# Handling Biases in Surface Pressure (Ps) Observations in Data Assimilation

**Drasko Vasiljevic**

*ECMWF, Shinfield Park, Reading, RG2 9AX, UK*

## 1. Introduction

It is well known that the surface pressure (Ps) observations reported by a large number of SYNOP (both land and sea based) and DRIBU stations are biased, by several hPa in many cases. The biases are mostly related to incorrect assumptions about the station altitude, and remain fairly constant in time. Several hundred stations currently appear on the ECMWF blacklist due to a significant long-term bias.

## 2. Bias Correction Methods

A few years ago Anders Persson and Peter Janssen each proposed suitable adaptive methods for bias correction of Ps time series. The two methods are based on linear estimation theory and are referred to as Kalman and OI methods, respectively. They are essentially similar in that they provide estimates of the bias and its confidence station by station, based on time series of observation-minus-background departures (O-M). Both methods rely on two assumptions. First, that the observational Ps bias is local for a given station (assuming no spatial correlation) and second that there is no, or small, model bias.

### 2.1. OI Method

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

$$B_n = W_p D_n + \left(1 - W_p\right) B_p$$

Where $W_p$ is the previous bias interpolation weight calculated at the previous observation departure occurrence $D_p$. The new bias interpolation weight $W_n$ 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 $\sigma_{bn}^{\ 2}$ and $\sigma_{on}^{\ 2}$ 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 $\sigma_{bg}^{\ 2}$ and observation $\sigma^2_{og}$ variances are found:

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

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

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

## 2.2. Kalman Method

Firstly, an intermediate or "first guess" bias estimate $B_g$ and its variance $\sigma_{bg}{}^2$ are found (one step ahead):

$$B_g = AB_p$$

Where, $B_p$ is the previous bias estimate and A is a coefficient (=0.999 if data missing and otherwise =1). Before the new bias estimate is calculated a "system performance constant" C is calculated:

$$C = F\left(D_n - B_g\right)^2$$

Where, $F$ is a "fading" term to simulate memory (=0.001) and $D_n$ is new observation departure. In the next step a "guess" value of the bias estimate variance is found:

$$\sigma_{bg}{}^2 = A^2 \sigma_{bp}{}^2 + C$$

Where, $\sigma_{bp}{}^2$ is the previous bias estimate variance. The new bias estimate $B_n$ is then found out:

$$B_n = B_g + W_n\left(D_n - B_g\right)$$

Where, $W_n$ the new bias interpolation weight:

$$W_n = \frac{\sigma_{bg}{}^2}{\left(\sigma_{bg}{}^2 + \sigma_o{}^2\right)}$$

Where, $\sigma_o{}^2$ is the observation variance (kept constant). And finally, the new bias estimate variance is calculated:

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

# 3. Practical Implementation

Based on the two above explained correction methods a practical scheme for estimating and correcting Ps bias has been developed and was introduced in the ECMWF operational analysis/forecasting system on the 5$^{th}$ of April 2005. A central component of this scheme is its database, PSBIAS. The PSBIAS is a hierarchical database modelled on the ECMWF operationally used ODB (Observation Data Base) database 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 a header and a 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. Stations for which either station altitude is missing or the height difference between station and model is more than 200m are omitted. Furthermore, Ps observation is discarded if observation departure is bigger then 15hPa or reported pressure is PMSL but station is above 500m. Bias estimates are calculated only if 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 a possible small model biases, bias estimates less then 1hPa are not considered. Also, too old bias estimates are not used. Currently, bias estimates older then 5 days are considered to be old. In that case they are recalculated from scratch (cold start). Furthermore, both Kalman and OI bias correction estimates as well as their related parameters are calculated and stored in the PSBIAS; both proposed schemes are run in parallel.

Figure 1 shows Ps departure (blue) and bias correction time series. Both Kalman (red) and OI (yellow) bias estimates are shown. This is for SYNOP station 65355 for February-April 2005 period. As it can be seen the station is about -11hPa biased. Both methods clearly identified the bias. However, the Kalman method bias estimate is “following” the Ps departure (blue) too closely; thus if applied would kill most of the signal. On the other hand, the OI method appears to be a somewhat smother operator. Based on examples like this it has been decided to use the OI as preferred method. It would be fair to add at this point that attempts to tune the Kalman method to be broader operator have been made. This was done by modifying the “system performance constant” C (see Section 2.2). However, the results obtained were still poorer compared with the OI method.

Figure 2 shows the resulting bias corrected Ps departure (red) along with the original Ps departure (blue) and applied Ps bias correct ion (yellow) for the same SYNOP station. As it can be seen, to start with, there is no bias correction because the sample size is not big enough (warm up period). Once the bias correction kicked in, it was correcting departures quite nicely. As it can also be seen at the beginning the scheme was turning itself on and off a few times before settling in. This station’s clear stable long-term bias was interrupted couple of times during this period. This was due to 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” runs, it was found out that in the “no Ps bias correction” run Ps was flagged and rejected, whereas in the “Ps bias correction” run 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, too. Blacklisted stations will be monitored after being bias corrected and eventually taken off the blacklist.

Figure 3 shows another example of Ps biased station. However, in this case the station height is thought to be correct and the real reason for being biased is not known. Again, after the warm up period the bias correction scheme turned itself on and was correcting Ps by about -3hPa. After closer examination of this station it was found out that in the “no Ps bias correction” run Ps was surviving (just) the first guess check but to be rejected by the analysis check. On the other hand in the “Ps bias correction” run it passed all the checks and was used in the analysis. However, there were two occasions when it failed the first guess check. The first time it 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 departure was still too big; hence rejected.

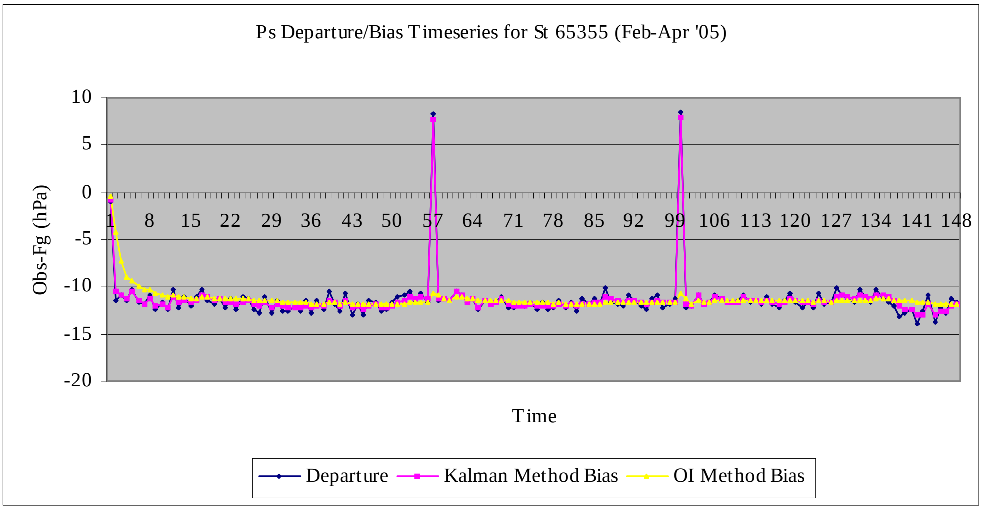


*Figure 1: Ps departure/bias time series (Feb-Apr 2005) for SYNOP station 65355; Departure (blue), Kalman method bias estimate (red) and OI method bias estimate (yellow).*

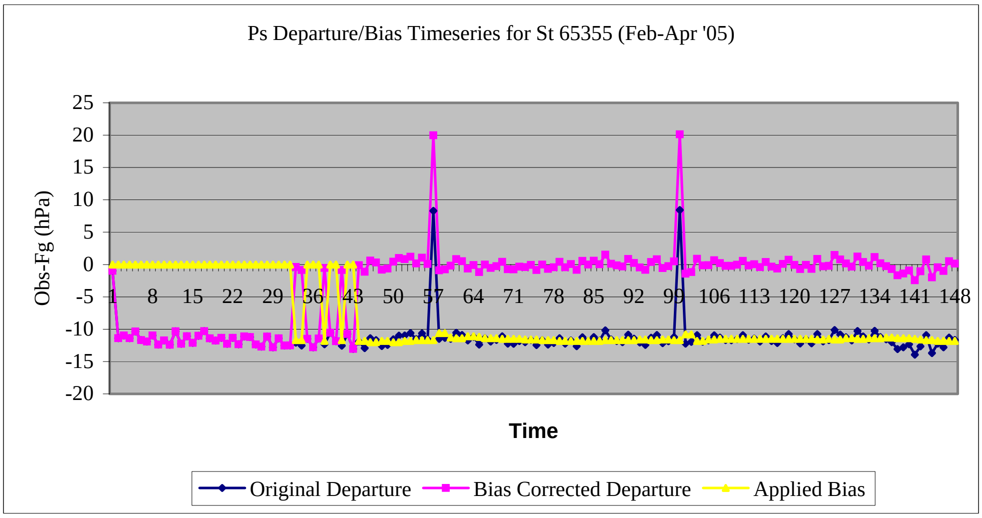


*Figure 2: Ps departure/bias time series (Feb-Apr 2005) for SYNOP station 65355; Original Ps departure (blue), bias corrected Ps departure (red) and applied Ps bias (yellow).*

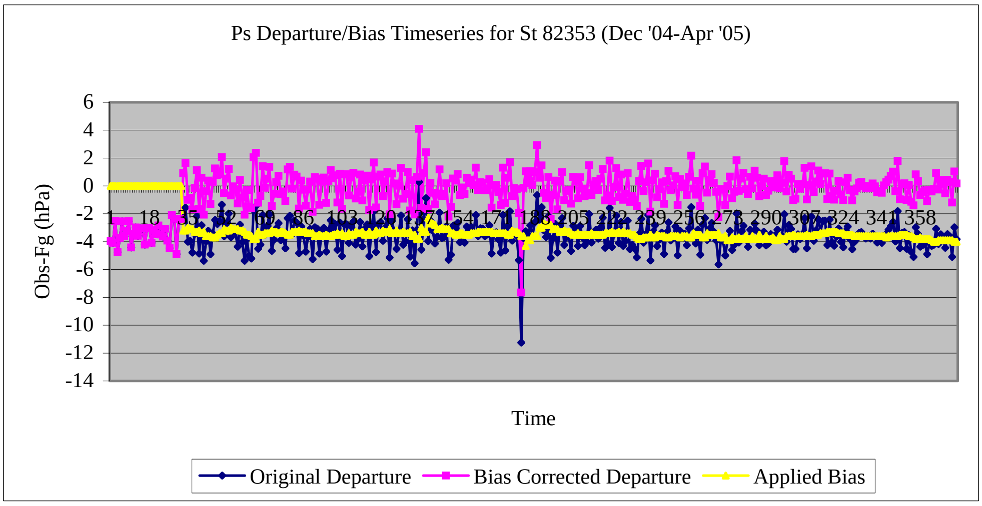


*Figure 3: Ps departure/bias time series (December 2004-April 2005) for SYNOP station 82353; Original Ps departure (blue), bias corrected Ps departure (red) and applied Ps bias (yellow).*

## 4. Impact on Analysis and Forecast

Prior to the Ps bias correction scheme operational implementation a number of shorter (one week long) experiments with and without Ps bias correction have been performed. They all have shown expected results. In brief, at the end of one week periods there would be several hundred bias corrected stations. Analysis increments would be somewhat smaller and the forecast scores would be neutral. Eventually, E-suite or E-suite type experiments were carried out. First, there was a long E-suite type experiment with Ps bias correction switched on. This experiment started on the 1st of August 2004 and went on till the 31st of December 2004. Furthermore, the E-suite was then run from the 1st of January 2005 till the operational implementation on the 5th of April 2005. Also, in August there was an experiment without the Ps bias correction shadowing the E-suite type experiment. Comparing these two experiments there were no surprises coming from the Ps bias correction. During August there were about 800-1000 identified biased stations and the correction scheme itself seemed to have performed satisfactorily. There was overall better fit to Ps observations with average analysis increments slightly reduced. The forecast impact in terms of 500hPa geopotential anomaly correlation scores for 10-day forecast for August is shown in Figure 4 for both the Northern (left) and Southern (right) hemispheres. As it can be seen they are neutral for the Northern and slightly positive in the Southern hemisphere.

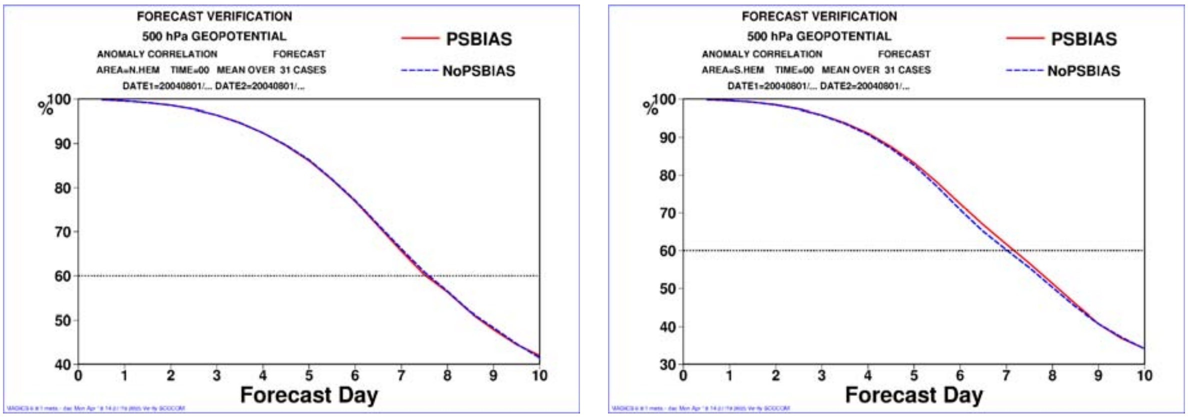


*Figure 4: "Ps Bias Correction" (red) and "No Ps bias Correction" (blue) 10-day forecast 500hPa geopotential anomaly correlation scores; 20040801-20040831 (00Z); Northern (left) and Southern (right) hemispheres.*

As already said the E-suite type experiment went on till the end of December 2004. Unfortunately, the forecast scores did not look that good for December. In the Southern hemisphere they were about neutral, but rather negative in the Northern hemisphere. The negative scores were obtained for the North Pacific, North America, North Atlantic and Europe, too. It is for that reason that a new December experiment with the Ps bias correction turned off was carried out. This new experiment was started from the already running E-suite type experiment. The PMSL analysis differences between these two experiments for the very first analysis cycle are shown in Figure 5. As it can be seen, most differences are small and local as expected. However, there is a rather large-scale positive difference over the North America (Canada/USA). This certainly was not expected and merited further investigation. Looking at the daily PMSL analysis difference between these two experiments, it was found out that just identified "**BLOB**" stayed there for about two weeks just moving around slightly. Furthermore, the "**BLOB**" survived in the ensuing forecasts and propagated downstream with the flow contributing to the bad forecast scores (not shown). 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 PBSIAS 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 1000 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 -1hPa to -3hPa. Thus, all stations were negatively biased. Bearing in mind that our first assumption when designing the Ps bias correction scheme was that observational Ps bias is not spatially correlated; hence, the large scale pattern we are seeing was not expected. 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.

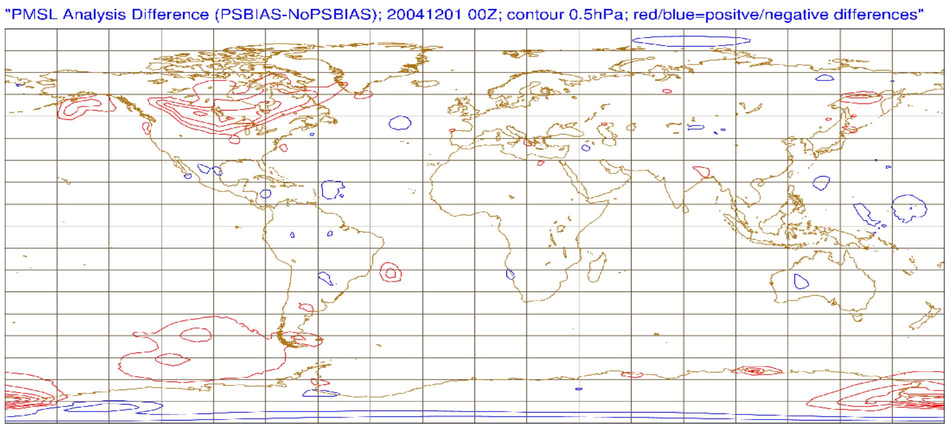


*Figure 5: PMSL analysis difference ("Ps Bias Correction"-"No Ps Bias Correction"); 20041201 00Z; positive/negative=red/blue; 0.5hPa contouring interval.*

Figure 6 shows the background and analysis fit to SYNOP Ps observations in the region. The black line represents "Ps bias correction" run fit to data, whereas the red line represents the "no Ps bias correction" run fit. From the "no Ps bias correction" run it is clear that there was about -2hPa bias. Furthermore, Figure 6 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 6 also shows is that both analyses drew closely for the observations. With one big difference, that "Ps bias correction" run resulting analysis creates, as earlier referred "**BLOB**" which subsequently impacts (negatively) the forecast quality. At this point it became clear that a further investigation was needed.

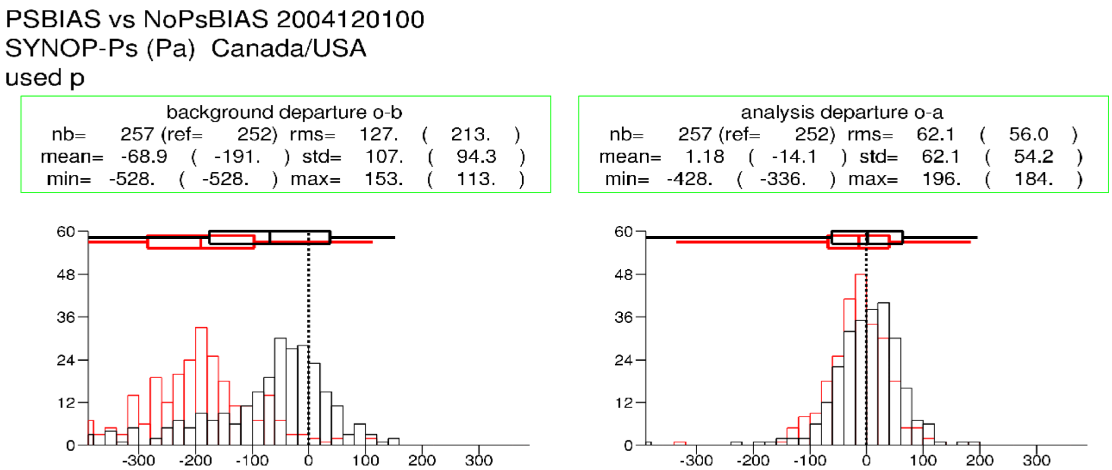


*Figure 6: Background (Bg) and Analysis (An) fit to SYNOP Ps data; "Ps Bias Correction" (black) and "No Ps Bias Correction" (red) experiments; 20041201 00Z; Canada/USA; unit Pa.*

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. Figure 7 is the E-suite type experiment (top) and the operational (bottom) time series for SYNOP data, from August till December 2004, respectively. Both time series are clearly showing that there was not much bias in the Canada/USA region to start with, from August till late September. However, as from late September bias started creeping in and growing. As seen earlier on, the bias correction scheme was recognising and correcting it. Since there was no bias correction in the operations at the time the size of the bias was somewhat bigger.

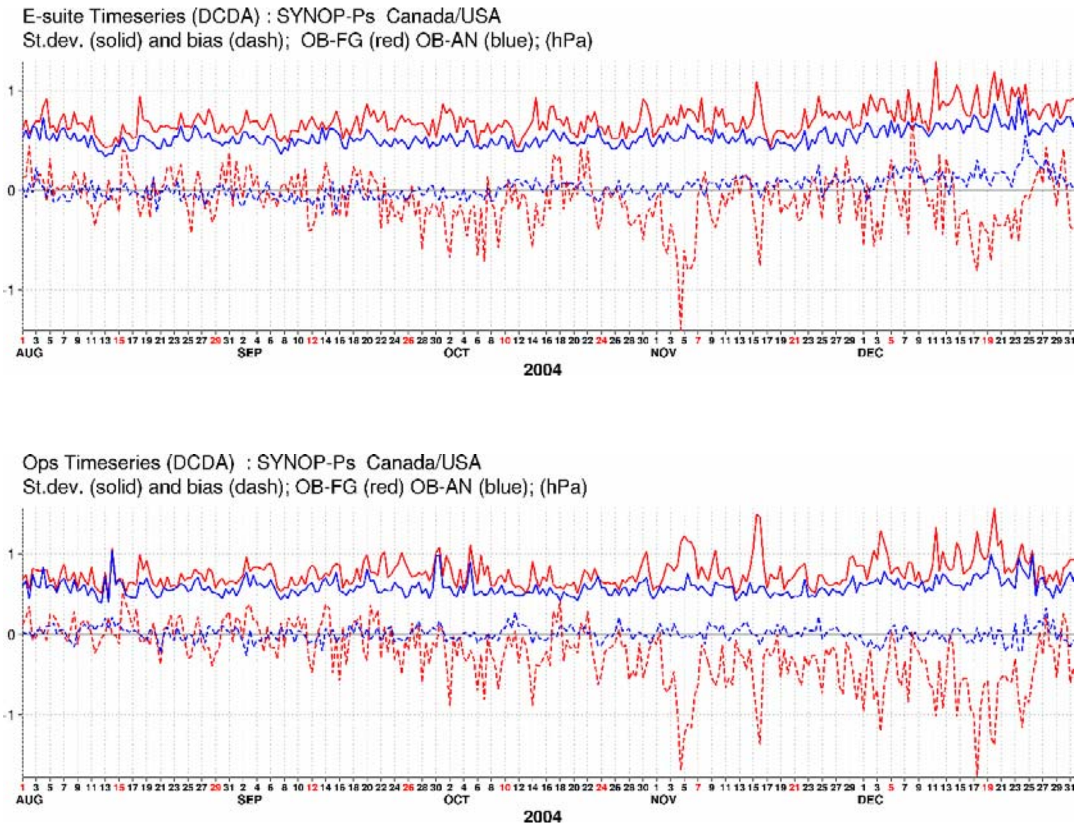


*Figure 7: SYNOP Ps time series; Background(red)/Analysis(blue) departures(dash)/standard deviation(solid); 2004080100-2004123112; Canada/USA; unit hPa; E-suite (top) and Operations (bottom).*

Now it started looking as we might be dealing with a possible model bias here. If true, the just described bias correction scheme should not be applied then. 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 1hPa limit when to apply 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 the Canada/USA where a similar problem could be found.

In order to facilitate understanding this problem further, a time series of max, min 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 max and min differences vary day by day, whereas the average differences are stable and around 0hPa. Time series of PMSL analysis differences between E-suite type experiment and operations for August-December period for the Northern hemisphere are shown in Figure 8. Red and blue lines are max and min differences, respectively. The black line is the average difference. As expected the max and min differences vary day by day, and the average difference stays rather stable and just hovering around 0hPa. A very similar picture (not shown) is for the Southern hemisphere. Furthermore, PMSL analysis differences for Europe also looked as good as the hemispheric ones. However, Figure 9 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 1hPa, but still reasonable. However, from late September till the end of December the average differences are far from stable and are as big as 2hPa, or even bigger. This result seems to be coinciding with the large-scale Ps bias noticed earlier. Unfortunately, we did no have "no Ps bias correction" experiment for the whole of August-December period. But there was one mentioned earlier during December. The average differences were not that stable in the "no Ps bias correction" run either, but the amplitudes are not as big as in the in the "Ps bias correction" run, just going over 1hPa.

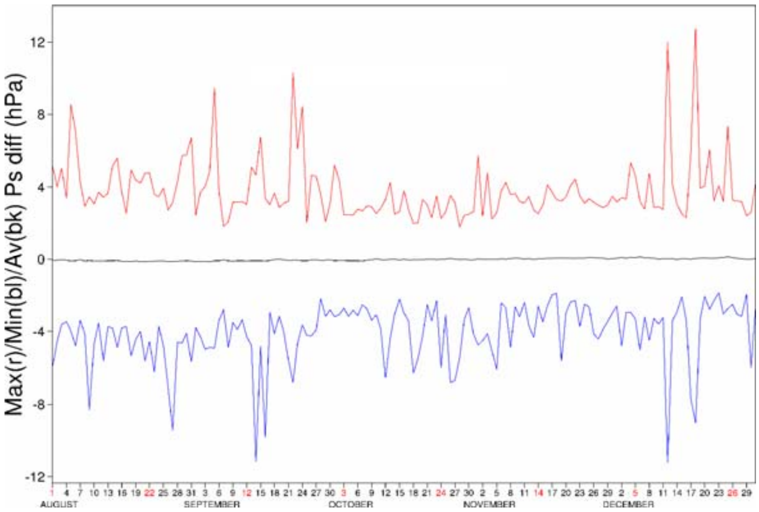


*Figure 8: Max (red), Min (blue) and Average (black) PMSL analysis difference ("Ps Bias Correction"-"Ops"); 20040801-20041231 (00Z); Northern Hemisphere; unit hPa.*

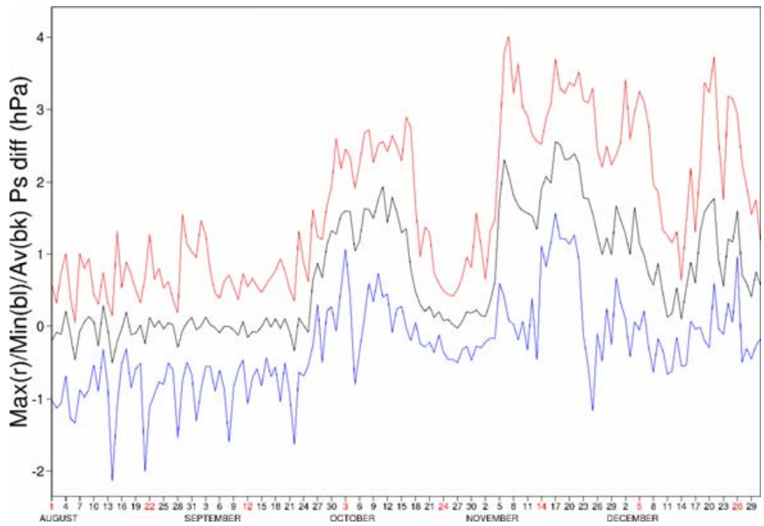


*Figure 9: Max (red), Min (blue) and Average (black) PMSL Analysis Difference ("Ps Bias Correction"-"Ops"); 20040801-20041231 (00Z); Canada/USA; unit hPa.*

Meanwhile, as the E-suite (started on the 1st 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, here, we are dealing with unexpected model bias. The bias correction scheme presented and introduced here is supposed to deal with uncorrelated observational Ps bias only. Thus in the presence of a larger scale model bias the correction should not be applied. The following Section 5 will deal with possible ways of identifying and separating model bias.

## 5. Model and Observational Bias Separation

As just discussed in the previous Section 4 correcting model bias by correcting observations leads to a rather poor result and it should not be done. The difficulty here is how to identify the model bias and subsequently separate it from the observational one. It is worth remembering that we introduced 1hPa limit when to apply the bias correction in order to avoid correcting for small model bias. Of course that limit could now be increased to lets say 2hPa. This rather quick fix should hopefully eliminate noticed problem over the 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 we should seek a little bit more selective solution to this problem.

The analysis of the problem presented in the previous Section 4 was clearly pointing out that when it happens a large number of stations are in agreement in terms of both the bias sign and the bias size. Now if we remember the old ECMWF OI analysis system where we used to have one quality control procedure, not used in the current ECMWF analysis system, called the “Buddy” check. 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 datum 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 +/- a multiple of standard deviation. Not to forget, 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. First, 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. Secondly, instead of using the mean as the analysed bias value a type of 2D univariate bias analysis can be done.

The just explained “Anti Buddy” check has been added to the Ps bias correction scheme. First, as originally, a list of potentially biased stations is compiled then for each of them the “Anti Buddy” check is applied. The circle radius around a given station is set to 300km. Number of influencing stations in order 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 analysed bias value +/- 2.0 standard deviations. Stations which do not have enough neighbours are not subjected to this check.

Now we reran the analysis cycle for which earlier we saw the large-scale positive PMSL analysis difference or the “**BLOB**”. Figure 10 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 it 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 impact only in the problematic region, and both the sign and the size of difference are good.

Furthermore, looking at Figure 11 one can see that we are not correcting as much of the bias as we used to do. The fit to both SYNOP and METAR (not shown) Ps observation in the area is now similar for two runs. Certainly this was the result we were hoping to see.

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? Such an experiment was run. Furthermore, an experiment where the limit of 1hPa when to apply bias correction was increased to 2hPa was run, too. Figure 12 shows December 500hPa geopotential 10-day forecast anomaly correlation scores for both hemispheres as well as for the North Atlantic and Europe.

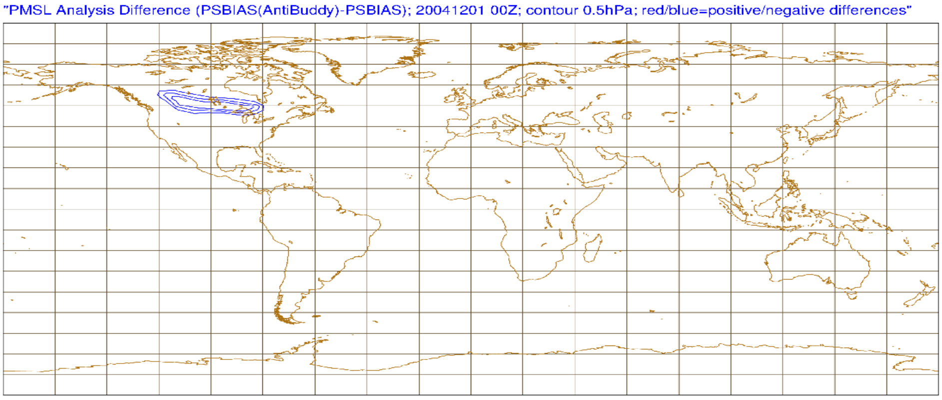


*Figure 10: PMSL analysis difference ("Ps Bias Correction plus Anti Buddy check"-original "Ps Bias Correction"); 20041201 00Z; positive/negative=red/blue; 0.5hPa contouring interval.*

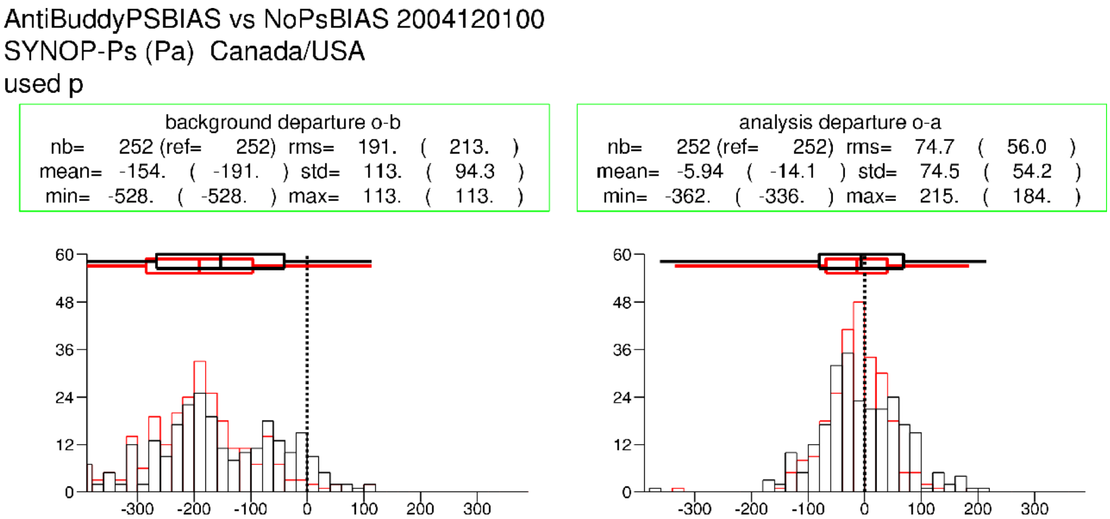


*Figure 11: Background (Bg) and Analysis(An) fit to Ps SYNOP data; "Ps Bias Correction plus Anti Buddy check" (black) and "No Ps Bias Correction" (red) experiments; 20041201 00Z.*

Depicted scores are for four experiments: the original "Ps bias correction" (green), "no "Ps bias correction" (brown), "Ps bias correction plus 2hPa limit" (blue) and the "Ps bias correction plus Anti Buddy check" (red). Now we can actually see how the original "Ps bias correction" underperformed the "no Ps bias correction" experiment. This is particularly the case in the Northern hemisphere, whereas in the Southern hemisphere it is more or less neutral in this period. The extent of the under performance in the Northern hemisphere can, also, easily be seen in all regional scores presented. The "Ps bias correction plus Anti Buddy check" and "Ps bias correction plus 2hPa limit" perform in the Northern hemisphere have done better then the original "Ps bias correction" experiment, and in the Southern hemisphere, where there was not much of the problem anyway, they are pretty much neutral. Despite the noticed improvement in the Northern hemisphere they both still underperformed the "no Ps bias correction" experiment. Turning our attention to the Northern hemisphere regional scores we can see that the "Ps bias correction plus Anti Buddy check" run came out on top. Thus, not only did it beat the "Ps bias correction plus 2hPa limit" experiment it was doing better then the "no Ps bias correction" experiment, too. Clearly, the "Anti Buddy" check had a positive impact on both the analysis and forecast. However, it did not go all the way to solving all of the December problems.

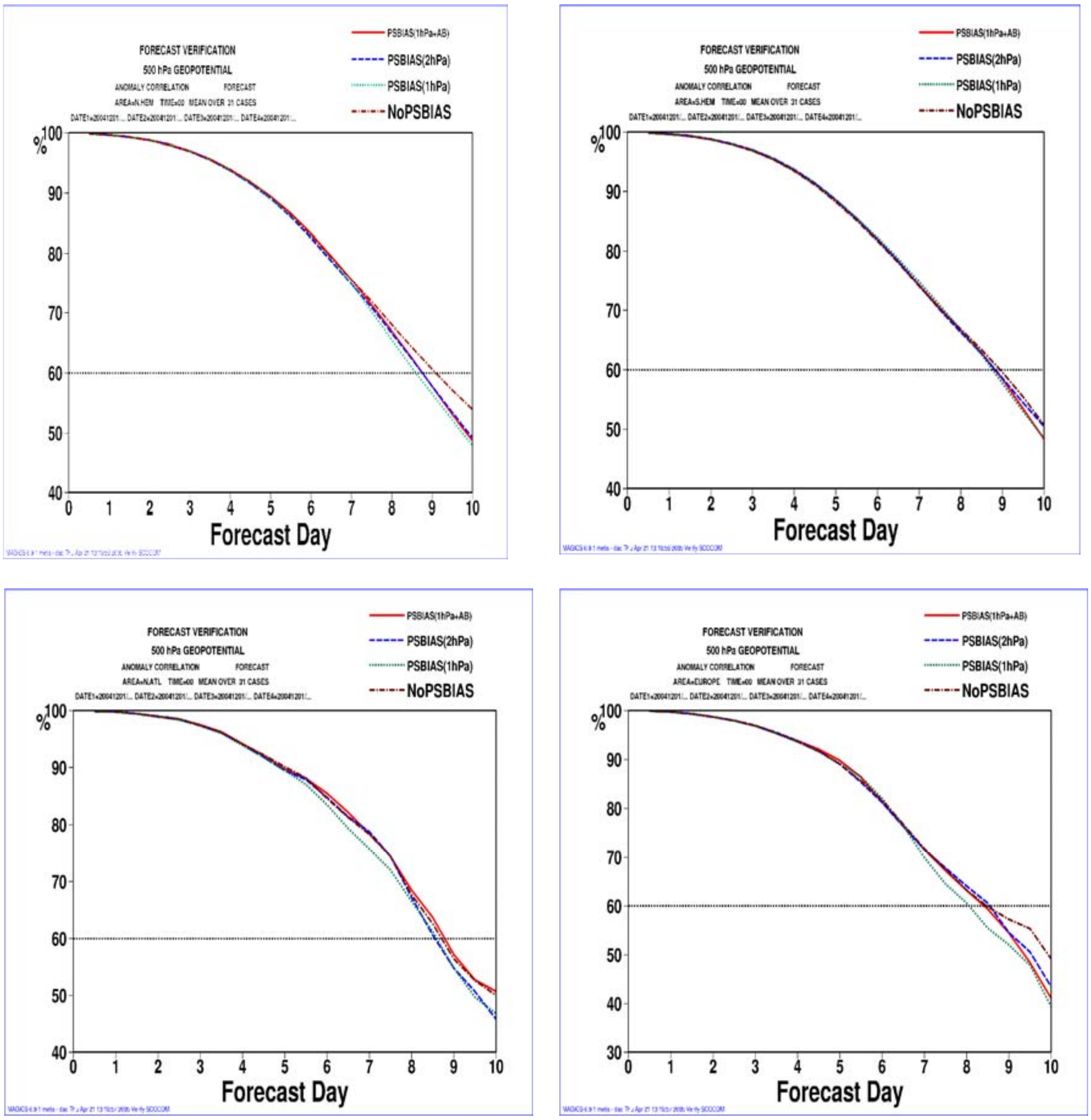


*Figure 12: "Ps Bias Correction plus Anti Buddy check" (red), "Ps Bias Correction plus 2hPa limit" (blue), "Ps Bias Correction plus 1hPa limit" (green) and "No Ps Bias Correction" (brown) 500hPa geopotential 10-day forecast anomaly correlation scores; 20040801-20040831 (00Z); Northern hemisphere (top left), Southern hemisphere (top right), North Atlantic (bottom left) and Europe (bottom right).*

## 6. Conclusion

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 between 800 and 1000 biased stations out of about 11000 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: (1) Ps bias is local (no spatial correlation) and (2) 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 a 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. It has been demonstrated that the proposed "Anti Buddy" check has managed to fulfil this requirement.