4.9 Verification

Forecasts made by the NWP models are routinely and systematically verified against analyses and/or observations. The verification procedure is based on the practical standard recommended by WMO/CBS and explained in the Manual on the GDPS. Furthermore, several other methods are also adopted to measure the forecast skill on particular phenomena or in specific areas. In this section the operational verifications are explained regarding GSM, RSM, MSM, TYM and One-week EPS.

The specifications of the operational verification are listed in Table 4.9.1 to Table 4.9.7. The verification results for GSM are described in 4.2.14. The categorical verification results for precipitation forecasts of RSM and MSM are presented in 4.4.11 and 4.5.11, respectively. The verification results of TYM for typhoon center position/intensity forecasts are shown in 4.6.7. The verification results for One-week EPS are described in 4.3.4 and 4.3.5. Major indices to score the forecast performance are explained in Table 4.9.8 to Table 4.9.10.

Table 4.9.1 Operational verification against analysis

	GSM	RSM	MSM
verification grid size	2.5 degree	80km	40km
element	Z,U,V and T at 850, 500 and 250 hPa; Psea; RH at 850 and 500 hPa; precipitable water		
score	mean error, root mean square error, tendency correlation, anomaly correlation, standard deviation, S1 skill score		
time interval	24 hours	12 hours	6 hours
verification area	NH(90N-20N), SH(20S-90S), TR(20N-20S), RSM core region	RSM core region	MSM core region

Table 4.9.2 Operational verification against radiosonde data

	GSM	RSM	MSM
element	Z,U,V and T at 850,500 and 250 hPa RH at 850 hPa and 500 hPa		
score	mean error, root mean square error, standard deviation, tendency correlation		
time interval	24 hours	12 hours	6 hours
verification area	NH,SH,TR North America, Europe, Asia, Australia, RSM core region	RSM core region	MSM core region

Table 4.9.3 Operational verification against AMeDAS rain gauge observation

	GSM	RSM	MSM
verification grid size	80km		
score	mean of observation, mean of forecast, root mean square error, correlation, standard deviation of observation, forecast and error		
time interval	12 hours	3, 6, and 12 hours	1, 3, and 6 hours
verification area	whole Japan and 6 districts in Japan		

Table 4.9.4 Operational verification against AMeDAS rain gauge observation (category)

	GSM	RSM	MSM
verification grid size	80km		
score	threat score, bias score		
time interval	12,24 hours	3, 6 and 12 hours	1, 3, and 6 hours
threshold	1, 5, 20, 50 mm		
verification area	whole Japan and 6 districts in Japan		

Table 4.9.5 Operational verification against AMeDAS temperature observation

	GSM	RSM	MSM
verification grid size	not applied	80km	
score	not applied	mean error, root mean square error	
time interval	not applied	3 hours maximum, minimum	3 hours maximum, minimum
verification area	not applied	whole Japan and 6 districts in Japan	

Table 4.9.6 Operational verification against Radar - Raingauge Analyzed Precipitation (category)

	GSM	RSM	MSM
verification grid size	not applied	20, 40 and 80km	5, 10, 20, 40 and 80km
score	not applied	threat score, bias score	
time interval	not applied	3 and 6 hours	1, 3 and 6 hours
threshold	not applied	1, 5, 10, 20 mm	1, 5, 10, 20, 50 mm
verification area	not applied	whole Japan and 6 districts in Japan	

Table 4.9.7 Operational verification for One-week EPS

	Deterministic verification	Probabilistic verification	
analysis	Global analysis on 2.5 x 2.5 degree grid	Global analysis on 2.5 x 2.5 degree grid	Dominant rain gauge observation
forecast	Ensemble mean and all members	Probability	Probability of precipitation
climatology	Climatological fields and standard deviations are calculated from JRA-25. The climatological probability is given by the monthly frequency derived from analysis fields.		
element	Z at 1000 and 500hPa; T at 850 and 500 hPa; U and V at 850 and 250hPa; Psea	Z500, T850, and Psea anomalies with thresholds ±1, ±1.5, ± 2 climatological standard deviation; Z500 anomalies with thresholds ±25m, ±50m; Psea with thresholds 1000hPa; 850 hPa wind speed with thresholds of 10, 15, 25 m/s	Precipitation with thresholds 0.5, 1, 2, 4, 5, 6, 10, 12, 24, 25, 48, and 72 mm/24 hours
score	Mean error, root mean square error, and anomaly correlation coefficient	Brier (skill) score, ROC-area (skill) score, and relative economic value	
time interval	12 hours	12 hours	24 hours
verification area	NH extratropics, East Asia, Japan, Tropics, Western Pacific, SH extratropics, NH, SH, North America, Europe, Asia		Japan

Limited area models have lateral boundaries and thus adjacent zones to the boundaries are strongly affected by an outer coarse mesh model. Therefore such adjacent areas are excluded from the verification of RSM and MSM, and the remaining areas are called core regions.

Observational data mainly used for verification are radiosonde data. The stations used are taken according to the WMO recommendation. Exceptionally, all stations available are used for RSM/MSM verification. All station data are subject to inspection and the data having passed quality control are used for the verification.

JMA operates a high-resolution surface observation network named the Automated Meteorological Data Acquisition System (AMeDAS), which consists of 1300 raingauges, 200 snow gauges and 800 thermometers, aerovanes and heliographs all over Japan. Its estimated grid spacing is about 17 km for raingauges and 20km for other facilities. The AMeDAS data have been used for verifying forecast performance on both precipitation and surface air temperature. The observational data are converted into a set of uniform grid data in 80 km mesh and the forecasts are compared with the gridded observations. This method is adopted to avoid discontinuity caused by changes in model resolution and to reduce sampling error of observation.

JMA produces a precipitation analysis over Japan by compounding radar reflections and rain gauge data. Since the precipitation analysis provides very dense (about 1km) information supplementing the raingauge data, it is employed to evaluate forecast skills on meso-scale disturbances.

The position and intensity of typhoons center is analyzed at the RSMC Tokyo Typhoon Center. It is applied to verify typhoon forecast.

JMA verifies probabilistic forecasts, calculated by counting up the number of members in an event eccurrence at every verification grid and then dividing by the ensemble size, using the method of probabilistic verification. Forecasts of each ensemble member and ensemble mean forecast are also verified using method of deterministic verification.

Table 4.9.8 Verification indices for categorical forecast.

FO, *FX*, *XO*, *XX* are defined as in Table 4.9.9.

Bias Score	$\frac{FO+FX}{FO+XO}$
Threat Score	$\frac{FO}{FO+FX+XO}$

Table 4.9.9 Category matrix for two-event situation

	Observed	Not Observed
Forecasted	*FO*	*FX*
Not Forecasted	*XO*	*XX*

Table 4.9.10 Verification indices

Mean Error	$\frac{1}{n}\sum_{i=1}^{n} D_{t,i}$
Root Mean Square Error	$\sqrt{\frac{1}{n}\sum_{i=1}^{n} {D_{t,i}}^2}$
Standard Deviation	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(D_{t,i}-\overline{D_{t,i}}\right)^2}$
Anomaly Correlation	$\frac{\sum_{i=1}^{n}\left(f_{t,i}-\overline{f_{t,i}}\right)\left(a_{t,i}-\overline{a_{t,i}}\right)}{\sqrt{\sum_{i=1}^{n}\left(f_{t,i}-\overline{f_{t,i}}\right)^2\sum_{i=1}^{n}\left(a_{t,i}-\overline{a_{t,i}}\right)^2}}$
Tendency Correlation	$\frac{\sum_{i=1}^{n}\left(\widetilde{F}_{t,i}-\overline{\widetilde{F}_{t,i}}\right)\left(\widetilde{A}_{t,i}-\overline{\widetilde{A}_{t,i}}\right)}{\sqrt{\sum_{i=1}^{n}\left(\widetilde{F}_{t,i}-\overline{\widetilde{F}_{t,i}}\right)^2\sum_{i=1}^{n}\left(\widetilde{A}_{t,i}-\overline{\widetilde{A}_{t,i}}\right)^2}}$
S1 Skill Score	$100\times\frac{\sum_{i=1}^{n}\left\{\left\lvert\partial_x D_{t,i}\right\rvert+\left\lvert\partial_y D_{t,i}\right\rvert\right\}}{\sum_{i=1}^{n}\left[\max\left(\left\lvert\partial_x F_{t,i}\right\rvert,\left\lvert\partial_y F_{t,i}\right\rvert\right)+\max\left(\left\lvert\partial_x A_{t,i}\right\rvert,\left\lvert\partial_y A_{t,i}\right\rvert\right)\right]}$
Brier Score	$\frac{1}{n}\sum_{i=1}^{n}\left(P_i-O_i\right)^2$

where $F_{t,i}$: forecast value, $A_{t,i}$: analysis or observation, $C_{t,i}$: climate,

P_i: forecasted probability, O_i: {observed:1, or not-observed:0}

$$D_{t,i}=F_{t,i}-A_{t,i},\quad f_{t,i}=F_{t,i}-C_{t,i},\quad \widetilde{F}_{t,i}=F_{t,i}-A_{o,i},$$
$$a_{t,i}=A_{t,i}-C_{t,i},\quad \widetilde{A}_{t,i}=A_{t,i}-A_{o,i},,$$
$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\qquad \partial_x X=\frac{\partial X}{\partial x},\qquad \partial_y X=\frac{\partial X}{\partial y},$$

i: point, t: time (o: t=0) , n: total number.

<u>Reference</u>

WMO, 1992: Manual on the global data-processing system. WMO Technical Document No.485, II.7-36 – II.7-39.