Report on

Surface moisture impact study

as deliverable 31 for the TOUGH project

by

Jana Sánchez-Arriola
Beatriz Navascués

Instituto Nacional de Meteorología, Madrid, Spain

1.- INTRODUCTION.

Because the moisture measured from ground based GPS ZTD is an integrated moisture value instead of a profile, it is useful to include the two meter relative humidity (2mRH) observations from the SYNOP stations in the HIRLAM assimilation system with the aim of helping the analysis to distribute in the vertical the error in the integrated water vapour content given by the information contained in GPS ZTD observations.

For this purpose, it is necessary to have an appropriate observation operator for 2mRH that translates the model variables into the measured quantities.

The 2mRH observation operator that currently exists in the HIRLAM assimilation system was developed following Geleyn (1988). But this operator has not been updated according to the development of new land surface introduced in the HIRLAM model.

Along 2004, INM has developed a new observation operator that has allowed assimilating two meter relative humidity observations in the HIRLAM variational data assimilation system, along with the corresponding tangent linear and its adjoint.

The impact of the assimilation of near surface humidity observations added to the rest of conventional observations, has been studied through a series of parallel experiments carried out over one month long period in spring 2004, when some heavy rain events took place in Spain.

2.- NEW 2mRH OBSERVATION OPERATOR.

An observation operator for near surface parameters was developed in the past and it is available in the HIRLAM three dimensional variational assimilation (3DVar) system (Gustafsson et al.1999).

This operator produces model 2mRH to be compared with the observations and it is based on Monin-Obukhov similarity theory for the surface layer, following the method proposed by Geleyn (1988) to interpolate dry static energy and specific humidity. It makes use of the Richardson number (Ri) as an stability parameter and different analytical expressions for the vertical interpolation are used for the stable

and unstable cases in the surface layer. Then, the procedure presents a discontinuity between the lowest model level and the surface at the neutral case, i.e. Ri=0.

By the time when this observation operator was developed, an older parametrization for the land surface processes was installed in the HIRLAM forecast model. But currently, this parametrization has been replaced by a mosaic scheme in which the land tiles follow the ISBA scheme. This fact has produced that the computation of model two meter relative humidity in the observation operator within the HIRLAM 3DVar assimilation system is not already valid, as it is mentioned in a previous TOUGH report by Ridal and Navascués (2004).

In order to allow the assimilation of surface relative humidity observations, new FORTRAN routines for a simple observation operator for two meter relative humidity (2mRH), the tangent linear and its adjoint have been coded, tested, and then introduced in the HIRLAM 3DVar assimilation system at INM.

This observation operator assumes that relative humidity is constant in the unstable surface layer, so 2mRH can be directly compared to relative humidity at the lowest model level, RH_{NLEV}. The smaller coupling between the near surface and the uppermost atmosphere in stable case is taken into account by increasing substantially the observation error. So, it produces the 2mRH observations to have no weight in the upper air analysis in this case.

Relative humidity depends on specific humidity q, pressure P, and temperature T, through vapour pressure e(q,P) and saturation vapour pressure Es(T):

$$e(q,P) = P \frac{q}{\varepsilon + q(1-\varepsilon)} \qquad \varepsilon = \frac{R_d}{R_v}$$

The Bolton's formula is used for saturation vapour pressure equation:

$$Es(T) = 100\left(d \exp\left(a\left(\frac{T-b}{T-b+c} \right) \right) \right)$$

With constants d=6,112 , a=17.67, b=273,16 and c=243,5 .

In our formulation, the model 2mRH is approached in the unstable surface layer by the relative humidity at the lowest model level RH_{NLEV} at each observation site:

$$2mRH \approx RH_{NLEV} = \frac{e(q_{NLEV}, P_{NLEV})}{Es(T_{NLEV})} = \frac{e_{NLEV}}{Es_{NLEV}}$$

In the HIRLAM 3DVar the horizontal resolution of the assimilation increment is smaller than the full model state due to the adopted incremental formulation. To calculate the interpolated values corresponding to analysis increments, a tangent-linear of observation operator is created, that is the linearization of the non-linear observation operator around the background field:

$$e_{NLEV}^{tl} = \frac{e_{NLEV}}{P_{NLEV}} P_{NLEV}^{tl} + \left(\frac{e_{NLEV}}{q_{NLEV}} - \left(\frac{(1-\varepsilon).e_{NLEV}}{\varepsilon + (1-\varepsilon)q_{NLEV}} \right) \right) q_{NLEV}^{tl}$$

$$Es_{NLEV}^{tl} = a.Es_{NLEV}\left[\frac{1}{T_{NLEV}-b+c} - \frac{T_{NLEV}-b}{(T_{NLEV}-b+c)^2} \right].T_{NLEV}^{tl}$$

$$2mRH^{tl} = RH^{tl}_{NLEV} = \frac{1}{Es_{NLEV}} e^{tl}_{NLEV} - \frac{e_{NLEV}}{Es_{NLEV}^2} E^{tl}_{NLEV}$$

Figure 1 shows the validity of the assumption of the new observation operator. It represents the difference between relative humidity at lowest model level and at two meters as a function of the Richardson number.

In unstable surface layer (Ri<0) the error of our simple observation operator is well below 10%, within the assumed observation error value. In the stable case (Ri>0), a direct comparison of model relative humidity at the lowest model level against the observations will be much less accurate, producing errors of observation operator up to 60%. In these cases, the lowest model level is much more uncoupled with the surface and then, humidity information contained in observations at 2m will not be useful to correct the atmosphere aloft.
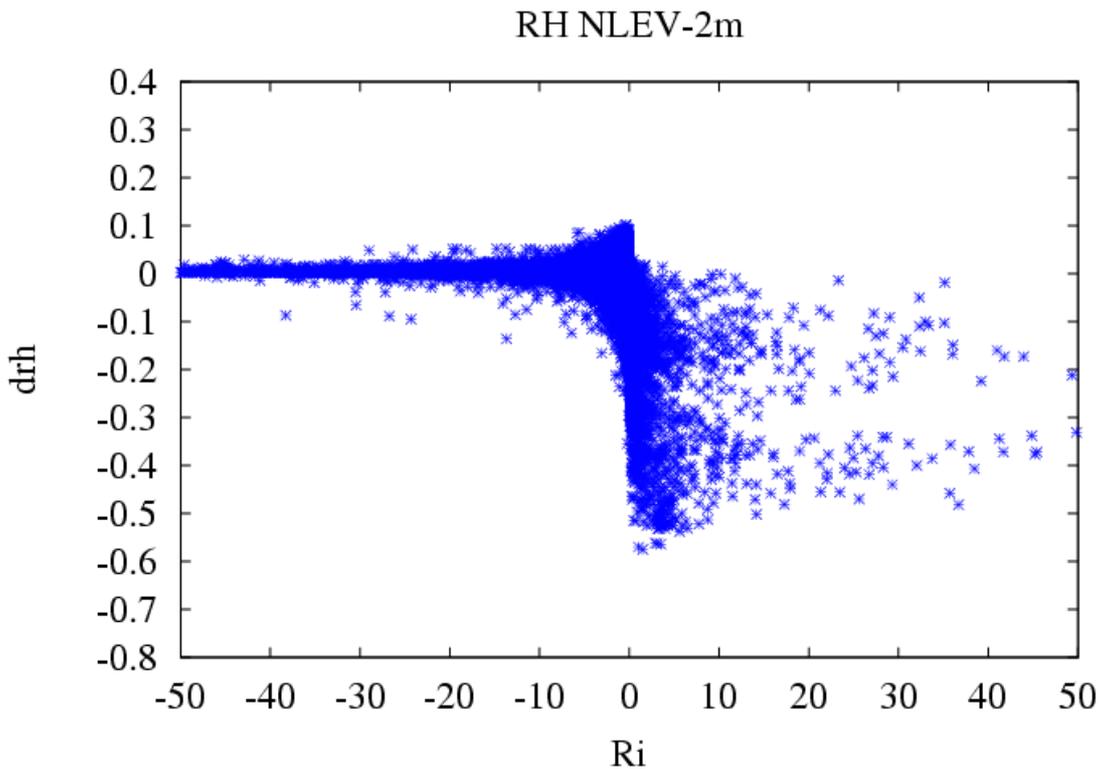
FIG 1.- Difference between relative humidity at the lowest model level and at two meter as a function of the Richardson number.

The new observation operator, its adjoint, and its tangent linear have been coded and introduced in the HIRLAM 3DVar assimilation system.

Observation error variance for 2mRH assimilation have been assigned similarly to radiosonde observations. An empirical regression of errors of relative humidity

dependent on temperature is used (Lindskog et al., 2001). It produces values of 2mRH error standard deviation around 10% over Europe in spring time.

The 2mRH background error standard deviation can be estimated using a Monte Carlo technique installed as option within (Schyberg et al 2003) the HIRLAM 3D-Var assimilation code that follows the method suggested by Fisher and Courtier (1995). A set of random perturbations with Gaussian distribution N(0,1) for the elements of control vector ξ are generated followed by a transformation to model space δx:

$$\delta x = U^{-1}\xi$$

For each model space perturbation δx, the tangent linear of the new observation operator for 2mRH is applied to obtain the model perturbation in terms of observed variable:

$$\delta y = H^{'}\delta x$$

This procedure produces an average background error standard deviation of 2mRH around 7% over land.

The observation error to background error ratio ratio for 2mRH has been also separately estimated from the HIRLAM surface analysis innovations. The obtained *sigmao* to *sigmab* ratio value for the error statistics calculated from the HIRLAM surface analysis at daytime is smaller than those used in the HIRLAM upper air variational assimilation system.

This new observation operator for 2mRH, the tangent linear and its adjoint have been tested in a series of parallel experiments over one month long period in spring 2004, as it is described in the next section.

3.- TESTS WITH THE NEW OBSERVATION OPERATOR FOR 2mRH

Some tests to the tangent linear and adjoint routines, and single observation experiments, have been performed to preliminary assess the new code developed. After that, some experiments have been carried out with the HIRLAM system (version 6.3.5), at 22 km horizontal resolution and 40 levels in the vertical using the computer facilities installed at ECMWF.

The time period chosen to carry out the experiments was from 15th April to 15th May 2004, a wet period in most of Spain with heavy rain events in Levante and Andalucia at the beginning of May.

Two experiments have been compared :
- *RE5*: Control experiment.
- *RH5*: Assimilation of 2mRH SYNOP observations.

Apart of 2mRH observations for *RH5*, both experiments have assimilated the same conventional observation types: surface geopotential observations from land stations, ships and buoys, aircraft temperature and wind measured data, and vertical profiles of wind (PILOT balloons and radiosondes), temperature and humidity (radiosondes). No satellite data from geostationary or polar platform instruments have been assimilated.

The control experiment, *RE5,* has been run with all available 2mRH observations introduced passively through the HIRLAM 3D-Var system. It makes

possible that simulated by the model 2mRH values can be compared to observations during the experiment run.

    RH5 experiment assimilates 2mRH observations from land surface stations over the model domain. The screening procedure for them includes a redundancy check that only retains the observation closest to the analysis time from all the available from the same station within the assimilation time window.

    The so called statistical balance formulation (Berre, 2000) has been used as background constraint in the HIRLAM variational assimilation system in both experiments. It implies a multivariate analysis of mass, wind and humidity.

Assimilation diagnostics

    Some histograms of 2mRH innovations and analysis residuals have been obtained (*Figure 2*). In case of 2mRH, the innovation distribution seems to be nearly Gaussian, although some small biases are observed in the histograms. The distance of observations to the analysis is not very much reduced according to the residuals histograms. A number of possibly bad observations might have been used in the analysis, due to a very large first guess check limit for 2mRH, and very small a priori probability of gross error in the variational quality control within the HIRLAM variational assimilation system. It has to be also taken into account that, as it has been mentioned, the *sigmao* to *sigmab* ratio for 2mRH is very large in the analysis.

    The root mean square (rms) of the humidity analysis increments for each experiment and the rms of the differences between the humidity analysis for the two experiments have been calculated at 12 UTC. Humidity analysis increments at 12UTC for experiments *RE5* (control) and *RH5* (control + 2mRH) at model level 30 (around 850 hPa) are shown in *Figure 3*, and at the lowest model level (40) in *Figure 4* (0.5g/kg contouring interval). Comparing control with *RH5,* larger humidity analysis increments mostly over central and northern Europe at the lowest model level are observed when 2mRH observations are assimilated. This effect is smoother at higher levels as e.g. around 850 hPa.

    Maps of rms analysis differences of specific humidity between the control experiment and *RH5* at model levels 30 and 40 are shown in *figures 5* and *6.* (1g/ kg contouring interval). As it can be observed, the differences appear not only over Europe but also over Africa and the Atlantic and the largest amplitudes are over the Ocean in both experiments around 850hPa. The effect of assimilating 2mRH observations can be again appreciated over central Europe at the lowest level.

    The fit of first guess and analysis to humidity radiosonde measurements has also been obtained. In the lowest atmosphere, the distance to the first guess from radiosonde humidity observations seems to be smaller in experiment *RH5* than in control. (*Figure* 7).

Objective verification of HIRLAM forecasts

    The bias and root mean square (rms) of 2mRH HIRLAM forecasts for the control and *RH5* experiments have been obtained using different observations datasets representative of several European geographical areas and against 2mRH observations from the EWGLAM surface stations list.

As it can bee seen in *Figure 8,* the assimilation of 2mRH data is improving the HIRLAM forecasts, mainly through the reduction of the bias observed in the control experiment. Impact of assimilation of surface humidity observations becomes smaller with increasing the forecast length (not shown).

Objective verification of HIRLAM forecasts against radiosonde relative humidity observations has shown a small or neutral impact of the assimilation of 2mRH observations, when all radiosondes in the model area are considered. The main impact is found at and below 850hPa and becomes smaller when increasing the forecast length. In some specific areas the impact of assimilation of 2mRH observations (**RH5** experiment) is positive, as it is the case of Scandinavia where the decrease in terms of rms near the surface is of the order of 5%. Over France the impact is found to extend more in the vertical, with a slight improvement of **RH5** experiment over control.

Verification of the model precipitation for the different experiments has revealed a clear sensitivity to the assimilation of near surface humidity. Validation of the model precipitation has been based on upscaled rain gauge from the high resolution INM Climate stations network in order to approach to model resolved scales. Superobservations of precipitation at each grid box have been created by averaging rain gauge at the available individual stations. In this way, around 1000 superobservations have been calculated every day over Spain that represent the upscaled rain gauge from the 3500 stations of the INM Climate Network to the 22km horizontal resolution of the HIRLAM experiments.

Contingency tables of observed/forecast precipitation for different forecast lengths corresponding to different precipitation thresholds (0.1, 0.3, 1., 3., 10., 30., 100. mm in 24 hours) have been obtained, as it is shown in Table 1. Some standard scores have been calculated like True Skill Score (TSS), Equitable Threat Score (ETS), Hit Rate, Probability of Detection, False Alarm Rate and Frequency Bias.

	Observed YES	Observed NO	Total
Forecasted YES	a (hit)	b (false alarm)	a+b
Forecasted NO	c (miss)	d (correct rejection)	c+d
Total	a+c	b+d	a+b+c+d=n

Table1: Contingency table for observed and forecasted precipitation categories (Wilks,1995).

The True Skill Score index TSS can be written as (TSS) =(ad-bc)/(a+c)(b+d). The TSS can also be written as the probability of detection (H=a/(a+c)) minus the probability of false detection (F=b/(b+d)): TSS=H-F. It ranges from 1 (perfect forecast) to -1.

The Equitable Threat Score ETS, is the TS rendered equitable by taking away the random forecast R(a)= (a+b)\cdot(a+c)/(a+b+c+d). ETS=(a-R(a))/(a-R(a)+b+c). It ranges from 1 (perfect forecast) to 0 (chance and constant forecast).

While Frequency Bias Index is (FBI) =(a+b)/(c+d) and it measures the event frequency and has value one for a perfect forecast, and larger (smaller) than one if the system is over (under)forecasting (Wilks,1995).

Standard scores obtained for H+24 precipitation forecasts are shown in *figure 9.*

In this figure it can be observed that for accumulated precipitation higher than 3 mm in 24 hours, **RH5** experiment shows higher TSS and ETS scores than control, indicating a better skill of the HIRLAM model to predict moderate to heavy precipitation events. This seems to be produced by a larger probability of detection for these events, although it is also observed that the model precipitation is still more overpredicted in **RH5** than in control experiment. The improvement in the HIRLAM precipitation due to the assimilation of surface humidity observations decreases when longer forecast lengths are considered (not shown).

Maps of model 24 hours accumulated precipitation in each experiment have been plotted to compare with INM Climate Stations Network rain gauge for two different case studies: 2nd may 2004 that was very rainy in southern the Iberian Peninsula (Andalucia), with precipitations that reached more than 60 mm in Cadiz, and 11th may 2004 that was rainy in eastern Spain, with more than 80 mm rainfall in a few hours in Levante.

Although differences in the forecasts are no very noticeables, it seems that in the first case (*figure 10*) **RH5** experiment could predict slightly better the position of the maximum of precipitation that occurred in Andalucia than the control experiment.

For the second case study (*figure 11*) , it seems that **RH5** was the experiment that best predicted the two events that took place: the storm over Levante (East), where it could cached better the maximum of precipitation, and the structure observed in Extremadura (south-western Spain).

4. - SUMMARY.

Assimilation of 2mRH observations from SYNOP stations in HIRLAM is necessary to complete the atmospheric moisture information, currently given by radiosondes and GPS ZTD observations in the future.

To make a proper assimilation of these type of observations, it has been needed to replace the old 2mRH observation operator in the HIRLAM variational assimilation system because it was not updated when a more recent parametrization for the land surface processes (ISBA) was installed in the HIRLAM model physics.

A simple observation operator for 2mRH observations for the HIRLAM 3DVar system has been developed and first tried to investigate the impact of near surface humidity observations themselves, and it will be also tested accompanied with GPS ZTD data.

This observation operator assumes that relative humidity is constant in the unstable surface layer, so 2mRH can be directly compared, in this cases, to RH at the lowest model level.

The results obtained with this observation operator have been promising in respect to the usage of 2mRH with the rest of conventional observations. Validation has been done verifying the model forecasts against radiosondes observations and the high resolution INM climate stations rain gauge. The assimilation of 2mRH observations seems to be beneficial.

Anyway, some tuning of the relative size of observation and background error standard deviations of 2mRH and of the quality control parameters seems to be needed to improve the assimilation of this moisture information.
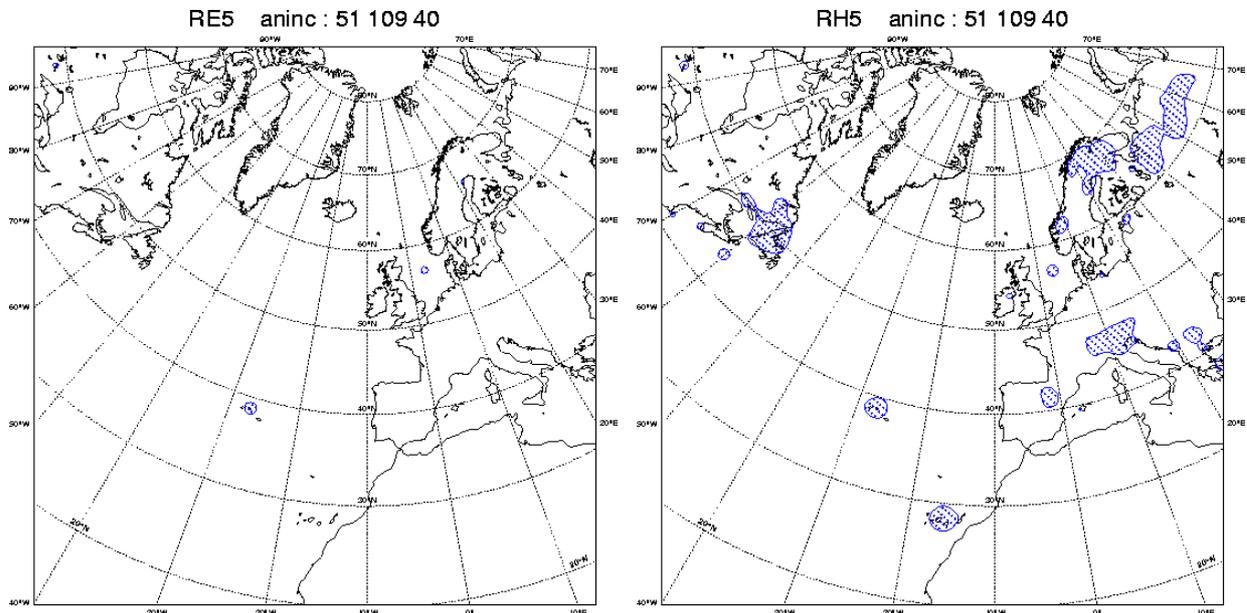
5. - REFERENCES.

1. Berre L. 2000. Estimation of synoptic and meso scale forecast error covariances in a limited area model. *Mon.Wea.Rev*. **128**, 644-667.

2. Fisher M., Courtier P. 1995. Estimating the covariance matrix of analisis and forecast errors in vatiational data assimilation. ECMWF Res.Dep. *Tech. Memorandum* No.**220**.

3. Geleyn J.F. 1988. Interpolation of wind, temperature and humidity values from the model levels to the heigh of measurement. *Tellus* **40**, 347-351.

4. Gustafsson N. 1999. The numerical scheme and lateral boundary conditions for the spectral HIRLAM and its adjoint. *ECMWF Seminar Proceedings*. Recent developments in numerical methods for atmospheric modelling. ECMWF, Reading, U.K. 7-11 September 1998, 335-363.

5. Lindskog M., Gustafsson N., Navascues B., Mogensen K.S., Huang X.Y., Yang X., Adrae U., Berre L., Thorsteinsson S., Ratakokko J. 2001. *Tellus* **53A**, 425-468.

6. Ridal M. and Navascués B. 2004. Surface moisture observation operator. *TOUGH deliverable Report* **D30**.

7. Schyberg,H., Landelius T., Thorsteinsson S., Tveter F.T., Vignes O., Amstrup B., Gustafsson N., Järvinen H., Lindskog M. 2003. Assimlation of ATOVS data in the HIRLAM 3D-VAR System. *HIRLAM Tech. Report*. **60**.

8. Wilks D.S. 1995. Statistical Methods in the Atmospheric Sciences. *Academic Press*.
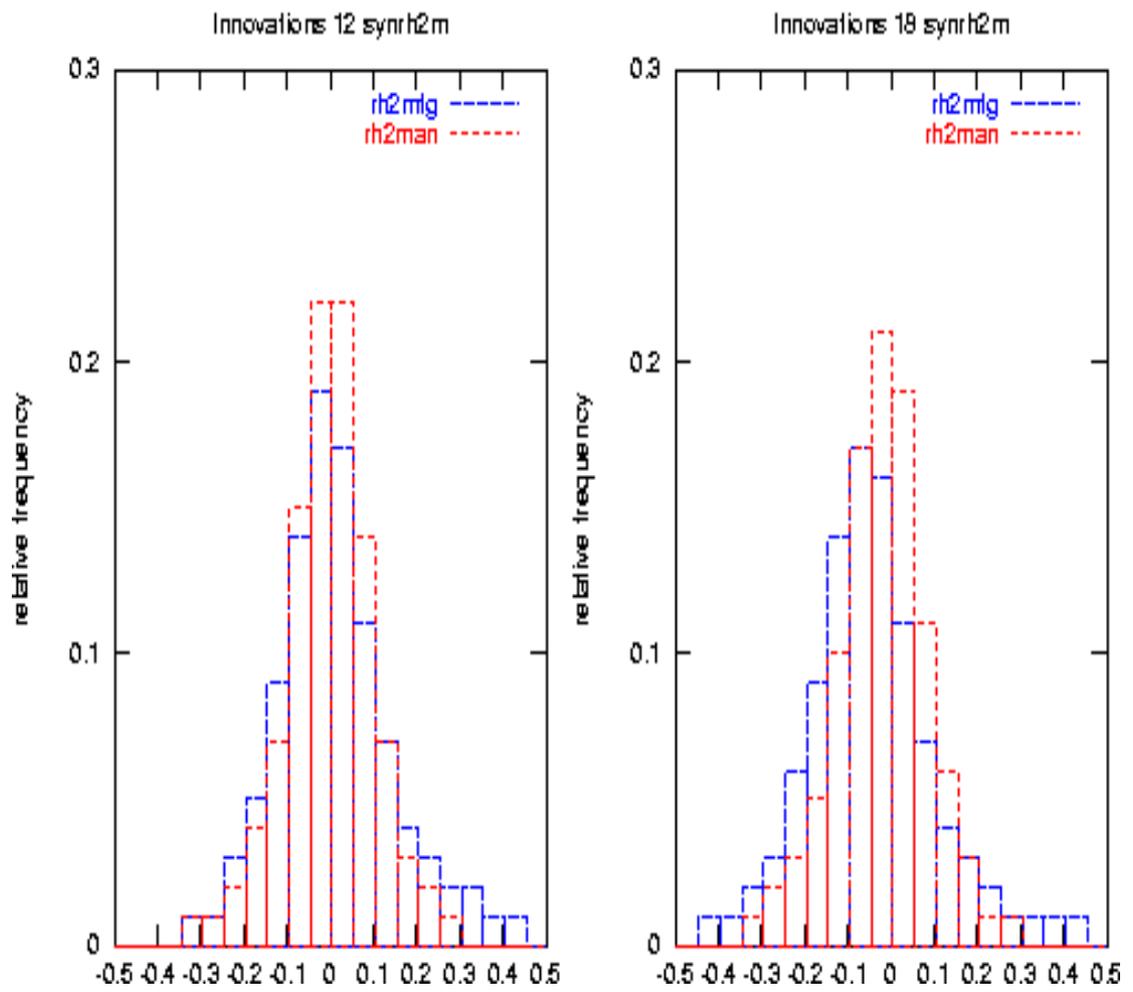
FIG.2. – Histograms of 2mRH from RH5 experiment at 12 and 18 UTC. Innovations are in blue and residuals in red.
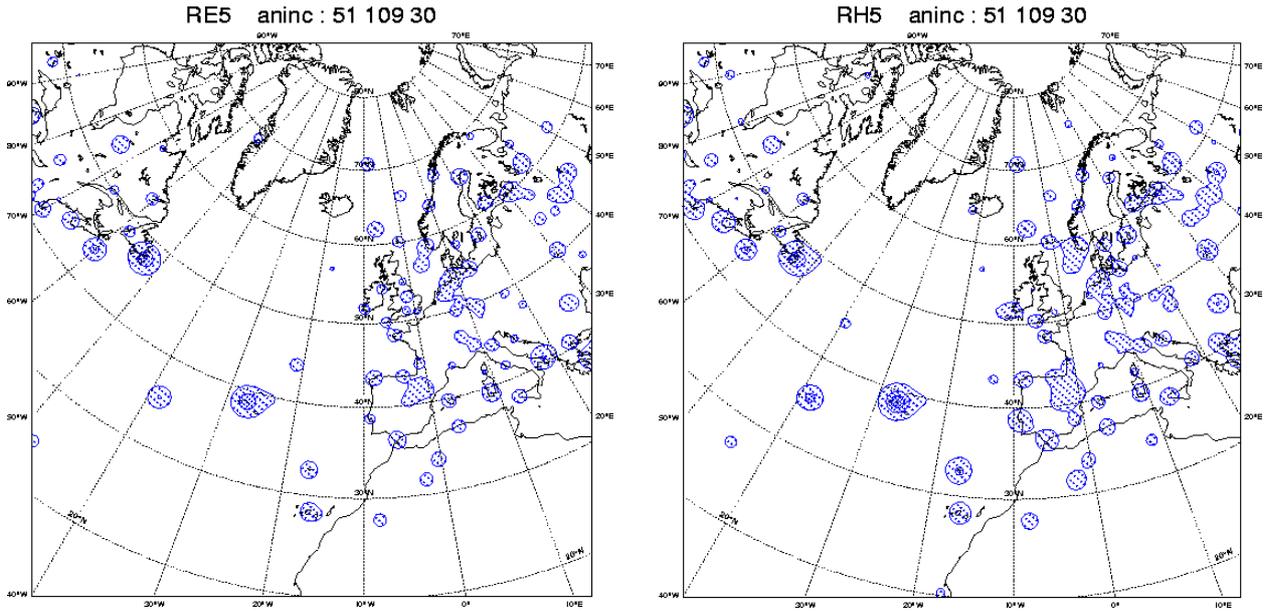
FIG.3 - Root mean square of humidity analysis increments at 12 UTC for experiments RE5 and RH5 at model level 30 (850hPa). (0,5g/kg contouring interval).

FIG. 4. - Root mean square of humidity analysis increments at 12 UTC for experiments RE5 and RH5, at the lowest model level 40. (0,5g/kg contouring interval).
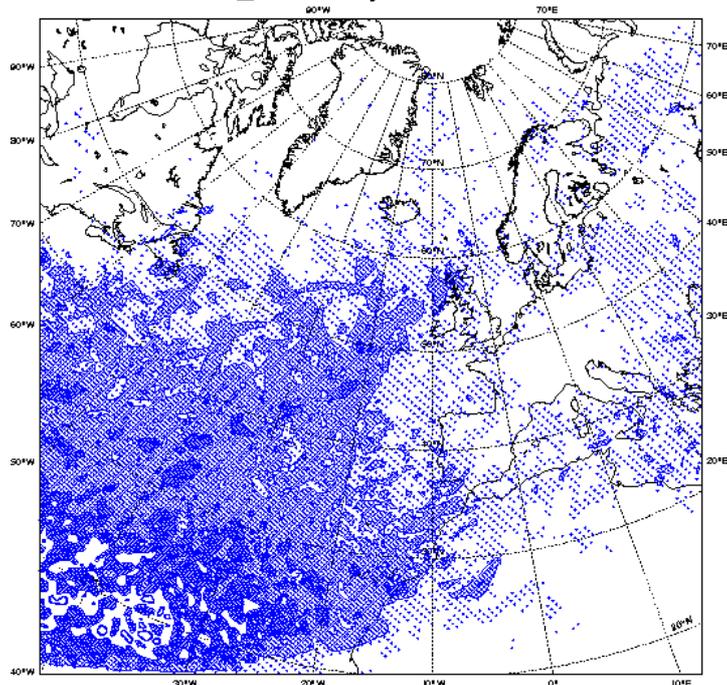
RE5_RH5 analysis diff.: 51 109 30

FIG. 5 - Root mean square of analysis differences of specific humidity between the control experiment RE5 and RH5 model level 30 (850hPa). (1g/kg contouring interval)

RE5_RH5 analysis diff.: 51 109 40
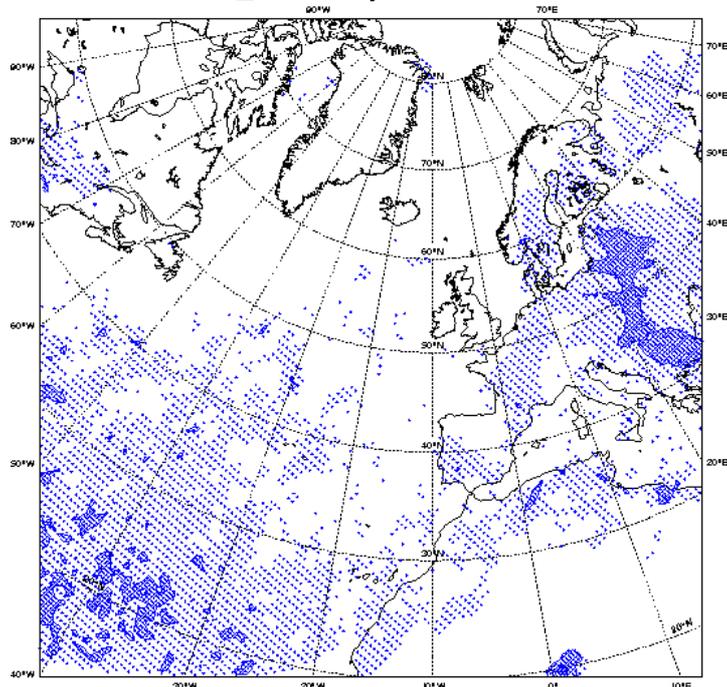
FIG. 6. - Root mean square of analysis differences of specific humidity between the control experiment RE5 and RH5 at model level 40. (1g/kg contouring interval).
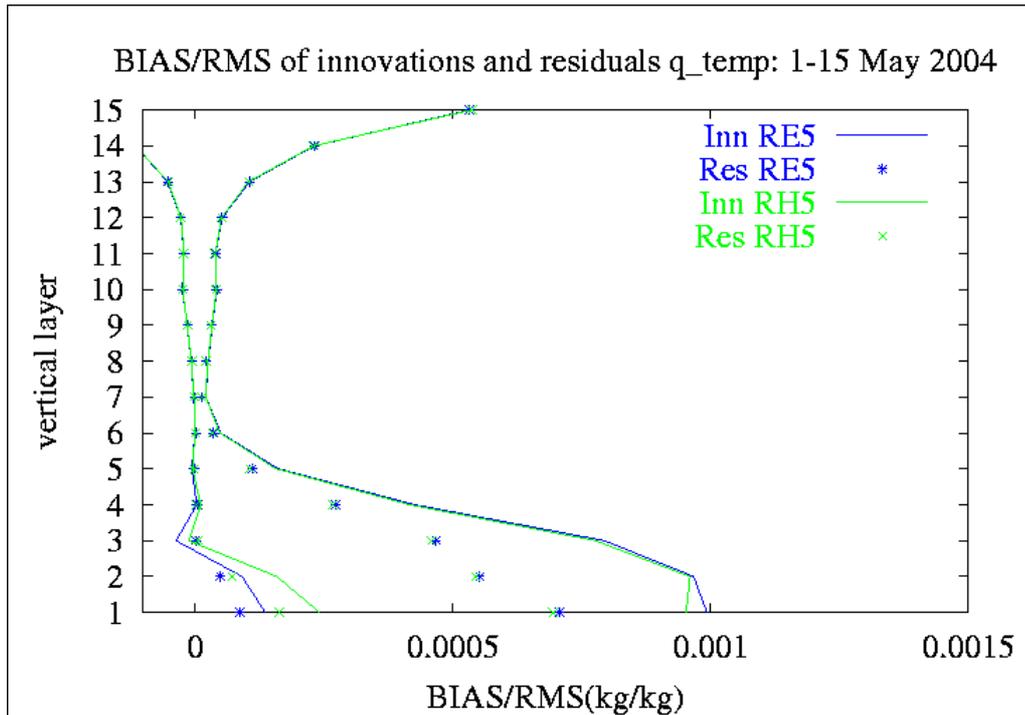
BIAS/RMS of innovations and residuals q_temp: 1-15 May 2004

FIG. 7. - Fit of first guess and analysis to humidity radiosonde measurements.Control experiment in blue and RH5 in green. Lines are innovations and points are residuals.
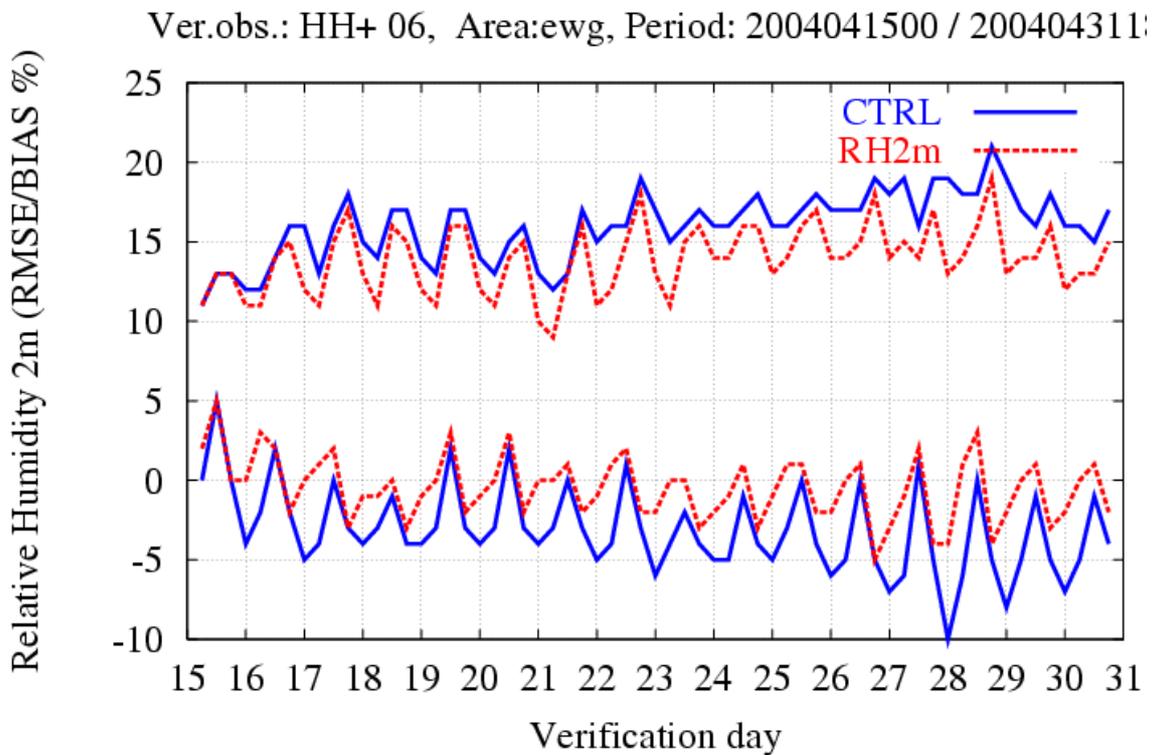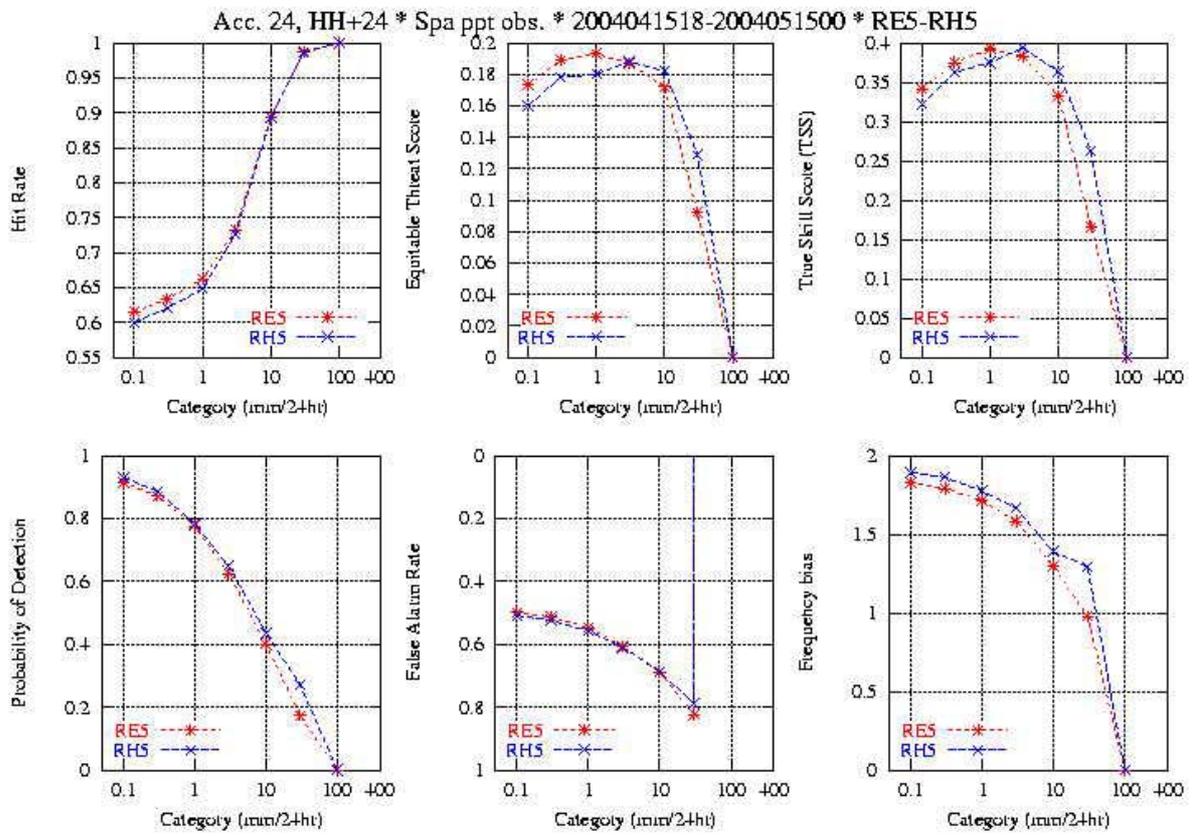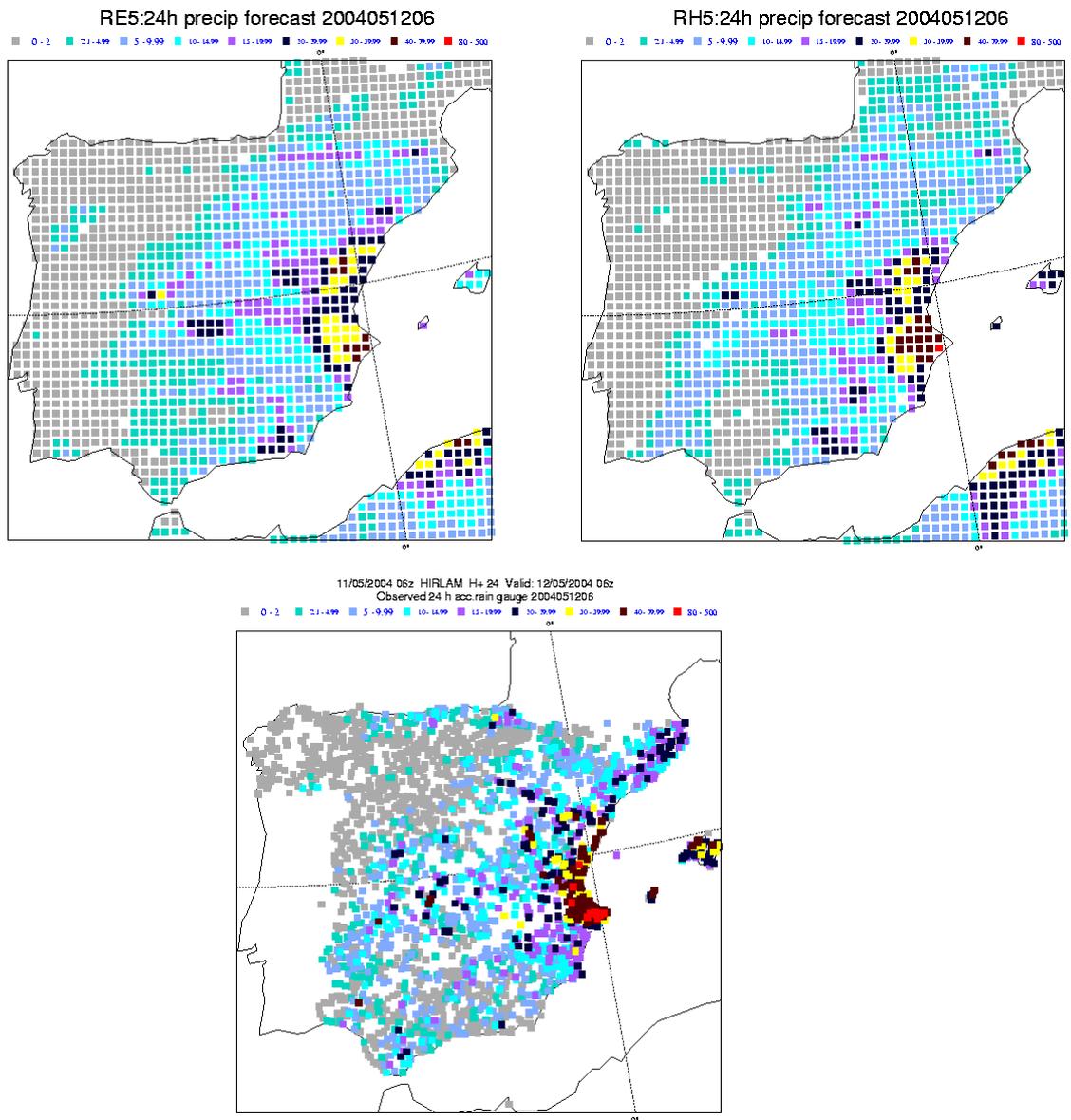

Ver.obs.: HH+ 06, Area:ewg, Period: 2004041500 / 200404311:

FIG. 8. – 2mRH RMS/BIAS (%) for the study period . CTRL is *RE5* experiment (blue) and RH2m is *RH5* experiment (red).

FIG. 9. - Contingency tables of observed/forecast precipitation for 24 forecast length corresponding to different precipitation thresholds (0.1, 0.3, 1., 3., 10., 30., 100. mm in 24 hours) for experiments RH5 (blue), and control RE5 (red).

13
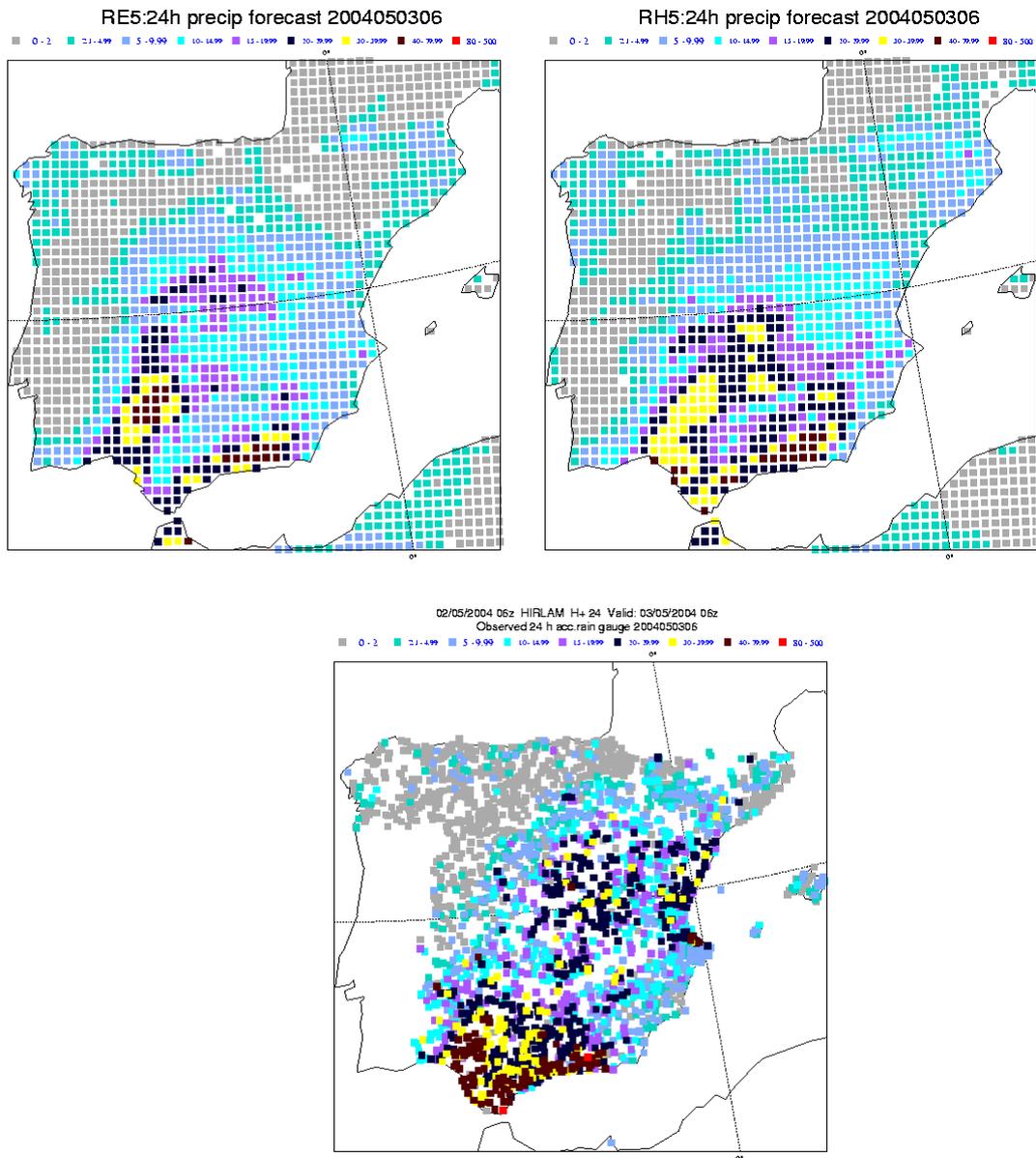
FIG. 10.- 24 hours forecast for RE5 and RH5 experiments and 24 hours accumulated precipitation at INM Climate stations Network for the first case study.

RE5:24h precip forecast 2004051206

RH5:24h precip forecast 2004051206

11/05/2004 06z HIRLAM H+ 24 Valid: 12/05/2004 06z
Observed 24 h acc.rain gauge 2004051206

FIG. 11.- 24 hours forecast for RE5, and RH5 experiments and 24 hours accumulated precipitation at INM Climate stations Network for the second case study.