

MODEL *OUTPUT STATISTICS* DENGAN *PRINCIPAL COMPONENT REGRESSION*, *PARTIAL LEAST SQUARE REGRESSION*, DAN *RIDGE REGRESSION* UNTUK KALIBRASI PRAKIRAAN CUACA JANGKA PENDEK

NISWATUL QONA'AH*, HASIH PRATIWI, YULIANA SUSANTI

*Universitas Sebelas Maret,
Jalan Ir. Sutami 36 Ketingan, Jebres, Surakarta, Jawa Tengah, Indonesia 57126
email : niswatulqonaah@staff.uns.ac.id, hasihpratiwi@staff.uns.ac.id,
yulianasusanti@staff.uns.ac.id*

Diterima 12 Juni 2021 Direvisi 22 Juni 2021 Dipublikasikan 26 Juli 2021

Abstrak. Penelitian ini merupakan upaya pengembangan *Model Output Statistics* (MOS) yang akan digunakan sebagai alat kalibrasi prakiraan cuaca jangka pendek. Informasi mengenai prakiraan cuaca yang akurat diharapkan dapat meminimalkan risiko kecelakaan yang disebabkan oleh cuaca, khususnya dalam bidang transportasi udara dan laut. Metode yang akan dikembangkan mencakup beberapa stasiun pengamatan cuaca di Indonesia. MOS merupakan sebuah metode berbasis regresi yang mengoptimalkan hubungan antara observasi cuaca dan luaran model *Numerical Weather Predictor* (NWP). Beberapa masalah yang muncul kaitannya dengan MOS adalah; mereduksi dimensi luaran NWP, mendapatkan variabel prediktor yang mampu menjelaskan variabilitas variabel respon, dan menentukan metode statistik yang sesuai dengan karakteristik data, sehingga dapat menggambarkan hubungan antara variabel respon dan variabel prediktor. Tujuan dari penelitian ini yaitu untuk mendapatkan pemodelan MOS yang sesuai untuk variabel respon suhu maksimum, suhu minimum, dan kelembapan udara. Metode regresi yang digunakan adalah *Principal Component Regression* (PCR), *Partial Least Square Regression* (PLSR), dan *ridge regression*. Selanjutnya, model MOS yang terbentuk divalidasi dengan kriteria *Root Mean Square Error* (RMSE) dan *Percentage Improval* (IM%). MOS mampu mengoreksi bias prakiraan NWP hingga lebih dari 50%. Berdasarkan RMSE terkecil pada penelitian ini, suhu maksimum lebih akurat diprakirakan menggunakan model PLSR, sementara suhu minimum dan kelembapan udara lebih akurat diprakirakan menggunakan *ridge regression*.

Kata Kunci: cuaca, MOS, NWP.

1. Pendahuluan

Informasi cuaca jangka pendek merupakan prakiraan cuaca untuk periode 1-7 hari kedepan. Bidang transportasi udara dan laut membutuhkan informasi kondisi cuaca

*penulis korespondensi

sebelum armada tersebut akan berangkat ke tempat tujuan. Informasi cuaca yang cepat dan akurat akan meminimalisir resiko kecelakaan akibat cuaca atau sering disebut dengan hidrometeorologi. Di Indonesia terdapat Badan Meteorologi Klimatologi dan Geofisika (BMKG) yang salah satu tugasnya adalah untuk membuat prakiraan cuaca harian. Unsur cuaca yang diramalkan diantaranya adalah suhu maksimum, suhu minimum, dan kelembapan udara. Namun, metode prakiraan yang saat ini digunakan masih bersifat subyektif atau bergantung pada kemampuan peramalnya.

Upaya untuk mengkaji pemodelan cuaca jangka pendek dengan pendekatan obyektif telah dilakukan oleh BMKG mulai tahun 2004 [1], [2]. Pendekatan yang dilakukan yaitu melalui pemodelan peramalan dengan memanfaatkan luaran *Numerical Weather Prediction* (NWP), sehingga meminimalisir unsur subjektivitas peramal dan diharapkan dapat meningkatkan akurasi hasil prakiraan. NWP adalah sekumpulan program komputer yang menggambarkan persamaan numerik atmosfer dalam waktu tertentu [3]. Negara-negara lintang tinggi sudah sejak lama mengembangkan prakiraan menggunakan luaran NWP, sementara di Indonesia masih tahap kajian. Saat ini telah tersedia banyak sumber data NWP dengan berbagai model, resolusi dan negara yang menghasilkannya, seperti CSIRO (Australia), GMS (Jepang), NCEP (USA), CMA (China), UKMO (UK).

Wilayah Indonesia dengan topografi dan vegetasi yang kompleks mengakibatkan luaran model NWP menjadi bias karena pendekatannya yang masih secara global. Untuk mengoptimalkan pemanfaatan luaran model NWP perlu dilakukan pra-pemrosesan sebelum digunakan sebagai prediktor dalam prakiraan cuaca. *Model Output Statistics* (MOS) merupakan salah satu metode yang seringkali digunakan dalam pra-pemrosesan NWP. Verifikasi MOS yang dikembangkan oleh BMKG sejak 2004 menunjukkan bahwa prakiraan suhu maksimum, suhu minimum, dan kelembapan udara masih belum memuaskan [1], [2]. Oleh karena itu, perlu pengembangan lebih lanjut dalam pemrosesan NWP menggunakan MOS.

MOS adalah metode prakiraan cuaca berbasis regresi yang memanfaatkan observasi cuaca sebagai variabel respon dan luaran NWP sebagai variabel prediktor. Tujuan dari metode ini adalah untuk mencari skala lokal dari informasi skala global melalui hubungan fungsional antara kedua skala tersebut, sehingga dapat mengurangi bias dari luaran NWP [4]. Penelitian sebelumnya [3], mengembangkan model MOS melalui *Projection Pursuit Regression* pada peramalan suhu udara minimum, suhu udara maksimum, dan kelembapan. Hasilnya, model MOS terbukti lebih akurat dibandingkan model NWP dan mampu mengoreksi bias sebesar 86%.

Permasalahan dalam MOS adalah kondisi data NWP yang memiliki dimensi besar, karena diukur dalam skala global (grid yang luas). Hal ini mengakibatkan adanya multikolinearitas yang tinggi antar variabel NWP. Beberapa metode regresi yang dapat mengatasi kasus multikolinearitas diantaranya adalah *Principal Component Regression* (PCR), *Partial Least Square Regression* (PLSR), dan *Ridge Regression*. Penelitian ini akan mengembangkan model MOS melalui pendekatan ketiga metode regresi tersebut. Hasil penelitian diharapkan mampu mendapatkan model MOS terbaik yang dapat mengoreksi bias dari prakiraan cuaca, sehingga diperoleh prakiraan kondisi cuaca yang lebih tajam dan akurat.

2. Landasan Teori

2.1. Numerical Weather Prediction (NWP)

Numerical Weather Prediction (NWP) merupakan kode pemrograman yang merepresentasikan persamaan-persamaan atmosfer secara numerik dan perubahan keadaan atmosfer pada satuan waktu tertentu menggunakan kemampuan komputer yang tinggi [5]. Prakiraan cuaca diperoleh dari penyelesaian persamaan perubahan atmosfer. NWP diukur pada domain skala global, sehingga informasi prakiraan cuaca yang dihasilkan menjadi bias pada daerah dengan skala lokal. Selain itu, luaran NWP juga bersifat deterministik, sehingga tidak mampu menjelaskan cuaca secara stokastik. Oleh karena itu, perlu diterapkan pemrosesan secara statistik (*statistical post-processing*) supaya dapat menjelaskan ketidakpastian tersebut [6].

2.2. Model Output Statistics (MOS)

Model Output Statistics (MOS) merupakan salah satu metode *statistical post-processing* untuk data NWP. MOS merupakan model berbasis regresi yang menganalisis hubungan antara variabel respon (hasil observasi cuaca) dengan variabel prediktor (luaran NWP) [3]. Hasil observasi cuaca yang digunakan sebagai respon berupa suhu maksimum, suhu minimum, dan kelembapan udara. Sementara luaran NWP yang dijadikan sebagai prediktor ada sebanyak 18 parameter NWP yang dikur pada grid dan level yang beragam. Secara umum, MOS dapat dituliskan dalam persamaan berikut.

$$\hat{y}_t = f_{MOS}(x_t) \quad (2.1)$$

\hat{y}_t adalah prakiraan cuaca pada saat t dan x_t adalah variabel-variabel luaran NWP pada saat t .

MOS memiliki dua fungsi utama. Pertama, MOS menghasilkan prakiraan cuaca kuantitatif kedepan dan mungkin tidak secara eksplisit diperoleh dari model. Kedua, MOS mereduksi rata-rata *error* dari prakiraan NWP dengan memperkecil bias dan mengoreksi model secara statistik.

2.3. Principal Component Regression (PCR)

Principal Component Regression (PCR) adalah salah satu metode yang dapat mengatasi kasus multikolinearitas tanpa harus mengeluarkan prediktor yang terlibat dalam hubungan kolinear. Dengan demikian, seluruh prediktor berkontribusi dalam model. Konsep utama dari PCR adalah 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 [7].

Misalkan $x' = [x_1, x_2, \dots, x_p]$ merupakan vektor dari matrik variabel prediktor X yang berukuran $n \times p$. Jika A merupakan matrik orthogonal berukuran $p \times p$

dengan kolom ke- m berisikan vektor eigen ke- m dari matrik $X'X$ dengan asumsi $m < p$, maka skor PC untuk tiap pengamatan seperti pada persamaan berikut.

$$Z = XA \quad (2.2)$$

dengan elemen $(i,)$ dari Z merupakan skor PC ke- m untuk pengamatan ke- i dimana $i = 1, 2, 3, \dots, n$. Berdasarkan sifat orthogonal A dimana $AA' = I$, regresi linear berganda dengan prediktor saling berkorelasi dapat dikonversi dalam PCR seperti pada persamaan berikut.

$$y = X\beta + \varepsilon = Z\gamma + \varepsilon \quad (2.3)$$

dimana $\gamma = A'\beta$ [8].

2.4. Partial Least Square Regression (PLSR)

Partial Least Square Regression (PLSR) merupakan metode yang memodelkan hubungan antara variabel respon dengan variabel prediktor berdasarkan laten secara simultan [9]. PLSR memiliki kemiripan dengan PCR yaitu membentuk matrik komponen laten T pada persamaan (2.4) yang berukuran $n \times c$ sebagai transformasi linear dari matrik prediktor X .

$$T = XW \quad (2.4)$$

dimana W adalah matrik pembobot berukuran $p \times c$. Indeks n menyatakan banyaknya observasi, p menyatakan banyaknya prediktor, dan c menyatakan banyaknya komponen laten. Komponen laten T berperan sebagai variabel random X yang digunakan untuk mengestimasi respon Y yang berukuran $q \times c$, dimana q menyatakan banyaknya respon. Setelah T terbentuk, maka Q' diperoleh melalui metode kuadrat terkecil berdasarkan persamaan berikut.

$$Q' = (T'T)^{-1}T'Y. \quad (2.5)$$

Berdasarkan [9], salah satu cara untuk memodelkan Y adalah sebagai berikut,

$$Y = XB + F. \quad (2.6)$$

Selanjutnya, dengan menerapkan substitusi yang melibatkan persamaan (2.6), maka diperoleh pembobot B pada persamaan berikut.

$$XB + F = TQ' + F \quad (2.7)$$

$$XB = XWQ' \quad (2.8)$$

$$B = WQ' = W(T'T)^{-1}T'Y. \quad (2.9)$$

Dengan demikian diperoleh persamaan berikut untuk mengestimasi respon Y .

$$\hat{Y} = X\hat{B} = [TW^{-1}W(T'T)^{-1}T'Y] = [T(T'T)^{-1}T'Y]. \quad (2.10)$$

2.5. Ridge Regression

Ridge regression merupakan salah satu metode yang digunakan untuk mereduksi pengaruh dari prediktor yang saling berkorelasi. Berdasarkan [8], ridge regression menggunakan konstanta non-negatif λ untuk menghitung koefisien regresi yang lebih efisien, sekaligus mereduksi singularitas akibat multikolinearitas. Koefisien ridge regression dihitung dengan persamaan berikut.

$$\hat{\beta}_{ridge} = (X^{-1}X + \lambda I)^{-1}X'y. \tag{2.11}$$

Dengan demikian, diperoleh model estimasi variabel respon pada persamaan berikut.

$$\hat{y}_{ridge} = X\hat{\beta}_{ridge}. \tag{2.12}$$

Konstanta λ dapat dipilih secara intuitif (visual) atau menggunakan teknik *cross-validation* berdasarkan kriteria *Root Mean Square Error* (RMSE) terendah. Semakin tinggi nilai λ , maka koefisien β_p semakin mendekati 0 atau mengakibatkan parameter p kurang berpengaruh terhadap variabel respon.

2.6. Evaluasi Model Kalibrasi

Kebaikan model tidak hanya dievaluasi melalui eror prakiraan yang diperoleh, namun juga koreksi bias prakiraan terhadap observasi [10]. Berdasarkan [11], *Root Mean Square Error* (RMSE) adalah indikator yang cukup untuk menilai kualitas prakiraan. RMSE merupakan salah satu indikator yang digunakan untuk mengevaluasi kebaikan prakiraan cuaca. RMSE merupakan akar kuadrat dari MSE, yaitu rata-rata jumlah kuadrat dari selisih antara observasi dengan nilai prakiraan. Formula untuk menghitung RMSE dituliskan pada persamaan berikut.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{2.13}$$

dimana n adalah banyaknya observasi. Nilai *RMSE* yang semakin kecil menunjukkan bahwa model semakin baik.

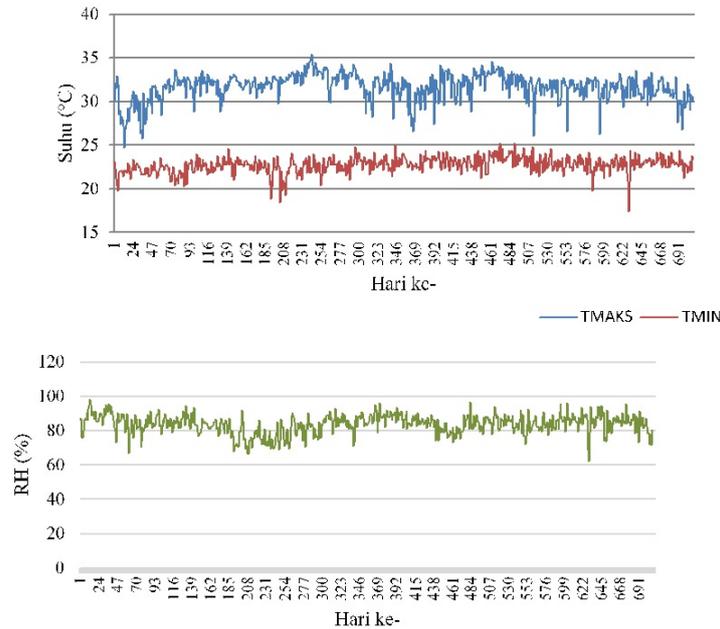
Ukuran pengoreksian bias diperlukan untuk mengetahui sejauh mana MOS dapat mengoreksi bias dari model NWP. Besarnya koreksi bias dapat diukur dengan *percentage improval* (%IM) dengan formula pada persamaan berikut.

$$\%IM = \frac{(RMSE_{NWP} - RMSE_{MOS})}{RMSE_{NWP}} \times 100\% \tag{2.14}$$

3. Pembahasan

3.1. Deskripsi dan Preprocessing Data

Sebelum memulai mengkalibrasi prakiraan cuaca menggunakan MOS, dilakukan analisis deskripsi secara visual untuk mendapatkan gambaran umum dari pola cuaca, yaitu suhu udara maksimum (TMAKS), suhu udara minimum (TMIN), dan kelembapan udara (RH).



Gambar 1. Observasi Cuaca di Stasiun Meteorologi Dermaga

Gambar 1 menunjukkan tren pergerakan suhu maksimum, suhu minimum, dan kelembapan udara di stasiun meteorologi Dermaga selama 708 hari pengamatan. Secara umum, pengamatan suhu maksimum, suhu minimum, dan kelembapan udara di stasiun meteorologi Dermaga tidak menunjukkan adanya gejala yang menyimpang cukup jauh, meskipun pada periode tertentu ketiga nilai observasi terlihat lebih rendah daripada periode sebelumnya.

Setelah menganalisis secara deskriptif terhadap observasi cuaca (suhu maksimum, suhu minimum, dan kelembapan udara), selanjutnya dilakukan standarisasi terhadap ketiga data observasi cuaca tersebut yang akan digunakan sebagai respon dan luaran NWP sebagai prediktor. Standarisasi dilakukan berdasarkan metode *scaled and centered*, yaitu nilai pengamatan dikurangi rata-rata kemudian dibagi dengan standar deviasi. Hal ini dilakukan untuk meminimalkan perbedaan skala antar variabel penelitian.

Setelah itu, dilakukan reduksi dimensi terhadap data luaran NWP menggunakan *Principal Component Analysis* (PCA). Luarannya NWP terdiri dari 32 parameter yang masing-masing diukur pada 9 grid, disajikan pada Tabel 1.

Setelah direduksi menggunakan PCA dengan kriteria nilai eigen > 1 dan kumulatif varians minimal 80%, didapatkan 41 komponen yang merepresentasikan seluruh parameter NWP di stasiun Dermaga, seperti yang disajikan pada Tabel 2.

3.2. Pembentukan Model Output Statistics (MOS)

Setelah dilakukan preprocessing terhadap data respon (observasi cuaca) dan data prediktor (luaran NWP), langkah selanjutnya adalah mendapatkan kalibrasi prakti-

Tabel 1. Luaran NWP

Parameter NWP (kode)	Level	X	Satuan
Surface Pressure Tendency (dpsdt)	Permukaan	X ₁	hPa
Water Mixing Ratio (mix)	1, 2, 4	X ₂ , X ₃ , X ₄	g/kg
Vertical Velocity (omega)	1, 2, 4	X ₅ , X ₆ , X ₇	Knot
PBL depth (pblh)	Permukaan	X ₈	Meter
Surface Pressure (ps)	Permukaan	X ₉	hPa
Mean Sea Level Pressure (psl)	Permukaan	X ₁₀	hPa
Screen Mixing Ratio (qgscm)	Permukaan	X ₁₁	g/kg
Relative Humidity (rh)	1, 2, 4	X ₁₂ , X ₁₃ , X ₁₄	%
Precipitation (rnd)	Permukaan	X ₁₅	Mm
Temperature	1, 2, 4	X ₁₆ , X ₁₇ , X ₁₈	Celcius
Maximum Screen Temperature (tmaxscr)	Permukaan	X ₁₉	Celcius
Maximum Screen Temperature (tminscr)	Permukaan	X ₂₀	Celcius
Pan Temperature (tpan)	Permukaan	X ₂₁	Celcius
Screen Temperature (tscm)	Permukaan	X ₂₂	Celcius
Zonal Wind (u)	1, 2, 4	X ₂₃ , X ₂₄ , X ₂₅	Knot
Friction Velocity (ustar)	Permukaan	X ₂₆	m/sec
Meridional Wind (v)	1, 2, 4	X ₂₇ , X ₂₈ , X ₂₉	Knot
Geopotential Height (zg)	1, 2, 4	X ₃₀ , X ₃₁ , X ₃₂	Meter

Tabel 2. Hasil PCA Luaran NWP di Stasiun Meteorologi 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%	tscm	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%
md	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%

raan cuaca menggunakan MOS berdasarkan PLSR, PCR, dan *ridge regression*. Sebelum memodelkan PLSR dan PCR, dilakukan penentuan banyak komponen yang optimal menggunakan emphCross Validation (CV). Sementara untuk *ridge regression*, dilakukan pemilihan koefisien penalti λ menggunakan metode grafis.

3.2.1. MOS Berdasarkan PLSR

Banyak komponen optimal model PLSR ditentukan oleh nilai CV terendah. Pada stasiun meteorologi Dermaga banyak komponen optimal untuk suhu maksimum

Tabel 3. Banyak Komponen Optimal Model PLSR pada 8 Stasiun Meteorologi

Banyak Komponen			CV terendah		
TMAKS	TMIN	RH	TMAKS	TMIN	RH
21	4	12	0,6411	0,8809	0,7708

Tabel 4. Koefisien Regresi Model PLSR

Parameter	TMAKS	TMIN	RH	Parameter	TMAKS	TMIN	RH
PC.dpsdt	-0.271	0.003	-0.045	PC.temp2	0.639	-0.043	-0.021
PC.mixr1	0.331	0.093	-0.004	PC.temp4	-0.007	-0.094	-0.114
PC.mixr2	0.291	-0.077	-0.111	PC.tmaxscr	-0.527	0.060	0.608
PC.mixr4	0.132	0.054	-0.006	PC.tminscr	-0.147	-0.073	0.072
PC1.omega1	0.354	0.012	-0.057	PC.tpan	0.089	-0.005	-0.280
PC2.omega1	0.372	0.005	-0.118	PC.tscrm	-0.141	-0.021	-0.008
PC1.omega2	-0.147	0.006	0.074	PC.u1	-0.400	-0.002	0.141
PC2.omega2	-0.022	0.008	0.139	PC.u2	-0.383	0.014	0.023
PC1.omega4	-0.025	-0.008	-0.015	PC.u4	0.175	0.023	0.103
PC2.omega4	-0.107	0.044	-0.026	PC1.ustar	0.029	0.051	-0.029
PC.pblh	-0.013	0.048	0.246	PC2.ustar	0.039	-0.033	0.053
PC.ps	0.018	0.020	-0.010	PC1.v1	-0.118	-0.008	-0.068
PC.psl	0.024	0.023	-0.004	PC2.v1	-0.030	-0.001	-0.010
PC1.qgscrm	-0.001	-0.089	0.056	PC1.v2	-0.053	-0.007	0.122
PC2.qgscrm	-0.054	0.003	0.062	PC2.v2	0.104	-0.007	-0.031
PC1.rh1	0.084	-0.060	-0.012	PC.v4	-0.005	-0.027	-0.013
PC2.rh1	-0.063	-0.059	0.008	PC.zg1	-0.286	0.020	0.090
PC.rh2	0.112	-0.045	-0.137	PC.zg2	0.079	0.012	0.041
PC.rh4	0.078	-0.009	0.020	PC1.zg4	0.010	-0.059	0.024
PC.md	0.068	-0.011	-0.039	PC2.zg4	0.235	0.060	0.054
PC.temp1	-0.293	-0.029	-0.114				

(TMAKS), suhu minimum (TMIN), dan kelembapan udara (RH) masing-masing adalah 21, 4, dan 12. Model PLSR untuk suhu maksimum diindikasikan optimal jika menggunakan komponen yang relatif banyak dibanding suhu minimum dan kelembapan udara. Banyaknya komponen optimal ini mengindikasikan bahwa kondisi cuaca yang memengaruhi suhu maksimum di stasiun Dermaga tidak saling berhubungan antar parameter, misal arah angin dan curah hujan hampir saling independen. Nilai CV terendah dan banyak komponen optimal untuk stasiun Dermaga, disajikan pada Tabel 3.

Setelah didapatkan banyak komponen yang optimal yang berkontribusi pada model, selanjutnya dilakukan pemodelan PLSR untuk prakiraan TMAKS, TMIN, dan RH. Koefisien regresi model PLSR untuk TMAKS, TMIN, dan RH di stasiun Dermaga disajikan dalam Tabel 4.

Koefisien regresi pada Tabel 4 merupakan koefisien yang terstandarkan sehingga lebih mudah membandingkan pengaruh masing-masing parameter NWP terhadap model. Model PLSR pada suhu minimum cenderung didominasi oleh nilai koefisien regresi yang kecil (mendekati 0). Hal ini mengindikasikan bahwa pengaruh tiap parameter NWP terhadap model suhu minimum relatif lebih seimbang.

Tabel 5. Koefisien Regresi Model PCR

Parameter	TMAKS	TMIN	RH	Parameter	TMAKS	TMIN	RH
PC.dpsdt	0.0308	-0.0094	-0.014	PC.temp2	-0.0551	-0.0382	0.000
PC.mixr1	0.0216	-0.0856	0.074	PC.temp4	-0.0320	-0.0554	-0.033
PC.mixr2	0.0067	-0.0619	-0.042	PC.tmaxscr	-0.0684	0.0020	0.052
PC.mixr4	0.0078	-0.0627	0.043	PC.tminscr	-0.0579	-0.0558	-0.002
PC1.omega1	0.0152	-0.0036	-0.001	PC.tpan	-0.0771	-0.0289	0.037
PC2.omega1	0.0415	0.0063	-0.012	PC.tscrn	-0.0687	-0.0138	0.034
PC1.omega2	0.0196	-0.0016	-0.002	PC.u1	0.0714	0.0284	0.030
PC2.omega2	0.0388	0.0042	-0.018	PC.u2	0.0703	0.0313	-0.027
PC1.omega4	0.0086	-0.0149	-0.005	PC.u4	0.0672	0.0190	-0.032
PC2.omega4	0.0077	0.0047	-0.009	PC1.ustar	-0.0087	0.0221	0.034
PC.pblh	-0.0496	0.0471	0.079	PC2.ustar	-0.0216	-0.0013	0.015
PC.ps	0.0169	-0.0184	-0.046	PC1.v1	0.0057	0.0038	-0.000
PC.psl	-0.0163	0.0167	-0.043	PC2.v1	-0.0043	0.0220	0.031
PC1.qgscrn	0.0093	-0.0815	-0.061	PC1.v2	0.0040	0.0014	-0.002
PC2.qgscrn	-0.0015	-0.0139	-0.013	PC2.v2	0.0130	-0.0236	-0.038
PC1.rh1	-0.0573	0.0588	-0.079	PC.v4	-0.0127	-0.0198	-0.015
PC2.rh1	0.0151	-0.0283	-0.046	PC.zg1	-0.0311	0.0175	0.059
PC.rh2	0.0319	-0.0343	-0.036	PC.zg2	-0.0394	0.0128	0.061
PC.rh4	0.0207	-0.0310	-0.025	PC1.zg4	-0.0625	-0.0445	0.018
PC.md	0.0564	-0.0062	-0.039	PC2.zg4	-0.0012	0.0455	0.058
PC.temp1	-0.0688	-0.0191	0.029				

3.2.2. MOS Berdasarkan PCR

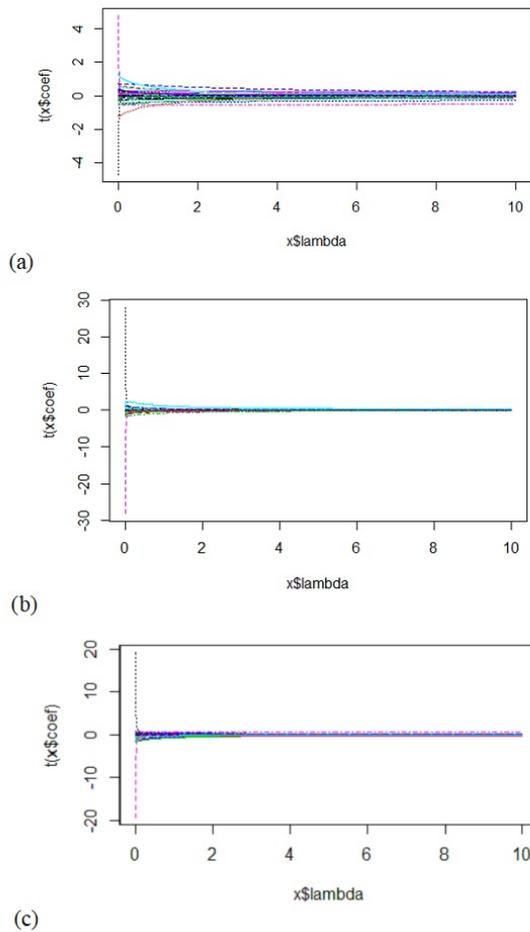
Sama seperti pada PLSR, langkah yang dilakukan sebelum membentuk MOS berdasarkan PCR adalah menentukan banyak komponen optimal dari 41 komponen di stasiun Dermaga. Namun, PCR memilih komponen optimal berdasarkan kemampuan merepresentasikan variabilitas prediktor minimal 80%. Persentase kumulatif varians yang merepresentasikan variabilitas prediktor model PCR suhu maksimum, suhu minimum, dan kelembapan udara adalah sama karena proses perhitungannya tidak melibatkan variabel respon. Banyak komponen optimal yang diperoleh adalah sebanyak 8 komponen yang mampu menjelaskan variabilitas prediktor sebesar 81,81%.

Setelah didapatkan banyak komponen yang optimal yang berkontribusi pada model, selanjutnya dilakukan pemodelan PCR untuk prakiraan TMAKS, TMIN, dan RH. Koefisien regresi model PCR TMAKS, TMIN, dan RH di stasiun Dermaga disajikan dalam Tabel 5.

Tabel 5 menunjukkan bahwa parameter model PCR baik untuk TMAKS, TMIN, dan RH lebih banyak didominasi oleh koefisien yang relatif kecil (mendekati 0). Hal ini mengindikasikan bahwa kontribusi tiap parameter NWP terhadap model PCR lebih rendah dibanding model PLSR. Dengan demikian, terdapat kemungkinan bahwa kondisi tersebut akan memengaruhi akurasi hasil prakiraan cuaca.

3.2.3. MOS Berdasarkan Ridge Regression

Setelah mendapatkan model PLSR dan PCR, selanjutnya dilakukan pembentukan MOS berdasarkan metode lain yang mampu mengatasi multikolinearitas yaitu ridge regression. Metode ini tidak memerlukan pemilihan komponen optimal, namun menggunakan suatu konstanta λ untuk meminimalkan dampak singularitas $X'X$.



Gambar 2. Konvergensi Koefisien *Ridge Regression*, (a) suhu maksimum, (b) suhu minimum, (c) kelembapan udara

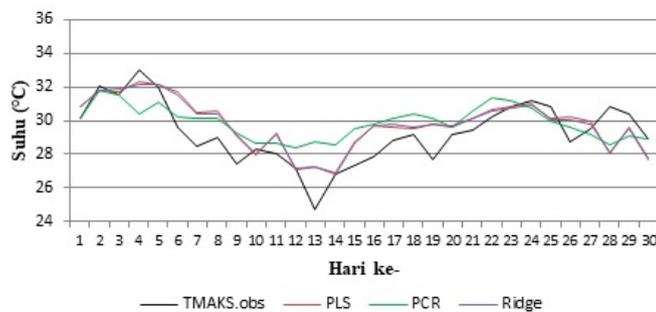
Nilai konstanta λ diharapkan tidak terlalu besar, karena bisa mengakibatkan nilai koefisien regresi yang semakin mendekati 0. Kondisi ini juga akan menyebabkan nilai estimasi respon mendekati 0. Penentuan nilai konstanta λ secara visual dapat dibantu oleh Gambar 2.

Penentuan nilai λ secara visual seharusnya dihindari karena bersifat subjektif. Akan tetapi, dalam kasus ini λ tetap tidak konvergen meski sudah mendekati 100. Oleh karena itu, diputuskan untuk menentukan nilai λ secara visual. Gambar 2 mengindikasikan bahwa koefisien regresi untuk TMAKS, TMIN, dan RH konvergen ketika nilai λ sebesar 9 atau lebih. Koefisien *ridge regression* untuk stasiun Dermaga dengan $\lambda = 9$ ditunjukkan pada Tabel 6.

Berdasarkan Tabel 6 dapat dikatakan bahwa model *ridge regression* untuk TMIN tidak lagi didominasi oleh koefisien regresi yang mendekati 0. Hal ini

Tabel 6. Koefisien Regresi Model Ridge Regression

Parameter	T _{MAKS}	T _{MIN}	RH	Parameter	T _{MAKS}	T _{MIN}	RH
PC.dpsdt	-0,075	-0,099	-0,008	PC.temp2	0,256	0,066	-0,072
PC.mixr1	0,123	-0,125	-0,001	PC.temp4	0,008	-0,219	-0,143
PC.mixr2	0,102	-0,129	-0,059	PC.tmaxscr	-0,513	0,149	0,574
PC.mixr4	0,030	0,008	-0,040	PC.tminscr	-0,108	-0,084	0,052
PC1.omega1	0,153	0,165	-0,045	PC.tpan	0,053	-0,026	-0,224
PC2.omega1	0,212	0,015	-0,173	PC.tscrn	-0,099	-0,010	0,004
PC1.omega2	-0,099	-0,051	0,040	PC.u1	-0,301	0,055	0,132
PC2.omega2	-0,029	-0,038	0,154	PC.u2	-0,236	0,016	0,052
PC1.omega4	-0,017	0,083	-0,008	PC.u4	0,141	0,124	0,087
PC2.omega4	-0,082	0,031	-0,029	PC1.ustar	0,048	0,014	-0,041
PC.pblh	-0,041	0,010	0,225	PC2.ustar	0,007	-0,088	0,058
PC.ps	0,024	0,037	-0,016	PC1.v1	-0,012	-0,075	-0,125
PC.psl	0,027	0,066	-0,004	PC2.v1	0,012	-0,049	0,039
PC1.qgscm	-0,005	-0,153	0,109	PC1.v2	-0,094	0,010	0,169
PC2.qgscm	-0,067	0,054	0,059	PC2.v2	0,089	0,053	-0,018
PC1.rh1	0,027	-0,021	-0,115	PC.v4	-0,032	-0,021	-0,032
PC2.rh1	-0,069	-0,172	-0,001	PC.zg1	-0,096	0,004	0,143
PC.rh2	0,031	-0,261	-0,206	PC.zg2	0,022	-0,166	-0,034
PC.rh4	0,037	0,009	-0,024	PC1.zg4	0,001	0,022	0,095
PC.md	0,071	-0,022	-0,041	PC2.zg4	0,079	0,290	-0,001
PC.temp1	-0,108	-0,083	-0,196				



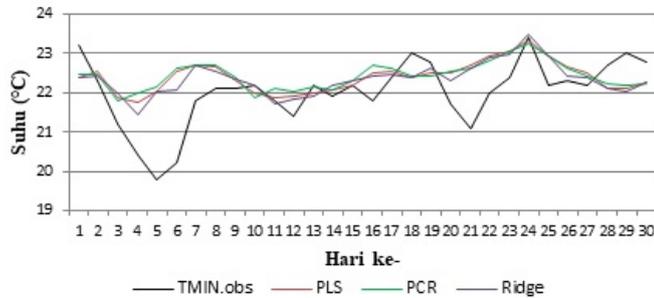
Gambar 3. Tren Prakiraan MOS dan Observasi Suhu Maksimum

mengindikasikan bahwa pengaruh parameter NWP terhadap model TMIN menggunakan *ridge regression* lebih signifikan dibanding PLSR dan PCR.

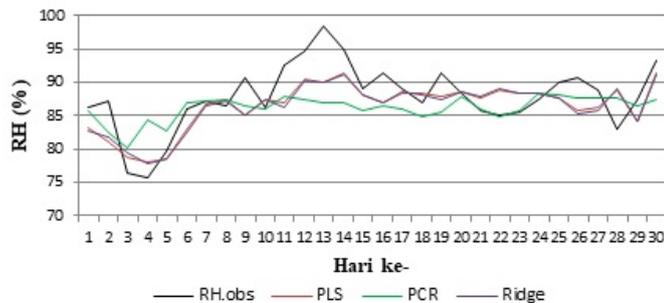
3.2.4. Deskripsi Hasil Kalibrasi Prakiraan Cuaca menggunakan MOS

Setelah mendapatkan model untuk kalibrasi prakiraan cuaca, maka langkah selanjutnya adalah membandingkan hasil prakiraan tiap model terhadap nilai observasi. Gambar 3, Gambar 4, dan Gambar 5 merupakan plot runtun waktu 30 hari pertama dari hasil prakiraan dan observasi cuaca stasiun Dermaga, masing-masing untuk suhu maksimum, suhu minimum, dan kelembapan udara.

Gambar 3 - Gambar 5 mengindikasikan bahwa tiap model MOS baik PLSR, PCR, maupun *ridge regression* telah mampu mengikuti pola umum dari observasi



Gambar 4. Tren Prakiraan MOS dan Observasi Suhu Minimum



Gambar 5. Tren Prakiraan MOS dan Observasi Kelembapan Udara

cuaca yaitu suhu maksimum, suhu minimum, dan kelembapan udara. Apabila tren observasi cuaca naik maka hasil prakiraan turut meningkat, begitu pula apabila observasi cuaca turun. Gambar 3 - Gambar 5 juga menunjukkan bahwa model PLSR dan *ridge regression* lebih mampu menghasilkan prakiraan yang mengikuti fluktuasi observasi cuaca daripada PCR. Hal ini memungkinkan bahwa model PLSR dan *ridge regression* memiliki akurasi prakiraan yang lebih tinggi daripada PCR. Selanjutnya, untuk mengetahui akurasi prakiraan dari masing-masing model dilakukan evaluasi menggunakan kriteria RMSE dan %IM.

3.2.5. *Evaluasi Model Kalibrasi Prakiraan Cuaca*

Langkah terakhir dalam analisis penelitian ini adalah evaluasi model kalibrasi prakiraan cuaca. Evaluasi ini dilakukan untuk mengetahui kebaikan prakiraan cuaca yang dihasilkan oleh MOS berdasarkan PLSR, PCR, dan *ridge regression*. Kriteria evaluasi yang digunakan adalah RMSE dan *percentage improval* (%IM) yang masing-masing dihitung melalui persamaan (2.13) dan (2.14). Tabel 7 menunjukkan hasil evaluasi kebaikan MOS dalam mengkalibrasi prakiraan cuaca.

Berdasarkan Tabel 7 dapat diketahui bahwa model MOS berdasarkan PLSR, PCR, dan *ridge regression* secara konsisten memiliki nilai RMSE yang lebih kecil dibandingkan dengan RMSE NWP. Tabel 7 juga menunjukkan bahwa suhu mak-

Tabel 7. Evaluasi Keباikan Model

Kriteria	Prakiraan	TMAKS	TMIN	RH
RMSE	NWP	2,745	2,010	13,626
	PLSR	0,939*	0,781	4,152
	PCR	1,148	0,809	4,747
	<i>Ridge Regression</i>	0,945	0,774*	4,132*
%IM	PLSR	65,79*	61,14	69,53
	PCR	58,18	59,75	65,16
	<i>Ridge Regression</i>	65,57	61,49*	69,68*

*Model terbaik dengan kriteria RMSE terendah dan %IM tertinggi

simum lebih akurat diprakirakan menggunakan model PLSR dengan RMSE terendah sebesar 0,939C. Sementara suhu minimum dan kelembapan udara lebih akurat diprakirakan menggunakan model *ridge regression* dengan nilai RMSE masing-masing sebesar 0,774C dan 4,132%. Model PCR memberikan hasil prakiraan cuaca yang kurang akurat dibanding PLSR dan *ridge regression* dengan nilai RMSE yang lebih tinggi. Hal ini juga telah diindikasikan pada Gambar 3-5, dimana tren prakiraan PCR tidak mengikuti fluktuasi observasi cuaca. Model MOS berdasarkan ketiga metode regresi yang digunakan mampu mengoreksi bias lebih dari 50%. Nilai %IM paling tinggi dimiliki oleh model *ridge regression* yang mampu mengoreksi bias hingga sebesar 69,68% dalam memprakirakan kelembapan udara.

4. Kesimpulan

Hasil evaluasi MOS berdasarkan PLSR, PCR, dan *ridge regression* menunjukkan bahwa ketiga model tersebut terbukti lebih akurat dari model NWP, dengan nilai RMSE yang konsisten lebih kecil. MOS mampu mengoreksi bias prakiraan NWP hingga lebih dari 50%. Model *ridge regression* mampu mengoreksi bias prakiraan NWP untuk kelembapan udara hingga sebesar 69,68%. Suhu maksimum lebih akurat diprakirakan menggunakan model PLSR karena memiliki nilai RMSE terkecil yaitu sebesar 0,939C. Sementara suhu minimum dan kelembapan udara lebih akurat diprakirakan menggunakan *ridge regression* dengan nilai RMSE untuk suhu minimum sebesar 0,774C dan untuk kelembapan udara sebesar 4,132%.

5. Ucapan Terima kasih

Penulis mengucapkan terima kasih kepada Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Sebelas Maret, Badan Meteorologi, Klimatologi, dan Geofisika yang telah memberikan dukungan perihal materi dan data yang dibutuhkan, serta semua pihak yang telah membantu, memberikan saran dan masukan sehingga paper ini dapat diselesaikan dengan baik.

Daftar Pustaka

- [1] BMKG, *Laporan Kegiatan Pengembangan Model Output Statistik (MOS) untuk Pemodelan Prakiraan Cuaca Jangka Pendek*, Jakarta, 2005.

- [2] BMKG, *Laporan Kegiatan Uji Operasionalisasi dan Validasi Model Output Statistik (MOS)*, Jakarta, 2006.
- [3] Safitri, R., Model Output Statistics dengan Projection Pursuit Regression untuk Meramalkan Suhu Minimum, Suhu Maksimum, dan Kelembapan, *Jurnal Sains dan Seni ITS*, vol.1, no.1, 2012
- [4] Wigena, A. H., Permodelan Statistical Downscaling dengan Regresi Projection Pursuit untuk Peramalan Curah Hujan Bulanan (Kasus Curah hujan bulanan di Indramayu), <https://123dok.com/document/dzx87rvq-permodelan-statistical-downscaling-regresi-projection-peramalan-bulanan-indramayu.html> (accessed Jun. 15, 2021).
- [5] Idowu, O. S. and Rautenbach, C. J., Model Output Statistics to improve severe storms prediction over Western Sahel, *Atmos. Res.*, 2008, doi: 10.1016/j.atmosres.2008.10.035.
- [6] Wilks, D. S., *Statistical Methods in the Atmospheric Sciences Second Edition*, vol. 14, no. 2. London: Elsevier, 2007.
- [7] Jolliffe, I. T., Principal Component Analysis, Second Edition, *Encycl. Stat. Behav. Sci.*, vol. 30, no. 3, p. 487, 2002, doi: 10.2307/1270093.
- [8] Draper, N. R. and Smith, H., *Applied Regression Analysis Third Edition*, John Wiley & Sons, Inc., New York, 1998.
- [9] Wold, S., Sjstrm, M., and Eriksson, L., PLS-regression: A basic tool of chemometrics, in *Chemometrics and Intelligent Laboratory Systems*, Oct. 2001, vol. 58, no. 2, pp. 109130, doi: 10.1016/S0169-7439(01)00155-1.
- [10] Feldmann, K. and Thorarinsdottir, T., Statistical Postprocessing of Ensemble Forecasts for Temperature: The Importance of Spatial Modeling, 2012, [Online]. Available: <http://www.nr.no/thordis/files/Feldmann2012.pdf>.
- [11] Mller, A., *Multivariate and spatial ensemble postprocessing methods submitted to the Combined Faculties for the Natural Sciences and for Mathematics of the Ruperto-Carola University of Heidelberg , Germany for the degree of Doctor of Natural Sciences put forward by*, no. July. Germany: Ruperto-Carola University of Heidelberg, 2018.