate the model.

Nguyen et al. [33] used 742 observations of monthly coffee price time series data from January 1960 to October 2021 from the Knoema platform to create a forecasting model, which starts with preprocessing the data from the 'Haar' family of wavelet transform to filter out noisy data. Data were trained with an echo state network (ESN) model and also with an optimized ESN based on the gray wolf optimization (GWO) technique; the best root mean square error (RMSE) achieved of 0.03 was for the optimized ESN.

Xu and Zhang utilized a broad dataset of futures contract data prices for coffee and other agricultural products over an extensive period of 50 to 63 years, covering eight agricultural commodities [19]. This included 12029 observations for coffee. They explored nonlinear autoregressive models for forecasting agricultural commodity prices; different configurations were tested, and a model with three hidden neurons and four lags for coffee was arrived at. The model could capture long-term price trends and market dynamics influencing coffee prices. After, they got a 2.47% RMSE.

Deina et al. [18] sourced their data from the website of the University of São Paulo, Brazil. This dataset is available at <http://cepea.esalq.usp.br>. The time range used is from 2001 to 2018 for Robusta Coffee (206 records) and from 1996 to 2018 for Arabica (271 records). The data comprises monthly coffee price figures in Brazilian currency (Reais) for 60 kg bags. They used the differencing method to make the time series stationary and experimented with different time series-based and neural network models. They used the MSE, the MAE, and the MAPE. They finally found that ELM performed better than the others.

In 2023, Mekala et al. [15] gathered historical coffee price data from the Intercontinental Exchanges, and the New York Stock Exchange (NYSE). Additionally, they incorporated weather and economic data relevant to coffee-producing countries. With that data, they created a system based on a combination of long and short-term bidirectional neural networks (BLSTM) and convolutional neural networks (CNN) to predict the price of coffee. They also trained other models for comparison purposes. The proposed model ranks first with 0.98 recall, last in accuracy with 0.74, and second to last in F1 Score with 0.77.

Merbah and Benito [14] collected a sample of 645 coffee packages sold across Spain's five largest supermarket chains between September and October 2021. From this dataset, they used a set of variables such as brand, quantity, type of business where the coffee is sold, coffee intensity, origin, weather, and certifications; the model was based on the one proposed by Rosen. Equation 5 shows it, where β is a dummy parameter to estimate the implicit price; ε_i is an error value. Finally, they evaluated the model with R^2 (0.77), adjusted R^2 (0.77), and the F-score(0); therefore, the model presents homoscedasticity.

$$\ln(P) = \text{Constant} + \beta_1 \ln(\text{Quantity}) + \beta_2(\text{Brand}) + \beta_3(\text{Hypermarket}) \\ + \beta_4(\text{Neighborhood}) + \beta_5(\text{Discount}) + \beta_6(\text{Fairtrade}) + \beta_7(\text{Rainforest}) + \beta_8(\text{UTZ}) \quad (5) \\ + \beta_9(\text{Organic}) + \beta_{10}(\text{Intensity}) + \beta_{11}(\text{Organic}) + \varepsilon_i$$

Chaovanapoonphol et al. [20] used secondary sources to analyze coffee price volatility, gathering 23 observations from 1999 to 2021; among other factors, they used the price of raw coffee in Thailand (TFP), the manufacturing demand in Thailand (TDD), and the export volume from key producers: Indonesia (INEX) and Brazil (BEX). Then, they used a Bayesian estimation technique in conjunction with a GARCH model with Covariates X (GARCH-X) model to analyze the volatility of the price of TFP. They concluded that the most influential variables in the volatility of coffee prices in Thailand are TDD with an SE of 0.0003, BEX with 0.0004, and INEX with an SE of 0.0015.

Le Ngoc et al. [25] employed data from 2021 through May 19, 2023. The study leveraged daily coffee prices. Additionally, they included the diesel fuel price and the rainfall and temperature data, resulting in a dataset of 869 records. The study considered the temperature because Robusta coffee thrives best between 24 and 30 °C. Data was used to train four-time series models, and other machine learning, being Random Forest (RF) the model that showed the best results when the six models were compared based on the RMSE, MAE, MAPE, and Mean Absolute Scaled Error (MASE). Results for RF were RMSE of 639.63, a MAE of 382.54, a MAPE of 0.01 and a MASE of 0.09.

3.2 Research Questions

What variables were used to calculate the specialty coffee price?

From the articles reviewed, it was found that the main variables used to estimate the price of specialty coffee through linear regression models are sensory test, ranking in the auction, country of origin, variety, lot size, altitude, and year of the auction; as well as whether it was organic and its certifications [8]-[11], [14], [31]. Researchers also discovered that time-series-based techniques can forecast the future value of coffee using solely coffee base price and the date of each price [18], [25], [33]; in addition to time series, neural network models can also be used for forecasting [15] and full sample (2004-2019 [22], [25], [26], [29]-[30]. Comparing the current variables with those used by the linear regression to estimate specialty coffee prices [8]-[11], it can be said that they are outdated and do not take into account some very desirable attributes of coffee currently produced in Colombia such as the Geisha and the pink Bourbon varieties. This factor negatively influences the accuracy of the models since applying the proposed models in current situations.

What open access data sources can be found to train models for calculating the price of specialty coffee?

Concerning the data used by those researchers, most of the authors who focused on linear regression used CoE [8]-[11] and data from supermarkets [14], [31]. It is essential to consider data from supermarkets because it compares the prices at which coffee is sold by coffee growers and the value at which it is purchased in consuming countries. Still, supermarket data are unrelated to the CoE as they target a different consumer.

Regarding other data sources, price data for some dates were found in Brazil from 1996 to 2018 (<http://cepea.esalq.usp.br>); the NYSE (<https://www.nyse.com/index>); ICO (<https://icocoffee.org>), the Intercontinental Exchange (<https://www.ice.com/index>), and other secondary data like the internal prices of coffee in different countries [15]-[18], [20]-[30], [33].

What models, techniques, or methods have been used to estimate the price of specialty coffee?

Based on the review, six authors were found that focus on the estimation of the price of specialty coffee through regression models [8]-[11], [14], [31]; in this research, they found that the country of origin dramatically influences the price of specialty coffee, it can be due to their quality reputation [9], [11] moreover, the specialty coffee certifications were essential to countries with high sustainability awareness like Swedish and Spain [14], [31]. Other important variable was the buyer origin because some countries prefer some attributes over others; for the North American market, the sensory quality score is crucial; for one score point, the increases are 40% concerning the European and Asian markets [10].

Other authors analyzed time series to forecast the price of coffee; the principal model used was ARIMA [17]-[18], [22], [25], [28]-[29]. However, Le Ngoc, who analyzed Robusta coffee data, found that Seasonal ARIMA (SARIMA) got an RMSE of 1086, ARIMA of 3482, and RF of 1063; for this reason, the author suggests that the best model is RF, and it can be said that the latter obtained a 69.47% lower RMSE than ARIMA [25]. Likewise, ELM had 27.29% better results than ARIMA for the Arabica time series, and MLP had 33.78% better than ARIMA for the Robusta time series. It is possible to say that Neural Networks could get better results than time series models like ARIMA [18], but new evaluations are required. Other models used for time series were NGBM, VECM, ECM, SARIMA GARCH-X, and ES [16] and full sample (2004-2019, [20]-[22], [25], [30]; nevertheless, in those researches, there is no benchmark between these models, except ARIMA vs. ELMs vs. MLP, where ELMs was the best [18]; and ARIMA vs. SARIMA vs. LSTM vs. SVM vs. GRU vs. RF [25], where RF was the best. To apply ARIMA and other time series models, the time series must be stationary [40]; after using the DF test in the most of the future and spot prices of coffee time series, it was found that it wasn't stationary and was necessary to use techniques to convert the data to stationary [22], [28], [30] so a method that able to obtain a good forecasting result from a non-stationary multivariate time series data is needed. Vector Error Correction Model (VECM, [32].

Mekala et al. [15] used artificial intelligence techniques such as CNN-BLSTM, based on their results, it can be said that while the model makes accurate predictions, it is susceptible to error in certain instances. Herrera et al. [26] used the LSTM model to forecast the Price of coffee, using prices from the FNC in Colombia; other neural network used was ESN [33].

What metrics are used to assess the quality of the prediction?

The main metrics used in linear regression models for the comparison in the quality prediction were AIC [9]-[11], and R2 was used in [8]-[9], [14], [31]. AIC was used to evaluate time series models [28]-[30] so a method that able to obtain a good forecasting result from a non-stationary multivariate time series data is needed. Vector Error Correction Model (VECM. In the time series model, the principal metric was MAPE [17], [18], [21], [25], [28], and to evaluate the artificial neural network-based model mainly used MSE [18], [23], [26], RMSE [19], [25], [33] and F Score [14]- [15].

4. CONCLUSIONS AND FUTURE WORK

Analysis of data sources reveals that supermarket data from importing countries and auction data can provide helpful information on consumer preferences. In addition, historical stock prices can determine the future basis value of specialty coffees.

Most studies conducted before 2018 rely on regression models to estimate the price of specialty coffee, using auction data as the primary data source. The most frequently recurring attributes are tree variety, sensory evaluation, auction ranking, country of origin, and lot size. After 2018, most studies focus on time series prediction models, utilizing historical coffee price data from different countries. Only a few authors used other models, such as ensembles like Random Forest, neural networks, and convolutional neural networks.

From 2018, the ARIMA model has been used in 5 of the 14 studies reported in this review. However, alternative models produced better results in two investigations; therefore, research using alternative models to ARIMA is recommended for future work.

This study identified R^2 as the principal metric for evaluating linear regression models. MSE was mainly used to assess artificial neural network-based models, and MAPE was used for time series models.

Future research should use data with attributes that accurately represent the production context and other factors that reflect the quality of the coffee. This could include growing site characteristics, production costs, seeds, country of origin and destination, climatic conditions, coffee tree variety, production volume, auction date, sensory evaluations, altitude of the crop, organic certification, processing methods, fair trade certification, the comments in social networks, trade policies, and economic changes. In addition, it is essential to explore linear and nonlinear artificial intelligence methods that provide accurate predictions and identify the most influential characteristics.

AUTHORS' CONTRIBUTION

Victor-Hugo Pinto-Rodriguez: Methodology; Investigation; Formal analysis; Writing – original draft; Writing – review & editing.

Carlos-Alberto Cobos-Lozada: Conceptualization; Formal analysis; Supervision; Writing – review & editing.

Adriana-Marcela Nieto-Muñoz: Conceptualization; Formal analysis; Supervision; Writing – review & editing.

ACKNOWLEDGMENTS

The University of Cauca partially supported this work. The Technological Development Center Creativ also supported this research through the project COD BPIN 2020000100538, funded by the Ministry of Science, Technology, and Innovation of Colombia.

REFERENCES

- [1] J. Norton et al., "A grand challenge for HCl: food + sustainability," *Interactions*, vol. 24, no. 6, pp. 50-55, Oct. 2017. <https://doi.org/10.1145/3137095>
- [2] Federación Nacional de Cafeteros, *Colombia celebra el día nacional del café y 93 años de la Federación Nacional de Cafeteros*, 2023. <https://federaciondecafeteros.org/wp/listado-noticias/colombia-celebra-el-dia-nacional-del-cafe-y-93-anos-de-la-federacion-de-cafeteros/>
- [3] S. G. Posada, *La economía del café: ¿Quién se está quedando el dinero?*, 2019. <https://quecafe.info/la-economia-del-cafe-quien-se-esta-quedando-el-dinero/>
- [4] I. Perfecto, R. Rice, R. Greenberg, M. Van der Voort, "Shade Coffee: A Disappearing Refuge for Biodiversity," *Bioscience*, vol. 46, no. 8, pp. 598-608. <https://doi.org/10.2307/1312989>
- [5] Firebird, *What is Specialty Coffee?*, 2024. https://www.firebirdcoffee.co.uk/blogs/roastersnotebook/what-is-specialty-coffee?srsitid=AfmBOoruVtKvWCcFi_RymlCvbbZusqz9YYHv3mWKEo2XHIKDF0k0YVCJ
- [6] Federación Nacional de Cafeteros, *Precio interno de referencia para la compra de café en Colombia*, 2023. https://federaciondecafeteros.org/app/uploads/2019/10/precio_cafe.pdf
- [7] S. Dávila-Hermeling, *¿Cómo se determina el precio del café? Introducción a la bolsa de valores y el mercado de futuros*, 2023. <https://perfectdailygrind.com/es/2021/03/04/como-se-determina-el-precio-del-cafe-introduccion-a-la-bolsa-de-valores-y-el-mercado-de-futuros/>
- [8] L. Donnet, D. Weatherspoon, J. Hoehn, "Price determinants in top-quality E-Auctioned specialty coffees," *Agricultural Economics*, vol. 38, pp. 267-276, May 2008. <https://doi.org/10.1111/j.1574-0862.2008.00298.x>
- [9] R. Teuber, R. Herrmann, "Towards a differentiated modeling of origin effects in hedonic analysis: An application to auction prices of specialty coffee," *Food Policy*, vol. 37, no. 6, pp. 732-740, 2012. <https://doi.org/10.1016/j.foodpol.2012.08.001>
- [10] A. P. Wilson, N. L. W. Wilson, "The economics of quality in the specialty coffee industry: Insights from the Cup of Excellence auction programs," *Agricultural Economics (United Kingdom)*, vol. 45, no. S1, pp. 91-105, 2014. <https://doi.org/10.1111/agec.12132>
- [11] T. Traore, N. Wilson, D. Fields, "What explains specialty coffee quality scores and prices: a case study from the cup of excellence program," *Journal of Agricultural and Applied Economics*, vol. 50, no. 3, pp. 349-368, Aug. 2018. <https://doi.org/10.1017/aae.2018.5>
- [12] B. Kitchenham, O. Pearl Brereton, D. Budgen, M. Turner, J. Bailey, S. Linkman, "Systematic literature reviews in software engineering-A systematic literature review," *Information and Software Technology*, vol. 51, no. 1, pp. 7-15, 2009. <https://doi.org/10.1016/j.infsof.2008.09.009>
- [13] W. Mengist, T. Soromessa, G. Legese, "Method for conducting systematic literature review and meta-analysis for environmental science research," *MethodsX*, vol. 7, e100777, Dec. 2019. <https://doi.org/10.1016/j.mex.2019.100777>

- [14] N. Merbah, S. Benito, "Sustainability labels in the Spanish coffee market: A hedonic price approach," *Spanish Journal of Agricultural Research*, vol. 21, no. 1, e19510, 2023.
<https://doi.org/10.5424/sjar/2023211-19510>
- [15] K. Mekala, V. Laxmi, H. Jagruthi, D. Shiv Ashish, R. Sridevi, Amar Prakash, "Coffee Price Prediction: An Application of CNN-BLSTM Neural Networks," in *International Conference on Advances in Computing, Communication and Applied Informatics*, 2023. <https://doi.org/10.1109/ACCAI58221.2023.10199369>
- [16] O. Gálvez-Soriano, M. Cortés, "Is there a pass-through from the international coffee price to the mexican coffee market?," *Studies in Agricultural Economics*, vol. 123, no. 2, pp. 86-94, 2021.
<https://doi.org/10.7896/j.2143>
- [17] R. R. Novanda *et al.*, "A comparison of various forecasting techniques for coffee prices," in *Journal of Physics: Conference Series*, 2018.
- [18] C. Deina *et al.*, "A methodology for coffee price forecasting based on extreme learning machines," *Information Processing in Agriculture*, vol. 9, no. 4, pp. 556-565, Dec. 2022.
<https://doi.org/10.1016/j.inpa.2021.07.003>
- [19] X. Xu, Y. Zhang, "Commodity price forecasting via neural networks for coffee, corn, cotton, oats, soybeans, soybean oil, sugar, and wheat," *Intelligent Systems in Accounting, Finance and Management*, vol 29, no. 3, pp. 169-181. Jul. 2022. <https://doi.org/10.1002/isaf.1519>
- [20] Y. Chaovanapoonphol, J. Singvejsakul, A. Wiboonpongse, "Analysis of exogenous factors to Thailand coffee price volatility: using multiple exogenous Bayesian Garch-X model," *Agriculture*, vol. 13, no. 10, e10, Oct. 2023. <https://doi.org/10.3390/agriculture13101973>
- [21] C.-N. Wang, M.-C. Yu, N.-N.-Y. Ho, T.-N. Le, "An integrated forecasting model for the coffee bean supply chain," *Applied Economics*, vol. 53, no. 28, pp. 3321-3333, Jun. 2021.
<https://doi.org/10.1080/00036846.2021.1887447>
- [22] A. Acevedo Amoroch, F. E. Ramírez Carreño, D. D. Salcedo Blanco, J. A. Román Ordoñez, "Pronóstico del precio del café: Una propuesta desde los modelos econométricos," *Revista Venezolana de Gerencia*, vol. 25, no. 4, pp. 564-578, 2020.
- [23] A. Rodríguez, M. Melgarejo, "Identification of Colombian coffee price dynamics," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 1, e013145, Jan. 2020.
<https://doi.org/10.1063/1.5119857>
- [24] L. Benavidez, T. Xia, "A Gravity Model of Central American Organic Coffee Trade with the United States," *Journal of Food Distribution Research*, vol. 53, no. 1, pp. 33-40, 2022.
<https://doi.org/10.22004/ag.econ.339667>
- [25] T. N. Le Ngoc *et al.*, "Machine Learning for agricultural price prediction: a case of coffee commodity in Vietnam market," in *IEEE/ACIS 8th International Conference on Big Data, Cloud Computing, and Data Science*, 2023, pp. 38-41. <https://doi.org/10.1109/BCD57833.2023.10466313>
- [26] Y. A. Herrera-Jaramillo, J. C. Ortega-Giraldo, A. Acevedo-Amoroch, D. Prada-Marin, "Colombian coffee price forecast via LSTM Neural Networks," in *Technological and Industrial Applications Associated with Intelligent Logistics*, 2021, pp. 501-517. https://doi.org/10.1007/978-3-030-68655-0_25

- [27] E. Figueroa-Hernández, F. Pérez-Soto, L. Godínez-Montoya, and R. A. Perez-Figueroa, "Los precios de café en la producción y las exportaciones a nivel mundial," *Revista Mexicana de Economía y Finanzas Nueva Época REMEF*, vol. 14, no. 1, Art. no. 1, 2019, <https://doi.org/10.21919/remef.v14i1.358>
- [28] D. G. Ababu, A. M. Getahun, "Time series analysis of price of coffee in case of Mettu town, Ilu Ababor zone, Oromia regional state, Ethiopia," *Asian Journal of Dairy and Food Research*, vol. 40, no. 3, pp. 279-284, Jul. 2021. <https://doi.org/10.18805/ajdfr.DR-204>
- [29] J. Crespo Cuaresma, J. Hlouskova, M. Obersteiner, "Fundamentals, speculation or macroeconomic conditions? Modelling and forecasting Arabica coffee prices," *European Review of Agricultural Economics*, vol. 45, no. 4, pp. 583–615, Sep. 2018. <https://doi.org/10.1093/erae/jby010>
- [30] W. S. Nugroho, A. B. Astuti, S. Astutik, "Vector error correction model to forecasting spot prices for coffee commodities during Covid-19 pandemic," *Journal of Physics: Conference Series*, vol. 1811, no. 1, e012076, Mar. 2021. <https://doi.org/10.1088/1742-6596/1811/1/012076>
- [31] L. Schollenberg, "Estimating the hedonic price for Fair Trade coffee in Sweden," *British Food Journal*, vol. 114, no. 3, pp. 428–446, Jan. 2012. <https://doi.org/10.1108/00070701211213519>
- [32] M. J. Holmes, J. Otero, "Psychological price barriers, El Niño, La Niña: New insights for the case of coffee," *Journal of Commodity Markets*, vol. 31, e100350, Sep. 2023. <https://doi.org/10.1016/j.jcomm.2023.100350>
- [33] D. T. Nguyen, C. Y. Hwang, B. L. Le Ngoc, S. Y. Sam, "Data analytics and optimised machine learning algorithm to analyse coffee commodity prices," *International Journal of Sustainable Agricultural Management and Informatics*, vol. 8, no. 4, e345, 2022. <https://doi.org/10.1504/IJSAMI.2022.126799>
- [34] Cup of Excellence, *Cup of Excellence winners in Colombia 2023 auction*, 2024. <https://allianceforcoffeexcellence.org/colombia-2023/>
- [35] G. C. Chow, "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, vol. 28, no. 3, pp. 591–605, 1960. <https://doi.org/10.2307/1910133>
- [36] S. Rosen, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, vol. 82, no. 1, pp. 34–55, Jan. 1974. <https://doi.org/10.1086/260169>.
- [37] Giacaphe, *Lam Dong coffee price*, 2024. <https://giacaphe.com/gia-ca-phe-lam-dong/>
- [38] T. Mayer, S. Zignago, *Notes on CEPII's Distances Measures: The GeoDist Database*, 2011. http://www.cepii.fr/PDF_PUB/wp/2011/wp2011-25.pdf
- [39] J. E. Houston, M. Santillan, J. Marlowe, "US demand for mild coffees: Implications for Mexican coffee," *Journal of Food Distribution Research*, vol. 34, no. 1, pp. 92–98, 2003. <https://doi.org/10.22004/ag.econ.27956>.
- [40] J. F. Sánchez Cifuentes, M. M. Cuellar Chaves, *Análisis de series de tiempo con métodos econométricos para el control de congestión en redes de telecomunicaciones*, Master Thesis, Unicauca, 2018. <https://elibro-net.acceso.unicauca.edu.co/es/lc/unicauca/titulos/128023>