Verification of probability forecast

Kohshiro Dehara

Forecast Unit

Climate Prediction Division

Japan Meteorological Agency

Outline

1. Purposes of verification
2. Verification methods
   - Reliability diagram
   - Brier Score, Brier Skill Score
3. Reliability diagrams of JMA's one-month forecast.

1. Purposes of verification

Verification is essential …

- To monitor the quality of forecasts
  - how accurate are forecasts and are they improving?
- To help us understand the characteristics such as reliability of the guidance

2. Verification methods

- Verification methods for probability forecast

  Reliability diagram

  Brier Skill Score (BSS)

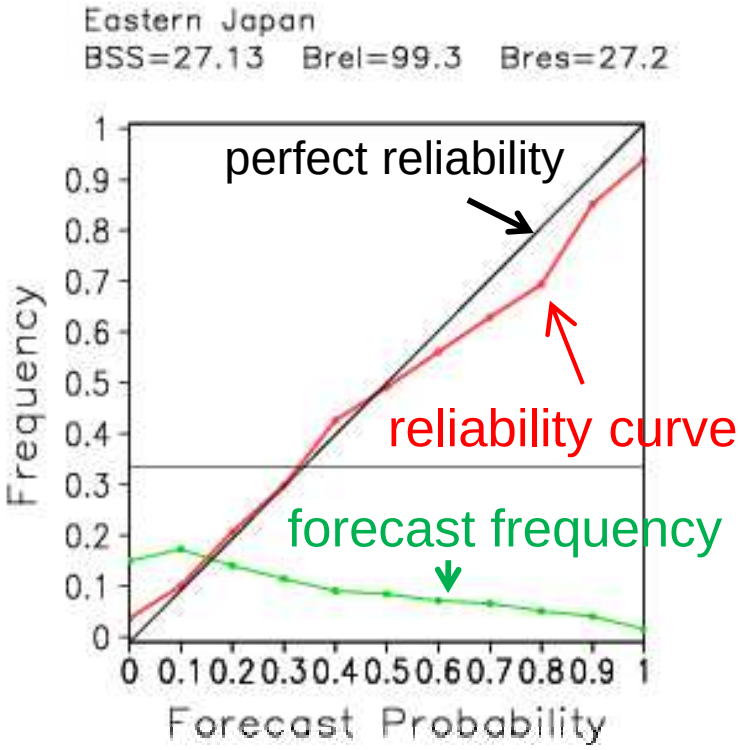
Reliability diagram

A reliability curve shows the occurrence frequency of an event against the forecast probability.

If a forecast has perfect reliability, the probability coincides with the observed frequency, and the curve will be a diagonal.

Points below the diagonal indicate over-forecasting and those above it indicate under-forecasting.

The forecast frequency indicates the tendency of a guidance.



Brier score

- Brier score measures mean squared error of the probability forecasts.

$$BS = \frac{1}{N}\sum_{i=1}^{N}(p_i - v_i)^2$$

p_i : forecast probability

v_i : observed occurrence (0 or 1)

N : total number of forecasts

Range: 0 to 1. Perfect score: 0 Climatology: $\bar{o}(1-\bar{o})$

$\bar{o}$: climatic occurrence

3 category forecast : 0.33

2 category forecast : 0.5

Brier skill score

= An improvement rate with respect to a climatic forecast

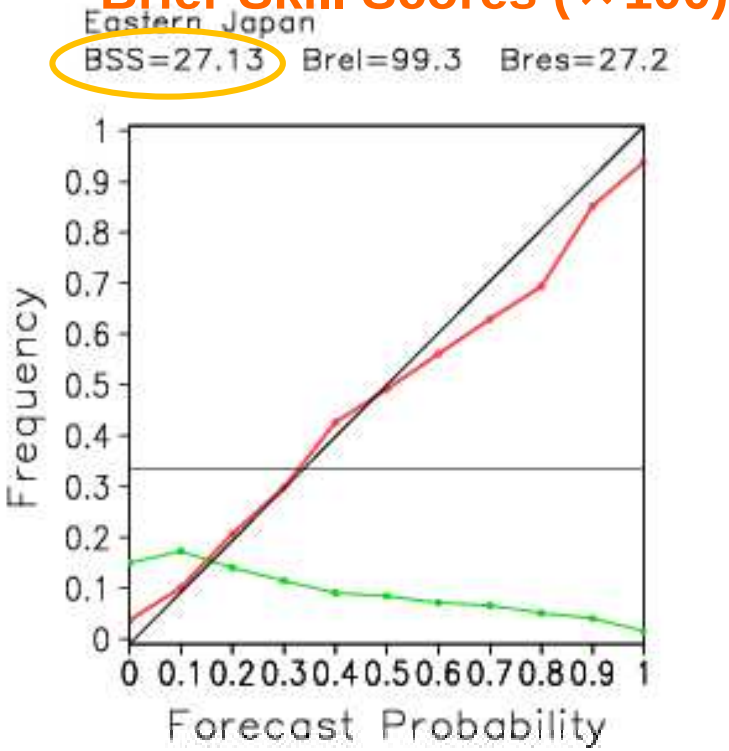
$$Skill\ score = \frac{score_{forecast} - score_{reference}}{score_{perfect\ forecast} - score_{reference}}$$

Brier skill score

$$BSS = \frac{BS - BS_{clim}}{0 - BS_{clim}} = 1 - \frac{BS}{BS_{clim}}$$

Range: minus infinity to 1. Perfect score: 1
BSS<0 : inferior to the climatic forecast.
BSS>0 : superior to the climatic forecast.

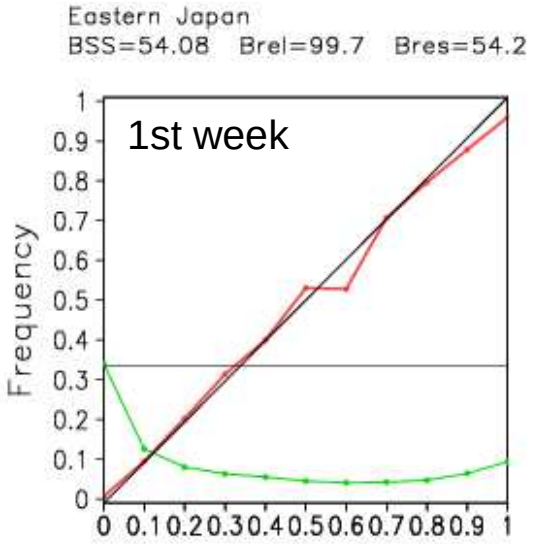
Brier Skill Scores (×100)

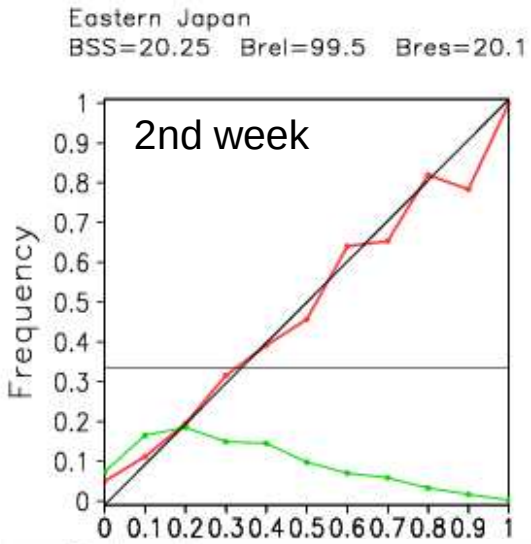


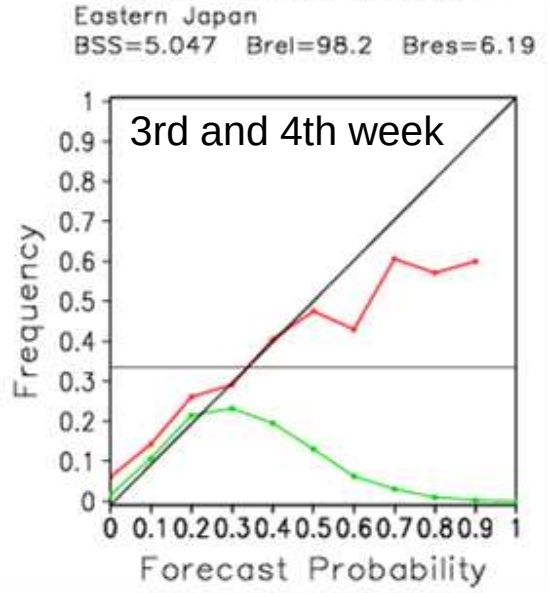
Reliability diagrams of JMA's one-month forecast

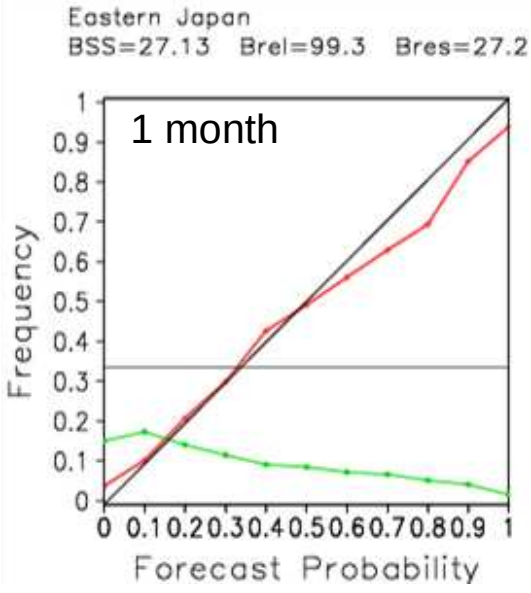
Reliability diagrams for temperature in Eastern Japan

Event: "above" and "below" in three-category



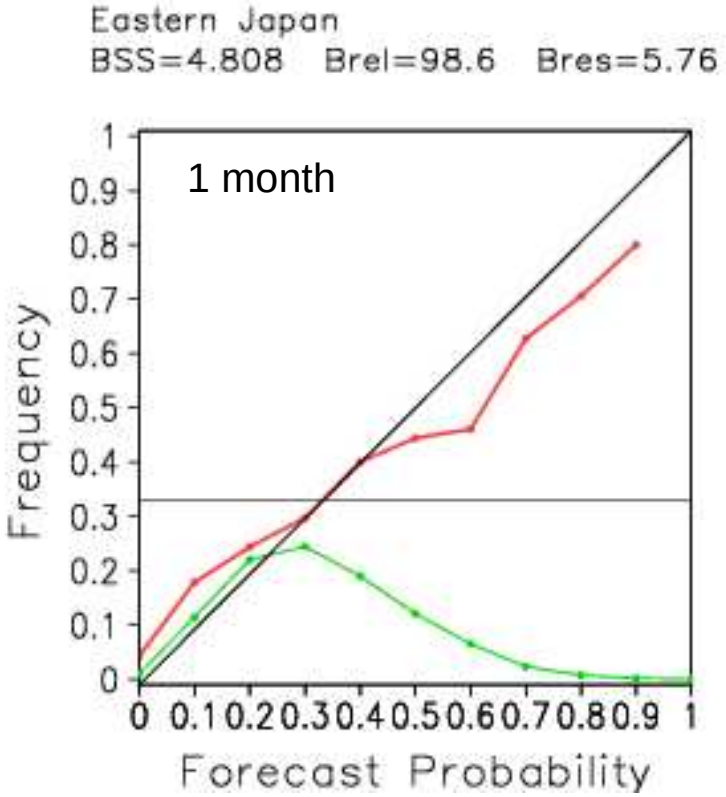






Reliability diagrams of JMA's one-month forecast

Reliability diagrams for precipitation in Eastern Japan
Event: "above" and "below" in three-category



Thank you
for your attention

Index	Climatology	Perfect
Reliability curve		Fit to the diagonal line
Brier Score	$\bar{o}(1-\bar{o})$	0
Brier Skill Score	0	+1

$\bar{o}$: climatic occurrence

Extra part

Relative Operating Characteristic (ROC)

Steps for making ROC diagram

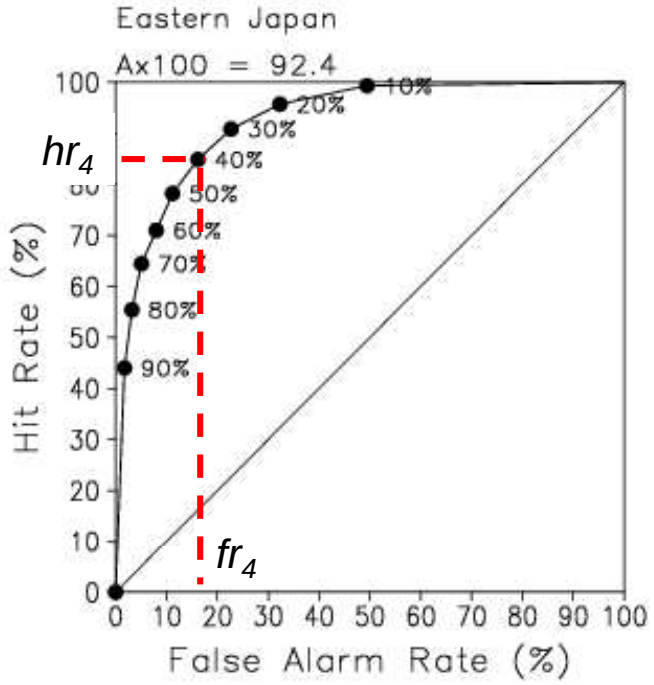
1. Transforming probability forecasts into yes/no forecasts.
2. Computing the hit rates (*hr*) and false alarm rates (*fr*) based on table 1.

$$hr_i = \frac{A_i}{A_i + C_i}, \qquad fr_i = \frac{B_i}{B_i + D_i}$$

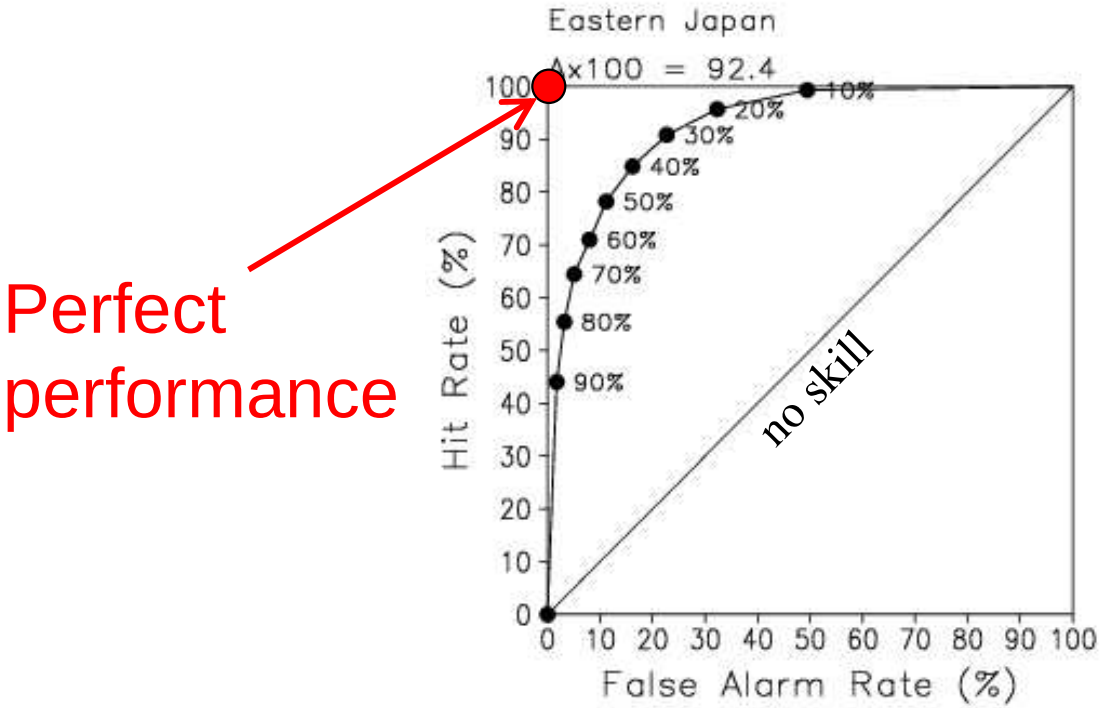
3. Plotting points based on the hit rates and false alarm rates, and draw the ROC curve.

Table 1 : contingency table

		Observation	
		Yes	No
Forecast	Threshold ≦	Ai	Bi
	Threshold >	Ci	Di



Interpretation of ROC curves



- A point in the upper left corner of an ROC diagram represents a perfect forecast (100% hit rate, 0% false alarm rate).
- The diagonal line where the hit rate and the false alarm rate are the same represents a forecast with no skill.
- If AUC>0.5, the forecast has a certain level of skill, and AUC is at its maximum (AUC=1) for a perfect forecast.

References

Murphy, A.H., 1973: A new vector partition of the probability score. J. Appl. Meteor., 12, 595-600.

Murphy, A.H., 1988: Skill scores based on the mean square error and their relationships to the correlation coefficient. Mon. Wea. Rev., 16, 2417-2424.

http://www.bom.gov.au/wmo/lrfvs/index.html

http://www.ecmwf.int/newsevents/meetings/workshops/2007/jwgv/index.html