



TESIS - SS142501

SPATIAL BAYESIAN MODEL AVERAGING
UNTUK PRAKIRAAN CUACA JANGKA PENDEK
TERKALIBRASI

NISWATUL QONA'AH
NRP. 06211650010031

DOSEN PEMBIMBING
Dr. Sutikno, M.Si
Dr. Purhadi, M.Sc

PROGRAM MAGISTER
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018



TESIS - SS142501

SPATIAL BAYESIAN MODEL AVERAGING
UNTUK PRAKIRAAN CUACA JANGKA PENDEK
TERKALIBRASI

NISWATUL QONA'AH
NRP. 06211650010031

DOSEN PEMBIMBING
Dr. Sutikno, M.Si
Dr. Purhadi, M.Sc

PROGRAM MAGISTER
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018



THESIS - SS142501

SPATIAL BAYESIAN MODEL AVERAGING TO CALIBRATE SHORT-RANGE WEATHER FORECAST

NISWATUL QONA'AH
NRP. 06211650010031

SUPERVISORS

Dr. Sutikno, M.Si

Dr. Purhadi, M.Sc

MAGISTER PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2018

**SPATIAL BAYESIAN MODEL AVERAGING UNTUK PRAKIRAAN
CUACA JANGKA PENDEK TERKALIBRASI**

**Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar
Magister Sains (M.Si)**

di

Institut Teknologi Sepuluh Nopember

oleh :

NISWATUL QONA'AH

NRP. 06211650010031

Tanggal Ujian : 18 Juli 2018

Periode Wisuda : September 2018

Disetujui oleh:


1. **Dr. Sutikno, M.Si.**

NIP: 19710313 199702 1 001

(Pembimbing I)


2. **Dr. Purhadi, M.Sc.**

NIP: 19620204 198701 1 001

(Pembimbing II)


3. **Santi Wulan Purnami, M.Si, P.hD.**

NIP: 19720923 199803 2 001

(Penguji)


4. **Prof. Drs. NUR Iriawan, M.Ikomp, P.hD.**

NIP: 19621015 198803 1 002

(Penguji)

**Dekan
Fakultas Matematika, Komputasi dan Sains Data
Institut Teknologi Sepuluh Nopember**


Prof. Dr. Basuki Widodo, M.Sc

NIP: 19650605 198903 1 002

SPATIAL BAYESIAN MODEL AVERAGING **UNTUK PRAKIRAAN CUACA JANGKA PENDEK** **TERKALIBRASI**

Nama Mahasiswa : Niswatul Qona'ah
NRP : 06211650010031
Dosen Pembimbing : Dr. Sutikno, M.Si
Dr. Purhadi, M.Sc

ABSTRAK

Hasil prakiraan cuaca menggunakan *ensemble* seringkali masih bersifat *underdispersive* dan tidak terkalibrasi. *Bayesian Model Averaging* (BMA) adalah metode *postprocessing* statistik untuk mengkalibrasi prakiraan *ensemble* dan membuat interval prediktif yang lebih reliabel. BMA menghasilkan ramalan terkalibrasi, tetapi, tidak mempertimbangkan korelasi spasial. Tidak seperti BMA, *Geostatistical Output Perturbation* (GOP) mempertimbangkan korelasi spasial di antara beberapa lokasi secara serentak. Metode ini memiliki parameter spasial yang memodifikasi output prakiraan untuk menangkap informasi spasial. Namun, GOP hanya berlaku untuk perkiraan deterministik tunggal. Spasial Bayesian Model Averaging (Spatial BMA) adalah metode yang menggabungkan BMA dan GOP. Metode ini memperluas BMA untuk menghasilkan perkiraan probabilistik yang dikalibrasi untuk beberapa lokasi secara bersamaan. Anggota *ensemble* Spatial BMA diperoleh dengan memakaikan ramalan cuaca dengan hasil simulasi eror spasial, dalam proporsi yang sesuai dengan bobot BMA. Suhu menjadi fokus penelitian ini karena cenderung memiliki korelasi yang relatif kuat dengan unsur-unsur lainnya. Analisis yang diterapkan untuk mengkalibrasi ramalan suhu di delapan lokasi meteorologi di Jabodetabek, Indonesia adalah BMA Spasial. Awalnya, Partial Least Square (PLS), Principal Component Regression (PCR), dan metode Regression Ridge menjadi anggota *ensemble* BMA. Prakiraan suhu BMA simultan digunakan untuk mendapatkan simulasi kesalahan yang secara spasial berkorelasi yang memodifikasi prakiraan suhu. Untuk masa pelatihan selama 30 hari, prakiraan suhu Spatial BMA di 8 stasiun meteorologi tergolong cukup baik. Metode ini mampu mengkalibrasi ramalan suhu lebih baik daripada *raw ensemble* yang nilai *coverage*-nya mendekati standar 50% tetapi tidak lebih baik daripada menggunakan BMA asli. Berdasarkan beberapa indikator ramalan cuaca, seperti RMSE dan CRPS, dalam hal ini BMA asli lebih baik dari Spatial BMA dalam hal akurasi dan presisi. Namun, jika dibandingkan dengan GOP, prakiraan suhu udara menggunakan Spatial BMA menunjukkan akurasi dan presisi yang lebih tinggi.

Kata Kunci : BMA, *Ensemble*, GOP, Spatial BMA

(Halaman ini sengaja dikosongkan)

SPATIAL BAYESIAN MODEL AVERAGING TO CALIBRATE SHORT-RANGE WEATHER FORECAST

Name : Niswatul Qona'ah
Registration Number : 06211650010031
Supervisors : Dr. Sutikno, M.Si
Dr. Purhadi, M.Sc

ABSTRACT

The output of weather forecasts using ensemble are often underdispersive and uncalibrated. Bayesian Model Averaging (BMA) is a statistical postprocessing method to calibrate the ensemble forecasts and create more reliable predictive interval. BMA producing the calibrated forecast, but, does not consider spatial correlation. Unlike BMA, Geostatistical Output Perturbation (GOP) consider spatial correlation among several locations altogether. It has spatial parameters that modifying the forecast output to capture spatial information. But, GOP only applies to single deterministic forecasts. Spatial Bayesian Model Averaging (Spatial BMA) is a method which combines BMA and GOP. It extends BMA to produce calibrated probabilistic forecasts for several location simultaneously. The members of the spatial BMA ensemble are obtained by dressing the weather forecasts with simulated spatially correlated error, in proportions that correspond to the BMA weights. Temperature becomes the focus of this study because it tends to have a relatively strong correlation with the other elements. Analysis applied to calibrate the temperature forecast at eight meteorological sites in Jakarta, Indonesia is Spatial BMA. Initially, Partial Least Square (PLS), Principal Component Regression (PCR), and Regression Ridge methods become ensemble members of BMA. Temperature forecasts of BMA are used to obtain simulated spatially correlated error that modifying temperature forecasts. For training period over 30 days, Spatial BMA temperature forecasts at 8 meteorological stations belongs to good enough. It is able to calibrate the temperature forecast better than raw ensemble which coverage comes closer to the standard 50% but it is not better than using original BMA. Based on some indicators of weather forecast, such as RMSE and CRPS, in this case original BMA is better than Spatial BMA in terms of accuracy and precision.

Keywords : BMA, Ensemble, GOP, Spatial BMA

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Segala puji bagi Allah yang telah memberikan petunjuk, kemudahan, kelancaran, dan kekuatan kepada penulis, sehingga dapat menyelesaikan buku Tesis dengan judul “***SPATIAL BAYESIAN MODEL AVERAGING UNTUK PRAKIRAAN CUACA JANGKA PENDEK TERKALIBRASI***” dengan baik.

Penulis menyadari bahwa dalam penyusunan Tesis ini tidak terlepas dari bantuan dan dukungan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih banyak kepada :

1. Bapak Dr. Sutikno, M.Si dan Bapak Dr. Purhadi, M.Sc selaku dosen pembimbing yang telah membagikan ilmu, nasehat, saran, dan waktu yang diluangkan untuk membimbing penulis dalam menyusun Tesis.
2. Ibu Santi Wulan P, M.Si., Ph.D dan Bapak Prof. Nur Iriawan, MIkomp., Ph.D selaku dosen penguji yang telah memberikan saran dan kritik yang membangun dalam penyusunan Tesis.
3. Bapak Dr. Suhartono, M.Sc selaku Kepala Departemen Statistika ITS, seluruh tenaga pengajar yang telah membagikan ilmu yang bermanfaat pada penulis, dan seluruh karyawan Tata Usaha dan Ruang Baca Statistika.
4. Bapak Dr.rer.pol. Heri Kuswanto, M.Si selaku Ketua Program Studi Pascasarjana Departemen Statistika ITS.
5. Ibu, Bapak, Kakak, Adik, serta segenap keluarga atas do'a yang selalu terucap nama penulis di dalamnya agar dapat meraih keberhasilan dan kesuksesan. Terimakasih pula atas motivasi dan inspirasinya. Ini hanya hadiah kecil dari penulis sebagai salah satu bukti bakti penulis.
6. Teman-teman seperjuangan di Pascasarjana S2 Statistika 2016, yang selama kuliah bersama-sama dengan penulis. Susah, senang, serta jatuh bangunnya selama perjuangan penulis hingga empat semester berada di tengah-tengah kalian.

7. Mas Luthfi dan Mbak Kartika sebagai partner diskusi terbaik yang telah turut membantu dan membimbing penulis dalam penyusunan tesis ini.
8. Yayasan Ukhuwah Islamiyah yang telah memberikan kepercayaan kepada penulis untuk turut berkarya serta menjadi bagian dari pengajar dan pendidik di SMA IT Al Uswah, sehingga penulis bisa merasakan profesi yang dicita-citakan sejak kecil.
9. Ustadz/ustadzah beserta siswa/siswi SMA IT Al-Uswah yang turut selalu memotivasi penulis untuk menyelesaikan tesis ini. Keluarga ini membuat warna tersendiri dalam kanvas kehidupan penulis selama kuliah di Pascasarjana S2 Statistika ITS.
10. Sahabat-sahabat fii sabilillah (Keluarga Fatimah KSI, Mutiara 4, KQ Az-Zahra, Tajul Karomah) yang selalu membersamai, menghibur, dan menyemangati.
11. Pihak-pihak lain yang telah mendukung, diantaranya BMKG dan Kemenristekdikti atas terselesaikannya Tesis ini.

Mengingat keterbatasan penulis, maka penulis merasa masih banyak kekurangan, baik pada teknis penulisan maupun materi. Oleh karena itu, penulis mengharapkan kritik dan saran dari pembaca agar ke depannya lebih baik. Semoga buku Tesis ini dapat bermanfaat bagi kita semua. Amin.

Surabaya, Juli 2018

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN	i
ABSTRAK	iii
ABSTRACT	v
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xv
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	6
1.4 Manfaat Penelitian	6
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	7
2.1 Distribusi Normal	7
2.1.1 Distribusi Normal Univariat	7
2.1.2 Distribusi Normal Multivariat	7
2.2 <i>Principal Component Analysis</i> (PCA)	8
2.3 <i>Partial Least Squares</i> (PLS)	11
2.4 <i>Principal Component Regression</i> (PCR)	13
2.5 Regresi Ridge	15
2.6 <i>Bayesian Model Averaging</i> (BMA)	16
2.7 <i>Geostatistical Output Perturbation</i> (GOP)	20
2.8 <i>Spatial Bayesian Model Averaging</i> (Spatial BMA)	23
2.9 Evaluasi Keباikan Model Kalibrasi	26
2.9.1 <i>Root Mean Square Error</i> (RMSE)	26

2.9.2	<i>Coverage</i>	27
2.9.3	<i>Continuous Rank Probability Score (CRPS)</i>	27
2.9	<i>Numerical Weather Prediction (NWP)</i>	29
2.10	Sistem Prediksi <i>Ensemble</i>	32
2.11	Suhu / Temperatur	34
BAB III	METODE PENELITIAN	35
3.1	Sumber Data	35
3.2	Variabel Penelitian	36
3.3	Kajian Teoritis	40
3.4	Kajian Aplikasi	42
BAB IV	HASIL DAN PEMBAHASAN	45
4.1	Penaksiran Parameter Spasial BMA	45
4.1.1	Estimasi Parameter Regresi	45
4.1.2	Estimasi Parameter Bobot dan Varians	49
4.1.3	Estimasi Parameter Kovarians Spasial	55
4.2	Aplikasi Spasial BMA untuk Prakiraan Cuaca Jangka Pendek Terkalibrasi di 8 Stasiun Meteorologi Jabodetabek	61
4.2.1	Analisis Deskripsi Suhu Udara dan Standardisasi	61
4.2.2	Pra-pemrosesan Parameter NWP dengan PCA	62
4.2.3	Pembentukan Model Prakiraan Anggota <i>Ensemble</i> untuk Suhu Maksimum dan Suhu Minimum	65
4.2.3.1	Regresi PLS untuk Suhu Maksimum dan Suhu Minimum	65
4.2.3.2	Regresi PCR untuk Suhu Maksimum dan Suhu Minimum	70
4.2.3.3	Regresi Ridge untuk Suhu Maksimum dan Suhu Minimum	74
4.2.4	Deskripsi Prakiraan Cuaca Anggota <i>Ensemble</i>	77

4.2.5 Kalibrasi Prakiraan Cuaca	
Menggunakan Spatial BMA	79
4.2.6 Evaluasi Prakiraan Spatial BMA dengan BMA	
dan GOP	90
BAB V KESIMPULAN DAN SARAN	93
5.1 Kesimpulan	93
5.2 Saran	94
DAFTAR PUSTAKA	95
LAMPIRAN.....	97
BIOGRAFI PENULIS	139

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

Tabel 3.1	Parameter NWP CCAM yang Digunakan	36
Tabel 3.2	Struktur Data untuk Metode BMA pada Stasiun s.....	39
Tabel 3.3	Struktur Data untuk Metode GOP pada Stasiun s.....	40
Tabel 4.1	Nilai Eigen dan Kumulatif Varians (%) PC Variabel Presipitasi (rnd)	62
Tabel 4.2	Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Dermaga	64
Tabel 4.3	Nilai PRESS Model PLS Stamet Dermaga.....	66
Tabel 4.4	Banyak Komponen Optimal Model PLS pada Delapan Stamet.....	67
Tabel 4.5	Koefisien Regresi Model PLS Stamet Dermaga.....	68
Tabel 4.6	Model PLS untuk Prakiraan Cuaca Stamet Dermaga	69
Tabel 4.7	Kumulatif Varians 41 Komponen Prediktor Stamet Dermaga	71
Tabel 4.8	Jumlah Komponen Optimal PCR pada Delapan Stasiun	71
Tabel 4.9	Koefisien Regresi Model PCR Stamet Dermaga	72
Tabel 4.10	Model PCR untuk Prakiraan Cuaca Stamet Dermaga	73
Tabel 4.11	Koefisien Regresi Ridge Stamet Dermaga	75
Tabel 4.12	Model Regresi Ridge untuk Prakiraan Cuaca Stamet Dermaga	76
Tabel 4.13	Estimasi Parameter dan Hasil Prakiraan BMA Simultan Stamet Kemayoran, 8 Februari 2009	81
Tabel 4.14	Matriks Jarak 8 Stasiun Meteorologi (dalam km).....	84
Tabel 4.15	Signifikansi Moran's I pada Observasi Cuaca di 8 Stamet..	84
Tabel 4.16	Estimasi Parameter Regresi Model Spatial BMA.....	85
Tabel 4.17	RMSE Prakiraan Spatial BMA dan Simultan BMA, 2 Maret 2009	86
Tabel 4.18	CRPS Prakiraan Spatial BMA dan Simultan BMA	89
Tabel 4.19	Evaluasi Prakiraan NWP, <i>Raw Ensemble</i> , GOP, dan Spatial BMA	91

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 2.1	Semivariogram Pra-Simulasi GOP.....	23
Gambar 2.2	CDF Observasi dan Prakiraan pada Waktu ke- <i>i</i>	28
Gambar 2.3	Ilustrasi Kondisi Cuaca Dunia Nyata dan NWP	31
Gambar 2.4	Prediksi 2 Model <i>Ensemble</i> untuk Temperatur Wilayah Jerman 5 Januari 2011.....	33
Gambar 3.1	Lokasi Stasiun Meteorologi	35
Gambar 3.2	Proyeksi Pengukuran NWP.....	37
Gambar 3.3	Tahapan Analisis Data	44
Gambar 4.1	Observasi Suhu Maksimum dan Suhu Minimum di Stamet Dermaga.....	61
Gambar 4.2	Konvergensi Koefisien Regresi Ridge	74
Gambar 4.3	Tren Prakiraan Anggota <i>Ensemble</i> dan Observasi Suhu Maksimum	77
Gambar 4.4	Tren Prakiraan Anggota <i>Ensemble</i> dan Observasi Suhu Minimum	78
Gambar 4.5	VRH <i>Raw Ensemble</i> Stamet Dermaga, 1 Jan '09 – 30 Jan '09	80
Gambar 4.6	Tahapan Mendapatkan Prakiraan Suhu Menggunakan Spatial BMA.....	83
Gambar 4.7	Semivariogram Empiris Suhu Udara Maksimum dan Minimum.....	85
Gambar 4.8	VRH Prakiraan Spatial BMA Suhu Maksimum dan Minimum, Jumlah <i>Ensemble</i> =99.....	87
Gambar 4.9	VRH Prakiraan Spatial BMA Suhu Maksimum dan Minimum, Jumlah <i>Ensemble</i> =3.....	88
Gambar 4.10	Tren Prakiraan Spatial BMA, BMA, GOP, dan Observasi Suhu Udara Stamet Dermaga.....	90

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

Lampiran 1. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Cengkareng.....	99
Lampiran 2. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Citeko	99
Lampiran 3. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Curug.....	100
Lampiran 4. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Kemayoran	101
Lampiran 5. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Pondok Betung	101
Lampiran 6. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Priok.....	102
Lampiran 7. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Tangerang	103
Lampiran 8. Nilai PRESS untuk Stamet Cengkareng.....	103
Lampiran 9. Nilai PRESS untuk Stamet Citeko	104
Lampiran 10. Nilai PRESS untuk Stamet Curug	105
Lampiran 11. Nilai PRESS untuk Stamet Kemayoran	106
Lampiran 12. Nilai PRESS untuk Stamet Pondok Betung	106
Lampiran 13. Nilai PRESS untuk Stamet Priok	107
Lampiran 14. Nilai PRESS untuk Stamet Tangerang	108
Lampiran 15. Model Regresi PLS Stamet Dermaga.....	109
Lampiran 16. Prakiraan Suhu Maksimum Model PLS	110
Lampiran 17. Prakiraan Suhu Minimum Model PLS	111
Lampiran 18. Kumulatif Varians Komponen Prediktor Stamet Cengkareng	112
Lampiran 19. Kumulatif Varians Komponen Prediktor Stamet Citeko	113

Lampiran 20. Kumulatif Varians Komponen Prediktor Stamet Curug	113
Lampiran 21. Kumulatif Varians Komponen Prediktor Stamet Kemayoran.....	114
Lampiran 22. Kumulatif Varians Komponen Prediktor Stamet Pondok Betung.....	114
Lampiran 23. Kumulatif Varians Komponen Prediktor Stamet Priok.....	115
Lampiran 24. Kumulatif Varians Komponen Prediktor Stamet Tangerang	115
Lampiran 25. Model PCR untuk Prakiraan Cuaca Stamet Dermaga.....	116
Lampiran 26. Prakiraan Suhu Maksimum Model PCR	117
Lampiran 27. Prakiraan Suhu Minimum Model PCR	118
Lampiran 28. Konstanta λ Terpilih untuk Regresi Ridge	119
Lampiran 29. Model Ridge untuk Prakiraan Cuaca Stamet Dermaga....	120
Lampiran 30. Prakiraan Suhu Maksimum Model Ridge	121
Lampiran 31. Prakiraan Suhu Minimum Model Ridge	122
Lampiran 32. Estimasi Bobot dan Koefisien Regresi Model Simultan BMA, Suhu Maksimum	123
Lampiran 33. Estimasi Bobot dan Koefisien Regresi Model Simultan BMA, Suhu Minimum.....	124
Lampiran 34. Prakiraan Suhu Maksimum Model Spatial BMA	124
Lampiran 35. Prakiraan Suhu Minimum Model Spatial BMA.....	125
Lampiran 36. Syntax Program	126

BAB I

PENDAHULUAN

1.1 Latar Belakang

Cuaca merupakan nilai sesaat (aktual) dari keadaan atmosfer, serta perubahan dalam jangka pendek (kurang dari 1 jam hingga 24 jam) di suatu tempat tertentu di bumi. Kehidupan manusia sangat dipengaruhi oleh kondisi cuaca. Dalam bidang pertanian, kondisi cuaca berpengaruh terhadap penentuan masa tanam dan panen. Di bidang perikanan dan kelautan, kondisi cuaca berpengaruh bagi nelayan dalam menentukan waktu pulang dan pergi ke laut untuk penangkapan hasil laut. Kondisi cuaca berpengaruh pula dalam bidang transportasi. Kelancaran transportasi baik transportasi laut maupun udara sangat dipengaruhi oleh faktor-faktor cuaca seperti pola angin dan curah hujan. Kondisi cuaca juga sangat berpengaruh terhadap bidang kehidupan lainnya, seperti telekomunikasi, pariwisata, industri, bangunan, budaya, persebaran penduduk, dan kekayaan hayati. Hal ini mengakibatkan informasi mengenai prakiraan cuaca kepada masyarakat menjadi suatu hal yang sangat diperlukan.

Kondisi cuaca di Indonesia diumumkan untuk jangka waktu sekitar 24 jam melalui hasil analisis prakiraan cuaca oleh Badan Meteorologi, Klimatologi, dan Geofisika (BMKG). BMKG merupakan sebuah lembaga pemerintah yang salah satu tugasnya adalah melakukan pengamatan dan prediksi terhadap unsur cuaca. Metode prakiraan cuaca yang digunakan oleh BMKG sebelumnya masih mengandalkan kemampuan para *forecaster* sehingga cenderung lebih subjektif. Saat ini, BMKG telah berupaya mengembangkan pengolahan peramalan cuaca secara numerik, yaitu menggunakan model *Numerical Weather Prediction* (NWP).

Teknik model *Numerical Weather Prediction* (NWP) diharapkan dapat memberikan informasi kondisi cuaca dengan akurasi optimal. Model NWP merupakan sekumpulan kode komputer yang mempresentasikan persamaan atmosfer secara numerik, diolah sistem informasi tertentu dan menghasilkan luaran (*output*) yang bersifat deterministik. Namun, NWP pada umumnya memberikan informasi cuaca yang homogen pada beberapa daerah dan diukur pada skala global (Wilks, 2006). Sementara itu, Indonesia merupakan negara yang kaya akan vegetasi

dan memiliki topografi yang beragam, sehingga prakiraan cuaca yang dihasilkan menjadi bias dan tidak reliabel. Selain itu, luaran (*output*) NWP menjadi bias karena keadaan atmosfer yang tidak pasti dan terbatasnya hitungan matematik untuk memodelkan keadaan fisik dan dinamika atmosfer (Idowu dan Rautenbach, 2009).

Pemrosesan data secara statistik (*statistical post-processing*) terhadap luaran (*output*) NWP dengan menggunakan *ensemble* dapat dilakukan untuk meningkatkan akurasi prakiraan. *Ensemble* merupakan gabungan luaran NWP dari beberapa metode individual (regresi *stepwise*, regresi komponen utama, ARIMA, dan sebagainya) dan juga metode *hybrid* (ARIMA-ANN, Wavelet-ANFIS, dan sebagainya). Pendekatan *ensemble* juga dapat menerapkan spesifikasi model dengan syarat inisiasi yang berbeda disesuaikan dengan kondisi atmosfer atau menggunakan sejumlah luaran (*output*) NWP yang dikeluarkan oleh otoritas resmi.

Namun, prakiraan *ensemble* seringkali masih bersifat *underdispersive* (nilai prakiraan cuaca terpusat pada suatu nilai dengan varians yang rendah), khususnya pada prakiraan cuaca jangka pendek, sehingga diperlukan kalibrasi *ensemble* (Schmeits dan Kok, 2010). Beberapa metode untuk mengkalibrasi *ensemble* diantaranya regresi logistik (Wilks, 2006), *Bayesian Model Averaging* (BMA) (Raftery, Gneiting, Balabdoui, dan Polakowski, 2005), dan *Ensemble Model Output Statistics* (EMOS) (Gneiting, Raftery, Westveld III, dan Goldman, 2005). Metode-metode tersebut memperoleh fungsi kepadatan peluang (PDF) untuk prakiraan cuaca dengan memadukan seluruh anggota *ensemble* (Devin., Clarke, Roulston, Ziehmann, Brocker dan Smith, 2004). Beberapa metode tersebut disesuaikan dengan karakteristik komponen cuaca yang ingin dikalibrasi apakah mengikuti distribusi normal, Gamma, dan sebagainya.

Raftery *et al.* (2005) mengusulkan metode BMA untuk mengkalibrasi ramalan *ensemble*. BMA adalah metode yang memadukan prakiraan seluruh anggota *ensemble* berdasarkan rata-rata terboboti, probabilitas posterior anggota *ensemble*, sehingga tidak mempertimbangkan kontribusi salah satu model saja seperti pada sebagian besar pemodelan statistika. Di bidang meteorologi, BMA merupakan metode yang paling banyak digunakan diantara metode kalibrasi *ensemble* lainnya, terutama setelah dikembangkan untuk model dinamis. BMA dapat diterapkan dengan beberapa teknik estimasi dan menunjukkan kinerja yang

memuaskan, baik pada data simulasi maupun data riil (Raftery dan Zheng 2003; Raftery *et al.*, 2005). Di Indonesia, penelitian sebelumnya yang menggunakan BMA untuk mengkalibrasi prakiraan *ensemble* jangka pendek diantaranya dilakukan oleh Kuswanto dan Sari (2013) serta Luthfi (2017) untuk prakiraan suhu udara.

Meskipun kinerja BMA cukup memuaskan, namun metode ini masih memiliki kelemahan. Tiap lokasi hanya dipertimbangkan secara individu (tidak serentak) dan korelasi spasial antar error diabaikan (Gel, Raftery, dan Gneiting, 2004). Akibatnya koreksi bias metode BMA belum optimal dan *coverage* prakiraan terkadang tidak mencapai presentase yang disyaratkan. Salah satu metode kalibrasi *ensemble* yang mempertimbangkan korelasi spasial antar error adalah *Geostatistical Output Perturbation* (GOP).

GOP adalah metode prakiraan cuaca yang dihasilkan hanya dari satu luaran deterministik saja, seperti NWP. Namun, GOP dapat memodifikasi luaran tersebut, sehingga mampu membangkitkan *ensemble* berukuran besar berdasarkan korelasi yang diidentifikasi dari error model (Gel *et al.*, 2004). Walaupun demikian, GOP belum banyak diaplikasikan di Indonesia, khususnya pada penelitian meteorologi. Hal ini disebabkan GOP membutuhkan data prakiraan NWP yang cukup besar dan melingkupi wilayah yang luas agar pengaruh cuaca antar lokasi dapat tertangkap dengan baik.

Intinya, metode BMA mempertimbangkan informasi *ensemble* tetapi mengabaikan korelasi spasial. Sementara metode GOP memperhitungkan dependensi spasial, namun hanya berlaku untuk prakiraan deterministik tunggal, bukan *ensemble* prakiraan cuaca, dan tidak dapat mempertimbangkan arah penyebaran dependensi yang berasal dari evolusi atmosfer yang nonlinear dan *ensemble* dinamis yang bersifat khas. *Spatial Bayesian Model Averaging* (Spatial BMA) mengatasi kekurangan ini dengan menggabungkan kedua metode (BMA dan GOP).

Seperti dalam metode BMA, PDF prediktif pada Spatial BMA adalah rata-rata terboboti dari PDF bersyarat yang terpusat pada koreksi bias dari model anggota *ensemble*, dengan bobot yang tergantung pada kontribusi masing-masing anggota. Pada model Spatial BMA, PDF bersyarat merupakan densitas multivariat

dengan struktur kovarians dibentuk untuk mempertimbangkan struktur spasial dari observasi cuaca. Sehingga, seperti pada BMA, parameter Spasial BMA dapat dibedakan menjadi 2 bagian, yaitu parameter regresi serta parameter bobot dan varians. Pada umumnya, parameter tersebut diestimasi menggunakan *Maximum Likelihood Estimation* (MLE). Namun, fungsi *joint likelihood* dari bobot dan varians tidak dapat diselesaikan secara analitik, sehingga digunakan algoritma *Expectation Maximum* (EM) untuk mendapatkan penaksir parameternya. Kemudian, estimator-estimator tersebut digunakan untuk membentuk PDF prediktif (Raftery *et al.*, 2005).

Selain itu, parameter model Spasial BMA juga harus menyesuaikan dengan parameter model GOP. Parameter kovarians spasial GOP diestimasi menggunakan *weighted least square* (WLS) yang diselesaikan secara iteratif dengan *Limited-Memory BFGS* (L-BFGS). Metode Spasial BMA dapat digunakan untuk menghasilkan *ensemble* statistik dari keseluruhan wilayah pengamatan cuaca secara simultan, dari berbagai ukuran, dan dengan biaya komputasi minimum. Pada lokasi secara individu, Spasial BMA dapat direduksi menjadi metode BMA (Berrocal *et al.*, 2007). Penelitian Feldmann *et al.* (2015) menyebutkan bahwa hasil prakiraan suhu di Jerman menggunakan Spasial BMA lebih reliabel daripada *Raw ensemble*, *Nonhomogeneous Gaussian Regression* (NGR), dan BMA.

Salah satu komponen cuaca yang seringkali menjadi fokus penelitian adalah suhu udara. Di sebagian besar Negara, suhu udara selalu dimuat dalam informasi prakiraan cuaca sebagai informasi pendamping. Selain itu, suhu udara memiliki keterkaitan dengan tekanan udara, angin dan kelembapan yang secara tidak langsung dapat berpengaruh terhadap siklus pembentukan awan hujan (Wilks, 2006).

Karakteristik prakiraan *ensemble* pada umumnya cenderung berpusat pada titik tertentu sehingga menyerupai distribusi normal. Namun, prakiraan tersebut seringkali memiliki varians yang rendah, sehingga kemungkinan untuk terjadi *underdispersive* cukup tinggi. Selain itu, kondisi cuaca di suatu lokasi dapat dipengaruhi oleh lokasi lain yang memiliki kedekatan. Hal ini disebabkan oleh pola atmosfer yang dinamis dan variatif. Untuk mengatasi hal tersebut, penelitian tesis

ini mengaplikasikan metode Spatial BMA yang merupakan kombinasi dari BMA dan GOP. Luthfi (2017) mengkalibrasi *ensemble* menggunakan BMA dan GOP dengan data NWP BMKG pada 8 stamet di Jabotabek, menyimpulkan bahwa akurasi dan presisi BMA lebih tinggi daripada GOP. Penelitian ini diharapkan juga dapat memperlihatkan perbandingan kinerja metode Spatial BMA dengan BMA dan GOP (berdasarkan penelitian sebelumnya) dalam mengoreksi bias prakiraan NWP.

1.2 Rumusan Masalah

BMA merupakan salah satu metode yang dapat digunakan untuk mengkalibrasi prakiraan *ensemble* yang lebih reliabel dan mampu memanfaatkan informasi *ensemble* sebanyak mungkin. Namun dalam mengkalibrasi prakiraan cuaca, BMA tidak mempertimbangkan aspek korelasi spasial. Korelasi spasial dalam meteorologi seringkali terjadi, karena kondisi cuaca suatu lokasi cenderung dipengaruhi oleh kondisi cuaca di lokasi lainnya yang berdekatan. GOP mengolah prakiraan cuaca dengan mempertimbangkan korelasi spasial seluruh lokasi secara serentak. Meskipun hanya didasari oleh satu model prakiraan deterministik saja, GOP memiliki parameter spasial yang berfungsi untuk memperbaiki hasil prakiraan cuaca yang awalnya belum mampu mengidentifikasi fenomena spasial. Spatial BMA merupakan kombinasi antara BMA dan GOP. PDF prediktif dari Spatial BMA merupakan densitas multivariat dengan struktur kovarians dibentuk untuk mempertimbangkan struktur spasial dari pengamatan cuaca.

Berdasarkan uraian di atas, maka rumusan masalah dalam penelitian ini adalah

1. Bagaimana bentuk penaksir parameter pada model Spatial BMA ?
2. Bagaimana hasil prakiraan terkalibrasi untuk suhu udara harian menggunakan Spatial BMA ?
3. Bagaimana perbandingan performansi metode Spatial BMA dengan BMA dan GOP pada kasus suhu udara harian ?

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mengkaji bentuk penaksir parameter Spatial BMA.
2. Menentukan prakiraan terkalibrasi untuk suhu udara harian menggunakan Spatial BMA.
3. Membandingkan performansi metode Spatial BMA dengan BMA dan GOP pada kasus suhu udara harian.

1.4 Manfaat Penelitian

Manfaat yang ingin dicapai pada penelitian ini adalah sebagai berikut.

1. Sebagai pengembangan keilmuan metode BMA, GOP, Spatial BMA khususnya mengenai metode statistika yang berhubungan dengan cuaca melalui pemanfaatan data luaran NWP dan prakiraan *ensemble*.
2. Sebagai alternatif peramalan cuaca bagi BMKG sebagai otoritas pemerintah yang berwenang menyebarkan informasi prakiraan cuaca maupun bagi institusi lain yang berhubungan dengan meteorologi. Metode ini memberikan gambaran seberapa baik koreksi bias yang dihasilkan untuk prakiraan cuaca jangka pendek, khususnya untuk suhu udara.

1.5 Batasan Masalah

Batasan permasalahan pada penelitian ini adalah sebagai berikut.

1. Ruang lingkup penelitian dibatasi pada stasiun meteorologi terpilih dengan anggota *ensemble* sebanyak 3 yang diperoleh melalui PLS, PCR, dan Regresi Ridge.
2. Variabel yang menjadi fokus penelitian adalah suhu udara, yaitu suhu maksimum dan suhu minimum.
3. Periode training yang digunakan adalah 30 hari karena memiliki akurasi yang cenderung lebih tinggi daripada periode yang lebih pendek maupun panjang.

BAB II TINJAUAN PUSTAKA

2.1 Distribusi Normal

Distribusi normal merupakan distribusi probabilitas kontinu yang terpenting dalam statistika karena dapat menggambarkan beberapa kejadian di alam, industri, ekonomi, dan lingkup lainnya dengan cukup baik. Beberapa komponen cuaca, seperti pengukuran suhu udara dan tekanan udara, umumnya mengikuti sebaran berdistribusi normal (Möller, 2014). Hal ini disebabkan error model kedua komponen cuaca tersebut jarang menunjukkan gejala penyimpangan yang sangat jauh dari deviasi standar atau dengan kata lain teridentifikasi sebagai pencilan (*outlier*).

2.1.1 Distribusi Normal Univariat

Menurut Casella dan Berger (2002), distribusi normal untuk variabel random kontinu y memiliki dua parameter, yaitu mean μ dan simpangan baku σ . Mean μ merupakan parameter lokasi (parameter yang menggeser densitas distribusi ke kiri atau ke kanan), sementara σ merupakan parameter skala (parameter yang memperlebar atau mempersempit densitas distribusi, namun tidak merubah bentuk distribusi). Fungsi probabilitas distribusi normal dapat dinotasikan seperti pada persamaan (2.1).

$$g(y | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2}, -\infty < y < \infty, -\infty < \mu < \infty, \sigma > 0 \quad (2.1)$$

2.1.2 Distribusi Normal Multivariat

Distribusi normal multivariat merupakan perluasan dari distribusi normal univariat, yang berdimensi lebih dari atau sama dengan 2. Dalam perspektif ini, normal multivariat dapat terjadi dalam kasus multi-respon yang saling berkorelasi, atau kasus yang mempertimbangkan korelasi antar ruang/lokasi (spasial). Mengacu pada persamaan (2.1), densitas normal univariat dapat diekstensi menjadi normal multivariat. Pada kasus ini, penelitian yang melibatkan banyak lokasi namun

berhingga, berlaku $s \in S$ dimana s merupakan himpunan lokasi pengamatan yang *fixed* dengan koordinat lintang-bujur tertentu. Dengan menggunakan *normalizing constant* multivariat $(2\pi)^{-s/2} |\Sigma|^{-1/2}$, densitas normal berdimensi s untuk vektor random $\mathbf{y}' = [y_1 \ y_2 \ \dots \ y_s]$ adalah

$$g(\mathbf{y}) = \frac{1}{(2\pi)^{s/2} |\Sigma|^{1/2}} e^{-\frac{(\mathbf{y}-\boldsymbol{\mu})' \Sigma^{-1} (\mathbf{y}-\boldsymbol{\mu})}{2}}, -\infty < y_i < \infty, i = 1, 2, \dots, s. \quad (2.2)$$

Densitas normal berdimensi s pada persamaan (2.2) dapat dinyatakan oleh $N_s(\boldsymbol{\mu}, \Sigma)$ dengan $\boldsymbol{\mu} = [\mu_1 \ \mu_2 \ \dots \ \mu_s]'$ merupakan vektor mean populasi berukuran $s \times 1$ dan Σ merupakan matrik varians-kovarian berukuran $s \times s$.

2.2 Principal Component Analysis (PCA)

Menurut Johnson dan Wichern (2007), *Principal Component Analysis* (PCA) adalah suatu metode tingkat menengah, digunakan untuk menjelaskan struktur varians-kovarian banyak variabel berdasarkan beberapa kombinasi linier yang terbentuk. PCA umumnya digunakan untuk mereduksi data dengan banyak variabel yang saling berkorelasi agar memudahkan interpretasi (Jolliffe, 2002). Hal tersebut dilakukan dengan mentransformasikan variabel-variabel asal ke beberapa komponen utama PC, yang diharapkan tidak lagi saling berkorelasi.

Mengacu pada perspektif aljabar linier, PC dipandang sebagai kombinasi linier dari p variabel random x_1, x_2, \dots, x_p . Kombinasi linier tersebut secara geometrik merepresentasikan pembentukan sistem koordinat yang baru dengan merotasi sistem koordinat asal dimana x_1, x_2, \dots, x_p berlaku sebagai sumbu koordinat. Struktur kovarian akan menjadi lebih sederhana dan interpretasi menjadi lebih mudah dengan adanya sumbu terbaru yang memiliki varians (keragaman) maksimal (Johnson dan Wichern, 2007). Syarat pembentukan PC sebagai kombinasi linier dari vektor random $\mathbf{x}' = [x_1 \ x_2 \ \dots \ x_p]$ agar mampu mewakili keragaman yang tinggi adalah dengan memilih vektor eigen $\mathbf{e}_i' = e_{i1} \ e_{i2} \ \dots \ e_{ip}$

sedemikian hingga $\text{var}(\mathbf{e}'_i \mathbf{x})$ bernilai maksimum dimana $\mathbf{e}'_i \mathbf{e}_i = 1$ dan $\text{cov}(\mathbf{e}'_i \mathbf{x}, \mathbf{e}'_m \mathbf{x}) = 0$ untuk $i \neq m$, dimana $i, m = 1, 2, \dots, p$.

Berdasarkan Johnson dan Wichern (2002), PC dapat dibentuk dari matrik kovarian dan matrik korelasi dimana matrik korelasi digunakan jika variabel penelitian memiliki satuan pengukuran yang beragam, sebaliknya matrik kovarian digunakan untuk variabel dengan satuan pengukuran yang relatif sama. Meski demikian, tidak banyak ditemukan variabel penelitian dengan satuan yang sama. Jolliffe (2002) menyatakan bahwa standardisasi variabel dapat menjadi alternatif untuk mengatasi problem satuan yang berbeda. Standarisasi data menyebabkan matrik varians-kovarian data yang distandardisasi akan sama dengan/menyerupai matrik korelasi data sebelum distandardisasi dan total varians PC akan sama besar dengan banyaknya variabel asal.

Secara umum, vektor random \mathbf{x}' memiliki matrik varians-kovarian Σ dengan akar ciri (*eigen value*) berturut-turut yaitu $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, sehingga kombinasi linier komponen utama adalah seperti pada persamaan (2.3) (Johnson dan Wichern, 2007).

$$\begin{aligned} PC_1 &= \mathbf{e}'_1 \mathbf{x} = e_{11}x_1 + e_{12}x_2 + \dots + e_{1p}x_p \\ PC_2 &= \mathbf{e}'_2 \mathbf{x} = e_{21}x_1 + e_{22}x_2 + \dots + e_{2p}x_p \\ &\vdots \\ PC_p &= \mathbf{e}'_p \mathbf{x} = e_{p1}x_1 + e_{p2}x_2 + \dots + e_{pp}x_p \end{aligned} \tag{2.3}$$

dengan :

PC_1 = PC pertama, yang mempunyai varians terbesar

PC_2 = PC kedua, yang mempunyai varians terbesar kedua

PC_p = PC ke- p , yang mempunyai varians terbesar ke- p

x_1 = variabel random pertama

x_2 = variabel random kedua

\vdots

x_p = variabel random ke- p

\mathbf{e}_p = vektor karakteristik (*eigen vector*) PC ke- p

Jadi, PC ke- i diberikan oleh persamaan berikut:

$$PC_i = \mathbf{e}_i' \mathbf{x}, \text{ dimana } : i = 1, 2, \dots, p,$$

dengan keterangan yang terkait dengan PC_i seperti ditunjukkan pada persamaan (2.4) dan (2.5).

$$\text{Var}(PC_i) = \mathbf{e}_i' \Sigma \mathbf{e}_i = \lambda_i, \quad i = 1, 2, \dots, p \quad (2.4)$$

$$\text{Cov}(PC_i, PC_m) = \mathbf{e}_i' \Sigma \mathbf{e}_m = 0, \quad i \neq m \quad (2.5)$$

PC yang terbentuk diharapkan tidak saling berkorelasi dan memiliki varians yang sama dengan akar ciri (*eigen value*) dari Σ , sehingga total varians (keragaman) terjelaskan dinotasikan pada persamaan (2.6),

$$\sum_{i=1}^p \text{Var}(x_i) = \sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} \quad (2.6)$$

$$\sum_{i=1}^p \text{Var}(PC_i) = \lambda_1 + \lambda_2 + \dots + \lambda_p$$

Persentase varians (keragaman) total yang dapat diterangkan oleh PC ke- i ditunjukkan pada persamaan (2.7), dimana

$$\text{proporsi varians PC ke-}i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \dots + \lambda_p}. \quad (2.7)$$

Apabila PC yang diambil sebanyak m dimana ($m < p$), maka persamaan (2.8) menjelaskan bahwa

$$\text{proporsi varians } m \text{ PC} = \frac{\lambda_1 + \lambda_2 + \dots + \lambda_m}{\lambda_1 + \lambda_2 + \dots + \lambda_p}. \quad (2.8)$$

Menurut Johnson dan Wichern (2007) dan Jolliffe (2002), terdapat beberapa acuan yang dapat digunakan untuk menentukan banyaknya PC, antara lain:

1. Melihat pola *scree plot* yang menggambarkan besarnya *eigenvalue* $\hat{\lambda}_i$. Banyaknya PC yang terpilih ditentukan oleh range terbesar antar titik dimana proses pemilihan PC akan dihentikan jika range antar titik makin rendah (terlihat kelandaian). Namun, pendekatan ini lebih bersifat subjektif, sehingga diperlukan pendekatan lain yang lebih objektif.
2. Pemilihan PC berdasarkan banyaknya *eigen value* yang lebih besar dari satu, yang merepresentasikan tingginya tingkat homogenitas antar variabel.
3. PC-PC yang terpilih diharapkan dapat memberikan kumulatif persentase varians 80% - 90%.

2.3 Partial Least Square (PLS)

Jika q variabel respon Y_1, \dots, Y_q diprediksi dengan p variabel prediktor X_1, \dots, X_p , dengan data terdiri atas n observasi, maka elemen pasangan respon dan prediktor dapat ditulis sebagai berikut, yaitu $(\mathbf{x}'_t, \mathbf{y}'_t)$, dimana $t = 1, \dots, n$. Metode PLS diawali dengan standarisasi data, baik variabel X maupun variabel Y . Salah satu teknik standarisasi adalah dengan mengurangi nilai tiap pengamatan dengan rata-rata dan dibagi dengan akar varians. Tujuan standarisasi data adalah untuk memperoleh pembobot dengan satuan relatif sama antar variabel berdasarkan metode perhitungan pada persamaan (2.9),

$$\begin{aligned} x_{tu}^* &= \frac{x_{tu} - \bar{x}_u}{S_u} \\ y_{tr}^* &= \frac{y_{tr} - \bar{y}_r}{S_r} \end{aligned} \quad (2.9)$$

dengan

$$\begin{aligned} \bar{x}_u &= \frac{1}{n} \sum_{t=1}^n x_{tu}; S_u = \sqrt{\frac{\sum_{t=1}^n (x_{tu} - \bar{x}_u)^2}{n-1}} \\ \bar{y}_r &= \frac{1}{n} \sum_{t=1}^n y_{tr}; S_r = \sqrt{\frac{\sum_{t=1}^n (y_{tr} - \bar{y}_r)^2}{n-1}} \end{aligned}$$

Pada persamaan (2.9), x_{tu}^* adalah data ke- t yang telah distandarisasi untuk variabel prediktor ke- u dimana $u = 1, 2, \dots, p$, sedangkan y_{tr}^* adalah data ke- t yang telah distandarisasi untuk variabel respon ke- r dimana $r = 1, 2, \dots, q$.

Kemudian, $\mathbf{x}'_t = (x_{t1} \cdots x_{tp})$ dihimpun dalam matriks \mathbf{X} yang berukuran $n \times p$, begitu juga dengan matriks \mathbf{Y} yang berukuran $n \times q$ berisikan $\mathbf{y}'_t = (y_{t1} \cdots y_{tq})$. Persamaan (2.10) menunjukkan himpunan matriks \mathbf{X} dan \mathbf{Y} tersebut

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}'_1 \\ \vdots \\ \mathbf{x}'_n \end{bmatrix} \text{ dan } \mathbf{Y} = \begin{bmatrix} \mathbf{y}'_1 \\ \vdots \\ \mathbf{y}'_n \end{bmatrix} \quad (2.10)$$

dimana n adalah banyak observasi, p adalah banyak prediktor dan q adalah banyaknya respon. Ketika $n < p$, metode regresi yang umum digunakan dengan teknik estimasi *Ordinary Least Square* (OLS) tidak dapat diaplikasikan karena matriks kovarians ($\mathbf{X}'\mathbf{X}$) adalah singular. Sebaliknya, PLS dapat diaplikasikan untuk kasus-kasus $n < p$ dan mampu menangani kasus ini dengan baik.

Dalam PLS, reduksi dimensi dan regresinya (pemodelan) dilakukan secara simultan. \mathbf{T} dinotasikan sebagai variabel laten atau *scores* yang merupakan hasil dekomposisi matriks dari variabel random observasi. \mathbf{P} dinyatakan sebagai *X-loadings* dan \mathbf{Q} dinyatakan sebagai *Y-loadings*. Regresi PLS didasarkan pada dekomposisi komponen laten seperti persamaan (2.11),

$$\begin{aligned} \mathbf{X} &= \mathbf{TP}' + \mathbf{E} \\ \mathbf{Y} &= \mathbf{TQ}' + \mathbf{F} \end{aligned} \quad (2.11)$$

dengan:

\mathbf{T} = matriks komponen laten (*scores matrices*) untuk n pengamatan yang berukuran $n \times c$

\mathbf{P} = matriks koefisien (*loadings matrices*) variabel X yang berukuran $p \times c$

\mathbf{Q} = matriks koefisien (*loadings matrices*) variabel Y yang berukuran $q \times c$

\mathbf{E} dan \mathbf{F} = matriks residual (Wold, Sjöström, dan Eriksson, 2001)

Menurut Wold, Sjöström, dan Eriksson (2001), PLS memiliki kesamaan dengan regresi PCR dimana PLS membentuk matriks komponen laten \mathbf{T} sebagai transformasi linier dari \mathbf{X} dan hubungan tersebut disajikan pada persamaan (2.12).

$$\mathbf{T} = \mathbf{XW} \quad (2.12)$$

\mathbf{W} adalah matriks *weight* (pembobot) yang berukuran $p \times c$ dengan c adalah banyaknya komponen laten, selanjutnya \mathbf{W} dan \mathbf{T} dinotasikan sebagai berikut :

$$\mathbf{w}_j = \begin{bmatrix} w_{1j} \\ \vdots \\ w_{pj} \end{bmatrix} \text{ dan } \mathbf{t}_j = \begin{bmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{bmatrix}, \text{ untuk } j = 1, 2, \dots, c.$$

Komponen laten berperan sebagai variabel asal \mathbf{X} dan digunakan untuk memprediksi \mathbf{Y} . Ketika \mathbf{T} sudah terbentuk, maka diperoleh \mathbf{Q}' melalui metode kuadrat terkecil dengan formulasi yang ditunjukkan pada persamaan (2.13).

$$\mathbf{Q}' = (\mathbf{T}' \mathbf{T})^{-1} \mathbf{T}' \mathbf{Y} \quad (2.13)$$

Dengan demikian, berdasarkan Wold, Sjöström, dan Eriksson (2001), matriks koefisien regresi untuk model $\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{F}$ (analog dengan model OLS) adalah matriks \mathbf{B} dimana estimator \mathbf{B} diperoleh dari penurunan singkat berikut.

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{F} = \mathbf{T}\mathbf{Q}' + \mathbf{F} \text{ (diperoleh dari persamaan (2.11))}$$

$$\Leftrightarrow \mathbf{X}\mathbf{B} = \mathbf{T}\mathbf{Q}'$$

$$\Leftrightarrow \mathbf{X}\mathbf{B} = \mathbf{X}\mathbf{W}\mathbf{Q}'$$

$$\Leftrightarrow \mathbf{B} = \mathbf{W}\mathbf{Q}' \text{ (}\mathbf{Q}' \text{ disubstitusi dari persamaan (2.13))}$$

Jadi, estimator untuk \mathbf{B} adalah seperti yang ditunjukkan persamaan (2.14),

$$\hat{\mathbf{B}} = \mathbf{W}(\mathbf{T}' \mathbf{T})^{-1} \mathbf{T}' \mathbf{Y} \quad (2.14)$$

dimana dapat disimpulkan bahwa dugaan \mathbf{Y} diperoleh berdasarkan persamaan (2.15).

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\mathbf{B}}$$

$$\hat{\mathbf{Y}} = [\mathbf{T}\mathbf{W}^{-1}\mathbf{W}(\mathbf{T}' \mathbf{T})^{-1} \mathbf{T}' \mathbf{Y}] \quad (2.15)$$

$$\hat{\mathbf{Y}} = [\mathbf{I}\mathbf{I}(\mathbf{T}' \mathbf{T})^{-1} \mathbf{T}' \mathbf{Y}] = [\mathbf{T}(\mathbf{T}' \mathbf{T})^{-1} \mathbf{T}' \mathbf{Y}]$$

Karakteristik berbagai pendekatan regresi PLS dapat dilakukan dengan menitikberatkan pada empat hal berikut (Wold, Sjöström, dan Eriksson, 2001).

1. Fungsi objektif yang dimaksimumkan oleh matriks pembobot \mathbf{W}
2. Nilai matriks pembobot \mathbf{W}
3. Matriks koefisien regresi \mathbf{B}
4. Algoritma yang digunakan untuk menghitung pembobot \mathbf{W} .

2.4 *Principal Component Regression (PCR)*

Principal Component Regression (PCR) adalah metode yang dapat digunakan untuk mengatasi multikolinieritas tanpa harus mengeluarkan prediktor yang terlibat hubungan kolinear. Dengan demikian, seluruh prediktor berkontribusi dalam pemodelan. Konsep utama dari PCR ialah menyederhanakan banyaknya

prediktor dengan cara mereduksi dimensinya. Hal ini dilakukan dengan cara mereduksi korelasi antar variabel bebas melalui transformasi prediktor asal ke variabel baru yang tidak lagi berkorelasi, dapat disebut sebagai komponen utama (PC). Setelah satu atau beberapa PC diperoleh, seluruh PC terpilih tersebut menjadi prediktor baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel respon dengan menggunakan analisis regresi (Jolliffe, 2002).

Cara pembentukan PCR melalui PCA ada dua cara, yaitu dengan menggunakan matrik kovarian atau matrik korelasi. Kedua cara tersebut digunakan berdasarkan tinggi rendahnya skala pengamatan antar prediktor (Johnson dan Wichern, 2002). Dalam penelitian ini, regresi komponen utama dibentuk berdasarkan matrik kovarian karena data terlebih dulu distandarisasi sebelum diolah/dimodelkan, sehingga perbedaan skala pengukuran antar variabel menjadi tidak signifikan atau relatif seimbang.

Dimisalkan bahwa $\mathbf{x}' = [x_1 \ x_2 \ \dots \ x_p]$ merupakan vektor prediktor dari matrik \mathbf{X} berukuran $n \times p$. Jika matrik \mathbf{A} adalah matrik ortogonal berukuran $p \times p$ dengan kolom ke- m berisikan vektor eigen ke- m dari $\mathbf{X}'\mathbf{X}$ dengan asumsi $m \leq p$, maka $\mathbf{A}'\mathbf{A} = \mathbf{A}\mathbf{A}' = \mathbf{I}$. Skor PC untuk tiap pengamatan adalah

$$\mathbf{Z} = \mathbf{X}\mathbf{A} \quad (2.16)$$

Dengan elemen (i,m) dari \mathbf{Z} merupakan skor dari PC ke- m untuk pengamatan ke- i dimana $i=1,2,\dots,n$.

Berdasarkan sifat ortogonal \mathbf{A} dalam Jolliffe (2002), regresi linear berganda dengan prediktor yang saling berkorelasi dapat dikonversikan ke dalam PCR, seperti pada persamaan (2.17),

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \mathbf{y} &= \mathbf{X}\mathbf{A}\mathbf{A}'\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ karena } \mathbf{A}\mathbf{A}' = \mathbf{I} \\ \mathbf{y} &= \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \end{aligned} \quad (2.17)$$

Dengan $\mathbf{X}\mathbf{A} = \mathbf{Z}$ dan $\mathbf{A}'\boldsymbol{\beta} = \boldsymbol{\gamma}$. Estimasi untuk $\boldsymbol{\gamma}$ dapat dilakukan dengan metode OLS, lalu mencari $\boldsymbol{\beta}$ berdasarkan persamaan (2.18).

$$\hat{\boldsymbol{\beta}} = \mathbf{A}\hat{\boldsymbol{\gamma}} \quad (2.18)$$

Karena data pada awalnya distandardisasi, maka diperlukan transformasi sedemikian hingga respon (Y) kembali ke satuan semula, yaitu dengan mengalikan Y terhadap akar variansnya, lalu dijumlahkan dengan rata-ratanya.

2.5 Regresi Ridge

Regresi Ridge merupakan salah satu dari sekian banyak metode yang digunakan untuk mereduksi pengaruh dari prediktor yang saling berkorelasi. Berdasarkan Rencher dan Schaalje (2008), singularitas $\mathbf{X}'\mathbf{X}$ akibat multikolinieritas dapat menyebabkan estimasi β menjadi tidak efisien (varians yang besar). Hal ini dikarenakan estimasi OLS ataupun estimasi MLE yang umumnya bergantung pada hasil $(\mathbf{X}'\mathbf{X})^{-1}$. Selain itu, regresi ini sesuai pada kasus dengan jumlah prediktor yang lebih banyak daripada jumlah observasi.

Estimator β , seperti pada persamaan (2.19) yang diperoleh melalui OLS, dapat menimbulkan permasalahan jika terjadi multikolinieritas sempurna atau mendekati sempurna yang berdampak pada tidak validnya hasil pengujian parameter regresi dan selang kepercayaan yang lebar. Dampak lainnya adalah berbedanya tanda koefisien regresi dari yang diharapkan.

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad (2.19)$$

Berdasarkan Draper dan Smith (1992), regresi Ridge menggunakan suatu konstanta non-negatif λ yang memodifikasi persamaan (2.19) untuk menghitung koefisien regresi yang lebih efisien, meskipun bias. Persamaan (2.20) merupakan persamaan yang lazimnya digunakan untuk mengestimasi koefisien regresi Ridge.

$$\hat{\beta}_{ridge} = (\mathbf{X}'\mathbf{X} + \lambda\mathbf{I})^{-1} \mathbf{X}'\mathbf{y} \quad (2.20)$$

Sama seperti PLS dan PCR, pada regresi Ridge disarankan untuk melakukan standarisasi terhadap prediktor dan respon sehingga perbedaan skala perhitungan antar prediktor tidak terlampaui signifikan. Konstanta λ dapat dipilih secara intuitif (*by default*) atau dengan teknik *cross-validation* berdasarkan RMSE terendah. Makin tinggi λ , maka koefisien β_p makin mendekati 0 atau menyebabkan parameter p kurang berpengaruh terhadap respon.

2.6 *Bayesian Model Averaging (BMA)*

Menurut Jolliffe dan Stephenson (2011), hasil prakiraan *ensemble* cenderung bersifat *underdispersive* atau *overdispersive* (nilai prakiraan cuaca terpusat pada suatu nilai dengan varians yang rendah atau tinggi). Keadaan yang demikian dapat menyebabkan hasil prakiraan menjadi kurang reliabel sebab *underdispersive* mengindikasikan interval prakiraan yang relatif sempit, sementara *overdispersive* mengindikasikan interval prakiraan yang relatif lebih lebar. Permasalahan tersebut dapat ditangani dengan kalibrasi prakiraan *ensemble* untuk menghasilkan fungsi padat probabilitas (PDF) prediktif yang lebih reliabel dan tajam. Jika interval prakiraan *ensemble* terkalibrasi memiliki standar deviasi dengan ukuran yang proporsional, maka hasil prakiraan tersebut mencapai keadaan *equidispersive*.

Salah satu metode yang dapat digunakan untuk mengkalibrasi prakiraan *ensemble* adalah *Bayesian Model Averaging* atau BMA, yang awalnya digunakan pada bidang sosial dan kesehatan. Aplikasi BMA untuk meteorologi diinisiasi oleh Raftery *et al.* (2005) yang mengembangkan metode ini berdasarkan konsep model dinamis. Kelebihan metode ini terletak pada kemampuan dalam memadukan beberapa model prakiraan *ensemble* dan mengkoreksi bias prakiraan agar mean dan varians makin mendekati kenyataan, sehingga prakiraan *ensemble* yang terkalibrasi dapat mendekati nilai observasi yang sesungguhnya. Dengan memanfaatkan informasi dari beberapa model *ensemble*, maka prakiraan yang dihasilkan tidak hanya bergantung pada salah satu model saja. Selain itu, estimator parameter BMA tidak stagnan dan selalu berubah mengikuti informasi cuaca dan atmosfer yang dinamis.

Fungsi padat probabilitas (PDF) prediktif BMA merupakan kombinasi linier beberapa model prakiraan dimana tiap model memiliki kontribusi atau bobot (*weight*) yang berbeda terhadap pembentukan PDF prediktif. Dalam BMA, penentuan bobot dilakukan berdasarkan konsep rata-rata (*model averaging*). Besaran bobot tergantung dari kemampuan model dalam melakukan prakiraan komponen cuaca dimana bobot akan semakin besar jika hasil prakiraan semakin mendekati nilai observasi riil. Dari sudut pandang Bayesian, bobot-bobot tersebut merupakan probabilitas posterior masing-masing anggota/model *ensemble* (Feldmann, 2012).

Dimisalkan bahwa pada suatu lokasi tertentu, y adalah nilai riil dari elemen cuaca yang menjadi fokus penelitian dengan f_1, f_2, \dots, f_M adalah kumpulan M -model prakiraan yang berbeda. Koreksi bias dari tiap anggota *ensemble* m dapat direpresentasikan ke persamaan (2.21).

$$y = \beta_{0,m} + \beta_{1,m}f_m + \varepsilon_m \quad (2.21)$$

Dalam kerangka kerja BMA, tiap prakiraan anggota *ensemble* f_m , $m = 1, 2, \dots, M$, memiliki PDF posterior $g_m(y | f_m)$ yang dapat diinterpretasikan sebagai densitas dari y jika f_m merupakan prakiraan anggota *ensemble* “terbaik” (Raftery et al., 2005).

Mengacu pada teorema probabilitas total, PDF posterior (prediktif) dari y berdasarkan data training $D = [y_1^{obs} \ y_2^{obs} \ \dots \ y_T^{obs}]$ adalah

$$g_m(y | f_1, \dots, f_M, D) = \sum_{m=1}^M g_m(y | f_m, D) g_m(f_m | D) \quad (2.22)$$

dimana persamaan (2.22) adalah PDF posterior hasil rata-rata dari PDF posterior tiap anggota *ensemble*, diboboti oleh probabilitas posterior dari tiap anggota *ensemble* (Raftery dan Zheng, 2003). Probabilitas posterior untuk anggota *ensemble* f_m adalah seperti pada persamaan

$$g(f_m | D) = \frac{g(D | f_m) g(f_m)}{\sum_{l=1}^M g(D | f_l) g(f_l)} \quad (2.23)$$

dimana $g(D | f_m) = \int g(D | \theta_m, f_m) g(\theta_m | f_m) d\theta_m$ merupakan likelihood anggota *ensemble* m yang diintegrasikan, θ_m adalah vektor parameter anggota *ensemble* m (misal $\theta = [\beta_0, \beta_1, \sigma^2]$), $g(\theta_m | f_m)$ adalah densitas prior θ_m untuk prakiraan *ensemble* f_m , $g(D | \theta_m, f_m)$ adalah likelihood, dan $g(f_m)$ adalah probabilitas prior untuk tiap *ensemble* m yang umumnya sebesar $1/M$.

Berdasarkan Raftery et al. (2005), persamaan (2.22) dapat dimodifikasi untuk kasus model dinamis, sehingga model prediktif BMA menjadi

$$g(y | f_1, f_2, \dots, f_M) = \sum_{m=1}^M w_m g_m(y | f_m) \quad (2.24)$$

dengan w_m merupakan probabilitas posterior dari anggota *ensemble* m yang dianggap sebagai model “terbaik” dimana w_1, \dots, w_M adalah nonnegatif dan $w_1 + w_2 + \dots + w_M = 1$. Persamaan (2.15) mengindikasikan bahwa PDF prediktif adalah model *mixture* dari beberapa densitas posterior. Bobot w_m pada persamaan (2.15) ditentukan oleh kemampuan peramalan tiap anggota *ensemble* selama periode *training*. PDF prediktif BMA untuk model dinamis pada persamaan (2.15) dirancang untuk kasus univariat, baik dari aspek banyak variabel respon, lokasi, dan *lead* waktu peramalan, misal dikhususkan untuk prakiraan 24 jam ke depan saja, sehingga korelasi spasial atau temporal antar error tidak dapat ditangani (Feldmann, 2012; Möller, 2014).

Pada BMA, koreksi bias dilakukan dengan menerapkan transformasi linier $\beta_{0,m} + \beta_{1,m}f_m$ untuk tiap anggota *ensemble* (Raftery *et al.*, 2005). Transformasi tersebut berkaitan dengan ekspektasi dari densitas bersyarat p_m . Parameter $\beta_{0,m}$ dan $\beta_{1,m}$ masing-masing adalah term koreksi bias adaptif dan multiplikatif yang diperoleh dari regresi linier sederhana antara observasi y dan f_m selama periode *training*. Jika prakiraan anggota *ensemble* f_m berdistribusi normal univariat, maka distribusi posterior dari y jika f_m merupakan anggota *ensemble* yang “terbaik” adalah

$$y | f_m \sim N(\beta_{0,m} + \beta_{1,m}f_m, \sigma^2) \quad (2.25)$$

dengan varians σ^2 yang diasumsikan konstan untuk seluruh anggota *ensemble* m .

Penentuan distribusi yang digunakan untuk menghitung densitas posterior y umumnya dilakukan dengan mengidentifikasi karakteristik data komponen cuaca. Suhu dan tekanan udara umumnya mengikuti distribusi normal seperti dinyatakan pada persamaan (2.16), sementara curah hujan dan kecepatan angin cenderung mengikuti distribusi Gamma (Möller, 2014). Berdasarkan Raftery *et al.* (2005), mean posterior (prediktif) BMA dinyatakan dalam persamaan (2.26), prakiraan deterministik, yaitu rata-rata terboboti dari prakiraan *ensemble* yang sudah terkalibrasi,

$$E(y | f_1, \dots, f_M) = \sum_{m=1}^M w_m (\beta_{0,m} + \beta_{1,m} f_m) \quad (2.26)$$

dan varians posterior (prediktif) BMA adalah

$$\text{var}(y | f_1, \dots, f_M) = \sum_{m=1}^M w_m \left((\beta_{0,m} + \beta_{1,m} f_m) - \sum_{m=1}^M w_m (\beta_{0,m} + \beta_{1,m} f_m) \right)^2 + \sigma^2 \quad (2.27)$$

Secara verbal, persamaan (2.27) dapat pula direpresentasikan ke dalam persamaan (2.28).

$$\text{Varians prediktif} = \text{Varians antar ensemble} + \text{Varians dalam ensemble} \quad (2.28)$$

Jika ditetapkan observasi riil y_t dan prakiraan *ensemble* sebanyak M f_{1t}, \dots, f_{Mt} untuk suatu waktu (jam, hari atau bulan) t selama periode *training* tertentu pada suatu lokasi, seluruh parameter BMA dapat diestimasi melalui 3 tahapan berikut:

- 1) Parameter $\beta_{0,m}$ dan $\beta_{1,m}$ untuk $m = 1, 2, \dots, M$ diestimasi dengan regresi linier sederhana antara observasi y_t dan f_{mt} .
- 2) Bobot w_m untuk $m = 1, 2, \dots, M$ dan varians diestimasi dengan memaksimumkan fungsi log-likelihood pada persamaan (2.29), dengan asumsi error prakiraan tidak saling berkorelasi antar lokasi dan waktu.

$$l(w_1, \dots, w_M, \sigma^2) = \sum_{t=1}^T \ln \left(\sum_{m=1}^M w_m g_m(y_t | f_m) \right) \quad (2.29)$$

- 3) Diperoleh bahwa persamaan (2.29) tidak dapat diselesaikan secara analitik. Terdapat beberapa pendekatan numerik yang digunakan untuk mengatasi permasalahan tersebut, salah satunya dengan algoritma *Expectation-Maximization* (EM) seperti pada Raftery *et al.* (2005) yang digagas oleh Dempster, Laird, dan Rubin (1977). Dimisalkan terdapat himpunan variabel laten yang tidak teramati $\mathbf{Z}_t = [z_{1t} \ z_{2t} \ \dots \ z_{Mt}]$, maka variabel laten yang tidak teramati z_{mt} akan bernilai 1 jika anggota *ensemble* m adalah prakiraan terbaik pada waktu t dan bernilai 0 jika bukan prakiraan terbaik. Himpunan z_t hanya memiliki satu elemen z_{mt} saja yang bernilai 1, sementara sisanya bernilai 0. Algoritma EM diawali oleh tahap *expectation* dimana nilai z_{mt} diestimasi secara

berulang berdasarkan bobot dan varians yang sudah ditentukan pada tahap inisiasi, dinyatakan dalam persamaan (2.30)

$$\hat{z}_{mt}^{(i)} = \frac{w_m g_m(y_t | f_{mt}, \sigma^{(i-1)})}{\sum_{l=1}^M w_l g_l(y_t | f_{lt}, \sigma^{(i-1)})} \quad (2.30)$$

dengan i adalah iterasi dan $g_m(y_t | f_{mt}, \sigma^{(i-1)})$ adalah PDF bersyarat dari anggota *ensemble* m yang dihitung pada observasi y_t dengan mean $\mu_{mt} = \beta_{0,m} + \beta_{1,m} f_{mt}$ dan standar deviasi $\sigma^{(i-1)}$ yang dievaluasi dari observasi cuaca y . Pada tahap berikutnya, tahap *maximization*, bobot dan varians diperbarui berdasarkan $z_{mt}^{(i)}$ yang dinyatakan dalam persamaan (2.31).

$$w_m^{(i)} = \frac{1}{T} \sum_{t=1}^T z_{mt}^{(i)}$$

$$\sigma^{2(i)} = \frac{1}{T} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2 \quad (2.31)$$

Konvergensi bobot dan varians tercapai jika selisih antar bobot, antar varians, dan antar \hat{z}_{mt} di tiap iterasi lebih kecil dari batas toleransi yang ditentukan.

2.7 Geostatistical Output Perturbation (GOP)

Teknologi dan metode yang digunakan untuk prakiraan cuaca saat ini sudah berkembang pesat, antara lain dengan memadukan beberapa metode, sehingga diperoleh data *ensemble* yang kemudian digunakan sebagai acuan untuk prakiraan cuaca. Namun, teknik ini tidak dapat diaplikasikan dengan optimal tanpa dukungan data atmosfer yang besar dan teknologi komputasi yang canggih. Permasalahan ini seringkali ditemui di banyak negara berkembang yang minim sumber daya. Untuk mengatasi hal tersebut, Gel *et al.* (2004) memodifikasi luaran NWP yang bersifat deterministik dengan memasukkan pengaruh error yang berdistribusi normal. Error tersebut dapat dimodifikasi dan dibangkitkan melalui simulasi model geostatistika hingga diperoleh prakiraan terkalibrasi yang reliabel dan tetap tajam. Selain itu, GOP juga mampu menangkap pola dependensi spasial yang seringkali teridentifikasi pada error model.

Berdasarkan Gel *et al.* (2004), dimisalkan S adalah himpunan lokasi pengamatan yang cukup besar, namun berhingga (*finite*) dimana waktu $t \in T$ dan periode *training* dianggap *fixed*. Prediktor yang digunakan umumnya hanya ada 1, misal luaran NWP pada domain grid yang bersesuaian/berdekatan dengan lokasi pengamatan. Pada periode *training* $t=1,2,\dots,T$ untuk seluruh lokasi s dimana $s \in S$, berlaku persamaan (2.32)

$$y_{s,t} = \beta_0 + \beta_1 x_{s,t} + \varepsilon_{s,t} \quad (2.32)$$

dimana $y_{s,t}$ adalah respon (observasi suhu udara) pada stasiun s hari ke- t , $x_{s,t}$ adalah prediktor (NWP suhu udara) pada grid terdekat dari stasiun s , dan $\varepsilon_{s,t}$ adalah error regresi. Untuk stasiun s yang sama, error $\varepsilon_{s,t}$ untuk suhu udara disyaratkan memenuhi asumsi error identik, independen, dan berdistribusi normal. Melalui estimasi dengan OLS atau MLE, diperoleh persamaan (2.33) bahwa

$$\beta_1 = \frac{n \sum_{j=1}^n x_j y_j - \left(\sum_{j=1}^n x_j \right) \left(\sum_{j=1}^n y_j \right)}{n \sum_{j=1}^n x_j^2 - \left(\sum_{j=1}^n x_j \right)^2} = \frac{\sum_{j=1}^n (x_j - \bar{x})(y_j - \bar{y})}{\sum_{j=1}^n (x_j - \bar{x})^2}; \beta_0 = \bar{y} - \beta_1 \bar{x} \quad (2.33)$$

dimana n adalah banyaknya pasangan $\{s,t\}$. Dapat dikatakan bahwa besar kemungkinan error antar lokasi tidak independen karena adanya pengaruh spasial.

Dengan mempertimbangkan aspek multivariat antar lokasi, untuk $t=1,2,\dots,T$, berlaku $\mathbf{y}'_t = [y_{1t} \ y_{2t} \ \dots \ y_{st}]$ vektor elemen cuaca yang diamati di seluruh lokasi pengamatan berukuran $s \times 1$ dan $\mathbf{x}'_t = [x_{1t} \ x_{2t} \ \dots \ x_{st}]$ vektor prakiraan cuaca berukuran $s \times 1$

$$\mathbf{y}_t = \beta_0 \mathbf{1} + \beta_1 \mathbf{x}_t + \boldsymbol{\varepsilon}_t \quad (2.34)$$

dengan $\mathbf{1}$ adalah vektor berukuran $s \times 1$ yang seluruh elemennya bernilai 1 dan $\boldsymbol{\varepsilon}_t = [\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{st}]'$. Error model GOP dari persamaan (2.34) berdistribusi normal dengan mean $\mathbf{0}$ dan kovarian $\boldsymbol{\Sigma}$ yang nilainya bergantung pada struktur kovarians dari error secara spasial.

Jika $C(\mathbf{s}_i, \mathbf{s}_j)$ pada persamaan (2.35) merupakan fungsi korelasi yang stasioner dan isotropi, maka nilai elemen (i,j) $\boldsymbol{\Sigma}$ diperoleh melalui

$$\frac{1}{2} \text{var}(\varepsilon(\mathbf{s}_i) - \varepsilon(\mathbf{s}_j)) = \rho^2 + \tau^2 (1 - C(\mathbf{s}_i, \mathbf{s}_j)) \quad (2.35)$$

Dalam terminologi geostatistika, ρ^2 adalah efek *nugget* yang dapat bermakna ganda, yaitu varians kesalahan pengukuran/instrumen/alat dan juga ukuran dari keragaman spasial pada jarak tertentu yang masih memberikan pengaruh, $\rho^2 + \tau^2$ adalah varians marginal dari residual atau *sill* (Cressie, 1993).

Terdapat beberapa fungsi korelasi yang dapat digunakan untuk memodelkan struktur spasial, antara lain fungsi korelasi eksponensial Gel *et al.* (2004), yaitu

$$C(\mathbf{s}_i, \mathbf{s}_j) = \exp\left(-\frac{\|\mathbf{s}_i - \mathbf{s}_j\|}{r}\right) \quad (2.36)$$

dimana $\|\cdot\|$ dan r pada persamaan (2.36) masing-masing adalah norm Euclidean dan range yang mengindikasikan rentang jarak dimana korelasi spasial error mulai berkurang secara signifikan dan dinyatakan dalam satuan km. Korelasi spasial bernilai 1 jika jarak antar lokasi d adalah 0 dan makin menurun hingga bernilai 0,5 saat d mencapai $3r$ (Cressie, 1993). Persamaan (2.36) mengindikasikan dependensi error dapat diidentifikasi melalui variogram eksponensial

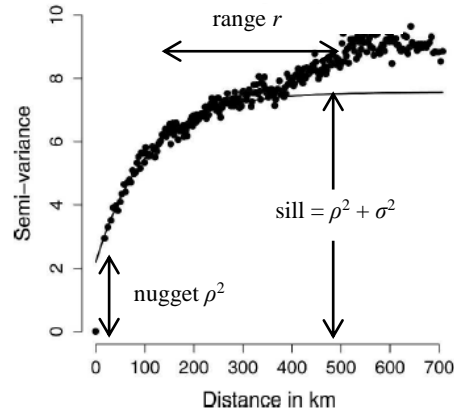
$$\gamma(d) = \rho^2 + \tau^2 \left(1 - \exp\left(-\frac{d}{r}\right)\right) \quad (2.37)$$

dimana d pada persamaan (2.37) merupakan $\|\mathbf{s}_i - \mathbf{s}_j\|$ yang menyatakan jarak Euclidean antara himpunan pasangan lokasi s_i dan s_j .

Untuk mendapatkan estimasi parameter, terdapat dua tahap yang dilakukan berdasarkan pendekatan menyerupai *maximum likelihood*, yaitu:

1. Mengestimasi koefisien β_0 dan β_1 menggunakan regresi linier sederhana dan menghitung residual model GOP.
2. Residual model GOP kemudian menjalani proses *binning* (pengelompokan) untuk mengestimasi parameter variogram ρ^2, τ^2 dan r . Metode yang digunakan untuk estimasi adalah kuadrat terkecil terboboti atau *weighted least squares* WLS berdasarkan Cressie (1993), dengan bobot sebesar banyaknya observasi dalam kelompok-kelompok *binning*.

Setelah estimasi ρ^2, τ^2 dan r diperoleh, maka dilakukan penyatuan informasi semivarians empiris dan teoritis ke dalam semivariogram, seperti yang diilustrasikan pada Gambar 2.1.



Gambar 2.1 Semivariogram Pra-Simulasi GOP (Berrocal *et al.*, 2007)

Gambar 2.1 menunjukkan bahwa semi-varians, yang mendeskripsikan kontinuitas spasial, makin tidak terlihat nyata dan bernilai relatif stabil ketika jarak menyentuh *range r* tertentu. Besarnya semi-varians atau disebut juga semivariogram pada range tersebut disebut *sill*. Kedua parameter tersebut, termasuk *nugget*, memiliki kontribusi terhadap hasil prakiraan cuaca dan dapat pula dipakai untuk menghasilkan prakiraan cuaca pada stasiun pengamatan yang tidak dilibatkan dalam pemodelan, yang disebut sebagai metode *kriging*. Kriging sendiri didasari dari persamaan linear terboboti dimana bobot bernilai makin besar jika jarak antara kedua lokasi makin dekat, begitu pula sebaliknya (Cressie, 1993).

2.8 Spatial Bayesian Model Averaging (Spatial BMA)

Postprocessing secara statistik dibutuhkan untuk mengoptimalkan prakiraan *ensemble*. Hal ini bertujuan untuk menghasilkan prakiraan probabilistik yang terkalibrasi dan tetap tajam. Metode untuk *postprocessing ensemble* secara statistik, diantaranya, *Bayesian Model Averaging* (BMA) (Raftery *et al.*, 2005) dan *Geostatistical Output Perturbation* (GOP) (Gel *et al.*, 2004), serta *Spatial Bayesian Model Averaging* (Spatial BMA) yang merupakan kombinasi antara BMA dan GOP (Berrocal *et al.*, 2007).

Metode Spatial BMA diharapkan mampu menutupi kekurangan dari metode BMA dan GOP. Seperti pada BMA, PDF prediktif Spatial BMA merupakan rata-

rata terboboti dari PDF prakiraan yang terpusat pada koreksi bias dari model anggota *ensemble*, dengan bobot yang berhubungan dengan kontribusi masing-masing anggota. Pada Spatial BMA, PDF prakiraan merupakan densitas multivariat dengan struktur kovarians dibentuk untuk mempertimbangkan struktur spasial dari pengamatan cuaca. Teknik Spatial BMA dapat digunakan untuk mendapatkan *ensemble* statistik dari seluruh wilayah cuaca secara simultan, dari berbagai ukuran, dan dengan biaya yang minimum.

Metode Spatial BMA mempertimbangkan wilayah cuaca $\mathbf{Y} = \{Y(s) : s \in S\}$, dimana S adalah himpunan lokasi yang cukup besar namun berhingga (*finite*) dan bersyarat pada suatu *ensemble*.

$$\mathbf{F}_1 = \{f_1(s) : s \in S\}, \dots, \mathbf{F}_M = \{f_M(s) : s \in S\}$$

dari M prakiraan cuaca secara simultan, bukan hanya prakiraan cuaca deterministik tunggal. PDF prediktif Spatial BMA untuk prakiraan cuaca adalah

$$g(\mathbf{y}_t | \mathbf{F}_1, \dots, \mathbf{F}_M) = \sum_{m=1}^M w_m g_m(\mathbf{y}_t | \mathbf{F}_m) \quad (2.38)$$

dimana w_m merupakan bobot BMA, sama dengan probabilitas bahwa anggota m merupakan yang terbaik diantara anggota *ensemble* prakiraan, dan $g_m(\mathbf{y}_t | \mathbf{F}_m)$ merupakan PDF bersyarat dari \mathbf{y}_t jika diketahui anggota m adalah yang terbaik. Dalam implementasinya, PDF bersyarat merupakan densitas multivariat yang terpusat pada koreksi bias anggota *ensemble* prakiraan, $\beta_{0,m} \mathbf{1} + \beta_{1,m} \mathbf{F}_m$, dan memiliki struktur kovarians matriks secara spasial Σ_m . Kondisi ini dapat dinotasikan dengan

$$(\mathbf{y}_t | \mathbf{F}_m) \sim MVN(\beta_{0,m} \mathbf{1} + \beta_{1,m} \mathbf{F}_m, \Sigma_m) \quad (2.39)$$

Pada persamaan (2.39),

$$\Sigma_m = \frac{\sigma^2}{\rho_m^2 + \tau_m^2} \Sigma, \quad (2.40)$$

dimana σ^2 adalah varians BMA, Σ adalah struktur matriks kovarians GOP secara spasial, ρ_m^2 dan τ_m^2 masing-masing adalah parameter kovarians GOP. Suatu kuantitas

$$\alpha_m = \frac{\sigma^2}{\rho_m^2 + \tau_m^2}$$

adalah ratio dari varians BMA dengan varians eror GOP, dan disebut sebagai faktor deflasi untuk anggota model m , dimana $m = 1, \dots, M$. Spatial BMA mengeneralisasikan metode BMA dan GOP. Hal ini dilakukan dengan mereduksi bentuk awal ketika himpunan S hanya terdiri dari *single* lokasi, dan kemudian mereduksi suatu *ensemble* berukuran $M = 1$, yaitu, prakiraan cuaca deterministik.

Sama seperti GOP, Spatial BMA memiliki PDF prediktif multivariat untuk prakiraan cuaca. Namun, PDF tersebut lebih bermanfaat untuk menghasilkan *ensemble* statistik dari prakiraan cuaca, dengan sampling dari PDF prediktif Spatial BMA. Syarat anggota *ensemble* m menjadi terbaik dapat dituliskan seperti persamaan (2.41).

$$\mathbf{y}_t | \mathbf{F}_m \sim \beta_{0,m} \mathbf{1} + \beta_{1,m} \mathbf{F}_m + \boldsymbol{\varepsilon}_{1m} + \boldsymbol{\varepsilon}_{2m} \quad (2.41)$$

dimana $\boldsymbol{\varepsilon}_{1m}$ dan $\boldsymbol{\varepsilon}_{2m}$ masing-masing menotasikan bagian kontinu dan diskrit eror bersyarat, dengan PDF multivariat normal, kecuali matriks kovarians *di-rescale* dengan faktor deflasi, α_k .

Berikut algoritma untuk mendapatkan anggota *ensemble* Spatial BMA :

- 1) Ambil sebanyak $m \in \{1, \dots, M\}$ sampel, dengan probabilitas diberikan oleh bobot BMA w_1, w_2, \dots, w_M . Hal ini dilakukan untuk mendapatkan anggota dari *ensemble* dinamis.
- 2) Simulasikan realisasi dari bagian kontinu dan diskrit, $\boldsymbol{\varepsilon}_{1m}$ dan $\boldsymbol{\varepsilon}_{2m}$, pada eror bersyarat dari masing-masing PDF bersyarat.
- 3) Gunakan ruas kanan persamaan (2.41) untuk mendapatkan bias terkoreksi prakiraan cuaca, $\beta_{0,m} \mathbf{1} + \beta_{1,m} \mathbf{f}_m$, dengan simulasi eror bersyarat, $\boldsymbol{\varepsilon}_{1m}$ dan $\boldsymbol{\varepsilon}_{2m}$.

Selanjutnya, didapatkan *ensemble* Spatial BMA dari prakiraan cuaca, dengan berbagai ukuran *ensemble* yang diinginkan, dan dengan biaya komputasi yang minimum (Berrocal et al., 2007).

2.9 Evaluasi Model Kalibrasi

Pada umumnya, kebaikan model untuk prakiraan cuaca dapat dievaluasi berdasarkan berbagai sudut pandang. Namun, evaluasi untuk model kalibrasi yang memanfaatkan *ensemble* sedikit berbeda daripada model-model prakiraan cuaca lainnya. Definisi kalibrasi adalah konsistensi statistik yang terbentuk antara observasi dan hasil ramalan metode *ensemble*, misal Spatial BMA. Suatu prakiraan dikatakan terkalibrasi jika kejadian yang diramalkan akan terjadi dengan probabilitas 70%, maka dalam jangka panjang probabilitas kejadian itu di masa mendatang berada pada kisaran 70%, misal 69% atau 71% (Raftery et al., 2005).

Kebaikan model tidak hanya dievaluasi dari eror prakiraan yang diperoleh, tetapi juga koreksi bias dan ketajaman prakiraan *ensemble* terhadap observasi (Feldmann, 2012). Ketajaman adalah kekuatan penyebaran distribusi prediktif, yang mengindikasikan bahwa makin tajam suatu distribusi prediktif yang terkalibrasi, maka estimator semakin stabil (mampu mengoreksi bias lebih baik) dan performa ramalan semakin baik. Misal diketahui bahwa suhu udara pada tanggal 1 Januari 2014 adalah 26,7° C, maka interval prediktif (26° C, 27° C) dinilai lebih tajam daripada interval prediktif (25° C, 28° C). Menurut Möller (2014), *Root Mean Square Error* (RMSE) adalah indikator yang cukup untuk menilai kualitas prakiraan. Sementara, koreksi bias dapat dievaluasi melalui *coverage* dan *Continuous rank probability score* (CRPS).

2.9.1 Root Mean Square Error (RMSE)

RMSE merupakan salah satu indikator yang digunakan untuk mengevaluasi kualitas prakiraan cuaca, khususnya untuk data *testing*. RMSE merupakan akar kuadrat dari MSE, yaitu rata-rata jumlah kuadrat dari selisih antara nilai prakiraan dan observasi. Persamaan (2.42) merupakan formula untuk menghitung RMSE.

$$\text{RMSE} = \sqrt{\text{MSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.42)$$

dengan n adalah banyak observasi. Kelebihan RMSE terletak pada satuan perhitungan yang sama dengan satuan variabel, sehingga mudah untuk dijadikan sebagai alat perbandingan ketimbang MSE atau MAE. Namun, jika terdapat penciran yang cukup banyak pada observasi yang diamati, RMSE cenderung kurang *robust* terhadap keadaan tersebut (Anggraeni, 2013).

2.9.2 Coverage

Ketajaman dari prakiraan *ensemble* dapat didefinisikan melalui *coverage*, yaitu indikator yang menggunakan lebar *ensemble range* sebagai tolak ukur. Jika observasi berada dalam *ensemble range*, maka observasi tersebut dikatakan berada dalam *coverage*. Coverage dihitung dari persentase banyaknya observasi yang berada dalam *ensemble range*, khususnya observasi yang berada dalam periode *testing* atau *out-sample* (Raftery et al., 2005).

Agar memiliki pembandingan antara nilai observasi dan ramalan dari *ensemble* dengan m anggota, maka standar *coverage* dihitung berdasarkan $\frac{M - 1}{M + 1} \times 100\%$. Jadi, standar *coverage* bergantung pada banyaknya anggota *ensemble* yang dilibatkan dalam pemodelan. Makin banyak anggota *ensemble*, maka *coverage* yang harus dicapai juga makin tinggi (Feldmann, 2012).

Coverage umumnya digunakan untuk mengidentifikasi apakah prakiraan *ensemble* masih bersifat *underdispersive* atau *overdispersive*. Jika *coverage* data jauh lebih rendah daripada standar *coverage*, maka prakiraan *ensemble* masih bersifat *underdispersive*, namun sebaliknya jika *coverage* data lebih tinggi daripada standar *coverage*, maka prakiraan *ensemble* diindikasikan bersifat *overdispersive*. Prakiraan *ensemble* yang terkalibrasi terlihat dari *coverage* data yang nilainya mendekati standar *coverage*.

2.9.3 Continuous Rank Probability Score (CRPS)

Salah satu bagian terpenting untuk mengevaluasi kebaikan model prakiraan cuaca, baik deterministik atau probabilistik, adalah melalui *scoring rules*. RMSE yang diuraikan pada subbab 2.9.1 merupakan salah satu *scoring rules*. Selain itu, CRPS juga dapat memverifikasi model prakiraan dari segi tingkat koreksi bias dan

ketajaman yang tercapai. Ketajaman mengindikasikan bahwa PDF prediktif model terkalibrasi mampu memperbaiki PDF prediktif model sebelum kalibrasi, yaitu interval PDF model kalibrasi lebih sempit ketimbang sebelum dikalibrasi. Makin sempit interval yang diperoleh dimana sebagian besar nilai observasi berada dalam interval tersebut, maka hasil prakiraan dikatakan reliabel (Feldmann, 2012).

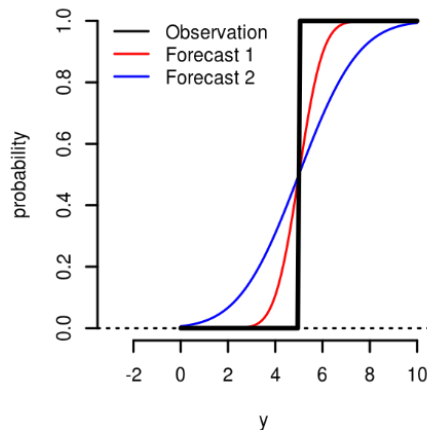
CRPS dirumuskan dalam persamaan (2.43),

$$CRPS = \frac{1}{n} \sum_{i=1}^n crps(F_i, y_i) = \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{\infty} [F_i^{forecast}(y) - F_i^{obs}(y)]^2 dy \quad (2.43)$$

dengan n adalah banyaknya observasi, i adalah periode waktu yang digunakan (jam, harian, bulanan, dan sebagainya), $F_i^{forecast}(y)$ adalah CDF prediktif pada waktu ke- i , dan $F_i^{obs}(y)$ adalah CDF empiris pada waktu ke- i (Anggraeni, 2013).

Khusus untuk $F_i^{obs}(y)$, berlaku bahwa $F_i^{obs}(y) = \begin{cases} 1, & y_{forecast} \geq y_{obs} \\ 0, & y_{forecast} < y_{obs} \end{cases}$

Makin rendah nilai CRPS, dalam hal ini mendekati 0, maka model tersebut makin baik dan reliabel dalam melakukan prakiraan cuaca. Gambar 2.2 merupakan CDF untuk prakiraan cuaca pada suatu waktu i dengan nilai observasi sebesar 5 satuan, dengan sumbu horizontal menyatakan rentang nilai prakiraan cuaca yang mungkin, dan sumbu vertikal menyatakan probabilitas dari CDF.



Gambar 2.2 CDF Observasi dan Prakiraan Waktu ke- i

Berdasarkan Gambar 2.2, terdapat dua CDF prakiraan yang diperoleh dari regresi PLS (berwarna merah), sementara CDF lainnya (berwarna biru diperoleh dari regresi PCR. Diasumsikan kedua CDF tersebut mengikuti distribusi normal. Jika diperhatikan, terlihat bahwa CDF PCR lebih landai ketimbang CDF PLS. Hal

ini mengindikasikan varians prakiraan PLS lebih kecil daripada varians prakiraan PCR. Meskipun observasi (CDF hitam) berada dalam interval prediktif kedua metode, namun interval PLS ternyata lebih sempit. Berdasarkan persamaan (2.32), hal ini mengakibatkan CRPS PLS menjadi lebih rendah daripada CRPS PCR yang mengindikasikan interval prediktif PLS lebih presisi atau reliabel daripada interval PCR, meskipun kedua metode memiliki akurasi yang hampir sama.

2.10 Numerical Weather Prediction (NWP)

Numerical Weather Prediction (NWP) adalah sekumpulan kode pemrograman yang mempresentasikan secara numerik persamaan-persamaan atmosfer berdasarkan sifat-sifat fisika dinamis, digunakan untuk memprediksi kondisi atau status NWP atmosfer yang akan datang dengan menggunakan kemampuan komputer yang tinggi (Idowu dan Rautenbach, 2009). Prakiraan cuaca dirumuskan dengan menyelesaikan persamaan pergerakan atmosfer. Persamaan-persamaan tersebut meliputi persamaan non-linier *time-dependent differential parsial* angin, temperatur, kelembapan dan tekanan.

NWP diukur dalam domain lokasi atau resolusi grid yang tinggi, berkisar antara 7 – 60 km, sehingga NWP diharapkan dapat menangkap fenomena cuaca yang sedang terjadi di domain tersebut. Namun, kondisi cuaca skala kecil atau skala lokal kurang terepresentasikan dengan baik. Oleh karena itu, luaran NWP memiliki sifat bias dalam meramalkan kondisi cuaca lokal sebab diukur pada domain yang tinggi. Selain itu, luaran NWP juga bersifat deterministik dan tidak bisa secara lengkap menjelaskan cuaca secara stokastik, sehingga perlu diterapkannya pemrosesan secara statistik (*statistical post-processing*) agar mampu menjelaskan ketidakpastian tersebut (Wilks, 2006).

Conformal Cubic Atmospheric Model (CCAM) adalah salah satu model aplikasi yang menghasilkan produk NWP. CCAM pertama kali dikembangkan oleh CSIRO (*Commonwealth Scientific and Industrial Research Organization*) Australia yang sebelumnya mengembangkan *Division of Atmospheric Research Limited Area Model* (DARLAM), kemudian CCAM diterapkan di Indonesia pada tahun 2007. Input yang diperlukan oleh CCAM adalah AVN/GFS. AVN/GFS adalah model spektral untuk prediksi cuaca global yang dijalankan oleh *National*

Centers for Environmental Prediction (NCEP). Model ini dapat memprediksi keadaan cuaca seluruh dunia hingga 2 minggu ke depan (BMKG, 2011). BMKG sendiri mulai mengembangkan penelitian tentang CCAM pada tahun 2007-2009.

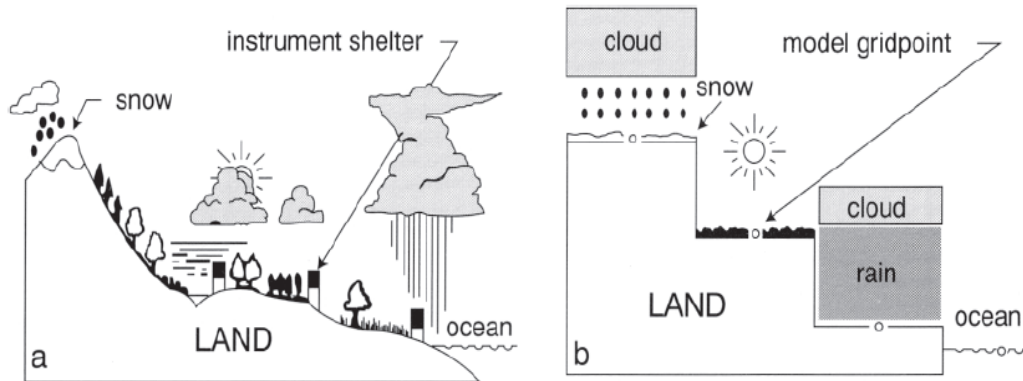
Menurut Raible, Bishop, dan Fraedrich (1998) dalam Arfianto (2008), secara umum model-model NWP cukup baik dalam peramalan jangka pendek (*short-term forecasting*) hingga 24 jam ke depan. NWP dicatat pada grid (kombinasi lintang-bujur) tertentu dengan deskripsi sebagai berikut.

1. Parameter NWP diantaranya *Surface Pressure tendency* (dpsdt), *Water Mixing Ratio* (mixr), *Geopotential Height* (Z), *Temperature* (T), *Relative Humidity* (RH), komponen U-V (komponen angin timur dan barat), *Mean Sea Level Pressure* (psl), *Vertical Velocity* (omega), *Maximum Screen Temperature* (tmaxscr), *Minimum Screen Temperature* (tminscr).
2. Level tekanan: 1000 mb, 950 mb, 925 mb, 900 mb, 850 mb, 800 mb, 700 mb, 600 mb, 500 mb, 400 mb, 350 mb, 300 mb dan 200mb. Level ketinggian: permukaan, 2 meter, dan 10 meter. NWP diukur pada level tekanan tertentu, dan dapat diukur pada level ketinggian : permukaan laut, 2 meter, dan 10 meter di atas permukaan laut.
3. Ramalan NWP dilakukan setiap 6 jam sekali, yaitu pada jam ke-00, 06, 12, 18, 24, 36, 42, 48, 54, 60, 66, dan 72.
4. Resolusi: grid lintang bujur $0,5^{\circ} \times 0,5^{\circ}$ ($55,5 \times 55,5$) km^2 . NWP diukur pada grid poin yang luas dengan ukuran lintang-bujur tertentu.

Grid merupakan proyeksi keadaan bumi dan atmosfer di atasnya yang berbentuk 3 dimensi. Grid pada umumnya menyerupai persegi atau persegi panjang, walaupun pada prosesnya grid dapat terbentuk segitiga atau segienam. Prakiraan NWP untuk tiap komponen cuaca, seperti suhu permukaan, kecepatan angin, tekanan udara, dapat dikeluarkan oleh tiap grid. Resolusi grid diatur sedemikian rupa oleh otoritas meteorologi yang disesuaikan dengan kepentingan peramalan jangka pendek hingga jangka panjang. Semakin tinggi resolusi yang digunakan, maka NWP akan menghasilkan peramalan jangka pendek yang lebih akurat, namun memakan waktu komputasi yang lebih lama (Wliks, 2006).

Hasil dari prediksi cuaca dengan NWP pada resolusi tinggi di suatu tempat (grid) seringkali menghasilkan bias yang besar, terutama untuk wilayah dengan

topografi dan tutupan vegetasi yang kompleks. Gambar 2.3 mengilustrasikan perbedaan kondisi cuaca riil di suatu wilayah dengan topografi yang beragam (a) dan grid model NWP (b).



Gambar 2.3 Ilustrasi Kondisi Cuaca (a) Dunia Nyata dan (b) NWP (Linacre dan Geerts, 2003)

Gambar 2.3 menunjukkan bahwa terdapat perbedaan kondisi cuaca yang signifikan antara dunia nyata dan model NWP. Hal ini tidak mengherankan sebab kondisi riil permukaan bumi dan atmosfer di suatu tempat bervariasi, baik dalam domain spasial (daerah), temporal (waktu) maupun elevasi (ketinggian). Sayangnya, model NWP justru belum menangkap fenomena di atas, padahal suatu area yang luas dipandang memiliki keragaman topografi dan keadaan atmosfer yang cenderung heterogen. Berdasarkan kenyataan ini, dapat dikatakan bahwa luaran NWP belum menjelaskan keragaman cuaca di wilayah yang lebih heterogen dari segi keadaan alam (Sari, 2013).

NWP menyederhanakan dan menyamaratakan kondisi permukaan agar homogen dengan merepresentasikan permukaan bumi ke dalam rangkaian titik grid. Dengan demikian, model NWP tidak merepresentasikan lokasi dan variabel yang dibutuhkan. Padahal, hubungan statistik justru dapat dibangun berdasarkan informasi yang dihasilkan NWP dan nilai prakiraan yang dimaksud. Informasi NWP yang digunakan bersamaan dengan metode statistik memungkinkan untuk mengkuantifikasi dan menjelaskan ketidakpastian dengan menghubungkannya berdasarkan kondisi prediksi yang berbeda atau probabilistik (Wilks, 2006).

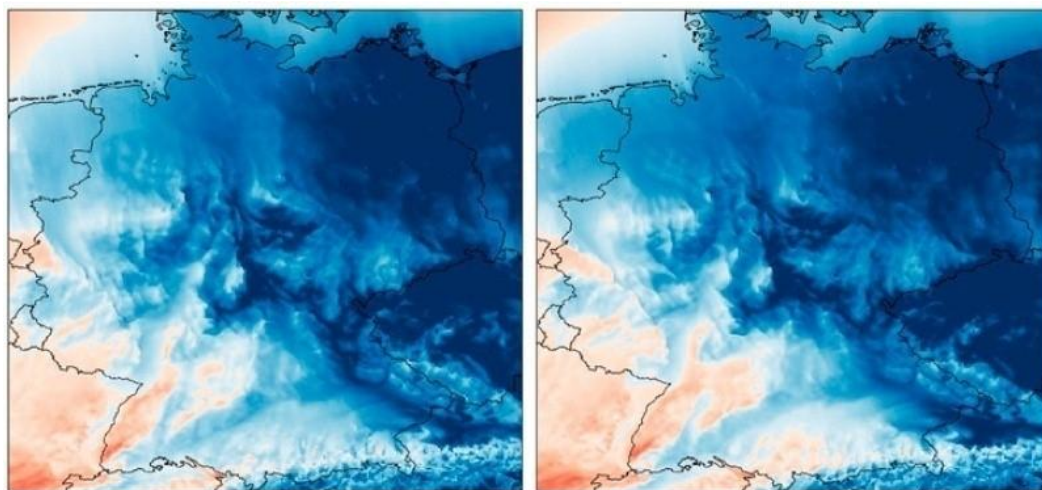
2.11 Sistem Prediksi *Ensemble*

Menurut Park (2006), sistem prediksi *ensemble* atau EPS (*ensemble prediction system*) adalah suatu sistem yang terdiri dari beberapa kombinasi model yang mengolah luaran tunggal deterministik (*deterministic forecast*) dalam waktu bersamaan. Salah satu contoh dari model prakiraan tunggal deterministik yang lazim dikembangkan oleh banyak negara adalah *Numerical Weather Prediction* (NWP). Model prediksi *ensemble* cukup banyak digunakan dalam bidang iklim dan sains atmosfer dimana hasil luaran model ini secara operasional memiliki kualitas yang cukup baik, meskipun kurang reliabel karena masih banyak nilai observasi yang jauh berada di luar *ensemble range* (*coverage* cenderung rendah). Prediksi *ensemble* dapat dieksekusi dengan beragam cara, antara lain dengan penggunaan model yang berbeda dan resolusi yang beragam, mulai resolusi rendah hingga resolusi tinggi (Viney, Croke, Breuer, Bormann, Bronstert, Freed, Graff, Hubrechts, Huisman, Jakeman, Kite, Lanini, Leavesley, Lettenmaier, Lindstrom, Seibert, Sivapalan, dan Willem, 2006). Selain itu, prediksi *ensemble* juga dapat dikonstruksi dengan penggunaan syarat awal (*initial condition*) perturbasi yang berbeda, sehingga akan membentuk beberapa anggota *ensemble* untuk prakiraan komponen cuaca tertentu.

Meskipun prediksi yang dihasilkan NWP saat ini makin mendekati observasi, namun tidak menutup kemungkinan diperoleh bias prediksi yang tinggi akibat keadaan atmosfer yang selalu dinamis dan berfluktuasi. Dalam Viney *et al.* (2005), penggunaan metode *ensemble* dalam prakiraan cuaca/iklim mampu memiliki performa yang lebih baik daripada NWP. Selain itu, Demeritt, Cloke, Pappenberger, Thielen, Bartholmes, dan Ramos (2007) menyimpulkan bahwa prakiraan *ensemble* lebih sesuai dengan kondisi riil. Namun demikian, terdapat dua sumber ketidakpastian yang dipertimbangkan agar model *ensemble* menghasilkan prakiraan yang reliabel dan tajam, yaitu syarat awal (*initial condition*) dan konstruksi model. Berdasarkan Anggraeni (2013), prakiraan *ensemble* diawali perumusan deterministik yang terdiri dari beragam variabel cuaca dan iklim (temperatur permukaan, tekanan udara, kelembapan, dan lain-lain) dimana masing-masing memiliki *initial condition* yang berlainan, kemudian disimulasikan secara iteratif. Pada tahap konstruksi model, terdapat beberapa teknik yang dapat

digunakan, antara lain skema perturbasi. Skema ini mengurangi unsur ketidakpastian model melalui modifikasi nilai antar anggota *ensemble*. Dengan demikian, prakiraan *ensemble* terkalibrasi dapat mendekati nilai observasi dan sesuai dengan kondisi riil yang sedang terjadi.

Saat ini, tidak banyak negara yang mampu dan konsisten dalam mengembangkan sistem prediksi *ensemble* (EPS) karena diperlukan sumber daya yang berkualitas, biaya yang besar dan teknologi yang canggih untuk membangkitkan data *ensemble* per satuan waktu ke depan (*lead time ahead*). Semakin singkat *lead* prediksi, maka proses komputasi dan spesifikasi pola atmosfer yang dilakukan akan makin kompleks (Demeritt *et al.*, 2007). Beberapa otoritas yang menerapkan sistem prediksi *ensemble* untuk komponen cuaca, diantaranya *National Centers for Environmental Prediction (NCEP) of US, Météo-France, European Centre for Medium-Range Weather Forecast (ECMWF)*, dan *Japan Meteorological Agency*. Gambar 2.4 adalah ilustrasi prediksi *ensemble* untuk elemen cuaca temperatur untuk wilayah Jerman yang dikeluarkan COSMO-DE, otoritas Jerman yang memiliki kewenangan mirip dengan BMKG Indonesia.



Gambar 2.4 Prediksi 2 Model *Ensemble* untuk Temperatur Wilayah Jerman 5 Januari 2011 (Feldmann, Scheurer, dan Thorarinsdottir, 2015)

Di Indonesia, sistem ini pernah dikembangkan oleh BMKG selama beberapa tahun, namun prakiraan yang dihasilkan tidak lebih baik dibandingkan beberapa metode lainnya. Biaya yang tinggi juga menjadi kendala bagi BMKG untuk menggunakan sistem prediksi *ensemble* secara berkelanjutan.

2.12 Suhu/Temperatur

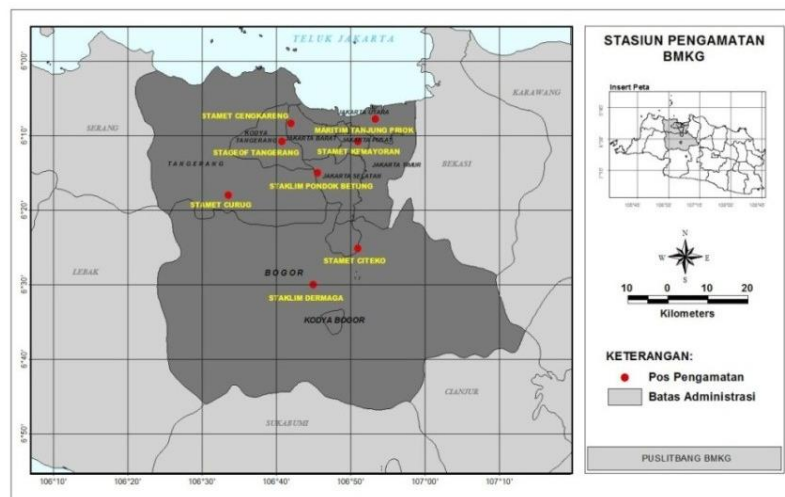
Suhu atau temperatur merupakan suatu ukuran yang menyatakan termalitas beberapa sistem baik secara parsial maupun integral atau dengan kata lain ukuran panas atau dinginnya suatu benda (Quinn, 1990). Sementara, untuk lingkup yang lebih spesifik, suhu atmosfer merupakan ukuran suhu pada berbagai tingkatan di atmosfer Bumi yang dapat dipengaruhi radiasi matahari, kelembapan serta ketinggian, sehingga mengakibatkan adanya hubungan kompleks antara biosfer, litosfer dan atmosfer. Energi secara konstan berpindah dari permukaan ke udara di atasnya.

Menurut Tanudidjaja (1993), keadaan suhu atmosfer pada suatu tempat di permukaan bumi ditentukan oleh beragam faktor, seperti lama penyinaran matahari, kemiringan sinar matahari, keadaan awan, dan keadaan permukaan bumi. Misal semakin lama matahari memancarkan sinarnya di suatu daerah, maka semakin banyak panas yang diterima oleh daerah tersebut. Keadaan atmosfer yang cerah sepanjang hari menyebabkan temperatur di daerah tersebut lebih panas daripada jika kondisi cuaca pada hari itu berawan atau hujan sejak pagi.

BAB III METODE PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari Badan Meteorologi, Klimatologi, dan Geofisika (BMKG), yaitu data observasi suhu udara (suhu maksimum dan suhu minimum) harian untuk periode 1 Januari 2009 sampai 31 Desember 2010 atau selama 708 hari. Untuk membentuk model BMA dan GOP, diperlukan data harian NWP model CCAM pada periode yang sama dengan observasi. Stasiun pengamatan yang menjadi fokus dalam penelitian ini adalah 8 stasiun meteorologi (stamet) di Jabotabek, yaitu stamet Kemayoran, Tanjung Priok, Cengkareng, Pondok Betung, Curug, Tangerang, Citeko, dan Dermaga yang ditunjukkan pada Gambar 3.1.



Gambar 3.1 Lokasi Stasiun Meteorologi (BMKG, 2011)

Seluruh stasiun meteorologi tersebut dipilih karena data histori pengukuran cuaca pada stasiun tersebut cukup baik dan lengkap. Selain itu, beberapa stasiun tersebut dianggap mewakili kota Jakarta (Kemayoran), kota Tangerang (Tangerang), kota Tangerang Selatan (Pondok Betung) dan kabupaten Bogor (Citeko). Karakteristik atmosfer yang dinamis memperbesar kemungkinan bahwa kondisi cuaca di DKI Jakarta juga dipengaruhi kondisi cuaca di daerah sekitar, yaitu Bogor dan Tangerang.

3.2 Variabel Penelitian

Variabel respon yang digunakan dalam penelitian ini adalah data observasi cuaca untuk suhu udara (Celcius), yaitu suhu maksimum (T_{MAKS}) dan suhu minimum (T_{MIN}). Prediktor yang digunakan pada Spatial BMA adalah prakiraan suhu udara dimodelkan dengan PLS, PCR, dan Regresi Ridge. Prediktor model *ensemble* diperoleh dari luaran parameter NWP yang direduksi menggunakan PCA.

Parameter NWP CCAM ditentukan oleh ahli meteorologi. Selain itu, BMKG dalam beberapa tahun terakhir sudah mengadakan penelitian yang intensif dengan CSIRO Australia sebagai lembaga yang mengembangkan CCAM. Parameter NWP CCAM untuk metode BMA, GOP, dan Spatial BMA disajikan pada Tabel 3.1.

Tabel 3.1 Parameter NWP CCAM yang Digunakan

Parameter NWP (kode)	Level	X	Satuan
<i>Surface Pressure Tendency</i> (dpsdt)	Permukaan	X_1	hPa
<i>Water Mixing Ratio</i> (mix)	1, 2, 4	X_2, X_3, X_4	g/kg
<i>Vertical Velocity</i> (omega)	1, 2, 4	X_5, X_6, X_7	Knot
<i>PBL depth</i> (pblh)	Permukaan	X_8	Meter
<i>Surface Pressure</i> (ps)	Permukaan	X_9	hPa
<i>Mean Sea Level Pressure</i> (psl)	Permukaan	X_{10}	hPa
<i>Screen Mixing Ratio</i> (qgscm)	Permukaan	X_{11}	g/kg
<i>Relative Humidity</i> (rh)	1, 2, 4	X_{12}, X_{13}, X_{14}	%
<i>Precipitation</i> (rnd)	Permukaan	X_{15}	Mm
<i>Temperature</i>	1, 2, 4	X_{16}, X_{17}, X_{18}	Celcius
<i>Maximum Screen Temperature</i> (tmaxscr)	Permukaan	X_{19}	Celcius
<i>Maximum Screen Temperature</i> (tminscr)	Permukaan	X_{20}	Celcius

Tabel 3.1 Parameter NWP CCAM yang Digunakan (Lanjutan)

Parameter NWP (kode)	Level	X	Satuan
<i>Pan Temperature</i> (tpan)	Permukaan	X ₂₁	Celcius
<i>Screen Temperature</i> (tscrn)	Permukaan	X ₂₂	Celcius
<i>Zonal Wind</i> (u)	1, 2, 4	X ₂₃ , X ₂₄ , X ₂₅	Knot
<i>Friction Velocity</i> (ustar)	Permukaan	X ₂₆	m/sec
<i>Meridional Wind</i> (v)	1, 2, 4	X ₂₇ , X ₂₈ , X ₂₉	Knot
<i>Geopotential Height</i> (zg)	1, 2, 4	X ₃₀ , X ₃₁ , X ₃₂	Meter

(Sumber : BMKG, 2011)

Ada sebanyak 18 parameter NWP yang akan digunakan di tiap stamet dengan 11 parameter diukur pada level permukaan, yaitu pada ketinggian ± 2 m di atas permukaan laut. Sementara 7 parameter lainnya diukur pada tiga level tekanan yang berbeda, yaitu level 1, 2, dan 4, dimana level 1 saat tekanan 1000 mb (milibar), level 2 saat tekanan 950 mb dan level 4 saat tekanan 850 mb. Jadi total parameter yang digunakan sebanyak (11 + 21) atau 32 parameter karena 7 parameter yang diukur pada 3 level berbeda akan menghasilkan (7 x 3) parameter atau 21 parameter. Kemudian, tiap 32 parameter diestimasi pada sembilan grid (3 x 3) pengukuran yang terdekat dari lokasi stasiun pengamatan. Proyeksi pengukuran parameter NWP dalam grid 3 x 3 diilustrasikan oleh Gambar 3.2.



Gambar 3.2 Proyeksi Pengukuran NWP

Pada Gambar 3.2, lingkaran merah menunjukkan grid terdekat dari stamet Kemayoran (diilustrasikan oleh balon udara), sedangkan lingkara berwarna hitam merupakan kombinasi grid di sekitar stasiun pengamatan. Jadi, tiap stasiun yang diamati memiliki total parameter sebanyak (32 x 9) parameter atau 288 parameter. Sehingga, tiap parameter akan memiliki korelasi yang kuat dengan dirinya sendiri karena diukur pada 9 grid yang saling berdekatan. Antar parameter NWP diduga

juga saling berkorelasi karena hubungan antar tiap komponen yang cenderung tinggi.

Berikut ini merupakan definisi parameter NWP yang diolah sebagai *input* prediktor (NOAA, 2015).

1. *Surface Pressure Tendency* atau kecenderungan tekanan permukaan merupakan suatu indikasi arah dan intensifikasi dari disturbansi siklon.
2. *Water Mixing Ratio* adalah rasio jumlah uap air di udara pada ketinggian tertentu.
3. *Vertical Velocity* adalah kecepatan angin vertikal yang diukur.
4. *PBL (Planetary Boundary Layer) depth* merupakan suatu skala ketinggian yang mendeskripsikan vertikal *mixing* dari turbulensi dan konveksi awan kumulus.
5. *Surface Pressure* atau tekanan permukaan diukur berdasarkan tekanan gaya pada permukaan dengan luas tertentu.
6. *Mean Sea Level Pressure* adalah suhu rata-rata di atas permukaan laut.
7. *Scree Mixing Ratio* adalah rasio jumlah uap air di udara pada level permukaan, yaitu 2 meter di atas permukaan air laut.
8. *Relative Humidity* atau kelembaban adalah konsentrasi uap air di udara.
9. *Precipitation* adalah setiap produk dari kondensasi uap air di atmosfer yang jatuh karena gravitasi.
10. *Temperature* atau suhu udara adalah derajat panas dan dingin udara di atmosfer, yang memiliki hubungan berbanding terbalik dengan tekanan udara.
11. *Maximum Screen Temperature* merupakan suhu tertinggi yang terukur pada grid-grid.
12. *Minimum Screen Temperature* merupakan suhu terendah yang terukur pada grid-grid.
13. *Pan temperature* adalah suhu udara di atas sumber air yang terdekat dengan grid-grid.
14. *Screen temperature* adalah derajat panas dan dingin udara pada grid-grid
15. *Zonal wind* atau zonal merupakan komponen angin yang bergerak dengan arah barat-timur.

16. *Friction Velocity* atau kecepatan gesekan adalah bentuk tegangan geser dan dapat ditulis dalam bentuk satuan kecepatan.
17. *Meridional Wind* atau angin meridian merupakan komponen angin yang bergerak dengan arah utara-selatan
18. *Geopotential Height* adalah koordinat vertikal yang direferensikan ke permukaan laut bumi dengan menggunakan variasi gravitasi dengan garis lintang dan ketinggian.

Berdasarkan Tabel 3.1 dan Gambar 3.2, diperoleh p sebanyak 32 parameter dan r sebanyak 9 grid. Ilustrasi struktur data yang akan diolah dengan metode BMA tersaji dalam Tabel 3.2. Struktur data untuk satu stamet s yang terletak pada koordinat lintang-bujur tertentu, dengan respon Y_t adalah observasi suhu udara pada hari ke t dan elemen variabel $X_{p,r,t}$ merupakan luaran NWP parameter p grid ke- r untuk hari ke- t .

Tabel 3.2 Struktur Data untuk Metode BMA pada Stasiun s

Tanggal	Y	X ₁			X ₂			...	X _p		
		X _{1,1}	...	X _{1,9}	X _{2,1}	...	X _{2,9}		X _{p,1}	...	X _{p,r}
1/1/2009	Y ₁	X _{1,1,1}	...	X _{1,9,1}	X _{2,1,1}	...	X _{2,9,1}	...	X _{p,1,1}	...	X _{p,r,1}
2/1/2009	Y ₂	X _{1,1,2}	...	X _{1,9,2}	X _{2,1,2}	...	X _{2,9,2}	...	X _{p,1,2}	...	X _{p,r,2}
3/1/2009	Y ₃	X _{1,1,3}	...	X _{1,9,3}	X _{2,1,3}	...	X _{2,9,3}	...	X _{p,1,3}	...	X _{p,r,3}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
29/12/2010	Y ₇₀₆	X _{1,1,706}	...	X _{1,9,706}	X _{2,1,706}	...	X _{2,9,706}	...	X _{p,1,706}	...	X _{p,r,706}
30/12/2010	Y ₇₀₇	X _{1,1,707}	...	X _{1,9,707}	X _{2,1,707}	...	X _{2,9,707}	...	X _{p,1,707}	...	X _{p,r,707}
31/12/2010	Y ₇₀₈	X _{1,1,708}	...	X _{1,9,708}	X _{2,1,708}	...	X _{2,9,708}	...	X _{p,1,708}	...	X _{p,r,708}

Sementara, data yang digunakan GOP untuk meramalkan suhu udara memiliki perbedaan struktur dengan data yang diolah dengan BMA. Tabel 3.3 merupakan struktur data untuk stasiun meterologi s yang berada pada koordinat lintang-bujur tertentu. Respon Y_t adalah observasi suhu udara pada hari ke- t dan variabel $X_{s,t}$ merupakan luaran/prakiraan model NWP untuk suhu udara pada grid terdekat atau grid yang bersesuaian dari stasiun meteorologi s pada hari ke- t .

Tabel 3.3 Struktur Data untuk Metode GOP pada Stasiun s

Tanggal	Y	X
1/1/2009	Y_1	$X_{s,1}$
2/1/2009	Y_2	$X_{s,2}$
3/1/2009	Y_3	$X_{s,3}$
\vdots	\vdots	\vdots
29/12/2010	Y_{706}	$X_{s,706}$
30/12/2010	Y_{707}	$X_{s,707}$
31/12/2010	Y_{708}	$X_{s,708}$

Struktur data yang digunakan untuk meramalkan suhu menggunakan Spatial BMA tidak berbeda dengan struktur data yang disajikan dalam Tabel 3.2. Namun setelah dilakukan pemilihan model *ensemble* terbaik berdasarkan bobot BMA, digunakan struktur data pada Tabel 3.3. Prediktor (X) yang digunakan hasil prakiraan model *ensemble* terpilih yang diolah secara simultan (dari semua stamet).

3.3 Kajian Teoritis

Langkah-langkah analisis yang diperlukan untuk mencapai tujuan penelitian teoritis pada Bab 1 adalah sebagai berikut.

- a. Mengestimasi parameter regresi untuk tiap anggota *ensemble* m , yaitu β_m , dengan regresi linier sederhana menggunakan *Maximum Likelihood Estimation* (MLE) distribusi normal multivariat. Berikut ini merupakan fungsi *likelihood* yang digunakan.

$$L(\beta_m, \Sigma_m) = \prod_{t=1}^T \frac{1}{(2\pi)^{\frac{s}{2}} |\Sigma_m|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} [\mathbf{y}_t - \mathbf{F}_m \beta_m]^T \Sigma_m^{-1} [\mathbf{y}_t - \mathbf{F}_m \beta_m] \right\}$$

- b. Mengestimasi parameter bobot untuk tiap anggota *ensemble* m dan varians menggunakan metode MLE dengan algoritma *Expectation-Maximization* (EM) dengan langkah berikut.

- 1) Jika $\theta = [w_1 \ w_2 \ \dots \ w_M, \sigma^2]^T$, mendapatkan fungsi likelihood untuk

$$\text{data lengkap } L(\theta; \mathbf{y}, \mathbf{z}) = \prod_{t=1}^T \sum_{m=1}^M \mathbf{I}_{(z_t=m)} w_m g_m(y_t | f_{mt}) \text{ dimana } \mathbf{I}_{(z_t=m)}$$

akan bernilai 1 jika $z_t = m$ dan bernilai 0 untuk anggota *ensemble* lainnya selain m .

- 2) Untuk iterasi ke- i , mencari nilai ekspektasi dari fungsi log-likelihood terhadap distribusi bersyarat dari \mathbf{Z} diberikan \mathbf{Y} berdasarkan estimasi parameter iterasi ke- i $\boldsymbol{\theta}^{(i)}$. (Tahap *Expectation*)

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)}) = E_{\boldsymbol{\theta}^{(i)}}(\ln L(\boldsymbol{\theta}; \mathbf{y}, \mathbf{z}))$$

Umumnya, nilai awal untuk $\boldsymbol{\theta}$, yaitu $\boldsymbol{\theta}^{(0)}$, didefinisikan sebagai

$$w_m^{(0)} = \frac{1}{M} \text{ dan } \sigma^{2(0)} = \frac{1}{T} \sum_{t=1}^T \sum_{m=1}^M (y_t - f_{mt})^2.$$

- 3) Memperoleh estimasi parameter $\boldsymbol{\theta}^{(i+1)}$ yang memaksimalkan $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})$ dimana $Q(\boldsymbol{\theta}^{(i+1)} | \boldsymbol{\theta}^{(i)}) \geq Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})$, sehingga diperoleh persamaan (2.22) dengan memanfaatkan observasi laten $z_{mt}^{(i+1)}$ berdasarkan persamaan (2.21). (Tahap *Maximization*)
- 4) Proses iterasi berhenti jika $L(\boldsymbol{\theta}^{(i+1)}; \mathbf{y}) \geq L(\boldsymbol{\theta}^{(i)}; \mathbf{y})$ sebesar nilai yang sangat rendah untuk serangkaian tahap iterasi.

c. Mengestimasi parameter kovarians spasial menggunakan metode *Weighted Least Square* (WLS)

- 1) Untuk tiap hari periode training, hitung parameter model yang diperoleh dengan mengurangi nilai observasi terhadap prakiraan $\varepsilon_{s,t} = y_{s,t} - \hat{y}_{s,t}$ dimana $y_{s,t} = \beta_0 + \beta_1 x_{s,t} + \varepsilon_{s,t}$.
- 2) Untuk tiap hari dalam periode training dan untuk tiap himpunan pasangan lokasi pengamatan pada hari tersebut, hitung jarak antar dua lokasi dengan jarak Euclidian $\|s_i - s_j\|$ dan hitung setengah dari selisih kuadrat antar error $\frac{1}{2} [\varepsilon(s_i) - \varepsilon(s_j)]^2$.
- 3) Mengelompokkan jarak-jarak yang sudah diperoleh ke dalam *bin* B_l dengan titik tengah d_l dan l adalah suatu jarak.
- 4) Menghitung nilai semivariogram empiris $\hat{\gamma}(d_l)$ pada jarak d_l dengan menghitung setengah dari ekspektasi kuadrat selisih antar error pada

$bin B_l, \frac{1}{2} E[\varepsilon(s_i) - \varepsilon(s_j)]^2$, dimana s_i dan s_j adalah himpunan lokasi berbeda yang berada dalam $bin B_l$.

- 5) Mengestimasi parameter ρ_m^2, τ_m^2 , dan r_m^2 menggunakan metode *Weighted Least Square* (WLS) dengan meminimumkan nilai

$$g(\rho_m^2, \tau_m^2, r_m^2) = \sum_{l=1}^k n_l \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right)^2, \rho_m^2, \tau_m^2, r_m^2 \geq 0$$

dengan n_l adalah banyak pasangan lokasi yang berada pada $bin B_l$ dan N adalah jumlah bin yang terbentuk.

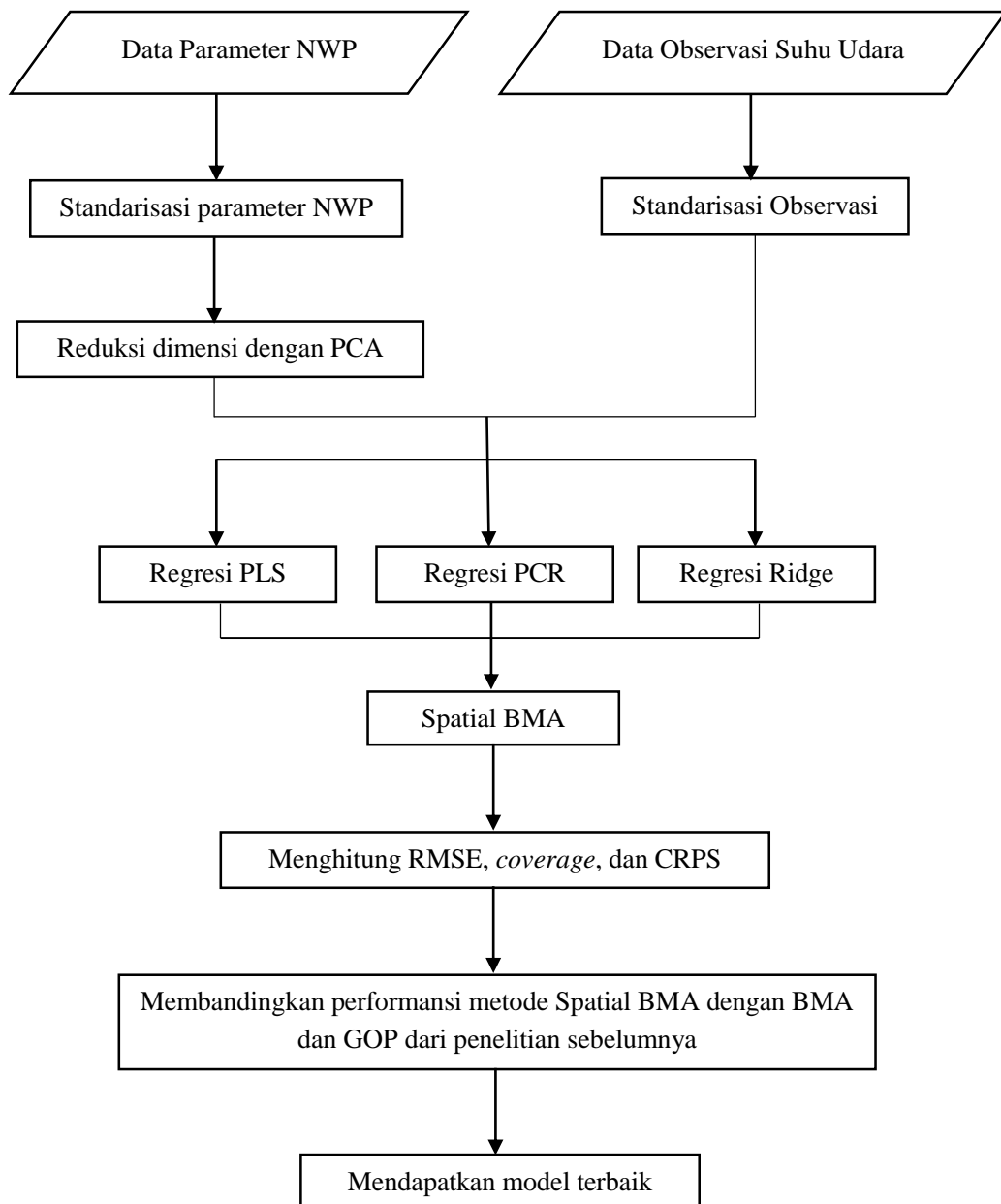
3.4 Kajian Aplikasi

Langkah-langkah analisis yang dibutuhkan untuk mencapai tujuan penelitian aplikatif pada Bab 1 adalah sebagai berikut.

- a. Melakukan standarisasi data, baik untuk data observasi maupun data parameter NWP. Teknik standarisasi yang diterapkan adalah mengurangi nilai tiap pengamatan dengan rata-rata, kemudian dibagi dengan akar variansnya (*scaled and centered*).
- b. Mereduksi dimensi tiap parameter NWP dengan PCA berdasarkan matriks kovarians sehingga diperoleh skor-skor PC.
- c. Melakukan prakiraan suhu udara dengan observasi riil sebagai respon dan skor PC sebagai prediktor berdasarkan pemodelan dengan PLS, PCR, dan Regresi Ridge.
- d. Mengkalibrasi prakiraan *ensemble* untuk suhu udara harian menggunakan Spatial BMA berdasarkan *training window* 30 hari dengan tahapan :
 - 1) Meregresikan observasi riil terhadap prakiraan *ensemble* menggunakan *Maximum Likelihood* untuk mendapatkan koefisien bias regresi, yaitu $\beta_{0,m}$ dan $\beta_{1,m}$. Langkah ini dilakukan untuk setiap anggota *ensemble m*.
 - 2) Mengestimasi parameter bobot untuk tiap anggota *ensemble* w_m dan varians σ^2 dengan algoritma *Expectation-Maximization* dan menginterpretasikan parameter tersebut, khususnya bobot w_m .

- 3) Menghitung mean, prakiraan BMA deterministik untuk suatu hari tertentu untuk seluruh lokasi secara simultan.
- 4) Menganalisis dan menguji signifikansi dependensi spasial suhu udara pada delapan stamet dengan Moran's I berdasarkan *cut-off distance* 30 km dan tingkat signifikansi α sebesar 0,05.
- 5) Meregresikan observasi riil terhadap luaran prakiraan BMA simultan menggunakan *Maximum Likelihood* untuk mendapatkan koefisien bias regresi anggota *ensemble* terpilih m , yaitu $\beta_{0,m}$ dan $\beta_{1,m}$.
- 6) Menganalisis semivariogram empiris berdasarkan residual regresi.
- 7) Mengestimasi parameter spasial pada variogram ρ_m^2 , τ_m^2 , dan r_m^2 berdasarkan residual yang telah di-*binning* (dikelompokkan), dengan metode kuadrat terkecil terboboti dimana bobot yang digunakan adalah banyaknya observasi dalam tiap kelompok *binning*.
- 8) Memperoleh prakiraan suhu udara untuk hari tertentu dan interval prediktif model Spatial BMA
- 9) Menganalisis kebaikan prakiraan di hari tersebut berdasarkan RMSE dan CRPS serta membandingkannya dengan anggota *ensemble*.
- 10) Membentuk *Verification Rank Histogram* (VRH) selama periode testing untuk melihat apakah prakiraan Spatial BMA sudah terkalibrasi atau masih bersifat *underdispersive* atau *overdispersive*.
- 11) Membandingkan performansi hasil prakiraan cuaca menggunakan kalibrasi *ensemble* Spatial BMA dengan BMA dan GOP.

Untuk mempermudah pemahaman, ilustrasi mengenai tahapan analisis untuk tujuan penelitian kedua dan ketiga disajikan pada Gambar 3.3.



Gambar 3.3 Tahapan Analisis Data

BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini, akan dibahas mengenai estimasi parameter *Spatial Bayesian Model Averaging* (Spatial BMA) menggunakan *Maximum Likelihood Estimation* (MLE) untuk parameter β_m dan algoritma *Expectation Maximum* untuk parameter bobot dan varians. Selain itu, dilakukan pula estimasi parameter kovarians spasial menggunakan *Weighted Least Square* (WLS) dengan iterasi L-BFGS. Kemudian, model tersebut digunakan untuk meramalkan suhu udara di 8 stasiun meteorologi (stamet) di Jabotabek secara simultan.

4.1 Penaksiran Parameter Spatial BMA

Parameter Spatial BMA dapat diklasifikasikan ke dalam 3 bagian, yaitu parameter regresi β_m , parameter bobot dan varians (w_m dan σ^2), dan parameter kovarians spasial (ρ_m^2 , τ_m^2 , dan r_m^2). Seluruh parameter tersebut berkontribusi terhadap koreksi bias yang dihitung dari tiap anggota *ensemble* dan digunakan untuk menghasilkan ramalan terkalibrasi beserta selang kepercayaan $(1 - \alpha) \times 100\%$ yang disesuaikan dengan distribusi dari data pengamatan. Berikut ini adalah penjabaran untuk memperoleh estimasi parameter-parameter tersebut.

4.1.1 Estimasi Parameter Regresi

Untuk tiap anggota *ensemble* m dimana $m = 1, 2, \dots, M$, $t = 1, 2, \dots, T$, $s = 1, 2, \dots, S$ dengan M , T dan S menyatakan banyaknya anggota *ensemble*, jumlah waktu/hari dalam periode training dan banyak lokasi, berlaku model pada persamaan (4.1) :

$$\mathbf{y} = \beta_{0,m} \mathbf{1} + \beta_{1,m} \mathbf{F}_{mt} + \boldsymbol{\varepsilon}_{1m} + \boldsymbol{\varepsilon}_{2m}. \quad (4.1)$$

Persamaan (4.1) merupakan model regresi multivariat, dimana $\mathbf{y}_t = \{y_t(s) : s \in S\}$ adalah variabel respon (observasi suhu udara) pada waktu ke- t dan lokasi ke- s , serta terdapat dependensi antar lokasi. Variabel dependen \mathbf{y}_t dinyatakan dalam bentuk matriks berukuran $s \times t$ sebagai berikut,

$s \backslash t$	1	2	...	708
1	$y_{1,1}$	$y_{1,2}$...	$y_{1,708}$
2	$y_{2,1}$	$y_{2,2}$...	$y_{2,708}$
\vdots	\vdots	\vdots	\vdots	\vdots
8	$y_{8,1}$	$y_{8,2}$...	$y_{8,708}$

$\mathbf{F}_{mt} = \{f_{mt}(s) : s \in S\}$, adalah variabel prediktor (prakiraan suhu udara maksimum dan minimum oleh anggota *ensemble m*), pada waktu ke- t dan lokasi ke- s . Variabel independen \mathbf{F}_{mt} dinyatakan dalam bentuk matriks berukuran $s \times t$ sebagai berikut,

$s \backslash t$	1	2	...	708
1	$f_{1,1}$	$f_{1,2}$...	$f_{1,708}$
2	$f_{2,1}$	$f_{2,2}$...	$f_{2,708}$
\vdots	\vdots	\vdots	\vdots	\vdots
8	$f_{8,1}$	$f_{8,2}$...	$f_{8,708}$

Demikian pula untuk $\boldsymbol{\varepsilon}_{1m} = \{\varepsilon_{1m}(s) : s \in S\}$ dan $\boldsymbol{\varepsilon}_{2m} = \{\varepsilon_{2m}(s) : s \in S\}$ adalah bagian kontinu dan diskrit dari eror bersyarat pada anggota *ensemble m* yang saling independen, seluruhnya pada waktu ke- t dan lokasi ke- s . Untuk kasus suhu udara, respon \mathbf{Y} atau eror $\boldsymbol{\varepsilon}_{1m}$ dan $\boldsymbol{\varepsilon}_{2m}$ masing-masing berdistribusi normal multivariat.

Persamaan (4.1) pada masing-masing pengamatan ke- t dan pada anggota *ensemble m*, dapat dituliskan sebagai model regresi multivariat sebagai berikut.

$$\begin{aligned}
y_{1t} &= \beta_{0,m} + \beta_{1,m} f_{m1t} + \varepsilon_{1m1t} + \varepsilon_{2m1t} \\
y_{2t} &= \beta_{0,m} + \beta_{1,m} f_{m2t} + \varepsilon_{1m2t} + \varepsilon_{2m2t} \\
&\vdots \\
y_{8t} &= \beta_{0,m} + \beta_{1,m} f_{m8t} + \varepsilon_{1m8t} + \varepsilon_{2m8t}
\end{aligned}$$

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{8t} \end{pmatrix}_{(8 \times 1)} = \begin{pmatrix} 1 & f_{m1t} \\ 1 & f_{m2t} \\ \vdots & \vdots \\ 1 & f_{m8t} \end{pmatrix}_{(8 \times 2)} \begin{pmatrix} \beta_{0,m} \\ \beta_{1,m} \end{pmatrix}_{(2 \times 1)} + \begin{pmatrix} \varepsilon_{1m1t} + \varepsilon_{2m1t} \\ \varepsilon_{1m2t} + \varepsilon_{2m2t} \\ \vdots \\ \varepsilon_{1m8t} + \varepsilon_{2m8t} \end{pmatrix}_{(8 \times 1)}$$

$$\mathbf{y}_t = \mathbf{F}_{mt} \boldsymbol{\beta}_m + \boldsymbol{\varepsilon}_{mt} \quad (4.2)$$

Berdasarkan Rencher dan Schaalje (2008), beberapa asumsi yang seharusnya dipenuhi adalah :

- a. $E(\boldsymbol{\epsilon}_{m_t}) = 0$ atau $E(\mathbf{y}_t) = \mathbf{F}_{m_t} \boldsymbol{\beta}_m$
b. $\text{cov}(\boldsymbol{\epsilon}_{m_t}) = \boldsymbol{\Sigma}_m$ atau $\text{cov}(\mathbf{y}_t) = \boldsymbol{\Sigma}_m$

Jadi, dapat dikatakan bahwa $\mathbf{y}_t \sim N_S(\mathbf{F}_{m_t} \boldsymbol{\beta}_m, \boldsymbol{\Sigma}_m)$ atau $\mathbf{E}_{m_t} \sim N_S(0, \boldsymbol{\Sigma}_m)$.

Setelah mengetahui distribusi dari respon atau eror beserta karakteristik variabel yang diamati, maka langkah selanjutnya adalah memperoleh estimator parameter regresi pada anggota *ensemble* m , yaitu $\hat{\boldsymbol{\beta}}_m$, berdasarkan fungsi probabilitas berikut yang mengacu persamaan (2.2).

$$g(\mathbf{y}_t) = \frac{1}{(2\pi)^{\frac{S}{2}} |\boldsymbol{\Sigma}_m|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m]^T \boldsymbol{\Sigma}_m^{-1} [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m] \right\}$$

Karakteristik $\hat{\boldsymbol{\beta}}_m$ adalah yang memaksimumkan likelihood $L(\boldsymbol{\beta}_m, \boldsymbol{\Sigma}_m)$ dimana

$$\begin{aligned} L(\boldsymbol{\beta}_m, \boldsymbol{\Sigma}_m) &= \prod_{t=1}^T g(\mathbf{y}_t) \\ &= \prod_{t=1}^T \frac{1}{(2\pi)^{\frac{S}{2}} |\boldsymbol{\Sigma}_m|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m]^T \boldsymbol{\Sigma}_m^{-1} [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m] \right\}. \quad (4.3) \\ &= \frac{\exp \left\{ -\frac{1}{2} \sum_{t=1}^T [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m]^T \boldsymbol{\Sigma}_m^{-1} [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m] \right\}}{(2\pi)^{\frac{TS}{2}} |\boldsymbol{\Sigma}_m|^{\frac{T}{2}}} \end{aligned}$$

Berikut merupakan fungsi ln-likelihood dari persamaan (4.2).

$$\begin{aligned} l(\boldsymbol{\beta}_m, \boldsymbol{\Sigma}_m) &= \ln L(\boldsymbol{\beta}_m, \boldsymbol{\Sigma}_m) \\ &= \ln \left\{ \frac{\exp \left\{ -\frac{1}{2} \sum_{t=1}^T [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m]^T \boldsymbol{\Sigma}_m^{-1} [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m] \right\}}{(2\pi)^{\frac{TS}{2}} |\boldsymbol{\Sigma}_m|^{\frac{T}{2}}} \right\} \quad (4.4) \\ &= -\frac{TS}{2} \ln(2\pi) - \frac{T}{2} \ln |\boldsymbol{\Sigma}_m| - \frac{\sum_{t=1}^T}{2} [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m]^T [\mathbf{y}_t - \mathbf{F}_{m_t} \boldsymbol{\beta}_m] \\ &= -\frac{TS}{2} \ln(2\pi) - \frac{T}{2} \ln |\boldsymbol{\Sigma}_m| - \frac{\sum_{t=1}^T}{2} [\mathbf{y}_t - \mathbf{F}_{m_t} \hat{\boldsymbol{\beta}}_m + \mathbf{F}_{m_t} \hat{\boldsymbol{\beta}}_m - \mathbf{F}_{m_t} \boldsymbol{\beta}_m]^T \\ &\quad [\mathbf{y}_t - \mathbf{F}_{m_t} \hat{\boldsymbol{\beta}}_m + \mathbf{F}_{m_t} \hat{\boldsymbol{\beta}}_m - \mathbf{F}_{m_t} \boldsymbol{\beta}_m] \end{aligned}$$

Berdasarkan Rencher dan Schaalje (2008), estimator dari $\boldsymbol{\beta}_m$, yaitu $\hat{\boldsymbol{\beta}}_m$, memaksimumkan fungsi likelihood atau fungsi ln-likelihood seperti persamaan (4.3).

$$\max_{\beta_m} l(\beta_m, \Sigma_m)$$

Untuk mendapatkan $\hat{\beta}_m$ diterapkan diferensial orde pertama dari persamaan (4.4) terhadap elemen β_m , kemudian disamadengankan nol. Jika $Q = l(\beta_m, \Sigma_m; y_t)$, maka $\hat{\beta}_m$ diperoleh melalui proses berikut.

$$\begin{aligned} \frac{\partial Q}{\partial \beta_m} &= \frac{\partial}{\partial \beta_m} \left\{ -\frac{TS}{2} \ln(2\pi) - \frac{T}{2} \ln|\Sigma_m| \right. \\ &\quad \left. - \frac{\Sigma_m^{-1}}{2} \sum_{t=1}^T [y_t - F_{mt} \hat{\beta}_m + F_{mt} \hat{\beta}_m - F_{mt} \beta_m]^T [y_t - F_{mt} \hat{\beta}_m + F_{mt} \hat{\beta}_m - F_{mt} \beta_m] \right\} \\ &= \frac{\partial}{\partial \beta_m} \left\{ -\frac{TS}{2} \ln(2\pi) - \frac{T}{2} \ln|\Sigma_m| \right. \\ &\quad \left. - \frac{\Sigma_m^{-1}}{2} \sum_{t=1}^T \left[(y_t - F_{mt} \hat{\beta}_m)^T (y_t - F_{mt} \hat{\beta}_m) + (\hat{\beta}_m - \beta_m)^T F_{mt}^T F_{mt} (\hat{\beta}_m - \beta_m) \right] \right. \\ &\quad \left. + 2(\hat{\beta}_m - \beta_m)^T (F_{mt}^T y_t - F_{mt}^T F_{mt} \hat{\beta}_m) \right\} \\ &= 0 - 0 + \frac{\Sigma_m^{-1}}{2} \sum_{t=1}^T [0 + 0 - 2(F_{mt}^T y_t - F_{mt}^T F_{mt} \hat{\beta}_m)] \\ &= \frac{\Sigma_m^{-1}}{2} \sum_{t=1}^T [0 + 0 - 2(F_{mt}^T y_t - F_{mt}^T F_{mt} \hat{\beta}_m)] = 0 \\ &= \sum_{t=1}^T [-2(F_{mt}^T y_t - F_{mt}^T F_{mt} \hat{\beta}_m)] = 0 \\ &= -2 \sum_{t=1}^T (F_{mt}^T y_t - F_{mt}^T F_{mt} \hat{\beta}_m) = 0 \\ &= -2 \left(F_{mt}^T \sum_{t=1}^T y_t - T F_{mt}^T F_{mt} \hat{\beta}_m \right) = 0 \end{aligned}$$

kedua ruas dibagi dengan $-2T$

$$F_{mt}^T y - F_{mt}^T F_{mt} \hat{\beta}_m = 0$$

$$F_{mt}^T F_{mt} \hat{\beta}_m = F_{mt}^T y$$

$$\hat{\beta}_m = (F_{mt}^T F_{mt})^{-1} F_{mt}^T y \quad (4.5)$$

Persamaan (4.5) adalah sama dengan hasil estimasi koefisien regresi multivariat yang terdapat dalam Rencher dan Schaalje (2008).

4.1.2 Estimasi Parameter Bobot dan Varians

Algoritma *Expectation Maximum* (EM) digunakan untuk mengestimasi parameter yang tidak dapat diestimasi menggunakan MLE dengan menerapkan konsep likelihood *complete-data* $L(\boldsymbol{\theta}; \mathbf{y}, \mathbf{z})$ berdasarkan likelihood *incomplete-data* $L(\boldsymbol{\theta}; \mathbf{y})$ yang tidak memiliki penyelesaian secara analitik (McLachlan dan Khrisnan, 2008). Berikut ini merupakan tahapan algoritma EM dalam mengestimasi bobot w_m untuk anggota *ensemble* m dan varians σ_m^2 pada kasus respon atau eror berdistribusi normal.

1. Mendapatkan likelihood *incomplete-data*

Berdasarkan McLachan dan Krishnan (2008), likelihood *incomplete-data* dibentuk untuk memastikan bahwa parameter yang ingin diestimasi tidak memiliki penyelesaian secara analitik. Untuk kasus ini, berlaku bahwa

$$\boldsymbol{\theta} = (w_1 \ w_2 \ \dots \ w_m \ \sigma_m^2)' = (\mathbf{w}' \ \sigma_m^2)'; \mathbf{y} = (y_{s1} \ y_{s2} \ \dots \ y_{sT})'; \mathbf{F}_{mt} = (f_{ms1} \ f_{ms2} \ \dots \ f_{msT})'$$

dimana $w_m \geq 0$ dan $\sum_{m=1}^M w_m = 1$ dengan \mathbf{y} dan \mathbf{F}_m adalah vektor yang berukuran

$T \times 1$, T merupakan panjang periode *training*, seluruhnya pada lokasi $s \in S$. Dengan fungsi padat probabilitas (PDF) prediktif BMA mengacu pada persamaan (2.38), dapat dibentuk likelihood

$$L(\boldsymbol{\theta}; \mathbf{y}) = \prod_{t=1}^T g(\mathbf{y}_t; \boldsymbol{\theta}) = \prod_{t=1}^T \left(\sum_{m=1}^M w_m g_m(\mathbf{y}_t | \mathbf{F}_{mt}) \right) \quad (4.6)$$

dimana \mathbf{F}_{mt} merupakan prakiraan tiap anggota *ensemble* m pada hari ke- t . Fungsi ln-likelihood dari persamaan (4.6) ditunjukkan pada persamaan (4.7)

$$l(\boldsymbol{\theta}; \mathbf{y}) = \ln L(\boldsymbol{\theta}; \mathbf{y}) = \sum_{t=1}^T \ln \left(\sum_{m=1}^M w_m g_m(\mathbf{y}_t | \mathbf{F}_{mt}) \right) \quad (4.7)$$

Estimasi untuk bobot w_m sebagai bagian dari parameter $\boldsymbol{\theta}$ diperoleh dari turunan

$$\text{pertama dengan cara } \frac{\partial l(\boldsymbol{\theta}; \mathbf{y})}{\partial \boldsymbol{\theta}} \sim \frac{\partial l(\boldsymbol{\theta}; \mathbf{y})}{\partial w_m} = 0.$$

Oleh karena $w_M = 1 - w_1 - w_2 - \dots - w_{M-1} = 1 - \sum_{m=1}^{M-1} w_m$, maka w_M dapat dicari setelah estimasi untuk w_1, w_2, \dots, w_{M-1} diperoleh dengan syarat $\sum_{m=1}^{M-1} w_m = 1$ (McLachlan dan Krishnan, 2008). Dengan demikian, untuk w_j dimana $j = 1, 2, \dots, M - 1$, maka turunan pertama untuk mendapatkan \hat{w}_j adalah sebagai berikut.

$$\begin{aligned} \frac{\partial l(\boldsymbol{\theta}; \mathbf{y})}{\partial w_j} &= \frac{\partial}{\partial w_j} \left[\sum_{t=1}^T \ln \left(\sum_{j=1}^{M-1} w_j g_j(\mathbf{y}_t | \mathbf{F}_{mt}) + w_M g_M(\mathbf{y}_t | \mathbf{F}_{mt}) \right) \right] \\ &= \sum_{t=1}^T \left[\frac{g_j(\mathbf{y}_t | \mathbf{F}_{mt})}{\sum_{j=1}^{M-1} w_j g_j(\mathbf{y}_t | \mathbf{F}_{mt}) + w_M g_M(\mathbf{y}_t | \mathbf{F}_{mt})} \right] \\ &= \sum_{t=1}^T \left[\frac{g_j(\mathbf{y}_t | \mathbf{F}_{mt})}{\sum_{m=1}^M w_m g_m(\mathbf{y}_t | \mathbf{F}_{mt})} \right] \end{aligned}$$

Untuk mendapatkan \hat{w}_j , maka $\frac{\partial l(\boldsymbol{\theta}; \mathbf{y})}{\partial w_j}$ disamadengankan nol, sehingga

$$\sum_{t=1}^T \left[\frac{g_j(\mathbf{y}_t | \mathbf{F}_{mt})}{\sum_{m=1}^M w_m g_m(\mathbf{y}_t | \mathbf{F}_{mt})} \right] = 0. \quad (4.8)$$

Pada persamaan (4.8) w_j terletak pada posisi penyebut, sehingga solusi untuk \hat{w}_j tidak dapat diselesaikan secara analitik. Untuk mendapatkan solusi dari \hat{w}_j , diperlukan pendekatan iteratif, dalam penelitian ini menggunakan algoritma EM yang dikembangkan oleh Dempster, Laird, dan Rubin (1977).

McLachlan dan Krishnan (2008) menyarankan untuk menggunakan variabel laten yang tidak diamati (*unobservable*) pada algoritma EM yang dinotasikan dengan $\mathbf{Z} = (\mathbf{Z}'_1 \mathbf{Z}'_2 \dots \mathbf{Z}'_T)$, dengan T menyatakan periode training ke- T . Pada $t = 1, 2, \dots, T$ berlaku pula $\mathbf{Z}_t = (z_{1t} z_{2t} \dots z_{Mt})$. Hanya ada salah satu elemen \mathbf{Z}_t

yang bernilai 1 dalam suatu hari ke- t , sedangkan yang lain bernilai 0. Kondisi tersebut dapat dinotasikan dalam fungsi indikator dimana $m = 1, 2, \dots, M$ sebagai berikut,

$$\mathbf{I}_{(z_{mt}=k)} = \begin{cases} 1, & z_{mt} = k \\ 0, & z_{mt} \neq k \end{cases}$$

dengan k menunjukkan prakiraan anggota *ensemble* ‘terbaik’.

Oleh karena tiap anggota *ensemble* m berdistribusi normal dengan $\mu_{mt} = \mathbf{F}_m \boldsymbol{\beta}_m$, sehingga dapat dikatakan bahwa

$$\begin{aligned} y_t | (z_{mt} = 1) &\sim N(\mathbf{F}_1 \boldsymbol{\beta}_1, \boldsymbol{\Sigma}_1) \\ y_t | (z_{mt} = 2) &\sim N(\mathbf{F}_2 \boldsymbol{\beta}_2, \boldsymbol{\Sigma}_2) \\ &\vdots \\ y_t | (z_{mt} = M) &\sim N(\mathbf{F}_M \boldsymbol{\beta}_M, \boldsymbol{\Sigma}_M) \end{aligned}$$

Dengan bobot w menunjukkan probabilitas variabel laten \mathbf{Z} pada waktu ke- t , yaitu

$$p(z_{mt} = 1) = w_1, p(z_{mt} = 2) = w_2, \dots, p(z_{mt} = M) = w_M = 1 - \sum_{m=1}^{M-1} w_m.$$

Bobot untuk tiap anggota *ensemble* dan tiap waktu akan selalu berubah, namun sama untuk tiap lokasi. Hal ini menunjukkan bahwa kebaikan ramalan tiap anggota ensembel selama proses *windowing* berkontribusi besar terhadap penentuan bobot.

2. Mendapatkan likelihood *complete-data*

Tahap ini dilakukan untuk mendapatkan likelihood *complete-data* yang digunakan pada tahap EM berikutnya. Likelihood ini didapatkan dengan memasukkan peran dari variabel laten \mathbf{Z} (McLachlan dan Krishnan, 2008). Dengan menggunakan indikator $\mathbf{I}_{(z_{mt}=k)}$ untuk merepresentasikan variabel laten, maka fungsi likelihood *complete-data* adalah

$$L(\boldsymbol{\theta}; \mathbf{y}, \mathbf{z}) = \prod_{t=1}^T \sum_{m=1}^M \mathbf{I}_{(z_{mt}=k)} w_m g_m(\mathbf{y}_t | \mathbf{F}_{mt})$$

atau

$$\begin{aligned}
L(\boldsymbol{\theta}; \mathbf{y}, \mathbf{z}) &= \exp\left(\sum_{t=1}^T \sum_{m=1}^M \mathbf{I}_{(z_{mt}=k)} \left[\ln w_m + \ln g_m(y_t | \mathbf{F}_{mt}) \right]\right) \\
&= \exp\left(\sum_{t=1}^T \sum_{m=1}^M \mathbf{I}_{(z_{mt}=k)} \left[\ln w_m - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi - \frac{1}{2\sigma^2} (y_t - \mathbf{F}_{mt})^2 \right]\right).
\end{aligned}$$

3. Melakukan tahap *Expectation*

Berdasarkan Lange (2010), tahap *Expectation* atau sering dituliskan E dilakukan secara iteratif untuk mendapatkan ekspektasi dari fungsi likelihood atau ln-likelihood *complete-data*. Iterasi ini diawali dengan *initial value* yang merupakan dugaan awal parameter varians sama dengan 1 dan bobot yang sama untuk masing-masing ensemble yaitu sebesar 1/M (Slougher *et al.*, 2007). Jika i menyatakan iterasi ke- i , maka

$$\begin{aligned}
Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)}) &= E_{\boldsymbol{\theta}^{(i)}} \left[\ln L(\boldsymbol{\theta}; \mathbf{y}, \mathbf{z}) \right] \\
&= E_{\boldsymbol{\theta}^{(i)}} \left[\sum_{t=1}^T \sum_{m=1}^M \mathbf{I}_{(z_{mt}=k)} \left(\ln w_m - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi - \frac{1}{2\sigma^2} (y_t - \mathbf{F}_{mt})^2 \right) \right].
\end{aligned}$$

Dengan demikian,

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)}) = \left[\sum_{t=1}^T \sum_{m=1}^M E_{\boldsymbol{\theta}^{(i)}} \mathbf{I}_{(z_{mt}=k)} \left(\ln w_m - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi - \frac{1}{2\sigma^2} (y_t - \mathbf{F}_{mt})^2 \right) \right]. \quad (4.9)$$

Dari persamaan (4.9), dapat dikatakan bahwa untuk suatu hari ke- t , hanya ada satu elemen \mathbf{Z}_t yang bernilai 1 atau salah satu anggota *ensemble* m dimana $z_{mt} = 1$. Jika dijabarkan, maka diperoleh

$$E_{\boldsymbol{\theta}^{(i)}} \mathbf{I}_{(z_{mt}=k)} = E_{\boldsymbol{\theta}^{(i)}} (z_{mt} | \mathbf{y}) = \Pr_{\boldsymbol{\theta}^{(i)}} (z_{mt} = 1 | \mathbf{y})$$

dengan

$$z_{mt}^{(i)} = \frac{w_m g_m(y_t | \mathbf{F}_{mt}, \sigma^{(i-1)})}{\sum_{l=1}^M w_l g_l(y_t | \mathbf{F}_{lt}, \sigma^{(i-1)})} \quad (4.10)$$

Persamaan (4.10) adalah probabilitas posterior dari observasi pada waktu ke- t untuk anggota *ensemble* m , sehingga,

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)}) = \left[\sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} \left(\ln w_m - \frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi - \frac{1}{2\sigma^2} (y_t - \mathbf{F}_{mt})^2 \right) \right].$$

4. Melakukan tahap *Maximization*

Tahap terakhir dalam algoritma EM adalah memilih $\boldsymbol{\theta}^{(i+1)}$ yang memaksimumkan $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})$ dimana

$$Q(\boldsymbol{\theta}^{(i+1)} | \boldsymbol{\theta}^{(i)}) \geq Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})$$

dengan kendala $w_1 + w_2 + \dots + w_m = 1$ dan $\boldsymbol{\theta} = (\mathbf{w}', \sigma^2)'$. Sama seperti tahap *Expectation*, tahap *Maximization* atau M juga dilakukan secara iteratif. Estimasi untuk $\boldsymbol{\theta}^{(i+1)}$ yang memaksimumkan $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})$ dapat diperoleh dari $\frac{\partial Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})}{\partial \boldsymbol{\theta}} = \mathbf{0}$.

a. Estimasi w_m pada iterasi ke- i w_m^i

Berdasarkan karakteristik observasi laten \mathbf{Z}_t dimana $z_{1t} + z_{2t} + \dots + z_{Mt} = 1$ untuk suatu hari t , maka hal ini mengakibatkan

$$\sum_{t=1}^T z_{1t} + \sum_{t=1}^T z_{2t} + \dots + \sum_{t=1}^T z_{Mt} = \sum_{t=1}^T (z_{1t} + z_{2t} + \dots + z_{Mt}) = T.$$

Estimasi untuk \mathbf{W} dapat dicari dengan $\frac{\partial Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})}{\partial \mathbf{w}} = \mathbf{0}$, dimana

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)}) = \left[\sum_{t=1}^T z_{1t}^{(i)} \right] \ln w_1 + \left[\sum_{t=1}^T z_{2t}^{(i)} \right] \ln w_2 + \dots + \left[\sum_{t=1}^T z_{Mt}^{(i)} \right] \ln w_M + C$$

$$\text{dengan } C = \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} \left(-\frac{1}{2} \ln \sigma^2 - \frac{1}{2} \ln 2\pi - \frac{1}{2\sigma^2} (y_t - \mathbf{F}_{mt})^2 \right).$$

Apabila dimisalkan $M = 2$, maka $w_1 + w_2 = 1$ dengan $w_2 = 1 - w_1$, sehingga diperoleh persamaan (4.11).

$$\begin{aligned} \frac{\partial Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})}{\partial w_1} &= \frac{\partial}{\partial w_1} \left(\left[\sum_{t=1}^T z_{1t}^{(i)} \right] \ln w_1 + \left[\sum_{t=1}^T z_{2t}^{(i)} \right] \ln w_2 + C \right) \\ &= \frac{\partial}{\partial w_1} \left(\left[\sum_{t=1}^T z_{1t}^{(i)} \right] \ln w_1 + \left[\sum_{t=1}^T z_{2t}^{(i)} \right] \ln (1 - w_1) + C \right) \\ &= \frac{\sum_{t=1}^T z_{1t}^{(i)}}{w_1} - \frac{\sum_{t=1}^T z_{2t}^{(i)}}{1 - w_1} \end{aligned} \quad (4.11)$$

Selanjutnya persamaan (4.11) disamadengankan 0, yaitu $\frac{\partial Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})}{\partial w_1} = 0$,

sehingga diperoleh

$$w_1^{(i)} \left(1 + \frac{\sum_{t=1}^T z_{1t}^{(i)}}{\sum_{t=1}^T z_{2t}^{(i)}} \right) = \frac{\sum_{t=1}^T z_{1t}^{(i)}}{\sum_{t=1}^T z_{2t}^{(i)}} \rightarrow w_1^{(i)} = \frac{\sum_{t=1}^T z_{1t}^{(i)}}{\sum_{t=1}^T z_{1t}^{(i)} + \sum_{t=1}^T z_{2t}^{(i)}} = \frac{1}{T} \sum_{t=1}^T z_{1t}^{(i)} .$$

$$\text{Jadi } w_2^{(i)} = 1 - w_1^{(i)} = 1 - \frac{1}{T} \sum_{t=1}^T z_{1t}^{(i)} = \frac{\sum_{t=1}^T z_{1t}^{(i)} + \sum_{t=1}^T z_{2t}^{(i)} - \sum_{t=1}^T z_{1t}^{(i)}}{T} = \frac{1}{T} \sum_{t=1}^T z_{2t}^{(i)} .$$

Sehingga secara umum, untuk *ensemble* berukuran M , berlaku

$$w_m^{(i)} = \frac{1}{T} \sum_{t=1}^T z_{mt}^{(i)}, m = 1, 2, \dots, M \quad (4.12)$$

b. Estimasi varians pada iterasi ke- i $\sigma^{2(i)}$

Dengan langkah yang sama seperti sebelumnya, estimasi $\sigma^{2(i)}$ didapatkan

dengan memaksimumkan $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})$ berdasarkan $\frac{\partial Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})}{\partial \sigma^2} = 0$. Berlaku bahwa

$$Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)}) = -\frac{1}{2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2 + D$$

$$\text{dengan } D = \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} \left(\ln w_m - \frac{1}{2} \ln 2\pi \right).$$

Turunan $Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})$ terhadap σ^2 adalah sebagai berikut.

$$\begin{aligned} \frac{\partial Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})}{\partial \sigma^2} &= \frac{\partial}{\partial \sigma^2} \left(-\frac{1}{2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2 + D \right) \\ &= -\frac{1}{2\sigma^2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} + \frac{1}{2(\sigma^2)^2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2 + 0 \end{aligned} \quad (4.13)$$

Selanjutnya persamaan (4.13) disamadengankan nol menjadi

$$\begin{aligned} \frac{\partial Q(\boldsymbol{\theta} | \boldsymbol{\theta}^{(i)})}{\partial \sigma^2} &= 0 \\ \frac{1}{2\sigma^{2(i)}} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} &= \frac{1}{2(\sigma^{2(i)})^2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2 \\ \frac{1}{2\sigma^{2(i)}} \left(\sum_{t=1}^T z_{1t}^{(i)} + \dots + \sum_{t=1}^T z_{Mt}^{(i)} \right) &= \frac{1}{2(\sigma^{2(i)})^2} \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2 \\ \sigma^{2(i)} T &= \sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2 \end{aligned}$$

sehingga

$$\sigma^{2(i)} = \frac{\sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2}{T}. \quad (4.14)$$

Persamaan (4.12) dan (4.14) merupakan estimasi untuk bobot w_m dan varians σ^2 yang sama dengan hasil estimasi yang terdapat dalam (Luthfi, 2017).

Konvergensi Iterasi :

Tahap E dan M terus dilakukan secara iteratif hingga konvergen. Untuk $i = 0, 1, 2, \dots$, iterasi dapat dihentikan (konvergen) ketika $\boldsymbol{\theta}^{(i+1)} - \boldsymbol{\theta}^{(i)}$ sebesar batas toleransi yang sangat rendah, misal 10^{-4} atau 10^{-6} . Sama dengan syarat konvergensi pada McLachlan dan Krishnan (2008) menyatakan bahwa konvergensi tercapai jika likelihood *incomplete-data* tidak menurun setelah iterasi ke-($i+1$), yaitu $L(\boldsymbol{\theta}^{(i+1)}; \mathbf{y}) \geq L(\boldsymbol{\theta}^{(i)}; \mathbf{y})$, setelah serangkaian proses iterasi.

4.1.3 Estimasi Parameter Kovarians Spasial

Selain parameter regresi serta parameter bobot dan varians, Spatial BMA juga memiliki parameter kovarians spasial (ρ_m^2 , τ_m^2 , dan r_m^2) berdasarkan model GOP. Berdasarkan persamaan (2.40) struktur matriks kovarians spasial pada model Spatial BMA adalah

$$\Sigma_m = \frac{\sigma^2}{\rho_m^2 + \tau_m^2} \Sigma$$

dimana σ^2 adalah varians model BMA yang telah diperoleh estimasinya pada persamaan (4.14). Berdasarkan Berrocal *et al.* (2007) Σ adalah struktur matriks kovarians spasial GOP, dimana elemennya adalah

$$\rho_m^2 \delta_{ij} + \tau_m^2 \exp\left(-\frac{\|s_i - s_j\|}{r_m}\right).$$

Dimana $\|s_i - s_j\|$ menotasikan jarak Euclidian antara masing-masing lokasi s_i dan s_j . Sedangkan δ_{ij} bernilai sama dengan 1 jika $s_i = s_j$ dan lainnya bernilai 0. Dalam terminologi geostatistik, ρ_m^2 disebut efek nugget yang merepresentasikan varians dari kesalahan pengukuran, $\rho_m^2 + \tau_m^2$ disebut sill, dan r_m merupakan jarak yang mengindikasikan tingkat dimana masih terdapat korelasi spasial.

Sebelum mengestimasi ρ_m^2 , τ_m^2 , dan r_m^2 , langkah pertama yang dilakukan adalah menghitung variogram $\hat{\gamma}_k(d)$ dengan tahapan sebagai berikut (Berrocal *et al.*, 2007).

- a. Untuk tiap hari dalam periode training, menghitung error model yang diperoleh dengan mengurangi nilai observasi terhadap prakiraan, yaitu $\varepsilon_{s,t} = y_{s,t} - \hat{y}_{s,t}$.
- b. Untuk tiap hari dalam periode training dan untuk tiap pasangan lokasi pengamatan pada hari tersebut, menghitung jarak antar lokasi $d(i,j)$ dengan jarak Euclidean $\|s_i - s_j\| = \sqrt{(\mathbf{u}_i - \mathbf{u}_j)^2 + (\mathbf{v}_i - \mathbf{v}_j)^2}$ dimana u dan v merupakan koordinat lintang dan bujur stasiun meteorologi s dan menghitung setengah dari selisih kuadrat antar error $\frac{1}{2} [\varepsilon(\mathbf{s}_i) - \varepsilon(\mathbf{s}_j)]^2$.
- c. Mengelompokkan jarak-jarak yang ada ke dalam *bin* B_l dengan titik tengah (*midpoint*) d_l , dimana l adalah suatu jarak, hingga diperoleh k atau banyaknya pasangan jarak antar lokasi yang berada pada bin B_l .
- d. Menghitung nilai semivariogram empiris $\hat{\gamma}_k(d_l)$ pada jarak d_l dengan menghitung setengah dari ekspektasi kuadrat selisih antar error pada *bin* B_l ,

$$\hat{\gamma}_k(d_l) = \frac{1}{2k} \sum_{i=1}^k (\varepsilon(\mathbf{s}_{i+d_l}) - \varepsilon(\mathbf{s}_i))^2 \text{ dimana } d_l = d_l(1), d_l(2), \dots$$

Langkah selanjutnya adalah mengaplikasikan metode kuadrat terkecil terboboti untuk mengestimasi parameter ρ_m^2 , τ_m^2 , dan r_m^2 . Estimasi ketiga parameter tersebut dapat diperoleh dengan meminimumkan persamaan (4.15) sebagai berikut.

$$g(\rho_m^2, \tau_m^2, r_m) = \sum_{l=1}^k n_l \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right)^2, \rho_m^2, \tau_m^2, r_m \geq 0 \quad (4.15)$$

1. *Nugget* ρ_m^2

$$\begin{aligned} \frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial \rho_m^2} &= \frac{\partial}{\partial \rho_m^2} \left(\sum_{l=1}^k n_l \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right)^2 \right) \\ &= \sum_{l=1}^k n_l \frac{\partial}{\partial \rho_m^2} \left(\hat{\gamma}(d_l) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-1} - 1 \right)^2 \\ &= \sum_{l=1}^k n_l 2 \left(\hat{\gamma}(d_l) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-1} - 1 \right) \times \\ &\quad \left(-\hat{\gamma}(d_l) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-2} \right) \\ &= -2 \sum_{l=1}^k n_l \frac{\hat{\gamma}(d_l)}{[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})]^2} \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right) \end{aligned} \quad (4.16)$$

Persamaan (4.16) mengindikasikan bahwa $\frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial \rho_m^2} = 0$ akan memberikan hasil

Syang tidak *close-form* karena $\hat{\rho}_m^2$ masih mengandung parameter τ_m^2 dan r_m .

2. *Partial sill* τ_m^2

$$\begin{aligned} \frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial \tau_m^2} &= \frac{\partial}{\partial \tau_m^2} \left(\sum_{l=1}^k n_l \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right)^2 \right) \\ &= \sum_{l=1}^k n_l \frac{\partial}{\partial \tau_m^2} \left(\hat{\gamma}(d_l) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-1} - 1 \right)^2 \end{aligned}$$

$$\begin{aligned}
&= \sum_{l=1}^k n_l 2 \left(\hat{\gamma}(d_l) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-1} - 1 \right) \times \\
&\quad \left(-\hat{\gamma}(d_l) (1 - e^{-d_l/r_m}) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-2} \right) \\
&= -2 \sum_{l=1}^k n_l \frac{\hat{\gamma}(d_l) (1 - e^{-d_l/r_m})}{[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})]^2} \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right) \quad (4.17)
\end{aligned}$$

Sama halnya dengan persamaan (4.16), persamaan (4.17) mengindikasikan bahwa

$\frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial \tau_m^2}$ akan memberikan hasil yang tidak *close-form* karena $\hat{\tau}_m^2$ masih

mengandung $\hat{\rho}_m^2$ dan \hat{r}_m .

3. Range r_m

$$\begin{aligned}
\frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial r_m} &= \frac{\partial}{\partial r_m} \left(\sum_{l=1}^k n_l \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right)^2 \right) \\
&= \sum_{l=1}^k n_l \frac{\partial}{\partial r_m} \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right)^2 \\
&= \sum_{l=1}^k n_l \frac{\partial}{\partial r_m} \left(\hat{\gamma}(d_l) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-1} - 1 \right)^2 \\
&= \sum_{l=1}^k 2n_l \left(-\hat{\gamma}(d_l) \sigma^2 \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-2} \right) \times \\
&\quad \left(\hat{\gamma}(d_l) \left[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m}) \right]^{-1} - 1 \right) \left(-d_l r_m^{-2} e^{-d_l/r_m} \right) \\
&= 2 \frac{\tau_m^2}{r_m^2} \sum_{l=1}^k n_l d_l e^{-d_l/r_m} \frac{\hat{\gamma}(d_l)}{[\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})]^2} \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right) \quad (4.18)
\end{aligned}$$

Sama seperti persamaan (4.16) dan (4.17), persamaan (4.17) juga mengindikasikan

bahwa $\frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial r_m} = 0$ akan memberikan hasil yang tidak *close-form* karena \hat{r}_m

masih mengandung $\hat{\rho}_m^2$ dan $\hat{\tau}_m^2$.

Sehingga diperlukan pendekatan iteratif untuk mendapatkan estimasi ketiga parameter tersebut. Berdasarkan Berrocal *et al.* (2007), pendekatan iteratif yang dapat digunakan diantaranya adalah *Limited-Memory* BFGS (L-BFGS). Pendekatan ini digunakan untuk kasus fungsi objektif non-linier berkendala maupun tidak berkendala dengan parameter yang cukup banyak, seperti pada persamaan (4.15). Berdasarkan Liu dan Nocedal (1989), berlaku vektor gradient

jika $\mathbf{z} = [\rho_m^2 \ \tau_m^2 \ r_m]'$,

$$\nabla g(\mathbf{z}) = \left[\frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial \rho_m^2} \quad \frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial \tau_m^2} \quad \frac{\partial g(\rho_m^2, \tau_m^2, r_m)}{\partial r_m} \right]' \quad (4.19)$$

dengan tahapan estimasi sebagai berikut dimana $k = 0, 1, 2, \dots$

1. Menentukan nilai awal untuk \mathbf{z} dan \mathbf{H} , yaitu \mathbf{z}_0 dan \mathbf{H}_0 . Pada iterasi $k = 0$, \mathbf{z}_0 bernilai non-negatif sesuai dengan kendala yang wajib dipenuhi seperti pada Cressie (1993), yaitu $\rho_m^2, \tau_m^2, r_m \geq 0$, dan \mathbf{H}_0 adalah matrik simetris definit positif, misalnya matrik identitas atau \mathbf{I} . Selain itu, menentukan m atau suatu bilangan integer yang positif untuk menentukan seberapa panjang informasi \mathbf{H}_0 digunakan dalam memperbarui iterasi. Lalu, menetapkan nilai β dan γ dimana

$$0 < \gamma < 0.5$$

$$\gamma < \beta < 1$$

Umumnya, m tidak lebih daripada 10 agar iterasi berjalan cukup singkat. Karena L-BFGS didasari konsep gradien konjugat, maka γ disarankan sangat rendah, misal 10^{-4} , dan β jauh lebih besar, misal 0,1.

2. Menghitung $\Delta \mathbf{z}_k = -\mathbf{H}_k \nabla g(\mathbf{z}_k)$ dan $\mathbf{z}_{k+1} = \mathbf{z}_k + \alpha_k \Delta \mathbf{z}_k$ dimana $\nabla g(\mathbf{z}_k)$ adalah gradien fungsi pada titik \mathbf{z}_k dan konstanta α_k memenuhi syarat Wolfe:

$$a. \quad g(\mathbf{z}_k + \alpha_k \mathbf{z}_k) \leq g(\mathbf{z}_k) + \gamma \alpha_k [\nabla g(\mathbf{z}_k)]^T \Delta \mathbf{z}_k$$

$$b. \quad [\nabla g(\mathbf{z}_k + \alpha_k \mathbf{z}_k)]^T \Delta \mathbf{z}_k \geq \beta [\nabla g(\mathbf{z}_k)]^T \Delta \mathbf{z}_k$$

Untuk $k = 0$, umumnya α_0 bernilai 1.

3. Jika $\|\nabla g(\mathbf{z}_{k+1}) - \nabla g(\mathbf{z}_k)\| < \varepsilon$, dimana ε merupakan bilangan yang sangat kecil, maka iterasi dihentikan. Jika tidak, iterasi dilanjutkan ke langkah 4.
4. Memperbarui \mathbf{H}_k berdasarkan informasi \mathbf{H}_0 sebanyak \hat{m} kali sehingga diperoleh \mathbf{H}_{k+1} dengan $\hat{m} = \min(k+1, m)$. Makin tinggi konstanta m , maka \mathbf{H}_0 makin konsisten digunakan untuk memperbarui \mathbf{H}_{k+1} .

$$\mathbf{H}_{k+1} = \mathbf{V}_k^T \mathbf{H}_k \mathbf{V}_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T \quad (4.20)$$

dimana

$$\rho_k = \frac{1}{\mathbf{y}_k^T \mathbf{s}_k}; \mathbf{V}_k = \mathbf{I} - \rho_k \mathbf{y}_k \mathbf{s}_k^T; \{\mathbf{s}_i, \mathbf{y}_i\} = \{\mathbf{z}_{i+1} - \mathbf{z}_i, \nabla g(\mathbf{z}_{i+1}) - \nabla g(\mathbf{z}_i)\}, i = k - \hat{m} + 1, \dots, k$$

sehingga persamaan (4.20) termodifikasi menjadi

$$\begin{aligned} \mathbf{H}_{k+1} &= \mathbf{V}_k^T \mathbf{H}_k \mathbf{V}_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T = \mathbf{V}_k^T (\mathbf{V}_{k-1}^T \mathbf{H}_{k-1} \mathbf{V}_{k-1} + \rho_{k-1} \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T) \mathbf{V}_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T \\ &= \mathbf{V}_k^T \mathbf{V}_{k-1}^T \mathbf{H}_{k-1} \mathbf{V}_{k-1} \mathbf{V}_k + \rho_{k-1} \mathbf{V}_k^T \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T \mathbf{V}_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T. \end{aligned} \quad (4.21)$$

Karena $\mathbf{H}_{k-1} = \mathbf{V}_{k-2}^T \mathbf{H}_{k-2} \mathbf{V}_{k-2} + \rho_{k-2} \mathbf{s}_{k-2} \mathbf{s}_{k-2}^T$, maka persamaan (4.21) menjadi

$$\begin{aligned} \mathbf{H}_{k+1} &= \mathbf{V}_k^T \mathbf{V}_{k-1}^T (\mathbf{V}_{k-2}^T \mathbf{H}_{k-2} \mathbf{V}_{k-2} + \rho_{k-2} \mathbf{s}_{k-2} \mathbf{s}_{k-2}^T) \mathbf{V}_{k-1} \mathbf{V}_k + \rho_{k-1} \mathbf{V}_k^T \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T \mathbf{V}_k + \\ &\quad \rho_k \mathbf{s}_k \mathbf{s}_k^T \\ &= \mathbf{V}_k^T \mathbf{V}_{k-1}^T \mathbf{V}_{k-2}^T \mathbf{H}_{k-2} \mathbf{V}_{k-2} \mathbf{V}_{k-1} \mathbf{V}_k + \rho_{k-1} \mathbf{V}_k^T \mathbf{V}_{k-1}^T \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T \mathbf{V}_{k-1} \mathbf{V}_k + \\ &\quad \rho_{k-1} \mathbf{V}_k^T \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T \mathbf{V}_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T. \end{aligned} \quad (4.22)$$

Berdasarkan persamaan (4.20), (4.21), dan (4.22), maka berlaku persamaan umum (4.23) untuk mencari \mathbf{H}_{k+1} .

$$\begin{aligned} \mathbf{H}_{k+1} &= (\mathbf{V}_k^T \mathbf{V}_{k-1}^T \dots \mathbf{V}_{k-\hat{m}+1}^T) \mathbf{H}_{k-\hat{m}+1} (\mathbf{V}_{k-\hat{m}+1} \dots \mathbf{V}_{k-1} \mathbf{V}_k) + \\ &\quad \rho_{k-\hat{m}+1} (\mathbf{V}_{k-1}^T \mathbf{V}_{k-2}^T \dots \mathbf{V}_{k-\hat{m}+2}^T) \mathbf{s}_{k-\hat{m}+1} \mathbf{s}_{k-\hat{m}+1}^T (\mathbf{V}_{k-\hat{m}+2} \dots \mathbf{V}_{k-2} \mathbf{V}_{k-1}) + \dots + \\ &\quad \rho_{k-1} \mathbf{V}_k^T \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T \mathbf{V}_k + \rho_k \mathbf{s}_k \mathbf{s}_k^T \end{aligned} \quad (4.23)$$

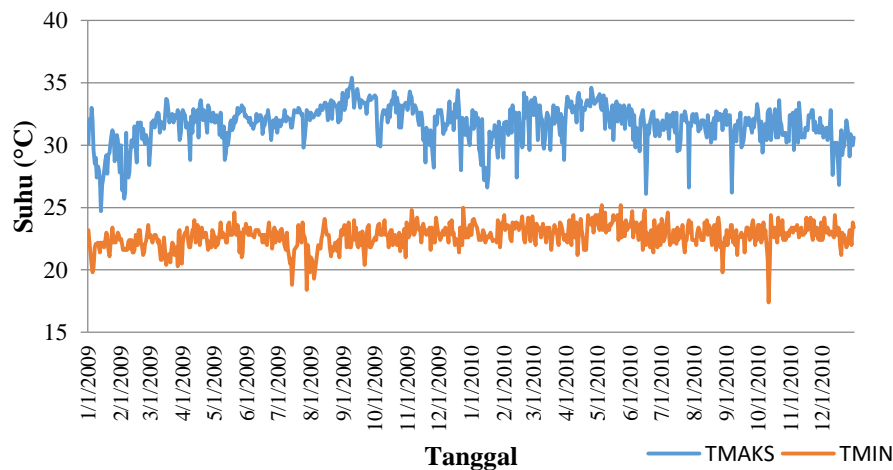
5. Kembali ke langkah 2 agar mendapatkan $\Delta \mathbf{z}_k$ dan memperbarui \mathbf{z}_{k+1} untuk memeriksa konvergensi iterasi dengan $k = k + 1$.

4.2 Aplikasi Spatial BMA untuk Prakiraan Cuaca Jangka Pendek Terkalibrasi di 8 Stasiun Meteorologi Jabodetabek

Setelah mengkaji proses untuk mendapatkan estimator parameter Spatial BMA, maka langkah selanjutnya adalah mengaplikasikan model tersebut dalam kasus prakiraan cuaca di 8 stasiun meteorologi (stamet) Jabodetabek. Pada sub-bab ini dijelaskan mengenai tahapan pra-pemrosesan, yaitu standarisasi dan reduksi dimensi dengan PCA, pembentukan anggota *ensemble* dengan regresi PLS, PCR, dan Ridge sebelum suhu udara dimodelkan dengan Spatial BMA.

4.2.1 Analisis Deskripsi Suhu Udara dan Standarisasi

Sebelum memulai pra-pemrosesan data, dilakukan analisis deskripsi secara visual untuk mendapatkan gambaran umum dari pola cuaca, yaitu suhu udara. Gambar 4.1 menunjukkan tren pergerakan suhu maksimum dan suhu minimum di stamet Dermaga selama periode 708 hari pengamatan, yaitu 1 Januari 2009 sampai 31 Desember 2010.



Gambar 4.1 Observasi Suhu Maksimum dan Suhu Minimum di Stamet Dermaga

Pada umumnya, suhu maksimum dan minimum di stamet Dermaga tidak menunjukkan adanya gejala yang menyimpang cukup jauh, meskipun pada hari-hari tertentu nilai keduanya berada pada angka jauh lebih rendah dibanding dengan hari sebelumnya. Pada Gambar 4.1, tren suhu maksimum (berwarna biru) dan tren

suhu minimum (berwarna merah) menunjukkan bahwa pergerakan suhu maksimum dan minimum harian berfluktuatif selama 708 hari pengamatan. Jika suhu minimum mengalami perubahan (kenaikan/penurunan) yang signifikan, maka terjadi pula hal yang sama pada suhu maksimum.

Setelah menganalisis secara deskriptif terhadap suhu maksimum dan suhu minimum harian, selanjutnya dilakukan pra-pemrosesan parameter NWP dengan PCA. Namun, sebelumnya dilakukan standarisasi terhadap respon (observasi suhu udara) dan parameter NWP berdasarkan metode *scaled and centered*, yaitu nilai pengamatan dikurangi mean lalu dibagi standar deviasi. Jadi banyak komponen utama yang digunakan dalam pemodelan kedepannya diperoleh dari data yang distandarkan (bukan data asli). Hal ini dilakukan untuk meminimalkan perbedaan skala pengukuran antar parameter NWP.

4.2.2 Pra-pemrosesan Parameter NWP dengan PCA

Tahap ini dilakukan untuk mengetahui keterkaitan grid sebab terdapat indikasi bahwa kondisi atmosfer dipengaruhi oleh hubungan spasial antar elemen cuaca pembentuk atmosfer. Misal kejadian cuaca tertentu di daerah Dermaga memiliki kemungkinan dipengaruhi oleh kejadian cuaca yang sedang atau baru saja berakhir di lokasi terdekat dengan Dermaga. Oleh karena itu, pra-pemrosesan parameter NWP dengan PCA bertujuan untuk mereduksi hubungan spasial antar grid ke dalam beberapa komponen utama. Selain itu, dengan mengetahui hubungan yang terbentuk antar grid, maka parameter yang ada dapat direduksi (reduksi dimensi), sehingga menyederhanakan pemodelan dan diharapkan mampu mempersingkat proses komputasi tanpa mengorbankan akurasi dan presisi ramalan. Tabel 4.1 menyajikan nilai eigen dan kumulatif varians (dalam persen) salah satu variable NWP yaitu presipitasi (rnd).

Tabel 4.1 Nilai Eigen dan Kumulatif Varians (%) PC Variabel Presipitasi

PC ke-	Cengkareng		Citeko		Curug		Dermaga	
	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians
1	7.350*	0.8166	7.920*	0.8800	8.003*	0.8892	7.359*	0.8176
2	0.995	0.9272	0.442	0.9291	0.548	0.9500	0.788	0.9052
3	0.348	0.9659	0.386	0.9719	0.299	0.9833	0.580	0.9696
4	0.170	0.9848	0.113	0.9845	0.075	0.9917	0.160	0.9874
5	0.077	0.9934	0.061	0.9913	0.034	0.9955	0.051	0.9930

Tabel 4.1 Nilai Eigen dan Kumulatif Varians (%) PC Variabel Presipitasi (Lanjutan)

PC ke-	Cengkareng		Citeko		Curug		Dermaga	
	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians
7	0.021	0.9991	0.020	0.9979	0.011	0.9988	0.023	0.9991
8	0.006	0.9998	0.016	0.9997	0.009	0.9999	0.007	0.9998
9	0.002	1.0000	0.003	1.0000	0.001	1.0000	0.001	1.0000
PC ke-	Kemayoran		Pondok Betung		Priok		Tangerang	
	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians	Nilai Eigen	Kumulatif Varians
1	7.682*	0.8536	7.943*	0.8826	7.085*	0.7872	7.856*	0.8729
2	0.655	0.9264	0.478	0.9357	1.213*	0.9220	0.645	0.9445
3	0.317	0.9617	0.262	0.9648	0.378	0.9639	0.323	0.9805
4	0.153	0.9787	0.148	0.9813	0.137	0.9792	0.091	0.9906
5	0.080	0.9876	0.103	0.9928	0.076	0.9876	0.046	0.9957
6	0.057	0.9939	0.030	0.9961	0.059	0.9941	0.017	0.9977
7	0.032	0.9975	0.023	0.9987	0.031	0.9976	0.012	0.9990
8	0.020	0.9997	0.010	0.9998	0.019	0.9997	0.007	0.9998
9	0.003	1.0000	0.002	1.0000	0.003	1.0000	0.002	1.0000

Keterangan : *) komponen dengan nilai eigen > 1 dan kumulatif varians minimal 80%

Berdasarkan Tabel 4.1 dapat diketahui bahwa variabilitas data presipitasi di Dermaga yang dapat dijelaskan oleh komponen pertama sebesar 81,76%. Tingginya kumulatif varians yang dicapai oleh hanya komponen pertama saja mengindikasikan bahwa presipitasi antar grid (lokasi) memiliki keterkaitan yang tinggi. Artinya, hujan yang terjadi di daerah Dermaga berpengaruh pada kejadian hujan pada lokasi-lokasi terdekat. Sebaliknya, hujan di daerah Dermaga dipengaruhi oleh kejadian hujan di daerah sekitar. Selain itu, variabilitas data presipitasi pada delapan stasiun meteorologi yang dapat dijelaskan oleh PCA berkisar antara 81,66% (Cengkareng) hingga 92,20% (Priok) dimana tingginya nilai variabilitas data telah dapat dijelaskan oleh satu atau dua komponen saja.

Kriteria untuk menentukan banyaknya komponen utama adalah berdasarkan nilai eigen yang lebih besar atau sama dengan 1 dan mencapai kumulatif varians minimal sebesar 80%. Selanjutnya, dengan menerapkan langkah yang sama pada Lampiran 36-a, diperoleh komponen-komponen yang merepresentasikan parameter NWP di stamet Dermaga dan disajikan pada Tabel 4.2. Sementara banyaknya PC parameter NWP yang terpilih di tujuh stamet lainnya disajikan pada Lampiran 1-7.

Tabel 4.2 Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Dermaga

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,998	99,98%	temp2	1	8,381	93,13%
mixr1	1	7,632	84,80%	temp4	1	8,753	97,26%
mixr2	1	8,367	92,96%	tmaxscr	1	8,640	96,00%
mixr4	1	8,618	95,76%	tminscr	1	7,542	83,80%
omega1	2	5,832; 2,139	88,56%	tpan	1	7,975	88,61%
omega2	2	5,597; 2,266	87,37%	tscrn	1	7,984	88,72%
omega4	2	6,299; 1,273	84,13%	u1	1	7,507	83,41%
pblh	1	7,483	83,15%	u2	1	8,012	89,02%
ps	1	8,529	94,77%	u4	1	8,877	98,63%
psl	1	8,995	99,95%	ustar	2	6,014; 1,393	82,29%
qgscrn	2	6,651; 1,104	86,17%	v1	2	6,371; 1,482	87,25%
rh1	2	7,186; 1,182	92,98%	v2	2	6,850; 1,136	88,74%
rh2	1	8,253	91,70%	v4	1	8,607	95,64%
rh4	1	8,591	95,45%	zg1	1	8,780	97,55%
rnd	1	7,359	81,76%	zg2	1	7,910	87,89%
temp1	1	7,943	88,25%	zg4	2	5,057; 3,814	98,57%

Tabel 4.2 memberikan informasi bahwa 1 komponen sudah sangat baik dalam merepresentasikan sebagian parameter NWP di stamet Dermaga. Namun juga terdapat cukup banyak parameter NWP yang direpresentasikan oleh 2 komponen, diantaranya seluruh parameter omega dan ustar. Dari hasil reduksi dimensi PCA diperoleh sebanyak 41 komponen dari 32 parameter dimana tiap satu parameter terdiri dari 10 grid.

Variabilitas parameter NWP yang dapat dijelaskan oleh komponen utama paling kecil adalah 81,76% hingga yang terbesar mendekati 100%. Hal ini mengindikasikan bahwa keterkaitan grid dalam satu parameter NWP relatif tinggi, sehingga reduksi dimensi merupakan langkah yang cukup tepat untuk mengoreksi bias luaran NWP. Komponen-komponen utama yang merepresentasikan parameter NWP pada 7 stamet lainnya disajikan dalam Lampiran 1-7. Selanjutnya, komponen

utama yang terbentuk dari seluruh parameter NWP digunakan sebagai variabel prediktor dalam pemodelan menggunakan regresi PLS, PCR, dan Ridge.

4.2.3 Pembentukan Model Prakiraan Anggota *Ensemble* untuk Suhu Maksimum dan Suhu Minimum

Setelah mengetahui banyaknya komponen utama yang digunakan sebagai prediktor, langkah selanjutnya adalah mendapatkan prakiraan anggota *ensemble* berdasarkan regresi PLS, PCR, dan Ridge. Sebelum memodelkan regresi PLS dan PCR, dilakukan penentuan banyak komponen yang optimal menggunakan PRESS. Sementara, untuk regresi Ridge, dilakukan pemilihan koefisien penalti λ menggunakan pendekatan intuitif grafis. Respon suhu udara dimodelkan dengan pendekatan univariat. Stamet yang dipilih sebagai ilustrasi untuk dijelaskan lebih rinci adalah Dermaga. Sementara hasil untuk stamet lainnya disajikan dalam Lampiran 8-31 karena memiliki langkah yang sama.

4.2.3.1 Regresi PLS untuk Suhu Maksimum dan Suhu Minimum

Sebelum melakukan pemodelan PLS untuk T_{MAKS} dan T_{MIN} , langkah pertama yang dilakukan adalah menentukan jumlah komponen yang optimal dari 41 komponen untuk membangkitkan model prakiraan cuaca pada stamet Dermaga. Selanjutnya, melakukan validasi silang untuk menentukan jumlah komponen yang optimal di setiap iterasi model. Proses validasi silang di setiap iterasi akan menghasilkan nilai PRESS, Nilai PRESS digunakan sebagai acuan dalam pemilihan jumlah komponen yang optimal untuk suatu model, PLS atau secara khusus PLS1 merupakan salah satu pemodelan univariat, sehingga validasi silang menggunakan PRESS dilakukan untuk tiap respon.

Ketika satu komponen dibangkitkan, nilai PRESS model tersebut dihitung. Proses ini berlanjut hingga mencapai banyaknya komponen yang maksimum dapat dibangkitkan. Setelah seluruh komponen dibangkitkan, jumlah komponen optimal ditetapkan berdasarkan nilai PRESS terendah. Nilai PRESS dari tiap komponen pada stamet Dermaga untuk suhu maksimum T_{MAKS} dan suhu minimum T_{MIN} dihitung berdasarkan algoritma validasi silang Kernel PLS dalam Lampiran 36-b, disajikan pada Tabel 4.3.

Tabel 4.3 menunjukkan nilai PRESS pada tiap iterasi proses regresi PLS menggunakan data validasi, Nilai PRESS terendah untuk suhu maksimum sebesar 0,6387 terdapat pada model PLS dengan 19 komponen. Hal ini menunjukkan bahwa model PLS yang optimal untuk meramalkan suhu maksimum diwakili oleh komponen yang cukup banyak. Pemilihan komponen yang relatif banyak sebaiknya dihindari karena menyebabkan reduksi dimensi sebagai tujuan utama justru tidak tercapai.

Tabel 4.3 Nilai PRESS Model PLS Stamet Dermaga

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MAKS}	1	0,7587	15	0,6412	29	0,6397
	2	0,6997	16	0,6409	30	0,6418
	3	0,6649	17	0,6402	31	0,6412
	4	0,6554	18	0,6397	32	0,6406
	5	0,6498	19	0,6387	33	0,6412
	6	0,6475	20	0,6390	34	0,6410
	7	0,6467	21	0,6416	35	0,6411
	8	0,6450	22	0,6420	36	0,6416
	9	0,6450	23	0,6415	37	0,6415
	10	0,6422	24	0,6408	38	0,6416
	11	0,6431	25	0,6409	39	0,6414
	12	0,6426	26	0,6407	40	0,6419
	13	0,6403	27	0,6410	41	0,6422
	14	0,6415	28	0,6436		

Tabel 4.3 Nilai PRESS Model PLS Stamet Dermaga (Lanjutan)

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MIN}	1	0,8971	15	0,8869	29	0,8848
	2	0,8862	16	0,8880	30	0,8864
	3	0,8810	17	0,8899	31	0,8878
	4	0,8792	18	0,8914	32	0,8842
	5	0,8798	19	0,8928	33	0,8842
	6	0,8800	20	0,8935	34	0,8858
	7	0,8808	21	0,8925	35	0,8852
	8	0,8831	22	0,8922	36	0,8867
	9	0,8841	23	0,8889	37	0,8871
	10	0,8842	24	0,8882	38	0,8864
	11	0,8845	25	0,8880	39	0,8865
	12	0,8842	26	0,8878	40	0,8861
	13	0,8846	27	0,8823	41	0,8838
	14	0,8847	28	0,8849		

Jika model PLS untuk suhu maksimum diindikasikan lebih optimal apabila menggunakan komponen yang relatif banyak, sebaliknya model PLS untuk suhu minimum hanya memerlukan 4 komponen saja. Namun tetap diputuskan jumlah komponen yang digunakan adalah 19 dan 4 untuk masing-masing suhu maksimum dan suhu minimum. Nilai PRESS untuk 7 stamet lainnya disajikan dalam Lampiran 8-14. Nilai PRESS terendah dan banyak komponen yang optimal untuk pemodelan regresi PLS di delapan stasiun disajikan pada Tabel 4.4.

Tabel 4.4 Banyak Komponen Optimal Model PLS pada Delapan Stamet

	Stamet	Banyak Komponen	PRESS terendah	Stamet	Banyak Komponen	PRESS terendah
	T_{MAKS}	Cengkareng	8	0,6990	Kemayoran	8
	Citeko	9	0,7299	Pd. Betung	22	0,7154
	Curug	19	0,7013	Priok	8	0,7520
	Dermaga	19	0,6441	Tangerang	8	0,7050
	Stamet	Banyak Komponen	PRESS terendah	Stamet	Banyak Komponen	PRESS terendah
	T_{MIN}	Cengkareng	6	0,8937	Kemayoran	4
	Citeko	8	0,7977	Pd. Betung	8	0,9072
	Curug	4	0,7603	Priok	4	0,8814
	Dermaga	7	0,8806	Tangerang	3	0,9483

Jumlah komponen laten yang digunakan untuk pemodelan PLS adalah yang terbanyak dibandingkan stamet lainnya. Berdasarkan Tabel 4.4, stamet Pondok Betung memiliki jumlah komponen terbesar yaitu sebanyak 22 komponen yang akan digunakan untuk menduga T_{MAKS} . Tingginya jumlah komponen yang optimal ini mengindikasikan bahwa kondisi cuaca yang memengaruhi suhu maksimum di stamet Pondok Betung tidak saling berhubungan antar parameter, misal arah angin dan curah hujan hampir saling independen.

Selain itu, dilihat dari rendahnya jumlah komponen yang optimal seperti pada stamet Priok dan Kemayoran, mengindikasikan bahwa korelasi yang tinggi antar parameter cuaca mempengaruhi observasi suhu maksimum pada stamet yang berada di Jakarta. Artinya, kondisi cuaca di kedua stamet tersebut cukup dinamis. Sedangkan dilihat dari tingginya komponen cuaca yang optimal, hampir seluruh stamet yang ada di Bogor dan Banten menunjukkan pola cuaca yang lebih statis.

Model PLS untuk suhu minimum tidak membutuhkan komponen laten yang cukup banyak jika dibandingkan dengan suhu maksimum, persebaran banyak komponen juga lebih beragam. Namun, nilai PRESS suhu minimum pada tiap

stamet relatif lebih tinggi dari suhu maksimum. Hal ini dapat mengakibatkan prakiraan suhu minimum menjadi sedikit tidak stabil apabila terdapat observasi cuaca baru. Setelah didapatkan jumlah komponen yang optimal pada tiap stamet dan digunakan sebagai komponen-komponen yang berkontribusi terhadap pemodelan, selanjutnya dilakukan regresi PLS untuk prakiraan T_{MAKS} dan T_{MIN} . Koefisien regresi dalam bentuk standarisasi untuk model PLS T_{MAKS} dan T_{MIN} di stamet Dermaga disajikan dalam Tabel 4.5 sebagai berikut.

Tabel 4.5 Koefisien Regresi Model PLS Stamet Dermaga

Var. Prediktor	T_{MAKS}	T_{MIN}	Var. Prediktor	T_{MAKS}	T_{MIN}
PC.dpsdt	-0,112	0,003	PC.temp2	0,318	-0,043
PC.mixr1	0,130	0,093	PC.temp4	0,017	-0,094
PC.mixr2	0,139	-0,077	PC.tmaxscr	-0,538	0,060
PC.mixr4	0,034	0,054	PC.tminscr	-0,128	-0,073
PC1.omega1	0,195	0,012	PC.tpan	0,092	-0,005
PC2.omega1	0,288	0,005	PC.tscrn	-0,169	-0,021
PC1.omega2	-0,144	0,006	PC.u1	-0,434	-0,002
PC2.omega2	-0,040	0,008	PC.u2	-0,415	0,014
PC1.omega4	-0,035	-0,008	PC.u4	0,169	0,023
PC2.omega4	-0,064	0,044	PC1.ustar	0,081	0,051
PC.pblh	-0,019	0,048	PC2.ustar	0,013	-0,033
PC.ps	0,022	0,020	PC1.v1	0,049	-0,008
PC.psl	0,026	0,023	PC2.v1	0,028	-0,0004
PC1.qgscrn	-0,038	-0,089	PC1.v2	-0,139	-0,007
PC2.qgscrn	-0,081	0,003	PC2.v2	0,106	-0,007
PC1.rh1	0,029	-0,060	PC.v4	-0,009	-0,027
PC2.rh1	-0,061	-0,059	PC.zg1	-0,086	0,020
PC.rh2	0,041	-0,045	PC.zg2	0,020	0,012
PC.rh4	0,029	-0,009	PC1.zg4	-0,0002	-0,059
PC.rnd	0,073	-0,011	PC2.zg4	0,067	0,060
PC.temp1	-0,115	-0,029			

Koefisien regresi pada Tabel 4.5 merupakan koefisien yang terstandarkan sehingga lebih mudah untuk membandingkan kontribusi masing-masing parameter NWP terhadap model. Model PLS T_{MIN} cenderung didominasi oleh koefisien yang mendekati nol. Hal ini mengindikasikan bahwa kontribusi tiap parameter NWP terhadap model T_{MIN} relatif lebih seimbang.

Selanjutnya, dibentuk model PLS berdasarkan koefisien regresi pada Tabel 4.5. Berikut merupakan model prakiraan elemen cuaca T_{MAKS} dan T_{MIN} menggunakan metode PLS di stamet Dermaga.

Tabel 4.6 Model PLS untuk Prakiraan Cuaca Stamet Dermaga

$\hat{y}_{T_{MAKS}}$	$= -0,179PC.dpsdt - 0,187PC.mixr1 + 0,193PC.mixr2 - 0,074PC.mixr4$ $+ 0,258PC1.omega1 + 0,307PC2.omega1 - 0,148PC1.omega2 - 0,019PC2.omega2$ $- 0,015PC1.omega4 - 0,100PC2.omega4 - 0,010PC.pblh + 0,021PC.ps$ $- 0,026PC.psl - 0,034PC1.qgscrn - 0,048PC2.qgscrn - 0,058PC1.rh1$ $- 0,088PC2.rh1 + 0,050PC.rh2 + 0,049PC.rh4 + 0,068PC.rnd - 0,174PC.temp1$ $+ 0,471PC.temp2 - 0,015PC.temp4 - 0,523PC.t max scr - 0,143PC.t min scr$ $+ 0,083PC.tpan - 0,181PC.tscrn + 0,444PC.u1 - 0,439PC.u2 + 0,202PC.u4$ $+ 0,052PC1.ustar + 0,024PC2.ustar - 0,010PC1.v1 - 0,036PC2.v1 - 0,116PC1.v2$ $+ 0,092PC2.v2 - 0,006PC.v4 - 0,153PC.zg1 + 0,044PC.zg2 + 0,005PC1.zg4$ $+ 0,123PC2.zg4$
$\hat{y}_{T_{MIN}}$	$= 0,002PC.dpsdt - 0,084PC.mixr1 - 0,093PC.mixr2 - 0,038PC.mixr4$ $+ 0,027PC1.omega1 - 0,014PC2.omega1 + 0,010PC1.omega2 - 0,004PC2.omega2$ $- 0,010PC1.omega4 + 0,073PC2.omega4 + 0,015PC.pblh + 0,054PC.ps$ $- 0,058PC.psl - 0,086PC1.qgscrn + 0,017PC2.qgscrn + 0,043PC1.rh1 - 0,073PC2.rh1$ $- 0,061PC.rh2 + 0,014PC.rh4 - 0,009PC.rnd - 0,048PC.temp1 - 0,041PC.temp2$ $- 0,122PC.temp4 + 0,109PC.t max scr - 0,081PC.t min scr + 0,008PC.tpan$ $- 0,039PC.tscrn - 0,014PC.u1 + 0,006PC.u2 + 0,040PC.u4 + 0,065PC1.ustar$ $- 0,059PC2.ustar - 0,011PC1.v1 - 0,036PC2.v1 - 0,008PC1.v2 + 0,018PC2.v2$ $- 0,026PC.v4 + 0,022PC.zg1 + 0,008PC.zg2 - 0,066PC1.zg4 + 0,070PC2.zg4$

Model pada Tabel 4.6 merupakan model yang masih memuat komponen utama (PC). Model yang mewakili sembilan grid hanya bisa diperoleh apabila tiap komponen utama pada Tabel 4.6 dipartisi, yaitu dengan mengalikan skor komponen utama terhadap tiap vektor eigennya. Matriks vektor eigen berukuran $s \times k$ dengan s adalah banyak grid dan k adalah banyak komponen utama. Berikut ini merupakan ringkasan vektor eigen untuk tiap 41 komponen pada stamet Dermaga dengan kolom pertama menunjukkan vektor eigen untuk $dpsdt$, begitu seterusnya hingga kolom terakhir untuk $zg4$.

$$\mathbf{E} = \begin{matrix} & \text{dpsdt} & \text{mixr1} & \text{mixr2} & \text{mixr4} & \text{omega1} & \text{omega1} & \text{zg4} & \text{zg4} \\ \left[\begin{array}{cccccccc} -0,3333 & -0,3411 & -0,3353 & -0,3337 & -0,2908 & -0,4402 & \dots & -0,2007 & 0,4516 \\ -0,3333 & -0,3220 & -0,3339 & -0,3346 & -0,2546 & -0,5264 & \dots & -0,3958 & -0,2278 \\ -0,3333 & -0,3131 & -0,3239 & -0,3301 & -0,2129 & -0,5401 & \dots & -0,4396 & 0,0187 \\ -0,3334 & -0,3494 & -0,3396 & -0,3370 & -0,3826 & 0,0026 & \dots & -0,0704 & 0,5045 \\ -0,3334 & -0,3565 & -0,3432 & -0,3395 & -0,3931 & 0,1631 & \dots & -0,4394 & -0,0633 \\ -0,3334 & -0,3478 & -0,3393 & -0,3355 & -0,3614 & 0,1186 & \dots & -0,4080 & 0,1937 \\ -0,3333 & -0,3270 & -0,3260 & -0,3233 & -0,3439 & 0,2151 & \dots & 0,1027 & 0,4975 \\ -0,3333 & -0,3226 & -0,3281 & -0,3337 & -0,3556 & 0,3005 & \dots & -0,2504 & -0,4224 \\ -0,3333 & -0,3174 & -0,3301 & -0,3323 & -0,3593 & 0,2454 & \dots & -0,4149 & 0,1481 \end{array} \right. \end{matrix}$$

Matriks eigen vektor **E** mengindikasikan bahwa mixr 1 dan mixr 2 memiliki pengaruh yang negatif terhadap suhu udara. Artinya, tingginya rasio jumlah uap air di udara pada level 1 dan level 2 cenderung mengakibatkan suhu udara berada pada kisaran yang rendah, demikian pula sebaliknya. Setelah memperoleh matriks eigen vektor dan mengalikannya dengan komponen utama yang bersesuaian, maka diperoleh model PLS untuk stamet Dermaga yang mewakili seluruh parameter NWP pada sembilan grid pada Lampiran 15. Untuk mendapatkan prakiraan cuaca, model PLS pada Tabel 4.6 atau model lebih detail pada Lampiran 15 harus dikembalikan ke bentuk awal dengan mengalikannya terhadap standar deviasi lalu dijumlahkan dengan mean (rata-rata). Secara ringkas, hasil prakiraan cuaca untuk 8 stamet disajikan pada Lampiran 16-17.

4.2.3.2 Regresi PCR untuk Suhu Maksimum dan Suhu Minimum

Sama halnya dengan regresi PLS, langkah yang dilakukan sebelum memodelkan cuaca dengan regresi PCR adalah menetapkan jumlah komponen yang optimal dari 41 komponen yang dimiliki stamet Dermaga. Walaupun terdapat kesamaan pada prosedur awal, namun PCR tidak memilih komponen optimal berdasarkan RMSE terendah seperti pada regresi PLS. Komponen terpilih adalah komponen yang mampu merepresentasikan variabilitas prediktor minimal 80%. Artinya, pemilihan komponen optimal pada PCR lebih didasari oleh persentase kumulatif varians daripada menggunakan PRESS.

Serupa dengan PLS, PCR juga dapat diklasifikasikan untuk respon multivariate atau univariat. Namun untuk kasus ini, ditetapkan untuk memodelkan suhu udara dengan PCR univariat agar konsisten dengan regresi PLS yang juga

dimodelkan dengan respon univariat. Tabel 4.7 menyajikan varians yang merepresentasikan variabilitas prediktor pada stamet Dermaga. Varians menjelaskan untuk prediktor pada model PCR suhu maksimum dan suhu minimum adalah sama karena proses perhitungannya yang tidak melibatkan variabel respon.

Tabel 4.7 Kumulatif Varians 41 Komponen Prediktor Stamet Dermaga

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	27,14	15	96,70	29	99,88
2	43,94	16	97,36	30	99,92
3	56,45	17	97,89	31	99,95
4	63,63	18	98,28	32	99,96
5	69,72	19	98,57	33	99,97
6	75,38	20	98,83	34	99,98
7	80,74	21	99,06	35	99,98
8	84,84	22	99,24	36	99,99
9	87,97	23	99,40	37	100
10	90,22	24	99,51	38	100
11	92,40	25	99,61	39	100
12	93,93	26	99,71	40	100
13	95,09	27	99,77	41	100
14	95,97	28	99,83		
Variabilitas T_{MAKS} menjelaskan				45,40%	
Variabilitas T_{MIN} menjelaskan				23,05%	

Berdasarkan Tabel 4.7, 7 dari 41 komponen mampu untuk mewakili 80% lebih variabilitas prediktor sehingga dapat disimpulkan bahwa banyak komponen yang optimal untuk model PCR suhu udara, baik T_{MAKS} maupun T_{MIN}, adalah 7. Ketujuh komponen tersebut dapat menjelaskan variabilitas suhu maksimum sebesar 45,40% , sedangkan untuk suhu minimum hanya separuhnya saja yaitu 23,05%. Hal ini berarti bahwa terdapat cukup banyak parameter cuaca di atmosfer atau di permukaan bumi yang memengaruhi suhu maksimum dan suhu minimum. Kumulatif varians untuk 7 stamet lainnya disajikan pada Lampiran18-24. Banyak komponen optimal berdasarkan keragaman prediktor yang direpresentasikan oleh kumulatif varians disajikan pada Tabel 4.8.

Tabel 4.8 Jumlah Komponen Optimal PCR pada Delapan Stasiun Meteorologi

Stamet	Banyak Komponen	%Varians Prediktor	Stamet	Banyak Komponen	%Varians Prediktor
Cengkareng	7	82,29	Kemayoran	7	82,32
Citeko	7	80,60	Pd. Betung	7	81,93
Curug	7	80,32	Priok	7	81,46
Dermaga	7	80,74	Tangerang	7	82,04

Berdasarkan Tabel 4.8, 7 komponen diindikasikan telah merepresentasikan minimal 80% variabilitas prediktor. Setelah memperoleh komponen hasil PCA yang bebas multikolinearitas, maka komponen tersebut menjadi prediktor baru yang akan diregresikan terhadap T_{MAKS} dan T_{MIN} . Koefisien regresi untuk model PCR T_{MAKS} dan T_{MIN} di stamet Dermaga, disajikan pada Tabel 4.9

Tabel 4.9 Koefisien Regresi Model PCR Stamet Dermaga

Var. Prediktor	T_{MAKS}	T_{MIN}	Var. Prediktor	T_{MAKS}	T_{MIN}
PC.dpsdt	0,0308	-0.0094	PC.temp2	-0,0551	-0.0382
PC.mixr1	0,0216	-0.0856	PC.temp4	-0,0320	-0.0554
PC.mixr2	0,0067	-0.0619	PC.tmaxscr	-0,0684	0.0020
PC.mixr4	0,0078	-0.0627	PC.tminscr	-0,0579	-0.0558
PC1.omega1	0,0152	-0.0036	PC.tpan	-0,0771	-0.0289
PC2.omega1	0,0415	0.0063	PC.tscrn	-0,0687	-0.0138
PC1.omega2	0,0196	-0.0016	PC.u1	0,0714	0.0284
PC2.omega2	0,0388	0.0042	PC.u2	0,0703	0.0313
PC1.omega4	0,0086	-0.0149	PC.u4	0,0672	0.0190
PC2.omega4	0,0077	0.0047	PC1.ustar	-0,0087	0.0221
PC.pblh	-0,0496	0.0471	PC2.ustar	-0,0216	-0.0013
PC.ps	0,0169	-0.0184	PC1.v1	0,0057	0.0038
PC.psl	-0,0163	0.0167	PC2.v1	-0,0043	0.0220
PC1.qgscrn	0,0093	-0.0815	PC1.v2	0,0040	0.0014
PC2.qgscrn	-0,0015	-0.0139	PC2.v2	0,0130	-0.0236
PC1.rh1	-0,0573	0.0588	PC.v4	-0,0127	-0.0198
PC2.rh1	0,0151	-0.0283	PC.zg1	-0,0311	0.0175
PC.rh2	0,0319	-0.0343	PC.zg2	-0,0394	0.0128
PC.rh4	0,0207	-0.0310	PC1.zg4	-0,0625	-0.0445
PC.rnd	0,0564	-0.0062	PC2.zg4	-0,0012	0.0455
PC.temp1	-0,0688	-0.0191			

Berdasarkan Tabel 4.9, dapat dikatakan bahwa parameter model PCR, baik untuk suhu maksimum maupun suhu minimum, lebih banyak didominasi oleh koefisien yang mendekati nol. Kondisi yang berbeda terlihat pada model PLS dimana koefisiennya untuk suhu minimum cenderung lebih seragam dan bernilai mendekati nol. Hal ini mengindikasikan bahwa kontribusi tiap parameter NWP terhadap model PCR lebih rendah daripada model PLS. Sehingga terdapat kemungkinan bahwa kondisi ini akan berpengaruh terhadap akurasi prakiraan cuaca.

Setelah memperoleh koefisien regresi, langkah selanjutnya adalah membentuk model PCR berdasarkan koefisien regresi pada Tabel 4.9. Model prakiraan elemen cuaca T_{MAKS} dan T_{MIN} menggunakan regresi PCR di stamet Dermaga disajikan dalam Tabel 4.10.

Tabel 4.10 Model PCR untuk Prakiraan Cuaca Stamet Dermaga

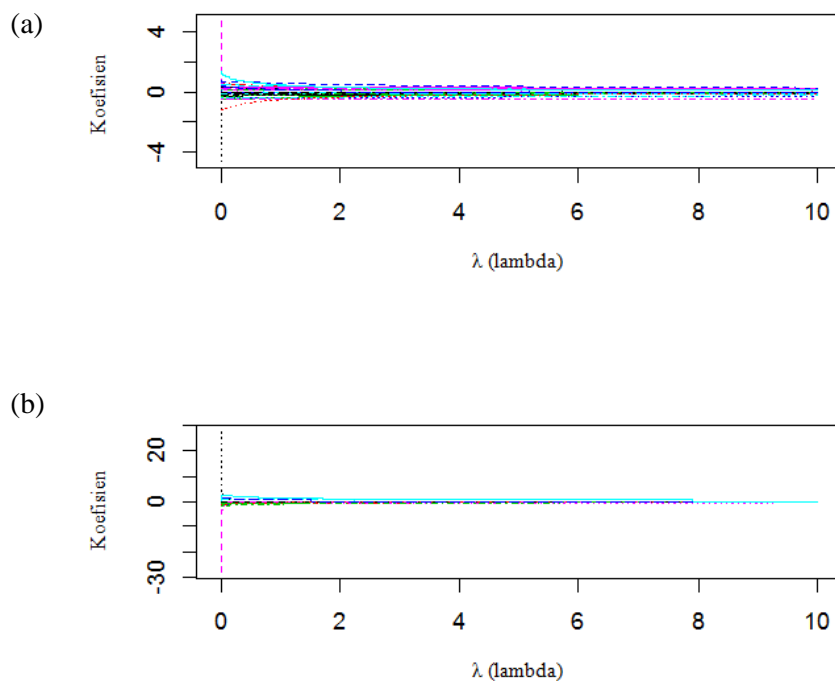
$\hat{y}_{T_{MAKS}}$	$= 0,0308PC.dpsdt + 0,0216PC.mixr1 + 0,0067PC.mixr2 + 0,0078PC.mixr4$ $+ 0,0152PC1.omega1 + 0,0415PC2.omega1 - 0,0196PC1.omega2 + 0,0388PC2.omega2$ $- 0,0086PC1.omega4 + 0,0077PC2.omega4 - 0,0496PC.pblh + 0,0169PC.ps$ $- 0,0163PC.psl + 0,0093PC1.qgscrn - 0,0015PC2.qgscrn - 0,0573PC1.rh1$ $+ 0,0151PC2.rh1 + 0,0319PC.rh2 + 0,0207PC.rh4 + 0,0564PC.rnd - 0,0688PC.temp1$ $- 0,0551PC.temp2 - 0,0320PC.temp4 - 0,0684PC.t max scr - 0,0579PC.t min scr$ $- 0,0771PC.tpan - 0,0687PC.tscrn + 0,0714PC.u1 + 0,0703PC.u2 + 0,0672PC.u4$ $- 0,0087PC1.ustar - 0,0216PC2.ustar + 0,0057PC1.v1 - 0,0043PC2.v1 + 0,0040PC1.v2$ $+ 0,0130PC2.v2 - 0,0127PC.v4 - 0,0311PC.zg1 - 0,0394PC.zg2$ $- 0,0625PC1.zg4 - 0,0012PC2.zg4,$
$\hat{y}_{T_{MIN}}$	$= -0,0094PC.dpsdt - 0,0856PC.mixr1 - 0,0619PC.mixr2 - 0,0627PC.mixr4$ $- 0,0036PC1.omega1 + 0,0063PC2.omega1 - 0,0016PC1.omega2 + 0,0042PC2.omega2$ $- 0,0149PC1.omega4 + 0,0047PC2.omega4 + 0,0471PC.pblh - 0,0184PC.ps$ $+ 0,0167PC.psl - 0,0815PC1.qgscrn - 0,0139PC2.qgscrn + 0,0588PC1.rh1$ $- 0,0283PC2.rh1 - 0,0343PC.rh2 - 0,0310PC.rh4 - 0,0062PC.rnd - 0,0191PC.temp1$ $- 0,0382PC.temp2 - 0,0554PC.temp4 + 0,0020PC.t max scr - 0,0558PC.t min scr$ $- 0,0289PC.tpan - 0,0138PC.tscrn + 0,0284PC.u1 + 0,0313PC.u2 + 0,0190PC.u4$ $+ 0,0221PC1.ustar - 0,0013PC2.ustar + 0,0038PC1.v1 + 0,0220PC2.v1 + 0,0014PC1.v2$ $- 0,0236PC2.v2 - 0,0198PC.v4 + 0,0175PC.zg1 + 0,0128PC.zg2$ $- 0,0445PC1.zg4 + 0,0455PC2.zg4$

Sama halnya dengan PLS, model PCR pada Tabel 4.10 memuat komponen utama. Model PCR yang mewakili sembilan grid hanya bisa diperoleh jika tiap komponen utama pada Tabel 4.10 dipartisi, yaitu dengan mengalikan skor komponen utama terhadap tiap vektor eigennya. Matriks eigen vektor model PCR adalah sama dengan eigen vektor model PLS. Setelah memperoleh matriks eigen vektor dan mengalikannya dengan komponen utama yang bersesuaian, maka model PCR untuk stamet Dermaga yang mewakili seluruh parameter NWP di sembilan grid disajikan pada Lampiran 25. Model PCR pada Tabel 4.10 atau model pada Lampiran 25 harus ditransformasi ke bentuk asal sehingga diperoleh suhu minimum pada kisaran 18°C-25°C dan suhu maksimum pada kisaran 25°C-35°C, seperti yang disajikan pada Lampiran 26-27.

4.2.3.3 Regresi Ridge untuk Suhu Maksimum dan Suhu Minimum

Setelah mendapatkan anggota *ensemble* dari model PLS dan PCR, selanjutnya dilakukan pemodelan cuaca dengan regresi lain yang mampu mereduksi efek multikolinearitas, yaitu regresi Ridge. Berbeda dengan dua metode sebelumnya, metode ini tidak memerlukan pemilihan banyak komponen optimal. Metode ini menggunakan suatu konstanta λ untuk meminimalkan dampak dari singularitas $\mathbf{X}^T\mathbf{X}$.

Konstanta λ diharapkan bernilai tidak terlalu besar. Hal ini untuk menghindari nilai koefisien regresi Ridge yang semakin mendekati 0. Jika terjadi, maka nilai dugaan dari respon juga turut mendekati 0. Padahal suhu udara di Indonesia hampir tidak pernah berada pada kisaran 0°C - 10°C . Gambar 4.2 membantu penentuan konstanta λ secara visual.



Gambar 4.2 Konvergensi Koefisien Regresi Ridge

(a) Suhu Maksimum; (b) Suhu Minimum

Penentuan konvergensi berdasarkan Gambar 4.2 seharusnya dihindari karena subjektif. Namun, pada kasus ini, konstanta λ bahkan tetap tidak konvergen saat λ mendekati 100 sehingga diputuskan untuk menggunakan bantuan visual saja dalam menentukan λ . Berdasarkan Draper dan Smith (1992), penentuan batas bawah dan batas atas untuk λ tidak ditetapkan dengan ketat. Tetapi, sebaiknya batas

atas tidak lebih besar daripada 10 atau 20 untuk menghindari koefisien regresi yang semakin tidak bermakna karena bernilai mendekati 0.

Dari Gambar 4.2, diindikasikan bahwa koefisien regresi untuk T_{MAKS} dan T_{MIN} konvergen saat λ sebesar 9 atau lebih. Tabel 4.11 adalah koefisien regresi Ridge untuk stamet Dermaga dengan $\lambda = 9$, sementara konstanta λ untuk tujuh stamet lainnya ditampilkan dalam Lampiran 28.

Tabel 4.11 Koefisien Regresi Ridge Stamet Dermaga

Var. Prediktor	T_{MAKS}	T_{MIN}	Var. Prediktor	T_{MAKS}	T_{MIN}
PC.dpsdt	-0,075	-0,099	PC.temp2	0,256	0,066
PC.mixr1	0,123	-0,125	PC.temp4	0,008	-0,219
PC.mixr2	0,102	-0,129	PC.tmaxscr	-0,513	0,149
PC.mixr4	0,030	0,008	PC.tminscr	-0,108	-0,084
PC1.omega1	0,153	0,165	PC.tpan	0,053	-0,026
PC2.omega1	0,212	0,015	PC.tscrn	-0,099	-0,010
PC1.omega2	-0,099	-0,051	PC.u1	-0,301	0,055
PC2.omega2	-0,029	-0,038	PC.u2	-0,236	0,016
PC1.omega4	-0,017	0,083	PC.u4	0,141	0,124
PC2.omega4	-0,082	0,031	PC1.ustar	0,048	0,014
PC.pblh	-0,041	0,010	PC2.ustar	0,007	-0,088
PC.ps	0,024	0,037	PC1.v1	-0,012	-0,075
PC.psl	0,027	0,066	PC2.v1	0,012	-0,049
PC1.qgscrn	-0,005	-0,153	PC1.v2	-0,094	0,010
PC2.qgscrn	-0,067	0,054	PC2.v2	0,089	0,053
PC1.rh1	0,027	-0,021	PC.v4	-0,032	-0,021
PC2.rh1	-0,069	-0,172	PC.zg1	-0,096	0,004
PC.rh2	0,031	-0,261	PC.zg2	0,022	-0,166
PC.rh4	0,037	0,009	PC1.zg4	0,001	0,022
PC.rnd	0,071	-0,022	PC2.zg4	0,079	0,290
PC.temp1	-0,108	-0,083			

Karena koefisien regresi pada Tabel 4.11 masih dalam bentuk standarisasi, maka akan lebih mudah untuk membandingkan kontribusi masing-masing parameter NWP terhadap model. Dapat dikatakan bahwa model Ridge T_{MIN} tidak lagi didominasi oleh koefisien yang mendekati nol, seperti pada regresi PLS dan regresi PCR. Hal ini mengindikasikan bahwa kontribusi beberapa parameter NWP terhadap model T_{MIN} relatif lebih bermakna pada Ridge daripada PLS atau PCR. Kemungkinan hal ini dapat meningkatkan akurasi prakiraan cuaca atau justru tidak berpengaruh sama sekali.

Setelah koefisien regresi diperoleh pada tahap sebelumnya, maka langkah selanjutnya ialah membentuk model berdasarkan koefisien regresi pada Tabel 4.11. Model prakiraan elemen cuaca T_{MAKS} dan T_{MIN} menggunakan regresi Ridge di stamet Dermaga disajikan pada Tabel 4.12.

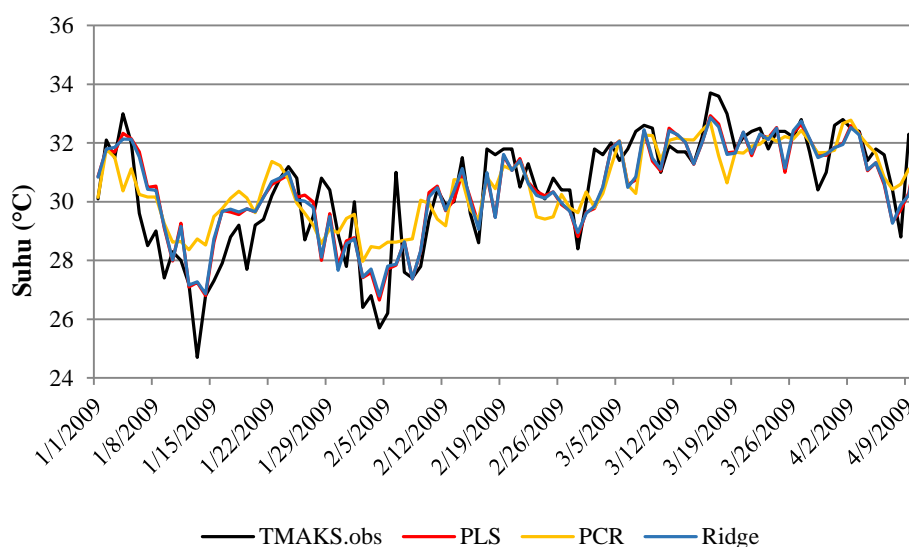
Tabel 4.12 Model Regresi Ridge untuk Prakiraan Cuaca Stamet Dermaga

$$\begin{aligned} \hat{y}_{T_{MAKS}} = & -0,075PC.dpsdt + 0,123PC.mixr1 + 0,102PC.mixr2 + 0,03PC.mixr4 + 0,153PC1.omega1 + \\ & + 0,212PC2.omega1 - 0,099PC1.omega2 - 0,029PC2.omega2 - 0,017PC1.omega4 + \\ & - 0,082PC2.omega4 - 0,041PC.pblh + 0,024PC.ps + 0,027PC.psl - 0,005PC1.qgscrn + \\ & - 0,067PC2.qgscrn + 0,027PC1.rh1 - 0,069PC2.rh1 + 0,031PC.rh2 + 0,037PC.rh4 + \\ & + 0,071PC.rnd - 0,108PC.temp1 + 0,256PC.temp2 + 0,008PC.temp4 - 0,513PC.tmaxscr + \\ & - 0,108PC.tminscr + 0,053PC.tpan - 0,099PC.tscrn - 0,301PC.u1 - 0,236PC.u2 + \\ & + 0,141PC.u4 + 0,048PC1.ustar + 0,007PC2.ustar - 0,012PC1.v1 + 0,012PC2.v1 + \\ & - 0,094PC1.v2 + 0,089PC2.v2 - 0,032PC.v4 - 0,096PC.zg1 + 0,022PC.zg2 + \\ & + 0,001PC1.zg4 + 0,079PC2.zg4 \\ \hat{y}_{T_{MIN}} = & -0,099PC.dpsdt - 0,125PC.mixr1 - 0,129PC.mixr2 + 0,008PC.mixr4 + 0,165PC1.omega1 + \\ & + 0,015PC2.omega1 - 0,051PC1.omega2 - 0,038PC2.omega2 + 0,083PC1.omega4 + \\ & + 0,031PC2.omega4 + 0,01PC.pblh + 0,037PC.ps + 0,066PC.psl - 0,153PC1.qgscrn + \\ & + 0,054PC2.qgscrn - 0,021PC1.rh1 - 0,172PC2.rh1 - 0,261PC.rh2 + 0,009PC.rh4 + \\ & - 0,022PC.rnd - 0,083PC.temp1 + 0,066PC.temp2 - 0,219PC.temp4 + 0,149PC.tmaxscr + \\ & - 0,084PC.tminscr - 0,026PC.tpan - 0,01PC.tscrn + 0,055PC.u1 + 0,016PC.u2 + \\ & + 0,124PC.u4 + 0,014PC1.ustar - 0,088PC2.ustar - 0,075PC1.v1 - 0,049PC2.v1 + \\ & + 0,01PC1.v2 + 0,053PC2.v2 - 0,021PC.v4 + 0,004PC.zg1 - 0,166PC.zg2 + \\ & + 0,022PC1.zg4 + 0,29PC2.zg4 \end{aligned}$$

Dari Tabel 4.12, model Ridge masih memuat komponen utama (PC) sehingga untuk memperoleh model yang mewakili sembilan grid, tiap komponen utama pada Tabel 4.12 dipartisi, yaitu dengan mengalikan skor komponen utama terhadap tiap vektor eigennya. Matriks vektor eigen berukuran $s \times k$ dengan s adalah banyak grid dan k adalah banyak komponen utama. Untuk stamet Dermaga, matrik vektor eigen berukuran 9×41 sesuai dengan matrik E pada halaman 69. Setelah memperoleh matriks vektor eigen dan mengalikannya dengan komponen utama yang bersesuaian, maka model Ridge untuk stamet Dermaga yang mewakili seluruh parameter NWP di sembilan grid disajikan pada Lampiran 29. Untuk mengembalikan ke bentuk awal, prakiraan T_{MAKS} dan T_{MIN} dari Tabel 4.12 harus ditransformasi menggunakan mean dan deviasi standar berdasarkan Lampiran 1 dimana hasil prakiraan cuaca untuk seluruh 8 stamet ditampilkan dalam Lampiran 30-31.

4.2.4 Deskripsi Prakiraan Cuaca Anggota *Ensemble*

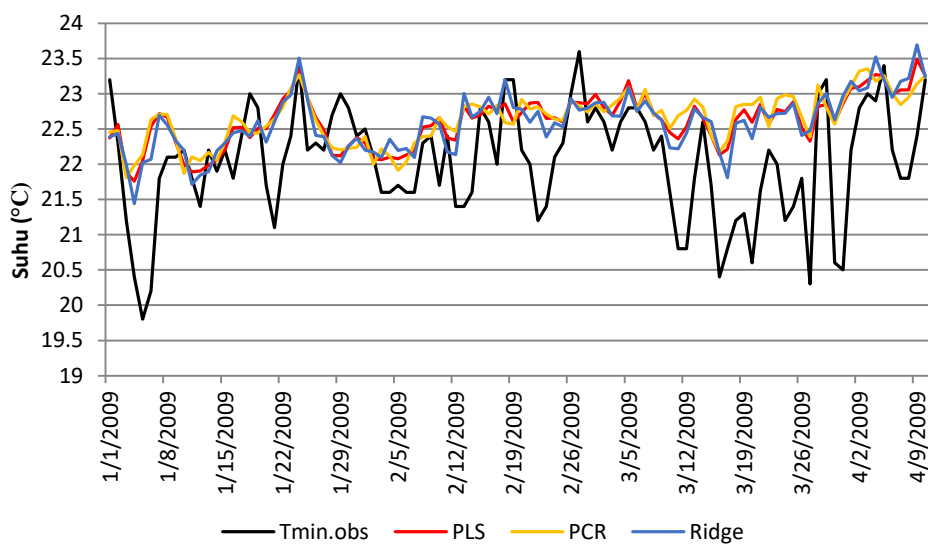
Setelah memperoleh model yang akan digunakan untuk meramalkan suhu udara, maka langkah berikutnya adalah membandingkan hasil ramalan tiap anggota *ensemble* dan nilai observasi. Pada kasus ini, anggota *ensemble* terdiri atas regresi PLS, PCR dan Ridge. Analisis deskriptif dilakukan dengan menggunakan plot runtun waktu untuk melihat seberapa baik ramalan anggota *ensemble* sebelum dikalibrasi oleh BMA. Gambar 4.3 dan Gambar 4.4 merupakan plot runtun waktu 100 hari pertama dari prakiraan *ensemble* dan observasi suhu udara untuk stamet Dermaga pada tahun 2009, masing-masing untuk suhu maksimum dan suhu minimum.



Gambar 4.3 Tren Prakiraan Anggota *Ensemble* dan Observasi Suhu Maksimum

Gambar 4.3 mengindikasikan bahwa tiap anggota *ensemble*, baik PLS, PCR maupun Ridge, telah dapat mengikuti pola umum dari suhu maksimum dimana jika tren suhu maksimum meningkat maka hasil prakiraan turut mengikuti naik, begitu pula sebaliknya jika tren suhu menurun. Namun, permasalahan yang timbul adalah kejadian *under-fitting* (ramalan berada di bawah nilai observasi) atau *over-fitting* (ramalan berada di atas nilai observasi) yang konsisten terjadi di hari yang sama. Sebagai contoh, observasi suhu maksimum pada tanggal 1 Februari 2009 adalah 30°C. Sementara, ramalan PLS, PCR dan Ridge pada tanggal yang sama adalah 28,8°C, 29,6°C, dan 28,7°C. Jika diperhatikan, error yang dihasilkan berada dalam skala yang sedang sebab berada pada kisaran 0,3°C.

Meskipun ramalan ketiganya sudah mendekati nilai observasi, namun *under-fitting* yang terjadi secara serentak ini akan menyebabkan proses pembobotan BMA condong kepada hasil yang paling dekat dengan observasi, dalam hal ini 29,6°C milik PCR. Dengan demikian, bobot yang diberikan BMA untuk PCR pada tanggal 1 Februari 2009 kemungkinan besar mendekati 0,9. Sementara, bobot untuk kedua anggota *ensemble* lainnya kemungkinan tidak lebih besar dari 0,1. Untuk memperkuat analisis awal tentang kebaikan ramalan anggota *ensemble*, diperlukan pula identifikasi secara visual untuk suhu minimum yang disajikan pada Gambar 4.4, khususnya pada stamet Dermaga dalam 100 hari pertama pada tahun 2009.



Gambar 4.4 Tren Prakiraan Anggota *Ensemble* dan Observasi Suhu Minimum

Gambar 4.4 mengindikasikan fakta yang sama dengan Gambar 4.3 dimana ramalan anggota *ensemble* mampu untuk menangkap pola dan tren fluktuasi dari suhu minimum. Selain itu, kejadian *under-fitting* atau *over-fitting* juga serentak terjadi pada PLS, PCR dan Ridge pada sebagian besar hari pengamatan. Tetapi, dapat dikatakan bahwa permasalahan yang lebih kompleks terjadi pada suhu minimum. Hal ini disebabkan musim penghujan tampaknya memiliki pengaruh yang cukup krusial terhadap tinggi rendahnya suhu minimum. Setelah musim beralih ke periode transisi, terlihat bahwa selisih yang lebar antara ramalan dan observasi yang sebelumnya cukup rutin terjadi sudah mengecil.

Meskipun ketiga anggota *ensemble* dapat menangkap pola suhu yang terjadi, namun ramalan yang dihasilkan cukup jauh dari observasi. Oleh karena itu,

diperlukan kalibrasi model agar dihasilkan prakiraan cuaca yang lebih akurat dan presisi. Proses pembobotan BMA yang dilakukan secara iteratif dengan menggunakan *rolling training window* diharapkan dapat meminimalkan dampak dari *under-fitting* atau *over-fitting*, atau bahkan dampak dari pola musiman.

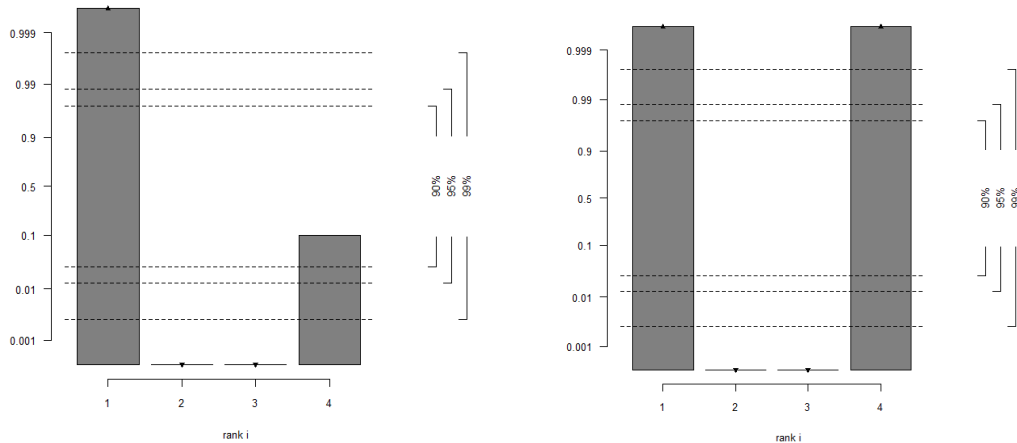
4.2.5 Kalibrasi Prakiraan Cuaca Menggunakan Spatial BMA

Berdasarkan sub-bab sebelumnya, diindikasikan bahwa hasil peramalan *ensemble*, khususnya untuk stamet Dermaga, masih memiliki akurasi yang cukup rendah. Oleh karena itu, diperlukan metode pemrosesan secara statistik untuk mengkalibrasi hasil peramalan agar bias ramalan menjadi lebih rendah. Kalibrasi dilakukan untuk melakukan penyesuaian terhadap varians, sehingga diperoleh ramalan yang lebih reliabel dengan varians yang proporsional dan memiliki interval prediksi yang lebih sempit.

Kalibrasi prakiraan cuaca pada 8 stamet secara individu menggunakan BMA pernah dilakukan oleh Luthfi (2017). Berdasarkan Raftery *et al.* (2005), dikatakan bahwa kalibrasi dengan BMA akan menghasilkan prakiraan yang lebih baik jika *ensemble range* memiliki korelasi yang signifikan dengan derajat kesalahan prakiraan. *Ensemble range* adalah selisih antara maksimum dan minimum dari suatu *ensemble*. Namun, korelasi yang signifikan tersebut belum tentu menjamin prakiraan *ensemble* sudah terkalibrasi dan tidak lagi *underdispersive* atau *overdispersive* yang dapat diidentifikasi dari *Verification Range Histogram* (VRH).

Penelitian Luthfi (2017) menunjukkan bahwa terdapat korelasi yang positif antara *ensemble range* dan eror mutlak. Korelasi untuk suhu maksimum adalah sebesar 0,11, sementara untuk suhu minimum sebesar 0,24. Namun, penggunaan korelasi saja belum cukup untuk dapat menyimpulkan kebakaran prakiraan *ensemble*. Sehingga dibutuhkan *Verification Range Histogram* (VRH) untuk mengidentifikasi apakah prakiraan *ensemble* bersifat *underdispersive*, *overdispersive*, atau *equidispersive*. Histogram yang menyerupai U mengindikasikan prakiraan *ensemble* yang *underdispersive*, sementara histogram yang menyerupai bentuk kurva distribusi normal mengindikasikan prakiraan *ensemble* yang *overdispersive* (Möller, 2014).

Pada Gambar 4.5, sumbu horizontal adalah rank ke- i dimana $i = 1, \dots, 4$ atau sesuai dengan banyak anggota *ensemble* ditambah satu observasi, sementara sumbu vertikal adalah kumulatif probabilitas Binomial dengan parameter n adalah total observasi dan p adalah $1/k$ dimana k adalah frekuensi observasi yang berada pada rank ke- i . Dalam kasus ini, *raw ensemble* beranggotakan PLS, PCR, dan Ridge.



Gambar 4.5 VRH *Raw Ensemble* Seluruh Stamet, 1 Jan '09 – 30 Jan '09
(kiri) suhu maksimum (kanan) suhu minimum

Berdasarkan Gambar 4.5, dapat diketahui bahwa prakiraan *ensemble*, baik suhu maksimum dan suhu minimum pada 8 stamet secara serentak, masih bersifat *underdispersive*. Hal ini ditunjukkan dengan bentuk histogram untuk masing-masing suhu maksimum dan suhu minimum menyerupai bentuk U, sehingga mengindikasikan bahwa masih banyak observasi suhu udara yang berada di luar *ensemble range*. Mekanisme VRH adalah dengan mengurutkan observasi suhu dan prakiraan anggota *ensemble* dari yang tertinggi hingga terendah (Hammill, 2001). Jika terdapat 3 anggota *ensemble* dimana observasi suhu pada 1 Februari 2009 lebih tinggi daripada prakiraan anggota *ensemble*, maka masuk ke dalam rank pertama, sebaliknya jika lebih rendah masuk kedalam rank keempat. Sementara, observasi terhitung masuk ke dalam rank kedua dan ketiga jika berada dalam *ensemble range*.

Berdasarkan Gambar 4.5, diketahui bahwa *coverage* prakiraan *ensemble* untuk T_{MAKS} dan T_{MIN} masing-masing 19,23% dan 9,57%, yang merupakan persentase banyaknya observasi yang berada di rank kedua dan rank ketiga. Nilai *coverage* yang berada jauh dibawah standar *coverage* 50%, memperkuat bahwa prakiraan *ensemble* MOS masih bersifat *underdispersive* sehingga interval prediksi

menjadi kurang presisi. Sebagai upaya untuk mengurangi dampak *underdispersive*, diperlukan kalibrasi prakiraan *ensemble*.

Berdasarkan penelitian Luthfi (2017), diindikasikan bahwa BMA sebagai pendekatan non-spasial dapat mengkalibrasi prakiraan *ensemble* dengan cukup baik. Dengan menerapkan pendekatan metode Spatial BMA, diharapkan dapat menangkap fenomena spasial yang terjadi dengan sekaligus dapat memanfaatkan informasi dari prakiraan *ensemble*, sehingga diperoleh prakiraan cuaca yang akurat dan reliabel.

Langkah pertama yang dilakukan dalam pemodelan Spatial BMA adalah meregresikan secara serentak seluruh stamet tiap anggota *ensemble* terhadap observasi sehingga diperoleh koefisien bias regresi, yaitu $\beta_{0,m}$ dan $\beta_{1,m}$, untuk tiap anggota m . Koefisien regresi dan bobot untuk tiap anggota m dan waktu t adalah sama untuk 8 stamet yang merupakan implementasi dari persamaan (4.12). Lampiran 32-33 menyajikan koefisien regresi dan bobot secara singkat. Sebagai contoh, Tabel 4.13 menyajikan koefisien regresi, bobot, hasil prakiraan yang akan dijadikan prediktor dalam model Spatial BMA.

Tabel 4.13 Estimasi Parameter dan Hasil Prakiraan BMA Simultan
Stamet Kemayoran, 8 Februari 2009

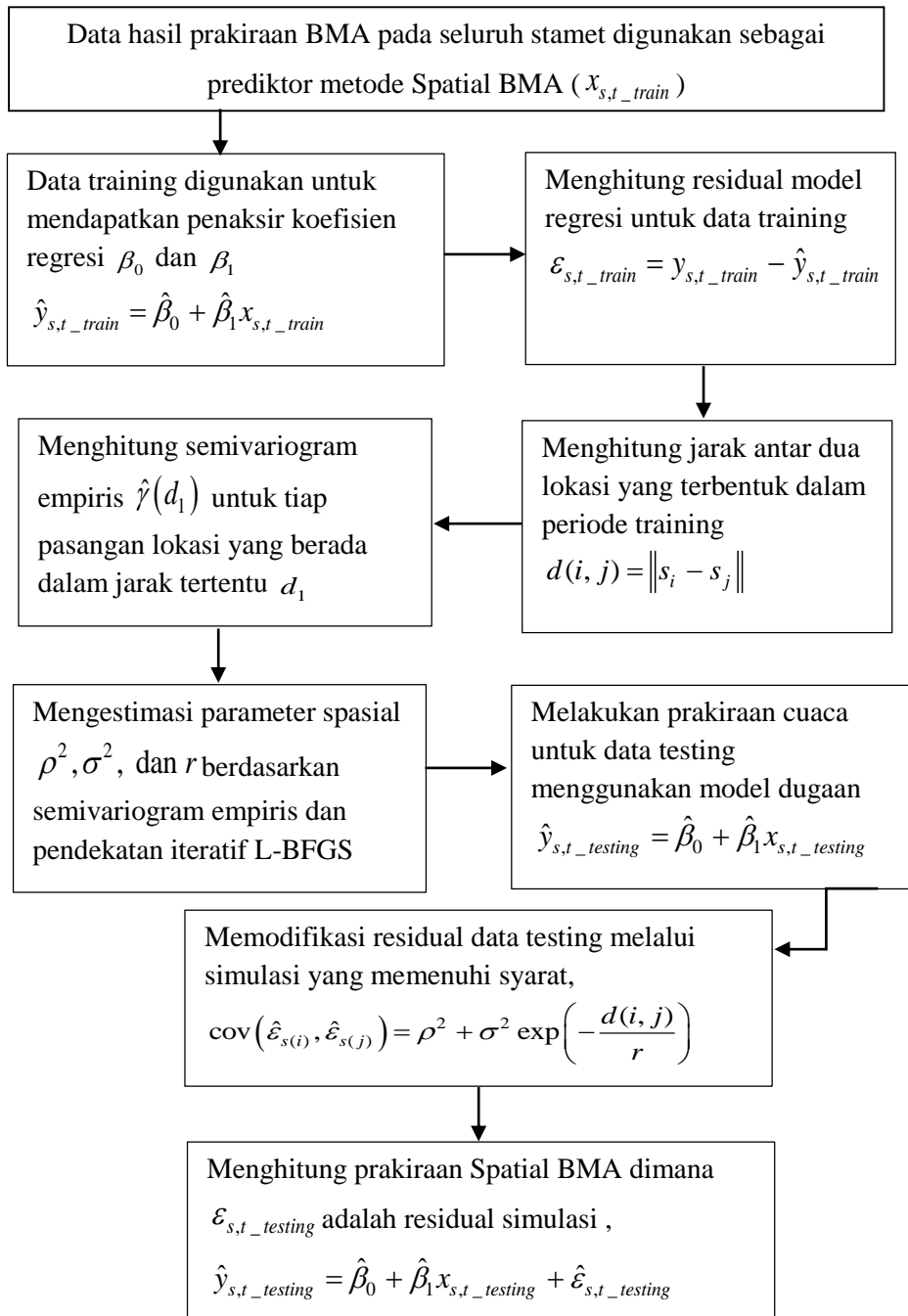
Respon	Model	β_0	β_1	w	Prakiraan Anggota Ensemble (°C)	Obs. (°C)	Simultan BMA (°C)
T _{MAKS}	PLS	0,27	0,97	0,771	29,26	29,3	28,60
	PCR	-0,77	0,99	0,095	28,20		
	Ridge	0,31	0,97	0,134	29,23		
T _{MIN}	PLS	0,43	0,98	0,001	23,97	25	23,89
	PCR	0,67	0,97	0,000	23,76		
	Ridge	0,26	0,99	0,999	23,98		

Berdasarkan Tabel 4.13, dapat diketahui hasil estimasi parameter dan besarnya bobot untuk tiap *ensemble* pada tanggal 8 Februari 2009 di stamet Kemayoran. Untuk respon suhu maksimum diindikasikan bahwa PLS memiliki kontribusi terbesar terhadap model dengan bobot sebesar 0,771, lebih tinggi daripada bobot PCR dan Ridge yang masing-masing hanya sebesar 0,095 dan

0,134. Namun sebaliknya, untuk respon suhu minimum Ridge memberikan kontribusi terbesar terhadap model dengan bobot 0,999, jauh lebih besar daripada bobot PLS dan PCR yang dapat dikatakan hampir tidak memberikan terbesar terhadap model dengan bobot masing-masing sebesar 0,001 dan 0,000.

Dari Tabel 4.13, diindikasikan bahwa akurasi prakiraan model simultan BMA pada 8 Februari 2009 di stamet Kemayoran tidak berbeda secara signifikan dibanding prakiraan *ensemble*. Pada respon suhu maksimum maupun suhu minimum justru hasil prakiraan PLS dan Ridge lebih mendekati nilai observasi daripada prakiraan model simultan BMA. Hal ini terjadi karena kasus *underfitting* yang serentak dialami oleh ketiga model anggota *ensemble*, sehingga berdampak pada berkurangnya akurasi prakiraan model simultan BMA.

Selanjutnya, hasil prakiraan model simultan BMA dijadikan prediktor untuk pembentukan model Spatial BMA. Gambar 4.6 merupakan ilustrasi tahapan yang ditempuh Spatial BMA untuk menghasilkan prakiraan cuaca, dengan respon adalah observasi elemen cuaca yang diteliti yaitu suhu maksimum dan suhu minimum, serta prediktor berupa luaran model simultan BMA.



Gambar 4.6 Tahapan Mendapatkan Prakiraan Suhu Menggunakan Spatial BMA

Pada penelitian Luthfi (2017) telah dilakukan pengujian Moran's I untuk memeriksa apakah terdapat dependensi spasial yang signifikan dari suhu maksimum dan suhu minimum antar 8 lokasi (stamet). Sebelumnya, telah dibentuk matriks jarak antar 8 stamet, diketahui bahwa stamet citeko dapat dikatakan terisolir karena memiliki jarak yang paling jauh dari 7 stamet lainnya. Dermaga merupakan stamet yang paling dekat dengan Citeko dengan jarak 25 km. Matriks jarak antar 8 stamet disajikan dalam Tabel 4.14

Tabel 4.14 Matriks Jarak 8 Stasiun Meteorologi (dalam km)

Stamet	1	2	3	4	5	6	7	8
1	0	6,810	21,694	15,479	22,929	53,593	60,575	39,634
2	6,810	0	24,091	22,091	28,231	56,130	66,818	46,372
3	21,694	24,091	0	19,615	12,98	32,108	68,598	44,080
4	15,479	22,091	19,615	0	11,609	46,443	50,027	26,550
5	22,929	28,231	12,980	11,609	0	34,907	56,467	31,683
6	53,593	56,130	32,108	46,443	34,907	0	85,158	61,212
7	60,575	66,818	68,598	50,027	56,467	85,158	0	24,859
8	39,634	46,371	44,080	26,550	31,683	61,212	24,859	0

Ket :

1 : Kemayoran

2 : Priok

3 : Cengkareng

4 : Pondok Betung

5 : Curug

6 : Tangerang

7 : Citeko

8 : Dermaga

Supaya dapat menguji dependensi spasial, dibutuhkan matriks jarak terboboti yang hanya berisi 0 dan 1. Lokasi yang bersisian diberi nilai 1 dan yang tidak bersisian diberi nilai 0. Karena antar 8 stamet tidak ada yang bersisian, maka digunakan *cut-off distance* sebagai jarak maksimal yang masih memberikan pengaruh terhadap dinamika cuaca. Berdasarkan BMKG (2011) dan dengan asumsi ketinggian stamet yang seragam, digunakan *cut-off distance* sebesar 30 km. Selanjutnya, dihitung nilai Moran's I beserta *p-value* yang digunakan untuk melihat signifikansi dependensi. Tabel 4.15 menyajikan nilai Moran's I beserta *p-value* berdasarkan matriks jarak terboboti.

Tabel 4.15 Signifikansi Moran's I pada Observasi Cuaca di 8 Stamet

Elemen cuaca	Moran's I	st.dev	p-value
T_{MAKS}	0,135	0,141	0,048
T_{MIN}	0,379	0,190	0,006

Nilai Moran's I dan *p-value* pada Tabel 4.15 mengindikasikan bahwa terdapat dependensi spasial untuk suhu maksimum dan suhu minimum yang signifikan pada tingkat signifikansi $\alpha = 0,05$. Nilai Moran's I yang positif menunjukkan bahwa suhu udara pada lokasi yang berdekatan cenderung memiliki hubungan yang lebih tinggi daripada lokasi yang berjauhan. Hubungan spasial suhu minimum lebih tinggi daripada suhu maksimum, ditunjukkan dengan nilai Moran's I suhu minimum lebih tinggi daripada suhu maksimum yaitu $0,379 > 0,135$.

Setelah dibuktikan adanya dependensi spasial antar 8 stamet, selanjutnya dapat diterapkan tahap pertama dalam pemodelan Spatial BMA, yaitu mendapatkan

estimasi parameter β_0 dan β_1 dan residual untuk membentuk semivariogram empiris. Model Spatial BMA yang berlaku dalam kasus ini adalah

$$\mathbf{y}_t = \beta_{0,m} \mathbf{1} + \beta_{1,m} \mathbf{F}_m + \mathbf{E}_{1m} + \mathbf{E}_{2m}$$

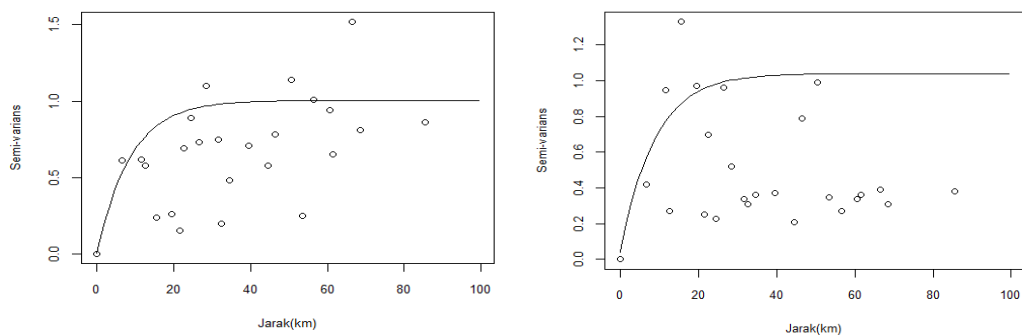
dimana \mathbf{F}_m merupakan luaran BMA berupa prakiraan suhu udara maksimum atau minimum pada hari ke- t dan lokasi ke- s dan pada anggota *ensemble* m yang terpilih. Pada model Spatial ini, digunakan panjang training 30 hari. Estimasi parameter $\beta_{0,m}$ dan $\beta_{1,m}$ disajikan dalam Tabel 4.16.

Tabel 4.16 Estimasi Parameter Model Spatial BMA

Elemen cuaca	$\hat{\beta}_0$	$\hat{\beta}_1$	SE($\hat{\beta}_0$)	SE($\hat{\beta}_1$)
T _{MAKS}	0,542	0,990	0,741	0,025
T _{MIN}	0,348	0,979	0,564	0,024

Berdasarkan Tabel 4.16, dapat diketahui bahwa parameter β_0 baik pada elemen cuaca T_{MAKS} maupun T_{MIN} tidak signifikan. Sedangkan parameter β_1 signifikan baik pada elemen cuaca T_{MAKS} maupun T_{MIN}. Walaupun terdapat parameter yang tidak signifikan, namun hal ini tidak berdampak pada *coverage* dan lebar interval prediktif model Spatial BMA.

Selanjutnya, dibentuk semivariogram empiris model eksponensial parameter spasial ρ^2 , τ^2 dan r yang diestimasi berdasarkan metode iteratif L-BFGS implementasi dari persamaan (4.16) hingga persamaan (4.18). Semivariogram dibentuk dari residual model Spatial BMA berdasarkan estimasi β_0 dan β_1 dari Tabel 4.15. Gambar 4.7 merupakan semivariogram yang terbentuk.



Gambar 4.7 Semivariogram Empiris Suhu Udara (kiri) suhu maksimum (kanan) suhu minimum

Pada Gambar 4.7, nilai semivariogram konstan setelah jarak r sebesar 8,69 km atau lebih, dengan nilai *sill* (*nugget* + *partial sill*) 1,005 untuk suhu maksimum dan 1,041 untuk suhu minimum. Cressie (1993) menyatakan bahwa *sill* yang makin besar dapat menyebabkan varians estimasi menjadi semakin besar. Hal ini berarti bahwa terdapat kemungkinan presisi prakiraan suhu maksimum lebih tinggi daripada suhu minimum sehingga berdampak pada interval model Spatial BMA yang dapat dievaluasi dengan CRPS.

Model Spatial BMA menjalankan proses simulasi dengan memodifikasi residual untuk mendapatkan prakiraan cuaca terkalibrasi. Proses tersebut dijalankan untuk mendapatkan 99 realisasi anggota *ensemble*. Alasan penggunaan 99 realisasi menurut Gel *et al.* (2004) adalah untuk mengakomodir nilai observasi bilamana Spatial BMA dikatakan terkalibrasi saat observasi suhu udara ke- t berada di antara kedua persentil. Tabel 4.17 menyajikan RMSE dan interval prediktif model Spatial BMA pada tanggal 2 Maret 2009. Hasil Prakiraan Spatial BMA secara ringkas disajikan dalam Lampiran 34-35.

Tabel 4.17 RMSE Prakiraan Spatial BMA dan Simultan BMA, 2 Maret 2009

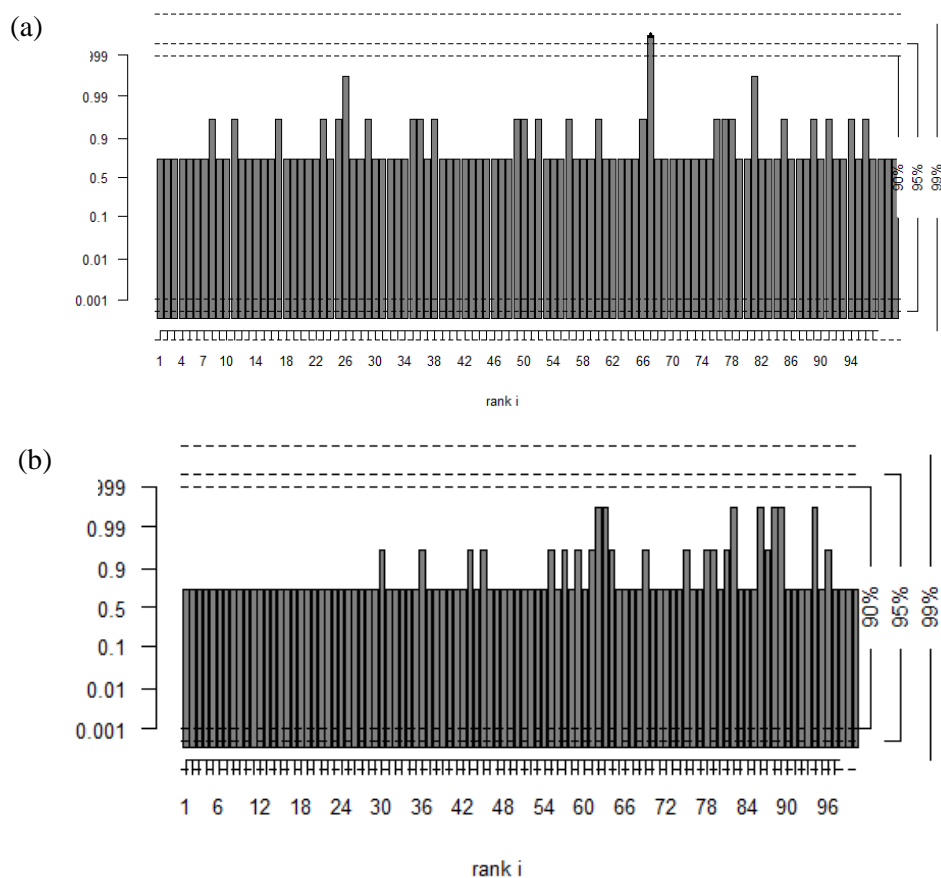
Suhu Udara	Stamet	Obs. (°C)	NWP (°C)	Simultan BMA (°C)	Spatial BMA (°C)	RMSE Simultan BMA	RMSE Spatial BMA	RMSE NWP
T_{MAKS}	Kemayoran	33	29,87	31,56	31,38	1,027°	1,399°	2,180°*
	Priok	32,4	29,58	31,53	31,18			
	Cengkareng	31,9	29,77	31,22	31,04			
	Pd. Betung	34	29,59	32,80	32,57			
	Curug	33,2	29,23	31,60	31,41			
	Tangerang	33	29,67	32,28	32,13			
	Citeko	26	29,65	25,19	24,84			
	Dermaga	31,8	29,69	30,02	29,76			
T_{MIN}	Kemayoran	25	24,18	24,77	24,12	0,811°	1,256°	1,570°*
	Priok	26	24,93	24,74	24,22			
	Cengkareng	24	24,35	23,86	23,67			
	Pd. Betung	23,8	24,14	23,54	23,13			
	Curug	23,6	23,92	23,19	22,51			
	Tangerang	25,4	24,28	23,79	23,24			
	Citeko	19,2	23,64	18,92	18,27			
	Dermaga	22,6	23,83	22,59	22,06			

*)Luthfi (2017)

Berdasarkan Tabel 4.17, dapat dikatakan bahwa Spatial BMA belum cukup mampu mengoreksi bias prakiraan *ensemble* yang dipilih berdasarkan bobot BMA. Hal ini ditunjukkan dengan nilai RMSE prakiraan Spatial BMA yang lebih tinggi daripada prakiraan *ensemble* BMA baik pada parameter suhu maksimum maupun

suhu minimum. Namun, jika dibandingkan dengan prakiraan NWP, Spatial BMA sudah mampu mengoreksi bias prakiraan suhu maksimum dan suhu minimum dengan nilai RMSE yang lebih rendah.

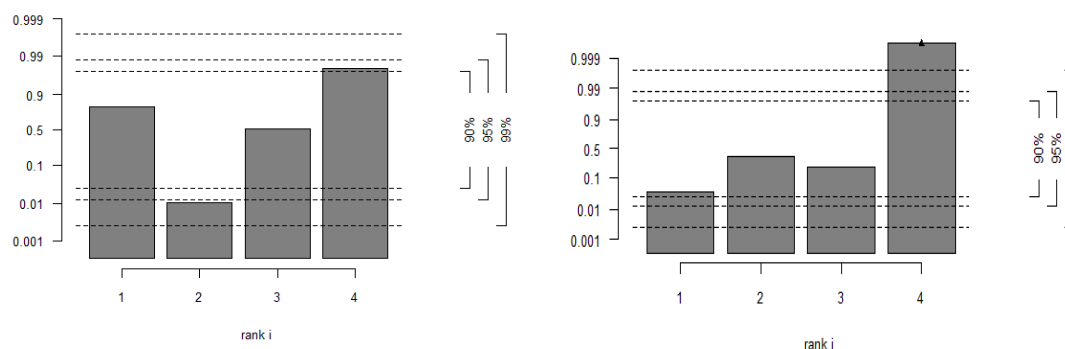
Selain dengan RMSE sebagai indikator keakuratan hasil prakiraan, digunakan nilai *coverage* sebagai indikator terkalibrasinya prakiraan. Hasil prakiraan dikatakan terkalibrasi jika nilai *coverage* mendekati nilai standar *coverage*. Standar *coverage* merupakan persentase perbandingan antara jumlah *ensemble* dikurangi 1 dengan jumlah *ensemble* dikurangi 1. Pada kasus ini, jumlah *ensemble* adalah sebanyak 99, sesuai dengan jumlah simulasi realisasi *ensemble* pada model Spatial BMA. Sehingga standar *coverage* yang digunakan adalah sebesar 98%. Indikator *coverage* ini dapat divisualisasikan dengan *Verification Rank Histogram* (VRH). Gambar 4.8 merupakan VRH prakiraan suhu maksimum dan suhu minimum menggunakan Spatial BMA.



Gambar 4.8 VRH Prakiraan Spatial BMA (a) suhu maksimum (b) suhu minimum, jumlah *ensemble*=99

Berdasarkan Gambar 4.8, diketahui bahwa *coverage* prakiraan Spatial BMA untuk T_{MAKS} dan T_{MIN} masing-masing 87,94% dan 95,16%, yang merupakan persentase banyaknya observasi yang berada di rank ke-2 sampai rank ke-98. Nilai *coverage* parameter T_{MAKS} dan T_{MIN} sudah mendekati standar *coverage* 98%. Hal ini mengindikasikan bahwa Spatial BMA cukup mampu mengkalibrasi ramalan baik parameter suhu udara maksimum maupun suhu udara minimum. Nilai *coverage* prakiraan T_{MIN} lebih mendekati standar *coverage* daripada nilai *coverage* T_{MAKS} . Artinya, Spatial BMA lebih mampu mengkalibrasi prakiraan suhu minimum daripada suhu maksimum. Hal ini juga ditunjukkan oleh Gambar 4.8, dimana VRH suhu minimum terlihat lebih seragam daripada VRH suhu maksimum.

Selanjutnya, agar dapat membandingkan langsung dengan hasil prakiraan sebelum dikalibrasi, digunakan jumlah simulasi realisasi *ensemble* sebanyak 3. Sesuai dengan jumlah anggota *ensemble* sebelum dikalibrasi. Dengan demikian, nilai standar *coverage* yang digunakan adalah 50%. Jika *coverage* prakiraan mendekati 50%, maka dapat dikatakan bahwa prakiraan tersebut telah terkalibrasi. Gambar 4.9 merupakan *Verification Rank Histogram* (VRH) dari hasil prakiraan T_{MAKS} dan T_{MIN} menggunakan Spatial BMA dengan jumlah *ensemble* sebanyak 3.



Gambar 4.9 VRH Prakiraan Spatial BMA (kiri) suhu maksimum (kanan) suhu minimum, jumlah *ensemble*=3

Gambar 4.9 menunjukkan bahwa *coverage* prakiraan Spatial BMA untuk parameter T_{MAKS} dan T_{MIN} masing-masing sebesar 48,63% dan 52,35%, yang merupakan persentase banyaknya observasi yang masuk dalam rank 2 dan rank 3. Nilai tersebut telah mendekati standar *coverage* sebesar 50%. Hal ini mengindikasikan bahwa Spatial BMA dengan jumlah *ensemble* sebanyak 3 telah cukup mampu mengkalibrasi hasil prakiraan baik suhu maksimum maupun

minimum. Indikasi ini diperkuat dengan nilai *coverage* prakiraan sebelum dikalibrasi yang masih jauh dari standar yaitu 19,23% dan 9,57% untuk suhu maksimum dan suhu minimum.

Berdasarkan kriteria RMSE dan standar *coverage*, dapat dikatakan bahwa metode Spatial BMA cukup mampu meningkatkan ketajaman prakiraan *ensemble* dengan membuat nilai observasi lebih banyak masuk dalam *range ensemble*. Namun dari sisi keakuratan, Spatial BMA belum cukup mampu mengoreksi bias prakiraan, sehingga hasil prakiraan menjadi kurang akurat dibanding dengan prakiraan *ensemble* sebelum dikalibrasi. Hal ini diperkuat dengan nilai CRPS model Spatial BMA yang lebih tinggi daripada model simultan BMA. Tabel 4.18 menunjukkan perbandingan nilai CRPS antara model Spatial BMA dan model simultan BMA (belum dikalibrasi secara spasial).

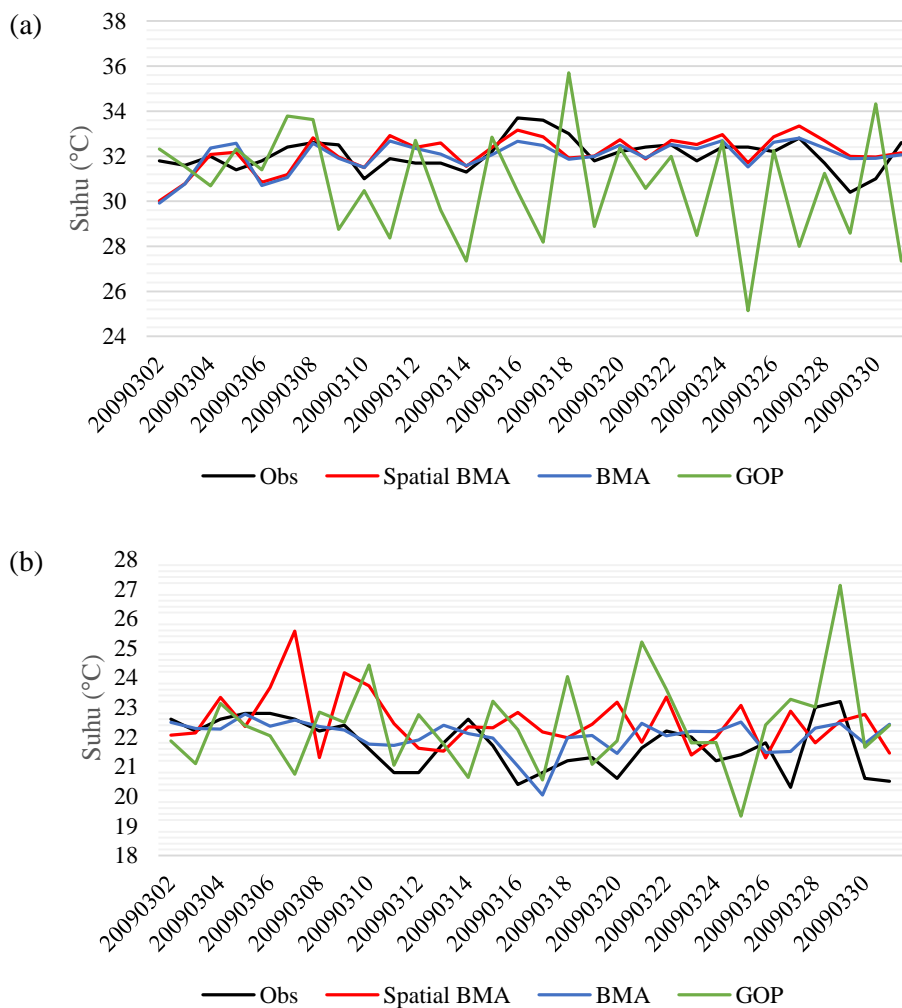
Tabel 4.18 CRPS Prakiraan Spatial BMA dan Simultan BMA

Respon	CRPS Simultan BMA	CRPS Spatial BMA	
		99 Ensemble	3 Ensemble
T_{MAKS}	0,562	0,580	0,763
T_{MIN}	0,440	0,468	0,662

Nilai CRPS dapat memverifikasi model prakiraan dari segi tingkat koreksi bias dan ketajaman yang tercapai. Ketajaman mengindikasikan bahwa PDF prediktif model terkalibrasi mampu memperbaiki PDF prediktif model sebelum kalibrasi, yaitu interval PDF model kalibrasi lebih sempit daripada sebelum dikalibrasi. Berdasarkan Tabel 4.18, nilai CRPS model Simultan BMA baik untuk suhu maksimum maupun suhu minimum lebih kecil daripada CRPS model Spatial BMA. Hal ini mengindikasikan bahwa model Spatial BMA belum cukup mampu mengoreksi bias prakiraan dan memperbaiki PDF prediktif model Simultan BMA. Tabel 4.18 juga menunjukkan bahwa nilai CRPS model Spatial BMA dengan 99 *ensemble* lebih kecil daripada CRPS model Spatial BMA dengan hanya 3 *ensemble*. Artinya, untuk model Spatial BMA lebih baik menggunakan jumlah simulasi realisasi *ensemble* yang besar agar dapat mengkalibrasi prakiraan secara optimal.

4.2.6 Evaluasi Prakiraan Spatial BMA dengan BMA dan GOP

Setelah mengetahui kebaikan prakiraan yang dihasilkan oleh Spatial BMA, langkah selanjutnya adalah membandingkan kebaikan prakiraan Spatial BMA dengan prakiraan BMA dan GOP yang diperoleh pada penelitian Luthfi (2017). Untuk melihat seberapa baik ketiga metode tersebut dalam menangkap pola cuaca yang terjadi di stamet tertentu, misalnya stamet Dermaga. Gambar 4.10 menunjukkan pola prakiraan suhu maksimum dan suhu minimum pada stamet Dermaga tanggal 2 Maret 2009 sampai 31 Maret 2009.



Gambar 4.10 Tren Prakiraan Spatial BMA, BMA, GOP dan Observasi Suhu Udara Stamet Dermaga (a) suhu maksimum (b) suhu minimum

Gambar 4.10 mengindikasikan bahwa ketiga model (Spatial BMA, BMA dan GOP) mampu untuk mengikuti pola dan tren suhu maksimum, walaupun dari segi akurasi model GOP kemungkinan besar jauh lebih rendah daripada BMA dan

Spatial BMA. Pada sebagian besar hari pengamatan, tampak prakiraan BMA dan Spatial BMA lebih mendekati nilai observasi. Sementara pada Gambar 4.10 (b), ketiga metode ini tampak belum cukup akurat dalam melakukan prakiraan suhu minimum di Stamet Dermaga. Pergerakan pola suhu minimum yang dinamis masih belum mampu diikuti oleh ketiga metode ini. Agar memperkuat analisis deskripsi, maka perlu dilakukan perbandingan dengan kriteria atau indikator yang lebih objektif.

Supaya BMA dapat dibandingkan dengan GOP dan Spatial BMA secara langsung, maka dapat dilakukan agregasi RMSE dan CRPS dari seluruh stamet untuk BMA (Luthfi, 2017). Perbandingan indikator kebaikan NWP, BMA, GOP, dan Spatial BMA disajikan dalam Tabel 4.19.

Tabel 4.19 Evaluasi Prakiraan NWP, *Raw Ensemble*, GOP, dan Spatial BMA

Indikator	Prakiraan	T _{MAKS}	T _{MIN}
RMSE (°C)	NWP	2,745*	2,010*
	<i>Mean ensemble</i>	1,058*	0,805*
	BMA	1,053*	0,819*
	GOP	3,070*	2,670*
	Spatial BMA	1,399	1,256
CRPS	<i>Raw ensemble</i>	0,726*	0,577*
	BMA	0,576*	0,451*
	GOP	1,540*	1,430*
	Spatial BMA	0,580	0,468

*) Luthfi (2017)

Tabel 4.19 menunjukkan bahwa suhu maksimum lebih akurat diramalkan dengan BMA, sedangkan suhu minimum lebih akurat diramalkan dengan *mean ensemble* yang merupakan rata-rata dari prakiraan PLS, PCR, dan Ridge. Berdasarkan kriteria CRPS terendah, BMA mampu menghasilkan prakiraan suhu maksimum dan suhu minimum yang lebih reliabel dibanding model lainnya. Jadi, dalam kasus ini, model Spatial BMA masih belum mampu mengungguli hasil prakiraan BMA dari segi keakuratan dan ketajaman. Namun, jika dibandingkan dengan GOP, hasil prakiraan Spatial BMA dapat dikatakan lebih akurat dan reliabel. Hal ini kemungkinan dapat terjadi dikarenakan oleh jumlah stasiun terlalu sedikit, karakter data yang kurang tepat dianalisis dengan metode spasial, dan sebagainya.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan analisis yang telah dilakukan dan dibahas sebelumnya pada Bab IV, diperoleh kesimpulan sebagai berikut.

1. Pada model Spatial BMA, penaksiran parameter regresi β_m untuk masing-masing anggota *ensemble* m dilakukan dengan metode *Maximum Likelihood Estimation* (MLE) distribusi normal multivariat. Persamaan penaksir parameter β_m adalah $\hat{\beta}_m = (\mathbf{F}_m^T \mathbf{F}_m)^{-1} \mathbf{F}_m^T \mathbf{y}$.

Penaksiran pembobot w_m dan varians σ^2 tidak dapat diselesaikan secara analitik menggunakan MLE, sehingga digunakan pendekatan iteratif *Expectation-Maximization* (EM) untuk mencari solusinya. Berikut ini adalah ringkasan persamaan bobot dan varians yang digunakan untuk mengkalibrasi prakiraan anggota *ensemble* dengan Spatial BMA.

$$w_m^{(i)} = \frac{1}{T} \sum_{t=1}^T z_{mt}^{(i)}, m = 1, 2, \dots, M$$

$$\sigma^{2(i)} = \frac{\sum_{t=1}^T \sum_{m=1}^M z_{mt}^{(i)} (y_t - f_{mt})^2}{T}$$

Penaksir parameter kovarians spasial ρ_m^2, τ_m^2 , dan r_m^2 yang digunakan untuk memodifikasi residual guna mendapatkan prakiraan cuaca tidak dapat diestimasi secara analitik akibat derivatif fungsi objektif

$$g(\rho_m^2, \tau_m^2, r_m^2) = \sum_{l=1}^k n_l \left(\frac{\hat{\gamma}(d_l)}{\rho_m^2 + \tau_m^2 (1 - e^{-d_l/r_m})} - 1 \right)^2 \text{ yang tidak } \textit{close-form} \text{ sehingga}$$

diperlukan pendekatan iteratif *Limited-Memory BFGS* (L-BFGS).

2. Berdasarkan kriteria RMSE dan standar coverage, dapat dikatakan bahwa metode Spatial BMA cukup mampu meningkatkan ketajaman prakiraan *ensemble* dengan membuat nilai observasi lebih banyak masuk dalam *range ensemble*. Coverage prakiraan Spatial BMA untuk suhu maksimum dan suhu

minimum meningkat menjadi lebih mendekati standar *coverage* 50%, yaitu 48,63% dan 52,35% dari yang sebelumnya hanya 19,23% dan 9,57%. Namun dari sisi keakuratan, Spatial BMA belum cukup mampu mengoreksi bias prakiraan, sehingga hasil prakiraan menjadi kurang akurat dibanding dengan prakiraan *ensemble* sebelum dikalibrasi. Hal ini diperkuat dengan nilai CRPS model Spatial BMA yang lebih tinggi daripada model simultan BMA.

3. Dalam kasus ini, suhu maksimum lebih akurat diramalkan dengan BMA, sedangkan suhu minimum lebih akurat diramalkan dengan *mean ensemble* (rata-rata dari prakiraan PLS, PCR, dan Ridge). Selain itu, BMA mampu menghasilkan prakiraan suhu maksimum dan suhu minimum yang lebih reliabel dibanding model lainnya karena memiliki nilai CRPS terendah yaitu sebesar 0,576 dan 0,451. Sehingga dalam kasus ini, Spatial BMA masih belum mampu mengungguli BMA dari segi keakuratan dan ketajaman. Namun, jika dibandingkan dengan GOP, hasil prakiraan Spatial BMA dapat dikatakan lebih akurat dan reliabel dengan nilai RMSE dan CRPS yang lebih rendah.

5.2 Saran

Berikut ini adalah saran yang dapat diberikan untuk penelitian selanjutnya, berdasarkan hasil analisis yang telah dilakukan pada penelitian ini.

1. Untuk mengetahui kriteria jenis, pola, dan karakteristik data yang sesuai dianalisis menggunakan metode Spatial BMA, perlu dilakukan simulasi dengan membangkitkan data yang diatur sedemikian rupa karakteristiknya.
2. Untuk melakukan analisis metode Spatial BMA dapat digunakan beberapa luaran NWP sebagai *ensemble* sehingga tidak memerlukan *preprocessing* yang memakan banyak waktu untuk komputasi.
3. Lokasi pengamatan yang dianalisis menggunakan Spatial BMA sebaiknya ditambah jumlahnya dan terdapat autokorelasi spasial yang tinggi.

DAFTAR PUSTAKA

- Anggraeni, D. (2013). Kalibrasi Peramalan Ensemble Data Curah Hujan Dengan Metode Ensemble Model Output Statistics (EMOS) dan Bayesian Model Averaging (BMA)". *Tesis*, Insitut Teknologi Sepuluh Nopember, Surabaya.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publisher.
- Arfianto, A. D. (2008). "Aplikasi Model Regresi Logistik untuk Perkiraan Kejadian Hujan". *Skripsi*, Institut Pertanian Bogor.
- Berrocal, V.J, Raftery, A.E., dan Gneiting, T. (2007). Combining Spatial Statistical and Ensemble Information in Probabilistic Weather Forecast. *Monthly Weather Review AMS*, Vol. 135, pp.1386-1402.
- Cliff, A.D. dan Ord, J.K. (1981). *Spatial Processes: Models & Applications*. London: Pion Limited.
- Cressie, N. (1985). Fitting Variogram Models by Weighted Least Squares. *Mathematical Geology*, Vol. 17, No. 5, pp. 563-585.
- Cressie, N. (1993). *Statistics for Spatial Data Revised Edition*. New Jersey: John Wiley and Sons, Inc.
- Dempster, A.P., Laird, N.M., dan Rubin, D.B. 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society*, Vol. 39, No. 1, pp. 1-38. College Article: Oxford, United Kingdom
- Devin, K., Clarke, L., Roulston, M., Ziehmman, C., Brocker, J. dan Smith, L. (2004). *Generalizing Model Output Statistics for Full Ensemble Forecasts*. Pembroke College Article: Oxford, United Kingdom.
- Draper, N.R. dan Smith, H. (1992). *Applied Regression Analysis Second Edition*. New York: John Wiley and Sons, Inc.
- Feldmann, K. (2012). "Statistical Postprocessing of Ensemble Forecasts for Temperature: The Importance of Spatial Modeling". *Tesis Diplomarbeit*, Ruperto-Carola University of Heidelberg, Germany.

- Feldmann, K., Scheuerer, M. dan Thorarinsdottir, T. L. (2015). Spatial Postprocessing of Ensemble Forecasts for Temperature Using Nonhomogeneous Gaussian Regression. *American Meteorological Society*, Vol. 143, pp. 955-971.
- Gel, Y., Raftery, A.E., dan Gneiting, T. (2004). Calibrated probabilistic mesoscale weather field forecasting : The Geostatistical Output Perturbation (GOP) method (with discussion). *Journal of the American Statistical Association* , Vol. 99 (467), pp. 575-583.
- Gneiting, T., Raftery, A.E, Westveld III, A.H., dan Goldman, T. (2005). Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *Monthly Weather Reviews AMS*, Vol. 133, pp. 1098-1118.
- Hammill, T. (2001). Interpretation of Rank Histograms for Verifying Ensemble Forecast. *Monthly Weather Review AMS*, hal. Vol. 129, pp. 550-560.
- Idowu, O. d. (2009). *Model Output Statistics to Improve Severe Storm Prediction Over Western Sahel*. South Africa: Departement of Geography, Geoinformatics and Meteorology, University of Pretoria.
- Johnson, R.A. dan Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis 5th Edition*. New Jersey: Prentice Hall.
- Jolliffe, I.T. (2002). *Principal Component Analysis, Second Edition*. New York: Springer-Verlag.
- Lange, K. 2010. *Numerical Analysis for Statisticians, Second Edition*. New York: Springer-Verlag.
- Linacre, E. dan Geerts, B. (2003). *Climate and Weather Explained*. New York: Routledge.
- Liu, D.C. dan Nocedal, J. 1989. On The Limited Memory BFGS Method for Large Scale Optimization. *Mathematical Programming*, Vol. 45, pp. 503-528
- Luthfi, M. (2017). Bayesian Model Averaging dan Geostatistical Output Perturbation untuk Prakiraan Cuaca Jangka Pendek Terkalibrasi. *Tesis*, Insitut Teknologi Sepuluh Nopember, Surabaya.

- McLachlan, G.J dan Krishnan, T. 2008. *The EM Algorithm and Extensions, Second Edition*. New Jersey: John Wiley & Sons, Inc.
- Möller, A. (2014). Multivariate and spatial ensemble postprocessing methods. *Disertasi*, Ruperto-Carola University of Heidelberg, Germany.
- Quinn, T. (1990). *Temperature :Second Edition*. Massachusetts: Academic Press.
- Raftery, A.E. dan Zheng, Y. (2003). Performance of Bayesian Model Averaging. *Journal of the American Statistical Association*, Vol. 98, pp. 931-938.
- Raftery, A.E., Gneiting, T., Balabdoui, F. dan Polakowski, M. (2005). Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Monthly Weather Review AMS*, Vol. 133, pp.1155-1174.
- Rencher, A.C. dan Schaalje, G.B. (2008). *Linear Models in Statistics Second Edition*. New Jersey: John Wiley and Sons, Inc.
- Sari, I. (2013). Pendekatan Regresi Partial Least Square Univariate dan Multivariate Response untuk Memprediksi Suhu dan Kelembapan. *Tesis*, Surabaya : Institut Teknologi Sepuluh Nopember.
- Slughter, J.M., Raftery, A.E., Gneiting Tilmann, dan Fraley, Chris. (2007). *Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging*. Washington: American Meteorological Society
- Schmeits, M.J dan Kok, K.J. (2010). A Comparison between Raw Ensemble Output, (Modified) Bayesian Model Averaging dan Extended Logistic Regression Using ECMWF Ensemble Precipitation Forecast. *Monthly Weather Review AMS*, Vol. 138, pp. 4199-4211.
- Tanudidjaja. (1993). *Ilmu Pengetahuan Bumi dan Antariksa*. Jakarta: Penerbit Departemen Pendidikan dan Kebudayaan.
- Wilks, D. (2006). *Statistical Methods in the Atmospheric Sciences Second Edition*. Boston: Elsevier.
- Wold, S., Sjöström, M., dan Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, Vol. 58, pp. 109-130.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Cengkareng

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,999	99,98%	temp2	1	8,609	95,65%
mixr1	1	7,775	86,39%	temp4	1	8,903	98,92%
mixr2	1	8,577	95,30%	tmaxscr	1	8,206	91,18%
mixr4	1	8,718	96,86%	tminscr	1	7,814	86,83%
omega1	1	8,854	98,37%	tpan	2	7,764; 1,092	98,40%
omega2	1	8,159	90,66%	tscrn	1	7,970	88,55%
omega4	1	7,880	87,56%	u1	1	8,582	95,35%
pblh	2	7,147; 1,378	94,73%	u2	1	8,823	98,03%
ps	1	8,998	99,98%	u4	1	8,949	99,44%
psl	1	8,998	99,98%	ustar	1	7,955	88,39%
qgscrn	1	7,749	86,10%	v1	1	8,029	89,21%
rh1	2	7,018; 1,434	93,91%	v2	1	8,456	93,96%
rh2	1	8,451	93,90%	v4	1	8,874	98,60%
rh4	1	8,687	96,52%	zg1	3	3,084; 2,271; 1,149	72,27%
rnd	1	7,350	81,66%	zg2	1	6,959	77,32%
temp1	1	7,946	88,29%	zg4	2	8,789	97,65%

Lampiran 2. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Citeko

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,998	99,98%	temp2	1	8,598	95,54%
mixr1	1	8,045	89,39%	temp4	1	8,864	98,48%
mixr2	1	8,509	95,54%	tmaxscr	1	8,814	97,93%
mixr4	1	8,692	96,57%	tminscr	1	8,217	91,30%
omega1	2	6,839; 1,138	98,37%	tpan	1	8,536	94,85%
omega2	2	6,658; 1,026	85,26%	tscrn	1	8,388	93,21%
omega4	1	7,242	80,46%	u1	1	8,357	92,86%
pblh	1	8,057	89,52%	u2	1	8,581	95,34%

Lampiran 2. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Citeko (Lanjutan)

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
ps	1	8,953	99,48%	u4	1	8,916	99,07%
psl	1	8,997	99,97%	ustar	2	6,797; 1,134	88,13%
qgscrn	1	7,498	83,31%	v1	2	6,817; 1,265	89,79%
rh1	1	7,783	86,47%	v2	2	6,820; 1,405	91,38%
rh2	1	8,416	93,51%	v4	1	8,658	96,20%
rh4	1	8,678	96,42%	zg1	1	7,658	85,08%
rnd	1	7,920	88,00%	zg2	2	6,602; 2,091	96,59%
temp1	1	8,367	92,96%	zg4	2	5,999; 2,893	98,79%

Lampiran 3. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Curug

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,998	99,98%	temp2	1	8,752	97,25%
mixr1	1	8,235	91,50%	temp4	1	8,907	98,96%
mixr2	1	8,696	96,62%	tmaxscr	1	8,840	98,22%
mixr4	1	8,764	97,38%	tminscr	1	8,332	92,58%
omega1	1	8,643	96,04%	tpan	1	8,753	97,25%
omega2	1	8,301	92,23%	tscrn	1	8,493	94,37%
omega4	1	8,008	88,98%	u1	1	8,648	96,09%
pblh	1	8,318	92,43%	u2	1	8,771	97,46%
ps	1	8,968	99,64%	u4	1	8,944	99,38%
psl	1	8,998	99,98%	ustar	1	8,092	89,92%
qgscrn	1	8,168	90,76%	v1	1	8,351	92,79%
rh1	1	7,972	88,58%	v2	1	8,347	92,74%
rh2	1	8,626	95,84%	v4	1	8,834	98,15%
rh4	1	8,741	97,13%	zg1	1	7,340	81,55%
rnd	1	8,003	88,92%	zg2	2	5,544; 2,997	94,91%
temp1	1	8,515	94,61%	zg4	2	6,823; 2,086	98,99%

Lampiran 4. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Kemayoran

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,998	99,98%	temp2	1	8,726	96,96%
mixr1	1	8,234	91,48%	temp4	1	8,910	99,00%
mixr2	1	8,670	96,34%	tmaxscr	1	8,754	97,26%
mixr4	1	8,753	97,25%	tminscr	1	8,315	92,39%
omega1	1	8,863	98,48%	tpan	1	8,617	95,75%
omega2	1	8,291	92,12%	tscrn	1	8,613	95,70%
omega4	1	8,083	89,82%	u1	1	8,758	97,31%
pblh	1	8,197	91,08%	u2	1	8,839	98,21%
ps	1	8,997	99,97%	u4	1	8,945	99,39%
psl	1	8,998	99,98%	ustar	1	8,182	90,91%
qgscrn	1	8,233	91,48%	v1	1	8,309	92,32%
rh1	1	7,984	88,71%	v2	1	8,330	92,55%
rh2	1	8,593	95,48%	v4	1	8,841	98,24%
rh4	1	8,736	97,07%	zg1	3	3,882; 2,051; 1,061	77,71%
rnd	1	7,682	85,36%	zg2	2	6,580; 1,255	87,06%
temp1	1	8,619	95,76%	zg4	1	8,740	97,12%

Lampiran 5. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Pondok Betung

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,998	99,98%	temp2	1	8,753	97,25%
mixr1	1	8,317	92,42%	temp4	1	8,909	98,99%
mixr2	1	8,667	96,30%	tmaxscr	1	8,866	98,51%
mixr4	1	8,766	97,40%	tminscr	1	8,516	94,62%
omega1	1	8,900	98,89%	tpan	1	8,753	97,26%
omega2	1	8,494	94,38%	tscrn	1	8,641	96,01%
omega4	1	8,061	89,57%	u1	1	8,724	96,93%
pblh	1	8,328	92,53%	u2	1	8,803	97,81%
ps	1	8,997	99,97%	u4	1	8,944	99,37%
psl	1	8,998	99,98%	ustar	1	8,177	90,86%
qgscrn	1	8,306	92,29%	v1	1	8,269	91,87%
rh1	1	8,214	91,27%	v2	1	8,210	91,22%
rh2	1	8,599	95,54%	v4	1	8,842	98,25%

Lampiran 5. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Pondok Betung (Lanjutan)

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
rh4	1	8,750	97,23%	zg1	3	4,009; 1,610; 1,176	75,51%
rnd	1	7,943	88,26%	zg2	2	6,843; 1,224	89,63%
temp1	1	8,665	96,28%	zg4	1	8,767	97,41%

Lampiran 6. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Priok

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,998	99,98%	temp2	1	8,543	94,92%
mixr1	1	7,681	85,34%	temp4	1	8,899	98,87%
mixr2	1	8,568	95,20%	tmaxscr	1	8,093	89,93%
mixr4	1	8,715	96,84%	tminscr	1	7,529	83,66%
omega1	1	8,746	97,18%	tpan	2	7,624; 1,116	97,13%
omega2	1	7,840	87,11%	tscrn	1	7,787	86,53%
omega4	1	7,771	86,34%	u1	1	8,614	95,71%
pblh	2	7,052; 1,451	94,48%	u2	1	8,822	98,02%
ps	1	8,997	99,97%	u4	1	8,948	99,42%
psl	1	8,998	99,98%	ustar	1	7,878	87,53%
qgscrn	1	7,744	86,04%	v1	1	7,907	87,86%
rh1	2	6,822; 1,601	93,59%	v2	1	8,330	92,56%
rh2	1	8,420	93,55%	v4	1	8,865	98,50%
rh4	1	8,692	96,57%	zg1	4	2,975; 2,625; 1,253; 1,017	87,44%
rnd	2	7,085; 1,213	92,20%	zg2	1	6,672	74,14%
temp1	1	7,772	86,36%	zg4	1	8,752	97,25%

Lampiran 7. Banyak PC, Nilai Eigen, dan Kumulatif Varians NWP di Stamet Tangerang

Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians	Variabel	Banyak PC	Nilai Eigen	Kumulatif Varians
dpsdt	1	8,998	99,98%	temp2	1	8,760	97,33%
mixr1	1	8,282	92,02%	temp4	1	8,914	99,04%
mixr2	1	8,655	96,17%	tmaxscr	1	8,833	98,14%
mixr4	1	8,756	97,29%	tminscr	1	8,441	93,78%
omega1	1	8,874	98,61%	tpan	1	8,728	96,98%
omega2	1	8,502	94,46%	tscrn	1	8,623	95,81%
omega4	1	8,090	89,89%	u1	1	8,754	97,27%
pblh	1	8,225	91,39%	u2	1	8,824	98,05%
ps	1	8,998	99,98%	u4	1	8,948	99,42%
psl	1	8,998	99,98%	ustar	1	8,159	90,66%
qgscrn	1	8,241	91,56%	v1	1	8,416	93,52%
rh1	1	8,106	90,07%	v2	1	8,469	94,10%
rh2	1	8,594	95,49%	v4	1	8,866	98,51%
rh4	1	8,736	97,06%	zg1	3	3,686; 2,028; 1,349	78,47%
rnd	1	7,856	87,29%	zg2	2	7,049; 1,015	89,60%
temp1	1	8,632	95,91%	zg4	1	8,791	97,68%

Lampiran 8. Nilai PRESS untuk Stamet Cengkareng

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
	T_{MAKS}	1	0,7969	14	0,7021	27
	2	0,7468	15	0,7028	28	0,7056
	3	0,7211	16	0,7031	29	0,7068
	4	0,7123	17	0,7031	30	0,7076
	5	0,7037	18	0,7027	31	0,7075
	6	0,7013	19	0,7035	32	0,7078
	7	0,7001	20	0,7034	33	0,7079
	8	0,6990	21	0,7035	34	0,7079
	9	0,6993	22	0,7041	35	0,7078
	10	0,6999	23	0,7043	36	0,7074
	11	0,7003	24	0,7043	37	0,7074
	12	0,7018	25	0,7043	38	0,7068
	13	0,7017	26	0,7044		

Lampiran 8. Nilai PRESS untuk Stamet Cengkareng (Lanjutan)

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MIN}	1	0,9063	14	0,9021	27	0,9122
	2	0,8981	15	0,9042	28	0,9119
	3	0,8956	16	0,9044	29	0,9122
	4	0,8949	17	0,9067	30	0,9129
	5	0,8943	18	0,9064	31	0,9131
	6	0,8937	19	0,9083	32	0,9137
	7	0,8948	20	0,9104	33	0,9142
	8	0,8973	21	0,9132	34	0,9148
	9	0,9006	22	0,9088	35	0,9149
	10	0,9057	23	0,9191	36	0,9151
	11	0,9077	24	0,9149	37	0,9152
	12	0,9067	25	0,9148	38	0,9167
	13	0,9049	26	0,9143		

Lampiran 9. Nilai PRESS untuk Stamet Citeko

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MAKS}	1	0,8135	14	0,7327	27	0,7363
	2	0,7742	15	0,7327	28	0,7352
	3	0,7487	16	0,7333	29	0,7358
	4	0,7432	17	0,7341	30	0,7358
	5	0,7369	18	0,7358	31	0,7347
	6	0,7324	19	0,7359	32	0,7360
	7	0,7311	20	0,7367	33	0,7360
	8	0,7300	21	0,7368	34	0,7360
	9	0,7299	22	0,7362	35	0,7362
	10	0,7305	23	0,7347	36	0,7362
	11	0,7313	24	0,7333	37	0,7357
	12	0,7318	25	0,7331	38	0,7352
	13	0,7326	26	0,7351	39	0,7350
T _{MIN}						

Lampiran 9. Nilai PRESS untuk Stamet Citeko (Lanjutan)

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MIN}	6	0,7990	19	0,8056	32	0,8060
	7	0,7978	20	0,8060	33	0,8055
	8	0,7977	21	0,8070	34	0,8058
	9	0,7989	22	0,8069	35	0,8055
	10	0,7999	23	0,8080	36	0,8053
	11	0,8023	24	0,8077	37	0,8054
	12	0,8040	25	0,8103	38	0,8054
	13	0,8016	26	0,8055	39	0,8059

Lampiran 10. Nilai PRESS untuk Stamet Curug

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MAKS}	1	0,8052	13	0,7064	24	0,7029
	2	0,7488	14	0,7054	25	0,7032
	3	0,7246	15	0,7045	26	0,7038
	4	0,7178	16	0,7035	27	0,7054
	5	0,7146	17	0,7029	28	0,7056
	6	0,7152	18	0,7027	29	0,7051
	7	0,7074	19	0,7013	30	0,7048
	8	0,7125	20	0,7081	31	0,7047
	9	0,7095	21	0,7030	32	0,7036
	10	0,7084	22	0,7027	33	0,7044
	11	0,7071	23	0,7028	34	0,7042
	12	0,7070				
T _{MIN}	1	0,7180	13	0,7648	24	0,7698
	2	0,7619	14	0,7651	25	0,7682
	3	0,7625	15	0,7649	26	0,7677
	4	0,7603	16	0,7652	27	0,7670
	5	0,7625	17	0,7643	28	0,7672
	6	0,7609	18	0,7643	29	0,7680
	7	0,7608	19	0,7760	30	0,7678
	8	0,7606	20	0,7676	31	0,7682
	9	0,7626	21	0,7692	32	0,7685
	10	0,7639	22	0,7667	33	0,7683
	11	0,7630	23	0,7678	34	0,7689
	12	0,7632				

Lampiran 11. Nilai PRESS untuk Stamet Kemayoran

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MAKS}	1	0,8144	13	0,7123	25	0,7158
	2	0,7603	14	0,7134	26	0,7161
	3	0,7313	15	0,7154	27	0,7153
	4	0,7225	16	0,7171	28	0,7164
	5	0,7145	17	0,7179	29	0,7159
	6	0,7119	18	0,7182	30	0,7163
	7	0,7109	19	0,7180	31	0,7166
	8	0,7100	20	0,7176	32	0,7167
	9	0,7103	21	0,7174	33	0,7168
	10	0,7106	22	0,7170	34	0,7169
	11	0,7111	23	0,7160	35	0,7174
	12	0,7117	24	0,7157		
	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MIN}	1	0,8898	13	0,8693	25	0,8790
	2	0,8723	14	0,8692	26	0,8771
	3	0,8654	15	0,8693	27	0,8752
	4	0,8647	16	0,8699	28	0,8746
	5	0,8649	17	0,8709	29	0,8742
	6	0,8648	18	0,8717	30	0,8740
	7	0,8661	19	0,8729	31	0,8734
	8	0,8657	20	0,8747	32	0,8731
	9	0,8665	21	0,8749	33	0,8739
	10	0,8682	22	0,8770	34	0,8739
	11	0,8689	23	0,8809	35	0,8741
	12	0,8695	24	0,8833		

Lampiran 12. Nilai PRESS untuk Stamet Pondok Betung

T _{MAKS}	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
	1	0,8215	13	0,7226	25	0,7170
	2	0,7640	14	0,7228	26	0,7167
	3	0,7350	15	0,7222	27	0,7167
	4	0,7293	16	0,7211	28	0,7168
	5	0,7249	17	0,7194	29	0,7168
	6	0,7245	18	0,7181	30	0,7167
	7	0,7226	19	0,7174	31	0,7165

Lampiran 12. Nilai PRESS untuk Stamet Pondok Betung (Lanjutan)						
T _{MAKS}	Jumlah	Nilai	Jumlah	Nilai	Jumlah	Nilai
	Komponen	PRESS	Komponen	PRESS	Komponen	PRESS
	8	0,7228	20	0,7156	32	0,7166
	9	0,7229	21	0,7160	33	0,7166
	10	0,7224	22	0,7154	34	0,7166
	11	0,7223	23	0,7156	35	0,7171
12	0,7224	24	0,7168			
T _{MIN}	Jumlah	Nilai	Jumlah	Nilai	Jumlah	Nilai
	Komponen	PRESS	Komponen	PRESS	Komponen	PRESS
	1	0,9273	13	0,9160	25	0,9127
	2	0,9126	14	0,9160	26	0,9135
	3	0,9095	15	0,9164	27	0,9152
	4	0,9082	16	0,9170	28	0,9161
	5	0,9079	17	0,9174	29	0,9173
	6	0,9074	18	0,9087	30	0,9157
	7	0,9074	19	0,9163	31	0,9147
	8	0,9072	20	0,9138	32	0,9150
	9	0,9077	21	0,9186	33	0,9149
	10	0,9085	22	0,9148	34	0,9149
	11	0,9109	23	0,9139	35	0,9155
12	0,9132	24	0,9115			

Lampiran 13. Nilai PRESS untuk Stamet Priok

T _{MAKS}	Jumlah	Nilai	Jumlah	Nilai	Jumlah	Nilai
	Komponen	PRESS	Komponen	PRESS	Komponen	PRESS
	1	0,8310	14	0,7583	27	0,7606
	2	0,7898	15	0,7581	28	0,7607
	3	0,7682	16	0,7586	29	0,7608
	4	0,7616	17	0,7588	30	0,7610
	5	0,7544	18	0,7591	31	0,7612
	6	0,7531	19	0,7597	32	0,7615
	7	0,7523	20	0,7601	33	0,7615
	8	0,7520	21	0,7598	34	0,7623
	9	0,7528	22	0,7594	35	0,7625
	10	0,7533	23	0,7595	36	0,7624
	11	0,7549	24	0,7599	37	0,7620
12	0,7564	25	0,7602	38	0,7620	
13	0,7580	26	0,7603	39	0,7631	

Lampiran 13. Nilai PRESS untuk Stamet Priok (Lanjutan)

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MIN}	1	0,9064	14	0,8822	27	0,8858
	2	0,8922	15	0,8834	28	0,8864
	3	0,8835	16	0,8826	29	0,8876
	4	0,8814	17	0,8826	30	0,8880
	5	0,8789	18	0,8833	31	0,8888
	6	0,8785	19	0,8859	32	0,8887
	7	0,8784	20	0,8863	33	0,8889
	8	0,8785	21	0,8861	34	0,8880
	9	0,8797	22	0,8858	35	0,8881
	10	0,8811	23	0,8857	36	0,8885
	11	0,8827	24	0,8858	37	0,8873
	12	0,8830	25	0,8854	38	0,8875
	13	0,8840	26	0,8853	39	0,8889

Lampiran 14. Nilai PRESS untuk Stamet Tangerang

	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
T _{MAKS}	1	0,8047	13	0,7090	25	0,7092
	2	0,7500	14	0,7122	26	0,7084
	3	0,7227	15	0,7157	27	0,7113
	4	0,7148	16	0,7156	28	0,7093
	5	0,7082	17	0,7127	29	0,7112
	6	0,7074	18	0,7114	30	0,7110
	7	0,7059	19	0,7088	31	0,7116
	8	0,7050	20	0,7076	32	0,7114
	9	0,7055	21	0,7074	33	0,7112
	10	0,7053	22	0,7084	34	0,7112
	11	0,7057	23	0,7099	35	0,7130
	12	0,7071	24	0,7097		
T _{MIN}	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS	Jumlah Komponen	Nilai PRESS
	1	0,9512	13	0,9593	25	0,9645
	2	0,9484	14	0,9603	26	0,9644
	3	0,9483	15	0,9599	27	0,9644
	4	0,9496	16	0,9600	28	0,9640
	5	0,9508	17	0,9599	29	0,9638
6	0,9517	18	0,9596	30	0,9638	

Lampiran 14. Nilai PRESS untuk Stamet Tangerang (Lanjutan)

T _{MIN}	Jumlah	Nilai	Jumlah	Nilai	Jumlah	Nilai
	Komponen	PRESS	Komponen	PRESS	Komponen	PRESS
	7	0,9524	19	0,9591	31	0,9650
	8	0,9533	20	0,9595	32	0,9651
	9	0,9538	21	0,9602	33	0,9653
	10	0,9541	22	0,9612	34	0,9654
	11	0,9559	23	0,9631	35	0,9670
	12	0,9572	24	0,9645		

Lampiran 15. Model Regresi PLS Stamet Dermaga

$$\begin{aligned}
 \hat{Y}_{T_{\text{MAX}}} = & -0,179 \sum_{i=1}^9 e_{i1} \text{dpsdt}_i - 0,187 \sum_{i=1}^9 e_{i2} \text{mixr1}_i + 0,193 \sum_{i=1}^9 e_{i3} \text{mixr2}_i - 0,074 \sum_{i=1}^9 e_{i4} \text{mixr4}_i \\
 & + 0,258 \sum_{i=1}^9 e_{i5} \text{omega1}_i + 0,307 \sum_{i=1}^9 e_{i6} \text{omega1}_i - 0,148 \sum_{i=1}^9 e_{i7} \text{omega2}_i - 0,019 \sum_{i=1}^9 e_{i8} \text{omega2}_i \\
 & - 0,015 \sum_{i=1}^9 e_{i9} \text{omega4}_i - 0,100 \sum_{i=1}^9 e_{i10} \text{omega4}_i - 0,010 \sum_{i=1}^9 e_{i11} \text{pblh}_i + 0,021 \sum_{i=1}^9 e_{i12} \text{ps}_i \\
 & - 0,026 \sum_{i=1}^9 e_{i13} \text{psl}_i - 0,034 \sum_{i=1}^9 e_{i14} \text{qgscrn}_i - 0,048 \sum_{i=1}^9 e_{i15} \text{qgscrn}_i - 0,058 \sum_{i=1}^9 e_{i16} \text{rh1}_i \\
 & - 0,088 \sum_{i=1}^9 e_{i17} \text{rh1}_i + 0,50 \sum_{i=1}^9 e_{i18} \text{rh2}_i + 0,049 \sum_{i=1}^9 e_{i19} \text{rh4}_i + 0,068 \sum_{i=1}^9 e_{i20} \text{rnd}_i \\
 & - 0,174 \sum_{i=1}^9 e_{i21} \text{temp1}_i + 0,471 \sum_{i=1}^9 e_{i22} \text{temp2}_i - 0,015 \sum_{i=1}^9 e_{i23} \text{temp4}_i - 0,523 \sum_{i=1}^9 e_{i24} \text{t max scr}_i \\
 & - 0,143 \sum_{i=1}^9 e_{i25} \text{t min scr}_i + 0,083 \sum_{i=1}^9 e_{i26} \text{tpan}_i - 0,181 \sum_{i=1}^9 e_{i27} \text{tscrn}_i + 0,444 \sum_{i=1}^9 e_{i28} \text{u1}_i \\
 & - 0,439 \sum_{i=1}^9 e_{i29} \text{u2}_i + 0,202 \sum_{i=1}^9 e_{i30} \text{u4}_i + 0,052 \sum_{i=1}^9 e_{i31} \text{ustar}_i + 0,024 \sum_{i=1}^9 e_{i32} \text{ustar}_i - 0,010 \sum_{i=1}^9 e_{i33} \text{v1}_i \\
 & - 0,036 \sum_{i=1}^9 e_{i34} \text{v1}_i - 0,116 \sum_{i=1}^9 e_{i35} \text{v2}_i + 0,092 \sum_{i=1}^9 e_{i36} \text{v2}_i - 0,006 \sum_{i=1}^9 e_{i37} \text{v4}_i - 0,153 \sum_{i=1}^9 e_{i38} \text{zg1}_i \\
 & + 0,044 \sum_{i=1}^9 e_{i39} \text{zg2}_i + 0,005 \sum_{i=1}^9 e_{i40} \text{zg4}_i + 0,123 \sum_{i=1}^9 e_{i41} \text{zg4}_i
 \end{aligned}$$

Lampiran 15. Model Regresi PLS Stamet Dermaga (Lanjutan)

$$\begin{aligned}
 \hat{y}_{T_{MIN}} = & 0,002 \sum_{i=1}^9 e_{i1} dpsdt_i - 0,084 \sum_{i=1}^9 e_{i2} mixr1_i - 0,093 \sum_{i=1}^9 e_{i3} mixr2_i - 0,038 \sum_{i=1}^9 e_{i4} mixr4_i \\
 & + 0,027 \sum_{i=1}^9 e_{i5} omega1_i - 0,014 \sum_{i=1}^9 e_{i6} omega1_i + 0,010 \sum_{i=1}^9 e_{i7} omega2_i - 0,004 \sum_{i=1}^9 e_{i8} omega2_i \\
 & - 0,010 \sum_{i=1}^9 e_{i9} omega4_i + 0,073 \sum_{i=1}^9 e_{i10} omega4_i + 0,015 \sum_{i=1}^9 e_{i11} pblh_i + 0,054 \sum_{i=1}^9 e_{i12} ps_i \\
 & - 0,058 \sum_{i=1}^9 e_{i13} psl_i - 0,086 \sum_{i=1}^9 e_{i14} qgscrn_i + 0,017 \sum_{i=1}^9 e_{i15} qgscrn_i + 0,043 \sum_{i=1}^9 e_{i16} rh1_i \\
 & - 0,073 \sum_{i=1}^9 e_{i17} rh1_i - 0,061 \sum_{i=1}^9 e_{i18} rh2_i + 0,014 \sum_{i=1}^9 e_{i19} rh4_i - 0,009 \sum_{i=1}^9 e_{i20} rnd_i \\
 & - 0,048 \sum_{i=1}^9 e_{i21} temp1_i - 0,041 \sum_{i=1}^9 e_{i22} temp2_i - 0,122 \sum_{i=1}^9 e_{i23} temp4_i + 0,109 \sum_{i=1}^9 e_{i24} t \max scr_i \\
 & - 0,081 \sum_{i=1}^9 e_{i25} t \min scr_i + 0,008 \sum_{i=1}^9 e_{i26} tpan_i - 0,039 \sum_{i=1}^9 e_{i27} tscrn_i - 0,014 \sum_{i=1}^9 e_{i28} u1_i \\
 & + 0,006 \sum_{i=1}^9 e_{i29} u2_i + 0,040 \sum_{i=1}^9 e_{i30} u4_i + 0,065 \sum_{i=1}^9 e_{i31} ustar_i - 0,059 \sum_{i=1}^9 e_{i32} ustar_i - 0,011 \sum_{i=1}^9 e_{i33} v1_i \\
 & - 0,036 \sum_{i=1}^9 e_{i34} v1_i - 0,008 \sum_{i=1}^9 e_{i35} v2_i + 0,018 \sum_{i=1}^9 e_{i36} v2_i - 0,026 \sum_{i=1}^9 e_{i37} v4_i + 0,022 \sum_{i=1}^9 e_{i38} zg1_i \\
 & + 0,008 \sum_{i=1}^9 e_{i39} zg2_i - 0,066 \sum_{i=1}^9 e_{i40} zg4_i + 0,070 \sum_{i=1}^9 e_{i41} zg4_i
 \end{aligned}$$

Lampiran 16. Prakiraan Suhu Maksimum Model PLS

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
1/1/2009	31,18	24,90	31,33	30,91
1/2/2009	32,68	26,98	32,87	31,63
1/3/2009	32,26	26,23	32,80	31,71
1/4/2009	32,39	26,25	32,69	32,34
1/5/2009	32,51	26,78	32,78	32,10
1/6/2009	30,85	26,15	31,20	31,65
1/7/2009	30,45	25,28	30,67	30,42
1/8/2009	30,70	24,90	30,67	30,55
1/9/2009	29,45	23,56	30,44	29,03
1/10/2009	28,60	22,80	28,54	27,98
1/11/2009	29,83	23,69	30,15	29,30
1/12/2009	28,86	22,12	28,04	27,05
⋮	⋮	⋮	⋮	⋮
12/25/2010	31,29	25,37	31,43	31,20

Lampiran 16. Prakiraan Suhu Maksimum Model PLS (Lanjutan)

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
12/26/2010	30,90	25,30	31,41	30,79
12/27/2010	30,72	23,97	30,04	29,11
12/28/2010	30,66	24,31	30,64	29,72
12/29/2010	31,11	25,67	31,01	31,14
12/30/2010	30,96	25,74	31,67	31,92
12/31/2010	31,08	25,12	30,94	30,52
Tanggal	Kemayoran (°C)	Pd, Betung (°C)	Priok (°C)	Tangerang (°C)
1/1/2009	31,48	32,20	31,02	32,06
1/2/2009	33,62	33,69	32,79	33,63
1/3/2009	33,59	33,87	32,65	33,41
1/4/2009	32,88	33,60	32,44	33,28
1/5/2009	33,27	33,75	32,60	33,62
1/6/2009	31,67	32,84	31,06	31,71
1/7/2009	31,20	32,31	30,74	31,47
1/8/2009	30,94	32,09	30,57	31,61
1/9/2009	30,03	31,41	29,83	30,27
1/10/2009	28,65	29,72	29,35	28,87
1/11/2009	29,79	30,86	30,20	30,47
1/12/2009	28,58	29,08	29,34	29,05
⋮	⋮	⋮	⋮	⋮
12/25/2010	32,19	32,85	31,59	32,34
12/26/2010	31,30	32,41	30,95	31,76
12/27/2010	30,98	31,51	30,64	31,46
12/28/2010	30,91	31,37	30,70	31,11
12/29/2010	31,70	32,25	31,32	31,73
12/30/2010	31,56	32,64	30,95	31,69
12/31/2010	31,89	31,92	31,32	31,55

Lampiran 17. Prakiraan Suhu Minimum Model PLS

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
1/1/2009	23,99	18,23	23,40	22,27
1/2/2009	23,87	18,40	23,01	22,46
1/3/2009	23,64	18,47	22,66	21,80
1/4/2009	23,81	17,96	22,89	21,42
1/5/2009	24,37	18,20	23,25	22,00
1/6/2009	24,01	18,60	22,97	22,34
1/7/2009	24,00	18,80	23,48	22,68
1/8/2009	23,73	18,55	23,33	22,55
1/9/2009	23,83	18,37	23,18	22,36
1/10/2009	23,84	17,94	22,73	22,17
1/11/2009	23,78	17,88	22,75	21,69
1/12/2009	23,38	17,79	22,82	21,82

Lampiran 17. Prakiraan Suhu Minimum Model PLS (Lanjutan)

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
∴	∴	∴	∴	∴
12/25/2010	24,62	18,97	24,17	22,84
12/26/2010	23,97	18,86	23,36	22,57
12/27/2010	23,75	18,45	23,45	22,46
12/28/2010	24,40	18,78	23,74	23,08
12/29/2010	24,48	18,88	23,81	22,97
12/30/2010	24,29	18,75	23,68	22,80
12/31/2010	24,22	18,90	23,79	22,77
Tanggal	Kemayoran (°C)	Pd, Betung (°C)	Priok (°C)	Tangerang (°C)
1/1/2009	24,77	23,60	25,18	23,24
1/2/2009	25,39	24,04	25,57	23,64
1/3/2009	25,21	23,39	25,60	22,79
1/4/2009	24,95	23,52	25,41	22,22
1/5/2009	25,26	24,01	25,74	22,97
1/6/2009	24,96	23,58	25,74	23,21
1/7/2009	24,78	23,85	25,32	23,92
1/8/2009	24,69	23,48	25,34	23,69
1/9/2009	24,27	23,06	24,78	23,51
1/10/2009	23,84	23,36	24,18	23,30
1/11/2009	24,23	23,54	24,56	22,56
1/12/2009	23,67	22,65	24,21	22,78
∴	∴	∴	∴	∴
12/25/2010	25,55	24,21	25,67	23,98
12/26/2010	25,03	23,66	25,49	23,58
12/27/2010	24,55	23,22	25,01	23,45
12/28/2010	24,94	23,81	24,78	24,47
12/29/2010	25,46	24,23	25,76	24,23
12/30/2010	25,31	23,90	25,62	24,08
12/31/2010	25,13	23,84	24,98	24,15

Lampiran 18. Kumulatif Varians Komponen Prediktor Stamet Cengkareng

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	23,29	14	95,29	27	99,81
2	42,10	15	96,12	28	99,86
3	55,55	16	96,79	29	99,90
4	64,33	17	97,38	30	99,94
5	71,72	18	97,93	31	99,96
6	77,75	19	98,32	32	99,98
7	82,29	20	98,68	33	99,99
8	85,48	21	98,95	34	99,99
9	88,19	22	99,16	35	100
10	90,20	23	99,33	36	100

Lampiran 18. Kumulatif Varians Komponen Prediktor Stamet Cengkareng (Lanjutan)

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
11	91,87	24	99,48	37	100
12	93,26	25	99,61	38	100
13	94,35	26	99,71		
Variabilitas T_{MAKS} terjelaskan				40,83%	
Variabilitas T_{MIN} terjelaskan				20,26%	

Lampiran 19. Kumulatif Varians Komponen Prediktor Stamet Citeko

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	26,01	14	96,06	27	99,86
2	43,50	15	96,91	28	99,89
3	56,17	16	97,54	29	99,92
4	64,08	17	98,12	30	99,94
5	70,66	18	98,56	31	99,96
6	75,91	19	98,83	32	99,97
7	80,60	20	99,08	33	99,98
8	84,27	21	99,29	34	99,99
9	87,60	22	99,45	35	99,99
10	90,23	23	99,56	36	100
11	92,39	24	99,66	37	100
12	93,99	25	99,74	38	100
13	95,14	26	99,81	39	100
Variabilitas T_{MAKS} terjelaskan				34,73%	
Variabilitas T_{MIN} terjelaskan				36,01%	

Lampiran 20. Kumulatif Varians Komponen Prediktor Stamet Curug

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	22,00	13	95,37	25	99,93
2	39,32	14	96,51	26	99,95
3	53,40	15	97,36	27	99,97
4	62,49	16	98,06	28	99,98
5	69,45	17	98,60	29	99,99
6	74,97	18	98,95	30	99,99
7	80,32	19	99,26	31	100
8	84,21	20	99,48	32	100
9	87,62	21	99,64	33	100
10	90,18	22	99,76	34	100
11	92,32	23	99,84		
12	94,04	24	99,89		

Lampiran 20. Kumulatif Varians Komponen Prediktor Stamet Curug (Lanjutan)

Variabilitas T_{MAKS} ter jelaskan	35,31%
Variabilitas T_{MIN} ter jelaskan	42,94%

Lampiran 21. Kumulatif Varians Komponen Prediktor Stamet Kemayoran

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	23,80	13	95,07	25	99,83
2	42,54	14	96,03	26	99,88
3	55,74	15	96,91	27	99,93
4	65,19	16	97,59	28	99,95
5	72,42	17	98,09	29	99,98
6	78,05	18	98,51	30	99,99
7	82,32	19	98,85	31	99,99
8	85,64	20	99,12	32	100
9	88,37	21	99,32	33	100
10	90,44	22	99,48	34	100
11	92,28	23	99,63	35	100
12	93,93	24	99,74		
Variabilitas T_{MAKS} ter jelaskan				39,98%	
Variabilitas T_{MIN} ter jelaskan				25,31%	

Lampiran 22. Kumulatif Varians Komponen Prediktor Stamet Pondok Betung

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	24,20	13	94,98	25	99,82
2	42,33	14	95,89	26	99,88
3	55,90	15	96,73	27	99,92
4	65,23	16	97,34	28	99,95
5	72,31	17	97,87	29	99,97
6	77,82	18	98,32	30	99,99
7	81,93	19	98,69	31	99,99
8	85,26	20	98,99	32	100
9	87,98	21	99,27	33	100
10	90,18	22	99,47	34	100
11	92,08	23	99,63	35	100
12	93,74	24	99,74		
Variabilitas T_{MAKS} ter jelaskan				39,00%	
Variabilitas T_{MIN} ter jelaskan				18,38%	

Lampiran 23. Kumulatif Varians Komponen Prediktor Stamet Priok

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	23,25	14	94,49	27	99,59
2	42,63	15	95,29	28	99,68
3	55,17	16	95,99	29	99,77
4	63,91	17	96,64	30	99,84
5	71,24	18	97,15	31	99,89
6	77,13	19	97,61	32	99,94
7	81,46	20	98,02	33	99,97
8	84,65	21	98,40	34	99,98
9	87,23	22	98,72	35	99,99
10	89,31	23	98,96	36	100
11	91,01	24	99,16	37	100
12	92,45	25	99,35	38	100
13	93,56	26	99,50	39	100
Variabilitas T_{MAKS} terjelaskan				35,72%	
Variabilitas T_{MIN} terjelaskan				20,45%	

Lampiran 24. Kumulatif Varians Komponen Prediktor Stamet Tangerang

Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians	Banyak Komponen	%Kumulatif Varians
1	23,45	13	94,66	25	99,79
2	41,51	14	95,66	26	99,86
3	55,27	15	96,59	27	99,91
4	64,73	16	97,23	28	99,95
5	71,91	17	97,78	29	99,97
6	77,61	18	98,24	30	99,99
7	82,04	19	98,62	31	99,99
8	85,26	20	98,94	32	100
9	87,93	21	99,20	33	100
10	90,03	22	99,41	34	100
11	91,75	23	99,57	35	100
12	93,41	24	99,70		
Variabilitas T_{MAKS} terjelaskan				39,13%	
Variabilitas T_{MIN} terjelaskan				11,69%	

Lampiran 25. Model PCR untuk Prakiraan Cuaca Stamet Dermaga

$$\begin{aligned}
 \hat{y}_{r_{\text{MAKS}}} &= 0,0308 \sum_{i=1}^9 e_{i1} \text{dpsdt}_i + 0,0216 \sum_{i=1}^9 e_{i2} \text{mixr1}_i + 0,0067 \sum_{i=1}^9 e_{i3} \text{mixr2}_i + 0,0078 \sum_{i=1}^9 e_{i4} \text{mixr4}_i \\
 &+ 0,0152 \sum_{i=1}^9 e_{i5} \text{omega1}_i + 0,0415 \sum_{i=1}^9 e_{i6} \text{omega1}_i - 0,0196 \sum_{i=1}^9 e_{i7} \text{omega2}_i \\
 &+ 0,0388 \sum_{i=1}^9 e_{i8} \text{omega2}_i - 0,0086 \sum_{i=1}^9 e_{i9} \text{omega4}_i + 0,0077 \sum_{i=1}^9 e_{i10} \text{omega4}_i \\
 &- 0,0496 \sum_{i=1}^9 e_{i11} \text{pblh}_i + 0,0169 \sum_{i=1}^9 e_{i12} \text{ps}_i - 0,0163 \sum_{i=1}^9 e_{i13} \text{psl}_i + 0,0093 \sum_{i=1}^9 e_{i14} \text{qgscrn}_i \\
 &- 0,0015 \sum_{i=1}^9 e_{i15} \text{qgscrn}_i - 0,0573 \sum_{i=1}^9 e_{i16} \text{rh1}_i + 0,0151 \sum_{i=1}^9 e_{i17} \text{rh1}_i + 0,0319 \sum_{i=1}^9 e_{i18} \text{rh2}_i \\
 &+ 0,0207 \sum_{i=1}^9 e_{i19} \text{rh4}_i + 0,0564 \sum_{i=1}^9 e_{i20} \text{rnd}_i - 0,0688 \sum_{i=1}^9 e_{i21} \text{temp1}_i - 0,0551 \sum_{i=1}^9 e_{i22} \text{temp2}_i \\
 &- 0,0320 \sum_{i=1}^9 e_{i23} \text{temp4}_i - 0,0684 \sum_{i=1}^9 e_{i24} \text{t max scr}_i - 0,0579 \sum_{i=1}^9 e_{i25} \text{t min scr}_i \\
 &- 0,0771 \sum_{i=1}^9 e_{i26} \text{tpan}_i - 0,0687 \sum_{i=1}^9 e_{i27} \text{tscrn}_i + 0,0714 \sum_{i=1}^9 e_{i28} \text{u1}_i + 0,0703 \sum_{i=1}^9 e_{i29} \text{u2}_i \\
 &+ 0,0672 \sum_{i=1}^9 e_{i30} \text{u4}_i - 0,0087 \sum_{i=1}^9 e_{i31} \text{ustar}_i - 0,0216 \sum_{i=1}^9 e_{i32} \text{ustar}_i + 0,0057 \sum_{i=1}^9 e_{i33} \text{v1}_i \\
 &- 0,0043 \sum_{i=1}^9 e_{i34} \text{v1}_i + 0,0040 \sum_{i=1}^9 e_{i35} \text{v2}_i + 0,0130 \sum_{i=1}^9 e_{i36} \text{v2}_i - 0,0127 \sum_{i=1}^9 e_{i37} \text{v4}_i \\
 &- 0,0311 \sum_{i=1}^9 e_{i38} \text{zg1}_i - 0,0394 \sum_{i=1}^9 e_{i39} \text{zg2}_i - 0,0625 \sum_{i=1}^9 e_{i40} \text{zg4}_i - 0,0012 \sum_{i=1}^9 e_{i41} \text{zg4}_i
 \end{aligned}$$

Lampiran 25. Model PCR untuk Prakiraan Cuaca Stamet Dermaga (Lanjutan)

$$\begin{aligned}
 \hat{y}_{T_{MN}} = & -0,0094 \sum_{i=1}^9 e_{i1} dpsdt_i - 0,0856 \sum_{i=1}^9 e_{i2} mixr1_i - 0,0619 \sum_{i=1}^9 e_{i3} mixr2_i - 0,0627 \sum_{i=1}^9 e_{i4} mixr4_i \\
 & -0,0036 \sum_{i=1}^9 e_{i5} omega1_i + 0,0063 \sum_{i=1}^9 e_{i6} omega1_i - 0,0016 \sum_{i=1}^9 e_{i7} omega2_i \\
 & +0,0042 \sum_{i=1}^9 e_{i8} omega2_i - 0,0149 \sum_{i=1}^9 e_{i9} omega4_i + 0,0047 \sum_{i=1}^9 e_{i10} omega4_i \\
 & +0,0471 \sum_{i=1}^9 e_{i11} pblh_i - 0,0184 \sum_{i=1}^9 e_{i12} ps_i + 0,0167 \sum_{i=1}^9 e_{i13} psl_i - 0,0815 \sum_{i=1}^9 e_{i14} qgscrn_i \\
 & -0,0139 \sum_{i=1}^9 e_{i15} qgscrn_i + 0,0588 \sum_{i=1}^9 e_{i16} rh1_i - 0,0283 \sum_{i=1}^9 e_{i17} rh1_i - 0,0343 \sum_{i=1}^9 e_{i18} rh2_i \\
 & -0,0310 \sum_{i=1}^9 e_{i19} rh4_i - 0,0062 \sum_{i=1}^9 e_{i20} rnd_i - 0,0191 \sum_{i=1}^9 e_{i21} temp1_i - 0,0382 \sum_{i=1}^9 e_{i22} temp2_i \\
 & -0,0554 \sum_{i=1}^9 e_{i23} temp4_i + 0,0020 \sum_{i=1}^9 e_{i24} t \max scr_i - 0,0558 \sum_{i=1}^9 e_{i25} t \min scr_i \\
 & -0,0289 \sum_{i=1}^9 e_{i26} tpan_i - 0,0138 \sum_{i=1}^9 e_{i27} tscrn_i + 0,0284 \sum_{i=1}^9 e_{i28} u1_i + 0,0313 \sum_{i=1}^9 e_{i29} u2_i \\
 & +0,0190 \sum_{i=1}^9 e_{i30} u4_i + 0,0221 \sum_{i=1}^9 e_{i31} ustari_i - 0,0013 \sum_{i=1}^9 e_{i32} ustari_i + 0,0038 \sum_{i=1}^9 e_{i33} v1_i \\
 & +0,0220 \sum_{i=1}^9 e_{i34} v1_i + 0,0014 \sum_{i=1}^9 e_{i35} v2_i - 0,0236 \sum_{i=1}^9 e_{i36} v2_i - 0,0198 \sum_{i=1}^9 e_{i37} v4_i \\
 & +0,0175 \sum_{i=1}^9 e_{i38} zg1_i + 0,0128 \sum_{i=1}^9 e_{i39} zg2_i - 0,0445 \sum_{i=1}^9 e_{i40} zg4_i + 0,0455 \sum_{i=1}^9 e_{i41} zg4_i
 \end{aligned}$$

Lampiran 26. Prakiraan Suhu Maksimum Model PCR

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
1/1/2009	30,60	24,47	30,60	30,27
1/2/2009	31,51	25,78	31,55	31,81
1/3/2009	31,17	25,29	31,93	31,43
1/4/2009	30,89	24,61	31,17	30,48
1/5/2009	31,04	25,18	31,26	31,19
1/6/2009	31,05	24,96	31,54	30,29
1/7/2009	30,74	24,85	31,18	30,24
1/8/2009	30,96	25,06	31,44	30,30
1/9/2009	30,19	23,88	30,75	29,54
1/10/2009	28,94	23,09	29,61	28,55
1/11/2009	28,90	23,50	30,01	28,63

Lampiran 26. Prakiraan Suhu Maksimum Model PCR (Lanjutan)

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
1/12/2009	29,14	23,25	30,19	28,73
⋮	⋮	⋮	⋮	⋮
12/25/2010	31,80	24,70	31,63	30,85
12/26/2010	30,88	24,27	31,15	29,83
12/27/2010	29,77	23,65	30,00	28,96
12/28/2010	30,54	24,31	31,04	29,66
12/29/2010	31,31	25,70	31,64	31,30
12/30/2010	31,07	25,57	31,43	31,39
12/31/2010	30,64	24,52	30,62	30,26

Tanggal	Kemayoran (°C)	Pd, Betung (°C)	Priok (°C)	Tangerang (°C)
1/1/2009	30,71	31,56	30,84	31,00
1/2/2009	31,80	32,94	31,75	32,16
1/3/2009	31,85	33,01	31,62	32,13
1/4/2009	30,93	32,07	31,01	31,44
1/5/2009	30,99	32,40	31,33	31,63
1/6/2009	31,54	32,27	31,15	31,76
1/7/2009	31,34	32,07	30,92	31,77
1/8/2009	31,80	32,21	31,06	31,98
1/9/2009	30,65	31,57	30,62	31,05
1/10/2009	29,56	30,00	29,73	29,61
1/11/2009	29,86	30,35	29,57	30,00
1/12/2009	29,72	30,88	29,65	30,29
⋮	⋮	⋮	⋮	⋮
12/25/2010	31,49	32,35	31,91	31,91
12/26/2010	30,97	31,89	30,99	31,55
12/27/2010	30,54	30,70	30,13	30,33
12/28/2010	31,09	31,66	31,06	31,20
12/29/2010	31,96	32,63	31,83	31,88
12/30/2010	31,64	32,26	31,44	31,57
12/31/2010	31,08	31,57	31,05	30,98

Lampiran 27. Prakiraan Suhu Minimum Model PCR

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
1/1/2009	24.01	18.37	23.40	22.44
1/2/2009	23.76	18.14	23.06	22.47
1/3/2009	23.35	17.52	22.63	21.82
1/4/2009	23.70	18.01	22.93	21.97
1/5/2009	24.12	18.12	23.24	22.13
1/6/2009	23.81	18.65	22.91	22.63
1/7/2009	24.07	18.82	23.45	22.70

Lampiran 27. Prakiraan Suhu Minimum Model PCR (Lanjutan)

Tanggal	Cengkareng (°C)	Citeko (°C)	Curug (°C)	Dermaga (°C)
1/8/2009	24.06	18.90	23.38	22.69
1/9/2009	24.05	18.29	23.17	22.34
1/10/2009	23.68	17.76	22.62	21.88
1/11/2009	23.76	18.12	22.67	22.11
1/12/2009	23.84	18.00	22.89	22.01
⋮	⋮	⋮	⋮	⋮
12/25/2010	24.48	18.83	24.03	22.81
12/26/2010	24.15	18.52	23.30	22.51
12/27/2010	23.57	17.98	23.14	22.08
12/28/2010	24.14	18.35	23.68	22.48
12/29/2010	24.37	18.85	23.66	22.94
12/30/2010	24.20	18.65	23.63	22.83
12/31/2010	24.10	18.45	23.65	22.55
Tanggal	Kemayoran (°C)	Pd, Betung (°C)	Priok (°C)	Tangerang (°C)
1/1/2009	24.67	23.62	25.05	23.83
1/2/2009	24.85	23.58	25.25	23.69
1/3/2009	24.50	23.15	24.94	23.29
1/4/2009	24.47	23.40	24.92	23.63
1/5/2009	24.66	23.60	25.19	23.89
1/6/2009	24.88	23.90	25.28	23.69
1/7/2009	24.91	23.94	25.27	23.97
1/8/2009	25.15	24.10	25.36	23.99
1/9/2009	24.58	23.49	25.02	23.74
1/10/2009	23.79	22.66	24.43	23.12
1/11/2009	24.03	22.93	24.54	23.19
1/12/2009	24.15	23.22	24.53	23.56
⋮	⋮	⋮	⋮	⋮
12/25/2010	25.22	24.09	25.62	24.33
12/26/2010	24.85	23.83	25.23	23.90
12/27/2010	24.29	22.97	24.68	23.22
12/28/2010	24.81	23.67	25.18	23.78
12/29/2010	25.39	24.14	25.71	24.02
12/30/2010	25.12	24.00	25.50	23.86
12/31/2010	24.84	23.67	25.12	23.78

Lampiran 28. Konstanta λ Terpilih untuk Regresi Ridge

Stamet	λT_{maks}	λT_{min}	Stamet	λT_{maks}	λT_{min}
Cengkareng	1	9	Kemayoran	3	3
Citeko	2	4	Pondok Betung	4	4
Curug	4	9	Priok	8	8
Dermaga	9	9	Tangerang	4	4

Lampiran 29. Model Ridge untuk Prakiraan Cuaca Stamet Dermaga

$$\begin{aligned}
 \hat{y}_{T_{MAXS}} = & -0,075 \sum_{i=1}^9 e_{i1} dpsdt_i + 0,123 \sum_{i=1}^9 e_{i2} mixr1_i + 0,102 \sum_{i=1}^9 e_{i3} mixr2_i + 0,03 \sum_{i=1}^9 e_{i4} mixr4_i \\
 & + 0,153 \sum_{i=1}^9 e_{i5} omega1_i + 0,212 \sum_{i=1}^9 e_{i6} omega1_i - 0,099 \sum_{i=1}^9 e_{i7} omega2_i - 0,029 \sum_{i=1}^9 e_{i8} omega2_i \\
 & - 0,017 \sum_{i=1}^9 e_{i9} omega4_i - 0,082 \sum_{i=1}^9 e_{i10} omega4_i - 0,041 \sum_{i=1}^9 e_{i11} pblh_i + 0,024 \sum_{i=1}^9 e_{i12} ps_i \\
 & + 0,027 \sum_{i=1}^9 e_{i13} psl_i - 0,005 \sum_{i=1}^9 e_{i14} qgscrn_i - 0,067 \sum_{i=1}^9 e_{i15} qgscrn_i + 0,027 \sum_{i=1}^9 e_{i16} rh1_i \\
 & - 0,069 \sum_{i=1}^9 e_{i17} rh1_i + 0,031 \sum_{i=1}^9 e_{i18} rh2_i + 0,037 \sum_{i=1}^9 e_{i19} rh4_i + 0,071 \sum_{i=1}^9 e_{i20} rnd_i - 0,108 \sum_{i=1}^9 e_{i21} temp1_i \\
 & + 0,256 \sum_{i=1}^9 e_{i22} temp2_i + 0,008 \sum_{i=1}^9 e_{i23} temp4_i - 0,513 \sum_{i=1}^9 e_{i24} tmaxscr_i - 0,108 \sum_{i=1}^9 e_{i25} tminscr_i \\
 & + 0,053 \sum_{i=1}^9 e_{i26} tpan_i - 0,099 \sum_{i=1}^9 e_{i27} tscrn_i - 0,301 \sum_{i=1}^9 e_{i28} u1_i - 0,236 \sum_{i=1}^9 e_{i29} u2_i + 0,141 \sum_{i=1}^9 e_{i30} u4_i \\
 & + 0,048 \sum_{i=1}^9 e_{i31} ustar_i + 0,007 \sum_{i=1}^9 e_{i32} ustar_i - 0,012 \sum_{i=1}^9 e_{i33} v1_i + 0,012 \sum_{i=1}^9 e_{i34} v1_i - 0,094 \sum_{i=1}^9 e_{i35} v2_i \\
 & + 0,089 \sum_{i=1}^9 e_{i36} v2_i - 0,032 \sum_{i=1}^9 e_{i37} v4_i - 0,096 \sum_{i=1}^9 e_{i38} zg1_i + 0,022 \sum_{i=1}^9 e_{i39} zg2_i \\
 & + 0,001 \sum_{i=1}^9 e_{i40} zg4_i + 0,079 \sum_{i=1}^9 e_{i41} zg4_i
 \end{aligned}$$

$$\begin{aligned}
 \hat{y}_{T_{MIN}} = & -0,099 \sum_{i=1}^9 e_{i1} dpsdt_i - 0,125 \sum_{i=1}^9 e_{i2} mixr1_i - 0,129 \sum_{i=1}^9 e_{i3} mixr2_i + 0,008 \sum_{i=1}^9 e_{i4} mixr4_i \\
 & + 0,165 \sum_{i=1}^9 e_{i5} omega1_i + 0,015 \sum_{i=1}^9 e_{i6} omega1_i - 0,051 \sum_{i=1}^9 e_{i7} omega2_i - 0,038 \sum_{i=1}^9 e_{i8} omega2_i \\
 & + 0,083 \sum_{i=1}^9 e_{i9} omega4_i + 0,031 \sum_{i=1}^9 e_{i10} omega4_i + 0,01 \sum_{i=1}^9 e_{i11} pblh_i + 0,037 \sum_{i=1}^9 e_{i12} ps_i \\
 & + 0,066 \sum_{i=1}^9 e_{i13} psl_i - 0,153 \sum_{i=1}^9 e_{i14} qgscrn_i + 0,054 \sum_{i=1}^9 e_{i15} qgscrn_i - 0,021 \sum_{i=1}^9 e_{i16} rh1_i \\
 & - 0,172 \sum_{i=1}^9 e_{i17} rh1_i - 0,261 \sum_{i=1}^9 e_{i18} rh2_i + 0,009 \sum_{i=1}^9 e_{i19} rh4_i - 0,022 \sum_{i=1}^9 e_{i20} rnd_i - 0,083 \sum_{i=1}^9 e_{i21} temp1_i \\
 & + 0,066 \sum_{i=1}^9 e_{i22} temp2_i - 0,219 \sum_{i=1}^9 e_{i23} temp4_i + 0,149 \sum_{i=1}^9 e_{i24} tmaxscr_i - 0,084 \sum_{i=1}^9 e_{i25} tminscr_i \\
 & - 0,026 \sum_{i=1}^9 e_{i26} tpan_i - 0,01 \sum_{i=1}^9 e_{i27} tscrn_i + 0,055 \sum_{i=1}^9 e_{i28} u1_i + 0,016 \sum_{i=1}^9 e_{i29} u2_i + 0,124 \sum_{i=1}^9 e_{i30} u4_i \\
 & + 0,014 \sum_{i=1}^9 e_{i31} ustar_i - 0,088 \sum_{i=1}^9 e_{i32} ustar_i - 0,075 \sum_{i=1}^9 e_{i33} v1_i - 0,049 \sum_{i=1}^9 e_{i34} v1_i + 0,01 \sum_{i=1}^9 e_{i35} v2_i \\
 & + 0,053 \sum_{i=1}^9 e_{i36} v2_i - 0,021 \sum_{i=1}^9 e_{i37} v4_i + 0,004 \sum_{i=1}^9 e_{i38} zg1_i - 0,166 \sum_{i=1}^9 e_{i39} zg2_i + 0,022 \sum_{i=1}^9 e_{i40} zg4_i \\
 & + 0,29 \sum_{i=1}^9 e_{i41} zg4_i
 \end{aligned}$$

Lampiran 30. Prakiraan Suhu Maksimum Model Ridge

Tanggal	Kemayoran (°C)	Priok (°C)	Cengkareng (°C)	Pd. Betung (°C)
1/1/2009	31,46	31,03	31,33	32,20
1/2/2009	33,48	32,83	32,76	33,74
1/3/2009	33,48	32,54	32,20	33,95
1/4/2009	32,89	32,44	32,33	33,53
1/5/2009	33,29	32,50	32,52	33,85
1/6/2009	31,77	31,10	30,95	32,90
1/7/2009	31,30	30,83	30,64	32,34
1/8/2009	31,05	30,76	30,98	32,13
1/9/2009	30,06	29,80	29,60	31,24
1/10/2009	28,77	29,37	28,64	29,75
1/11/2009	29,90	30,24	29,87	30,89
1/12/2009	28,54	29,30	28,84	29,23
⋮	⋮	⋮	⋮	⋮
12/25/2010	32,15	31,60	31,41	32,91
12/26/2010	31,45	30,97	31,05	32,49
12/27/2010	30,99	30,54	30,83	31,47
12/28/2010	30,89	30,79	30,74	31,44
12/29/2010	31,70	31,32	31,07	32,43
12/30/2010	31,51	30,94	30,88	32,70
12/31/2010	31,81	31,26	31,11	32,02

Tanggal	Curug (°C)	Tangerang (°C)	Citeko (°C)	Dermaga (°C)
1/1/2009	31,36	32,03	24,96	30,83
1/2/2009	32,84	33,56	26,96	31,80
1/3/2009	32,74	33,31	26,16	31,85
1/4/2009	32,63	33,25	26,56	32,14
1/5/2009	32,77	33,49	26,80	32,12
1/6/2009	31,36	31,79	26,44	31,52
1/7/2009	30,77	31,55	25,52	30,42
1/8/2009	30,80	31,76	25,11	30,39
1/9/2009	30,40	30,44	23,58	29,16
1/10/2009	28,64	29,04	22,71	28,02
1/11/2009	30,17	30,66	23,80	29,16
1/12/2009	28,13	28,94	22,05	27,16
⋮	⋮	⋮	⋮	⋮
12/25/2010	31,56	32,39	25,26	31,39
12/26/2010	31,49	31,88	25,52	30,73
12/27/2010	30,05	31,41	23,72	29,15

Lampiran 30. Prakiraan Suhu Maksimum Model Ridge (Lanjutan)

Tanggal	Curug (°C)	Tangerang (°C)	Citeko (°C)	Dermaga (°C)
12/28/2010	30,72	31,16	24,18	29,68
12/29/2010	31,20	31,71	25,70	31,14
12/30/2010	31,79	31,67	25,84	31,87
12/31/2010	30,95	31,60	25,14	30,60

Lampiran 31. Prakiraan Suhu Minimum Model Ridge

Tanggal	Kemayoran (°C)	Priok (°C)	Cengkareng (°C)	Pd. Betung (°C)
1/1/2009	24,60	25,23	24,09	23,53
1/2/2009	25,30	25,42	23,82	24,02
1/3/2009	25,06	25,52	23,55	23,31
1/4/2009	24,78	25,19	23,75	23,42
1/5/2009	25,23	25,87	24,46	23,94
1/6/2009	24,97	25,90	24,12	23,55
1/7/2009	24,74	25,34	24,07	23,84
1/8/2009	24,50	25,30	23,83	23,54
1/9/2009	24,22	24,73	23,86	23,01
1/10/2009	23,86	24,20	23,72	23,42
1/11/2009	24,35	24,73	23,78	23,50
1/12/2009	23,67	24,35	23,45	22,64
⋮	⋮	⋮	⋮	⋮
12/25/2010	25,33	25,78	24,64	24,20
12/26/2010	24,87	25,42	24,09	23,60
12/27/2010	24,48	24,85	23,72	23,11
12/28/2010	25,10	24,75	24,21	23,79
12/29/2010	25,40	25,65	24,50	24,24
12/30/2010	25,45	25,62	24,26	23,89
12/31/2010	25,31	25,06	24,12	23,89

Tanggal	Curug (°C)	Tangerang (°C)	Citeko (°C)	Dermaga (°C)
1/1/2009	23,30	24,00	18,24	22,39
1/2/2009	22,82	23,91	18,37	22,44
1/3/2009	22,38	23,76	18,45	21,99
1/4/2009	22,73	23,93	17,91	21,44
1/5/2009	23,24	24,33	18,23	22,02
1/6/2009	23,09	24,13	18,65	22,07
1/7/2009	23,48	23,75	18,78	22,70
1/8/2009	23,28	23,91	18,54	22,56

Lampiran 31. Prakiraan Suhu Minimum Model Ridge (Lanjutan)

Tanggal	Curug (°C)	Tangerang (°C)	Citeko (°C)	Dermaga (°C)
1/9/2009	23,21	23,61	18,39	22,34
1/10/2009	22,65	23,33	17,98	22,20
1/11/2009	22,73	23,16	17,86	21,71
1/12/2009	22,72	22,71	17,74	21,85
⋮	⋮	⋮	⋮	⋮
12/25/2010	24,08	24,63	18,92	22,80
12/26/2010	23,34	23,90	18,85	22,46
12/27/2010	23,34	23,21	18,46	22,42
12/28/2010	23,68	23,87	18,69	23,22
12/29/2010	23,97	24,43	18,80	22,93
12/30/2010	23,89	24,37	18,76	22,99
12/31/2010	23,74	24,10	18,83	22,93

Lampiran 32. Estimasi Bobot dan Koefisien Regresi Model Simultan Ensemble BMA, Suhu Maksimum

Tanggal	Stamet	Bobot			PLS		PCR		Ridge	
		PLS	PCR	Ridge	β_0	β_1	β_0	β_1	β_0	β_1
20090131	Kemayoran	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
20090131	Priok	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
20090131	Cengkareng	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
20090131	Pd. Betung	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
20090131	Curug	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
20090131	Tangerang	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
20090131	Citeko	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
20090131	Dermaga	0,99	0,00	0,01	-2,21	1,06	-2,91	1,08	-2,21	1,06
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
20110101	Kemayoran	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95
20110101	Priok	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95
20110101	Cengkareng	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95
20110101	Pd. Betung	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95
20110101	Curug	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95
20110101	Tangerang	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95
20110101	Citeko	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95
20110101	Dermaga	0,50	0,21	0,29	1,68	0,95	1,38	0,96	1,67	0,95

Lampiran 33. Estimasi Bobot dan Koefisien Regresi Model Simultan Ensemble BMA, Suhu Minimum

Tanggal	Stamet	Bobot			PLS		PCR		Ridge	
		PLS	PCR	Ridge	β_0	β_1	β_0	β_1	β_0	β_1
20090131	Kemayoran	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
20090131	Priok	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
20090131	Cengkareng	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
20090131	Pd. Betung	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
20090131	Curug	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
20090131	Tangerang	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
20090131	Citeko	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
20090131	Dermaga	0,00	0,01	0,99	0,01	0,99	0,05	0,99	-0,07	0,99
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
20110101	Kemayoran	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00
20110101	Priok	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00
20110101	Cengkareng	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00
20110101	Pd. Betung	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00
20110101	Curug	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00
20110101	Tangerang	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00
20110101	Citeko	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00
20110101	Dermaga	0,00	0,24	0,76	-0,23	1,01	0,43	0,98	-0,15	1,00

Lampiran 34. Prakiraan Suhu Maksimum Model Spatial BMA

Tanggal	Stamet	obs_test	fore Spatial BMA	fore_test	v	u
20090302	1	33	31.5614761	31.37594	31.743	31.49468
20090302	2	32.4	31.52862393	31.18044	31.58769	31.30497
20090302	3	31.9	31.21840707	31.03794	31.43628	31.15417
20090302	4	34	32.80031865	32.56835	32.93442	32.67552
20090302	5	33.2	31.59564197	31.40638	31.77421	31.50998
20090302	6	33	32.28267664	32.12521	32.49723	32.20521
20090302	7	26	25.18666194	24.8408	25.26122	24.98685
20090302	8	31.8	30.01727757	29.7607	30.16528	29.90362
⋮	⋮	⋮	⋮	⋮	⋮	⋮
20101231	1	33.2	32.33699516	32.04207	32.45336	32.10485
20101231	2	33.5	31.60645094	31.38071	31.77233	31.49757
20101231	3	32	31.843163	31.60708	31.97935	31.69864
20101231	4	34	33.00738505	32.80289	33.18191	32.85533
20101231	5	32.5	32.03346097	31.88914	32.257	31.99087

Lampiran 34. Prakiraan Suhu Maksimum Model Spatial BMA (Lanjutan)

Tanggal	Stamet	obs_test	fore Spatial BMA	fore_test	v	u
20101231	6	33.4	32.73947795	32.57639	32.93058	32.68373
20101231	7	25	25.58897292	25.34508	25.80213	25.51961
20101231	8	31.6	30.78376798	30.47516	30.87447	30.60428

Lampiran 35. Prakiraan Suhu Minimum Model Spatial BMA

Tanggal	Stamet	obs_test	fore Spatial BMA	fore_test	v	u
20090302	1	25	24.12424	24.77478	26.3573	23.20047
20090302	2	26	24.21661	24.73765	26.11175	23.2778
20090302	3	24	23.66827	23.85797	25.32754	22.27089
20090302	4	23.8	23.12191	23.54193	25.07573	22.04918
20090302	5	23.6	22.51094	23.19503	24.7666	21.6876
20090302	6	25.4	23.23532	23.79436	25.34771	22.37471
20090302	7	19.2	18.26633	18.91921	20.54681	17.37164
20090302	8	22.6	22.06046	22.59305	24.07641	21.09535
⋮	⋮	⋮	⋮	⋮	⋮	⋮
20101231	1	26.2	25.67258	25.18389	26.79025	23.28391
20101231	2	27	25.86682	25.01234	26.5465	23.08785
20101231	3	24.6	24.96469	24.04911	25.85069	22.05857
20101231	4	24.8	23.86435	23.82165	25.40242	22.03234
20101231	5	24	24.00879	23.66657	25.43449	22.12132
20101231	6	23.8	24.4471	24.01389	25.65585	22.11906
20101231	7	18.4	19.61208	18.73814	20.43281	17.06889
20101231	8	23.4	23.22737	22.79171	24.38787	20.91084

Lampiran 36. Syntax Program

a. Syntax untuk Reduksi Dimensi Parameter NWP dengan PCA

```
pca1<-read.csv("D:/kuliah s2/smt 3/Thesis/New folder/Dermaga/parameter/dpsdt.csv",header=T)
pca=scale(pca1)
kovarian<-cov(pca)
eigenvalue<-eigen(kovarian)$value
eigenvalue
fit_pca<-princomp(pca,cor=FALSE)
sdev<-fit_pca$sdev
sumvar<-sum(eigenvalue)
p<-ncol(pca)
var <- matrix(c(0),p,1)
cumvar <- matrix(c(0),p,1)
for(i in 1:p){
  var[i] <- eigenvalue[i]/sumvar
  if (i == 1) {cumvar[i] <- var[i]}
  else
    cumvar[i] <- var[i] + cumvar[i-1]
}
write.csv(cbind(eigenvalue,sdev,var,cumvar),"D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/eigenvalue_dpsdt.csv")
loadings<-loadings(fit_pca)
write.csv(cbind(loadings),"D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/koefisien_dpdst.csv")
score<-fit_pca$scores
write.csv(cbind(score),"D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/score_dpsdt.csv")
```

Keterangan : Syntax di atas hanya untuk parameter dpsdt stamet Dermaga

b. Syntax untuk Menghasilkan Prakiraan Ensemble

```
#PLS#
library(pls)
library(plsdepot)
press<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/pca_drmg.csv",header=T)
tr<-press[,1:44]
#mencari komponen optimal Tmaks#
lm<-
mvr(TMAX~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.omega2+
PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,data=as.data.frame(tr),ncomp=41,
method=pls.options())$plsr)
summary(crossval(lm,length.seg=1))
#mencari komponen optimal Tmin#
lm<-
mvr(TMIN~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.omega2+
PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,data=as.data.frame(tr),ncomp=41,
method=pls.options())$plsr)
summary(crossval(lm,length.seg=1))

pca<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/pca_drmg.csv",header=T)
t.maks<-pca[,42]
t.min<-pca[,43]
Xtr<-pca[,1:41]
dmg1.tr=plsreg1(Xtr,t.maks,comps=19)
dmg2.tr=plsreg1(Xtr,t.min,comps=4)
dmg1.tr$reg.coefs
dmg2.tr$reg.coefs
Ytr1.hat=as.matrix(Xtr)%*%(dmg1.tr$reg.coefs[-1])
Ytr2.hat=as.matrix(Xtr)%*%(dmg2.tr$reg.coefs[-1])
obs<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/obs_drmg.csv",header=T)
a=vector(mode="numeric",length=2)
b=vector(mode="numeric",length=2)
for (i in 1:2){
a[i]=mean(obs[,i])
b[i]=sd(obs[,i])
}
```

```

PLS1.Tmaks=(Ytr1.hat*b[1])+a[1]
PLS1.Tmin=(Ytr2.hat*b[2])+a[2]
y1.hat=cbind(PLS1.Tmaks,PLS1.Tmin)
write.csv(y1.hat,"D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/ramalanPLS1drmg.csv")

#PCR#
library(pls)
library(plsdepot)
press<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/pca_drmg.csv",header=T)
tr<-press[1:708,1:44]
lm<-
mvr(TMAX+TMIN+RH.AVERAGE~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+P
C2.omega1+PC1.omega2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1
.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,data=as.data.frame(tr),ncomp=41,
method=pls.options()$pcr)
summary(crossval(lm,length.seg=1))

tr1<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/pca_drmg.csv",header=T)
x1=tr1[,1:41]
y=tr1[,42:44]
x=scale(x1)
tr=cbind(x,y)

tmax.tr=pcr(TMAX~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.
omega2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,7,data=tr)

tmin.tr=pcr(TMIN~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.o
mega2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,7,data=tr)

tmax.tr.fit=tmax.tr$fitted.values[,7]
tmin.tr.fit=tmin.tr$fitted.values[,7]
tmax.tr$coefficients
tmin.tr$coefficients

```

```

obs<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/obs_drmg.csv",header=T)

a=vector(mode="numeric",length=3)
b=vector(mode="numeric",length=3)
for (i in 1:2){
a[i]=mean(obs[,i])
b[i]=sd(obs[,i])
}
tmaks=(tmax.tr.fit*b[1])+a[1]
tmin=(tmin.tr.fit*b[2])+a[2]
y.hat=cbind(tmaks,tmin)
write.csv(y.hat,"D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/ramalanPCR.csv")

#Ridge#
library(pls)
library(plsdepot)
library(MASS)

pca<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/pca_drmg.csv",header=T)
tr<-pca[1:708,1:44]
y<-tr[1:708,42:44]
x<-tr[1:708,1:41]

#mencari penalti optimal TMAX Ridge#
ridge.tmaks<-
lm.ridge(TMAX~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.ome
ga2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+PC2.qgscrn+
PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+PC.temp4+PC.tmaxscr+PC.tmi
nscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+P
C2.v2+PC.v4+PC.zg1+PC.zg2+PC1.zg4+PC2.zg4,data=tr,lambda=seq(0,10,.001))
plot(ridge.tmaks)

#mencari penalti optimal TMAX Ridge#
ridge.tmin<-
lm.ridge(TMIN~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.ome
ga2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+PC2.qgscrn+
PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+PC.temp4+PC.tmaxscr+PC.tmi
nscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+P
C2.v2+PC.v4+PC.zg1+PC.zg2+PC1.zg4+PC2.zg4,data=tr,lambda=seq(0,10,.001))
plot(ridge.tmin)

pca<-read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/pca_drmg.csv",header=T)
tr<-pca[1:708,1:44]
y<-tr[1:708,42:44]
x<-tr[1:708,1:41]

```

```

ridge <-
lm.ridge(TMAX~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.ome
ga2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,data=tr,lambda = seq(0,30,.001))

ridge <-
lm.ridge(TMAX~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.ome
ga2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,data=tr,lambda=22.11)
select(ridge)
summary(ridge)

ridge1 <-
lm.ridge(TMIN~PC.dpsdt+PC.mixr1+PC.mixr2+PC.mixr4+PC1.omega1+PC2.omega1+PC1.ome
ga2+PC2.omega2+PC1.omega4+PC2.omega4+PC.pblh+PC.ps+PC.psl+PC1.qgscrn+
PC2.qgscrn+PC1.rh1+PC2.rh1+PC.rh2+PC.rh4+PC.rnd+PC.temp1+PC.temp2+
PC.temp4+PC.tmaxscr+PC.tminscr+PC.tpan+PC.tscrn+PC.u1+PC.u2+PC.u4
+PC1.ustar+PC2.ustar+PC1.v1+PC2.v1+PC1.v2+PC2.v2+PC.v4+PC.zg1+PC.zg2+
PC1.zg4+PC2.zg4,data=tr,lambda = 22.11)
summary(ridge1)

ridge.hat<-coef(ridge)[1] + coef(ridge)[2]*tr[,1] +
coef(ridge)[3]*tr[,2] + coef(ridge)[4]*tr[,3] + coef(ridge)[5]*tr[,4]+
coef(ridge)[6]*tr[,5] + coef(ridge)[7]*tr[,6] + coef(ridge)[8]*tr[,7]+
coef(ridge)[9]*tr[,8] + coef(ridge)[10]*tr[,9] + coef(ridge)[11]*tr[,10]+
coef(ridge)[12]*tr[,11]+coef(ridge)[13]*tr[,12]+coef(ridge)[14]*tr[,13]+
coef(ridge)[15]*tr[,14]+coef(ridge)[16]*tr[,15]+coef(ridge)[17]*tr[,16]+
coef(ridge)[18]*tr[,17]+coef(ridge)[19]*tr[,18]+coef(ridge)[20]*tr[,19]+
coef(ridge)[21]*tr[,20]+coef(ridge)[22]*tr[,21]+coef(ridge)[23]*tr[,22]+
coef(ridge)[24]*tr[,23]+coef(ridge)[25]*tr[,24]+coef(ridge)[26]*tr[,25]+
coef(ridge)[27]*tr[,26]+coef(ridge)[28]*tr[,27]+coef(ridge)[29]*tr[,28]+
coef(ridge)[30]*tr[,29]+coef(ridge)[31]*tr[,30]+coef(ridge)[32]*tr[,31]+
coef(ridge)[33]*tr[,32]+coef(ridge)[34]*tr[,33]+coef(ridge)[35]*tr[,34]+
coef(ridge)[36]*tr[,35]+coef(ridge)[37]*tr[,36]+coef(ridge)[38]*tr[,37]+
coef(ridge)[39]*tr[,38]+coef(ridge)[40]*tr[,39]+coef(ridge)[41]*tr[,40]+
coef(ridge)[42]*tr[,41]

```

```

ridge1.hat<-coef(ridge1)[1] + coef(ridge1)[2]*tr[,1] +
coef(ridge1)[3]*tr[,2] + coef(ridge1)[4]*tr[,3] + coef(ridge1)[5]*tr[,4]+
coef(ridge1)[6]*tr[,5] + coef(ridge1)[7]*tr[,6] + coef(ridge1)[8]*tr[,7]+
coef(ridge1)[9]*tr[,8] + coef(ridge1)[10]*tr[,9] + coef(ridge1)[11]*tr[,10]+
coef(ridge1)[12]*tr[,11]+coef(ridge1)[13]*tr[,12]+coef(ridge1)[14]*tr[,13]+
coef(ridge1)[15]*tr[,14]+coef(ridge1)[16]*tr[,15]+coef(ridge1)[17]*tr[,16]+
coef(ridge1)[18]*tr[,17]+coef(ridge1)[19]*tr[,18]+coef(ridge1)[20]*tr[,19]+
coef(ridge1)[21]*tr[,20]+coef(ridge1)[22]*tr[,21]+coef(ridge1)[23]*tr[,22]+
coef(ridge1)[24]*tr[,23]+coef(ridge1)[25]*tr[,24]+coef(ridge1)[26]*tr[,25]+
coef(ridge1)[27]*tr[,26]+coef(ridge1)[28]*tr[,27]+coef(ridge1)[29]*tr[,28]+
coef(ridge1)[30]*tr[,29]+coef(ridge1)[31]*tr[,30]+coef(ridge1)[32]*tr[,31]+
coef(ridge1)[33]*tr[,32]+coef(ridge1)[34]*tr[,33]+coef(ridge1)[35]*tr[,34]+
coef(ridge1)[36]*tr[,35]+coef(ridge1)[38]*tr[,37]+
coef(ridge1)[39]*tr[,38]+coef(ridge1)[40]*tr[,39]+coef(ridge1)[41]*tr[,40]+
coef(ridge1)[42]*tr[,41]

a=mean(obs[,1])
b=sd(obs[,1])
tmaks=ridge.hat*b+a
c=mean(obs[,2])
d=sd(obs[,2])
tmin=ridge1.hat*d+c

write.csv(cbind(tmaks,tmin),"D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/ramalanRidge.csv")

```

Keterangan : Syntax hanya untuk stamet Dermaga

c. Syntax untuk Menghasilkan Prakiraan BMA Secara Simultan

```

library(ensembleBMA)
ensBMAtest=read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/BMA/ens_BMA
(gab).csv",header=T)
data=ensBMAtest
T=30
n=3
tgl=data$vdate
start=which(data$vdate==tgl)
jmlh=length(start)
test=jmlh-8*T
data1=data[,6:8]
y1=data[,5]
x=matrix(c(0),test,n)
y=vector(mode="numeric",length=test)
x=as.matrix(data1[(8*T+1):jmlh])
y=y1[(8*T+1):jmlh]
ensMemNames = c("PLS","PCR","Ridge")
obs = paste("Tmaks","obs", sep = ".")

```

```

ens = paste("Tmaks", ensMemNames, sep = ".")
BMADData = ensembleData(forecasts = data[,ens],dates = data[,"vdate"],observations =
data[,obs],station = data[,"id.stat"],forecastHour = 24,initializationTime = "00")
BMAfit=ensembleBMA(BMADData,model="normal",trainingDays=T)
write.csv(cbind(BMAfit$weights),"D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/BMA/weight2.csv")
miss=c(186,192,196,198,199,211,217,218,235,237,259,271,272,273,380,385,406,458,524,586,65
3,722,731)
aa=miss-T
aa1=3*aa
aa2=aa1-1
aa3=aa1-2
a=sort(c(aa3,aa2,aa1))
bobot1=as.vector(BMAfit$weights)
bobot=bobot1[-a]
write.csv(cbind(bobot),"D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/BMA/bobot.csv")
bobot=read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/BMA/bobot.csv",header=T)
bobot=as.vector(bobot$bobot)
w1=vector(mode="numeric",length=test)
w2=vector(mode="numeric",length=test)
w3=vector(mode="numeric",length=test)
k=1
for (l in 1:test){
  w1[l]=bobot[k]
  w2[l]=bobot[k+8]
  w3[l]=bobot[k+16]
  k=k+24
}
w=read.csv("D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/BMA/bobot_30_gab.csv",header=T)
w1=w$w1
w2=w$w2
w3=w$w3
aa4=6*aa
aa5=aa4-1
aa6=aa4-2
aa7=aa4-3
aa8=aa4-4
aa9=aa4-5
az=sort(c(aa4,aa5,aa6,aa7,aa8,aa9))
koef=BMAfit$biasCoefs
koef=as.matrix(koef)
koef1=koef[-az]
write.csv(cbind(koef1),"D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/BMA/koef2.csv")
koef1=read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/BMA/koef2.csv",header=T)
koef1=as.vector(koef1$koef1)

```



```

i=1
a1=vector(mode="numeric",length=test)
b1=vector(mode="numeric",length=test)
a2=vector(mode="numeric",length=test)
b2=vector(mode="numeric",length=test)
a3=vector(mode="numeric",length=test)
b3=vector(mode="numeric",length=test)
for (j in 1:test){
  a1[j]=koef1[i]
  b1[j]=koef1[1+i]
  a2[j]=koef1[2+i]
  b2[j]=koef1[3+i]
  a3[j]=koef1[4+i]
  b3[j]=koef1[5+i]
  i=i+6
}
coef=read.csv("D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/BMA/koefisien_30_gab.csv",header=T)
a1=coef$a1
a2=coef$a2
a3=coef$a3
b1=coef$b1
b2=coef$b2
b3=coef$b3
sdev1=as.vector(BMAfit$sdev)
sdev=sdev1[-aa]
write.csv(cbind(sdev),"D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/BMA/sdev.csv")
sdev=read.csv("D:/kuliah s2/smt 3/Thesis/olah data niswa/Dermaga/BMA/sdev1.csv",header=T)
sdev=sdev$sdev
obs=y
fore1=as.vector(a1+b1*(x[,1]))
fore2=as.vector(a2+b2*(x[,2]))
fore3=as.vector(a3+b3*(x[,3]))
mean_BMA=w1*fore1+w2*fore2+w3*fore3
mean_ens=as.vector((x[,1]+x[,2]+x[,3])/n)
var_BMA=(w1*(fore1-mean_BMA)^2+w2*(fore2-mean_BMA)^2+w3*(fore3-
mean_BMA)^2)+sdev

#nilai_CRPS=CRPS(BMAfit,BMAData)
res_PLS=vector(length=test)
res_PCR=vector(length=test)
res_Ridge=vector(length=test)
res_BMA=vector(length=test)

```

```

for (i in 1:test){
  res_PLS[i]=(y[i]-x[i,1])^2
  res1=as.vector(res_PLS)
  res_PCR[i]=(y[i]-x[i,2])^2
  res2=as.vector(res_PCR)
  res_Ridge[i]=(y[i]-x[i,3])^2
  res3=as.vector(res_Ridge)
  res_BMA[i]=(y[i]-mean_BMA[i])^2
  res4=as.vector(res_BMA)
}
RMSE_PLS=sqrt(sum(res1)/test)
RMSE_PCR=sqrt(sum(res2)/test)
RMSE_Ridge=sqrt(sum(res3)/test)
RMSE_BMA=sqrt(sum(res4)/test)
RMSE=list(RMSE_PLS=RMSE_PLS,RMSE_PCR=RMSE_PCR,RMSE_Ridge=RMSE_Ridge,R
MSE_BMA=RMSE_BMA)
l=mean_BMA-1.96*sqrt(var_BMA)
u=mean_BMA+1.96*sqrt(var_BMA)
pls=vector(mode="numeric",length=test)
pcr=vector(mode="numeric",length=test)
ridge=vector(mode="numeric",length=test)
c4=vector(mode="numeric",length=test)
bma=matrix(0,test,3)
for (i in 1:test){
  s=100000
  pls[i]=sample(rnorm(s,fore1[i],sdev[i]),1)
  pcr[i]=sample(rnorm(s,fore2[i],sdev[i]),1)
  ridge[i]=sample(rnorm(s,fore3[i],sdev[i]),1)
  bma[i,]=cbind(pls[i],pcr[i],ridge[i])
  if (y[i]<min(bma[i,])|y[i]>max(bma[i,])){c4[i]=0}
  else c4[i]=1
}
coverage=(sum(c4)/test)*100
RMSE
nilai_CRPS
coverage
library(SpecsVerification)
rank.hist1=Rankhist(bma,y)
PlotRankhist(rank.hist1,mode="prob.paper")
write.csv(cbind(w1,w2,w3),"D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/BMA/bobot_30_gab.csv")
write.csv(cbind(a1,b1,a2,b2,a3,b3),"D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/BMA/koeffisien_30_gab.csv")
write.csv(cbind(x,mean_ens,mean_BMA,obs),"D:/kuliah s2/smt 3/Thesis/olah data
niswa/Dermaga/BMA/forecast_30_gab.csv")

```

Keterangan : Syntax hanya untuk suhu maksimum

d. Syntax untuk Menghasilkan Prakiraan Spatial BMA

```
library(fields)
library(RandomFields)
library(ProbForecastGOP)
library(gstat)
library(SpecsVerification)
library(fields)
library(RandomFields)
library(ProbForecastGOP)
library(gstat)
library(SpecsVerification)
library(automap)

datasbma=read.csv("D:/kuliah s2/smt 3/Thesis/spaBMA/datasbma.csv",header=T)
koord=read.csv("D:/kuliah s2/smt 3/Thesis/spaBMA/tmaks_utm.csv",header=T)
utm=koord[,-3]
train=30
day <-datasbma$date.obs[1:(8*train)]
id <- datasbma$id.stat[1:(8*train)]
obs1 <- datasbma$obs[1:(8*train)]
forecast <-datasbma$forecast[1:(8*train)]
col1=cbind(id,obs1)
col2=cbind(id,forecast)
obs=aggregate(col1[,2],list(col1[,1]),mean)
fore=aggregate(col2[,2],list(col2[,1]),mean)
obs=obs[,-1]
fore=fore[,-1]
data=cbind(utm,obs,fore)

coordinates(data) =~ x+y
variog= autofitVariogram(obs~fore,data,model="Exp")
param=variog$var_model
sill=param$psill
range=(param$range[2])/1000

datasbma=read.csv("D:/kuliah s2/smt 3/Thesis/spaBMA/datasbma.csv",header=T)
train=30
day <-datasbma$date.obs[1:(8*train)]
id <- datasbma$id.stat[1:(8*train)]
coord1 <- datasbma$lon.stat[1:(8*train)]
coord2 <- datasbma$lat.stat[1:(8*train)]
obs <- datasbma$obs[1:(8*train)]
forecast <-datasbma$forecast[1:(8*train)]
lt=length(datasbma$obs)/8
ntest=lt-train
```

```

lm.tmaks<- lm(obs ~ forecast)
lm.tmaks$coeff

variogram <-
  Emp.variog(day=day,obs=obs,forecast=forecast,id=id,coord1=coord1,coord2=coord2,
    cut.points=seq(0,100,by=1),max.dist=NULL,nbins=NULL)

plot(variogram$bin.midpoints,variogram$empir.variog,xlab="Jarak(km)",ylab="Semi-variars")
lines(variogram$bin.midpoints,linesmodel(distance=variogram$bin.midpoints,
  variog.model="exponential",param=c(sill[1],sill[2],range)))

nsim=99
a=array(rep(0,8*ntest*nsim),dim=c(8,ntest,nsim))
for (z in 1:nsim){
  for (i in 1:ntest){
    model <- RMexp(var=sill[2],scale=range) + RMnugget(var=sill[1])
    simu <- RFsimulate(model,x=coord1[1:8],y=coord2[1:8])
    a[,i,z]=simu$variable1
  }
}

j=1
z=1
s=matrix(c(0),(ntest*8),nsim)
for (z in 1:nsim){
  for (i in 1:ntest){
    k=j+1
    l=j+2
    m=j+3
    n=j+4
    o=j+5
    p=j+6
    q=j+7
    s[j,z]=a[1,i,z]
    s[k,z]=a[2,i,z]
    s[l,z]=a[3,i,z]
    s[m,z]=a[4,i,z]
    s[n,z]=a[5,i,z]
    s[o,z]=a[6,i,z]
    s[p,z]=a[7,i,z]
    s[q,z]=a[8,i,z]
    j=j+8
  }
  j=1
  z=z+1
}

```

```

obs_test <-datasbma$obs[-(1:(8*train))]
fore_test <-datasbma$forecast[-(1:(8*train))]
b0=as.vector(lm.tmaks$coeff[1])
b1=as.vector(lm.tmaks$coeff[2])
foresbma=matrix(c(0),(ntest*8),nsim)
for (j in 1:nsim){
  foresbma[,j]=(b0+b1*fore_test)+s[,j]
}

res=matrix(c(0),(ntest*8),nsim)
for (j in 1:nsim){
  res[,j]=(obs_test-foresbma[,j])^2
}

sta=c(seq(from=1,to=8,by=1))
col=factor(rep(sta,ntest))
resi=cbind(col,res)
MSE_sbma=aggregate(resi[,2:(nsim+1)],list(resi[,1]),mean)
RMSE_sbma=vector(mode="numeric",length=nsim)
for (i in 1:nsim){
  RMSE_sbma[i]=mean(sqrt(MSE_sbma[,i+1]))
}

res_bm=(obs_test-fore_test)^2
resi_bm=cbind(col,res_bm)
MSE_bm=aggregate(resi_bm[,2],list(resi_bm[,1]),mean)
RMSE_bm=mean(sqrt(MSE_bm[,2]))
RMSE_bm
min(RMSE_sbma)
select=which.min(RMSE_sbma)
forecastsbma=foresbma[,select]

v=vector(mode="numeric",length=(ntest*8))
u=vector(mode="numeric",length=(ntest*8))
c1=vector(mode="numeric",length=(ntest*8))
for (i in 1:(ntest*8)){
  v[i]=quantile(foresbma[i,],probs=95/100)
  u[i]=quantile(foresbma[i,],probs=5/100)
  if (obs_test[i]<u[i]|obs_test[i]>v[i]){c1[i]=0}
  else c1[i]=1
}
coverage=(sum(c1)/(ntest*8))*100
coverage

nilai_CRPS=mean(EnsCrps(foresbma,obs_test))
nilai_CRPS

```

```

comp=cbind(datasbma$date.obs[-(1:(8*train))],datasbma$id.stat[-
(1:(8*train))],datasbma$lon.stat[-(1:(8*train))],datasbma$lat.stat[-
(1:(8*train))],obs_test,fore_test,forecastsbma,v,u)
write.csv(comp,"D:/kuliah s2/smt 3/Thesis/spaBMA/sbma_tmaks_nsim3.csv")

#for ensemble --> nsim=99# -----#
fore_sbma.tmaks=cbind(obs_test,foresbma)
write.csv(fore_sbma.tmaks,"D:/kuliah s2/smt 3/Thesis/spaBMA/fore_sbma.tmaks.csv")

#-----VRH-----#
library(SpecsVerification)
tmaks.dt=read.csv("D:/kuliah s2/smt 3/Thesis/spaBMA/fore_sbma.tmaks.csv",header=T)
T=30
ens=tmaks.dt[1:T,3:101]
obs=tmaks.dt[1:T,2]
rank.hist=Rankhist(ens,obs)

PlotRankhist(rank.hist,mode="prob.paper")

```

Keterangan : Syntax hanya untuk suhu maksimum

BIOGRAFI PENULIS



Niswatul Qona'ah, lahir di Bojonegoro pada tanggal 25 Juli 1994, anak ketiga dari 4 bersaudara dari pasangan Sumadi dan Munisri. Penulis menempuh jenjang pendidikan formal di TK Tunas Rimba Sukorejo-Bojonegoro, SDN Pacul III Bojonegoro, SMPN 2 Bojonegoro, dan SMAN 1 Bojonegoro, kemudian tahun 2012 ia diterima di Jurusan Statistika ITS melalui jalur Undangan - Bidik Misi. Penulis lulus dari program sarjana Statistika ITS pada tahun 2016 dengan Tugas Akhir berjudul “Analisis Pola Persebaran Infeksi Saluran Pernafasan Akut sebagai Dampak Industri Migas di Kabupaten Bojonegoro menggunakan *Spatial Pattern Analysis* dan *Flexibly Shaped Spatial Scan Statistic*”. Pada tahun 2016 pula, penulis diterima menjadi mahasiswa pasca sarjana di Statistika ITS melalui jalur *fresh graduate* ber beasiswa. Semasa kuliah penulis aktif di organisasi Forum Silaturrahim Mahasiswa Muslim Pasca Sarjana (FORSIMMPAS) ITS sebagai staff Departemen Syiar. Selain aktivitas kuliah dan organisasi di kampus, penulis juga tergabung sebagai salah satu pengajar matematika (honorar) di SMA IT Al-Uswah Surabaya. Apabila pembaca memiliki saran, kritik, atau ingin berdiskusi dengan penulis, dapat menghubungi melalui email: niswatulqonaah@gmail.com.

(Halaman ini sengaja dikosongkan)