ALADIN NEWSLETTER 25

&

ALATNET NEWSLETTER 8

July–December 2003

ALADIN Newsletter 25 & ALATNET Newsletter 8

1. EDITORIAL

1.1. Introduction

These two joint Newsletters are quite short and published very late. So there will be no distinction between the different types of contributions (ALATNET or not) in the parts dedicated to research and development.

This situation may be explained as follows. First, many publications during the last six months of 2003 and the first ones of 2004 : late delivery of the previous Newsletters, proceedings of the last ALATNET seminar (October 2003) with extended contributions from the Young researchers, of the mini-workshop on data assimilation (October 2003) and of the 13th ALADIN workshop (November 2003). Second a lot of worries with the definition of the ALADIN-2 project and the closure of ALATNET (with the joint l**ast-year and final reports available** on the dedicated web site : http://www.cnrm.meteo.fr/alatnet/) during the first months of 2004.

The ALADIN Newsletter 26 should be more complete and ready in time.

1.2. Events

1.2.1. ALATNET seminar

The last ALATNET event was organised in Kiralyret (70 km N from Budapest) at 15-17 October, 2003. The idea of the seminar was born in the mid-term review in Brussels (April, 2002). The main objective of the workshop was to give an opportunity to the young researchers to present their work before the end of the project. The occasion was also used to discuss potential continuation possibilities of such research and training networking.

Beside the young researchers, supervisors, ALATNET coordinators also each SRNWP consortia obtained invitation to the seminar. Finally the SRNWP coordinator (Jean Quiby), the HIRLAM Project Leader (Per Unden) and a representative of the Met' Office (Nigel Wood) accepted the invitation and took part in the meeting.

The details of the programme together with the presentations and proceedings can be found in the ALATNET home page and at the homepage of the Hungarian Meteorological Service:

http://omsz.met.hu/ismeretterjesztes/rendezvenyek/alatnet2003/alatnet.php

1.2.2. Mini-workshop on data assimilation

The LACE community had an idea to organise a small workshop on data assimilation, where the progress in the field of ALADIN data assimilation can be discussed and future plans can be drawn. The workshop was organised in Budapest at the headquarters of the Hungarian Meteorological Service (HMS) at 20-22 October (just after the ALATNET seminar). The workshop was proved to be very useful and efficient. All the details (presentations) can be found at the homepage of HMS:

http://omsz.met.hu/ismeretterjesztes/rendezvenyek/aladin_ws2003/aladin_ws2003.php.

1.2.3. 8th Assembly of ALADIN Partners

The annual Assembly of partners took place in Cracow, on October 31th. Unluckily the Minutes won't be available. The following points may be underlined :

- Algeria was accepted as an Associated Member.
- The form of an enhanced cooperation with HIRLAM, especially as concerns very high resolution modelling, was discussed; however there was no firm HIRLAM position at that time.
- The strategic document on ALADIN-2 promised at the ALADIN-AROME meeting of April 2003

was presented to directors, but they asked for more, especially as concerns the impact on operations. A first revised version is available in the proceedings of the 13th ALADIN workshop.

1.2.4. 13th ALADIN workshop

This workshop, focussing on "ALADIN applications in very high resolution" was held in Prague, from 24 to 28 November 2003. The program was quite dense, and the proceedings are available by the Czech ALADIN team (filip.vana@chmi.cz).

1.3. Announcements

Too late !

1.4. Gossip

No time !

2. OPERATIONS

2.1. Introduction

The sparse informations provided here do not fully reflect the operational changes during the last semester of 2003. More details are available in the proceedings of the annual EWGLAM (presentations available on the SRNWP site http://srnwp.cscs.ch/ , ☛ Annual Meetings ☛ 2003) and ALADIN workshops. Or by writing to the corresponding contact points of course (**recently updated**).

2.2. Changes in the operational version of ARPEGE

(more details joel.stein@meteo.fr)

A few "neutral" changes first :

- ✔ 2003, June 30th : ***New computer (1)***
  - VPP 5000 ➔ VPP 5000,
  - slight changes in SST ⇒ non neutral,
  - computer failure 2 days later !
- ✔ 2003, July 28th : ***New library***
  - update of the source code version (CY26T1),
  - slight improvements in post-processing,
  - semi-Lagrangian advection : from 2 to 3 iterations,
  - use of TOVS data : variable emissivity over sea
- ✔ 2003, October 6th : ***New computer (2)***
  - VPP 5000 ➔ VPP 5000
  - improvements in post-processing: filtering, wind gusts (no longer over-estimated),
  - safety bug fixes

A major modification on 2003, December 8th : *New observations*

- ✔ improved use of raw AMSUA data (NOAA15, NOAA16, ~~NOAA17~~)
- ✔ use of raw HIRS data (NOAA16, NOAA17) ⇒ more humidity data !
- ✔ use of "GEOWIND" data (Meteosat 5-7 SATOBs in BUFR) (higher spatial and temporal (× 4) resolution, quality index
- ✔ real height for "10 m" wind SHIP observations (24 m on average !)
- ✔ far more observations used (~ + 17 % for the assimilation cycle), but impact mainly in the Tropics and Southern hemisphere

2.3. Austria

(more details thomas.haiden@zamg.ac.at)

2.4. Belgium

(more details olivier.latinne@oma.be)

2.5. Bulgaria

(more details andrey.bogatchev@meteo.bg)

2.6. Croatia

(more details ivateks@cirus.dhz.hr, tudor@cirus.dhz.hr)

ALADIN is operationally run twice a day, for 00 and 12 UTC on LACE and Croatian domains.

Characteristics of the LACE domain

- ✔ horizontal resolution 12.2 km
- ✔ 37 vertical levels
- ✔ time-step 514
- ✔ 229x205 grid points (240x216 with extension zone)
- ✔ latitude and longitude of corners : SW (34.00, 2.18), NE (55.62, 39.08).

Characteristics of the Croatian (HRn8) domain

- ✔ horizontal resolution 8 km
- ✔ 37 vertical levels
- ✔ time-step 327 sec
- ✔ 169x149 grid points (180x160)
- ✔ corners : SW (39.00, 25), NE (49.57, 22.30).

Initialisation of ALADIN on LACE domain is provided by Digital Filter Initialisation (DFI). Coupling frequency and frequency of output files for LACE and HRn8 domains are 3 hours. When the 48 hours forecast on LACE domain is finished 48 hours forecast for HRn8 starts without initialisation with coupling files from LACE.

6 domains are used for the dynamical adaptation of the wind field in lower troposphere to 2-km resolution orography. Four of them cover mountainous parts of Croatia (Karlovac, Senj, Maslenica, Split, Dubrovnik and Osijek). Dynamical adaptation is run sequentially, for each output file, with 3 hour interval. In the dynamical adaptation meteorological fields are first interpolated from the input 8-km resolution to the dynamical adaptation 2-km resolution. The same file is used as initial file and as coupling file that contains boundary conditions for the model. Dynamical adaptation is run on 15 levels, the 10 lowest levels are the same as for the Croatian or LACE domains.

Visualisation of numerous meteorological fields is done on LINUX PC. Comparison of forecasts with SYNOP data is performed hourly for today's and yesterday's forecast. The products are available on the Intranet & Internet. Internet address with some of the ALADIN products, total precipitation and 10 m wind :

http://prognoza.hr/aladin_prognoza_e.html &
http://www.tel.hr/dhmz/prognoza/aladin_prognoza_e.html

2 m temperature, surface pressure and 10 m wind speed data from 42 SYNOP stations are compared to model forecast data on all domains for today's and yesterday's runs every hour indicating when the state of the atmosphere starts to evolve differently than predicted by forecast.

Machine: SGI ORIGIN 3400, 16 x 400 MHz IP35 Processors, Main memory size: 12288 Mb, OS IRIX 6.5. Coupling files are retrieved from via Internet.

Changes in the second half of 2003 :

- ✔ From 30th of July 2003 12 UTC, operational version of ALADIN is AL25T1_op2.
- ✔ From 21st of December 2003, new domain for the high-resolution dynamical adaptation of surface wind is operational.

2.7. Czech Republic

(more details filip.vana@chmi.cz)

2.7.1. Evolution of the ALADIN/CE application.

There was no change in the operational application in the second half of 2003. Two modifications were under preparation, i.e. switch to higher resolution and a small modification in the diagnostics of clouds. It was decided that these two changes would be switched to the

operational service at the same time, on 13 January 2004.

2.7.2. Parallel Suites

The following parallel tests were launched to assess the impact of different modifications:

- Suite ACN: this suite was a major test of increasing the spatial resolution of ALADIN/CE. In horizontal the mesh size decreased from 12.2 km to 9 km; the number of vertical levels increased from 37 to 43. The spacing of vertical levels followed the previous strategy: increasing resolution in low troposphere and avoiding high stratosphere. Further, spectral resolution was increased by switching to a « linear grid ». These changes required new tuning of blending parameters (cut-off truncation and digital filter settings) and of horizontal diffusion. For example the tuning of horizontal diffusion was found in agreement with values used for ALADIN/France using also the linear grid. The scores, computed for winter and summer periods, were quite satisfactory, namely those computed against SYNOP observations. In altitude only geopotential scores with respect to TEMP measurements were slightly worse for summer testing period. The «August 2002 flood case» was rerun and results were found slightly more realistic compared to the 12km version as well as to the physics used operationally at that time. The suite was declared ready for the operational switch.
- Suite ACO: it provided a reference where mean orography is used instead of the envelope orography for experiments with new version of orographic lift and form drag.
- Suites ACP, ACQ: test of the first version of the lift and form drag, with various tunings. The results were not quite satisfactory.
- Suites ACR, ACS: in these tests the formulation of form drag was further modified and gave better results. The remaining problem is bias of 10m wind (too weak) and 2m temperature (too cold).
- Suite ACT: it was a fix of the compilation problem discovered for the routine *suehdf.F90*. The optimization of the code was changing results and therefore the routine was compiled without optimization. A parallel test was made for safety reasons but the scores remained perfectly neutral.
- Suites ACU, ACV, ACW, ACX, ACZ: in all these tests further improvements of the form drag and tunings were made. There were no significant changes compared to the previous results.
- Suite ACY: test of the switch LREWS, using higher resolution version (the base is ACN suite). It was found to have a significant impact on the scores. This suite was a part of the preparatory steps toward full COCONUT physics.
- Suite ADA: here the full version of COCONUT physics was tested at high resolution (ACN) with an additional modification of the cloud diagnostics (proposed by F. Bouyssel and tested in ARPEGE parallel suite cy26t1_op6). The purpose was to have more realistic amount of low level cloudiness. Badly represented (too few) stratus clouds in ALADIN forecast represent quite a serious problem in winter season, leading to large errors in 2m temperature forecast. The modification helped a bit, since it improved 2m temperature scores for the typical «low level clouds season»; more realistic amount of low level clouds could have been noticed on two selected situations. The suite ADA, merged with ACN, was declared ready for the operational switch. It was scheduled for 13 January 2004. On the other hand, the cloudiness scheme of COCONUT version based on another formula of Xu and Randall, produces a lot of 0/1 cloud amounts, with very few intermediate values. The additional modification in the cloud diagnostics improving the total cloud amounts alleviates a bit this shortcoming but not sufficiently. Therefore further research in this area will be needed.
- Suite ADB: this suite is based on ADA suite, where more consistent expression is used in the radiation scheme. This suite is part of research made on the radiation scheme currently. The results of this particular change were neutral.

The results of parallel tests may be consulted on :

www.chmi.cz/meteo/ov/lace/aladin_lace/partests

2.7.3. ALADIN/MFSTEP configuration

CHMI is one of members of MFSTEP (Mediterranean Forecasting System Toward Environmental Predictions) project financed by the 5th Framework Program of European Commission. Its role is to provide atmospheric forcing data for basin and shelf models in the Mediterranean Sea, near Atlantic Ocean and Black Sea, in near real time regime. Therefore a special configuration of ALADIN, named ALADIN/MFSTEP, has been created, having an horizontal resolution of 9.5km with 589x309 points (Figure 1). The Lambert projection is used with the reference point at [2.58E; 46.47N]. The domain centre is at [9.81E; 41.95N]. The elliptical spectral truncation is 299/159 on a linear grid. Vertically there are 37 irregularly distributed levels from 17m above the surface at the bottom to 5hPa on the model top. The model time step is 400s. ALADIN/MFSTEP runs in DF blending 'assimilation' mode and a production up to 120h takes place once a week on Wednesdays. The results from blending cycle and production are post-processed every hour on two latitude/longitude regular grids corresponding to each sea basin (green and red frame on Figure 1).

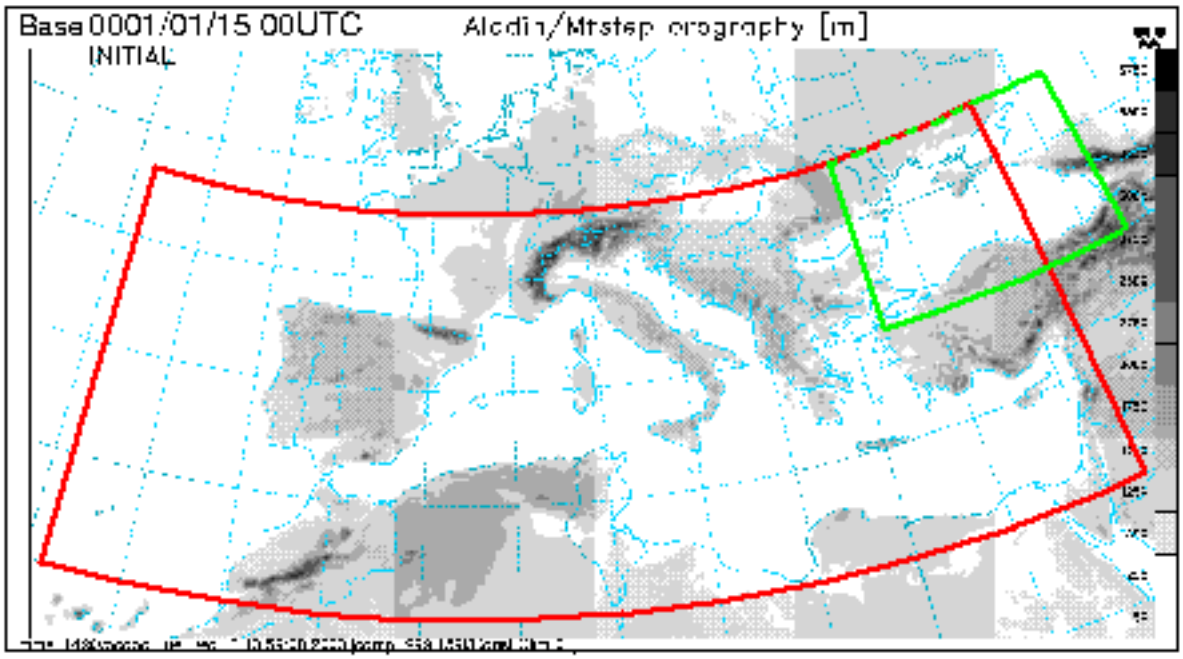


Figure 1: ALADIN/MFSTEP domain with red and green frames of post-processing areas.

The ALADIN/MFSTEP is different from the reference operational version in Prague. Couple of developments was made to improve the model performance over the sea with the emphasis on the surface fluxes and other near surface data which are the input for the ocean modelers. Substantial part of those tests was carried out on the so-called 'Black sea case' (+48h forecast from 12/09/2003), when an intensive cyclogenesis occurred with the minimum value of the forecast mean sea level pressure at 986hPa. This low pressure system was observed in reality, but its predicted minimum value was too deep. This case is hence very interesting and it was reported by Romanian colleagues, too. Many tests were performed with the aim to improve this forecast by keeping the cyclone with a more realistic pressure drop. The most important steps leading to the ALADIN/MFSTEP system improvement were:

- Introduction of a modified computation of the thermal roughness length over the sea;
- Activation of an already existing 'moist gustiness' parameterization;
- Introduction of several 'safety' constants in physics;

- Tuning of blending parameters.

Final version of the tunings and modifications was also carefully tested in the parallel suite environment. This setup was then used for scientific validation period of MFSTEP project as well as to start the preliminary near real time suite.

The selective semi-Lagrangian horizontal diffusion (SLHD) was introduced into the ALADIN/MFSTEP system but more validation and tuning has to be done before activating it in the nominal configuration. Some attempts were made to get-rid of the envelope orography (including also some tests of the semi-envelope orography) but the obtained results were not convincing so far. However this work will continue, as well as research on improving and optimizing the radiation scheme. Useful results of this research will be ported to the ALADIN reference.

2.7.4. DiagPack

During the second part of 2003 a (pre)operational version of CANARI - DiagPack was implemented. The analysis of the mean sea level pressure, the 2 meters' temperature and humidity, the 10 meters' wind, the diagnostic of KO-index, CAPE and MOCON are performed every hour. Up to now only SYNOP observation are used and following figures show the 2 meters' temperature analysis with used observations and the guess.

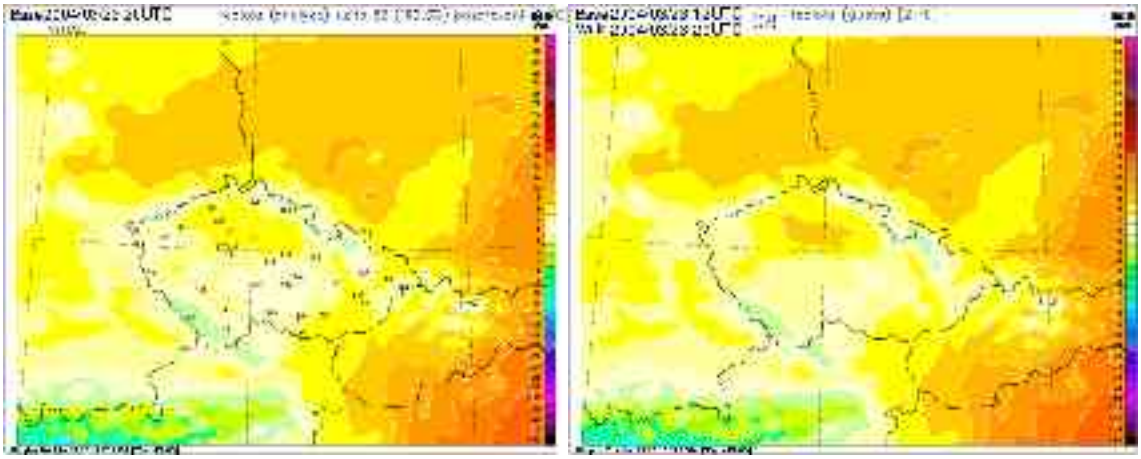


Figure 2: DiagPack analysis and guess of the 2 meters' temperature with observations included.

2.8. France

(more details joel.stein@meteo.fr)

Same model changes for ALADIN-France as for ARPEGE.

2.9. Hungary

(more details kertesz.s@met.hu)

There were no changes at all regarding the operational suite of the ALADIN/HU model.

Nevertheless the intensive tests of the ALADIN 3d-var data assimilation scheme with different observations continued in the second half of 2003.

It is also mentioned here that the CY26T1 version of the code was installed (but not used operationally) and the post-phasing of the B-level parallelization was performed on it.

2.10. Morocco

(more details ajjaji@marocmeteo.ma)

2.11. Poland

(more details zijerczy@cyf-kr.pl)

2.12. Portugal

(more details margarida.belo@meteo.pt)

During the second half of 2003 no relevant changes took place on the Portuguese operational suite (AL12_bf_CYCORA_bis). However, significant effort has been put on the last year objective verification of the model. As a main result we mention the detection of an abnormal increase of the 2m dew-point temperature bias, mostly during the summer season. Moreover a new approach for the verification tools has been attempted in order to make them more interactive to the user of ALADIN/Portugal products. New diagnostic tools have been built (Q divergence vector) and a local manual on the usage of those tools was started. The validation work of the diagnostic tools is in progress using 2001 particular weather cases of deep convection. Local documentation of CANARI was also started. As a final remark we mention the well-succeeded home workshop to promote and discuss NWP issues with forecasters and other NWP users.

2.13. Romania

(more details doina.banciu@meteo.inmh.ro)

2.14. Slovakia

(more details oldrich.spaniel@shmu.sk)

The main event during the last quarter of 2003 was related to finishing the process of public procurement for a new high performance computing system. The computer by IBM was chosen with following configuration:
IBM @server pSeries 690, Typ 7040 Model 681, 32 processors POWER 4+ 1,7 GHz , 32 GB RAM of memory, IBM FASt T600 Storage Server + EXP700 - 1,5 TB.

The benchmark test was based on cycle AL25T1 with a domain of 320x288 points, horizontal resolution 8.3 km and 51 vertical layers, with the time-limit 28 minutes for configuration *morgane*. The operational suite is expected from the middle of 2004 year. By now the current operational state is the same as in the previous report.

2.15. Slovenia

(more details neva.pristov@rzs-hm.si)

After a testing period, the new operational suite (see details in the previous Newsletter) was declared as operational from 1st June. The performance was checked regularly since then and few improvements of hardware and software were done:
- hyper-threading on computing boxes was switched-off due to degraded performance when switched on,
- massive storage disk array was installed (with the target capacity of 4 TB),
- transfer of LBC files was included into sms operational suite ,
- LBC files archive on DVD media started ,
- announcement of delays via email to users was introduced.

During this period the internal intranet pages were improved, mostly those about the monitoring of the products. Script for users to run ALADIN model in operational configuration in SMS system was prepared. Required model products for PEPS project were sent to DWD.

In the following text a short overview about the availability of LBC files and of 48-hour model integration along last year is presented.

The coupling files from ARPEGE model are transfered via Internet from Toulouse. The average time for transfer of one file is one minute (transfer rate 128kB/s max. 153kB/s). Files were significantly delayed 15 times (4.1%) in the morning (after 4:30 UTC) and 11 times (3.0 %) in the afternoon (after 16:30 UTC). Main reasons are that connection to *sirius1* or *sirius2* was not possible or files appeared late in the database.

Performance of the model integration is presented in the following table and for the new

operational suite also in Fig. 2. The old operational suite was still running on our old cluster for a whole year.

	old operational suite		new operational suite
model integration finished	1.1.2003 - 31.12.2003	1.1.2003 - 31.5.2003	1.6.2003 - 31.12.2003
00 run - after 6 UTC	46 (12.6%)	22 (14.6%)	26 (12.1%)
12 run - after 18 UTC	28 (7.7%)	15 (10.0%)	16 (7.5%)

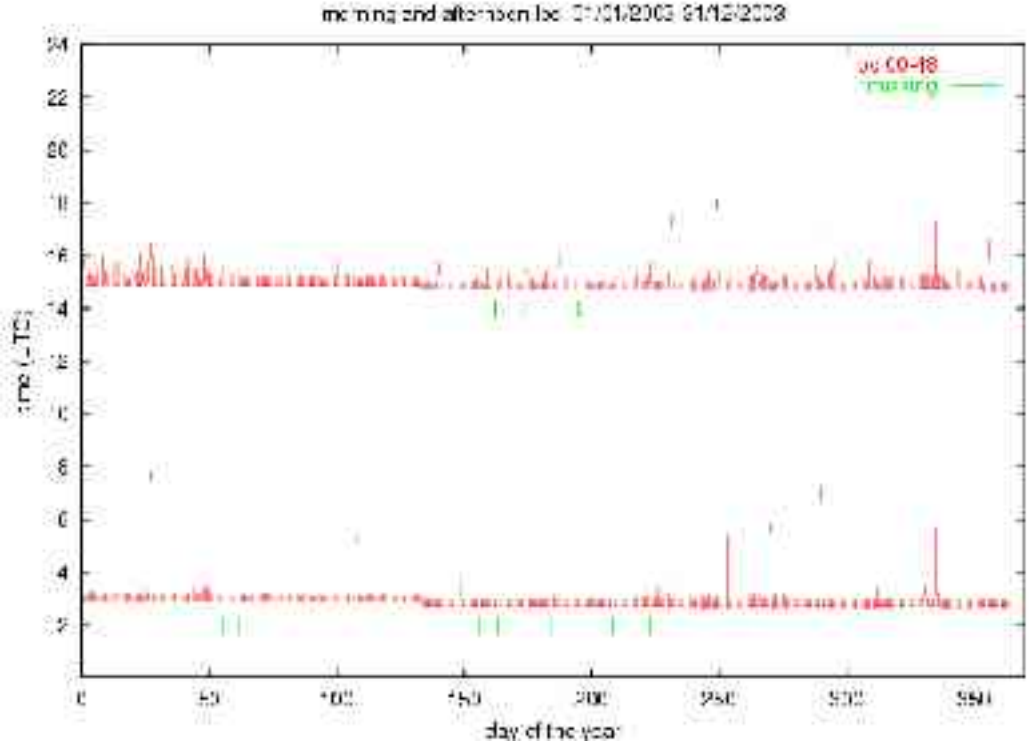


Figure 1: Time intervals for transferring LBC files, red bars indicate time between end of transfer of the first and last LBC files, green bar indicate that file(s) were transfered later manually.

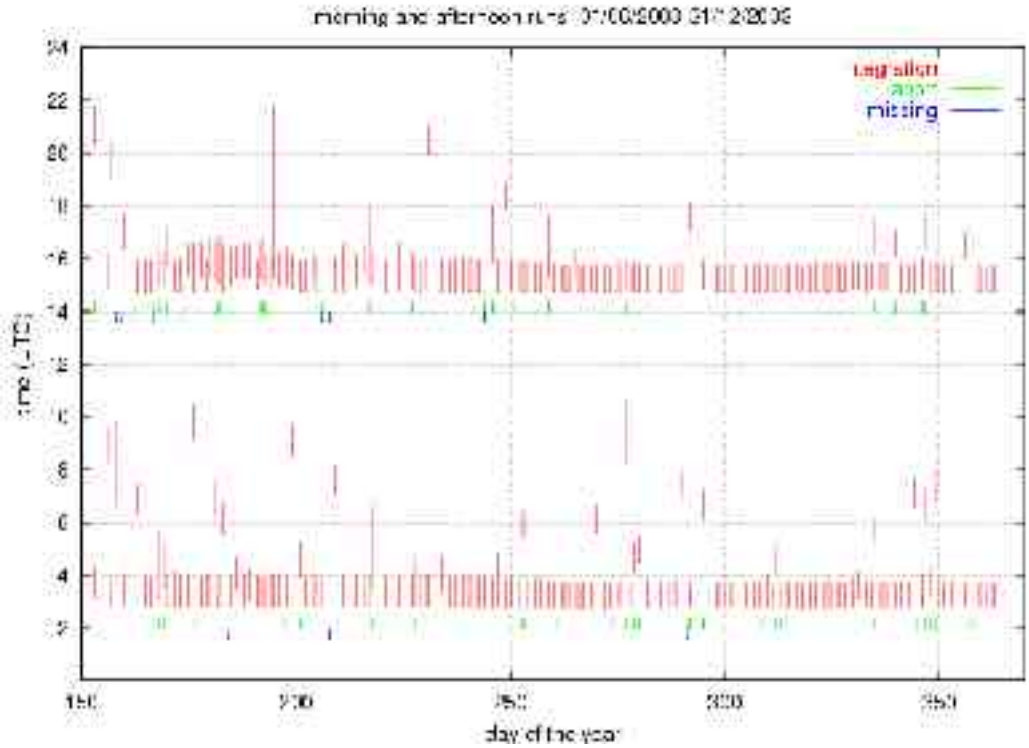


Figure 2: Model integration time intervals for the model ALADIN/SI (new configuration) runs in year

2003 (Jun-Dec). Red bars indicate time between start and end of model integration, green bar indicate that abort message appeared, blue bar indicate that model was available that day.

The main unresolved problem in the new operational suite are non-repeatable aborts during integration of the model. They are of two types and we haven't find solution yet.

- While writing history or fullpos files, integration can abort with message : *"MPI_BSEND : Insufficient space available in user-defined buffer"*. This happens rare during 48 hour model integration (4% in 2003), while in the integration of dynamical wind and precipitation adaptation more often (17% in 2003, which means in average 1.9 times out of 11 DADA runs per one whole suite).
- The other message which occurs is *"In DSINH(dx) or SINH(dx) or DCOSH(dx) or COSH(dx), DABS(dx).ge.710.475 (dx=nan) Error occurs at g_dcosh"*. This kind of abort happens on different places : before DFI, in the end of DFI or somewhere during integration (in 3.6% cases in year 2003). The error itself happens in physics and it is not repeatable.

Occasionally ALADIN binary was not started in SCore. This was solved with better monitoring of SCore program.

If there are other jobs running at the same time, aborts are more frequent. There is a simple queuing system which takes care about preempting jobs but it looks that it has some problems and we are thinking about a more proper queuing system.

The number of cases of missed or delayed products has not been decreased in the new operational suite. Automatic rerun of the script in SMS is beneficial but there are still cases where human intervention is needed. This can be improved with monitoring of the suite while it is active.

2.16. Tunisia

(more details nmiri@meteo.nat.tn)

3. RESEARCH AND DEVELOPMENTS

3.1. Austria

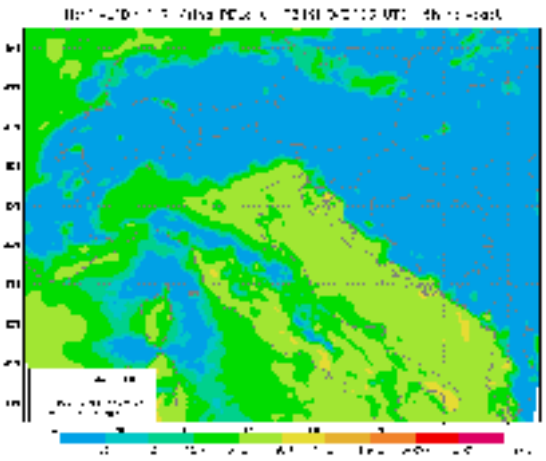
3.2. Belgium

3.3. Bulgaria

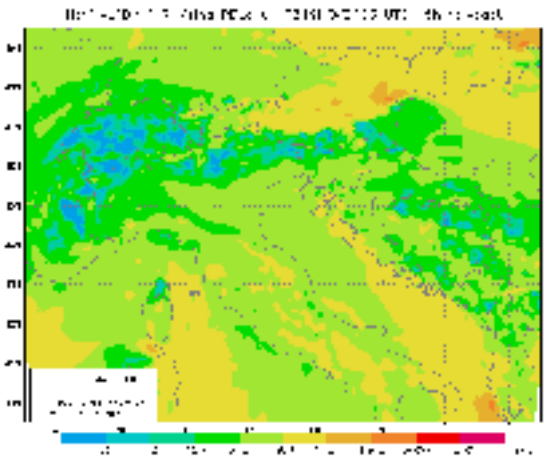
3.4. Croatia

3.4.1. PBL height determination

tudor@cirus.dhz.hr



before modification



after modification

Operational version of ALADIN used in Zagreb is AL25T1_op2. This one calculates PBL height using a modified formula from Ayotte (1996) where PBL height is determined using virtual potential temperature vertical integral. PBL height is determined using the vertical integral of virtual potential temperature. If the value at the level *z* is significantly larger than its vertical integral value $1/z \int_0^z \theta_v(z)\,dz$, PBL height is detected. However, using this formula, PBL height is detected very close to the ground in statically stable situation. The effect of wind shear on stability is neglected.

The virtual potential temperature profile is modified to include wind shear impact on stability. The value at surface remains the same, above it the values are modified by wind shear :

$$\theta^*_{Vl} = \theta^*_{Vl+1} + (\theta_{Vl} - \theta_{Vl+1}) - \frac{\theta_{Vl} + \theta_{Vl+1}}{2\,\Delta\Phi}\,(\Delta u^2 + \Delta v^2)$$

The new values θ^*_v are used in the integral. If the wind on the lowest model level is strong, the diagnosed PBL height are high. In statically stable situations with weak wind, the diagnosed values are low. The vertical integral makes the method robust to oscillations the model variables might have in the vertical. The figures show forecasted PBL height before (left) and after (right) introduction of the wind shear impact.

3.4.2. New meteograms and pseudo EPS

drvar@cirus.dhz.hr

In order to take full benefit of the mesoscale ALADIN model the method of «Pseudoensembles» was developed by analogy with Ensemble Prediction System (EPS) that is operationally running in European Centre for Medium-range Weather Forecasts (ECMWF).

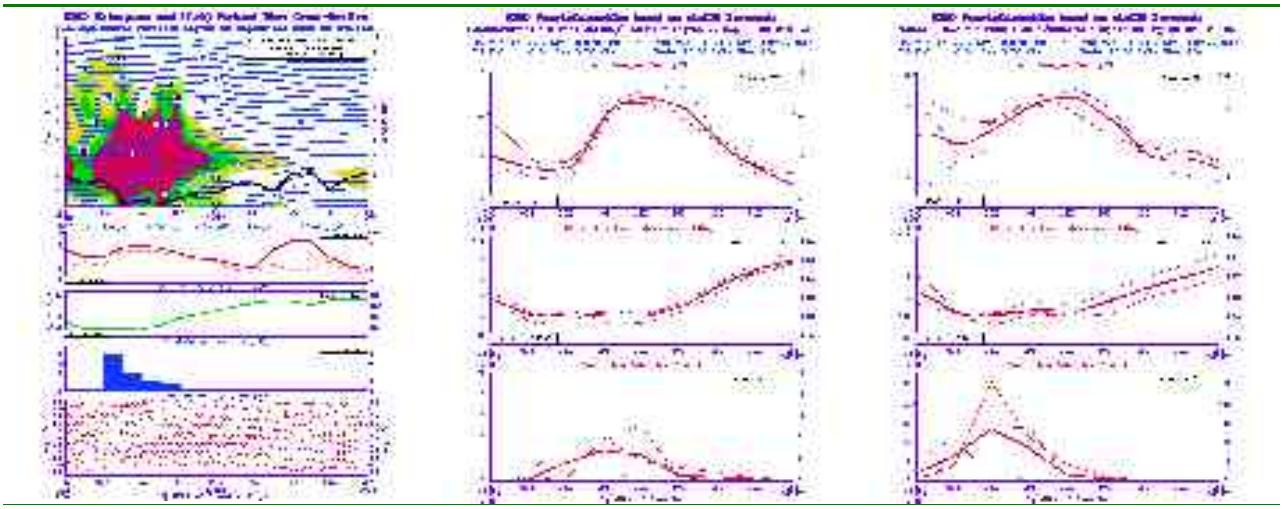
Conceptually, instead of the perturbation method used in producing fifty ensemble members,

the different initial states for the «Pseudoensembles»are provided by three successive ALADIN runs based on ARPEGE coupling files.

The main goal in elaborating this method was to examine the consistency of the successive forecasts daily produced in MHS of Croatia with two resolutions of the mesoscale ALADIN model (not shown here). In fact, in most cases the changes are quite small, but occasionally there might be large differences in dynamically sensitive areas, where successive forecasts can be in conflict with each other. Having such an objective information either on more or less consistent solutions in various parts of the model domain, or even on the possible inconsistencies, the forecaster is able to draw conclusion on the most likely development. On the other hand, it might be useful feedback providing information on the model behaviour.

The «Pseudoensembles» make a special part of the HRID diagnostic package working on ALADIN direct model output. For the time being, they have been applied quite locally at some particular grid-points. It includes: "*peudoensemble mean"* of the vertical time cross-sections based on several elementary and derived parameters, *"pseudoplumes"* of the geopotential, temperature and relative humidity at several standard pressure levels as well as for the 2m temperature and relative humidity, wind speed, mean sea level pressure and total precipitation and *"pseudoensemble spread"* information.

The example ilustrated on the figures corresponds to a cold front passage over Zagreb and the mesoscale vortex in the Northern Adriatic with extreme precipitation event.



3.4.3. Calculation of PBL height and stability in the lower layer of the atmosphere for special purpose

vidic@cirus.dhz.hr , jericevic@cirus.dhz.hr & spoler@cirus.dhz.hr

Forecast values of some atmospheric parameters are calculated for input to the dispersion model for Croatian Oil refinery Urinj, near Rijeka;

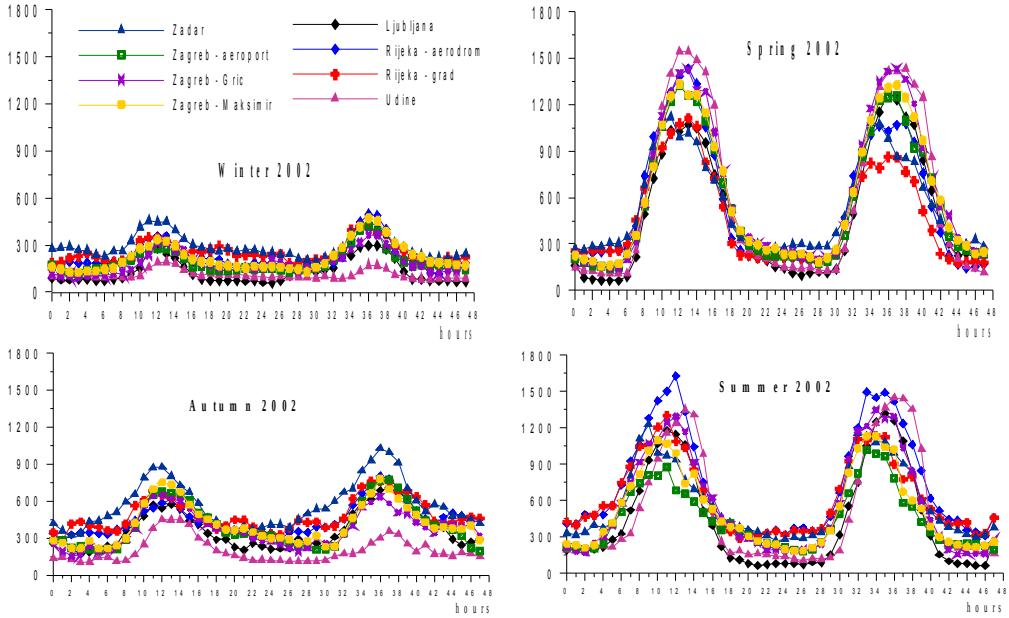
To calculate ground-level concentrations in the vicinity of the oil refinery, located in complex coastal terrain, two procedures have been developed:

- Procedure to determine atmospheric stability parameter in a layer of vertical plume transport (0-150 m above ground) and dispersion, and
- Procedure to calculate hourly values of mixing layer height (0-4500 m above ground).

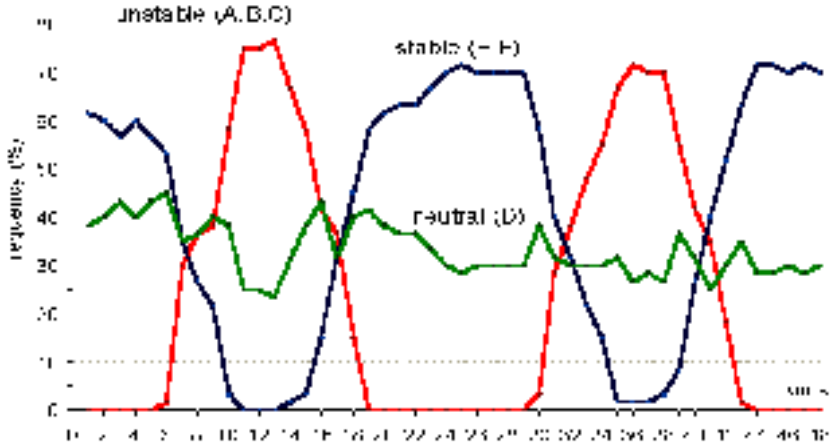
Both procedures are developed on the basis of ALADIN output: wind and temperature fields at 8 km horizontal resolution and 20 layers in the vertical.

Mixing-layer height is determined on the basis of critical bulk Richardson number. Critical value of 0.1 is used to separate turbulent mixing layer from free troposphere. Results were tested on

four seasonal experiments (winter, spring, summer and autumn) and for eight different locations (see below).



Mixing layer height (m)



Stability parameter

Stability parameter is determined within first 150 m above ground level. Five methods have been tested: Pasquill-Gifford method (P-G), vertical temperature gradient method (VTG), VTG method that takes into account wind speed in the middle of the layer (VOGT), Richardson and Bulk-Richardson number methods. As a result, the method that was sensitive enough to explain hour-to-hour stability parameter fluctuations was the method that combines wind speed and temperature gradient (VOGT method). Results were tested on four seasonal experiments (not shown here).

3.4.4. Reduction of envelope and optimisation of subgrid orographic drag parameterisation in ARPEGE

drvar@cirus.dhz.hr

The work aimed at reducing the envelope and optimizing the gravity wave drag parametrization. The envelope was reduced using the spectral cost function as described in (Marguinaud and Moudden, 1999) and gravity wave drag coefficients were tuned with respect to climate model and theoretical values. The most satisfying results in terms of scores were gained

with the combination of lift coefficient set to 0.5 (which is the value used in climate model), form drag coefficient set to 2.0 (which is the theoretical value), and total coefficient 0.5E-2. The tests were performed on a period of 8 winter days, the scheme needed to be tested on longer period, and some other times of the year with different global circulations. Further work was done by Francois Bouyssel and Jean Francois Geleyn and can be referenced at Météo France.

3.4.5. Other Work

dijana@cirus.dhz.hr , *ivateks@cirus.dhz.hr* , *ttrosic@irb.hr* , *tudor@irb.hr*

The following research work was presented on the 13th ALADIN workshop and will be described in the proceedings:

Martina Tudor	ALADIN operational suite in Croatia
Stjepan Ivatek-Šahdan	Coupling frequency - two time nesting
Martina Tudor	Robustness of the physical parameterization schemes
Stjepan Ivatek-Šahdan	Smoothing of SWI (soil wetness index)
Dijana Klari?	Test of DFI Blending on several MAP cases (IOP2b,IOP3, IOP5, IOP8)
Martina Tudor	High resolution dynamical adaptation using hydrostatic ALADIN
Tanja Troši?	Ad hoc comparison of two atmospheric model predictions - a case of February 2003 bora

3.5. Czech Republic

3.5.1. Theoretical aspects of non-hydrostatism

Top and bottom boundary conditions (R. Brožková, J. Mašek, P. Bénard)

While the "chimney" problem of the semi-Lagrangian advection scheme is practically solved, recent experiments confirmed that "chimneys" can be generated also by other mechanisms. For example, linear spectral diffusion applied on vertical divergence also generates this syndrome. Such effects of horizontal diffusion were suspected due to experimental results obtained already last spring. However only after removing the chimneys from the semi-Lagrangian scheme it could be confirmed. Further, a linear analysis made by P. Bénard for a stationary mountain flow has also shown that the linear spectral diffusion applied on vertical divergence creates a spurious term, affecting the surface vertical velocity, responsible for the chimney. The chimney problem is explained more generally in an extended abstract (Brožková et al., 2003).

There were also other problems studied, for example whether the use of vertical divergence cannot lead to multiple solutions of the equation system (perhaps to those with chimneys) in contrary to the use of vertical velocity as a prognostic variable. It was shown that the equation system using vertical divergence has the same solution as when using vertical velocity variable, due to the chosen top and bottom boundary conditions applied on the pressure field.

A new problem has appeared which however is probably not linked to the boundary conditions. The results of density current (bubble) test are more correct when using semi-Lagrangian advection of vertical velocity (introduced at first to cure the bottom boundary treatment in the semi-Lagrangian scheme) instead of vertical divergence. As one can see, the question of using vertical velocity or vertical divergence comes back to the table. On the other hand the Eulerian test, using vertical divergence, provides also correct result although it suffers due to a higher level of noise. This suggests that there is still a hook in the semi-Lagrangian advection of vertical divergence, creating this time a much weaker but still not completely negligible syndrome.

Finally, there is nothing really new about the so-called Z term. One alternative discretization of this term was tried in the bubble test, without any impact.

Predictor-Corrector scheme (J. Vivoda)

We can consider this topic as successfully finished by now. There is a comprehensive Young Researcher report available.

3.5.2. Data assimilation related issues

Digital Filter Blending and Explicit Blending (H. Tóth)

In order to understand better the role of individual steps made within the blending procedure, a diagnostic tool was developed based on the so-called Lönnberg-Hollingsworth statistics of forecast errors. These statistics were computed not only for guess, but also for intermediate states used in the blending algorithm (filtered ARPEGE analysis, filtered ALADIN guess, blending analysis). Results obtained so far are preliminary and should be further examined to draw conclusions. At the same time it is very useful to have at disposal such a diagnostic tool.

Besides, an explicit blending method was developed and compared to DF blending. The explicit method is very simple; spectra of ALADIN guess and ARPEGE analysis are blended directly; a linear function is used up to the cut-off truncation Tc. The cut-off truncation Tc was taken the same as in the DF blending setup for recent ALADIN/LACE application. Results of both methods were compared with help of objective scores and also some case studies were made. The scores of forecasts starting from either explicitly blended state or DF blended state were quite the same. It should be noted, however, that incremental digital filter was used and that in case of explicit blending it seemed to act significantly on the initial state, as it could have been noticed on the vertical velocity fields. For this reason the scores were computed for guess in the blending cycle. There some differences occurred; scores were a bit worse for the explicit blending in case of mean sea level pressure. A case study was made and precipitations were looked at, but it was difficult to draw a conclusion. What can be said at this stage is that any blending step needs a filtering. In case of DF blending it is built in-core and makes the blending function implicit; the digital filter applied on increments is not a necessity but security. In case of explicit blending the resulting increments need a filter to recover the balance; as seen from the results it would be dangerous to skip the incremental filtering within the blending cycle.

Technical Aspects (R. Brožková, A. Trojáková)

Data assimilation configurations of screening and minimization were successfully tested with cycle AL25T1 for conventional observations, when using sets of standard and lagged statistics. Also the optimal interpolation based diagnostic analyses works correctly for SYNOP observations.

3.5.3. Horizontal diffusion (F. Vá?a)

The Semi-Lagrangian Horizontal Diffusion (SLHD) scheme was tested on specially selected real cases in ALADIN, including proper choice of domains. The scheme either improved forecast or at least didn't worsen the reference model performance.

At the end of September 2003, the PhD study profiting from the SLHD research was defended at the Charles University in Prague.

Starting from August 2003, the SLHD scheme was phased to the recent model cycle. Simultaneously, it has been extended in such a way to be available in various configurations of ARPEGE/ALADIN like the iterative (P/C) scheme, global model (uniform and stretched grid) or even tangent-linear and adjoint model configurations. Basic SLHD code should be available from cycles CY28T1/AL28T1.

During the continuous real case experiment of the SLHD scheme some problems have been detected. Mainly they concern tuning when the resolution and model time-step are changed. There is also some potential to further reduce the computational cost. Hence some further work is planned for 2004 to finalise the SLHD related source and answer remaining open questions around it.

3.5.4. Effects of the orographic lift (J.-F. Geleyn, R. Brož ková, Mária Derková)

Many beta-tests were made to find the optimal formulation of the orographic lift and form-drag parameterizations, including the tunings. The goal is to possibly get rid of the use of envelope orography, where its volume effect is compensated by the lift and form-drag. Latest results are

more or less satisfactory; the remaining problem is too weak wind in the valleys as it comes up from the bias of 10m wind forecast computed against SYNOP observations. There is an associated cold bias of 2m temperature. A comprehensive description of this work may be found in 13th ALADIN Workshop Proceedings (Geleyn et al., 2003).

References

Brožková, R., P. Smolíková and C. Smith, 2003: Bottom Boundary Condition Formulation and Associated Problems. *13th ALADIN Workshop Proceedings on "ALADIN applications in very high resolution", Prague 24-28 November 2003, in print.*

Geleyn, J-F., F. Bouyssel, R. Brožková, B. Catry, M. Derková and D. Drvar, 2003: Foreseen changes of the mountain drag/lift effects in the 'ACDRAG' parameterization scheme. *13th ALADIN Workshop Proceedings on "ALADIN applications in very high resolution", Prague 24-28 November 2003, in print.*

3.6. France

The description of the work of ALADIN visitors in Toulouse is shared between Newsletters, so that some contributions were already mentioned in Newsletter 24 while the present report also covers studies performed in 2004 (as a continuation of a 2003 research work).

3.6.1. Phasing

This represented once again a very significant part of the work performed in Toulouse along the last four months of the year, gathering the efforts of 10 persons from GMAP, veterans and newcomers (Jean-Marc Audoin, Gérald Desroziers, Ryad El Khatib, Claude Fischer, Gwenaëlle Hello, Patrick Moll, Dominique Puech, Patrick Saez, Yann Seity, Karim Yessad), the GCO team, and 9 ALADIN experts (Andrey Bogatchev, Chantal Moussy, Gabor Radnoti, Oldrich Spaniel, Martina Tudor, Rashyd Zaaboul and lately Filip Vana) under the supervision of Claude Fischer, as usual.

This led to the building of cycle 27, just followed by cycle 28 after an "automatic" cleaning process. The new libraries include significant changes : new data flow ("GFL", assumed to make the introduction of new fields easier), new coding rules and interfaces (described by R. El Khatib in a paper now available on the ALADIN web site, under item "Documentation"), latest developments in non-hydrostatic dynamics (NH). A detailed description was sent to ALADIN partners by Claude Fischer. A summary by Ryad El Khatib is also available in the proceedings of the 14th ALADIN workshop : http://www.zamg.ac.at/workshop2004/ibk_2004.html.

3.6.2. Dynamics and coupling

A huge step was performed with CY27 (28) : definition and test of the new data flow for coupling and NH variables, introduction (and debugging) of the latest developments in non-hydrostatic dynamics and in the semi-Lagrangian horizontal diffusion scheme, ...

Besides analytical studies were successfully pursued. The requirements for a semi-implicit time-stepping in case of mass- versus height-type vertical coordinates were examined in cooperation with CMC (Pierre Bénard and Claude Girard). This resulted in the correction of a weakness in the MC2 model and the proposal of new objectives for the SRNWP NT network.

The origin of the "chimney" problem in NH dynamics when horizontal diffusion is introduced ("diffusive chimneys") was identified by Jan Masek (in Prague) and Pierre Bénard. As for "semi-Lagrangian chimneys" it is due to an inconsistency between dynamical equations and the formulation of the lower boundary condition : not taking into account contributions from horizontal diffusion and physics here.

As concerns the radiative upper boundary condition (RUBC), Martin Janousek investigated the potential impact of the semi-implicit treatment (changing the phase-speed of fast waves) on the efficiency of RUBC. This answer is quite optimistic : no theoretical obstacle to combine both. Practical experimentation should start soon, once the code stabilized.

Alena Trojakova compared the exact introduction of diabatic forcing in non-hydrostatic dynamics (Eq. 1), when both temperature and pressure fields are modified,

$$\frac{dp}{dt} = -\frac{c_p p}{c_v} D_3 + \frac{\boldsymbol{Q\, p}}{\boldsymbol{c_v T}} \quad ; \quad \frac{dT}{dt} = -\frac{RT}{c_v} D_3 + \frac{Q}{\boldsymbol{c_v}} \quad (1),$$

to the previously used "hydrostatic adjustment" (Eq. 2), when the diabatic tendency for pressure is neglected to limit the generation of acoustic waves by the instantaneous heating Q :

$$\frac{dp}{dt} = -\frac{c_p p}{c_v} D_3 \quad ; \quad \frac{dT}{dt} = -\frac{RT}{c_v} D_3 + \frac{Q}{\boldsymbol{c_p}} \quad (2).$$

Equation (1) is not only "exact" but also simpler, being consistent with the definition of NH variables and the continuity equation. It was coded in the model, then academic 2D experiments were performed to compare the two formulations. Both solutions converge very quickly, and the fields obtained with (1) look more sensible along the first time steps.

As concerns AROME, the choice of the *d*4 variable with a semi-implicit semi-Lagrangian two-time-level advection scheme was confirmed (Yann Seity and Gwenaëlle Hello). The work on the definition of a general set of equations including NH dynamics and a detailed set of humidity variables, was pursued (Pierre Bénard, Joël Stein, Sylvie Malardel).

3.6.3. Physics

The FMR15 radiation scheme, i.e. a former "frozen" version of the Morcrette model, was successfully tested by Yves Bouteloup and François Bouyssel in ARPEGE. There is a 10% increase in cost when calling the radiation scheme every 3h (instead of every time-step for the previous one) but a noticeable improvement in low-cloudiness situations. The other schemes used at ECMWF were also tested. For more details, see the proceedings of the 13th ALADIN workshop. Besides a retuning of cloudiness was performed, in order to get more "intermediate" values between 0. and 1.

The dependence on horizontal resolution in the deep-convection parameterization was once again retuned, and two problems in updraughts were identified (but not yet solved) by Jean-Marcel Piriou. Besides the KFB scheme was introduced in the 1D version of AROME by Gwenaëlle Hello and Yann Seity.

Bart Catry pursued his PhD work on the impact of resolution on orographic forcing, in the quasi-academic ALPIA framework, which provided key informations for the improvement of the corresponding parameterizations. See the presentations on this topic at the 13th (proceedings available at CHMI) and 14th (http://www.zamg.ac.at/workshop2004/ibk_2004.html) ALADIN workshops.

3.6.4. Surface

Gianpaolo Balsamo (mainly), François Bouyssel, Eric Bazile, Olivier Latinne and Dominique Giard further investigated the potential use of the ECOCLIMAP databases, in the framework of the ELDAS project. Some more weaknesses were identified, and the results of assimilation experiments still show a deterioration of the forecast skill.

The initialization of SST in ARPEGE was improved, with a new sea-ice mask, computed using SSM/I observations, and a finer "background" SST field from NCEP (resolution 0.5° instead of 1°).

Lora Taseva studied how to use the ECMWF snow analysis to improve the initialization of snow cover in ARPEGE, for the global version of the corresponding O.I. analysis is far from ready. This doesn't prevent from implementing a fine-scale analysis in ALADIN to further improve this raw first guess, of course. The following steps were performed :

- extraction of available data (equivalent water content and density of snow) on initial grid,
- conversion into a snow-depth field considering the discontinuities related to the land-sea and ice-

cap masks,
- interpolation on to the ARPEGE grid with the comparison of 3 methods (4-points interpolation was chosen) ,
- further corrections according to other surface fields.
Besides a quality control was designed for the corresponding SYNOP observations (the density of which shows a high time- and space- variability, with measures concentrated in the morning). A first evaluation was performed, on three 10-day winter periods.

3.6.5. 3D-Var

Jb

Improvements in the formulation of the background cost-function were proposed, following the work of Loïk Berre, Claude Fischer (isotropy, in cooperation with Gergö Bölöni in Budapest), Simona Stefanescu (use of an ensemble approach to compute background error statistics for a limited-area domain), Alex Deckmyn (definition in a wavelet space), and Rachida El Ouaraini (from a f-plane to a β-plane approximation in 3D-Var). See the four dedicated papers.

Jk

The introduction of an additional cost-function in 3D-Var, to relax the LAM analysis towards the coupling large-scale one, is described in the PhD report of Vincent Guidard.

Jo

Marian Jurasek adapted and tuned the IFS variational quality control (VarQC) of observations for ARPEGE (first step before ALADIN). The first experiments with a quadratic VarQC show a slight positive impact.

Next

A work plan for implementing an operational 3D-Var assimilation for ALADIN France was drawn by Claude Fischer.

3.6.6. Observations

Work focussed on the use of satellite data : HIRS (Elisabeth Gérard), AMSU-B (Zahra Sahlaoui), "local" AMSU-A observations (Nadia Fourrié), MSG/SEVIRI (Thibault Montmerle), AIRS (Malgorzata Szczech, Thomas Auligné), Geowind and Quikscat (Christophe Payan), in the framework of ARPEGE first. Only MSG/SEVIRI data were tested in ALADIN 3D-Var. More details may be found in the last-year ALATNET report (http://www.cnrm.meteo.fr/alatnet/)

3.6.7. Tools (of many kinds !)

Siham Sbii and Jan Masek worked at improving the model-to-satellite tool : enhanced portability, especially as concerns observation handling, and use of RTTOV-6 (cheaper than the previously used Morcrette code) as radiative transfer model in the computation of brightness temperatures.

Jadwiga Woyciechowska started to clean the first part of configuration 923, i.e. the computation of spectral orography : gathering all calls to spectral transforms and having the ARPEGE and ALADIN code closer, in order to improve legibility, make maintenance easier, and allow the introduction of more sensible definitions of semi-enveloppe.

Tomislav Kovacic introduced diagnostics on physical tendencies in ARPEGE/ALADIN :
- frequency distribution of (the decimal logarithm of) absolute tendencies for wind, enthalpy, and specific humidities for the 3 water phases, for the identification of relevant warning thresholds;
- identification and storing of information on gridpoints where at least one of the tendencies is above the threshold fxed by namelist (NAMCHET), at each time-step (for use in operations);
- extraction and storing of vertical profiles (i.e. those data required for the 1D version) at the corresponding points.

This work is to be completed in Zagreb (porting to distributed memory).

Some more diagnostic tools for the radiative budget were developed by Eric Bazile.

Yong Wang started downscaling experiments from IFS re-analyses or ARPEGE to ALADIN-Vienna on MAP cases.

The PALADIN package is still regularly updated by Jean-Daniel Gril.

The ALADIN/ALATNET database was ported under Linux and My_SQL by Patricia Pottier and Jean-Daniel Gril. The porting and update of the ALADIN and ALATNET web sites started too, with the help of Ildiko Bujdoso.

3.6.8. Else

The latest developments in the PhD work of Radi Ajjaji, Margarida Belo Pereira, Karim Bergaoui, Andre Simon, Cornel Soci, Simona Stefanescu and Malgorzata Szczech-Gajewska, are described in the dedicated part of this Newsletter, or in the final ALATNET report.

Three new persons joined the French team : Gwenaëlle Hello is responsible for the ALARO-10 km prototype, Ludovic Auger will work on Var-Pack and Eric Wattrelot on radar observations.

3.7. Hungary

3.7.1. ALADIN developments

The main area of development during the second half of 2003 was the further assessment of the 3d-var data assimilation scheme used for the ALADIN/HU model version. This topic is in fact carried out in the framework of the ALATNET project, therefore the main conclusions of this work will be elaborated at the ALATNET part of this Newsletter.

The other main development areas are as follows:

The LAM EPS activities started with the precision and discussion of the working plan basically together with the French colleagues (the plan is enclosed). The first activities on the one hand were related to the adaptation of singular vector computation (601) of the ARPEGE model. It is needed in order to start the sensitivity studies of 601 with respect to the integration time and target domain. All the components of the creation of initial global perturbations are now working in Budapest as it is in Toulouse. On the other hand the first ensemble runs with the ALADIN model was carried out: the members of the ARPEGE ensemble (PEACE) were used as initial and boundary conditions of the ALADIN model. One randomly chosen case was tried and the corresponding verification and visualisation softwares were validated. All the LAM EPS related work will be detailed more in the next Newsletter.

The continuous control of the operational model version was continued and some case studies were carefully investigated. Among them the most interesting one was a winter case, when the ALADIN model strongly overestimated the minimum temperature, when other models proved to be much better. Different cures were tried from modified radiation scheme through snow analysis until new physical package, nevertheless the improvements were still moderate (more details can be read in the article of Helga Toth in the next newsletter).

The activities related to the physical parametrisation of the model were performed in the framework of the LACE physics working group with the leadership of Thomas Haiden. The subjects of investigation are radiation scheme and the adaptation of Lopez scheme.

3.7.2. ALATNET developments

The most important ALADIN-related activities at the Hungarian Meteorological Service are concentrating on the scientific topics defined in the ALATNET research plan. Our Service was active in the following ALATNET sub-topics (in parenthesis the topic number refers to the ALATNET research plan) during this semester :

- specific coupling problems (topic 6),
- use of new observations (topic 10),
- 3d-var analysis and variational applications (topic 11).

Hereafter the main activities in these subtopics will be briefly described.

Specific coupling problems

The ALATNET stay of Raluca Radu finished in Budapest with a 2-month stay at the end of the year. The results obtained in Ljubljana were discussed and finally presented at the ALATNET seminar (see more at the proceedings of the seminar and at the young researchers' presentation at this newsletter).

The work regarding the inter-comparison of explicit and DFI blending was completed in the framework of a Prague stay during the summer. The results are rather neutral, i.e. hardly any difference was found between the two techniques (an internal LACE report is available on that work).

Use of new observations

We have further continued the application of new types of data sources in the course of ALADIN 3d-var data assimilation. The tests with ATOVS (satellite) data were continued and reported at the EWGLAM and ALADIN workshops (Lisbon and Prague). The latest experiments were dealing with some modifications of the bias correction method and some other small modifications. As it was the case for the previous experiments, no significant improvement were found while using additional ATOVS data on top of the existing surface and radiosonde data.

The other important and investigated data source is the aircraft (AMDAR) data. After the preliminary work (shortly mentioned in the previous newsletter) more deep investigations were carried out. Similarly to the ATOVS data only slight improvements in some of the variables were shown with the experimentation (more details can be found in a separate article in the same newsletter).

3d-var analysis and variational applications

Our ALATNET young researcher Steluta Alexandru continued and finished her altogether 20 month stay in Budapest. Her subject was "*Scientific strategy for the implementation of a 3d-var data assimilation scheme for a double nested limited area model*". The latest results of Steluta can be found as a separate paper in the newsletter.

The parallel testing of the 3d-var scheme continued in the second half of 2003. The conclusions of such comparison were rather the same as before : the performance of the model with 3d-var scheme is rather neutral with respect to the dynamical adaptation one.

Some further case studies were conducted : now some winter cases were chosen (inversion and heavy snow fall event). The 3d-var version was behaving rather similarly than that of the reference for these cases (more details are in the paper of Steluta Alexandru in this newsletter).

3.7.3. LAMEPS experiments with ALADIN at HMS for 2004-2005

The main objective of LAMEPS experiments with ALADIN is the better understanding and prediction of local extreme events like phenomena causing heavy precipitation, windstorms etc. on short time range. At the moment this work would primarily address scientific questions without considering operational feasibility.

The main questions to be answered are as follows

- What is the sensitivity (in terms of target domain and integration time) of a global singular vector (SV) computation for a Central-European EPS application?
- Are the PEACE provided initial and boundary conditions convenient for the local EPS run?
- What are the perspectives to compute singular vectors with ALADIN and generate

perturbations for the local EPS run?

- What are the boundary conditions to be used for an ensemble system with local perturbations (ARPEGE or PEACE members)?

Three major experiments will be carried out in the near future :

1. Investigation of the impact of target domain and target time window of the global singular vector computation to ALADIN EPS.

Typical values are : some Central European target domain (size to be specified by preliminary testings), 12 to 24 hours integration time. These global (singular vector) runs will be performed at HMS.
Deadline: March 2004

2. Running of ALADIN EPS coupled with quasi-operational PEACE ensemble members.

The PEACE system is run at Meteo-France once a day to create a 10 member global ensemble forecast. The coupling domain for the production of coupling files from the Meteo-France PEACE must be also defined in the experiments. This setting should take into account the following :
- the "transmission coupling domain" should also provide other possible ALADIN members with ensemble coupling data
- the size and required intermediate storage space in Toulouse
Deadline: September 2004

3. Running of ALADIN EPS based on ALADIN native SV perturbations.

ALADIN 601 will be run on a low resolution large size domain (part of the Atlantic and Europe, possibly with a resolution not much better than 50 to 100 km). Then a forecast will be performed on this large domain (coupled to ARPEGE) and then the outputs of this forecasts will be served as initial and boundary conditions for the EPS running on a smaller domain with higher resolution. In case of obtaining insufficient spread an ensemble of LBC (based on PEACE) should be used. Other perturbation mechanisms could be considered in later steps (bred modes, perturbed observations).
Deadline: March 2005

The tasks above altogether require the definition of four domains :

- The target domain for the PEACE singular vector computation (and sensitivity experiments for the time window simultaneously).
- Common post-processing (output) domain for the PEACE results (for all the ALADIN partners for transmission purposes; "transmission coupling domain")
- Large-sized low resolution "singular vector domain": where native singular vectors to be computed (if any).
- Smaller-sized higher resolution local EPS domain for providing the LAMEPS forecasts.

Other issues for setting up a research LAMEPS system at HMS :

- High cost in both memory and computational time (ALADIN 601, PEACE)
- Development of research tools: evaluation and installation some classical EPS performance and visualisation products.
- Note: in case of computing global singular vectors locally in a routine way, it should be checked with ECMWF.

3.8. Morocco

3.9. Poland

3.10. Portugal

See the report on operations (improvement of verification package, *margarida.belo@meteo.pt*)

3.11. Romania

3.12. Slovakia

3.13. Slovenia

3.14. Tunisia

4. PHD THESES

4.1. Introduction

A new doctor since September 30th : Filip Vana !

Research topic : **The dynamical and physical control of kinetic energy spectra in a NWP spectral semi-Lagrangian model**

For more details about the work of the ALATNET PhD students, the presentations at the last ALATNET seminar and the ALATNET final report are available (see Part 1).

4.2. Radi AJJAJI : Incrementality deficiency in ARPEGE 4d-var assimilation scheme

The scheduled paper for this Newsletter is still under revision. A paper to be submitted to QJRMS is under preparation.

4.3. Steluta ALEXANDRU : Scientific strategy for the implementation of a 3D-Var data assimilation scheme for a double-nested limited-area model

Introduction

The research work on the 3D-Var data assimilation scheme for ALADIN/Hungary model was continued with another two case studies. Since the first three cases happened in the summertime, it was decided to choose also some "winter" cases, one when the operational model failed to predict a strong temperature inversion, and another one, when the model forecasted a snow storm.

The 3D-Var scheme for the ALADIN model consists in minimizing a cost function, which is a sum of two terms, the background term and the observational term. For the background error covariances, the standard NMC statistics (Parrish and Derber, 1992) are used, and for the observational term, only the surface (SYNOP) and upper-air (TEMP) data are considered in the assimilation cycle.

The best framework for the 3D-Var experiments has been established during the general evaluation of the scheme (Alexandru, 2003; Alexandru, 2002). Thus the experiments for these two case studies were carried out using the operational lateral boundary conditions (LBC) for ALADIN/Hungary, from the ARPEGE model, and the SYNOP and TEMP data available at the Hungarian Meteorological Service (HMS). From SYNOP observations, geopotential measurements are considered, and from TEMP data, wind, geopotential, temperature and relative humidity information. The time-consistency coupling technique is chosen both in cycling and in production. This technique means that the lateral boundary data are coming from the same integration of the coupling model, thus the corresponding large-scale information is consistent in time. The initial condition is the 3D-Var analysis, and the first two LBCs are coming from the coupling model (ARPEGE). Digital filter initialization (DFI) is applied at the beginning of the integration, both in cycling and in production.

The "inversion" case (the 27th of December 2002)

The forecasters proposed a case when a strong temperature inversion occurred in the north-western part of Hungary, on the 27th of December 2002, at noon. In the previous days to this situation, an anticyclone extended over Hungary, so the air was very cold and dry. A warm advection from the Mediterranean Sea replaced the cold air in altitude, causing a temperature contrast around 8^{o}C between the 980 hPa and 850 hPa levels.

More than 12 hours before the event, both models (in dynamical adaptation and using the 3D-Var scheme) provided a similar forecast, predicting a temperature inversion, in that region, but not so strong (around 4^{o}C). The main difference appeared in the analysis from 27.12 12 UTC run, when

the operational model was very close to reality.

Figure 1 represents two cross-sections of the temperature field through Hungary (from west to east, through Budapest), from the analysis of the operational model and of the model using 3D-Var scheme. One can see that the operational analysis showed a strong inversion. Also the model with 3D-Var scheme predicted an inversion, only that the variation of the temperature in altitude is smaller.

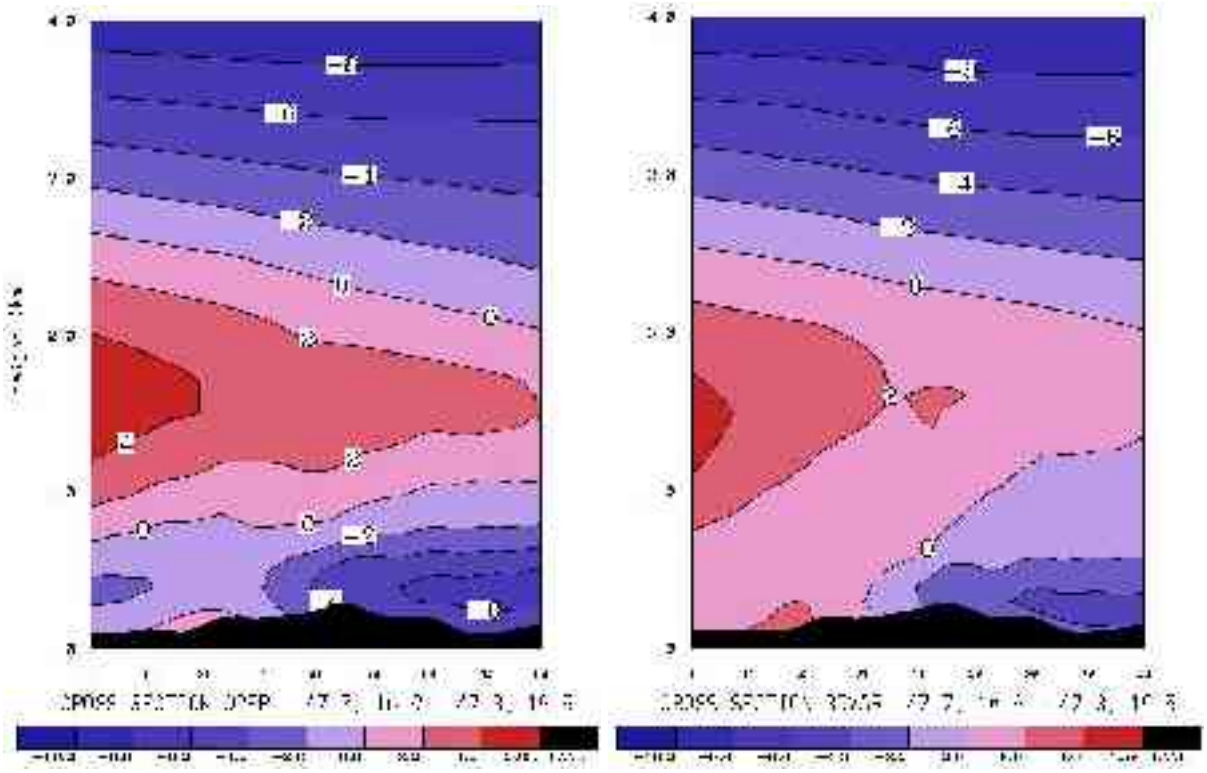


Figure 1: The vertical cross-sections of the temperature field, through Hungary, from the analysis of the operational model (left) and using 3D-Var scheme (right), from 27.12 12 UTC run

At first sight, these results appeared as being "strange", mainly because the idea of data assimilation procedure is to improve the analysis, and in this case it seems that it was getting worse. Being in the region where the inversion occurred, Budapest station has been selected in order to compare in altitude the temperature measurements (*the green line* in Fig. 2) with the temperature field from the analysis of the two models (with and without data assimilation). Figure 2 (left hand side) shows the vertical profile (from the surface till 700 hPa) of the temperature from the operational analysis (*the blue line*) and using 3D-Var scheme (*the red line*), for Budapest station. It can be seen that the operational model was closer to reality.

In order to understand the "behaviour" of the 3D-Var scheme, the temperature field from the different information sources used in the integration of the model (the first guess, the 3D-Var analysis, the lateral boundary conditions, the analysis after the digital filter initialization has been applied) was investigated. One can see in Fig. 2 (right hand side) that the temperature forecast from the first guess (*the violet line*) was the farthest from reality. The 3D-Var scheme succeeded to bring the analysis closer to observations (*the turquoise line*), but not as much as the operational analysis was. The model in dynamical adaptation had more accurate information contained in the initial condition, from the global model. One can see that the curve of the temperature from the ARPEGE analysis (*the pink line*) is very close to the real measurements. After the digital filter initialization has been applied, the analysis of the model using the 3D-Var scheme (*the red line*) was getting farther from the observations.

Other experiments have been performed in order to check the reasons for the misforecast of the model using the 3D-Var scheme. First, it was tested whether or not the information from the observations was sufficient. So the observations used for the 4D-Var cycling in the ARPEGE model (from Météo-France), which include SYNOP, AIREP, SATOB, TEMP and PILOT measurements were assimilated. The same variables were considered from SYNOP and TEMP observations as in the previous experiments, besides, from AIREP, wind and temperature, and from SATOB and PILOT, wind. The cross-section of the temperature field from this set of experiments is represented in Fig. 3 (left hand side). It can be seen that the analysis is looking almost the same as with the data from HMS, which means that the reason of the bad analysis is not the amount of observations in the region.

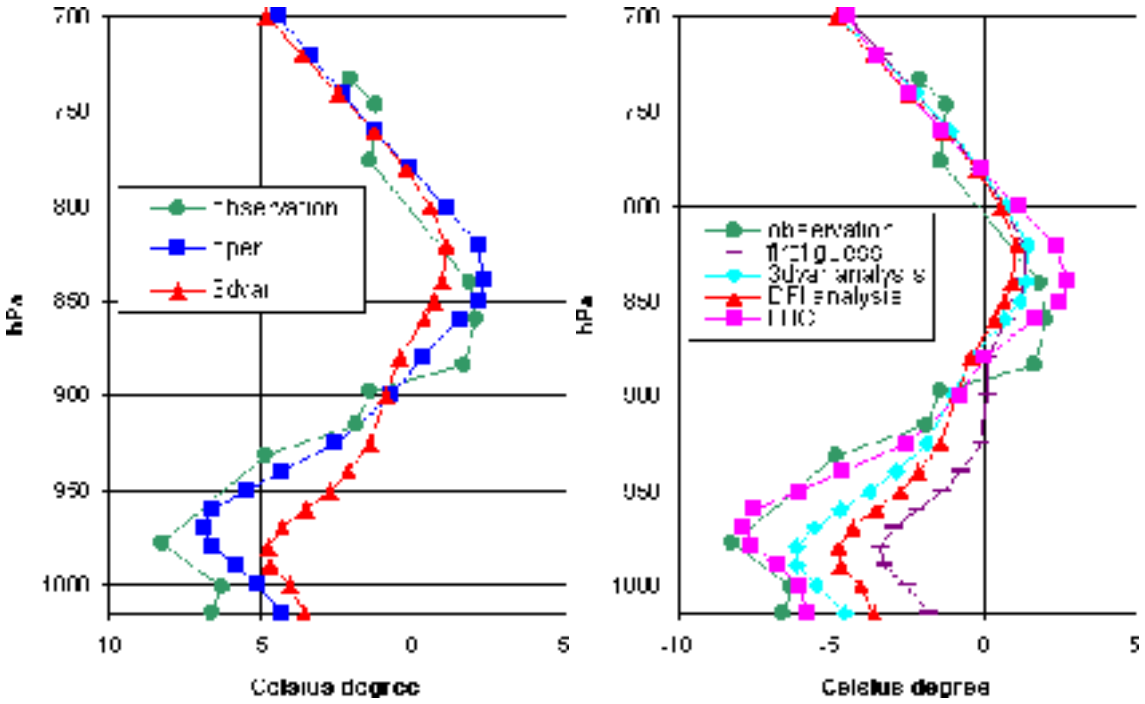


Figure 2: The vertical profile of the temperature field, for Budapest station, from the analysis of the operational model and using 3D-Var scheme (left), and from different other files (first guess, 3D-Var analysis, DFI analysis, LBC) (right), from 27.12 12 UTC run

Then, another experiment was performed, using the first guess from the ARPEGE model (the 6h forecast from the 27.12 06 UTC run), instead of that one of the ALADIN/Hungary 3D-Var system. We can see in Fig. 3 (right hand side) that the results showed an improved analysis, almost similar with the one of the operational analysis. So it seems that for this case, the global model was able to predict a better first guess than the limited-area model. This is probably due to the advantage of the 4D-Var scheme to use the observations at the "correct time".

So this case study showed that the data assimilation procedure worked properly, but due to the fact that the information from the first guess was too far from observations, the model could not obtain a better result. This means that further it should be investigated which part of the physics of the limited-area model has to be improved, in order not to loose, already after 6h integration, the good features introduced by the 3D-Var procedure. (Alexandru, [1] 2004)

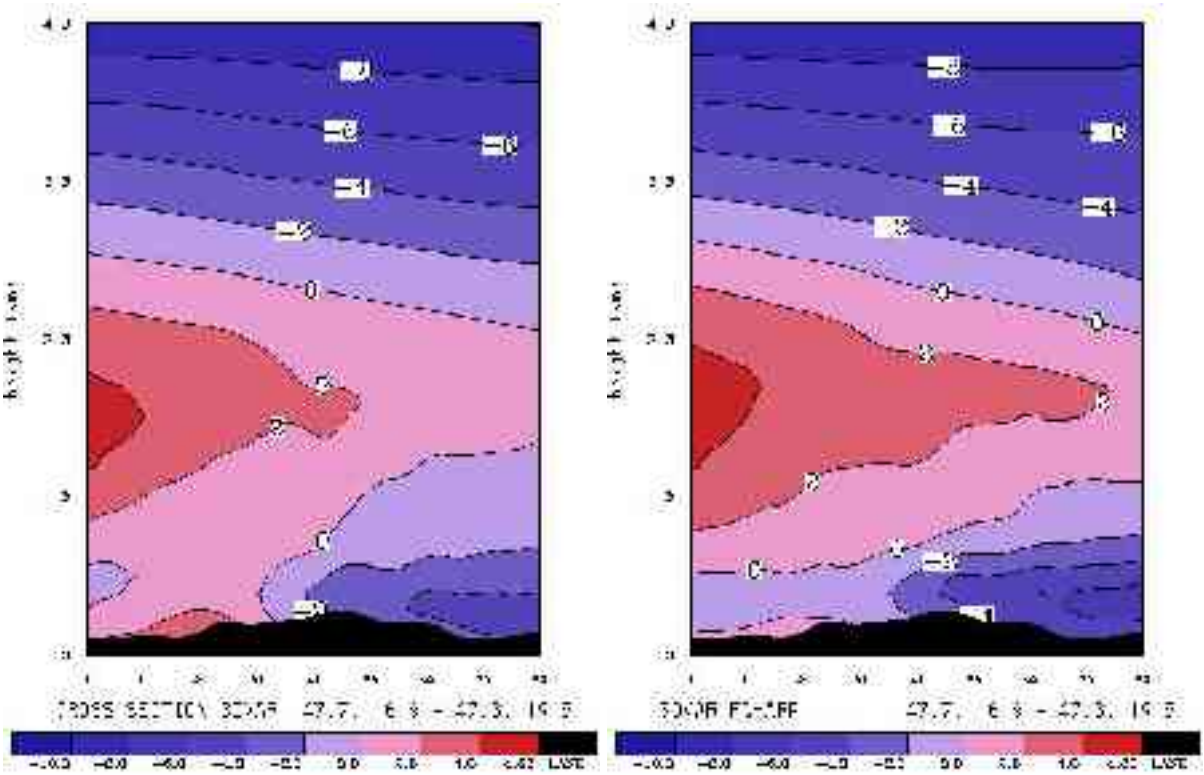


Figure 3: The vertical cross-sections of the temperature field, through Hungary, of the model using the 3D-Var scheme and data from Météo-France (left) and the first guess from the ARPEGE model (right), from 27.12.2002 12 UTC run

The "snow" case (the 7th of January 2003)

Another chosen "winter" case is from the 7th of January 2003, when a significant amount of snow fell over Hungary. The synoptic situation can be described by a high-pressure zone placed in the northern part of Europe, and by cyclones moving over the Mediterranean Sea, in the south. So Hungary was at the edge between the two weather systems. As an effect of the air with two contrasting characteristics there were sharp differences in temperature and intensive frontal zone over the Carpathian Basin, causing snowfall almost all over the country on the 7th of January (Fig. 4).

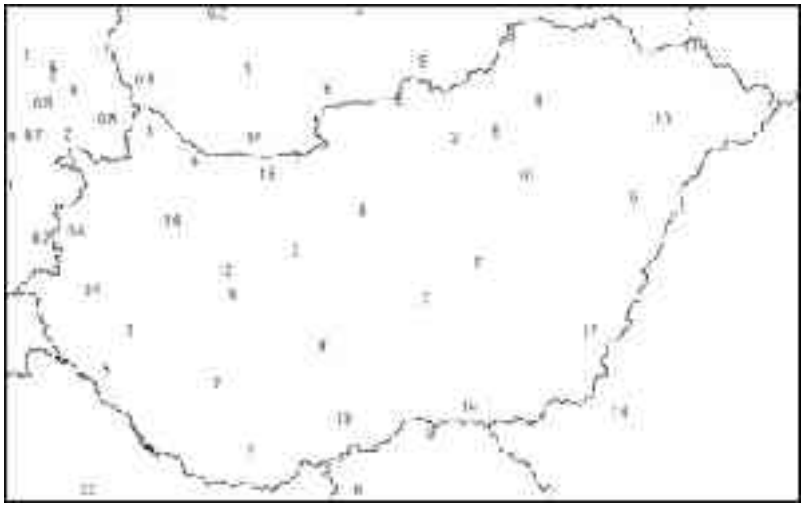


Figure 4: The quantity of precipitation (mm/12h) measured over Hungary, between 07.01 06 UTC - 18 UTC

The forecasts of different experiments, with and without data assimilation, have been examined. The area of high humidity is covering the entire country (not shown). The air has an

ascending motion, especially in the south-eastern part of Hungary (also not shown), where the quantity of precipitation is significant. Comparing with the real measurements, one can say that both models predicted the same pattern of the areas where the precipitation is expected (Fig. 5). Only that there is an overestimation of the maximum of precipitation fall in the northern part of Serbia and also the area in the north-western part of the country, predicted by the models as being covered with snow, is smaller than in reality.

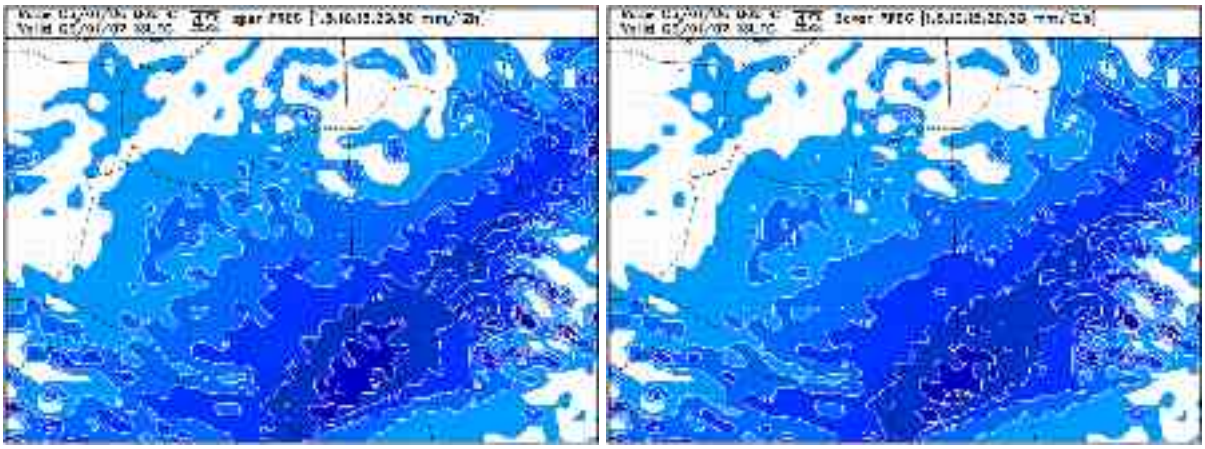


Figure 5: The quantity of precipitation (mm/12h) forecasted by the operational model (left) and using 3D-Var scheme (right), between 07.01 06 UTC - 18 UTC, from 06.01 00 UTC run

Considering the limitations of the models in defining the precise location of precipitation, these forecasts showed the possibility of the occurrence of snowfall almost over the entire country. The good prediction of the model running in dynamical adaptation is partly due to the accurate information provided by the global model through the initial and lateral boundary conditions. For the model using the 3D-Var scheme, both the information from the observations and from the lateral boundary conditions "played" an important part in the forecasting of the snowfall. (Alexandru, [2] 2004).

Conclusions.

During my ALATNET stays, the research work was concentrated on to determination of the best strategy for the application of the 3D-Var scheme for ALADIN/Hungary, a double-nested limited-area model. The experiments were realized in two parts. At the beginning, a general evaluation was carried out, in order to get a rough idea about the applicability of the ALADIN 3D-Var system. Then, some case studies have been investigated, to study the scheme in more detail.

It was concluded that the optimal choices for the 3D-Var scheme for the ALADIN/Hungary model are as follows: digital filter initialization to be applied in cycling and production, lateral boundary conditions to come from the same integration of the model (thus the information to be consistent in time), and the choice of the lateral boundary conditions, from the global or limited-area model, is not so important.

The case studies showed the good performance of the scheme, but also some limitations. So the assimilation of surface data together with the radiosoundings observations is necessary. The information from the lateral boundary conditions is also important to be accurate, in order to obtain a good short-range forecast. For some situations, the amount of applied data was proven to be too poor for providing a good forecast. So more other types of observations should be used in the future. Also the physics in the limited-area model has to be improved, especially for an accurate very short range forecast (6h).

So far, we could conclude that the present framework of the ALADIN/Hungary 3D-Var

system is appropriate to obtain a good forecast. But the scheme is still the "subject" of improvements. In the future, it is planned to try to test beside others, new background error statistics and the explicit blending technique for improving the first-guess of the scheme.

References:

Alexandru, S., [1] 2004: "A temperature inversion case study performed using the 3D-Var scheme for the ALADIN/Hungary limited area model". *ALATNET Internal Note.*

Alexandru, S., [2] 2004: "3D-Var experiments for the ALADIN/Hungary model: a case study (III)". *ALATNET Internal Note.*

Alexandru, S., 2003: "Further experiments with the 3D-VAR data assimilation scheme for ALADIN/Hungary model". *ALATNET Newsletter 6.*

Alexandru, S., 2002: "3D-VAR data assimilation experiments for the double-nested limited area model ALADIN/Hungary". *ALATNET Internal Note.*

Parrish, D., and Derber, J., 1992: "The National Meteorological Center's spectral statistical interpolation analysis system". *Mon. Weather Rev.*, 120, 1747—1763.

4.4. Margarida BELO-PEREIRA : Estimation and study of forecast error covariances using an ensemble method in a global NWP model

Nothing new (operational duties at home, see the corresponding report).

4.5. Karim BERGAOUI : Further improvement of a simplified 2d variational soil water analysis

Introduction

Actually, in the operational model ARPEGE, we are using the optimal interpolation method (OI) to analyse the soil water content by assimilating observed temperature and relative humidity data ($T_{2\mathrm{m}}$ and $HU_{2\mathrm{m}}$). The idea is to convert, with statistical regressions, the 2m forecast errors of temperature and relative humidity into soil state corrections (here for W_{p}) :

$$\Delta W_p = \alpha_p^T \Delta T_{2\mathrm{m}} + \alpha_p^{HU} \Delta HU_{2\mathrm{m}} \quad (1),$$

where α_p^T and α_p^{HU} are the statistical coefficients of the O.I.

Balsamo and al. (2004), applied a simplified variational 2D-Var method to analyse W_p . More precisely, they applied a dynamical optimal interpolation (DYNOI) scheme using the same observed 2m data, with some simplification hypotheses :
- horizontal decoupled gridpoints,
- unbiased observed and forecasted errors,
- linear observation operator **H** (with assimilation window 6h, 24h or 48h).

The idea is to calculate **H** with a finite-difference method. At each gridpoint the analysis procedure is the following. We start with a reference state (guess) : W_p^G. We add a small perturbation δW_p^{prt1} to obtain a first perturbed state W_p^{prt1}. We deduct the same quantity δW_p^{prt2} to obtain a second perturbed state W_p^{prt2} . When the observations of $T_{2\mathrm{m}}$ and $HU_{2\mathrm{m}}$ are available (each 6 hours), we calculate the differences between the perturbed runs for $T_{2\mathrm{m}}$ and $HU_{2\mathrm{m}}$. Those differences are divided by the initial humidity perturbation δW_p to get a dynamical estimation of **H** components at the gridpoints. The gain matrix is deduced : $\mathbf{K} = \mathbf{B}\,\mathbf{H}^{\mathrm{T}}\,(\mathbf{H}\,\mathbf{B}\,\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}$.

In a 1D-Var case : $\mathbf{B} = \left(\sigma^2_{W_p}\right)$ $\mathbf{R} = \begin{pmatrix} \sigma^2_{T_{2m}} & 0 \\ 0 & \sigma^2_{HU_{2m}} \end{pmatrix}$ $\mathbf{H^T} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ $\mathbf{K} = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix}$

The modified soil state (analysis) is written as :

$$W_p^A = W_p^G + k_1 \Delta T_{2\,\mathrm{m}} + k_2 \Delta HU_{2\,\mathrm{m}} \quad (2)$$

We obtain the same formulation as for O.I. (Eq. 1). The advantage is that the coefficients are calculated dynamically. The relation between observations and analysed fields better integrates the meteorological and physiographical context of each gridpoint. In DYNOI, we don't need to place a set of sophisticated regressions like in statistical O.I.

Objectives

The results obtained by Balsamo show the superiority of DYNOI (using a 6 h assimilation window) over the statistical O.I. used operationally. However, DYNOI cannot be used operationally in ARPEGE because it is very expensive in CPU time : in fact, we have to perform 4 forecasts of 6 h each to obtain the observation operator **H**.

In our study, we propose to reduce this time by using an assimilation window of 1 h only. We hope that the model can reproduce the essential shape of the physical signal. The obtained $\mathbf{H}^{1h}$ and $\mathbf{K}^{1h}$ will be multiplied by a constant (by 2 here) to find $\mathbf{H}^{6h}$ and $\mathbf{K}^{6h}$.

Experiment and results

To obtain an important effect of the initial soil moisture over the boundary layer, we choose the day of 16 June 2000 at 12 h for our experience. Then we performed correction runs of soil humidity with a variational assimilation window of 1 hour and we multiplied the components of the H operator by 2. Figure 1 shows the results.

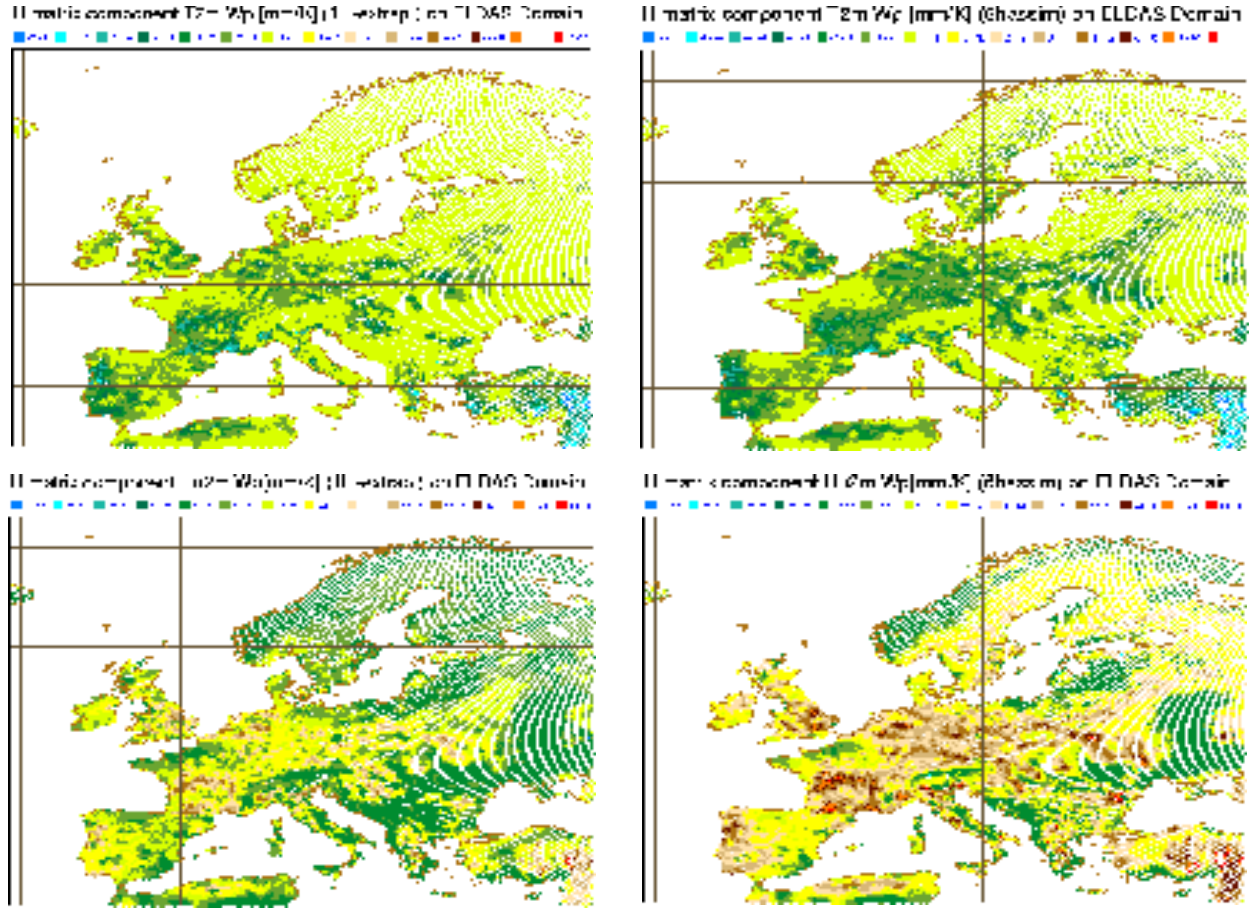


Figure 1 : $\mathbf{H}_1$ (upper) and $\mathbf{H}_2$ (lower) matrix components for the experiences of 16 June 2000 at 12h UTC (H1 is the component for T2m and H2 is the component for HU2m, for the two time-windows : 1 hour (left) and 6 hours (right)

This shows that the 1 hour time-window is adapted to our problem. It reconstitute the main characteristics of the physical signal but the multiplication by 2 is not enough to reproduce the intensity of the signal.

Conclusion and perspectives

The simplification of the variational soil water analysis in ARPEGE was considered. We assimilated real T_{2m} and HU_{2m} observations with an assimilation window of 1 hour. The component of **H** were multiplied by 2 and compared to $\mathbf{H}^{6h}$. The results were closely comparable to the analysis with 6 hour assimilation window. The shape of the signal was reproduced but not the intensity. We have to validate this method over a full period of many days and many assimilation cycles to study the effect of such an initialization on the forecast.

4.6. Vincent GUIDARD : Evaluation of assimilation cycles in a mesoscale limited-area model

Introduction of a large-scale cost-function in a LAM 3d-var

Introduction

The use of the DFI-blending procedure [Brožkova et al., 2001] to update the large scales in ALADIN using the ARPEGE analysis is commonly known in the ALADIN community. Its combination with the 3d-var assimilation scheme, called Blendvar [Siroka et al., 2001], has already been evaluated in ALADIN. The DFI-blending step reduces the discrepancies between the model state and the ARPEGE analysis, thus the discrepancies between the model state and the observations. Some drawbacks of this technique are : its empiric formulation, the impossibility of controlling the amplitude of the correction and the inconsistency between its strong-constraint formulation and the variational framework of 3d-var. Which is why the introduction of the ARPEGE analysis as a new source of information in 3d-var has been developed.

Formalism

In order to select the scales of the ARPEGE analysis (x^{AA}) to be introduced in the 3d-var, an operator H_1 is used to put x^{AA} on to a subspace (which can be a low-truncation spectral space or a low-resolution physical space). The new information vector can be written:

$$\begin{pmatrix} x^b \\ y \\ H_1(x^{AA}) \end{pmatrix} ,$$

where x^b is the background state (short-range ALADIN forecast) and y is the observation vector. The error vectors can be defined :

$$\varepsilon^b = x^b - x^t, \quad \varepsilon^o = y - H(x^t) \text{ and } \varepsilon^k = H_1(x^{AA}) - H_2(x^t) ,$$

where x^t is the true state vector in the ALADIN geometry, H is the observation operator and H_2 is an operator from the ALADIN nominal space to the previously described subspace.

In the variational formalism, the matrix of covariances between the error vectors is needed. It can be written as follows:

$$W = \begin{pmatrix} B & E(\varepsilon^b \varepsilon^{o\mathrm{T}}) & E(\varepsilon^b \varepsilon^{k\mathrm{T}}) \\ E(\varepsilon^o \varepsilon^{b\mathrm{T}}) & R & E(\varepsilon^o \varepsilon^{k\mathrm{T}}) \\ E(\varepsilon^k \varepsilon^{b\mathrm{T}}) & E(\varepsilon^k \varepsilon^{o\mathrm{T}}) & V \end{pmatrix},$$

where B, R and V are respectively the background, the observation and the “large-scale” covariances error matrices. A classical hypothesis in data assimilation is that background errors and

observation errors are not correlated, thus : $E(\varepsilon^b \varepsilon^{o\mathrm{T}})=0$ and $E(\varepsilon^o \varepsilon^{b\mathrm{T}})=0$.

If the observations used in the limited-area model (LAM) are new in comparison with the observations used in the global model, one can assume that the "large-scale" errors and the observation errors are not correlated, i.e. : $E(\varepsilon^k \varepsilon^{o\mathrm{T}})=0$ and $E(\varepsilon^o \varepsilon^{k\mathrm{T}})=0$.

The cross-covariances $E(\varepsilon^k \varepsilon^{b\mathrm{T}})$ and $E(\varepsilon^b \varepsilon^{k\mathrm{T}})$ seem impossible to simplify. Only a computation of their weight with respect to the diagonal square matrices would allow their nullification. If these quantities are negligible, adding this new source of information in the 3d-var is simply adding a new term in the cost function: $J(x)=J_b(x)+J_o(x)+J_k(x)$, where J_b and J_o are the classical terms measuring the distance to the background and to the observations, and $J_k(x)=(H_1(x^{AA})-H_2(x))^T\, V^{-1}\,(H_1(x^{AA})-H_2(x))$ is the new term, which measures the distance to the global analysis in a specific subspace.

Three kinds of 3d-var assimilation schemes are now available : an analysis using the background and the observations (classical 3d-var, called hereafter BO), an analysis using the background and the global analysis (called BK, which is an alternative to the DFI-blending technique) and an analysis using the three sources of information (called BOK, an alternative to the Blendvar technique).

Application to simplified models

This method has been implemented in a 1D academic, shallow-water, model to evaluate its abilities. Ilian Gospodinov implemented a spectral, periodic 1D version of this model [Gospodinov, 2000]. The LAM counterpart was implemented by Piet Termonia. The system follows the ARPEGE-ALADIN formulation: spectral, semi-Lagrangian, semi-implicit, Davies coupling. A gridpoint analysis system has been added to these forecast systems, which is not a variational analysis, but the optimal solution with respect to the least-square method. The three kinds of analysis are available, plus the dynamical adaptation of the global analysis (AD).
A true state is obtained thanks to a high-resolution "global" model, from which the statistics of the various errors are computed through an ensemble method.

AD versus BK

The BK analysis is closer to the global analysis than to the LAM background (which is of a rather poor quality). The differences between AD and the BK analysis are slight.

BO versus BOK

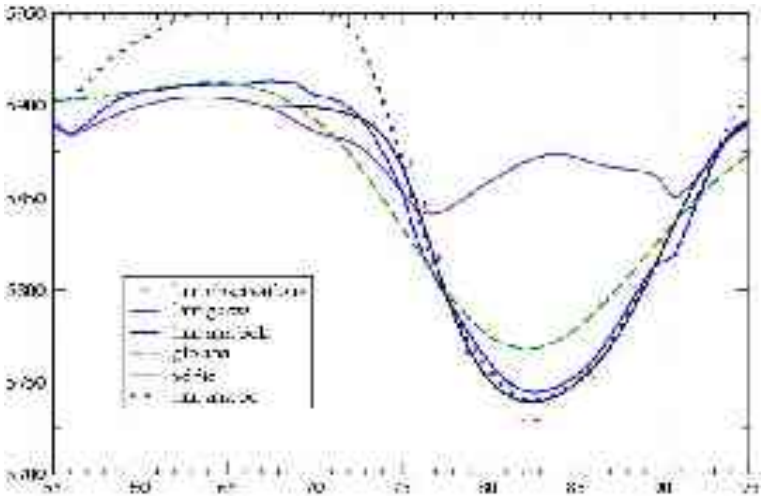


Fig. 1 : Comparison between a BO and a BOK analyses, using a band of observations.

When the observations are regularly spaced within the domain, the differences between the

BO and the BOK analyses are negligible. But it can happen that the observations to analyse are only over a part of the domain. To mimic such a case, one set of experiments using a band of observations has been performed. The well-known problem of wrapping-around increments in ALADIN is observed in the BO analyses (Fig.1). The large-scale correction smooths this effect, thanks to a new source of information present even in the area without observation.

Despite the encouraging results obtained in the particular framework of bands of observations, the other results are rather neutral. The LAM does not generate smaller scales than the global model. The LAM forecasts are less accurate than those of the global model, which shows the limits of using ALADIN-like methods in a simplified system.

Similar methods have been used in a 1D gridpoint Burger model. No new result has been obtained from this system.

Towards the implementation in ARPEGE-ALADIN

An ensemble evaluation of the various statistics has already begun in ARPEGE-ALADIN. It is based on the ensemble work in ARPEGE by Margarida Belo-Pereira and Loïk Berre, and its implications in ALADIN studied by Simona Stefanescu. The main objectives of this evaluation phase are:

- Implementation of diagnoses (from the simplified models);
- Mapping of the innovations ($d^k = H_1(x^{AA}) - H_2(x^b)$) and of the variances (σ_b^2 and σ_k^2);
- Evaluating the weight of the non-diagonal terms of the matrix *W*;
- Evaluating the possibility for *V* to be gridpoint and/or spectral;
- Evaluating the possibility for *V* to be multivariate.

References

Brožkova, R., D. Klaric, S. Ivatek-Sahdan, J.-F. Geleyn, V. Cassé, M. Siroka, G. Radnoti, M. Janousek, K. Stadlbacher and H. Seidl, 2001. DFI-blending: An alternative tool for preparation of the initial conditions for LAM. In H. Ritchie (Ed.), Research activities in atmospheric and oceanic modelling, pp.1.7-1.8.

Siroka, M., G. Bölöni, R. Brozkova, A. Dziedzic, C. Fischer, J.-F. Geleyn, A. Horanyi, W. Sadiki and C. Soci, 2001. Innovative developments for a 3D-VAR analysis in a limited area model : scale selection and blended cycling. In H. Ritchie (Ed.), Research activities in atmospheric and oceanic modelling, pp.1.53-1.54.

Gospodinov, I., 2000. La méthode semi-Lagrangienne à deux pas de temps appliquée à une modèle de prévision atmosphérique : Amélioration de la conservation et de la stabilité. Ph.D. thesis, Université Toulouse III - Paul Sabatier.

4.7. Jean-Marcel PIRIOU : Correction of compensating errors in physical packages; validation with special emphasis on cloudiness representation

No report this time.

4.8. Raluca RADU : Extensive study of the coupling problem for a high-resolution limited-area model

Introduction

It was proposed a strategy of coupling in spectral space the large-scale model with the limited-area one. The updating time is particularly an issue for those schemes running with boundary conditions provided from another centre. Most European forecast centres use Davies' type relaxation techniques in limited-area models, which seem to be responsible for the forecast failures from 25-27 December 1999 ("Christmas storm"). The improvement of the treatment of boundary conditions is, however, an increasingly important issue, as is the nesting ratio, as the resolution increases. The idea of using a coupling method based on the spectral representation of

the fields as an additional step to the operational coupling method (see previous ALATNET reports [1]) with the aim of gaining the possible missing information in the LAM, is supposed to lead to a substantial reduction of errors in forecast. The accuracy of this boundary method is studied by integrating the shallow-water one-dimensional model developed by I. Gospodinov and the ALADIN/LACE operational version. When the spectral coupling method was applied, the signal coming from the large scales is kept and solved by the limited-area model, which is very promising, but one should take care that only a suitable tuning of the proposed method will allow a correct evolution of the large scales inside the LAM.

Tuning of the spectral coupling.

Supplementary experiments for finding the best interval for relaxation were carried out and the results were very similar to the results with 1D tests. The lower and upper part of the wavenumber band, in which the relaxation is done and where the large-scale fields are mixed with the small scale fields, was defined by certain wavenumbers used in the model as namelist parameters (K0 and K1). The most suitable is to choose the parameters for relaxation in the lowest wavenumber band of the spectrum, as the energy of the cyclone is carried in this zone (see previous reports). The proper relaxation interval should be related to the wavenumber of the largest and smallest disturbance entering the domain during an operational coupling interval. A suitable result close to the reference was found when the lowest parameter was set to 2 and the upper one to 10.

The preliminary one-dimensional tests with the shallow-water model showed a positive expected impact of the spectral coupling method, but revealed at first look some difficulties:

- suddenly appearance of the cyclone into the LAM when the spectral coupling method is applied at the fixed operational time-interval, 3 h, of updating the boundary information at the same moment of the operational coupling,

- corruption represented by a dipole structure of locally growing and shrinking lows, a direct linear time-interpolation effect, produced by increasing the frequency of the spectral coupling.

The sudden appearance of the cyclone into the limited domain can be explained by the strong influence of the large-scale information provided to the LAM at the spectral coupling times. The LAM is forced at a certain fixed time, the coupling time to couple its spectral fields with the large-scale ones. The model itself will suffer shock by the introduction of large scales and the dynamical evolution of the LAM is seriously affected. This behaviour brings the idea to minimize the forcing by a smoother introduction of the large-scale information into the limited domain which permits a more realistic evolution of the cyclone inside the LAM. After one-dimensional tests were performed with the shallow-water model [2], it was underlined that the temporal relaxation is the solution for eliminate the first difficulty listed and the removal of time-interpolation by using the information from the second coupling file as a solution for the second problem. In this frame several 3D tests with ALADIN/LACE were carried out. The tuning tests were performed using ALADIN-LACE with AL25 version, LBCs every 6 h, and spectral coupling method as additional step to operational Davies' relaxation coupling. All tests were done in order to improve the "Christmas storm" case.

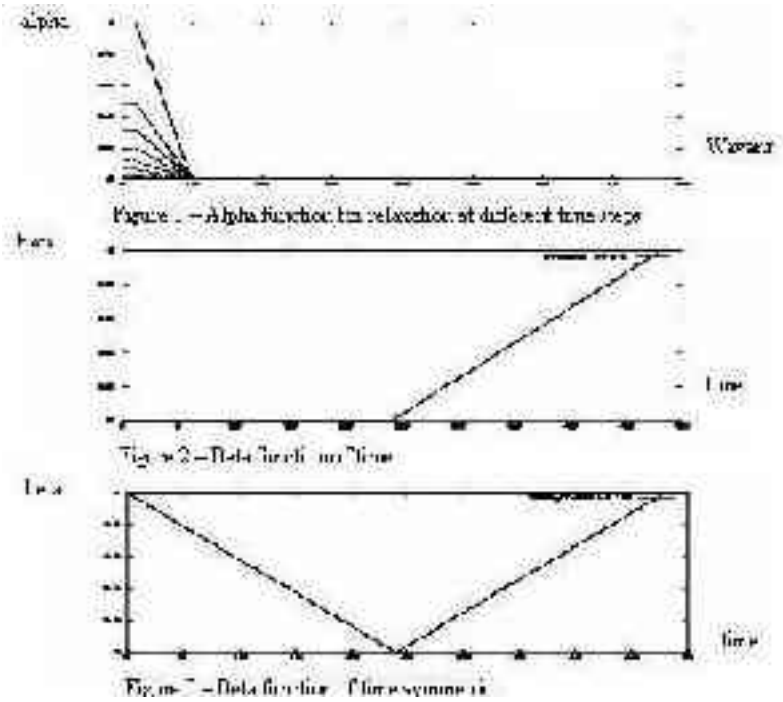
Tuning the alpha function for temporal relaxation.

In order to produce an introduction of the large-scale information very smoothly into the domain, the spectral coupling should be done is such a way that the alpha function, function depending on wavenumbers, is multiplied by a beta function depending on time. The equation used in the proposed method is very similar with the operational coupling equation, where alpha is a linear function which changes smoothly (Figure 1). It was proposed that :

$$\text{alpha}\ (m, n, t) = \text{alpha}\ (m, n) * \text{beta}\ (t),$$

(where m, n are the wavenumbers and t is the time). Evolution of alpha function during a coupling time-interval is very important in this point. Observing that the cyclone enters the limited domain after half time of the first coupling interval, the beta function was built so that at the first part of the

coupling time interval it is kept constant at the zero value in order to obtain an unaffected solution of the LAM. After the relative threshold TSTARTSC is reached, the function is growing to its maximum value at the end of the coupling time-interval, being reset to zero at the next coupling time. This means that exactly at the coupling time the LAM receives full large scales, with the next time-step the small scales provided by LAM are predominant and after the relative threshold is touched, small scales are mixed with large-scale information by decreasing the percent of the first ones and increasing the percent of the latest, see Fig. 2 (TSTARTSC = 0.5).



Optimal parameters for the relative threshold and for beta exponent.

For the relative threshold TSTARTSC, it was found out that the smaller it is, the smoother is the information entering the LAM. Additionally by introducing the beta exponent, the large-scale information is provided even more smoothly. Several tests performed indicated similar results in using different values for the relative threshold. The same conclusions were found using different exponents for beta function. By choosing an exponent with a value of 2 , we studied the impact of temporal relaxation on the MSLP field (Figure 4) : the cyclone is appearing one hour sooner inside the domain than using an exponent 4 (Figure 5). It can be seen that an earlier introduction of large-scale information with a higher exponent brings similar effect as combining bigger threshold with small exponent. It can be concluded with all this that using the same settings as in the 1D tests the most reasonable for producing a good forecast close to real situation would be the values TSTARTSC = 0.5 and EXPONENT = 4.

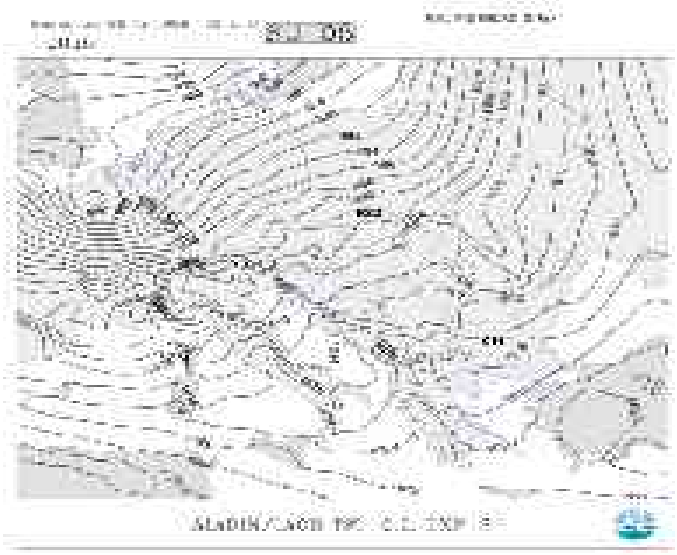


Figure 4 - Impact of spectral coupling on MSLP field by using exponent 2 of beta function.

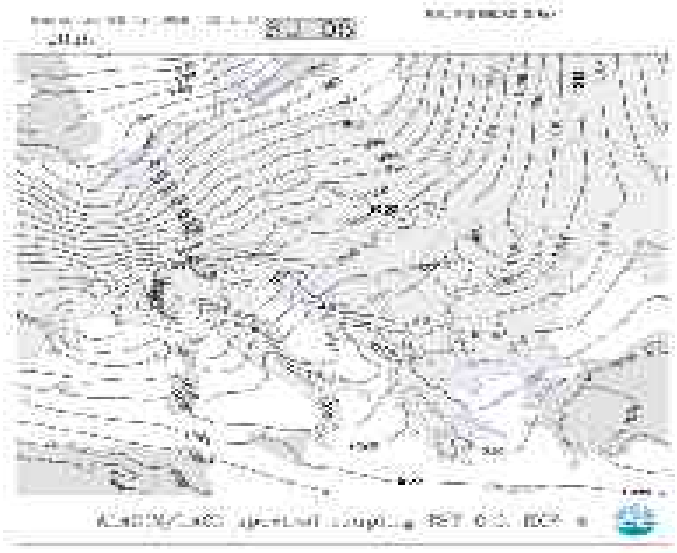


Figure 5 - Impact of spectral coupling on MSLP field by using exponent 4 of beta function.

Further experiments with ALADIN-LACE continued, with the proposal of a symmetric shape of the beta function with respect to the coupling time-interval. Interesting to see is if this does not lead to a degradation of the forecast. The beta function is decreasing from the maximum value at the coupling time to zero value until the relative threshold is reached and starts to grow until it arrives to its final shape at the end of the coupling time-interval (Figure 3). The large scales are fully considered at the coupling time, following a decreasing of their percent until the relative threshold. At this point just small scales of the LAM are taken into account and, evolving further in time, small scales are replaced by large scales until the end of the coupling time-interval. This tuning will provide the large-scale information sooner into LAM. The results using the efficient settings (TSTARTSC = 0.5. EXPONENT = 4, K0 = 2, K1 =1 0) and a symmetric function show an interesting evolution of the cyclone which seems to oscillate inside the domain. The cyclone appears after the first time-step, moves further and jumps back exactly at the operational Davies relaxation coupling moment. The symmetric function approach is far to be considered as positive result as the solution of the LAM is seriously disturbed. It can be concluded that a constant shape of function at the beginning followed by a linear smoothly growing shape to produce a temporal

relaxation in the second part seems to be the most suitable tuning for the selected case using the spectral coupling.

In our trial to observe the impact of the proposed method, several tests were performed with ALADIN operational versus ALADIN with spectral coupling on the SLOVENIA domain on a usual case with a zonal weather situation without strong fronts. At first look it can be observed that spectral coupling does not impede the forecast ability of the LAM.

Further plans.

It has to be underlined that the spectral coupling method, tested on the "Christmas storm" case, gives a promising impulse, but with some limitations of the applicability of the idea. It is needed and very necessary to continue the validation work. By studying the method on some more cases it has to be proved that the mesoscale features are not destroyed in situations in which a "Christmas storm"-like problem is not present, but when ALADIN provides extra mesoscale information with respect to the coupling model. Another point that has to be solved is to find the proper moment to use the proposed coupling method. The idea launched was that the spectral coupling should be applied just in extreme situations when the operational mode seems to be unable to give a reasonable forecast. P. Termonia effort was directed in the identification of such situations by estimating the quality of the interpolated lateral-boundary data.

It should be mentioned that among different treatment for lateral boundaries it can be studied the implication of the extension zone width in the context of transparent LBCs.

References

[1] R. Radu, ALADIN/ALATNET Newsletters 22-24/4-7

[2] M. Tudor, 2003, ALADIN Workshop on Coupling - internal note

4.9. Wafaa SADIKI : A posteriori verification of analysis and assimilation algorithms and study of the statistical properties of the adjoint solutions

Nothing new.

4.10. Andre SIMON : Study of the relationship between turbulent fluxes in deeply stable PBL situations and cyclogenetic activity

See the ALATNET final report for the most recent results.

4.11. Cornel SOCI : Sensitivity study at high resolution using a limited-area model and its adjoint for the mesoscale range

Finalization and translation in Romanian of the PhD manuscript.

4.12. Klaus STADLBACHER : Systematic qualitative evaluation of high-resolution non-hydrostatic model

Nothing new (operational duties).

4.13. Simona STEFANESCU : The modelling of the forecast error covariances for a 3D-Var data assimilation in an atmospheric limited-area model

See the joint paper. A communication was also sent to the WGNE "Blue Book".

4.14. Malgorzata SZCZECH-GAJEWSKA : Use of IASI/AIRS observations over land.

See the ALATNET final report for the most recent results.

4.15. Jozef VIVODA : Application of the predictor-corrector method to non-hydrostatic dynamics

Writing the PhD manuscript hopefully.

5. ARTICLES

5.1. Plans in 2004 for developing radar data assimilation in the ALADIN, AROME and Méso-NH models

François Bouttier & Claude Fischer *(Météo-France)*

5.1.1. Abstract

Radar reflectivity data assimilation is being developed simultaneously in the ALADIN/AROME and Méso-NH mesoscale models used at Météo-France. The priority is on reflectivity data, either at single levels or in volumic scans. Although the absolute reflectivity simulation may involve along-the-beam effects (refraction and attenuation), the analysis scheme will be a 3D-Var minimization that regards reflectivities as single-column measurements, involving a vertical convolution between model micro-physical fields and the radar beam shape. The relationship between increments of micro-physical variables and other model variables (temperature and water vapour), as well as the physical balancing between cloud and precipitation fields in the vertical column, will be implemented as a physical module inside the observation operator itself. This is because the current background error term (Jb) and the 3D-Var initialization cannot provide a satisfactory initialization of the model's cloud multivariate vertical structure. The physical module will act as a 1D column retrieval, either as an independent 1D-Var retrieval, or as a physically-based observation operator inside the 3D-Var minimization. This physical 1D retrieval will be adapted to each physics package, its evaluation may require several time-steps of the parametrizations in the case of prognostic micro-physics schemes.

This paper is an overview of the planned work on radar reflectivity data assimilation at Météo-France (CNRM, Toulouse) in 2004. The objective is to make significant technical progress in the 3D-Var assimilation of radar reflectivities for operational applications, and to provide an experimental framework for further scientific work.

5.1.2. Tasks

The project can be broken down into well-separated tasks, which are listed below, with the names of the likely involved people (they are from Météo-France or from participating ALADIN countries).

Radar data production:

This is the processing to be done at data production level, mainly data formating and production of data quality indicators. The definition of these aspects requires some iteration between data producers and data assimilation teams. The organization of data production is specific to each country. In France, this involves the teams that manage the ARAMIS operational network and the PANTHER project of upgrading this network. Contact points: J. Parent du Châtelet and P. Tabary (Météo-France, observing systems department).

Instrumental aspects and physical aspects of radar simulation:

This is the development of the radar-oriented aspects of the "observation operator", including data pre-processing that cannot be done at data production level. Contact points: O. Caumont and V. Ducrocq (mesoscale research group, Météo-France), who exchange information directly with scientists from the radar instrument community, including research-oriented radar teams inside and outside of Météo-France.

Data ingestion into the observational database (ODB) of the 3D-Var code:

This requires the definition of a runtime ODB radar data template, with appropriate selection of relevant meta-data, grouping and geometrical location of reflectivities, and expandability to future data layouts: volumic scans, polarization information, Doppler wind data. Specialists: R. Zaaboul (Météo Maroc) and S. Kertesz (Hungarian Met Service), coordination by C. Fischer (Météo-France).

Observation operator software:

This is the heavy technical part of putting the radar data into the ALADIN 3D-Var, with interpolation of ALADIN model fields, plugging in of the radar simulation model, Jo computation, tangent linear and adjoint code, logistics of the interface with physics, parallelization aspects. Specialists: M. Jurasek (Slovak Met Service) and E. Wattrelot (Météo-France), coordination by C. Fischer.

Interfacing with the model physics:

This is the part of the observation processing that interfaces the radar simulation model with the atmospheric model physics, in order to balance the micro physical information with the other model variables, locally. It will be specific to the atmospheric model used, so there are two distinct subtasks : in ARPEGE/ALADIN, specialists: D. Banciu (Romanian Met Service) and E. Bazile (Météo-France); in AROME/Méso-NH, specialist: O. Caumont. Coordination by F. Bouttier and V. Ducrocq.

Data assimilation experiments:

In Méso-NH: managed by V. Ducrocq with her team; in ALADIN/AROME: E. Wattrelot and ALADIN partners

To be defined once the ALADIN 3D-Var can assimilate radar data.

An additional package may be needed for the development of the background term Jb and the physical initialization in ALADIN/AROME, depending on the outcome of preliminary 3D-Var experiments.

5.1.3. Basic choices

Since radar data assimilation is a rather new area in NWP, there are many open scientific and technical issues, so we need to make choices in order to start the development. Some of them are arbitrary and based on intuition (mainly experience with satellite data over the past 20 years), others are based on recent scientific data. Of course they will be revised in the light of future experience.

The emphasis is on operational data assimilation using real operational radars in 2006-2012. Some radars are more advanced than others. In order to have a sufficient coverage of the forecast domains, one needs to focus on the most widely available radar data types, which means that some features of the most advanced ones may not be used in the first version of the data assimilation: polarization, volumic scan, Doppler information, etc. Still, newly developed software shall be designed in order to be upgradeable to future radar technologies, in operational networks as well as in research (e.g. field experiments). But research radars will not be specifically considered in this project; they may be implemented as options to the operational code using outside manpower.

Doppler wind information is potentially very valuable. It is almost disconnected from the assimilation of reflectivities, which, today, are more widely available than Doppler winds. Furthermore, there is already some foreign activity on Doppler winds which will be used as a reference later. For planning reasons, it is decided to give priority to the assimilation of reflectivities, and to postpone the development of the assimilation of Doppler winds until 2006, by that time we should be in a position to assimilate them within about one year.

Operational assimilation means that the whole radar processing system must be robust with respect to any problem that can occur at the instrumental level or in the assimilation software: unusable data (because it is wrong or it cannot be meaningfully used in the assimilation) must be automatically rejected, otherwise it could compromise the whole NWP system. The NWP system must be able to cope with any weather situation, including stable boundary layers. (see section 4)

There are strong technical advantages in presenting the radar data as pixels that are horizontally local. If there is a need to perform interactive computations along the beam path, this has to be supported by good reasons i.e. do we need to represent the impact of the analysis on the

simulation of perturbations of reflectivities ? It does not preclude the use of the 3D model fields to simulate background reflectivity values. Only very sensitive, non-local atmosphere-dependent influences are an issue. Two have been identified (see section 5) :

Beam refraction:

The beam shape may be sensitive to the vertical stability of the atmosphere near the radar, leading to downward refraction (abnormal propagation). This is a common phenomenon when the boundary layer is very stable.

Attenuation by precipitation:

Strong precipitation along a beam will cause an apparent decrease in the reflectivity further away from the radar.

The difficulty is to decide if these effects should be handled by the data producers (who have access to extra instrumental data to diagnose spurious signals) or at the data assimilation stage (where the model background fields, or independent e.g. satellite data can be used). Unfortunately, models may not provide sufficient information about the boundary-layer stratification or the presence of strong rain along the beam. The involved error budget needs to be investigated. There are two approaches to affected data: either correct it (which assumes we can compute a precise correction factor), or reject it from the assimilation (which is usually done in similar cases with satellite radiances).

Basic choices of the data assimilation in ALADIN and AROME system mean that 3D-Var will be used at an horizontal resolution between 2 and 10 km, a vertical resolution of the order of 300 m. 3D-Var may use the data at a resolution of one minute, but only one datum from the same radar every hour or so. The effective radar pixel width is currently 1 km, with a vertical beam extent of up to 800 m or so. Re-interpolation of polar reflectivities into pixels on a grid clearly introduces horizontal error correlations, but it is probably not a problem at the considered resolutions (pixels will be sub-sampled at the actual 3D-Var resolution). In the horizontal, assuming a solution will be found to the along-the-beam effects mentioned in the previous section, a 1-km pixel will be handled as a dimensionless measurement in the horizontal. Along the vertical, the vertical beam width cannot be neglected, so the observation operator will involve a vertical averaging function.

The radar simulation software will use fields of precipitating rain, snow and graupel, interpolated at the relevant latitude/longitude, vertically weighted using the computed beam shape, for each available beam at this location (within a reasonable time-interval).

In its initial version, the 3D-Var control variable will not include condensed water; it may be introduced later if there are good reasons to do so. But it is not believed that assimilating condensed water only is a viable strategy for two reasons :

1. It would not make physical sense to attempt a correction of condensed water species in the models without a consistent correction of temperature and water vapour: otherwise the model integration could not sustain a coherent cloud structure.

2. Radars often observe precipitation below the clouds, meaning that the relevant model correction will not be at the observed level, but somewhere in the air-column above it.

For the same reasons, it is not believed that radar data can be assimilated without taking into account the complete vertical profile of condensed species, temperature and humidity (and perhaps wind convergence) in the troposphere. Therefore, a specific, physics-dependent, software module needs to be developed in order to spatialise the observed information above and below the beam level. In theory, the spatialisation could be provided by the 3D model of background error covariances, but this is a highly nonlinear non-Gaussian problem, and there are few scientific ideas on how to build cloud-structure information into the 3D Jb. It sounds like a very difficult mathematical problem. Until there are advances in this area, a 1D column approach will be used. The idea is that a specific module will convert single-level (or multi-level, in the case of volumic scan) information on precipitating species into column retrievals of temperature, humidity, and

possibly other parameters. (see section 7)

The advocated retrieval technique is a 1D-Var approach that will include calls to the parametrizations of cloud micro-physics (Méso-NH or ARPEGE/ALADIN condensation), and appropriate Jo and Jb terms in order to compute atmospheric profiles that match the observed reflectivities as well as possible. Several 1D time-steps are possible in order to put some time-dimension in the 3D-Var approach (i.e. a "3D and a half"-Var). Alternatively, some simplified micro-physics packages may be considered. This framework is consistent with previous experiences with precipitation assimilation at ECMWF and CMC (J.-F. Mahfouf and L. Fillion), and with soil moisture assimilation (F. Bouyssel and G. Balsamo at CNRM, and U. Callies et al at DWD). See section 8.

In a first stage, the 1D-Var radar retrieval will be computed off-line (before the 3D-Var minimization), in order to gain experience with it. Then it may be converted into an observation operator for interactive reflectivity assimilation, or into a cheaper statistical retrieval if the retrieved increments prove to have a simple structure.

The following sections provide details on the work to be done in each package, using the above arguments as justification.

5.1.4. Radar data production

(This section is mainly written as an expression of requirements for the French radar network. Other ALADIN countries have different networks, probably each with their own corresponding issues.)

Data will be processed from each radar individually (no composites, but it is permitted to use neighbouring data to help in the pre-processing). The measured data should be as pure as possible, except for corrections, which can only be done either at instrument level (e.g. electronic corrections, variance-based removal of ground echoes), or which are completely documented in the transmitted data (e.g. static attenuation factor). Ideally it should be full-resolution polar reflectivities for each radar site and elevation (every hour for operations, and at the maximum frequency for research), but some compromising with the data producers are acceptable if it proves to be too expensive in terms of telecommunications.

The same applies to neighbouring countries (Spain, Great Britain and Italy in particular), but we will have to use their composites until the radar-data exchange practices improve.

The highest priority will be given in assimilating the single-level C- and S-band reflectivities since it is available now in the LUNAIRS archive, and since the development work has already started. There is nevertheless a high interest in volumic-scan data and polarimeter's information for these radars. Sample data-files will be needed quickly: from the time they are first available, it will take take 12 to 18 months of R/D work to make these data-types usable in NWP systems.

Volumic-scan data assimilation does not require significant work to implement on top of single-level data assimilation. Indeed it is probably easier, because the vertical distribution information will facilitate the retrieval of cloud profiles from reflectivity profiles. It is recommended that experimental volumic-scan data should be available to the data assimilation teams as soon as possible, so that its usefulness can be demonstrated in experiments. It is important to proceed so in order to request future funding for more volumic-scan radars.

The usefulness of dual polarization information for data assimilation is not clear yet. At the time of writing, the main expected advantage is better radar calibration and characterization of the time of hydro-meteors. This is clearly desirable in an absolute sense, but more information is needed to decide whether polarization is more or less important (and expensive ?) than volumic scan and Doppler capability.

Doppler wind data is potentially very important for NWP, because :

(1) there is little competing wind information with this level of coverage,

(2) it is algorithmically easy to implement in 3D/4D-Var despite ambiguities in the measurement,

(3) its use is not limited to precipitating areas.

Moreover, there is proof of its usefulness in the scientific literature, with abundant experimentation in the USA and pre-operational testing in the HIRLAM system. The plan for AROME is to assimilate raw, high-quality Doppler line-of-sight wind components into 3D-Var (not retrieved wind vectors such as VAD profiles). The problem is that the corresponding observation operator is going to be quite independent from the reflectivity one. We have not enough available manpower to develop both simultaneously. Since radar Doppler coverage is going to be very sparse for another year or two , compared to reflectivity data coverage, it has been decided to postpone the work on Doppler wind until 2006. In the meantime, if Doppler winds prove to be valuable in other NWP centres, higher priority will be given to the work in this field.

In 2004, the work in Météo-France will concentrate on the operational radar data, with some limited manpower on more advanced data in the research teams (O. Caumont and V. Ducrocq).

5.1.5. Physical radar simulation

This is a highly specialized and complex scientific field. The aim is to use model information to simulate as well as possible the observed radar measurements. This is one topic of the Ph.D work of O. Caumont. Most of the work consists in compiling the existing literature and specialist advices, extracting only the aspects relevant to :

1. the types of radar and data that will actually be considered in operations (C and X-band reflectivities), and

2. the class of models that will be used : mesoscale models - ALADIN, AROME and Méso-NH- at horizontal resolutions between 1 and 10 km, using simplified micro-physics (ARPEGE diagnostic micro-physics or Méso-NH prognostic one).

Some scientific aspects that are being studied by O. Caumont are : (1) radar antenna properties, (2) beam propagation and shape, including "anomalous propagation", (3) gate function, (4) attenuation, (5). signal interaction with cloud species, including Rayleigh and Mie diffusion. In addition, masking effects shall be considered. On all these topics, useful data and software from other labs will be used as much as possible.

The objective is to develop a radar simulation software that includes all the processes that are important for an accurate simulation of most reflectivity data, plus (ideally) indications on how to spot reflectivity data that cannot be accurately simulated. At the time of writing, most effects can be satisfactorily managed, either by direct physical simulation (with further work necessary on how to compute Mie diffusion cheaply, since it seems more accurate than the Rayleigh diffusion hypothesis), or by using information from the data producers (e.g. on the time-averaged part of the attenuation). O. Caumont has identified two troublesome effects:

Anomalous propagation:

Although beam refraction by the atmosphere is fairly easy to predict in most weather situations, if the atmosphere is stably stratified close to the radar station (which is not a rare event in the morning) the refraction can be very strong and presumably difficult to predict, given the weaknesses in the models' representation of boundary layers. Basically there are three situations from the point of view of data assimilation: *(a)* weak refraction that can be computed statically using the model background, *(b)* very strong refraction where "anaprop" occurs i.e. the beam is reflected towards the ground which makes the data unusable, *(c)* strong refraction where the data is not worthless but the beam height is difficult to predict using the model (and presumably different if one considers the atmospheric background or analysis). Cases *(a)* and *(b)* are easy to handle as long as they can be safely identified at the quality-control level. Case *(c)* is trickier because we do not know whether it is worth trying to use this data or not. One needs to study how often case *(c)* occurs and whether the beam height computed by the model background is acceptable or not: are the uncertainties on the model refractivity cancelling or not the usefulness of the observation ?

Rain-induced attenuation:

Strong precipitation events along the beam will cause reflectivities beyond these events to appear

underestimated. The intensity of the involved precipitation must be accurately known in order to correct for the attenuation. It seems unlikely that the model background will provide acceptable estimates of the precipitation location and intensity. The model analysis may be better, which would lead to a non-local observation operator (the interpretation of a reflectivity pixel will change during the minimization). Or the information can perhaps be inferred from the measurement itself: if the rain closest to the radar is correctly observed, then the implied attenuation can be computed. This last approach seems to be the best, because, though it requires non-local computations, it can be applied at pre-processing stage once before the 3D-Var minimization, perhaps using only radar data along each beam. More scientific information is needed to confirm whether this is true. Anyway, it is clear that pixels that are likely to be attenuated must have an inflated observation error variance in 3D-Var, i.e. less weight in the analysis.

Both effects have potentially important consequences for the technical development of the observation operator. Although we assume here that they do not imply a non horizontally local observation operator in 3D-Var, this shall be checked as soon as possible. In the vertical, any weighting function is possible in order to match the beam shape : the model vertical resolution will be in the 100-400 m range in the mid-troposphere, and radar beams are known to have up to 1000 m vertical extension within the used ranges. The adjoint of these weighting processes will be used in the observation operator, so that reflectivities are actually used as vertical averages.

The portability of the radar-simulation code is a special problem since the concerned team works with the Méso-NH model. How can we ensure that this work can be reused in ALADIN/AROME ? It should be possible if some software modularity rules are applied:

1. The radar pre-processing, outside the model, and the purely 0D part of the reflectivity simulation are not specific to the model, so they should be written as independent software (modules or subroutines) that are easy to compile outside Méso-NH.

2. The interfacing of the local reflectivity computation with the local model fields is physics- and grid-dependent, so the software will be model-specific (the 3D interpolations in particular), this shall be designed as a clearly identifiable interface : when going to ALADIN, only the interface shall need to be rewritten.

3. The physics-dependent part of the interface is going to be similar in AROME since the same micro-physics will be used. Thus, the grid-dependent part of the interface (i.e. geometrical aspects) should be clearly separated from the physics-dependent part (i.e. local micro-physics aspects).

4. The observation data structure already has a precise, complex layout in ALADIN (ODB); in research mode this can be ignored, it will have to be rewritten completely for ALADIN/AROME.

The last question is : how does the assimilation aspects (the observation operator) interfere with the scientific design of the simulation operator ? Ideally there should be no interference, the simulation operator should not have to worry about how 3D-Var will use it. In practice, there is the technical constraint of parallelization in 3D-Var, which is a strong motivation for making early choices about the geometrical aspect of the simulation operator, and which model variables it will use as input. And there is the scheduling constraint that 3D-Var requires stable versions of the simulation operator (say, one update per year) so that they both can be developed and improved in parallel. Apart from that, the simulation operator design should only worry about getting the most accurate reflectivity computation at a reasonable computer price.

5.1.6. Observational database

This is a very technical part of 3D-Var. For efficiency purposes all observation-related data in 3D-Var must be presented in ODB, which is a relational database. In ODB, the data is organized in a hierarchical way, including meta-data, location of the data, the measurement, its expected error characteristics, information on the observing instrument, and pre-processing flags. It is important to

define the data layout early, some meta-data can be added later, but the table structure must clearly identify information that is shared between several measurements, or individual to each pixel.

Examples of a hierarchical organization:

1. data about the observing radar: frequency, name, location, error level, correction factor, blacklisting information...

2. data about each beam: elevation, polarization, angle and time, occurrence of anomalous propagation, range, distance of mask...

3. data about each pixel: height, latitude/longitude, vertical weight function...

4. datum information: type (reflectivity, polarity, Doppler wind component), quality control flags, observation error variance and bias...

For parallel optimization purposes, data transposition can be used, e.g. to present all pixels data relevant to a given (lat-lon) column for the interfacing with the physics, as opposed to presenting them beam by beam (which is better for computations related to the beam propagation). In the ODB, these would be two different views on the same data.

The ODB needs interfacing with the outside world : ingest software to feed it from the existing databases (operational, or research field measurements, or simulated data), pre-processing software to apply before 3D-Var (blacklisting, insertion of quality control meta-data and masking information), plotting software for input and output ("feedback") data, feedback statistics on output from 3D-Var, i.e. statistics of obs-background and obs-analysis, amounts of rejected data, all stratified according to radar station, elevation, region, type of weather, time of day, etc... The plotting aspect is very important in order to help finding bugs visually.

5.1.7. Variational observation operator

This part is not particularly complex, but it is sensitive to the choices made in the other sections. The observation operator works in two modes, direct observation operator (which is scientifically identical to the radar simulation operator), and linearized operator. The key of the radar assimilation is in the design of the linearisation, because this is what determines how radar observations are converted into model state corrections by the 3D-Var minimization. In summary (refer to data assimilation courses for more explanation) this conversion works in several stages:

Jo computation

The departure between model-simulated and observed reflectivities determines the amplitude of the correction, which tends to produce an analysed reflectivity that is a weighted average between background-simulated and observed values. The weight is determined by the assumed background and observation errors.

From observed parameter to observation-operator input

The reflectivity departure determines a correction to the local input of the observation operator i.e. local micro-physical species.

From observation-operator input to model control variables

The micro-physical species need to be converted into parameters that are handled by 3D-Var and the Jb term. It is assumed that we primarily need a conversion from micro-physics increments to corrections of the vertical profile of temperature and vapour humidity (and perhaps divergent wind) at the lat-lon of the measurement. This conversion can be based on physical computations inside the column (e.g. simplified cloud-scheme for perturbations of micro-physics), or on a statistical method (linear regression, nonlinear methods, profile lookup).

From local model correction to spatialized increment

3D corrections of the model fields are built from 1D column increments using multivariate smoothing functions implied by the Jb term of 3D-Var. Jb is not specific to any observation, it depicts the likely structures of errors in the model forecast used as a background. It is expressed in terms of the 3D-Var control variables. Currently the ALADIN 3D-Var does not correct micro-

physical species, but it would not be difficult to add them : the real problem is to define a Jb that will ensure physical consistency between them and the other variables. This explains why little priority is given in putting micro-physics into the control variable, whereas the emphasis is put on the local variable conversion (justification: a model forecast is more affected by a change in air temperature and humidity than by a correction of the micro-physics fields distribution)

All these operations are actually done using operators linearized around a reference state (normally, the model background, re-linearisation can be done if some operators are very nonlinear, but it is numerically better to redefine the manipulated quantities so that the resulting mathematical operators are as linear as possible). The conversion stages are actually done in an adjoint way, using sensitivity of the output of each simulation stage with respect to its input. The Jb part of the spatialisation is implicitly done by the 3D-Var minimization.

Technically, the important part is the interpolation of the model fields to the observed columns, their conversion into observed reflectivities using both 1D conversion to micro-physics and 0D conversion to reflectivities, and the linearisation and adjointing of all these operations. Everything needs to be computationally efficient, fitting the existing 3D-Var software organization. It is believed that the 1D conversion is the one component that is sensitive to the model physics, so one version will be written for each parametrization package.

In development stage, the 1D conversion can be coded as a preliminary 1D-Var retrieval before doing the 3D-Var minimization, i.e. reflectivities would be converted into 1D retrievals of temperature and humidity, which would then be assimilated into 3D-Var. This will facilitate the study of the observation operator independently from the Jb aspects.

5.1.8. Physics interfacing

A physics-based approach is believed to be the best long-term strategy for interpreting the reflectivity data into corrections of the model state. The main reason is, when a radar observes rain, we want not only to put rain into the model, but also the cloud that produces the rain. The process that links cloud with precipitation is a complex nonlinear non-instantaneous one (except in ALADIN), that can be simulated using physics. The main hypothesis here is that this process is mostly local to each model column, and can be simulated using a few time-steps of the physics.

Actually, the relevant process is the link between *perturbations* of cloud variables and *perturbations* of observable precipitation, which we hope is simpler to approximate and linearize than the prediction of the absolute fields. The link between perturbations can be calculated in at least three ways:

1. coding the tangent-linear of the model physics (possible but tedious, and a tricky linearisation of discontinuous functions is required), or

2. coding simplified physics for the perturbations and linearizing them (already partly done in the IFS/ARPEGE 4D-Var), or

3. computing the link statistically using ensembles of 1D physics computations (computationally expensive but easy to implement if we know which distribution of perturbations to use; already used with some success for soil moisture assimilation).

It seems that the best way is to start with (3) in research mode, then to investigate more efficient incremental physics along (2) if CPU cost proves to be a problem.

Another issue is the horizontal interpolation of model fields. If cloud fields have complex horizontal variations at the scale of the model grid, then averaging several model columns at the lat-lon of a radar pixel may not result in a physically meaningful atmospheric column (e.g. sub saturated with condensed water). It may be necessary to consider the closest model column instead. Also, since several radar pixels may be available for use in each model column, the pixel selection algorithm may require some thinking. ECMWF has developed a new version of the observation operators which minimizes horizontal interpolation of model fields for satellite radiance computation ("marriage of observations and physics"), and it is desirable to use it for radars too.

In the ARPEGE/ALADIN physics, the time dimension is not a problem, since cloud and precipitation variables are diagnostic ones, the physics computes an equilibration of all processes at each time-step. The challenge consists "only" in finding an atmospheric profile that will produce the correct instantaneous micro-physical fields to match the reflectivities. E. Bazile has started studying the problem in the framework of the large-scale precipitation scheme, which is easier to manipulate than the convection scheme (but the latter will have to be considered later unless we can discriminate between observations of large-scale and convective precipitation).

In the Méso-NH/AROME physics, the fall-time of precipitation is taken into account, it can be quite long (40 minutes, i.e. dozens of time-steps of AROME). It seems that an efficient way of equilibrating air and precipitation parameters needs to be developed, otherwise we will need to call the parametrization many times at each observation point. Perhaps there is some advantage in considering neighbouring model columns in the computation : they may provide cheap indications of alternate micro-physics responses to a similar atmosphere.

5.1.9. Data assimilation experiments

With such a complex data-stream, validation will have to be careful. There are two basic questions to solve: (1) can we correct the instantaneous model state to match the radar data, and (2) how does the correction evolve during the subsequent model forecast. As for any observation operator, there is a mandatory sequence of checks to carry out:

1. check the simulated data is consistent with the observations, except for differences related to errors in the model background.

2. monitor large samples of simulated vs. observed data, and refined the quality control to remove all data that are wrong or poorly simulated.

3. use the monitoring to derive parametrizations of the observation error statistics, using e.g. ensembles of forecasts and ad-hoc mathematical methods (B. Chapnik). A bias correction scheme may be needed, too.

4. check the 3D-Var gradient with one single observation of each relevant type, in several weather situations; check the linearisation of the observation operator around the background, for perturbations that could occur in data assimilation (e.g. using ensembles of forecasts).

5. check the 3D-Var increment with one single observation for each of these cases. Does it look physical ? Is the increment retained during a forecast ?

6. check the 3D-Var increment with one set of observations; in particular, check the column retrieval in the presence of single-level and volumic-scan data.

7. run a data assimilation experiment: Are there systematic biases in the increments ? e.g. if only rainy pixels are assimilated, we will moisten the model in an unrealistic way. Do the backgrounds get closer to the data when it is assimilated ? Do they get further from the other data types (e.g. synops and radiances), which would a bad sign ?

8. run one single analysis and a forecast: Is the forecast improved ? Start again with several cycles of assimilation before the same forecast.

9. At which ranges are the forecasts improved or degraded ?

10. run observing-systems experiments to determine the relative weights of all the observing systems, and to assess the usefulness of different types of radar data (for future funding decisions).

11. refine the Jb and initialization algorithms; recalibrate the assimilation statistics using the radar data.

This will hopefully demonstrate that radar reflectivities does carry benefit in data assimilation.

5.1.10. Tentative schedule

Radar data production :

Already available as single-level ("PPI") imagery. Samples of volumic data is available on limited periods, more samples will be produced during the summer of 2004, along with Doppler data samples. Volumic and Doppler data is expected in real-time on specific French radar from 2005 onwards.

Physical radar simulation :

Some modules are already working in Méso-NH. The remaining work consists mainly in the detection and/or the treatment of masks, "anaprop", attenuation. The RSM German simulation software is being studied as well. The Méso-NH simulation software should be nearly complete around June 2004, ready for plugging into ALADIN/AROME.

Observational Database :

The data structures and codes are being specified in liaison with ECMWF. The specification should be complete early in April, at which time the ODB development work will start, with completion of a first working version in June, at which time more effort will be needed on the pre-processing and monitoring tools of the data.

Variational observation operators:

It will be complete early in summer 2004.

Physics interfacing :

Preliminary studies have started in February 2004, but a satisfactorily working solution is only expected early in autumn 2004.

Data assimilation experiments:

The work can start in July, first real data assimilations should be working around November 2004. Operational radar reflectivity data assimilation will probably be implemented in a first, simple form into the ALADIN 3D-Var data assimilation in winter 2005, followed by substantial improvements during 2005 and 2006 (implementation of volumic and Doppler data, better use of the AROME prognostic cloud physics in the cloud/precipitation retrieval).

5.2. (An)Isotropy of background error structure functions

Bölöni Gergely *(HMS)*, Loïk Berre *(MF)*, Claude Fischer *(MF)*

5.2.1. Introduction

In ALADIN 3d-var, background error structure functions are defined as horizontally homogeneous and isotropic in physical space. This means, that horizontal covariances are supposed to be invariant to an horizontal translation or rotation, they depend only on the horizontal separation distance (Berre, 2000). However, due to the spectral representation of background errors, some anisotropy is implemented into the structure functions, that we illustrate on Fig. 1a. The picture shows a simple graphical check of the isotropy, namely the resulting analysis increment (i.e. observation minus background field) of a single-observation experiment. The increment isolines are not fully spherical, especially at a larger distance from the observation's location, which obviously indicates anisotropy. Our short writing is dealing with one possible reason for the deviation of structure functions from the full isotropy and describes the experiment we have performed in order to get rid of the problem.

5.2.2. Investigating anisotropy

(a) Discretization of the isotropy assumption

First of all, one should think over, how the homogeneity and especially the isotropy assumption is applied to the spectral background error covariances in ALADIN. The covariance computations between the (*m*,*n*) spectral coefficients of the background errors are greatly simplified because of the homogeneity assumption. For example, for a given variable *x* the covariance is computed as:

$$cov(x_{m,n}, x_{m',n'}) = \delta_m^{m'} \delta_n^{n'} cov(x_{m,n}, x_{m,n}) = \sigma^2(x_{m,n})$$

where δ_m^n is the Kronecker delta. It means, that only the covariances between the same (*m, n*) pairs are considered (Berre, 2000). The isotropy assumption is applied then by an averaging of the $\sigma^2(x_{m,n})$ spectral variances over the (*m, n*) pairs corresponding to the same $k^*_{m,n}$ total wavenumber, that is:

$$k^*_{m,n} = N \sqrt{\left(\frac{m}{M}\right)^2 + \left(\frac{n}{N}\right)^2} \quad (1)$$

where *M* and *N* are the maximum wavenumbers in X and Y directions. As a consequence, the covariances are not any more dependent on *m* and *n* separately, which ensures the invariance to the horizontal direction in physical space. The problem here is that the different (*m, n*) pairs never correspond exactly to the same real $k^*_{m,n}$ value, so within the averaging a discretization of the real $k^*_{m,n}$ total wavenumber and the real $\sigma^2(x_{m,n})$ variance spectrum is introduced as follows next. Let's denote by k^*_i the nearest integer to the real $k^*_{m,n}$ total wavenumber $(k^*_i = 0, 1, \ldots, N)$. The average mentioned here above determines the isotropic spectral variance spectrum as a function of the k^*_i integer total wavenumber :

$$\sigma^2(x_{k^*_i}) = 1/J \sum_{m=0}^{M} \sum_{n=0}^{N} \sigma^2(x_{m,n}) \quad (2)$$

for which (*m, n*) satisfies :

$$k^*_{m,n} = k^*_i \pm \epsilon \quad (3)$$

In (2), *J* is the number of (*m, n*) pairs for which (3) is true and ϵ is a constant real number specifying the half interval for which the isotropic average above is done. As a consequence the **B** matrix finally consists of a set of isotropic spectral variances representing horizontal scales

corresponding to the k_i^* integer total wavenumber.

(b) The normalization

On the other hand, in the J_b part of the 3d-var code, the change of variable

$$\chi_{m,n} = \mathbf{B}^{-1/2}\,\delta x$$

(Fischer, 2002) is done for the usual spectral coefficients corresponding to (*m*, *n*) wave pairs and not to the integer total wavenumbers k_i^* . It is clear then, that an inevitable estimation of the real $\sigma^2(x_{m,n})$ variances should be done from the available $\sigma^2(x_{k_i^*})$ isotropic variances, in order to do the

$$\chi_{m,n} = \delta x_{m,n} / \hat{\sigma}(x_{m,n}) \quad (4)$$

normalization, where $\hat{\sigma}(x_{m,n})$ is the square-root of $\hat{\sigma}^2(x_{m,n})$, which is itself the estimation of the real $\sigma^2(x_{m,n})$. The presently used estimation (all cycles up to now) is very simple, namely

$$\hat{\sigma}^2(x_{m,n}) = \sigma^2(x_{k_i^*}) \quad (5)$$

where k_i^* is the nearest integer to the given $k_{m,n}^*$.

(c) Rectangular domains

An other thing to be considered is how the normalization above acts in case of a rectangular domain. If the domain is rectangular, that is $M \neq N$, then :

$$k_{m,n}^* \neq k_{n,m}^* \quad \text{whenever} \quad m \neq n \qquad m = 0,1,\ldots,M \qquad n = 0,1,\ldots,N \ .$$

Substituting into (1), one can easily see that $k_{0,1}^* = 1$ and $k_{1,0}^* = N/M$ for example, which means that if we have only one wave in X or Y directions, we have got total wavenumbers that differ by the N/M ratio. On the other hand, using the estimation (5) in (4), we will normalize both $\delta x_{m,n}$ and $\delta x_{n,m}$ with the same $\sigma^2(x_{k_i^*})$ variance, because $k_{m,n}^*$ and $k_{n,m}^*$ are corresponding to the same k_i^* value for not too elongated domains ($N/M < 3/2$ at least). Our statement was that this is an approximation of the isotropy assumption, because we assume x to have the same variance on both of the horizontal scales corresponding to (*m*, *n*) and (*n*, *m*), however the exact isotropy would rather mean that only two wave pairs that correspond exactly to the same $k_{m,n}^*$ have exactly the same variance.

5.2.3. The experiment

Our proposal was to try to make the approximation of the isotropy assumption more realistic by applying the following estimation of $\hat{\sigma}^2(x_{m,n})$ instead of (5):

$$\hat{\sigma}^2(x_{m,n}) = \sigma^2(x_{k_i^*})\frac{k_{i+1}^* - k_{m,n}^*}{k_{i+1}^* - k_i^*} + \sigma^2(x_{k_{i+1}^*})\frac{k_{m,n}^* - k_i^*}{k_{i+1}^* - k_i^*} \quad (6)$$

which is a linear interpolation between the two neighbouring $\sigma^2(x_{k_i^*})$ variances of the given $\sigma^2(x_{m,n})$. Note, that taking into account that $k_{i+1}^* - k_i^* = 1$, the interpolation above can be written in a simpler form:

$$\hat{\sigma}^2(x_{m,n}) = \sigma^2(x_{k_i^*})(k_{i+1}^* - k_{m,n}^*) + \sigma^2(x_{k_{i+1}^*})(k_{m,n}^* - k_i^*)$$

On Fig. 1b one can see the result of the same single observation experiment as shown on Fig. 1a, but using the interpolation above instead of the original nearest integer estimation. It is noticeable, that the increment isolines are closer to be spherical at a larger distance from the observation, in comparison with Fig. 1a. However the shape of the increments is not visually changed near the observation point. The same features were observed making comparison of vertical cross-sections. In order to be able to make a more reliable comparison of the isotropic properties of the analysis increments, we have prepared a simple tool that measures quantitatively

the degree of anisotropy. We have defined the measure of the anisotropy by the absolute value of the ratio $\delta x_{d\mathrm{X}} / \delta x_{d\mathrm{Y}}$, where $\delta x_{d\mathrm{X}}$ and $\delta x_{d\mathrm{Y}}$ are the analysis increment values, at the same d horizontal distance from the observation point, in X and Y directions (Fig. 2). The result of this diagnostic, comparing the (5) and (6) estimations is shown on Fig. 3. The two most obvious things that one can see on the plot, are that the degree of anisotropy is increasing with the distance for both experiments and that for a given distance the degree of anisotropy is smaller for the experiment using the linear interpolation of the $\sigma^2(x_{k_i^*})$ variances than for the one using the nearest integer estimation.

As a summary we can say, that this more realistic approximation of the isotropy assumption could decrease the unwanted anisotropy but in a quite moderate extent.

5.2.4. Figures

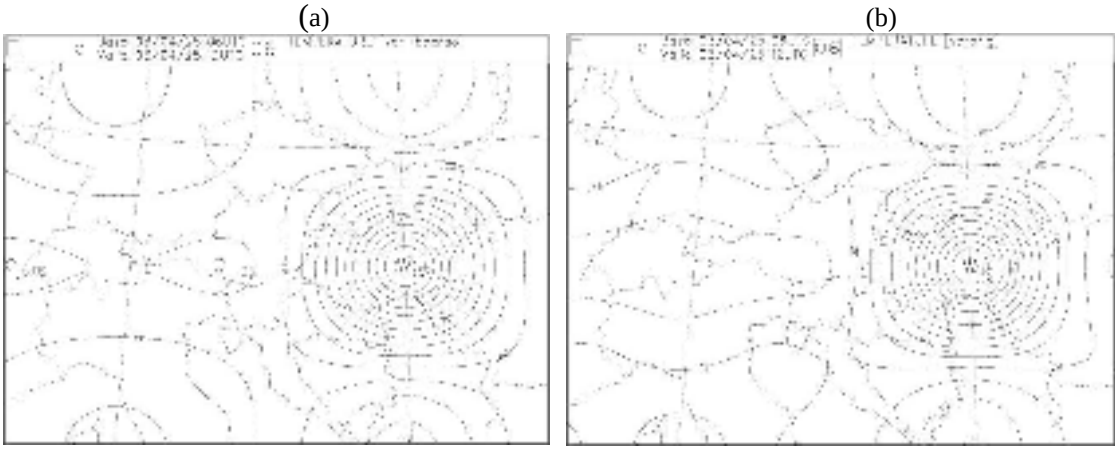


Figure 1: 3d-var analysis increment fields on model level 16, due to a temperature single observation at 500 hPa.
(a) using the nearest-integer estimation (5), (b): using the interpolation (6).
The domain is the former Hungarian operational domain with 200 points in X and 144 points in Y directions including the extension zone.

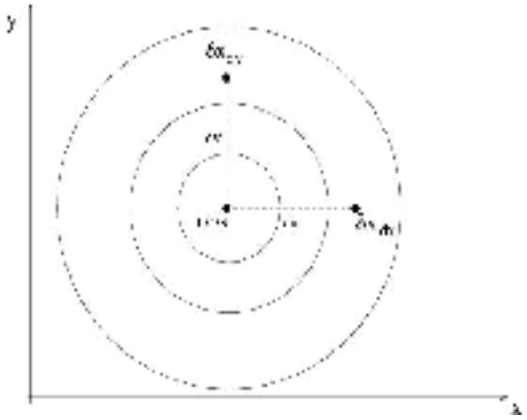


Fig. 2 : Definition of a measure of anisotropy

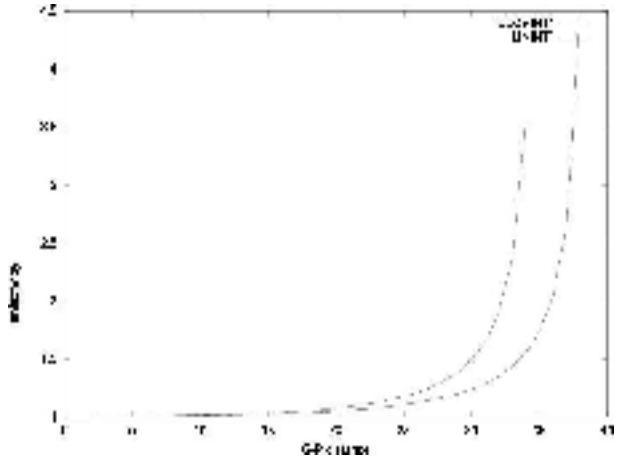


Fig. 3 : Degree of anisotropy for the two formulations : CLOSINT, nearest-integer estimation (5) and LININT, linear interpolation (6)

5.2.5. References:

Berre, L., 2000 : Estimation of Synoptic and Mesoscale Forecast Error Covariances in a Limited-Area Model. *Mon. Wea. Rev.*, 128, 644-667

Fischer, C., 2002 : The variational computation inside ARPEGE/ALADIN cycle CY25T1. *ALADIN Technical documentation*, 59 pp.

5.3. A model intercomparison for heavy precipitation with special focus on the flood event 2002 in Austria.

Alexander Kann *(ZAMG)*

5.3.1. Introduction

This article tries to evaluate the performance of NWP models for heavy precipitation events and gives special regard to the flood event of August 2002 in Austria. For hydrological purposes the computation of areal precipitation means for defined catchment areas is more useful than the use of the direct model output on single gridpoints. ZAMG operationally analyses and predicts areal precipitation amounts for 26 regions in Austria and adjacent surroundings. Observations from TAWES stations, KLIMA stations and stations from hydrological networks are archived as well as the model output from ALADIN-Vienna. In case of the flood event 2002, the dataset of the German limited-area model LM and of the global ECMWF model are used to estimate model uncertainties, too.

5.3.2. Verification of precipitation time series

In order to quantify the error of areal precipitation forecasts, it is necessary to define specific areas (Figure 1). With regard to the flood event 2002, areas 9 to 13 in the provinces of Upper and Lower Austria were particularly affected by heavy precipitation.

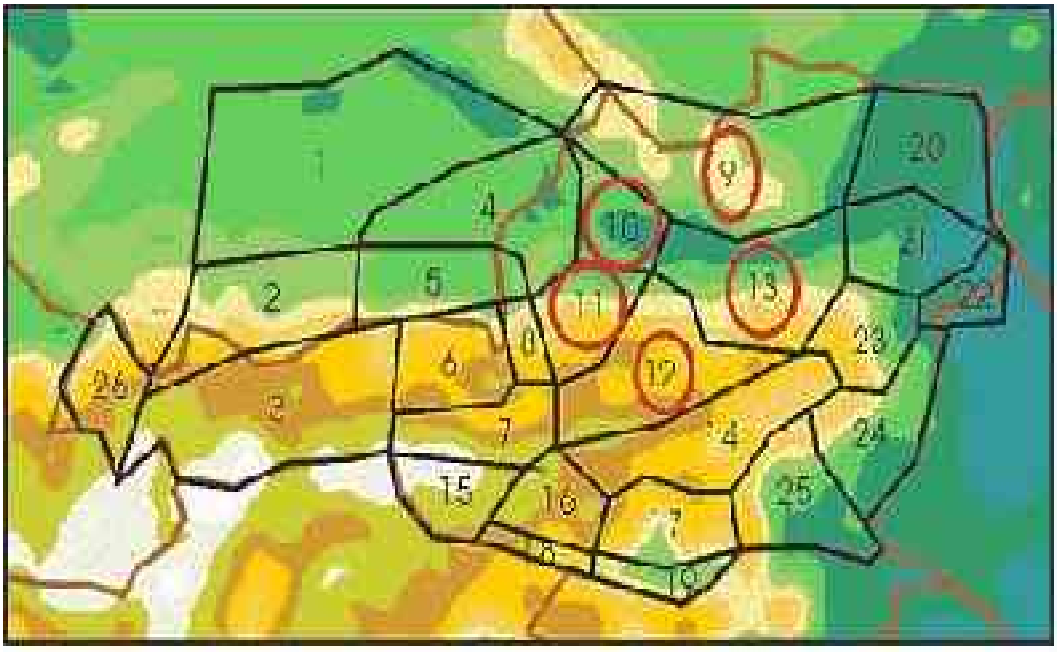


Figure 1: Definition of catchment-type areas for operational precipitation analyses and forecasts.
Areas marked in red were heavily affected by precipitation in August 2002 and chosen for more detailed studies.

Figures 2a and 2b give an overview of observed precipitation amounts during the August 2002 flood event, based on TAWES observations.

During the first part of the event, the centre of precipitation was located in the area of Mühl- and Waldviertel (area **9** in Figure 1). During the second part, these areas were hit again, but this time the more classical heavy precipitation regions along the northern Alpine rim were affected as well.

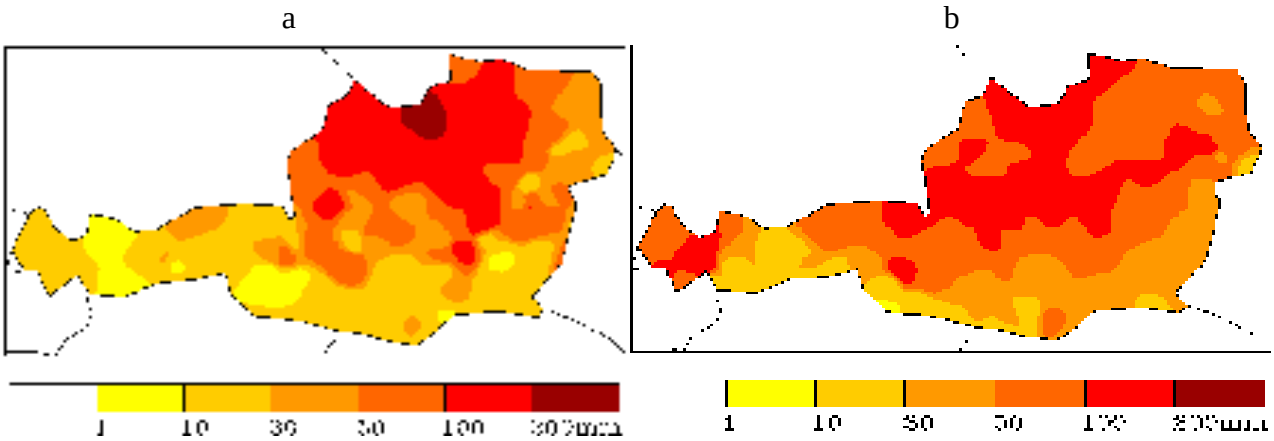


Figure 2 : TAWES precipitation sum, interpolated on a regular 10 x 10km grid. Left (a) : first part of the event (6.8.2002 12UTC - 8.8.2002 12UTC). Right (b) : second part of the event (11.8.2002 12UTC - 13.8.2002 12UTC).

Figure 3 shows time series of forecasted and observed precipitation rates for the August 2002 event. During the first part of the flood event (7.-8.8.2002, top row) maximum intensities and total rainfall amounts were underestimated by both LAMs, ALADIN and LM. The rapid increase of rainfall intensity shortly after onset, and the bimodal temporal structure of the whole first part of the event were not captured. Note that the second peak within the first part of the event was not forecasted at all. Almost as important as the prediction of the onset of heavy rainfall is the prediction of its end. The upper right panel in Figure 3 shows that the end of the rainfall episode was more or less satisfactorily forecasted, with ALADIN giving a somewhat better indication of the actual ending than LM.

The second event (bottom row of Figure 3), which was associated with a much larger low-pressure system than the first one, shows generally better model results. ALADIN was able to forecast the hourly maximum at 11.8.2002 18UTC almost at the correct time, and with roughly the correct intensity. Other predicted intensity maxima, however, show little correspondence with observed intensity peaks. Again, the end of the episode is predicted more realistically by ALADIN than LM.

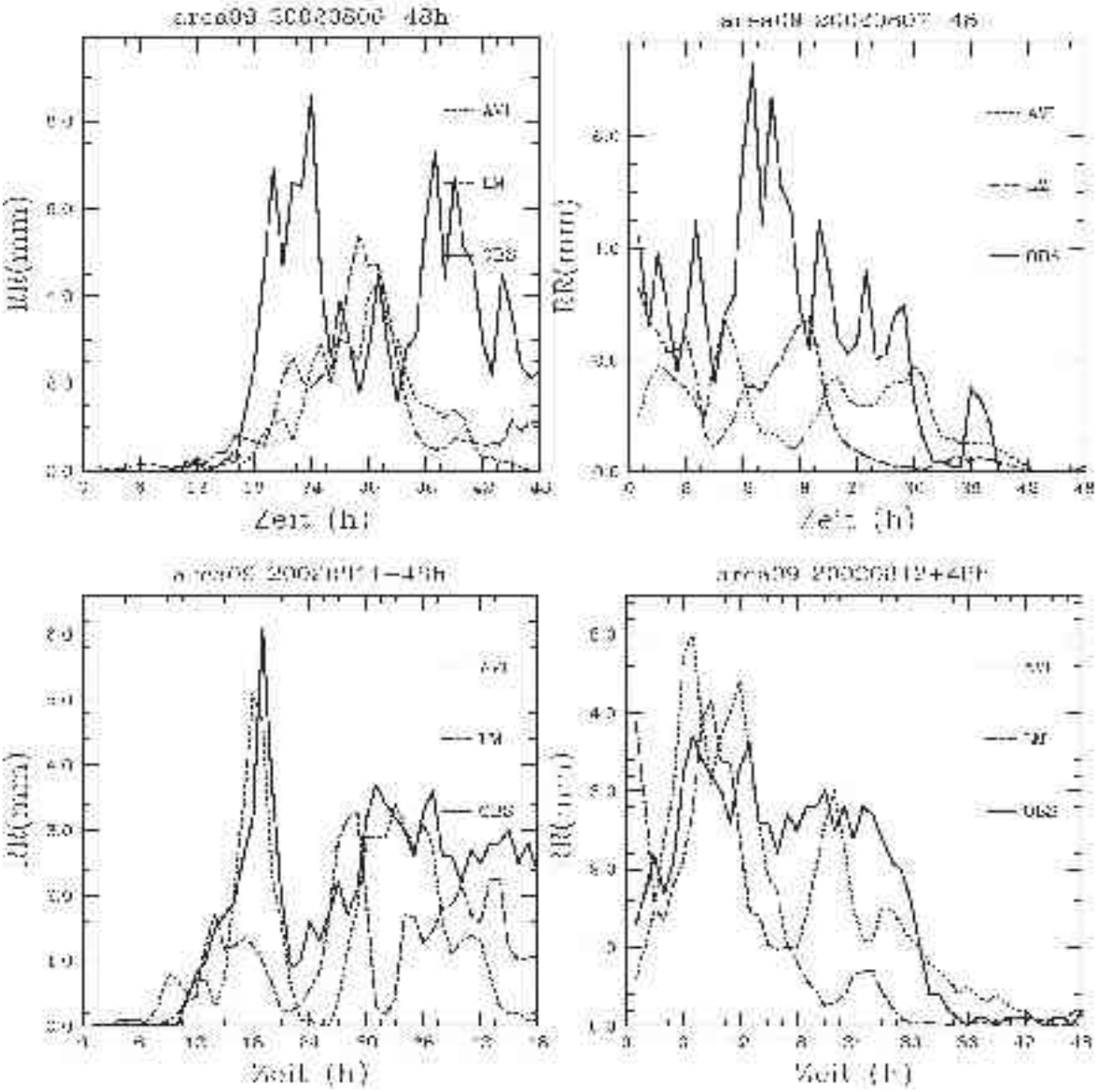


Figure 3: Hourly intensity of spatially averaged precipitation for area 9 (Mühl-, Waldviertel).
Solid line gives observations (OBS), dotted line the ALADIN-Vienna forecast (AVI), and dashed-dotted line the forecast of the Lokal Modell of DWD. Upper row shows results for the first flood event, for analysis times 6 and 7 August 2002, 00 UTC. Bottom row shows results for the second part of the event, for analysis times 11 and 12 August 2002, 00 UTC.

Regarding cumulative rainfall amounts instead of intensity diagrams, there is a general underestimation of the 48 hour precipitation total, especially on the 7th and 8th of August where both models end up with roughly 50% of the observed precipitation. The second part of the event was forecasted more accurately, and errors after 48 h amount to about 10-30% for ALADIN and 30-50% for LM. Generally, heavy precipitation is more difficult to predict in lowland regions than in mountainous areas, especially along the Alpine rim. This is because the blocking effect of the topography on the airflow introduces a deterministic element into the precipitation formation process. This tendency for reduced rainfall forecast errors is illustrated in Figure 4 which gives results for the catchment area Traisen (area **13** in Figure 1).

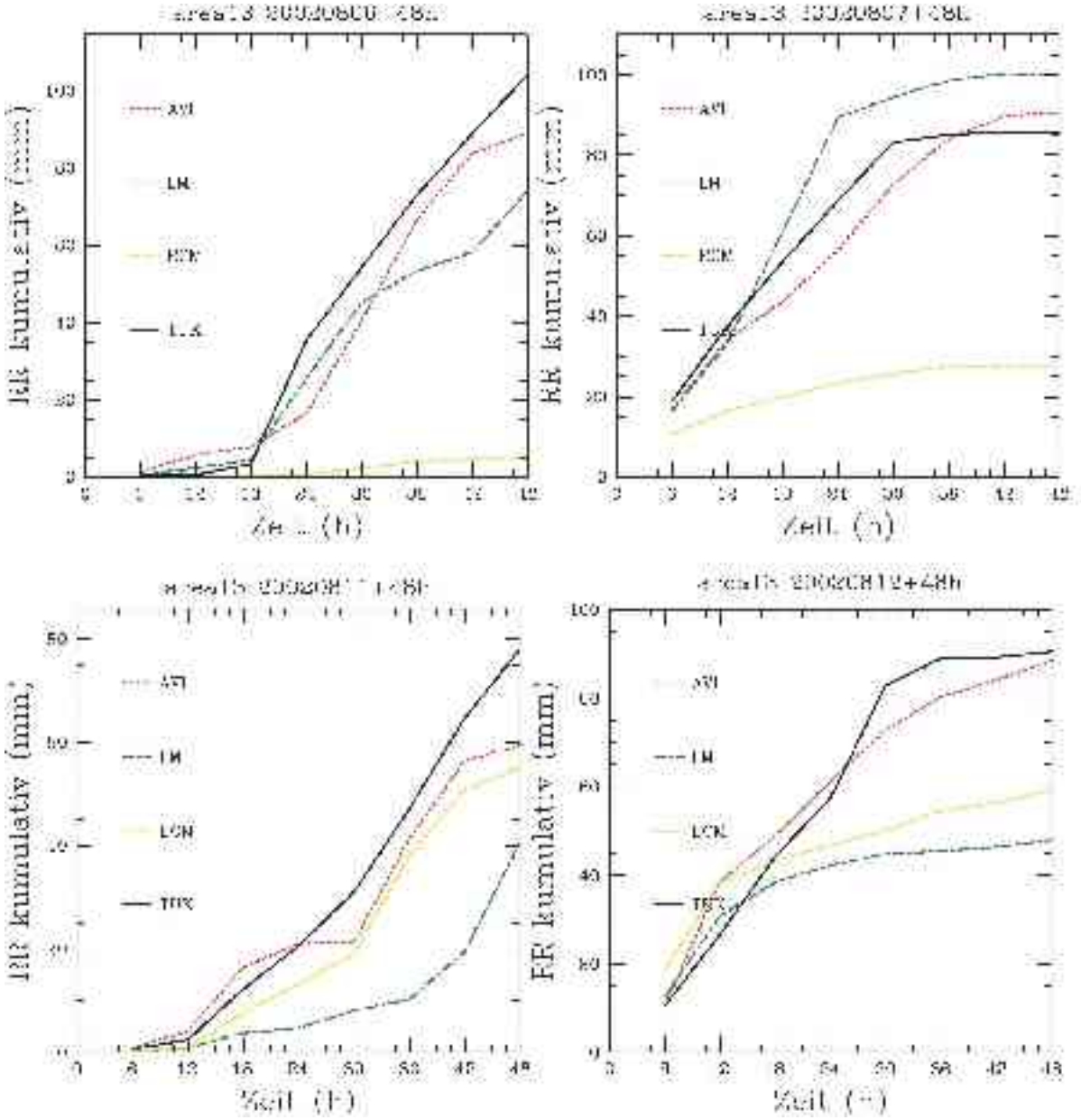


Figure 4: Cumulative, spatially averaged precipitation for area 13 (Traisen). Black solid line gives observations derived from a combination of HZB and TAWES data (TUK), red dotted line the ALADIN-Vienna forecast (AVI), green dashed-dotted line the forecast of the Lokal Modell of DWD, and yellow dashed line the ECMWF forecast. Upper row shows results for the first flood event, for analysis times 6 and 7 August 2002, 00 UTC. Bottom row shows results for the second part of the event, for analysis times 11 and 12 August 2002, 00 UTC.

In Figure 4 a comparison of the three models with observations is shown. In contrast to the Kamp area, even the first part of the event (top row) was relatively well forecasted by the limited area models ALADIN and LM (relative errors in the range 10-30%), on the contrary ECMWF failed completely in this case. During the second part of the event ECMWF model shows better results than LM, but ALADIN simulates cumulative amounts most precisely.

5.3.3. Error statistics

If errors in the forecast of precipitation intensity change sign within an event, it can be expected that there is a tendency for error compensation as we go from shorter to longer durations (Figure 5). It was investigated to what extent the prediction of 48-h totals has smaller relative errors than the prediction of 6-h totals (and totals for durations in between).

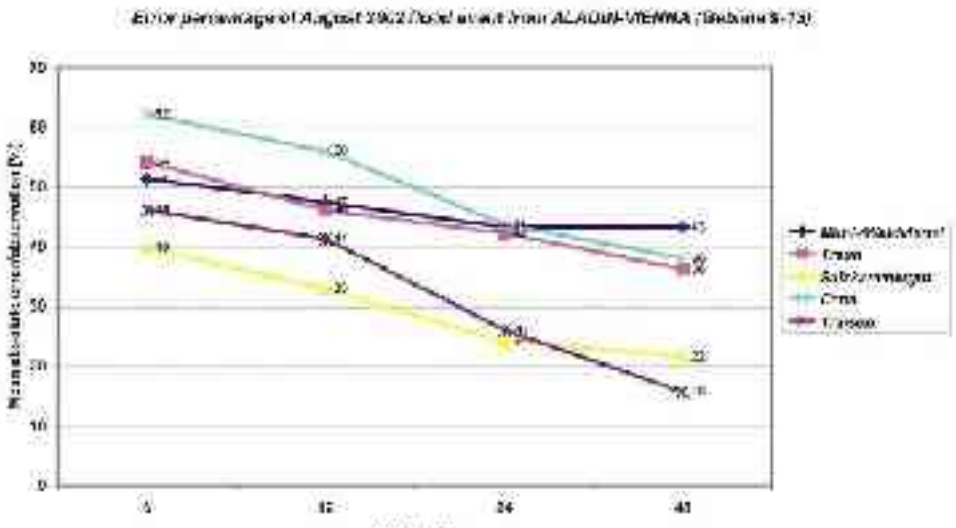


Figure 5: Mean absolute forecast error in percent of precipitation forecast of ALADIN (August 2002 flooding) as a function of duration for different areas.

For all areas the mean absolute error decreases with increasing duration (meaning the duration for which a forecast is done). Typically, the error drops from 40-60% at 6-hourly duration to 20-40% at 48-hourly duration. This is especially obvious for the areas Traisen and Enns, which contain large mountainous areas. Another area where orographic blocking effects play an essential role is the region Salzkammergut, where the model shows the best results for short forecast periods. On the other hand the model output does not significantly improve with increasing duration in this area. The region Mühl-/Waldviertel shows the smallest temporal compensation effect.

Figure 6 indicates similar tendencies for the LM model, apart from the error in the region Salzkammergut which is much higher than in ALADIN. A closer look at the precipitation intensities on single forecast runs showed that LM extremely overestimated the 48-h precipitation sum at the beginning of the flood event.

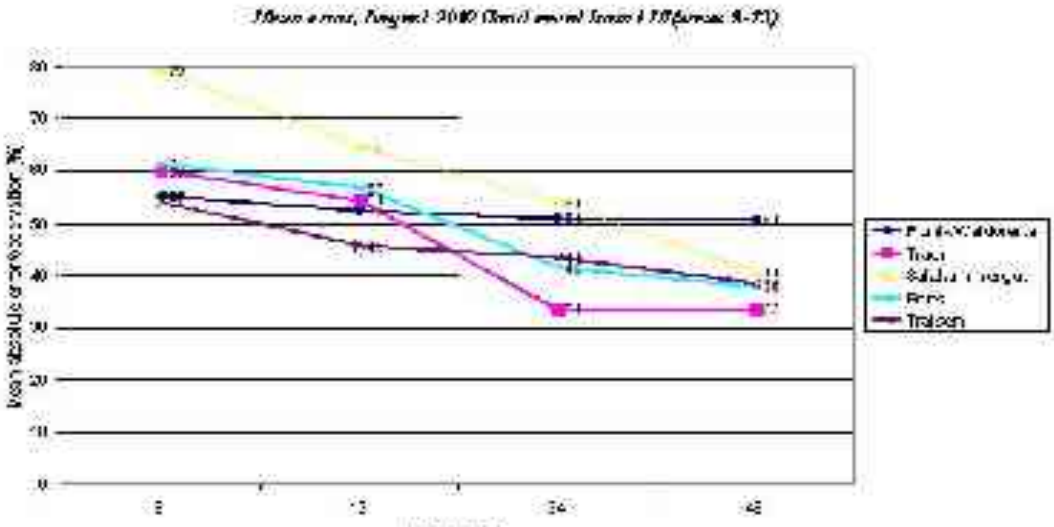


Figure 6: Mean absolute forecast error in percent of precipitation forecast of LM (August 2002 flooding) as a function of duration for different areas.

In contrast to LAMs, the global model of the European Center is not able to simulate the actual precipitation amounts occurring during such severe events. This is mainly due to the lower horizontal resolution which leads to smoother precipitation fields. For example, heavy precipitation due to blocking effects is smoothed out and therefore locally underestimated. As a result, the mean

absolute error does not vary much with duration and location (Figure 7). Even for a duration of 48 hours the model error does not drop below 40%.

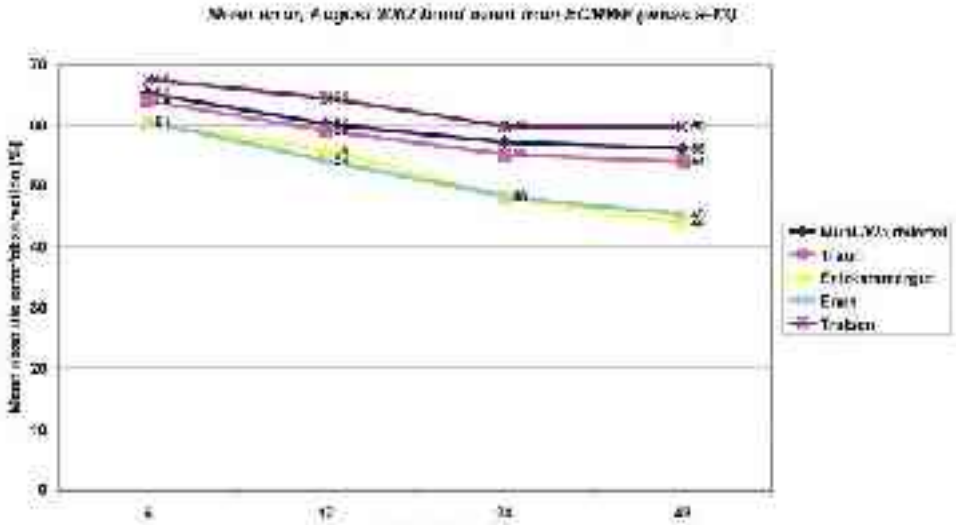


Figure 7: Mean absolute forecast error in percent of precipitation forecast of ECMWF operational run (August 2002 flooding) as a function of duration for different areas.

During the flood event 2002, the median, or 50% percentile, of ECMWF-EPS gives little indication for a heavy precipitation event and produces roughly the same error values as the reference run. For the August 2002 event we must increase the percentile up to 90% in order to gain a signal for extreme precipitation and obtain reduced errors. This percentile also gives a much more pronounced reduction of error with increasing duration. The mean absolute errors of the 90% percentiles are about 55-75% for a 6 hour duration, but decrease up to 15-32% for 48 hours. An exception is the region Enns, where the 90% percentile significantly overestimates the intensity of precipitation amount (50% error).

5.3.4. Summary/Conclusion

Comparing the performance of ALADIN, LM and the operational models of ECMWF during severe precipitation events in August 2002, weaknesses of NWP models highlighting the state-of-the-art turn out. Although the overall timing, the onset and offset of the event correspond qualitatively well to observations, single hourly peaks are rarely simulated by both LAMs. Lowland regions, that are mostly affected by convective precipitation events, show worse results than regions at the Alpine rim, although especially LM tends to overestimate precipitation amounts. A decrease of forecast errors when increasing the duration is recognized in both LAMs, whereas the ECMWF model does not show this compensation effect. ECMWF EPS forecasts hardly give a hint for heavy precipitation, especially mean, median or even 75% percentile underestimated precipitation amounts. Only regarding high percentiles (~90% percentile) the forecast error rapidly decreases.

5.4. First tests of the AROME prototype.

Yann Seity *(MF, CNRM/GMAP)* and Patrick Jabouille *(MF, CNRM/GMME)*

5.4.1. Introduction

The AROME project consists in developing a complete convection-resolving Numerical Weather Prediction (NWP) system over mainland France, and is expected for operational production around 2008. Méso-NH is a research model developed by Météo-France and the Laboratoire d' Aérologie (LA), with a very complete physics. AROME will integrate the ALADIN NH dynamics and the Méso-NH physics.

This paper describes the first 2D comparison tests performed with the AROME prototype. We run the latest fully-validated version of ALADIN NH (pre-cycle 27), with the Méso-NH physics interfaced. The objective is to verify that the AROME prototype is able to reproduce the main features of the squall line, which was well simulated by Méso-NH (Caniaux et al., 1994).

5.4.2. Description of the run

For this academic test, the physical package only includes the Méso-NH micro-physical scheme "ICE3", and the Méso-NH 1D turbulence scheme. We do not use surface nor radiation schemes, but a simplified friction is applied to the wind at the surface.

The initial profile is based on the observation of a tropical squall line that occurred on 23 June during the COPT 81 experiment. It provides horizontally constant wind, temperature and humidity.

To run the Méso-NH physics in ALADIN NH, we have added 6 new prognostic variables in ALADIN :

- 5 micro-physical specific humidity fields for :
  rain (qr),
  cloud droplets (qc),
  ice crystals (qi),
  snow (qs),
  graupel (qg),
- 1 turbulent variable : the Turbulent Kinetic Energy (TKE)

The domain consists in 180 horizontal gridpoints in AROME (160 in Méso-NH) with a grid-length of 2.5 km (the expected grid-length of AROME). The vertical grid uses 46 levels up to 22 km height.

The length of the run is 8 hours, with a time-step of 7.5 seconds. The AROME dynamics uses the ($d4$, $P2$) variables and a two-time-level semi Lagrangian scheme, whereas Méso-NH uses an Eulerian leap-frog scheme. We still do not use a P/C scheme in AROME.

The convection is initiated by a cold pool produced with a 0.01 K/s cooling rate applied for 10 minutes in a low-level area. A Galilean transformation of -16 m/s is applied to keep the system in the simulation frame.

5.4.3. Results

Figures 1 and 2 show some AROME and Méso-NH fields, respectively, after 3 hours of integration. The two models are in good agreement.

(a) (c)

(b) (d)

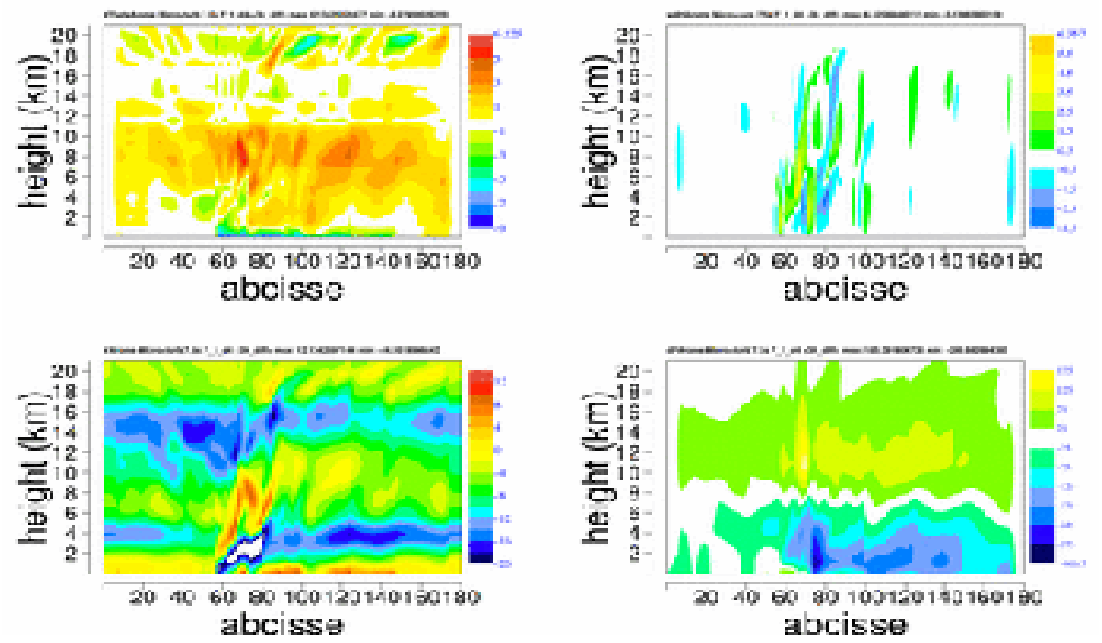


Figure 1 : AROME (time step = 7.5 s) after 3h : a) θ(3 h)-θ(0 h) (K), b) vertical velocity (m/s), c) horizontal wind field relative to the squall line (m/s), d) pressure(3 h)-pressure(0 h) (Pa)

(a) (c)

(b) (d)

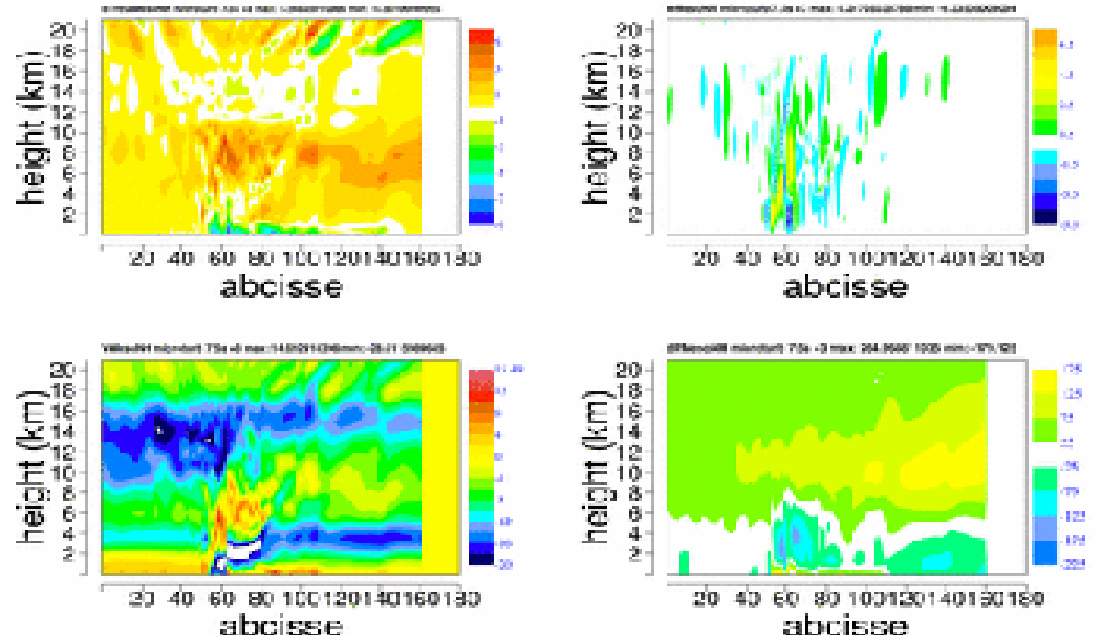


Figure 2 : Méso-NH (time step = 7.5 s) after 3h : a) θ(3 h)-θ(0 h) (K), b) vertical velocity (m/s), c) horizontal wind field relative to the squall line (m/s),d) pressure(3 h)-pressure(0 h) (Pa)

The moving speed of the squall line is the same in the two models, which is a proof of a good dynamics / micro-physics interaction. The dynamical structure of the squall line is well established. The cold pool is clearly visible on Figures 1c and 2c, with a relative wind speed lower than -16m/s. The maximal vertical velocities are 6.2 m/s in Méso-NH, and 6.3m/s in AROME (Figures 1b et 2b). The heating in the highest part of the cells and in the stratiform part of the system, associated with latent heat exchange, are comparable (in red on Figures 1a and 2a). The cooling associated with the evaporation of precipitations is also visible near the ground (in blue on Figures 1a and 2a). We can also notice a slight difference on the front of the squall line with a cooling of about -2K

between 2 and 4 km height in AROME, which do not exist in Méso-NH. The explanation for this behaviour would need further diagnostics which have not yet been performed.

Concerning the perturbation of pressure (Figures 1d and 2d), AROME gives a positive perturbation in the top of the cells of about 2 hPa, which is only 1hPa in Méso-NH. This difference could be related to the system of equations used (anelastic in Méso-NH, and fully compressible in AROME). Moreover, the precipitating species are not taken into account exactly the same manner in the two models. But once again, it would need further diagnostics to be well established. Figure 3 displays the cumulated prognostic micro-physical species (rain, ice, graupel, snow and cloud droplets). At first view, the two models are in good confidence. If we look at more details, we can see that the maximum is about 5.7 g/kg in AROME, and 6.7 g/kg in Méso-NH. Most of this maximum is due to the graupel content. At lower levels, Méso-NH seems to produce a little more rain than AROME. The secondary cell in the right part of the system seems to contain more condensed water in AROME.

(a) (b)

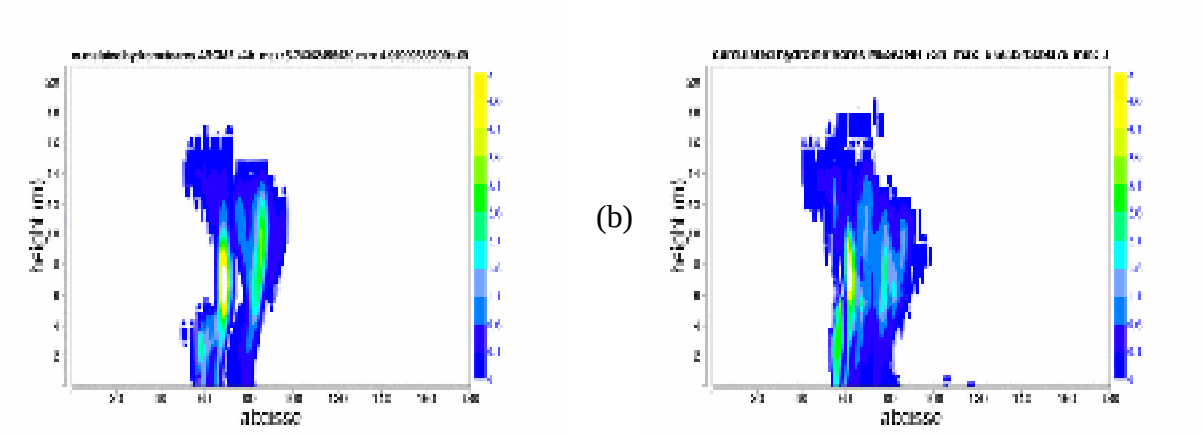


Figure 3 : Cumulated prognostic micro-physical species (g/kg) after 3h, in a) AROME b) Méso-NH (time-steps 7.5 s)

The results after 4 hours of simulations could not be compared with Méso-NH because of different lateral boundary conditions which perturbed the inside domain. Lateral boundary is opened in Méso-NH. This kind of boundary condition is not coded in ALADIN and so the boundaries are relaxed toward the initial profile with a Davies' coupling. We have to try to enlarge the horizontal size of the domain in AROME in order to decrease the influence of the boundary conditions in the squall-line evolution.

In AROME, the semi-implicit semi-Lagrangian scheme allows us to increase the time-step. Figure 4 and Figure 5 show the same fields as Figure 1, but with time-steps of 30 s and 60 s respectively. The results with 30 s are quite similar to those with 7.5 s. The maximum vertical velocity is 6.2 m/s. With a 60 s time-step, the squall line still exists, but is less vigorous. There are fewer convective cells, and for example, the maximum vertical velocity is only 3.6 m/s.

5.4.4. Conclusions

As one of the goals of AROME is to improve the prediction of thunderstorms over France, this first case study with the AROME 2D prototype is encouraging. It shows that AROME is able, as Méso-NH, to reproduce the dynamics and micro-physics of a squall line. The tests of increasing the time-steps, needed for operational production, are also encouraging. The expected time-step for AROME should be in the range between 30 s and 60 s.

Further tests of the AROME prototype are performed on a 3D academic case of a hot bubble, before trying to perform a real case. The ALARO-10km prototype, devoted to coarser resolution but based on the same physics/dynamics interface, Méso-NH physics and parametrized convection of Méso-NH, is also under tests. We will probably hear soon other news about AROME.

(a) (c)

(b) (d)

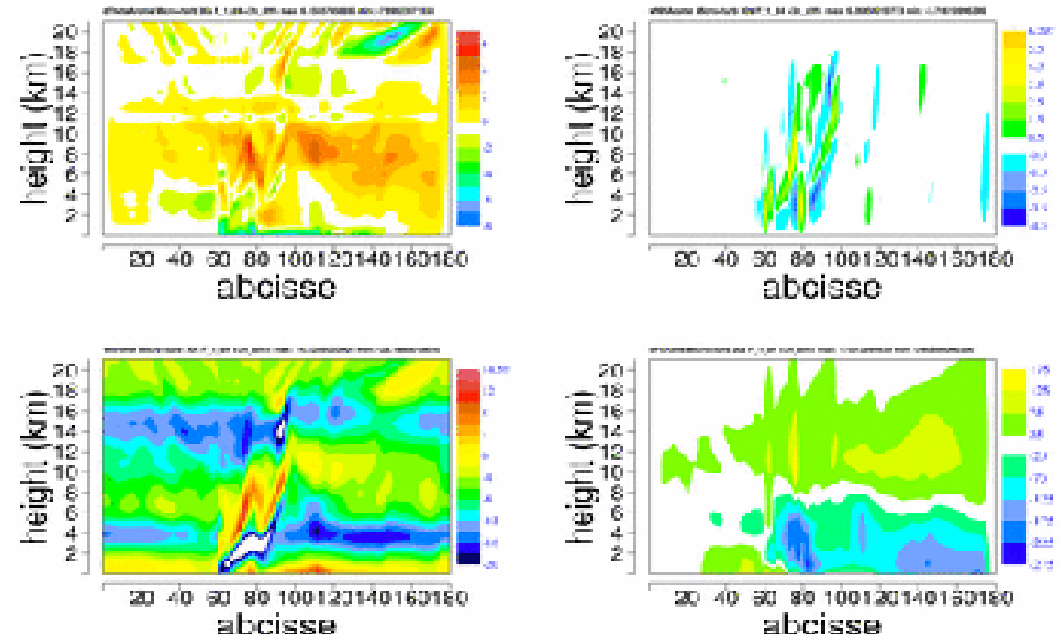


Figure 4 : AROME (time step 30 s) after 3h : a) θ(3 h)-θ(0 h) (K), b) vertical velocity (m/s), c) horizontal wind field relative to the squall line (m/s), d) pressure(3 h)-pressure(0 h) (Pa)

(a) (c)

(b) (d)

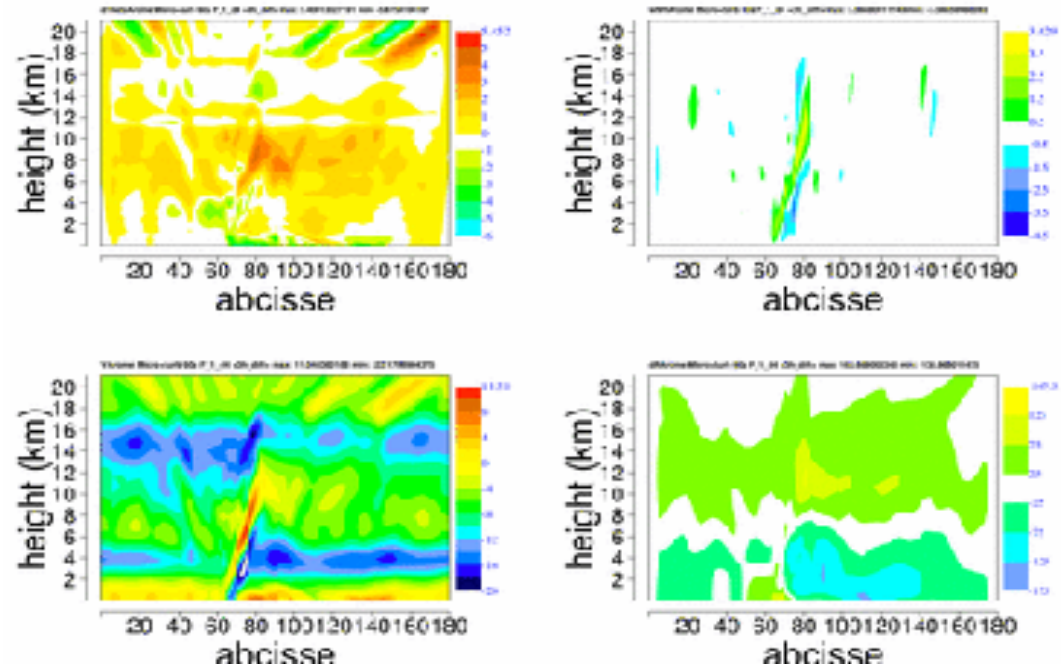


Figure 5 : AROME (time step 60 s) after 3h : a) θ(3 h)-θ(0 h) (K), b) vertical velocity (m/s), c) horizontal wind field relative to the squall line (m/s), d) pressure(3 h)-pressure(0 h) (Pa)

5.4.5. References

Caniaux, G., J.-L. Redelsperger, and J.-P. Lafore, 1994, A Numerical Study of the Stratiform Region of a Fast-Moving Squall Line. Part I: General Description and Water and Heat Budgets.J. Atmos. Sci., 51, 2046-2074.

5.5. Wavelet representation of background error covariances

Alex Deckmyn *(RMI)* and Loïk Berre *(MF)*

5.5.1. Introduction

In 3D-Var, the cost function to be minimized is

$$J(\mathbf{x}) = \tfrac{1}{2} (\mathbf{x} - \mathbf{x}^b)^* \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}^b) + \tfrac{1}{2} (\mathbf{Hx} - \mathbf{y})^* \mathbf{R}^{-1} (\mathbf{Hx} - \mathbf{y})$$

where $\mathbf{x}$ is the analysis, $\mathbf{x}^b$ the background and $\mathbf{y}$ the observations. $\mathbf{H}$ is the observation operator, $\mathbf{B}$ the matrix of background error covariances and $\mathbf{R}$ contains the covariances of the observation errors (* denotes the adjoint).

To improve the representation of local variability and heterogeneity, we have used a wavelet basis. The experiments are for horizontal covariances of temperature errors.

Details of the work are in Deckmyn & Berre (2004).

5.5.2. Hybrid 3-basis approach

In ALADIN, the covariance matrix $\mathbf{B}$ is represented in spectral space :

$$\mathbf{B} = \mathbf{F}^{-1} \mathbf{B_f} \mathbf{F}.$$

The structure functions are assumed to be homogeneous and with this simplification $\mathbf{B_f}$ becomes diagonal in spectral space. An additional simplification of isotropy is also imposed: all diagonal entries corresponding to the same absolute wavenumber $|k|$ are equal.

To introduce heterogeneity in the model of $\mathbf{B}$, we introduced a hybrid approach that utilises 3 bases and 3 diagonal matrices :

$$\mathbf{B} = \mathbf{D_g} \mathbf{F}^{-1} \mathbf{D_f} \mathbf{F} \mathbf{W}^{-1} \mathbf{B_w} (\mathbf{W}^{-1})^* \mathbf{F}^* \mathbf{D_f}^* (\mathbf{F}^{-1})^* \mathbf{D_g}^* ,$$

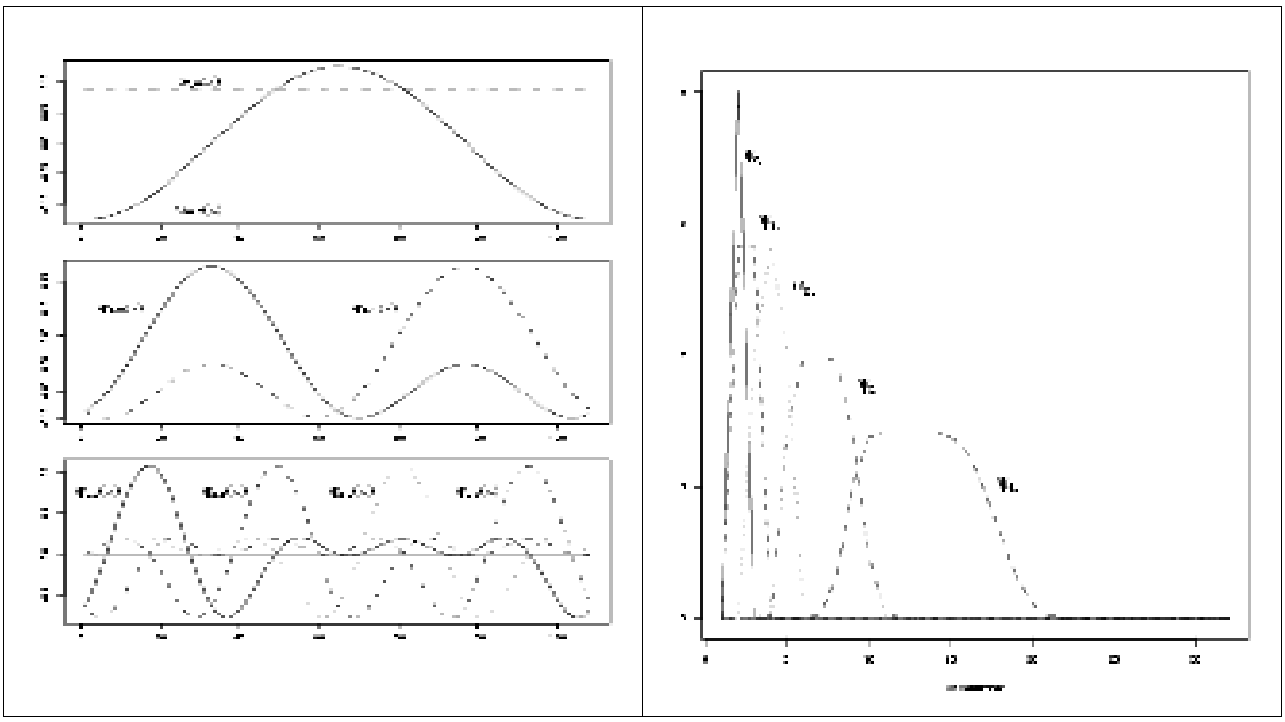
where $\mathbf{F}$ and $\mathbf{W}$ signify the Fourier and wavelet transforms, respectively, and $\mathbf{D}_.$ is a diagonal matrix in one of the bases. $\mathbf{D_g}$ contains the standard deviations in grid space and $\mathbf{D_f}$ describes the mean correlation function. $\mathbf{B_w}$ describes the local deviation from the implied covariance functions (i.e. implied by the local variances and by the average correlation function) at different scales. Note that setting $\mathbf{B_w}$ to unity reduces this formula back to the common spectral approach. In short, we could say that *we use the 3 different coordinate systems to model that aspect of the covariance matrix they are best suited for*. Fourier space models the average structure function and the average tilt, wavelet space the local and scale-dependent heterogeneities that remain and grid space the local variance (which could also be seen as the heterogeneity at the smallest possible scale).

5.5.3. Orthogonal Wavelet Transforms

Orthogonal wavelets form a class of basis function that combine properties of gridpoint and spectral bases. They are localized in both grid and Fourier space, which makes them a very powerful tool for analysis. A comprehensive overview of orthogonal wavelet analysis in the field of meteorology is given by Fournier (2000).

We have used the smooth, symmetric wavelets introduced by Meyer (1992). In the figure below, the Meyer wavelet is plotted at 3 scales (index j). This illustrates how the wavelets are localized in grid space and how the distance between 2 neighbouring basis functions (index k) depends on the scale j. On the right are the spectra of the wavelets, illustrating the spectral localization (band-pass).

The standard Meyer wavelet transform requires the domain size to be a power of 2. This is rarely the case. To define The wavelet transform of domains of any size, we constructed generalizations of Meyer wavelets that scale by other factors (e.g. 3 or 5).



5.5.4. 2D wavelets

The common technique to construct 2D wavelets is as the tensor products of 1D wavelets. Such wavelets give a special status to the X and Y directions, which adds a directional component to the transform, in addition to scale and location. Directional resolution, though, is very coarse. Its effect will be shown in the modelling of local anisotropy.

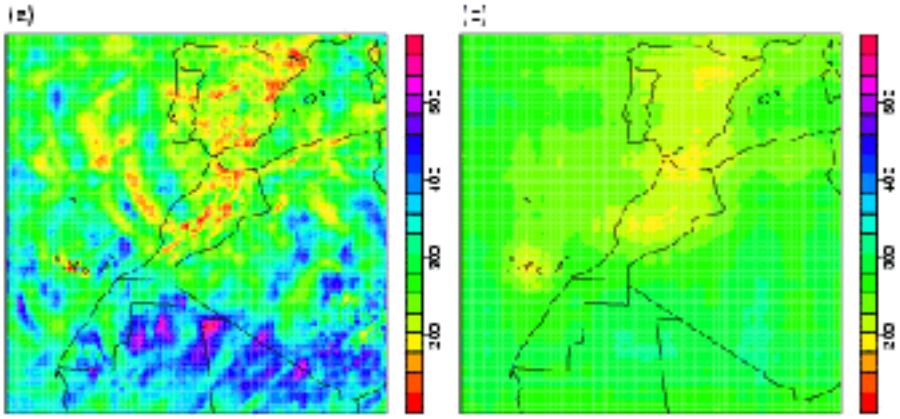
5.5.5. Data

The first data set is an ensemble of 87 winter days on the ALADIN/Morocco domain (128x128 including the extension zone and 31 levels). Since $128=2^7$ this domain can be described with traditional dyadic wavelets. This same data set was used by Raouindi (2001) who studied latitudinal heterogeneity using off-diagonal terms in the spectral covariance matrix.

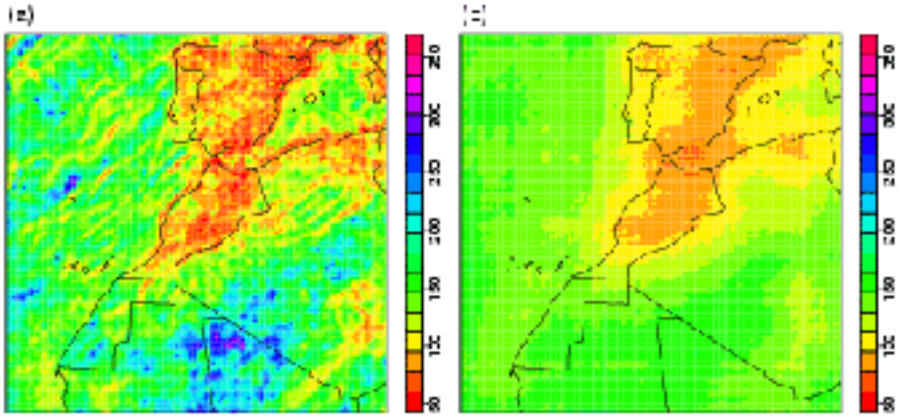
The other domain is ALADIN/France (300x300 and 41 levels). The factorization $300=2^2.3.5^2$ shows that this domain requires M=3 and M=5 wavelets in addition to the dyadic M=2 wavelet.

5.5.6. Length scales

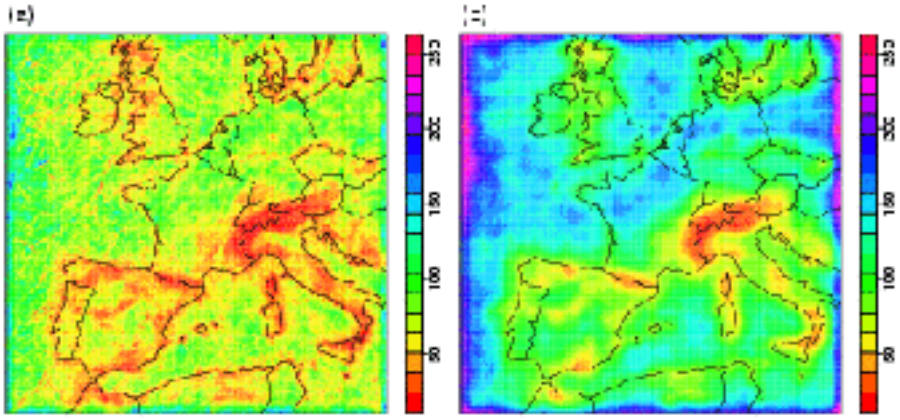
Length scales are defined as in Belo-Pereira et al. (2002). The results are shown for the ALADIN-Morocco domain at level 13 and 31 and for the ALADIN-France domain at level 41. The wavelet formulation appears to represent well the geographical variations of the 2D length-scale (with e.g. small length-scales over Spain and the Atlas region for the Moroccan domain, and small length-scales in mountain ranges - Alps, Pyrénées - for the ALADIN-France domain). The raw results are considerably smoothed, while the salient features are retained.



Length scale at level 13 (a) from raw data (b) from wavelet **B**.



Length scale at level 31 (a) from raw data (b) from wavelet **B**.



Length scale at level 41 (a) from raw data (b) from wavelet **B**.

5.5.7. Anisotropy

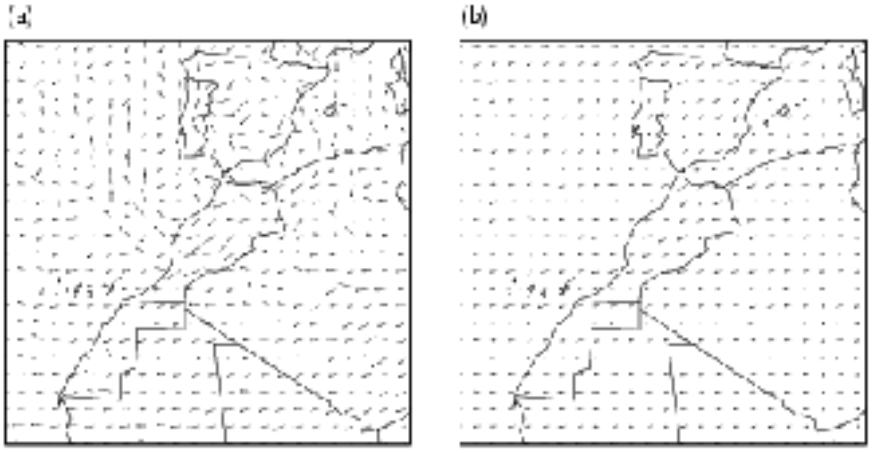
Our definition of anisotropy is based on the different length-scales in the *x* and *y* direction.

On the Moroccan domain e.g. the South-West/North-East elongation near the Atlas mountains and the zonal elongation in the Southern part (mostly at level 13) seem to be well captured. The modelled variations appear to be smoother, and some local anisotropies are not represented.

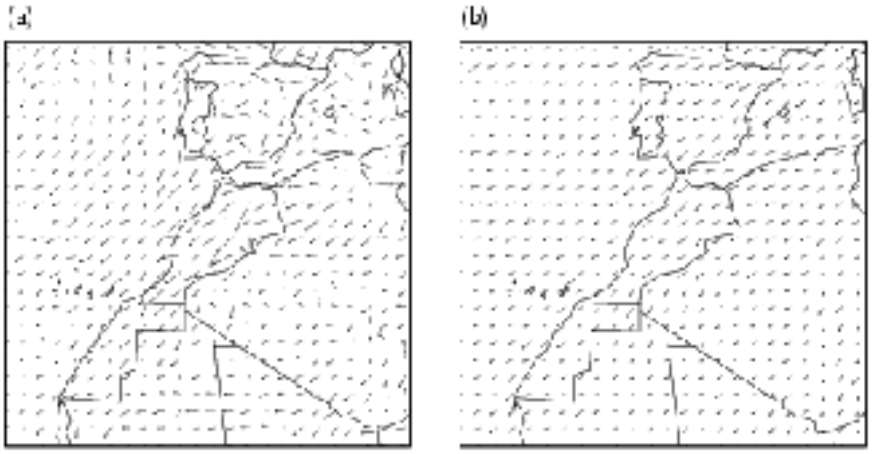
This smoothing of local anisotropies is also visible on the French domain. Particularly interesting is the representation of the anisotropy around the Alpine region.

These results reflect a lack of anisotropy that will be also illustrated in the next section. In the

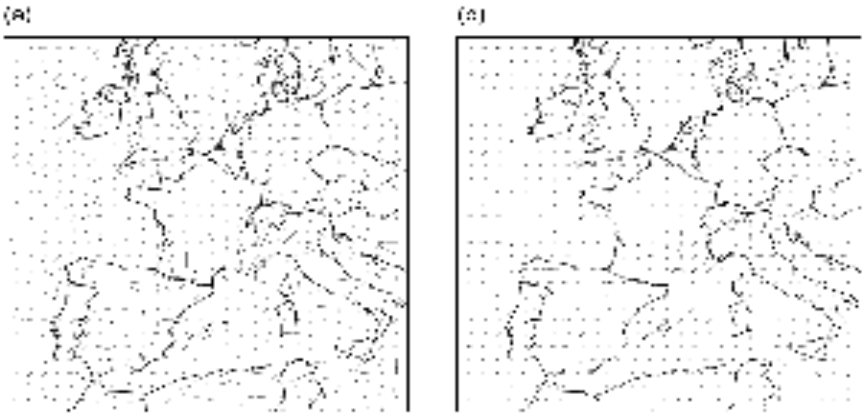
figures, only the direction of the axes is important, while the orientation of the arrows is arbitrary.



Local anisotropy axes at model level 13; (a) from raw data and (b) from **B**w



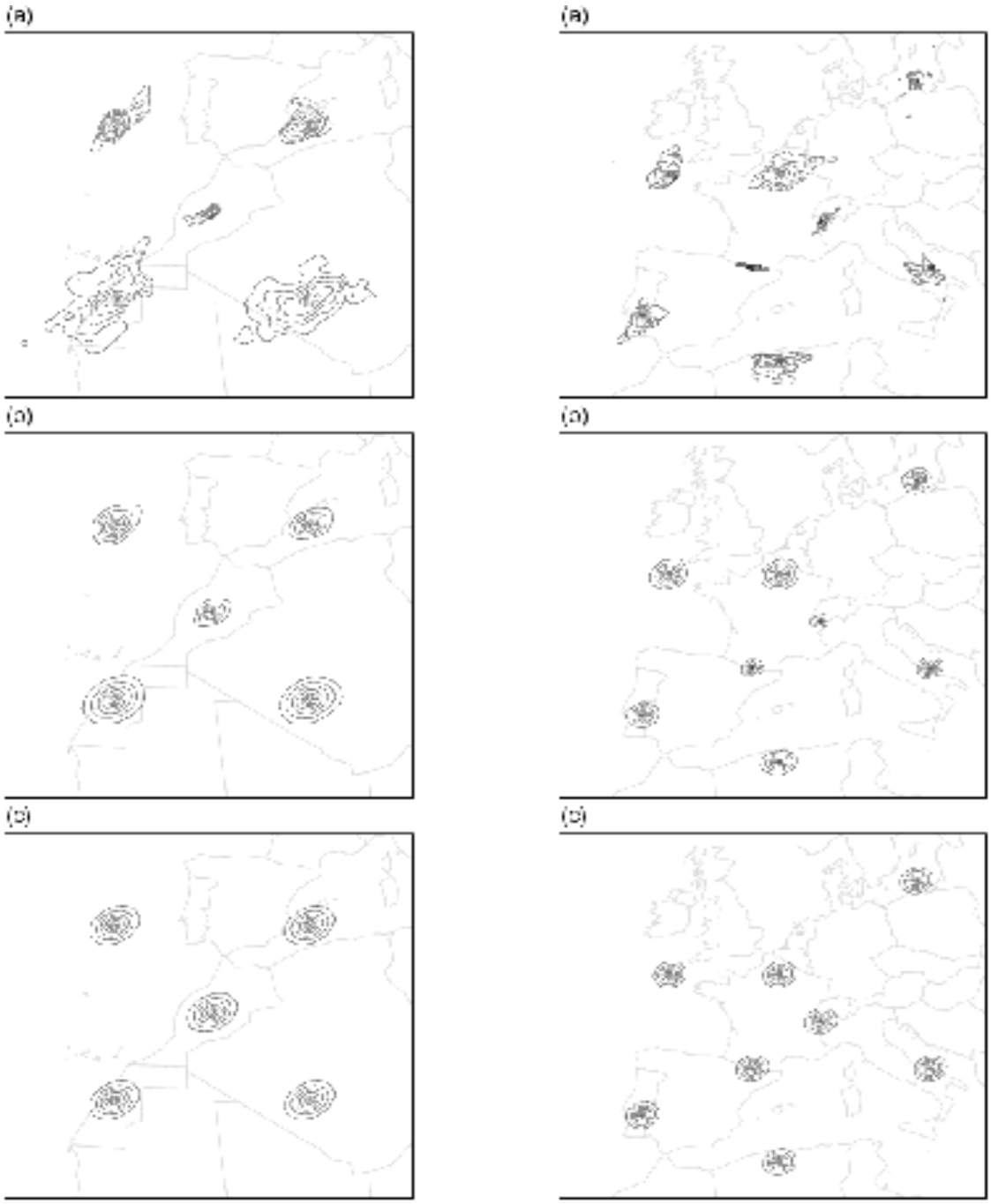
Local anisotropy axes at model level 31: (a) from raw data and (b) from **B**w



Local anisotropy axes at model level 41 (a) from raw data and (b) from **B**w

5.5.8. Correlation functions

Plotting some examples of local correlation functions confirms therefore that on one hand the chosen formulation is able to represent a large part of the geographical variations of the 2D length-scale, while on the other hand the achieved degree of local anisotropy is relatively small.



Local correlation functions at the lowest model level : (a) from raw data, (b) from**Bw**, and (c) from **Bf**. Contours are at 0.5,0.6...0.9,0.95,0.98

level 31, Moroccan domain level 41, French domain

5.5.9. Conclusions and outlook

At this stage, our hybrid wavelet approach is especially successful at reproducing local variations of length scales.

The average anisotropy is well represented in spectral space, but the local anisotropies of the correlation functions appear to be smaller than in the raw data. To improve the representation of these local anisotropies, one might either relax the diagonality of $\mathbf{B}_w$ or introduce wavelets that are more directional.

The extension of our results to 3 dimensions does not pose theoretical problems. We allowed the correlation function in spectral space to be anisotropic, and a similar freedom in vertical

correlations allows for the modelling of the mean tilt of the correlation functions. In that case the block-diagonal $\mathbf{B_f}$ describing the vertical correlations between the corresponding Fourier components at 2 different levels, will no longer be real-valued.

5.5.10. References

Belo-Pereira, M., L. Berre, and G. Desroziers, 2002 : Estimation et études des covariances d'erreur de prévision d'Arpège/Aladin. Ateliers de modélisation de l'atmosphère, Météo France/CNRM, 17-20.

Deckmyn, A. and L. Berre, 2004 :A Wavelet Approach to Representing Background Error Covariances in a LAM, submitted to *Mon. Wea. Rev.*

Fournier, A., 2000 : Introduction to orthonormal wavelet analysis with shift invariance: Application to observed atmospheric blocking spatial structure. J. Atmos. Sci., 57, 3856-3880.

Raouindi, M., 2001 : Etude et représentation spectrales de la variabilité latitudinale des covariances spatiales des erreurs de prévision sur une aire limitée. work report, ENM.

5.6. Ensemble dispersion spectra and the estimation of error statistics for a limited-area model analysis

Simona ?tefãnescu, *(NIMH, Romania)* and Loïk Berre *(Météo-France)*

5.6.1. Introduction

Deriving background error statistics can be done using an ensemble of perturbed assimilation systems. Houtekamer et al. (1996) combined two complementary approaches into a system simulation experiment, namely an ensemble prediction based on a perfect-model approach, for which only the observations that enter in the assimilation cycle are randomly perturbed, and model sensitivity experiments. The analysis ensemble approach was also implemented at the ECMWF (Fisher, 1999) and Météo-France (Belo Pereira, 2002).

In the present study, the ensemble approach is used to sample the forecast error covariances to be used in a 3D-Var data assimilation for the limited-area model ALADIN. A generalized formulation of this 3D-Var is considered, in which the analysis of the coupling model, namely ARPEGE, is included as an additional source of information (Bouttier, 2002). This is related to the idea of relying on the "fresh" ARPEGE analysis for the large scales, while still extracting the small-scale information of the ALADIN background (i.e. the 6h forecast).

The evolution of dispersion spectra in the perfect-model framework has been investigated. The ARPEGE/ALADIN model differences have been also evaluated, and a decomposition of them is proposed. Finally, the implications for the specification of the error statistics in the generalized formulation of the ALADIN 3D-Var data assimilation are pointed out. A comparison with the statistics derived through the NMC method has been carried out too.

5.6.2. Contributions to the evolution of dispersion spectra in a perfect-model framework

From the ARPEGE ensemble of perturbed assimilation cycles, it is possible to run the operational ALADIN limited-area system (currently in dynamical adaptation mode): this provides an ensemble of limited-area states, whose evolution of dispersion can be studied.

The effect of ARPEGE analysis is to reduce the error variance, especially in the large scales (see Figure 1). The reduction of the ARPEGE first guess variance is about 30 % for wavenumber 1.

After applying a digital filter initialization (DFI), a reduction of the error variance for ARPEGE analysis and first guess, especially in the small scales, has been observed. This can be explained by the fact that DFI removes some unbalanced components of the error variance, namely the structures artificially created in the small scales by the horizontal interpolation of the ARPEGE fields into the ALADIN grid.

Compared with the ARPEGE 6h forecast, the effect of ALADIN 6h forecast is to increase the error variance in the small scales. The ALADIN and ARPEGE 6h forecast error variance curves start to depart significantly at wavenumbers greater than 13-17, which corresponds to length scales smaller than 220-170 km, as can be seen from Figure 2. This means that the limited-area model builds up its own structures beyond these wavenumbers.

The smaller errors of ARPEGE (than those of ALADIN) that are suggested in the small scales can be understood on one hand knowing its low resolution and the effect of diffusion, but it can also be seen as a paradox on the other hand, as we would expect that the higher-resolution model (ALADIN) should give a better solution in the small scales.

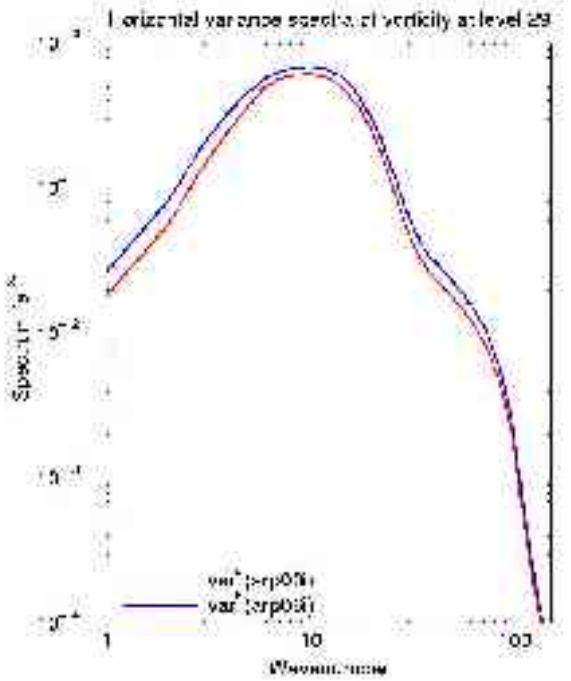


Figure 1: Dispersion spectra related to the initial condition perturbations for initialized ARPEGE analysis and first guess fields

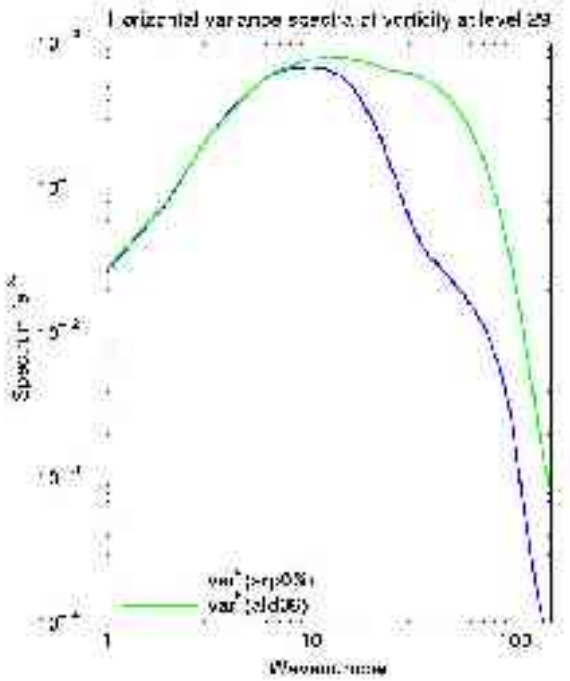


Figure 2: Dispersion spectra related to the initial condition perturbations for initialized ARPEGE first guess and ALADIN background fields

5.6.3. Model difference evaluation and decomposition

We have calculated the differences between the ARPEGE and ALADIN models, when these two models are subject roughly to the same initial state. The potential of these model differences is to give informations about some of the involved model errors.

The ARPEGE/ALADIN model differences appear to be of relatively small scale compared with the differences that are related to the initial condition perturbations, but there are also some significant contributions in the large scales. This suggests that the model differences arise not only from the differences in resolution, but possibly also from the coupling inaccuracies and from the interactions between the small and large scales. One may wonder therefore if it could be possible to distinguish these different possible contributions.

A parameter α has been defined, in order to estimate the part of the model differences that is related to the resolution differences. α is defined as the percentage of ALADIN dispersion that is unrepresented by ARPEGE :

$$\alpha = \frac{\mathrm{var}(ald06) - \mathrm{var}(arp06)}{\mathrm{var}(ald06)}$$

Further, the model differences variance $\mathrm{var}(\epsilon^m)$ can be decomposed as follows :

$$\mathrm{var}(\epsilon^m) = \underbrace{\alpha \cdot \mathrm{var}(\epsilon^m)}_{\mathrm{var}(\epsilon^{ss})} + \underbrace{(1-\alpha) \cdot \mathrm{var}(\epsilon^m)}_{\mathrm{var}(\epsilon^{ls})}$$

$\mathrm{var}(\epsilon^{ss})$ corresponds to the small-scale structures that are represented by ALADIN and not by ARPEGE : they may be interpreted as some ARPEGE model errors, with respect to the truth at the ALADIN resolution.

The residual $\mathrm{var}(\epsilon^{ls})$ corresponds to some large-scale structures, that are related e.g. to some coupling inaccuracies and to some small scale/large scale interactions. In the future, a refined decomposition of this residual could be obtained by comparing some global and limited-area models with similar resolutions, and also by comparing global models with different resolutions.

The decomposition of the ARPEGE/ALADIN model differences variance into small-scale and

large-scale parts is represented in Figure 3.

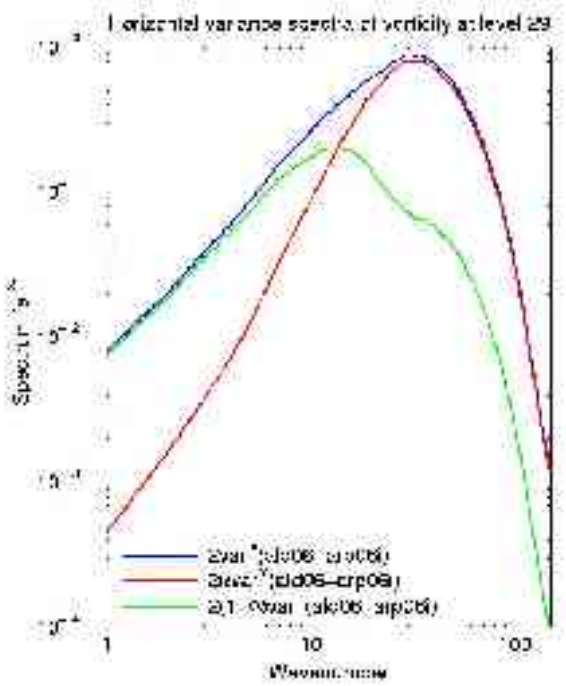


Figure 3: Decomposition of initialized ARPEGE first guess / ALADIN background differences variance

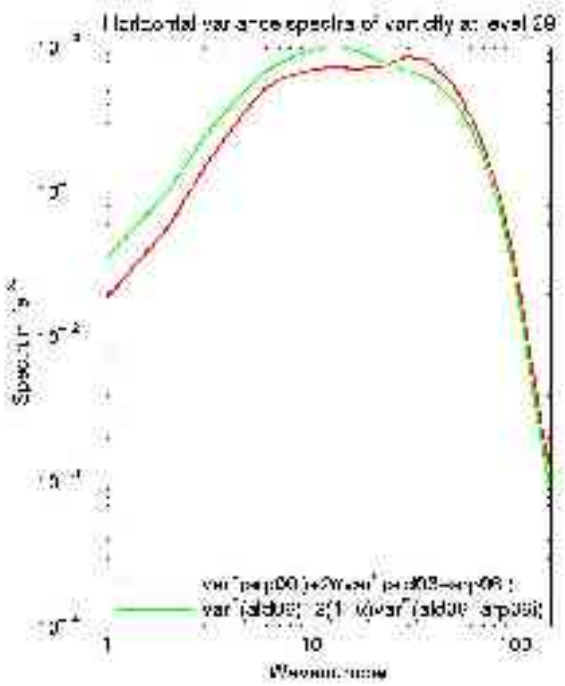


Figure 4 : The final dispersion spectra for initialized ARPEGE analysis and ALADIN background fields

5.6.4. Implications for the specification of the error statistics in the generalized formulation of the ALADIN 3D-Var

The variances of the decomposed model differences have been added to the respective dispersion variances of the ARPEGE analysis and of the ALADIN background (see Figure 4). A multiplying factor 2 is used for these model difference variances, to be consistent with the corresponding factor 2 that is implicit in the variance estimates provided by the ensemble of analyses with a perfect model.The variance of the small scale structures that are unrepresented by the ARPEGE model has been added to the variance of the ARPEGE analysis dispersion. This increases strongly the small scale dispersion, while leaving the large scale dispersion mainly unchanged.

The variance of the large scale model differences may be added to the variance of the ALADIN background dispersion, if they are interpreted as being caused by ALADIN (due to coupling errors for instance); this increases mostly the large scale dispersion.

The comparison between the two final dispersion spectra suggests that in the large scales, the ARPEGE analysis errors are smaller than those of the ALADIN background, while the reverse holds in the small scales.

This ensemble approach, based on some perturbed assimilation cycles and model differences, appears therefore to be a good framework for the evaluation of the error statistics that are involved in the generalized formulation of the ALADIN 3D-Var.

5.6.5. References

Belo Pereira M., 2002 : Improving the assimilation of water in a NWP model.*ALATNET Newsletters 4 and 5.*

Bouttier F., 2002 : Strategies for kilometric-scale LAM data assimilation.*Hirlam workshop report on variational data assimilation and remote sensing* 21-23 January 2002, Helsinki, p.9-17.

Fisher M., 1999 : Background error statistics derived from an ensemble of Analyses.*ECMWF Research Department Technical Memorandum,* **79**, September 1999, 12pp.

Houtekamer et al., 1996 : A system simulation approach to ensemble prediction.*Mon. Wea. Rev.,* 124. 1225-1242.

5.7. Introduction of the β-plane into the horizontal balance equation of ALADIN Jb

Rachida El Ouaraini *(Maroc-Météo)* and Loïk Berre *(Météo-France)*

5.7.1. Introduction

The principal objective of this study is to code and to test the β-plane approximation into the horizontal balance equation, in the software of covariance calculation ("festat") and in the ALADIN 3D-Var, and to compare it with the previous approach (based on the f-plane approximation).

Indeed, the adoption of the f-plane approximation in the horizontal balance equation remains reasonable on small domains far from the equator, knowing that the more one moves away from the equator, the less the variations of f are fast. The situation becomes worrying when great domains are treated, in particular those close to the tropics where the variations of the Coriolis parameter are not negligible. Model ALADIN-NORAF is an example.

The horizontal balance equation currently used in ALADIN adopts the f-plane approximation and is close to the following equation :

$$\phi = f_0 \Delta^{-1} \zeta = f_0 \psi$$

ϕ : linearized geopotential.

f_0 : mean value of the Coriolis parameter.

Δ^{-1} : inverse of the Laplacian.

ζ : relative vorticity.

ψ : stream function.

The equation used in this study to express the horizontal balance in β-plane in "festat" and the ALADIN 3D-Var is :

$$\phi = \Delta^{-1}(f\zeta)$$

We took into account in this equation the main part of the horizontal variations of f (which, implicitly, amounts writing $f = f_0 + \beta y$, even if the formulation and the tested balance code do not utilize explicitly the parameter β).

5.7.2. Bi-Fourier development of the horizontal balance equation

The coding of the horizontal balance equation in β-plane in "festat" and the ALADIN 3D-Var is carried out in the bi-Fourier complex space (over a half ellipsis) rather than in the space of real quadruplets (over a quarter of ellipsis). The choice was made taking into account the two following elements :

The equations are definitely simpler in the bi-Fourier complex space than in the space of real quadruplets.

The equations connecting these two spaces are simple. The fact of using them explicitly in the code does not generate an additional cost, and improves the legibility of the code.

The development of the equation (2) in bi-Fourier is :

$$\hat{\phi}(m,n) = \widehat{(f\psi)}(m,n) = \sum_l \sum_p (\hat{f})(l,p)\,\hat{\psi}(m-l, n-p)$$

If it is supposed that f depends only on the direction y of the ALADIN model in grid points (and thus of the meridional wave number n in the spectral representation), one has :

$$f(x,y) = \sum_p (\hat{f})(0,p)\, e^{2\pi i p y}$$

$$\hat{\phi}(m,n) = \sum_p (\hat{f})(0,p)\,\hat{\psi}(m, n-p)$$

Within this framework, Loïk Berre had made beforehand a study of truncation of f and had

found that a truncated representation of the Coriolis parameter around T10 and T20 is a good approximation of f over the NORAF domain (Figure 1). He also checked that truncation can be reduced even more when one increases the size of the extension zone : thus a truncation of T05 is enough when the extension zone accounts for example for 25 % of the field (Figure 2), instead of 6 % in Fig. 1.

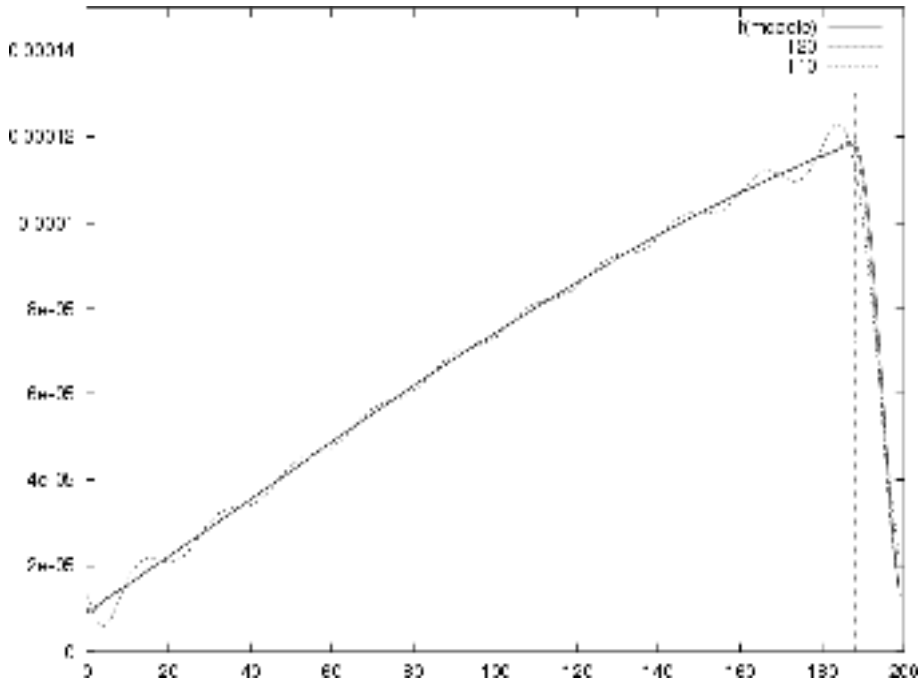


Figure 1 : Representation of the Coriolis parameter with a truncation T10 and T20 and comparison with f resulting from the model.

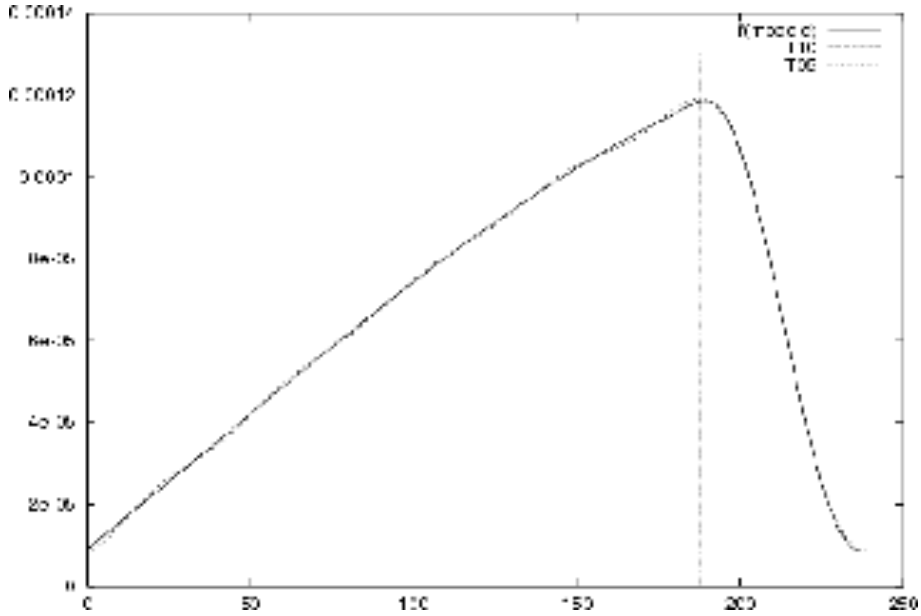


Figure 2 : Truncation can be reduced even more when the size of the extension zone E increases.

5.7.3. Coding in the software of covariance calculation ("festat") and diagnosis of the explained variance

We were interested in the study of the β-plane on the level of error statistics on the-NORAF domain.

To test the contribution of the β-plane approximation in the horizontal balance equation, one

plotted α, the variance percentage of linearized geopotential ϕ, which is explained by the horizontal balance equation used in this study.

We plotted the parameter α (explained variance) in the case of the f-plane and the β-plane T15 (Figure 3).

We note that the explained variance clearly increased while passing from the f-plane (maximum around 43 %) to the β-plane (maximum around 54 %); this is a first validation of the approach of the β-plane in the horizontal balance equation. We were then interested in coding in the 3D-Var ALADIN.

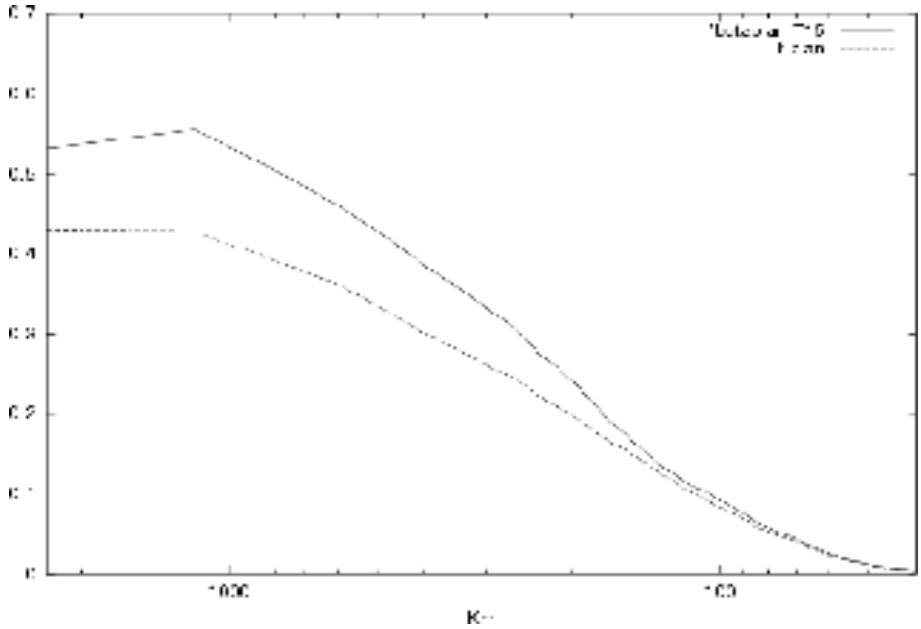


Figure 3 : Percentage of variance of the geopotential which is explained by the horizontal balance, in the case of f-plane and β-plane T15.

5.7.4. Experiments with an observation at two different latitudes

The goal of this experiment is to check that the variations of the temperature and wind increments are coherent with the effects of f on the guess error statistics. One varies the position of an observation, while keeping constant the innovation amplitude. The increment equations (Berre 2001), which we use to validate our results, are detailed in the following section.

Increment equations and variations awaited

Impact of the β-plane on the guess error statistics of ϕ :

$$\phi = f\psi + \phi_u \quad [1]$$

$$\sigma_\phi{}^2 = (f\sigma_\psi)^2 + \sigma_{\phi_u}{}^2 \qquad \left(cov(f\psi, \phi_u) = 0\right)$$

$$cov(\phi, \psi) = f\sigma_\psi{}^2$$

$$cor(\phi, \psi) = \frac{1}{1 + (\sigma_{\phi_u} / f\sigma_\psi)^2}$$

σ_ϕ, $cov(\phi,\psi)$ and $cor(\phi,\psi)$ are respectively the standard deviation of ϕ, the cross-covariance between ϕ and ψ, and the cross-correlation between ϕ and ψ. These quantities decrease when the latitude gets lower[2].

1 This equation is a simplified version of the coded exact equation, which is : $\phi = \gamma\Delta^{-1}(f\zeta) + \phi_u$

2 σ_ψ and σ_{ϕ_u} are constants (assumption of horizontal homogeneity)

Impact of an observation of ϕ over ϕ :

$$d\phi = \frac{1}{1+(\sigma_\phi^o / \sigma_\phi)^2}\ \delta\phi = \frac{1}{1+(\sigma_\phi^o)^2/[(f\sigma_\psi)^2+(\sigma_{\phi_u})^2]}\ \delta\phi \quad (1)$$

$d\phi$: increment value;

$\delta\phi$: innovation value;

σ_ϕ^o : standard deviation of the observation error;

σ_ϕ : standard deviation of the guess error.

Impact of an observation of ϕ over ψ :

$$d\psi = \frac{cov(\phi,\psi)}{\sigma_\phi^2}\ d\phi = \frac{f\sigma_\psi^2}{f^2\sigma_\psi^2+\sigma_{\phi_u}^2}\ d\phi = \underbrace{\frac{1}{1+(\sigma_{\phi_u}/f\sigma_\psi)^2}}_{1\downarrow}\ \underbrace{\frac{1}{f}}_{2\uparrow}\ \underbrace{d\phi}_{3\downarrow} \quad (2)$$

The product of terms 1 and 2 increases slightly when the latitude drops (this result is valid in the mid-latitudes, whereas it is the reverse in the tropics) (Berre 2001).

Term 3 ($d\phi$) decreases when the latitude drops.

The product of terms 1, 2 and 3 should decrease when the latitude drops, because the variations of term 3 have the largest effect in the mid-latitudes.

Impact of an observation of ψ over ψ :

$$d\psi = \frac{1}{1+(\sigma_\psi^o / \sigma_\psi)^2}\ \delta\psi$$

As $\sigma_\psi{}^o$ and σ_ψ are constants, $d\psi$ should remain constant.

Impact of an observation of ψ over ϕ :

$$d\phi = \frac{cov(\phi,\psi)}{\sigma_\psi^2}\ d\psi = \frac{f\sigma_\psi^2}{\sigma_\psi^2}\ d\psi = f\ d\psi$$

As $d\psi$ remains constant, $d\phi$ should decrease when f decreases.

The two selected positions are :
- in the Channel (Lat = 50° , Lon = 0°).
- in Spain (Lat = 40° , Lon = 0°).

Observation of temperature

For the validation of our results using the preceding equations of increments, one considers that the increments of temperature T vary in a way similar to those of geopotential ϕ (T and ϕ are connected by the hydrostatic relation), and the variations of the increments of the wind $\vec{u}$ are close to those of the increments of ψ (knowing that $\vec{u} \approx \vec{k} \times \vec{\nabla}\psi$ in the mid-troposphere of the mid-latitudes).

We will show here the results obtained when using a single observation of temperature. Similar consistent results were obtained when using a single observation of wind, in accordance with the increment equations of the previous subsection.

Temperature increment

As the equation (1) suggests, one notices that the amplitude of the increment of the temperature drops towards the low latitudes : the increment in the Channel has a larger amplitude from approximately 10 % than the increment in Spain (Figs 4 and 5).

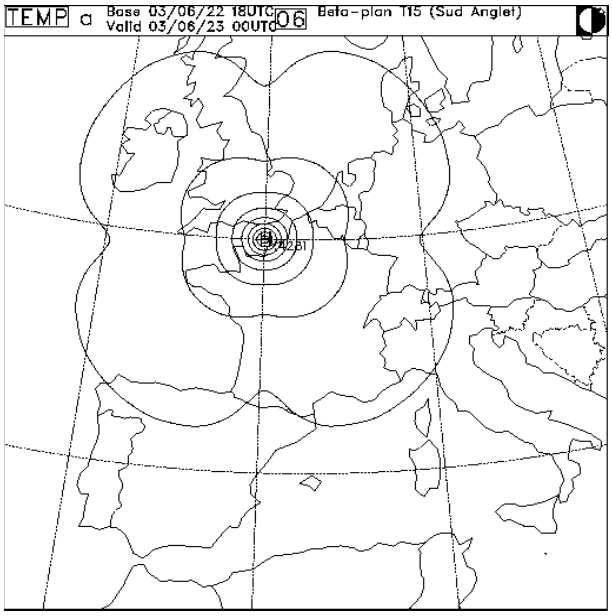


Figure 4 : Temperature increment: Channel

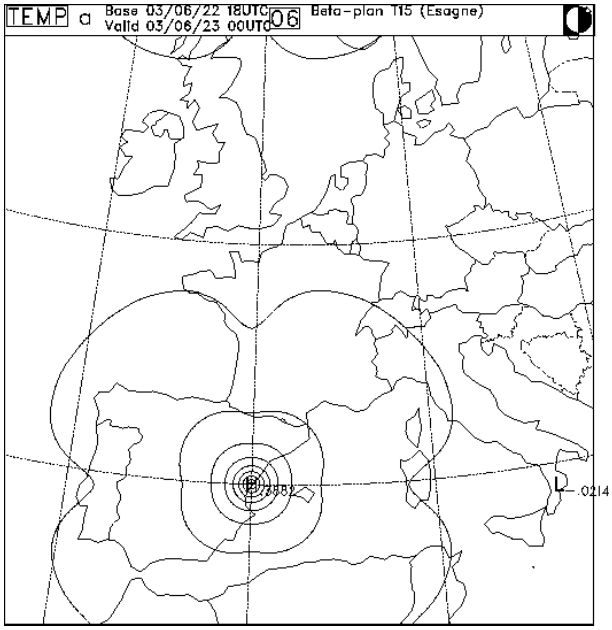


Figure 5 : Temperature increment: Spain

Wind increment

As the equation (2) suggests, the amplitude of the meridional wind increment drops towards the low latitudes. The variation is approximately 10 % between the two positions of observations.

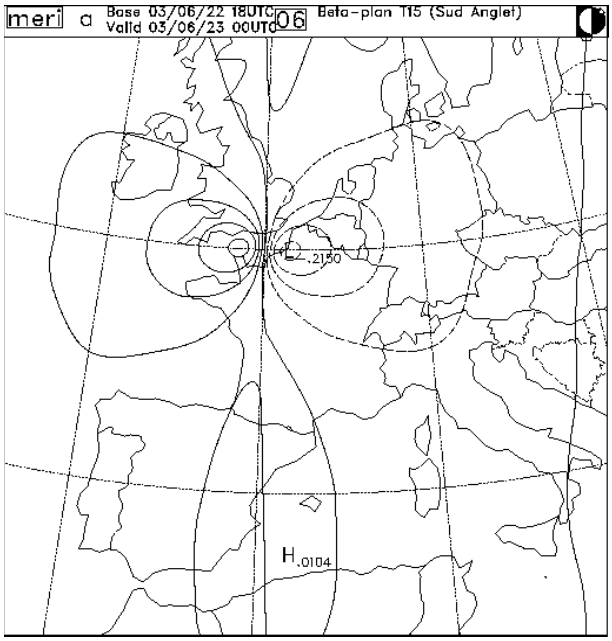


Figure 6 : Meridional wind Increment: Channel

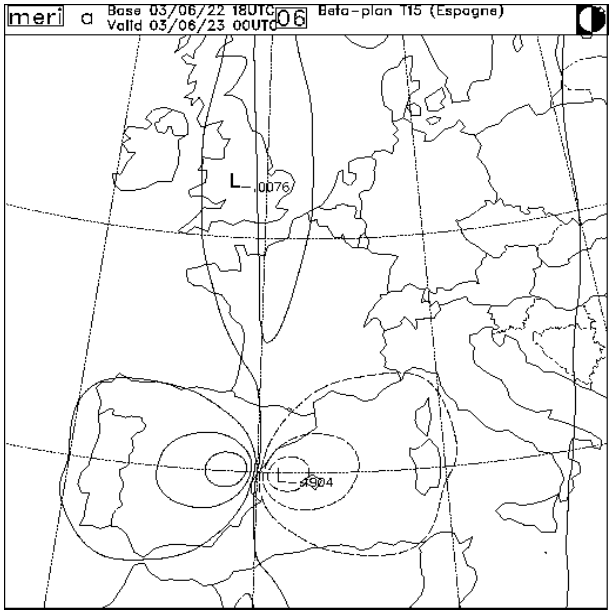


Figure 7 : Meridional wind increment: Spain

5.7.5. Conclusions and perspectives

A β-plane balance has been developed and coded for the Jb term of the ALADIN 3D-Var and for the associated software of error covariance calculations.

The formulation is based on a truncated spectral expansion of the meridional variations of the Coriolis parameter. It can be seen as a multi-diagonal approach, in contrast with the purely diagonal approach of the f-plane balance.

This approach was first validated by examining, over the ALADIN NORAF domain, the increase of explained variance by the β-plane balanced geopotential, compared with the f-plane balanced geopotential.

The formulation was then coded in the ALADIN 3D-Var, and it was validated by using in particular some single observation experiments.

A natural prolongation of this work would be to develop and test other new formulations such as the nonlinear and omega balances (Fisher 2003).

5.7.6. References

Berre, L., 2001 : Représentation des covariances spatiales des erreurs de prévision pour une assimilation variationnelle dans un modèle atmosphérique à aire limitée. PhD Thesis of Paul Sabatier University, Toulouse, France, 250 pp.

Fisher, M., 2003 : Background error covariance modelling. ECMWF seminar on recent developments in data assimilation for atmosphere and ocean, 8-12 September 2003, 45-63.

5.8. Assimilation of the AMDAR data in the ALADIN 3D-Var system

Roger Randriamampianina, Gabriella Csima and Regina Szoták (*Hungarian Meteorological Service)*

5.8.1. Introduction

Aircraft Meteorological Data Reporting (AMDAR) observation system provides a good coverage of measurements in time and space in some region over the globe. The domain of ALADIN Hungary (ALADIN/HU) covers the Western part of Europe where the coverage of the AMDAR data is very high. Impact studies (e.g. Pailleux and Böttger, 2000) show that AMDAR data have positive impact, in general, on the short range forecast.

The pre-processing and the implementation of AMDAR data into the ALADIN three-dimensional variational (3D-Var) data assimilation system at the Hungarian Meteorological Service (HMS) was described in Randriamampianina and Csima (2003).

This report presents the first results of the study on the impact of the AMDAR data on the analysis and forecasts of the ALADIN model.

Section 2 gives brief description of the characteristics of the ALADIN/HU model. Section 3 introduces the pre-processing of AMDAR data. Description of the experiments done for the impact study is shown in Section 4. Section 5 presents the results of the impact study and discusses the efficiency of the assimilation of the AMDAR data in the 3D-Var system. In Section 6 we draw some conclusions and discuss further tasks.

5.8.2. Main characteristics of the ALADIN/HU model and its assimilation system

The hydrostatic version of the ALADIN model was used in this study. The horizontal resolution of the ALADIN/HU is 6.5 km. ALADIN/HU has 37 vertical levels from surface up to 5 hPa. We use the 3D-Var technique in our assimilation system. From the AMDAR observation we assimilate the temperature and the wind components. In this preliminary study the observation errors were the same as in ARPEGE model. Assimilation systems require a good estimation of background error covariance , the so-called "B" matrix, which was computed using the "standard NMC method" (Parrish and Derber, 1992). 6-hour assimilation cycling was chosen, consequently 3D-Var is running 4 times a day at 00, 06, 12 and 18 UTC. We perform a 48 hour forecast once a day from 00 UTC.

5.8.3. 3. Pre-processing of the AMDAR data

The AMDAR data are pre-processed for 3D-Var every 6 hour. The pre-processing interval is ±3 hours. Thus, for producing the 12-hour analysis, for instance, we consider AMDAR data received between 9 and 15 UTC. Figure 1. shows those airports in the ALADIN/HU domain where some AMDAR data are available. Bold dots indicate places (Frankfurt, Cologne, Hamburg, Berlin, Hanover, Bremen, Rome, Paris, Amsterdam, Venice, Istanbul and Budapest) where a big amount of measurements was observed during the study period (2003.02.25 - 2003.03.01). Figures 2a-b present the spatial locations of all measurements corresponding to two (12 and 18 UTC) analysis times. As can be seen in the figures, most of the aircraft measurements are performed over Western Europe.

We use the *oulan-bator-obsort-to-ODB* pre-processing chain to insert the AMDAR data into ODB.

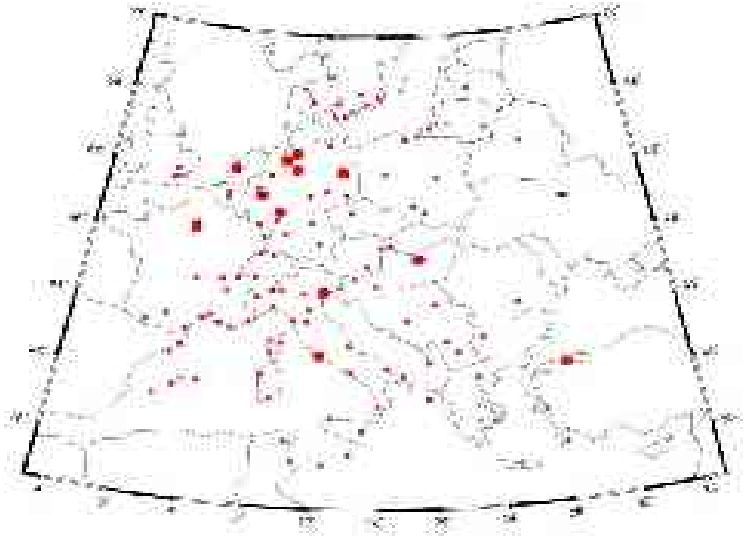
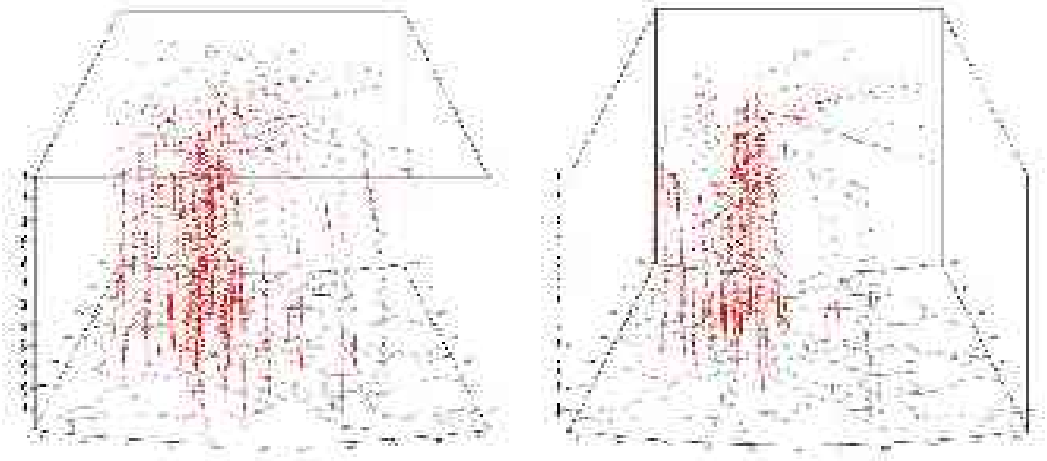


Figure 1.: Location of airports in the ALADIN/HU domain, where AMDAR are available. Bold dots indicate places with big amount of AMDAR measurements.



(a) (b)

Figure 2.: Spatial three-dimensional distribution of AMDAR measurements (during flying, landing and take-off) over ALADIN/HU domain within a ±3 hours interval at 12 UTC (a) and 18 UTC (b) (2003.02.25) assimilation time.

5.8.4. Design of the experiments

In the experiments, two thinning distances (170 and 50 km) were investigated. Thinning distance is the minimum allowed horizontal distance between the location of active observations. 170 km is the default thinning distance in the ARPEGE model and 50 km resolution was chosen based on the results of the investigation on the thinning of the aircraft data described in (Kertész and Fischer, 2001) (Fig. 3.). It is shown on this graph that, at best, we can use more than 50 percent of the incoming AMDAR data by reducing the thinning distance, at least, to 50km. The impact of AMDAR data was studied over a period of about three-week (from 2003.04.18 to 2003.05.07). Surface (SYNOP) and radiosonde (TEMP) observations were used in the control run. The impact was evaluated at both thinning resolutions, comparing the control run and runs with TEMP, SYNOP and AMDAR data. Due to the problem related to the assimilation of the humidity, described in Randriamampianina et al. (2003), experiments were done assimilating the specific humidity in univariate form and together with all control variables (multivariate formulation).

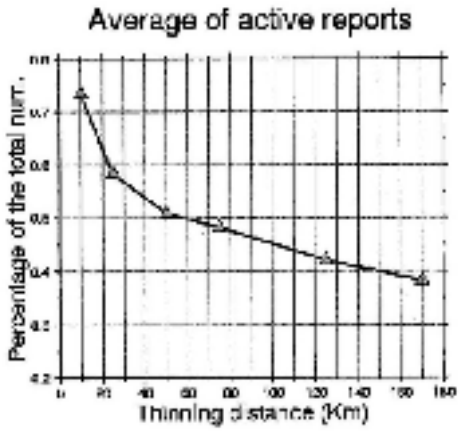


Figure 3.: Percentage of active AIREP data at different thinning distances over the ALADIN France domain.

The following experiments were carried out:

- **am170** - TEMP, SYNOP and AMDAR data were assimilated. The AMDAR data were thinned at 170 km resolution. The multivariate formulation was used for all control variables.
- **amd50** - TEMP, SYNOP and AMDAR data were assimilated. The AMDAR data were thinned at 50 km resolution. The multivariate formulation was used for all control variables.
- **aladt** - TEMP and SYNOP were assimilated - control run. It is our 3D-Var cycling running in parallel suite. The multivariate formulation was used for all control variables.
- **uamd** or **amduh or uamdh** - TEMP, SYNOP and AMDAR data were assimilated. The AMDAR data were thinned at 170 km resolution. The specific humidity was assimilated in form of univariate control variable, when the other control variables were assimilated using multivariate formulation.
- **aluhu** or **auhu** - TEMP and SYNOP data were assimilated. The specific humidity was assimilated as an univariate control variable, when the other control variables were assimilated using multivariate formulation.

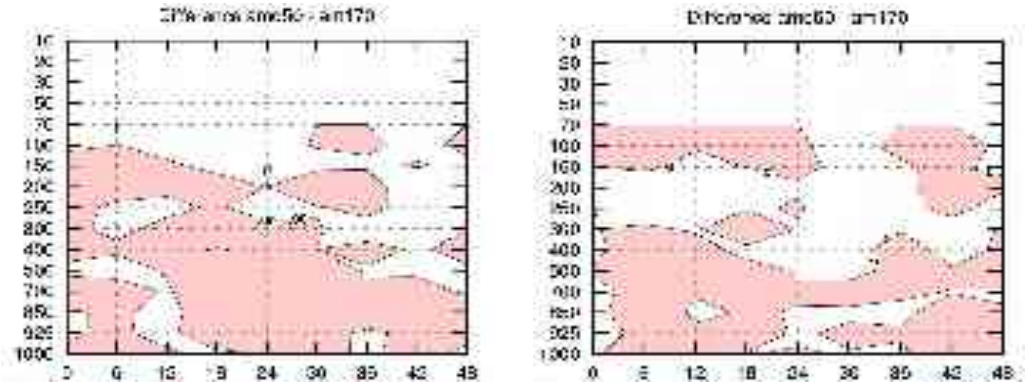
5.8.5. Most important results

5.1 Using the multivariate formulation

Using the multivariate formulation, two impact studies at 170 km and 50 km resolutions were performed with the AMDAR data. Neutral impacts on the analysis and forecast were found in both cases. However, the impact of AMDAR data was slightly better when using AMDAR data at finer (50 km) resolution (Figs. 4a-b).

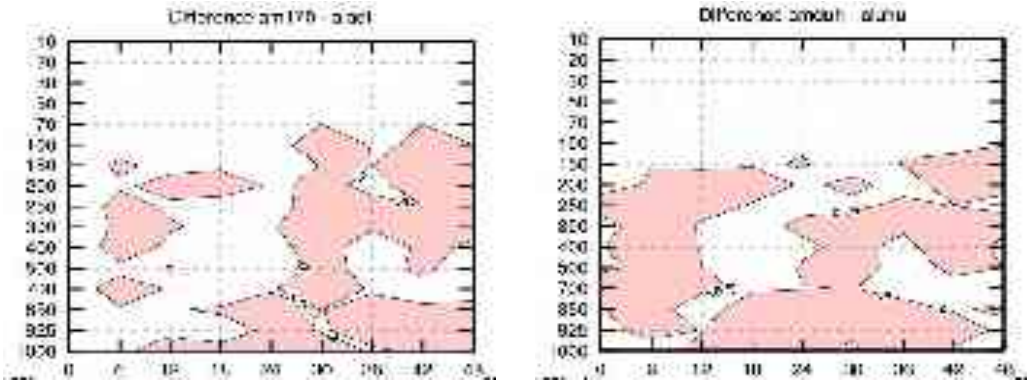
5.2 Assimilating the specific humidity in univariate form

A preliminary experiment was performed to estimate the impact of AMDAR data on the analysis and forecast when assimilating the specific humidity in univariate form. We found a slightly better impact in this case compared to the multivariate one (Figs. 5a-b). It is important to mention that the tinning distance used in this experiment is 170 km, but we found already a big impact in the very short range (6-hour) forecast (Fig. 5b).



(a) (b)

Figure 4.: Difference between the root-mean-square errors (RMSE) for temperature (a) and wind speed (b) : $RMSE_{amd50}$ - $RMSE_{amd170}$.
Coloured areas represent negative values. X and Y axes present the forecast ranges and the model levels respectively.



(a) (b)

Figure 5.: Comparison of the impact of AMDAR data assimilated in 170 km resolution on wind speed, when assimilating the specific humidity with all control variables (multivariate formulation) (a) and in univariate form (b).
These graphs show the difference between the root-mean-square errors (with minus without AMDAR data).
Coloured areas represent negative values. X and Y axes present the forecast ranges and the model levels respectively.

5.3 Efficiency of the assimilation of AMDAR data

Comparing the distance between the observations and first-guess (background) with the distance between the observations and analysis, we can estimate the impact of the observation and the background errors in the analysis. These errors play a role of weighting function in the analysis. Distances were estimated comparing the root-mean-square errors (RMSE) of the observations minus first-guess and that of the observations minus analysis (Fig. 6). The bigger the distance between the two curves, the bigger the impact of the given observation (parameter) in the analysis. One can see, that the impact of the AMDAR data in the analysis of temperature and wind fields is big enough.

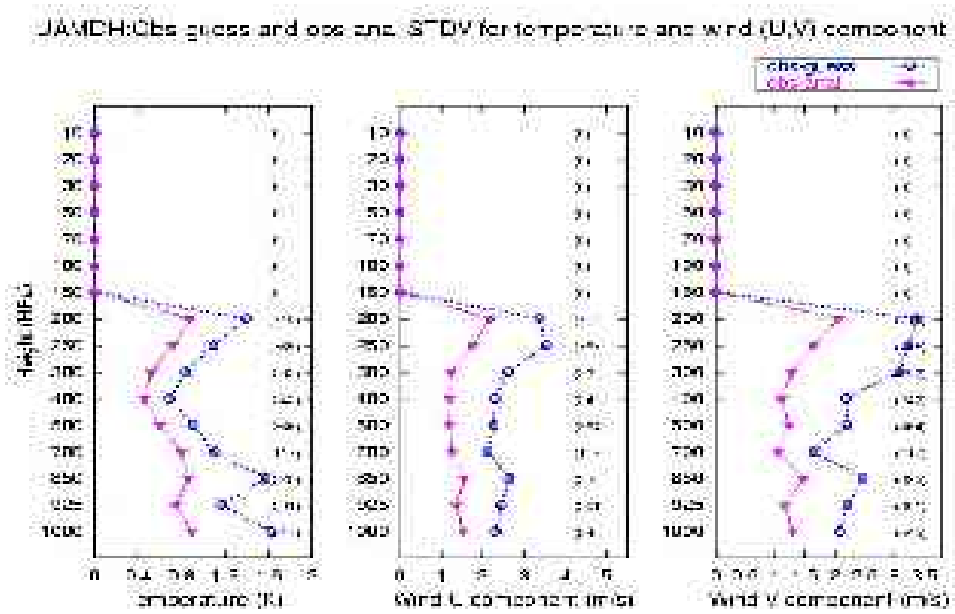


Figure 6.: Standard deviation (STDV) between the observation minus first-guess (line with triangle) and the observation minus analysis (line with circle). Statistics computed for all assimilation times (00, 06, 12 and 18 UTC) during a period of 4 days (2003.04.20 - 2003.04.23) when the AMDAR data were thinned in 170km resolution.

We have also investigated the quality of the first-guess fields available at each assimilation time. We can conclude that, in general, the AMDAR data provide quite good 6-hour forecasts (first-guesses). Figures 7-8 show the improvement in the temperature and wind fields at 06 UTC assimilation time. We have neutral impact of the AMDAR data in the 6-hour forecast from 00 UTC (Figs. 9-10). Note that the quality of the first-guesses was investigated comparing the forecasts with the corresponding ARPEGE long cut-off analyses.

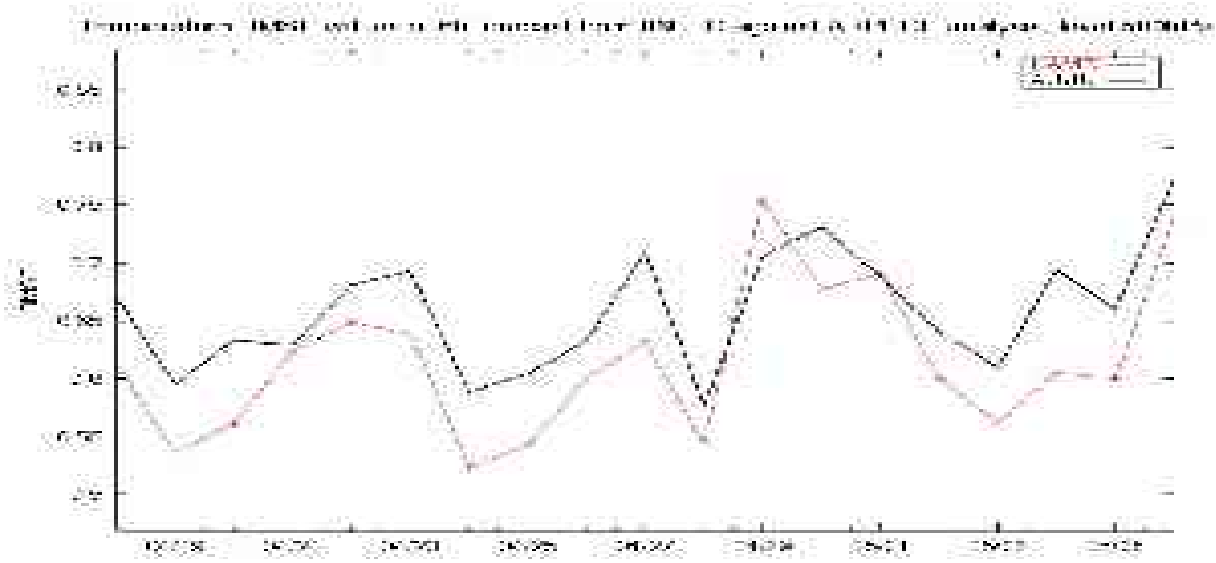


Figure 7.: Root-mean-square error (°C) of 6-hour forecast from 06 UTC of temperature field observed every day. The forecasts of run with TEMP and SYNOP (AUHU – red, thick line) and run with TEMP, SYNOP and AMDAR (UAMD – black, thin line) were compared with the ARPEGE long cut-off analysis.

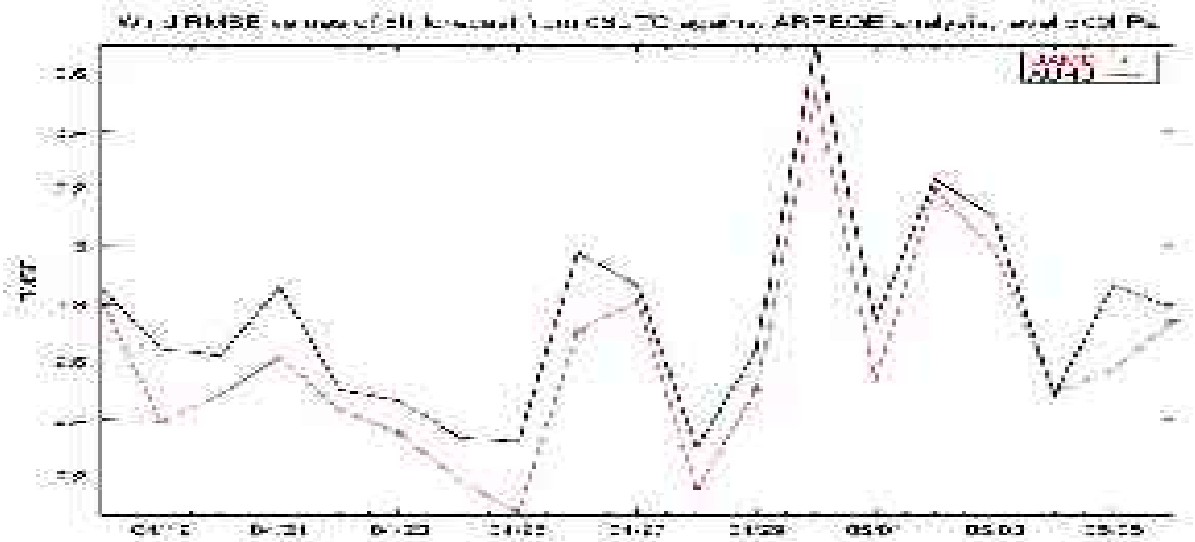


Figure 8.: Root-mean-square error (m/s) of 6-hour forecast from 06 UTC of wind field observed every day. Legend as in Fig. 7.

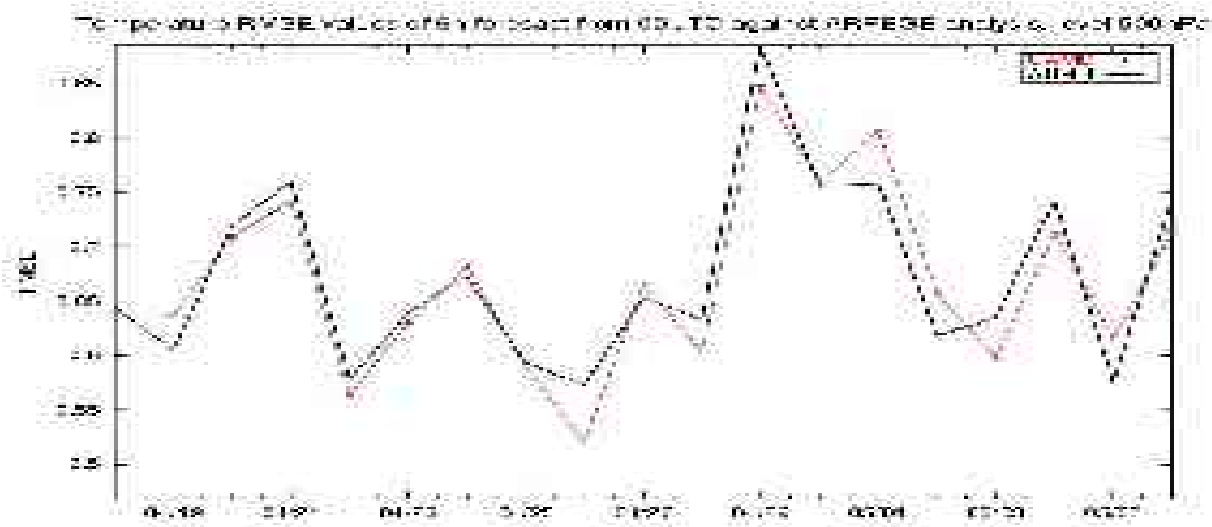


Figure 9.: Root-mean-square error of 6-hour forecast from 00 UTC of temperature field observed every day. Legend as in Fig. 7.

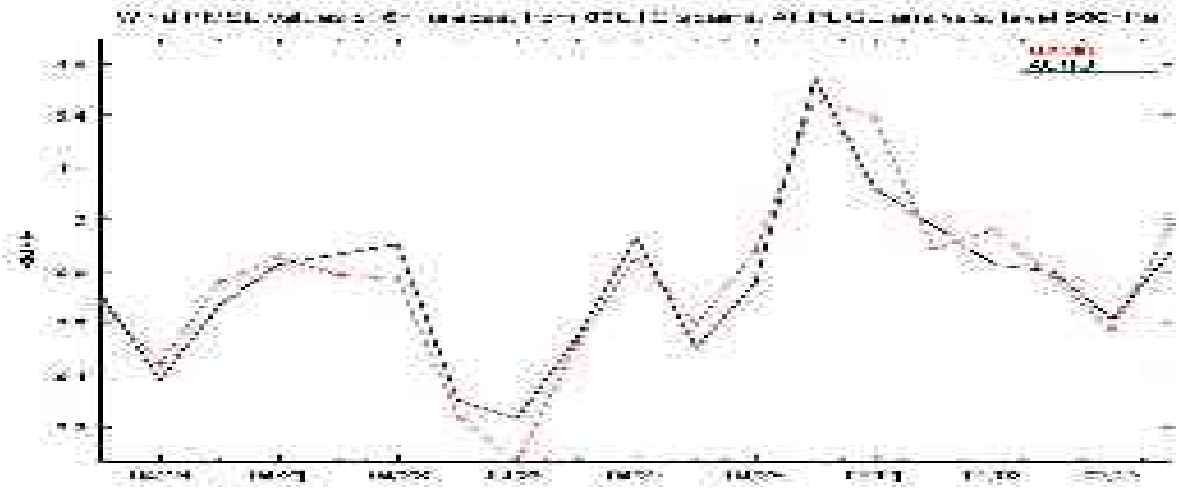


Figure 10.: Root-mean-square error of 6-hour forecast from 00 UTC of wind field observed every day. Legend as in Fig. 8.

We investigated the amount of AMDAR profiles at each assimilation time compared to radiosonde (TEMP) ones. The smallest amount of AMDAR data was found at 00 UTC (Table 1.) for the study period. Thus, the reason of the neutral impact of the AMDAR data could be the small amount of AMDAR profiles at the chosen prediction time.

Analysis time	***00 UTC***	***06 UTC***	***12 UTC***	***18 UTC***
Relative amount of AMDAR data	0.15	4.5	0.80	5.00

Table 1. Relative amount of AMDAR profiles compared to TEMP ones for the period from 2003.04.18 to 2003.05.07. One can see that we have the smallest amount of AMDAR data at 00 UTC.

5.8.6. 6. Summary, further suggestions and experiments

- We concluded that the AMDAR data could provide additional information mainly at 06, 12 and18 UTC analysis times.
- The preliminary impact study showed neutral or slightly positive impact of AMDAR data on the analysis and forecast.
- Slightly positive impact of the AMDAR data on the analysis and forecast was found when assimilating the specific humidity in univariate form.
- It would be very important to study the influence of reduced thinning distances. It is also recommended to investigate the extraction of the AMDAR data within shorter time-interval or assimilate them with shorter (for example 3-hour) cycling.

5.8.7. Acknowledgements

The authors wish to thank Florence Rabier, Élisabeth Gérard, Philippe Caille and Jean-Marc Audoin for their help and fruitful discussions. Thanks are also contributed to Szenyán Ildikó and our colleagues from the System Management Division, HMS, for their help. Partial support for this work has been provided by the Hungarian Scientific Joint Foundation OTKA T031865 and T032466.

5.8.8. References

Pailleux, Jean and Horst Böttger, 2000: Proceedings of the second CGS/WMO workshop on the impact of the various observing systems on numerical weather prediction, 6-8 March 2000, Toulouse, France.

Kertész, Sándor and Claude Fischer, 2001: Observation management for ALADIN,*ALADIN internal report,* available at Météo-France and at HMS.

Randriamampianina, Roger and Regina Szoták, 2003: Impact of the ATOVS data on the Mesoscale ALADIN/HU Model, *ALADIN Newsletter* ***N. 24***.

Randriamampianina, Roger and Gabriella Csima, 2003: Pre-processing of the AMDAR data at HMS, *ALADIN Newsletter* ***N. 24***.

Parrish, D. F. and J. C. Derber, 1992: The National Meteorological Centre's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

CONTENT