Factors Affecting Life Expectancy in East Java: Predictions with A Bayesian Model Averaging Approach

Harun Al Azies and Vivi Mentari Dewi

Department of Statistics, Faculty of Science and Data Analytics,

Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Abstract

This study predicts the factors that influence life expectancy in East Java, Indonesia. In particular, this study compares the prediction results between the linear regression model and the Bayesian Model Averaging (BMA). The study used a 2015 data set from the Central Bureau of Statistics (BPS) of the province of East Java. The results of data exploration show that the life expectancy in East Java is 70.68 years, the Bondowoso regency is the region with the lowest life expectancy at 65.73 years and the city of Surabaya is the area with the highest life expectancy value in East Java, which is 73.85 years. The results of the inference study show that the factors affecting life expectancy in East Java are infant mortality and illiteracy rates among the population aged ten and over. Based on the standard error value, the BMA is a better model for predicting the factors that affect life expectancy in East Java compared to the regression model because the BMA model can estimate parameters more efficiently.

Keywords: Bayesian Model Averaging; Life Expectancy; Regression.

Factors Affecting Life Expectancy in East Java: Predictions with A Bayesian Model Averaging Approach

Harun Al Azies and Vivi Mentari Dewi

I.   Introduction

Life expectancy is a significant summary measure of a population's health and well-being. A country's life expectancy reflects its social and economic conditions and the quality of its public health and health infrastructure, among other factors (Ho and Hendi, 2018). Man's most outstanding achievement in the last century was the tremendous increase in life expectancy. Life expectancy has increased throughout the lifespan, of course up to 100 years. For example, life expectancy at ages 65 and 85 increased by about 50% during this century (Bell and Miller, 2005). Initially, this increase in length of life was due to reducing infectious diseases and deaths concentrated among young people. After the majority of deaths from infectious diseases declined, cardiovascular conditions and cancer dominated the causes of death. It became the target of science and medicine in the second half of the last century. Due largely to the decline in deaths from heart disease, life expectancy has steadily increased in the last decades of the 20th century. Because heart disease primarily causes death among older adults, a recent increase in life expectancy has occurred at older ages. Life expectancy has increased throughout the lifespan, of course up to 100 years. For example, life expectancy at ages 65 and 85 increased by about 50% during this century (Bell and Miller, 2005). The massive improvement in life expectancy had been a signicant trend for high-income developed countries during the 20th and 21st centuries. The actual decline in life expectancy is very worrying. Stagnation or decline in life expectancy can signal a decline in health profiles driven by adverse socioeconomic trends, decreased provision or quality of health care, or deteriorating behavioral factors (Ho and Hendi, 2018).

According to the results of the Indonesian Central Bureau of Statistics (BPS) report, life expectancy in Indonesia in 2015 reached 70.78 years, and life expectancy in Indonesia can be said to have increased from 69.1 years (2005 -2010). However, it is known from the value of life expectancy by the province in Indonesia that 29 provinces have a lower life expectancy than Indonesia's life expectancy. In the island of Java, the province of East Java is an area with a lower life expectancy value than other provinces of the island of Java, where the life expectancy value for East Java is 70.68. Based on the description results, the researcher wants to predict what factors affect life expectancy in East Java Province. Research on factors that affect life expectancy in East Java was conducted by Budiantara et al. (2014) using the semi-parametric spline regression method. Nisa' (2020) also examines the factors that influence life expectancy in East Java with the semiparametric regression mixed spline truncated approach. Both adopt a semi-parametric approach in their research. Meanwhile, Aidi (2014) also studied life expectancy in districts/cities of East Java in 2014. However, they involved spatial effects, notably with Geographical Weighted Regression, while Nazhifah (2017) examined the life expectancy in East Java using regression.

The model results obtained in the above research are a unique model that does not involve model uncertainty. This causes inaccurate estimates to appear in the model selection. The case of life expectancy is a case involving model uncertainty, so we need a method that involves model uncertainty in selecting the best model. The Bayesian Averaging Model

(BMA) provides a coherent mechanism for calculating the uncertainty of this model and provides improved out-of-sample predictive performance. Several methods for implementing BMA have recently emerged. We discuss this method and provide instructions for several applications (Madigan and Raftery, 1994). Kuswanto et al. (2013) conducted a study on the application of BMA regarding the calibration of temperature forecasts in Indonesia in 2013. In addition, Hustianida and Kuswanto (2012) regarding infant mortality cases in Indonesia. The research shows that BMA provides higher accuracy of prediction results than linear regression. Therefore, in this study, the researchers used the Bayesian Model Averaging (BMA) method to predict the factors that affect life expectancy in East Java that will be compared to the best linear regression model. Research on the application of BMA has never been done to model the case of life expectancy in Indonesia.

II. Methods/Methodology

This section describes the types and sources of data used in this study. This section also describes the methods used in this research and the analysis process.

2.1. Data Sources and Research Variables

The type of data in this study is secondary data from the Central Bureau of Statistics (BPS) of the province of East Java. The object of observation used in this study was 38 municipalities and cities of East Java. The research variables used in this study are the variable of life expectancy by municipality/city in 2015 as dependent variable (Y) and five independent variables (X), which include infant mortality rate (X_1), illiteracy rates of the population aged ten years and over (X_2), percentage of households that have access to adequate sources of drinking water (X_3), percentage of the population that has health problems (X_4) and percentage of the last lifeguard child at the time of birth (X_5).

a. Life expectancy (Y) is the estimated average number of years a person can live in their lifetime. This figure can be obtained from the number of deaths in a certain age group divided by the number of people at that age in a year.

b. The infant mortality rate (X_1) is the total number of deaths of infants under one year of age (aged 0-1 years) recorded during one year per 1,000 live births in the same year.

c. The illiteracy rate of the population aged ten and over is the proportion of the population aged ten and over (X_2) who cannot read and write Latin letters and other letters to the population aged ten and over.

d. Households with access to drinking water sources (X_3) are the number of households with access to drinking water sources at a given time divided by the total number of households during the same period expressed as a percentage (%).

e. Percentage of the population with health problems (X_4) describes health problems that are disturbances in physical and mental conditions, primarily due to accidents or other things that interfere with daily activities.

f. Percentage of the last lifeguard child at the time of birth (X_5) is the number who attended the birth of mothers who gave birth.

2.2. Data Analysis Methods

The analytical technique used is a descriptive analysis used to obtain a general description of the state of life expectancy in East Java using map visualization. In addition, an inferential analysis approach was also performed in this study, namely multiple linear regression analysis and the Bayesian Model Averaging (BMA), which will be compared with these two methods to obtain the best prediction model of life expectancy in East Java. The steps of data analysis techniques in research are as follows:

1. Describe the life expectancy in East Java and the factors that influence it.

2. Obtain a model from multiple linear regression analysis by regressing the factors that affect the life expectancy variables. Linear regression modeling aims to determine the relationship between life expectancy and the factors that influence it (Sarstedt and Mooi, 2014). The regression equation model used is as follows:

$$Y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon \qquad (1)$$

description:

Y = Life expectancy.

α = constant.

β = regression coefficient ($\beta_1,\beta_2,\beta_3,\beta_4,\beta_5$).

x_1 = Infant mortality rate.

x_2 = illiteracy rate of the population aged 10 years and over.

x_3 = Percentage of households that have access to proper drinking water sources.

x_4 = Percentage of population who have health complaints.

x_5 = Percentage of the last rescuer under five during the birth process.

ε = error.

3. Testing the regression assumptions, the residual assumptions that must be met include the normality test using the Jarque–Bera Test (Thadewald and Büning, 2007), the heteroscedasticity assumption using the Breusch-Pagan test (Klein et al., 2016), and the last assumption is autocorrelation test using Durbin Watson Test (King and Harris, 1995). Next, detect multicollinearity using Variance Inflation Factors (VIF) (Daoud, 2017).

4. Obtain the model from the analysis using the Bayesian Model Averaging (BMA). There have been many discussions about BMA, including (Madigan and Raftery, 1994); (Raftery, Madigan, and Hoeting, 1997); (Hoeting et al., 1999); (Viallefont, Raftery, and Richardson, 2001); (Liang, Troung, and Wong, 2001); and (Montgomery and Nyhan, 2010). The basic concept of modeling with BMA is to select the model so that the best model is obtained by considering all possible models, with the working system averaging the posterior distribution of all possible models formed. BMA's principle combines uncertain models to obtain a good model (Montgomery and Nyhan, 2010). The Bayesian Model Mean Equation (BMA) model used is as follows:

$$E(\Delta|\mathbf{Y}) = \sum_{s=1}^{q} E(\Delta|\mathbf{Y}, M_s) P(M_s|\mathbf{Y}) \qquad (2)$$

description:

Δ   = value to be predicted.

M   = the set of models that can be formed by **M** with$(M_1, M_2, \ldots, M_q)$.

$E(\Delta|Y)$ = expected value of coefficient Δ conditional on **Y** data.

$E(\Delta|\boldsymbol{Y}, M_s)$ = the expected value of the conditional Δ coefficient of **Y** data in the *s*-th **M** model.

$P(M_s|\boldsymbol{Y})$ = weighting value in the *s*-th **M** model with *s*= 1,2, … ,*q*.

s   = number of best selected models.

After obtaining all the possible BMA models trained, then selecting the best model by selecting the model that will be included in equation (2), the method used in this study is the Occam's Window method (OW) which is based on its posterior probability (Madigan and Raftery, 1994). Models that can be included in equation (2) BMA modeling must meet equation (3) (Jeffreys, 1961) below.

$$\Lambda' = \left\{ M_s : \frac{\text{maksimum}_l\left(P(M_l|\boldsymbol{Y})\right)}{P(M_s|\boldsymbol{Y})} \leq c \right\} \qquad (3)$$

description:

Λ' = posterior odds of the s-th model.

$P(M_l|\boldsymbol{Y})$ = the *lth* model in **M**, which has the highest posterior probability value.

The value of *c* has a limit of 20, which is the limit value of the *s*-th model selection that can be included or not in the BMA modeling in equation (2). The value of *c* is equivalent to alpha (α) 0.05 for p-value (sig) (Jeffreys, 1961). If a model has a Λ'-value greater than $c = 20$, then the model is not selected in the BMA calculation and must be selected or cannot fit into equation (2) and vice versa (Madigan and Raftery, 1994) (Hustianda, 2012).

III. Results and Discussions

This section describes and discusses the results of the analysis. Authors can use the table or picture to present the results of the analysis. The results of the analysis and discussion are presented in a coherent with good systematics so easy to follow.

3.1. Descriptive Statistics

Life expectancy (e0) describes the average age that a person can take during his life, which is calculated from birth. The following is an overview of the life expectancy (e0) in East Java which is presented in the form of a regional map as shown in Figure 1 below.
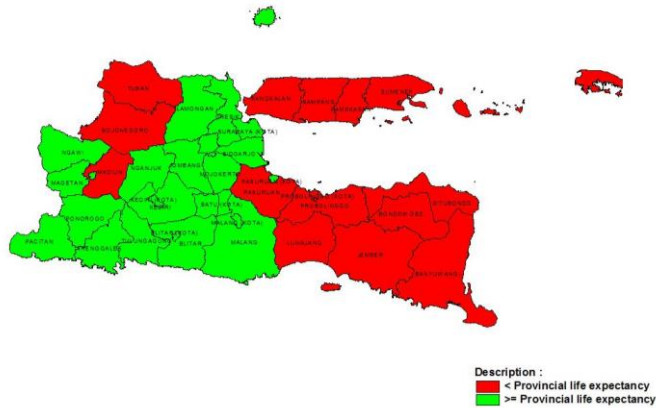
Figure 1. Overview of life expectancy in municipalities/cities of East Java

Source*: Central Bureau of Statistics of East Java, with processed*

Figure 1 based on data released by BPS East Java in 2015, East Java's life expectancy is 70, 68 years old. Based on the visualization in Figure 1, there are 15 municipalities/cities in East Java province which have lower life expectancy than provincial life expectancy rate; these areas include the municipality of Bondowoso, which is also an area in East Java with the lowest life expectancy of 65.73 years.In addition, other areas which are included as areas with lower life expectancy value than provincial life expectancy are Probolinggo, Pamekasan, Sampang, Jember, Situbondo, Lumajang, Bangkalan, Probolinggo City, Pasuruan, Banyuwangi, Madiun, Sumenep, Bojonegoro, and Tuban. The remaining 23 regions of East Java have a higher life expectancy than the provincial life expectancy. The city of Surabaya is one of them and is also an area in East Java with the highest life expectancy value of 73.85 years, which signifies the average age of the population. Surabaya is 73.85 years old. In addition, Table 1 presents descriptive statistics on the factors that influence life expectancy in East Java.

Table 1. Descriptive Statistics Factors Affecting Life Expectancy

Variable	Mean	Variance	Minimum	Maximum
Infant Mortality Rate (X_1)	30.93	146.03	17.27	60.51
Illiteracy Rate of Population Age 10 years and over (X_2)	7.507	27.794	1.140	21.970
% Household Access to Adequate Drinking Water Sources (X_3)	94.47	38.08	70.63	100.00
% Population Has Health Complaints (X_4)	34.746	28.429	24.740	45.590
% of births under five with medical assistance (X_5)	96.40	47.59	68.50	100.01

Table 1 shows that the average infant mortality rate in East Java is 30.93, with quite a diversity of infant mortality rate with a value of 146.03, meaning that the infant mortality rate value varies between municipalities/cities. The lowest infant mortality rate is in Blitar City, which is 17.27, while the highest infant mortality rate is in Probolinggo Regency, which is 60.51. The average illiteracy rate of the population aged ten and over in East Java Province in 2015 was 7,507. This figure shows the value of the welfare level indicator for the province of East Java which is relatively low. The variance value indicates a high value, which is 27.794, this figure explains the imbalance in the level of reading ability of the population of East Java Province across municipalities/cities. The lowest illiteracy rate is in

Sidoarjo Regency, which is 1.12. In comparison the highest is in Sampang Regency, which is 21.97.

The average percentage of households with access to adequate drinking water sources in East Java Province in 2015 was 94.47 percent, and diversity was relatively low at 38.08 percent. Pacitan Regency is the area with the lowest percentage of households with access to drinking water sources in East Java, which is 70.63 percent, unlike Batu city, where 100 percent of households in the region have access to reputable drinking water sources. In East Java, an average of 96.40 percent of deliveries under five by medical personnel, Sampang Regency is the area with the lowest average number of children under five who are assisted by medical personnel during childbirth in East Java. At the same time, there are 18 districts/cities in Java. Timur has 100 percent of toddlers who are assisted by medical personnel during the birth process.

3.2. Multiple Linear Regression Model

The regression model obtained is as follows.

$$\hat{y} = 76.511 - 0.118\, x_1 - 0.126\, x_2 + 0.036\, x_3 - 0.055\, x_4 - 0.025\, x_5 \quad (4)$$

Linear regression modeling aims to determine the relationship between life expectancy and the factors that influence it. To discover the relationship simultaneously and identify the significance of the pattern using the F test. The results of the F test indicate that the value of sig (0.000) $<\alpha$ (0.05), which means that the model is significant or that the predictor variable simultaneously affects the response variable. Table 2 shows the results of the concurrent test.

Table 2. The Simultaneous Test Results of Multiple Linear Regression Modeling

Source of Variance	Sum Square (SS)	F-test	P-value (sig)
Regression	125.819	23.51	0.000*
Error	34.254		
Total	160.073		

Note:*) Significant α (0.05)

After being tested simultaneously. Then individually tested using the t-test. the aim is to find out the predictor variables explaining the variance of the response variables. The individual test results show that there are two significant predictor variables with a p-value smaller than 0.05. namely the Infant Mortality Rate (X_1) and the population illiteracy rate aged ten years and over (X_2). The R^2 value of 78.6% indicates that the resulting model is good because the predictor variable can explain the life expectancy variable accordingly.

Table 3. The Individual Test Results of Multiple Linear Regression Modeling

Variable	Coefficient	Standard Error (SE)	P-value (sig)
(Intercept)	76.51	530.765	1.51e-1
Infant Mortality Rate (X_1)	-0.117	0.0235	1.93e-0*
Illiteracy Rate of Population Age 10 years and over (X_2)	-0.12579	0.06041	0.0454*
% Household Access to Adequate Drinking Water Sources (X_3)	0.03555	0.03065	0.2547
% Population Has Health Complaints (X_4)	-0.05503	0.03446	0.1202
% of births under five with medical assistance (X_5)	-0.02505	0.03957	0.5312
R^2	0.786		

Residual assumptions that must be satisfied include tests for normality, heteroskedasticity, and autocorrelation. Whereas the assumption on the variable is that there is no multicollinearity.

3.2.1.   Assumption of Normality/Normality Test

Testing assumptions to find out whether the residuals meet the normal distribution or not. The test results show that the resulting p-value (sig) is more significant than (0.05). Moreover, and the Jarque-Bera test value of 2.7567 is less than the D-table value of 0.242. The decision is that the residual data meet the normal distribution. Determination of meeting the normal distribution assumption can also be done visually through QQ-Plot as follows.
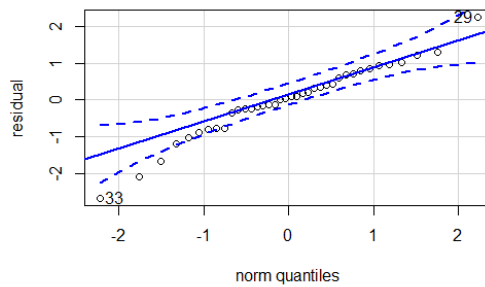

Figure 2. Graph of QQ-Plot Assumption of Normality

Source*: Central Bureau of Statistics of East Java. with processed

Based on Figure 2. it is known that the residual data has a normal distribution. This is because the plot is close to the straight blue line. The conclusion is that the residual data on the factors that affect life expectancy in East Java meet the normal distribution.

3.2.2.   Assumption of Heteroscedasticity

The heteroskedasticity test checks the variance of the error variance. the heteroskedasticity test uses the statistics of the Breusch-Pagan test. The test statistics show that the p (sig) value produced is 0.1128 greater than α (0.05). Moreover, And the Breusch-Pagan value is 8.9074, it means that the residual data does not indicate the existence of heteroskedasticity or is identical.

3.2.3.   Assumption of Autocorrelation

Autocorrelation test to test whether there is a relationship between errors using the Durbin Watson test. The test statistic shows that the Durbin-Watson statistic or "d" value is 1.9716 greater than, the lower limit value "dl" (1.2042) or the value "4-d" (2.0284) is greater than the lower limit "dl" (1.2042). Moreover, and the value of d (1.9716) is greater than the upper limit of "du" (1.7916). So, it means that the residual data meets the independent assumption or there is no autocorrelation.

3.2.4.   Multicollinearity detection

To determine the presence or absence of multicollinearity using an indication of the value of the variance inflation factor (VIF). Table 4 shows the VIF value of each predictor

variable. The test results show that the VIF value for all the predictor variables is less than ten, so that there is no multicollinearity in this case.

Table 4. Multicollinearity Detection Results

Variable	VIF
Infant Mortality Rate (X_1)	2.786937
Illiteracy Rate of Population Age 10 years and over (X_2)	3.506210
% Household Access to Adequate Drinking Water Sources (X_3)	1.236469
% Population Has Health Complaints (X_4)	1.167125
% of births under five with medical assistance (X_5)	2.575864
Infant Mortality Rate (X_1)	2.786937

3.3. Bayesian Model Averaging (BMA)

The posterior probabilities of each predictor variable in Table 5 with parameter coefficient values and the standard error of the predictor variables.

Table 5. Posterior Probability of Predictor Variables. Parameter Coefficients and Standard Error (SE)

Variable	Posterior Probability	Coefficient	Standard Error (SE)
(Intercept)	100	74.765	3.298
Infant Mortality Rate (X_1)	100*	–0.115	0.027
Illiteracy Rate of Population Age 10 years and over (X_2)	76.5*	–0.097	0.071
% Household Access to Adequate Drinking Water Sources (X_3)	28.8	0.012	0.026
% Population Has Health Complaints (X_4)	34.0	–0.017	0.031
% of births under five with medical assistance (X_5)	12.6	–0.0008	0.016

Note:*) Influential variable

Table 5 shows that two predictor variables have a very high posterior probability percentage namely the infant mortality rate (X_1) and the illiteracy rate of the population aged ten years and over (X_2). The other three predictor variables have a posterior probability of less than 50% they are indicating that these variables do not affect the model. Therefore, the predictor variables that affect the life expectancy variable are the infant mortality rate (X1) and the illiteracy rate of the population aged ten and over (X2). BMA to predict factors that affect life expectancy by averaging the posterior probabilities of each predictor variable in each of the selected models. The criteria for determining the best selection model use Occam's window selection method by implementing equation (3) to obtain the best BMA model.

Table 6. BMA Model Selection Results using the Occam's Window Method.

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
(Intercept)	75.097	77.096	71.903	69.923	75.456
Infant Mortality Rate (X_1)	–0.102	–0.110	–0.106	–0.139	–0.145
Illiteracy Rate of Population Age 10 years and over (X_2)	–0.130	–0.131	–0.108	–	–
% Household Access to Adequate Drinking Water Sources (X_3)	–	–	0.033	0.056	–

% Population Has Health Complaints (X₄)	-	–0.051	-	-	-
% of births under five with medical assistance (X₅)	-	-	-	-	-
R²	0.760	0.775	0.768	0.741	0.714
posterior model probability (PMP).	0.323	0.178	0.099	0.075	0.072

Note:*) Influential variable

Posterior probability value above by summing the posterior model probability (PMP) from all models for each predictor. The BMA results show that based on the results of the Occam's window selection, there are five best models out of 32 trained models. Model 1 has the largest PMP, which is 0.323. Which means it affects 32.3% of the total posterior probability. Meanwhile, model 2 gives an effect of 17.8% of the total posterior probability. The cumulative posterior probability of the top five models above is 0.747, which means that it has an effect of 74.7% of the total posterior probability. They are considered from the influence of each predictor variable. the variable X_1 (infant mortality rate) contributes to the five best models so that it has a significant influence on the response variable in the model. At the same time, the variable X_2 (Illiteracy rate of the population aged ten and over) contributes to the three models. Therefore, both predictor variables have a significant posterior probability. Based on the posterior probability above. the Bayesian mean model for predicting factors that affect life expectancy in East Java is as follows.

$$E(\Delta|Y) = 0.323M_1 + 0.178\,M_2 + 0.099M_3 + 0.075M_4 + 0.072M_5 \qquad (5)$$

3.4. Comparison of the Bayesian Model Averaging (BMA) and Multiple Linear Regression

The comparison of the parameter estimates, and standard errors generated to predict the factors that affect life expectancy between the Bayesian model averaging method and the regression is in Table 6. The parameters of the predictor variables estimated between the BMA and the regression are the same namely that two predictor variables influence the model including the Infant mortality rate (X_1) and the population illiteracy rate aged ten and over (X_2). However, when viewed from the standard error of each parameter generated. BMA has a more minor standard error than the regression on all predictor variables. This figure shows that BMA provides parameter estimates that are more effective in predicting factors that affect life expectancy.

Table 7. Comparison of Parameter Significance Value and Standard Error with BMA and Regression

Variable	BMA		Regression	
	probabilities (%)	SE	P-value	SE
(Intercept)	100	3.298	1.5e-1	530.765
Infant Mortality Rate (X₁)	100*	0.027	1.9e-0*	0.0235
Illiteracy Rate of Population Age 10 years and over (X₂)	76.5*	0.071	0.0454*	0.06041
% Household Access to Adequate Drinking Water Sources (X₃)	28.8	0.026	0.2547	0.03065
% Population Has Health Complaints (X₄)	34.0	0.031	0.1202	0.03446
% of births under five with medical assistance (X₅)	12.6	0.016	0.5312	0.03957

Note:*) Influential variable

It was viewed from the coefficients of the parameters. the BMA. and the regression have coefficient values that are not too far apart. Based on the values in Table 8 below.

Table 8. Comparison of parameter coefficients with BMA and regression

Variable	BMA Coefficient	Regression Coefficient
(Intercept)	74.765	76.51
Infant Mortality Rate (X_1)	-0.115	-0.117
Illiteracy Rate of Population Age 10 years and over (X_2)	-0.097	-0.12579
% Household Access to Adequate Drinking Water Sources (X_3)	0.012	0.03555
% Population Has Health Complaints (X_4)	-0.017	-0.05503
% of births under five with medical assistance (X_5)	-0.0008	-0.02505

Table 8 shows that all the predictor variables have the same sign of the coefficient for the coefficients of the BMA and regression parameters. The difference in coefficients between the two methods is not too different. This figure shows that there is no significant difference between the coefficients of the BMA and the regression results.

IV. Conclusion

The life expectancy rate in East Java is 70.68 years. There are 15 municipalities/cities in East Java Province with lower life expectancy than the provincial life expectancy rate. Bondowoso Municipality is in the East Java region with the lowest life expectancy of 65.73 years, and the city of Surabaya is the East Java region with the highest life expectancy value of 73.85 years. The factors that affect life expectancy in East Java with BMA and regression are the same namely infant mortality rate (X_1) and an illiteracy rate of the population aged ten and over (X_2). The results of the comparison between BMA and Regression are based on the standard error of each parameter produced. the BMA model has a standard error more minor than the regression for all predictor variables, this figure shows that the most efficient parameter estimate is the model produced by BMA as regression.

References

Aidi. M. N.. Sumertajaya. I. M.. & Yusuf. L. M. (2014). Geographical Weighted Regression with Kernel Gaussian Weighted Function in Life Expectancy Rate (Case Study: Life Expectancy Rate of Regencies / Cities in East Java Province). *International Journal of Statistics and Applications.* Vol. 4 No. 3. 2014. pp. 144-152.

Badan Pusat Statistik. (2015a). Angka Harapan Hidup Menurut Provinsi. Jakarta: Badan Pusat Statistik.

Badan Pusat Statistik Provinsi Jawa Timur. (2015b). Angka Harapan Hidup Menurut Kabupaten/Kota. Surabaya: Badan Pusat Statistik Provinsi Jawa Timur.

Budiantara. I. N.. Sugiantari. A.. Ratnasari. V.. Ratna. M.. & Zain. I. (2014). Analysis of Factors that Influence Life Expectancy in East Java (Indonesia) Using Semiparametric Spline Regression Approach. *International Jornal of Basic & Applied Sciences. 14(2).*

Daoud. J. (2017). Multicollinearity and Regression Analysis. *Journal of Physics: Conf. Series* . 949.

Ho. J.. & Hendi. A.S. (2018). Recent trends in life expectancy across high income countries: retrospective observational study. *The BMJ. 362.*

Hoeting. J. A.. Madigan. D.. Raftery. A. E.. & Volinsky. C. T. (1999). Bayesian Model Averaging: A Tutorial (with Discussion). *Statistical Science* . 382-417.

Hustianda. V. F. (2012). *Perbandingan Bayesian Model Averaging dan Regresi Linear Berganda dalam Memprediksi Faktor-Faktor yang Mempengaruhi Jumleh Kematian Bayi di Jawa Timur.* Institut Teknologi Sepuluh Nopember. Statistika. Surabaya: Institut Teknologi Sepuluh Nopember.

Jeffreys. H. (1961). *Theory of Probability. 3rd ed.* Oxford: Press Oxford.

King. M.. & Harris. D. (1995). The application of the durbin-watson test to the dynamic regression model under normal and non-normal errors. *Econometric Reviews* . *4* (14). 487-510.

Klein. A. G.. Gerhard. C.. Büchner. R. D.. Diestel. S.. & Schermelleh-Engel. K. (2016). The detection of heteroscedasticity in regression models for psychological data. *Psychological Test and Assessment Modeling* . *4* (58). 567–592.

Kuswanto. H.. & Sari. M. R. (2013). Bayesian model averaging with Markov chain monte Carlo for calibrating temperature forecast from combination of time series models. *Journal of Mathematics and Statistics. 9*(4). 349.

Liang. F. M.. Troung. Y.. & Wong. W. H. (2001). Automatic Bayesian Model Averaging for Linear Regression and Applications in Bayesian Curve Fitting. *Statistical Science* . *4* (11). 1005-1029.

Madigan. D.. & Raftery. A. E. (1994). Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window. *Journal of the American Statistical Association* . *89* (428). 1535-1546.

Miller. B. d. (2005). Lifespan and Healthspan: Past. Present. and Promise. *National Library of Medicine* .

Montgomery. J.. & Nyhan. B. (2010). Bayesian Model Averaging: Theoretical Developments and Practical Applications. *Political Analysis* . 245-270.

Nazhifah. N. (2017). Modeling Of Life Expectancy At Birth In East Java 2015 With Multiple Linear Regression. Institut Teknologi Sepuluh Nopember. Statistika. Surabaya: Institut Teknologi Sepuluh Nopember.

Nisa'. K.. & Budiantara. I. N. (2020). Modeling East Java Indonesia Life Expectancy Using Semiparametric Regression Mixed Spline Truncated And Fourier Series. *Media Statistika. 13*(2). 149-160.

Othman. A.. & Gloaguen. R. (2013). River Courses Affected by Landslides and Implications for Hazard Assessment: A High Resolution Remote Sensing Case Study in NE Iraq–W Iran. *Remote Sensing* . *5* (3). 1024–1044.

Raftery. A. E.. Madigan. D.. & Hoeting. J. (1997). Bayesian Model Averaging for Linear Regression Models. *Journal of the American StatisticaL Association* . *92* .

Sarstedt. M.. & Mooi. E. (2014). *Regression Analysis. In: A Concise Guide to Market Research. Springer Texts in Business and Economics.* Berlin: Springer.

Thadewald. T.. & Büning. H. (2007). Jarque–Bera Test and its Competitors for Testing Normality – A Power Comparison. *Journal of Applied Statistics . 1* (34). 87-105.

Viallefont. V.. Raftery. A.. & Richardson. S. (2001). Variable Selection and Bayesian Model Averaging in Case- Control Studies. *Statistical in Medicine. .* 3215-3230.