

BAB II

TINJAUAN PUSTAKA

2.1 Statistical Downscaling

Teknik *downscaling* adalah suatu proses transformasi data dari unit skala besar menjadi data pada unit skala yang lebih kecil. Menurut Wilby & Igley, *Downscaling* adalah suatu cara menginterpolasi peubah-peubah prediktor atmosfer berskala besar terhadap peubah-peubah berskala lebih kecil. Terdapat dua jenis pendekatan *downscaling*, yaitu *dynamical downscaling* dan *statistical downscaling*. *Dynamical downscaling* dilakukan dengan cara menetapkan model *General Circulation Model* (GCM) tersarang dengan resolusi spasial yang lebih tinggi, sedangkan *statistical downscaling* dilakukan berdasarkan hubungan fungsional antara variabel prediktor dengan skala besar dan variabel respon dengan skala kecil. Di samping itu juga terdapat pendekatan *statistical dynamical downscaling* yang merupakan kombinasi dari kedua metode *downscaling*. Pendekatan *statistical downscaling* menggunakan data regional atau global untuk memperoleh hubungan fungsional yang dimaksud adalah seperti model regresi.

Pendekatan *Statistical Downscaling* memanfaatkan data *General Circulation Model* (GCM) untuk peramalan iklim lokal (Fruentes & Heiman 2000). *Statistical Downscaling* menjelaskan hubungan antara skala global dan lokal dengan lebih memperhatikan keakuratan model penduga untuk mempelajari dampak perubahan iklim (Yurnal *et al.* 2001). Dalam pendekatan *Statistical Downscaling* perlu dilakukan pemilihan peubah yang akan dijadikan sebagai prediktor dan penentuan domain (lokasi dan jumlah grid), karena kedua hal ini merupakan faktor kritis yang

akan mempengaruhi kestabilan peramalan (Wilby & Wigley 2000). Dengan demikian dalam hal peramalan curah hujan, pemilihan peubah prediktor sebaiknya berdasarkan pada adanya korelasi yang kuat antara peubah tersebut dengan curah hujan.

Statistical Downscaling didefinisikan sebagai upaya menghubungkan antara sirkulasi peubah skala global sebagai peubah penjas dan peubah skala lokal sebagai peubah respon (Sutikno 2008). *Statistical Downscaling* (SD) merupakan model statistik yang menggambarkan hubungan antara data pada grid-grid berskala besar *General Circulation Model* (GCM) dengan data pada grid berskala lebih kecil. *General Circulation Model* (GCM) merupakan alat prediksi utama iklim dan cuaca secara numerik dan sebagai sumber informasi primer untuk menilai pengaruh perubahan iklim. Pendekatan *Statistical Downscaling* (SD) disusun berdasarkan adanya hubungan antara grid skala besar prediktor dengan grid skala lokal respon yang dinyatakan dengan model statistik yang dapat digunakan untuk menterjemahkan anomali-anomali skala global yang menjadi anomali dari beberapa peubah iklim lokal.

Pendekatan *statistical downscaling* disusun berdasarkan adanya hubungan antara variabel prediktor dengan grid skala besar dan variabel respon dengan grid skala lokal yang dinyatakan dengan model statistik sehingga dapat digunakan untuk menterjemahkan penyimpangan-penyimpangan yang terjadi pada skala global menjadi penyimpangan dari beberapa variabel iklim lokal (Zorita dan Storch, 1999). Namun, data *General Circulation Model* (GCM) tidak dapat digunakan secara langsung sebagai variabel prediktor pada model regresi biasa. Data luaran

General Circulation Model (GCM) secara umum merupakan data spasial dan temporal yang memiliki kemungkinan besar untuk terjadi korelasi spasial antar data pada grid yang berbeda dalam satu domain dan Autokorelasi pada data deret waktu. Dalam pemodelan *statistical downscaling*, domain adalah sebagai variabel prediktor yang berdimensi banyak yang besar kemungkinan terjadi *curse of dimensionality*, korelasi spasial antar grid dalam domain dan multikolinearitas antar variabel prediktor. Semakin besar domain dan semakin banyak variabel yang digunakan akan mengakibatkan model semakin kompleks.

Pendekatan *statistical downscaling* memanfaatkan data *General Circulation Model* (GCM) untuk melakukan peramalan iklim lokal. Hasil dari model *statistical downscaling* secara langsung berkaitan dengan statistik iklim pada waktu sebelumnya yang dapat memberikan ramalan untuk waktu yang akan datang dalam berbagai perubahan iklim. Menurut Busuioc, Chen dan Hellstrom (2001) model *statistical downscaling* akan memberikan hasil yang baik apabila memenuhi beberapa syarat, antara lain:

1. Hubungan yang kuat antara variabel respon dan variabel prediktor yang menjelaskan keragaman iklim lokal dengan baik.
2. Variabel prediktor disimulasikan dengan baik oleh model *General Circulation Model* (GCM)
3. Hubungan antara variabel respon dan variabel prediktor tidak berubah seiring dengan perubahan waktu dan tetap sama meskipun terdapat perubahan iklim.

Tabel 2.1 Pengelompokan Statistical Downscaling

No	Kategori	Teknik-Teknik Statistical Downscaling
1	a. Berbasis Regresi	MOS, Perfect Prognosis, Analog, PCR, CCA, MARS, Regresi Bertatar, ANN, PPR
	b. Berbasis Klasifikasi	CART (TSR), NHMM
2	a. Model Linier	MOS, Perfect Prognosis, Analog, Regresi Bertatar, PCR, CCA, CART (TSR)
	b. Model Nonlinier	ANN, MARS, PPR
3	a. Model Parametrik	MOS, Regresi Bertatar, PCR, CCA, CART, TSR
	b. Model Nonparametrik	ANN, MARS, PPR
4	a. Berbasis Proyeksi	PCR, PPR
	b. Berbasis Seleksi	CART (TSR), Regresi Bertatar
5	a. Model Driven	MOS, PCR, CCA, Regresi Bertatar
	b. Data Driven	MARS, CART, (TSR), PPR

Teknik-teknik *Statistical Downscaling* dapat dikelompokkan menjadi lima kategori seperti tercantum pada Tabel 2.1. Kategori ini berdasarkan teknik berbasis Regresi atau Klasifikasi, teknik dengan model Linear atau model Nonlinear, teknik dengan model Parametrik atau model Nonparametrik, teknik berbasis Proyeksi atau Seleksi, dan teknik berbasis model driven atau data driven. Sailor & Li (1999) menyatakan bahwa ada dua kelompok teknik *Statistical Downscaling*, yaitu Teknik *Statistical Downscaling* berbasis Regresi, dan Teknik *Statistical Downscaling* berbasis Klasifikasi tipe cuaca. Kelompok teknik pertama digunakan untuk memprediksi peubah lokal yang bersifat kontinu berdasarkan waktu, sedangkan kelompok teknik kedua dapat digunakan untuk peubah lokal yang kontinu maupun diskrit. Zorita & Storch (1999) mengklasifikasikan teknik-teknik *Statistical*

Downscaling menjadi tiga kategori, yaitu Teknik *Statistical Downscaling* berbasis model linear, Teknik *Statistical Downscaling* berbasis model nonlinear, dan Teknik *Statistical Downscaling* berbasis klasifikasi tipe cuaca.

Model *statistical downscaling* merupakan suatu fungsi yang menggambarkan hubungan fungsional sirkulasi atmosfer global dengan unsur-unsur iklim lokal. Secara umum model *statistical downscaling* adalah sebagai berikut.

$$Y_{(t,p)} = f(X_{(t,q,s,g)}) \quad (2.1)$$

Dimana:

- Y = Variabel iklim lokal
- X = Variabel GCM
- t = Waktu
- p = Banyak Variabel Y
- q = Banyak Variabel X
- s = Banyak Lapisan Atmosfer
- g = Grid Domain GCM

Model *Statistical Downscaling* secara umum menggunakan data deret waktu dan data spasial. Banyaknya variabel Y, variabel X, dan lapisan atmosfer dalam model serta adanya multikolinieritas dan autokorelasi pada variabel Y maupun variabel X menunjukkan tingkat kompleksitas model. Semakin banyak variabel Y dan variabel X yang di gunakan dalam pemodelan, maka model *Statistical Downscaling* juga semakin kompleks.

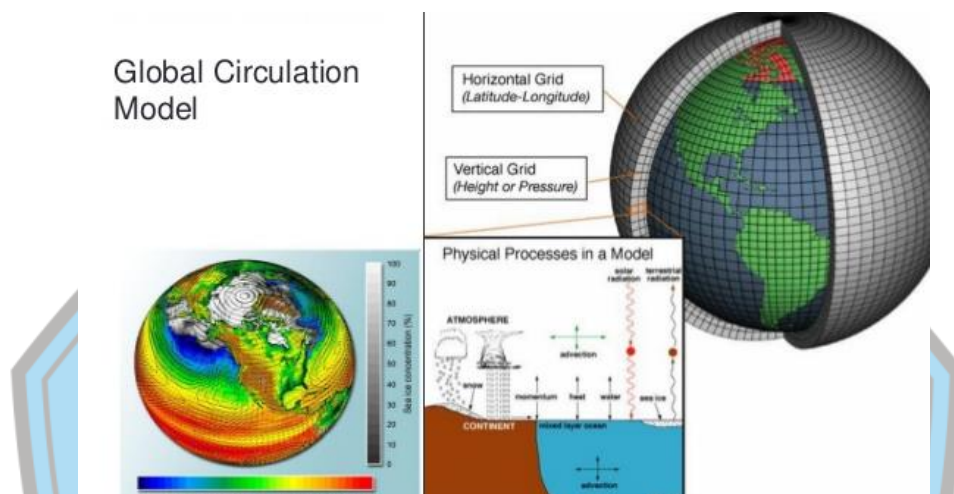
Secara umum permasalahan dalam pemodelan *Statistical Downscaling* adalah sebagai berikut:

1. Luasan dan lokasi domain *General Circulation Model* (GCM), yaitu jumlah grid dalam domain dan lokasi domain di mana peubah prediktornya berkorelasi tinggi dengan peubah respon.
2. Peubah prediktor (luaran *General Circulation Model*) yang bersifat *curse of dimensionality*, nonlinear, *non Gaussian*, bahkan tidak mengikuti sebaran statistik yang baku, dan multikolinearitas.
3. Peubah respon juga bersifat nonlinear, *non Gaussian*, bahkan tidak mengikuti sebaran statistik yang baku.
4. Panjang data historis, di mana umumnya teknik *Statistical Downscaling* memerlukan data historis yang relatif panjang.
5. Data peubah respon yang ekstrim atau pencilan.

2.2 General Circulation Model (GCM)

General Circulation Model (GCM) merupakan model numerik yang bersifat deterministik dengan simulasi komputer yang rumit yang dapat menggambarkan kondisi iklim dengan berbagai macam komponennya yang berubah sepanjang waktu. Model ini menggambarkan hubungan matematik dari berbagai interaksi fisika, kimia dan dinamika atmosfer bumi. Model ini merupakan model yang mampu memahami iklim di masa lampau, sekarang dan masa mendatang. *General Circulation Model* (GCM) mampu mensimulasikan variabel-variabel iklim global pada setiap grid berukuran 300 km^2 di setiap lapisan atmosfer, yang dapat digunakan untuk memprediksi pola iklim dalam jangka waktu yang panjang, yaitu

jangka waktu tahunan. *General Circulation Model* (GCM) merupakan suatu alat yang penting dalam penelitian keragaman dan perubahan iklim (Zorita dan Storch, 1999). Model ini mampu menggambarkan berbagai macam subsistem iklim di bumi, seperti proses-proses di lautan, daratan dan atmosfer.



Gambar 2.1 General Circulation Model (GCM)

General Circulation Model (GCM) juga mampu mensimulasi kondisi iklim berskala besar atau beresolusi rendah, misalnya dapat memproduksi dengan baik pola-pola keragaman atmosfer dan temperatur permukaan laut. Namun, *General Circulation Model* (GCM) tidak dirancang untuk menghasilkan informasi penting dengan resolusi yang lebih tinggi atau skala yang lebih kecil, misalnya untuk temperatur dan curah hujan skala lokal. Dengan demikian, *General Circulation Model* (GCM) dapat melakukan simulasi dengan baik untuk variabel iklim skala besar, tetapi tidak untuk variabel dengan skala yang lebih kecil (lokal atau regional). Menurut Zorita dan Storch (1999) alasan mengapa *General Circulation Model*

(GCM) tidak dapat menghasilkan informasi untuk skala lokal atau regional adalah sebagai berikut.

1. Deskripsi solusi spasial tentang struktur permukaan bumi tidak jelas, khususnya topografi.
2. Hidrodinamika atmosfer bersifat nonlinier sehingga antara grid skala kecil.
3. Terlalu banyak parameter yang tidak mungkin tepat digunakan untuk proses-proses pada skala kecil.

General Circulation Model (GCM) perlu disempurnakan secara terus menerus dan bertahap, meskipun model ini belum memadai untuk fenomena berskala kecil di kawasan topografi yang kompleks. Oleh karena itu, *General Circulation Model* (GCM) tidak dapat digunakan secara langsung untuk melakukan penelitian terhadap dampak iklim secara regional maupun lokal. Adapun langkah-langkah yang dilakukan dalam melakukan penyempurnaan terhadap *General Circulation Model* (GCM) untuk mengatasi perbedaan skala adalah meningkatkan kemampuan komputer terutama untuk skala yang lebih kecil, sehingga dapat melakukan *Regional Circulation Model* (RCM) dari luaran *General Circulation Model* (GCM). Disamping itu juga perlu menerapkan teknik *statistical downscaling*.

Pada umumnya, *General Circulation Model* (GCM) telah digunakan dan dikembangkan sejak lama di sebagian besar pusat peramalan cuaca dan iklim utama di dunia. Seperti *Goddard Institute for Space Studies* (GISS) dari NASA, *United Kingdom Meteorological Office* (UKMO), *Geophysical Fluid Dynamic*

Laboratory (GDFL) dari NOAA, National Centers for Environmental Prediction (NCEP) dari Amerika, Commonwealth Scientific and Industrial Research Organization (CSIRO) dari Australia, dan sebagainya. Masing-masing dari General Circulation Model (GCM) tersebut memiliki perbedaan dalam bentuk resolusi spasial dan persamaan-persamaan untuk membangkitkan parameter-parameter atmosfer seperti pada tabel berikut :

Tabel 2.2 Model General Circulation Model (GCM)

No	Nama GCM	Resolusi	Jumlah Lapisan
1	GISS	4,0°X5,0°	20
2	UKMO	2,80°X3,75°	19
3	GFDL	2,8°X2,8°	18
4	CSIRO	1,875°X1,875°	18
5	NCEP	2,5°X2,5°	17
6	ECHAM	2,8125°X2,789°	18

General Circulation Model (GCM) telah dikembangkan dan digunakan di Indonesia untuk simulasi, prediksi dan pembuatan skenario iklim. General Circulation Model (GCM) juga telah dimanfaatkan untuk mempelajari variabilitas iklim dan mengkaji dampak perubahan iklim (Ratag 2001; Mole et al. 2001). Siswanto dan Ratag (2001) memprediksi curah hujan dan temperatur permukaan bulanan berbasis GCM CSIRO-9. Prediksinya dilakukan dengan menjalankan model sirkulasi global dari tahun 1949 sampai dengan 1999, menggunakan data pengamatan temperatur permukaan untuk memprediksi curah hujan tahun 2000, dengan kriteria di bawah normal, normal, dan di atas normal.

2.3 Principal Component Analyst (PCA)

Analisis peubah ganda seringkali memerlukan adanya reduksi dimensi dikarenakan himpunan data yang besar, sehingga diperlukan adanya *Principal Component Analysis (PCA)*. *Principal Component Analysis (PCA)* merupakan suatu teknik statistik yang secara linier mengubah bentuk sekumpulan variabel asli menjadi menjadi kumpulan variabel lebih kecil yang tidak berkorelasi sehingga dapat mewakili informasi dari kumpulan variabel asli (Dunteman,1989). *Principal Component Analysis (PCA)* dapat digunakan untuk menjelaskan struktur variansi-kovariansi dari sekumpulan variabel melalui beberapa variabel baru dimana variabel baru ini saling bebas, dan merupakan kombinasi linier dari variabel asal. Selanjutnya variabel baru ini dinamakan komponen utama. Morrison (1978) menyarankan agar memilih jumlah komponen utama sampai komponen-komponen utama tersebut mempunyai keragaman kumulatif 75%. Secara umum tujuan dari analisis komponen utama adalah mereduksi dimensi data dan untuk kebutuhan interpretasi. Tujuan khusus *Principal Component Analysis (PCA)* yaitu:

1. Untuk meringkas pola korelasi antar variabel yang diobservasi
2. Mereduksi sejumlah besar variabel menjadi sejumlah kecil komponen
3. Memberikan sebuah definisi oprasional (sebuah persamaan regresi) dimensi pokok penggunaan variabel yang diobservasi (Tabachnick,2001)

2.4 Principal Component Regression (PCR)

Principal component regression (pcr) merupakan teknik statistika yang dapat digunakan untuk menjelaskan pengaruh variabel prediktor terhadap variabel

respon. Metode ini akan menghasilkan komponen-komponen utama yang tidak berkorelasi. Perbedaan diantara kedua metode ini adalah jika *Principal component regression (pcr)* untuk mengetahui ada atau tidaknya hubungan variabel dependent dan independen, sedangkan *Principal component analysis (pca)* untuk meng sederhanakan variabel yang diamati dengan cara menyusutkan (mereduksi) dimensi. Perlu diketahui jika komponen utama diikut sertakan dalam model regresi, maka model yang dihasilkan sama seperti metode kuadrat terkecil. Sebaliknya jika beberapa komponen utama yang diikut sertakan maka akan diperoleh penduga dengan koefisien regresi yang bias namun memiliki *variance* yang minimum. Menurut Dreper dan Smith (1992) *Principal component regression* adalah analisis regresi dari variabel-variabel dependent atau terikat terhadap komponen utama yang tidak saling berkorelasi, dimana semua komponen utama merupakan kombinasi linier dari semua variabel independent atau bebas. Hal ini dilakukan untuk menghilangkan korelasi diantara variabel melalui transformasi variabel asal ke variabel baru (merupakan kombinasi linier dari variabel-variabel asal) yang tidak saling berkorelasi dari p hasil variabel asal dapat dibentuk p hasil komponen utama, dipilih k hasil dari komponen utama ($k < p$) yang telah menerangkan keragaman data yang cukup tinggi (antara 80% sampai 90%). Komponen utama yang dipilih adalah (k hasil) yang dapat menggantikan p hasil variabel asal tanpa mengurangi informasi.

Pembentukan regresi utama melalui analisis komponen utama ada dua cara yaitu komponen utama yang berdasarkan matriks kovarians dan komponen utama yang berdasarkan matriks korelasi. Matriks korelasi dari data yang telah

distandarisasi (bentuk baku Z) digunakan jika variabel yang diamati tidak memiliki satuan pengukuran yang sama. Sedangkan matrik kovarians digunakan jika semua variabel yang diamati mempunyai satuan pengukuran yang sama.

Analisis *Principal component regression (pcr)* adalah analisis regresi variabel dependent atau terikat terhadap komponen utama yang tidak saling berkorelasi, dengan regresi komponen utama yang dinyatakan sebagai berikut :

$$Y = w_0 + w_1K_1 + w_2K_2 + \dots + w_mK_m + \varepsilon$$

Dimana :

Y : Variabel Dependent

K : Komponen Utama

W : Parameter Regresi Utama

$K_1, K_2, K_3, \dots, K_m$ menunjukkan komponen utama yang dilibatkan dalam analisis regresi komponen utama, dengan besaran m lebih kecil dari banyaknya variabel bebas yaitu p , serta Y sebagai variabel terikat.

Komponen utama adalah gabungan linier variabel baku Z, sehingga diperoleh

$$K_1 = a_{11}Z_1 + a_{21}Z_2 + \dots + a_{p1}Z_p$$

$$K_2 = a_{12}Z_1 + a_{22}Z_2 + \dots + a_{p2}Z_p$$

⋮

$$K_m = a_{1m}Z_1 + a_{2m}Z_2 + \dots + a_{pm}Z_p$$

Jika $K_1, K_2, K_3, \dots, K_m$ dalam persamaan didistribusikan kembali kedalam persamaan regresi komponen utama, maka persamaan tersebut diperoleh :

$$\begin{aligned}
 Y &= W_0 + W_1(a_{11}Z_1 + a_{21}Z_2 + \dots + a_{p1}Z_p) + W_2(a_{12}Z_1 + a_{22}Z_2 + \dots + \\
 &\quad a_{p2}Z_p) + \dots + W_m(a_{1m}Z_1 + a_{2m}Z_2 + \dots + a_{pm}Z_p) + \varepsilon \\
 &= W_0 + W_1a_{11}Z_1 + W_1a_{21}Z_2 + \dots + W_1a_{p1}Z_p + W_2a_{12}Z_1 + W_2a_{22}Z_2 \\
 &\quad + \dots + W_2a_{p2}Z_p + W_ma_{1m}Z_1 + W_ma_{2m}Z_2 + W_ma_{pm}Z_p + \varepsilon \\
 &= W_0 + (W_1a_{11} + W_2a_{12} + \dots + W_ma_{1m})Z_1 + (W_1a_{21} + W_2a_{22} + \dots + \\
 &\quad W_ma_{2m})Z_2 + \dots + (W_1a_{p1} + W_2a_{p2} + \dots + W_ma_{pm})Z_p + \varepsilon
 \end{aligned}$$

Maka dari persamaan diatas diperoleh persamaan regresi dengan komponen utama sebagai berikut :

$$Y = b_0 + b_1Z_1 + \dots + b_pZ_p$$

Dengan :

$$b_0 = W_0$$

$$b_1 = W_1a_{11} + W_2a_{12} + \dots + W_ma_{1m}$$

$$b_2 = W_1a_{21} + W_2a_{22} + \dots + W_ma_{2m}$$

⋮

$$b_p = W_1a_{p1} + W_2a_{p2} + \dots + W_ma_{pm}$$

Syarat Principal component regression (pcr) adalah sebagai berikut :

1. Menghitung nilai eigen vektor dan nilai eigen value pada matrik korelasi atau kovarians
2. Terdapat p komponen utama yang orthogonal dan tidak terkorelasi
3. Dipilih komponen yang eigen value >1 atau yang mampu menerangkan keragaman cukup tinggi (70%-80%)
4. Regresi variabel dependent atau terikat dengan komponen-komponen utama yang telah terpilih.

2.5 Curah Hujan

Hujan merupakan gejala meteorologi dan juga unsur klimatologi. Hujan adalah hydrometer yang jatuh berupa partikel-partikel air yang mempunyai diameter 0,5 mm atau lebih. Hydrometer yang jatuh ke tanah disebut hujan sedangkan yang tidak sampai ke permukaan tanah dapat di ukur dengan jalan mengukur tinggi air hujan tersebut dengan berdasarkan volume air hujan per satuan luas. Curah hujan merupakan salah satu unsur cuaca yang datanya diperoleh dengan cara mengukurnya dengan menggunakan alat penakar hujan, sehingga dapat diketahui jumlahnya dalam satuan milimeter (mm). Curah hujan 1 mm adalah jumlah air hujan yang jatuh ke permukaan per satuan luas (m^2) dengan catatan tidak ada yang menguap, meresap atau mengalir. Jadi, curah hujan sebesar 1 mm setara dengan 1 liter/ m^2 (Aldrian,E.dkk,2011). Curah hujan dibatasi sebagai tinggi air hujan yang diterima di permukaan sebelum mengalami aliran permukaan, evaporasi dan peresapan ke dalam tanah.

Curah hujan diartikan sebagai jumlah hujan dalam satu satuan waktu yang dinyatakan dalam satuan milimeter per jam (mm/jam) (LAPAN, 2014). Menurut Sosrodarsono dan Takeda (2003), hujan dapat digolongkan menjadi beberapa golongan berdasarkan intensitas curah hujannya. Penggolongan hujan berdasarkan intensitas curah hujannya dapat dilihat pada Tabel berikut :

Tabel 2.3 Curah Hujan

Deret Hujan	Intensitas Curah Hujan
Hujan sangat lemah	$<0,02$
Hujan lemah	$0,02 < CH \leq 0,05$
Hujan normal	$0,05 < CH \leq 0,25$
Hujan deras	$0,25 < CH \leq 1,00$
Hujan sangat deras	$>1,00$

Intensitas curah hujan adalah jumlah curah hujan yang dinyatakan dalam tinggi hujan atau volume hujan tiap satuan waktu air hujan terkonsentrasi (Wesli, 2008). Besarnya intensitas curah hujan berbeda-beda tergantung dari lamanya curah hujan dan frekuensi kejadiannya. Secara umum curah hujan di Indonesia dipengaruhi oleh beberapa fenomena alam, antara lain gelombang El-Nino, monsun Asia-Australia, sirkulasi atmosfer lokal, regional dan global, seperti sirkulasi utara-selatan, sirkulasi barat-timur dan sistem angin lokal. Apabila salah satu dari sistem sirkulasi tersebut terganggu, maka secara otomatis akan mempengaruhi cuaca dan iklim di Indonesia (LAPAN, 2014).

Keragaman curah hujan berdasarkan skala ruang dipengaruhi oleh dua faktor, yaitu ketinggian tempat dan angin. Semakin rendah ketinggian suatu tempat maka potensi curah hujan yang diterima akan semakin banyak dikarenakan tempat tersebut memiliki suhu yang rendah. Adapaun faktor angin dibedakan menjadi dua,

yaitu angin zonal dan angin meridional. Angin zonal merupakan komponen angin barat-timur yang berpengaruh terhadap fenomena *Dipole Mole* (DM) yang terjadi dikarenakan adanya pergerakan massa udara dari barat ke timur Samudera Hindia ataupun sebaliknya. Angin meridional berpengaruh terhadap fenomena *Dipole Mole* (DM) sebagai akibat dari adanya aliran udara antara wilayah India bagian selatan dengan wilayah Australia bagian selatan barat. Adapun *Dipole Mole* (DM) merupakan gejala naiknya suhu permukaan laut dari kondisi normal di sepanjang ekuator Samudera Hindia khususnya di sebelah selatan India yang diikuti dengan menurunnya suhu permukaan laut pada kondisi tidak normal di perairan Indonesia khususnya di wilayah barat Pulau Sumatera (LAPAN, 2014).

Penurunan curah hujan yang cukup besar pada musim kemarau termasuk akibat El Nino telah membawa dampak terhadap produksi tanaman pangan. Mundurnya awal musim hujan selama 30 hari pada tahun El Nino telah menurunkan produksi beras pada musim hujan (Januari-April) sekitar 6,5% di Jawa Barat dan Jawa Tengah, dan 11,0% di Jawa Timur dan Bali. Menurut Las anomali iklim dan curah hujan berkorelasi di sebagian besar wilayah di mana lebih dari 67% diantaranya merupakan sentra produksi padi. Berdasarkan data historis tentang dampak El Nino terhadap produksi beras nasional mengindikasikan bahwa produksi beras nasional sangat rentan terhadap kejadian iklim ekstrim. Kekeringan periode 1980-1990 telah menyebabkan kehilangan produksi beras tiga kali lipat dibandingkan pada periode 1991-2000, yaitu penurunan produksi dari 100 ribu ton per tahun per kabupaten menjadi 300 ribu ton per tahun di 4 kabupaten.

Berdasarkan data historis tentang dampak El Nino terhadap produksi beras nasional mengindikasikan bahwa produksi beras nasional sangat rentan terhadap kejadian iklim ekstrim. Kekeringan periode 1980-1990 telah menyebabkan kehilangan produksi beras tiga kali lipat dibandingkan pada periode 1991-2000, yaitu penurunan produksi dari 100 ribu ton per tahun per kabupaten menjadi 300 ribu ton per tahun per kabupaten. El-Nino merupakan fenomena global dari sistem interaksi laut dan atmosfer yang ditandai dengan memanasnya suhu permukaan air laut di Pasifik ekuator, dimana suhu permukaan air laut di daerah tersebut lebih panas dari suhu rata-ratanya. Namun, tidak seluruh wilayah Indonesia terkena dampak fenomena El-Nino dikarenakan luasnya wilayah Indonesia dan posisi topografisnya yang dikenal sebagai benua maritim. Sebaliknya, fenomena La-Nina ditandai dengan naiknya suhu permukaan laut di Samudera Pasifik tengah dan timur, dimana suhu permukaan air laut di daerah tersebut lebih panas dari suhu rata-ratanya, sehingga mengakibatkan curah hujan bertambah di wilayah Indonesia.

Sebagai salah satu wilayah tropis yang unik dinamika atmosfernya dimana banyak dipengaruhi oleh kehadiran angin pasat, angin monsun, iklim maritim, dan penaruh berbagai kondisi lokal, maka cuaca dan iklim di Indonesia diduga memiliki karakteristik khusus. Secara umum curah hujan di Indonesia didominasi oleh adanya pengaruh unsur iklim lain seperti temperatur atau suhu yang menyebabkan perbedaan suhu daratan dan perairan, apabila suhu tanah atau daratan lebih tinggi daripada perairan maka hujan akan sering terjadi di perairan dan sebaliknya. Selain itu arah angin juga menjadi faktor pengaruh bagi curah hujan. Angin berperan dalam memindahkan awan dari satu tempat ke tempat lain. Wilayah

yang memiliki angin yang lemah memiliki kemungkinan jarang terjadi hujan dan wilayah yang memiliki curah hujan terbanyak adalah wilayah yang berada di lintang rendah atau mendekati garis khatulistiwa.

2.6 Validasi Model

Validasi model merupakan suatu proses untuk mengetahui apakah model yang terbentuk menghasilkan nilai prediksi yang akurat. Dalam melakukan pemodelan data dibagi menjadi dua bagian, yaitu data *in-sample* (*data training*) dan data *out-sample* (*data testing*). Data *in-sample* digunakan untuk mendapatkan model, sedangkan data *out-sample* digunakan untuk melakukan validasi terhadap model tersebut. Pemilihan model terbaik berdasarkan data *in-sample* dilakukan menggunakan *Root Mean Square Error* (RMSE), sedangkan data *out-sample* dilakukan menggunakan *Root Mean Square Error Prediction* (RMSEP). Selain itu pemilihan model terbaik berdasarkan data *out-sample* juga dapat dilakukan menggunakan $R^2_{\text{Prediction}}$ atau koefisien determinasi. Semakin kecil nilai RMSE, RMSEP menunjukkan bahwa model yang didapatkan semakin baik. Sebaliknya, semakin besar nilai $R^2_{\text{Prediction}}$ menunjukkan bahwa model yang didapatkan semakin baik.

Root Mean Square Error (RMSE) merupakan akar kuadrat dari rata-rata kuadrat eror, dimana error adalah selisih dari data *in-sample* dan data hasil pemodelan. Kriteria *Root Mean Square Error* (RMSE) yaitu membandingkan data *in-sample* dengan taksiran data hasil pemodelan. Adapaun *Root Mean Square Error* (RMSE) secara matematis dapat dituliskan sebagai berikut :

$$RMSE = \sqrt{\frac{1}{n_1} \sum_{i=1}^{n_1} (Y_i - \hat{Y}_i)^2} \quad (2.14)$$

Dimana Y_i merupakan data *in-sample* ke- i , \hat{Y}_i merupakan taksiran data hasil pemodelan data *in-sample* ke- i dan n merupakan jumlah data *in-sample*.

Root Mean Square Error Prediction (RMSEP) merupakan akar kuadrat dari rata-rata kuadrat *error*, dimana *error* adalah selisih dari data *out-sample* dan data hasil peramalan. Sedangkan $R^2_{\text{Prediction}}$ merupakan proporsi keragaman total nilai-nilai variabel respon yang dapat diterangkan oleh variabel prediktor dalam model yang digunakan. Kriteria RMSEP dan $R^2_{\text{Prediction}}$ dilakukan dengan membandingkan data *out-sample* dengan data hasil peramalan. Adapun *Root Mean Square Error Prediction* (RMSEP) dan $R^2_{\text{Prediction}}$ secara berturut-turut dapat dihitung menggunakan rumus sebagai berikut.

$$RMSEP = \sqrt{\frac{1}{n_2} \sum_{j=1}^{n_2} (Y_j - \hat{Y}_j)^2} \quad (2.15)$$

$$R^2_{\text{Prediction}} = \left[\frac{\sqrt{\frac{1}{n_2} \sum_{j=1}^{n_2} (Y_j - \hat{Y}_j)^2}}{\sqrt{\frac{1}{n_2} \sum_{j=1}^{n_2} (Y_j - \bar{Y})^2}} \right] \times 100\% \quad (2.16)$$

Dimana Y_j merupakan data *out-sample* ke- j , \hat{Y}_j merupakan taksiran data hasil pemodelan data *out-sample* ke- j , \bar{Y} merupakan rata-rata variabel respon dan n_2 merupakan jumlah data *out-sample*.