Data Assimilation by Field Alignment. Testing the Theory.

Carlos Geijo

2011 HIRLAM ASM / 21th ALADIN Workshop. Norrköping, April 5-8

INTRODUCTION (I)

- DAbyFA , a method proposed by a group of MIT scientists and published in November 2006

- Classical formulations of DA, whether sequential, ensemble-based or variational,are “amplitude adjustment methods”. Such methods can perform poorly when forecast locations of weather systems are displaced from their observations

- Characterization of position errors is complex, yet very important for forecasting weather of strong and localized phenomena (tropical cyclones, thunderstorms, squall lines, etc...). Position errors introduce bias between observations and forecasts

- The issue is not new. For years, “ad-hoc” techniques (“bogussing”) have been used operationally in Tropical Cyclone Forecasting

INTRODUCTION (II)

- In the last 20 years different objective methods to tackle this problem have been proposed and tested.

a) Mariano A.J (1990) : contour analysis and melding fields

b) Hoffman R.N et al (1995,1996): a variational technique accounting for “distortion errors”, proved on ECMWF analyses using microwave satellite data. More recently, Nehrkorn T. et al (2003) on calibration of this method

c) Alexander G.D et. al (1998) : image warping using microwave satellite data to improve forecasts of ***mesoscale marine cyclones***

d) Brewster K.A (2003): a different field alignment algorithm to the one that this presentation refers to, tested on ***storm-scale NWP*** with simulated data

INTRODUCTION (III)

- Ravela S. et al, at MIT (2006), start off from the Bayesian formulation of the DA problem, which under a number of hypotheses gives for the inference of the model state this expression (ignoring the normalizing constant $P(Y_n)$)

$$P(X_n \mid Y_{0:n}) \quad \alpha \quad P(Y_n \mid X_n) \quad P(X_n^f)$$

"a posteriori" or "inferred" model state X at t=n, conditioned on observations from 0 to n

Data likelihood fot t=n. It implies the existence of model equivalents to the observations

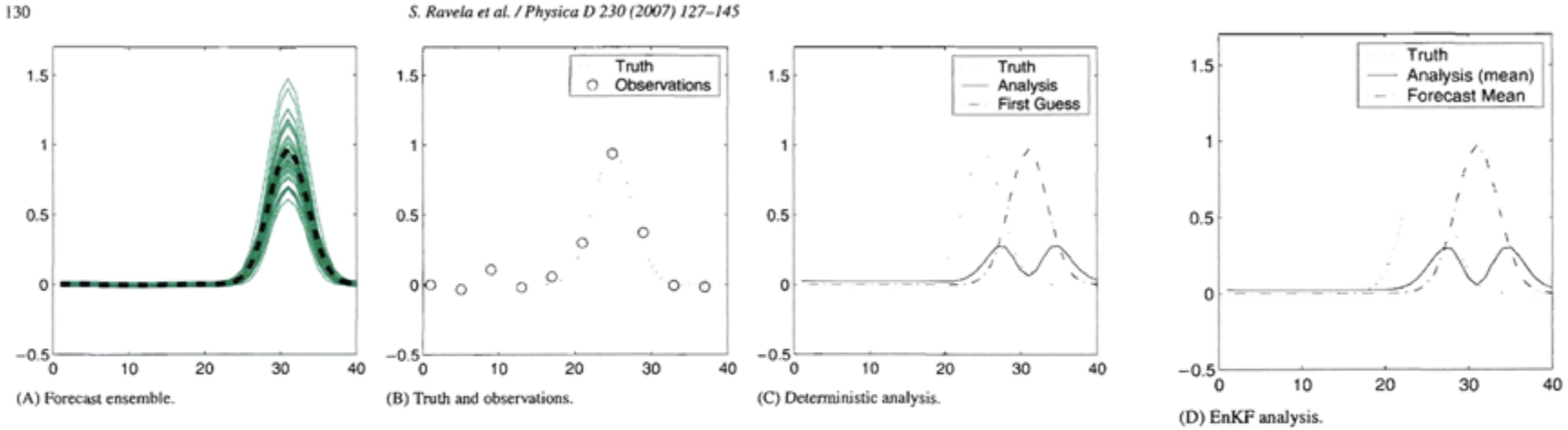
Forecast Model Error or "Forecast Prior"

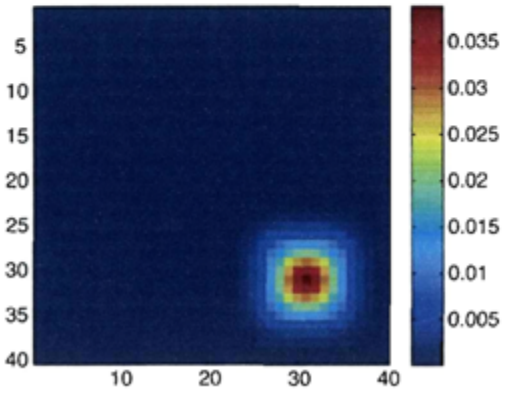
- Under further assumptions (Gaussianity,obs op linear) this formulation leads immediately to the 3Dvar (in a deterministic context) or the EnKF (in a probabilistic context) quadratic equations for the analysis objective J

INTRODUCTION (IV)

- Both schemes, 3DVar ***and EnKF***, can perform bad in the presence of position errors

3Dvar EnKF



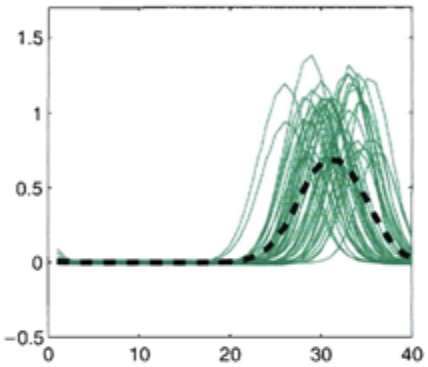


1-D example built with a 40 members ensemble, perturbed only in amplitude. B-matrix shown down left. "Truth" displaced left about 3*δ, where δ is the width of the "front". 3DVar analysis and EKF mean analysis appear both distorted. σ_O is substantially less than σ_b (about 1/5). The observation density is 1/10.

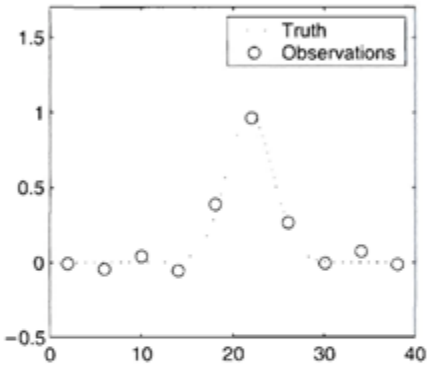
INTRODUCTION (V)

- Both schemes, 3DVar ***and EnKF***, can perform bad in the presence of position errors

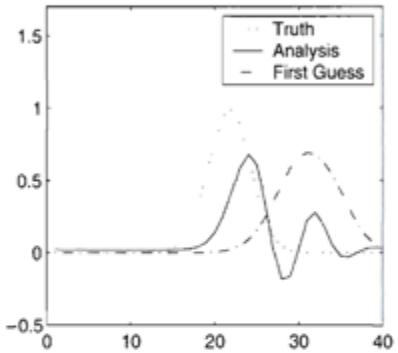
3Dvar EnKF



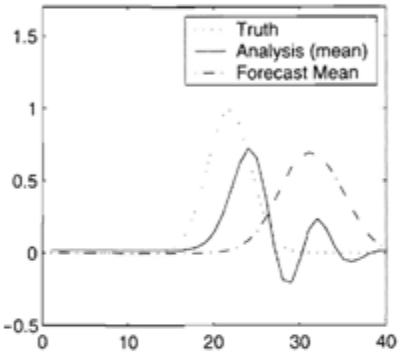
(A) Forecast ensemble.



(B) Truth and observations.



(C) Deterministic analysis – 1.



(E) EnKF analysis.

The same 1D-example, but with perturbations in position as well. The "truth" is displaced to the left about 3*δ , where δ is the perturbation in position. B-matrix is computed from the 40 members ensemble. The distortion in the 3DVar and EKF mean analyses is still important.

DAbyFA: the method (I)

- The method explicitly represents position errors by introducing in the analysis control space a displacement vector field **q**, defined in each analysis grid point, that gives the deformation necessary to minimize these position errors

- In the Bayesian framework sketched previously, the inference for the model state now becomes (ommiting some indexes)

$$P(X, \mathbf{q} \mid Y) \quad \alpha \quad P(Y \mid X, \mathbf{q}) \quad P(X^f \mid \mathbf{q}) \quad P(\mathbf{q})$$

"Data likelihood". Connects observations to the *displaced* model state

The "amplitude prior". Says that the forecast statistics are conditioned on the displacement field **q** (e.g. B(**q**))

"displacement prior", enables the introduction of smoothness constraints on the **q** field

DAbyFA: the method (II)

In the usual assumption of gaussian statistics for these component PDFs, we have:

a) Data Like. $P(Y \mid X, \mathbf{q}) \; \alpha \; exp\ -1/2 \; (Y - H\, X\,(\mathbf{p}))^T \, R^{-1} \, (Y - H\, X\,(\mathbf{p}))$

where X ($\mathbf{p} = \mathbf{r} - \mathbf{q}$) represents X displaced by **q**

b) **Amp. prior** $P(X^f \mid \mathbf{q}) \; \alpha \; |B(\mathbf{q})|^{-1/2} \; exp\ -1/2 \; (X(\mathbf{p}) - X^f\,(\mathbf{p}))^T \, B(\mathbf{q})^{-1} \, (X(\mathbf{p}) - X^f\,(\mathbf{p}))$

the forecast error is assumed Gaussian in the position corrected space. Note the dependence of B on **q**

c) Displa. prior $P(\mathbf{q}) \; \alpha \; exp - L\,(\mathbf{q})$; where:

$$L(q) = w_1 / 2 \sum_{j \in \Omega} tr\left[\left[\vec{\nabla} \mathbf{r}_{q_j} \right] \left[\vec{\nabla} \mathbf{r}_{q_j} \right]^T \right] + w_2 / 2 \sum_{j \in \Omega} \left[div \mathbf{r}_{q_j} \right]^2$$

This term expresses the smoothness or “regularization” constraints imposed on the solution for **q**. The parameters w_1 and w_2 are free and weight both terms of the constraint. This formulation (Tikhonov type formulation) is inspired in the theory of viscid fluids.

DAbyFA: the method (III)

With these definitions of probabilities, the *Field Alignment Cost Function* becomes:

$$2J_{FA} = \left(X(\vec{p}) - X^f(\vec{p})\right)^T B(\vec{q})^{-1} \left(X(\vec{p}) - X^f(\vec{p})\right) + \left(Y - HX(\vec{p})\right)^T R^{-1} \left(Y - HX(\vec{p})\right) + 2L(\vec{q}) - \ln\left(\left|B(\vec{q})\right|\right)$$

The solution of this problem is complicated. It is not clear how to compute B(**q**) and the gradients of J_{FA} are not easy to compute either. Ravela et al. present two ways of overcoming these difficulties by making several approximations.

a) The **"one-step algorithm".** An iterative procedure that works with ensembles. The denomination refers to the fact that in this case the minimun is searched simultaneously in amplitude and position. This algorithm can be very expensive in computational terms because it does not scale well with ensemble size.

b) The "**sequential solution**". It can be utilized in probabilistic and deterministic approaches alike.

DAbyFA: the method (IV)

In the "**sequential solution**" or "**two-step algorithm**", the idea is to solve the two equations:

$$\frac{\partial J}{\partial X} = 0 \quad (1) \quad ; \frac{\partial J}{\partial \mathbf{q}} = 0 \quad (2)$$

sequentially. **In the first step**, X is fixed to X^f in (2) and then a solution for **q** is found. This smooth and continous deformation is used to correct the position errors in X^f . **In the second step,** X^f (**q**) (the aligned forecast) is used to obtain an analysis from (1).

The first step amounts to solving the "**alignment equation**"

$$w_1 \Delta \vec{q} + w_2 \nabla\left(\nabla \cdot \vec{q}\right) +$$

$$(\nabla X^f{}_{|\vec{p}})^T H^T R^{-1}\left(H X^f(\vec{p}) - Y\right) = 0$$

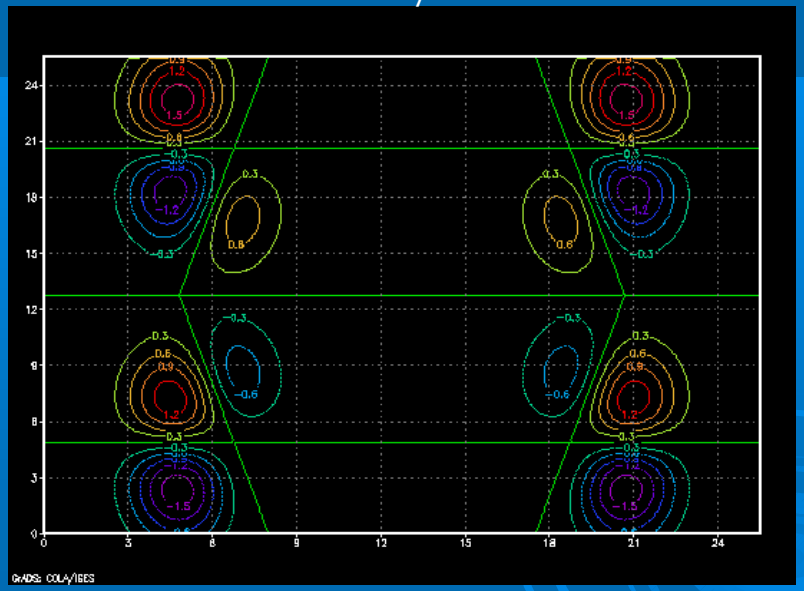
which, due to the dependence of the forcing on **q**, is non-linear and has to be solved iteratively. The forcing term is based on the residual between FG and observations, modulated by the local gradient of the FG.

DAbyFA: Implementation of the method (I)

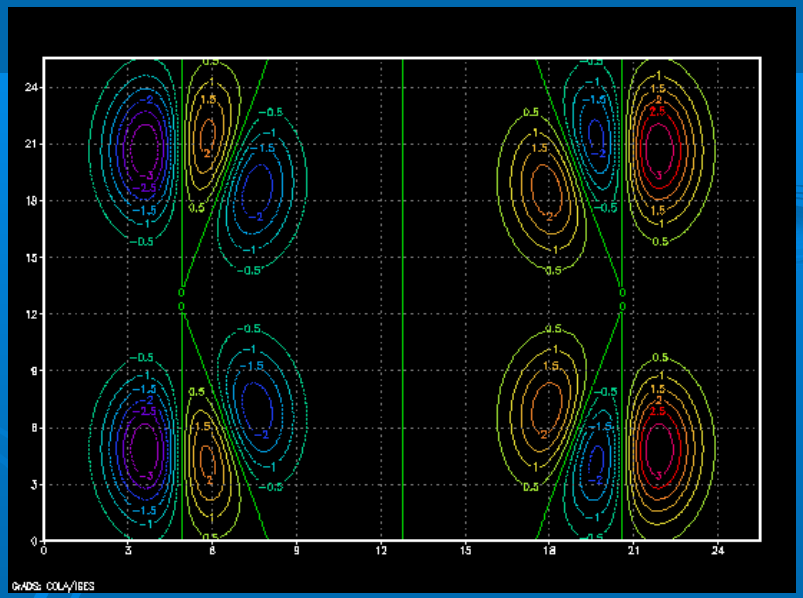
- It is not hard to solve the “alignment equation”. For “natural boundary conditions” ($q_n = 0$) on a rectangle, it is found that a very convenient way of solving it is by using spectral methods on an extended domain (2x2)

- We reflect the increment field through the horizontal and vertical central axes of the extended domain and get in this way even periodic functions. The forcing terms are the product of this increment field and the gradients, therefore they are odd functions along the corresponding direction (i.e.,along the x direction for F_x and along the y direction for F_y) and even along the other direction

F_y



F_x



DAbyFA: Implementation of the method (II)

- Functions like these satisfy the q_n=0 lateral boundary condition. Therefore, if the symmetry properties of the PDE are such that the solutions have the same symmetries as the forcing terms, solutions with the required LBCs are readily found

- And this is the case. The PDE diagonalized in (k,l) space has the simple form:

$$C_x(k,l)\,\mathbf{Q_x}(k,l) + S(k,l)\,\mathbf{Q_y}(k,l) = \mathbf{F_x}(k,l)$$

$$S(k,l)\,\mathbf{Q_x}(k,l) + C_y(k,l)\,\mathbf{Q_y}(k,l) = \mathbf{F_y}(k,l)$$

$C_i(k,l)$, $S(k,l)$ real and

$C_i(-k,l) = C_i(k,-l) = C_i(-k,-l) = C_i(k,l)$

$S(-k,l) = S(k,-l) = -S(-k,-l) = -S(k,l)$

After some elementary algebra we find that for:

$\mathrm{Re}[\mathbf{F_x}(k,l)] = 0$; $\mathrm{Im}[\mathbf{F_x}(k,l)] = -\mathrm{Im}[\mathbf{F_x}(-k,l)]$; $\mathrm{Im}[\mathbf{F_x}(k,l)] = \mathrm{Im}[\mathbf{F_x}(k,-l)]$

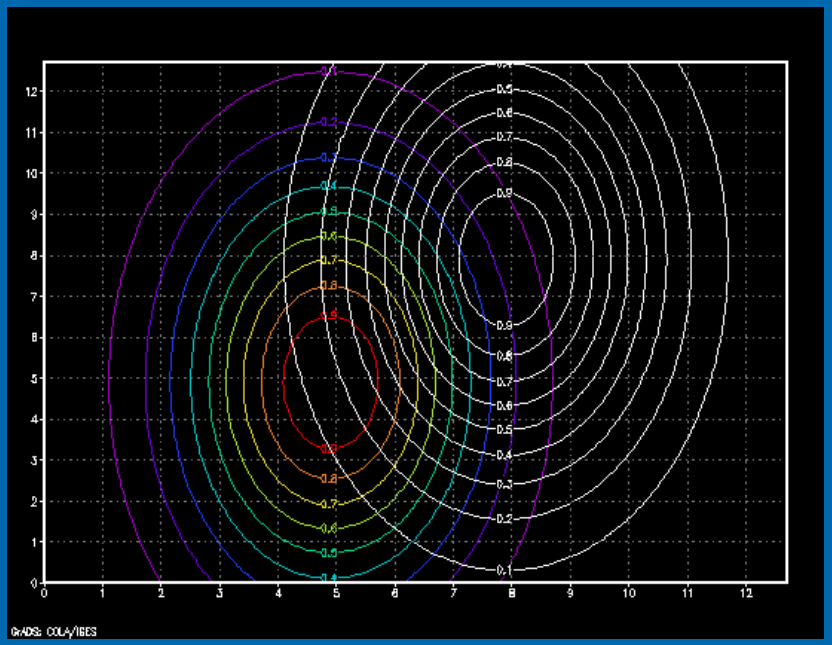
and

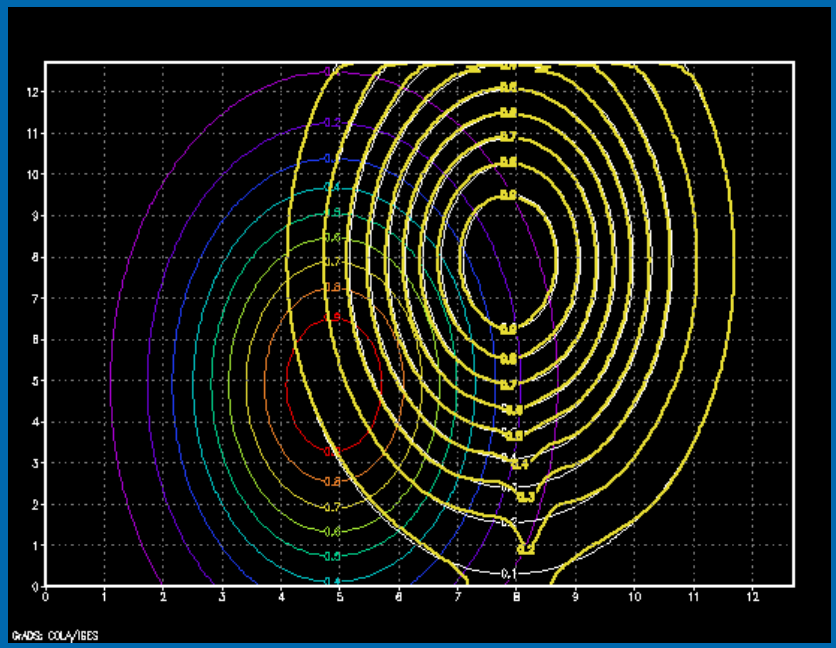
$\mathrm{Re}[\mathbf{F_y}(k,l)] = 0$; $\mathrm{Im}[\mathbf{F_y}(k,l)] = \mathrm{Im}[\mathbf{F_y}(-k,l)]$; $\mathrm{Im}[\mathbf{F_y}(k,l)] = -\mathrm{Im}[\mathbf{F_y}(k,-l)]$

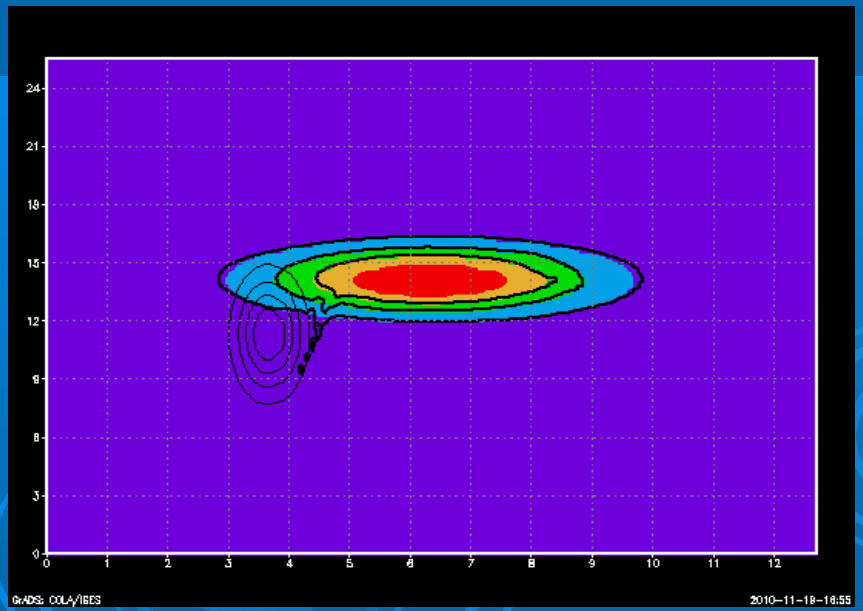
The $\mathbf{Q_x}(k,l)$ and the $\mathbf{Q_y}(k,l)$ satisfy the same symmetries

- In addition, the use of the extended zone allows to get solutions with $<\mathbf{q}> \neq 0$ in the area of interest in spite of the fact that the above equation is not invertible for k = l = 0

DAbyFA: Implementation of the method (III)



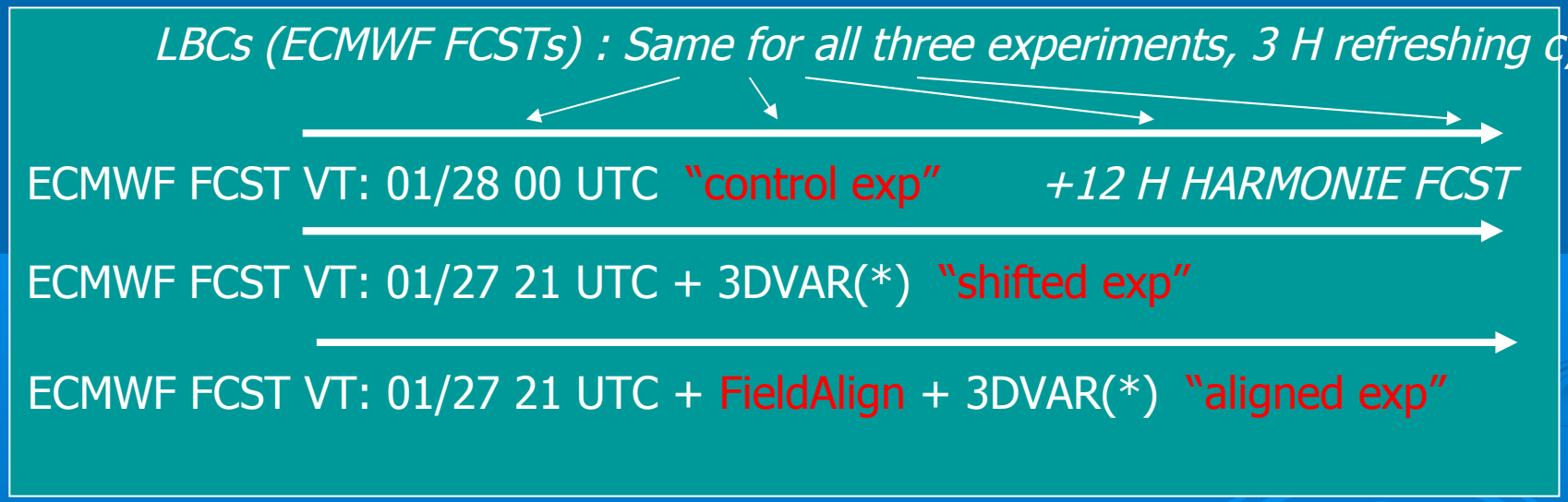




DAbyFA: Testing the Method

- This scheme has been tested using HARMONIE 36h1.3 (2.5Km,60L) fields as surrogates of FGs and observations.

- The testing exercise comprises so far a single case with some weather activity over the Western Mediterranean off the Iberian coast. The exercise consists of three runs of the HARMONIE 3Dvar (no surface ass) + NH + AROME physics NWP system

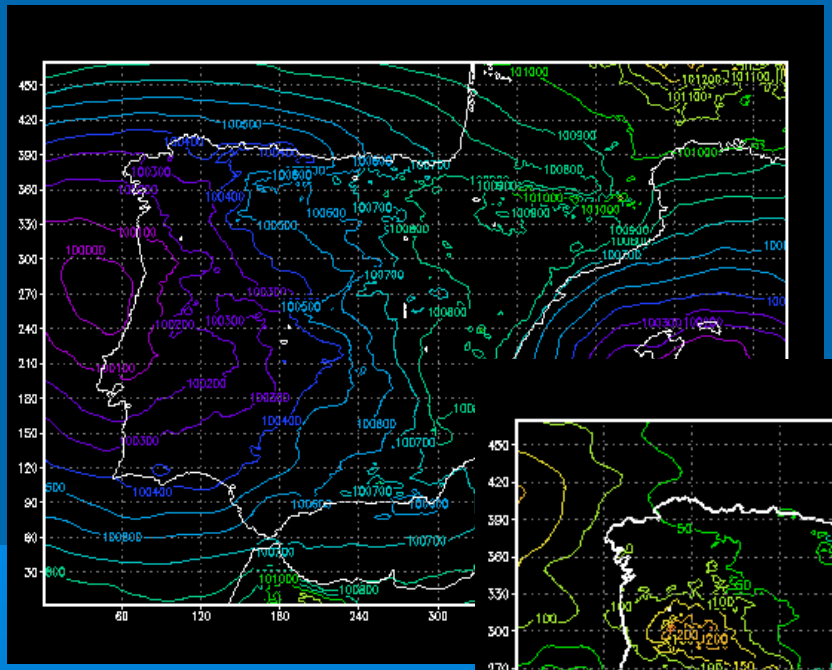


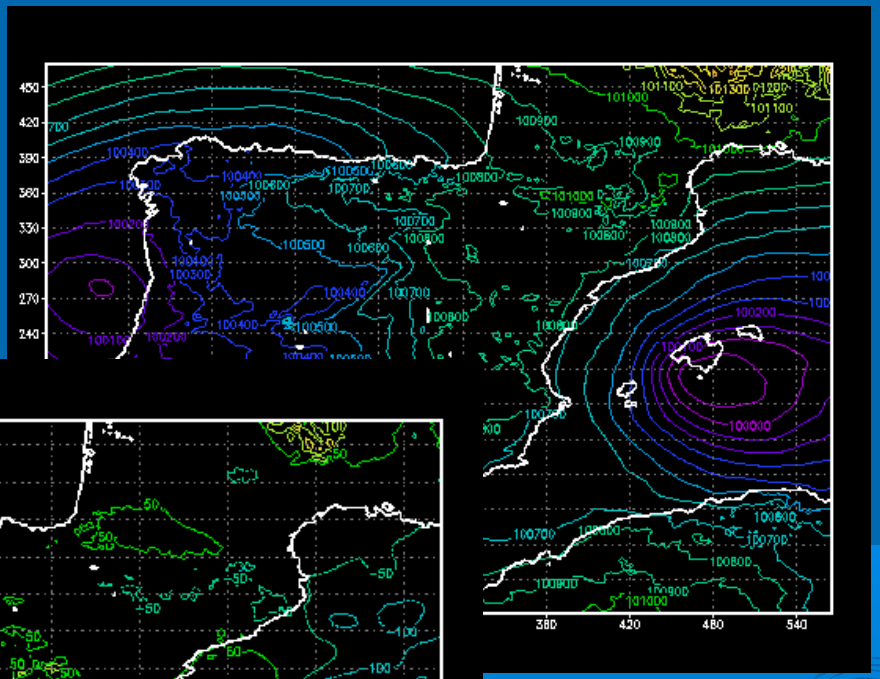
(*) Observations for the 3DVar analyses are "bogus obs". They are read off the fields used as initial conditions (i.e. the ECMWF FCST VT: 01/28 00 UTC)

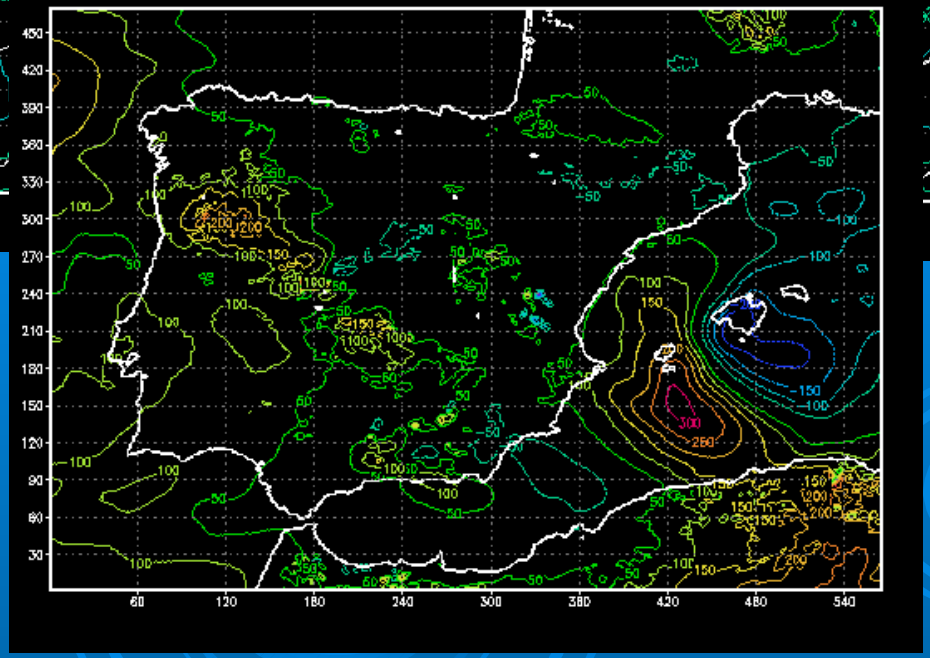
DAbyFA: Testing the Method

MSLP field valid for 27th 21 UTC
It is a +15H ECMWF forecast, with analysis time 27th 06 UTC

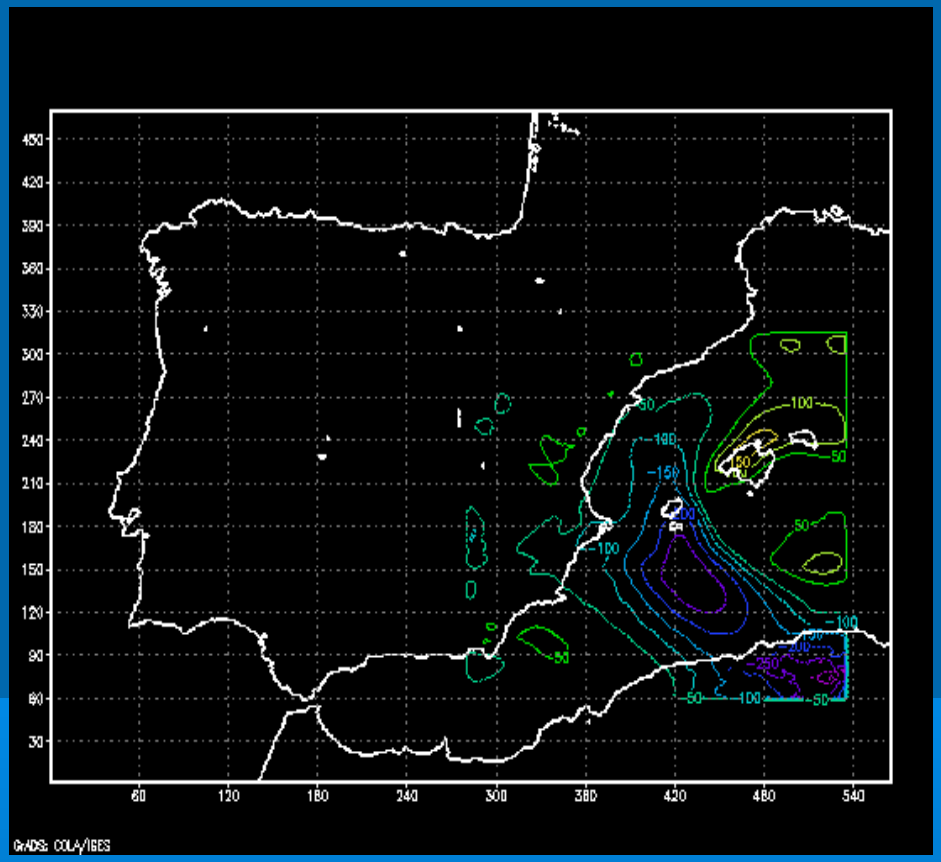
MSLP field valid for 28th 00 UTC
It is a +6H ECMWF forecast, with analysis time 27th 18 UTC

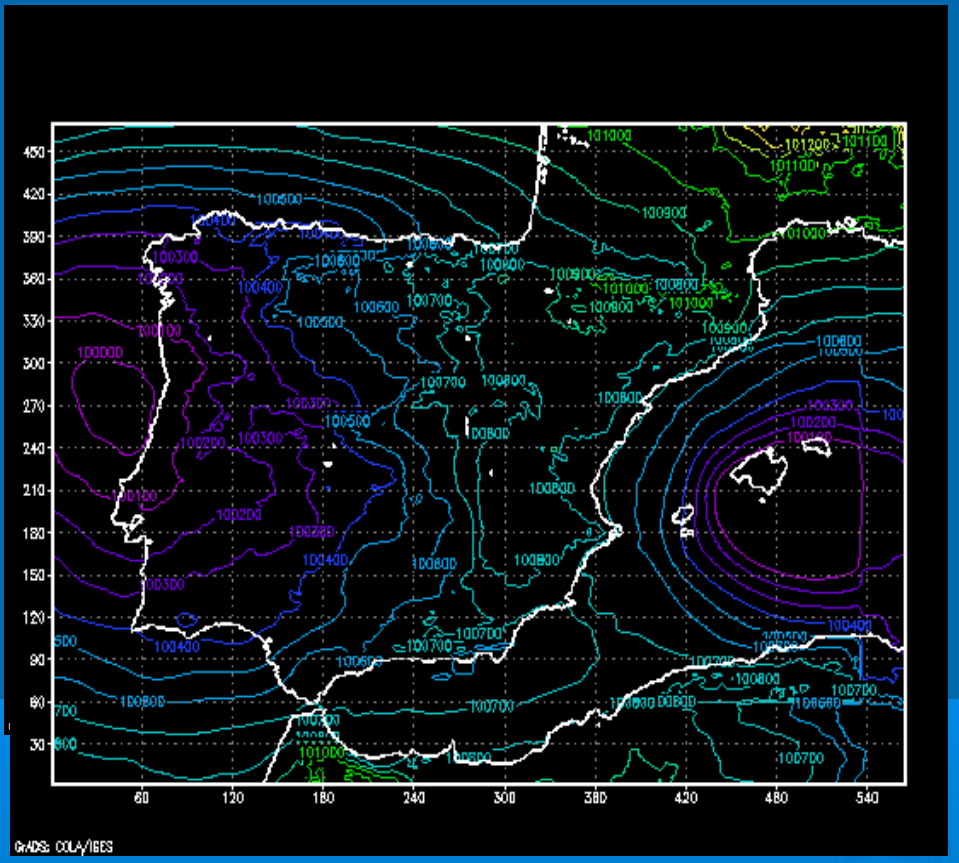


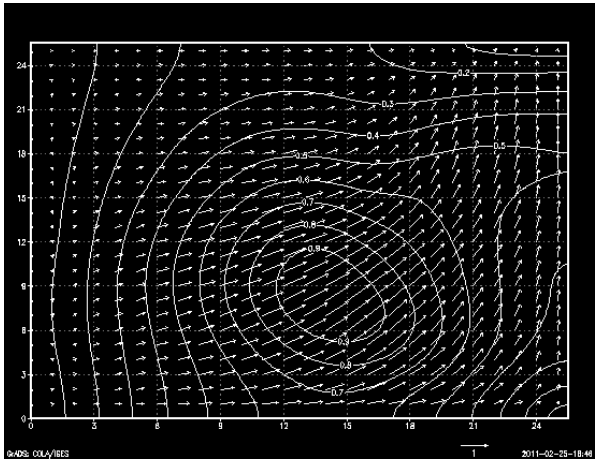


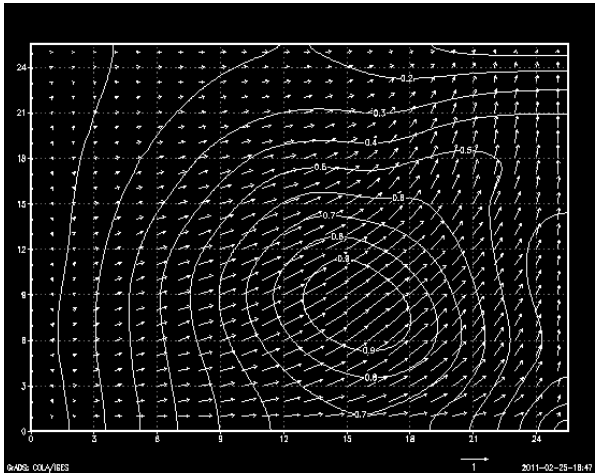


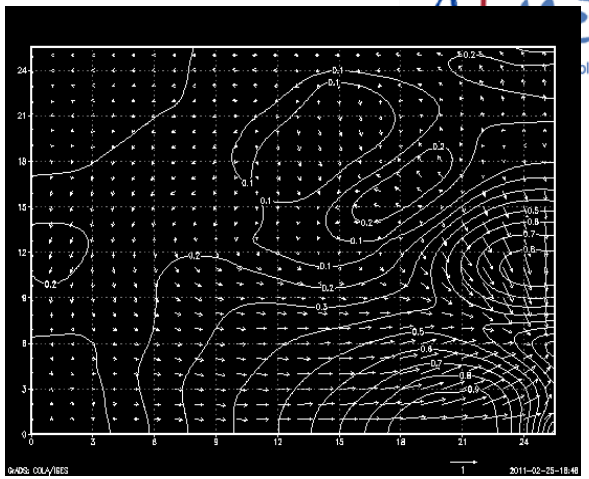
DAbyFA: Testing the Method

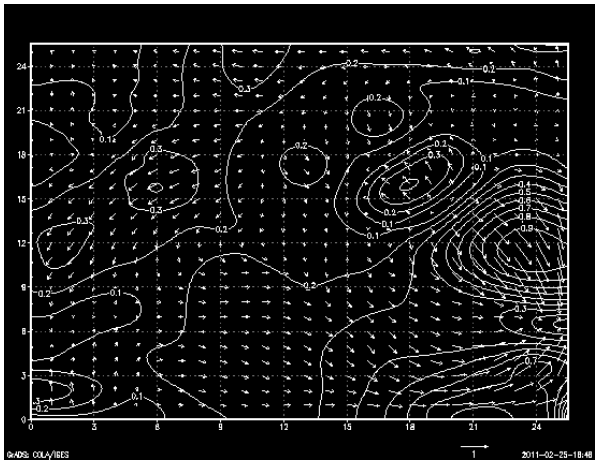


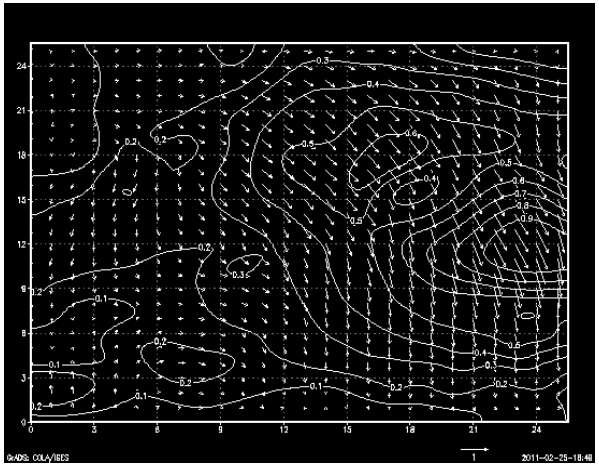


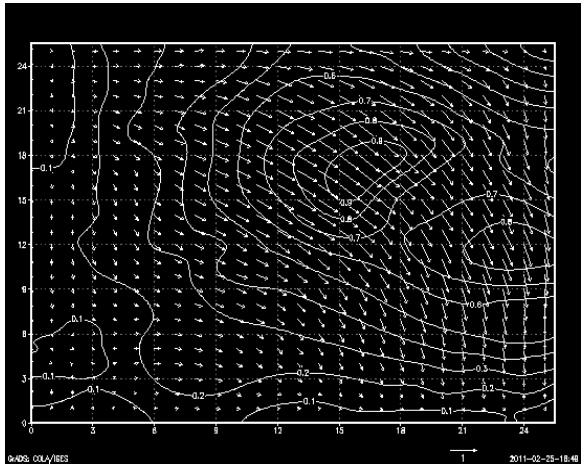

GrADS: COLA/IGES
2011-02-25-18:46

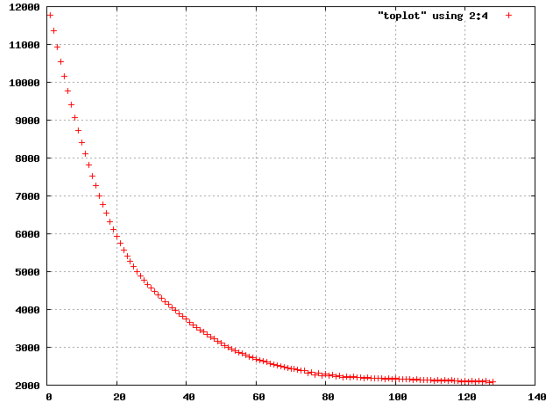

GrADS: COLA/IGES
2011-02-25-18:47


GrADS: COLA/IGES
2011-02-25-18:48

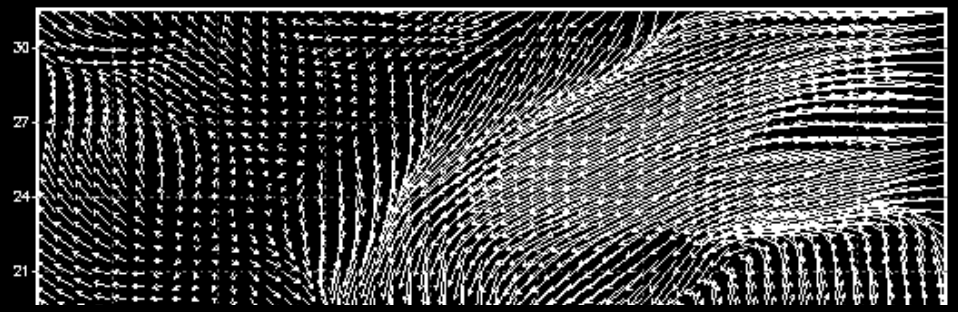

GrADS: COLA/IGES
2011-02-25-18:48

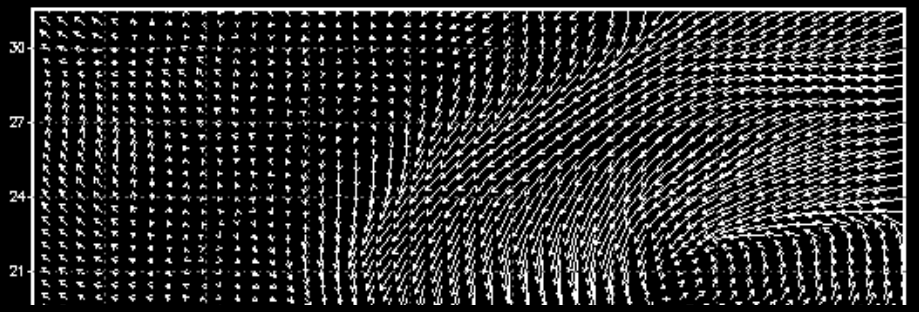

GrADS: COLA/IGES
2011-02-25-18:48

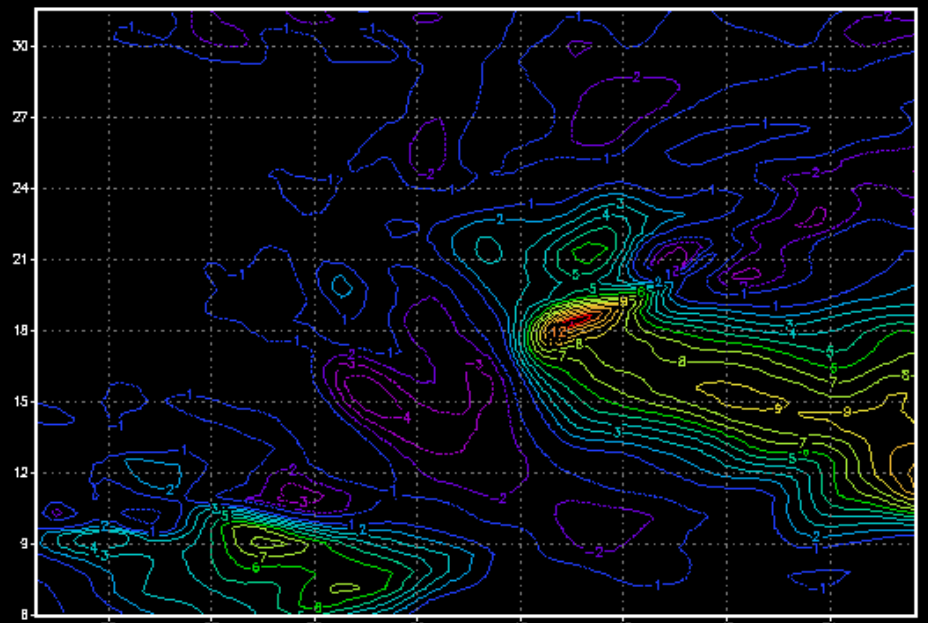

GrADS: COLA/IGES
2011-02-25-18:48

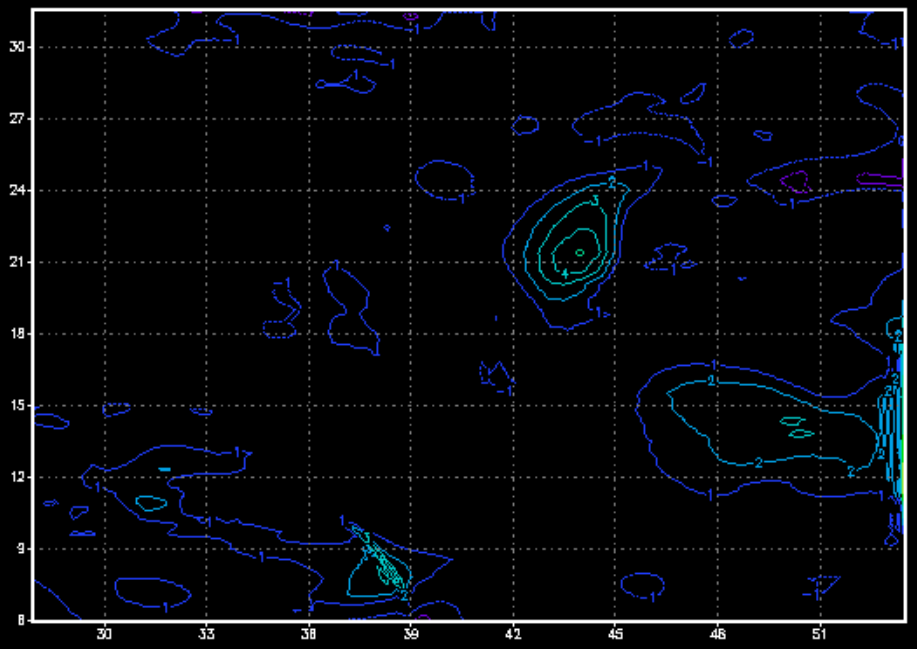

"toplot" using 2:4

DAbyFA: Testing the Method

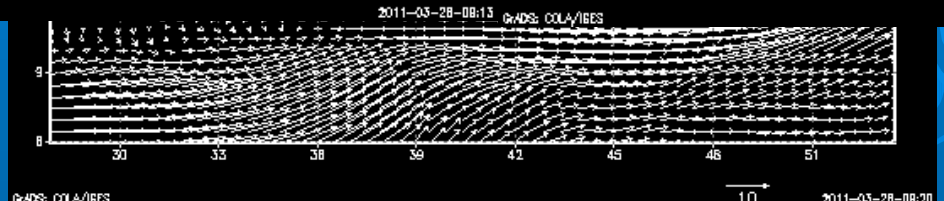








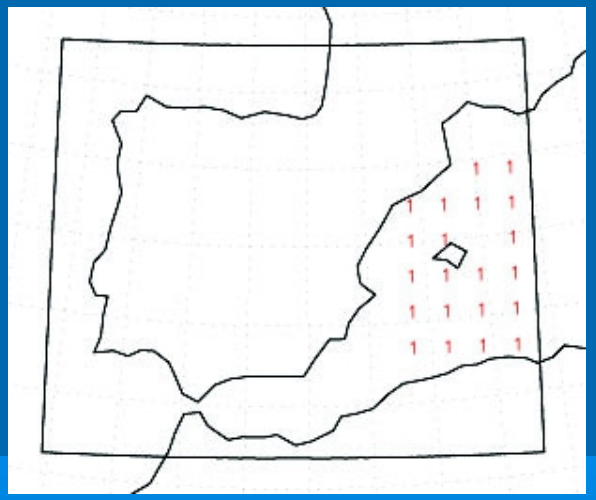
GrADS: COLA/IGES



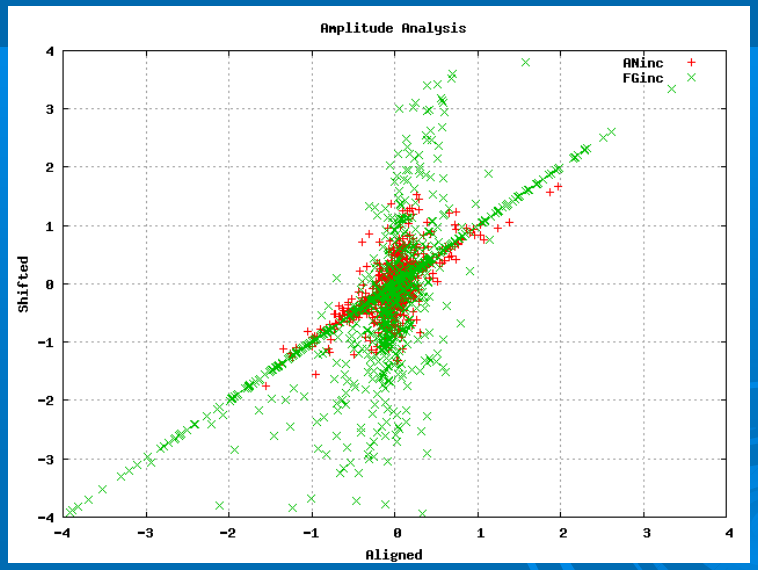
2011-03-28-08:18

DAbyFA: Testing the Method

• The scheme is a two-step one. After the correction in position, a correction in Amplitude is still required. To include this step in this test, a fairly dense network Of in-situ observations over the area of interest was defined and the value of the ps, T, u, v and q parameters read off the initial conditions fields acting Here as "truth"

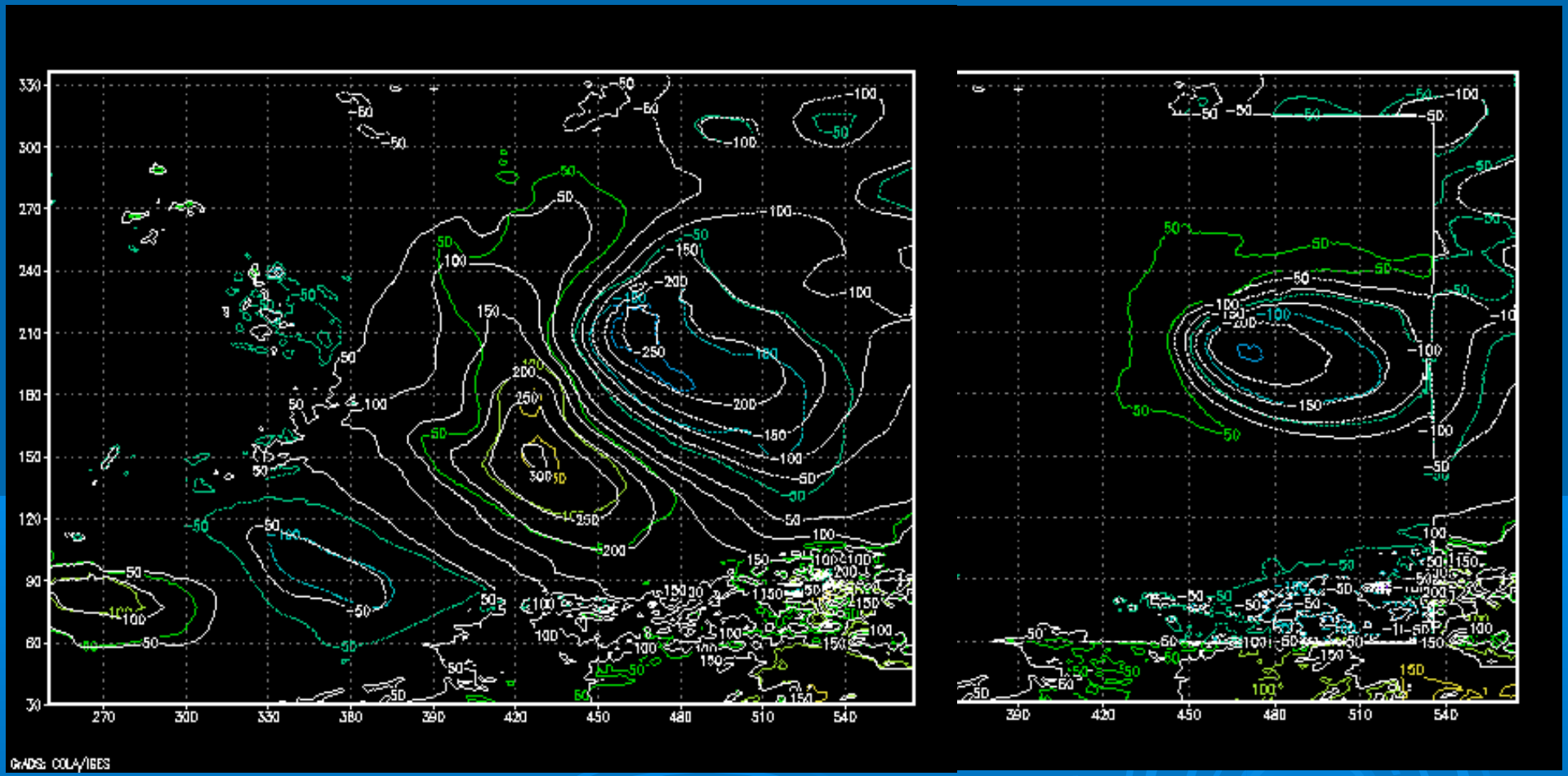


• Each position contains: Ps observations plus ten levels (up to 200 hPa approx.) of u, v, q and T. In total, 816 observations were assimilated. Some "ad-hoc" quality constrains currently implemented in the system (i.e., reduction zone) were removed.



• No observation (in either run) was screened out by the QC module, that is, all fell within Tolerance limits. The distribution of obs Increments, particularly for the FG inc., looks however quite different in both cases. *The same calibration parameters were used in both cases*

DAbyFA: Testing the Method



Evaluation of Results

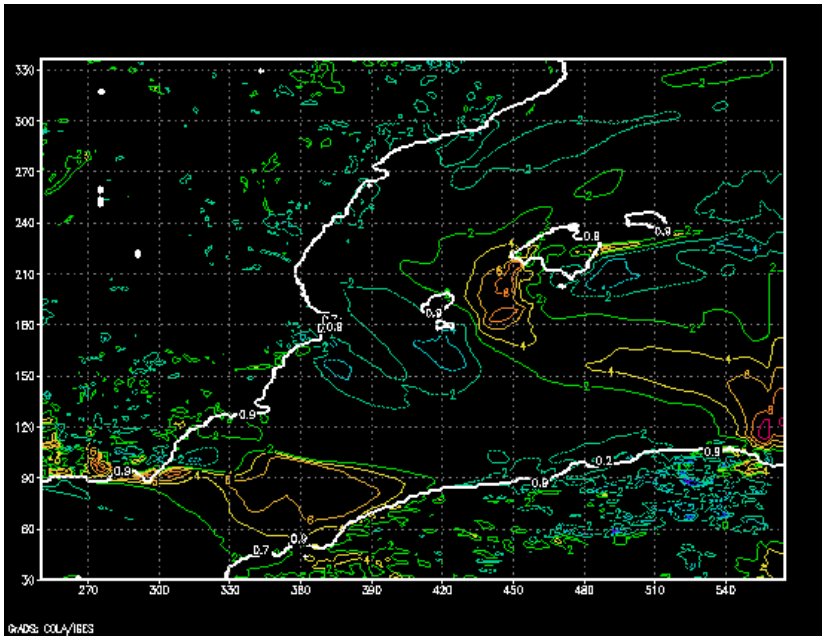
• The results of the two experiments: “shifted” and “aligned” have been compared with the control run in order to gauge the impact of the alignment

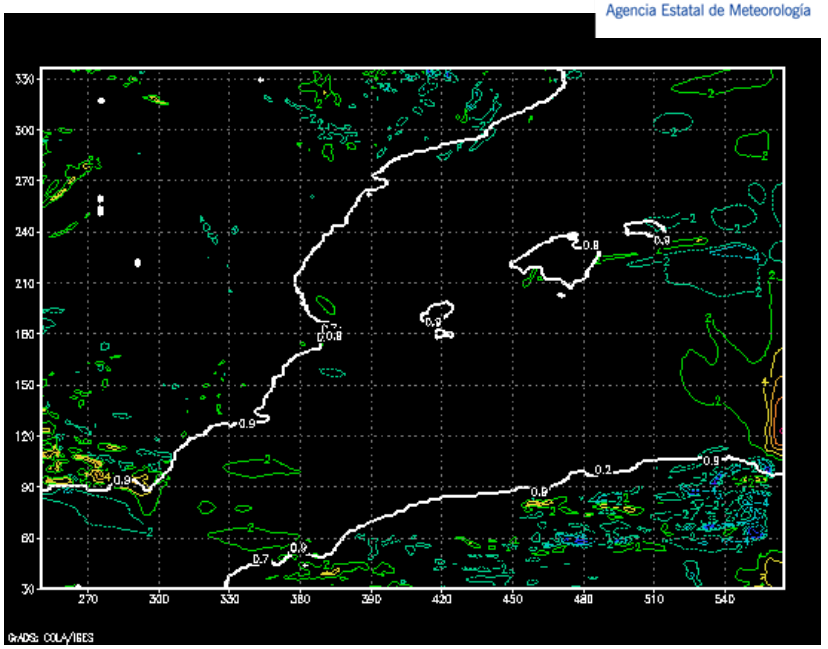
• The results are very positive. During the first hours of the integration the impact on important parameters like precipitation and wind near the surface is apparent. This impact dilutes afterwards, in this case after 4-5 hours in the wind field and 2-3 hours in the precipitation field

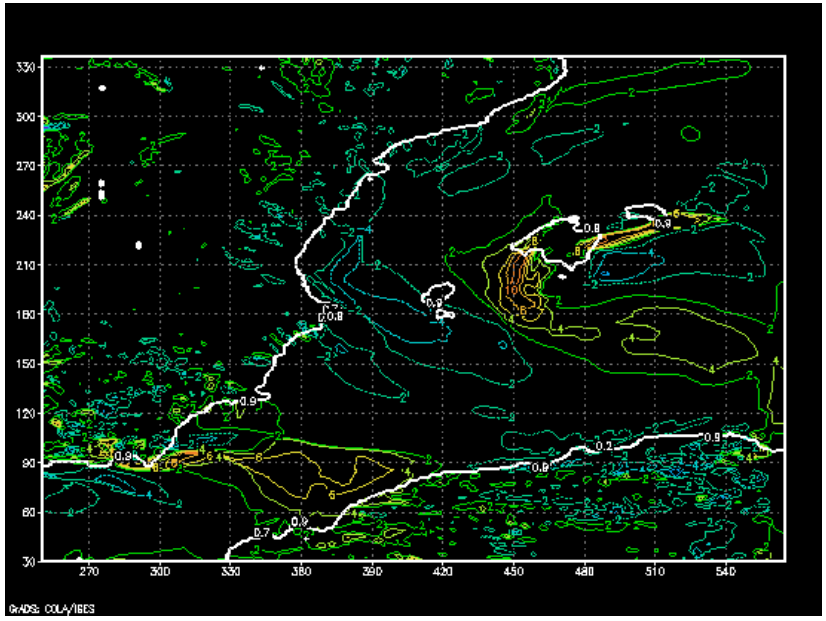
• This result, “short-range impact”, which is well documented in other “comparable” impact studies (e.g., assimilation and/or blending with radar data), is enhanced in this study due to its own specific characteristics, namely, the proximity of the low dominating the meteorological situation to the border of the domain. It is clear that after a few hours, the LBCs take over control of the run

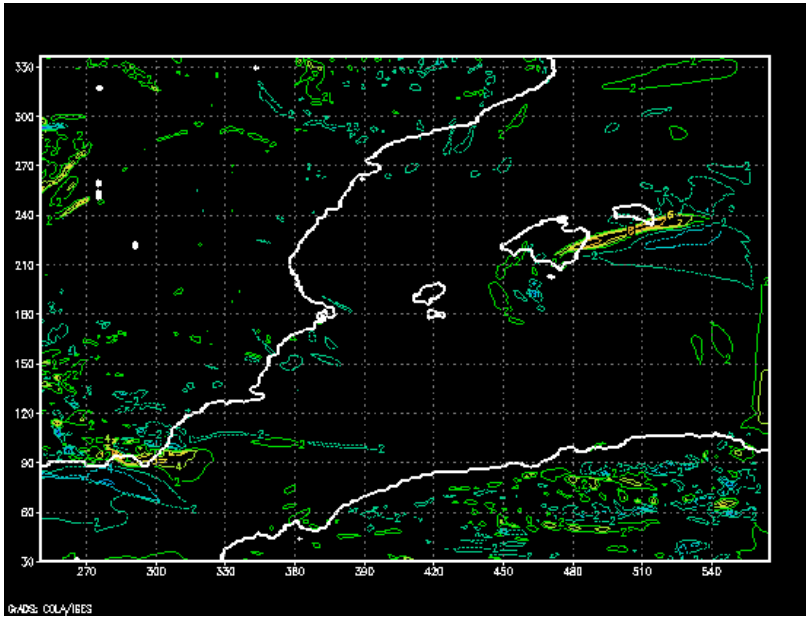
Shifted WIND SPEED (@30m) Error (cntl – exp)

Aligned



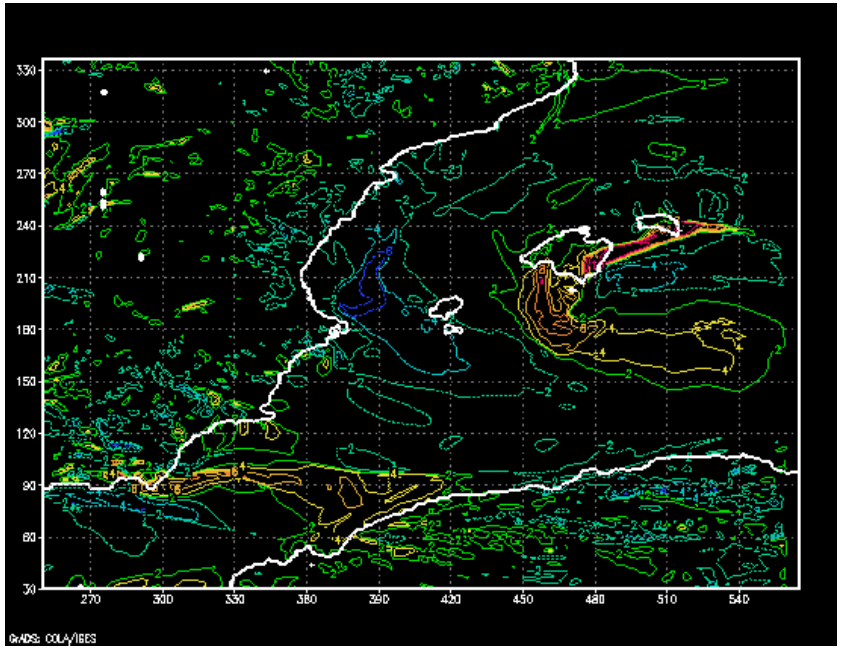


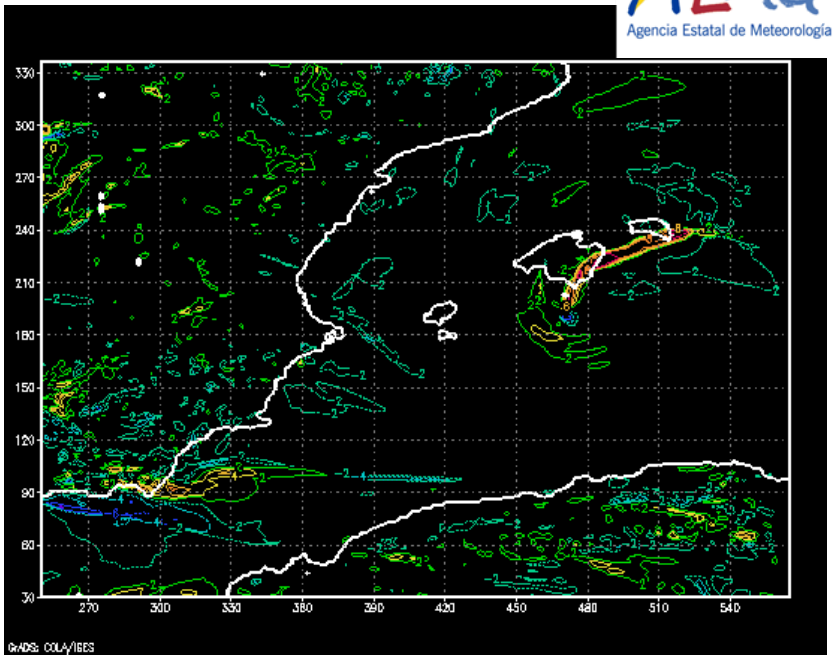


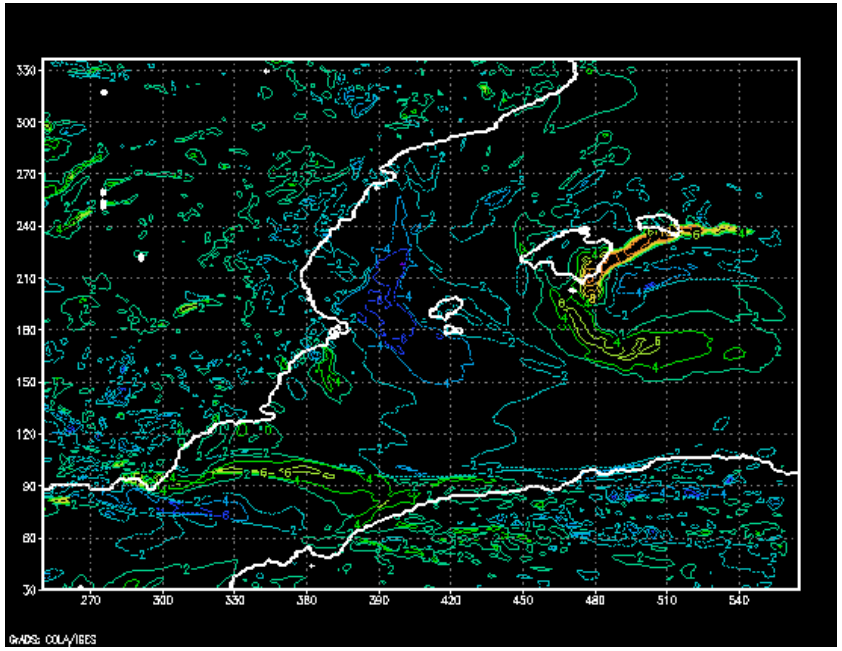


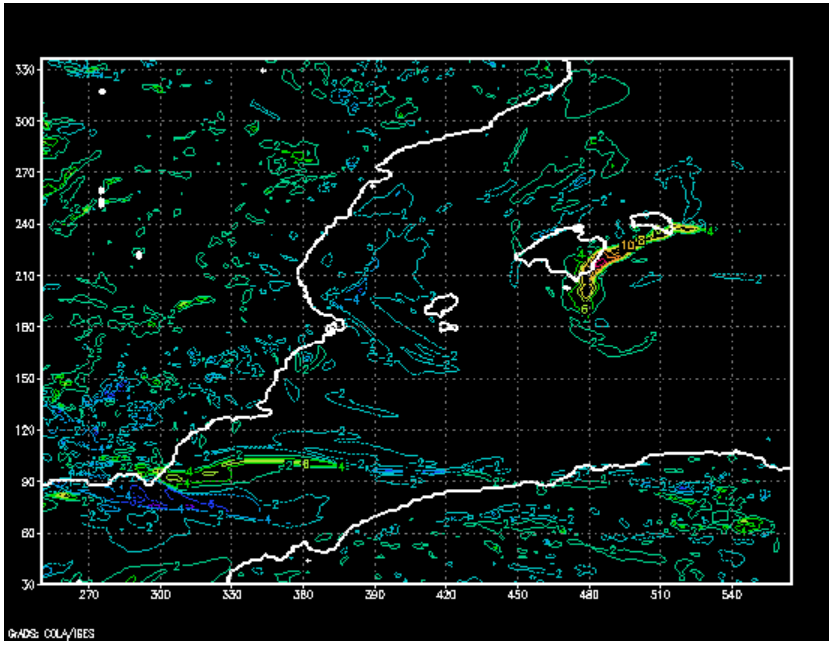
Shifted WIND SPEED (@30m) Error (cntl –exp)

Aligned



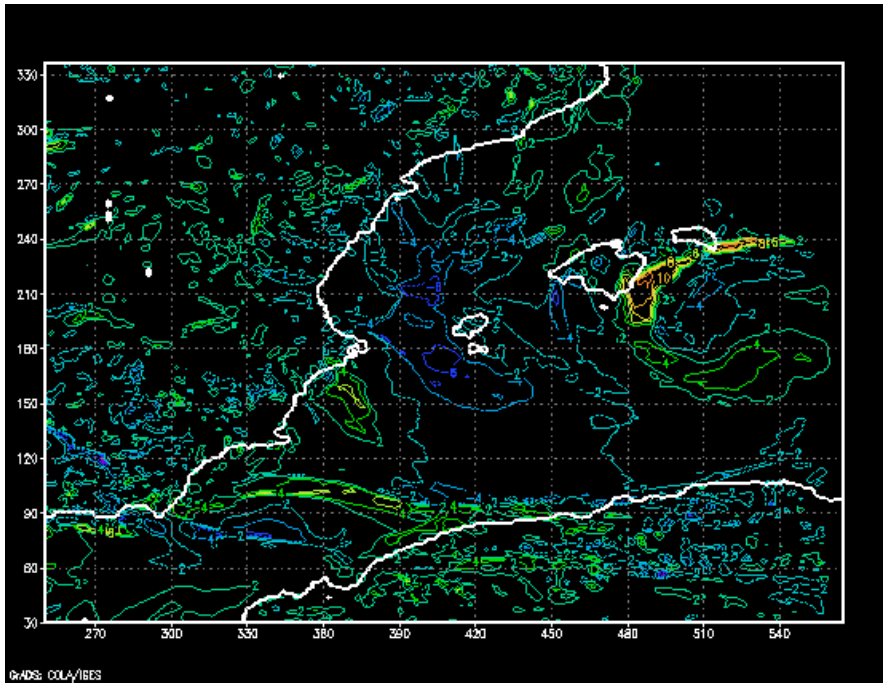


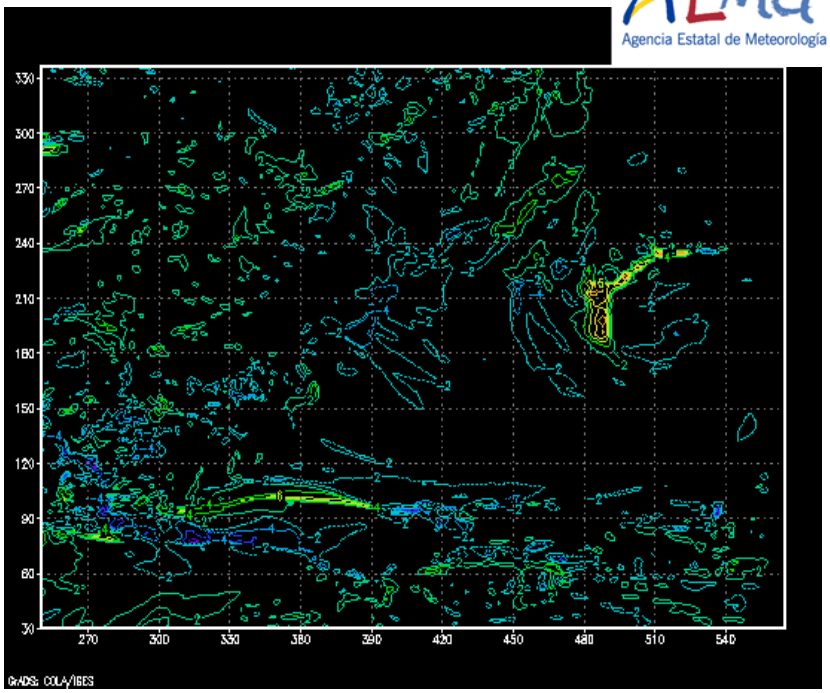


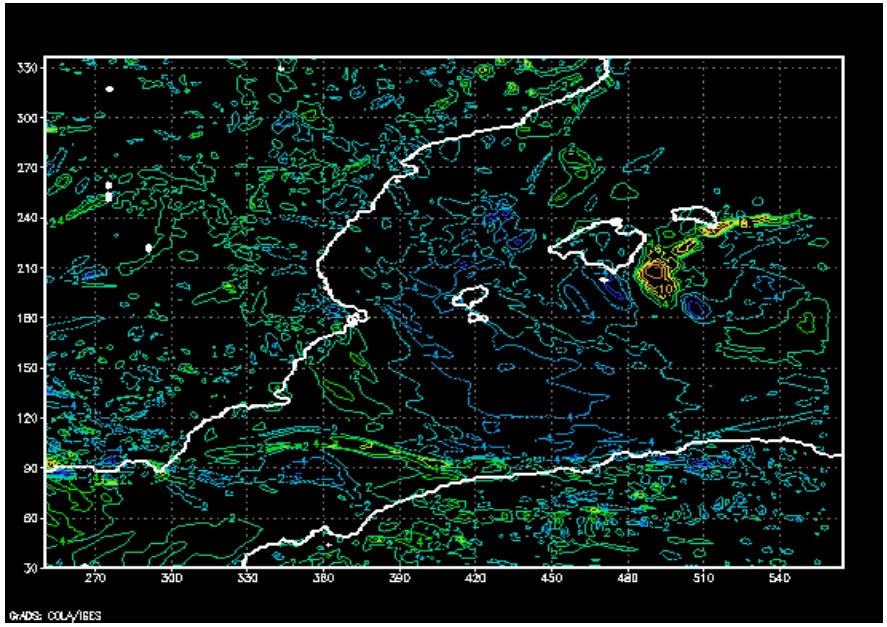


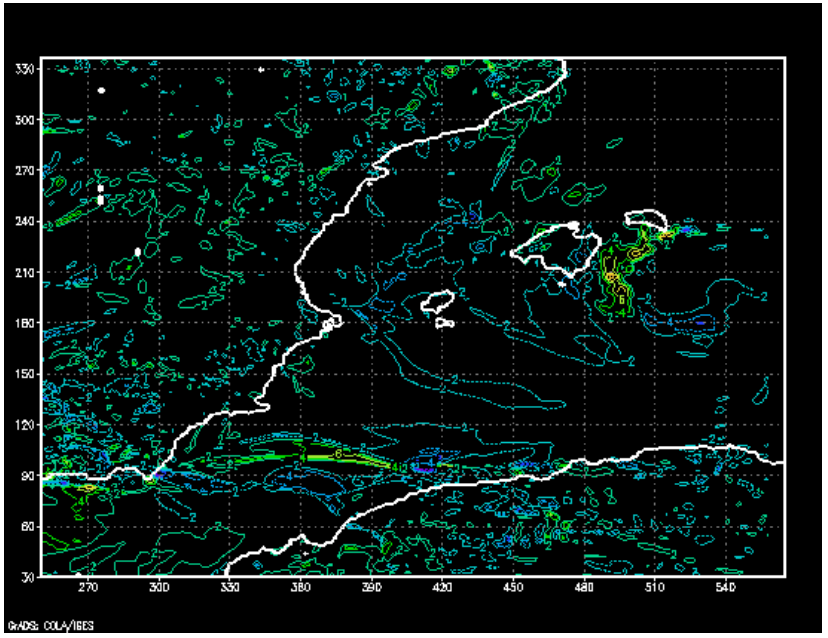
Shifted WIND SPEED (@30m) Error (cntl – exp)

Aligned



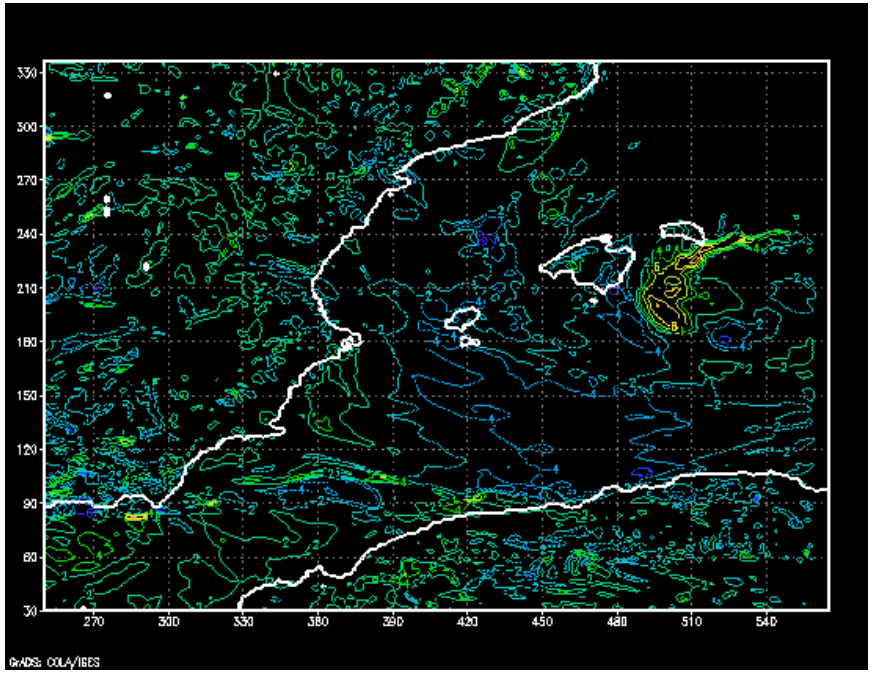


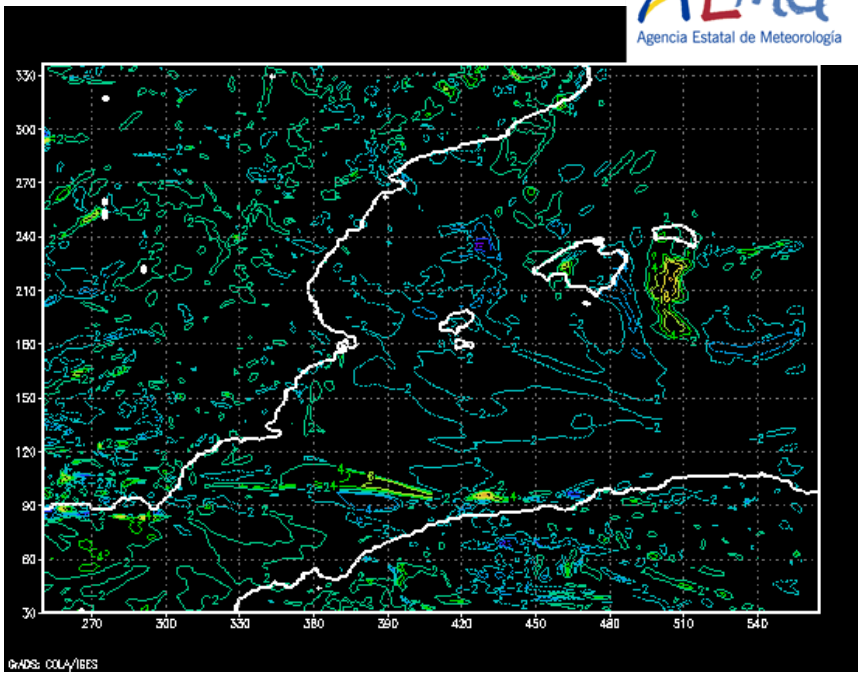


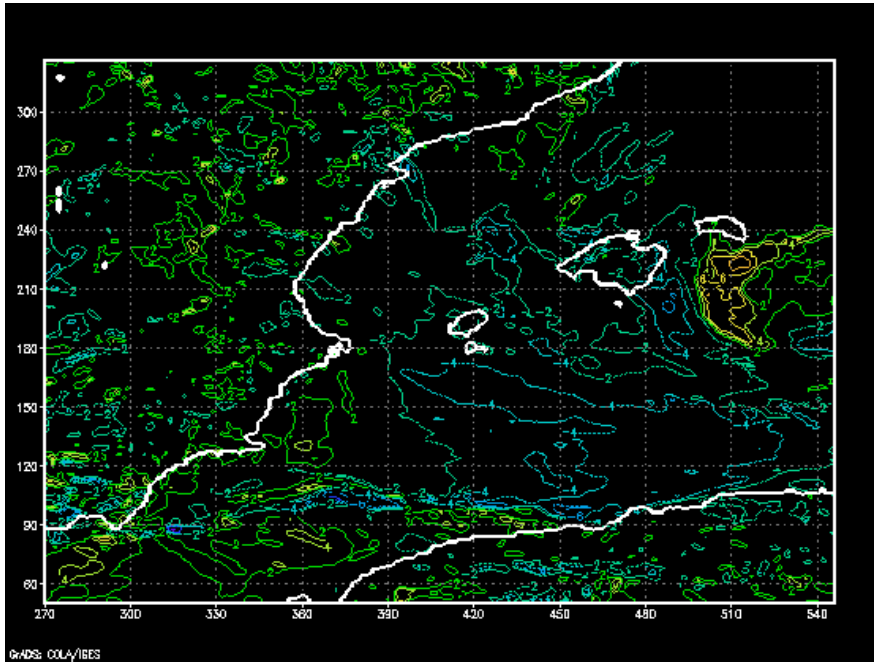


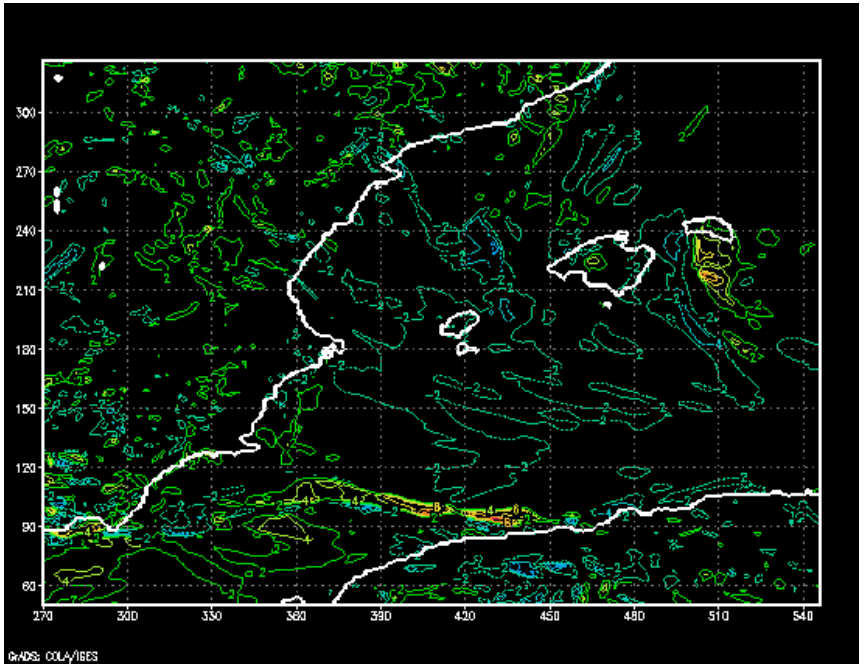
Shifted WIND SPEED (@30m) Error (cntl – exp)

Aligned



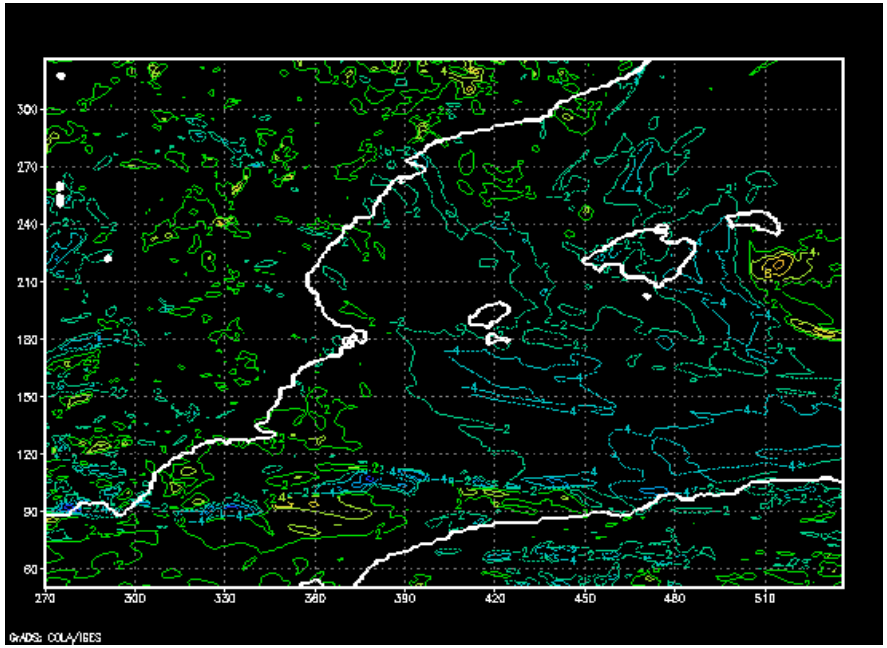


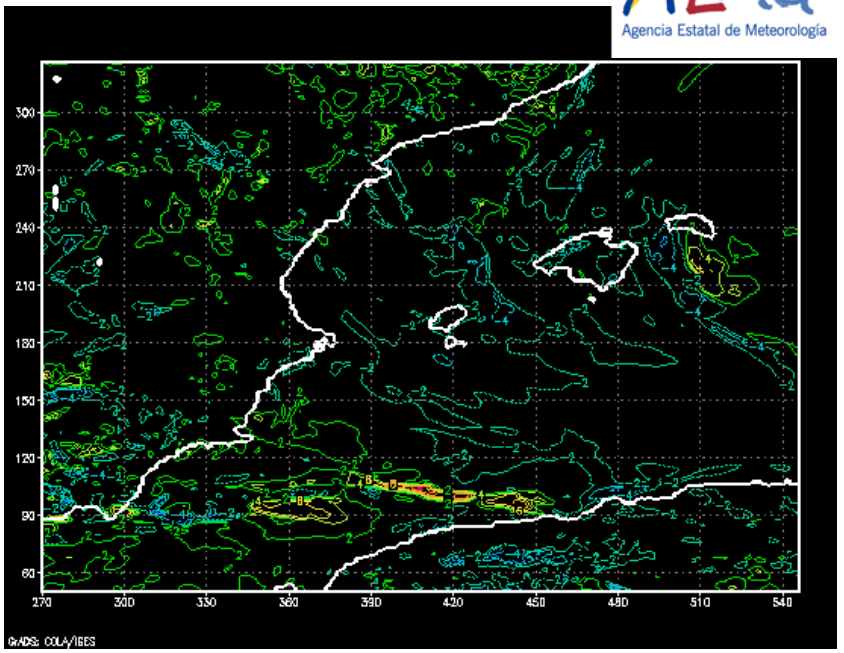


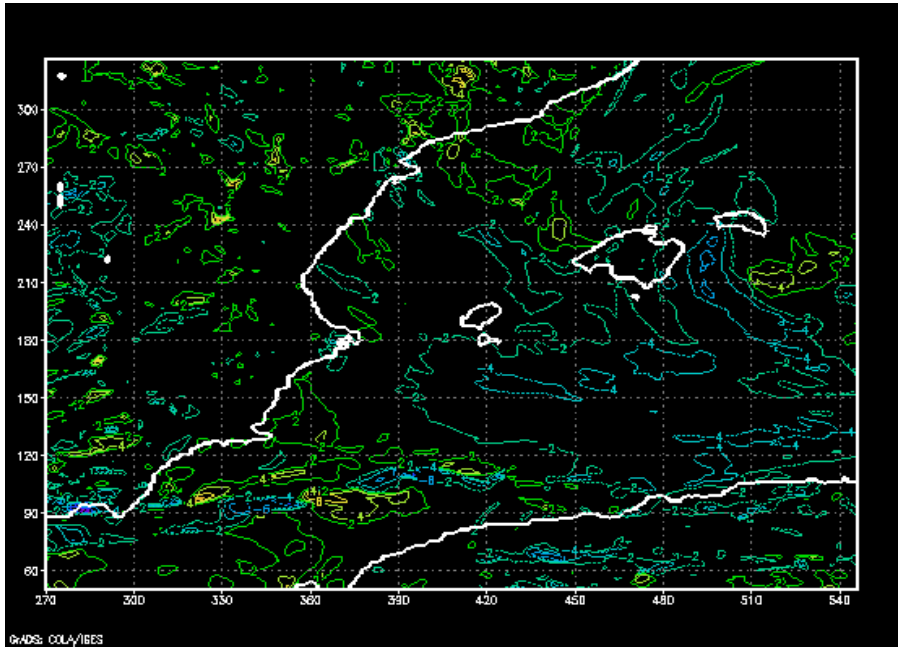


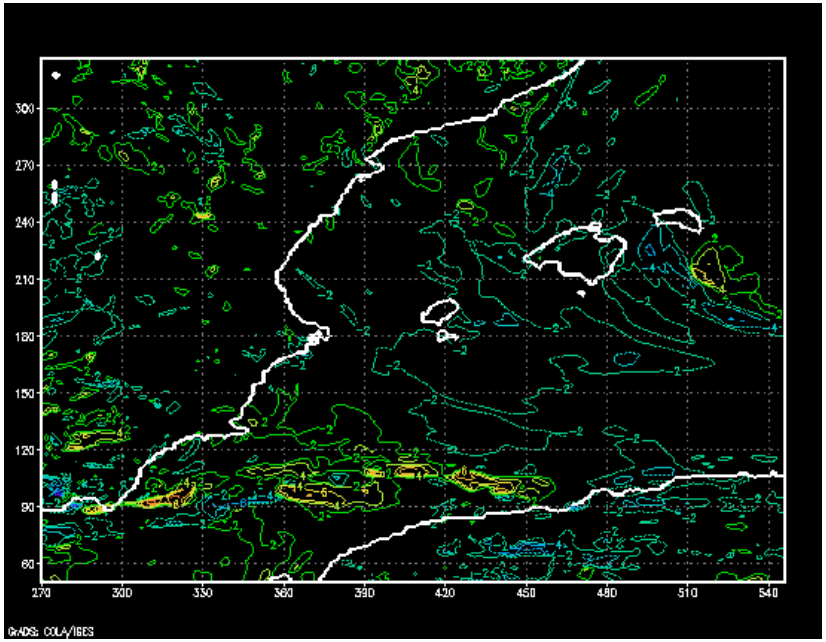
Shifted WIND SPEED (@30m) Error (cntl – exp)

Aligned

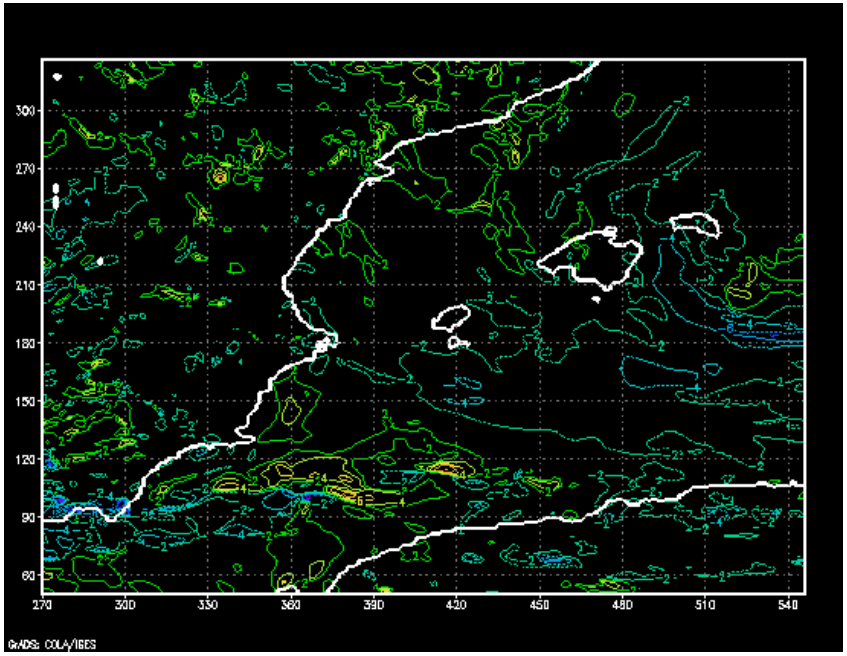


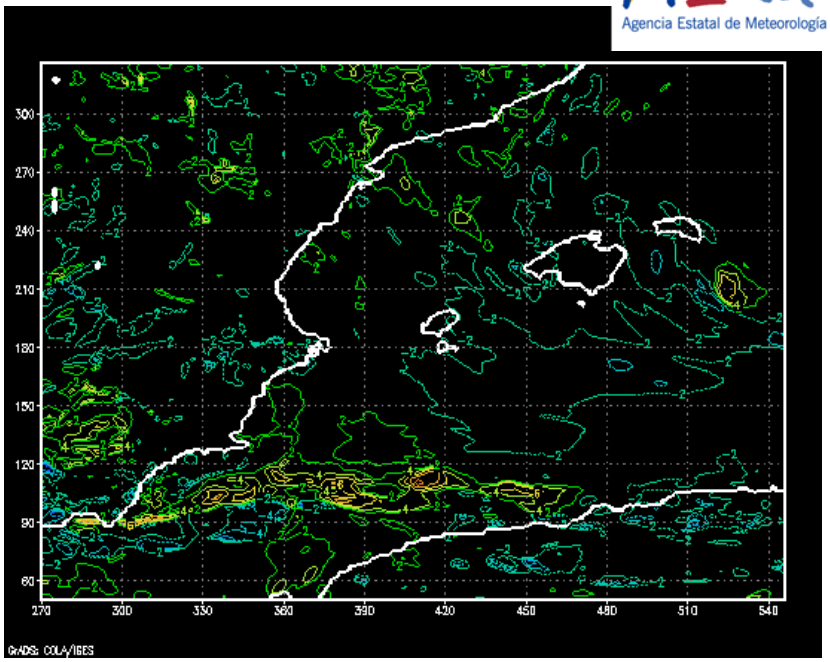


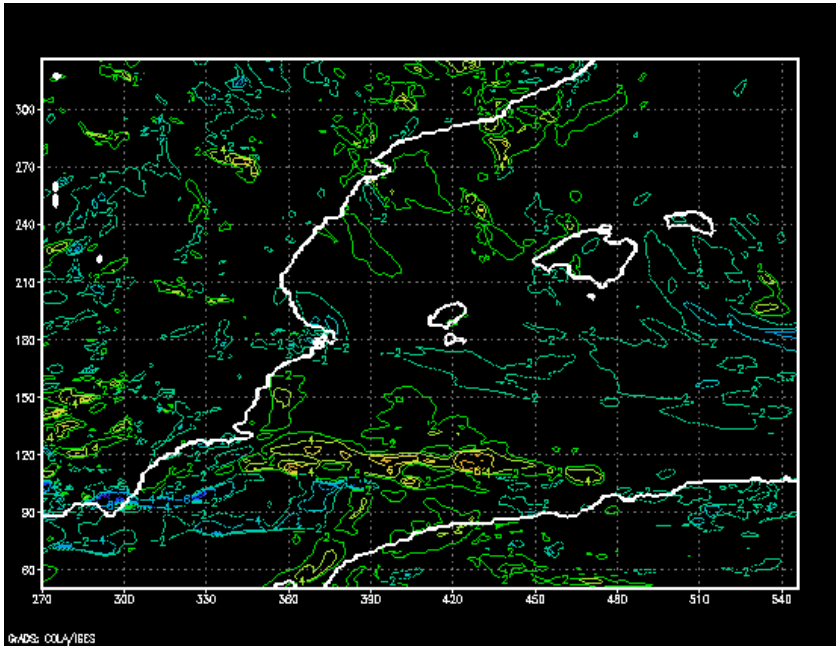


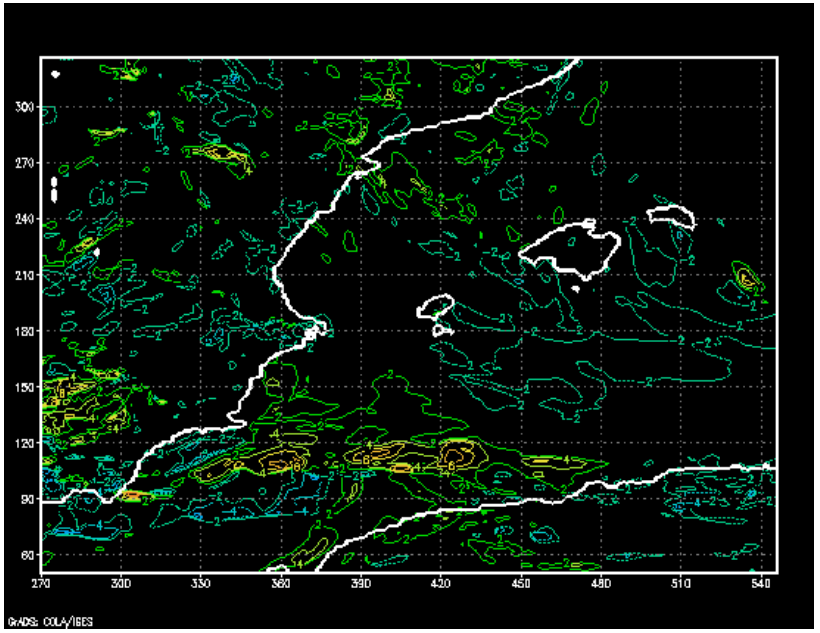


Shifted WIND SPEED (@30m) Error (cntl –exp) Aligned



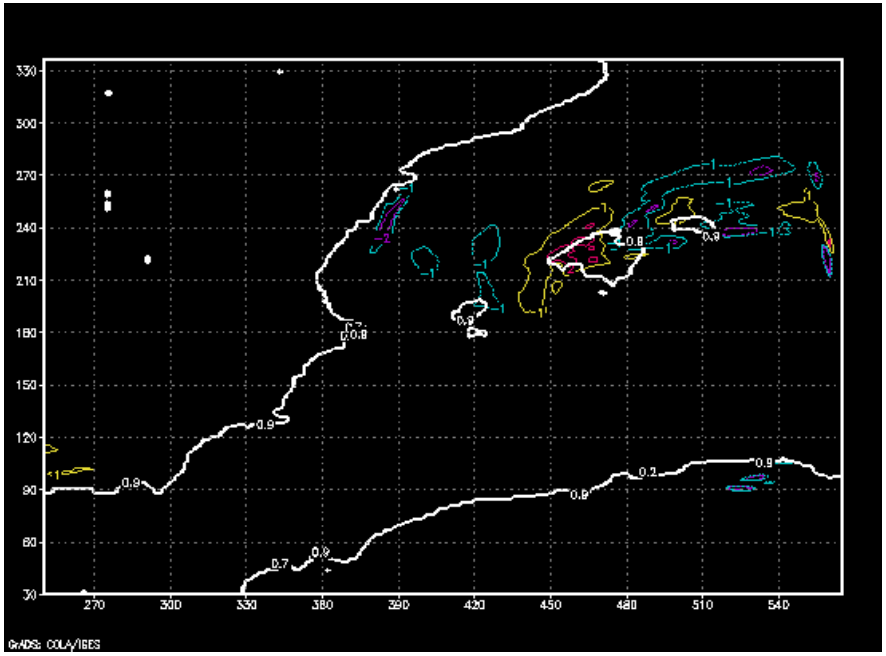


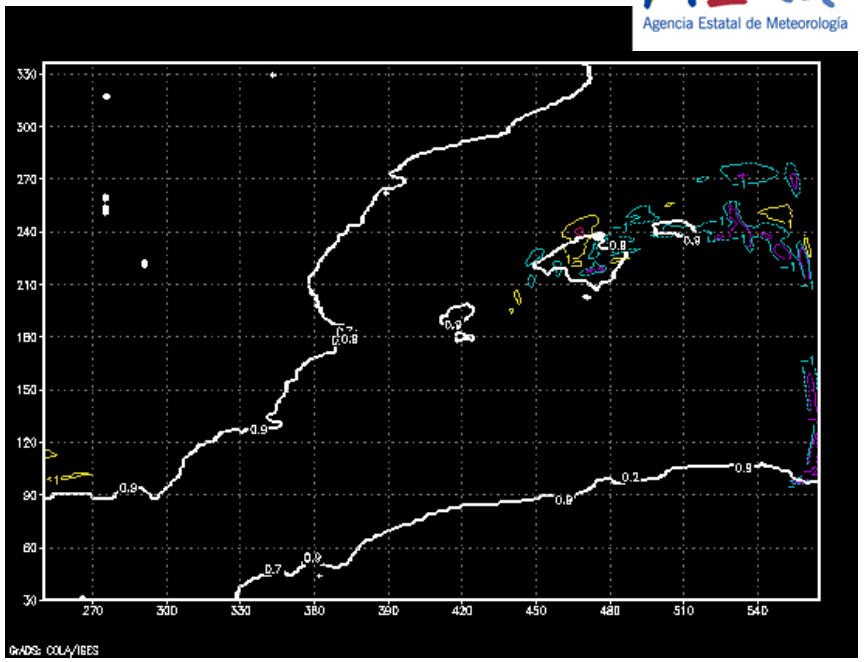


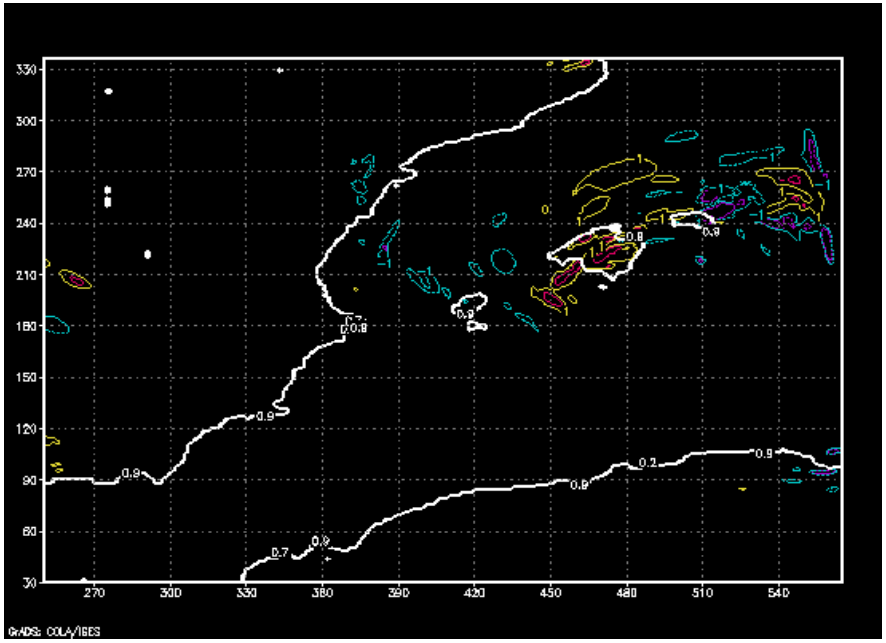


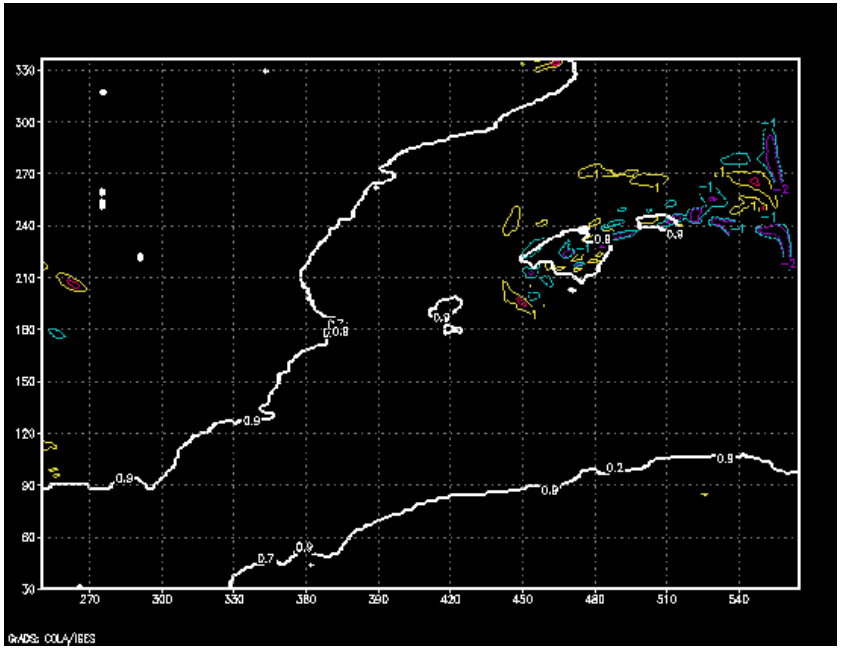
Shifted SIM Z (@30m) Error(cntl - exp)

Aligned



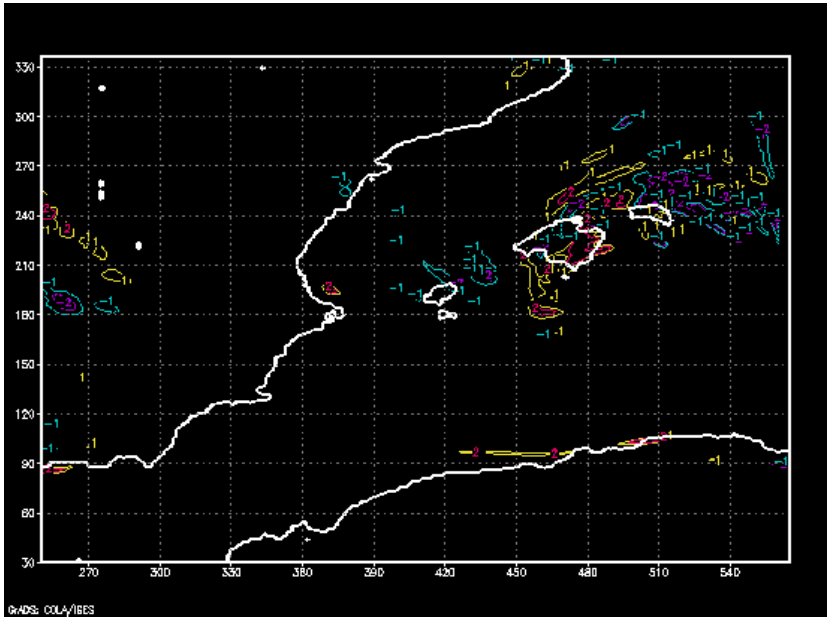


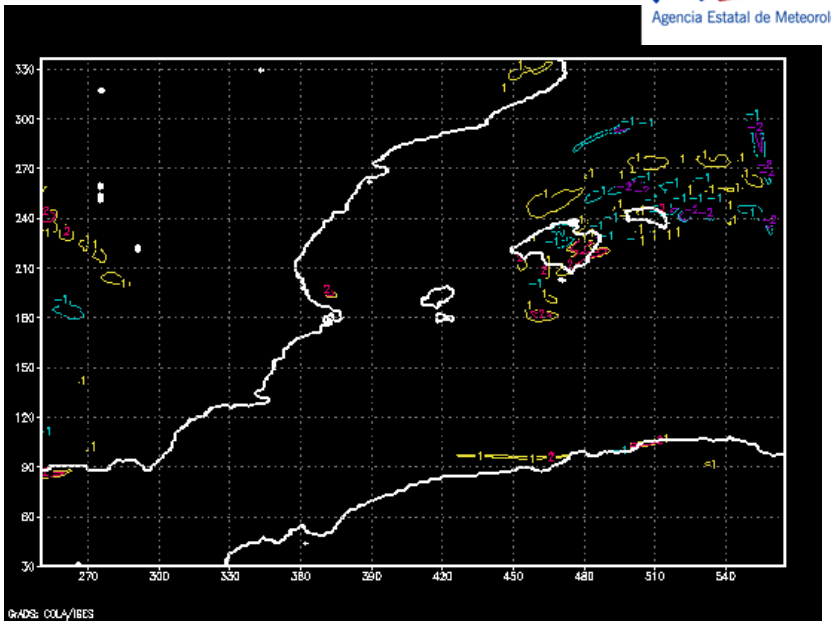


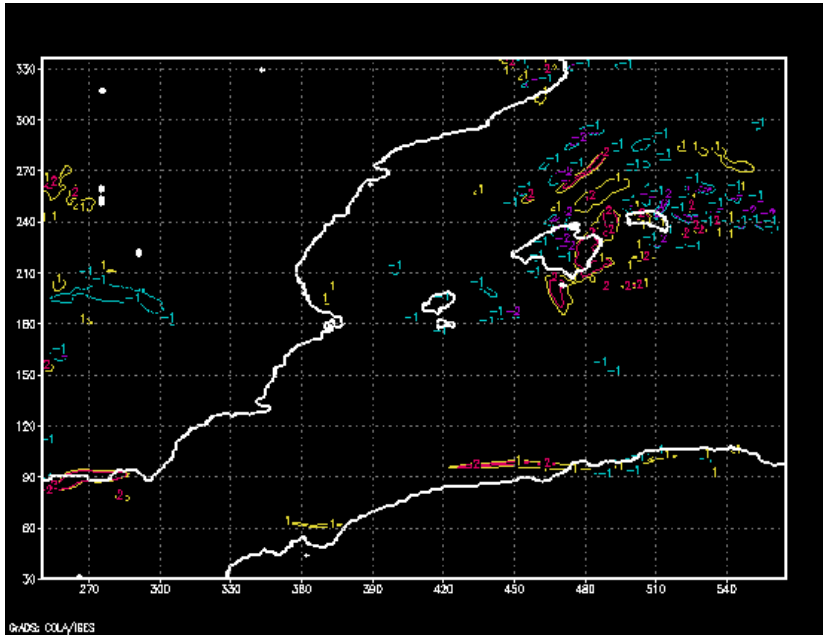


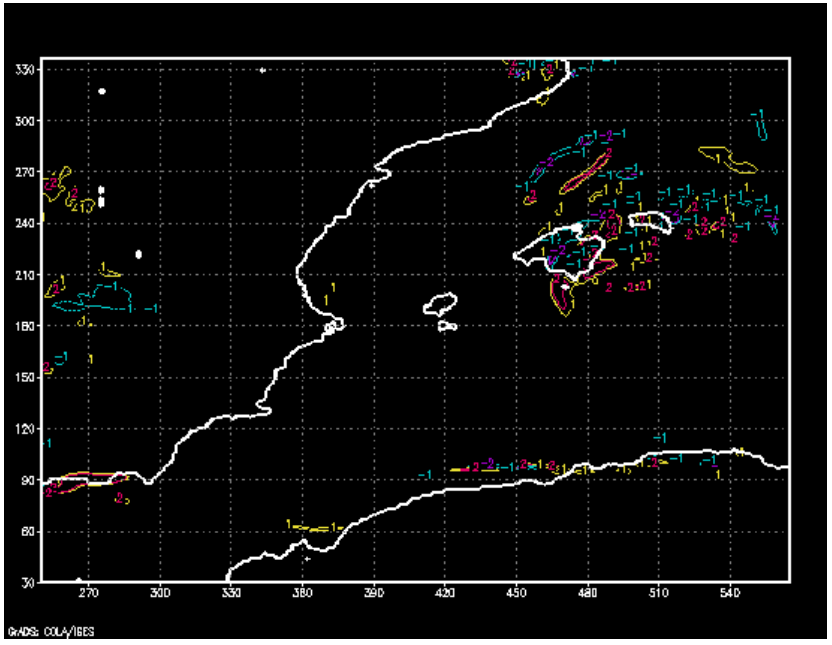
Shifted SIM Z (@30m) Error (cntl - exp)

Aligned



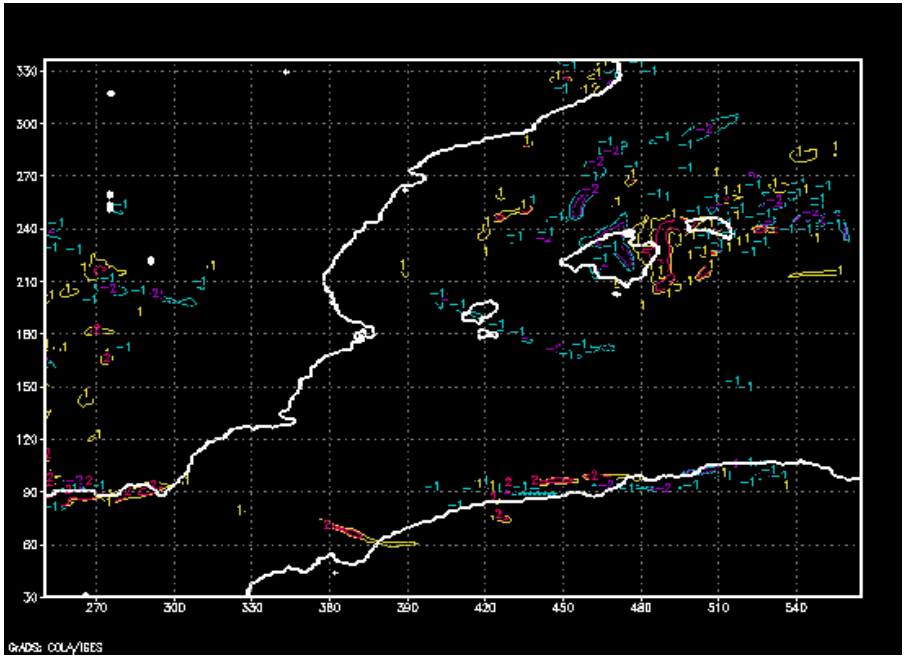


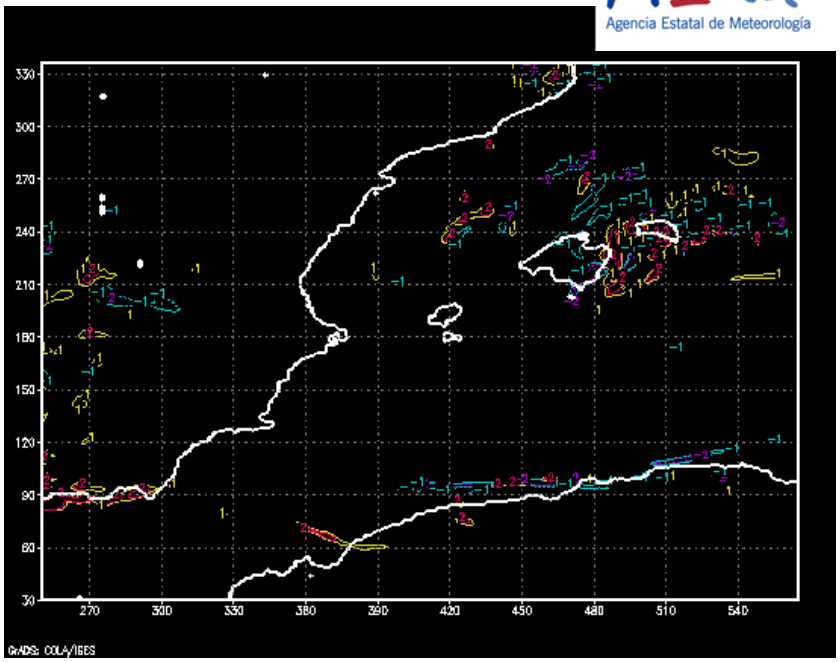


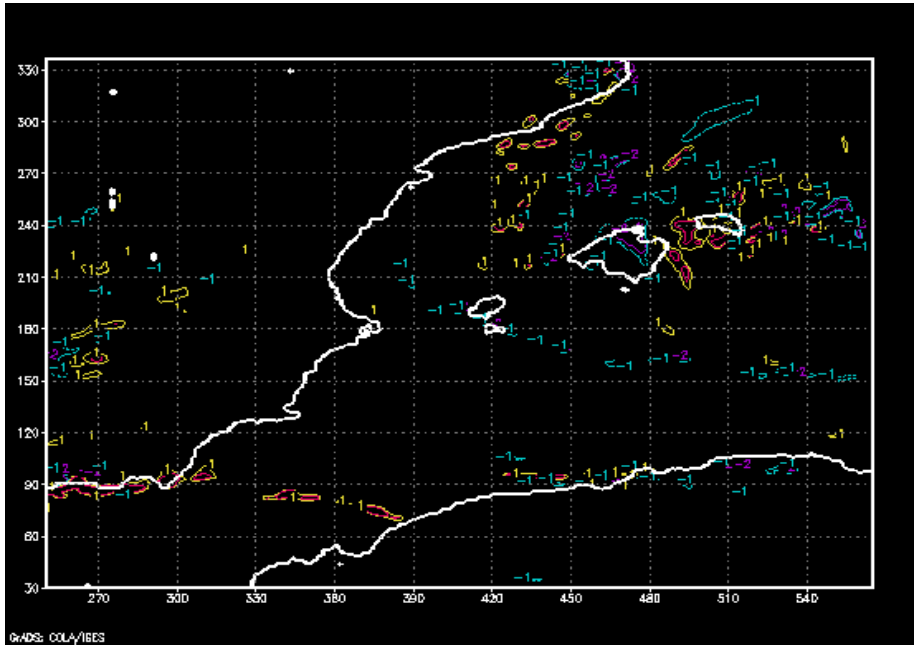


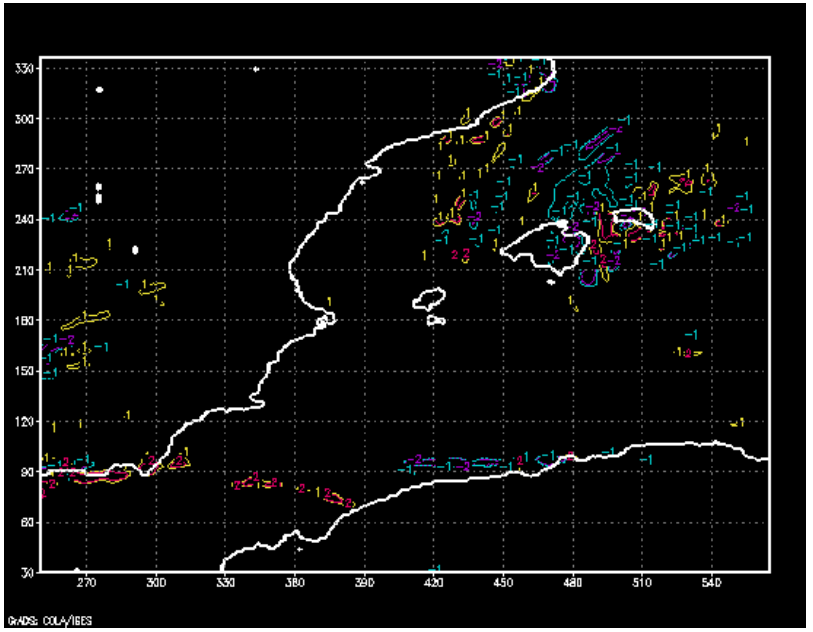
Shifted SIM Z (@30m) Error (cntl - exp)

Aligned









CONCLUSIONS

- A new method for correction of position errors in weather analyses has been tested with the HARMONIE NWP system. The test however was done in an ideal setting and only for one case. This case was not one of hazardous or extreme weather.

- The results are good. The impact on the weather forecasts for wind and precipitation is clear. No "shocks" or rejection problems were found in the only case considered in this work. The multivariate aspect of the issue (having patterns for some parameters but not for others) was however not addressed here.

- There are, of course, many pieces still missing before an eventual explotation of this, or equivalent, method can be a reality.

- The first one clearly is the lack of observations required. The extension of this method to indirect measurements (radar, satellite) conceptually presents no problem

- There are too quite a number of "AI" related issues like "detection" and "correspondence" that would have to be sorted out.