

# Usage of SYNOP 10 meters wind observations in HIRLAM

Alberto Cansado Auría<sup>1</sup>

*Instituto Nacional de Meteorología INM. Numerical Weather Prediction Department. Madrid.*

Beatriz Navascués Fernández-Victorio<sup>2</sup>

*Instituto Nacional de Meteorología INM. Numerical Weather Prediction Department. Madrid.*

## ABSTRACT

It is not easy to deal with 10 meters wind observations over land in NWP models. This is due mainly to the difficulty of the models to represent the real orography and physiography and the high degree of dependence of near surface wind observations with both, physiography and orography. Other source of errors in the analyzed 10 meters wind field is the usage of observations which are not representative of the neighbouring terrain. Here, we describe the method we used to, subjectively, identify *good* and *bad* wind stations, the way to introduce SYNOP 10 meters wind observations in HIRLAM using whitelists and the analysis of the experiments and case studies carried out.

---

## 1.- Introduction

Dealing with 10 meters wind observations over land in Numerical Weather Prediction Models is difficult. The main reason is the misrepresentation of both, physiography and orography in the models and the high degree of dependence of the near surface wind observations with the neighbouring orography and physiography. Obviously, the problem should be less important as models increase their resolution, but still high differences between model and observed 10 meters winds are found, especially in mountainous regions or in areas with complicated orography. In these cases, local winds used to be driven in the direction imposed by the orography and the speed of the wind is controlled by the orography and the physiography of the terrain. Another source of error is that the lowest model level are quite above 10 meters and there are, obviously, errors coming from the observation operator, not only in the interpolation algorithm (exchange coefficients) but also in the model stability, the values of the Richardson number ( $Ri$ ) or the roughness length which might be wrong.

Besides, some meteorological stations are situated in zones which are not representative of the neighbouring terrain (top of mountains, bottom of valleys, etc). In this case, they supply 10 meters wind information which, being correct, should not be assimilated. Finally, some of the analyzed stations contained errors (sometimes very evident) in wind direction, wind speed or both.

Big differences between model and observed SYNOP 10 meters wind will make the observations to be very easily rejected in the quality controls. The problem comes when innovations are not big enough to be rejected in the quality controls and the information of the observations is assimilated in the model. In this case, the quality of the forecasts can be compromised.

To prevent this to happen, the quality of the SYNOP 10 meters wind data have been controlled subjectively and only stations which reach a minimum quality are allowed to be assimilated in the model. The way to do it is by using the so-called whitelists.

In this paper, we first analyze the SYNOP 10 meters wind observations, provide a method to subjectively measure their quality and suggest an objective method (measuring the correlation coefficient and the slope of the linear regression) to determine their quality. Then, we describe the observation operator which interpolates from the model gridpoints to the observation location. The following section describes how to run HIRLAM (version 6.4.4) to assimilate SYNOP 10 m wind observations. The last sections cover the description of the experiments, the analysis of the results using the verification against observations, the analysis of the case studies and a brief description of a bias correction scheme to improve the analyzed wind field to end up with some conclusions.

---

<sup>1</sup> *Corresponding Author address:* Alberto Cansado Auría. Instituto Nacional de Meteorología (INM) Servicio de Modelización Numérica del Tiempo. Leonardo Prieto Castro 8. Ciudad Universitaria. E-28040 Madrid. E-mail: alberto.cansado@inm.es

<sup>2</sup> *Current adscription:* Instituto Nacional de Meteorología (INM). Área de Proyectos. Madrid.

## 2.- Analysis of the quality of the SYNOP 10-meters wind observations

Our first step consisted in identifying a set of stations which accomplish a minimum quality to be assimilated in the model.

To do this, we used all the observations ingested in the INM surface analysis during a period of two months (from 10<sup>th</sup> February to 15<sup>th</sup> April, 2005). This surface analysis has a resolution of 20 km and runs every three hours. It has been developed from the existing HIRLAM surface analysis and extended to MSL Pressure and 10m wind and it is intended as a diagnostic tool to support operational forecasting. We build three kinds of scatterplots for every SYNOP station (674 in total). In the first one, we plot direction of the model wind against direction of the observed wind. What we expect in a station that behaves well is that most of the points will be distributed in the diagonal. In the second one, we plot speed of the model wind against speed of the observed wind and, again, we expect a distribution of the points over the diagonal. Finally in the third one, we plot differences of the directions of the model wind and the observed wind against observed wind speed. In a good station we expect high differences in the directions of the model and observed winds when the speed is fairly low and decreasing differences in the wind directions as wind speed raises.

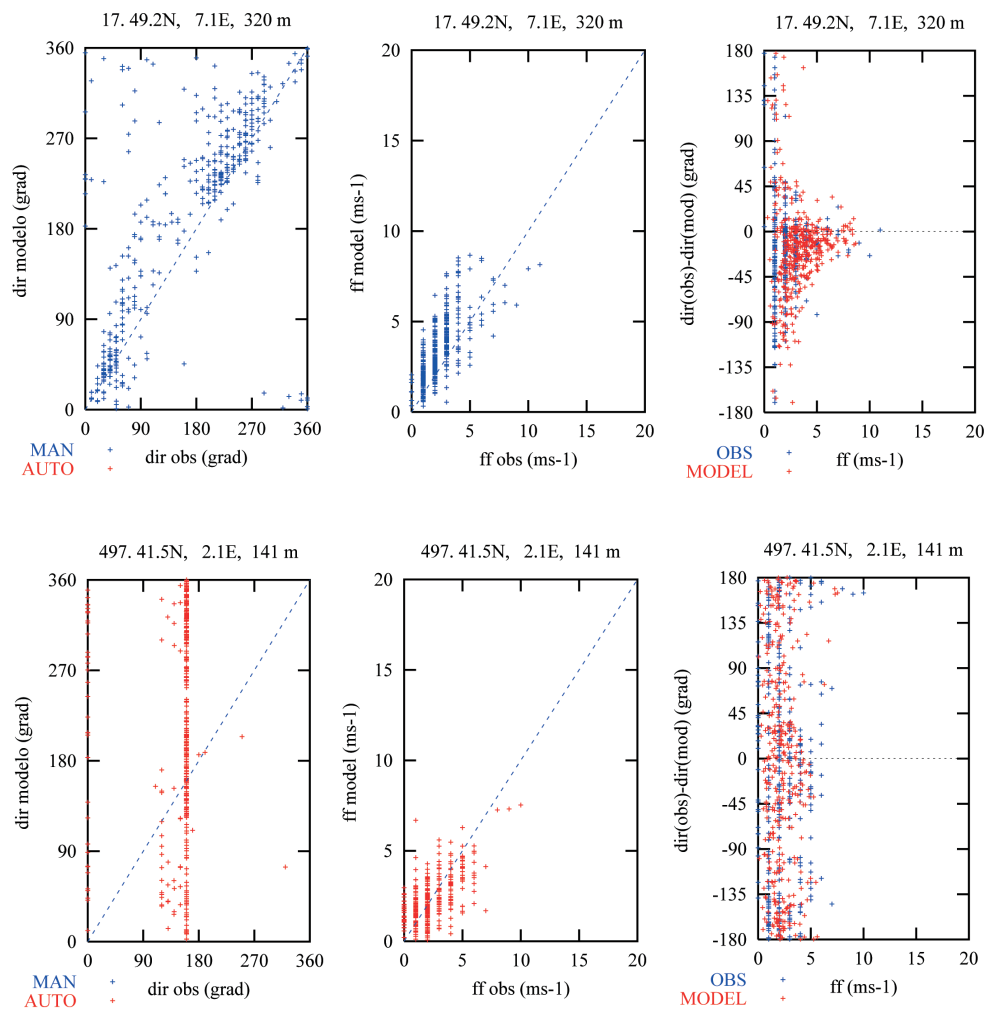


Figure 1. The three kind of scatterplots that we have built for every analyzed SYNOP station. Here we present two cases: A good station (17. Saarbruecken/Ensheim 10708) with an acceptable behaviour which has been included in the whitelist and another station (497. Sabadell 08192) that has been excluded from the whitelist

One thing we have observed is that the model seems to overestimate the wind speed when the wind is weak whereas the model underestimates the wind speed when the wind is strong. This seems to be a general behaviour in

most of the wind stations.

After having analyzed (visually, subjectively) the total set of plots, 290 SYNOP stations were selected to be assimilated by using a whitelist. Whitelists in HIRLAM work forcing to assimilate all the stations included in the whitelist even if they were rejected in the quality controls. The rest of the stations are assimilated only if they are able to pass all the quality controls of the screening steps. To assimilate only the stations included in the whitelist, what we did is setting the first guess check tolerances to zero to force all the stations to be rejected in the first guess check. This way, we only assimilate the stations included in the whitelist.

We are concerned that the right way to make a selection of the stations is through an objective procedure. That is the reason why we have started to test an objective procedure using correlation coefficients and the slope of the linear regression between observed and model wind components.

The linear regressions and correlation coefficients corresponding to three different stations (initially classified by our subjective procedure as good, intermediate and bad quality stations) have been calculated. We have done all the same using a) all the information from radiosondes at 1000 hPa assimilated on 16<sup>th</sup> January 2006 at 12 UTC in our operational run and b) all the information at 1000 hPa from the Coruña radiosonde (situated in the North Western coast of Spain) corresponding to the month of June 2005. We looked this radiosonde information as an upper limit in the quality of data (that is, high correlation coefficients between model and observed data and slope near to one).

What we have found is that we cannot expect correlation coefficients ( $R^2$ ) much above 0.7 which is, roughly, the limit we have found in radiosonde data at 1000 hPa. Another point is that when a station is subjectively classified as intermediate quality, the correlation coefficient ( $R^2$ ) lowers until 0.3. Bad stations have values of  $R^2$  near to zero. The values of the slope of the linear regressions are more variable but, roughly, they seem to be near one in the case of radiosonde and good behaved wind stations whereas seems to be further from one as the quality of the stations decreases. The results are shown in the following table.

	$R^2$ (u component)	Slope (u component)	$R^2$ (v component)	Slope (v component)
All Radiosondes 16 <sup>th</sup> Jan, 2006 12 UTC	0.6	0.83	0.66	1.06
Coruña Radiosonde June 2005	0.68	0.89	0.58	0.87
Good Quality Station (10708)	0.77	1.2	0.58	0.97
Intermediate Quality Station (07299)	0.32	0.57	0.52	0.61
Bad Quality Station (08192)	0.01	0.15	0.01	0.15

Table 1. Values of the correlation coefficient and of the slope of the linear regressions of observed and model wind components for the analyzed cases

We think that this method, together with the analysis of histograms of innovations to check if the errors are gaussian, can provide an objective procedure that is worth to be studied and applied in the future to the determination of the wind stations quality. Anyway, we recommend to build a new whitelist adapted to the domain of each user including quality SYNOP stations with 10m wind data before running an experiment.

### 3.- The observation operator

The observation operator is used to interpolate model values corresponding to the observed 10m wind ( $u_{10m}$ ,  $v_{10m}$ ). Here we used the observation operator for near surface parameters available in the HIRLAM Three Dimensional Variational Assimilation (3DVar) system. The observation operator is described in Gustafsson et al (1999) and it is based on the method described by Geleyn (1988) which is physically coherent with the Monin-Obukhov similarity theory for the surface layer.

The model estimates wind at the lowest model level ( $u_{nlev}$ ,  $v_{nlev}$ ) and at the model surface ( $u_s = 0$ ,  $v_s = 0$ ) and they interpolate, using the observation operator, to find the model estimated value at the measurement height ( $z$ ). The non-linear model equations for 10m wind are:

$$u(z) = z_{rat} u_{nlev}$$

$$v(z) = z_{rat} v_{nlev}$$

where  $z_{rat}$  depends on the Richardson number Ri (that is, on the stability) and on the Von Karman constant through the surface exchange coefficients for momentum  $C_m$  and

$$z_{rat} = \frac{1}{b_m} \left[ \ln \left( 1 + \frac{z}{z_{nlev}} (e^{b_n} - 1) \right) - f_v(z) \right]$$

with  $f_v(z)$  having two different functional definitions depending on the atmospheric stability

$$f_v(z) = \frac{z}{z_{nlev}} (b_n - b_m) \quad (stable)$$

$$f_v(z) = \ln \left[ 1 + \frac{z}{z_{nlev}} (e^{b_n - b_m} - 1) \right] \quad (unstable)$$

$b_n$  and  $b_m$  are defined as

$$b_n = \frac{\kappa}{\sqrt{C_n}}$$

$$b_m = \frac{\kappa}{\sqrt{C_m}}$$

$\kappa$  being the Von Karman constant ( $\approx 0.4$ ),  $C_n$  constant is defined as

$$C_n = \left( \frac{k}{\ln \left( 1 + \frac{z_{nlev}}{z_0} \right)} \right)^2$$

and the  $C_m$  constant, defined as depending on the stability

$$C_m = C_n \left( \frac{1}{1 + 2b \left( \frac{Ri}{\sqrt{1 + d Ri}} \right)} \right) \quad (stable)$$

$$C_m = C_n \left( 1 - \frac{2b Ri}{1 + 2b c C_n \sqrt{(1 + z_{nlev} / z_0)(-Ri)}} \right) \quad (unstable)$$

Where  $z_0$  is the roughness length;  $z_{nlev}$  is the height of the lowest model level;  $b$ ,  $c$  and  $d$  are constants ( $b=5$ ,  $c=5$  and  $d=5$ ) and the Richardson number ( $Ri$ ) is defined as

$$Ri = \frac{g \Delta \theta \Delta Z}{\theta (u^2 + v^2)}$$

being  $\theta$  the potential temperature,  $g$  the acceleration of gravity,  $u$  and  $v$  the wind components and  $Z$  the height.  $\Delta$  in front of a variable denotes the difference between the lowest model level and at the surface.

To create the interpolated values corresponding to the analysis increments a tangent linear model is created which is the linearization of the non-linear operator around the background field.

$$u(z)^{il} = z_{rat}^{il} u_{nlev} + z_{rat} u_{nlev}^{il}$$

$$v(z)^{il} = z_{rat}^{il} v_{nlev} + z_{rat} v_{nlev}^{il}$$

#### 4.- How to run HIRLAM with 10-meter wind observations

The only thing a HIRLAM user must do if he/she wants to assimilate 10 meter wind observations is to modify the variable LWIND10M at Env\_expdesc script. The default value for LWIND10M is *no*, meaning that no 10-meters wind observation would be used. To use them, LWIND10M must be set to *yes*. Changes are included in the recent 6.4.4 release

The default whitelist includes those 290 stations, but it is possible to modify the set of stations in the whitelist. In this case, it is necessary to leave the new list (called synop\_uv10m\_whitelist.dat) under the directory \$HL\_LIB/hirlam/data/hirvda/station\_data. When HIRLAM is running, it checks if there is a whitelist file in this directory and only if no whitelist is found, the default one will be used.

The format of the new whitelist must be the same that the default one: The first line contains the number of stations included in the whitelist. Below there is a list of the indicatives (one per line). The number of characters of every indicative is eight, leaving blanks on the left if necessary. Please, note that it is absolutely necessary respecting the format. Otherwise, the run will fail.

```
290
08238
10763
...
10655
10675
```

The code is ready to ingest as much as 500 stations. In the case that somebody needs to assimilate more than 500 wind stations, it will be necessary to modify the scrobs subroutine **scrcom\_default.F** increasing the value of the variable maxwhitelist.

#### 5.- Description of the experiments

3DVAR and 4DVAR experiments have been carried out using SYNOP 10 meter wind observations in both passive and active mode to check the impact of that kind of observations in the forecasts. All of them cover the period from 2004 April 25<sup>th</sup> until 2004 May 15<sup>th</sup> and have been done with 290 stations in the whitelist.

Experiment	Assimilation	Characteristics
S25	3DVAR	Active 10 meters Wind Observations. Statistical Balance
S28	3DVAR	Passive 10 meters Wind Observations. Statistical Balance
S23	4DVAR	Active 10 meters Wind Observations. Analytical Balance
S27	4DVAR	Passive 10 meters Wind Observations. Analytical Balance

Table 2. List of the main characteristics of the experiments carried out to test the SYNOP 10m Wind assimilation

The domain of integration is SMHI area S22 (NLAT=306; NLON=306; NLEV=40. 0.2 degrees resolution in both, latitude and longitude) modified to cover Southern Europe. The domain of integration of the experiments is shown in figure 2. All of them are performed using HIRLAM version 6.3.8.

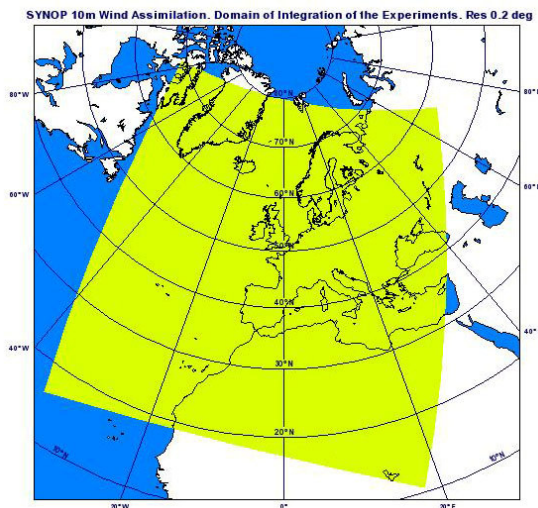


Figure 2. Domain of integration of the SYNOP 10m Wind Assimilation Experiments

The 3DVAR experiments have been done using the so-called statistical balance (Berre, 2000) in the background constraint term of the cost function. The statistical balance includes multivariate relations between forecast errors of humidity and those of mass and wind. It means that corrections in the wind field will affect the humidity field.

### 6. Analysis of the verification results

The result of the verification of the 3DVAR experiments shows improvements in wind speed at all levels (EWGLAN and ALL) and slight improvements in relative humidity in high and medium levels (EWGLAN). Very slight enworsening of scores in two meters temperature and relative humidity are observed. Finally, geopotential presents very slight changes (Figure 3). Surface parameters do not show any impact of assimilation of 10 m wind, but it has to be taken into account that no selection has been done in 10m winds during the verification.

## Verification against observations EXP: S25 S28

Time: 2004042500 - 2004051518 Domain: EWG Forecast from 00 06 12 18

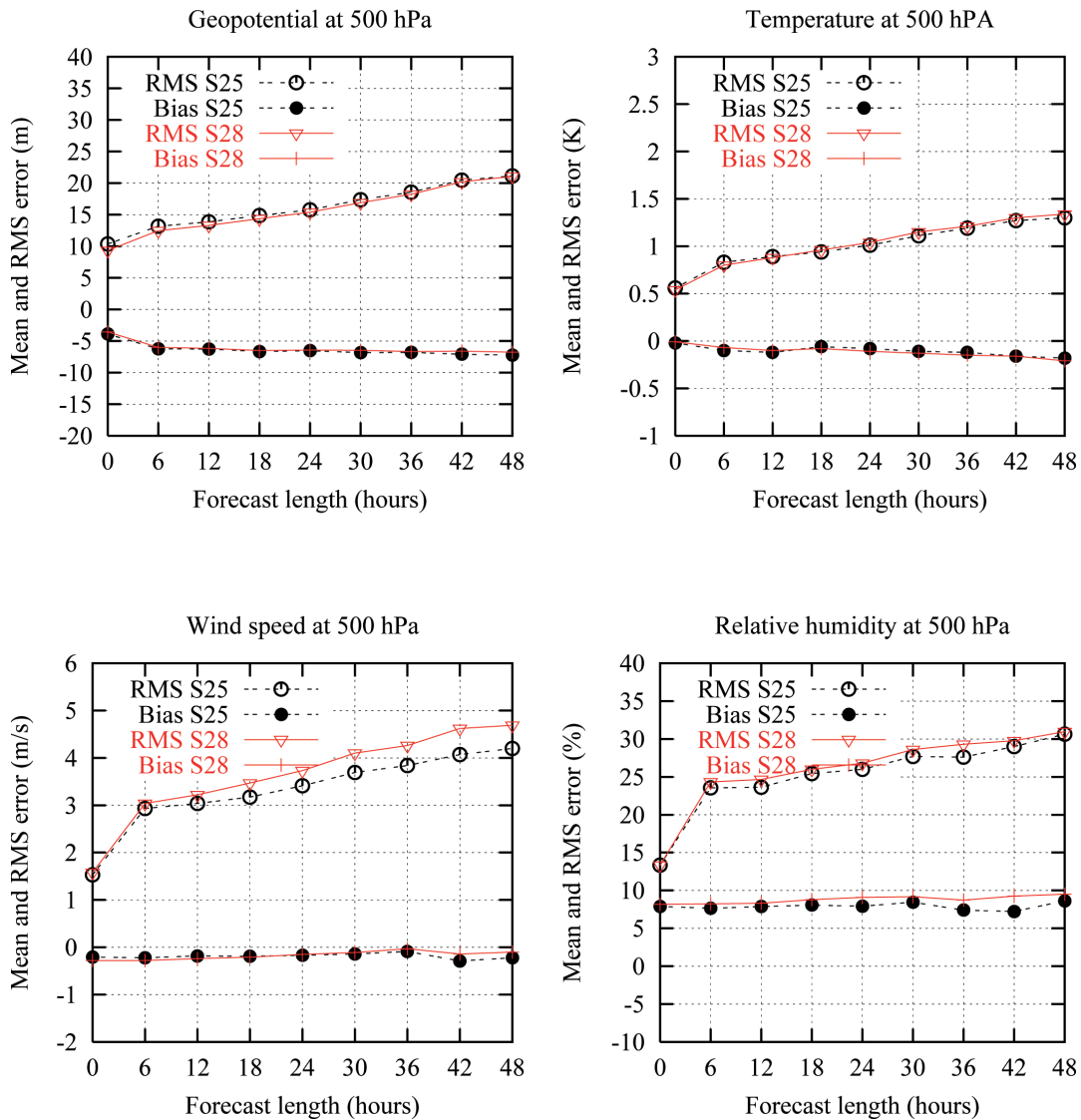


Figure 3. Results of the verification of 3DVAR experiments of 500 hPa parameters.

Histograms in both 3DVAR and 4DVAR, show a slight improvement in the innovations when 10 meters wind observations are used. We present below (Figures 4, 5, 6, and 7) the histograms of u and v components (separately) for the 3DVAR experiments (S25 and S28) and the 4DVAR experiments (S23 and S27)

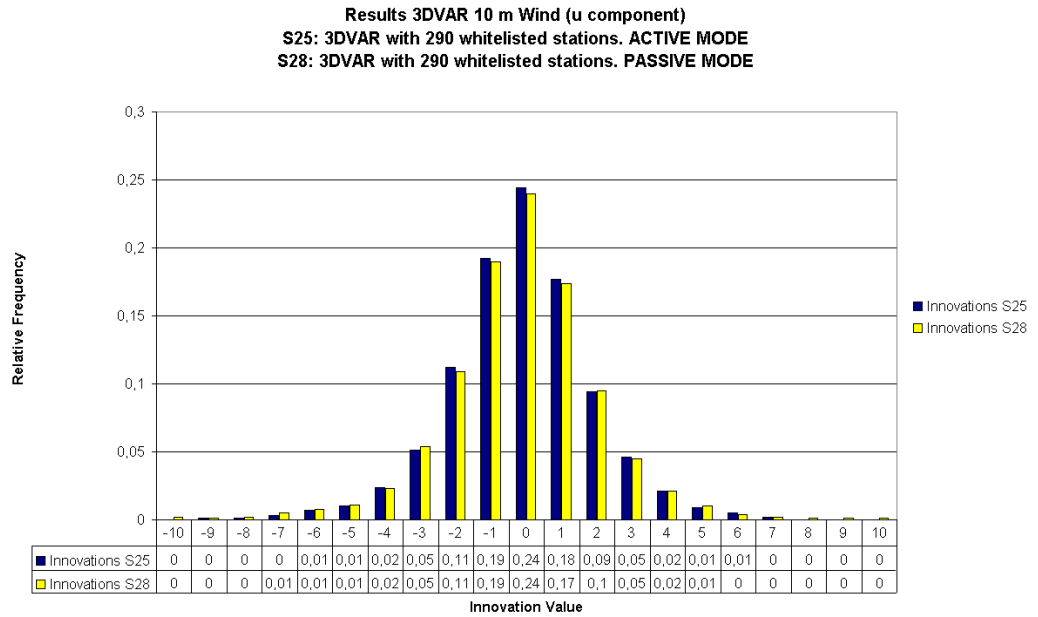


Figure 4. Histograms comparing relative frequency of innovations in the case of passive and active SYNOP 10m wind assimilation (u component) in the 3DVAR experiments.

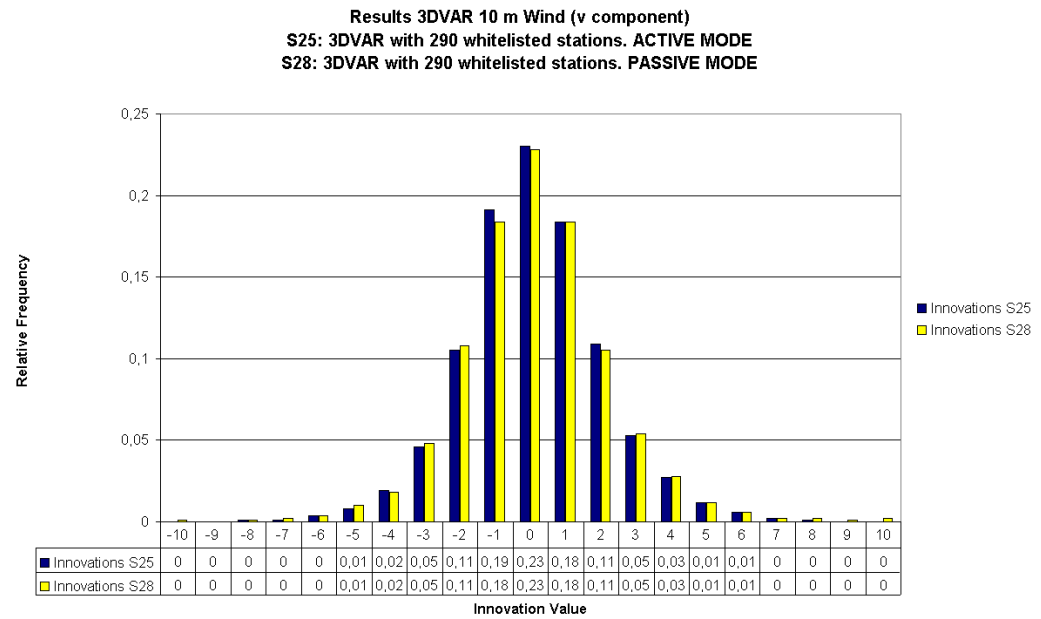


Figure 5. Histograms comparing relative frequency of innovations in the case of passive and active SYNOP 10m wind assimilation (v component) in the 3DVAR experiments.

**Results 4DVAR 10 m Wind (u component)**  
**S23: 4DVAR with 290 whitelisted stations. ACTIVE MODE**  
**S27: 4DVAR with 290 whitelisted stations. PASSIVE MODE**

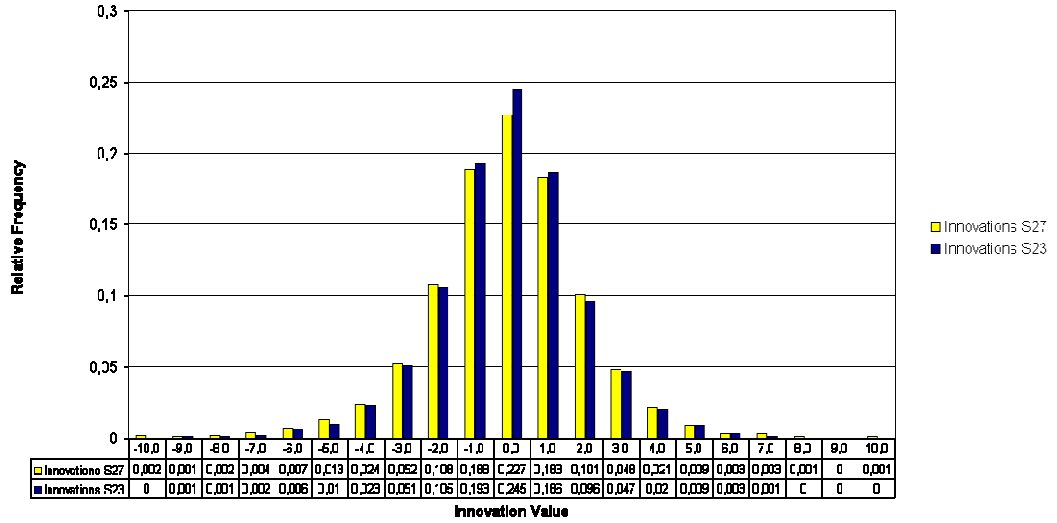


Figure 6. Histograms comparing relative frequency of innovations in the case of passive and active SYNOP 10m wind assimilation (u component) in the 4DVAR experiments.

**Results 4DVAR 10 m Wind (v component)**  
**S23: 4DVAR with 290 whitelisted stations. ACTIVE MODE**  
**S27: 4DVAR with 290 whitelisted stations. PASSIVE MODE**

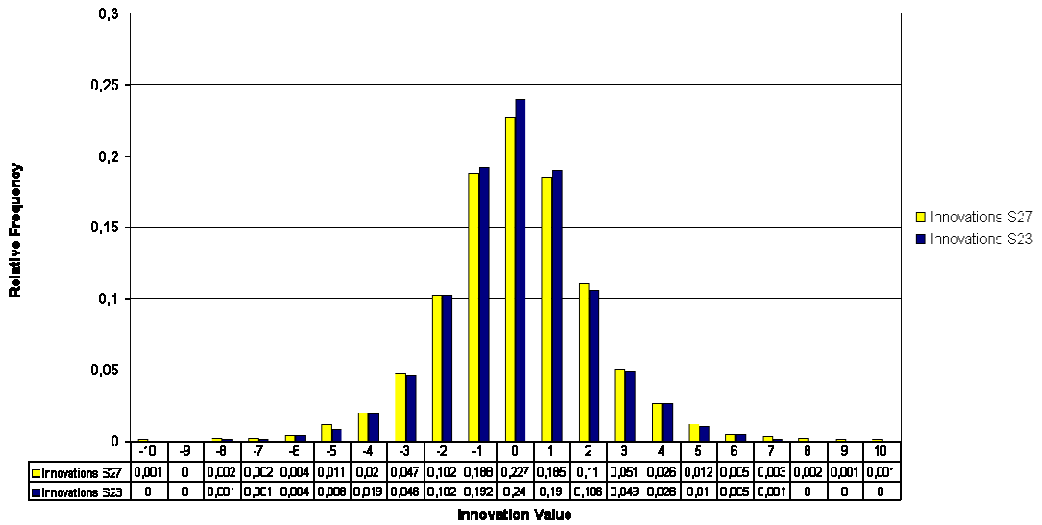


Figure 7. Histograms comparing relative frequency of innovations in the case of passive and active SYNOP 10m wind assimilation (v component) in the 4DVAR experiments.



## 7.- Case Studies

Two case studies with heavy precipitations were analyzed comparing H+24 HIRLAM forecasts and the observations from climatological INM rain gauge network. Cold colors (purple, blue) indicate no or few precipitation in 24 hours. Hot colors (orange, red) indicate much accumulated precipitation. The intervals are 0-2 (purple), 2-5 (dark blue), 5-10 (light blue), 10-15 (dark green), 15-20 (green), 20-30 (light green), 30-40 (light orange), 40-80 (dark orange) and 80-200 (red)

The first one had place on 2<sup>nd</sup> May 2004. The second one on 11<sup>th</sup> May 2004. Both of them lead to similar conclusions. Here, we present the second case study which focus on the heavy precipitation occurred on 11<sup>th</sup> May 2004. In this case, precipitations focused on Eastern Iberia.

On the left of figure 8, we have the information from the climatological INM rain gauge network from the climatological day 12<sup>th</sup> May 2004 covering the precipitation from 11<sup>th</sup> May at 7 UTC to 12<sup>th</sup> May at 7 UTC . The climatological rain gauge network is attended by volunteers which gather the daily information and supply it to the INM on a monthly basis, so it is not a real-time information. The graphic shows the average of all raingauges which are situated within a model gridbox, but the data is situated not in the center of the model gridbox but in the average position of all the raingauges taken into account and that is the reason of the irregularity in the graphic.

The center of the figure represents the accumulated precipitation from 11<sup>th</sup> May at 6 UTC to 12<sup>th</sup> May at 6 UTC with a 3DVAR analysis and including SYNOP 10m wind observations (S25). On the right, it is the same but without using SYNOP 10m wind observations in the analysis (S28).

The time interval in the rain gauge information and in the model did not match exactly, but we assume it because it is not possible to get hourly precipitation information from the climatological rain gauge network daily data and because the differences should be small.

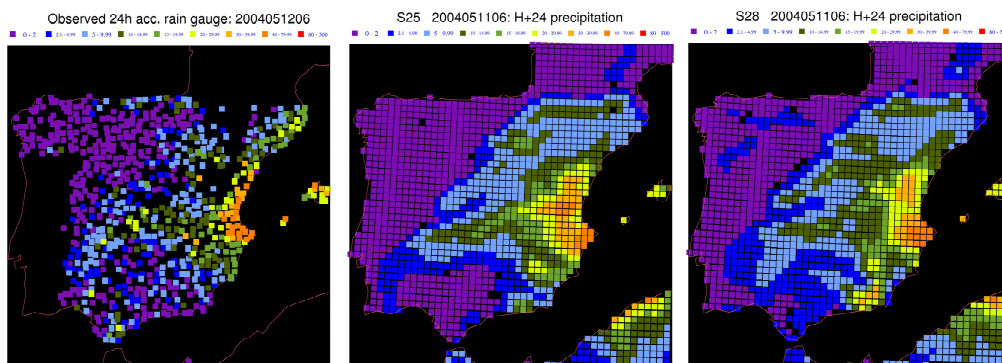


Figure 8. Results of the case study analyzed, with heavy precipitations on Eastern Iberia. On the left, climatological data. In the center, 24 hours forecast accumulated precipitation from S25 experiment (3DVAR, including SYNOP 10 m wind observations). On the right, the same but without SYNOP 10 m wind observations in the analysis (S28).

In both, 4DVAR (not shown) and 3DVAR cases, we see how the precipitation patterns are more realistic in the experiments which include SYNOP 10 m wind observations. See for instance the differences between S25 and S28 (Figure 8) and how S25 removes a lot of the spurious precipitation from the South-East corner of Spain and in the Guadalquivir Valley in the South-West area. Nevertheless, none of the 3DVAR experiments are able to foresee the precipitation in the North-East coast of Spain (Catalonia). However, neither 3DVAR nor 4DVAR experiments are able to forecast that most of the heavy precipitation remains in the Eastern coast of Spain leaving the inland with much less intense precipitations or the precipitations registered on Balear Islands.

## 8.- Bias correction of the surface wind

Some studies show that using wind near surface observations (e.g. SYNOP 10 m wind observations) does not always guarantee an improvement in the quality of the analysis. This is due to the terrain effects on surface winds and the differences between the model description of the terrain and the actual Earth surface. These differences between the real and model terrain induce differences in the analyzed wind field. The solution suggested by some authors [Nishijima et al, 2005] consists in making a statistical bias correction of the surface wind. In every observation point, a set of coefficients are calculated to estimate the surface wind taking into account the real orography and physiography from the model surface wind.

The estimated surface wind ( $u, v$ ) is calculated as a linear combination of the model surface wind components ( $u_m, v_m$ ) and the coefficients are different at each observation point.

$$u = C_1 + C_2 \cdot u_m + C_3 \cdot v_m$$

$$v = C_4 + C_5 \cdot u_m + C_6 \cdot v_m$$

In this way, it is possible to reduce systematic errors, due to terrain effects, of the model wind and the results obtained by Nishijima et al indicates that the analyzed surface wind field is much more realistic. We think that it can be worth exploring this way to improve the wind field analysis and forecast.

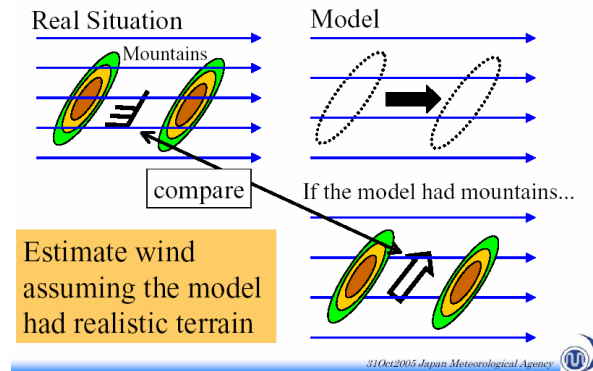


Figure 9. Illustration of the effect of bias correction of the surface wind (taken from the presentation by Nishijima et al (2005) at the 6<sup>th</sup> SRNWP workshop on non-hydrostatic modelling).

## 9.- Conclusions

A simple procedure to introduce SYNOP 10 m wind observations into HIRLAM has been developed. A first selection of quality stations has been carried out and a method to measure objectively its quality envisaged. Experiments using a whitelist which included all the stations selected has shown improvements in wind at all levels and near neutral impact in the rest of the parameters. A more adequate selection of the SYNOP stations is left as a next step in the process of assimilating SYNOP 10 m wind observations and is supposed to improve the scores that we got in our first attempt. The case studies show an improvement on the forecasted precipitation patterns.

## 10.- Acknowledgements

The authors want to thank Makoto Nishijima and Yuki Honda from the Japan Meteorological Agency for their kind permission to reproduce figure 9 from their presentation and Ms Jana Sánchez (INM) for providing the climatological data of precipitation in the case studies.

## References

- Berre L. (2000) Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Monthly Weather Review*, Vol 128, pp 644-667.
- Geleyn J.F. (1988) Interpolation of wind, temperature and humidity values from model levels to the height of measurement. *Tellus 40A* pp 347-351.
- Gustafsson N., Hörnquist S., Lindskog M., Berre L., Navascués B., Thorsteinsson S., Huang X., Mogensen K., Rantakokko J. (1999) Three-dimensional variational data assimilation for a high resolution limited area model (HIRLAM). HIRLAM Technical Report No. 40. January 1999.
- Nishijima M., Honda Y. (2005) Assimilation of surface observations with non-hydrostatic 3DVAR. Presentation in the 6<sup>th</sup> International SRNWP-Workshop on Non-Hydrostatic Modelling. Available at: [http://srnwp.cscs.ch/Lead\\_Centres/2005-Nonhydrostatic/SP-Nishijima.pdf](http://srnwp.cscs.ch/Lead_Centres/2005-Nonhydrostatic/SP-Nishijima.pdf)