Environnement Canada
Centre météorologique canadien

Environment Canada
Canadian Meteorological Centre

Improvements to the Global Deterministic Prediction System (GDPS) (from version 2.2.2 to 3.0.0), and related changes to the Regional Deterministic Prediction System (RDPS) (from version 3.0.0 to 3.1.0)

Meteorological Research Branch, National Predictions Development Division, National Predictions Operations Division
at the Canadian Meteorological Centre of Environment Canada

Document co-authors :
S. Antonopoulos[2], B. Archambault[3], A.Beaulne[2], N. Bois[3], M. Buehner[1], A. Giguère[3], J. Marcoux[2], F. Petrucci[2], L. Poulin[3], M. Reszka[2], T. Robinson[3], J. St-James[2], A. Rahill[3], A. Zadra[1]

1 – S&T Meteorological Research Branch
2 – MSC National Predictions Development Division
3 – MSC National Predictions Operations Division

Revisions			
Version	**Date**	**Author**	**Comments**
1.0		Ayrton Zadra	First draft, model portion.
2.0		Mark Buehner	Second draft, assimilartion portion
3.0	2012-12-14	Alain Beaulne	Inclusion de la passe mineure à la section 2.4
4.0	2013-01-11	Jacques Marcoux	Added title page, revision table and table of contents; interchanged assimilation and modeling sections.
5.0	2013-01-15	Ayrton Zadra	Added introduction
6.0 6.1	2013-01-18 2013-01-21	André Giguère Jacques Marcoux	- Miscellaneous corrections throughout the document - Added header and footer, list of figures and cross references in the text..
7.0 7.1 7.2 7.3 7.4	2013-01-21 2013-01-26 2013-01-28 2013-01-28 2013-01-29	André Giguère Tom Robinson S. Antonopoulos Franco Petrucci Mateusz Reszka Judy St-James Allan Rahill	- Addition of an introduction - Addition of the objective verification scores and related figures. - Addition of UMOS and Scribe verification, text and figures. - Addition of sections related to changes in the RDPS. - Addition of subjective evaluation portion, text, tables and figures.
7.5	2013-01-31	Jacques Marcoux	– Updated list of figures and added cross references to new figures. – Created list of figures added cross references to it.
7.6	2013-02-11	Nicole Bois et Lewis Poulin	Minor editorial corrections
7.7	2013-09-23	Lewis Poulin	Editorial adjustments
7.8	2013-12-19	B. Archambault	Put back Draft watermark since section 5.3.3 not complete.
7.9	2014-01-27	André Giguère	Addition of 5.3.3; correction of authors lists (main authors of the document).

Table of contents

List of figures

List of tables

Improvements to the Global Deterministic Prediction System (GDPS) from version 2.2.2 to 3.0.0

Introduction

This document describes the project leading to the implementation of the new Canadian Global Deterministic Prediction System, hereinafter referred to as GPDS version 3.0.0. The new system includes changes to the global forecast model and its data assimilation. The main improvements are related to an increase of horizontal resolution, to the new vertical coordinate and its new vertical discretization, as well as changes to the physics, notably to the orographic blocking and the boundary layer schemes. In this document, the details of these changes will be described and evaluation scores will be shown.

The new system became operational on February 13, 2013 and replaced the previous GDPS version 2.2.4. This accomplishment represents the cumulative effort of many people from the analysis and modelling sections in research and development divisions, and operational sections of EC, all co-located at the Canadian Meteorological Centre (CMC) of Environment Canada in Dorval, Québec, and the list of the co-authors of this document in front page is only a subset of all the people involved in this endeavor.

1. Changes to the data assimilation system

1.1. 4D-Var configuration

1.1.1. Outer loop configuration

Analyses are now generated on the new staggered vertical coordinate. The impact of this change can be seen in **Figure 1** (page 30), mainly as an improvement to the geopotential bias in the 24-hour forecasts. Both of the experiments shown in **Figure 1** use the model with the new staggered vertical coordinate to produce the medium-range forecasts at the same resolution of approximately 33km. Consequently, the impact seen is due only to the use of the model with the new vertical coordinate to produce the background state and within the 4D-Var outer loop within the assimilation cycle.

The outer-loop model uses the same dynamics, physics and resolution as the new forecast model, except for a slightly shorter time-step – i.e. 9 min instead of 12 min, to allow synchronization with the inner-loop time-stepping (see below). This increase in resolution in the outer-loop was shown to produce a small but significant reduction of middle-range forecast errors, as exemplified by the upper-air scores in **Figure 2** page30.

1.1.2. Inner loop configuration

The horizontal resolution of the inner-loop model was increased, from 240x120 to 400x200 grid points. As opposed to the outer-loop model, the non-linear, tangent-linear and adjoint models used in the inner loop are still based on the non-staggered version of GEM. The background

error covariances continue to be defined on non-staggered levels as well, but their (spectral) resolution was extended from T108 to T180. This implies an increase in the effective resolution of the analysis increments, and the impact of this improvement on the forecast errors can be seen in **Figure 3 (**page 30).

The time step of the inner-loop model was reduced from 45 min to 18 min which translates into 21 bins in a 6-h assimilation window, as opposed to 9 bins in the previous system. This augmented number of bins has increased the total amount of assimilated data – notably a 13% increase for AMSU-A and 18% for aircraft data – which also led to significant improvements in the middle-range forecasts (**Figure 4** page 31).

Sensitivity studies on the cost-function convergence revealed that the new system could benefit from an increase in the number of iterations in the inner-loop. Instead of the original number of 55 iterations (30 in the first loop + 25 in the second), the GDPS now uses 65 (35 + 30) which led to significant improvements in the middle-range forecast (**Figure 5** page 31).

1.2. Bias correction coefficients

To fully realize the expected improvements to the stratospheric temperature bias resulting from the new definition of the model vertical coordinate, the bias correction coefficients for the assimilated satellite radiance observations were updated. This was done following an iterative procedure in which the coefficients for all channels (including those that are normally held static) were allowed to evolve dynamically in a 4D-Var cycling experiment. The main adjustment was achieved after this experiment had run over a winter period for approximately 3 weeks. As a result, the stratospheric temperature bias was significantly reduced (relative to radiosonde observations) and the coefficients for AMSU-A ch.11-14 at the end of that period extracted. After running two additional 1-week experiments (1-week summer and 1-week winter), the resulting coefficients were used in all subsequent experiments for the channels that normally use a static set of coefficients (AMSU-A ch.11-14) and also as the initial values for all of the other radiance observations, for which the coefficients are dynamically computed. **Figure 6 (**page 31**)** shows the impact on the global 24-hour forecasts of using these final set of bias correction coefficients relative to using those from the previously operational system. The main impact is a significant reduction in the geopotential height and temperature bias in the stratosphere.

1.3. GPS-RO data

Following research described in Aparicio et al. (2009) and Aparicio and Laroche (2011), the thermodynamic relationship of GPSRO data was updated. Notably, this includes a more accurate compressibility factor for moist air, and a newly developed expression that describes the refractivity of air at given thermodynamic conditions. The main impact is an improved agreement between GPSRO and radiosonde temperatures, which are both used as anchor measurements for the system, and particularly for the radiance bias correction. The smaller bias between technologies has an immediate impact in the overall thermal and geopotential bias of the system. This slowly propagates to moderate improvements in the standard deviations (shown in **Figure 7** page 32), and to the winds.

1.4. New observation data

New observation data were added in order to maintain the continuity and robustness of the data assimilation system:

- A new source of polar Satwinds data was added with the assimilation of winds produced by the AVHRR instrument onboard the various NOAA satellites and of METOP-2. This addition doubles the amount of total polar winds assimilated which were previously only provided by data produced by the MODIS instrument onboard the AQUA and TERRA satellites.
- A greater amount of ATOVS (AMSUA, AMSUB/MHS) data is now assimilated with the addition of RARS. These are data received and processed in various regional centres and are retransmitted to us more rapidly. When the same data is received from both the global and RARS network, the global data is kept. RARS data from NOAA satellites are currently assimilated whereas data from METOP-A are not assimilated at this time because they require a data recalibration to make them consistent with the global METOP-A data, which is not yet implemented at CMC. During a global model « g2 » run, the increase in the data volume assimilated due to the RARS is generally very small for AMSUA (<1%) but is around 20% for AMSUB/MHS. These ratios increase to around 10% and 30% respectively for a forecast « g1 » run, which has a shorter cutoff.
- Radiance data from the water vapour channel of the GOES-13 (E), GOES-15 (W) and MTSAT-2 satellites are now assimilated, completing the network of five geostationary satellites. An update has been done on the RTTOV radiative transfer model coefficients for the MeteoSat-7 (E), MeteoSat-9 (W) and MTSAT-1R satellites.
- Radiance data of the 7 SSM/I-like microwave channels from the SSMIS instrument onboard the DMSP17 and DMSP18 satellites are now added to the already assimilated data from the DMSP16 satellite.

1.5. Surface analysis

No significant changes were made to the surface analysis except for some adjustments to take into account the increased resolution of the model. In particular, the analysis of surface temperature, valid at the level of the model topography, was modified to use the new 25-km topography field.

1.6 Related changes to the Regional Deterministic Prediction System

Several minor modifications were introduced into the Regional Deterministic Prediction System to maintain consistency with the GDPS-3.0.0, thus the RDPS version is increased from RDPS-3.0.0 to. RDPS-3.1.0. The GDPS analysis provided to the RDPS now has to be interpolated vertically from staggered to unstaggered levels. As in the previous version, both the driver and limited area model (LAM) subsystems of the RDPS utilize the bias correction coefficients produced by the GDPS. However, since the GDPS bias correction has now changed (as described in section 1.2 above), there is a marked (generally positive) impact on the bias correction performed in RDPS-3.1.0. The new observation sources added in the GDPS-3.0.0 (as described in section 1.4 above) have also been added in RDPS-3.1.0.

2. Changes to the forecast model

2.1. Increased resolution

The global model's horizontal (non-rotated, latitude-longitude) grid went from 800x600 grid points (i.e. a resolution of approximately 33km at mid-latitudes) to 1024x800 (i.e. approximately 25km). The time step was reduced accordingly, from 15min to 12min. The number of prognostic vertical levels (79) and the approximate position of the model lid (0.1 hPa) remained the same, but the vertical coordinate and the vertical grid discretization were both changed, as briefly described in subsection 2.2 below.

Preliminary tests performed in forecast mode showed that the impact of this resolution increase on forecast errors was minor. It was only in full analysis+forecast cycles that the benefits from the enhanced resolution became noticeable and significant (**Figure 8** page 32).

2.2. Main changes to the dynamics

2.2.1. New vertical coordinate

The model's terrain-following vertical coordinate was changed to ζ, a coordinate of the type *log-hydrostatic pressure*. The precise definition of the new coordinate, its motivation and the associated benefits may be found in Girard and Plante (2012).

2.2.2. New vertical grid

While the original discretization of the old *vertical* coordinate was on a regular (non-staggered) grid, the new ζ coordinates are discretized on a *Charney-Phillips* (staggered) grid. Details on the implementation of the new grid may also be found in Girard and Plante (2012).

The combination of the new coordinate, vertical staggering and other numerical changes (e.g. logarithmic differencing whenever appropriate) led to an overall reduction of noise, as well as an enhanced numerical stability, improved conservation properties, and a significant reduction of forecast errors in the stratosphere (**Figure 9** page 32).

2.3. Main changes to the physics

2.3.1. Enhanced drag coefficient in the orographic blocking scheme

Low-level drag associated with upstream blocking of the flow over unresolved mountains is parametrized in the GEM model according to the formulation proposed by Lott and Miller (1997). This formulation includes a bulk drag coefficient (C_d), an adjustable non-dimentional parameter that controls the overall amplitude of the blocking tendency,

$$\left(\frac{\partial \vec{U}}{\partial t}\right)_{blocking} = -C_d \cdot \mu \cdot \frac{1}{2}\vec{U}\left|\vec{U}\right|,$$

where $\vec{U}$ is the incident wind, and the factor μ depends on subgrid-scale orography properties (its height, slope and direction) and on background flow properties (stability and wind direction). Since its implementation in GEM in 2001 (Zadra et al., 2003), the value of C_d has been set to 1.

Recent studies by Wells et al. (2008) and by Vosper et al. (2009) suggested that instead of a constant, the parameter C_d should vary according to the non-uniform static stability of the flow, and might be further amplified by non-linear effects depending on the incident wind direction.

The implementation of this modulated drag coefficient in the GDPS (**Figure 10** page 33) was shown to have a large impact on the general circulation of the model, especially in the winter hemisphere, leading to a significant reduction of forecast errors with the exception of a slight increase of the tropospheric wind bias (**Figure 11** page 34). This small bias degradation was eventually offset by the introduction of turbulent hysteresis in the vertical diffusion parametrization, described below.

2.3.2. Turbulent hysteresis in the boundary layer scheme

For several years, forecasters have been reporting on an important error mode in the GEM model, associated with warm advection in winter continental cyclones. Early in the integration, the advectively-induced warm nose at the top of the planetary boundary layer (PBL) was eroded and near-surface temperatures rose. The resulting misrepresentation of the temperature profile was particularly problematic when surface conditions were near 0°C and the possibility of freezing rain existed. The problem was shown to arise from over-activity in the PBL scheme, resulting in enhanced vertical diffusion that unrealistically homogenized the temperature, moisture and wind profiles in the lower layers. Simple reductions of the vertical diffusion coefficients in the turbulent kinetic energy (TKE)-based PBL scheme could successfully reduce the over-mixing problem, but yielded important degradations of forecast quality across the remainder of the forecast domain.

The introduction of a physically-based hysteretic effect in the PBL scheme (**Figure 12** page 35) was shown to reduce the amplitude of errors in forecast profiles without adversely affecting guidance in other regions. In this technique, the source terms in the TKE equation are adjusted based on the turbulent regime of the flow, which is itself a function of both the flux Richardson number and the current regime. In both medium range forecasts and assimilation cycles the introduction of turbulent hysteresis was shown to improve guidance quality through a deep layer of the troposphere (

Figure 13 page 36). Turbulent hysteresis has since been successfully implemented in the regional system (RDPS) as well.

2.4. Other changes

Various additional changes in the model dynamics and physics are summarized in **Table 1** page 26. Most changes can be described as minor corrections to existing formulas or minor improvements to the numerical code, which generally rendered the model more robust but had little impact on the forecasts. Two notable exceptions are the changes in the non-orographic gravity wave drag which helped improve the circulation near the model lid, and the optimized

sub-terrain extrapolation algorithm which led to less noisy fields such as the sea-level pressure (PN) over the continents. No changes were brought to the model configuration of the RDPS.

3. Changes to the climatological and geophysical fields

3.1. New set of geophysical fields

Using the GENESIS generator and the operational geophysical databases (i.e. same databases previously used by the 33-km global model), two new sets of geophysical fields were generated, the one set on a 1024x800 grid for the outer-loop model, and the other on a 400x200 grid for the inner-loop.

3.2. New climatology of sea-ice thickness

A new and more realistic monthly global climatology of sea-ice thickness was implemented. This new climatology was recently implemented in the regional system (RDPS) and was obtained by merging three sources of data. The main source is the Global Ocean Reanalysis and Simulations Project (Mercator Ocean Quarterly Newsletter #36, January 2010, pages 15-28) over the years 2002-2008. The second source is CMC's operational ice-ocean analysis and forecasting system for the Gulf of St-Lawrence (Smith et al. 2012). The sea-ice thickness analyses from this system were time averaged over the years 2006-2011 to produce a monthly climatology for the Gulf of St-Lawrence. The third source of data is the old climatology which continues to be used over the freshwater lakes.

With respect to the previous climatology around Canada, the new sea-ice thickness is reduced everywhere except in the northwest portion of the Canadian Arctic archipelago. In most of the Arctic Ocean, in Hudson's Bay and in the Gulf of St-Lawrence, the ice thickness is reduced by 40 cm or more.

4. Objective evaluation of final cycles

Two analysis+forecast cycles of the GDPS 300 were produced, the one over a winter (20-Jan to 31-Mar 2011) and the other over a summer period (15-Jun to 31-Aug 2011) and a full set of verification scores were generated. For evaluation against the operational GDPS 2.2.2, scores are here compared with those of the strato2b-R final cycles, i.e. the control cycles of the global system that became operational in Nov 2011 (note: in all figures, scores from the control cycles are shown in blue, and those from the new system in red).

4.1. Upper-air scores (against radiosondes)

With respect to the previous system, the new GDPS has warmer temperatures in the upper-troposphere and the stratosphere, which translates into a better agreement with radiosonde data and a general reduction of temperature and geopotential biases, over most regions and on both seasons (see **Figure 14** page 37 to **Figure 21** page44). This improvement is mainly due to the new vertical coordinate and grid.

Regarding the standard deviation of forecast errors, a significant reduction was also observed throughout the troposphere, most notably over the extra-tropics in the corresponding winter seasons (**Figure 16** page 39, **Figure 17** page 40 , **Figure 18** page 41 and **Figure 19** page 42). This error reduction is attributed to changes in the near-surface physics of the model – i.e. modifications in the orographic blocking and boundary layer schemes – that were further enhanced in the analysis cycles by the various improvements in the data assimilation system.

Signs of the above improvements are already visible in trial fields, shown as P-O 6h scores in **Figure 14** to **Figure 21**. Most improvements are maintained or amplified in the middle-range forecasts as exemplified by the P-O 120h scores, also shown in **Figure 14** to **Figure 21**.

In particular over North America (**Figure 14** and **Figure 15**), a significant reduction of day-5 errors (bias and standard deviation) for temperature and geopotential is found at practically all levels during the winter season. The maximum improvement in the temperature standard deviation is found in the lower troposphere, below 850 hPa. A bias reduction is also observed for tropospheric winds and dew-point depression. In the summer, the impact of the new system on day 5 is relatively small, except for the generalized warming of the whole column and a minor reduction of the error standard deviation for temperature and zonal wind.

4.2. Surface scores

Verification of surface variables – sea-level pressure (PN), surface pressure (P0), 2m temperature (TT) and 2m dew-point depression (ES) – against surface station observations are shown in **Figure 22** page 45 to **Figure 25** page 48.

Regarding the surface pressure, the systematic reduction of errors is simply explained by the better-resolved topography of the new (higher-resolution) model.

In the winter (Figure 22 page 45 and Figure 23 page 46) the new systems shows a significant reduction of the error standard deviations for temperature, which is most visible over North America and the Northern Hemisphere. In the summer (Figure 24 page 47 and Figure 25 page 48), what stands out is the reduction of dew-point depression biases over North America as well as the Northern Hemisphere. A small but systematic reduction of the dew-point depression error standard deviation can also be seen in all domains and in both seasons.

4.3. Scores against analyses

Verification against analyses generally confirms the improvements measured by the scores against radiosondes.

During the winter season, the new system exhibits systematically smaller RMS errors for temperature, geopotential and winds in the troposphere over the Northern Hemisphere, and for temperature in the Tropics (**Figure 26** page 49, **Figure 27** page 50 and **Figure 28** page 51). An increase in anomaly correlations for tropospheric temperature is also observed in all domains and almost all lead times (**Figure 29** page 52). A similar increase in geopotential anomaly correlations is found in the N. Hemisphere, but less so in other regions (**Figure 30** page 53).

For the summer cycle a reduction of tropospheric RMS errors is still observed, mostly in the Tropics and to a lesser extend in the extra-tropics (**Figure 31** page 54 , **Figure 32** page 55 and **Figure 33** page 56). This time, the positive impact of the new system on the anomaly correlations is more evident in the South Hemisphere and in the Tropics (**Figure 34** page 57 and **Figure 35** page 58).

4.4. Precipitation scores

Precipitation forecasts from the winter and summer cycles were verified against the Synoptic network (at 00 and 12Z over North America) and SHEF network (at 12Z over USA only). Categorical scores for the periods 00-24h, 48-72h and 96-120h are shown in Figure 36 page 60, Figure 37 page 60 and Figure 38 page 61.

In the summer, the impact of the new system on precipitation is mostly neutral; the only systematic change being a slight increase in the biases for classes 0.2 to 5.0mm. In the winter, the signal of the impact varies depending on the network used, on the lead time and on the precipitation class. Given this uncertainty and the questionable quality of the observations in the winter season (especially for large amounts) the evaluation of the winter cycle precipitation was not conclusive.

4.5. Other scores

4.5.1. USTAT scores

Surface scores over North America were also generated using the experimental tool USTAT, which exhibits interesting features such as verification at higher temporal frequency (compared to the traditional surface scores discussed in section 4.2), a height-dependent quality control for stations, and uncertainty estimates obtained with the bootstrapping method. Results are shown in **Figure 39** page 62 and **Figure 40** page 63 for the winter and summer cycles, respectively.

In the winter, the new GDPS shows a significant reduction of the error standard deviation for 2m temperature, 2m dew-point depression and 10m winds at practically all lead times. A systematic improvement of 10m wind bias is also observed up to day 4.

During the summer, the impact on the 10m wind bias is still visible but the amplitude of the gain is less than the one observed in the winter. The impact on the 2m dew-point depression bias depends on the time of the day but was improved in average. The reduction of error standard deviations remain true for both the 2m temperature and dew-point depression up to day 5, but here the impact is neutral for 10m winds.

4.5.2. Tropical cyclones (TCs)

Using the automated TC detection and tracking algorithm proposed by Caron et al. (2010), TCs from the summer cycle forecasts were tracked over the Atlantic, East Pacific and West Pacific basins, and compared to best track data (provided by the US National Hurricane Centre and the Joint Typhoon Warning Centre). During the period covered by the summer cycle, 12 TCs were observed in the Atlantic basin, 8 in the East Pacific and 12 in the West Pacific.

A summary of results is shown in **Figure 41** page 64. Overall, the new GDPS shows a significant increase in TC hit rate and a reduction in the intensity bias. These improvements are attributed to the increased resolution – both the resolution of the forecast model and the resolution of the analyses – and are therefore more evident in the verification over the Atlantic and East Pacific basins, where small-scale TCs occur more frequently than in the West Pacific. No significant changes were found in the average TC position errors or in the number of TC false alarms.

4.6 Scores for the final experiments of the Regional Deterministic Prediction System

Figure 42 page 65 shows time-mean 48-hour forecast verification scores against the RAOB network for the RDPS LAM over North America during winter (Jan. 22 - Mar. 30, 2011), comparing two experiments performed in development mode: one very similar to the operational system and one which received its bias correction coefficients and (vertically interpolated) GDPS analyses from the GDPS-3.0.0 final cycle. There is a very significant improvement in the geopotential bias above approximately 600 hPa, and the separation between the curves increases markedly with height. The temperature bias is also improved at the same levels, although with a smaller separation and fewer levels where the improvement is statistically significant. A slightly positive (although not significant) impact is also visible in the wind speed standard deviation at jet levels. It should be noted that these development experiments comprised only 40 cases.

In **Figure 43** page 66 we present the same scores as in **Figure 42**, but comparing the operational (RDPS-3.0.0) and parallel (RDPS-3.1.0) runs during the period Nov. 14, 2012 - Jan. 14, 2013 (140 cases). Unlike the experiment in **Figure 42**, the parallel run also included the additional observation types during the second half of this period. The geopotential bias is improved at almost all levels and the geopotential standard deviation also shows a small but statistically significant improvement in the troposphere. Standard deviations of the temperature, zonal wind and wind modulus at jet levels are also improved, and the curve separation is more apparent than in **Figure 42**. Finally, some change (both positive and negative) in the wind biases appears at upper levels. In general, the impact of the changes described in this document appears to be qualitatively similar in both the GDPS 3.0.0 and RDPS-3.1.0 although in the former case the improvements are more robust (i.e. larger curve separation and more levels where statistical significance is achieved).

5. Objective evaluation of the parallel runs

The new GDPS-3.0.0 was installed in a parallel run during the third week of October 2012, beginning with the assimilation runs (g2) on October 23rd, and then with the forecast runs (g1) with the 1200 UTC cycle on October 26th. The parallel run period ended on January 27th, 2013 with the 0000 UTC cycle.

5.1. GDPS-3.0.0

5.1.1. Upper air fields

Objective verification scores against radiosonde observations and against model analyses are calculated for geopotential height, temperatures and winds at all model pressure levels, over a set of strictly defined domains based on the international verification data exchange program of the World Meteorological Organization / Commission for Basic Systems (WMO / CBS). Scores for the GDPS-3.0.0 parallel run were compiled over the period from October 29th, 2012 to January 21st, 2013.

Improvement in verification scores during the parallel run is very significant for all variables, levels and regions. Reduction in geopotential height and temperature forecast errors is particularly strong in the upper levels of the atmosphere (above 150 hPa), but is also impressive within the troposphere. Statistically significant improvement is noted at all levels out to day 9 to 10 over the Northern Hemisphere (20-90N) and out to day 6 to 8 over North America (**Figure 44** page 67 to **Figure 47** page 68). Improvement in the wind forecasts is relatively smaller, but still statistically significant at most levels out to day 8 to 10 over the Northern Hemisphere (**Figure 48** page 69) and day 6 to 8 over North America.

Temperature forecasts over the Tropics (20N-20S) are improved at almost all levels out to day 10 (**Figure 49** page 69), though deteriorated at 200 and 250 hPa. Forecast winds over the Tropics are strongly improved in the upper atmosphere, but little changed at and below the tropopause (**Figure 50** page 70). Over the Southern Hemisphere (20-90S), the pattern is similar to the North, with very strong improvement for the geopotential heights at the upper levels but with smaller, though still generally significant, improvement throughout the troposphere (**Figure 51** page 70).

Anomaly correlations for both geopotential heights and temperatures are higher at all forecast times and levels in both the northern and southern extra-tropics (**Figure 52** page 71).

5.1.2. Quantitative Precipitation Forecasts (QPF)

Model 24-hour QPF are scored against both the North American synoptic network and the high density U.S. SHEF network. For the GDPS-3.0.0 parallel run, scores were calculated over the period November 18th, 2012 to January 17th, 2013.

For the short-term forecasts (day 1 and 2), equitable threat scores calculated against the synoptic observations (**Figure 53** page 71) show a slight improvement over the operational model, except in the middle categories (between 5 and 25 mm), where the signal is neutral. Against the SHEF observations (**Figure 54** page 72), improvement is noted in all categories from 5 mm and up, with a stronger improvement at the higher levels than scores against the synoptic

observations. In both cases the parallel run has a larger bias in the higher amounts, indicating a somewhat larger over-forecast of precipitation in those categories.

At the longer term, out to day 5 (**Figure 55** page 73), the signal is more neutral, except that the slightly higher bias in the higher precipitation categories remains.

5.2. RDPS-3.1.0

5.2.1. Upper air fields

The RDPS-3.1.0 parallel run was scored over the period November 14th, 2012 to January 21st, 2013 (136 cases).

As with the GDPS, geopotential height forecasts for the RDPS-3.1.0 parallel run are strongly improved in the upper atmosphere (**Figure 56** page 73). Throughout the troposphere, there is a small but statistically significant improvement in scores at all levels and forecast times except for day one forecasts (out to 24 hours) at 850 hPa and below, where the signal is neutral.

Temperature forecasts (**Figure 57** page 74) are improved at all levels except at 1000 hPa, though the deterioration noted at that level is not statistically significant. Forecast winds (**Figure 58** page 75) are improved at all levels and forecast times, though this is most evident between 700 and 250 hPa.

5.2.2. Quantitative Precipitation Forecasts (QPF)

As for the GDPS-3.0.0, 24-hour QPF scores for the RDPS-3.1.0 parallel run are calculated against the synoptic and SHEF observation networks. Scores were calculated over the period November 7th, 2012 to January 17th, 2013

Differences between the parallel and operational models are generally neutral, except for a slight improvement in the higher precipitation amounts for scores against the SHEF network (**Figure 59** page 76).

5.3. Performance of dependent systems

5.3.1. UMOS global

The training of the UMOS system with the new GDPS-3.0.0 has been carried out using the hindcasts (final series of forecasts) for the periods of January 20th to March 31st 2011 for the winter season and June 15th to August 25th 2011 for the summer season (runs of 0000 and 1200 UTC). In addition, the training for the winter season has advanced during the parallel run for all runs.

Verifications performed in both the summer and winter seasons have shown that the accumulated cases ensure a stable and robust production of regression equations (see Figure 60 page 77). There exist a few individual cases where the performance of the new system varies, with some improvement or deterioration over the old model. However, in terms of bias, root mean square error and variance reduction, the averaged overall performances of the UMOS forecasts are maintained and are equivalent to the previous system.

5.3.2. Scribe matrices

The GDPS-3.0.0 parallel run was verified with the Scribe verification system for the period extending from November 2nd 2012 to January 7, 2013. The Scribe matrices were post-processed, for the temperature, Probability of precipitation (POP) and cloud forecasts. UMOS was used to derive the temperatures (using the retrained equations as explained in section 5.3 above) while the same low resolution Perfect Prog system as in the operational GDPS was used for the POP and cloud forecasts. Finally, the winds were extracted directly from the model, as is done in the operational GDPS. The precipitation amounts and types were not verified as these weather elements are not mentioned often in Scribe in days 3 to 5, and also because the amounts are verified directly for the model.

The results of the verifications are greatly in favor of the parallel model. For the maximum and minimum temperature forecasts, both the bias and RMSE (**Figure 61** page 78) scores are improved significantly, especially from days 2 to 6. For the POPs, significant improvements can also be seen beyond day 2 or 3 in the Brier scores, and more markedly in the CSI (Critical Success Index) scores for POPs of 50% or less (**Figure 62** page 78). For POPs greater than 50%, it is seen to remain neutral. This means that it is mainly in situations with low probabilities of precipitation that the forecasts are now better with the parallel system. For the cloud forecasts, the verification doesn't show any improvement in the parallel run.

For the wind forecasts, there is a slight deterioration in the bias scores for the 40 to 60 km/h category (**Figure 63** page 78). This is because the parallel model tends to slightly increase higher winds compared with the operational at the vertical level where Scribe extracts wind gusts from the model. Since Scribe will often use these gusts to specify a range in the winds, and because it already increases slightly too much the stronger winds coming from the model, it therefore overcompensates a little more the winds coming from the model in parallel. The frequency of these forecast winds (in the 40-60 km/h category) however still remains close in absolute terms (5.7%) compared to their observed frequency (4.1%). Lastly, in the other categories, 0-20 km/h and 20-40 km/h, there isn't much difference between the operational and the parallel. Both models forecast too often the 0-20 km/h winds and forecast not often enough 20-40 km/h winds. In both cases, Scribe corrects the model winds very well and the frequency of its forecasts for these two classes of winds corresponds to their frequency of occurrence.

5.3.3. Other dependent systems

Regional Deterministic Wave Prediction System (RDWPS)

Subjective verifications were performed by the regional forecast offices who use the Wave Model (WAM) output on a regular basis. The wave forecasts obtained using the parallel GDPS 3.0.0 (25km) derived winds were compared against those produced by the the GDPS 2.2.2 (33km) derived winds. No negative comments were brought forward from these offices.

The main lead for the WAM project (Jamie McLean) also did his own subjective analysis during the parallel run, and from his observations, he did not notice any significant changes in the WAM outputs with the new GDPS 3.0.0 winds either. These conclusions are what one would expect, given none of the WAM binaries were changed, just the model winds (due to the increased model resolution) driving the wind-wave portion of the WAM output.

Canadian Seasonal and Inter-Seasonal prediction System (CanSIPS)

A parallel run of the long range forecast system CanSIPS using the analysis from the GDPS 3.0.0 (25 km) was run from November 1st 2012 to January 22, 2013. No significant difference was observed in the forecasts issued from this parallel run. Statistical tests between these forecasts and operational forecasts did not showed statistically significant differences.

6. Subjective evaluation of the parallel runs

A subjective evaluation was done twice daily by meteorologists of the Analysis and Prognosis operational section (hereafter "A&P") at CMC over a significant portion of the course of the parallel run of the GDPS-3.0.0 (from November 5th to January 6th). Mass fields (geopotential heights at 500 hPa – GZ500 - and mean sea level pressure - MSLP) and quantitative precipitation forecasts (QPF) were evaluated by visually comparing sets of images from the two different models (as in **Figure 64** and **Figure 65)**, at lead times of 48, 72, 96, and 120 hours. The meteorologists had then to indicate whether any of the models had performed better over specific regions (**Figure 67**), or if both had performed equally well or poorly. The results are displayed in **Table 2** page 26 and **Table 3** page 28, and are additionally shown graphically by region (all mass fields) in **Figure 68** and by lead time (MSLP, GZ500, and QPF) in **Figure 69** page 82**, Figure 70** page 83 and Figure 71 page 84.

To have a measure of the impact of the new GDPS on its dependent RDPS, a similar subjective evaluation was done during the course of the RDPS-3.1.0 parallel run (from November 22nd to January 7th), but reduced in scope as only verification images of forecasts with a lead time of 48 hours were examined. Results for the RDPS evaluation are summarized in **Table 4**.

6.1. GDPS

In the case of the Global Deterministic Prediction System, when a difference in mass field forecasts was found between the two systems (more than 40% of the time), the parallel system was generally preferred over the operational one by a ratio of about 2:1, slightly higher for GZ500 than SLP. When a difference in the quantitative precipitation forecasts (QPF) was found (46,4% of the time) the ratio of preference was only slightly in favor of the GDPS-3.0.0, at 5:4.

On a region per region basis differences were more frequently found along the Pacific and Atlantic coast (50.4 and 46.2% of the time) than over the continent or the Arctic (38.1 and 30.9% of the time, respectively). However, when the systems differed, preference for the new system was stronger over the Arctic area (by a ratio of 3:1) than over the Pacific (by a ratio of 3:2) or elsewhere (by a ratio of about 2:1).

Looking at it by lead time, not surprisingly the frequency of preference for one system or another increases significantly with lead time (forecasts being more likely to diverge), but the ratio of preference in favor of the new system remains fairly constant at 2:1 or better for any lead time, peaking at 72h at about 3:1.

6.1.1. Comments and observations from A&P meteorologists

During the parallel run evaluation, meteorologists from the A&P section provided an abundance of comments on the comparative performance of the two systems. A thorough review

of those comments and of the evaluation images was done at the end the evaluation period, to detect any pattern in the new system with respect to the one it replaces.

When cyclogenesis occurs to the lee of the Canadian Rockies, CMC guidance has a reputation of displaying a negative bias with heights and surface pressure over the high plains. There was no typical strong zonal circulation ("Alberta clipper") during this time period but there were two cases where the upper circulation was light to moderate from the west and southwest, with cyclogenesis along the Foothills. The developing lows verified better and further south in the new system, which is an improvement, and also the inverted surface troughs north of these systems verified better. One example of this is shown in **Figure 72** page 85.

It was also mentioned that the new system verified significantly better with short waves evolving in a closed circulation over the Arctic region. This may explain the high ratio of preference in the subjective evaluation for this area. It was noted as well that the new GDPS-3.0.0 could be too slow or too deep with upper trough nearing the Pacific coast. On the other hand, the previous system often appeared to be too slow or too deep with systems off the Atlantic coast. Another noted improvement in the GDPS-3.0.0 is a faster or earlier correction of persistent errors in the longer term forecast. An example of this is shown in **Figure 73** page 86.

One major improvement in the GDPS-3.0.0 is the introduction of a physically-based hysteretic effect in the boundary layer scheme (see sub-section 2.3.2 and Figure 12), identical to that implemented in the RDPS-3.0.0 in October 2012. A&P meteorologists reported on several occasion the benefits of this feature, as shown in **Figure 74** page 87. It is noted however that its effects are not as marked in the GDPS-3.0.0 as they were seen in the RDPS because of the coarser horizontal resolution.

Fortunately, the parallel run was installed just in time to evaluate the behavior of the GDPS-3.0.0 with the major weather event that was hurricane Sandy. As shown as an example in **Figure 75** page 87 with the 72 hour forecast valid October 30, 0000 UTC, the GDPS-3.0.0 gave a closer estimate of the depth and trajectory of the storm for most forecast runs than the GDPS operational version at the time.

On the negative side, a possible problem was noted with the forecasts of "cutoff" upper lows over the Southwestern U.S. and adjacent Eastern Pacific waters. Many of them were forecast too slow or developed too fast in the parallel as opposed to the operational runs. One such case is illustrated in **Figure 76** page 88. A similar case was also seen over the Great Lakes region in early November. On the other hand the GDPS-3.0.0 seemed to fare better with similar systems off the Atlantic coast.

With respect to QPF, the GDPS was found to deliver more precipitation over the mountainous terrain of the West coast, too much according to some, but with better positioned maxima, a direct result of the increase in horizontal resolution from 33 to 25 km.

6.2. RDPS

In the case of the RDPS-3.1.0 parallel run, when a difference in mass field forecasts was found between the two systems (about 40% of the time) at 48 hours lead time, the system in parallel was generally preferred over the operational one by a ratio of about 3:2. This ratio was similar when considering separately the fields GZ500 or MSLP. When a difference in the quantitative precipitation forecasts (QPF) was found (24% of the time) the ratio of preference was also in favor of the RDPS-3.1.0 at about 3:2, as with the mass fields.

On a region per region basis, differences were found significantly more frequently along the Pacific coast (58 % of the time) than over the continent or the Atlantic coast (36 and 40 % of

the time respectively), and much less often over the Arctic (29 %). When a preference was expressed, the two systems were chosen about equally often, except over the Atlantic coast region, where the preference was markedly in favor of the new system, with a surprisingly high ratio of 5:2.

With this parallel run as well, comments were made by the A&P meteorologists. They found that, while there is a problem with meteorological entities being too deep in both operational and parallel prediction systems, especially upper troughs over the Eastern Pacific and North America systems (there were more than twice as many mentions overall of troughs/lows being too deep as there were not deep enough), the new RDPS-3.1.0 fared better during the parallel run, slightly fewer problems being reported for it. The new system was also reported to better forecast upper ridges over the Pacific coast and adjacent waters, and surface ridges over the continent. The forecast of surface lows was also slightly improved. With respect to QPF, several A&P meteorologists found it to be somewhat excessive in both systems.

An interesting point to make is that a good performance of the GDPS-3.0.0 at the same 48 hour lead time would not necessarily ensure that the new RDPS-3.1.0 would perform better than its operational counterpart (even though the GDPS-3.0.0 provided the analysis for the RDPS-3.1.0 guess fields).

7. Products availability

Although this major upgrade of the GDPS does not come with new analysis or forecast products, the increase in resolution for both the assimilation and the model portions of the new system are more demanding in computer resources. Therefore, to minimize the impact on the time of availability of the forecast products, the data acquisition cut-off time has been shortened to 2 hours 30 minutes for the first minimization of the inner loop, and to 3 hours 10 minutes for the second minimization, the latter being more critical for the quality of the forecasts. These new cut-off times are fairly close to those that were in use prior to the migration to the current supercomputer upgrade (IBM P7 in May 2012). These changes in data acquisition cut-off time are expected and are done in a balance between providing forecast products as soon as possible, taking advantage of additional computational resources, and implementing improvements to NWP systems which often require more computing capacities.

In short, the reduction of cut-off time ensures in this case that most post-processed products (SCRIBE outputs, bulletins, charts, images) will be available at roughly the same time than prior to implementation of the new system. Note however, that GRIB data files on the MSC Datamart will be available 10 to 15 minutes later than prior to this implementation. Ongoing code optimization, combined with further upgrades in computing capacity will eventually allow us to increase again the data acquisition cut-off times .

Acknowledgements

Many thanks to all collaborators of GDPS3.0.0 R&D team: Michel Roch, Martin Charron, Claude Girard, André Plante, Ayrton Zadra, Stéphane Bélair, Anne-Marie Leduc, Jocelyn Mailhot, Ron McTaggart-Cowan, Paul Vaillancourt, Mark Buehner, Cécilien Charette, Stéphane Laroche, Josée Morneau, Monique Tanguay, Nicloas Wagneur, Josep Aparicio, Louis Garand, Sylvain Heilliette, Stephen MacPherson, Bruce Brasnett, Jean-Francois Lemieux, Greg Smith, Stéphane Chamberland, Michel Desgagné, Bin He, Erwig Lapalme, Vivian Lee and Lubos Spacek.

References

Aparicio, J. M., G. Deblonde, L. Garand, and S. Laroche (2009), *Signature of the atmospheric compressibility factor in COSMIC, CHAMP, and GRACE radio occultation data*, J. Geophys. Res., 114, D16114.

Aparicio, J. M. and S. Laroche (2011), *An evaluation of the expression of the atmospheric refractivity for GPS signals*, J. Geophys. Res., 116, D11104.

Girard C. and A. Plante, 2013: *GEM4.2 – A non-hydrostatic atmospheric model in terrain-following vertical coordinate of the log-hydrostatic-pressure type vertically discretized on a Charney-Phillips grid*. RPN internal report.

Lott, F. and M. J. Miller, 1997: *A new subgrid-scale orographic drag parametrization: Its formulation and testing*. Q. J. R. Meteorol. Soc., 123, 101-127.

Smith, G.C., F. Roy and B. Brasnett : *Evaluation of an Operational Ice-Ocean Analysis and Forecasting System for the Gulf of St. Lawrence*. QJRMS, accepted.

Vosper, S.B., H. Wells and A. R. Brown, 2009: *Accounting for non-uniform static stability in orographic drag parametrization.* J. R. Meteorol. Soc., 135, 815-822.

Wells, H., S. B. Vosper, A. N. Ross, A. R. Brown and S. Webster, 2008: *Wind direction effects on orographic drag*. Q. J. R. Meteorol. Soc., 134, 689-701.

Zadra, A., M. Roch, S. Laroche and M. Charron, 2003: *The subgrid-scale orographic blocking parametrization of the GEM Model*. Atmosphere and Ocean, 41, 155-170.

Table 1: List of minor model changes

item	description	impact
sub-surface extrapolation	- Liebmann calculation was parallelized - number of levels and iterations were optimized to insure convergence in the sub-surface extrapolation of the sea-surface pressure (PN), temperature (TT) and geopotential (GZ) fields	- little impact on scores, but much better (less noisy) fields such as PN
semi-Lagrangian calculation	- smoothing of winds ($u, v, \dot{\varsigma}$) used in the calculation of trajectories	- elimination of noise in temperature field near the poles
height field used by the physics	- heights are now taken directly from the dynamics	- little impact on scores - no need to recalculate heights - avoid errors in model configurations where vertical resolution is enhanced near the surface
radiation transfer scheme	- isothermal (instead of linear) extrapolation of temperature above the model lid - new minimum value (1.5e-6) of the specific humidity seen by the radiation code - new value of parameter epsilon (1e-10 instead of 1e-15) used to avoid numerical singularities in the short-wave calculations	- little impact on scores - solves known cases of model crash
orbital parameters	- corrected orbital declination	- little impact on scores
cloud scheme	- new threshold (5hPa instead of 1hPa) in the calculation of clouds and convective cloud tendencies	- little impact on scores
type of precipitation	- adopt same scheme (bourge3D) used by the regional model	- little impact on scores
boundary layer scheme	- use implicit (instead of explict) treatment of surface (sensible and latent) heat fluxes - corrected initialization of friction velocity (u*) - correct missing beta (Prandl) factor in stability functions - generalize definition of parameter epsilon (taking floating point precision into account) for neutral case - correct height of lowest model level	- implicit treatment led to a slight reduction of heat fluxes over the oceans - otherwise, little impact on scores
surface fluxes, parameters and diagnostics	- code clean up - quasi-neutral branch eliminated - correct calculation of stability function gradients - maximum of 0.2m imposed on thermal roughness length for glaciers	- little impact on scores
orographic gravity-wave drag	- generalized near-surface averages, to avoid dependence on the vertical discretization	- little impact on scores
non-orographic gravity wave drag	- generalized calculation of gravity-wave emission level - vertical smoothing of tendencies near the model top - correct centered calculation of gradients	- little impact on scores in the troposphere - improved representation of the circulation near the model lid

Table 2: Summary of the subjective evaluation of the GDPS-3.0.0 parallel run by CMC's "Analysis and prognosis" (A&P) operational section. **123 evaluations were done over 126 possible, from November 5th to January 6th, 0000 and 1200 UTC runs. Letter "o" and "p" indicate a preference (O, P, a strong preference) for one model. Letter "e" indicates if they are alike and good (E, very good), and X, if they are different and that neither is good.**

GDPS-3.0.0 parallel run – Passe parallèle SGPD-3.0.0				
Toutes régions , toutes échéances / All regions, all lead times	**Aucune préfér./ No preference**	**Préférence OP Preference OP**	**Préférence PAR / Preference PAR**	**Rapport de / Pref. ratio:**
Champs de masse / Mass fields				
GZ500 + PNM / SLP	e : 26,7% E : 29,6% X : 2,3 % 58,6%	O : 0,9% o : 12,4% 13,3%	P : 2,1 % p : 26,0% 28,1%	**P > O** **> 2:1** **(28,1 / 13,3)**
GZ500 seulement / only,	e : 27,8% E : 30,3% X : 2,0% 60,1%	O : 0,9% o : 11,3% 12,2%	P : 2,4% p : 25,3% 27,7%	**P > O** **> 2:1** **(27,7 / 12,2)**
PNM / SLP seulement / only,	e : 25,7% E : 28,8% X : 2,6% 57,1%	O : 0,9% o : 13,5% 14,4%	P : 1,8% p : 26,7% 28,5%	**P > O** **≈ 2:1** **(28,5 / 14,4)**
QPP / QPF				
	e : 15,7% E : 35,5% X : 2,2% 53,4%	O : 1,1% o : 19,7% 20,8%	P : 1,2% p : 24,6% 25,8%	**P ≈ O** **> 5:4** **(25,8 / 20,8)**

Table 3: Result of the subjective evaluation of the GDPS-3.0.0 parallel run by A&P, per region and per forecast lead time. Letter coding similar as in Table 2.

GDPS-3.0.0 parallel run – Passe parallèle SGPD-3.0.0				
GZ500 + PNM / SLP	**Aucune préfér./ No preference**	**Préférence OP Preference OP**	**Préférence PAR / Preference PAR**	**Rapport de / Pref. ratio:**
Région / Region				
Pacifique / Pacific	e : 25,7% E : 21,2% X : 2,7% 49,6%	O : 1,9% o : 17,0% 18,9%	P : 1.7% p : 29.7% 31,4%	**P > O** **> 3:2** **(31,4 / 18,9)**
Amérique du Nord / North America	e : 27,9% E : 32,5% X : 1,5% 61,9%	O : 0,7% o : 12,6% 13,3%	P : 2,3% p : 22,5% 24,8%	**P > O** **≈ 2:1** **(24,8 / 13,3)**
Atlantique / Atlantic	e : 23,7% E : 26,7% X : 3,4% 53,8%	O : 0,6% o : 12,3% 12,9%	P : 2,8% p : 30,5% 33,3%	**P > O** **> 2:1** **(33,3 / 12,9)**
Arctique / Arctic	e : 29,7% E : 37,8% X : 1,6% 69,1%	O : 0,3% o : 7,8% 8,1%	P : 1,6% p : 21,1% 22,7%	**P > O** **≈ 3:1** **(22,7 / 8,1)**
Échéance de prévision Forecast lead time				
48 h	e : 16,8% E : 48,6% X : 0,1% 65,5%	O : 0,2% o : 7,20% 7,4%	P : 0,1% p : 17,1% 17,2%	**P > O** **> 2:1** **(17,2 / 7,4)**
72 h	e : 25,6% E : 32,4% X : 1,4% 59,4%	O : 0,8% o : 11,1% 11,9%	P : 0,9% p : 27,8% 28,7%	**P > O** **< 3:1** **(28,7 / 11,9)**
96 h	e : 32,9% E : 17,77% X : 3,0% 53,6%	O : 0,7% o : 14,3% 15,0%	P : 2,7% p : 28,7% 31,4%	**P > O** **> 2:1** **(31,4 / 15,0)**
120 h	e : 31,8% E : 9,1% X : 4,8% 45,7%	O : 1,9% o : 17,1% 19,0%	P : 4,8% p : 30,6% 35,4%	**P > O** **< 2:1** **(35,4 / 19,1)**

Table 4: Summary of the subjective evaluation of the RDPS-3.1.0 parallel run by A&P, **107 evaluations were done over 108 possible from November 22nd to January 7th, 0000 and 1200 UTC runs. Letter coding similar as in Table 2.**

RDPS-3.1.0 parallel run – Passe parallèle RGPD-310				
Champs de masse / Mass fields : GZ500 + PNM / SLP :				
Toutes régions / All regions:	e : 25,3% E : 33,2% X : 0,9 % 59,4%	O : 1,1% o : 16,6% 17,7%	P : 0,9 % p : 22,1% 23,0%	**P > O** **< 3:2** **(23,0/ 17,7)**
GZ500 seulement / only, Toutes rég. / All reg.	e : 24,9% E : 32,2% X : 1,0% 58,1%	O : 1,5% o : 16,6% 18,1%	P : 1,0% p : 22,9% 23,9%	**P > O** **< 3:2** **(23,9/ 18,1)**
PNM / SLP seulement / only, Toutes rég. / All reg.	e : 25,6% E : 34,2% X : 0,8% 60,6%	O : 0,8% o : 16,6% 17,4%	P : 0,8% p : 21,4% 22,2%	**P > O** **< 3:2** **(22,4/ 17,4)**
Par Région / Per Region	**Aucune préfér./ No preference**	**Préférence OP Preference OP**	**Préférence PAR / Preference PAR**	**Rapport de / Pref. ratio:**
Pacifique / Pacific	e : 19,2% E : 22,2% X : 1,0% 42,4%	O : 2,5% o : 25,8% 28,3%	P : 1,5% p : 27,8% 29,3%	**P ≈ O** **≈ 1:1** **(29,3/ 28,3)**
Amérique du Nord / North America	e : 30,8% E : 27,8% X : 1,0% 59,6%	O : 0,5% o : 18,2% 18,7%	P : 1,5% p : 20,2% 21,7%	**P ≈ O** **>1:1** **(21,7/ 18,7)**
Atlantique / Atlantic	e : 24,8% E : 39,1% X : 0,0% 63,9%	O : 1,0% o : 9,4% 10,4%	P : 0,5% p : 25,2% 25,7%	**P > O** **> 5:2** **(25,7/ 10,4)**
Arctique / Arctic	e : 26,3% E : 43,4% X : 1,5% 71,2%	O : 0,5% o : 13,1% 13,6%	P : 0,0% p : 15,2% 15,2%	**P ≈ O** **> 1:1** **(15,2/ 13,6)**
QPP / QPF				
Toutes régions / All regions:	e : 26,1% E : 49,5% X : 0,3% 75,9%	O : 0,7% o : 8,7% 9,4%	P : 0,0% p : 14,7% 14,7%	**P > O** **≈ 3:2** **(14,7/ 9,4)**

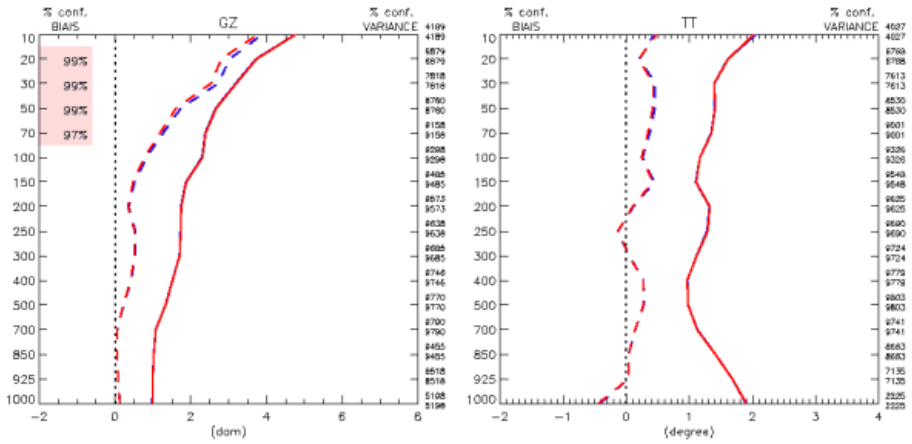


Figure 1: impact of using the model with the new vertical coordinate (staggering) in the outer-loop of the 4DVar data assimilation cycle

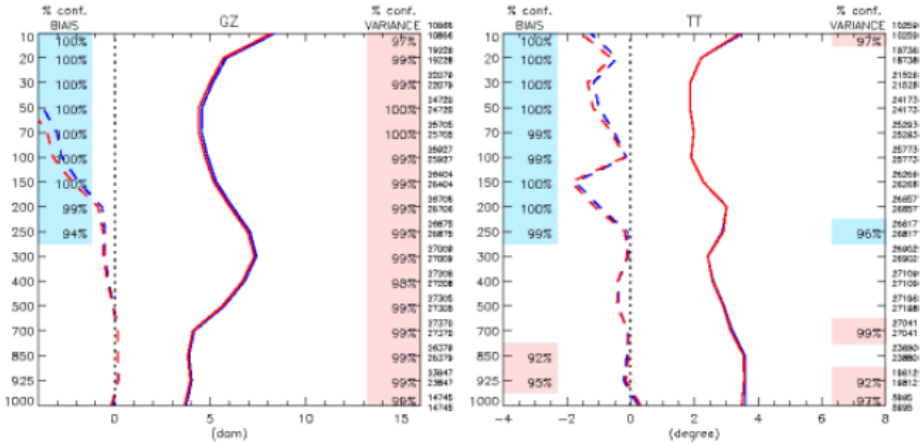


Figure 2: impact of increased resolution in outer loop of the 4DVar.

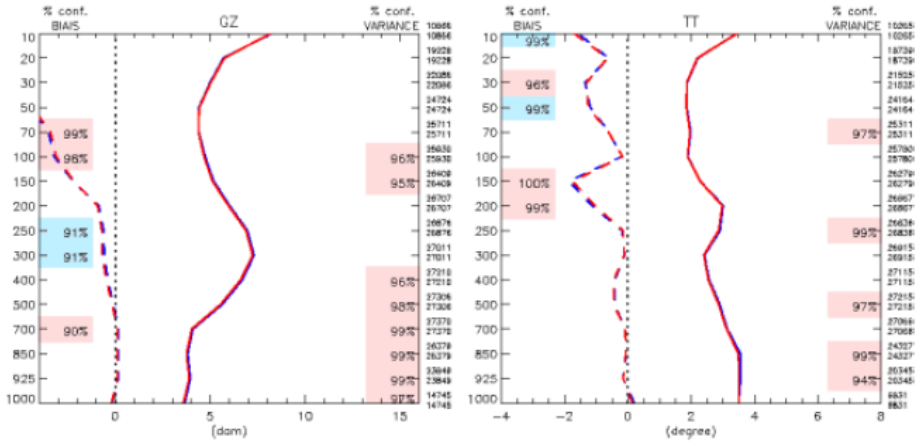


Figure 3: impact of increased resolution in inner loop of the 4DVar.

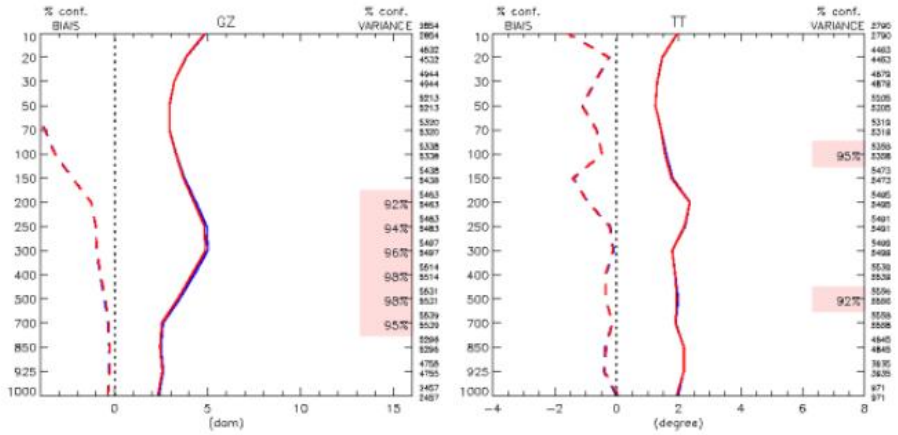


Figure 4: impact of increased number of temporal bins.

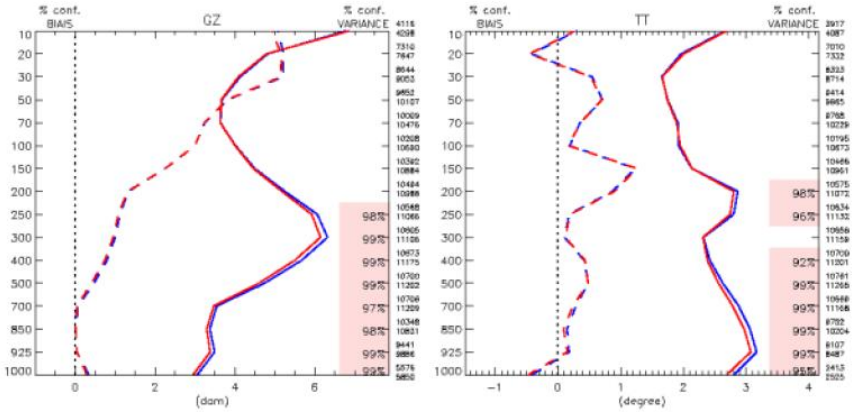


Figure 5: impact of increased number of iterations in the inner loop.

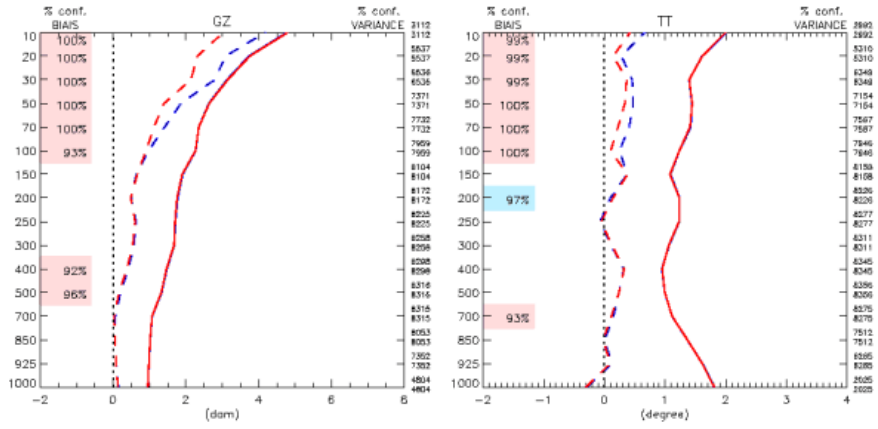


Figure 6: impact of using updated bias correction coefficients for all assimilate satellite radiance observations.

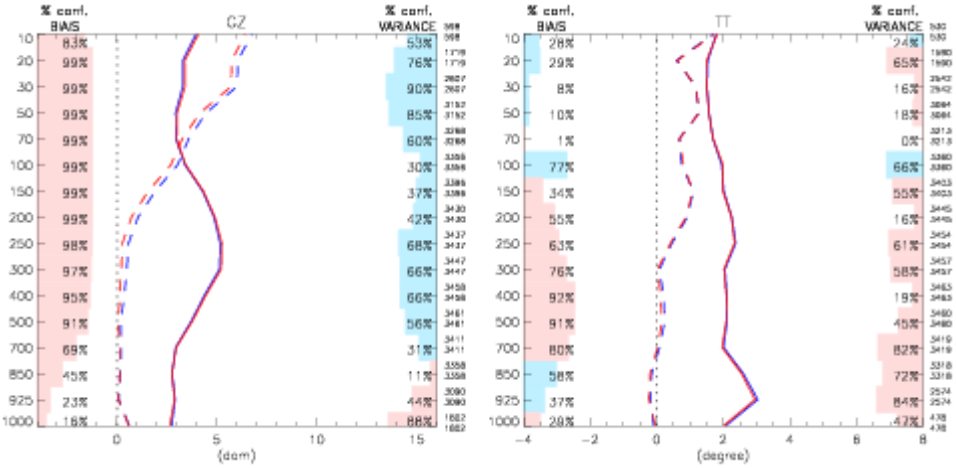


Figure 7: Impact of improved assimilation of GPS-RO observations (here for the southern extra-tropics)

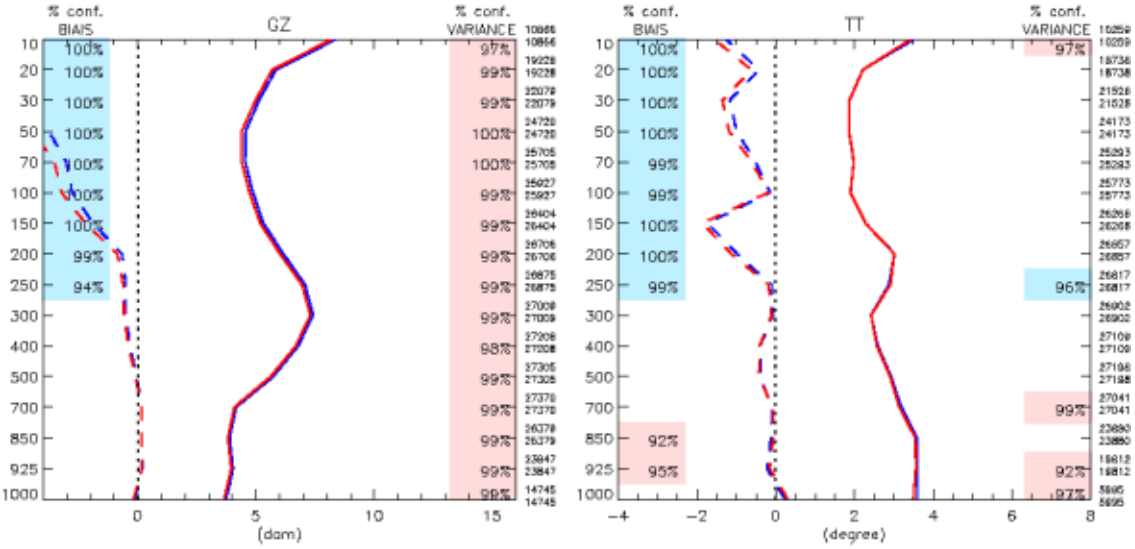


Figure 8: Impact of the increase in resolution (from 33km system in blue, to the 25km system in red) in analysis+forecast mode, during the winter season over the N. Hemisphere: 120-h forecast errors – bias (dashed) and standard deviation (solid) – against radiosondes, for geopotential (left) and temperature (right).

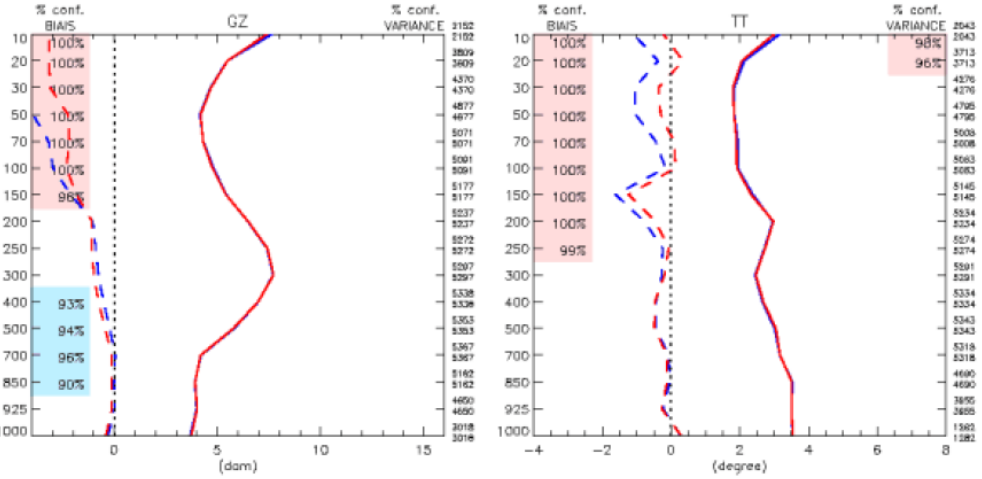


Figure 9: Same as Figure 8, impact of new vertical coordinate + new vertical grid.

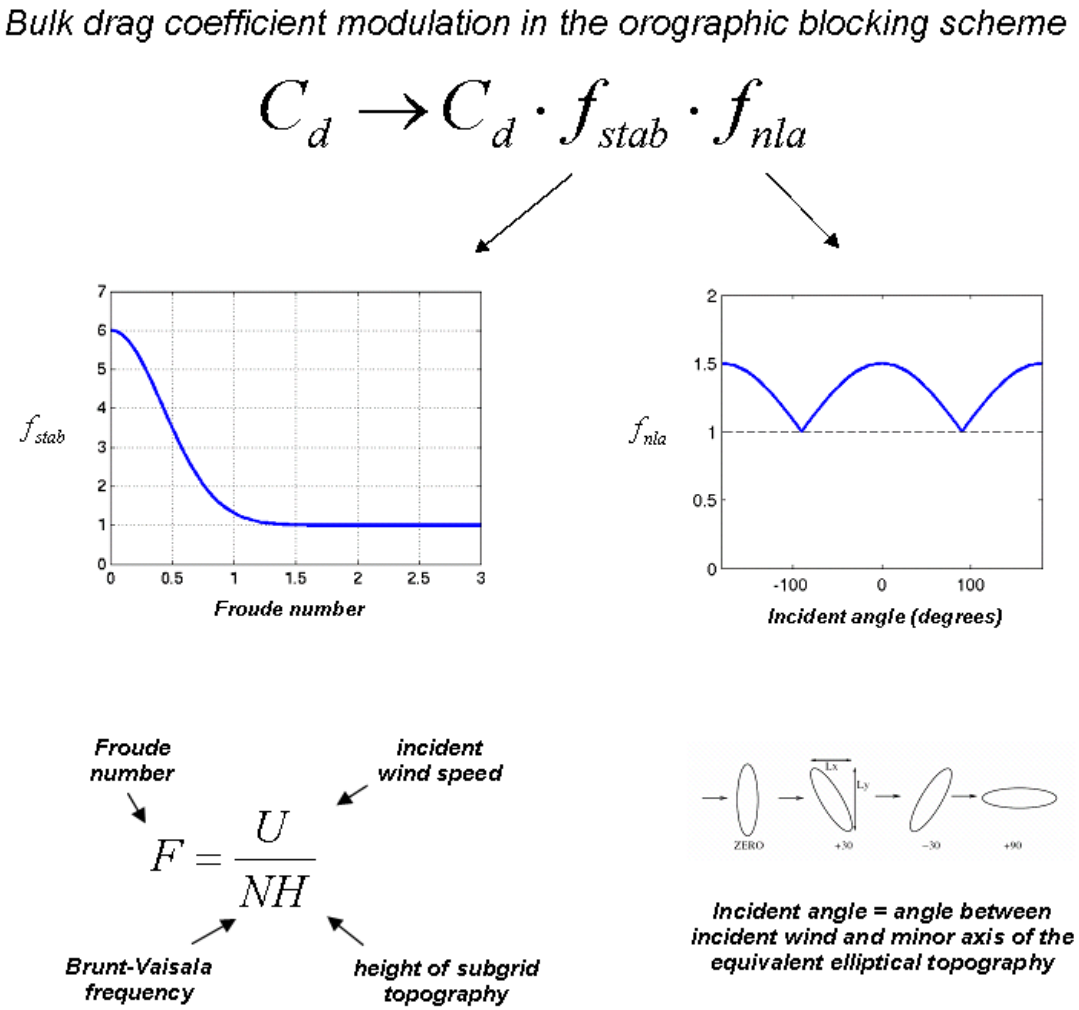


Figure 10: Schematics of the enhancement factors for the bulk drag coefficient of the orographic blocking scheme.

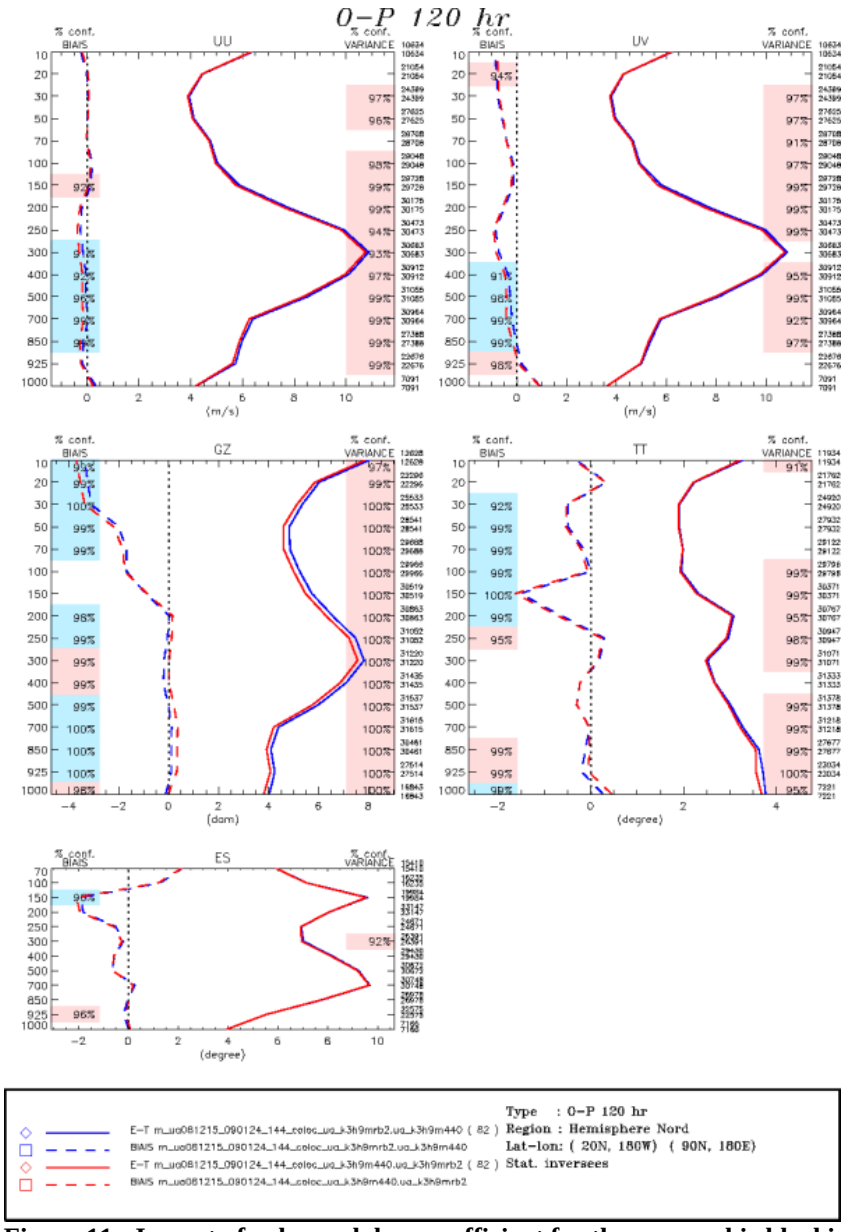


Figure 11: Impact of enhanced drag coefficient for the orographic blocking: 120-h upper-air scores for the winter season over the North Hemisphere; constant drag in blue, enhanced drag in red.

Diffusion coefficient: $K = 0.516 \frac{\lambda_{mix}}{\phi(Ri)} \sqrt{E}$

λ_{mix} = mixing length
ϕ = stability function
Ri = Richardson number
E = turbulent kinetic energy (TKE)

TKE equation: $\frac{\partial E}{\partial t} = BE^{1/2} - CE^{3/2} + \frac{\partial}{\partial z}\left(K \frac{\partial E}{\partial z}\right)$

production/destruction term — *dissipation term* — *diffusion term*

TKE production/destruction coefficient B (in the original scheme):

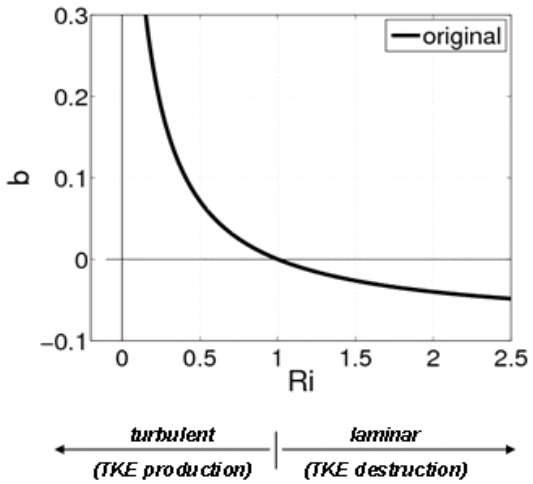
wind shear

$$B = 0.516 \cdot \lambda_{mix} \cdot \left|\frac{\partial V}{\partial z}\right|^2 \cdot b(Ri)$$

$$b(Ri) = \frac{1 - Ri/\Pr}{\phi_m(Ri)}$$

$$\Pr = \frac{\phi_t(Ri)}{\phi_m(Ri)}$$

thermal and momentum stability functions



New hysteresis effect

- Within a transition range, the coefficient b now depends on Ri and on its past state (hysteresis), as shown in the figure beside.
- Turbulent-laminar transition: instead of a single critical value (Ri = 1), the transition is now defined by a pair of Ri values
  - turbulent-to-laminar transition occurs at Ri = 2
  - laminar-to-turbulent transition occurs at Ri = 0.25

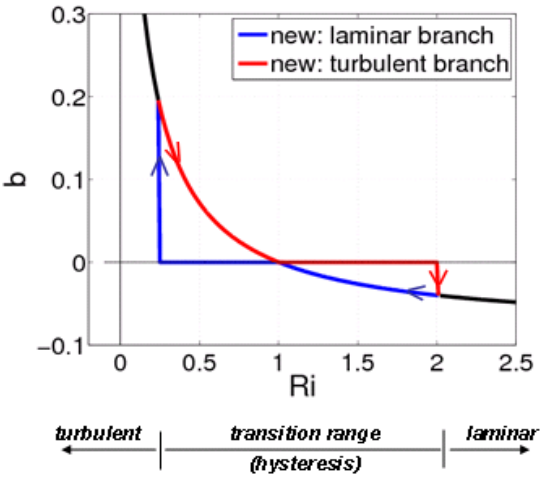


Figure 12: Implementation of turbulent hysteresis in the boundary layer scheme.

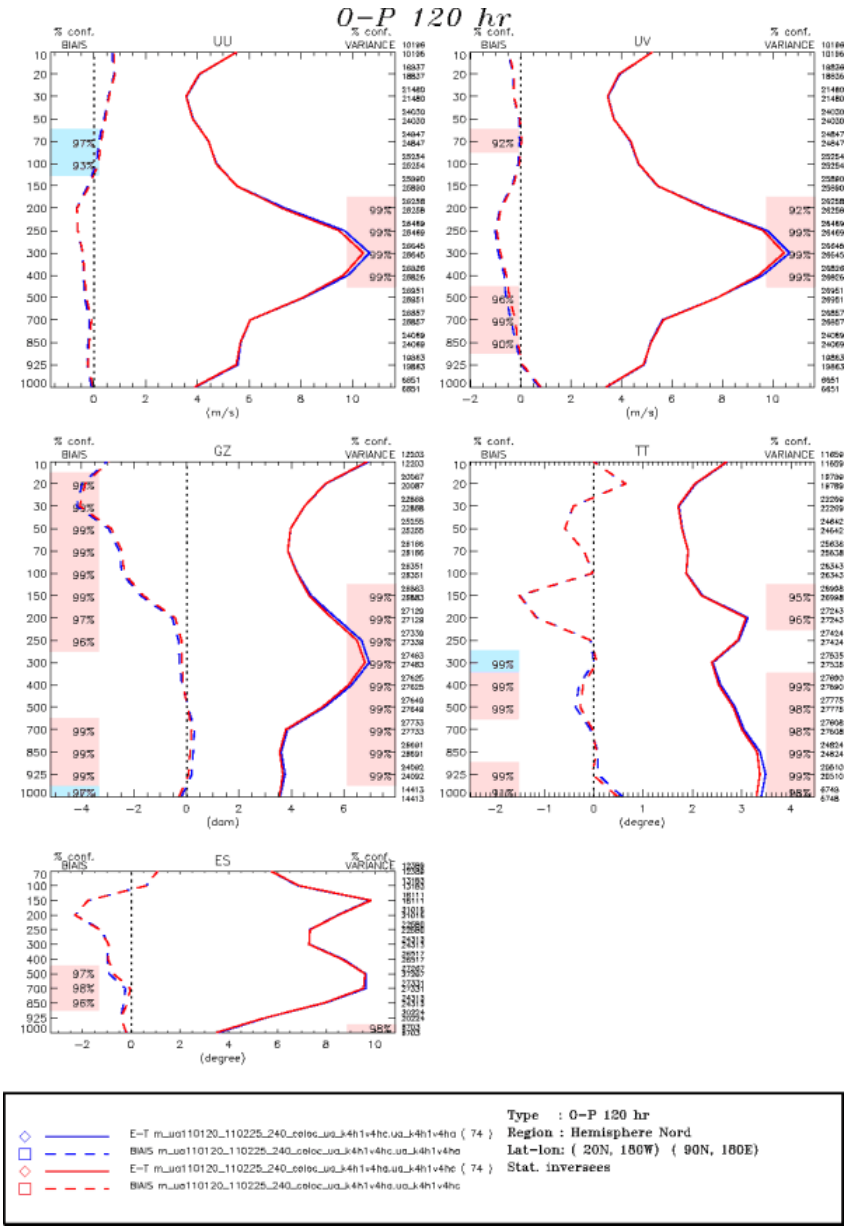


Figure 13: Impact of turbulent hysteresis (as in Figure 11)

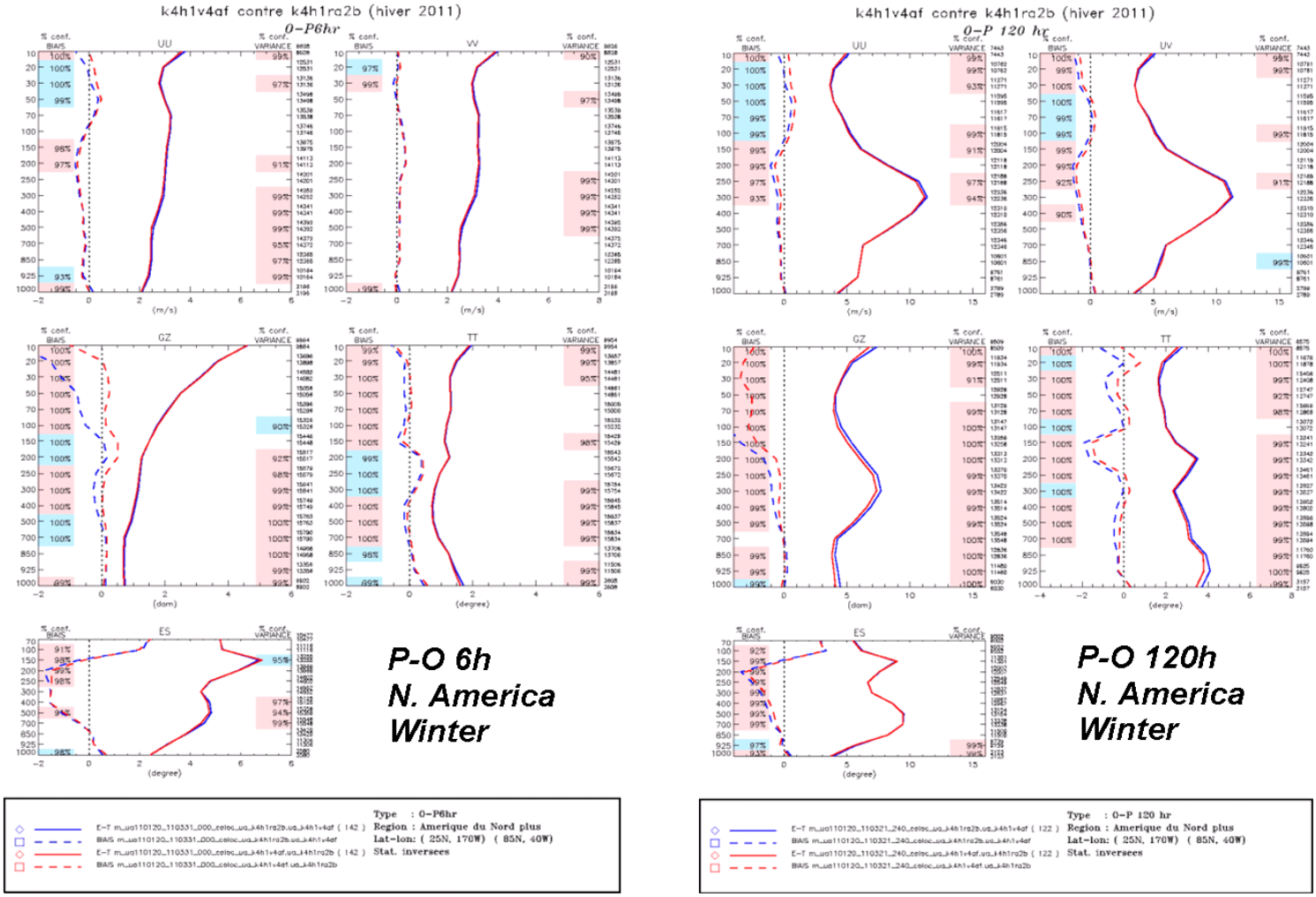


Figure 14: **Upper-air scores over North America, for the winter cycle at t=6h (left) and t=120h (right). Control system in blue, new system in red.**

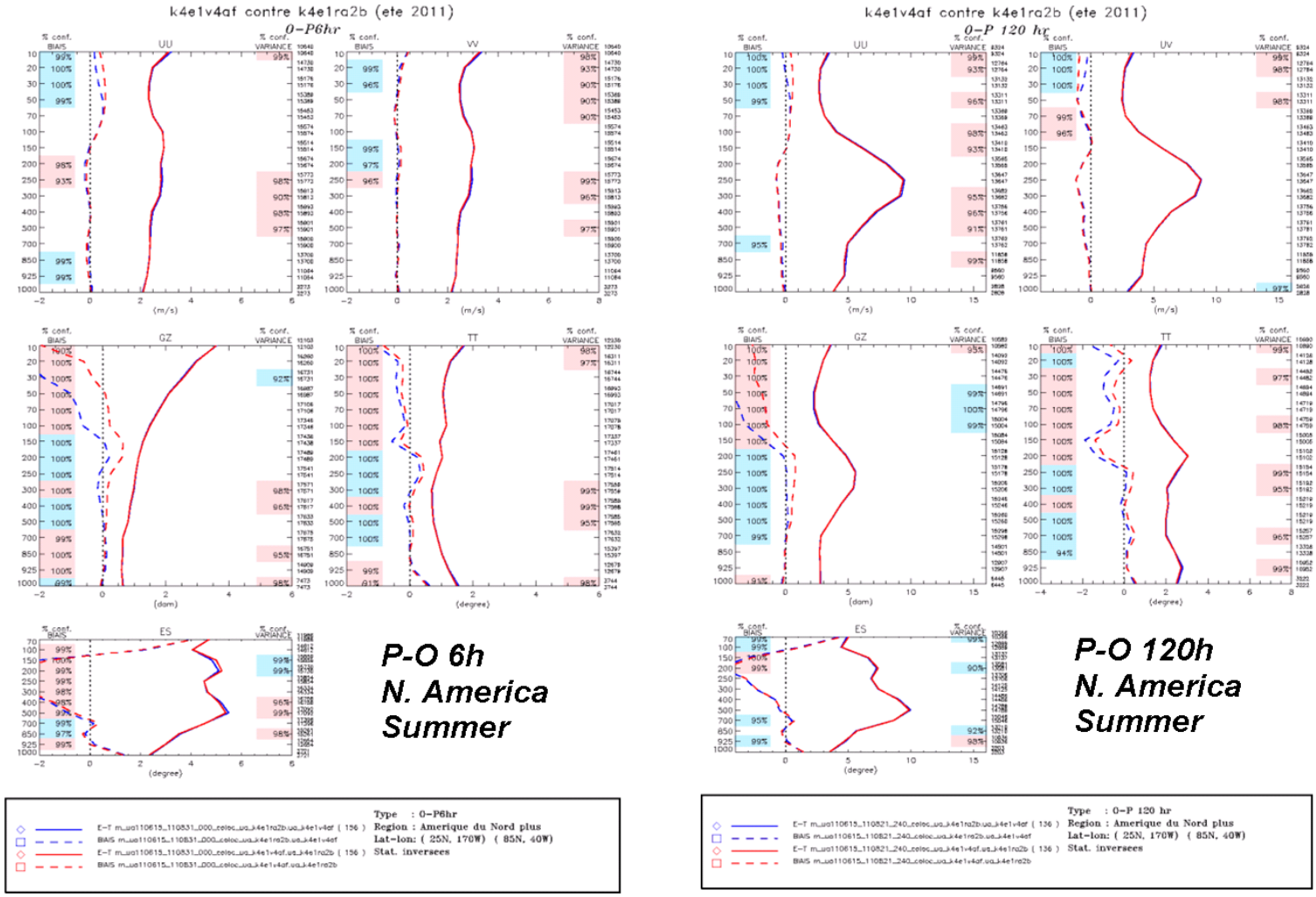


Figure 15: Upper-air scores over North America, for the summer cycle, at t=6h (left) and t=120h (right). Control system in blue, new system in red.

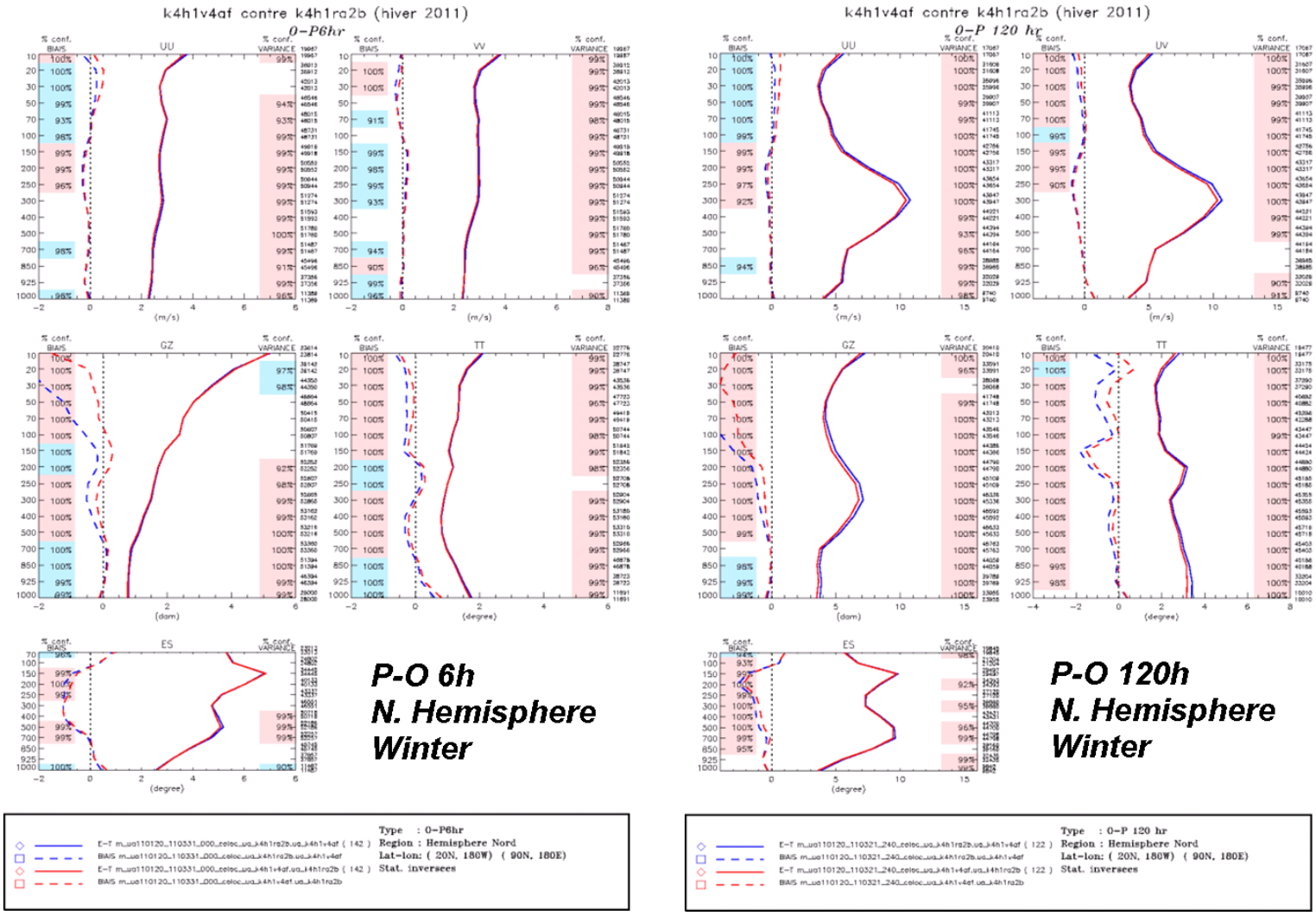
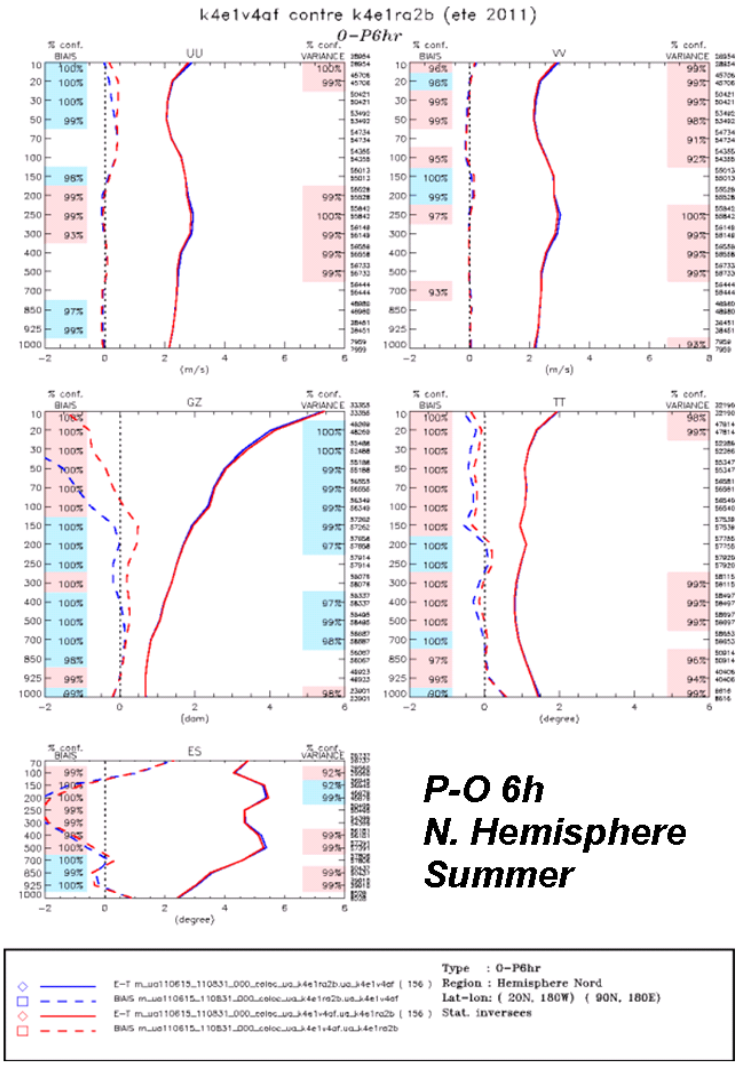


Figure 16: **Upper-air scores over the North Hemisphere, for the winter cycle at t=6h (left) and t=120h (right). Control system in blue, new system in red.**



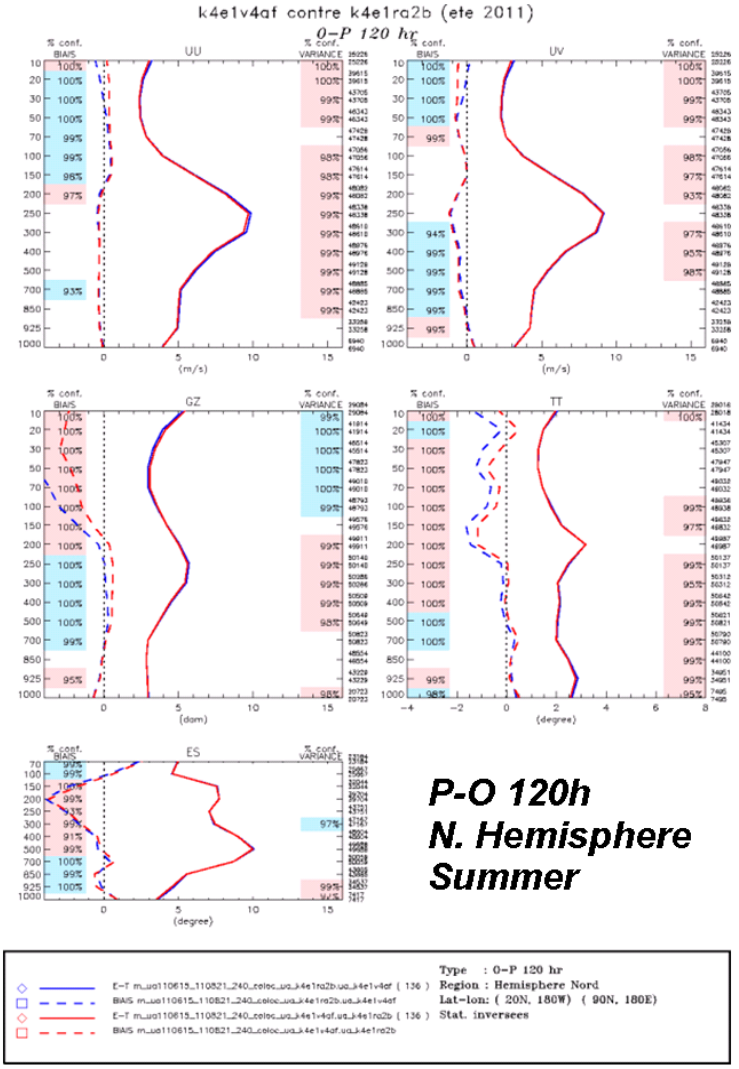


Figure 17: Upper-air scores over the North Hemisphere, for the summer cycle at t=6h (left) and t=120h (right). Control system in blue, new system in red

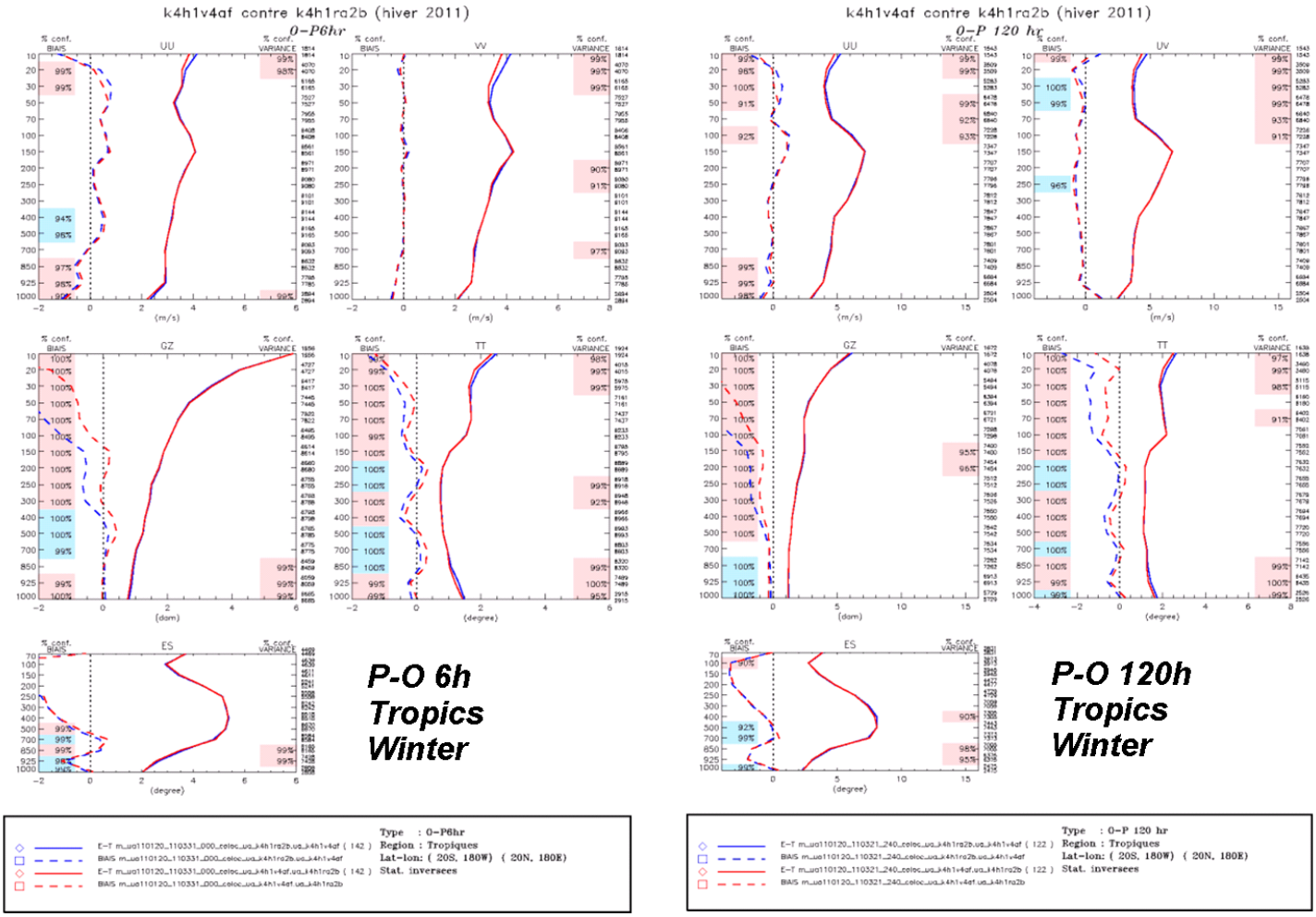


Figure 18: Upper-air scores over the Tropics, for the winter cycle at t=6h (left) and t=120h (right). Control system in blue, new system in red.

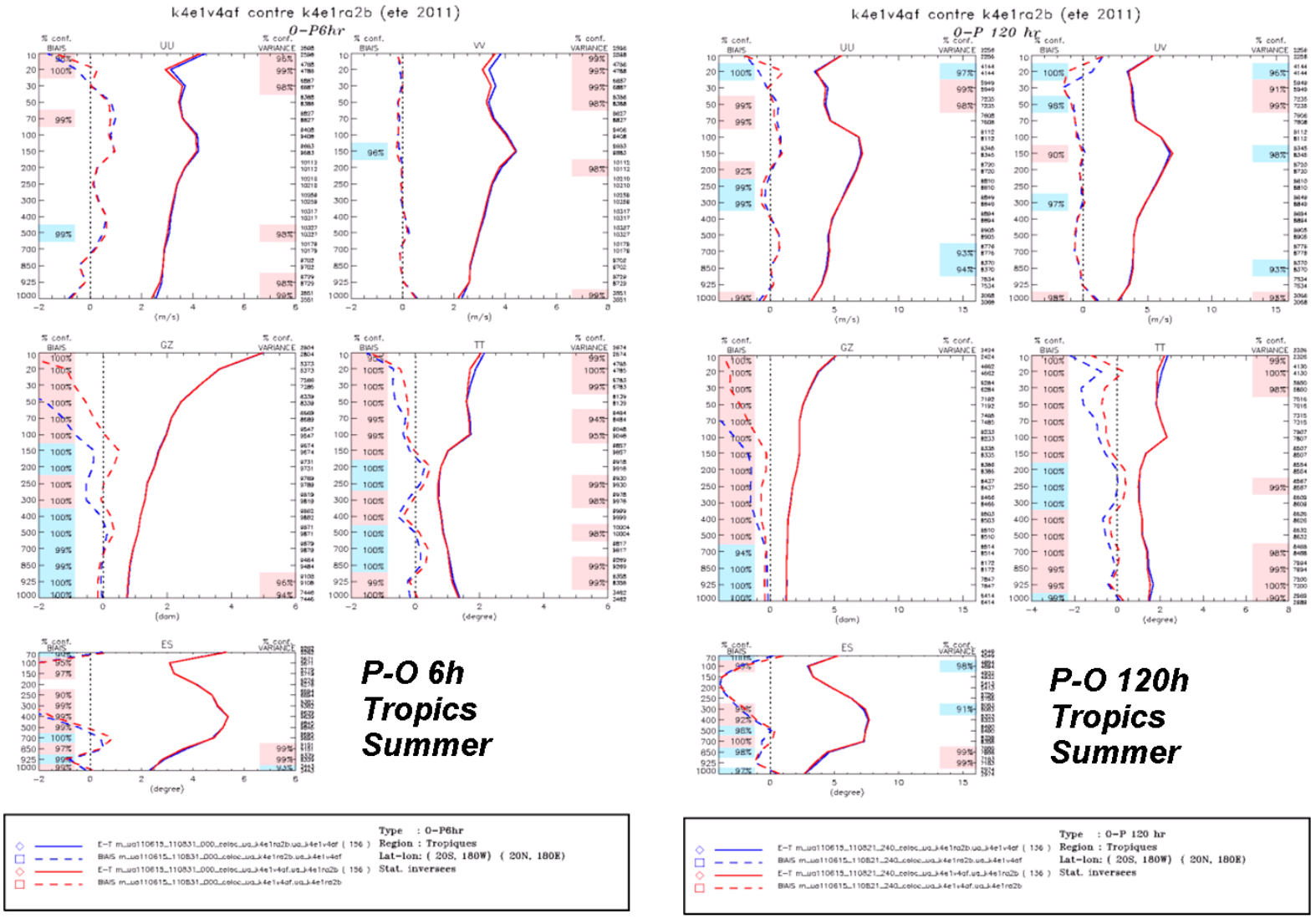


Figure 19: Upper-air scores over the Tropics, for the summer cycle at t=6h (left) and t=120h (right). Control system in blue, new system in red.

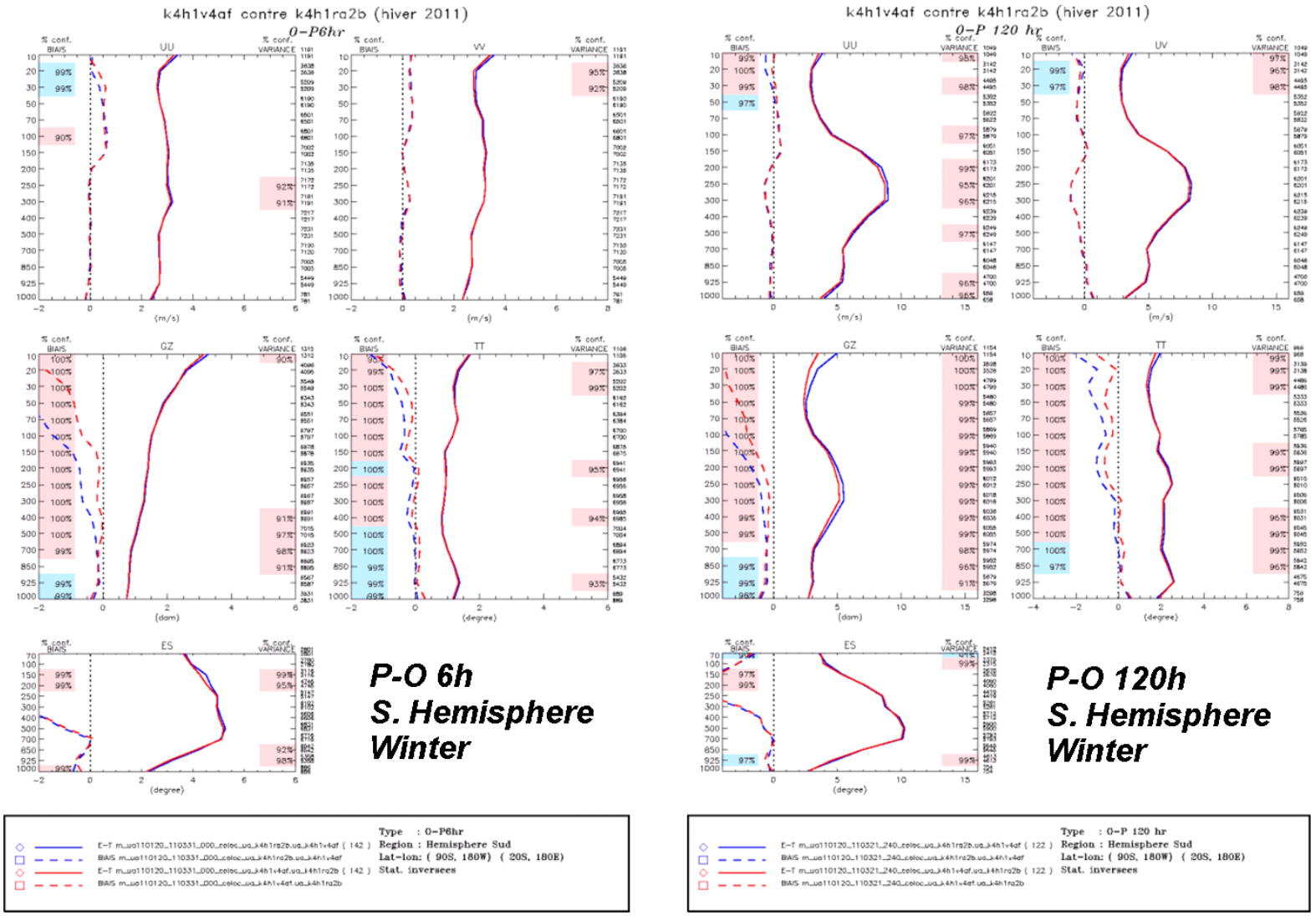
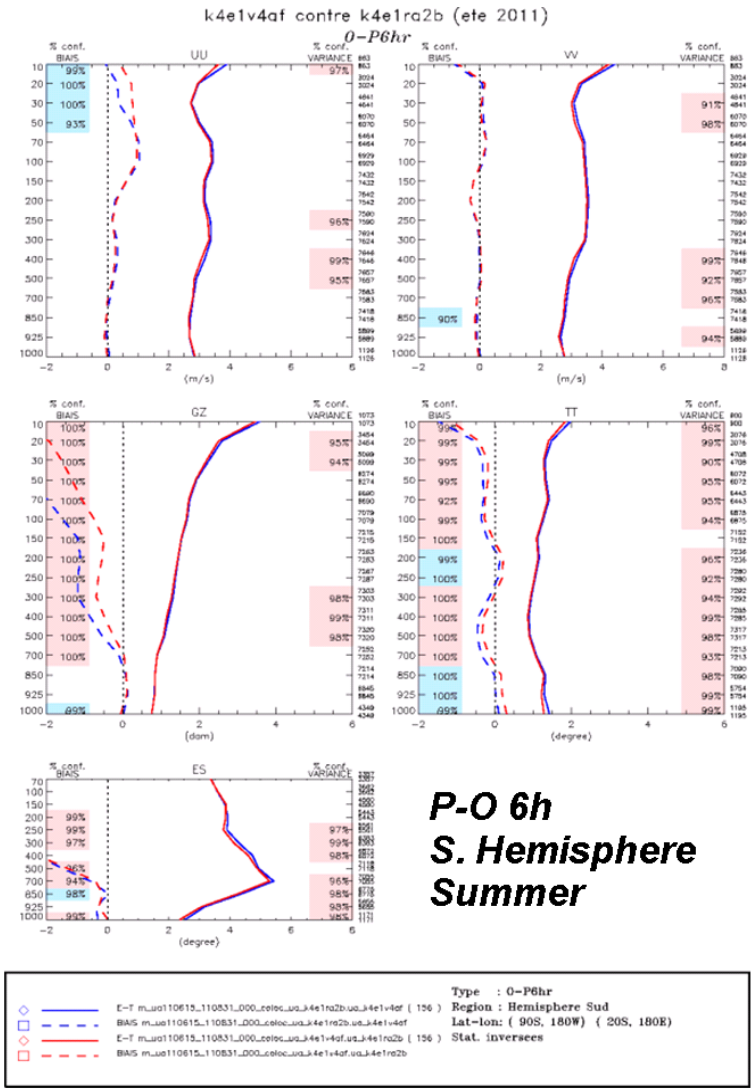


Figure 20: Upper-air scores over the S. Hemisphere, for the winter cycle at t=6h (left) and t=120h (right). Control system in blue, new system in red.



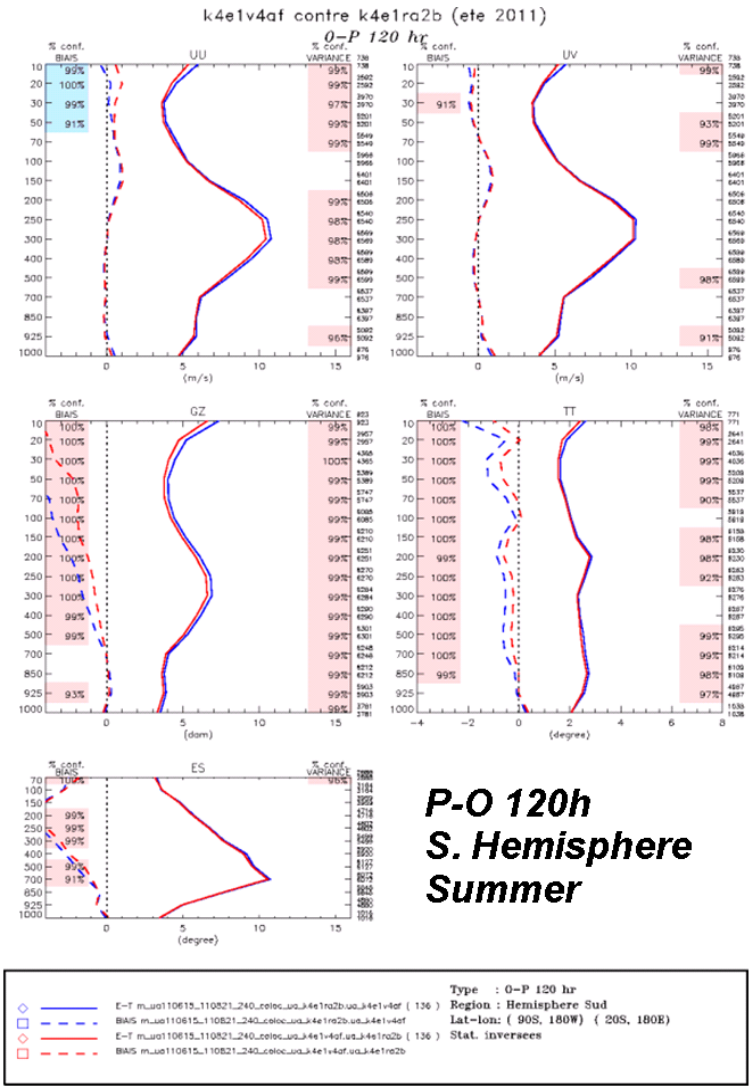
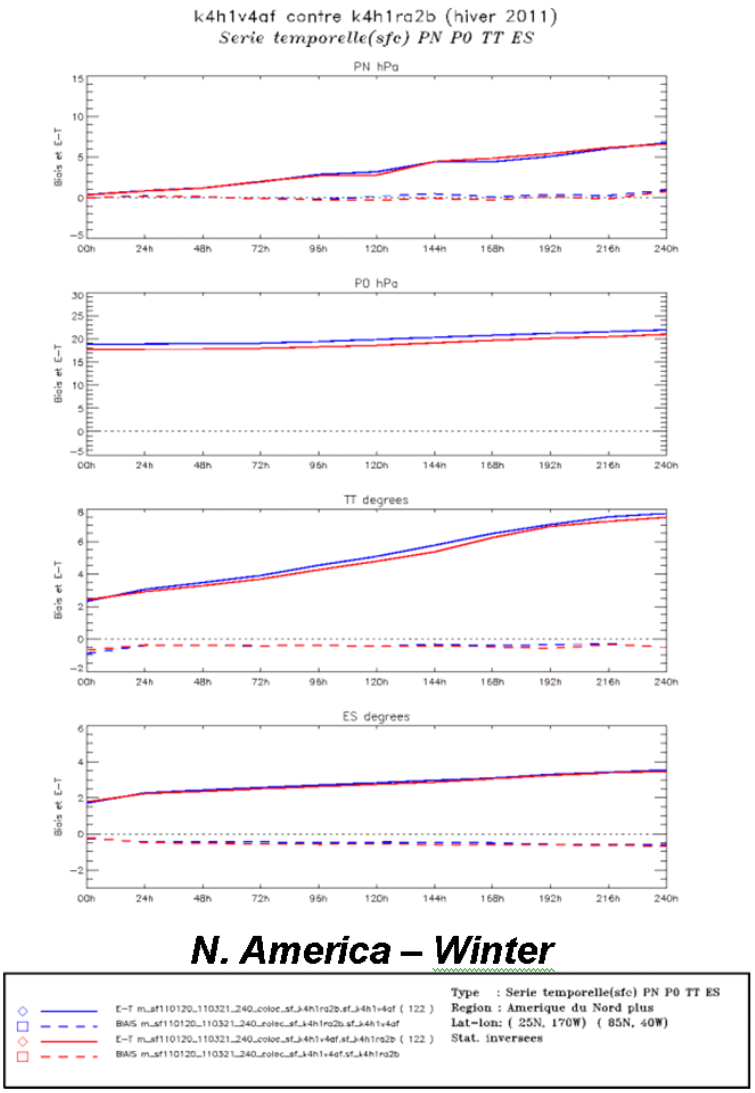


Figure 21: Upper-air scores over the S. Hemisphere, for the summer cycle at t=6h (left) and t=120h (right). Control system in blue, new system in red.



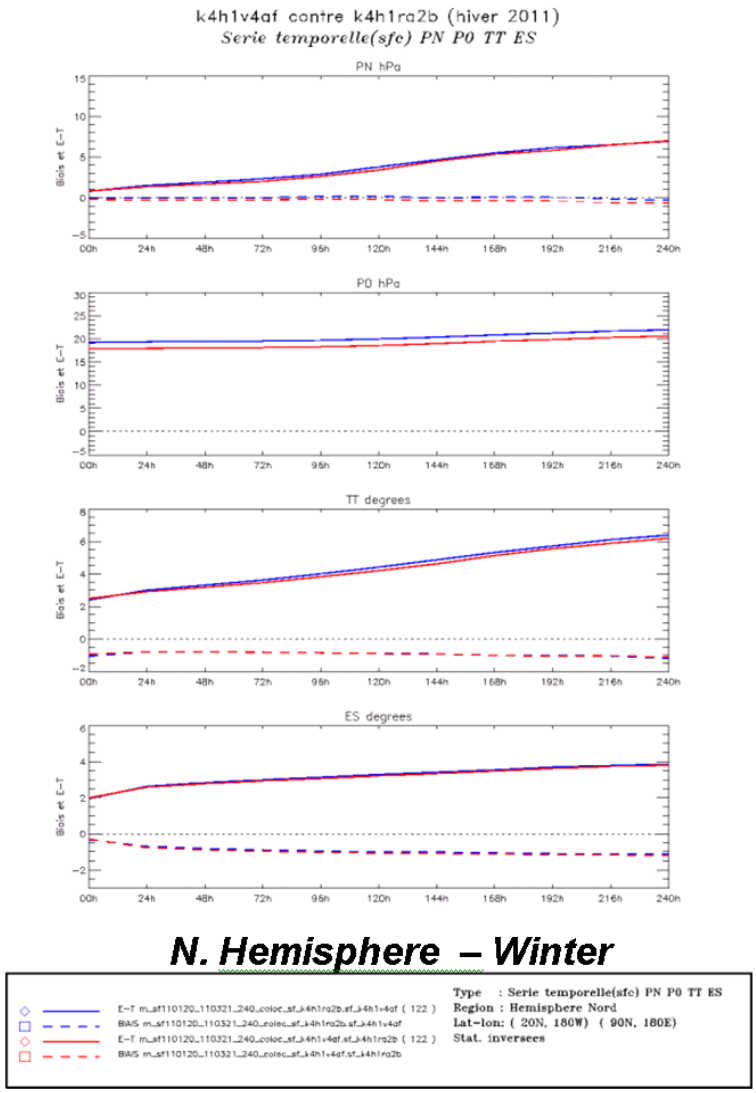
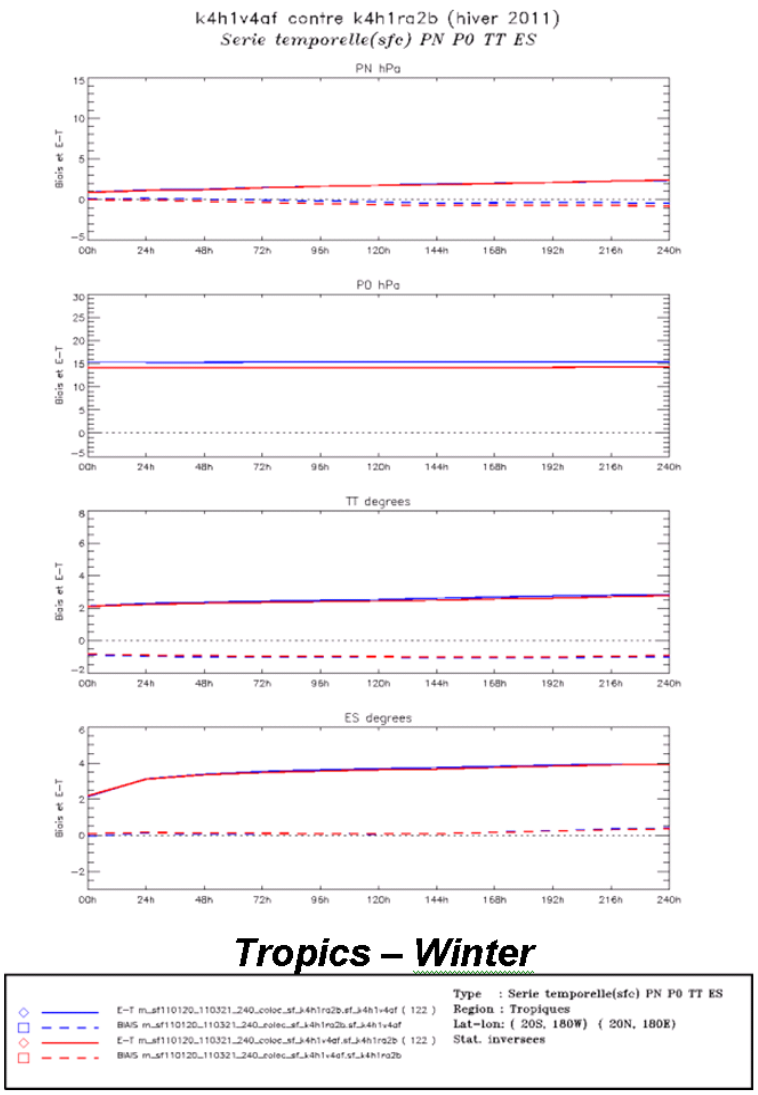


Figure 22: Surface scores over N. America and North Hemisphere for the winter cycle. Control system in blue, new system in red.



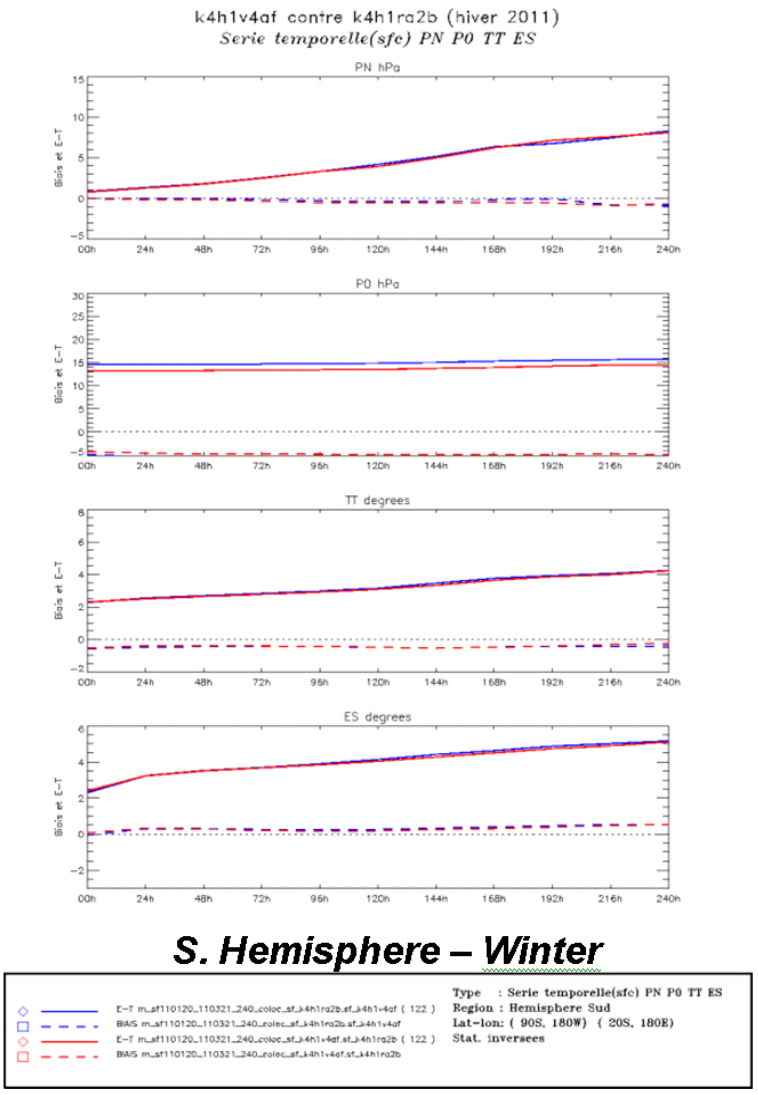
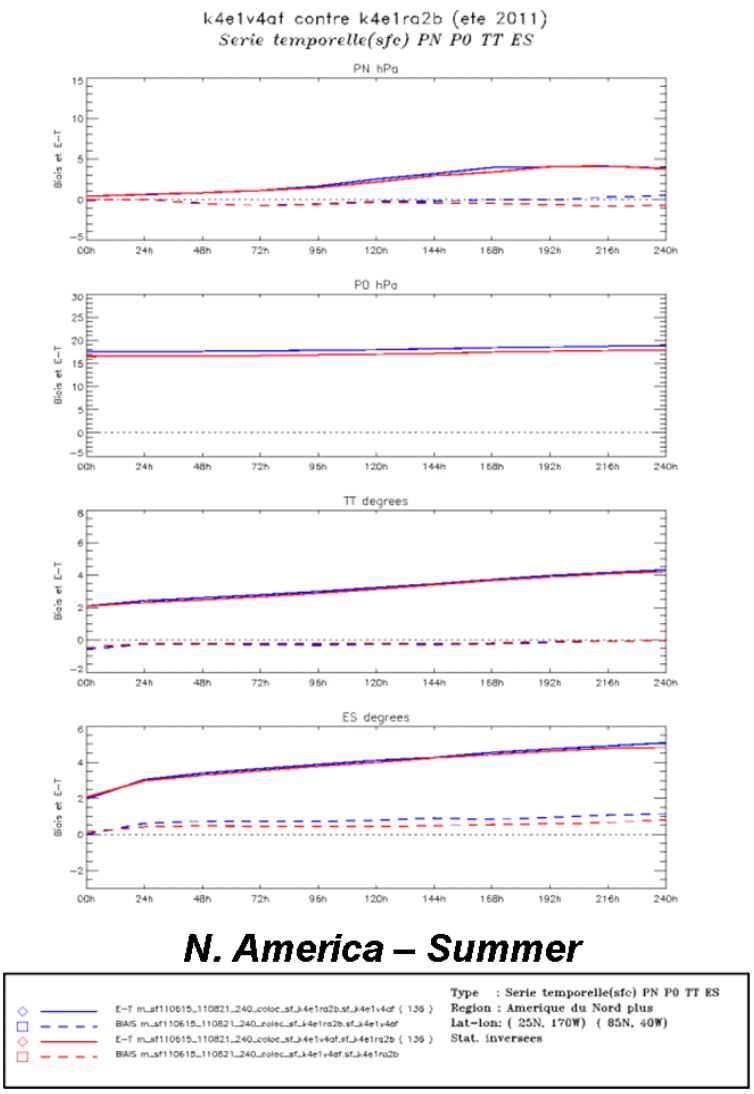


Figure 23: Surface scores over Tropics and South Hemisphere for the winter cycle. Control system in blue, new system in red.



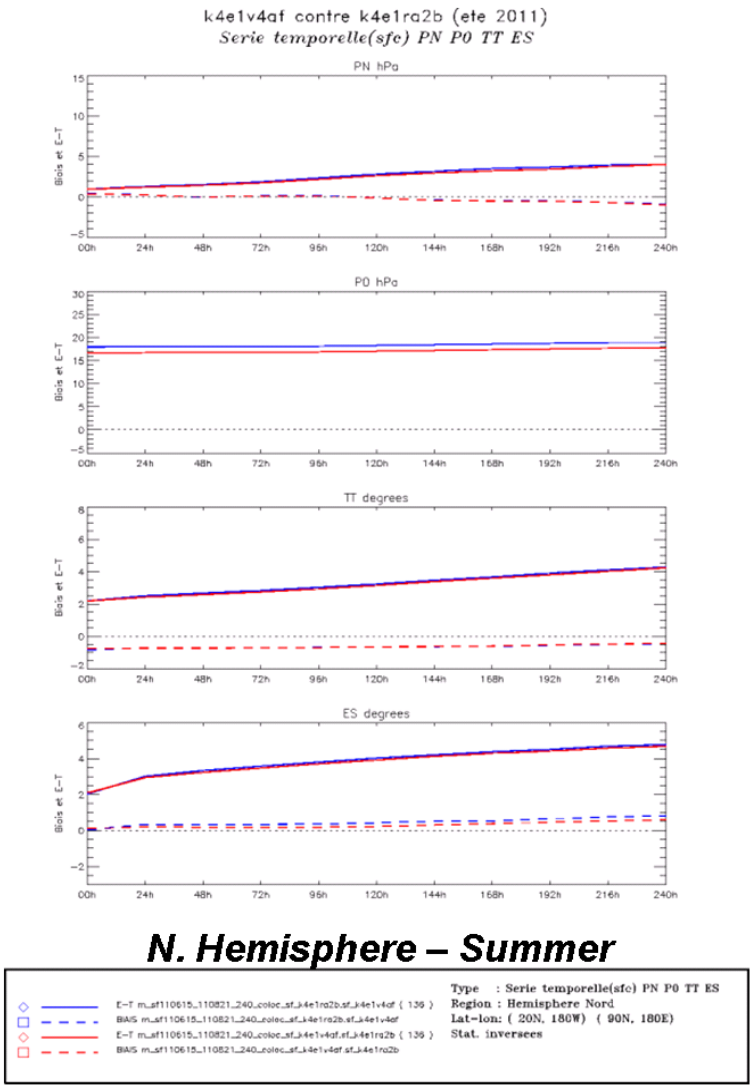


Figure 24: Surface scores over N. America and North Hemisphere for the summer cycle. Control system in blue, new system in red.

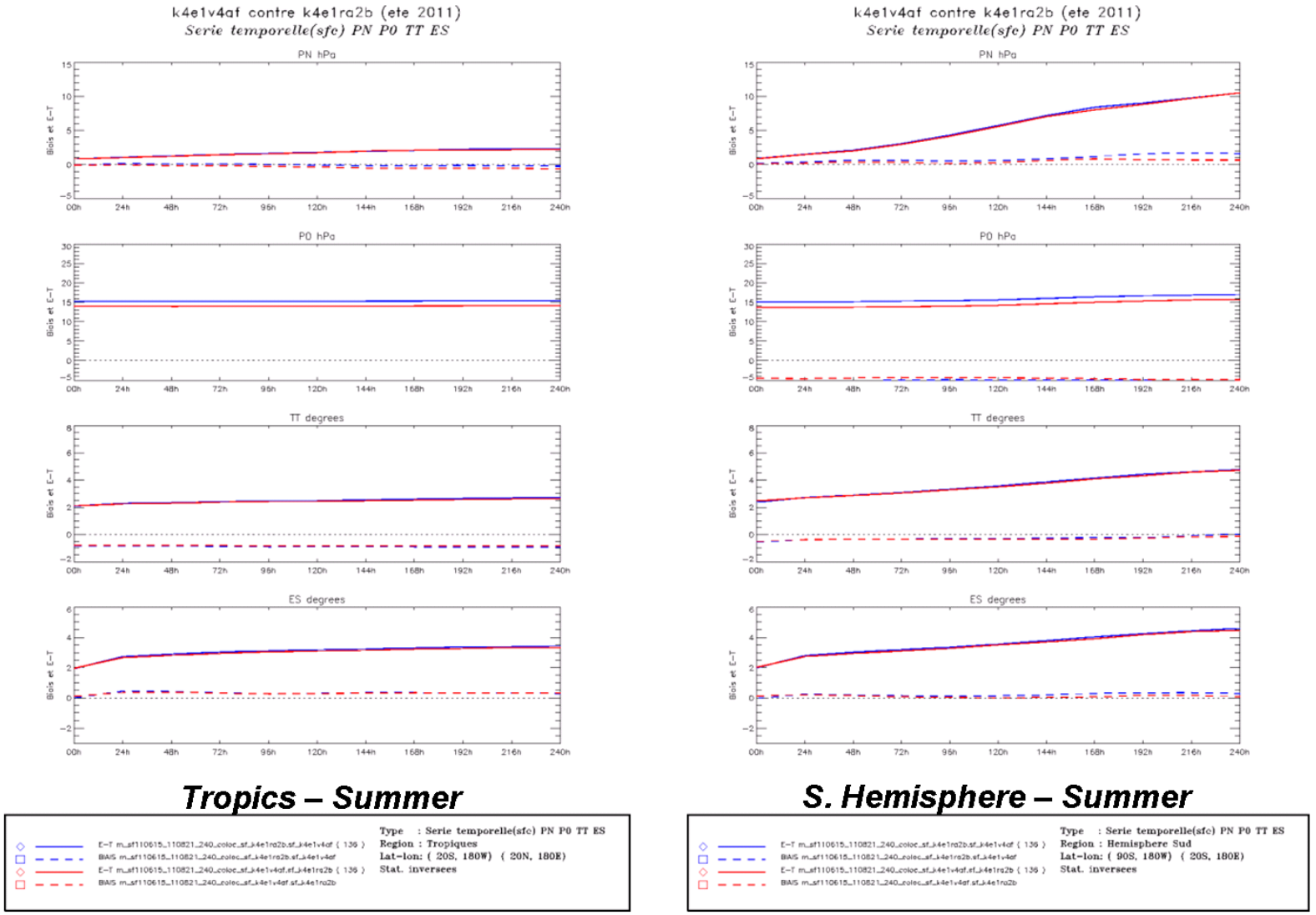


Figure 25: : Surface scores over Tropics and South Hemisphere for the summer cycle. Control system in blue, new system in red.

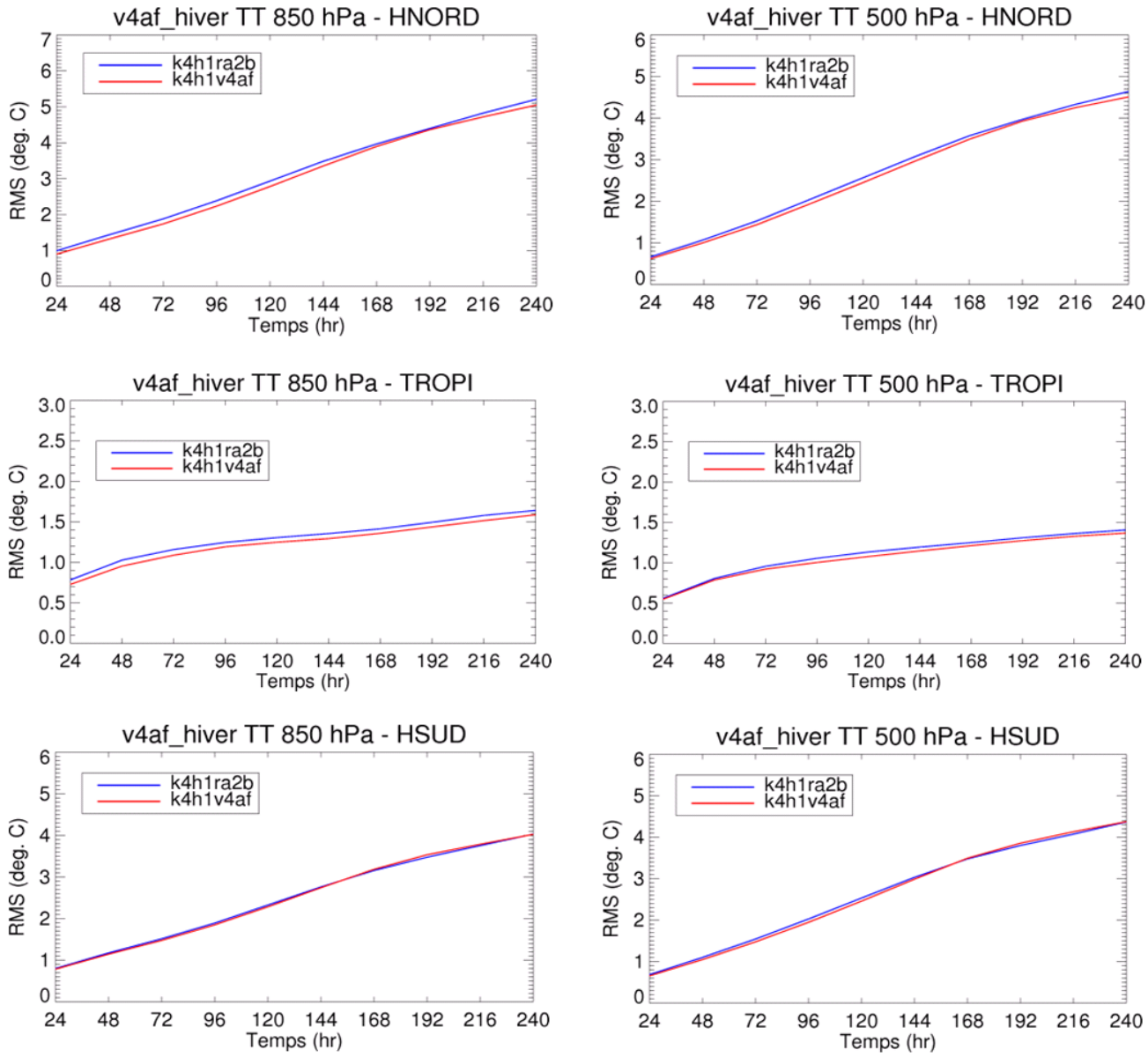


Figure 26: Winter scores against analyses: temperature RMSE at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

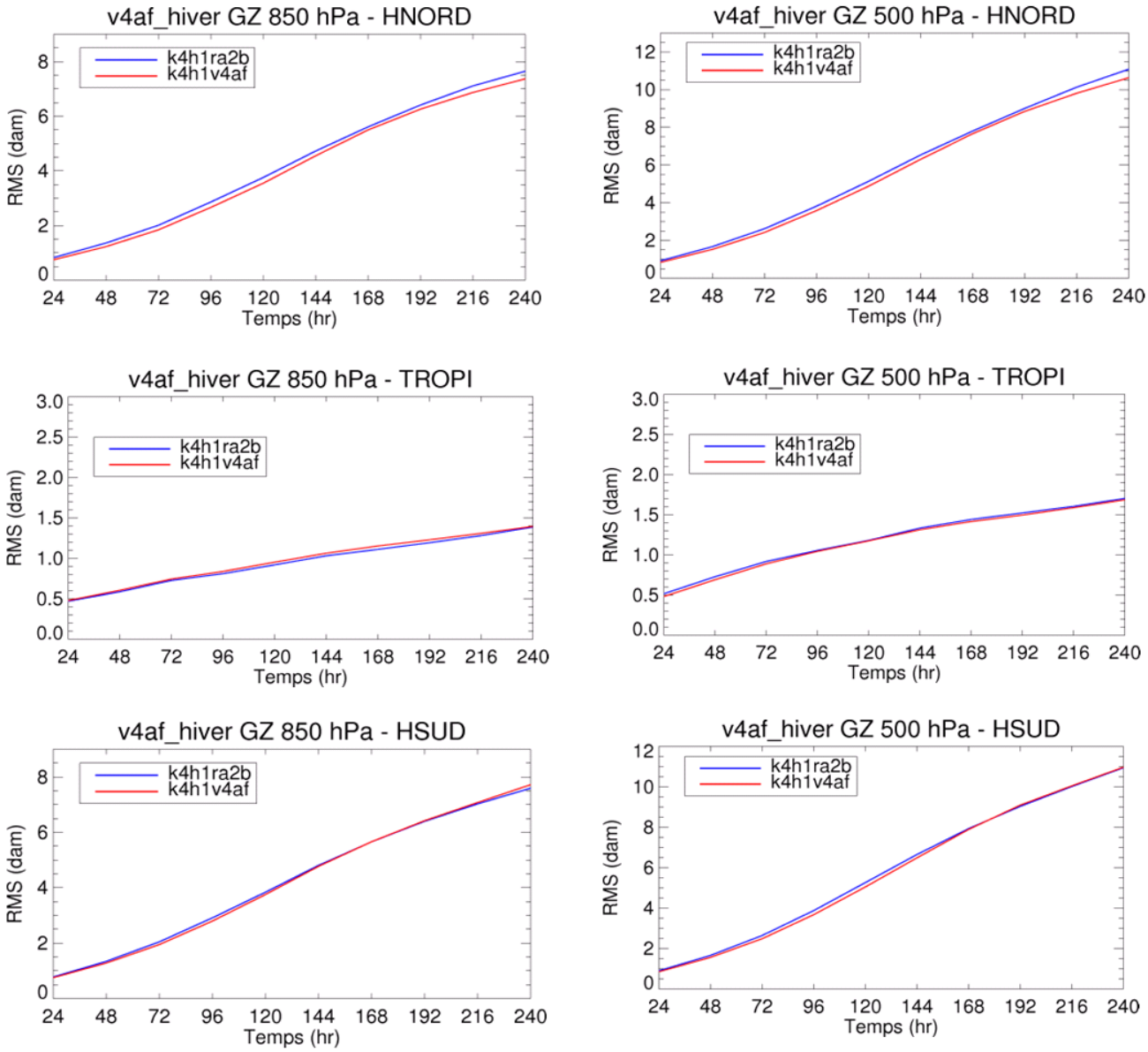


Figure 27: Winter scores against analyses: geopotential RMSE at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

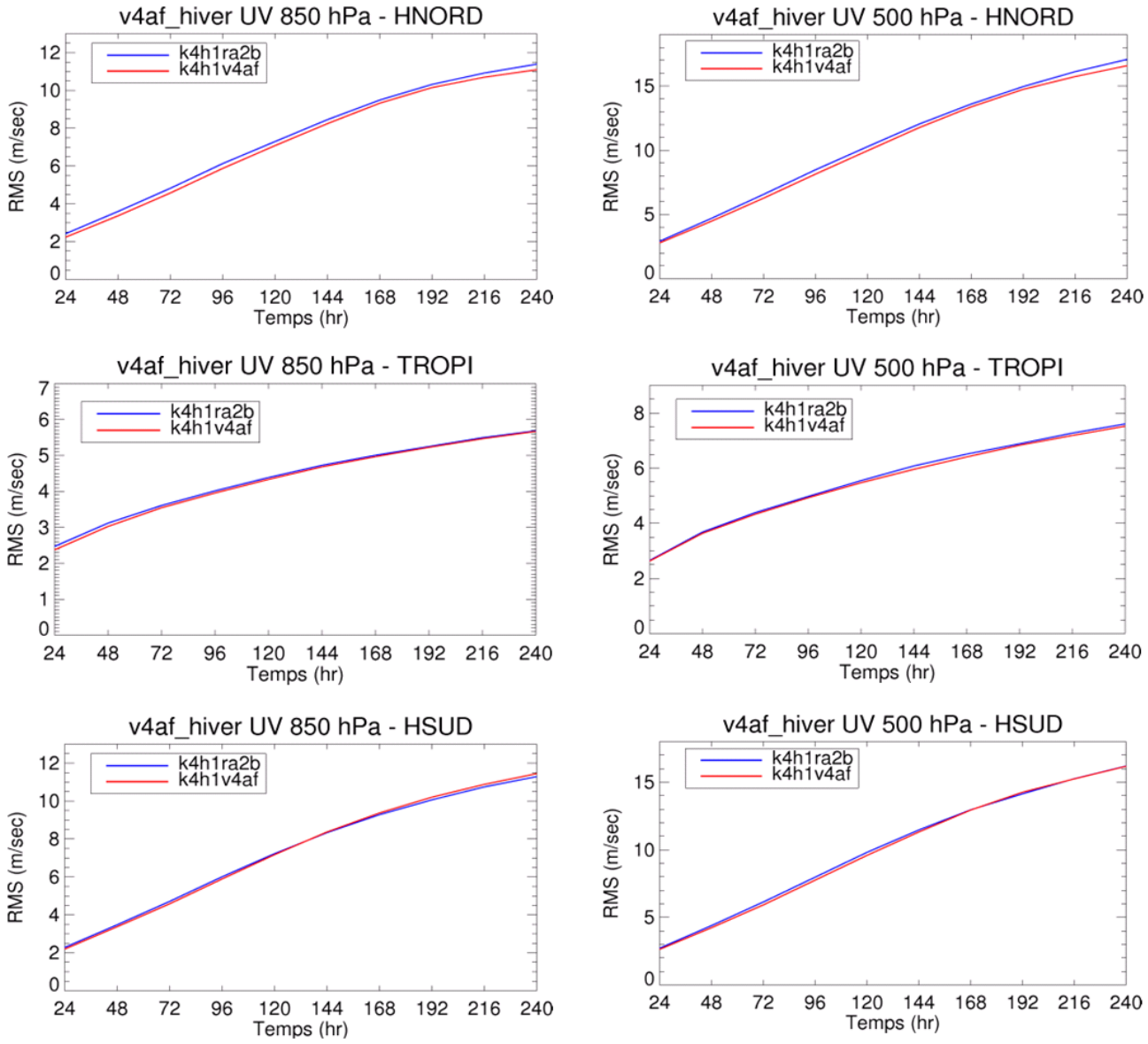


Figure 28: Winter scores against analyses: wind speed RMSE at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

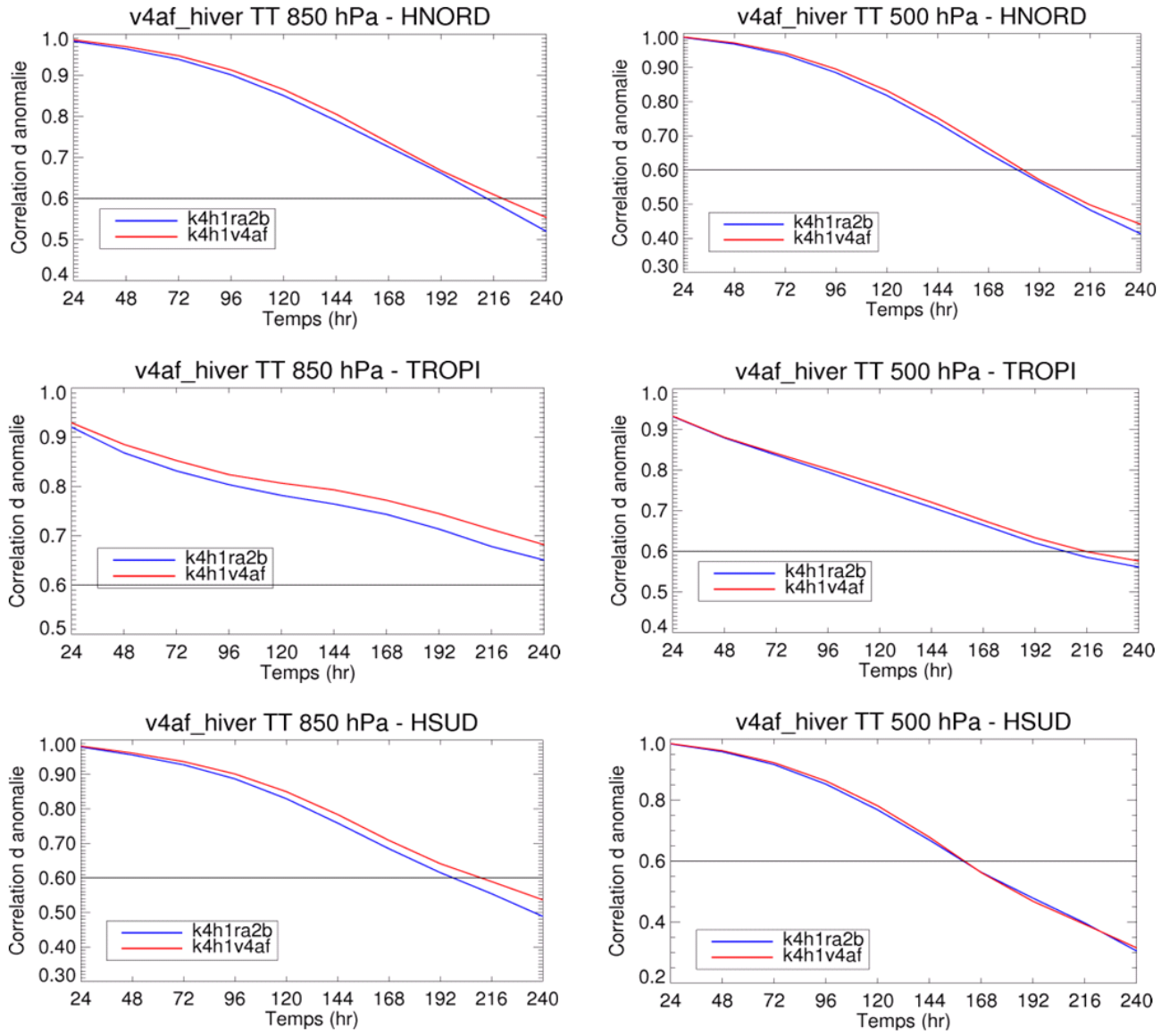


Figure 29: Winter scores against analyses: temperature anomaly correlations at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

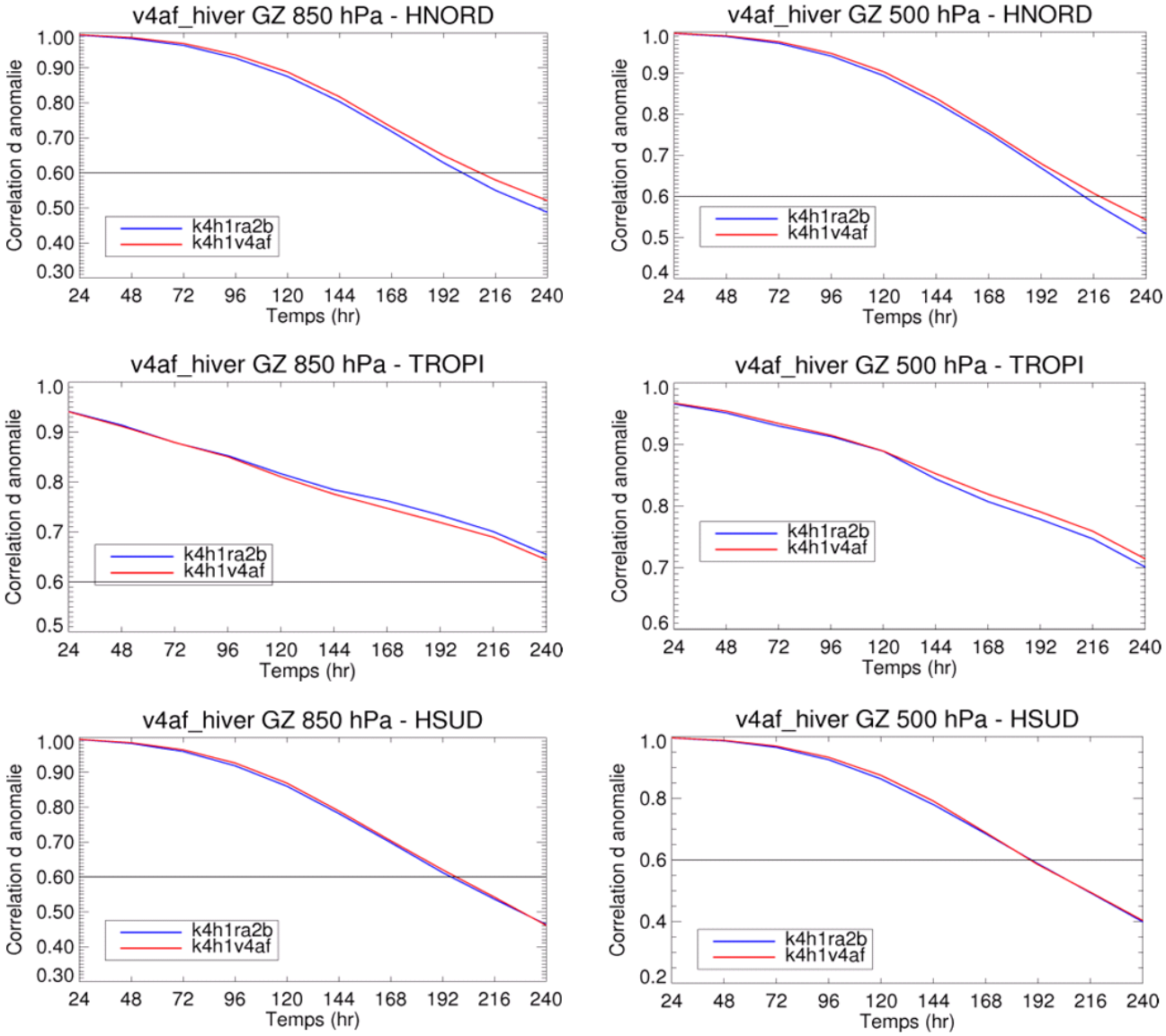


Figure 30: Winter scores against analyses: geopotential anomaly correlations at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

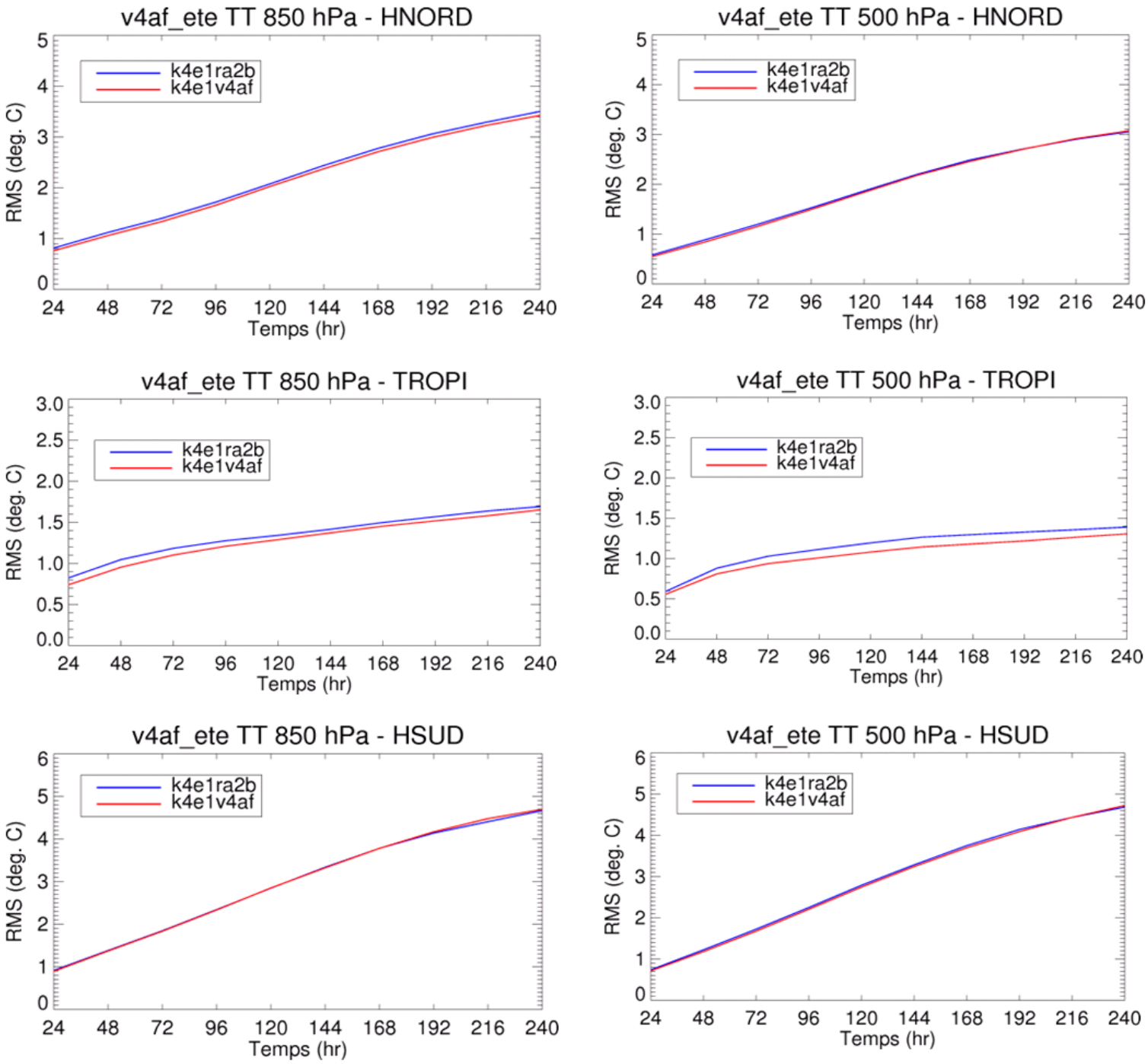


Figure 31: Summer scores against analyses: temperature RMSE at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

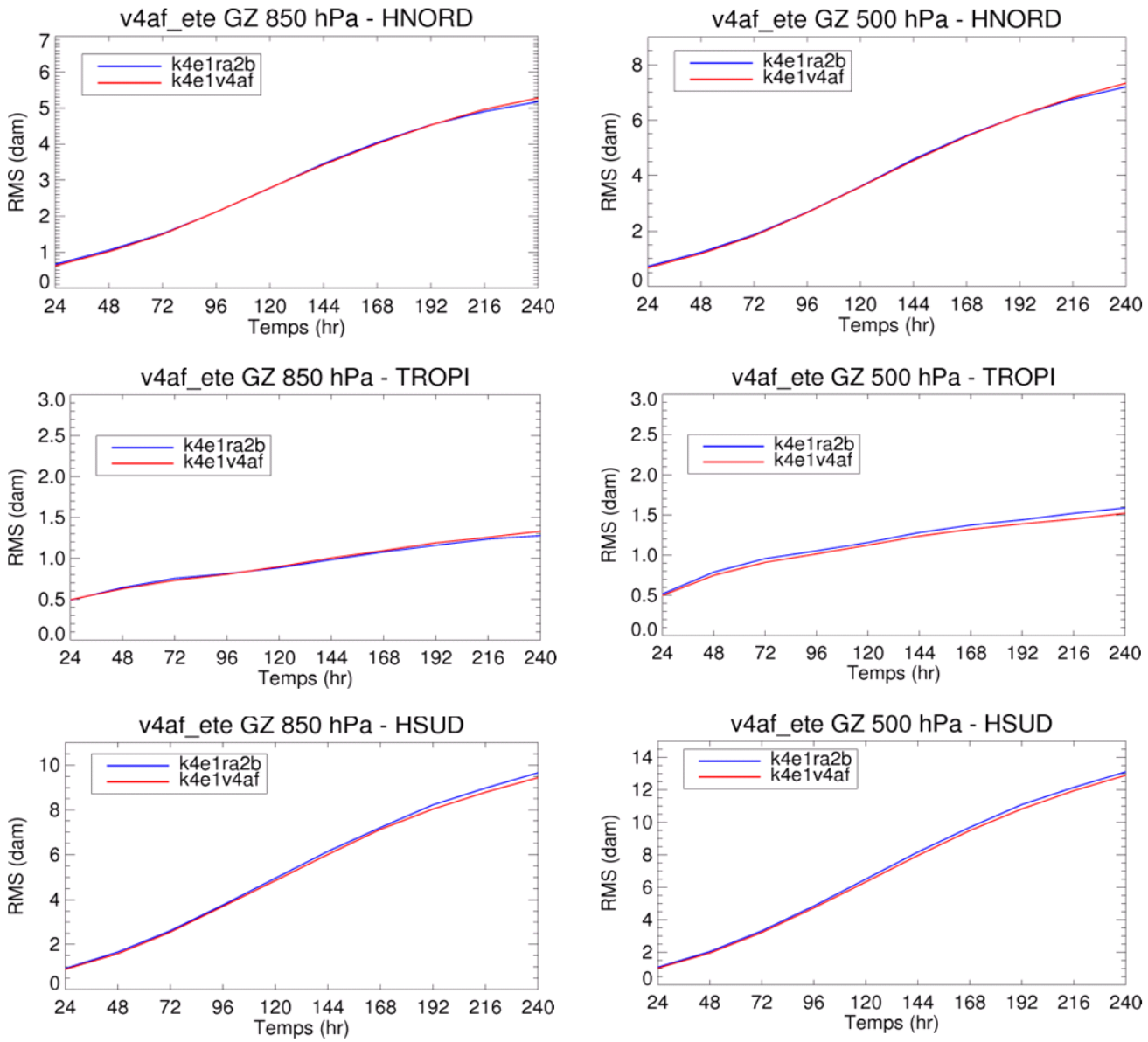


Figure 32: Summer scores against analyses: geopotential RMSE at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

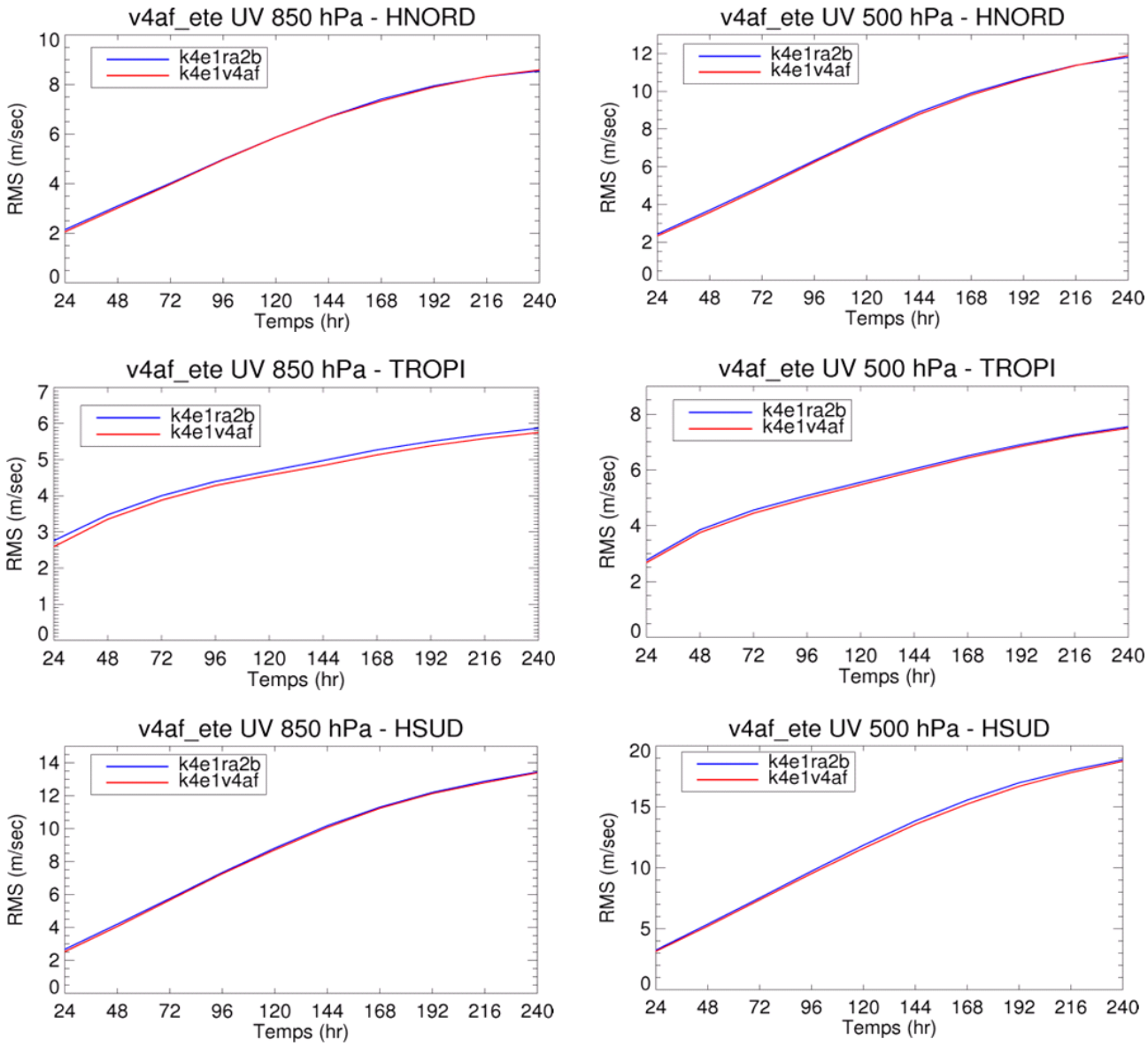


Figure 33: Summer scores against analyses: wind speed RMSE at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

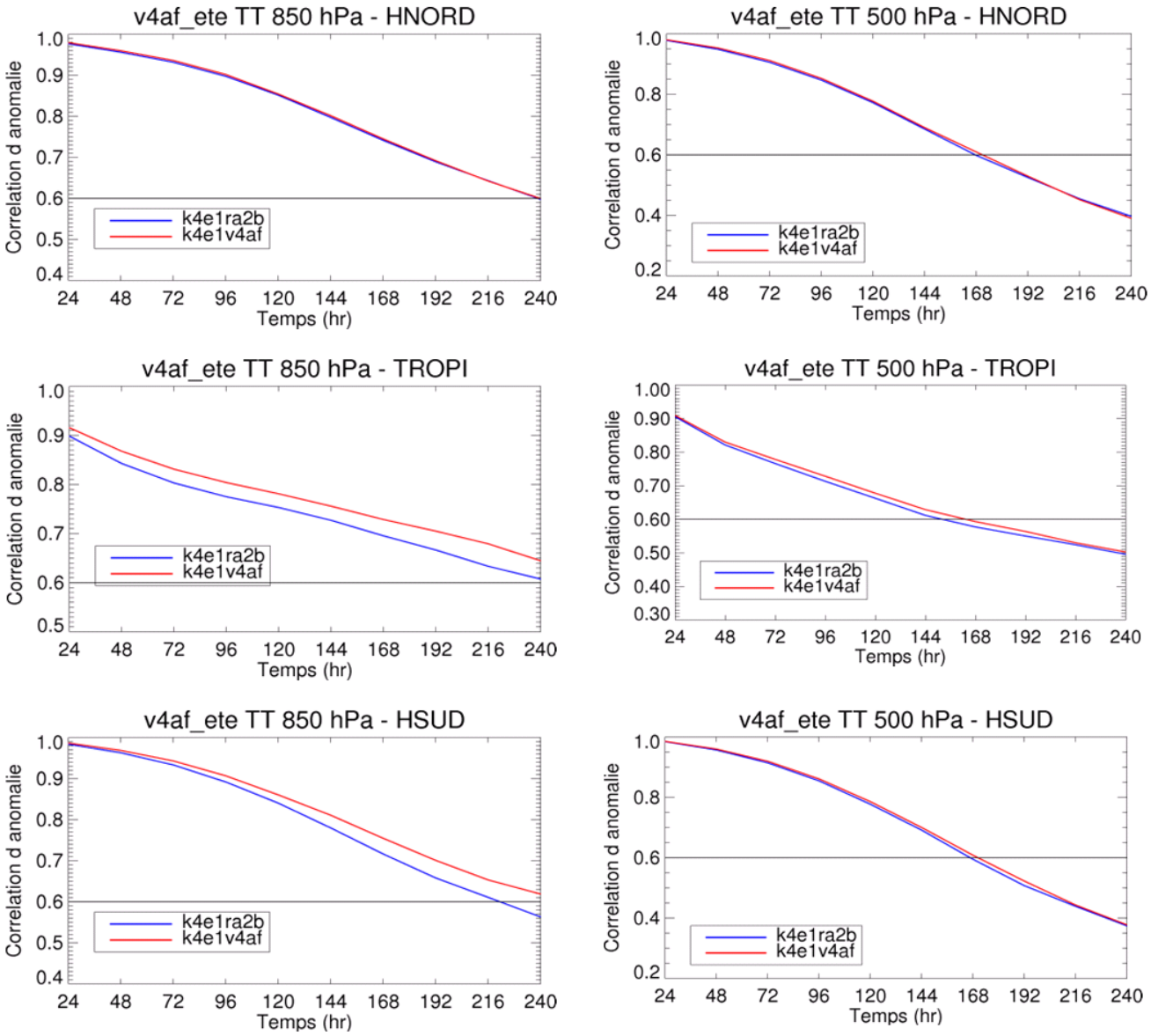


Figure 34: Summer scores against analyses: temperature anomaly correlations at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.

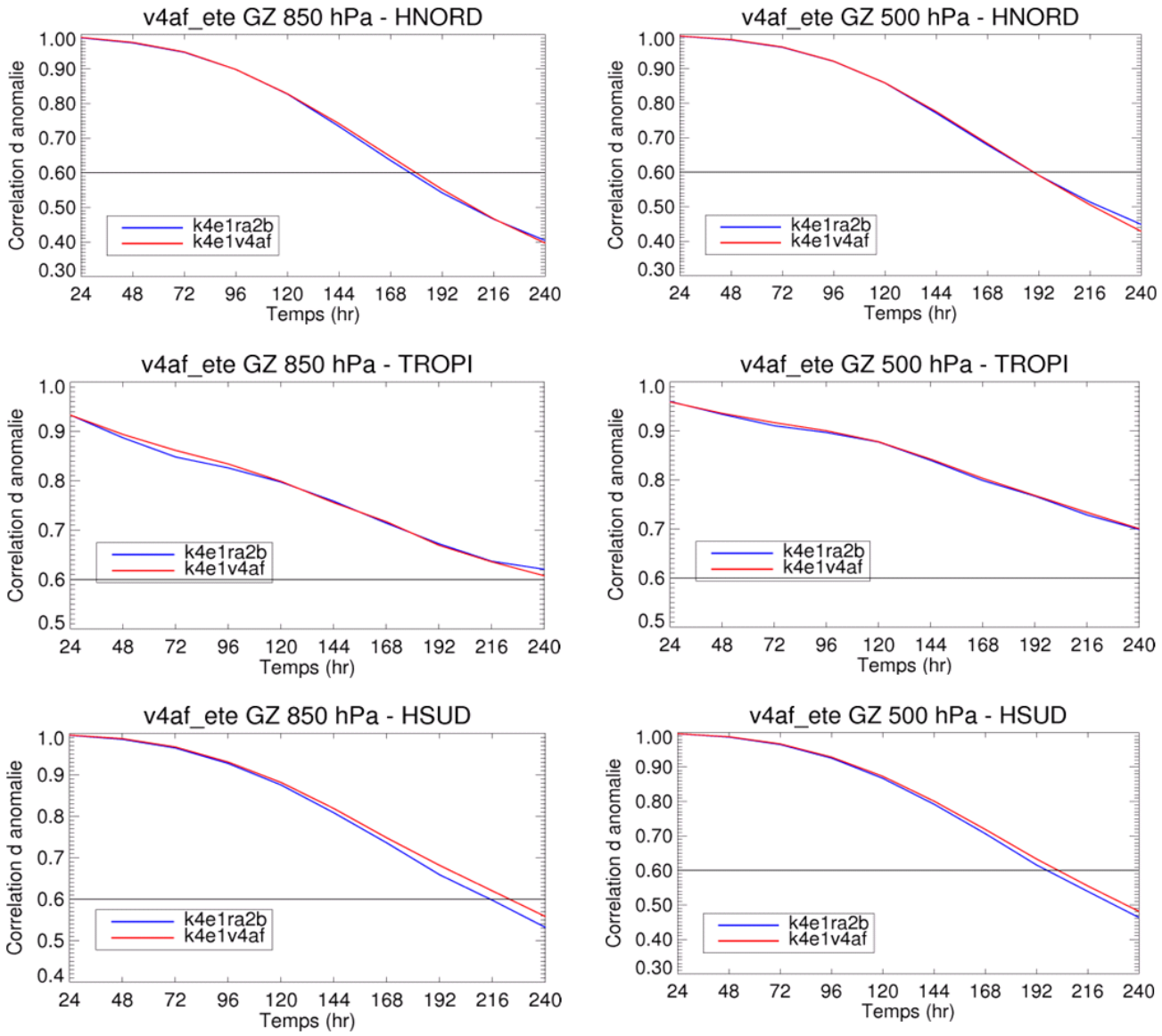
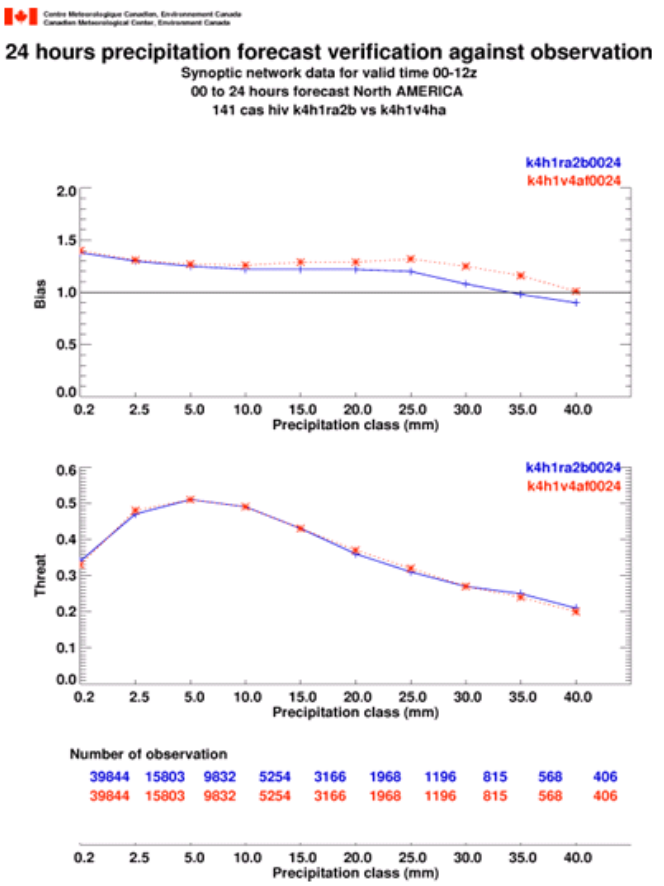
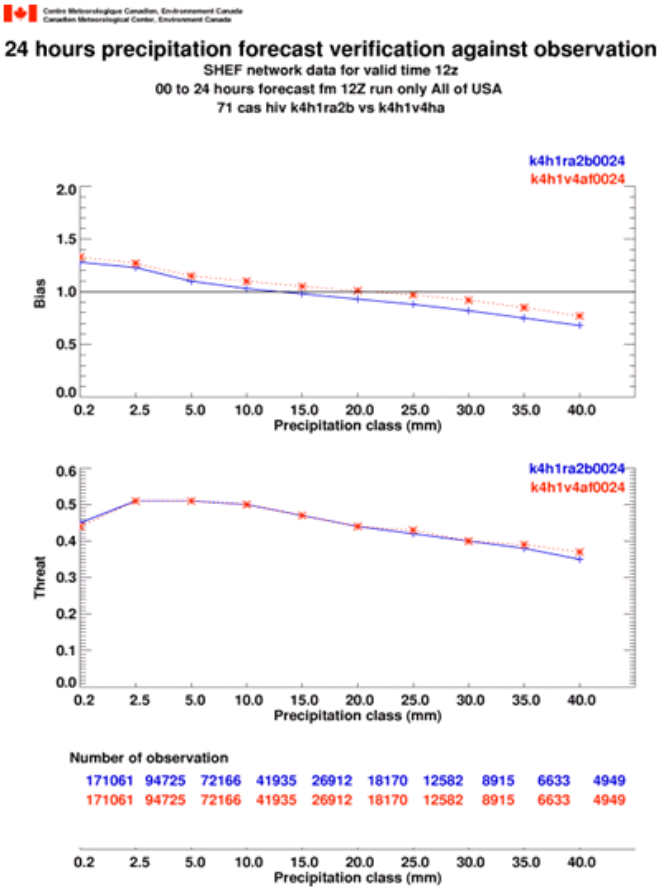
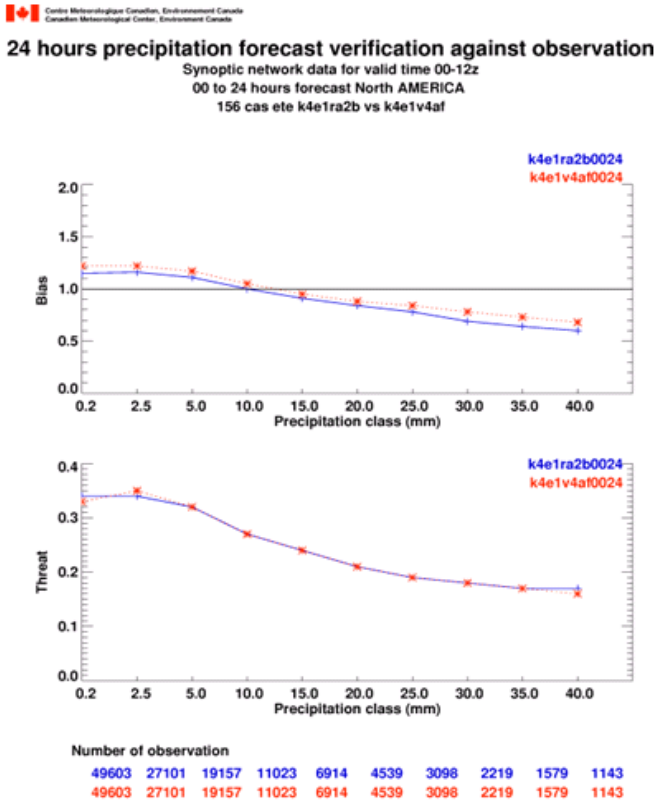


Figure 35: Summer scores against analyses: geopotential anomaly correlations at 850hPa (left) and 500hPa (right) over the North Hemisphere (top), Tropics (middle) and South Hemisphere (bottom). Control system in blue, new system in red.


Centre Meteorologique Canadien, Environnement Canada
Canadian Meteorological Center, Environment Canada
24 hours precipitation forecast verification against observation
Synoptic network data for valid time 00-12z
00 to 24 hours forecast North AMERICA
141 cas hiv k4h1ra2b vs k4h1v4ha
k4h1ra2b0024
k4h1v4af0024
Bias
Precipitation class (mm)
Threat
Number of observation
39844 15803 9832 5254 3166 1968 1196 815 568 406
39844 15803 9832 5254 3166 1968 1196 815 568 406


Centre Meteorologique Canadien, Environnement Canada
Canadian Meteorological Center, Environment Canada
24 hours precipitation forecast verification against observation
SHEF network data for valid time 12z
00 to 24 hours forecast fm 12Z run only All of USA
71 cas hiv k4h1ra2b vs k4h1v4ha
k4h1ra2b0024
k4h1v4af0024
Bias
Precipitation class (mm)
Threat
Number of observation
171061 94725 72166 41935 26912 18170 12582 8915 6633 4949
171061 94725 72166 41935 26912 18170 12582 8915 6633 4949


Centre Meteorologique Canadien, Environnement Canada
Canadian Meteorological Center, Environment Canada
24 hours precipitation forecast verification against observation
Synoptic network data for valid time 00-12z
00 to 24 hours forecast North AMERICA
156 cas ete k4e1ra2b vs k4e1v4af
k4e1ra2b0024
k4e1v4af0024
Bias
Precipitation class (mm)
Threat
Number of observation
49603 27101 19157 11023 6914 4539 3098 2219 1579 1143
49603 27101 19157 11023 6914 4539 3098 2219 1579 1143

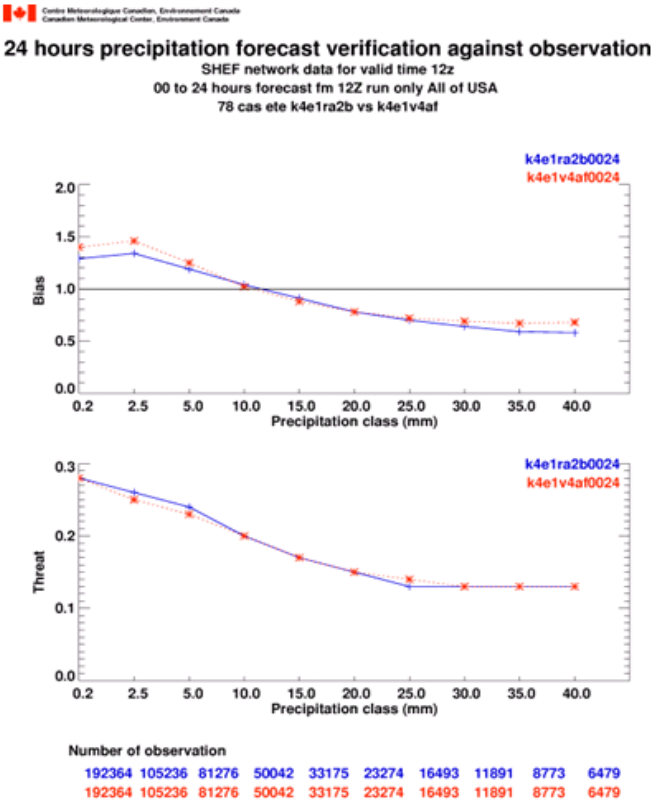
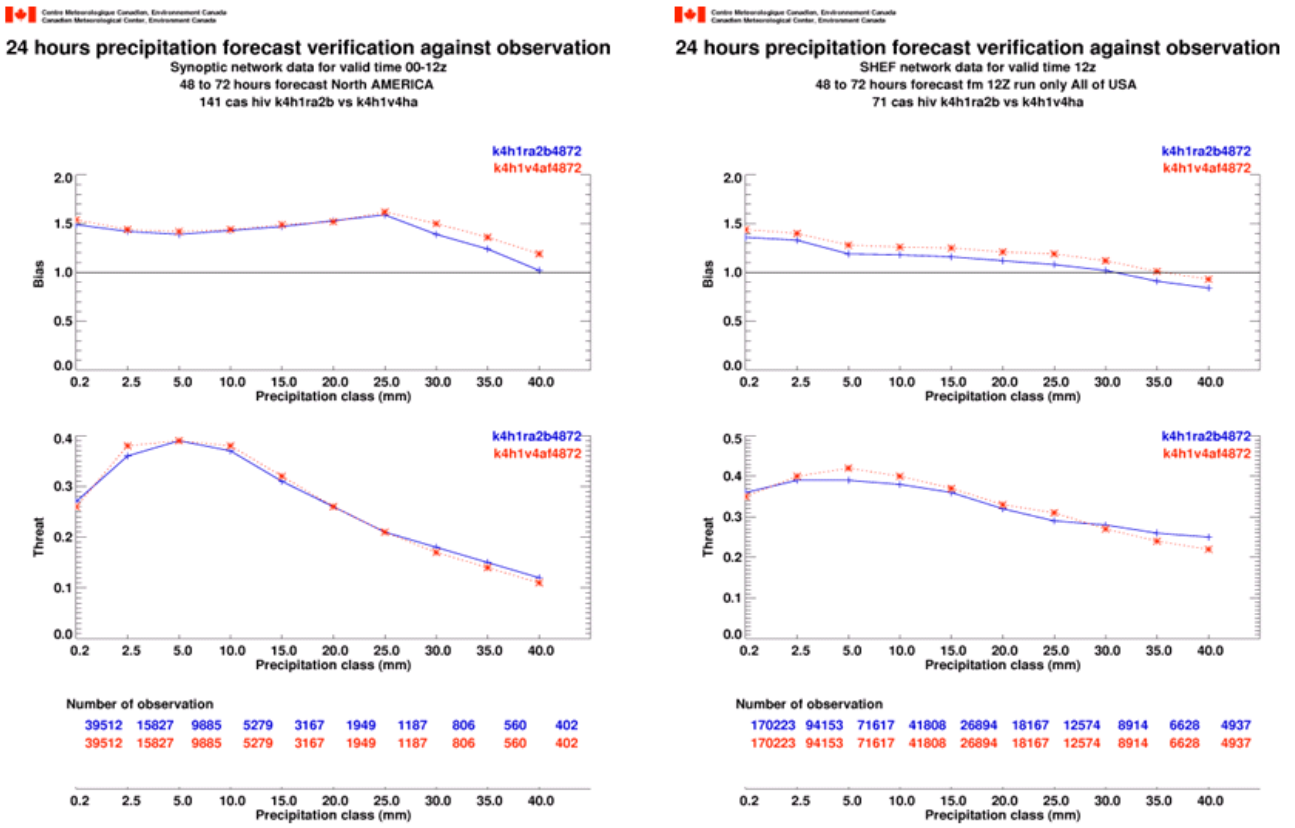

Centre Meteorologique Canadien, Environnement Canada
Canadian Meteorological Center, Environment Canada
24 hours precipitation forecast verification against observation
SHEF network data for valid time 12z
00 to 24 hours forecast fm 12Z run only All of USA
78 cas ete k4e1ra2b vs k4e1v4af
k4e1ra2b0024
k4e1v4af0024
Bias
Precipitation class (mm)
Threat
Number of observation
192364 105236 81276 50042 33175 23274 16493 11891 8773 6479
192364 105236 81276 50042 33175 23274 16493 11891 8773 6479

Figure 36: Precipitation scores for the 00-24h period: winter (top row) and summer cycles (bottom row), against SYNOP (left column) and SHEF network (right column). Control system in blue, new system in red.



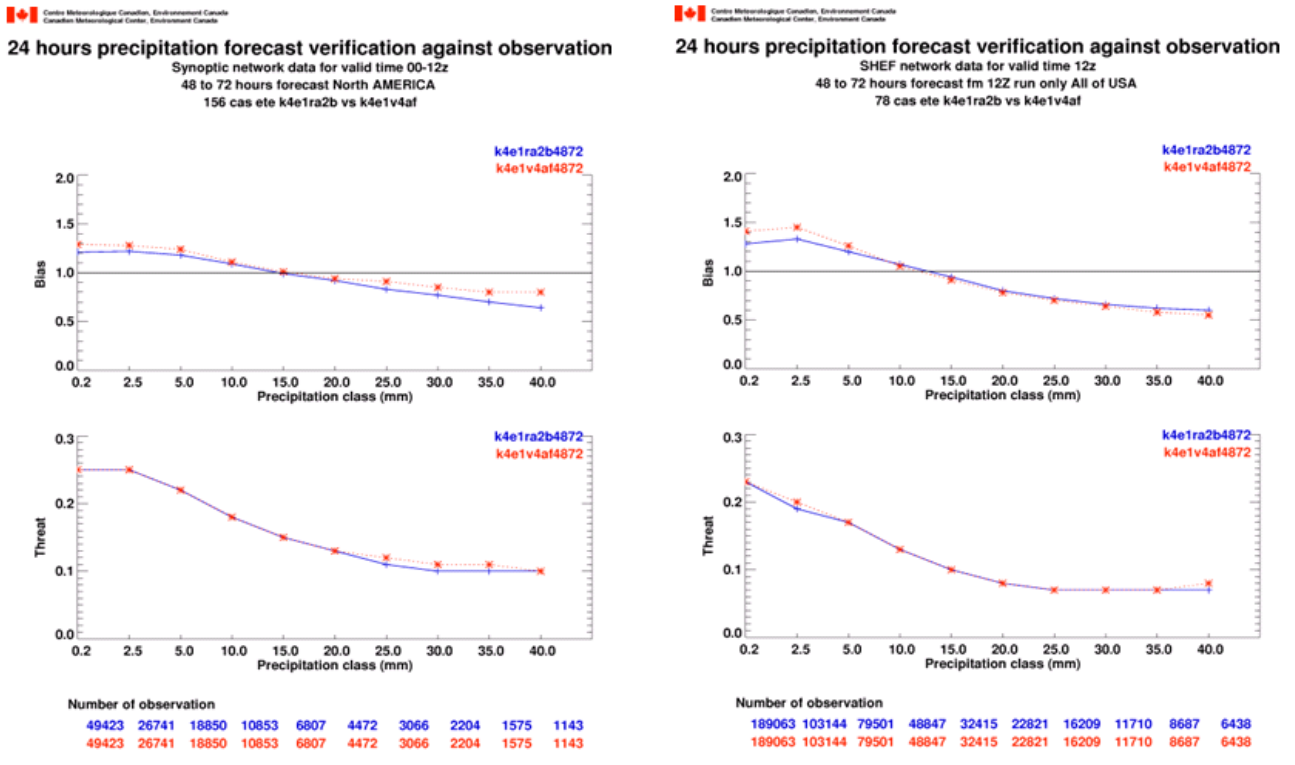


Figure 37: Precipitation scores for the 48-72h period: winter (top row) and summer cycles (bottom row), against SYNOP (left column) and SHEF network (right column). Control system in blue, new system in red.

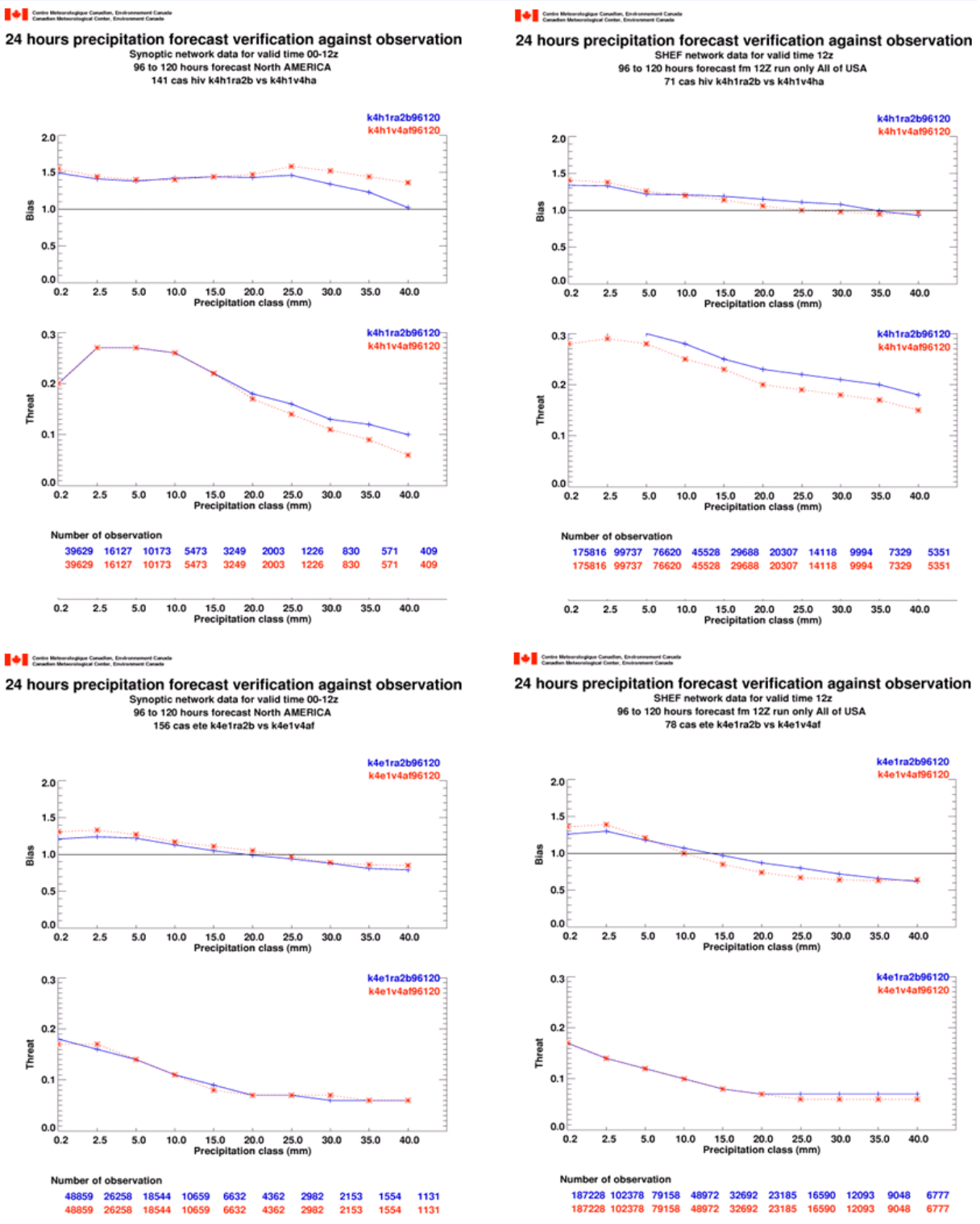
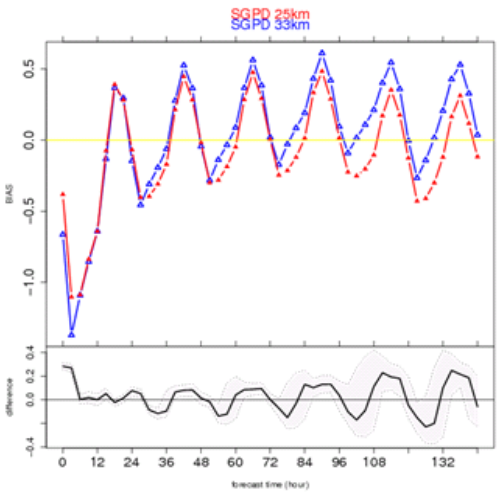
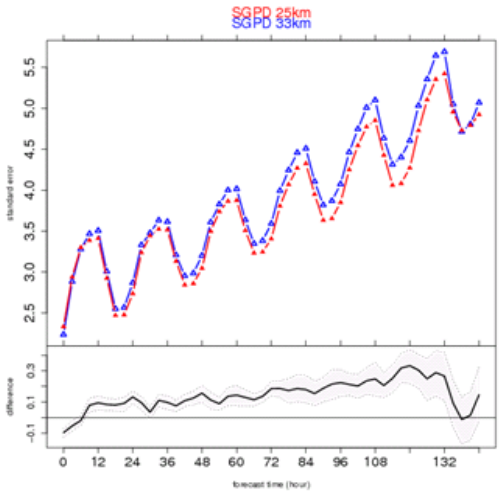


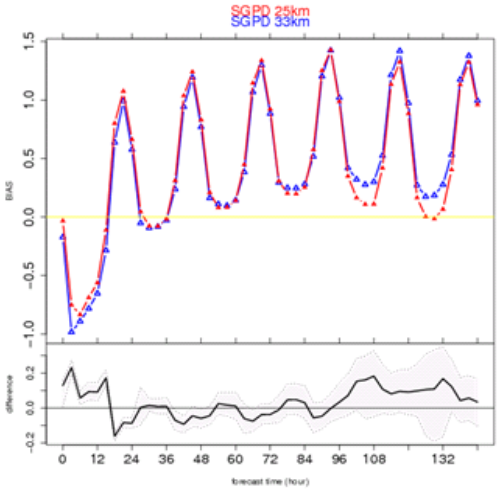
Figure 38: Precipitation scores for the 96-120h period: winter (top row) and summer cycles (bottom row), against SYNOP (left column) and SHEF network (right column). Control system in blue, new system in red.



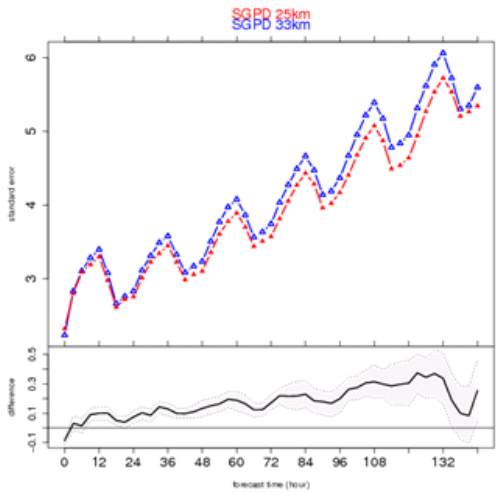
T 2m Bias Winter



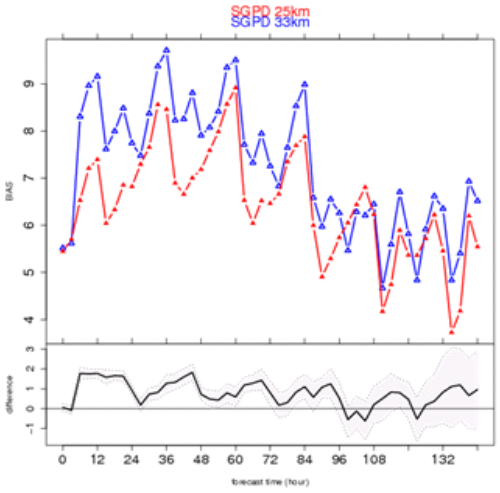
T 2m Standev Winter



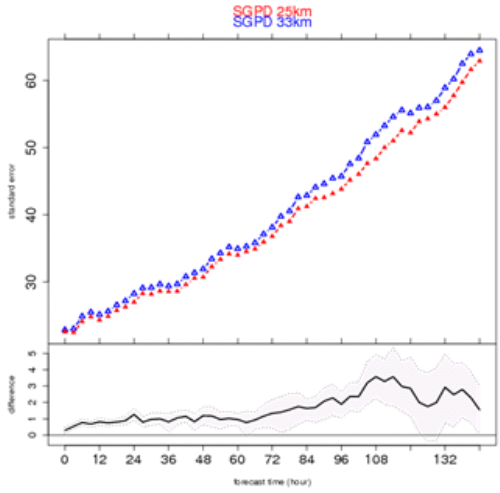
T-Td 2m Bias Winter



T-Td 2m Standev Winter



Wind 10m Bias Winter



Wind 10m Standev Winter

Figure 39: Surface (USTAT) scores for the winter cycle over North America: bias (left) and standard deviation (right) for 2m temperature (top), 2m dew-point depression (middle) and 10m wind speed (bottom). Results for the control system in blue, the new system in red.

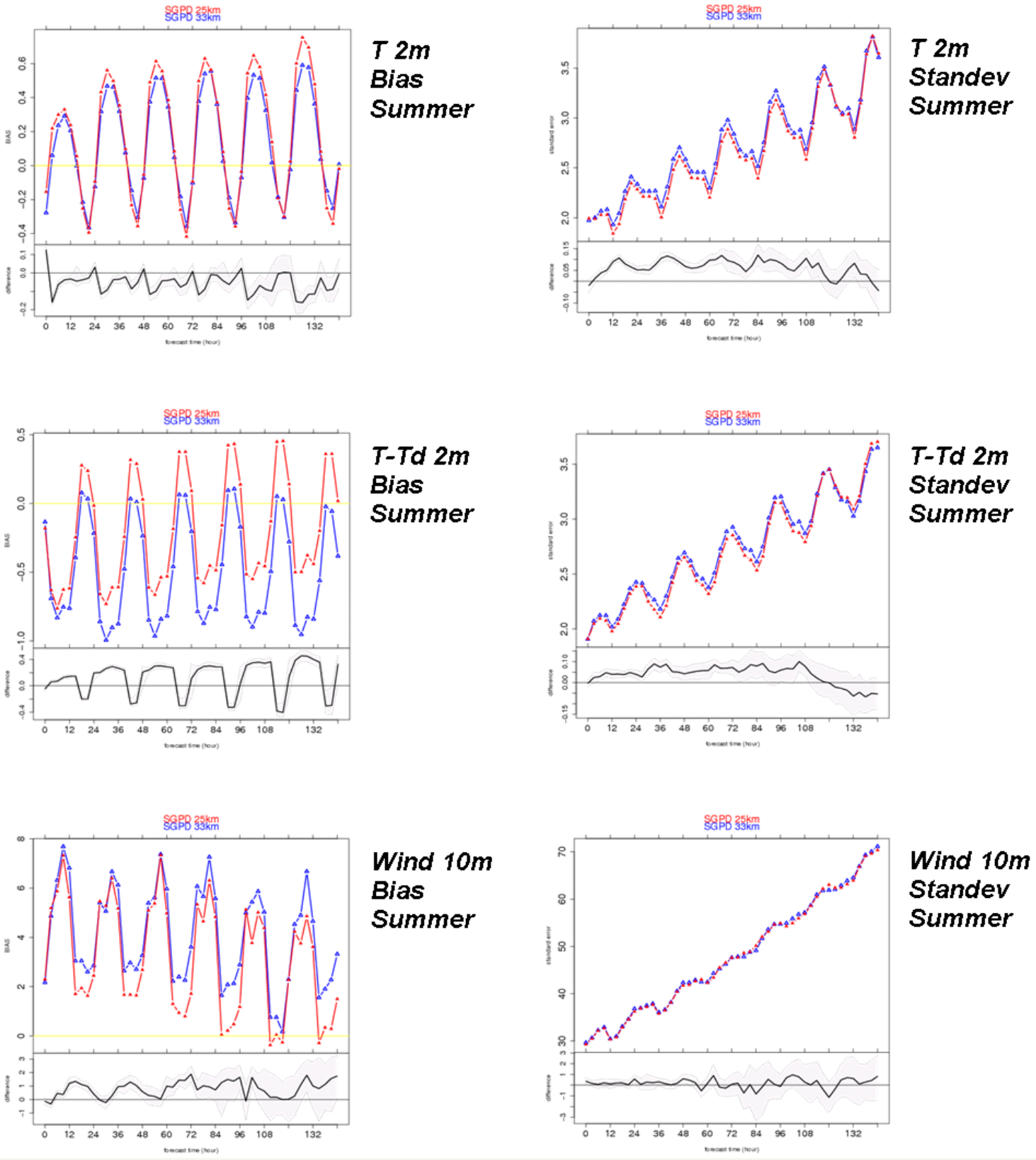


Figure 40: Surface (USTAT) scores for the summer cycle over North America: bias (left) and standard deviation (right) for 2m temperature (top), 2m dew-point depression (middle) and 10m wind speed (bottom). Results for the control system in blue, the new system in red.

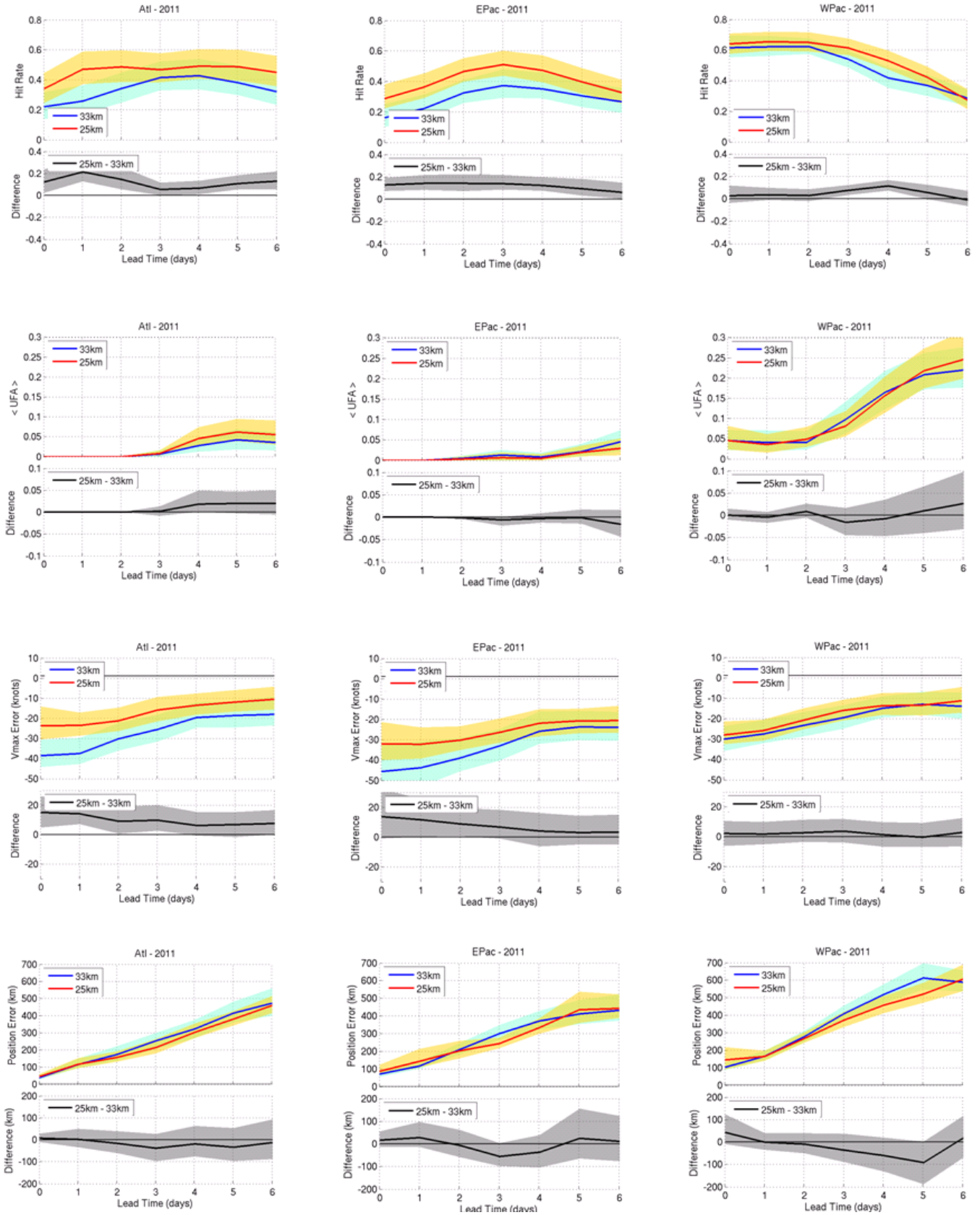


Figure 41: Tropical cyclone verification for the summer cycle: hit rate (top row), average number of false alarms (2nd row), intensity errors based on maximum wind speed (3rd row) and track position errors (bottom row) over the Atlantic (left column), East Pacific

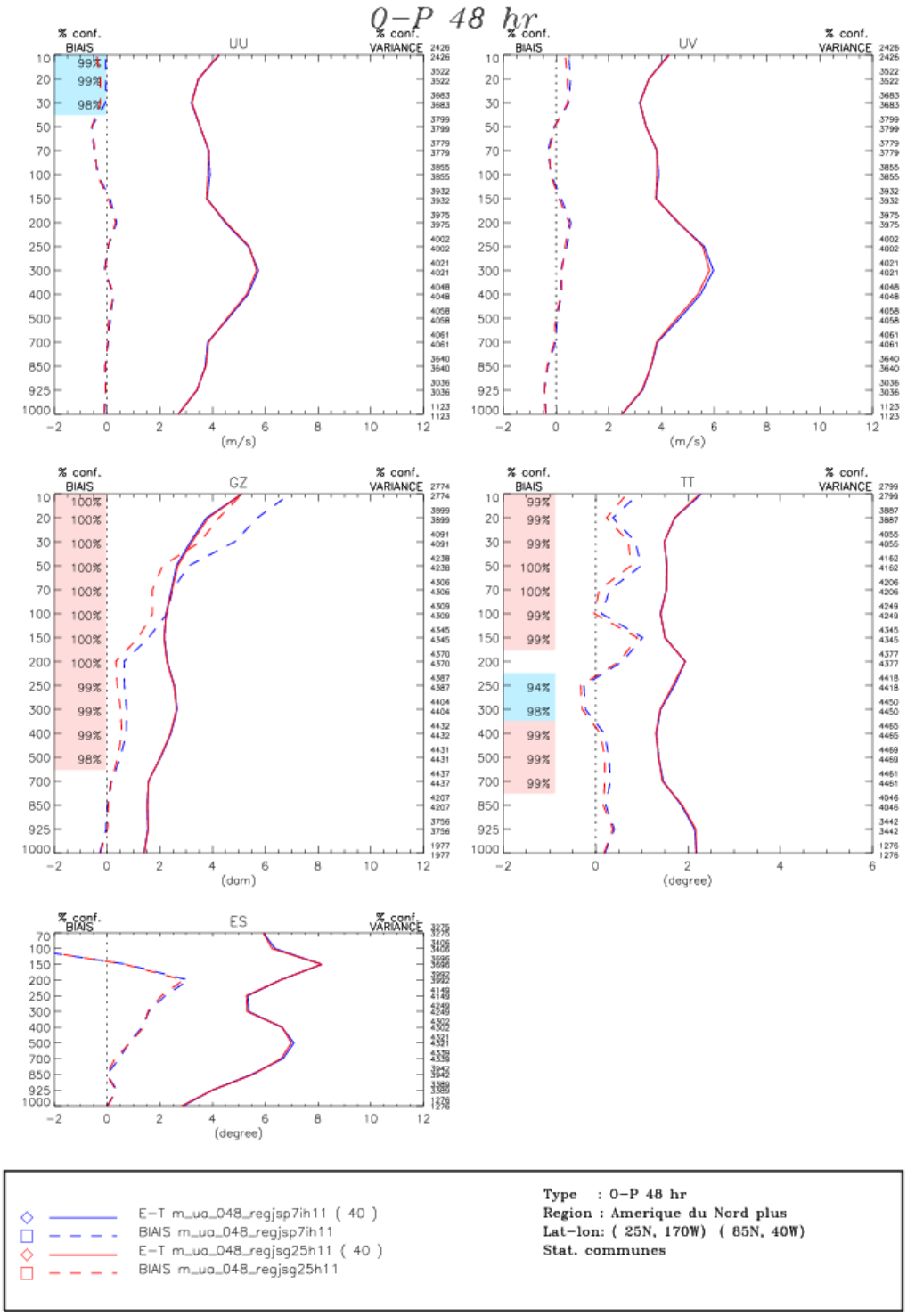


Figure 42: Time-mean upper-air forecast verification scores against radiosondes for the LAM component of the RDPS over North America at a lead time of 48 hours, comparing two development experiments: (blue) an experiment very similar to the operational system and (red) an experiment which obtained its GDPS analyses and bias correction coefficients from the GDPS-3.0.0 final cycle. The period was Jan. 22 - Mar. 30, 2011 every 36 hours, giving a total of 40 cases. The dashed (continuous) lines correspond to the bias (standard deviation).

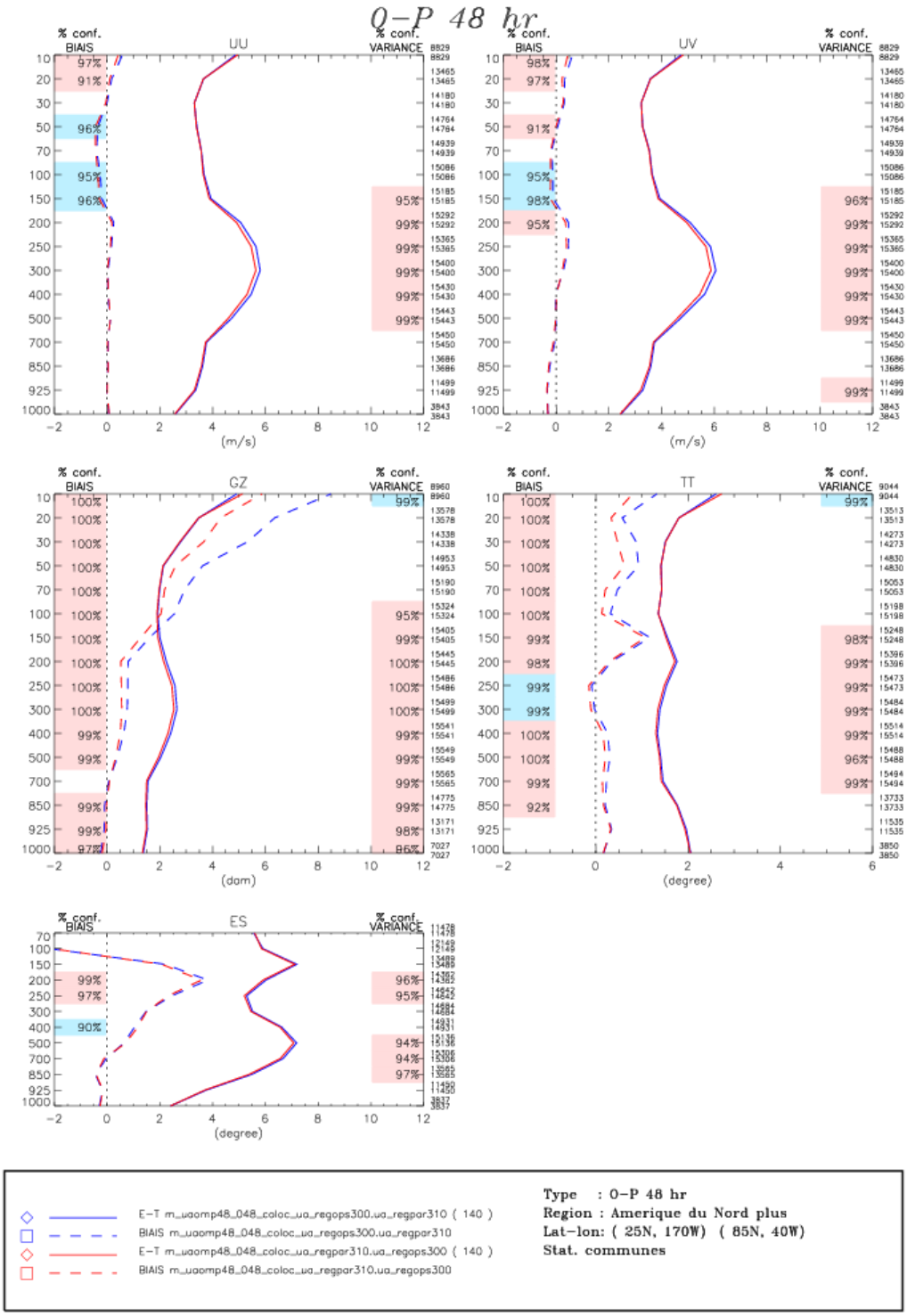


Figure 43: Time-mean upper-air forecast verification scores against radiosondes for the LAM component of the RDPS over North America at a lead time of 48 hours, comparing the (blue) operational and (red) parallel runs. The period considered here is Nov. 14, 2012 - Jan. 14, 2013 every 12 hours, giving a total of 140 cases. Besides using the new GDPS analyses and bias correction coefficients, the parallel run also included new observation types during the second half of this period. The dashed (continuous) lines correspond to the bias (standard deviation).

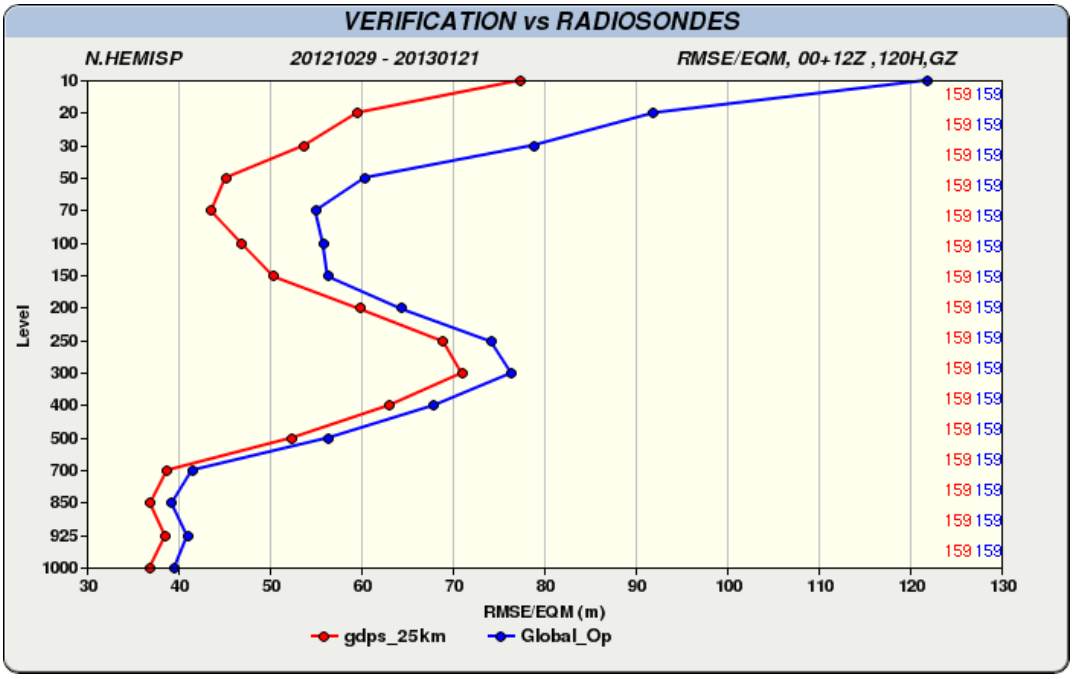


Figure 44: RMS (root mean square) errors of geopotential height, in metres, for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for 120-hour forecasts at all model levels, against the North Hemisphere radiosonde observation network, for the period October 29, 2012 – January 21, 2013 (159 cases).

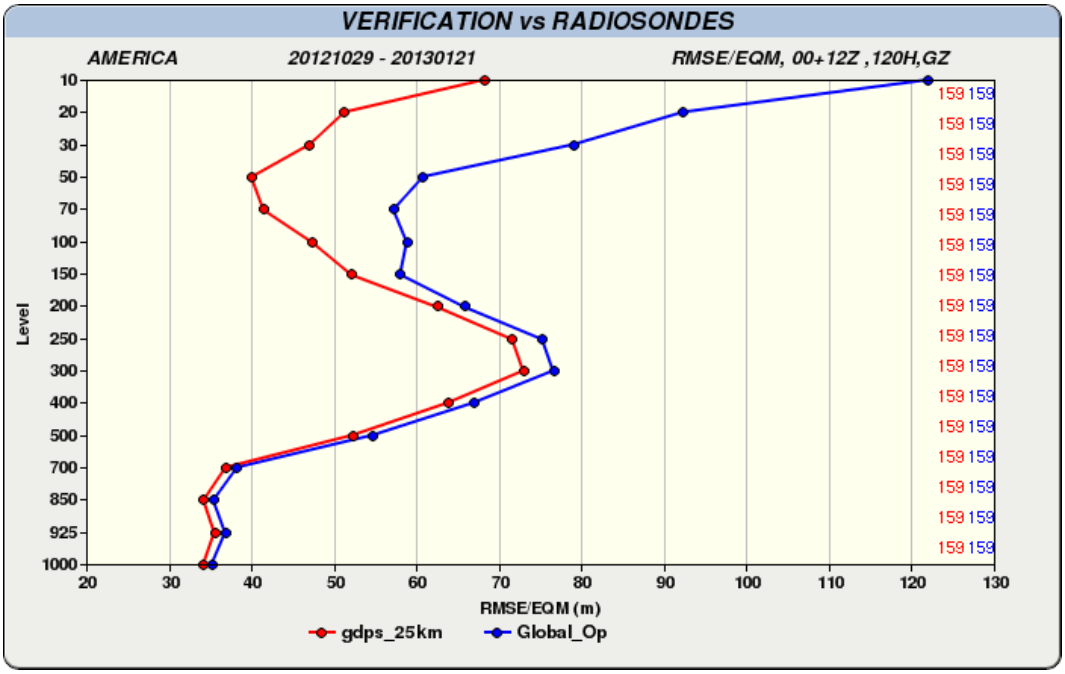


Figure 45: RMS (root mean square) errors of geopotential height, in metres, for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for 120-hour forecasts at all model levels, against the North American radiosondes observation network, for the period October 29, 2012 – January 21, 2013 (159 cases).

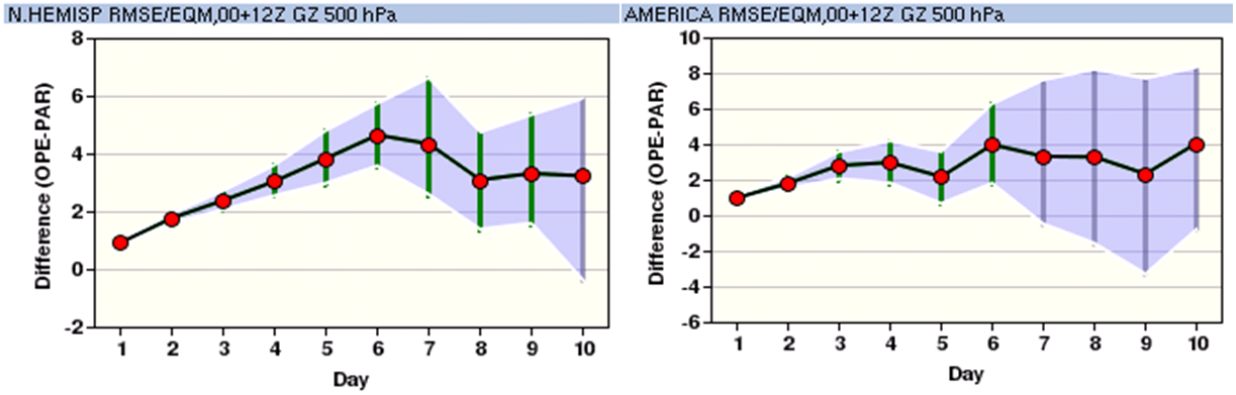


Figure 46: Confidence intervals (5-95 %) calculated on the difference between the operational GDPS and parallel run GDPS-3.0.0 models, for 500 hPa geopotential height forecasts from day 1 to day 10, against the North Hemisphere (left) and North America (right) radiosonde networks. Green bars refer to a statistically significant improvement of the GDPS 300 over the operational GDPS model, while grey bars indicate the difference is not statistical significant.

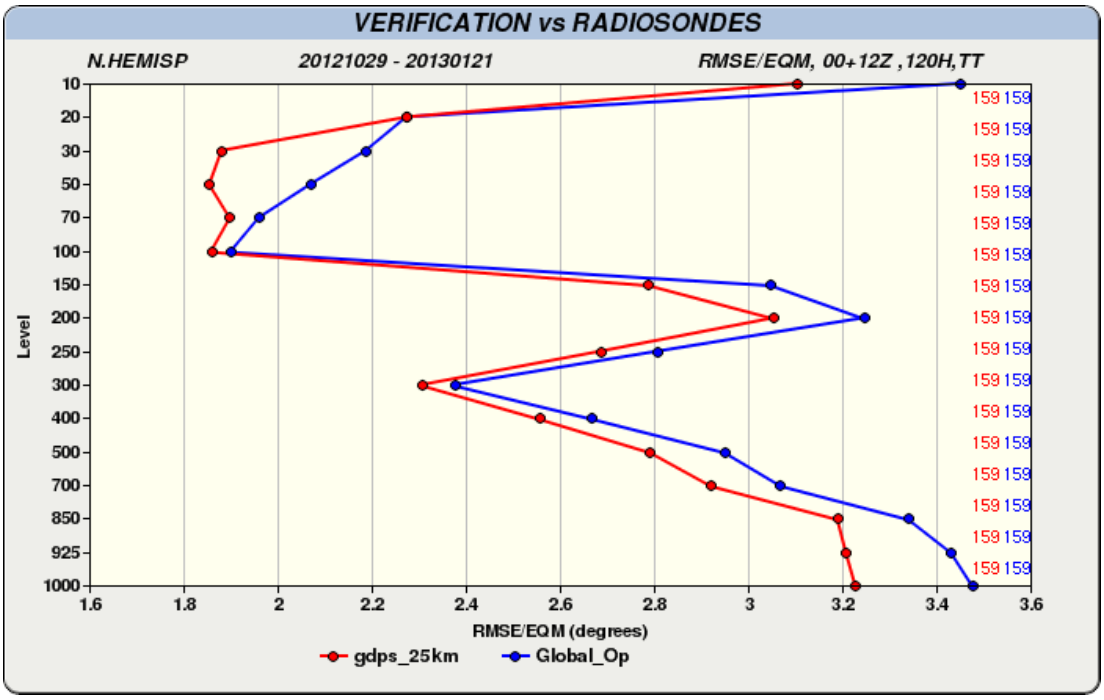


Figure 47: RMS (root mean square) errors of model temperature forecasts, for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for 120-hour forecasts at all model levels, against the North Hemisphere radiosonde observation network, for the period October 29, 2012 – January 21, 2013 (159 cases).

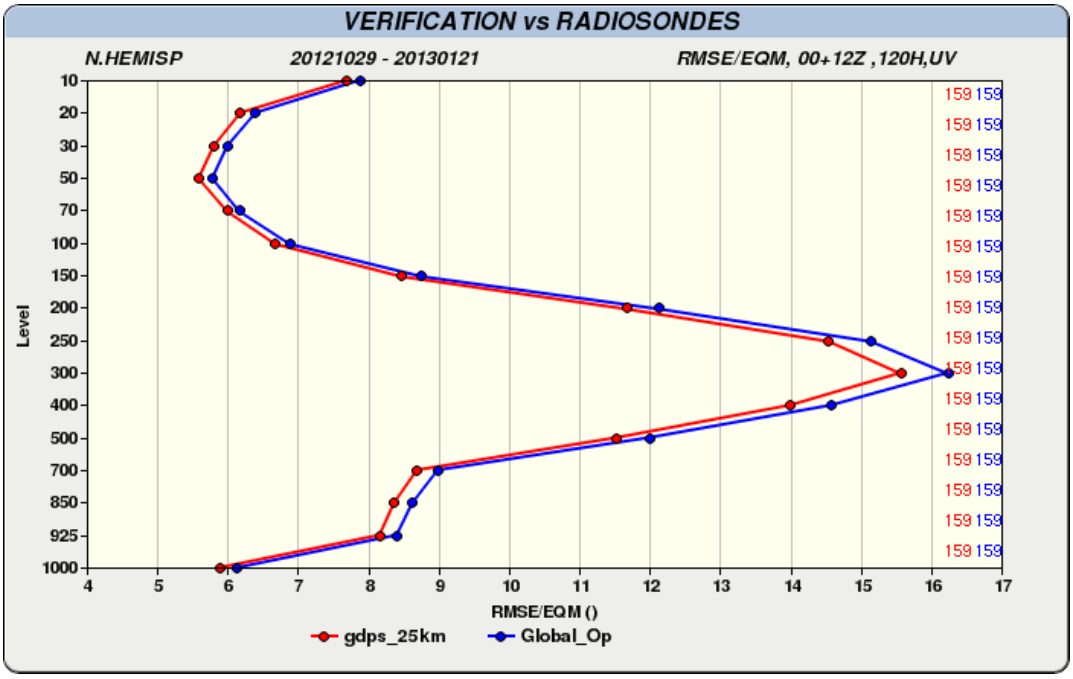


Figure 48: RMS (root mean square) errors of model wind forecasts, for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for 120-hour forecasts at all model levels, against the North Hemisphere radiosonde observation network, for the period October 29, 2012 – January 21, 2013 (159 cases).

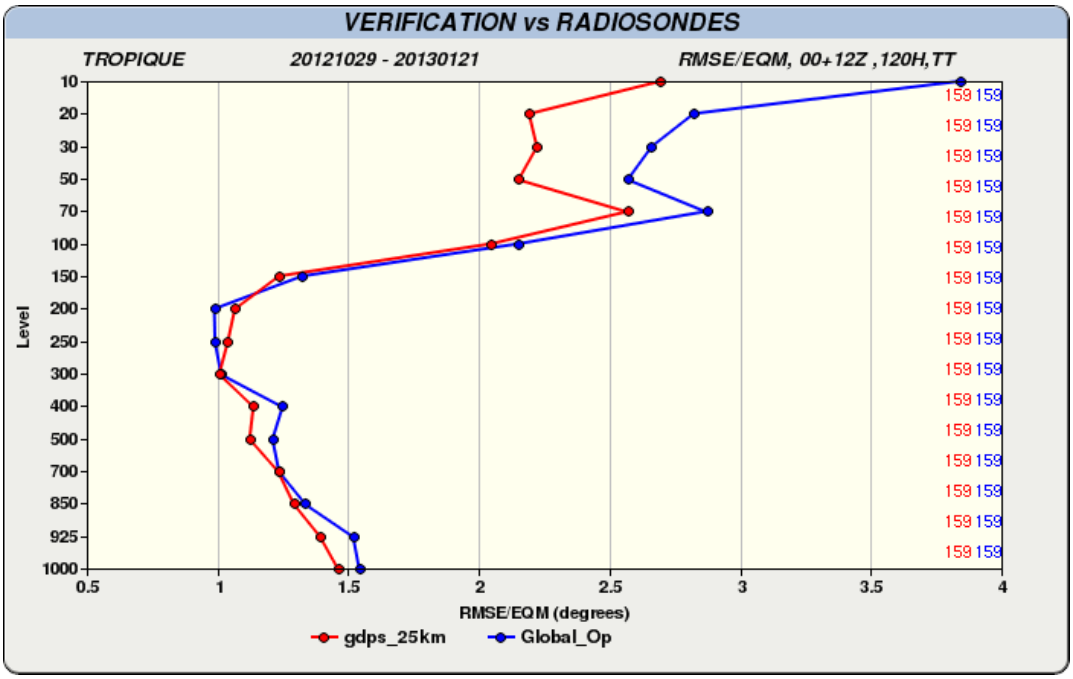


Figure 49: RMS (root mean square) errors of model temperature forecasts, for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for 120-hour forecasts at all model levels, against the Tropics radiosonde observation network, for the period October 29, 2012 – January 21, 2013 (159 cases).

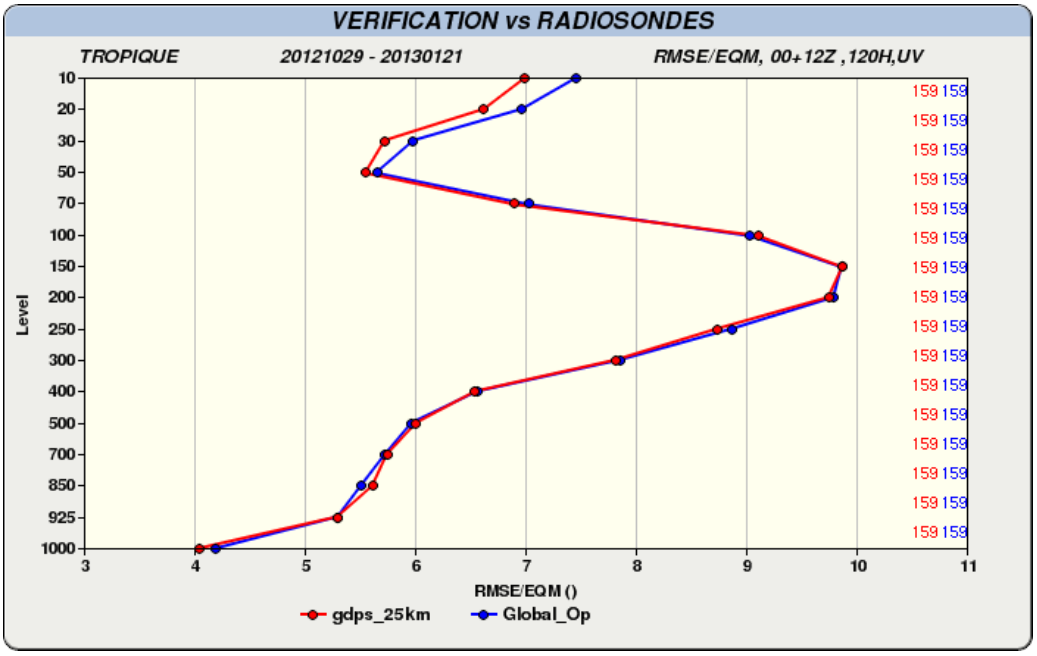


Figure 50: RMS (root mean square) errors of model wind forecasts, for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for 120-hour forecasts at all model levels, against the Tropics radiosonde observation network, for the period October 29, 2012 – January 21, 2013 (159 cases).

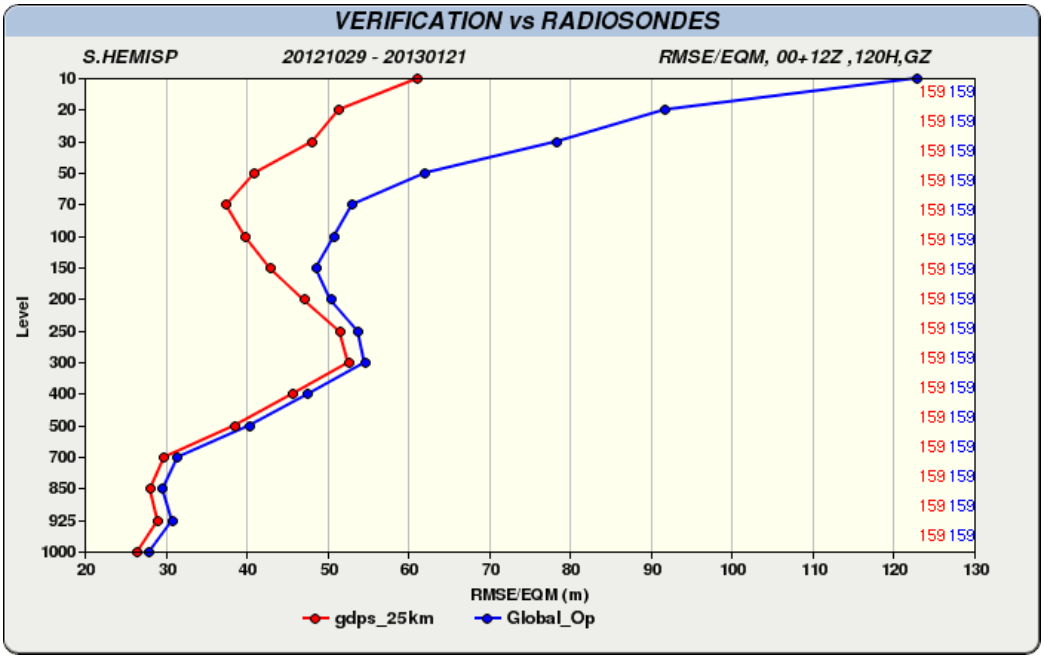


Figure 51: RMS (root mean square) errors of geopotential height, in metres, for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for 120-hour forecasts at all model levels, against the South Hemisphere radiosonde observation network, for the period October 29, 2012 – January 21, 2013 (159 cases).

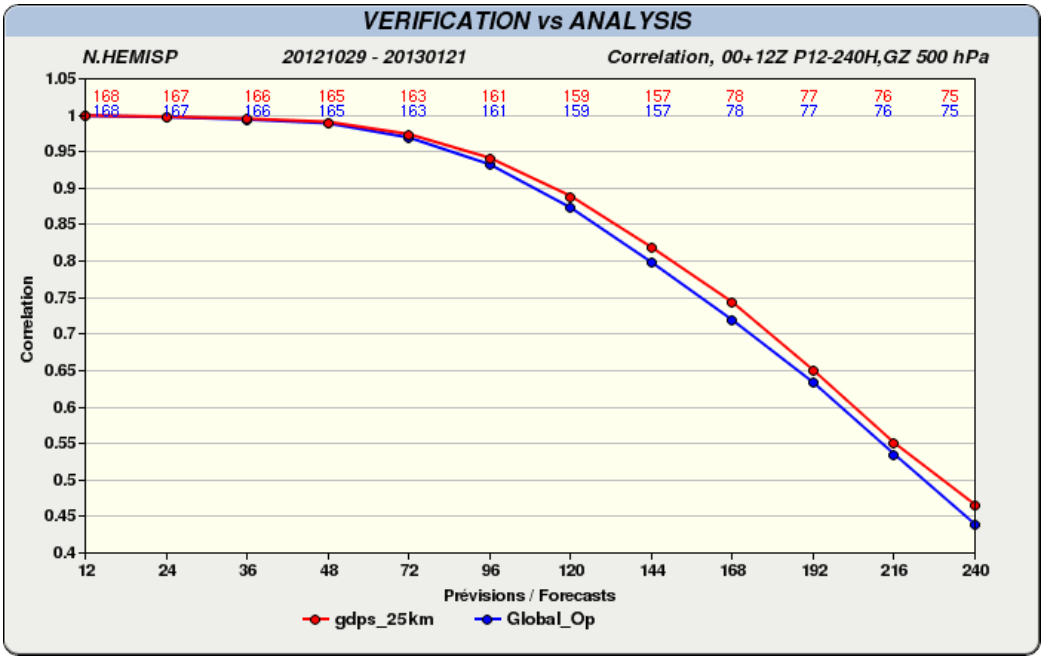


Figure 52: Anomaly correlation of 500 hPa geopotential heights for the operational GDPS (blue) and parallel run GDPS-3.0.0 (red) forecasts of 12-240 hours over the North Hemisphere. Scores are calculated using each model's own analysis.

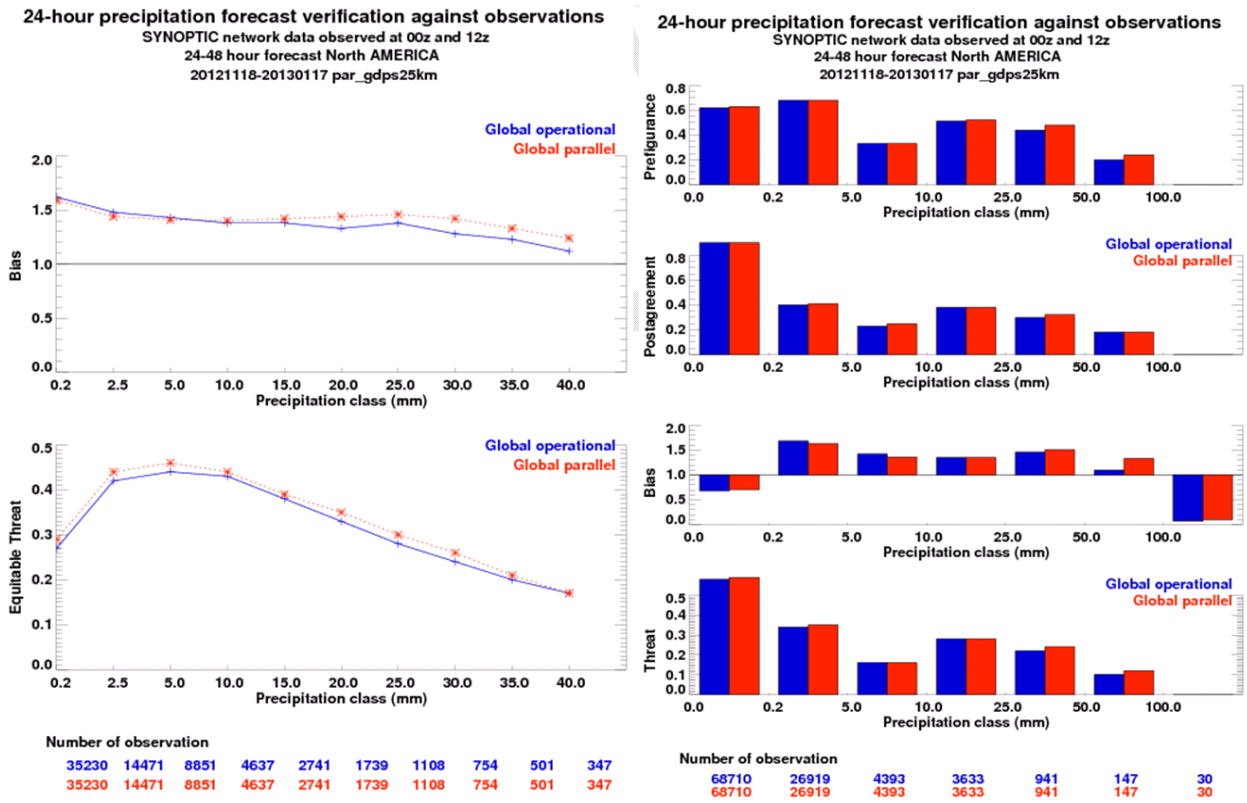


Figure 53: Equitable Threat Scores and bias for the GDPS operational (blue) and GDPS 300 parallel run (red) models for the forecast period 24-48 hours, scored against the North American synoptic observation network. The graph on the left is a threshold diagram, wherein scores are calculated for all categories above the given threshold.

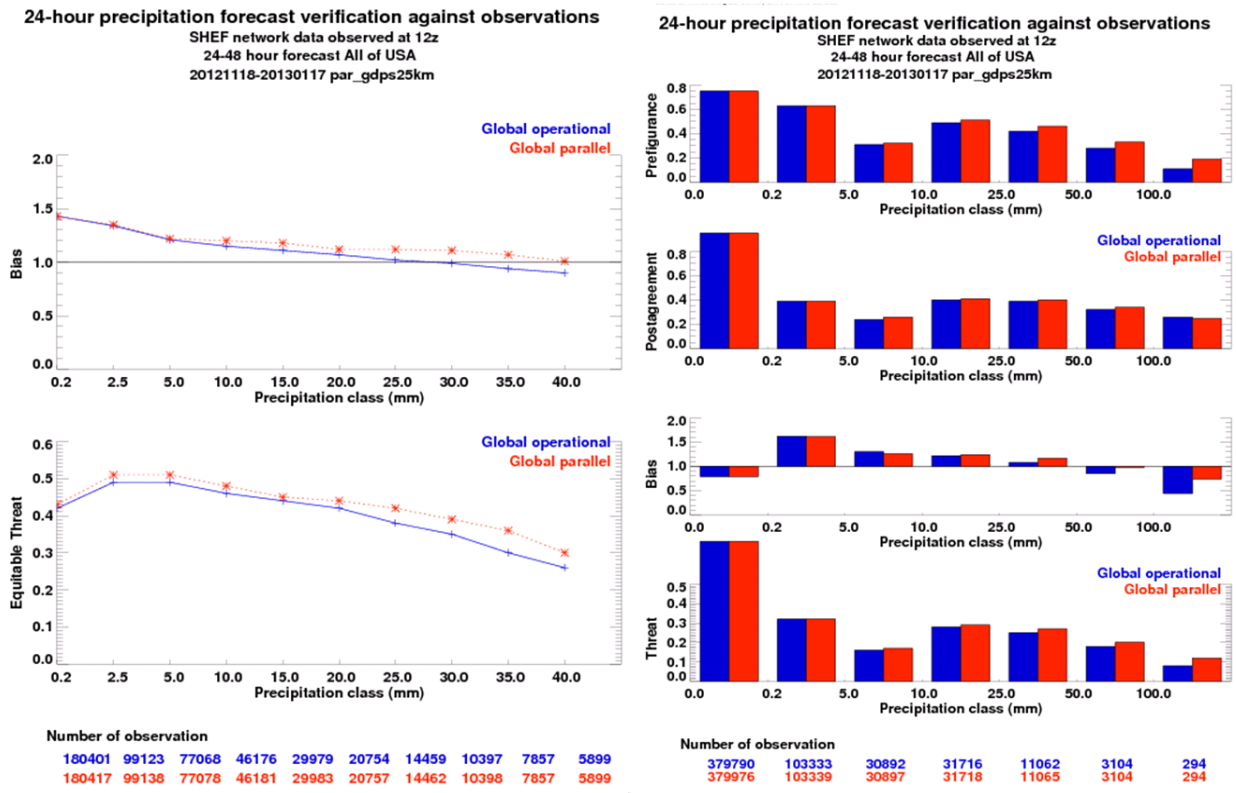


Figure 54: Equitable Threat Scores and bias for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for the forecast period 24-48 hours, scored against the U.S. SHEF observation network. The graph on the left is a threshold diagram, wherein scores are calculated for all categories above the given threshold.

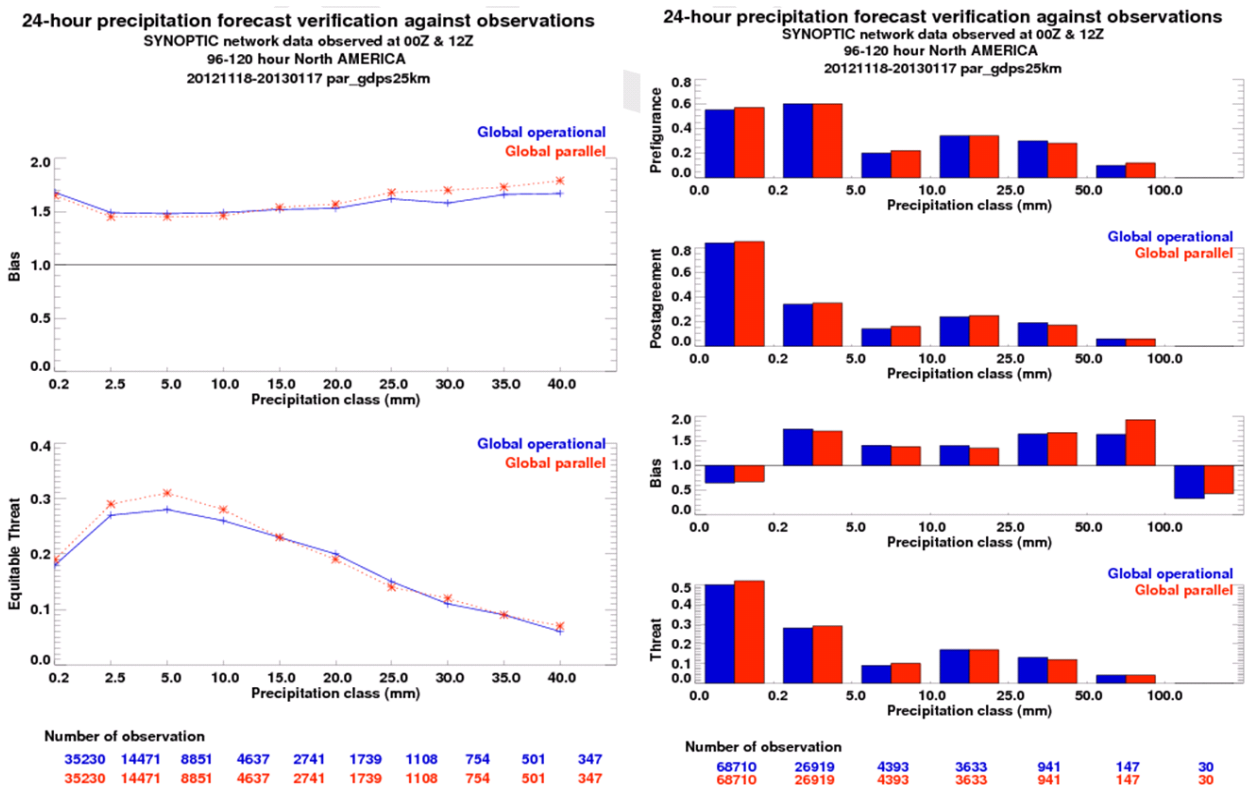


Figure 55: Equitable Threat Scores and bias for the GDPS operational (blue) and GDPS-3.0.0 parallel run (red) models for the forecast period 96-120 hours, scored against the North American synoptic observation network. The graph on the left is a threshold diagram, wherein scores are calculated for all categories above the given threshold.

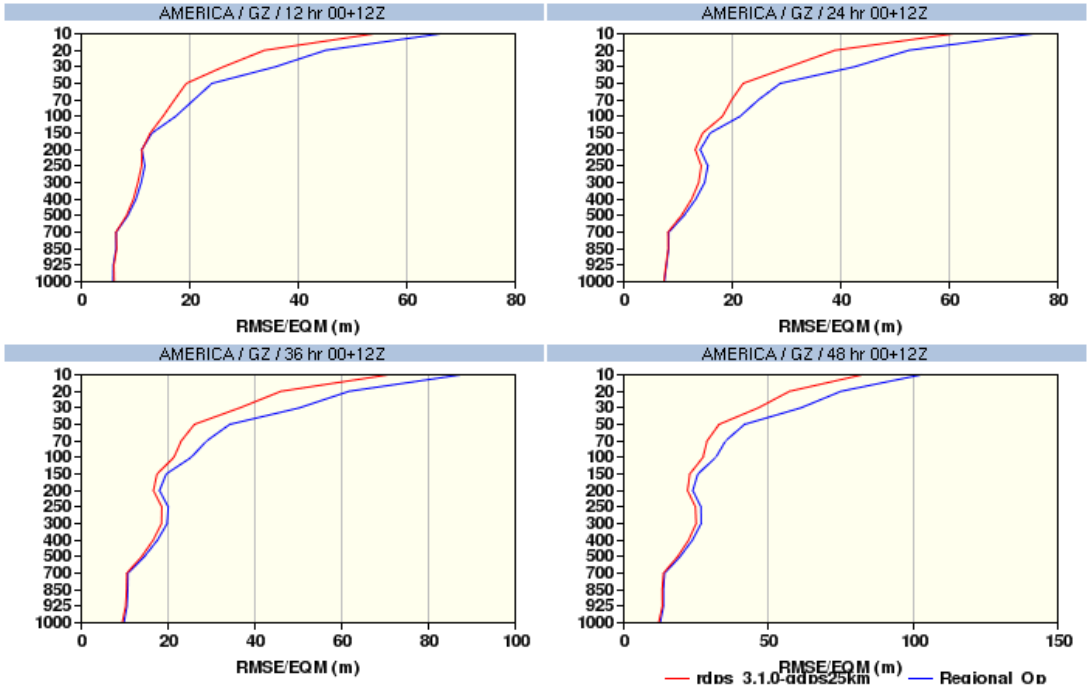


Figure 56: RMS (root mean square) errors of geopotential height forecasts for the RDPS operational (blue) and RDPS 3.1.0 parallel run (red) models for forecast leads times of 12 to 48 hours, against the North American radiosonde observation network, for the period November 14, 2012 – January 21, 2013 (136 cases).

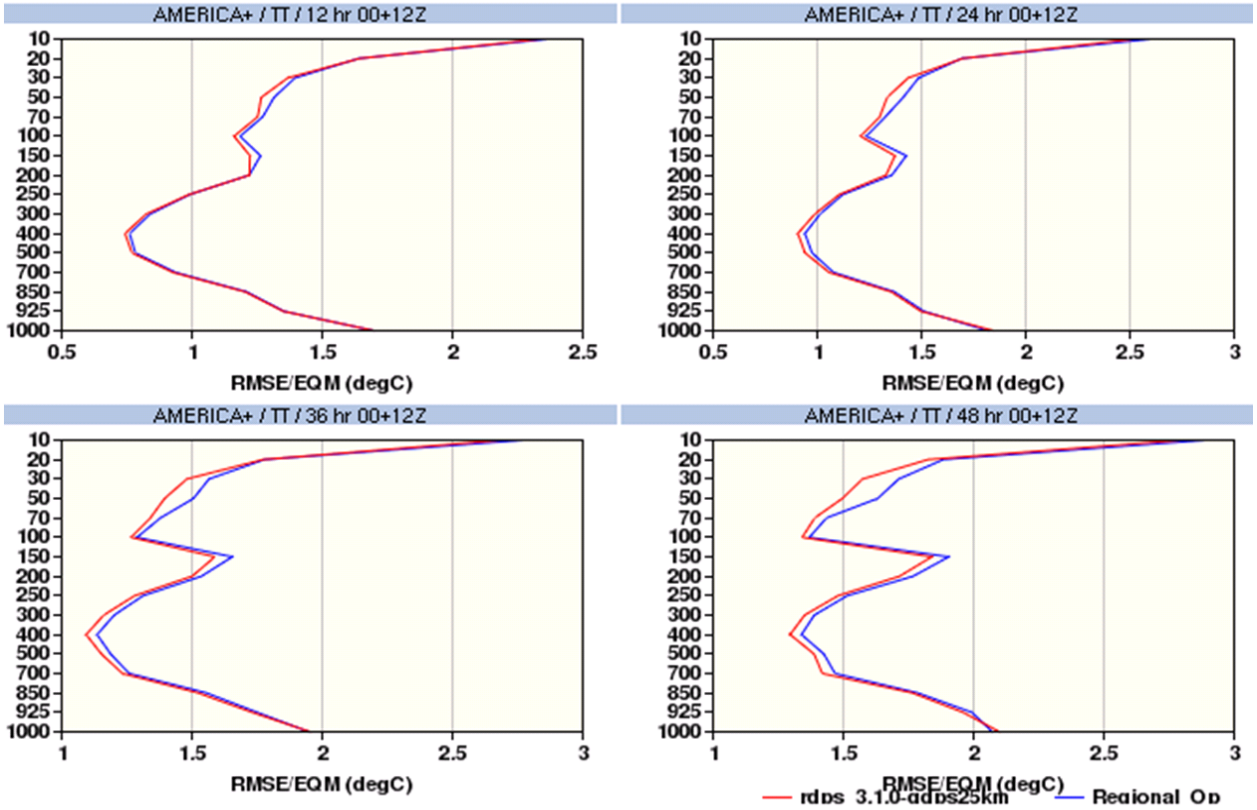


Figure 57: RMS (root mean square) errors of temperatures forecasts for the RDPS operational (blue) and RDPS-3.1.0 parallel run (red) models for forecast leads times of 12 to 48 hours, against the North American radiosonde observation network, for the period November 14, 2012 – January 21, 2013 (136 cases).

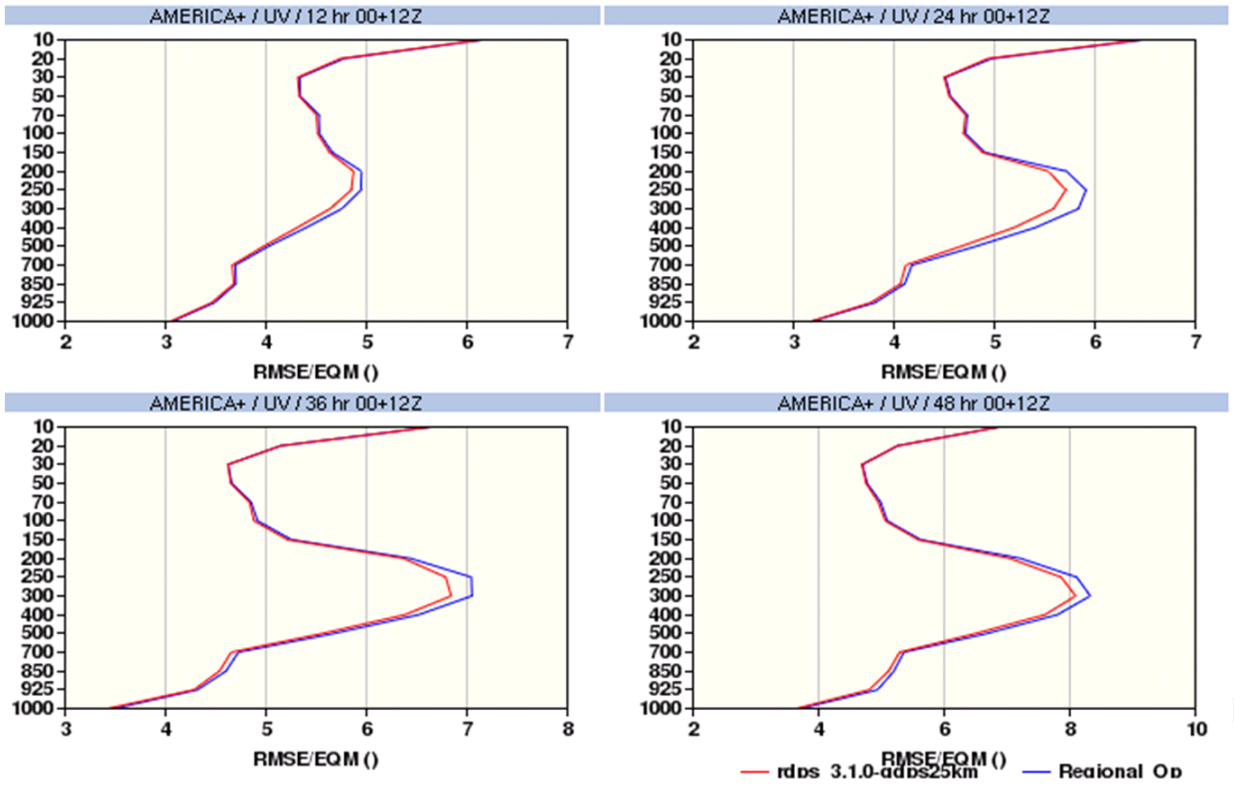


Figure 58: RMS (root mean square) errors of winds forecasts for the RDPS operational (blue) and RDPS-3.1.0 parallel run (red) models for forecast leads times of 12 to 48 hours, against the North American radiosonde observation network, for the period November 14, 2012 – January 21, 2013 (136 cases).

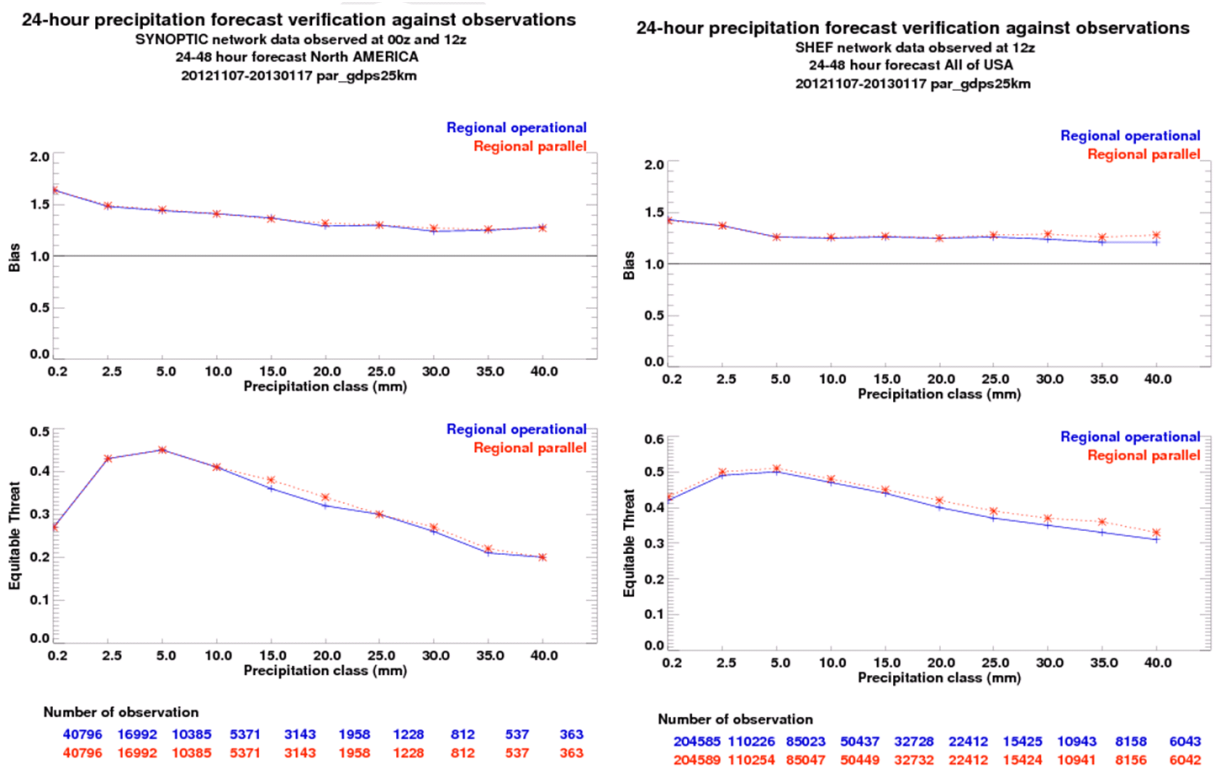


Figure 59: Equitable Threat Scores and bias for the RDPS operational (blue) and RDPS 310 parallel run (red) models for the forecast period 24-48 hours, scored against the North American synoptic observation network (left) and U.S. SHEF network (right).

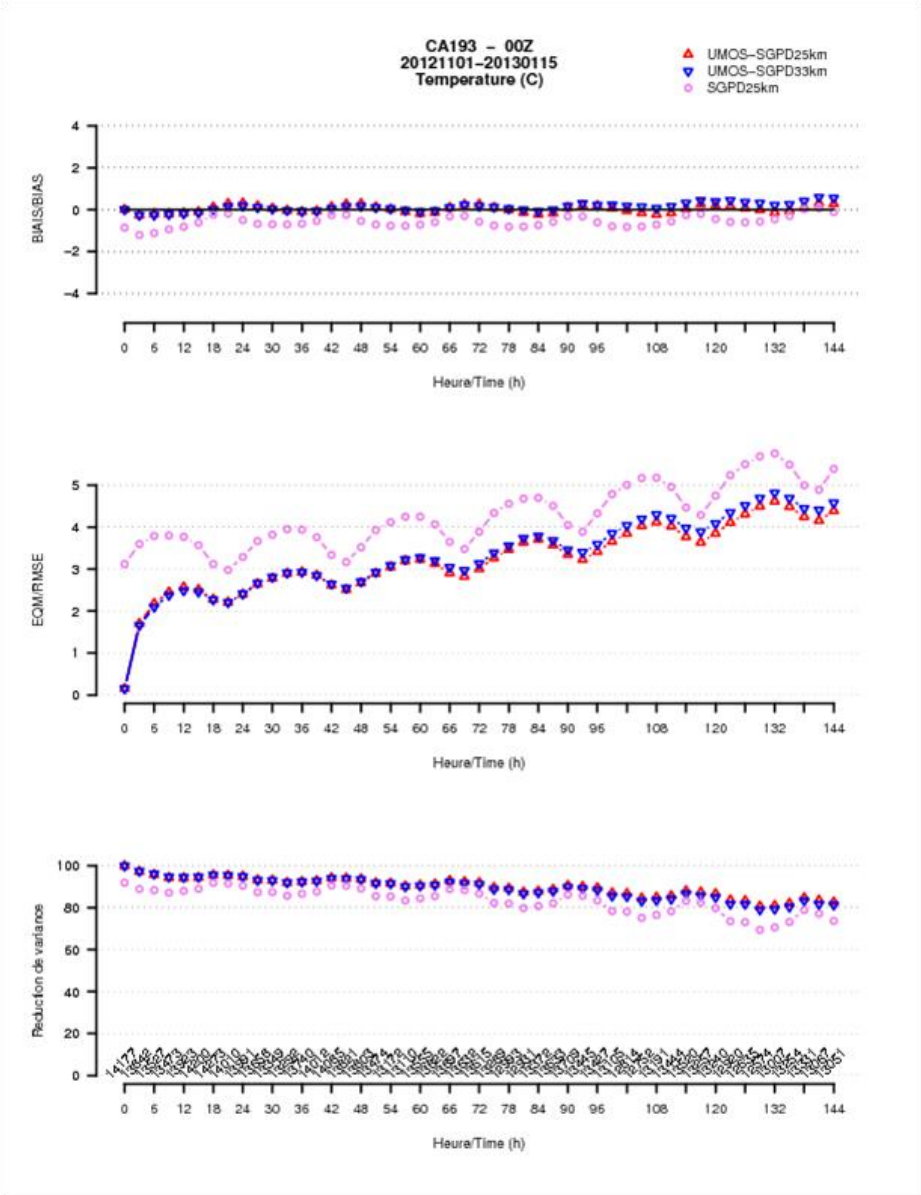


Figure 60: Temperature verification scores (BIAS, RMSE and Reduction of Variance), averaged over 193 Canadian stations, that clearly demonstrate the performance of both the parallel and the operational version of UMOS over the model.

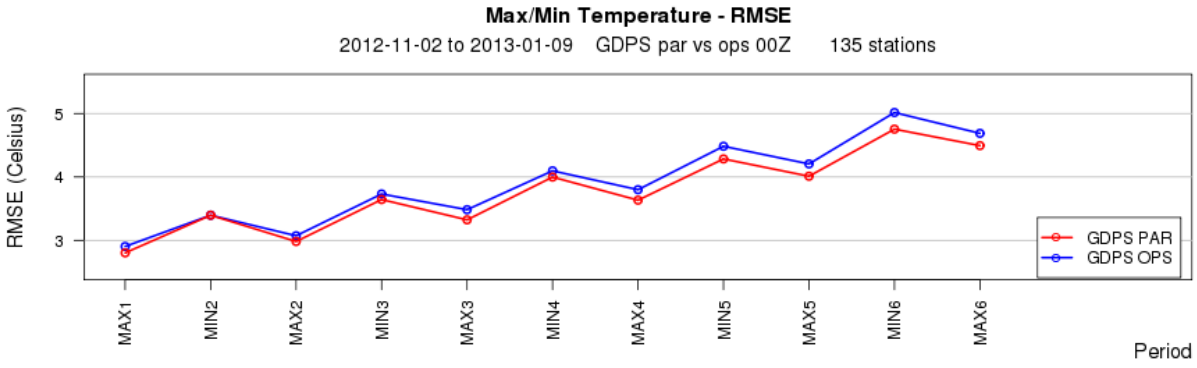


Figure 61: Scribe maximum and minimum temperature forecasts RMSE

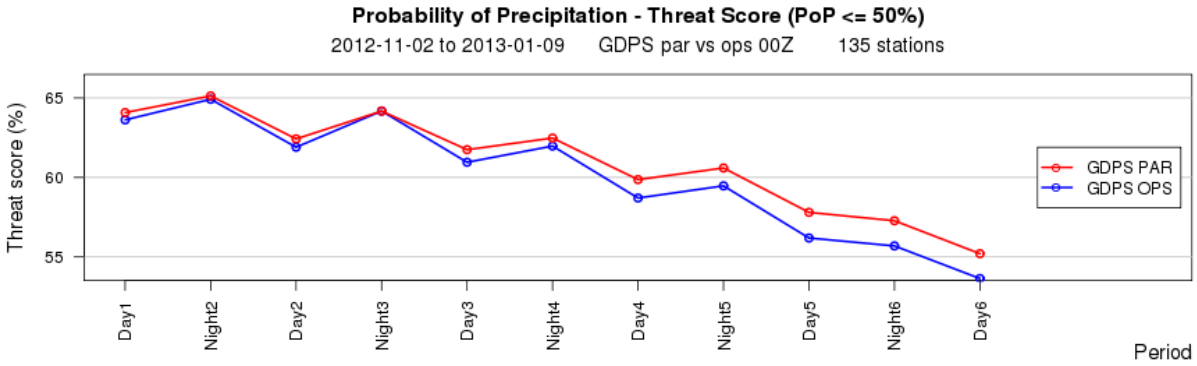


Figure 62: Scribe probability of precipitation forecasts threat score

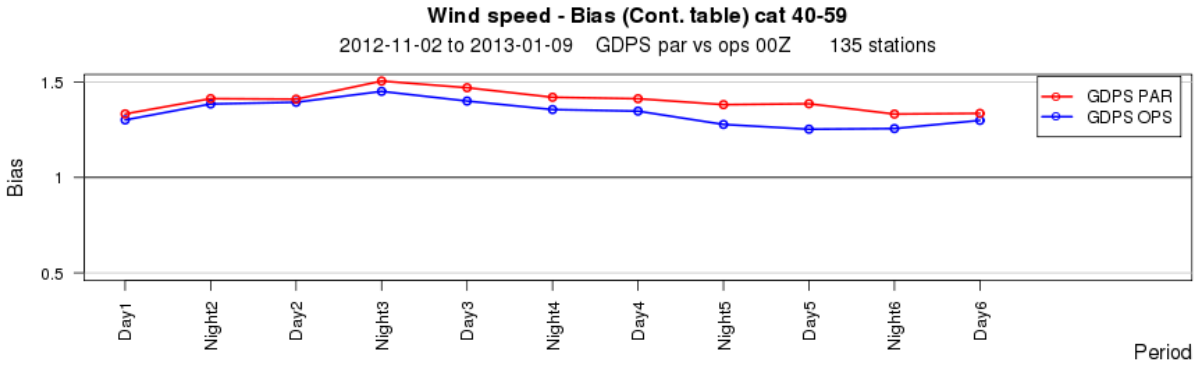


Figure 63: Scribe wind speed forecasts bias

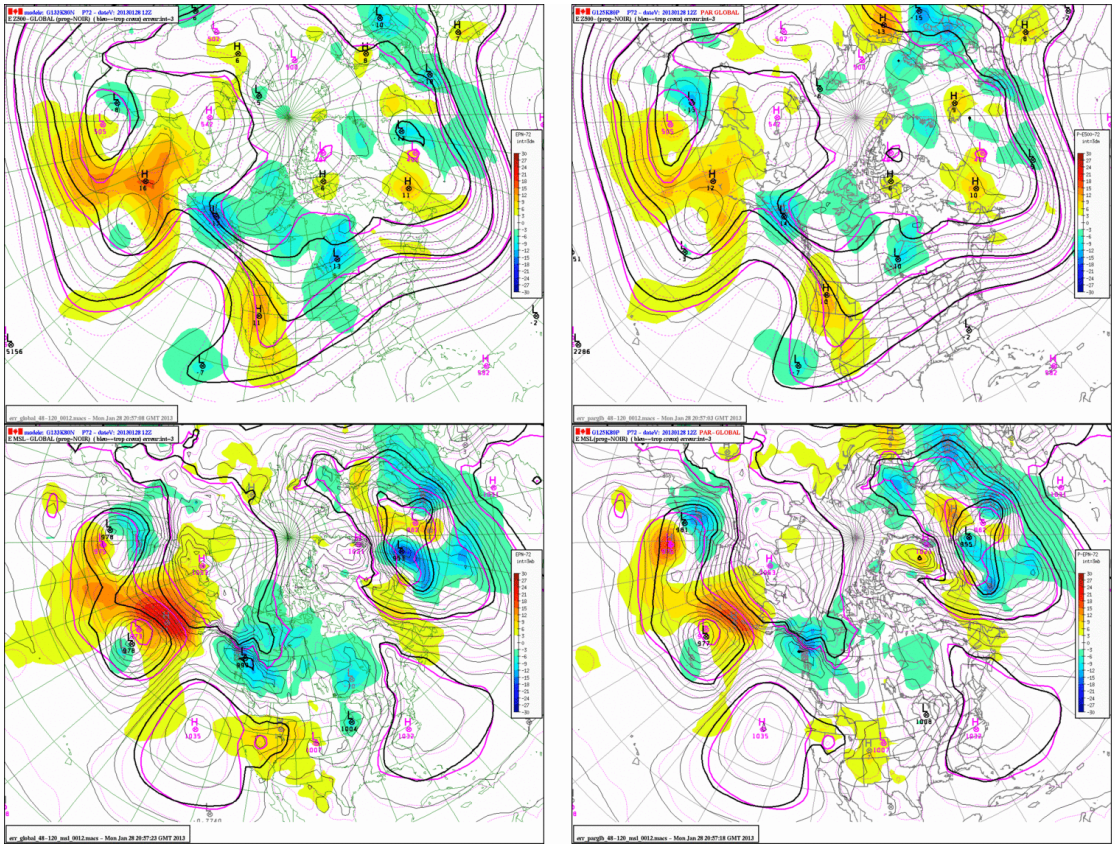


Figure 64: Difference between the 72h forecast from one system and its respective analysis, valid 28 January 2013 at 1200 UTC. The operational system is at left, parallel (GDPS-3.0.0) at right, GZ500 at top, and MSLP at bottom. Positive errors are in yellowish shades, negative errors in bluish shades.

a) b)

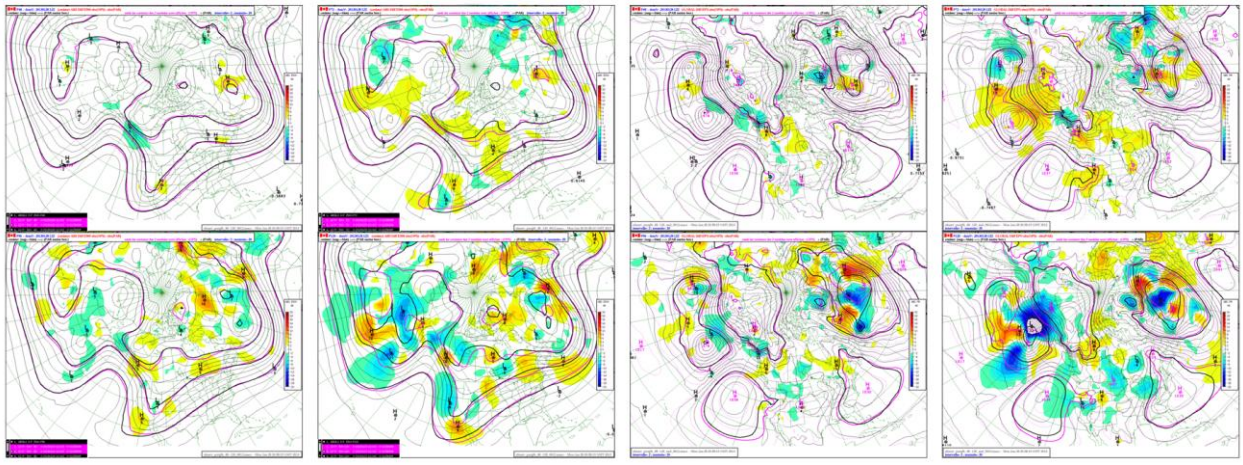


Figure 65: Difference between the absolute values of the errors (as seen in Figure 64) of the two systems, valid 28 January 2013 at 1200 UTC, for a) geopotential heights at 500 hPa and b) mean sea level pressure, for lead times of 48, 72, 96 and 120 hours from upper left to lower right corner in each four panel composite. Yellowish shades indicate greater errors from the operational system, bluish shades larger errors from the parallel (GDPS-3.0.0) system.

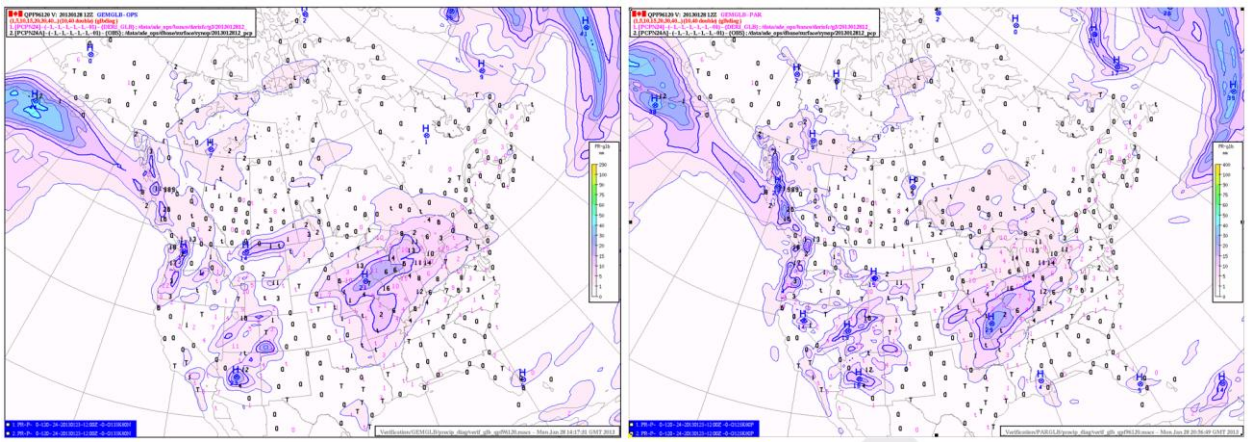


Figure 66: Quantitative precipitation forecast for the period 72 to 96 hours from the operational GDPS, left, and the GDPS-3.0.0 in parallel, right, for the 72 to 96 hours lead time over North America, valid 1200 UTC on January 28th, 2013, showing a system approaching the Great Lakes region and expected to spread significant freezing rain over portions of Ontario and Québec.

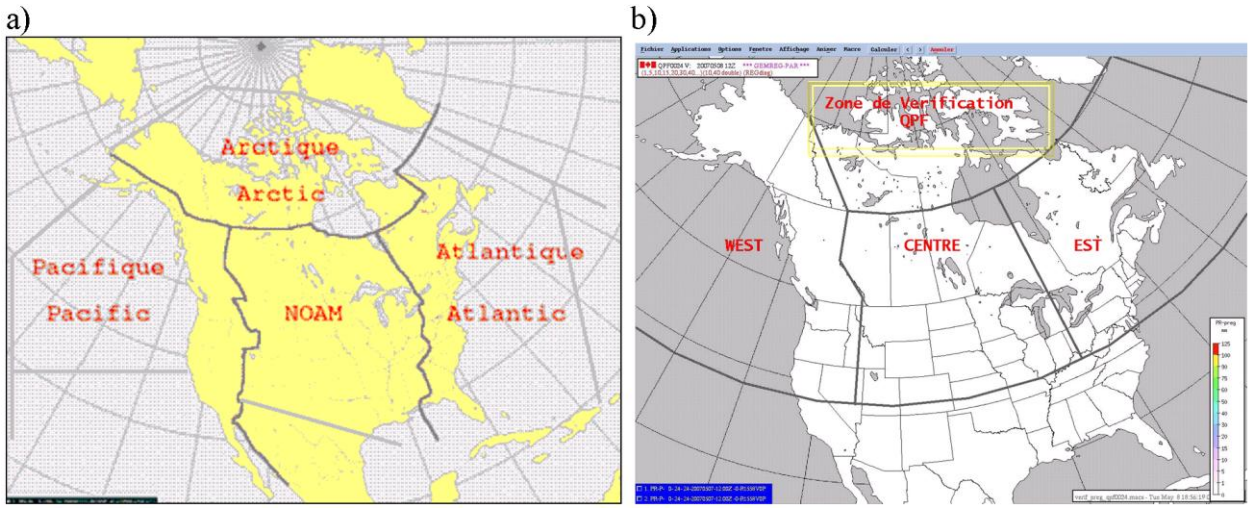


Figure 67: Evaluation regions for the CMC subjective evaluation, for a) mass fields (left) and b) QPF (right)

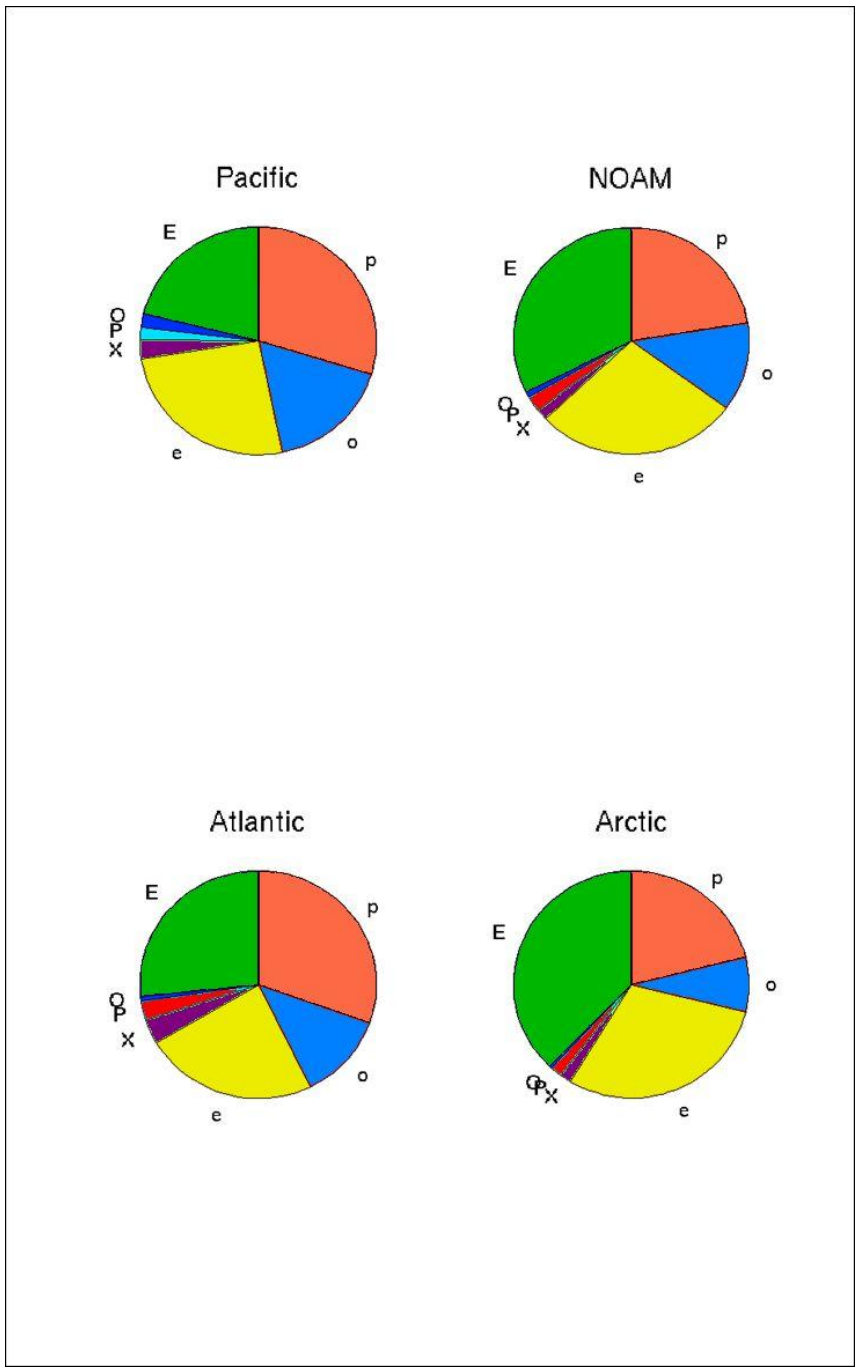


Figure 68: Proportion, for each evaluation region (as in Figure 67a) of the A&P meteorologists subjective evaluations for which they indicated what system had performed better, (parallel GDPS-3.0.0, letter "p", operational GDPS, letter "o", or much better (letter P and O). When both systems were alike, it they both performed equally very well this was indicated by letter "E", if both performed not so well or poorly, letter "e" was used. When both were fairly different and both performed poorly, the letter "X" was used. The exact percentages can be found in Table 3.

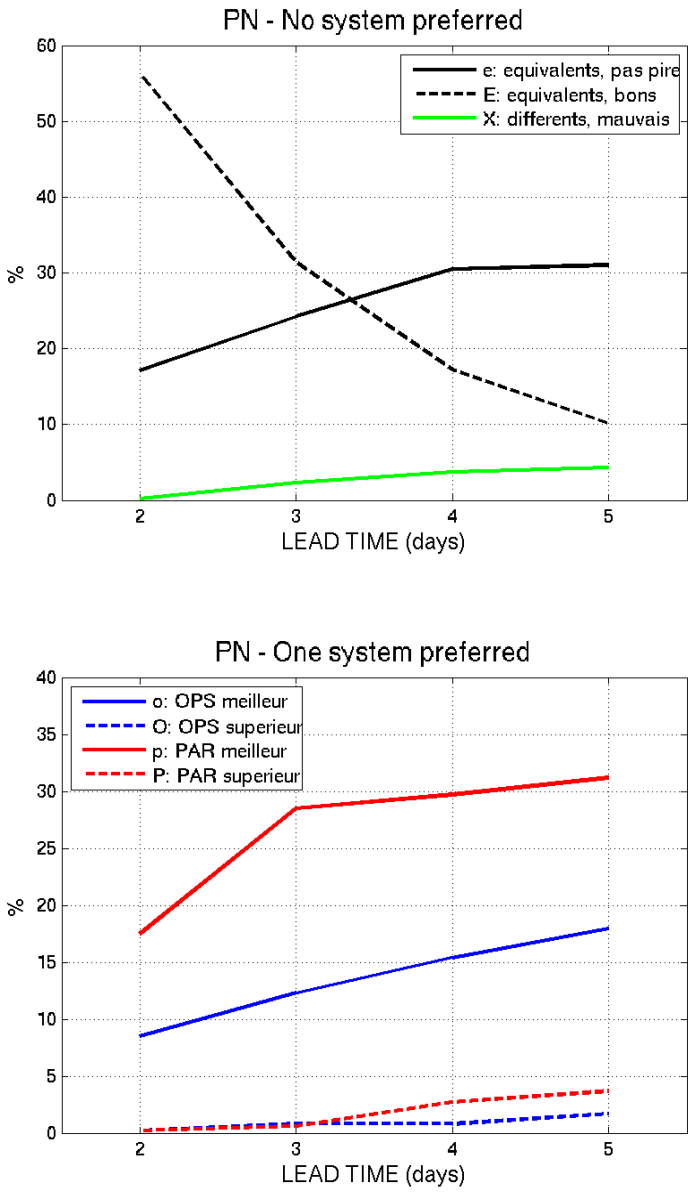


Figure 69: results of the A&P meteorologists subjective evaluations with respect to lead time (2 to 5 days) for MSLP. Bottom panels, percentage in favour of one system or another;
- "p", red line ("o", blue line), parallel (operational) system performed better;
- "P", dotted red line ("O" dotted blue line), parallel (operational) system was clearly superior;
Top panels:, percentage when no system was preferred;
- "E", dotted black line ("e" thick black line), both systems were alike and performed very well (or not so well);
- "X", green line, both systems were different and both performed poorly.

Done over 123 forecast runs, for 000 and 1200 UTC initial time.

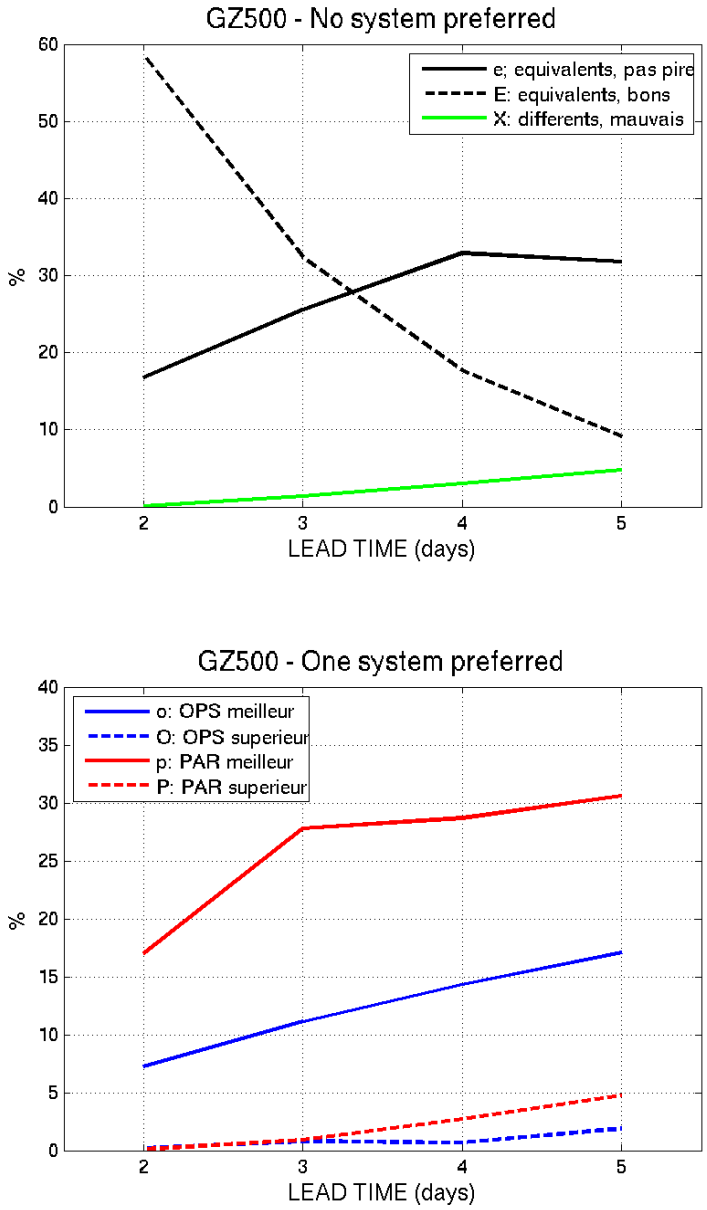


Figure 70: results of the A&P meteorologists subjective evaluations with respect to lead time (2 to 5 days) for GZ500. Bottom panels, percentage in favour of one system or another;
- "p", red line ("o", blue line), parallel (operational) system performed better;
- "P", dotted red line ("O" dotted blue line), parallel (operational) system was clearly superior;
Top panels:, percentage when no system was preferred;
- "E", dotted black line ("e" thick black line), both systems were alike and performed very well (or not so well);
- "X", green line, both systems were different and both performed poorly.

Done over 123 forecast runs, for 000 and 1200 UTC initial time.

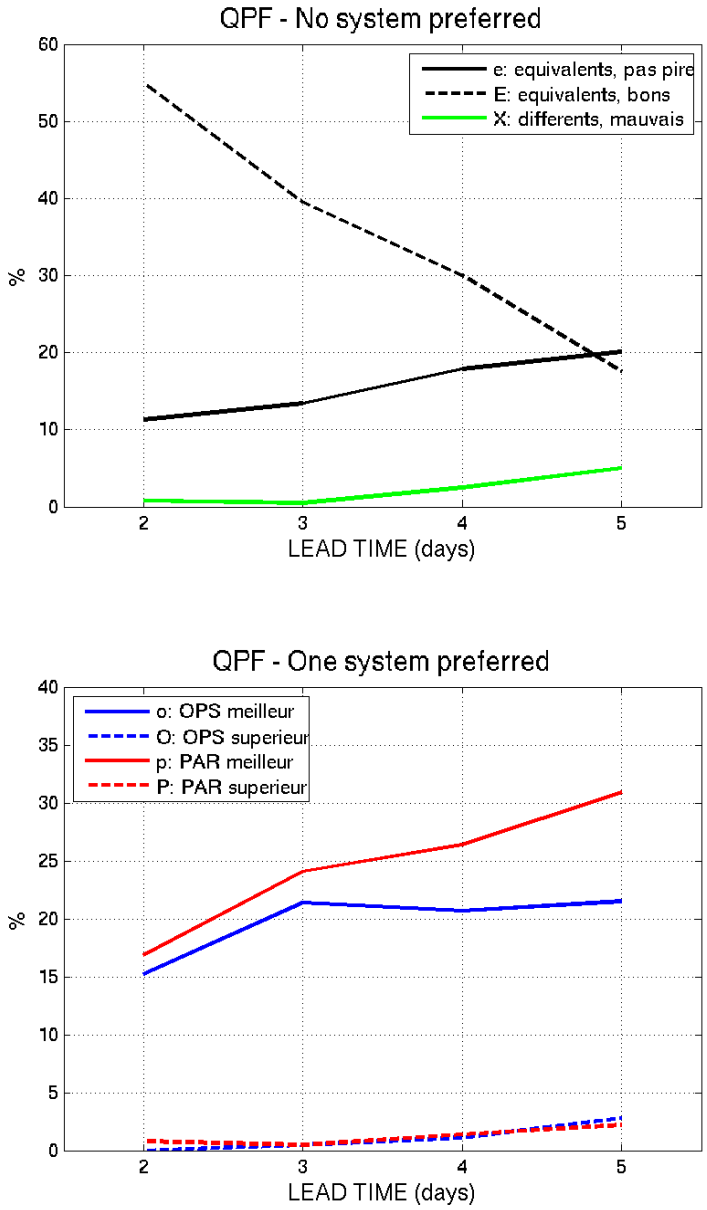


Figure 71: results of the A&P meteorologists subjective evaluations with respect to lead time (2 to 5 days) for QPF. Bottom panels, percentage in favor of one system or another;
- "p", red line ("o", blue line), parallel (operational) system performed better;
- "P", dotted red line ("O" dotted blue line), parallel (operational) system was clearly superior;
Top panels:, percentage when no system was preferred;
- "E", dotted black line ("e" thick black line), both systems were alike and performed very well (or not so well);
- "X", green line, both systems were different and both performed poorly.

Done over 123 forecast runs, for 000 and 1200 UTC initial time.

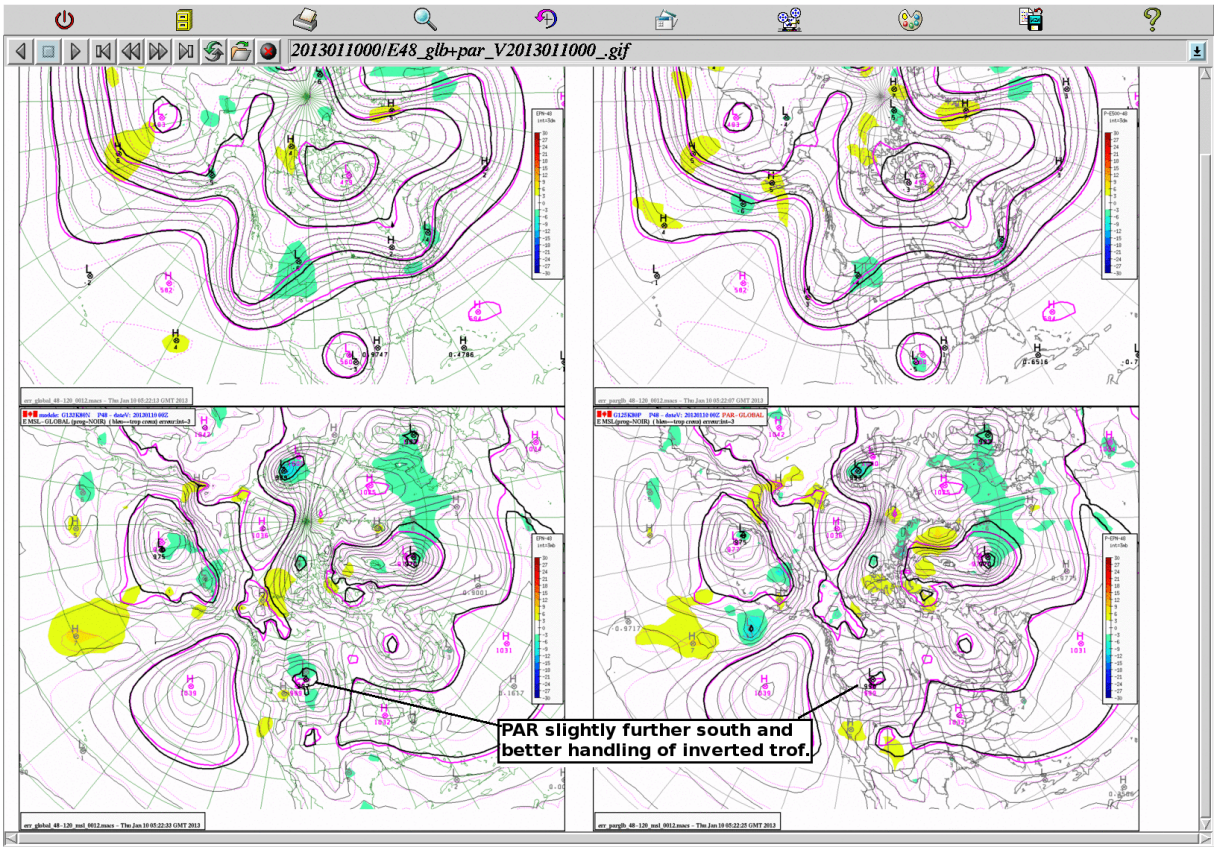


Figure 72: Difference between the 48h forecast from one system and its respective analysis, valid 10 January 2013, 0000 UTC. Operational system is at left, GDPS-3.0.0 at right; GZ500 at top, and MSLP at bottom. Positive errors are in yellowish shades, negative errors in bluish shades. This case illustrates a better handling by GDPS-3.0.0 of low developing to the lee of the Canadian Rockies.

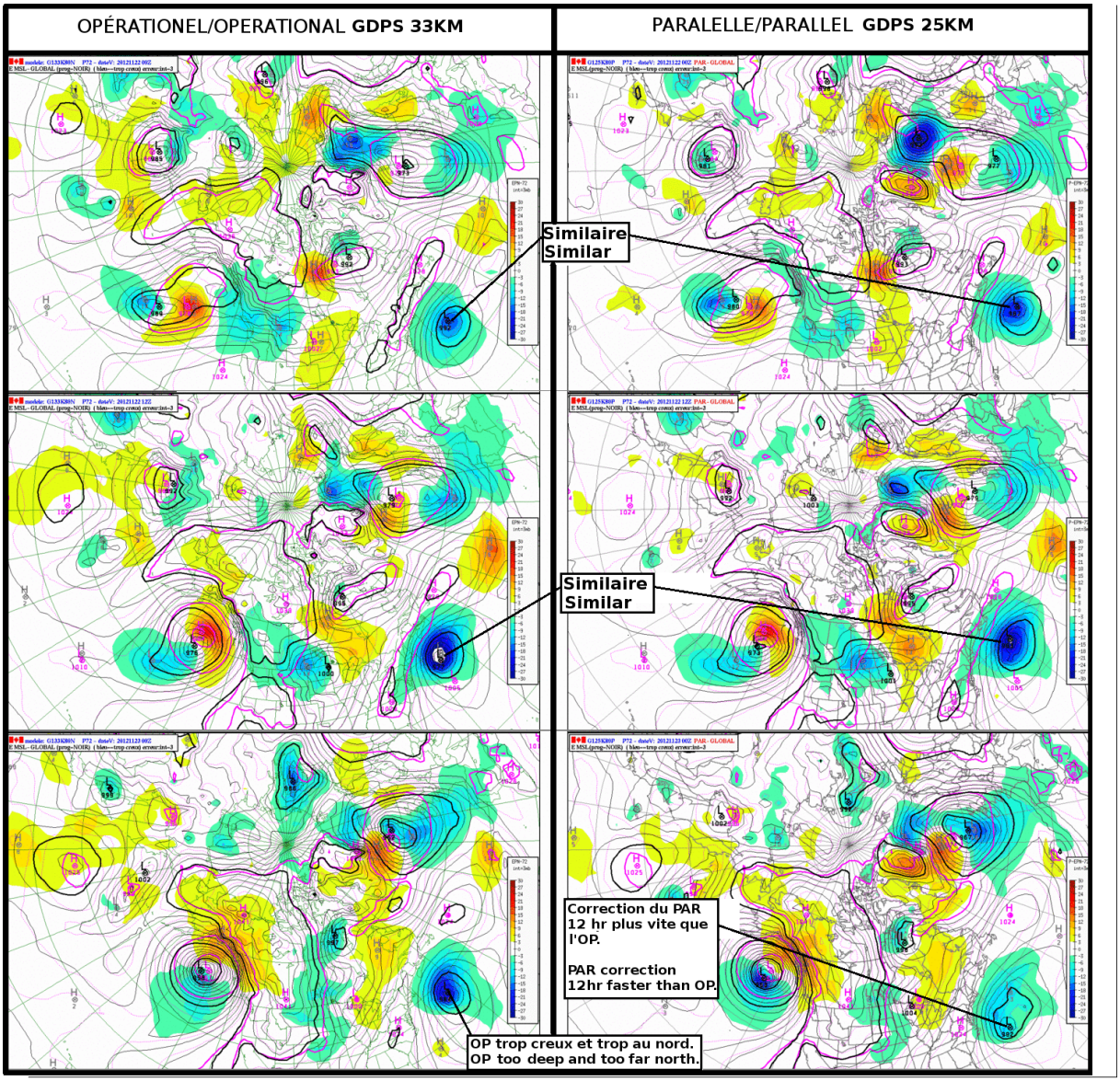


Figure 73: Sequence of 3 forecast runs for MSLP with 72 hr lead time, valid from November 22nd 0000 UTC to 23rd 0000 UTC, with verifications against each system analysis, operational to the left and parallel GDPS-3.0.0 to the right. Color fields show regions in error, negative (too deep) errors in blue and positive (not deep enough or too strong) in yellow. This illustrates the GDPS-33km faster or earlier corrections for persistent errors in the mid-range forecast compared to the GDPS-25km. In this case there was a systematic error in the forecast over the Atlantic of a deep storm well south of Nova Scotia. A much weaker system actually developed than forecast by both GDPS system, but the new GDPS-3.0.0 was faster than the operational runs to forecast a weaker, more southern system.

Figure 74: forecast tephigrams for Ottawa (YOW site), with lead-time of 21 (left) and 24 (right) hours, 18 December 1200 UTC forecast run; it shows the effect of the hysteretic effect in the planetary boundary layer (PBL) scheme introduced in GDPS-3.0.0. The GDPS-3.0.0 thermal profile is in red, that of the operational system in blue. Winds (blue wind barbs) are those of the operational system. The observed surface weather is in the blue ribbon at the bottom. This clearly shows how the inversion created by the warm air advection at mid-level and the near surface cold air advection is maintained the improved PBL scheme, producing a correct forecast of below freezing temperature and freezing rain.

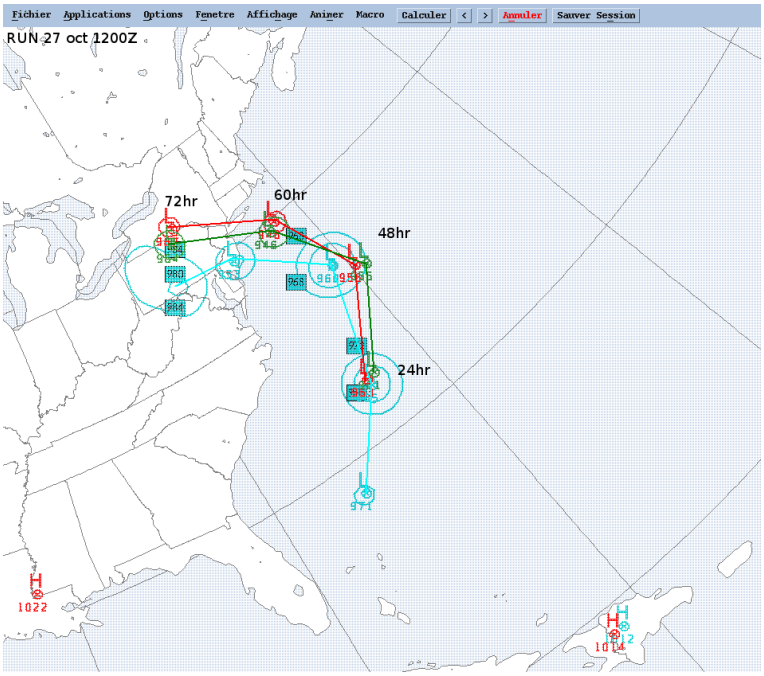


Figure 75: trajectory of the centre of hurricane Sandy in cyan at time of landfall, from 27 October 1200 UTC to 30 October 1200 UTC, with 24 to 72 hours forecast from 27 October 1200 UTC model run, GDPS-3.0.0 (green) and operational GDPS (red, opposite to usual convention). The new system forecast a trajectory closer to that observed at 60 and 72 hours lead time, after being too far east during the first 48 hours.

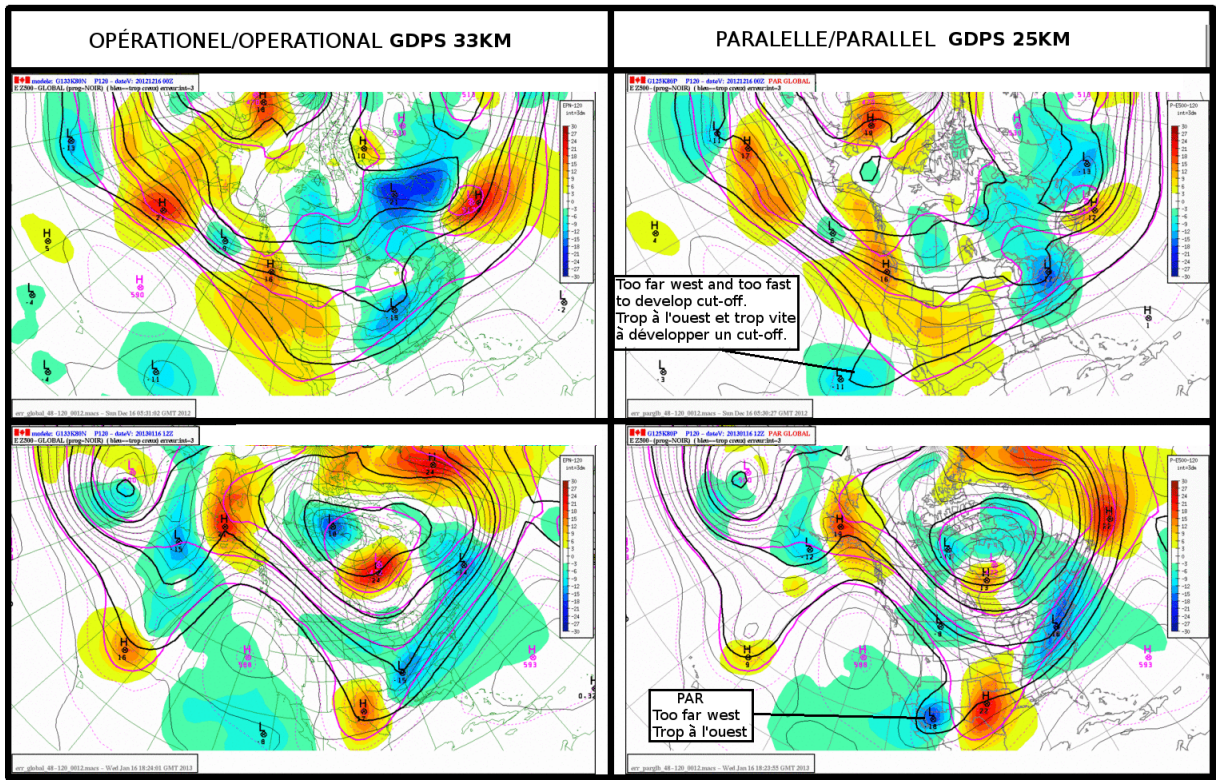


Figure 76: sequence of two 120 hours forecast valid 16 December 0000 UTC and 1200 UTC, from the operational system (left) and the GDPS-3.0.0 system (right), showing greater errors in the parallel system in relation with the formation of an upper low over the Southern U.S.. Negative errors are in blue and positive in yellow colors. The feature was forecast too far west and to close too early, and errors were propagated downstream (short wave impulse moving too fast over the Northeast U.S.