Study of potential spread from LBCs in the context of gamma-mesoscale EPS

Phase 2: Notes about experiments using 5 global models from TIGGE

Juanma Sancho

Phase 2, Part 1

A brief description and main results of a spread-skill study using available fields in MARS ECMWF archive for several variables and 5 TIGGE global models is shown in this document. Different experiments have been carried out for the period June to December 2011, for 00 and 12 runs, and 36h forecast length. The variables used were mslp, t2m, sfcWind, (500,100)gh and (300,700)Wind. The domain of the ensemble experiments covered the entire Iberian Peninsula region [Hirlam HNR] (Fig. 1) with 0.1°x0.1° spatial resolution. In all cases, ECMWF analysis has been used as reference for Root Mean Square Error (RMSE) calculations.

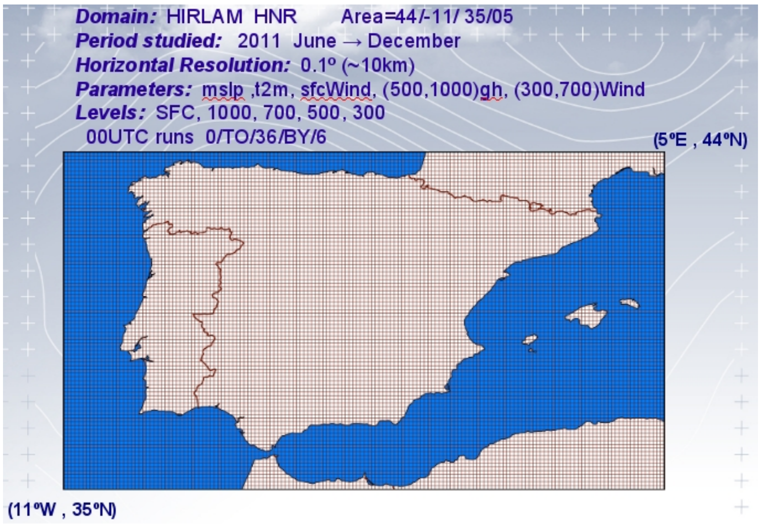

Figure 1: Domain of the spread-skill experiments

During Phase 1 of this study, a comparison of different spread-skill relationships was performed designing ensembles using the 51 ECMWF members and also a random selection of 10 of them. In addition, ensembles constructed with a selection of 10 random perturbed members for each of 5 selected TIGGE global models archived in MARS were analyzed. The selected global models were: ECMWF, NCEP, CMA, CMC, UKMO (Table 1). In summary, the main result of these experiments indicates the better performance of ECMWF EPS system when compared with any mono-model ensemble using the rest of the selected global models, being all of them underdispersive. Nevertheless, one ensemble formed using only the 5 control members for each of the models (including no perturbed members) shows a clearly better performance concerning the ensemble reliability than any mono-model approach (even better than ECMWF ensemble). A slightly overdispersive behaviour is remarkable.

TIGGE, the THORPEX Interactive Grand Global Ensemble

Different characteristics and methodologies of the TIGGE models used and not used in the study:

Centre	Initial Pert. method (area)	Model error Simul	Horizon. Res.	Vert res	Fcst length (days)	# pert mem	#runs per day (UTC)	# mem per Day	Operational & in TIGGE since
Models used for this study:									
ECMWF [ec]	SV(NH, SH, TC) + EDA	YES	TL639 (0.28°x0.28°)	62	0-10	50	2(00/12)	102	11 Mar 08
			TL319 (0.56°x0.56°)	62	10-15 (32)				
NCEP(USA) [kwbc]	ETR(globe)	YES	T190 (0.90°x0.90°)	28	16	20	4(00/06/12/18)	84	27 Mar 07
CMA (China) [babj]	BV(globe)	NO	T213 (0.84°x0.84°)	31	16	14	2(00/12)	30	15 May 07
CMC (Canada) [cwao]	EnKF(globe)	YES	600x300 (0.6°x0.6°)	40	16	20	2(00/12)	42	3 Oct 07
UKMO(UK) [egrr]	ETKF (globe)	YES	(0.55°x0.83°)	70	15	23	2(00/12)	48	1 Oct 06
Models NOT used for this study:									
BOM(Australia)	SV(NH,SH)	NO	TL119 (1.5°x1.5°)	19	10	32	2(00/12)	66	Sep07 to 2010
CPTEC(Brazil)	EOF(40S: 30N)	NO	T126 (1.0°x1.0°)	28	15	14	2(00/12)	30	1 Feb 08
JMA (Japan)	SV(NH, TR, SH)	YES	TL319 (0.56°x0.56°)	60	9	50	1(12)	51	21 Nov 07
KMA(Korea)	BV(NH)	NO	T213 (0.84°x0.84°)	40	10	16	2(00/12)	34	28 Dec 07
Meteo France	SV(local)+ EDA	YES	TL358 (stretched 2.4)	65	3 (4.5)	34	2(06/18)	70	28 Jan 08

Table 1: TIGGE global models used and not used in the study

The encouraging results of Phase 1 concerning the good performance of multi-model experiments, seems to recommend exploring additional different ensembles combinations using the 5 TIGGE global models insisting in the multi-model approach. New ensembles experiments were designed in order to compare spread and skill relationships using only analysis (or control members if analysis not available), analysis members adding some perturbed members and LAF (Lagged Average Forecasting) and SLAF (Scaled Lagged Average Forecasting) approaches which include previous forecasts as well as analysis members. See Table 2 for a description of these proposed experiments.

Experiment	Ensemble Description	Members
Tigge_5	***5 control members (T+00), one for each global model***	***5 control***
Tigge_15	***5 control members (T+00) and 2 perturbed members (T+00) for each model***	***5 control + 2x5 perturbed =15***
TiggeSlaf_10	***5 control members (T+00) and 5 control forecast (T+12) forming a Lagged Average Forecasting (LAF)***	***5 control + 5 fc(T+12) =10***
TiggeSlaf_15	***5 control members (T+00) and 10 control forecast (T+12, T+24) forming a LAF***	***5 control + 5 fc(T+12) + 5 fc(T+24) =15***
TiggeFc_5	***5 control forecast (T+06) LAF***	***5 fc(T+06)***
TiggeFc_10	***10 control forecast (T+06, T+18) LAF***	***5 fc(T+06) + 5 fc(T+18) =10***

Table 2: Additional multimodel ensemble experiments.

Experiment TIGGE_5 consists on a 5 member ensemble formed using only control members (low resolution analysis at forecast length 0h) of the 5 global models and runs 0 and 12 UTC. It would be desirable to use high resolution analysis fields every 6h in this case but only control members every 12h are available in MARS at the moment. Experiment TIGGE_15 is formed adding 2 perturbed members from each of the global models to the previous ensemble, forming a 15 member ensemble. The experiment TIGGESLAF_10 is the result of adding 5 forecasts from the previous 12h control members runs to the TIGGE_5 obtaining a 10 member LAF ensemble as a result. It is important to remark that this experiment is not a real SLAF ensemble in the sense that no age-dependent scale factors are applied which could take into account the degradation of the members with the forecast length. Experiment TIGGESLAF_15 simply adds h+24 forecasting members to the previous experiment, but it is neither a proper SLAF ensemble and so does not apply any decreasing factors to the forecasting members in order to avoid the ensemble to be tainted by older forecasts.
The extent of this study with additional SLAFS experiments constructing ensembles by adding members symmetrically around the control members and properly scaled depending on the forecasting age is addressed in Part 2 of this phase. Finally, experiments TIGGE_FC_5 and TIGGE_FC_10 are formed using h+6h control members forecasts in the case of TIGGE_FC_5 and adding h+18 forecasts for TIGGE_FC_10 ensemble.

RESULTS (Part 1):

Spread-skill relationships for 500hPa geopotential and for each of the experiments described are shown in Figure 2. In this figure each point in different color, depending of the experiment, is the result of the calculation of the Standard Deviation, as a measure of the spread, and the Root Mean Square Error for the ensemble mean with respect to the ECMWF analysis used as reference for 00 and 12UTC runs. The estimated values used for the computation of each point in the graph is approximately 16.000.000 values (14651grid points. x 182days x 5models + ECMWF analysis (14651x182)). The black line indicate the spread-skill evolution for 36h forecasting length and 6h timestep, in the case of a 5 control member multi-model ensemble formed using the run 0UTC for each global model. The first and second point of this black line obviously coincide with points T=0h and T=6h of TIGGE_5 and TIGGEFC_5 experiments respectively.
The 0 and 12h observed differences in RMSE are due to systematic error or bias of global model members with respect to ECMWF analysis values. For instance, in the case of 2 meter temperature, most of the global models (except NCEP and UKMO) show at 12UTC systematically lower temperature values than ECMWF analysis (in yellow), whereas the distribution is approximately symmetric around ECMWF analysis at 0UTC (see Figure 3a). Nevertheless, spread differences are not important between 0UTC and 12UTC (Figure 3b). These differences have been observed also in other parameters (not shown).
In the case of 500hPa geopotential, the experiment showing the lowest spread and error is TIGGE_5 formed with 5 global models control members (T+00). An important increase of the spread without any important associated increase of error is observed in experiment TIGGE_15, which adds 2 perturbed members for each global model.
LAF experiments show somehow higher errors and intermediate spread values. (Results are much more encouraging in the second part of this phase when LAFs experiments are converted to real SLAFS, resulting in a remarkable error reduction).
Experiments including only forecasting members (TIGGEFC_5 and TIGGEFC_10) are the ones showing the highest spread.

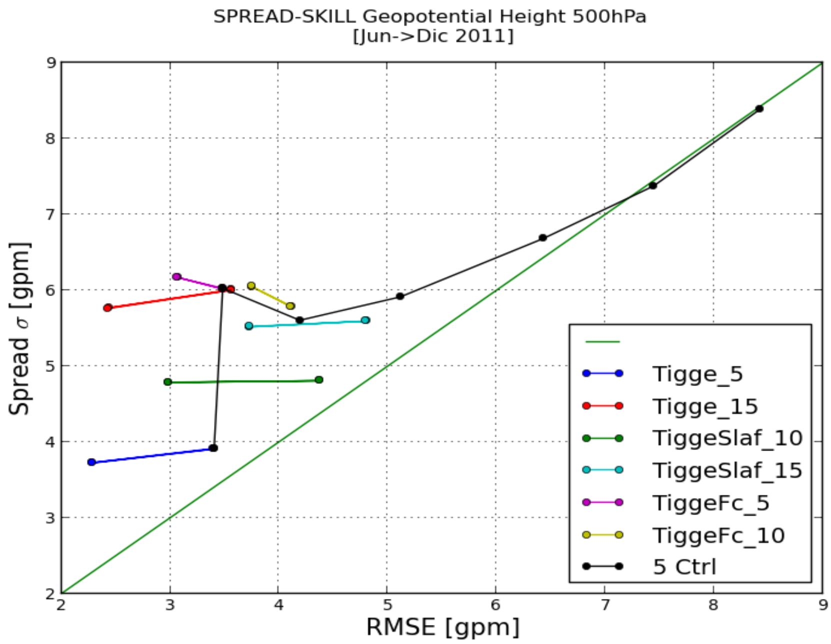

Figure 2: Spread-Skill relationships for Phase2 (Part1) experiments

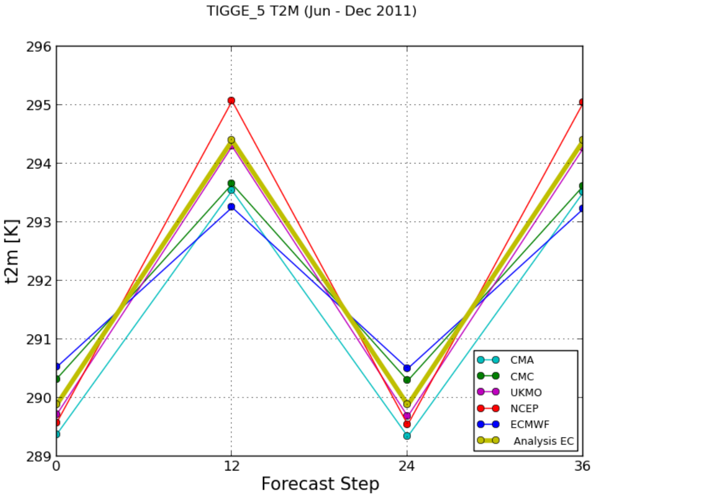

Figure 3a: T2m Global models vs. ECMWF analysis

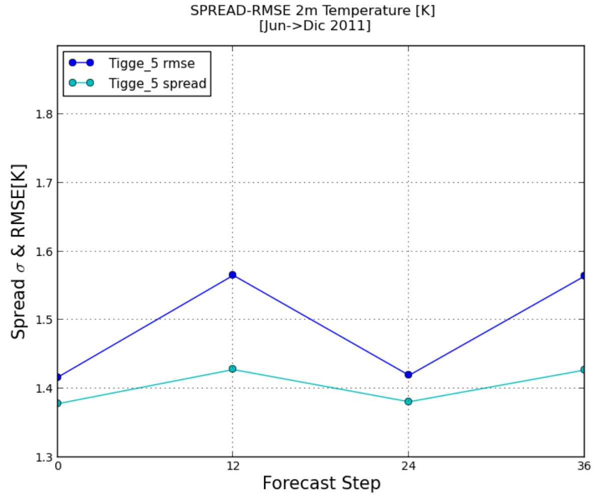

Figure 3b: TIGGE_5 t2m Spread-Skill

Phase 2, Part 2

As it was described before, experiments TIGGESLAF_10 y TIGGESLAF_15 are the result of adding to TIGGE_5 new extra control forecasting members ([T+12] y [T+12,T+24] respectively) forming a LAF (Lagged Average Forecasting). Therefore, they are not real SLAF (Scaled Lagged Average Forecasting) because no age-dependent scaling factors are applied to the forecasting members to account for the degradation of the members with forecast length.
In order to explore more deeply the SLAFs approach, and possible benefits in scaling properly members coming from forecasts, a new set of 4 experiments named TIGGESLAF_10R, TIGGESLAF_15R, TIGGESLAF10R2 y TIGGESLAF15R2 are designed. Table 3 shows a description of each of these experiments.

LAFs simple Lagged Average Forecasting	SLAFs: Scaled perturbations are added to and substracted from control	K_{12},K_{24} Age-dependent scale factors	Ensemble Description 5 ctrl[00] for each model + SLAFs	Number of members
TiggeSlaf_10 (uses only T+12)	TiggeSlaf_10R	K_{12}=0.75	5ctrl[00] + ctrl[00] ± K_{12}(ctrl[00]-fc(12))	*5 control* + *2x5 fc(T+12)* *=15*
	TiggeSlaf_10R2	K_{12}=1		
TiggeSlaf_15 (uses T+12&T+24)	TiggeSlaf_15R	K_{12}=0.75 K_{24}=0.25	5ctrl[00] + ctrl[00] ± K_{12}(ctrl[00]-fc(12)) + ctrl[00] ± K_{24}(ctrl[00]-fc(24))	*5 control* + *2x5 fc(T+12)* + *2x5 fc(T+24)* *=25*
	TiggeSlaf_15R2	K_{12}=1 K_{24}=0.5		

Table 3: SLAF Experiments Phase 2 (Part 2)

The main idea is perturbing symmetrically the control values for each model adding and subtracting perturbations from previous forecasts, but applying scaling factors depending on the age of the forecast in order to avoid the ensemble to be tainted by older forecasts. For this purpose, 2 sets of scale factors have been applied which are shown in Table 3. The values K_{12}=0.75 and K_{24}=0.25 used in experiments 10R and 15R, are the usual factors more frequently applied in these kind of ensembles whereas the higher values K_{12}=1 and K_{24}=0.5 are the factors used in SAMEX ' 98 experiment (Hou et. al, 2000). Another consequence of these SLAFs experiments is the increase of the number of ensemble members, obtaining 15 members in the case of TIGGESLAF10R and 25 members in TIGGESLAF15R experiment.

RESULTS (Part 2):

There are some remarkable effects in spread-skill relationships as a consequence of the change from LAF to SLAFs experiments. For all the studied variables (Z500, mslp, t2m, viento[sfc,700]) the spread (Standard Deviation) increases while the error (RMSE) drops for SLAFs ensembles formed using T+12 forecasts along with control members T+0 (TIGGESLAF10R y TIGGESLAF10R2 experiments).

Obviously, this increase in spread is dramatically higher when the greater scale factors are used (TIGGESLAF10R2 experiment), but the error values are not affected at all, since the members are symmetrically distributed around the control values.

These positive changes in spread and error are much more important for surface and 700hPa winds and 2 meter temperature. An overdispersive behaviour in spread-skill diagrams and the highest spreads are observed for these experiments (Figures 4, 5 and 6).

In the case of 500hPa and mslp variables (Figures 7 and 8) higher spread values and lower error are also observed although the changes are not so important for these parameters and spread is still lower that the value for TIGGE_15 experiment (containing perturbed members) with similar error.

Experiments TIGGEFC_5 y TIGGEFC_10 (with forecast members) provide high spread values but high errors as well.

Concerning to the SLAFs experiments including T+24 forecasts, a decrease in error is also evident compared with LAFs experiments but spread values drop despite of the increase in the number of members, possibly due to the low values of factor K_{24} applied to the oldest forecasts.

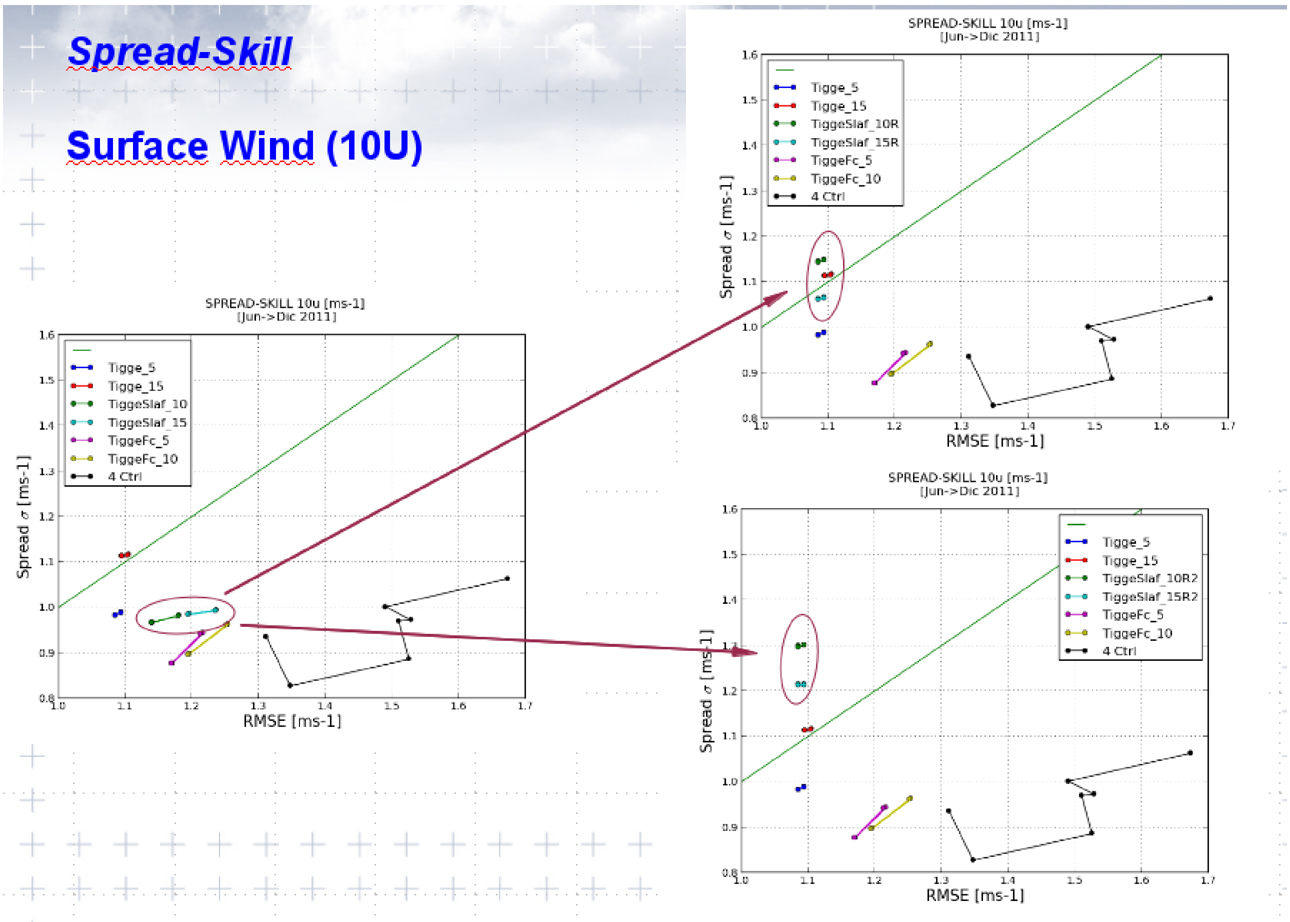

Figure 4

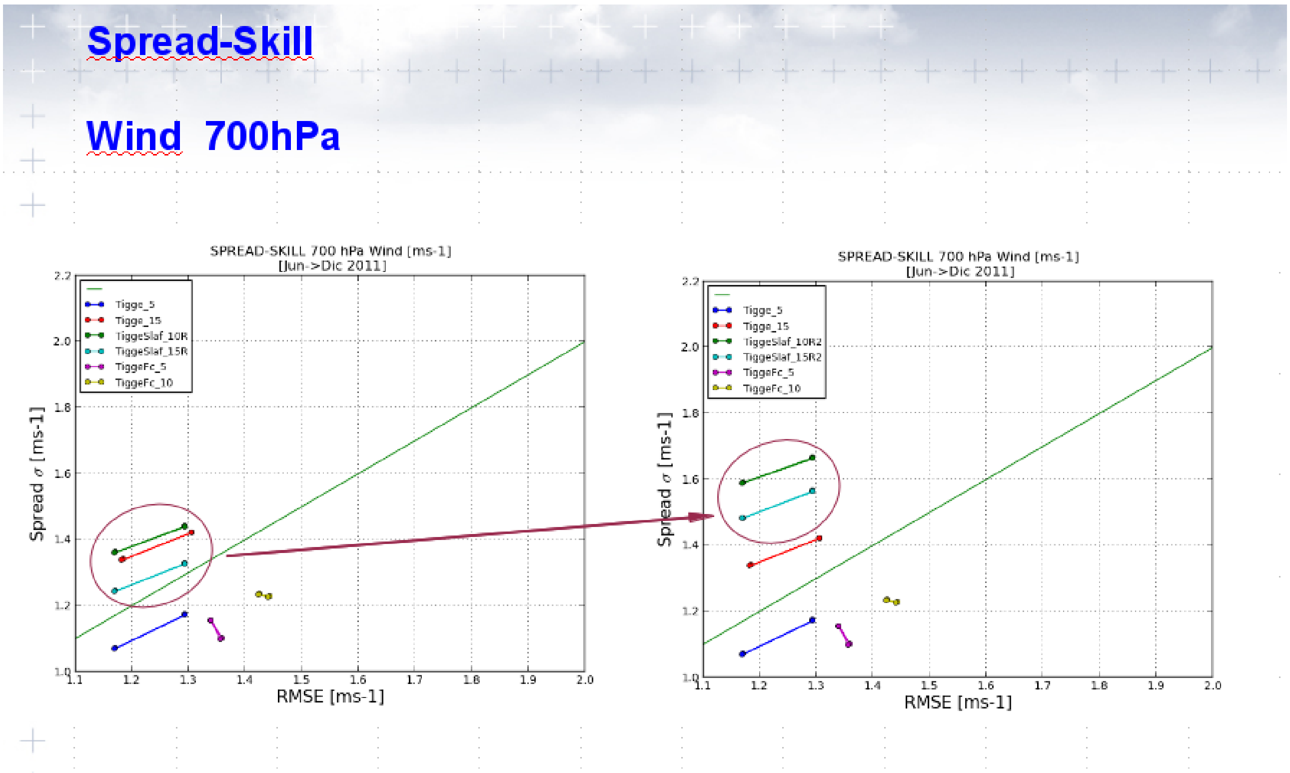

Figure 5

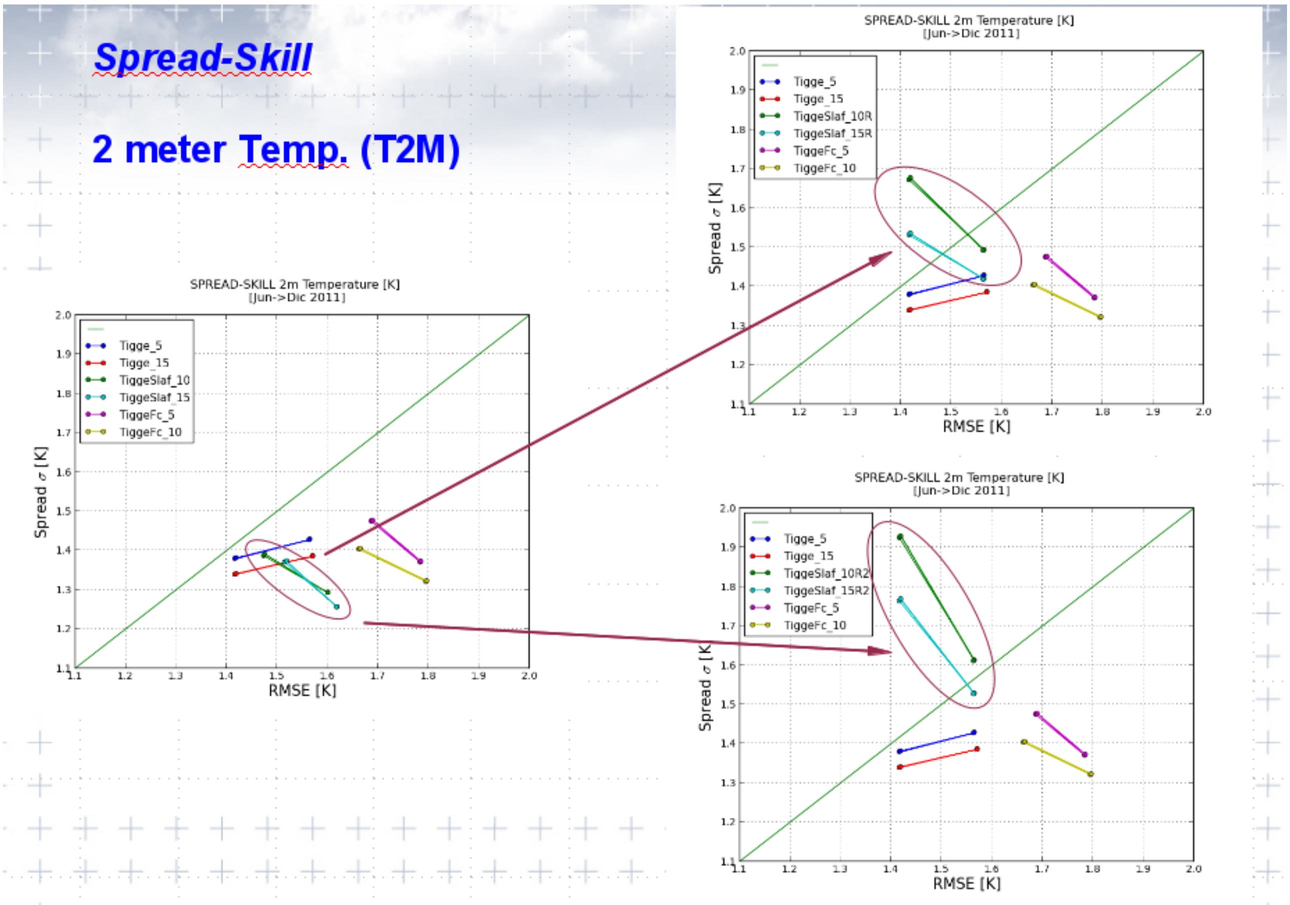

Figure 6

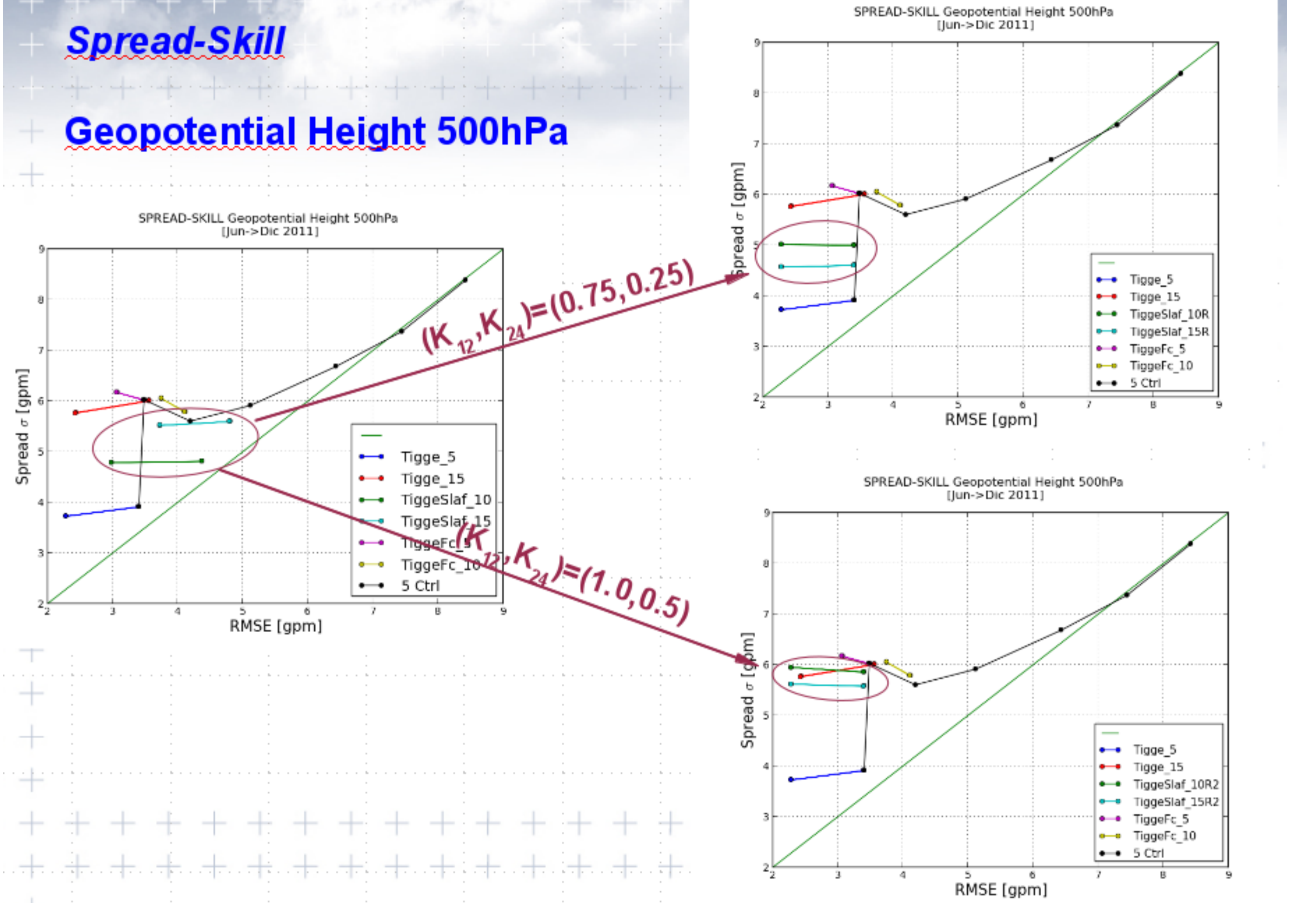

Figure 7

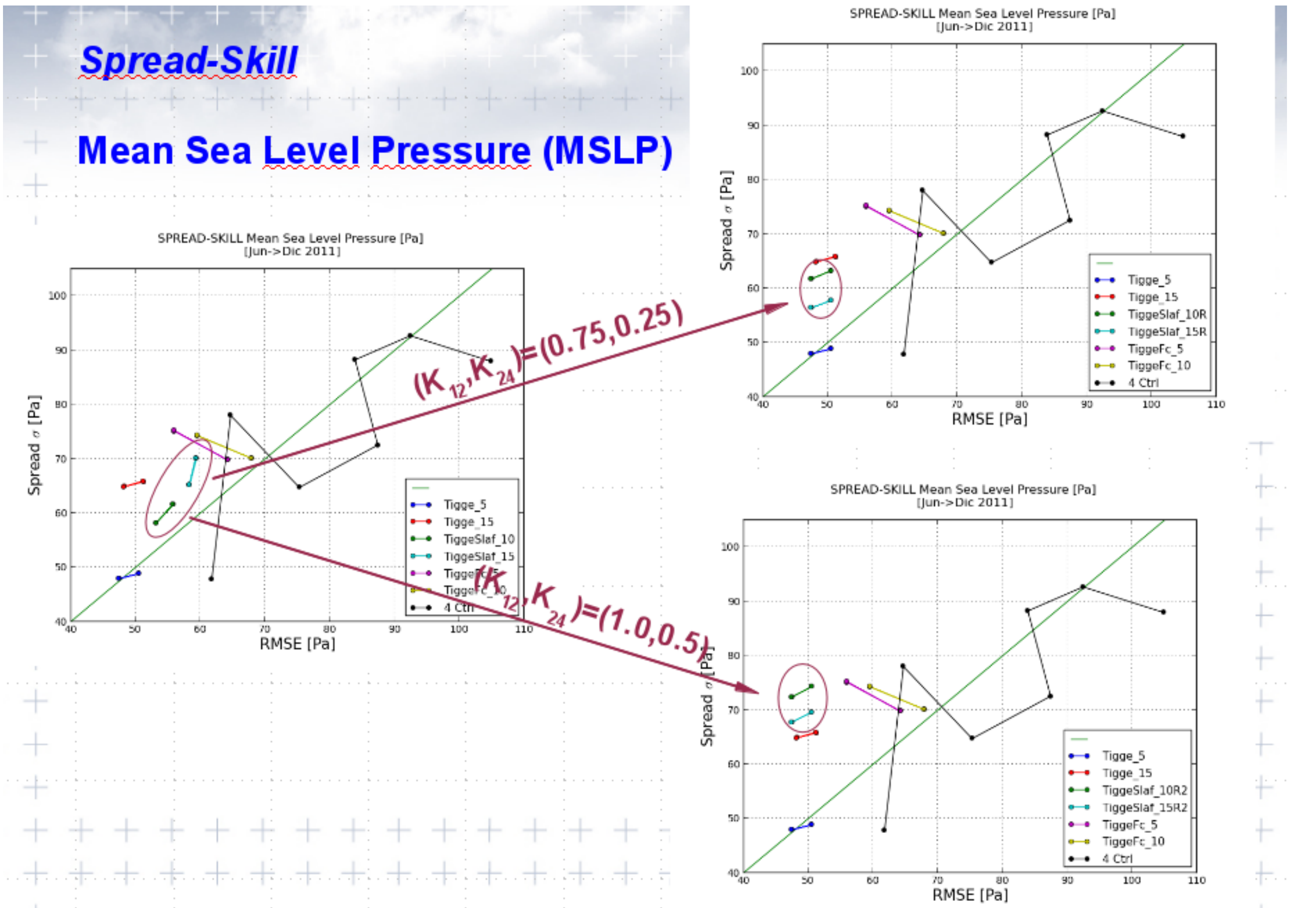

Figure 8

CONCLUSIONS:

The first phase of this study concluded that ECEPS performs better than the rest of mono-model ensembles using CMA, CMC, UM and GFS TIGGE global model separately and a subdispersive spread-skill behaviour was observed in all cases. Furthermore, a multimodel approach forming an ensemble using only the control members for each of the 5 global models verify better than ECEPS and it is slightly overdispersive.

During the second phase, the multimodel approach has been explored more deeply and new experiments have been analyzed combining control members for T+0, T+12 and T+24 forecast lengths with perturbed members (T+0), obtaining the more encouraging results for the proposed SLAFs (Scaled Lagged Average Forecasting) experiments providing high spreads (Standard Deviation) and low errors (RMSE)

The only experiment using perturbed members along with control members (TIGGE_15) provides promising results as well resulting in a good spread-skill relationship.

Obviously, the age-dependent scale factors K_{12} y K_{24} affect importantly to the spread (but not to the error) and special care should be taken in the chosen values in order not to damage the ensembles.

It could be interesting to extent the study designing additional SLAFs experiments, centering, for instance, the ensembles in ECMWF control members instead than in the control members for each of the global models. This would allow us to have model levels available. Furthermore it might be interesting to use some perturbed members from TIGGE models in SLAFs experiments given the reasonable results shown by TIGGE_15 experiment.