Objective verification of the GEM 4DVAR parallel runs
December to February 2005

1. **GEM Global model**

1a. Upper air fields

Verification scores presented here are based on an international verification data exchange program through the World Meteorological Organization (WMO) Commission for Basic Systems. Objective verification scores against radiosonde observations are calculated for geopotential height, temperatures and winds at 850, 500 and 250 hPa, over a number of strictly defined domains. Scores are also calculated against each model's own analysis. These scores presented here were calculated over the period of the official parallel run, from December 7, 2004 to March 10, 2005.

RMS (root mean square) errors for the model run using the 4DVAR assimilation system were generally lower than the operational model, for all three variables and over most observation domains.

Over the Northern Hemisphere, the largest improvements occurred over North America (Fig. 1a), in the 2-5 day period, where reduction in errors was generally between 3-9%. Geopotential height errors were reduced more than the wind and temperature errors. Over Asia and Europe, the signal was more neutral, but with somewhat of an improvement in the 8-10 day period over the latter domain.

Over the Southern Hemisphere, there was a relatively strong reduction in height errors, and more so at 500 and 250 hPa than at 850 hPa. Wind and temperature errors were also reduced more in the mid and upper levels, but overall the extent of the reduction was not as great as for the geopotential heights.

Overall, performance of the new system is illustrated in Fig. 2, which shows a slight gain in 500 hPa anomaly correlation over the Northern Hemisphere, but a strong gain over the Southern Hemisphere.

1b. Quantitative Precipitation Forecasts (QPF)

Precipitation forecasts are scored against both the North American synoptic network and the high density U.S. SHEF network. Scores were calculated over the period Dec. 2, 2004 to Feb. 28, 2005.

For the day 1-2 forecasts, equitable threat scores against the synoptic network were neutral to slightly positive in favour of the parallel run model in almost all categories and forecast time periods (Fig. 2). Against the SHEF network there was a slight deterioration in the 24-48 hour forecasts of precipitation amounts above 5mm/24 hour (Fig. 3). The parallel run model systematically scored better at discriminating precipitation from no precipitation cases.

In the 3-5 day period, scores were generally neutral to slightly positive in favour of the parallel run model.

2. GEM Regional model

2a. Upper air fields

A modest gain (less than 5%) was noted in RMS errors for all variables, levels and forecast time periods over the North American radiosonde observation domain (Fig. 1b).

2b. Quantitative Precipitation Forecasts (QPF)

The signal from the equitable threat scores was generally neutral to slightly positive in favour of the parallel run model against both the North American synoptic and U.S. SHEF networks, for amounts up to 25mm/24 hour. Against the synoptic network there was a slight deterioration for amounts in excess of 25mm.   However, the number of cases was relatively small, and these were dominated by Pacific coast orographic precipitation events during the latter half of January.
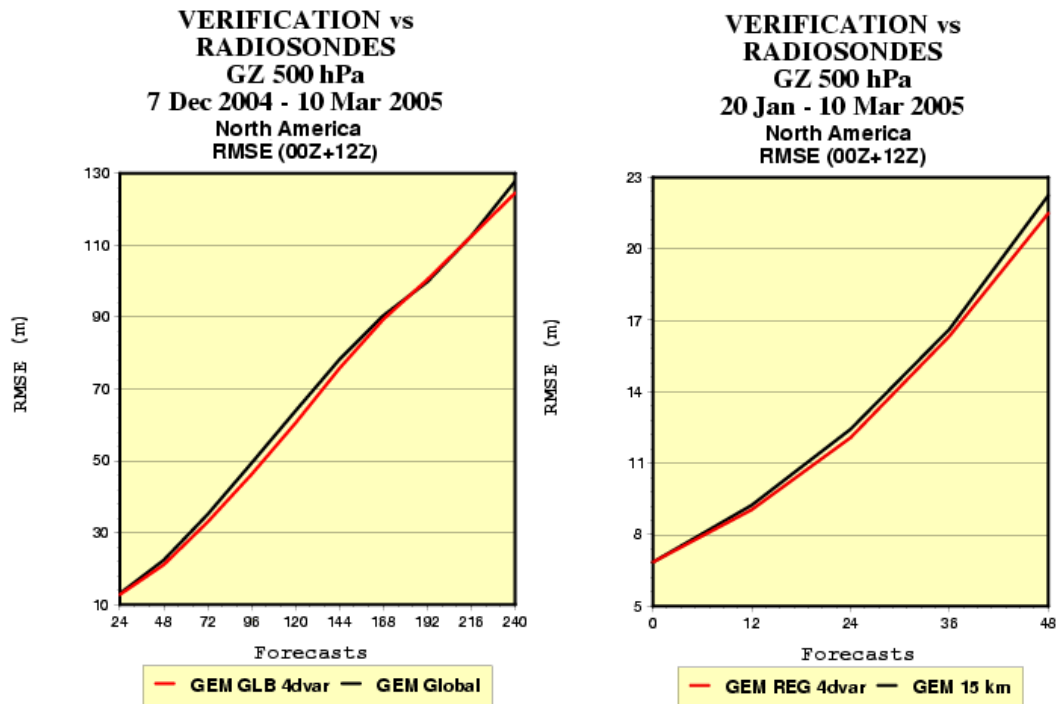
Figure 1: 500 hPa RMS (root mean square) errors for a. GEM Global operational (black) and parallel run (red) models, and b. GEM Regional operational and parallel run models, against the North American radiosonde observation network. Global model data includes 185 cases, while 100 cases were used for the Regional model.
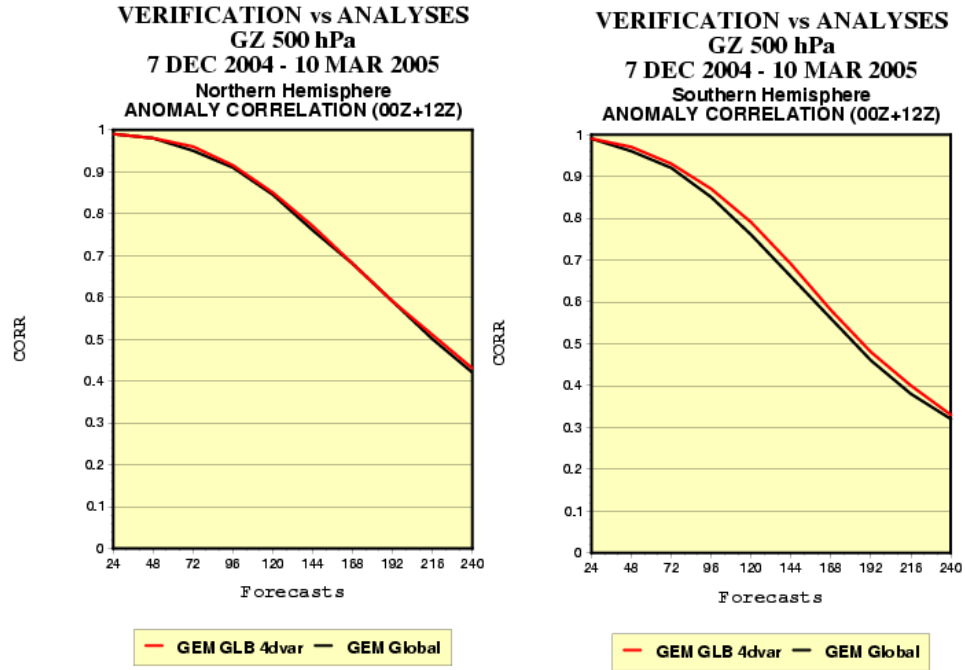
Figure 2: 500 hPa anomaly correlation over the Northern (left) and Southern (right) hemispheres.
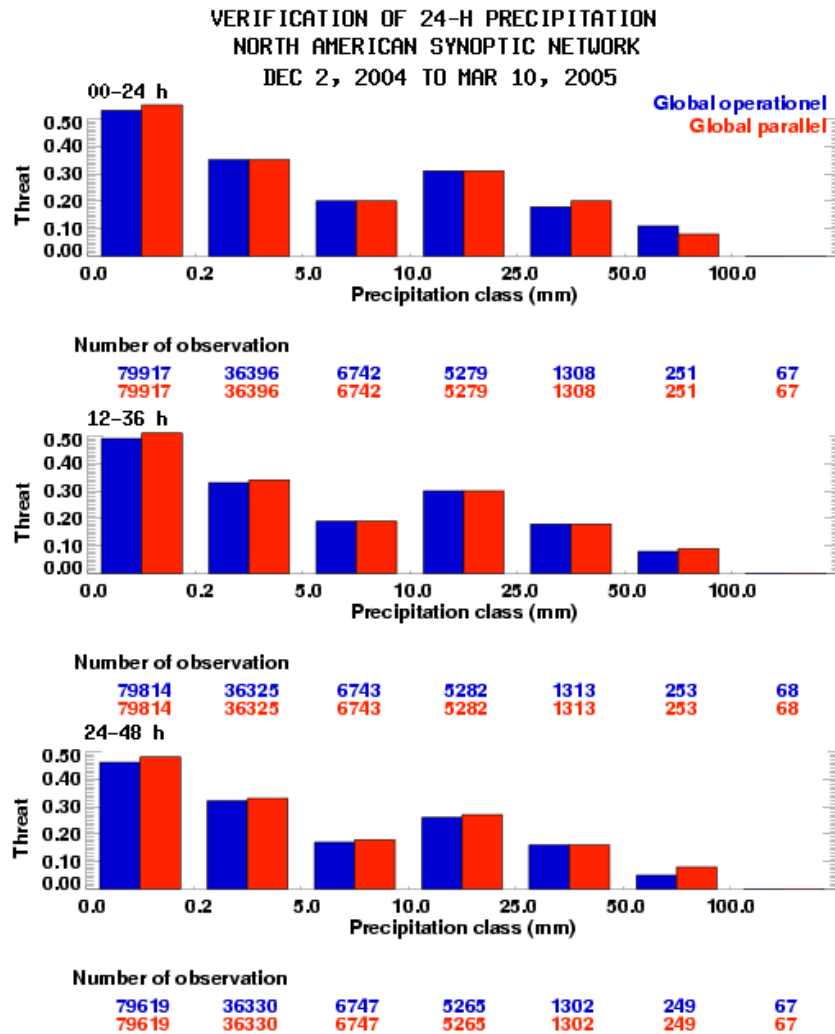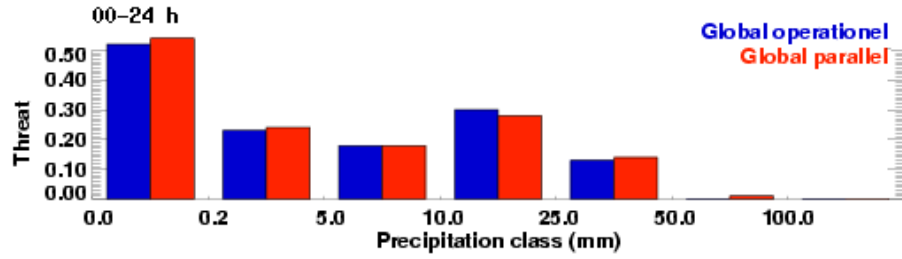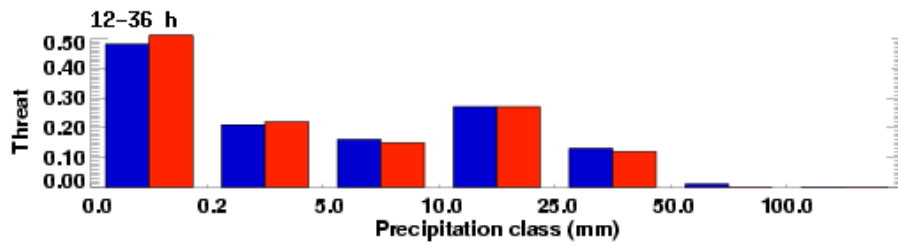
Figure 3: Equitable Threat Scores for the Global operational (blue) vs. parallel run (red) model for three forecast time periods (00-24 h, 12-36 h and 24-48 h) against the North American synoptic observation network.

VERIFICATION OF 24-H PRECIPITATION
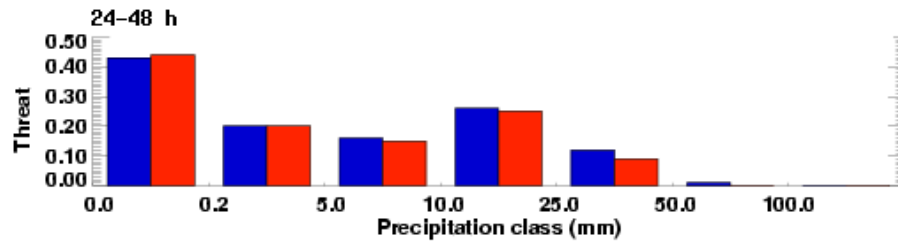**U.S. SHEF NETWORK**
DEC 2, 2004 TO MAR 10, 2005

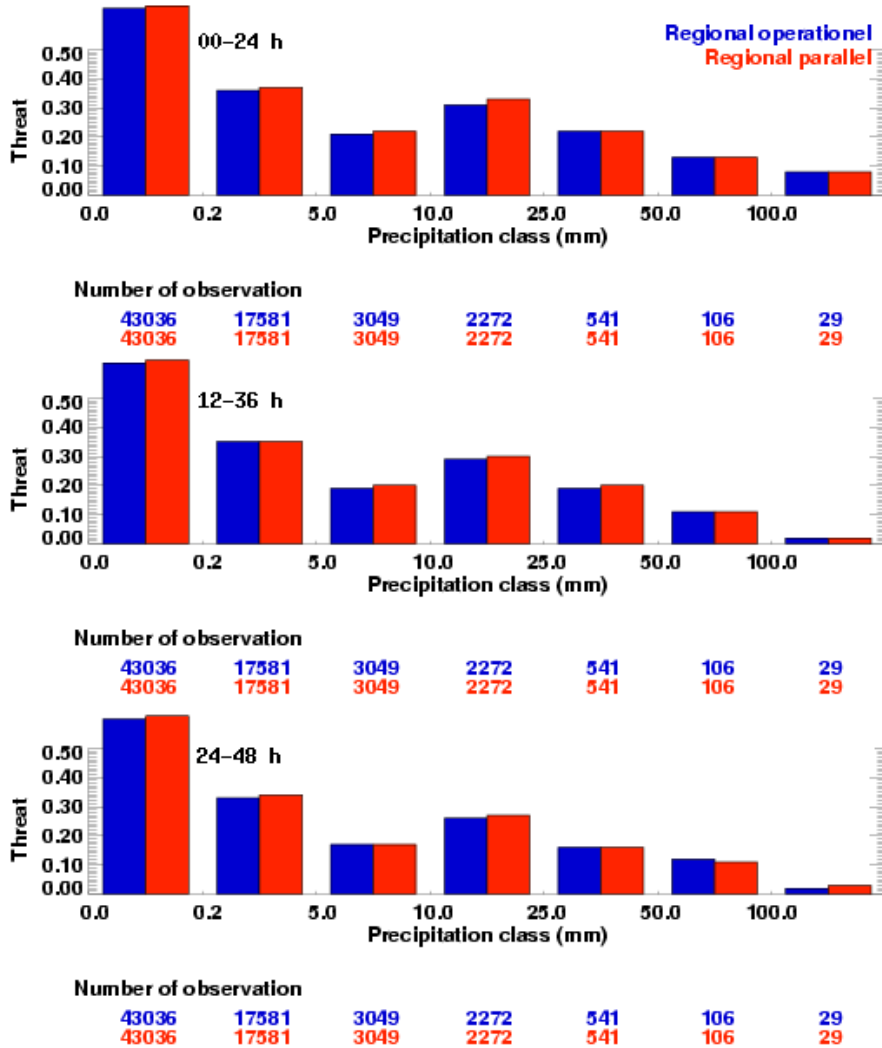**Figure 4: Same as Fig. 3, against the U.S. SHEF network**

Figure 5: Equitable Threat Scores for the Regional operational (blue) vs. parallel run (red) model for three forecast time periods (00-24 h, 12-36 h and 24-48 h) against the North American synoptic observation network.
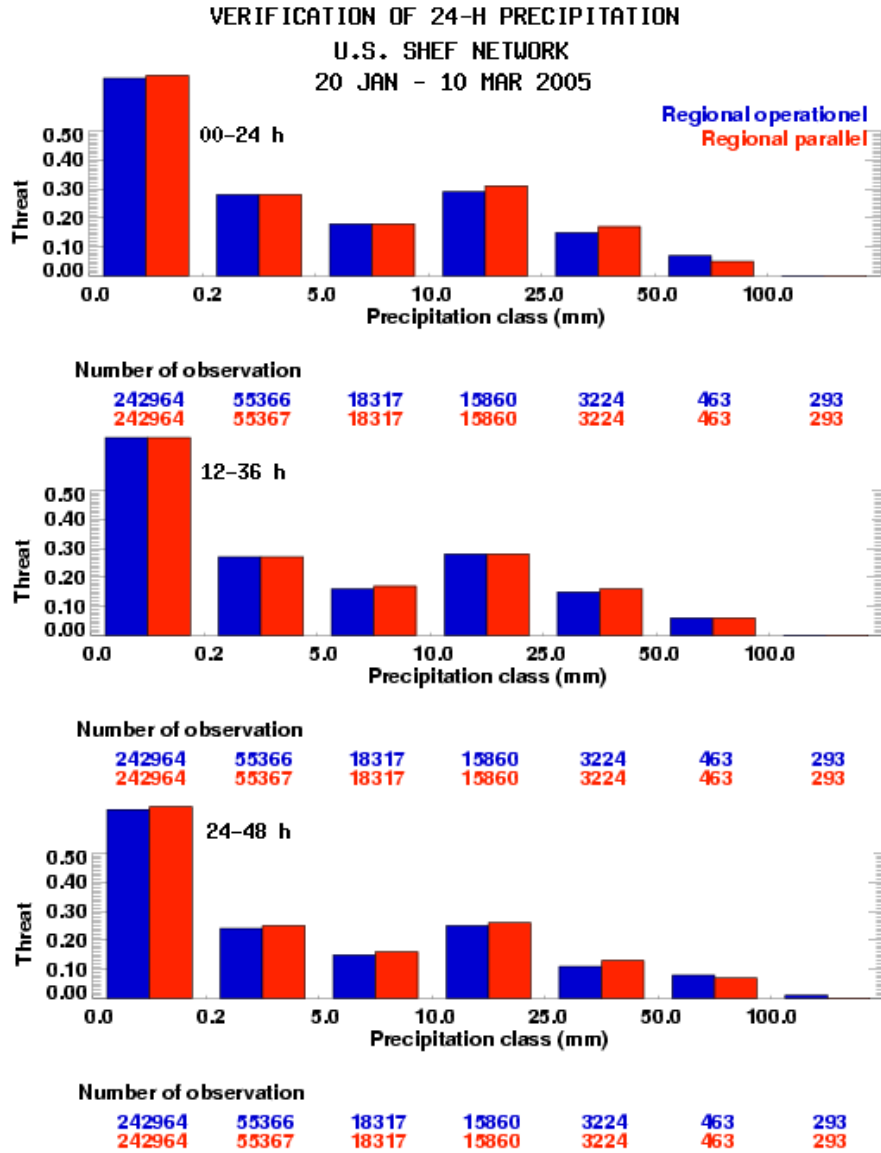
Figure 6: Same as Fig. 5, against the U.S. SHEF network