

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

*Forecasting* atau peramalan merupakan kegiatan guna memperkirakan peristiwa di masa depan dengan mengandalkan analisis data historis. Peramalan juga dapat dikerjakan dengan menerapkan metode deret waktu. Analisis deret waktu bertujuan untuk menganalisis hubungan antara periode waktu yang berbeda. Tujuannya adalah untuk menemukan cara atau model yang tepat guna menggambarkan pola hubungan waktu yang terstruktur antara beberapa data atau peristiwa, sehingga memungkinkan evaluasi atau prediksi terhadap satu atau lebih data di masa depan (Deviana et al., 2021). Analisis data deret waktu mencakup dua jenis model analisis yaitu model deret waktu *univariate* (satu variabel observasi) yang hanya memperhitungkan unsur waktu tanpa mempertimbangkan aspek wilayah, sedangkan untuk memodelkan masalah yang melibatkan data deret waktu *multivariate* (dua atau lebih observasi) seperti pada penelitian ini digunakan model *Vector Autoregressive Integrated Moving Average* (VARIMA). Model VARIMA merupakan sistem persamaan dengan pendugaan suatu variabel dalam periode tertentu bergantung pergerakan variabel tersebut serta variabel-variabel lain yang terkait dalam periode sebelumnya (Safira et al., 2024).

Model yang memperluas model VARIMA yang tidak hanya mempertimbangkan unsur waktu tetapi juga melibatkan keterkaitan dengan wilayah, dikenal dengan istilah model *space time*. Model *space time* merupakan salah satu model yang mengelompokkan unsur ketergantungan waktu maupun wilayah dalam data deret waktu *multivariate* (Wutsqa et al., 2012). Model *space time* pertama kali diperkenalkan oleh Pfeifer dan Deutsch pada tahun 1980 yang dikenal sebagai model *Space Time Autoregressive Moving Average* (STARMA) (Pfeifer & Deutch, 1980). Perbedaan utamanya terdapat pada penambahan matriks pembobot wilayah yang mencirikan adanya pengaruh interaksi wilayah atau spasial.

Pada model STARMA hanya dapat diaplikasikan pada wilayah yang berkarakteristik homogen dengan asumsi bahwa parameter sama di seluruh wilayah. Kondisi aktual menunjukkan suatu wilayah dengan wilayah lain memiliki karakteristik heterogen yang menjadikan penggunaan model STARMA dianggap kurang fleksibel, sehingga (Borovkova et al., 2008) dalam (Prameshti & Saputro, 2023) mengembangkan model STARMA menjadi *Generalized Space Time Autoregressive Moving Average* (GSTARMA). Dalam model GSTARMA diharuskan terpenuhinya asumsi kestasioneran data serta harus mempunyai orde *autoregressive* dan *moving average* dengan orde spasial pada umumnya yang ditentukan terbatas pada *orde 1*. Orde spasial 1 (satu) ditetapkan karena orde spasial lebih dari 1 (satu) sulit untuk diinterpretasikan dalam model (Sjahid et al.,

2020). Orde spasial satu merepresentasikan bahwa kejadian pada suatu lokasi dipengaruhi oleh kejadian pada lokasi yang paling dekat dengan lokasi tersebut. Oleh karena itu, pada penelitian ini orde spasial yang digunakan hanya sampai pada orde satu.

Untuk memenuhi asumsi kestasioneran data perlu dilakukan *differencing*, sehingga model GSTARMA dikembangkan menjadi *Generalized Space Time Autoregressive Integrated Moving Average* (GSTARIMA). Model GSTARIMA sesungguhnya sama dengan model STARIMA hanya saja nilai parameter pada lag spasial yang sama antar lokasi diperbolehkan berlainan. Model GSTAR adalah model GSTARIMA tanpa adanya *differencing* dan nilai parameter *moving average* bernilai nol, sedangkan model GSTMA juga merupakan model GSTARIMA tanpa diferensiasi dan nilai parameter *autoregressive* bernilai nol dan untuk efek spasial dinyatakan dengan bobot wilayah yang menunjukkan besarnya hubungan antar wilayah (Salsabila et al., 2024). GSTARIMA dapat digunakan untuk data yang memuat unsur *integrated* sesudah data stasioner dan memuat unsur dengan elemen *autoregressive* dan *moving average*.

Seiring dengan perkembangannya, banyak peneliti yang mengembangkan konsep pada data *time series* (deret waktu) menjadi data *space time* (ruang waktu). Pada data *space time*, kejadian pada suatu lokasi dianggap dipengaruhi oleh faktor lokasi dan waktu. Menurut hukum Tobler, segala sesuatu saling berhubungan. Semakin dekat jarak suatu lokasi maka keberpengaruhannya semakin besar (Salsabila et al., 2024). Keterkaitan kedekatan secara spasial dapat diinterpretasikan dalam matriks bobot spasial. Matriks pembobot invers jarak adalah salah satu bentuk umum dari matriks ini, di mana kekuatan hubungan antar lokasi berbanding terbalik dengan jaraknya; semakin dekat dua lokasi, semakin besar bobot pengaruhnya. Pendekatan ini mencerminkan asumsi bahwa lokasi yang berdekatan memiliki pengaruh yang lebih besar satu sama lain dibandingkan lokasi yang berjauhan. Di sisi lain, korelasi silang (*cross-correlation*) digunakan untuk mengukur keterkaitan temporal antar lokasi dalam seri waktu yang berbeda. Korelasi silang membantu dalam identifikasi adanya keterlambatan (lag) dan pola hubungan antar wilayah, yang kemudian digunakan untuk membentuk struktur model GSTARIMA yang optimal. Kombinasi antara matriks pembobot invers jarak dan analisis korelasi silang memungkinkan model menangkap dinamika ruang-waktu dengan lebih akurat (Sjahid et al., 2020).

Informasi tentang kondisi cuaca dan iklim sangat penting bagi berbagai pihak. Data ini digunakan untuk berbagai keperluan, seperti sebagai bahan pertimbangan dalam perencanaan pembangunan yang mencakup sektor transportasi, pertanian, ekonomi, serta sektor kesehatan dan lainnya. Adapun hasil penelitian terkait peramalan iklim yang dilakukan oleh (Anwar, 2017) menyatakan bahwa data unsur cuaca pada umumnya bersifat tidak stasioner. Model *space time* dapat diterapkan pada semua data yang menunjukkan keterkaitan antar waktu dan wilayah. Suhu udara adalah salah satu contoh fenomena iklim yang memperlihatkan hubungan tersebut. Suhu udara merefleksikan jumlah energi radiasi matahari dalam bentuk panas yang dirasakan, yang berkontribusi

pada pemanasan atmosfer. Selain itu, suhu udara juga menggambarkan kondisi panas atau dinginnya udara, yang umum sebagai temperatur (Prasetyo et al., 2021).

Perubahan iklim dapat berdampak pada berbagai sektor secara global termasuk pertanian, kesehatan, perikanan, pariwisata dan lainnya (Uni Eropa, 2023) dalam (Salsabila et al., 2022). Sebagai salah satu komponen utama dalam sektor pertanian, suhu udara berperan vital dalam memajukan pertumbuhan ekonomi dan produksi tanaman pangan di Indonesia terutama di wilayah Nusa Tenggara Barat, dikenal sebagai negara agraris karena mayoritas penduduknya bermata pencaharian di sektor pertanian dan bercocok tanam (BPS Nusa Tenggara Barat, 2022). Suhu udara merupakan variabel penting dalam peramalan cuaca, membantu memprediksi hujan, badai, dan fenomena ekstrem lainnya. Membantu mitigasi bencana seperti gelombang panas, kebakaran hutan, dan cuaca ekstrem lainnya.

Berdasarkan data dari Stasiun Klimatologi BMKG Nusa Tenggara Barat tercatat bahwa suhu udara rata-rata harian bulan Juli 2021 di Nusa Tenggara Barat berkisar antara 25,0 – 27,0 derajat celcius dan mengalami peningkatan dari awal November 2023 berada di kisaran 20 kemudian mengalami peningkatan hingga 36 derajat celcius yang menyebabkan setiap tahun mengalami kekeringan ekstrem dimana setiap kenaikan suhu udara 2 celcius akan menurunkan produksi pangan pertanian yang berimplikasi terhadap potensi gagal panen.

Penelitian yang membahas tentang VARIMA telah dilakukan oleh (Karim & Ahmed, 2023) dalam penelitiannya yang membahas tentang meramalkan harga mata uang kripto menggunakan optimasi parameter baru model varima. Hasil penelitian menunjukkan bahwa model terbaik adalah VARIMA (2,1,2), yang berisi sembilan parameter signifikan dan tiga parameter tidak signifikan, meskipun tidak ada korelasi silang dalam residual berdasarkan hasil matriks kovarian residual. Dengan kata lain, residual bersifat acak, tidak ada autokorelasi dalam data, dan model telah memenuhi syarat *white noise*. Akibatnya, model VARIMA (2, 1, 2) dapat digunakan untuk meramalkan covid-19 dan harga minyak.

Penelitian selanjutnya oleh (Andriani et al., 2024) yaitu membahas tentang penerapan model VARIMA untuk peramalan tingkat hunian kamar hotel di provinsi Sumatera Selatan. Dalam penelitian ini dilakukan analisis data time series multivariat untuk meramalkan nilai impor dan ekspor non migas di Jawa Barat dengan menggunakan metode VARIMA. Berdasarkan data impor dan ekspor non migas di Jawa Barat pada Januari 2013-April 2022 diperoleh model peramalan yang didapatkan dan sesuai adalah VARIMA (2,1,1). Dari model dapat diketahui bahwa nilai ekspor dipengaruhi oleh nilai impor dan dari hasil peramalan impor dan ekspor non migas tertinggi akan terjadi pada bulan Juni 2022 dan terendah akan terjadi pada bulan Mei 2022.

Kemudian penelitian yang dilakukan oleh (I. D. Putri & Oktaviarina, 2024) penerapan VARIMA pada prediksi indeks standar pencemaran udara di Kabupaten Gresik. Hasil dari penelitian ini diperoleh bahwa model terbaik berdasarkan nilai RMSE dan MAD terkecil

model VARIMA (2,1,1), dimana konsentrasi  $PM_{2.5}$  dan  $O_3$  pada saat ket dipengaruhi oleh ketiga variabel pada saat ke (t-1), (t-2), dan (t-3) serta dipengaruhi oleh nilai residual ketiga variabel pada saat (t-1). Sedangkan variabel  $NO_2$  pada saat ke-t hanya dipengaruhi oleh variabel  $PM_{2.5}$  dan  $O_3$  pada saat ke (t-1), (t-2), dan (t-3) serta dipengaruhi oleh nilai residual ketiga variabel pada saat (t-1).

Beberapa penelitian yang berkaitan dengan model STARIMA telah dilakukan oleh (Awwad et al., 2021) tentang meramalkan kasus Covid-19 di wilayah Makkah Arab Saudi dengan model STARIMA. Hasil penelitian menunjukkan bahwa model STARIMA lebih andal dalam memperkirakan epidemi COVID-19 di masa mendatang daripada model ARIMA. Penelitian yang dilakukan oleh (Rathod et al., 2018) membahas tentang pemodelan STARMA untuk pemodelan dan peramalan data deret waktu spasio-temporal. Hasilnya mengungkap bahwa metode model STARMA yang diusulkan mengungguli ARIMA univariat dan model STARMA orde spasial pertama untuk pemodelan dan peramalan, baik untuk data simulasi maupun data aktual.

Penelitian STARIMA juga dilakukan oleh (Joebaedi et al., 2018) meneliti tentang model stasioner STAR orde pertama pada data perminyakan. Hasil penelitian menyimpulkan bahwa model STAR (1;1) bersifat stasioner dan dapat digunakan untuk meramalkan pengamatan di masa mendatang pada suatu lokasi berdasarkan satu waktu sebelumnya dari lokasi itu sendiri dan lingkungan spasialnya. Hal ini menunjukkan bahwa model STAR (1;1) pada data produktivitas minyak bumi di Balongan, Indramayu, Jawa Barat memenuhi syarat stasioner.

Penelitian yang berkaitan dengan model GSTARIMA yaitu (A. Salsabila et al., 2024) dalam penelitiannya tentang pengembangan tatanan model GSTARIMA (1,1,1) untuk peramalan data iklim, yang dimana dalam penelitiannya membandingkan model GSTARIMA (1,1,1) dengan model GSTARIMA (3,1,1) sebagai pengembangan dari model GSTARIMA (1,1,1) dan diterapkan pada data iklim. Berdasarkan hasil penelitiannya, model GSTARIMA (3,1,1) memiliki nilai MAPE sebesar 9% untuk data *out-sample* dan 11% untuk data *in-sample*. Sebaliknya model GSTARIMA (1,1,1) memiliki nilai MAPE sebesar 11% untuk data *out-sample* dan 12% untuk data *in-sample*. Sehingga model GSTARIMA (3,1,1) memberikan hasil peramalan yang lebih akurat.

Kemudian penelitian yang dilakukan oleh (Aryani et al., 2020) yang meneliti tentang penerapan model GSTAR pada data nilai tukar petani di tiga provinsi Pulau Sumatera. Berdasarkan analisis data, model GSTAR terbaik diperoleh dengan menggunakan pembobotan terbaik dengan model tersebut adalah GSTAR  $(1_1) - I(1)$  dengan menggunakan normalisasi korelasi silang karena asumsi *white noise* normal dan multivariat terpenuhi dengan nilai RMSE sebesar 1.097775. Model GSTAR terbaik menjelaskan bahwa nilai tukar petani Sumbar hanya waktu sebelumnya, nilai tukar petani Bengkulu adalah waktu sebelumnya dan merupakan nilai tukar petani Sumbar dan Jambi, sedangkan untuk nilai tukar petani Jambi adalah nilai tukar petani Bengkulu dan Sumatera Barat serta dipengaruhi oleh masa-masa sebelumnya.

Penelitian selanjutnya dikerjakan oleh (Yulianti et al., 2018) tentang pemodelan harga beras di pulau Sumatera dengan menggunakan model *Generalized Space Time Arima*. Hasil pemodelan menunjukkan bahwa model terbaik adalah model GSTARIMA (1,1,0) dengan matriks pembobot *queen contiguity* yang mempunyai nilai *mean absolute percentage error* (MAPE) yang terkecil yaitu sebesar 1.17817%

Penelitian sebelumnya yang melakukan perbandingan dengan GSTARIMA diantaranya penelitian yang dilakukan oleh (Safira et al., 2024) tentang dampak spasial terhadap inflasi pulau Jawa menggunakan prediksi ARIMA dan GSTARIMA menunjukkan bahwa hasil kedua model akan dibandingkan untuk mengetahui model mana yang memiliki akurasi lebih baik. Berdasarkan nilai RMSE, model GSTARIMA memiliki nilai RMSE rata-rata paling kecil, yaitu sebesar 0,113 dibandingkan dengan model ARIMA yang memiliki nilai RMSE rata-rata 0,319 sehingga dapat disimpulkan bahwa penambahan faktor spasial dapat meningkatkan akurasi prediksi inflasi di Pulau Jawa.

Selanjutnya penelitian yang dilakukan oleh (Zebua et al., 2022) perbandingan pemodelan dan prediksi nilai ekspor di Jawa Barat, Banten, dan Jawa Tengah menggunakan STIMA dan GSTIMA (1, 1, 1). Hasil penelitian menghasilkan bahwa model GSTIMA (1,1,1) menghasilkan prediksi yang lebih baik dibandingkan STIMA (1,1,1) karena memiliki nilai MAPE yang lebih kecil, dimana nilai MAPE masing-masing sebesar 14,23% untuk STIMA dan 11,38% untuk GSTIMA.

Model VARIMA adalah model yang cocok untuk menganalisis data *time series* dengan lebih dari satu variabel seperti data suhu udara serta digunakan untuk mengetahui keterkaitan antara beberapa variabel pada suatu waktu dengan variabel periode sebelumnya. Sedangkan model GSTARIMA merupakan perluasan dari model VARIMA yang tidak hanya mempertimbangkan unsur waktu tetapi juga memiliki penambahan matriks pembobot wilayah yang mencirikan adanya pengaruh interaksi spasial/wilayah dalam data deret waktu *multivariate*. Suhu udara adalah salah satu fenomena iklim yang memperlihatkan hubungan keterkaitan antara unsur waktu dan wilayah.

Interaksi spasial suhu udara antara wilayah memiliki dampak yang luas dan saling berhubungan, mempengaruhi hampir setiap aspek kehidupan manusia dan lingkungan. Ini mencakup bagaimana perbedaan suhu di berbagai wilayah mempengaruhi kondisi cuaca, ekosistem, dan aktivitas manusia di masing-masing wilayah. Di sisi lain, pengaruh interaksi waktu terhadap suhu udara mencakup perubahan suhu udara dari waktu ke waktu, baik dalam jangka pendek (sehari-hari), bulanan, maupun tahunan, terhadap berbagai aspek lingkungan dan kehidupan manusia. Selain itu adanya heterogenitas antar wilayah pada data suhu udara, karena suhu yang dirasakan di setiap ketiga wilayah dapat bervariasi yang melatar belakangi pentingnya melakukan peramalan pada data. Misalnya, perbedaan suhu yang signifikan antara musim hujan dan musim kemarau, dimana selama musim kemarau suhu di Sumbawa NTB bisa sangat tinggi, sedangkan di daerah pegunungan Lombok NTB cenderung lebih sejuk.

Dalam penelitian sebelumnya yang membandingkan model VARIMA dan GSTARIMA Affandy (2014) dan Setiawan (2015). Berdasarkan penelitian yang dilakukan oleh Affandy (2014) tentang hasil perbandingan model VAR dengan GSTAR menunjukkan bahwa peramalan dengan model VAR lebih akurat. Sedangkan berdasarkan penelitian Setiawan (2015) menunjukkan bahwa model GSTAR merupakan model terbaik.

Data suhu udara di Nusa Tenggara Barat mencapai 31 hingga 37 derajat Celsius. Hasil pengamatan BMKG Stasiun Meteorologi Zaenudin Abdul Madjid (ZAM) Lombok, suhu di wilayah Lombok Tengah mencapai 31-33 derajat Celsius. Kemudian hasil pengamatan stasiun Meteorologi Sultan Muhammad Kaharudin Sumbawa suhu udara mencapai 33 hingga 37 derajat Celsius. Sedangkan Stasiun Meteorologi Sultan Muhammad Salahuddin Bima suhu udara mencapai 33 hingga 36 derajat Celsius (BPS Nusa Tenggara Barat, 2023). Sehingga menurut pengamatan suhu udara di NTB menunjukkan bahwa provinsi NTB menunjukkan rentang suhu yang seragam di atas 30°C. Rentang suhu yang konsisten dan tinggi ini mempertegas bahwa kondisi suhu saat ini adalah ekstrem dan meskipun terdapat variasi suhu antara wilayah semua pengamatan menunjukkan suhu di atas ambang batas normal, mengonfirmasikan bahwa suhu di NTB secara keseluruhan berada dalam kondisi ekstrem dan kering yang berdampak pada kesehatan, lingkungan, dan ekosistem secara keseluruhan.

Oleh karena itu berdasarkan uraian sebelumnya, maka pada penelitian ini akan mengkaji estimasi parameter menggunakan *Ordinary Least Squares* (OLS). Metode OLS menjadi pilihan utama dalam estimasi parameter model GSTARIMA dan VARIMA karena model ini dapat direpresentasikan dalam bentuk persamaan regresi linear (Alawiyah et al., 2021). Selain itu, OLS memiliki sifat *Best Linear Unbiased Estimator* (BLUE), yang memastikan estimasi parameter bersifat efisien dan tidak bias. OLS juga meminimalkan jumlah kuadrat error, sehingga menghasilkan parameter yang mendekati nilai sebenarnya (Ruchjana dkk., 2012), serta pemodelan suhu udara rata-rata bulanan di Provinsi Nusa Tenggara Barat dengan pendekatan VARIMA dan GSTARIMA dan juga belum terdapat penelitian sebelumnya yang melakukan secara *multivariate* dalam mengidentifikasi keterkaitan wilayah antara suhu udara di Nusa Tenggara Barat. Salah satu analisis statistika yang dapat digunakan untuk melakukan peramalan yang tidak hanya menggunakan unsur waktu akan tetapi juga melihat unsur wilayah atau disebut *space time* yaitu GSTARIMA. Adapun judul yang di angkat oleh peneliti yaitu “**Pemodelan Vector Autoregressive Integrated Moving Average dan Generalized Space Time Autoregressive Integrated Moving Average Berbasis Matriks Invers Jarak dan Normalisasi Korelasi Silang pada Kasus Suhu Udara di Nusa Tenggara Barat**”.

## 1.2 Rumusan Masalah

Mengacu pada paparan latar belakang yang telah diuraikan, masalah yang menjadi fokus penelitian ini dapat dirumuskan sebagai berikut:

1. Bagaimana mengestimasi parameter VARIMA dan GSTARIMA?

2. Bagaimana memodelkan suhu udara rata-rata bulanan di Provinsi Nusa Tenggara Barat dengan pendekatan VARIMA dan GSTARIMA Berbasis Pembobot Invers Jarak dan Normalisasi Korelasi Silang?
3. Bagaimana hasil prediksi model VARIMA dan GSTARIMA Berbasis Pembobot Invers Jarak dan Normalisasi Korelasi Silang untuk peramalan suhu udara rata-rata di Provinsi Nusa Tenggara Barat?

### **1.3 Batasan Penelitian**

Berdasarkan rumusan masalah, maka ruang lingkup dalam penelitian ini dibatasi pada beberapa hal sebagai berikut:

1. Data suhu udara rata-rata bulanan di Nusa Tenggara Barat dengan rentang waktu Februari 2017 hingga November 2024
2. Peramalan dilakukan di Provinsi Nusa Tenggara Barat dengan 3 stasiun BMKG yaitu stasiun Meteorologi Sultan Muhammad Kaharuddin, stasiun Meteorologi Zainuddin Abdul Madjid, stasiun Meteorologi Sultan Muhammad Salahuddin.
3. Dalam penelitian ini, digunakan model VARIMA dan GSTARIMA dengan bobot invers jarak dan normalisasi korelasi silang. Orde spasial dibatasi pada orde 1 untuk memudahkan interpretasi dan perbandingan antara kedua model, mengingat orde yang lebih tinggi dapat menjadi sulit untuk diinterpretasikan (Wutsqa et al., 2012).

### **1.4 Tujuan Penelitian**

Berikut adalah tujuan yang ingin dicapai dalam penelitian ini:

1. Untuk mengestimasi parameter VARIMA dan GSTARIMA
2. Untuk memodelkan suhu udara rata-rata bulanan di Provinsi Nusa Tenggara Barat dengan pendekatan VARIMA dan GSTARIMA Berbasis Pembobot Invers Jarak dan Normalisasi Korelasi Silang.
3. Untuk menganalisis hasil prediksi metode VARIMA dan GSTARIMA Berbasis Pembobot Invers Jarak dan Normalisasi Korelasi Silang untuk peramalan suhu udara rata-rata di Provinsi Nusa Tenggara Barat.

### **1.5 Manfaat Penelitian**

1. Dapat memberikan gambaran kepada masyarakat maupun pemerintah terkait peramalan suhu udara rata-rata bulanan di provinsi Nusa Tenggara Barat.
2. Sebagai bahan acuan untuk pengembangan penelitian selanjutnya maupun pada kasus atau wilayah lain.
3. Menambah wawasan bagi penulis terkait analisis runtun waktu multivariat menggunakan metode VARIMA dan GSTARIMA.

## 1.6 Analisis Deret Waktu

Deret waktu mengacu pada bagian penting dalam pengambilan keputusan karena efektivitas keputusan seringkali bergantung pada beberapa faktor yang belum dapat diketahui saat pengambilan keputusan. Selain itu, deret waktu adalah teknik yang diterapkan untuk memperkirakan suatu nilai pada titik waktu dimasa depan dengan mempertimbangkan data historis maupun sekarang. Tujuan dari analisis deret waktu meliputi meramalkan (*forecasting*) kondisi yang akan datang, mengetahui hubungan antar variabel, dan kepentingan kontrol, yaitu untuk mengetahui apakah proses terkendali atau tidak (Aswi & Sukarna, 2017).

Suatu rangkaian observasi mempunyai model deret waktu apabila terpenuhi dua syarat berikut ini (Aswi & Sukarna, 2006):

1. Interval waktu antar indeks waktu  $t$  dapat dinyatakan dalam satuan waktu yang seragam (identik).
2. Terdapat ketergantungan antara pengamatan  $Z_t$  dengan  $Z_{t+k}$  yang dipisahkan oleh jarak waktu berupa kelipatan  $\Delta_t$  sebanyak  $k$  kali (diistilahkan sebagai lag  $k$ ).

## 1.7 Autokorelasi Spasial

Untuk melihat ada atau tidaknya ketergantungan spasial antar stasiun maka dilakukan uji autokorelasi spasial pada data suhu udara di Nusa Tenggara Barat. Autokorelasi spasial adalah taksiran dari korelasi antar nilai amatan yang berkaitan dengan lokasi pada variabel yang sama (Yasin et al., 2020). Autokorelasi spasial adalah korelasi antara variabel dengan dirinya sendiri berdasarkan ruang atau bisa dikatakan kemiripan objek dalam suatu ruang, baik jarak, waktu maupun wilayah. Besaran autokorelasi spasial dapat digunakan untuk mengidentifikasi hubungan spasial (Anselin, 1988). Autokorelasi spasial dapat di uji pada suatu wilayah secara global maupun secara lokal. Pengujian autokorelasi melibatkan suatu bobot yang disebut matriks pembobot spasial (*spatial weight matrix*) yang menggambarkan kedekatan hubungan antar lokasi. Matriks pembobot spasial disebut juga sebagai matriks yang menggambarkan interaksi antar lokasi.

Menurut (Lee & Wong, 2001) autokorelasi spasial secara global umumnya menggunakan statistik *Moran's I* yang diusulkan oleh ahli statistik Patrick A.P. Moran pada tahun 1950. Menurut Kosfeld dalam (Wuryandari et al., 2014), *Moran's I* dapat digunakan untuk mendeteksi permulaan dari keacakan spasial. Perhitungan autokorelasi spasial dengan metode Indeks Moran yaitu:

$$I = \frac{n}{W} \times \frac{\sum_{i=1}^n \sum_{j=1}^n W_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_{i=1}^n (X_i - \bar{X})^2} \quad (1)$$

dengan:

$I$  : Indeks Moran

$n$  : Banyaknya wilayah pengamatan



$X_i$  dan  $X_j$  : Nilai pada wilayah  $i$  dan wilayah  $j$   
 $\bar{X}$  : rata-rata dari jumlah variabel  
 $W_{ij}$  : Pembobot spasial yang diterapkan antar wilayah ke- $i$  dan ke- $j$

Uji hipotesis untuk Indeks Moran adalah sebagai berikut:

$H_0: I = 0$  (tidak ada autokorelasi antar wilayah)

$H_1: I \neq 0$  (ada autokorelasi antar wilayah)

Tingkat signifikansi :  $\alpha = 5\%$

Statistik Uji:

$$Z(I) = \frac{I - E(I)}{\sqrt{\text{var}(I)}} \approx N(0,1) \quad (2)$$

dengan:

$$E(I) = -\frac{1}{n-1}$$

$$\text{Var}(I) = \frac{n^2 \cdot S_1 - n \cdot S_2 + 3 \cdot S_0^2}{(n^2 - 1)S_0^2} - [E(I)]^2$$

$$S_0 = \sum_{i=1}^n \sum_{j=1}^n W_{ij}$$

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (W_{ij} + W_{ji})^2$$

$$S_2 = \sum_{i=1}^n \left( \sum_{j=1}^n W_{ij} + \sum_{j=1}^n W_{ji} \right)^2$$

Kriteria uji:

Tolak  $H_0$  pada taraf signifikansi  $\alpha$  jika  $Z(I) > Z_{1-\alpha}$  dengan  $Z_{1-\alpha}$  adalah  $(1 - \alpha)$  kuantil dari distribusi normal standar atau tolak  $H_0$  jika  $p\text{-value} < \alpha$ .

Interpretasi dari output nilai *Moran's I* adalah sebagai berikut (Wuryandari et al., 2014):

- Nilai *Moran's I* semakin mendekati nilai +1 menyatakan bahwa nilai autokorelasi spasial semakin positif kuat. Ini berarti daerah-daerah yang berdekatan memiliki nilai efek spasial yang mirip dan cenderung berkelompok.
- Nilai *Moran's I* mendekati 0 menyatakan bahwa autokorelasi spasial acak atau random (tidak ada korelasi spasial).
- Nilai *Moran's I* semakin mendekati nilai -1 menyatakan bahwa autokorelasi spasial semakin negatif kuat dimana polanya menyebar, artinya daerah-daerah yang berdekatan memiliki nilai efek spasial yang berbeda.

## 1.8 Uji Keheterogenan Wilayah dengan Indeks Gini

Indeks Gini adalah suatu koefisien yang menunjukkan tingkat ketidakmerataan suatu pengamatan untuk menguji heterogenan wilayah. Indeks Gini diperkenalkan oleh Corrado Gini (1884-1965). Nilai indeks gini berkisar antara 0 hingga 1, suatu wilayah dikatakan memiliki pemerataan sempurna (homogen) jika nilai indeks gini mendekati 0

dan sebaliknya (Aryani et al., 2020). Salah satu syarat dalam melakukan analisis model GSTARMA yaitu karakteristik wilayah penelitian bersifat heterogen sehingga data yang digunakan dapat berupa data time series maupun data spasial (wilayah). Oleh karena itu, indeks Gini dapat digunakan untuk memeriksa keheterogenitas wilayah. Berikut langkah-langkah menguji heterogenitas wilayah menggunakan indeks Gini.

1. Hipotesis

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2$  (Wilayah homogen atau pemerataan sempurna)

$H_1: \sigma_i^2 \neq \sigma^2; i = 1,2,3, \dots n$  (Wilayah heterogen atau pemerataan tidak sempurna)

2. Tingkat Signifikansi

$$\alpha = 5\%$$

3. Statistik Uji:

$$G = 1 + \frac{1}{n} - \frac{2}{n^2 \bar{y}_i} \sum_{i=1}^N y_i \tag{3}$$

dimana:

$G$  : Indeks Gini

$n$  : Jumlah data

$y_i$  : Nilai variabel yang diamati berdasarkan indeks sampel  $i$

$\bar{y}_i$  : Rata-rata seluruh variabel yang diamati

$N$  : Jumlah data pada wilayah ke-  $i$

$i$  : Indeks sampel

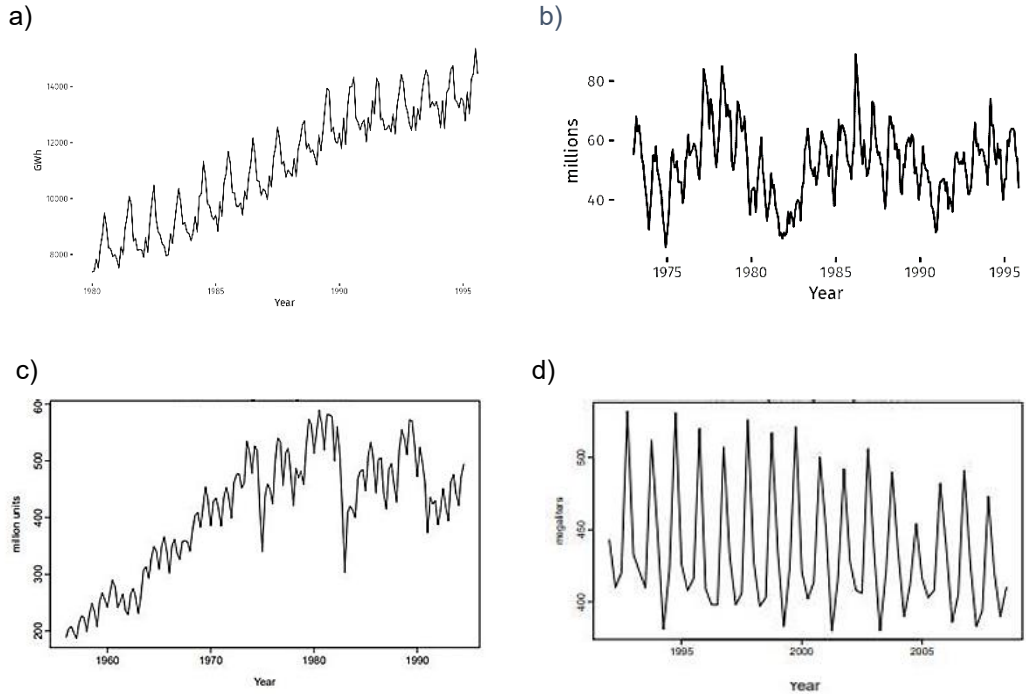
4. Daerah Kritis

Tolak  $H_0$  jika nilai indeks  $G \geq 1$ , disimpulkan bahwa wilayah diasumsikan heterogen.

5. Keputusan dan Kesimpulan

### 1.9 Stasioner Data

Suatu data dianggap layak untuk dimodelkan secara deret waktu apabila ketentuan kestasioner yang digunakan terpenuhi. Dengan kata lain, stasioneritas berarti data tidak terjadi pertumbuhan atau penurunan seiring waktu. Dalam arti tertentu, data berada dalam keseimbangan statistik. Singkatnya, suatu data dikatakan stasioner dalam rata-rata jika data tersebut secara stabil atau tetap berfluktuasi di sekitar garis yang sejajar sumbu waktu atau dinamakan berada di sekitar suatu nilai rata-ratanya, sedangkan suatu data dikatakan stasioner dalam variansi jika struktur data dari waktu ke waktu memiliki fluktuasi data yang tetap atau konstan dan tidak berubah ubah (Aswi & Sukarna, 2017). Berikut adalah beberapa contoh grafik deret waktu yang stasioner dan non stasioer dalam rata-rata dan variansi.



**Gambar 1.** Data Stasioner dan Tidak Stasioner

Sumber gambar: (Hyndman & Athanasopoulos, 2018).

### 1.9.1 Stasioner dalam Variansi

Menurut (Tutupoho et al., 2024) data dianggap stasioner dalam variansi jika tidak menunjukkan fluktuasi pada data dan tetap konstan sepanjang waktu. Bentuk visual representasi data *time series* seringkali cukup meyakinkan bahwa data tersebut adalah stasioner atau tidak stasioner. Namun, untuk mengidentifikasi stasioner terhadap variansi suatu data dapat dideteksi menggunakan plot Box-Cox. Apabila nilai parameter transformasi ( $\lambda$ ) sama atau mendekati satu, maka dapat dikatakan bahwa data telah stasioner dalam variansi. Sebaliknya, ketika nilai parameter transformasi ( $\lambda$ ) tidak sama dengan satu, yang berarti mengindikasikan data tidak stasioner dalam variansi, maka perlu dilakukan Box-Cox untuk menstasionerkan variansi. Secara umum, parameter transformasi ( $\lambda$ ) yang diperkenalkan oleh Box & Cox (1964) dikutip dalam (Yunita, 2020) untuk mencapai stasioneritas dalam variansi dapat dilihat pada persamaan (4).

$$Z_t^{(\lambda)} = \frac{Z_t^{(\lambda)} - 1}{\lambda} \quad (4)$$

Dimana  $Z_t^{(\lambda)}$  merupakan data yang mengalami transformasi dan  $\lambda$  adalah parameter transformasi. Nilai  $\lambda$  dan bentuk transformasi Box-Cox yang sesuai dapat dilihat pada Tabel 1.

**Tabel 1. Transformasi Box-Cox**

Nilai pendugaan $\lambda$	-1,0	-0,5	0,0	0,5	1,0
Transformasi	$\frac{1}{Z_t}$	$\frac{1}{\sqrt{Z_t}}$	$\ln(Z_t)$	$\sqrt{Z_t}$	$Z_t$

Sumber: (Wei, 2006) dalam (Yunita, 2020).

Parameter  $\lambda$  didapatkan dengan cara estimasi menggunakan metode *maximum likelihood*. Adapun langkah-langkah untuk menentukan  $\lambda$  sebagai berikut:

1. Menentukan kisaran  $\lambda$ , biasanya diambil dari kisaran  $[-1,1]$  dan diperluas jika diperlukan.
2. Untuk  $\lambda$  yang terpilih hitung:

$$L_{maks}(\lambda) = -\frac{n}{2} \ln \hat{\sigma}^2(\lambda) + (\lambda - 1) \sum_{i=1}^n \ln Z_i \quad (5)$$

dengan:

$n$  : Banyaknya pengamatan

$\hat{\sigma}^2$  :  $\frac{1}{n} \sum_{i=1}^n (Z_i(\lambda) - \bar{Z}(\lambda))^2$

$\bar{Z}(\lambda)$  :  $\frac{1}{n} \sum_{i=1}^n (Z_i(\lambda)) = \frac{1}{n} \sum_{i=1}^n \left( \frac{Z_i^\lambda - 1}{\lambda} \right)$

3. Setelah memperoleh nilai  $\lambda$  dalam kisaran yang ditetapkan, dipilih  $\lambda$  yang menghasilkan  $L_{maks}(\lambda)$  yang terbesar. Nilai ini adalah estimasi kemungkinan maksimum  $\hat{\lambda}$  (maximum likelihood estimator) untuk parameter  $\lambda$ .

Nilai  $\lambda$  yang diperoleh dibulatkan terhadap nilai dalam barisan  $\dots, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \dots$

Misalnya didapatkan  $\hat{\lambda} = 1,21$  maka digunakan  $\lambda = 1$ .

### 1.9.2 Stasioner dalam Rata-rata

Pada data yang stasioner terhadap rata-rata, nilai autokorelasi pada data stasioner terhadap rata-rata biasanya tidak signifikan setelah *lag time* kedua atau ketiga. Sebaliknya, pada data yang tidak stasioner memiliki nilai autokorelasi yang signifikan setelah *lag time* ketiga atau lebih. Pemeriksaan kestasioneran terhadap rata-rata dapat dilakukan secara grafis dengan diagram ACF. Data dikatakan stasioner terhadap rata-rata jika nilai autokorelasi pada tiga lag pertama tidak melampaui batas  $\pm \frac{2}{\sqrt{n}}$ . Apabila data tidak memenuhi kriteria ini, proses *differencing* perlu diterapkan hingga data menjadi stasioner terhadap rata-rata. Berikut adalah bentuk *differencing* pertama ( $d = 1$ ) (S. M. Putri & Arliani, 2022):

$$\Delta Z_t = Z_t - Z_{t-1} \quad (6)$$

Keterangan:

$\Delta Z_t$  : Hasil *differencing* waktu ke-t

$Z_t$  : Hasil pengamatan waktu ke-r

$Z_{t-1}$  : Hasil pengamatan waktu ke ( $t - 1$ )

Selain secara grafis, kestasioneran terhadap rata-rata dapat diketahui dengan uji *Augmented Dickey Fuller* (ADF). Misalkan terdapat model AR (1) sebagai berikut:

$$Z_t = \Phi Z_{t-1} - e_t \quad (7)$$

Jika persamaan (7) dikurangi dengan  $Z_{t-1}$  pada masing masing ruas, maka persamaan tersebut akan menjadi:

$$\begin{aligned} Z_t - Z_{t-1} &= \Phi Z_{t-1} - Z_{t-1} - e_t \\ \Delta Z_t &= (\Phi - 1)Z_{t-1} - e_t \\ \Delta Z_t &= \Phi^* Z_{t-1} - e_t \end{aligned} \quad (8)$$

Berdasarkan persamaan (8) dapat ditarik hipotesis:

$H_0 : \Phi^* \geq 0$  (data tidak stasioner terhadap rata-rata)

$H_1 : \Phi^* < 0$  (data stasioner terhadap rata-rata)

Statistik uji:

$$\tau = \frac{\hat{\phi}^*}{Se(\hat{\phi}^*)} \quad (9)$$

Keterangan:

$\hat{\phi}^*$  : Nilai duga parameter *Autoregressive* (AR)

$Se(\hat{\phi}^*)$  : Standar error dari  $\hat{\phi}^*$

Ketika menggunakan nilai *p-value*, kriteria yang diterapkan adalah apabila *p-value*  $\geq \alpha$  berarti hipotesis nol diterima yang mengindikasikan bahwa data belum stasioner. Sedangkan, apabila *p-value*  $< \alpha$ , hipotesis nol ditolak yang menunjukkan bahwa data sudah stasioner.

## 1.10 Identifikasi Model VARIMA

Untuk mengidentifikasi model VARIMA, dapat menggunakan *Matrix Autocorrelation Function* (MACF) dan *Matrix Partial Autocorrelation Function* (MPACF). MACF berfungsi untuk menentukan orde Vektor MA( $q$ ) sementara MPACF membantu untuk menetapkan orde Vektor AR( $p$ ) (Anggraeni et al., 2018).

Menurut (Anggraeni et al., 2018) fungsi matriks korelasi sampel untuk vektor deret waktu dengan  $n$  pengamatan  $Z_1, Z_2, \dots$ , dan  $Z_n$  dapat dihitung menggunakan rumus:

$$\hat{\rho}(k) = [\hat{\rho}_{ij}(k)]$$

dengan  $\hat{\rho}_{ij}(k)$  adalah korelasi silang sampel dari komponen deret ke- $i$  dan ke- $j$  pada lag ke- $k$  yaitu:

$$\hat{\rho}_{ij}(k) = \frac{\sum_{t=1}^{n-k} (Z_{i,t} - \bar{Z}_i)(Z_{j,t+k} - \bar{Z}_j)}{\sqrt{\sum_{t=1}^n (Z_{i,t} - \bar{Z}_i)^2 \sum_{t=1}^n (Z_{j,t} - \bar{Z}_j)^2}} \quad (10)$$

Sementara itu, rumus untuk menghitung fungsi matriks korelasi parsial adalah sebagai berikut

$$\Phi_{kk} = \frac{\text{Cov}[(Z_t - \hat{Z}_t), (Z_{t+k} - \hat{Z}_{t+k})]}{\sqrt{\text{var}(Z_t - \hat{Z}_t)}\sqrt{\text{var}(Z_{t+k} - \hat{Z}_{t+k})}} \quad (11)$$

dengan  $\hat{Z}_t$  dan  $\hat{Z}_{t+k}$  adalah estimasi regresi linier dengan *Mean Squared Error* (MSE) terkecil. MPACF digunakan dalam identifikasi model VAR. Jika MPACF terpotong pada lag ke- $p$  maka model yang bersesuaian adalah VAR( $p$ ).

Matriks dan grafik menjadi semakin rumit ketika dimensi vektor meningkat, yang membuat identifikasi model deret waktu menjadi lebih sulit. Untuk mengatasi hal ini, Tiao dan Box mengembangkan metode dengan menandai matriks korelasi sampel ke  $(i,j)$  menggunakan simbol (+), (-), dan (.), yaitu:

1. Simbol (+) menyatakan bahwa  $\hat{\rho}_{ij}(k) > 2$  kali standar error dan mengindikasikan adanya hubungan korelasi positif.
2. Simbol (-) menyatakan bahwa  $\hat{\rho}_{ij}(k) < -2$  kali standar error dan menunjukkan hubungan korelasi negatif.
3. Simbol (.) menyatakan bahwa  $\hat{\rho}_{ij}(k)$  berada diantara  $\pm 2$  kali standar error dan menunjukkan tidak adanya korelasi.

### 1.11 Model Vector Autoregressive Integrated Moving Average

Model VARIMA merupakan model *time series* multivariat pengembangan dari model ARIMA dengan variabel dependen secara bersamaan. Model VARIMA berfungsi sebagai sistem persamaan dinamis yang memungkinkan perkiraan nilai suatu variabel pada periode tertentu dengan mempertimbangkan pergerakan variabel tersebut serta variabel lainnya dalam sistem dari periode-periode sebelumnya.

Model VARIMA ( $p, d, q$ ) dengan  $p$  dan  $q$  masing-masing adalah ordo *autoregressive* dan *moving average* didefinisikan sebagai berikut (Anggraeni et al., 2018):

$$\Phi_p(B)D(B)Z_t = \Theta_q(B)e_t \quad (12)$$

dimana:

$$\begin{aligned} \Phi_p(B) &= \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p \\ \Theta_q(B) &= \Theta_1 B - \Theta_2 B^2 - \dots - \Theta_q B^q \end{aligned}$$

dengan:

$Z_t$  : Vektor pengamatan dengan  $Z_t = [Z_{1,t}, Z_{2,t}, \dots, Z_{n,t}]$  berukuran  $(n \times 1)$

$\Phi_p$  : Matriks parameter vektor *autoregressive* ordo ke- $p$  berukuran  $(n \times n)$

$\Theta_q$  : Matriks parameter vektor *moving average* ordo ke- $q$  berukuran  $(n \times n)$

$B$  : Operator shift mundur

$D(B)$  : Operator *differencing*

$e_t$  : Vektor acak white noise dengan  $e_t \sim NM(0, \Sigma)$  berukuran  $(n \times 1)$

Ketika  $d = 0$  dan  $q = 0$  maka menjadi model VAR dituliskan sebagai berikut:

$$\Phi_p(B)Z_t = e_t \quad (13)$$

Metode *Ordinary Least Square* (estimasi kuadrat terkecil) digunakan untuk mengestimasi parameter VARIMA. Memilih model berdasarkan lag signifikan pada MACF dan MPACF akurat namun tidak selalu praktis. Sebagai alternatif, kriteria nilai AIC terkecil (*Akaike Information Criterion*) diterapkan sebagai dasar untuk menentukan panjang *lag* yang tepat. Persamaannya dapat dituliskan sebagai berikut (Nasution & Cipta, 2023):

$$AIC = m \log \left( \frac{SSE}{m} \right) \quad (14)$$

dengan:

$m$  : Banyaknya wilayah pengamatan

SSE : Jumlah kuadrat galat

## 1.12 Matriks Pembobot Spasial

Matriks pembobot spasial  $\mathbf{W} = (W_{ij})$  mencerminkan suatu keterkaitan spasial, dengan  $W_{ij}$  merupakan hubungan spasial antara lokasi  $i$  dengan lokasi  $j$  (Borovkova et al., 2008), (Setiawan & Prastuti, 2016). Matriks pembobot  $\mathbf{W} = (W_{ij})$  merupakan matriks persegi berukuran  $N \times N$  dengan elemen diagonal utamanya 0 yaitu:

$$\mathbf{W} = (W_{ij}) = \begin{bmatrix} 0 & W_{12} & W_{13} & W_{14} & W_{15} & \cdots & W_{1j} \\ W_{21} & 0 & W_{23} & W_{24} & W_{25} & \cdots & W_{2j} \\ W_{31} & W_{32} & 0 & W_{34} & W_{35} & \cdots & W_{3j} \\ W_{41} & W_{42} & W_{43} & 0 & W_{45} & \cdots & W_{4j} \\ W_{51} & W_{52} & W_{53} & W_{54} & 0 & \cdots & W_{5j} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ W_{i1} & W_{i2} & W_{i3} & W_{i4} & W_{i5} & \cdots & W_{ij} \end{bmatrix}$$

Pada penelitian bobot yang dipilih adalah dua pembobotan spasial. Model GSTARIMA dengan dua metode pembobotan spasial yang digunakan diharapkan mampu menjelaskan dan membandingkan model yang terbaik yang dapat digunakan untuk pemodelan suhu udara. Pada GSTARMA terdapat beberapa bobot spasial yang telah dikembangkan antara lain:

### 1.12.1 Bobot Wilayah *Invers Jarak*

Beberapa peneliti melakukan perhitungan matriks invers jarak secara berbeda. Diantaranya (Hadi et al., 2018), (Utami et al., 2021) menghitung jarak berdasarkan jarak sebenarnya. Perhitungan matriks invers jarak menggunakan jarak sebenarnya dapat dilakukan, namun seringkali acuan pusat lokasi satu dengan lainnya masih menjadi pertimbangan. Selain itu, jika jarak sebenarnya yang dimaksudkan adalah jarak nyata di lapangan, maka akan terjadi banyak pertimbangan seperti jalan yang tidak lurus, jalan yang menanjak, jalan yang dipisahkan oleh laut, dan lain sebagainya (Utami et al., 2024). Oleh karena itu, pada penelitian ini perhitungan matriks jarak didasarkan pada rumus *Euclidian distance*. Dalam konteks pengukuran antara dua titik atau objek dalam ruang Euclidean, jarak Euclidean merupakan panjang garis lurus yang menghubungkan dua titik

tersebut. Misalnya, jika dua titik merepresentasikan lokasi geografis dua kota, maka jarak Euclidean antara kota-kota tersebut akan memberikan ukuran garis lurus terpendek antara keduanya.

Penghitungan nilai bobot invers jarak dilakukan berdasarkan jarak aktual antara wilayah. Perhitungan bobot invers jarak melibatkan jarak koordinat lintang dan bujur antara pusat-pusat wilayah yang diamati. Misalnya, jika  $u$  dan  $v$  menyatakan koordinat lintang dan bujur wilayah dan  $d_{ij}$  menyatakan jarak wilayah ke- $i$  terhadap wilayah ke- $j$ , maka diperoleh persamaan (15).

$$d_{ij} = \sqrt{(u_i - u_j)^2 + (v_i - v_j)^2} \quad (15)$$

Menurut (Utami et al., 2024) bahwa setelah diperoleh nilai  $d_{ij}$ , langkah selanjutnya yaitu menghitung bobot invers jarak untuk masing-masing wilayah. Pembobotan wilayah invers jarak dapat dilihat pada persamaan (16).

$$W_{ij} = \begin{cases} \frac{1}{d_{ij}} & j \neq i \\ \frac{1}{\sum_{j=1}^n \frac{1}{d_{ij}}} & j = i. \end{cases} \quad (16)$$

Keterangan:

$W_{ij}$  : Bobot invers jarak dari wilayah  $i$  ke  $j$

$d_{ij}$  : Jarak sebenarnya dari wilayah  $i$  ke  $j$

### 1.12.2 Bobot Normalisasi Korelasi Silang

Penggunaan matriks korelasi silang pada pemodelan GSTAR banyak dilakukan oleh beberapa peneliti, diantaranya (Utami et al., 2024). Matriks korelasi silang merupakan matriks yang menyajikan koefisien korelasi antara dua variabel pada lag waktu tertentu. Nilai estimasi dari korelasi silang untuk variabel di lokasi ke- $i$  dan ke- $j$  dapat dihitung dengan persamaan (17) (Zewdie et al., 2018).

$$r_{ij}(k) = \frac{\sum_{t=k+1}^T [y_{i,t} - \bar{y}_i][y_{j,t-k} - \bar{y}_j]}{\sqrt{(\sum_{t=1}^T [y_{i,t} - \bar{y}_i]^2) (\sum_{t=1}^T [y_{j,t} - \bar{y}_j]^2)}} \quad (17)$$

dengan  $\bar{y}_i$  dan  $\bar{y}_j$  masing-masing menyatakan rata-rata pengamatan pada lokasi ke- $i$  dan lokasi ke- $j$ . Perhitungan bobot korelasi silang dirumuskan dengan (Fina et al., 2016):

$$W_{ij} = \frac{r_{ij}(k)}{\sum_{k \neq i} |r_{ik}(k)|} \quad (18)$$

dengan  $\sum_{j \neq i} |w_{ij}| = 1$ .



### 1.13 Model *Generalized Space Time Autoregressive Integrated Moving Average*

Model *space time* merupakan sebuah pendekatan deret waktu multivariat yang mempertimbangkan unsur ketergantungan waktu dan wilayah. Keterbatasan model STARIMA terhadap pengasumsian karakteristik wilayah yang seragam (homogen) menjadikan model ini tidak fleksibel. Akibatnya, dilakukan pengembangan menjadi model GSTARIMA. Sebagai generalisasi dari model STARIMA, model GSTARIMA tidak mengharuskan parameter yang sama di seluruh wilayah (Triangga et al., 2024).

Perbedaan mendasar antara model STARIMA dan GSTARIMA terletak pada parameternya, pada model STARMA  $\Phi_{kl}$  dan  $\Theta_{kl}$  yang merupakan konstanta, sedangkan untuk model GSTARIMA berupa matriks  $\Phi_{kl}$  dan  $\Theta_{kl}$  GSTARMA  $(p, d, q)_1$  yang menggabungkan komponen waktu *autoregressive* dengan orde  $p$  dan *moving average* dengan orde  $q$ . Model GSTARIMA dapat ditemukan dalam persamaan (19) Wei (2019) seperti yang tercantum dalam (Ilmi et al., 2023):

$$\nabla \mathbf{Z}_t = \sum_{k=1}^p \sum_{l=0}^{\lambda_p} \Phi_{kl} \mathbf{W}^{(l)} (\mathbf{1} - \mathbf{B})^d \nabla \mathbf{Z}_{t-k} - \sum_{k=1}^q \sum_{l=0}^{m_q} \Theta_{kl} \mathbf{W}^{(l)} \mathbf{e}_{t-k} + \mathbf{e}_t \quad (19)$$

Keterangan:

- $\nabla \mathbf{Z}_t$  : Vektor amatan pada waktu ke- $t$  wilayah ke- $i$  yang berukuran  $N \times 1$
- $p$  : Orde *autoregressive* (AR)
- $q$  : Orde *moving average* (MA)
- $\lambda_p$  : Orde spasial pada keadaan *autoregressive* ke- $p$
- $m_q$  : Orde spasial pada keadaan *moving average* ke- $q$
- $(\mathbf{1} - \mathbf{B})^d$  : *Differencing* orde  $d$
- $\mathbf{B}$  : Operator mundur (*blackshift* operator)
- $N$  : Banyaknya wilayah yang diteliti
- $\Phi_{kl}$  : Matriks diagonal parameter *autoregressive* pada waktu ke- $k$  dan *lag* spasial ke- $l$  berukuran  $N \times N$  dengan elemen diagonal  $(\phi_{kl}^1, \dots, \phi_{kl}^N)$
- $\Theta_{kl}$  : Matriks diagonal parameter *moving average* pada *lag* waktu ke- $k$  dan *lag* spasial ke- $l$  berukuran  $N \times N$  dengan elemen diagonal  $(\theta_{kl}^1, \dots, \theta_{kl}^N)$
- $\mathbf{W}^{(l)}$  : Matriks pembobot wilayah dengan ukuran  $(N \times N)$  pada *lag* spasial  $l, l = 0, 1, 2, \dots, N$ .
- $\mathbf{e}_t$  : Vektor acak *white noise* berukuran  $N \times 1$  yang diasumsikan berdistribusi normal dengan rata-rata nol dan variansi konstan.

Matriks pembobot untuk *lag* spasial nol  $\mathbf{W}^{(0)}$  membentuk matriks identitas berukuran  $N \times N$  ( $\mathbf{W}^{(0)} = \mathbf{I}_N$ ) karena pada *lag* spasial nol tidak mempunyai hubungan ketetangaan. Model GSTARMA yang tidak mengandung komponen *autoregressive*, dapat digunakan model GSTMA ( $q$ ) untuk pemodelannya, sebagaimana diperlihatkan dalam persamaan (20).

$$Z_t = e_{(t)} - \sum_{k=1}^q \sum_{l=0}^{m_q} \Theta_{kl} \mathbf{W}^{(l)} e_{t-k} \quad (20)$$

Di sisi lain, model GSTARMA yang tidak mengandung elemen *moving average* dapat dimodelkan menggunakan model GSTAR ( $p$ ). Model GSTAR ( $1_1$ ) dijelaskan dalam persamaan (21).

$$Z_t = \Phi_{10} Z_{t-1} + \Phi_{11} \mathbf{W}^{(1)} Z_{t-1} + e_t \quad (21)$$

Pada persamaan (21), apabila dibentuk dalam notasi matriks, maka diperoleh :

$$\begin{bmatrix} Z_1(t) \\ Z_2(t) \\ \vdots \\ Z_N(t) \end{bmatrix} = \begin{bmatrix} \Phi_{10}^1 & 0 & \cdots & 0 \\ 0 & \Phi_{10}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_{10}^N \end{bmatrix} \begin{bmatrix} Z_1(t-1) \\ Z_2(t-1) \\ \vdots \\ Z_N(t-1) \end{bmatrix} + \begin{bmatrix} \Phi_{11}^1 & 0 & \cdots & 0 \\ 0 & \Phi_{11}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Phi_{11}^N \end{bmatrix} \begin{bmatrix} Z_1(t-1) \\ Z_2(t-1) \\ \vdots \\ Z_N(t-1) \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} \quad (22)$$

Misalkan terdapat 3 wilayah, maka :

$$\begin{bmatrix} Z_1(t) \\ Z_2(t) \\ Z_3(t) \end{bmatrix} = \begin{bmatrix} \Phi_{10}^1 & 0 & 0 \\ 0 & \Phi_{10}^2 & 0 \\ 0 & 0 & \Phi_{10}^3 \end{bmatrix} \begin{bmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{bmatrix} + \begin{bmatrix} \Phi_{11}^1 & 0 & 0 \\ 0 & \Phi_{11}^2 & 0 \\ 0 & 0 & \Phi_{11}^3 \end{bmatrix} \begin{bmatrix} 0 & W_{12} & W_{13} \\ W_{21} & 0 & W_{23} \\ W_{31} & W_{32} & 0 \end{bmatrix} \begin{bmatrix} Z_1(t-1) \\ Z_2(t-1) \\ Z_3(t-1) \end{bmatrix} + \begin{bmatrix} e_1(t) \\ e_2(t) \\ e_3(t) \end{bmatrix} \quad (23)$$

Jika dijabarkan menjadi:

$$Z_1(t) = \Phi_{10}^1 Z_1(t-1) + \Phi_{11}^1 w_{12} Z_2(t-1) + \Phi_{11}^1 w_{13} Z_3(t-1) + e_1(t)$$

$$Z_2(t) = \Phi_{10}^2 Z_2(t-1) + \Phi_{11}^2 w_{21} Z_1(t-1) + \Phi_{11}^2 w_{23} Z_3(t-1) + e_2(t)$$

$$Z_3(t) = \Phi_{10}^3 Z_3(t-1) + \Phi_{11}^3 w_{31} Z_1(t-1) + \Phi_{11}^3 w_{32} Z_2(t-1) + e_3(t)$$

Estimasi parameter pada model GSTARIMA dilakukan dengan *metode Ordinary Least Square (OLS)* atau metode kuadrat terkecil. Metode OLS dapat digunakan dalam pendugaan parameter *vector autoregressive dan vector moving average* (Pasaribu et al., 2021; Wutsqa et al., 2012). Sebagai ilustrasi, model yang digunakan adalah GSTAR ( $1_1$ ). Estimasi parameter model GSTAR menggunakan metode OLS yaitu dengan meminimumkan jumlah kuadrat error. Model GSTAR ( $1_1$ ) dinyatakan sebagai berikut:

$$Y(t) = \Phi_{10} Y(t-1) + \Phi_{11} \mathbf{W}^{(1)} Y(t-1) + e(t)$$

dengan  $\Phi_{10}$  merupakan parameter *autoregressive* yang mengukur ketergantungan waktu,  $\Phi_{11}$  merupakan parameter regresi spasial dan  $\mathbf{W}$  merujuk pada matriks pembobot spasial.

Jika data deret waktu  $\{Z_t: t = 0, 1, 2, \dots, T; i = 1, 2, \dots, n\}$  dengan model umum GSTAR ( $p_1$ ), model lokasi ke- dapat dilihat pada persamaan (24).

$$Z_{i(t)} = X_{i(t)} + \beta_{i(t)} + e_{i(t)} \quad (24)$$

dengan,

$$Z_{i(t)} = \begin{bmatrix} Z_{i(t)} \\ \vdots \\ Z_{i(T)} \end{bmatrix}; X_{i(t)} = \begin{bmatrix} X_{i(0)} & V_{i(0)} \\ \vdots & \vdots \\ X_{i(T-1)} & V_{i(T-1)} \end{bmatrix}; \beta_{i(t)} = \begin{bmatrix} \Phi_{i0}^1 \\ \Phi_{i1}^1 \end{bmatrix}; e_{i(t)} = \begin{bmatrix} e_{i(1)} \\ \vdots \\ e_{i(T)} \end{bmatrix}; V_{i(t)} = \sum_{j \neq i} w_{ij} Z_{j(t)}$$

Keterangan :

$Z_{i(t)}$  : Pengamatan ke- $t$  untuk wilayah ke- $i$

$X_{i(t)}$  : Matriks variabel independen ke-  $t$  untuk wilayah ke- $i$

$\beta_{i(t)}$  : Vektor pengamatan ke- $t$  untuk wilayah ke- $i$

$e_{i(t)}$  : Residual ke- $t$  untuk wilayah ke- $i$

$V_{i(t)}$  : Jumlah hasil kali antara pembobot dengan pengamatan ke- $t$  untuk wilayah ke- $i$

$W_{ij}$  : Pembobot wilayah ke- $i$  dan ke- $j$

## 1.14 Pemeriksaan Diagnostik Model

Pengujian diagnostik model bertujuan untuk proses evaluasi dalam menilai dan membuktikan seberapa baik model layak digunakan untuk peramalan. Uji ini diterapkan pada residual model untuk mengevaluasi apakah residual memenuhi kriteria *white noise* dan distribusi normal *multivariate*.

### 1.14.1 Uji *White Noise*

Asumsi dasar bahwa residual bersifat *white noise* artinya tidak terdapat korelasi antar residual dengan mean sama dengan nol dan varian konstan. Uji untuk mengecek independensi residual (*white noise*) dapat dilakukan menggunakan statistik uji Ljung-Box (Aswi & Sukarna, 2017; Handayani et al., 2019).

Hipotesis yang digunakan sebagai berikut:

$H_0$  :  $\rho_1 = \rho_2 = \dots = \rho_K = 0$  (Residual bersifat *white noise* (tidak berkorelasi))

$H_1$  : Minimal ada satu  $\rho_j \neq 0, j = 0, 1, 2, \dots, K$  (Residual tidak bersifat *white noise* (berkorelasi))

Statistik Uji :

$$Q = n(n+2) \sum_{k=1}^K \frac{\rho_k^2}{n-k} \quad (25)$$

dengan:

$Q$  : Statistik uji *Ljung Box*

$n$  : Jumlah residual

$k$  : Banyak *lag* yang di uji 1, 2, ...,  $K$

$\rho_k^2$  : Nilai koefisien residual pada lag ke- $k$

Kriteria pengujian yaitu tolak  $H_0$  jika  $|Q| \geq X^2_{(\alpha; df=K-m)}$ ,  $K$  berarti pada lag-  $K$  dan  $m$  adalah jumlah parameter yang ditaksir dalam model atau juga menggunakan  $p$ -value, yakni menolak  $H_0$  jika  $p$ -value  $< \alpha$ .

### 1.14.2 Uji Normal *Multivariate*

Asumsi kedua yang harus dipenuhi adalah residual berdistribusi normal multivariat. Pemeriksaan distribusi normal dapat dilakukan dengan uji jarak Mahalanobis (Ilmi et al., 2023). Tujuan dari uji hipotesis ini adalah untuk mengevaluasi apakah data memenuhi asumsi normalitas.

Hipotesis yang digunakan yaitu:

$H_0$  : residual berdistribusi normal *multivariate*

$H_1$  : residual tidak berdistribusi normal *multivariate*

Jarak Mahalanobis dinyatakan dalam bentuk matriks dan vektor, dan perhitungannya dapat dilakukan dengan merujuk pada persamaan 26 (Ilmi et al., 2023).

$$D^2 = (x_i - me)C^{-1}(x_i - me)^T \quad (26)$$

dengan:

$x_i$  : Variabel setiap data ke- $i$

$me$  : Nilai pusat dari kelompok variabel

$C^{-1}$  : Invers matriks kovarian

### 1.15 Pemilihan Model Terbaik

Pemilihan model yang paling sesuai didasarkan pada akurasi peramalan menggunakan data *out sample* (Setiawan & Prastuti, 2016). Kriteria utama dalam memilih model terbaik adalah dengan *Root Mean Square Error* (RMSE), di mana model dengan nilai RMSE terendah (semakin mendekati nol) dianggap sebagai model yang paling akurat (Ilmi et al., 2023). Persamaan RMSE dirumuskan pada persamaan (27) sebagai berikut:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Z_{(t)} - \hat{Z}_{(t)})^2} \quad (27)$$

dengan:

$n$  : banyaknya data ramalan yang dilakukan

$Z_{(t)}$  : data sebenarnya pada waktu ke- $t$

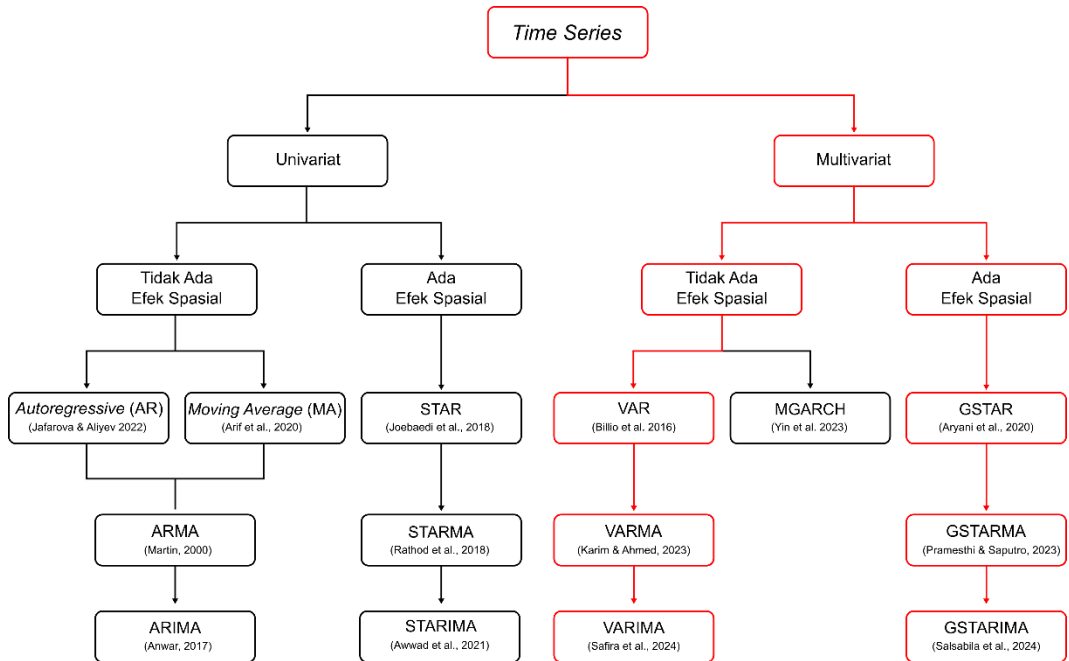
$\hat{Z}_{(t)}$  : data hasil ramalan pada waktu ke- $t$ .

### 1.16 Suhu Udara

Suhu menunjukkan jumlah energi radiasi matahari dalam bentuk panas yang dirasakan (*sensible heat*) dan berperan dalam pemanasan atmosfer. Suhu udara di kenal sebagai temperatur, mencerminkan kondisi panas atau dinginnya udara. Suhu udara harian rata-rata dihitung berdasarkan rata-rata suhu pada beberapa kali pengamatan

dalam setiap periode 24 jam (sehari semalam). Frekuensi pengamatan dapat dilakukan hingga 8 kali dalam sehari, setiap 3 jam sekali dan dimulai dari tengah malam (Prasetyo et al., 2021). Suhu udara maksimum dan minimum diukur dengan menggunakan thermometer maksimum dan thermometer minimum. Variasi suhu di kepulauan Indonesia dipengaruhi oleh ketinggian tempat, suhu udara cenderung lebih rendah di daerah yang lebih tinggi dan sebaliknya (Falgenti & Kahfi, 2023).

### 1.17 Kerangka Konseptual



**Gambar 2.** Kerangka Konseptual

## **BAB II**

### **METODE PENELITIAN**

#### **2.1 Jenis Penelitian**

Jenis penelitian yang diterapkan adalah penelitian kuantitatif. Penelitian kuantitatif merupakan metode di mana peneliti menggunakan data sebagai dasar untuk menyimpulkan pengetahuan atau fakta, serta menganalisis informasi untuk mendapatkan wawasan yang lebih mendalam mengenai topik yang sedang diteliti.

#### **2.2 Sumber Data**

Data sekunder yang digunakan dalam penelitian ini diperoleh dari situs Web Badan Meteorologi, Klimatologi, dan Geofisika. Data yang di gunakan meliputi suhu rata-rata bulanan dari periode Januari 2017 hingga November 2024 untuk Provinsi Nusa Tenggara Barat. Data tersebut diambil dari tiga stasiun, yaitu Stasiun Meteorologi Sultan Muhammad Kaharuddin, Stasiun Zainuddin Abdul Madjid, Stasiun Meteorologi Sultan Muhammad Salahuddin dengan total jumlah data mencapai sebanyak 95.

#### **2.3 Definisi Operasional Variabel**

Penelitian ini mengkaji variabel yang meliputi data harian suhu udara rata-rata ( $^{\circ}\text{C}$ ) dari tiga stasiun, yaitu Stasiun Meteorologi Sultan Muhammad Kaharuddin, Stasiun Zainuddin Abdul Madjid, Stasiun Meteorologi Sultan Muhammad Salahuddin, yang berada di Provinsi Nusa Tenggara Barat.

#### **2.4 Teknik Analisis Data**

Proses analisis yang digunakan dalam penelitian ini terdiri dari tahapan berikut:

1. Mengumpulkan data sekunder mengenai suhu udara rata-rata bulanan dari situs Web Badan Meteorologi, Klimatologi, dan Geofisika.
2. Membagi data menjadi dua kelompok, yaitu:
  - a. Data in sample: Januari 2017 – April 2023 (76 pengamatan).
  - b. Data out sample: Mei 2023 – November 2024 (19 pengamatan).  
Data in sample digunakan untuk memodelkan data VARIMA dan GSTARIMA, sedangkan data out sample berfungsi untuk menguji keakuratan dan validasi model.

3. Mendeskripsikan data suhu udara pada ketiga stasiun BMKG yang terdapat di Provinsi NTB dengan menggunakan statistika deskriptif.
4. Untuk memenuhi syarat dalam penerapan model VARIMA dan GSTARIMA, uji autokorelasi spasial untuk GSTARIMA dan uji heterogenitas wilayah untuk VARIMA dan GSTARIMA diperlukan pada data. Uji autokorelasi spasial menggunakan Statistik *Moran's I* pada persamaan (2) untuk mendeteksi adanya autokorelasi spasial antar pengamatan. Sementara itu, uji heterogenitas wilayah pada persamaan (3) berfungsi untuk mengevaluasi apakah terdapat perbedaan signifikan dalam pola suhu udara antara berbagai wilayah.
5. Proses perhitungan matriks bobot lokasi untuk model GSTARIMA melibatkan dua jenis bobot, yaitu bobot berdasarkan invers jarak yang mengacu pada persamaan (16), dan normalisasi korelasi silang yang tertera pada persamaan (18).
6. Melakukan pemodelan suhu udara menggunakan metode VARIMA dan GSTARIMA dengan langkah-langkah sebagai berikut:
  - a. Mengidentifikasi atau memeriksa data dengan melakukan uji stasioneritas dalam *mean* dan *varians*. Kestasioneran data suhu udara terhadap varians dapat dilihat dari nilai estimasi parameter transformasi *lambda* ( $\lambda$ ), jika nilai  $\lambda$  mendekati satu, dapat dianggap telah stasioner, begitupun sebaliknya. Jika data tidak stasioner, terapkan transformasi Box-Cox sesuai persamaan (4) hingga data stasioner terhadap varians. Untuk memeriksa kestasioneran data terhadap rata-rata, periksa plot MACF dari data yang telah stasioner terhadap *varians* atau dapat diketahui melalui uji *Augmented Dickey Fuller* (ADF) sesuai persamaan (6). Jika data tidak stasioner maka lakukan diferensiasi hingga data stasioner terhadap *mean*.
  - b. Merancang dan menganalisis model VARIMA dengan melalui tahap-tahap berikut:
    - a) Model ditentukan berdasarkan orde yang menghasilkan nilai AIC minimum, sebagaimana dijelaskan dalam persamaan (14).
    - b) Mengestimasi parameter model VARIMA sesuai dengan orde yang terbaik.
    - c) Untuk menguji kecocokan model, perlu dilakukan diagnosa di mana model dianggap tepat jika sisaan menunjukkan *white noise* berdasarkan persamaan (25) sedangkan *d-square* digunakan untuk memeriksa apakah data menyebar secara normal multivariat sesuai persamaan (26).

- c. Membentuk dan menganalisis model GSTARIMA melalui tahapan berikut:
  - a) Menentukan orde spasial dan waktu. Orde spasial dibatasi hanya berorde satu, kemudian orde waktu ditentukan berdasarkan orde hasil identifikasi model VARIMA
  - b) Mengestimasi parameter model dari GSTARIMA dengan metode penduga *Ordinary Least Square* (OLS).
  - c) Melakukan uji asumsi terhadap sisaan dari model GSTARIMA yaitu dengan uji Ljung-Box melalui persamaan (25) untuk melihat apakah data telah bersifat *white noise* sedangkan *d-square* digunakan untuk melihat apakah data telah menyebar normal multivariat sesuai persamaan (26).
7. Melakukan perbandingan model VARIMA dan GSTARIMA serta hasil prediksinya untuk peramalan suhu udara melalui langkah-langkah berikut:
  - a. Melakukan peramalan dengan menggunakan model VARIMA dan GSTARIMA.
  - b. Membandingkan dan memilih model terbaik, serta menilai keakuratan hasil peramalan berdasarkan nilai RMSE terkecil pada data *out sample* sesuai persamaan (27).
  - c. Menginterpretasi hasil peramalan yang diperoleh.