# Verification of Quantitative Precipitation Forecast from Operational Numerical Weather Prediction Models over Japan (WGNE precipitation forecast intercomparison project)

Masayuki Hirai, Takuya Sakashita
Numerical Prediction Division, Japan Meteorological Agency
(E-mail: m-hirai@met.kishou.go.jp, sakashita@met.kishou.go.jp)

## 1. Introduction

In 1995 the WGNE initiated the verification of quantitative precipitation forecasts (QPFs) from operational NWP models over different areas of the globe. A number of results of this project have been already reported (e.g., Goeber et al. 2002; Ebert et al. 2003). We also undertook the QPFs verification over Japan last year (2002). This paper reports the findings of our verification results until September 2003.

## 2. Verification Methods

Table 1 indicates the specifications of the QPFs data sent by each NWP centers as of December 2003. The observational precipitation data are referred to the operational high-dense (17 × 17 km) rain gauge observation network. Both the observation data and the model forecasts data are interpolated into the verification grid, whose resolution is 80km.

## 3. Verification Results

### (1) 24-h QPFs Verification Results

Figure1 shows the frequency bias score (BS) for 24-h precipitation in day 3 (FT48~72) in summer of 2003. All models tend to overestimate the frequency of light precipitation, though there are differences in degree. Similar features are also reported in other regions (Goeber et al. 2002; Ebert et al. 2003). Most models underestimate the frequency of moderate or intense (>20mm/24h) precipitation.

Monthly time series of BS and the equitable threat score (ETS) for 24-h precipitation in day 3 (FT48~72) are shown in Figure 2. BS for the threshold of 1mm/24h is larger than 1.0 in most models all year around. Since meso-scale convective systems are dominant in precipitation associated with Asian summer monsoon in Japan region, all models tend to decrease ETS in summer. It is also found that some models show low ETS in winter monsoon season due to the overestimation of frequency (high BS).

### (2) 6-h QPFs Verification Results

Although a number of investigations have been made on accuracy of 24-h precipitation forecasts, there is little report on precipitation forecasts in shorter timescale (6 or 12-h). It is expected that the verification of 6-h or 12-h QPFs reveals characteristics on diurnal variation.

Figure 3 indicates BS and ETS for 6-h forecasts in summer of 2003. BS for each model is larger in daytime (00~06UTC or 06~12UTC) than nighttime. BS for some models at FT00~06 is high despite nighttime (12~18UTC) indicating these models have so called spin-down problem at the beginning of forecast. ETSs for these models, therefore, are lower at FT00~06 than FT06~12.

## References

Ebert, E. E., U. Damrath, W. Wergen and M. E.Baldwin, 2003: The WGNE assessment of Short-term Quantitative Precipitation Forecasts. Bull. Am. Meteorol. Soc., **84**, 481-492.

M. Goeber, S.F. Milton and C.A. Wilson, 2002: WGNE assessment of Quantitative Precipitation Forecasts from Operational Numerical Weather Prediction Models over the U.K. WMO/TD **No.1105**, WMO, 02-07 - 02-08.

Table 1. The specifications of the QPFs data sent by NWP centers as of December 2003.

| NWP center | horizontal resolution of data(°) | forecast time (h) | verified since | |
|---|---|---|---|---|
| ABoM | 1.25×1.25 | 12,24,36,…,120 | Aug 2002 | *1 |
| DWD | 0.75×0.75 | 24,48,72 | Jul 2002 | *2 |
| ECMWF | 0.50×0.50 | 6,12,18,…,72 | Apr 2002 | *3 |
| NCEP | 1.00×1.00 | 6,12,18,…,72 | Aug 2002 | *4 |
| UKMO | 0.83×0.56 | 6,12,18,…,96 | Oct 2001 | *5 |
| JMA | 0.56×0.56 | 3,6,9,12,…,72 | Apr 2002 | *6 |

*1: Australian Bureau of Meteorology
*2: Deutscher Wetterdienst
*3: European Centre for Medium-Range Weather Forecasts
*4: National Centers for Environment Prediction (Aviation model)
*5: United Kingdom Meteorological Office
12-h accumulated QPFs data received until Sep 2002.
*6: Japan Meteorological Agency

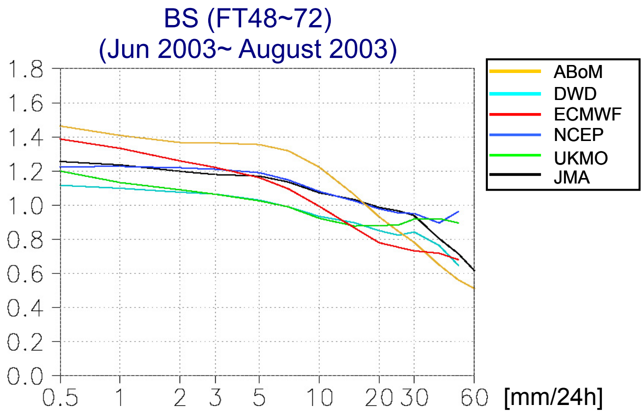


Fig. 1. Bias Score as a function of precipitation threshold for 24-h precipitation in day 3 (FT48~72) during June to August 2003. Initial time for each model is 12 UTC. The score is not plotted when the number of event in either observation or forecast is less than 450 in case of high threshold values.

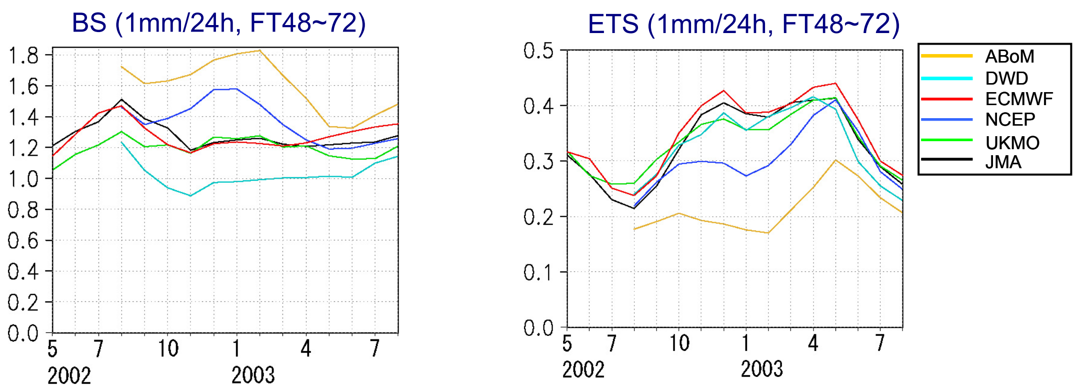


Fig. 2. Monthly time series of BS (left) and ETS (right) for 24-h precipitation in day 3 (FT48~72) from May 2002 to August 2003. The threshold is 1[mm/24h]. Scores are calculated for 3 consecutive months (from the previous month to the next).

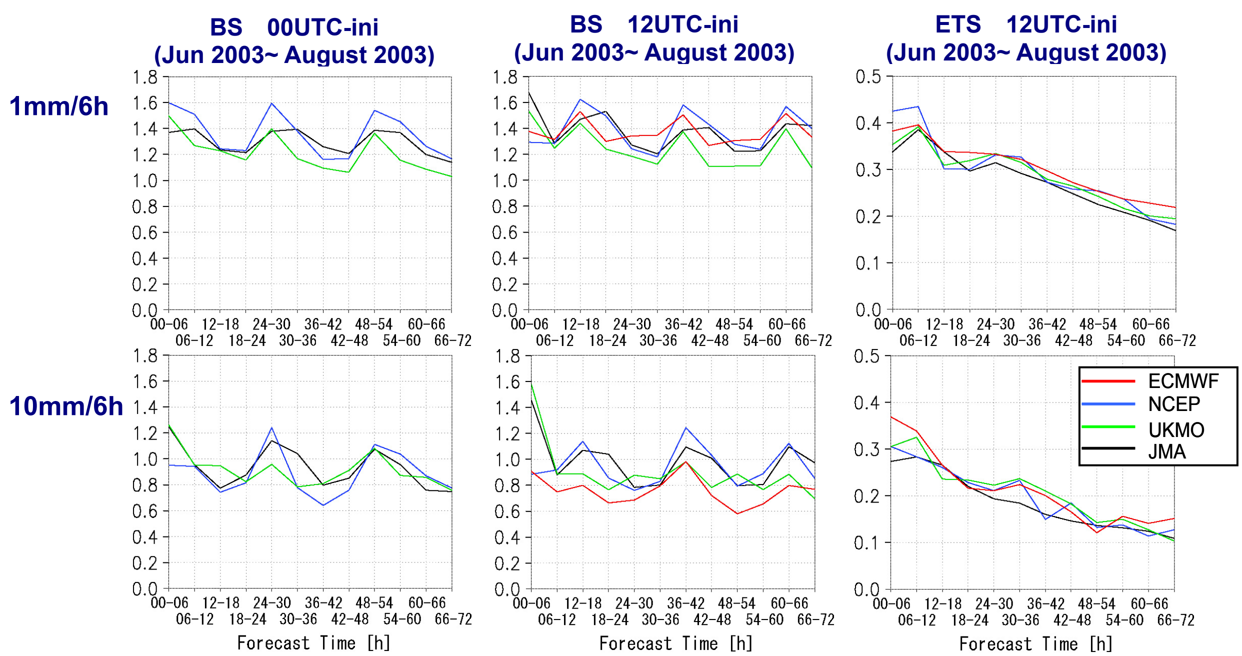


Fig. 3. Monthly time series of BS (left and middle) and ETS (right) during June 2003 to August 2003 as the functions of forecast time. Precipitation threshold is 1[mm/6h] (above) and 10[mm/6h] (below).

**WGNE NWP stratospheric prediction comparison**

Greg Roff[1], Agathe Untch[2], John McCormack[3], Mark Iredell[4] and Mike Keil[5]

[1] *BMRC*, [2] *ECMWF*, [3] *NRL*, [4] *NCEP*, [5] *UKMO*

Contact: *Australian Bureau of Meteorology Research Centre, 150 Lonsdale St, Melbourne, Australia*

Email: g.roff@bom.gov.au

In this study we examine how well NWP models simulate the stratosphere when the polar vortex undergoes large changes. To do this we compare analyses for the period from 15 Sep.-15 Oct. 2002 (Days 0-30 in this study) and forecasts from 20 Sep.- 3 Oct. 2002 (Days 5-18) during the southern hemisphere major sudden warming of 2002 from five current NWP models: the Australian BMRC Atmospheric Model (BAM); the ECMWF IFS; the NCEP MRF; the NRL NOGAPS and the UKMO model. These models provided forecasts out to 8, 10, 10, 5 and 10 days, respectively. TOMS plots (not shown) indicate that the vortex started to deform on 20 Sep. (Day 5), split in two by 24 Sep. (Day 9) and had a single vortex centre again by 30 Sep. (Day 15)

Figure 1 shows the 30 hPa temperature RMSE between the model 5-day forecasts and their respective analyses starting at 00UTC on 20 September (Day 5) for 14 days to 12UTC 3 October 2002 (Day 18.5) averaged over latitudes $60^oS$ to $90^oS$ for the four models BAM, ECMWF, NCEP and NOGAPS. For forecasts initiated on Days 5-10 (20-25 September), when the vortex was splitting, all the models have almost continuously increasing RMSE for any given forecast. This implies that over this period there is a steady reduction of forecast skill and that this is an increasingly difficult period for all the models. From initialization Days 10-12 (25-27 September) the skill in all the models is seen to improve.

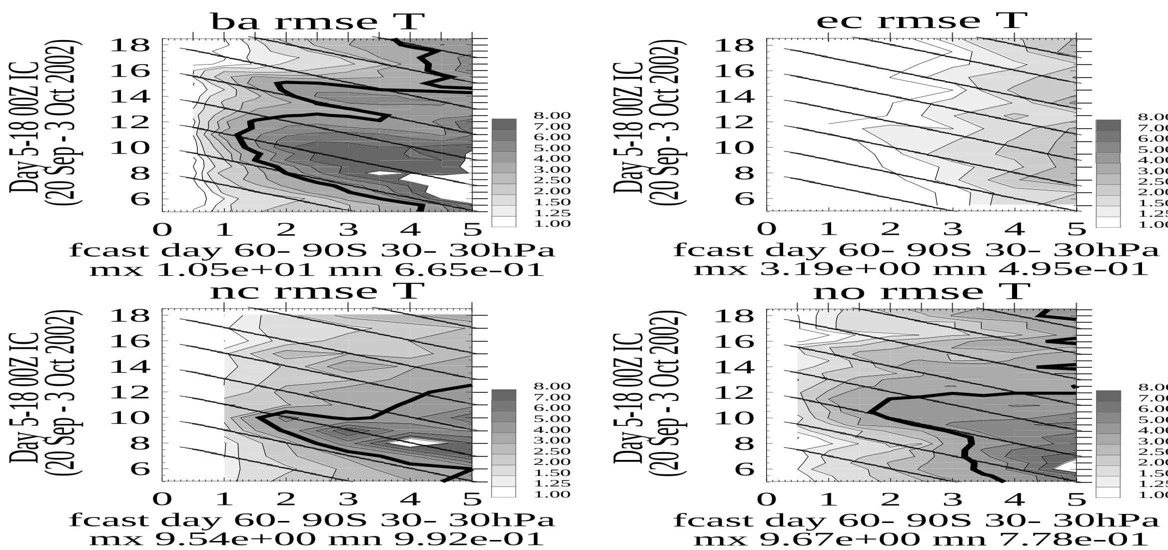


Figure 1: *Plots of forecast day (0-5) against initialization analysis day from 00UTC 20 September to 12UTC 3 October 2002 (Days 5-18.5) of the 30 hPa temperature field (K) RMSE between the BAM, ECMWF, NCEP and NOGAPS forecasts and their respective analyses averaged over latitudes $60^oS$ to $90^oS$. The maximum and minimum are below each plot, the diagonals are lines of constant verification day and contour intervals are the same and are indicated. The thick contour is the $4K$ contour line.*

The average of these RMSE values over the initialization and forecast days are seen in Figure 2. All models have errors increasing with forecast time and show that initialization Days 10-12 (25-27 Sep. 2002) separate an early period with large RMSE from a later period with smaller errors.

Can these RMSE difficulties be related to particular days? If this is true then there should be a strong dependence of the RMSE on the verification day, where lines of constant verification day are indicated by the diagonals in Figure 1. In this figure there are contoured regions which extend along the verification diagonals encompassing one or several verification days eg all the models show that the periods 27-28 September (Days 12-13) and around 2 October (Day 17) are dynamical situations which they have difficulty with forecasting. These periods are when the split vortex is decaying and when the reformed vortex is moving westward, respectively.

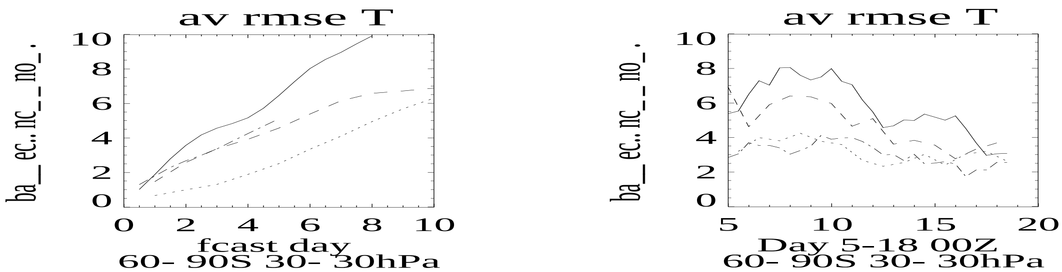


Figure 2: *Plots of the mean 30 hPa BAM, ECMWF, NCEP and NOGAPS temperature fields (K) RMSE and (left) averaged over initiation day for forecast days (0-10) and (right) averaged over forecast day for initiation Days 5-18.5. The line style for each model is indicated.*

What do these forecast errors look like? We plot in Figure 3 the 50 hPa polar stereo plot of the geopotential height field for the first and last forecast from the BAM , ECMWF and NCEP models initiated on 23 September. We can see the large changes that the vortex undergoes over this period, going from a two cell vortex to a single cell, but all the models show a final forecast vortex which is smaller, more circular, more poleward and more westerly displaced and with a more easterly orientation, though the latter is not as obvious in the ECMWF case. The creation of a smaller, more circular and more polewardly displaced vortex indicates that all the models are trying to create weaker and less disturbed vortices.

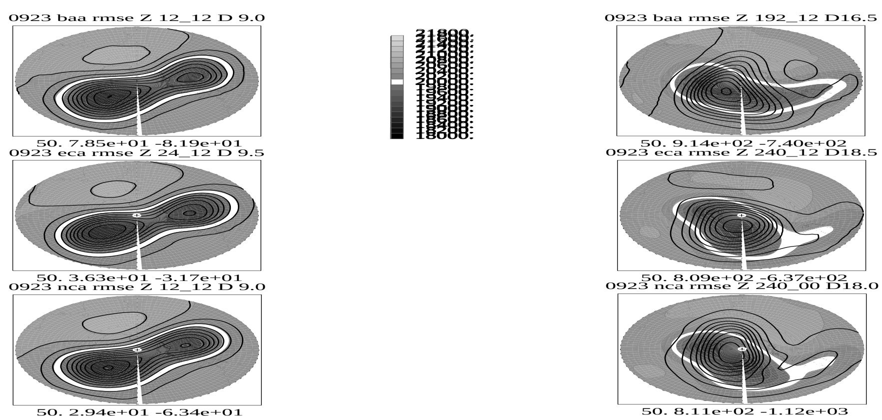


Figure 3: *50 hPa polar stereo plot of the geopotential height (m) field for the (left) first and (right) last forecast from each of the models (top) BAM, (middle) ECMWF and (bottom) NCEP initiated on 23 September.The forecasts are the line contours while the corresponding analyses are the filled contours. The forecast time (in hours) and corresponding verification Day (eg D 9.0 in the first plot) are at the top of each diagram.*

The forecasting ability of the BAM, ECMWF, NCEP and NOGAPS NWP models has been studied during this sudden warming and we find that if the vortex undergoes rapid changes after forecast initialization all the models have some degree of difficulty in capturing this event. There are certain verification days, common to all the forecast models, which the model forecasts have difficulty, and these errors are largest in the stratosphere. These characteristic errors were that the forecast vortex was seen to be: smaller; displaced westward; displaced poleward; have a faster easterly rotation of its orientation and to be more circular.

We also found that the BAM, ECMWF, NCEP, NOGAPS and UKMO NWP models analyses are well correlated over the period of our study when the vortex is quasi-stationary but that when the polar vortex is undergoing rapid changes these analyses are seen to have larger RMSE differences and to become less correlated. Also during these active periods the model analyses correlations with TOMS total column ozone decreases dramatically from the very high values found when the vortex is quiescent.

# Seasonal Climate Signatures in the FSU Climate Model Coupled to the CLM2

D. W. Shin, S. Cocke, T. E. LaRow, and J. J. O'Brien
Center for Ocean-Atmosphere Prediction Studies, The Florida State University, Tallahassee, Florida
(shin@coaps.fsu.edu)

The recently released community land model (CLM2) is coupled to the Florida State University (FSU) climate model (Cocke and LaRow, 2000) to improve land surface properties and investigate its role in the seasonal climate studies. The previously used FSU simplified land surface scheme includes a 3 layer soil temperature model based on the force-restore method. Surface characteristics are determined from the USGS 24 category land use/land cover survey. Seasonally varying climatological values for soil moisture, land albedo and surface roughness are prescribed based on the USGS data. Meanwhile, the CLM2 is a new and advanced land surface model (Bonan et. al, 2002 and Zeng et. al, 2002). With improved physical parameterizations, it uses five primary subgrid land cover types (glacier, lake, wetland, urban, vegetated) in each grid. The vegetated portion of a grid is further divided into patches of plant functional types obtained from satellite data.

Simulations of 10-yr length (1987-1996) were performed with each land model and four convective schemes (NCEP/SAS: moisture flux, only one cloud type, NCAR/ZM: similar to the AS but three significant assumptions, NRL/RAS: handing of detrainment, MIT/EMANUEL: buoyancy-sorting hypothesis, mixing hypothesis, and a stochastic coalescence model) coupled to the FSU climate model at a resolution of T63 ($\sim$ 1.86 °) with 17 vertical levels. The integrations commence on 1 January, 1987. Only the last 5 yr of the simulations (i.e., 1992-1996) were analyzed to allow a 5-yr spinup of soil water and temperature for the FSUCLM run.

Simulations with the atmospheric model coupled to the CLM2 (hereafter, CLM) are compared to the control (the original FSU model, FSUc). In Fig. 1, surface (2 m) air temperatures (°K) of FSUc and FSUCLM are compared to the Willmott and Matsuura (2002) observations for the DJF (upper left 3 panels) and JJA (bottom left 3 panels). Meanwhile, precipitation (mm/d) of FSUc and FSUCLM are compared to the Willmott and Matsuura observations for the same season in the right panel. Results from the NCEP scheme are only shown here. As evident from the figure, the FSUCLM experiment improves the seasonal simulation of both surface air temperature and precipitation compared to the control. The FSUCLM reduced much of the surface temperature cold bias noted in the FSUc run. The wet bias in the FSUc was reduced as well especially over the Eurasia during the JJA. Figure 2 shows skill scores in terms of RMSE for surface air temperature in the upper panel and precipitation in the bottom panel. Each of two versions of land models and four versions of convective schemes are compared. Noticeable improvements are evident in the simulation of both variables.

---

Computations were performed on the IBM SP4 at the FSU. COAPS receives its base support from the Applied Research Center, funded by NOAA Office of Global Programs awarded to

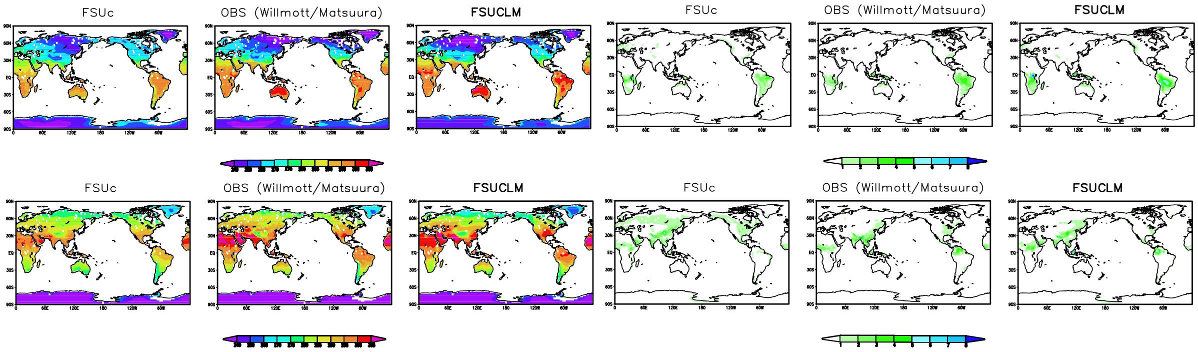


Figure 1:

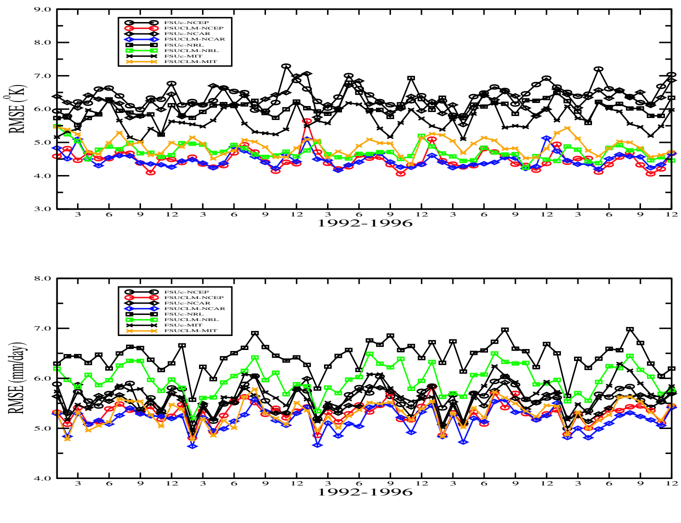


Figure 2:

Dr. James J. O'Brien.

## REFERENCES

Bonan G. B., et. al, 2002: The land surface climatology of the community land model coupled to the NCAR community climate model. *J. Climate*, **15**, 3123-3149.
Cocke, S., and T. E. LaRow, 2000: Seasonal predictions using a regional spectral model embedded within a coupled ocean-atmosphere model. *Mon. Wea. Rev.*, **128**, 689-708.
Zeng X., et. al, 2002: Coupling of the common land model to the NCAR community climate model. *J. Climate*, **15**, 1832-1854.
Willmott C. J., and K. Matsuura, 2002: Terrestrial air temperature and precipitation. [Available on-line from http://climate.geog.udel.edu/∼climate]

# Monthly forecasting at ECMWF

Frederic VITART
ECMWF
nec@ecmwf.int

## 1. The ECMWF monthly forecasting system

The main goal of the ECMWF monthly forecasting system is to produce forecasts from day 10 to day 30, in order to fill the gap between the medium-range forecasts and the seasonal forecasts. Therefore, the monthly forecasting system has been built as a combination of the medium-range EPS and the seasonal forecasting system. It contains features of both systems and, in particular, is based on coupled ocean-atmosphere integrations, as is the seasonal forecasting system.

The ECMWF monthly forecasts are based on an ensemble of 51 coupled ocean-atmosphere integrations (one control and 50 perturbed forecasts). The length of the coupled integrations is 32 days, and the frequency of the monthly forecasts is currently every 2 weeks. The atmospheric component is IFS, with the same cycle as ECMWF operational forecast. Currently, the atmospheric model is run at TL159 resolution with 40 vertical levels in the vertical. The oceanic component is the same as for the current ECMWF seasonal forecasting system based on HOPE from MPI. The ocean and atmosphere communicate with each other through a coupling interface (OASIS from CERFACS). The atmospheric fluxes of momentum, heat and fresh water are passed to the ocean every hour.

The 51-member ensemble is generated by perturbing the atmospheric initial conditions using singular vectors (in the extratopics but also in some tropical regions) and the oceanic initial conditions by applying SST perturbations to 5 different ocean analyses. In addition, stochastic perturbations are applied throughout the atmospheric integrations. In order to calibrate the system, a 5-member ensemble hindcast is run with the same starting day and month as the real time forecast for each of the past 12 years.

## 2. Verification

The ECMWF monthly forecasting system is running every 2 weeks since March 2002. Products include anomaly, probability and tercile maps of 2-meter temperature, surface temperature, mean sea-level pressure and precipitation averaged over 4 weekly periods (days 5-11, days 12-18, days 19-25 and days 26-32). 30 real-time cases have been verified.

For all 30 real-time cases, the anomaly correlation and RMS scores of the ensemble mean have been calculated, along with probabilistic scores such as Brier skill scores, ROC areas or potential economic value.

Results suggest that during the 10 first days of the forecast, the skill of the monthly forecasting system is close to that of the EPS. Over the period days 12-18, the monthly forecasting system produces forecasts that are generally better than climatology or persistence (see example in Figures 1 and 2). Therefore, the monthly forecasting system is probably useful for forecasts at this time-range. Summer seems to be a difficult season as in the medium-range and probabilistic scores over Europe are generally lower than over other regions like North America or Asia.

During the two following weeks (from day 19 to day 32), the coupled model performs generally better than persistence. At this time range, the model's skill increases with higher thresholds. The model displays some skill over some areas like North America and the Southern Extratropics.

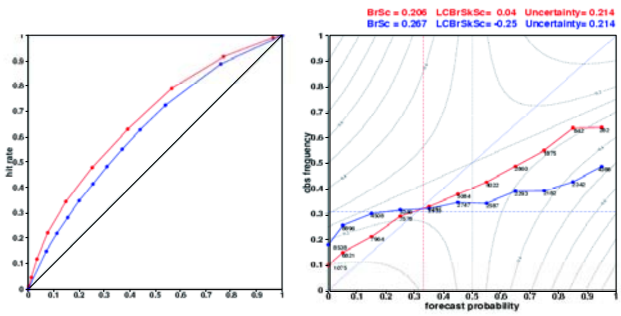


Figure 1: ROC (left panel) and reliability (right panel) diagrams of the probability that the 2-metre temperature is in the upper tercile. Only land points in the Northern Hemisphere have been considered. The red curves represent the diagrams obtained with the monthly forecasting system. The blue curves (closest to the diagonal in the left panel and the most horizontal in the right panel) correspond to the diagrams obtained by persisting the anomalies from the previous week (days 5-11). For the ROC diagram, the closer the curve is to the top left corner, the better is the forecast. For the reliability diagram, the closer the curve is to the diagonal, the most is reliable the forecast.

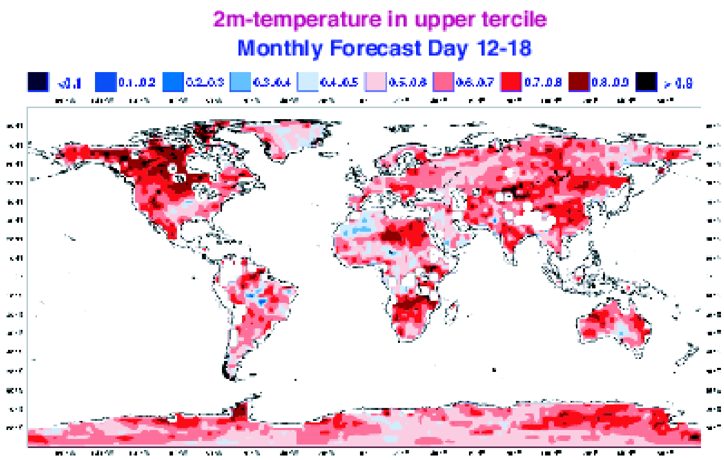


Figure 2: Map of ROC areas of probability that the 2-metre temperature anomaly is in the upper tercile. The verification period is March 2002-May 2003. The red color-scale corresponds to ROC scores higher than 0.5 (better than climatology). The blue color-scale corresponds to ROC scores lower than 0.5 (worse than climatology). In this figure, the red colour is largely dominating, indicating that the model performs better than climatology at this time-range( for colour graphics, see the version of this paper on the web).