

# Meningkatkan Kualitas Prediksi Curah Hujan Musiman saat Fase ENSO Menggunakan Metode *Bayesian Model Averaging* (BMA), Studi Kasus: Pulau Jawa

## *Improving Quality of Seasonal Rainfall Prediction during ENSO Phase using Bayesian Model Averaging (BMA) method, case study: Java Island*

Robi Muharsyah\*, Dian Nur Ratri

Pusat Informasi Perubahan Iklim, Badan Meteorologi Klimatologi dan Geofisika, Indonesia

### INFORMASI ARTIKEL

#### Riwayat artikel:

Diterima: 29 November 2020  
Disetujui: 3 Februari 2021  
Dipublikasi online: 7 Februari 2021

#### Kata Kunci:

Prediksi curah hujan musiman  
ECMWF System 4  
Post-processing statistik  
Reliabilitas

#### Keywords:

Seasonal rainfall prediction  
ECMWF System 4  
Statistical post-processing  
Reliability

#### Direview oleh:

Yayan Apriyana, Elza Surmaini

**Abstrak.** Ketika fase *El Niño Southern Oscillation* (ENSO), *El Niño* atau *La Niña* terjadi, Curah Hujan (CH) musiman di Pulau Jawa cenderung mengalami kondisi di Bawah Normal (BN) atau Atas Normal (AN). Oleh karena itu, prediksi CH musiman yang akurat menjadi sangat penting pada kedua fase tersebut. Model *ECMWF System 4* (S4) adalah salah satu model prediksi musim yang dapat dipakai untuk menghasilkan prediksi probabilistik kejadian BN atau AN. Akan tetapi, kualitas dari keluaran langsung (RAW) model S4 masih buruk. Metode *Bayesian Model Averaging* (BMA) dipilih sebagai *post-processing* statistik untuk memperbaiki kualitas tersebut. *Probability Density Function* (PDF) prediktif BMA mampu menghasilkan prediksi deterministik dan probabilistik lebih akurat dari RAW model S4. Tingkat akurasi tersebut diketahui dari *Root Mean Square Error* (RMSE) dan *Brier Score* (BS) BMA lebih rendah dari RAW; *RMSE Skill Score* (RMSS), *Brier Skill Score* (BSS) dan *Relative Operating Characteristic Skill Score* (ROCSS) BMA lebih besar dari RAW, serta reliabilitas BMA menjadi kategori “sempurna” dan “sangat berguna” dari sebelumnya “tidak berguna” pada RAW model S4. Penerapan BMA mampu memperbaiki kualitas prediksi CH RAW model S4 sehingga prediksi CH musiman menjadi bermanfaat jika dipakai dalam pengambilan keputusan terkait kondisi iklim beberapa bulan ke depan, khususnya pada fase *El Niño* atau *La Niña* di Pulau Jawa.

**Abstract.** When *El Niño Southern Oscillation* (ENSO), *El Niño* or *La Niña* phase is occurred, seasonal rainfall over Java Island tend to experience *Below Normal* (BN) or *Above Normal* (AN) conditions. Therefore, more accurate seasonal rainfall predictions are essential for both phases. The *ECMWF System 4* (S4) is a seasonal prediction model which can be used to generate probabilistic predictions BN or AN event. However, the direct output (RAW) of S4 models such as rainfall prediction has poor quality. The *Bayesian Model Averaging* (BMA) is selected as one of the *post-processing* statistics to improve its quality. The predictive *Probability Density Function* (PDF) of BMA is able to produce deterministic and probabilistic prediction more accurately than RAW S4 models. The accuracy is known from *Root Mean Square Error* (RMSE), *Brier Score* (BS) BMA that is lower than RAW; *RMSE Skill Score* (RMSS), *Brier Skill Score* (BSS) and *Relative Operating Characteristic Skill Score* (ROCSS) BMA that is greater than RAW. Hence, the reliability of the BMA is changing to the “perfect” and “very useful” category from the previous “not useful” in RAW model S4. The implementation of BMA is able to improve the prediction quality of CH RAW model S4. As the result, seasonal rainfall prediction will be useful in making decisions related to climate conditions for the coming months, especially in the *El Niño* or *La Niña* phase over Java Island.

## Pendahuluan

Layanan informasi iklim seperti prediksi curah hujan (CH) beberapa bulan ke depan, atau dikenal sebagai prediksi musim, menjadi hal penting bagi pengguna di sektor pertanian, pengairan dan perencanaan pembangunan suatu wilayah. Informasi prediksi musim yang akurat dapat membantu mereka dalam merancang strategi dan mengambil keputusan yang lebih baik untuk menghadapi

masalah terkait iklim. Oleh karena itu, beberapa penelitian terbaru terus berupaya untuk meningkatkan kualitas prediksi musim tersebut (Ratri *et al.* 2019; Manzanos *et al.* 2020; Hemri *et al.* 2020).

BMKG selaku *National Meteorological and Hydrological Services* (NMHS) di Indonesia, secara rutin memproduksi informasi prediksi musim setiap bulan termasuk prakiraan awal musim hujan atau kemarau (BMKG 2018) dengan basis model prediksi musim berasal

\* Corresponding author: [robi.muhsyah@gmail.com](mailto:robi.muhsyah@gmail.com)

dari *European Center for Medium-Range Weather Forecasts* (ECMWF) (BMKG 2020). ECMWF adalah salah satu lembaga internasional yang telah menjalankan prediksi musim untuk keperluan operasional informasi iklim dengan cakupan global dan jangkauan prediksi hingga tujuh bulan ke depan (Tompkins *et al.* 2017). Prediksi musim dari ECMWF dibuat dalam bentuk Sistem Prediksi *Ensemble* (SPE) (Molteni *et al.* 2011; Johnson *et al.* 2019), namun, parameter seperti CH dari keluaran langsung SPE ECMWF ini yang disebut *raw* model memiliki keterbatasan terkait dengan kualitas prediksi. Salah satu keterbatasan tersebut karena bias dan reliabilitas yang masih buruk untuk wilayah tropis termasuk Indonesia (Kim *et al.* 2012; Weisheimer and Palmer 2014; Ratri *et al.* 2019; Gubler *et al.* 2019)

Masalah bias dan reliabilitas CH dari keluaran SPE prediksi musim ECMWF telah diteliti pada beberapa kajian (Ratri *et al.* 2019; Manzananas *et al.* 2020; Hemri *et al.* 2020). Mereka menggunakan *post-processing* statistik untuk memperbaiki kekurangan yang terdapat dalam *raw* model. Ratri *et al.* (2019), melakukan bias koreksi menggunakan *Empirical Quantile Mapping* (EQM) terhadap CH model SEAS 5 (versi terbaru dari prediksi musim ECMWF *System 4*) di Pulau Jawa. Ratri *et al.* (2019), fokus pada kemampuan EQM untuk mengurangi bias *raw* SEAS 5 pada periode kemarau Juni-September. Manzananas *et al.* (2020), mengkalibrasi ECMWF *System 4* untuk wilayah Asia Tenggara dengan beberapa metode diantaranya *Bias Adjustment* (BA), *Perfect Prognosis* (PP), *Piani Quantile Mapping* (PQM). Kajiannya fokus pada periode Desember-Januari-Februari (DJF) atau saat terjadinya puncak musim hujan di sebagian besar wilayah Indonesia. Terbaru, dari Hemri *et al.* (2020), mengkalibrasi multi-model prediksi musim, termasuk SEAS 5, untuk cakupan global menggunakan beberapa metode kalibrasi seperti BA dan *Ensemble Model Output Statistic* (EMOS). Berdasarkan beberapa kajian tersebut, belum satupun kajian yang fokus pada evaluasi kemampuan SPE prediksi musim model ECMWF pada fase *El Niño Southern Oscillation* (ENSO): *El Niño* dan *La Niña*, baik *raw* ataupun hasil *post-processing* menggunakan metode koreksi bias ataupun kalibrasi prediksi *ensemble*.

ENSO adalah salah satu faktor yang mempengaruhi CH di Indonesia. Banyak kajian yang telah membahas hubungan fenomena ENSO, fase *El Niño* dan *La Niña*, terhadap CH di Indonesia dan semuanya mempunyai kesimpulan yang sama bahwa pada saat *El Niño* (*La Niña*) wilayah Indonesia umumnya mengalami pengurangan (penambahan) CH (Supari *et al.* 2018; Alsepan and

Minobe 2020). Sebagai contoh *El Niño* Kuat tahun 2015 menyebabkan sebagian besar wilayah Indonesia, terutama Pulau Jawa, mengalami CH kecil dari *percentile* ke-33 atau berada pada kategori Bawah Normal (BN) (Supari *et al.* 2016).

Penelitian ini bertujuan untuk mengkalibrasi CH dari keluaran SPE prediksi musim ECMWF *System 4* menggunakan *post-processing* statistik *Bayesian Model Averaging* (BMA). Menurut Kolachian dan Saghafian, (2019) dan Muharsyah *et al.* (2020), BMA mampu memperbaiki kualitas prediksi *ensemble*, baik prediksi yang disajikan dalam bentuk deterministik ataupun probabilistik. Penelitian ini merupakan lanjutan dari Muharsyah *et al.* (2020), namun lebih menekankan hasil kalibrasi BMA terhadap *raw* ECMWF *System 4* pada fase ENSO, *El Niño* dan *La Niña* karena belum menjadi fokus pada penelitian sebelumnya. Harapan penelitian ini agar hasil yang diperoleh dapat memberikan kontribusi untuk mengoptimalkan kemampuan prediksi model global seperti ECMWF *System 4*, sehingga mampu memperbaiki akurasi prediksi musim, khususnya pada saat kejadian ekstrim *El Niño* dan *La Niña*.

## Bahan dan Metode

### Bahan

Data observasi yang digunakan adalah CH bulanan yang diintegrasikan menjadi CH musiman atau jumlah CH per tiga bulan. Jika terdapat satu bulan dengan data kosong diantara tiga bulan yang dijumlahkan maka CH musim tersebut dianggap tidak ada. Data kosong diisi dengan nilai CH klimatologi musim yang bersesuaian. Data observasi diambil dari pengamatan pada 26 stasiun BMKG di Pulau Jawa dengan sebaran spasial disajikan pada Gambar 1 dan ketersediaan data pada Tabel 1.

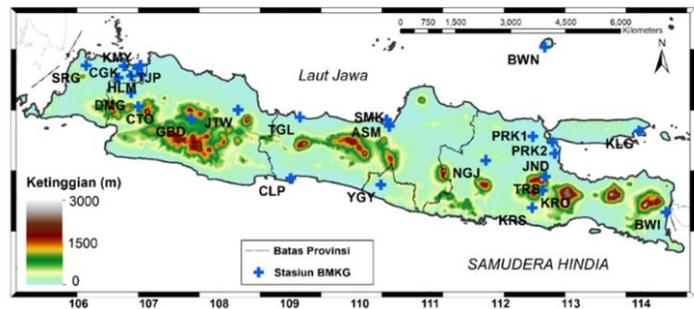
Data model yang dikalibrasi adalah CH dari keluaran SPE model prediksi musim ECMWF *System 4* (selanjutnya S4) (Molteni *et al.* 2011). Data ini dapat diunduh secara gratis di [http://climexp.knmi.nl/selectfield\\_sea.cgi?id=someone@somewhere](http://climexp.knmi.nl/selectfield_sea.cgi?id=someone@somewhere) dengan ketersediaan *reforecast* untuk periode 1981-2015. Penelitian ini menggunakan periode kajian 1981-2018, sisa tiga tahun digunakan data *realtime forecast* 2016-2018 diambil dari arsip operasional BMKG yang selama ini sudah menggunakan S4 sebagai basis produk informasi prakiraan hujan bulanan dan musiman. Model S4 merupakan prediksi *ensemble* dengan 15 (51) anggota untuk *reforecast* (*realtime forecast*). Pada proses pengolahan hanya 15 dari 51 anggota pada *realtime*

Tabel 1. Daftar 26 stasiun di Pulau Jawa

Table 1. List of 26 station over Java Island

No.	ID WMO	Nama Stasiun	Kode	LON	LAT	Elev (mdpl)	Miss (%)	Zero (%)	Ann, mean
1	96733	Pd, Betung	PDB	106,75	-6,26	27	1,8	0,2	2301
2	96737	Serang	SRG	106,11	-6,11	100	0,4	0,2	1618
3	96739	Curug	CRG	106,57	-6,29	42	5,5	0,0	2261
4	96741	Tj Priuk	TJP	106,88	-6,11	3	2,2	0,2	1761
5	96745	Kemayoran	KMY	106,84	-6,16	4	0,4	0,0	1888
6	96747	Halim	HLM	106,90	-6,25	4	12,1	0,4	2236
7	96749	Cengkareng	CGK	106,65	-6,12	11	11,0	0,0	1703
8	96751	Citeko	CTO	106,85	-6,70	920	8,3	0,0	3114
9	96753	Darmaga	DMG	106,75	-6,50	207	1,1	0,0	3822
10	96783	(G) Bandung	GBD	107,60	-6,88	791	5,0	0,7	2192
11	96791	Jatiwangi	JTW	108,27	-6,75	85	5,7	0,4	2690
12	96797	Tegal	TGL	109,15	-6,85	1	0,4	0,0	1788
13	96805	Cilacap	CLP	109,02	-7,73	8	0,4	0,0	3481
14	96835	Semarang (K)	SMK	110,42	-6,98	6	3,5	0,2	2289
15	96839	Semarang (M)	ASM	110,38	-6,89	5	5,0	1,5	2322
16	96845	Yogyakarta	YGY	110,30	-7,82	153	0,4	1,8	2436
17	96925	Bawean	BWN	112,63	-5,85	3	0,4	2,4	2339
18	96933	Perak I	PRK1	112,46	-7,13	3	1,1	5,0	1659
19	96935	Juanda	JND	112,77	-7,37	3	0,4	3,3	2143
20	96937	Perak II	PRK2	112,74	-7,21	3	19,7	0,4	1562
21	96943	Karangploso	KRO	112,60	-7,90	590	7,7	0,2	1856
22	96945	Tretes	TRS	112,64	-7,70	832	19,7	1,1	3783
23	96949	Karangates	KRS	112,45	-8,15	285	1,1	3,5	2174
24	96973	Kalianget	KLK	113,97	-7,05	1	24,1	2,9	1415
25	96975	Nganjuk	NGJ	111,79	-7,47	723	2,6	0,0	3011
26	96987	Banyuwangi	BWI	114,35	-8,21	52	2,4	0,2	1395

Keterangan: *Elev* = Ketinggian stasiun dalam meter diatas permukaan laut (mdpl), *Miss* = Jumlah data kosong (dalam persen), *Zero* = Jumlah CH musim sama dengan nol (dalam persen), *Ann Mean* = rata-rata jumlah CH per tahun (dalam milimeter), inisial pada nama stasiun Pd (Pondok), Tj (Tanjung), G (Geofisika), K (Klimatologi), M (Meteorologi)



Gambar 1. Sebaran spasial 26 stasiun BMKG di Pulau Jawa

Figure 1. Spatial distribution of 26 BMKG's station over Java Island

forecast yang dipilih secara acak untuk digabung dengan produk reforecast. Model S4 mempunyai resolusi spasial sekitar  $0,7^{\circ} \times 0,7^{\circ}$ . Nilai CH model S4 pada tiap stasiun diperoleh dengan metode *nearest neighbour* antara koordinat stasiun terhadap grid S4. Penelitian ini hanya fokus pada *lead time* nol (LT0) untuk dikalibrasi, karena prediksi pada LT0 memiliki bias lebih kecil dari LT yang

lain (Ratri *et al.* 2019) sehingga dianggap LT0 memiliki tingkat kesalahan paling rendah terhadap observasi. Prediksi CH model S4 yang belum dikalibrasi (selanjutnya disebut sebagai RAW model S4).

Penelitian ini juga menggunakan indeks ONI (*Ocean Nino Index*) ([https://origin.cpc.ncep.noaa.gov/products/analysis\\_monitoring/ensostuff/ONI\\_v5.php](https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php)). Indeks ONI

merupakan nilai Anomali *Seas Surface Temperature* (ASST) di samudera Pasifik tengah atau NINO34 [5N-5S,170W-120W]. Indeks ONI menunjukkan tahun kejadian La Niña (El Niño), jika terdapat nilai ASST  $\leq -0.5$  ( $\geq 0.5$ ) selama lebih atau sama dari lima musim berturut-turut, selain itu didefinisikan sebagai tahun Netral. Selama periode JFM 1981 hingga OND 2018 terdapat 119 (124) dari 454 bulan yang dikategorikan sebagai periode El Niño (La Niña) dengan rincian: 34 (17) kategori Kuat, 31 (38) Sedang dan 58 (75) Lemah.

### Metodologi

BMA adalah salah satu dari metode *dressing ensemble* yang dapat mengkalibrasi prediksi *ensemble* yang berasal dari SPE seperti ECMWF (Buizza 2018). Hasil kalibrasi BMA berbentuk fungsi kerapatan peluang atau *Probability Density Function* (PDF) yang dapat dipakai untuk menghitung peluang suatu kejadian (Raftery *et al.* 2005; Vrugt *et al.* 2008; Slougher *et al.* 2009). PDF yang dihasilkan BMA merupakan PDF prediktif. PDF ini diperoleh berdasarkan informasi di masa lalu (periode *training*) antara masing-masing anggota prediksi *ensemble* dengan nilai observasi. Masing-masing anggota *ensemble* mempunyai PDF *posterior* yang berkontribusi pada PDF prediktif tersebut (Raftery *et al.* 2005; Vrugt *et al.* 2008; Slougher *et al.* 2009). Ketersediaan *reforecast* data ECMWF *System 4* (Molteni *et al.* 2011; Manzananas *et al.* 2020) dapat dipakai untuk membangun PDF *posterior* antara prediksi *ensemble* dan observasi, sehingga diperoleh PDF prediktif untuk target musim yang dikalibrasi.

Pada penelitian ini, kemampuan BMA dievaluasi pada saat kejadian ENSO, oleh karena itu kalibrasi BMA dilakukan pada skala CH musiman berupa akumulasi CH per tiga bulan seperti JFM (Januari-Ferburai-Maret), FMA (Februari-Maret-April), dan seterusnya. PDF prediktif untuk tiap musim tersebut di peroleh dengan skema *Conditional Training Window*. Lebih jelas tentang skema ini dapat ditemukan pada Erickson *et al.* (2012) dan Muharsyah *et al.* (2020).

Agar *robust* secara statistik, maka konsep *Leave One Out Cross Validation* (LOOCV) (Themeßl *et al.* 2011) diterapkan untuk mendapatkan PDF prediktif pada tiap stasiun dan tiap musim dari JFM 1981 hingga DJF 2018. Prinsip kerja LOOCV adalah satu tahun dikeluarkan sebagai *testing* sedangkan sisanya digunakan sebagai periode *training* dan dilakukan berulang hingga semua tahun pada musim tersebut diperlakukan sebagai tahun *testing*. Selanjutnya, kemampuan BMA pada periode ENSO dianalisis dengan memisahkan PDF prediktif untuk musim yang diidentifikasi sebagai El Niño dan La Niña berdasarkan indeks ONI. Kemudian, seperti yang

dijelaskan diawal bahwa dampak El Niño (La Niña) cenderung mengurangi (meningkatkan) CH di Pulau Jawa, maka penelitian ini fokus pada kemampuan BMA untuk prediksi peluang kejadian ekstrem di Bawah Normal (BN) [peluang Atas Normal (AN)] pada saat indeks ONI menunjukkan kategori El Niño [La Niña].

### Bayesian Model Averaging

Tujuan utama dari metode BMA adalah menghasilkan prediksi probabilistik (berpeluang) yang dinyatakan dalam bentuk PDF prediktif (Raftery *et al.* 2005; Slougher *et al.* 2007; Wang *et al.* 2012). Misalkan  $y_{nt}$  adalah observasi CH musiman di stasiun ke- $n$  ( $n = 1, 2, \dots, 26$ ), pada musim ke- $t$  ( $t = \text{JFM, FMA, } \dots, \text{DJF}$ ) dan  $f = (f_{1nt}, f_{2nt}, \dots, f_{knt})$  adalah prediksi *ensemble* RAW model S4 dengan  $k = 15$  anggota pada tiap stasiun dan musim yang bersesuaian. Setiap anggota *ensemble*  $f_{knt}$  mempunyai PDF *posterior*  $g_{knt}(y_{nt}|f_{knt})$ , sedemikian sehingga model PDF prediktif BMA (selanjutnya disebut PDF BMA), dapat dinyatakan pada persamaan (1)

$$p(y_{nt}|f_{1nt}, f_{2nt}, \dots, f_{knt}) = \sum_{k=1}^{15} w_{knt} g_{knt}(y_{nt}|f_{knt}) \quad (1)$$

Penelitian ini menggunakan distribusi *Gamma* untuk mendapatkan PDF BMA. Distribusi *Gamma* hanya berlaku untuk CH lebih besar dari nol, untuk mengatasi hal ini jika terdapat CH musiman dengan nilai nol maka sebelum kalibrasi BMA diterapkan, terlebih dahulu dilakukan transformasi dengan mengubah CH sama dengan nol mm menjadi 0,1 mm. Nilai CH sebelum dan setelah diganti dengan 0,1 mm telah diuji dengan *t-test* dan diperoleh kedua nilai tersebut tidak memiliki perbedaan yang signifikan dengan tingkat kepercayaan 99%. Penjelasan rinci terkait prinsip statistik pada metode BMA yang dipakai pada penelitian ini dapat ditemukan di Muharsyah *et al.* (2020).

### Evaluasi Hasil Kalibrasi BMA

Pada tahap evaluasi, kemampuan BMA dibandingkan dengan RAW untuk mengetahui apakah hasil kalibrasi BMA mampu memperbaiki kualitas RAW model S4. Kualitas tersebut ditandai oleh peningkatan akurasi berdasarkan ukuran rata-rata kesalahan, *skill* dan reliabilitas. Setelah diperoleh PDF BMA untuk 454 periode musim (JFM1981 - NDJ2018), kemudian dipisahkan berdasarkan indeks ONI, masing-masing 119 (124) musim untuk El Niño (La Niña). Kemampuan BMA pada dua fase ENSO tersebut diukur melalui dua arah: deterministik dan probabilistik. Semua skor yang digunakan dalam tahapan evaluasi merujuk pada *Forecast Verification methods Across Time and Space*

*Scales* (<https://www.cawcr.gov.au/projects/verification/>) dan *Forecast Verification A Practitioner's Guide in Atmospheric Science* (Jolliffe and Stephenson 2003).

Pertama secara deterministik, kemampuan BMA dievaluasi dengan membandingkan rata-rata kesalahan antara *median* PDF BMA dan *median* PDF RAW. *Root Mean Square Error* (RMSE) antara kedua median tersebut dihitung terhadap observasi. Jika RMSE BMA lebih kecil dari RMSE RAW maka dikatakan terdapat peningkatan akurasi. *RMSE Skill Score* (RMSS) adalah ukuran *skill* dari RMSE. Semakin besar atau bernilai positif RMSS berarti menunjukkan adanya perbaikan dalam sistem prediksi dibandingkan kondisi klimatologi.

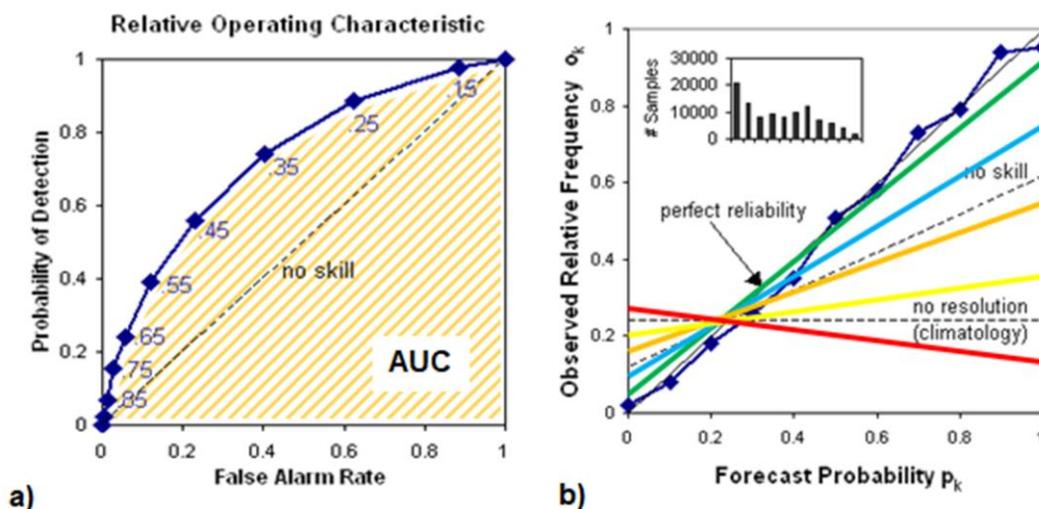
Kedua secara probabilistik, kemampuan PDF BMA dievaluasi dalam memprediksi peluang BN (AN) saat El Niño (La Niña). Peluang BN (AN) dihitung dari luas area di bawah kurva PDF BMA dan PDF RAW yang lebih kecil (besar) dari P33 (P66). Verifikasi prediksi probabilistik BN(AN) diukur menggunakan *Brier Score* (BS) dan *Brier Skill Score* (BSS). Tahapan verifikasi ini dilakukan untuk stasiun-per-stasiun sehingga dapat diketahui kemampuan BMA di tiap stasiun yang dikalibrasi.

*Area Under Curve* (AUC) of *Relative Operating Characteristic* (ROC) (Gambar 2a), *ROC Skill Score* (ROCSS), dan *Reliability Diagram* (RD) (Gambar 2b) juga dipakai untuk mengevaluasi prediksi probabilistik BN

(AN). Akan tetapi untuk ROC dan RD membutuhkan jumlah *sample* atau jumlah kejadian yang cukup besar (biasanya lebih dari 100), agar setiap kelas (*bin*) probabilistik atau *axis-x* pada Gambar 2 dapat terwakili. Semakin jarang probabilistik suatu kejadian terjadi maka semakin besar *sample* atau jumlah kejadian yang diperlukan untuk menghasilkan skor yang stabil (Jolliffe and Stephenson, 2003). Oleh karena itu kedua ukuran tersebut tidak dihitung stasiun-per-stasiun melainkan sekaligus keseluruhan stasiun di Pulau Jawa. Alasan lain penghitungan skor untuk seluruh daratan Pulau Jawa dapat dilakukan karena menurut Aldrian and Susanto (2003), Pulau Jawa termasuk wilayah dengan tipe hujan yang sama. Pengaruh ENSO cenderung hampir sama pada tiap daerah di Pulau Jawa (Supari *et al.*, 2018; Alsepan and Minobe, 2020).

Berikut ini penjelasan singkat tentang BS, BSS, AUC ROC, ROCSS:

- a) *BS dan BSS*. BS adalah rata-rata kuadrat kesalahan prediksi probabilistik suatu kejadian (Jolliffe and Stephenson 2003). BS mengukur besar tingkat kesalahan prediksi kejadian yang bersifat dikotomi, misal “terjadi BN” atau “tidak terjadi BN”, nilai BS semakin baik jika mendekati nol.
- b) *AUC ROC dan ROCSS*. ROC berbentuk kurva (Gambar 2a), biasanya dipakai untuk mengevaluasi kemampuan sistem prediksi dalam mendiskriminasi



Gambar 2. (a) AUC adalah luas area (arsiran kuning) di bawah kurva ROC dan (b) lima kategori *Weisheimer Score* (WS) pada *Reliability Diagram*, yaitu kelas 1- “berbahaya jika digunakan” (merah), 2- “tidak berguna” (kuning), 3- “cukup berguna” (oranye), 4- “sangat berguna” (biru), 5- “sempurna” (hijau) (dimodifikasi dari <https://www.cawcr.gov.au/projects/verification/>)

Figure 2. (a) AUC is the area under ROC curve (yellow shading area), and (b) Five WS category based on *Reliability diagram* curve: class 1- “dangerous useless” (red), 2- “not useful” (yellow), 3- “marginally useful” (orange), 4- “very useful” (blue) and 5- “perfect” (green) (modified from <https://www.cawcr.gov.au/projects/verification/>)

suatu kejadian (Jolliffe and Stephenson 2003). Jika  $AUC ROC > 0,5$  maka disebut sistem prediksi mampu mendiskriminasi atau membedakan, kejadian “BN” atau “bukan BN” (begitu juga untuk AN) dan sebaliknya jika  $AUC ROC \leq 0,5$ . ROCSS adalah *skill score* dari ROC yang dihitung dengan  $2(AUC ROC) - 1$ .

- c) *Reliability Diagram*. Sistem prediksi disebut reliabel (handal) jika prediksi probabilistik yang dihasilkan sebanding dengan frekuensi kejadian yang sebenarnya. Reliabilitas diukur dari kemiringan kurva RD terhadap *perfect reliability line* (Jolliffe and Stephenson, 2003). Terdapat lima kategori reliabilitas berdasarkan kemiringan kurva RD (Gambar 2 b) atau dikenal sebagai *Weisheimer Score* (WS) (Gubler *et al.* 2019). Lima kelas WS menggambarkan manfaat dari prediksi probabilistik BN atau AN untuk dipakai dalam pengambilan keputusan terkait prediksi musim di suatu wilayah, kelas-1: “berbahaya jika digunakan”, -2: “tidak berguna”, -3: “cukup berguna”, -4: “sangat berguna” dan -5 “sempurna” (Weisheimer and Palmer 2014).

Sama seperti RMSS, BSS dan ROCSS adalah ukuran *skill* atau ketepatan relatif dari model prediksi (BMA atau RAW), terhadap prediksi referensi (biasanya nilai klimatologi observasi). BSS dan ROCSS semakin baik jika lebih besar dari nol atau mendekati satu, sebaliknya dianggap buruk jika bernilai negatif. *Skill* positif menandakan adanya peningkatan akurasi karena membaiknya sistem prediksi. Sebaliknya *skill* negatif menunjukkan sistem prediksi lebih buruk dari prediksi klimatologi nya (Jolliffe and Stephenson 2003).

## Hasil dan Pembahasan

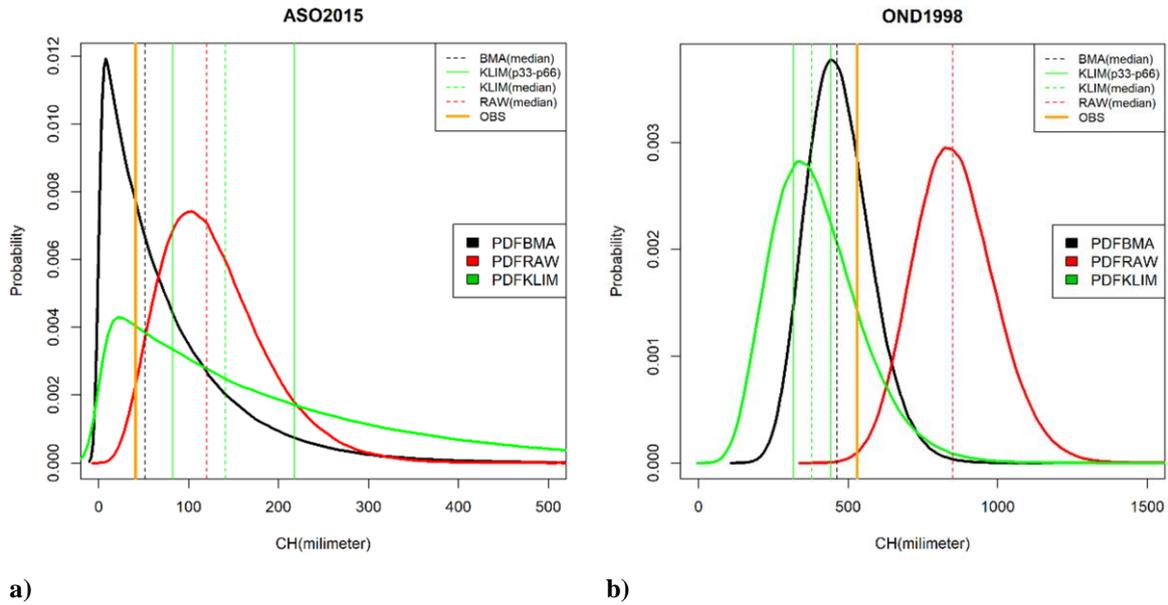
### Kemampuan BMA untuk Prediksi Deterministik saat El Niño dan La Niña

Semua daratan Pulau Jawa masuk pada kategori tipe A dalam pewilayahan pola hujan oleh Aldrian and Susanto (2003). Tipe A mempunyai puncak hujan di sekitar DJF dan puncak kemarau di sekitar JAS. Terdapat perbedaan jumlah CH yang besar antara kedua musim tersebut, dengan kata lain variasi CH pada periode musim hujan lebih besar dari musim kemarau. Akibatnya, bentuk PDF BMA yang dihasilkan pada saat mengkalibrasi periode musim kemarau umumnya lebih sempit dibandingkan pada musim hujan. Selain itu, meskipun seluruh Pulau Jawa memiliki tipe hujan yang sama, namun faktor topografi yang kompleks, terdiri dari pegunungan dan pantai, berkontribusi pada keragaman CH di Pulau Jawa. Ragam

topografi ini tampak dari ketinggian 26 stasiun yang beragam dari 1 meter (KLG) hingga > 900 meter (CTO) di atas permukaan laut (mdpl). Stasiun yang berada di topografi lebih tinggi cenderung mempunyai rata-rata hujan tahunan lebih besar dibandingkan stasiun yang berada di sekitar pantai (Gambar 1 dan Tabel 1). Oleh karena itu, bentuk PDF BMA stasiun-per-stasiun bervariasi bergantung pada ruang (dimana stasiun itu berada) dan waktu (musim apa) yang dikalibrasi.

Ada sebanyak 454 PDF BMA selama periode JFM1981 hingga OND2018 untuk tiap stasiun yang dihasilkan dari LOOCV. Masing-masing PDF BMA disandingkan dengan PDF RAW dan PDF Klimatologi (Gambar 3). PDF BMA pada stasiun ke-*n* dan musim ke-*t* mempunyai bentuk yang beragam, lebar (varians) ataupun tingginya (besarnya nilai *probability*). Lebar dan tinggi PDF BMA tidak menunjukkan akurasi dari PDF BMA tersebut. Secara sederhana dapat diamati pada Gambar 3 bahwa PDF BMA yang akurat jika nilai *median* dari PDF BMA tersebut berada di sekitar nilai observasi. Gambar 3 adalah contoh PDF BMA hasil kalibrasi untuk stasiun SRG pada periode musim ASO 2015 (Gambar 3a), mewakili kejadian El Niño Kuat dan OND 1998 (Gambar 3b) mewakili kejadian La Niña Kuat. Pada dua kejadian ENSO tersebut, PDF BMA lebih baik dari PDF RAW karena *median* PDF BMA lebih dekat ke nilai observasi dibandingkan *median* PDF RAW. Bahkan, pada kedua gambar tersebut, *median* PDF BMA juga lebih dekat ke nilai observasi dibandingkan *median* PDF Klimatologi, atau dengan kata lain PDF BMA tidak hanya akurat tetapi juga mempunyai *skill* (Murphy, 1993; Jolliffe and Stephenson 2003).

Akurasi PDF BMA pada saat fase El Niño dan La Niña untuk stasiun-per-stasiun di Pulau Jawa diberikan pada Tabel 2. Pada fase La Niña, PDF BMA mempunyai akurasi lebih baik dibandingkan PDF RAW. Selisih RMSE antara *median* PDF BMA terhadap PDF RAW bernilai negatif pada tiap stasiun. Rata-rata dari 124 kali kejadian La Niña menunjukkan *median* PDF BMA lebih dekat ke nilai observasi dibandingkan *median* PDF RAW. Begitu juga dari 119 kali kejadian El Niño, umumnya *median* PDF BMA lebih unggul dari *median* PDF RAW kecuali pada stasiun DMG. Selisih RMSE yang semakin kecil (semakin negatif), menyebabkan semakin tinggi nya selisih RMSS antara BMA dan RAW. Ini menandakan selain akurasi, *median* PDF BMA mempunyai *skill* yang lebih baik karena hampir semua RMSS bernilai positif dan lebih tinggi dari RMSS RAW, kecuali pada stasiun DMG pada fase El Niño. RMSS positif menunjukkan adanya peningkatan akurasi dengan menerapkan metode BMA



Gambar 3. Perbandingan PDF BMA (hitam), PDF RAW (merah) saat kejadian: (a) El Niño Kuat, ASO 2015 dan (b) La Niña Kuat, OND 1998 di stasiun meteorologi SRG. PDF Klimatologi (1981-2010) pada kedua musim tersebut ditunjukkan oleh kurva hijau. Garis vertikal *dash* adalah nilai *median* masing-masing PDF tersebut, sedangkan garis vertikal *solid* adalah nilai observasi (oranye), serta persentil-33 (P33) [P66] (hijau) dari PDF Klimatologi sebagai batas kejadian BN dan AN

Figure 3. Comparison of PDF BMA (black color), PDF RAW (red) when: (a) Strong El Niño ASO 2015 and (b) Strong La Niña OND 1998, at meteorology station of SRG. PDF Climatology (1981-2010) for both of season shown by green curve. Vertical dash line indicate median for respective PDF meanwhile vertical solid line indicate observed value (orange), and percentile-33 (P33) [P66] of PDF Climatology (green) as the threshold of BN(AN)

sebagai sistem prediksi dibandingkan penggunaan RAW model sebagai sistem prediksi, bahkan jauh lebih unggul dari penggunaan nilai klimatologi untuk memprediksi CH musiman pada kedua fase ENSO tersebut.

**Kemampuan BMA untuk Prediksi Probabilistik BN saat El Niño dan AN saat La Niña**

Selama 124 (119) kali kejadian El Niño (La Niña), dibandingkan PDF RAW, umumnya PDF BMA mampu menghasilkan peluang BN (AN) saat El Niño (La Niña) dengan persentase mendekati kejadian sebenarnya (100%). Gambar 4 adalah contoh prediksi probabilistik untuk kategori BN, N dan AN serta kategori observasi pada saat El Niño Kuat musim ASO 1997 dan ASO 2015. Pada kedua musim tersebut, PDF RAW cenderung lebih banyak memprediksi peluang N dan AN dibandingkan peluang BN untuk tiap stasiun di Pulau Jawa, sedangkan PDF BMA memprediksi peluang BN di semua stasiun dengan kisaran persentase 50-80%. Pada contoh ini, sangat jelas BMA mampu memperbaiki kekurangan yang terdapat di dalam prediksi *ensemble* RAW model S4. Selanjutnya, Gambar 5 adalah contoh prediksi probabilistik untuk BN, N dan AN serta kategori observasi pada saat La Niña Kuat pada musim OND 1998 dan SON 2010. Pada contoh ini,

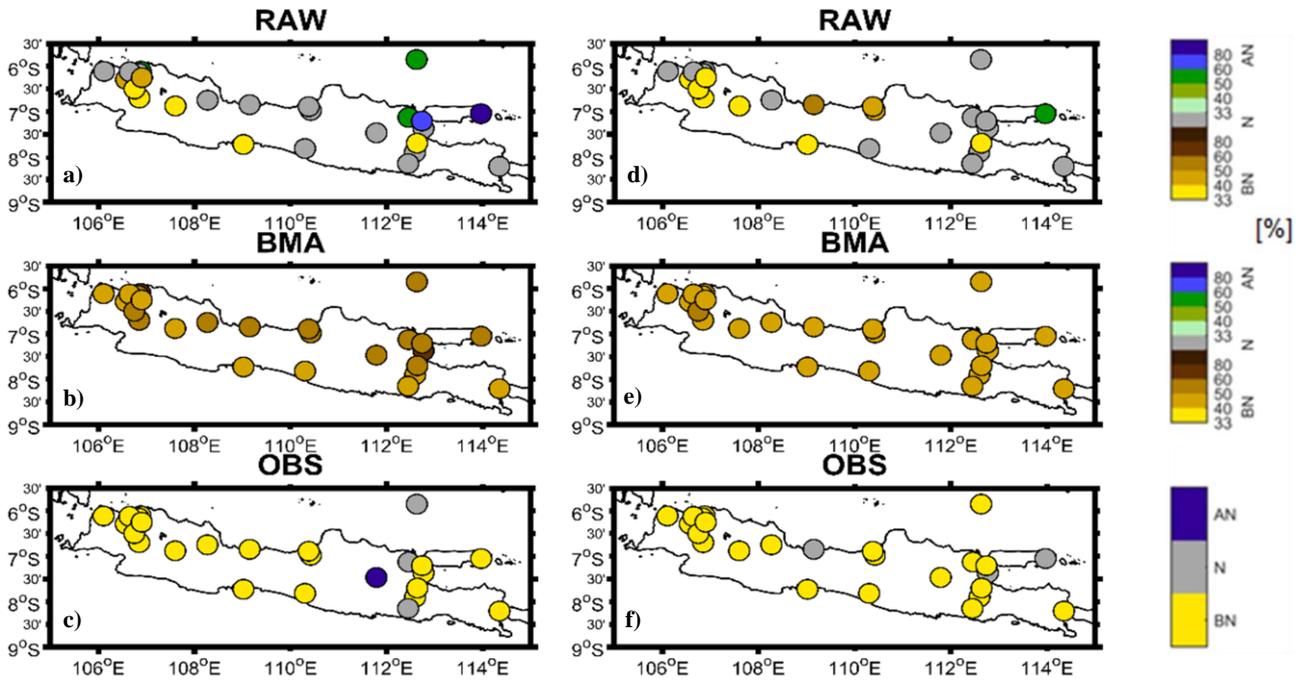
umumnya PDF RAW dan PDF BMA hampir sama dalam memprediksi peluang kejadian AN pada saat terjadi nya La Niña Kuat di kedua musim tersebut. Akan tetapi pada beberapa stasiun PDF BMA mampu memprediksi peluang AN sedangkan PDF RAW memprediksi peluang N bahkan BN.

Keunggulan PDF BMA terhadap PDF RAW dalam menghasilkan peluang BN (AN) pada saat El Niño (La Niña) dengan persentase mendekati kejadian sebenarnya, terukur dari skor BS dan BSS pada Tabel 2. Pada fase La Niña, selisih BS BMA terhadap RAW bernilai negatif di setiap stasiun. Hasil yang hampir sama juga ditemukan pada fase El Niño, namun terdapat satu stasiun, yaitu DMG dengan selisih BS positif. Kemudian pada skor BSS, sangat jelas bahwa RAW model S4 memiliki sistem prediksi dengan *skill* buruk karena hampir semua BSS bernilai negatif pada kedua fase ENSO tersebut. Penggunaan metode BMA mampu memperbaiki hal tersebut, dari *skill* buruk menjadi *skill* yang lebih baik ditandai dengan BSS positif hampir di semua stasiun di Pulau Jawa, kecuali di stasiun CTO pada saat El Niño, meski sebenarnya di stasiun tersebut BMA mampu meningkatkan nilai BSS dari -1.6 menjadi -0.04.

Tabel 2. Nilai RMSE, RMSS, BS dan BSS saat kejadian La Niña dan El Niño, kolom (1) BMA, kolom (2) RAW dan kolom (3) selisih BMA terhadap RAW. Tulisan cetak miring pada kolom (3) menandakan BMA tidak lebih baik dari RAW. Teks miring pada BSS kolom (1) menunjukkan *skill* negatif atau *skill* buruk

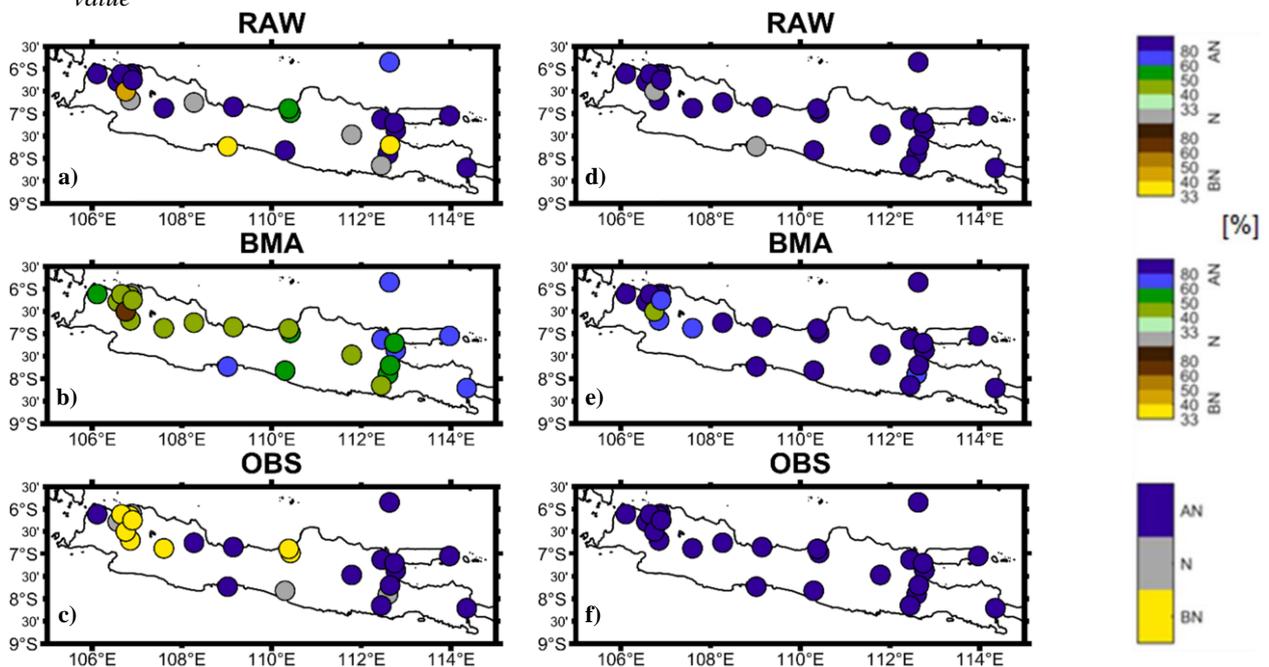
Table 2. RMSE, RMSS, BS and BSS when La Niña dan El Niño event, column with number (1) is BMA, (2) RAW, (3) difference BMA to RAW. *Italic text in (3) indicate BMA not better than RAW. Italic text in BSS for column number (1) indicate negative or poor skill*

Kode Stasiun	RMSE						MSS						BS						BSS					
	El Nino			La Nina			El Nino			La Nina			El Nino (BN)			La Nina (AN)			El Nino (BN)			La Nina (AN)		
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
PDB	188	187	<b>-1</b>	213	195	<b>-18</b>	0,44	0,45	<b>0,01</b>	0,14	0,21	<b>0,07</b>	0,24	0,22	<b>-0,03</b>	0,24	0,19	<b>-0,06</b>	0,12	0,22	<b>0,10</b>	<i>-0,11</i>	0,15	<b>0,26</b>
SRG	231	118	<b>-113</b>	280	127	<b>-153</b>	0,12	0,55	<b>0,43</b>	-0,2	0,43	<b>0,68</b>	0,44	0,23	<b>-0,21</b>	0,51	0,20	<b>-0,31</b>	<i>-0,50</i>	0,20	<b>0,70</b>	<i>-1,14</i>	0,15	<b>1,28</b>
CRG	149	140	<b>-9</b>	213	188	<b>-26</b>	0,49	0,52	<b>0,03</b>	0,05	0,17	<b>0,11</b>	0,23	0,23	<b>0,00</b>	0,35	0,20	<b>-0,15</b>	0,24	0,25	<b>0,01</b>	<i>-0,56</i>	0,12	<b>0,68</b>
TJP	249	168	<b>-81</b>	283	203	<b>-81</b>	0,34	0,55	<b>0,21</b>	0,12	0,37	<b>0,25</b>	0,45	0,26	<b>-0,18</b>	0,40	0,21	<b>-0,19</b>	<i>-0,57</i>	0,08	<b>0,65</b>	<i>-0,78</i>	0,07	<b>0,85</b>
KMY	232	181	<b>-51</b>	244	193	<b>-51</b>	0,37	0,51	<b>0,14</b>	0,16	0,33	<b>0,18</b>	0,43	0,24	<b>-0,19</b>	0,38	0,22	<b>-0,17</b>	<i>-0,52</i>	0,15	<b>0,67</b>	<i>-0,67</i>	0,06	<b>0,73</b>
HLM	233	157	<b>-76</b>	281	186	<b>-96</b>	0,26	0,50	<b>0,24</b>	0,02	0,35	<b>0,33</b>	0,36	0,22	<b>-0,14</b>	0,45	0,20	<b>-0,25</b>	<i>-0,42</i>	0,12	<b>0,54</b>	<i>-1,03</i>	0,09	<b>1,12</b>
CGK	317	198	<b>-120</b>	320	263	<b>-56</b>	0,29	0,56	<b>0,27</b>	0,23	0,37	<b>0,13</b>	0,39	0,22	<b>-0,17</b>	0,21	0,21	<b>0,00</b>	<i>-0,67</i>	0,04	<b>0,71</b>	<i>-0,09</i>	0,09	<b>0,18</b>
CTO	485	251	<b>-235</b>	352	264	<b>-88</b>	-0,35	0,30	<b>0,65</b>	-0,3	0,01	<b>0,33</b>	0,60	0,21	<b>-0,39</b>	0,21	0,17	<b>-0,04</b>	-1,96	-0,04	<b>1,92</b>	<i>-0,11</i>	0,09	<b>0,20</b>
DMG	177	203	<b>26</b>	209	158	<b>-52</b>	0,49	0,42	<b>-0,07</b>	0,05	0,29	<b>0,23</b>	0,19	0,20	<b>0,01</b>	0,31	0,18	<b>-0,13</b>	0,22	0,19	<b>-0,03</b>	<i>-0,66</i>	0,03	<b>0,69</b>
GBD	293	151	<b>-142</b>	406	238	<b>-168</b>	0,40	0,69	<b>0,29</b>	0,17	0,51	<b>0,34</b>	0,37	0,24	<b>-0,14</b>	0,39	0,24	<b>-0,16</b>	<i>-0,26</i>	0,20	<b>0,46</b>	<i>-0,46</i>	0,12	<b>0,58</b>
JTW	176	140	<b>-36</b>	227	150	<b>-77</b>	0,44	0,56	<b>0,11</b>	0,31	0,55	<b>0,23</b>	0,36	0,27	<b>-0,10</b>	0,32	0,23	<b>-0,09</b>	<i>-0,30</i>	0,04	<b>0,35</b>	<i>-0,16</i>	0,16	<b>0,32</b>
TGL	482	364	<b>-118</b>	534	337	<b>-197</b>	0,13	0,35	<b>0,21</b>	-0,1	0,33	<b>0,39</b>	0,33	0,20	<b>-0,14</b>	0,26	0,19	<b>-0,07</b>	<i>-0,29</i>	0,24	<b>0,53</b>	<i>-0,18</i>	0,15	<b>0,33</b>
CLP	159	153	<b>-7</b>	212	206	<b>-6</b>	0,52	0,54	<b>0,02</b>	0,43	0,45	<b>0,02</b>	0,28	0,27	<b>0,00</b>	0,30	0,25	<b>-0,06</b>	0,16	0,18	<b>0,01</b>	<i>-0,03</i>	0,16	<b>0,19</b>
SMK	163	158	<b>-5</b>	206	201	<b>-5</b>	0,49	0,51	<b>0,01</b>	0,45	0,47	<b>0,01</b>	0,30	0,28	<b>-0,01</b>	0,29	0,25	<b>-0,04</b>	0,07	0,11	<b>0,04</b>	0,00	0,14	<b>0,14</b>
ASM	242	234	<b>-8</b>	331	308	<b>-23</b>	0,45	0,47	<b>0,02</b>	0,34	0,39	<b>0,05</b>	0,33	0,25	<b>-0,07</b>	0,27	0,20	<b>-0,07</b>	<i>-0,09</i>	0,16	<b>0,25</b>	<i>-0,30</i>	0,04	<b>0,33</b>
YGY	191	180	<b>-11</b>	302	217	<b>-85</b>	0,56	0,59	<b>0,03</b>	0,34	0,53	<b>0,19</b>	0,34	0,19	<b>-0,14</b>	0,28	0,21	<b>-0,06</b>	<i>-0,15</i>	0,34	<b>0,49</b>	0,01	0,24	<b>0,23</b>
BWN	142	122	<b>-20</b>	191	142	<b>-49</b>	0,57	0,63	<b>0,06</b>	0,44	0,59	<b>0,14</b>	0,30	0,22	<b>-0,08</b>	0,41	0,24	<b>-0,17</b>	<i>-0,18</i>	0,14	<b>0,32</b>	<i>-0,51</i>	0,11	<b>0,63</b>
PRK1	263	167	<b>-96</b>	280	219	<b>-61</b>	0,43	0,64	<b>0,21</b>	0,32	0,47	<b>0,15</b>	0,41	0,22	<b>-0,19</b>	0,27	0,21	<b>-0,06</b>	<i>-0,47</i>	0,23	<b>0,70</b>	<i>-0,09</i>	0,14	<b>0,23</b>
JND	148	127	<b>-21</b>	186	136	<b>-50</b>	0,53	0,60	<b>0,07</b>	0,42	0,57	<b>0,16</b>	0,32	0,25	<b>-0,07</b>	0,34	0,24	<b>-0,10</b>	<i>-0,15</i>	0,10	<b>0,25</b>	<i>-0,16</i>	0,16	<b>0,33</b>
PRK2	151	134	<b>-17</b>	205	149	<b>-56</b>	0,58	0,62	<b>0,05</b>	0,37	0,54	<b>0,17</b>	0,26	0,22	<b>-0,04</b>	0,37	0,23	<b>-0,15</b>	<i>-0,15</i>	0,03	<b>0,18</b>	<i>-0,63</i>	0,01	<b>0,63</b>
KRO	615	285	<b>-330</b>	626	265	<b>-362</b>	0,13	0,60	<b>0,46</b>	0	0,58	<b>0,58</b>	0,39	0,26	<b>-0,12</b>	0,27	0,20	<b>-0,07</b>	<i>-0,33</i>	0,09	<b>0,42</b>	<i>-0,12</i>	0,18	<b>0,30</b>
TRS	190	179	<b>-11</b>	215	193	<b>-21</b>	0,49	0,52	<b>0,03</b>	0,4	0,46	<b>0,06</b>	0,26	0,25	<b>-0,01</b>	0,27	0,21	<b>-0,06</b>	<i>-0,02</i>	0,01	<b>0,03</b>	<i>-0,27</i>	0,00	<b>0,28</b>
KRS	156	130	<b>-26</b>	233	174	<b>-59</b>	0,44	0,54	<b>0,09</b>	0,23	0,43	<b>0,20</b>	0,35	0,24	<b>-0,10</b>	0,30	0,24	<b>-0,06</b>	<i>-0,25</i>	0,13	<b>0,37</b>	<i>-0,02</i>	0,20	<b>0,22</b>
KLG	440	225	<b>-215</b>	463	276	<b>-186</b>	0,29	0,64	<b>0,35</b>	0,19	0,52	<b>0,33</b>	0,47	0,22	<b>-0,25</b>	0,23	0,20	<b>-0,03</b>	<i>-1,08</i>	0,02	<b>1,10</b>	<i>-0,02</i>	0,12	<b>0,14</b>
NGJ	229	144	<b>-86</b>	247	141	<b>-106</b>	0,01	0,38	<b>0,37</b>	-0,1	0,39	<b>0,46</b>	0,36	0,27	<b>-0,09</b>	0,32	0,24	<b>-0,08</b>	<i>-0,11</i>	0,18	<b>0,29</b>	<i>-0,16</i>	0,14	<b>0,30</b>
BWI	204	200	<b>-4</b>	198	181	<b>-17</b>	0,44	0,45	<b>0,01</b>	0,27	0,33	<b>0,06</b>	0,23	0,21	<b>-0,01</b>	0,27	0,19	<b>-0,08</b>	0,10	0,15	<b>0,05</b>	<i>-0,35</i>	0,03	<b>0,38</b>



Gambar 4. Prediksi probabilistik (dalam persen) untuk kategori BN, N dan AN pada dua kali kejadian El Niño Kuat ASO 1997 (kolom kiri) dan ASO 2015 (kolom kanan) untuk (a, d) RAW model S4, (b, e) hasil kalibrasi BMA serta (c,f) nilai observasi yang sudah di ubah dalam kategori BN, N dan AN

Figure 4. Probability prediction (in percent) of BN, N and AN in two event of Strong El Niño, ASO 1997 (left column) and ASO 2015 (right column) of (a, d) RAW model S4, (b, e) BMA output and (c, f) categorical event of observed value



Gambar 5. Sama seperti Gambar 4 namun untuk periode La Niña Kuat OND 1998 (kolom kiri) dan SON2010 (kolom kanan)

Figure 5. Sama as figure 4, but for two Strong La Niña event in OND 1998 (left column) and SON2010 (right column)

Gambar 6 merangkum keseluruhan hasil pada Tabel 1 untuk nilai BS dan BSS, dan ditambah dengan skor AUC ROC dan ROCSS. Perbedaannya dengan Tabel 1,

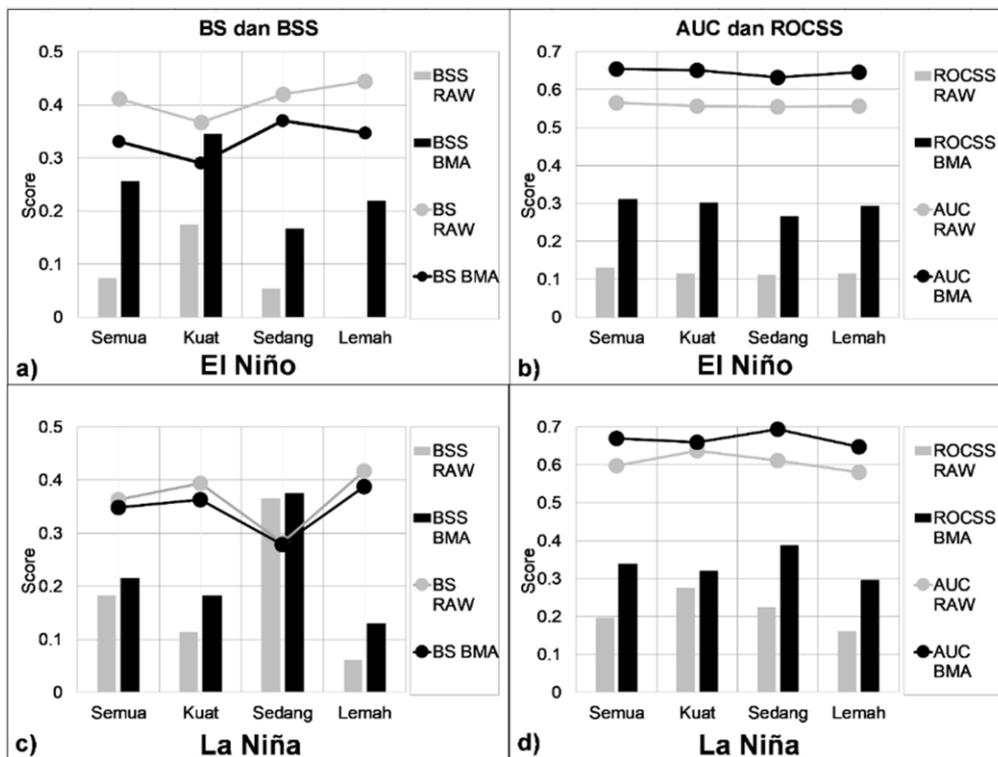
perhitungan skor melibatkan semua stasiun di Pulau Jawa tanpa membedakan lokasi geografis tiap stasiun. Masing-masing skor juga dihitung pada tiap kategori ENSO:

“Semua”, “Kuat”, “Sedang” dan “Lemah”. Hasil skor-skor tersebut bervariasi, namun pada tiap kategori ENSO, BMA selalu mengungguli RAW. Terdapat peningkatan pada skor BSS dan ROCSS, dan peningkatan pada fase El Niño sedikit lebih besar dibandingkan La Niña. Bahkan pada kategori El Niño Lemah, BMA mampu memperbaiki *skill* RAW model S4 yang sebelumnya nol menjadi 0,22. Ini menunjukkan *skill* BMA sangat baik memprediksi kejadian BN saat El Niño.

ROCSS BMA yang lebih tinggi dari RAW menandakan adanya peningkatan kemampuan dalam hal diskriminasi suatu kejadian, dari sebelumnya RAW model yang sulit membedakan kejadian BN saat El Niño atau sebaliknya AN saat La Niña. Penggunaan BMA mampu memperbaiki akurasi RAW model S4 untuk mengidentifikasi munculnya kejadian BN (AN) saat El Niño (La Niña) di Pulau.

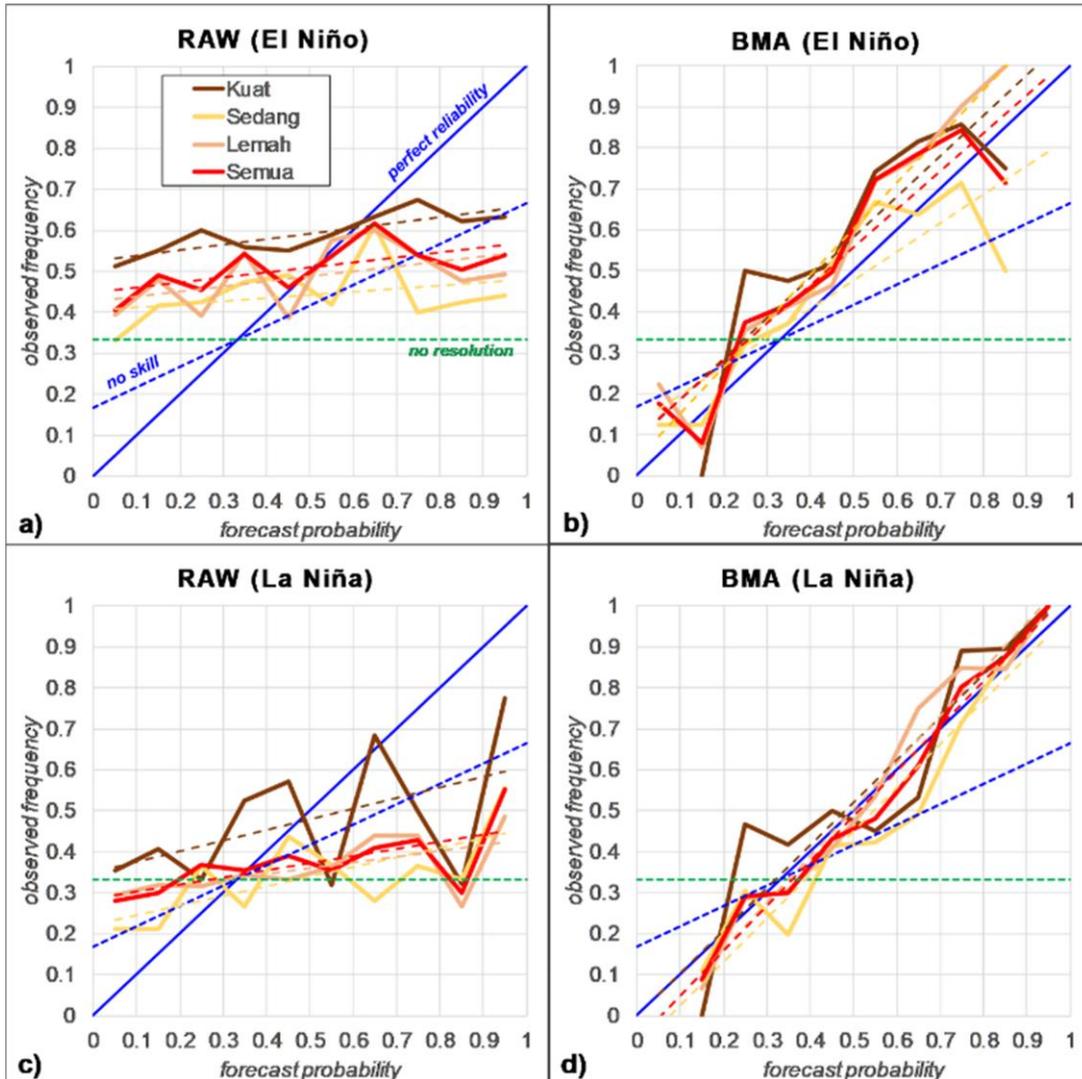
Tidak hanya *skill* yang meningkat, BMA juga memperbaiki reliabilitas RAW model. Gambar 7

menunjukkan, pada fase El Niño dan La Niña kurva *reliability* BMA mendekati garis *perfect reliability* dibanding RAW yang cenderung berada pada garis *no skill* dan *no resolution*. BMA mampu memberikan konsistensi statistik antara prediksi probabilitas kejadian BN (AN) saat El Niño (La Niña) terhadap frekuensi kejadian BN (AN) sebenarnya (Wilks 2018). Kemudian, berdasarkan kategori reliabilitas menurut Weisheimer and Palmer, (2014), kurva *reliability* BMA pada Gambar 7b (El Niño) dan Gambar 7c (La Niña) memiliki *Weisheimer Score* (WS) pada kelas “sempurna” dan “sangat berguna”, sedangkan RAW memiliki WS pada kelas “tidak berguna”. Ini berarti setiap informasi peluang BN (AN) saat El Niño (La Niña) yang dihasilkan dari kalibrasi dengan metode BMA, sangat berguna dan bahkan disarankan dipakai dalam pengambilan keputusan terkait iklim, khususnya pada kondisi ekstrem saat El Niño (La Niña) yang cenderung menyebabkan Pulau Jawa lebih kering (basah) (Supari *et al.* 2018).



Gambar 6. Perbandingan rata-rata nilai (a, c) BS dan BSS; (b, d) AUC ROC dan ROCSS saat El Niño dan La Niña untuk semua stasiun di Pulau Jawa. Nilai BS dan AUC ROC (BSS dan ROCS) dinyatakan dalam grafik batang (garis). Warna abu-abu (hitam) untuk RAW (BMA). Kategori ENSO “Semua” berarti 119 (124) kejadian El Niño (La Niña) dilibatkan dalam perhitungan semua skor, sedangkan untuk kategori “Kuat”, “Sedang”, “Lemah” berturut-turut 34 (17), 31 (38) dan 58 (75) kejadian El Niño (La Niña)

Figure 6. Comparison of average of (a, c) BS and BSS; (b, d) AUC ROC and ROCSS when El Niño and La Niña for all station over Java Island. BS and AUC ROC (BSS and ROCS) shown by bar (line) graph. Grey (black) color indicates RAW (BMA). “All” ENSO category means 119 (124) El Niño (La Niña) events are involved for calculation all score, meanwhile “Strong”, “Moderate” and “Weak” category, respectively 34 (17), 31 (38) dan 58 (75) of El Niño (La Niña) events are involved



Gambar 7. Perbandingan *Reliability Diagram* untuk (a,c) RAW dan (b,d) BMA saat El Niño dan La Niña. Garis *dash* biru (hijau) menunjukkan posisi reliabilitas *no skill (no resolution)* sedangkan garis *solid* biru adalah posisi *perfect reliability*. Kurva *Reliability* tiap kategori ENSO Kuat (coklat tua), Sedang (kuning), Lemah (krem) dan Semua (merah). Garis *dash* yang menempel pada kurva *Reliability* adalah *fitting* regresi linier setiap kurva sebagai representasi *Weisheimer Score*

Figure 7. Comparison of *Reliability Diagram* for (a, c) RAW and (b, d) BMA when El Niño and La Niña. Dash blue (green) line show *no skill (no resolution)* reliability line meanwhile solid blue line line show *perfect reliability*. Reliability curve for each ENSO category “Strong”, “Moderate”, “Weak”, “All” indicate by colors dark brown, yellow, cream and red respectively. Dash line over the reliability curve is a *fitting linear regression* line to identify *Weisheimer Score*

### Pengaruh Topografi di Pulau Jawa terhadap Kemampuan BMA

Pada Tabel 2, nilai BS dan BSS menunjukkan bahwa BMA mampu mengurangi rata-rata kesalahan untuk prediksi probabilistik kejadian BN (AN) saat El Niño (La Niña). BMA juga memperbaiki *skill* prediksi *ensemble* RAW model S4. Akan tetapi pada beberapa stasiun seperti CTO dan DMG, BMA tidak mampu memperbaiki kualitas RAW model S4 di stasiun tersebut. CTO dan DMG terletak di Jawa Barat (Gambar 1), kedua stasiun ini

berada pada topografi tinggi (pegunungan), dengan ketinggian berturut-turut 920 m dpl dan 207 m dpl. Pada kenyataannya, tidak dua stasiun itu saja, namun stasiun lain di Jawa Timur, seperti KRO, TRS, KRS dan NGJ juga mempunyai ketinggian >200 m dpl. Secara visual pada Gambar 1 jelas bahwa kompleksitas topografi di Jawa Barat lebih besar dibandingkan Jawa Timur. Oleh karena itu diasumsikan bahwa kegagalan BMA untuk stasiun CTO dan DMG dipengaruhi kompleksitas topografi di Jawa Barat.

Hasil ini selaras dengan temuan Ratri *et al.* (2019), pada kajiannya menggunakan metode EQM sebagai *post-processing*. Terdapat perbedaan *skill* antara wilayah Jawa Timur dibandingkan Jawa Barat. Pada kajian Ratri *et al.* (2019), metode kalibrasi EQM masih menyisakan *skill* negatif untuk wilayah Jawa Barat, khususnya di sekitar lokasi stasiun CTO dan DMG. Padahal RAW model yang digunakan pada Ratri *et al.* (2019), sudah menggunakan model prediksi musim ECMWF SEAS5 yang merupakan versi terbaru dari S4 di penelitian ini. Artinya pemutakhiran model dinamis yang dipakai untuk meningkatkan kemampuan prediksi musim S4 ke SEAS5 (Johnson *et al.* 2019), tidak memberikan dampak pada wilayah dengan topografi kompleks. Hal ini juga dikonfirmasi pada hasil verifikasi produk ECMWF yang dilaporkan oleh Hewson (2018). Pada laporan tersebut dijelaskan bahwa model dinamis yang dikembangkan ECMWF untuk prediksi jangka menengah (~46 hari) hingga musiman (~ 7 bulan) tidak mampu mensimulasikan parameter cuaca yang terlibat pada proses konveksi skala lokal seperti suhu dipermukaan (T2m), arah dan kecepatan angin 10 m, khususnya wilayah dengan kompleks topografi dan memiliki pantai. Meskipun laporan dari Hewson (2018) ditujukan untuk negara-negara sub-tropis di Eropa, namun kondisi topografi negara-negara yang didiskusikan pada laporan tersebut relatif mirip dengan pulau Jawa yang memiliki pegunungan dan pantai. Hewson (2018) juga melaporkan penggunaan *dynamical downscaling* yang disebut *Limited Area Model* (LAM) dengan resolusi hingga 2 Km<sup>2</sup> mampu memberikan hasil prediksi yang lebih baik dibanding penggunaan keluaran langsung produk ECMWF ataupun metode *post-processing* statistik.

## Kesimpulan

Hasil kalibrasi BMA bervariasi pada tiap stasiun di Pulau Jawa, namun secara umum BMA dapat digunakan sebagai metode alternatif *post-processing* statistik untuk meningkatkan kualitas prediksi RAW model S4, khususnya pada saat akan diprediksi terjadinya El Niño ataupun La Niña. Prediksi deterministik dan probabilistik yang dihasilkan oleh PDF prediktif BMA lebih akurat, rata-rata kesalahan lebih rendah, memiliki *skill* lebih tinggi serta reliabilitas “sangat berguna” dan “sempurna”. Oleh karena itu, penerapan BMA pada RAW model S4 sangat disarankan agar prediksi hujan musiman yang dihasilkan mempunyai kualitas yang lebih baik sehingga prediksi tersebut dapat membantu berbagai sektor, seperti pertanian, perkebunan dan pengairan dalam proses pengambilan keputusan yang berkaitan dengan

penggunaan informasi iklim pada waktu yang akan datang.

## Ucapan Terimakasih

Penulis mengucapkan terimakasih kepada Dr. Tri Wahyu Hadi, Sapto Wahyu Indratno, Ph. D dan Dr. rer.pol. Heri Kuswanto, M.Si atas bimbingan serta diskusi terkait metode BMA. Penulis juga mengucapkan terimakasih kepada *anonim reviewer* atas saran dan masukan terhadap tulisan ini. Kedua penulis adalah kontributor utama dalam penulisan naskah ini.

## Daftar Pustaka

- Aldrian E, Susanto R. 2003. Identification of three dominant rainfall regions within Indonesia and their relationship to sea surface temperature. *Int. J. Climatol.* 23:1435–1452. Doi: 10.1002/joc.950.
- Alespan G, Minobe S. 2020. Relations between Interannual Variability of Regional-Scale Indonesian Precipitation and Large-Scale Climate Modes during 1960–2007. *J. Clim.* 33:5271–5291. Doi: 10.1175/JCLI-D-19-0811.1.
- As-syakur AR, Adnyana IWS, Mahendra MS, Arthana IW, Merit IN, Kasa IW, Ekayanti NW, Nuarsa IW, Sunarta IN. 2014. Observation of spatial patterns on the rainfall response to ENSO and IOD over Indonesia using TRMM Multisatellite Precipitation Analysis (TMPA) *Int. J. Climatol.* 34:3825–3839. Doi: 10.1002/joc.3939.
- Buizza R. 2018. Ensemble Forecasting and the Need for Calibration (Chapter 2), in: *Statistical Postprocessing of Ensemble Forecasts*. Elsevier. 15–48. Doi: 10.1016/B978-0-12-812372-0.00002-9.
- BMKG. 2018. *Buku Prakiraan Musim Kemarau 2018 untuk wilayah Indonesia*. Badan Meteorologi Klimatologi dan Geofisika. 130 Hlm
- Erickson MJ, Colle BA, Charney JJ. 2012. Impact of Bias-Correction Type and Conditional Training on Bayesian Model Averaging over the Northeast United States. *Weather Forecast.* 27:1449–1469. Doi: 10.1175/WAF-D-11-00149.1.
- Gubler S, Sedlmeier K, Bhend J, Avalos G, Coelho CAS, Escajadillo Y, Jacques-Coper M, Martinez R, Schwierz C, de Skansi M, Spirig Ch. 2019. Assessment of ECMWF SEAS5 Seasonal Forecast Performance over South America. *Weather Forecast.* 35:561–584. Doi: 10.1175/WAF-D-19-0106.1.
- Hemri S, Liniger MA, Manzanar R. 2020. How to create an operational multi-model of seasonal forecasts? *Clim. Dyn.* 55:1141–1157. Doi: 10.1007/s00382-020-05314-2.

- Hewson T. 2018. Use and Verification of ECMWF products in Member and Co-operating States (2018). ECMWF Tech. Memo. 840, 40p.
- Johnson SJ, Stockdale TN, Ferranti L, Balmaseda MA, Molteni F, Magnusson L, Tietsche S, Decremet D, Weisheimer A, Balsamo G, Keeley SPE, Mogensen K, Zuo H, Monge-Sanz BM. 2019. SEAS5: the new ECMWF seasonal forecast system. *Geosci. Model Dev.* 12:1087–1117. Doi: 10.5194/gmd-12-1087-2019.
- Jolliffe IT, Stephenson DB. 2003. *Forecast verification: a practitioner's guide in atmospheric science*. J. Wiley, Chichester, West Sussex, England. Hoboken, NJ. 240p.
- Kim HM, Webster PJ, Curry JA. 2012. Seasonal prediction skill of ECMWF System 4 and NCEP CFSv2 retrospective forecast for the Northern Hemisphere Winter. *Clim. Dyn.* 39:2957–2973. Doi: 10.1007/s00382-012-1364-6.
- Kolachian R, Saghafian B. 2019. Deterministic and probabilistic evaluation of raw and post processed sub-seasonal to seasonal precipitation forecasts in different precipitation regimes. *Theor. Appl. Climatol.* 137:1479–1493. Doi: 10.1007/s00704-018-2680-5.
- Manzanas R, Gutiérrez JM, Bhend J, Hemri S, Doblas-Reyes FJ, Penabad E, Brookshaw A. 2020. Statistical adjustment, calibration and downscaling of seasonal forecasts: a case-study for Southeast Asia. *Clim. Dyn.* 54:2869–2882. Doi: 10.1007/s00382-020-05145-1.
- Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T, Vitart F. 2011. The new ECMWF seasonal forecast system (System 4). ECMWF Tech. Memo. 656, 51p.
- Muharsyah R, Hadi TW, Indratno SW. 2020. Configuration of Bayesian Model Averaging Training Window to Improve Seasonal Rainfall Ensemble Prediction. *IOP Conf. Ser. Earth Environ. Sci.* 572 (012034). Doi: 10.1088/1755-1315/572/1/012034.
- Murphy AH. 1993. What Is A Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather Forecast.* 8:281–293. Doi: 10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2.
- Ogutu GEO, Franssen WHP, Supit I, Omondi P, Hutjes RWA. 2017. Skill of ECMWF system-4 ensemble seasonal climate forecasts for East Africa. *Int. J. Climatol.* 37:2734–2756. Doi: 10.1002/joc.4876.
- Raftery AE, Gneiting T, Balabdaoui F, Polakowski M. 2005. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Mon. Weather Rev.* 133. 20p. Doi: 10.1175/MWR2906.1.
- Ratri DN, Whan K, Schmeits M. 2019. A Comparative Verification of Raw and Bias-Corrected ECMWF Seasonal Ensemble Precipitation Reforecasts in Java (Indonesia). *J. Appl. Meteorol. Climatol.* 58:1709–1723. Doi: 10.1175/JAMC-D-18-0210.1.
- Sloughter, Gneiting T, Raftery AE. 2009. Probabilistic Wind Speed Forecasting Using Ensembles and Bayesian Model Averaging. *J. Am. Stat. Assoc.* 105:25–35. Doi: 10.1198/jasa.2009.ap08615.
- Sloughter, Raftery AE, Gneiting T, Fraley C. 2007. Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging. *Mon. Weather Rev.* 135:3209–3220. Doi: 10.1175/MWR3441.1.
- Supari, Muharsyah R, Niken W. 2016. Impact of the 2015 Godzilla El Niño event on the Indonesian rainfall. *Scientific Journal of PPI-UKM.* 3:26–31. Doi: 10.21752/sjppi-ukm/se/a18072016.
- Supari, Tangang F, Salimun E, Aldrian E, Sopaheluwakan A, Juneng L. 2018. ENSO modulation of seasonal rainfall and extremes in Indonesia. *Clim. Dyn.* 51:2559–2580. Doi: 10.1007/s00382-017-4028-8.
- Themeßl MJ, Gobiet A, Leuprecht A. 2011. Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int. J. Climatol.* 31: 1530–1544. Doi: 10.1002/joc.2168.
- Tompkins AM, Ortiz De Zárate MI, Saurral RI, Vera C, Saulo C, Merryfield WJ, Sigmund M, Lee WS, Baehr J, Braun A, Butler A, Déqué M, Doblas-Reyes FJ, Gordon M, Scaife AA, Imada Y, Ishii M, Ose T, Kirtman B, Kumar A, Müller WA, Pirani A, Stockdale T, Rixen M, Yasuda T. 2017. The Climate-System Historical Forecast Project: Providing Open Access to Seasonal Forecast Ensembles from Centers around the Globe. *Bull. Am. Meteorol. Soc.* 98:2293–2301. Doi: 10.1175/BAMS-D-16-0209.1.
- Vrugt JA, Diks CGH, Clark MP. 2008. Ensemble Bayesian model averaging using Markov Chain Monte Carlo sampling. *Environ. Fluid Mech.* 8:579–595. Doi: 10.1007/s10652-008-9106-3.
- Wang QJ, Schepen A, Robertson DE. 2012. Merging Seasonal Rainfall Forecasts from Multiple Statistical Models through Bayesian Model Averaging. *J. Clim.* 25:5524–5537. Doi: 10.1175/JCLI-D-11-00386.1.
- Weisheimer A, Palmer TN. 2014. On the reliability of seasonal climate forecasts. *J. R. Soc. Interface* 11:20131162. Doi: 10.1098/rsif.2013.1162.
- Wilks DS. 2018. Enforcing calibration in ensemble postprocessing: Enforcing Calibration in Ensemble Postprocessing. *Q. J. R. Meteorol. Soc.* 144:76–84. Doi: 10.1002/qj.3185.

