

BAB I PENDAHULUAN

1.1 Latar Belakang

Curah hujan di Indonesia memiliki intensitas yang bervariasi pada waktu-waktu tertentu, dengan beberapa bulan menunjukkan curah hujan yang tinggi di berbagai wilayah. Sementara itu, wilayah lain mengalami curah hujan yang rendah, bahkan ada daerah yang tidak hujan sama sekali (Septiani 2024). Sebagai negara yang terletak di garis khatulistiwa, Indonesia mengalami dua musim utama, yaitu musim kemarau dan musim hujan yang keduanya mempengaruhi aktivitas masyarakat (Rupa et al., n.d.).

Musim kemarau dapat menyebabkan kekeringan, mengurangi ketersediaan air untuk pertanian dan kebutuhan sehari-hari serta meningkatkan risiko kebakaran lahan. Sedangkan musim hujan dengan intensitas yang tinggi dapat menimbulkan masalah seperti banjir, tanah longsor, dan kerusakan infrastruktur (Popi 2014). Curah hujan suatu lokasi dengan lokasi lain yang berdekatan memiliki kecenderungan jumlah curah hujan yang mirip karena letak daerahnya yang tidak terpaut jauh (Siswanto dkk 2019).

Dalam beberapa tahun terakhir, curah hujan di Sulawesi Selatan menunjukkan pola fluktuasi signifikan dengan periode musim hujan intensif dan musim kemarau yang jelas, dimana sejak 2020-2023, BMKG mencatat peningkatan fenomena cuaca ekstrem di wilayah ini, dengan intensitas hujan tinggi pada bulan November-Februari dan penurunan signifikan pada Juni-September (BMKG, 2023). Pola ini berkaitan erat dengan fenomena global La Niña dan El Niño yang mempengaruhi distribusi curah hujan di Indonesia.

Data dari Badan Penanggulangan Bencana Daerah (BPBD) Sulawesi Selatan mencatat bahwa dalam periode 2018-2023, terjadi 187 kejadian banjir dan 56 kejadian tanah longsor yang mengakibatkan kerugian ekonomi mencapai Rp 1,2 triliun dan mempengaruhi lebih dari 200.000 jiwa (BPBD Sulawesi Selatan, 2023). Di sisi lain, kekeringan yang terjadi pada musim kemarau telah berdampak pada penurunan produksi pertanian hingga 30% di beberapa wilayah, yang secara langsung mempengaruhi ketahanan pangan dan pendapatan masyarakat lokal (Dinas Pertanian Sulawesi Selatan, 2022). Hal ini menunjukkan pentingnya prediksi curah hujan yang akurat untuk mitigasi bencana dan perencanaan ekonomi regional.

Kondisi geografis Sulawesi Selatan yang beragam menyebabkan variasi signifikan antar wilayah. Untuk menangkap variasi spasial ini, digunakan data dari tiga lokasi pengamatan strategis yang memiliki karakteristik geografis berbeda di Sulawesi Selatan, yaitu: Stasiun Maritim Paotere (mewakili wilayah pesisir utara), Stasiun Hasanuddin (mewakili wilayah tengah), dan Stasiun Makassar (mewakili wilayah selatan). Ketiga stasiun ini dipilih



karena memiliki kelengkapan data yang baik dan merepresentasikan karakteristik geografis yang berbeda di provinsi ini.

Mengingat kompleksitas pola curah hujan yang dipengaruhi oleh faktor spasial dan temporal secara simultan, diperlukan model statistik yang mampu menganalisis keterkaitan antar lokasi curah hujan yang berdekatan sekaligus memperhitungkan dinamika temporal. Model prediksi konvensional yang hanya fokus pada aspek temporal seperti ARIMA sering kali tidak mampu menangkap dinamika spasial curah hujan, sehingga diperlukan pendekatan yang mengintegrasikan kedua aspek tersebut untuk meningkatkan akurasi peramalan.

Model *Space Time Autoregressive* (STAR) yang diperkenalkan oleh Ruchjana (2019) merupakan salah satu pendekatan yang mengaitkan dimensi waktu dan lokasi dalam pemodelan data. Namun, model STAR memiliki keterbatasan karena mengasumsikan bahwa parameter autoregressive dan parameter *spasio-temporal* bersifat homogen di seluruh lokasi. Asumsi ini mengimplikasikan bahwa hubungan antar variabel tidak berubah dari satu lokasi ke lokasi lainnya (Agnesya dkk, 2023), yang tidak selalu sesuai dengan realitas di lapangan, terutama untuk wilayah dengan karakteristik geografis beragam seperti Sulawesi Selatan. Keterbatasan ini membuat model STAR kurang fleksibel ketika diterapkan pada lokasi-lokasi dengan karakteristik yang berbeda (Islamiyah et al., 2018).

Untuk mengatasi keterbatasan tersebut, Borovkova dan Nurani (2002) merancang model *Generalized Space Time Autoregressive* (GSTAR) sebagai perluasan dari model STAR. Keunggulan utama model GSTAR adalah kemampuannya mengakomodasi heterogenitas parameter antar lokasi, sehingga menghasilkan model dengan parameter yang berbeda untuk setiap lokasi pengamatan. Komponen ini memungkinkan model GSTAR untuk menangkap dinamika yang kompleks dan variasi dalam keterkaitan waktu dan lokasi dengan lebih efektif, sehingga memberikan hasil yang lebih akurat dan relevan dalam analisis data *spasio-temporal* (Bstract, 2024).

Selain aspek spasial dan temporal, curah hujan di Sulawesi Selatan juga menunjukkan pola musiman (*seasonal*) yang jelas. Oleh karena itu, penambahan elemen *seasonal* pada model GSTAR menjadi penting untuk analisis data curah hujan. Dengan adanya komponen *seasonal*, model GSTAR-*Seasonal* dapat menangkap dan memprediksi pola curah hujan yang berulang seiring waktu dengan lebih akurat. Model ini memiliki keunggulan dibandingkan model prediksi konvensional karena kemampuannya mempertimbangkan tiga dimensi secara simultan, yaitu lokasi (spasial), pola perubahan dari waktu ke waktu, dan pola musiman yang berulang (*seasonal*). Kemampuan ini sangat diperlukan untuk menganalisis dan memprediksi curah hujan di Sulawesi Selatan yang menunjukkan variasi musiman dan lokasi geografis.



Dalam implementasi model *GSTAR-Seasonal*, pemilihan matriks pembobot lokasi yang tepat menjadi faktor krusial karena mencerminkan hubungan antar lokasi pengamatan. Terdapat tiga jenis pembobot lokasi yang umum digunakan dalam model *GSTAR*, yaitu bobot seragam, bobot invers jarak, dan bobot korelasi silang. Bobot seragam didasarkan pada asumsi bahwa semua lokasi dalam penelitian memiliki karakteristik yang sama (Arini dkk, 2023). Sementara itu, bobot invers jarak memberikan nilai pembobot yang lebih tinggi pada lokasi yang berdekatan, dengan asumsi bahwa lokasi yang secara geografis lebih dekat memiliki hubungan yang lebih kuat (Gusnadi dkk, 2015). Adapun bobot normalisasi korelasi silang disesuaikan dengan karakteristik data yang digunakan dan menurut Andriyani et al. (2018), bobot ini sering memberikan hasil yang lebih baik dibandingkan dengan jenis pembobot lainnya.

Beberapa penelitian telah dilakukan dengan menggunakan metode *GSTAR-Seasonal* diantaranya adalah Ahmad dan Prahutama (2021) yang membandingkan performa model *Generalized Space Time Autoregressive - Seemingly Unrelated Regression Seasonal* dengan model *VARIMA* untuk peramalan curah hujan. Hasil menunjukkan model *GSTAR-SUR Seasonal* menunjukkan performa yang lebih baik dengan nilai AIC yang lebih kecil dan akurasi peramalan yang lebih tinggi (R,Ahmad., 2021). Selanjutnya (Permata 2023) mengembangkan model *GSTAR-Seasonal* untuk meramalkan debit air di beberapa titik Daerah Aliran Sungai Brantas. Hasilnya menunjukkan model mampu menangkap pola musiman debit air dengan baik dan menghasilkan peramalan dengan akurasi sebesar 87.6%. Kemudian (Satria Utama dkk, 2025) mengembangkan model *S-GSTAR* untuk memprediksi data suhu di Provinsi Sulawesi Selatan. Hasil evaluasi menunjukkan bahwa model dengan urutan AR (4), PS (6), dan musiman ($s=12$) memberikan kinerja terbaik dengan nilai RMSE terendah.

Berdasarkan karakteristik curah hujan di Sulawesi Selatan yang menunjukkan pola spasial, temporal, dan musiman yang kompleks, penelitian ini menggunakan model *GSTAR-Seasonal* untuk pendugaan curah hujan di tiga stasiun pengamatan strategis selama periode 2018-2023. Pendekatan ini dipilih karena kemampuannya mempertimbangkan ketiga dimensi (spasial, temporal, dan musiman) secara simultan, yang diharapkan dapat menghasilkan peramalan yang lebih akurat, presisi, dan relevan.

1.2 Batasan Masalah



Batasan pada penelitian ini adalah data bulanan curah hujan di Makassar yaitu Stasiun Meteorologi Sultan Hasanuddin, Stasiun Meteorologi Maritim Paotere, dan Stasiun Klimatologi Sulawesi Januari 2018 hingga Desember 2023.

2. Model *spatio temporal* yang digunakan adalah model *Generalized Space Time Autoregressive Seasonal* dengan bobot lokasi seragam, invers jarak, dan korelasi silang serta orde spasial dibatasi hanya berorde satu.
3. Menentukan model terbaik dengan menggunakan *Root Mean Square Error*.
4. Pembentukan Model *Generalized Space Time Autoregressive* dengan *Matrix Autocorrelation Function* (MACF), *Matrix Partial Autocorrelation Function* (MPACF), dan Nilai *Akaike's Information Criterion* (AIC).

1.3 Tujuan dan Manfaat

Tujuan dari penelitian ini adalah sebagai berikut :

1. Mendapatkan model *Generalized Space Time Autoregressive Seasonal* untuk data curah hujan tiga stasiun di Makassar
2. Memperoleh hasil ramalan model *Generalized Space Time Autoregressive Seasonal* data curah hujan tiga stasiun di Makassar.

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut :

1. Secara teoritis, penelitian ini dapat memberikan pemahaman yang lebih baik tentang bagaimana pola curah hujan dipengaruhi oleh faktor ruang dan waktu serta menambah wawasan mengenai metode *Generalized Space Time Autoregressive Seasonal* dan penerapannya dalam kehidupan sehari-hari.
2. Secara praktis, penelitian dapat digunakan dalam kehidupan sehari-hari, terutama bagi pihak-pihak yang bergantung pada informasi cuaca. Misalnya: Bagi petani, hasil penelitian ini bisa membantu mereka mengetahui pola hujan sehingga bisa menentukan waktu tanam dan panen yang lebih tepat. Informasi ini dapat diakses melalui penyuluhan pertanian di dinas pertanian setempat, aplikasi cuaca pertanian, atau kerjasama dengan BMKG untuk penyebaran data iklim kepada kelompok tani.

1.4 Teori

1.4.1 Analisis Runtun Waktu

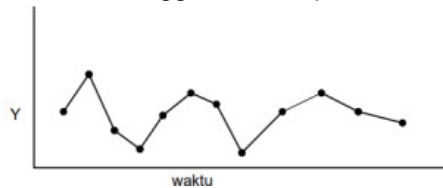
Runtun waktu adalah rangkaian data yang berupa nilai pengamatan yang diukur selama kurun waktu pengamatan yang dilakukan harus dalam unit selang waktu yang sama (Pawestri et al., 2019). Analisis runtun waktu mulai diperkenalkan oleh George E.P.Box dan Gwilym M. Jenkins pada tahun 1970. Analisis deret waktu merupakan metode yang mempelajari data deret waktu, baik dari segi teori maupun membuat peramalan. Menurut Markridakis dkk (1999), terdapat langkah



akan dalam memilih suatu metode *time-series* secara tepat, pertimbangkan jenis pola data. Beberapa pola data yang sebagai berikut.

1. Pola Data Musiman

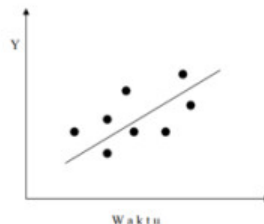
Pola data yang memiliki kecenderungan berulang pada jangka waktu dan periode tertentu. Gambar 1 menggambarkan pola data musiman



Gambar 1. Pola Data Musiman

2. Pola Data Trend

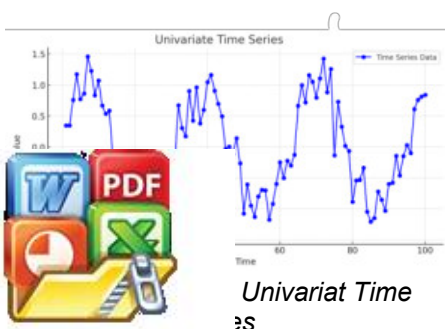
Pola data jangka Panjang yang didasari atas pertumbuhan atau penurunan. Pergerakannya cenderung sedikit demi sedikit mengalami naik atau turun. Gambar 2 menggambarkan pola data *trend*.



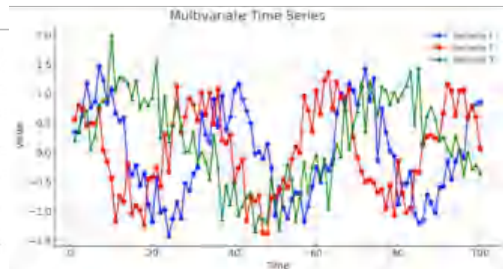
Gambar 2. Pola Data *Trend*

1.4.2 Univariate dan Multivariate Time Series

Model Univariate *Time Series* adalah model yang hanya melibatkan satu variabel runtun waktu, yang dapat diidentifikasi melalui plot *Autocorelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF), seperti di tunjukkan pada Gambar 3. Sedangkan Model *Multivariate Time Series* terdiri dari pengamatan beberapa variabel yang disajikan dalam *vector time series*. Kestasioneritasan model *Multivariate Time Series* dapat dilihat dari MACF, dimana plot yang menurun perlahan (tails off) menandakan data tidak stasioner, seperti ditunjukkan pada Gambar 4.



Univariate Time Series



Gambar 4. Grafik Multivariate Time Series

Berdasarkan Gambar 3 *Univariate Time Series* adalah sekumpulan data yang dikumpulkan berdasarkan waktu dengan hanya satu variabel yang diamati. Dari gambar sumbu X (*time*) menunjukkan urutan waktu, yang bisa berupa detik, menit, hari, bulan, atau tahun dan sumbu Y (*value*) menunjukkan nilai dari variabel yang diukur dalam periode waktu tertentu. Garis berwarna biru menggambarkan perubahan nilai variabel tersebut seiring dengan waktu.

Berdasarkan Gambar 4 grafik menampilkan tiga variabel (biru, merah, dan hijau) dengan fluktuasi berbeda tetapi bergerak dalam pola yang dapat saling berkorelasi atau independent. Perbedaan pola antar variabel menunjukkan bahwa setiap variabel memiliki dinamika sendiri, tetapi mungkin terdapat keterkaitan atau hubungan antar variabel dalam periode tertentu. *Multivariate Time Series* seperti ini sering digunakan dalam analisis keuangan, cuaca, dan sistem kompleks lainnya.

1.4.3 Stasioneritas Time Series

Runtun waktu dikatakan stasioner adalah relatif tidak terjadi kenaikan ataupun penurunan nilai secara tajam pada data atau fluktuasi data berada pada sekitar nilai rata-rata yang konstan (Fadliani et al., 2021). Data secara kasarnya harus horizontal sepanjang sumbu waktu. Terdapat dua macam stasioneritas data yaitu stasioner terhadap ragam dan stasioner terhadap rata-rata. Kestasioneran pada rata-rata dapat dilakukan dengan pengujian terhadap ada tidaknya akar unit (*unit root*) dalam variabel dengan uji *Augmented Dickey Fuller* (ADF). Pemeriksaan stasioneritas pada rata-rata dapat menggunakan uji *Augmented Dickey Fuller* (ADF). Untuk memudahkan pemahaman mengenai uji *Augmented Dickey Fuller* (ADF), diberikan model AR (1):

$$Z_t = \phi Z_{t-1} + e \quad (1)$$

Jika Persamaan (1) dikurangi Z_{t-1} pada sisi kanan dan kiri, maka:

$$Z_t - Z_{t-1} = \phi Z_{t-1} - Z_{t-1} + e_t \quad (2)$$

$$\Delta Z_t = \{\phi - 1\} Z_{t-1} + e_t$$

$$\Delta Z_t = \phi^* Z_{t-1} + e_t$$

Dimana $e_t \sim IID(0, \sigma^2)$

Dari Persamaan (2) dapat di buat hipotesis:

$$H_0 : \phi^* = 0 \text{ (data tidak stasioner)}$$

$$H_1 : \phi^* < 0 \text{ (data stasioner)}$$

Statistika uji yang digunakan adalah:

$$t = \frac{\hat{\phi}^*}{SE(\hat{\phi}^*)} \quad (3)$$



ilai duga parameter autoregressive (AR)

tandar error (SE) $\hat{\phi}^*$

$$SE(\hat{\phi}^*) = \frac{\sigma(\hat{\phi}^*)}{\sqrt{t}} \quad (4)$$

Dengan $\sigma(\hat{\phi}^*)$ adalah simpangan baku dari $(\hat{\phi}^*)$ dan t adalah banyaknya pengamatan. Nilai statistic uji t dibandingkan dengan table *Dickey Fuller* dengan Tingkat kesalahan sebesar α dan derajat bebas $t \sim t_{(\alpha,t)}$ berdasarkan kriteria uji sebagai berikut:

- Tidak ada cukup bukti untuk menerima H_0 atau data stasioner jika $t > t_{(\alpha,t)}$
- Cukup bukti untuk menerima H_0 atau data tidak stasioner jika $t \leq t_{(\alpha,t)}$

Pada $\alpha = 0,05$ jika $p\text{-value}$ kurang dari α maka H_0 ditolak, artinya deret stasioner, sedangkan penerimaan H_0 menunjukkan bahwa deret tidak stasioner, sehingga perlu dilakukan *differencing*.

Differencing

Jika suatu data deret waktu tidak stasioner terhadap rata-rata, maka data dapat diubah menjadi deret stasioner melalui *differencing* (pembedaan). Banyaknya *differencing* yang dilakukan untuk mencapai keadaan stasioner dinotasikan dengan d . Proses *differencing* ini bisa dilakukan hingga orde $ke - d$. Jika proses *differencing* dilakukan sebanyak satu kali dinamakan *differencing* dengan orde $d = 1$, bila dilakukan *differencing* sebanyak dua kali maka dinamakan *differencing* dengan orde ($d = 2$) dan seterusnya (Suhendra et al., 2023). *Differencing* pertama $d = 1$ dari Z_t dapat dituliskan sebagai berikut (Cryer dan Chan, 2008):

$$\Delta Z = Z_t - Z_{t-1} \quad (5)$$

di mana:

- Z_t = nilai pengamatan pada waktu $ke t$
- Z_{t-1} = nilai pengamatan pada waktu $ke t - 1$
- ΔZ = hasil *differencing* pertama pengamatan Z_t

Sehingga bentuk *differencing* kedua ($d = 2$) adalah:

$$\Delta^2 Z = \Delta Z_t - \Delta Z_{t-1} \quad (6)$$

Diferencing dilakukan hingga data hasil *differencing* menunjukkan keadaan yang stasioner terhadap rata-rata.

1.4.4 Data Spasial

Data Spasial dapat dikategorikan sebagai data yang memiliki ketergantungan karena proses pengumpulannya dilakukan di berbagai titik lokasi berbeda, dimana terdapat keterkaitan antara hasil pengukuran dengan posisi geografisnya. Hal ini mengindikasikan bahwa data spasial merupakan kumpulan pengamatan yang diperoleh dari lokasi-lokasi spesifik dengan pola saling mempengaruhi satu sama lain (Jainudin et al., 2019). Berdasarkan tipe data, dapat dibedakan menjadi pemodelan dengan pendekatan titik elan spasial dengan pendekatan titik antara lain yaitu *ighted Regression* (GWR), *Geographically Weighted Poisson* $\mathcal{R}$), *Space Time Autoregressive* (STAR), dan *Generalized gressive* (GSTAR).



1.4.5 Indeks Gini

Indeks gini adalah suatu koefisien yang menunjukkan tingkat ketidakmerataan suatu pengamatan untuk menguji keheterogenan lokasi. Nilai indeks gini berkisar antara 0 hingga 1, suatu lokasi dikatakan memiliki pemerataan sempurna (homogen) jika nilai indeks gini mendekati 0 dan sebaliknya (Nurfadila & Intan, 2023). Berikut rumus untuk menghitung indeks gini:

$$G = 1 + \frac{1}{T} - \frac{2}{(T^2 \bar{z}_t)} \sum_{i=1}^N \sum_{t=1}^T z_t \quad (7)$$

keterangan:

z_t = nilai variable yang diamati berdasarkan indeks sampel t

$\bar{z}_t$ = rata-rata nilai seluruh variable yang diamati

T = banyaknya pengamatan (waktu)

N = banyaknya lokasi pengamatan

1.4.6 Seasonal Time Series

Suatu runtun waktu terkadang mengandung fenomena *seasonal* atau musiman yang berulang pada periode waktu tertentu. Periode waktu terkecil untuk fenomena berulang ini disebut *seasonal* period atau periode musiman (Gusnadi et al., 2015). Model ARIMA musiman multiplikatif dalam prosedur Box-Jenkins adalah sebagai berikut:

$$\phi_P(B^s)\phi_P(B)(1-B)^d(1-B^s)^D Z_t = \theta_Q(B)\theta_Q(B^s) \quad (8)$$

Secara Umum, komponen *autoregressive* musiman dan non musiman memiliki PACF *cutting off* atau terpotong pada *lag* musiman dan non musiman.

1.4.7 Matrix Autocorrelation Function

Matrix Autocorrelation Function (MACF) merupakan alat yang digunakan untuk mengidentifikasi pola hubungan dalam data deret waktu multivariat, yang menunjukkan korelasi antara pengamatan pada waktu yang berbeda untuk setiap komponen dan antar komponen dalam vektor deret waktu (Nugroho, 2022). Jika diberikan suatu *vector time series* sebanyak T observasi yaitu $Z_1, Z_2, \dots, Z_T$ maka permasalahan matriks korelasi sampel dapat dihitung sebagai berikut:

$$\hat{\rho}(k) = [\hat{\rho}_{ij}(k)] \quad (9)$$

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

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

dan $\bar{z}_i$ dan $\bar{z}_j$ merupakan rata-rata sampel dari komponen deret yang



untuk fungsi matriks korelasi parsial dituliskan sebagai berikut:

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

Bentuk plot akan semakin kompleks jika dimensi dan vektornya semakin besar. Menurut Tiao dan Box (1981), terdapat metode lain untuk memudahkan pengidentifikasian yang menjelaskan korelasi sampel dengan menggunakan simbol yang dinotasikan (+), (-), dan (.) pada matriks korelasi sampel ke (i, j) yaitu:

1. Simbol (+) menunjukkan bahwa nilai $\hat{\rho}_{ij}(k)$ lebih besar dari 2 kali standar error dari $\hat{\rho}_{ij}(k)$ dan menunjukkan hubungan korelasi positif antara komponen (i, j) .
2. Simbol (-) menunjukkan bahwa nilai $\hat{\rho}_{ij}(k)$ lebih kecil dari -2 kali standar error dari $\hat{\rho}_{ij}(k)$ dan menunjukkan hubungan korelasi negatif antara komponen (i, j) .
3. Simbol (.) menunjukkan bahwa nilai $\hat{\rho}_{ij}(k)$ berada diantara ± 2 kali standar error dari $\hat{\rho}_{ij}(k)$ dan menunjukkan tidak ada korelasi antara komponen (i, j) .

Standar error dari nilai $\hat{\rho}_{ij}(k)$ dapat dihitung dengan persamaan sebagai berikut:

$$S_{\hat{\rho}_{ij}(k)} = \sqrt{\frac{1}{T} \left(1 + 2\hat{\rho}_{ij}^2(1) + 2\hat{\rho}_{ij}^2(2) + \dots + 2\hat{\rho}_{ij}^2(k-1) \right)} \quad (12)$$

Dimana T adalah banyaknya pengamatan (observasi).

1.4.8 Matriks Partial Autocorrelation Function

Fungsi matriks autokorelasi parsial sangat berguna dalam mengidentifikasi orde dari sebuah model AR (p). *Matrix Partial Autocorrelation Function* (MPACF) pada lag ke- k yang dinotasikan dengan $P(k)$. merupakan koefisien matriks terakhir ketika data yang dimasukkan ke dalam suatu proses *vector time series* dari orde (p). $P(k)$ merupakan persamaan untuk $\phi_{k,k}$ dalam regresi linier multivariat.

Persamaan untuk matriks autokorelasi parsial adalah sebagai berikut:

$$P(k) = \begin{cases} \Gamma'(1)[\Gamma(0)]^{-1} & k = 1 \\ \{\Gamma'(k) - \mathbf{c}'(k)[\mathbf{A}(k)]^{-1}\mathbf{b}(k)\} \{\Gamma'(0) - \mathbf{b}'(k)[\mathbf{A}(k)]^{-1}\mathbf{b}(k)\}^{-1}, & k > 1 \end{cases} \quad (13)$$

Untuk $k \geq 2$, maka persamaan untuk nilai $\mathbf{A}(k)$, $\mathbf{b}(k)$ dan $\mathbf{c}(k)$ adalah sebagai berikut:

$$\mathbf{A}(k) = \begin{bmatrix} \Gamma(0) & \Gamma'(1) & \dots & \Gamma'(k-2) \\ \Gamma(1) & \Gamma(0) & \dots & \Gamma'(k-3) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma(k-2) & \Gamma(k-3) & \dots & \Gamma'(0) \end{bmatrix}$$

$$\mathbf{b}(k) = \begin{bmatrix} \Gamma'(k-1) \\ \Gamma'(k-2) \\ \vdots \\ \Gamma'(1) \end{bmatrix} \quad (14)$$

$$\mathbf{c}(k) = \begin{bmatrix} \Gamma(1) \\ \Gamma(2) \\ \vdots \\ \Gamma(k-2) \end{bmatrix}$$



1.4.9 Model Generalized Space Time Autoregressive

Model *Generalized Space Time Autoregressive* (GSTAR) merupakan suatu model yang lebih fleksibel sebagai generalisasi dari model *Space Time Autoregressive* (STAR) yang mampu mengungkapkan keterkaitan linear dari waktu dan lokasi. Model GSTAR memungkinkan nilai-nilai parameter *autoregressive* bervariasi untuk setiap lokasi. Pada model *Generalized Space Time Autoregressive* (GSTAR), keterikatan spasial dinyatakan oleh matriks pembobot.

Jika kita memiliki data *time series* $\{Y(t): t = 0, 1, 2, \dots; 1 = 1, 2, \dots, N\}$ yang merupakan sebuah *multivariate time series* dengan N pengamatan, maka model GSTAR dengan ordo waktu AR (p) dan ordo spasial $(\lambda_1, \lambda_2, \dots, \lambda_p)$ dapat dinyatakan sebagai GSTAR $(p, \lambda_1, \lambda_2, \dots, \lambda_p)$ seperti yang dijelaskan oleh ruchjana (2002):

$$Y_t = \sum_{k=1}^p \left(\Phi_{k0} Y_{t-k} + \sum_{l=1}^{\lambda_p} \Phi_{kl} W^{(l)} Y_{t-k} \right) + \varepsilon_t \quad (15)$$

Keterangan:

Y_t = vector acak berukuran $(N \times 1)$ pada waktu t

Φ_{k0} = matriks kosfisien parameter waktu *diag* $(\phi_{k0}^1, \dots, \phi_{k0}^N)$

Φ_{kl} = matriks kosfisien parameter spasial *diag* $(\phi_{kl}^1, \dots, \phi_{kl}^N)$

$W^{(l)}$ = nilai matriks pembobot berukuran $(N \times N)$ pada *lag* spasial ke- l dengan

syarat $w_{ii}^{(l)} = 0$ dan $w_{i \neq j}^{(l)} w_{ij}^{(l)} = 1$

$\varepsilon^{(l)}$ = vector galat yang memenuhi iid dan berdistribusi normal multivariate

Sebagai contoh, model GSTAR dengan ordo waktu 1 dan ordo spasial 1 pada lokasi yang berbeda, atau GSTAR (1_1) dapat dinyatakan sebagai berikut sesuai dengan Persamaan (18):

$$Y_t = \Phi_{10} Y_{t-1} + \Phi_{11} W^{(1)} Y_{t-1} + \varepsilon_t \quad (16)$$

Berdasarkan Persamaan (16), model untuk tiga lokasi yang berbeda dapat dituliskan sebagai berikut :

$$Y_{1,t} = \phi_{10}^1 Y_{1,t-1} + \phi_{11}^1 w_{12} Y_{2,t-1} + \phi_{11}^1 w_{13} Y_{3,t-1} + \varepsilon_{1,t}$$

$$Y_{2,t} = \phi_{10}^2 Y_{2,t-1} + \phi_{11}^2 w_{21} Y_{1,t-1} + \phi_{11}^2 w_{23} Y_{3,t-1} + \varepsilon_{2,t}$$

$$Y_{3,t} = \phi_{10}^3 Y_{3,t-1} + \phi_{11}^3 w_{31} Y_{1,t-1} + \phi_{11}^3 w_{32} Y_{2,t-1} + \varepsilon_{3,t}$$

Atau dapat dinyatakan dalam notasi matriks seperti yang ditunjukkan pada persamaan berikut:



$$= \begin{bmatrix} \phi_{10}^1 & 0 & 0 \\ 0 & \phi_{10}^2 & 0 \\ 0 & 0 & \phi_{10}^3 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \\ Y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \phi_{11}^1 & 0 & 0 \\ 0 & \phi_{11}^2 & 0 \\ 0 & 0 & \phi_{11}^3 \end{bmatrix} \begin{bmatrix} Y_{1,t-1} \\ Y_{2,t-1} \\ Y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{bmatrix} \quad (17)$$

Dalam menentukan ordo model GSTAR, ordo waktu dapat dipilih menggunakan AIC (Wei,2006), sementara ordo spasial umumnya dibatasi hingga ordo satu karena ordo yang lebih tinggi sulit untuk diinterpretasikan. Penentuan ordo model juga dapat dilakukan dengan melihat plot MCCF dan MPCCF yang terbentuk. Jika data yang digunakan memiliki pola musiman, maka digunakan model GSTAR musiman. Secara matematika, model GSTAR $(p; \lambda_1, \lambda_2, \dots, \lambda_p)^s$ dapat ditulis dalam bentuk sebagai berikut:

$$Y_t = \sum_{k=1}^p \left(\Phi_{k0}^s Y_{t-k} + \sum_{l=1}^{\lambda_p} \Phi_{kl}^{\lambda_p} W^{(l)} Y_{t-k} \right) + \varepsilon_t \quad (18)$$

Untuk model GSTAR dengan ordo musiman 12 dan ordo spasial 1 ditulis dengan persamaan sebagai berikut:

$$Y_t = \Phi_{10}^{12} Y_{t-12} + \Phi_{11}^{12} W^{(1)} Y_{t-12} + \varepsilon_t \quad (19)$$

a. Bobot Lokasi Seragam

Bobot lokasi seragam memberikan nilai yang sama untuk setiap lokasi dan mengasumsikan jara k antar lokasi sama (Widyastuti et al., 2016). Bobot lokasi seragam dapat dihitung dengan rumus:

$$w_{ij} = \frac{1}{n} \quad (20)$$

$$w_{ij} = \begin{cases} \frac{1}{n}, i \neq j \\ 0, i = j \end{cases} ; \sum_{j=1}^N w_{ij} = 1, \forall i \text{ dan } \sum_{i=1}^N \sum_{j=1}^N w_{ij} = N$$

b. Bobot Invers Jarak

Menggunakan jarak sebenarnya antar lokasi, perhitungan bobot diperoleh dari hasil invers jarak sebenarnya yang kemudian dinormalisasi. Bobot lokasi invers jarak memberikan nilai koefisien yang kecil untuk jarak yang lebih jauh dan sebaliknya (Homepage et al.,2023). Bobot lokasi invers jarak dapat dihitung dengan rumus sebagai berikut (Siregar,2015):

$$d_{ij} = d_{ji} = \sqrt{(x_i(u_i) - x_j(u_j))^2 + (x_i(v_i) - x_j(v_j))^2} \quad (21)$$

$$w_{ij} = w_{ji} = \frac{1}{d_{ij}} = \frac{1}{d_{ji}} \quad (22)$$

dan normalisasi bobot lokasi invers jarak dapat dihitung menggunakan rumus berikut:

$$w_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j=1}^n \frac{1}{d_{ij}}}, j \neq i \quad (23)$$



lokasi ke- i dengan $i = 1, 2, \dots, N$

lintang lokasi

bujur lokasi

jarak lokasi ke- i terhadap lokasi lainnya (lokasi ke- j)

c. Bobot Normalisasi Korelasi Silang

Pada tahun 2006, Suhartono dan Subanar memperkenalkan suatu metode penentuan bobot baru yaitu dengan menggunakan bobot normalisasi korelasi silang antar lokasi pada *lag* waktu yang bersesuaian. Pembobot normalisasi korelasi silang tidak mensyaratkan aturan tertentu, seperti bergantung pada jarak antar lokasi. Secara umum korelasi silang antar lokasi ke-*i* dan ke-*j* pada *lag* waktu ke-*k* didefinisikan sebagai berikut:

$$\rho_{ij}(k) = \frac{\gamma_{ij}(k)}{\sigma_i \sigma_j} \quad (24)$$

dengan $\gamma_{ij}(k)$ merupakan kovarian silang antar amatan pada lokasi ke-*i* dan lokasi ke-*j* pada *lag* waktu ke-*k*, σ_i dan σ_j adalah simpangan baku dari amatan pada lokasi ke-*i* dan lokasi ke-*j*. Penduga korelasi silang pada data contoh adalah sebagai berikut:

$$r_{ij}(k) = \frac{\sum_{t=k+1}^n [z_i(t) - \bar{z}_i][z_j(t-k) - \bar{z}_j]}{\sqrt{(\sum_{t=1}^n [z_i(t) - \bar{z}_i]^2)(\sum_{t=1}^n [z_j(t) - \bar{z}_j]^2)}} \quad (25)$$

Penentuan bobot normalisasi korelasi silang dapat diselesaikan dengan menormalisasikan korelasi silang antar lokasi pada *lag* yang sesuai. Proses ini secara umum menghasilkan pembobot lokasi untuk model GSTAR (1₁) sebagai berikut:

$$W_{ij} = \frac{|r_{ij}(k)|}{\sum_{k \neq 1} |r_{ik}(k)|}, i \neq j \quad (26)$$

harus memenuhi $\sum_{i \neq j} |W_{ij}| = 1$.

1.4.10 Estimasi Parameter *Generalized Space Time Autoregressive*

Suatu model GSTAR dapat direpresentasikan sebagai sebuah model linear dan parameter-parameter *autoregressive* model dapat diestimasi menggunakan metode kuadrat terkecil atau metode *least square* (Irawati et al., 2015). Dengan ordo *autoregressive* $p = 1$ dan ordo spasial $\lambda_p = 1$ maka Persamaan GSTAR (1₁) dapat dituliskan sebagai berikut:

$$Y_t = \Phi_{10} Y_{t-1} + \Phi_{11} W^{(1)} Y_{t-1} + \epsilon_t \quad (27)$$

Metode *least square* sering digunakan untuk melakukan estimasi parameter model linear, sehingga metode ini dapat diterapkan pada model GSTAR (1₁) untuk beberapa lokasi berikut:

$$Y_{i,t} = X_i \beta_i + \epsilon_i \quad (28)$$

banyaknya pengamatan ke-*t* ($t = 1, 2, \dots, T$) untuk lokasi ke-*j* $\beta_i = (\phi_{i0}^1, \phi_{i1}^1)'$ yang dapat dijabarkan dalam bentuk matriks



$$Y = \begin{bmatrix} Y_{i,t} \\ Y_{i,t} \\ \vdots \\ Y_{i,t} \end{bmatrix}, X = \begin{bmatrix} Y_{i,0} & V_{i,0} \\ Y_{i,1} & V_{i,1} \\ \vdots & \vdots \\ Y_{i,T-1} & V_{i,T-1} \end{bmatrix}, \beta = \begin{bmatrix} \phi_{i0}^1 \\ \phi_{i1}^1 \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \vdots \\ \varepsilon_{i,T} \end{bmatrix}$$

dimana $V_{i,(t)} = j_{i \neq j} w_{ij} Y_{j,T}$ yang mengindikasikan bahwa β_i dapat dihitung pada masing-masing lokasi tetapi tetap bergantung pada nilai Y_t di lokasi yang lain. Jika $\beta_i = (\phi_{10}^1, \phi_{11}^1, \phi_{20}^1, \phi_{21}^1, \dots, \phi_{N0}^1, \phi_{N1}^1)$, maka Persamaan (24) dapat dijabarkan dalam bentuk matriks yang lebih rinci seperti pada Persamaan berikut:

$$\begin{bmatrix} Y_{i,1} \\ Y_{i,2} \\ \vdots \\ Y_{i,T} \\ \vdots \\ Y_{N,1} \\ Y_{N,2} \\ \vdots \\ Y_{N,T} \end{bmatrix} = \begin{bmatrix} Y_{i,0} & V_{i,1} & \dots & 0 & 0 \\ Y_{i,1} & V_{i,2} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ Y_{i,T-1} & V_{i,T-1} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & Y_{N,0} & V_{N,1} \\ 0 & 0 & \dots & Y_{N,1} & V_{N,2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & Y_{N,T-1} & Y_{N,T-1} \end{bmatrix} \begin{bmatrix} \phi_{10}^1 \\ \phi_{11}^1 \\ \vdots \\ \phi_{N0}^1 \\ \phi_{N1}^1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \vdots \\ \varepsilon_{i,T} \\ \vdots \\ \varepsilon_{N,1} \\ \varepsilon_{N,2} \\ \vdots \\ \varepsilon_{N,T} \end{bmatrix} \quad (29)$$

Pada prinsipnya metode OLS mengestimasi parameter β_i dengan meminimumkan jumlah kuadrat *error* dengan cara sebagai berikut:

$$\begin{aligned} Y &= X\beta + \varepsilon \\ \varepsilon &= Y - X\beta \\ \varepsilon' \varepsilon &= (Y - X\beta)'(Y - X\beta) \\ &= (Y' - \beta' X')(Y - X\beta) \\ &= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta \\ &= Y'Y - 2\beta'X'Y + \beta'X'X\beta \end{aligned} \quad (30)$$

Persamaan (30) diturunkan terhadap β menjadi:

$$\frac{\partial(\varepsilon' \varepsilon)}{\partial(\beta)} = -2X'Y + 2X'X\beta \quad (31)$$

Kemudian Persamaan (31) diminimumkan dengan :

$$\begin{aligned} -2X'Y + 2X'X\beta &= 0 \\ 2X'X\beta &= 2X'Y \\ \beta &= (X'X)^{-1}X'Y \end{aligned}$$

Sehingga diperoleh estimator $\hat{\beta}$ sebagai berikut :

$$\hat{\beta} = [X'X]^{-1}X'Y \quad (32)$$



Pemilihan Model

Urutan *autoregressive* p dari semua kemungkinan model digunakan metode *Akaike's Information Criterion* (AIC). Kriteria pemilihan model peramalan terbaik ditentukan dengan nilai *Root Mean Square Error* (RMSE) (Arini et al., 2023). *Information Criterion*

Perhitungan nilai AIC (p) diperoleh melalui persamaan berikut:

$$AIC(p) = \ln \left(|\hat{\Sigma}_p| + \frac{2K^2p}{T} \right) \quad (33)$$

dimana K adalah banyaknya parameter dalam model, T merupakan banyak pengamatan, dan $\hat{\Sigma}_p$ merupakan matriks dugaan varian-kovarian residual. Orde *autoregressive* p dari model GSTAR ditentukan oleh nilai $AIC(p)$ terkecil dari semua kemungkinan orde model

b. *Root Mean Square Error* (RMSE)

Pemilihan model terbaik dilakukan apabila terdapat lebih dari satu model. Indikator yang dapat digunakan untuk pemilihan model terbaik adalah RMSE (*Root Mean Square Error*). Tujuan dari model peramalan adalah untuk meramalkan nilai yang akan datang dengan *error* sekecil mungkin. RSME mengukur perbedaan antara nilai peramalan dari model dengan nilai amatan, berikut adalah rumus untuk perhitungan RMSE (Wei, 2006):

$$RMSE = \sqrt{\frac{1}{m} \sum_{l=1}^m (Y_{n+l} - \hat{Y}_n(l))^2} \quad (34)$$

keterangan:

m = banyaknya ramalan

n = banyaknya data *training*

y_l = data *testing* ke- l

$\hat{Y}_l$ = data hasil ramalan ke- l

1.4.12 Pengujian Asumsi Residual

Asumsi dasar yang harus dipenuhi adalah *error vector* bersifat *white noise*. Residual bersifat *white noise* mengartikan bahwa residual dari masing-masing data adalah saling independen. Berdasarkan model estimasi, penduga untuk residual dapat ditulis sebagai berikut:

$$\hat{a}(t) = Z(t) - \hat{Z}(t)$$

Pengecekan asumsi *white noise* residual dapat dilihat menggunakan ACF atau PACF dari residual. Jika setiap plot ACF tidak signifikan secara statistik terhadap 2 standar *error* dari ACF residual, maka residual bersifat *white noise*. Cara lain yang dapat digunakan adalah dengan kriteria minimum AIC dari residual. Jika nilai AIC residual terkecil berada pada *lag* ke-0 maka dapat dikatakan bahwa tidak ada korelasi antar masing-masing residual, yang berarti residual bersifat *white noise* (Yuliyanti & Arliani, 2022).



Akurasi

percentage error (MAPE) adalah pengukuran statistic tentang (prediksi) pada metode peramalan. Metode MAPE memberikan a besar kesalahan peramalan dibandingkan dengan nilai series tersebut. Semakin kecil nilai presentasi kesalahan

(*percentage error*) maka MAPE semakin akurat hasil peramalan tersebut. MAPE dapat dihitung dengan rumus berikut:

$$MAPE = \frac{\sum_{t=1}^n |x_t - \hat{x}_t|}{n} \times 100\% \quad (35)$$

dimana,

x_t = Data actual ke- t

$\hat{x}_t$ = Hasil peramalan ke- t

n = Banyaknya data



BAB II METODE PENELITIAN

2.1. Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari *website* Badan Meteorologi Klimatologi dan geofisika (BMKG) yang dapat diakses melalui <https://dataonline.bmkg.go.id/>. Data tersebut berupa data curah hujan tiga stasiun di Makassar yaitu Pertama, Stasiun Meteorologi Sultan Hasanuddin. Kedua, Stasiun Meteorologi Maritim Paotere, dan Ketiga Stasiun Klimatologi Sulawesi Selatan.

Data dibagi menjadi dua bagian, yaitu data *training* mulai dari Januari 2018 hingga Desember 2022 dengan total 60 pengamatan, dan data *testing* dari Januari 2023 hingga Desember 2023 dengan 12 pengamatan.

2.2. Prosedur Penelitian

Berikut adalah langkah-langkah pembentukan model *GSTAR Seasonal*

1. Eksplorasi Data Awal
 - Statistik deskriptif curah hujan 3 stasiun.
 - Korelasi antar lokasi untuk mengetahui hubungan spasial.
 - Perhitungan Indeks Gini untuk mengetahui Tingkat heterogenitas spasial antar lokasi, menggunakan Persamaan (7).
2. Identifikasi Orde Model
 - Melihat pola MACF dan MPACF untuk menentukan struktur orde waktu dan spasial model multivariat.
 - Digunakan Persamaan (9) dan (10) untuk MACF dan Persamaan (13) untuk MPACF.
 - Model dengan nilai AIC terkecil dipilih sebagai model optimal, dihitung dengan Persamaan (13).
3. Uji dan Transformasi Stasioneritas
 - Menggunakan uji ADF untuk menilai kestasioneritas, dengan dasar pada Persamaan (1)-(4).
 - Jika tidak stasioner, maka dilakukan *differencing* sesuai Persamaan (5) dan (6).
4. Penentuan Matriks Pembobot Lokasi
 - Bobot Seragam dihitung menggunakan Persamaan (20).



$$W = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

- Bobot invers jarak menggunakan jarak antar lokasi Persamaan (21)-(23).

$$w_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j=1}^n \frac{1}{d_{ij}}}, \quad j \neq i,$$

- Bobot Korelasi Silang menggunakan Persamaan (24)-(26).

$$W_{ij} = \frac{|r_{ij}(1)|}{\sum_{j \neq i} |r_{ij}(1)|},$$

5. Estimasi Parameter Model

- Estimasi dilakukan menggunakan metode *Least Squares* pada model GSTAR, sebagaimana dijelaskan dalam Persamaan (28)-(32).

6. Diagnostik Model

- Pemeriksaan *white noise* residual dilakukan melalui plot ACF dan nilai AIC residual Persamaan (33)

7. Evaluasi Akurasi Model

- Digunakan dua ukuran kesalahan peramalan:
 - *Root Mean Square* (RMSE) Persamaan (34).
 - *Mean Absolute Percentage Error* Persamaan (35).

8. Melakukan Peramalan curah hujan tiga stasiun di Sulawesi Selatan menggunakan data *test* tahun 2023.

