LACE report

Singular vectors in ALADIN model

Richard Mládek, mladek@chmi.cz

14.5. - 13.6. 2007
Stay in Hungarian Meteorological Service, Budapest
Supervised by András Horányi & Edit Hágel

1. Introduction

The work was done according to Draft working plan for singular vector computations with ALADIN model prepared by Edit Hagel and Andras Horany in February 2007 (see below annexed to this document) . Task1 of the working plan - Testing and informatic validation of ALADIN configuration 601 in Reading etc. was completed during the stay and also the tasks 2 and 3 were partially tackled (comparison of limited area singular vectors on different resolutions and some sensitivity tests of ALADIN singular vectors to the optimization time). Two test cases for preliminary validation were chosen in the study June 28, 2006 [Hagel, 2006] and March 5, 2006 [Barkmeijer, 2007]. Singular vectors are computed for 00 UTC and two target times +12 and +24 hour. Two domains are tested GLAMEPS and "LACE like" both with two horizontal resolutions 22 and 44 km. The LACE domain and all vertical levels are used for targetting, coupling frequency is 3 hour.

2. Technical testing

The first aim was to verify ALADIN configuration 601 for singular vector computation on HPCE supercomputer at ECMWF. The results were compared to the same case (June 28, 2006, LACE domain and 20km resolution) computed on TORA at Météo-France. There was found no problem when running 601 on HPCE so the work could continue with computer cost testing with various settings.

Before starting such experimentation decisions about possible initial setup had to be made. GLAMEPS working plans and already used configurations (table 1) of global ensemble systems (PEARP, ECMWF) have to be considered. Because of planned forecast range for GLAMEPS up to +60 hour only two optimization times has been finally tested: +12 and +24 hour. Further LACE area for both GLAMEPS and LACE domains and all vertical levels are used for targeting, coupling frequency is 3 hour. The most important results of the computer cost testing are summarized in table 3.

	number of SV	***resolution***	***optimization time***	***vertical optimization***	***optimization area***
PEARP	16 (64 iter.)	T95L41	12h	all levels	SW=[30, 280] NE=[65, 40]
ECMWF	~50 (>150 iter.?)	T42L62	48h	all levels	*

Table 1 Global ensemble setup for PEARP and ECMWF.

** optimization area fo ECMWF:*
1) Extra-tropics: 50 SVs for N.-Hem. (30 90 N) + 50 for S.-Hem.(30 90 S). Tangent-linear model with vert. diffusion and surf. friction only.
2) Tropical cyclones: 5 singular vectors per region targeted on active tropical depressions/cyclones. Up to 6 such regions. Tangent-linear model with representation of diabatic processes (large-scale condensation, convection, radiation, gravity-wave drag, vert. diff. and surface friction).

The original idea was to compute high resolution singular vectors (~22 km) on the whole GLAMEPS area. Unfortunately it proved to be unrealistic because of extra high computational costs even with 44 km horizontal resolution (see exp. F002, F003 in table 3). That's why we had to define smaller LACE-like area for our sensitivity tests called glac in the following text. Thus two smaller domains with different resolutions 22 and 44 km were defined. Domain glac22 resp. glac44 is exact cut from original GLAMEPS domain glam22 resp. glam44. Reference latitude and longitude of Lambert projection are the same for all defined domains. For details see table 2 and figure 1. Defined glac and glam domains with different resolutions are geographically not exactly the same as they could be because the domains with lower resolution have indeed half of gridpoints at both directions but including ALADIN extension zone too. For our purposes it shouldn't be a problem.

Around 30 integrations were finished for sensitivity tests. The various combinations of number of nodes, tasks per node and memory per task were used to find the most effective HPCE setup for ALADIN singular vectors computation with given resolution, domain and optimization time. We remind that there are available 144 nodes and 32 tasks with 25 GB of total memory per node on HPCE supercomputer (for more technical info see http://www.ecmwf.int). However the number of asked nodes mustn't be too high because of the high expenses otherwise expressed in SBU (HPCE billing units). Normally 1-2 nodes should be sufficient for our type of tasks. It was not fully understood how SBU are exactly computed for the jobs using more nodes. For single node integrations it seems that formula used is:

SBU = 717000 / (2288*86400) * "CPU time in seconds"

If you use more nodes you pay much more as can be seen e.g. from experiments I101, I002 or when comparing two various setups of H001 (table 3). But there has to be considered also other factors than number of nodes and CPU time regarding SBU job price as can be seen from comparison of experiments I101 and I001 which have the same model setting and both use two nodes but with different numbers of tasks per node and defined memory per task. Maybe bigger efficiency of memory usage for job I001 (60%) is the reason why this job costs only 779 SBU instead of 1145 SBU as for I102 with only 35 % of ratio between maximum used and allocated memory.

	NMSMAX	*NSMAX*	*NDGL*	*NDLON*	*ELAT1*	*ELAT2*	*ELON1*	*ELON2*	*RESOL [km]*
glam22	106	99	300	320	16.1	75.1	-20.1	87.3	22
glam44	53	49	150	160	17.5	75.2	-19.8	81.4	44
glac22	47	42	128	144	33.3	56.3	3.4	40.6	22
glac44	23	21	64	72	34.6	55.4	4.4	38.0	44

Table 2 Domain definition

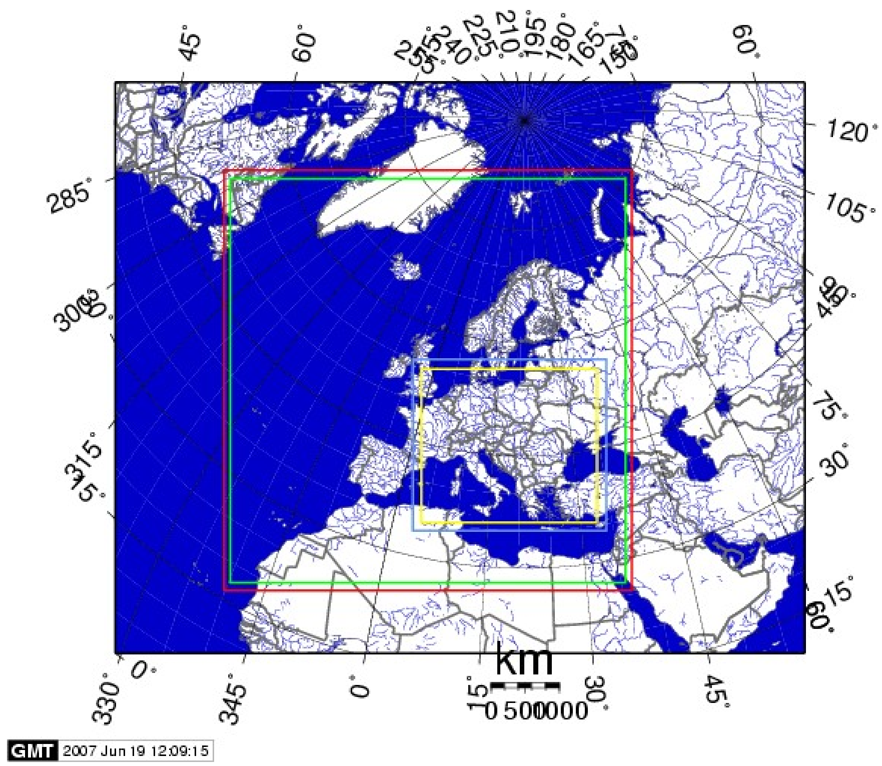

Figure 1 Defined domains: glam22 - red, glam44 - green, glac22 - blue, glac44 - yellow

The optimal found HPCE setup for each defined singular vector computation on glac domain has gray background in table 3. So one can see that for 44 km resolution and 12 resp. 24 hour optimization time the elapsed time and billing units are 1826 sec, 73 SBU resp. 3615 sec, 159 SBU (exp. H001 resp. H102). With 22 km horizontal resolution the jobs are already very expensive even on glac domain with elapsed time and billing units 6716 sec, 779 SBU resp. 22030 sec, 780 SBU for 12 resp. 24 hour optimization time and experiments I001 resp. I102. The reasonable SBU cost for experiment I102 is achieved due to one used node only. However total elapsed time is then very long. The another finding is that one cannot use more than cca 10 tasks per node using only one node and glac22 domain to speed up the integration (and pay reasonable SBU price) because the job will be killed due to memory lack (12 and 16 tasks per one used node were unsuccessfully tested with maximum allocable memory available on one node). So e.g. HPCE setting used for I102 seems to be the best possible. There is also a limitation of maximal used number of tasks because of the model code parallelization. If you ask for more than 24 tasks you need to have switched on b-level parallelization [Radnoti et al., 2006] in the namelist but it didn't work for 601 configuration. It might be only a namelist problem (some missing additional switches related to given cycle) but the real code problems in ALADIN are of course possible too. The only changes which have been applied in the namelist related to the number of asked cpus/total tasks on all nodes are:

1. for cpus < 24
   NBPROC=cpus ; NBPROC2=$NBPROC ; NBPROC3=1
2. for cpus > 24
   NBPROC2=xxx; ((NBPROC= $NBPROC2*$NBPROC3)); NBPROC3=$NBPROC2
   tested xxx = 4, 8, 12 (6 nodes/ 24 tasks per node), 16 (8 nodes/ 32 tasks per node)
   what is equivalent to 16, 64, 144 and 256 cpus together.

```
&NAMPAR0
 NOUTPUT=1,
 NPROC=NBPROC,
 NPRTRW=NBPROC2,
 NPRTRV=NBPROC3,
 NPRGPNS=NBPROC2,
 NPRGPEW=NBPROC3,
 MP_TYPE=2,
 MBX_SIZE=50000000,
 LIMP=.FALSE.,
 LIMP_NOOLAP=.FALSE.,
/
&NAMPAR1
 LSPLIT=.FALSE.,
 NSTRIN=NBPROC,
 NSTROUT=NBPROC,
 NCOMBFLEN=1800000,
 LSLONDEM=.FALSE.,
/
```

The time steps found to be stable are 150 resp. 240 sec for 22 resp. 44 km resolution for glac domain. Tested time steps when model blew up were also 180 resp. 360 sec for glac22 resp. glam44 domains. For glam44 domain was used without problem also time step 280 sec. Shorter time step 120 sec (experiment I002) was used when searching for the reason of repeated model crashes (finally the length of time step was not guilty) and the costs are presented here to show the most expensive tested setup (high resolution, short time step, 24 hour opt. time and 3 nodes/8 tasks per node with low efficiency of allocated memory).

Beside technical testing of 601 on HPCE supercomputer run with Eulerian scheme also the first attempts with semi-Lagrangian advection scheme on SX68 supercomputer in Prague were done. Semi-Lagrangian code in tangent-linear and adjoint version of ALADIN model will be first officially introduced in cycle al32t2. Our tested version was based on new development of Filip Vana in cycle al31t1 coded in Prague. After debugging of the code 60 % of CPU time reduction was reached using semi-Lagrangian instead of Eulerian scheme with six times larger time step. The results were similar but surely not the same. Further investigation of the differences between singular vectors computed both ways is needed in near future. Our promising results with distinct cost reduction are valid for vector supercomputer SX68 mainly due to the good vectorization of debugged code and the situation can be different on scalar machines like HPCE.

experiment	***domain***	***resolution [km]***	***optimization time [hour]***	***time step [sec]***	***number of nodes***	***number of task per node***	***memory per task [GB]***	***SBU***	***memory usage [%]***	***elapsed time [sec]***	***CPU time [day+hh:mm:ss]***	***number of singular vectors***	***notice***
G001	glam	22	12	180	2	5	5	**2090**	78	18009	2+01:51:56	12 it.	X
F002	glam	44	12	280	2	5	5	**758**	75	6533	1+11:46:47	25	
F003	glam	44	24	280	2	5	5	**1271**	80	10959	2+12:18:27	30	
H001	glac	44	12	240	2	5	5	**203**	10	1753	0+04:46:40	28	
H001	glac	44	12	240	1	16	0.78	**73**	59	1826	0+05:37:21	28	
H101	glac	44	12	240	1	24	0.6	**75**	73	1522	0+05:44:05	24	
H002	glac	44	24	240	1	16	0.78	**162**	63	3799	0+12:25:23	36	
H102	glac	44	24	240	1	16	0.78	**159**	65	3615	0+12:09:48	33	
I101	glac	22	12	150	2	5	5	**1145**	35	9871	1+03:03:58	28	
I001	glac	22	12	150	2	10	2.5	**779**	60	6716	1+12:31:24	30	
I002	glac	22	24	120	3	8	3.12	**2333**	54	13404	3+16:31:41	35	
I102	glac	22	24	150	1	10	2.5	**780**	79	22030	2+11:44:18	32	

Table 3 Computer cost testing (X mean killed because of CPU time limit)

3. Sensitivity testing

After introductory technical validation the singular vectors of two cases were studied in more details. The notation of experiments with above described resolutions and optimization times are shown in table 4. First nine singular vectors (temperature u- and v-components) were visualized and examined on model levels for each experiment. Only temperature perturbations are discussed below. Wind perturbations are much bigger at final time then at initial one so the different contour intervals would have had to be used. The patterns and vertical structure of perturbations are qualitatively analogical for both temperature and wind perturbations.

	5.3. 2006		***28.6. 2006***	
resol. \ opt.time	**12 h**	**24 h**	**12 h**	**24 h**
22 km	I001	I002	I101	I102
44 km	H001	H002	H101	H102

Table 4 Experiment names

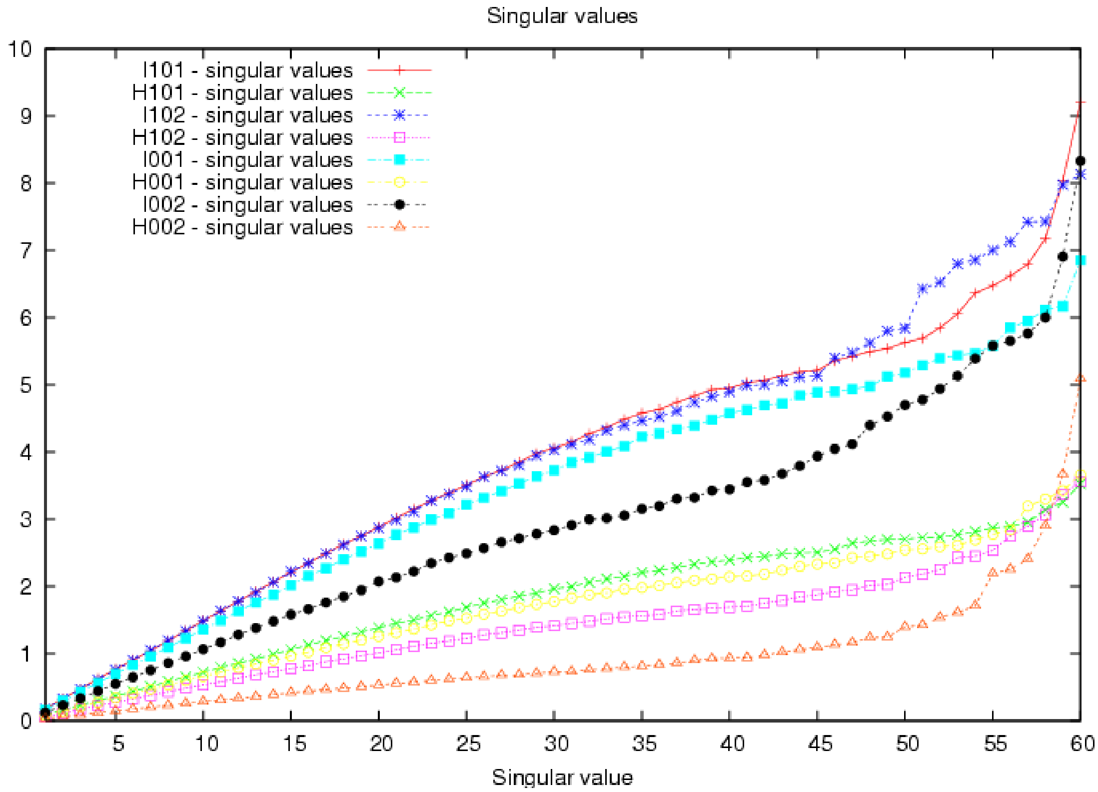

Figure 2: Singular values of first 60 singular vectors (first singular vector is denoted here as 60). Only approximately a halve of the singular vectors with the highest singular values is numerically stabilized after 60 iterations and with given precision.

The singular values computed in experiments with higher horizontal resolutions are around 2-3 bigger then in experiments with 44 km as expected (Figure 2). The biggest decrease of energy is within first 3-4 singular vectors then it seems to be more linear. Last approximately 30 smallest singular values (denoted as 1-30 in figure 2) are not too realistic because of limited number of iterations (corresponding singular vectors are not numerically stabilized).

3.1 First case

Synoptic situation of the first “HIRLAM” case is shown in Figure 3. This situation was chosen mainly to have comparable results from ALADIN and HIRLAM models and not because of some very specific or dangerous weather phenomenon.

The singular vectors structure is pretty complicated even exploring only first nine of them. We found four resp. two different types of singular vectors at initial time, 12 hour optimization time and high (exp. I001) resp. lower resolution (exp. H001). After twelve hour integration one can see the same group of singular vectors evolved over optimization period. While at beginning the biggest perturbation were localized between model levels 21-27 (320-540 hPa) at the end they spread much more - from the surface to the highest model levels. One group of singular vectors which can be found at starting time using 12 hour optimization time for both resolutions show vertical north-west tilt which is conserved at the end of integration too. Other singular vectors have no such vertical tilt. It is interesting that the first singular vector with the biggest singular value of exp. I001 is completely different from the first singular vector of exp. H001 (Fig. 4). Similar patterns one can found in singular vectors 7-9 of exp. I001 (combined with other types of singular vectors) but not among first nine singular vectors of exp. H001.

For optimization time 24 hour the situation is different. For higher horizontal resolution (I002) there are only three groups of similar singular vectors which are even less diverse one by one then in exp. I001. For lower resolution 44 km (exp. H002) all first nine singular vectors are very similar. At optimization time +24 hour one can see very similar character of all nine evolved singular vectors for both experiments I002 and H002. Much stronger gradients of perturbations in first singular vector comparing rest eight in H002 is remarkable and this is not so expressed in I002. Similar distinctive decline of second and other singular values can be seen in Fig. 2 for exp. H002.

Generally we can say that rather similar gradients of perturbations at initial time for both horizontal resolutions lead to much stronger gradients of evolved perturbations at final optimization time in case of higher (22 km) resolution. This is clearly seen especially for 24 hour optimization time when the perturbations are in addition much bigger then for 12 hour optimization time (not shown).

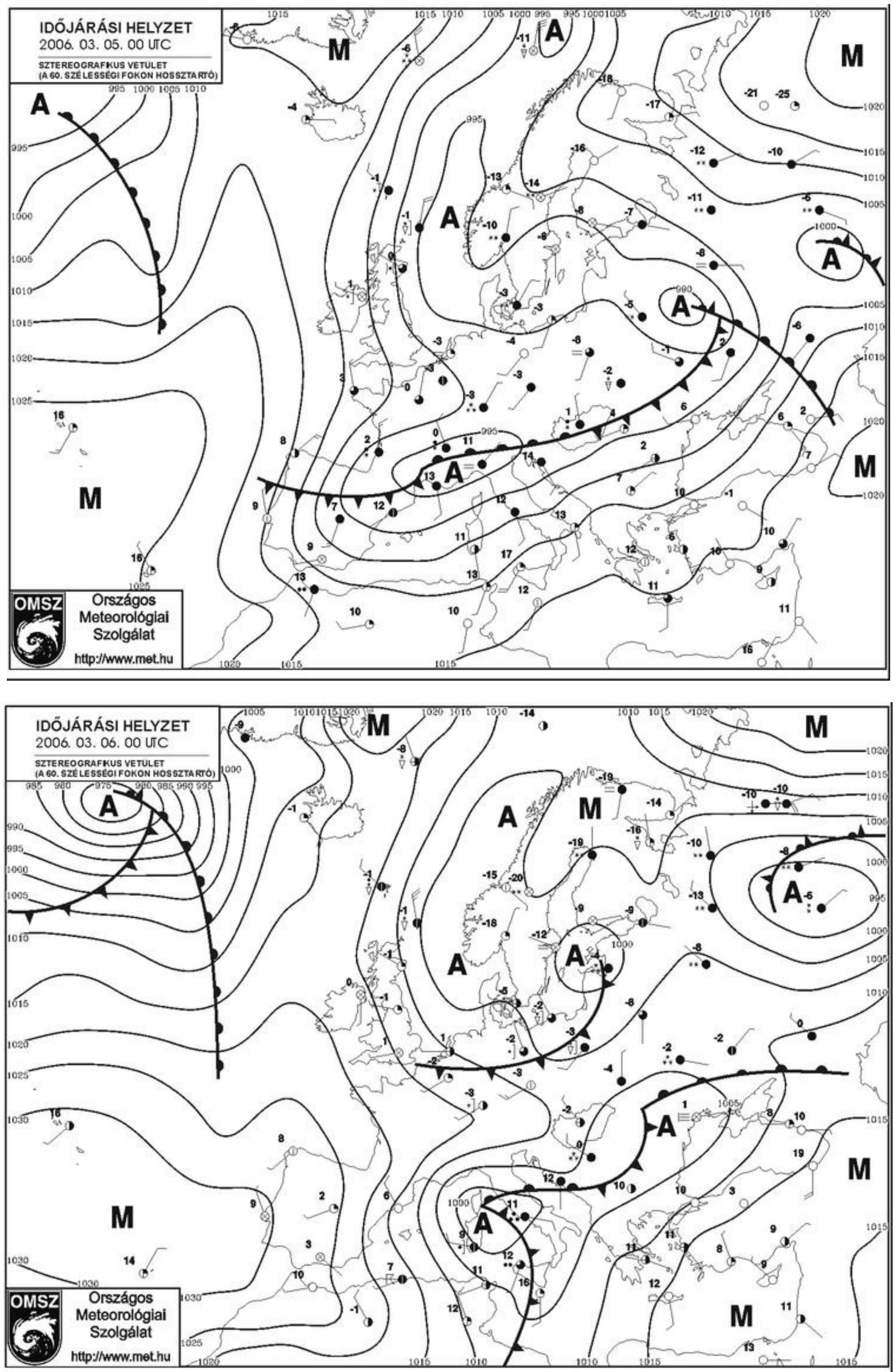

Figure 3: First case from 5th March, 2006. Top synoptic situation at 00 UTC, March 5; bottom at 00 UTC, March 6.

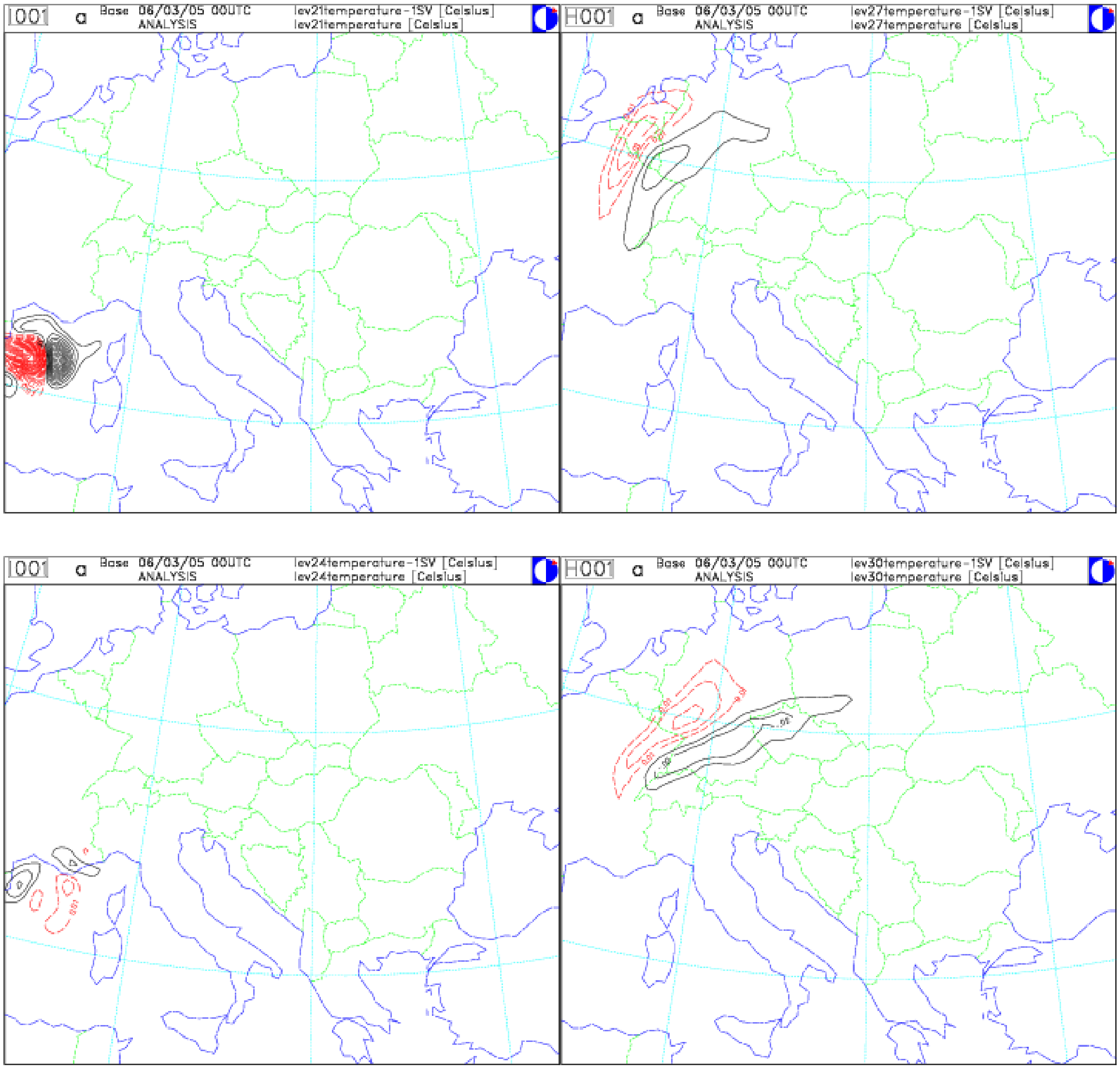

Figure 4: First singular vectors of exp. I001 (left) and H001 (right) for 5th March, 2006, 12 hour optimization time, temperature perturbation only. Contour interval is 0.01 K. Notice please different vertical model levels for both experiments. There are no other perturbations than between model levels 21-24 in exp. I001. Similar perturbations in lower levels can be seen in H001 (with vertical NW tilt when going surface to the top).

3.2 Second case

Synoptic situation of the second case is shown in Fig. 5. The structure of singular vectors is much simpler then in previous case and they are not so spread over big area as in many singular vectors before (not shown). As far as 12 hour optimization time is considered the bigest perturbation are localised in SW of the integration domain (to the west from Sardinia and over Corsica and Sardinia) for both resolutions. The centre of the most perturbed area is somewhat shifted to the east in H101 (44 km) then in I101 (22 km) as can be seen in Fig. 6 (valid for initial time). At final time +12 hour this centre is moved in NE direction (northern Italia and northern Adriatic see) what is shown in Fig. 7. The second evolved centre of perturbations is between Sardinia and Corsica in exp. I101 (4, 6 and 8th singular vector) which is not pesent in H101 (not shown).

At initial time for 24 hour optimization time the similar centre of the biggest perturbation as for 12 hour optimization time still can be found among the first nine singular vectors but there is another one localised over Middle Europe (1, 4 and 6th resp. 3-6 and 9th singular vector in I102 resp. H102). At final time those two centres are moved again in north-east or east direction in accordance with prevailing average flow.

There is no evident vertical tilt in singular vector structure for both optimization times.

3.3 Conclusions

With higher resolution more types of singular vectors can be found which are not present in lower horizontal resolution experiments. Even if the singular vectors are sometimes very similar for both resolutions their geographical or vertical placement is mostly somewhat changed. The gradients of perturbation are much stronger with higher resolution especially at final time. The structure of at least first nine singular vectors depends on optimization time too. Even first singular vectors with the biggest singular value cannot be qualitatively similar in different horizontal resolutions as was shown for first case.

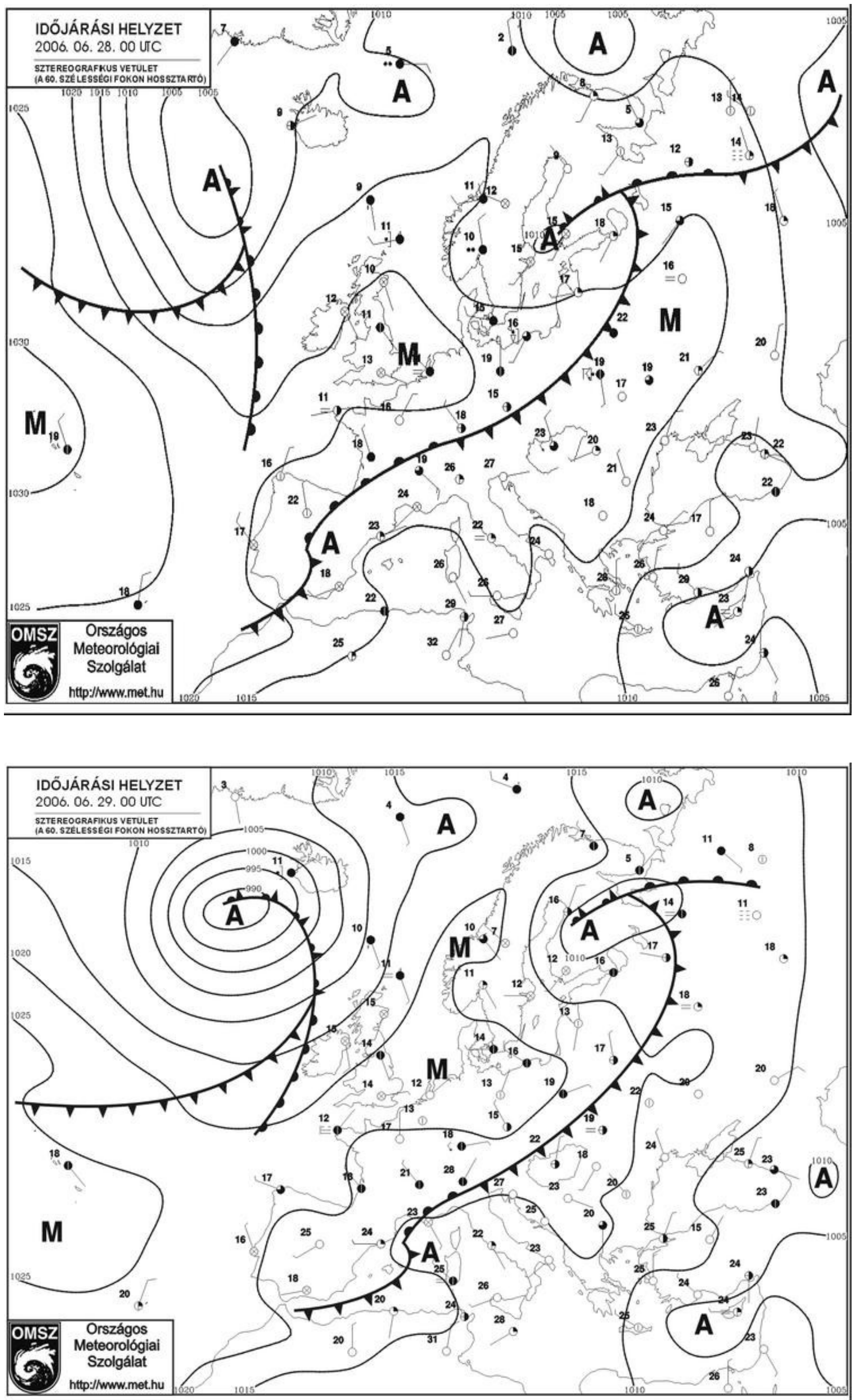

Figure 5: Second case from 28th June, 2006. Top synoptic situation at 00 UTC, June 28; bottom at 00 UTC, June 29.

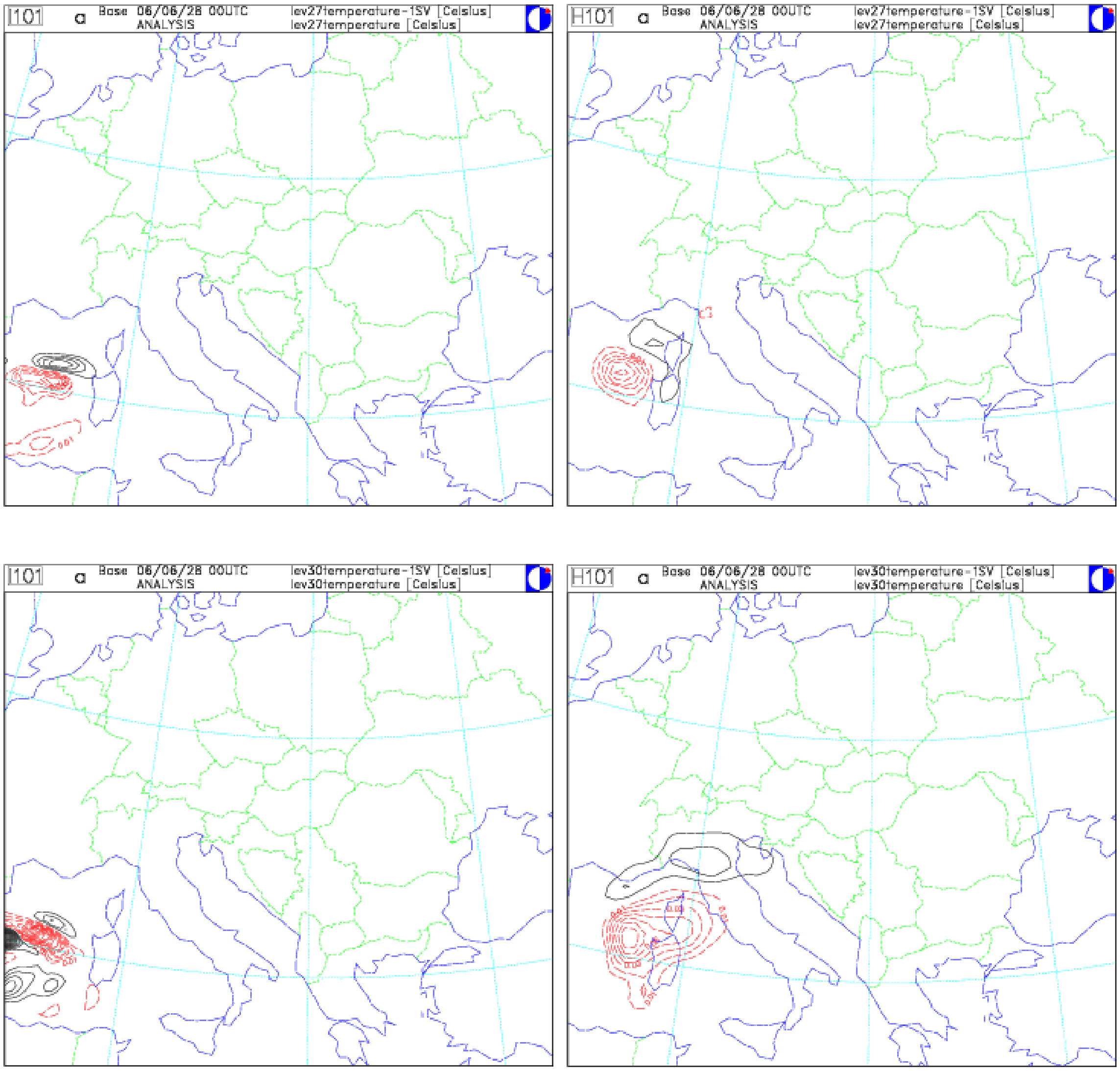

Figure 6: First singular vectors of exp. I101 (left) and H101 (right) at initial time 28th June, 2006, 12 hour optimization time. Temperature perturbation shown only, contour interval 0.01 K.

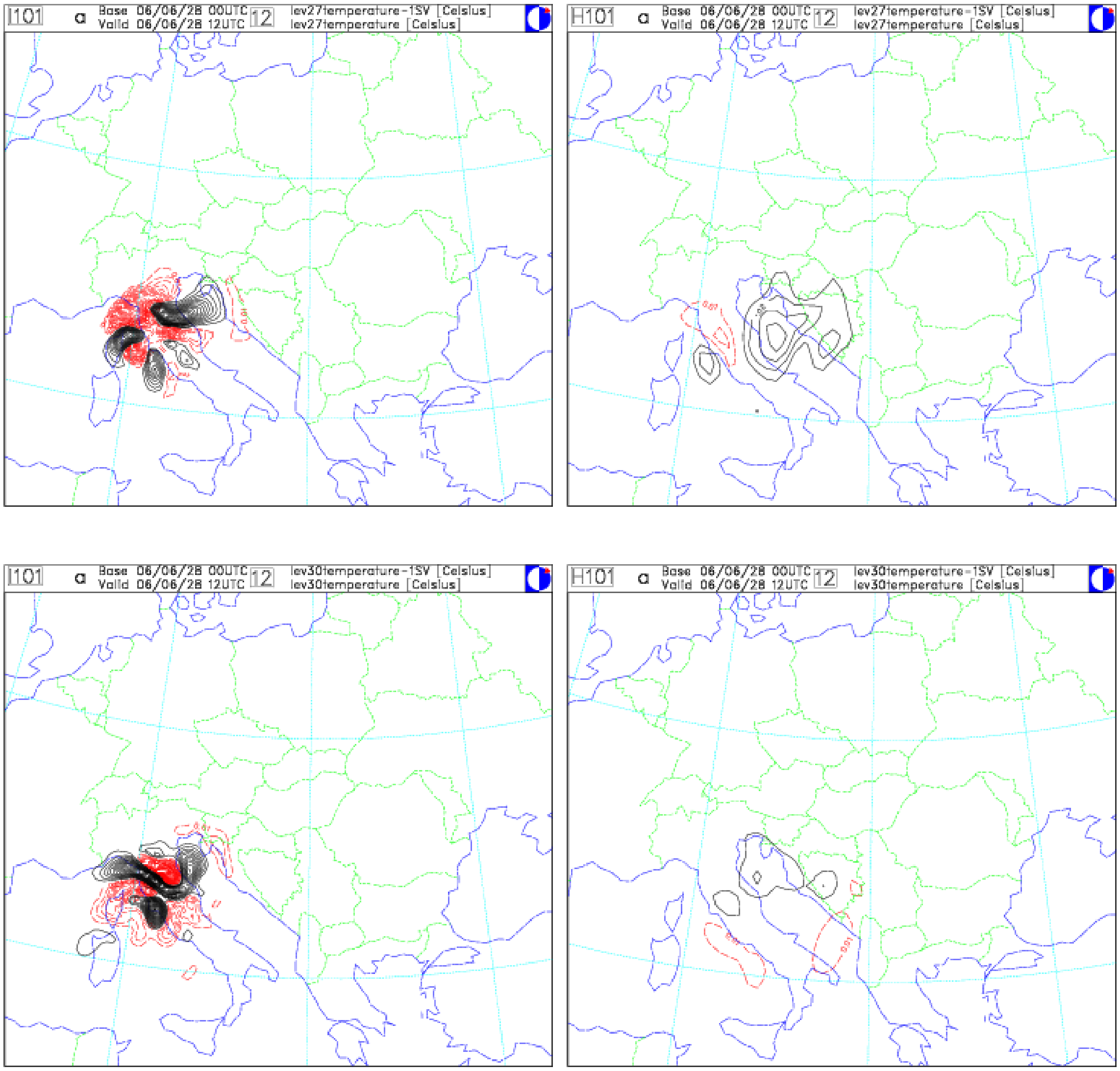

Figure 7: As in Fig. 6. but situation at final time +12 hour.

4. Next work

More investigation of singular vector structures on more cases is needed because of their complexity and dependance on weather regime. The discussion of the connection of the shape and overall character of singular vectors to the synoptic situation should be attempted. It is not easy or even maybe possible to fully understand and objectively evaluate the influence of each individual singular vector perturbation to the final weather change during optimization time. So after more studies of singular vectors in a general manner as suggested above an effort to create the initial perturbations based on local and global singular vectors should follow together with objective verification of the resulting forecast.

5. Data & scripts locations

- **regatta:**

chagal scripts:
/clwork/workdir/mladek/sv/chagal/chagal.sh

- **hpce:**

/home/ms/cz/czr/sv_ald/job/601.job
/home/ms/cz/czr/sv_ald/nam/cy30_601_phys.nam

- **ecfs:**

singular vector data:
DOMAIN=/glam, glac, lace/; RESOLUTION=/22km, 44km/
ec:/czr/sv_ald/testLBC/$DOMAIN/$RESOLUTION

log outputs:
ec:/czr/sv_ald/sv_ald_log.tar.gz
ec:/czr/sv_ald/sv_ald_log.tar.txt .. tar listing

experiments results:
DATE=/20060305r0 , 20060628r0/; EXP=/H001, H002, H101, H102, I.., F,G../

- **mladek@pc2264:**

html generation:
/vol1/home/mladek/sv_ald/html/genhtml.sh .. with t,u,v base maps
/vol1/home/mladek/sv_ald/html/genhtml.sv.sh .. first X singular vectors on one page

- **mladek@pcRogger:**

html results:

H:/mladek/sv_ald/html

6. References

- Bakrmeijer Jan, Stappers Roel: Singular vectors in HIRLAM, Oslo ALADIN-HIRLAM meeting 2007, (https://hirlam.org/HX/organisation/reports/ASM2007/s5_barkmeijer.pdf)

- Fischer Claude, Guidard Vincent, Payet Pierre-Luc: Study of singular vectors in ALADIN.. Aladin Workshop, Toulouse, 7-8 June 2001(http://www.cnrm.meteo.fr/aladin/meetings/Toulouse2001/RAPPORTS/Rapport_CF.aw.ps.gz)

- Hágel Edit, Horányi András: Sensitivity experiments of global singular vectors at the Hungarian Meteorological Service, ALADIN Newsletter 29, 2006,(http://www.cnrm.meteo.fr/aladin/newsletters/news29/N29WEB/HAGEL_HORANYI.pdf)

- Hágel Edit, Horányi András: Computation of Aladin singular vectors: Very first results, Vienna meeting 2006, (http://hirlam.org/open/publications/HLworkshops/GLAMEPS200611/Becs_ALADIN_SVfinal.ppt)

- Žagar Mark, Fischer Claude, Radnoti Gabor, Janoušek Martin, Trojáková Alena, Seity Yann: The ARPEGE/ALADIN Tech'Book: Implication of LAM aspects on the global model code, CY32/AL32, December 2006

ANNEX

Draft working plan for singular vector computations with ALADIN

Prepared by Edit Hagel and Andras Horanyi

February, 2007

Introduction

The singular vectors were always considered as powerful vehicles for the computations of perturbations for an ensemble prediction system. However for the time being their use is natural and proven to be beneficial only in the medium range. Recently the interest had been increased to use singular vectors for short range ensemble applications as well. It is a rather new area of interest not only for the ALADIN, but also for the HIRLAM project. It was decided that the development and application of the singular vector technique will be pursued in close cooperation between HIRLAM and ALADIN scientists (the main actors in that field are Jan Barkmeijer from HIRLAM and Edit Hagel from ALADIN). Hereafter a brief working plan is proposed especially taken into account the interest coming from Richard Mladek in Czech Republic (based on discussions with Radmila Brozkova in Budapest and Richard Mladek in Reading). At that stage, the preliminary validation (cy30, on the tora machine in Toulouse) of the ALADIN singular vector computations (configuration 601) was already performed by Edit Hagel. The first results were presented in the last ALADIN-HIRLAM LAMEPS workshop in Vienna (available at
http://hirlam.org/open/publications/HLworkshops/GLAMEPS_planmeeting_200611.html
).

The main objective of the work on singular vectors would be to deeply validate the singular vector computations for ALADIN (from the scientific point of view) and to inter-compare their results to global (possibly ARPEGE and IFS) counterparts with possibly different horizontal resolutions. At the next stage perturbations should be computed from the singular vectors and those perturbations should be used as initial conditions for a limited area ensemble prediction system. The proposed work is fitting to the GLAMEPS project's activities, therefore it is recommended that the work should be performed on the ECMWF machines under the GLAMEPS account (this would be also useful for accessing the outputs by every partners including HIRLAM).

The work should be performed in close cooperation with Edit Hagel, who is at the moment in Reading (until the end of March). This latter fact should be used, while the installations at ECMWF will be started (therefore an early start of the work would be welcome). At the same time it is mentioned that the proposed work needs significant persistent working capacity, i.e. as a strict minimum at least 3 man-months work devoted to that work should be considered for 2007. The bold faced letters indicate such subtopics, where the most help would be appreciated.

Tasks to be performed

1. **Technical preparations:**
    a. **Investigation of the accessibility of the global (ARPEGE and IFS) singular vectors to be used for the comparison: results are available at Meteo France and ECMWF as well (care should be taken to address the possibility of re-computation of singular vectors in Toulouse and Reading respectively).**
    b. **Testing and informatic validation of ALADIN configuration 601 in Reading (comparison to the results obtained in Toulouse – on tora with cy30 – by Edit Hagel).**
    c. **Testing and basic validation of ALADIN model integration in Reading.**

    Estimated workload: 4-8 weeks

2. Detailed comparison of global and limited area singular vectors on the same and different horizontal resolutions: limited area singular vectors on different resolutions and limited area and global singular vectors on the same resolution (beside ALADIN, ARPEGE, IFS and HIRLAM singular vectors should be used, the latter to be coordinated with Jan Barkmeijer). This task also requires the choice of interesting (from dynamical meteorological point of view) weather situations and evaluate and inter-compare the obtained results.

    Estimated workload: 4-8 weeks

3. Sensitivity of ALADIN singular vectors to the optimisation time and domain (horizontal and vertical extension), the role of simplified physics etc.

Estimated workload: 8-12 weeks

4. **Computation of perturbations for an ensemble prediction system based on the singular vectors computed with ALADIN. The software used in ECMWF for the combination of singular vectors into perturbations should be applied and adapted for that purpose.**
   Estimated workload: 4-8 weeks

5. First tests with ALADIN singular vector based ensemble prediction system.
   Estimated workload: 4-8 weeks

Outlook

The proposed work above can give a very solid basis for the ALADIN singular vector computations with special emphasis on their use for ensemble forecasting. It would be nice if more manpower could be devoted to that project in order to have significant progress in that field.