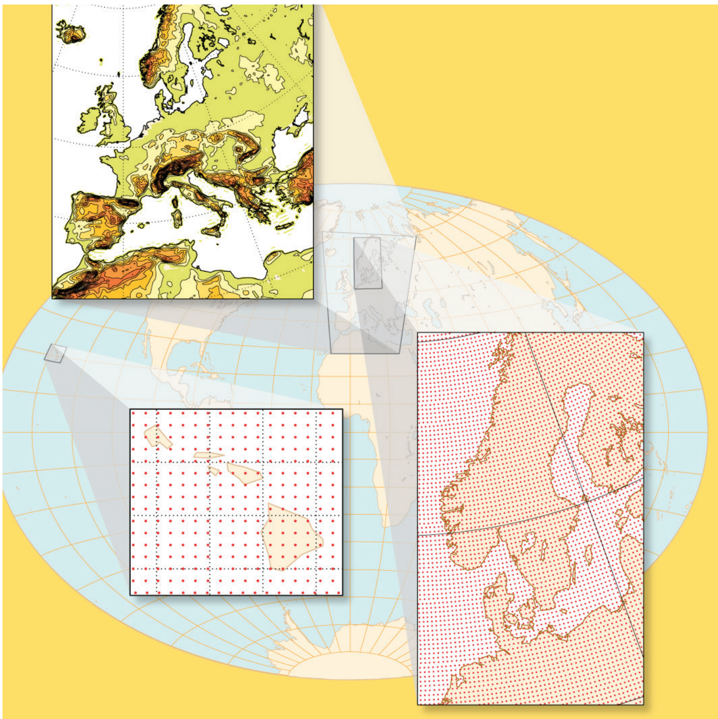
ECMWF Feature article

from Newsletter Number 108 – Summer 2006

METEOROLOGY

Surface pressure bias correction in data assimilation



www.ecmwf.int/en/about/news-centre/media-resources

doi:10.21957/uv295rfmx5

This article appeared in the Meteorology section of ECMWF Newsletter No. 108 – Summer 2006, pp. 20–27.

Surface pressure bias correction in data assimilation

Drasko Vasiljevic, Erik Andersson, Lars Isaksen, Anotonio Garcia-Mendez

It is well known that the surface pressure (Ps) observations reported by a large number of SYNOP land and sea (SHIP) stations as well as drifting buoy (DRIBU) stations are biased, and in many cases by several hPa. These biases are mostly related to incorrect assumptions about the station heights, and remain fairly constant in time. Therefore, several hundred stations would normally appear on the ECMWF blacklist due to a significant long-term bias.

A new scheme for estimating and correcting Ps bias, based on an adaptive correction method, was introduced into the ECMWF operational analysis/forecasting system on 5 April 2005. The scheme has had a positive impact on both the analysis and forecast, and has reduced the number of stations on the blacklist.

Bias correction method

A few years ago Peter Janssen proposed an adaptive method for bias correction of Ps time series. The method is based on linear estimation theory and is referred to as the OI method. This provides estimates of the bias and its confidence, station by station, based on a time series of observation-minus-background departures. Furthermore, the method relies on two assumptions: (a) the observational Ps bias is local for a given station (i.e. no spatial correlation) and (b) there is no, or only small, model bias. For more details see the Appendix.

Experiments have been carried out to examine the behaviour of the OI method. Figure 1 shows Ps departures and the bias correction time series using the OI bias estimates. This is for SYNOP station 65355 (Niamtougou, Togo) for February–April 2005. It can be seen that the station has a bias of about -11 hPa. The OI method clearly identified the bias and appears to be a smooth operator.

Practical aspects

A central component of the bias correction scheme is its database, PSBIAS. The PSBIAS is a hierarchical database modelled on ECMWF's operational ODB (Observation Data Base) and it plays a key role in providing the cycling mechanism for the Ps bias correction scheme. The PSBIAS is structured in such a way that the database main entry points are stations. Each station, or entry point, is further divided into two parts: (a) header and (b) body, which are appropriately linked. All relevant bias correction parameters needed to be carried forward in time to perform the bias correction are kept in the header, whereas the body part keeps the station time record. The time record is up to one month long for hourly observations and longer for less frequent observations.

Not all Ps observing stations are considered for bias correction. For example, stations for which either station altitude is missing or the height difference between station and model is more than 200 m are omitted. Furthermore, a Ps observation is discarded if the observation departure is bigger than 15 hPa or reported pressure is mean-sea level pressure but the station is above 500 m. Bias estimates are calculated only if the station sample size is big enough. At the moment the sample size limit is set to 30. Once this number is reached the station is assumed to be purely biased only if its bias estimate is bigger than one standard deviation. Additionally, in order to avoid correcting for possible small model biases, any bias estimates less then 1 hPa are not considered. Also, bias estimates which are too old are not used. Currently, bias estimates older than 5 days are considered to be old. In that case they are recalculated from scratch (cold start), once a sufficient new sample is available.

As an illustration of the impact of the bias correction scheme, Figure 2 shows the bias corrected Ps departure (red) along with the original Ps departure (blue) and applied Ps bias correction (green) for the same SYNOP station as used in Figure 1. To start with there is no bias correction because the sample size is not big enough (the warm up period). Once the bias correction kicked in, it was correcting departures quite nicely. It can also be seen that at the beginning the scheme was turning itself on and off a few times before settling in. This station's stable long-term bias was interrupted a couple of times during this period. This was a result of two out of sequence departures. Obviously in these two cases the bias correction scheme made it even worse but recovered quickly. As it turned out, this station is RDB (Report Data Base) flagged because of its wrong altitude; hence it was not used even after bias correction. However, by closer examination of the first guess check flags in both the *"no Ps bias correction"* and the *"Ps bias correction"*

 doi:10.21957/uv295rfmx5

runs, it was found out that in the *"no Ps bias correction"* run the Ps was flagged and rejected, whereas in the *"Ps bias correction"* run the Ps was not flagged and would have been used in the analysis if it was not for the overriding RDB station height flag. It would be quite possible to come up with a practical scheme whereby biased stations, like this one, could have their altitude "corrected" by using the long-term bias. Likewise, a similar mechanism could be applied for blacklisted biased stations. Blacklisted stations have been monitored after bias correction was introduced and many stations have been taken off the blacklist.

Figure 3 shows another example of a Ps biased station (SYNOP 82353, Altamira, Brazil). However, in this case the station height is thought to be correct and the real reason for being biased was not known. Again, after the warm up period, the bias correction scheme turned itself on and was correcting the Ps by about -3 hPa. After closer examination of this station it was found that in the *"no Ps bias correction"* run Ps was surviving (just) the first guess check but only to be rejected by the analysis check. On the other hand in the *"Ps bias correction"* run it passed all the checks and it was used in the analysis. However, there were two occasions when it failed the first guess check. It first happened when the original Ps departure became positive (in a long sequence of negative ones) and after bias correction it became even bigger. The second time the original departure happened to be about twice its usual negative size. The bias correction scheme did correct it, but the departure was still too big – hence it was rejected.

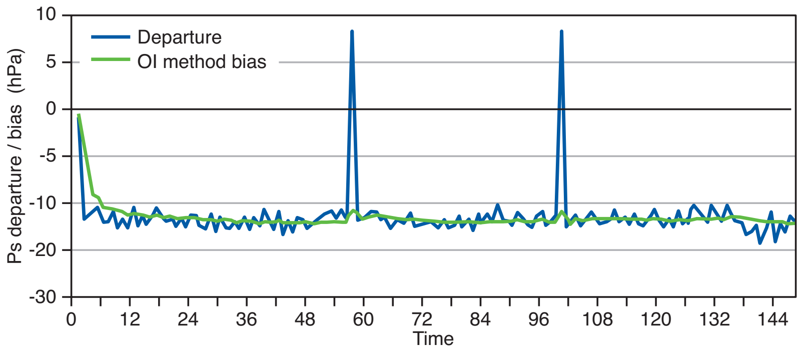


Figure 1 Ps departure/bias time series (February–April 2005) for SYNOP station 65355: Ps departure (blue) and OI method bias estimate (green).

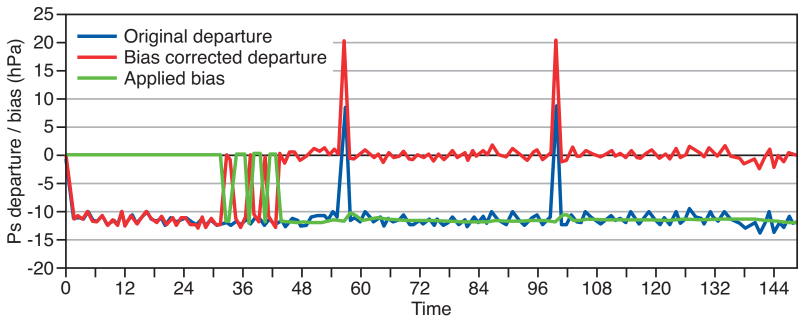


Figure 2 Ps departure/bias time series (February-April 2005) for SYNOP station 65355: original Ps departure (blue), bias corrected Ps departure (red) and applied Ps bias (green).

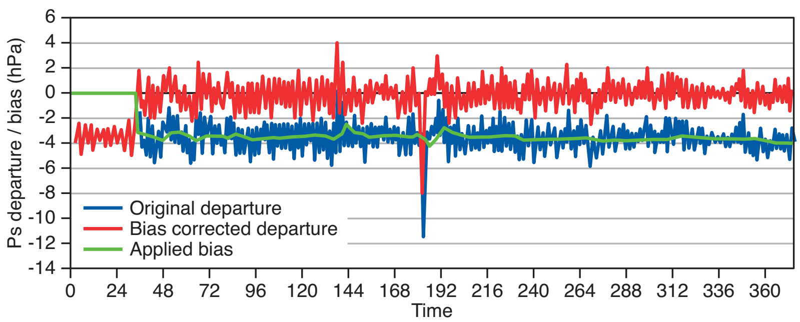


Figure 3 Ps departure/bias time series (December 2004–April 2005) for SYNOP station 82353: original Ps departure (blue), bias corrected Ps departure (red) and applied Ps bias (green).

Analysis and forecast impacts

Experiments with the Ps bias correction scheme

Prior to the operational implementation of the Ps bias correction scheme a number of experiments with and without Ps bias correction have been performed. They have all shown the expected results. In these experiments there were several hundred bias corrected stations. Analysis increments were somewhat smaller and the impact on forecast scores was about neutral. Eventually, E-suite or E-suite type experiments were carried out (the E-suite is used to test possible operational changes).

There was a long E-suite type experiment with Ps bias correction switched on from 1 August 2004 till 31 December 2004. Furthermore, the E-suite was then run from the 1 January 2005 till the operational implementation on 5 April 2005. Additionally, just for August there was another E-suite type experiment without the Ps bias correction shadowing the above mentioned E-suite type experiment. Comparing these two August E-suite type experiments there were no surprises coming from the Ps bias correction. At the end of the month about 800–1000 biased stations were identified and the correction scheme itself seemed to have performed satisfactorily. Overall there was a better fit to Ps observations with average analysis increments slightly reduced. The impact on the forecast in terms of geopotential anomaly correlation scores for 10-day forecasts was neutral for the northern hemisphere and slightly positive in the southern hemisphere.

The E-suite type experiment with Ps bias correction went on till the end of December 2004. Unfortunately, the forecast scores did not look that good in December. In the southern hemisphere the impact was still about neutral, but rather negative in the northern hemisphere. Negative results were obtained also for North Pacific, North America, North Atlantic and Europe. It is for that reason that a new E-suite type experiment for December with the Ps bias correction turned off was carried out. This new experiment was started from the already running E-suite type experiment. The mean-sea level pressure analysis differences between these two experiments for the very first analysis cycle are shown in Figure 4.

As it can be seen in Figure 4, most differences are small and local as expected. However, there is a rather large-scale positive difference over the North America (Canada/ USA). This was not expected and merited further investigation.

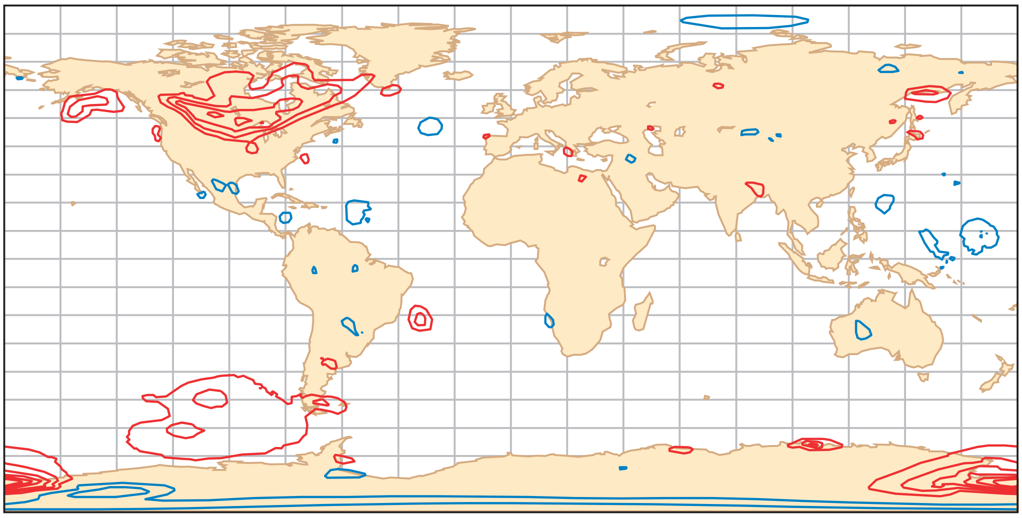


Figure 4 PMSL analysis difference (*"Ps bias correction"* minus *"no Ps bias correction"*) for 00 UTC on 1 December 2004: positive (red) and negative (blue); 0.5 hPa contouring interval.

doi:10.21957/uv295rfmx5

Investigation of the large difference between experiments over North America

Looking at the daily PMSL analysis difference between the two experiments, it was found that the *"BLOB"* stayed over North America for about two weeks, though it did move around slightly. Furthermore, the *"BLOB"* survived in the ensuing forecasts and propagated downstream with the flow contributing to the bad forecast scores. There was no immediate answer to what went wrong.

The first thing which came to mind was that the Ps bias correction scheme somehow went wrong. Checking the PSBIAS database it was found out that for this particular analysis cycle about 80 stations in the area were bias corrected. These 80 stations accounted for about 600 out of 1,000 reports for the 12-hour data window. Out of these 600 reports about 200 were SYNOP and 400 were METAR reports. The size of the Ps bias ranged from -1 hPa to -3 hPa. Thus, all stations were negatively biased. The large-scale pattern was not expected as our first assumption when designing the Ps bias correction scheme was that observational Ps bias is not spatially correlated. Looking at the bias time series for these stations leading to the analysis cycle in question one could not see anything unusual in the scheme's behaviour.

Figure 5 shows the background and analysis fit to SYNOP Ps observations in the region. The black line represents the fit to data for the *"Ps bias correction"* run, whereas the red line represents the fit for the *"no Ps bias correction"* run. From the *"no Ps bias correction"* run it is clear that there was a bias of about -2 hPa. Furthermore, Figure 5 shows that in the case of the *"Ps bias correction"* run the scheme did a rather good job by, to a large extent, removing the bias. What Figure 5 also shows is that both analyses were a close fit to the observations. However, the analysis resulting from the *"Ps bias correction"* run creates the *"BLOB"* which subsequently impacts (negatively) the forecast quality. At this point it became clear that further investigation was needed.

First of all, none of the just described behaviour was noticed during the August runs, thus raising the question what happened with Ps biases from September till December. In order to try to answer this question Ps departure time series for the Canada/USA region were looked at. The time series are shown for the E-suite type experiment with Ps bias correction (Figure 6(a)) and the operational version (Figure 6(b)) for SYNOP data from August till December 2004. Both time series clearly show that from August till late September there was not much bias in the Canada/USA region to start with. However, as from late September, bias started creeping in and growing. As discussed earlier, the bias correction scheme was recognising and correcting it. Since there was no bias correction in the operational system at the time the size of the bias was somewhat bigger.

Now it started looking as we might be dealing with a possible model bias. If true, the bias correction scheme should not be applied. This is because it would go against the second bias correction scheme assumption that there is no, or small, model bias. As mentioned earlier, in anticipation of something like this, although on a smaller scale, the 1 hPa limit on when to apply the bias correction had been introduced in an attempt to avoid correcting for possible small model biases. Also, what came as a surprise here was that there was no other region like Canada/USA where a similar problem could be found.

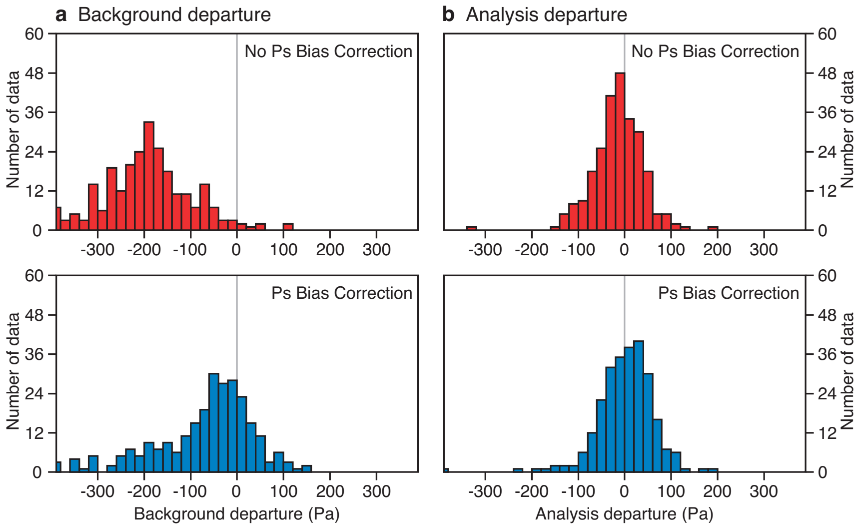


Figure 5 Background and analysis fit to SYNOP Ps data for 00 UTC on 1 December 2004 for Canada/USA: (a) background departure and (b) analysis departure; *"no Ps bias correction"* (red) and *"Ps bias correction"* (blue); units Pa.

In order to facilitate understanding this problem further, a time series of maximum, minimum and average PMSL analysis differences for various scenarios and regions were looked at. What one expects to see from this type of time series is that both maximum and minimum differences vary day by day, whereas the average differences are stable and around 0 hPa. Time series of PMSL analysis differences between an E-suite type experiment with Ps bias correction and operational system for August–December period for the northern hemisphere are shown in Figure 7(a). Red and blue lines are maximum and minimum differences, respectively, and the black line is the average difference. As expected the maximum and minimum differences vary day by day, and the average difference stays rather stable and just hovers around 0 hPa.

Figure 7(b) shows differences for the Canada/USA region, and here we have a surprise. From August till late September the average differences are not as stable as previously seen with amplitudes of about 1 hPa, but still reasonable. However, from late September till the end of December the average differences are far from stable and are as big as 2 hPa, or even bigger. This result seems to coincide with the large-scale Ps bias noticed earlier. Unfortunately, we did not have the *"no Ps bias correction"* experiment for the whole of August– December. But there was one during December which was mentioned earlier. The average PMSL differences for *"no Ps bias correction"* run (not shown) are not that stable in the December either, though the amplitudes are not as big as in the in the *"Ps bias correction"* run, just going over 1 hPa.

Meanwhile, as the E-suite (started on 1 January 2005) was going on, the forecast scores improved and the December problem in the Canada/USA region was less and less evident.

All this was suggesting that we are dealing with unexpected model bias. The bias correction scheme presented and introduced here is only supposed to deal with uncorrelated observational Ps bias. Thus in the presence of a larger-scale model bias the correction should not be applied. The following section will deal with possible ways of identifying and dealing with model bias.

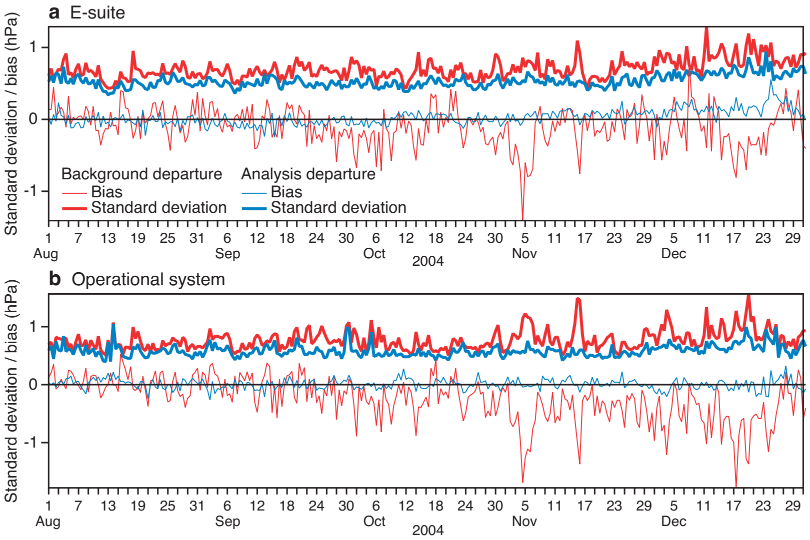


Figure 6 SYNOP Ps time series for 00 UTC on 1 August to 12 UTC on 31 December 2004 for Canada/USA for (a) E-suite and (b) operational system. Bias of background departures (thin red) and analysis departures (thin blue) with corresponding standard deviations in thick red and thick blue; units hPa.

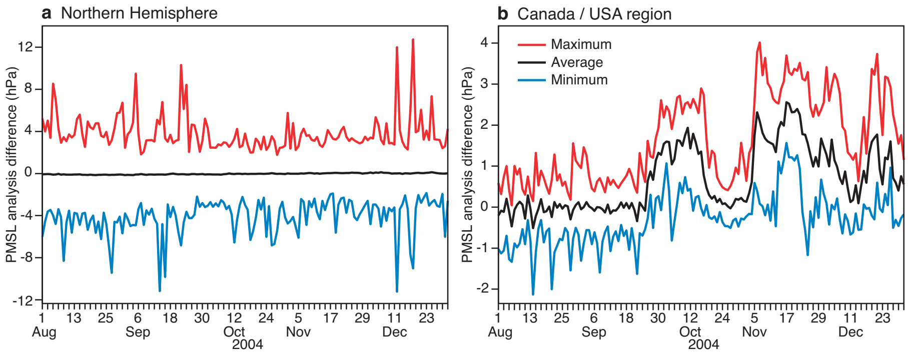


Figure 7 Maximum (red), minimum (blue) and average (black) PMSL analysis difference (*"Ps bias correction"* minus *"operational system"*) for 00 UTC on 1 August to 00 UTC on 31 December 2004 for (a) the northern hemisphere and (b) Canada/USA region; units hPa.

doi:10.21957/uv295rfmx5

Model versus observational bias

As just discussed correcting model bias by correcting observations leads to a rather poor result and it should not be done. The difficulty is how to identify the model bias and subsequently separate it from the observational one. It is worth remembering that we introduced a 1 hPa limit on when to apply the bias correction in order to avoid correcting for small model bias. Of course that limit could be increased to say 2 hPa. This rather quick fix should hopefully eliminate the problem which occurred over North America. However, since the limit is applied globally, the increase would unjustly exclude a number of genuinely biased stations from being corrected, therefore killing positive effects of the bias correction scheme. Thus a more selective solution to this problem should be sought.

The analysis of the problem presented in the previous section clearly indicated that when it happens a large number of stations are in agreement in terms of both the bias sign and the bias size. Now, remember that in the old ECMWF OI analysis system we used to have a quality control procedure called the *"buddy"* check which is not used in the current ECMWF analysis system. In brief, the idea was that in order to quality control a given observation one could actually do the analysis at that point without the observation itself being used. Then if analysed and observed values at that point agree within some limits one assumes that the observation is probably correct. Now if we turn this idea around, and if for a given biased station its bias value agrees, within limits, with the analysed bias from its neighbours without using the station itself, then one should not apply bias correction at that station. The neighbouring stations to be considered should be within a circle of a certain radius. Also, as an agreement limit, we could use for example the analysed bias value plus or minus a multiple of the standard deviation. But do not forget that there should be a limit on how many stations ought to be found in the vicinity. Since this type of check is very similar to the original *"buddy"* check but in the opposite sense, naming it the *"anti-buddy"* check sounded appropriate. Furthermore, there were at least two possibilities on how to perform the bias analysis from the neighbouring stations.

- One could do a simple statistical analysis, calculate the mean and standard deviation and use the mean as analysed bias value along with the standard deviation to perform the *"anti-buddy"* check.
- Instead of using the mean as the analysed bias value a type of two-dimensional univariate bias analysis could be performed.

The *"anti-buddy"* check has been added to the Ps bias correction scheme. First, as in the original scheme, a list of potentially biased stations is compiled and then for each of them the *"anti-buddy"* check is applied. The circle radius around a given station is set to 300 km. The number of influencing stations used to perform the *"anti-buddy"* check is set to 10 or more. The analysed bias value is assumed to be the mean bias and the agreement limit is set to an analysed bias value of ±2.0 standard deviations. Stations which do not have enough neighbours are not subjected to this check.

Figure 8 shows the PMSL analysis difference between what we now call the *"Ps bias correction plus anti-buddy check"* run and the original *"Ps bias correction"* run. As can be seen, the map is mainly void except for the Canada/USA region where we experienced the problem before. This was a very good result. The *"anti-buddy"* check clearly had an impact only in the problematic region, and both the sign and the size of difference are good.

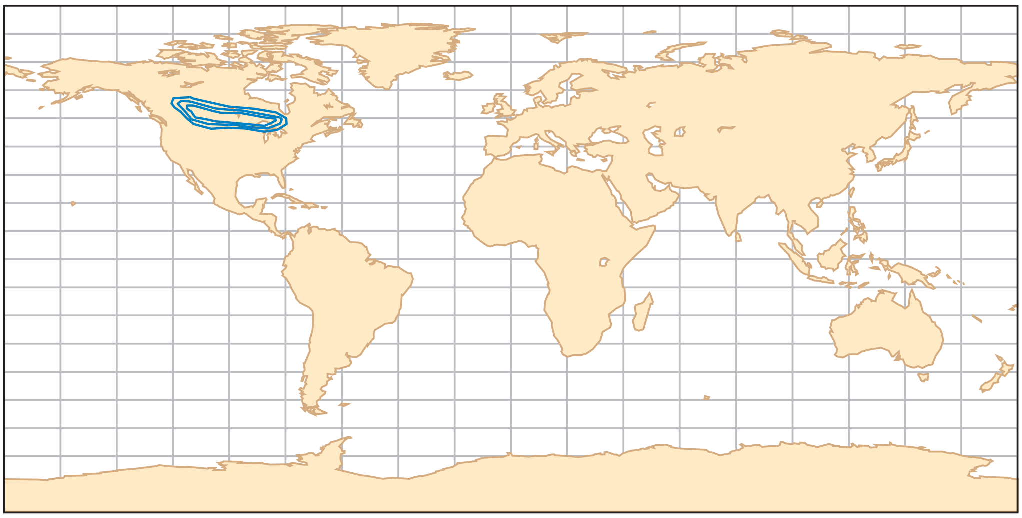


Figure 8 PMSL analysis difference (*"Ps bias correction plus anti-buddy check"* minus original *"Ps bias correction"*) for 00 UTC on 1 December 2004: positive (red) and negative (blue); 0.5 hPa contouring interval.

Now comes the all important question: if we were to rerun the E-suite type experiment with the *"anti-buddy"* check included for December, would that improve the forecast scores?

As mentioned earlier the original *"Ps bias correction"* run underperformed the *"Ps bias correction"* experiment particularly in the northern hemisphere. The extent of the under performance in the northern hemisphere was also evident in the regional scores. The *"Ps bias correction plus anti-buddy check"* experiment has done a lot better in the northern hemisphere than the original *"Ps bias correction"* experiment. In the southern hemisphere, where there was not much of the problem anyway, there was little difference between the experiments. The improvement in the northern hemisphere scores was also found in the regional scores.

The *"anti-buddy"* check had a positive impact on both the analysis and forecast. Consequently it was included in the operational implementation of the Ps bias correction scheme.

To sum up

The Ps bias correction scheme based on the OI method for estimating and correcting Ps bias is now a part of the ECMWF operational system. The scheme identifies about 1,000 biased stations out of about 11,000 surface stations. The biases are mostly related to incorrect station height and remain more or less constant in time. The scheme is based on two assumptions: (a) Ps bias is local (no spatial correlation) and (b) no model bias, or very small model bias. When both of these assumptions are satisfied the scheme had a positive impact on both the analysis and forecast. However, in the presence of a larger model bias (both spatially and size wise) the scheme was not performing as well. Thus, an adjustment to the scheme was needed to recognise a possible model bias and separate it for the observational bias. The proposed *"anti-buddy"* check has managed to fulfil this requirement.

The Ps bias correction scheme has now been used operational for more than a year and has been performing well. Also, the scheme is being used in the re-analysis experiments and from the first runs it appears to be doing well there too.

The current situation is illustrated by Figure 9 which shows all operationally Ps bias corrected stations for the 12 UTC analysis cycle on 29 May 2006. There are 1,260 bias corrected stations out 15,444 available stations for this analysis cycle. Furthermore, since the Ps bias correction scheme became operational about 150 biased stations have been taken off the blacklist.

We would like to take this opportunity to thank Anders Persson for his valuable contributions and for bringing the issue of surface-pressure station bias to our attention.

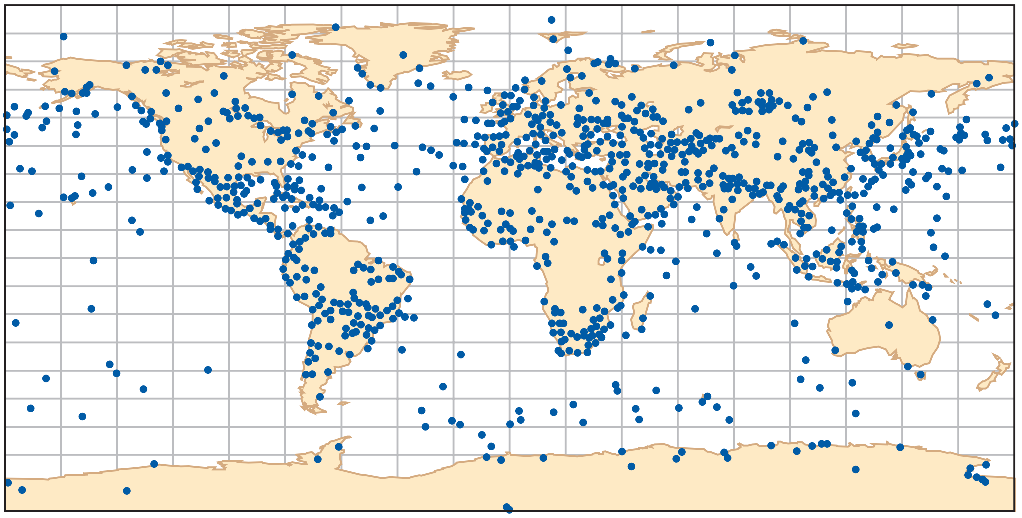


Figure 9 Ps bias corrected stations for the 12 UTC on 29 May 2006 analysis cycle.

 doi:10.21957/uv295rfmx5

Appendix

The OI adaptive bias correction method

The new bias estimate B_n is found as a linear combination between previously estimated bias B_p and new observation departure D_n so that:

$$B_n = W_p D_n + (1 - W_p) B_p$$

where W_p is the previous bias interpolation weight calculated at the previous observation departure occurrence D_p. The new bias interpolation weight Wn to be used for the next departure occurrence is calculated as:

$$W_n = \frac{\sigma_{bn}^2}{\left(\sigma_{bn}^2 + \sigma_{on}^2\right)}$$

where σ_{bn}^2 and σ_{on}^2 are new bias estimate and observation variances, respectively. They are calculated in a two step procedure. In the first step an intermediate or "guess" bias estimate variance σ_{bg}^2 and observation variance σ_{og}^2 are found from:

$$\sigma_{bg}^2 = \frac{\left(D_n - B_n\right) - \left(D_p - B_p\right)^{\ 2}}{C}$$

$$\sigma_{og}^2 = \min\ D_n^2, \left(D_n - B_n\right)^2$$

where C is a constant (=16). In the second step the final variances are calculated:

$$\sigma_{bn}^2 = W_p \sigma_{bg}^2 + \left(1 - W_p\right) \sigma_{bp}^2$$

$$\sigma_{on}^2 = W_c \sigma_{og}^2 + \left(1 - W_c\right) \sigma_{op}^2$$

where σ_{bp}^2 and σ_{op}^2 are the previous bias estimate and observation variances, and W_c is a constant interpolation weight (=0.010).

doi:10.21957/uv295rfmx5