IFS/Arpège Memorandum

From: Peter Lean (ECMWF)

To: (ECMWF) DR, RD Division & Section Heads

To: (Météo-France) Arpège diffusion list

To: (ALADIN) Piet Termonia

To: (HIRLAM) Ulf Andræ

File: RD15-xxx

Subject: Minutes of the IFS/Arpège coordination meeting held at ECMWF on 15 June 2015.

Participants:

Météo-France: François Bouyssel, Claude Fischer, Ryad El Khatib (Yves Bouteloup by video-conference)

ECMWF: Stephen English, Deborah Salmond, Yannick Trémolet, Peter Lean, Sylvie Malardel, Robin Hogan, Richard Forbes, Alessio Bozzo

ALADIN: Piet Termonia

HIRLAM: Ulf Andrae

1. Adoption of Agenda

The agenda was adopted.

2. Approval of Minutes of meeting of 12 March 2015

Approved.

3. Review of list of actions from last meeting

1. Recommendations about Fortran/C++/Boost standards for benchmarking and call for tender specifications => for FORTRAN, F90 is mandatory, but support for F2003 is highly recommended for any new HPC for IFS-Arpège-LAM partners. A list of specific F2003 features, agreed in the technical IFS-Arpège videoconferences, is

mandatory for recent cycles (the list is available by Deborah or Claude). For C++ and Boost, Claude to liaise with Yannick. (action)

- Claude confirmed that email exchanges had taken place with Yannick.
- Recommend to keep Fortran document alive and update as necessary. ***Action closed.***

2. A regular cross-exchange of information and possibly coordination decisions should be ensured between the OOPS Project per se, and the IFS-Arpège-LAM code collaboration. It was suggested that this aspect could be discussed at one of the forthcoming OOPS Board meetings. Action on Claude and Deborah.
    - Several issues of coordination will require IFS/Arpège and OOPS cross-cutting discussions, in order to minimize risks of conflicts or code divergence: (1) keep updated a list of ongoing actions in IFS, (2) keep a thorough plan and overview of the refactoring work in the IFS (and sometimes liaise with code experts of specific topics), (3) timing of cycles, (4) general assessment of risk of code divergence. Both sides need to continue to communicate plans which affect the evolution of the two libraries, and be mindful of the risks of code divergence.
3. Steve to send MF material in preparation by Elias Holm, about the impact of the changes in the EDA on the performances of the CY41r2 IFS E-suite.
    - Steve emailed Claude with these details. ***Action closed.***
4. Sylvie to send participants the slides of her presentation about the cubic octahedral grid in IFS. Claude to disseminate in GMAP.
    - Sylvie did send the presentations but it seems likely that they were blocked by a spam filter. **Sylvie will resend or make the presentations available on a shared file system on ecaccess.**
5. Deborah to check with Olivier Marsden for the status of the Python script and the tests for passing FIELDSETs (GMVs, GFLs) by list of arguments in the model code. Contacts at MF: Claude and Alexandre Mary.
    - Now complete and in CY42. ***Action closed***

4. MF information about progress and plans of E-suites and cycles

- Esuite CY40_op2 (operational since 13 April, 2015)
    - Arpege:

- Resolution upgrade from Tl798c2.4 L70 (10km on W Europe) to Tl1198c2.2 L105 (7.5km on W Europe).
- 4DVar : from Tl107c1L70 & Tl323c1L70, to Tl399c1 L105 & Tl479c1 L105
- Forecast length: from t+102h to t+114h
- Calibration in EDA and background error variance filtering.
- Increased number of iterations in minimisations (now 40/40, previously 25/30)
- "Jc_dfi" revision (on surface pressure and divergence only)
- 30 minute timslots instead of 1 hour.
- Increased density of radiances by a factor of 2 (in screening)
- 6 SSMI/S sounding channels of DMSP-F17 and F18
- Use of edge of swath ATMS data
- Assimilation of 6 sounding channels of SAPHIR on Megha-Tropiques
- GPSRO: increased vertical extension, reduced sigmaO and new satellites introduced.
- New CrIS tropospheric channels (extra 22 over sea and 8 over land)
- Radiosoundings in BUFR format
- Radiation computations every hours (instead of 3h)

◦ AEARP:

- Resolution upgrade from Tl399c1 L70; 6 members to Tl479c1 L105 25 members
- 4DVar from Tl107c L70 to Tl149c1 L105
- Background covariances now averaged over 1.5 days and updated every 6h.

◦ PEARP:

- from tl538c2.4 L65 (15km over W Europe) to Tl798c2.4 L90 (10ok over W Europe).
- Perturbations changed from using 6 EDA members and singular vectors, to 17 EDA members and singular vectors. Also introduced a new set of 10 different physics packages (including a new convection scheme "PCMT").

◦ AROME:

        - resolution increase from 2.5km L60 to 1.3km L90
        - data assimilation cycling increased from 3 hourly 3DVar to hourly
            - results in more observations being assimilated
        - forecast length increased from t+36 to t+42h
        - dynamics:
            - introduced new predictor-corrector temporal scheme with one iterations
            - modified semi-lagrangian advection scheme “COMAD”
            - numeriocal diffusion tunings
            - off-centred semi-lagrangian discretization to correct numerical resonance (VESL)
        - physics: auto-conversion, orographic surface drag
        - new orography database (GMTED 2010 at 250m resolution)
    - Forecast impact scores for ARPEGE in CY40_op2 look strongly positive.
    - A new ARPEGE Index (similar to that used at the Met Office) has been introduced:
        - details: Based on 6 normalized EQM scores for T850, Z500, V250 at t+48 and t+72 verified against radiosondes over Europe.
    - PEARP Ensemble verification scores show improvements
    - AROME forecast verification:
        - reduced precipitation frequency bias (previously overestimated convective precipitation at t+18h)
- E-suite CY41T1_op1 (second semester 2015)
    - ARPEGE plans:
        - RTTOV-11 with internal interpolation activated (using new RTTOV coefficients on 54 levels for MW,HIRS,CSR and 101 levels for hyperspectral IR)
        - revised vertical thinning for HR TEMP
        - assimilation of winds from ISS-RapidScat
        - improvements to SSMIS assimilation (including new stratospheric channels 6,7)

        - assimilation of 5 additional CrIS water vapour channels
        - increased density of CSR from geostationary satellites (from 250km to 100km)
        - use of VarBC for ground based GPS observations
        - OSI-SAF sea ice production
        - monitoring of GMI
        - migration from MTSAT-2 to Himawari-8
        - optimisations: new compiler, new options, code modifications
        - Preliminary forecast scores are generally neutral but with some regions showing a slight negative impact.
    - AROME plans:
        - same updates as ARPEGE for RTTOV-11, TEMP, RapidScat, ground based GPSRO varbc
        - increased density of radar reflectivities and winds (from 15km to 8km)
        - monitoing of 2 X-band radars
        - parameterisation of radiation effects linked to orography (slopes, shadows, sky-view factor)
        - updated diagnostics for T2m, snow depth, MSL pressure
        - optimisations (as with ARPEGE)
        - verificaiton:
            - improve first guess departure fits for relative humidity.
    - New regional NWP systems:
        - AROME-Nowcasting:
            - hourly forecasts to t+6h at 1.3km L90 using 10 minute cutoff (pre-operational mid-2015)
        - AROME-Overseas:
            - 2.5km L90 to replace current ALADIN systems (autumn 2015 or 2016)
        - AROME-EPS:
            - 2.5km L90 coupled with PEARP

        - SSPT for model error
        - 12 members, twice per day up to t+42h (pre-operational mid-2016)
- Plans further ahead:
    - E-suite CY42T1 starting mid/end 2016
    - Next resolution upgrade planned for 2017/2018
        - ARPEGE (~5km over W Europe, ~30km over antipodes)
        - PEARP (~8km over W Europe)
        - AROME-France (~1km)
        - AROME-EPS (4 times per day, more members or higher resolution?)
    - Assimilation upgrades in 2017/2018:
        - AROME-EDA (new system based on 3DVar)
        - ARPEGE-EDA (~100 members)
        - AROME and ARPEGE move towards EnVar algorithms.

Technical aspects of co-ordinating the cycles between Météo-France and ECMWF were discussed. Since FLUBs are often not complete they don't provide enough information determine where changes came from. For example, the recent issue of 4DVar running slowly in CY42 took Deborah a week to track down which change had caused it as the FLUB did not provide complete information of which branches modified each routine. It was recommended that the addition of change information in the code comments at the top of each routine should be enforced more strictly and that the FLUBS provide a complete description of the changes going into each cycle.

Action on Deborah to update the coding standards / rules to enforce addition of a comment line at the top of each routine for "non-trivial" modifications.

It was suggested that a Confluence wiki page to share information about known issues, bug fixes and issues with build systems etc would greatly help co-ordination in the joint cycles. **Action on Steve to prototype a new Confluence page to share information on known issues and bug fixes for IFS co-ordination. Claude to test that he can access and use the site once set up. This can be discussed further at the next video-conference.**

The possibility of setting up a toy IFS 4DVar system for testing purposes was discussed, but the general consensus was that maintaining such a system was not feasible. Also, most problems won't show up in low resolution (T21) toy systems.

5. EC information about progress and plans of E-suites and cycles

41r1 went operational on May 12th, 2015.

- Plans for 2016:

  The resolution upgrade is planned to be implemented in January 2016. This could be based on either 41r2 or 42r1. Testing so far has been based on 41r2. Changes in 42r1 and largely technical and not likely to impact the results and so little difference would be expected if 42r1 was used. The resolution upgrade will use the cubic truncation and the octahedral grid in the outer loops (TCo1279). Stability issues with the cubic octahedral grid in the tangent linear model mean that the current grid will still be used for the inner loops in the initial implementation but at a higher resolution than before (TL255-319-399).

  42r2 will be a scientific cycle. The summary below highlights a number of possible contributions.

  - Model:
    - multi-layer snow
    - 2nd set climate fields (new aerosols)
    - ¼ degree ocean (if affordable)
    - unstructured grid (coasts) for wave model
    - revision to SSPT/SKEB
    - improvements to boundary layer shallow convection
  - DA:
    - Tco399 innter loop (once stability issues in TL for octahedral grid are resolved)
    - land surface analysis in 4DVar outer loop
    - improvements to aircraft bias-correction
    - stratospheric balance constraint
    - observation error retuning
    - re-introduction of stratospheric weak constraint
    - Aeolus capability
    - VarBC changes for reanalysis
  - System configuration:

        - HRES: 4DVar Tco1279/ TL255-319-399 (t+10d) [41r2 or 42r1]
        - EDA: Tco639 / TL191-191 [41r2 or 42r1]
        - Ens: TCo639 (t+10d) / TCo319 (t+10 - 46d) [41r2 or 42r1]
        - 50 member EDA or similar reconfiguration of Ens in DA
        - Optimisation e.g. I/O server

- Plans for 2017/2018

  A number of new system configurations are being investigated including having four overlapping data assimilation windows per day, weak constraint 4dvar, long window 4dvar and changes to the ensemble data assimilation. The testing of these configurations will be made substantially easier with OOPS and so many of the decisions in this area are awaiting developments in that project.

    - Model:
        - CHTESSEL
        - Interactive aerosol
    - DA:
        - Cloud control variable
        - Weak constrain (random error)
        - Seamless EDA-EPS
        - Ocean DA weak-coupling
        - COSMIC-2, SMAP, FY-3C, Metop-C, Aeolus, EarthCARE, Jason-CS, JPSS-1
    - System configuration:
        - HRES to t+15d?
        - 4 overalapping DA windows per day?
        - S5 seasonal prediction system

The collaboration between MeteoFrance and ECMWF on the use of SAPHIR data was noted.

6. HIRLAM comments

Problems in CY40 were tracked down to changes in a linear algebra routine (MINV). Previously, when values close to zero were detected the routine gave a warning, but now it

aborts. This happened in the optimal interpolation scheme used by the surface assimilation. It is not clear if this issue was introduced in CY39 or CY40.

E-suite CY40 has now started and is planned to go operational in late Autumn 2015.

The next cycle will be CY42T1.

Several people will be contributing to the OOPS refactoring primarily on observation operators associated with observations relevant for mesoscale models (e.g. radar, GNSS, MODE-S).

Claude mentioned that some of the Hirlam suggested contributions to the next T1 cycle should be described and discussed within the ECMWF/MF/Aladin/Hirlam coordination meetings. For instance, the time step first order extrapolation (or not) option (by M. Hortal), the radiation code changes for ice aspects (Emily Gleeson, K. Pagh Nielsen, L. Rontu), the changes in the B-matrix normalized perturbations in obs space (J. Bojarova).

7. ALADIN comments

Work is ongoing to clean up the number of configurations that are used for testing.

The interface with the radiation scheme was discussed.

It was noted that good co-ordination had taken place between members of the dynamics teams.

8. Specific issue:

8.1 OOPS (preparatory discussion for project board only)

Yannick presented an update on the latest OOPS developments.

SPAM; Olivier Marsden has produced a python script which automatically passes many variables by argument to remove the current globals from modules. So far this had been done for the 5 'field' derived types in CY42. Next to do are Geometry and Model derived types.

MKGLOBSTAB : Tomas Wilhelmsson has been working on refactoring the code which interpolates model fields to observation locations; passing by argument.

In the next phase Tomas will focus on Non-linear aspects and Olivier will focus on the TL & AD and trajectory re-factoring.

Plans need to be made on whether to refactor the current trajectory code or write a new one from scratch. If new trajectory code is written, will this also be put back into the IFS? How will code divergence be managed?

July meeting in Toulouse to make detailed plans for the observation operator refactoring work.

.

8.2 Pre-OOPS cleaning/refactoring of observation processing code

Peter Lean gave a recap of the current effort under way to clean and refactor the observation operator code prior to the OOPS refactoring. The aim is to untangle the model and observation code.

Currently, the observation code calls lots of model routines related to the model grid (for interpolation purposes). This dependence of the observation code on the model is removed with a new GOM_PLUS derived type which stores everything needed from the model by the observation code. This allows the PREINT routines to be deleted from HOP.

Similarly, the model code currently uses ODB in the derivation of the GOMS. This dependence on the observation database can be removed by simply passing the latitudes, longitudes, pressures and times to the model; it has no need for direct ODB access.

Further code cleaning (e.g. KDLEN / ILEN, using varnos throughout instead of a combination of varno, NVNUMB and CLU variables) and refactoring making a new observation operator subroutines (obsop_rad.F90) has been undertaken.

Many of these changes (all bit-comparable) have been accepted into CY42R1.

An improved ODB/IFS interface has been developed to replace the existing GETDB/ROBHDR/ROBODY access. New database and view derived types encapsulate the data making it much clearer where the observation data is coming from. A python script is used to upgrade to the new interface and pass the ODB data around by argument list removing the current ROBDHR/ROBODY globals.

Further work remains to be completed by the September deadline when the Météo-France OOPS observation operator refactoring work begins. These changes will be a branch on top of CY42. The planned meeting in Toulouse in July will be used to co-ordinate these changes; in particular care will be needed to ensure that CY42T1 developments do not break any of the changes in this preOOPS cleaning.

8.3 Use of COMAD / SLHD_HEAT in the IFS

Sylvie Malardel described the instability in the tangent linear code for the cubic octahedral grid. SLHD_HEAT applies diffusion to the grid point humidity field and had been used in sensitivity tests to help understand this issue, but would not be used in operations.

“Rain bombs” in 41r1:

- since the release of 41r1, some cases of localised very high precipitation (more than 500mm/24hr) have been seen in situations of light winds over moderate orography.
- the use of COMAD largely fixes these grid point storms
- however, these “rain bombs” aren't seen in TCo1279 forecasts and so the issue disappears and COMAD is no longer needed.

8.4 Current issues with radiation in the IFS

Discussions were held related to a recent bug in the new radiation scheme update of CY41R1 which only manifested itself in the configuration used by Météo-France. This has now been resolved by Alessio Bozzo and Yves Boutleoup; the latest Météo-France results are in line with those expected from the new scheme.

Robin Hogan described recent improvements in the radiation calculations near coastlines:

- the radiation scheme isn't run at every grid point
- near coasts this can lead to inaccuracies
- example presented where excessive surface cooling occurred in clear skies near the coast caused by spatial interpolations of skin temperature and surface emissivity
- new scheme updates long and shortwave fluxes every timestep and every grid point in response to surface albedo and skin temperature.

Current developments:

- aerosol from MACC climatology
- revised chemistry climatology from MACC simulations
- better surface albedo (solar zenith angle dependence)

8.5 Medium term plans for a new radiation code in the IFS

Robin Hogan presented future plans for the radiation scheme:

- SPARTACUS: “Speedy Algorithm for Radiative TrAnsfer through CloUd Sides”
    - a model of 3d radiation effects which can lead to biases
    - approx x1.5 cost of a standard 2d radiation scheme, but with capable of modelling many of the 3d effects related to cloud sides.
- Flexibility of radiation code:

- ECMWF uses RRTM-G which is expensive but accurate; radiation is only calculated every 6th grid point and not every timestep. Other users may want a different configuration, but current scheme has lots of these things hard coded.
- Aim to make the code more configurable in terms of:
  - gas optics
  - cloud optics
  - aerosol optics
  - solver (McICA, SPARTACUS, Triple clouds etc)
- More object oriented / modular
- Open source release planned.

- Timescales:
  - 42r1 : new aerosol climatology
  - Summer 2015:
    - modular radiation code
    - investigate global impact of 3d radiation effects
  - Spring 2016:
    - test faster RRTMGP (better exploitation of parallel architectures)
  - Summer 2016:
    - Open source radiation code release

Given the now significant plans of evolution of the radiation codes, and since MF and other LAM partners do use parts of the IFS radiation codes in their own applications, regular scientific and technical exchanges of information between ECMWF, MF and the LAM groups are required. This would generally take place through direct email exchanges between the contact persons (for MF, Yves Bouteloup, for Hirlam, Laura Rontu, for Aladin, Jan Mašek[1]), but occasionally, the topic of radiation changes could be brought up again at coordination discussions.

8.6 Tests for the next resolution upgrade with cubic octahedral grid

1 Yves.bouteloup@meteo.fr, Laura.Rontu@fmi.fi, jan.masek@chmi.cz

Richard Forbes presented the latest forecast impact results.

The TCo1279 has a resolution of 8km at the equator and around 10km near the poles. The octahedral grid brings much of the benefit of the TC1279 model but at an affordable cost; the number of grid points is around 5M compared with 8M for the TC1279. The kinetic energy spectrum of the TCo1279 model is very similar to that of the full TC1279 model, with substantially more energy at high wavenumbers than TL1279, particularly for scales below 60km.

There is an issue with the stability of the cubic octahedral grid in the tangent linear model, consequently the inner loops will remain on the linear grid albeit at a higher resolution than before (TL255/TL319/TL399 compared with TL255/TL255/TL255 before). The aim is to move the inner loops to TCo255/TCo319/TCo399 in a later cycle once the tangent linear instability issues have been resolved.

Forecast impact looks very positive; with reductions in Z500 RMS error on the order of 5-10% in the first 5 days of the forecast. These improvements come from the combination of science upgrades in 41r2, higher resolution data assimilation and the higher resolution model itself. The impact of the higher resolution model alone were still positive, but less than the combined impact.

One negative impact was in the stratospheric temperature; a cold bias in the model has become even colder.

The impact on forecasts of tropical cyclones, near surface variables, ensemble and seasonal forecasts is still ongoing.

8.7 FFTW code in the IFS

Issues relating to the GNU GPL license in the FFTW library were discussed. There are concerns that use of FFTW obliges us to release the rest of the IFS source code under the GNU GPL license. In addition, ECMWF products on the cubic octahedral grid would also require partners and customers to use FFTW to transform from the spectral to grid point space.

The ECMWF legal team is studying the details of the license, but have yet to report back. There are concerns that this issue could delay the implementation of the resolution upgrade. A possibility remains to pay a license fee to MIT to avoid these conditions.

Action: Steve to follow up with the ECMWF legal team and report back in light of his advice.

8.8 MF optimisation work for BULL or Cray

Ryad El Khatib presented benchmarks of IFS for binaries built with different compilers; the Cray, Intel and GNU compilers were compared. These comparisons were performed on the Cray XC30. Prior to the benchmarks, optimisations to minimise the use of the Lustre file system for small files were made. This reduced the variability of the run times to within +/- 2% allowing reliable benchmarks to be performed.

Preliminary results showed that the Cray compiler was about 20% faster than the Intel compiler, which itself was 2% faster than the GNU compiler.

Then, the impact of specific optimisations was investigated:

i) array syntax: the Cray compiler was good at optimising array syntax, but others were not. Re-writing in Fortran 77 syntax led to a 10-15% speedup in the Intel and GNU compilers, helping to close the gap with the Cray compiler.

ii) vectorization of loops helped get speedups of 20 to 300%

iii) removal of allocation and deallocation cycle led to >200% speedup

An optimised ARPEGE cycle (41T1+) was made using the optimisations described above; this led to a 5% speedup with the Cray compiler, 16% speedup with Intel and 8% speedup with gfortran.

The overall conclusion was that the compiler does make a difference. The Cray compiler allowed the best performance with the least optimisation effort.

MF further confirmed that the first of two BULL Phase 2 clusters was to be installed in Toulouse in January-February 2016, followed by a verification of aptitude period of two months (presumably March-April, if no specific delay occurs). It was expected that during the construction time of the Phase 2 cluster, MF scientists and developers would have very little access to the HPC resources, since only one Phase 1 cluster would be operating, including the full operational suites. The second Phase 2 cluster would be installed in July-September 2016.

8.9 Cross-exchange of information about MF and EC Phase 2 HPC upgrades

Likely configurations for the Phase 2 HPC upgrades were discussed.

9. Content and timing of cycles

Joint cycle	ECMWF	MF	Deadline	Declaration	Misc. / Oper plans
CY42				June 2015	
	CY42R1		15 June 2015	July 2015	Could be implemented with resolution upgrade but not essential
	CY42R2		Sep-Oct 2015	TBC	
		CY42T1		Mid-Oct (start 12/10 ?) through mid-Dec 2015	Dates to be confirmed
	CY42R3		March 2016	TBC	
CY43			March 2016	April-June 2016	Start of merge to be confirmed; MF will need to complete the verification of aptitude of its BULL Phase 2 cluster by end of April

For CY42T1, co-ordination will be required in Jan/Feb to make sure that any changes won't break the automated scripts associated with the preOOPS refactoring.[2]

10. AOB

11. Next meetings

Next technical video-conferences:

- Not yet planned

Next Coordination video conferences:

- Thursday 15 October 2015, 1.30pm UK / 14h30 MEST
- beginning of March 2016 (date to be discussed later)

Next Coordination Meeting in Reading:

- Monday 13 June 2016 in Toulouse (provisional)

2 **As an important post-meeting information, an alternative scenario was discussed between EC and MF after the OOPS Board meeting the next day (16 June). In this scenario, a CY43 would be built in the autumn 2015, thus much earlier than in the plan presented at the coord meeting and reported in the Table above.**

List of actions

1. **Sylvie to resend her presentations on cubic octahedral grid or make them available on a shared file system on ecaccess.**

2. **Deborah to update the coding standards / rules to enforce addition of a comment line at the top of each routine for non-trivial modifications.**

3. **Steve to prototype a new Confluence page to share information on known issues and bug fixes for IFS co-ordination. Claude to test that he can access and use the site once setup. This can be discussed further at the next video-conference.**

4. **Steve to follow up with the ECMWF legal team about licensing issues for FFTW and report back to the project board and partners in light of his advice.**