

## BAB III

### METODE PENELITIAN

#### 3.1 Metode Penelitian

Metode yang digunakan dalam penelitian ini adalah metode *maximal Overlap Discrete wavelet Transform SARIMA*. Metode MODWT-SARIMA merupakan gabungan antara proses MODWT dan proses SARIMA, dimana proses MODWT sebagai pre-processing data yang akan di modelkan, dari hasil dekomposisi MODWT akan dilakukan pemodelan dengan menggunakan proses SARIMA untuk mencari model terbaik dari setiap hasil dekomposisi MODWT.

#### 3.2 Jenis dan Sumber Data Penelitian

Jenis data yang digunakan untuk penelitian ini adalah data yang diperoleh dari suatu sumber yang sudah tersedia sebelumnya, yaitu data jumlah penumpang kereta api Jabodetabek yang diperoleh dari website resmi BPS ( Badan Pusat Statistik) dengan alamat resmi <https://www.bps.go.id/indicator/17/72/7/jumlah-penumpang-kereta-api.html>. Data diunduh pada tanggal 14 Juni 2021.

#### 3.3 Teknik Analisis data

Langkah awal yang dilakukan adalah menganalisis karakteristik data yang meliputi sifat kestasioneran dan pola data. Adapun penjelasan dari flowchart sebagai berikut.

1. Dekomposisi MODWT, dekomposisi MODWT digunakan sebagai proses *pre-processing* data yang akan dimodelkan. Proses MODWT dilakukan terhadap data dengan menggunakan filter Haar dan 3 level. Proses MODWT dilakukan dengan bantuan *software* MatlabR2018a. Hasil dari proses ini berupa 4 barisan data berupa 3 barisan data koefisien wavelet dan 1 barisan data koefisien skala, secara berturut-turut dinamakan  $D_1$  ,  $D_2$ ,  $D_3$  dan  $S_3$ .

Amalya Fatonah, 2021

**Penerapan Metode *Maximal Overlap Dicrete Wavalet Transform Sarima* Untuk Peramalan Banyaknya Penumpang Kereta Api Jabodetabek (Berdasarkan Data Periode Januari 2015 – April 2021)**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

2. Identifikasi model, proses ini dilakukan terhadap  $D_1$ ,  $D_2$ ,  $D_3$  dan  $S_3$ . Identifikasi secara sederhana dilakukan dengan melihat plot dari masing-masing  $D_1$ ,  $D_2$ ,  $D_3$  dan  $S_3$ . Hal ini bermanfaat untuk mengetahui adanya *trend* atau adanya pengaruh musiman pada masing-masing runtun waktu  $D_1$ ,  $D_2$ ,  $D_3$  dan  $S_3$ . Pengecekan stasioneritas secara formal dilakukan dengan Uji *Unit Root* dengan metode *Augmented Dickey-Fuller Test* (Uji ADF), dengan menguji hipotesis  $H_0$ : terdapat akar unit atau data tidak stasioner dan  $H_1$ : tidak terdapat akar unit atau data stasioner. Hipotesis  $H_0$  ditolak jika nilai statistik uji ADF kurang dari daerah kritisnya dan berlaku sebaliknya. Jika data tidak stasioner, maka biasanya bisa dikonversi menjadi stasioner, dengan proses *differencing* dan transformasi.
3. Estimasi parameter model, analisis overfitting juga dilakukan dalam pemodelan, dengan cara mengkaji dan menganalisis model runtun waktu yang memiliki orde yang lebih tinggi dari pada model yang teridentifikasi. Uji signifikansi juga dilakukan terhadap parameter model terestimasi. Pengujian signifikansi didasarkan pada statistik uji t yang berdistribusi student-t dengan derajat bebas  $n - 1$ ,  $n$ : banyaknya sampel.
4. Cek Diagnostik dan Pemilihan Model Terbaik. Cek diagnostik dimaksudkan untuk memverifikasi kesesuaian model dengan sifat-sifat data. Jika model merupakan model yang tepat maka data yang dihitung dengan model (*fitted value*) akan memiliki sifat-sifat yang mirip dengan data asli. Dengan demikian, residual yang dihitung berdasarkan model yang terestimasi bersifat *white noise*. Sifat *white noise* dari residual model dapat dilihat dari plot FAK dan FAKP residual dan melakukan uji korelasi serial. Plot FAK dari residual model yang bersifat *white noise* ditandai dengan tidak adanya lag  $k \geq 1$  yang keluar dari garis batas interval. Sedangkan uji korelasi serial dilakukan dengan dengan statistik uji Ljung-Box  $Q = n(n+2) \sum_{k=1}^{n-j} \hat{\rho}(j)^2$ , yang akan berdistribusi  $\chi^2(k - (p + q))$ ,  $k > (p + q)$  dimana  $\hat{\rho}(j)$  menunjukkan nilai sampel FAK residual pada lag- $j$ , sedangkan  $p$  dan  $q$  menunjukkan orde dari model sarima. Hipotesis uji korelasi serial ini adalah  $H_0: \rho_1 = \rho_2 = \dots = \rho_k, k < n$ , dimana

Amalya Fatonah, 2021

**Penerapan Metode *Maximal Overlap Discrete Wavelet Transform Sarima* Untuk Peramalan Banyaknya Penumpang Kereta Api Jabodetabek (Berdasarkan Data Periode Januari 2015 – April 2021)**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

tidak terdapat korelasi serial dalam residual sampai lag- $k$ ,  $k < n$ . Hipotesis ini diterima apabila nilai  $p$ -value dari statistik uji Ljung-Box di atas garis batas 5% dan berlaku sebaliknya. Pemilihan model terbaik didasarkan pada nilai *Root Mean Square Error* (RMSE) yang paling kecil mendekati nol.

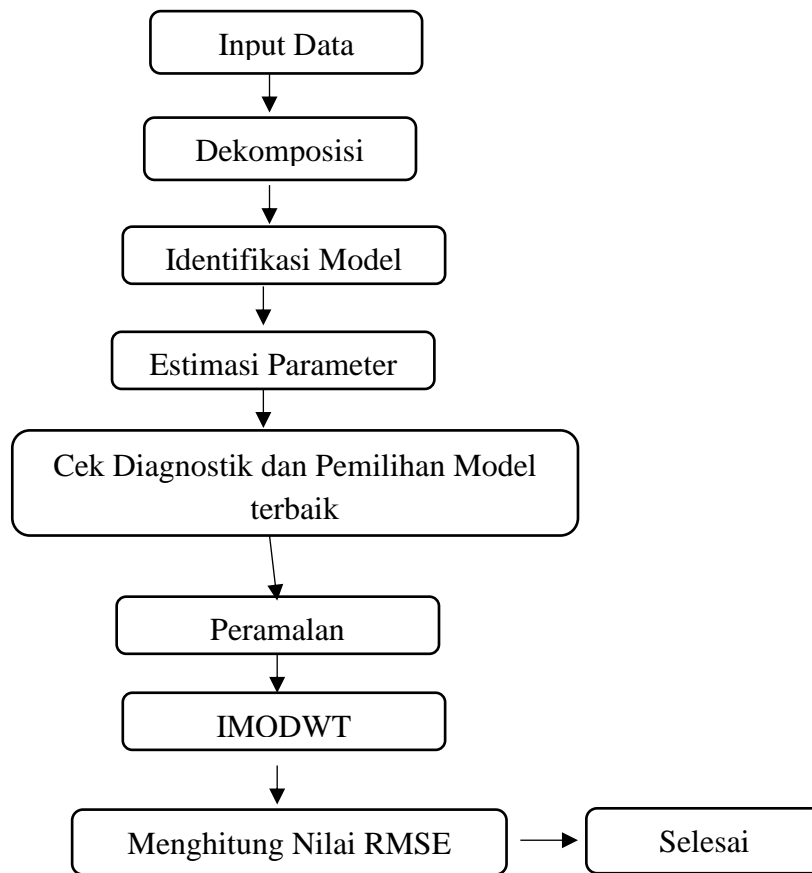
5. Peramalan, proses peramalan dilakukan terhadap masing-masing  $D_1$ ,  $D_2$ ,  $D_3$  dan  $S_3$  dengan menggunakan model terbaik untuk masing-masing runtun waktu tersebut.
6. IMODWT, *Invers Maximal Overlap Discrete Wavelet Transform* dilakukan untuk merekonstruksi data asli ( $Z_t$ ) dari data hasil dekomposisi  $D_1$ ,  $D_2$ ,  $D_3$  dan  $S_3$ . Proses ini dilakukan dengan filter yang ketika melakukan proses MODWT.
7. Menghitung RMSE, proses ini dimaksudkan untuk mengetahui keakuratan hasil peramalan dengan model tersebut.
8. Akhir dari proses analisis data ini adalah diperolehnya nilai hasil ramalan dan nilai RMSE nya.

Proses dekomposisi MODWT dan IMODWT dalam penelitian ini dilakukan dengan bantuan *software* MatlabR2018a sedangkan proses pemodelan Sarima dilakukan dengan bantuan *software* R-3.5.3.

**Amalya Fatonah, 2021**

**Penerapan Metode *Maximal Overlap Discrete Wavelet Transform* Sarima Untuk Peramalan Banyaknya Penumpang Kereta Api Jabodetabek (Berdasarkan Data Periode Januari 2015 – April 2021)**

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu



**Gambar 3.1** *flowchart* Analisis MODWT-SARIMA