Introduction of One-month Forecast Guidance

Hiroshi OHNO
Tokyo Climate Center (TCC)/
Climate Prediction Division of
Japan Meteorological Agency (JMA)

Outline

- Outline of Guidance
  - Objective of Guidance
  - MOS Technique
  - Regression Model
  - Estimation of Probability
- Verification
  - Verification Score

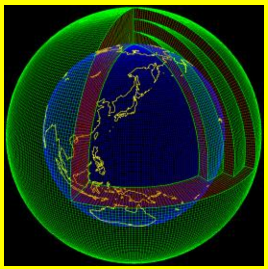
Outline of Guidance

Guidance

INPUT

OUTPUT

Numerical model



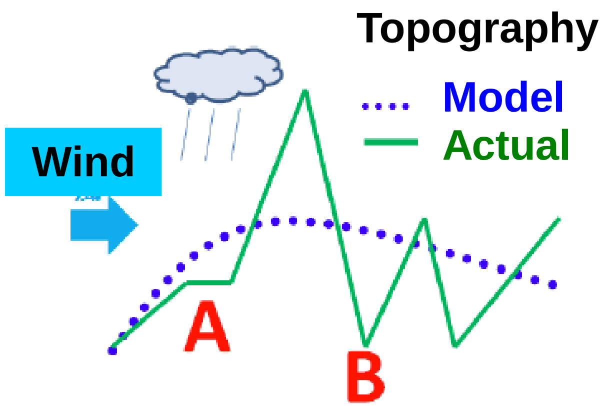
Statistical downscale

Probabilistic forecast

- "Guidance" is an application to translate model output values into target of forecasting.
- Principle of guidance is based on statistical relationship using model forecasts and observation data for past cases.

Role of Guidance

- To extract effect of sub-grid scale topography
  - Model does not necessarily reproduce effect of local topography due to limited resolution.
- To reduce imperfection of the model, such as systematic error (bias error).
- To estimate degree of uncertainty, considering prediction skill



A: Upwind side
  ➢ Model may underestimate precipitation

B: Bottom of the valley
  ➢ Model may have warming bias

"Guidance" enable to improve prediction skill, compared with the direct model output.

Principle of Guidance – MOS Technique

MOS (Model Output Statistics):

To derive statistical relationship between observation and model forecast from past cases, and apply it to the real-time forecast

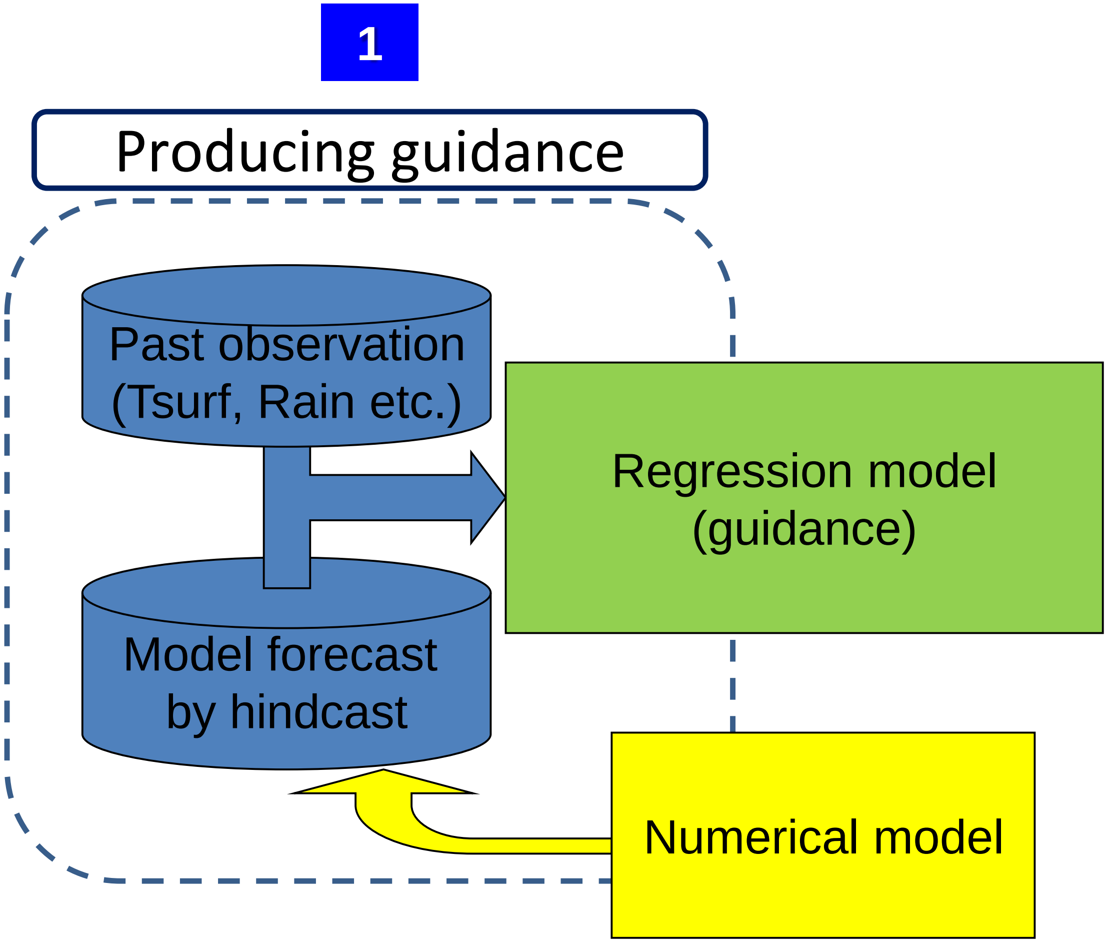
- Two types of the time series data are needed in order to produce guidance.
  - 1. Past observation — Prepared by users

    (Variable to be predcited; Predictand)
  - 2. Past model forecast by hindcast (Predictor)

Available on TCC-HP or The Guidance Tool

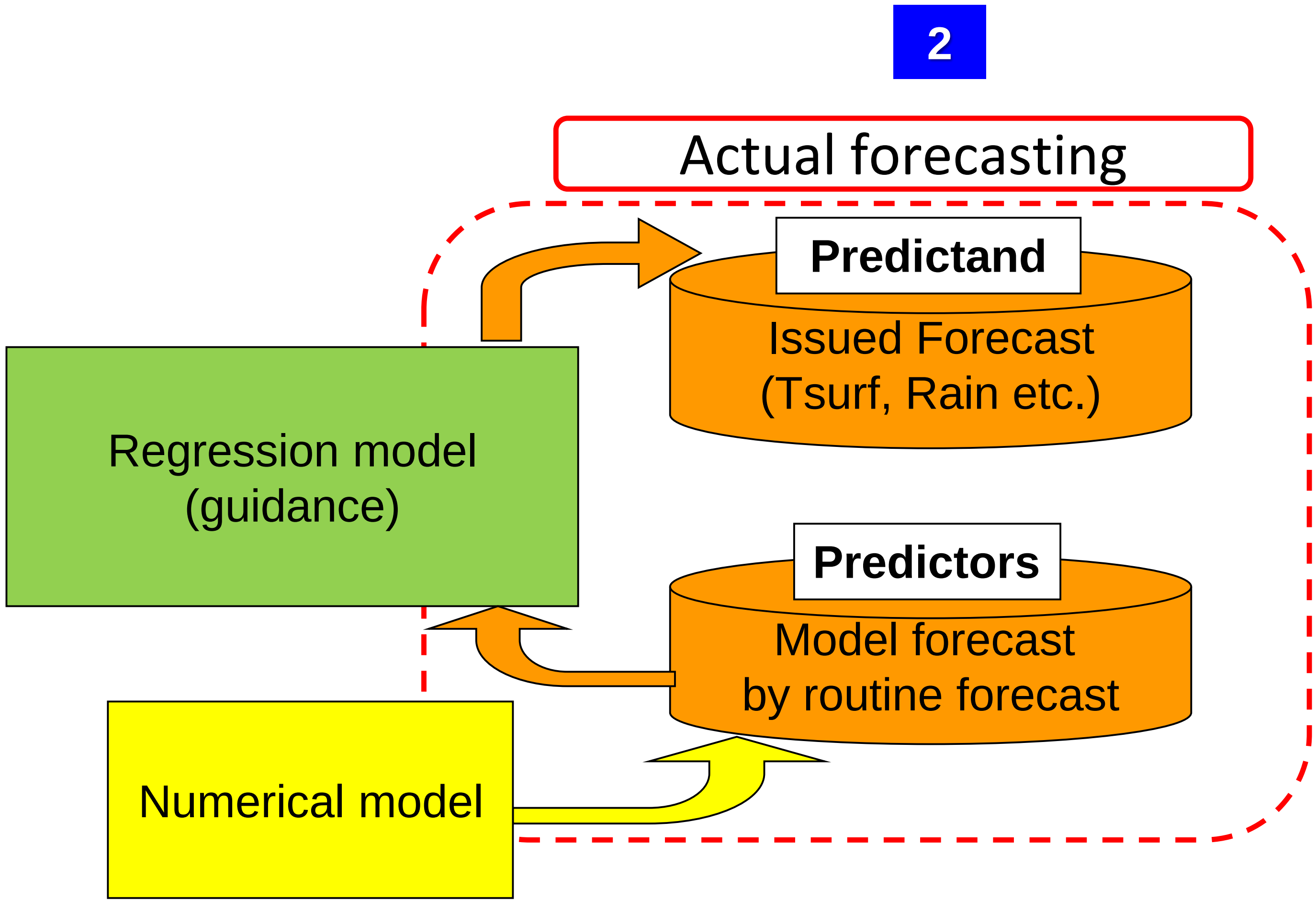
Concept of MOS Technique (1)

- **Statistical relationship** is estimated using observation and model forecast for past cases.



Concept of MOS Technique (2)

- In the real-time forecast, model results are applied to the **statistical relationship** to obtain variable to be precited



Single Regression

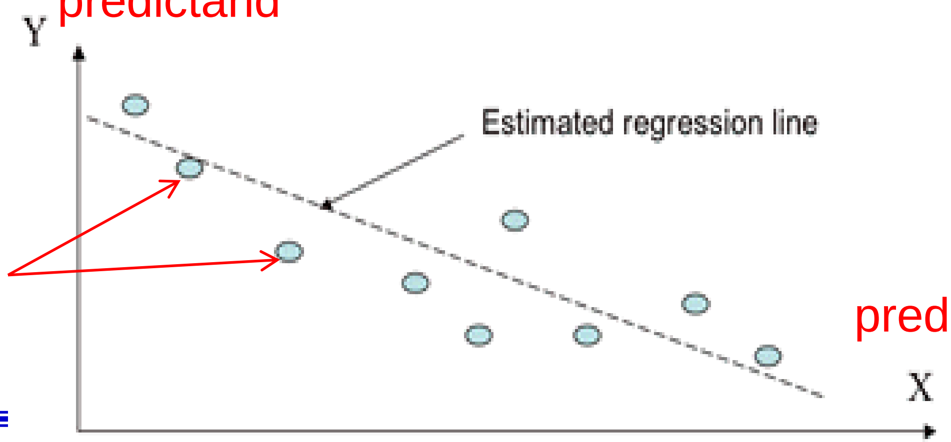
- Single regression is the relationship between **one explanatory variable** (predictors) and **variable to be precidted** (predictand, ex., temp. rainfall).
- Single regression model is written as

$$Y = aX + b + \varepsilon$$

Y: predictand　X: predictor
a: regression coefficient　b: constant,
ε: error term

Predictand (e.g., temp., precipitation)

Predictor (i.e., model output)



Multiple Regression

- More than one predictors are employed in multiple regression.
- It is assumed that the predictand is the sum of a linear combination of predictors.

Example: **two** predictors

a_1, a_2: regression coefficient
b: constant
ε: error term

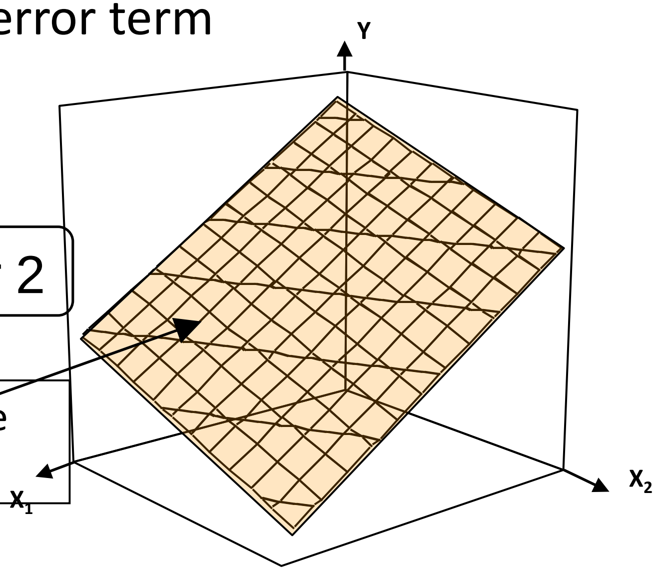
$$Y = a_1X_1 + a_2X_2 + b + \varepsilon$$

Predictand (e.g., temp., precipitation)

Predictor 1 (i.e., model output)

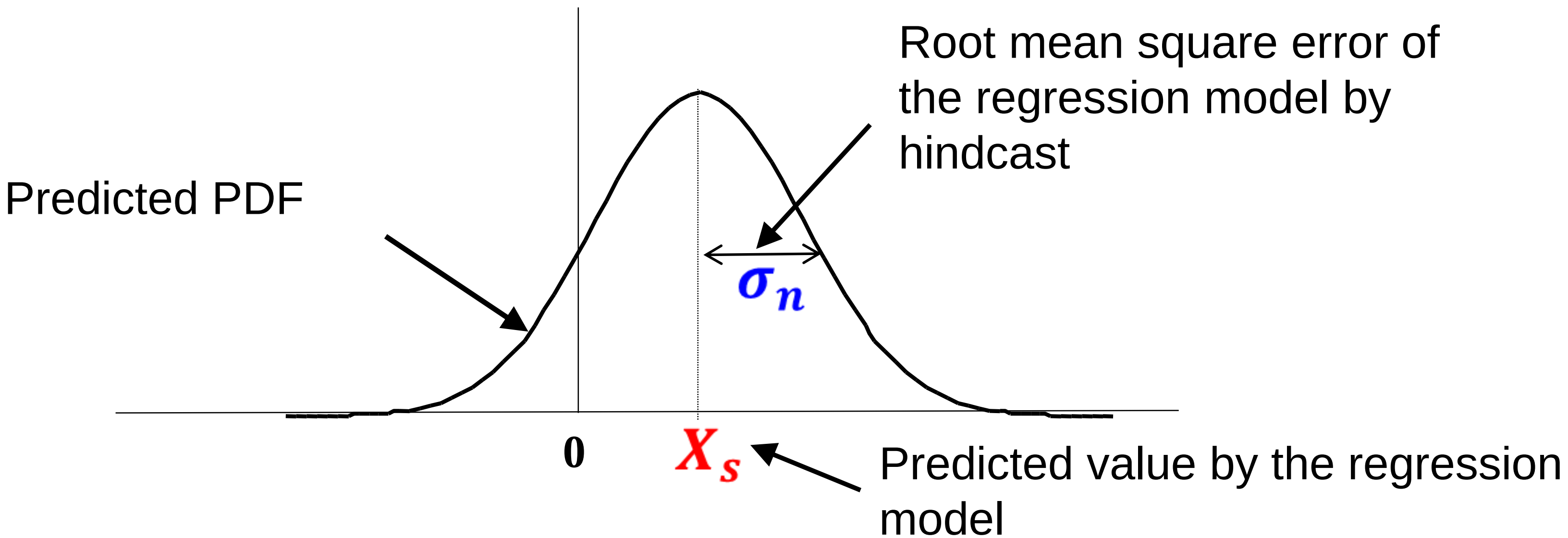
Predictor 2

Predictand will be near this plane.



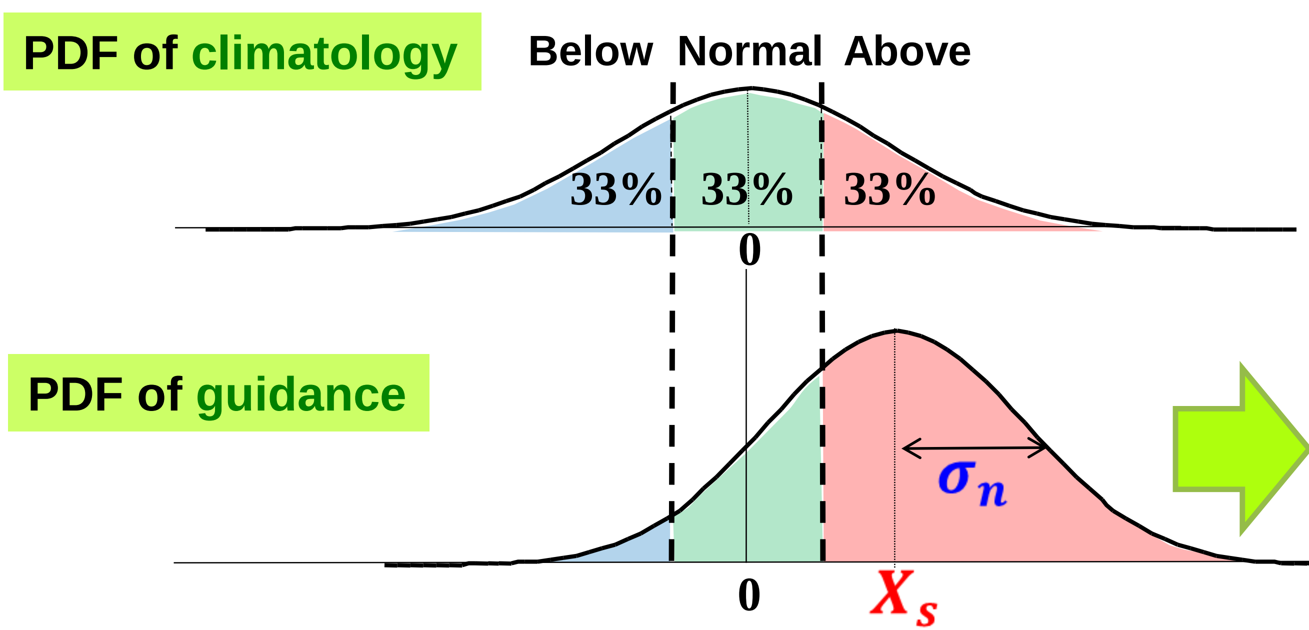
From Regression Model to Probability

- Probability Density Function (PDF) is assumed to be a **normal distribution.**
  - **Mean (X_s):** Predicted value by the regression model
  - **Standard deviation (σ_n):** RMSE of the regression model.



Estimation of Probability for 3-category

- The threshold values for 3 categories are determined from the past observation (1981 to 2010).
- Probability for each category (below-, near-, above-normal) is calculated by PDF and the threshold values.



Issued Forecast		
Below	Normal	Above
10%	30%	60%

Normalization of Precipitation Data

- **Normal distribution** is assumed in the regression model.
- As for temperature, its distribution is generally approximated by a normal distribution.

Meanwhile,

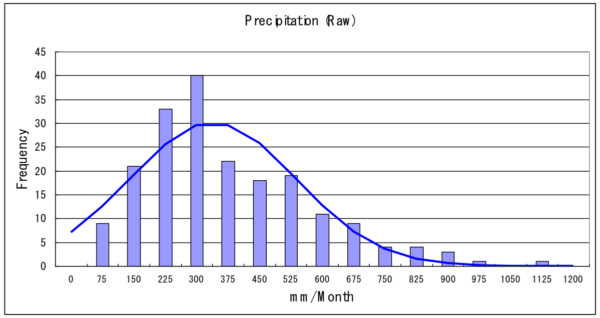
- As for precipitation, its distribution does not represent a normal distribution, and it's usually approximated by a gamma distribution.

- In order to create guidance, precipitation data need to be normalized.

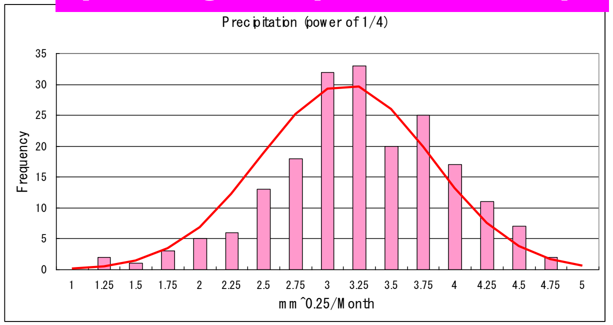
- Power of 1/4 for precipitation ($RAIN^{1/4}$) is approximated by a normal distribution.

Ex. Precipitation over Japan

(Raw value)



(Taking the power of 1/4)



Verification

Verification for Deterministic Forecast

- Root Mean Square Error (RMSE)

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)^2}$$

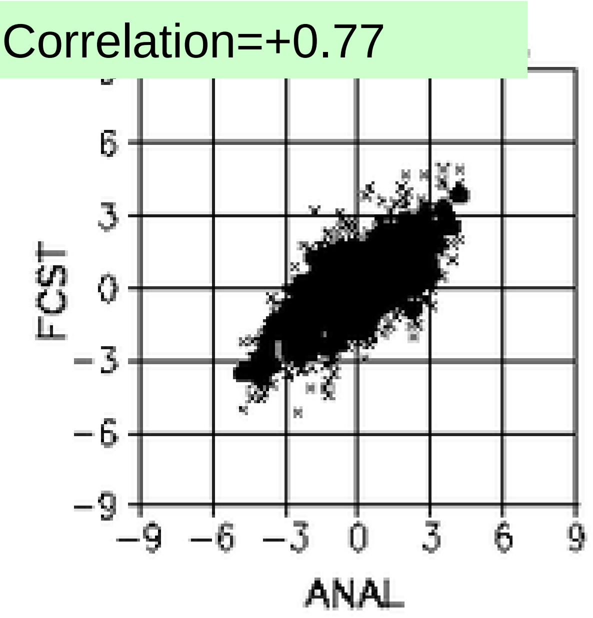
F_i: Forecast
O_i: Observation
C_i: Climatology
N: Sample size

Perfect score: 0

- Anomaly Correlation Coefficient (ACC)

$$ACC = \frac{\sum_{i=1}^{N}(F_i - C_i)(O_i - C_i)}{\sqrt{\sum_{i=1}^{N}(F_i - C_i)^2}\sqrt{\sum_{i=1}^{N}(O_i - C_i)^2}}$$

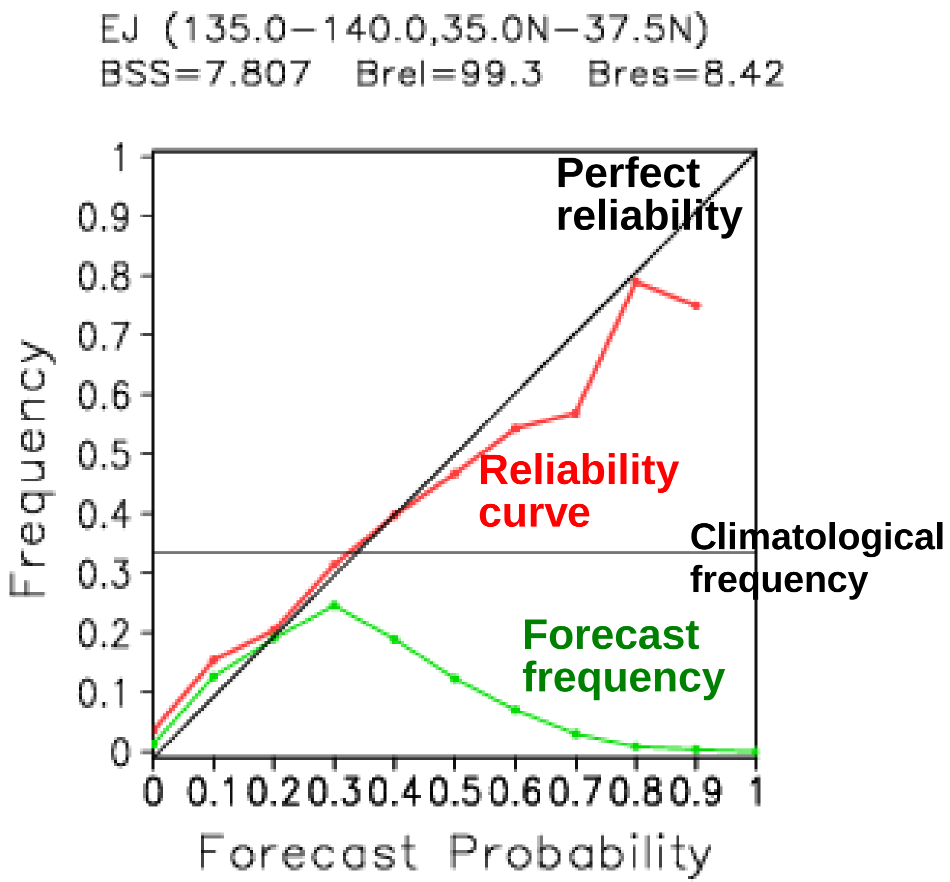
Range: -1 to 1.
Perfect score: 1



Reliability Diagram

- Red line (reliability curve); plotted the observed frequency(Y-axis) against the forecast probability(X-axis)

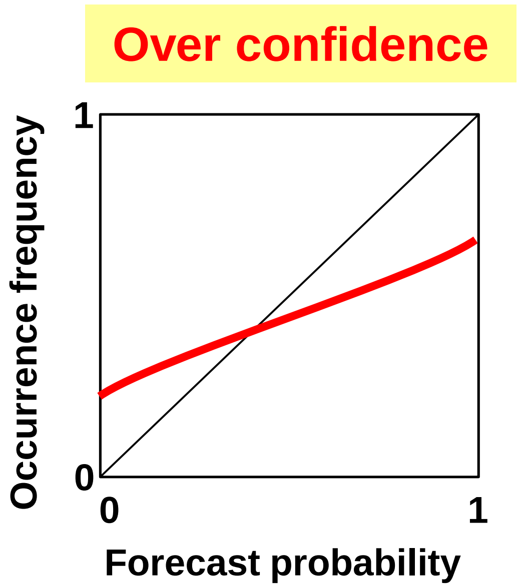
Probabilistic forecast becomes better the more the reliability curve fit to 45° line (perfect reliability).

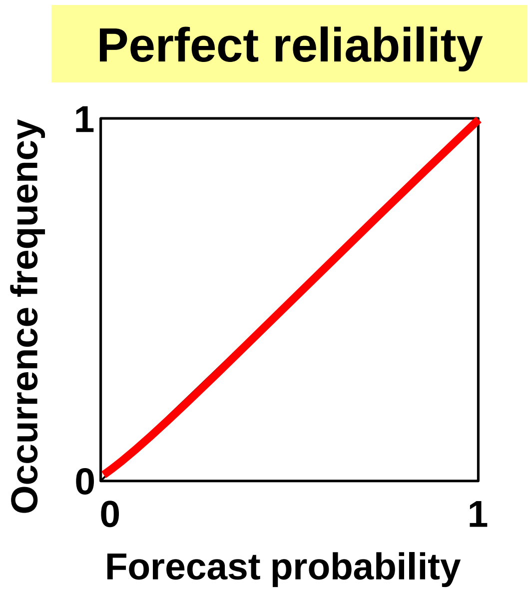


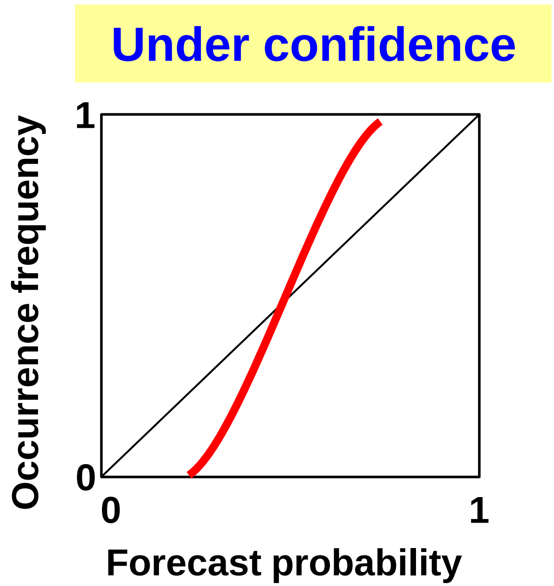
- Green line denotes forecast frequency (sharpness diagram);
  - If most of the forecast probabilities are near the climatological frequency = unsharp
  - If probabilities near 0 and 1 (100%) are often used = sharp

Probabilistic forecast

Over/under Confidence





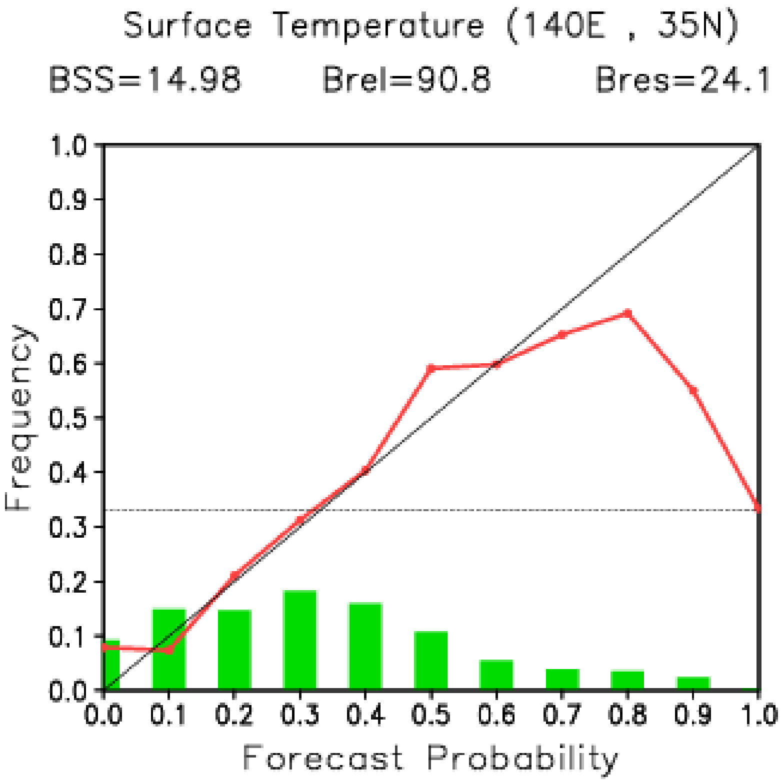


- ✓ Predicted probabilities are overestimated as compared with actual
- ✓ Predicted probabilities are underestimated as compared with actual

Probabilistic forecast

Interpretation of Reliability Diagram

Example



- The forecast is generally reliable for below 60%, while over-confident over 70%.

✓ Maximum probability should be suppressed under 60%

Brier Score (BS)

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

$$BS = \frac{1}{2N}\sum_{i=1}^{N}\sum_{m=1}^{3}(p_i^m - o_i^m)^2$$

p_i^m : forecast probability

o_i^m : observed occurrence (0 or 1)

N : forecast frequency

m : category

Range: 0 to 1

Smaller score indicates better forecast (Perfect score: 0)

Forecast (Below, Near, Above): (0.1, 0.3, 0.6)

Observation: Above normal (0, 0, 1)

BS: $\{(0.1-0)^2+(0.3-0)^2+(0.6-1)^2\}/2 = 0.13$

Brier Skill Score (BSS)

- Brier skill score is ***skill*** relative to a reference forecast (usually climatology).

$$BSS = 1 - \frac{BS}{BS_{referenc}}$$

$$BSr = \frac{1}{3}$$

- Perfect score: 1
- BSS>0 : better than the climatological forecast.
- BSS<0 : worse than the climatological forecast.

Tomorrow's Exercise

- We will make 1-month forecast guidance (multi-regression model) using the internet-based tool.
- Predictors are JMA's 1-month forecast model's output such as temperature, rainfall and atmospheric circulation variables (ex. geopotential height, wind) around your forecast point.
- You have to find a better combination of predictors with higher prediction skill.