Verification of NCMRWF Global Data Assimilation System

Sujata Pattanayak and V. S. Prasad
National Centre for Medium Range Weather Forecasting
Ministry of Earth Sciences
A-50, Sector-62, Noida-201309, UP, INDIA
sujata@ncmrwf.gov.in (corresponding author)

Abstract

The National Centre for Medium Range Weather Forecasting (NCMRWF) is the mandate to provide observational data analysis to the leading operational centers in India. For this purpose, the Global Data Assimilation system (GDAS), based on National Centers for Environmental Prediction (NCEP) Global Forecasting System (GFS), is integrated in operational mode at NCMRWF since 1994. The performances of numerical models are enhanced with the increase in observational network in recent years with some limitations. Thus, the verification of the analysis and forecast products is crucial to validate and improve the competency of the Numerical Weather Prediction (NWP) systems. This study illustrates the verification statistics of GDAS analysis and the GFS forecast globally and region specific like Northern Hemisphere (NH), Southern Hemisphere (SH), Asia (AS), Europe (EU), North America (NA) and Tropics (TR) as per the Commission for Basic Systems (CBS). This will be helpful in inter-comparison of the models performance used in the operational centers worldwide.

1 Introduction

The National Centre for Medium Range Weather Forecasting (NCMRWF) is the mandate to provide observational data analysis to the leading operational centers in India. The Global Data Assimilation system (GDAS), based on National Centers for Environmental Prediction (NCEP) Global Forecasting System (GFS), is integrated in operational mode at NCMRWF since 1994, though there is a lot of transition in the new version of the system since then.

With the advancement of observational network for both conventional and non-conventional observations, the performances of numerical models are enhanced with some limitations. However, the verification of the analysis and forecast products is crucial to validate and improve the competency of the Numerical Weather Prediction (NWP) systems. In the present study, the verification statistics of GDAS analysis and the GFS forecast is illustrated. The brief on the GDAS is presented in section 2, data used and methodology in section 3 with glimpses of results in section 4 with conclusions.

2 GDAS Analysis System

The GDAS analysis system is initially implemented at a horizontal resolution of T80 with 18 vertical layers (T80L18) on CRAY supercomputer with Spectral Statistical Interpolation (SSI) and continued to provide operational analysis product till 2007 (Rajagopal et al., 2007; Prasad et al., 2011). In 2007, it has been upgraded to T254L64. The system is upgraded with Grid Point Statistical (GSI) assimilation technique in 2008 and since then continuous effort is made to improve the performance of the analysis system. In the recent years, the analysis system upgraded with the hybrid-GSI and 4-Dimentional Variational data assimilation techniques.

3 Data and Methodology

In the present study, the validation of GDAS analysis product is examined with the radiosonde (RS/RW) observations received at NCMRWF. The performance is evaluated on daily basis and monthly. The mean error, standard deviation, correlation coefficient, anomaly correlation, and skill etc. are evaluated for temperature, geopotential height, u- and v-wind components at standard pressure levels. The model performance is evaluated globally and region specific like Northern Hemisphere (NH), Southern Hemisphere (SH), Asia (AS), Europe (EU), North America (NA) and Tropics (TR) as per the Commission for Basic Systems (CBS).

The error statistics are calculated as follows;

Mean Error (ME)

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(f_i - o_i) = \overline{f} - \overline{o}$$

Forecast standard deviation

$$SD_f = \sqrt{\frac{1}{T+1}\sum_{i=1}^{T}(f_i - \bar{f})^2} \qquad \overline{f} = \frac{1}{n}\sum_{i=1}^{n} f_i$$

Observation standard deviation

$$SD_o = \sqrt{\frac{1}{T+1}\sum_{i=1}^{T}(o_i - \bar{o})^2} \qquad \overline{o} = \frac{1}{n}\sum_{i=1}^{n} o_i$$

The mean error in the analysis gives the bias in the analysis. The Pearson correlation coefficient measures the linear association between forecast and observations. It is defined as,

$$r = \frac{\sum_{i=1}^{T}(f_i - \overline{f})(o_i - \overline{o})}{\sqrt{\sum(f_i - \overline{f})^2}\sqrt{\sum(o_i - \overline{o})^2}}$$

Also, the anomaly correlation coefficient is estimated which is the equivalent to the Pearson correlation coefficient, but both the observation and forecast are first adjusted to the climatological value. It measures the strength of linear association between forecast anomalies and observed anomalies. It is defined as,

$$\mathrm{ANOM\_CORR} = \frac{\sum(f_i - c)(o_i - c)}{\sqrt{\sum(f_i - c)^2}\sqrt{\sum(o_i - c)^2}}$$

The skill of the analysis evaluated through Gilbert Skill Score (GSS) and Heidke Skill Score (HSS). It is defined as follows.

$$GSS = \frac{a - \frac{(a+b)(a+c)}{a+b+c+d}}{a+b+c-\frac{(a+b)(a+c)}{a+b+c+d}}$$

$$HSS = \frac{a+d-K}{T-K}$$

$$K = \frac{(a+b)(a+c)+(c+d)(b+d)}{a+b+c+d}$$

The value of a, b, c, d are measures as in Table 1.

Table 1. Observation and Forecast measures to evaluate skill

Forecast	observation		Total
	o = 1 (Yes)	o = 0 (No)	
f = 1 (Yes)	a	b	a + b
f = 0 (No)	c	d	c + d
Total	a + c	b + d	T = a + b + c + d

4 Results and Discussion

The glimpses of the results are presented in this section. Figure 1 shows the Pearson Correlation coefficient and the anomaly correlation.

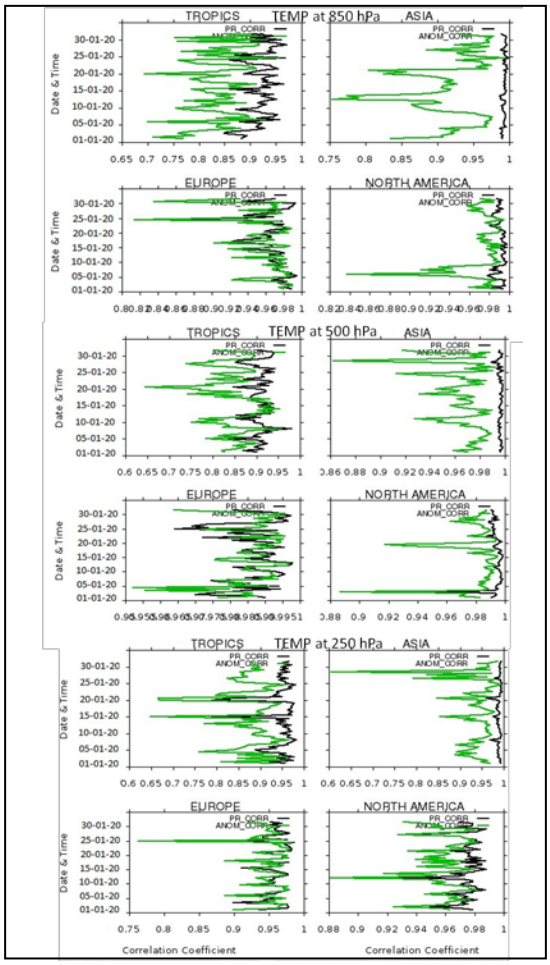

Figure 1. Pearson and anomaly Correlation of temperature at 850, 500 and 250 hPa.

The vertical profile of bias and RMSE shows less errors in the low troposperic level and increases in middle and upper troposperic levels. Figure 2 demonstrates the skill of the analysis in global and region specific. The skill (Gilbert skill score) of the model also shows more than 0.9 in GL, NH, SH, EU, NA domains; with some deviations in few days over AS and TR regions.

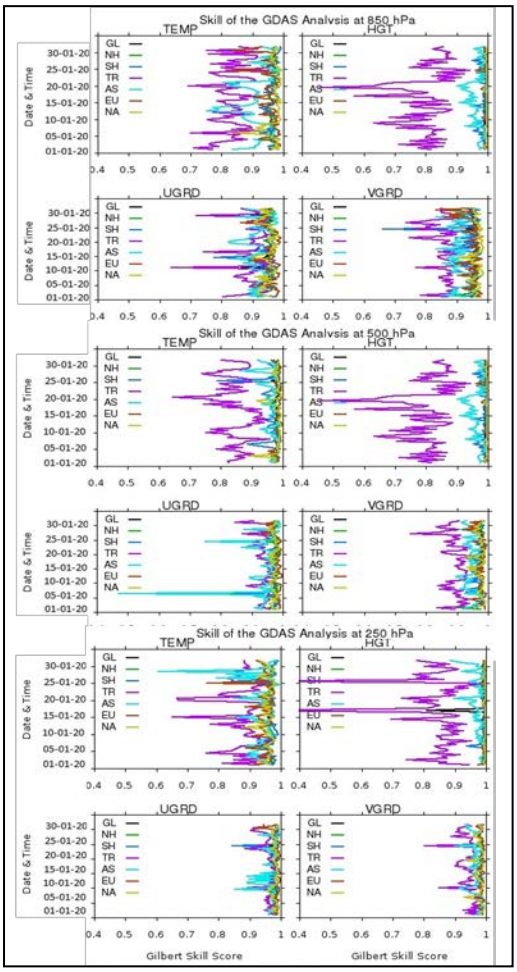

Figure 2. Skill of the analysis system in GL, NH, SH, AS, TR, EU and NA regions.

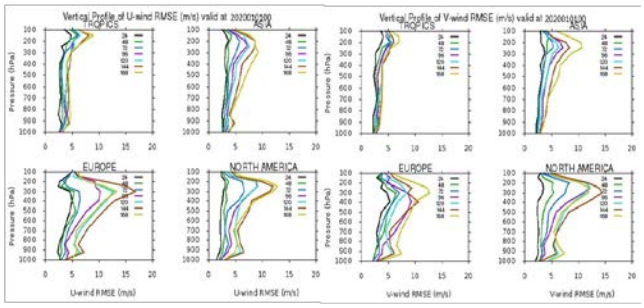

Figure 3. Validation of the model forecast for U- and V-wind components.

The analysis is validated up to 7 day and Figure 3 demonstrates the RMSE of wind component. The model bias and RMSE increase with time, but AS and TR regions show less error than other regions.

References

Rajagopal et al (2007): Implementation of T254L64 Global Forecast System at NCMRWF.

V.S. Prasad, et al (2011): Implementation of Upgraded Global Forecasting Systems (T382L64 and T574L64) at NCMRWF.