

MODEL FISHER'S Z AUTOREGRESSIVE

STUDI KASUS PADA PERAMALAN PRODUK DOMESTIK REGIONAL BRUTO ATAS DASAR HARGA KONSTAN PROVINSI KALIMANTAN TIMUR

(Fisher's z Autoregressive Model: A Case Study on Forecasting Gross Regional Domestic Products at Constant Prices in East Kalimantan Province)

Arifatus Solikhah

Badan Pusat Statistik Provinsi Kalimantan Timur

E-mail: arifatus.solikhah@bps.go.id

Submitted : 23 Mei 2022 | Accepted : 29 Juni 2022

ABSTRAK

Dalam penelitian ini, diperkenalkan model Fisher's z Autoregressive (ZAR) beserta penerapannya dalam peramalan dengan menggunakan metode Bayesian algoritma *Hamiltonian Monte Carlo* (HMC) sebagai metode estimasinya. Model ZAR merupakan model Autoregressive (AR) yang *error term*-nya berdistribusi Fisher's z. Bentuk distribusi Fisher's z yang bisa miring maupun simetris, menjadikan distribusi ini lebih fleksibel dalam menangkap pola *fat-tailed* dalam data, dibandingkan distribusi Gaussian. Model ZAR digunakan untuk meramalkan Produk Domestik Regional Bruto Atas Dasar Harga Konstan 2010 (PDRB ADHK) Provinsi Kalimantan Timur (Kaltim) Triwulan I tahun 2022. Sebagai pembandingan, dalam penelitian ini juga dilakukan peramalan dengan menggunakan model Gaussian Autoregressive (GAR). Hasil penelitian menunjukkan bahwa model ZAR tanpa intersep mengungguli model GAR dengan intersep maupun tanpa intersep, seperti yang ditunjukkan oleh kriteria minimum dari *Pareto-Smoothed Important Sampling Leave-One-Out cross-validation* (PSIS-LOO).

Kata kunci: model Fisher's z autoregressive, fat-tailed, peramalan PDRB ADHK, metode Bayesian, Stan

ABSTRACT

The Fisher's z Autoregressive (ZAR) model is introduced. Its application in forecasting using the Bayesian method of the Monte Carlo Hamiltonian (HMC) algorithm as the estimation method. The ZAR model is an Autoregressive (AR) model whose error term has a Fisher's z distribution. The form of the Fisher's z distribution, which can be skewed or symmetrical, makes this distribution more flexible in capturing fat-tailed patterns in the data, compared to the Gaussian distribution. The ZAR model is used to forecast the Gross Regional Domestic Product (GRDP) at constant prices 2010 in East Kalimantan Province in the first quarter of 2022. As a comparison, in this study also uses the Gaussian Autoregressive (GAR) model to forecast. The results showed that the ZAR model without intercepts outperformed the GAR models with and without intercepts, as indicated by the minimum criteria of the Pareto-Smoothed Important Sampling Leave-One-Out cross-validation (PSIS-LOO).

Keywords: Fisher's z autoregressive model, fat-tailed, GRDP at constant prices forecasting, Bayesian method, Stan

1. PENDAHULUAN

Analisis *time series* yang linier biasanya mengasumsikan bahwa *error term* model berdistribusi Gaussian. Implikasi dari asumsi tersebut, baik distribusi marginal maupun bersyarat dari *time series* juga mengikuti distribusi Gaussian. Namun, karakteristik distribusi Gaussian yang tidak mampu menangkap pola data yang berekor tebal (*fat-tailed*) dan tidak simetri (*asymmetry*) mengakibatkan masih ditemukannya nilai *squared residual* yang besar ketika asumsi Gaussian digunakan sebagai *error term* dalam pemodelan.

Beberapa peneliti telah banyak yang mengembangkan distribusi yang tidak simetri dan berekor tebal. Fisher (1924) memperkenalkan distribusi z sebagai distribusi dari suatu fungsi $Y = \frac{1}{2} \ln \mathbb{F}$ dengan dua parameter bentuk d_1 dan d_2 , yang variabel acak \mathbb{F} mengikuti distribusi F . Selanjutnya,

distribusi tersebut disebut dengan distribusi Fisher’s z standar yang *unimodal*, *leptokurtic*, memiliki bentuk simetri ketika $d_1 = d_2$ dan tidak simetri ketika $d_1 \neq d_2$, serta memiliki modus pada angka 0 (Aroian, 1941). Namun, distribusi tersebut belum adaptif jika digunakan dalam pemodelan karena modulusnya masih tetap berada pada angka 0. Dalam rangka memperbaiki kekurangan dari distribusi Fisher’s z standar, Solikhah dkk. (2021a; 2021b) mengembangkan distribusi Fisher’s z dengan cara menambahkan parameter lokasi μ dan parameter skala σ pada distribusi Fisher’s z standar.

Distribusi Fisher’s z telah digunakan sebagai *error term* pada beberapa model, yaitu model Fisher’s z *Error Regression* (ZER) (Solikhah dkk., 2021a), model Fisher’s z *Mixture Autoregressive* (ZMAR) (Solikhah dkk., 2021b), serta model Fisher’s z *Markov Switching Autoregressive* (ZMSAR) (Solikhah dkk., 2022). Berdasarkan kajian tersebut, terbukti bahwa distribusi Fisher’s z lebih baik daripada distribusi Gaussian dalam menangkap pola data yang tidak simetris dan berekor tebal.

Pada penelitian ini, kami menggunakan distribusi Fisher’s z sebagai *error term* pada model *autoregressive* untuk menangkap pola data yang tidak simetris dan berekor tebal. Model ini disebut sebagai model Fisher’s z *Autoregressive* (ZAR). Estimasi parameter model ZAR dilakukan dengan menggunakan metode Bayesian algoritma *Hamiltonian Monte Carlo* (HMC). Selanjutnya, model ZAR digunakan untuk meramalkan Produk Domestik Regional Bruto Atas Dasar Harga Konstan 2010 (PDRB ADHK) Provinsi Kalimantan Timur Triwulan I tahun 2022. Sebagai pembandingan, dalam penelitian ini juga dilakukan peramalan dengan menggunakan model Gaussian *Autoregressive* (GAR). Adapun pemilihan model terbaik dilakukan dengan menggunakan *Leave-One-Out cross-validation* (LOO) yang estimasinya menggunakan *Pareto Smoothed Importance Sampling* (PSIS) (Vehtari dkk., 2017).

2. MODEL FISHER’S Z AUTOREGRESSIVE

Misal y_t merupakan nilai dari suatu pengamatan pada waktu t ; $t = 1, 2, \dots, T$ dengan T menyatakan banyaknya data pengamatan. Misal, ϵ_t merupakan variabel acak yang identik dan independen berdistribusi Fisher’s z standar, $z(d_1; d_2; 0; 1)$, sedemikian rupa sehingga ϵ_t independen terhadap $\{y_{t-i}, i > 0\}$ Model Fisher’s z *Autoregressive* dengan orde p dinyatakan dengan persamaan:

$$y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + \sigma \epsilon_t, \tag{1}$$

yang mana ϕ_0 merupakan intersep, ϕ_i merupakan koefisien AR *lag* ke- i ; $i = 1, 2, \dots, p$, dan σ merupakan parameter skala. Selanjutnya, model ini disingkat ZAR(p) dengan parameter $\theta = (d_1, d_2, \sigma, \phi_0, \phi)'$; $\phi = (\phi_1, \phi_2, \dots, \phi_p)'$; yang mana ϕ diasumsikan memenuhi kondisi berikut:

$$1 - \sum_{i=1}^p \phi_i a^i \neq 0 \text{ untuk } |a| \leq 1. \tag{2}$$

Misal, θ merupakan vektor parameter yang tidak diketahui dari suatu model dan y merupakan vektor data. Dengan menggunakan pendekatan Bayesian, kesimpulan tentang parameter θ dinyatakan dalam bentuk probabilitas yang bergantung pada nilai y yang diamati. Pernyataan probabilitas tersebut biasanya dinotasikan dengan $f(\theta|y)$. Misal, informasi awal tentang parameter tersebut dinyatakan dengan distribusi *joint prior*, yang dilambangkan dengan $f(\theta)$. Fungsi *likelihood* suatu model berdasarkan pada data yang ada dinyatakan dengan $\mathcal{L}(y|\theta)$. Maka $f(\theta|y)$, atau biasanya disebut dengan distribusi *joint posterior* dari parameter θ , secara proporsionalnya dapat dinyatakan sebagai berikut:

$$f(\theta|y) \propto \mathcal{L}(y|\theta)f(\theta). \tag{3}$$

Adapun fungsi *likelihood* untuk model ZAR dinyatakan sebagai berikut:

$$\mathcal{L}(y|\theta) = \prod_{t=p+1}^T \frac{2(d_2/d_1)^{\frac{1}{2}d_2} \exp(-d_2(a_t/\sigma))}{\sigma B\left(\frac{1}{2}d_1, \frac{1}{2}d_2\right) [1 + \exp(-2(a_t/\sigma) + \ln[d_2/d_1])]^{(d_1+d_2)/2}}, \tag{4}$$

yang mana $B(\cdot)$ merupakan fungsi beta, $\exp(\cdot)$ merupakan fungsi inversi dari logaritma natural, $\ln(\cdot)$ merupakan fungsi logaritma natural, d_1, d_2 merupakan dua parameter bentuk, σ merupakan parameter skala, dan $a_t = y_t - (\phi_0 + \sum_{i=1}^p \phi_i y_{t-i})$. Selanjutnya, jika diasumsikan bahwa parameter dalam model ZAR saling bebas maka distribusi *joint prior* dari model ZAR memenuhi kondisi berikut:

$$f(\theta) = f(d_1)f(d_2)f(\sigma)f(\phi_0) \prod_{i=1}^p f(\phi_i), \tag{5}$$

yang mana lambang $f(d_1)$, $f(d_2)$, $f(\sigma)$, $f(\phi_0)$ dan $f(\phi_i)$ masing-masing merupakan distribusi *prior* dari parameter d_1 , d_2 , σ , ϕ_0 dan ϕ_i pada model ZAR, dan indeks i menyatakan *lag* yang ke- i ; $i = 1, 2, \dots, p$. Adapun distribusi *prior* dari masing-masing parameter tersebut ditetapkan sebagai berikut.

1. Parameter $d_1 \sim t_{v_1}(m_1, g_1^2)I(0, \infty)$ dan memiliki f.k.p sebagai berikut:

$$f(d_1) = \frac{1}{1 - F_{d_1}(0)} \frac{\Gamma\left(\frac{v_1 + 1}{2}\right)}{g_1 \sqrt{\pi v_1} \Gamma\left(\frac{v_1}{2}\right)} \left(1 + \frac{(d_1 - m_1)^2}{v_1 g_1^2}\right)^{-\frac{(v_1 + 1)}{2}}; \tag{6}$$

$$d_1 > 0; g_1 > 0; v_1 > 0; -\infty \leq m_1 \leq \infty,$$

yang mana $\Gamma(\cdot)$ merupakan fungsi gamma, $F_{d_1}(0)$ menyatakan fungsi distribusi kumulatif dari variabel acak d_1 pada batas bawah titik 0, dan π merupakan konstanta *pi* yang dibulatkan bernilai 3,14.

2. Parameter $d_2 \sim t_{v_2}(m_2, g_2^2)I(0, \infty)$ dan memiliki f.k.p sebagai berikut:

$$f(d_2) = \frac{1}{1 - F_{d_2}(0)} \frac{\Gamma\left(\frac{v_2 + 1}{2}\right)}{g_2 \sqrt{\pi v_2} \Gamma\left(\frac{v_2}{2}\right)} \left(1 + \frac{(d_2 - m_2)^2}{v_2 g_2^2}\right)^{-\frac{(v_2 + 1)}{2}}; \tag{7}$$

$$d_2 > 0; g_2 > 0; v_2 > 0; -\infty \leq m_2 \leq \infty,$$

yang mana $F_{d_2}(0)$ menyatakan fungsi distribusi kumulatif dari variabel acak d_2 pada batas bawah titik 0.

3. Parameter $\sigma \sim t_{v_3}(m_3, g_3^2)I(0, \infty)$ dan memiliki f.k.p sebagai berikut:

$$f(\sigma) = \frac{1}{1 - F_{\sigma}(0)} \frac{\Gamma\left(\frac{v_3 + 1}{2}\right)}{g_3 \sqrt{\pi v_3} \Gamma\left(\frac{v_3}{2}\right)} \left(1 + \frac{(\sigma - m_3)^2}{v_3 g_3^2}\right)^{-\frac{(v_3 + 1)}{2}}; \tag{8}$$

$$\sigma > 0; g_3 > 0; v_3 > 0; -\infty \leq m_3 \leq \infty,$$

yang mana $F_{\sigma}(0)$ menyatakan fungsi distribusi kumulatif dari variabel acak σ pada batas bawah titik 0.

4. Parameter $\phi_0 \sim N(u_0, g_0^2)$ dan memiliki f.k.p sebagai berikut:

$$f(\phi_0) = \frac{1}{g_0 \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\phi_0 - u_0}{g_0}\right)^2\right); \tag{9}$$

$$-\infty \leq \phi_0 \leq \infty; -\infty \leq u_0 \leq \infty; g_0 > 0.$$

5. Parameter $\phi_i \sim N(u_i, g_i^2)$; $i = 1, 2, \dots, p$ dan memiliki f.k.p sebagai berikut:

$$f(\phi_i) = \frac{1}{g_i \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\phi_i - u_i}{g_i}\right)^2\right); \tag{10}$$

$$-\infty \leq \phi_i \leq \infty; -\infty \leq u_i \leq \infty; g_i > 0.$$

Memperhatikan distribusi *joint posterior* secara umum untuk parameter θ yang dinyatakan pada Persamaan (3), fungsi *likelihood* model ZAR yang dinyatakan pada Persamaan (4), dengan parameternya diasumsikan saling bebas sehingga distribusi *joint prior*-nya dapat dinyatakan seperti Persamaan (5) dan *prior* masing-masing parameternya dinyatakan pada Persamaan (6) s.d. Persamaan (10), maka fungsi ln dari distribusi *joint posterior* model ZAR dalam bentuk proporsionalnya dinyatakan sebagai

$$\ell(\theta|\mathbf{y}) = \ln f(\theta|\mathbf{y}) \propto \left(\sum_{t=p+1}^T \Lambda_1 + \Lambda_2 \right) + \Lambda_3 + \Lambda_4 + \Lambda_5 + \Lambda_6 + \sum_{i=1}^p \Lambda_7 \tag{11}$$

$$\Lambda_1 = d_2 \ln \left[\frac{d_2}{d_1} \right] + \ln \left[\Gamma \left(\frac{d_1 + d_2}{2} \right) \right] - \ln \left[\Gamma \left(\frac{d_1}{2} \right) \right] - \ln \left[\Gamma \left(\frac{d_2}{2} \right) \right] - \ln \sigma ;$$

$$\Lambda_2 = -d_2 \left(\frac{a_t}{\sigma} \right) - \frac{d_1 + d_2}{2} \ln \left[1 + \exp \left[-2 \frac{a_t}{\sigma} + \ln \left[\frac{d_2}{d_1} \right] \right] \right] ;$$

$$\Lambda_3 = -\frac{1}{2} (v_1 + 1) \ln \left[1 + \frac{1}{v_1} \left(\frac{d_1 - m_1}{g_1} \right)^2 \right] ;$$

$$\Lambda_4 = -\frac{1}{2} (v_2 + 1) \ln \left[1 + \frac{1}{v_2} \left(\frac{d_2 - m_2}{g_2} \right)^2 \right] ;$$

$$\Lambda_5 = -\frac{1}{2} (v_3 + 1) \ln \left[1 + \frac{1}{v_3} \left(\frac{\sigma - m_3}{g_3} \right)^2 \right] ;$$

$$\Lambda_6 = -\frac{1}{2} \left(\frac{\phi_0 - u_0}{g_0} \right)^2 ; \quad \Lambda_8 = -\frac{1}{2} \left(\frac{\phi_i - u_i}{g_i} \right)^2 ,$$

yang mana $a_t = y_t - (\phi_0 + \sum_{i=1}^p \phi_i y_{t-i})$.

Distribusi *posterior* berisi semua informasi terkini tentang parameter θ . Idealnya, seluruh distribusi *posterior* $f(\theta|y)$ ditampilkan secara grafis. Namun, untuk tujuan praktisnya, perlu diberikan berbagai ringkasan numerik dari distribusi tersebut (Gelman dkk., 2014). Ringkasan numerik tersebut biasanya berupa ringkasan lokasi atau titik, seperti *mean*, median, dan modus. Selain itu, juga bisa diberikan ringkasan interval yang biasanya disebut dengan *Credible Interval* (CI).

CI dari distribusi *posterior* $f(\theta|y)$ tidaklah unik, terdapat beberapa metode untuk menentukannya. Beberapa metode tersebut di antaranya (Edwards dkk., 1963; Gelman dkk., 2014) sebagai berikut.

1. Memilih interval yang memiliki probabilitas sama besar antara wilayah yang berada di bawah *posterior median* dan wilayah yang berada di atasnya. Interval yang terbentuk menggunakan metode ini dikenal dengan istilah *Central Posterior Interval* (CPI). Dalam praktiknya, CPI lebih banyak digunakan karena CPI bisa diinterpretasikan sebagai kuantil $\alpha/2$ dan $1 - \alpha/2$ dari distribusi *posterior* dan sudah langsung dihitung menggunakan *posterior simulation*.
2. Memilih interval terpendek, yang mana *posterior mode* termasuk di dalamnya. Interval yang dibentuk menggunakan metode tersebut dikenal dengan istilah *Highest Posterior Density Interval* (HPDI). Cakupan HPDI identik dengan CPI jika distribusi *posterior* berbentuk simetris dan *unimodal*.
3. Memilih interval secara geometris, yang mana *posterior mean* sebagai titik pusatnya.

Inferensi statistik dengan simulasi sering diterapkan pada analisis Bayesian karena sampel dari distribusi *posterior* $f(\theta|y)$ relatif mudah untuk dibangkitkan. Misal, untuk memperkirakan persentil ke-95 dari distribusi *posterior* $f(\theta|y)$, dapat dilakukan dengan membangkitkan sampel acak berukuran \mathcal{R} dari $f(\theta|y)$ dan menggunakan statistik orde ke- $0,95 \times \mathcal{R}$ sebagai perkiraan persentilnya (Gelman dkk., 2014).

Markov Chain Monte Carlo (MCMC) merupakan metode yang populer untuk membuat simulasi dari distribusi *posterior*. Strategi pengambilan sampel MCMC menghasilkan Markov *Chain* yang *positive recurrent*, *irreducible*, dan *aperiodic*, yang berarti nilai parameter yang dibangkitkan konvergen pada distribusi stasioner. Jika kondisi ini terpenuhi, distribusi parameter yang dihasilkan MCMC sama dengan distribusi *posterior* yang diinginkan (Robert, 1996). Algoritma Gibbs *sampler* dan Metropolis merupakan algoritma MCMC yang banyak dipakai. Perilaku langkah acak (*random walk*) dalam algoritma Gibbs *sampler* dan Metropolis menjadi salah satu penyebab simulasi yang dilakukan sering memakan waktu lama (Gelman dkk., 2014). Hamiltonian Monte Carlo (HMC) merupakan modifikasi dari algoritma Metropolis sehingga langkah acak dalam algoritma tersebut dapat ditekan.

Adapun algoritma HMC untuk membuat simulasi distribusi *posterior* model ZAR adalah sebagai berikut:

1. Menentukan nilai awal parameter θ^0 , matriks massa diagonal \mathbf{M} , faktor skala ε , jumlah *leapfrog step* per iterasi \mathcal{N} , dan banyaknya iterasi \mathcal{R} .
2. Untuk setiap iterasi ke- r ; $r = 1, 2, \dots, \mathcal{R}$ dengan \mathcal{R} menyatakan banyaknya iterasi, dilakukan tahapan berikut.
 - a) Membangkitkan vektor momentum φ dengan $\varphi \sim \text{Normal}(\mathbf{0}, \mathbf{M})$;
 - b) Untuk setiap iterasi $n = 1, 2, \dots, \mathcal{N}$, dilakukan tahapan berikut.
 - i. Membuat setengah langkah untuk *update* momentum φ_n ;

$$\varphi_{n+0,5} \leftarrow \varphi_n + \frac{1}{2} \varepsilon \frac{\partial}{\partial \theta_n} \ln f(\theta_n | \mathbf{y}),$$
 yang mana $\frac{\partial}{\partial \theta} \ln f(\theta | \mathbf{y})$ merupakan turunan pertama dari fungsi \ln distribusi *joint posterior* model ZAR seperti yang terdapat pada Persamaan (11), terhadap masing-masing parameternya.
 - ii. Memperbarui vektor θ_n dengan menggunakan vektor $\varphi_{n+0,5}$;

$$\theta_{n+1} \leftarrow \theta_n + \varepsilon \mathbf{M}^{-1} \varphi_{n+0,5}.$$
 - iii. Memperbarui setengah langkah berikutnya untuk φ ;

$$\varphi_{n+1} \leftarrow \varphi_{n+0,5} + \frac{1}{2} \varepsilon \frac{\partial}{\partial \theta_{n+1}} \ln f(\theta_{n+1} | \mathbf{y}).$$
 - c) Memberi label $\theta^{(r-1)}$ dan $\varphi^{(r-1)}$ sebagai nilai dari vektor parameter dan momentum pada awal proses *leapfrog* dan $\theta^\#$, $\varphi^\#$ sebagai nilai keduanya setelah \mathcal{N} langkah.
 - d) Menghitung

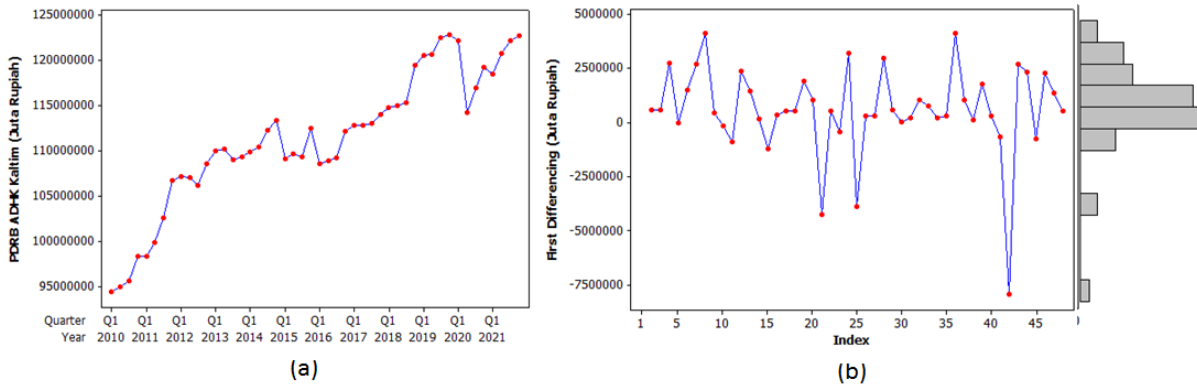
$$b = \frac{f(\theta^\# | \mathbf{y}) f(\varphi^\#)}{f(\theta^{(r-1)} | \mathbf{y}) f(\varphi^{(r-1)})}.$$
 - e) Menentukan $\theta^{(r)}$;

$$\theta^{(r)} = \begin{cases} \theta^\# & \text{dengan probabilita } \min(b, 1), \\ \theta^{(r-1)} & \text{lainnya.} \end{cases}$$
 - f) Menyimpan $\theta^{(r)}$; $r = 1, 2, \dots, \mathcal{R}$.

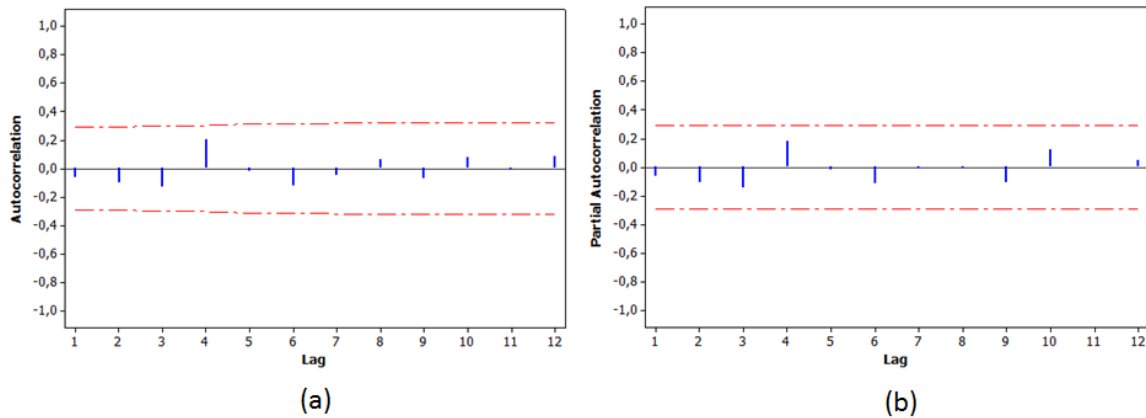
3. HASIL DAN PEMBAHASAN

Sebagai terapan model ZAR, kami menggunakan data PDRB ADHK Kaltim Triwulan I-2010 s.d. Triwulan IV-2021 (BPS Prov. Kalimantan Timur, 2022a; BPS Prov. Kalimantan Timur, 2022b). PDRB ADHK Kaltim oleh BPS Provinsi Kalimantan Timur setiap triwulan. Data triwulanan PDRB ADHK Kaltim mengandung pola tren, seperti yang terlihat pada Gambar 1.a. Oleh karena itu, data tersebut perlu dilakukan *differencing* terlebih dahulu. Berdasarkan plot *first differencing* yang terdapat pada Gambar 1.b, data tersebut memiliki distribusi marginal yang miring dan berekor tebal. Jika menggunakan model GAR, maka data tersebut lebih sesuai jika dimodelkan dengan model GAR yang memiliki intersep. Tetapi jika menggunakan model ZAR, data tersebut bisa dimodelkan dengan menggunakan model ZAR yang memiliki intersep maupun tanpa intersep. Oleh karena itu, pada penelitian ini perlu dibandingkan kebaikan model ZAR dan GAR yang menggunakan intersep maupun tanpa intersep, dalam meramalkan angka triwulanan PDRB ADHK Kaltim.

Seperti peneliti lain, di antaranya Hamilton (1989), pada penelitian ini kami juga melakukan transformasi pada angka PDRB ADHK Kaltim, dengan transformasi sebagai berikut. Jika variabel X_t menyatakan angka PDRB ADHK Kaltim pada triwulan t , maka $Y_t = 100(\ln X_t - \ln X_{t-1})$. Setelah dilakukan transformasi, langkah selanjutnya yang perlu dilakukan adalah melakukan identifikasi orde AR yang nantinya digunakan dalam model, dengan cara memperhatikan plot *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF). Gambar 2 mengilustrasikan tidak ada *lag* dari plot ACF dan PACF variabel Y yang signifikan pada tingkat signifikansi 5%. Oleh karena itu, peramalan variabel Y pada penelitian ini menggunakan model AR tanpa orde *autoregressive* atau ordenya 0. Empat model AR yang dibandingkan pada penelitian ini meliputi model ZAR dengan intersep dan tanpa intersep, serta model GAR dengan intersep dan tanpa intersep.



Gambar 1. Plot dan distribusi marginal PDRB ADHK triwulan I-2010 s.d. triwulan IV-2021 Kaltim; (a) data asli, (b) data *first differencing*.



Gambar 2. Plot dan batas signifikansi 5 persen untuk ACF dan PACF variabel Y_t ; (a) ACF, (b) PACF.

Tabel 1. Perbandingan nilai PSIS-LOO model ZAR dengan intersep, ZAR tanpa intersep, GAR dengan intersep, dan GAR tanpa intersep pada variabel Y .

Model	PSIS-LOO
ZAR dengan intersep	195,60
ZAR tanpa intersep	188,50
GAR dengan intersep	189,00
GAR tanpa intersep	190,80

Fitting keempat model tersebut menggunakan *software* R dengan *package* *rstan*, *warm-up* sebesar 1500 iterasi dari sejumlah 5000 iterasi *sampling* yang dilakukan. Nilai *thin* ditetapkan sebesar 1, *chains* sebesar 3, *max_treedepths* ditetapkan sebesar 15, dan *adapt_delta* sebesar 0,99. Kode Stan untuk *fitting* dan meramal angka PDRB ADHK Kaltim triwulan I-2022 dengan model ZAR terdapat pada Lampiran A.

Nilai PSIS-LOO masing-masing model terdapat pada Tabel 1, yang mana model ZAR tanpa intersep memiliki nilai PSIS-LOO yang terkecil, yaitu sebesar 188,50. Sedangkan model lainnya memiliki nilai PSIS-LOO yang lebih besar, yaitu sebesar 195,60 untuk model ZAR dengan intersep, sebesar 189,00 untuk model GAR dengan intersep, dan sebesar 190,80 untuk model GAR tanpa intersep. Oleh karena itu, bisa disimpulkan bahwa model ZAR tanpa intersep adalah yang terbaik.

Tabel 2. Perbandingan nilai *posterior mean*, *posterior standard error of mean* (se mean), *posterior standard deviation* (sd), dan *Central Posterior Interval* (CPI) dari parameter model ZAR dengan intersep, ZAR tanpa intersep, GAR dengan intersep, dan GAR tanpa intersep.

Parameter model	mean	se_mean	sd	2,5%	25%	50%	75%	97,5%	n_{eff}	\hat{R}
ZAR dengan intersep										
σ	3,43	0,00	0,13	3,18	3,36	3,43	3,50	3,71	6484	1
d_1	4,03	0,00	0,15	3,69	3,96	4,04	4,11	4,32	4868	1
d_2	10,55	0,00	0,20	10,22	10,47	10,55	10,63	10,90	2139	1
ϕ_0	0,90	0,00	0,09	0,72	0,84	0,90	0,96	1,08	6446	1
ZAR tanpa intersep										
σ	4,92	0,00	0,14	4,63	4,85	4,92	5,00	5,21	4865	1
d_1	7,93	0,00	0,16	7,62	7,85	7,93	8,01	8,24	4160	1
d_2	8,98	0,00	0,16	8,68	8,90	8,98	9,05	9,28	5059	1
GAR dengan intersep										
σ	1,85	0,00	0,10	1,66	1,79	1,85	1,91	2,06	4653	1
ϕ_0	0,56	0,00	0,26	0,04	0,39	0,57	0,73	1,07	4579	1
GAR tanpa intersep										
σ	1,92	0,00	0,10	1,71	1,86	1,92	1,99	2,13	2193	1

Simulasi pada data ini telah mencapai konvergen yang ditunjukkan dengan nilai $\hat{R} < 1,01$ dan $n_{eff} > 400$ untuk semua parameter model, seperti yang terdapat pada Tabel 2. Nilai estimasi titik dan interval untuk masing-masing parameter model tersebut juga terdapat pada Tabel 2, berupa *posterior mean* dan CPI. Pada model ZAR tanpa intersep, angka *posterior mean* parameter σ , d_1 , d_2 masing-masing sebesar 4,92; 7,93; 8,98 dan 95% CPI-nya berada pada interval 4,63 s.d. 5,21 untuk parameter σ , interval 7,62 s.d. 8,24 untuk parameter d_1 , dan interval 8,68 s.d. 9,28 untuk parameter d_2 . Adapun model ZAR untuk variabel Y tersebut adalah sebagai berikut:

$$y_t = z(7,93; 8,98; 0; 4,92).$$

Model untuk variabel Y tersebut selanjutnya ditransformasi ke model data asli X , model ZAR untuk data triwulanan PDRB ADHK Kaltim, yaitu:

$$x_t = \exp\left(\frac{z(7,93; 8,98; 0; 4,92)}{100} + \ln x_{t-1}\right), \tag{12}$$

Setelah mengestimasi parameter semua model, langkah selanjutnya adalah melakukan peramalan. Angka *posterior mean* dan HPDI ramalan PDRB ADHK Kaltim pada triwulan I-2022 dari masing-masing model terdapat pada Tabel 3.

Tabel 3. Perbandingan nilai *posterior mean*, *posterior standard error of mean* (se_mean), *posterior standard deviation* (sd), Batas Bawah (BB) HPDI, dan Batas Atas (BA) HPDI dari model ZAR dengan intersep, ZAR tanpa intersep, GAR dengan intersep, dan GAR tanpa intersep pada peramalan angka PDRB ADHK Kaltim, triwulan I-2022 (juta rupiah).

Model	mean	se_mean	sd	BB 90% HPDI	BA 90% HPDI
ZAR dengan intersep	123.498.730,60	21.006,95	1.968.062,34	120.219.330,29	126.558.918,61
ZAR tanpa intersep	122.733.773,78	23.612,40	2.232.582,77	118.936.587,72	126.248.050,76
GAR dengan intersep	123.503.219,47	25.081,33	2.334.791,66	119.604.224,59	127.292.829,52
GAR tanpa intersep	122.779.171,69	24.924,04	2.375.268,27	119.032.322,70	126.846.269,85

Jika menggunakan model ZAR tanpa intersep, diperkirakan nilai PDRB ADHK Kaltim pada triwulan I-2022 berada dalam kisaran 118.936.587,72 juta rupiah sampai dengan 126.248.050,76 juta rupiah, yang mana estimasi titiknya diramalkan berada pada angka 122.733.773,78 juta rupiah.

4. KESIMPULAN

Kami telah memperkenalkan model ZAR, yaitu model *autoregressive* yang *error term*-nya berdistribusi Fisher's *z*. Metode Bayesian merupakan metode yang kami gunakan untuk mengestimasi parameter model ZAR, dengan menggunakan algoritma HMC. Selanjutnya, model ZAR diaplikasikan untuk meramal PDRB ADHK Kaltim triwulan I-2022, dan terbukti model ZAR tanpa intersep lebih baik daripada model GAR dengan intersep maupun tanpa intersep. Ramalan interval PDRB ADHK Kaltim Triwulan I-2022 diperkirakan berada dalam kisaran 118,94 triliun rupiah sampai dengan 126,25 triliun rupiah. Jika dibandingkan dengan hasil rilis BPS Kalimantan Timur, PDRB ADHK Kaltim pada Triwulan I-2022 sebesar 120,74 triliun rupiah berada dalam ramalan interval model ZAR.

DAFTAR PUSTAKA

- Aroian, L.A. (1941), "A Study of R. A. Fisher's *z* Distribution and the Related F Distribution", *The Annals of Mathematical Statistics*, Vol. 12, No. 4, pp. 429-448, diakses dari: <http://www.jstor.org/stable/2235955>.
- BPS Prov. Kalimantan Timur (2022a), *BPS Prov. Kalimantan Timur*. [Online] diakses dari: <https://kaltim.bps.go.id/indicator/52/88/1/-seri-2010-pdrb-atas-dasar-harga-konstan-menurut-lapangan-usaha.html> [pada 17 June 2022].
- BPS Prov. Kalimantan Timur (2022b), Berita Resmi Statistik Pertumbuhan Ekonomi Kalimantan Timur Triwulan I-2022, *Triwulanan*, 9 May.
- Edwards, W., Lindman, H. and Savage, L.J. (1963), "Bayesian Statistical Inference for Psychological Research", *Psychological Review*, Vol. 70, No. 3, pp. 193-242.
- Fisher, R.A. (1924), "On a Distribution Yielding the Error Functions of Several Well Known Statistics", *Proceedings of the International Congress of Mathematics*, Toronto, pp. 2: 805-813.
- Franses, P.H. and Legerstee, R. (2010), "A unifying view on multi-step forecasting using an autoregression", *Journal of Economic Surveys*, Vol. 24, No. 3, pp. 389-401.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (2014), *Bayesian Data Analysis*, 3rd ed., Chapman and Hall/CRC Press, Boca Raton, Fla.
- Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. and Rubin, D.B. (2014), *Bayesian Data Analysis*, 3rd ed., Chapman and Hall/CRC Press, Boca Raton, Fla.
- Gelman, A. and Tuerlinckx, F. (2000), "Type S error rates for classical and Bayesian single and multiple comparison procedures", *Computational Statistics*, Vol. 15, pp. 373-390.
- Hamilton, J.D. (1989), "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle", *Econometrica*, Vol. 57, No. 2, pp. 357-384.
- Robert, G.O. (1996), "Markov Chain Concept Related to Sampling Algorithm", dalam *Markov Chain Monte Carlo in Practice*, eds. W.R. Gilks and S. Richardson, CRC Press, Boca Raton, Florida, pp. 45-57.
- Solikhah, A., Kuswanto, H., Iriawan, N., Fithriasar, K. and Choir, A.S. (2021a), "Extending Runjags: A tutorial on adding Fisher's *z* distribution to Runjags", *AIP Conference Proceedings*, AIP Publishing LLC, 2021, p. 060005.
- Solikhah, A., Kuswanto, H., Iriawan, N. and Fithriari, K. (2021b), "Fisher's *z* Distribution-Based Mixture Autoregressive Model", *Econometrics*, Vol. 9, No. 27, pp. 1-35.
- Solikhah, A., Kuswanto, H., Iriawan, N. and Fithriari, K. (2022), *Pengembangan Model Mixture Autoregressive (Studi Kasus: Peramalan Harga Sumatran Light Crude)*, Disertasi, Institut Teknologi Sepuluh Nopember, Surabaya.

- Tsay, R.S. (2010), *Analysis of Financial Time Series*, Third ed., John Wiley & Sons, Inc., Hoboken, New Jersey.
- Vehtari, A., Gelman, A. and Gabry, J. (2017), "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC", *Statistics and Computing*, Vol. 27, No. 5, pp. 1413–1432.
- Wong, C.S. and Li, W.K. (2000), "On a mixture autoregressive model.", *Journal of the Royal Statistical Society B*, Vol. 62, pp. 95–115.

LAMPIRAN A. KODE STAN UNTUK *FITTING* DAN MERAMAL ANGKA PDRB ADHK KALTIM TRIWULAN I-2022, MODEL ZAR DENGAN INTERSEP

```
library(rstan)
auto_write = TRUE
fitAR1 = "
functions{
  real fisher_z_lpdf(real x, real d1, real d2, real mu, real sigma){
    return (log(2)+0.5*d2*(log(d2)-log(d1))-d2*(x-mu)/sigma
      -log(sigma)-lbeta(0.5*d1,0.5*d2)-(d1+d2)/2*log1p_exp((-2*(x-mu)/sigma)+log(d2)
      -log(d1)));
  }
  real fisher_z_rng(real d1, real d2, real mu, real sigma){
    return(mu+sigma*0.5*log((chi_square_rng(d1)*d2)/(chi_square_rng(d2)*d1)));
  }
}
data {
  int T;
  vector[T] y;
  int U;
  vector[U] x;
  int U_p;
}
parameters {
  real<lower = 0> sigma;
  real<lower = 0> d1;
  real<lower = 0> d2;
  real phi0;
}
transformed parameters {
}
model {
  real lf[T];
  // priors
  sigma ~ student_t(3,3.42,0.1);
  d1 ~ student_t(3,4.05,0.1);
  d2 ~ student_t(3,10.55,0.1);
  phi0 ~ normal(0.90,0.1);
  // model ZAR
  for(t in 1:T) {
    lf[t]= fisher_z_lpdf(y[t]|d1,d2,phi0,sigma);
  }
  // likelihood
  target += sum(lf);
}
generated quantities {
  real x_p[U_p];
  real log_lik[T];
  for(t in 1:T) {
    log_lik[t] = fisher_z_lpdf(y[t]|d1,d2,0,sigma);
  }
  for(t in 1:U_p) {
    if(t<=1) {
      x_p[t] = x[t];
    } else {
      x_p[t] = exp((fisher_z_rng(d1,d2,phi0,sigma))/100+log(x[t-1]));
    }
  }
}
}
```