



TUGAS AKHIR - KS184822

***STATISTICAL DOWNSCALING* OUTPUT GCM  
DENGAN PENDEKATAN REGRESI GAMMA UNTUK  
PREDIKSI CURAH HUJAN SEBAGAI DASAR  
PREDIKSI BANJIR GENANGAN DI SURABAYA**

RISDA IKFINA PUTRI  
NRP 062115 4000 0065

Dosen Pembimbing  
Santi Puteri Rahayu, S.Si., M.Si., Ph.D.  
Dr.rer.pol. Dedy Dwi Prastyo, S.Si., M.Si.

PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019





**TUGAS AKHIR - KS184822**

***STATISTICAL DOWNSCALING* OUTPUT GCM  
DENGAN PENDEKATAN REGRESI GAMMA  
UNTUK PREDIKSI CURAH HUJAN SEBAGAI  
DASAR PREDIKSI BANJIR GENANGAN DI  
SURABAYA**

**RISDA IKFINA PUTRI  
NRP 062115 4000 0065**

**Dosen Pembimbing  
Santi Puteri Rahayu, S.Si., M.Si., Ph.D.  
Dr. rer. pol. Dedy Dwi Prastyo, S.Si., M.Si.**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**





**FINAL PROJECT - KS184822**

**STATISTICAL DOWNSCALING OF GCM  
OUTPUT USING GAMMA REGRESSION TO  
PREDICT RAINFALL FOR FLOOD PREDICTION  
IN SURABAYA**

**RISDA IKFINA PUTRI  
SN 062115 4000 0065**

**Supervisors**

**Santi Puteri Rahayu, M.Si., Ph.D**

**Dr. rer. pol. Dedy Dwi Prastyo, S.Si., M.Si.**

**UNDERGRADUATE PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2019**



**LEMBAR PENGESAHAN**

**STATISTICAL DOWNSCALING OUTPUT GCM DENGAN  
PENDEKATAN REGRESI GAMMA UNTUK PREDIKSI  
CURAH HUJAN SEBAGAI DASAR PREDIKSI BANJIR  
GENANGAN DI SURABAYA**

**TUGAS AKHIR**

Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Statistika  
pada  
Program Studi Sarjana Departemen Statistika  
Fakultas Matematika, Komputasi, dan Sains Data  
Institut Teknologi Sepuluh Nopember

Oleh:

**Risda Ikfina Putri**  
NRP.062115 4000 0065

Disetujui oleh Pembimbing:

**Santi Puteri Rahayu, S.Si., M.Si., Ph.D.**

NIP. 19750115 199903 2 003

**Dr. rer. pol. Dedy Dwi Prastyo, S.Si., M.Si.**

NIP. 19831204 200812 1 002

(*[Signature]*)  
(*[Signature]*)

Mengetahui,  
Kepala Departemen Statistika



**Dr. Suhartono**

NIP. 19710929 199512 1 001

SURABAYA, JULI 2019





# **STATISTICAL DOWNSCALING OUTPUT GCM DENGAN PENDEKATAN REGRESI GAMMA UNTUK PREDIKSI CURAH HUJAN SEBAGAI DASAR PREDIKSI BANJIR GENANGAN DI SURABAYA**

**Nama** : Risa Ikfina putri  
**NRP** : 062115 4000 0065  
**Departemen** : Statistika  
**Dosen Pembimbing** : Santi Puteri Rahayu, M.Si., Ph.D  
Dr. rer. pol. Dedy Dwi Prastyo, M.Si

## **Abstrak**

*Analisa banjir dapat dilakukan dengan memperkirakan potensi terjadinya hujan lebat (curah hujan tinggi) yang diturunkan dari data variabel-variabel yang mempengaruhi curah hujan yang diperoleh dari data GCM (General Circulation Model) yang menyediakan data iklim secara global. Namun informasi GCM sifatnya global dan tidak berlaku untuk informasi skala kecil, sehingga untuk menjembatani skala gcm ke skala kecil dipakai teknik Statistical Downscaling (SD). Output gcm yang digunakan adalah pada grid 3x3, maka digunakan metode reduksi dimensi untuk mendapatkan prediktor yang lebih sederhana. Metode reduksi dimensi yang sering digunakan adalah Principal Component Analysis (PCA). Dari hasil reduksi dimensi domain grid tersebut selanjutnya diregresikan dengan variabel respon yaitu data curah hujan harian kota Surabaya periode 2009-2018 menggunakan regresi gamma untuk mendapatkan model SD, kemudian dilakukan validasi model SD dengan menggunakan kriteria kebaikan model RMSEP. Prediksi curah hujan terbaik di wilayah Keputih, Wonorejo, Wonokromo, dan Gubeng terdapat pada grid 3x3. Sedangkan Prediksi curah hujan terbaik di wilayah Kedung Cowek dan Gubengterdapat pada grid 3x3. Prediksi curah hujan yang paling tinggi di setiap wilayah digunakan untuk melakukan prediksi banjir. Prediksi banjir pada semua saluran di enam wilayah tidak berpotensi banjir.*

**Kata Kunci** : *Banjir, Curah Hujan, GCM, Regresi Gamma, Statistical Downscaling*

*(Halaman ini sengaja dikosongkan)*

# STATISTICAL DOWNSCALING OF GCM OUTPUT USING GAMMA REGRESSION TO PREDICT RAINFALL FOR FLOOD PREDICTION IN SURABAYA

**Name** : Rida Ikfina Putri  
**Student Number** : 062115 4000 0065  
**Department** : Statistics  
**Supervisors** : Santi Puteri Rahayu, M.Si., Ph.D.  
Dr. rer. pol. Dedy Dwi Prastyo, M.Si.

## Abstract

*Flood analysis can be done by estimating heavy rain (high rainfall) predicted from potential variables that affect rainfall. The potential predictor are obtained from GCM (General Circulation Model) that provide climate data globally. However, GCM information is global in nature and does not apply to small-scale information. Therefore, to bridge the GCM scale to small scale, Statistical Downscaling (SD) techniques are used. The gcm output used is on the 3x3 grid, then the dimension reduction method is used to get a simpler predictor. The dimension reduction method that is often used is Principal Component Analysis (PCA). From the results obtained from dimension reduction is regressed with the response variable (rainfall). This research employs daily rainfall data for the period 2009-2018 using gamma regression to obtain the SD model, whereas the model validation is performed using the RMSEP criterion. The best rainfall predictions in the Keputih, Wonorejo, Wonokromo, and Gubeng areas are obtained by considering on the 3x3 grid of GCM output or predictor. While the best rainfall predictions in the Kedung Cowek and Gubeng areas are on the 3x3 grid. The highest rainfall predictions in each region are based to predict floods. Flood prediction on all channels in six regions has no potential for flooding.*

**Keywords** : *Downscaling Statistics, Flood, Gamma Regression, GCM Rainfall*

*(Halaman ini sengaja dikosongkan)*

## KATA PENGANTAR

Puji Syukur penulis panjatkan atas rahmat dan hidayah yang telah diberikan Allah SWT sehingga penulis dapat menyelesaikan Laporan Tugas Akhir yang berjudul **“Statistical Downscaling Output GCM dengan Pendekatan Regresi Gamma untuk Prediksi Curah Hujan sebagai dasar Prediksi Banjir di Surabaya”** dengan tepat waktu.

Penulis menyadari dalam penyusunan Tugas Akhir ini tidak terlepas dari bantuan maupun dukungan dari berbagai pihak. Pada kesempatan ini penulis menyampaikan terima kasih kepada:

1. Santi Puteri Rahayu, S.Si., M.Si., Ph.D selaku dosen pembimbing utama yang telah meluangkan waktu memberikan bimbingan, saran, serta motivasi selama penyusunan Tugas Akhir berlangsung.
2. Dr. rer. pol. Dedy Dwi Prastyo S.Si., M.Si. selaku dosen pembimbing Tugas Akhir yang telah meluangkan waktu dengan sangat sabar memberikan bimbingan, saran, dukungan serta motivasi selama penyusunan Tugas Akhir.
3. Dr. R.Mohamad Atok, M.Si. dan Dr. rer. pol. Heri Kuswanto S.Si., M.Si. selaku dosen penguji yang telah memberikan masukan dan bantuan dalam menyelesaikan Tugas Akhir.
4. Orang tua, saudara, serta keluarga besar penulis yang selalu memberikan doa, kasih sayang, serta dukungan sehingga menjadi semangat dalam menghadapi kesulitan selama penyusunan Tugas Akhir.
5. Teman-teman seperjuangan dan satu bimbingan, Henidar, Devita, Icha, Afifah, dan Angga yang selalu memberikan bantuan, dukungan, dan semangat dalam penyusunan Tugas Akhir.
6. Sahabat-sahabat, Fonda, Shindy, dan Esti yang senantiasa memberikan dukungan dalam menyelesaikan Tugas Akhir.
7. Terimakasih kepada Angga Mahendra yang selalu membantu dan memberikan semangat kepada penulis dalam penyusunan Tugas Akhir.

8. Seluruh pihak yang turut membantu dalam penyelesaian laporan Tugas Akhir ini baik secara langsung maupun tidak langsung.

Penulis menyadari masih banyak kekurangan dalam pembuatan laporan Tugas Akhir ini. Penulis berharap semoga laporan Tugas Akhir ini bermanfaat dan menambah wawasan bagi pembaca. Kritik dan saran sangat diperlukan untuk perbaikan di masa yang akan datang.

Surabaya, Juli 2019

Penulis

# DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>COVER PAGE</b> .....	iii
<b>ABSTRAK</b> .....	vii
<b>ABSTRACT</b> .....	ix
<b>LEMBAR PENGESAHAN</b> .....	v
<b>KATA PENGANTAR</b> .....	xi
<b>DAFTAR ISI</b>	xiii
<b>DAFTAR GAMBAR</b> .....	xvii
<b>DAFTAR TABEL</b> .....	xxi
<b>DAFTAR LAMPIRAN</b> .....	xxiii
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	6
1.3 Tujuan.....	7
1.4 Manfaat Penelitian.....	7
1.5 Batasan Masalah.....	7
<b>BAB II TINJAUAN PUSTAKA</b> .....	<b>9</b>
2.1 Statistika Deskriptif.....	9
2.2 Konsep Dasar <i>Downscaling</i> .....	9
2.3 <i>Principal Components Analysis</i> (PCA).....	12
2.4 <i>Generalized Linear Model</i> (GLM) .....	14
2.4.1 Keluarga Eksponensial.....	14
2.4.2 Pendugaan Parameter pada GLM.....	16
2.5 Regresi Gamma .....	17
2.5.1 Fungsi Gamma .....	17
2.5.2 Distribusi Gamma .....	18
2.5.3 Model Regresi Gamma .....	19
2.5.4 Penaksiran Parameter .....	22
2.5.5 Pengujian Parameter .....	24
2.5.6 Koefisien Korelasi.....	26

2.5.7	Multikolinearitas .....	26
2.5.8	Uji <i>Goodness of Fit</i> .....	27
2.6	Ukuran Kebaikan Model .....	28
2.7	Hubungan Empiris Curah Hujan dan Limpasan Menggunakan Metode Rasional .....	28
2.8	Metode Mononobe.....	29
2.9	Curah Hujan dan Banjir.....	32
2.8.1	Curah Hujan .....	32
2.8.2	Banjir.....	33
<b>BAB III</b>	<b>METODOLOGI PENELITIAN.....</b>	<b>37</b>
3.1	Sumber Data .....	37
3.2	Definisi Variabel .....	38
3.3	Variabel Penelitian .....	40
3.4	Langkah Analisis .....	43
<b>BAB IV</b>	<b>ANALISIS DAN PEMBAHASAN.....</b>	<b>45</b>
4.1	Identifikasi Karakteristik .....	45
4.1.1	Statistika Deskriptif.....	45
4.1.2	<i>Time Series Plot</i> .....	48
4.1.3	<i>Box Plot</i> .....	49
4.1.4	<i>Scatterplot</i> .....	51
4.2	Prediksi Curah Hujan .....	56
4.2.1	Keputih.....	56
4.2.2	Kedung Cowek.....	73
4.2.3	Gubeng.....	80
4.2.4	Wonorejo.....	88
4.2.5	Wonokromo.....	96
4.2.6	Gunung Sari .....	104
4.2.7	Model Terbaik .....	112
4.3	Prediksi Banjir .....	113
4.3.1	Keputih.....	114
4.3.2	Kedung Cowek.....	114
4.3.3	Gubeng .....	115
4.3.4	Wonorejo.....	116



4.3.5 Wonokromo .....	117
4.3.6 Gunung Sari .....	119
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>121</b>
5.1 Kesimpulan.....	121
5.2 Saran.....	121
<b>DAFTAR PUSTAKA</b>	<b>123</b>
<b>LAMPIRAN 129</b>	
<b>BIODATA PENULIS</b>	<b>145</b>

*(Halaman ini sengaja dikosongkan)*

## DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Pemodelan Statistical Downscaling.....	11
Gambar 2.2 Distribusi Gamma.....	21
Gambar 3.1 Lokasi Pengamatan Penelitian.....	38
Gambar 3.2 Rancangan variabel penelitian .....	40
Gambar 3.3 Diagram Alir Penelitian Secara Umum .....	44
Gambar 4.1 <i>Time Series Plot</i> Data Curah Hujan (CH).....	49
Gambar 4.2 <i>Box Plot</i> Data Curah Hujan (CH) .....	50
Gambar 4.3 <i>Scatterplot</i> Curah Hujan (Vertical) dengan Presipitasi (Horizontal) .....	51
Gambar 4.4 <i>Scatterplot</i> Curah Hujan (Vertical) dengan Kelembapan (Horizontal) .....	52
Gambar 4.5 <i>Scatterplot</i> Curah Hujan (Vertical) dengan Suhu (Horizontal) .....	53
Gambar 4.6 <i>Scatterplot</i> Curah Hujan (Vertical) dengan Kecepatan Angin (Horizontal) .....	54
Gambar 4.7 <i>Scatterplot</i> Curah Hujan (Vertical) dengan Lag Curah Hujan (Horizontal) .....	55
Gambar 4.8 Prediksi Curah Hujan Keputih dengan Empat Variabel Klimatologi .....	59
Gambar 4.9 Hasil Prediksi dengan Variabel Dummy, Suhu, dan Presipitasi.....	61
Gambar 4.10 Hasil Prediksi dengan Empat Prediktor dan Dummy .....	61
Gambar 4.11 PACF Model dengan Dummy dan Semua Variabel .....	62
Gambar 4.12 Hasil Prediksi Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 .....	64
Gambar 4.13 Hasil Prediksi Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2.....	65

<b>Gambar 4.14 Hasil Prediksi Curah Hujan Menggunakan Data Testing .....</b>	<b>66</b>
<b>Gambar 4.15 Prediksi Curah Hujan Keputih dengan 8 Variabel PC dan Variabel Dummy pada Grid 3x3 .....</b>	<b>68</b>
<b>Gambar 4.17 Prediksi Curah Hujan Keputih dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 pada Grid 3x3.....</b>	<b>71</b>
<b>Gambar 4.18 Prediksi Curah Hujan Keputih dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 pada Grid 3x3...</b>	<b>71</b>
<b>Gambar 4.20 Hasil Prediksi Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 .....</b>	<b>74</b>
<b>Gambar 4.21 Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2.....</b>	<b>75</b>
<b>Gambar 4.22 Hasil Prediksi Curah Hujan Menggunakan Data Testing .....</b>	<b>76</b>
<b>Gambar 4.23 Prediksi Curah Hujan Kedung Cowek dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 pada Grid 3x3.....</b>	<b>78</b>
<b>Gambar 4.24 Prediksi Curah Hujan Kedung Cowek dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 pada Grid 3x3...</b>	<b>79</b>
<b>Gambar 4.25 Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3.....</b>	<b>80</b>
<b>Gambar 4.26 Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1.....</b>	<b>82</b>
<b>Gambar 4.27 Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2.....</b>	<b>83</b>

<b>Gambar 4.28 Hasil Prediksi Curah Hujan Menggunakan Data Testing .....</b>	<b>84</b>
<b>Gambar 4.29 Prediksi Curah Hujan Gubeng dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 .....</b>	<b>86</b>
<b>Gambar 4.30 Prediksi Curah Hujan Gubeng dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 pada Grid 3x3 ...</b>	<b>87</b>
<b>Gambar 4.31 Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3.....</b>	<b>88</b>
<b>Gambar 4.32 Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 .....</b>	<b>90</b>
<b>Gambar 4.33 Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2 .....</b>	<b>91</b>
<b>Gambar 4.34 Hasil Prediksi Curah Hujan Menggunakan Data Testing .....</b>	<b>92</b>
<b>Gambar 4.35 Prediksi Curah Hujan Wonorejo dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 .....</b>	<b>94</b>
<b>Gambar 4.37 Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3.....</b>	<b>95</b>
<b>Gambar 4.38 Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 .....</b>	<b>98</b>
<b>Gambar 4.39 Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2 .....</b>	<b>99</b>
<b>Gambar 4.40 Hasil Prediksi Curah Hujan Menggunakan Data Testing .....</b>	<b>100</b>
<b>Gambar 4.41 Prediksi Curah Hujan Wonokromo dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 .....</b>	<b>102</b>

<b>Gambar 4.42</b>	<b>Prediksi Curah Hujan Wonokromo dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 .....</b>	<b>102</b>
<b>Gambar 4.43</b>	<b>Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3.....</b>	<b>104</b>
<b>Gambar 4.44</b>	<b>Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1....</b>	<b>106</b>
<b>Gambar 4.45</b>	<b>Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2.....</b>	<b>107</b>
<b>Gambar 4.46</b>	<b>Hasil Prediksi Curah Hujan Menggunakan Data Testing .....</b>	<b>108</b>
<b>Gambar 4.47</b>	<b>Prediksi Curah Hujan Gunung Sari dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 .....</b>	<b>110</b>
<b>Gambar 4.48</b>	<b>Prediksi Curah Hujan Gunung Sari dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 .....</b>	<b>111</b>
<b>Gambar 4.49</b>	<b>Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3.....</b>	<b>112</b>

## DAFTAR TABEL

Tabel 3.1 Penentuan Lokasi Stasiun.....	38
Tabel 3.2 Variabel Penelitian .....	41
Tabel 3.3 Struktur Data di setiap Stasiun Grid 1x1 .....	41
Tabel 3.4 Struktur Data di setiap Stasiun Grid 3x3.....	42
Tabel 4.1 Statistika Deskriptif Data Curah Hujan.....	46
Tabel 4.2 Statistika Deskriptif Data Luaran GCM.....	46
Tabel 4.3 Pemodelan Curah Hujan Keputih dengan masing- masing variabel.....	57
Tabel 4.4 Pemodelan Curah Hujan Keputih dengan Empat Variabel Klimatologi.....	58
Tabel 4.5 Nilai VIF Model dengan Empat Variabel Klimatologi.....	59
Tabel 4.6 Pemodelan dengan Dummy, Suhu, dan Presipitasi 60	
Tabel 4.7 Pemodelan dengan Semua Variabel dan Dummy .	60
Tabel 4.8 Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1.....	63
Tabel 4.9 Keragaman Kumulatif Variabel GCM.....	67
Tabel 4.10 Pemodelan Curah Hujan pada Grid 3x3.....	67
Tabel 4.11 Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3.....	70
Tabel 4.12 Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1.....	73
Tabel 4.13 Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3.....	77

<b>Tabel 4.14</b>	<b>Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1.....</b>	<b>81</b>
<b>Tabel 4.15</b>	<b>Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3.....</b>	<b>85</b>
<b>Tabel 4.16</b>	<b>Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1.....</b>	<b>89</b>
<b>Tabel 4.17</b>	<b>Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3.....</b>	<b>93</b>
<b>Tabel 4.18</b>	<b>Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1.....</b>	<b>97</b>
<b>Tabel 4.19</b>	<b>Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3.....</b>	<b>101</b>
<b>Tabel 4.20</b>	<b>Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1.....</b>	<b>105</b>
<b>Tabel 4.21</b>	<b>Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3.....</b>	<b>109</b>
<b>Tabel 4.22</b>	<b>Nilai RMSE Semua Model .....</b>	<b>112</b>
<b>Tabel 4.23</b>	<b>Curah Hujan Maksimum.....</b>	<b>113</b>
<b>Tabel 4.24</b>	<b>Perhitungan Banjir Keputih .....</b>	<b>114</b>
<b>Tabel 4.25</b>	<b>Perhitungan Banjir Kedung Cowek.....</b>	<b>114</b>
<b>Tabel 4.26</b>	<b>Perhitungan Banjir Gubeng .....</b>	<b>115</b>
<b>Tabel 4.27</b>	<b>Perhitungan Banjir Wonorejo.....</b>	<b>116</b>
<b>Tabel 4.28</b>	<b>Perhitungan Banjir Wonokromo.....</b>	<b>118</b>
<b>Tabel 4.29</b>	<b>Perhitungan Banjir Gunung Sari.....</b>	<b>119</b>



## DAFTAR LAMPIRAN

<b>Lampiran 1</b>	Syntax Cropping Grid Data GCM.....	129
<b>Lampiran 2</b>	Syntax Cropping Grid Data GCM (lanjutan).....	130
<b>Lampiran 3</b>	<i>Syntax</i> Pemodelan Curah Hujan pada Grid 1x1.....	131
<b>Lampiran 4</b>	<i>Syntax</i> Pemodelan Curah Hujan pada Grid 1x1 (lanjutan) .....	132
<b>Lampiran 5</b>	<i>Syntax</i> Pemodelan Curah Hujan pada Grid 1x1 (lanjutan) .....	133
<b>Lampiran 6</b>	<i>Syntax</i> Pemodelan Curah Hujan pada Grid 1x1 (lanjutan) .....	134
<b>Lampiran 7</b>	<i>Syntax</i> Pemodelan Curah Hujan pada Grid 3x3.....	135
<b>Lampiran 8</b>	<i>Syntax</i> Pemodelan Curah Hujan pada Grid 3x3 (Lanjutan).....	136
<b>Lampiran 9</b>	<i>Syntax</i> Pemodelan Curah Hujan pada Grid 1x1 (lanjutan) .....	137
<b>Lampiran 10</b>	<i>Syntax</i> Prediksi Data Testing pada Grid 1x1.....	138
<b>Lampiran 11</b>	<i>Syntax</i> Prediksi Data Testing pada Grid 3x3.....	139
<b>Lampiran 12</b>	<i>Syntax</i> Prediksi Data Testing pada Grid 3x3 (lanjutan) .....	140
<b>Lampiran 13</b>	<i>Syntax</i> Output Pemodelan .....	141
<b>Lampiran 14</b>	Surat Keterangan Pengambila Data Curah Hujan.....	142
<b>Lampiran 15</b>	Surat Keterangan Pengambila Data Iklim.....	143



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Indonesia terletak diantara dua samudra yaitu Samudra Hindia dan Samudra Pasifik, dan diantara dua benua yaitu Benua Asia dan Benua Pasifik. Letak geografis tersebut menyebabkan Indonesia memiliki tiga iklim utama yaitu iklim monsun (musim), iklim panas (tropis), dan iklim laut. Iklim monsun atau musim terjadi karena adanya pengaruh angin monsun yang bertiup berganti arah setiap setengah tahun sekali. Angin monsun merupakan gerakan massa udara yang terjadi karena adanya perbedaan tekanan udara yang mencolok antara daratan dan juga samudra. Angin monsun terbagi menjadi dua yaitu, angin monsun barat dan angin monsun timur. Angin monsun barat atau yang juga disebut sebagai angin monsun musim dingin timur laut yang bertiup dari kawasan benua Asia ke kawasan benua Australia. Angin bertiup melalui samudra Hindia yang megandung banyak uap air sehingga pada bulan Oktober hingga April Indonesia mengalami musim penghujan. Angin monsun timur atau yang juga disebut sebagai angin monsun musim panas barat daya yang bertiup dari kawasan benua Australia ke kawasan benua Asia. Angin bertiup melewati daerah gurun pasir yang terletak di bagian utara benua Australia yang bersifat kering dan melalui lautan yang sempit. Akibatnya angin tidak mengandung uap air sehingga pada bulan April hingga Oktober Indonesia akan mengalami musim kemarau.

Munculnya angin monsun ditandai dengan adanya curah hujan yang tinggi. Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap dan tidak mengalir. Curah hujan satu millimeter berarti dalam luasan satu meter persegi pada tempat yang datar tertampung air setinggi satu liter. Indonesia merupakan negara yang memiliki angka curah hujan yang bervariasi dikarenakan daerahnya yang berada pada ketinggian yang berbeda-beda salah

satunya daerah Jawa Timur, tepatnya di Kota Surabaya. Surabaya merupakan kota terbesar yang berada di Jawa Timur dan sekaligus menjadi pusat pemerintahan provinsi. Menurut Badan Pusat Statistik Surabaya (2014) secara geografis Surabaya memiliki luas wilayah 326,81 km<sup>2</sup> dan rata-rata curah hujan berkisar 1.321 mm pertahun. Kota Surabaya mengalami musim kemarau pada umumnya pada bulan Mei sampai Oktober, dan musim hujan dari bulan November hingga April. Hujan lebat biasanya terjadi antara bulan Desember dan Januari (BPS 2014)

Kejadian hujan merupakan proses stokastik, sehingga untuk keperluan analisa dan menjelaskan proses stokastik tersebut digunakan teori probabilitas dan analisa frekuensi. Dalam statistik dikenal empat macam distribusi frekuensi yang banyak digunakan dalam hidrologi, yaitu distribusi Normal, Log-Normal, Gumbel dan Log Pearson III. Masing-masing distribusi mempunyai sifat yang khas, sehingga data curah hujan harus diuji kecocokannya dengan sifat statistik masing-masing distribusi tersebut. Pemilihan jenis distribusi yang tidak benar dapat menimbulkan kesalahan perkiraan yang cukup besar, baik overestimated maupun underestimated (Sri Harto 1993). Pengujian distribusi probabilitas dapat menggunakan metode goodness of fit test, yaitu uji chi-squared dan uji kolmogorov-smirnov. Susilowati pada tahun 2010 melakukan pengujian distribusi pada data curah hujan di Lampung dan didapatkan distribusi yang sesuai dengan semua stasiun pengamatan adalah distribusi log Pearson III. Togani pada tahun 2016 melakukan pemilihan distribusi probabilitas pada analisa hujan dan didapatkan distribusi log normal sebagai distribusi terbaik.

Dampak negatif yang ditimbulkan akibat intensitas curah hujan berlebih di Surabaya adalah banjir. Banjir didefinisikan sebagai peristiwa meluapnya air sungai melebihi palung sungai (Anonim, 2011). Hujan bukan merupakan factor utama yang menyebabkan meluapnya air sungai, tetapi bisa disebabkan banyaknya sampah yang menyumbat aliran air. Namun apabila hujan dapat diprediksi hal ini akan dapat membantu pihak

pemerintah dan masyarakat untuk bersiap diri sebagai langkah antisipasi bila terjadi hujan dengan intensitas tinggi. Yuksek, Kankal, dan Ucuncu (2013) menyatakan bahwa banjir merupakan bencana yang umum terjadi dan sering menyebabkan kerugian material yang besar. Beberapa peristiwa banjir yang terjadi di wilayah Surabaya yang bersumber dari SurabayaTribunNews.com (2019) yang memberitakan bahwa banjir di Surabaya barat mengakibatkan banyak mobil penghuni perumahan elit tenggelam. Banjir yang lain terjadi pada Januari 2019 di wilayah yang sama hingga menyebabkan seorang anak meninggal dunia. Hasil survei Surabaya Survey Center (SSC) menyatakan bahwa banjir merupakan masalah utama yang dihadapi masyarakat Surabaya selain kemacetan.

Salah satu penyebab terjadinya banjir adalah adanya genangan air yang terjadi pada suatu tempat dalam kurun waktu tertentu. Salah satu cara untuk dapat mengelola resiko terjadinya genangan adalah dengan memperkirakan kapan suatu daerah akan berpotensi terkena genangan. Analisa ini dapat dilakukan dengan memperkirakan potensi terjadinya hujan lebat (curah hujan tinggi) yang diturunkan dari data variabel-variabel yang mempengaruhi curah hujan yang diperoleh dari data GCM (General Circulation Model) yang menyediakan data-data iklim secara global. GCM merupakan suatu model berbasis computer yang terdiri atas berbagai persamaan numerik dan deterministik yang terpadu dan mengikuti kaidah-kaidah fisika. GCM juga merupakan alat prediksi utama cuaca dan iklim secara numerik dan sebagai sumber informasi primer untuk menilai pengaruh perubahan iklim. Pada beberapa kasus, korelasi diantara variabel independen terjadi dengan pola yang khusus (tertentu), misalnya pada model curah hujan dengan data luaran GCM. Namun informasi GCM sifatnya global dan tidak berlaku untuk informasi skala kecil, sehingga untuk menjembatani Skala GCM ke Skala Kecil dipakai Teknik *Downscaling* (Wigena, 2006) yang merupakan teknik pereduksian dimensi. Metode reduksi dimensi dalam pra-pemrosesan yang

dapat digunakan antara lain: Principal Component Analysis (PCA), Transformasi Wavelet Diskrit (TWD), Kernel PCA, dan ROBPCA

Teknik downscaling dapat dikelompokkan menjadi dua model pendekatan, yaitu : *Dynamic Downscaling* (DD) dan *Statistical Downscaling* (SD). Pendekatan model DD merupakan Regional Iklim Model (RCM) yang merujuk pada kondisi batas secara fisik GCM pada skala regional. Pendekatan ini memerlukan desain yang kompleks dan biaya komputasi yang tinggi (Crane dan Hewitson, 1998 ; Tripathi et.al., 2006) pendekatan model SD merupakan pendekatan yang relatif sederhana dan murah, dengan cara menentukan fungsi transfer (empiris) yang menghubungkan antara variabel sirkulasi atmosfer (prediktor) dengan variabel iklim lokal (prediktan) (Gusfan 2014). Ada tiga asumsi yang digunakan dalam model SD (Hewitson dan Crane, 1996), yaitu : a) prediktor adalah variabel luaran GCM yang relevan dan realistis, b) hubungan empiris ini berlaku juga pada kondisi iklim yang berubah, c) prediktor yang digunakan sepenuhnya mewakili signal iklim.

Secara umum teknik SD terdiri atas tiga kategori yaitu a) pengklasifikasian pola cuaca, b) pembangkitan cuaca, dan c) analisis regresi (Berliana, 2007). Keuntungan model SD ini adalah tidak diperlukannya computer yang banyak dan peralatan penunjang real time lainnya. Selain itu, model SD ini telah memberikan kontribusi terhadap kemampuan RCM dalam menghasilkan prediktor yang relevan (Goodess *et.al.*, 2006). Data luaran GCM bersifat curse of dimensionality. Disamping itu, data GCM merupakan data spasial dan temporal sehingga memungkinkan terjadinya korelasi spasial antara data pada grid yang berbeda dalam satu domain (multikolinearitas) oleh karena itu perlu dilakukan pra-pemrosesan data GCM dengan mereduksi dimensi data GCM.

Beberapa penelitian pemodelan SD yang telah yang telah dikembangkan antara lain menggunakan regresi kontinum (Purnomoadi *et.al.*, 2009), *Markov Chain Monte Carlo* (Lembang, 2013), regresi komponen utama (RKU), jaringan syaraf tiruan

(JST), dan regresi splines adaptif berganda (RSAB) (Berliana, 2007), *projection pursuit regression* (Wigena, 2006), (Generalized Linear Model (Soleh, 2015). Penelitian pemodelan SD tersebut menggunakan data luaran GCM meliputi variabel *precipitable water* (prw), tekanan permukaan laut (slp), komponen angin meridional (va), komponen zonal (ua), ketinggian geopotensial (zg), dan kelembaban spesifik (hus). Hasil penelitian Purnomoadi *et.al.* (2009) diperoleh kesimpulan bahwa tidak ada kekonsistenan pada regresi kontinum, karena tidak ada perbedaan yang signifikan berdasarkan nilai RMSEP dan  $R^2$ -nya diantara kedua stasiun yang diteliti yaitu stasiun hujan di Pontianak dan Ambon. Wigena (2006) menggunakan metode PPR untuk peramalan curah hujan di Kabupaten Indramayu dengan satu variabel GCM yaitu presipitasi. Penelitian tersebut menyimpulkan bahwa pendugaan curah hujan bulanan dengan model PPR lebih akurat dan pola nilai dugaannya lebih mendekati pola data aktualnya daripada model PCR. Lembang (2013) melakukan pemodelan *statistical downscaling* dengan pendekatan *markov chain monte carlo* PCA (studi kasus : data GCM stasiun ambon). Hasil penelitian didapatkan bahwa semakin besar domain grid maka semakin baik pula ketepatan validasi model SD untuk kriteria kebaikan model  $R^2_{pred}$ . Keakuratan prediksi curah hujan untuk model SD menggunakan metode MCMC dominan pada periode bulan Januari-April dan Bulan Oktober. Sedangkan untuk bulan yang lain tingkat akurasi prediksi belum memuaskan, Berliana (2007) mengkaji Dampak Pemanasan Global Terhadap Pola Curah Hujan Indonesia dengan Menggunakan Statistik Downscaling menggunakan metode regresi komponen utama (RKU), jaringan syaraf tiruan (JST), dan regresi splines adaptif berganda (RSAB) Berdasarkan hasil validasi model, metode RSAB mempunyai tingkat akurasi yang relatif stabil tinggi diberbagai wilayah kajian, Soleh (2015) melakukan pemodelan *statistical downscaling* untuk menduga curah hujan bulanan menggunakan model linier terampat sebaran gamma. Pemodelan linier terampat sebaran gamma dapat dianggap lebih baik dibanding metode PCR karena memberikan dugaan curah

hujan non-negatif. Beberapa peneliti sebelumnya juga pernah menggunakan teknik SD seperti model regresi linier berganda (Kidson dan Thompson, 1998 ; Kysely J., 2002), analisis korelasi kanonikal (Busuioc et.al., 2001), jaringan saraf tiruan (Coulibaly dan Dibike, 2004 ; Tolika et.al., 2007) dan support vector machine (Tripathi et.al., 2006).

Banjir yang terjadi perlu dipelajari, pengetahuan mengenai keadaan suatu tempat yang terkena banjir dapat membantu menyiapkan masyarakat setempat untuk dapat menghadapi bencana banjir (Brody, Kang, dan Bernhardt, 2010), disamping meningkatkan kewaspadaan dan kesadaran masyarakat terhadap bencana (Susanti, Miardini, dan Harjadi, 2017). Maka dari itu pada penelitian ini akan dilakukan *Statistical Downscaling* output GCM dengan pendekatan Generalized Linear Model untuk prediksi curah hujan dan banjir di Surabaya. Pengujian distribusi dilakukan untuk melihat pola sebaran data curah hujan tahun 2009 hingga tahun 2018 di kota Surabaya pada enam stasiun yaitu Keputih, Kedungcowek, Gubeng, Wonokromo, Wonorejo, dan Gunungsari. Prediksi curah hujan dengan *Statistical Downscaling* variabel-variabel data luaran GCM yang dihasilkan akan digunakan untuk memprediksi banjir. Variabel prediktor yang digunakan pada penelitian ini adalah *precipitation* (pr), suhu (tas), kecepatan angin (sfwind) dan kelembapan spesifik (huss). Kriteria kebaikan model yang digunakan adalah berdasarkan nilai RMSEP. Nilai RMSEP merupakan nilai dari error hasil taksiran sehingga model dengan RMSEP minimum yang menandakan nilai taksiran mendekati nilai sebenarnya (Draper dan Smith, 1996).

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan, bahwa Surabaya memiliki rata-rata curah hujan yang cukup tinggi berkisar 1.321 mm pertahun. Hal ini menyebabkan kota Surabaya sering mengalami banjir yang menimbulkan banyak kerugian bagi



masyarakat. Berdasarkan hal tersebut, munculah permasalahan sebagai berikut.

1. Bagaimana karakteristik data curah hujan dan variabel luaran GCM?
2. Bagaimana prediksi curah hujan dengan *Statistical Downscaling* variabel-variabel data luaran GCM?
3. Bagaimana prediksi banjir genangan di Surabaya berdasarkan analisis *Statistical Downscaling* curah hujan?

### **1.3 Tujuan**

Berdasarkan rumusan masalah yang ada, tujuan yang akan dicapai dalam melakukan penelitian ini adalah:

1. Mengetahui karakteristik data curah hujan dan variabel luaran GCM.
2. Memperoleh prediksi curah hujan dengan *Statistical Downscaling* variabel-variabel data luaran GCM.
3. Memperoleh prediksi banjir genangan di Surabaya berdasarkan analisis *Statistical Downscaling* curah hujan.

### **1.4 Manfaat Penelitian**

Manfaat yang diharapkan pada penelitian ini diantaranya:

1. Memberikan wawasan mengenai metode *Statistical Downscaling* data dari GCM dalam kasus prediksi curah hujan.
2. Memberikan wawasan keilmuan statistika tentang metode reduksi dimensi menggunakan Principal Component Analysis (PCA) dan metode Generalized Linear Model dengan sebaran Gamma.
3. Memberikan informasi dan rekomendasi bagi profesional dalam bidang klimatologi dan meteorology mengenai penelitian yang membahas curah hujan dan banjir genangan.

### **1.5 Batasan Masalah**

Batasan masalah yang digunakan pada penelitian ini adalah prediksi banjir hanya dipengaruhi oleh curah hujan dan variabel-variabel terkait. Sedangkan variabel lain yang mempengaruhi

terjadinya banjir seperti sampah dan saluran drainase dianggap konstan. Antar variabel presipitasi, kelembapan, suhu, dan kecepatan angin pada grid 3x3 diasumsikan tidak terjadi multikolinearitas

## **BAB II**

### **TINJAUAN PUSTAKA**

Tinjauan pustaka dalam penelitian ini memuat statistika deskriptif, *Statistical Downscaling*, *Principal Components Analysis* (PCA), dan Regresi Gamma serta pengetahuan mengenai curah hujan dan banjir yang akan menjadi sebuah landasan dalam melakukan analisis dan pembahasan pada Bab IV.

#### **2.1 Statistika Deskriptif**

Statistika deskriptif merupakan metode-metode yang berkaitan dengan pengumpulan dan penyajian data sehingga memberikan informasi yang berguna. Metode ini bertujuan untuk menguraikan tentang sifat-sifat atau karakteristik dari suatu keadaan dan membuat gambaran yang sistematis dan akurat mengenai fakta-fakta dan sifatsifat dari fenomena yang diselidiki. Statistika deskriptif hanya untuk memberikan informasi mengenai seputar data tanpa mengambil keputusan atau menarik kesimpulan (inferensia) dari data tersebut (Walpole, 2007). Penyajian data dapat dilakukan dalam bentuk tabel, grafik dan gambar. Ukuran pemusatan memberikan ukuran mengenai nilai maksimum, nilai minimum, rata-rata, median, modus. Ukuran penyebaran meliputi simpangan rata-rata, standar deviasi, jangkauan kuartil dan jangkauan persentil. Grafik yang akan digunakan sebagai visualisasi dari analisa statistika deskriptif adalah Box Plot.

#### **2.2 Konsep Dasar *Downscaling***

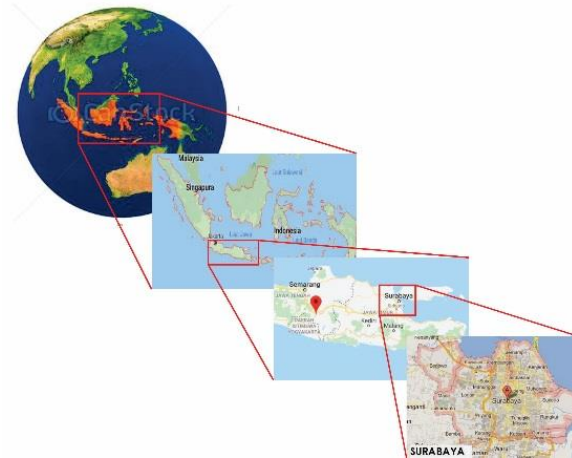
*Downscaling* didefinisikan sebagai upaya untuk menghubungkan antara sirkulasi skala global (variabel prediktor) dengan skala lokal (variabel prediktan). *Downscaling* lebih menunjukkan pada proses perpindahan dari prediktor ke prediktan, yaitu perpindahan dari skala besar (resolusi rendah) ke skala kecil (resolusi tinggi). Von Storch (1999) menyatakan bahwa *downscaling* didasarkan pada asumsi bahwa iklim regional dipengaruhi oleh iklim skala global (benua).

Teknik *downscaling General Circulation Model (GCM)* dapat dibedakan dalam dua pendekatan, yaitu: model *Dynamic Downscaling Model (DD)* dan model *Statistical Downscaling (SD)*. Pendekatan model DD merupakan proses *downscaling* berdasarkan pendekatan proses (*process based*) yang dirunning secara terus menerus, yang dilakukan pada grid dengan skala yang lebih kecil mengikuti perubahan variabel prediktor yang sama pada grid berskala besar (GCM). Pendekatan ini memerlukan desain yang kompleks dan biaya komputasi yang tinggi (Crane dan Hewitson, 1998; Tripathi et.al., 2006).

*Statistical Downscaling (SD)* adalah suatu proses *downscaling* yang bersifat statik dimana data pada grid-grid berskala besar dalam periode dan jangka waktu tertentu digunakan sebagai dasar untuk menentukan data pada grid berskala lebih kecil (Wigena, 2006).

Pendekatan SD menggunakan data regional atau global untuk memperoleh hubungan fungsional antara skala lokal dengan skala global GCM, seperti model regresi. Pendekatan SD disusun berdasarkan adanya hubungan antara grid skala besar (prediktor) dan grid skala lokal (respon) yang dinyatakan dengan model statistik yang dapat digunakan untuk menterjemahkan anomali-anomali skala global menjadi anomali dari beberapa variabel iklim lokal (Zorita dan Storch, 1999). SD menjelaskan hubungan antara skala global dan lokal dengan lebih memperhatikan keakuratan model penduga untuk mempelajari dampak perubahan iklim (Yarnal *et.al.*, 2001 dalam Wigena, 2006). Pendekatan SD menggunakan data GCM untuk peramalan iklim lokal (Fuentes dan Heimann, 2000 dalam Wigena, 2006). Dalam pendekatan ini perlu dilakukan pemilihan variabel-variabel 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 dan Wigley, 2000). Dengan demikian dalam hal peramalan curah hujan, pemilihan variabel prediktor (data GCM) sebaiknya berdasarkan pada adanya korelasi yang kuat antara variabel tersebut dengan curah hujan.

Metode SD didasarkan pada asumsi bahwa iklim regional dikendalikan oleh dua faktor yaitu: kondisi iklim skala besar (resolusi rendah) dan kondisi/karakteristik fisiografik regional/lokal (misal: topografi distribusi daratan-lautan dan tataguna lahan) (Von Stroch *et.al.*, 1999 dalam Sutikno, 2008).



**Gambar 2.1** Ilustrasi Pemodelan Statistical Downscaling

Secara umum bentuk hubungan tersebut dinyatakan dengan:

$$Y = f(X) + \varepsilon \quad (2.1)$$

Dimana :

- |                       |                                                           |
|-----------------------|-----------------------------------------------------------|
| $Y_{(txp)}$           | : variabel respon (curah hujan)                           |
| $X_{(txq \times xl)}$ | : variabel prediktor (luaran GCM)                         |
| $p$                   | : banyaknya variabel y                                    |
| $q$                   | : banyaknya variabel x                                    |
| $g$                   | : banyaknya grid domain GCM                               |
| $l$                   | : banyaknya level ketinggian                              |
| $t$                   | : banyaknya waktu (missal: harian, dasarian atau bulanan) |
| $\varepsilon$         | : error                                                   |

### 2.3 *Principal Components Analysis (PCA)*

PCA adalah suatu prosedur untuk mereduksi dimensi data dengan cara mentransformasi variabel-variabel asal yang berkorelasi menjadi sekumpulan variabel baru yang tidak berkorelasi. Variabel-variabel baru itu dikatakan sebagai *Principal Component* (PC) (Johnson dan Wichern, 2002).

PCA dapat dibentuk dengan dua cara, yaitu dengan matrik kovarian dan matrik korelasi (Jolliffe, 2002). Standardisasi data dilakukan terlebih dahulu jika satuan data antar variabel tidak sama. Standardisasi data perlu dilakukan supaya dominansi satu atau dua variabel dalam PC dapat dihindari. Misalkan vector random  $X' = [X_1, X_2, \dots, X_p]$  yang terdiri atas sejumlah observasi sebanyak  $p$  variabel, maka PC adalah kombinasi linear dari  $p$  variabel tersebut yang merupakan sistem koordinat baru yang didapat dari hasil rotasi sistem asal  $X_1, X_2, \dots, X_p$  sebagai sumbu koordinat. Sumbu baru  $(Z_1, Z_2, \dots, Z_p)$  merupakan arah dengan variabilitas maksimum yang memberikan struktur kovariansi yang lebih sederhana dan  $Z_1, Z_2, \dots, Z_p$  adalah PC yang tidak berkorelasi (Johnson dan Winchern, 2002). PC dapat diperoleh dari pasangan eigenvalue-eigenvektor matriks kovarian maupun matriks korelasi. Selanjutnya bila  $\Sigma$  adalah matriks varian-kovarian dari vektor random  $X' = [X_1, X_2, \dots, X_p]$ ,  $\Sigma$  didapatkan berdasarkan rumus,

$$\Sigma = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T \quad (2.2)$$

dengan

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

$x_i$  = observasi ke- $i$

$n$  = jumlah observasi

dan  $\Sigma$  memiliki pasangan eigen value-eigen vektor  $(\lambda_1 e_1), (\lambda_2 e_2), \dots, (\lambda_p e_p)$  dengan  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$  Maka model PC dapat ditulis sebagai berikut :

$$Z_1 = e_1' X = e_{11} X_1 + e_{21} X_2 + \dots + e_{p1} X_p \quad (2.3)$$

$$Z_2 = e_2' X = e_{12} X_1 + e_{22} X_2 + \dots + e_{p2} X_p \quad (2.4)$$

.

.

.

$$Z_p = e_p' X = e_{1p} X_1 + e_{2p} X_2 + \dots + e_{pp} X_p \quad (2.5)$$

dengan

$Z_1$ = PC pertama, yang mempunyai varians terbesar

$Z_2$ = PC kedua, yang mempunyai varians terbesar kedua

$Z_j$ = PC ke- $p$ , yang mempunyai varians terbesar ke- $j$

$X_1$ = variabel asal pertama

$X_2$ = variabel asal kedua

$X_j$ = variabel asal ke- $j$

dan diperoleh:

$$\text{Var}(Z_j) = e_j' \Sigma e_j \quad j = 1, 2, \dots, p \quad (2.6)$$

$$\text{Cov}(Z_j, Z_k) = e_j' \Sigma e_k \quad j \neq k \quad (2.7)$$

PC tidak berkorelasi dan mempunyai varians yang sama dengan eigenvalue dari  $\Sigma$ , sehingga:

$$\sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \sum_{i=1}^p \text{Var}(X_i) = \text{tr}(\Sigma) = \lambda_1 + \lambda_2 + \dots + \lambda_p \quad (2.8)$$

Apabila total variansi populasi adalah

$\sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \lambda_1 + \lambda_2 + \dots + \lambda_p$ , maka:

$$\text{Proporsi varian ke } i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad (2.9)$$

Apabila PC yang diambil sebanyak  $k$  dengan ( $k < p$ ), maka:

$$\text{Proporsi varian } k \text{ PC} = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad (2.10)$$

Apabila yang digunakan di awal adalah matriks kovariansi dari data yang distandarkan, karena diagonal utama matriks berisi nilai satu,

maka total variansi populasi untuk variabel distandarkan adalah  $p$ , yang merupakan jumlah elemen diagonal matriks  $\rho$ . Sehingga:

$$\text{Proporsi varian ke } i = \frac{\lambda_i}{p} \quad (2.11)$$

## 2.4 Generalized Linear Model (GLM)

Bila peubah respons  $Y$  tidak lagi mengikuti sebaran normal namun seperti Gamma atau *Inverse Gaussian* (asalkan termasuk dalam keluarga eksponensial) dan ragam  $Y$  merupakan fungsi dari nilai tengahnya sehingga dapat dipastikan bahwa ragam tidak homogen maka digunakanlah suatu model yang disebut model linear terampat (*Generalized Linear Model/GLM*).

### 2.4.1 Keluarga Eksponensial

Peubah pengamatan  $Y$  diasumsikan memiliki fungsi peluang keluarga eksponensial yang dapat dimodelkan sebagai berikut (Jong dan Heller, 2008):

$$f(y | \theta, \phi) = c(y, \phi) \exp\left(\frac{y\theta - b(\theta)}{a(\phi)}\right) \quad (2.12)$$

Dengan  $\theta$  dan  $\phi$  adalah parameter kanonik dan parameter dispersi dan  $a(\phi)$ ,  $b(\theta)$ , dan  $c(y, \phi)$  adalah suatu fungsi yang diketahui.

Fungsi log-likelihood adalah  $\log f(y | \theta, \phi)$  ditulis sebagai  $l(\theta, \phi | y)$  merupakan suatu fungsi dari  $\theta$  dan  $\phi$  dengan  $y$  diketahui. Nilai harapan dan ragam  $Y$  dapat ditentukan dengan mengevaluasi fungsi turunan dari  $l(\theta, \phi | y)$  yang bersesuaian yaitu :

$$E(Y) = \mu = b'(\theta) \text{ dan } \text{Var}(Y) = \sigma^2 = b''(\theta)a(\phi)$$

Dengan  $b'(\theta)$  dan  $b''(\theta)$  adalah turunan pertama dan kedua dari  $b(\theta)$  Beberapa distribusi peluang yang termasuk dalam distribusi keluarga eksponensial adalah distribusi binomial, Poisson, normal, gamma, *inverse Gaussian* dan *negative binomial*.



**Tabel 2.1** Distribusi Keluarga Eksponensial dan Parameternya

Distribusi	$\theta$	$b(\theta)$	$\phi$	$E(Y)$	$Var(\mu)$
Binomial, $B(n, \pi)$	$\ln \frac{\pi}{1-\pi}$	$n \ln(1 + e^\theta)$	1	$n\pi$	$n\pi(1-\pi)$
Poisson, $P(\mu)$	$\ln(\mu)$	$e^\theta$	1	$\mu$	$\mu$
Normal, $N(\mu, \sigma^2)$	$\mu$	$\frac{1}{2}\theta^2$	$\sigma^2$	$\mu$	1
Gamma, $G(\mu, \nu)$	$-\frac{1}{\mu}$	$-\ln(-\theta)$	$\frac{1}{\nu}$	$\mu$	$\mu^2$
Inverse Gaussian, $IG(\mu, \sigma^2)$	$-\frac{1}{2\mu^2}$	$-\sqrt{-2\theta}$	$\sigma^2$	$\mu$	$\mu^3$
Negative Binomial, $NB(\mu, k)$	$\ln \frac{k\mu}{1+k\mu}$	$-\frac{1}{k} \ln(1 - ke^\theta)$	1	$\mu$	$\mu(1+k\mu)$

*GLM* merupakan pengembangan dari model linear klasik dengan peubah respons  $Y$  merupakan suatu komponen yang bebas dengan nilai tengah  $\mu$ .

Ada tiga komponen utama dalam *GLM* (McCullagh dan Nelder, 1989):

1. Komponen acak, yaitu komponen dari  $Y$  yang bebas dan fungsi sebaran peluang  $Y$  termasuk dalam keluarga sebaran eksponensial dengan  $E(Y) = \mu$
2. Komponen sistematis, yaitu  $X_1, X_2, \dots, X_p$  yang menghasilkan penduga linear  $\eta$  dimana  $\eta = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$ .

3. Fungsi penghubung (*link function*)  $g(\cdot)$ , menggambarkan hubungan antara penduga linear  $\eta$  dengan nilai tengah  $\mu$ .

Hubungan ini dapat ditulis dengan  $\eta = g(\mu)$

Dalam model linear klasik, komponen (1) menyebar normal dan komponen (3) merupakan fungsi identitas. Sedangkan dalam *GLM*, komponen (1) mungkin berasal dari salah satu anggota keluarga sebaran eksponensial lainnya dan komponen (3) merupakan fungsi monoton lainnya. Berikut Tabel 2 tentang fungsi hubung kanonik untuk berbagai distribusi.

**Tabel 2.2** Fungsi Hubung

Fungsi Hubung	$g(\mu)$	Distribusi
<i>Identity</i>	$\mu$	Normal
Log	$\ln \mu$	Poisson
Power	$\mu^p$	Gamma (p = -1) <i>Inverse</i> Gaussian (p = -2)
Square root	$\sqrt{\mu}$	
Logit	$\ln \frac{\mu}{1-\mu}$	Binomial

### 2.4.2 Pendugaan Parameter pada GLM

Pendugaan parameter  $\beta$  dan  $\theta$  dapat dilakukan dengan metode *Maximum Likelihood Estimation (MLE)* yaitu dengan memaksimumkan fungsi log-likelihood sebagai berikut :

$$\ell(\beta, \phi) = \sum_{i=1}^n \ln f(y_i; \beta, \phi) = \sum_{i=1}^n \left\{ \ln c(y_i, \phi) + \frac{y_i \theta_i - a(\theta_i)}{\phi} \right\} \quad (2.13)$$

dengan mengasumsikan bahwa  $y_i$  adalah peubah respons keluarga eksponensial yang bebas. Lalu *MLE* untuk masing-masing parameter  $\beta_j$  adalah dengan menurunkan  $\ell(\beta, \phi)$  terhadap  $\beta_j$  yaitu

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^n \frac{\partial \ell}{\partial \theta_i} \frac{\partial \theta_i}{\partial \beta_j},$$

dengan

$$\frac{\partial \ell}{\partial \theta_i} = \frac{y_i - a(\theta_i)}{\phi} = \frac{y_i - \mu_i}{\phi}, \quad \frac{\partial \ell}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \eta_i} x_{ij}$$

diketahui bahwa  $\eta_i = x_i' \beta$  dan  $x_{ij}$  adalah komponen  $i$  dari  $x_j$ , kemudian  $\partial \ell / \partial \beta_j = 0$  (yang merupakan kondisi orde pertama dalam memaksimalkan fungsi *likelihood*) maka akan didapatkan penduga dari  $\beta$  dan  $\phi$ . Namun pada kondisi order pertama tersebut biasanya sulit diperoleh secara langsung kecuali untuk kasus normal dengan fungsi hubung identitas. Untuk mengatasi hal tersebut dapat digunakan dengan iterasi *Fisher Scoring*.

## 2.5 Regresi Gamma

Penelitian ini akan menggunakan regresi gamma dimana sebelum melakukan analisis terlebih dahulu perlu diketahui mengenai distribusi gamma dan generalized linear model

### 2.5.1 Fungsi Gamma

Fungsi gamma pertama kali dikenalkan oleh Leonhard Euler (1707 – 1783) seorang matematikawan kebangsaan Swiss yang bertujuan untuk mengeneralisasi faktorial ke nilai non integer. Seiring dengan berkembangnya ilmu pengetahuan maka fungsi gamma mulai dikembangkan oleh ilmuwan lainnya, seperti Adrien-Marie Legendre (1752-1833), Carl Friedrich Gauss (1777 – 1855), Christoph Gudermann (1798 – 1833), Joseph Liouville (1809 – 1982) dan lain-lain.

Fungsi gamma memiliki parameter  $\alpha$  dan dinotasikan sebagai berikut:

$$\Gamma(\alpha) = \int_0^{\infty} e^{-y} y^{\alpha-1} dy \quad \text{dengan } \alpha > 0 \quad (2.14)$$

Sebagai fungsi yang bersifat kontinyu dan memiliki parameter yang positif, fungsi gamma memiliki karakteristik sebagai berikut (Randservices, 2018):

1.  $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1), \alpha > 0$

$$2. \Gamma(\alpha) = (\alpha - 1)!, \alpha > 1$$

$$3. \Gamma(1/2) = \sqrt{\pi}$$

Distribusi gamma merupakan model probabilitas yang banyak diaplikasikan untuk waktu tunggu misalnya dalam pengamatan yang berhubungan dengan lamanya waktu menyelesaikan pekerjaan, klaim asuransi, bidang meteorologi untuk memodelkan curah hujan, bidang komunikasi, maupun bidang kesehatan dan lain-lain yang berhubungan dengan teori antrian dan teori keandalan atau reliabiliti.

### 2.5.2 Distribusi Gamma

Distribusi gamma merupakan keluarga dari distribusi probabilitas kontinu. Distribusi gamma berasal dari fungsi gamma yang memiliki fungsi kepadatan peluang sama dengan satu. Distribusi gamma menghasilkan kurva yang memiliki *skewness* atau *unbalance*. Terbentuknya distribusi Gamma berawal dari distribusi dengan tiga parameter yaitu  $\alpha, \theta, \lambda$ . Parameter  $\alpha$  disebut sebagai parameter bentuk (*shape*) karena mempengaruhi puncak keruncingan (*peaked-ness*) dari kurva distribusi gamma, sedangkan parameter  $\theta$  disebut sebagai parameter skala karena berpengaruh terhadap sebaran distribusi. Parameter  $\lambda$  disebut sebagai parameter lokasi yang secara grafis hanya menggeser kurva distribusi gamma ke kanan atau ke kiri sejauh  $\lambda$  tanpa mengubah bentuk dan sebaran distribusinya. Fungsi kepadatan peluang dari distribusi gamma tiga parameter adalah sebagai berikut (Ewemoje dan ewemooje, 2011) :

$$f(y | \alpha, \theta, \lambda) = \begin{cases} \frac{1}{\theta^\alpha \Gamma(\alpha)} (y - \lambda)^{\alpha-1} e^{-(y-\lambda)/\theta} & , \alpha > 0, \theta > 0, \lambda < y < \infty \\ 0 & , \text{ untuk } y \text{ yang lainnya} \end{cases} \quad (2.15)$$

Distribusi  $y$  tersebut biasanya dinotasikan dengan  $y \sim \text{Gamma}(\alpha, \theta, \lambda)$  dengan nilai harapan (koefisien rata-rata), varian dan *skewness* sebagai berikut :

$$\mu_y = \lambda + \alpha\theta \quad ; \quad \sigma_y^2 = \alpha\theta^2 \quad ; \quad g_y = \frac{2}{\sqrt{\alpha}} \quad (2.16)$$

Distribusi gamma yang paling banyak dikenal adalah distribusi gamma dengan dua parameter yaitu  $\alpha$  dan  $\theta$ . Distribusi ini terbentuk ketika parameter lokasi  $\lambda$  bernilai nol. Fungsi kepadatan peluang dari distribusi gamma dengan dua parameter yaitu  $\alpha$  dan  $\theta$  dimana  $y \sim \text{Gamma}(\alpha, \theta)$  adalah sebagai berikut (Hogg, McKean, dan Craig, 2013):

$$f(y | \alpha, \theta) = \begin{cases} \frac{1}{\theta^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-y/\theta} & , y > 0, \alpha > 0, \theta > 0 \\ 0 & , \text{ untuk } y \text{ yang lainnya} \end{cases} \quad (2.17)$$

dengan  $\Gamma(\alpha)$  merupakan fungsi gamma yang bernilai positif. Nilai harapan dan varian dari distribusi gamma dengan dua parameter  $(\alpha, \theta)$  adalah sebagai berikut:

$$\mu = \alpha\theta \text{ dan } \sigma^2 = \alpha\theta^2 \quad (2.18)$$

Bentuk lain dari fungsi kepadatan peluang distribusi gamma dengan dua parameter yaitu  $\alpha$  dan  $\gamma$  dimana  $y \sim \text{Gamma}(\alpha, \gamma)$  adalah sebagai berikut (Nadarajah dan Gupta, 2006):

$$f(y | \alpha, \gamma) = \begin{cases} \frac{\gamma^\alpha y^{\alpha-1} e^{-\gamma y}}{\Gamma(\alpha)} & , y > 0, \alpha > 0, \gamma > 0 \\ 0 & , \text{ untuk } y \text{ yang lainnya} \end{cases} \quad (2.19)$$

dengan nilai harapan dan variannya adalah sebagai berikut:

$$\mu = \frac{\alpha}{\gamma} \text{ dan } \sigma^2 = \frac{\alpha}{\gamma^2} \quad (2.20)$$

### 2.5.3 Model Regresi Gamma

Model regresi gamma adalah sebagai berikut:

$$\mu_i = E(Y_i) = \exp(\beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ip}) \quad (2.21)$$

dengan

$$\mathbf{x}_i = [1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}]_{1 \times (1+p)}^T ; \quad \boldsymbol{\beta} = [\beta_0 \quad \beta_1 \quad \cdots \quad \beta_p]_{1 \times (1+p)}^T$$

Berdasarkan persamaan (2.20) maka diperoleh:

$$\mu_i = \frac{\alpha}{\gamma_i} = \exp(\mathbf{x}_i^T \boldsymbol{\beta}) \text{ sehingga:}$$

$$\gamma_i = \frac{\alpha}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \quad (2.22)$$

Dengan melakukan substitusi persamaan (2.22) ke persamaan (2.19), maka diperoleh fungsi kepadatan peluang dari model regresi gamma adalah sebagai berikut:

$$f(y_i | \alpha, \gamma) = \begin{cases} \left( \frac{\alpha}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \right)^\alpha y_i^{\alpha-1} e^{-\left( \frac{\alpha}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \right) y_i} & , y_i > 0, \alpha > 0, \gamma > 0 \\ 0 & , \text{ untuk } y \text{ yang lainnya} \end{cases} \quad (2.23)$$

Peubah respons dalam model regresi gamma merupakan peubah kontinu non negatif  $Y_i > 0$  dengan ukuran kemiringan ke kanan yang ditentukan oleh besarnya  $v$  dengan nilai harapan sebagai berikut:

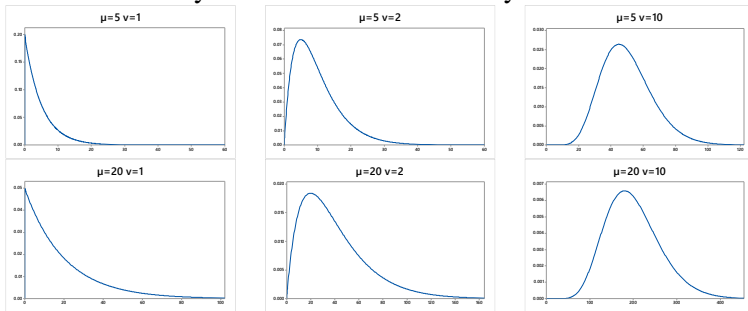
$$\begin{aligned} \mu_i &= \frac{v}{\gamma_i} \text{ sehingga} \\ \gamma_i &= \frac{v}{\mu_i} \end{aligned} \quad (2.24)$$

Dengan melakukan substitusi persamaan (2.24) ke persamaan (2.19), maka fungsi padat peluangnya sebagai berikut  $Y_i \sim \text{Gamma}(\mu_i, v)$ , (McCullagh dan Nelder, 1989):

$$f(y_i) = \frac{1}{\Gamma(v) y_i} \left( \frac{v y_i}{\mu_i} \right)^v \exp\left( -\frac{v y_i}{\mu_i} \right) \quad (2.25)$$

Dimana  $y_i > 0, v > 0, \mu_i > 0, i = 0, 1, 2, \dots, n$

Berdasarkan fungsi padat peluang tersebut diperoleh  $E(Y_i) = \mu_i$  dan  $Var(Y_i) = \mu_i^2 / v = \mu_i^2 \sigma^2$ . Berikut Gambar 1 tentang parameter  $v$  menyatakan bentuk distribusinya



Gambar 2.2 Distribusi Gamma

Pada model regresi gamma, parameter  $v = \sigma^{-2}$  diasumsikan konstan untuk semua pengamatan, sehingga fungsi densitasnya mempunyai bentuk yang sama untuk semua pengamatan. Model regresi gamma merupakan suatu model yang tergolong dalam model linear terampat (*Generalized Linear Model*), dengan mengasumsikan bahwa peubah respons  $Y_i \sim \text{Gamma}(\mu_i, v)$ ,  $E(Y_i) = \mu_i = X_i\beta$ . Fungsi hubung dikatakan fungsi penghubung kanonik bila parameter kanoniknya sama dengan fungsi hubungannya, yaitu  $\eta = \theta$  dengan  $\theta$  adalah parameter kanonik.

Pada regresi gamma diketahui bahwa  $\theta = -\frac{1}{\mu_i}$ , sehingga karena nilai  $\mu_i > 0$  maka digunakan fungsi hubung  $g(\mu_i) = \mu_i^{-1}$ , kadangkala juga menggunakan fungsi hubung log natural (Czado, 2004).

Apabila menggunakan fungsi hubung “sqrt” maka model regresi gamma yang erbentuk adalah sebagai berikut.

Diketahui bahwa :

$$\eta = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (2.26)$$

Dengan fungsi link “sqrt” adalah  $\sqrt{\mu}$  maka,

$$\eta = \sqrt{\mu} \quad (2.27)$$

$$\eta^2 = \mu \quad (2.28)$$

Sehingga model yang terbentuk untuk fungsi link “sqrt” adalah sebagai berikut.

$$\hat{\mu} = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon_i)^2 \quad (2.29)$$

Dengan :

$\beta_j$  : Koefisien parameter

$X_j$  : Variabel prediktor,  $j = 1, 2, 3, \dots, p$

$\varepsilon_i$  : Residual dari model

#### 2.5.4 Penaksiran Parameter

Penaksiran parameter dilakukan dengan metode *maximum likelihood estimator* (MLE). Metode MLE bertujuan untuk memaksimumkan nilai fungsi *lnlikelihood* (Casella dan Berge, 2002). Berdasarkan fungsi kepadatan peluang bersama dari Y pada persamaan(2.23), maka diperoleh fungsi *lnlikelihood* sebagai berikut:

$$\ln L(\alpha, \beta, i = 1, 2, \dots, n) = \ln \left( \prod_{i=1}^n \left( \frac{\left( \frac{\alpha}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \right)^\alpha}{\Gamma(\alpha)} y_i^{\alpha-1} \exp\left( -\frac{y_i \alpha}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \right) \right) \right) \quad (2.30)$$

$$= -n \ln \Gamma(\alpha) - \alpha \sum_{i=1}^n (\mathbf{x}_i^T \boldsymbol{\beta}) + \alpha (n \ln \alpha) + (\alpha - 1) \ln \sum_{i=1}^n y_i - \sum_{i=1}^n \left( \frac{y_i \alpha}{\exp(\mathbf{x}_i^T \boldsymbol{\beta})} \right) \quad (2.31)$$

Penaksir parameter pada regresi gamma yaitu  $\hat{\alpha}$  dan  $\hat{\boldsymbol{\beta}}$  diperoleh dengan memaksimumkan fungsi  $L(\alpha, \boldsymbol{\beta})$ . Maksimum dari fungsi  $L(\alpha, \boldsymbol{\beta})$  dapat diperoleh dengan cara mencari turunan pertama fungsi *lnlikelihood* pada persamaan (2.27), untuk parameter  $\alpha$  dan  $\boldsymbol{\beta}$  kemudian disamadengankan nol.

Turunan parsial fungsi  $\ln L(\alpha, \boldsymbol{\beta})$  terhadap parameter  $\boldsymbol{\beta}$  :



$$\frac{\partial \ln L(\alpha, \beta)}{\partial \beta} = -\alpha \sum_{i=1}^n \mathbf{x}_i^T + \alpha \sum_{i=1}^n \mathbf{x}_i^T y_i (\exp(\mathbf{x}_i^T \beta))^{-1} \quad (2.32)$$

Turunan parsial fungsi  $\ln L(\alpha, \beta)$  terhadap parameter  $\alpha$  :

$$\frac{\partial \ln L(\alpha, \beta)}{\partial \alpha} = -n\psi(\alpha) - \sum_{i=1}^n (\mathbf{x}_i^T \beta) + n(\ln \alpha + 1) + \sum_{i=1}^n y_i - \sum_{i=1}^n \left( \frac{y_i}{\exp(\mathbf{x}_i^T \beta)} \right) \quad (2.33)$$

Persamaan (2.27) dan (2.28) menghasilkan penyelesaian yang tidak *close form*, sehingga untuk penyelesaiannya digunakan pendekatan numerik yaitu algoritma *Fisher Scoring*. Menurut Thomas, dkk (2001) algoritma adalah prosedur komputasi dimana mengambil sebuah nilai atau menentukan nilai sebagai *input* dan menghasilkan beberapa nilai sebagai *output*. Sebuah algoritma adalah sebuah urutan langkahlangkah komputasi yang dapat merubah sebuah input menjadi output. Menurut Lawless (2003) misalkan terdapat fungsi-fungsi  $\frac{\partial l(\beta)}{\partial \beta_j} = 0$  dengan  $j = 1, 2, \dots, p$ ,

maka nilai-nilai  $\beta_j$  yang memenuhi fungsi implisit tersebut dapat diperoleh melalui iterasi *Newton-Raphson* sebagai berikut :

$$\hat{\beta}^{(r+1)} = \hat{\beta}^{(r)} - H(\hat{\beta}^{(r)})^{-1} D(\hat{\beta}^{(r)}), \quad r = 0, 1, 2, \dots \quad (2.34)$$

Dengan

$$\beta = (\beta_1, \beta_2, \dots, \beta_p)^T, \quad D(\beta) = \left( \frac{\partial l(\beta)}{\partial \beta_j}, \dots, \frac{\partial l(\beta)}{\partial \beta_j} \right)^T, \quad H(\beta) = \left[ \frac{\partial^2 l(\beta)}{\partial \beta_j \partial \beta_k} \right], \text{ dan}$$

$$j, k = 0, 1, 2, \dots, p.$$

Adapun langkah-langkah dalam algoritma *Newton-Raphson* adalah sebagai berikut :

1. Menentukan nilai awal  $\beta^0$
2. Menentukan  $D(\beta^0)$  dan  $H(\beta^0)$
3. Menghitung estimator parameter untuk  $r = 0, 1, 2, \dots$  dengan menggunakan persamaan (2.29)
4. Mengulangi iterasi sampai diperoleh nilai yang konvergen, yaitu  $\max |\hat{\beta}^{(r+1)} - \hat{\beta}^{(r)}| \leq \varepsilon$  dengan  $\varepsilon$  adalah konstanta positif yang ditentukan.

Algoritma *Fisher-Scoring* Menurut Smyth (2002) adalah salah satu bentuk pengembangan dari metode *Newton-Raphson* dengan mengganti  $H(\beta)$  dengan  $I(\beta)$  Bentuk persamaan iterasi *Fisher-Scoring* adalah sebagai berikut :

$$\hat{\beta}^{(r+1)} = \hat{\beta}^{(r)} + I(\hat{\beta}^{(r)})^{-1} D(\hat{\beta}^{(r)})$$

dengan  $I(\beta) = -E[H(\beta)]$  dan  $I(\beta)$  adalah matriks informasi *fisher* berukuran  $(p+1) \times (p+1)$ .

### 2.5.5 Pengujian Parameter

Pengujian serentak parameter model regresi gamma dilakukan menggunakan metode *Maximum Likelihood Ratio Test* (MLRT) dengan hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{paling sedikit ada satu } \beta_j \neq 0 \quad ; j = 1, 2, \dots, p$$

Himpunan parameter dibawah populasi adalah:

$$\Omega = \{\beta_0, \beta_1, \dots, \beta_p, \alpha\} = \{\boldsymbol{\beta}, \alpha\}$$

Himpunan parameter dibawah  $H_0(\omega)$  adalah :

$$\omega = \{\beta_0, \alpha\}$$

$L(\hat{\Omega})$  adalah nilai maksimum *likelihood* untuk model lengkap dimana melibatkan variabel prediktor (dibawah populasi) sedangkan  $L(\hat{\omega})$  adalah nilai maksimum *likelihood* untuk model sederhana tanpa melibatkan variabel prediktor (dibawah  $H_0$ ) .

Nilai penaksir parameter dibawah populasi untuk  $\boldsymbol{\beta}$  dan  $\hat{\alpha}$  diperoleh berdasarkan penaksiran parameter di subbab (2.4.7).

Sedangkan nilai penaksir parameter  $\hat{\beta}_0$  dan  $\hat{\alpha}$  dibawah  $H_0(\omega_{GW})$  diperoleh dari penaksiran parameter fungsi *likelihood* dibawah  $H_0(\omega_{GW})$ .

Statistik uji yang digunakan adalah:

$$G^2 = -2 \ln \Lambda = -2 \ln \left( \frac{L(\hat{\omega})}{L(\hat{\Omega})} \right) = -2 \ln(L(\hat{\omega}) - L(\hat{\Omega})) \quad (2.35)$$

dimana

$$\ln L(\hat{\Omega}) = -n \ln \Gamma(\hat{\alpha}) - \hat{\alpha} \sum_{i=1}^n (\mathbf{x}_i^T \hat{\boldsymbol{\beta}}) + \hat{\alpha} (n \ln \hat{\alpha}) + (\hat{\alpha} - 1) \ln \sum_{i=1}^n y_i - \sum_{i=1}^n \left( \frac{y_i \hat{\alpha}}{\exp(\mathbf{x}_i^T \hat{\boldsymbol{\beta}})} \right)$$

$$\ln L(\hat{\omega}) = -n \ln \Gamma(\hat{\alpha}) - \hat{\alpha} n \hat{\beta}_0 + \hat{\alpha} (n \ln \hat{\alpha}) + (\hat{\alpha} - 1) \ln \sum_{i=1}^n y_i - \sum_{i=1}^n \left( \frac{y_i \hat{\alpha}}{\exp(\hat{\beta}_0)} \right)$$

Apabila  $n \rightarrow \infty$  maka *likelihood* ratio  $G^2$  asymptotik berdistribusi  $\chi^2$  (Cameron dan Trivedi, 2005), sehingga tolak  $H_0$  jika  $G_{hitung}^2 > \chi_{\alpha, k}^2$ .

Apabila keputusan pengujian secara serentak adalah tolak  $H_0$ , maka pengujian dilanjutkan dengan uji parsial yang bertujuan untuk mengetahui variabel prediktor mana yang berpengaruh terhadap model. Hipotesis yang digunakan adalah:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 \quad ; j = 1, 2, \dots, p$$

Statistik uji yang digunakan :

$$Z = \frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \quad \text{dimana } se(\hat{\beta}_j) = \sqrt{\text{var}(\hat{\beta}_j)}$$

$\text{var}(\hat{\beta}_j)$  merupakan elemen diagonal ke  $j+1$  dari  $-\mathbf{H}^{-1}(\hat{\boldsymbol{\gamma}})$ .

Berdasarkan konsep CLT maka apabila taksiran  $\hat{\beta}_j$  memiliki nilai  $E(\hat{\beta}_j) = \beta_j$  dan  $\text{var}(\hat{\beta}_j)$  maka  $Z$  akan berdistribusi normal asymptotik untuk  $n \rightarrow \infty$  (Gupta, 2008). Tolak  $H_0$  jika  $|Z_{hitung}| > Z_{\alpha/2}$ .

### 2.5.6 Koefisien Korelasi

Koefisien korelasi merupakan suatu indikator untuk mengetahui hubungan linier antara dua variabel. Nilai koefisien korelasi berada pada rentang  $-1 \leq r_{y_1, y_2} \leq 1$ . Jika nilai korelasi mendekati angka 1 maka kedua variabel itu memiliki hubungan secara linier sedangkan jika nilai koefisien korelasi bernilai 0, maka dapat dikatakan bahwa kedua variabel tersebut tidak memiliki hubungan secara linier. Nilai positif dan negatif pada koefisien korelasi hanya membedakan arah hubungannya saja. Jika nilai koefisien korelasi bernilai positif maka hubungan kedua variabel itu memiliki hubungan linier yang searah demikian juga sebaliknya. Koefisien korelasi didefinisikan sebagai berikut (Hogg, McKean, dan Craig, 2013):

$$r_{y_1, y_2} = \frac{\sum_{i=1}^n (y_{1i} - \bar{y}_1)(y_{2i} - \bar{y}_2)}{\sqrt{\left(\sum_{i=1}^n (y_{1i} - \bar{y}_1)^2\right)\left(\sum_{i=1}^n (y_{2i} - \bar{y}_2)^2\right)}} \quad (2.36)$$

Hipotesis untuk pengujian korelasi dari variabel respon adalah sebagai berikut:

$H_0 : \rho_{y_1, y_2} = 0$  (tidak terdapat hubungan antara  $Y_1$  dan  $Y_2$ )

$H_1 : \rho_{y_1, y_2} \neq 0$  (terdapat hubungan antara  $Y_1$  dan  $Y_2$ )

Dengan statistik uji yang digunakan adalah : (Gujarati D. N., 2003)

$$t = \frac{r_{y_1, y_2} \sqrt{n-2}}{\sqrt{1 - (r_{y_1, y_2})^2}} \quad (2.37)$$

Tolak  $H_0$  jika  $|t| > t_{(\alpha/2; (n-2))}$

### 2.5.7 Multikolinieritas

Syarat yang harus dipenuhi dalam analisis regresi yang menggunakan lebih dari satu variabel prediktor adalah tidak adanya multikolinieritas. Multikolinieritas adalah suatu kondisi dimana terjadi korelasi tinggi antar variabel prediktornya. Adanya

multikolinieritas mengakibatkan penaksir parameter yang diperoleh menjadi tidak akurat. Hal ini dikarenakan penaksir parameter pada kasus multikolinieritas menghasilkan *standart error* yang besar dengan signifikansi yang kecil bahkan bisa mengakibatkan pengujian parameter secara individu tidak signifikan tetapi pengujian serentak signifikan. Salah satu cara mengidentifikasi adanya multikolinieritas yaitu dengan *Variance Inflation Factors* (VIF) yang dinyatakan sebagai berikut (Gujarati DN., 2003):

$$VIF_j = \frac{1}{(1 - R_j^2)} \quad ; j = 1, 2, \dots, p \quad (2.38)$$

dimana  $R_j^2$  adalah koefisien determinasi antara  $x_j$  dengan variabel prediktor lainnya. Nilai VIF yang lebih besar dari 10 menunjukkan adanya kolinieritas antar variabel prediktor. Jika terjadi multikolinieritas maka ada beberapa cara untuk mengatasinya antara lain dengan mengeluarkan variabel prediktor yang berkorelasi tinggi, melakukan transformasi data, menambah data, menggunakan analisis komponen utama (*principal component analysis*), menggunakan analisis *ridge regression* dan lain-lain.

### 2.5.8 Uji Goodness of Fit

Uji *goodness of fit* digunakan untuk mengetahui apakah model yang digunakan layak atau tidak. Berikut langkah-langkah pengujian (Jong dan Heller, 2008)

Hipotesis :

$H_0$  : Model layak

$H_1$  : Model tidak layak

Taraf nyata :  $\alpha$

Statistik Uji :

$$Deviance = 2v \sum_{i=1}^n \left\{ -\ln \left( \frac{y_i}{\hat{\mu}_i} \right) + \left( \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i} \right) \right\} \quad (2.39)$$

Kriteria Keputusan :  $H_0$  ditolak jika  $Deviance > \chi_{a(n-p)}^2$

$n$  : banyaknya pengamatan,  $p$  : banyaknya parameter.

Kesimpulan : Jika nilai  $Deviance \leq \chi_{a(n-p)}^2$ , maka dapat

disimpulkan bahwa model layak

## 2.6 Ukuran Kebaikan Model

Terdapat ukuran kebaikan hasil prediksi data validasi (*out of sample*):

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}} \quad (2.40)$$

$RMSEP$  = Root Mean Square Error Prediction

$n$  = jumlah sampel

$Y_i$  = nilai aktual ke- $i$

$\hat{Y}_i$  = nilai prediksi ke- $i$

## 2.7 Hubungan Empiris Curah Hujan dan Limpasan Menggunakan Metode Rasional

Terdapat beberapa metode dalam menentukan banjir rencana salah satunya adalah dengan menggunakan metode rasional. Metode Rasional merupakan pemodelan hidrologi sederhana yang sering digunakan untuk mengestimasi debit puncak suatu DAS. Konsep yang terdapat pada metode Rasional terbilang canggih karena membutuhkan pengetahuan teknik yang sangat dalam terutama dalam karakteristik hidrologi seperti waktu konsentrasi (Hayes dan Young, 2005). Metode ini sudah banyak digunakan sejak pertengahan abad 19 dan telah banyak digunakan untuk perencanaan banjir di wilayah perkotaan (Chow et.al., 1998). Parameter-parameter banjir yang digunakan dalam metode ini adalah koefisien *runoff* / limpasan, intensitas curah hujan, dan luas DAS. Metode rasional ini digunakan untuk daerah perkotaan dengan luas DAS kurang lebih 200 acres atau  $\pm 81$  ha (Subarkah, 1980).

Rumus matematis :

$$Q = 0.278 C I A \quad (2.41)$$

Dimana :

- $C$  = koefisien *runoff* / limpasan  
 $I$  = intensitas curah hujan (mm/jam)  
 $A$  = luas daerah aliran (km<sup>2</sup>)  
 $Q$  = debit maksimum (m<sup>3</sup>/detik)

Beberapa asumsi dasar untuk menggunakan metode rasional adalah :

1. Curah hujan terjadi dengan intensitas yang tetap dalam jangka waktu tertentu, setidaknya sama dengan waktu konsentrasi.
2. Limpasan langsung mencapai maksimum ketika durasi hujan dengan intensitas tetap sama dengan waktu konsentrasi.
3. Koefisien *run off* dianggap tetap selama durasi hujan.
4. Luas DAS tidak berubah selama durasi hujan (Wanielista, 1990)

## 2.8 Metode Mononobe

Dalam perencanaan bangunan pengairan (misalnya drainase), debit rencana sangat diperlukan untuk mengetahui kapasitas yang seharusnya dapat ditampung oleh sebuah drainase, agar semua debit air dapat ditampung dan teralirkan. Metode mononobe adalah metode yang digunakan untuk menghitung intensitas curah hujan. Persamaannya adalah sebagai berikut.

$$I = \frac{R_{24}}{24} \left( \frac{24}{t_c} \right)^{\frac{2}{3}} \quad (2.42)$$

dimana :

- $I$  : Intensitas curah hujan (mm/jam)  
 $t_c$  : Lamanya curah hujan / durasi curah hujan (jam)  
 $R$  : Curah hujan maksimum

Dapat diketahui untuk memperoleh nilai  $t_c$  dapat dirumuskan sebagai berikut.

$$t_c = t_0 + t_f \quad (2.43)$$

## Keterangan

- $t_0$  : waktu yang dibutuhkan untuk mengalir di permukaan untuk mencapai inlet (*overland flow time, inlet time*)  
 $t_f$  : waktu yang diperlukan untuk mengalir di sepanjang saluran

Untuk menentukan harga  $t_0$ , diasumsikan permukaan dengan penutup homogen sehingga untuk perhitungan dapat menggunakan perumusan Kerby sebagai berikut.

$$t_0 = 1,44 \left( n_d \times \frac{l}{\sqrt{s}} \right)^{0,467} \quad (2.44)$$

## Keterangan

- $l$  : jarak dari titik terjauh ke inlet (m)  
 $n_d$  : koefisien setara koefisien kekasaran  
 $s$  kemiringan saluran

Pada saat  $Q_{hidrolika} < Q_{hidrologi}$  maka dapat dikatakan saluran air atau sungai tidak dapat menampung air hujan dengan baik yang natinya dapat menjikan air meluap dan terjadinya banjir. Agar mendapat desain yang ekonomis maka seharusnya  $Q_{hidrologi} \approx Q_{hidrolika}$ . Q hidrolika dapat dirumuskan sebagai berikut.

$$Q_{hidrolika} = V_{(m/detik)} \times A_{(m^2)} = \frac{1}{n} \times R^{2/3} \times S^{1/2} \times A$$

Dimana A merupakan luas penampang sungai yang diperoleh dari

$$A_{(m^2)} = b_{(m)} \times h_{(m)}$$

## Keterangan

- $b_{(m)}$  : lebar sungai  
 $h_{(m)}$  : tinggi sungai

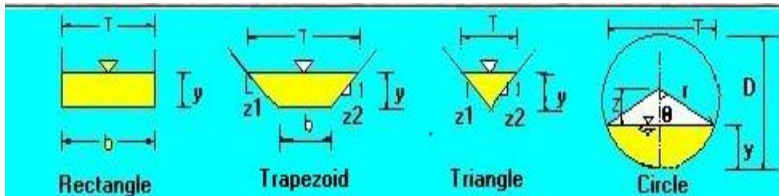
$b_{(m)}$  diperoleh dari rata-rata lebar sungai minimum dan maksimum yang diperoleh data PU Binamarga dan Pematusan. Sedangkan untuk  $h_{(m)}$  dari tabel u-ditch tergantung nilai  $b_{(m)}$ -nya. Untuk memudahkan dalam analisis, maka dapat diketahui  $h_{(m)}$  suatu



saluran adalah sama dengan  $b(m)$  jika  $b_{(m)}$  lebih besar sama dengan

12. Jika  $b_{(m)} < 12$  maka  $h_{(m)} = b_{(m)} \times \frac{10}{6}$ . Hal ini dikarenakan,

keseluruhan saluran di Kota Surabaya menggunakan *u-ditch* atau saluran *rectangle*. Berikut beberapa bentuk saluran yang sering digunakan.



**Gambar 2.1** Bentuk Saluran

Sedangkan untuk kecepatan saluran dapat menggunakan Rumus Manning sebagai berikut.

$$V_{(m/detik)} = \frac{1}{n} R^{2/3} S^{1/2}$$

Keterangan

- $n$  : koefisien kekerasan saluran
- $R$  : jari-jari hidrolis
- $S$  : kemiringan salur

Untuk jari-jari hidrolis dapat dirumuskan sebagai berikut.

$$R_{(m)} = \frac{A_{(m^2)}}{P_{(m)}}$$

Dimana  $P_{(m)}$  atau penampang yang terkena air dapat dirumuskan sebagai berikut.

$$P_{(m)} = b_{(m)} + 2.h_{(m)}$$

Untuk menentukan kecepatan aliran di saluran tergantung material saluran yang digunakan. Semakin halus permukaan saluran maka kecepatan semakin besar. Saluran juga harus dapat menahan kecepatan agar tidak tergerus aliran. Kekasaran juga merupakan

hambatan pada aliran, berikut tabel koefisien kekasaran saluran yang sering digunakan.

**Tabel 4. 1** Tabel Kekasaran Manning untuk Saluran

Saluran	Keterangan	<i>n</i> Manning
Tanah	Lurus, baru, seragam, landai dan bersih	0,016-0,033
	Berkelok, landai dan berumput	0,023-0,040
	Tidak terawat dan kotor	0,050-0,140
	Tanah berbatu, kasar dan tidak teratur	0,035-0,045
Pasangan	Batu kosong	0,023-0,035
	Pasangan batu belah	0,017-0,030
Beton	Halus, sambungan baik dan rata	0,014-0,018
	Kurang halus dan sambungan kurang rata	0,018-0,030

## 2.9 Curah Hujan dan Banjir

Penelitian ini akan dilakukan analisa curah hujan untuk memprediksi banjir di kota Surabaya. Sebelum dilakukan analisa lebih lanjut, terlebih dahulu perlu diketahui mengenai curah hujan dan banjir.

### 2.8.1 Curah Hujan

Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, dan tidak mengalir. Curah hujan 1 (satu) milimeter artinya adalah luasan satu meter persegi pada tempat yang datar tertampung air sebanyak satu liter. Menurut pengertian klimatologi, satu hari hujan adalah periode 24 jam dimana terkumpul curah hujan setinggi 0.5 mm atau lebih. Kurang dari ketentuan ini hujan dinyatakan nol, meskipun tinggi curah hujannya tetap diperhitungkan. Curah hujan di suatu daerah tidaklah selalu sama dengan di daerah lain. Ada suatu daerah yang pada akhir tahun hujannya mulai meningkat tinggi dan mencapai puncaknya dan pertengahan tahun mencapai titik terendahnya. Sebaliknya, di daerah lain pada akhir tahun hujannya mencapai titik terendah, sedangkan pada pertengahan tahun mencapai titik tertingginya.

Curah hujan menurut BMKG dibagi menjadi empat kelompok, yaitu:

1. Curah hujan rendah: 0-20 mm, 21-50 mm, 51-100 mm.

2. Curah hujan menengah: 101-150 mm, 151-200 mm, 201-300 mm.
3. Curah hujan tinggi: 301-400mm.
4. Curah hujan sangat tinggi: 401-500 mm >500 mm.

Salah satu faktor yang mempengaruhi curah hujan adalah kelembaban udara. Kelembaban udara adalah perbandingan antara massa uap dalam suatu satuan volum dengan massa uap yang jenuh dalam satuan volum itu pada suhu yang sama. Secara umum kelembaban menyatakan banyaknya kadar air yang ada di udara. Banyaknya uap yang bergerak di dalam atmosfer berpengaruh terhadap besarnya hujan, lamanya hujan, dan intensitas curah hujan. Kelembaban tertinggi umumnya terjadi pada musim penghujan dan paling rendah pada musim kemarau. Variasi kelembaban bergantung dari suhu udara, jika pagi hari suhu rendah, maka kelembaban akan lebih tinggi jika dibanding pada siang hari saat suhu tinggi.

### **2.8.2 Banjir**

Banjir merupakan bencana yang disebabkan oleh fenomena alam yang terjadi selama musim hujan yang meliputi potensi daerah, terutama sungai yang relatif landai. Selain itu, banjir juga bisa disebabkan oleh naiknya air yang disebabkan oleh hujan deras di atas normal, perubahan suhu, tanggul yang rusak, dan obstruksi aliran air di lokasi lain. Banjir dapat menyebabkan kerusakan besar pada kehidupan sosial dan ekonomi masyarakat. Menurut Undang-Undang No. 4 tahun 2008 bencana merupakan peristiwa atau rangkaian peristiwa yang mengancam dan mengganggu kehidupan dan penghidupan masyarakat yang disebabkan baik oleh faktor alam dan non alam maupun factor manusia sehingga mengakibatkan timbulnya korban jiwa manusia, kerusakan lingkungan, kerugian harta benda dan dampak psikologis. Bencana banjir dapat diartikan suatu peristiwa yang mengancam dan mengganggu kehidupan dan penghidupan masyarakat yang disebabkan oleh bencana banjir yang berdampak timbulnya korban manusia, kerusakan lingkungan, kerugian harta benda dan dampak psikologis.

Menurut Richards, 1955 dalam Suherlan, 2001, *Flood Estimation and Control* bahwa banjir memiliki dua arti yaitu (1) meluapnya air sungai disebabkan oleh debit sungai yang melebihi daya tampung sungai pada keadaan curah hujan yang tinggi dan (2) banjir merupakan genangan pada daerah rendah yang datar yang biasanya tidak tergenang. Kodoatie dan Sugiyanto (2002) menyebutkan bahwa banjir terdiri atas dua peristiwa, pertama banjir terjadi di daerah yang tidak biasa terkena banjir dan kedua banjir terjadi karena limpasan air dari sungai karena debitnya yang besar sehingga tidak mampu dialirkan oleh alur sungai. Menurut Suripin (2004) penyebab banjir dapat dibedakan menjadi tiga yaitu:

1. Banjir kiriman  
Aliran banjir yang datang dari daerah hulu di luar kawasan yang tergenang. Hal ini terjadi jika hujan yang terjadi di daerah hulu menimbulkan aliran banjir yang melebihi kapasitas sungainya atau banjir kanal yang ada, sehingga mengakibatkan terjadinya limpasan.
2. Banjir lokal (banjir genangan)  
Genangan air yang timbul akibat hujan yang jatuh di daerah itu sendiri. Hal ini dapat terjadi apabila hujan yang terjadi melebihi kapasitas sistem drainase yang ada. Pada banjir lokal, ketinggian genangan air antara 0.2-0.7 m dan lama genangan 1-8 jam. Banjir ini terdapat pada daerah yang rendah.
3. Banjir rob (banjir air pasang)  
Banjir yang terjadi baik akibat aliran langsung air pasang dan/ atau air balik dari saluran drainase akibat terhambat oleh air pasang.

Bencana banjir dapat terjadi karena faktor alamiah maupun pengaruh perlakuan masyarakat terhadap alam dan lingkungannya. Sedikitnya ada lima faktor penting penyebab banjir di Indonesia yaitu faktor hujan, faktor hancurnya retensi daerah aliran sungai (DAS), faktor kesalahan perencanaan pembangunan alur sungai, faktor pendangkalan sungai dan faktor kesalahan tata wilayah dan

pembangunan sarana dan prasarana (Maryono, 2005). Faktor alamiah yang utama dari terjadinya banjir adalah curah hujan.

*(Halaman ini sengaja dikosongkan)*

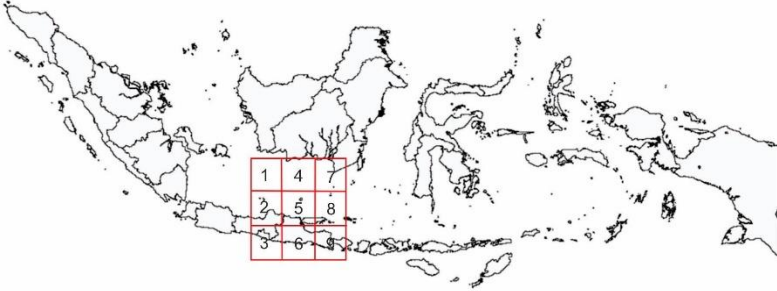
## BAB III METODOLOGI PENELITIAN

Metodologi penelitian ini akan menjelaskan tentang sumber data, definisi variabel, variabel penelitian, langkah analisis beserta diagram alirnya sebagai penunjang dalam melakukan analisis dan pembahasan.

### 3.1 Sumber Data

Sumber data yang digunakan pada penelitian ini merupakan data sekunder yang diperoleh dari sumber yang berbeda pada masing-masing variabel. Data yang digunakan merupakan data curah hujan harian mulai tahun 2010 hingga tahun 2018 yang diperoleh dari Balai PSAWS Buntung Peketingan Surabaya. Selain itu, variabel-variabel yang diduga mempengaruhi curah hujan dapat diperoleh dari General Circulation Model (GCM). GCM merupakan model matematis yang menggambarkan sirkulasi umum atmosfer di bumi. Model ini menduga perubahan unsur-unsur cuaca dalam bentuk luaran grid-grid yang berukuran 100-500 km menurut lintang dan bujur (von Storch *et.al.* 1993 dalam Sutikno, 2008). GCM merupakan suatu alat penting dalam studi keragaman iklim dan perubahan iklim (Zorita dan Storch, 1999). Namun informasi GCM masih berskala global, sehingga sulit untuk memperoleh langsung informasi berskala lokal dari GCM. Tetapi GCM masih mungkin digunakan untuk memperoleh informasi skala lokal atau regional bila teknik *downscaling* digunakan (Fernandez, 2005 dalam Wigena, 2006). Data yang digunakan adalah peubah-peubah luaran CSIRO Mk3 dari Australia dengan resolusi lintang-bujur 1.8650 x 1.8750 pada tahun 2009-2018 dengan lokasi grid berada ditengah tengah Kota Surabaya. Grid yang digunakan adalah 1x1 dan 3x3 dengan domain 3x3 yang digunakan adalah -4.66<sup>0</sup> LS -8.39<sup>0</sup> LS, 110.6<sup>0</sup> BT-114.3<sup>0</sup> BT. Variabel prediktor dari luaran GCM yang digunakan yaitu *precipitation* (pr), suhu (tas), kecepatan angin (sfcwind) dan kelembapan spesifik (huss). Variabel prediktan yang digunakan yaitu data curah hujan harian pada 6 stasiun

(Gunungsari, Wonokromo, Wonorejo, Keputih, Kedungcowek, dan Gubeng) di Surabaya. Lokasi pengamatan pada penelitian ini terletak pada koordinat *latitude* dan *longitude* : -7.25, 112.76 yaitu tepat pada lokasi nomor 5 seperti pada gambar 3.1



**Gambar 3.1** Lokasi Pengamatan Penelitian

Penentuan lokasi stasiun berdasarkan *latitude* dan *longitude* dapat dilihat pada tabel 3.1

**Tabel 3.1** Penentuan Lokasi Stasiun

		Longitude		
		110.625	112.5	114.375
Latitude	-8.3936	1	4	7
	-6.5284	2	5	8
	-4.6631	3	6	9

### 3.2 Definisi Variabel

Variabel yang digunakan dalam penelitian ini terdiri dari variabel respon dan empat variabel prediktor. Adapun variabel-variabel tersebut adalah sebagai berikut.

#### a. Variabel Respon

Penelitian ini menggunakan variabel respon yaitu Curah Hujan. Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, dan tidak mengalir. Curah hujan 1 (satu) milimeter artinya adalah luasan satu meter persegi pada tempat yang datar tertampung air sebanyak satu liter. Menurut pengertian klimatologi, satu hari hujan adalah



periode 24 jam dimana terkumpul curah hujan setinggi 0.5 mm atau lebih. Curah hujan diukur menggunakan alat yang dinamakan ombrometer. Berdasarkan data dari UPT PSDA WS Brantas, alat pengukur curah hujan diletakkan di enam wilayah yang ada di Kota Surabaya.

#### **b. Variabel Prediktor**

Penelitian ini menggunakan empat variabel yang diduga berpengaruh terhadap Curah Hujan. Adapun keempat variabel prediktor tersebut adalah sebagai berikut:

##### 1. Presipitasi

Presipitasi adalah curah hujan atau turunnya air dari atmosfer ke permukaan bumi dan laut dalam bentuk yang berbeda, yaitu curah hujan di daerah tropis dan curah hujan serta salju dan di daerah beriklim sedang. Presipitasi adalah faktor utama yang mengendalikan berlangsungnya daur hidrologi dalam suatu wilayah (merupakan elemen utama yang perlu diketahui mendasari pemahaman tentang kelembaban tanah, proses peresapan air tanah, dan debit aliran).

##### 2. Kelembapan

Kelembapan adalah konsentrasi uap air di udara. Angka konsentrasi ini dapat diekspresikan dalam kelembapan absolut, kelembapan spesifik, atau kelembapan relatif. Kelembapan spesifik adalah metode untuk mengukur jumlah uap air di udara dengan rasio terhadap uap air diudara kering. Kelembapan berbanding lurus dengan curah hujan, semakin tinggi kelembapan udara maka akan semakin tinggi pula curah hujan yang terjadi. Sebaliknya semakin rendah kelembapan udara maka akan semakin rendah pula curah hujan yang terjadi.

##### 3. Suhu

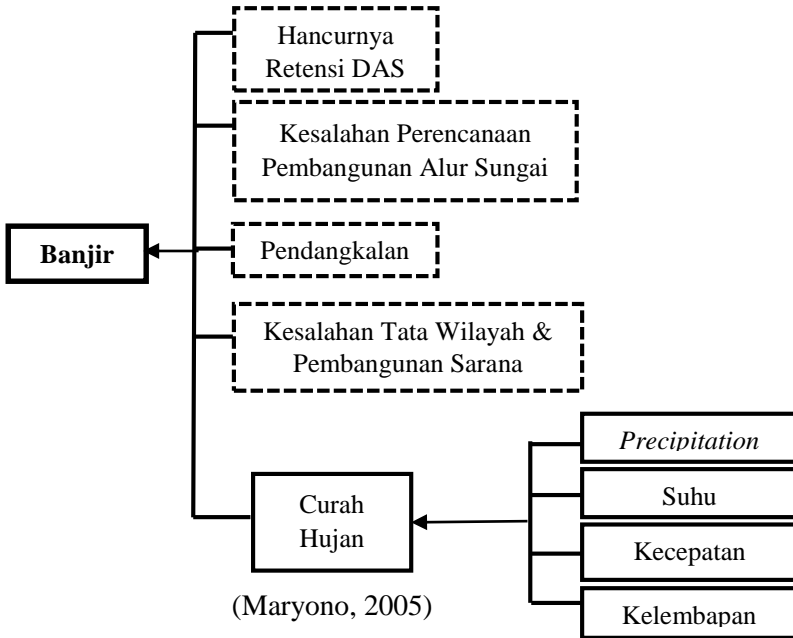
Suhu merupakan suatu besaran yang menunjukkan derajat panas suatu wilayah. Semakin rendah suhu atau dingin udara diatmosfer menyebabkan terbentuknya es dan air, karena angin tidak kuat menopang air pada awan tersebut maka air jatuh dan jadilah hujan atau yg disebut presipitasi

4. Kecepatan Angin

Kecepatan Angin atau velositas gelombang angin adalah sebuah kuantitas atmosferik fundamental. Kecepatan angin disebabkan oleh pergerakan angin dari tekanan tinggi ke tekanan rendah, biasanya terjadi karena perubahan suhu. Kecepatan angin berbanding lurus dengan curah hujan, semakin tinggi kecepatan angin maka akan semakin tinggi pula curah hujan yang terjadi. Sebaliknya semakin rendah kecepatan angin maka akan semakin rendah pula curah hujan yang terjadi.

3.3 Variabel Penelitian

Rincian variabel yang akan dianalisis pada penelitian ini dapat dijelaskan sebagai berikut.



Gambar 3.2 Rancangan variabel penelitian

Banjir pada umumnya disebabkan oleh curah hujan tinggi. Selain itu banjir juga dapat dipengaruhi oleh faktor lain seperti

pembuangan sampah sembarangan, kurangnya saluran drainase, dan pasangannya air laut. Namun data-data tersebut sedikit sulit untuk diperoleh. Maka dari itu pada penelitian ini digunakan variabel yang datanya mudah untuk diperoleh seperti data curah hujan. Curah hujan sendiri dipengaruhi oleh variabel-variabel klimatologi yang disimbolkan sebagai variabel prediktor X pada Tabel 3.1.

**Tabel 3.2** Variabel Penelitian

No.	Variabel	Keterangan	Satuan	Sumber Data
1	$Y_t$	Curah Hujan	mm	Balai PSAWS
2	$X_1$	<i>Precipitation</i>	$\text{Kgm}^{-2}\text{s}^{-1}$	GCM
3	$X_2$	Kelembapan	ltr	GCM
4	$X_3$	Suhu	K	GCM
5	$X_4$	Kecepatan Angin	$\text{ms}^{-1}$	GCM

Variabel penelitian yang digunakan pada penelitian ada empat variabel klimatologi. Apabila hasil prediksi kurang bisa menyerupai data asli curah hujan maka akan ditambahkan variabel dummy dan lag dari curah hujan. Variabel dummy dikategorikan 0 apabila tidak terjadi hujan, yaitu pada bulan Mei sampai dengan Oktober dan untuk hari-hari hujan dikategorikan 1 yaitu terjadi pada bulan November sampai dengan April. Struktur data secara umum yang digunakan dalam penelitian ini terdapat pada Tabel 3.2. Struktur data terdiri atas data pengukuran yang berasal dari setiap rayon dan data luaran GCM

**Tabel 3.3** Struktur Data di setiap Stasiun Grid 1x1

Tanggal	t	$Y_t$	Data Pengukuran dari GCM			
			$X_1$	$X_2$	$X_3$	$X_4$
1/1/2009	1	$Y_{t_1}$	$X_{1,1}$	$X_{2,1}$	$X_{3,1}$	$X_{4,1}$
2/1/2009	2	$Y_{t_2}$	$X_{1,2}$	$X_{2,2}$	$X_{3,2}$	$X_{4,2}$
3/1/2009	3	$Y_{t_3}$	$X_{1,3}$	$X_{2,3}$	$X_{3,3}$	$X_{4,3}$
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	t	$Y_{t,t}$	$X_{1,t}$	$X_{2,t}$	$X_{3,t}$	$X_{4,t}$
:	:	:	:	:	:	:
:	:	:	:	:	:	:
31/12/2018	n	$Y_{n,n}$	$X_{1,n}$	$X_{2,n}$	$X_{3,n}$	$X_{4,n}$

**Tabel 3.4** Struktur Data di setiap Stasiun Grid 3x3

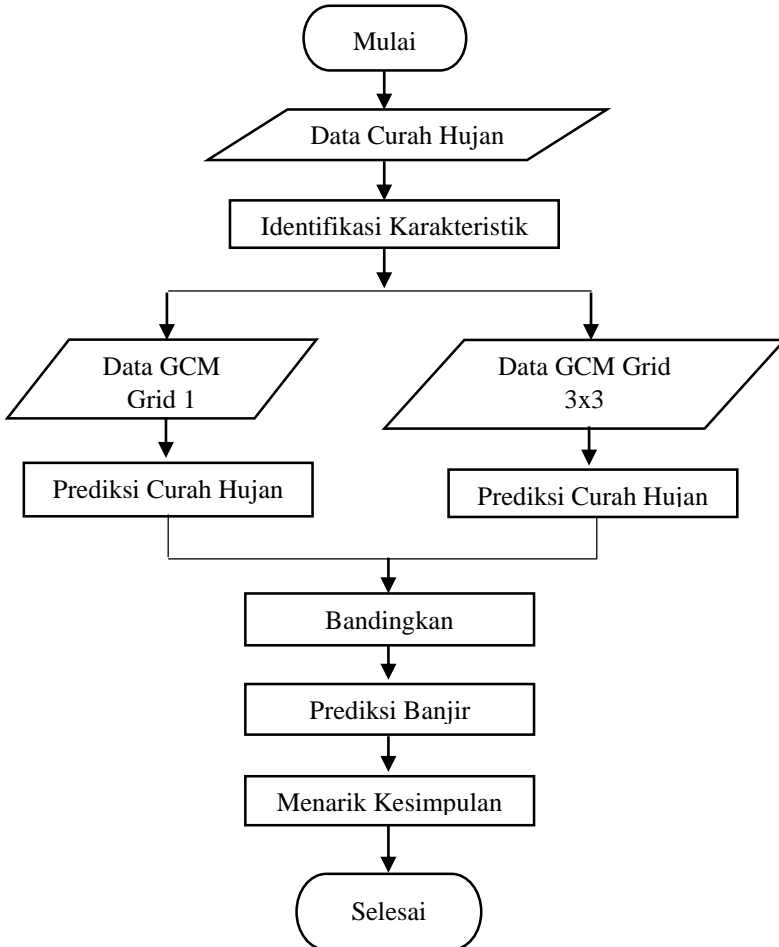
Tanggal	t	Curah Hujan	Presipitasi			Kelembapan			Suhu			Kecepatan Angin		
			X <sub>11</sub>	...	X <sub>19</sub>	X <sub>21</sub>	...	X <sub>29</sub>	X <sub>31</sub>	...	X <sub>39</sub>	X <sub>41</sub>	...	X <sub>49</sub>
1/1/2009	1	Y <sub>t1</sub>	X <sub>11,1</sub>	:	X <sub>19,1</sub>	X <sub>21,1</sub>	:	X <sub>29,1</sub>	X <sub>31,1</sub>	:	X <sub>39,1</sub>	X <sub>41,1</sub>	:	X <sub>49,1</sub>
2/1/2009	2	Y <sub>t2</sub>	X <sub>11,2</sub>	:	X <sub>19,2</sub>	X <sub>21,2</sub>	:	X <sub>29,2</sub>	X <sub>31,2</sub>	:	X <sub>39,2</sub>	X <sub>41,2</sub>	:	X <sub>49,2</sub>
3/1/2009	3	Y <sub>t3</sub>	X <sub>11,3</sub>	:	X <sub>19,3</sub>	X <sub>21,3</sub>	:	X <sub>29,3</sub>	X <sub>31,3</sub>	:	X <sub>39,3</sub>	X <sub>41,3</sub>	:	X <sub>49,3</sub>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	t	Y <sub>t,t</sub>	X <sub>11,t</sub>	:	X <sub>19,t</sub>	X <sub>21,t</sub>	:	X <sub>29,t</sub>	X <sub>31,t</sub>	:	X <sub>39,t</sub>	X <sub>41,t</sub>	:	X <sub>49,t</sub>
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
31/12/2018	n	Y <sub>n,n</sub>	X <sub>11,n</sub>	:	X <sub>19,n</sub>	X <sub>21,n</sub>	:	X <sub>29,n</sub>	X <sub>31,n</sub>	:	X <sub>n</sub>	X <sub>41,n</sub>	:	X <sub>49,n</sub>

### 3.4 Langkah Analisis

Berikut ini adalah langkah-langkah analisis yang dilakukan untuk mencapai tujuan dari penelitian.

1. Mendeskripsikan karakteristik data curah hujan dan data variabel-variabel luaran GCM
  - a. Melakukan *pre-processing* pada data curah hujan dan membagi data menjadi data *Training* dan *Testing*
  - b. Melakukan analisis statistika deskriptif secara umum meliputi rata-rata, deviasi standar, maksimum, dan minimum.
  - c. Melakukan analisis pola hubungan variabel-variabel luaran GCM dengan data curah hujan dengan menggunakan *scatterplot*
  - d. Menyajikan data dalam bentuk *box plot* dan *time series plot*.
2. Melakukan prediksi curah hujan dengan *Statistical Downscaling* pada data keseluruhan dan pada data dengan periode tertentu
  - a. Melakukan reduksi dimensi domain grid untuk mengatasi multikolinearitas pada variabel luaran GCM menggunakan analisis komponen utama (PCA)
  - b. Melakukan pemodelan variabel hasil reduksi dimensi dengan variabel curah hujan di 6 stasiun menggunakan regresi gamma dengan variabel prediktor pada grid satu saja dan dengan variabel prediktor pada grid 3x3
  - c. Membandingkan akurasi prediksi curah hujan yang dihasilkan pada pemodelan GLM dengan variabel prediktor pada grid satu saja dan dengan variabel prediktor pada grid 3x3.
3. Memprediksi banjir berdasarkan hasil prediksi curah hujan yang terbaik
  - a. Menyiapkan data luas daerah aliran sungai (DAS) dan hasil prediksi curah hujan
  - b. Melakukan perhitungan banjir dengan metode rasional

Langkah analisis dapat digambarkan dalam diagram alir pada Gambar 3.2.



**Gambar 3.3** Diagram Alir Penelitian Secara Umum

## **BAB IV**

### **ANALISIS DAN PEMBAHASAN**

Pada penelitian ini, analisis dan pembahasan yang berisi mengenai metode yang digunakan untuk *Statistical Downscaling* yaitu dengan metode PCA (*Principal Components Analysis*) dan Regresi Gamma. Pada penelitian ini dilakukan pemodelan di masing-masing stasiun curah hujan dan pemilihan model terbaik dengan membandingkan nilai RMSE setiap model pada data *Training*. Selanjutnya dilakukan prediksi curah hujan untuk satu bulan kedepan dengan menggunakan model terbaik.

#### **4.1 Identifikasi Karakteristik**

Banjir merupakan suatu bencana yang menimbulkan banyak kerugian bagi masyarakat. Penyebab utama dari banjir ialah curah hujan tinggi yang terjadi dalam kurun waktu tertentu. Salah satu cara untuk dapat mengelola resiko terjadinya banjir adalah dengan memperkirakan kapan suatu daerah akan berpotensi mengalami banjir. Analisa ini dapat dilakukan dengan memperkirakan potensi terjadinya hujan lebat atau curah hujan tinggi. Sebelum dilakukan analisa banjir terlebih dahulu perlu dilakukan analisa data curah hujan dan variabel-variabel yang mempengaruhinya yang diperoleh dari data GCM. Karakteristik data curah hujan dan variabel-variabel luaran GCM ditunjukkan melalui visualisasi data serta pola data yang menunjang pemodelan selanjutnya.

##### **4.1.1 Statistika Deskriptif**

Analisis data curah hujan pada tahun 2009 sampai tahun 2018 di Kota Surabaya pada enam stasiun yaitu stasiun Keputih, stasiun Kedung Cowek, stasiun Gubeng, stasiun Wonorejo, stasiun Wonokromo, dan stasiun Gunungsari untuk mengetahui karakteristik data serta mengetahui kondisi dari ukuran pemusatan data dan ukuran penyebaran data. Berdasarkan tabel statistika deskriptif dapat diperoleh informasi yaitu rata-rata curah hujan paling tinggi terdapat pada stasiun Gubeng yaitu sebesar 6.487 mm per tahun dengan nilai variansi yang cukup tinggi dibandingkan

dengan stasiun-stasiun lainnya yaitu sebesar 196, rata-rata curah hujan tinggi selanjutnya terdapat pada stasiun Wonokromo dan Gunungsari dengan nilai variansi yang paling tinggi dibandingkan dengan stasiun-stasiun lainnya yaitu sebesar 198 dan 197. Sedangkan rata-rata curah hujan paling rendah terdapat pada stasiun Kedung Cowek 5.132 mm per tahun. Curah hujan sangat tinggi sebesar 164 mm terjadi di daerah Keputih dengan nilai variansi sebesar 174. Pada stasiun Wonokromo terjadi curah hujan paling tinggi yaitu sebesar 114 mm yang merupakan curah hujan maksimum yang paling rendah jika dibandingkan dengan stasiun Keputih, Kedung Cowek, Gubeng, Wonorejo, dan Gunung Sari. Daerah dengan variansi curah hujan yang paling rendah terdapat pada stasiun Kedung Cowek.

**Tabel 4.1** Statistika Deskriptif Data Curah Hujan

<i>Variable</i>	<i>Mean</i>	<i>Variance</i>	<i>Minimum</i>	<i>Maximum</i>
CH_Keputih	5.513	174.097	0	164
CH_Cowek	5.132	129.972	0	123
CH_Gubeng	6.487	196.242	0	116
CH_Wonorejo	5.860	189.957	0	122
CH_Wonokromo	6.419	198.935	0	114
CH_GnSari	6.361	197.468	0	120

Curah hujan yang terjadi pada suatu daerah dapat dipengaruhi oleh beberapa variabel seperti Presipitasi, Kelembapan, Suhu, dan Kecepatan Angin. Berikut merupakan statistika deskriptif variabel tersebut yang merupakan output luaran GCM.

**Tabel 4.2** Statistika Deskriptif Data Luaran GCM

<i>Variable</i>	<i>Mean</i>	<i>Variance</i>	<i>Minimum</i>	<i>Maximum</i>
PR1	0.000054	0.000000	0.000000	0.001037
PR2	0.000108	0.000000	0.000000	0.001129
PR3	0.000132	0.000000	0.000000	0.001438
PR4	0.000057	0.000000	0.000000	0.001379
PR5	0.000112	0.000000	0.000000	0.001373
PR6	0.000122	0.000000	0.000000	0.001037
PR7	0.000055	0.000000	0.000000	0.001418



<i>Variable</i>	<i>Mean</i>	<i>Variance</i>	<i>Minimum</i>	<i>Maximum</i>
PR8	0.000114	0.000000	0.000000	0.001374
PR9	0.000118	0.000000	0.000000	0.001046
TAS1	301.78	1.56	297.50	305.71
TAS2	302.35	0.876	300.04	305.07
TAS3	302.23	0.533	300.47	304.45
TAS4	302.12	2.23	297.63	306.71
TAS5	302.50	0.848	300.37	305.54
TAS6	302.37	0.501	300.81	304.76
TAS7	302.23	2.19	297.51	306.50
TAS8	302.49	0.772	299.91	305.45
TAS9	302.27	0.502	300.52	304.74
HUSS1	0.016552	0.000009	0.008902	0.021616
HUSS2	0.023598	0.000039	0.016402	0.050250
HUSS3	0.020812	0.000001	0.016612	0.033648
HUSS4	0.016580	0.000010	0.008862	0.021953
HUSS5	0.022967	0.000032	0.015675	0.050250
HUSS6	0.020921	0.000001	0.016709	0.035347
HUSS7	0.016270	0.000009	0.008259	0.021893
HUSS8	0.022539	0.000023	0.016448	0.050250
HUSS9	0.021014	0.000001	0.017332	0.028495
SFCWIND1	3.2585	1.6567	0.3360	9.1509
SFCWIND2	3.9360	2.2925	0.6250	11.7070
SFCWIND3	5.7065	5.0528	0.4820	13.6431
SFCWIND4	2.7179	1.1058	0.3500	8.2264
SFCWIND5	4.5238	3.2233	0.3624	13.1878
SFCWIND6	5.7654	5.4836	0.5302	15.1690
SFCWIND7	2.6167	1.0200	0.3733	6.6881
SFCWIND8	4.5703	3.4113	0.6203	12.7388
SFCWIND9	5.1912	4.5972	0.6592	13.3292

Presipitasi adalah curah hujan atau turunnya air dari atmosfer ke permukaan bumi dan laut dalam bentuk yang berbeda, yaitu curah hujan di daerah tropis dan curah hujan serta salju dan di daerah beriklim sedang. Rata-rata presipitasi paling tinggi yaitu

presipitasi lokasi tiga dengan nilai sebesar  $0.000132 \text{ Kg m}^{-2} \text{ s}^{-1}$ , sedangkan rata-rata presipitasi paling rendah terdapat pada lokasi satu.

Rata-rata suhu paling tinggi terdapat pada lokasi lima yaitu tepat pada lokasi Surabaya. Suhu berbanding terbalik dengan curah hujan, semakin tinggi suhu semakin rendah curah hujan yang terjadi di suatu daerah. Hal ini berarti jika suhu di suatu daerah semakin tinggi, maka resiko terjadinya banjir di daerah tersebut semakin rendah. Sebaliknya, semakin rendah suhu di suatu daerah maka semakin tinggi curah hujan yang mengakibatkan resiko resiko terjadinya banjir di daerah tersebut juga semakin tinggi. Rata-rata suhu paling rendah terdapat pada lokasi satu yang relatif jauh dari lokasi Surabaya.

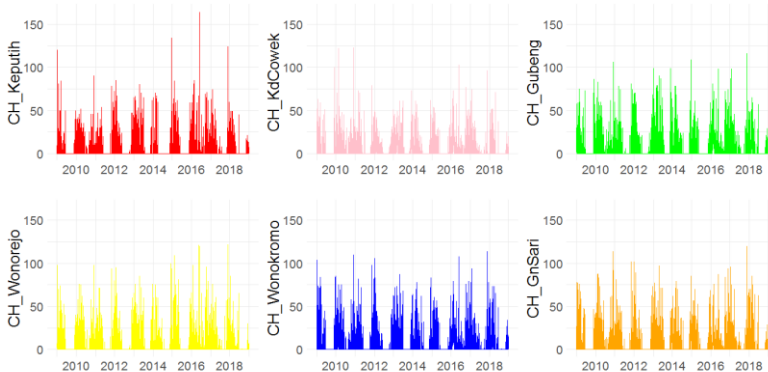
Kelembapan adalah konsentrasi uap air di udara. Rata-rata kelembapan paling tinggi terdapat pada lokasi lima yaitu tepat pada lokasi Surabaya. Kelembapan berbanding lurus dengan curah hujan, semakin tinggi kelembapan udara maka akan semakin tinggi pula curah hujan yang terjadi. Sebaliknya semakin rendah kelembapan udara maka akan semakin rendah pula curah hujan yang terjadi. Kelembapan paling rendah terdapat pada posisi tujuh yaitu tepat pada sebelah timur Surabaya.

Kecepatan angin merupakan pergerakan angin dari tekanan tinggi ke tekanan rendah yang disebabkan oleh perubahan suhu. Rata-rata kecepatan angin paling tinggi terjadi pada lokasi enam yaitu tepat di sebelah selatan lokasi Surabaya. Kecepatan angin berbanding lurus dengan curah hujan, semakin tinggi kecepatan angin maka akan semakin tinggi pula curah hujan yang terjadi. Sebaliknya semakin rendah kecepatan angin maka akan semakin rendah pula curah hujan yang terjadi. Rata-rata kecepatan angin paling rendah terjadi pada lokasi tujuh.

#### **4.1.2 Time Series Plot**

Sebelum dilakukan analisis lebih lanjut, terlebih dahulu perlu dilakukan analisis trend untuk melihat kecenderungan pergerakan data pada periode-periode waktu sebelumnya. Untuk analisis tersebut dapat digunakan time series plot, yaitu diagram

scatter antara data terhadap waktu. Berikut ini time series plot data curah hujan di enam stasiun pada tahun 2009-2018.



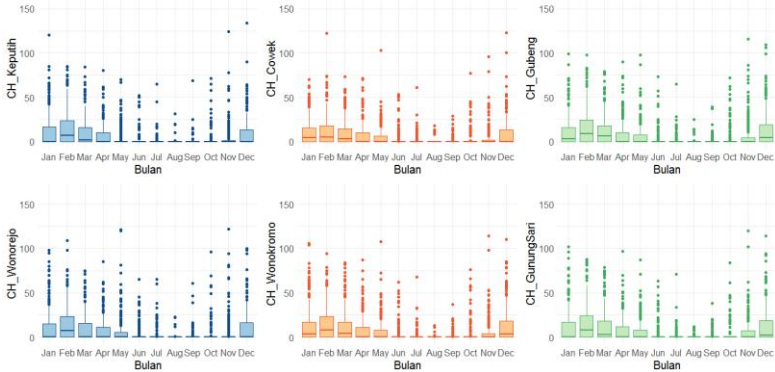
**Gambar 4.1** Time Series Plot Data Curah Hujan (CH)

Berdasarkan time series plot data curah hujan dapat diketahui bahwa data curah hujan pada enam stasiun yaitu Keputih, Kedung Cowek, Gubeng, Wonorejo, Wonokromo, dan Gunung Sari memiliki pola musiman. Hujan di kota Surabaya terjadi pada bulan November sampai April. Hal ini dipengaruhi oleh Angin monsun barat atau yang juga disebut sebagai angin monsun musim dingin timur laut yang bertiup dari kawasan benua Asia ke kawasan benua Australia. Angin bertiup melalui samudra Hindia yang mengandung banyak uap air sehingga pada bulan Oktober hingga April Indonesia mengalami musim penghujan. Curah hujan tertinggi terjadi pada tahun 2016 pada wilayah Keputih.

#### 4.1.3 Box Plot

Boxplot merupakan ringkasan distribusi sampel yang disajikan secara grafis yang bisa menggambarkan bentuk distribusi data (*skewness*), ukuran tendensi sentral, dan ukuran penyebaran (keragaman) data pengamatan. Terdapat lima ukuran statistik yang bisa diketahui dari boxplot yaitu nilai minimum, kuartil terendah/pertama, median, kuartil tertinggi/ketiga, dan nilai maksimum. Selain itu boxplot juga dapat menunjukkan ada

tidaknya nilai outlier dan nilai ekstrim dari data pengamatan. Berikut ini disajikan Boxplot data curah hujan pada keenam stasiun pengamatan dari tahun 2009 sampai tahun 2018

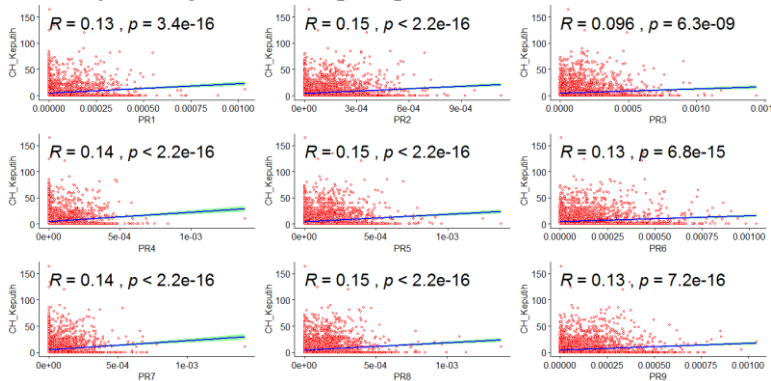


**Gambar 4.2** Box Plot Data Curah Hujan (CH)

Berdasarkan boxplot curah hujan dapat diketahui bahwa secara keseluruhan data curah hujan pada keenam stasiun tidak simetris. Hal ini terlihat dari median yang tidak berada di tengah box dan whisker atas yang lebih panjang dari whisker bawah. Adanya outlier di bagian atas boxplot yang disertai whisker yang lebih panjang menunjukkan bahwa distribusi data cenderung menjulur ke arah kanan (*positive skewness*). Berdasarkan boxplot curah hujan dapat diperoleh informasi bahwa pada keenam stasiun di musim hujan yaitu pada bulan November sampai April, hujan paling banyak terjadi pada bulan februari dan paling sedikit terjadi pada bulan April. Hal ini dapat disebabkan karena April merupakan bulan terakhir pada musim hujan sebelum masuk di musim kemarau pada bulan Mei. Namun pada bulan Mei terlihat masih mengalami hujan di semua stasiun kecuali Keputih meskipun sangat jarang. Begitu juga pada bulan November pada beberapa stasiun seperti Keputih, Kedung Cowek, dan Wonorejo terlihat belum mengalami musim hujan.

#### 4.1.4 Scatterplot

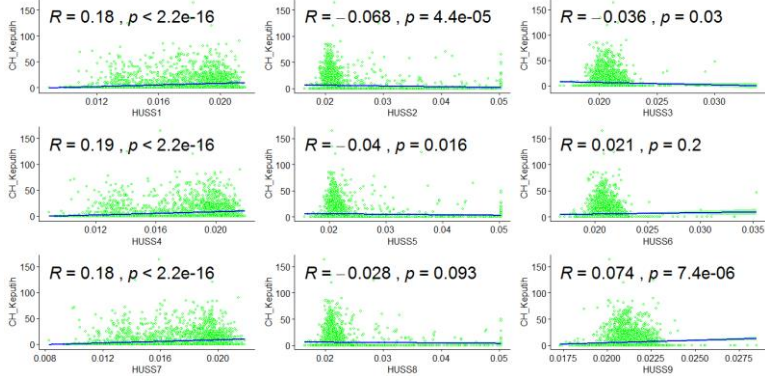
Scatterplot merupakan sebuah grafik yang biasa digunakan untuk melihat suatu pola hubungan antara dua variabel. Sebelum dilakukan pemodelan, terlebih dahulu perlu diketahui apakah terdapat hubungan yang signifikan antara variabel prediktor dengan variabel respon. Berikut ini disajikan scatterplot antara curah hujan dengan variabel presipitasi.



**Gambar 4.3** Scatterplot Curah Hujan (Vertical) dengan Presipitasi (Horizontal)

Berdasarkan scatterplot data curah hujan dengan variabel presipitasi di Sembilan lokasi (grid 3x3) dapat diketahui bahwa nilai koefisien korelasi positif yang berarti curah hujan memiliki hubungan yang berbanding lurus dengan variabel presipitasi. Nilai koefisien korelasi paling tinggi sebesar 0.15 terdapat pada lokasi 2, 5, dan 8. Lokasi 3 memiliki nilai koefisien korelasi yang terendah yaitu sebesar 0.096. Nilai koefisien korelasi yang jauh dari satu mengindikasikan bahwa antara variabel presipitasi dengan curah hujan memiliki hubungan yang rendah.  $P\text{-value} < \text{nilai } \alpha (5\%)$  berarti terdapat hubungan yang signifikan antara variabel presipitasi dengan curah hujan. Sebaliknya jika  $P\text{-value} > \text{nilai } \alpha (5\%)$  berarti tidak terdapat hubungan yang signifikan antara variabel presipitasi dengan curah hujan. Scatterplot data curah hujan menunjukkan bahwa  $P\text{-value} < \text{nilai } \alpha (5\%)$  di semua lokasi. Sehingga dapat disimpulkan bahwa presipitasi di masing-masing lokasi memiliki hubungan yang signifikan dengan curah

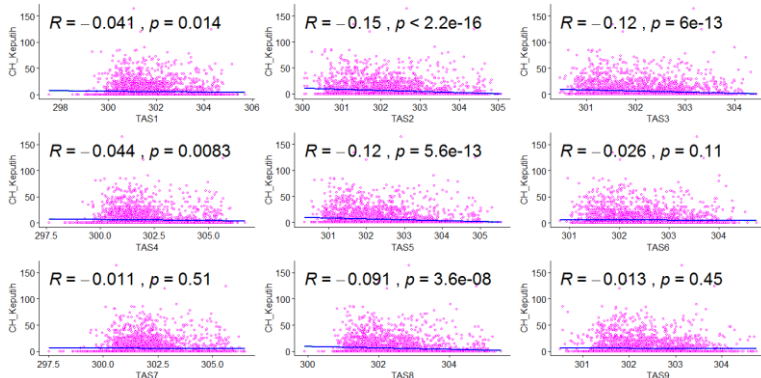
hujan di Keputih. Selain variabel presipitasi, curah hujan juga dapat dipengaruhi oleh variabel kelembapan. Berikut ini disajikan scatterplot antara curah hujan dengan variabel kelembapan.



**Gambar 4.4** Scatterplot Curah Hujan (Vertical) dengan Kelembapan (Horizontal)

Scatterplot data curah hujan dengan variabel kelembapan di Sembilan lokasi (grid 3x3) dapat diketahui bahwa terdapat nilai koefisien korelasi positif dan negative. Nilai koefisien korelasi positif terdapat pada lokasi 1,4,6,7, dan 9. Hal ini berarti bahwa variabel kelembapan pada lokasi 1,4,6,7, dan 9 memiliki hubungan yang berbanding lurus dengan variabel curah hujan. Sebaliknya nilai koefisien korelasi negatif terdapat pada lokasi 2,3,5, dan 8. Hal ini berarti bahwa variabel kelembapan pada lokasi 2,3,5, dan 8 memiliki hubungan yang berbanding terbalik dengan variabel curah hujan. Nilai koefisien korelasi tertinggi terdapat pada lokasi 4 dengan nilai sebesar 0.19. Sedangkan nilai koefisien korelasi terendah terdapat pada lokasi 6 dengan nilai sebesar 0.021. Nilai koefisien korelasi yang jauh dari satu mengindikasikan bahwa antara variabel kelembapan dengan curah hujan memiliki hubungan yang rendah.  $P\text{-value} < \text{nilai } \alpha (5\%)$  berarti terdapat hubungan yang signifikan antara variabel kelembapan dengan curah hujan. Sebaliknya jika  $P\text{-value} > \text{nilai } \alpha (5\%)$  berarti tidak terdapat hubungan yang signifikan antara variabel kelembapan dengan curah hujan. Data curah hujan di wilayah

Keputih dengan variabel kelembapan memiliki hubungan yang signifikan pada semua lokasi kecuali pada lokasi 3, 6, dan 8. Selain variabel presipitasi dan kelembapan, curah hujan juga dapat dipengaruhi oleh variabel suhu. Berikut ini disajikan scatterplot antara curah hujan dengan variabel suhu.

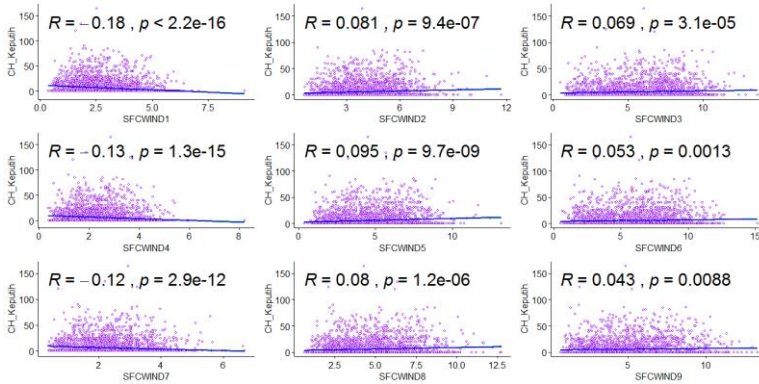


**Gambar 4.5** Scatterplot Curah Hujan (Vertical) dengan Suhu (Horizontal)

Nilai koefisien korelasi di semua lokasi pada scatterplot data curah hujan keputih dengan variabel suhu memiliki tanda negatif. Hal ini berarti bahwa variabel suhu dengan curah hujan di wilayah keputih memiliki hubungan yang berbanding terbalik. Lokasi 2 memiliki nilai koefisien korelasi yang paling tinggi diantara lokasi-lokasi lain yaitu sebesar 0.15, lokasi ini tepat berada di sebelah barat kota Surabaya. Sedangkan nilai koefisien korelasi terendah terdapat pada lokasi 7 yaitu sebesar 0.011. Nilai koefisien korelasi yang jauh dari nilai 1 mengindikasikan bahwa antara variabel suhu dengan curah hujan Keputih memiliki hubungan yang rendah.  $P\text{-value} < \text{nilai } \alpha$  (5%) berarti terdapat hubungan yang signifikan antara variabel suhu pada semua lokasi kecuali pada lokasi 6, 7, dan 9 dengan curah hujan di wilayah Keputih. Selain variabel presipitasi, kelembapan, dan suhu, curah hujan juga dapat dipengaruhi oleh variabel kecepatan angin.

Seperti analisa scatterplot data curah hujan dengan variabel kelembapan, scatterplot data curah hujan dengan variabel kecepatan angin juga terdapat nilai koefisien korelasi positif dan

negative di sembilan lokasi (grid 3x3). Nilai koefisien korelasi positif terdapat pada lokasi 2,3,5,6,8 dan 9. Hal ini berarti bahwa variabel kecepatan angin pada lokasi 2,3,5,6,8 dan 9 memiliki hubungan yang berbanding lurus dengan variabel curah hujan di wilayah Keputih. Sebaliknya nilai koefisien korelasi negatif terdapat pada lokasi 1, 4, dan 7. Hal ini berarti bahwa variabel kelembapan pada lokasi 1, 4, dan 7 memiliki hubungan yang berbanding terbalik dengan variabel curah hujan di wilayah keputih. Nilai koefisien korelasi paling tinggi yaitu sebesar 0.18 terdapat pada lokasi 1. Sedangkan lokasi 9 memiliki nilai koefisien korelasi yang paling rendah yaitu sebesar 0.043.

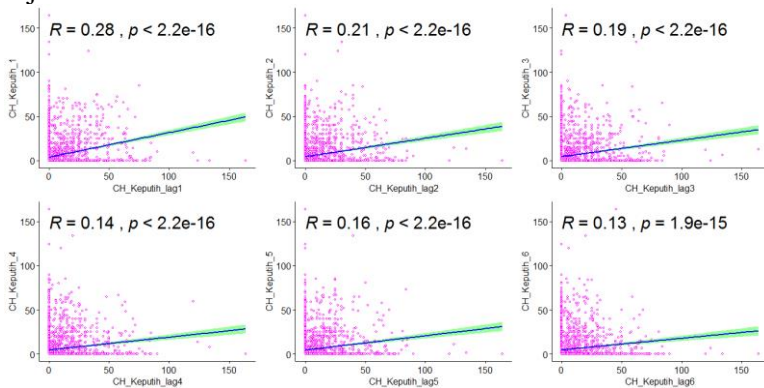


**Gambar 4.6** Scatterplot Curah Hujan (Vertical) dengan Kecepatan Angin (Horizontal)

Nilai koefisien korelasi yang jauh dari nilai 1 mengindikasikan bahwa antara variabel kecepatan angin dengan curah hujan Keputih memiliki hubungan yang rendah.  $P\text{-value} < \text{nilai } \alpha$  (5%) berarti terdapat hubungan yang signifikan antara variabel kecepatan angin dengan curah hujan. Sebaliknya jika  $P\text{-value} > \text{nilai } \alpha$  (5%) berarti tidak terdapat hubungan yang signifikan antara variabel kecepatan angin dengan curah hujan. Scatterplot data curah hujan menunjukkan bahwa  $P\text{-value} < \text{nilai } \alpha$  (5%) di semua lokasi. Sehingga dapat disimpulkan bahwa kecepatan angin di masing-masing lokasi memiliki hubungan yang signifikan dengan curah hujan di wilayah Keputih. Secara



keseluruhan variabel yang memiliki koefisien korelasi paling tinggi dengan curah hujan di wilayah Keputih adalah variabel kelembapan pada lokasi 4. Sedangkan variabel yang memiliki koefisien korelasi paling rendah dengan curah hujan di wilayah Keputih adalah variabel suhu pada lokasi 7. Selain variabel presipitasi, kelembapan, suhu, dan kecepatan angin, curah hujan juga dapat dipengaruhi oleh lag dari curah hujan sendiri. Berikut ini disajikan scatterplot antara curah hujan dengan lag dari curah hujan.



**Gambar 4.7** Scatterplot Curah Hujan (Vertical) dengan Lag Curah Hujan (Horizontal)

Nilai koefisien korelasi dari curah hujan keputih dengan lag nya sendiri terlihat lebih tinggi daripada Nilai koefisien korelasi dari curah hujan keputih dengan variabel lain. Curah hujan dengan curah hujan lag 1 memiliki koefisien korelasi sebesar 0.28, nilai ini merupakan yang paling tinggi jika dibandingkan dengan lag 2, lag 3, lag 4, lag 5, dan lag 6. Terlihat pada grafik scatterplot bahwa nilai koefisien korelasi curah hujan dengan curah hujan lag 1 sampai dengan lag 4 terus mengalami penurunan, namun nilai koefisien korelasi curah hujan dengan curah hujan lag 5 terlihat lebih tinggi dari lag 4 dan kembali mengalami penurunan pada lag 6. Nilai koefisien korelasi positif berarti bahwa curah hujan memiliki hubungan yang berbanding lurus dengan lag curah hujan.

$P$ -value < nilai  $\alpha$  (5%) berarti terdapat hubungan yang signifikan antara curah hujan dengan lag curah hujan. Scatterplot data curah hujan menunjukkan bahwa  $P$ -value < nilai  $\alpha$  (5%) pada lag 1 sampai dengan lag 6 sehingga dapat disimpulkan bahwa curah hujan di wilayah Keputih memiliki hubungan yang signifikan dengan lag curah hujan pada lag 1 sampai dengan lag 6.

## 4.2 Prediksi Curah Hujan

Prediksi curah hujan dengan memodelkan variabel luaran GCM yang telah dibahas sebelumnya yaitu variabel presipitasi, kelembapan, suhu, dan kecepatan angin dengan variabel respon yaitu curah hujan. Pemodelan dilakukan dengan menggunakan Regresi Gamma pada keenam stasiun pengamatan curah hujan yaitu stasiun Keputih, stasiun Kedung Cowek, stasiun Gubeng, stasiun Wonorejo, stasiun Wonokromo, dan stasiun Gunungsari. Pemodelan regresi gamma pada keenam stasiun tersebut akan dilakukan pada grid 1x1 dan grid 3x3. Pemodelan dengan regresi gamma dilakukan untuk memberikan hasil dugaan curah hujan non negatif. Berikut ini pemodelan dan prediksi curah hujan di masing-masing stasiun pengamatan.

### 4.2.1 Keputih

Curah hujan di wilayah keputih akan dimodelkan dengan variabel-variabel luaran GCM dengan teknik *statistical downscaling* menggunakan Regresi Gamma. Apabila model yang dihasilkan kurang baik, akan ditambahkan variabel prediktor seperti dummy dan lag dari curah hujan keputih. Penambahan variabel dummy diharapkan dapat memperbaiki hasil prediksi. Variabel dummy dikategorikan 1 sebagai musim hujan yaitu pada bulan November sampai April dan dikategorikan 0 sebagai musim kemarau yaitu pada bulan Mei sampai Oktober. Penambahan variabel lag diharapkan dapat meningkatkan hasil prediksi karena seperti yang sudah dibahas sebelumnya bahwa lag dari curah hujan keputih memiliki nilai koefisien korelasi yang lebih tinggi

dibandingkan variabel-variabel prediktor luaran GCM. Pemodelan curah hujan akan dilakukan pada grid 1x1 dan grid 3x3.

#### 4.2.1.1 Grid 1x1

Variabel yang digunakan pada pemodelan grid 1x1 adalah variabel pada lokasi 5. Digunakan data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE. Berikut ini merupakan pemodelan variabel curah hujan dengan variabel masing-masing variabel luaran GCM yaitu presipitasi, kelembapan, suhu, dan kecepatan angin dengan regresi gamma menggunakan *link inverse*, *identity*, dan "*sqrt*".

**Tabel 4.3** Pemodelan Curah Hujan Keputih dengan masing-masing variabel

Variabel	Link	Intercept		Prediktor	
		Estimate	P-value	Estimate	P-value
Presipitasi	<i>Inverse</i>	0.20	0.00	-147.31	0.00
	<i>Identity</i>	3.23	0.00	22230	0.00
	"sqrt"	1.86	0.00	4201	0.00
Kelembapan	<i>Inverse</i>	0.08	0.05	4.44	0.02
	<i>Identity</i>	7.14	0.00	-70.22	0.01
	"sqrt"	2.75	0.00	-17.24	0.01
Suhu	<i>Inverse</i>	-18.48	0.00	0.06	0.00
	<i>Identity</i>	286.18	0.00	-0.93	0.00
	"sqrt"	75.59	0.00	-0.24	0.00
Kecepatan Angin	<i>Inverse</i>	0.28	0.00	-0.02	0.00
	<i>Identity</i>	3.16	0.00	0.52	0.00
	"sqrt"	1.80	0.00	0.12	0.00

$P\text{-value} < \text{nilai } \alpha$  (5%) berarti terdapat pengaruh yang signifikan dari variabel prediktor terhadap curah hujan. Sebaliknya jika  $P\text{-value} > \text{nilai } \alpha$  (5%) berarti tidak terdapat pengaruh yang signifikan dari variabel prediktor terhadap curah hujan. Berdasarkan pemodelan curah hujan keputih dengan masing-masing variabel presipitasi, kelembapan, suhu, dan kecepatan angin memberikan pengaruh yang signifikan terhadap variabel curah hujan di wilayah Keputih. Hal ini dapat diketahui  $P\text{-value}$  yang kurang dari nilai  $\alpha$  (5%) pada pemodelan curah hujan

keputih dengan masing-masing variabel prediktor baik menggunakan link *inverse*, *identity*, maupun “sqrt”. Selanjutnya dilakukan pemodelan curah hujan di wilayah Keputih dengan semua variabel prediktor pada lokasi 5. Pemodelan semua variabel luaran GCM dengan curah hujan di wilayah Keputih menggunakan link *Inverse* dan link “sqrt” menghasilkan *intercept* dan dua variabel prediktor yang memberikan pengaruh signifikan terhadap curah hujan yaitu variabel presipitasi dan suhu. Hal ini dapat dilihat dari *P-value* < nilai *alpha* (5%), sedangkan variabel kelembapan dan kecepatan angin memiliki *P-value* yang lebih dari nilai *alpha* (5%) sehingga variabel kelembapan dan kecepatan angin tidak memberikan pengaruh yang signifikan terhadap curah hujan di wilayah Keputih. Namun berbeda halnya dengan menggunakan link *identity*, karena hanya menghasilkan satu variabel yang signifikan yaitu variabel presipitasi sedangkan ketiga variabel lainnya dan juga *intercept* tidak memberikan pengaruh yang signifikan terhadap curah hujan di wilayah Keputih. Dalam hal ini link *inverse* dan link “sqrt” memberikan hasil pemodelan yang lebih baik.

**Tabel 4.4** Pemodelan Curah Hujan Keputih dengan Empat Variabel Klimatologi

	<i>Inverse</i>		<i>Identity</i>		“sqrt”	
	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>
<i>Intercept</i>	-15.98	0.00	147.4	0.08	54.94	0.01
<b>x15</b>	-163.3	0.00	21224	0.00	3766	0.00
<b>x25</b>	-1.98	0.31	33.29	0.37	12.44	0.19
<b>x35</b>	0.05	0.00	-0.48	0.09	-0.18	0.01
<b>x45</b>	-0.0002	0.93	-0.12	0.39	-0.01	0.65

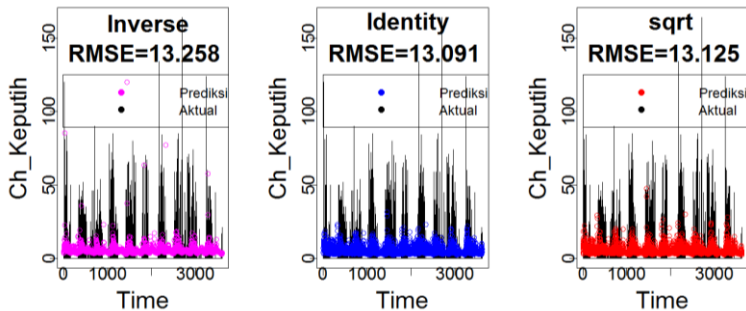
Selanjutnya perlu dilakukan analisa VIF (*Variance Inflation Factor*) untuk mendeteksi adanya multikolinearitas karena salah satu syarat melakukan pemodelan menggunakan GLM yaitu tidak adanya multikolinearitas. Berikut ini disajikan nilai VIF masing-masing variabel.

Berdasarkan pemodelan menggunakan link *Inverse*, *Identity*, maupun “sqrt” memberikan nilai VIF yang kurang dari 10.

**Tabel 4.5** Nilai VIF Model dengan Empat Variabel Klimatologi

Link	VIF			
	x15	x25	x35	x45
<i>Inverse</i>	2.06	1.34	1.34	2.14
<i>Identity</i>	1.01	1.51	1.62	1.28
“sqrt”	1.04	1.46	1.64	1.33

Sehingga dapat disimpulkan bahwa baik variabel presipitasi, kelembapan, suhu, maupun kecepatan angin tidak saling berhubungan atau tidak terdapat multikolinieritas pada semua model dengan penggunaan link yang berbeda. Berikut nilai RMSE dan hasil prediksi dari ketiga model dengan penggunaan link yang berbeda.

**Gambar 4.8** Prediksi Curah Hujan Keputih dengan Empat Variabel Klimatologi

Nilai RMSE yang dihasilkan oleh ketiga model tidak berbeda signifikan. Namun RMSE terkecil dihasilkan oleh model dengan link *identity* meskipun pemodelan dengan menggunakan link *identity* hanya menghasilkan satu variabel yang signifikan yaitu presipitasi, variabel lain dan juga *intercept* tidak signifikan. Nilai RMSE yang dihasilkan dari pemodelan menggunakan link *inverse* dan link “sqrt” hanya selisih 0.133 dengan pemodelan menggunakan link “sqrt” yang lebih kecil. Seperti terlihat pada gambar prediksi bahwa curah hujan pada musim kemarau yaitu pada bulan Mei sampai dengan Oktober yang seharusnya bernilai nol tetapi tidak bernilai nol sehingga perlu ditambahkan variabel dummy untuk menghasilkan prediksi bernilai nol pada musim

kemarau. Variabel dummy bernilai nol pada musim kemarau yaitu pada pada bulan Mei sampai dengan Oktober dan bernilai 1 ketika musim hujan yaitu pada pada bulan November sampai dengan April. Selanjutnya dilakukan pemodelan curah hujan di wilayah Keputih dengan variabel dummy dan semua variabel prediktor. Selain itu juga dilakukan pemodelan curah hujan di wilayah Keputih dengan variabel dummy dan dua variabel prediktor yang signifikan yaitu suhu dan presipitasi, diperoleh hasil sebagai berikut.

**Tabel 4.6** Pemodelan dengan Dummy, Suhu, dan Presipitasi

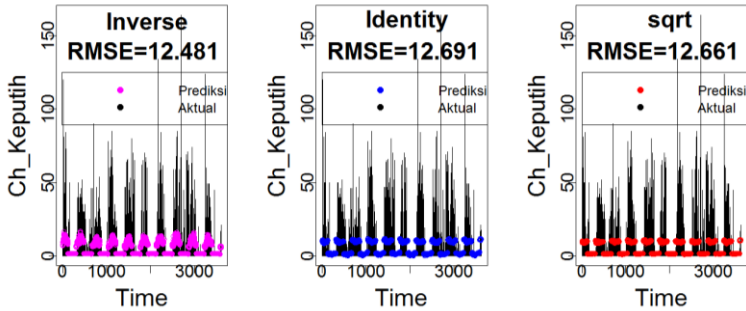
	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
<i>Intercept</i>	-7.10	0.02	-176.15	0.00	-14.90	0.44
dummy	-0.52	0.00	8.68	0.00	1.92	0.00
x15	-8.02	0.89	-963.21	0.00	-233.68	0.62
x35	0.03	0.01	0.59	0.00	0.05	0.40

**Tabel 4.7** Pemodelan dengan Semua Variabel dan Dummy

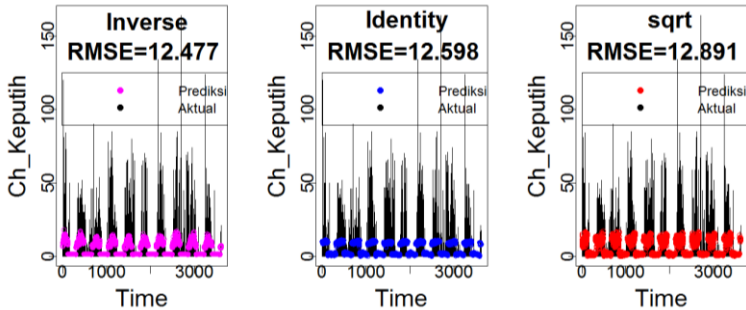
	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
<i>Intercept</i>	-8.74	0.03	190.5	0.00	26.31	0.24
dummy	-0.52	0.00	8.15	0.00	2.29	0.00
x15	-12.80	0.84	-1095	0.00	-657.4	0.00
x25	-1.72	0.49	-21.44	0.08	-4.73	0.67
x35	0.03	0.02	-0.62	0.00	-0.08	0.28
x45	0.001	0.85	-0.29	0.00	-0.19	0.00

Pemodelan curah hujan dengan dummy dan dua variabel menggunakan link *inverse* memberikan hasil semua variabel signifikan kecuali variabel presipitasi. Apabila menggunakan link *identity* menghasilkan semua variabel signifikan. Namun pemodelan dengan menggunakan link "sqrt" tidak ada satu pun variabel prediktor yang signifikan kecuali hanya dummy dan *intercept*. Pemodelan curah hujan dengan dummy dan semua variabel menggunakan link *inverse* hanya menghasilkan variabel suhu yang berpengaruh signifikan terhadap curah hujan selain daripada *intercept* dan variabel dummy. *Intercept*, variabel prediktor kelembapan dan suhu tidak memberikan pengaruh yang

signifikan terhadap curah hujan di wilayah Keputih apabila pemodelan dilakukan dengan menggunakan link “sqrt”. Dalam hal ini pemodelan curah hujan menggunakan link *identity* memberikan hasil yang lebih baik karena menghasilkan lebih banyak variabel yang signifikan. Berikut ini nilai RMSE yang dihasilkan oleh ketiga model dengan penggunaan link yang berbeda.



**Gambar 4.9** Hasil Prediksi dengan Variabel Dummy, Suhu, dan Presipitasi

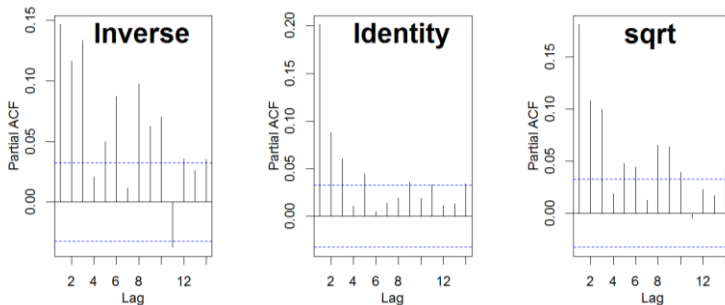


**Gambar 4.10** Hasil Prediksi dengan Empat Prediktor dan Dummy

Hasil prediksi pemodelan dengan penambahan variabel dummy terlihat memberikan dugaan curah hujan yang lebih baik terlihat pada bulan-bulan musim kemarau yaitu dari bulan Mei sampai dengan Oktober bernilai nol. Hasil prediksi pemodelan dengan penambahan variabel dummy dan dua variabel prediktor menghasilkan nilai RMSE terendah dari pemodelan menggunakan link *inverse*. Pemodelan dengan link *identity* menghasilkan nilai

RMSE yang paling besar, namun lebih banyak menghasilkan variabel predictor yang signifikan. Hal ini seperti yang dibahas sebelumnya bahwa pemodelan dengan variabel dummy, suhu, dan presipitasi menggunakan link *identity* menghasilkan semua variabel berpengaruh signifikan terhadap curah hujan di wilayah Keputih. Jumlah variabel yang signifikan lebih banyak daripada pemodelan dengan menggunakan link *inverse* maupun “sqrt”.

Nilai RMSE dari hasil prediksi pemodelan dengan penambahan variabel dummy dan semua variabel predictor menggunakan link *inverse* merupakan yang paling rendah dibandingkan dengan menggunakan link *identity* maupun “sqrt”. Pemodelan dengan menggunakan link *inverse* dan *identity* menghasilkan nilai RMSE yang lebih rendah pada pemodelan dengan semua variabel predictor dibandingkan pada pemodelan dengan dua variabel predictor yaitu sebesar 12.477 dan 12.598. Sebaliknya pada pemodelan curah hujan dengan semua variabel predictor menggunakan link “sqrt” menghasilkan nilai RMSE yang lebih besar dibandingkan pada pemodelan dengan dua variabel predictor yaitu sebesar 12.891. Dalam hal ini pemodelan curah hujan dengan semua variabel predictor memberikan dugaan yang lebih baik jika menggunakan link *inverse* dan link *identity*.



**Gambar 4.11** PACF Model dengan Dummy dan Semua Variabel

Namun penambahan variabel dummy kurang bisa untuk menaikkan hasil prediksi, seperti yang telah dibahas sebelumnya



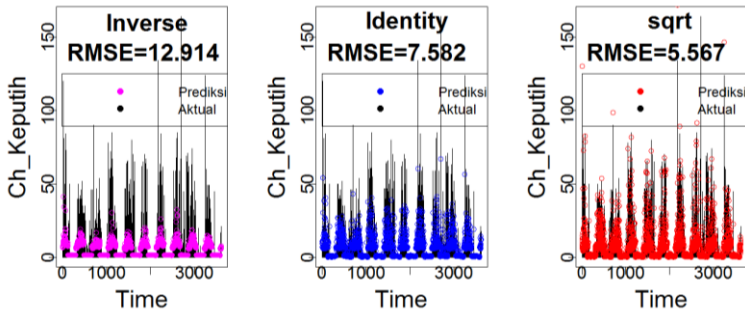
bahwa nilai koefisien korelasi antara curah hujan di wilayah Keputih dengan lag nya lebih tinggi daripada korelasi curah hujan dengan variabel prediktor. Maka dari itu, untuk mengetahui lag berapa yang harus dimasukkan ke dalam model agar dapat meningkatkan hasil prediksi dapat diketahui dengan melihat PACF dari ketiga model. Berdasarkan PACF pemodelan curah hujan di wilayah Keputih dengan semua variabel dan penambahan variabel dummy dapat diketahui bahwa lag 1, lag 2, lag 3 signifikan pada model dengan menggunakan link *inverse*, *identity*, maupun “sqrt”. Berikut ini dilakukan pemodelan curah hujan di wilayah Keputih dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Keputih untuk meningkatkan hasil prediksi.

**Tabel 4.8** Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1

	<i>Inverse</i>		<i>Identity</i>		“sqrt”	
	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>
<i>Intercept</i>	-6.86	0.09	44.20	0.00	68.16	0.00
dummy	-0.51	0.00	6.01	0.00	1.85	0.00
x15	16.40	0.80	-1523	0.00	-599.5	0.00
x25	-1.33	0.58	-31.81	0.00	-26.42	0.00
x35	0.02	0.06	-0.14	0.00	-0.22	0.00
x45	-0.0007	0.95	-0.13	0.00	-0.17	0.00
lag1	-0.0003	0.00	0.40	0.00	0.08	0.00

Pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse* hanya menghasilkan dua variabel yang berpengaruh signifikan terhadap curah hujan di wilayah Keputih yaitu variabel dummy dan lag 1. Sedangkan semua variabel prediktor seperti presipitasi, kelembapan, suhu, dan kecepatan angin tidak memiliki pengaruh yang signifikan terhadap curah hujan di wilayah Keputih. Hal ini dapat diketahui dari *P-value* yang kurang dari nilai *alpha* (5%). Lain halnya pada pemodelan menggunakan link *identity* maupun link “sqrt” memberikan hasil bahwa semua variabel berpengaruh signifikan terhadap curah hujan di wilayah Keputih baik variabel prediktor, variabel dummy,

maupun variabel lag 1. Hal ini dapat diketahui dari  $P$ -value yang lebih dari nilai  $\alpha$  (5%). Berikut ini nilai RMSE yang dihasilkan oleh ketiga model dengan penggunaan link yang berbeda.



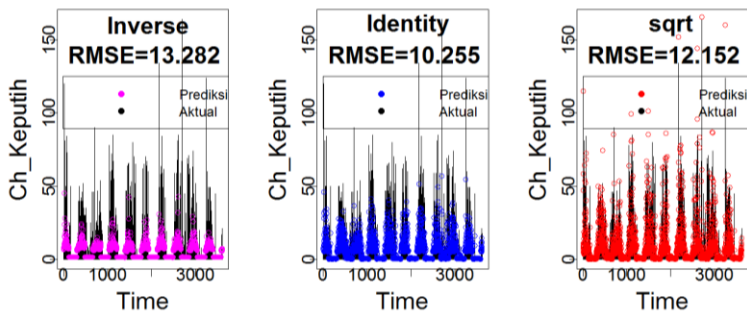
**Gambar 4.12** Hasil Prediksi Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1

Nilai RMSE dari pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang paling kecil yaitu sebesar 5.567. Hal ini berarti hasil prediksi data curah hujan di wilayah Keputih menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Keputih dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu sebesar 12.914. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Keputih menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Keputih dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Keputih dengan penambahan lag 1 dan lag 2 memberikan hasil yang lebih besar daripada nilai RMSE yang dihasilkan dari

pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 10.225. Sedangkan pemodelan menggunakan link invers menghasilkan nilai RMSE yang paling besar yaitu 13.282. Dalam hal ini pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{1,t} = (68.16 + 1.85dummy_t - 599.5X_{1,t} - 26.42X_{2,t} - 0.22X_{3,t} - 0.17X_{4,t} + 0.08Y_{1,t-1})^2$$

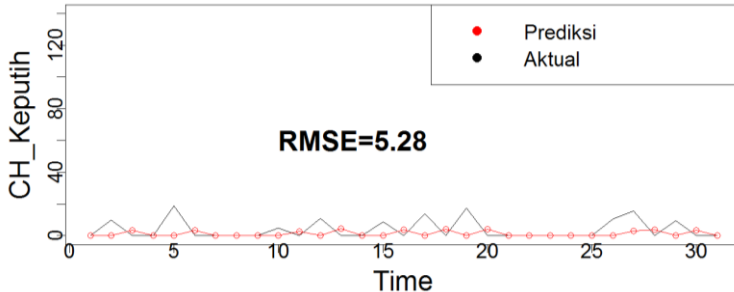


**Gambar 4.13** Hasil Prediksi Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Keputih karena memiliki nilai RMSE terendah yaitu 5.567. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Keputih dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.

Perhitungan RMSE data testing berdasarkan pada lag 1 dari rata-rata curah hujan bulan Desember periode 2009 sampai 2018. Hasil prediksi pemodelan curah hujan di wilayah Keputih dengan semua variabel predictor dan penambahan variabel dummy

dan lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 5.28 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 5.567. tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing.



**Gambar 4.14** Hasil Prediksi Curah Hujan Menggunakan Data Testing

Hal ini berarti bahwa pemodelan curah hujan di wilayah Keputih dengan semua variabel predictor dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.

#### 4.2.1.2 Grid 3x3

Variabel yang digunakan pada pemodelan grid 3x3 adalah variabel GCM pada 9 lokasi. Data yang digunakan adalah data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE. Pemodelan curah hujan di wilayah Keputih dengan variabel prediktor luaran GCM pada grid 3x3 dilakukan dengan teknik reduksi dimensi menggunakan *Principal Components Analysis* atau yang biasa disebut dengan PCA yang selanjutnya dapat dimodelkan dengan regresi Gamma. Berikut ini hasil analisis menggunakan PCA dari variabel predictor luaran GCM.

**Tabel 4.9** Keragaman Kumulatif Variabel GCM

<b>Variabel</b>	<b>PC1</b>	<b>PC2</b>	<b>PC3</b>	<b>PC4</b>	<b>PC5</b>	<b>PC6</b>
Presipitasi	0.65	0.83	0.90	0.95	0.97	0.99
Kelembapan	0.66	0.85	0.95	0.98	0.99	1.00
Suhu	0.78	0.94	0.97	0.99	0.99	1.00
Kecepatan Angin	0.80	0.90	0.95	0.98	0.99	0.99

Berdasarkan keragaman kumulatif keempat variabel prediktor luaran GCM diperoleh informasi bahwa pada variabel presipitasi, jumlah komponen utama yang terbentuk dengan keragaman yang dapat dijelaskan sebesar >80% pada grid 3x3 adalah dua komponen utama dengan keragaman kumulatif 83%. Sedangkan pada variabel prediktor kelembapan, suhu, dan kecepatan angin menghasilkan nilai keragaman kumulatif yang lebih tinggi dari variabel presipitasi yaitu masing-masing sebesar 85%, 94%, dan 90% dengan jumlah komponen utama yang terbentuk sebanyak dua. Berikut ini merupakan pemodelan variabel curah hujan dengan delapan komponen utama yang terbentuk dari masing-masing variabel luaran GCM yaitu presipitasi, kelembapan, suhu, dan kecepatan angin dengan regresi gamma menggunakan *link inverse*, *identity*, dan "*sqrt*".

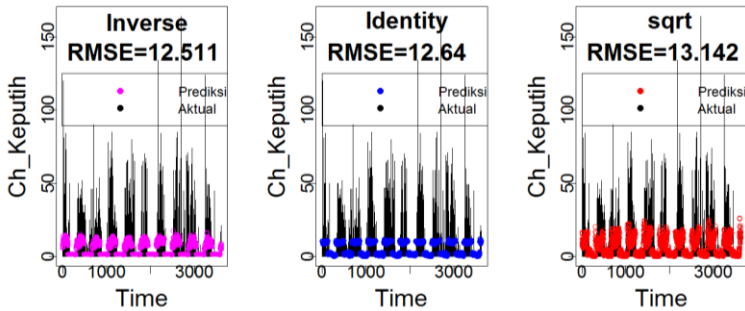
**Tabel 4.10** Pemodelan Curah Hujan pada Grid 3x3

	<i>Inverse</i>		<i>Identity</i>		<i>"sqrt"</i>	
	<b>Estimate</b>	<b>P-value</b>	<b>Estimate</b>	<b>P-value</b>	<b>Estimate</b>	<b>P-value</b>
Intercept	0.63	0.00	1.30	0.00	1.03	0.00
dummy	-0.51	0.00	8.67	0.00	2.48	0.00
pc11	0.92	0.98	-770.7	0.00	-436.6	0.00
pc12	-27.60	0.56	148.9	0.02	378.4	0.14
pc21	-0.59	0.76	23.84	0.10	25.42	0.00
pc22	3.51	0.38	-42.32	0.00	-2.99	0.82
pc31	-0.01	0.14	0.11	0.00	0.08	0.00
pc32	-0.004	0.67	-0.23	0.00	-0.23	0.00
pc41	-0.0002	0.91	-0.07	0.00	-0.06	0.00
pc42	0.0016	0.84	-0.03	0.00	-0.06	0.01

Delapan variabel komponen utama (PC) yang terbentuk dari PCA dan variabel dummy dimodelkan dengan curah hujan di wilayah Keputih menggunakan regresi Gamma dengan penggunaan link *inverse* dapat diketahui bahwa tidak ada satu pun

variabel PC yang berpengaruh signifikan terhadap curah hujan di wilayah Keputih kecuali hanya variabel dummy dan intercept.

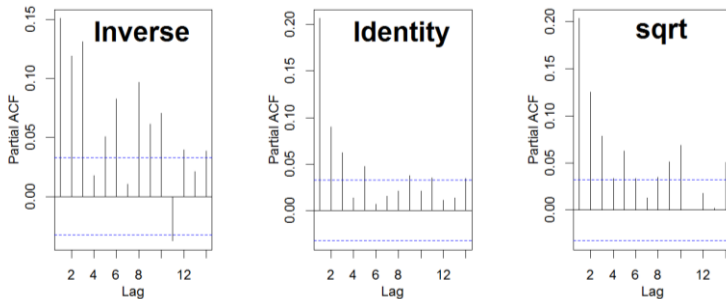
Regresi Gamma menggunakan link *identity* memberikan hasil bahwa semua variabel signifikan kecuali pc12 yaitu PC 2 dari variabel presipitasi. Sedangkan regresi Gamma menggunakan link “sqrt” lebih sedikit menghasilkan variabel signifikan jika dibandingkan dengan menggunakan link *identity*. Semua variabel signifikan berpengaruh terhadap curah hujan di wilayah Keputih pada regresi Gamma menggunakan link “sqrt” kecuali pc12 dan pc22. Pc 22 merupakan PC 2 dari variabel kelembapan. Berikut ini nilai RMSE yang dihasilkan dari ketiga model dengan penggunaan link yang berbeda.



**Gambar 4.15** Prediksi Curah Hujan Keputih dengan 8 Variabel PC dan Variabel Dummy pada Grid 3x3

Nilai RMSE yang dihasilkan dari pemodelan menggunakan link *inverse* merupakan yang paling kecil namun tidak ada variabel PC yang signifikan terhadap curah hujan di wilayah Keputih. Pemodelan menggunakan link *identity* dan menghasilkan nilai RMSE sebesar 12.64. Sedangkan pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang terbesar yaitu 13.142. Pemodelan curah hujan di wilayah Keputih dengan delapan PC yang terbentuk dan variabel dummy menggunakan link *identity* memberikan hasil yang lebih baik dibandingkan dengan menggunakan link lain. Namun penambahan variabel dummy kurang bisa untuk menaikkan hasil prediksi, seperti yang telah

dibahas sebelumnya bahwa nilai koefisien korelasi antara curah hujan di wilayah Keputih dengan lag nya lebih tinggi daripada korelasi curah hujan dengan variabel prediktor. Maka dari itu, untuk mengetahui lag berapa yang harus dimasukkan ke dalam model agar dapat meningkatkan hasil prediksi dapat diketahui dengan melihat PACF dari ketiga model.



**Gambar 4.16** PACF Model pada Grid 3x3

Berdasarkan PACF pemodelan curah hujan di wilayah Keputih dengan delapan komponen utama dan variabel dummy dapat diketahui bahwa lag 1, lag 2, lag 3 signifikan pada model dengan menggunakan link *inverse*, *identity*, maupun “sqrt”. Berikut ini dilakukan pemodelan curah hujan di wilayah Keputih dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Keputih untuk meningkatkan hasil prediksi.

Pemodelan curah hujan di wilayah Keputih dengan delapan komponen utama yang terbentuk dan variabel dummy serta penambahan lag 1 menggunakan link *inverse* diperoleh hasil variabel yang signifikan hanyalah variabel dummy dan lag 1 dari curah hujan. Sedangkan pada link *identity* dihasilkan semua variabel signifikan kecuali pc21 yaitu PC 1 dari variabel kelembapan. Regresi Gamma menggunakan link “sqrt” diperoleh hasil semua variabel signifikan kecuali pc11, pc12, pc31, dan pc42. Pc11 dan pc12 merupakan PC 1 dan PC 2 dari variabel Presipitasi. Pc 31 merupakan PC 1 dari variabel suhu dan pc42 merupakan PC 2 dari variabel kecepatan angin.

**Tabel 4.11** Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3

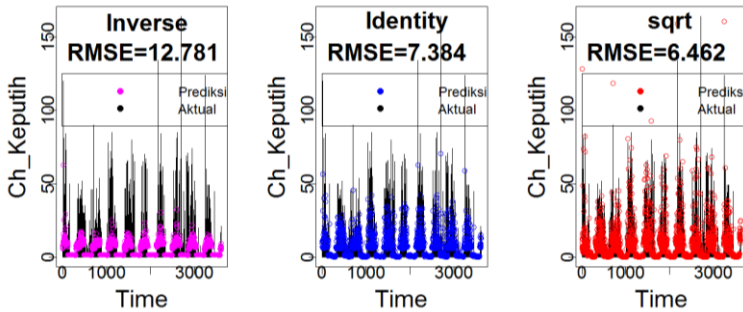
	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
Intercept	0.63	0.00	0.81	0.00	0.79	0.00
dummy	-0.50	0.00	6.04	0.00	2.13	0.00
pc11	15.25	0.64	-577.9	0.00	61.78	0.64
pc12	-21.84	0.65	407.9	0.00	-339.5	0.22
pc21	-0.08	0.97	-12.07	0.20	7.87	0.22
pc22	3.74	0.34	-24.53	0.00	20.44	0.02
pc31	-0.01	0.25	0.04	0.00	-0.01	0.40
pc32	-0.003	0.77	-0.16	0.00	-0.27	0.00
pc41	-0.0005	0.81	-0.06	0.00	-0.06	0.00
pc42	-0.0007	0.93	-0.02	0.00	-0.02	0.39
lag 1	-0.0007	0.00	0.42	0.00	0.08	0.00

Nilai RMSE dari pemodelan curah hujan dengan delapan komponen utama yang terbentuk dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan "sqrt" memberikan hasil yang berbeda signifikan. Nilai RMSE pemodelan regresi Gamma dengan menggunakan link "sqrt" merupakan yang paling rendah yaitu sebesar 6.462. Hal ini berarti hasil prediksi data curah hujan di wilayah Keputih menggunakan link "sqrt" merupakan yang paling mendekati dengan data asli curah hujan di wilayah Keputih dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Pemodelan regresi Gamma dengan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 12.781. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Keputih menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Keputih dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Keputih pada grid 3x3 dengan penambahan lag 1 dan lag 2 menghasilkan nilai RMSE yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE



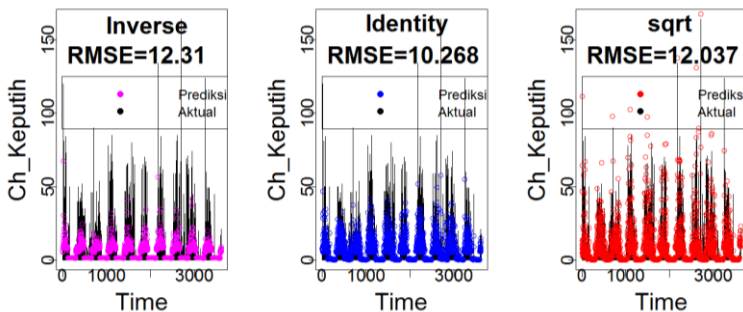
yang paling kecil yaitu sebesar 10.268. Sedangkan pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 12.31.



**Gambar 4.17** Prediksi Curah Hujan Keputih dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 pada Grid 3x3

Pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

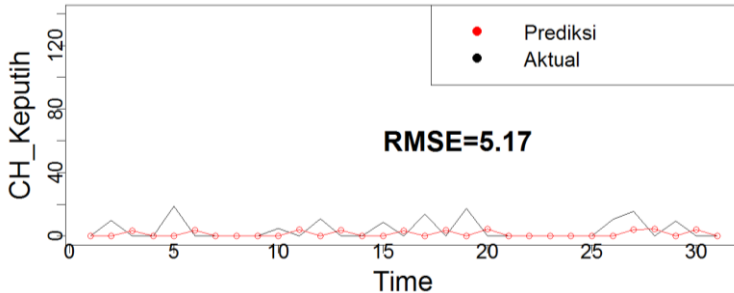
$$\hat{Y}_{1,t} = (0.79 + 2.13dummy_t + 61.78PC1X_{1,t} - 339.5PC2X_{1,t} + 7.87PC1X_{2,t} + 20.44PC2X_{2,t} - 0.01PC1X_{3,t} - 0.27PC2X_{3,t} - 0.06PC1X_{4,t} - 0.02PC2X_{4,t} + 0.08Y_{1,t-1})^2$$



**Gambar 4.18** Prediksi Curah Hujan Keputih dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 pada Grid 3x3

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Keputih karena memiliki nilai RMSE terendah yaitu

6.462. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Keputih dengan delapan komponen utama yang terbentuk pada grid 3x3 dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.



**Gambar 4.19** Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3

Hasil prediksi pemodelan curah hujan di wilayah Keputih dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 5.17 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 6.462. tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Keputih dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing. Hasil prediksi dari pemodelan curah hujan di wilayah keputih pada grid 1x1 lebih baik daripada Hasil prediksi dari pemodelan curah hujan di wilayah keputih pada grid 3x3 karena menghasilkan nilai RMSE yang lebih kecil.

## 4.2.2 Kedung Cowek

Curah hujan di wilayah Kedung Cowek akan dimodelkan dengan variabel-variabel luaran GCM dengan teknik *statistical downscaling* menggunakan Regresi Gamma. Apabila model yang dihasilkan kurang baik, akan ditambahkan variabel prediktor seperti dummy dan lag dari curah hujan keputih. Penambahan variabel dummy diharapkan dapat memperbaiki hasil prediksi. Variabel dummy dikategorikan 1 sebagai musim hujan yaitu pada bulan November sampai April dan dikategorikan 0 sebagai musim kemarau yaitu pada bulan Mei sampai Oktober. Penambahan variabel lag diharapkan dapat meningkatkan hasil prediksi karena seperti yang sudah dibahas sebelumnya bahwa lag dari curah hujan keputih memiliki nilai koefisien korelasi yang lebih tinggi dibandingkan variabel-variabel prediktor luaran GCM. Pemodelan curah hujan akan dilakukan pada grid 1x1 dan grid 3x3.

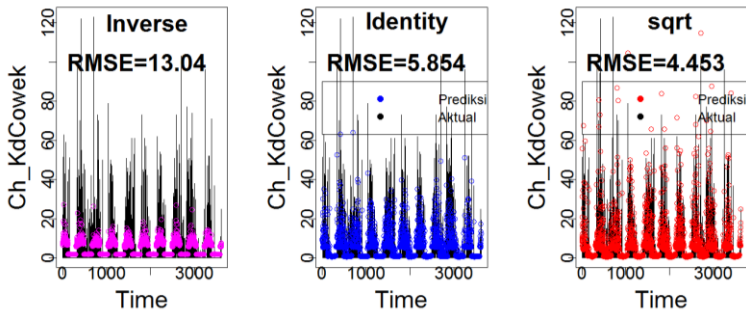
### 4.2.2.1 Grid 1x1

Variabel yang digunakan pada pemodelan grid 1x1 adalah variabel pada lokasi 5. Digunakan data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Berikut ini dilakukan pemodelan curah hujan di wilayah Kedung Cowek dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Kedung Cowek untuk meningkatkan hasil prediksi.

**Tabel 4.12** Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	<i>Estimate</i>	<i>P-value</i>	<i>Estimate</i>	<i>P-value</i>	<i>Estimate</i>	<i>P-value</i>
<i>Intercept</i>	-4.83	0.15	-15.49	0.67	-21.72	0.24
<i>dummy</i>	-0.40	0.00	4.78	0.00	1.33	0.00
<b>x15</b>	-43.29	0.39	874.19	0.41	874.98	0.048
<b>x25</b>	-1.07	0.60	-10.59	0.64	3.82	0.66
<b>x35</b>	0.02	0.11	0.06	0.64	0.08	0.22
<b>x45</b>	-0.0005	0.91	-0.08	0.14	-0.04	0.14
<b>lag1</b>	-0.0010	0.00	0.47	0.00	0.10	0.00

Pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” hanya menghasilkan dua variabel yang berpengaruh signifikan terhadap curah hujan di wilayah Kedung Cowek yaitu variabel dummy dan lag 1. Sedangkan semua variabel prediktor seperti presipitasi, kelembapan, suhu, dan kecepatan angin tidak memiliki pengaruh yang signifikan terhadap curah hujan di wilayah Kedung Cowek. Hal ini dapat diketahui dari *P-value* yang kurang dari nilai *alpha* (5%). Lain hal nya pada pemodelan menggunakan link “sqrt” memberikan hasil bahwa variabel dummy, lag 1, dan presipitasi berpengaruh signifikan terhadap curah hujan di wilayah Kedung Cowek. Hal ini dapat diketahui dari *P-value* yang lebih dari nilai *alpha* (5%). Berikut ini nilai RMSE yang dihasilkan oleh ketiga model dengan penggunaan link yang berbeda.



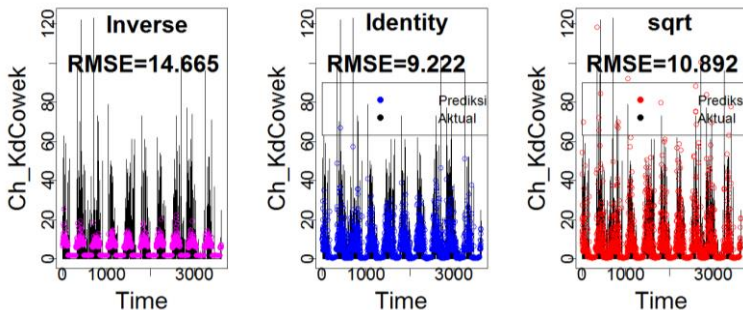
**Gambar 4.20** Hasil Prediksi Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1

Nilai RMSE dari pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang paling kecil yaitu sebesar 4.453. Hal ini berarti hasil prediksi data curah hujan di wilayah Kedung Cowek menggunakan link “sqrt” merupakan yang paling

mendekati dengan data asli curah hujan di wilayah Kedung Cowek dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*.

Pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu sebesar 13.04. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Kedung Cowek menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Kedung Cowek dibandingkan dengan hasil prediksi pemodelan menggunakan kedua link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek dengan penambahan lag 1 dan lag 2.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek dengan penambahan lag 1 dan lag 2 memberikan hasil yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 9.222. Sedangkan pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 14.665.

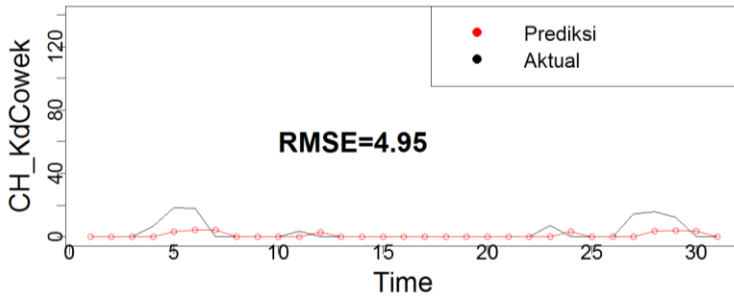


**Gambar 4.21** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2

Pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{2,t} = (-21.72 + 1.33dummy_t + 874.98X_{1,t} + 3.82X_{2,t} + 0.08X_{3,t} + 0.04X_{4,t} + 0.1Y_{1,t-1})^2$$

Pemodelan terbaik menggunakan “link “sqrt”” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Kedung Cowek karena memiliki nilai RMSE terendah yaitu 4.453. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.



**Gambar 4.22** Hasil Prediksi Curah Hujan Menggunakan Data Testing

Hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 4.95 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 4.453. Tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Kedung Cowek dengan semua variabel predictor dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.

#### 4.2.2.2 Grid 3x3

Variabel yang digunakan pada pemodelan grid 3x3 adalah variabel GCM pada 9 lokasi. Data yang digunakan adalah data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Pemodelan curah hujan di wilayah Kedung Cowek dengan variabel prediktor luaran GCM pada grid 3x3 dilakukan dengan teknik reduksi dimensi menggunakan *Principal Components Analysis* atau yang biasa disebut dengan PCA yang selanjutnya dapat dimodelkan dengan regresi Gamma. Berikut ini dilakukan pemodelan curah hujan di wilayah Kedung Cowek dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Kedung Cowek untuk meningkatkan hasil prediksi.

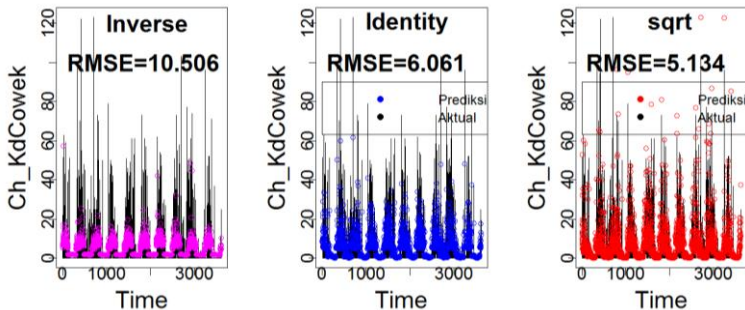
**Tabel 4.13** Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
Intercept	0.52	0.00	1.03	0.00	0.94	0.00
dummy	-0.36	0.00	4.39	0.00	1.52	0.00
pc11	-0.88	0.97	150.08	0.32	25.02	0.69
pc12	-46.27	0.23	123.31	0.74	332.05	0.18
pc21	-0.18	0.91	11.96	0.00	29.06	0.00
pc22	5.79	0.07	-14.40	0.00	5.76	0.10
pc31	-0.004	0.41	-0.02	0.00	0.01	0.43
pc32	-0.001	0.80	-0.26	0.00	-0.29	0.00
pc41	-0.001	0.49	0.02	0.00	-0.01	0.00
pc42	0.01	0.08	-0.18	0.00	-0.15	0.00
lag 1	-0.001	0.00	0.45	0.00	0.09	0.00

Pemodelan curah hujan di wilayah Kedung Cowek dengan delapan komponen utama yang terbentuk dan variabel dummy serta penambahan lag 1 menggunakan link *inverse* diperoleh hasil variabel yang signifikan hanyalah intercept, variabel dummy dan lag 1 dari curah hujan. Sedangkan pada link *identity* dihasilkan semua variabel signifikan kecuali pc11 dan pc12 yang merupakan PC 1 dan PC 2 dari variabel presipitasi. Regresi Gamma menggunakan link "sqrt" diperoleh hasil semua variabel signifikan

kecuali pc11, pc12, pc22, dan pc31. Pc 22 merupakan PC 2 dari variabel kelembapan dan pc 31 merupakan PC 1 dari variabel suhu.

Nilai RMSE dari pemodelan curah hujan dengan delapan komponen utama yang terbentuk dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Nilai RMSE pemodelan regresi Gamma dengan menggunakan link “sqrt” merupakan yang paling rendah yaitu sebesar 5.134. Hal ini berarti hasil prediksi data curah hujan di wilayah Kedung Cowek menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Kedung Cowek dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Sedangkan pemodelan regresi Gamma dengan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 10.506.



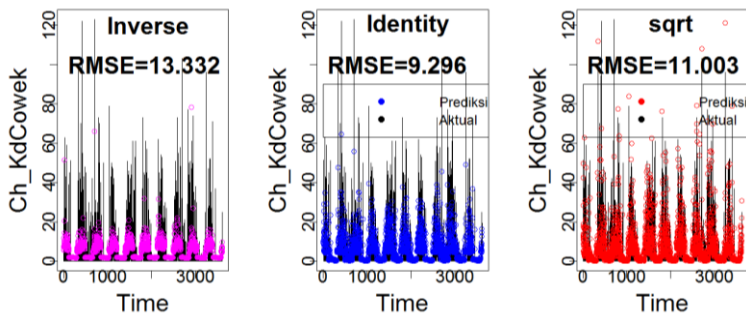
**Gambar 4.23** Prediksi Curah Hujan Kedung Cowek dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 pada Grid 3x3

Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Kedung Cowek menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Kedung Cowek jika dibandingkan dengan hasil prediksi pemodelan menggunakan kedua link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek dengan penambahan lag 1 dan lag 2.



Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek pada grid 3x3 dengan penambahan lag 1 dan lag 2 menghasilkan nilai RMSE yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 9.296. Sedangkan pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 13.332. Dalam hal ini pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

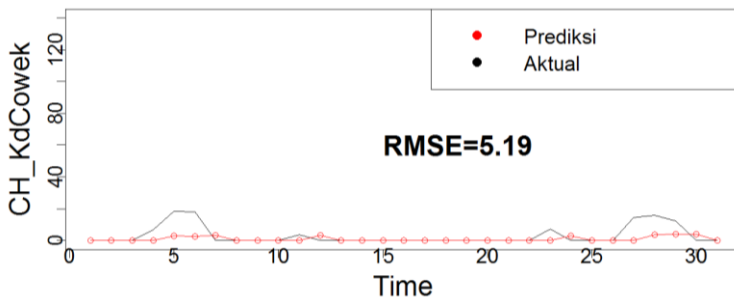
$$\hat{Y}_{2,t} = (0.94 + 1.52dummy_t + 25.02PC1X_{1,t} + 332.05PC2X_{1,t} + 29.06PC1X_{2,t} + 5.76PC2X_{2,t} + 0.01PC1X_{3,t} - 0.29PC2X_{3,t} - 0.01PC1X_{4,t} - 0.15PC2X_{4,t} + 0.09Y_{2,t-1})^2$$



**Gambar 4.24** Prediksi Curah Hujan Kedung Cowek dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 pada Grid 3x3

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Kedung Cowek karena memiliki nilai RMSE terendah yaitu 5.134. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek dengan delapan komponen utama yang terbentuk pada grid 3x3 dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.

Hasil prediksi pemodelan curah hujan di wilayah Kedung Cowek dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 5.19 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 5.134. Tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing.



**Gambar 4.25** Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3

Hal ini berarti bahwa pemodelan curah hujan di wilayah Kedung Cowek dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing. Hasil prediksi dari pemodelan curah hujan di wilayah Kedung Cowek pada grid 1x1 lebih baik daripada Hasil prediksi dari pemodelan curah hujan di wilayah Kedung Cowek pada grid 3x3 karena menghasilkan nilai RMSE yang lebih kecil yaitu selisih 0.431.

### 4.2.3 Gubeng

Curah hujan di wilayah Gubeng akan dimodelkan dengan variabel-variabel luaran GCM dengan teknik *statistical downscaling* menggunakan Regresi Gamma. Apabila model yang dihasilkan kurang baik, akan ditambahkan variabel prediktor

seperti dummy dan lag dari curah hujan keputih. Penambahan variabel dummy diharapkan dapat memperbaiki hasil prediksi. Variabel dummy dikategorikan 1 sebagai musim hujan yaitu pada bulan November sampai April dan dikategorikan 0 sebagai musim kemarau yaitu pada bulan Mei sampai Oktober. Penambahan variabel lag diharapkan dapat meningkatkan hasil prediksi karena seperti yang sudah dibahas sebelumnya bahwa lag dari curah hujan keputih memiliki nilai koefisien korelasi yang lebih tinggi dibandingkan variabel-variabel prediktor luaran GCM. Pemodelan curah hujan akan dilakukan pada grid 1x1 dan grid 3x3.

#### 4.2.3.1 Grid 1x1

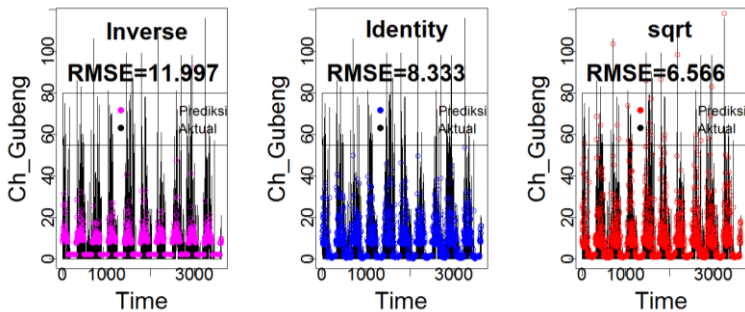
Variabel yang digunakan pada pemodelan grid 1x1 adalah variabel pada lokasi 5. Digunakan data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Berikut ini dilakukan pemodelan curah hujan di wilayah Gubeng dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Gubeng untuk meningkatkan hasil prediksi.

**Tabel 4.14** Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
<i>Intercept</i>	-2.46	0.32	-91.65	0.00	-47.96	0.01
<b>dummy</b>	-0.35	0.00	6.91	0.00	1.58	0.00
<b>x15</b>	-22.43	0.55	-156.78	0.81	952.26	0.04
<b>x25</b>	-0.69	0.65	-37.93	0.01	-15.04	0.045
<b>x35</b>	0.01	0.23	0.31	0.00	0.16	0.01
<b>x45</b>	-0.002	0.56	-0.19	0.00	-0.06	0.06
<b>lag1</b>	-0.001	0.00	0.38	0.00	0.07	0.00

Pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse* hanya menghasilkan dua variabel yang berpengaruh signifikan terhadap curah hujan di wilayah Gubeng yaitu variabel dummy dan lag 1. Sedangkan semua variabel prediktor seperti presipitasi, kelembapan, suhu, dan kecepatan angin tidak memiliki pengaruh yang signifikan terhadap curah

hujan di wilayah Gubeng. Hal ini dapat diketahui dari  $P$ -value yang kurang dari nilai  $\alpha$  (5%). Lain halnya pada pemodelan menggunakan link *identity* memberikan hasil bahwa semua variabel berpengaruh signifikan terhadap curah hujan di wilayah Gubeng kecuali variabel presipitasi. Kecepatan angin pada pemodelan menggunakan link “sqrt” tidak memberikan pengaruh yang signifikan terhadap curah hujan di wilayah Gubeng. Hal ini dapat diketahui dari  $P$ -value yang lebih dari nilai  $\alpha$  (5%). Sedangkan intercept, dummy, dan ketiga variabel prediktor lainnya berpengaruh signifikan terhadap curah hujan di wilayah Gubeng.

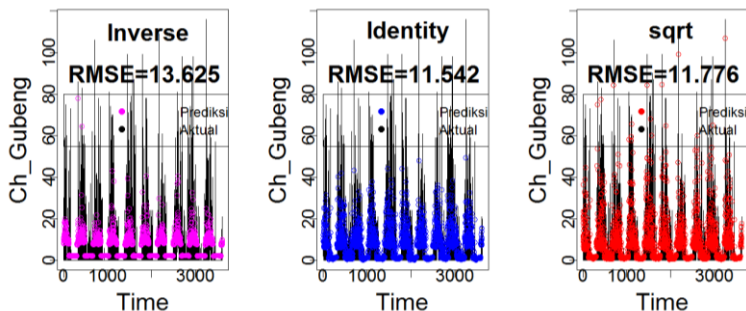


**Gambar 4.26** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1

Nilai RMSE dari pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang paling kecil yaitu sebesar 6.566. Hal ini berarti hasil prediksi data curah hujan di wilayah Gubeng menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Gubeng dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu sebesar 11.997. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah

Gubeng menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Gubeng dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Gubeng dengan penambahan lag 1 dan lag 2.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Gubeng dengan penambahan lag 1 dan lag 2 memberikan hasil yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja.



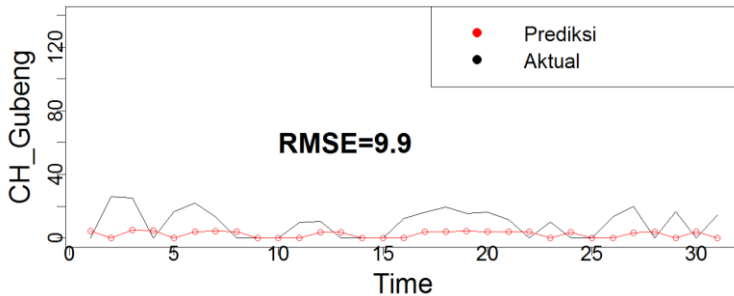
**Gambar 4.27** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2

Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu 11.542. Sedangkan pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 13.625. Pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{3,t} = (-47.96 + 1.58dummy_t + 952.26X_{1,t} - 15.04X_{2,t} + 0.16X_{3,t} - 0.06X_{4,t} + 0.07Y_{3,t-1})^2$$

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Gubeng karena memiliki nilai RMSE terendah yaitu

6.566. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Gubeng dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.



**Gambar 4.28** Hasil Prediksi Curah Hujan Menggunakan Data Testing

Hasil prediksi pemodelan curah hujan di wilayah Gubeng dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 9.9 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 6.566. Tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Gubeng dengan semua variabel predictor dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.

#### 4.2.3.2 Grid 3x3

Variabel yang digunakan pada pemodelan grid 3x3 adalah variabel GCM pada 9 lokasi. Data yang digunakan adalah data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Pemodelan curah hujan di wilayah Gubeng dengan

variabel prediktor luaran GCM pada grid 3x3 dilakukan dengan teknik reduksi dimensi menggunakan *Principal Components Analysis* atau yang biasa disebut dengan PCA yang selanjutnya dapat dimodelkan dengan regresi Gamma. Berikut ini dilakukan pemodelan curah hujan di wilayah Gubeng dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Gubeng untuk meningkatkan hasil prediksi.

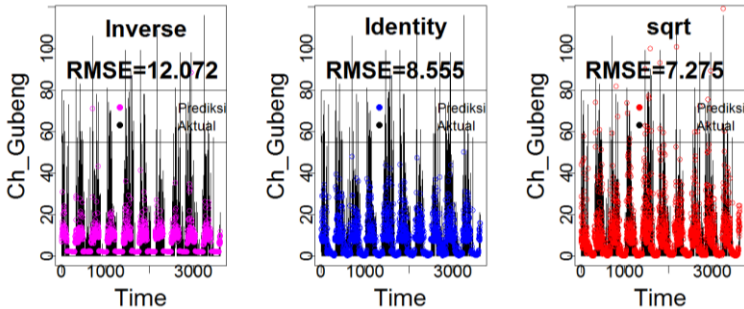
**Tabel 4.15** Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
Intercept	0.45	0.00	1.34	0.00	1.07	0.00
dummy	-0.32	0.00	6.55	0.00	1.94	0.00
pc11	-1.07	0.95	-33.90	0.81	232.86	0.36
pc12	-13.23	0.65	-52.38	0.58	-221.11	0.62
pc21	0.18	0.88	19.70	0.03	12.19	0.10
pc22	4.60	0.08	-10.94	0.00	11.92	0.38
pc31	-0.002	0.64	-0.06	0.00	-0.02	0.25
pc32	0.000	0.95	-0.45	0.00	-0.36	0.00
pc41	-0.001	0.37	0.01	0.00	-0.03	0.02
pc42	0.003	0.55	-0.18	0.00	-0.06	0.05
lag 1	-0.001	0.00	0.36	0.00	0.07	0.00

Pemodelan curah hujan di wilayah Gubeng dengan delapan komponen utama yang terbentuk dan variabel dummy serta penambahan lag 1 menggunakan link *inverse* diperoleh hasil variabel yang signifikan hanyalah intercept, variabel dummy dan lag 1 dari curah hujan. Sedangkan pada link *identity* dihasilkan semua variabel signifikan kecuali pc11 dan pc12 yang merupakan PC 1 dan PC 2 dari variabel presipitasi. Regresi Gamma menggunakan link "sqrt" diperoleh hasil variabel yang signifikan yaitu intercept, dummy, lag 1, pc32, dan pc 41. Pc 32 merupakan PC 2 dari variabel suhu dan pc41 merupakan PC 1 dari variabel kecepatan angin. Berikut nilai RMSE yang dihasilkan ketiga model.

Nilai RMSE dari pemodelan curah hujan dengan delapan komponen utama yang terbentuk dan penambahan variabel dummy

serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Nilai RMSE pemodelan regresi Gamma dengan menggunakan link “sqrt” merupakan yang paling rendah yaitu sebesar 7.257. Hal ini berarti hasil prediksi data curah hujan di wilayah Gubeng menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Gubeng dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*.



**Gambar 4.29** Prediksi Curah Hujan Gubeng dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1

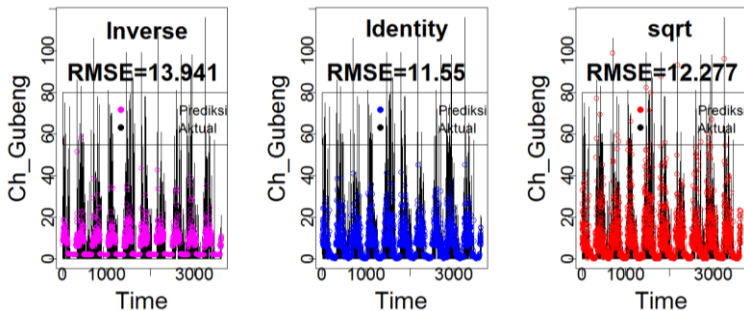
Pemodelan regresi Gamma dengan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 12.072. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Gubeng menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Gubeng dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Gubeng dengan penambahan lag 1 dan lag 2.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Gubeng pada grid 3x3 dengan penambahan lag 1 dan lag 2 menghasilkan nilai RMSE yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 11.55. Sedangkan pemodelan



menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 13.941. Pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{3,t} = (1.07 + 1.94dummy_t + 232.86PC1X_{1,t} - 221.11PC2X_{1,t} + 12.19PC1X_{2,t} + 11.92PC2X_{2,t} - 0.02PC1X_{3,t} - 0.36PC2X_{3,t} - 0.03PC1X_{4,t} - 0.06PC2X_{4,t} + 0.07Y_{3,t-1})^2$$

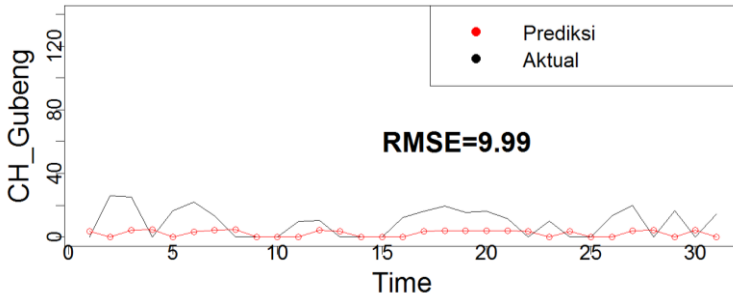


**Gambar 4.30** Prediksi Curah Hujan Gubeng dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2 pada Grid 3x3

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Gubeng karena memiliki nilai RMSE terendah yaitu 6.462. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Gubeng dengan delapan komponen utama yang terbentuk pada grid 3x3 dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.

Hasil prediksi pemodelan curah hujan di wilayah Gubeng dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 9.99 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 7.275. Tidak terdapat perbedaan yang

signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing.



**Gambar 4.31** Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3

Hal ini berarti bahwa pemodelan curah hujan di wilayah Gubeng dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing. Hasil prediksi dari pemodelan curah hujan di wilayah Gubeng pada grid 1x1 lebih baik daripada hasil prediksi dari pemodelan curah hujan di wilayah Gubeng pada grid 3x3 karena menghasilkan nilai RMSE yang lebih kecil yaitu selisih 0.057.

#### 4.2.4 Wonorejo

Curah hujan di wilayah Wonorejo akan dimodelkan dengan variabel-variabel luaran GCM dengan teknik *statistical downscaling* menggunakan Regresi Gamma. Apabila model yang dihasilkan kurang baik, akan ditambahkan variabel prediktor seperti dummy dan lag dari curah hujan keputih. Penambahan variabel dummy diharapkan dapat memperbaiki hasil prediksi. Variabel dummy dikategorikan 1 sebagai musim hujan yaitu pada bulan November sampai April dan dikategorikan 0 sebagai musim kemarau yaitu pada bulan Mei sampai Oktober. Penambahan variabel lag diharapkan dapat meningkatkan hasil prediksi karena seperti yang sudah dibahas sebelumnya bahwa lag dari curah hujan

keputih memiliki nilai koefisien korelasi yang lebih tinggi dibandingkan variabel-variabel prediktor luaran GCM. Pemodelan curah hujan akan dilakukan pada grid 1x1 dan grid 3x3.

#### 4.2.4.1 Grid 1x1

Variabel yang digunakan pada pemodelan grid 1x1 adalah variabel pada lokasi 5. Digunakan data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Berikut ini dilakukan pemodelan curah hujan di wilayah Wonorejo dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Wonorejo untuk meningkatkan hasil prediksi.

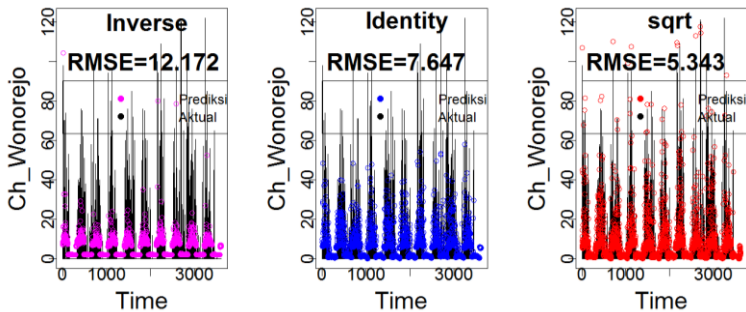
**Tabel 4.16** Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>
<i>Intercept</i>	-4.23	0.17	184.50	0.00	85.40	0.00
<b>dummy</b>	-0.33	0.00	5.50	0.00	1.48	0.00
<b>x15</b>	-29.51	0.49	-1116.00	0.34	69.68	0.90
<b>x25</b>	-0.09	0.97	-5.58	0.29	-25.25	0.00
<b>x35</b>	0.02	0.13	-0.60	0.00	-0.27	0.00
<b>x45</b>	0.0014	0.77	-0.09	0.25	-0.13	0.00
<b>lag1</b>	-0.0010	0.00	0.43	0.00	0.08	0.00

Pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse* hanya menghasilkan dua variabel yang berpengaruh signifikan terhadap curah hujan di wilayah Wonorejo yaitu variabel dummy dan lag 1. Sedangkan semua variabel prediktor seperti presipitasi, kelembapan, suhu, dan kecepatan angin tidak memiliki pengaruh yang signifikan terhadap curah hujan di wilayah Wonorejo. Hal ini dapat diketahui dari *P-value* yang kurang dari nilai *alpha* (5%). Lain hal nya pada pemodelan menggunakan link "sqrt" memberikan hasil bahwa semua variabel berpengaruh signifikan terhadap curah hujan di wilayah Wonorejo kecuali presipitasi. Hal ini dapat diketahui dari *P-value* yang lebih dari nilai *alpha* (5%). Sedangkan pemodelan menggunakan link

*identity* hanya variabel suhu yang berpengaruh signifikan terhadap curah hujan selain daripada intercept, dummy, dan lag 1. Berikut ini nilai RMSE yang dihasilkan oleh ketiga model dengan penggunaan link yang berbeda.

Nilai RMSE dari pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang paling kecil yaitu sebesar 5.343. Hal ini berarti hasil prediksi data curah hujan di wilayah Wonorejo menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Wonorejo dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*.

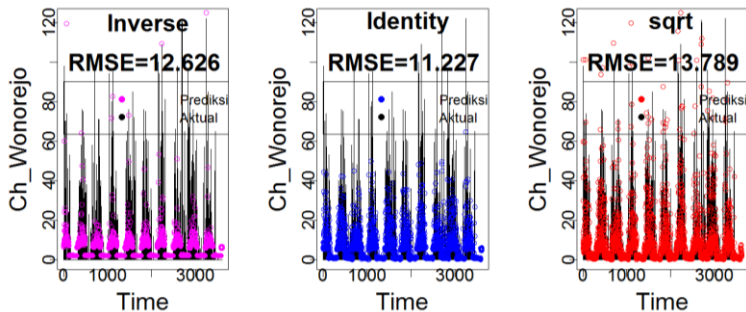


**Gambar 4.32** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1

Pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu sebesar 12.172. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Wonorejo menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Wonorejo dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonorejo dengan penambahan lag 1 dan lag 2.

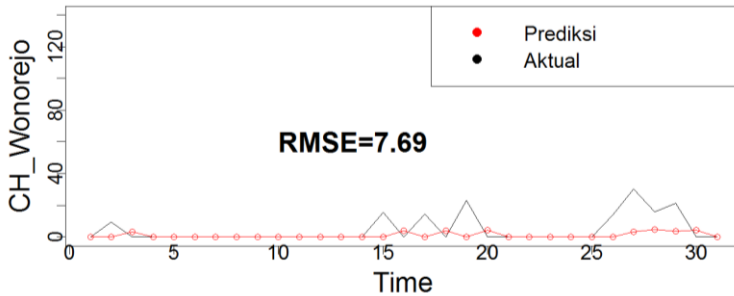
Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Wonorejo dengan penambahan lag 1 dan lag 2 memberikan hasil yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 11.227. Sedangkan pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang paling besar yaitu 13.789. Pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{4,t} = (85.4 + 1.48dummy_t + 69.68X_{1,t} - 25.25X_{2,t} - 0.27X_{3,t} + -0.13X_{4,t} + 0.08Y_{4,t-1})^2$$



**Gambar 4.33** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Wonorejo karena memiliki nilai RMSE terendah yaitu 5.567. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonorejo dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.



**Gambar 4.34** Hasil Prediksi Curah Hujan Menggunakan Data Testing

Hasil prediksi pemodelan curah hujan di wilayah Wonorejo dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 7.69 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 6.566. tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Wonorejo dengan semua variabel predictor dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.

#### 4.2.4.2 Grid 3x3

Variabel yang digunakan pada pemodelan grid 3x3 adalah variabel GCM pada 9 lokasi. Data yang digunakan adalah data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Pemodelan curah hujan di wilayah Wonorejo dengan variabel predictor luaran GCM pada grid 3x3 dilakukan dengan teknik reduksi dimensi menggunakan *Principal Components Analysis* atau yang biasa disebut dengan PCA yang selanjutnya dapat dimodelkan dengan regresi Gamma. Berikut ini dilakukan pemodelan curah hujan di wilayah

Wonorejo dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Wonorejo untuk meningkatkan hasil prediksi.

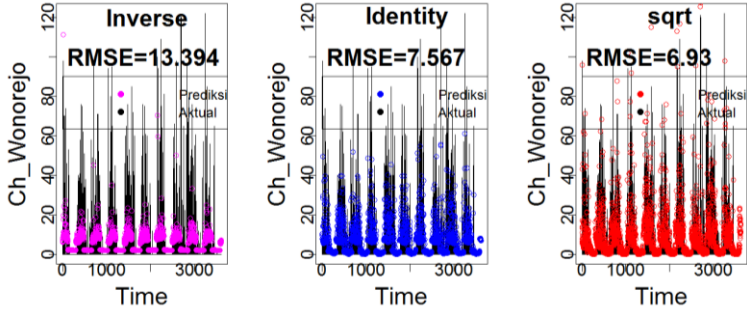
**Tabel 4.17** Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
Intercept	0.45	0.00	1.07	0.00	1.02	0.00
dummy	-0.31	0.00	5.99	0.00	1.93	0.00
pc11	-1.16	0.96	53.06	0.64	-176.68	0.04
pc12	-1.73	0.96	-228.40	0.00	-170.56	0.33
pc21	0.64	0.70	-0.37	0.96	19.35	0.01
pc22	3.57	0.25	36.72	0.00	20.62	0.00
pc31	-0.002	0.67	-0.03	0.02	-0.03	0.04
pc32	0.001	0.94	-0.39	0.00	-0.40	0.00
pc41	-0.001	0.75	-0.07	0.00	-0.02	0.03
pc42	0.004	0.57	-0.12	0.00	-0.12	0.00
lag 1	-0.001	0.00	0.44	0.00	0.08	0.00

Pemodelan curah hujan di wilayah Wonorejo dengan delapan komponen utama yang terbentuk dan variabel dummy serta penambahan lag 1 menggunakan link *inverse* diperoleh hasil variabel yang signifikan hanyalah intercept, variabel dummy dan lag 1 dari curah hujan. Sedangkan pada link *identity* dihasilkan semua variabel signifikan kecuali pc11 dan pc21. Regresi Gamma menggunakan link "sqrt" diperoleh hasil semua variabel signifikan kecuali pc12. Pc11 dan pc12 merupakan PC 1 dan PC 2 dari variabel Presipitasi. Pc 21 merupakan PC 1 dari variabel kelembapan.

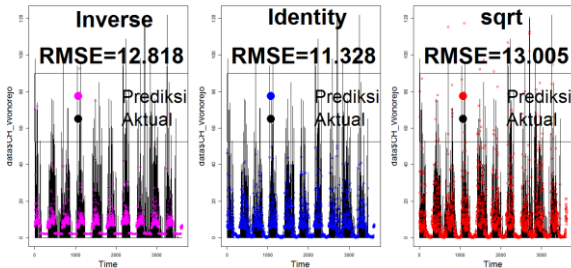
Nilai RMSE dari pemodelan curah hujan dengan delapan komponen utama yang terbentuk dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan "sqrt" memberikan hasil yang berbeda signifikan. Nilai RMSE pemodelan regresi Gamma dengan menggunakan link "sqrt" merupakan yang paling rendah yaitu sebesar 6.93. Hal ini berarti hasil prediksi curah hujan di wilayah Wonorejo menggunakan link "sqrt" merupakan yang paling mendekati dengan data asli curah hujan di wilayah Wonorejo dibandingkan dengan hasil

prediksi pemodelan menggunakan link *identity* dan link *inverse*. Sedangkan pemodelan regresi Gamma dengan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 13.394.



**Gambar 4.35** Prediksi Curah Hujan Wonorejo dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1

Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Wonorejo menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Wonorejo dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonorejo dengan penambahan lag 1 dan lag 2.



**Gambar 4.36** Prediksi Curah Hujan Wonorejo pada Grid 3x3 dengan Penambahan Lag 1, 2

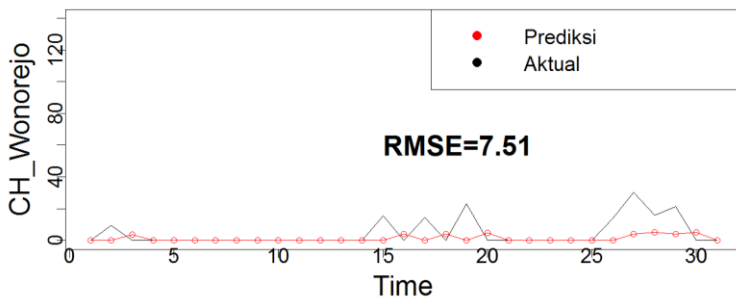
Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Wonorejo pada grid 3x3 dengan penambahan lag 1 dan lag 2 menghasilkan nilai RMSE yang lebih besar daripada nilai RMSE



yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 11.328. Sedangkan pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 12.818. Dalam hal ini pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{4,t} = (1.02 + 1.93dummy_t - 176.68PC1X_{1,t} - 170.56PC2X_{1,t} + 19.35PC1X_{2,t} + 20.62PC2X_{2,t} + 0.03PC1X_{3,t} - 0.4PC2X_{3,t} - 0.02PC1X_{4,t} - 0.12PC2X_{4,t} + 0.08Y_{4,t-1})^2$$

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Wonorejo karena memiliki nilai RMSE terendah yaitu 6.462. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonorejo dengan delapan komponen utama yang terbentuk pada grid 3x3 dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.



**Gambar 4.37** Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3

Hasil prediksi pemodelan curah hujan di wilayah Wonorejo dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar

7.51 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 6.93. Tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Wonorejo dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing. Hasil prediksi dari pemodelan curah hujan di wilayah Wonorejo pada grid 1x1 lebih baik daripada Hasil prediksi dari pemodelan curah hujan di wilayah Wonorejo pada grid 3x3 karena menghasilkan nilai RMSE yang lebih kecil.

#### **4.2.5 Wonokromo**

Curah hujan di wilayah Wonokromo akan dimodelkan dengan variabel-variabel luaran GCM dengan teknik *statistical downscaling* menggunakan Regresi Gamma. Apabila model yang dihasilkan kurang baik, akan ditambahkan variabel prediktor seperti dummy dan lag dari curah hujan keputih. Penambahan variabel dummy diharapkan dapat memperbaiki hasil prediksi. Variabel dummy dikategorikan 1 sebagai musim hujan yaitu pada bulan November sampai April dan dikategorikan 0 sebagai musim kemarau yaitu pada bulan Mei sampai Oktober. Penambahan variabel lag diharapkan dapat meningkatkan hasil prediksi karena seperti yang sudah dibahas sebelumnya bahwa lag dari curah hujan keputih memiliki nilai koefisien korelasi yang lebih tinggi dibandingkan variabel-variabel prediktor luaran GCM. Pemodelan curah hujan akan dilakukan pada grid 1x1 dan grid 3x3.

##### **4.2.5.1 Grid 1x1**

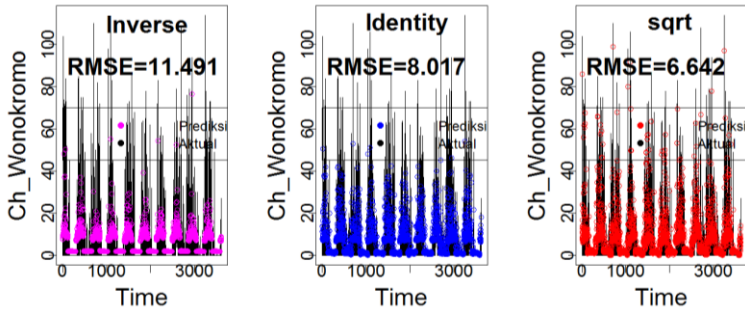
Variabel yang digunakan pada pemodelan grid 1x1 adalah variabel pada lokasi 5. Digunakan data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil.

Berikut ini dilakukan pemodelan curah hujan di wilayah Wonokromo dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Wonokromo untuk meningkatkan hasil prediksi.

**Tabel 4.18** Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>
<i>Intercept</i>	-4.82	0.06	86.24	0.00	67.82	0.00
<b>dummy</b>	-0.37	0.00	6.47	0.00	1.61	0.00
<b>x15</b>	-12.46	0.72	1609.17	0.27	-110.21	0.86
<b>x25</b>	-0.73	0.66	-29.44	0.00	-33.54	0.00
<b>x35</b>	0.02	0.04	-0.28	0.00	-0.22	0.00
<b>x45</b>	0.001	0.79	-0.22	0.00	-0.12	0.00
<b>lag1</b>	-0.001	0.00	0.41	0.00	0.06	0.00

Pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse* hanya menghasilkan dummy, lag 1 dan variabel suhu yang berpengaruh signifikan terhadap curah hujan di wilayah Wonokromo. Sedangkan ketiga variabel prediktor seperti presipitasi, kelembapan, dan kecepatan angin tidak memiliki pengaruh yang signifikan terhadap curah hujan di wilayah Wonokromo. Hal ini dapat diketahui dari *P-value* yang kurang dari nilai *alpha* (5%). Lain hal nya pada pemodelan menggunakan link *identity* maupun link "sqrt" memberikan hasil bahwa semua variabel berpengaruh signifikan terhadap curah hujan di wilayah Wonokromo kecuali variabel presipitasi. Hal ini dapat diketahui dari *P-value* yang lebih dari nilai *alpha* (5%). Berikut ini nilai RMSE yang dihasilkan oleh ketiga model dengan penggunaan link yang berbeda.



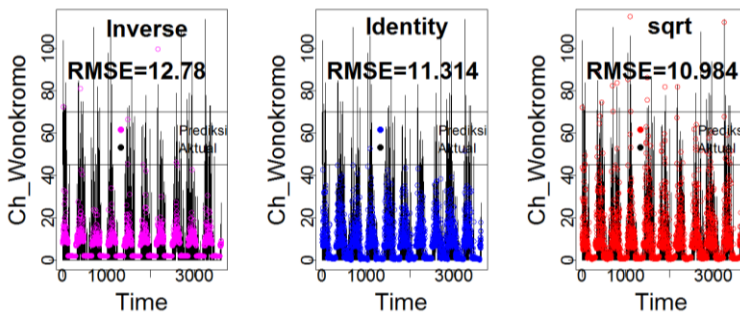
**Gambar 4.38** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1

Nilai RMSE dari pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang paling kecil yaitu sebesar 6.642. Hal ini berarti hasil prediksi data curah hujan di wilayah Wonokromo menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Wonokromo dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu sebesar 11.491. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Wonokromo menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Wonokromo dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonokromo dengan penambahan lag 1 dan lag 2.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Wonokromo dengan penambahan lag 1 dan lag 2 memberikan hasil yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE

yang paling kecil yaitu sebesar 10.984. Sedangkan pemodelan menggunakan link invers menghasilkan nilai RMSE yang paling besar yaitu 12.78. Dalam hal ini pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{5,t} = (67.82 + 1.61dummy_t - 110.21X_{1,t} - 33.54X_{2,t} + 0.22X_{3,t} - 0.12X_{4,t} + 0.06Y_{5,t-1})^2$$



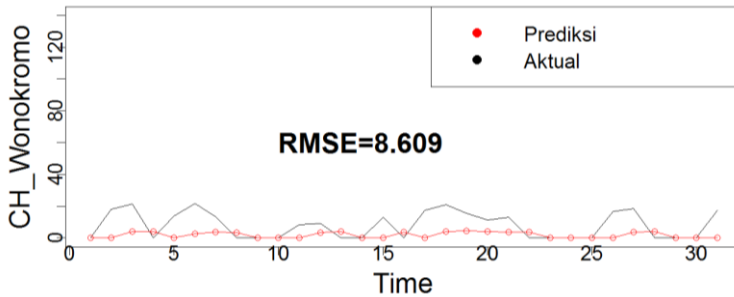
**Gambar 4.39** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Wonokromo karena memiliki nilai RMSE terendah yaitu 6.642. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonokromo dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.

Hasil prediksi pemodelan curah hujan di wilayah Wonokromo dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 8.609 pada data Testing.

Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 6.642. Tidak terdapat perbedaan yang signifikan dari hasil

prediksi dengan data Training maupun hasil prediksi dengan data Testing.



**Gambar 4.40** Hasil Prediksi Curah Hujan Menggunakan Data Testing

Hal ini berarti bahwa pemodelan curah hujan di wilayah Wonokromo dengan semua variabel predictor dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.

#### 4.2.5.2 Grid 3x3

Variabel yang digunakan pada pemodelan grid 3x3 adalah variabel GCM pada 9 lokasi. Data yang digunakan adalah data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Pemodelan curah hujan di wilayah Wonokromo dengan variabel predictor luaran GCM pada grid 3x3 dilakukan dengan teknik reduksi dimensi menggunakan *Principal Components Analysis* atau yang biasa disebut dengan PCA yang selanjutnya dapat dimodelkan dengan regresi Gamma. Berikut ini dilakukan pemodelan curah hujan di wilayah Wonokromo dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Wonokromo untuk meningkatkan hasil prediksi.

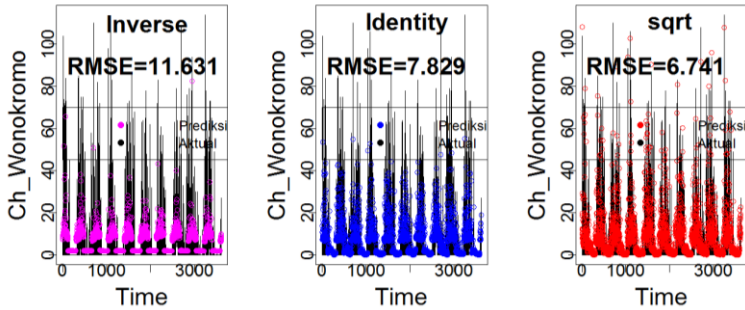
**Tabel 4.19** Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
Intercept	0.48	0.00	1.46	0.00	1.17	0.00
dummy	-0.36	0.00	6.10	0.00	1.72	0.00
pc11	-3.80	0.83	1568.00	0.00	341.60	0.11
pc12	11.43	0.68	-302.90	0.26	-137.10	0.77
pc21	-0.01	1.00	11.33	0.19	28.01	0.00
pc22	2.90	0.26	32.74	0.00	34.33	0.00
pc31	-0.0030	0.44	0.08	0.00	0.06	0.00
pc32	-0.0039	0.56	-0.25	0.00	-0.33	0.00
pc41	-0.0003	0.83	-0.04	0.00	-0.04	0.00
pc42	0.0003	0.95	-0.22	0.00	-0.20	0.00
lag 1	-0.0009	0.00	0.42	0.00	0.07	0.00

Pemodelan curah hujan di wilayah Wonokromo dengan delapan komponen utama yang terbentuk dan variabel dummy serta penambahan lag 1 menggunakan link *inverse* diperoleh hasil variabel yang signifikan hanyalah intercept, dummy, dan lag 1 dari curah hujan. Sedangkan pada link *identity* dihasilkan semua variabel signifikan kecuali pc12 dan pc21. Regresi Gamma menggunakan link "sqrt" diperoleh hasil semua variabel signifikan kecuali pc11 dan pc12. Pc11 dan pc12 merupakan PC 1 dan PC 2 dari variabel Presipitasi. Pc21 dan pc22 merupakan PC 1 dan PC 2 dari variabel kelembapan. Pc 31 dan pc32 merupakan PC 1 dan PC 2 dari variabel suhu. pc 41 dan pc42 merupakan PC 1 dan PC 2 dari variabel kecepatan angin. Berikut nilai RMSE yang dihasilkan ketiga model.

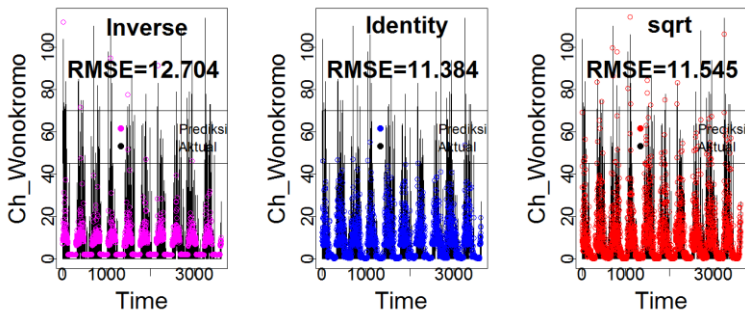
Nilai RMSE dari pemodelan curah hujan dengan delapan komponen utama yang terbentuk dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan "sqrt" memberikan hasil yang berbeda signifikan. Nilai RMSE pemodelan regresi Gamma dengan menggunakan link "sqrt" merupakan yang paling rendah yaitu sebesar 6.741. Hal ini berarti hasil prediksi data curah hujan di wilayah Wonokromo menggunakan link "sqrt" merupakan yang paling mendekati dengan data asli curah hujan di wilayah Wonokromo dibandingkan

dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Sedangkan pemodelan regresi Gamma dengan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 11.631.



**Gambar 4.41** Prediksi Curah Hujan Wonokromo dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1

Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Wonokromo menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Wonokromo dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonokromo dengan penambahan lag 1 dan lag 2.



**Gambar 4.42** Prediksi Curah Hujan Wonokromo dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2



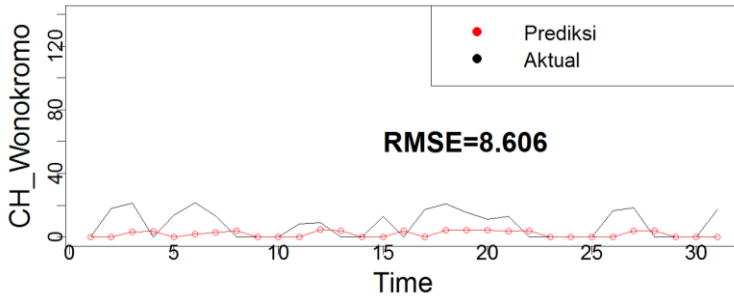
Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Wonokromo pada grid 3x3 dengan penambahan lag 1 dan lag 2 menghasilkan nilai RMSE yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 11.384. Sedangkan pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 12.704. Pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{5,t} = (1.17 + 1.72dummy_t - 341.6PC1X_{1,t} - 137.1PC2X_{1,t} + 28.01PC1X_{2,t} + 34.33PC2X_{2,t} + 0.06PC1X_{3,t} - 0.33PC2X_{3,t} - 0.04PC1X_{4,t} - 0.2PC2X_{4,t} + 0.07Y_{5,t-1})^2$$

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Wonokromo karena memiliki nilai RMSE terendah yaitu 6.462. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Wonokromo dengan delapan komponen utama yang terbentuk pada grid 3x3 dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.

Hasil prediksi pemodelan curah hujan di wilayah Wonokromo dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 8.606 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 6.741. Tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Wonokromo dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan

menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.



**Gambar 4.43** Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3

Hasil prediksi dari pemodelan curah hujan di wilayah Wonokromo pada grid 3x3 lebih baik daripada Hasil prediksi dari pemodelan curah hujan di wilayah Wonokromo pada grid 1x1 karena menghasilkan nilai RMSE yang lebih kecil.

#### 4.2.6 Gunung Sari

Curah hujan di wilayah Gunung Sari akan dimodelkan dengan variabel-variabel luaran GCM dengan teknik *statistical downscaling* menggunakan Regresi Gamma. Apabila model yang dihasilkan kurang baik, akan ditambahkan variabel prediktor seperti dummy dan lag dari curah hujan keputih. Penambahan variabel dummy diharapkan dapat memperbaiki hasil prediksi. Variabel dummy dikategorikan 1 sebagai musim hujan yaitu pada bulan November sampai April dan dikategorikan 0 sebagai musim kemarau yaitu pada bulan Mei sampai Oktober. Penambahan variabel lag diharapkan dapat meningkatkan hasil prediksi karena seperti yang sudah dibahas sebelumnya bahwa lag dari curah hujan keputih memiliki nilai koefisien korelasi yang lebih tinggi dibandingkan variabel-variabel prediktor luaran GCM. Pemodelan curah hujan akan dilakukan pada grid 1x1 dan grid 3x3.

#### 4.2.6.1 Grid 1x1

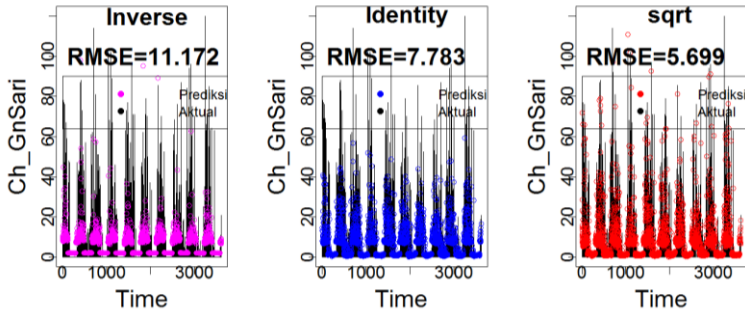
Variabel yang digunakan pada pemodelan grid 1x1 adalah variabel pada lokasi 5. Digunakan data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Berikut ini dilakukan pemodelan curah hujan di wilayah Gunung Sari dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan Gunung Sari untuk meningkatkan hasil prediksi. Pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse* hanya menghasilkan satu variabel prediktor yang berpengaruh signifikan terhadap curah hujan di wilayah Gunung Sari yaitu suhu selain daripada intercept, dummy dan lag 1. Sedangkan semua variabel prediktor seperti presipitasi, kelembapan, dan kecepatan angin tidak memiliki pengaruh yang signifikan terhadap curah hujan di wilayah Gunung Sari.

**Tabel 4.20** Pemodelan Curah Hujan dengan Empat Variabel Klimatologi, Variabel Dummy, dan Curah Hujan Lag 1

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>	Estimate	<i>P-value</i>
<i>Intercept</i>	-4.94	0.04	-59.82	0.13	-8.30	0.67
<b>dummy</b>	-0.37	0.00	6.37	0.00	1.59	0.00
<b>x15</b>	-44.04	0.17	1587.05	0.18	826.35	0.08
<b>x25</b>	-0.97	0.51	-27.60	0.04	-10.02	0.23
<b>x35</b>	0.02	0.03	0.21	0.12	0.03	0.61
<b>x45</b>	0.005	0.20	-0.17	0.00	-0.11	0.00
<b>lag1</b>	-0.001	0.00	0.43	0.00	0.07	0.00

Hal ini dapat diketahui dari *P-value* yang kurang dari nilai *alpha* (5%). Lain hal nya pada pemodelan menggunakan link *identity* memberikan hasil bahwa variabel dummy, kelembapan, kecepatan angin, dan lag 1 berpengaruh signifikan terhadap curah hujan di wilayah Gunung Sari. Kecepatan angin, dummy, lag 1 memberikan pengaruh yang signifikan terhadap curah hujan di wilayah Gunung Sari jika

pemodelan dilakukan dengan menggunakan link “sqrt”. Hal ini dapat diketahui dari *P-value* yang kurang dari nilai *alpha* (5%).



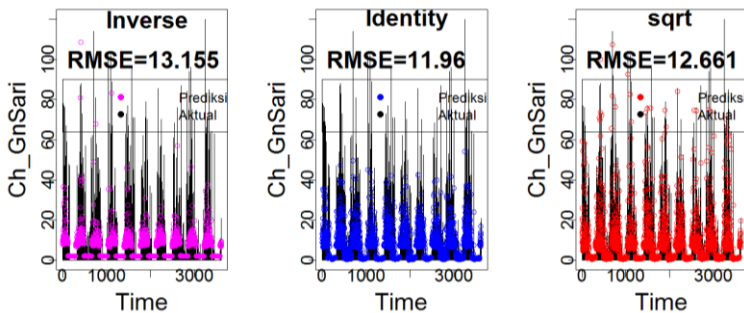
**Gambar 4.44** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1

Nilai RMSE dari pemodelan curah hujan dengan semua variabel dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan “sqrt” memberikan hasil yang berbeda signifikan. Pemodelan menggunakan link “sqrt” menghasilkan nilai RMSE yang paling kecil yaitu sebesar 5.699. Hal ini berarti hasil prediksi data curah hujan di wilayah Gunung Sari menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Gunung Sari dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu sebesar 11.172. Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Gunung Sari menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Gunung Sari dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Gunung Sari dengan penambahan lag 1 dan lag 2.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Gunung Sari dengan penambahan lag 1 dan lag 2

memberikan hasil yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 11.96. Sedangkan pemodelan menggunakan link inverse menghasilkan nilai RMSE yang paling besar yaitu 13.155. Dalam hal ini pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.

$$\hat{Y}_{6,t} = (-8.30 + 1.59dummy_t + 826.35X_{1,t} - 10.02X_{2,t} + 0.03X_{3,t} - 0.11X_{4,t} + 0.07Y_{6,t-1})^2$$

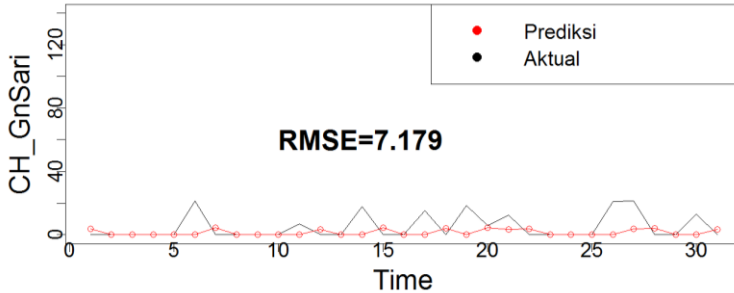


**Gambar 4.45** Curah Hujan dengan Predictor Empat Variabel Klimatologi dan Curah Hujan pada Lag 1 dan Lag 2

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Gunung Sari karena memiliki nilai RMSE terendah yaitu 5.567. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Gunung Sari dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.

Hasil prediksi pemodelan curah hujan di wilayah Gunung Sari dengan semua variabel predictor dan penambahan variabel dummy dan lag 1 dari curah hujan menggunakan link “sqrt”

menghasilkan nilai RMSE sebesar 7.179 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 5.699.



**Gambar 4.46** Hasil Prediksi Curah Hujan Menggunakan Data Testing

Tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Gunung Sari dengan semua variabel predictor dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.

#### 4.2.6.2 Grid 3x3

Variabel yang digunakan pada pemodelan grid 3x3 adalah variabel GCM pada 9 lokasi. Data yang digunakan adalah data training yaitu pada 1 Januari 2009 sampai dengan 30 November 2018. Ukuran kebaikan model yang digunakan adalah nilai RMSE. Pemilihan model terbaik berdasarkan pada nilai RMSE yang terkecil. Pemodelan curah hujan di wilayah Gunung Sari dengan variabel predictor luaran GCM pada grid 3x3 dilakukan dengan teknik reduksi dimensi menggunakan *Principal Components Analysis* atau yang biasa disebut dengan PCA yang selanjutnya dapat dimodelkan dengan regresi Gamma. Berikut ini dilakukan pemodelan curah hujan di wilayah Gunung Sari dengan semua variabel dan penambahan variabel dummy serta

lag 1 dari curah hujan Gunung Sari untuk meningkatkan hasil prediksi.

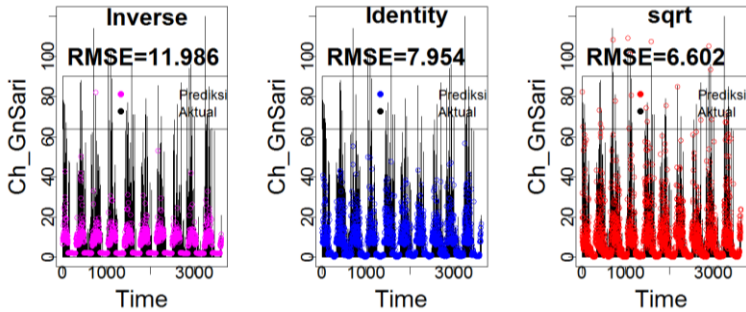
**Tabel 4.21** Pemodelan Curah Hujan dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan Lag 1 pada Grid 3x3

	<i>Inverse</i>		<i>Identity</i>		"sqrt"	
	Estimate	P-value	Estimate	P-value	Estimate	P-value
Intercept	0.47	0.00	1.22	0.00	1.11	0.00
dummy	-0.35	0.00	6.23	0.00	1.73	0.00
pc11	-11.79	0.53	1244.00	0.00	596.93	0.01
pc12	-17.15	0.57	163.90	0.87	417.70	0.27
pc21	0.55	0.66	9.63	0.30	20.41	0.00
pc22	5.61	0.04	32.48	0.00	25.61	0.00
pc31	-0.0022	0.57	0.02	0.26	0.01	0.10
pc32	-0.01	0.29	-0.30	0.00	-0.32	0.00
pc41	0.0005	0.75	-0.05	0.00	-0.04	0.00
pc42	0.0002	0.96	-0.14	0.00	-0.15	0.00
lag 1	-0.0008	0.00	0.41	0.00	0.07	0.00

Pemodelan curah hujan di wilayah Gunung Sari dengan delapan komponen utama yang terbentuk dan variabel dummy serta penambahan lag 1 menggunakan link *inverse* diperoleh hasil variabel yang signifikan hanyalah intercept, variabel dummy, pc22 dan lag 1 dari curah hujan. Sedangkan pada link *identity* dihasilkan semua variabel signifikan kecuali pc12, pc21, dan pc31. Regresi Gamma menggunakan link "sqrt" diperoleh hasil semua variabel signifikan kecuali 12 dan pc31. Pc11 dan pc12 merupakan PC 1 dan PC 2 dari variabel Presipitasi. Pc21 dan pc22 merupakan PC 1 dan PC 2 dari variabel kelembapan. Pc 31 dan pc32 merupakan PC 1 dan PC 2 dari variabel suhu. pc 41 dan pc42 merupakan PC 1 dan PC 2 dari variabel kecepatan angin. Berikut nilai RMSE yang dihasilkan ketiga model.

Nilai RMSE dari pemodelan curah hujan dengan delapan komponen utama yang terbentuk dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link *inverse*, *identity*, dan "sqrt" memberikan hasil yang berbeda signifikan. Nilai RMSE pemodelan regresi Gamma dengan menggunakan link "sqrt" merupakan yang paling rendah yaitu sebesar 6.602. Hal ini berarti hasil prediksi data curah hujan di wilayah Gunung Sari

menggunakan link “sqrt” merupakan yang paling mendekati dengan data asli curah hujan di wilayah Gunung Sari dibandingkan dengan hasil prediksi pemodelan menggunakan link *identity* dan link *inverse*. Sedangkan pemodelan regresi Gamma dengan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 11.986.



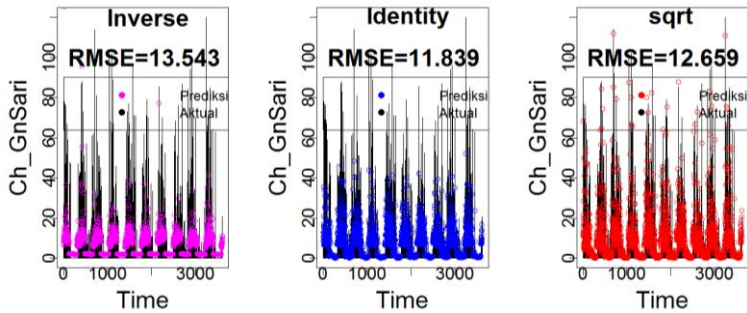
**Gambar 4.47** Prediksi Curah Hujan Gunung Sari dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1

Hal ini berarti bahwa hasil prediksi data curah hujan di wilayah Gunung Sari menggunakan link *inverse* merupakan yang paling jauh dari data asli curah hujan di wilayah Gunung Sari dibandingkan dengan hasil prediksi pemodelan menggunakan link yang lain.

Nilai RMSE hasil prediksi pemodelan curah hujan di wilayah Gunung Sari pada grid 3x3 dengan penambahan lag 1 dan lag 2 menghasilkan nilai RMSE yang lebih besar daripada nilai RMSE yang dihasilkan dari pemodelan dengan penambahan lag 1 saja. Pemodelan menggunakan link *identity* menghasilkan nilai RMSE yang paling kecil yaitu sebesar 11.839. Sedangkan pemodelan menggunakan link *inverse* menghasilkan nilai RMSE yang paling besar yaitu 13.543. Pemodelan dengan penambahan lag 1 memberikan hasil prediksi yang lebih baik daripada pemodelan dengan penambahan variabel lag 1 dan lag 2. Berikut ini hasil pemodelannya.



$$\hat{Y}_{6,t} = (1.11 + 1.73dummy_t + 596.93PC1X_{1,t} + 417.7PC2X_{1,t} + 20.41PC1X_{2,t} + 25.62PC2X_{2,t} + 0.01PC1X_{3,t} - 0.32PC2X_{3,t} - 0.04PC1X_{4,t} - 0.15PC2X_{4,t} + 0.07Y_{6,t-1})^2$$

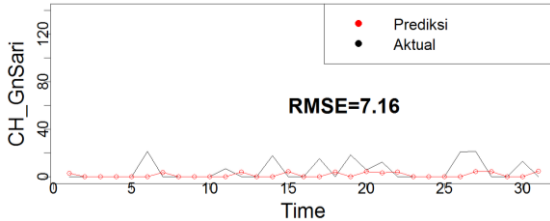


**Gambar 4.48** Prediksi Curah Hujan Gunung Sari dengan 8 Variabel PC, Variabel Dummy, dan Curah Hujan pada Lag 1 dan Lag 2

Pemodelan terbaik menggunakan link “sqrt” pada pemodelan dengan penambahan variabel lag 1 dari curah hujan di wilayah Gunung Sari karena memiliki nilai RMSE terendah yaitu 6.462. Berikut ini hasil prediksi pemodelan curah hujan di wilayah Gunung Sari dengan delapan komponen utama yang terbentuk pada grid 3x3 dan penambahan variabel dummy dan lag 1 menggunakan data Testing yaitu pada 1 Desember 2018 sampai dengan 31 Desember 2018.

Hasil prediksi pemodelan curah hujan di wilayah Gunung Sari dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan menggunakan link “sqrt” menghasilkan nilai RMSE sebesar 7.16 pada data Testing. Sedangkan nilai RMSE yang dihasilkan dari data Training sebesar 6.602. tidak terdapat perbedaan yang signifikan dari hasil prediksi dengan data Training maupun hasil prediksi dengan data Testing. Hal ini berarti bahwa pemodelan curah hujan di wilayah Gunung Sari dengan delapan komponen utama yang terbentuk dari PCA pada grid 3x3 dan penambahan variabel dummy serta lag 1 dari curah hujan

menggunakan link “sqrt” telah sesuai digunakan untuk memprediksi curah hujan pada data Testing.



**Gambar 4.49** Hasil Prediksi Curah Hujan Menggunakan Data Testing pada Grid 3x3

Hasil prediksi dari pemodelan curah hujan di wilayah Gunung Sari pada grid 1x1 lebih baik daripada Hasil prediksi dari pemodelan curah hujan di wilayah Gunung Sari pada grid 3x3 karena menghasilkan nilai RMSE yang lebih kecil.

#### 4.2.7 Model Terbaik

Model terbaik dari pemodelan pada grid 1x1 dan 3x3 terhadap enam wilayah pengamatan curah hujan yaitu Keputih, Kedung Cowek, Gubeng, Wonorejo, Wonokromo, Gunung Sari dapat diketahui dari nilai RMSE sebagai berikut.

**Tabel 4.22** Nilai RMSE Semua Model

Wilayah	Grid 1x1	Grid 3x3
Keputih	5.28	5.17
Kedung Cowek	4.95	5.19
Gubeng	9.9	9.99
Wonorejo	7.69	7.51
Wonokromo	8.609	8.606
Gunung Sari	7.179	7.16

Berdasarkan nilai RMSE pemodelan pada grid 1x1 dan 3x3 terhadap enam wilayah pengamatan curah hujan yaitu Keputih, Kedung Cowek, Gubeng, Wonorejo, Wonokromo, Gunung Sari dapat diketahui bahwa di wilayah Keputih, Wonorejo, Wonokromo, dan Gunung Sari pemodelan curah hujan lebih baik dilakukan pada Grid 3x3 terlihat dari nilai RMSE yang lebih kecil pada grid 3x3. Hal ini berarti curah hujan di wilayah Keputih,

Wonorejo, Wonokromo, dan Gunung Sari tidak hanya dipengaruhi oleh presipitasi, kelembapan, suhu, dan kecepatan angin di wilayah Surabaya saja tetapi juga dipengaruhi oleh presipitasi, kelembapan, suhu, dan kecepatan angin di sekitar wilayah Surabaya. Sedangkan di wilayah Kedung Cowek dan Gubeng, pemodelan curah hujan lebih baik dilakukan pada Grid 1x1 terlihat dari nilai RMSE yang lebih kecil pada grid 1x1. Hal ini berarti curah hujan di wilayah Kedung Cowek dan Gubeng hanya dipengaruhi oleh presipitasi, kelembapan, suhu, dan kecepatan angin di wilayah Surabaya saja.

### 4.3 Prediksi Banjir

Dampak negatif yang ditimbulkan akibat intensitas curah hujan berlebih di Surabaya adalah banjir. Salah satu penyebab terjadinya banjir adalah adanya genangan air yang terjadi pada suatu tempat dalam kurun waktu tertentu. Salah satu cara untuk dapat mengelola resiko terjadinya genangan adalah dengan memperkirakan kapan suatu daerah akan berpotensi terkena genangan. Analisa ini dapat dilakukan dengan memperkirakan potensi terjadinya hujan lebat (curah hujan tinggi) yang diturunkan dari data variabel-variabel yang mempengaruhi curah hujan yang diperoleh dari data GCM (General Circulation Model) yang menyediakan data-data iklim secara global. Setelah didapatkan prediksi curah hujan pada analisis sebelumnya maka, selanjutnya data prediksi curah hujan yang dihasilkan akan digunakan untuk melakukan perhitungan banjir. Perhitungan banjir dilakukan dengan metode rasional dan metode mononobe. Berikut curah hujan maksimum yang dihasilkan pada masing-masing wilayah.

**Tabel 4.23** Curah Hujan Maksimum

Stasiun	Curah Hujan Maksimum
Keputih	4.307555
Kedung Cowek	4.346684
Gubeng	4.92691
Wonorejo	5.160887
Wonokromo	4.323248
Gunung sari	4.570338

### 4.3.1 Keputih

Hasil prediksi curah hujan yang digunakan untuk melakukan perhitungan banjir di wilayah Keputih adalah prediksi curah hujan yang dihasilkan pada grid 3x3. Berikut hasil perhitungan banjir berdasarkan hasil prediksi curah hujan di wilayah Keputih pada grid 3x3.

**Tabel 4.24** Perhitungan Banjir Keputih

Saluran	Q Hidrolika	Q Hidrologi
Saluran HKSAN / Kalijudan	0.865	0.064
Saluran Kalijudan	8.251	0.041
Saluran Kalisari Timur	8.251	0.068
Saluran laban Sari	21.426	0.072
Saluran Larangan	63.549	0.079
Saluran Manyar Kertoadi	8.251	0.060
Saluran Manyar Sabrangan	8.251	0.066
Saluran Mulyorejo	119.168	0.042
Saluran Ring Road ITS	35.023	0.070
Saluran Sutorejo	3.413	0.058

Berdasarkan perhitungan Q Hidrolika dan Q Hidrologi dapat diketahui bahwa apabila Q Hidrologi lebih besar dari Q Hidrolika, maka diprediksi akan terjadi banjir. Dari tabel perhitungan banjir dapat diketahui bahwa di semua saluran wilayah Keputih tidak berpotensi terjadi banjir karena Q Hidrologi kurang dari Q Hidrolika.

### 4.3.2 Kedung Cowek

Hasil prediksi curah hujan yang digunakan untuk melakukan perhitungan banjir di wilayah Kedung Cowek adalah prediksi curah hujan yang dihasilkan pada grid 1x1. Berikut hasil perhitungan banjir berdasarkan hasil prediksi curah hujan di wilayah Kedung Cowek pada grid 1x1.

**Tabel 4.25** Perhitungan Banjir Kedung Cowek

Saluran	Q Hidrolika	Q Hidrologi
Saluran Gading	89.007	0.068
Saluran Jeblokan	103.401	0.060
Saluran Kenjeran AL	103.401	0.082
Saluran Kenjeran Lama	176.320	0.081
Saluran Kenjeran Pantai Ria	53.032	0.079
Saluran Pegirian	307.968	0.050

<b>Saluran</b>	<b>Q Hidrolika</b>	<b>Q Hidrologi</b>
Saluran Kapas Madya	103.401	0.093
Saluran Kedung Sroko	21.426	0.097
Saluran Putra Agung	63.549	0.088
Saluran Tambaksari	5.503	0.066
Saluran Tambak Segaran	43.214	0.089
Saluran Sidotopo Wetan	7.739	0.048
Saluran Tambak Wedi	862.397	0.067

Berdasarkan perhitungan Q Hidrolika dan Q Hidrologi dapat diketahui bahwa apabila Q Hidrologi lebih besar dari Q Hidrolika, maka diprediksi akan terjadi banjir. Dari tabel perhitungan banjir dapat diketahui bahwa di semua saluran wilah Kedung Cowek tidak berpotensi terjadi banjir karena Q Hidrologi kurang dari Q Hidrolika.

### 4.3.3 Gubeng

Hasil prediksi curah hujan yang digunakan untuk melakukan perhitungan banjir di wilayah Gubeng adalah prediksi curah hujan yang dihasilkan pada grid 1x1. Berikut hasil perhitungan banjir berdasarkan hasil prediksi curah hujan di wilayah Gubeng pada grid 1x1.

**Tabel 4.26** Perhitungan Banjir Gubeng

<b>Saluran</b>	<b>Q Hidrolika</b>	<b>Q Hidrologi</b>
Saluran Jeblokan	89.007	0.055
Saluran Juwangan	16.247	0.078
Saluran Kalibokor	284.374	0.048
Saluran Kalidami	1281.770	0.079
Saluran Kalisari THR	5.503	0.087
Saluran Kamboja	5.503	0.083
Saluran Kampung Seng	11.088	0.077
Saluran Kedung Tarukan	176.320	0.090
Saluran Kertajaya Viaduck	21.426	0.110
Saluran Menur	16.247	0.083
Saluran Mojo	0.300	0.069
Saluran Pacar	5.503	0.105
Saluran Pucang Adi	3.413	0.087
Saluran Pucang Jajar	11.932	0.074
Saluran Pucang Rinenggo	1.842	0.079
Saluran Simokerto / Donorejo	16.247	0.093

Saluran	Q Hidrolika	Q Hidrologi
Saluran Simolawang	8.251	0.097
Saluran Slamet	3.413	0.076
Saluran Srikana	8.251	0.075
Saluran Sawah Pulo	5.503	0.091
Saluran Sidoluhur dsk	0.370	0.070
Saluran Sidotopo Lor	53.032	0.089
Saluran Tengggung karya	0.865	0.047
Saluran Wonosari Lor	4.381	0.078

Berdasarkan perhitungan Q Hidrolika dan Q Hidrologi dapat diketahui bahwa apabila Q Hidrologi lebih besar dari Q Hidrolika, maka diprediksi akan terjadi banjir. Dari tabel perhitungan banjir dapat diketahui bahwa di semua saluran wilah Gubeng tidak berpotensi terjadi banjir karena Q Hidrologi kurang dari Q Hidrolika.

#### 4.3.4 Wonorejo

Hasil prediksi curah hujan yang digunakan untuk melakukan perhitungan banjir di wilayah Wonorejo adalah prediksi curah hujan yang dihasilkan pada grid 3x3. Berikut hasil perhitungan banjir berdasarkan hasil prediksi curah hujan di wilayah Wonorejo pada grid 3x3.

**Tabel 4.27** Perhitungan Banjir Wonorejo

Saluran	Q Hidrolika	Q Hidrologi
Saluran Wonorejo III - IV	1.773	0.053
Saluran Jatisari	3.413	0.093
Saluran Kedung Turi	0.370	0.052
Saluran Kedungsari	5.503	0.069
Saluran Indrapura	17.315	0.093
Saluran Rembang	21.426	0.093
Saluran Pelemahan	5666.167	0.119
Saluran Jl.Semarang	1.042	0.052
Saluran Keputran	4.381	0.074
Saluran Wonorejo I - II	4.615	0.081
Saluran Sumber Mulyo	12.933	0.097
Saluran Tembok Dukuh	304463.214	0.139
Saluran Jalan Opak	37311.287	0.135
Saluran Jepara	48.755	0.111
Saluran Dupak Ps.Turi	7.498	0.081

<b>Saluran</b>	<b>Q Hidrolika</b>	<b>Q Hidrologi</b>
Saluran Kutilang PA.Pesapen	3396.223	0.121
Saluran Ronggolawe	6.790	0.069
Saluran Mojopahit	1586.754	0.104
Saluran Bubutan	0.669	0.064
Saluran Embong Malang	1.296	0.076
Saluran Komering	4.381	0.115
Saluran Petemon Kuburan	3.413	0.096
Saluran Pregolan Bunder	1.842	0.065
Saluran Krembangan Jaya Sel	4.381	0.093
Saluran Ikan Mungsing	4.381	0.079
Saluran Juwono	8.251	0.098
Saluran Embong Kemiri	1.842	0.065
Saluran Keputran	4.381	0.115
Saluran Cempaka	4.381	0.108
Saluran Ciliwung	5.503	0.091
Saluran Ikan Cucut	4.381	0.078
Saluran Anwari	8.251	0.087
Saluran Sriwijaya	121.513	0.137
Saluran Brawijaya	3.722	0.078
Saluran Serayu	156.053	0.128
Saluran Greges	307.968	0.067
Saluran Kali Mas	1737.545	0.058

Berdasarkan perhitungan Q Hidrolika dan Q Hidrologi dapat diketahui bahwa apabila Q Hidrologi lebih besar dari Q Hidrolika, maka diprediksi akan terjadi banjir. Dari tabel perhitungan banjir dapat diketahui bahwa di semua saluran wilayah Wonorejo tidak berpotensi terjadi banjir karena Q Hidrologi kurang dari Q Hidrolika.

#### **4.3.5 Wonokromo**

Hasil prediksi curah hujan yang digunakan untuk melakukan perhitungan banjir di wilayah Wonokromo adalah prediksi curah hujan yang dihasilkan pada grid 1x1. Berikut hasil perhitungan banjir berdasarkan hasil prediksi curah hujan di wilayah Wonokromo pada grid 1x1.

**Tabel 4.28** Perhitungan Banjir Wonokromo

<b>Saluran</b>	<b>Q Hidrolika</b>	<b>Q Hidrologi</b>
Saluran Kebon Agung	176.320	0.033
Saluran Avoor Wonorejo	156.053	0.027
Saluran Medokan Ayu	35.023	0.036
Saluran Dukuh Menanggal	16.247	0.046
Saluran Gayungsari	16.247	0.062
Saluran Gayung Kebonsari	18.725	0.060
Saluran Wonokromo Tangkis	24.682	0.048
Saluran Jetis Kulon	1.842	0.054
Saluran Margorejo	1.842	0.053
Saluran Bendul Merisi Timur	16.247	0.071
Saluran Bendul Merisi Besar	16.247	0.078
Saluran Tepi Jl. A. Yani	2.592	0.053
Saluran Jemursari Prapen	24.682	0.046
Saluran Jemur Wonosari	4.381	0.084
Saluran Kutasari	4.381	0.069
Saluran Saroni Jiwo	18.725	0.063
Saluran Kali Rejo	8.251	0.066
Saluran Kali Rungkut	35.023	0.067
Saluran Tenggilis Mejoyo	4.381	0.066
Saluran Rungkut Kidul	21.426	0.082
Saluran Rungkut Asri Lor	35.023	0.077
Saluran Kedung Asem	1.842	0.056
Saluran Penjaringan Sari	6.790	0.050
Saluran Medokan Semampir	35.023	0.043
Saluran Semolowaru	35.023	0.052
Saluran Manyar	27.879	0.075
Saluran Manyar Rejo	13.985	0.074
Saluran Ngagel Wasono	2.592	0.081
Saluran Kalisumo	137.308	0.079
Saluran Ngagel Jaya Selatan	16.247	0.069
Saluran Krukah	1.296	0.058

Berdasarkan perhitungan Q Hidrolika dan Q Hidrologi dapat diketahui bahwa apabila Q Hidrologi lebih besar dari Q Hidrolika, maka diprediksi akan terjadi banjir. Dari tabel perhitungan banjir dapat diketahui bahwa di semua saluran wilayah Wonokromo tidak berpotensi terjadi banjir karena Q Hidrologi kurang dari Q Hidrolika.



### 4.3.6 Gunung Sari

Hasil prediksi curah hujan yang digunakan untuk melakukan perhitungan banjir di wilayah Gunung Sari adalah prediksi curah hujan yang dihasilkan pada grid 3x3. Berikut hasil perhitungan banjir berdasarkan hasil prediksi curah hujan di wilayah Gunung Sari pada grid 3x3.

**Tabel 4.29** Perhitungan Banjir Gunung Sari

<b>Saluran</b>	<b>Q Hidrolika</b>	<b>Q Hidrologi</b>
Lidah wetan - Kulon	89.0066	0.086885
Jeruk	16.24698	0.075346
Babadan Indah	1.842482	0.060767
Babadan Karang	11.93152	0.069119
Wiyung	27.87941	0.069225
Menganti	3.413415	0.055162
Gemol Kedurus	156.0525	0.095611
Gogor	43.2144	0.096073
jajar Tunggal	31.32418	0.081123
Gunung Sari	53.03195	0.070095
Bogangin I	5.503122	0.064928
Bogangin Baru	2.592149	0.051658
Kebraon	6.790135	0.059511
Kebraon Manis	6.790135	0.066395
Prima Kebraon	11.93152	0.072332
Pondok Maritim	6.790135	0.07307
Karang Klumprik Barat	11.93152	0.070282
Balas Klumprik	406711.8	0.120016

Berdasarkan perhitungan Q Hidrolika dan Q Hidrologi dapat diketahui bahwa apabila Q Hidrologi lebih besar dari Q Hidrolika, maka diprediksi akan terjadi banjir. Dari tabel perhitungan banjir dapat diketahui bahwa di semua saluran wilayah Gunung Sari tidak berpotensi terjadi banjir karena Q Hidrologi kurang dari Q Hidrolika.

*(Halaman ini sengaja dikosongkan)*

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan analisis yang telah dilakukan, diperoleh kesimpulan sebagai berikut:

1. Analisa statistika deskriptif dapat diketahui bahwa dari keenam stasiun, curah hujan tertinggi terjadi pada stasiun Keputih. Berdasarkan *time series plot* dapat diketahui bahwa data curah hujan pada enam stasiun memiliki pola musiman. Dari *boxplot* curah hujan dapat diketahui bahwa di bulan Mei-Oktober merupakan bulan yang jarang hujan. Berdasarkan scatterplot data curah hujan dengan masing-masing variabel predictor dapat diketahui bahwa lag dari curah hujan memiliki korelasi yang lebih besar terhadap curah hujan daripada variabel predictor luaran GCM.
2. Prediksi terbaik curah hujan di wilayah Keputih, Wonorejo, Wonokromo, dan Gunung Sari yaitu pada grid 3x3 diperoleh dari memodelkan curah hujan dengan semua variabel predictor dan penambahan dummy dan lag 1. Prediksi terbaik curah hujan di wilayah Kedung Cowek dan Gubeng yaitu pada grid 1x1 diperoleh dari memodelkan curah hujan dengan delapan komponen utama yang dihasilkan dari PCA dan penambahan dummy serta lag 1.
3. Prediksi banjir pada enam wilayah dapat diketahui bahwa seluruh saluran yang ada pada wilayah Keputih, Kedung Cowek, Gubeng, Wonokromo Wonorejo, dan Gunung Sari tidak berpotensi mengalami banjir berdasarkan dari nilai  $Q$  Hidrolikan yang lebih besar dari  $Q$  Hidrologi.

#### **5.2 Saran**

Saran untuk penelitian selanjutnya, agar dihasilkan prediksi yang memiliki nilai RMSE yang kecil dan untuk lebih meningkatkan ketepatan hasil prediksi dapat digunakan metode

nonparametrik. Berdasarkan perhitungan  $Q$  Hidrologi dapat dijadikan bahan pertimbangan bagi pembuat kebijakan dalam menentukan ukuran dari saluran agar  $Q$  hidrolika tidak kurang dari  $Q$  hidrologi, dengan demikian banjir dapat dicegah.

## DAFTAR PUSTAKA

- Anonim. Peraturan Pemerintah No. 38 tentang Sungai (2011). Indonesia.
- Badan Pusat Statistik Kota Surabaya. (2014). *Kota Surabaya dalam Angka*. Surabaya: Badan Pusat Statistik Kota Surabaya.
- Berliana, S. (2007). Kajian Dampak Pemanasan Global Terhadap Pola Curah Hujan Indonesia dengan Menggunakan Statistik Downscaling. *Statistika*, 7(2), 5-12.
- Brody, S. D., Kang, J. E., and Bernhardt, S. (2010). Identifying factors influencing flood mitigation at the local level in Texas and Florida: The role of organizational capacity. *Natural Hazards*, 52(1), 167-184. <https://doi.org/10.1007/s11069-009-9364-5>
- Busuioac A, Deliang C, Hellstrom C. (2001). Performance of statistical Downscaling models in GCM validation and regional climate change estimates: application for Swedish precipitation. *International Journal of Climatology* 21: 557-578.
- Chow, Van Te, Maidment, David R., Mays, Larry W. (1988). *Applied Hydrology*, McGraw Hill International Editions, Singapore
- Coulibaly P. and Dibike Y.B., (2004). Downscaling Precipitation and Temperature with Temporal Neural Networks. *American Meteorology Society* (6) : 483-496.
- Crane R.G., Hewitson B.C., (1998). Doubled CO<sub>2</sub> Precipitation Change for the Susquehanna Basin : Downscaling from the Genesis General Circulation Model. *International Journal of Climatology* (18) : 65-76.
- Czado, C. (2004). Gamma regression. [terhubung berkala]. <http://www.m4.ma.tum.de/courses/GLM/lec8.pdf>.
- Draper, N.R. and Smith, H., (1992). *Analisis Regresi Terapan, Edisi kedua*. Jakarta: PT. Gramedia Pustaka Utama.

- Goodess CM, Anagnostopoulou C, B`ardossy A, Frei C, Harpham C, Haylock MR, Hundecha Y, Maheras P, Ribalaygue J, Schmidli J, Schmith T, Tolika K, Tomozeiu R, Wilby RL. (2006). An intercomparison of statistical downscaling methods for Europe and European regions-assessing their performance with respect to extreme temperature and precipitation events. *Climatic Research Unit Research Publication 11 (CRU RP11)*.
- Hayes, D.C. and R.L. Young. (2005). *Comparison of Peak Discharge and Runoff Characteristic Estimates from the Rational Method to Field Observations for Small Basins in Central Virginia, Scientific Investi-gation Report 2005-5254*. USGS.
- Hewitson, B.C., and R.G. Crane. (1996). *Climate Downscaling: Techniques and Application*. Climate Research (7), pp. 85–95
- Johnson, R.A. and Wichern, D.W. (2002). *Applied Multivariate Statistical Analysis*, 5th End. Prentice Hall. New Jersey.
- Jolliffe. (2002). *Principal Component Analysis*. Second Edition, Springer-Verlag.
- Kidson JW, Thompson CS. (1998). A comparison of statistical and model – based downscaling techniques for estimating local climate variations. *Journal of Climate* (11) : 735–753.
- Kodoatie, R. J., dan Sugiyanto. (2002). *Banjir, Beberapa Penyebab dan Metode Pengendaliannya dalam Prespektif Lingkungan*. Yogyakarta: Pustaka Pelajar.
- Kysely J., (2002). Comparison of extremes in GCM-simulated downscaled and observed central – European temperature series. *Climate Research* 20: 211–222.
- Lembang, F. K. (2013). *Pemodelan Statistical Downscaling dengan Pendekatan Markov Chain Monte Carlo PCA*. *Prosiding FMIPA Universitas Pattimura*.

- Maryono, A. (2005). Menangani Banjir, Kekeringan dan Lingkungan. *Gadjah Mada University Press*.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models 2nd Edition*. London: Chapman dan Hall.
- Randomservices. (2018). Dipetik Maret 16, 2019, dari [randomservices.org](http://www.randomservices.org/):  
<http://www.randomservices.org/random/special/Gamma.html>
- Soleh, A. M. (2015). Pemodelan Statistical Downscaling untuk Menduga Curah Hujan Bulanan Menggunakan Model Linier Terampat Sebaran Gamma. *Informatika Pertanian*, 24(2), 215-222.
- Sri Harto. (1993). "Analisis Hidrologi". PT Gramedia. Jakarta
- Subarkah, I. (1980). *Hidrologi untuk Perencanaan Bangunan Air*. Bandung : Idea Dharma.
- Suherlan. (2001). *Zonasi Tingkat Kerentangan Banjir Kabupaten Bandung Menggunakan System Informasi Geografis*. (skripsi). Bogor.
- SurabayaTribunNews (2018). Diambil pada 13 Februari 2019, dari <http://m.tribunnews.com/amp/section/2019/01/31/kabar-terkini-banjir-di-surabaya-barat-penyebab-banjir-hingga-seorang-anak-meninggal-dunia?page=4>
- SurabayaTribunNews (2019). Diambil pada 13 Februari 2019, dari <http://surabaya.tribunnews.com/2019/01/11/surabaya-survey-center-banjir-dan-kemacetan-masih-jadi-pr-besar-pemerintah-kota-surabaya>.
- SurabayaTribunNews (2019). Diambil pada 13 Februari 2019, dari <http://www.tribunnews.com/regional/2019/01/31/banjir-di-surabaya-barat-banyak-mobil-penghuni-perumahan-elit-tenggelam>.
- Suripin. (2004). *Sistem Drainase Perkotaan yang Berkelanjutan*, Andi, Yogyakarta.

- Susanti, P. D., Miardini, A., dan Harjadi, B. (2017). Analisis kerentanan tanah longsor sebagai dasar mitigasi di Kabupaten Banjarnegara. *Jurnal Penelitian Pengelolaan Daerah Aliran Sungai*, 1(1), 49–59.
- Susilowati. (2010). Analisa Karakteristik Curah Hujan dan Kurva Intensitas Durasi Frekuensi (IDF) di Provinsi Lampung. *Rekayasa*, 14(1).
- Sutikno. (2008). *Statistical Downscaling Luaran GCM dan Pemanfaatannya untuk Peramalan Produksi Padi*. Disertasi. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Tolika K., P. Maheras, M. Vafiadis, H.A. Flocas, and A.A. Papadimitriou. (2007). *Simulation of Seasonal Precipitation and Raindays Over Greece: A Statistical Downscaling Technique Based on Artificial Neural Networks (ANNs)*. *Int. J. Climatol.* 27, pp. 861–881.
- Tripathi R C, Gupta C R, and Pair K P. (1993). Statistical Test Involving Several Independent Gamma Distribution *Ann. Inst. Statist. Math.* 45(4) pp 773-786.
- Tripathi V., V. V. Srinivas, and R. S. Nanjundiah. (2006). *Downscaling of Precipitation for Climate Change Scenarios : A Support Vector Machine Approach*. *Journal of Hydrology* 330, pp. 621-640.
- Upomo, T. C. (2016). Pemilihan Distribusi Probabilitas pada Analisa Hujan dengan Metode Goodness of Fit Test. *Teknik Sipil dan Perencanaan*, 18(2), 139-148.
- Von Storch H., (1999). On the Use of Inflation in Statistical Downscaling. *Journal Of Climate* (12) : 3505-3506.
- Walpole, R. (2007). *Pengantar Statistika Edisi ke-3 Terjemahan Bambang Sumantri*. Jakarta: Gramedia.
- Wigena, A.H., (2006). *Pemodelan Statistical Downscaling dengan Regresi Projection Pursuit untuk Peramalan Curah Hujan*



- Bulanan*. Disertasi. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Wilby R.L., and T.M. Wigley. (2000). *Precipitation Predictors for Downscaling : Observed and General Circulation Model Relationships*. Int. J. Climatol. 20 : 641-661.
- Wilks, D.S., Wilby, R.L. (1999). The weather generation game: a review of stochastic weather models. Prog. Phys. Geogr. 23, 329–357.
- Wanielista, M.P. (1990). *Hydrologi and Water Quantity Control*, John Wiley and Sons New York
- Yukse, O., Kankal, M., and Ucuncu, O. (2013). Assessment of big floods in the Eastern Black Sea Basin of Turkey. *Environment Monitoring Assessment*, 185, 797–814. <https://doi.org/10.1007/s10661-012-2592-2>
- Zorita, E. and von Storch, H. (1999). The Analog Method as a Simple Statistical Downscaling technique: comparison with more complicated method. *Journal of Climate*, 12, 2474-2489.



## LAMPIRAN

### Lampiran 1 Syntax Cropping Grid Data GCM

```
library(xlsx)
#kecepatan angin
datapr<-read.csv("C:/Users/Os 8.1/Desktop/sfcWind.csv", header=FALSE)
datapr_1<-datapr[,-1:-50]
datapr_2<-datapr_1[,-1:-9]
datapr_3<-datapr_2[,-4:-100]
datapr_4<-datapr_3[,-4:-36]
data<-datapr_4[70081:455808,]
databaru=matrix(nrow=4018,ncol=9)
for(i in 1:4018){
  databaru[i,]=as.matrix(data[(96*i-52):(96*i-50),])}
databaru<-write.csv(databaru, "C:/Users/Os 8.1/Desktop/datasfcwind.csv")

#presipitasi
datapr<-read.csv("C:/Users/Os 8.1/Desktop/pr.csv", header=FALSE)
datapr_1<-datapr[,-1:-50]
datapr_2<-datapr_1[,-1:-9]
datapr_3<-datapr_2[,-4:-100]
datapr_4<-datapr_3[,-4:-36]
data<-datapr_4[70081:455808,]
databaru=matrix(nrow=4018,ncol=9)
for(i in 1:4018){
  databaru[i,]=as.matrix(data[(96*i-52):(96*i-50),])}
databaru2<-write.csv(databaru, "C:/Users/Os 8.1/Desktop/datapr.csv")

#Suhu
datapr<-read.csv("C:/Users/Os 8.1/Desktop/tas.csv", header=FALSE)
datapr_1<-datapr[,-1:-50]
datapr_2<-datapr_1[,-1:-9]
datapr_3<-datapr_2[,-4:-100]
datapr_4<-datapr_3[,-4:-36]
data<-datapr_4[70081:455808,]
databaru=matrix(nrow=4018,ncol=9)
for(i in 1:4018){
  databaru[i,]=as.matrix(data[(96*i-52):(96*i-50),])}
databaru3<-write.csv(databaru, "C:/Users/Os 8.1/Desktop/datatas.csv")

#kelembapan
datapr<-read.csv("C:/Users/Os 8.1/Desktop/huss.csv", header=FALSE)
datapr_1<-datapr[,-1:-50]
```

**Lampiran 2** Syntax Cropping Grid Data GCM (lanjutan)

```
datapr_2<-datapr_1[,-1:-9]
datapr_3<-datapr_2[,-4:-100]
datapr_4<-datapr_3[,-4:-36]
data<-datapr_4[70081:455808,]
databaru=matrix(nrow=4018,ncol=9)
for(i in 1:4018){
  databaru[i,]=as.matrix(data[(96*i-52):(96*i-50),])
}
databaru4<-write.csv(databaru, "C:/Users/Os 8.1/Desktop/datahuss.csv")
```

**Lampiran 3** *Syntax* Pemodelan Curah Hujan pada Grid 1x1

```
#### model Keputih####  
library(DescTools)  
library(readr)  
library(car)  
library(DataCombine)  
library(ggplot2)  
data <- read_csv("C:/Users/Os 8.1/Desktop/data.csv")  
data[3:8][data[3:8]==0]<-0.00001  
train=data[1:3621,]  
data=train  
data$dummy=as.factor(data$dummy)  
####prediktor####  
x11 = data$PR1  
x12 = data$PR2  
x13 = data$PR3  
x14 = data$PR4  
x15 = data$PR5  
x16 = data$PR6  
x17 = data$PR7  
x18 = data$PR8  
x19 = data$PR9  
x21 = data$HUSS1  
x22 = data$HUSS2  
x23 = data$HUSS3  
x24 = data$HUSS4  
x25 = data$HUSS5  
x26 = data$HUSS6  
x27 = data$HUSS7  
x28 = data$HUSS8  
x29 = data$HUSS9  
x31 = data$TAS1  
x32 = data$TAS2  
x33 = data$TAS3  
x34 = data$TAS4  
x35 = data$TAS5  
x36 = data$TAS6  
x37 = data$TAS7  
x38 = data$TAS8  
x39 = data$TAS9
```

**Lampiran 4** *Syntax* Pemodelan Curah Hujan pada Grid 1x1 (lanjutan)

```

x41 = data$SFCWIND1
x42 = data$SFCWIND2
x43 = data$SFCWIND3
x44 = data$SFCWIND4
x45 = data$SFCWIND5
x46 = data$SFCWIND6
x47 = data$SFCWIND7
x48 = data$SFCWIND8
x49 = data$SFCWIND9
#respon
Y1= data$CH_Keputih
Y2 = data$CH_Cowek
Y3 = data$CH_Gubeng
Y4 = data$CH_Wonokromo
Y5 = data$CH_Wonorejo
Y6 = data$CH_GnSari
dummy=data$dummy
####model####
library(car)
model= glm(Y1~x15,family =Gamma(link = "inverse") ,data = data)
model1= glm(Y1~x15,family =Gamma(link = "identity") ,data = data)
model2= glm(Y1~x15,family =Gamma(link = "'sqrt'") ,data = data)
summary(model)
summary(model1)
summary(model2)
model= glm(Y1~x25,family =Gamma(link = "inverse") ,data = data)
model1= glm(Y1~x25,family =Gamma(link = "identity") ,data = data)
model2= glm(Y1~x25,family =Gamma(link = "'sqrt'") ,data = data)
summary(model)
summary(model1)
summary(model2)
model= glm(Y1~x35,family =Gamma(link = "inverse") ,data = data)
model1= glm(Y1~x35,family =Gamma(link = "identity") ,data = data)
model2= glm(Y1~x35,family =Gamma(link = "'sqrt'") ,data = data)
summary(model)
summary(model1)
summary(model2)
model= glm(Y1~x45,family =Gamma(link = "inverse") ,data = data)
model1= glm(Y1~x45,family =Gamma(link = "identity") ,data = data)

```

**Lampiran 5** *Syntax* Pemodelan Curah Hujan pada Grid 1x1 (lanjutan)

```

model2= glm(Y1~x45,family =Gamma(link = ""sqrt""),data = data)
summary(model)
summary(model1)
summary(model2)
model= glm(Y1~x15+x25+x35+x45,family =Gamma(link = "inverse") ,data =
data)
model1= glm(Y1~x15+x25+x35+x45,family =Gamma(link = "identity") ,data =
data)
model2= glm(Y1~x15+x25+x35+x45,family =Gamma(link = ""sqrt""),data =
data)
summary(model)
summary(model1)
summary(model2)
vif(model)
vif(model1)
vif(model2)
####model dummy####
model= glm(Y1~dummy+x15+x35,family =Gamma(link = "inverse") ,data =
data)
model1= glm(Y1~dummy+x15+x35,,family =Gamma(link = "identity") ,data =
data)
model2= glm(Y1~dummy+x15+x35,family =Gamma(link = ""sqrt""),data =
data)
summary(model)
summary(model1)
summary(model2)
model= glm(Y1~dummy+x15+x25+x35+x45,family =Gamma(link = "inverse")
,data = data)
model1= glm(Y1~dummy+x15+x25+x35+x45,family =Gamma(link =
"identity") ,data = data)
model2= glm(Y1~dummy+x15+x25+x35+x45,family =Gamma(link = ""sqrt""
),data = data)
summary(model)
summary(model1)
summary(model2)
win.graph()
par(mfrow=c(1,3))
inverse=model$residuals
identity=model1$residuals

```

**Lampiran 6** *Syntax* Pemodelan Curah Hujan pada Grid 1x1 (lanjutan)

```

“sqrt”=model2$residuals
pacf(inverse,lag.max = 14)
pacf(identity,lag.max = 14)
pacf(“sqrt”,lag.max = 14)
#### model lag####
#LAG 1
data1=cbind.data.frame(data$CH_Keputih,data$dummy,x15,x25,x35,x45)
colnames(data1)=c("CH_Keputih","dummy","x15","x25","x35","x45")
data1=slide(data1, Var = "CH_Keputih", slideBy = -c(1))
colnames(data1)=c("Y1","dummy","x15","x25","x35","x45","lag1")
head(data1)
data1=data1[2:3621,]
model= glm(Y1~.,family =Gamma(link = "inverse") ,data = data1)
model1= glm(Y1~.,family =Gamma(link = "identity") ,data = data1)
model2= glm(Y1~.,family =Gamma(link = ""sqrt"" ) ,data = data1)
summary(model)
summary(model1)
summary(model2)
win.graph()
par(mfrow=c(2,2))
plot(model2)
#LAG 2
data1=cbind.data.frame(data$CH_Keputih,x15,x25,x35,x45,data$dummy)
colnames(data1)=c("CH_Keputih","x15","x25","x35","x45","dummy")
data1=slide(data1, Var = "CH_Keputih", slideBy = -c(1,2))
colnames(data1)=c("Y1","x15","x25","x35","x45","dummy","lag1","lag2")
head(data1)
data1=data1[3:3621,]
model= glm(Y1~.,family =Gamma(link = "inverse") ,data = data1)
model1= glm(Y1~.,family =Gamma(link = "identity") ,data = data1)
model2= glm(Y1~.,family =Gamma(link = ""sqrt"" ) ,data = data1)
summary(model)
summary(model1)
summary(model2)

```



**Lampiran 7** *Syntax* Pemodelan Curah Hujan pada Grid 3x3

```

#grid keputih 3x3
#grid per variabel
data <- read_csv("C:/Users/Os 8.1/Desktop/data.csv")
data$dummy=as.factor(data$dummy)
data[3:8][data[3:8]==0]<-0.00001
train=data[1:3621,]
data=train
Y1= data$CH_Keputih
Y2 = data$CH_Cowek
Y3 = data$CH_Gubeng
Y4 = data$CH_Wonokromo
Y5 = data$CH_Wonorejo
Y6 = data$CH_GnSari
dummy=data$dummy
data.pca <- prcomp(data[,9:17])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc11=grid$PC1
pc12=grid$PC2
data.pca <- prcomp(data[,18:26])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc31=grid$PC1
pc32=grid$PC2
data.pca <- prcomp(data[,27:35])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc21=grid$PC1
pc22=grid$PC2
data.pca <- prcomp(data[,36:44])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc41=grid$PC1
pc42=grid$PC2
####model grid####

```

**Lampiran 8** *Syntax* Pemodelan Curah Hujan pada Grid 3x3 (Lanjutan)

```

data1=cbind.data.frame(dummy,pc11,pc12,pc21,pc22,pc31,pc32,pc41,pc42)
model= glm(Y1~.,family =Gamma(link = "inverse") ,data = data1)
model1= glm(Y1~.,family =Gamma(link = "identity") ,data = data1)
model2= glm(Y1~.,family =Gamma(link = ""sqrt"" ) ,data = data1)
summary(model)
summary(model1)
summary(model2)
vif(model)
vif(model1)
vif(model2)
win.graph()
par(mfrow=c(1,3))
inverse=model$residuals
identity=model1$residuals
"sqrt"=model2$residuals
pacf(inverse,lag.max = 14)
pacf(identity,lag.max = 14)
pacf("sqrt",lag.max = 14)
#lag 1 grid 3
data1=cbind.data.frame(data$CH_Keputih,dummy,pc11,pc12,pc21,pc22,pc31,pc32,pc41,pc42)
colnames(data1)=c("CH_Keputih","pc11","pc12","pc21","pc22","pc31","pc32","pc41","pc42","dummy")
data1=slide(data1, Var = "CH_Keputih", slideBy = -c(1))
colnames(data1)=c("Y1","dummy","pc11","pc12","pc21","pc22","pc31","pc32","pc41","pc42","lag1")
head(data1)
data1=data1[2:3621,]
model= glm(Y1~.,family =Gamma(link = "inverse") ,data = data1)
model1= glm(Y1~.,family =Gamma(link = "identity") ,data = data1)
model2= glm(Y1~.,family =Gamma(link = ""sqrt"" ) ,data = data1)
summary(model)
summary(model1)
summary(model2)
vif(model)
vif(model1)
vif(model2)
win.graph()
par(mfrow=c(2,2))

```

**Lampiran 9** *Syntax* Pemodelan Curah Hujan pada Grid 1x1 (lanjutan)

```

plot(model2)
#lag 2 grid 3
data1=cbind.data.frame(data$CH_Keputih,dummy,pc11,pc12,pc21,pc22,pc31,pc32,pc41,pc42)
colnames(data1)=c("CH_Keputih","pc11","pc12","pc21","pc22","pc31","pc32","pc41","pc42","dummy")
data1=slide(data1, Var = "CH_Keputih", slideBy = -c(1,2))
colnames(data1)=c("Y1","dummy","pc11","pc12","pc21","pc22","pc31","pc32","pc41","pc42","lag1","lag2")
head(data1)
data1=data1[3:3621,]
model= glm(Y1~.,family =Gamma(link = "inverse"),data = data1)
model1= glm(Y1~.,family =Gamma(link = "identity"),data = data1)
model2= glm(Y1~.,family =Gamma(link = "sqrt"),data = data1)
summary(model)
summary(model1)
summary(model2)
vif(model)
vif(model1)
vif(model2)

```

**Lampiran 10** *Syntax* Prediksi Data Testing pada Grid 1x1

```

#testing
#prediksi lag
data <- read_csv("C:/Users/Os 8.1/Desktop/data.csv")
data$dummy=as.factor(data$dummy)
data[3:8][data[3:8]==0]<-0.00001
lag1=data$CH_Keputih[3621]
data=data[3622:3652,]
x15=data$PR5
x25=data$HUSS5
x35=data$TAS5
x45=data$SFCWIND5
Y1=data$CH_Keputih
dummy=data$dummy
baru = cbind.data.frame(Y1,x15,x25,x35,x45,dummy)
datatest=cbind.data.frame(baru[1,],lag1)
pred2 = matrix(ncol=1,nrow=31)
test2 = predict(model2, datatest)
pred2[1,]= test2
for (i in 2:31){
  Model = cbind(baru[i,],test2)
  datatest = Model
  test2 = predict(model2, datatest)
  pred2[i,]= test2
}
data=cbind.data.frame(Y1,dummy,x15,x25,x35,x45,pred2)
colnames(data)=c("Y1", "dummy", "x15", "x25", "x35", "x45", "lag1")
test2=predict(model2,data)
databaru<-write.csv(test2, "C:/Users/Os 8.1/Desktop/PrediksiKeputih.csv")
win.graph()
data <- read_csv("C:/Users/Os 8.1/Desktop/data.csv")
data$dummy=as.factor(data$dummy)
data[3:8][data[3:8]==0]<-0.00001
ts.plot(data$CH_Keputih[3622:3652])
yhat2=test2
points(test2,col="red")
text(15,12,font = 2, adj = 0, cex = 5,
      paste0("RMSE=", round(RMSE(test2,Y1),3)))

```

**Lampiran 11** *Syntax* Prediksi Data Testing pada Grid 3x3

```

#prediksi grid
data <- read_csv("C:/Users/Os 8.1/Desktop/data.csv")
data$dummy=as.factor(data$dummy)
data[3:8][data[3:8]==0]<-0.00001
lag1=data$CH_Keputih[3621]
data=data[3622:3652,]
Y1=data$CH_Keputih
dummy=data$dummy
data.pca <- prcomp(data[,9:17])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc11=grid$PC1
pc12=grid$PC2
data.pca <- prcomp(data[,18:26])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc31=grid$PC1
pc32=grid$PC2
data.pca <- prcomp(data[,27:35])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc21=grid$PC1
pc22=grid$PC2
data.pca <- prcomp(data[,36:44])
summary(data.pca)
data.pca
grid=as.data.frame(data.pca$x)
pc41=grid$PC1
pc42=grid$PC2
baru = cbind.data.frame(Y1,pc11,pc12,pc21,pc22,pc31,pc32,pc41,pc42,dummy)
datatest=cbind.data.frame(baru[1,],lag1)
pred2 = matrix(ncol=1,nrow=31)
test2 = predict(model2, datatest)
pred2[1,]= test2
for (i in 2:31){
  Model = cbind(baru[i,],test2)

```

**Lampiran 12** *Syntax* Prediksi Data Testing pada Grid 3x3 (lanjutan)

```

datatest = Model
  test2 = predict(model2, datatest)
  pred2[i,]= test2
}
data=cbind.data.frame(Y1,dummy,pc11,pc12,pc21,pc22,pc31,pc32,pc41,pc42,p
red2)
colnames(data)=c("Y1","dummy","pc11","pc12","pc21","pc22","pc31","pc32","
pc41","pc42","lag1")
test2=predict(model2,data)
databaru<-write.csv(test2, "C:/Users/Os 8.1/Desktop/PrediksiKeputihGrid3.csv")
win.graph()
data <- read_csv("C:/Users/Os 8.1/Desktop/data.csv")
data$dummy=as.factor(data$dummy)
data[3:8][data[3:8]==0]<-0.00001
ts.plot(data$CH_Keputih[3622:3652])
yhat2=test2
points(test2,col="red")
text(15,12,font = 2, adj = 0, cex = 5,
  paste0("RMSE=", round(RMSE(test2,Y1),3)))

```

**Lampiran 13** *Syntax* Output Pemodelan

```
####Keputih####
library(DescTools)
win.graph()
par(mfrow=c(1,3))
ts.plot(data$CH_Keputih)
yhat=model$fitted.values
points(yhat,col="magenta")
text(110,160,font = 2, adj = 0, cex = 5,
     paste0("RMSE=", round(RMSE(model$fitted.values,Y1),3)))
ts.plot(data$CH_Keputih)
yhat1=model1$fitted.values
points(yhat1,col="blue")
text(110,160,font = 2, adj = 0, cex = 5,
     paste0("RMSE=", round(RMSE(model1$fitted.values,Y1),3)))
ts.plot(data$CH_Keputih)
yhat2=model2$fitted.values
points(yhat2,col="red")
summary(yhat2)
text(110,160,font = 2, adj = 0, cex = 5,
     paste0("RMSE=", round(RMSE(model2$fitted.values,Y1),3)))
```

**Lampiran 14 Surat Keterangan Pengambil Data Curah Hujan****SURAT KETERANGAN**

Saya yang bertanda tangan di bawah ini menerangkan bahwa :

1. Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

Nama : Risdia Ikfina Putri

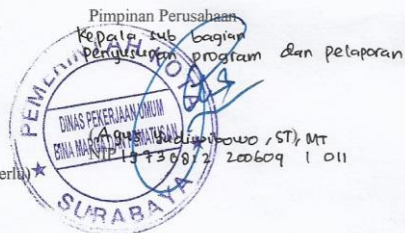
NRP : 0621154000065

Telah mengambil data di instansi/perusahaan kami :

Nama Instansi : Dinas PU Bina Marga dan Pematusan

Divisi/ bagian : *Penyusunan Program dan Pelaporan*  
 sejak tanggal *1 Maret 2019* sampai dengan *30 April 2019* untuk  
 keperluan Tugas Akhir/ ~~Thesis~~ Semester Gasal/Genap\* 20 / 20 .

2. Tidak Keberatan/Keberatan\* nama perusahaan dicantumkan dalam Tugas Akhir/ ~~Thesis~~ mahasiswa Statistika yang akan di simpan di Perpustakaan ITS dan dibaca di lingkungan ITS.
3. Tidak Keberatan/Keberatan\* bahwa hasil analisis data dari perusahaan dipublikasikan dalam E journal ITS yaitu Jurnal Sains dan Seni ITS.



\*(coret yang tidak perlu)



**Lampiran 15** Surat Keterangan Pengambil Data Iklim**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS:

Nama : Risda Ikfina Putri

NRP : 0621154000065

menyatakan bahwa data yang digunakan dalam Tugas Akhir/ Thesis ini merupakan data sekunder yang diambil dari penelitian / buku / Tugas Akhir/ Thesis/ publikasi / lainnya yaitu:

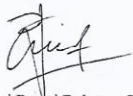
Sumber : <https://cera-www.dkrz.de/WDCC/ui/cerasearch/>

Keterangan : Data iklim luaran GCM model CSIRO-Mk3 (Presipitasi, Temperatur, Kelembaban Spesifik, dan Kecepatan Angin) dari tahun 1 Januari 2009 sampai dengan 31 Desember 2018.

Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui  
Pembimbing Tugas Akhir

Surabaya,



(Santi Puteri Rahayu, S.Si., M.Si., Ph.D.)  
NIP. 1978 0406 200112



(Risda Ikfina Putri)  
NRP. 0621154000065

\*(coret yang tidak perlu)

*(Halaman ini sengaja dikosongkan)*

## BIODATA PENULIS



Penulis dengan nama lengkap Risda Ikfina Putri dilahirkan di Kabupaten Blitar pada 20 Mei 1997. Penulis menempuh pendidikan formal di SDN Bendosari I selama 4 tahun dan SDN Pakunden 2 selama 2 tahun, dilanjutkan menempuh pendidikan di SMP Negeri 1 Blitar selama 3 tahun, dan di SMA Negeri 1 Blitar selama 3 tahun. Kemudian penulis diterima sebagai Mahasiswa Departemen Statistika ITS melalui jalur SNMPTN pada tahun 2015. Selama masa perkuliahan, penulis pernah berkesempatan menjadi pengajar SISI HIMASTA ITS periode 2016/2017 di kampung-kampung binaan di sekitar Institut Teknologi Sepuluh Nopember. Selama menjadi pengajar SISI, penulis mengajar dari mulai jenjang SD sampai dengan SMP. Selain itu penulis juga pernah berkesempatan menjadi pengajar ABC yang diadakan oleh Jamaah Manarul Ilmi ITS (JMII). Selama menjadi pengajar banyak hal yang bisa dipelajari oleh penulis seperti bagaimana menjadi pengajar yang baik, melatih kesabaran, dan bagaimana cara menghadapi anak-anak dengan berbagai macam perilaku yang berbeda. Penulis juga pernah menjadi panitia pada acara STATION sebagai Sie Acara. Selain itu, Penulis juga pernah menjadi panitia pada acara LKMM pra-td yang diselenggarakan oleh BEM FMKSD ITS. Apabila pembaca ingin memberikan kritik dan saran terkait Laporan Tugas Akhir ini, dapat menghubungi penulis melalui email [risdaikfina20@gmail.com](mailto:risdaikfina20@gmail.com) atau melalui nomor 082334211820.

*(Halaman ini sengaja dikosongkan)*