On the Flexible Neo-Normal MSAR MSN-Burr Control Chart in Air Quality Monitoring

Authors: Rasyid, Dwilaksana Abdullah, Iriawan, Nur, and Mashuri, Muhammad

Source: Air, Soil and Water Research, 17(1)

Published By: SAGE Publishing

URL: https://doi.org/10.1177/11786221241272391

On the Flexible Neo-Normal MSAR MSN-Burr Control Chart in Air Quality Monitoring

Air, Soil and Water Research
Volume 17: 1–15

DOI: 10.1177/11786221241272391
Sage

Dwilaksana Abdullah Rasyid, Nur Iriawan and Muhammad Mashuri
Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

ABSTRACT: Air quality significantly influences human health and the environment, necessitating a robust monitoring to detect abnormalities. This paper aims to develop a new model to accurately capture air quality data's structural changes and asymmetrical patterns. We introduce the neo-normal Markov Switching Autoregressive (MSAR) Modified Skew Normal Burr (MSN-Burr) model, called neo-normal MSAR MSN-Burr. This model extends the MSAR normal framework, handling symmetrical and asymmetrical patterns in air quality data. The MSN-Burr distribution is employed for accurate estimation of skewed and symmetric data. The model efficiency is demonstrated through simulation studies generating symmetric data with normal, double exponential, and Student-*t* distributions, followed by application to real air quality data using Stan language. The proposed model successfully adapts to asymmetric structural changes, as evidenced by creating the Highest Posterior Distribution (HPD) for upper and lower limits. The model identifies two regimes representing normal and abnormal air quality conditions, with modes of 8 and 19 μg/m^3, respectively. The MSAR-MSN-Burr model exhibits a 32.27% RMSE improvement in simulations and a 16.4% RMSE improvement in real air quality data over the normal-MSAR model. The proposed neo-normal MSAR MSN-Burr model is significantly enhancing the accuracy of air quality monitoring, providing a more efficient tool for detecting air quality abnormalities.

KEYWORDS: Air quality, highest posterior distribution (HPD), Markov switching autoregressive (MSAR), modified skew normal Burr (MSN Burr)

RECEIVED: February 4, 2024. **ACCEPTED:** July 15, 2024.

TYPE: Research Article

CORRESPONDING AUTHOR: Nur Iriawan, Statistics Department, Faculty of Science and Data Analytics, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia. Email: nur_i@statistika.its.ac.id

Introduction

Air pollution is a complex mixture of solid particles, liquid droplets, and gases. It can come from many sources, for example: domestic fuel combustion, industrial chimneys, motor vehicles, power plants, open waste burning, agricultural activities, dust, and many other sources (Ouyang et al., 2022). According to the World Health Organization (WHO), air pollution is measured by many variables, namely PM2.5 and PM10 (particles with an aerodynamic diameter equal to or less than 2.5, also called fine, and every 10μm), ozone (O_3), nitrogen dioxide (NO_2), carbon monoxide (CO), and sulfur dioxide (SO_2) (Yang et al., 2020). Fine particles (PM10 and PM2.5) can through the lungs and then enter the body through the bloodstream, affecting all major organs (Thangavel et al., 2022; Yang et al., 2020). This can cause illness in both the cardiovascular and respiratory systems, leading to diseases such as stroke, lung cancer, and chronic obstructive pulmonary disease (Choung & Kim, 2019; Ren & Tong, 2008; Wright et al., 2023). Recent research also shows a link between prenatal exposure to high levels of air pollution and developmental delays in 3-year-old children, as well as psychological and behavioral problems later in life, including symptoms of attention deficit hyperactivity disorder (ADHD), difficulty focusing, anxiety and depression (Johnson et al., 2021; Kaur et al., 2023).

Several analytical methods are useful for monitoring air quality, including machine learning methods. Natarajan et al. (2024) applies several machine learning methods to monitor air pollution in several cities in India using k-nearest neighbor, random forest regressor, and support vector regressor models. Other analytical methods such as Markov switching models are used to understand when switching between anomalous and non-anomalous conditions in air quality occurs. The Markov switching model is a statistical analysis tool for identifying regime shifts in time series data (Franke, 2012). In the context of air quality, this can help identify periods when air pollution reaches abnormal or dangerous levels. This model allows us to group data into two or more different regimes, each with different statistical characteristics (Gao, 2020; Zakaria et al., 2019).

Numerous research studies have examined the regime-switching model specifically the Markov Switching Autoregressive (MSAR) model, applied in various field. These several studies show the versatility and effectiveness of the MSAR model in analyzing different type of data. Table 1 explaining the summary of key research employing the MSAR model, including their objective and estimation methods. These studies demonstrate the MSAR model's ability to handle different type of data, from financial market to wind speed and mortality rates, highlighting its flexibility and robustness in various applications. The unique characteristic of MSAR is the ability to handle the regime switching without defining the threshold area first like the TAR model (Zhang et al., 2023). The original estimation method in MSAR is using EM. then the estimation method was developed using the Bayesian method to obtain optimal estimation results was proposed by Kim and Nelson (2000) and Hamilton and Raj (2002).

Table 1. Summary Key Research Employing the MSAR Model.

AUTHORS	APPLICATION	OBJECTIVE	ESTIMATION METHOD	PROS	CONS
Adejumo et al. (2021)	Nigeria's stock market (All-Share Index)	Determine bear and bull phases of stock market volatility	Expectation Maximization (EM)	Effective in identifying market phases	Limited by normality assumption
Ailliot and Monbet (2012)	Wind speed data	Predict wind speed and direction for optimal energy production	Expectation Maximization (EM)	Accurate prediction of wind patterns	Requires normality assumption
Troug and Murray (2021)	Hong Kong and Tokyo stock markets	Crisis identification and financial contagion analysis	Bayesian estimation	Handles non-normal data very well	Computationnally intensive
Fu et al. (2023)	Mortality data	Capture transient variations in mortality for risk management	Bayesian estimation	Provides insightful quantitative mortality data	More complex extimation process

There are several methods for estimating model parameters in Bayesian modeling. Some of the studies mentioned above used the Bayesian method coupled with the Gibbs sampling algorithm to estimate the parameters developed by Sims et al. (2008). Apart from that, there are also several developments in the Gibbs Sampler algorithm, including Hamiltonian Monte Carlo (HMC). HMC is an estimation method that uses the same Markov Chain Monte Carlo (MCMC) as the Gibbs sampling (Duane et al., 1987; Neal, 2011). The performance of HMC effectively mitigates the random walk behavior and correlated parameter sensitivity issues common in MCMC methods by employing first-order gradient information-based steps. The disadvantage of HMC is the number of leapfrogs, if it is too small, the algorithm will show undesirable random walk behavior, while if it is too large the computation will take a long time. To overcome that problem Hoffman and Gelman (2014) developed an extension of HMC called No U-Turn Sampler (NUTS). This algorithm can simplify the problem of HMC by automating the number of leapfrogs. NUTS also automatically stops simulation iterations if it approaches a U-Turn pattern, helping to avoid inefficient sampling of the posterior distribution.

Meanwhile, the development of the MSAR model using Bayesian estimation has been completed by Li et al. (2022) using Just Another Gibbs Sampler (JAGS) software. Their development combines the zero-inflated multilevel Poisson distribution with an autoregressive model which applied to longitudinal data. JAGS allows users to write their functions, distributions, and samplers (Wabersich & Vandekerckhove, 2014). JAGS is a development of Bayesian Inference Using Gibbs Sampling (BUGS), both of which use the MCMC algorithm for estimation. Adding new distributions to the BUGS program is complicated and requires other programs such as BlackBox Component Builder (Wetzels et al., 2010). Just like BUGS, adding new distributions to JAGS is also complicated because of testing and validation requirements, and the necessity for a clear document (Wabersich & Vandekerckhove, 2014). Different from the two previous software, in the Stan language, users can use several features to create distribution-based models that allow researchers to build based on their creativity (Annis et al., 2017). Therefore, researchers can build various models based on data-driven analysis. Modeling with Stan is widely available in several interfaces in several software. The most popular and widely used are RStan in R and PyStan in Python. With Stan available on many interfaces, it will be easier for researchers to apply the proposed method (Annis et al., 2017). In this study, we used RStan.

In real cases, not all data has a normal pattern, especially data that is suspected to have anomalous events. These anomalies can distort statistical analyses and avoid accurate modeling. In response to this challenge, researchers have developed innovative approaches such as replacing the normative Gaussian-based models with other distributions. This study approach was once carried out by Deschamps (2006) by replacing the error in the MSAR model with a Student-t distribution. Different from the approach taken by Deschamps (2006), a skewed normal Azzalini distribution which is intended by Azzalini (1985) is used to replace the error distribution carried out by Lhuissier (2019). Their studies both used the Gibbs sampling algorithm to estimate parameters. Handling skewed pattern data in this research, we propose a simulation-based model estimation using NUTS which is applied to neo-normal data distribution, namely Modified Skew Normal Burr (MSN-Burr).

This study aims to create a user-defined neo-normal Markov Switching Autoregressive Modified Skew Normal Burr (neo-normal MSAR MSN-Burr) model. Next, the proposed model is compared with the normal MSAR model to determine its ability to analyze simulation data and PM10 data. The simulation was made from three scenarios with three different distributions with the aim of finding out in general

that the proposed model is able to deal with symmetric and asymmetric data conditions. The neo-normal MSAR MSN-Burr model is able to demonstrate the flexibility of the adaptive MSAR model for various data-driven distributions. The RMSE is used as the evaluation metric to better understand which model performs more effectively. Furthermore, adaptive control limits are created for each regime which are built using the highest posterior distribution for air quality mapping. A more mathematical and in-depth explanation of the MSN-Burr distribution can be seen in Iriawan (2000).

The rest of this paper is organized as follows. The next section introduces the general MSAR model. The Stan code for the general MSAR model can be seen in Osmundsen et al. (2021). The following section describes the MSN-Burr distribution and the neo-normal MSAR MSN-Burr model and demonstrates the Stan code according to the mathematical model description. The next section is comparison between the HMC and NUTS algorithm, in this section we explain the efficiency of NUTS over the HMC. After that we explain how the proposed model is estimated using a combination of EM and NUTS which we call EM-NUTS estimation. We have also provided a combination of the two estimation methods in the Stan language by adding the MSN-Burr distribution to Stan. The next section shows a simulation study comparing normal MSAR with neo-normal MSAR MSN-Burr. Simulation studies are carried out by generating data that has characteristics such as regime switching with errors using normal, double exponential, and Student-t distributions. After that we applied the neo-normal MSAR MSN-Burr model on Yogyakarta air quality data in 2021 and showing the result of the neo-normal MSAR MSN-Burr model that has converged and applied the HPD for each regime for several levels of significance. The conclusions are given in the last section.

Markov Switching Autoregressive Model

Markov switching models can be combined with time series models such as autoregressive models used to identify changes in conditions or time series data patterns (Hamilton, 1989). The forms of a Markov switching autoregressive (MSAR) model can be written in the following Equation 1.

$$y_t - \mu_{s_t} = \beta_1\left(y_{t-1} - \mu_{s_{t-1}}\right) + \ldots + \beta_p\left(y_{t-p} - \mu_{s_{t-p}}\right) + e_t, \quad (1)$$

with e_t is residual as $e_t \sim N(0,\sigma_{s_t}^2)$, s_t is regime (unobserved random variable), $y_t,\ldots,y_{t-p}$ are observation data, $\beta_1,\ldots,\beta_p$ are the autoregressive coefficient of order p, $\sigma_{s_t}^2$ is the variance that is influenced by regime changes, and $\mu_{s_t},\ldots,\mu_{s_{t-p}}$ are mean that is influenced by changes in regime.

There are two types of modeling steps in MSAR, which are regime transition estimation and parameter estimation for each regime (Frühwirth-Schnatter, 2009). Regime transfer is an unknown condition, which is why it is called a latent variable. However, the number of regimes can be determined by various combinations. A comprehensive discussion of the MSAR model has been conducted by Kim and Nelson (1999). In this study, we refer to the normal MSAR model in Stan as a first introduction.

Neo-Normal Markov Switching Autoregressive Modified Skew Normal Burr (Neo-Normal MSAR MSN-Burr) Model

The MSN Burr distribution is a relaxation of the normal distribution developed by Iriawan (2000) from the Burr II distribution (Burr, 1942). Its cumulative distribution function (CDF) and probability density function (pdf) are

$$F\left(y_t^*\right) = \left(1 + \frac{e^{-y_t^*}}{\lambda}\right)^{-\lambda}, \quad (2)$$

$$f\left(y_t^*|\lambda\right) = e^{-y_t^*}\left(1 + \frac{e^{-y_t^*}}{\lambda}\right)^{-(\lambda+1)}, \quad (3)$$

By performing the transformation, the CDF in Equation 2 and the pdf in Equation 3 are called the Modified Stable Burr or MS-Burr distribution. The mode of the MS-Burr distribution is stable at any value. However, the adjustment is needed because when compared with the standard normal distribution, $N(0.1)$, the mode of the MS-Burr pdf value is lower. Furthermore, a transformation to fit the normal distribution such that the CDF and pdf were obtained in the form of Equations 4 and 5.

$$F(y_t) = \left(1 + \frac{\exp\left(-k\left(\frac{y_t-\mu}{\sigma}\right)\right)}{\lambda}\right)^{-\lambda}, \quad (4)$$

$$f\left(y_t|k,\lambda,\mu,\sigma\right) = \frac{k}{\sigma}\exp\left(-k\left(\frac{y_t-\mu}{\sigma}\right)\right)\left(1 + \frac{\exp\left(-k\left(\frac{y_t-\mu}{\sigma}\right)\right)}{\lambda}\right)^{-(\lambda+1)}, \quad (5)$$

where $-\infty < y_t < \infty, -\infty < \mu < \infty, k > 0, \lambda > 0$, and $\sigma > 0$ with the μ is mode, σ is variance and λ is the skewness parameter of this distribution. The k value is obtained from the difference between the pdf of the standard normal distribution, $N(0,1)$, and the pdf of the MSN-Burr distribution. Then the difference between the two distributions is equalized to zero, then the mode and scale parameter values in the MSN-Burr pdf are the same as in the standard normal pdf. The k can be written in Equation 6.

$$k = \frac{1}{\sqrt{2\pi}}\left(1 + \frac{1}{\lambda}\right)^{(\lambda+1)}, \quad (6)$$

By substituting the k into Equation 5, we obtain the MSN-Burr pdf which is detailed in Equation 7.

$$f(y_t \mid \lambda,\mu,\sigma) = \frac{\frac{1}{\sqrt{2\pi}}\left(1+\frac{1}{\lambda}\right)^{(\lambda+1)}}{\sigma}\exp\left(-\frac{1}{\sqrt{2\pi}}\left(1+\frac{1}{\lambda}\right)^{(\lambda+1)}\left(\frac{y_t-\mu}{\sigma}\right)\right)\left(1+\frac{\exp\left(-\frac{1}{\sqrt{2\pi}}\left(1+\frac{1}{\lambda}\right)\left(\frac{y_t-\mu}{\sigma}\right)^{(\lambda+1)}\right)}{\lambda}\right)^{-(\lambda+1)}, \quad (7)$$

This distribution is then added to the MSAR model in Equation 1 by replacing the residuals of the normal distribution with the MSN-Burr distribution. The definition of neo-normal MSAR MSN-Burr model is explained in Equation 8 including how to estimate its parameter.

Parameter Estimation

The combination of MSAR with MSN-Burr distribution can lead to complex estimation methods. To overcome this problem, we estimate this model using the combination of EM algorithm which proposed by Dempster et al. (1977) with NUTS which proposed by Hoffman and Gelman (2014). To enable a better understanding, we include pseudocode for both the HMC and NUTS methods, as well as a comparison of their differences. The pseudocode of HMC can be seen in Algorithm 1, meanwhile the NUTS in Algorithm 2.

Algorithm 1. HMC algorithm.

Input $\theta^{(0)}$, ε, L, and N
Initialize $\theta = \theta^{(0)}$
for i to N do
  sample momentum $\tau^{(0)}$ from Normal (0,1)
  set $(\theta', \tau') = (\theta, \tau^{(0)})$
  for j = 1 to L do
    $\tau' = i = \tau' - (\varepsilon/2) * \nabla U(\theta')$
    $\theta' = \theta' + \varepsilon * \tau'$
    $\tau' = \tau' - (\varepsilon/2) * \nabla U(\theta')$
  end for
  compute acceptance probability α
  sample u from Unifrom (0,1)
  if u<α then
    $\theta = \theta'$
  end if
  save θ as a sample
end for

Algorithm 2. NUTS algorithm.

Input $\theta^{(0)}$, ε, L, and N
Initialize $\theta = \theta^{(0)}$
for i = 1 to N do
  sample momentum $\tau^{(0)}$ from Normal (0,1)
  initialize θ^+, $\theta^- = \theta$, τ^+, $\tau^- = \tau^{(0)}$
  initialize path length to 0
  while not U-turn do
    if random choose direction = forward then
      θ^+, τ^+ = Leapfrog (θ^+, τ^+ ε)
    else
      θ^-, τ^- = Leapfrog (θ^-, τ^- ε)
    end if
    check U-turn comdition
    increase depth
    save θ if it improves the trajectory
  end while
  compute acceptance probability α
  sample u from Unifrom (0,1)
  if u<α then
    θ = new θ
  end if
  adapt ε if in warm up phase
  save θ as sample
end for

The main difference between HMC and NUTS lies in how they determine the length of the trajectory for sampling. HMC use pre-define leapfrog steps to simulate the Hamiltonian dynamics which is very good for carefully estimating parameters. In contrast, NUTS automatically define the leapfrog steps which will have an impact on trajectory length. This automatic determination is assisted by binary tree calculations which are useful for checking the U-turns. The addition of the automatic leapfrog determination step and the binary tree calculation in the NUTS algorithm prevents it from getting trapped in local optima, enabling it to achieve optimal parameter estimates more efficiently.

Combination EM-NUTS Estimation

The key combination of the EM and NUTS is only using the Expectation step in EM then the Maximization step using NUTS. Before we calculate the expectation step, we need to find the likelihood using the joint density of y_t, s_t and s_{t-1}.

1. Deriving the joint density function of y_t, s_t and s_{t-1} conditionally on past information ψ_{t-1}

$$f(y_t, s_t, s_{t-1} \mid \psi_{t-1}) = f(y_t \mid s_t, s_{t-1}, \psi_{t-1}) \Pr[s_t, s_{t-1} \mid \psi_{t-1}]$$

Where Ψ_{t-1} is the observation value on $t-1$ of time, while y_t is the observation value at the time t which follow the MSAR model using MSN-Burr distribution on Equation 7 with parameter $\theta_{s_t} = (\lambda_{s_t}, \mu_{s_t}, \sigma_{s_t}, \beta_{0,s_t}, \beta_{1,s_t})$ is given by Equation 8.

$$f(y_t \mid s_t, s_{t-1}, \Psi_{t-1}) = \frac{k_{s_t}}{\sigma_{s_t}} exp\left(-k_{s_t}\left(\frac{\beta_{0,s_t} + \beta_{1,s_t} + (y_t - \mu_{s_t})}{\sigma_{s_t}}\right)\right)$$
$$\left(1 + \frac{exp\left(-k_{s_t}\left(\frac{\beta_{0,s_t} + \beta_{1,s_t} + (y_t - \mu_{s_t})}{\sigma_{s_t}}\right)\right)}{\lambda_{s_t}}\right)^{-(\lambda_{s_t}+1)} . \quad (8)$$

Since the s_t parameter was added to the MSN-Burr distribution, the MSN-Burr MSAR model equation has parameters in each regime. Mode is μ_{s_t}, variance is σ_{s_t}, skewness is λ_{s_t}, autoregressive parameter is β_{0,s_t} and β_{1,s_t}.

2. Find the function $f(y_t|\psi_{t-1})$ by integrating s_t *and* s_{t-1} *from the joint density over all possible values of* s_t *and* s_{t-1}

$$f(y_t \mid s\Psi_{t-1}) = \sum_{S_t=1}^{M}\sum_{S_{t-1}=1}^{M} f(y_t, s_t, s_{t-1} \mid \Psi_{t-1})$$
$$= \sum_{s_t=1}^{M}\sum_{s_{t-1}=1}^{M} f(y_t \mid s_t, s_{t-1}, \Psi_{t-1}) \Pr[_t, s_{t-1} \mid \Psi_{t-1}]$$

where $f(y_t \mid \Psi_{t-1})$ is the marginal density which is the average of conditional density weighted by

$$\Pr[s_t = j, s_{t-1} = i \mid \Psi_{t-1}], i = 1,2,...,M; j = 1,2,...,M$$

where M is the number of regimes inside the model. The likelihood function is written below.

$$ln\, L = \sum_{t=1}^{T} ln\{\sum_{s_t=1}^{M}\sum_{s_{t-1}=1}^{M} f(Z_t \mid s_t, s_{t-1}, \Psi_{t-1}) \Pr[s_t, s_{t-1} \mid \Psi_{t-1}]\}$$

To calculate $\Pr[s_t, s_{t-1} \mid \Psi_{t-1}]$ is going to be solved using the filtering process. In short, the filtering process aims to get the $\Pr[s_t = j, s_{t-1} = i \mid \Psi_{t-1}]$ value.

$$\Pr[s_t = j, s_{t-1} = i \mid \Psi_{t-1}] = \frac{f(y_t \mid s_t = j, s_{t-1} = i, \Psi_{t-1}) \times \Pr[s_t = j, s_{t-1} = i \mid \Psi_{t-1}]}{\sum_{s_t=1}^{M}\sum_{s_{t-1}=1}^{M} f(y_t \mid s_t = j, s_{t-1} = i, \Psi_{t-1}) \times \Pr[s_t = j, s_{t-1} = i \mid \Psi_{t-1}]}$$

After the Expectation step, we move to Maximization step using the NUTS algorithm. In this step, we aim to optimize the parameter $\theta_{s_t} = (\lambda_{s_t}, \mu_{s_t}, \sigma_{s_t}, \beta_{0,s_t}, \beta_{1,s_t})$.

1. Initialize the parameter value in each regime θ_{s_t} and step size ε on the first iteration.
2. Draw initial momentum of $\tilde{\tau}$ at every iteration from $\tau_d \sim MultiNormal\,(0, \sum_{dd})$
3. Use the leapfrog integration method to simulate Hamiltonian dynamics. There are three main steps:
   a. Update momentum (half-step):
   $$\tilde{\tau}_{t+\frac{1}{2}} = \tilde{\tau}_t - 2\varepsilon\, \nabla U(\theta_{s_t})$$
   b. Update position:
   $$\theta_{s_{t-1}} + 1 = \theta_t + \varepsilon\, \tilde{\tau}_{t+\frac{1}{2}}$$
   c. Update momentum (half-step):
   $$\tilde{\tau}_{t+1} = \tilde{\tau}_{t+\frac{1}{2}} - \frac{\varepsilon}{2} \nabla U(\theta_{s_{t+1}})$$
4. Build the binary tree
   a. Expand both forward and backward in time:
      - Forward:
      $$\theta^+, \tilde{\tau}^+ \leftarrow Leapfrog\,(\theta^+, \tilde{\tau}^+, +\varepsilon)$$
      - Backward:
      $$\theta^-, \tilde{\tau}^- \leftarrow Leapfrog\,(\theta^-, \tilde{\tau}^-, -\varepsilon)$$
   b. Check the U-Turn
   $$(\theta^+ - \theta^-)^\top\, \tilde{\tau}^+ < 0 \; or \; (\theta^+ - \theta^-)^\top\, \tilde{\tau}^- < 0$$
5. Calculate the acceptance rejection
   a. Calculate the Hamiltonian for the initial and proposed states:
   $$H(\theta_{s_t}, \tilde{\tau}) = U(\theta_{s_t}) + \frac{1}{2} \tilde{\tau}^\top\, \tilde{\tau}$$
   b. Calculate the acceptance probability
   $$\alpha = min(1, exp(H(\theta_{s_t,0}, \tilde{\tau}_0) - H(\theta_{s_t}', \tilde{\tau}')))$$
   c. Accept or reject the proposed θ_{s_t}':
   $$\theta_{s_t} new = \begin{cases} \theta_{s_t}' \text{ with probability } \alpha \\ \theta_{s_t,0} \text{ with probability } 1-\alpha \end{cases}$$
6. Repeat the procedures for the desired number of iterations to produce samples from the specified distribution

The NUTS algorithm is computationally intensive because it requires calculating the binary tree structure at each iteration in order to dynamically adjust the trajectory length. This process involves multiple leapfrog integration phases for both forward and backward expansion, as well as repeated checks for U-turn condition. This algorithm will make the parameter estimation process in the MSN-Burr MSAR model in Equation 8 converge more quickly. We briefly summarize the overall combination of the two estimation methods which can be seen in Algorithm 3.

Algorithm 3. Combined EM-NUTS estimation method.

Initialization: Initialize all the parameter list $\theta_{s_t}^{(0)}$ and $s_t^{(0)}$
Repeat until convergence:
  E-Step: Compute the expected log-likelihood on Equation 8
  M-Step using NUTS:
    Set up the potential energy function $\tilde{\tau}_d \sim MultiNormal(0, \Sigma_{dd})$
    Perform NUTS to draw samples from the posterior distribution of θ_{s_t}
    Update the parameter estimates θ_{s_t}' with NUTS samples
Check for convergence: Determine if the algorithm has converged based on predefine criteria and update s_t'.

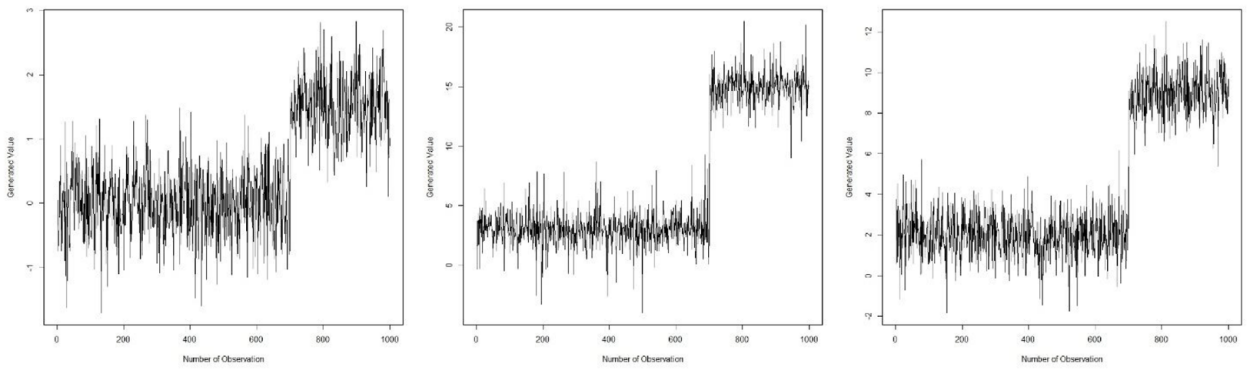


Figure 1. Time series plot of generated data using (a) normal, (b) double exponential, and (c) Student-*t* distributions.

To simplify the estimation process using the proposed algorithm, we use the Stan language available in R. The Stan language was developed to overcome convergence problems that commonly occur in Bayesian inference using Gibbs sampling (BUGS) languages (Gelman et al., 2015). The key stages of the Stan Language are data and model input, calculating the log of the pdf and their gradients, a warm-up phase for parameter tuning, applying NUTS, monitoring convergence, and calculating summary inference (Hoffman & Gelman, 2014). Standard distributions such as normal, Poisson, binomial, etc. are already available in the booth. However, Stan is very flexible in adding new distributions by writing a log of pdf from that distribution. Stan uses a numerical auto-differentiation method by utilizing reverse-mode automatic differentiation to automatically perform function reduction (Carpenter et al., 2017).

Adding the MSN-Burr Distribution in Stan

Adding a new distribution to the BUGS software is very complex and requires other programs such as BlackBox Component Builder, as stated in the introduction (Wetzels et al., 2010). Similarly, JAGS is also complicated for adding distribution because it has many complicated steps that must be followed, as stated by Wabersich and Vandekerckhove (2014). The steps for adding a new distribution in Stan are relatively simple, users only need to know the mathematical form of the distribution to be added. This convenience gives researchers an advantage for adding new distributions, like MSN-Burr. Instructions for adding new distributions are explained in Annis et al. (2017). Based on Equation 7, the addition of the MSN-Burr distribution syntax in the Stan code can be seen in Github https://github.com/Rasyid/MSAR_MSN-Burr.

Adding the Neo-Normal MSAR MSN-Burr Distribution in Stan

Adding a normal distribution to the MSAR model can be seen in Osmundsen et al. (2021). In this study, the proposed user-defined Stan code for the neo-normal MSAR MSN-Burr was created. Based on Equation 1, the model of MSAR was declared then the error distribution was replaced by using Equation 7. The addition of the neo-normal MSAR MSN-Burr model syntax in the Stan code can be seen in Github https://github.com/DwilaksanaAbdullahRasyid/MSAR_MSN-Burr.

Simulation Study

From Equation 7, when the value of the skewness parameter $\lambda = 1$, it can be seen that the MSN-Bur distribution becomes a normal distribution, which is symmetric and bell-shaped. Therefore, normally distributed data are used to validate the user-defined MSN-Burr distribution in the Stan program, allowing Stan to use the user-defined MSN-Burr distribution for estimation. The fact shows that MSN-Burr can accurately and consistently estimate μ and σ in accordance with the generator parameters serves as evidence of its validity in detecting normal data.

We then present comparative evidence between the standard conventional MSAR model and the neo-normal MSAR MSN-Burr model in Stan. We created three different scenarios involving the normal distribution, the double exponential, and the Student-*t* distribution. Three distributions are used to demonstrate that neo-normal MSAR MSN-Burr is capable of collecting symmetric distribution characteristic data. These scenarios contain a challenging scheme, which is used to compare the capabilities of the neo-normal MSAR MSN-Burr model with the standard MSAR normal model to detect data with different variances and zero-centered data using leptokurtic properties. In each scenario, we generate two different time series observations based on the selected distribution. As many as 700 observations are generated based on the selected distribution, then 300 observations with different parameters to obtain a total of 1,000 observations with a proportion of 0.7 in regime 1 and 0.3 in regime 2. We then estimate each regime in each scenario using a standard autoregressive model to find the target parameter. A visualization of the generated non-normal time series data is displayed in Figure 1. The target parameters are then derived based on the autoregressive models of each

Table 2. The Target Parameter from Three Scenarios for Normal, Double Exponential, and Student-*t* Distribution.

PARAMETER	GENERATED DATA BASED ON DISTRIBUTIONS		
	NORMAL	DOUBLE EXPONENTIAL	STUDENT-*T*
β_{01}	2.0293	3.0558	2.0168
β_{02}	6.0048	14.9457	8.9786
β_{11}	0.0341	−0.0248	−0.0464
β_{12}	0.0168	0.0957	0.0196
ρ_1	0.7000	0.7000	0.7000
ρ_2	0.3000	0.3000	0.3000
σ_1	1.0046	1.4133	1.0620
σ_2	0.9854	1.3564	1.0899
λ_1	1.0000	1.0000	1.0000
λ_2	1.0000	1.0000	1.0000

scenario and are available in Table 2. Target parameter in Table 2, we assign a value to the slope parameter $\lambda = 1$ to mimic the symmetric conditions that would be estimated by neo-normal MSAR MSN-Burr.

From the three scenarios above, we obtain the 95% credible interval for each parameter as seen in Table 3. We can use the 95% credible interval in Table 3 as a parameter significance test by providing a range of values for the parameter of interest that are considered plausible given the observed data. In the context of hypothesis testing, if the target parameter value in Table 3 falls within the interval, it suggests that the data is consistent with that value, supporting the null hypothesis. Conversely, if the target parameter is outside the credible interval, it leads to reject the null hypothesis (Martinez & Martinez, 2016).

Each scenario was replicated 200 times. In each replication, estimation is carried out using the NUTS algorithm which is available on Stan through 10,000 iterations. Since each scenario was replicated 200 times and in every one of them was calculated using the sampler from NUTS, we computed the root-mean-square error (RMSE) to evaluate and compare the quality of parameter estimates within each scenario. The RMSE for estimated $\hat{\theta}$ is defined in Equation 9 (Walther & Moore, 2005):

$$\text{RMSE}_{\hat{\theta}} = \sqrt{\frac{\sum_i^{n_{rep}} (\theta - \hat{\theta}_i)^2}{n_{rep}}}, \quad (9)$$

where θ is the list of target parameter from Table 2, $\hat{\theta}$ is the estimation result in each replication, and n_{rep} is the number of replications. Since the θ is the list of target parameter, the RMSE will be calculated for each parameter. RMSE performance can be seen from a value close to zero. When the RMSE approaches zero, the parameter estimate is better.

All the target parameter in Table 3 lies within the 95% credible interval. Likewise, for that the MSN-Burr distribution in the neo-normal MSAR MSN-Burr model is able to estimate data generated from symmetric distributions, namely normal, double exponential, and Student-t distributions as can be seen at the λ value in each regime that is close to one. For the autoregressive parameters, namely β_{11} and β_{12}, they have their own hypothesis, namely that they must be inside unity, which means they do not have a zero value in the credible interval. If the autoregressive parameter is zero then the model essentially becomes a random walk. The evidence that none of the intervals for the parameters β_{11} and β_{12} have a value of zero is presented in Table 3. So, it can be concluded from the results of the 95% credible interval the neo-normal MSAR MSN-Burr model has succeeded in identifying random data points with a symmetrical distribution.

The results of the RMSE from each parameter calculation are shown in Table 4. This RMSE was calculated through 200 replications, with each replication generate estimated value from each parameter. This approach allows us to assess the error for each parameter across all replications. In the generated normal distribution scenario, the RMSE value of the sigma parameter in the neo-normal MSAR MSN-Burr model is slightly higher than in the normal MSAR model. In the generated double-exponential and generated Student-t distribution scenarios, the RMSE value of the p parameter in the neo-normal MSAR MSN-Burr model is slightly higher than normal MSAR. This doesn't mean that the neo-normal MSAR MSN-Burr has poor performance. From several explanations given, it can be concluded that the MSAR model that uses MSN-Burr distributed error can detect symmetric random data points from the normal, double exponential, and Student-*t* distribution.

We also calculate the RMSE for each observation to find out whether the predicted value is around what the observation should be or not. This RMSE comparison was carried out by the normal MSAR model with the MSN-Burr MSAR to find out which model is better at capturing symmetrical patterns in the three simulation scenarios. RMSE calculations for predicted values from observations are presented in Table 5.

The observation RMSE calculation produces a percentage value of the difference between the normal MSAR model and the neo-normal MSAR MSN-Burr in the three scenarios. Table 5 shows that neo-normal MSAR MSN-Burr is 14.0408% better in modeling the generated normal data scenario and 82.7683% in modeling the generated student-t data compared with the normal MSAR model. However, both MSAR models seem unable to model data that has very high leptokurtic properties in the Generated Double Exponential scenario. The percentage difference in this scenario is only 0.0046%, which shows that the two models provide the same estimation results for estimating the generated data with high leptokurtic properties.

Table 3. The 95% Credible Interval for Estimated Parameters of Three Scenario Simulations for Normal, Double Exponential, and Student-*t* Distributions.

SCENARIO	PARAMETER	95% CREDIBLE INTERVAL MSAR NORMAL MODEL		95% CREDIBLE INTERVAL NEO-NORMAL MSAR MSN-BURR MODEL	
		LL	UL	LL	UL
Generated normal data	β_{01}	1.4918	4.4641	0.3986	2.4720
	β_{02}	2.5360	3.0465	6.2793	6.4983
	β_{11}	0.3161	0.4060	0.0015	0.0489
	β_{12}	0.3184	0.4191	0.0130	0.0497
	ρ_1	0.6284	0.7095	0.0132	0.7500
	ρ_2	0.6345	0.7171	0.9813	0.9899
	σ_1	0.9466	1.0884	0.1570	1.0946
	σ_2	0.9380	1.0870	1.0968	1.1000
	λ_1	–	–	0.0753	1.0916
	λ_2	–	–	0.9941	1.0999
Generated double exponential data	β_{01}	4.4641	6.0707	0.0874	2.1843
	β_{02}	3.6188	4.8912	14.2539	14.9941
	β_{11}	3.6188	4.8912	−0.0459	−0.0004
	β_{12}	0.3746	0.5101	0.1504	0.1996
	ρ_1	0.8493	0.9357	0.0096	0.6184
	ρ_2	0.8482	0.9335	0.9828	0.9899
	σ_1	1.5150	1.8433	0.9801	1.4966
	σ_2	1.5245	1.8613	1.4982	1.5000
	λ_1	–	–	0.0121	1.0887
	λ_2	–	–	0.9933	1.0999
Generated student-*t* data	β_{01}	3.5964	4.2958	0.2587	2.1465
	β_{02}	3.8055	4.5659	8.7158	8.9979
	β_{11}	0.1950	0.2921	−0.0482	−0.0010
	β_{12}	0.1993	0.3018	0.0189	0.0497
	ρ_1	0.7821	0.8523	0.0120	0.7251
	ρ_2	0.7811	0.8515	0.9822	0.9899
	σ_1	1.0419	1.1970	0.4714	0.0961
	σ_2	1.0450	1.2027	1.0981	1.1000
	λ_1	–	–	0.0191	1.0063
	λ_2	–	–	0.9939	1.0999

Note. LL and UL are the lower limit and upper limit respectively for a 95% credible interval.

Application

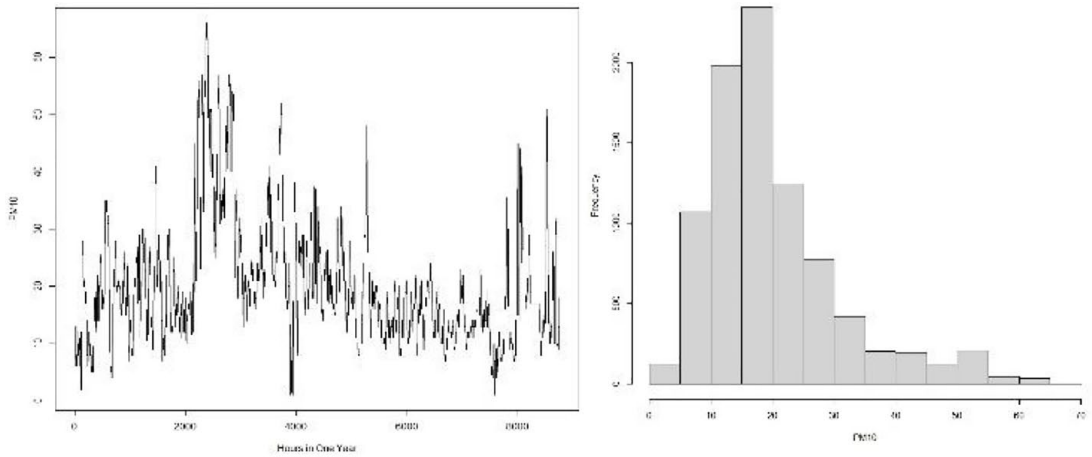
This session discusses the application of the neo-normal MSAR MSN-Burr model using an air quality dataset from Yogyakarta, Indonesia, collected in 2021 and recorded hourly. This dataset is openly available from the Yogyakarta City Environmental Service. This dataset has 8,760 observations and has several missing values, we handle it by employ moving average. We use this method by taking the average of the previous 24 observations, because the data used is data taken hourly so the assumption of taking the average of the previous 24 hr.

Table 4. Root-Mean-Square Error (RMSE) for the Estimated Parameter of Three Scenario Simulations for Normal, Double Exponential, and Student-*t* Distributions.

SCENARIO	PARAMETER	RMSE MSAR NORMAL MODEL	RMSE NEO-NORMAL MSAR MSN-BURR MODEL
Generated normal data	β_{01}	2.0319	0.3116
	β_{02}	3.9929	0.4341
	β_{11}	0.5543	0.0603
	β_{12}	0.5457	0.0562
	ρ_1	0.4133	0.3997
	ρ_2	0.5486	0.2878
	σ_1	0.0296	0.1700
	σ_2	0.0364	0.1137
	λ_1	–	0.3232
	λ_2	–	0.0015
Generated double exponential data	β_{01}	6.4698	2.2550
	β_{02}	6.8951	0.1528
	β_{11}	6.5646	0.0083
	β_{12}	0.5976	0.0906
	ρ_1	0.2613	0.4702
	ρ_2	0.2653	0.2880
	σ_1	0.8702	0.0504
	σ_2	0.8463	0.1430
	λ_1	–	0.6497
	λ_2	–	0.0017
Generated student-*t* data	β_{01}	3.9816	0.7928
	β_{02}	5.9158	0.0566
	β_{11}	0.4623	0.0236
	β_{12}	0.4348	0.0217
	ρ_1	0.3581	0.4157
	ρ_2	0.3606	0.2879
	σ_1	0.0946	0.1192
	σ_2	0.0891	0.0095
	λ_1	–	0.5282
	λ_2	–	0.0016

Table 5. Root-Mean-Square Error (RMSE) for Predicted Observation of Three Scenario Simulation for Normal, Double Exponential, and Student-*t* Distributions.

SCENARIO	MSAR NORMAL	NEO-NORMAL MSAR MSN-BURR	PERCENTAGE DIFFERENCE
Generated normal data	1.7355	1.4918	14.0408
Generated double exponential data	7.8336	7.8333	0.0046
Generated student-*t* data	4.7477	0.8179	82.7683



(a) (b)

Figure 2. (a) Time series plot of PM10 and (b) histogram of PM10.

As in the introduction, there are several variables in air quality, namely PM2.5 and PM10 (particles with an aerodynamic diameter equal to or less than 2.5, also called fine, and every 10 μm), ozone (O_3), nitrogen dioxide (NO_2), carbon monoxide (CO) and sulfur dioxide (SO_2) (Manisalidis et al., 2020). In this study, we use fine particle PM10 variable to demonstrate the neo-normal MSAR MSN-Burr model. The detailed PM10 movement patterns are recorded hourly and its histogram which looks skewed to the right can be seen in Figure 2.

This application compares the standard PM10 level published by WHO with the actual accidents that occurred in Yogyakarta city in 2021. Clean air typically contains PM10 levels of less than 15 micrograms per cubic meter (μg/m^3) (World Health Organization, 2021). The results obtained from the neo-normal MSAR MSN-Burr model are in two states: normal and abnormal.

Results and Discussions

This section shows the results obtained from the neo-normal MSAR MSN-Burr model written in the Stan code available on Github https://github.com/DwilaksanaAbdullahRasyid/MSAR_MSN-Burr. The normal distribution was chosen as the prior over several parameters because it can approximate many other distributions. This makes it a good choice for modeling a wide range of phenomena (Kruschke, 2015). Bayesian methods are very effective when used on small amounts of data. Indeed, Bayesian uses Monte Carlo simulation (Beer et al., 2013). Therefore, if this estimation method is applied to big data, including air quality data, a computer device with large capacity random access memory (RAM) is required. This happened in this study when the neo-normal MSAR MSN-Burr model parameter estimation for the PM10 variable on the Yogyakarta 2021 air quality dataset was implemented on a computer with a RAM capacity of only 12 GB and the sampling scenario was repeated four times. The program is run for 100,000 iterations in each replication, consuming 11 GB of memory, so it is difficult for the parameters to reach convergent conditions. To overcome this problem, a sampling scenario is applied by reducing the domain range of each parameter only in the predicted area with a feasible pdf in the next replication. This scenario provides excellent parameter estimation results, it is found that the slope parameter of the MSN-Burr distribution shows a dominant role in achieving convergence. The membership of each observation after all parameters converge is shown in Figure 3, with regime 1 representing PM10 under normal conditions and regime 2 representing PM10 under abnormal conditions.

In the fourth repetition, the parameters converge, as indicated by the perfect "Rhat" value of all estimated parameters equal to one. "Rhat" as a tool to monitor MCMC convergence was discovered by Gelman and Rubin (1992). This value measures the ratio of the average variance of the samples in each chain to the variance of the pooled samples across the chain. According to Gelman and Rubin, independent Markov chains should be sampled until all values of "Rhat" are equal to one after being initialized with diffuse initial values for the parameters. According to Gelman and Rubin, independent Markov chains should be sampled until all values of "Rhat" are equal to one after being initialized with diffuse initial values for the parameters. Some of these parameters are $p_i, i = 1,2$ for the probability regime shifts, $\beta_{0i}, i = 1,2$ for the intercepts, $\beta_{1i}, i = 1,2$ for the autoregressive constants, $\sigma_i, i = 1,2$ for the variances, and $\lambda_i, i = 1,2$ for the skewness. A summary of the model estimation results can be seen in Table 6 and the estimation results of the MSAR normal model using original EM algorithm as a baseline can be seen in Table 7.

The MSAR parameters presented in Table 7 indicate that the estimate for the autoregressive parameter, β_{12}, is 1. This estimation suggests that β_{12} parameter leads to a random walk, which implies that the model's predictions do not revert to a mean but instead exhibit continuous and irregular drift over time. As a result, regimes are separated irregularly across the

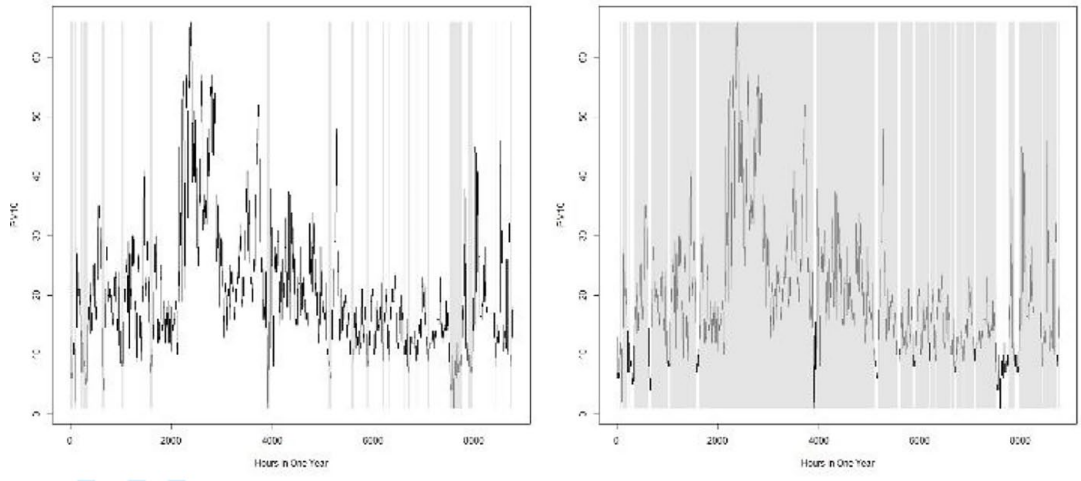


Figure 3. Membership distribution from observations (a) for regime 1 and (b) for regime 2.

Table 6. The Estimated Parameter of Neo-Normal MSAR MSN-Burr with Two Regimes.

PARAMETER	*M*	SE_ MEAN	*SD*	2.5%	25%	50%	75%	97.5%	N_EFF	RHAT
β_{01}	2.49	0.01	0.81	0.95	1.92	2.47	3.03	4.11	12,473	1
β_{02}	0.39	0.00	0.36	0.01	0.12	0.28	0.54	1.34	9,018	1
β_{11}	0.95	0.00	0.04	0.84	0.94	0.97	0.99	1.00	11,985	1
β_{12}	0.94	0.00	0.05	0.81	0.91	0.95	0.98	1.00	10,887	1
ρ_1	0.98	0.00	0.00	0.97	0.98	0.98	0.98	0.99	11,928	1
ρ_2	0.99	0.00	0.00	1.00	1.00	1.00	1.00	1.00	9,441	1
σ_1	5.61	0.00	0.16	5.31	5.51	5.61	5.72	5.92	10,468	1
σ_2	7.19	0.00	0.07	7.05	7.15	7.19	7.24	7.33	11,948	1
λ_1	1.19	0.00	0.01	1.18	1.19	1.20	1.20	1.20	14,419	1
λ_2	17.59	0.00	0.11	17.36	17.52	17.59	17.67	17.80	11,661	1

Table 7. MSAR Normal Estimated Parameter.

PARAMETER	
β_{01}	1.85
β_{02}	0.02
β_{11}	0.94
β_{12}	1
ρ_1	0.74
ρ_2	0.99
σ_1	3.54
σ_2	0.58

model, indicating a high level of uncertainty and unpredictability in regime changes.

To get a deeper understanding of the impact of the random walk on the model, we may look at the regime shifts in the conventional MSAR model, as seen in Figure 4. The MSAR normal regime changes are contrasted to those of the neo-normal MSAR MSN-Burr model. The main difference is the separation of observations within each regime. The separation of the MSAR normal model in Figure 4(a) is highly irregular, showing unpredictability between regimes. In contrast, the proposed model in Figure 4(b) shows the separation of regime and can explain skewness in the data. The separation to regime 1 explaining in-control conditions, whereas regime 2 explaining out of control conditions. This separation demonstrates the capability of the neo-normal MSAR MSN-Burr in

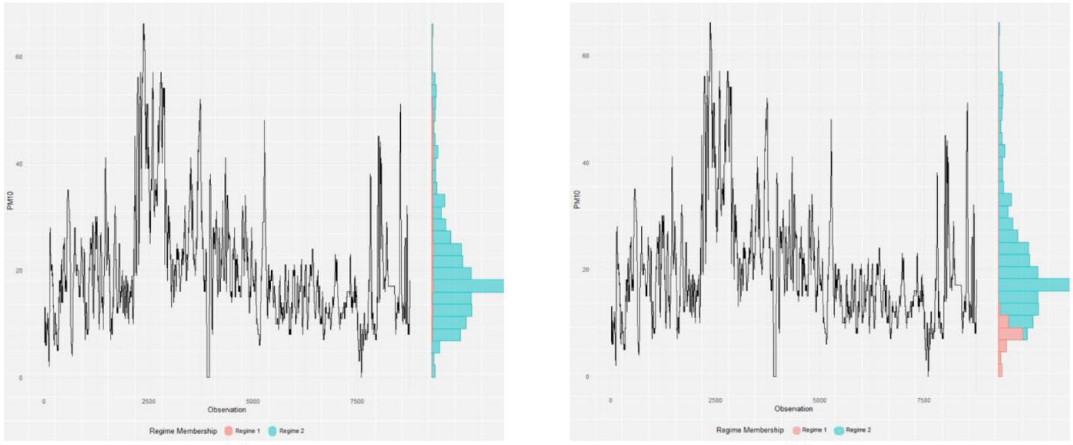


Figure 4. Regime membership (a) MSAR normal (b) neo-normal MSAR MSN-Burr.

distinguishing different regime, leads to reliable and interpretable results.

The performance of the models is compared using the Root Mean Square Error (RMSE) metric. The normal MSAR model has an RMSE of 11.87909, whereas the MSN-Burr MSAR model achieves a significantly lower RMSE of 9.9303. With the resulting RMSE the normal MSAR is 11.87909 and the MSN-Burr MSAR is 9.9303. This significant reduction in RMSE with 16.4049% improvement illustrates the proposed greater accuracy and effectiveness in capturing underlying data patterns when compared to the MSAR normal model.

After estimating using EM-NUTS, calculate the Upper Control Limit (UCL) and Lower Control Limit (LCL) using the Highest Posterior Distribution (HPD). HPD is a method for estimating confidence intervals from asymmetric distributions (Gelman et al., 2014). Chen and Shao (1999) estimated HPD intervals using Monte Carlo procedures. The basic idea for creating a $(1-\alpha)$ HPD interval (also called as a credible interval) is to use the concept of density equilibrium by solving two simultaneous equations. For example, suppose that $f(x)$ is a pdf of random variable X. The values a and b are the lower and upper limits of the HPD interval in the domain $-\infty < x < \infty$ so that $f(a) = f(b)$. Then the values of a and b respectively as the lower limit and upper limit of the credible interval must cover a certain area of the pdf, so that $\int_a^b f(x)dx = 1-\alpha$, for a certain level of significance α. This concept can be stated that the choice of a and b which maintains both the highest density and the area under the density between the chosen boundaries, can be obtained by solving the following two equations simultaneously.

$$f(a) - f(b) = 0 \tag{10}$$

$$P(a < x < b) = 1-\alpha \tag{11}$$

In general, the calculation of ULC and LCL on HPD can be seen in Algorithm 4.

Algorithm 4. Control Limit Estimation Algorithm in HPD.

1. Determine the distribution function and probability density function from the data.
2. Calculate the mode which is then defined as the center line (CL).
3. Solving Equations 10 and 11 simultaneously, solutions a and b are the LCL and UCL of HPD, respectively.

The HPD calculation is applied to the MSN-Burr distribution in each regime. The aim adding the HPD in this study is to know the early warning limits before a regime change occurs. The results of the control limit estimation using Algorithm 2 are the upper and lower limits of each regime. Several significance levels were tested to be applied to Algorithm 2. We use three different levels of significance to see how sensitive the specified control limits are. The significance levels used are 1%, 5%, and 10%, as shown in Table 8.

Figure 5 provides a visualization of the boundaries of each regime using HPD with three different significance levels. As described at the beginning of Application section, regime 1 is a normal PM10 condition in Figure 5(a), and regime 2 is an abnormal PM10 condition in Figure 5(b). In addition to the visualization, Figure 5 also shows the different modes for each regime. The normal condition mode of PM10 in Yogyakarta is 8 μg/m^3 and the abnormal condition mode is 19 μg/m^3.

High levels of air pollution, particularly the PM10 variable in the air quality measurements, can be seen in Figure 5(b), especially under abnormal conditions. High levels of PM10 can cause health problems, as described in Introduction section. Therefore, a large level of significance is required to create highly sensitive control limits. Under normal conditions, however, all of the upper control limit compliance the WHO air quality guideline. This approach aligns better with WHO guidelines and ensures more sensitive detection of deviations. To address abnormalities, adaptive measures such as real-time

Table 8. Several Levels of Significance in Regime 1 and Regime 2.

α (%)	REGIME 1		REGIME 2	
	LOWER	UPPER	LOWER	UPPER
1	1	13	8.0000	56
5	3	12	8.8750	47
10	4	11	8.8750	36

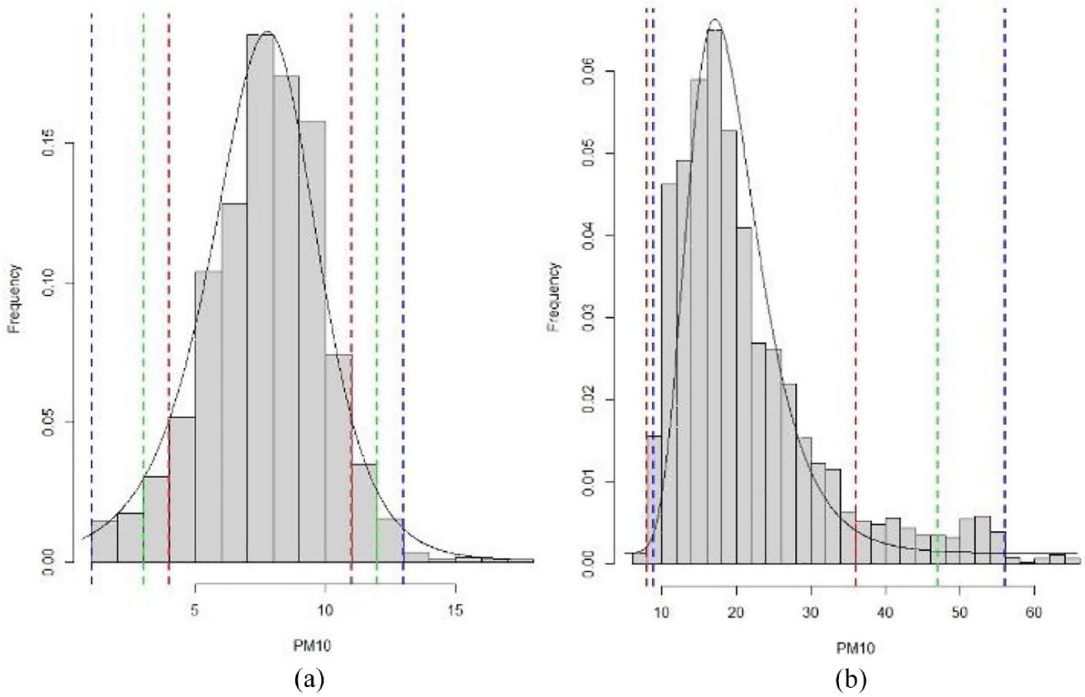


Figure 5. HPD for several significance level, red line for 10%, green line for 5%, and blue line for 1%, which are overlaid with a histogram for each regime (a) regime 1 and (b) regime 2.

monitoring adjustments are essential for maintaining air quality within safe limits.

Limitations

The proposed model is highly effective at capturing data with skewed characteristics. However, when applied to data with high leptokurtic characteristics, it is less representative of all observations. This limitation is evident in the simulation study section, particularly with the Double Exponential distributed generation data scenario. The model's ineffectiveness in this context highlights the need for further refinement when dealing with such datasets.

Conclusions

The MSAR model was created to capture the movement between certain conditions and other conditions. However, with anomalies in real-world conditions that occur successively, the standard MSAR model cannot capture them because the normal residual assumption is violated. Therefore, the proposed model contains residuals with a neo-normal distribution, namely MSN-Burr. With the help of the Stan programming language using EM-NUTS as an estimator, we obtain a new model that is mathematically complex but capable of being developed.

This paper presents the simulation study using normal, double exponential, and Student-t distribution. All three have something in common, which is a symmetrical pattern. The 95% credible interval results, the neo-normal MSAR MSN-Burr model can estimate data generated from the normal, double exponential, and Student-t distributions which can be seen λ value in each regime that is close to one. We then also provide bias and RMSE calculations in each parameter to compare the neo-normal MSAR MSN-Burr model and normal MSAR. As a result, the neo-normal MSAR MSN-Burr model can capture the conditions of symmetric pattern data generation.

This paper also explained the calculation of RMSE from predicted value in simulation study. The double exponential scenario, the RMSE improvement percentage of neo-normal MSAR MSN-Burr is not much different from MSAR normal. This is because the double exponential distribution generates a

high leptokurtic pattern. However, in the other two scenarios, namely the normal and Student-t scenarios, neo-normal MSAR MSN-Burr has better performance than MSAR normal in producing RMSE.

The neo-normal MSAR MSN-Burr model can capture shifts in PM10 conditions, where each condition exhibits an asymmetric pattern. PM10 is a type of particle with an aerodynamic diameter of 10 μm, which can penetrate the lungs and then enter the body through the bloodstream, affecting all major organs. This can cause diseases of both the cardiovascular and respiratory systems. In Yogyakarta, the PM10 in normal conditions is 8 μg/m^3 which is classified as slightly higher than WHO standards.

In addition, we also calculate the upper and lower limits for each condition using HPD. All of the significance level are advised to be implemented in Yogyakarta due to its lower limits more settle to the WHO air quality guideline. The higher significance level is produces control limits that are more sensitive than the lowest significance level. This approach could detect the abnormalities and the significance level which compliance to the guidelines.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The author(s) received financial support for the research, authorship, and/or publication of this article from the Directorate for Research and Community Service (DRPM) - Institut Teknologi Sepuluh Nopember (ITS), under research grant number 1738/PKS/ITS/2023.

ORCID iDs

Dwilaksana Abdullah Rasyid https://orcid.org/0009-0002-1551-0567
Nur Iriawan https://orcid.org/0000-0003-2833-6115

REFERENCES

Adejumo, O. A., Albert, S., & Asemota, O. J. (2021). Markov regime-switching autoregressive model of stock market returns in Nigeria. *Central Bank of Nigeria Journal of Applied Statistics*, 11(2), 65–83. https://doi.org/10.33429/Cjas.11220.3/8

Ailliot, P., & Monbet, V. (2012). Markov-switching autoregressive models for wind time series. *Environmental Modelling & Software*, 30, 92–101. https://doi.org/10.1016/j.envsoft.2011.10.011

Annis, J., Miller, B. J., & Palmeri, T. J. (2017). Bayesian inference with Stan: A tutorial on adding custom distributions. *Behavior Research Methods*, 49(3), 863–886. https://doi.org/10.3758/s13428-016-0746-9

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*,12, 171–178.

Beer, M., Ferson, S., & Kreinovich, V. (2013). Imprecise probabilities in engineering analyses. *Mechanical Systems and Signal Processing*, 37(1–2), 4–29. https://doi.org/10.1016/j.ymssp.2013.01.024

Burr, I. W. (1942). Cumulative frequency functions. *The Annals of Mathematical Statistics*, 13(2), 215–232. https://doi.org/10.1214/aoms/1177731607

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Chen, M.-H., & Shao, Q.-M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *Journal of Computational and Graphical Statistics*, 8(1), 69. https://doi.org/10.2307/1390921

Choung, Y.-J., & Kim, J.-M. (2019). Study of the relationship between urban expansion and PM10 concentration using multi-temporal spatial datasets and the machine learning technique: Case study for Daegu, South Korea. *Applied Sciences*, 9(6), 1098. https://doi.org/10.3390/app9061098

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39, 1–22.

Deschamps, P. J. (2006). A flexible prior distribution for Markov switching autoregressions with Student-t errors. *Journal of Econometrics*, 133(1), 153–190. https://doi.org/10.1016/j.jeconom.2005.03.012

Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, 195(2), 216–222. https://doi.org/10.1016/0370-2693(87)91197-X

Franke, J. (2012). Markov switching time series models. In T. S. Rao, S. S. Rao, & C. R. Rao (Eds.), *Handbook of statistics* (Vol. 30, pp. 99–122). https://doi.org/10.1016/B978-0-444-53858-1.00005-3

Frühwirth-Schnatter, S. (2009). Finite mixture and Markov switching models. *Psychometrika*, 74(3), 559–560. https://doi.org/10.1007/s11336-009-9121-4

Fu, W., Smith, B. R., Brewer, P., & Droms, S. (2023). Markov-switching Bayesian vector autoregression model in mortality forecasting. *Risks*, 11(9), 152. https://doi.org/10.3390/risks11090152

Gao, Y. (2020). A Markov chain model of air quality index: Modelling and simulation. *Journal of Physics: Conference Series*, 1575(1), 012209. https://doi.org/10.1088/1742-6596/1575/1/012209

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. S., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). CRC Press.

Gelman, A., Lee, D., & Guo, J. (2015). Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics*, 40(5), 530–543. https://doi.org/10.3102/1076998615606113

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472. https://doi.org/10.1214/ss/1177011136

Hamilton, J. D. (1989). A new approach to the economic analysis of non-stationary time series and the business cycle. *Econometrica*, 57(2), 357–384.

Hamilton, J. D., & Raj, B. (2002). *Advances in Markov-switching models*. Physica-Verlag.

Hoffman, M. D., & Gelman, A. (2014). The no-U-turn sampler: Adaptively setting path lengths in hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15, 1593–1623.

Iriawan, N. (2000). *computationally intensive approaches to inference in neo-normal linear models*. CUT-Australia.

Johnson, N. M., Hoffmann, A. R., Behlen, J. C., Lau, C., Pendleton, D., Harvey, N., Shore, R., Li, Y., Chen, J., Tian, Y., & Zhang, R. (2021). Air pollution and children's health—A review of adverse effects associated with prenatal exposure from fine to ultrafine particulate matter. *Environmental Health and Preventive Medicine*, 26(1), 72. https://doi.org/10.1186/s12199-021-00995-5

Kaur, S., Morales-Hidalgo, P., Arija, V., & Canals, J. (2023). Prenatal exposure to air pollutants and attentional deficit hyperactivity disorder development in children: A systematic review. *International Journal of Environmental Research and Public Health*, 20(8), 5443. https://doi.org/10.3390/ijerph20085443

Kim, C.-J., & Nelson, C. R. (1999). Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics*, 81(4), 608–616. https://doi.org/10.1162/003465399558472

Kim, C.-J., & Nelson, C. R. (2000). *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. The MIT Press.

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd ed.). Elsevier.

Lhuissier, S. (2019). *Bayesian inference for Markov-switching skewed autoregressive models*. Banque de France Eurosysteme (working papers 726).

Li, Y., Oravecz, Z., Zhou, S., Bodovski, Y., Barnett, I. J., Chi, G., Zhou, Y., Friedman, N. P., Vrieze, S. I., & Chow, S.-M. (2022). Bayesian forecasting with a regime-switching zero-inflated multilevel Poisson regression model: An application to adolescent alcohol use with spatial covariates. *Psychometrika*, 87(2), 376–402. https://doi.org/10.1007/s11336-021-09831-9

Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and health impacts of air pollution: A review. *Frontiers in Public Health*, 8, 14. https://doi.org/10.3389/fpubh.2020.00014

Martinez, W. L., & Martinez, A. R. (2016). *Computational statistics handbook with MATLAB* (3rd ed.). CRC Press; Taylor & Francis Group.

Natarajan, S. K., Shanmurthy, P., Arockiam, D., Balusamy, B., & Selvarajan, S. (2024). Optimized machine learning model for air quality index prediction in major cities in India. *Scientific Reports*, 14(1), 6795. https://doi.org/10.1038/s41598-024-54807-1

Neal, R. (2011). *MCMC using Hamiltonian dynamics*. Chapman and Hall/CRC, https://doi.org/10.1201/b10905-6

Osmundsen, K. K., Kleppe, T. S., & Oglend, A. (2021). MCMC for Markov-switching models—Gibbs sampling vs. marginalized likelihood. *Communications in Statistics - Simulation and Computation*, 50(3), 669–690. https://doi.org/10.1080/03610918.2019.1565580

Ouyang, H., Tang, X., Kumar, R., Zhang, R., Brasseur, G., Churchill, B., Alam, M., Kan, H., Liao, H., Zhu, T., Chan, E. Y. Y., Sokhi, R., Yuan, J., Baklanov, A., Chen, J., & Patdu, M. K. (2022). Toward better and healthier air quality: Implementation of WHO 2021 global air quality guidelines in Asia. *Bulletin of the American Meteorological Society*, 103(7), E1696–E1703. https://doi.org/10.1175/BAMS-D-22-0040.1

Ren, C., & Tong, S. (2008). Health effects of ambient air pollution – recent research development and contemporary methodological challenges. *Environmental Health*, 7(1), 56. https://doi.org/10.1186/1476-069X-7-56

Sims, C. A., Waggoner, D. F., & Zha, T. (2008). Methods for inference in large multiple-equation Markov-switching models. *Journal of Econometrics*, 146(2), 255–274. https://doi.org/10.1016/j.jeconom.2008.08.023

Thangavel, P., Park, D., & Lee, Y.-C. (2022). Recent insights into particulate matter (PM2.5)-mediated toxicity in humans: An overview. *International Journal of Environmental Research and Public Health*, 19(12), 7511. https://doi.org/10.3390/ijerph19127511

Troug, H., & Murray, M. (2021). Crisis determination and financial contagion: An analysis of the Hong Kong and Tokyo stock markets using an MSBVAR approach. *Journal of Economic Studies*, 48(8), 1548–1572. https://doi.org/10.1108/JES-03-2020-0095

Wabersich, D., & Vandekerckhove, J. (2014). Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behavior Research Methods*, 46(1), 15–28. https://doi.org/10.3758/s13428-013-0369-3

Walther, B. A., & Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. *Ecography*, 28(6), 815–829. https://doi.org/10.1111/j.2005.0906-7590.04112.x

Wetzels, R., Lee, M. D., & Wagenmakers, E.-J. (2010). Bayesian inference using WBDev: A tutorial for social scientists. *Behavior Research Methods*, 42(3), 884–897. https://doi.org/10.3758/BRM.42.3.884

World Health Organization. (2021). *WHO global air quality guidelines*.

Wright, N., Newell, K., Chan, K. H., Gilbert, S., Hacker, A., Lu, Y., Guo, Y., Pei, P., Yu, C., Lv, J., Chen, J., Li, L., Kurmi, O., Chen, Z., Lam, K. B. H., & Kartsonaki, C. (2023). Long-term ambient air pollution exposure and cardio-respiratory disease in China: Findings from a prospective cohort study. *Environmental Health*, 22(1), 30. https://doi.org/10.1186/s12940-023-00978-9

Yang, L., Li, C., & Tang, X. (2020). The impact of PM2.5 on the host defense of respiratory system. *Frontiers in Cell and Developmental Biology*, 8, 91. https://doi.org/10.3389/fcell.2020.00091

Zakaria, N. N., Othman, M., Sokkalingam, R., Daud, H., Abdullah, L., & Abdul Kadir, E. (2019). Markov chain model development for forecasting air pollution index of Miri, Sarawak. *Sustainability*, 11(19), 5190. https://doi.org/10.3390/su11195190

Zhang, J., Zhu, F., & Chen, H. (2023). Two-threshold-variable integer-valued autoregressive model. *Mathematics*, 11(16), 3586. https://doi.org/10.3390/math11163586