Reliability Evaluation of HyBMG by Using ROC Curve

BMKG

Kadarsah

Meteorological Climatological and Geophysical Agency (BMKG)

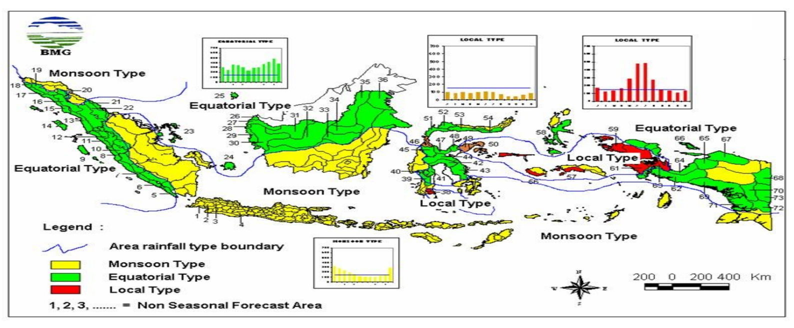
Jl.Angkasa I No.2 Jakarta

Abstract

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 34 cities over 10 years from 1998 to 2007. The result is ROC's curve that describes the reliability HyBMG to predict the rainfall in a particular region. HyBMG has a reliability to predict the rainfall in a particular region.

Keyword:*HyBMG,ROC, False Alarm Rate,Hit Rate*

Introduction



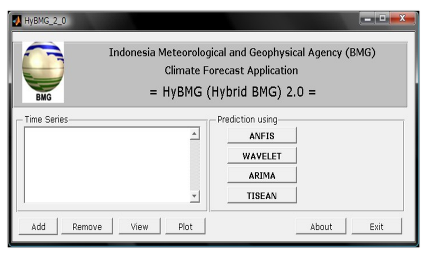
The three main climate regions (monsoon,equatorial and local type),220 seasonal forecast area (without number) and 73 non seasonal forecast area

The rainfall climate of this region is potentially predictable on monthly and seasonal scales but only for limited and specific periods and regions. The study shows a using ROC Curve to verification of HyBMG model. For this purpose, we used rainfall data only 34 cities over 10 years from 1998 to 2007.

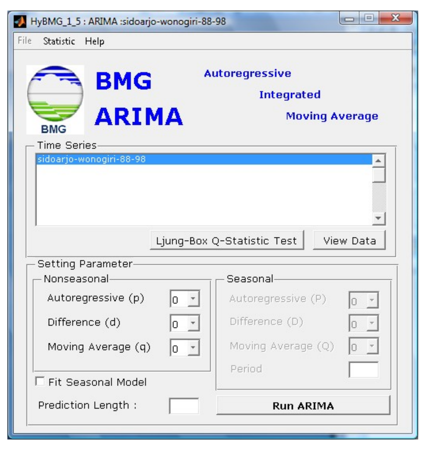
Model Description

HyBMG (Hybrid BMG) as Ensemble Prediction System (Statistical model)

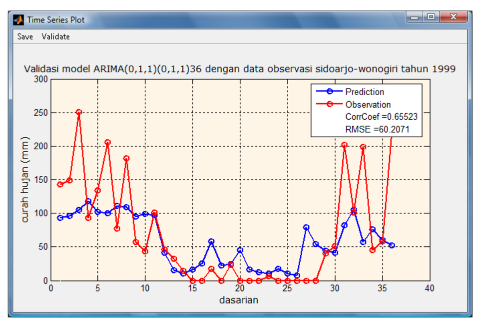
HyBMG



Arima (Autoregressive Integrated Moving Average)



Results from Arima



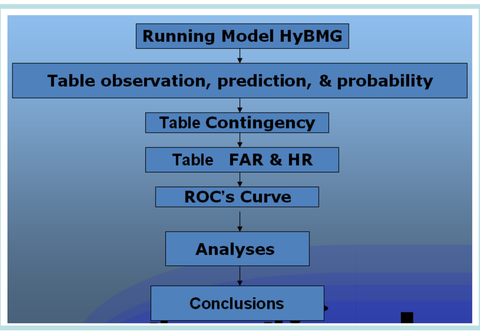
Data

•Data: Average rainfall (1980-2007 34 cities) for standar data
•Data input: Average rainfall (1998-2007 34 cities)

Research Area

No	Cities	No	Cities
1	BANDA ACEH	18	MAKASSAR
2	BANDUNG GEOF	19	MANADO
3	BANYUWANGI	20	MAUMERE
4	BAWEAN	21	PACITAN
5	BIMA	22	PALEMBANG
6	BMG JAKARTA	23	PANGKAL PINANG
7	CILACAP	24	POLONIA
8	DENPASAR	25	RENGAT
9	JATIWANGI	26	RUTENG
10	JUANDA	27	S E R A N G
11	KALIANGET	28	SABANG
12	KLIMAT SEMARANG	29	SAMARINDA
13	KOTABARU	30	SORONG
14	LAMPUNG	31	SUMBAWA
15	LARANTUKA	32	TANJUNG PRIOK
16	MADIUN	33	TEGAL
17	MAJENE	34	WAINGAPU

Methodology



Contingency table

Two-by-two contingency table for verification of a binary forecast system.

	Forecasts		
Observations	Warning, W	No warning, W′	Total
Event, E	*h*	*m*	*e*
Nonevent, E′	*f*	*c*	*e*′
Total	*w*	*w*′	*n*

Contingency tables are highly flexible methods that can be used to estimate the quality of deterministic and probabilistic forecast systems that express output in continuous, categorical, or binary mode. In their simplest form, contingency tables indicate the quality of a forecast system by considering its ability to anticipate correctly the occurrence or nonoccurrence of predefined events that are expressed in binary terms.

ROC Curve

The relative operating characteristic, ROC (Mason 1982), is being considered by the World Meteorological Organization as a recommended method of indicating the skill of probabilistic weather and climate forecasts. The ROC is a highly flexible system that can be used to assess the skill level of dichotomous, categorical, continuous, and probabilistic forecasts. It is based on a 2 x 2 contingency table and compares the proportion of events that were forewarned (the hit rate) with the proportion of none vents that occurred after a warning (the false-alarm rate). Given an ensemble of forecasts, it is useful to construct an ROC curve showing different combinations of hit and false-alarm rates given different forecast probabilities. The ROC curve is useful for identifying an optimal strategy for issuing warnings, by indicating the trade-off between false alarms and misses.

Results

Table observation, prediction, & probability

Area-averaged observations and ensemble-mean predictions of January-December rainfall over Banda Aceh for the period 1998–2007. The observations and ensemble-mean simulations are expressed in tercile format, with "B" representing below-normal, "N" near-normal, and "A" above-normal rainfall. Also presented are the percentages of the individual ensemble members that simulated rainfall in each of the three categories.

NO	Observation	Prediction	Probability B	Probability N	Probability 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

Table Contingency

Probability (%)	Observation	Prediction A	Prediction TA	Total
100%	A	0	8	8
	TA	0	28	28
90	A	0	8	8
	TA	0	28	28
80	A	2	6	8
	TA	0	28	28
70	A	2	6	8
	TA	1	27	28
60	A	6	2	8
	TA	1	27	28
50	A	6	2	8
	TA	1	27	28
40	A	7	1	8
	TA	2	26	28
30	A	7	1	8
	TA	2	26	28
20	A	7	1	8
	TA	2	26	28
10	A	7	1	8
	TA	2	26	28
0	A	8	0	8
	TA	28	0	28

Probability (%)	Observation	Prediction B	Prediction TB	Total
100%	B	0	14	14
	TB	0	22	22
90	B	2	12	14
	TB	4	18	22
80	B	2	12	14
	TB	4	18	22
70	B	4	10	14
	TB	6	16	22
60	B	6	8	14
	TB	7	15	22
50	B	7	7	14
	TB	8	14	22
40	B	8	6	14
	TB	8	14	22
30	B	8	6	14
	TB	8	14	22
20	B	8	6	14
	TB	8	14	22
10	B	8	6	14
	TB	8	14	22
0	B	14	0	14
	TB	22	0	22

Contingency tables for the ensemble mean prediction of Januari–Desember rainfall over Banda Aceh Station for the period 1998–2007. Tables are provided for the simulation of (B) below-normal,(TB) not below-normal and (A) above-normal,(TA) not above-normal rainfall.

Table FAR & HR

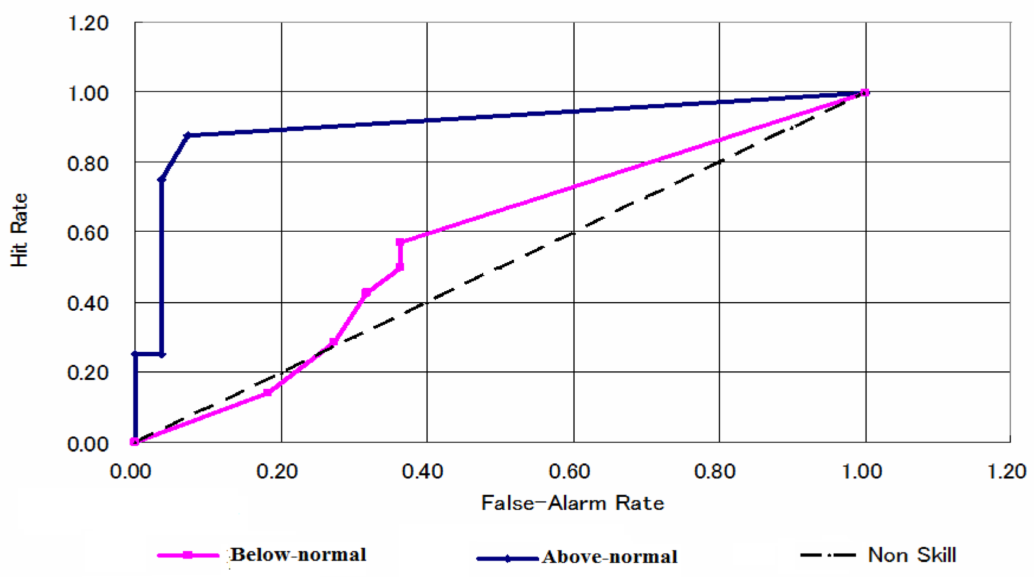
Summary of *False Alarm (FAR) and Hit Rate (HR) for Banda Aceh*

Value (%)	Above-limit FAR	Above-limit HR	Below-limit FAR	Below-limit HR
100	0.00	0.00	0.00	0.00
90	0.00	0.00	0.18	0.14
80	0.00	0.25	0.18	0.14
70	0.04	0.25	0.27	0.29
60	0.04	0.75	0.32	0.43
50	0.04	0.75	0.36	0.50
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	1.00	1.00	1.00	1.00

ROC's Curve

(Relative Operating Characteristics)

Banda Aceh



Interpretation of ROC

The ROC curve lie above the 45 degree line from the origin: the forecast system is skillful and the total area under the curve will be greater than 0.5.

HyBMG model has skill to predict Indonesia rainfall for above-normal (skill score=0.88 and below-normal (skill score=0.52). HyBMG model prediction for above normal better than below normal.

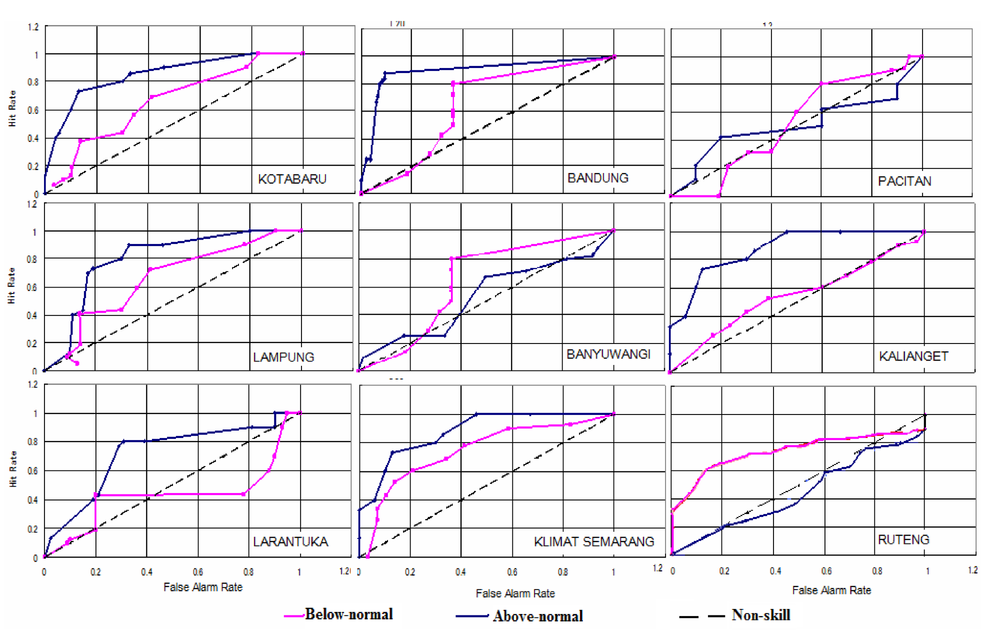
A simple transformation of the ROC score can be suggested so its range is from 1.0 (for a perfect forecast system) to -1.0 (for a perfectly bad forecast system), with 0.0 indicating no skill:

$$S = 2 \times (A - 0.5)$$

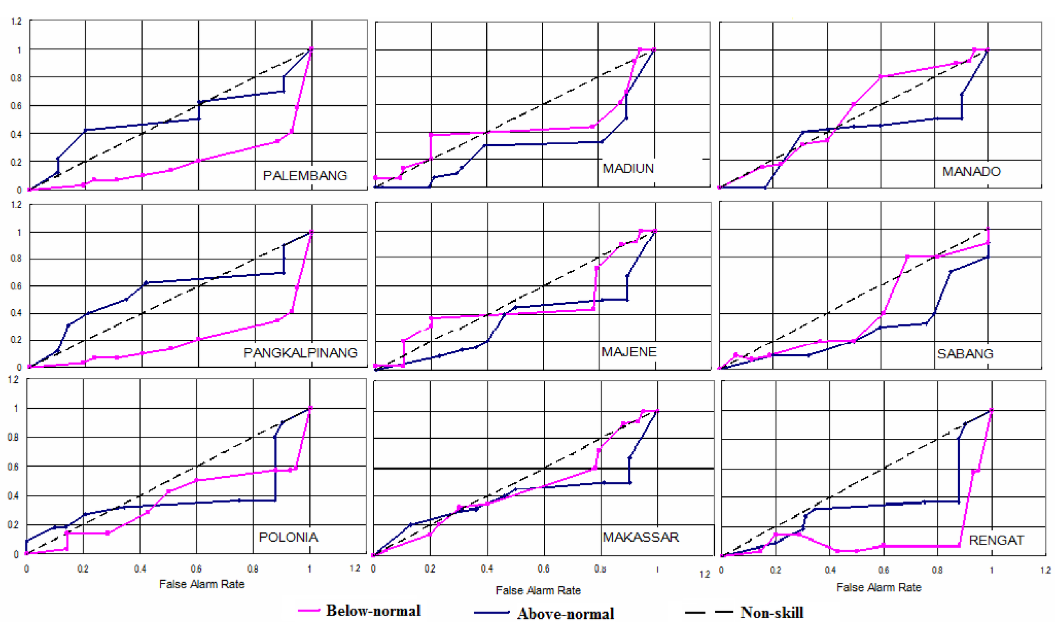
S=skill score
A=the area under the curve

Hit rates vs false-alarm rates (skill positif,negatif and non skill) for 9 cities in Indonesia from 1998 to 2007. The hit and false-alarm rates were calculated using rainfall prediction by the HyBMG model forced with observed rainfall and using 10 ensemble members. Results are shown for the prediction of rainfall in the above-normal (blue-line) and below-normal (red-line) terciles.

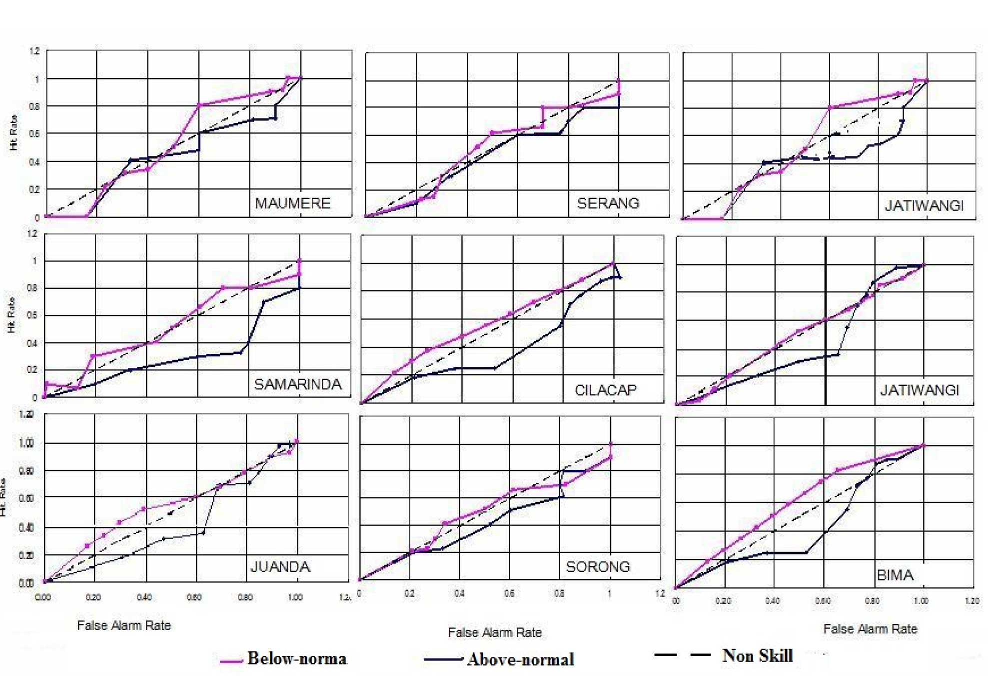
Skil Positif



Skill Negatif



Non- Skill



For the efficacy reason, we display the ROC curve result based on skill score. That means we show the skill positif result of ROC Curve (left) , skill negatif (right) and non-skill (left-below).

The ROL curve can be used by the forecaster to help compare levels of predictability of events of differing magnitude,each region and by the forecast user to estimate probabilities of events other than that defined by the forecaster.

HyBMG has differences of reliability for each climate region in Indonesia. Although this method is good in describing the differences, a better presentation is to show also the ability of HyBMG model to predict rainfall on each region.

The ROC curve for the above-normal rainfall indicates that a useful number of wet events potentially could be forewarned successfully, with a minimal threat of a false alarm, if warnings are issued only when there is high confidence. If the cost of a miss, rather than of a false alarm, is prohibitively high, then it would be desirable to increase the number of warnings by relaxing the warning criterion. Issuing more warnings should hopefully ensure that the number of hits is increased at the expense of the number of misses, but with the penalty of issuing more false alarms. The ROC curve is useful in helping to identify an optimum warning criterion, by indicating the trade-off between misses and false alarms. (Mason, And Graham, 1999)

For a probabilistic system, the ROC curve illustrates the varying quality of the forecast system at different levels of confidence in the warning (the forecast probability). It is not necessarily the case that a forecast system demonstrates greatest value at the point at which the likelihood ratio is maximized: instead, each user has a specific cost–loss operating structure, and hence the relative frequencies of hits, false alarms, and misses have to be optimized. The ROC curve can be used in helping to identify this optimum strategy in any specific application (Harvey et al. 1992).

The ROC curve is useful in helping to identify an optimum warning criterion, by indicating the trade-off between misses and false alarms.

Conclusions

The relative operating characteristic (ROC) is being considered by the World Meteorological Organization as a recommended method of indicating the skill of probabilistic weather and climate forecasts. The ROC is a highly flexible system that can be used to assess the skill level of dichotomous, categorical, continuous, and probabilistic forecasts.

In an operational environment, particularly in BMKG(Indonesia Met Office), the warning is provided in advance of the outcome, and so there is additional value in knowing the probability of an event occurring, contingent upon the forecast probability by using HyBMG model.

ROC curve can be generated by plotting the hit and false-alarm rate for the forecast system, together with the hit and false-alarm rates obtained for perpetual warnings (for which the hit and false-alarm rates equal 1.0) and no-warnings (for which the hit and false-alarm rates equal 0.0)

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 cost–loss table.

HyBMG has a reliability to predict the rainfall in a particular region.

【 References 】

- 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.
- Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30,** 291–303.
- Mason, S.J. And Graham, N.E. 1999. Conditional probabilities, Relative operating characteristics and relative operating levels. Weather and Forecasting 14:713-725
- Mason, S.J., 2004: On using "climatology" as a reference strategy in the Brier and ranked probability skill scores. Mon. Wea. Rev., 1891-1895
- Murphy, A.H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather Forecasting* 8:281-293
- Nurmi, P., 2003: Recommendations on the verification of local weather forecasts (at ECWMF member states). ECMWF Operations Department, October 2003.\
- von Storch, H. and F.W. Zwiers, 1999: Statistical Analysis in Climate Research. Cambridge University Press, Cambridge.
- Wilks, D.S., 1995: Statistical Methods in the Atmospheric Sciences. An Introduction. Academic Press, San Diego, 467 pp.

[Acknowledgements]

This poster presentation has been supported by WMO (World Meteorological Organization)

This poster is presented in the 4th International Verification Methods Workshop 4-10 June 2009