ALADIN - HIRLAM NEWSLETTER
No. 3, September 2014
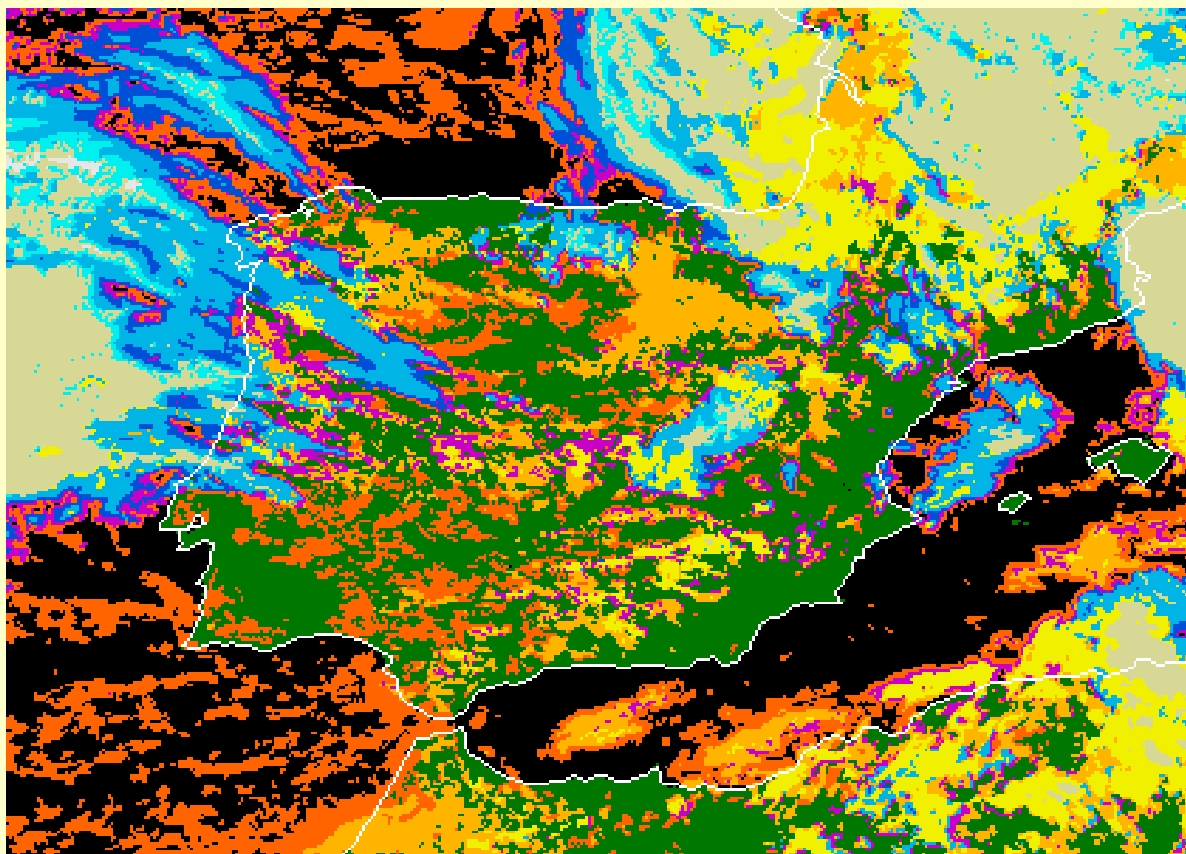
NWC-SAF Cloud type product used by D. Martin et al.

Contents

Scientific contributions

Organisational information

Introduction

Tilly Driesenaar and Patricia Pottier

We are glad to present the third edition of the combined Newsletter of the HIRLAM and ALADIN consortium. This edition is dedicated to the joint ALADIN 24th Workshop / HIRLAM All Staff Meeting 2014. We thank the authors for their contributions!

Also a list of events for the near future is added.

For any additional information please visit the ALADIN and HIRLAM websites, or just ask the authors.

Enjoy reading!

http://www.cnrm.meteo.fr/aladin/

http://hirlam.org

Conference diner

Experiences with scatterometer assimilation in HARMONIE in a severe storm case

Teresa Valkonen
The Norwegian Meteorological Institute, Oslo, Norway

1 Introduction

The fellowship project *Scatterometer winds in rapidly developing storms* (SCARASTO) of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) aims to take better benefit of scatterometer winds in situations of rapidly developing storms. The focus is on weather systems with rapid or unexpected development, strong winds or large damage potential. This includes both synoptic-scale mid-latitude cyclones and small-scale and short-lived low pressure systems, such as polar lows, with prevailing Atlantic or Northern flow direction.

Several weather forecasting centres apply successfully scatterometer wind data assimilation in their global numerical weather prediction (NWP) models. Positive impacts have been reported also from limited area models, such as from the High Resolution Limited Area Model (HIRLAM) [3, 5]. Despite the option to use scatterometer winds in the model system of the Hirlam Aladin Regional/Mesoscale Operational NWP In Europe (HARMONIE), the scatterometer wind assimilation is not operationally applied at the Norwegian Meteorological Institute (MET Norway). This work gains preliminary experience needed to progress towards the operational use of scatterometer winds in Norway using the HARMONIE system.

This short study describes one selected storm case on 16 November 2013. The storm event is considered to be severe and one of the strongest reported in Norway [2]. Scatterometer surface wind data over oceans from MetOp-A and MetOp-B satellites are applied in the three dimensional variational (3DVAR) data assimilation system available in the HARMONIE model. Quality and coverage of scatterometer wind data, and impact on forecast performance are studied. Specifically, the aim is to

- investigate the temporal and spatial scatterometer data usage in the HARMONIE model system,

- learn about observation and first guess accuracies by monitoring the scatterometer departures from the model first guess, and

- investigate the impact of scatterometer wind assimilation on the forecast in this storm case.

2 ASCAT scatterometer ocean surface winds

Advanced Scatterometer (ASCAT) is an active microwave instrument on-board satellite MetOp-A and MetOp-B, launched on 19 October 2006 and on 17 September 2012, respectively, by the European Space Agency (ESA) and operated by the EUMETSAT. The ASCAT instrument was designed to measure wind speed and wind direction over the oceans [4]. The ASCAT system covers two 550 km-wide swaths, which are separated by 670 km. The orbit period is approximately 90 minutes. The paths of MetOp-A and MetOp-B partly collocate, and they are 50 minutes apart in time.

The ASCAT products describe a horizontal equivalent neutral wind vector at 10 m height. The data range is 0-50 m/s, however, the wind speeds over 25 m/s are generally known to be less reliable [1]. Generally, the wind component Root-Mean-Square errors (RMSE) are smaller than 2 m/s and the bias less than 0.5 m/s. There are 5 different ASCAT scatterometer wind products available through the Ocean and Sea Ice Satellite Application Facility (OSI SAF). Global OSI SAF coastal wind product is used in this study.

3   Experimental setup

The model experiments for the period 11/11/2013 – 18/11/2013 were made applying the HARMONIE model system, version cy37h1.2, which was the operational model version at MET Norway at the time. Results are based on two experiments. The main experiment included data assimilation of conventional observations and ASCAT winds, and the second one only conventional observations. The model domain was the operational model domain used in Norway, covering roughly Scandinavia, Finland and the Baltic states (Figure 1). The grid size of the model is 2.5 km. The ECMWF forecasts were used as lateral boundary conditions. The 3DVAR data assimilation method was applied in 3-hourly assimilation cycling with data window of 3 hours. No First Guess at Appropriate Time (FGAT) scheme is available, so all observations in the data window are handled as if they were taken at the analysis nominal time. Data latency was not taken into account in this study. The option with blending of 3-hour HARMONIE forecast and the ECMWF forecast to obtain a background field was applied.
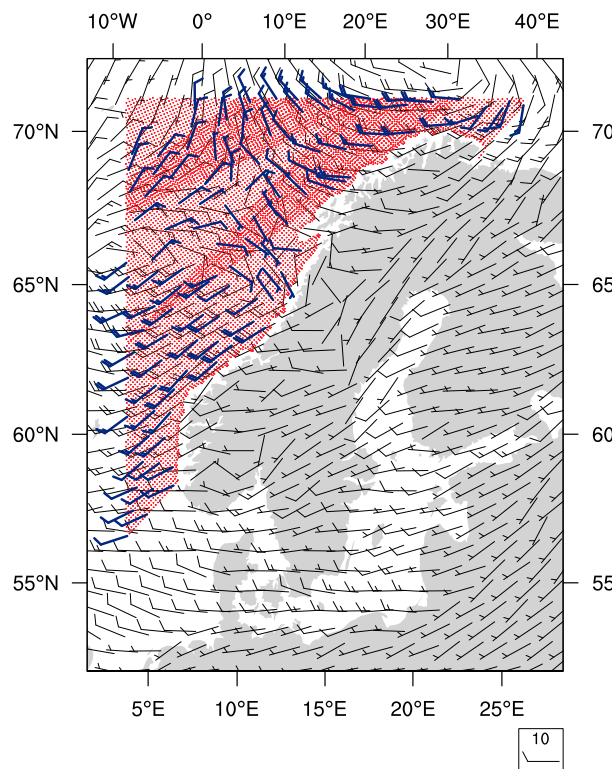

Figure 1: An example of ASCAT data coverage over the model domain at 16 November 2013 12 UTC cycle. The red dots are the ASCAT data points, the blue wind barbs are the used ASCAT winds, and the black wind barbs are the background winds.

4   Results

4.1   Data usage

The available ASCAT data coverage in the domain varied a lot depending on the time of the day (Figure 2a). Most ASCAT data were available for cycles 09UTC, 12UTC, 18UTC and 21UTC, whereas no data were available for 00UTC, 03UTC and 06UTC cycles. As the nominal MetOp orbit repeat cycle is 29 days, there are some day-to-day variations in the swath locations and thus also in the number of data hitting the model domain. The daily variations were, however, rather small and the diurnal distribution of data amount in this study gives a good estimate of the data amount in the domain in general.

Observations are thinned before they are taken into the data assimilation scheme. The default thinning factor for ASCAT is set to 4 leading to the thinning distance of approximately 100 km. Besides thinning, there are first guess check and quality control procedures, which lead to additional data reduction. The total fraction of used observations was about 1% of all available ASCAT observations in this experiment (Figure 2). On cycles 09 UTC and 21 UTC, the fraction of used data was clearly smaller than on other cycles. The reason for the excessive data rejection on these cycles is still under investigation.

Figure 1 illustrates the ASCAT data usage over the domain when the severe storm had arrived into the model domain. The strong thinning and data rejection can be seen, as well as an individual ambiguity selection problem (around 65°N, 10°E). The found ambiguity selection problems were mostly related to proceeding fronts as the 3DVAR data assimilation system cannot make use of the time of observations. Gross error problems, for example due to heavy precipitation or instrumental technical issues, were not found in the assimilated ASCAT data during this study period, which indicates that the data rejection procedure works effectively.

a) b)
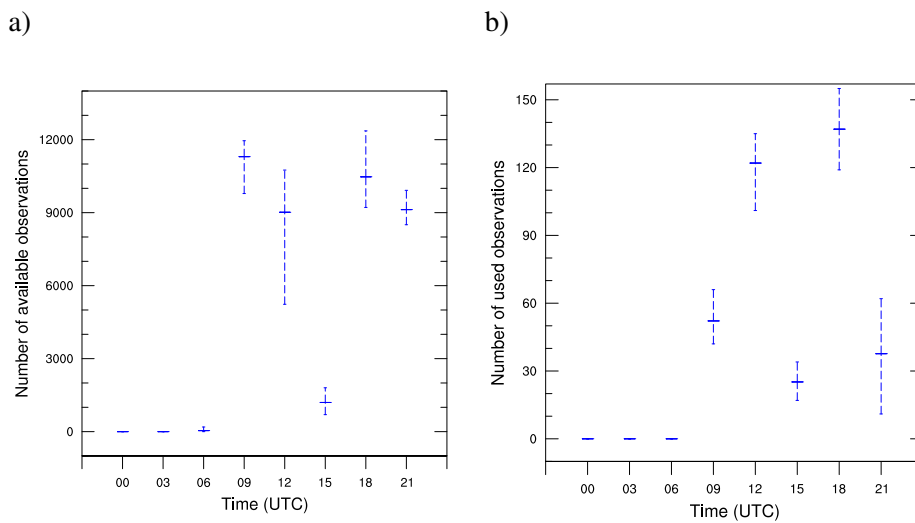
Figure 2: a) Number of ASCAT data points read into the model and b) number of ASCAT data points used in the data assimilation on each cycle. The horisontal line marks the average number, and the vertical dashed bars the range for each model cycle.

4.2   Observation departures

Influence of adding ASCAT winds in the model system can be assessed in terms of the model fit to the scatterometer observations. Figure 3 shows the histograms of background and analysis departures for assimilated
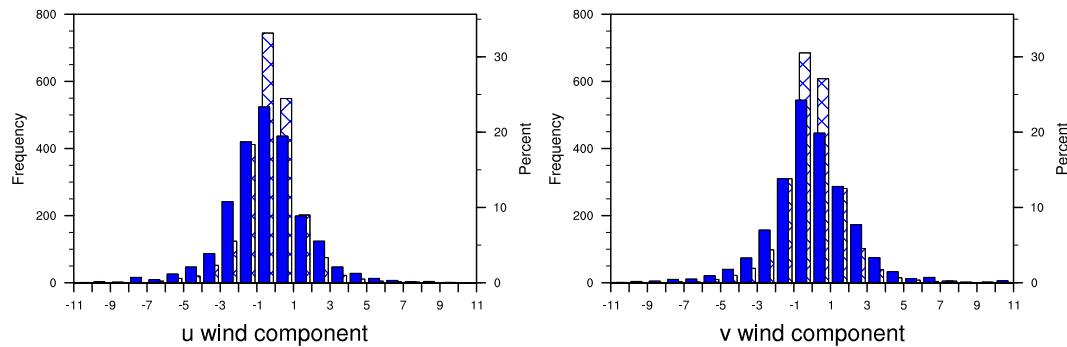
Figure 3: Frequency distribution of background departure and analysis departure for u and v wind component (m/s). The blue bars indicate background departures and the hatched bars analysis departures.

ASCAT winds for u and v wind component. There is a reasonable reduction of bias (from -0.51 m/s and -0.09 m/s to -0.34 m/s and -0.05 m/s for u and v respectively) and standard deviation (from 2.23 m/s and 2.48 m/s to 1.58 m/s and 1.74 m/s for u and v respectively). The standard deviation of background and analysis departure varies with time which means that the model fit to scatterometer observations is most likely dependent on the flow pattern. The standard deviation of background departure up to 3.7 m/s was found at the time of the strongest cyclone activity.

Mid-latitude cyclones, such as the storm of the interest, are associated with high wind speeds. It is therefore essential to be able to model the ocean surface winds accurately. In this storm event, the model background winds were always higher than the assimilated scatterometer winds when background wind speeds were higher than 23 m/s. This could indicate problems in the model parameterisation of the boundary layer over oceans. However, the number of assimilated observations with such high wind speeds were limited in this study, and longer simulation periods are needed to find out if the model winds are persistently lower than scatterometer winds in case of strong winds.

4.3   Impact of ASCAT wind assimilation on the forecast

Figure 4a shows the mean sea level pressure (MSLP) statistics by forecast length from the two experiments averaged over all stations in the model domain. The experiments with and without ASCAT wind assimilation are very close to each others. Normalised RMSE difference, up to 0.01 hPa, between the experiments is, however, statistically significant on the forecast lengths of 6-23 hours (figure not shown). Time series of 12-h forecasts (Figure 4b) shows that the forecast errors are largest during the most intense storm event (16-17 Nov). During that time, the data assimilation of ASCAT winds improves the RMSE and bias of MSLP. This demonstrates that ASCAT winds have a potential to improve the forecasting of storms.

In addition to MSLP, other surface variables did not show any significant changes. Near surface temperature and wind fields over land where the verification stations are located are strongly driven locally, whereas MSLP and atmosphere well above the surface are influenced by larger scale circulation which can be improved by scatterometer winds over oceans. There were minor improvements in humidity and temperature fields at 700 hPa level but those changes were not statistically significant. Longer study periods are needed to find out the general impact of ASCAT winds on higher levels of atmosphere.

All available SYNOP and sounding observations in the model domain were used for the impact evaluation. It is known that the impact of ASCAT data assimilation is most pronounced in the areas near the storm. It is, therefore, expected that an evaluation with a smaller selection of stations would lead to clearer impact of ASCAT winds.

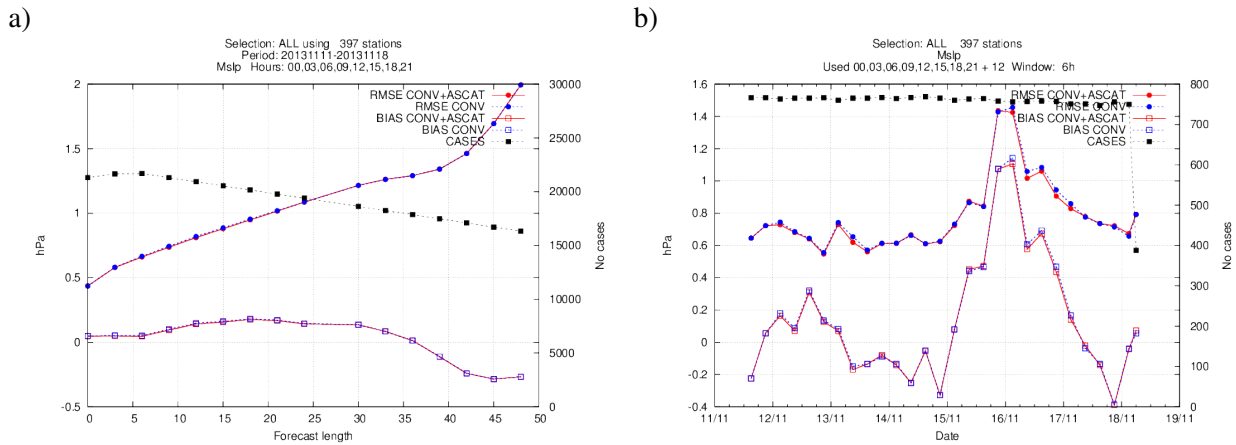a)   b)

Figure 4: RMSE and bias of observation minus forecast for MSLP from two experiments a) as a function of forecast length and b) as a time series of 12-h forecasts averaged over all the SYNOP stations in the domain. The main experiment using both conventional and ASCAT winds is denoted with CONV+ASCAT on red, and the reference experiment with conventional observations only with CONV on blue.

5   Conclusions and outlook

The Norwegian seas are well covered by scatterometer observations during morning and evening cycles (09, 12, 18 and 21 UTC). However, the sea region inside the domain is relatively narrow, and a rapidly proceeding storm can be observed in the domain area only a few times before it hits the land. Scatterometer winds are quality checked and thinned before entering the data assimilation of the model. In this experiment only about 1% of the incoming ASCAT data was used. The thinning procedure was originally designed for coarser resolution data. One might expect that use of more data could be beneficial in a model with such a high resolution, and this should be investigated. The presently applied ASCAT assimilation in HARMONIE is based on an initial configuration not believed to be optimal for the present resolution and configuration of the model.

Individual ambiguity selection problems occur close to proceeding weather fronts. They are mainly related to the time difference between the observation and analysis time. 3DVAR with FGAT or four dimensional variational (4DVAR) data assimilation method would likely reduce the problem. In the current system, FGAT or 4DVAR are not available options. Instead, a shorter assimilation window is planned to be tested. A shorter window would let less data in but they would be more representative for the analysis time, which would likely lead to less ambiguity selection problems.

The experiments described here make use of a static background-error covariance matrix (**B** matrix) based on average structures. This is probably not optimal near strong low pressure systems, where smaller spatial scales and baroclinic structures are expected. Revisiting the current **B** matrix estimation might improve the results to some extent. Furthermore, 4DVAR and hybrid assimilation methods could help on longer term.

Despite the strong thinning and data rejection, verification results from the model experiments show a weak positive impact of scatterometer winds on the MSLP forecasts on the forecast lengths of 6-23 hours. It is worth to note that the period of interest here is only one week, and results cannot be interpreted to apply generally. Nevertheless, achieving neutral and positive impacts in this case demonstrates the potential of scatterometer wind assimilation in high resolution NWP models. Longer experimentation periods are planned for finding out the overall impact of scatterometer winds on forecasts.

Acknowledgements

References

[1] ASCAT Wind Product User Manual version 1.13. Technical Report Reference: SAF/OSI/CDOP/KNMI/TEC/MA/126, Ocean and Sea Ice SAF, 2013.

[2] Report of extreme weather event: Case Hilde 16-17 November 2013. [Ekstremvêrrapport. Hending: Hilde, 16.-17.11.2013.]. Technical Report MET info no. 15/2013, The Norwegian Meteorological Institute, Bergen, 2013.

[3] S. de Haan, G.-J. Marseille, P. de Valk, and J. de Vries. Impact of ASCAT Scatterometer Wind Observations on the High-Resolution Limited-Area Model (HIRLAM) within an Operational Context. *Wea. Forecasting*, 28:489–503, 2013.

[4] J. Figa-Saldana, J.J.W. Wilson, E. Attema, R. Gelsthorpe, M.R. Drinkwater, and A. Stoffelen. The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers. *Can. J. Remote Sensing*, 28(3):404–412, 2002.

[5] P. Ollinaho. Feasibility of assimilating ASCAT surface winds into a limited area. Master's thesis, Dept. of Physics, University of Helsinki, 2010. Available online at https://helda.helsinki.fi/bitstream/handle/10138/20968/feasibil.pdf.

Collection and assimilation of Mode-S MRAR observations in Slovenia

Benedikt Strajnar

Introduction

This article describes the use of Mode-S Meteorological Routine Air Report (MRAR) data, one type of recently available high-frequency aircraft-derived data (ADD). Such observations of wind and temperature are routinely collected at Ljubljana airport and cover the air space around Slovenia. They have been available in experimental mode to the Slovenian Environment Agency since 2011 and in operational mode after extensive validation in 2014. The impact on mesoscale weather prediction by data assimilation into a 3-hourly ALADIN assimilation cycle is assessed.

Mode-S observations

The basic components of a Mode-S system are ground secondary surveillance radars (SSR) and aircraft transponders. The Mode-S ("Mode Selective") system enables selective communication between a specific aircraft and radar with the possibility to transmit various 56-bit data registers, up to 5 for a standard system. The system was designed to improve air-traffic control and safety but can be exploited to transmit meteorological observations.

Mode-S MRAR

Sensors on board the aircraft directly measure temperature and wind as a difference between air- and ground-speed, taking into account the orientation of the aircraft. MRAR (BDS 4,4) is a dedicated meteorological data register within Mode-S to transmit these observations. Such data from one radar site located near Ljubljana were validated by Strajnar (2012) and Hrastovec and Solina (2013). It was demonstrated by Strajnar (2012) that the overall quality of the Mode-S data for AMDAR-equipped aircraft is the same as the quality of AMDAR data, both for wind and temperature. It was also observed in these studies and also recently by de Haan (2014) that a relatively small subset of all aircraft transmit meteorological data through the MRAR data register. This depends on the transponder type and configuration. This percentage is around 5%, but for the case of Slovenia, it increases in the lower atmosphere (the most frequent aircraft type flying to Ljubljana is Canadair Regional Jet, which provides data). The definition of Mode-S data registers also includes capabilities to transmit turbulence, humidity, wind shear, microburst and icing (Hrastovec and Solina, 2013), but no aircraft currently provides these parameters. An example of horizontal and vertical data coverage over Slovenia for an arbitrary date is shown in Fig. 1.

Mode-S EHS

Another approach to extract meteorological information from aircraft is by using data registers available within the Mode-S Enhanced Surveillance (EHS) protocol. The transmission of these data is mandatory for all Mode-S equipped aircraft with the aim of improving air-traffic safety. The parameters include ground and true air speeds of the aircraft, magnetic heading and Mach number. de Haan (2011) provides a method to extract wind and (indirectly) temperature from EHS parameters. While this method offers the advantage of using all aircraft in the visual range of a radar, the quality of

EHS derived temperature was shown to be smaller than the quality of AMDAR temperature. Assimilation experiments at KNMI (de Haan and Stoffelen, 2012) which used data collected at Amsterdam Airport showed positive impact on wind and temperature for the first few hours of the forecast.

Preprocessing and quality control

The raw Mode-S MRAR data used in this study are provided by air-traffic control with a frequency of 4 s. They are first subject to temporal smoothing over a period of 12 s during ascent or descent and 60 s during constant-level flights.

The observation selection procedure is based on a comparison of Mode-S against operational NWP over a period of 22 months, separately for each aircraft. Based on mean and standard deviation difference, a white list of acceptable aircraft was prepared. This comparison showed that the quality of observations depends on aircraft type and that it can be problematic especially for some smaller private aircraft (which also respond with MRAR) and these need to be removed from the analysis.

After smoothing and data selection, the observations are coded into the OBSOUL format and enter the data assimilation together with AMDAR observations. The same error characteristic are also assumed.

Data assimilation experiments

Model setup

The data assimilation experiments were carried out with the latest operational setup at the Slovenian Environment Agency as of spring 2014. This includes ALADIN/ALARO cy38t1, 4.4 km horizontal resolution and 87 vertical sigma levels. The atmospheric analysis is provided by 3-hourly three-dimensional variational assimilation (3d-Var) , and uses Synop and automatic surface observations, radiosondes, Meteosat Second Generation (MSG) atmospheric motion vectors, AMDAR and (recently) Mode-S, and satellite radiances from MSG, the National Oceanic and Atmospheric Administration (NOAA) and MetOp satellites. The European Centre for Medium-Range Weather Forecasts (ECMWF) global model is used for lateral boundary conditions.

In the experimentation, two identical 3-hourly assimilation cycles are run for winter and summer periods and selected cases. The only difference is that Mode-S are assimilated in the experimental and not in the reference runs. Longer, 24-hourly forecast runs are made every 6 hours.

The verification is done against Mode-S MRAR data itself because it is the only available source of high-frequency vertical profiles over the region.

Winter period

The winter evaluation period extends from 18 December 2013 to 10 January 2014. The weather during the first half of this period was relatively stable with frequent temperature inversions. Later on it was more dynamic, with excessive rain amounts for this part of the year. Figure 3 shows how Mode-S data improve the local analysis at the beginning of the period. The temperature profile is modified to match better the local Mode-S observations. Additionally, part of the improvement with respect to the reference setup comes already from the first guess. On average, the root mean square (RMS) error reduction of the temperature forecast due to assimilated Mode-S (left-hand side of Fig. 3) is visible in the first two hours of the forecast range in all atmospheric layers, up to around 5 hours in the planetary boundary layer (PBL), and even longer near the surface. The latter can be explained by the persistent inversions in the Ljubljana basin. For wind (right-hand side of Fig. 3), a positive impact is visible up to 5 hours in the mid-troposphere.

Summer period

The weather during the summer evaluation period (18 June 2013 – 10 July 2013) was quite typical for this part of the year, with periods of unstable and convective conditions and dry, warmer periods. Unlike for the winter period, the RMS difference between experimental and reference forecasts (Fig. 4) does not everywhere indicate improvements. It is negative (error reduced) in the first 1-2 hours of the forecast but is mixed further in the forecast range. Especially for the temperature in the PBL, predominantly detrimental impact can be seen near the ranges of 1 day and even longer in the forecast range(not shown). This interesting result is under further investigation, but is most probably related to problematic humidity initialization. At analysis time, the humidity (not contained in Mode-S MRAR) relies mostly on the (climatological) multivariate balance in the forecast error covariances.

Freezing rain case

Around the end of January 2014, almost the whole inland of Slovenia was hit by unusually intense freezing rain storm, causing extreme damage on woods and infrastructure. Figure 5 shows the vertical cross section of temperature and hydrometeors over Slovenia (from coast to inland) for 1 February 19 UTC (1-hourly forecast). It is evident that the positive temperature layer (white or yellow colors) is better described in the experimental cycle and that it extends further into the continental part of Slovenia. The probability of freezing rain (see caption of Fig. 5) is therefore much better analysed for the region of Ljubljana. The positive local impact of Mode-S is lost quickly further in the forecast range, because the situation was quite dynamic.

Conclusions

In this article, the use and evaluation of Mode-S MRAR observations from one radar site near Ljubljana was described. The comparisons to AMDAR, radiosondes and NWP outputs proved that Mode-S MRAR are on average very good observations. However, observational selection and quality control is mandatory due to possible biases for some aircraft types. There is a clear local impact on the nowcasting time scale which is much easier to be demonstrated during winter conditions. During summer, the impact is more complex and indicates strong sensitivity to humidity analysis. The observations showed to be beneficial for analysis and nowcasting during the extreme freezing rain case.

The Mode-S MRAR data have great potential for operational nowcasting and short-range NWP once their collection would become more widespread. First samples of Mode-S MRAR are already available also from Czech Republic and Denmark. The possibility of future coordinated dissemination of Mode-S MRAR over Europe, together with other ADD, are currently also explored by an expert group within the network of European National Meteorological Services (EUMETNET).

References

de Haan, S., Availability and quality of Mode-S MRAR (BDS4.4) in the MUAC area : a first study KNMI Internal Report: IR-1, 2014.

de Haan, S., and A. Stoffelen, Assimilation of High-Resolution Mode-S Wind and Temperature Observations in a Regional NWP model, Weath. Forec., 27:918–937, 2012.

de Haan S., High-resolution wind and temperature observations from aircraft tracked by Mode-S air traffic control radar, J. Geophys. Res., 116:D10111–, 2011.

Hrastovec, M., and F. Solina, Obtaining meteorological data from aircraft with Mode-S radars, IEEE Journals & Magazines, 28, 12-24, 2013.

Strajnar, B., Validation of mode-s meteorological routine air report aircraft observations, J. Geophys. Res., 117(D23), 2012.
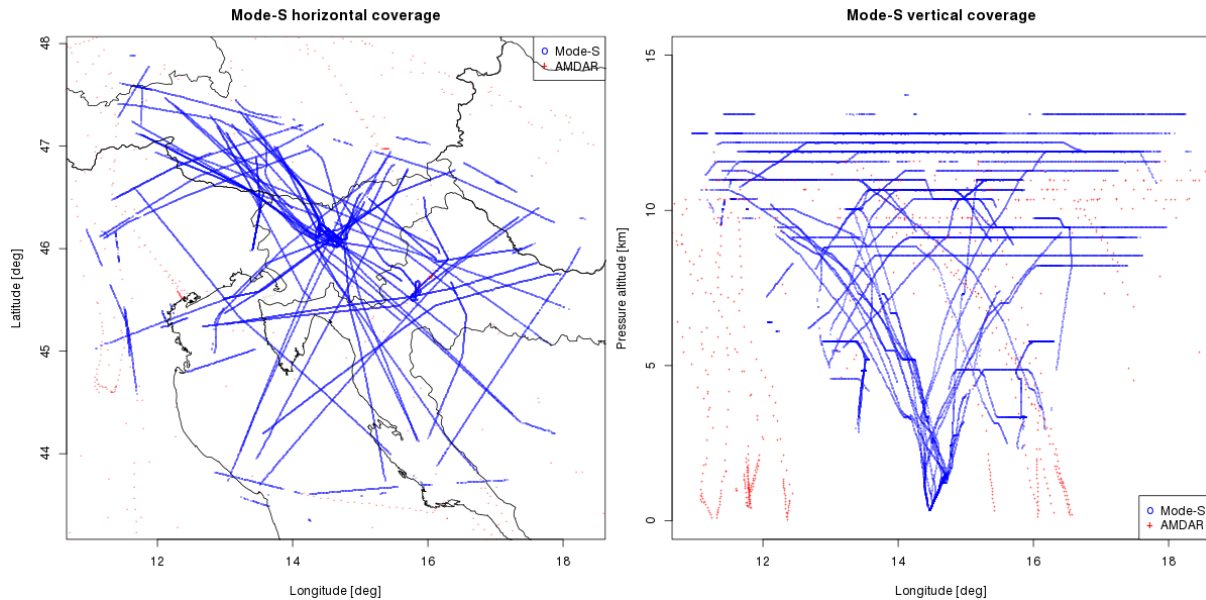
Figure 1:   An example of Mode-S MRAR horizontal and vertical coverage in 6 hours. The number of AMDAR observations is also shown.
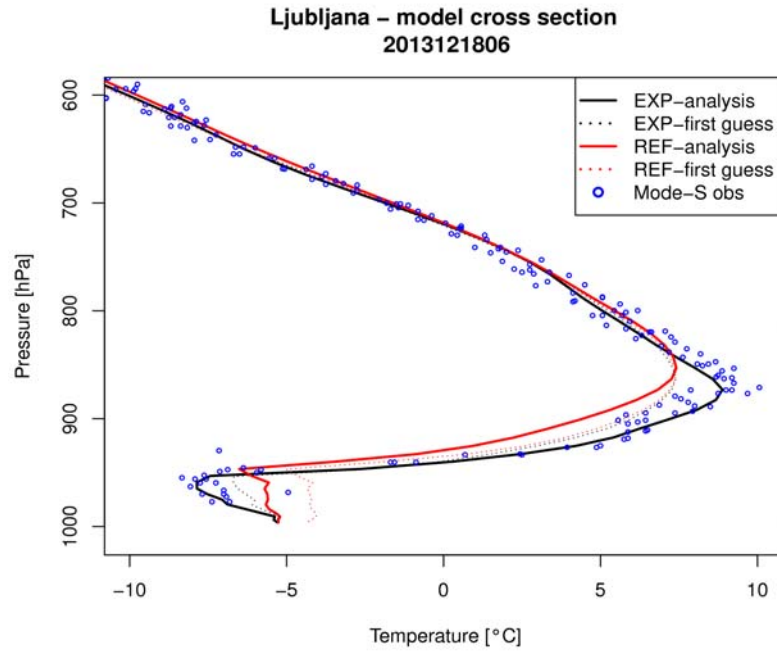
Figure 2:   *Impact of Mode-S MRAR on vertical temperature profile analysis. EXP is experiment with Mode-S observations, which are also shown as blue circles. Full lines are analysis profiles and dotted lines are first guess for the same analysis.*
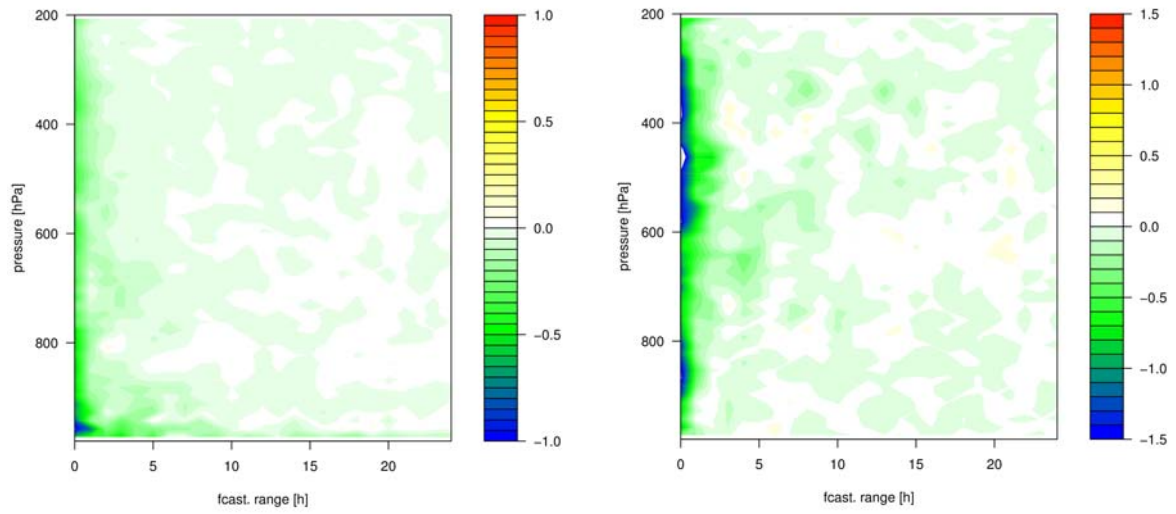
Figure 3:   RMS error reduction due to assimilated Mode-S MRAR during winter evaluation period. Negative values (greenish colors) indicate improvements.
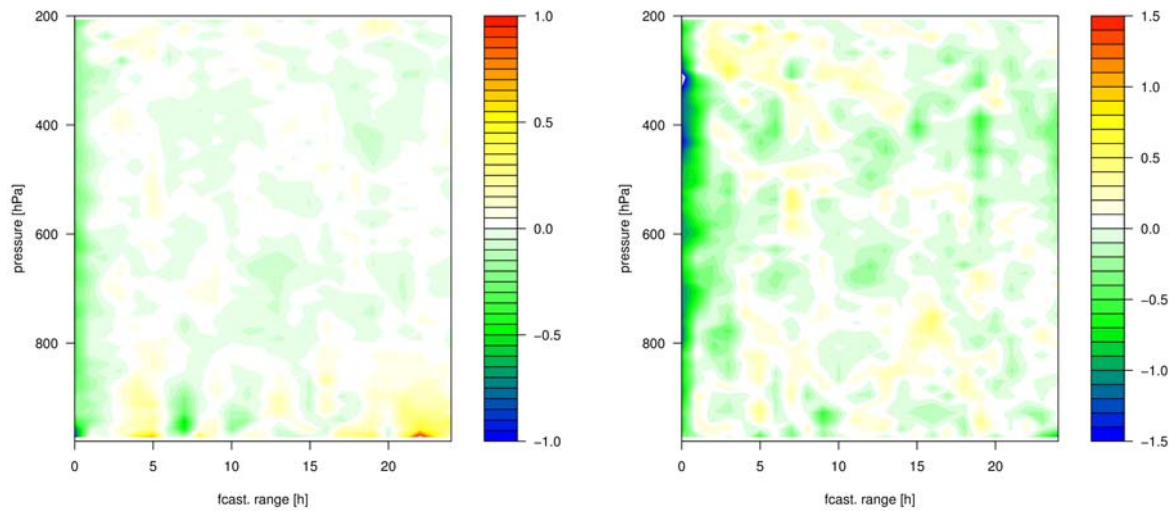
Figure 4:   Same as Fig. 3 but for the summer evaluation period.
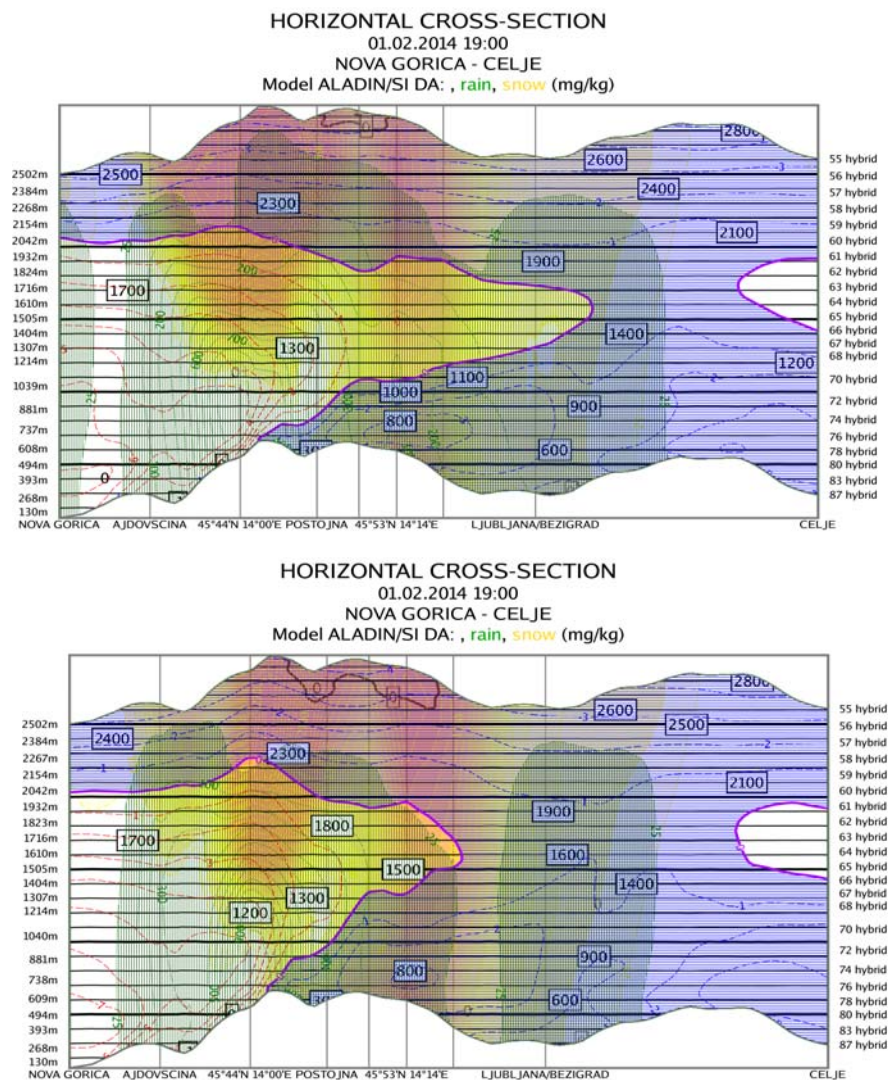
Figure 5: Vertical cross section of temperature and hydrometeors over Slovenia on 1 February 2014 19 UTC for experimental (top) and reference (bottom) cycles. Positive temperatures are plotted with red and negative with blue isolines and additionally, marked with blue horizontal lines. Snow is plotted in yellow fill and rain with green vertical lines. The squared areas where rain and negative temperature overlap indicate danger of freezing rain.

Aliasing and upper level explosions. Two fixes

Mariano Hortal
AEMET
P.L. for dynamics in HIRLAM-B

Introduction

In cycle 38r2 of IFS, Nils Wedi from ECMWF introduced a reduction of the aliasing in vorticity coming from the pressure-gradient term. This reduction of aliasing, as well as an aesthetic effect on the forecast fields, produced a remarkable improvement in the conservation of the total mass of dry air (Tomas Morales *private communication*) on the global model. The improvement didn't come from an a-posteriori fix resulting from a computation of the global norm, as is the case with the mass-conservation fix existing already in the model, but from a local improvement in the accuracy of the computations.

The beneficial effect happens at all resolutions tested and should therefore be beneficial for the conservation of dry air in HARMONIE, although the conservation in the limited area model is very difficult to assess.

The way of achieving this reduction of aliasing is based on the fact that the aliasing is tied to orography and therefore very stationary. The pressure-gradient term computed at the present time step is transformed to spectral space together with all the other dynamical fields and filtered in spectral space. The result of the filtering is added to the computation of the pressure-gradient term at the following time step.

There is therefore a part of the procedure which is applied in grid-point space. This part is usable straight away in the limited area version. The filtering is done in spectral space and therefore has to be adapted if we want to use the procedure in the HARMONIE model, as the two spectral spaces are different.

Adaptation of the aliasing elimination

In order to check that the aliasing elimination works properly in the limited-area version, the pressure gradient term was postprocessed at the hybrid model levels in a run using the area covering the Iberian peninsula and the Balearic Islands with a resolution of 2.5 km.

Figure 1 shows a plot of the x-component of the pressure gradient field at level 30 in a run using 65 levels. One can see a large noise not related with orography. The noise is not stationary as the one coming from the aliasing in vorticity and is not reduced when the elimination of aliasing in vorticity is applied. The aliasing in vorticity related with orography is not apparent in this plot, presumably because the (spectrally smoothed) orography in the area is not too steep. It nevertheless can be seen (not shown here) in the runs at 1 km resolution on the Canary Islands referred to in the next section. After some investigation, the noise was traced back as coming from the field "pressure departure". The same filter applied to the vortical part of the pressure gradient term in the IFS procedure was applied to the pressure departure field but with an amplitude reduction 5000 times weaker. The resulting pressure gradient field at the same time and same level is shown in Fig 2.

The needed amplitude of the filter is smaller than in the case of the pressure-gradient term because it is applied to one of the dynamic fields, namely the pressure departure, and therefore its effect accumulates from one time step to the next, while this is not the case with the pressure-gradient field which is "freshly" computed at every time step.

Upper nesting boundary conditions

Some runs at 1 km resolution run over the Canary Islands area had been failing due to too strong winds at the uppermost few levels.
Once the aliasing elimination on vorticity as well as the filtering of the pressure departure field were implemented, these cases were rerun with the filters applied.
The runs didn't fail any more with too strong wind but, after a few more time steps, failed in subroutine "condensation", an indication that the temperature field had crazy values at some points. These points belonged also to the uppermost few levels of the model.
It was clear therefore that the dynamics was not behaving properly in the upper part of the model. A change in the "sponge layer" increasing the linear horizontal diffusion there or introducing a Rayleigh friction term didn't help.
It was therefore decided to include an upper boundary condition similar to the lateral Davies relaxation boundary condition. This implementation has been in the plan of the dynamics for some time but up to now its need had not been felt. The fields forecasted by the limited area model were averaged with the ones coming from the nesting model, using (as a first fast and dirty solution) the same Davies relaxation coefficients used in the lateral procedure. This procedure will be called "upper nesting boundary conditions".
The forecast ran stably without any sign of strange noise in the upper part of the model.
There is a relaxation mechanism already implemented in the limited area model called spectral nudging, but this is applied only to the largest scales and it does not prevent the instabilities seen, which happen at the shortest scales.
Another way of stabilizing that run is to use the predictor-corrector time stepping procedure. One disadvantage of that time stepping is that, using the same size of the time step, the run of the model becomes some 50% more expensive in terms of computing time.
In order to check the quality of the predictor-corrector scheme against the standard SETTLS semi-implicit scheme, both with and without upper nesting boundary conditions, the run which failed over Canary Islands using 1 km resolution with 30 s time step was repeated with the three configurations and the spectra of the dynamic fields plotted at every model level.
Fig 3 shows the spectra of horizontal divergence at level 10 for the three runs. The run labelled "SETTLS" exploded after 17 hours of integration and the other two ran stably to the end of 24 hours. Nevertheless the run using the predictor-corrector scheme shows a large peak of divergent energy at around wavenumber 100, an effect also shown in the presentation of Petra Smolikova at the ASM in Bucharest. Fig 4 shows the evolution of the spectral norm of horizontal divergence vertically averaged. Clearly the evolution shown by the predictor-corrector scheme looks much less healthy than the other two. The vertical shoot-up of the green line corresponds to the explosion of the run using the standard SETTLS scheme without pressure-gradient dealiasing and without upper nesting boundary conditions.
The standard verification scores (not shown) didn't show any significant quality difference between the three runs.

Conclusions

The elimination of the aliasing in vorticity coming from the pressure-gradient term, implemented in cy38r2 of the global model, has been adapted to the limited area (HARMONIE) version of the forecast model, together with a filtering of the pressure departure field.
An upper nesting boundary condition has been implemented for the limited area forecast model, which eliminates most of the stability problems seen so far at very high resolutions.
The predictor-corrector time stepping scheme suffers from some noise in the upper part of the model and a strange behaviour of the evolution of some spectral norms. I wouldn't recommend its use for stabilizing the HARMONIE model runs.

Friday 15 November 2013 12UTC t   t+6 VT: Friday 15 November 2013 18UTC S
0.01 (0.04)
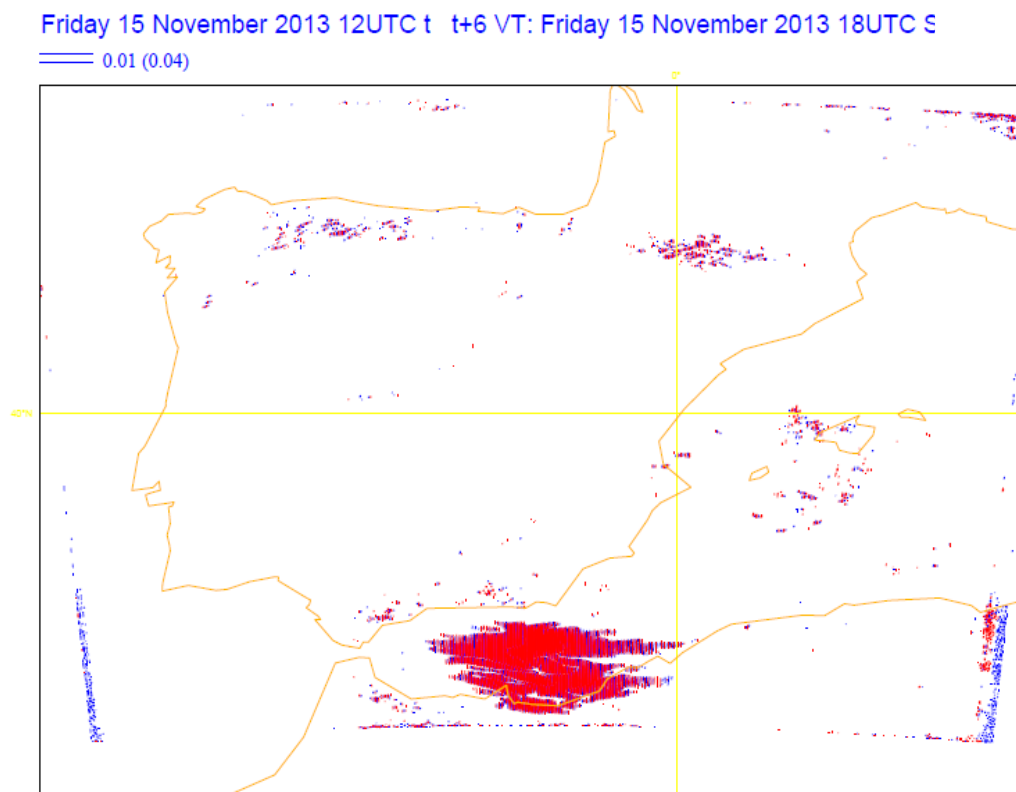
Fig. 1: Pressure gradient term at level 30 using the standard cy38h1.1

Friday 15 November 2013 12UTC t   t+6 VT: Friday 15 November 2013 18UTC Surface:
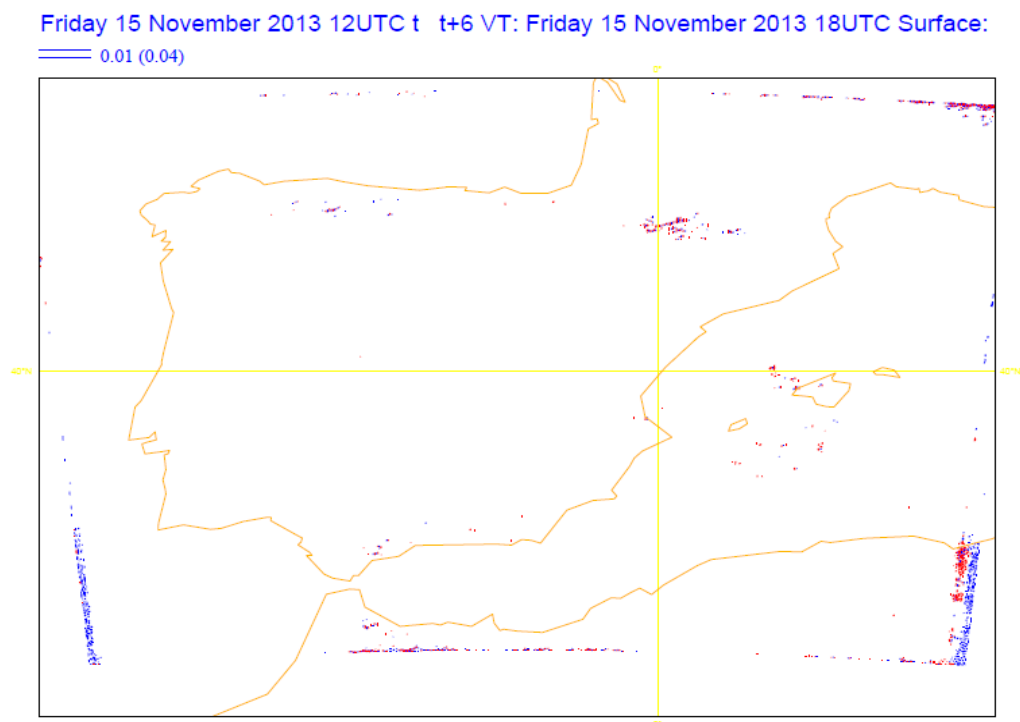0.01 (0.04)

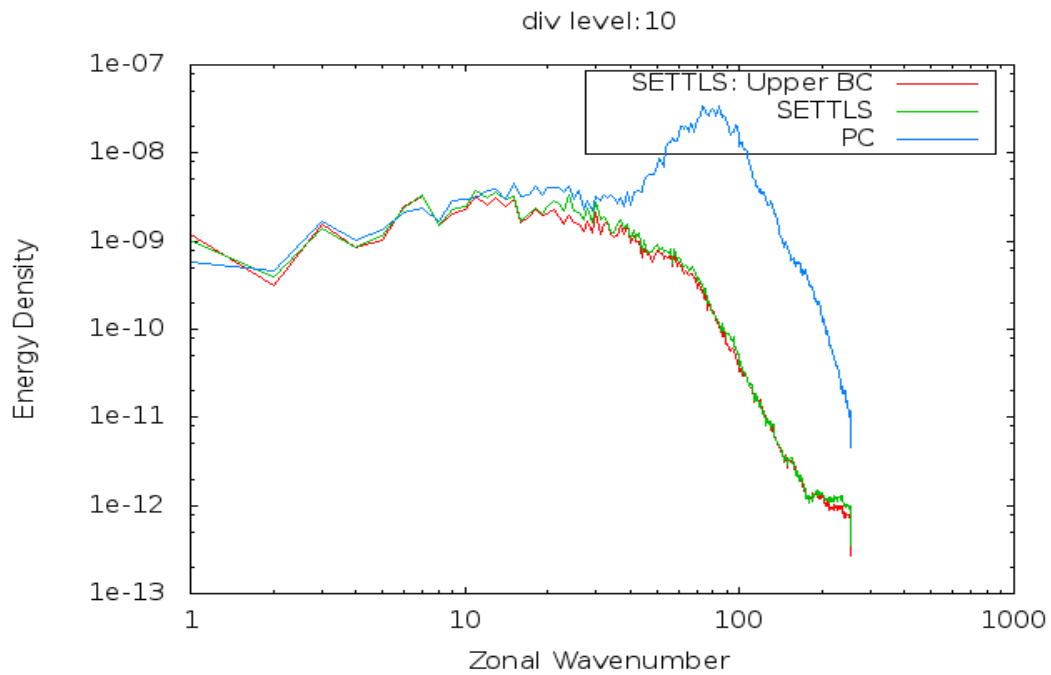*Fig 2: As Fig 1 but with a filter applied to the pressure departure field*

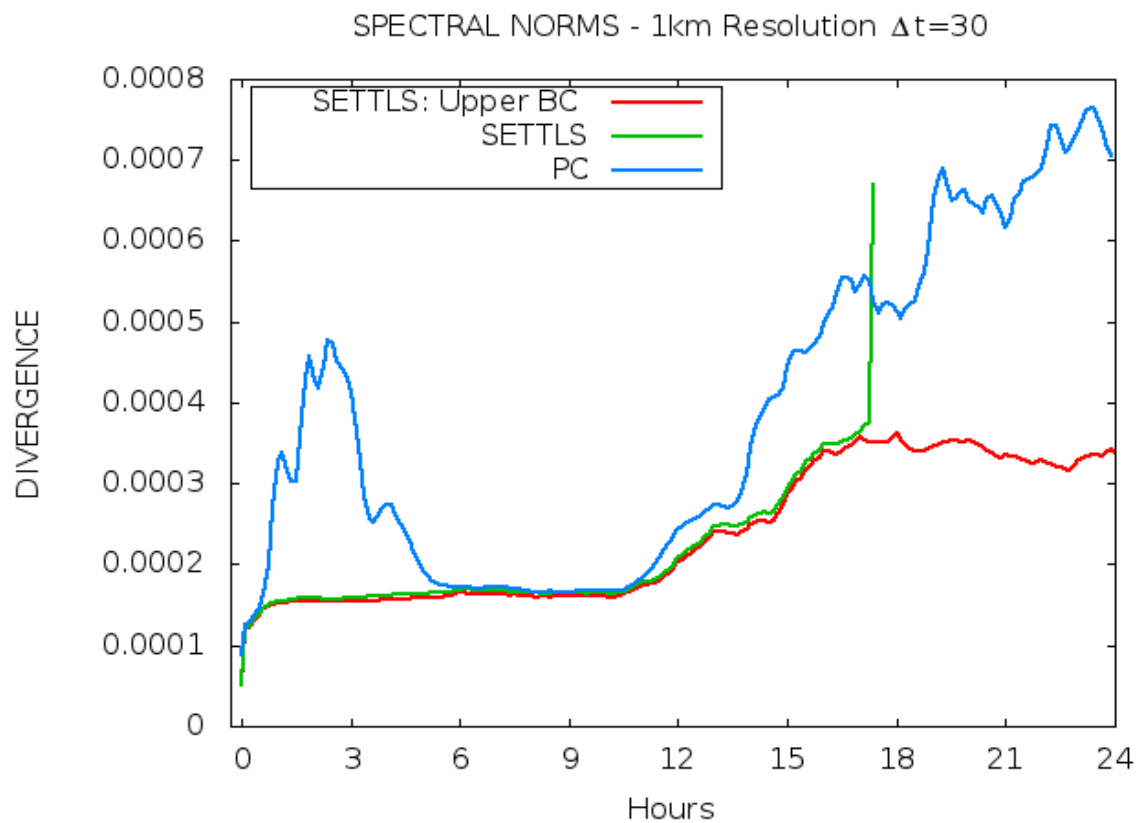*Fig 3.: Spectra of horizontal divergence at level 10 after 15 hours of integration*

Fig. 4: Evolution of the vertically averaged horizontal divergence spectral norms for the three runs.

Could there be a need for a reformulation of the spectral horizontal discretization?

Steven Caluwaerts, Piet Termonia, Daan Degrauwe, Fabrice Voitus and Pierre Bénard

1   Introduction

The spectral semi-implicit semi-Lagrangian (SISL) discretization forms the fundament of the dynamical cores of the ALADIN/ARPEGE/IFS family of models. This approach combines efficiency with accuracy: the method can be used with large timesteps thanks to the SISL time-discretization and the calculation of derivatives and the Helmholtz equation in spectral space guarantees a high order of accuracy. The success of the ALADIN/ARPEGE/IFS models during the last decades forms the best proof of its merits.

However, recently some questions (both technical and scientific) were posed about the durability of this spectral approach. The trend in high performance computing (HPC) manufacturing is clearly in favour of massively parallel machines on which global communication between all cores is very slow and thus needs to be avoided. The Fourier transforms and Legendre transforms (only for the global models) necessary for the shuttling between spectral and gridpoint space would thus become a bottleneck for the scalability of our models. Some also estimate that spectral methods would suffer problems above steep orography, a claim founded on the Gibbs phenomenon[1].

For a better understanding of the limitations of the spectral method, a study of and comparison with local horizontal discretization methods (e.g. finite differences (FD), finite elements (FE)...) would be of great interest for the ALADIN/ARPEGE/IFS community and this work can be considered as a first step.

In this document we discuss first the Z-grid method which could be used to study local discretization methods and fits within the current SISL timestep organization. Furthermore we investigate in a strongly simplified context the claim about orography. And finally we present our conclusions and future plans.

2   The Z-grid method as a testbed for local discretization methods

2.1   Local methods and IG wave dispersion

In order to obtain a useful dynamical core, many conditions (e.g. energy conservation, no grid imprinting,...) need to be satisfied as outlined by Staniforth and Thuburn (2012). In this section we will focus on one constraint, namely appropriate geostrophic adjustment because it is well known that when leaving the spectral discretization a careful treatment is needed in order to still satisfy this property. The atmosphere is always aiming at a balance between mass and velocity fields, the so-called geostrophic balance. An unbalanced atmosphere will try to restore balance by redistributing energy via the propagation of IG-waves. Therefore modellers are interested in designing numerical schemes whose IG-wave dispersion is realistic.

Mesinger and Arakawa (1976) have shown that schemes based on a finite difference horizontal discretization on the A-grid (= a grid where all prognostic variables are defined in the same gridpoints) are not able to represent appropriately the dispersion behaviour of short IG-waves. The group velocity $v_g = \frac{\partial \omega}{\partial k}$ which

---

[1]Approximating a discontinuous function by a Fourier series will result in significant over -and undershooting in the vicinity of the discontinuity.
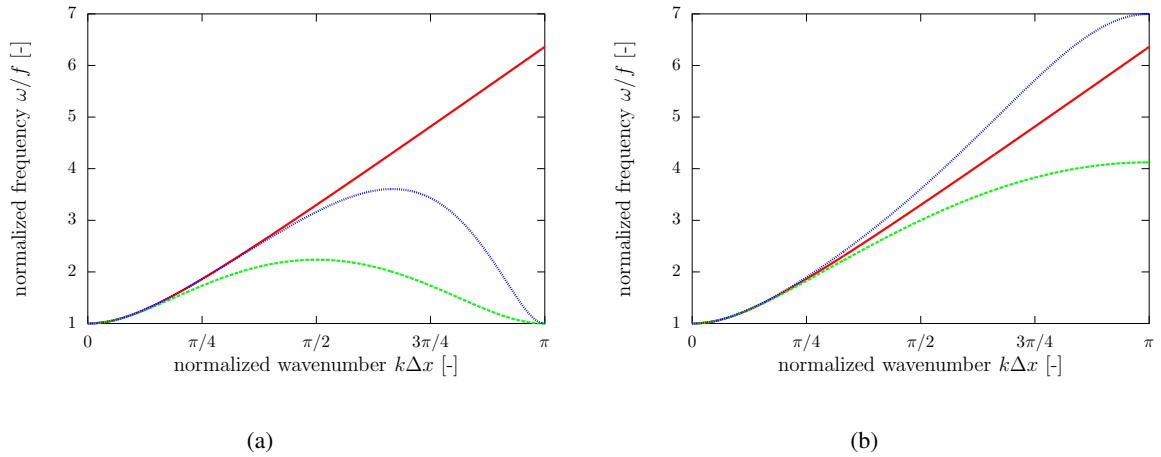
Figure 1: *Numerical dispersion of waves is typically characterized by a dispersion relation which relates the frequency ω of a wave to its wavenumber k (f and Δx are the Coriolis number and the resolution respectively). This figure shows the 1D numerical dispersion of the IG waves for the linearized shallow water equations assuming exact time derivatives. Figure (a) shows dispersion for A-grid discretizations: spectral (red), second order FD (green) and linear FE (blue) discretization. Figure (b) presents results for Z-grid discretizations: spectral (red), second order FD (green) and linear FE (blue) discretization.*

represents the velocity at which energy is propagating, has in this type of discretization schemes a negative sign for the shortest waves whereas the analytical group velocity always remains positive. This behaviour is illustrated in the dispersion diagrams of Figure 1(a); the red line represents the exact, spectral dispersion relation whereas the green one stands for the finite difference A-grid discretization. A similar behaviour (blue line in Figure 1(a)) is found when a linear finite element discretization is used on the A-grid. A spectral scheme does not suffer from this problem and reproduces the analytical relation because the spectral method has an infinite order of accuracy.

One could argue about the relevance of this issue: there is only a problem for the waves with wavelength $\lambda < 4\Delta x$ and it is well known that these waves are below the resolved resolution (due to diffusion, spectral truncation,...). It would be interesting to study the impact of the dispersion behaviour in this part of the spectrum because if it turns out that this does not deteriorate the model run one could consider using an A-grid discretization. However untill proven irrelevant, we assume that a numerical scheme must result into appropriate[2] dispersion behaviour.

2.2   Z-grid SISL approach

There exist different solutions to use a local spatial discretization method and still obtain appropriate IG-wave dispersion relations. The most common method uses a so-called staggered grid on which mass and velocity variables are no longer defined at the same gridpoints. However, this method does possess some drawbacks (e.g. more gridpoints means more trajectory calculations if a semi-Lagrangian scheme is used) and the use of a new type of grid would have a huge impact on the complete ALADIN/ARPEGE/IFS code. Therefore the so-called Z-grid method seems more promising.

The basis of the Z-grid method consists of solving the equations formulated in vorticity ζ and divergence D instead of wind components u and v. The relation between both is given by

---

[2]Appropriate meaning no negative group velocity.

22

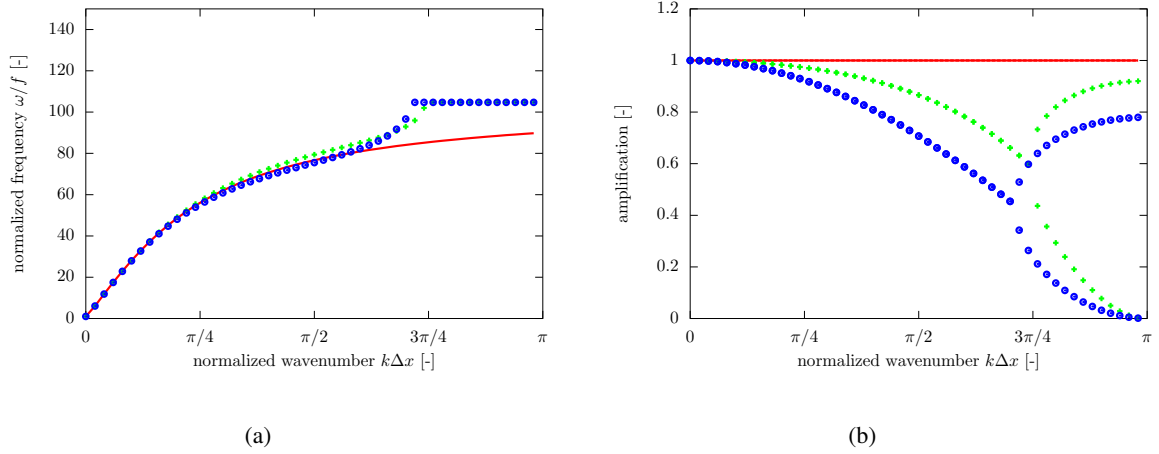(a)   (b)

Figure 2: IG wave normalized frequency plot (left) and amplification factor factor (right) of 2TL SISL Z-grid schemes for $\Delta t = 300$s. The other parameters are: $\Delta x = 10$km and $\Phi = 9000 \text{m}^2/\text{s}^2$. The used spatial discretization schemes are: spectral scheme (red cont. line), a linear FE scheme (green crosses) and a second order FD scheme (blue circles). It can be clearly seen that the FE and FD Z-grid schemes have inappropriate dispersion and strong damping. Remark that in the normalized frequency plots only the IG wave propagating in the positive x-direction is shown, whereas in the amplification plot both IG waves are shown.

$$D = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \tag{1}$$

$$\zeta = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}. \tag{2}$$

The dispersion relations of a Z-grid second order finite difference discretization (green) and linear finite element discretization (blue) are shown in Figure 1(b).

Despite the excellent geostrophic adjustment in case of perfect time derivatives, Caluwaerts et al (2014) found a surprisingly poor dispersion relation for the Z-grid method once time discretization was included (Figure 2). Rewriting the continuous equations in terms of vorticity/divergence results in some extra non-linear terms which is not suitable. To avoid this, a Z-grid method typically starts from a time-discretization of the equations formulated in (u,v) and only then is rewritten into a vorticity/divergence formulation. The shallow water equations linearized around a reference state ($u_0 = 0$, $v_0 = 0$ and $\phi_0 = \Phi$) can be written as

$$\frac{du}{dt} = -\frac{\partial \phi}{\partial x} + fv \tag{3}$$

$$\frac{dv}{dt} = -\frac{\partial \phi}{\partial y} - fu \tag{4}$$

$$\frac{d\phi}{dt} = -\Phi \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right), \tag{5}$$

with $\frac{d}{dt}$ the total time derivative, ϕ the geopotential perturbation with respect to the reference geopotential Φ. For simplicity only an x-dependency of the fields is assumed in what follows, so $\frac{\partial}{\partial y} = 0$. Semi-implicit semi-Lagrangian 2TL time discretization results in:

$$u^+ + \frac{\Delta t}{2} \frac{\partial \phi^+}{\partial x} - \frac{f \Delta t}{2} v^+ = R_u \tag{6}$$

$$v^+ + \frac{f\Delta t}{2}u^+ = R_v \tag{7}$$

$$\phi^+ + \frac{\Phi\Delta t}{2}D^+ = R_\phi \tag{8}$$

where ^{0} and $^+$ represent an evaluation at the previous time t^0 and the next time $t^+ = t^0 + \Delta t$ (with timestep Δt). The right hand sides are evaluated at time t^0 at the departure position x_d by using a suitable interpolation formula:

$$R_u = \left(u^0 - \frac{\Delta t}{2}\frac{\partial \phi^0}{\partial x} + \frac{f\Delta t}{2}v^0\right)_{x_d} \tag{9}$$

$$R_v = \left(v^0 - \frac{f\Delta t}{2}u^0\right)_{x_d} \tag{10}$$

$$R_\phi = \left(\phi^0 - \frac{\Phi\Delta t}{2}D^0\right)_{x_d}. \tag{11}$$

Based on (6) and (7) one can now construct the discrete divergence and vorticity equations, and combined with (8) these can be rewritten into a Helmholtz equation. This is exactly the same procedure as the one used in the current dynamical core.

It is sufficient to have a look at the discrete divergence equation to understand the observed dispersion problem of the Z-grid method (Figure 2). By taking $\frac{\partial(9)}{\partial x} + \frac{\partial(10)}{\partial y}$ one finds the discrete divergence equation

$$D^+ + \frac{\Delta t}{2}\frac{\partial^2 \phi^+}{\partial x^2} - \frac{f\Delta t}{2}\zeta^+ = \frac{\partial R_u}{\partial x}. \tag{12}$$

The previous equation is in general not symmetric between the implicit evaluations at time t^+ and the explicit evaluations at t^0. As an example, the second order derivative of the geopotential is defined as $\frac{\partial^2}{\partial x^2}$ in the implicit part, and as $\frac{\partial}{\partial x} \cdot \frac{\partial}{\partial x}$ for the explicit calculations. Taking once a second order derivative or twice a first order derivative is analytically identical but not necessarily numerically.

Figure 2 shows the dispersion relation for a Z-grid scheme discretized with second order finite differences, linear finite elements and a spectral discretization. The earlier discussed asymmetries do not only result in inappropriate dispersion, but the scheme is also no longer neutral. The spectral Z-grid scheme does not have this problem because the spectral discretization has an infinite order of accuracy. In Caluwaerts et al (2014) it is shown that the asymmetry problem can be solved by introducing some slightly changed numerical operators which restore symmetry as illustrated in Figure 3.

2.3   Analogy timestep organization ALADIN/ARPEGE/IFS and Z-grid approach

Finally it is interesting to remark that the described Z-grid approach is nearly identical to the currently used spectral SISL algorithm. In the latter, the time discretized (u,v) equations are also transformed into a discrete vorticity and divergence equation, before constructing the Helmholtz equation. The algorithmic similarity between the current timestep organization and the proposed Z-grid method makes it feasible to test the Z-grid method within the ALADIN/ARPEGE/IFS framework.

3   Spectral versus FE/FD discretization over orography

NWP models are known to have limitations when simulating flow over orography, a problem which becomes more important if one goes to higher resolutions and thus steeper slopes. This problem is mainly related to the vertical discretization and model level distribution. However, the choice of the horizontal discretization also plays a role in this, for 2 reasons:

(a)　　　　　　　　　　　　　　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　　　　　　　　　　　　　　(d)

(e)　　　　　　　　　　　　　　　　　　　　　　　　　　(f)
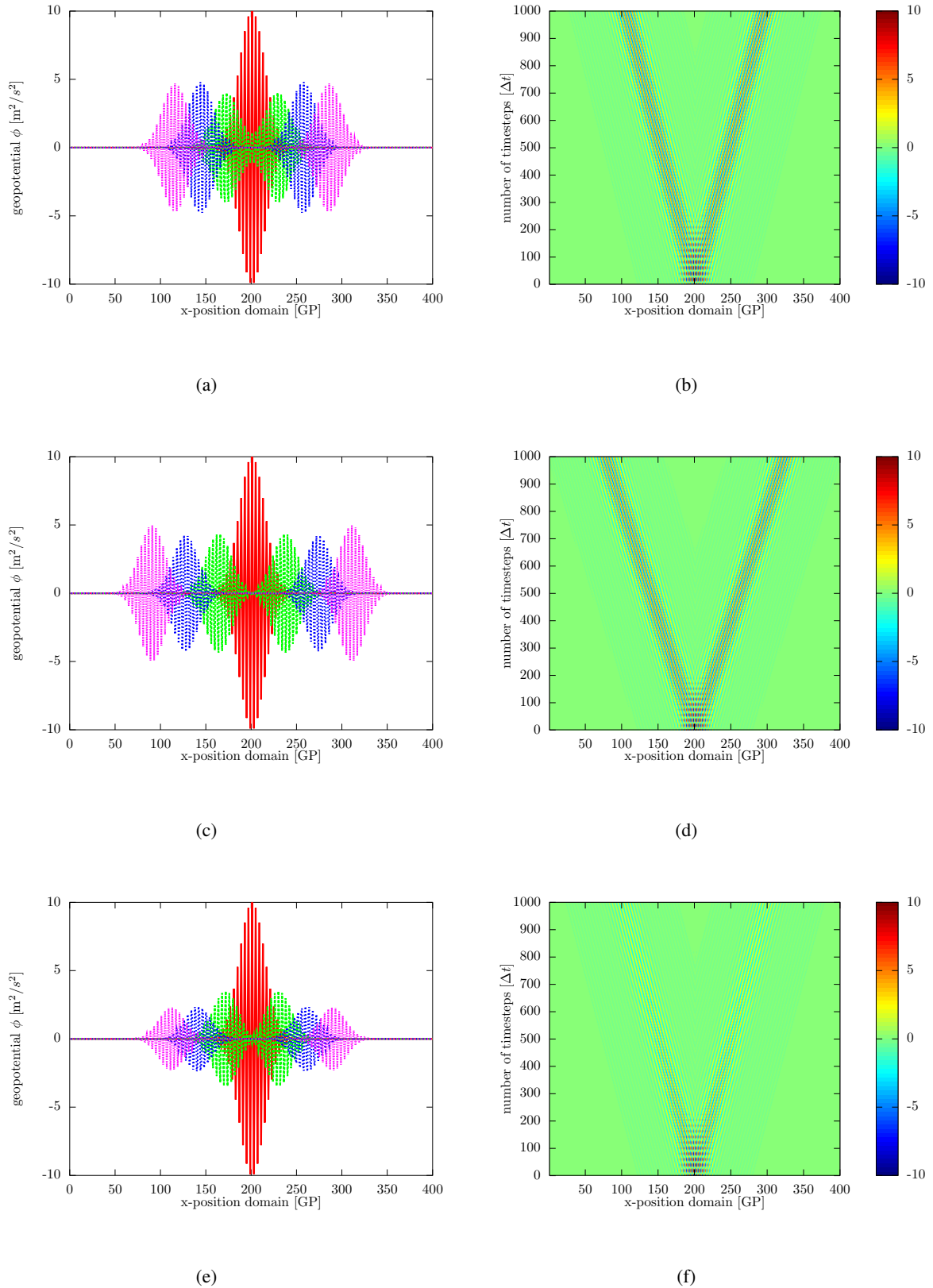
Figure 3: 1D IG wave propagation test similar to the one in Staniforth (2013) . From top to bottom the results are shown for a spectral scheme, a linear FE symmetric Z-grid scheme and a linear FE asymmetric Z-grid scheme. The figures in the left column show the geopotential field ϕ at the initial time (red), after 300 (green), 600 (blue) and 900 (magenta) time integrations. The figures in the right column plot the evolution of the geopotential field in a space-time diagram. The parameter values of the test are: $\Delta x = 10 \text{km}$, $\Delta t = 10 \text{s}$ *and* $\Phi = 9000 \text{m}^2/\text{s}^2$.

25

1. The presence of orography will result in fields undergoing large changes over short distances (e.g. geopotential of the surface, temperature,...). Approximating these fields by a spectral expansion will result in significant over and undershootings (Gibbs phenomenon) and this will have a negative impact especially on the spatial derivatives. This problem does only exist if a spectral spatial discretization is used.

2. Non-linear terms are not treated in a semi-implicit way, they need an explicit evaluation and it is known that these non-linear terms can influence the stability. In finite difference and finite element schemes it is however possible to include the non-linear products with orography (e.g. $\phi_{oro} \cdot D$ in the continuity equation with ϕ_{oro} the surface geopotential and D divergence) in the semi-implicit operator whereas this is impossible if a spectral discretization is used.

To numerically study the previous, a 1D orography test was done with the shallow water equations. Given a certain orography one can calculate analytically stationary states over this topography. Using this stationary state as an initial state for an experiment, the quality of a discretization can then be measured as the deviation of this run from the initial state.

Experimentally we found that over very steep orography both spectral and FD/FE methods do a poor job, however for different reasons. As explained earlier spectral methods suffer from the Gibbs phenomenon, whereas FD and FE methods are simply not able to give a realistic approximations for the derivatives in the shortest scale part of the spectrum. So the Gibbs phenomenon can not be considered as a handicap of the spectral method over FE/FD methods.

However over less steep orography the possibility to include the orography terms in the SI part of the calculations instead of considering them as non-linear terms was found to be advantageous. It was shown that in FE/FD schemes with $\phi_{oro} \cdot D$ included in the semi-implicit calculations, the predictor corrector iteration was redundant.

4   Conclusions and future work

It was shown that the Z-grid method is well suited as a framework to test local discretization methods within the current SISL ALADIN/ARPEGE/IFS framework. However a detailed study revealed that one must be careful not to break the symmetry between the explicit and implicit calculations. The claim that spectral methods are not able to represent well flow over steep orography is correct, however FD and FE methods did not do better. An advantage of local methods over spectral methods is that more terms can be added into the SI part of the calculations.

In the following months, we would like to study with the ALADIN-model the importance of the numerical dispersion properties of the shortest scale waves (waves with a wavelength smaller than $4\Delta x$). There is some debate within the NWP community about whether the dispersion of these waves is really relevant within a complete NWP model (physics, diffusion,...). It is interesting to study this because some very simple and efficient numerical methods (e.g. FE/FD on an A-grid) are not considered for the time being precisely because of their poor dispersion of the shortest scale waves.

References

Mesinger F. and Arakawa A. Numerical methods used in atmospheric models. GARP Pub. Series No. 17 1. World Meteorological Organisation, Geneva, 1976.

Staniforth A. and Thuburn J., Horizontal grids for global weather and climate prediction models: a review, Q. J. R. Meteorol. Soc., 138, 1-26, 2012.

Caluwaerts S., Degrauwe D., Termonia P., Voitus F., Bénard P. and Geleyn JF, Importance of temporal symmetry in spatial discretization for geostrophic adjustment in semi-implicit Z-grid schemes, Q. J. R. Meteorol. Soc., in press

Staniforth A., Melvin T., Cotter C. , Analysis of a mixed finite-element pair proposed for an atmospheric dynamical core, Q. J. R. Meteorol. Soc., 139, 1239-1254, 2013.

Single-column experiments with the EFB turbulence closure

Carl Fortelius and Evgeny Kadantsev

1   Introduction

A turbulence closure model accounting for the potential energy associated with thermal fluctuations in a stably stratified fluid has been presented by Zilitinkevich et al. (2013, Z13, hereafter). Among the advantages brought about by this "Energy Flux Budget (EFB) model" the authors mention the consistent energetics based on prognostic equations for both kinetic and potential energy of turbulence, generally non-gradient turbulent transport based on budget equations for turbulent fluxes, advanced treatment of the inter-component exchange of kinetic energy, and an advanced concept of turbulent dissipation time scale. The EFB model was introduced into the AROME physics package of HARMONIE and tested in single-column mode to assess its potential for improving the simulation of the stable boundary layer.

A brief outline of the EFB model is given in section 2, and its implementation into HARMONIE in section 3. The test cases are presented in section 4, and the results in section 5. The final section 6 contains a summary and a few ideas for further work.

2   EFB theory

According to Z13, the potential energy associated with turbulent fluctuations of density plays an essential role in the evolution of turbulence in a stably stratified flow. This is reflected in the evolution-equations for second moments in a dry turbulent flow, as derived in Z13 applying approximations of the characteristics of geophysical flows:

$$\frac{DE_k}{Dt} - \frac{\partial}{\partial z} K_E \frac{\partial E_k}{\partial z} = \overline{u_i'w'}\frac{\partial U_i}{\partial z} + \beta\overline{\theta'w'} - \frac{E_k}{t_T} \tag{1}$$

$$\frac{DE_p}{Dt} - \frac{\partial}{\partial z} K_E \frac{\partial E_p}{\partial z} = -\beta\overline{\theta'w'} - \frac{E_p}{C_P t_T} \tag{2}$$

$$\frac{D\overline{u_i'w'}}{Dt} - \frac{\partial}{\partial z} K_{FM} \frac{\partial \overline{u_i'w'}}{\partial z} = -2A_z E_k \frac{\partial U_i}{\partial z} - \frac{\overline{u_i'w'}}{C_\tau t_T} \quad i = 1,2 \tag{3}$$

$$\frac{D\overline{\theta'w'}}{Dt} - \frac{\partial}{\partial z} K_{FH} \frac{\partial \overline{\theta'w'}}{\partial z} = -2(A_z E_k - C_\theta E_p)\frac{\partial\Theta}{\partial z} - \frac{\overline{\theta'w'}}{C_F t_T} \tag{4}$$

(U_1, U_2) and Θ refer to the Reynolds-averaged wind and potential temperature, (u_1', u_2', w') and θ' are the corresponding fluctuating quantities, $E_k = \overline{u_i'^2}/2$ is the turbulent kinetic energy,

$$E_p \equiv \frac{\beta}{2}(\frac{\partial\Theta}{\partial z})^{-1}\overline{\theta'^2} \tag{5}$$

is the turbulent potential energy, β is the buoyancy-parameter, the exchange-coefficients $K_x = C_x A_Z E_k t_T$, $A_z \equiv \overline{w'^2}/(2E_k)$ measures the anisotropy of turbulence, t_T is a turbulent time scale, and C_x are empirical constants.

Taking the sum of 1 and 2 shows that total turbulent energy is maintained by shear-production (the first term on the LHS of 1), and destroyed by dissipation (the last terms in 1 and 2), while the buoyancy-term only acts

to convert one form of energy to the other without affecting the total amount. Hence turbulence cannot be destroyed by buoyancy, so there should be no critical Richardson-number, above which turbulence could not be maintained by the shear production. Even in a situation where all the kinetic energy vanishes momentarily, a positive heat flux will be generated via the second term on the LHS of 4, and this, in turn, will serve to convert some of the turbulent potential energy back to kinetic energy.

Z13 derives a diagnostic expression for the factor of anisotropy A_z as a function of the stratification. This function decreases monotonically from a value close to 0.2 in the neutral limit, to an asymptotic value of the order of 0.03 for very high stability. For the turbulent time scale Z13 proposes a relaxation equation of the form:

$$\frac{Dt_T}{Dt} - \frac{\partial}{\partial z} K_T \frac{\partial t_T}{\partial z} = -C_R(\frac{t_T}{t_{TE}} - 1), \tag{6}$$

where t_{TE} is an equilibrium time-scale characterizing stationary turbulence, and associated with a length scale $l = t_{TE}E_k^{1/2}$. An expression for t_{TE} is derived in Z13, based on considering the stability-dependence of wind shear:

$$t_{TE} = \frac{\kappa z}{E_k^{1/2} + C_\Omega \Omega z}(\frac{E_k}{\tau})^{3/2}(1 - \frac{\Pi}{\Pi_\infty}), \tag{7}$$

where κ is the von Kármán konstant, Ω is the frequency of the earth's rotation, τ is the magnitude of the momentum-flux, $\Pi \equiv E_p/E_k$ measures the static stability, approaching the asymptotic limit of Π_∞ in very strong stratification. Finally,

$$(\frac{E_k}{\tau})^2 = \frac{C_p}{2C_\tau(C_p + \Pi)A_z}, \tag{8}$$

where C_p is a constant characterizing the dissipation rate of E_p. Alternatively, 7 and 8 can be written in terms of the flux Richardson number (Ri_f), using the relationship:

$$\Pi = \frac{C_p Ri_f}{1 - Ri_f} \tag{9}$$

Z13 suggests that the EFB model 1- 4 can be simplified without loss of the advantages outlined in the introduction, by replacing 3 and 4 by diagnostic flux-gradient relationships obtained for the steady state:

$$\overline{u_i'w'} = -K_M \frac{\partial U_i}{\partial z}, \quad K_M = 2C_\tau A_z E_k t_T, \quad i = 1, 2 \tag{10}$$

$$\overline{\theta'w'} = -K_H \frac{\partial \Theta}{\partial z}, \quad K_H = 2C_F A_z E_k t_T(1 - C_\theta \frac{E_p}{A_z E_k}) \tag{11}$$

3   Implementation in HARMONIE MUSC

The EFB model has been implemented into the AROME physics package of HARMONIE cycle 38.h1 in its flux-gradient form, based on equations: 1, 2, 10, 11, 6, 7. This involved carrying out the following tasks:

1. Introduction of new prognostic variables

   For simulating a dry atmosphere, it is sufficient to introduce one new prognostic equation describing the potential energy. For realistic flows, moisture must be included, and the EFB-model must be extended

accordingly. Code updates introducing the required prognostic equations into the turbulence-scheme of Meso-NH were kindly provided by Valèry Masson, and updates introducing the corresponding variables into the GFL structure of the IFS by Yves Bouteloup.

In a dry atmosphere, the equation for $\overline{\theta'^2}$ reads:

$$\frac{D\overline{\theta'^2}}{Dt} + \frac{\partial}{\partial z}\overline{w'\theta'^2} = -2\frac{\partial\Theta}{\partial z}\overline{w'\theta'} - 2C_{\epsilon\theta}\frac{E_k^{1/2}}{L_\epsilon}\overline{\theta'^2}, \tag{12}$$

where L_ϵ is a dissipative length-scale, and $\overline{w'\theta'}$ is given by the flux-gradient relationship:

$$\overline{w'\theta'} = -\frac{L}{C_{p\theta}E_k^{1/2}}[A_z E_k \frac{\partial\Theta}{\partial z} - \frac{2}{3}\beta\overline{\theta'^2}]. \tag{13}$$

Taking into account the definition of E_p (5), 12 and 13 can be made equivalent to 2 and 4 simply by setting $L_\epsilon = L = E_k^{\frac{1}{2}} t_T$, and selecting appropriate values of the relevant constants. In a moist atmosphere, additional terms appear in 12 and 13, and prognostic equations for variance of specific humidity and the covariance of humidity and potential temperature need to be included.

2. Treatment of mixing length and dissipative length scale

HARMONIE AROME makes use of the mixing length designed by Bougeault and Lacarrère (1989), which is also used as a length scale for the dissipation. An attempt to include only a prognostic E_p, without changing the formulations of the length-scale proved to be unsuccessful: the change compared to the original model was very small. For neutral to stable stratification the mixing length of Bougeault and Lacarrère (1989) was simply replaced by $t_T E_k^{1/2}$, but the EFB model does not provide a length scale for unstable stratification. Simply retaining the original mixing length in these situations would be problematic, because of the discontinuity introduced at the transition from stable to unstable. The problem was resolved heuristically by setting $t_{TE}(unstable) = t_{TE}(neutral)$. (The length scale may still vary with stability because of its dependence on E_k).

Having to carry a full prognostic variable for the time scale t_T was avoided by ignoring the advective and diffusive terms in 7, and writing the relaxation-term as follows:

$$t_T^{i\Delta t} = t_T^{(i-1)\Delta t} + min[0.2, \frac{\Delta t}{t_{TE}}](t_{TE} - t_T^{(i-1)\Delta t}), \tag{14}$$

where ΔT is the time step and the limiting factor 0.2 is included for numerical stability. When computing t_{TE}, Ri_f was used to measure stability, rather the Π-variable. This was to avoid numerical problems evaluating 5 for very small values of $\partial\Theta/\partial z$.

3. Setting empirical constants

The empirical constants appearing in the EFB-model are defined by specifying four parameters determined from data or LES/DNS models, as described by Z13. These parameters and the values applied in this study are given in Table 1.

Two further simplifications were introduced at this point. Firstly, the earth's rotation was ignored, by setting the constant $C_\Omega = 0$ in 7. Secondly, the anisotropy-factor A_z was kept constant at its neutral value of 0.2. The latter modification reduced the number of coding modifications needed to implement the EFB-scheme, and seemed justified for the range of Richardson-numbers encountered in the experiments performed

Table 1: Empirical parameters defining the constants of the EFB-model

Parameter	Value	Source
$\lim\limits_{Ri\to\infty} Ri_f$	0.2	fig. 4 in Z13
$\lim\limits_{Ri\to\infty} \dfrac{E_p}{E_k + E_p}$	0.13	fig. 7 in Z13
$A_z(Ri = 0)$	0.2	fig. 3 in Z13
$\lim\limits_{Ri\to\infty} A_z$	0.03	fig. 3 in Z13

4   Experiments and validation material

The EFB turbulence closure was tested in single column runs with the HARMONIE MUSC cycle 38h1. Experiments were carried out for three cases:

1. The stably stratified idealized GABLS1 case

   In GABLS1, (GABLS: GEWEX Atmospheric Boundary-Layer Study; Holtslag, 2003, 2006) an initially 100 m deep mixed layer topped by a stably stratified atmosphere, having a uniform zonal wind in geostrophic balance is evolving in response to a constant geostrophic forcing, surface drag, and a surface temperature cooling at a constant rate. No moisture is present, and no parametrized processes other than turbulent mixing, surface heat flux, and surface friction are active.

   For this case, the MUSC-domain consisted of 64 isobaric levels with the lowest full level at about 3 m, the highest at 400 m above ground, and a spacing of about 6 m.

2. A "conventionally neutral" idealized case

   This case is similar to GABLS1, but the surface heat flux has been deactivated. Hence, the initially well mixed layer evolves solely in response to geostrophic forcing, surface drag, and the heat being transported downwards from the free atmosphere by turbulence.

   For this case, the MUSC-domain consisted of 90 isobaric levels with the lowest full level at about 3 m, the highest at 900 m above ground, and a spacing of about 10 m.

3. The realistic GABLS3 case

   GABLS3 consists of a clear-sky mid-latitude summertime diurnal cycle at Cabauw in the Netherlands, starting from a convective boundary layer at 12 UTC, close to mid day LST, on the 1st of July 2006. Initial conditions and time-dependent geostrophic/advective forcing are specified based on local observations, radio soundings at De Bilt, and 3-D modelling. In the simulations for GABLS3, the full (AROME) physics are active in MUSC.

   For this case, the MUSC-domain consisted of the 65 hybrid levels used operationally at FMI, with the lowest full level at about 12 m, the highest at about 30 km above ground, and 20 levels within the lowest 1000 m.

For the two idealized cases MUSC is compared to LES-results provided by Andrey Glazunov. In the LES-runs, a grid spacing of about 3 m was used for GABLS1, and one of about 6 m for the conventionally neutral experiment. For GABLS3, results are compared to data.

5   Results

5.1   GABLS1

Results from the GABLS1 experiments are shown in Fig. 1. After nine hours of integration, the temperature near the surface has dropped by about 2 degrees, and the initial profiles have evolved into a deeper moderately stable layer bounded by a shallow inversion (seen as a steep increase of Θ with height) near the surface, and a deeper inversion around 200 m. The constant wind speed has evolved into a nearly logarithmic profile, with a supergeostrophic maximum associated with the upper inversion-layer, and the hodographs show the characteristic spiral-like veering inside the boundary-layer. Heat and momentum flow downward within the boundary layer, and the magnitude of the fluxes decrease linearly with height.

Like the majority of the NWP and research models examined for the GABLS1-case by Holtslag et al (2013), MUSC cy38h1 is seen to overestimate the downward transport of heat and momentum, compared with LES, and consequently develops a too deep boundary layer. The wind speed near the surface slightly too high in MUSC, and unlike in the models examined by Hotlsag et al. (2013), the cross-isobaric angle too wide. Introducing the EFB-model is seen to bring improvements to all these features. In particular, the Θ-profile and the heat flux agree almost perfectly with LES.
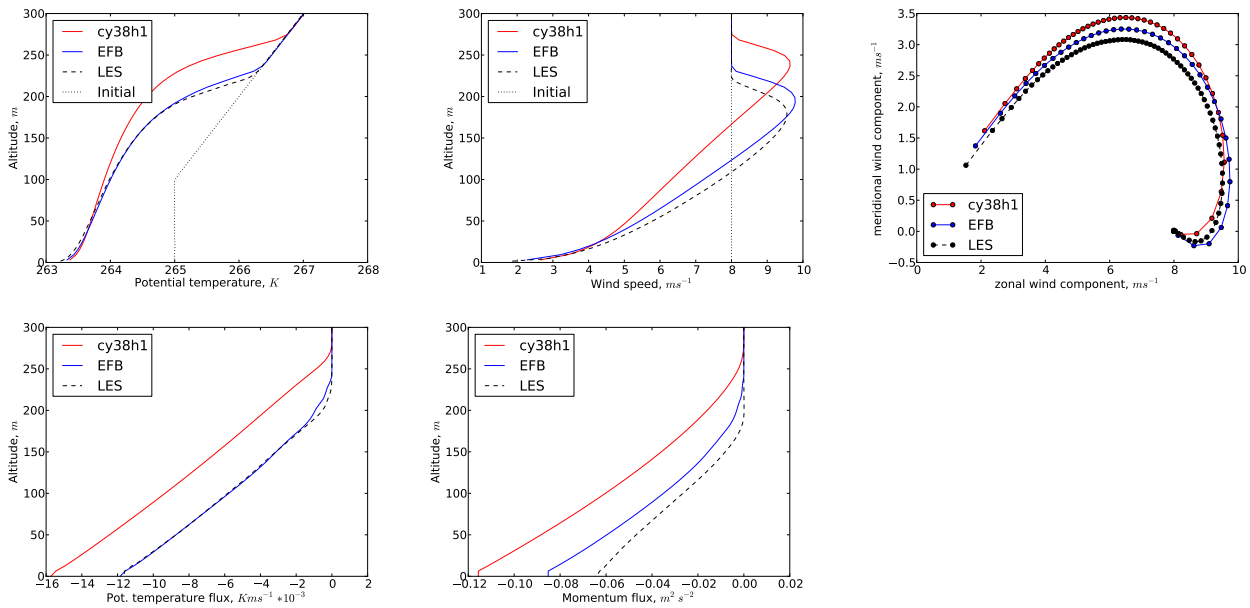

Figure 1: Profiles of potential temperature, wind speed, heat flux, momentum flux, and a hodograph, plotted nine hours after the initial time in GABLS1. MUSC cy38h1 is shown in red, MUSC with the EFB-closure in blue, and LES in black. The initial profiles of potential temperature and wind speed are drawn as dotted lines.

5.2   The conventionally neutral boundary layer

Results for the conventionally neutral experiments are shown in Fig. 2, 20 hours of integration after the initial time. In these experiments the mixed layer grows in depth and warms with time, in response to the heat flux from above. The heat flux attains a maximum close to the inversion layer, and generally decreases towards the surface, while the downward momentum flux is again strongest at the surface, and decreases linearly with height.

MUSC cy38h1 overestimates the downward transport of heat by a factor of 3 near the top of the mixed layer, which becomes consequently too deep and too warm. The heat flux changes sign within the lowest 100 m, possibly in response to the heating from dissipation of E_k close to the surface. The downward momentum flux is overestimated by MUSC cy38h1 by a factor of 2. The wind speed profile has a pronounced point of inflection within the mixed layer, and consequently the wind maximum is too narrow compared with LES, and the shape of the hodograph distorted.

Introducing the EFB-model yields only a modest improvement on the results. The fluxes of heat and momentum are reduced in magnitude, leading to an improvement in the depth and temperature of the mixed layer, but the characteristic discrepancies compared with LES remain.
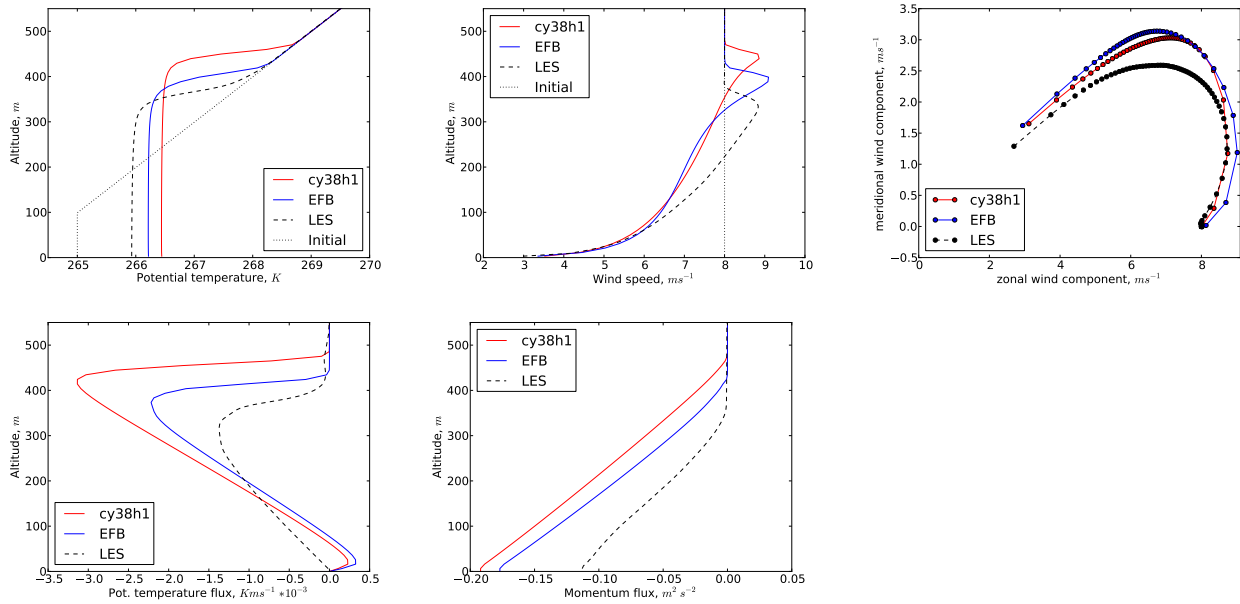

Figure 2: *Profiles of potential temperature, wind speed, heat flux, momentum flux, and a hodograph, plotted 20 hours after the initial time of the conventionally neutral experiment. MUSC cy38h1 is shown in red, MUSC with the EFB-closure in blue, and LES in black. The initial profiles of potential temperature and wind speed are drawn as dotted lines.*

5.3 GABLS3

Fig. 3 shows the evolution of wind speed and friction velocity for GABLS3. At a height of 200 m above ground, a typical nocturnal maximum develops in the evening, in response to frictional decoupling from the surface seen as a drop in the friction velocity. In MUSC cy38h1 this drop takes place with a delay compared with the observations, and $u_\star$ remains too strong. Accordingly, the onset of the nocturnal maximum is likewise delayed, and the intensity reduced, in common with the models examined for this case by Holtslag et al. (2013). Introducing EFB gives a partial improvement.

Fig. 4 shows profiles of potential temperature and wind speed at 12 UTC on the 2nd July 2006, 24 hour after the initial time. The sounding from De Bilt shows an unstable surface layer under a well mixed layer extending to height of nearly 2000 m. The same features are present in MUSC, and in particular, the impact of the EFB-model is small, despite the crude treatment of the turbulent length scale in unstable stratification. Probably the vertical diffusion in this case is dominated by the shallow convection scheme, while turbulence plays a secondary part. The profiles of wind speed are in approximate agreement with the observations in the lower part of the boundary layer. Higher up, the profiles given by the local profiler and the sounding at De Bilt disagree.
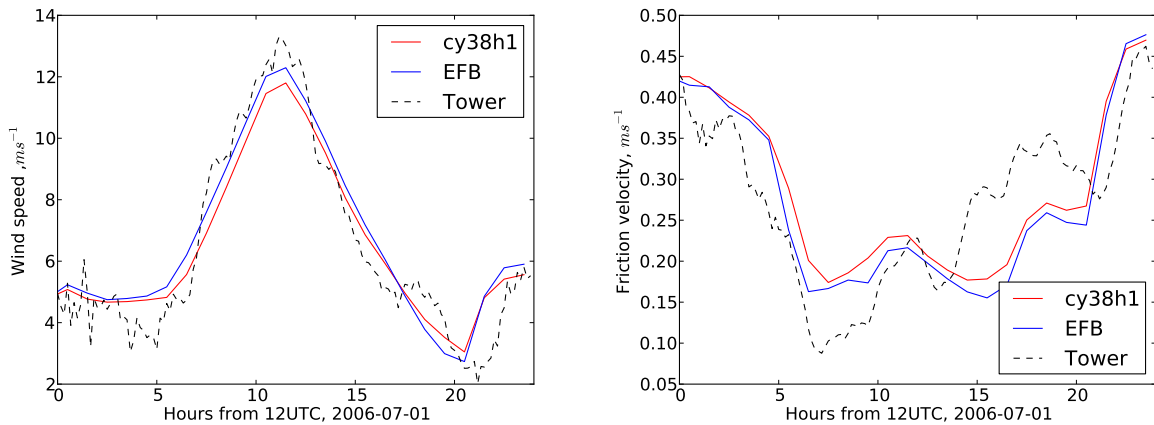
Figure 3: Wind speed at 200 m above ground (left) and surface friction velocity (right) in the GABLS3 experiments. MUSC cy38h1 is shown in red, MUSC with the EFB-closure in blue, and observations in black.

6 Summary and conclusions

The Energy Flux Budget turbulence closure model of Z13 has been implemented in the AROME physics parametrizations of HARMONIE, and tested in idealized and realistic single-column experiments with good results. 3-dimensional experiments need to be performed, to assess the possible impact in the presence of interactive dynamics, and hence, on the forecast scores.

The current implementation must be regarded as preliminary at least with respect to the treatment of unstable stratification, and extension of the (dry) EFB-model to a moist atmosphere. The calculation of the turbulent time scale (length scale) also needs further attention. The equilibrium time scale (eq. 7) is sensitive to its argument, and large changes from one time step to the next may easily lead to numerical instability unless some damping is applied. The formulation shown in 14 represents a compromise between simplicity and stability, but even applying this relaxation, the time step still had to be reduced from the original 60 s to 10 s to maintain stability when EFB was in use. The length scale obtained in this way showed clearly more variability in the vertical than the mixing length of Bougeault and Lacarrère (1989), so spatial diffusion should probably be taken into account as well, as recommended in Z13. These issues need to be resolved before the new closure can be applied in 3-dimensional experiments.
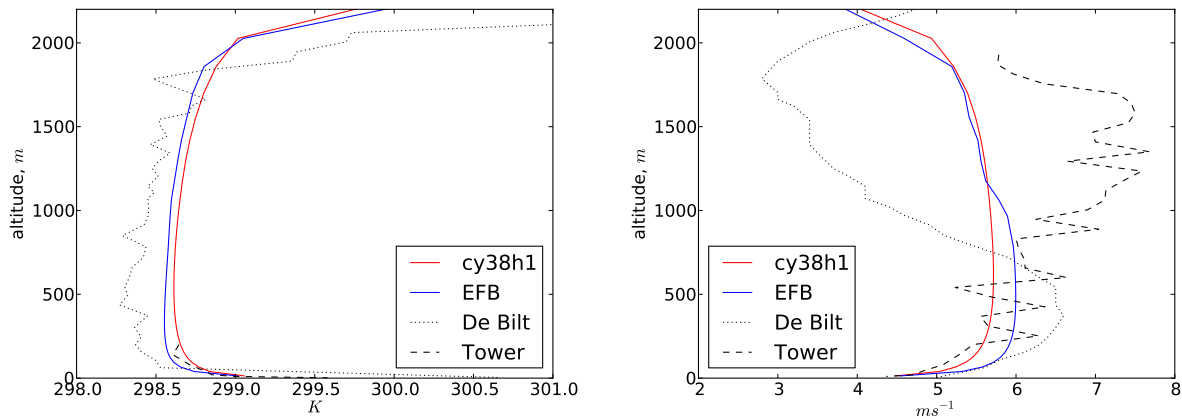
Acknowledgements

Figure 4: Potential temperature (left) and wind speed (right) at 12UTC on the 2nd of July 2006, 24 hours after the initial time in the GABLS3 experiments. MUSC cy38h1 is shown in red, MUSC with the EFB-closure in blue. Black dashed lines show data from the Cabauw tower (left) or wind profiler (right), dotted lines show the sounding at De Bilt some 30 km away.

References

Bougeault, P., and P. Lacarrère, 1989: Parameterization of orography-induced turbulence in a meso-beta scale model. Mon. Wea. Rev., 117, 1872-1890.

Holtslag, A. A. M., 2003: GABLS initiates intercom- parison for stable boundary layers. GEWEX News, No. 13, International GEWEX Project Office, Silver Spring, MD, 7âĂŞ8.

Holtslag, A. A. M. 2006: GEWEX Atmospheric Boundary-Layer Study (GABLS) on stable boundary layers. Bound.- Layer Meteor., 118, 243âĂŞ246.

Holtslag, A. A. M., G. Svensson, P. Baas, S. Basu, B. Beare, A. C. M. Beljaars, F. C. Bosveld, J. Cuxart, J. Lindvall, G. J. Steeneveld, M. TjernstrÃűm and B. J. H. Van De Wiel, 2013: Stable Atmospheric Boundary Layers and Diurnal Cycles: Challenges for Weather and Climate Models. Bull. Amer. Meteor. Soc. 94 1691-1706 http://dx.doi.org/10.1175/BAMS-D-11-00187.1

Zilitinkevich, S.S., T. Elperin, N. Kleeorin, I. Rogachevskii and I. Esau, 2013: A Hierarchy of Energy- and Flux-Budget (EFB) Turbulence Closure Models for Stably-Stratified Geophysical Flows. Boundary-Layer Meteorol 146, 341âĂŞ373 DOI: 10.1007/s10546-012-9768-8

Statistical condensation scheme:
Vertical variation of the saturation and convolution pdf

Daniel Martin Perez

Gema Morales Martin, Javier Calvo Sanchez

1   Introduction

In order to obtain the cloud water content and the cloud fraction inside a grid volume, the statistical subgrid condensation scheme considers a function that gives us the probability distribution of a certain parameter, related with the condensation, inside that grid volume. That function is called pdf (probability density function). In our case the parameter considered is the saturation deficit, that is the difference between the total water mixing ratio and the saturation mixing ratio. This parameter is normalized by its variance. So, in this scheme, the two important features to know are the pdf function and the variance of the saturation deficit.

Section 2 of this article is devoted to the inclusion of a new parameter for the variance that takes into account the vertical dependence of the saturation deficit. In the third section a new pdf function obtained from the convolution of two other functions is considered. Some verification results are presented in section 4, and finally the conclusions are drawn in section 5.

2   Saturation deficit variance

The problem of obtaining the variance of the saturation deficit has been studied by several authors (see, for example, Chaboureau and Bechtold, 2002). The variance depends on the turbulence and the convective activity, but it was shown that these terms were not enough to characterise it, so an extra term proportional to q_{sat} is used in the AROME physics of HARMONIE (de Rooy et al., 2010).

In this study, another term is going to be added to the variance. This term depends on the vertical variation of the saturation deficit. The idea behind it is that instead of a uniform distribution in the vertical within a model layer, a variance depending on the vertical variation of the saturation deficit is considered. For every grid volume a linear vertical variation of the saturation deficit inside it is considered exclusively, its valuesl ranging from a minimum value (S_{Vmin}) to a maximum (S_{Vmax}). In order to describe the distribution of the values of the saturation deficit, a rectangular function is chosen such that all the values of the saturation deficit between the maximum and the minimun inside the grid volume will have the same probability. The variance of a rectangular distribution is given by: $\frac{1}{2\sqrt{3}}(S_{Vmax} - S_{Vmin})$. This term has been included in the total variance.

It will be shown that the variance due to this term can be more important than the rest of the terms for the cells where there is a strong vertical variation of the saturation deficit.

Different ways of introducing this term in the total variance have been considered. Just by adding it to the rest of the terms:

$$\sigma = 2\sqrt{\sigma_{turb}^2 + \sigma_{conv}^2} + 0.02q_{sat} + \frac{1}{2\sqrt{3}}(S_{Vmax} - S_{Vmin}) \tag{1}$$

or including it inside the square root:

$$\sigma = \sqrt{(2\sqrt{\sigma_{turb}^2 + \sigma_{conv}^2})^2 + (0.02q_{sat})^2 + \left(\frac{1}{2\sqrt{3}}(S_{Vmax} - S_{Vmin})\right)^2} \qquad (2)$$

The first expression (equation 1) has been used in the 3D case and the verification with model version 37h12. The second (equation 2) produces smaller differences than expected in the verification, even though, it has been tested in a MUSC experiment (using v. 38h11) based on the ASTEX case with good results.

In general, when the variance increases, the cloud fraction also increases, although this isl not always the case.

2.1 ASTEX case

The proposed term (equation 2) has been tested with the single column version of HARMONIE v38h11 (MUSC) for the ASTEX [1] case. It has been run with the EDMF configuration and the statistical cloud scheme (CMF_UPDRAF='DUAL', CMF_CLOUD='STAT'). The result for the cloud water content and the cloud fraction after 40 hours of simulation are shown in figure 1. It is noticed that when the vertical variation of the saturation is considered in the variance, the cloud layer rises, the maximum values are slightly lower than those for the reference case and there appear two relative maximum as in the LES. Nevertheless, the result is still far from the LES values.

(a) Cloud water content   (b) cloud fraction
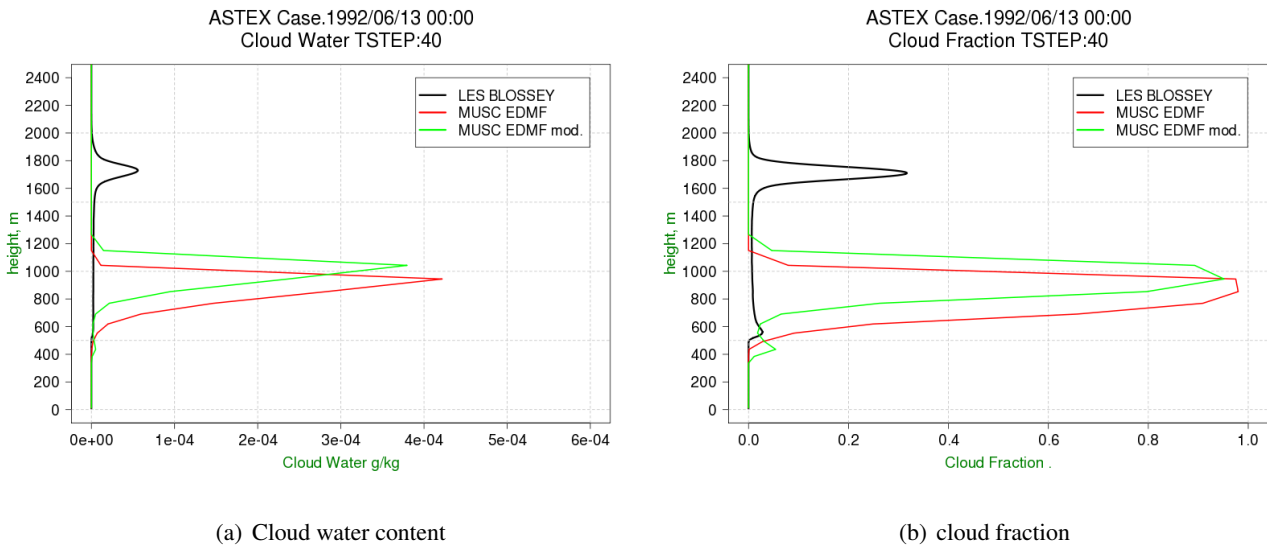
Figure 1: *Profiles of the cloud water content 1(a) and the cloud fraction 1(b) after 40 hours of simulation. Comparison of the MUSC reference (red lines), modified variance (green line) and LES (black line).*

The contribution of the vertical variation term is especially important at the top of the cloud where the vertical variation of the other terms is bigger. After 40 hours of simulation, the other terms of the variance, especially the convective term, have changed as well (figure 2).

---

[1]Atlantic Stratocumulus Transition EXperiment

(a) Case without added term　　　　　　　　　　　　　　(b) With added term

Figure 2: Profiles of the different terms that contribute to the variance after 40 h of simulation. 2(a) no modification, 2(b) with the term proposed. The shadowed area represents the cloud layer

In figure 2 it can be seen that the values of the vertical variation dependent term are higher than the rest of terms at the top of the cloud and above, and it also has important values under the cloud base. When this term is considered, the mass flux term also increases while small differences are observed for the rest of the terms.
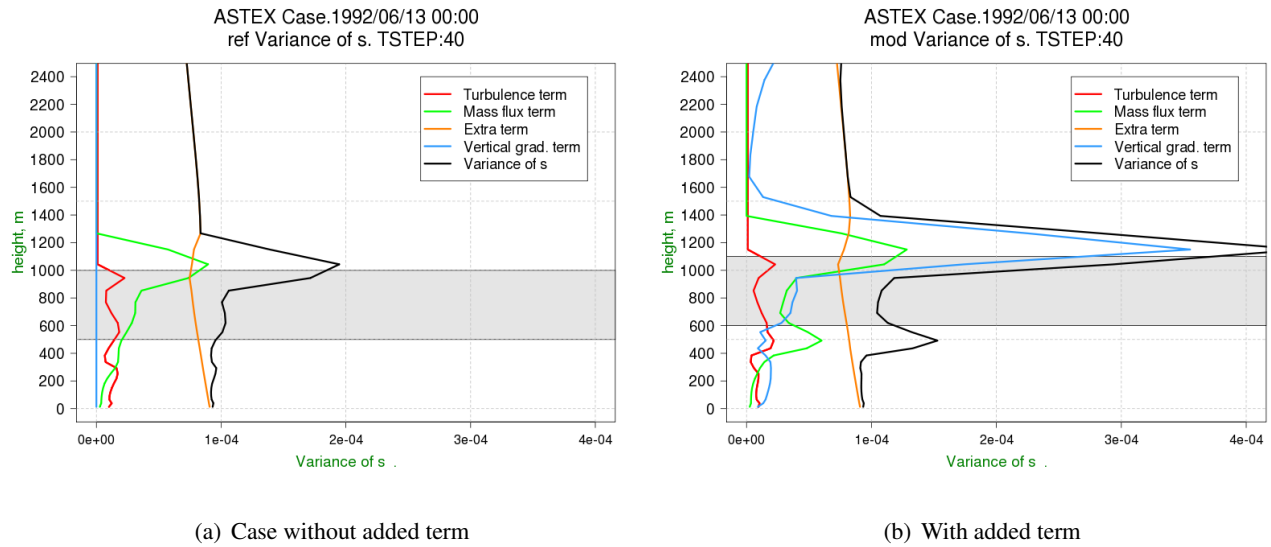
2.2　3D example

A 3D Harmonie case has been run with and without the modification of the variance to illustrate the effects that the addition of this new term can produce. The version of the model was 37h12 and the area covers the Iberian peninsula.

The cloud fraction forecasted is compared with the cloud type product of NWCSAF[2](http://www.nwcsaf.org). Figure 3(b) shows the cloud type product corresponding to the 8th of March 2013 at 00H, where the white and light blue colours represent the high clouds, while figures 3(a) and 3(c) are the plots of the high cloud fraction obtained for time step H+24 of the 07/03/2013 00H run of HARMONIE model.

---

[2]Support to Nowcasting and Very Short Range Forecasting. Satellite Application Facilities

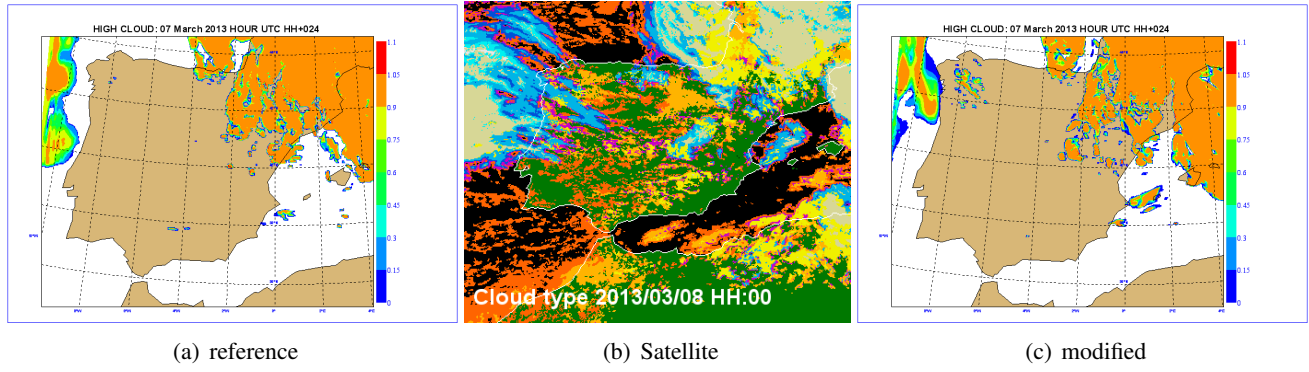| (a) reference | (b) Satellite | (c) modified |

Figure 3:   *Figure 3(b) shows a Cloud type NWCSAF product where white and light blue colours represent high clouds. Figures 3(a) and 3(c) shows the cloud fraction of high clouds for the reference and modified cases respectively*

It can be seen that new clouds appear in the center and in the north west of the Iberian peninsula (fig 3(c)), with respect to the reference case (fig 3(a)), which are in better agreement with the observations. For the clouds near the western coast, the analysis is not so straightforward, because despite there is a bigger extension in the north there is a reduction of the clouds near the coast of Portugal. Another effect that can be observed is that the contours of the cloud are not so sharp, there is a gradual variation in the cloud fraction in the borders.

3   Probability density function

In this section, a probability density function for the normalized saturation deficit is going to be introduced. In the AROME physics the cloud water content and the cloud fraction functions have been obtained empirically from several test cases (Chaboureau and Bechtold, 2002), but there is no pdf from which those functions could be obtained. It is presumed that using functions for the cloud fraction and cloud water obtained from a unique pdf will give better results as far as they are not too much different from the experimental ones.

The pdf proposed is the result of the convolution of a gaussian function (eq. 3) and a exponential function (eq. 4). The reason for choosing these two functions is that they were among the first proposed for parameterization of cloud fraction and also because their simplicity. With the convolution we try to keep the characteristics of both functions in the final result, so that one can have more weight in the final pdf than the other just by choosing the value of two parameters. This will give us a high flexibility for future research of the cloud scheme.

Given the mean value of the normilized saturation deficit in a grid volume (s_f) the pdf corresponding to the exponential function has the expression:

$$f(s) = \begin{cases} \frac{1}{\sigma_f} e^{-\frac{(s-s_f)}{\sigma_f} - 1} & s \in (-\sigma_f + s_f, \infty) \\ 0 & s \in (-\infty, -\sigma_f + s_f) \end{cases} \tag{3}$$

Being s the normalized saturation deficit. While the gaussian function is given by:

$$g(s) = \frac{1}{\sqrt{2\pi}\sigma_g} e^{-\frac{(s-s_g)^2}{2\sigma_g^2}} \tag{4}$$

where s_g is the mean value.

The convolution has interesting properties. As both are unit area functions, the convolution will also be an unit area function.

$$\phi(s) = \int_{-\infty}^{\infty} f(s - s')g(s')ds' \tag{5}$$

The expression obtained is given by equation (6)

$$\phi(s_\phi, s) = \frac{1}{2\sigma_f} e^{\left(-\frac{s-s_\phi}{\sigma_f} + \frac{\sigma_g^2}{2\sigma_f^2} - 1\right)} \left[1 - erf\left(-\frac{s - s_\phi}{\sqrt{2}\sigma_g} - \frac{\sigma_f}{\sqrt{2}\sigma_g} + \frac{\sigma_g}{\sqrt{2}\sigma_f}\right)\right] \tag{6}$$

Where $s_\phi = s_f + s_g$.

Figure 4(a) shows an example of an exponential and a gaussian pdf together with the convolution of the two of them, while in figure 4(b) it is shown two pdfs obtained for differente paramenters of σ_f and σ_g.

(a) Comparison of pdfs

(b) Probability density function
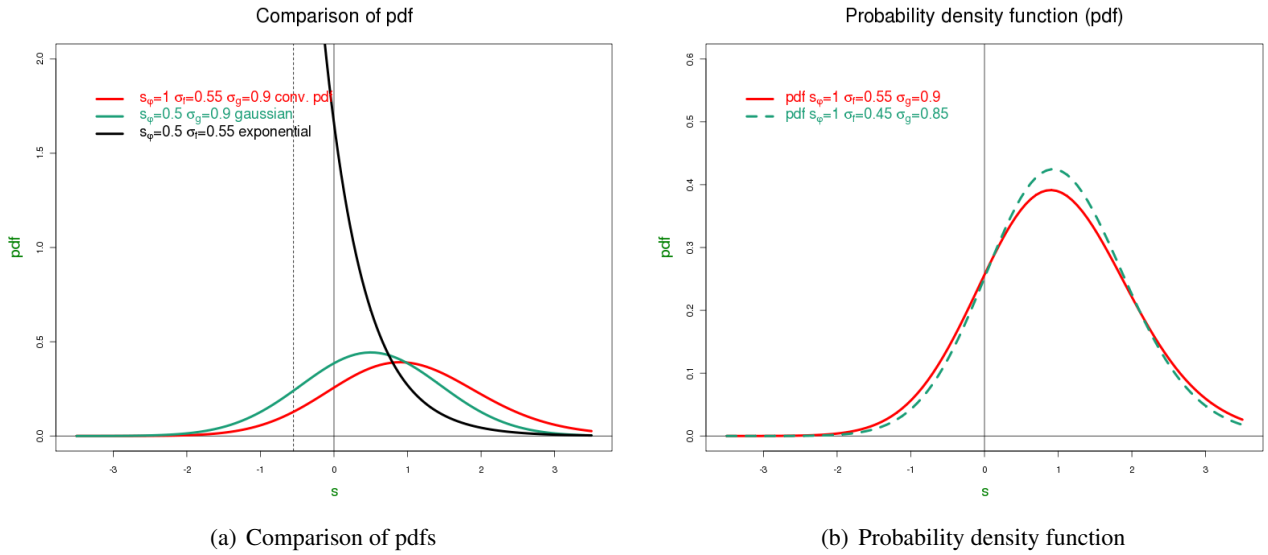
Figure 4: *Plots of the probability density function. In 4(a) it is shown the exponential and the gaussian function as well as the convolution function of both of them. In 4(b) two pdf obtained by the convolution for different paramenters.*

The cloud fraction, N, is obtained solving the integral (7), the result is given by the equation (8).

$$N(s_\phi) = \int_0^\infty \phi(s)ds \tag{7}$$

$$N(s_\phi) = \frac{1}{2}\left[1 + erf\left(\frac{s_\phi - \sigma_f}{\sqrt{2}\sigma_g}\right)\right] + \frac{1}{2}e^{\left(\frac{s_\phi}{\sigma_f} + \frac{\sigma_g^2}{2\sigma_f^2} - 1\right)}\left[1 - erf\left(\frac{s_\phi - \sigma_f}{\sqrt{2}\sigma_g} + \frac{\sigma_g}{\sqrt{2}\sigma_f}\right)\right] \tag{8}$$

The cloud water content, $\frac{q}{\sigma_s}$, is obtained from the integral (9).

$$\frac{q(s_\phi)}{\sigma_s} = \int_0^\infty s\phi(s)ds \tag{9}$$

$$\frac{q(s_\phi)}{\sigma_s} = \frac{1}{2}s_\phi \left[1 + erf\left(\frac{s_\phi - \sigma_f}{\sqrt{2}\sigma_g} \right) \right] + \frac{\sigma_g}{\sqrt{2\pi}} e^{-\left(\frac{s_\phi - \sigma_f}{\sqrt{2}\sigma_g} \right)^2}$$
$$+ \frac{1}{2}\sigma_f e^{\left(\frac{s_\phi}{\sigma_f} + \frac{\sigma_g^2}{2\sigma_f^2} - 1 \right)} \left[1 - erf\left(\frac{s_\phi - \sigma_f}{\sqrt{2}\sigma_g} + \frac{\sigma_g}{\sqrt{2}\sigma_f} \right) \right]$$

(10)

The cloud fraction (eq. 8) and the cloud water content (eq. 10) depend on the parameters σ_f and σ_g that are present in the exponential and the gaussian functions respectively. Our purpose is to get expressions for the cloud fraction and the cloud water content that don't differ too much from those obtained experimentally. The values of σ_f and σ_g have been calculated so that they minimise the integral of the square root of the square of the difference between the cloud water content from the pdf and the expression used in AROME. The values obtained are: $\sigma_f = 0.55$, $\sigma_g = 0.9$. In a similar way those parameters can be calculated using the cloud fraction instead of the cloud water. In this case: $\sigma_f = 0.45$, $\sigma_g = 0.85$. The fact that they are similar values shows that the formulae are close to come from the same pdf.

The cloud fraction is plotted in figure 5(a) and the cloud water content in figure 5(b).

(a) Cloud water content   (b) cloud fraction
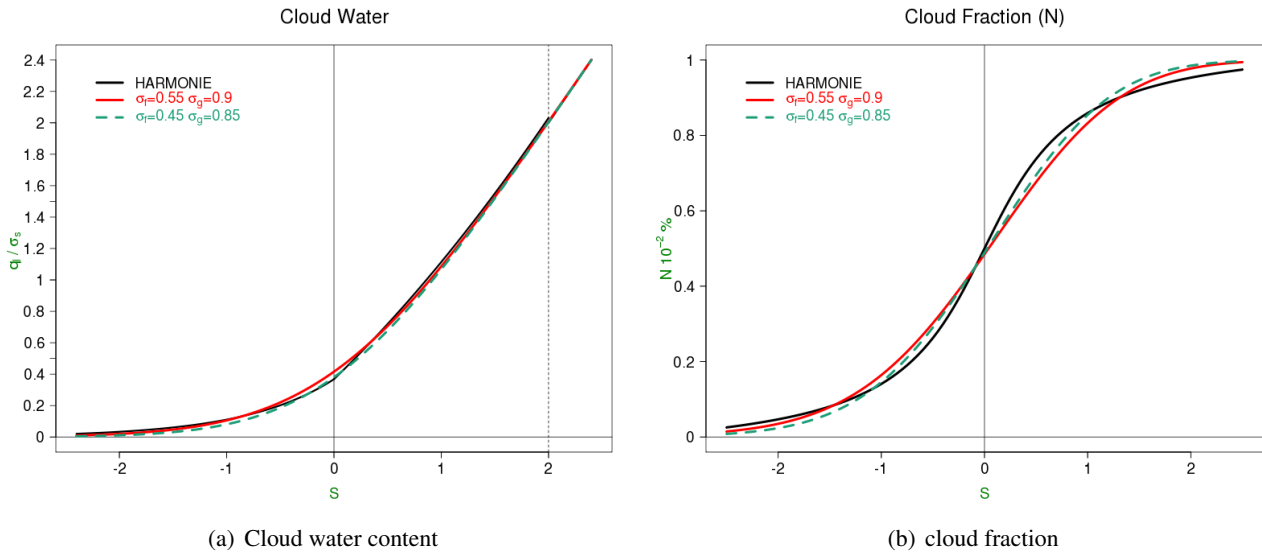
Figure 5: Plots of the cloud water content and the cloud fraction. In black the plot corresponding to the model, while in red and green the ones corresponding to the convolution pdf with $\sigma_f = 0.55$, $\sigma_g = 0.9$ (red) and $\sigma_f = 0.45$, $\sigma_g = 0.85$ (green)

In general, the curves are very similar to the one of the model (in black). There are more differences in the cloud fraction than in the cloud water content, specially when the number of octas is higher than 5.

4   Verification results

The verification has been done for two months, March and August 2013. From the plots of the frequency distribution for cloud cover (figures 6(a) and 6(b)), it can be concluded that the model behaves strongly as all or nothing scheme for the cloud cover, overestimating the overcast and clear sky situations and underestimating the partial cloudiness.

(a) March 2013

(b) August 2013

Figure 6: Frequency for the cloud cover. The reference experiment (aib_37h12, red line), with the modified variance (cond_37h12_op, green line), new pdf function (nwc_37h12_A, blue line), considering both modifications (nwcond_37h12_op, pink line).
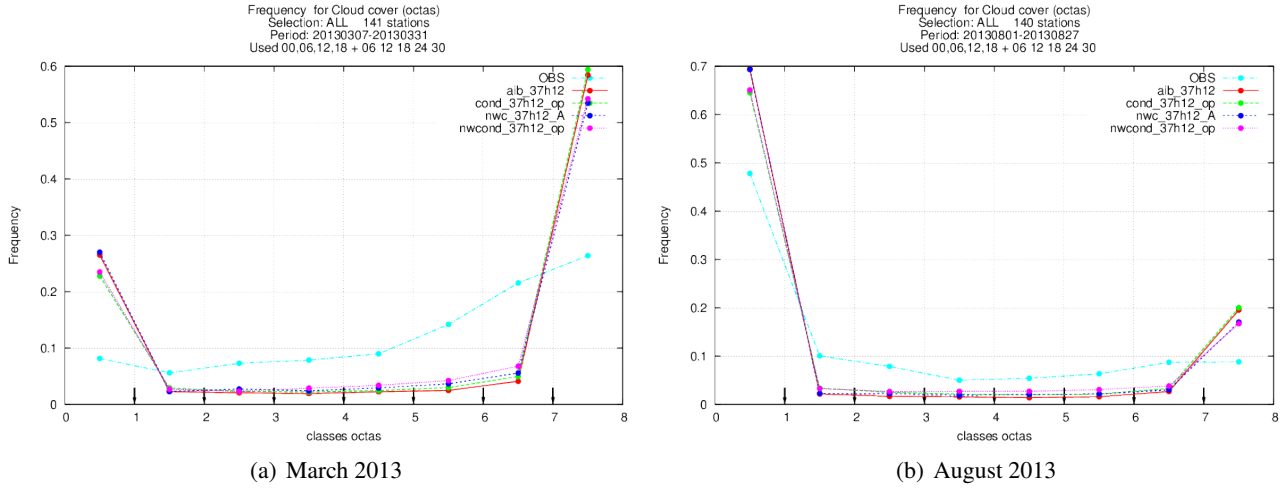
March was a very cloudy month. For the reference case, the bias was small (figure 1). This means that for this case the overestimation of the cloud coverage given by the overcast cases compensates the underestimation for the partial cloudiness cases. August was a different case, not as cloudy as March. The bias was negative, -0.633. This result can be explain because the frequency of overcast cases was very low and, although the model overestimates them, the relative importance was not as high as for March, so there is an underestimation of the cloud coverage. Despite the difference in the bias, there is not such a difference in the RMSE between the two months.

When the modification of the variance proposed in section 2 is considered (experiment cond_37h12_op), the cloud cover forecasted increases, as well as the bias, while the RMSE decreases for both months. The experiment nwc_37h12_A corresponds to the introduction of the pdf described in section 3 (with $\sigma_f = 0.55$, $\sigma_g = 0.9$). The results are not as good as for the previous experiment, although the RMSE for August is better than for the previous experiments. There is a third experiment, nwcond_37h12_op, that combines both modifications (variance and new pdf) with the best results for the RMSE and better results for the bias than the reference case.

Table 1: Bias and RMSE for cloud cover for all the forecast lenghts

	March 2013		August 2013	
Experiment	**Bias**	**RMSE**	**Bias**	**RMSE**
aib_37h12 (ref.)	-0.173	2.850	-0.633	2.763
cond_37h12_op (var term)	0.036	2.732	-0.386	2.676
nwc_37h12_A (new pdf)	-0.335	2.835	-0.702	2.610
nwc_37h12_op (var term + new pdf)	-0.115	2.699	-0.532	2.606

In figure 7 the equitable threat score (ETS) for the cloud cover and the precipitation are shown. The score for the precipitation (figures 7(c) and 7(d)) shows a neutral impact (compared with the reference case, red line), while the best results are obtained for the case that considers only the modification of the variance term. The ETS for the cloud cover (figures 7(a) and 7(b)) is positive for all the experiments. In the case of the modification of the variance the score is better for low thresholds while the pdf improves the score for high thresholds. Only when both modifications are considered together, the score of the cloud cover is improved for
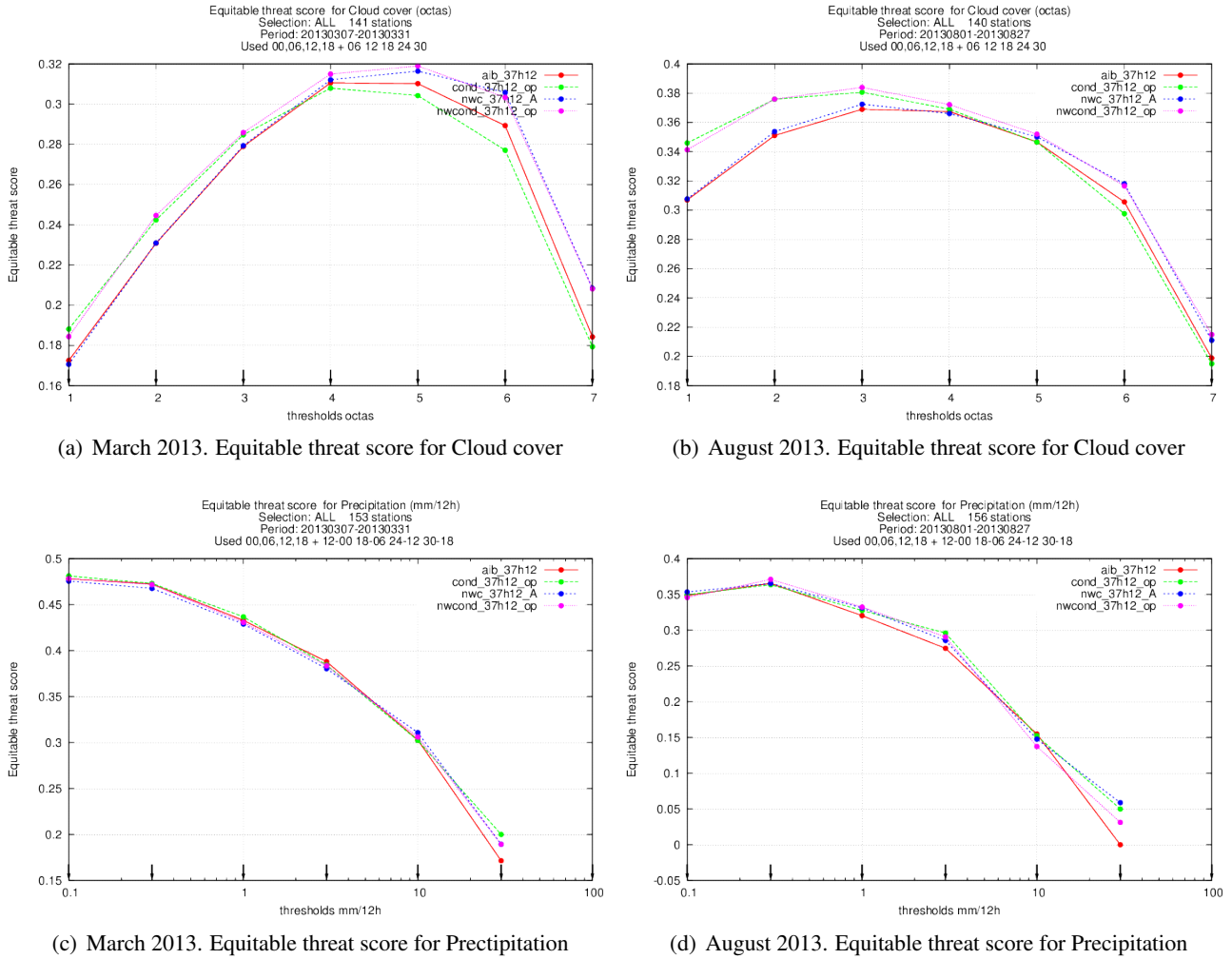
every threshold.

(a) March 2013. Equitable threat score for Cloud cover

(b) August 2013. Equitable threat score for Cloud cover

(c) March 2013. Equitable threat score for Prectipitation

(d) August 2013. Equitable threat score for Precipitation

Figure 7: Plots of the equitable threat score for cloud cover and precipitation for March and August 2013.

5   Conclusions

Two different modifications of the statistical cloud scheme in HARMONIE have been proposed. The first one (section 2) is the inclusion of another term in the variance of the saturation deficit. This term depends on the vertical variation of the saturation deficit. It has been tested with the ASTEX case using the MUSC model with good results, showing that the term can be important at the top and the bottom of the clouds and can modify other terms of the variance like the one corresponding to the mass flux. In test with the complete model, it has been shown that this term causes a smoother variation of the cloud fraction in the border of the clouds. The RMSE of the cloud cover also improves for a summer month as well as for a rainy spring month, and no negative impact for the precipitation and the rest of the variables have been observed.

The second modification proposed is a new pdf (section 3) that is the result of the convolution of a exponential and a gaussian function. It depends on two parameters that have been calculated such that the cloud fraction and the cloud water content stay close to the reference. Therefore it is considered a big difference, but it provides room for future improvements. In fact, the bias of the cloud cover is worse for this case, although the RMSE improves. The ETS of the cloud cover gives better results for thresholds over 5 octas and there is a

neutral impact for the precipitation.

When both modifications are considered together, we obtain the best scores for the cloud cover. The ETS for the cloud cover is better for every threshold while the RMSE is lower for this case than for the rest of the cases and for the months considered, March and August.

Although the expressions obtained for the cloud water content and the cloud fraction from this new pdf are more complex, their dependence on two parameters permits some further research in order to characterise cloud types with different values of the parameters.

References

Bougeault Ph., Modeling the Trade-Wind Cumulus boundary Layer. Part I: Testing the Ensemble Cloud Relations Against Numerical Data, J. Armos. Sci., 28, 2414-2428, 1981

Chaboureau J.P., P. Bechtold, A Simple Cloud Parameterization Derived from Cloud Resolving Model Data: Diagnostic and Prognostic Applications, J. Atmos. Sci., 59, 2362-2372, 2002.

de Rooy W., C. de Bruijn, T. Sander, R. Neggers, P. Siebesma, J. Barkmeijer, Experiences with Harmonie at KNMI, HIRLAM Newsletter, 56, 2010.

Mellor G.L., The Gaussian Cloud Model Relations, J. Atmos. Sci., 34, 356-358, 1977.

Naumann A.K., A. Seifert, J.P. Mellado, A refined statistical cloud closure using double-Gausisian probability density functions, Geosci. Model Dev., 6, 1641-1657, 2013

Improving wintertime low cloud forecasts in AROME: sensitivity experiments and microphysics tuning

Balázs Szintai[1], Eric Bazile[2], Yann Seity[2]

[1]Hungarian Meteorological Service

[2]Meteo-France

Introduction

In this paper the performance of the AROME model is investigated during winter anti-cyclonic cases. During strong winter anticyclones, cold air resides near the surface and no significant fronts occur which could sweep out the cold air from the Pannonian Basin. As solar irradiation is quite low in this season the morning fog is not dissolved, it is only elevated to about 300-500 m above ground level and a stratus layer is formed which stays constant during daytime. The elevation of the fog layer is basically caused by the high albedo of the cloud top, which causes the cloud top to cool even after sunrise. Consequently, condensation occurs at the cloud top and the fog starts to elevate. This kind of situation can typically last for 7-10 days over Hungary. Due to low wind speeds and constant cloudiness the mixing height of pollutants is relatively low and the concentration of air pollutants can rise significantly. Generally, the AROME model – similarly to other NWP models applied at the Hungarian Meteorological Service (OMSZ) – is not very successful in simulating this weather phenomenon. The stratus layer tends to be dissolved by the model by early afternoon and consequently afternoon temperatures are overestimated and night temperatures are underestimated by the model. The difficulty of this situation from the modelling point of view is that several processes (radiation, turbulence, microphysics) are interacting to form the constant stratus layer.

To tackle these modelling challenges, a French—Hungarian bilateral project was initiated in autumn 2012 for two years. The project is financed by the government of France and Hungary and consisted of eight research stays.

Experimental setup and case study

For the experiments, the AROME non-hydrostatic model was used at 2.5 km horizontal resolution and 60 vertical levels, with similar settings of the physical parameterizations as used operationally at OMSZ (cy36_t1). The selected case study (30th November 2011) was a late autumn day with stratus cover of nearly the whole Pannonian Basin. The simulation started at 00 UTC on 30th November 2011 and lasted 14 hours. In the reference simulation initial and lateral boundary conditions for AROME were interpolated from the operational ARPEGE run.
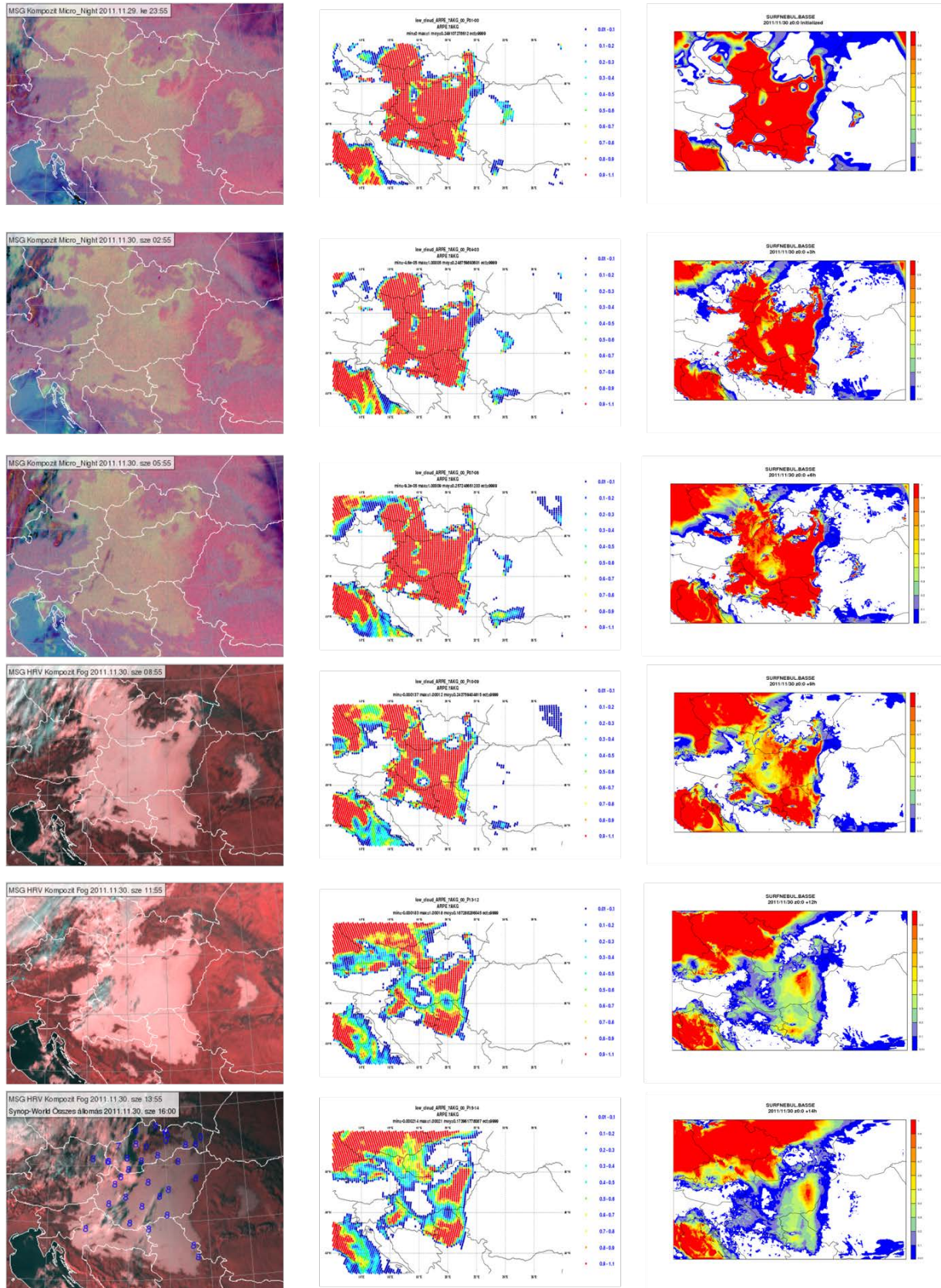
Figure 1. Reference runs of the selected case study (30th November 2011). Satellite observations (first column), ARPEGE (second column) and AROME (third column) forecasts of low cloud cover. Each row refers to a single forecast time: 00 UTC (initial time), 03 UTC, 06 UTC, 09 UTC, 12 UTC and 14 UTC.

Figure 1 shows the ARPEGE and AROME forecasts and the satellite observations of low cloud cover. At the start of the simulations both models diagnose the spatial extension of the fog well. As the simulation proceeds, both models erroneously dissolve the fog over the western part of Hungary. Over the eastern part of the country, ARPEGE keeps most of the low clouds while AROME dissolves a considerable part of the fog.

First, the performance of ARPEGE and AROME over the western part of Hungary was investigated. Over this area both models dissipate the fog and give wrong low cloud forecasts. Vertical profiles of wind speed (Fig. 2) in AROME show that there is a substantial difference between the western and the eastern part of Hungary. Over the eastern part, wind at higher levels (i.e. at and above the height of the stratus layer) is constant or even decreasing during the simulation, while over the western part wind speed is increasing to about 14 m/s above the boundary layer. It is assumed that over the western part this strong wind increases the mixing in the model in the boundary layer and dissolves the fog/stratus. As this effect is considered to be a large scale phenomenon, in the following the investigation is focused on the eastern part of Hungary.
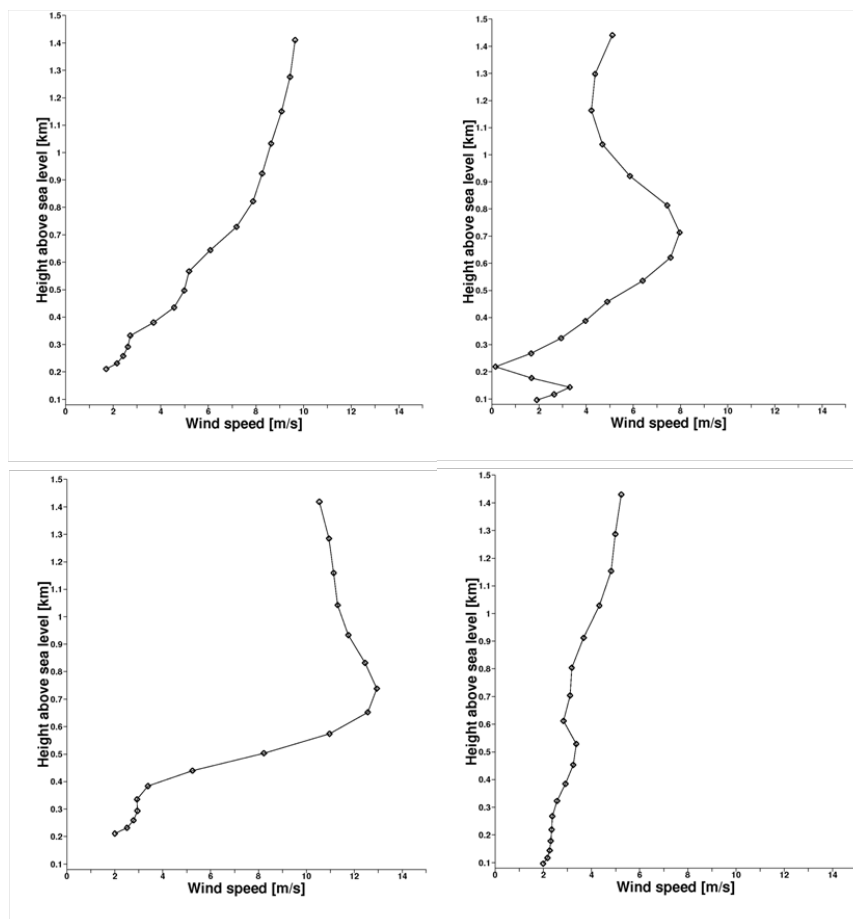

Figure 2. Profiles of wind speed in AROME for a grid point in the western (left column) and in the eastern part (right column) of Hungary for 30th November 2011. Forecast range is +1h (01 UTC, upper row) and +14h (14 UTC, lower row).

Sensitivity tests

Several sensitivity tests were completed to diagnose which part of the AROME model is responsible for the wrong low cloud forecast.

The reference run was not using data assimilation, AROME was run in dynamical adaptation mode, i.e. the initial state was produced by interpolating the ARPEGE analysis to the AROME grid. In the first sensitivity test the impact of data assimilation was tested. An assimilation cycle with 5 day spin-up was run for the selected case. In the atmosphere, the 3DVAR method was applied using conventional (synop and radiosounding) observations, while on the surface the optimal interpolation method (OI_MAIN) was applied. The application of data assimilation has an overall neutral but spatially variable impact on the low cloud forecast in AROME (Fig. 3a).

The second sensitivity test aimed at quantifying the impact of lateral boundary conditions. In this experiment boundary conditions from the ECMWF/IFS global model were used instead of the ARPEGE global model. Impact on the stratus cloud cover over Hungary was rather small (Fig. 3b).

The choice of the subgrid statistical cloud scheme has also been investigated. At the Hungarian Meteorological Service a diagnostic formulation is used operationally (`LOSIGMAS=.FALSE.`, used also in the reference run), while at Meteo-France a prognostic one (`LOSIGMAS=.TRUE.`). As experienced for several other case studies the prognostic formulation gives much less low clouds and tends to produce a low cloud cover value of "zero or one". For the given stratus case, the prognostic formulation gives much worse results than the reference (Fig. 3c).

The impact of turbulence parameterization has also been tested. It was supposed that Turbulent Kinetic Energy (TKE) is too high in the boundary layer and too strong mixing dissipates the fog in AROME. To test this assumption, the dissipation rate for TKE was increased. Surprisingly, this resulted in even less stratus (not shown). In a second experiment, the dissipation rate was decreased (from 0.85 to 0.2) which resulted in slightly more low clouds in AROME (Fig. 3d). This model behaviour is still under investigation. A preliminary explanation is the following: the lower dissipation rate results in higher TKE, which decreases the cloud cover due to stronger mixing, however, higher TKE values increase the subgrid variability in the statistical cloud scheme which in turn creates more clouds. Apparently, there seem to be two compensating mechanisms in the model when TKE is increased, which results in a modest change in cloud cover.

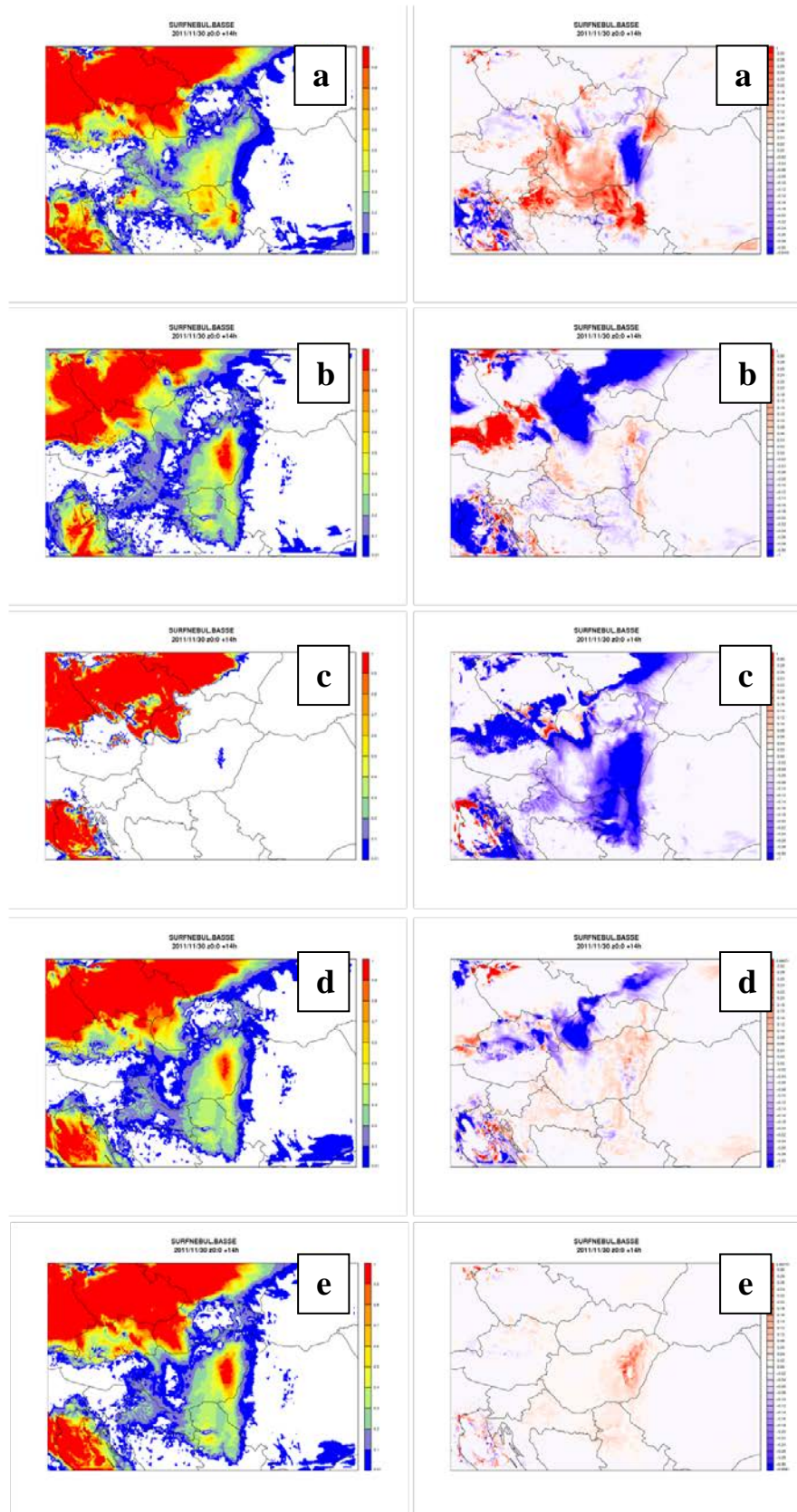*Figure 3. Sensitivity experiments for AROME on 30th November 2011 at 14 UTC: (a) data assimilation, (b) lateral boundary conditions, (c) statistical cloud scheme, (d) turbulence parameterization, (e) cloud inhomogeneity factor in the radiation parameterization. Left column: low cloud cover; right column: low cloud cover difference (experiment-reference; red colours indicate more low clouds in the experiment).*

It was assumed that one reason for the dissipation of stratus in AROME could be that after sunrise more short wave radiation is transmitted through the fog layer in the model than in reality. Consequently, surface downward short wave radiation simulations were verified with the radiation measurement network of the Hungarian Meteorological Service, which consists of 39 stations. For the present case study, only four stations were selected in the eastern part of Hungary, where stratus was present the whole day both in reality and in ARPEGE, but not in AROME. Time series show that ARPEGE slightly overestimates the short wave radiation after sunrise, and the reference version of AROME simulates even higher values than ARPEGE. One of the main differences in the radiation settings of the two models is the value of the long wave inhomogeneity factor, which accounts for an increased radiation transfer in clouds. This parameter is set to 0.9 in ARPEGE and 0.7 in AROME. To test the impact of this parameter both the short wave (RSWINHF) and long wave (RLWINHF) inhomogeneity factors were set to 1.0 in AROME. These settings result in a lower downward short wave radiation flux at the surface in AROME (Fig. 4). The increased optical thickness of clouds (RSWINHF=LSWINHF=1.0) and the consequently slower heating of the boundary layer slightly increases low cloud cover over the eastern part of Hungary (Fig. 3e). In areas with frontal activity (north of Hungary) the modification has no impact on cloud cover.
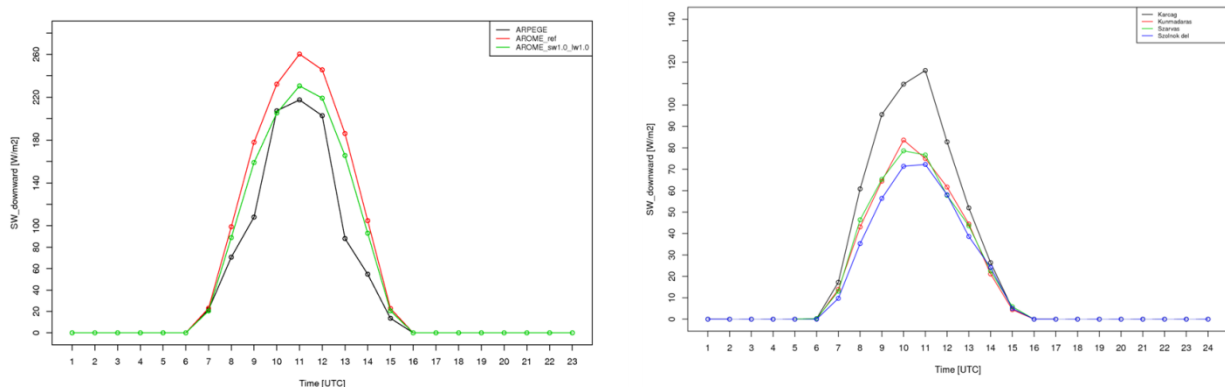

Figure 4. Simulated (left) and observed (right) time series of downward short wave radiation on the surface on 30th November 2011. Model values are horizontally averaged on a 1ºx1º box, which includes the ground measurements plotted on the right panel. Note the scale difference of the vertical axes between the two plots.

As shown in Fig. 4, even with the modification of the inhomogeneity factor the simulated shortwave downward radiation of AROME is about three times as large as the measurements around noon (220 W/m^2 in AROME as compared to 70 W/m^2 in the measurements). To validate whether this overestimation is caused by the error in the radiation or in the microphysics scheme, the one-dimensional version of the AROME model (MUSC) was used. It was found that if the thickness of the cloud is enlarged by a factor of two (a value assumed to be close to reality) and the cloud liquid water content is increased from 0.1 g/kg to 0.3 g/kg (the latter is a typical value for continental stratus according to Hoffmann and Roth, 1989) then at 12 UTC the simulated downward shortwave radiation is close to observations. Consequently, the radiation parameterization is most likely not responsible for the wrong low cloud simulation of AROME.
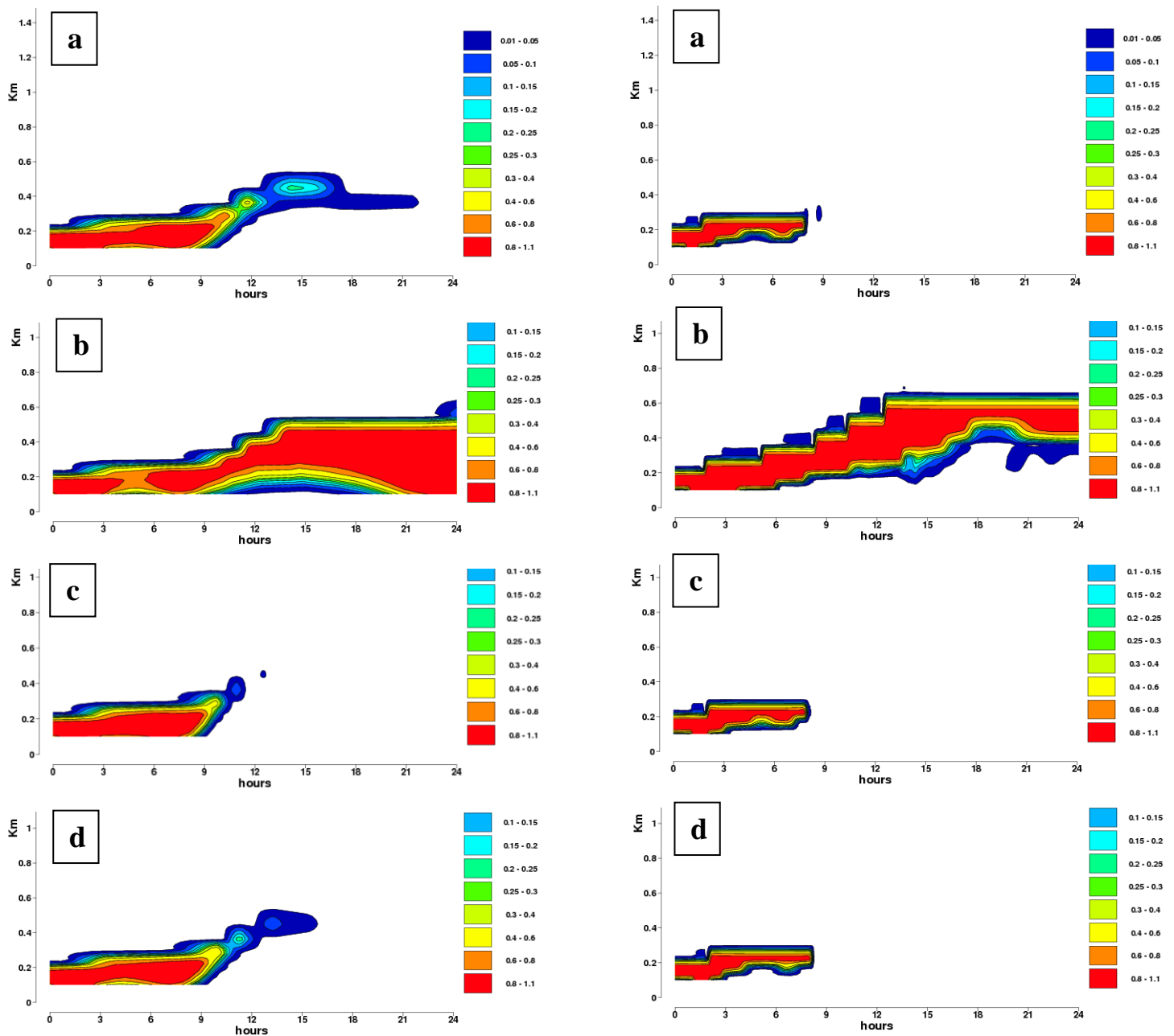
Figure 5. 1D experiments : sensitivity tests on surface. Cloudiness time evolution in ARPEGE (left) and in AROME (right). (a) VEG=100% SWI=0.8, (b) VEG=0, SWI=0.8, (c) VEG=100% SWI=0, (d) VEG=0 SWI=0.

Sensitivity tests on surface parameters were performed in 1D. AROME and ARPEGE 1D use the same initial vertical profiles. Only a few variables are changed in the initial surface file. Figure 5 shows that this case is very sensitive to the vegetation fraction (VEG) and soil wetness index (SWI). With a bare (VEG=0) and moist (SWI=0.8) ground, AROME is not able to dissipate the cloud even if the cloud base and top are higher than in ARPEGE. In that case, moisture is easily taken from the first ground layer. With VEG=100% (a and c), it is more difficult to feed the atmosphere with moisture coming from the ground. Indeed water has to be taken by the roots and is then provided to the atmosphere via plant evapotranspiration. Moreover, a too dry soil (SWI = 0) dissolves the fog during the afternoon,

51

both in ARPEGE and in AROME (c and d). Consequently, differences in soil moisture and/or vegetation fraction may partly explain differences observed between ARPEGE and AROME 3D runs.

Tuning of microphysics

After the above described sensitivity experiments the microphysics parameterization of AROME has also been investigated. It was found that for wintertime stratus cases the AROME model often produces light precipitation. In the model, if the temperatures close to the surface are above 0 °C, the precipitation phase is liquid and the amount (typically around 0.3—0.5 mm/12h) is close to observations and other operational models. Otherwise, for the negative temperature the precipitation phase is solid and AROME gives higher values than the observations or other models (Fig. 6a).
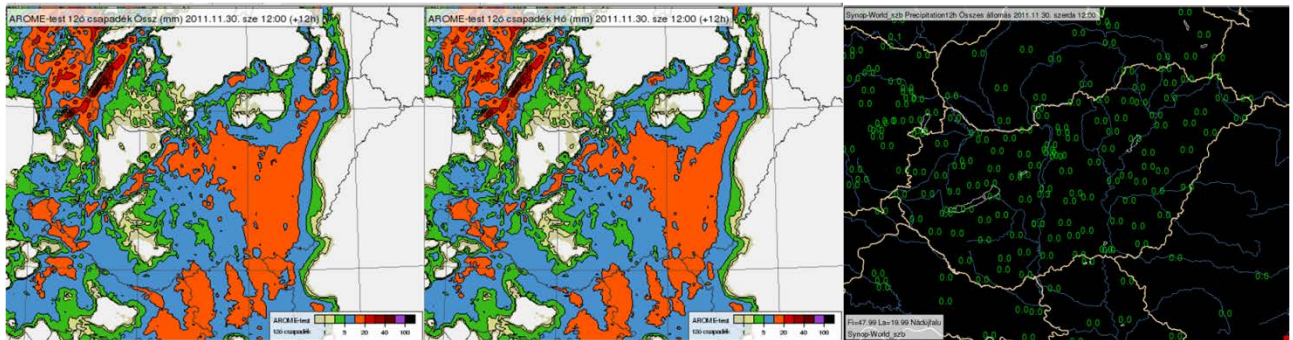
Figure 6a. Accumulated 12h precipitation from AROME at 30th November 2011 at 12 UTC (+12h forecast) compared to synop observations. Left: total precipitation from AROME (for readability, values are multiplied by 100), middle: snow from AROME (for readability, values are multiplied by 100), right: synop observations (in mm/12h).
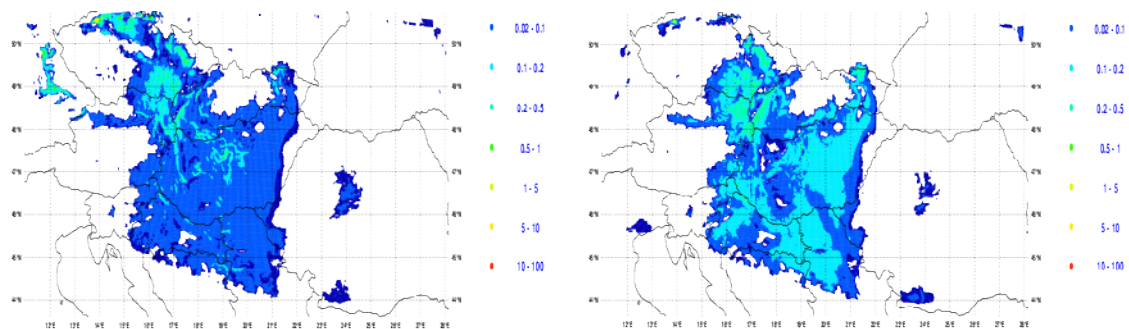

Figure 6b. Accumulated 12h snowfalls at 30th November 2011 at 12 UTC (+12h forecast) : left: ARPEGE physics at 2,5km, right: AROME.

This behaviour is not due to dynamics or horizontal/vertical resolution differences, as it is not present in a forecast using ARPEGE physics but AROME dynamics on the AROME grid (Figure 6b). It may come from differences in microphysics processes (Lopez scheme in ARPEGE, ICE3 in AROME). To understand the problem, DDH budget profiles of microphysics were investigated from AROME (Fig. 7).
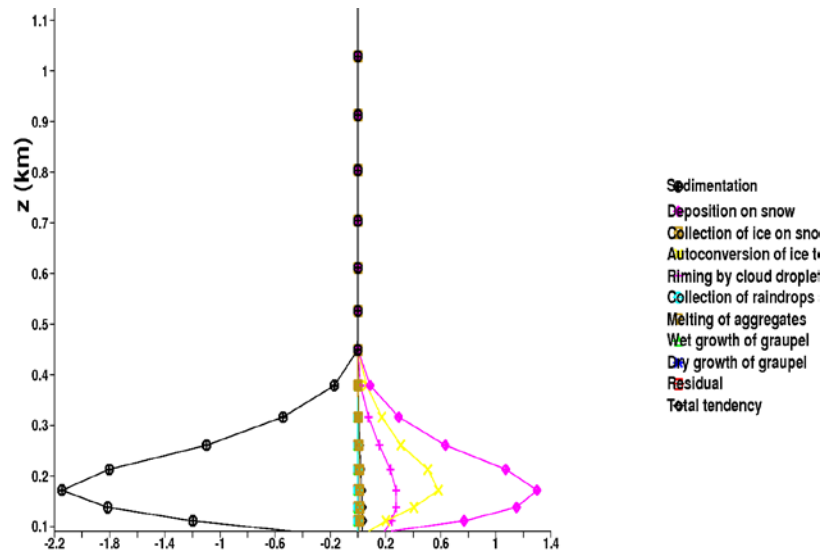
Figure 7. DDH budget profiles from the AROME reference run between 06 and 00 UTC on 30th November 2011. The DDH averaging was made on a 1 degree x 1 degree domain over the eastern part of Hungary.

It was found that for the given case three processes are responsible for the generation of falling snow: deposition, autoconversion and riming. The deposition process does not exist in ARPEGE Lopez scheme.   That may explain differences shown in Figure 6b. It was assumed that the largest uncertainty lies in the parameterization of autoconversion, so this process was selected for a tuning exercise. To reduce the solid precipitation in the model, the critical specific humidity value above which autoconversion could occur was increased from 0.02 g/m^3 (as in (Chaboureau et al. 2002) to 1 g/m^3 (as in Lin et al. (1983)), by keeping the temperature dependency of the critical value. Due to reduced solid precipitation the low cloud cover increases for the selected case (Fig. 8).

Figure 8. Forecasted low cloud cover by AROME with the modified autoconversion for 30th November 2011 at 14 UTC. Left: low cloud cover, right: difference form the reference run.
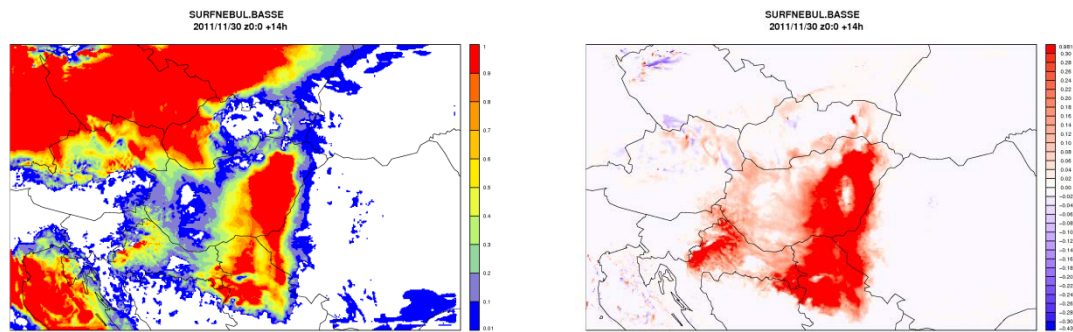
The proposed modification of the microphysics was tested on a two-weeks period in late autumn 2011. This period was characterized by a frequent occurrence of daytime stratus due to a persistent anticyclone over Hungary. For every day of the period a 24 hour forecast was run starting at 00 UTC. The configuration of this experimental chain was similar to the operational AROME-Hungary configuration in 2014, so an atmospheric 3DVAR data assimilation system was used with three hourly assimilation cycle. For the 3DVAR only conventional observations (synop stations, radiosoundings and AMDAR data) were used. The initial condition of the soil was produced by interpolating the operational soil analysis of the ALADIN model (run at 8 km horizontal resolution). Three model configurations were compared: the reference run ("ref") which has similar settings to the operational AROME-Hungary (LOSIGMAS=.FALSE., original autoconversion); the first experiment ("exp1") with modified autoconversion as described above (with LOSIGMAS=.FALSE.); and a second experiment ("exp2") (LOSIGMAS=.TRUE., modified autoconversion). Significant differences were only detected in cloud cover and 2 metre temperature (Fig. 9). By comparing "ref" and "exp1" it can be noticed that the modified autoconversion slightly increases cloud cover thus decreases the model bias and it also improves the temperature forecast. The comparison of "exp1" and "exp2" shows the impact of the saturation deficit variance parameterization. It is apparent that the prognostic parameterization (LOSIGMAS=.TRUE.) gives much less clouds and also deteriorates the temperature forecast, especially during night for the selected time period.
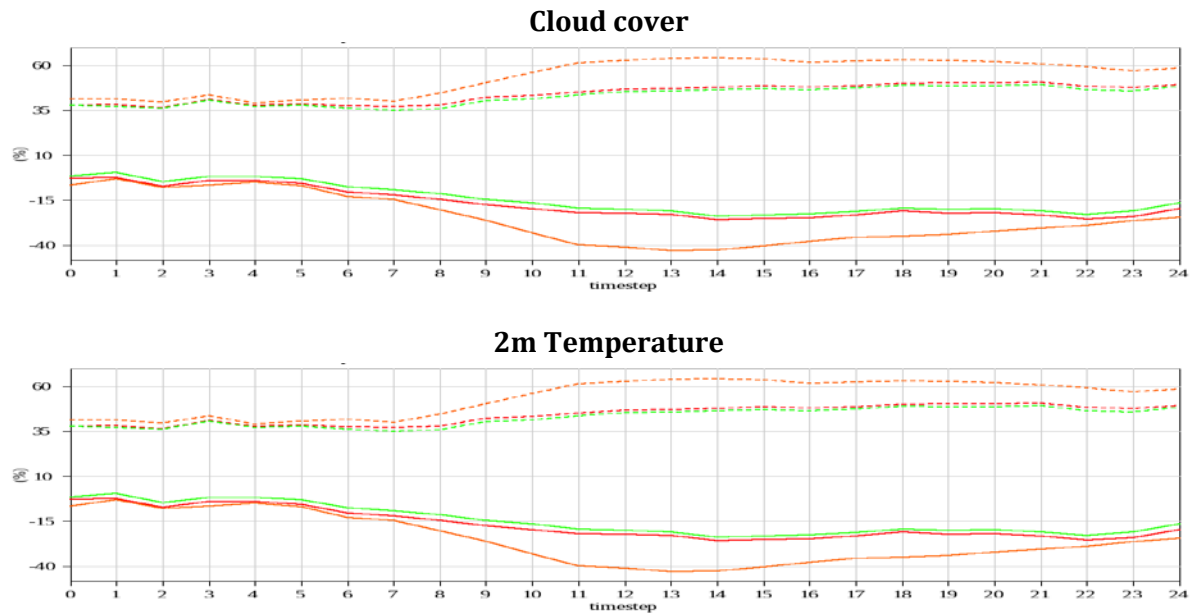
Cloud cover

2m Temperature

Figure 9. Verification scores as a function of lead time for different AROME configurations for the period between 18th November 2011 and 2nd December 2011. For the verification synop stations below 400 m were used over the AROME-Hungary domain. Upper panel: cloud cover, lower panel: 2 metre temperature. Solid lines: bias, dashed lines: RMSE. Red: reference run (LOSIGMAS=.FALSE., original autoconversion); green: first experiment (LOSIGMAS=.FALSE., modified autoconversion); orange: second experiment (LOSIGMAS=.TRUE., modified autoconversion).

Control cases

The proposed modification of the microphysics parameterization has also been tested on so-called control cases. These cases should be characterized by different synoptic conditions than the one for which the modification have been developed. As the original case was a winter anticyclonic case, for the first control case a winter cyclonic situation with heavy snowfall was chosen while for the second control case a summer convective event was selected.

On the 24th January 2014 a Mediterranean cyclone was passing over Hungary causing heavy snowfall over the south-western part of the country. According to synop measurements the depth of fallen snow exceeded 20 cm in large areas and in mountainous regions also values over 30 cm were measured. Operational models (both hydrostatic and non-hydrostatic) forecasted the event successfully, however, the high values associated to orography were only present in the non-hydrostatic models (AROME and WRF). The run with the modified microphysics ("exp1") simulated a snowfall pattern similar to the reference AROME run, with slightly lower values of snowfall (Fig. 9). No significant differences were detected for other variables, like temperature, cloud cover or wind.

Figure 10. Model forecasts and synop observations for the first control case. All models were initialized at 00 UTC on 24th January 2014 and +30h forecasts were made. Depicted is the 30h cumulated snowfall (except for the ALADIN model) both for models and synop measurements.

On the 24th June 2013 a convective event occurred over Hungary. A cold front was approaching the region from north-west and before the front in the warm sector thunderstorms developed which caused heavy precipitation (locally over 50 mm/24h). Figure 11 compares the precipitation forecasts of the reference AROME version and the one with modified microphysics.
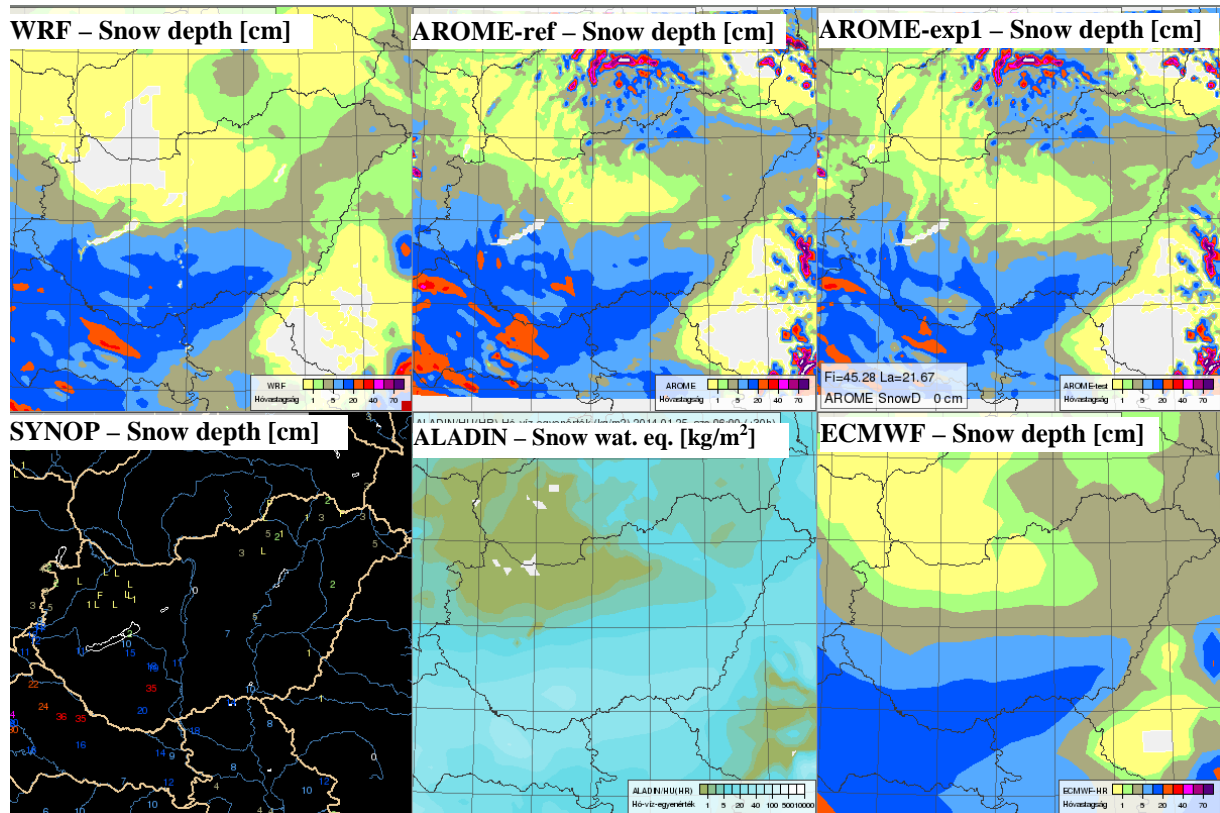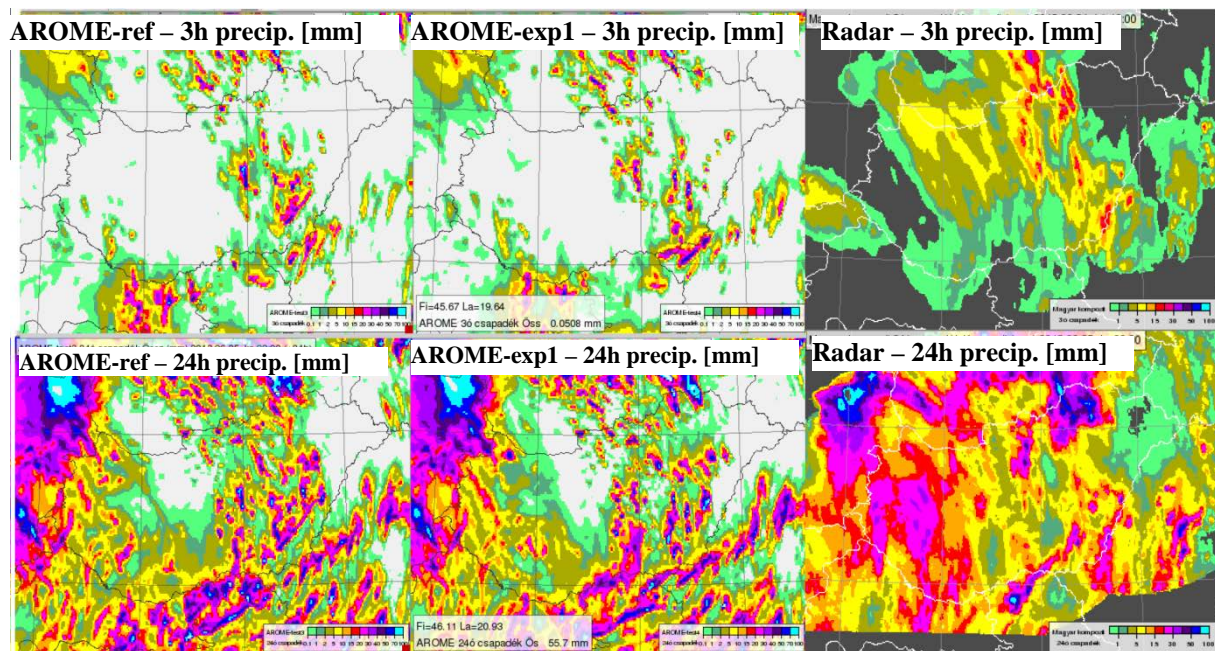
Figure 11. Model forecasts and radar observations for the second control case. Both models were initialized at 00 UTC on 24th June 2013 and +24h forecasts were made. First row: 3h accumulated precipitation for the +18h lead time; second row: 24h accumulated precipitation for the +24h lead time.

Both versions were fairly successful in forecasting the thunderstorm activity over the eastern part of Hungary, however, over the western part the precipitation amounts were underestimated by both versions. Regarding the timing of convection and the precipitation amounts no significant difference was detected between the two versions of AROME.

Conclusions and plans

In the present paper, wintertime low cloud cases were investigated which are associated with anticyclonic conditions over Central Europe. It was found that most operational numerical weather forecasting models face difficulties when simulating these situations. The most common problem is the underestimation of low cloud cover. Several sensitivity experiments were performed with the AROME model over Hungary, which indicated that the cause of the inadequate low cloud forecast can be traced back to the microphysics parameterization. By increasing the critical threshold of autoconversion, the overestimation of solid drizzle could be decreased and consequently the low cloud cover forecast improves.

In the near future it is planned to test the proposed modification on longer (summer and winter) periods. If these experiments are successful, an operational use might be possible. Finally, this kind of meteorological situation has also to be tested with a more advanced microphysical scheme based on a two moments approach (should be available in AROME in 2015). Indeed, such a scheme with prognostic number of ice particles is able to better handle snow formation.

Acknowledgement

References

Chaboureau, J.-P., J.-P. Cammas, P. J. Mascart, J.-P. Pinty, and J.-P. Lafore, 2002: Mesoscale model cloud scheme assessment using satellite observations. J. Geophys. Res., 107(D16), 4301, doi:10.1029/2001JD000714.

Hoffmann, H.-E. and Roth, R., 1989: Cloudphysical Parameters in Dependence on Height Above Cloud Base in Different Clouds. Meteorol. Atmos. Phys. 41,247-254.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of snow field in a cloud model. J. Climate Appl. Meteor., 22, 1065-1092.

The fog above sea problem in Harmonie Part II: Experiences with the RACMO turbulence scheme

Wim de Rooy

Geert Lenderink, Emiel van der Plas, Jan Barkmeijer

1.   Implementation

Implementation of the new turbulence scheme in Harmonie was a long and tough process, partly due to bugs and complexity of the code. Despite the time involved with it, these matters will not be discussed here. Other complicating factors concern the sensitivity to vertical interpolation, settings of turbulence parameters, as well as strong feedbacks between the turbulence, cloud and convection scheme. These issues will be briefly discussed here before we present some results. Note that the RACMO turbulence scheme was implemented in cycle 35t1 for the 1D ASTEX case and cycle 37h1.2 for 3D experiments.

First the sensitivity to vertical interpolation: Vertical interpolation of turbulent length scales (from hereon L) and TKE is inevitable because they are needed at full and half levels. Probably most important, the tendencies of the prognostic variables are determined by the vertical divergence of the turbulent fluxes at half levels. Therefore TKE and L are needed, and in the RACMO turbulence scheme determined, at half levels. However, in the original CBR turbulence scheme of Harmonie [1], TKE and L are calculated at full levels so for the tendencies of the prognostic variables they have to be vertically interpolated. On the other hand, the turbulent contribution to the cloud scheme variance term (see part I) for example is needed at full levels as cloud fraction and liquid and ice water are needed at full levels. Note that this issue seems subtle but due to feedbacks the impact can be substantial. For example, suppose we use the turbulent length scale at half levels and apply it at the full level below for calculating the variance contribution due to the turbulence. As a result of this, liquid water is produced slightly lower in the inversion layer and so is the radiative cooling related to it. Consequently, there is almost no top entrainment as shown in Fig. 3d for the ASTEX case (see part I for a description of ASTEX).

Secondly, in the default (edmfm) Harmonie setting, a full statistical cloud scheme is used with contributions to the variance from turbulence and convection (apart from the additional variance term). Consequently, there are strong links between the cloud, turbulence and convection scheme. A clear example is the convection cloud feedback: convection produces variance => produces liquid water => reduces water vapor => no moist updraft is diagnosed by the convection scheme => less variance etc. etc. Although, realistic from a physical point of view, such feedbacks can result in some on-off behavior as shown in Fig. 1. Nevertheless, overall the scheme finally performed smoothly. For the ASTEX case this is illustrated in Fig. 2 showing a time series of the vertically integrated TKE being as smooth as with the EDKF option and the default CBR turbulence scheme (although with generally some higher values).
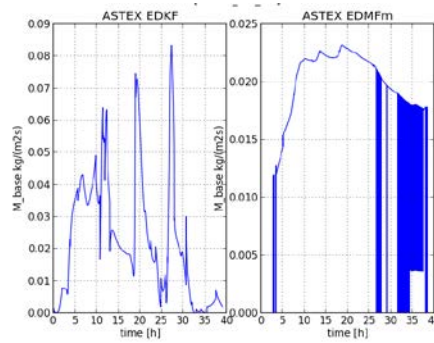
Fig. 1   The cloud base mass flux as a function of the simulation time during ASTEX. The left panel shows a version with edkf whereas the right panel shows the edmfm version with RACMO turbulence.
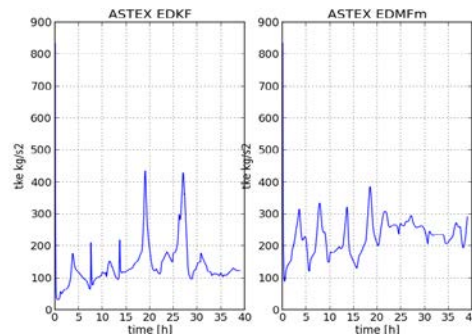
Fig. 2   As Fig 1 but with the vertically integrated TKE [kg/s^2].

2.   Results

ASTEX
In Fig. 3b the increased top entrainment with the RACMO turbulence scheme can be clearly recognized in the ASTEX case. Although the top entrainment is still underestimated and the stratocumulus does not break up as it should (see Fig. 6 in part 1 of the accompanying paper). Note that the latter not only involves the turbulence scheme but especially the cloud scheme. This is illustrated by Fig. 3c where the contribution of the variance due to turbulence is calculated with the dissipation length scale times the length scale for heat whereas in Fig 3b we have put both length scales equal for this particular purpose (like in the original code). While Fig. 3c does show the breaking up of the stratus layer it also shows less top entrainment. We will therefore proceed with the settings as in Fig 3b.

Based on the results of ASTEX we expect an increased boundary layer and cloud base height with the new turbulence scheme. Additionally, some reduction of stratus could be observed.
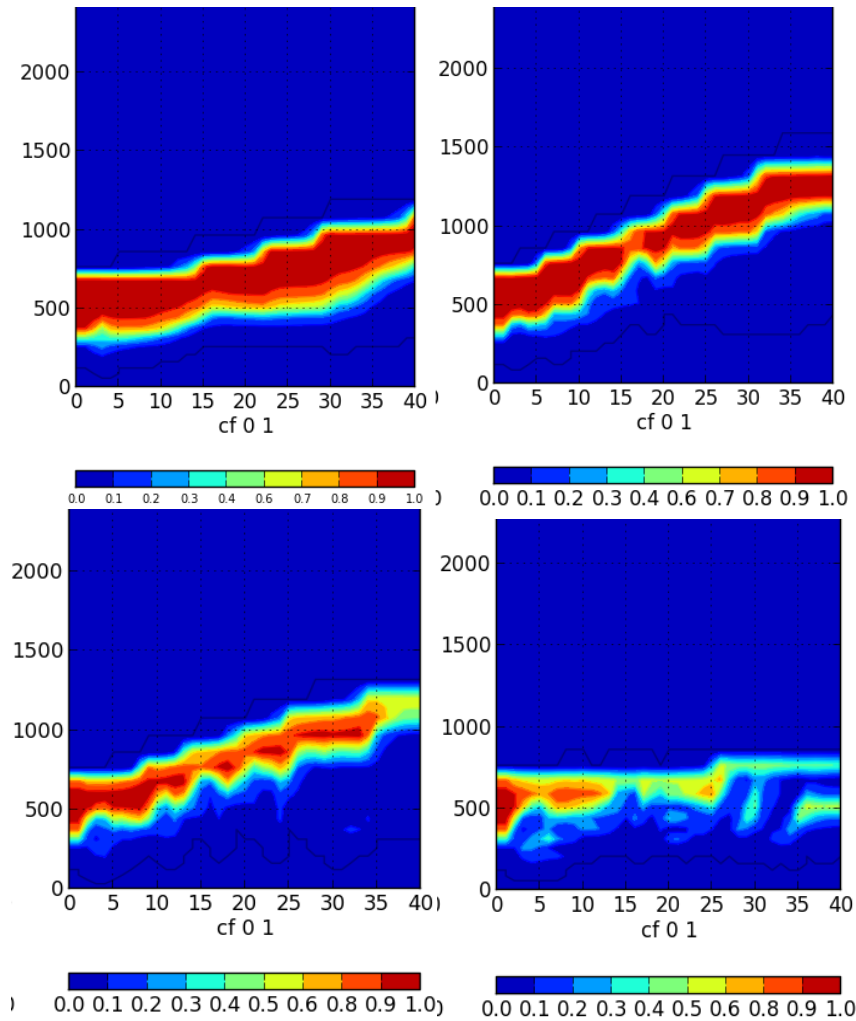
Fig. 3   *ASTEX: Contour plots of cloud fraction as a function of the simulation time for the reference edmfm option (a) upper left panel) and the edmfm version with RACMO turbulence scheme (b)upper right panel). The lower left panel (c) is the result of using a different length scale for dissipation and heat in the calculation of the contribution to the variance of the turbulence. The lower right panel (d) shows the result if the turbulent length scale is calculated at half levels and applied at the half level below for the calculation of the contribution to the variance of the turbulence.*

North Sea fog case

The obvious next step is to investigate the impact of the new turbulence scheme on the North Sea fog case as described in part I. Therefore, the RACMO turbulence scheme is implemented in Harmonie cycle 37h1.2. First we present results of a cold start, so the initial conditions are primarily determined by the ecmwf model. The racmo turbulence scheme is run with two options, namely *zch=0.15* (default in RACMO) and *zch=0.2* (as in the original paper [2]). Parameter *zch* determines the mixing in stable conditions and influences the top entrainment in the inversion. The results in Fig. 4 reveal that after 24h the already existing fog field in the northern part has become less dense and holes appear in it, especially with *zch=0.2*. The fog field in the southern part which develops during the run is virtually absent in the version with *zch=0.2*. The run with *zch=0.15* seems to behave in between the standard and the *zch=0.2* run. The development of the fog in the southern North Sea is clearly delayed

when looking at all individual forecast hours (see [3]). Interesting is that with *zch=0.15* the dense fog in the southern part visibly develops first above land (see Fig. 5a) and is subsequently advected above the cold sea where it "explodes". On the other hand the related fog above land is quickly dissolved due to the increased mixing during day time (Fig. 5b). This behavior was also shown by Fig. 5 of Part I of this paper, a run with strongly reduced variance. We will return to this issue later on.
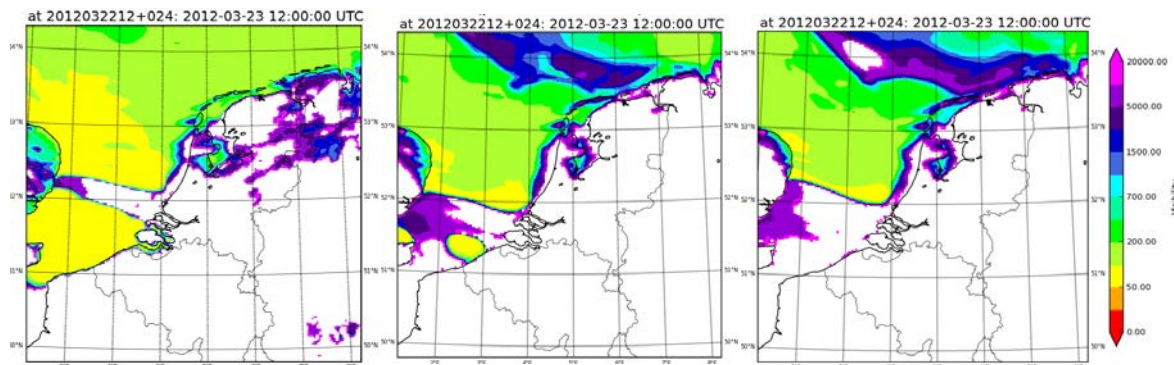
Fig. 4   Harmonie visibility forecasts (+24h) [m] with the default (a) left panel), with the RACMO turbulence scheme and zch=0.15 (b) middle panel) and as b) but with zch=0.2 (c) right panel).
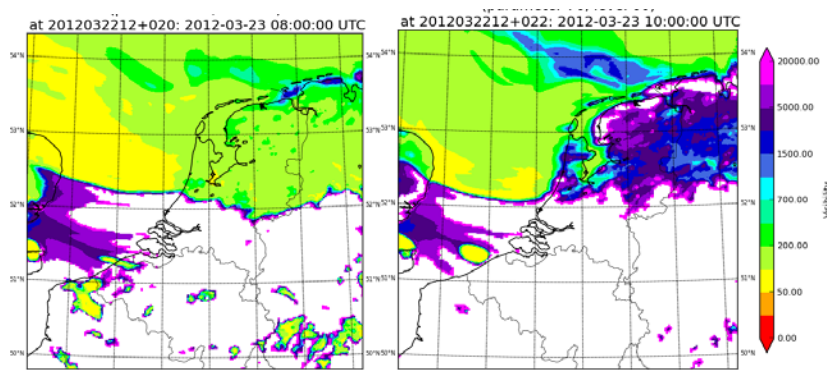
Fig. 5   Harmonie visibility [m] forecasts with the RACMO turbulence scheme and zch=0.15 showing the +20h (a) left panel) and +22h forecast (b) right panel).

The new turbulence scheme is also applied in an assimilation cycle started at 15 March 2012 to allow for spinning up of the model. In this way the model climatology and first guess are influenced by the modified turbulence. Potentially, this could result in a higher impact for the North Sea case. For example, in the assimilation cycle with the new turbulence scheme, the northern fog field might never be developed. However, in practice the difference between a warm and cold start turned out to be quite small and insignificant.

So two fog fields can be distinguished: An already existing northern fog field and a developing southern fog field. There are strong indications that these fog fields have different causes as described below.

Southern fog field.
The southern fog field is primarily caused by too moist air above land which is advected above the relatively cold sea water where it gives rise to fog. Due to the stable conditions above sea this fog is difficult to dissolve. The too moist air above land is caused by insufficient top entrainment by the turbulence scheme as revealed by the ASTEX case. Insufficient top entrainment leads to too little ventilation of the boundary layer with relatively dry and warm air from above.

There are several indications for the aforementioned hypothesis concerning the main cause of the southern erroneous fog field. Firstly, Yann Seity performed a North Sea experiment with AROME where the soil moisture was initialized according to ARPEGE (see [3]). As a result the soil moisture and therewith the surface latent heat flux was significantly reduced and consequently the southern fog field disappeared. Note that the soil moisture itself is not the cause as in The Netherlands it is extremely rare that soil moisture is a limiting factor for the latent heat flux. However, this experiment does support the hypothesis that too moist air above land causes the erroneous fog field.

Secondly, as first noted by Sander Tijm, Harmonie sometimes produces fog fields above land that "explode" and become very persistent when being advected above a cold sea (stable conditions). Two examples are shown. In the first example (Fig. 5 of Part I) the reduced variance prevents the forming of liquid water and therewith reduced visibility in the first half of the run although too moist air was already advected from the land. In the second example (Fig. 5) the delayed development of sea fog is caused by the air coming from land being less moist due to somewhat (though probably not enough) increased top entrainment. However, in both experiments the too moist air above land finally becomes visible as dense fog being advected above sea where it spreads out and could not be dissolved due to persistent stable conditions.

Finally, the hypothesis can be investigated more directly by making vertical slices of relative humidity (RH) through the developing southern fog field along the line presented in Fig. 6.
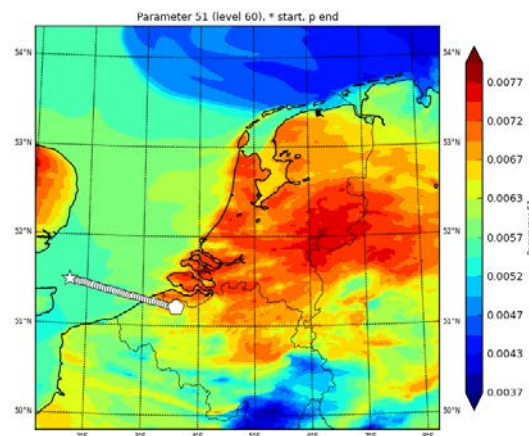
Fig. 6.   Line along which the vertical slice of the RH is taken. Starting at the star above sea (left hand side of the vertical slice plot) and ending at the pentagon above land. Note that the flow is from the land (right) to the left (sea). Plotted is the specific humidity [kg/kg] at the lowest model level.

In Fig. 7 we present some snapshots during the forecast run of the vertical slices and the corresponding fog field. For a movie of the complete forecast we refer to [3]. As clearly visible in Fig. 7, the RH values are lower in the runs with RACMO turbulence. When high RH values (> 0.9) develop and are advected above sea, the visibility starts to decrease (see left panels in the second and third row of Fig. 7). At sun rise the RH decreases due to enhanced mixing and top entrainment of the boundary layer. This process starts earlier (see Fig. 7 +21h forecast) and is more intense in the run with RACMO turbulence. Consequently, the high visibility levels in the run with modified turbulence change to completely clear when this relatively dry air is advected above sea in the following forecast hours (see Fig. 4c after 24h).

+0h

+11h

+12h

+21h
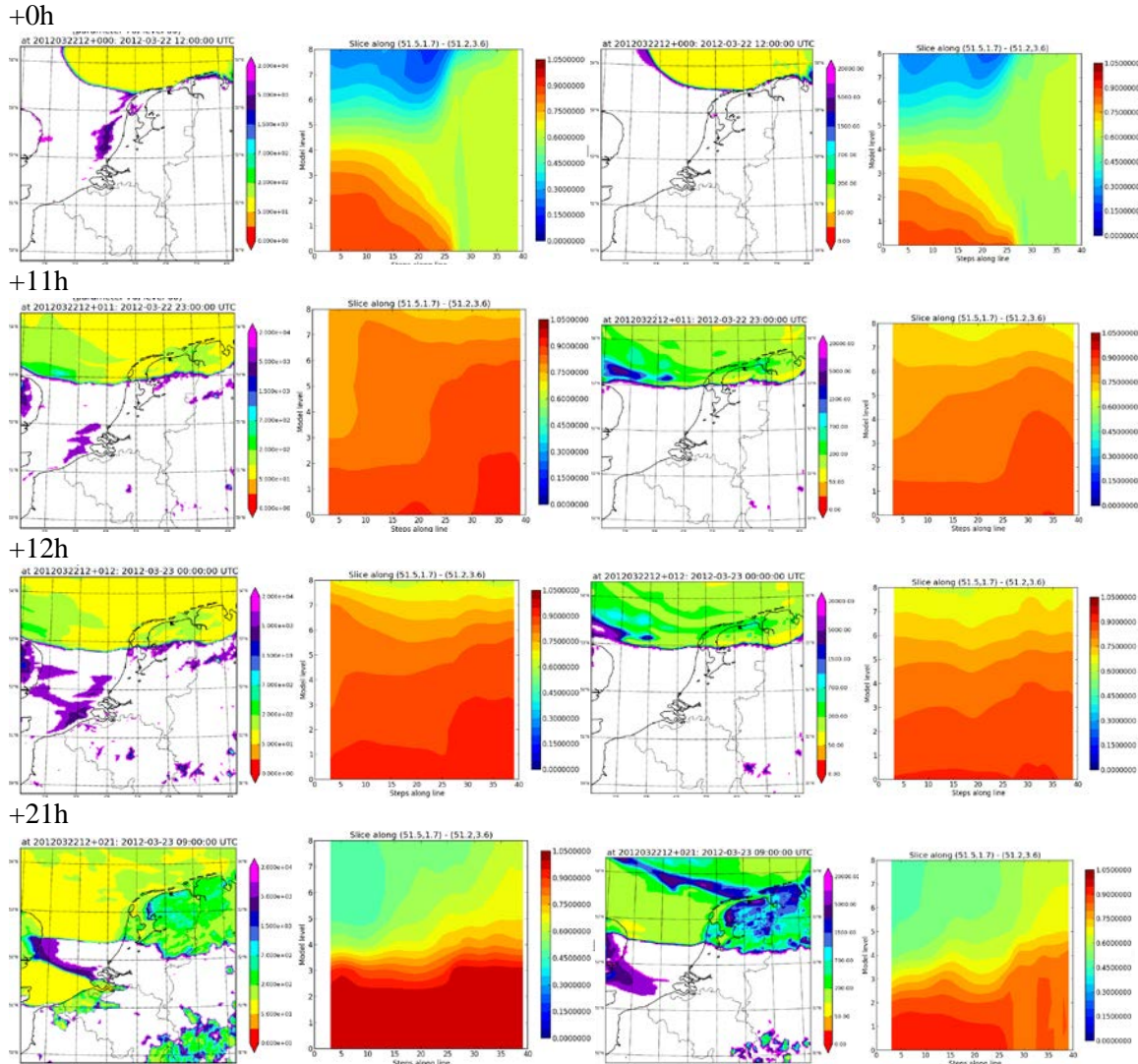
*Fig. 7   Visibility plots [m] and vertical slices (lowest 8 model levels) of RH [0-1]along the line
           shown in Fig. 6. The left side of the vertical slice plots is located above sea and the right
           above land. The two columns on the left refer to standard runs and the two right columns
           refer to runs including the RACMO turbulence scheme and zch=0.2. The rows correspond
           with the forecast period as indicated.*

Northern fog field.
Our hypothesis is that the northern fog field is primarily caused by deficiencies in the initialization
and boundaries. This hypothesis is supported by:

- Runs of AROME nested in ARPEGE as provided by Yann Seity. The relevant physics of
  these runs are the same as in our Harmonie runs and consequently the southern fog field also
  appear in these runs. However, the initialization and boundaries are different because
  AROME is nested in ARPEGE. Consequently, the northern fog field is completely absent in
  the AROME runs.

- The experiment by Sibbo van der Veen. In this experiment the specific humidity field is modified based on a satellite cloud mask and synop observations of cloud base height. In Fig. 8 we see that the southern fog field, related to deficiencies in the physics, is still present. However, with the modified initialization the northern fog field disappeared.
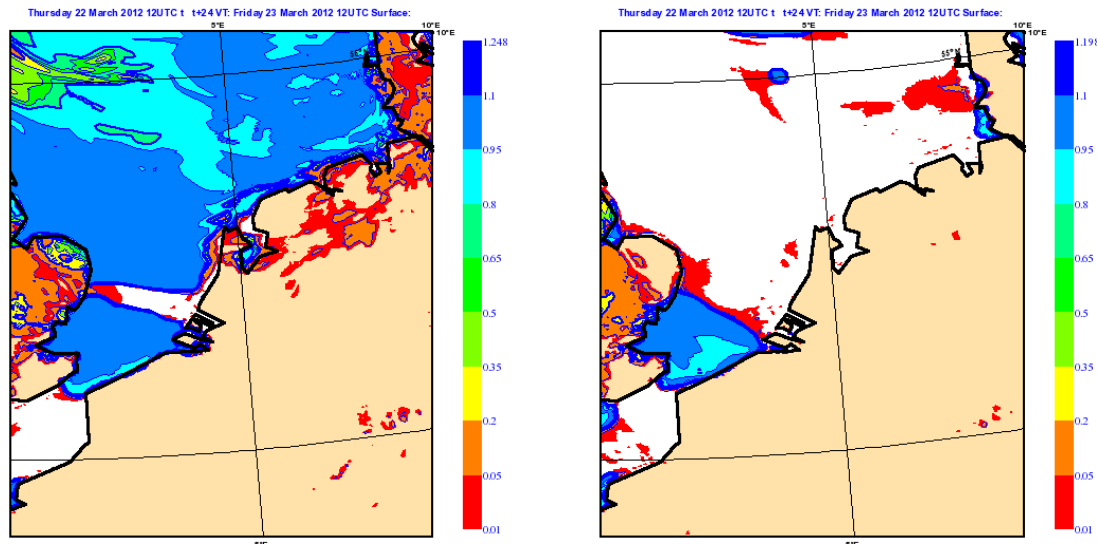
Fig. 8   The cloud cover at the lowest model level for the North Sea fog case after +24h. On the LHS the standard run and on the RHS the result of an experimental run with humidity initialization based on satellite information. Although cloud cover is shown instead of visibility, the northern and southern fog fields can clearly be recognized.

Are the good fog forecasts retained?

With the modified turbulence scheme we prevent some of the erroneous fog forecasts but can we also retain the good fog forecasts? To investigate this we have rerun a case where the fog, especially above land, was captured very well including the gradual retreat towards the north western part of the Netherlands. The run starts at 22 October 2012 00UTC and we focus on the +8h forecast. Fig. 9 reveals that the sea fog north east of the Netherlands is reduced too much but grosso modo the good forecast is retained with the modified turbulence scheme. Obviously, one case is not enough to draw firm conclusions but our experience with a longer term experimental suite seem to confirm the reproduction of accurate fog forecasts with the RACMO turbulence scheme. Nevertheless, longer term systematic verification is necessary.
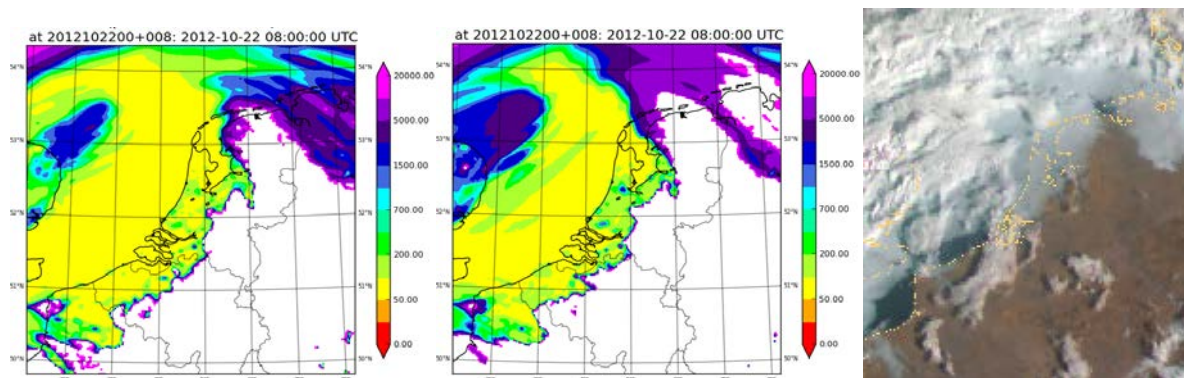
Fig 9.   Example of a good fog forecast, started at 22 October 2012. Here presented is the +8h forecast. The left panel is the standard run, the middle panel the run with RACMO turbulence and the right panel is the corresponding visible satellite image showing a combination of fog and higher level clouds.

Objective verification.

From mid February 2014 on we started a Harmonie suite version 37h1.2 including RACMO turbulence scheme. This suite together with experimental daily runs with 3DVAR, MSG initialization, RUC (Rapid Update Cycle) could be compared with the operational run cycle 36h1.4. Our subjective experiences with the RACMO run were generally positive with less stratus and less erroneous sea fog (as expected) as well as more smooth rain fields. Moreover, the accurate fog forecasts seem to be retained. On the other hand we sometimes observed more fog in the RACMO and other parallel runs that we could not explain. The output of all runs were processed in the standard verification package of Xiaohua Yang. Thanks to this verification we gradually discovered a large discrepancy between the operational version 36h1.4 run and the experimental 37h1.2 runs, with colder and more humid boundary layers in version 37h1.2. This can e.g. be illustrated by the increased positive bias in q_{2m} (see Fig.10). It became evidently to start a standard Harmonie cycle 37h1.2 as a controlrun (green lines in Fig. 10) to see the impact of the modifications only. Unfortunately, Fig. 10 covers only a short verification period and moreover some of the parallel runs (including the RACMO run) experienced a cold start in this period. Nevertheless, it can be concluded that the impact of the modifications in the parallel runs on the near-surface verification are dominated by the transition from cycle 36h1.4 to 37h1.2. With this transition, probably dominated by the modifications in SURFEX, the bowen ratio seems to change substantially, with more latent and less sensible heat flux. In this way the change in the turbulence scheme is counteracted strongly by the surface scheme. As a result we have decided to continue our experiments with cycle 38.

Note that apart from the comparison against the operational cycle 36h1.4 in this section, all other 3D results are based on version 37h1.2, showing the impact of the turbulence scheme only.
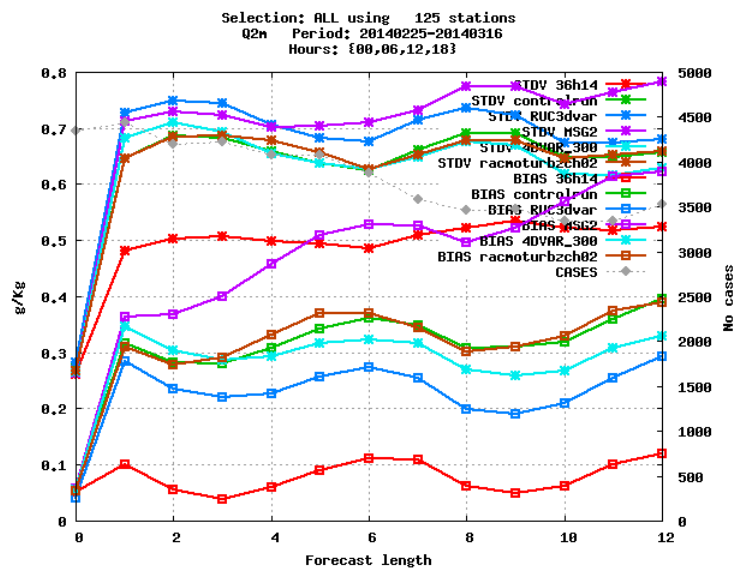
Fig. 10   Verification of q$_{2m}$ [g/kg] of different Harmonie runs against EWGLAM stations. Shown are the bias and standard deviation as a function of the forecast period for the period from 20140225 to 20140316. Most important is the difference between the operational cycle 36h1.4 run (red) and the control run cycle 37h1.2 (green). Note that this period includes a cold start for the parallel runs.

There is however one effect of the turbulence scheme that might not be dominated by the version number, namely the effect on the wind speed bias. With the RACMO turbulence scheme the negative bias in the 10m wind speed is increased roughly from -0.1 m/s to -0.3 m/s. Further investigation is needed to establish if this is the result of a real deterioration or simply a representation mismatch between grid points and observation sites. Verification of wind speeds above sea and against radiosonds reveal an improvement with the modified turbulence scheme indicating the influence of a representation mismatch.

3.   Conclusions and outlook

The present study increased our understanding of processes related to the production of erroneous sea fog. Two processes seem to play a dominant role. First of all a deficiency in the turbulence scheme. From the ASTEX intercomparison case we know that the current turbulence scheme does not provide enough top entrainment. As a consequence the boundary layer and cloud base height are too low, problems we also experience in every day practice. Moreover, the boundary layer will be too moist due to insufficient ventilation with relatively warm and dry air from above. This effect is also responsible for some of the erroneous fog fields, e.g. the southern fog field in the North Sea fog case. When the too moist boundary layer air is advected above relatively cold sea water, as in late winter and spring, it can give rise to not observed fog fields that are difficult to dissolve because of the stable conditions above sea. With increased top entrainment as in the RACMO turbulence scheme the drier boundary layer air can prevent the development of the erroneous southern fog field. Based on results for ASTEX the modified turbulence scheme is expected to produce higher cloud base and boundary layer heights and probably less stratus, all addressing known Harmonie problems. Here we want to emphasize the importance of intercomparison studies like ASTEX

because they provide a controlled environment in which we are able to really pinpoint an observed deficiency to a particular process (here turbulence). This is possible because some processes are simply circumvented (e.g. through prescribed SST and initial conditions) as well as the availability of many dedicated observations to check the individual components (e.g. surface fluxes).

On the other hand not all erroneous sea fog was solved by modifying the turbulence scheme. For some cases, e.g. the northern fog field in the North Sea case, the initialization and or boundaries seem to be the main cause. So at the same time we should put effort in an improved (humidity) initialization, using e.g. radar, GPS or satellite cloud images.

Replacing a turbulence scheme will have strong impact on many aspects of the model. Consequently, a lot of systematic verification is necessary. Although we have run Harmonie including RACMO turbulence on a daily basis for several weeks, the objective near-surface verification seems to be dominated by the transition from cycle 36 to 37 i.o. modifying the turbulence scheme. Therefore, we will restart the parallel RACMO run in cycle 38 and the results will be processed in the standard verification package as well as in Cloudnet [4]. In addition, we will perform a 1D reforecasting within the KNMI parameterization testbed. In this reforecasting several Harmonie versions, including one with RACMO turbulence will be run. Subsequently, the results will be compared with advanced observations in Cabauw as well as with output from LES runs.

The way the RACMO code is currently incorporated in Harmonie still needs a proper recoding before operational implementation. In a separate routine (vdfexcuhl.F90) the turbulent length scale and TKE are determined and used as input for the default turbulence code. For example the tendencies of the prognostic variables and the variance for the cloud scheme are calculated in the original code. We expect the new turbulence code to be faster as the new length scale does not need an iteration procedure anymore. The code is distributed to some participants of the cloud fog task force (Sweden, Spain and France). Experimenting with the code is encouraged and all feedback is welcome.

References

[1] Cuxart J., P. Bougeault, J.-L. Redelsperger, 2000: A turbulence scheme allowing for mesoscale and large-eddy simulations. *Q. J. R. Meteor. Soc*., **126**, 1-30

[2] Lenderink, G. and A.A.M. Holtslag, 2004: An updated length scale formulation for turbulent mixing in clear and cloudy boundary layers. *Q. J. R. Meteor. Soc*., **130**, 3405-3427

[3] https://hirlam.org/trac/wiki/HarmonieWorkingWeek/Clouds201210 under "Preliminary results North Sea fog case"

[4] http://www.cloud-net.org/radar/cloudnet/index.html

Running HARMONIE on Xeon Phi Coprocessors

Enda O'Brien

1 Introduction

For HARMONIE users looking for that extra boost in speed, the expanding collection of hardware "accelerators" offers tempting prospects of large performance gains to any application that can exploit their characteristically massive parallelism. Such accelerators include both general-purpose graphical processing units (GP-GPUs) and "Many Integrated Core" (MIC) processors (including Intel's Xeon Phi coprocessors). The catch, especially in the case of GPUs, is that applications must also be adaptable to a more restrictive programming environment.

In the case of Xeon Phi, however, it is possible to avoid those restrictions, since Xeon Phi is based on a more conventional X86 architecture, which means that applications can be compiled to run entirely (or "natively") on Xeon Phis, something that is not possible with GPUs. Xeon Phi coprocessors may not possess the raw computational power of the most powerful GPUs (as measured by peak FLOPs), but because they are more "programmable", their performance potential is easier to harness.

Some questions that naturally arise when porting HARMONIE to accelerators are:

1. How hard is it? In particular, what source-code changes, if any, are needed, and how portable, maintainable, or generally disruptive will these be?

2. Should porting efforts focus on GPUs or on Xeon Phi coprocessors?

3. Last but not least, how does HARMONIE perform with accelerators?

Ultimately, it only makes sense to use hardware accelerators (such as Xeon Phi) to run applications like HARMONIE if the price/performance benefits make it more attractive to buy accelerators to attach to (standard) host nodes rather than simply buy more nodes.

This article reports on how HARMONIE currently runs on Xeon Phi coprocessors, in particular the Xeon Phi 5110P, as installed with two such MICs per node on a 16-node cluster at the Irish Centre for High-End Computing (ICHEC).

2 HARMONIE and Accelerators

Broadly speaking there are three options for porting HARMONIE to accelerators like GPUs or MICs:

1. For GPUs, at least, one option is to translate HARMONIE to C, and thence to CUDA (or OpenCL), which are the most developed and most powerful programming interfaces to GPUs.

2. A second option, equally available for both GPUs and MICs, is to insert accelerator directives into the source, similar to the OpenMP directives already there. Such directives identify variable and code sections that may be offloaded to accelerators at run-time.

3. A third option, currently only available for MICs, is to leave HARMONIE source essentially unchanged and simply compile the code for deployment directly on the MICs.

The first option above (translating to C/CUDA) is not really viable in practise, since HARMONIE is written primarily in Fortran, and is under constant development with a frequent update cycle. Taking this option would change HARMONIE into non-standard software, making it very difficult to port and maintain. Moreover it would be very difficult to implement in any case.

The second option (insertion of compiler offload directives in the source), is more attractive since HARMONIE already contains OpenMP directives — very efficiently implemented, too. In practise, OpenACC or MIC directives are relatively easy to implement for many low-level "kernel"-type routines in HARMONIE, and preliminary tests with GPUs even show reasonably good speedups of selected routines. To use directives effectively, however, they need to be inserted at high-levels in the source call-tree (as was done for OpenMP directives).

In practise, many features in HARMONIE run foul of the coding restrictions that must be observed for OpenACC or MIC "offload" directives to work at all. In particular, variables "of derived type with pointer or allocatable components" cannot be used in offload regions. This poses major problems for HARMONIE, since it contains many variables like this throughout the entire code. On the other hand, this is quite a fundamental restriction, given the separate memory systems on each host and accelerator device. Furthermore, any "offload" acceleration strategy would still suffer from the universal accelerator problem of having to transfer data between host and device at regular intervals (usually every time-step).

The third option above (compiling HARMONIE to run natively on Xeon Phi) is not possible with GPUs, but is the natural place to start with MICs since it requires little or no source-code changes and avoids the need for periodic data transfers between host and accelerator. To use Intel terminology, using compiler offload directives corresponds to running on Xeon Phi in "offload" mode, while compiling to run directly on Phi corresponds to "native" mode. Another variant of native mode is "symmetric" mode, where each host and each MIC are essentially separate nodes in a heterogenous cluster, and parallel applications run with "native" executables on each such node.

To summarize, host applications can use accelerators in one of three different modes:

1. "Offload mode" (using language extensions like CUDA or compiler directives);

2. "Native mode" (where the application runs entirely on the device), or

3. "Symmetric mode" (where separate instances of the application run natively on both hosts and devices in a heterogeneous cluster).

Offload mode is available on both GPUS and Xeon Phis, but requires a certain amount of source-code modification (which is hard, at least in the case of HARMONIE), and a certain amount of data transfer between host and device (which is generally slow). On the other hand, native and symmetric modes require no source code changes, and no data extra transfers beyond the usual inter-node MPI messages that are necessary on any cluster. Both native and symmetric modes are available on Xeon Phis but not on GPUs, which makes Xeon Phis particularly attractive as a platform for accelerating HARMONIE, even if the peak raw performance of Xeon Phi may be less than that of GPUs. Consequently, the easiest way to try to accelerate HARMONIE (at least initially), is using native and symmetric mode on Xeon Phi.

3   Xeon Phi Coprocessor Characterization Tests and Kernel Benchmarks

Before attempting to run an "industrial-strength" application like HARMONIE on a new platform like Xeon Phi, it is worthwhile to first run some smaller "kernel" tests to find out how the system works, and to

Table 1: Test cluster details, and kernel benchmark performance results

	E5-2660 2.2 GHz	Xeon Phi 5110P
Cores (per node)	20	61
Threads (per node)	40	240
Clock Frequency	2.2 GHz	1.05 GHz
Memory	64 GB/node	8 GB ($\times$ 2 MICs) =16 GB
Max. Stream Triad	91 GB/s	137 GB/s
Linpack	316 Gflop/s (max)	720 Gflop/s
	288 Gflop/s (16 GB)	
IMB PingPong 0 Byte latency	$< 2\mu$sec	$5 - 12\mu$sec
IMN PingPong 4 MB bandwidth	> 4 GB/s	$0.22 - 4$ GB/s

characterize the performance and capability of the system. On a standard linux cluster, four numbers that provide concise characterisation of overall performance are:

- Stream (e.g., the Triad component), as a measure of memory bandwidth (in MB/s);

- Linpack, as a measure of raw CPU compute power (in GFlops/s);

- Small-message latency from the Intel MPI Benchmark (IMB) "PingPong" test, as a measure of inter-node network speed (in μsec);

- Large-message bandwidth, also from the IMB PingPong test, as a measure of inter-node network capacity (in GB/s).

The tests described here were run on a cluster of Ivy Bridge (Xeon E5-2660 2.2 GHz) dual-processor (20-core) nodes, each with 2 Xeon Phi 5110P coprocessors attached, connected with a "fourteen data-rate" (FDR) Infiniband network. The Stream and Linpack tests were run separately on a single host node and a single coprocessor, while the IMB tests were run between different *pairs* of processes on hosts and coprocessors, in various permutations across the cluster. Table 1 shows the vital statistics of the cluster, along with representative results of the kernel benchmark tests.

Despite its relatively slow clock-speed, Xeon Phi can exceed the Stream-Triad and Linpack performance of a 20-core host node by relying on its greater parallelism. Note that the Stream and Linpack performance shown in Table 1 for the Xeon Phi is contingent on using *all* available cores or threads.

MPI PingPong performance depends heavily on whether communication is between two MICs on the same node, two MICs on different nodes, or between a MIC and a host (on the same or different nodes). Results for all possible permutations are shown in Tables 2 and 3, for 0 B latency and 4 MB bandwidth, respectively. As a rule of thumb, host-to-host communication has low latency and high bandwidth, while any communication involving a MIC has higher latency by at least a factor of two, and lower bandwidths — up to an order of magnitude lower! Note too that the two coprocessors attached to each host are not equivalent: one is "closer" to the host and the other more "remote", as measured by PingPong latencies and bandwidths.

Low latency for inter-node communications is one of the main attributes of a network that defines it as "high-performance". It is one of the principal advantages that Infiniband has over Ethernet. As shown by Table 2, however, latency between MIC coprocessors on different nodes (or even between the different MICs on the same node) is approx. 5 times higher than between two different hosts. Similarly, bandwidth between two different MICs (esp. attached to different nodes) is severely limited compared with inter-host bandwidth over FDR Infiniband. These results amount to something of a warning flag for the Xeon Phi, and suggest that

Table 2: IMB PingPong 0 byte message latency (μsec) between "senders" on the leftmost column and "receivers" on the top row.

	host0	host0-mic0	host0-mic1	host1	host1-mic0	host1-mic1
host0	0.36	5.24	6.40	1.96	6.43	7.05
host0-mic0	5.24	2.28	9.08	6.43	8.96	9.71
host0-mic1	6.40	9.08	2.37	7.05	9.71	10.99

Table 3: IMB PingPong 4 MB message bandwidth (MB/s), between "senders" on the leftmost column and "receivers" on the top row.

	host0	host0-mic0	host0-mic1	host1	host1-mic0	host1-mic1
host0	4067	4923	5193	5870	4156	505
host0-mic0	4923	2020	1269	4156	3539	494
host0-mic1	5193	1269	1951	505	494	266

MPI communications between (or even within) MICs could turn out to be a major performance bottleneck for larger applications like HARMONIE.

A very simple Fortran "stencil-test" program was written to complement the more standard tests mentioned above, and to be roughly representative of the LAITRI routine in HARMONIE. Results from this test were shown in last year's ALADIN-HIRLAM newsletter no. 1 [1].

The main lessons to be drawn from all these simple kernel tests is that OpenMP multi-threading scalability is impressively good on the Xeon phi coprocessor once the problem size is large enough. In that case, approx. 60 or more threads are needed to get worthwhile performance, i.e., performance better than what is available on the host. Evidence from both the Linpack and the stencil benchmarks suggests that the best performance possible from a Xeon Phi coprocessor is a factor of 3 or so better than that of a host node - though clearly that factor could vary a lot depending on the nature of the host.

As a corollary, Xeon Phi coprocessor performance is likely to be poor if the problem cannot fit snugly into the 8 GB memory available, if it cannot scale to 60 or more threads, if it depends on data transfers between host and coprossor - or even if it uses MPI message-passing between coprocessors and hosts on different nodes.

With results from these simple kernel benchmark tests as a guide, the following sections report on how HARMONIE runs on the Xeon Phi.

4   HARMONIE Performance on Xeon Phi Coprocessors

4.1   Building HARMONIE and Utility Libraries for Xeon Phi

In principle, any application can be built to run on Xeon Phi coprocessors by simply adding "$-$âĂŞmmic" to the list of compiler options. This is essentially all that had to be done to compile HARMONIE source. The only minor change required was to remove the SEQUENCE keyword from the derived-type declarations in `xrd/module/eggpack.F90` and in `arp/module/yomdb.F90`. When compiling for âĂIJstandardâĂİ host platforms (i.e., without the $-$âĂŞmmic option), the SEQUENCE keyword (as used in those routines) provoked a warning, but when compiled with âĂŞ-mmic, it generated a full error message instead (at least,

with Intel composer_xe_2013_sp1.2.144). Removing the SEQUENCE keyword allowed the compiler to "pad" components of the derived variable to conserve alignment.

The utility libraries needed by HARMONIE at link-time (including zlib, hdf5, netcdf, and grib-api), must also be built with −âĂŞmmic. Typically, these use the standard linux package sequence of "configure", "make", and "make install". If −âĂŞmmic is included in the CFLAGS or FCFLAGS macros, this sequence invariably fails at "configure" stage as the script attempts to build and run small test executables. Those executables were built to run natively on Xeon Phi, but are launched on the standard host (where the compilers are installed). A normal cross-compile doesn't work in these cases because MIC is not a recognized architecture.

One workaround is to edit the "configure" scripts to skip running all test executables. Another is to leave out any mention of −âĂŞmmic during the configure phase, run configure as on a standard host, then edit the resulting config.status and/or Makefiles to insert −âĂŞmmic where needed in CFLAGS or FCFLAGS before running "make" and "make install". Useful online resources for solving these issues include: `http://linux4hippos.blogspot.ie/2013/05/cross-compile-hdf5-for-intel-xeon-phi.html`. At least, building utility libraries for Xeon Phi compatibility is a job that need only be done once.

4.2   Configuring HARMONIE to run on Xeon Phi

The main MASTERODB executable, built with −âĂŞmmic to run on the Xeon Phi, can be renamed to MASTERODB.MIC and placed alongside the standard MASTERODB in the otherwise standard HARMONIE installation. The HARMONIE `Forecast` (or `config_exp.h`) script is then modified to set the environment variables and issue the appropriate command to launch MASTERODB.MIC "natively" on a logical "cluster" consisting exclusively of Xeon Phi coprocessors.

The HARMONIE configuration chosen for most experiments on Xeon Phi was the IRELAND55 case, which consists of a domain spanning Ireland and the UK with a grid resolution of 5.5km. The full domain has 300×300 horizontal grid-points and 65 vertical levels. When run as a sequential, single-threaded process, it uses approx. 20 GB memory. As more MPI processes and more OpenMP threads are used to run this case, memory consumption increases quickly âĂŞ to the point of almost saturating the 96 GB total memory available when 12 MPI processes over 12 MIC cards are used with 64 threads per process.

While the IRELAND55 case is not large by modern standards, it is still more than big enough for current purposes, since in practise it needs at least 4 separate 5110P Xeon Phi coprocessors (each with 8 GB memory) in order to run at all. All results shown below are run-times for a 6-hr forecast.

4.3   MPI processes, or OpenMP threads?

HARMONIE is well parallelised using both MPI and OpenMP, so on shared-memory systems there is a lot of flexibility as to how total parallelism can be partitioned between MPI processes and OpenMP threads. Table 4 shows how HARMONIE performance (i.e., run-time for a 6-hr forecast of the IRELAND55 configuration) varies when run on a single Ivy Bridge E5-2660 v2 node with 20 physical cores (or 40 logical, "hyperthreadedâĂŹâĂŹ cores) and 64 GB memory, as *total* thread-count remains constant while the MPI/OpenMP "mixâĂŹâĂŹ is varied.

According to Table 4, the most efficient way to run HARMONIE on this host node is to use 40 MPI processes on the 40 logical cores, with just 1 OpenMP thread per process. In practise, HARMONIE is more commonly run on nodes like this to use as many MPI processes as there are physical cores (i.e., 20 in this case), with two OpenMP threads per task to soak up the "hyperthreads". Evidently, intra-node shared-memory message passing is very efficient while the overhead of starting and closing multiple OpenMP threads is relatively expensive and inefficient.

Table 4: E5-2660 v2 host node performance sensitivity to MPI/OpenMP mix, for constant total thread-count, (IRELAND55 case).

MPI Processes	Without HyperThreads			Using HyperThreads		
	OMP_NUM _THREADS	Total Threads	Forecast Time (s)	OMP_NUM _THREADS	Total Threads	Forecast Time (s)
2	10	20	1570	20	40	940
5	4	20	1445	8	40	814
10	2	20	1384	4	40	727
20	1	20	769	2	40	687
40				1	40	668

Table 5: Xeon Phi coprocessor performance sensitivity to MPI/OpenMP mix, for constant total thread-count (IRELAND55 case).

		Total Threads	Run-time (s)
MPI Only	12 MICs, 16 MPI tasks/MIC	192	3779
MPI/OpenMP	12 MICs, 16 OpenMP threads/task	192	1448
MPI/OpenMP	12 MICs, 50 OpenMP threads/task	600	931

This experiment cannot be directly replicated on a single Xeon Phi coprocessor because it doesn't have enough memory. However, Table 5 shows roughly analogous results from some HARMONIE runs across 12 MIC cards (i.e., 6 nodes). Using 16 OpenMP threads per MIC is more than twice as fast as using 16 MPI processes per MIC, even though a total of 192 threads are used in both cases. OpenMP scalability continues to 50 threads per MIC (at least), as shown in the last row of Table 5.

The moral is clear that when running on Xeon Phi coprocessors, it is better to use as many OpenMP threads as possible along with as few MPI processes as possible.

4.4   HARMONIE "Native-Mode" Performance on Xeon Phi

The overall run-time performance of the IRELAND55 configuration, run natively on Xeon Phi, is summarized in Fig. 1. Results are shown using 8 or 12 MICs (i.e., 4 or 6 nodes), as a function of OpenMP threads per MIC along the abscissa. Also shown, as the coloured horizontal lines near the bottom, are host-only performance using either 1, 2 or 4 E5-2660 v2 20-core nodes. Wall-time for the 4-node (host-only) run was just 253s, which may be viewed as a reference performance target for the "completeâĂŹâĂŹ 4-node system — including the 8 Xeon Phi coprocessors.

It is clear from Fig. 1, however, that Xeon Phi native performance is still a long way from 253s. The best result achieved using 4 nodes (8 MICs) was 948s, using 64 threads/MIC (as represented by a green triangle at that point). By that metric, 8 MICs are approx. 4 times slower than 4 host nodes.

The "MxN" labels in the legend in Fig. 1 refer to the HARMONIE horizontal domain decomposition; i.e., M processes in the east-west direction, with N processes in the north-south direction. Normally, when running on a cluster of standard host nodes, the optimal domain decomposition consists of subdomains that are approximately "square", or with M slightly smaller than N. Such a decomposition minimizes the total length of the interface or "shadow zone" between neighbouring processes, and thus minimizes MPI communication between them. When running over 12 MICs then, it was natural to choose a 3×4 domain decomposition.
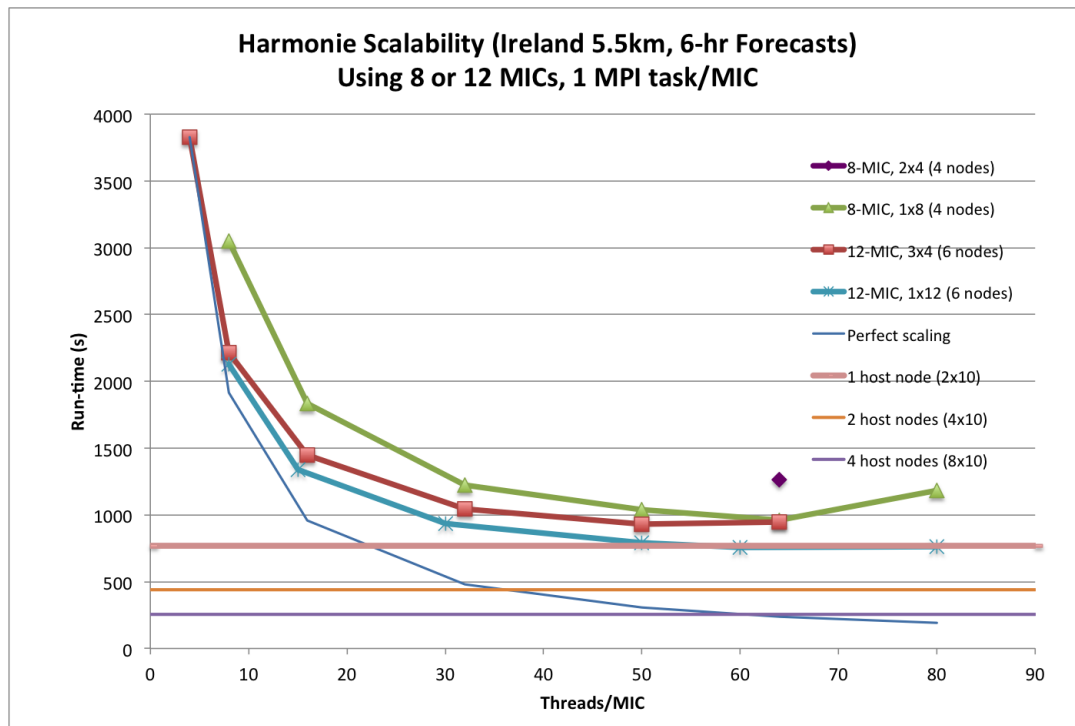
Figure 1: Run-times (s) for a 6-hr forecast of the HARMONIE IRELAND55 case, using 8 or 12 MICs, with 1 MPI task/MIC.

However, when the total number of MPI processes is much smaller than the number of north-south grid points, it makes sense to revert to a 1-dimensional domain decomposition by setting $M = 1$, since then each process "owns" a complete set of east-west points and can perform FFTs in that dimension without the need for any data transposing beforehand. With a 2-dimensional domain decomposition, data transposition is required to give each process a complete set of zonal points so they can each perform their FFTs, and that transpose is computationally expensive.

One of the performance opportunities provided by Xeon Phi is that it permits the use of more efficient 1-dimensional domain decompositions, since it makes sense on Xeon Phi to use very few MPI processes but as many threads as possible. The value of the 1-dimensional decomposition is apparent in Fig. 1, since over 12 MICs, a 1×12 decomposition (blue asterisks) is consistently faster than the 3×4 decomposition (red squares) for the same number of threads. Over 8 MICs, only one run was done with a 2×4 decomposition (purple diamond at 64 threads/MIC), since the corresponding 1×8 result (green triangle) was so much faster that only the 1×8 decomposition was used thereafter.

A final feature to note in Fig. 1 is that no run was possible using more than 80 threads/MIC — using xsany more threads caused jobs to fail due to lack of memory. Admittedly none of the performance curves in Fig. 1 look like they will scale past 80 threads/MIC in any case. However, the fact that the 12-MIC runs scale better than the corresponding 8-MIC runs (which have fewer total threads) suggests that the lack of scalability might itself be an artefact of the lack of memory, instead of the usual reason that each thread has too little work to do.

75

4.5   HARMONIE in "Symmetric Mode" on Xeon Phi

Given a HARMONIE executable (MASTERODB) that runs on host nodes, and a MASTERODB.MIC version, built with "`-mmic`", that runs on Xeon Phi MICs, it should be possible to configure a single job to place MASTERODB processes on host nodes, and MASTERODB.MIC processes on MICs, in such a way that all processes cooperate — just as they do on a homogeneous distributed-memory cluster — and all available resources are productively used.

With no changes to HARMONIE source, it is only possible to run HARMONIE in this way if the same number of threads (OMP_NUM_THREADS) is used on both hosts and MICs. Thread-specific MPI-communicators are created in `mpl_init_mod.F90`, and used to send messages between counterpart threads on each MPI task. If OMP_NUM_THREADS is not the same for all nodes, both hosts and MICs, some "sends" will have no matching "receives", some "receives" will have no matching "sends", and the job will fail.

Ideally, however, OMP_NUM_THREADS should be different between the host processes and the MIC processes. Tables 4 and 5 suggests that all physical cores on each host should each run a single MASTERODB process, with OMP_NUM_THREADS set to 1 or 2 (for hyper-threading), while the Xeon Phis should run just a handful of MASTERODB.MIC processes (even just 2 or 4), with OMP_NUM_THREADS set to as large a number as allowed by the memory or the number of logical cores on the device.

In order for HARMONIE to run in symmetric mode with these different OMP_NUM_THREADS settings, `mpl_init_mod.F90` was modified to create only as many communicators as the *minimum* value of OMP_NUM_THREADS across all MPI tasks. When only these communicators are used in message-passing, each "send" has a corresponding "receive", and vice versa. This was done with help from Sami Saarinen of CSC, Finland. Source code changes are minimal: essentially the introduction of just one extra variable (the global minimum of OMP_NUM_THREADS), and one call to MPI_Allreduce to find it. It is also important that OML_MY_THREAD remains fixed at 1 in `oml_mod.F90`.

With those changes, HARMONIE runs in symmetric mode, with arbitrary numbers of threads on both host and MIC nodes. As an example, given a host node with 20 physical cores and 2 Phi cards attached, a job can be launched with a total of 24 MPI processes in a 4×6 domain decomposition; 20 of these on the host (each with 2 threads), and 2 on each of the 2 Phis (each using say 40 threads). This is achieved using a "configuration" file of the form:

```
-n 20 -env OMP_NUM_THREADS=2  -host ${host}      MASTERODB     -maladin
-n 2  -env OMP_NUM_THREADS=40 -host ${host}-mic0 MASTERODB.MIC -maladin
-n 2  -env OMP_NUM_THREADS=40 -host ${host}-mic1 MASTERODB.MIC -maladin
```

Here "`-maladin`" represents the long list of HARMONIE arguments; `${host}` is the name of the host node, and for simplicity, other environment variables (e.g., `KMP_STACKSIZE`) are not shown. The job is then launched with, e.g., "`mpiexec.hydra -configfile $config`", where `$config` is the name of the configuration-file above.

A 20-process job (4×5 decomposition) runs in 663s on a single 20-core IvyBridge host node (using OMP_NUM_THREADS=2). The challenge when using Xeon Phi is to reduce that 663s run-time. The way to do this in "symmetric" mode is to partition the problem to run partly on the 2 Xeon Phi cards available, thus reducing the amount of work to be done on the host. A 4×6 decomposition allows a 20 + 4 MPI process "split"; 20 on the host, with 2 on each of the 2 MICs. This places quite a modest demand on the MICS: only one-sixth of the total problem is being run on them, so even "perfect" scaling would reduce run-time by only about 17%.

Fig. 2 shows performance of "symmetric mode" runs like this as a function of MIC thread-count (blue diamonds). The best run-time achieved was 690s, using 32 MIC threads per MPI task, though performance is essentially saturated with 24 threads. Using more than about 40 threads is not possible anyway due to lack of
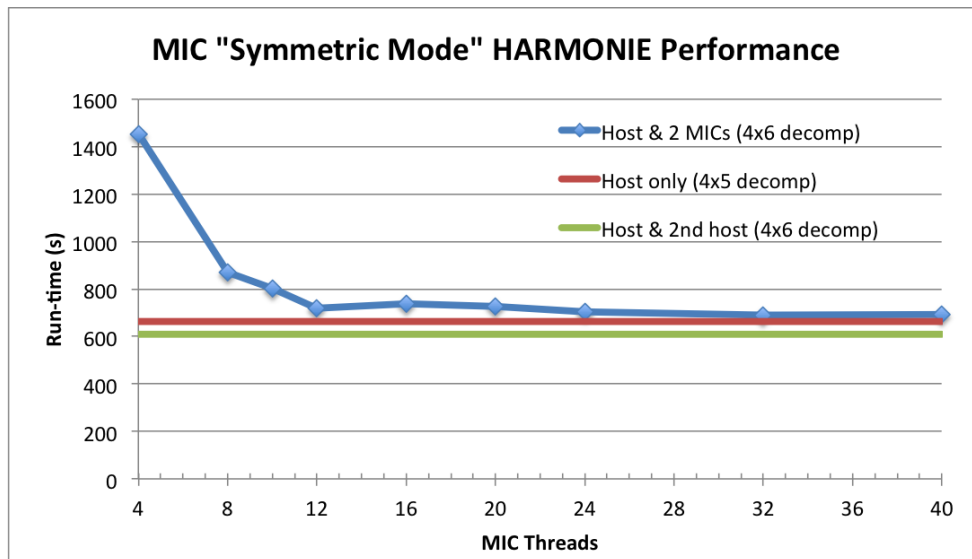
Figure 2: Total run-times 6-hr forecasts of the Ireland55 case, as in Fig. 1, when run in "symmetric" mode (diamonds), with 20 MPI tasks on the host node, and 2 on each of 2 MICs, as a function of MIC threads. Also shown is the host-only result (horizontal red line), and the result with 4 MPI tasks on a second host instead of on the 2 MICs (green line).

memory. So MIC cards do not work as "accelerators" in this case at all: placing part of the problem on them actually slows everything down relative to the "host only" run (horizontal red line). However, if the 4 extra MPI tasks are placed on a second host instead of on the MICs, total run-time is indeed reduced to 609s (horizontal green line). This represents a parallel speedup efficiency of 91%, which is quite respectable, and typical for HARMONIE scalability on cases like this.

The overall message to be drawn from Fig. 2 is that the only way to accelerate HARMONIE is still to use extra host nodes. Placing even a small part of the work on Xeon Phi coprocessors actually slows everything down — partly because of the poor MPI latency and bandwidths shown in Tables 2 and 3, partly because of the "small problem" saturation effect. Placing larger fractions of the total problem on the Xeon Phis means that not enough MIC threads can be used to provide scalability, due to the 8GB memory limit of the cards.

5   Conclusions

While it is relatively easy to port HARMONIE to the Xeon Phi (in contrast with porting HARMONIE to GPUs), it is difficult to harness all the computing power of the MICs to obtain performance speedups that would make Xeon Phi a more attractive platform than a standard Xeon node.

While speedup factors of $3\times$ or more are possible from standard Xeon nodes to Xeon Phi coprocessors, such speedups are crucially dependent on the problem size being large enough to scale to 60 or more threads, while remaining small enough to fit in the 8 GB memory available per device. The HARMONIE configurations tried so far cannot satisfy these contradictory requirements, since each additional thread that is used consumes substantial incremental memory (through variables listed as "private" for OpenMP, and through local variables in subroutines called within OpenMP parallel regions).

Apart from the default position of waiting for Xeon Phis with more memory, the obvious approach with future work will be to see if the HARMONIE memory footprint can be reduced, possibly through elimination of

some "local" variables within OpenMP parallel regions.

Other avenues to be explored in future work include the use of vectorization (i.e., the "AVX" extensions). AVX has typically not been used so far because it destroys bit-wise reproducibility from run to run. Nevertheless, if use of vector extensions provided enough performance gains, they could be used for operational runs as long as bit-wise reproducibility could be enforced again for debugging purposes.

Reference

[1] O'Brien, E.W., and A. Ralph, "Can HARMONIE be accelerated with GPUs or coprocessors?", ALADIN-HIRLAM Newsletter No. 1, 64-71, 2013.

FROST-2014 – performance of Harmonie 1km during Sochi Olympics

Sami Niemelä, Sigbritt Näsman and Pertti Nurmi,
Finnish Meteorological Institute

Introduction

The 2014 winter Olympic games were held in Sochi, Russia, in the period 8 to 23 February 2014. The World Meteorological Organization (WMO) organized both Research Development Projects (RDP) and Forecast Demonstration Projects (FDP) during the Olympics. The purpose of the projects were to advance and demonstrate the status and benefits of the state-of-the-art forecasting and nowcasting systems in a highly demanding operational environment. The FROST-2014 (Forecast and Research: the Olympic Sochi Testbed) project focuses in particular on the development/demonstration of modern short-term Numerical Weather Prediction (NWP) and nowcasting systems over a mountainous region in winter-time conditions and on the assessment of the practical use of this information.

FMI contributed to the FROST-2014 FDP by running the high resolution mesoscale NWP-model Harmonie with a 1-km horizontal grid size over the Sochi region, on a daily basis, before and during the Olympic games. The preparation phase of Harmonie FDP took place in winter 2013. The main goal of the preparation was to compare the Harmonie model with existing operational models, to find out the effect of alternative orographic datasets and to study the effect of the model grid size in the Harmonie (Niemelä et al.; 2013). The purpose of this study is to evaluate the performance of the Harmonie model during the Sochi Olympics. The evaluation is based on i) user feedback and ii) verification with the observations provided by the local organizers.

Observations in Sochi area

The characteristics of the Sochi region are explained in detail in Niemelä et al. (2013). The area is located on the eastern coast of the Black Sea (43.4º N, 40.0º E). The events took place in two specific areas: in the coastal (Adler) and the mountain cluster. Here we focus on the mountain cluster since all weather sensitive sporting events took place there.

Figure 1a shows the locations of the sporting events in the mountain cluster. The Olympic park was in the valley, in the village of Krasnaya Polyana, which is located about 560 m above the sea surface. Most of the ski events took place on the northern slope of the valley, in Laura ski stadium, approximately 1400m above the sea level. Alpine skiing and freestyle events were held on the southern side of the valley, in the Rosa Khutor skiing area. There the different events took place at heights between 1000 and 2100 m above sea level. Some skiing events were held below 1000 m. Ski jump and Nordic Combined competitions took place in the valley, in the vicinity of the Ski Jump stadium at 650m above the sea level.

Figure 2b shows the Harmonie orographic height map with 1-km grid size in the mountain cluster area. Furthermore, the locations of the observation stations used in this study are shown. The verification analysis is done separately based on the altitudes of the selected stations: with respect to valley (dots) and mountain (cross and circles) stations.
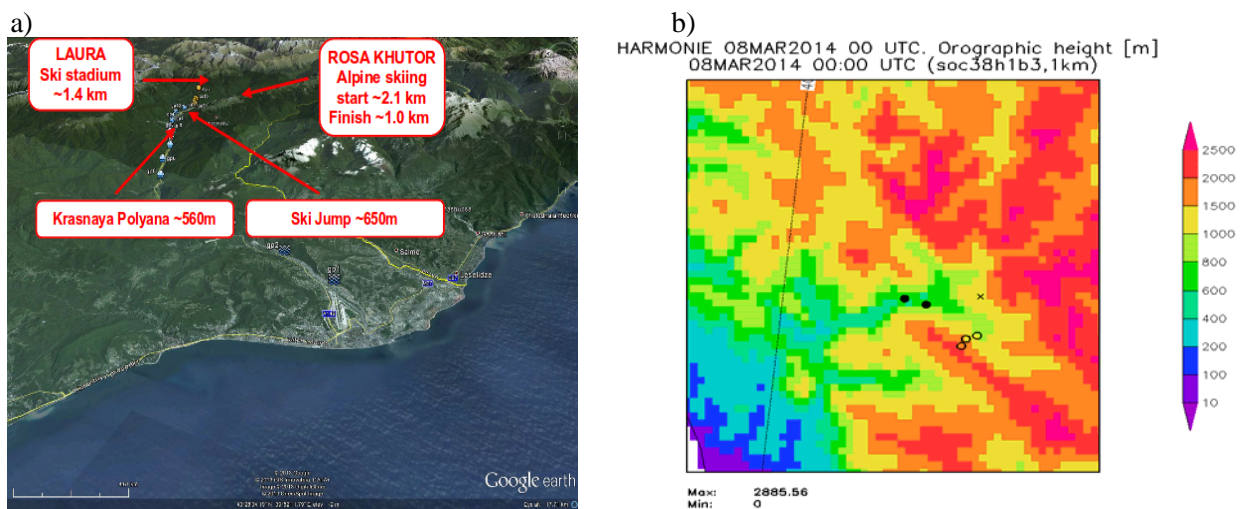
Figure 1:   *Characteristics of the Olymic area showing a) the main sporting event locations and their altitude from the sea level, b) Harmonie orographic height in the mountain cluster area with main observation locations (dot = valley observations, Karsnaya Polyana and Ski Jump stadium, cross = ski stadium in the mountains and circle = alpine skiing route, start, middle and finish).*

Model configuration

Harmonie is a non-hydrostatic model based on fully compressible Euler equations. The configuration of the Harmonie model for the Sochi FDP was based on the experience gained in Niemela et al. (2013).

The model system over Sochi was based on the latest (at the time) Harmonie cy38h1.beta.3 with 1 km horizontal grid size and with 65 levels in the vertical (soc38h1b3). Figure 2 shows the domain of the Harmonie setup (640x500 points). This configuration uses the Arome-physics package (Seity et al. 2011). Upper air data assimilation is handled by 3D-Var, whereas optimum interpolation is used for surface variables. Background error statistics were created by using an ensemble method. Forecasts were provided from December 2013 to March 2014, +36 hour forecast 4 times per day with a 6-hourly assimilation interval including conventional observations only. Hourly boundary conditions are taken from the ECMWF global model. The differences between the FDP experiment and the reference Harmonie configuration were the following:

     i.   use of Predictor-Corrector (PC) time stepping scheme (with 30s time step) instead of Stable Extrapolating Two Time Level Scheme (SETTLS) as explained in Niemela et al. (2013),

     ii.   use of Ecoclimap version 1 instead of version 2, and

     iii.   excluding LSMIXBC option, where large scale features from the boundary conditions are blended into the first guess of the assimilation process.

Results

The evaluation of the model performance is twofold, based on both the subjective user feedback and the verification against observations. Therefore, the comparison presented in this paper focuses mainly on the period from the start of the Olympic Games (7.2.2014) to the end of Paralympic Games (15.3.2014). In the FDP, it was agreed that all the participating institutes send their forecasting data into the FROST ftp-site, where the local organizers can use the data in the actual forecasting work as

they wish. In addition, FMI provided the Harmonie forecasts directly to Finnish Olympic Committee (FOC) ski maintenance team as an exchange of possible user feedback and comments.
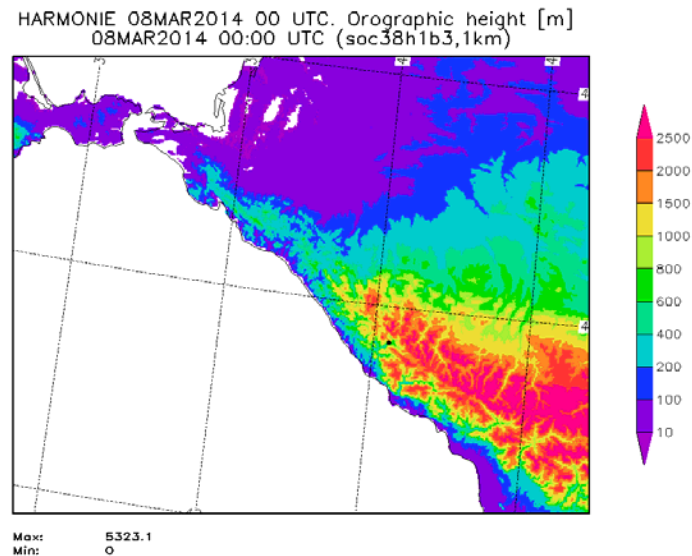

*Figure 2:   The domain of the Harmonie model with 1km grid-size (640x500 points). The location of
              the mountain cluster is marked as a dot.*

User feedback

Figure 3 shows an example of a five panel meteogram representation produced from the Harmonie direct model output. FOC was provided with meteograms of 20 locations around the mountain and coastal clusters. FOC listed the most interesting products as follows:

    i.   The 1st panel was used to detect the cloudy vs. cloud free days and the timing of the fog (cloud present on the lowest model levels),

    ii.   The 2nd and 3rd panels were used to detect the timing and the phase of the precipitation,

    iii.   The 4th and 5th panels were used for getting a general overview on temperature and wind conditions. Although the 3rd panel gives the phase of surface precipitation directly, the 2nd panel was used for detecting the height of the melting layer. This was useful for estimating the uncertainty of the precipitation phase especially in cases of a melting layer close to the surface.

All the following comments are based on meteogram representation only.

In general, the user feedback was exceptionally positive. FOC praised Harmonie being "superior" compared to operational forecasting products given by the local organizers and FMI (based on ECMWF and HIRLAM). Furthermore, the fog (occurrence and timing) and cloud height forecast were said to be "excellent". Fog is considered to be one of the most challenging phenomenon for any atmoshperic model. Therefore, the positive feedback was surprising. However, the good performance for fog conditions can be partly explained by the geographical features of the target area. NWP model are typically good in predicting the larger scale low pressure systems and the related cloud/precipitation areas. In the Sochi mountain cluster the well forecasted fog situation were connected to low pressure systems moving from the Black Sea. Well predicted cloud areas (height about 1000-2500 m above sea level) were seen as fog, when the low pressure system hit the mountains. The visual representation of the data was considered "very good" by FOC.

Observation verification

The main focus of the verification is on the performance of the Harmonie model. For comparison purposes, other available (from FROST ftp-site) model data was used as well. Models were chosen on the basis of grid size and forecast length to be close to or equivalent with Harmonie. Other models in the comparison are COSMO-RU (2.2 km, see COSMO documentation) , NMMB (1 km, see NMMB documentation) and GEM (1km, see GEM documantation).
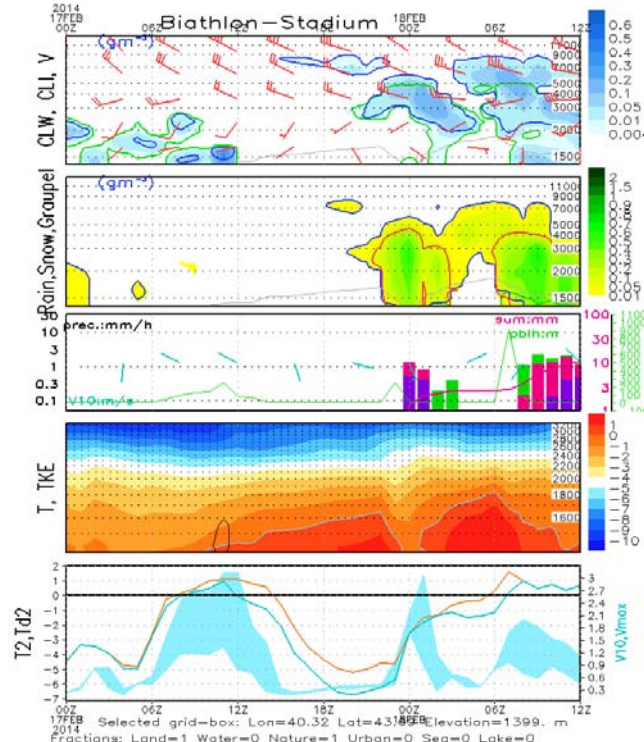

Figure 3:   An example of a meteogram product for the Biathlon stadium based on Harmonie model output. 1st panel: total cloud water content [gm^{-3}] in colour, location of cloud water (green)/ice (blue) and wind barbs at different heights. 2nd panel: total precipitating hydrometeor content [gm^{-3}] in colour, location of rain (green), snow (blue) and graupel (red). 3rd panel: Hourly precipitation [mm] (bars) with different phase (green=rain, purple=snow, red=graupel). Total accumulated precipitation [mm] is given as a red solid line. 4th panel: temperature (colour) and TKE (black contour) within boundary layer. 5th panel: 2m temperature (red) and dew point (cyan). The lower edge of the colour field shows the 10m wind speed, whereas the higher edge gives 10m gust wind speed. The dashed line in 1,2 and 4 represent the height above the model surface.

Figure 4 shows the average diurnal cycle of the 10m wind speed in both valley and mountains. Overall, the whole study period was characterized with relatively weak wind conditions. In the valley Harmonie seems to be able to capture the weak diurnal cycle of wind speed better than other models. The GEM model generally underestimates the wind speed and the NMMB and COSMO seem to have opposite diurnal cycle than observed. In the mountains the Harmonie and GEM models give the best wind speeds on average, whereas NMMB and COSMO overestimates the wind speed. From the wind speed point of view, Harmonie performs the best.

Figure 5 shows the average 2m temperature in both valley and mountains. In the valley Harmonie has the largest diurnal cycle (similar as observed). All models tend to underestimate the temperatures in the valley. In the mountains all model predict well the day-time temperatures. During night-time all

models tend to underestimate the temperature. However, the underestimation in Harmonie is very large, on the average 4 – 5 °C. Overall, Harmonie is among the best models during the day-time, but is the worst during night-time.
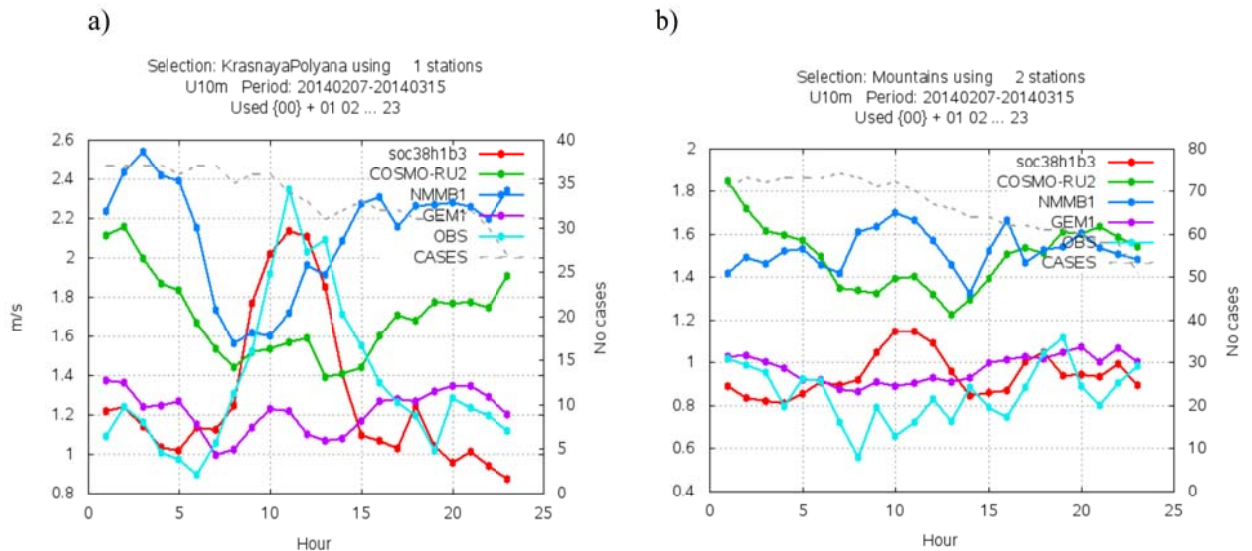

Figure 4:   The average diurnal cycle of wind speed in a) valley and b) mountains. Red is Harmonie, Green is COSMO, blue is NMMB and purple is GEM. Observations are given in light blue.
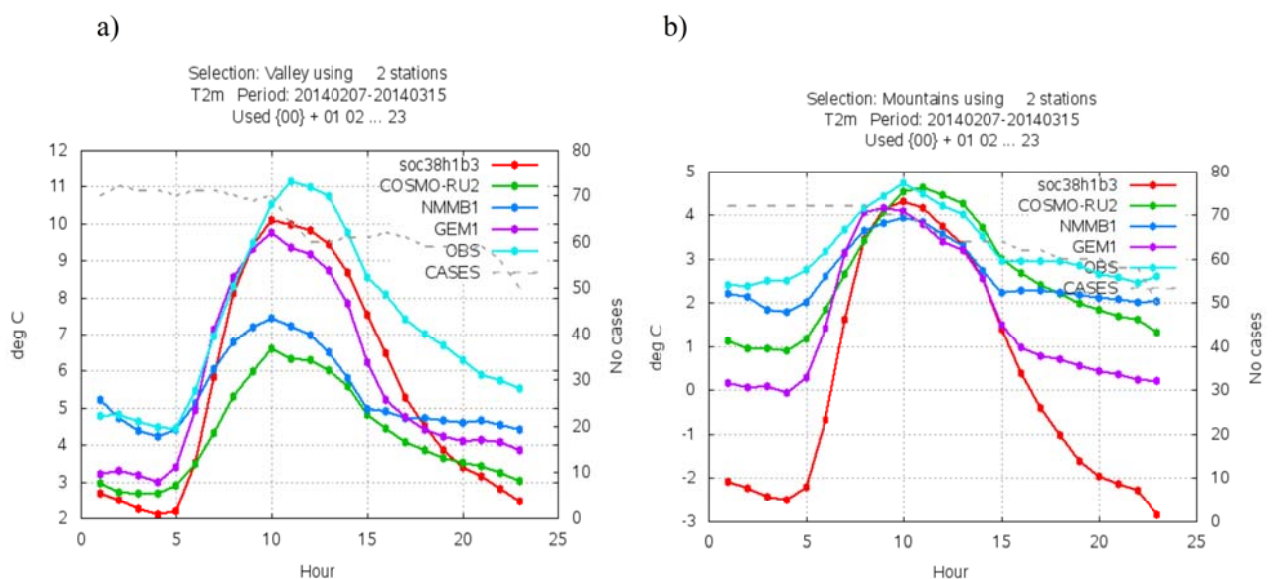

Figure 5:   The average diurnal cycle of temperature in a) valley and b) mountains. Colours are the same as in Fig. 4.

Temperature problems in Harmonie

Figure 6 shows the scatter plots of Harmonie temperature forecasts in the mountains. Underestimation is well present and can be as large as 10 °C in individual cases (Fig. 6a). Furthermore, splitting the period in moist-and-windy (Fig. 6b) and dry-and-calm (Fig. 6c) cases it is evident that most of the underestimation cases occur in dry and calm mountain conditions.

Figure 7a shows a temperature structure in a single Harmonie forecast representing typical dry and calm conditions in the mountains. During the night-time the 2m temperature is well below -10 °C

whereas the observation is -1 °C (not shown). However, the lowest model (~12m height) level temperatures (0 °C) are very close to the observed 2m value indicating that the surface is somehow detached from the atmosphere. During the day-time Harmonie 2m temperature is +1°C and corresponds the observed value (+2 °C, not shown) very well.

a)   b)   c)
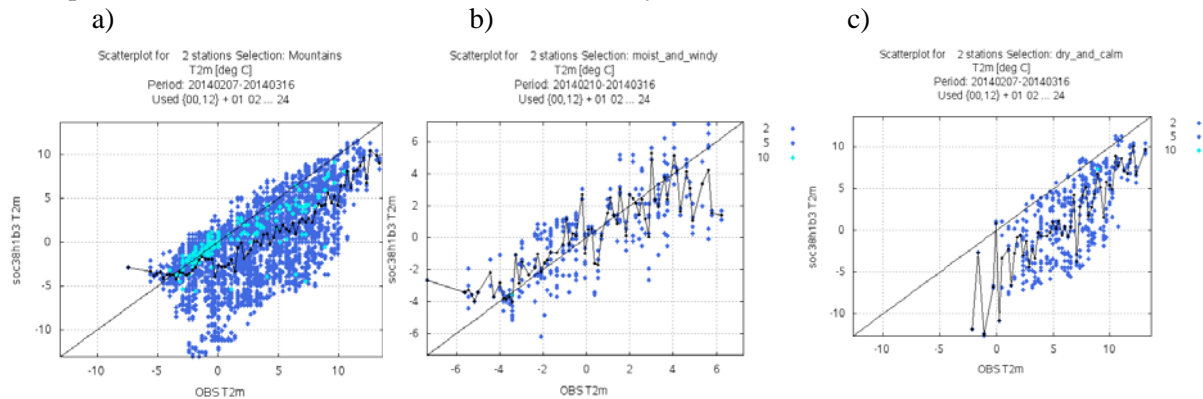

Figure 6: *The scatter plots of Harmonie temperature (observations vs. model). a) All cases in the mountains. b) Cases with relative humidity over 70% and wind over 1ms^{-1}. c) Cases with relative humidity below 70% and wind below 1ms^{-1}.*

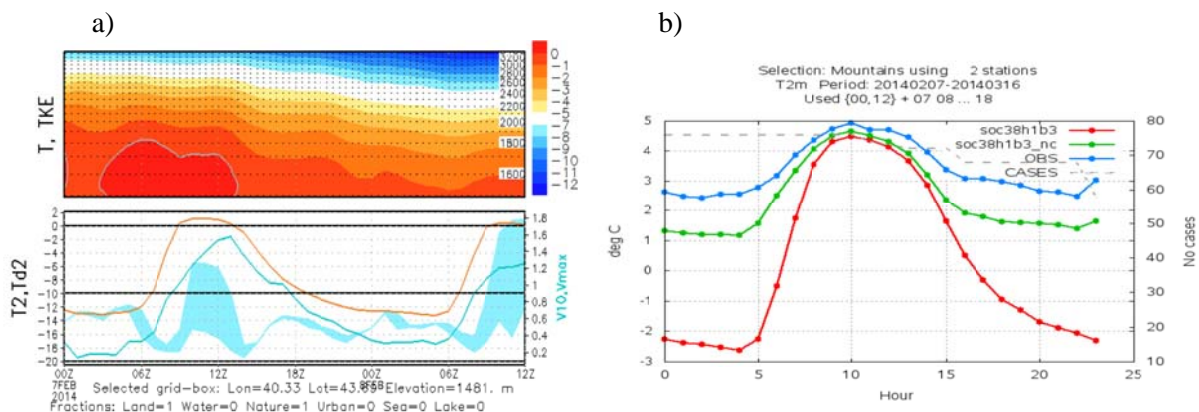a)   b)

Figure 7: *a) Boundary layer temperature structure (upper) and 2m temperature (lower, red) timeseries from Harmonie model cycle 2014020700. b) The average diurnal cycle of 2m temperature in the mountains. Red is original Harmonie, green is Harmonie without CANOPY scheme and blue is observation.*

Screen level temperature and wind are diagnosed by the so called CANOPY-scheme (Masson and Seity; 2009) in Hamonie. CANOPY is a prognostic 1D sub-layer turbulence scheme between the lowest model level and the surface with 6 vertical levels. The scheme predicts the evolution of TKE, wind, specific humidity and temperature and provides diagnostics of the screen level values for each of them. A model experiment exploring the role of the CANOPY in the seen temperature problems was conducted by switching off the CANOPY scheme. Figure 7b shows the average diurnal cycle of 2 m temperature over mountains with and without the CANOPY-scheme. During day-time the impact is very small. However, during night-time the CANOPY-scheme is responsible for a large part of the underestimation. Without the CANOPY scheme the temperature is still being underestimated, but at a much more reasonable level 1-2 °C.

84

Issues with numerical instability

Niemela et al. (2013) faced several model crashes during the preparation phase of the FDP. The model turned out to be more stable after using the PC-scheme instead of SETTLS. However, during FDP Harmonie suffered still some instability issues. Sometimes the pressure fields were very noisy with "delta-2-x" type of features.

a)   b)
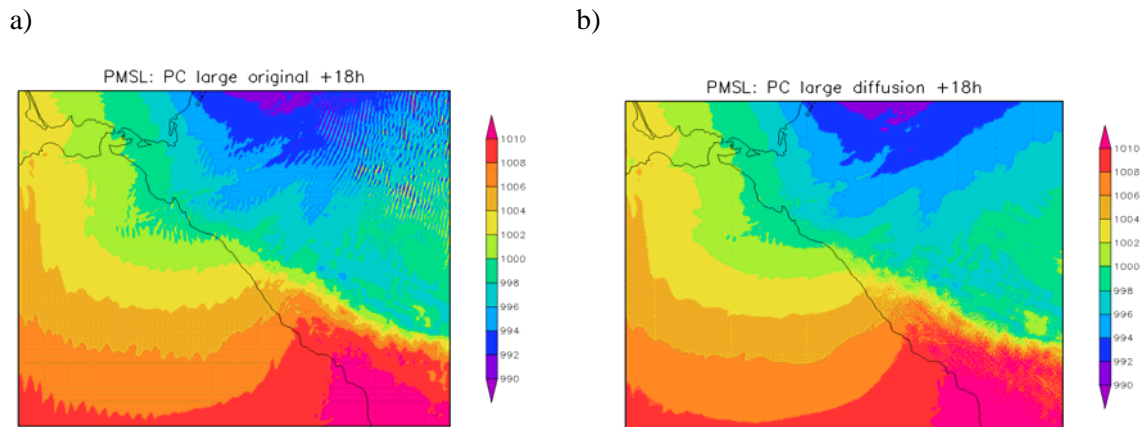

Figure 8:   Mean sea level pressure field with a) default and b) increased horizontal diffusion.

An additional model experiment was prepared to study the effect of horizontal diffusion (Mariano Hortal, personal communication). Figure 8 shows the mean sea level pressure field with default settings (RDAMPDIV=RDAMPVD=RDAMPVOR=RDAMPT=20) and increased settings (RDAMPDIV=RDAMPVD=RDAMPVOR=RDAMPT=1) horizontal diffusion settings. Increasing the horizontal diffusion reduces the noisy pressure features considerably. Effect of the diffusion on other parameters has not been studied. Therefore, one should be careful in applying stronger diffusion in km-scale model setups.

Conclusions

This paper presents the results of the Harmonie 1-km model over the region of Sochi Olympics during the Forecasting Demonstration Project part of FROST-2014. In addition, the performance of other participating high-resolution models are studied as well for the comparison purpose. The main results are following:

- Based on subjective user feedback by the Finnish Olympic Committee ski maintenance team, the Harmonie forecasts were considered highly useful. Especially, the forecasting capability of the fog, cloud height and precipitation (phase) were praised.
- Overall, in the verification intercomparison (wind and temperature) the Harmonie model was among the best models in FROST-2014.
- However, Harmonie suffered of a large cold bias in dry and calm conditions in the mountains. Experiments without the CANOPY-scheme improved the temperature scores considerably in these particular conditions.
- Harmonie had some numerical instability problems occasionally which occurred as noisy mean sea level pressure pattern. Increasing horizontal diffusion resulted in less noisy fields.

References

COSMO-documentation, available online: http://www.cosmo-model.org/content/model/documentation/core/default.htm

GEM-documentation, available online: https://www.earthsystemcog.org/projects/dcmip-2012/gem-latlon

Masson V. and Y. Seity, 2009: Including atmospheric layers in vegetation and urban offline surface schemes. *J. Appl. Meteor. Climatol.*, **48,** 1377–1397.

Niemelä S., E. Atlaskin, S. Näsman and P. Nurmi, 2013: FROST-2014 – First experiences of deterministic Harmonie in Sochi region. *ALADIN-HIRLAM Newsletter*, **1**, 72—77. http://www.hirlam.org/index.php/meeting-reports-and-presentations/doc_download/1458-aladin-hirlam-newsletter-no-1-september-2013

NMMB-documantation, available online: http://www.bsc.es/earth-sciences/nmmbbsc-project

Seity, Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, V. Masson, 2011: The AROME-France Convective-Scale Operational Model. *Mon. Wea. Rev.*, **139**, 976–991. doi: http://dx.doi.org/10.1175/2010MWR3425.1

AROME-France 1,3 km status and plans (Model part)

Yann Seity, Météo-France CNRM-GAME

Introduction

In 2014, computing power at Météo-France significantly increased thanks to 2 IBM scalar machines. As for the previous calculators updates, model resolutions will be increased. In 2015, the ARPEGE stretched global model will have an horizontal resolution around 7 km over France, with 105 vertical levels. AROME horizontal and vertical resolution will switch from 2,5 km and 60 levels to 1,3 km and 90 levels. Daily AROME experiments at these resolutions have started on a reduced domain since June 2012. In order to prepare the operational configuration, it has been completed by full domain experiments without data assimilation by mid-2013. This paper presents this new AROME configuration (in cy40_op) : in the first part both in terms of dynamics, physics and technical choices. Some objective scores are provided in the second part, whereas the last part will give the status of the work concerning a new microphysics scheme developed for Meso-NH and AROME.

Overview of the operational configuration

Preparatory work started in 2012 with the choice of the 90 vertical levels. The lowest model level altitude (5m versus 10m currently) was the best compromise between a reduced time step and the quality of model forecasts, particularly in fog events. After some tests, a regular vertical resolution increase between L60 and L90 AROME has been chosen. For stability reasons, the model top will be at 10 hPa (versus 1hPa currently). AROME will use 33 levels below 2000 m height (versus 21 in AROME - 2,5 km L60).

Concerning the horizontal grid, the model domain at 1,3 km resolution (Figure 1) contains 1440x1536 points on a Lambert projection centred at (2°E;47,5°N). It has been slightly enlarged to the north compared to the current operational one. With this horizontal resolution increase, the orography is more realistic. Moreover, a new orographic data base is used (GMTED2010 at 250 m, instead of GTOPO30 at 1 km). Over the mountains, it provides deeper valleys and higher peaks. The maximal slope is 38° (versus 23° currently).

The differences between the dynamic model settings of AROME 1,3 km and AROME 2,5 km are listed in Table1. The model time step (45 s) has been reduced, and for numerical stability reasons, an iterative centred implicit (ICI) time-scheme with one iteration (Bénard, 2003) has been applied, with ND4SYS=2 and RDAMPPD=20.

The SLHD diffusion scheme applied on hydrometeors has been re-tuned by using a new setup (SLHD_OLD=F). Coupling area size has been enlarged up to 16 points, and top spectral relaxation has been adapted to the new set of vertical levels. More conservative new Semi-Lagrangian weights has been used (Malardel and Ricard, 2014) on temperature, wind, specific moisture, surface pressure, pressure departure and vertical divergence.

For the physics, the following modifications are implemented:

- on surface : surfex cycle update from version v6+ to version 7.2, with a retuning of surface orographic drag due to lowest model level height change, and the SBL scheme (Masson and Seity, 2009) switched off.

- on microphysics: the snow auto-conversion threshold has been increased in order to fix problems of dissipation of stratus in negative temperature situations (see B. Szintaï paper in this newsletter). The new threshold used is 0,2g/kg instead of 0,02 g/kg currently.

Technical optimisations have also been performed such as the use of a dedicated I/O server. Meso-NH physics used in operational AROME is now able to run with the ARPEGE/ALADIN vertical levels ordering and most of it without additional top and bottom levels (except for the turbulence scheme). It allowed a significant cleaning of the interface routine apl_arome.F90.
274 of our 1000 BULL nodes will be required to perform AROME 1,3 km 24h forecast in 30 minutes elapse time with mixed MPI/OpenMP parallelisation.
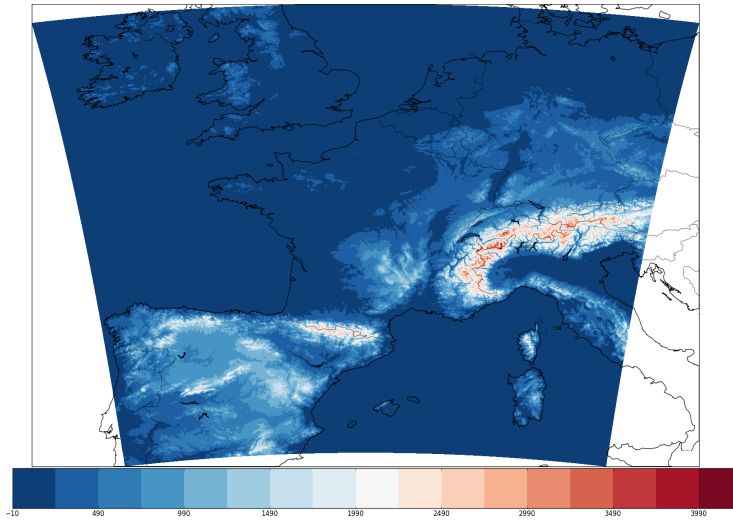

Figure 1: AROME-1,3km domain

	AROME 1,3km L90	AROME 2,5km L60
Time step	45s	60s
ICI scheme (NSITER=1)	T	F
PC_CHEAP	T	F
LGWADV/LRDBBC	T/F	F/T
ND4SYS	2	1
LSLHD_OLD	T	F
COMAD	T	F
Coupling area	16 points	8 points
Top spectral relaxation	T (retuned)	T

Table 1: Comparison of AROME dynamics settings
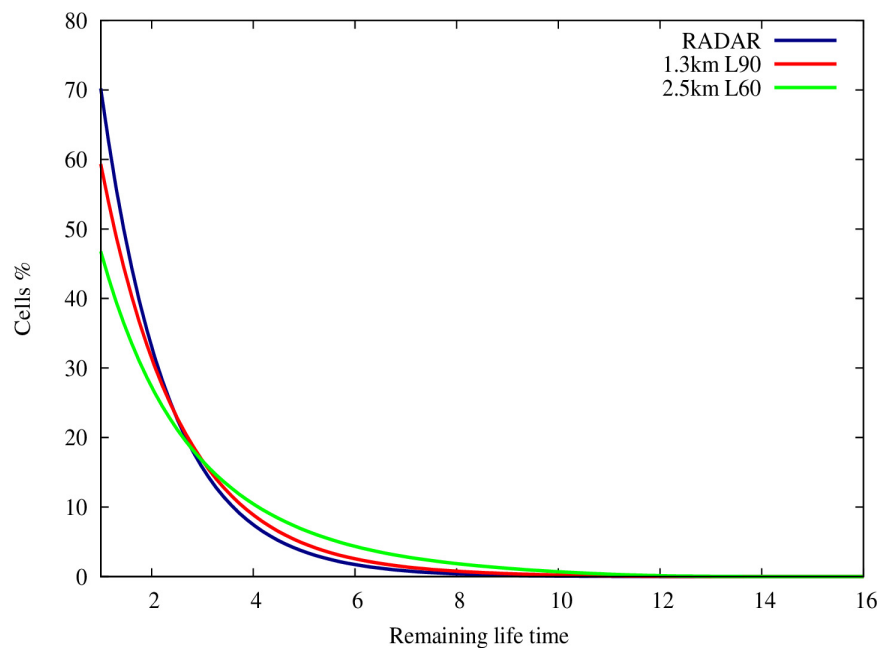
Objective evaluation

Figure 2: Statistics on convective cells life-time. 21St June 2012. On the horizontal axis, the remaining life time (in minutes) has been divided by 15.

By the statistical study of 48 convective days with the use of a cell tracking algorithm (Morel et al., 2002), we have shown (Seity et al., 2014) that AROME - 1,3 km L90 provides more realistic convective cells than AROME – 2,5 km L60. They are more numerous and more intense at 1,3 km. At 2,5 km, the AROME model produced too large cells.

Statistics of the convective cells remaining life time (remaining life time since they are detected on a radar or a model simulated radar reflectivity field) have been calculated by Rahma Ben Romdhane for 2 cases (21st June and 4th July 2012) at 15 minutes fields temporal frequency. The results are very close between the 2 days. It shows that AROME 2,5km over-estimated the convective cells life time, whereas AROME 1,3km is closer to radar observations (Figure 2).
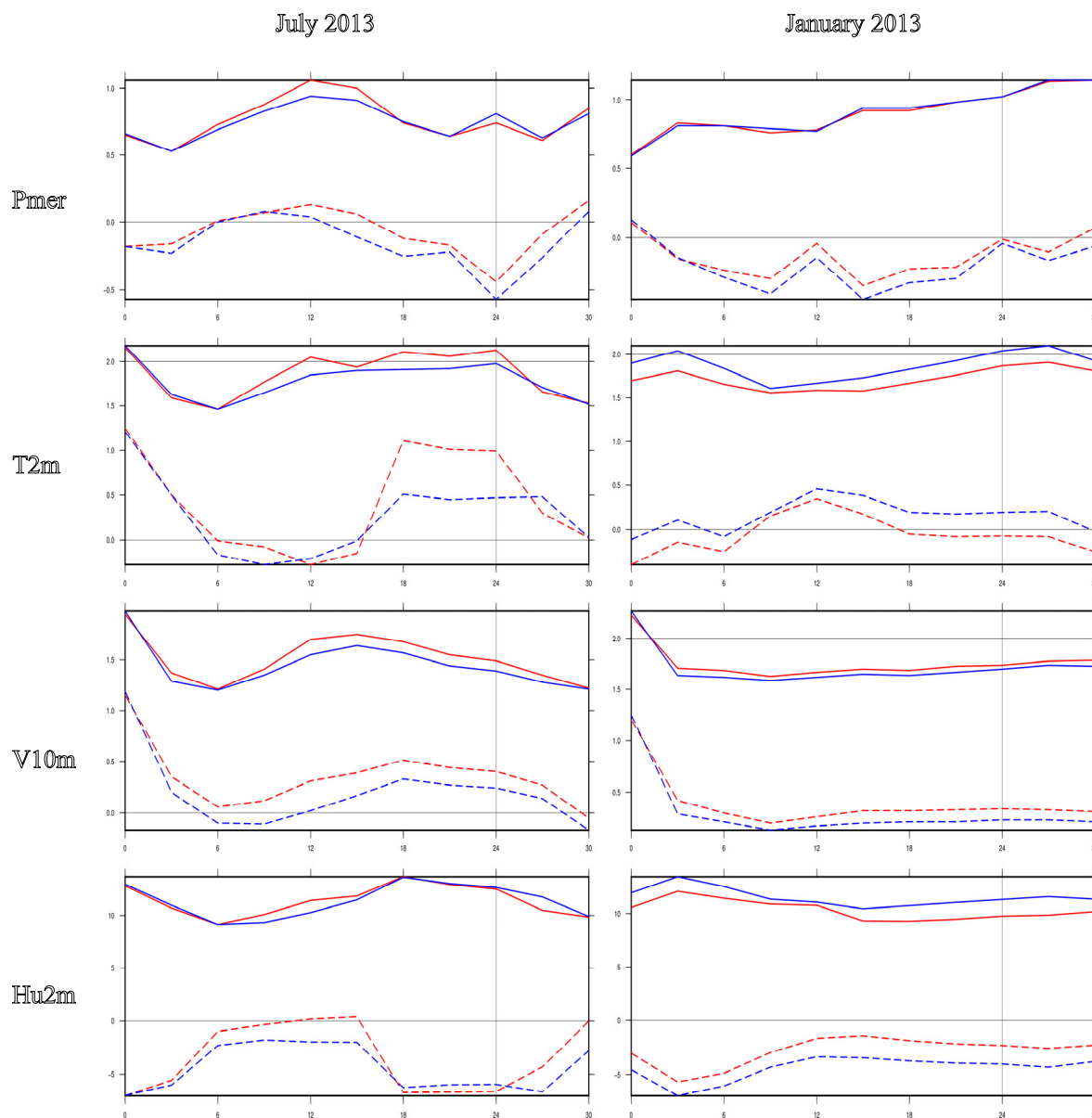
Figure 3: AROME surface scores of surface pressure(~150 obs), 2m temperature(~1000 obs), 10m wind speed (~700 obs) and 2m relative humidity (~800 obs) as a function of forecast step (in hours). AROME forecasts starts at 3TU, and are initialized and coupled with ARPEGE 0 TU forecasts. AROME-2,5km is in red, AROME-1,3km in blue. Left column July 2013, right column January 2013. Bias in dotted lines and RMSE in continuous lines.
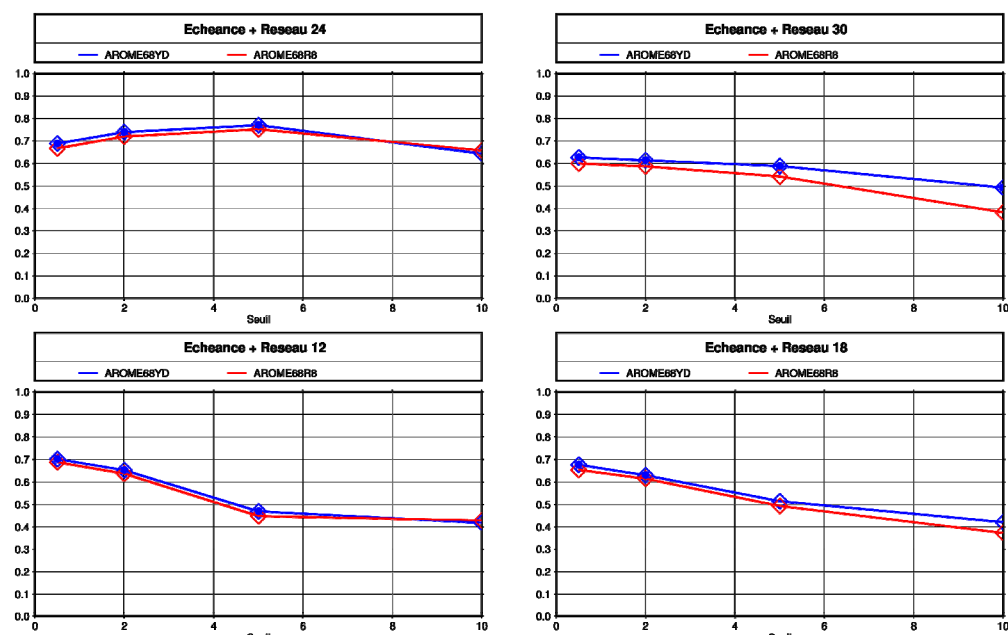
Figure 4a: Brier Skill Scores of 6h cumulated rainfalls with a neighbourhood of 20 km, as a function of precipitation thresholds (>0,2>2,>5 and >10 mm/6h), for 4 forecasts times (12 TU, 18 TU, 24 TU and 30=6 TU the next day).31 daily forecasts starting at 3 TU in January 2013 are used.   AROME1,3km in blue, AROME 2,5km in red.
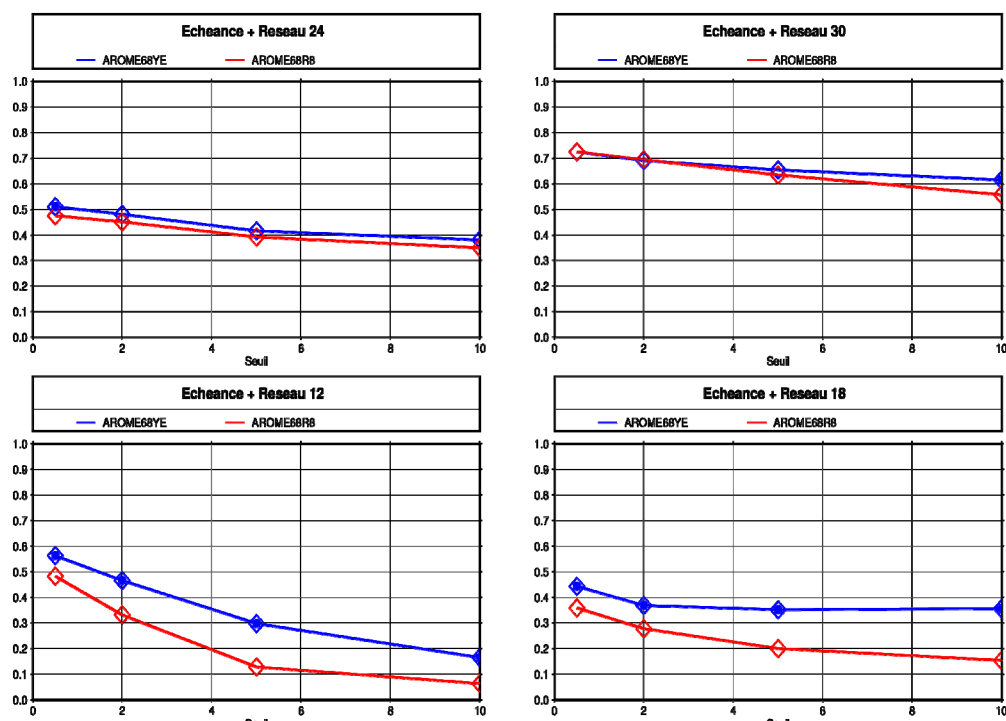*Blue dots indicates statistically significant differences between the 2 curves.*

Figure 4b: Same as Figure 4a, for July 2013.

Figure 3 displays the classical surface scores over 2 months, January and July 2013. These 2 months are very contrasting and contain different kinds of winter / summer weather conditions observed over France.

January 2013 has been 0,3 °C colder than the climatology, with a very low sunshine duration (50 % of the climatology). Precipitation is very contrasting, with 200 % of the climatology over the south west part of France, and a deficit in the north. Numerous snow events occurred between 13 and 26 January with very cold temperatures (for instance -16,1 °C at Rouvroy-en-Santerre in the north of France 16th January) whereas between 3rd and 8th, in anticyclonic conditions, fog occurs but no significant precipitations.

July 2013 was the third warmest July month since 1900, 1,9°C warmer than the climatology. Cumulative rainfalls are quite normal, mostly due to convective events during the second half of the month. Hail was frequently reported especially on 23rd. Very strong convective wind gusts (> 120 km/h) are reported on 27th July.

For both months, partly due to its more realistic orography, surface pressure and 10m winds (also 10m wind gusts, not shown)) are improved by AROME 1,3 km.

Concerning 2m temperature and humidity, the significant improvement observed during night time in summer is due to the SBL scheme which has been switched off. On the same time, we noticed a degradation in winter. This degradation would have been even stronger if SBL scheme would still have been used.

Precipitations scores based on fuzzy approach (Amodei and Stein 2009) are also improved (Figure 4a and b), mainly for 5mm and 10mm per 6h thresholds. The improvement is stronger in summer because Semi-Lagrangian COMAD new interpolators have more impact on convective events.

Simulated satellite radiances (not shown) are also improved, corresponding to an improved model cloudiness, even if an unknown part of this improvement may be due to changes in RTTOV code between CY38 and CY40.

LIMA

A new microphysics scheme, Liquid Ice Multiple Aerosols (LIMA), has been developed in MesoNH in order to improve the representation of complex aerosols/clouds/precipitations interactions (Vié et al., 2014). Aerosols fields acting as Cloud Condensation Nuclei (CCN) or Ice Freezing Nuclei (IFN) are taken from MACC products at ECMWF. MACC mixing ratios have to be converted into number concentrations (kg-1). Sulphate and sea salts fields are used as CCN whereas dusts, hydrophobic organic matter and hydrophobic black carbon are used as IFN. Hydrophilic organic matter and hydrophilic black carbon (called coated IFN) can act either as CCN or as IFN depending on their environment. A parametrisation of the aerosols activation is used to well represent nucleation of pristine ice and cloud droplets. With such a scheme, auto-conversion thresholds (cf part 1) are no longer used. Aerosols are transported by the model dynamics. Their main sink is the impaction scavenging by precipitation. New 3D prognostic variables are the following: concentration of cloud droplets (Nc), pristine ice (Ni) and rain (Nr) with a split between activated and non-activated ones for each mode of aerosols. With this kind of scheme, over-saturations remain more easily than in ICE3. It is interesting to point out that ICE3/ICE4 cold processes concerning hail/snow/graupel remain untouched. Thus, improvements in ICE4 for instance will automatically benefit to LIMA.

The scheme is now technically working in MesoNH. The validation/tuning work is on-going using observations from HYMEX. Thinkings concerning its implementation in AROME will start in autumn 2014. Questions on numerical efficiency, optional simplifications of the code and adaptation to long time-steps will be tackled at that time.

Conclusions and Outlook

The AROME 1,3 km model configuration is chosen and validated both on case studies and on monthly scores. Improvements are observed in summer and in winter in most of the classical meteorological parameters. For instance, convective precipitations positive bias is significantly reduced.
The model configuration improvements are now mixed with a 1,3 km data assimilation cycle based on a 1h Rapid Update Cycle. The E-suite will start this summer, with a switch to operational mode expected in the very beginning of 2015. It will be complemented by a ensemble AROME forecast system at 2,5 km (on the same geographic domain as AROME 1,3 km).
In order to prepare future AROME physics, thanks to a collaboration with ZAMG and FMI, the surface / radiation interaction will be updated with more advanced parametrisations. Moreover, in parallel with the LIMA 2-moments microphysics scheme implementation, an 'intermediate' strategy for fog forecasts will be evaluated. It consists in a daily initialisation of the Nc in the current ICE3 microphysics scheme, using input information from an aerosol model (MOCAGE or MACC).

References :

Amodei and Stein., 2009 : Deterministic and fuzzy verification methods for a hierarchy of numerical models,Meteorol. Appl., 16:191-203.

Bénard, P., 2003 : Stability of semi-implicit and iterative Centred-implicit time discretizations for various equations systems used in NWP, Monthly Weather Review, 131, 2479-2491.

Malardel S, Ricard D. 2014: More conservative interpolation for pointwise semi-Lagrangian transport scheme. Quart. J. Roy. Meteor. Soc. In revision.

Masson V. and Y. Seity, 2009 : Including Atmospheric Layers in Vegetation and Urban Offline Surface Schemes, J. of Appl. Meteorol. And Clim., 1377-1397.

Morel, C., S. Sénési, and F. Autones, 2002: Building upon SAF-NWC products: Use of the Rapid Developing Thunderstorms (RDT) product in Météo-France nowcasting tools. Proceedings, The 2002 Meteorological Satellite Data Users' Conference,Eumetsat and Met Eirean, Dublin, Ireland, 248-255.

Seity, Y., D. Ricard, J. Léger and M. Pietrisi, 2014 : Evaluation of increased horizontal and vertical resolutions in AROME-France deterministic forecasts for convective events, ALADIN-HIRLAM Newsletter N°2,16-19.

Vié B., J.-P. Pinty, 2014 : LIMA: A 2-moment scheme driven by a multimodal population of CCN and IN, 14th Conference on cloud physics, 7-11 July, Boston

List of events

[36th EWGLAM and 21st SRNWP meeting, 29/09-2/10/2014, Offenbach (Germany)](#)

[Regular 19th General Assembly and 1st joint ALADIN GA and HIRLAM Council, Reading, UK, December 2, 2014](#)

[25th ALADIN Workshop & HIRLAM All Staff Meeting 2015, 13-17/04/2015, Copenhagen, Denmark](#)

[3rd joint ALADIN PAC and HIRLAM HAC meeting and 12th ALADIN Policy Advisory Committee (PAC) meeting, Helsinki, Finland, May 21, 2015](#)

For updated event information please look at the consortia websites:

http://www.cnrm.meteo.fr/aladin/

http://hirlam.org

24th Wk/ASM 2014 in Bucarest