Proceedings of 8 ICSHMO, Foz do Iguaçu, Brazil, April 24-28, 2006, INPE, p. 1769-1771.

A SENSIBILITY STUDY OF THE INCLUSION OF THE CLOUD DRIFT WIND DATA IN THE CPTEC GLOBAL DATA ASSIMILATION SYSTEM

Sérgio H. S. Ferreira
José A. Aravéquia

Centro de Previsão do Tempo e Estudos Climáticos, Cachoeira Paulista, Brazil

1. Introduction

Since 2004, the Physical-Space Statistical Analysis System (PSAS) (Guo et al.,1998) has been operationally tested at CPTEC along with the Spectral Atmospheric Global Circulation Model CPTEC/COLA (AGCM). In this system, designated as GPSAS, was first used a data base composed by conventional meteorological data and Advanced TOVS data. However the data coverage was not sufficient to produce Analysis and Forecasts with the desired quality. In this work, Cloud Drift Wind Data (CDW) was added to the initial data based to improve quality analysis and forecast with GPSAS. The results with and without CDW data are compared in term of Root Mean Square Errors (RMS), BIAS errors, fields of pressure and winds and has been showed that the CDW data improves the representation of the eddies, troughs and ridges in the analysis and forecasts.

2. GPSAS overview

The GPSAS system obtains analysis and forecast fields in a cycle process that combines observational data with a short time forecast field (first-guess) from AGCM. The analysis obtained is used as initial condition in the next cycle. The equations below represent part of PSAS formulation.

$$\mathbf{w}^a = \mathbf{w}^f + \mathbf{K}\left(\mathbf{w}^o - \mathbf{H}\mathbf{w}^f\right)$$

$$\mathbf{K} = \mathbf{P}^f\mathbf{H}^t\left(\mathbf{H}\mathbf{P}^f\mathbf{H}^t + \mathbf{R}\right)^{-1}$$

In these equations $\mathbf{w}^a$ is the analysis matrix, $\mathbf{w}^f$ is the first-guess matrix, $\mathbf{K} \in \Re^n \times \Re^p$ is the gain matrix, which ascribes appropriate weights to the observations by acting on the innovation vector $\mathbf{w}^o - \mathbf{H}\mathbf{w}^f$. The $\mathbf{w}^o$ is the observation matrix and $\mathbf{H}$ represent an generic interpolate operator. The $\mathbf{K}$ matrix depends on the forecasts and observation matrixes errors ($\mathbf{P}$ and $\mathbf{R}$ respectively). An complete formulation of PSAS is made by Cohn et. All(98) and Da Silva and Guo (95). In this impact study is important emphasize that the matrix $\mathbf{P}$ is multivariate with the geopotencial height - wind. In that way, the inclusion of CDW, not only affected winds field but also the representation of geopotential height and sea level pressure fields. The PSAS assimilate all conventional meteorological data in terms of geopotencial height, wind vectors, sea level pressure and the specific humidity variables. The sea level pressure is used to transform ATOVS thickness data into heights and also to provide pseudo 1000 hPa heights for the upper air analysis. Some details about ATOVS thickness data are discussed by Ferreira (2004). The surface wind analyses are not currently used.

3. The additional CDW Data base

The additional CDW data base used in this work was received and processed at CPTEC during October of 2005. One part of this data is available through IDD (Internet Data Distribution) and is produced by J.M.A (Japan Meteorological Agency) and NOAA (Nacional Oceanic & Atmospheric Administration). The other part of this is produced locally at CPTEC Satellite Division (CPTEC-DSA). The figure 1 shows the coverage of each data source.

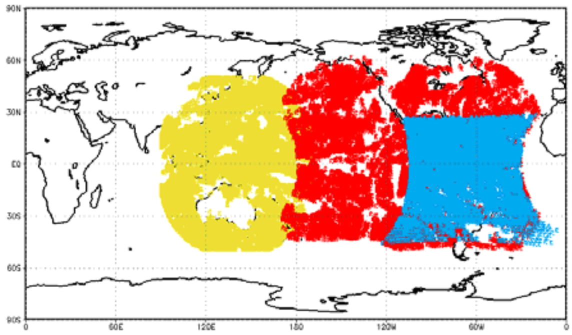

Figure 1 – CDW data assimilated in GPSAS at Dec 10th 2005 from a) JMA (Yellow) from NOAA (red) , CPTEC-DSA(blue)

The inclusion of CDW data represents 103681 new wind data assimilated by GPSAS during October of 2005. Most part of this data (60805) were assimilated in 850 hPa level and represent 73% of total wind data in this level, including radiosondes and AIREP data.

4. Quantitative impact

Many test of impact of data in data assimilation systems have been performed through comparison forecasts and analysis and supposing that the analysis fields represent the best estimative of real atmospheric state. It can be valid in regions where there is good coverage of observed data, but in Southern Hemisphere, where there is a lack observed data, the Analysis fields represents, for the most part, the first-guess fields from AGCM. It can product a biased impact evaluation.

In this work, we preferred to use observations data to representing the real state of atmosphere instead of analysis fields. The coverage of observed data is usually sparse and not uniform. Consequently the impact evaluation can not be representative of extensive area, but just representative around of geographic data points. On the other hands the errors involved in observation data are unbiased and uncorrelated with forecast models, providing in this way, an unbiased impact evaluation.

To perform this Impact evaluation, the models field was interpolated to the same point of observed data and the differences between observed data and interpolated values was used to calculate Bias Errors and RMS errors during October of 2005 in two regions: The first region is limited to 20N – 60N and corresponds the region of world that has the better coverage by conventional data. The second region is limited between 20N – 60S and represent the part of world where the coverage by conventional data is usually sparse.

The Impact results over RMS were obtained through differences between analysis and forecast which were produced with and without CDW data inclusion as is illustrated in the equation

$$I = 100 \cdot \frac{RMS(\text{withCDW}) - RMS(\text{withoutCDW})}{MSE(\text{withoutCDW})}$$

In this equation, if *RMS* (with CDW) is less than *RMS* (without CDW) the impact *I* is positive and indicate that CDW contributes to decrease the forecast errors. If *RMS* (with CDW) is bigger than *RMS*(without CDW) the impact *I* is negative and indicate that CDW contributes to increase the forecast errors.

The results was predominantly positives and more expressive in second region, showing that CDW data is an important kind of information in southern Hemisphere. As an example, the picture 1 shows the Impact of CDW data in RMS of geopotencial heights, winds vectors and specific humidity in the second region.

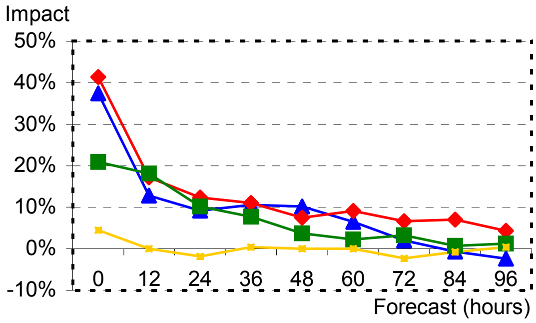

Figure1 - Impact of CTW inclusion in geopotential at 850hPa (green), wind speed at 850 hpa (blue), wind speed at 250 hPa (red), specifc humidity at 850 hPa (yellow)

Figure1 indicate that the impact begins around 40% for wind speed (850 and 250 hPa), 20% for geopotencial height and 5% for humidity at analyses field (0 hours). The positive impact in analysis wind fields was expected. The most important results was that this impact keeps positive in the next 72 hours of forecasts, indicating that CDW improve de quality of this forecasts. The positive impact in geopotencial heights confirm the positive impact of wind, and indicated that CDW data improve the quality of pressure and geopotential fields too. The impact was very small (positive or negative) in specific humidity filed. This field was not directly dependent of wind field as geopotential height is, so a smaller impact was expected in this field. However CDW should improve the moisture fluxes in AGCM. It was not observed probably because the coverage of moisture data is actually very sparse in GPSAS. The bias errors was estimate just for analysis and first-guess and showed that CDW decreases the Bias errors of wind in 41% in analysis and 50% in first-guess at 850hPa.

5. Qualitative impact in geopotential fields

The qualitative impact was estimated through many subjective comparisons between analysis with CDW data and without CTW data. As an example, the figure 2 shows an analysis of GPSAS and observed wind data over South America. In that figure can be seen that geopotential height isolines has good coherence with CDW and radiosonde winds. It well describes the troughs at 20S/40W, the height pressure center in the Pacific Ocean (at 35S/97W) and the Cyclone in the Atlantic (at 45S/45W). The figure 3, represent the analysis of GPSAS at same region and date, but without CDW data. In this case, can be seen that trough in the Atlantic was not represented. The aspect and position of the height pressure center at the Pacific and Cyclone in the Atlantic has been modified. These results confirm the positive impact observed in RMS errors and Bias errors

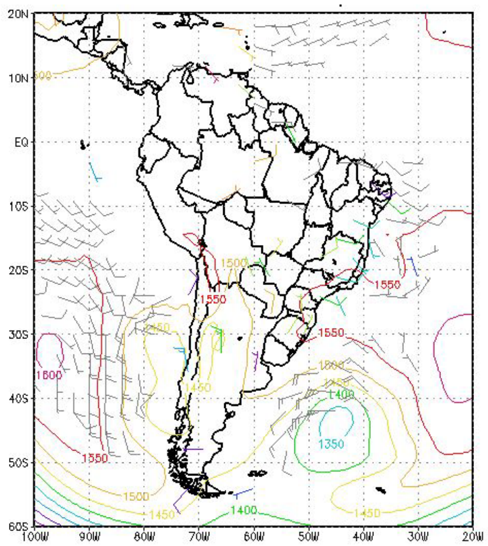

Figure 2 – GPSAS Analysis of Geopotencial Height at 850 hPa (isolines) and observed winds assimilated by GPSAS from radiosondes station (barb if colors) and CDW (gray barb)

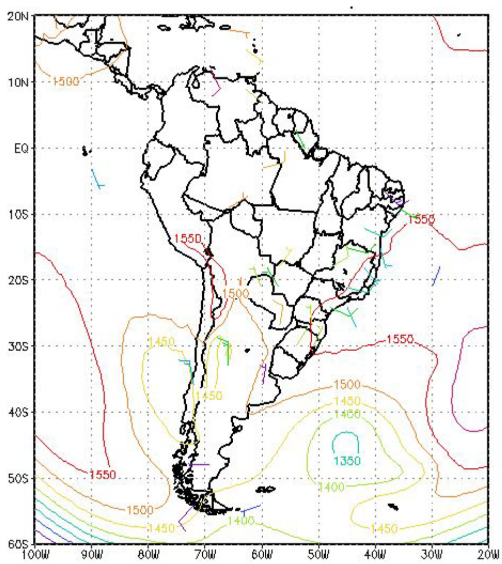

Figure 3 – GPSAS analise of Geopotential Height at 850 hPa (isolines) and radiossondes winds assimilated by GPSAS

5. Quality control and computational const aspects

In terms of quality of assimilate data, PSAS performs a statistical quality control that compare each suspected observational data with neighboring data. It process excludes those data that is probably incorrect or not represent synoptic characteristic. During October of 2005, just 1% of CTW data was rejected over oceans and 7% was rejected over continents. It's a good result if was consider the rejection of radiosondes data that was 5%. It was other indication of quality of CDW data.

The computational costs were estimate in term of CPU run times in supercomputer NEC SX6 using 4 processors. The inclusion of CDW increases the CPU run time from 20 to 30 minutes in mean terms.

In this preliminary tests ATOVS, surface and CDW from Water Vapor Channel data was thinned to decreases the CPU run time. Before this test another test was performed using the ATOVS in their full resolution. The CPU time was around 180 minute and the difference between analysis with and without thinning was not significantly to justify the use of those that in full resolution. In case of CDW, the use of this data in full resolution can increase the quality of analysis and forecast. At moment others tests with other thinner data option for CDW are been doing to establish the most adequate configuration in GPSAS

6. Conclusions

This paper presents the preliminary impact test with CDW inclusion in GPSAS data assimilation system. This test showed that CDW data have a large and positive impact in quality of analysis and forecasts. It decreases the RMS errors and improve the representation of the eddies, troughs and ridges. At moment more tests with CDW in GPSAS are been doing for operational use. The plans for near future includes the massively use if CDW data at higher resolution in GPSAS and RPSAS (regional PSAS) system, and the expansion of CDW coverage areas, produced by CPTEC DSA, to Africa region where other winds data are not available. In parallel of this study, QuikScat, ERS and GPS data are been tested to complete the data base necessary to operational use of GPSAS in CPTEC.

References

Cohn, S. E.; Silva, A. D.; Guo, J.; Sienkewicz, M.; Lamich, D. Assessing the effects of data selection with the DAO Phisical-space Statistical Analysis System. Monthly Weather Review, v. 126, p. 2913-2926, 1998.

Ferreira, S.H.S, 2004: Análise do procedimento de Assimilação de Sondagens Derivadas de Satélite. (INPE-114820TDI/956)

Guo, J., J. Larson, G. Gaspari, A. daSilva, P. Lyster, 1998. Documentation of the PSAS Part III: The Software Implementation. DAO *Office Note*, 98-05. [Available at http://dao.gsfc.nasa.gov/subpages/office- notes.html].