

BAB I

PENDAHULUAN

1.1. Latar Belakang

Inflasi adalah kenaikan harga barang dan jasa yang berlangsung secara terus-menerus. Inflasi merupakan permasalahan ekonomi yang tidak bisa diabaikan karena dapat menimbulkan dampak yang sangat luas sehingga inflasi sering menjadi target kebijakan pemerintah. Permasalahan ekonomi yang banyak terjadi di beberapa negara baik negara maju maupun negara berkembang seperti Indonesia adalah Inflasi. Menurut Bank Indonesia (BI) (2020) menjelaskan bahwa inflasi merupakan kenaikan harga barang dan jasa secara umum dan terus-menerus dalam jangka waktu tertentu. Inflasi yang rendah dan stabil diartikan sebagai prasyarat bagi pertumbuhan ekonomi yang berkesinambungan dimana pada akhirnya memberikan manfaat bagi peningkatan kesejahteraan masyarakat.

Menurut Badan Pusat Statistik (2016) bahwa inflasi juga dapat diartikan sebagai penurunan nilai mata uang terhadap nilai barang dan jasa secara umum. Angka inflasi sebelumnya akan mempengaruhi angka inflasi saat ini. Suatu perubahan pada angka inflasi terkadang dapat diprediksi dan kadang juga secara tak terduga. Misalnya kenaikan inflasi yang dapat diprediksi adalah ketika permintaan terhadap barang dan jasa di masyarakat meningkat pada waktu-waktu tertentu seperti pada saat hari raya. Kenaikan inflasi juga dapat terjadi secara tak terduga misalnya bencana alam seperti bencana banjir, tsunami, dan sebagainya. Jika laju inflasi suatu negara atau wilayah dapat terkendali dengan baik, maka akan berdampak baik pula pada perekonomian di suatu negara atau wilayah tersebut. Begitupun jika laju inflasi tidak dapat dikontrol dengan baik, maka dapat memberikan dampak yang negatif bagi suatu negara atau wilayah tersebut dalam perekonomiannya. Terdapat beberapa faktor yang dapat meningkatkan atau menurunkan inflasi diantaranya adalah adanya daya tarik yang tinggi pada suatu barang atau jasa, meningkatnya biaya produksi yang dapat mempengaruhi harga barang atau jasa, dsb.

Inflasi dapat digunakan sebagai acuan informasi dasar dalam pengambilan keputusan di berbagai bidang misalnya pada bidang tingkat ekonomi makro atau tingkat ekonomi mikro, ataupun dalam kebijakan pemerintah yang dapat mengatur perekonomian melalui pengelolaan pendapatan dan pengeluaran suatu negara oleh pemerintah, serta kebijakan moneter oleh Bank Indonesia. Sehingga, sangat penting bagi masyarakat atau negara dalam pengambilan keputusan di berbagai bidang baik bidang ekonomi agar dapat meramalkan inflasi.

Penyebab inflasi disebabkan oleh ekspektasi inflasi, volatilitas, dan juga dapat disebabkan oleh nilai tukar dan ketidakstabilan pasokan dan permintaan karena jumlah barang atau jasa suatu masyarakat tidak stabil pada suatu wilayah atau lokasi tersebut. Sehingga, setiap wilayah memerlukan wilayah lainnya agar memenuhi kebutuhan barang atau jasa yang belum terpenuhi sendiri di wilayah yang bersangkutan tersebut. Oleh karena itu, dapat menimbulkan ketergantungan antar

wilayah dalam hal pemenuhan barang atau jasa misalnya pada wilayah provinsi Sulawesi Selatan.

Provinsi Sulawesi Selatan merupakan salah satu wilayah yang penting di Indonesia dengan dinamika ekonomi yang cukup kompleks dan membutuhkan model peramalan inflasi yang handal dalam pengambilan keputusan yang akurat. Inflasi di provinsi Sulawesi Selatan seringkali dipengaruhi oleh beberapa faktor seperti perubahan harga komoditas utama, distribusi barang, kondisi pasar regional, dan kebijakan pemerintah daerah. Kabupaten Bulukumba, kabupaten Watampone, kota Makassar, dan kota Palopo merupakan bagian dari wilayah provinsi Sulawesi Selatan. Wilayah tersebut memiliki peran penting dalam mendukung perekonomian wilayah. Agar dapat mengetahui serta memahami dan meramalkan inflasi empat lokasi wilayah tersebut, diperlukan suatu analisis yang mampu menangkap hubungan dinamis serta mampu meramalkan inflasi tiap wilayah.

Menurut Badan Pusat Statistika (BPS) merilis inflasi pada periode setiap bulan atau tahun. Keempat lokasi wilayah tersebut menunjukkan bahwa adanya perubahan harga dari setiap waktu ke waktu. Oleh karena itu, dengan permasalahan tersebut terdapat metode yang cocok digunakan untuk menganalisis data inflasi di wilayah tersebut yaitu analisis time series. Data time series merupakan sebuah kumpulan observasi yang dibuat secara beruntun atau berurut dari waktu ke waktu. Menurut Cryer, et al., (2008) mengemukakan bahwa tujuan dari analisis time series yaitu untuk memprediksi nilai periode yang akan datang berdasarkan nilai periode sebelumnya serta beberapa faktor yang saling berkaitan.

Salah satu pemodelan time series yang pernah digunakan sebelumnya adalah *Vector Autoregressive Moving Average* (VARMA), model ini menjelaskan keterkaitan antar pengamatan pada variabel itu sendiri pada waktu sebelumnya, dan juga keterkaitannya dengan pengamatan pada variabel lain pada waktu sebelumnya. Wutsqa, et.al. (2010) telah melakukan penelitian mengenai perbandingan model VARMA dan model *Vector Autoregressive-Generalized Space Time Autoregressive* (VAR-GSTAR). Hasil penelitiannya menunjukkan bahwa model VAR-GSTAR lebih baik dibandingkan dengan model VARMA. Dimana VARMA memiliki kelemahan yaitu tidak dapat mengakomodasikan lag-lag seasonal pada orde model tersebut. Adapun model *Vector Autoregressive* (VAR) adalah suatu sistem persamaan yang dinamis yang mana pendugaan suatu variabel pada waktu tertentu bergantung pada pergerakan variabel tersebut serta variabel lain yang terlibat dalam waktu sebelumnya. Model VAR merupakan pemodelan time series yang bersifat *multivariate* dan model yang tidak melibatkan faktor lokasi. Keunggulan dari model VAR adalah kemampuannya dalam menangkap hubungan yang dinamis antar variabel, dan fleksibilitas dalam berbagai jenis data. Oleh karena itu, dalam penelitian ini menggunakan salah satu model yaitu model VAR. Sebagaimana penelitian-penelitian yang telah dilakukan oleh Firdos, et al., (2020) mengenai pemodelan dan peramalan pada COVID-19 dengan menggunakan model VAR di Pakistan. Hasil yang diperoleh adalah hasil prediksi menunjukkan jumlah kasus baru 5.363/hari dengan interval

kepercayaan 95% pada tanggal 3 juli. Kematian 167/hari dengan interval kepercayaan 95% dan pemulihan maksimum selama 4.016/hari dengan interval kepercayaan 95% selama 10 hari kedepan. Temuan penelitian ini dapat membantu pemerintah dan lembaga lain untuk membentuk kembali strategi sesuai dengan situasi prediksi tersebut.

Namun, model VAR tersebut terbatas pada faktor lokasi, dimana belum memperhatikan pengaruh faktor lokasi geografisnya. Sehingga, diperlukan metode tambahan yang mampu mengakomodasikan keterkaitan waktu dan lokasi pada data *time series multivariate*. Metode tersebut diperkenalkan oleh Preifer, et.al. (1980) yaitu model *Space Time Autoregressive* (STAR). Akan tetapi, walaupun model tersebut mampu memperhatikan faktor lokasi, namun memiliki kelemahan yaitu model ini hanya mampu diterapkan pada lokasi yang bersifat homogen sehingga parameter waktu dan lokasi untuk semua lokasi konstan. Terlihat di kehidupan nyata, lebih banyak ditemukan lokasi yang cenderung berbeda. Kelemahan tersebut dibenahi oleh Ruchjana, et.al. (2002), dimana diperkenalkan model yaitu model *Generalized Space Time Autoregressive* (GSTAR). Model ini mampu menerapkan data dengan variabel lokasi yang heterogen dengan memberikan bobot setiap lokasi. Terdapat penelitian yang dilakukan oleh Noverlina, et al., (2022) mengenai penerapan model GSTAR untuk peramalan Indeks Harga Konsumen (IHK) di tiga kota Jawa Tengah. Hasil penelitian yang diperoleh menunjukkan bahwa model GSTAR terbaik adalah model GSTAR (1,1,1). Model tersebut memenuhi asumsi homoskedastisitas, *white noise*, dan multivariat normal. Nilai *Mean Absolute Percentage Error* (MAPE) yang diperoleh dengan menggunakan pembobot *Inverse Distance Weighting* (IDW) dan *Normalisasi Cross-Covariance* (NCC) masing-masing adalah 0,2922% dan 0,2914%. Dari hasil tersebut dapat disimpulkan bahwa model GSTAR (1,1,1) terbaik untuk peramalan IHK di tiga lokasi jawa tengah adalah pembobot NCC karena mempunyai nilai MAPE minimum. Kemudian penelitian oleh Feby, et al., (2023) mengenai implementasi model GSTAR untuk prediksi Inflasi. Hasil penelitian menunjukkan bahwa model terbaik untuk memprediksi inflasi adalah model GSTAR(5,1)I(1) dengan menggunakan bobot korelasi silang ternormalisasi yang memiliki nilai *Root Mean Square Error* (RMSE) sebesar 0,5743, nilai RMSE tersebut lebih kecil dibandingkan model GSTAR (2,1).

Selain itu, terdapat metode time series dengan pendekatan *nonlinear* yang dapat digunakan yaitu pendekatan *Support Vector Regression* (SVR). Salah satu penelitian terkait SVR yang telah dilakukan oleh I Nengah, et al., (2022) mengenai peramalan transaksi uang elektronik dengan menggunakan model SVR dan VARMA. Hasil penelitian menunjukkan bahwa hasil peramalan model tersebut dibandingkan dengan menggunakan MAPE. Berdasarkan hasil penelitiannya bahwa model SVR memiliki hasil yang lebih baik dibandingkan dengan model VARMA, dengan nilai MAPE sebesar 3,577%. Hal ini menunjukkan bahwa data peramalan model SVR mendekati data sebenarnya.

Hal yang mendasar dari model SVR yaitu memaksimalkan *hyperplane* untuk memperoleh data-data dari yang menjadi *support vector*. Kelebihan dari model SVR ini adalah mampu menangani atau mengatasi *overfitting*. Sehingga, dalam analisis ekonomi khususnya pada inflasi, metode VAR salah satu metode yang umum digunakan dalam memodelkan hubungan antar variabel dalam sistem multivariat. VAR unggul dalam menganalisis interaksi dinamis antar variabel, tetapi cenderung kurang optimal dalam menangkap pola *nonlinear* yang sering muncul. Oleh karena itu untuk mengatasi keterbatasan tersebut, pendekatan SVR dapat digunakan dalam penelitian ini. Gabungan VAR-SVR memungkinkan integrasi penggabungan kekuatan analisis hubungan dinamis VAR dengan kemampuan SVR dalam memodelkan pola *nonlinear*.

Namun, inflasi di wilayah provinsi Sulawesi Selatan tidak hanya bergantung pada aspek temporal tetapi juga dipengaruhi oleh ketergantungan spasial antar wilayah tersebut. Sehingga mendorong memerlukan penggunaan model *spatio-temporal* misalnya GSTAR, yang mampu menangkap pola hubungan antar lokasi dan waktu secara simultan. Hal ini dapat mengintegrasikan pendekatan SVR kedalam GSTAR, maka model GSTAR-SVR mampu memodelkan kompleksitas *spatio temporal* serta mempertimbangkan pola *nonlinear* dalam data inflasi di wilayah tersebut. Terdapat penelitian yang dilakukan oleh Intan, et al., (2024) tentang peramalan radiasi gelombang pendek untuk tiga lokasi di Provinsi Lampung menggunakan hybrid GSTARX-SVR. Hasil penelitiannya menunjukkan bahwa model GSTARX-SVR terbaik adalah GSTARX(6,1)-SVR dengan bobot spasial invers jarak. Hasil tersebut menunjukkan bahwa dengan model GSTARX-SVR, rentang perkiraan radiasi gelombang pendek matahari adalah 15 dan 35 MJ/m^2 untuk ketiga lokasi tersebut.

Oleh karena itu, penelitian ini menggunakan pendekatan Model VAR-SVR dan GSTAR-SVR dengan menggunakan variabel lag yang signifikan dalam model VAR dan GSTAR sebagai variabel prediktor dalam pendekatan model SVR. Jadi dengan uraian diatas, maka penelitian ini meneliti mengenai “Peramalan Inflasi di Provinsi Sulawesi Selatan dengan Model *Hybrid Vector Autoregressive-Support Vector Regression* dan *Generalized Space Time Autoregressive-Support Vector Regression*. Hasil yang diperoleh nantinya adalah hasil peramalan Inflasi di Provinsi Sulawesi Selatan.

1.2. Rumusan Masalah

Berdasarkan uraian pada latar belakang tersebut, maka rumusan masalah dalam penelitian ini adalah sebagai berikut

1. Bagaimana hasil peramalan Inflasi di Kabupaten Bulukumba, Kabupaten Watampone, Kota Makassar, dan Kota Palopo dengan model Hybrid VAR-SVR dan GSTAR-SVR?

2. Bagaimana akurasi model Hybrid VAR-SVR dan GSTAR-SVR dalam meramalkan Inflasi di Kabupaten Bulukumba, Kabupaten Watampone, Kota Makassar, dan Kota Palopo?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka diperoleh tujuan dalam penelitian ini adalah sebagai berikut

1. Mendapatkan hasil peramalan Inflasi di Kabupaten Bulukumba, Kabupaten Watampone, Kota Makassar, dan Kota Palopo dengan model Hybrid VAR-SVR dan GSTAR-SVR
2. Memperoleh hasil akurasi model Hybrid VAR-SVR dan GSTAR-SVR dalam meramalkan Inflasi di Kabupaten Bulukumba, Kabupaten Watampone, Kota Makassar, dan Kota Palopo

1.4. Batasan Masalah

Adapun batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut

1. Data yang digunakan dalam penelitian ini adalah inflasi di provinsi Sulawesi Selatan dengan membatasi hanya empat lokasi yaitu Kabupaten Bulukumba, Kabupaten Watampone, Kota Makassar, dan Kota Palopo
2. Periode waktu yang digunakan dalam penelitian ini adalah periode waktu bulan Januari 2014 sampai bulan Desember 2023
3. Metode yang digunakan untuk meramalkan data Inflasi dalam penelitian ini adalah model Hybrid VAR-SVR dan GSTAR-SVR

1.5. Manfaat Penelitian

Adapun manfaat yang akan diperoleh dalam penelitian ini adalah sebagai berikut

1. Memberikan tambahan informasi sebagai bahan pertimbangan terkait data inflasi khususnya di Provinsi Sulawesi Selatan pada empat lokasi
2. Menambah kontribusi keilmuan pada bidang statistika dalam menerapkan metode VAR-SVR dan GSTAR-SVR
3. Dapat digunakan untuk menambah bahan referensi penelitian selanjutnya.

1.6. Teori

Pada bagian ini, diuraikan teori yang membahas beberapa sub teori yang mendukung penyelesaian permasalahan yang digunakan dalam penelitian ini adalah sebagai berikut

1.6.1 Time Series Analysis

Time series merupakan suatu himpunan pengamatan data yang terurut dalam waktu, serta metode ini merupakan metode peramalan dengan menggunakan analisis pola hubungan antara peubah yang akan di estimasi dengan peubah waktu, yang mana perlu memperhatikan pola atau tipe datanya. Time series analysis adalah suatu

metode statistika di implementasikan dengan tujuan meramalkan peluang suatu keadaan yang akan terjadi masa yang akan datang dalam bentuk pengambilan keputusan.

Menurut Hanke, et al., (2014) mengemukakan bahwa terdapat empat jenis pola data time series diantaranya yaitu tren, siklis, musiman, dan irregular. Secara umum, terdapat dua jenis pemodelan data analisis time series yaitu univariat dan multivariat. Dimana time series univariat merupakan model peramalan suatu peubah yang hanya didasarkan pada nilai peubah di masa lampau. Time series multivariat merupakan model peramalan dengan menambahkan peubah lain yang memiliki hubungan jangka panjang untuk mendapatkan keakuratan dalam peramalan.

1.6.2 Time Series Multivariat Analysis

Time series multivariate adalah himpunan data yang lebih dari satu atau beberapa variabel yang diambil dari waktu ke waktu, kemudian disajikan secara berurutan menurut waktu kejadian dengan interval waktu yang tetap. Penting dilakukan pengujian dalam membangun model multivariat time series untuk menentukan dan menemukan interaksi yang ada antara suatu variabel time series dengan satu atau lebih variabel lain.

Menurut Cromwell, et al. (1994) mengatakan bahwa prosedur paling sulit dalam membangun model yaitu proses identifikasi model, dimana proses pembentukan modelnya yaitu proses identifikasi, estimasi, dan cek diagnosa. Pada analisis multivariat time series $Z_i(t) = Z_1(t), Z_2(t), \dots, Z_m(t)$ yaitu suatu fenomena yang mana ketergantungannya tidak hanya ada dalam variabel i tetapi juga dalam waktu t , oleh karena itu variabel i dan t saling berkaitan.

1.6.3 Stasioneritas Model

Salah satu asumsi yang harus terpenuhi yaitu asumsi stasioneritas data. Asumsi stasioneritas perlu diperhatikan dan penting dalam analisis time series. Stasioneritas merupakan tidak terdapat perubahan yang signifikan pada data. Pada model multivariate time series, kestasioneran data terbagi menjadi dua yaitu stasioner dalam mean, dan stasioner dalam varian.

Kestasioneran data dalam mean secara visual dapat terlihat melalui plot *Matrix Autocorrelation Function* (MACF). Plot MACF yang turun secara perlahan atau lambat menunjukkan bahwa data belum stasioner dalam mean. Sehingga, perlu dilakukan proses *differencing* data agar data tersebut stasioner. Kemudian untuk menstasionerkan data dalam varian, maka digunakan cara yaitu transformasi *Box-Cox*. Jika nilai λ (λ dapat dilihat dari plot Box-Cox) mendekati atau sama dengan 1, maka data stasioner dalam varian. Adapun salah satu cara untuk melihat kestasioneran data tersebut yaitu dengan uji stasioner dengan uji akar unit (*unit root test*) dengan memperhatikan nilai *Augmented Dickey Fuller* (ADF). Adapun caranya yaitu uji hipotesisnya adalah

$$H_0 : \text{data tidak stasioner}$$

H_1 : data stasioner

Daerah kritis:

$p - value < \alpha$ maka tolak H_0

Ketika telah di uji kestasioneritas datanya dan ternyata masih belum stasioner, maka perlu dilakukan *differencing* data. Adapun kestasioneran data terbagi menjadi dua yaitu stasioner dalam mean, dan stasioner dalam varian berikut

a. Stasioner dalam mean

Dikatakan stasioner dalam mean jika $E(Y_t) = \mu$ artinya nilai mean konstan terhadap waktu. Analisis time series yang tidak stasioner dalam mean artinya dipengaruhi oleh waktu pengamatan yaitu $E(Y_t) = \mu_t$, sehingga dapat diatasi dengan melakukan pembeda atau melakukan differencing. Maka rumus *differencing* yaitu sebagai berikut

$$W_t = Y_t - Y_{t-1} \quad (2)$$

Dimana

W_t = Differencing orde ke-1
 Y_t = Data pengamatan ke- t
 Y_{t-1} = Data pengamatan ke- ($t - 1$)
 t = indeks waktu

b. Stasioner dalam varian

Data time series yang tidak stasioner dalam varian bisa dilihat dari nilai varian Y_t yang akan berubah sejalan dengan perubahan level $var(Y_t) = \sigma_t = cf(\mu_t)$, yang mana konstanta c yang positif dan fungsi f diperlukan transformasi $T(Y_t)$ memiliki variansi konstanta disebut transformasi varians. Secara garis besar, untuk menstabilkan varians maka dilakukan transformasi Box-Cox.

$$T(Y_t) = \frac{Y_t^\lambda - 1}{\lambda} \quad (3)$$

Dimana

$T(Y_t)$ = Transformasi data ke- t
 λ = Nilai koefisien dari transformasi Box-Cox

1.6.4 Model Vector Autoregressive (VAR)

Model VAR salah satu model peramalan *time series* dimana model ini merupakan gabungan beberapa model dari *Autoregressive* (AR), yang mana model ini membentuk suatu vektor yang variabelnya saling mempengaruhi. Menurut enders (2004), model VAR adalah sebuah persamaan dinamis yang pendugaan variabel pada periode tertentu bergantung pada pergerakan variabel tersebut dan variabel lain yang terlibat dalam sistem terhadap periode sebelumnya.

Terdapat asumsi dalam model VAR diantaranya stasioneritas, serta error harus *white noise*. Sehingga, model *VAR* (p) dapat ditulis dalam bentuk rumus berikut

$$Z_t = \tau + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t \quad (4)$$

Keterangan:

Z_t = Vektor Z pada waktu ke- t

a_t = Vektor residual pada waktu ke- t

τ = Vector konstanta

ϕ_1 = Matriks parameter ke-1

ϕ_p = Matriks parameter ke- p

Kemudian persamaan tersebut membentuk suatu operator *back-shift* B berikut

$$(I - \phi B - \dots - \phi_p B^p) \dot{Y}_t = \phi_0 + a_t \quad (5)$$

Dimana I merupakan matriks identitas yang berukuran $m \times m$. Ketika nilai pengamatan disuatu lokasi pada waktu ke- t dipengaruhi oleh nilai pengamatan di satu lag sebelumnya serta pengamatan dilokasi lain maka modelnya membentuk model *VAR*(1). Model *VAR* (1) merupakan model yang berorde 1 yang mana variabel *independen* dari model tersebut hanyalah satu nilai lag dari variabel *dependen*. Sebuah *multivariate time series* dikatakan suatu model VAR proses dari orde 1 ketika mengikuti persamaan berikut

$$Z_t = \phi_1 Z_{t-1} + a_t$$

Kemudian membentuk operator *back-shift* berikut

$$(I - \phi_1 B) Z_t = a_t$$

Keterangan :

Z_t = Vector pengamatan pada waktu ke- t lokasi ke- n yang berukuran $n \times 1$

ϕ_1 = Matriks parameter VAR ordo ke-1 yang berukuran $n \times n$

a_t = Vektor *white noise*, dimana $a_t \sim MN(0, \Sigma)$ yang berukuran $n \times 1$

1.6.5 Model *Generalized Space Time Autoregressive* (GSTAR)

Model GSTAR adalah salah satu model yang digunakan untuk memprediksi data deret waktu yang memiliki hubungan antara waktu dan lokasi. Model tersebut merupakan pengembangan dari model STAR, dan cenderung lebih fleksibel dibandingkan model STAR. Model STAR memiliki kelemahan pada asumsi parameter *autoregressive* dimana memiliki nilai yang sama pada semua lokasi. Oleh karena itu, model STAR lebih cocok digunakan untuk lokasi yang memiliki karakteristik yang sama (*homogen*).

Model ini dapat diterapkan pada lokasi yang memiliki karakteristik *heterogen* yang ditunjukkan dalam bentuk matriks pembobot.

Model GSTAR dengan model beberapa observasi $Z_i(t)$ yang ditemukan disetiap lokasi n dalam ruang $i = 1, 2, \dots, n$, selama periode t . Efek waktu diformulasikan sebagai model deret waktu dan efek spasial diformulasikan sebagai matriks pembobot lokasi. GSTAR yang dimodelkan oleh orde autoregresif (p) dan orde spasial $\lambda_1, \lambda_2, \dots, \lambda_k$ dilambangkan dengan GSTAR (p, λ_k) . Jika data tidak stasioner dalam mean maka lakukan differencing maka model GSTAR yang terbentuk adalah GSTAR $(p, \lambda_k)I(d)$. Sehingga model GSTAR dapat dirumuskan dalam persamaan berikut

$$Z_i(t) = \sum_{k=1}^p \left[\phi_{k0} + \sum_{l=0}^{\lambda_k} \phi_{kl} W^{(l)} \right] Z_i(t-k) + e_i(t) \quad (6)$$

Keterangan :

$Z_i(t)$ = Vektor pengamatan pada waktu t dan lokasi i dengan ukuran $(n \times 1)$

p = Orde *autoregressive*

λ_k = Orde spasial ke- k

N = Jumlah lokasi

$W^{(l)}$ = Matrik pembobot lokasi yang berukuran $(n \times n)$

ϕ_{k0} = Diagonal matriks parameter *autoregressive* pada lag waktu ke- k dan lag spasial ke-0

ϕ_{kl} = Diagonal matriks parameter *autoregressive* pada lag waktu ke- k dan lag spasial ke- l

$e_i(t)$ = Vektor residual berukuran $(N \times 1)$ pada waktu ke- t .

1.6.6 Penentuan Bobot Lokasi Model GSTAR

Penentuan bobot lokasi pada model GSTAR merupakan salah satu tahap permasalahan yang harus dihadapi dalam pemodelan GSTAR. Terdapat beberapa jenis metode yang digunakan dalam menentukan bobot lokasi dalam pemodelan GSTAR yaitu sebagai berikut

a. Bobot Lokasi Seragam

Bobot lokasi seragam memiliki asumsi yaitu lokasi yang digunakan bersifat *homogen* dalam penelitian sehingga memberi nilai bobot yang sama untuk masing-masing lokasi. Bobot lokasi ini sering digunakan pada data yang lokasinya *homogen* dan mempunyai jarak antar lokasi yang sama. Berikut persamaan yang digunakan adalah

$$W_{ij} = \frac{1}{n_i} \quad (7)$$

Dimana

n_i = Banyaknya lokasi yang berdekatan

W_{ij} = Bobot lokasi ke- i dan ke- j

b. Bobot Lokasi Invers Jarak

Bobot lokasi invers jarak dihitung dengan menggunakan jarak sebenarnya antar lokasi. Nilai bobot tersebut memberikan nilai bobot yang besar untuk jarak yang lebih dekat dan memberikan nilai bobot yang kecil dengan jarak yang jauh. Perhitungan bobot pada lokasi ini dengan cara metode invers jarak diperoleh dari normalisasi hasil invers jarak sebenarnya. Berikut persamaan yang digunakan adalah

$$W_{ij} = \frac{\frac{1}{d_{ij}}}{\sum_{j=1}^n \frac{1}{d_{ij}}}, \quad j \neq i \quad (8)$$

Dimana

$$\sum_{i \neq j} W_{ij} = 1$$

W_{ij} = Nilai bobot dari lokasi i dan j

d_{ij} = Jarak dari lokasi i ke j

c. Bobot Lokasi Biner

Bobot lokasi biner hanya bergantung pada kondisi yang terbatas. Berikut persamaan pembobot dengan metode biner memiliki pembobot yaitu

$$W_{ij} = 0 \text{ atau } W_{ij} = 1 \quad (9)$$

Dengan kata lain, bobot ini menggunakan nilai kategori 1 atau 0

d. Bobot Lokasi Normalisasi Korelasi Silang

Pembobot lokasi ini menggunakan hasil normalisasi korelasi silang antar lokasi pada lag yang bersesuaian. Korelasi silang antar lokasi ke- i dan ke- j pada lag waktu ke- k , $\text{corr}[Z_i(t), Z_j(t - k)]$ didefinisikan sebagai berikut

$$\rho_{ij}(k) = \frac{\gamma_{ij}(k)}{\sigma_i \sigma_j}, \quad k = 0, 1, 2, \dots \quad (10)$$

Dimana $\gamma_{ij}(k)$ merupakan kovarians silang antara kejadian dilokasi ke- i dan ke- j . Adapun penaksiran dari lokasi ini pada sampel dapat dihitung melalui persamaan berikut

$$r_{ij}(k) = \frac{\sum_{t=k+1}^T [Y_i(t) - \bar{Y}_i][Y_j(t-k) - \bar{Y}_j]}{\sqrt{\sum_{t=1}^T [Y_i(t) - \bar{Y}_i]^2 \sum_{t=1}^T [Y_j(t) - \bar{Y}_j]^2}} \quad (11)$$

Dimana $Y_i(t)$ merupakan data waktu ke- t pada daerah i . $Y_j(t)$ adalah data waktu ke- t pada daerah j dan k adalah lag waktu ke- k . Kemudian penentuan bobot lokasi dapat dilakukan melalui normalisasi dari hasil besaran-besaran korelasi silang antar lokasi pada waktu yang bersesuaian. Proses ini secara umum menghasilkan bobot lokasi untuk model *GSTAR* (1) dengan persamaan berikut

$$W_{ij} = \frac{r_{ij}(1)}{\sum_{k \neq i} |r_{ik}(1)|} \quad (12)$$

Dimana $i \neq j$ dan $\sum_{i \neq j} |W_{ij}| = 1$.

1.6.7 Support Vector Regression (SVR)

Model SVR bagian dari metode *support vector machine* (SVM) dalam regresi dan prediksi. SVM mempunyai konsep yaitu *e-insentive loss function* yang dapat digeneralisasikan dengan tujuan melakukan pendekatan fungsi yang disebut SVR. Konsep SVM menggunakan *hyperplane* tunggal pada ruang berdimensi banyak yang pada akhirnya partisi tersebut dapat diselesaikan secara nonlinear. Kegunaan SVR untuk menemukan suatu fungsi yang memiliki deviasi paling besar ε dari target aktual y_i . Konsep SVR berdasarkan pada *structural risk minimization* dengan tujuan mengestimasi suatu fungsi dengan meminimalkan batas atas dari generalisasi error. Oleh karena itu, SVR mampu mengatasi *overfitting*. Adapun tujuan dari SVR yaitu untuk memperoleh fungsi dengan tingkat kesalahan paling kecil sehingga menghasilkan suatu prediksi yang baik.

Pada model SVR, dapat menangani kasus regresi dengan menggunakan penerapan konsep yang dikemukakan oleh Vapnik (1995) yaitu *ε -insensitive loss function*. Fungsi regresi dari metode SVR akan sempurna apabila batas deviasinya sama dengan 0, maka dapat dituliskan persamaan berikut

$$f(x) = w^T \varphi(x) + b \quad (13)$$

Koefisien w dan b berfungsi untuk meminimalkan fungsi resiko. Meminimalkan fungsi resiko tersebut akan membuat suatu fungsi menjadi setipis mungkin, sehingga kapasitas fungsi dapat terkontrol, hal ini dinamakan regularisasi. Koefisien w dan b diestimasi dengan cara meminimalkan fungsi resiko (*risk function*) yang didefinisikan dalam persamaan berikut

$$R(f(x)) = \frac{C}{T} \sum_{i=1}^T L_{\varepsilon}(y_i, f(x_i)) + \frac{1}{2} \|w\|^2 \quad (14)$$

Dimana

$$L_{\varepsilon}(y_i, f(x_i)) = \begin{cases} 0 & |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon & \text{lainnya} \end{cases}$$

Dimana L_{ε} merupakan ε -insensitive loss function, y_i merupakan vector dari nilai sebenarnya, C dan ε merupakan *hyperparameter* yang telah ditentukan. Adapun fungsi SVR dapat ditulis pada persamaan berikut.

$$f(x, \beta_i, \beta_i^*) = \sum_{i=1}^n (\beta_i^* - \beta_i) K(x_i, x) + b \quad (15)$$

Terdapat ide yang mendasar dari SVR dengan menentukan beberapa set data yaitu data set *training* dan data set *testing*. Dari data set *training* ditentukan suatu fungsi regresi dengan batasan deviasi sehingga menghasilkan prediksi yang mendekati dari target aktual, sedangkan data set *testing* yaitu data yang akan atau sedang terjadi dan digunakan sebagai bahan uji untuk mengukur sejauh mana model berhasil melakukan prediksi dengan baik.

Kelebihan SVR adalah menghasilkan tren data yang bergelombang mengikuti jalur data yang terbentuk sehingga diharapkan hasil prediksi yang diperoleh lebih akurat. Pada *machine learning* terdapat istilah metode kernel yang digunakan dalam permasalahan nonlinear. Salah satu fungsi kernel yang paling sederhana adalah linear kernel dan yang paling sering digunakan adalah *Radial Basis Function* (RBF). Fungsi liner kernel merupakan fungsi kernel yang biasa digunakan ketika data dapat dipisahkan secara linear, sedangkan fungsi RBF kernel digunakan ketika data tidak terpisah secara linear. Terdapat dua *hyperparameter* dalam pemodelan SVR menggunakan fungsi linear kernel yaitu *epsilon*, *cost*. *Hyperparameter* pada fungsi RBF kernel sama seperti linear kernel hanya saja terdapat satu *hyperparameter* tambahan yaitu *gamma*. Ketiga *hyperparameter* tersebut penting untuk akurasi prediksi. Berikut fungsi kernel yang dapat digunakan dalam metode SVR adalah sebagai berikut

a. Linear

$$K(X, X^T) = X^T X \quad (16)$$

b. Polinomial

$$K(X, X^T) = (X^T X + 1)^p \quad (17)$$

c. *Radial Basis Function* (RBF)

$$K(X_t, X_u) = \exp\left(-\frac{1}{2\sigma^2} \|x_t - x_u\|^2\right) \quad (18)$$

Cara menentukan parameter yang optimal dalam model SVR yaitu menggunakan metode *grid search*. Optimasi *grid search* adalah kombinasi parameter yang diujikan pada suatu model SVR untuk mencari nilai *error* terkecil, dengan tujuan mengidentifikasi parameter optimal dalam data *training* sehingga model tersebut mampu secara akurat memprediksi data *testing*. Proses tersebut dilakukan secara *trial* dan *error* sampai memperoleh parameter optimal dengan RMSE paling kecil.

1.6.8 Algoritma Grid Search

Grid search adalah kombinasi *hyperparameter* yang diujikan pada suatu model SVM untuk mencari nilai error dalam klasifikasi atau prediksi. Salah satu algoritma untuk menentukan *hyperparameter* optimal pada model SVR yaitu dengan menggunakan algoritma *grid search*. Algoritma tersebut membagi jangkauan *hyperparameter* yang akan dioptimalkan kedalam *grid* dan melintasi semua titik untuk mendapatkan *hyperparameter* optimal dalam data *training*. Sehingga model tersebut mampu secara akurat memprediksi data *testing*. Menurut maharesi, (2013) mengemukakan bahwa *hyperparameter* fungsi kernel terbaik diperoleh menggunakan metode *grid search*.

1.6.9 Estimasi Parameter

a. Estimasi Parameter GSTAR

Estimasi parameter yang digunakan adalah *Maximum Likelihood Estimation* (MLE). Metode MLE merupakan pengembangan dari metode *Ordinary Least Square* (OLS) dengan mempertimbangkan residual yang berkorelasi antar persamaan dengan residual yang diperoleh dari estimator dengan OLS. Teknis estimasi parameter dengan *maximum likelihood* dengan meminimumkan jumlah kuadrat sisa (JKS). Misalkan diberikan persamaan regresi berikut

$$Y = X\phi + \varepsilon \quad (19)$$

Adapun asumsi model regresi pada persamaan (19) yaitu error terdistribusi secara normal *multivariate* $e \sim N(X\phi, \Sigma)$ dengan matriks definit positif Σ berukuran $m \times m$ yang fungsi peluang bersamanya adalah variabel random error dapat didefinisikan pada persamaan berikut

$$\iota(Y|\phi, \Sigma) = \prod_{j=1}^m f(\phi, \Sigma|Y) \quad (20)$$

Sehingga diperoleh fungsi *likelihood* berdasarkan persamaan (20), yaitu sebagai berikut

$$\iota(\Phi, \Sigma|Y) = (2\pi)^{-\frac{m}{2}} |\Sigma|^{-\frac{m}{2}} \exp\left(-\frac{1}{2}(Y - X\Phi)^T \Sigma^{-1}(Y - X\Phi)\right) \quad (21)$$

Maka diperoleh fungsi *log-likelihood* berikut

$$\begin{aligned} L(\Phi, \Sigma|Y) &= \ln(\iota(\Phi, \Sigma|Y)) \\ &= \ln(2\pi)^{-\frac{m}{2}} \ln|\Sigma|^{-\frac{m}{2}} \ln(\exp)\left(-\frac{1}{2}(Y - X\Phi)^T \Sigma^{-1}(Y - X\Phi)\right) \\ &= -\frac{m}{2} \ln(2\pi) - \frac{m}{2} \ln|\Sigma| - \frac{1}{2}(Y^T \Sigma^{-1} Y - Y^T \Sigma^{-1} X\Phi - \Phi^T X^T \Sigma^{-1} Y \\ &\quad + \Phi^T X^T \Sigma^{-1} X\Phi) \end{aligned} \quad (22)$$

Selanjutnya, $(\Phi, \Sigma|Y)$ atau persamaan (22) terhadap Φ^T , dapat dijabarkan sebagai berikut

$$\begin{aligned}
 & \frac{\partial(L(\Phi, \Sigma|Y))}{\partial(\Phi^T)} \\
 &= \frac{\partial(-\frac{m}{2}\ln(2\pi) - \frac{m}{2}\ln|\Sigma| - \frac{1}{2}(Y^T\Sigma^{-1}Y - Y^T\Sigma^{-1}X\Phi - \Phi^T X^T\Sigma^{-1}Y + \Phi^T X^T\Sigma^{-1}X\Phi)}{\partial(\Phi^T)} \\
 &= -\frac{1}{2}(-X^T\Sigma^{-1}Y + X^T\Sigma^{-1}X\Phi) \\
 &= \frac{1}{2}X^T\Sigma^{-1}Y - \frac{1}{2}X^T\Sigma^{-1}X\Phi
 \end{aligned} \tag{23}$$

Kemudian meminimumkan $\frac{\partial(L(\Phi, \Sigma|Y))}{\partial(\Phi^T)} = 0$, maka diperoleh

$$\begin{aligned}
 & \frac{1}{2}X^T\Sigma^{-1}Y - \frac{1}{2}X^T\Sigma^{-1}X\Phi = 0 \\
 & -\frac{1}{2}X^T\Sigma^{-1}X\Phi = -\frac{1}{2}X^T\Sigma^{-1}Y \\
 & X^T\Sigma^{-1}X\Phi = X^T\Sigma^{-1}Y \\
 & \Phi = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}Y
 \end{aligned} \tag{24}$$

Sehingga diperoleh estimasi parameter GSTAR adalah sebagai berikut

$$\hat{\Phi}_{ml} = (X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1}Y \tag{25}$$

b. Estimasi Parameter Model VAR

Pada estimasi parameter model VAR, dilakukan dengan menggunakan metode *conditional least square*. Sehingga, persamaan dengan model VAR(p) yang telah dinyatakan stasioner adalah

$$y_t = \delta + \sum_{i=1}^p \Phi_i y_{t-i} + a_t \tag{26}$$

Dimana $y_{t-p+1}, \dots, y_0$ diasumsikan datanya tersedia untuk memudahkan dalam penotasian. Persamaan (26) dapat dinyatakan dalam bentuk umum model linear multivariat berikut

$$Y = XB + A \tag{27}$$

Atau

$$y = (X^T \otimes I_m)\beta + a \tag{28}$$

Dimana

$\otimes$ = Kronecker product

$Y = (y_1, \dots, y_T)_{(m \times T)}$

$B = (\delta, \Phi_1, \dots, \Phi_p)_{(m \times (mp+1))}$

$X = (X_0, \dots, X_t, \dots, X_{T-1})_{((mp+1) \times T)}$

$X_t = \begin{pmatrix} 1 \\ y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}_{((mp+1) \times 1)}$

$A = (a_1, \dots, a_T)_{(m \times T)}$

$y = (vec(Y))_{(mT \times 1)}$

$\beta = (vec(B))_{((m^2p+m) \times 1)}$

$a = (vec(A))_{(mT \times 1)}$

Vec menyatakan operator *stacking* kolom. Sehingga estimasi parameter β dapat dituliskan dalam persamaan berikut:

$$\hat{\beta} = ((X'X)^{-1}X' \otimes I_m)y \quad (29)$$

Sifat konsistensi dan asimtotik normalitas dari penaksir *conditional least square* $\hat{\beta}$ adalah

$$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Gamma_p^{-1} \otimes \Sigma) \quad (30)$$

Dimana $X'X/T$ konvergen dalam probabilitas menuju Γ_p dan $\xrightarrow{d}$ menyatakan konvergen dalam distribusi. Penaksir untuk Σ adalah sebagai berikut

$$\widehat{\Sigma} = (T - (mp + 1))^{-1} \sum_{t=1}^T \hat{a}_t \hat{a}_t' \quad (31)$$

Dimana $\hat{a}_t$ merupakan vector residual.

1.6.10 Uji Linearitas

Uji linearitas digunakan untuk mengetahui apakah memiliki hubungan linear atau tidak serta digunakan sebagai dasar penentuan fungsi kernel pada model SVR. Salah satu uji yang digunakan adalah uji *Ramsey's Reset*. Adapun hipotesis pengujian linearitasnya adalah sebagai berikut

H_0 : Hubungan linear

H_1 : Hubungan nonlinear

Keputusan :

Jika $p - value < \alpha$ maka tolak H_0 artinya hubungan antara variabel memiliki hubungan nonlinear.

1.6.11 Uji Diagnosis Model

Pada pengujian diagnostik model, dilakukan uji kesesuaian model yang terdiri dari uji asumsi multivariat *white noise*, dan distribusi normal multivariat dari residual. Pada pengujian kesesuaian model mencakup kecukupan model yaitu uji apakah residual memenuhi multivariate *white noise*, dan uji asumsi normal multivariat dari residual. Setelah memperoleh estimasi parameter model yang signifikan, model tersebut sebaiknya dilakukan pengecekan apakah asumsi-asumsi model telah terpenuhi.

Asumsi dasar yang harus dipenuhi yaitu error vektor bersifat *white noise*, dan terdistribusi secara normal. Pada model, uji residual (*normalitas*, dan *independensi* (tidak ada efek *autokorelasi*)) harus diperiksa. Secara visual, pada uji multivariat normal dapat dilihat dari plot residual. Dan uji *white noise* juga dapat dilakukan dengan melihat plot MACF dan MPACF dari error.

a. Uji White Noise

Uji *white noise* dilakukan dengan tujuan melihat apakah residual dari model tersebut sudah tidak berkorelasi satu dengan yang lainnya yaitu dengan uji *portmanteau*. Berikut uji hipotesis yang digunakan adalah

$$\begin{aligned} H_0 &: \rho_1 = \rho_2 = \dots = \rho_k = 0 \text{ (residu yang memenuhi asumsi white noise)} \\ H_1 &: \exists \rho_1 \neq 0 \text{ (residu tidak memenuhi asumsi white noise)} \end{aligned}$$

Statistik uji:

$$Q = n(n+2) \sum_{k=1}^n \frac{\hat{\rho}_k^2}{n-k} \quad (32)$$

Dimana

$\hat{\rho}_k^2$ adalah residual *autokorelasi* lag ke- k

Kesimpulan :

Tolak H_0 jika $Q > \chi_{\alpha, df}^2$ atau $p - value < 0.05$ artinya residualnya sudah *white noise*.

b. Uji distribusi normal multivariat

Pemeriksaan uji tersebut dapat dilakukan dengan cara membuat *Q-Q Plot* dari nilai d_f^2 yang mana hipotesisnya dapat diperhatikan sebagai berikut

$$\begin{aligned} H_0 &: \text{residual berdistribusi normal multivariat} \\ H_1 &: \text{residual tidak berdistribusi normal multivariat} \end{aligned}$$

Statistik uji:

$$d_j^2 = (x_j - \bar{x})' \Sigma^{-1} (x_j - \bar{x}), \quad j = 1, 2, \dots, n \quad (33)$$

Keterangan:

x_j = Pengamatan ke- j

Σ^{-1} = Invers matriks varians kovarians

Kesimpulan:

Gagal tolak H_0 jika lebih dari 50% jumlah data memiliki nilai $d_j^2 \leq \chi_{n,0.5}^2$.

1.6.12 Model Vector Autoregressive-Support Vector Regression (VAR-SVR)

Model VAR-SVR merupakan pendekatan hybrid analisis deret waktu multivariat dengan machine learning atau menggabungkan model VAR dan SVR. Pada model VAR dan SVR telah mencapai keberhasilan dalam linear atau nonlinear. Namun, tidak satupun dari kedua model tersebut adalah model universal yang cocok untuk semua keadaan. Pendekatan model VAR untuk kondisi nonlinear kompleks mungkin tidak cukup. Model SVR merupakan salah satu model deret waktu yang mampu mengatasi noise, nonlinearitas, dan overfitting.

Model VAR-SVR adalah model kombinasi antara model linear VAR dengan model nonlinear SVR. Konsep model VAR-SVR yaitu menjadikan variabel yang terbentuk pada model VAR sebagai input dan output pada model SVR. Model VAR-SVR dapat dikatakan layak untuk digunakan jika tidak mengalami overfitting atau underfitting. Overfitting adalah keadaan ketika kinerja model bernilai baik untuk data training, tapi buruk untuk data testing. Underfitting adalah keadaan ketika kinerja model bernilai buruk pada data training maupun testing.

1.6.13 Model Generalized Space Time Autoregressive-Support Vector Regression (GSTAR-SVR)

Model GSTAR-SVR adalah sebuah pendekatan hybrid yang menggabungkan metode statistik GSTAR dengan SVR. Model ini dirancang untuk memprediksi data yang terikat oleh waktu (deret waktu) dan lokasi (spasial), terutama ketika data tersebut memiliki pola nonlinear. Struktur spasio temporal dari GSTAR digunakan sebagai input, lalu hubungan antara input-output dimodelkan secara nonlinear dengan SVR. Artinya, GSTAR menentukan variabel prediktor dan SVR mempelajari fungsi nonlinear.

Model GSTAR memiliki keterbatasan yaitu mengasumsikan bahwa hubungan dalam data bersifat linear. Namun, dalam realitanya yang terdapat pada beberapa karakteristik data, yang mana pola seringkali bersifat nonlinear. Dalam model hybrid ini, GSTAR digunakan untuk menangkap komponen linear dan ketergantungan

spasialnya. Sedangkan SVR digunakan untuk memodelkan residual atau pola rumit yang tidak bisa ditangkap oleh GSTAR.

1.6.14 Model Hybrid

Model hybrid pertama kali diperkenalkan oleh Zhang (2003) dengan tujuan meningkatkan akurasi peramalan. Model hybrid menggabungkan komponen linear dan komponen nonlinear. Berikut persamaannya adalah sebagai berikut

$$Y_t = L_t + N_t \quad (34)$$

Dimana L_t merupakan komponen linear dan N_t merupakan komponen nonlinear. Pertama yang dilakukan adalah memodelkan data pada level satu dengan model linear dan diperoleh residual e_t berikut

$$e_t = Y_t - \hat{L}_t \quad (35)$$

Dimana $\hat{L}_t$ merupakan nilai ramalan model linear pada level satu. Kemudian e_t dimodelkan menggunakan model nonlinear pada level dua dengan p input sebagai berikut

$$e_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-p}) + \varepsilon_t \quad (36)$$

Dimana f merupakan fungsi nonlinear yang ditentukan oleh model nonlinear dan ε_t merupakan kesalahan acak. Adapun hasil ramalan pada persamaan (36) dinotasikan dengan $\hat{N}_t$. Berikut persamaannya adalah sebagai berikut

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t \quad (37)$$

1.6.15 Pemilihan Model Terbaik

Pemilihan model terbaik dapat dilakukan dengan menggunakan nilai *Mean Square Error* (MSE) atau *Root Mean Square Error* (RMSE). RMSE adalah salah satu model ukuran yang digunakan untuk mengukur perbedaan antar nilai yang diramalkan suatu model dengan nilai sebenarnya dari suatu pengamatan. Adapun dalam suatu peramalan akan menghasilkan ramalan yang optimum dengan *error* yang terkecil. Dengan demikian, dapat dirumuskan dalam bentuk sebagai berikut

$$RMSE = \sqrt{MSE}$$

$$\sqrt{MSE} = \sqrt{\frac{1}{T} \sum_{i=1}^T (Y_i - \hat{Y}_i)^2} \quad (38)$$

Dengan T merupakan banyaknya data *training*, Y_i merupakan data *training* ke- i , dan $\hat{Y}_i$ merupakan data hasil ramalan ke- i . Sedangkan RMSE *testing* dapat dirumuskan sebagai berikut

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{T_o} \sum_{j=1}^{T_o} (Y_j - \hat{Y}_j)^2} \quad (39)$$

Keterangan:

T_o = Banyaknya Pengamatan
 Y_j = Data *testing* ke- j
 $\hat{Y}_j$ = Data hasil ramalan ke- j

Sehingga, model peramalan yang menghasilkan nilai RMSE yang paling optimum atau yang paling kecil maka model tersebut yang lebih akurat.

1.6.16 Inflasi

Inflasi merupakan kenaikan harga barang atau jasa secara umum, dan terus menerus dalam suatu perekonomian pada periode waktu tertentu. Inflasi yang tinggi dapat menimbulkan beberapa akibat buruk terhadap individu, masyarakat, ataupun kegiatan perekonomian secara agregat misalnya pada penurunan pendapatan riil dan nilai kekayaan bagi individu dan masyarakat. Inflasi dapat disebabkan oleh beberapa faktor seperti peningkatan permintaan penurunan pasokan, atau faktor ekonomi dan moneter lainnya. Tidak dapat disebut inflasi ketika harga satu atau dua barang saja yang mengalami kenaikan, kecuali kenaikan tersebut meluas atau menyebabkan harga barang lain meningkat secara signifikan. Terdapat beberapa komponen aspek yang harus terpenuhi sehingga dapat dikatakan terjadi inflasi, sebagai berikut

- Kenaikan harga (*Tendency*), yaitu dimana harga suatu komoditas dikatakan naik jika menjadi lebih tinggi daripada harga periode sebelumnya
- Bersifat umum (*General level of price*), artinya kenaikan harga suatu komoditas belum dapat dikatakan inflasi ketika kenaikan tersebut tidak menyebabkan harga-harga secara umum naik
- Berlangsung terus-menerus (*Sustained*), artinya kenaikan harga yang bersifat umum belum akan memunculkan inflasi jika hanya terjadi sesaat. Oleh karena itu, dilakukan dalam jangka waktu minimal bulanan. Maka dalam sebulan akan terlihat apakah kenaikan harga bersifat umum serta terus menerus. Jangka waktu yang lebih panjang yaitu pada triwula dan tahunan.

Kondisi inflasi yang tinggi dan bergejolak menggambarkan kondisi ketidakstabilan dalam aspek ekonomi yang berdampak pada kenaikan harga barang dan jasa secara keseluruhan yang sedang berlangsung. Sehingga, mengakibatkan tingginya angka kemiskinan. Tingkat inflasi yang tinggi dapat menghadirkan tantangan secara signifikan bagi individu dan bisnis dalam membuat keputusan mengenai tingkat konsumsi dan produksi. Dampak yang lebih luas yaitu potensi penurunan pertumbuhan ekonomi. Ketika harga barang atau jasa naik, maka daya beli masyarakat menurun sehingga barang yang ditawarkan tidak banyak yang dibeli. Apabila barang tidak banyak yang dibeli maka jumlah produksi menurun, sehingga

dapat mengurangi tenaga kerja yang akan mengakibatkan pengangguran, maka berdampak pada pertumbuhan ekonomi yang menurun.

Inflasi dipandang sebagai parameter yang penting dalam menganalisis perekonomian, dengan upaya yang terus dilakukan untuk menjaga laju pertumbuhan inflasi pada tingkat yang rendah dan stabil sehingga dapat menghindari munculnya gejala ekonomi makro yang berpotensi merusak stabilitas dalam perekonomian. Adapun jenis-jenis inflasi yaitu sebagai berikut

- a. *Domestik Inflation* artinya inflasi yang berasal dari dalam negeri, yaitu terjadi karena adanya tekanan dari variabel makro dalam negeri sehingga mendorong terjadinya kenaikan harga-harga barang.
- b. *Imported Inflation* artinya inflasi yang berasal dari luar negeri, yaitu inflasi yang terjadi karena adanya pengaruh dari luar negeri. Pengaruh tersebut dapat berupa kejadian inflasi atau kenaikan harga dinegara lain yang mempunyai hubungan erat sehingga harga barang-barang import menjadi lebih mahal. Oleh karena itu, menyebabkan indeks harga konsumen meningkat dan secara tidak langsung akan menaikkan indeks harga konsumen melalui kenaikan biaya produksi.

Menurut Badan Pusat Statistik, mengemukakan bahwa inflasi merupakan persentase kenaikan harga barang dan jasa yang dikonsumsi rumah tangga secara umum. Inflasi dihitung berdasarkan IHK atau *consumer price indeks* (CPI). Inflasi dapat terjadi karena naiknya harga sebagian besar indeks kelompok pengeluaran. Sementara itu, penurunan harga disebut deflasi.

Badan Pusat Statistik Provinsi Sulawesi Selatan merilis dan mempublikasikan bahwa terjadi inflasi di provinsi Sulawesi Selatan pada bulan april 2024 sebesar 2,61 persen dengan IHK sebesar 106,23. Inflasi tertinggi terjadi diwilayah Kab. Sidenreng Rappang sebesar 3,59 persen dengan IHK sebesar 103,20. Inflasi terjadi karena adanya kenaikan harga yang ditunjukkan oleh naiknya indeks kelompok pengeluaran yaitu kelompok makanan, minuman, dan tembakau sebesar 5,82 persen, kelompok pakaian dan alas kaki sebesar 0,33 persen, kelompok perumahan, air, listrik, dan bahan bakar rumah tangga sebesar 0,47 persen. Kelompok perlengkapan, peralatan dan pemeliharaan rutin rumah tangga sebesar 0,89 persen, kelompok kesehatan sebesar 2,15 persen, kelompok transportasi sebesar 1,95 persen, kelompok rekreasi, olahraga, dan budaya sebesar 1,14 persen, kelompok pendidikan sebesar 1,77 persen, kelompok penyediaan makanan dan minuman atau restoran sebesar 1,25 persen, kelompok perawatan pribadi dan jasa lainnya sebesar 3,04 persen. Adapun kelompok informasi, komunikasi, dan jasa keuangan tidak mengalami perubahan indeks.

BAB II

METODE PENELITIAN

2.1. Data

Pada penelitian ini menggunakan data sekunder yang diperoleh dari *website* Badan Pusat Statistik yaitu <https://sulsel.bps.go.id> . Data yang digunakan adalah data Inflasi di empat lokasi yaitu Kabupaten Bulukumba, Kabupaten Watampone, Kota Palopo, dan Kota Makassar. Periode waktu yang digunakan pada data tersebut adalah periode bulan Januari 2014 sampai bulan Desember 2023.

2.2. Variabel Penelitian

Variabel dalam penelitian ini menggunakan data Inflasi di Kabupaten Bulukumba, Kabupaten Watampone, Kota Palopo, dan Kota Makassar. Data tersebut merupakan data yang dipublikasikan oleh Badan Pusat Statistik pada periode bulanan dalam jangka waktu bulan Januari 2014 sampai bulan Desember 2023.

Data inflasi tersebut dibagi menjadi dua kategori yaitu data *training* dan data *testing*. Data *training* digunakan pada periode bulan Januari 2014 sampai Desember 2021 dengan tujuan untuk membentuk suatu model. Sedangkan data *testing* digunakan pada periode bulan Januari 2022 sampai Desember 2023 dengan tujuan untuk mengetahui nilai akurasi dalam peramalan.

2.3. Metode Analisis

Metode yang digunakan dalam penelitian ini adalah VAR-SVR dan GSTAR-SVR. Berikut tahapan analisis yang dilakukan dalam penelitian ini adalah

1. Menyiapkan data yang akan dianalisis dalam penelitian ini adalah data Inflasi di empat lokasi yaitu Kabupaten Bulukumba, Kabupaten Watampone, Kota Palopo, dan Kota Makassar
2. Membagi data Inflasi menjadi dua kategori yaitu data *training* dan data *testing*. Data *training* merupakan data Inflasi pada bulan Januari 2014 sampai bulan desember 2021 di Kabupaten Bulukumba, Kabupaten Watampone, Kota Palopo, dan Kota Makassar, sedangkan data *testing* merupakan data Inflasi pada periode bulan Januari 2022 sampai bulan desember 2023
3. Melakukan analisis statistika deskriptif pada data Inflasi di Kabupaten Bulukumba, Kabupaten Watampone, Kota Palopo, dan Kota Makassar
4. Melakukan uji stasioneritas pada data inflasi di Kabupaten Bulukumba, Kabupaten Watampone, Kota Palopo, dan Kota Makassar
5. Melakukan tahapan analisis pada model VAR-SVR sebagai berikut
 - a. Mengidentifikasi orde model VAR dari data Inflasi pada lokasi Kabupaten Bulukumba, Kabupaten Watampone, Kota Palopo, dan Kota Makassar
 - b. Melakukan estimasi parameter dari model VAR. Kemudian melakukan uji linearitas

- c. Integrasi pada VAR ke SVR. Melakukan pemodelan SVR dengan menjadikan variabel *lag* yang signifikan oleh model VAR sebagai input pada model SVR, kemudian menentukan parameter terbaik dengan menggunakan algoritma *grid search*, dengan tahapan menentukan parameter cost, parameter ε , dan gamma.
 - d. Melakukan evaluasi pemilihan model pada data inflasi dengan menggunakan nilai RMSE
6. Melakukan tahapan analisis pada GSTAR-SVR. Berikut proses analisisnya
- a. Melakukan penentuan pemilihan matriks bobot lokasi
 - b. Mengestimasi parameter dari model GSTAR
 - c. Integrasi GSTAR ke SVR. Melakukan pemodelan SVR dengan menjadikan variabel *lag* yang signifikan dari model GSTAR sebagai input pada model SVR, kemudian menentukan parameter terbaik menggunakan algoritma *grid search*, dengan tahapan melakukan penentuan nilai-nilai parameter yaitu parameter cost, parameter ε dan gamma.
 - d. Melakukan evaluasi pemilihan model pada data inflasi dengan menggunakan nilai RMSE.
7. Melakukan peramalan pada data inflasi dengan menggunakan model terbaik dari model Hybrid VAR-SVR dan GSTAR-SVR.

2.4. Diagram Alir

