eksis jurnal riset ekonomi & bisnis
E-ISSN : 2549-6018
P-ISSN : 1907-7513

FORECASTING THE HUMAN DEVELOPMENT INDEX BASED ON SOCIAL AND ECONOMIC FACTORS IN INDONESIA USING THE ARIMAX MODEL

Fairuz Azizah Nurfayza, Kevin Simamora, Nisrina Luthfia, Fitri Kartiasih*
Program Studi DIV Statistika, Politeknik Statistika STIS
Correspondence : fkartiasih@stis.ac.id

Submitted: 23 March 2025, Accepted: 18 April 2025, Published: 23 April 2025

Abstract

The Human Development Index (HDI) plays an important role in Indonesia's development policy. Through the HDI, related social and economic factors can be identified, in addition to its main dimensions, namely long and healthy life, level of knowledge, and decent standard of living. This study aims to analyze the influence of social and economic factors outside the dimensions of the HDI on increasing the HDI and predict its impact in the future. The analysis method used is the ARIMAX model, which allows forecasting by considering exogenous variables that affect the variables being analyzed. The results of this study conclude that social and economic factors have a significant influence on human development with the formation of the ARIMAX(3,0,0) model. After forecasting, the model is able to capture historical patterns well, although there are fluctuating patterns. The forecast shows a downward trend in 2026 and 2033, which is likely influenced by the pattern of the previous year, 2020.

Keywords: Human Development Index (HDI), Infant Mortality Rate (IMR), Inflation, Open Unemployment Rate (OPR), Poor Population

A. INTRODUCTION

The Human Development Index (HDI), developed by the United Nations Development Programme (UNDP), is a tool used to measure the level of human development in a country. The three main dimensions used to measure the Human Development Index include a long and healthy life, knowledge, and a decent standard of living (Statistics Indonesia, 2018). HDI provides a comprehensive picture of people's welfare and quality of life. The importance of studying HDI lies in its ability to reflect the success and failure of development policies, as well as an evaluation tool that assists governments and international organizations in formulating strategies to improve welfare and alleviate poverty. Understanding and analyzing the HDI allows us to identify areas for improvement and direct resource allocation more effectively in order to achieve sustainable and inclusive development goals.

The progress of human development in Indonesia can be seen from the significant increase in the Human Development Index (HDI) in recent years. In 2024, Indonesia's HDI reached 75.02, an increase of 0.63 points or 0.85 percent compared to the previous year (Statistics Indonesia, 2024). This increase covers all aspects of the HDI, namely health and longevity, education, and a decent standard of living. The importance of HDI in Indonesia is very large. According to Statistics Indonesia, HDI has several uses in Indonesia, including its role as the main parameter in measuring the success of efforts to improve people's living standards. In addition, HDI also serves as a reflection of a country's level of progress in development. In Indonesia, HDI has a strategic role because it is used as an indicator of the performance of state managers and also to determine the amount of the General Allocation Fund (Statistics Indonesia, 2023).

One of the dimensions of HDI is longevity and healthy living, which is closely related to the infant mortality rate (IMR). Infant mortality rate is the number of deaths of infants under one year of age per 1,000 live births in a population. The World Health Organization (WHO) established this indicator to measure the health and well-being of the Community. Previous research states that infant mortality is considered a barometer of the welfare of a community or country (Alfandi et al., 2022). Infant mortality rates provide an overview of socioeconomic development as well as policy implications (Gonzalez & Gilleskie, 2017). Infant mortality is one of the indicators of successful development in the health sector, so reducing infant mortality is one of the main targets in the Long-Term Development Plan (KEMENKES RI, 2018). So, it can be concluded that the Infant Mortality Rate (IMR) serves as a measure of public health. The level of success of health services and regional health development programs can be reflected through IMR (Statistics Indonesia, 2016). Based on research conducted by Zuhairoh in East Java province, the infant mortality rate is significant and has a negative effect on the Human Development Index (Zuhairoh, 2018). In addition, the open unemployment rate also affects infant mortality; the infant mortality rate (IMR) will increase by 0.467 deaths per 1000 live births if the open unemployment rate increases by 1 percent. This is due to economic and health mechanisms, as well as a decrease in people's income, which impacts access to healthy and nutritious food, especially during the first thousand days of life. This leads to an increase in infant mortality (Putri et al., 2024).

The Human Development Index (HDI) is strongly influenced by the unemployment rate (Sarkodie & Adams, 2020; Jin et al., 2020). One of the most significant economic problems that has a direct impact on people is unemployment. For most people, job loss means psychological distress and a decline in living standards (Mankiw, 2003). Research results show that the levels of unemployment and poverty are closely related (Ningrum et al., 2020; Parolin & Wimer, 2020; Feng et al., 2024). High unemployment will lead to a reduction in people's income, which in turn increases poverty and reduces people's welfare. Unemployment has a negative and significant effect on HDI in Indonesia; when unemployment falls, HDI in Indonesia increases (Hasibuan & Rujiman, 2020). According to data from the Indonesian Statistics, the Open Unemployment Rate (OPR) in August 2024 reached 4.91 percent. Although it has decreased from the previous year, this unemployment rate still indicates a big challenge in efforts to improve the quality of life and welfare of people in Indonesia, which in turn affects the Human Development Index.

According to data from Badan Pusat Statistik, the percentage of poor people in Indonesia in March 2024 was 9.03%. This poverty has a direct impact on the decent living dimension of the HDI, where low income and access to basic needs directly affect the quality of life of individuals. The results of the tests conducted show that poverty has a significant influence on the Human Development Index (Ningrum et al., 2020). The relationship between the Human Development Index (HDI) and poverty is very close, as poverty has a significant impact on various aspects of human development. According to Putra in his book, there is a negative relationship between poverty and welfare (W. Putra, 2019). Poverty includes a lack of assets, social and political organization, knowledge and skills, social networks, financial resources, and information. These deficiencies are seen in the form of malnutrition, lack of access to clean water, inadequate health services, and low levels of education, which ultimately harm welfare. These results are consistent with Pangestika & Widodo (2017) research, which shows that poverty affects HDI. According to research conducted by Tarumingkeng et al. (2021), there is a linear influence between the independent variable poverty and the dependent variable Human Development Index (HDI). In other words, the study found that poverty has a negative and significant impact on HDI, or the value of HDI falls along with the level of poverty, indicating that people's quality of life and welfare become worse.

Other economic variables, such as inflation, also play a significant role in determining people's quality of life. High inflation can reduce purchasing power, increase the cost of living, and hinder people's access to basic needs such as education and health. Previous research conducted in East Nusa Tenggara Province shows that inflation has a negative impact on the Human Development Index (HDI) because inflation reflects problems in the economy, characterized by an increase in the general price of goods and services over a long period. Meanwhile, the supply of goods does not increase or even decrease due to inadequate distribution. This condition causes people's purchasing power to decrease because the price of goods is too high, so people's welfare decreases (Kiha et al., 2021). Thus, inflation has a significant effect on the Human Development Index (HDI), so understanding its impact is very important in an effort to improve people's welfare.

In contrast to previous studies that discuss the relationship of each variable separately, this study presents an update by evaluating the relationship between the variables of infant mortality rate, percentage of poor people, inflation, and open unemployment rate with its effect on the Human Development Index simultaneously, as well as forecasting the Human Development Index (HDI) in Indonesia. This topic is important to analyze because an in-depth understanding of how each of these variables interact and influence the HDI can help the government and policymakers formulate more effective strategies to improve the quality of life. The infant mortality rate and the percentage of poor people reflect health and social welfare conditions, while the open unemployment rate and inflation reflect economic conditions. By conducting comprehensive Human Development Index-related research and forecasting, we can identify significant patterns and trends and take appropriate preventive and corrective actions to achieve sustainable and equitable human development.

This study aims to explore the relationship between the infant mortality rate (IMR), inflation, percentage of poor people, and open unemployment rate (OPR) to the Human Development Index (HDI) in Indonesia, using the ARIMAX model for forecasting. By understanding the variables that affect HDI, which is the basis for determining the level of development achievement of a country's population, we can make more efficient policies to improve the quality of life of the people (Azfirmawarman et al., 2023). This research includes the analysis of economic and social variables that have a significant influence on HDI. Through comprehensive study and forecasting, we can identify important patterns and trends and formulate appropriate policies to address the problems found so as to improve the quality of life and welfare of the people.

B. LITERATURE REVIEW

Indonesia adopts the UNDP's human development measurement by calculating the same index, the Human Development Index (HDI) (Statistics Indonesia, 2024). According to the United Nations Development Programme (UNDP) in 1990, human development is defined as a process that aims to expand choices for each individual in living a life that is considered meaningful. Important dimensions of human development include:

a. The ability to enjoy a long and healthy life (health),
b. know (education),
c. gaining access to the resources necessary to fulfill the needs of a decent life (decent

standard of living).

Therefore, a composite index is created, which consists of various indicators calculated using certain methods to produce the HDI. This composite index is then used to measure the extent of achievement in improving the quality of human life in a measurable and representative manner. The Human Development Index (HDI), also known as the Human Development Index (HDI), has a range of values from 0 to 100. HDI values closer to 100 indicate a better level of human development. The UNDP has placed countries or regions into three categories of human development (HDI). HDI below 50 is considered low; HDI between 51 and 80 is considered medium/medium, and HDI above 81 is considered high.

Statistics Indonesia (2024), in the publication "Human Development Index 2023," referring to UNDP (2015), has summarized the indicators forming the HDI, which include:

a. In the dimension of longevity and healthy living, the indicator used is Life Expectancy at Birth (UHH), which is an estimate of the average age (in years) that a person can live throughout his life.
b. In the knowledge dimension, there are two indicators, namely: i) Expected Years of Schooling (HLS), which reflects people's opportunities to obtain formal education, and ii) Average Years of Schooling, which describes the level of education that human resources have achieved in a region.
c. For the decent standard of living dimension, the indicator used is not the same as the UNDP standard, namely Gross National Income (GNI) per capita (US$ PPP), because the data is not available at the provincial or district/city level. After a lengthy study, an adjusted real expenditure per capita indicator was used. This indicator was chosen after an in-depth study because it is considered capable of reflecting the level of income and welfare of the community across regions and time.

By utilizing these dimensions in the calculation of HDI, Pratowo (2012) emphasizes that HDI is a composite index calculated as a simple average of three indices that describe basic human capacity in expanding their choices. The general formula used is as follows (UNDP, 2004) :

$$HDI=1/3\ (Y_1+Y_2+Y_3)$$

HDI = Human Development Index, Y_1 = Long and Healthy Life Dimension, Y_2= Knowledge Dimension, Y_3 = Decent Standard of Living Dimension

UNDP developed this index to highlight the important role of humans and the resources they possess in the development process. Therefore, this research will examine the influence of social and economic factors, which are part of the important dimensions of development, on the Human Development Index (HDI).

Social and economic factors, which are important dimensions in the formulation of the Human Development Index (HDI), are obtained through various indicators that reflect the condition of a region. In the dynamic table on the Statistics Indonesia website, for example, health and education indicators fall under the category of social factors. In contrast, the decent living standard indicator represents economic factors. The research conducted by Setianingtias et al. (2019) supports this by formulating indicators grouped into social and economic dimensions based on the principles of the Sustainable Development Goals (SDGs), using factor loading analysis as the basis for their study.

Supporting this, D. M. Putra & Ratnasari (2016) examined variables such as the infant mortality rate (IMR), percentage of health complaints, number of health facilities, illiteracy rate, Senior High School Enrollment Rate (APS SMA), teacher-student ratio, school-student ratio, percentage of poor population, per capita GDP, economic growth, and labor force participation rate. Variables such as infant mortality rate (IMR), health complaints, health facilities, illiteracy rate, high school enrollment rate, teacher-student ratio, and school-student ratio are classified as social factors, while other variables, such as the percentage of poor population, GDP per capita, economic growth, and open unemployment rate, fall under economic factors.

In addition, other dependent variables, such as prosperity, the size of local government, inflation, intergovernmental revenue (government revenue from external sources), and population poverty, also show similar clustering. Based on Manik (2013) analysis, these variables are grouped as economic factors, except for prosperity, which is the only social factor. A more advanced approach is seen in the research by Chalid, N. & Yusuf, Y. (2014), where the variables used, such as the poverty rate, unemployment rate, district/city minimum wage, and economic growth rate, are all classified as economic factors.

From these studies, some variables yield significant values affecting the independent variable of the Human Development Index (HDI) and represent the HDI development indicators in terms of social and economic factors used in various important analyses in the research. These variables are exogenous variables which are not IPM building variables, including the following:

The Infant Mortality Rate (IMR) can represent the dimension of longevity and health as a social factor because it is used in the calculation of Life Expectancy (LE) (D. M. Putra & Ratnasari, 2016). IMR refers to the death of infants before they reach the age of one year. It is an important measure of a country's health and well-being (Arini et al., 2024). The calculations of IMR are the total number of infant deaths per 1,000 live births, used to evaluate the quality of healthcare services, particularly in addressing maternal and child health issues. This figure is an important indicator in assessing the success of public health programs in reducing infant mortality, which is greatly influenced by the quality of prenatal and postnatal care. Therefore, this variable is indirectly not an indicator of HDI development, so it can be considered an exogenous variable.

Next, the Open Unemployment Rate (OPR) can represent the dimension of a decent standard of living as an economic factor Chalid & Yusuf, (2014). Unemployment greatly affects the economy, impacting not only individuals and families but also the overall stability of a nation's economy Widyart et al., 2024). Measures are taken by the percentage of the workforce that is not employed but is actively seeking work. The unemployment rate (OPR) has become an important indicator in analyzing the labor market conditions in Indonesia and illustrates the level of workforce engagement in economic activities. The increase in the unemployment rate reflects a significant disparity between the number of job seekers and the availability of job opportunities, which has the potential to affect social and economic stability. Because this variable is indirectly not an indicator of HDI development, it can be considered an exogenous variable.

Inflation can represent the dimension of a decent standard of living as an economic factor because it is used in the calculation of Real Per Capita Expenditure (Manik, 2013). Supported by Kumalasari, M., & Poerwono, D. (2011) per capita expenditure is expenditure adjusted with the consumer price index (CPI), where inflation is calculated using the CPI (Pangesti & Susanto, 2018). Inflation affects the purchasing power of the community and economic stability, where fluctuating price changes can negatively impact social welfare and worsen economic inequality, especially for low-income communities. Therefore, indirectly, this variable is not an indicator of HDI development, but it can be used as an exogenous variable.

According to D. M. Putra & Ratnasari (2016), the Percentage of Poor Population (PPP) can represent the dimension of a decent standard of living as an economic factor because it is used in the calculation of Real Per Capita Expenditure. Apriliyah (2007) added that per capita consumption has a negative and significant impact on the number of poor people. According to Statistics Indonesia, the percentage of the poor population is calculated by including the number of poor people. This indicator is important for measuring socioeconomic inequality, where a high poverty rate reflects significant differences in access to resources and basic services. Therefore, indirectly, this variable is not an indicator of HDI development, but it can be used as an exogenous variable.

Previous research was conducted by D. M. Putra & Ratnasari (2016), who analyzed various variables such as the Infant Mortality Rate (IMR), Senior High School Enrollment Rate (SHSER), the number of health facilities, the percentage of health complaints, the Percentage of Poor Population (PPP), and the Open Unemployment Rate (OPR). The data were grouped by social factors such as the Infant Mortality Rate (IMR), percentage of health complaints, number of health facilities, illiteracy rate (aged 10 and above), school participation rate (high school), teacher-student ratio (high school), and school-student ratio (high school), as well as economic factors such as the Percentage of Poor Population (PPP), per capita GDP, economic growth, percentage of the population aged 15 and above who are working, Open Unemployment Rate (OPR), and labor force participation rate. This study uses the Reg-Log Ridge method and finds that through the backward elimination technique, the best model is produced with an accuracy rate of 97.37%. This model identifies and provides an overview of the relationship between social and economic factors and HDI, which will also be analyzed in this study.

There is an impact of poverty levels and economic growth on the HDI, considering individual heterogeneity and geographical differences in Indonesia, according to Syofya (2018). The impact was analyzed using Multiple Linear Regression (MLR), which showed that poverty and economic growth have a significant influence on the HDI. This research supports the importance of economic factors such as inflation and the unemployment rate in predicting the HDI.

Baihawafi dan Sebayang (2023) also analyzed the influence of the district/city minimum wage (UMK), HDI, and economic growth rate on the unemployment rate in West Java. The analysis adopted panel data regression with the Fixed Effect Model (FEM) approach and then found that the HDI negatively affects the OPR. These findings provide a basis for using the OPR as one of the economic factors to be analyzed in this study, which is in line with the objective of forecasting the HDI using economic and social variables.

These studies provide valuable insights into how social and economic factors influence the HDI and serve as a primary reference in developing the ARIMAX model for HDI forecasting in Indonesia.

C. METHOD

1. Data

This research uses data collected by the Central Bureau of Statistics of Indonesia. This research involves five variables to be analyzed using the ARIMAX model. The Human Development Index (HDI) is the dependent variable in this study, which encompasses aspects of health, education, and per capita income to measure overall quality of life. Meanwhile, four other variables will be used as exogenous variables, namely: Infant Mortality Rate (IMR), Open Unemployment Rate (OPR), Inflation, measured through the Consumer Price Index (CPI), and the Percentage of Poor Population (PPP), calculated based on the poverty line.

Table 1. Variable description

Variable	Variable Name	Unit	Source
HDI	Human Development Index	Percentage	Statistics Indonesia
IMR	Infant Mortality Rate	Death per 1000 Live Births	Statistics Indonesia
OPR	Open Unemployment Rate	Percentage	Statistics Indonesia
Inflation	Inflation	Percentage	Statistics Indonesia
PPP	Percentage of the Poor Population	Percentage	Statistics Indonesia

This research uses annual data covering the period from 1995 to 2024, allowing for a more comprehensive analysis of trends and relationships between variables in the long term, as well as providing a deeper understanding of the socioeconomic factors affecting the Human Development Index in Indonesia.

This research uses the ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables) model, which is an extension of the ARIMA model. The main difference between the ARIMA model and the ARIMAX model lies in the use of exogenous variables.

The ARIMA (Autoregressive Integrated Moving Average) model is a time series forecasting method that combines three main components: autoregressive (AR), integration (I), and moving average (MA). ARIMA only considers the internal variables of the time series in the forecasting process (Riestiansyah et al., 2022). The ARMA model is used to analyze stationary data by combining autoregressive (AR) and moving average (MA) components without integration. On the other hand, the ARIMA model develops ARMA by adding an integration (I) component to handle non-stationary data. The ARMA and ARIMA models can be formulated as follows:

$$Yt=\phi1Yt-1+\cdots+\phi pYt-p+et-\theta1et-1-\cdots-\theta qet-q \quad \textit{(ARMA)}$$

$$\Delta Yt=\mu+\phi1\Delta Yt-1+\cdots+\phi p\Delta Yt-p+et-\theta1et-1-\cdots-\theta qet-q \quad \textit{(ARIMA)}$$

Constraint:

if d = 0, so $\Delta Yt = Yt$
if d = 1, so $\Delta Yt = Yt - Yt - 1$
if d = 2, so $\Delta Yt = \Delta Yt - \Delta Yt - 1 = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2})$

With:

ΔYt : showing the change (differencing) of the variable Yt, which aims to make the data stationary,
μ : constant,
ϕi : parameter for the AR component,
θ_j : parameter for the MA component,
et : error or residual at time t, which is usually assumed to follow a normal distribution.

On the other hand, ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables) combines exogenous variables (independent variables) with internal variables to make more accurate predictions. Research has shown that the use of the ARIMAX model can improve forecasting accuracy by considering relevant external factors. The ARIMAX model is used to analyze time series data that has temporal dependence on the dependent variable and the influence of exogenous variables. This model allows for capturing the relationship between the dependent variable and the exogenous factors that influence it. The ARIMAX model can be formulated as follows (Ahn, E., & Hur, J., 2023) :

$$Y_t=\mu+\sum_{i=1}^{p}\phi_i Y_{t-i}+\sum_{j=1}^{q}\theta_j \epsilon_{t-j}+\epsilon_t \quad \textit{(ARIMA)}$$

ARIMA model with exogenous variables:

$$Y_t = \mu + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{r} \beta_k X_{t-k} + \epsilon_t \quad \textit{(ARIMAX without lag)}$$

$$Y_t = \mu + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{K} \beta_k X_{k,t} + \sum_{l=1}^{s} \gamma_l X_{t-l} + \epsilon_t \quad \textit{(ARIMAX with lag)}$$

With:

Yt : dependent variables such as HDI
X_{t-k} / X_{t-l} : exogenous variables such as IMR, OPR, Inflation, and PPP,
μ : constant,
ϕ_i : parameter for the AR component,
θ_j : parameter for the MA component,
β_k : parameter for the exogenous variable,
$X_{k,t}$: exogenous variable at time t,
s : *lagged order for exogenous variables*,
γ_l : coefficient of the lagged exogenous variable,
et : error or residual at time t, which is usually assumed to follow a normal distribution.

Form the ARIMAX model with the variables used:

$$HDI_t = \mu + \sum_{i=1}^{p} \phi_i HDI_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \beta_1 IMR_t + \beta_2 OPR_t + \beta_3 Inflation_t + \beta_1 PPP_t + \epsilon_t$$
*(ARIMAX without lag)**

$$HDI_t = \mu + \sum_{i=1}^{p} \phi_i HDI_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{l=0}^{s} (\beta_l IMR_{t-1} + \gamma_l OPR_{t-1} + \delta_l Inflation_{t-1} + \zeta_l PPP_{t-1}) + \epsilon_t$$
(ARIMAX with lag)

With:

$\beta_l, \gamma_l, \delta_l, \zeta_l$: lag coefficients for each exogenous variable,
$l = 0$: using the value of the exogenous variable at time t,
$1 > 0$: using the value of the exogenous variable from the previous time.

*) The lag-free model is simpler and remains valid if the lagged effects are not significant in the data (Majka, n.d. and several developments from researchers)

Model Selection and ARIMA (or ARIMAX) Forecasting

According to Sugiantari & Budiantara (2013), as well as several analytical developments from researchers, the steps for ARIMA and ARIMAX analysis are as follows:

1. ARIMA Stationarity Identification: checking whether the data is stationary in mean, variance, and autocorrelation.
2. Identification of Autoregressive (AR) Components, Moving Average (MA), and Exogenous Influence: exogenous variables are included to capture the direct influence on the dependent variable.
3. Identification of ARIMAX Stationarity: checking whether the data is stationary in mean, variance, and autocorrelation. If the data is non-stationary,

transformations such as differencing or logarithm are performed to achieve stationarity.

4. Differencing ARIMAX: performing a differential transformation on non-stationary data to eliminate trends and make it stationary. This step is carried out until the data meets the stationarity requirements.
5. The selection of ARIMAX Model Parameters is done by testing various combinations of AR (p), differencing (d), and MA (q) parameters to determine the most optimal model. The model selection process is carried out by comparing the values of the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), where a model with lower values is considered better.
6. ARIMAX Model Diagnostic Test: evaluating the model by examining the residuals using ACF analysis and statistical tests such as Ljung-Box to ensure the residuals are random and do not exhibit autocorrelation.
7. Forecasting ARIMAX: using a validated ARIMA model to generate future value predictions based on historical data.
8. Model Optimization with AIC and BIC: ensuring that the model used is the most optimal by re-evaluating the AIC and BIC values. The best model is the one with the lowest AIC and BIC values.

The steps can be illustrated in a flowchart. According to Chi L. Y. et. al. (2019), along with several modifications from researchers, the following is a flowchart for ARIMA (or ARIMAX) modeling and forecasting.

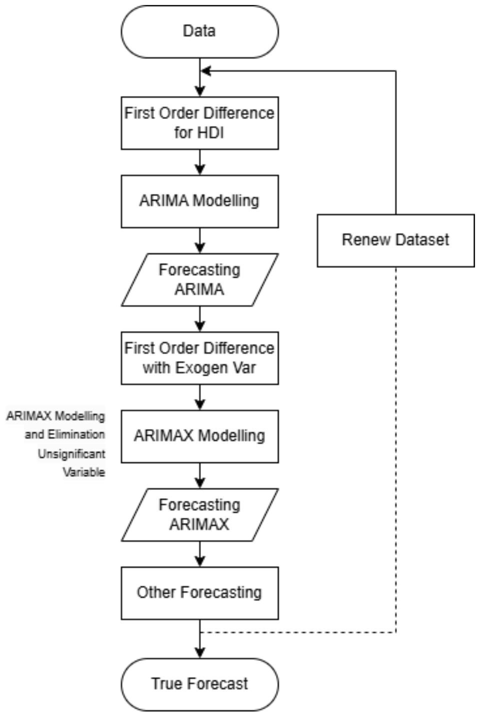

Figure 1. Flowchart of the steps in ARIMA (or ARIMAX) modeling and forecasting

D. RESULTS AND DISCUSSION

1. Overview of human development index, inflation, percentage of poor people, open unemployment rate, and infant mortality rate in Indonesia 1995 - 2024

The method used to collect, process, present, and interpret data quantitatively or in percentages can be referred to as descriptive analysis (Walpole, 1995). Various descriptive forms can be created from different types of data, ranging from visual forms, such as tables and graphs, to data measures, such as measures of central tendency, measures of position, and measures of dispersion (Martias, 2021).

Table 2. Statistics of the analyzed variables

Variable	Min	Max	Mean	Standard Deviation	Coefficient of Variation
HDI	64,3	74,2	70,49	2,3	5,31
Inflation	1,57	77,63	8,48	13,59	184,65
PPP	9,03	24,2	14	4,19	17,55
OPR	4,68	10,28	6,89	1,71	2,93
IMR	16	55	31,5	12,6	157,58

Descriptive statistics of all research variables are presented in Table 1. Of all the variables studied, inflation has the highest data variability, with a coefficient of variation of 184.65%. Meanwhile, the variable with the lowest variability is the open unemployment rate variable, with a coefficient of variation of 2.93%.

Unit Root Test

The stationarity testing of time series data can be conducted through the unit root test. Stationary data has a mean, variance, and autocovariance that remain constant over time. One of the commonly used unit root tests is the Augmented Dickey-Fuller (ADF) test, developed by David Dickey and Wayne Fuller. This test serves to avoid spurious regression that often occurs when two non-stationary variables are regressed against each other (Basuki, 2017). Thus, this unit root test can help improve the accuracy and reliability of time series analysis and forecasting results.

The initial step in the analysis with the ARIMAX model is to conduct a unit root test to determine whether the time series data is stationary. From Table 2, it is known that the ADF test shows that all variables are non-stationary at the level. However, after being transformed into first differences, the KPSS and PP tests indicate that all variables have achieved stationarity. This indicates that the data needs to be transformed into first differences to meet the stationarity assumption in the ARIMAX model.

Table 3. Unit root test results

Variable	Level			First Difference		
	Test	t-stat	prob	t-stat	prob	Result
HDI	ADF	-1,669	0,436	-9,184	0,000	Stationary
Inflation	ADF	-1,216	0,652	-7,709	0,000	Stationary
PPP	ADF	-0,726	0,825	-4,376	0,002	Stationary
OPR	ADF	-1,210	0,656	-5,120	0,000	Stationary
IMR	ADF	-0,714	0,829	-6,728	0,000	Stationary

ARIMAX Modelling

The ARIMAX modeling process begins by dividing the data that has undergone the first differentiation process into two parts: the training data, which includes the first 19 observations out of 29, and the testing data, which consists of the last 10 observations. This division aims to ensure that the built model can be evaluated on data that is not used in the training phase. Training data (training) is used to identify and build the model, while testing data (test) serves to evaluate the accuracy and predictive capability of the model on data that has never been seen before. This approach aims to ensure that the model can perform well not only on the data used during training but also on new data.

After data partitioning, the next step is the identification of the ARIMAX model. At this stage, the search for the optimal combination of AR (AutoRegressive), MA (Moving Average), and exogenous variables parameters is conducted. The models tested include various combinations of AR and MA orders as well as the selection of relevant exogenous variables, such as inflation, IMF, and OPR.

Table 4. Identification of the ARIMA model

Model	**AIC**	**BIC**
ARIMA(0,0,0)	-75.78096	-70.11433
ARIMA(1,0,0)	-79.08692	-72.47584
ARIMA(2,0,0)	-81.37033	-73.81482
ARIMA(3,0,0)	-85.62015	-75.13056

Based on the AIC and BIC evaluation results in Table 4, the ARIMA(3,0,0) model was selected as the best model. Therefore, the ARIMA(3,0,0) model will be used as the basis for developing the ARIMAX model by incorporating exogenous variables. The parameters of the ARIMAX model used are p=3, d=0, and q=0, in accordance with the results of the previous model selection. The estimation results of the ARIMAX(3,0,0) model for various combinations of exogenous variables are presented in the table below:

Table 5. Combination of Exogenous Variables for ARIMAX Model Estimation

Exogenous Variable	AIC	BIC	Parameter Significance (α=0.05)
diff(inf), diff(ppp)	-82.11398	-75.50291	not significant
diff(inf), diff(opr)	-81.98782	-75.37674	not significant
diff(inf), diff(imr)	-87.25980	-80.64873	all significant
diff(ppp), diff(opr)	-80.08119	-73.47012	not significant
diff(ppp), diff(imr)	-82.55375	-75.94267	diff(imr) significant
diff(opr), diff(imr)	-82.51577	-75.90470	diff(imr) significant
diff(inf), diff(ppp), diff(opr)	-80.42753	-72.87201	not significant
diff(inf), diff(ppp), diff(imr)	-85.52164	-77.96613	diff(inf), diff(imr) significant
diff(inf), diff(opr), diff(imr)	-85.33482	-77.77930	diff(inf), diff(imr) significant
diff(ppp), diff(opr), diff(akb)	-80.55806	-73.00255	diff(imr) significant
diff(inf), diff(ppp), diff(opr), diff(imr)	-83.63051	-75.13056	diff(inf), diff(imr) significant

Table 5 (five) shows that the ARIMAX(3,0,0) model, with the addition of the exogenous variables of inflation and infant mortality rate, provides the best estimation results. This model has the lowest AIC and BIC values, indicating that it is more parsimonious and has better predictive ability compared to the other models that have been developed. Therefore, the ARIMAX(3,0,0) model with the two exogenous variables was chosen as the final model and can be expressed as follows:

$$\log(\widehat{HDI}_t) = 0.0002 + \log(HDI_{t-1}) - 1.4788(\log(HDI_{t-1}) - \log(HDI_{t-2})) - 1.4391(\log(HDI_{t-2}) - \log(HDI_{t-3})) - 0.6405(\log(HDI_{t-3}) - \log(HDI_{t-4})) - 0.0127(\log(inf_t) - \log(inf_{t-1})) - 0.1086(\log(imr_t) - \log(imr_{t-1})) + \varepsilon_t$$

Where $\log(HDI_t)$ is the logarithm of the Human Development Index (HDI) value at time t, while $\log(HDI_{t-1})$, $\log(HDI_{t-2})$, $\log(HDI_{t-3})$, and $\log(HDI_{t-4})$ are the logarithms of the HDI values at previous times up to lag 4, the variables $\log(inf_t)$, and $\log(inf_{t-1})$ are the logarithm of the inflation value at time t and lag 1. In contrast, $\log(imr_t)$ and $\log(imr_{t-1})$ are the logarithm of the Infant Mortality Rate at time t and lag 1, and ϵt is the error at period t.

The coefficients in the ARIMAX model indicate that the logarithm of the HDI is significantly influenced by lag values up to the third lag, with the largest impact coming from the first lag (-1.4788). The inflation coefficient (-0.0127) indicates that a 1% increase in inflation will decrease the logarithm of the HDI by 0.0127, while the baby birth rate coefficient (-0.1086) shows that a 1% increase in the baby birth rate will reduce the logarithm of the HDI by 0.1086. This emphasizes that inflation and the birth rate have a negative relationship with the development of the HDI.

After obtaining the optimal ARIMAX(3,0,0) model, the next step is to evaluate the quality of the model. In addition to conducting the Ljung-Box test to detect autocorrelation and the Kolmogorov-Smirnov test to check for normality, visualizing the residual plot is also very important. By observing the residual plot, we can visually examine whether any unusual patterns or outliers may indicate problems with the model. If the residual plot shows a random pattern and is distributed around zero, then the model can be considered to have met the classical assumptions.

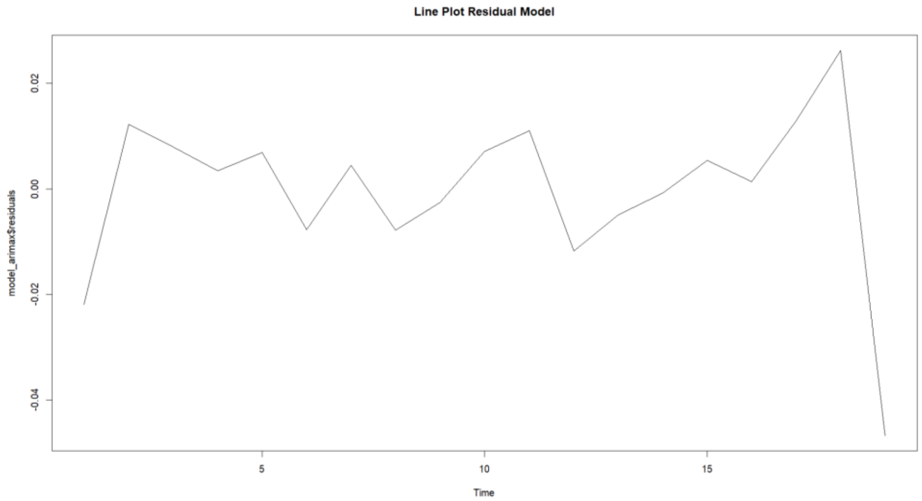

Figure 3. ARIMAX model residual line graph

Residual plot analysis shows that the residuals are randomly scattered around zero without any systematic pattern. This indicates that the ARIMAX(3,0,0) model has successfully captured all deterministic components in the data, resulting in residuals that are white noise. Thus, the model has met the classical assumptions and can be considered a good model for explaining the observed phenomenon.

The results of this visualization can then be reinforced with the Ljung-Box test, which statistically tests for the presence of autocorrelation in the residuals. The diagnostic test yields a p-value from the Ljung-Box test of 0.3329, which is greater than the significance level; thus, it can be concluded that there is no evidence of autocorrelation in the residuals. Additionally, the Kolmogorov-Smirnov test is used to test the normality of the residuals. The p-value from the Kolmogorov-Smirnov test obtained was 0.3805, which is greater than the significance level. Therefore, it can be concluded that the residuals follow a normal distribution. (To further confirm the goodness of the model, evaluation can be conducted by examining the forecasting errors of the model built with the test data, such as Mean Error (ME), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Scaled Error (MASE). Based on the model evaluation results using the test data, the Mean Error (ME) value is 0.000115118, the Root Mean Squared Error (RMSE) is 0.001745426, the Mean Absolute Error (MAE) is 0.00114217, and the Mean Absolute Scaled Error (MASE) is 0.5829929. The low error values indicate that the ARIMAX(3,0,0) model performs well in forecasting. Therefore, if the visualization of the residual plot, the results of the statistical tests, and the error evaluation meet the classical assumptions, then the obtained ARIMAX(3,0,0) model can be considered valid and can be used for forecasting purposes.

Forecasting

Here are the results of the HDI forecast for the next 10 years using the ARIMAX (3,0,0) model with the addition of the exogenous variables of inflation and infant mortality rate:

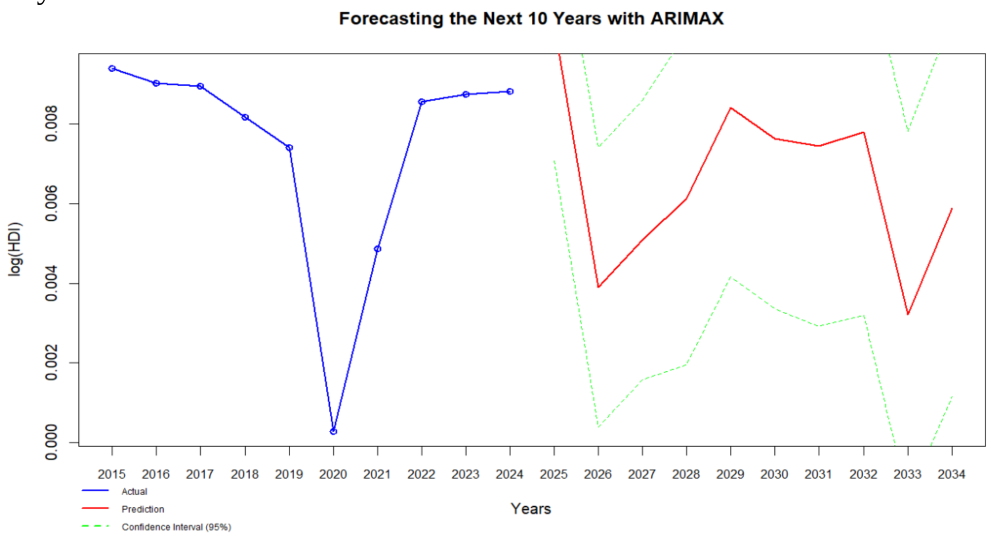

Figure 4. Line graph of the forecast with ARIMAX

The forecast graph in Figure 4 shows that the ARIMAX(3,0,0) model is capable of capturing the general pattern in the historical data. The forecasting results show significant fluctuation patterns in certain years, and the forecasting results also indicate a significant decline around the years 2026 and 2033, which is likely influenced by the patterns observed in the previous year, 2020. The widening confidence interval indicates an increase in uncertainty as the time horizon extends. Nevertheless, this model provides a fairly good picture of medium-term trends and can be used as a tool to support strategic decision-making.

E. CONCLUSION

The Human Development Index (HDI) is an indicator that illustrates the extent to which the population can enjoy the benefits of development, particularly in terms of income, health, education, and other aspects. Three basic dimensions are used as the basis for measuring the Human Development Index (HDI), namely a long and healthy life, a knowledge level, and a decent standard of living (Statistics Indonesia, 2018). Thus, the HDI provides a more comprehensive picture of the overall quality of life of the community, not just from an economic perspective. The ARIMAX(3,0,0) model, with the addition of exogenous variables such as inflation and infant mortality rate, provides the best estimation results. This model has the lowest AIC and BIC values, indicating that it is more parsimonious and has better predictive capability compared to other models. Therefore, the ARIMAX(3,0,0) model with two exogenous variables was chosen as the final model. The forecast graph in Figure 4 shows that the ARIMAX(3,0,0) model is able to capture the general pattern in the historical data.

The forecasting results indicate significant fluctuations in certain years, with a sharp decline predicted around 2026 and 2033. This pattern is suspected to be related to the conditions that occurred in 2020. The widening confidence interval indicates increasing uncertainty along with longer time projections. However, this forecasting model still provides fairly accurate insights into medium-term trends and can be used as an aid in strategic decision-making.

If the government want to simultaneously address the infant mortality rate (IMR), open unemployment rate (OPR), and poverty rate, the government needs to adopt a coordinated and comprehensive approach. These steps include policy integration involving the health, education, economy, and infrastructure sectors, as well as data-driven policies to identify areas most in need of intervention. In addition, improving health and nutrition services, economic development and empowerment, education, and training, as well as cooperation and partnerships between the government, the private sector, and international organizations, are also important. These measures enable the government to address poverty, unemployment, and infant mortality rates effectively and to improve the overall human development index.

REFERENCES

Ahn, E., & Hur, J. (2023). Short-term forecasting of wind power outputs using the enhanced wavelet transform and arimax techniques. *Renewable Energy*, *212*, 394-402. https://doi.org/10.1016/j.renene.2023.05.048

Alfandi, Z., Aryawati, W., & Yanti, D. E. (2022). Hubungan Antara Capaian Indikator Kesehatan Bayi dengan Kematian Bayi. *Jurnal Ilmiah Kesehatan*, *11*(1), 133–143. https://doi.org/10.52657/jik.v11i1.1605

Arini, R. P. ., Mumtazah, S. A. ., Siahaan, R. M. ., & Kartiasih, F. (2024). The Role of Socioeconomic and Female Indicators on Infant Mortality in West Nusa Tenggara: A Panel VECM Analysis. *Journal of Developing Economies*, 9(1), 1–26. https://doi.org/10.20473/jde.v9i1.53416

Azfirmawarman, D., Magriasti, L., & Yulhendri, Y. (2023). Indeks Pembangunan Manusia Di Indonesia. *Jurnal Pendidikan Dan Konseling (JPDK)*, 5(5), 117–125. https://doi.org/10.31004/jpdk.v5i5.22864.

Badan Pusat Statistik (2018). Indeks Pembangunan Manusia Sumatera Utara.

Badan Pusat Statistik. (2023). Perkembangan Beberapa Indikator Sosial Ekonomi Indonesia, Statistik Indonesia.

Badan Pusat Statistik. (2024). Berita Resmi Statistik No. 85/11/Th. XXVII, 15 November 2024.

Badan Pusat Statistik. (2024). "Indeks Pembangunan Manusia 2023". Jakarta, Indonesia. www.bps.go.id.

Badan Pusat Statistik. Tabel Dimensi. https://www.bps.go.id/id/query-builder.

Baihawafi, M., & Sebayang, A. F. (2023). Pengaruh Upah Minimum, Indeks Pembangunan Manusia dan Laju Pertumbuhan Ekonomi terhadap Pengangguran Terbuka. *Jurnal Riset Ilmu Ekonomi Dan Bisnis*, 39-44. https://doi.org/10.29313/jrieb.v3i1.1911

Basuki, A. T., & Prawoto, N. (2017). *Analisis Regresi dalam Penelitian Ekonomi dan Bisnis*. PT RajaGrafindo Persada.

Chalid, N., & Yusuf, Y. (2014). Pengaruh tingkat kemiskinan, tingkat pengangguran, upah minimum kabupaten/kota dan laju pertumbuhan ekonomi terhadap indeks pembangunan manusia di Provinsi Riau. *Jurnal ekonomi*, *22*(2), 1-12.

Chiu, L. Y., Rustia, D. J. A., Lu, C. Y., & Lin, T. T. (2019). Modelling and forecasting of greenhouse whitefly incidence using time-series and ARIMAX analysis. fitrihttps://doi.org/10.1016/j.ifacol.2019.12.521

Feng, Y., Lagakos, D., & Rauch, J. E. (2024). Unemployment and development. *The Economic Journal*, *134*(658), 614-647.

Gonzalez, R. M., & Gilleskie, D. (2017). Infant Mortality Rate as a Measure of a Country's Health: A Robust Method to Improve Reliability and Comparability. Demography, 54(2), 701–720. https://doi.org/10.1007/s13524-017-0553-7.

Jin, H., Qian, X., Chin, T., & Zhang, H. (2020). A global assessment of sustainable development based on modification of the human development index via the entropy method. *Sustainability*, *12*(8), 3251.

Kementrian Kesehatan RI. (2018). Rencana Strategis Kementrian Kesehatan Tahun 2015-2019 Revisi 1 - th. 2017. In Kementerian Kesehatan RI.

Kiha, E. K., Seran, S., & Seuk, G. (2021). Pengaruh Inflasi, Produk Domestik Regional Bruto Dan Upah Minimum Regional Terhadap Indeks Pembangunan Manusia Propinsi Nusa Tenggara Timur. *INVEST : Jurnal Inovasi Bisnis Dan Akuntansi*, *2*(1), 41-56. https://doi.org/10.55583/invest.v2i1.128

Kumalasari, M., & Poerwono, D. (2011). Analisis Pertumbuhan Ekonomi, Angka Harapan Hidup, Angka Melek Huruf, Rata Rata Lama Sekolah, Pengeluaran Perkapita dan Jumlah Penduduk terhadap Tingkat Kemiskinan Di Jawa Tengah (Doctoral dissertation, Universitas Diponegoro).

Majka, M. ARIMAX: Time Series Forecasting with External Variables.

Mankiw, N. G. (2003). Makro Ekonomi. Terjemahan: Fitria Liza. Jakarta: Penerbit Erlangga.

Martias, D. (2021). Statistika Deskriptif Sebagai Kumpulan Informasi. *Jurnal Adab: Fakultas Adab dan Ilmu Budaya, UIN Sunan Kalijaga. https://doi.org/10.14421/fhrs.2021.161.40-59*

Naibaho, M. M., & Nabila, U. (2021). Pengaruh produk domestik regional bruto (PDRB) dan tingkat pengangguran terbuka terhadap indeks pembangunan manusia di Kabupaten Langkat. *Jurnal Gamma-Pi*, *3*(2), 21-26. http://dx.doi.org/10.33059/jgp.v3i2.3684

Ningrum, J. W., Khairunnisa, A. H., & Huda, N. (2020). Pengaruh Kemiskinan, Tingkat Pengangguran, Pertumbuhan Ekonomi dan Pengeluaran Pemerintah Terhadap Indeks Pembangunan Manusia (IPM) di Indonesia Tahun 2014-2018 dalam Perspektif Islam. *Jurnal Ilmiah Ekonomi Islam*, *6*(2), 212. https://doi.org/10.29040/jiei.v6i2.1034

Pangesti, I., & Susanto, R. (2018). Pengaruh inflasi terhadap indeks pembangunan manusia (IPM) di Indonesia. *Journal of Applied Business and Economics*, *5*(1), 70-81. http://dx.doi.org/10.30998/jabe.v5i1.3164

Pangestika, M., & Widodo, E. (2017). Analisis Regresi Panel terhadap Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia di Kabupaten/Kota D.I.Yogyakarta. Seminar Nasional dan The 4th Call for Syariah Paper.

Parolin, Z., & Wimer, C. (2020). Forecasting estimates of poverty during the COVID-19 crisis. *Center on Poverty and Social Policy at Columbia University–Режим доступа: https://static1. squarespace. com/static/5743308460b5e922a25a6dc7*, *5*, 1586988788821.

Pratowo, N. I. (2012). Analisis faktor-faktor yang berpengaruh terhadap Indeks Pembangunan Manusia. *Jurnal Studi Ekonomi Indonesia*, *1*(1), 15-31.

Putra, Windhu. 2019. Perekonomian Indonesia (Penerapan Beberapa Teori Ekonomi Pembangunan di Indonesia). PT. Raja Grafindo Persada: Jakarta.

Putra, D. M., & Ratnasari, V. (2016). Pemodelan Indeks Pembangunan Manusia (IPM) Provinsi Jawa Timur Dengan Menggunakan Metode Regresi Logistik Ridge. *Jurnal Sains dan Seni ITS*, *4*(2). https://doi.org/10.12962/j23373520.v4i2.10450

Putri, A., Agustina, S., Sagita, F., & Kartiasih, F. (2024). Socioeconomic and Health Analysis of Infant Mortality in East Java 2022. *Seminar Nasional Official Statistics*, *2024*(1), 341-352. https://doi.org/10.34123/semnasoffstat.v2024i1.2205

Riestiansyah, F., Damayanti, D., Reswara, M., & Susetyoko, R. (2022). Perbandingan metode ARIMA dan ARIMAX dalam Memprediksi Jumlah Wisatawan Nusantara di Pulau Bali. *Jurnal Infomedia: Teknik Informatika, Multimedia, dan Jaringan*, 7, 123-135. http://dx.doi.org/10.30811/jim.v7i2.3336

Rujiman, Sukardi, L. S. H. (2020). Analisis Determinan Indeks Pembangunan Manusia (IPM) di Indonesia. *Jurnal Penelitian Pendidikan Sosial Humaniora*, *5*(2), 139-141. https://doi.org/10.32696/jp2sh.v5i2.470.

Sarkodie, S. A., & Adams, S. (2020). Electricity access, human development index, governance and income inequality in Sub-Saharan Africa. *Energy Reports*, *6*, 455-466.

Setianingtias, R., Baiquni, M., & Kurniawan, A. (2019). Pemodelan Indikator Tujuan Pembangunan Berkelanjutan Di Indonesia Modeling Indicators of Sustainable Development Goals in Indonesia. *Jurnal Ekonomi Dan Pembangunan*, *27*(2). https://dx.doi.org/10.14203/jep.27.2.2019.61-74

Sugiantari, A. P., & Budiantara, I. N. (2013). Analisis Faktor-faktor yang mempengaruhi angka harapan hidup di Jawa Timur menggunakan Regresi Semiparametrik Spline. *Jurnal sains dan Seni ITS*, *2*(1), D37-D41. http://dx.doi.org/10.12962/j23373520.v2i1.3132

Syofya, H. (2018). Pengaruh tingkat kemiskinan dan pertumbuhan ekonomi terhadap indeks pembangunan manusia Indonesia. *Jurnal Ilmiah Ekonomi Dan Bisnis*, *15*(2), 177-185. https://doi.org/10.31849/jieb.v15i2.1153

Tarumingkeng, W. A., Rumate, V. A., & Rotinsulu, T. O. (2021). Pengaruh belanja modal dan tingkat kemiskinan terhadap indeks pembangunan manusia (IPM) di Provinsi Sulawesi Utara. *Jurnal Pembangunan Ekonomi Dan Keuangan Daerah*, *19*(2), 82-95. https://dx.doi.org/10.35794/jpekd.19789.19.6.2018

Tumpal Manik. (2013). ANALISIS PENGARUH KEMAKMURAN, UKURAN PEMERINTAH DAERAH, INFLASI, INTERGOVERNMENTAL REVENUE DAN KEMISKINAN TERHADAP PEMBANGUNAN MANUSIA DAN PERTUMBUHAN EKONOMI. Jurnal Organisasi Dan Manajemen, 9(2), 107–124. https://doi.org/10.33830/jom.v9i2.41.2013.

United Nation Development Program (UNDP). (1990). Human Development Report 1990. New York: Oxford University Press.

United Nations Development Programme. (2004). *Human Development Index* (HDI). Retrieved January 12, 2025, from https://www.undp.org.

Walpole, R. E. (1995). *Statistical Methods in Research and Production*.

Wati, L., & Solichin, A. (2024). Prediksi Nilai Pengadaan Barang Dan Jasa Pada Sebuah Perusahaan Pariwisata Menggunakan Metode Arima dan Fuzzy Time Series. *Jurnal Inovtek Polbeng Seri Informatika*, *9*(1). https://doi.org/10.35314/isi.v9i1.4041

Widyarta, I. K. P., Samosir, C. P. D., Aysyah, P., & Kartiasih, F. (2024). Revisiting Unemployment in Indonesia: Error Correction Model (ECM) Analysis. Jurnal Akuntansi, Manajemen, Bisnis dan Teknologi, 4(2), 222-241. https://doi.org/10.56870/4q9ctm59

Zuhairoh, Z. A. (2018). Pengaruh Angka Kematian Bayi, Angka Partisipasi Murni, Rasio Ketergantungan terhadap Indeks Pembangunan Manusia Provinsi Jawa Timur. *Jurnal Biometrika Dan Kependudukan*, *7*(1), 87–95. https://doi.org/10.20473/jbk.v7i1.2018.87-95.