

PENERAPAN MODEL ARIMA DALAM PERAMALAN ANAK USIA 5–14 Th YANG TERINFEKSI HIV DI INDONESIA

Wigid Hariadi, Sulantari

Pendidikan Matematika, FP. MIPA, IKIP PGRI Jember

wigid.hariadi@gmail.com

Abstract:

Human Immunodeficiency Virus (HIV) is dangerous diseases for humans, and until now has not found a cure. Virus HIV is attacks the human immune system so that someone is susceptible to disease. This causes if someone is infected with HIV, then the person can experience an danger condition, it will even effect is death. In recent years, the number of children aged 5 – 14 years old that infected with HIV continues to increase. Therefore the author was moved to write about the application of the ARIMA model in forecasting the number of children aged 5 – 14 years old that infected with HIV in Indonesia by 2023. With the hope that the public or the government can find out the potential dangers of HIV disease, especially in children aged 5 – 14 years old. So that the public and government can jointly eradicate the spread of the HIV virus, especially in children. The result are obtained that the model that is suitable for use in forecasting is the ARIMA(0,1,2) models, with error value obtained is 0.057429. with the forecast value of the number of children aged 5 – 14 years old that infected with HIV in Indonesia from 2019 – 2023 in a row is : 570.82, 647.12, 734.14, 823.85, 944.83.

Keywords: ARIMA, Forecasting, HIV, children aged 5 – 14 years old.

PENDAHULUAN

Model *Autoregresif Integreted Moving Average* (ARIMA) merupakan salah satu model ststistika dalam analisis runtun waktu yang popular digunakan dalam aplikasi peramalan data. Model ini ditulis sebagai model ARIMA (p, d, q). Model ARIMA ini merupakan gabungan antara model *Autoregresif* (AR(p)) dan model *Moving Average* (MA(q)). Prosedur penggunaan model ARIMA (p, d, q) ini tidaklah rumit. Dalam metode ini, akan dihasilkan sebuah model matematis yang merepresentasikan data serta dari model tersebut akan dapat dilakukan perhitungan peramalan data untuk beberapa periode kedepan. Jadi pada model ini, peramalan data yang dilakukan berdasarkan pada data sebelumnya (empiris). Aplikasi model ARIMA (p, d, q) sangat luas, artinya model ini dapat diterapkan bada berbagai bidang (As'ari, Tohir, Valentino, Imron, & Taufiq, 2014). Dalam analisis runtun waktu, seringkali data dibagi menjadi dua bagian yang disebut data in-sample yakni data yang digunakan untuk memilih model terbaik dan data out-sample yakni bagian data yang digunakan untuk memvalidasi keakuratan peramalan dari model terbaik yang diperoleh berdasarkan data in-sample (Eksiandayani, 2016; Rosadi, 2012).

Saat sekarang, terdapat banyak masalah kesehatan yang menghantui umat manusia, dan tidak sedikit dari jenis penyakit tersebut dapat mengancam nyawa. Diantara sekian banyak penyakit yang berbahaya, salah satunya adalah



Content from this work may be used under the terms of the [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/) that allows others to share the work with an acknowledgment of the work's authorship and initial publication in this journal.

HIV/AIDS yang mana sampai sekarang bahkan belum ditemukan obatnya. Menurut Noer (Sdityo Nadis & Prabawati, 2018), menjelaskan bahwa HIV (*Human Immunodeficiency Virus*) merupakan virus yang menyerang daya tahan tubuh manusia sehingga seseorang mudah terserang penyakit. Orang yang terinfeksi HIV, cepat atau lambat (2 sampai 10 tahun) akan menderita AIDS (*Acquired Immuno Deficiency Syndrome*) jika tidak berobat secara teratur. Sementara AIDS merupakan kumpulan gejala penyakit dengan karakteristik defisiensi imun yang berat dan merupakan manifestasi stadium akhir infeksi HIV. Selain itu, menurut Doktersehat (2016), virus HIV/AIDS dapat merusak atau menghancurkan sel kekebalan tubuh, sehingga sel kekebalan tubuh tidak mampu berperang melawan infeksi atau kanker. Sekitar 3,2 juta anak-anak di bawah 15 tahun hidup dengan AIDS pada akhir 2013, menurut data Organisasi Kesehatan Dunia (WHO).

Saat ini, HIV/AIDS tidak hanya mengancam orang dewasa, namun sudah banyak menjangkiti anak-anak dan remaja. Menurut Kompasdata (2019), Kementerian kesehatan mencatat, dari 79,6 juta anak di Indonesia pada 2018, tercatat 2.881 anak adalah Anak Dengan HIV/AIDS (ADHA). Kondisi ADHA cukup rentan, selain karena faktor penyakitnya, beban yang mereka alami juga cukup berat karena mereka secara psikologis belum dewasa. Beban kian berat akibat problem diskriminasi di sejumlah daerah. dengan semakin mengkhawatirkannya perkembangan dan pertumbuhan jumlah penderita HIV/AIDS dari tahun ke tahun, khususnya Anak-anak, maka penulis tergerak untuk menulis tentang penerapan model ARIMA dalam peramalan jumlah anak usia 5 – 14 tahun yang terinfeksi HIV di Indonesia pada tahun 2023. Semoga melalui tulisan ini, setelah diketahui proyeksi pertumbuhan jumlah anak terinfeksi HIV untuk beberapa tahun kedepan, penulis berharap tulisan ini dapat digunakan sebagai dasar pertimbangan dalam upaya pemberantasan dan pencegahan penyakit HIV/AIDS di Indonesia terutama dikalangan anak-anak.

METODE PENELITIAN

2.1 Jenis Penelitian dan Sumber Data

Menurut tujuannya, penelitian dalam tulisan ini merupakan penelitian terapan dan kuantitatif. Karena data yang digunakan adalah data kuantitatif. Menurut Sugiyono (2017) mengatakan bahwa metode kuantitatif dapat diartikan sebagai metode penelitian yang berlandaskan pada filsafat positivisme, digunakan untuk meneliti pada populasi atau sampel tertentu, pengumpulan data menggunakan instrument penelitian, analisis data bersifat kuantitatif atau statistik, dengan tujuan untuk menguji hipotesis yang telah ditetapkan. Sedangkan menurut metodenya, tulisan ini merupakan penelitian historis. Penulis menggunakan data sekunder yang berasal dari data Litbang Kompas yang diunggah dalam @kompasdata. Dimana data tersebut berisi tentang jumlah anak Indonesia yang terinfeksi HIV dari tahun 2010 sampai 2018.

2.2 Analisis Data

2.2.1 Model AR (p)

Proses *autoregressive* (AR), Menurut (Wei, 2006), proses AR(p) didefinisikan sebagai:

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + a_3 X_{t-3} + \dots + a_p X_{t-p} + \varepsilon_t \quad (1)$$

Dengan:

a_i untuk $i = 1, 2, \dots, p$ adalah parameter model AR dengan $\varepsilon_t \sim WN(0, \sigma^2)$.

2.2.2 Model MA (q)

Proses Moving average (MA), menurut (Wei, 2006), mengatakan bahwa jika series yang stasioner merupakan fungsi linear dari kesalahan peramalan sekarang dan masa lalu yang berurutan, persamaan itu dinamakan Moving Average Model (MA). Proses MA didefinisikan sebagai berikut:

$$X_t = \varepsilon_t - b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} - \dots - b_q \varepsilon_{t-q} \quad (2)$$

Dengan:

b_j untuk $j = 1, 2, \dots, q$ adalah parameter model MA dengan $\varepsilon_t \sim WN(0, \sigma^2)$.

2.2.3 Model ARIMA (p,d,q)

Menurut (Soejoeti, 1987), model ARMA secara umum merupakan model campuran antara proses AR dan proses MA, dimana modelnya didefinisikan sebagai berikut:

$$X_t = a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} + \varepsilon_t - b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} - \dots - b_q \varepsilon_{t-q} \quad (3)$$

Atau dapat juga ditulis dengan :

$$a_p (B)X_t = b_q (B)\varepsilon_t \quad (4)$$

$$\text{dengan : } \begin{aligned} a_p (B) &= 1 - a_1 B - \dots - a_p B^p \\ b_q (B) &= 1 - b_1 B^1 - \dots - b_q B^q \end{aligned}$$

Sehingga berdasarkan model ARMA diatas, diperoleh pengembangan model ARIMA (p,d,q) yang secara matematis, modelnya dituliskan sebagai berikut:

$$a_p (B)(1-B)^d X_t = b_0 + b_q (B)\varepsilon_t \quad (5)$$

dengan : d = adalah koefisien *differencing*

2.2.4 Metode Estimasi Maksimum Likelihood (MLE)

Definisi (Bain & Engelhardt, 1992): Fungsi densitas bersama dari n variabel random X_1, X_2, \dots, X_n diestimasi dengan x_1, x_2, \dots, x_n dilambangkan dengan $f(x_1, x_2, \dots, x_n; \theta)$ dengan θ adalah parameter yang tidak diketahui, maka fungsi *likelihood* dari θ adalah :

$$L(\theta) = f(x_1; \theta)f(x_2; \theta) \dots f(x_n; \theta) \quad (6)$$

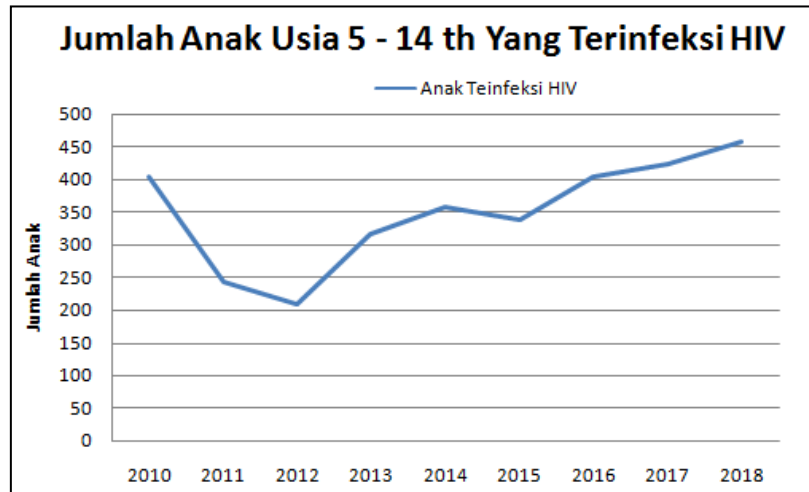
Menurut (Subanar, 2013), metode estimasi maksimum likelihood adalah metode estimasi parameter yang paling populer dalam menghasilkan estimator. Dimana jika X_1, \dots, X_n adalah i.i.d dan sampel berasal dari populasi yang memiliki fungsi densitas $f(x | \theta_1 \dots \theta_k)$. Maka fungsi dari kemungkinan (likelihood) dapat didefinisikan sebagai berikut:

$$L(\underline{\theta} | \underline{x}) = L(\theta_1 \dots \theta_k | x_1 \dots x_n) = \prod_{i=1}^n f(x_i | \theta_1 \dots \theta_k) \quad (7)$$

Jika fungsi likelihood terdeferensialkan (dalam θ_i), maka nilai calon MLE yang mungkin terjadi adalah harga-harga dari $(\theta_1 \dots \theta_k)$.

HASIL DAN PEMBAHASAN

Besarnya jumlah Anak-anak usia 5 – 14 th yang terinfeksi HIV dari tahun ke tahun terus mengalami pergerakan naik. Meskipun sempat mengalami penurunan, namun kini angka tersebut perlahan mengalami kenaikan. Seperti tersaji pada Gambar 1 dibawah ini:



Gambar 1: Grafik Jumlah Anak Indonesia Usia 5 – 14 Yang Terinfeksi HIV tahun 2010 – 2018. (Sumber : @Kompasdata).

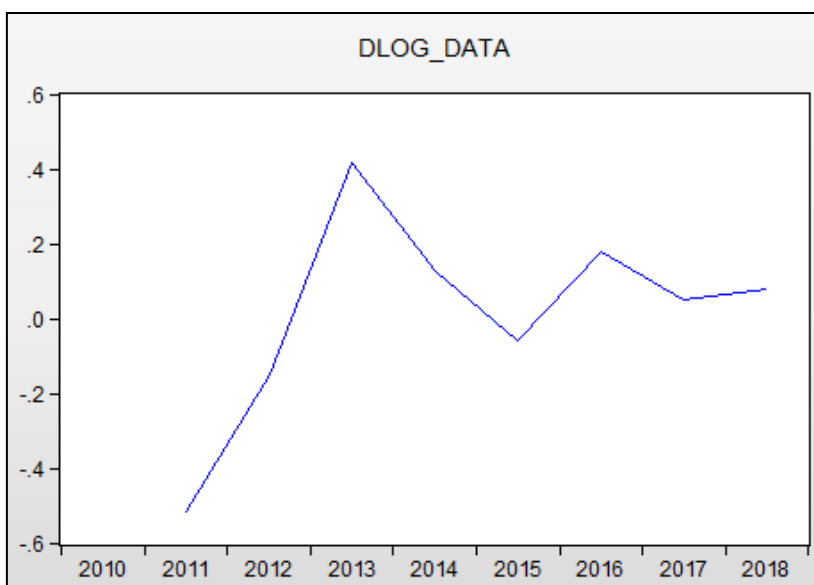
Dari Gambar 1 diatas, dapat dijelaskan bahwa Jumlah Anak usia 5-14 th yang terinfeksi HIV mengalami kenaikan dari ke tahun. karena datanya cenderung mengalami trend naik, maka penulis berpendapat bahwa model yang tepat untuk memprediksi jumlah anak usia 5-14 th yang terinfeksi HIV pada tahun 2023 adalah model ARIMA (p,d,q). dalam tulisan ini penulis menggunakan software Eviews7. Untuk melakukan proses perhitungan menggunakan model ARIMA, pertama-tama perlu dilakukan pengecekan stasioneritas data. Hal ini untuk mengetahui apakah data pada Gambar 1 sudah stasioner atau belum. Adapun output dari uji stasioneritas datanya adalah sebagai berikut:

Null Hypothesis: DATA has a unit root		
Exogenous: Constant		
Lag Length: 0 (Automatic - based on SIC, maxlag=1)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.106967	0.6548
Test critical values:		
1% level	-4.582648	
5% level	-3.320969	
10% level	-2.801384	

Gambar 2 : Output Uji Stasioneritas Data.

Dengan menggunakan tingkat signifikansi $\alpha = 5\%$, maka dari Gambar 2, dapat terlihat bahwa nilai proabilitas ADF test nya sebesar 0.6548, hal ini menunjukkan

bahwa nilai probabilitas ADF test $> \alpha = 5\%$. sehingga dapat di simpulkan bahwa data belum stasioner baik dalam mean ataupun variansi. Maka selanjutnya adalah melakukan transformasi dan differencing. Untuk transformasi data, penulis menggunakan transformasi Logaritma. Setelah dilakukan transformasi data, kemudian dilakukan differencing data. Dimana grafik data hasil transformasi dan differencing dapat terlihat pada Gambar 3 dibawah ini:



Gambar 3 : Grafik Hasil Differencing Setelah Data ditransformasi

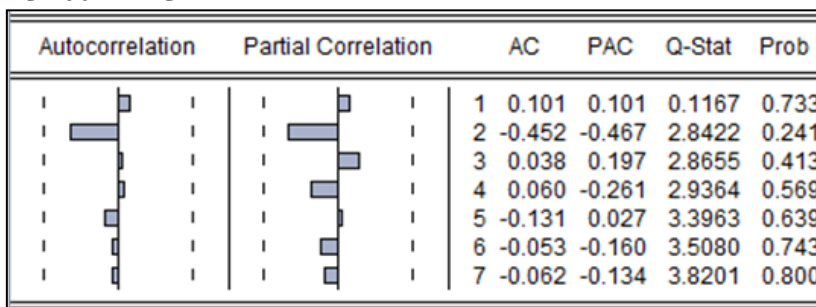
Pada Gambar 3 diatas, terlihat bahwa grafik data hasil differencing satu kali sudah berada di sekitar rata-rata. Hal ini menunjukkan adanya indikasi bahwa datanya sudah stasioner. Maka selanjutnya adalah melakukan uji stationeritas data hasil differencing. Adapun output uji stasioneritasnya adalah sebagai berikut:

Null Hypothesis: DLOG_DATA has a unit root		
Exogenous: Constant		
Lag Length: 1 (Automatic - based on SIC, maxlag=1)		
	t-Statistic	Prob.*
<hr/>		
Augmented Dickey-Fuller test statistic		
	-12.15448	0.0001
Test critical values:		
1% level	-5.119808	
5% level	-3.519595	
10% level	-2.898418	

Gambar 4 : Output Uji Stasioneritas Data Setelah dilakukan Differencing.

Dengan menggunakan tingkat signifikansi $\alpha = 5\%$, dan dengan hipotesis nol (H_0) dalam uji ini adalah data belum stasioner dalam mean maupun variansi, dan daerah kritisnya adalah tolak H_0 jika p-value (prob) $< \alpha$. Sehingga dari Gambar 4 diatas, dapat terlihat bahwa nilai probabilitas ADF test nya sebesar 0.0001, hal ini menunjukkan bahwa nilai p-value ADF test $< \alpha = 5\%$. Sehingga dapat di simpulkan bahwa data sudah stasioner baik dalam mean ataupun variansi. Oleh

karena itu langkah selanjutnya adalah melakukan penentuan model ARIMA (p,d,q) melalui plot ACF dan PACF.



Gambar 5 : Plot ACF dan PACF

Pada Gambar 5, terlihat bahwa pada plot ACF dan PACF tidak terdapat lag yang keluar dari garis batas. Sehingga ada dugaan tidak terdapat model ARIMA(p,d,q). Namun, identifikasi visual bukanlah satu-satunya dasar penentuan model ARIMA. Yang menjadi penentu model ARIMA adalah hasil dari estimasi model. Karena pada Gambar 5 tidak dapat diketahui model dasarnya. Maka penulis langsung melakukan pengujian estimasi pada model overfitting. Pada model ini, model overfitting yang diperoleh yakni ARIMA (0,1,1), dan ARIMA (0,1,2). Kemudian berdasarkan dua model tersebut, dilakukan estimasi model ARIMA, kemudian setelah dilakukan estimasi model, dilanjutkan dengan uji asumsi pada residual. Uji asumsi ini meliputi uji normalitas, uji Noautokorelasi, dan uji Homoskedastik. dimana rangkuman hasil estimasi model tersaji pada Tabel 1 dibawah ini.

Tabel 1 : Rangkuman Estimasi Model ARIMA (p,d,q)

		Model ARIMA	
		ARIMA (0,1,1)	ARIMA (0,1,2)
Estimasi	C	0	0.126157
	p-value		0.0044
	Ma(1)	0.936364	0
	p-value	0.0000	
	Ma(2)	0	-0.954991
	p-value		0.0000
	R-squared	0.519578	0.889742
	S.E of regression	0.189070	0.097834
	Sum Squared Resid	0.250232	0.057429
	Akaike info criterion	-0.376932	-1.598775
Schwarz criterion	-0.367002	-1.732725	
Uji Asumsi	Normalitas	Ya	Ya
	p-value	0.945306	0.181968
	No Autokorelasi	Ya	Ya
	Homoskedastik	Ya	Ya

Menggunakan tingkat signifikansi $\alpha = 5\%$, maka pada Tabel 1 terlihat bahwa model ARIMA (0,1,1) dan ARIMA (0,1,2) sama-sama merupakan model yang signifikan. Artinya kedua model ini merupakan model yang bagus untuk merepresentasikan data, dan model ini dapat digunakan untuk melakukan peramalan data. Karena terdapat dua model yang signifikan, maka perlu dilakukan pemilihan model, tujuannya untuk memilih model yang terbaik. Kriteria model yang terbaik adalah model yang signifikan, memenuhi uji asumsi, serta memiliki nilai error yang terkecil.

Berdasarkan Tabel 1, terlihat bahwa kedua model tersebut sama-sama model yang signifikan, kemudian sama-sama memenuhi uji asumsi normalitas, uji asumsi no autokorelasi (tidak terdapat autokorelasi pada residual model), serta uji homoskedastik. Sehingga pada akhirnya tinggal melihat model mana yang memiliki nilai error yang terkecil. Dalam model ini satuan error yang digunakan adalah SSR (Sum Squared Residual). Pada model ARIMA (0,1,1) nilai SSR nya sebesar 0.250232, dan pada model ARIMA (0,1,2) nilai SSR nya sebesar 0.057429. Dari kedua model tersebut, terlihat bahwa nilai error model ARIMA (0,1,2) merupakan error yang paling kecil. Sehingga berdasarkan signifikansi model, uji asumsi, dan pengecekan nilai error, maka dapat diputuskan bahwa model ARIMA (p,d,q) yang terbaik untuk merepresentasikan data Anak Usia 5 -14 th yang terinfeksi HIV tahun 2010-2018 adalah model ARIMA (0,1,2).

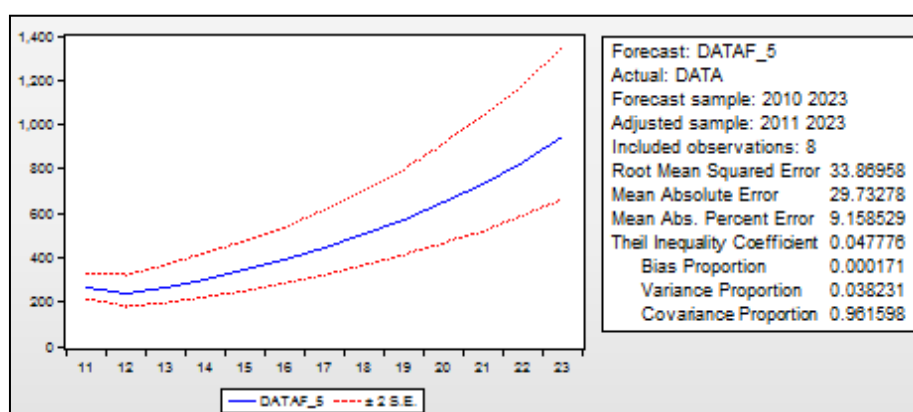
Dikarenakan telah diputuskan bahwa model terbaiknya adalah model ARIMA(0,1,2), maka secara matematis modelnya dapat dituliskan sebagai berikut:

$$(1-B)^1 X_t = b_0 + \varepsilon_t - b_2 \varepsilon_{t-2}$$

$$(1-B)^1 X_t = 0.126157 + \varepsilon_t - (-0.954991)\varepsilon_{t-2}$$

$$(1-B)^1 X_t = 0.126157 + 0.954991)\varepsilon_{t-2} + \varepsilon_t$$

Selanjutnya berdasarkan model diatas, dapat dilakukan peramalan data (forecasting) untuk beberapa periode kedepan. Dalam tulisan ini penulis mencoba untuk melakukan peramalan data 5 periode kedepan. Yakni peramalan data sampai tahun 2023. Dimana grafik peramalan datanya terlihat pada Gambar 6.



Gambar 6 : Grafik Peramalan Data Sampai 5 Periode Kedepan.

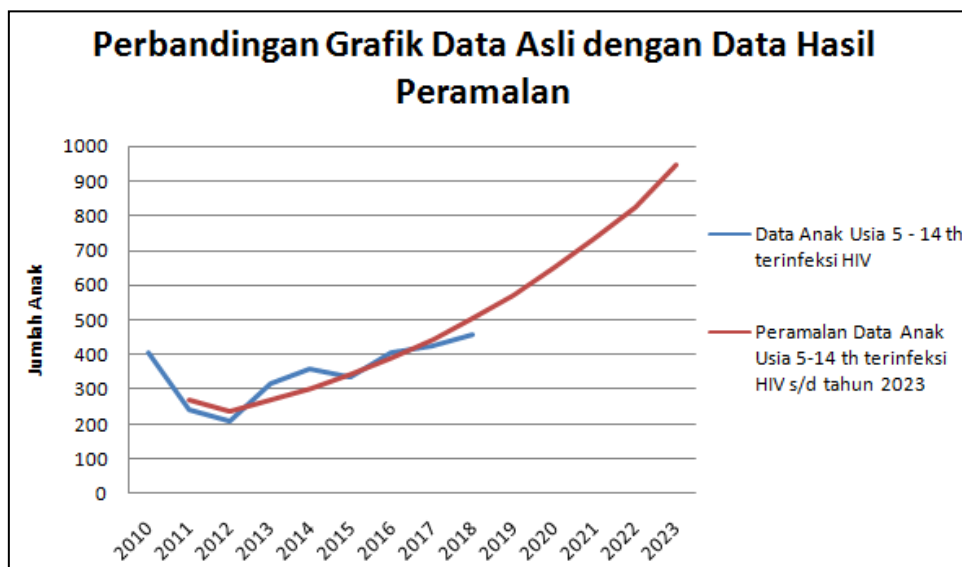
Penerapan Model Arima dalam Peramalan Anak Usia....

Adapun hasil perhitungan peramalan data Anak usia 5-14 tahun yang terinfeksi HIV di Indonesia untuk 5 periode kedepan atau sampai tahun 2023, dapat terlihat pada Tabel 2.

Tabel 2 : Hasil Peramalan Data Anak Usia 5-14 th Yang Terinfeksi HIV di Indonesia Sampai Tahun 2023.

Tahun	Anak Usia 5-14 th Terinfeksi HIV	
	Data Asli	Peramalan Data
2010	405	
2011	242	266.41
2012	208	235.87
2013	316	267.59
2014	358	303.57
2015	338	344.38
2016	405	390.69
2017	425	443.22
2018	459	502.82
2019		570.43
2020		647.12
2021		734.14
2022		823.85
2023		944.83

Berdasarkan hasil perhitungan peramalan data yang telah dilakukan, maka selanjutnya dapat dilakukan perbandingan grafik data antara data asli dengan data hasil ramalan. Hal ini dilakukan agar masyarakat dengan mudah dapat melihat prediksi pertumbuhan jumlah Anak usi 5-14 tahun yang terinfeksi HIV di Indonesia sampai tahun 2023. Dimana grafik tersebut tersaji dalam Gambar 7.



Gambar 7 : Grafik Nilai Perbandingan Data Asli dengan Hasil Peramalan

Berdasarkan Gambar 7 diatas, terlihat bahwa pergerakan jumlah Anak usia 5-14 tahun yang terinfeksi HIV di Indonesia terus mengalami kenaikan dari tahun ketahun. Hal ini patut menjadi perhatian pihak pemerintah, agar dapat lebih

menggiatkan program pencegahan penularan penyakit HIV pada anak-anak. Dikarenakan anak-anak merupakan tunas muda harapan bangsa yang akan menjadi ujung tombak bangsa Indonesia dimasa depan. Tentu kita semua tidak ingin masa depan generasi penerus ini hancur akibat terinfeksi HIV. Tentu saja tujuan akhirnya adalah untuk menjaga masa depan bangsa Indonesia tetap cemerlang.

KESIMPULAN

1. Model ARIMA(p,d,q) yang baik dalam merepresentasikan data anak usia 5-14 tahun yang terinfeksi HIV di Indonesia adalah model ARIMA(0,1,2) dengan nilai error (SSR) sebesar 0.057429.
2. Hasil peramalan data memperlihatkan Anak usia 5-14 tahun yang terinfeksi HIV di Indonesia terus mengalami kenaikan.
3. Hasil peramalan data Anak usia 5-14 tahun yang terinfeksi HIV di Indonesia untuk 5 periode kedepan (dari tahun 2019 – 2023) secara berturut-turut adalah sebagai berikut: 570.82, 647.12, 734.14, 823.85, 944.83.

DAFTAR PUSTAKA

- As'ari, A. R., Tohir, M., Valentino, E., Imron, Z., & Taufiq, I. (2014). *Matematika*. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Bain, L. J., & Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*. Duxbury. *Pacific Grove, CA*.
- Doktersehat. (2016). *HIV/AIDS pada Anak – Penyebab, Gejala, dan Pengobatan*. doktersehat.com: <https://doktersehat.com/hiv-aids-pada-anak/> [Diakses, 10 November 2019].
- Eksiandayani, S. (2016). *Pemodelan Peramalan Inflasi Umum dan Inflasi Menurut Kelompok Pengeluaran di Indonesia dengan Metode Hibrida Arimax-NN*. Institut Teknologi Sepuluh Nopember.
- Kompasdata. (2019). Jumlah Anak Terinfeksi HIV. In *Instagram Kompasdata*. Jakarta: Litbang Kompas.
- Rosadi, D. (2012). *Ekonometrika & Analisis Runtun Waktu Terapan dengan EViews*. Yogyakarta: Graha Ilmu.
- Sdityo Nadis, B., & Prabawati, I. (2018). Implementasi Program Online SMS, Delivery Access, High Access LINK (ODHALINK) di Rumah Sakit Umum Daerah Bangil Kabupaten Pasuruan. *Publika*, 6(9).
- Soejoeti, Z. (1987). *Analisis Runtun Waktu*. Jakarta: Universitas Terbuka.
- Subanar. (2013). *Statistika Matematika*. Yogyakarta: Graha Ilmu.
- Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif dan R&D*. Bandung: Alfabeta.
- Wei, W. W. S. (2006). *Time Series Analysis: Univariate and Multivariate Methods: Second Edition*. California: Addison-Wesley Publishing Company.