

Engineering Journal: IJOER

Volume-11, Issue-3, March 2025

www.ijoer.com

[Download Now](#)



Preface

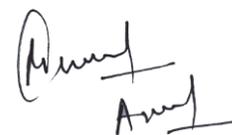
We would like to present, with great pleasure, the inaugural volume-11, Issue-3, March 2025, of a scholarly journal, *International Journal of Engineering Research & Science*. This journal is part of the AD Publications series *in the field of Engineering, Mathematics, Physics, Chemistry and science Research Development*, and is devoted to the gamut of Engineering and Science issues, from theoretical aspects to application-dependent studies and the validation of emerging technologies.

This journal was envisioned and founded to represent the growing needs of Engineering and Science as an emerging and increasingly vital field, now widely recognized as an integral part of scientific and technical investigations. Its mission is to become a voice of the Engineering and Science community, addressing researchers and practitioners in below areas:

Chemical Engineering	
Biomolecular Engineering	Materials Engineering
Molecular Engineering	Process Engineering
Corrosion Engineering	
Civil Engineering	
Environmental Engineering	Geotechnical Engineering
Structural Engineering	Mining Engineering
Transport Engineering	Water resources Engineering
Electrical Engineering	
Power System Engineering	Optical Engineering
Mechanical Engineering	
Acoustical Engineering	Manufacturing Engineering
Optomechanical Engineering	Thermal Engineering
Power plant Engineering	Energy Engineering
Sports Engineering	Vehicle Engineering
Software Engineering	
Computer-aided Engineering	Cryptographic Engineering
Teletraffic Engineering	Web Engineering
System Engineering	
Mathematics	
Arithmetic	Algebra
Number theory	Field theory and polynomials
Analysis	Combinatorics
Geometry and topology	Topology
Probability and Statistics	Computational Science
Physical Science	Operational Research
Physics	
Nuclear and particle physics	Atomic, molecular, and optical physics
Condensed matter physics	Astrophysics
Applied Physics	Modern physics
Philosophy	Core theories

Chemistry	
Analytical chemistry	Biochemistry
Inorganic chemistry	Materials chemistry
Neurochemistry	Nuclear chemistry
Organic chemistry	Physical chemistry
Other Engineering Areas	
Aerospace Engineering	Agricultural Engineering
Applied Engineering	Biomedical Engineering
Biological Engineering	Building services Engineering
Energy Engineering	Railway Engineering
Industrial Engineering	Mechatronics Engineering
Management Engineering	Military Engineering
Petroleum Engineering	Nuclear Engineering
Textile Engineering	Nano Engineering
Algorithm and Computational Complexity	Artificial Intelligence
Electronics & Communication Engineering	Image Processing
Information Retrieval	Low Power VLSI Design
Neural Networks	Plastic Engineering

Each article in this issue provides an example of a concrete industrial application or a case study of the presented methodology to amplify the impact of the contribution. We are very thankful to everybody within that community who supported the idea of creating a new Research with IJOER. We are certain that this issue will be followed by many others, reporting new developments in the Engineering and Science field. This issue would not have been possible without the great support of the Reviewer, Editorial Board members and also with our Advisory Board Members, and we would like to express our sincere thanks to all of them. We would also like to express our gratitude to the editorial staff of AD Publications, who supported us at every stage of the project. It is our hope that this fine collection of articles will be a valuable resource for *IJOER* readers and will stimulate further research into the vibrant area of Engineering and Science Research.



Mukesh Arora
(Chief Editor)

Board Members

Mr. Mukesh Arora (Editor-in-Chief)

BE (Electronics & Communication), M.Tech (Digital Communication), currently serving as Assistant Professor in the Department of ECE.

Prof. Dr. Fabricio Moraes de Almeida

Professor of Doctoral and Master of Regional Development and Environment - Federal University of Rondonia.

Dr. Parveen Sharma

Dr Parveen Sharma is working as an Assistant Professor in the School of Mechanical Engineering at Lovely Professional University, Phagwara, Punjab.

Prof. S. Balamurugan

Department of Information Technology, Kalaignar Karunanidhi Institute of Technology, Coimbatore, Tamilnadu, India.

Dr. Omar Abed Elkareem Abu Arqub

Department of Mathematics, Faculty of Science, Al Balqa Applied University, Salt Campus, Salt, Jordan, He received PhD and Msc. in Applied Mathematics, The University of Jordan, Jordan.

Dr. AKPOJARO Jackson

Associate Professor/HOD, Department of Mathematical and Physical Sciences, Samuel Adegboyega University, Ogwa, Edo State.

Dr. Ajoy Chakraborty

Ph.D.(IIT Kharagpur) working as Professor in the department of Electronics & Electrical Communication Engineering in IIT Kharagpur since 1977.

Dr. Ukar W. Soelistijo

Ph D, Mineral and Energy Resource Economics, West Virginia State University, USA, 1984, retired from the post of Senior Researcher, Mineral and Coal Technology R&D Center, Agency for Energy and Mineral Research, Ministry of Energy and Mineral Resources, Indonesia.

Dr. Samy Khalaf Allah Ibrahim

PhD of Irrigation &Hydraulics Engineering, 01/2012 under the title of: "Groundwater Management under Different Development Plans in Farafra Oasis, Western Desert, Egypt".

Dr. Ahmet ÇİFCİ

Ph.D. in Electrical Engineering, Currently Serving as Head of Department, Burdur Mehmet Akif Ersoy University, Faculty of Engineering and Architecture, Department of Electrical Engineering.

Dr. M. Varatha Vijayan

Annauniversity Rank Holder, Commissioned Officer Indian Navy, Ncc Navy Officer (Ex-Serviceman Navy), Best Researcher Awardee, Best Publication Awardee, Tamilnadu Best Innovation & Social Service Awardee From Lions Club.

Dr. Mohamed Abdel Fatah Ashabrawy Moustafa

PhD. in Computer Science - Faculty of Science - Suez Canal University University, 2010, Egypt.

Assistant Professor Computer Science, Prince Sattam bin AbdulAziz University ALkharj, KSA.

Prof.S.Balamurugan

Dr S. Balamurugan is the Head of Research and Development, Quants IS & CS, India. He has authored/co-authored 35 books, 200+ publications in various international journals and conferences and 6 patents to his credit. He was awarded with Three Post-Doctoral Degrees - Doctor of Science (D.Sc.) degree and Two Doctor of Letters (D.Litt) degrees for his significant contribution to research and development in Engineering.

Dr. Mahdi Hosseini

Dr. Mahdi did his Pre-University (12th) in Mathematical Science. Later he received his Bachelor of Engineering with Distinction in Civil Engineering and later he Received both M.Tech. and Ph.D. Degree in Structural Engineering with Grade "A" First Class with Distinction.

Dr. Anil Lamba

Practice Head – Cyber Security, EXL Services Inc., New Jersey USA.

Dr. Anil Lamba is a researcher, an innovator, and an influencer with proven success in spearheading Strategic Information Security Initiatives and Large-scale IT Infrastructure projects across industry verticals. He has helped bring about a profound shift in cybersecurity defense. Throughout his career, he has parlayed his extensive background in security and a deep knowledge to help organizations build and implement strategic cybersecurity solutions. His published researches and conference papers has led to many thought provoking examples for augmenting better security.

Dr. Ali İhsan KAYA

Currently working as Associate Professor in Mehmet Akif Ersoy University, Turkey.

Research Area: Civil Engineering - Building Material - Insulation Materials Applications, Chemistry - Physical Chemistry – Composites.

Dr. Parsa Heydarpour

Ph.D. in Structural Engineering from George Washington University (Jan 2018), GPA=4.00.

Dr. Heba Mahmoud Mohamed Afify

Ph.D degree of philosophy in Biomedical Engineering, Cairo University, Egypt worked as Assistant Professor at MTI University.

Dr. Aurora Angela Pisano

Ph.D. in Civil Engineering, Currently Serving as Associate Professor of Solid and Structural Mechanics (scientific discipline area nationally denoted as ICAR/08"-"Scienza delle Costruzioni"), University Mediterranea of Reggio Calabria, Italy.

Dr. Faizullah Mahar

Associate Professor in Department of Electrical Engineering, Balochistan University Engineering & Technology Khuzdar. He is PhD (Electronic Engineering) from IQRA University, Defense View, Karachi, Pakistan.

Prof. Viviane Barrozo da Silva

Graduated in Physics from the Federal University of Paraná (1997), graduated in Electrical Engineering from the Federal University of Rio Grande do Sul - UFRGS (2008), and master's degree in Physics from the Federal University of Rio Grande do Sul (2001).

Dr. S. Kannadhasan

Ph.D (Smart Antennas), M.E (Communication Systems), M.B.A (Human Resources).

Dr. Christo Ananth

Ph.D. Co-operative Networks, M.E. Applied Electronics, B.E Electronics & Communication Engineering Working as Associate Professor, Lecturer and Faculty Advisor/ Department of Electronics & Communication Engineering in Francis Xavier Engineering College, Tirunelveli.

Dr. S.R.Boselin Prabhu

Ph.D, Wireless Sensor Networks, M.E. Network Engineering, Excellent Professional Achievement Award Winner from Society of Professional Engineers Biography Included in Marquis Who's Who in the World (Academic Year 2015 and 2016). Currently Serving as Assistant Professor in the department of ECE in SVS College of Engineering, Coimbatore.

Dr. Balasubramanyam, N

Dr.Balasubramanyam, N working as Faculty in the Department of Mechanical Engineering at S.V.University College of Engineering Tirupati, Andhra Pradesh.

Dr. PAUL P MATHAI

Dr. Paul P Mathai received his Bachelor's degree in Computer Science and Engineering from University of Madras, India. Then he obtained his Master's degree in Computer and Information Technology from Manonmanium Sundaranar University, India. In 2018, he received his Doctor of Philosophy in Computer Science and Engineering from Noorul Islam Centre for Higher Education, Kanyakumari, India.

Dr. M. Ramesh Kumar

Ph.D (Computer Science and Engineering), M.E (Computer Science and Engineering).

Currently working as Associate Professor in VSB College of Engineering Technical Campus, Coimbatore.

Dr. Maheshwar Shrestha

Postdoctoral Research Fellow in DEPT. OF ELE ENGG & COMP SCI, SDSU, Brookings, SD Ph.D, M.Sc. in Electrical Engineering from SOUTH DAKOTA STATE UNIVERSITY, Brookings, SD.

Dr. D. Amaranatha Reddy

Ph.D. (Postdoctoral Fellow, Pusan National University, South Korea), M.Sc., B.Sc. : Physics.

Dr. Dibya Prakash Rai

Post Doctoral Fellow (PDF), M.Sc., B.Sc., Working as Assistant Professor in Department of Physics in Pachhunga University College, Mizoram, India.

Dr. Pankaj Kumar Pal

Ph.D R/S, ECE Deptt., IIT-Roorkee.

Dr. P. Thangam

PhD in Information & Communication Engineering, ME (CSE), BE (Computer Hardware & Software), currently serving as Associate Professor in the Department of Computer Science and Engineering of Coimbatore Institute of Engineering and Technology.

Dr. Pradeep K. Sharma

PhD., M.Phil, M.Sc, B.Sc, in Physics, MBA in System Management, Presently working as Provost and Associate Professor & Head of Department for Physics in University of Engineering & Management, Jaipur.

Dr. R. Devi Priya

Ph.D (CSE), Anna University Chennai in 2013, M.E, B.E (CSE) from Kongu Engineering College, currently working in the Department of Computer Science and Engineering in Kongu Engineering College, Tamil Nadu, India.

Dr. Sandeep

Post-doctoral fellow, Principal Investigator, Young Scientist Scheme Project (DST-SERB), Department of Physics, Mizoram University, Aizawl Mizoram, India- 796001.

Dr. Roberto Volpe

Faculty of Engineering and Architecture, Università degli Studi di Enna "Kore", Cittadella Universitaria, 94100 – Enna (IT).

Dr. S. Kannadhasan

Ph.D (Smart Antennas), M.E (Communication Systems), M.B.A (Human Resources).

Research Area: Engineering Physics, Electromagnetic Field Theory, Electronic Material and Processes, Wireless Communications.

Mr. Amit Kumar

Amit Kumar is associated as a Researcher with the Department of Computer Science, College of Information Science and Technology, Nanjing Forestry University, Nanjing, China since 2009. He is working as a State Representative (HP), Spoken Tutorial Project, IIT Bombay promoting and integrating ICT in Literacy through Free and Open Source Software under National Mission on Education through ICT (NMEICT) of MHRD, Govt. of India; in the state of Himachal Pradesh, India.

Mr. Tanvir Singh

Tanvir Singh is acting as Outreach Officer (Punjab and J&K) for MHRD Govt. of India Project: Spoken Tutorial - IIT Bombay fostering IT Literacy through Open Source Technology under National Mission on Education through ICT (NMEICT). He is also acting as Research Associate since 2010 with Nanjing Forestry University, Nanjing, Jiangsu, China in the field of Social and Environmental Sustainability.

Mr. Abilash

M.Tech in VLSI, BTech in Electronics & Telecommunication engineering through A.M.I.E.T.E from Central Electronics Engineering Research Institute (C.E.E.R.I) Pilani, Industrial Electronics from ATI-EPI Hyderabad, IEEE course in Mechatronics, CSHAM from Birla Institute Of Professional Studies.

Mr. Varun Shukla

M.Tech in ECE from RGPV (Awarded with silver Medal By President of India), Assistant Professor, Dept. of ECE, PSIT, Kanpur.

Mr. Shrikant Harle

Presently working as a Assistant Professor in Civil Engineering field of Prof. Ram Meghe College of Engineering and Management, Amravati. He was Senior Design Engineer (Larsen & Toubro Limited, India).

Zairi Ismael Rizman

Senior Lecturer, Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM) (Terengganu) Malaysia Master (Science) in Microelectronics (2005), Universiti Kebangsaan Malaysia (UKM), Malaysia. Bachelor (Hons.) and Diploma in Electrical Engineering (Communication) (2002), UiTM Shah Alam, Malaysia.

Mr. Ronak

Qualification: M.Tech. in Mechanical Engineering (CAD/CAM), B.E.

Presently working as a Assistant Professor in Mechanical Engineering in ITM Vocational University, Vadodara. Mr. Ronak also worked as Design Engineer at Finstern Engineering Private Limited, Makarpura, Vadodara.

Table of Contents

Volume-11, Issue-3, March 2025

S. No	Title	Page No.
1	Digital Inclusive Finance and Individual Commercial Pension Insurance Participation Authors: YuFei Zhang  DOI: https://dx.doi.org/10.5281/zenodo.15105154  Digital Identification Number: IJOER-MAR-2025-1	01-19
2	Determinants of Childbearing Conceptions in Married Women: A Machine Learning Comparative Study Authors: Li Zhu, Qian Long  DOI: https://dx.doi.org/10.5281/zenodo.15105168  Digital Identification Number: IJOER-MAR-2025-5	20-32
3	Assessment of Antibiotic Residues in Commercial Aquaculture Feeds in Selected Farms in Awka, Anambra State, Nigeria Authors: Toochukwu Franklin OSEGBO; Chika Florence IKEOGU; Chidimma Gift NWANKWO  DOI: https://dx.doi.org/10.5281/zenodo.15113627  Digital Identification Number: IJOER-MAR-2025-7	33-40

Digital Inclusive Finance and Individual Commercial Pension Insurance Participation

YuFei Zhang

School of Big Data Statistics, Guizhou University of Finance and Economics, Guiyang, Guizhou Province

Received: 01 March 2025/ Revised: 08 March 2025/ Accepted: 15 March 2025/ Published: 31-03-2025

Copyright © 2025 International Journal of Engineering Research and Science

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted Non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract— *With the increasingly severe problem of social aging, commercial pension insurance, as an important supplement to personal pension security, has become the focus of public attention. The rise of digital financial inclusion provides new opportunities for individuals to participate in commercial pension insurance, and financial literacy and financial development play a crucial role in this process. This paper explores the relationship between digital financial inclusion and the participation rate of individual commercial pension insurance. The results show that digital inclusive finance significantly improves the probability of residents participating in commercial pension insurance by providing convenient and low-cost financial services. At the same time, financial literacy and financial development play an important intermediary role in this process, helping residents better understand financial products, thereby increasing their willingness to participate in commercial pension insurance. In addition, personal characteristics, family status and regional economic factors also have a significant impact on the participation rate of commercial endowment insurance, healthy individuals, individuals participating in basic endowment insurance, women, individuals with more real estate and vehicles are more inclined to buy commercial endowment insurance, families with more children are less demanding for commercial endowment insurance, and the increase in government public budget expenditure also prompts residents to be more inclined to buy commercial endowment insurance. Further analysis shows that digital inclusive finance has a more significant promoting effect among non-party members, residents who often use the Internet, agricultural residents, urban residents, and residents in the eastern region. Overall, the rapid development of digital inclusive finance has provided a strong impetus for the participation of individual commercial pension insurance, and financial literacy and financial development have played a bridge role in this process, promoting residents' awareness and participation in commercial pension insurance.*

Keywords—*Commercial Endowment Insurance, Digital Financial Inclusion, Financial Literacy, Financial Development.*

I. INTRODUCTION

With the intensification of the global aging problem, the traditional old-age security system is facing huge challenges. In many countries, especially China, the rate of population aging is accelerating, and the pressure on old-age security is increasing day by day. In order to cope with this social challenge, commercial endowment insurance, as an important supplement to personal endowment security, has gradually attracted widespread attention. Compared with the traditional pension model, commercial pension insurance can provide individuals with more flexible and personalized protection options, and has become an indispensable part of the aging society. However, despite the obvious advantages of commercial pension insurance, there are still a large number of residents who have not fully participated in it, especially in low-income groups and in some areas [2]

In recent years, with the rapid development of digital technology, digital inclusive finance, as an innovative financial service model, provides a new opportunity for individuals to participate in commercial pension insurance. Digital inclusive finance has lowered the threshold for financial services through Internet technology, enabling more residents, especially those in remote areas, to obtain financial products at a lower cost and participate in commercial pension insurance [1]. This emerging financial

model not only increases the penetration of financial services, but also improves the financial availability of residents, thus providing them with more pension security options[3]. However, although digital inclusive finance has brought a positive impact on the promotion of commercial pension insurance, the actual effect is constrained by many factors [4]

In this context, financial literacy refers to an individual's ability to understand and apply financial products, which directly affects whether residents can correctly understand and choose commercial pension insurance products suitable for themselves. Individuals with high financial literacy are often better able to make rational financial decisions, so they are more willing to participate in commercial pension insurance. At the same time, the degree of financial development, including the maturity of the financial market and the popularity of financial services, will also affect residents' participation behavior in the context of digital financial inclusion.

The purpose of this paper is to explore how digital inclusive finance can promote residents' willingness to participate in commercial pension insurance by improving financial literacy and promoting financial development. Through empirical analysis, this paper will deeply reveal the mechanism of digital financial inclusion, financial literacy and financial development in promoting residents' participation in commercial pension insurance, and explore the influence of multiple factors such as individual characteristics, family status and regional economy on this process. The study will also examine the differences in the participation of different groups in commercial pension insurance, especially in terms of the impact of digital financial inclusion in different regions, income levels, frequency of Internet use, and other socio-economic factors.

II. LITERATURE REVIEW AND RESEARCH HYPOTHESIS

2.1 Research on digital financial inclusion:

As an emerging financial innovation, digital inclusive finance is profoundly changing the financial behavior of Chinese residents. The research shows that digital inclusive finance significantly improves the possibility of residents participating in various financial products, including commercial pension insurance, by providing convenient and low-cost financial services. In the existing literature, digital financial inclusion is often regarded as the core explanatory variable, and its impact on financial participation and insurance participation rate has been widely explored.

Yu et al. [5] (2022) pointed out that digital inclusive finance has significantly promoted the transformation of rural residents' consumption structure, enabling them to more easily access and purchase various financial products, including commercial pension insurance. Similarly, Li et al. [6] (2022) also found that digital financial inclusion enhances residents' financial access capacity, especially among economically weaker groups, and further promotes their possibility of participating in commercial pension insurance. Li et al. [7] (2020) discussed how digital inclusive finance can improve residents' awareness and acceptance of products such as pension insurance by simplifying financial processes and access to information.

Financial literacy as a mediating variable has also received attention in several studies. Hasan et al. [8] (2021) pointed out that financial literacy indirectly affects residents' acceptance of long-term security products such as commercial pension insurance by enhancing their awareness of investment and savings products. In particular, the popularization of digital inclusive finance has improved residents' financial literacy, which in turn has affected their insurance purchase behavior. Hasan et al. [8] (2021) and Yi et al. [9] (2024) both showed that financial literacy can enhance residents' understanding of financial products and make them more willing to participate in commercial pension insurance.

2.2 Research on the participation rate of commercial pension insurance:

The participation rate of commercial pension insurance is affected by a variety of factors, involving multiple levels such as individuals, families and regions. The study shows that factors at the individual level, such as age, marital status, education level, participation in basic pension insurance and medical insurance, as well as economic status and demographic structure at the family level, have an important impact on whether residents participate in commercial pension insurance. In addition,

external environmental factors such as price levels and policy support at the regional level also affect residents' participation in pension insurance decision-making to a certain extent.

At the individual level, age and gender were the control variables that were widely studied. Age and gender have been selected in several studies as factors influencing participation rates in commercial pension insurance [10][11]. In addition, marital status was also included as one of the influencing factors, and there were differences in the participation rate between married and unmarried residents [11]. Education level is also generally regarded as an individual-level variable in the study, and it is believed that people with higher education level have different attitudes towards commercial pension insurance [12]. The participation of basic pension insurance is also a control variable in many literatures, and researchers believe that residents who participate in basic pension insurance may have a lower demand for commercial pension insurance [13].

At the household level, garage ownership is widely used to analyze households' participation behavior as an indicator of economic status [14]. In addition, the number of family members and the number of children are also commonly studied variables, and it has been found that the number of family members may influence the intention to participate [15]. These factors reflect the family's financial situation and needs.

Among the control variables at the regional level, the price level (CPI) has been mentioned many times, and the price level affects residents' spending ability and willingness to save, so it has become an important factor influencing the decision to participate in insurance [16]. Another common regional-level variable is local government policy support, such as the proportion of fiscal expenditure in each province, and studies have shown that the level of local government fiscal expenditure is closely related to the likelihood of residents participating in commercial pension insurance [17].

To sum up, the decision-making of residents' participation in commercial endowment insurance is the result of the joint influence of multiple factors. Individuals' age, marital status, educational background, participation in basic security, family wealth status and member composition, as well as regional price levels and government policy support, all affect the participation rate of commercial pension insurance to varying degrees. Understanding the impact of these control variables on insurance decision-making will help to better design targeted policies and promote the healthy development of the commercial pension insurance market.

2.3 Research on the participation rate of digital inclusive finance and commercial pension insurance:

The impact of digital inclusive finance on the participation rate of commercial pension insurance has become a research hotspot in recent years. Wang et al. [18] (2020) argue that digital inclusive finance effectively promotes the popularization of social security and increases the participation rate of commercial pension insurance by improving the availability of financial services, especially for low-income groups. Liu Dongjiao and Zhuang Pengtao [19] (2021) found that digital inclusive finance significantly increased households' willingness to purchase commercial insurance by simplifying the insurance application process and reducing transaction costs. Li et al. [20] (2021) further pointed out that the popularity of digital finance has enhanced consumers' trust in insurance products, especially in economically underdeveloped regions. Wang Renzeng and Huang Xiaoying [21] (2021) showed that the convenient online services and information transparency provided by digital inclusive finance have further reduced the complexity of insurance products and promoted the participation of commercial pension insurance. Xie et al. [22] (2024) analyzed that digital inclusive finance promotes the participation of low-income families in pension insurance by reducing information asymmetry.

International research has also confirmed the positive impact of digital financial inclusion. Zhang and Tang [23] (2022) found that digital financial inclusion not only drives the health insurance market, but also has a positive impact on participation in the pension insurance market [8]. Hasan et al. [24] (2022) highlighted that improving financial inclusion through digital financial services can enable more uninsured groups to access financial services and further promote the popularization of pension insurance. Hou et al. [25] (2024) showed that the differences in digital financial inclusion in different regions and groups have contributed to the popularization of commercial insurance, especially in countries with large regional differences.

2.4 Research hypotheses:

With the continuous progress of digital technology, digital inclusive finance has become a key direction for the development of financial services in China. Through technologies such as the Internet, big data, and artificial intelligence, digital inclusive finance has lowered the threshold for financial services and provided more convenient and low-cost financial products and services [8]. Previous studies have shown that digital financial inclusion not only promotes financial inclusion in rural areas, but also improves the availability of financial services for low-income groups in urban areas, especially in areas and populations that are difficult to be covered by traditional financial services [9][26].

In this context, digital inclusive finance is expected to increase the participation rate of China's individual commercial pension insurance. By providing low-barrier financial products, digital inclusive finance enables residents to access financial services such as commercial pension insurance at a lower cost. At the same time, the optimization of information dissemination channels and the popularization of financial literacy have further reduced the barriers to participation [6]. Therefore, the development of digital inclusive finance is expected to significantly increase the participation rate of commercial pension insurance. In summary:

Hypothesis H1: Digital inclusive finance can promote the participation rate of commercial pension insurance in China.

Previous studies have shown that the level of financial literacy directly affects residents' cognition and use of financial products, and then affects their enthusiasm to participate in the financial market [8]. Especially in China, financial literacy is generally low, and there are still great obstacles for residents to understand complex financial products, especially commercial pension insurance [6]. With the development of the financial market, improving financial literacy has become an important way to improve residents' participation in the financial market.

Digital financial inclusion can help improve residents' financial literacy by providing financial education resources, information popularization, and online consultation [26]. Through the digital platform, residents can not only easily access relevant financial knowledge, but also better understand the features and advantages of various financial products. In addition, the digital financial inclusion platform can also guide residents to participate in financial products that meet their needs through precise recommendations[5]. Therefore, digital inclusive finance can not only improve the financial literacy of residents, especially by increasing the number of investment products, enhance their tendency to participate in the financial market, and further promote the purchase of commercial pension insurance. In summary:

Hypothesis H2: Digital inclusive finance can improve residents' financial literacy and enhance their tendency to participate in the financial market, thereby promoting residents' purchase of commercial pension insurance.

Financial development plays a key intermediary role between digital inclusive finance and the participation rate of commercial pension insurance. Specifically, the popularization of digital inclusive finance has promoted financial development, thereby promoting the innovation of financial products and the improvement of market efficiency. Financial development is measured by the added value of the financial industry, which improves the accessibility of financial services and enhances consumer trust in participation, especially in the field of pension insurance, and promotes wider participation. Therefore, financial development has become a bridge for digital inclusive finance to affect the participation rate of commercial pension insurance by improving the quality and popularity of financial services.

The added value of the financial sector is an important indicator to measure the contribution of the financial sector to the economy, reflecting the gross domestic product of financial services and their role in resource allocation and economic growth. With the development of digital financial inclusion, financial services have become more accessible through the Internet and mobile technology, especially in low-income groups and rural areas. The promotion of digital inclusive finance not only enhances the added value of the financial industry, but also promotes the innovation and popularization of financial products, provides convenient financial services for more people, and creates favorable conditions for the participation of long-term financial products such as commercial pension insurance. In summary:

Hypothesis H3: As an indicator of financial development, the added value of the financial industry plays an intermediary role between digital inclusive finance and the participation rate of commercial pension insurance.

III. RESEARCH DESIGN

3.1 Data sources:

The data used in this paper is mainly composed of three parts: First, the data of commercial pension insurance comes from the survey data of the Chinese general social survey (CGSS) project in 2021, which is the earliest micro survey project based on comprehensive and continuous nationwide in China, CGSS has collected data since 2003, covering multiple dimensions including society, community, family and individual. It is of great scientific and practical significance. Second, the data on digital financial inclusion comes from a joint research group composed of the Digital Finance Research Center of Peking University and the Ant Technology Group Research Institute, which is used to measure the development of digital financial inclusion. Third, the data at the provincial level come from the provincial statistical yearbooks, which ensures the authenticity, integrity and representativeness of the data. In this paper, the data of the above three aspects are matched and combined, and a total of 5810 samples are involved after excluding outliers and missing values.

3.2 Variable selection:

3.2.1 Explanatory variables:

Commercial pension insurance participation rate (CEI), this paper uses the answers of "whether you currently participate in the following social security programs - commercial pension insurance" in the 2021 CGSS questionnaire to measure the commercial pension insurance participation rate, and assigns the explanatory variable to 0 or 1 according to the respondents' choice from the options "participated, not participated, not applicable, do not know, refuse to answer", in which no participation, not applicable, do not know, refuse to answer 0, participation is 1.

3.2.2 Core explanatory variables:

Digital Financial Inclusion Index (Ln_DFI). The logarithm of the Peking University Digital Inclusion Index, compiled by a joint research group composed of the Peking University Digital Finance Research Center and the Ant Technology Group Research Institute, is used to measure the development of digital financial inclusion. In addition, this paper also uses Breadth, one of the sub-dimensions of digital inclusive finance, to further explore the impact of different dimensions of digital inclusive finance on the participation rate of individual commercial pension insurance, and the coverage breadth is selected as a substitute variable for the robustness analysis of the model.

3.2.3 Intermediary variables:

Financial literacy (Totle_Investment). Financial literacy is selected as the mediating variable, and financial literacy is measured by the number of investments involved (e.g., stocks, bonds, real estate, etc.). The number of investments can reflect the knowledge and behavioral ability of residents in financial decision-making, and then reflect their ability to understand and use different financial products.

Financial Development (Ln_Finance). With the continuous development of the financial industry, the popularity and quality of financial services have been significantly improved, which makes it easier for residents to access and use digital inclusive financial products. By improving the accessibility and diversity of financial services, financial development has effectively promoted residents' participation in commercial pension insurance.

3.2.4 Control variables:

The control variables in this paper mainly include three levels: individual, family and region, including age (Ln_Age), marital status (Marriage), basic pension insurance (Basic_pesion) and medical insurance (Medical_pention), education level (Edu), household registration status (Accounts), Political, and Internet usage; At the household level, it includes car and house ownership, children and total family size. The regional level, including the price level, is measured as the logarithm (Ln_CPI) of the consumer price index of each province, and policy support (Gov) is measured as the logarithm of the general public budget expenditure of each provincial government as a share of the province's GDP. The definitions, symbols, and descriptive results of each variable are shown in Table 1.

TABLE 1
VARIABLE DEFINITIONS AND DESCRIPTIVE STATISTICS

Variable		Valid sample size	Mean	Variance	Min	Max	Variable definitions
Explanatory variables	Commercial_Pension	5,810	0.078	0.269	0	1	Whether or not to participate in commercial pension insurance (yes=1, no=0)
Core explanatory variables	Ln_DFI	5,810	3.858	0.311	3.3	4.45	The logarithm of the Master Index for Digital Financial Inclusion
	Ln_DFI20	5,810	5.867	0.089	5.7	6.04	
	Ln_DFI19	5,810	5.814	0.092	5.64	5.99	
	Ln_DFI18	5,810	5.736	0.091	5.57	5.91	
Mediation variables	Total_Investment	5,810	0.137	0.433	0	2	Number of investment types purchased (stocks,funds,bonds,futures,warrants,property speculation,foreign exchange investments)
	Ln_Finance	5,810	8.051	0.683	5.62	9.11	The logarithm of value added in the financial sector
Control variables	Children	5,810	0.884	0.768	0	3	Number of children
	Gender	5,810	0.505	0.5	0	1	Gender (0=Female 1=Male)
	Ln_Age	5,810	3.84	0.367	3	4	The logarithm of age
	Edu	5,810	1.785	1.586	0	4	Highest education (0=elementary school and below 1=junior high school, high school, technical secondary school and technical school 2=college degree or above)
	Health	5,810	3.494	1.061	1	5	Health (1=Very Good 2=Very Good 3=Good 4=Fair 5=Poor)
	Marriage	5,810	2.437	1.054	0	3	Marital status (0=cohabiting/unmarried, 1=divorced, 2=widowed, 3=having a legal spouse)
	Basic_Pension	5,810	0.765	0.424	0	1	Participation in basic pension insurance (0=no 1=yes)
	Medical_Pesion	5,810	0.947	0.224	0	1	Whether or not to participate in basic medical insurance (0=no 1=yes)
	Gov	5,810	2.899	0.271	2.35	4	The logarithm of the general public budget expenditure of each provincial government as a percentage of the GDP of the province
	Size	5,810	3.317	1.736	1	9	Total household population
	Property	5,810	1.26	0.748	0	5	Number of properties
	Car	5,810	0.429	0.495	0	1	Car ownership (0=no 1 = yes)
	Ln_CPI	5,810	4.615	0.004	4.61	4.62	The logarithm of the consumer price index
	Accounts	5,810	1.437	0.52	0	2	Hukou registration status (0=Other, 1=Rural 2=Urban)
	Political	5,810	0.146	0.354	0	1	Political outlook (0=non-party member 1=party member)
Internation	5810	1.731	0.443	1	2	Frequent internet use (1=No 2=Yes)	

3.3 Model setting:

3.3.1 Benchmark model:

In order to test the impact of digital inclusive finance on the participation of commercial pension insurance, since the questionnaire of the relevant dependent variable is set to "whether you currently participate in the following social security programs - commercial pension insurance", which is a binary variable, it is necessary to use a probabilistic model for fitting and estimating, and the benchmark model is set as follows:

$$CEI_{i,j} = \alpha_i + \beta_1 \ln_DFI_j + \gamma_i X_{i,j} + \varepsilon_{i,j} \quad (1)$$

In Eq. (1), the explanatory variable $CEI_{i,j}$ indicates whether the individual i in province j participates in commercial pension insurance, the explanatory variable \ln_DFI_j indicates the degree of digital inclusive financial development of province j , and $X_{i,j}$ represents the control variables at the individual, family and regional levels: at the individual level, including age, marital status, participation in basic pension insurance and medical insurance, education level, political outlook, and household registration status; At the household level, it includes garage ownership, individual children and the total family population; At the regional level, it includes the price level and related policy support (the proportion of the general public budget expenditure of each provincial government in the GDP of the province), which ε is a random error term, and α_i, β_1 and γ_i are the parameters to be estimated.

3.3.2 Mediation effect model:

In order to test the mediating effect of the benchmark regression model, this paper draws on the method proposed by Hasan et al. [8] to construct the model with financial literacy (whether or not to engage in investment activities) as an intermediate variable. Therefore, the following intermediary effect model is set up to test the impact mechanism of digital inclusive finance on the participation rate of commercial pension insurance:

$$CEI_{i,j} = \alpha_1 + \beta_1 \ln_DFI_j + \gamma_1 X_{i,j} + \varepsilon_{i,j} \quad (2)$$

$$M_{i,j} = \alpha_2 + \beta_2 \ln_DFI_j + \gamma_2 X_{i,j} + \varepsilon_{i,j} \quad (3)$$

$$CEI_{i,j} = \alpha_3 + \beta_3 \ln_DFI_j + \vartheta M_{i,j} + \gamma_3 X_{i,j} + \varepsilon_{i,j} \quad (4)$$

In equations (2), (3) and (4), $M_{i,j}$ is the mediating variable, which represents the financial literacy of the i th individual located in province J , $CEI_{i,j}$ represents the participation of individual i in the commercial pension insurance located in province J , the explanatory variable DFI_j represents the degree of digital inclusive financial development of province J , $X_{i,j}$ is the control variable, the ε is the random error term, and the β_1 value represents the total utility of the digital inclusive finance index level on the participation rate of commercial pension insurance, $\beta_2 * \vartheta$ represents the mediating effect of financial literacy on the participation rate of commercial pension insurance, and $\frac{\beta_2 * \vartheta}{\beta_1}$ represents the proportion of the mediating effect of financial literacy on the participation rate of commercial pension insurance.

IV. EMPIRICAL RESULTS AND ANALYSIS

4.1 Benchmark regression results:

First, the benchmark model was analyzed, and the results are shown in Table 2.

From the results of the benchmark regression model in Table 2, it can be seen that the digital inclusive finance index has a significant positive effect on the participation rate of commercial pension insurance, so the hypothesis H1 is verified. The coefficients of health status, participation in basic pension insurance, number of real estate ownership, car ownership and policy support were significantly positive at the 1% significance level, gender was significantly negative at the 5% significance level, and the coefficient of basic medical insurance purchase and the number of children was negative at the 1% significance level.

In summary, individuals with healthier physical conditions are more inclined to purchase commercial endowment insurance, which may be related to their strong sense of self-protection and risk management ability. Healthy individuals often expect to live longer and have a higher quality of life, so they tend to plan financially ahead of time to ensure that their quality of life in retirement is not compromised. In addition, healthy individuals often enjoy lower premiums when purchasing insurance, which also incentivizes them to accumulate wealth through commercial pension insurance. In summary, the behavior of healthy

individuals purchasing commercial endowment insurance reflects their active response to future uncertainty and the importance they attach to financial security.

TABLE 2
BENCHMARK MODEL REGRESSION RESULTS

Variable	(1)	(2)	(3)
	Commercial_Pension	Commercial_Pension	Commercial_Pension
Ln_DFI	0.077***	0.074***	0.087***
	-0.011	-0.011	-0.016
Gender		-0.012*	-0.014**
		-0.007	-0.007
Ln_Age		0.005	0.017
		-0.011	-0.012
Edu		0.002	0
		-0.002	-0.002
Health		0.020***	0.016***
		-0.003	-0.003
Marriage		-0.001	0
		-0.004	-0.004
Basic_Pension			0.036***
			-0.009
Medical_Pension			-0.043***
			-0.016
Size			-0.003
			-0.002
Property			0.014***
			-0.005
Car			0.046***
			-0.008
Children			-0.015***
			-0.005
Ln_CPI			0.687
			-0.95
Gov			0.063***
			-0.017
_cons	-0.219***	-0.290***	-3.721
	-0.044	-0.062	-4.357
N	5810	5810	5810
R ²	0.008	0.014	0.033

Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.

Individuals who participate in basic social pension insurance are more likely to purchase commercial pension insurance, which reflects the need for multi-level protection. The limited nature of basic endowment insurance prompts them to seek commercial insurance to fill the protection gap, and at the same time, their strong risk aversion awareness and financial planning ability make them inclined to achieve long-term economic security and asset appreciation through commercial insurance.

In contrast, individuals participating in basic medical social insurance have a lower demand for commercial pension insurance. Due to the unpredictability of disease and the priority of medical insurance, medical insurance expenditure often replaces the investment of commercial pension insurance, thereby reducing the willingness to participate in insurance.

Women are more likely to buy commercial pension insurance, mainly because they have a longer life expectancy and pay more attention to long-term financial security. At the same time, women take on more family care responsibilities and have relatively low or unstable incomes, making them more reliant on commercial insurance to ensure financial security in retirement. In addition, the socio-cultural and policy environment may also encourage women to choose insurance as a risk management tool.

Individuals with more real estate tend to buy commercial pension insurance because they have more economic capital and risk aversion needs. As a symbol of wealth, real estate not only enhances purchasing power, but also reflects the focus on economic security and asset preservation and appreciation after retirement.

Individuals who own vehicles also show a higher propensity to purchase commercial pension insurance, reflecting their stronger economic strength and financial planning capabilities. As the embodiment of economic capital, the vehicle is consistent with a higher sense of risk aversion, prompting it to protect economic security after retirement through insurance.

On the contrary, families with more children are more dependent on family pension, reducing the demand for commercial pension insurance. At the same time, the increase in the number of children has led to the dispersion of resources and reduced investment in insurance products.

In addition, the increase in the proportion of general public budget expenditure of the provincial government in GDP (Gov) is positively correlated with the tendency of residents to purchase commercial pension insurance. Stronger policy support has boosted residents' confidence in future pension security and encouraged them to actively participate in commercial insurance schemes.

4.2 Mediation effect test:

4.2.1 Analysis of the mediating effect of financial literacy:

In order to further test the promotion effect of the digital inclusive financial index on the participation rate of commercial pension insurance in each province, this paper tests the mediating effect based on the results of the previous paper. The results of the mediating effect of financial literacy are shown in Table 3.

According to the results in Table 3, the β_1 value, θ value and β_3 value all reached significant levels, indicating that there was a significant mediating effect. At the same time, the digital inclusive finance index has a significant positive impact on the participation rate of commercial pension insurance, with a total effect value of 0.087, which further supports the positive role of digital inclusive finance in promoting the participation rate of commercial pension insurance.

TABLE 3
RESULTS OF THE MEDIATING EFFECT OF FINANCIAL LITERACY

Variable	(1)	(2)	(3)
	Commercial_Pension	Total_Investment	Commercial_Pension
Ln_DFI	0.087***	0.450***	0.045***
	-0.016	-0.024	-0.016
Gender	-0.014**	0.022**	-0.016**
	-0.007	-0.011	-0.007
Ln_Age	0.017	-0.102***	0.027**
	-0.012	-0.018	-0.012
Edu	0	-0.004	0
	-0.002	-0.003	-0.002
Health	0.016***	0.011**	0.015***
	-0.003	-0.005	-0.003
Marriage	0	-0.009	0.001
	-0.004	-0.006	-0.004
Basic_Pension	0.036***	0.044***	0.031***
	-0.009	-0.013	-0.009
Medical_Pension	-0.043***	-0.011	-0.041***
	-0.016	-0.024	-0.016
Size	-0.003	-0.015***	-0.002
	-0.002	-0.003	-0.002
Property	0.014***	0.053***	0.009*
	-0.005	-0.007	-0.005
Car	0.046***	0.102***	0.037***
	-0.008	-0.012	-0.008
Children	-0.015***	-0.039***	-0.012**
	-0.005	-0.008	-0.005
Ln_CPI	0.687	-3.554**	1.018
	-0.95	-1.445	-0.94
Gov	0.063***	0.197***	0.045***
	-0.017	-0.026	-0.017
Total_Investment			0.093***
			-0.009
_cons	-3.721	14.548**	-5.078
	-4.357	-6.629	-4.315
N	5810	5810	5810
R ²	0.033	0.139	0.053

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

TABLE 4
RESULTS OF THE SOBEL TEST OF FINANCIAL LITERACY

	Mediator effect size	Standard error	z
Sobel	0.042***	0.004	9.44
Aroian	0.042***	0.004	9.43
Goodman	0.042***	0.004	9.45

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

According to the analysis provided in Table 4, three different methods (Sobel, Aroian, Goodman) were used to test the significance of the mediating effect, and the results showed that the mediating effect estimates of the three methods were 0.042, the standard error was 0.004, and the z-statistic was 9.439, 9.430, and 9.449, respectively, and the corresponding P values were 0.000, which was significantly less than 0.05. This suggests that the mediating effect is significant.

TABLE 5
INDIRECT, DIRECT AND TOTAL EFFECTS (FINANCIAL LITERACY)

	Coefficient	Standard error	z
a_coefficient	0.450***	0.024	18.81
b_coefficient	0.093***	0.009	10.91
Indirect_effect_aXb	0.042***	0.004	9.439
Direct_effect_c'	0.045***	0.016	2.784
Total_effect_c	0.087***	0.016	5.508

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

According to the results in Table 5, the indirect effect was 0.042, the direct effect was 0.045, and the total effect was 0.087. The P values for all effects were less than 0.05, indicating that these effects were statistically significant. At the same time, financial literacy has a significant role in promoting the participation rate of commercial pension insurance, and the intermediary effect accounts for 4.2%. Specifically, financial literacy plays a significant mediating role in the impact of the participation rate of commercial pension insurance, and about 48.5% of the effect is transmitted through the mediating variable of financial literacy. Therefore, it can be concluded that the mediating variable plays a key role between the independent variable and the dependent variable, and a considerable part of this effect is transmitted through financial literacy, so the hypothesis H2 is validated.

4.2.2 Analysis of the intermediary effect of financial development:

In order to further test the promotion effect of the digital inclusive financial index on the participation rate of commercial pension insurance in each province, this paper tests the mediating effect based on the results of the previous paper. The results of the intermediary effect of financial development (added value of the financial sector) are shown in Table 6.

TABLE 6
RESULTS OF THE INTERMEDIARY EFFECT OF FINANCIAL DEVELOPMENT

Variable	(1)	(2)	(3)
	Commercial Pension	ln_Finance1	Commercial Pension
Ln_DFI	0.087***	1.208***	0.056***
	-0.016	-0.018	-0.021
Gender	-0.014**	-0.019**	-0.013*
	-0.007	-0.008	-0.007
Ln_Age	0.017	0.053***	0.016
	-0.012	-0.014	-0.012
Edu	0	0.006**	0
	-0.002	-0.003	-0.002
Health	0.016***	-0.002	0.016***
	-0.003	-0.004	-0.003
Marriage	0	-0.012***	0
	-0.004	-0.005	-0.004
Basic_Pension	0.036***	0.01	0.035***
	-0.009	-0.01	-0.009
Medical_Pension	-0.043***	0.013	-0.043***
	-0.016	-0.019	-0.016
Size	-0.003	-0.003	-0.003
	-0.002	-0.002	-0.002
Property	0.014***	0.016***	0.014***
	-0.005	-0.006	-0.005
Car	0.046***	0.008	0.046***
	-0.008	-0.009	-0.008
Children	-0.015***	-0.023***	-0.015***
	-0.005	-0.006	-0.005
Ln_CPI	0.687	-3.766***	0.781
	-0.95	-1.11	-0.95
Gov	0.063***	-1.095***	0.091***
	-0.017	-0.02	-0.021
ln_Finance1			0.025**
			-0.011
_cons	-3.721	23.761***	-4.319
	-4.357	-5.093	-4.363
N	5810	5810	5810
R ²	0.033	0.796	0.034

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

From the analysis in Table 6, it can be seen that the β_1 value, θ value, and β_3 value are all significant, indicating that the mediating effect exists and is significant. At the same time, the digital inclusive finance index also showed a significant positive impact on the participation rate of commercial pension insurance, with a total effect size of 0.087. This shows that digital inclusive finance has a positive role in promoting the participation of commercial pension insurance.

TABLE 7
RESULTS OF THE SOBEL TEST FOR FINANCIAL DEVELOPMENT

	Mediator effect size	Standard error	z
Sobel	0.030**	0.014	2.24
Aroian	0.030**	0.014	2.24
Goodman	0.030**	0.014	2.24

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

According to the results of Table 7, the three methods (Sobel, Aroian, Goodman) were used to test the significance of the mediating effect, and the estimated values of the mediating effect of the three methods were 0.030, the standard error was 0.014, and the z-statistic was 2.239, 2.238 and 2.239, respectively, and the corresponding P values were all 0.025, which was significantly lower than 0.05, which further proved the significance of the mediating effect.

TABLE 8
INDIRECT, DIRECT AND TOTAL EFFECTS (FINANCIAL DEVELOPMENT)

	Coefficient	Standard error	z
a_coefficient	1.208***	0.018	65.7
b_coefficient	0.025**	0.011	2.24
Indirect_effect_aXb	0.030**	0.014	2.239
Direct_effect_c'	0.056***	0.021	2.708
Total_effect_c	0.087***	0.016	5.508

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

As can be seen from Table 8, the indirect effect is 0.030, the direct effect is 0.056, and the total effect is 0.087. The P values for all effects were less than 0.05, indicating that they were statistically significant. At the same time, financial development (added value of the financial industry) has a certain role in promoting the participation rate of commercial pension insurance, of which the indirect effect accounts for 3%. Specifically, the added value of the financial industry plays a significant mediating role in the impact of the participation rate of commercial pension insurance, and about 35.1% of the effect is transmitted through the intermediary variable of the added value of the financial industry. Therefore, it can be concluded that the value added of the financial industry plays an important mediating role between the independent variable and the dependent variable, and about one-third of the effect is transmitted through this intermediary, so the hypothesis H3 is validated.

4.3 Robustness test:

4.3.1 Replace the explanatory variable:

In order to alleviate the influence of endogeneity on the estimation results, this paper uses the digital inclusive financial index with a lag of one period, the next two periods, and the third lag period as the core explanatory variables for regression, in order to reduce the estimation bias caused by reverse causality and prove that the development of digital inclusive finance has a long-term impact on the purchase of commercial pension insurance.

Since the development of digital inclusive finance can affect the purchase of commercial pension insurance for a long time, but the purchase behavior of commercial pension insurance may only affect the digital inclusive financial index of the current year, this paper replaces the core explanatory variables with the total index of digital inclusive finance development in 2020, 2019 and 2018 respectively to solve the possible reverse causality and prove the long-term impact of the development of digital inclusive finance on the purchase of commercial pension insurance, and the regression results are shown in Table 9.

TABLE 9
REPLACE THE REGRESSION RESULTS OF THE CORE EXPLANATORY VARIABLES

Variable	(1)	(2)	(3)
	Commercial_Pension	Commercial_Pension	Commercial_Pension
Ln_DFI20	0.312***		
	-0.056		
Ln_DFI19		0.287***	
		-0.053	
Ln_DFI18			0.290***
			-0.052
Gender	-0.014**	-0.014**	-0.014**
	-0.007	-0.007	-0.007
Ln_Age	0.017	0.017	0.017
	-0.012	-0.012	-0.012
Edu	0	0	0
	-0.002	-0.002	-0.002
Health	0.016***	0.016***	0.016***
	-0.003	-0.003	-0.003
Marriage	0	0	0
	-0.004	-0.004	-0.004
Basic_Pension	0.036***	0.036***	0.036***
	-0.009	-0.009	-0.009
Medical_Pension	-0.043***	-0.042***	-0.042***
	-0.016	-0.016	-0.016
Size	-0.003	-0.003	-0.003
	-0.002	-0.002	-0.002
Property	0.014***	0.014***	0.014***
	-0.005	-0.005	-0.005
Car	0.046***	0.046***	0.046***
	-0.008	-0.008	-0.008
Children	-0.015***	-0.015***	-0.015***
	-0.005	-0.005	-0.005
Ln_CPI	0.809	0.965	1.073
	-0.94	-0.934	-0.928
Gov	0.067***	0.063***	0.062***
	-0.017	-0.017	-0.017
_cons	-5.792	-6.339	-6.826
	-4.244	-4.229	-4.207
N	5810	5810	5810
R ²	0.034	0.033	0.033

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

The results in Table 9 show that the lag period of the total index of digital inclusive finance has a significant impact on the participation rate of commercial pension insurance, indicating that the total index of digital inclusive finance has a significant positive impact on the participation rate of commercial pension insurance. $0.312 > 0.290 > 0.287$, the 2019 Digital Inclusive Finance Development Index is slightly lower than that in 2018, which may be affected by the early stage of the new crown epidemic, financial activities are restricted, and the financial demand of enterprises and individuals has weakened. At the same time, the epidemic has led to a slowdown in the construction of digital financial infrastructure and the implementation of inclusive financial policies in some regions, affecting the overall development. Among them, the coefficient of the total index of digital inclusive finance development in 2020 is 0.312, indicating that if the total development index increases by 1% in 2020, the probability of individuals purchasing commercial pension insurance in 2021 will increase by 31.2 percentage points, but factors such as the impact of the epidemic, the constraints of economic conditions and the low penetration rate of the pension insurance market limit the actual play of this effect. Therefore, the increase in the probability of purchasing commercial endowment insurance does not fully meet the theoretical expectations.

4.3.2 Replace the model:

In order to verify the robustness of the regression results of the linear regression model, this paper refers to the study of Peng Weizhujia [29] (2020), and uses another model, the Probit model and the Logit model, to perform regression analysis on the same sample. By comparing the regression results of models based on different distribution assumptions (linear regression vs. Probit and Logit regression), the consistency of the models is tested to ensure the reliability and robustness of the research conclusions. The regression results of the Probit and Logit models are shown in Table 10, and it can be found that the regression results of Probit and Logit are consistent with the regression results of the linear regression model, and the results show that the total index of digital inclusive financial development has a significant positive effect on the participation rate of commercial pension insurance, which indicates that the benchmark model is relatively stable.

TABLE 10
LOGIT AND PROBIT REGRESSION RESULTS OF THE BENCHMARK MODEL

Variable	Logit	Probit
	Commercial_Pension	Commercial_Pension
Ln_DFI	0.943***	0.477***
	-0.194	-0.101
Control variables	YES	YES
Sample size	5810	5810
Pseudo R2	0.0607	0.0599

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

4.4 Heterogeneity analysis:

The previous empirical results show that the digital financial inclusion index and the digital financial inclusion index have a significant promoting effect on the participation rate of commercial pension insurance. However, due to the differences in the characteristics of different individuals and families, there may be heterogeneity in the facilitation effect. In order to explore the heterogeneity of the impact of the Digital Financial Inclusion Index on the participation of commercial pension insurance, this paper conducts group regression of the samples according to the individual's political outlook, urban and rural conditions, region, and whether they often use the Internet. The results are shown in Table 11, Table 12 and Table 13.

TABLE 11
THE HETEROGENEITY OF THE TOTAL INDEX OF DIGITAL FINANCIAL INCLUSION ON THE PARTICIPATION OF COMMERCIAL PENSION INSURANCE

Variable	Non-party members	Party	Rarely use the internet	Use the internet a lot
	Commercial_Pension	Commercial_Pension	Commercial_Pension	Commercial_Pension
Ln_DFI	0.098***	0.028	0.011	0.091***
	-0.017	-0.041	-0.021	-0.02
Control variables	YES	YES	YES	YES
N	4959	851	1934	3876
R ²	0.037	0.026	0.012	0.035

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

TABLE 12
HETEROGENEITY OF THE DIGITAL INCLUSIVE FINANCE GENERAL INDEX ON THE PARTICIPATION OF COMMERCIAL PENSION INSURANCE

Variable	Other	Agricultural	Urban
	Commercial_Pension	Commercial_Pension	Commercial_Pension
Ln_DFI	-0.057	0.043*	0.085***
	-0.16	-0.025	-0.023
Control variables	YES	YES	YES
N	70	3130	2610
R ²	0.126	0.022	0.038

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

TABLE 13
HETEROGENEITY OF THE GENERAL INDEX OF DIGITAL INCLUSIVE FINANCE ON THE PARTICIPATION OF COMMERCIAL PENSION INSURANCE III

Variable	Eastern	Central	Western
	Commercial_Pension	Commercial_Pension	Commercial_Pension
Ln_DFI	0.075***	-0.027	0.173
	-0.023	-0.068	-0.121
Control variables	YES	YES	YES
N	2228	1716	1556
R ²	0.035	0.033	0.025

*Note: ***, **, and * represent the significance levels of 1%, 5%, and 10%, respectively, and the robustness standard error is in parentheses.*

The results show that the promotion effect of digital inclusive finance on the decision-making of commercial pension insurance exists in the samples of non-party members, the samples of agricultural residents, the samples of urban residents and the samples of eastern regions, and the probability of non-party members purchasing commercial pension insurance increases by 9.8%, 9.1% and 4.3% for agricultural residents with each unit of digital inclusive finance index. The probability of urban residents purchasing commercial endowment insurance increased by 8.5%, and that of residents in eastern China increased by 7.5%.

Non-Party members are more sensitive and involved in digital inclusive finance, perhaps because they lack the policy support and information resources of Party organizations, and rely more on the popularization and innovative products of digital inclusive finance. On the contrary, the financial inclusion index of the party member group is not significant, probably because

they have obtained sufficient social security information and resources through other channels, and the promotion of digital financial inclusion is relatively small.

Residents who regularly use the Internet are more inclined to buy commercial pension insurance, mainly because they can more easily obtain relevant information about pension insurance through the Internet platform. The internet provides users with a wealth of resources that make it easy to compare different insurance products, learn about what is covered, and how much it costs. The availability of this information has increased their interest in and willingness to purchase commercial pension insurance. In addition, the convenience of the internet allows residents to quickly purchase insurance online, saving time and effort. As a result, residents who regularly use the Internet are generally more willing to participate in commercial pension insurance because they can easily and quickly access information and make decisions.

Urban dwellers generally have better access to infrastructure and financial services than rural dwellers, and access to financial services is higher. As a result, the impact of digital financial inclusion on them is likely to be more immediate and significant, as they have greater access to digital financial services and a higher level of acceptance of these services.

The economic and financial infrastructure in the eastern region is relatively developed, and residents have a high acceptance of digital inclusive finance, so digital inclusive finance can significantly promote the participation of commercial pension insurance. On the contrary, due to the slow economic development, insufficient financial services, and low financial literacy of residents in the central and western regions, the promotion of digital inclusive finance on commercial pension insurance is relatively limited.

V. CONCLUSIONS AND POLICY RECOMMENDATIONS

Through empirical analysis, the impact of digital inclusive finance on the participation rate of individual commercial pension insurance is deeply discussed, and the mediating role of financial literacy and financial development in it is investigated. The results of the study found that, first, the health status of individuals, their participation in basic pension insurance, the number of houses and cars they owned, and policy support had a significant impact on the decision to purchase commercial pension insurance. Individuals in better health tend to purchase commercial pension insurance for long-term financial planning and risk management, while individuals in lower health tend to rely more on insurance due to their health risk sensitivity. Individuals who participate in basic pension insurance are more likely to purchase commercial pension insurance to fill the protection gap; Women, individuals with more properties and cars show higher willingness to buy due to their greater economic strength and risk aversion needs. Policy support also has a positive impact on purchasing decisions, reflecting residents' trust in the security system. In general, these factors work together to promote the willingness to purchase commercial pension insurance.

Second, financial literacy plays a significant mediating role between digital inclusive finance and the participation rate of commercial pension insurance. Specifically, digital inclusive finance enhances residents' understanding of financial products by improving their financial literacy, which in turn increases their willingness to participate in commercial pension insurance. The mediating effect of financial literacy accounted for 48.5%, reflecting its key role in promoting the participation rate of commercial pension insurance. At the same time, financial development also plays a significant intermediary role between digital inclusive finance and the participation rate of commercial pension insurance, accounting for 35.1%, indicating that the promotion of financial development on the participation rate of commercial pension insurance cannot be ignored.

Third, the role of digital inclusive finance in promoting the decision-making of commercial pension insurance is reflected in many groups, especially among non-party members, residents who often use the Internet, agricultural residents, urban residents and residents in the eastern region. Non-party members and internet-frequent residents are more sensitive to digital financial inclusion because they have easier access to information and purchase decisions through digital platforms. Urban residents are more receptive to digital inclusive finance because they enjoy better infrastructure and financial services, and the promotion effect is more obvious. In the eastern region, due to its developed economic and financial infrastructure, the influence of digital inclusive finance is particularly prominent. In contrast, due to lagging economic development, insufficient financial services and low financial literacy of residents in the central and western regions, the promotion of digital inclusive finance to commercial pension insurance is relatively weak.

Based on the conclusions of the paper, this paper puts forward the following policy suggestions: First, strengthen health management and risk awareness training, and find that the health status of individuals has a significant impact on the purchase decision of commercial pension insurance. Individuals in better health tend to use commercial pension insurance for long-term financial planning, while individuals in lower health may be more reliant on insurance due to their greater sensitivity to health risks. Therefore, it is suggested that policymakers should strengthen health management and risk awareness training, especially

for groups with poor health status, and encourage them to achieve better risk sharing and financial security through commercial pension insurance. At the same time, it can be combined with the comprehensive product design of health insurance and pension insurance to enhance residents' willingness to participate.

Second, to promote the construction of financial infrastructure in the eastern and central and western regions, considering that the eastern region has stronger acceptance and influence on digital inclusive finance, it is recommended to increase the construction of financial infrastructure in the central and western regions to improve residents' awareness and accessibility of digital financial products. This will not only narrow the gap between regions, but also promote the overall access to financial services and enhance the social inclusion of digital financial inclusion.

Third, the design and marketing of diversified financial products, due to the differences in the acceptance of digital inclusive finance by different groups, it is recommended that financial institutions consider more diversified demand characteristics when designing commercial pension insurance products, such as launching financial products with low thresholds and strong flexibility. At the same time, the marketing of financial products should be accurately launched through different channels and platforms, combined with the specific conditions of each group.

Fourth, to enhance policy support and incentives, local governments should strengthen policy support for commercial pension insurance, especially in terms of fiscal expenditure. By increasing incentives such as subsidies or tax incentives for commercial pension insurance products, residents will be able to enhance their confidence in the future pension security, encourage them to actively participate in commercial insurance plans, and improve the level of pension security for the whole society.

Through the implementation of these policy recommendations, we can effectively promote the development of digital inclusive finance and enhance residents' financial participation, thereby promoting the popularization and development of commercial pension insurance and helping more people achieve economic security and retirement security.

REFERENCES

- [1] Li R. Development status, problems, and countermeasures of digital inclusive finance in Shaanxi Province[J]. *Finance*, 2024, 14: 451.
- [2] Zhou T Y, Chen M X. Digital Penetration, Inclusive Finance and Household Wealth Growth[J]. *Journal of Finance and Economics*, 2021, 47(7): 33-47.
- [3] Zhang X, Wan G H, Zhang J J, et al. Digital economy, financial inclusion and inclusive growth[J]. *China Economist*, 2020, 15(3): 92-105.
- [4] LUSARDI A, MITCHELL O S. The economic importance of financial literacy: Theory and evidence[J]. *American Economic Journal: Journal of Economic Literature*, 2014, 52(1): 5-44.
- [5] Yu C, Jia N, Li W, et al. Digital inclusive finance and rural consumption structure—evidence from Peking University digital inclusive financial index and China household finance survey[J]. *China Agricultural Economic Review*, 2022, 14(1): 165-183.
- [6] LI K, MENG MENG H, HUO J. Digital inclusive finance and asset allocation of Chinese residents: Evidence from the China Household Finance Survey[J]. *Plos one*, 2022, 17(5): e0267055.
- [7] LI J, WU Y, XIAO J J. The impact of digital finance on household consumption: Evidence from China[J]. *Economic modelling*, 2020, 86: 317-326.
- [8] HASAN M, LE T, HOQUE A. How does financial literacy impact on inclusive finance?[J]. *Financial innovation*, 2021, 7(1): 40.
- [9] Yi S, Qi Y, Ya Y, et al. The impact of China's digital inclusive financial development gap on the optimization of rural consumption structure[J]. *Plos one*, 2024, 19(8): e0308412.
- [10] Zhang L. Empirical analysis on the factors influencing the demand of commercial endowment insurance—Data analysis based on CGSS 2021[J]. *Operations Research and Fuzziology*, 2023, 13: 6750.
- [11] QIN M, ZHUANG Y, LIU H. Old age insurance participation among rural-urban migrants in China[J]. *Demographic research*, 2015, 33: 1047-1066.
- [12] Li D, Ding J, Ma S. The effect of social interaction on household commercial insurance purchases: Evidence from the China Household Finance Survey[J]. *Financial Research*, 2019, 469(7): 96-114.
- [13] Xu B C, Xu X N, Zhao J C, et al. Influence of internet use on commercial health insurance of Chinese residents[J]. *Frontiers in Public Health*, 2022, 10: 907124.
- [14] Xie Y, Zhang B, Yao Y, et al. Mechanism of human capital influence on household commercial insurance holding behavior—An empirical analysis based on China household finance survey (CHFS) data[J]. *Frontiers in Environmental Science*, 2022, 10: 961184.
- [15] WANG Q, LIU C, LAN S. Digital literacy and financial market participation of middle-aged and elderly adults in China[J]. *Economic and Political Studies*, 2023, 11(4): 441-468.
- [16] Xu H, Zhang C, Huang Y. Social trust, social capital, and subjective well-being of rural residents: Micro-empirical evidence based on the Chinese General Social Survey (CGSS)[J]. *Humanities and Social Sciences Communications*, 2023, 10(1): 1-13.
- [17] Zhou M, Wang Y, Liang Y, et al. The effect of subjective life expectancy on the participation in commercial pension insurance of Chinese elderly[J]. *Frontiers in Psychology*, 2022, 13: 969719.

- [18] Wang Y, Tan Z, Zheng L. The impact of digital inclusive finance on social security[J]. *Journal of Quantitative & Technological*, 2020, 37(7): 92-112.
- [19] Liu D J, Zhuang P T. Digital financial inclusion and household commercial insurance purchases[J]. *Consumer Econ*, 2021, 37: 67-78.
- [20] Li X, Wu Y, Li J. Digital finance development and household commercial insurance participation[J]. *Statistical Research*, 2021, 38(05): 29-41.
- [21] Wang R Z, Huang X Y. The empirical study of the influence of digital inclusive finance on household demand for commercial insurance[J]. *Journal of Northwest Minzu University (Philosophy and Social Science Edition)*, 2021(04): 123-137.
- [22] Xie M M, Yan W J, Yang Y F. Effects of digital inclusive finance on household demand for commercial insurance: Based on mechanism analysis and threshold effect testing[J]. *Financial Theory and Practice*, 2024(05): 108-118.
- [23] ZHANG Y, TANG X, OTHERS. Research on the impact of digital inclusive finance on the income of China's commercial health insurance[J]. *Academic Journal of Business & Management*, 2022, 4(17): 72-79.
- [24] HASAN M M, YAJUAN L, KHAN S. Promoting China's inclusive finance through digital financial services[J]. *Global Business Review*, 2022, 23(4): 984-1006.
- [25] Hou, Zaikun, et al. "The Impact of Digital Financial Inclusion on Household Commercial Insurance for Sustainable Governance Mechanisms under Regional Group Differences." *Sustainability* 16.9 (2024): 3596.
- [26] PENG P, MAO H. The effect of digital financial inclusion on relative poverty among urban households: a case study on China[J]. *Social Indicators Research*, 2023, 165(2): 377-407.

Determinants of Childbearing Conceptions in Married Women: A Machine Learning Comparative Study

Li Zhu¹, Qian Long^{2*}

¹Associate Professor, School of Big Data Statistics, Guizhou University of Finance and Economics

²Master's student in the School of Big Data Statistics, Guizhou University of Finance and Economics

*Corresponding Author

Received: 05 March 2025/ Revised: 14 March 2025/ Accepted: 20 March 2025/ Published: 31-03-2025

Copyright © 2025 International Journal of Engineering Research and Science

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted Non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract— In recent years, with the continuous progress of society and the continuous change of people's way of thinking, the inclusiveness of society towards women, especially married women, is increasing, and married women have gained more rights and choices, but the fertility rate in China is declining. As a result, the study of demographics and family structure has been gaining traction, especially on the decline in the natural growth rate of the population shown in the national census data, which is also related to the overall characteristics of women, such as whether they are employed or not. This paper uses multiple regression model, decision tree, random forest, gradient boosted regression tree model, support vector machine, XGboost and other machine learning models to train and compare the data, and draws the following conclusions: (1) The comprehensive characteristics of married women such as employment, age, and mother's education level are negatively correlated with their willingness to have children; (2) The influence of married women's husband's husbands and family characteristics on the intention to have children was positively significant. (3) Based on the results of multiple regression models, it warns us that we should consider multiple aspects when doing research or model selection, and that the selected model is good but not necessarily the most consistent with the characteristics of the data; (4) Based on the comparison of various models, it is found that the gradient boosted regression tree model, K-nearest neighbor model, XGBoost model and random forest model are better than the linear regression model.

Keywords— *Comprehensive Characteristics of Married Women; Fertility Concept; Machine Learning.*

I. INTRODUCTION

In recent years, with the continuous progress of society and the profound transformation of people's thinking patterns, the tolerance and respect for women, especially married women, have been increasing. This change has given married women more choices and rights, however, China's fertility rate has shown a downward trend in this context, and changes in demographics and family structure have attracted more and more attention from all walks of life. In particular, the declining natural population growth rate revealed by the national census data is closely related to individual characteristics such as women's employment status, which is reflected in the fact that married women begin to consider whether to have children. Among them, the employment status, age, education level, and the relationship between various personal and family characteristics of married women are of great significance for understanding the dynamics of social fertility, promoting women's career development, and optimizing family policy formulation. In this context, it is crucial to use modern machine learning techniques to dig deeper into this topic. In addition, this study not only helps us to comprehensively examine and improve the social reproductive environment from the unique perspective of married women, but also has great significance for the exploration of women's career development, the change of family structure, the optimization of social labor structure, and the improvement of the overall level of China's economic and social development.

In order to comprehensively and deeply explore this problem, this paper constructs a dataset covering a variety of variables based on abundant social survey data resources, and on this basis, it is proposed to use a variety of machine learning models for simulation and comparison, including but not limited to multiple regression, decision trees, random forests, XGboost and logistic regression, and support vector machines. The optimal prediction model is found by comparing the accuracy metrics of each model. Based on the training and comparison of these machine learning models, it can not only provide a comprehensive

analysis of the influencing factors of married women's reproductive age, but also reveal the performance differences and applicability of different models in dealing with such complex problems, so as to provide a solid scientific basis and reference for the formulation and practical operation of relevant policies.

This paper combs and summarizes the existing literature, and finds that many scholars at home and abroad have a lot of relevant research on the relationship between the characteristics of married women and whether they have the youngest child, and the research scope is also very broad. Among them, Huang Qian (2022) et al. believe that when married women have the youngest child, their employment status is related to the age of the child. The presence of the youngest children of different ages has different effects on women's employment. Based on the comprehensive evaluation and analysis of CGSS survey data, Song et al. (2022) concluded that family characteristics such as parents' education level, spouse's occupation, and family economic status have a significant impact on the employment quality of married women. Yan Yu (2022) argues that the youngest child is not significant when married women are employed, but the youngest age of 0-3 years old has a significant impact on lowering the employment threshold for married women. Luo Yuan (2009) concluded that age, health status, and children's age have significant negative effects on married women's labor supply decisions. The education level factor has a significant positive effect on the labor supply decision of married women. Wang Guanghui (2010) argues that as married women become more educated, their income also increases significantly. Dou Zhang Tianzi (2023) established the Logit model to conduct an empirical analysis of the married female labor force from the aspects of personal factors and family factors, and believed that education level and spouse annual income have a significant impact on the willingness of married female labor force to work in the nearest place. Zhang Kang et al. (2021) said that with the improvement of urbanization, the increase in the number of young children has an increasing negative impact on the employment and self-employment of married women, and the increase in the number of children aged 0~2 has a greater negative impact on the employment and self-employment of married women than the increase in the number of children aged 3~6.

In addition, by comparing various models in machine learning, this paper comprehensively selects multiple regression models, decision trees, random forests, XGboost, logistics, support vector machines and other models for training and simulation, which have their own advantages, among which multiple regression models are used to quantify the linear relationship between multiple independent variables (married female characteristics) and dependent variables (the age of the youngest child). Wu et al. (2011) showed that decision trees are a common machine learning algorithm, and their structure is presented as a tree-like structure, in which each internal node represents an attribute judgment, each branch represents a possible value of the property, and each leaf node represents a category or a numeric value. Zhang Guijie et al. (2008) believe that decision tree algorithm is the most effective method to solve classification problems. Zhang Guijie et al. (2023) believe that decision trees have the advantages of discovering nonlinear relationships of latent variables, dealing with multi-classification problems, and selecting features when using decision trees to study the factors affecting the mental health of college students. It is more conducive to our next step of analysis and has a high comparability compared with multiple regression models. Yao et al. (2014) argue that although random forests may produce generalization errors within a certain limit, random forests do not produce overfitting problems with the increase of decision trees. Random forest (RF) Breiman is an ensemble machine learning method that uses the random resampling technology bootstrap and node random splitting technology to construct multiple decision trees and obtain the final classification result through voting. XGboost is an efficient gradient boosting algorithm that is able to handle complex nonlinear relationships and provide ranking of feature importance. Although the logistic model is mainly used for binary classification problems, in this study, it can be applied to the prediction of whether or not to have the youngest child by setting thresholds, for example.

Based on this, there are significant differences between the research methods in this paper and those in the previously mentioned literature. This article is not limited to the application of a single machine learning model, but also compares the advantages and disadvantages of multiple machine learning models and their characteristics. Specifically, this paper carefully selects multiple regression models, decision tree models, random forest models, XGboost models, logistic regression models, and support vector machine models for systematic research and comparison. Rather than only pursuing a major conclusion, this paper focuses more on the simulation effect of various models when processing the data selected in the article. Through detailed comparison and analysis, this paper aims to reveal the differences between different models in terms of prediction performance, data adaptability, and interpretability, so as to find the best prediction model. In this process, this study not only focuses on the prediction accuracy of the model, but also pays attention to the feasibility and reliability of the model in practical application. By comparing the simulation effects of various models, a series of conclusions can be drawn, which will help to deeply understand the influencing factors of married women's reproductive decisions, provide a scientific basis for the formulation of relevant policies, and then promote the optimization of population structure and the sustainable development of society.

II. DATA DESCRIPTION AND ANALYSIS

2.1 Data description:

This study uses the data table in the SATA learning package specifically for the employment of married women as the core data source, and selects the excel format for data storage after downloading, the dataset contains 23 different attributes, covering multiple aspects related to the employment of married women, and contains a total of 753 records, where the meaning of the attributes is shown in Table 1.

TABLE 1
DATA SOURCES AND THEIR MEANINGS

Attribute	Meaning
inlf	Whether a married woman is employed (1 means employed, 0 means unemployed)
hours	Number of hours worked
kidslt6	Whether you have a child younger than 6 years old (1 means yes, 0 means none)
kidsge6	The age of the child under 6 years old (the value here is from 0 to 6)
age	Married female age
edu	Educational attainment of married women
wage	Salary level for married women
repwage	Average wage
hushrs	Working hours of a married woman's husband
husage	The age of the married woman's husband
husedu	Educational attainment of the married woman's husband
huswage	The level of salary of a married woman's husband
faminc	Household income
mtr	Average productivity
matheduc	Married female mother's level of education
fatheduc	Married female father's level of education
unem	Number of years of unemployment
city	Whether the married woman is a town resident (1 means yes, 0 means no)
nwifeinc	Wife's income
lwage	Minimum wage level
exper	Work experience
expersq	Work experience squared
college	University education (1 means yes, 0 means no)

Based on the research in this paper, kidslt6 was selected as the dependent variable of the study, and the rest was used as the independent variable of model training to construct the training set and test set of the research data on the influence relationship between the characteristics of married women and the age of whether they have the youngest child. At the beginning of this paper, multiple regression models were selected to construct regression prediction models. It is compared with decision tree, random forest, xgboost, support vector machine, logistic and other models to observe whether the multiple regression model is the best prediction model for this dataset.

2.2 Descriptive statistical analysis of data:

2.2.1 Matrix diagram:

In the Python environment, in order to deeply explore the correlation between the variables in the dataset, especially the potential influence of the independent variable on the dependent variable, the correlation drawing package in Python was used to draw a matrix plot of the correlation scatter plot between the variables. In the matrix chart, the horizontal and vertical coordinates are composed of individual features, and in each cell of the matrix chart is a scatter plot between the corresponding abscissa variable and the corresponding ordinate variable, and the icon reflected by the main diagonal of the matrix chart is a column chart display of the changes of each variable.

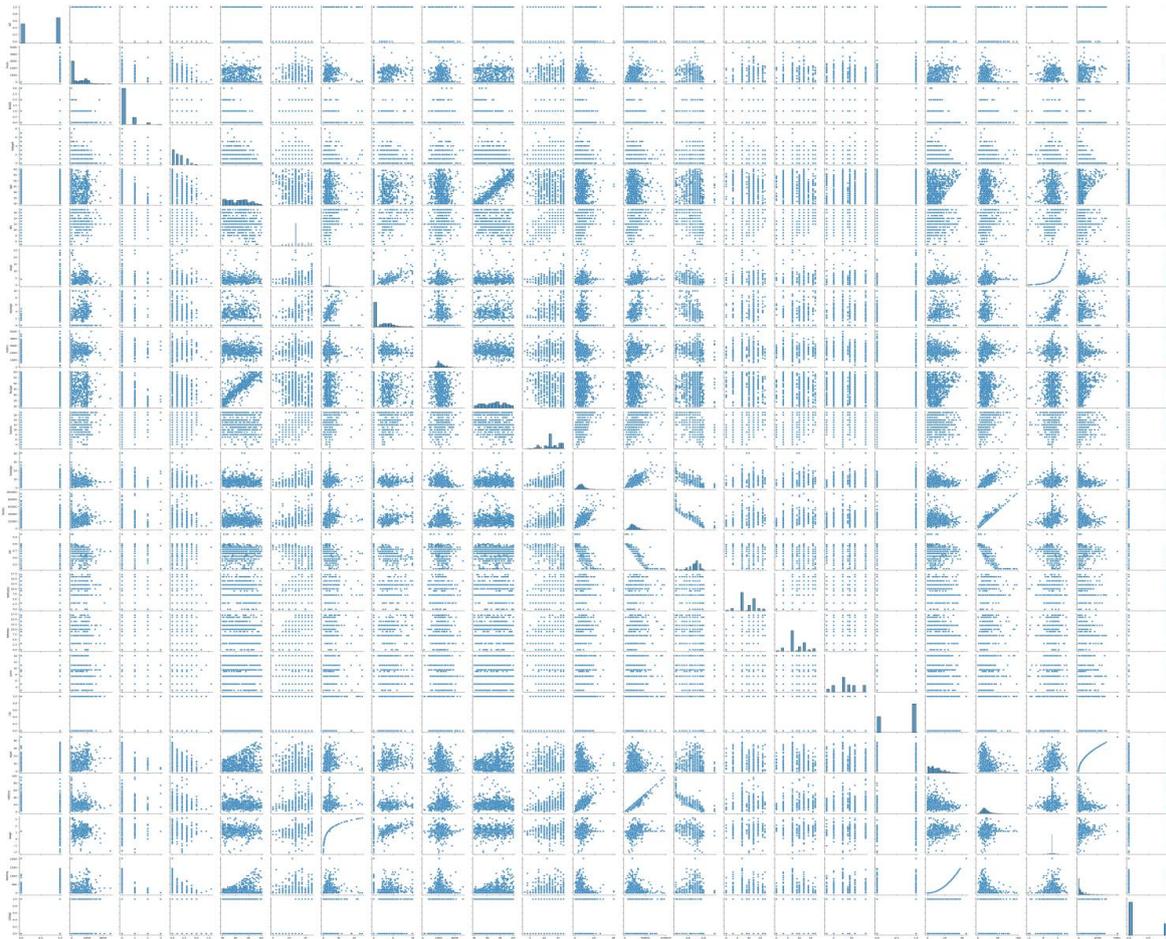


FIGURE 1: Correlation between variables, scatter plot matrix plot

You can zoom in on the graph, first look at the information displayed by each column chart on the main diagonal of the matrix chart, the change form of each variable is different, such as `inlf` variable, because it is classified data, so the column chart shown in the graph is the corresponding 1 and 0 of a data distribution, the same is true for the graphics of `kidslt6`, `kidsge6`, `city`, `college` and other variables, and then look at `age`, `edu`, `wage`. In addition to categorical variables, the column charts of the data show different shapes, from which we can see the distribution of each data, or there is a left-biased distribution, a right-biased distribution, and most of them are normally distributed, which are all information reflected in the data. In addition to observing the graph of the main diagonal, you can also randomly view the relationship between any two variables, some scatter plots are exponentially distributed, some are concentrated distributed, and some show obvious positive correlation trends and so on. Based on the nature of the dependent variable, the relationship between the graph and other independent variables is not obvious, and the correlation between them is very low, but more data information can be viewed through the matrix chart, which is convenient for later prediction and interpretation, and can also reflect certain practical problems and data trends.

2.2.2 Heat map:

The `df.corr()` code is used to view the correlation coefficients between the variables, and based on this, a heat map of the correlation coefficients between the variables is drawn, and the following figure 2 is obtained.

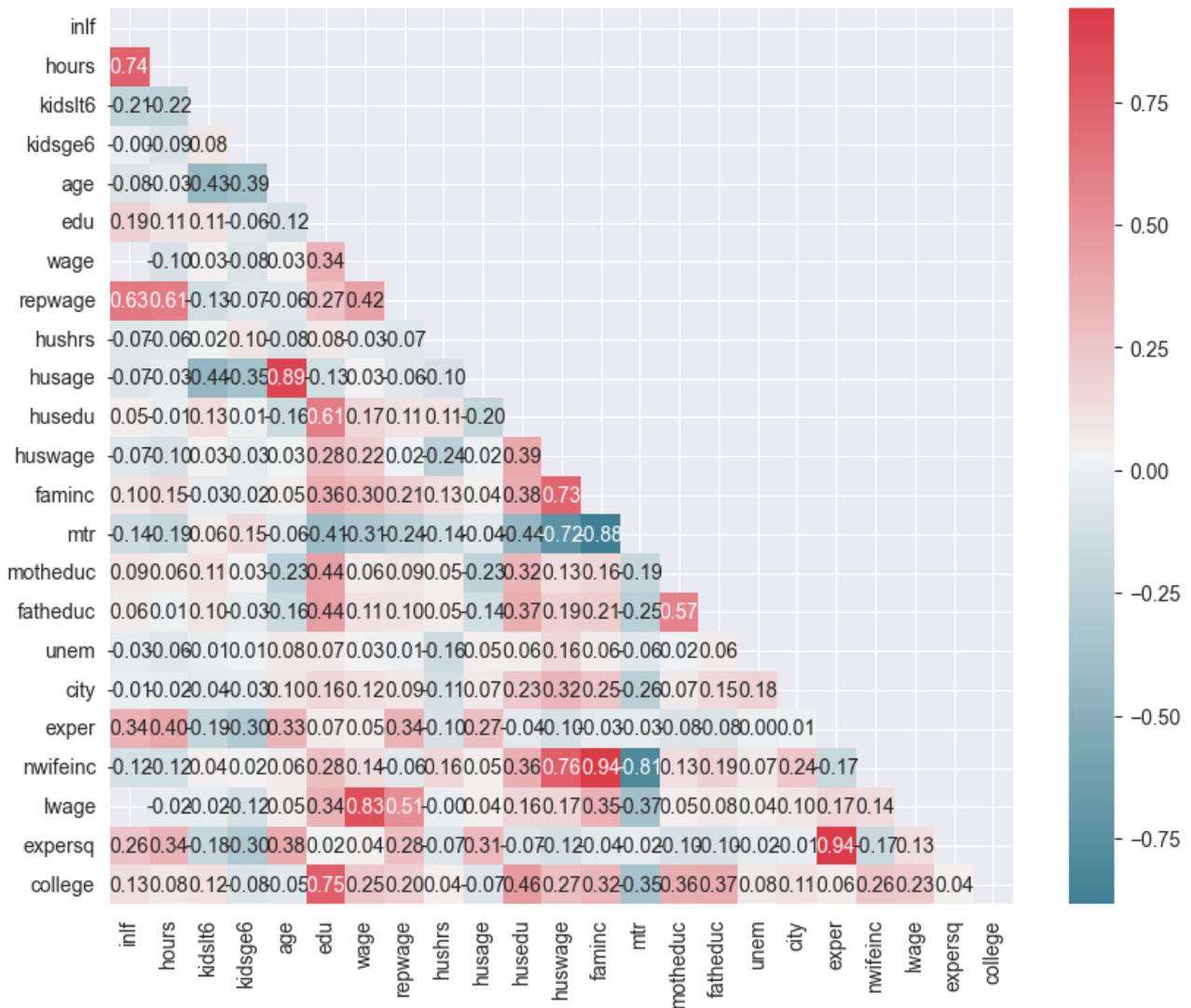


FIGURE 2: Heat map of the correlation between variables

Heatmaps are a highly intuitive visualization tool that clearly reveals the strength of correlations between variables in a dataset. In the heat map, it is possible to visualize which variables have significant correlations with each other by varying shades of color. Specifically, cells marked in red and blue represent a high degree of positive and negative correlation, while light blue may indicate a weak correlation (see Figure 2), which allows you to look at the relationship between the dependent and independent variables to set up a model to train and test the data, to test the research hypothesis or find the best predictive model.

III. THEORETICAL EXPLANATION OF THE MODEL

3.1 Theoretical analysis and research hypothesis:

3.1.1 Multiple regression model theory and advantages:

Multiple regression is a statistical analysis method used to explore the relationship between two or more variables, based on simple linear regression analysis, and can be used to predict the correlation between multiple independent and dependent variables. In the multiple regression model, the change of the dependent variable can be explained by the change of the independent variable and the disturbance of random factors, so the results of the multiple regression model analysis can intuitively show the change relationship between the dependent variable and the respective variable and random factors, which is more convenient for empirical analysis. Based on the research question of the article, kidslt6 was set as the dependent variable and the rest of the features as its independent variables.

3.1.2 Research hypotheses:

It is well known that the birth of a child is strongly influenced by a combination of family characteristics and a number of factors that make a married woman personal. From a theoretical point of view, with the increase in the proportion of married women in employment and the increase in family income, the number of children they raise may increase correspondingly, because the improvement of economic conditions may provide material security for raising more children, and the improvement of the social status of working women may also prompt them to be more willing and qualified to have more children. Realistically, however, the situation may not be so straightforward. As society continues to progress and develop, the improvement in the employment situation of married women may mean that they invest more time and energy in the workplace, which may in turn show a negative correlation with having the youngest children. Similarly, the relationship between other independent variables and dependent variables can be analyzed. Here are some assumptions made in this article:

H1: *inlf*, *hours*, *age*, *kidsge6*, *matheduc*, *city* were negatively correlated with *kidslt6*

H2: *hushrs*, *husage*, *husedu*, *mtr*, *faminc*, *wage*, *expersq*, *college*, etc. were positively correlated with *kidslt6*

3.1.3 Model theory:

Based on the selection of data features, the equations of the multiple regression model are as follows:

$$\begin{aligned} \hat{Y}(\textit{kidslt6}) = & \beta_0 \textit{inlf} + \beta_1 \textit{hours} + \beta_2 \textit{kidsge6} + \beta_3 \textit{expersq} + \beta_4 \textit{age} + \beta_5 \textit{wage} \\ & + \beta_6 \textit{repwage} + \beta_7 \textit{hushrs} + \beta_8 \textit{husage} + \beta_9 \textit{husedu} + \beta_{10} \textit{faminc} + \beta_{11} \textit{mtr} \\ & + \beta_{12} \textit{mathedu} + \beta_{13} \textit{fathedu} + \beta_{14} \textit{unem} + \beta_{15} \textit{city} + \beta_{16} \textit{exper} \\ & + \beta_{17} \textit{nwifeinc} + \beta_{18} \textit{college} + \mu \end{aligned} \quad (1)$$

The variables are interpreted as follows: *inlf* indicates whether or not employed, where "*inlf*=1" indicates employment and "*inlf*=0" indicates unemployment. *Kidsge6* indicates the number of children under the age of 6; *hours* indicates the number of hours worked; *expersq* denotes the square of the experience; *age* denotes age; *wage* indicates the level of wages; *Repwage* indicates average salary; *hushrs* indicate the number of hours worked by married female husbands; *husage* indicates the age of the married woman's husband; *husedu* indicates the level of education of the married female husband; *faminc* indicates the status of household income; *MTR* stands for Household Productivity; *Matheduc* indicates the education of a married female mother, and *fatheduc* indicates the education of a father; *unem* indicates unemployment status; *City* means city, 1 means in city, 0 means not in city; *Exper* means work experience. *Nwifeinc* indicates the wife's income; *College* indicates whether you have attended a university, where "*college*=1" indicates that you have attended a university and "*college*=0" indicates that you have not attended a university; μ denotes the random perturbation term of the model, which is the random error term of the model.

Through the construction of the model, more direct results can be obtained in the subsequent training and testing, and the data can be regressed by using the least squares method (originally the data was initially processed, including the test and correction of the multicollinearity of the model, the test and correction of heteroskedasticity, autocorrelation, etc., and the stationarity of the model, but in view of the needs of this paper, these processes are directly ignored, and the original data is directly used for regression testing and the results are used as the basis for comparison between models). The results of each variable and parameter are obtained.

IV. FEATURE PROCESSING AND MODEL TRAINING

4.1 Feature processing:

In the data preprocessing stage, this study uses the code `user_info.info()` to view the information of the data and observe whether there are potential problems such as outliers and missing values in the data. It is important to note that none of these data with missing values are categorical data, but continuous numerical data. Considering that the lack of categorical data may involve data integrity and classification accuracy, and the absence of continuous data may be compensated by a reasonable filling strategy, this study decided to use the mean filling method for the variables with missing values, and the results in Table 3 were obtained.

TABLE 2
CLASS 'PANDAS.CORE.FRAME.DATAFRAME

Number	Column	Non-Null Count	Dtype
0	inlf	753 non-null	int64
1	hours	753 non-null	int64
2	kidslt6	753 non-null	int64
3	kidsge6	753 non-null	int64
4	age	753 non-null	int64
5	edu	753 non-null	int64
6	wage	753 non-null	int64
7	repwage	753 non-null	int64
8	hushrs	753 non-null	int64
9	husage	753 non-null	int64
10	husedu	753 non-null	int64
11	huswage	753 non-null	int64
12	faminc	753 non-null	int64
13	mtr	753 non-null	int64
14	matheduc	753 non-null	int64
15	fatheduc	753 non-null	int64
16	unem	753 non-null	int64
17	city	753 non-null	int64
18	nwifeinc	753 non-null	int64
19	lwage	753 non-null	int64
20	exper	753 non-null	int64
21	expersq	753 non-null	int64
22	college	753 non-null	int64

4.2 Model training:

4.2.1 Multiple regression model:

Based on the data obtained by feature processing of the data, the multiple regression model was constructed, and the least squares (OLS regression) regression analysis was realized by the code, and the following least squares regression result table was obtained.

TABLE 4
OLS REGRESSION RESULTS

Dep. Variable:	kidslt6	R-squared (uncentered):	0.425
Model:	OLS	Adj. R-squared (uncentered):	0.409
Method:	Least Squares	F-statistic:	27.11
Date:	Thu, 28 Dec 2023	Prob (F-statistic):	1.32e-74
Time:	22:12:20	Log-Likelihood:	-443.38
No. Observations:	753	AIC:	926.8
Df Residuals:	733	BIC:	1019
Df Model:	20		
Covariance Type:	nonrobust		

Variable	Coefficient	Std. Error	t-value	P> t	[0.025	0.975]
inlf	-0.1820	0.053	-3.413	0.001	-0.287	-0.077
hours	-6.307e-06	4.24e-05	-0.149	0.882	-8.95e-05	7.69e-05
expersq	0.0003	0.000	1.418	0.157	-0.000	0.001
kidse6	-0.0538	0.014	-3.787	0.000	-0.082	-0.026
age	-0.0186	0.005	-4.020	0.000	-0.028	-0.010
edu	0.0134	0.012	1.073	0.284	-0.011	0.038
wage	0.0102	0.009	1.151	0.250	-0.007	0.027
repwage	0.0068	0.011	0.595	0.552	-0.016	0.029
hushrs	-1.802e-05	2.79e-05	-0.647	0.518	-7.27e-05	3.67e-05
husage	-0.0155	0.004	-3.522	0.000	-0.024	-0.007
husedu	0.0073	0.007	1.021	0.308	-0.007	0.021
faminc	-9.608e-06	9.89e-06	-0.971	0.332	-2.9e-05	9.81e-06
mtr	2.0558	0.206	9.964	0.000	1.651	2.461
motheduc	-0.0033	0.006	-0.538	0.591	-0.015	0.009
fatheduc	0.0011	0.006	0.195	0.845	-0.010	0.013
unem	0.0024	0.005	0.439	0.661	-0.008	0.013
city	-0.0186	0.037	-0.508	0.612	-0.090	0.053
exper	-0.0042	0.006	-0.690	0.491	-0.016	0.008
nwifeinc	0.0220	0.010	2.233	0.026	0.003	0.041
college	0.1043	0.055	1.893	0.059	-0.004	0.212

Omnibus:	224.371	Durbin-Watson:	1.964
Prob(Omnibus):	0.000	Jarque-Bera (JB):	677.117
Skew:	1.455	Prob(JB):	9.25e-148
Kurtosis:	6.621	Cond. No.:	3.38e+05

Note: Marked in red is significant at 0.1 and marked in blue is significant at 0.5

By calculating the specific formula of the regression model in Python and combining the results obtained by the least squares method, the specific form of the model can be obtained:

$$Y = -0.182 \text{ inlf} - 0.000 \text{ hours} - 0.053 \text{ kidsge6} + 0.000 \text{ expersq} - 0.019 \text{ age} + 0.013 \text{ edu} + 0.010 \text{ wage} + 0.007 \text{ repwage} - 0.000 \text{ hushrs} - 0.015 \text{ husage} + 0.007 \text{ husedu} - 0.000 \text{ faminc} + 2.055 \text{ mtr} - 0.003 \text{ matheduc} + 0.001 \text{ fatheduc} + 0.002 \text{ unem} - 0.018 \text{ city} - 0.004 \text{ exper} + 0.022 \text{ nwifeinc} + 0.104 \text{ college} + 1.031 \quad (2)$$

4.2.2 Decision tree model:

A decision tree algorithm is a common machine learning algorithm used to solve classification and regression problems. The main idea is to construct a tree structure by recursively segmenting the dataset, where each node represents a feature, each branch represents the value of the feature, and each leaf node represents a category or a numeric value. The Classification and Regression Tree algorithm is suitable for modeling complex data with multiple feature variables, and has the advantages of simple extraction rules, high accuracy and strong interpretability. In view of this, this paper uses a simple decision tree model to predict that married women will have children under 6 years of age. The generation of decision trees is a binary recursive partitioning process, which adopts a splitting criterion that minimizes the square error. The splitting process starts from the root node, and each time the feature attribute with the smallest square error and its feature attribute value are taken as the optimal splitting attribute and the optimal splitting attribute value, all the features and eigenvalues are traversed, and the input is divided into two regions in turn. Then, the above partitioning process is repeated for each sub-region until the stop condition is satisfied, and a regression decision tree is generated.

The model compares the magnitude of the loss function error by the squared error, and uses the squared error minimization criterion to determine the optimal segmentation point and the predicted value. According to the principle of least squares, the mean of all outputs on the subspace can be expressed as, based on which the sharding attribute is selected to divide the input space, so that it can go through all the input eigenvariables and their eigenvalues to find the optimal sharding point, and then divide the space into two subregions according to the slicing point, until the two subregions cannot be divided, and the corresponding optimal output value can be obtained. The decision tree constructed in this paper uses the square error minimization as the splitting criterion for the selection of characteristic variables, the maximum depth is 10, the minimum number of samples required for internal node division is 2, and the maximum leaf node is 10. The following values about the accuracy of the model are obtained by the code implementation Accuracy_TAET which is 0.786, the mean result of cross-validation of the decision tree, and the cross-validation scores of the decision tree model are [0.74834437 ,0.79470199 ,0.69536424 ,0.7 ,0.64], and the average cross-validation score is 0.715.

After training and testing the model data, the algorithm related to cross-validation scores was used to perform five-fold cross-validation, and each cross-validation used a different subset of data to train and test the model, and the results are shown above, and the decision tree cross-validation results can be known, where the first cross-validation score is 0.74834437, the second score is 0.79470199, the third score is 0.69536424, and the fourth score is 0.7, with a score of 0.64 for the fifth time. According to the results of the five cross-validations, the average cross-validation score is obtained, and the value is the average cross-validation score = $(0.74834437 + 0.79470199 + 0.69536424 + 0.7 + 0.64) / 5 = 0.71568$. This average cross-validation score represents the average performance of the model over multiple cross-validations. A higher score usually indicates a better performance of the model. In this example, the average cross-validation score is 0.715, which means that the model has achieved an average accuracy of about 71.6% across multiple cross-validations, and the model performs well.

4.2.3 Random forest model:

Random forest is an ensemble learning model, and each decision tree in the random forest is built based on random samples and random features, which can reduce the risk of overfitting and improve the generalization ability of the model. Random forests can combine multiple decision trees for classification or regression, which can reduce the error of a single decision tree and improve the accuracy of the model. Random forests can process a large number of features and samples, making them suitable for datasets of all sizes. Random forests can assess the importance of features and help with feature selection. Random forests can be computed in parallel, speeding up model training. In the process of training, the data of the training set and the test set were divided into 8:2 and a random forest classifier containing 100 trees was created, and the seed of the random number generator was set to 42. Finally, the accuracy of the model is 0.814.

4.2.4 Gradient boosting regression tree:

Gradient Boosting Regression Tree is a machine learning method for regression problems. Unlike random forests, gradient boosted regression trees train models by fitting residuals stepwise rather than by building multiple parallel models to integrate predictions. The basic idea of gradient boosted regression trees is to combine multiple weak learners into one strong learner. At each iteration step, the new decision tree attempts to correct the errors of the previous tree. At the heart of this approach is the use of a gradient boosting strategy, which optimizes the gradient of the loss function with respect to the parameters. Based on this, a model was also selected for comparison in the problem study, and the following settings were made in the training

process, and the data of the training set and the test set were divided into a ratio of 8:2, and the number of random seeds was 2. After training, the accuracy on training set is 0.99003, and the accuracy on test set is 0.834.

4.2.5 XGboost model:

XGBoost is a machine learning algorithm based on gradient boosting decision trees. It fits the data by iteratively adding trees and uses a gradient boosting strategy to optimize the loss function. XGBoost uses a forward distribution algorithm for greedy learning during training, and each iteration learns a CART tree to fit the residuals of the prediction results of the previous $t-1$ trees and the true values of the training samples. The core principle of XGBoost is that weak classifiers are integrated together to form strong classifiers. XGBoost adds a regular term to the cost function to control the complexity of the model. XGBoost supports multi-threading (parallel) gain calculation of each feature when splitting nodes, so the algorithm is faster and has a relatively higher accuracy. During the training process, the following settings were made to divide the data of the training set and the test set in an 8:2 ratio. 100 trees are selected with a learning rate of 0.1, a random sampling ratio of 0.3 for each tree training, and a maximum depth of 5 for each tree. After training, the accuracy of the model is 0.827.

4.2.6 Logistic model:

Logistic Regression is a machine learning method for classification problems. It uses the logistic function (also known as the Sigmoid function) to convert the results of a linear regression into a probability value, enabling the classification of data. The basic principle of logistic regression is to construct a regression equation from training data, and then use this equation to predict the category of new unknown data. During training, the algorithm looks for the best combination of parameters to minimize the error between the predicted and true categories. In the problem study, the following settings were made in the training process, and the data of the training set and the test set were divided into a ratio of 8:2, and the number of iterations was set to 1000, and finally the accuracy of the model was 0.79.

4.2.7 K nearest neighbor model:

The K-Nearest Neighbor (KNN) model is a very simple and intuitive model, and it is important to choose an appropriate value of k (k is an integer) in the K-nearest neighbor model. The choice of k -value will determine how the model classifies or predicts. In the problem study, the test data is set to 0.25 of the original data, the number of random seeds is 10, and the number of neighbors in the KNN regressor is 502. After the model is trained, the accuracy of the model is 0.830.

4.2.8 Support vector machine model:

A Support Vector Machine (SVM) is a supervised learning model used for classification and regression analysis. It does this by finding decision boundaries that maximize the separation of different categories of data points. The basic idea of the SVM is to find a hyperplane that maximizes the separation of different classes of data points. In the case of nonlinear separability, SVMs can map the input space to a high-dimensional feature space by using a kernel function, and then look for a hyperplane in that feature space that maximizes the separation of data points. In model training, the test set is set to the original 0.3, the random number generator seed is 42, and the regularization parameter is 1. The model is trained and the accuracy of the model is 0.778.

V. MODEL COMPARISON AND ANALYSIS

5.1 Model evaluation index standards:

The following evaluation indicators are used in problem research: (1) R^2 (R-squared) is a statistical index used to measure the goodness of fit of the regression model. R^2 represents the proportion of the variation in the dependent variable that can be explained by the independent variable. It has a value between 0 and 1, with closer to 1 indicating a better fit for the model and the more explanatory the independent variable is to the dependent variable. (2) The Average Cross-Validation Score is a metric that evaluates the performance of a machine learning model, and the average score is calculated by dividing the dataset into multiple subsets and training and validating the model multiple times on these subsets. This score reflects the model's performance on unknown data and helps to understand the model's generalization ability. (3) Accuracy is an important evaluation metric used to measure the proximity between the result of a measurement, calculation, or observation and the true value. Accuracy is a key evaluation criterion in fields such as scientific experiments, engineering, statistics, and more. (4) Accuracy on training set is limited to using the accuracy on the training set as an evaluation indicator. In machine learning, accuracy on the training set usually refers to the classification or regression accuracy of the model on the training data, which does not fully reflect the performance of the model in real-world applications. (5) Accuracy on test set is one of the important

metrics to evaluate the performance of machine learning models. A test set is a separate dataset from the training set that evaluates how the model performs on unknown data. By calculating the accuracy on the test set, you can get an idea of how the model performs in a real-world scenario.

5.2 Model comparison:

TABLE 3
THE VALUE OF THE EVALUATION INDEX OF EACH MODEL

Model	Evaluation indicators	Value
Multiple regression models	R ²	0.452
	Average cross-validation score	0.102
Decision tree model	Accueacy	0.781
	Average cross-validation score	0.715
Random forest model	Accueacy	0.814
Gradient boosting regression tree	Accuracy on training set	0.99
	Accuracy on test set	0.834
XGboost model	Accueacy	0.827
Logistic model	Accueacy	0.79
K-nearest neighbor model	Accueacy	0.83
Support vector machine models	Accueacy	0.778

From the above table, it can be seen that the accuracy of gradient boosted regression tree model, K-nearest neighbor model, XGBoost model and random forest model is not much different, the accuracy of the improved regression tree model is the highest, the model is the best, the accuracy of the decision tree model, the support vector machine model and the logistic model is also higher, and the performance of the model is also better, but the fitting effect of the multiple regression model is the worst. It shows that the data used in this paper are better than the linear regression model compared with the regression machine learning algorithm.

5.3 Model analysis and interpretation:

Based on the training results obtained in the model training process, for the regression model, there are several significant factors in the model of the relationship between the comprehensive characteristics of married women and whether they have children aged 0-6 years, and the characteristics of married women are inlf, expersq, kidsge6, age, edu, wage, husage, husedu, faminc, mtr, exper, nwifeinc, college The relationship of this independent variable to the dependent variable is significant at different levels of significance, and it can be concluded that both hypothesis H1 and hypothesis H2 are true.

In addition, the accuracy of other models is due to the degree of fit of multiple regression models, and the best model is the gradient boosted regression tree model, as it is a machine learning method for regression problems. Unlike random forests, gradient boosted regression trees train models by fitting residuals stepwise rather than by building multiple parallel models to integrate predictions. The basic idea of gradient boosted regression trees is to combine multiple weak learners into one strong learner. At each iteration step, a new decision tree corrects the mistakes of the previous tree. At the heart of this approach is the use of a gradient boosting strategy, which optimizes the gradient of the loss function with respect to the parameters. In turn, a high accuracy rate can be obtained during the training process. The K-nearest neighbor model is second, and the XGboost model also performs well, i.e., the classification or ensemble model is better than the linear regression model. There are several reasons why the linear regression model is not well fitted: first, there is no multicollinearity, heteroskedasticity, autocorrelation and other tests and corrections for each variable during model training, and there is no stationarity test for the model; Second, the dependent variable is a categorical data, which is regressed with continuous data, and the effect is not significant, and there

is no good reference in the selection of variables, and the multiple linear regression model does not perform as well as the gradient to improve the regression tree and other machine learning models.

VI. CONCLUSIONS AND RECOMMENDATIONS

Based on the research on the comprehensive characteristics of married women and whether they have children aged 0-6 years old and other methods such as machine learning, this paper draws the following conclusions: (1) A series of comprehensive characteristics such as employment status, age level, and mother's education level of married women have a significant negative correlation with the decision of whether to have children aged 0-6. In other words, these characteristics will influence to some extent whether married women choose to have children. Specifically, as married women increase in employment, age, and mothers become more educated, their willingness to have children appears to be decreasing, leading to a corresponding decrease in the number of children aged 0-6 years. (2) The characteristics of married women's husbands and the characteristics of the family as a whole had a positive and significant impact on whether they had children aged 0-6 years. This means that when married women have good husbands and good family characteristics, they are more inclined to want married women to have children aged 0-6 years, and it is worth noting that the higher the education level of the father of a married woman, the positive impact on his or her childbearing. This finding seems to hint at a general psychological tendency that men, whether as husbands or fathers, expect married women to be able to have young children to some extent. (3) Based on the results of multiple regression model, it is found that when conducting research or selecting models, it is necessary to fully consider various possible influencing factors and conduct a comprehensive analysis. Avoid subjective selection of models, but there is a real possibility that they do not fully match the characteristics of the data. Therefore, before making a decision, it is necessary to carefully evaluate the pros and cons of various models to ensure that the research or prediction can achieve better results. (4) After comparing various models, it is found that the gradient boosted regression tree model, K-nearest neighbor model, XGBoost model and random forest model are significantly better than linear regression models. The results suggest that when dealing with complex data problems, more diverse models and algorithms should be tried to find the most suitable solution for the data characteristics, so as to better capture the nonlinear relationships and latent features in the data, so as to provide deeper and more comprehensive insights for research. The research in this paper has the following three policy implications:

First, policies tend to encourage childbearing. In order to effectively encourage childbearing, policymakers should focus on improving the family situation of married women, thereby motivating them to have children. Specifically, a series of measures such as information platform sharing and sharing sessions can be used to promote married women and their husbands to receive a higher level of university education, so as to improve the education level of the whole family. In order to better broaden their horizons and enhance the efficiency and competitiveness of families in production and life. At the same time, by improving family productivity, such as providing a family-friendly working environment and a flexible work system, the pressure on married women between home and the workplace can be further reduced, and a more relaxed and suitable environment for them to have children can be created.

Second, to promote women's employment, we should take practical and effective measures to protect women's equal rights and interests in the workplace based on the reality of their family situation. This includes, but is not limited to, promoting equity in employment in society and ensuring that women are not discriminated against or excluded on the basis of their gender. At the same time, the principle of fairness should be promoted within the family to avoid placing the responsibility of raising young children entirely on the shoulders of women, thereby limiting their career choices and career development. Men should be encouraged to participate more in family life, share childcare responsibilities and create a more equal and free employment environment for women.

Third, married women should fully consider their age when choosing to have children. Because as you get older, the risks and difficulties you face in having children also increase. Therefore, it is advisable for married women to plan their family plans as early as possible within the appropriate age range to ensure the health and safety of themselves and their children. At the same time, society should also provide more reproductive support and protection for married women, such as providing high-quality medical resources and parenting guidance, so as to help them better cope with various challenges in the reproductive process.

Fourth, when analyzing the data used in this paper, we should take a comprehensive and detailed comparative analysis approach. By comparing the similarities and differences between different data, we can gain a deeper understanding of the nature and laws of the problem. At the same time, in order to find the most suitable analysis model, we also need to explore

and try in many aspects based on the actual situation. The same principle applies to other research work. We should always maintain objectivity and rigor when analysing and processing data to ensure the accuracy and reliability of the research results.

REFERENCES

- [1] Huang Qian, Cao Shurui. The impact of the number of children on the employment choice of married women[J].China Economic Issues,2022(06):67-81.)
- [2] Song Alun, Zhang Feng. Analysis of the influence of family characteristics on the employment quality of married women[C]//Sichuan Labor and Social Security Magazine Publishing Co., Ltd. Proceedings of the Labor and Social Security Research Conference (15), 2022:4.
- [3] Yan Yu. Age of the youngest child and married women's labor force participation[D].Southwestern University of Finance and Economics,2022.
- [4] Luo Yuan. An empirical study on the labor supply decision of married women[D].Nanjing Agricultural University,2009.
- [5] Wang Guanghui, Zhang Shiwei. An empirical analysis of the impact of education on the employment and income of married women in China[J].Science, Economy & Society,2010,28(04):83-87+93.
- [6] Dou Zhangtianzi. Analysis of influencing factors of local and nearby employment of rural married female labor force: Based on Logit model[J].Tianjin Agricultural Sciences,2023,29(01):45-48+61.
- [7] Zhang Kangzi, Wang Yadi. Research on the influence of childbearing on married women's employment choice[J].Public Management Review,2021,3(03):53-75.
- [8] Wu Xiaogang, Zhou Ping, Peng Wenhui. Application of decision tree algorithm in the evaluation of college students' mental health[J].Computer Applications and Software,2011,28(10): 240-244.
- [9] Zhang Guijie, Wang Shuai. Research on ID3 algorithm for decision tree classification[J].Journal of Jilin Normal University(Natural Science Edition),2008,29(3):135-137.
- [10] Zhang Guijie, Wang Xiaocan, Xing Weikang, et al. Research on psychological assessment model based on decision tree classification algorithm[J].Journal of Jilin Normal University(Natural Science Edition),2023,44(04):123-130.
- [11] Yao Dengju, Yang Jing, Zhan Xiaojuan. Feature selection algorithm based on random forest[J].Journal of Jilin University(Engineering Science),2014,44(01):137-141.
- [12] BreimanL.Randomforests[J]. Machine Learning,2001,45(1):5-32.
- [13] Wu Xiaogang, Zhou Ping, Peng Wenhui. Application of decision tree algorithm in the evaluation of college students' mental health[J].Computer Application and Software,2011,28(10) : 240-244.
- [14] Zhang Guijie, Wang Shuai. Research on ID3 algorithm for decision tree classification[J].Journal of Jilin Normal University(Natural Science Edition) ,2008,29(3) : 135-137.
- [15] Hu Nan, Xue Fujing, Wang Haonan. Does Short-sightedness of Managers Affect Long-Term Investment? Management World,2021,37(05):139-156+11+19-21.
- [16] Yang Jianfeng, Qiao Peirui, Li Yongmei, et al. Statistics and Decision,2019,35(06):36-40.
- [17] Wu Changqi, Zhang Kunxian, Zhou Xinyu, et al. Digital transformation, competitive strategy choice and high-quality development of enterprises: Evidence based on machine learning and text analysis[J].Economic Management,2022,44(04):5-22.
- [18] Li Bin, Shao Xinyue, Li Yueyang. China Industrial Economics,2019,(08):61-79.
- [19] Huang Bowen, Liu Tianli. Population and Development,2024,30(03):38-50.
- [20] Huang Qian, Cao Shurui. The impact of the number of children on the employment choice of married women[J].China Economic Issues,2022,(06):67-81.
- [21] Zhang Kangzhi, Wang Yadi. Research on the influence of childbearing on married women's employment choice[J].Public Management Review,2021,3(03):53-75.
- [22] Zheng Yuxin, Wu Yuxia, Mi Hong. Analysis of influencing factors of married female floating population having two children: Based on data analysis of Jiangsu, Zhejiang and Shanghai[J].Northwest Population,2023,44(06):100-114.

Assessment of Antibiotic Residues in Commercial Aquaculture Feeds in Selected Farms in Awka, Anambra State, Nigeria

Toochukwu Franklin OSEGBO^{1*}, Chika Florence IKEOGU²;
Chidimma Gift NWANKWO³

Department of Fisheries and Aquaculture Management, Faculty of Agriculture, Nnamdi Azikiwe University, Awka, Nigeria

*Corresponding Author

Received: 06 March 2025/ Revised: 11 March 2025/ Accepted: 23 March 2025/ Published: 31-03-2025

Copyright © 2025 International Journal of Engineering Research and Science

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted Non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract— Aquaculture feeds are often compounded with antibiotics to promote growth and prevent diseases in fish. However, the use of antibiotics in aquaculture feeds has raised concerns about the presence of antibiotic residues in fish which subsequently leads to antimicrobial resistance. This study therefore investigated antibiotic residues in Nine (9) different fish feeds used in three selected fish farms located in Awka, Anambra state, Nigeria. One gram (1g) of the feed samples were used for antibiotics residues analysis. The fish feed samples were collected in three replicates each, making it twenty-seven treatment samples. The targeted antibiotics were the commonly used antibiotics which include Tetracycline, Ciprofloxacin, Amoxicillin and Azithromycin. The feed samples were analyzed for antibiotics residues concentration using Uv-Visible Spectrophotometry. The results showed that the mean antibiotic residue concentration from the three farms were higher in the following order; Tetracycline > Ciprofloxacin > Amoxicillin > Azithromycin, with mean concentrations of 85.40 ± 1.78 ug/g, 0.031 ± 0.002 ug/g, 0.019 ± 0.002 ug/g and 0.010 ± 0.001 ug/g respectively. The mean concentration of Tetracycline was higher than the European Union (EU) maximum residue limit of 0.1 ug/g but Ciprofloxacin, Amoxicillin and Azithromycin were all below the EU maximum residue limits of 0.05 ug/g in feed. This study concluded that the analyzed feed samples contained low dose of antibiotics which may contribute to the development of antibiotic-resistant bacteria. It is therefore recommended that antibiotics should not be used in fish feed formulation, so as to help the farmer to make responsible use of antibiotics when there is a genuine need.

Keywords— Antibiotics resistance, Drug residues, Aquaculture, Fish feeds, Awka-Nigeria.

I. INTRODUCTION

Nutrition is a major factor determining the potential of cultured fish to exhibit their genetic capability for growth and reproduction. The increasing costs and limited availability of fish feed have prompted the need for research into alternative feed sources (Ogunkalu, 2019). It is essential to provide cultured fish with sufficient nutritional feed to enhance growth and immunity, ultimately leading to a successful harvest. In order to meet the nutritional requirements of cultured fish, fish feeds are often enriched with additives. Feed additives are dietary ingredients that are incorporated into feed formulations not only for the usual provision of basic nutritional requirements as offered by traditional feed, but also to improve growth, immune responses, and disease resistance in aquatic animals (Onomu and Okuthe, 2024). These additives include probiotics, prebiotics, enzymes, antibiotics, medicinal plants and other feed additives. Antibiotics can be used as feed additives in animal production in order to maintain animal health (Ghimpețeanu *et al.*, 2022). In the aquaculture industry, the health of aquatic animals is of utmost importance in order to ensure optimal production and profitability. Therefore antibiotics have been commonly used to combat diseases and parasites in aquatic animals (Lulijwa *et al.*, 2020). Antibiotics have been included in aquatic feed to prevent or treat bacterial diseases and promote growth (Onomu and Okuthe, 2024). The most often used antibiotics in aquaculture are oxytetracycline, florfenicol, macrolides, fluoroquinolones, and sulfadiazine (Lulijwa *et al.*, 2020). The most

common route for the administration of antibiotics in aquaculture is through aquaculture feed (Liu *et al.*, 2017). Pathmalal (2018) reported that antimicrobials fed to fish are not efficiently metabolized by fish and are eliminated through urine and feces into the pond water. While antibiotics have been effective in controlling diseases, their use has raised significant concerns regarding fish health, environmental impact, and human health (Chowdhury *et al.*, 2022). A consequence of the use of the antibiotics in food-producing animals is the presence of drug residues, even in a very low concentrations, in the edible tissues of the treated animal (Monteiro *et al.*, 2018). The presence of antibiotic residues in aquaculture products could result in the development of bacterial resistance and toxicity to consumers that can lead to morbidity and/or death (Okocha *et al.*, 2018). Consumption of such products may result in many health problems in humans. Chiefly among the health concerns is the development and propagation of antimicrobial resistance along the food chain (Tang *et al.*, 2020). Several studies have reported the presence of antibiotic residues in various types of animal feed (Manyi-Loh *et al.*, 2023). Currently, more antibiotics are used in food animal production to promote growth and prevent disease than in the entire human population. Although millions of human lives have benefited from antimicrobials, the majority (73%) of such antimicrobials are those utilised for animal production (Van Boeckel *et al.*, 2017). Among intensive types of livestock production that use antimicrobial agents, particular attention should be paid to fish farms because of their direct impact on the aquatic environment (Shao *et al.*, 2021). Antibiotics in fish culture are usually administered through feed and the intestinal environment of fish is an optimal site for the selection of antibiotic resistant bacteria and through fish feces, antibiotic resistant bacteria are dispersed to the water column or sediments (Bajorski *et al.*, 2020). Thus, utilising a diverse range of these antimicrobial agents in fish feed, including those used in human medicine, guarantees their persistence in the aquatic environment for an extended period (Okon *et al.*, 2022) which could lead to the development of antimicrobial resistance. Large quantities of antimicrobials are used in aquaculture feed in low- and middle-income countries, often without professional supervision with consequences for development and spread of resistance and global public health.

Okeke *et al.* (2022) reported that antibiotic usage increased from 21.1 to 34.8 billion defined daily doses in 76 countries between 2000 and 2015, and it is anticipated to reach 126 billion defined daily doses in 2030 if no policy changes are made. There are strict laws on the usage of antibiotics and chemotherapeutics in the aquaculture feed industry (Bojarski *et al.*, 2020). In the absence of effective regulations, there is a risk for large number of antibiotics including medically important antibiotics (Chowdhury *et al.*, 2022) becoming available for use in fish feed formulation. The United States Food and Drug Administration (FDA) enforces rigorous monitoring practices for antibiotics, such as oxytetracycline and florfenicol, with a zero-tolerance policy for certain high-risk residues (Miranda *et al.*, 2018). Meanwhile, antibiotics used in aquaculture feed help to improve fish health by killing or inhibiting the growth of pathogenic bacteria. On the other hand, it causes more side effects. Antibiotics have helped to reduce mortality rates in developing nations by reducing deaths from infections, but the costs of antibiotic contamination continue to be a major issue (Okeke *et al.*, 2022). In order to ensure food safety for consumers, more and more studies have attempted to find effective and rapid methods for the detection of antibiotic residues in feed and food products (Long *et al.*, 2018). Consumers are also increasingly interested in consuming quality food, and are turning to organic products, which provide them with more safety relative to conventional products (Ghimpețeanu *et al.*, 2022). Therefore, this study was aimed to assess the levels of antibiotics in commercial aquaculture feeds used in selected fish farms in Awka to ensure the safety of aquaculture products and protect human health.

II. MATERIAL AND METHODS

2.1 Sample Collection and Preparation:

A total number of nine (9) different fish feeds samples were collected from three different fish farms in Awka, Awka South local government Area, Anambra state. Awka south is located within Latitude 6°21'N to 6°08'N, and between Longitude 7°07'N to 7°44'E. The study sites where the fish feed samples were collected are Fish Farm 1 (latitude 7°48'33"N, Longitude 6°91'75"E), with fish feed samples (A,B,C), Fish Farm 2 (Latitude 6°21'06"N, Longitude 7°07'60"E) with fish feed samples (D,E,F) and Fish Farm 3 (Latitude 6°21'06"N, Longitude 7°07'09"E) (Source Google), with fish feed samples (G,H,I). The fish feed samples were collected in three replicates each, making it twenty-seven treatment samples. The fish feed samples collected depends on the type of fish feed used by those fish farmers. 9mm feed samples was used for this analysis. The feed samples

were collected in clean plastic bags and taken to Springboard laboratory for analysis. Springboard Laboratory is located at Udoka Estate in Awka, Anambra state.

2.2 Preparation and Methods:

2.2.1 Antibiotics Assay:

One gram of the feed samples was weighed with digital weighing balance (Model OHAUS M44Y) and then grounded using a mortar and pestle. Tetracycline was extracted with McIlvaine buffer, Quinolones was extracted with 0.1% formic acid in acetonitrile, Beta-Lactams were extracted with a solution consisting of 0.1 M Phosphate buffer (pH 7) and methanol (1:1) while Macrolides were extracted using 0.1 M Phosphate buffer (pH 6.9), to enhance their solubility. After extraction, Tetracycline was vortexed and then centrifuged before determining the absorbance at 355 nm, Quinolones were vortexed and then centrifuged before determining the absorbance at 278 nm, Beta-Lactams were vortexed and then centrifuged before determining the absorbance at 230 nm, and Macrolides were vortexed and then centrifuged before determining the absorbance at 210 nm using a search tech UV-Visible spectrophotometer as described by Granados-Chinchilla and Rodríguez (2017). Tetracycline standards were prepared by using a gradient of known concentrations of tetracycline. The results were represented in micrograms of Tetracycline per gram of feed (ug TET/g). Similarly, Ciprofloxacin standards were prepared using a gradient of known concentrations of Ciprofloxacin, with results represented in micrograms of Ciprofloxacin equivalent per gram of feed (ug CIPRO Eq/g). Amoxicillin standards were also prepared using a gradient of known concentrations of Amoxicillin, with results represented in micrograms of Amoxicillin equivalent per gram of feed (ug AMOX Eq/g). Lastly, Azithromycin standards were prepared using a gradient of known concentrations of Azithromycin, with results represented in micrograms of Azithromycin equivalent per gram of feed (ug AZITHRO Eq/g).

2.2.2 Farm Record Data Collection

In this study, farm records from the three different fish farms were collected to assess the level of antibiotics residues in the feeds consumed by the fish over a production cycle of 6 months. Farm 1 stocked 2000 catfish and fed them with 2400kg of feed, which is equivalent to 160 bags. Farm 2, on the other hand, stocked 3000 catfish and fed them with 3600kg of feed, equivalent to 240 bags. Lastly, Farm 3 stocked 2500 catfish and fed them with 3000kg of feed, equivalent to 200 bags.

2.3 Data Analysis:

Data obtained from this study were expressed as mean \pm SEM and subjected to One-way Analysis of Variance (ANOVA) using the IBM Statistical package for Social Science (SPSS) version 29.0.2, Differences were considered significant at $p < 0.05$.

III. RESULTS AND DISCUSSION

3.1 Antibiotic Residues in the nine (9) Fish Feed Samples from the Three Fish Farms:

The results of the antibiotics residues in the sampled fish feeds are shown in Table 1. The results revealed that all the nine fish feed samples from the three different fish farms contained varying concentrations of commonly used antibiotics. Antibiotics detected in the feed samples were Tetracycline, Ciprofloxacin, Amoxicillin and Azithromycin which are widely used in aquaculture. The results showed that the mean antibiotic residue concentrations were higher in the order; tetracycline > ciprofloxacin > amoxicillin > azithromycin with mean concentrations of 85.40 ± 1.78 ug/g, 0.031 ± 0.002 ug/g, 0.019 ± 0.002 ug/g and 0.010 ± 0.001 ug/g respectively. In this study, Tetracycline residue were the highest among the antibiotics tested. Samples G and H contained the highest levels of Tetracycline residues and sample I contained the lowest level of Tetracycline residue level. However, sample D contained the highest level of Quinolone residue level while sample I contained the lowest level of Quinolone residue level. However, sample F contained the highest level of Beta-lactam residue level while sample C contained the lowest level of Beta-lactam residues. Sample E contained the highest level of Macrolide residue level while sample I contained the lowest level of Macrolide residue level.

TABLE 1
THE LEVELS OF ANTIBIOTICS DETECTED IN THE NINE (9) FISH FEED SAMPLE

Treatment	Tetracycline residue (ug TeT Eq/g)	Quinolone residues (ug CIPRO Eq/g)	Beta-lactams residues (ug AMOX Eq/g)	Macrolides residues (ug AZITHRO Eq/g)	Total antibiotics residue
Sample A	87.46± 0.08	0.030± 0.05	0.013±0.01	0.007±0.03	87.51±0.14
Sample B	87.79± 0.01	0.032± 0.02	0.026±0.01	0.013±0.06	87.86±0.01
Sample C	84.85± 0.08	0.027±0.08	0.010±0.03	0.007±0.09	84.89±0.17
Sample D	82.60± 1.91	0.042± 0.01	0.013±0.05	0.006±0.04	82.66±1.93
Sample E	78.63 ±1.91	0.029± 0.03	0.016±0.04	0.019±0.02	78.69±1.93
Sample F	87.37 ±1.38	0.036±0.02	0.032 ±0.02	0.009±0.05	87.44±1.41
Sample G	92.45 ±0.48	0.033±0.05	0.022±0.02	0.016±0.06	92.52±0.50
Sample H	90.92 ±1.03	0.029±0.01	0.019±0.01	0.012±0.05	90.98±1.07
Sample I	76.51 ±1.52	0.022±0.02	0.018±0.01	0.002±0.01	76.55±1.54
Mean	85.40±1.78	0.031±0.002	0.019±0.002	0.010±0.001	
MRL Values by (EU)	0.1ug/g	0.05ug/g	0.05ug/g	0.05ug/g	

MRL = Maximum Residues Limit; EU = European Union (2010).
MRL = Maximum Residues Limit; EU = European Union (2010)

It was also observed in Table 1 that tetracycline had the highest residue concentrations in all the analyzed fish feed samples and this may be as a result of its frequent use in aquaculture feed while macrolide (Azithromycin) had the lowest residue concentrations. Tetracyclines are widely used for both the treatment and prevention of diseases in animals as well as for the promotion of rapid animal growth and weight gain (Pérez-Rodríguez *et al.*, 2018). The antibiotic type detected in the fish feeds were very similar to those reported by Tang *et al.*, (2020) for cultured fish in China indicating that antibiotics were used in aquaculture production. Antibiotics are commonly used in animal husbandry as growth promoters, with Tetracyclines being one of the most widely utilized antibiotic. Tetracyclines are of particular importance due to their extensive use in the industry, surpassing the quantities applied of almost every other antibiotic group (Granados-Chinchilla and Rodríguez, 2017). Tetracyclines are a family of compounds frequently employed due to their broad spectrum of activity as well as their low cost, compared with other antibiotics. However, Lui *et al.* (2017) reported that macrolides are also commonly used as feed additives in fish cultivation for growth promotion and disease prevention and treatment. The mean concentrations of Tetracycline were higher than the maximum residue limit (MRL) while Ciprofloxacin, Amoxicillin and Azithromycin observed in the fish feed samples were all below the MRL as shown in Table 1. Antibiotics can cause serious effects on human health which have led to the introduction of maximum residue limits (MRL) in food safety legislation (EU, 2019). The wide and sometimes inappropriate uses of antibiotics in aquaculture feed have resulted in the presence of the drug residue above the maximum residue limits (MRLs) in fish produced with potential health implications such as nausea, sore throat, diarrhoea and serious illnesses (Okocha *et al.*, 2018). Furthermore, the lack of Maximum Residue Limits (MRLs) for some antibiotics complicates the assessment of antibiotic residues in aquaculture products. The results of this study can serve as a reference for the development of quality standards for aquatic products, particularly in the absence of established MRLs for certain antibiotics. It is important for fish farmers and feed manufacturers to adhere to MRLs to ensure the safety of consumers and to comply with regulatory requirements. Monitoring and testing of fish feeds for antibiotic residues should be conducted regularly to ensure compliance with MRLs and to prevent the presence of harmful residues in fish products.

TABLE 2
MEAN ANTIBIOTICS RESIDUES IN FARM 1, FARM 2, AND FARM 3, IN DIFFERENT FEED SAMPLES FOR ONE PRODUCTION CYCLE OF 6 MONTHS

Sample Sites	Samples	Tetracycline	Ciprofloxacin	Amoxicillin	Azithromicin
	(Treatment)	(ug/g)	(ug/g)	(ug/g)	(ug/g)
Farm 1	A	209.09±1.01	0.72±0.31	0.0312±0.03	0.0168±0.01
	B	198.24±0.23	0.1008±0.12	0.0312±0.03	0.0144±0.07
	C	188.712±0.32	0.0696±0.01	0.0384±0.05	0.0456±0.02
Farm 2	D	254.55±1.32	0.081±0.03	0.03±0.03	0.021±0.05
	E	262.11±1.44	0.108±0.05	0.096±0.04	0.027±0.03
	F	277.35±0.45	0.099±0.07	0.066±0.03	0.048±0.04
Farm 3	G	316.008±1.32	0.115±0.16	0.094±0.04	0.047±0.03
	H	327.312±1.97	0.1044±0.12	0.0684±0.07	0.0432±0.02
	I	275.436±0.54	0.0792±0.04	0.0648±0.01	0.0072±0.03

Tables 2 shows the concentration of antibiotics that cultured fish consumed in six months of production cycle. The presence of antibiotics residues in feed samples observed in this study posed a significant risk to the quality and safety of fish. The finding of this study suggested that these feed samples should not be used for a long period in fish production, as they may lead to the accumulation of antibiotics in cultured fish. The continuous ingestion of antibiotic-laden feed enhances build-up of drugs residues in edible tissues of fish. Long-term exposure of fish to feed containing antibiotics could be systemically toxic to consumers and has a direct negative impact on the complex microflora that inhabits the human gastrointestinal system with potentially adverse implications (Monteiro and Andrade, 2018). The consumption of such fish may result in human health risks and conditions such as drug hypersensitivity reactions, the disruption of normal intestinal flora and carcinogenic, mutagenic and teratogenic effects (Okocha *et al.*, 2018). The daily consumption of aquatic products may lead to exposure to various levels of low-dose antibiotics which can lead to antimicrobial resistance in fish. The presence of antibiotics residues in fish feeds is a major concern as it can contribute to the development and transmission of antibiotic-resistant bacteria through the food chain (Chowdhury *et al.*, 2022). Antibiotic resistance is of great public health concern because the antibiotic-resistant bacteria associated with fish may be pathogenic to humans, easily transmitted to humans via food chains, and widely disseminated in the environment via their wastes (Manyi-Loh *et al.*, 2023). These may cause complicated, untreatable, and prolonged infections in humans, leading to higher healthcare cost and sometimes death. Residues of antimicrobials in feed have received much attention in recent years because of growing food safety and public health concerns. Their presence in food of animal origin constitutes socioeconomic challenges in international trade in animal and animal products (Lee *et al.*, 2023). The major public health significance of antimicrobial residues includes the development of antimicrobial drug resistance, hypersensitivity reaction, carcinogenicity, mutagenicity, teratogenicity, bone marrow depression, and disruption of normal intestinal flora. In order to ensure food safety and avoid exposure to these substances, national and international regulatory agencies have established tolerance levels for authorized veterinary drugs (Pérez-Rodríguez *et al.*, 2018). In order to ensure the correct usage and dosage of antibiotics in food and feedstuffs (Granados-Chinchilla and Rodríguez, 2017), the accurate assessment of antibiotic residues is essential for determining the levels of these substances in feed and food products, as well as for evaluating potential risks to human health.

TABLE 3
MEAN ANTIBIOTICS RESIDUES IN FARM 1, FARM 2 AND FARM 3 IN DIFFERENT SAMPLES OF FEED THAT A FISH WILL CONSUME FOR A PRODUCTION CYCLE OF 6 MONTHS

Sample sites	Samples (Treatment)	Tetracycline (ug/g)	Ciprofloxacin (ug/g)	Amoxicillin (ug/g)	Azithromycin (ug/g)
Farm 1	A	0.105±0.03	0.000036±0.02	0.000016±0.08	0.000084±0.02
	B	0.099±0.05	0.000052±0.05	0.000016±0.08	0.000072±0.04
	C	0.094±0.07	0.000034±0.01	0.000019±0.05	0.000022±0.05
Farm 2	D	0.105±0.14	0.000052±0.09	0.000031±0.03	0.000016±0.01
	E	0.109±0.16	0.000034±0.01	0.000023±0.09	0.000014±0.02
	F	0.092±0.05	0.000024±0.02	0.000021±0.03	0.000024±0.06
Farm 3	G	0.102±0.15	0.000032±0.03	0.000012±0.04	0.000084±0.02
	H	0.105±0.17	0.000043±0.06	0.000038±0.07	0.000018±0.05
	I	0.111±0.13	0.000039±0.05	0.000026±0.01	0.000019±0.07

The levels of antibiotics detected in the feeds were relatively low, however, the daily consumption of these feeds by the fish as shown in Table 3 and can lead to the accumulation of antibiotics in their tissues. This continuous exposure to antibiotics can potentially result in the development of antibiotic resistance in the fish and in the environment. In many countries, aquaculture production systems are not separated from the environment leading to the accumulation of antimicrobial residues in the waters used for fish farming and adjacent waters affecting wild fish, plankton and sediments (Lulijwa *et al.*, 2020). Yang *et al.* (2021) reported that antibiotics have become widespread in the environment due to their extensive and long-term use, influencing both human health and environment, due to the emergence of antibiotic resistance. Studies by Erofeeva *et al.* (2021) have shown that a significant percentage of all manufacturers add antibiotics to healthy animals' feed to prevent, rather than to cure diseases. The presence of antibiotics residues in fish feeds is a major concern as it can contribute to the development of antibiotic-resistant bacteria (Ghimpețeanu *et al.*, 2022). A number of factors can determine the best alternatives to antibiotics to be used within an aquaculture system especially when compared to FAO/WHO and European Union maximum acceptable limit. Good animal husbandry practices as well as the use of alternatives to antibiotics such as vaccination, probiotics, phage therapy, and essential oils are recommended panaceas to reducing the use of antimicrobial residues in aquaculture feed (Okocha *et al.*, 2018). Strategies to reduce or limit the therapeutic use of antibiotics in animals through improved animal nutrition, improved living conditions and waste management, biosecurity measures, and improvement in animals' natural immunity can result in infection prevention and control (Manyi-Loh *et al.*, 2023). Ultimately, the use of antibiotics in food-producing animals should be limited by incorporating possible alternative substances, including probiotics, prebiotics, and plant-derived or crude plant extracts for the treatment and prevention of diseases (Guetiya-Wadoum *et al.*, 2016). To assure food safety, continuous monitoring and identification of risk factors in aquatic feeds and products are crucial, particularly in underdeveloping and developing nations. This therefore calls for strengthening of regulations that direct antibiotic manufacture, distribution, dispensing, and prescription, hence fostering antibiotic stewardship (Manyi-Loh *et al.*, 2023). Feed manufacturers and authorities should establish procedures and instructions for the effective and safe use of authorized and prescribed veterinary medicinal products. EU (2019) established rules for the authorization of use of veterinary medicinal products such as antibiotics in feeding stuffs, including the manufacture, distribution, advertising, and surveillance of such products. Feed business operators, which handle manufacturing, storing, transporting, or placing medicated feed and intermediate products on the market, must be authorized by the competent authority, in accordance with the authorization system, to ensure both the safety of the feed and the traceability of the products (Ghimpețeanu *et al.*, 2022).

IV. CONCLUSION AND RECOMMENDATION

Antibiotics residues in fish feeds used for aquaculture have become a growing concern due to their potential impact on human health and the environment. This study evaluated the levels of antibiotics residues in nine fish feeds used in three aquaculture farms in Awka, Anambra State. The results revealed that all the nine fish feed samples from the three different fish farms contained varying concentrations of commonly used antibiotics. The result showed that the mean antibiotic residue concentrations were higher in the order; Tetracycline >Ciprofloxacin > Amoxicillin > Azithromycin. It was also observed that

Tetracycline had the highest residue concentrations in all the analyzed fish feed samples and this may be as a result of its frequent use in aquaculture feed while Azithromycin had the lowest residue concentrations. The mean concentrations of Tetracycline were higher than the maximum residue limit (MRL) while Ciprofloxacin, Amoxicillin and Azithromycin observed in the fish feed samples were below the MRL. To address this issue, it is recommended that fish farmers should carefully monitor the amount of antibiotics in the feeds they use and practice responsible antibiotic use when necessary. Fish feed producers should also adhere to standards for antibiotics in fish feeds to protect the safety of the food chain and minimize antimicrobial residues. Responsible use of antibiotics in aquaculture under veterinary supervision is crucial to ensure the safety of aquaculture products. Additionally, education and training should be provided to fish farmers on the proper use of antibiotics, including dosage and withdrawal periods, to reduce the risk of antibiotic resistance. Research on novel feed additives, such as the inclusion of probiotics, prebiotics and medicinal plants in fish feeds, should be conducted to reduce feed costs, improve digestibility, and prevent residual effects of antibiotics in fish tissues that may impact human health.

ACKNOWLEDGEMENT

The authors wish to express their gratitude to the fish farmers in Awka, Anambra State, for their cooperation and support in allowing the collection of fish feed samples and farm records from their fish farms. We also acknowledge the assistance of the laboratory technicians and research assistants who helped with sample preparation and analysis.

Ethical Approval: Not Applicable.

Conflict of Interest: The study was carried out without any conflict of interest.

AUTHORS' CONTRIBUTIONS

OTF conducted the fieldwork and sample collection, in various fish farm locations in Awka, Awka South Anambra State and prepared the initial manuscript. OTF and NCG performed the laboratory procedures, data analysis and interpretation. ICF designed the study, supervised the research and corrected the initial manuscript. All authors reviewed and approved the manuscript for submission.

ORCID NUMBERS

Osegbo Toochukwu Franklin NO. 0009-0004-7938-009X

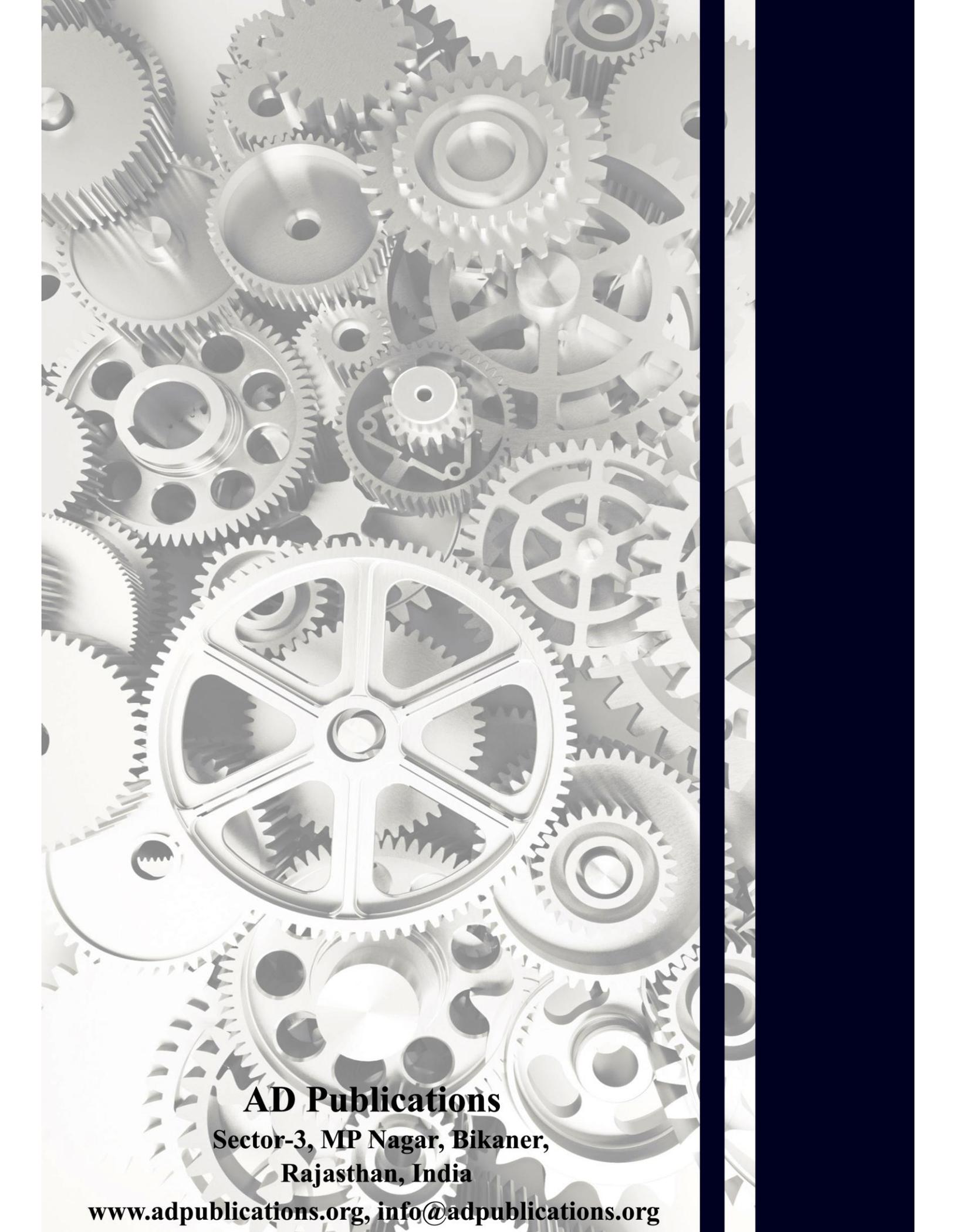
Ikeogu Chika Florence NO.0000-0002-2564-2785

Nwankwo Chidimma Gift NO. 0009-0005-3251-721X

REFERENCES

- [1] Bojarski, B., Kot, B. and Witeska, M. (2020). Antibacterials in Aquatic Environment and Their Toxicity to Fish. *Pharmaceuticals*, 13, 189. doi:10.3390/ph13080189www.
- [2] Chowdhury, S., Rheman, S., Debnath, N., Delamare-Deboutteville, J., Akhtar, Z., Ghosh, S., Parveen, S., Islam, K., Islam, A., Rashid, R., Khan, Z.H., Rahman, M., Chadag, V.M. and Chowdhury, F. (2022). Antibiotics usage practices in aquaculture in Bangladesh and their associated factors. *One Health*, 15, 2352-7714. <https://doi.org/10.1016/j.onehlt.2022.100445>.
- [3] Erofeeva, V., Zakirova, Y., Yablochnikov, S., Prys, E. and Prys, I. (2021). The Use of Antibiotics in Food Technology: The Case Study of Products from Moscow Stores. *E3S Web Conf.*, 311, 10005.
- [4] European Union (EU). (2010). Commission Regulation (EU) No 37/2010 of 22 December 2009 on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin. *Official Journal of the European Union*, 50, 1–72.
- [5] EU. (2019). Regulation of the European Parliament and of the Council of 11 December 2018 on Veterinary Medicinal Products and Repealing Directive 2001/82/EC (Text with EEA Relevance); European Commission: Brussels, Belgium, 2018; Volume 4.
- [6] Ghimpețeanu, O.M., Pogurschi, E.N., Popa, D.C., Dragomir, N., Drăgotoiu, T., Mihai, O.D. and Petcu, C.D. (2022). Antibiotic Use in Livestock and Residues in Food—A Public Health Threat. *A Review. Foods*, 11(10), 1430. <https://doi.org/10.3390/foods11101430>.
- [7] Granados-Chinchilla, F. and Rodríguez, C. (2017). Tetracyclines in Food and Feedingstuffs: From Regulation to Analytical Methods, Bacterial Resistance, and Environmental and Health Implications. *J Anal Methods Chem.*, 2017, 1315497. doi: 10.1155/2017/1315497.
- [8] Guetiya-Wadoum, R.E., Zambou, N.F., Anyangwe, F.F., Njimou, J.R., Coman, M.M., Verdenelli, M.C., Cecchini, C., Silvi, S., Orpianesi, C., Cresci, A. and Colizzi, V. (2016). Abusive use of antibiotics in poultry farming in Cameroon and the public health implications. *Br. Poult. Sci.*, 57, 483–493. doi: 10.1080/00071668.2016.1180668.
- [9] Lee, P.T., Liao, Z.H., Huang, H.T., Chuang, C.Y. and Nan, F.H. (2020). β -glucan alleviates the immunosuppressive effects of oxytetracycline on the non-specific immune responses and resistance against *Vibrio alginolyticus* infection in *Epinephelus fuscoguttatus* × *Epinephelus lanceolatus* hybrids. *Fish and Shellfish Immunology*, 100, 467-475.

- [10] Liu, X., Steele, J.C. and Meng, X.Z. (2017). Usage, residue, and human health risk of antibiotics in Chinese aquaculture: A review. *Environmental Pollution*, 223, 161–169. <https://doi.org/10.1016/J.ENVPOL.2017.01.003>.
- [11] Long, Y., Li, B. and Liu, H. (2018). Analysis of Fluoroquinolones Antibiotic Residue in Feed Matrices Using Terahertz Spectroscopy. *Appl. Opt.*, 57, 544–550.
- [12] Lulijwa, R., Rupia, E.J. and Alfaro, A.C. (2020). Antibiotic use in aquaculture, policies and regulation, health and environmental risks: A review of the top 15 major producers. *Review in Aquaculture*, 12, 640–663. <https://doi.org/10.1111/raq.12344>.
- [13] Miranda, C.D., Godoy, F.A. and Lee, M.R. (2018). Current Status of the Use of Antibiotics and the Antimicrobial Resistance in the Chilean Salmon Farms. *Front. Microbiol.*, 9, 1284.
- [14] Monteiro, S. and Andrade, G.M. (2018). Antibiotic residues and resistant bacteria in aquaculture. *Pharmaceutical and Chemical Journal*, 5, 127–147.
- [15] Monteiro, S.H., Andrade, G.M., Fabiana Garcia, F. and Pilarski, F. (2018). Antibiotic Residues and Resistant Bacteria in Aquaculture. *The Pharmaceutical and Chemical Journal*, 5(4), 127-147.
- [16] Ogunkalu, O.A. (2019). Effects of Feed Additives in Fish Feed for Improvement of Aquaculture.
- [17] Onomu, A.J. and Okuthe, G.E. (2024). The Role of Functional Feed Additives in Enhancing Aquaculture Sustainability. *Fishes* 2024, 9, 167. <https://doi.org/10.3390/fishes9050167>.
- [18] Okeke, E.S., Chukwudozie, K.I., Nyaruaba, R., Ita, R.E., Oladipo, A., Ejeromedoghene, O., Atakpa, E.O., Agu, C.V. and Okoye, C.O. (2022). Antibiotic resistance in aquaculture and aquatic organisms: a review of current nanotechnology applications for sustainable management. *Environmental Science Pollution Research*, 29, 69241–69274. <https://doi.org/10.1007/s11356-022-22319-y>.
- [19] Okocha, R.C., Olatoye, I.O. and Adedeji, O.B. (2018). Food safety impacts of antimicrobial use and their residues in aquaculture. *Public Health Review*, 39, 21. <https://doi.org/10.1186/s40985-018-0099-2>.
- [20] Okon, E.M., Okocha, R.C., Adesina, B.T., Ehigie, J.O., Alabi, O.O., Bolanle, A.M., Matekwe, N., Falana, B.M., Tiamiyu, A.M., Olatoye, I.O. and Adedeji, O.B. (2022). Antimicrobial resistance in fish and poultry: Public health implications for animal source food production in Nigeria, Egypt, and South Africa. *Frontier Antibiotic*, 1, 1043302. doi: 10.3389/frabi.2022.1043302.
- [21] Pathmalal, M. (2018). Heavy use of antibiotics in aquaculture: emerging human and animal health problems: a review Sri Lanka. *Journal of Aquatic Science*, 23, 13–27. <https://doi.org/10.4038/sljas.v4023i4031.7543>.
- [22] Pérez-Rodríguez, M., Pellerano, R.G., Pezza, L. and Pezza, H.R. (2018). An overview of the main foodstuff sample preparation technologies for tetracycline residue determination. *Talanta*, 182, 1-21. doi: 10.1016/j.talanta.2018.01.058
- [23] Shao, Y., Wang, Y., Yuan, Y. and Xie, Y. (2021). A systematic review on antibiotics misuse in livestock and aquaculture and regulation implications in China. *Science of the Total Environment*, Vol 798, 149205. doi: 10.1016/j.scitotenv.2021.149205.
- [24] Tang, J., Wang, S., Tai, Y., Tam, N. F., Su, L., Shi, Y., Luo, B., Tao, R., Yang, Y. and Zhang, X. (2020). Evaluation of Factors Influencing Annual Occurrence, Bioaccumulation, and Biomagnification of Antibiotics in Planktonic Food Webs of a Large Subtropical River in South China. *Water Research*, 170, 115302. doi:10.1016/j.watres.2019.115302.
- [25] Van Boeckel, T.P., Glennon, E.E., Chen, D., Gilbert, M., Robinson, T.P. and Grenfell, B.T. (2017). Reducing antimicrobial use in food animals. *Science*, 357(6358), 1350–1352. doi: 10.1126/science.aao1495.
- [26] Yang, X., Chen, Z., Zhao, W., Liu, C., Qian, X., Zhang, M., Wei, G., Khan, E., Hau Ng, Y. and Sik Ok, Y. (2021). Recent Advances in Photodegradation of Antibiotic Residues in Water. *Chem. Eng. J.*, 405, 126806.



AD Publications

**Sector-3, MP Nagar, Bikaner,
Rajasthan, India**

www.adpublications.org, info@adpublications.org