

# APLIKASI ROC UNTUK UJI KEHANDALAN MODEL HYBMG

Kadarsah

Puslitbang BMKG

Email:kadarsah@yahoo.com

## ABSTRAK

*Relative Operating Characteristics (ROC) merupakan metode yang direkomendasikan oleh World Meteorological Organization (WMO) sebagai metode yang dapat mengindikasikan kemampuan probabilistik prediksi cuaca dan iklim. Metode ini diterapkan pada model HyBMG untuk menguji kehandalannya dengan cara memplotkan hit dan false-alarm rate. Uji kehandalan model ini menggunakan data curah hujan 10 tahun (1998-2007) pada 28 Daerah Prakiraan Musim (ZOM). Hasilnya berupa kurva ROC yang menunjukkan tingkat kehandalan model HyBMG dalam memprediksi curah hujan. Model HyBMG memiliki kehandalan dalam memprediksi curah hujan di beberapa daerah ZOM. Kurva ROC menunjukkan variasi kualitas sistem prediksi pada tingkat kepercayaan peringatan yang berbeda (probabilistik prediksi) dan dapat digunakan untuk optimalisasi nilai prediksi tertentu dari table kontigensi.*

Kata kunci: ROC, verifikasi prediksi, tabel kontingensi

## ABSTRACT

*The relative operating characteristic (ROC), is being considered by the World Meteorological Organization as a recommended method to indicate the skill of probabilistic weather and climate forecasts. Reliability evaluation of HyBMG model has been done by using Relative Operating Characteristics (ROC) which is created by plotting the hit and false-alarm rate. The Evaluation model is use rainfall data from only 28 climate regions over 10 years from 1998 to 2007. The result is ROC's curve that describes the reliability HyBMG to predict rainfall. HyBMG has a reliability to predict the rainfall in a particular region. The ROC curve illustrates the varying quality of the forecast system at different levels of confidence in the warning (the forecast probability) and can be used to optimize forecast value given the specifics of an individual user's contingency table.*

Key words: ROC, forecast verification, contingency table

---

Naskah masuk : 13 April 2010

Naskah diterima : 20 Juni 2010

---

## I. PENDAHULUAN

Model prediksi curah hujan banyak dikembangkan dalam dunia meteorologi diseluruh dunia, secara garis besar terdapat dua jenis model prediksi berdasarkan metode atau teknik yang digunakan yaitu, model dinamik dan model statistik. Model dinamik lebih menfokuskan pada prediksi berdasarkan proses fisis yang terjadi di atmosfer dengan memodelkannya dan men-downscaling hingga

resolusi tinggi. Teknik ini cukup menggambarkan kondisi sebenarnya di atmosfer, akan tetapi diperlukan sumber daya yang sangat besar karena model yang dijalankan memerlukan superkomputer yang mahal. Sedangkan model statistik lebih murah dibandingkan dengan model dinamik karena memerlukan sumberdaya yang relatif kecil sehingga metode statistik banyak digunakan dalam penelitian iklim<sup>1,2</sup>. Data yang digunakan pada model statistik adalah data historis

misalnya data curah hujan untuk memprediksi intensitas curah hujan di masa yang akan datang<sup>1</sup>. Beberapa contoh model dinamik, diantaranya MM5, DARLAM, WRF, ARPS, Eta Model, CCAM, NASA GMAO model, JMA model, Lamont-Doherty model, KMA model, ECHAM/MOM, COLA ANOM, METEO FRANCE model, COLA CCSM3 model, POAMA, ECMWF model, UKMO model, dan model lainnya. Contoh model statistik diantaranya ANFIS, Transformasi Wavelet, ARIMA, Kalman Filter, Markov model, Linear Inverse, CLIPER, Regression, Constructed Analog dan Neural Network. Sedangkan model statistik yang sedang dikembangkan oleh Puslitbang BMKG adalah ARIMA, Wavelet, ANFIS dan Tisean yang kemudian diintegrasikan menjadi Ensemble Prediction System (EPS) yang disebut HyBMG. Saat ini versi HyBMG yang terbaru adalah versi 2.0.7 yang masih terus dikembangkan dengan dilengkapi metode statistik lainnya yang sesuai.

### 1.1. Model HyBMG

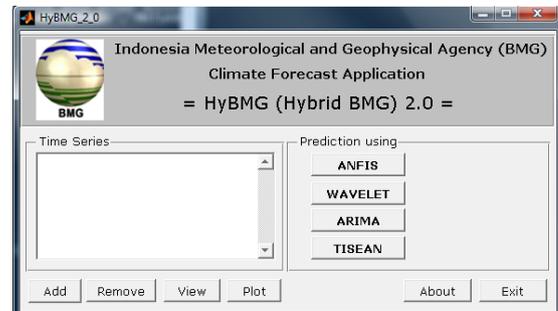
Model yang digunakan dan masih terus di kembangkan di Puslitbang BMKG dalam bidang iklim adalah HyBMG. HyBMG (Hybrid BMG) merupakan aplikasi antarmuka windows berbasis perangkat lunak MATLAB (MATrix LABoratory) menggunakan PC tunggal. HyBMG adalah kompilasi model-model statistik non-konvensional yang menggabungkan beberapa teknik prakiraan time series diantaranya neural network (ANFIS), transformasi wavelet, AutoRegressive Integrated Moving Average (ARIMA) dan non-linear dynamics (teori chaotic)/TISEAN. Data input yang digunakan HyBMG adalah data time series curah hujan. HyBMG digunakan untuk prakiraan jangka panjang (1 tahun ke depan) yang didalamnya juga terdapat fitur untuk melakukan validasi model. Tampilan awal ketika model ini digunakan seperti ditunjukkan Gambar.1. Tampilan ini juga disertai tombol-tombol yang menunjukkan metode yang digunakan. Model HyBMG merupakan sistem prediksi ensemble (EPS/*Ensemble Prediction System*) karena terdiri dari empat model prediksi yang terintegrasi. Pengembangan HyBMG akan terus dilakukan dengan menambahkan metode-metode lain misalnya CCA (Canonical Correlation Analysis, PCA (Principal

Component Analysis serta menambahkan konsep multivariat.

## 1.2. Model-Model HyBMG

### 1.2.1. ARIMA

Metode ARIMA adalah metode statistik yang menggabungkan teknik Autoregressive (AR) dan Moving Average (MA) dengan melakukan



Gambar 1. Tampilan utama HyBMG

diferensiasi pada data awal untuk membuat data awal menjadi stasioner (mempunyai mean dan varian yang stabil). Metode ARIMA digunakan untuk prediksi runtun waktu (time series) yang mempunyai korelasi diantara anggotanya (tidak random) dan runtun waktu yang bersifat stasioner, jika belum stasioner maka perlu adanya stasionerisasi dengan melakukan diferensiasi biasa atau musiman (seasonal). Data curah hujan yang digunakan, biasanya diamati dalam selang waktu tertentu sehingga jarak waktu yang memisahkan dua observasi adalah sama. Istilah time-lag digunakan untuk menunjukkan selang waktu antara dua observasi. Lag-1 berarti selang waktu yang memisahkan dua observasi adalah satu satuan waktu, lag-2 adalah untuk dua satuan waktu, dan seterusnya. Hal ini mengisyaratkan bahwa data tersebut tidak bebas, artinya observasi ke-i berkorelasi dengan observasi ke-j. ARIMA melakukan prediksi dengan menggunakan metode time-series yang didasarkan pada nilai-nilai peubah atau penyimpangannya yang telah terjadi di waktu yang lalu. Tujuannya adalah menentukan pola historis data kemudian mengekstrapolasikan pola tersebut ke masa yang akan datang berdasarkan asumsi kontinuitas. Metode Arima telah umum digunakan di Indonesia khususnya BMKG dalam memprediksi curah hujan.

### 1.2.2. WAVELET

Transformasi wavelet merupakan metode untuk mendeteksi fluktuasi periodik yang bersifat transien berikut parameternya, dapat fokus pada suatu rentang waktu tertentu dari data yang ada<sup>3</sup> dan dapat menggambarkan proses dinamik nonlinear kompleks yang ditunjukkan oleh interaksi gangguan dalam skala ruang dan waktu<sup>4</sup>.

Transformasi wavelet dapat digunakan untuk menganalisa deret waktu yang mengandung daya non-stasioner pada frekwensi yang berbeda<sup>5</sup>. Curah hujan dapat dianggap sebagai sinyal yang benar-benar non-stasioner sehingga analisis curah hujan bertujuan mengetahui periodesitas dan informasi tentang kapan waktu terjadinya. Untuk itu diperlukan suatu transformasi yang dapat memberikan tampilan waktu-frekwensi dari sinyal. Transformasi wavelet dari  $x(t)$  atau sinyal hujan menghasilkan *approximation coefficient* dan *detail coefficient* dari data hujan. *Approximation coefficient* berhubungan dengan skala tinggi dan komponen frekwensi rendah, sedangkan *detail coefficient* berhubungan dengan skala rendah dan komponen frekwensi tinggi. Koefisien-koefisien ini selanjutnya direntang ke depan (*extend*) sepanjang selang waktu tertentu.

### 1.2.3. ANFIS

ANFIS merupakan suatu teknik optimasi yang menggabungkan konsep neural network dengan fuzzy logic. Neural-network mengenal pola-pola dan menyesuaikan pola terhadap perubahan lingkungan, sedangkan fuzzy logic menggabungkan pengetahuan manusia dan mencari kesimpulan untuk membuat suatu keputusan. ANFIS juga dapat diterapkan secara langsung dalam bidang pemodelan, pengambilan keputusan, pengolahan sinyal, dan kontrol<sup>6</sup>.

### 1.2.4. TISEAN

Tisean merupakan metode untuk analisis data time series berdasarkan teori sistem dinamik deterministik nonlinear atau teori chaos. Metode ini masih dalam tahap pengembangan dalam HyBMG.

### 1.2.5. Kurva ROC dan Skor ROC

ROC adalah sebuah metode untuk menguji kemampuan sistem prediksi berdasarkan tabel

kontingensi<sup>7,8,9</sup>. ROC menampilkan skill sistem prediksi dengan membandingkan hit rate dan false alarm rate<sup>7</sup>. Definisi false alarm rate harus dibedakan dengan false alarm ratio<sup>10,11,12</sup>, tetapi dalam beberapa literature keduanya mengacu pada false alarm rate<sup>13</sup>. Derivasi dari ROC didasarkan pada tabel kontingensi memberikan *hit rate* dan false alarm rate untuk deterministik atau prediksi probabilistik. Suatu kejadian didefinisikan sebagai biner, yang berarti bahwa hanya ada dua hasil yang mungkin, suatu kejadian atau non-kejadian. Selain ROC juga digunakan metode atau cara lain untuk verifikasi<sup>14,15,16</sup> yang semuanya berdasarkan konsep bagaimana prediksi yang baik<sup>17</sup> serta dengan melibatkan pengetahuan di bidang atmosfer<sup>18</sup> serta fenomena alam lainnya, misal *sea breeze*<sup>19</sup>.

Peluang prediksi adalah kemungkinan suatu peristiwa akan terjadi, sedangkan probabilitas skor ROC hanya dihitung menggunakan *Hit Rate* dan *False Alarm Rate*. HR dan FAR dihitung untuk setiap rentang probabilitas. Dalam hal ini suatu peristiwa dikatakan diprediksi terjadi pada suatu titik jika peluang prediksi untuk sebuah event terjadi dalam rentang peluang. (misalnya, sebuah prediksi curah hujan atas normal memiliki 43% kemungkinan akan jatuh dikisaran peluang 40-50%) Ditinjau dari kejadian (yakni, *Hits*) berapa kali peluang prediksi berada dalam rentang peluangnya, sedangkan peristiwa yang tidak terjadi (yaitu, *Miss*) adalah berapa kali suatu prediksi dibuat untuk peluangnya bahwa prediksi tersebut salah. Hit rate sering disebut POD (*probability of detection*)<sup>9,10,13,20,21,22</sup> dan memberikan perkiraan probabilitas bahwa suatu kejadian akan diperingatkan dan menghasilkan suatu estimasi probabilitas kejadian yang akan diperingatkan. Untuk peluang prediksi, peringatan dapat diterbitkan ketika probabilitas prediksi untuk kejadian yang sudah ditetapkan melebihi beberapa ambang batas<sup>23</sup>. Kurva ROC dapat digunakan dalam membantu untuk mengidentifikasi strategi optimal tertentu dalam berbagai bidang yang memungkinkan<sup>9</sup> tentunya dengan mempertimbangkan interaksi antara riset dan

operasional yang melibatkan verifikasi model numerik secara subjektif<sup>24</sup>.

HR dan FAR dihitung untuk setiap peluang kritis, kemudian hasilnya berupa plot titik-titik pada grafik HR (sumbu vertikal) terhadap FAR (sumbu horisontal) sehingga terbentuk kurva ROC. Kurva ini harus melalui titik (0,0) dan (1,1). Prediksi yang tidak memiliki skill ditandai dengan garis diagonal (HR = FAR); semakin jauh kurva ke arah atas pojok kiri (di mana HR = 1 dan FAR = 0) menunjukkan skill yang lebih baik. Luas dibawah kurva ROC sering digunakan secara statistik untuk menunjukkan kemampuan sistem prediksi. Luas area dibawah kurva merupakan ukuran kemampuan skill sistem prediksi. Sehingga sistem prediksi yang sempurna memiliki luas wilayah satu, kurva yang terletak sepanjang diagonal memiliki luas wilayah 0,5 dan sistem prediksi dikatakan tidak memiliki skill sebab ketepatan prediksi sama banyaknya dengan kegagalan prediksi. Jika kurva ROC berada dibawah garis diagonal sistem prediksi disebut memiliki skill negatif artinya bahwa sistem tersebut memiliki kegagalan prediksi lebih banyak dibanding keberhasilannya.

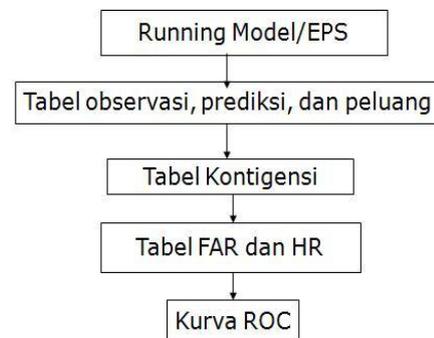
### 1.2.6. Plot Kurva ROC

Terdapat dua cara untuk memplot grafik *Hit Rate False vs alarm Rate* untuk suatu probabilitas ramalan. Pertama, dengan membuat tabel kontingensi untuk setiap probabilitas ambang batas (yaitu satu peluang untuk satu rentang tertentu, > 90%, > 80%, > 70% dll) kemudian plot masing-masing Hit Rate vs False Alarm Rate pada grafik yang sama. Kedua metode yang lebih cepat, dengan cara menghitung jumlah perkiraan benar peristiwa yang terjadi di masing-masing dari beberapa interval probabilitas (misalnya, 90-100%, 80-90%, 70-80% dll) dan juga menghitung jumlah "nonprediksi" kejadian dalam interval probabilitas yang sama. Kemudian *hit rate* untuk probabilitas tertentu (misalnya 80%) dihitung dengan cara jumlah kejadian diatas tingkat tersebut dibagi dengan total jumlah kejadian yang diprediksi. Demikian pula, *false alarm rate* adalah jumlah

"non prediksi" kejadian dengan kriteria diatas ambang dibagi dengan total jumlah kejadian yang non-prediksi. Jika pada awalnya metode kedua ini tidak muncul identik secara intuitif dengan metode pertama, hanya ingat bahwa, misalnya, untuk *Hit Rate* (yang sama dengan  $Hits / (Hits + Misses)$ ) setiap *Misses* di setiap probabilitas interval (katakanlah 80 -- 90%) dihitung pada interval probabilitas yang lebih rendah. Karena itu *Misses* dimasukkan dalam penyebut secara default, dan karena itu dua metode memberikan jawaban yang sama.

## II. DATA DAN METODE

Data yang digunakan terdiri dari data curah hujan (CH) dasarian 10 tahun (1998-2007) pada 28 Zona Musim (ZOM) (Tabel 1) dan data CH dasarian selama 30 tahun (1971-2000) dijadikan data normal sebagai pembanding keluaran model HyBMG dari masing-masing ZOM. HyBMG dijadikan sebagai *Ensemble Prediction System* (EPS). Pemilihan 28 ZOM ini berdasarkan kelengkapan data CH. Sedangkan penggunaan data CH dasarian berdasarkan kenyataan bahwa data tersebut digunakan untuk prakiraan bulanan, tiga bulanan dan musim HyBMG. Kriteria sifat hujan ditentukan berdasarkan ketentuan BMKG yang menyatakan bahwa CH diatas normal jika CH keluaran model > 115 % CH normalnya, CH dibawah normal jika CH keluaran model < 85 % CH normalnya sedangkan CH keluaran model normal jika CH berkisar 85-115 % dari CH normalnya. Diagram alir pembuatan kurva ROC ditunjukkan Gambar 2. Langkah pertama adalah menjalankan HyBMG dengan masukan data CH dasarian di 28 ZOM, hasilnya kemudian dibuat tabel observasi, prediksi dan peluang seperti yang ditunjukkan Tabel 2 dengan N=Normal,



Gambar 2. Diagram alir metodologi pembuatan Kurva ROC

A= atas normal, B= bawah normal. Kolom prediksi menunjukkan rata-rata hasil prediksi sedangkan kolom peluang menunjukkan besarnya masing-masing peluang untuk B,N dan A. Jadi pada dasarian ke-1, observasi dan prediksi menunjukkan CH yang normal sedangkan peluangnya menunjukkan 30 % untuk peluang CH bawah normal dan normal sedangkan untuk atas normal sebesar 40 %.

Tabel 1. Nama 28 ZOM

ZOM	N	ZOM
Banda Aceh	1	Kotabaru
Bandung Geof	2	Pacitan
Banyuwangi	3	Ruteng
Bawean	4	Palembang
Bima	5	Pangkal Pinang
BMG Jakarta	6	Polonia
Cilacap	7	Madiun
Denpasar	8	Majene
Jatiwangi	9	Makassar
Juanda	10	Manado
Kalianget	11	Sabang
Lampung	13	Rengat
Larantuka	14	Samarinda

Tabel 2 tersebut hanya menunjukkan untuk satu ZOM saja sehingga untuk 28 ZOM maka diperlukan 28 tabel dan tahapan yang berbeda. Langkah selanjutnya adalah membuat tabel kontigensi 2 x 2. Tabel kontigensi 2 x 2 diperuntukkan untuk prediksi yang memiliki kategori dikotomi. Terdapat hanya dua kemungkinan misal hujan atau tidak hujan, angin kuat dan tidak kuat. Dalam penelitian ini kategori yang digunakan dalam tiap ZOM adalah CH atas normal-tidak atas normal serta CH bawah normal-tidak bawah normal.

Tabel 2. Tabel data rata-rata dasarian 10 tahun (1998-2007) : observasi, prediksi dan peluang untuk ZOM Banda Aceh

Dasarian	Observasi	Prediksi	Peluang		
			B	N	A
1	N	N	30	30	40
2	N	A	20	40	40
3	A	A	0	20	80
4	B	A	10	20	70
5	A	N	40	30	30
6	N	B	60	30	10
7	N	B	50	30	20
8	N	N	30	30	40
9	A	A	10	30	60
10	A	A	20	40	40
11	B	B	50	40	10
12	A	A	10	30	60
13	N	N	20	60	20
14	N	B	90	10	0
15	B	B	100	0	0
16	B	B	70	20	10
17	N	B	100	0	0
18	B	N	20	70	10
19	B	N	30	50	20
20	N	B	70	30	0
21	N	N	40	30	30
22	B	B	40	60	0
23	B	B	60	30	10
24	B	N	20	60	20
25	A	A	10	30	60
26	A	A	0	20	80
27	B	B	60	40	0
28	A	A	10	30	60
29	N	N	20	60	20
30	N	B	90	10	0
31	B	B	100	0	0
32	B	B	70	20	10
33	N	B	100	0	0
34	B	N	20	70	10
35	B	N	30	50	20
36	N	B	70	30	0

Bentuk dasar tabel kontigensi seperti Tabel 3. Tabel tersebut memuat kriteria yang disebut *hit* (jika prediksi tepat dengan observasi, dilambangkan a), *false alarm* (jika prediksi menyatakan terjadi/ya sedangkan observasi tidak terjadi, b), *miss* (jika prediksi menyatakan tidak terjadi sedang observasi menyatakan terjadi,c) dan *correct non-event* (jika prediksi dan observasi menyatakan tidak terjadi, d). Nilai-nilai inilah yang dijadikan dasar untuk membuat tabel FAR (False Alarm

Rate) dan HR (Hit Rate) ditunjukkan masing-masing persamaan 1 dan 2.

$$FAR = b / (a+b) \dots \dots \dots (1)$$

$$HR = a / (a+c) \dots \dots \dots (2)$$

Tabel 3. Tabel kontigensi 2 x 2

Kriteria		Observasi		Total
		Ya	Tidak	
Prediksi	Ya	a	b	a+b
	Tidak	c	d	c+d
Total		a+c	b+d	a+b+c+d=n

Tabel 4 menunjukkan tabel kontigensi untuk ZOM Banda Aceh untuk kategori bawah dan atas normal.

Tabel 4. Tabel kontigensi ZOM Banda Aceh a. atas normal

Atas Normal		Observasi		Total
		A	TA	
Prediksi	A	7	2	9
	TA	1	26	27
Total		8	28	36

b. bawah normal

Bawah normal		Observasi		Total
		B	TB	
Prediksi	B	8	8	16
	TB	6	14	20
Total		14	22	36

Langkah selanjutnya adalah menyusun tabel kontigensi pada tiap peluang kritis 0-100 % pada kategori bawah dan atas normal. Tabel kontigensi saat peluang kritis 80 % ditunjukkan Tabel 5 untuk atas dan bawah normal. Hasil tiap peluang kritis tersebut dicari nilai FAR dan HR kemudian ditabelkan untuk atas dan bawah normal sehingga dihasilkan tabel 6. Kurva ROC dibuat dengan memplotkan antara nilai *hit rate* dengan nilai *false alarm rate* dimana nilai hit dan false alarm rate disusun dari satu set nilai peluang kritis yang digunakan dalam menentukan apakah suatu peringatan akan diberikan atau tidak, misalnya peringatan akan diberikan jika peluang terjadinya sama atau di atas 80%. Tabel 5 menunjukkan bahwa pada peluang kritis 80 % nilai FAR, HR berturut-turut adalah 0 dan 0.25 untuk batas atas curah hujan sedangkan FAR, HR untuk batas bawah curah hujan adalah 0.67 dan 0.14. Kondisi tersebut dapat diartikan bahwa pada

peluang kritis 80 % curah hujan yang sesuai dengan observasi lebih besar pada saat memprediksi curah hujan atas normal disbanding saat memprediksi curah hujan dibawah normal.

Tabel 5. Tabel kontigensi ZOM Banda Aceh saat peluang kritis 80 %

Nilai Peluang Kritis 80%		Observasi		Total
		A	TA	
Prediksi	A	2	0	2
	TA	6	28	34
Total		8	36	36

Nilai Peluang Kritis 80%		Observasi		Total
		B	TB	
Prediksi	B	2	4	6
	TB	12	18	30
Total		14	22	36

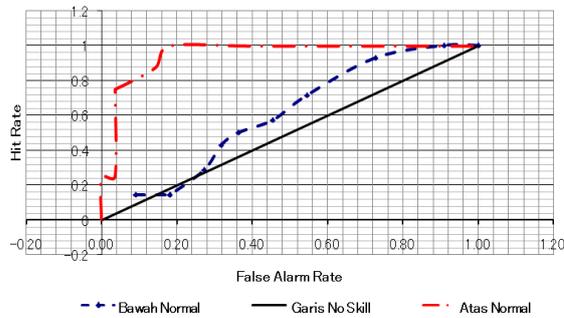
Nilai FAR dan HR pada Tabel 6 menunjukkan bahwa prediksi curah hujan atas normal lebih baik dibanding saat memprediksi curah hujan bawah normal. Kondisi yang lebih jelas terlihat saat nilai FAR dan HR di plotkan menjadi suatu kurva yang disebut Kurva ROC sehingga analisisnya bisa lebih mudah. Kurva ROC dalam Gambar 3 menunjukkan bahwa saat memprediksi curah hujan atas normal (garis merah) tingkat ketepatannya lebih besar dibanding saat memprediksi curah hujan bawah normal (garis biru). Garis hitam merupakan garis non skill yang menyatakan bahwa nilai FAR sama dengan HR. Kondisi tersebut menyatakan model prediksi tidak memiliki kemampuan sebab ketepatan prediksi sama dengan ketidaktepatannya.

Tabel 6. Tabel kontigensi ZOM Banda Aceh

Nilai Peluang Kritis	Batas Atas		Batas Bawah	
	FAR	HR	FAR	HR
100	0	0	0	0
90	0	0	0.6	0.14
80	0	0.25	0.67	0.14
70	0.04	0.25	0.6	0.29
60	0.04	0.75	0.32	0.43
50	0.04	0.75	0.36	0.5
40	0.07	0.88	0.36	0.57
30	0.07	0.88	0.36	0.57
20	0.07	0.88	0.36	0.57
10	0.07	0.88	0.36	0.57
0	1	1	1	1

### III. HASIL DAN DISKUSI

Kurva ROC hasil plot FAR dan HR ditunjukkan Gambar 3. Kurva ROC tersebut menggambarkan tingkat kehandalan model HyBMG saat memprediksi curah hujan atas dan bawah normal di ZOM Banda Aceh seperti yang pernah dilakukan oleh Mason<sup>25</sup> di daerah berbeda.



Gambar 3. Kurva ROC ZOM Banda Aceh

Model prediksi memiliki kemampuan dalam memprediksi jika kurva ROC berada diatas garis non skill atau memiliki skill positif seperti yang terjadi pada kurva ROC yang berwarna merah (prediksi curah hujan atas normal). Semakin jauh posisi kurva dari garis non skill semakin tinggi kemampuan model prediksi tersebut. Artinya bahwa model prediksi memiliki kemampuan prediksi dengan ketepatan yang tinggi saat memprediksi curah hujan di atas normal dibanding saat memprediksi curah hujan dibawah normal (warna biru). Rendahnya kemampuan model dalam memprediksi curah hujan dibawah normal terlihat saat kurva yang berwarna biru lebih dekat ke garis non skill bahkan terdapat nilai yang negatif saat kurvanya berada dibawah garis non skill. Hal tersebut menunjukkan bahwa model mengalami kegagalan yang lebih besar dibanding ketepatan saat memprediksi curah hujan dibawah normal pada peluang kritis yang tinggi. Tingkat kemampuan model prediksi dapat diukur dengan nilai skill yang dimiliki oleh model tersebut. Nilai skill model dapat dihitung dengan menghitung luas dibawah kurva<sup>26</sup>. Kurva ROC merah memiliki nilai skill sama dengan  $22/25 \times 100\% = 88\%$ . Sedangkan kurva ROC biru memiliki nilai skill  $16/25 \times 100\% = 64\%$ . Tabel 6 dan 7 berisi tentang parameter verifikasi untuk prediksi curah hujan atas normal dan bawah normal. Parameter tersebut dapat digunakan untuk mengukur skill atau

kehandalan suatu model sebagai bagian dari verifikasi model prediksi.

Parameter verifikasi pada Tabel 6 dan 7 merupakan parameter yang sering digunakan dalam proses verifikasi prediksi. Parameter B adalah indeks bias, dengan rentang 0 sampai tak terhingga, indeks bias yang sempurna saat nilainya mendekati satu dan akan menjadi *overforecasting* jika lebih dari satu dan *underforecasting* jika kurang dari satu. Untuk parameter indeks bias prediksi curah hujan atas normal lebih baik dibanding prediksi curah hujan dibawah normal. Parameter selanjutnya adalah PC atau proportion correct /proporsi benar dengan rentang 0-1 dengan skor sempurna mencapai satu saat  $b=c=0$ . Parameter ini menunjukkan proporsi tepat suatu prediksi dibanding total observasi. Nilai PC prediksi atas normal lebih besar (0,916) dari nilai PC bawah normal (0,611) menunjukkan bahwa proporsi tepat suatu prediksi lebih banyak pada prediksi curah

Tabel 6. Tabel nilai parameter verifikasi atas normal ZOM Banda Aceh

Parameter Verifikasi Atas Normal	Nilai
$B = (a+b)/(a+c)$	1.125
$PC = (a+d)/n$	0.916666667
$POD = a/(a+c)$	0.875
$FAR = b/(a+b)$	0.222222222
$PAG = a/(a+b)$	0.777777778
$F = b/(b+d)$	0.071428571
$KSS = POD-F$	0.803571429
$TS = a/(a+b+c)$	0.7
$ETS = (a-ar)/(a+b+c-ar)$	0.625
$HSS = 2(ad-bc)/[(a+c)(c+d)+(a+b)(b+d)]$	0.769230769
$OR = ad/bc$	91
$ORSS = (OR-1)/(OR+1)$	0.97826087
$ar$	2

Tabel 7. Tabel nilai parameter verifikasi bawah normal ZOM B.anda Aceh

Parameter Verifikasi Bawah Normal	Nilai
$B = (a+b)/(a+c)$	1.142857143
$PC = (a+d)/n$	0.611111111
$POD = a/(a+c)$	0.571428571
$FAR = b/(a+b)$	0.5
$PAG = a/(a+b)$	0.5
$F = b/(b+d)$	0.363636364
$KSS = POD-F$	0.207792208
$TS = a/(a+b+c)$	0.363636364
$ETS = (a-ar)/(a+b+c-ar)$	0.112676056
$HSS = 2(ad-bc)/[(a+c)(c+d)+(a+b)(b+d)]$	0.202531646
$OR = ad/bc$	84
$ORSS = (OR-1)/(OR+1)$	0.976470588
$ar$	6.222222222

hujan atas normal dibanding bawah normal. Parameter selanjutnya adalah H (*hit rate*)/POD (*Probability of Detection*), PAG (*Post Agreement*), dan FAR (*False Alarm Ratio*), F (*False Alarm Rate/ Probability of False Detection*) keduanya memiliki rentang 0 - 1 dengan nilai sempurna POD/FAG, FAR/F berturut-turut adalah 0 dan 1. Empat parameter tersebut menunjukkan bahwa prediksi curah hujan atas normal lebih baik dibanding prediksi curah hujan bawah normal. Parameter KSS (*Hanssen & Kuipper's Skill Score/True Skill Statistic (TSS)/Pierce's Skill Score*) merupakan hasil pengurangan POD-F. Rentang nilai -1 s.d 1, nilai sempurna 1 dan nilai 0 menunjukkan model tidak memiliki skill. Parameter ini erat berhubungan dengan ROC dan sangat tepat untuk memisahkan prediksi saat yes atau no. Nilai KSS yang tinggi menunjukkan bahwa prediksi curah hujan atas normal pun menunjukkan lebih baik dibanding bawah normal. Nilai TS (*Threat Score*)/CSI (*Critical Success Index/Gilbert Score/rasio verifikasi, Gilbert 1884*) merupakan parameter yang sangat sensitif dan sering digunakan dalam klimatologi. Parameter ini memiliki range 0 s.d 1 (non skill =0, sempurna=1). Nilai ini pun menunjukkan bahwa prediksi curah hujan atas normal lebih baik dibanding bawah normal. Nilai ETS (*Equitable Threat Score, Gilbert Skill Score (GSS)*), memiliki rentang -1/3 s.d 1 (no skill=0, skor sempurna=1). Nilai ETS ini sering digunakan untuk mengukur atau meramalkan kejadian yang diperkirakan benar, disesuaikan dengan frekuensi hits yang akan diharapkan terjadi hanya dengan peluang acak (misalnya, lebih mudah untuk prediksi terjadinya hujan dengan tepat pada iklim basah daripada dalam iklim kering).

Nilai GSS/ETS sering digunakan dalam verifikasi curah hujan dalam model NWP karena "pemerataan" memungkinkan skor menjadi lebih sesuai dibandingkan dengan metode lain yang berbeda namun tidak benar-benar merata. Peka terhadap *hits* karena prediksi yang meleset dan *false alarm* dengan cara yang sama, itu tidak membedakan sumber kesalahan prediksi. Harus digunakan dalam kombinasi dengan

setidaknya satu tabel kontingensi statistik (misalnya, Bias).

Skor HSS (*Heidke Skill Score*) memiliki rentang  $-\infty$  sampai 1 (skor sempurna 1 dan non skill 0). Nilai negatif menunjukkan peningkatan prediksi ke arah lebih baik. Skor HSS mengukur peningkatan prediksi fraksional berdasarkan standard prediksi yang ditentukan. Seperti sebagian besar skor lainnya, HSS dinormalkan dengan total rentang perbaikan berbagai kemungkinan di atas standar, yang berarti nilai Skill Heidke dapat dengan aman dibandingkan pada dataset yang berbeda. HSS adalah skor populer, sebagian karena relatif mudah untuk menghitung dan mungkin juga karena prediksi standar, kemungkinan, relatif mudah untuk mengalahkan. Skor standar lainnya mungkin, seperti *persistensi* atau klimatologi, tetapi ini membutuhkan informasi tambahan untuk menghitung, dalam bentuk tabel kontingensi yang terpisah. Prediksi curah hujan atas normal masih menunjukkan lebih baik dengan tingginya nilai skor HSS dibanding skor HSS pada Tabel 6 (prediksi curah hujan bawah normal).

Odds Ratio<sup>27</sup> digunakan mengukur probabilitas prediksi(odds) skor hit (H) dibandingkan dengan false alarm (F), independen dari bias memiliki rentang 0 s.d  $\infty$  (skor sempurna  $\infty$ , non skill 1). Hubungan OR dengan F dan H dapat ditentukan dari persamaan 3.

$$OR = \frac{\left[ \frac{H}{1-H} \right]}{\left[ \frac{F}{1-F} \right]} \quad (3)$$

Skor yang semakin tinggi menunjukkan prediksi yang lebih baik seperti ditunjukkan Tabel 5 jika dibanding Tabel 6. Skor lain yang sering digunakan dalam klimatologi tapi tidak dalam meteorologi adalah Odds Ratio Skill Score (ORSS). Skor ini menghasilkan nilai skill absolut yang sangat tinggi, memiliki rentang -1 s.d 1 dengan skor sempurna mencapai 1. Untuk menghitung ORSS dapat digunakan OR sesuai dengan persamaan 4.

$$ORSS = \frac{(ad - bc)}{(ad + bc)} = \frac{OR - 1}{OR + 1} \quad (4)$$

Nilai ORSS prediksi curah hujan atas normal masih menunjukkan lebih tinggi dibanding ORSS prediksi curah hujan bawah normal sehingga dapat disimpulkan bahwa prediksi curah hujan atas normal lebih baik dibanding bawah normal. Parameter lainnya adalah  $a_r$  yang dihitung

$$a_r = \frac{(a+b)(a+c)}{n} \quad (5)$$

berdasarkan persamaan 5. Parameter  $a_r$  digunakan untuk menghitung ETS. Parameter  $a_r$  ini digunakan saat mengukur ETS jadi hasilnya akan berguna jika ditampilkan dalam nilai ETS.

Dari proses verifikasi prediksi menggunakan HyBMG pada satu daerah ZOM yaitu Banda Aceh diketahui bahwa HyBMG memiliki kemampuan yang sangat baik lebih dari 80 % ketika memprediksi curah hujan diatas normal sedangkan saat memprediksi curah hujan dibawah normal memiliki kelemahan yaitu tidak sebaik saat memprediksi curah hujan diatas normal. Kemampuan HyBMG saat memprediksi ZOM lain juga bervariasi dalam memprediksi curah hujan atas normal dan bawah normal.

Skor ROC banyak dibahas oleh ilmuwan pada berbagai bidang khususnya untuk peluang suatu prediksi<sup>28</sup>. Proses verifikasi prediksi merupakan proses yang terus berkembang dan memerlukan metode serta teknik yang terus diperdalam<sup>29</sup> dengan memperhatikan hal-hal yang mungkin bisa mempengaruhi kualitas prediksi<sup>30, 31</sup>. Panduan yang sangat baik untuk proses verifikasi prediksi telah banyak dilakukan, misalnya oleh Jolliffe<sup>32</sup>, Mason<sup>33</sup>, Woodcock<sup>34</sup>, Wilks<sup>35</sup> sehingga hasil prediksi bisa bernilai ekonomi dan dijadikan pijakan pengambilan keputusan<sup>36, 37, 38</sup>.

#### IV. KESIMPULAN

- Relative Operating Characteristics (ROC) merupakan metode yang

direkomendasikan oleh World Meteorological Organization (WMO) sebagai metode yang dapat mengindikasikan kemampuan probabilistik prediksi cuaca dan iklim

- Kurva ROC menunjukkan variasi kualitas sistem prediksi pada tingkat kepercayaan peringatan yang berbeda (probabilistik prediksi) dan dapat digunakan untuk optimasi nilai prediksi tertentu dari table kontigensi.
- ROC dapat diterapkan pada model, ensemble model baik statistik maupun dinamik.
- Keandalan HyBMG berbeda pada tiap ZOM baik secara spasial maupun temporal dan saat memprediksi curah hujan atas normal dan bawah normal.

#### V. UCAPAN TERIMAKASIH

Penelitian ini dapat terlaksana dengan perbaikan atas saran dan masukan Pertti Nurmi, Finnish Meteorological Institute, Finlandia dan Laurie Wilson, Meteorological Research Division, Canada serta rekan-rekan Puslitbang BMKG yang telah mengembangkan HyBMG.

#### VI. DAFTAR PUSTAKA

- <sup>1</sup> Von Storch, H. and F.W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge.
- <sup>2</sup> Murphy, A.H. and D.S. Wilks, 1998: *A case study of the use of statistical models in forecast verification: Precipitation probability forecasts. Wea. Forecasting*, **13**, 795-810.
- <sup>3</sup> Torrence and G. Compo, 'A practical guide to wavelet analysis', *Bull. Amer. Met. Soc.* 79 (1998), 61--78.
- <sup>4</sup> Astaf'eva N M: "Wavelet analysis: basic theory and some applications" *Phys. Usp.* **39** 1085.
- <sup>5</sup> Daubechies, I *The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. Information Theory*, 36, 1990.
- <sup>6</sup> Jang, J.S.R., C.T. Sun and E. Mizutani, 1997. *Neuro-fuzzy and Soft*

- Computing: A Computational Approach to Learning and Machine Intelligence*. 1st Edn. Prentice Hall, Upper Saddle River, NJ. USA. ISBN: 0132610663.
- <sup>7</sup> Swets, J. A., 1973: *The relative operating characteristic in psychology*. *Science*, **182**, 990–1000.
- <sup>8</sup> Mason, I., 1982: *A model for assessment of weather forecasts*. *Aust. Meteor. Mag.*, **30**, 291–303.
- <sup>9</sup> Harvey, L. O., K. R. Hammond, C. M. Lusk, and E. F. Mross, 1992: *The application of signal detection theory to weather forecasting behavior*. *Mon. Wea. Rev.*, **120**, 863–883.
- <sup>10</sup> Doswell, C.A. III, R. Davies-Jones, and D.L. Keller, 1990: *On summary measures of skill in rare event forecasting based on contingency tables*. *Wea. Forecasting*, **5**, 576–585.
- <sup>11</sup> Schaefer, J. T., 1990: *The critical success index as an indicator of warning skill*. *Wea. Forecasting*, **5**, 570–575.
- <sup>12</sup> Harvey, L. O., K. R. Hammond, C. M. Lusk, and E. F. Mross, 1992: *The application of signal detection theory to weather forecasting behavior*. *Mon. Wea. Rev.*, **120**, 863–883.
- <sup>13</sup> Wilks, D.S., 1995: *Statistical Methods in the Atmospheric Sciences. An Introduction*. Academic Press, San Diego, 467 pp.
- <sup>14</sup> Hamill, T.M., 1997: *Reliability diagrams for multicategory probabilistic forecasts*. *Wea. Forecasting*, **12**, 736–741.
- <sup>15</sup> Hamill, T.M., 1999: *Hypothesis tests for evaluating numerical precipitation forecasts*. *Wea. Forecasting*, **14**, 155–167.
- <sup>16</sup> Briggs, W.M. and R.A. Levine, 1997: *Wavelets and field forecast verification*. *Mon. Wea. Rev.*, **125**, 1329–1341.
- <sup>17</sup> Murphy, A.H., 1993: *What is a good forecast? An essay on the nature of goodness in weather forecasting*. *Weather Forecasting* **8**:281–293.
- <sup>18</sup> Murphy, A.H and H. Daan, 1985: *Forecast evaluation. Probability, Statistics, and Decision Making in the Atmospheric Sciences*, A. H. Murphy and R. W. Katz, Eds., West view Press, 379–437.
- <sup>19</sup> Case, J.L., J. Manobianco, J. E. Lane, C.D. Immer, and F.J. Merceret, 2004: *An objective technique for verifying sea breezes in high-resolution numerical weather prediction models*. *Wea. Forecasting*, **19**, 690–705.
- <sup>20</sup> Olson, R. H., 1965: *On the use of Bayes' theorem in estimating false alarm rates*. *Mon. Wea. Rev.*, **93**, 557–558.
- <sup>21</sup> Panofsky, H. A., and G. W. Brier, 1968: *Some Applications of Statistics to Meteorology*. The Pennsylvania State University, 224 pp.
- <sup>22</sup> Murphy, A.H. and R.L. Winkler, 1987: *A general framework for forecast verification*. *Mon. Wea. Rev.*, **115**, 1330–1338.
- <sup>23</sup> Mason, I., 1979: *On reducing probability forecasts to yes/no forecasts*. *Mon. Wea. Rev.*, **107**, 207–211.
- <sup>24</sup> Kain, J.S., M.E. Baldwin, P.R. Janish, S.J. Weiss, M.P. Kay and G.W. Carbin, 2003: *Subjective verification of numerical models as a component of a broader interaction between research and operations*. *Wea. Forecasting*, **18**, 847–860.
- <sup>25</sup> Mason, S.J., and N.E. Graham, 1999. *Conditional Probabilities, Relative Operating Characteristics, and Relative Operating Levels.*, *Weather and Forecasting*, **14**, 713–725.
- <sup>26</sup> Mason, S.J., and N.E. Graham, 2002. *Areas beneath the relative operating characteristics (ROC), and relative operating levels (ROL) curves: Statistical significance and interpretation.*, *Q.J.R. Meteorol. Soc.*, **128**, 2145–2166.
- <sup>27</sup> Stephenson, D.B., 2000: *Use of the "odds ratio" for diagnosing forecast skill*. *Wea. Forecasting*, **15**, 221–232.
- <sup>28</sup> Kharin, V., and F. Zwiers. 2003. *On the ROC score of Probability Forecasts*. *Journal of Climate*, **16**, 4145–4150.
- <sup>29</sup> Casati, B., L.J. Wilson, D.B. Stephenson, P. Nurmi, A. Ghelli, M. Pocerlich, U. Damrath, E.E. Ebert, B.G. Brown and S. Mason, 2008: *Forecast verification: current status and future directions*. *Meteorol. Appl.*, **15**, 3–18.
- <sup>30</sup> Doblas-Reyes, F.J., C. A. S. Coelho, D. B. Stephenson, 2008: *How much does*

*simplification of probability forecasts reduce forecast quality?* Meteorol. Appl., **15**.

- <sup>31</sup>Murphy, A.H., 1995: *The coefficients of correlation and determination as measures of performance in forecast verification*. Wea. Forecasting, **10**, 681–688
- <sup>32</sup>Jolliffe, I.T., and D.B. Stephenson, 2003: *Forecast Verification. A Practitioner's Guide in Atmospheric Science*. Wiley and Sons Ltd, 240 pp.
- <sup>33</sup>Mason, S.J., 2008: *Understanding forecast verification statistics*. Meteorol. Appl., **15**.
- <sup>34</sup>Woodcock, F., 1976: *The evaluation of yes/no forecasts for scientific and administrative purposes*. Mon. Wea. Rev., **104**, 1209-1214.
- <sup>35</sup>Wilks, D.S., 2005: *Statistical Methods in the Atmospheric Sciences. 2nd Edition*. Elsevier, 627 pp.
- <sup>36</sup>Katz, R.W. and A.H. Murphy (eds), 1997: *Economic Value of Weather and Climate Forecasts*. Cambridge University Press, Cambridge.
- <sup>37</sup>Richardson, D.S., 2000: *Skill and relative economic value of the ECMWF ensemble prediction system*. Quart. J. Royal Met. Soc., **126**, 649-667.
- <sup>38</sup>Wilks, D.S., 2001: *A skill score based on economic value for probability forecasts*. Meteorol. Appl., **8**, 209-219

