Revista Agri-Environmental Sciences, Palmas-TO, v. 12, e026003, 2026
DOI: https://doi.org/10.36725/agries.v12i1.10296
https://revista.unitins.br/index.php/agri-environmental-sciences/index Artigo Científico
ISSN 2525-4804

STRUCTURAL ECONOMIC ANALYSIS AND PREDICTIVE MODELING OF FISH FINGERLING PRODUCTION IN BRAZIL: A MULTIVARIATE INVESTIGATION BASED ON THE IBGE/SIDRA DATABASE (2013–2023)

Murilo Henrique Tank Fortunato[1], Karoline Moreira Barbuio [2]

ABSTRACT:

Fingerling production, the breeding and commercialization of young fish, represents the main biological and economic link in Brazilian fish farming. This study aims to analyze the economic structure of this sector using a Data Science approach applied to Table 3940 from the SIDRA/IBGE system. Through a rigorous analytical protocol integrating Exploratory Data Analysis (EDA), Partition-Based Clustering (K-means), and Multiple Linear Regression with regional error correction, the article indicates that value generation in the sector is governed by a pronounced dichotomy between large-scale operational efficiency and the economic value of regional niche markets. The study concludes that the long-term sustainability of the sector depends on correcting pricing asymmetries in the Northeast region and maintaining productive stability in the Center-South of Brazil.

Keywords: Machine learning, fish farming, economy.

ANÁLISE ECONÔMICA ESTRUTURAL E MODELAGEM PREDITIVA DA ALEVINAGEM DE PEIXES NO BRASIL: UMA INVESTIGAÇÃO MULTIVARIADA SOBRE A BASE DE DADOS IBGE/SIDRA (2013–2023)

RESUMO

A alevinagem, correspondente à produção e comercialização de peixes jovens (alevinos), constitui o elo biológico e financeiro primário da piscicultura brasileira. Este estudo tem como objetivo dissecar a estrutura econômica deste setor, utilizando uma abordagem de Ciência de Dados aplicada à Tabela 3940 do sistema SIDRA/IBGE. Através de um protocolo rigoroso que integra Análise Exploratória de Dados (AED), Agrupamento por Partição de Médias (K-means) e Regressão Linear Múltipla com correção de erros regionais, o artigo indicou que a formação de valor no setor é governada por uma dicotomia acentuada entre a eficiência operacional de escala e a valorização de nichos regionais. O estudo conclui que a sustentabilidade do setor depende da correção das assimetrias de precificação no Nordeste e da manutenção da estabilidade produtiva no Centro-Sul.

Palavras-chave: Aprendizado de máquina, piscicultura, economia.

---

[1] Biologist, Master in Environmental Sciences, and PhD in Sustainable Agriculture. Academic advisor in the MBA program at USP/ESALQ, Piracicaba, São Paulo; mtank6691@gmail.com; https://orcid.org/0000-0002-1695-8120. [2] Master in Animal Science and Pastures, University of São Paulo, Piracicaba, São Paulo; karolinembarbuio@usp.br; https://orcid.org/0009-0002-1467-4092.

Received 21 Jan, 2026 • Accepted 09 Mar, 2026

INTRODUCTION

Brazilian fish farming has evolved from a subsistence activity into a complex agro-industrial chain, intensive in technology, capital, and scientific knowledge. Embedded within the concept of the blue agribusiness, this activity accounts for an increasing share of national aquaculture production, with emphasis on species such as Nile tilapia (*Oreochromis niloticus*), tambaqui (*Colossoma macropomum*), pacu (*Piaractus mesopotamicus*), and their hybrids, which together sustain most of the country's production volume. The sector also plays a strategic role in food security, job creation, and the productive use of continental freshwater resources (FAO, 2022; Peixe BR, 2025).

At the core of aquaculture development lies fingerling production, a fundamental stage distinct from larviculture, as it encompasses the phase in which fish already adapted to exogenous feeding undergo initial growth, size standardization, and physiological strengthening until suitable for stocking in grow-out systems such as earthen ponds, cage systems, or recirculating aquaculture systems (Cyrino et al., 2010; Baldisserotto, 2013). During this period, essential atributes, including zootechnical performance, health status, stress resistance, and genetic potential expression, are consolidated, directly influencing production efficiency and profitability, particularly in widely farmed species such as Nile tilapia and Amazonian round fishes (Tave, 2011; de Melo et al., 2024). Failures related to genetic quality, management, nutrition, or sanitary protocols can compromise the entire production cycle, increasing mortality, reducing growth, and elevating costs (Miranda et al., 2025; FAO, 2022), which is why fingerling quality is now recognized not merely as an input, but as a strategic asset essential to the economic viability of modern fish farming.

In Brazil, the production and distribution of fingerlings reflect pronounced regional inequalities in infrastructure, access to technology, technical capacity, and organization of production chains. Although aquaculture activities are present in all regions of the country, their intensity and degree of integration vary according to socioeconomic conditions and the availability of technical assistance and marketing services (Cordeiro, 2024). Preliminary evidence from official statistics reinforces these disparities. Data from the Brazilian Institute of Geography and Statistics (IBGE/SIDRA) indicate that fingerling production is strongly concentrated in the Center-South of the country, with the South region accounting for approximately 35% of national production (around 5.09 million kilograms), followed by the Center-West and Southeast regions with 2.83 and 2.73 million kilograms, respectively, while the North region records the lowest production volume, about 1.16 million kilograms. This uneven spatial distribution highlights structural differences in infrastructure, technological adoption, and market integration across Brazilian regions.

Despite Brazil's vast freshwater resources, this natural advantage does not automatically translate into sustainable growth without proper planning and management, particularly regarding the efficient use of water in aquaculture production systems (Valenti et al., 2021; FAO, 2022).The rational use of water is essential to ensure the sustainability of fish farming in the face of growing demand for aquatic protein, while also contributing to the mitigation of environmental impacts, optimization of production efficiency, and reduction of conflicts with other competing water uses such as irrigation, human consumption, and energy generation (Oliveira & Santos, 2024; FAO, 2022). Strategies such as recirculating aquaculture systems, integration with agriculture, and the reuse of effluents for fertigation stand out as viable alternatives to enhance water-use efficiency and the productive resilience of aquaculture systems (de Oliveira & Santos, 2024).

Within this context, economic and productive analyses of fingerling production are essential to understanding regional development patterns in Brazilian fish farming. This study aims to analyze the economic structure and regional disparities of fingerling production in Brazil between 2013 and 2023, using official data from the Brazilian Institute of Geography and Statistics (IBGE/SIDRA). Specifically, the research investigates how production volume and production value vary across the country's five geographic regions and whether distinct regional production profiles can be identified through multivariate statistical methods. By integrating exploratory data analysis, clustering techniques, and regression modeling, the study seeks to identify structural patterns in the formation of economic value within the fingerling production chain. The results contribute to a better understanding of regional asymmetries in Brazilian aquaculture and provide analytical insights to support decisionmaking by producers, technicians, policymakers, and investors, thereby contributing to the sustainable strengthening of blue agribusiness in Brazil.

MATERIALS AND METHODS

Data Source and Statistical Curation

The primary database used in this study was obtained from the IBGE Automatic Recovery System (SIDRA), maintained by the Brazilian Institute of Geography and Statistics (IBGE), with specific extraction from Table 3940 – Aquaculture Production, available at: https://sidra.ibge.gov.br/tabela/3940.

The temporal scope covers the period from 2013 to 2023, a timeframe that enables the observation of structural changes and regional dynamics in the Brazilian aquaculture sector over the last decade. The variables selected for analysis were fingerling production, expressed in kilograms, used as an indicator of physical output and installed production capacity, and the value of production, expressed in Brazilian reais, adopted as a proxy for gross revenue and market liquidity over the analyzed period.

All procedures related to data treatment, organization, and analysis, as well as the codes used to reproduce the statistical and graphical analyses, are publicly available in a GitHub repository accessible at: https://github.com/mtank6691-coder/Artigo-alevinagem.

Computational Processing in R Software

The analyses were conducted using R software, following the principles of scientific reproducibility. Data cleaning and standardization were performed using the *janitor* package, while relational data manipulation relied on the *tidyverse* grammar. To mitigate scale differences between variables, given that monetary values in reais are numerically much larger than production volumes in kilograms, Z-score normalization was applied, centering the mean at zero and the standard deviation at one.

K-means Clustering

To segment the national market, the K-means clustering algorithm was applied. This unsupervised learning technique is widely used to group observations into clusters with similar characteristics based on the minimization of within-cluster distances between observations and their respective centroids. The choice of the number of clusters (k = 3) was validated using the Elbow method, a heuristic procedure that estimates the optimal k by identifying the inflection point in a plot of within-cluster sum of squares (WCSS) as a function of the number of clusters. Beyond this point, additional increases in k yield only marginal reductions in internal variability, indicating a balance between model complexity and clustering quality.

Based on this criterion, three distinct sectoral profiles were identified: **High Performance**, representing regions with greater technical and commercial maturity; **Value Efficiency**, characterized by lower physical output but high unit profitability; and **Base Volatility**, describing regions with large recorded production volumes but inconsistent pricing patterns. The heuristic application of the Elbow method follows established practices in cluster analysis literature (Yuan & Yang, 2019).

Predictive Modeling

The relationship between fingerling production and production value was analyzed using multiple linear regression, a method widely employed to model associations between quantitative variables while controlling for structural effects. Production value, expressed in Brazilian reais, was defined as the dependent variable, whereas the physical production of fingerlings, in kilograms, was considered the main explanatory variable. To capture structural differences among Brazilian regions, categorical variables representing the country's major geographic regions were included and operationalized through dummy variables.

The model was estimated using the ordinary least squares method, assuming linearity, independence of residuals, homoscedasticity, and normality of errors (Wooldridge, 2016). This approach allowed for the evaluation of the marginal effect of production on production value while controlling for regional asymmetries within the fingerling production chain.

RESULTS AND DISCUSSION

Production Ranking and Regional Representativeness

Descriptive results reveal a strong concentration of fingerling production in the Center-

South axis of Brazil (Figure 1). The South Region stands out as the dominant production hub, accounting for approximately 35% of national output, with a total volume of 5.09 million kilograms. This performance reflects high installed capacity, greater production standardization, and higher technical maturity of the sector. The Center-West and Southeast regions follow, with production levels of 2.83 million and 2.73 million kilograms, respectively, followed by the Northeast with 2.70 million kilograms, and the North with the lowest physical output, totaling 1.16 million kilograms. This spatial pattern confirms the historical concentration of commercial fish farming in regions with superior infrastructure, access to inputs, logistics, and consumer markets (Cordeiro, 2024), a pattern also highlighted in broader structural analyses of Brazilian aquaculture development (Pedroza Filho et al., 2020).

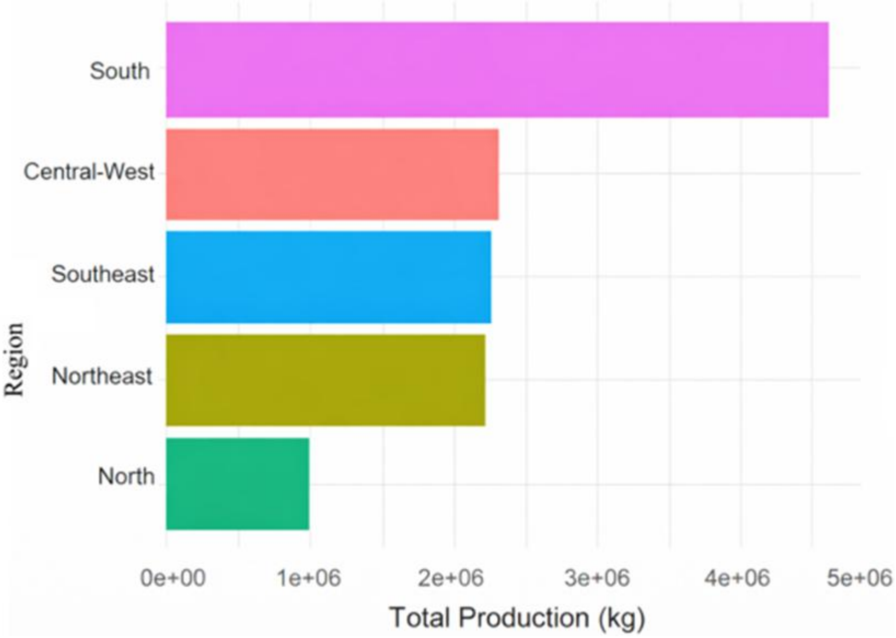

Figure 1. Fingerling production ranking by region

However, when the economic dimension is incorporated into the analysis, a relevant paradox emerges between the North and Northeast regions. Although the Northeast exhibits production volumes comparable to those of the Southeast and Center-West, it records the lowest revenue nationwide, totaling only BRL 376,823, resulting in an average price of approximately BRL 0.13 per kilogram. This outcome suggests a scenario of intense competitive pressure, potential local oversupply, and low product differentiation, characterizing a process of commoditization of fingerling production, in which sale prices tend to converge toward marginal production costs (Pincinato, 2021).

In contrast, the North Region, despite producing less than half the volume of the Northeast, achieves significantly higher revenue (BRL 552,870), sustaining the highest average price in the country, approximately BRL 0.47/kg. This pattern reflects a distinct market structure, likely associated with the production of native species with higher added value, lower regional competition, and relatively scarce supply, factors that raise unit prices. Studies in aquaculture economics indicate that markets with restricted supply and lower elasticity tend to sustain higher prices, especially for differentiated inputs and products such as native fingerlings, highlighting that market value is closely linked to scarcity and product differentiation (FAO, 2022).

This regional contrast reinforces that the economic performance of fingerling production depends not exclusively on production scale, but also on supply chain organization, market positioning, and differentiation strategies adopted in each region, as discussed in the Brazilian aquaculture literature (Valenti et al., 2021).

Cluster Interpretation

The multivariate mapping of the data, presented in Figure 2, clearly reveals the existence of distinct regional production profiles within Brazilian fingerling production. The first cluster is characterized by an industrialized profile, primarily represented by the South Region and parts of the Center-West. In these areas, the dispersion pattern indicates a more stable and predictable relationship between production volume and economic value, suggesting the presence of more integrated production systems, higher levels of technological standardization, and stronger linkages with consumer markets. This behavior is consistent with regions where fish farming has developed more intensively, supported by infrastructure, logistics, and well-organized production chains.

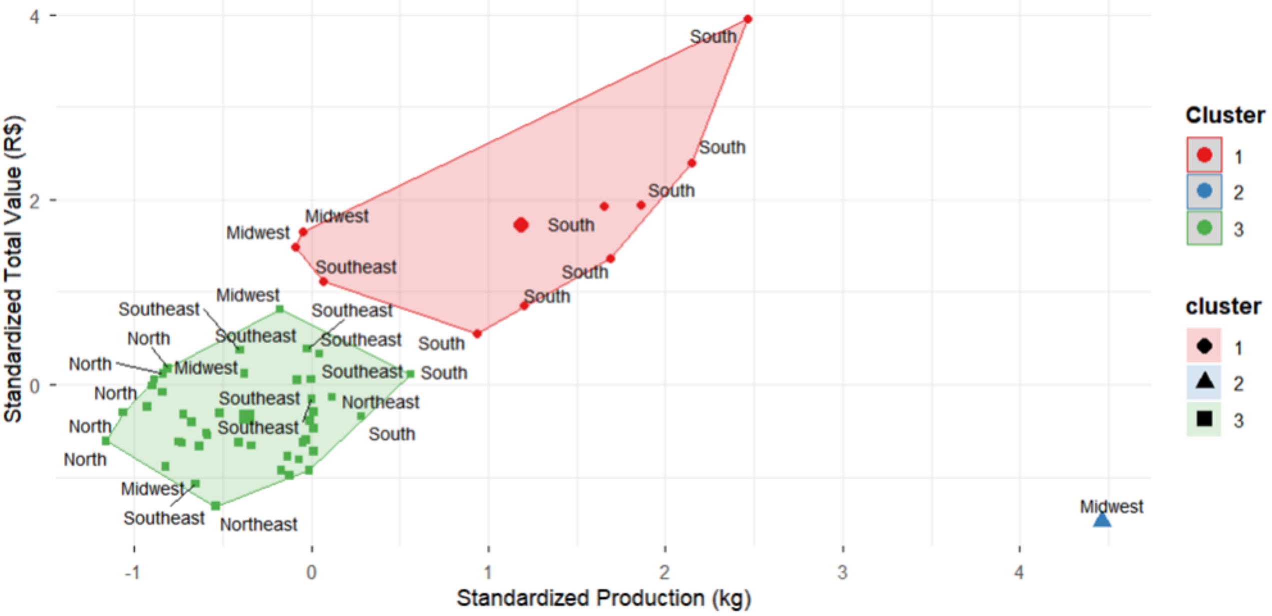

Figure 2. Clustering: Production vs Value (Identified by Region)

Similar dynamics have been described by authors who emphasize the structural transformation of Brazilian aquaculture over the last decades. For example, Sidonio et al. (2012) highlight the growing professionalization and vertical integration of aquaculture production chains in Brazil, particularly in regions with stronger agro-industrial traditions. Likewise, Baccarin et al. (2017) point out that technological adoption, scale gains, and improved coordination between hatcheries, grow-out farms, and processing units have contributed to greater economic efficiency and stability in more consolidated production hubs.

In addition, Pedroza Filho et al. (2020) argue that regions with better access to credit, technical assistance, and logistics infrastructure tend to present more predictable production patterns and stronger market integration, reinforcing the idea that the observed industrialized cluster reflects structural advantages rather than merely higher production volumes. Together, these studies support the interpretation that the South and parts of the Center-West exhibit a more mature and coordinated fingerling production system, aligned with broader processes of intensification and modernization in Brazilian aquaculture.

The second cluster exhibits a **market niche profile,** predominantly associated with observations from the North Region. In this group, relatively lower production volumes are linked to higher economic returns, reflecting unit prices above the national average. This pattern suggests the exploitation of native species with higher added value and less saturated markets, where relative supply scarcity and product differentiation allow for greater profitability per unit produced. The literature on aquaculture in the Amazon region emphasizes the economic potential of native species production and the servicing of specific regional markets, in which scale is not the primary determinant of economic performance (FAO, 2022).

In the third cluster, a pattern emerges that points to structural challenges, concentrating mainly on observations from the Northeast and Southeast regions, where the high dispersion in the production–value space indicates price instability and lower economic predictability. This type of price volatility is widely documented in the agricultural market literature and is associated with supply and demand shocks, climatic factors, short-term production constraints, and complex interconnections between local and global markets, which can exacerbate the risks faced by producers exposed to market fluctuations and information asymmetries (Huchet-Bourdon, 2011). Empirical studies show that price volatility in agricultural value chains adversely affects value chain actors, resulting in yield uncertainties and riskier production decisions, particularly in contexts with lower market coordination and insufficient infrastructure (Mustafa et al., 2023).

Moreover, reviews on economic and supply chain risks highlight that structural vulnerabilities, such as limited market information access, weak logistical integration, and dependence on external conditions, increase farmers' exposure to price shocks and can reduce productive efficiency and economic equity in the agri-food system (Nyamah et al., 2017). Such evidence reinforces the idea that, in contexts of increasing commoditization logic, production coordination strategies, value-adding initiatives, and risk mitigation mechanisms are essential to reduce producers' economic vulnerability and improve market stability, complementing the arguments of Valenti et al. (2021) and FAO (2022).

Regression Model and Predictability

The multiple linear regression model, Figure 3, yielded an adjusted coefficient of determination of 0.2996, indicating that physical fingerling production explains approximately 30% of the observed variation in production value, while the remaining 70% is associated with regional effects and factors not explicitly captured by the model, such as genetic quality, market structure, commercial strategies, and seasonal fluctuations in demand. The estimated coefficient for physical production (0.0565) indicates that, on average, each additional kilogram of fingerlings produced in Brazil is associated with an increase in revenue of approximately BRL 0.06, evidencing a positive but moderate relationship between production scale and value generation.

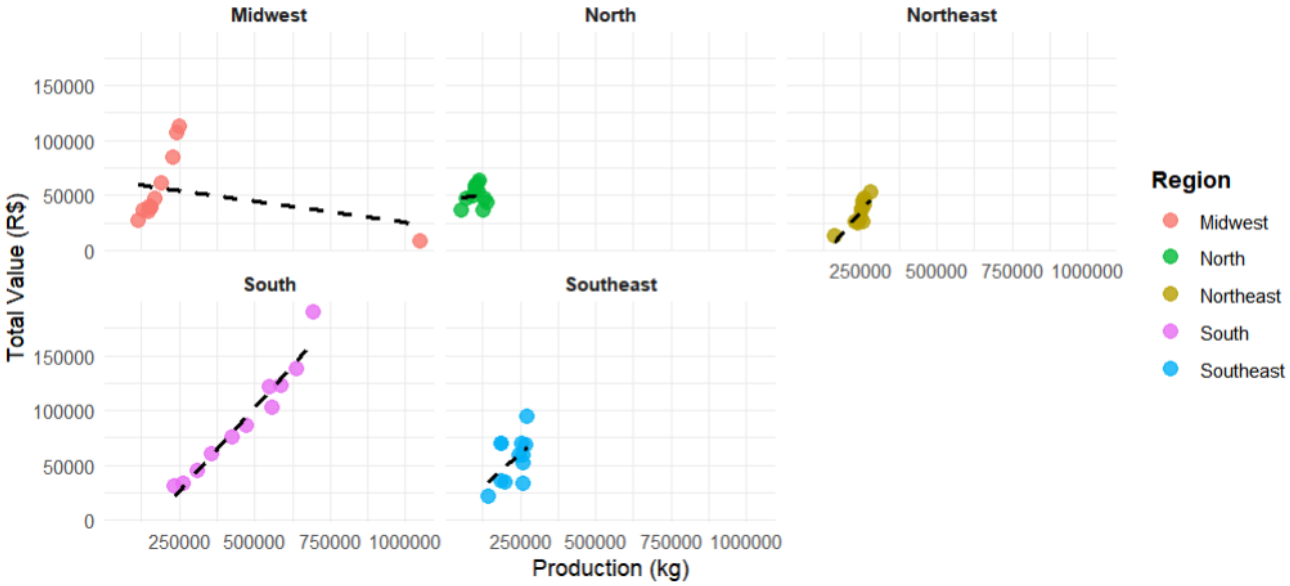

Figure 3. Regression: Production vs Real Value

In turn, the relatively high intercept, estimated at BRL 39,580, suggests the presence of substantial fixed costs and significant financial barriers to entry for new producers in the fingerling segment, reflecting the need for initial investments in infrastructure, technology, and human capital. This pattern is consistent with the classical interpretation of linear regression models in economic contexts, in which the intercept may represent minimum structural costs that are independent of the level of production (Wooldridge, 2016).

CONCLUSION

This study contributes to the understanding of the Brazilian fingerling production sector by integrating official IBGE data with data science techniques to reveal structural and economic asymmetries across regions. By combining descriptive analysis, clustering, and multiple linear regression, the results demonstrate that value generation in fingerling production is not driven solely by production scale, but by a clear dichotomy between industrialized regions with stable, large-scale output and niche-oriented regions capable of generating higher unit value. These findings provide strategic insights for producers, policymakers, and investors seeking to strengthen the sustainability and competitiveness of Brazilian aquaculture.

Despite its contributions, this study presents some limitations. The analysis relies on aggregated secondary data, which may mask intra-regional heterogeneity and firm-level differences in management, genetics, and cost structures. Additionally, relevant determinants of value generation, such as feed costs, survival rates, genetic lineage, contractual arrangements, and market access, were not directly observed and therefore could not be explicitly incorporated into the econometric model. As a result, the explanatory power of the regression, while informative, remains partial.

Future research should focus on integrating micro-level data from hatcheries and farms to better capture technological, biological, and managerial drivers of economic performance in fingerling production. Expanding the analytical framework to include dynamic models, spatial econometrics, or panel data approaches could further improve predictive accuracy and deepen understanding of regional interactions. Moreover, studies evaluating the impacts of genetic improvement programs, water-use efficiency, and biosecurity practices would contribute substantially to the formulation of more effective public policies and private investment strategies in the Brazilian aquaculture sector.

REFERENCES

Andrade, D.R.; Yasui, G.S. Manejo da reprodução natural e artificial e sua importância na produção de peixes no Brasil. **Revista Brasileira de Reprodução Animal**, Belo Horizonte, v. 27, n. 2, p. 166-172, 2003.

Baccarin, J. G.; Triches, R.M.; Teo, C.R.P.A.; Silva, D.B.P.D. Indicadores de avaliação das compras da agricultura familiar para alimentação escolar no Paraná, Santa Catarina e São Paulo. **Revista de Economia e Sociologia Rural**, Brasília, v. 55, n. 1, p. 103-122, 2017.

Baldisserotto, B. **Fisiologia de peixes aplicada à piscicultura**. Santa Maria: Editora UFSM, 2002.

Cordeiro, P. H. Aquicultura no Brasil: análise da evolução do setor entre os anos de 2013 a 2023. **ARACÊ**, v. 7, n. 7, p. 37184-37208, 2025.

Cyrino, J.E.P.; Bicudo, Á.J.D.A.; Sado, R.Y.; Borghesi, R.; Dairik, J.K. A piscicultura e o ambiente: o uso de alimentos ambientalmente corretos em piscicultura. **Revista Brasileira de Zootecnia, Viçosa**, v. 39, p. 68-87, 2010.

Melo, C.L.; Carvalho, I.; Fortunato, M.H.T.; Natel, A.S.; Nascimento, A.F.; Pedreira, M.M. Growth performance, hematological and histological parameters of Nile tilapia larvae fed diets supplemented with β-glucans and nucleotides. **Acta Biologica Brasiliensia**, v. 7, n. 2, p. 254-278, 2024.

Oliveira, E.G.; Santos, F.J.S. Piscicultura e os desafios de produzir em regiões com escassez de água. **Ciência Animal**, Fortaleza, v. 25, n. 1, p. 133-154, 2015.

Food and agriculture organization of the united nations (FAO). **The state of world fisheries and aquaculture 2022**. Rome: FAO, 2022.

Huchet-Bourdon, M. Agricultural Commodity Price Volatility: An Overview. **OECD Food, Agriculture and Fisheries Papers**, Paris, n. 52, OECD Publishing, 2011. Disponível em: https://doi.org/10.1787/5kg0t00nrthc-en. Acesso em: 2026

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produção da aquicultura**. SIDRA – Sistema IBGE de Recuperação Automática. Tabela 3940. Disponível em: https://sidra.ibge.gov.br. Acesso em: 2024.

Miranda, J.P.S.; Ferreira, A.H.C.; Araújo, A.Â.F.; Rodrigues, E.M.; Pereira, A.V.G.; Freitas, A.K.S.; Santos, D.C. Ácido ascórbico em rações para alevinos de tilápias do Nilo. **Ciência Animal**, Fortaleza, v. 35, n. 1, p. 21-32, 2025.

Mustafa, Z.; Vitali, G.; Huffaker, R.; Canavari, M. A systematic review on price volatility in agriculture. **Journal of Economic Surveys**, [S.l.], v. 38, n. 1, p. 268-294, 2024.

Nyamah, E.Y.; Jiang, Y.; Feng, Y.; Enchill, E. Agri-food supply chain performance: an empirical impact of risk. **Management Decision**, [S.l.], v. 55, n. 5, p. 872-891, 2017.

Pedroza Filho, M.X.; Ribeiro, V. S.; Rocha, H.S.; UMMUS, M. E. **Caracterização da cadeia produtiva da tilápia nos principais polos de produção do Brasil.** [S.l.: s.n.], 2020.

PEIXE BR – **Associação Brasileira da Piscicultura. Anuário Peixe BR da piscicultura 2025**. São Paulo: Peixe BR, 2025.

Pincinato, R.B.M. Market aspects and external economic effects of aquaculture. **Aquaculture Economics & Management**, [S.l.], v. 25, n. 2, p. 127-134, 2021.

R CORE TEAM. R: A language and environment for statistical computing. Vienna: **R Foundation for Statistical Computing**, 2023. Disponível em: https://www.R-project.org/.

Sidonio, L.; Cavalcanti, I.; Capanema, L.; Morch, R.; Magalhães, G.; Lima, J.; Burns, V.; Alves Júnior, A. J.; Mungioli, R. **Panorama da aquacultura no Brasil: desafios e oportunidades**. Agroindústria, [S.l.], v. 35, p. 421-463, 2012.

Tave, D.; Jo, J.Y.; Kim, D.S. Gross abnormalities in tilapia. **Fisheries and Aquatic Sciences**, v. 14, n. 2, p. 148-160, 2011.

Valenti, W.C.; Barros, H.P.; Moraes-Valenti, P.; Bueno, G.W.; Cavalli, R.O. Aquaculture in Brazil: past, present and future. **Aquaculture Reports**, v. 19, p. 100611, 2021. DOI: https://doi.org/10.1016/j.aqrep.2021.100611.

Wooldridge, J.M. **Introductory econometrics: a modern approach**. 6. ed. Boston: South-Western Cengage Learning, 2016.

Yuan, C.; Yang, H. **Research on K-value selection method of K-means clustering algorithm**. Journal, v. 2, n. 2, p. 226-235, 2019. DOI: https://doi.org/10.3390/j2020016.