

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Hasil Penelitian

##### 4.1.1 *Business Understanding*

Tahap pertama dalam proses penelitian ini adalah *Business Understanding*. Pada tahap ini dilakukan pemahaman terhadap permasalahan yang menjadi dasar penelitian, tujuan yang ingin dicapai, serta konteks penerapan metode yang digunakan.

##### 4.1.1.1 Identifikasi Masalah

Identifikasi masalah dalam penelitian ini berfokus pada keterbatasan data *General Circulation Models* (GCM) yang memiliki resolusi spasial rendah dan tingkat bias yang tinggi terhadap data observasi, sehingga memerlukan proses *downscaling*. Meskipun penelitian terdahulu menunjukkan bahwa *Super Resolution Generative Adversarial Network* (SRGAN) memiliki performa unggul dibandingkan metode *deep learning* lain dalam tugas *downscaling*, sebagian besar studi masih menggunakan data *low resolution* sintetis hasil *downsampling*. Oleh karena itu, terdapat ketidakpastian apakah performa SRGAN tetap optimal ketika diimplementasikan pada data *low resolution* asli dari GCM.

##### 4.1.1.2 Tujuan Penelitian

Berdasarkan permasalahan yang telah diidentifikasi, penelitian ini bertujuan untuk mengimplementasikan algoritma SRGAN untuk melakukan *statistical downscaling* pada data curah hujan *Coupled Model Intercomparison Project Phase 6* (CMIP6) di wilayah Jawa Barat. Penelitian ini bertujuan mengurangi keterbatasan resolusi spasial dan bias pada data GCM dengan mentransformasikan data beresolusi rendah (*low resolution*) menjadi data beresolusi tinggi (*high resolution*) yang setara dengan *Climate Hazards Group InfraRed Precipitation with Station data Version 2* (CHIRPS2), sehingga dapat dimanfaatkan untuk kebutuhan mitigasi bencana hidrometeorologi di tingkat regional.

#### 4.1.1.3 Studi Kebutuhan Data & Lingkup Wilayah

Jawa Barat merupakan salah satu wilayah dengan frekuensi bencana hidrometeorologi tertinggi di Indonesia, sehingga kebutuhan akan data proyeksi iklim resolusi tinggi penting untuk menunjang akurasi strategi mitigasi bencana di tingkat daerah.

Situasi geografis Jawa Barat yang memiliki banyak pegunungan memerlukan variabel tambahan selain data curah hujan yakni data topografi. Oleh karena itu, penelitian ini menetapkan kebutuhan tiga dataset utama:

1. CMIP6: Sebagai input utama sirkulasi iklim global *low resolution*.
2. CHIRPS2: Sebagai referensi observasi atau *ground truth*.
3. *Digital Elevation Model Nasional* (DEMNAS): Sebagai variabel prediktor topografi untuk membantu model memahami efek orografis pada curah hujan lokal.

#### 4.1.1.4 Rencana Eksperimen

Penelitian dilakukan melalui beberapa tahapan utama, yaitu:

1. *Business Understanding*: Tahap awal dimulai dengan Identifikasi Masalah terkait rendahnya resolusi data GCM, penetapan Tujuan Penelitian, Studi Kebutuhan Data & Lingkup Wilayah di Jawa Barat, serta penyusunan Rencana Penelitian.
2. *Data Understanding*: Melakukan Pengumpulan Dataset yang terdiri dari data curah hujan CMIP6 (input), data observasi CHIRPS (target), dan data elevasi DEMNAS. Selanjutnya, dilakukan Eksplorasi Dataset untuk memahami karakteristik awal data tersebut.
3. *Data Preparation*: Tahap ini melibatkan rangkaian proses teknis yang intensif, meliputi Penggabungan Data, *Crop Area* Jawa Barat, *Regridding* untuk penyesuaian grid, Konversi Satuan CMIP6, serta Penanganan Nilai Null & Sinkronisasi Waktu. Data kemudian diproses melalui Normalisasi *Min Max Scaling* sebelum dilakukan *Split Data Train/Val*.
4. *Modelling*: Mengimplementasikan Arsitektur SRGAN untuk melatih model dalam merekonstruksi detail curah hujan beresolusi tinggi. Proses ini diakhiri dengan Simpan Model Terbaik berdasarkan performa selama pelatihan.

5. *Evaluation*: Tahap akhir untuk melakukan Evaluasi Model menggunakan metrik statistik (seperti RMSE dan korelasi) dan Analisis Hasil guna memastikan bahwa data *downscaling* yang dihasilkan telah akurat dan representatif sebelum penelitian dinyatakan selesai (*End*).

#### 4.1.2 *Data Understanding*

Setelah tahap *Business Understanding* dilakukan untuk memahami permasalahan dan tujuan penelitian, langkah selanjutnya adalah tahap *Data Understanding*. Tahap ini bertujuan untuk memahami karakteristik data yang digunakan dalam penelitian, termasuk sumber data, struktur data, resolusi spasial dan temporal, serta kualitas data yang tersedia. Pada tahap ini dilakukan proses identifikasi dan eksplorasi dataset yang digunakan, yaitu data curah hujan observasi CHIRPS2 sebagai data target beresolusi tinggi, data proyeksi iklim global CMIP6 sebagai data input beresolusi rendah, serta data topografi DEMNAS sebagai variabel prediktor tambahan. Pemahaman terhadap karakteristik dataset ini penting untuk memastikan bahwa data yang digunakan sesuai dengan kebutuhan pemodelan *statistical downscaling* yang dilakukan dalam penelitian ini

##### 4.1.2.1 Pengumpulan Dataset

Dataset yang dikumpulkan untuk kebutuhan penelitian ini adalah:

1. CMIP6

Data curah hujan harian diperoleh melalui portal *Copernicus Climate Data Store* (CDS) yang menyediakan akses resmi terhadap dataset CMIP6. Konfigurasi pengunduhan data ditunjukkan pada Tabel 4.1.

Tabel 4.1 Konfigurasi Unduhan Data CMIP6

| NO | Temp Solution | Experiment | Variable      | Model        | Year      | Month         |
|----|---------------|------------|---------------|--------------|-----------|---------------|
| 1  | Daily         | Historical | Precipitation | CNRM CM6 1   | 2000 2014 | Select<br>All |
| 2  | Daily         | Historical | Precipitation | CanESM5      | 2000 2014 | Select<br>All |
| 3  | Daily         | Historical | Precipitation | IPSL CM6A LR | 2000 2014 | Select<br>All |

Berdasarkan konfigurasi yang ditunjukkan pada tabel 4.1, Pengumpulan data difokuskan pada variabel *precipitation* dengan resolusi temporal harian (*daily temporal resolution*). Pemilihan model dilakukan pada tiga *General Circulation Models (GCM)*, yaitu:

- a. CNRM CM6 1
- b. CanESM5
- c. IPSL CM6A LR

Pemilihan tiga model tersebut didasarkan pada ketersediaan data harian yang lengkap pada skenario Historical serta kemudahan akses melalui portal Copernicus Climate Data Store. Data yang digunakan mencakup skenario *Historical* dengan rentang waktu 2000–2014. Seluruh dataset dapat diakses secara resmi melalui portal Copernicus Climate Data Store pada tautan : [https://cds.climate.copernicus.eu/datasets/projections\\_cmip6?tab=download](https://cds.climate.copernicus.eu/datasets/projections_cmip6?tab=download)

## 2. CHIRPS2

Data curah hujan observasi diperoleh dari dataset Climate Hazards Group InfraRed Precipitation with Station data Version 2 (CHIRPS2) yang dikembangkan oleh Climate Hazards Center (CHC), University of California, Santa Barbara (UCSB).

Dataset CHIRPS2 menyediakan data curah hujan dengan resolusi spasial  $0,05^\circ$  dan resolusi temporal harian. Data diunduh dalam format tahunan sesuai dengan rentang waktu 2000–2014.

Seluruh dataset dapat diakses melalui portal resmi Climate Hazards Center pada tautan: [https://data.chc.ucsb.edu/products/CHIRPS\\_2.0/global\\_daily/netcdf/p05/](https://data.chc.ucsb.edu/products/CHIRPS_2.0/global_daily/netcdf/p05/)

## 3. DEMNAS

Data topografi diperoleh dari Digital Elevation Model Nasional (DEMNAS) yang disediakan oleh Badan Informasi Geospasial (BIG). Dataset DEMNAS menyediakan informasi elevasi permukaan tanah dalam format raster dengan resolusi spasial tinggi.

Dataset DEMNAS dapat diakses melalui portal resmi Badan Informasi Geospasial pada tautan : <https://tanahair.indonesia.go.id/portal/web/login?page=/unduh/demnas>

Setelah proses pengunduhan selesai, seluruh dataset disimpan dalam Google Drive dengan struktur direktori yang sistematis untuk mendukung proses pemrosesan selanjutnya. Akses terhadap dataset penelitian dapat diperoleh melalui tautan berikut: [https://drive.google.com/drive/u/4/folders/1J1aM6Wn4zkgmZA14bEk6mSIT0VW\\_vdC?usp=drive\\_link](https://drive.google.com/drive/u/4/folders/1J1aM6Wn4zkgmZA14bEk6mSIT0VW_vdC?usp=drive_link)

#### 4.1.2.2 Eksplorasi Dataset

Eksplorasi dataset dilakukan untuk memahami karakteristik awal data yang digunakan dalam penelitian sebelum memasuki tahap pemodelan. Pada tahap ini dilakukan analisis terhadap struktur data, resolusi spasial, dimensi data, serta periode waktu dari masing masing dataset yang digunakan. Informasi ini penting untuk memastikan kesesuaian antara data input, data target, serta variabel tambahan yang digunakan dalam proses *statistical downscaling*. Ringkasan struktur dataset yang digunakan dalam penelitian ini ditunjukkan pada Tabel 4.2.

Tabel 4.2 Struktur Dataset

| Dataset      | Variabel  | Dimensi          | Grid Size        | Resolusi                     | Lat Range     | Lon Range       | Periode                 | Jumlah Hari | Format  |
|--------------|-----------|------------------|------------------|------------------------------|---------------|-----------------|-------------------------|-------------|---------|
| CHIRPS2      | precip    | time × lat × lon | 2000 × 7200      | 0.05° × 0.05° (~5.5 km)      | 49.97 — 49.97 | 179.98 — 179.98 | 2000 01 01 — 2014 12 31 | 5479        | NetCDF  |
| CNRM CM6 1   | pr        | time × lat × lon | 128 × 256        | 1.40° × 1.41° (~156 km)      | 88.93 — 88.93 | 0.00 — 358.59   | 2000 01 01 — 2014 12 31 | 5479        | NetCDF  |
| CanESM5      | pr        | time × lat × lon | 64 × 128         | 2.79° × 2.81° (~312 km)      | 87.86 — 87.86 | 0.00 — 357.19   | 2000 01 01 — 2014 12 31 | 5475        | NetCDF  |
| IPSL CM6A LR | pr        | time × lat × lon | 143 × 144        | 1.27° × 2.50° (~278 km)      | 90.00 — 90.00 | 0.00 — 357.50   | 2000 01 01 — 2014 12 31 | 5479        | NetCDF  |
| DEMNAS       | Elevation | lat × lon        | 3333 × 3333/tile | 0.000075° × 0.000075° (~8 m) | 7.8 — 5.8     | 106.0 — 108.8   | Statis                  |             | GeoTIFF |

Tabel 4.3 Karakteristik Dataset di Wilayah Jawa Barat

| Dataset      | Grid             | Total Pixel   | NaN / Null Pixel   | Valid Pixel   | Min | Max     | Std Dev | Satuan                            |
|--------------|------------------|---------------|--------------------|---------------|-----|---------|---------|-----------------------------------|
| CHIRPS2      | 40 x 56          | 12272960      | 3,758,594 (30.63%) | 8514366       | 0   | 331,905 | 11,52   | mm/day                            |
| CNRM CM6 1   | 2 x 2            | 21916         | 0 (0.00%)          | 21916         | 0   | 0,001   | 0       | kg m <sup>2</sup> s <sup>-1</sup> |
| CanESM5      | 1 x 1            | 5475          | 0 (0.00%)          | 5475          | 0   | 0,002   | 0       | kg m <sup>2</sup> s <sup>-1</sup> |
| IPSL CM6A LR | 2 x 1            | 10958         | 0 (0.00%)          | 10958         | 0   | 0,003   | 0       | kg m <sup>2</sup> s <sup>-1</sup> |
| DEMNAS       | 3333 x 3333/tile | 11108889/tile | 0                  | 11108889/tile | 6,4 | 3032.8  | 40,6    | meter                             |

Dataset CHIRPS2 memiliki dimensi  $time \times latitude \times longitude$  dengan ukuran grid global  $2000 \times 7200$  dan resolusi spasial sekitar  $0.05^\circ \times 0.05^\circ$  (~5.5 km). Data ini mencakup periode 1 Januari 2000 hingga 31 Desember 2014 dengan total 5479 hari dan disimpan dalam format NetCDF. Resolusi spasial yang relatif tinggi membuat dataset ini mampu merepresentasikan variasi curah hujan dengan lebih detail pada skala regional.

Dataset proyeksi iklim dari CMIP6 terdiri dari beberapa model iklim global dengan resolusi spasial yang lebih kasar dibandingkan CHIRPS2. Model CNRM CM6 1 memiliki resolusi sekitar  $1.40^\circ \times 1.41^\circ$  (~156 km) dengan grid sebanyak  $128 \times 256$ , model CanESM5 memiliki resolusi  $2.79^\circ \times 2.81^\circ$  (~312 km) dengan grid sebanyak  $64 \times 128$ , sedangkan model IPSL CM6A LR memiliki resolusi sekitar  $1.27^\circ \times 2.50^\circ$  (~278 km) dengan grid sebanyak  $143 \times 144$ . Seluruh model memiliki dimensi data  $time \times latitude \times longitude$ , periode waktu yang sama dengan CHIRPS2, serta disimpan dalam format NetCDF. Perbedaan resolusi spasial antara dataset CHIRPS dan CMIP6 ini menunjukkan adanya kesenjangan skala data yang cukup signifikan.

DEMNAS memiliki dimensi  $latitude \times longitude$  dengan resolusi spasial sangat tinggi, yaitu sekitar  $0.000075^\circ$  (~8 meter) dan ukuran grid  $3333 \times 3333$  pada setiap tile. Data ini disimpan dalam format GeoTIFF dan menggambarkan variasi elevasi wilayah yang dapat mempengaruhi distribusi curah hujan pada skala lokal.

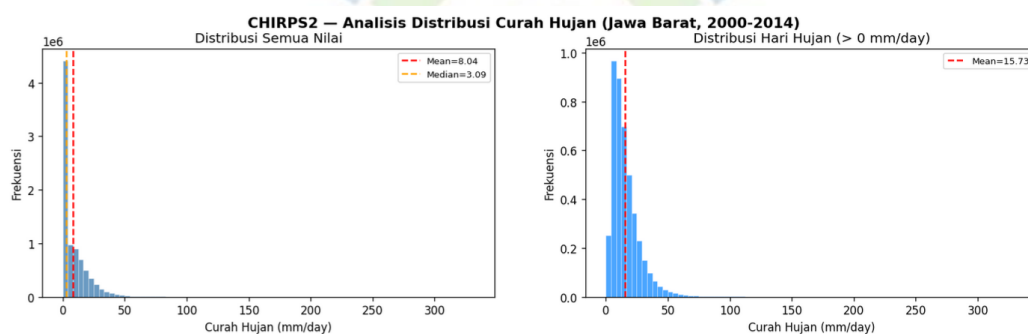
Selanjutnya, Tabel 4.3 menampilkan karakteristik dataset setelah dilakukan proses *subset* pada wilayah Jawa Barat. Pada wilayah ini, data CHIRPS2 memiliki grid sebesar  $40 \times 56$  dengan total 12.272.960 piksel, di mana sekitar 30.63% merupakan nilai NaN. Nilai NaN tersebut merupakan daerah lautan. Statistik distribusi curah hujan menunjukkan bahwa curah hujan maksimum mencapai 331.905 mm/hari, dengan nilai rata-rata 8.044 mm/hari dan standar deviasi 11.52 mm/hari, yang menunjukkan adanya variasi curah hujan yang cukup tinggi di wilayah Jawa Barat.

Sebaliknya, data CMIP6 pada wilayah Jawa Barat memiliki jumlah grid yang jauh lebih sedikit akibat resolusi spasial yang lebih kasar. Model CNRM CM6 1 memiliki grid  $2 \times 2$ , model CanESM5 memiliki grid  $1 \times 1$ , sedangkan model IPSL CM6A LR memiliki grid  $2 \times 1$ . Tidak ditemukan nilai NaN pada dataset CMIP6

dalam wilayah tersebut. Nilai curah hujan yang dihasilkan relatif kecil karena masih berada dalam satuan  $\text{kg m}^{-2} \text{s}^{-1}$ , yang berbeda dengan satuan data observasi CHIRPS.

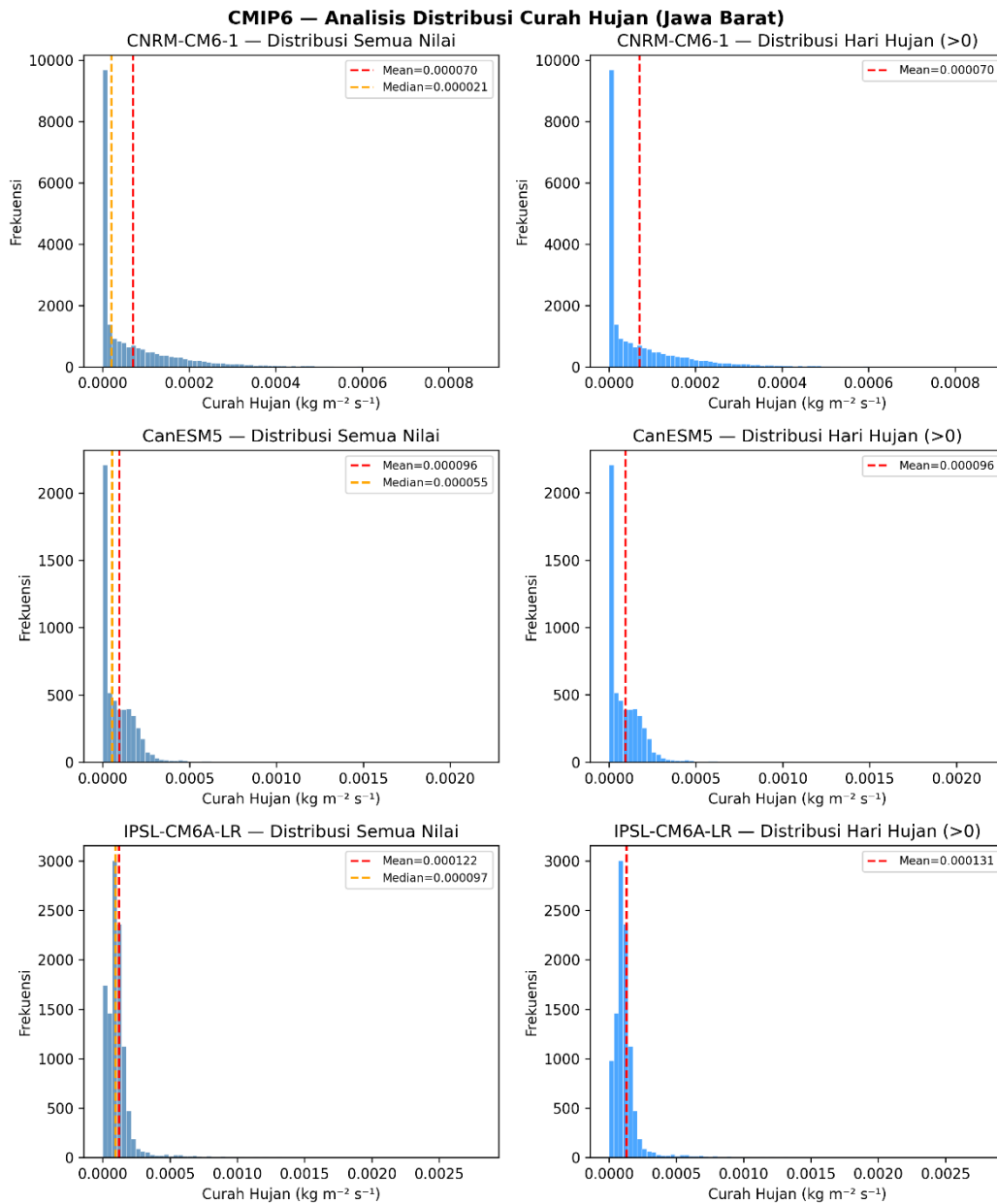
Data DEMNAS pada wilayah studi memiliki resolusi spasial yang sangat tinggi dengan ukuran tile  $3333 \times 3333$  piksel dan total sekitar 11.108.889 piksel. Nilai elevasi minimum tercatat sebesar  $-6.3$  meter, sedangkan nilai maksimum mencapai 3032.8 dengan rata-rata elevasi 12.4 meter dan standar deviasi 40.6 meter. Variasi elevasi ini menggambarkan kondisi topografi wilayah Jawa Barat yang cukup beragam, mulai dari dataran rendah hingga daerah dataran tinggi seperti perbukitan dan pegunungan.

Untuk memahami pola distribusi curah hujan secara lebih rinci, dilakukan analisis distribusi nilai curah hujan menggunakan histogram pada data observasi dan data model iklim. Untuk melihat pola penyebaran nilai curah hujan pada data observasi dan data model iklim.



Gambar 4.1 Histogram Distribusi Curah Hujan CHIRPS Jawa Barat

Gambar 4.1 menampilkan distribusi curah hujan berdasarkan data CHIRPS2 pada periode 2000–2014. Histogram pertama menunjukkan distribusi seluruh nilai curah hujan, sedangkan histogram kedua hanya menampilkan distribusi pada hari hujan (*rainy days*) dengan nilai curah hujan lebih besar dari 0 mm/hari. Berdasarkan grafik tersebut, distribusi curah hujan CHIRPS menunjukkan pola right skewed distribution, di mana sebagian besar nilai curah hujan berada pada kisaran rendah dengan frekuensi yang tinggi, sementara kejadian curah hujan sangat tinggi memiliki frekuensi yang relatif kecil.



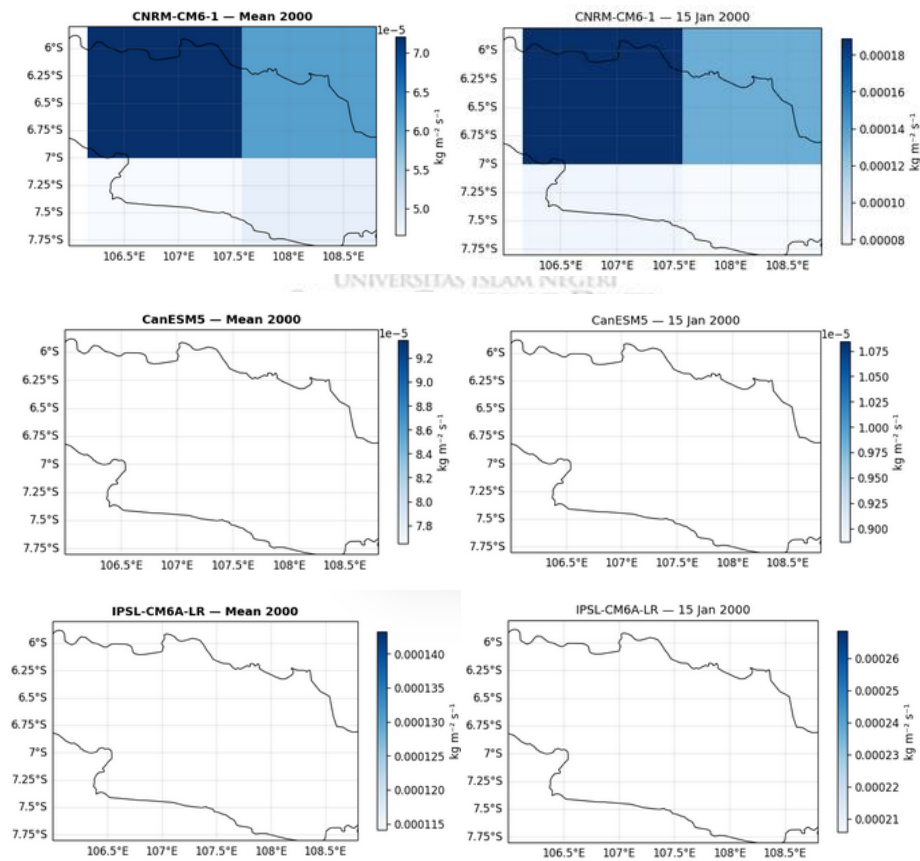
Gambar 4.2 Histogram Distribusi Curah Hujan CMIP6 Jawa Barat

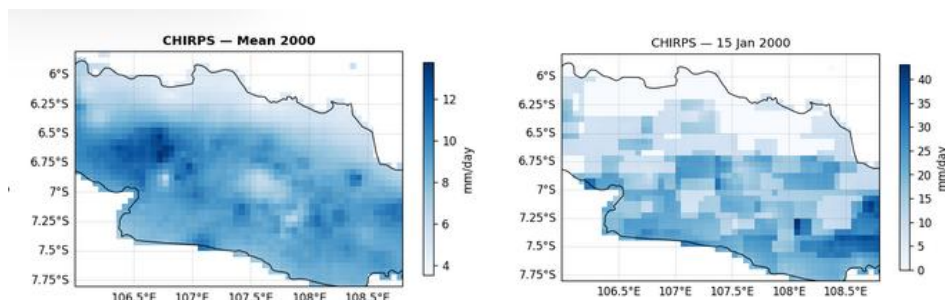
Gambar 4.2 menunjukkan distribusi curah hujan dari tiga model iklim global CMIP6, yaitu CNRM CM6 1, CanESM5, dan IPSL CM6A LR. Setiap model ditampilkan dalam dua histogram, yaitu distribusi seluruh nilai dan distribusi hari hujan. Secara umum, distribusi curah hujan dari ketiga model menunjukkan pola yang mirip dengan data observasi, yaitu distribusi yang condong ke kanan dengan sebagian besar nilai berada pada kisaran curah hujan rendah. Namun, nilai curah hujan yang dihasilkan oleh model CMIP6 berada pada skala yang jauh lebih kecil

karena masih menggunakan satuan  $\text{kg m}^{-2} \text{s}^{-1}$ , sehingga nilai rata-ratanya berada pada orde  $10^{-4}$ .

Selain itu, variasi distribusi antar model juga terlihat cukup berbeda. Model CNRM CM6 1 menunjukkan nilai rata-rata sekitar  $7.0 \times 10^{-5} \text{ kg m}^{-2} \text{ s}^{-1}$ , sedangkan model CanESM5 memiliki nilai rata-rata sekitar  $9.6 \times 10^{-5} \text{ kg m}^{-2} \text{ s}^{-1}$ . Model IPSL CM6A LR menunjukkan nilai rata-rata yang sedikit lebih tinggi, yaitu sekitar  $1.2 \times 10^{-4} \text{ kg m}^{-2} \text{ s}^{-1}$ . Perbedaan distribusi ini mencerminkan variasi karakteristik simulasi presipitasi antar model iklim global.

Secara umum, histogram distribusi curah hujan menunjukkan bahwa data observasi dan data model memiliki pola distribusi yang tidak simetris dengan dominasi nilai curah hujan rendah. Perbedaan skala nilai dan resolusi spasial antara data observasi dan model iklim global menunjukkan adanya kesenjangan karakteristik data yang perlu diperhatikan dalam analisis curah hujan pada skala regional.





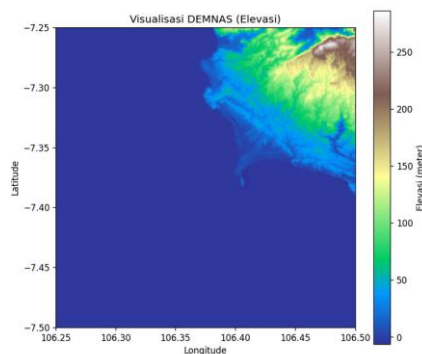
Gambar 4.3 Visualisasi spasial CHIRPS dan CMIP6 Jawa Barat

Distribusi spasial curah hujan pada wilayah Jawa Barat ditampilkan pada Gambar 4.3, yang memperlihatkan perbandingan antara data proyeksi curah hujan dari model CMIP6 (CNRM CM6 1, CanESM5, dan IPSL CM6A LR) dengan data observasi CHIRPS. Visualisasi tersebut menampilkan dua jenis informasi, yaitu distribusi rata-rata curah hujan harian selama tahun 2000 serta contoh snapshot curah hujan pada tanggal 15 Januari 2000.

Pada baris pertama ditampilkan peta rata-rata curah hujan tahunan, sedangkan baris kedua menunjukkan kondisi curah hujan pada salah satu hari pengamatan. Berdasarkan visualisasi tersebut terlihat bahwa distribusi curah hujan pada wilayah Jawa Barat memiliki variasi spasial yang cukup jelas. Data CHIRPS menunjukkan pola curah hujan yang lebih detail dengan variasi intensitas yang berbeda di setiap lokasi, sedangkan data CMIP6 terlihat lebih homogen karena resolusi spasialnya lebih kasar.

Selain itu, pola distribusi curah hujan pada data observasi menunjukkan bahwa beberapa wilayah memiliki intensitas curah hujan yang lebih tinggi dibandingkan wilayah lainnya. Variasi spasial ini mengindikasikan bahwa distribusi curah hujan tidak merata pada seluruh wilayah penelitian dan dapat dipengaruhi oleh berbagai faktor geografis, seperti kondisi topografi dan dinamika atmosfer regional.

Perbedaan tingkat detail antara data CHIRPS dan CMIP6 juga terlihat jelas pada visualisasi ini. Data CHIRPS mampu menampilkan variasi curah hujan yang lebih kompleks pada skala lokal, sedangkan data CMIP6 hanya merepresentasikan kondisi curah hujan secara lebih umum akibat keterbatasan resolusi grid. Perbedaan karakteristik spasial tersebut menunjukkan adanya kesenjangan resolusi antara data observasi dan model iklim global.



Gambar 4.4 Visualisasi 1 Tile DEMNAS

Visualisasi topografi wilayah ditampilkan pada Gambar 4.4 yang menunjukkan distribusi elevasi berdasarkan data DEMNAS. Peta elevasi tersebut menggambarkan variasi ketinggian permukaan wilayah dalam satuan meter terhadap permukaan laut. Skala warna pada peta menunjukkan perbedaan ketinggian, di mana warna biru merepresentasikan wilayah dengan elevasi rendah, sedangkan warna hijau hingga coklat menunjukkan wilayah dengan elevasi yang lebih tinggi.

Visualisasi pada Gambar 4.4 hanya menampilkan satu tile DEMNAS sebagai contoh representasi data topografi. Setiap tile DEMNAS memiliki ukuran grid sekitar  $3333 \times 3333$  piksel dengan resolusi spasial sekitar 8 meter. Untuk mencakup seluruh wilayah Jawa Barat, diperlukan sekitar 83 tile DEMNAS yang kemudian digabungkan (*mosaic*) sehingga membentuk satu peta topografi yang utuh.

Berdasarkan hasil eksplorasi dataset, terlihat adanya perbedaan resolusi spasial yang signifikan antara data observasi CHIRPS dan data model iklim CMIP6. Selain itu, distribusi curah hujan menunjukkan pola yang tidak simetris dengan dominasi curah hujan intensitas rendah. Perbedaan karakteristik ini menunjukkan perlunya proses penyesuaian resolusi spasial agar data model iklim dapat merepresentasikan kondisi curah hujan pada skala regional dengan lebih baik.

#### 4.1.3 *Data Preparation*

Berdasarkan hasil eksplorasi dataset, terdapat perbedaan karakteristik antar data seperti resolusi spasial, struktur grid, dan satuan data sehingga belum dapat langsung digunakan dalam proses pemodelan. Oleh karena itu dilakukan

tahap *data preparation* untuk menyesuaikan struktur dan karakteristik data agar konsisten dan dapat digunakan dalam proses *statistical downscaling*. Tahapan yang dilakukan meliputi penggabungan data, pemotongan wilayah penelitian (*crop area*), *regridding*, konversi satuan data CMIP6, penanganan nilai *null* serta sinkronisasi waktu, dan normalisasi data.

#### 4.1.3.1 Penggabungan Data

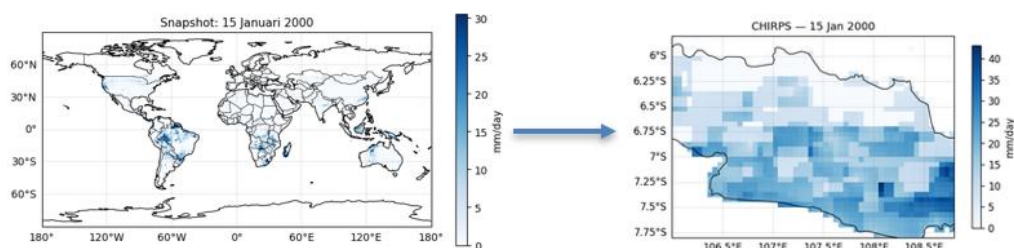
Penggabungan data dilakukan untuk menyatukan beberapa file dataset menjadi satu struktur data yang utuh sebelum digunakan pada tahap pengolahan selanjutnya. Data curah hujan CHIRPS tersedia dalam beberapa file berformat NetCDF yang masing masing merepresentasikan periode waktu satu tahun. Seluruh file tersebut terlebih dahulu diidentifikasi pada direktori data, kemudian digabungkan menggunakan fungsi `open_mfdataset` dari pustaka *xarray* dengan metode `combine='by_coords'`. Proses ini memungkinkan penggabungan beberapa file NetCDF berdasarkan kesesuaian koordinat spasial dan temporal sehingga membentuk satu dataset kontinu dengan dimensi  $\text{time} \times \text{latitude} \times \text{longitude}$ .

Sementara itu, data topografi DEMNAS tersedia dalam beberapa file raster (*tile*) berformat GeoTIFF yang masing masing merepresentasikan bagian wilayah tertentu. Seluruh tile DEMNAS terlebih dahulu diidentifikasi pada direktori data kemudian digabungkan menggunakan metode *mosaic* dengan memanfaatkan fungsi `merge` dari pustaka *rasterio*. Proses ini menghasilkan satu raster DEMNAS yang mencakup wilayah yang lebih luas dengan sistem koordinat yang konsisten. Hasil penggabungan tersebut kemudian disimpan kembali dalam format GeoTIFF sebagai dataset DEMNAS yang telah terintegrasi.

#### 4.1.3.2 Crop Area Jawa Barat

Setelah proses penggabungan data, langkah selanjutnya adalah melakukan pemotongan wilayah penelitian (*crop area*) agar dataset hanya mencakup wilayah Jawa Barat. Proses ini dilakukan untuk mengurangi ukuran data serta memfokuskan analisis pada area studi yang relevan. Batas wilayah yang digunakan ditentukan berdasarkan rentang koordinat latitude dan longitude wilayah Jawa Barat. Pemotongan wilayah ini ditunjukkan pada Gambar 4.5, di mana data global

dipersempit menjadi wilayah Jawa Barat sehingga hanya area Jawa Barat yang digunakan dalam analisis.



Gambar 4.5 *Cropping* Data Curah Hujan Skala Global ke Wilayah Jawa Barat

Pada dataset CHIRPS, proses pemotongan dilakukan menggunakan metode seleksi koordinat pada *xarray*. Proses ini memilih subset data berdasarkan rentang koordinat latitude dan longitude yang telah ditentukan sehingga menghasilkan dataset curah hujan yang hanya mencakup wilayah Jawa Barat.

Pada dataset DEMNAS, proses pemotongan dilakukan dengan membaca bagian raster yang berada dalam batas koordinat wilayah penelitian menggunakan fungsi *window from bounds* dari pustaka *rasterio*. Metode ini memungkinkan pengambilan subset raster secara langsung berdasarkan batas koordinat geografis tanpa perlu memproses seluruh raster.

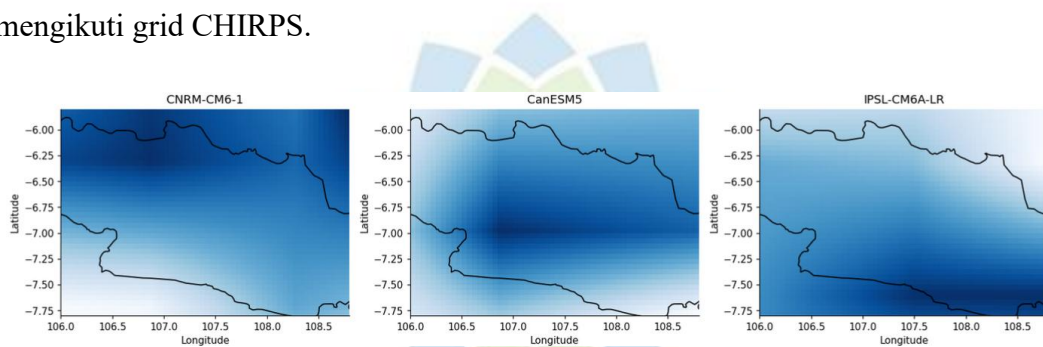
Sementara itu, pada dataset CMIP6, proses pemotongan wilayah tidak dilakukan secara langsung pada tahap ini. Pemotongan wilayah dilakukan bersamaan dengan proses *regridding* pada tahap berikutnya sehingga data model iklim dapat disesuaikan dengan grid wilayah penelitian secara lebih efisien.

#### 4.1.3.3 *Regridding*

Setelah proses pemotongan wilayah penelitian, langkah selanjutnya adalah melakukan *regridding* untuk menyamakan resolusi spasial antara dataset curah hujan observasi dan data model iklim. Pada penelitian ini, grid CHIRPS digunakan sebagai grid target, sehingga data curah hujan dari model iklim CMIP6 perlu disesuaikan agar memiliki resolusi spasial dan struktur grid yang sama dengan CHIRPS.

Proses *regridding* dilakukan menggunakan metode *bilinear interpolation*. Metode ini menghitung nilai pada grid target dengan melakukan interpolasi linier terhadap empat titik grid terdekat pada data sumber.

Sebelum proses interpolasi dilakukan, data CMIP6 terlebih dahulu dipotong menggunakan batas wilayah yang diperluas dengan margin sekitar  $\pm 2^\circ$  dari batas wilayah penelitian. Langkah ini bertujuan untuk mengurangi ukuran data yang diproses sekaligus memastikan bahwa titik titik grid yang diperlukan untuk proses interpolasi tetap tersedia di sekitar wilayah penelitian. Selanjutnya, proses interpolasi dilakukan untuk setiap langkah waktu pada dataset CMIP6 sehingga menghasilkan dataset baru dengan dimensi  $\text{time} \times \text{latitude} \times \text{longitude}$  yang mengikuti grid CHIRPS.



Gambar 4.6 Hasil *Regridding* Data Curah Hujan CMIP6 ke Grid CHIRPS pada wilayah Jawa Barat

Hasil proses *regridding* ditunjukkan pada Gambar 4.6, yang menampilkan contoh distribusi curah hujan dari tiga model CMIP6 yaitu CNRM CM6 1, CanESM5, dan IPSL CM6A LR setelah dilakukan penyesuaian grid. Pada gambar tersebut terlihat bahwa setiap model telah memiliki struktur grid yang seragam mengikuti resolusi CHIRPS pada wilayah Jawa Barat. Dengan demikian, data curah hujan dari model iklim global dapat direpresentasikan pada skala spasial yang lebih rinci sehingga dapat digunakan secara konsisten pada tahap analisis dan pemodelan selanjutnya.

#### 4.1.3.4 Konversi Satuan CMIP6

Data curah hujan dari model iklim CMIP6 umumnya disajikan dalam satuan  $\text{kg m}^{-2} \text{s}^{-1}$ , yang merepresentasikan laju presipitasi per satuan luas dan waktu. Agar dapat dibandingkan secara langsung dengan data observasi CHIRPS yang menggunakan satuan  $\text{mm/hari}$  (*mm/day*), diperlukan proses konversi satuan.

Konversi dilakukan dengan memanfaatkan hubungan bahwa  $1 \text{ kg m}^{-2}$  setara dengan  $1 \text{ mm}$  presipitasi, sehingga nilai curah hujan dalam satuan  $\text{kg m}^{-2} \text{ s}^{-1}$  dapat dikonversi ke *mm/day* dengan mengalikan nilai tersebut dengan jumlah detik dalam satu hari, yaitu 86.400 detik.

Proses konversi diterapkan pada seluruh dataset CMIP6 yang telah melalui tahap *regridding*. Selain itu, nilai curah hujan negatif yang muncul akibat proses interpolasi atau kesalahan numerik kemudian ditangani dengan melakukan *clipping*, yaitu mengubah nilai negatif menjadi nol. Langkah ini dilakukan karena secara fisik curah hujan tidak mungkin memiliki nilai negatif.

Setelah proses konversi selesai, seluruh dataset CMIP6 memiliki satuan *mm/day*, sehingga konsisten dengan satuan data curah hujan CHIRPS dan dapat digunakan secara langsung pada tahap pengolahan data dan pemodelan selanjutnya

#### **4.1.3.5 Penanganan Nilai *Null* dan Sinkronisasi Waktu**

Setelah proses konversi satuan, langkah selanjutnya adalah melakukan penanganan nilai null (NaN) serta sinkronisasi waktu antar dataset. Penanganan nilai null diperlukan untuk memastikan bahwa seluruh data yang digunakan pada tahap pemodelan tidak mengandung nilai kosong yang dapat menyebabkan kesalahan pada proses pelatihan model.

Pada dataset CHIRPS, nilai NaN berasal dari area lautan yang berada di luar wilayah daratan Jawa Barat. Nilai tersebut kemudian ditangani dengan metode *fill zero*, yaitu mengganti nilai NaN dengan nilai  $0 \text{ mm/hari}$ . Pendekatan ini digunakan karena nilai tersebut merepresentasikan kondisi tanpa presipitasi pada lokasi yang tidak relevan terhadap analisis wilayah daratan.

Selanjutnya dilakukan sinkronisasi waktu antara dataset CHIRPS dan CMIP6. Proses ini bertujuan untuk memastikan bahwa seluruh dataset memiliki rentang waktu yang sama sehingga dapat dibandingkan secara langsung pada setiap langkah waktu. Sinkronisasi dilakukan dengan memilih irisan tanggal yang sama pada seluruh dataset tanpa melakukan interpolasi waktu. Dengan pendekatan ini, hanya tanggal yang tersedia secara bersamaan pada seluruh dataset yang dipertahankan untuk tahap analisis berikutnya.

Hasil dari proses ini adalah dataset curah hujan CHIRPS dan CMIP6 yang telah bersih dari nilai null serta memiliki dimensi waktu yang konsisten dengan

jumlah hari sebanyak 5475 hari. Dataset yang telah disinkronkan tersebut kemudian digunakan pada tahap persiapan data selanjutnya sebelum memasuki proses pemodelan.

#### 4.1.3.6 Normalisasi Data

Normalisasi dilakukan untuk menyamakan skala nilai antar dataset sehingga seluruh data berada pada rentang yang sama sebelum digunakan pada proses pelatihan model. Pada penelitian ini normalisasi dilakukan menggunakan metode *Min-Max scaling* sehingga nilai data ditransformasikan ke dalam rentang 0 hingga 1.

Nilai minimum dan maksimum yang digunakan dalam normalisasi dihitung hanya dari data pelatihan (*training set*), kemudian nilai tersebut digunakan kembali untuk menormalisasi data validasi dan data pengujian. Pendekatan ini digunakan untuk menjaga konsistensi distribusi data selama proses pelatihan dan evaluasi model.

Proses normalisasi diterapkan pada data curah hujan CHIRPS sebagai data target serta pada data curah hujan CMIP6 untuk setiap model iklim yang digunakan sebagai data input. Selain itu, data elevasi DEMNAS juga terlebih dahulu disesuaikan ke grid CHIRPS melalui proses interpolasi spasial sebelum dilakukan normalisasi. Nilai minimum dan maksimum dari setiap dataset kemudian disimpan dalam berkas scalers sehingga dapat digunakan kembali pada proses denormalisasi hasil prediksi model pada tahap evaluasi.

#### 4.1.3.7 Data Split

Setelah seluruh proses preprocessing selesai, dataset kemudian dibagi menjadi tiga subset data yaitu data pelatihan (*train*), data validasi (*validation*), dan data pengujian (*test*). Pembagian data dilakukan berdasarkan tahun pengamatan untuk menjaga kontinuitas temporal pada data curah hujan. Pada penelitian ini, data tahun 2000–2010 digunakan sebagai data pelatihan, tahun 2011–2012 digunakan sebagai data validasi, dan tahun 2013–2015 digunakan sebagai data pengujian.

Data CHIRPS disimpan sebagai target resolusi tinggi (high resolution), sedangkan data curah hujan dari model CMIP6 disimpan sebagai input resolusi

rendah (*low resolution*). Selain itu, data topografi DEMNAS disimpan secara terpisah karena bersifat statis dan tidak memiliki dimensi waktu.

Seluruh dataset kemudian disimpan dalam format NetCDF agar struktur data spasial dan temporal tetap terjaga. Untuk setiap subset data disimpan file CHIRPS sebagai target serta file CMIP6 untuk masing masing model sebagai input. Selain itu, file `scalers.json` menyimpan parameter normalisasi yang digunakan pada proses preprocessing, sedangkan `split_info.json` menyimpan informasi pembagian data berdasarkan tahun.

Struktur direktori dataset hasil preprocessing ditunjukkan pada Gambar 4.7, yang memperlihatkan organisasi file dataset ke dalam direktori *train*, *validation*, dan *test* beserta file pendukung yang digunakan pada tahap pemodelan.

```

├─ netcdf/
│  ├─ demnas_topo.nc
│  ├─ train/
│  │  ├─ chirps_hr_train.nc
│  │  ├─ cmip6_CNRM-CM6-1_lr_train.nc
│  │  ├─ cmip6_CanESM5_lr_train.nc
│  │  └─ cmip6_IPSL-CM6A-LR_lr_train.nc
│  │
│  ├─ val/
│  │  ├─ chirps_hr_val.nc
│  │  ├─ cmip6_CNRM-CM6-1_lr_val.nc
│  │  ├─ cmip6_CanESM5_lr_val.nc
│  │  └─ cmip6_IPSL-CM6A-LR_lr_val.nc
│  │
│  └─ test/
│     ├─ chirps_hr_test.nc
│     ├─ cmip6_CNRM-CM6-1_lr_test.nc
│     ├─ cmip6_CanESM5_lr_test.nc
│     └─ cmip6_IPSL-CM6A-LR_lr_test.nc
│
├─ scalers/
│  └─ scalers.json
│
└─ splits/
   └─ split_info.json

```

Gambar 4.7 Struktur Direktori Dataset Hasil *Preprocessing*

#### 4.1.4 Modelling

Setelah seluruh dataset melalui tahap *data preparation* dan pembagian data, tahap selanjutnya adalah membangun model untuk melakukan proses *statistical downscaling*. Proses pemodelan dilakukan menggunakan arsitektur Super Resolution Generative Adversarial Network (SRGAN) yang dirancang untuk merekonstruksi data beresolusi tinggi dari input beresolusi rendah.

#### 4.1.4.1 Desain Arsitektur

Model yang digunakan pada penelitian ini merupakan implementasi dari Super Resolution Generative Adversarial Network (SRGAN) yang terdiri dari dua komponen utama, yaitu *Generator* dan *Discriminator*. *Generator* bertugas untuk merekonstruksi data curah hujan resolusi tinggi dari input resolusi rendah, sedangkan *Discriminator* berfungsi untuk membedakan antara data curah hujan hasil rekonstruksi model dan data observasi sebenarnya.

- a. *Input Layer*: Model menerima input dengan dimensi  $(2, H, W)$  yang terdiri dari dua kanal data, yaitu data curah hujan CMIP6 sebagai variabel utama dan data elevasi DEMNAS sebagai variabel tambahan yang merepresentasikan kondisi topografi wilayah. Kedua variabel tersebut digabungkan sebagai *multichannel input* sehingga model dapat mempelajari hubungan antara pola curah hujan dan kondisi elevasi permukaan.
- b. *Feature Extraction Layer*: Lapisan awal *Generator* menggunakan *Convolution Layer* dengan *kernel*  $9 \times 9$  yang diikuti oleh fungsi aktivasi PReLU. Lapisan ini bertujuan mengekstraksi fitur spasial awal dari data input serta memetakan informasi curah hujan dan topografi ke dalam representasi fitur yang lebih tinggi.
- c. *Residual Blocks*: Setelah tahap ekstraksi fitur awal, data diproses melalui 16 residual blocks yang berfungsi untuk mempelajari representasi spasial yang lebih kompleks. Setiap residual block terdiri dari dua lapisan konvolusi berukuran  $3 \times 3$  yang diikuti oleh Batch Normalization dan fungsi aktivasi PReLU, serta dilengkapi dengan *skip connection* untuk mempertahankan informasi fitur dari lapisan sebelumnya dan meningkatkan stabilitas proses pelatihan.
- d. *Global Skip Connection*: Selain *skip connection* pada setiap residual block, arsitektur *Generator* juga menggunakan global skip connection yang menghubungkan fitur dari lapisan awal dengan fitur setelah residual blocks. Mekanisme ini bertujuan mempertahankan informasi spasial penting dari input serta mempercepat proses konvergensi model.

- e. *Reconstruction Layer*: Setelah proses ekstraksi fitur, jaringan melakukan rekonstruksi data melalui beberapa lapisan konvolusi yang bertujuan menghasilkan data curah hujan resolusi tinggi. Lapisan ini terdiri dari Convolution Layer  $3 \times 3$  yang diikuti oleh fungsi aktivasi PReLU, kemudian Convolution Layer  $9 \times 9$  untuk menghasilkan output akhir.
- f. *Output Layer*: Lapisan terakhir menggunakan fungsi aktivasi Sigmoid untuk menghasilkan nilai curah hujan dalam rentang 0–1, sesuai dengan data target yang telah melalui proses normalisasi pada tahap *data preparation*.
- g. *Discriminator Network*: *Discriminator* menggunakan arsitektur PatchGAN, yang terdiri dari beberapa lapisan konvolusi dengan fungsi aktivasi LeakyReLU. Jaringan ini bertugas membedakan antara data curah hujan hasil rekonstruksi *Generator* dan data observasi CHIRPS. Pendekatan PatchGAN memungkinkan model mengevaluasi keaslian data pada tingkat *patch* sehingga mampu menangkap pola distribusi spasial curah hujan secara lebih detail.

#### 4.1.4.2 Training Model

```

BEGIN
# =====
# 1. Inisialisasi Parameter
# =====
SET N_RESBLOCKS = 16
SET N_FILTERS   = 128
SET BATCH_SIZE  = 16
SET N_EPOCHS    = 200
SET LR_G        = 1e 4
SET LR_D        = 1e 4
SET LAMBDA_ADV  = 1e 4
SET PATIENCE    = 40

SET daftar_model = [CanESM5, CNRM CM6 1, IPSL CM6A LR]
# =====
# Definisi Loss Function
# =====
FUNCTION PixelLoss(y_pred, y_true):
    mse ← MSE(y_pred, y_true)
    mae ← MAE(y_pred, y_true)
    corr ← CorrelationLoss(y_pred, y_true)

```

```

freq ← FFTLoss(y_pred, y_true)

RETURN 0.3*mse + 0.2*mae + 0.3*corr + 0.2*freq

FUNCTION AdversarialLoss(pred, label):
    RETURN BCEWithLogits(pred, label)

# =====
# Training per Epoch
# =====
FUNCTION TrainOneEpoch(G, D, DataLoader):

    SET G.train = TRUE
    SET D.train = TRUE

    FOR setiap batch (x, y_real) DO

        #
        # (1) Train Discriminator
        #
        y_fake ← G(x) (tanpa grad)

        real_out ← D(y_real)
        fake_out ← D(y_fake)

        loss_D ← 0.5 * (
            BCE(real_out, 0.9) + # label smoothing
            BCE(fake_out, 0.0)
        )

        BACKPROP loss_D
        UPDATE parameter D

        #
        # (2) Train Generator
        #
        REPEAT 2 kali:

            y_fake ← G(x)
            fake_out ← D(y_fake)

            loss_pix ← PixelLoss(y_fake, y_real)
            loss_adv ← BCE(fake_out, 1.0)

            loss_G ← loss_pix + LAMBDA_ADV * loss_adv

            BACKPROP loss_G
            UPDATE parameter G

```

```

        END REPEAT
    END FOR
    RETURN rata rata loss_G dan loss_D

# =====
# Validasi
# =====
FUNCTION Validate(G, DataLoader):
    SET G.eval = TRUE
    FOR setiap batch (x, y_real) DO
        y_pred ← G(x)
        hitung loss ← PixelLoss(y_pred, y_real)
    END FOR
    RETURN rata rata loss_validasi
# =====
# Training Utama
# =====
FOR setiap model_name dalam daftar_model DO
    PRINT "Training model:", model_name
    LOAD DataLoader_train, DataLoader_val
    INITIALIZE:
        G ← Generator
        D ← Discriminator
        Optimizer_G ← Adam(G, LR_G)
        Optimizer_D ← Adam(D, LR_D)
        Scheduler_G ← CosineAnnealingLR
        Scheduler_D ← CosineAnnealingLR
    SET best_val_loss = ∞
    SET no_improve = 0
    FOR epoch = 1 TO N_EPOCHS DO
        loss_G, loss_D ← TrainOneEpoch(G, D, train_loader)

        val_loss ← Validate(G, val_loader)
        UPDATE Scheduler_G
        UPDATE Scheduler_D

        #
        # Save model terbaik
        #
        IF val_loss < best_val_loss THEN
            best_val_loss ← val_loss
            no_improve ← 0

        SAVE checkpoint:
            state Generator
            state Discriminator

```

```

        epoch
        val_loss
        hyperparameter

    ELSE
        no_improve ← no_improve + 1
    END IF

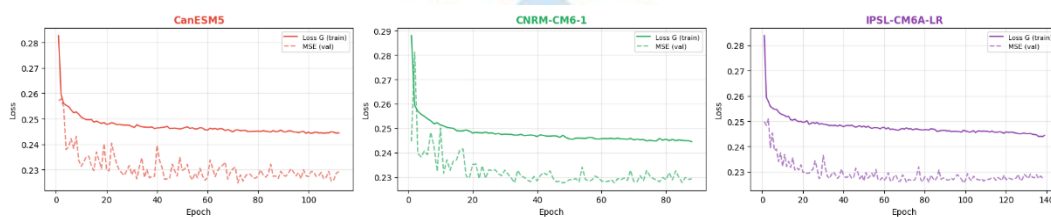
    #
    # Early stopping
    #
    IF no_improve ≥ PATIENCE THEN
        PRINT "Early stopping"
        BREAK
    END

```

Proses pelatihan model dilakukan menggunakan arsitektur SRGAN yang terdiri dari dua jaringan utama, yaitu *Generator* dan *Discriminator*. *Generator* bertugas merekonstruksi data curah hujan resolusi tinggi dari input resolusi rendah, sedangkan *Discriminator* berfungsi membedakan antara data curah hujan hasil rekonstruksi dan data observasi. Pada tahap ini, data CMIP6 yang telah diproses digunakan sebagai input bersama dengan data DEMNAS sebagai variabel tambahan, sedangkan data CHIRPS digunakan sebagai target resolusi tinggi. Proses pelatihan dilakukan menggunakan *optimizer* Adam dengan *learning rate* sebesar  $1 \times 10^{-4}$  untuk *Generator* dan *Discriminator*, serta batch size sebesar 16. Model dilatih hingga maksimum 200 epoch dengan mekanisme *early stopping* untuk menghentikan pelatihan apabila tidak terjadi peningkatan performa pada data validasi selama sejumlah 40 epoch. Fungsi loss pada *Generator* menggunakan kombinasi beberapa komponen, yaitu Mean Squared Error (MSE), Mean Absolute Error (MAE), correlation loss, dan frequency loss, yang bertujuan agar model mampu mempelajari kesamaan nilai, pola spasial, serta karakteristik frekuensi dari data curah hujan. Selain itu, digunakan juga *adversarial loss* yang berasal dari *Discriminator* untuk mendorong *Generator* menghasilkan distribusi curah hujan yang lebih realistis.

Proses pelatihan dilakukan secara terpisah untuk tiga model iklim CMIP6 yaitu CanESM5, CNRM CM6 1, dan IPSL CM6A LR, di mana model terbaik pada setiap eksperimen disimpan berdasarkan nilai validation loss terendah yang diperoleh selama proses pelatihan. Perkembangan proses pelatihan dan konvergensi

model untuk ketiga eksperimen dapat diamati melalui kurva *training* pada Gambar 4.8. Grafik tersebut menunjukkan bahwa nilai *loss* Generator pada data *training* serta MSE pada data validasi mengalami penurunan secara signifikan pada *epoch* awal dan kemudian melandai secara stabil hingga mencapai titik konvergensi. Secara spesifik, model IPSL-CM6A-LR memerlukan waktu pelatihan paling lama hingga mencapai sekitar 139 *epoch*, disusul oleh model CanESM5 yang stabil pada sekitar 112 *epoch*, dan model CNRM-CM6-1 yang mencapai konvergensi tercepat pada 88 *epoch*. Proses pelatihan dilakukan secara terpisah untuk ketiga model iklim CMIP6 tersebut, di mana model terbaik pada setiap eksperimen disimpan berdasarkan nilai *validation loss* terendah yang diperoleh selama proses pelatihan.



Gambar 4.8 Kurva Training

#### 4.1.4.3 Simpan Model Terbaik

Selama proses pelatihan model, dilakukan mekanisme penyimpanan model terbaik (*model checkpoint*) untuk memastikan bahwa model yang digunakan pada tahap evaluasi merupakan model dengan performa terbaik pada data validasi. Pada setiap *epoch* pelatihan, nilai *validation loss* dihitung menggunakan data validasi yang tidak digunakan pada proses pelatihan. Jika nilai *validation loss* yang diperoleh lebih kecil dibandingkan nilai terbaik sebelumnya, maka parameter model *Generator* dan *Discriminator* akan disimpan sebagai model terbaik.

Proses penyimpanan model dilakukan dengan menyimpan parameter jaringan (*state dictionary*), informasi *epoch* pelatihan, nama model CMIP6 yang digunakan, serta konfigurasi hiperparameter seperti jumlah filter, jumlah residual block, dan bobot fungsi loss. Model disimpan dalam format *.pth* sehingga dapat digunakan kembali pada tahap evaluasi maupun inferensi tanpa perlu melakukan proses pelatihan ulang.

Selain itu, untuk mencegah proses pelatihan berlangsung terlalu lama tanpa peningkatan performa, digunakan mekanisme *early stopping*. Jika tidak

terjadi perbaikan nilai validation loss dalam sejumlah epoch tertentu, maka proses pelatihan akan dihentikan secara otomatis. Pendekatan ini membantu menghindari overfitting serta memastikan model yang diperoleh memiliki kemampuan generalisasi yang lebih baik terhadap data yang belum pernah dilihat sebelumnya.

#### 4.1.5 *Evaluation*

Setelah proses pelatihan model selesai dan model terbaik telah disimpan, tahap selanjutnya adalah melakukan evaluasi untuk menilai kinerja model dalam merekonstruksi data curah hujan resolusi tinggi. Evaluasi dilakukan dengan membandingkan hasil prediksi model terhadap data observasi CHIRPS pada data pengujian yang tidak digunakan selama proses pelatihan.

##### 4.1.5.1 **Evaluasi Model**

Evaluasi model dilakukan menggunakan data uji (test set) untuk menilai kemampuan model SRGAN dalam merekonstruksi curah hujan resolusi tinggi. Model terbaik yang telah diperoleh pada tahap pelatihan dimuat kembali dan digunakan untuk menghasilkan prediksi curah hujan berdasarkan data input CMIP6 dan DEMNAS.

Sebelum dilakukan evaluasi, hasil prediksi yang masih berada dalam rentang normalisasi  $[0, 1]$  terlebih dahulu dikembalikan ke nilai aslinya melalui proses denormalisasi sehingga satuan curah hujan kembali menjadi *mm/day*. Selanjutnya, nilai prediksi dibandingkan dengan data observasi CHIRPS pada periode pengujian.

Evaluasi dilakukan melalui beberapa tahap, yaitu inferensi pada data uji, denormalisasi hasil prediksi, serta perhitungan metrik evaluasi yang meliputi RMSE, MAE, dan PCC. Perhitungan metrik dilakukan untuk seluruh piksel serta rata-rata temporal guna menilai tingkat kesalahan prediksi dan kesesuaian pola curah hujan. Selain itu, dilakukan pula perbandingan dengan baseline, yaitu data CMIP6 asli tanpa *downscaling*, untuk melihat peningkatan performa setelah diterapkan model SRGAN. Ringkasan hasil evaluasi model dan perbandingan dengan baseline ditampilkan pada Tabel 4.4.

Tabel 4.4 Tabel Evaluasi

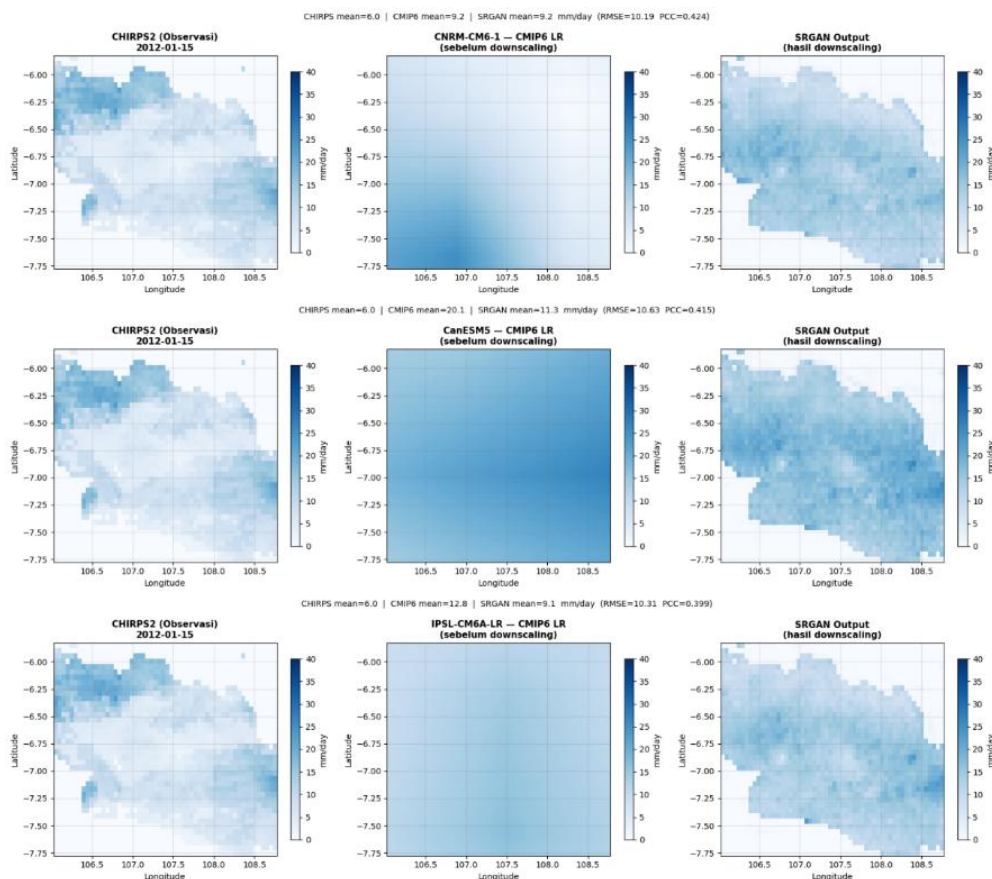
| Model        | RMSE    | RMSE    | MAE     | MAE    | PCC    | PCC    |
|--------------|---------|---------|---------|--------|--------|--------|
|              | CMIP6   | SRGAN   | CMIP6   | SRGAN  | CMIP6  | SRGAN  |
| CanESM5      | 21,7957 | 10,6318 | 12,5288 | 6,8983 | 0,0562 | 0,4145 |
| CNRM CM6 1   | 28,2427 | 10,1938 | 16,5766 | 6,3886 | 0,1279 | 0,4239 |
| IPSL CM6A LR | 18,0907 | 10,3085 | 11,5679 | 6,4725 | 0,0693 | 0,3988 |

#### 4.1.5.2 Analisis Hasil

Berdasarkan hasil evaluasi pada Tabel 4.4, model SRGAN menunjukkan peningkatan kinerja yang signifikan dibandingkan dengan data CMIP6 asli (baseline) dalam merekonstruksi curah hujan resolusi tinggi di wilayah Jawa Barat. Peningkatan tersebut ditunjukkan oleh penurunan nilai RMSE dan MAE, serta kenaikan nilai PCC pada seluruh model CMIP6 yang digunakan.

Pada model CanESM5, nilai RMSE menurun dari 21.7957 *mm/day* menjadi 10.6318 *mm/day* dan MAE dari 12.5288 *mm/day* menjadi 6.8983 *mm/day*, sementara nilai PCC meningkat dari 0.0562 menjadi 0.4145. Pada model CNRM CM6 1, penurunan RMSE terjadi dari 28.2427 *mm/day* menjadi 10.1938 *mm/day* dan MAE dari 16.5766 *mm/day* menjadi 6.3886 *mm/day*, dengan PCC meningkat dari 0.1279 menjadi 0.4239. Sementara itu, pada model IPSL CM6A LR, nilai RMSE menurun dari 18.0907 *mm/day* menjadi 10.3085 *mm/day* dan MAE dari 11.5679 *mm/day* menjadi 6.4725 *mm/day*, serta PCC meningkat dari 0.0693 menjadi 0.3988.

Perbedaan performa antar model CMIP6 ini kemungkinan berkaitan erat dengan *rasio upsampling* yang berbeda-beda. CNRM-CM6-1 dengan resolusi ~156 km memiliki rasio *upsampling* sekitar 28× terhadap target resolusi CHIRPS2 ~5.5 km, sementara CanESM5 (~312 km) dan IPSL-CM6A-LR (~278 km) memiliki rasio yang jauh lebih besar, yaitu sekitar 57× dan 51×. Semakin besar rasio *upsampling*, semakin banyak informasi skala kecil yang harus direkonstruksi oleh model dari data input yang sangat kasar, sehingga tingkat kesulitan rekonstruksi meningkat secara signifikan. Hal ini menjelaskan mengapa meskipun CMIP6 raw CNRM-CM6-1 memiliki RMSE awal yang paling buruk (28.24 mm/hari), model SRGAN justru menghasilkan rekonstruksi terbaik pada model tersebut



Gambar 4.9 Perbandingan CHIRPS, CMIP6, dan Hasil *Downscaling* SRGAN

Secara visual, perbedaan kualitas rekonstruksi curah hujan dapat diamati pada Gambar 4.9, yang menyajikan perbandingan tiga sumber data untuk model CNRM CM6 1, CanESM5, dan IPSL CM6A LR pada tanggal 15 Januari 2012. Interpretasi intensitas curah hujan pada visualisasi ini didasarkan pada skala warna (*colorbar*) warna terang seperti putih menunjukkan intensitas curah hujan yang rendah atau tidak ada hujan, sedangkan warna yang semakin biru pekat (biru tua) mengindikasikan bahwa intensitas curah hujan di wilayah tersebut semakin tinggi. Selain itu, seluruh model hasil rekonstruksi SRGAN terlihat telah mampu membedakan wilayah daratan Jawa Barat dan wilayah lautan secara jelas, mengikuti pola yang ada pada data observasi CHIRPS2.

Pada baris pertama untuk model CNRM CM6 1, panel kiri menampilkan data observasi CHIRPS2 sebagai referensi dengan nilai rata-rata sekitar 6,0 mm/hari, yang memperlihatkan variasi spasial curah hujan yang heterogen di wilayah Jawa Barat. Panel tengah menampilkan data CMIP6 resolusi rendah

sebelum *downscaling* dengan nilai rata-rata sekitar 9,2 mm/hari, yang tampak sangat halus dan homogen sehingga gagal merepresentasikan variabilitas spasial curah hujan yang sesungguhnya. Panel kanan menampilkan hasil *downscaling* SRGAN dengan nilai rata-rata sekitar 9,2 mm/hari (RMSE = 10,19; PCC = 0,424), yang secara visual menunjukkan pola spasial yang lebih detail dan lebih mendekati distribusi data CHIRPS2 dibandingkan data CMIP6 asli.

Pada baris kedua untuk model CanESM5, panel kiri kembali menunjukkan data CHIRPS2 sebagai acuan dengan pola curah hujan yang heterogen. Panel tengah menampilkan data CMIP6 resolusi rendah dengan nilai rata-rata sekitar 12,8 mm/hari yang terlihat sangat halus dan cenderung homogen, sehingga tidak mampu menangkap variasi spasial secara detail. Panel kanan menunjukkan hasil *downscaling* SRGAN dengan nilai rata-rata sekitar 9,1 mm/hari (RMSE = 10,63; PCC = 0,413), yang menunjukkan peningkatan detail spasial serta pola distribusi yang lebih mendekati data observasi dibandingkan data CMIP6 sebelum *downscaling*.

Pada baris ketiga untuk model IPSL CM6A LR, panel kiri menampilkan data CHIRPS2 sebagai referensi. Panel tengah menunjukkan data CMIP6 resolusi rendah dengan nilai rata-rata sekitar 12,8 mm/hari yang tampak sangat homogen dan memiliki variasi spasial yang sangat terbatas. Panel kanan menampilkan hasil *downscaling* SRGAN dengan nilai rata-rata sekitar 9,1 mm/hari (RMSE = 10,31; PCC = 0,359), yang menunjukkan peningkatan detail spasial dibandingkan data CMIP6, meskipun pola yang dihasilkan masih relatif lebih halus dibandingkan hasil pada model lainnya.

## 4.2 Pembahasan

### 4.2.1 Rumusan Masalah 1 Perancangan Model SRGAN untuk *Downscaling* Curah Hujan CMIP6

Rumusan masalah pertama dalam penelitian ini adalah “Bagaimana merancang dan mengimplementasikan model SRGAN untuk melakukan *statistical downscaling* pada data curah hujan CMIP6 yang memiliki resolusi spasial rendah sehingga dapat menghasilkan data curah hujan dengan resolusi yang setara dengan data observasi CHIRPS2 di wilayah Jawa Barat”. Untuk menjawab permasalahan tersebut, penelitian ini membangun sebuah model *deep learning* berbasis SRGAN

yang memanfaatkan data curah hujan CMIP6 sebagai input utama dan data topografi DEMNAS sebagai variabel tambahan. Kedua data tersebut digabungkan sebagai *multichannel input* sehingga model dapat mempelajari hubungan antara pola curah hujan dan kondisi topografi wilayah.

Proses perancangan model dimulai dengan tahap data preparation, yang meliputi penggabungan dataset, pemotongan wilayah studi Jawa Barat, proses *regridding*, konversi satuan, penanganan nilai *null*, sinkronisasi waktu, normalisasi data, serta pembagian dataset menjadi data pelatihan, validasi, dan pengujian. Setelah tahap persiapan data selesai, model SRGAN dirancang dengan arsitektur *Generator* berbasis SRResNet yang terdiri dari beberapa lapisan konvolusi dan 16 residual blocks untuk mempelajari representasi spasial curah hujan secara lebih mendalam. Model ini juga menggunakan skip connection untuk mempertahankan informasi spasial penting selama proses pelatihan. Sementara itu, jaringan *Discriminator* menggunakan arsitektur PatchGAN yang berfungsi untuk membedakan antara data curah hujan hasil prediksi *Generator* dan data observasi CHIRPS.

Dalam proses pelatihan, model dilatih menggunakan kombinasi beberapa fungsi *loss* yang dirancang untuk meningkatkan kualitas hasil prediksi, termasuk *pixel loss* dan *adversarial loss*. Model kemudian dilatih menggunakan data pelatihan hingga mencapai konvergensi dan model terbaik disimpan berdasarkan nilai kesalahan validasi terendah. Dengan pendekatan ini, model SRGAN mampu mempelajari hubungan spasial antara data curah hujan resolusi rendah dan data referensi resolusi tinggi sehingga dapat menghasilkan data curah hujan dengan resolusi yang lebih detail dan mendekati data observasi CHIRPS2 seperti yang ditunjukkan pada Tabel 4.5.

Tabel 4.5 Perbandingan Resolusi Awal dan Resolusi Akhir Data CMIP6

| Data CMIP6   | Resolusi Awal                                | Resolusi Akhir                               |
|--------------|--|--|
| CNRM CM6 1   | $1.40^{\circ} \times 1.41^{\circ}$ (~156 km) | $0.05^{\circ} \times 0.05^{\circ}$ (~5.5 km) |
| CanESM5      | $2.79^{\circ} \times 2.81^{\circ}$ (~312 km) | $0.05^{\circ} \times 0.05^{\circ}$ (~5.5 km) |
| IPSL CM6A LR | $1.27^{\circ} \times 2.50^{\circ}$ (~278 km) | $0.05^{\circ} \times 0.05^{\circ}$ (~5.5 km) |

#### 4.2.2 Rumusan Masalah 2 Evaluasi Performa Model SRGAN dalam Rekonstruksi Curah Hujan Resolusi Tinggi

Rumusan masalah kedua dalam penelitian ini adalah “Seberapa baik performa model SRGAN dalam merekonstruksi data curah hujan resolusi tinggi dari data proyeksi CMIP6 jika dievaluasi menggunakan parameter *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), dan *Pearson Correlation Coefficient* (PCC)?”. Berdasarkan hasil evaluasi, baik secara kuantitatif maupun secara visual pada Gambar 4.9, metode *statistical downscaling* berbasis SRGAN terbukti mampu meningkatkan kualitas estimasi curah hujan secara signifikan dibandingkan dengan data CMIP6 asli.

Tabel 4.6 Hasil Evaluasi Statistik

| Model        | Metrik | CMIP6   | SRGAN   | Improvement |
|--------------|--------|---------|---------|-------------|
| CanESM5      | RMSE   | 21,7957 | 10,6318 | 51,2%       |
|              | MAE    | 12,5288 | 6,8983  | 44,9%       |
|              | PCC    | 0,0562  | 0,4145  | 0,3584      |
| CNRM CM6 1   | RMSE   | 28,2427 | 10,1938 | 63,9%       |
|              | MAE    | 16,5766 | 6,3886  | 61,5%       |
|              | PCC    | 0,1279  | 0,4239  | 0,296       |
| IPSL CM6A LR | RMSE   | 18,0907 | 10,3085 | 43%         |
|              | MAE    | 11,5679 | 6,4725  | 44%         |
|              | PCC    | 0,0693  | 0,3988  | 0,3295      |

Peningkatan ini terlihat jelas dari penurunan tingkat kesalahan prediksi dan peningkatan korelasi pada seluruh model. Seperti yang terlihat di Tabel 4.6 pada model CanESM5, penerapan SRGAN berhasil memperbaiki (*improvement*) nilai RMSE sebesar 51,2% dan MAE sebesar 44,9%, serta mencatat kenaikan nilai absolut korelasi PCC sebesar 0,3584. Hasil serupa terlihat pada model IPSL CM6A LR, di mana nilai RMSE membaik sebesar 43% dan MAE sebesar 44%, dengan peningkatan nilai korelasi PCC mencapai 0,3295. Dari ketiga model yang diuji, CNRM CM6 1 menghasilkan performa terbaik dan peningkatan paling optimal; nilai kesalahan RMSE berhasil diturunkan drastis sebesar 63,9% (dari 28,2427 menjadi 10,1938) dan MAE menurun sebesar 61,5% (dari 16,5766 menjadi 6,3886), dengan peningkatan PCC sebesar 0,296. Dengan peningkatan akurasi yang signifikan pada metrik metrik tersebut, hasil *downscaling* yang diperoleh berpotensi

besar untuk digunakan sebagai informasi pendukung dalam analisis iklim regional serta kajian mitigasi bencana hidrometeorologi pada skala lokal.

#### 4.2.3 Limitasi Penelitian

Meskipun model SRGAN menunjukkan hasil *generate* dengan nilai RMSE, MAE, dan PCC yang lebih mendekati data observasi dibandingkan dengan data CMIP6 asli, penelitian ini masih memiliki beberapa keterbatasan yang perlu diperhatikan dan diperbaiki agar hasil *generate* yang dihasilkan dapat menjadi lebih optimal.

##### 1. Keterbatasan Variabel Input

Penelitian ini hanya memanfaatkan variabel curah hujan CMIP6 dan data topografi DEMNAS sebagai input model, sehingga belum mempertimbangkan faktor meteorologis lain yang juga mempengaruhi pembentukan curah hujan, seperti suhu udara, kelembapan, tekanan atmosfer, dan pola sirkulasi angin. Penambahan variabel atmosfer tambahan berpotensi meningkatkan kemampuan model dalam merepresentasikan proses fisik pembentukan hujan. Hal ini sejalan dengan penelitian Lin dkk. yang menunjukkan bahwa pendekatan *deep learning* untuk *downscaling* dapat memanfaatkan berbagai variabel meteorologi seperti suhu, kelembapan, dan angin dalam menghasilkan data resolusi tinggi [52]

##### 2. Rasio Upsampling yang Sangat Besar

Penelitian ini menghadapi keterbatasan terkait rasio peningkatan resolusi *downscale* yang sangat besar antara data CMIP6 dan data CHIRPS dari ~5 km ke ~100 300 km. Kondisi ini dapat menyebabkan model kesulitan dalam merekonstruksi detail spasial secara akurat karena sebagian informasi skala kecil tidak tersedia pada data input resolusi rendah. Hal ini terlihat dari oleh hasil penelitian, dimana model SRGAN menunjukkan performa yang lebih baik pada data CNRM CM6 1 yang memiliki rasio *upsampling* lebih kecil dibandingkan dengan dua dataset lainnya. Meskipun nilai RMSE, MAE, dan PCC dari data CMIP6 raw pada CNRM CM6 1 relatif lebih buruk, hasil *downscaling* yang dihasilkan oleh SRGAN justru lebih baik. Hal ini mengindikasikan bahwa rasio *upsampling* yang lebih kecil mempermudah model dalam merekonstruksi detail spasial, karena kehilangan informasi pada data resolusi rendah tidak sebesar pada dataset dengan rasio *upsampling* yang lebih tinggi. Permasalahan serupa

juga dilaporkan dalam berbagai penelitian *downscaling* berbasis *deep learning* yang menyatakan bahwa keterbatasan resolusi pada data GCM menyebabkan hilangnya informasi skala kecil, sehingga menyulitkan model dalam merekonstruksi detail spasial secara akurat, terutama pada fenomena curah hujan yang kompleks [53].

