2

Concept of numerical guidance

Masayuki Hirai

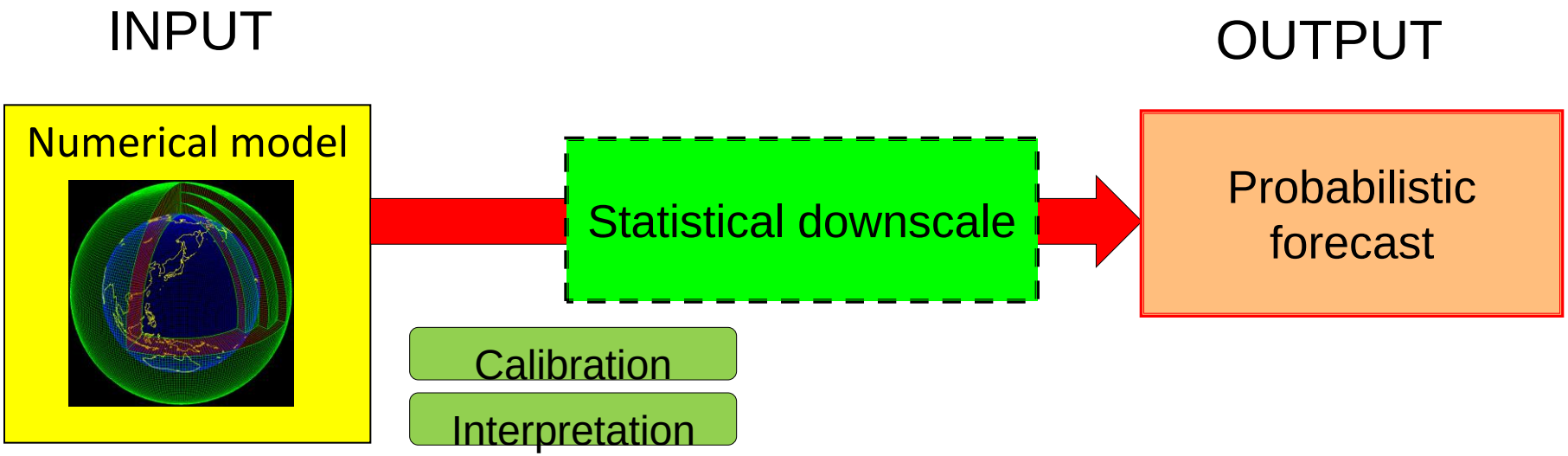
Tokyo Climate Center (TCC)/
Climate Prediction Division of
Japan Meteorological Agency (JMA)

Outline

- Outline of guidance
  - Role
  - Principle
    - Regression model
    - Estimation of probability
- Verification
  - Verification score

Outline of guidance

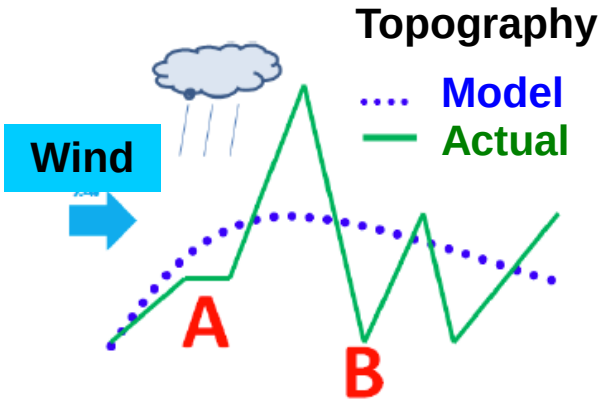
Guidance



- "Guidance" is an application to translate model output values into target of forecasting.
- Principle of guidance is to predict future values 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 may not reproduce effect of local topography due to limited resolution, while enable to reproduce large-scale field.
- 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);

Calculation of statistical relationship between observation and model forecast for past cases, and apply to the real-time forecast

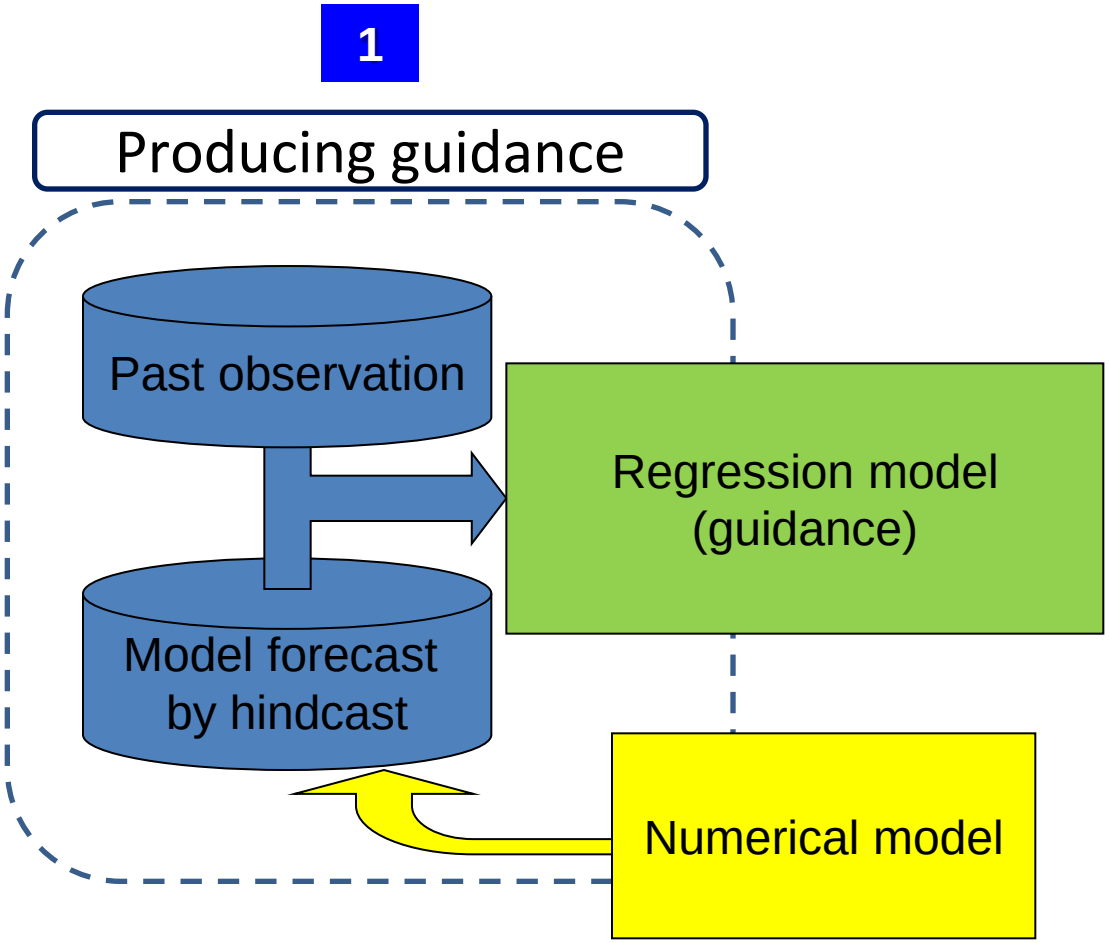
- Two types of the time series data are needed in order to produce guidance.
  - 1. Past observation (i.e., Predictands)
  - 2. Past model forecast by hindcast (i.e., Predictors)

Create by users

On ITACS

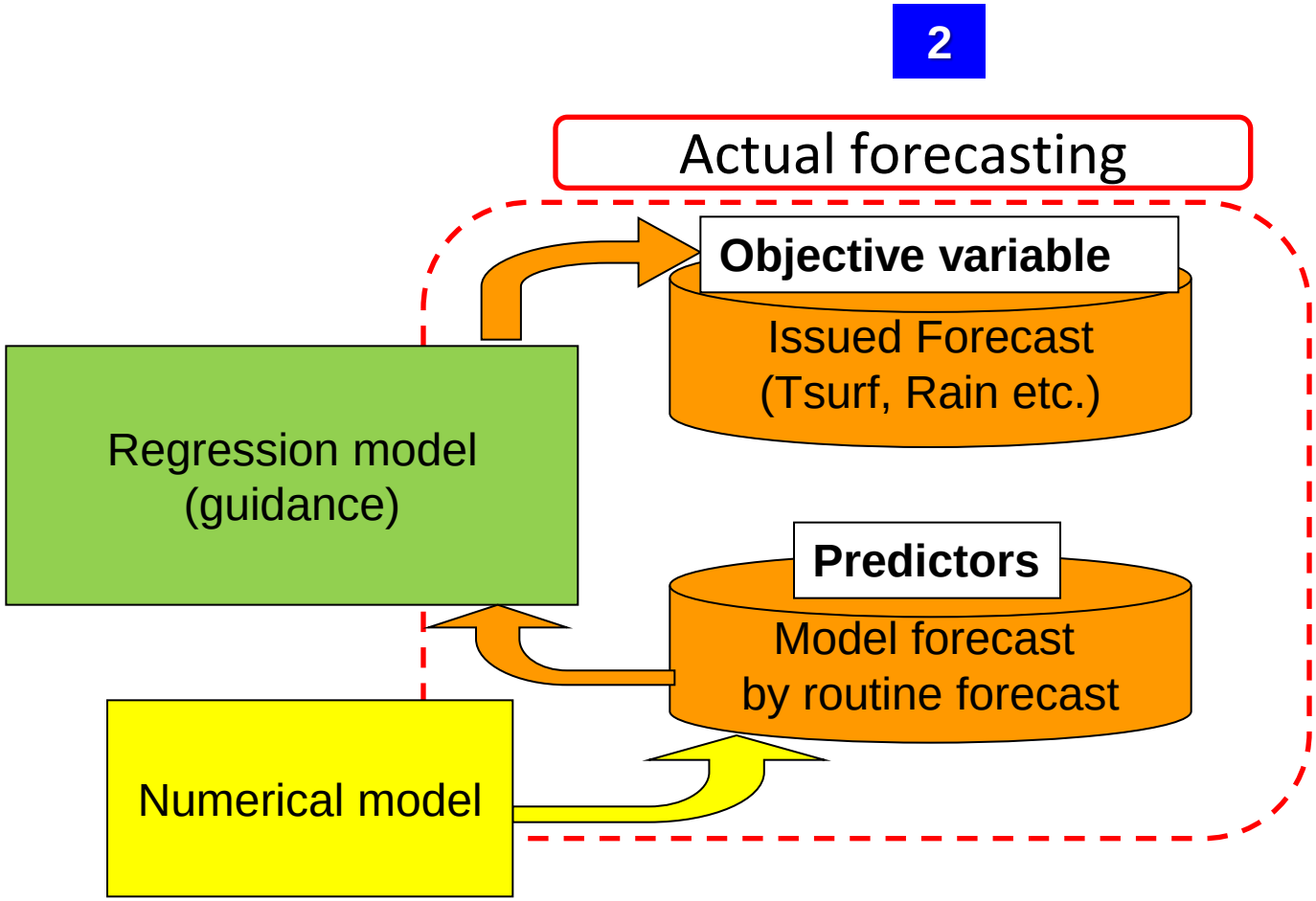
Concept of the guidance adopting MOS technique (1)

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



Concept of the guidance adopting MOS technique (2)

- In the real-time forecast, predicted value is calculated applying to the statistical relationship.



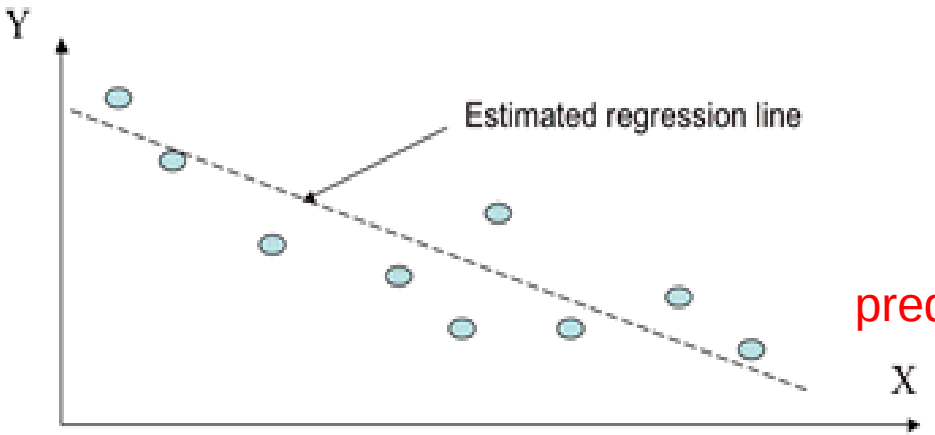
Single regression

- Single regression is modeled the relationship between **one explanatory variable** (predictors) and **objective variable** (ex. temp. rainfall).
- Single regression model is written as

$$Y = a_x + b + \varepsilon$$

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

objective variable



predictors

Multiple regression

- Multiple regression is assumed that the objective variable is the sum of a linear combination of plural predictors.
- Multiple regression model is written as

$$Y = \sum_{k} (a_k X_k) + b + \varepsilon$$

k=1,2,…,n

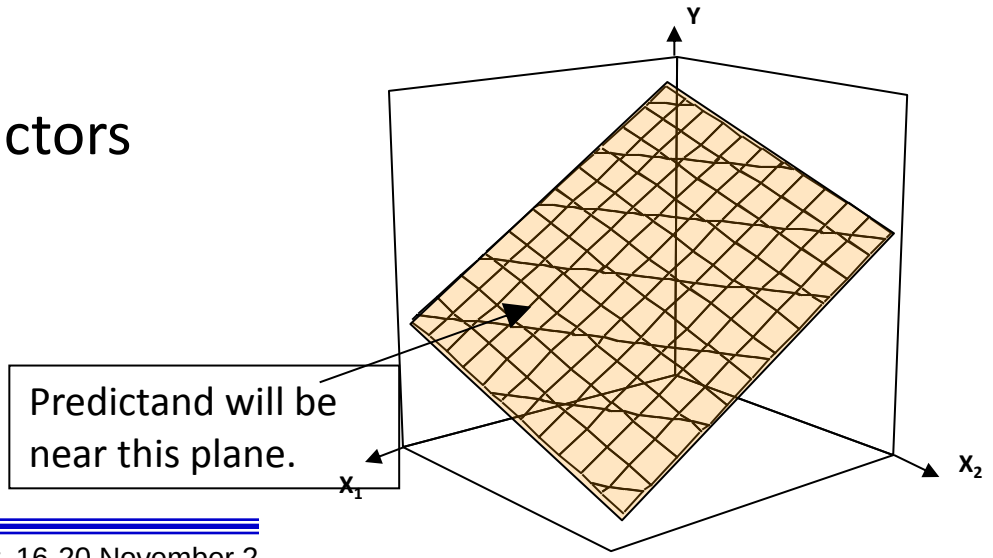
Y: predictand　X: predictors

a: regression coefficient

b: constant

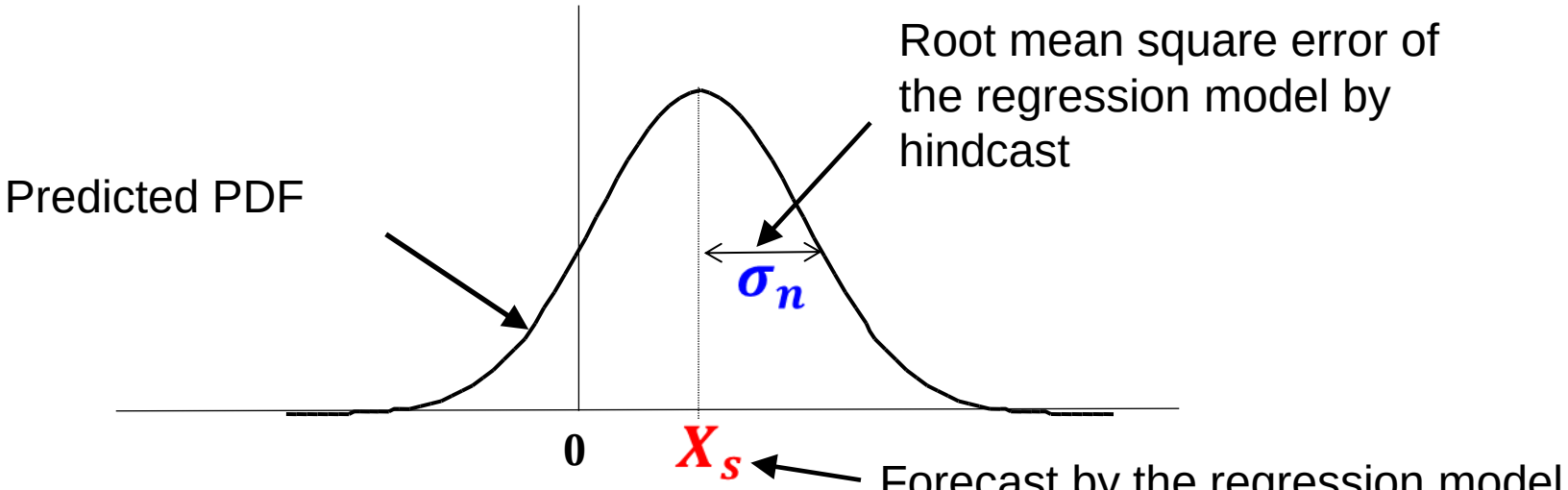
ε: error term

Example: two predictors



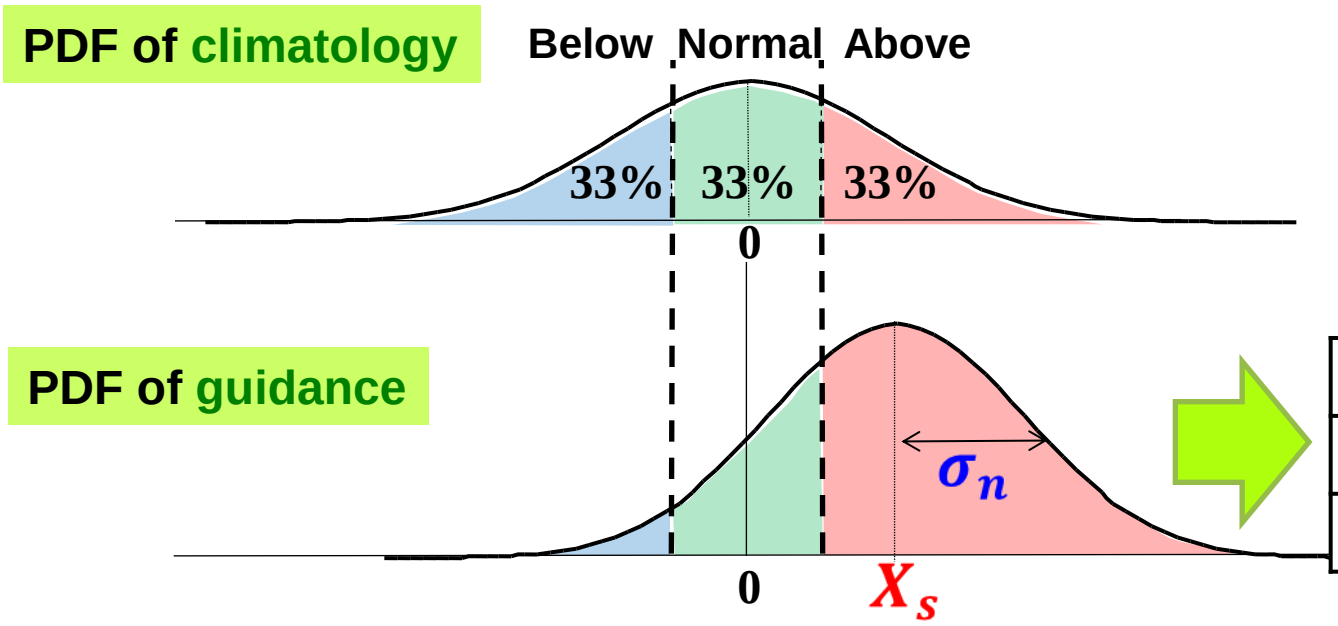
Translation to PDF in the regression model

- In the guidance tool, Probability Density Function (PDF) is assumed to be a **normal distribution.**
  - **Mean (x_s)** : prediction value by the regression model
  - **Standard deviation (σ_n)** is error of the regression model, which is assumed to be RMSE of the regression model using hindcast.



Estimation of Tercile probability with regression model

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



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,

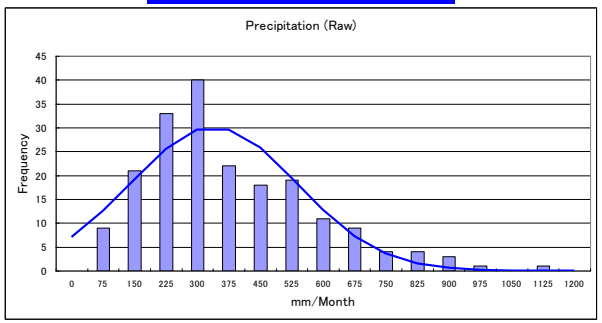
- As for precipitation, its distribution does not represent a normal distribution, and it's usually approximated by a gamma distribution.

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

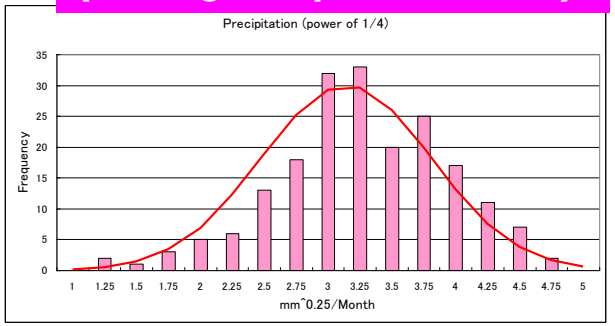
- Power of 1/4 for precipitation ($RAIN^{1/4}$) is approximated by a normal distribution.

Ex. Precipitation over Japan

(Row value)



(Taking the power of 1/4)



Verification

Verification for deterministic forecast (1 of 2)

- Mean Square Error (MSE)
- Root Mean Square Error (RMSE)

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(F_i - O_i\right)^2$$

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

F : forecast

O : observation

N : sample size

- MSSS

Skill score of MSE, comparing with climatology

Verification for deterministic forecast (2 of 2)

- Anomaly Correlation Coefficient (ACC)

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

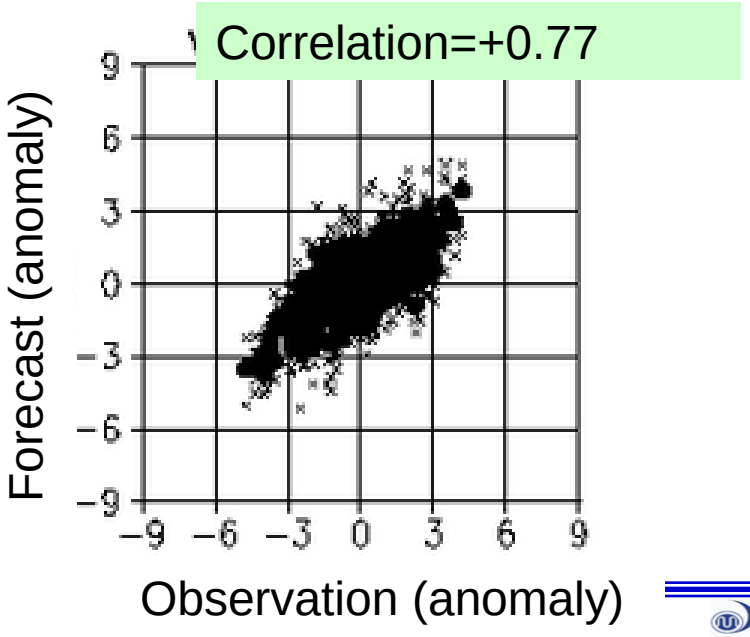
F :**forecast**

O :**observation**

C :**climatology**

Range: -1 to 1.
Perfect score: 1.

Forecast data is calibrated

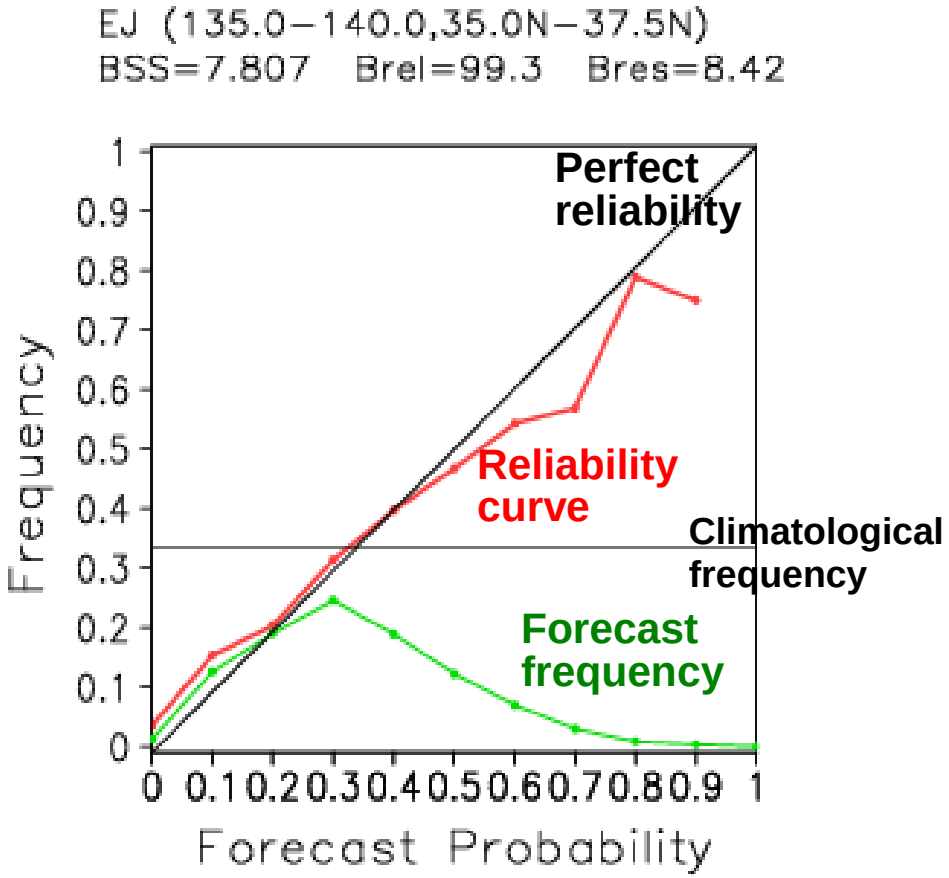


Verification for probabilistic forecast (1) Reliability diagram

- 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).

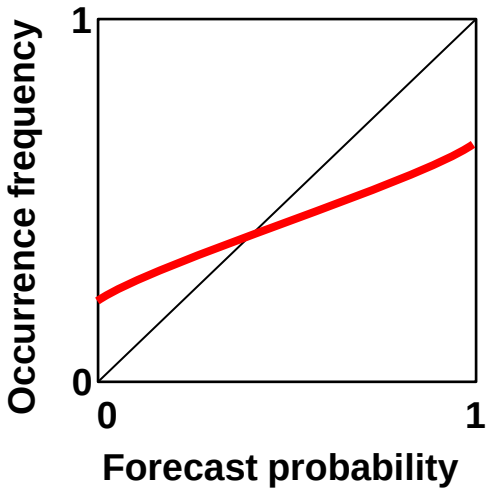
- Green line denotes forecast frequency (sharpness diagram);
  - •If most of the forecast probabilities are near the climatological frequency = unsharp
  - •If probabilities near 0 and 1 (100%) are often used = sharp



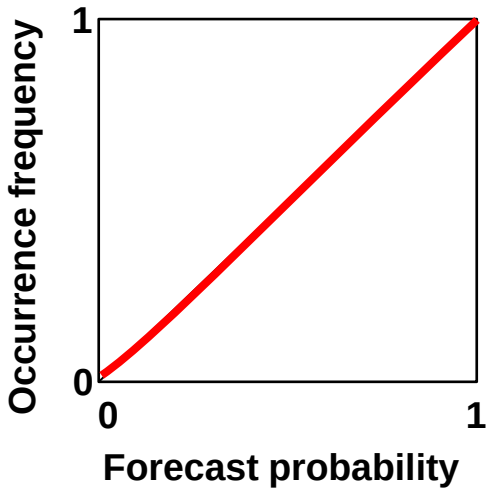
- probabilistic forecast

Over/under confidence

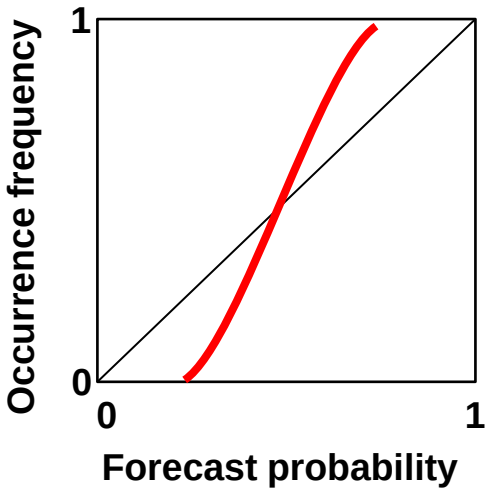
Over confidence



Perfect reliability



Under confidence



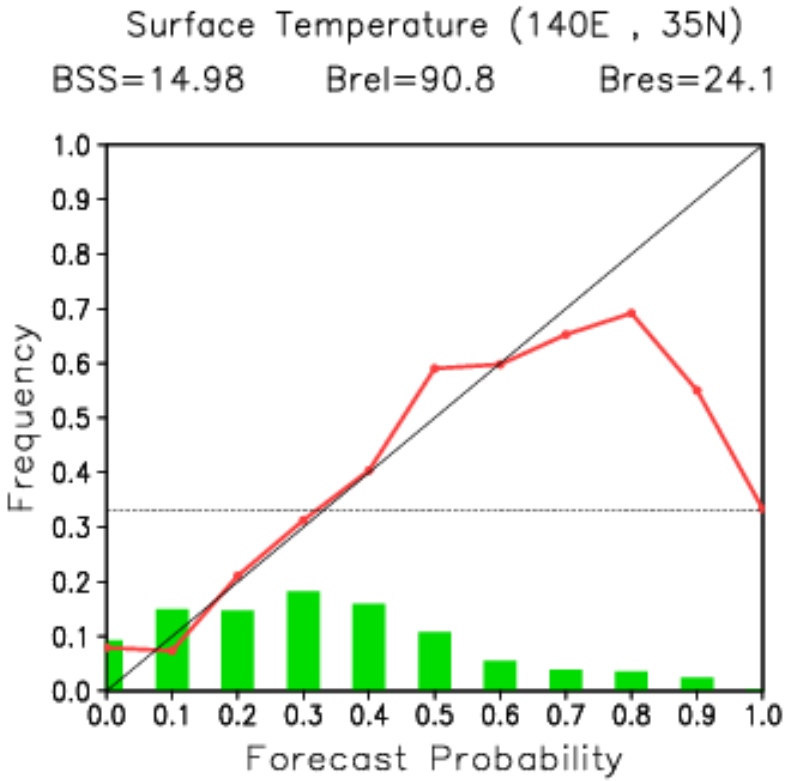
✓ Predicted probabilities are overestimated as compared with actual

✓ 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%

- probabilistic forecast

Brier skill score (BSS)

- 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. Perfect score: 0

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

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

- Perfect score: 1
- BSS>0 : better than the climatological forecast.
- BSS=0 indicates no skill when compared to the climatological forecast.
- BSS<0 : worse than the climatological forecast.