WORLD METEOROLOGICAL ORGANIZATION

COMMISSION FOR BASIC SYSTEMS
OPAG on DPFS

MEETING OF THE CBS (DPFS) TASK TEAM ON SURFACE VERIFICATION

GENEVA, SWITZERLAND
20-21 OCTOBER 2014

DPFS/TT-SV/Doc. 4.1d

(16.X.2014)

____

Agenda item : 4.1

ENGLISH ONLY

Development of verification application at the Canadian Meteorological Centre

(Submitted by Tom Robinson)

Summary and purpose of document

This document provides background information on a project to consolidate verification activities in the Research, Development and Operational section of the CMC. The prototype of the EMET verification application was used to develop a surface verification system for the High Resolution Deterministic Prediction System (HRDPS) and as such, it indicates a certain readiness on the part of CMC for the production of the required data for an international an exchange of surface verification data.

The document also shows results taken from the EMET system with respect to two issues to be discussed at the meeting, correction for station altitude and use of nearest neighbor for interpolation to verifying observation points

Action Proposed

The meeting is invited to take note of the steps undertaken at the CMC regarding the development of an NWP surface verification system and the results of the two studies.

Over the last few years the Canadian Meteorological Centre has engaged in the development of an application designed to consolidate verification efforts within the Research, Development and Operational components of the organization. The prototype application, called EMET (a Hebrew word for “truth”), was used to develop a verification system for the High Resolution Deterministic Prediction System (HRDPS, a 2.5 km grid model). As such, it naturally fit with the development of a model surface verification system.

The system uses a PostgreSQL database to store forecasts and observations at points. The database can then be queried to produce verification data and graphics. An early prototype using SQLite was abandoned based on performance requirements.

The system operates on an HP Z620 Workstation with 6 CPU. Four 500 GB Solid State Drives and a controller are used for storing the observations while a 3 TB hard drive for storing the forecasts. The observations are managed by the system itself in tables common to all users, so that the users do not have to concern themselves with this aspect. Meanwhile, the forecasts are in user or project-defined tables.

The observations currently in the database include temperatures (2m, max/min), wind (direction and speed), MSLP, cloud (opacity, total, ceiling) and visibility from both METAR and SYNOP reports (land stations) over the entire globe, as well as precipitation observations (3 h, 6 h, 12 h and 24 h accumulations), both from SYNOP reports and from the **Ca**nadian **P**recipitation **A**nalysis (CaPA). The system has the capacity to store at least 3 years of data, which will eventually include marine observations (bouys, ships, platforms) and radiosondes (either as fixed points or with displacement). A model-independent quality control system for the observations (described in the document for agenda item 6) is being developed for the system.

For the forecasts, any deterministic model output can be inserted into the database. Currently the system includes surface temperature, dew-point, winds and accumulated precipitation. The system accepts any standard type of interpolation (linear, cubic, nearest neighbor, etc) to the observation points.

Dynamic matching of forecast-observation (*F, O*) pairs is done upon insertion into the database. There is no need for static station lists as such, though this can be accommodated. New stations added to a domain are verified automatically. Forecasts are matched to observations from static (e.g. land stations) or moving (ships, radiosondes) platforms. Specific strategies were required to allow the matching of forecasts and observations to be done within a reasonable time-frame, given the enormous amount of data involved, as well as to ensure that the “best” observation is matched to the forecast in cases of multiple observations valid at the same time.

The database can be queried using either python or SQL scripts or via a web-interface, currently under development. The application uses the “OpenSource” concept to allow the development of modules for creating verification scores accessible to any user. Using the web-interface, users can select from pre-defined domains or define a domain of their own, creating either simple or complex polygons. The user chooses the forecast model, hours, variables, start/end dates, stations, scores and plot type. Manual input of SQL commands is accommodated as well.

Regarding system performance, uploading one year’s worth of METAR and SYNOP observations over the entire globe takes about 12 hours. For the forecasts, ingesting one month worth of Regional model scores (oo and 12 UTC runs out to 48 hours, matching with hourly METARs from temperature, dew-point, wind direction and speed, requires about 15 minutes. This includes interpolation of the forecasts to the observation points.

Once all the data is in place, querying the DB and producing scores and graphs is quite quick. Queries over the forecast tables take seconds to minutes depending on the size of the table containing the (*F,O*) pairs. Calculation of scores, including the confidence intervals via bootstrap takes no more than a few minutes or so. Without the bootstrap it is a matter of seconds.

Currently the system can produce the following scores:

1. Continuous scores
   - Mean forecast error
   - Mean Absolute Error
   - Standard deviation of the forecast error
   - Root Mean Squared Error
   - Correlation coefficient
2. Categorical scores
   - Hit rate
   - Frequency bias
   - Equitable threat score
   - Proportion correct
   - Heidke and Peirce Skill score

However, any type of score, either continuous or categorical, which can mathematically be described by f(F,O), can easily be added to the system. As well, the data can be displayed in any type of graph (time series, error growth, scatter plot, Q-Q plot, error distribution diagrams, etc).

Future work includes the

- Verification of clouds and precipitation type;
- Upper air verification against radiosondes (either as static points or following the motion);
- Ensembles (a strategy is required to handle the large volume of data involved);
- Post- processed data : the data format and matching of forecasts and observations are issues to be resolved;
- Continued development of a quality control system
- Continued development of the web-interface

For the purposes of this meeting, two relevant studies are presented here based on data from the EMET verification system:

1. Correction for station altitude between model and actual station elevation

   A simple standard atmosphere (+6.5 ^{0}C / 1000m) correction was applied to the model position with respect to the station height. Since stations are generally in valleys this results in an increase in the temperature bias, which produces a favourable result on the scores in winter where the GDPS model has a cold bias, but a negative effect in summer when the model has a slightly warm bias.

   On the other hand the correction results in a positive impact on the RMS error.

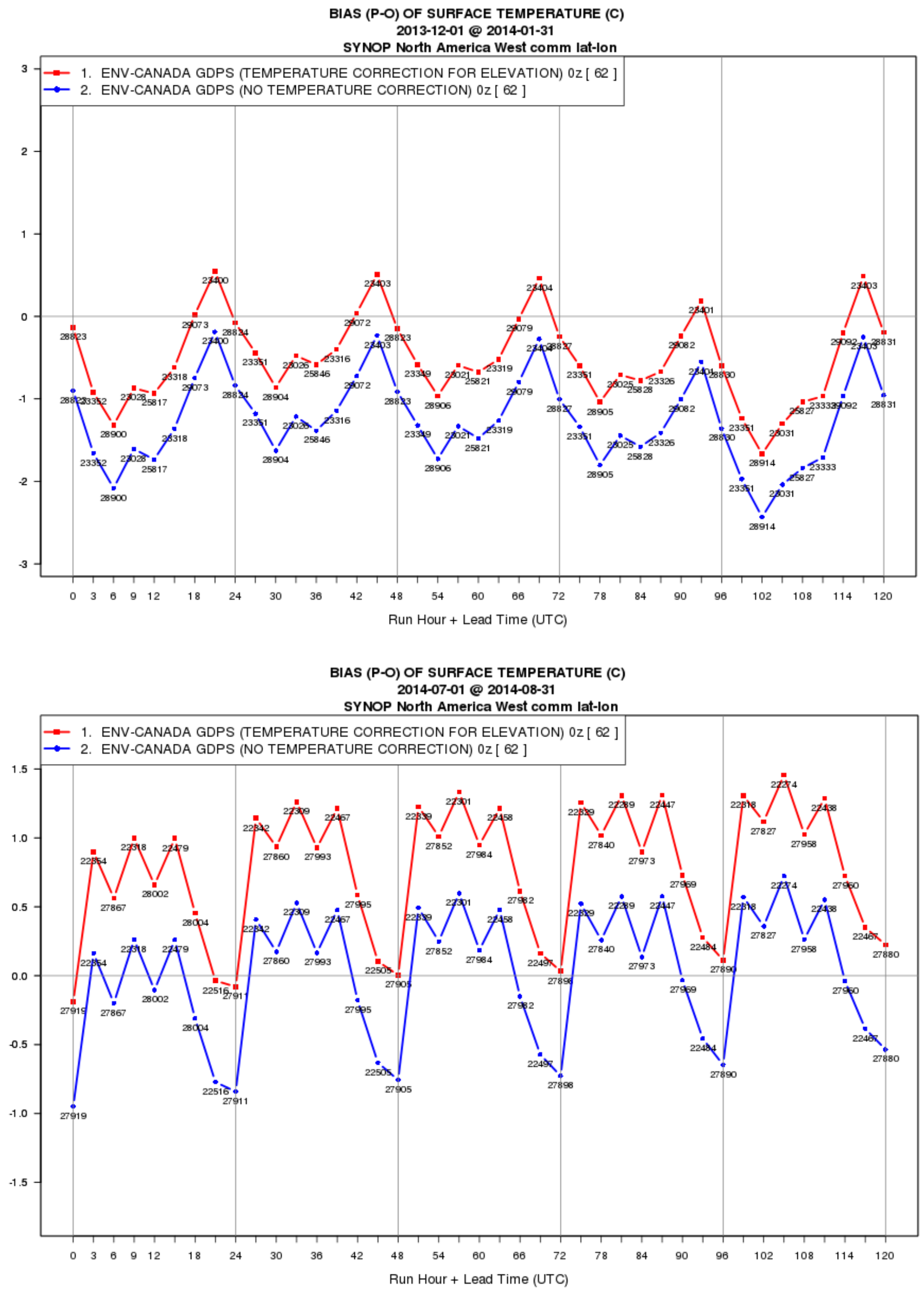

Figure 1: Mean forecast error with and without correction for altitude between model point and observation site of the CMC Global Deterministic Prediction System surface temperature forecasts scored against the North America SYNOP observation network over the period December 1, 2013 to January 31, 2014 (top) and July 1 to August 31, 2014 (bottom)

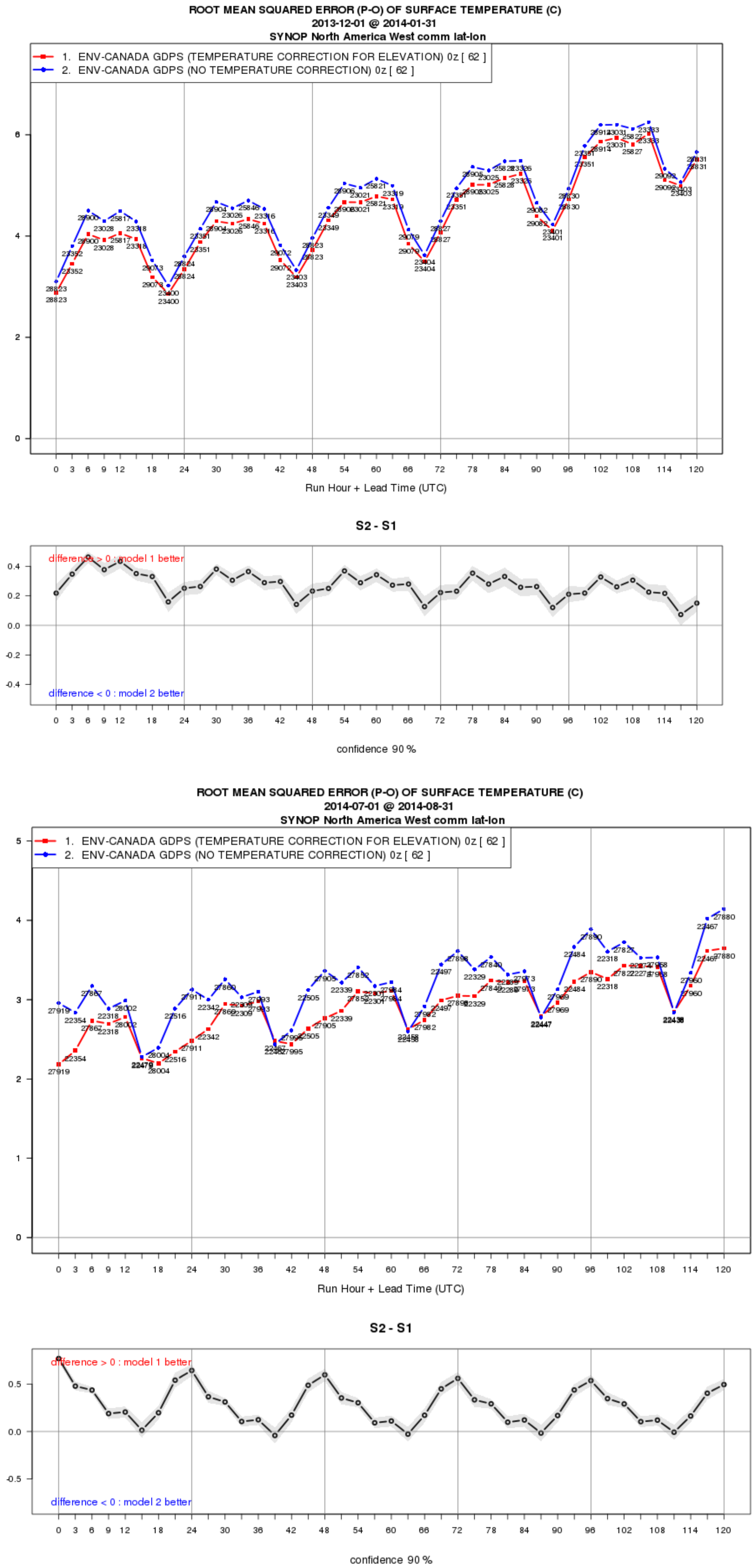

Figure 2: Same as figure 1, but for the root mean squared error. Confidence intervals are calculated using a block bootstrap on the difference between the two cases

2. Use of nearest neighbour model point vs. nearest land point

Verifying forecasts at coastal observations sites is problematic as the nearest neighbouring model grid point may be over water rather than land, thus leading to unacceptable biases.

For this study a domain (fig. 3) was chosen over eastern North America in order to maximise the number of coastal or shore stations along the eastern seaboard, Gulf of St. Lawrence and the Great Lakes. A winter period was chosen in order to maximise the difference between land and sea conditions.

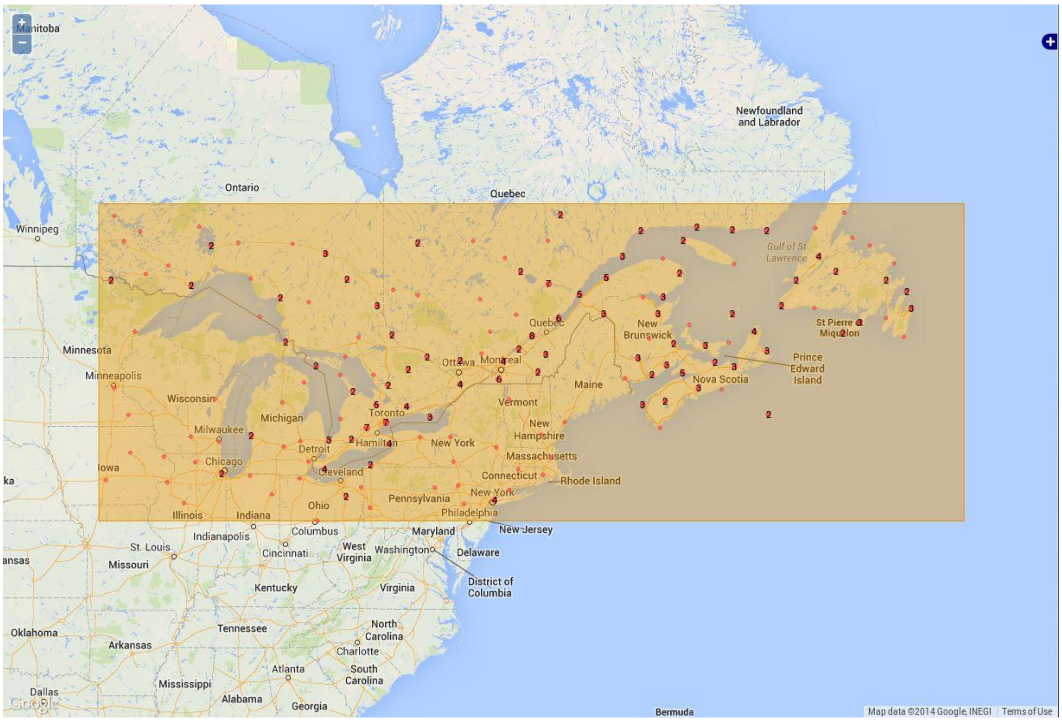

Figure 3: domain chosen for the study on nearest neighbour

Note first of all that CMC models do not use a strict land-sea mask. Instead, they use an aggregate surface where the grid-points are characterized by a percentage of land and a percentage of sea. This allows the model to combine in an intelligent manner the effects of land and sea which more realistically reflects these effects on coastal observations sites.

In order to test the effect of using land values only, a scheme was used by which the land surface was artificially extrapolated in such a way that the grid-points surrounding an observation point was entirely land.

The results indicate a generally significant degradation in temperature scores when only land stations are taken into consideration (fig. 4). The mean forecast errors (fig. 5) indicate that the forecasts over the strictly land points are cooler than when using a land-sea hybrid, which makes sense since the sea in winter would have a warming effect. In the overnight period the land points become unrealistically cold. During the course of the day, the effect on the bias is somewhat positive insofar as it reduces the model warm bias.

The conclusion drawn from the study is that, given the use of an aggregate surface at the interface between land and water, the use of nearest land point to verify against the coastal observations would overly penalize the CMC GDPS in comparison with the use of nearest neighbour.

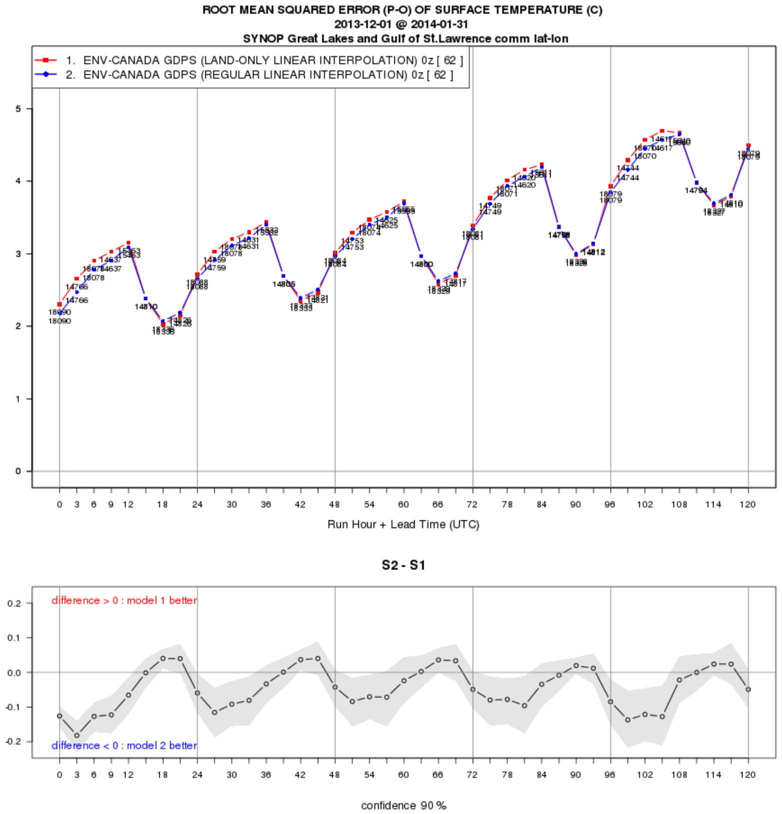

Figure 4: Root mean squared error of CMC GDPS surface temperature forecasts using linear interpolation of forecast fields to observation points form nearest land point and from nearest grid point, for the period December 1, 2013 to January 31, 2014, and corresponding confidence intervals.

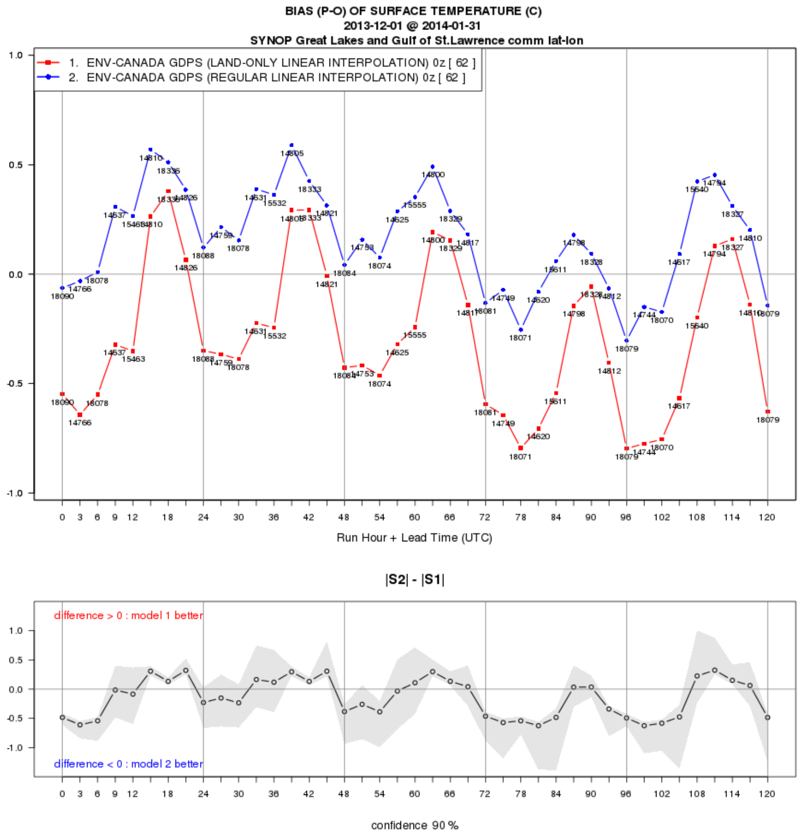

Figure 5: Same as figure 4, for the mean forecast error.