

## VERIFIKASI MODEL PEMBANGKIT DATA HUJAN DAN SUHU UDARA HARIAN DI TANGERANG SELATAN VERIFICATION GENERATE DAILY RAINFALL AND AIR TEMPERATURE MODEL IN TANGERANG SELATAN

**Qurrata A'yun Kartika<sup>1\*</sup>**

1) Tangerang Selatan, Jl. Raya Kodam Bintaro No. 82 Kel. Pondok Betung  
Kec. Pondok Aren, Kota Tangerang Selatan, 15221

\*Email: [ayun.kartika@gmail.com](mailto:ayun.kartika@gmail.com)

### ABSTRAK

Beberapa kesalahan dan kekosongan dapat terjadi dalam pengumpulan data. Data suhu udara dan curah hujan harian sangat penting dalam menunjang perencanaan dan evaluasi aktivitas manusia. Kekosongan pada data iklim dapat dilakukan dengan menggunakan model pembangkit data iklim. Model pembangkit dapat dilakukan dengan metode stokastik sesuai dengan sebaran data. Suhu udara dan curah hujan harian dapat diestimasi dengan membangkitkan nilai berdasarkan sebaran gamma dari data bulanan dan harian. Tangerang Selatan sebagai kota dengan kepadatan penduduk yang tinggi membuat pentingnya kelengkapan data. Berdasarkan hasil pengolahan, scatter plot suhu udara cenderung memiliki pola yang acak terhadap observasi dan estimasi curah hujan overestimate terhadap curah hujan observasi. suhu udara hasil estimasi memiliki korelasi sebesar 0,589 hingga 0,596 sedangkan untuk curah hujan sebesar 0,414 hingga 0,543 yang mana dalam kategori korelasi cukup hingga kuat, RMSE untuk suhu sebesar 0,765 hingga 0,785 dan curah hujan sebesar 10,131 hingga 11,174 dan RSTD suhu udara sebesar 0,355 hingga 0,526 dan curah hujan sebesar 1,563 hingga 2,024.

Kata kunci: Model Pembangkit data Iklim, scatterplot, korelasi, RMSE, dan RSTD.

### ABSTRACT

Some errors and voids can occur in data collection. Data on air temperature and daily rainfall are very important in supporting the planning and evaluation of human activities. Blanks in climate data can be done using climate data generator models. The generator model can be done using the stochastic method according to the distribution of the data. Air temperature and daily rainfall can be estimated by generating values based on the gamma distribution of monthly and daily data. South Tangerang as a city with a high population density makes data completeness important. Based on the results of the processing, the scatter plot of air temperature tends to have a random pattern to the observations and the estimated rainfall overestimates the observed rainfall. air temperature estimation results have a correlation of 0,589 to 0,596 while for rainfall it is 0,414 to 0,543 which is in the category of moderate to strong correlation, RMSE for temperature is 0,765 to 0,785 and rainfall is 10,131 to 11,174 and RSTD air temperature is 0,355 to 0,526 and rainfall of 1,563 to 2,024.

Keywords: Climate data generation model, scatterplot, correlation, RMSE, and RSTD.

## 1. Pendahuluan.

Data iklim harian seringkali digunakan untuk berbagai kepentingan, seperti perencanaan pembangunan, perencanaan kegiatan, klaim asuransi dan lain sebagainya. Data iklim tersebut dapat berupa keadaan cuaca, curah hujan, suhu udara dan lain-lain. Namun demikian, seringkali kelengkapan data menjadi rendah akibat berbagai faktor, seperti bencana alam, *error* alat, kesalahan paralaks dan sebagainya. Dengan demikian dibutuhkan suatu metode yang dapat digunakan untuk menangani data *error* atau data kosong, terutama data suhu udara dan hujan harian yang cukup sensitif terhadap suatu fenomena dengan durasi yang singkat.

Pembangkit data sering digunakan untuk mengisi data kosong. Di Eropa, berbagai usaha telah dilakukan dalam melakukan pembangkitan data termasuk *Climate Data Record* (CDR) untuk berbagai kepentingan [1]. Setiap model memiliki unsur *error*, skill dan ketidakpastian. Ketidakpastian parameter harian menjadi masalah yang menantang [2].

Kota Tangerang Selatan dengan kepadatan penduduk yang tinggi dan pembangunan yang cepat membuat informasi iklim harian sangat diperlukan. Namun diperlukan verifikasi model pembangkit iklim untuk mengetahui bagaimana kesesuaian estimasi dari model pembangkit iklim terhadap observasinya. Validasi bertujuan untuk mengetahui sejauh mana kesesuaian atau kemampuan model terhadap observasi [3].

## 2. Data dan Metode.

### 2.1. Data.

Data yang digunakan disini ialah data suhu udara harian dan bulanan tahun 2019 dan curah hujan harian dan

bulanan observasi di Stasiun Klimatologi Tangerang Selatan selama 40 tahun, yaitu tahun 1981-2020. Stasiun Klimatologi Tangerang sendiri berada pada koordinat 106,76°BT dan 6,25°LS pada elevasi 26 m yang mewakili Kota Tangerang Selatan.

### 2.2. Metode.

Metode yang digunakan dalam penelitian ini dilakukan dengan dua cara dengan pada suhu udara harian yang terdistribusi secara normal dan curah hujan harian dengan sifat distribusi gamma. Pengolahan data dilakukan dengan menggunakan Microsoft excel. Model pembangkit data suhu udara dan curah hujan yang digunakan disini sesuai dengan yang dilakukan oleh Boer [4].

#### 2.2.1. Suhu Udara.

Estimasi suhu udara dilakukan berdasarkan persamaan (1) dan (2).

$$Y_{suhu} = a_0 + \sum_{k=1}^n [a_k \sin(kt') + b_k \cos(kt')] \quad (1)$$

$$t' = \frac{2\pi t}{12}; t = (T - 0,5) + \frac{(m-0,5)}{D} \quad (2)$$

dengan

$Y_{suhu}$	= suhu udara
$n$	= jumlah harmonik
$a_0$	= <i>intercept</i>
$a_k, b_k$	= koefisien
$T$	= Bulan
$m$	= tanggal
$D$	= jumlah hari dalam bulan

Koefisien dan *intercept* dibuat dengan membuat persamaan regresi linear berganda.

#### 2.2.2. Curah Hujan.

Curah hujan harian terdistribusi secara gamma, sehingga pendekatan dapat dilakukan dengan pendekatan distribusi gamma. Estimasi data hujan

harian dilakukan dengan langkah-langkah sebagai berikut:

- [1] Membuat peluang kejadian hujan dan transformasi peluang ke bentuk fungsi logit sesuai persamaan (3), (4), (5) dan (6)

$$p_{01}(i) = \frac{n_{01}(i)}{n_{01}(i)+n_{00}(i)} \quad (3)$$

$$p_{11}(i) = \frac{n_{11}(i)}{n_{11}(i)+n_{10}(i)} \quad (4)$$

$$g_{01}(i) = \ln\left(\frac{p_{01}(i)}{1-p_{01}(i)}\right) \quad (5)$$

$$g_{11}(i) = \ln\left(\frac{p_{11}(i)}{1-p_{11}(i)}\right) \quad (6)$$

Dengan:

$n_{00}(i)$  = banyaknya kejadian tidak hujan pada hari ke  $i$  dan hari sebelumnya tidak hujan

$n_{01}(i)$  = banyaknya kejadian tidak hujan pada hari ke  $i$  dan hari sebelumnya hujan

$n_{10}(i)$  = banyaknya kejadian hujan pada hari ke  $i$  dan hari sebelumnya tidak hujan

$n_{11}(i)$  = banyaknya kejadian hujan pada hari ke  $i$  dan hari sebelumnya hujan

$p_{01}(i)$  = peluang kejadian hujan pada hari ke  $i$  dan hari sebelumnya tidak hujan

$p_{11}(i)$  = kejadian hujan pada hari ke  $i$  dan hari sebelumnya hujan

$g_{01}(i)$  = nilai logit untuk kejadian hujan pada hari ke  $i$  dan hari sebelumnya tidak hujan

$g_{11}(i)$  = nilai logit untuk kejadian hujan pada hari ke  $i$  dan hari sebelumnya hujan

- [2] Membuat kurva penyesuaian untuk nilai peluang yang ditransformasi sesuai persamaan (7) dan (8).

$$g_{jl}(i) = a_{jlo}(i) + \sum_{k=1}^m [a_{jlk} \sin(ki') + b_{jlk} \cos(ki')] \quad (7)$$

$$i' = \frac{2\pi t}{365}; t = 1, 2, 3, \dots, 365/366 \quad (8)$$

Dengan:

$g_{jl}(i)$  = nilai logit untuk  $jl$

$m$  = jumlah harmonik

$a_{jlo}$  = *intercept*

$a_{jlk}, b_{jlk}$  = koefisien

$jl$  = notasi kejadian bersyarat (0 untuk kejadian tidak hujan dan 1 untuk kejadian hujan)

- [3] Transformasi nilai dugaan ke bentuk nilai peluang sesuai persamaan (9)

$$p_{jl}(i) = \frac{1}{\{1 + \exp[-g_{jl}(i)]\}} \quad (9)$$

- [4] Transformasi data peluang kejadian hujan ke bentuk kejadian hujan dengan menggunakan sebaran Uniform (0,1) dengan ketentuan jika  $p_{jl}(i) < u(0,1)$  maka tidak terjadi hujan dan jika  $p_{jl}(i) > u(0,1)$  maka terjadi hujan.

- [5] Mengestimasi data hujan dengan sebaran gamma dengan persamaan untuk model kepekatannya yaitu sebagai berikut :

$$f(x; \alpha, \beta) = \frac{\left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)}{\beta \Gamma(\alpha)} \quad (10)$$

Dengan:

$$y = \ln(x) - 1/n \sum_{i=1}^n \ln(x_i) \quad (11)$$

Jika  $0 \leq y \leq 0,5772$  maka  $\alpha = \frac{(0,5000876+0,1648852y-0,0544276y^2)}{y} \quad (12)$

Jika  $0,5772 < y \leq 17$  maka  $\alpha = \frac{(8,898919+9,05995y-0,9775373y^2)}{y(17,79728+11,968477y+y^2)} \quad (13)$

Jika  $y > 17$  maka  $\alpha = 1/y \quad (14)$

$$\beta = \bar{x}/\alpha \quad (15)$$

- $\alpha$  = parameter skala
- $\beta$  = parameter bentuk
- $\Gamma$  = fungsi gamma

$\alpha$  dan  $\beta$  disini dibagi sesuai musim hujan dan musim kemarau. Musim hujan dalam penelitian ini diasumsikan terjadi pada Januari hingga maret dan oktober hingga desember, sedangkan musim kemarau diasumsikan terjadi pada April hingga September. Dalam penelitian ini, estimasi curah hujan dilakukan dengan menggunakan *Crystal Ball*.

### 2.2.3. Analisis kesesuaian.

Dalam memestimasi data suhu udara dan curah hujan, diperlukan adanya uji kesesuaian terkait bagaimana kemiripan dan *error* yang dihasilkan. Dalam penelitian ini, uji kesesuaian dilakukan dengan metode sebagai berikut:

a. *Scatter Plot*

*Scatter plot* juga dapat digunakan untuk menganalisis bagaimana penilaian estimasi suhu udara dan curah hujan secara visual. Ketika estimasi suhu udara dinilai lebih besar dari suhu udara observasi maka disebut dengan *Overestimate*, sedangkan ketika estimasi suhu udara dinilai lebih kecil dari suhu udara observasi maka disebut dengan *Underestimate*. *Overestimate* dan *Underestimate* juga dapat diamati dari sebaran *plot* data terhadap garis linear pada masing-masing parameter yang bernilai sama.

b. Korelasi Pearson

Korelasi digunakan untuk mengetahui hubungan statistik

antara data estimasi dan data observasi stasiun. Korelasi yang digunakan dalam penelitian ini ialah korelasi Pearson. Menurut Sugiyono [5], klasifikasi dari besaran nilai korelasi Pearson ditunjukkan pada tabel 1.

**Tabel 1.** Klasifikasi korelasi pearson.

Koefisien korelasi	Kekuatan Hubungan
0,000-0,199	Sangat lemah
0,200-0,399	lemah
0,400-0,599	Cukup Kuat
0,600-0,799	Kuat
0,800-1,00	Sangat kuat

c. *Root Mean Square Error (RMSE)*

RMSE atau akar rata-rata *error* kuadrat digunakan untuk melihat sejauh mana *error* data estimasi menyimpang terhadap nilai observasi stasiun.

d. *Ratio Standard Deviation (RSTD)*

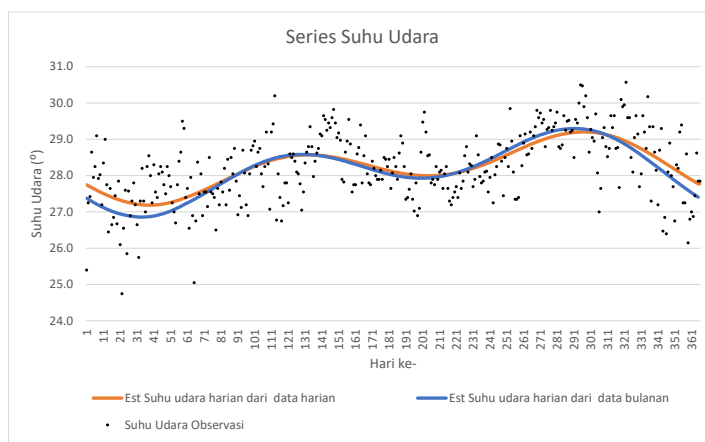
RSTD adalah rasio standar deviasi yang menggambarkan perbandingan variasi antara data estimasi dan observasi. Nilai sempurna RSTD ialah 1 yang menggambarkan kesamaan variasi dua data.

$$RSTD = \frac{\sigma_{Estimasi}}{\sigma_o}$$

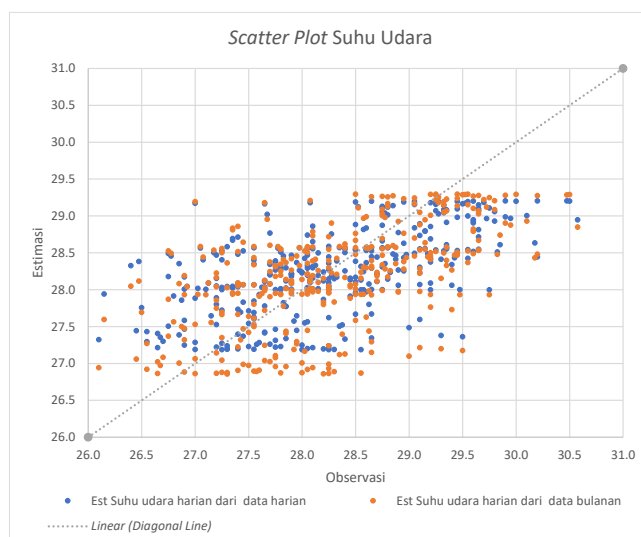
$\sigma_{Estimasi}$  = standar deviasi estimasi  
 $\sigma_o$  = standar deviasi observasi

### 3. Pembahasan dan Diskusi.

#### 3.1. Suhu udara.



**Gambar 1.** Estimasi Data Suhu Udara Harian di Stasiun Klimatologi Tangerang Selatan.



**Gambar 2.** Sebaran suhu udara estimasi terhadap observasi di Stasiun Klimatologi Tangerang Selatan.

Berdasarkan gambar 1 dapat diketahui suhu udara yang dihasilkan dari estimasi cukup menggambarkan pola umum suhu udara observasi. Berdasarkan estimasi suhu udara dari data bulanan dan harian di Stasiun Klimatologi Tangerang Selatan terdapat dua puncak dan dua lembah baik dari estimasi bulanan maupun harian. Puncak pertama terjadi pada awal mei dan puncak kedua terjadi pada pertengahan oktober. Gambar 2

menggambarkan sebaran estimasi suhu udara terhadap suhu udara observasi. Berdasarkan sebaran *scatter plot* diketahui bahwa estimasi suhu udara memiliki pola antara *underestimate* atau *overestimate* yang tidak terlalu jelas. Kesesuaian antara data estimasi dengan data observasi dijelaskan pada tabel 2. Estimasi suhu udara dari data harian lebih baik daripada bulanan berdasarkan korelasi pearson dan RMSE dengan

perbedaan yang kecil, sedangkan berdasarkan RSTD estimasi suhu udara dari data bulanan lebih baik. Hal ini menunjukkan bahwa secara pola

dan *error* estimasi suhu udara dari data harian lebih baik, namun untuk menggambarkan variasi estimasi suhu udara dari data bulanan lebih baik.

**Tabel 2.** Analisis Kesesuaian Estimasi Data Suhu Udara dengan suhu udara harian di Stasiun Klimatologi Tangerang Selatan.

Analisis Kesesuaian	Estimasi suhu udara harian dari data harian	Estimasi suhu udara harian dari data bulanan
Korelasi Pearson	0,596	0,589
RMSE	0,765	0,785
RSTD	0,355	0,526

### 3.2. Curah Hujan

Curah hujan di Stasiun Klimatologi Tangerang Selatan memiliki pola hujan muson (Staklim Tangsel, 2020). Dalam estimasi curah hujan harian diperlukan koefisien regresi dari fungsi logit yang ditunjukkan pada tabel 3 untuk Stasiun Klimatologi Tangerang Selatan. Fungsi ini kemudian disubstitusikan kembali pada persamaan 7 sebagai penyesuaian fungsi logit sehingga didapat pola yang

ditunjukkan pada gambar 3. Berdasarkan gambar 3 dapat diketahui variasi nilai logit untuk kondisi tidak hujan sebelum hari hujan lebih besar daripada variasi fungsi nilai untuk kondisi hujan sebelum hari hujan. Jika diamati, pola nilai logit mirip dengan pola hujan di Stasiun Klimatologi Tangerang Selatan. Senada dengan penelitian yang dilakukan oleh Sudiar [6] dimana pola hujan ekuatorial di Padang memiliki pola yang sama dengan nilai logit.

**Tabel 3.** Koefisien Regresi Estimasi Curah Hujan Harian di Stasiun Klimatologi Tangerang Selatan.

	variabel	$a_0$	$b_1$	$c_1$	$b_2$	$c_2$
Harian	g01	-0,88	0,646	0,767	-0,24	-0,02
	g11	0,002	0,253	0,649	0,019	0,033
Bulanan	g01	-0,825	0,324	0,469	-0,086	-0,033
	g11	0,371	0,213	0,308	-0,057	-0,021



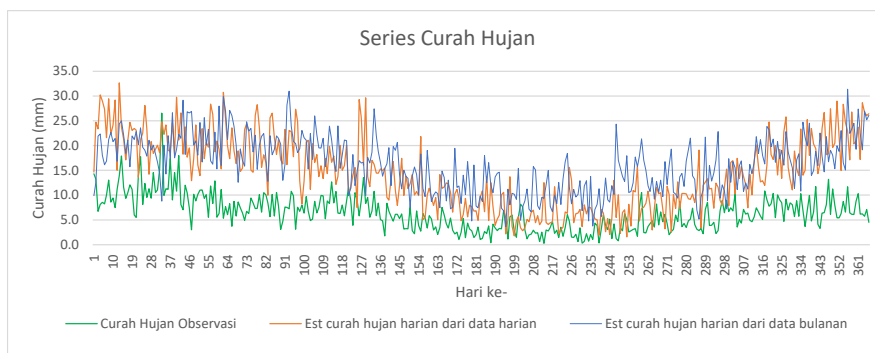
**Gambar 3.** Kurva Fitting g01 dan g11 Curah Hujan Harian di Stasiun Klimatologi Tangerang Selatan.

Tabel 4 menggambarkan  $\alpha$  dan  $\beta$  sebagai parameter skala dan bentuk yang digunakan untuk mengestimasi sebaran gamma hujan dengan menggunakan crystal ball. Dengan menggunakan crystal ball didapat estimasi curah hujan harian. Curah hujan harian rata-rata selama 40 tahun yang ditunjukkan pada gambar 4. Pola hujan harian cenderung menggambarkan

musiman. Curah hujan observasi tertinggi terjadi pada 2 februari sebesar 26,6 mm dan terendah terjadi pada 1 agustus dan 19 agustus sebesar 0,3 mm. Curah hujan estimasi dari data harian terjadi pada 13 Januari sebesar 32,7 mm dan terendah terjadi pada 9 September sebesar 1,6 mm. Curah hujan estimasi dari data bulanan terjadi pada 22 Desember sebesar 31,4 mm dan terendah terjadi pada 4 Juni sebesar 3,7 mm.

**Tabel 4.** Parameter sebaran gamma curah hujan harian di Stasiun Klimatologi Tangerang Selatan.

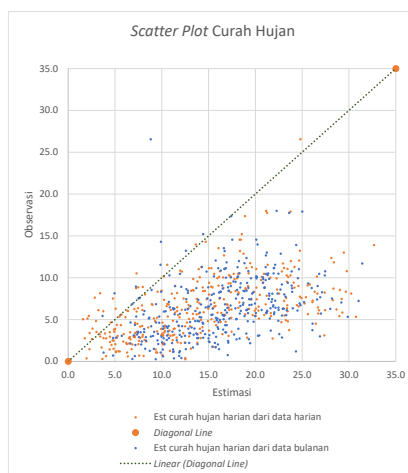
Parameter	Musim	g01	g11
$\alpha$	Hujan	0,86	0,91
	Kemarau	17,48	18,46
$\beta$	Hujan	0,86	0,88
	Kemarau	19,73	18,78



**Gambar 4.** Grafik Estimasi Data Curah Hujan Harian di Stasiun Klimatologi Tangerang Selatan

Gambar 5. Menggambarkan *scatter plot* data estimasi dengan data curah hujan. Berdasarkan gambar tersebut diketahui bahwa hampir secara

keseluruhan data estimasi *overestimate* terhadap data observasi yang dilihat berdasarkan sebaran data yang berada di bawah garis diagonal.



**Gambar 5.** *Scatter Plot* Estimasi Curah Hujan Harian terhadap Observasi di Stasiun Klimatologi Tangerang Selatan.

Analisis kesesuaian data estimasi curah hujan dari data harian dan bulanan ditunjukkan pada Tabel 5. Berdasarkan korelasi dan RMSE estimasi curah hujan dari data harian lebih baik daripada estimasi curah hujan dari data bulanan. Namun

demikian korelasi yang termasuk kuat untuk estimasi curah hujan dari data harian dariada estimasi curah hujan dari data bulanan yang berada pada kategori cukup. Namun berdasarkan RSTD estimasi curah hujan dari data bulanan lebih baik daripada estimasi curah hujan dari data harian.

**Tabel 5.** Analisis Kesesuaian estimasi Data curah hujan dengan curah hujan harian di Stasiun Klimatologi Tangerang Selatan.

Analisis Kesesuaian	Estimasi curah hujan harian dari data harian	Estimasi curah hujan harian dari data bulanan
Korelasi Pearson	0,543	0,414
RMSE	10,131	11,174
RSTD	2,024	1,563

#### 4. Kesimpulan.

Estimasi suhu udara dari model pembangkit cukup menggambarkan pola observasi dengan dua puncak dan dua lembah. Hal ini dikuatkan oleh korelasi pearson sebesar 0,6 yang termasuk dalam kategori kuat, dengan RMSE berkisar antara 0,765 hingga 0,785 namun RSTD yang cukup jauh. Secara umum, estimasi suhu udara dari data harian lebih baik daripada estimasi suhu udara dari data bulanan.

Estimasi curah hujan dari model pembangkit cukup menggambarkan pola observasi dengan pola musonal. Hal ini dikuatkan oleh korelasi pearson sebesar 0,4 hingga 0,5 yang termasuk dalam kategori cukup hingga kuat, dengan RMSE berkisar antara 10,131 hingga 11,174 namun RSTD yang sangat jauh terutama untuk estimasi curah hujan dari data harian dimana variasinya mencapai dua kali curah hujan observasi. Namun demikian, secara umum, estimasi curah hujan dari data harian masih lebih baik daripada estimasi curah hujan dari data bulanan.

#### Ucapan Terima Kasih.

Saya ucapkan terimakasih kepada Prof. Dr. Rizaldi Boer yang telah pengajaran dan bimbingan sehingga penulisan ini dapat diselesaikan, Bapak Kepala Stasiun dan Bapak Kepala Seksi Observasi dan Informasi serta rekan-rekan di Tangerang Selatan yang telah memberi informasi data serta dukungan penuh kepada penulis.

#### Daftar Pustaka.

[1] Su, Z., Timmermans, W., Zeng, Y., Schulz, J., John, V. O., Roebeling, R. A., ... & Wen, J. (2018). An overview of European efforts in generating climate data records. *Bulletin of the American*

*Meteorological Society*, 99(2), 349-359.

- [2] Srikanthan, R., & McMahon, T. A. (2001). Stochastic generation of annual, monthly and daily climate data: A review. *Hydrology and earth system sciences*, 5(4), 653-670.
- [3] Hargreaves, J. C. (2010). Skill and uncertainty in climate models. *Wiley Interdisciplinary Reviews: Climate Change*, 1(4), 556-564.
- [4] Boer, R. (tanpa tahun). Model Pembangkit Data Iklim. Kapita Selektta Agroklimatologi. Dirjen Dikti. Dikbud.
- [5] Sugiyono. (2012). Statistika Untuk Penelitian. CV Alfabeta: Bandung.
- [6] Sudiar (2015). Model Pembangkit Data Curah Hujan: Studi Kasus Stasiun Simpang Alai Kota Padang. Jurnal Saintek.