Regional Cooperation
for Limited Area modeling
in Central Europe

Prague, 23.01.2023

FRAME APPROACH IN COUPLING – REPORT

Author: Gabriel Stachura, IMGW
Supervisor: Petra Smolíková, CHMI

Outline

1. Introduction
   - Work done during the previous stay
   - The aim of this stay
2. Methodology
   - Reference LBC files
   - Experimental LBC files
3. Experiments
   - Description of experiments
   - Verification method
   - Verification results
4. Compression
   - Size reduction estimate
   - Compression methods
   - Compression results
5. Summary

1. Introduction

The report presents results of the continuation of work on the subject of frame approach in coupling. In general, the objective was to test whether compression of LBC files, which are needed to run a LAM model, is possible. The task was a part of the RC-LACE Workplan for

Dynamics and Coupling and was completed during two stays in CHMI in Prague in November 2021 and October 2022. We followed up the work done previously at Météo-France and summarized in [MF]. The main idea of the work is that in the ACCORD system coupling is realized in the grid-point representation and only in an outside frame of a domain, hence only this information needs to be contained in the LBC files. This frame is subdivided into (Fig.1):

- extension zone (E) – an artificial zone created for biperiodizing a meteorological field with data coming completely from LBC files
- intermediate zone (I) - a zone where data are a combination of a forcing from LBCs and inner evolution of the local model

In the remaining inner part of the domain - so called central zone (C) – physical processes are evolved freely without being directly affected by LBCs. When framing the files (or making the hollow) only data from the central zone may be removed, but not necessarily all of them. The size of the hollow is controlled by the parameter ICPLSIZE, where ICPLSIZE=0 means the biggest hole possible (only I and E zones remain). The new frame zone F (grey, orange and blue in Fig.1) consists from the I+E zones plus the frame of ICPLSIZE grid points from each side taken from the central zone. Its size is equal to *w·(NDLON+NDGL-w)* for *w=2(NBZONL+ICPLSIZE)+11* where 11 is the width of E zone.

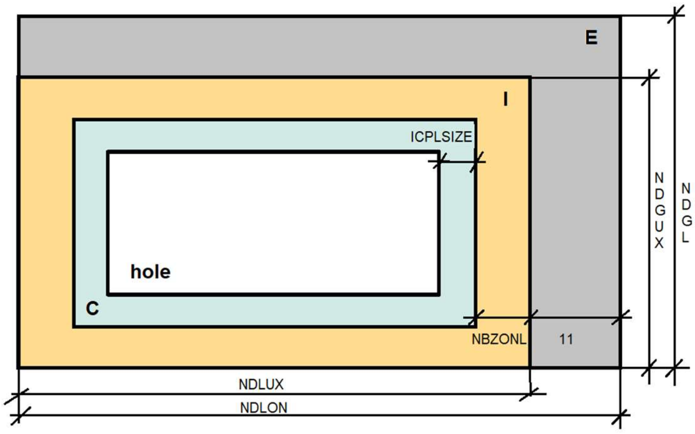

Figure 1 Scheme of coupling zones and sizes.

Work done during the previous stay

During the first stay, we prepared first the full LBC files in the fine resolution (2.3km) over the target domain (Czech operational) and in grid-point representation of all fields. Then we removed the inner part and filled it with artificial values. Then we compared the results of forecast run with LBC files prepared in the described way with those run with the reference coarse resolution LBC files. The correspondence of the forecasts was good, no degradation of its quality was detected even for ICPLSIZE=0, but the main target, which was the compression of the LBC files, was not reached. The reasons are twofold: first, the number of points in the coupling zone I in the fine resolution was big, and second the grid-point representation of fields in LBC files gives bigger file size then the spectral one used originally. As the consequence, the fine-scaled LBC file with framing had bigger size then the coarse-scaled reference LBC file. For further details see [GS].

The aim of this stay

The aim of the second stay was to check whether it is possible to frame LBC files in grid-point representation already at an earlier stage of processing, when they are still in coarse resolution. We would like to prepare the LBC files in grid-point representation, remove the inner part of the domain (so called *framing* or *hollowing*), transmit the LBC files to the final destination (a national meteorological service, NMS) and reconstruct fields in the whole domain in such a way that the verification scores of the forecast run using these LBC files are not worse than for the reference experiment with the original LBC files. Some challenge comes from the fact that the model variables in the ACCORD system are represented in spectral coefficients and hence global information is used when the coupling process is finished.

2. Methodology

Reference LBC files

We consider a LAM model application ALARO being run at a NMS for which the lateral boundary conditions are furnished by the global model ARPEGE and produced in Météo France (MF) over the LACE telecom domain (DOM1, see Tab.1) with 8.0 km horizontal resolution, quadratic truncation and 105 vertical levels. The LBC files prepared at MF are downloaded to NMS. We further consider the current Czech operational setting in which the LBC files from ARPEGE are operationally processed to the final domain (DOM3) with 2.3 km horizontal resolution, linear truncation and 87 vertical levels.

Table 1 Geometry characteristics of the domains used during the experiment.

domain	NDLON x NDGL	horizontal resolution	number of vertical levels	spectral truncation	size/2 of I zone (NBZONL)
DOM1	384 x 360	8.0 km	105	Quadratic	8
DOM2	324 x 270	8.0 km	105	- (all upper-air fields are GP)	8
DOM3	1080 x 864	2.3 km	87	Linear	16

Experimental LBC files

Since the telecom domain is much bigger then the target Czech operational domain, the first step in our modification process was to run e927 configuration to crop LBC files fetched from MF (DOM1) roughly to the size of DOM3 without changing its horizontal resolution. Obviously, the new domain (DOM2) must completely cover DOM3, so in fact its area needs to be slightly bigger. We chose its size to be 324x270. The domain extends a bit more to the south and north. Next, we transform the fields from spectral to grid-point representation and rename SPECSURFGEOPOTEN with SURFGEOPOTEN using fa-tools available in CHMI. After that, we apply another fa-tool called *fa_zero*, which allows us to make a hole in the fields and fill it with a constant value. Only upper air fields were hollowed since they are the subject of coupling. The other fields which are needed for physics, surface scheme or other purposes as orography, roughness etc. are kept unchanged. Initially we set ICPLSIZE=0, however it turned out that due to different grid-point width of I zone, the northern and southern edge of the hole was out of the hole made in DOM3 (Fig. 2 and 3). Therefore, we had to increase the size of ICPLSIZE from the north and the south to 3 gridpoints. So finally, we made a hole within DOM2 which extends 8-gridpoint from W and E and 8+3 gridpoints from N and S (counting from domain edges). At this stage, the files are ready to be compressed and sent to the national NWP center. Compression experiments and their results are described in chapter 5.

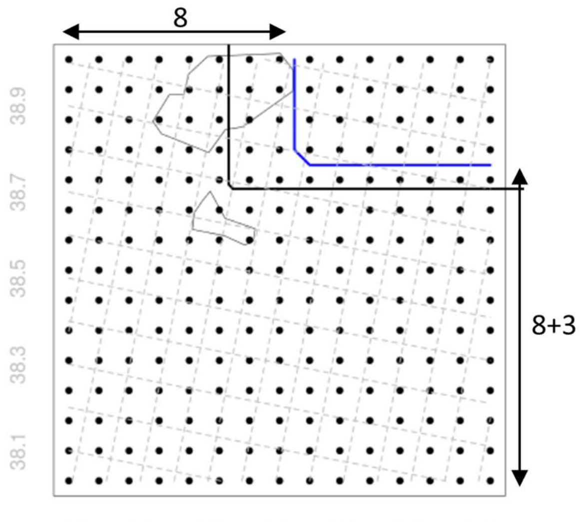

Figure 2: The SW corner of the DOM2 domain. Thick lines represent the edge of the hole in DOM2 (blue) compared to the hole in DOM3 (black). Solid grey lines represent two of the Balearic Islands (Ibiza and Formentara). Axis are scaled in degrees of latitude and longitude.

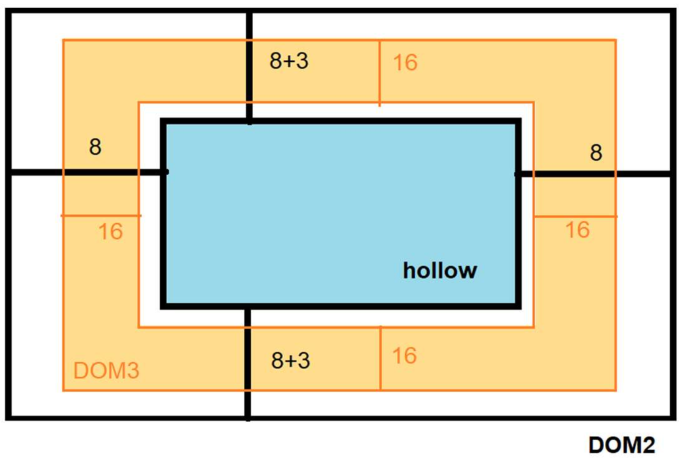

Figure 3: The internal zone of DOM3 (orange) must be inside the frame zone of DOM2 (black-white). The extension zones are omitted here.

However, we know from previous experiments that filling a hole with a constant value will cause problems when producing a forecast. Hence, the next step in preparation of LBC files is application of fa_zero once again, this time to fill the hole with an inverse distance-weighted interpolated values from inner edges of a frame (Fig. 4). This was inspired by the solution used in the routine *fadcpl.F90* within the model code. Having smooth fields, we run e927 once again – this time to downscale hollow LBC files to 2.3 km and 87 vertical levels, i.e. to DOM3. Necessary modifications in the namelist which allow reading input files in GP representation are:

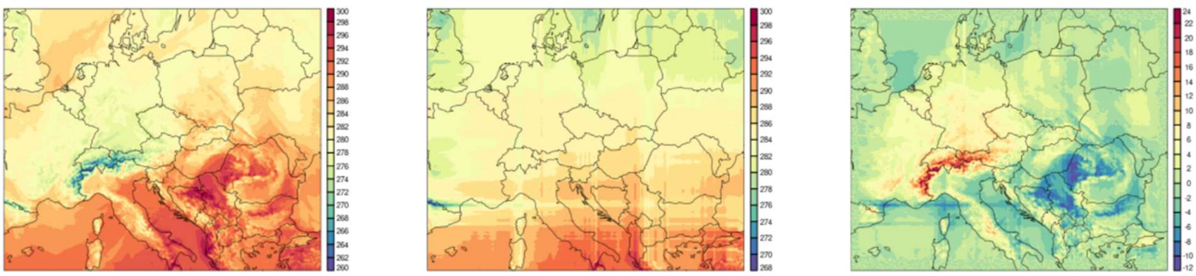

Figure 4: S105TEMPERATURE in a reference (left) and a hollow LBC file where the inner part is filled with the interpolated values (center). On the right is the difference between them.

LSUSPECA_GP=.T. and LSUSPECA_GP_UV =.T. in NAMCT0. Controlling the type of output is done by LWRSPECA_GP. Setting it to the default value (which is FALSE) produces all upper-air variables in spectral representation. LBC files are now ready to be used in forecast calculation. The whole procedure is graphically summarized in Fig. 5.

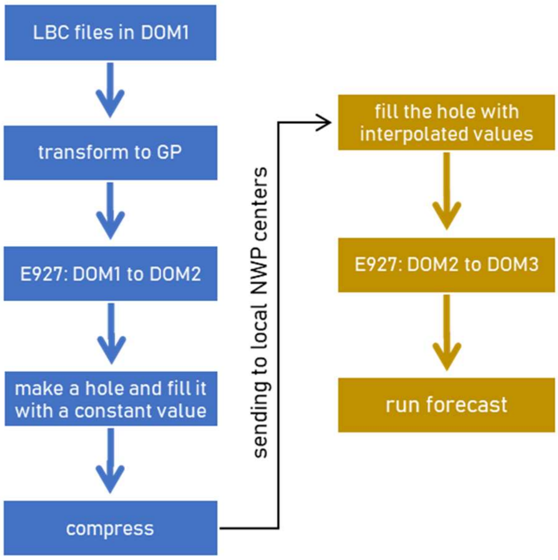

Figure 5: Steps of preparation of LBC files.

3. Experiments

Description of experiments

Based on modified LBC files, forecasts were produced and then compared with the reference ones, which were based on the reference LBC files. Coupling frequency was set to 1h. The forecast length is 72 hours. We run from 00 UTC. Initial conditions created with the Czech operational application through a data assimilation loop in 2.3km resolution were used. Other settings and the model version correspond to the Czech operational application (from November 2022) as well: non-hydrostatic dynamics with SL horizontal diffusion and PC time scheme with one iteration, ALARO-1 physical package with TOUCANS, ACRANEB-2 and 3MT, parametrizations of deep convection and GWD are not used. The time step is 90s. The model version is based on CY43t2.

Verification method

Verification spans a period from 1st to 14th November 2021 (14 days). The period was chosen since it involves a low pressure system passing through the central part of the domain on the 4th November 2021. Forecasts were evaluated point-wise against observations using objective scores (e.g. bias, RMSE, standard deviation). For that purpose, we use a dedicated piece of software called VERAL.

Verification results

Verification results are depicted in Fig. 6-9. The metrics are pretty much alike. There are some small differences in case of precipitation (Fig. 6) as well as wind direction and wind speed (Fig. 7), however they seem to be random, not showing a systematic degradation of results.

In addition to mean value of accuracy metrics, we "visually" checked spatial distribution of bias for some parameters (temperature, wind components and precipitation) at different model levels. Two examples are shown in Fig. 10. Some differences occur mainly within the front zones that are near centers of low pressure systems (e.g. in the central Italy and at the North Sea), however they are local in space and thus do not affect the overall accuracy of the forecast measured by the verification scores.

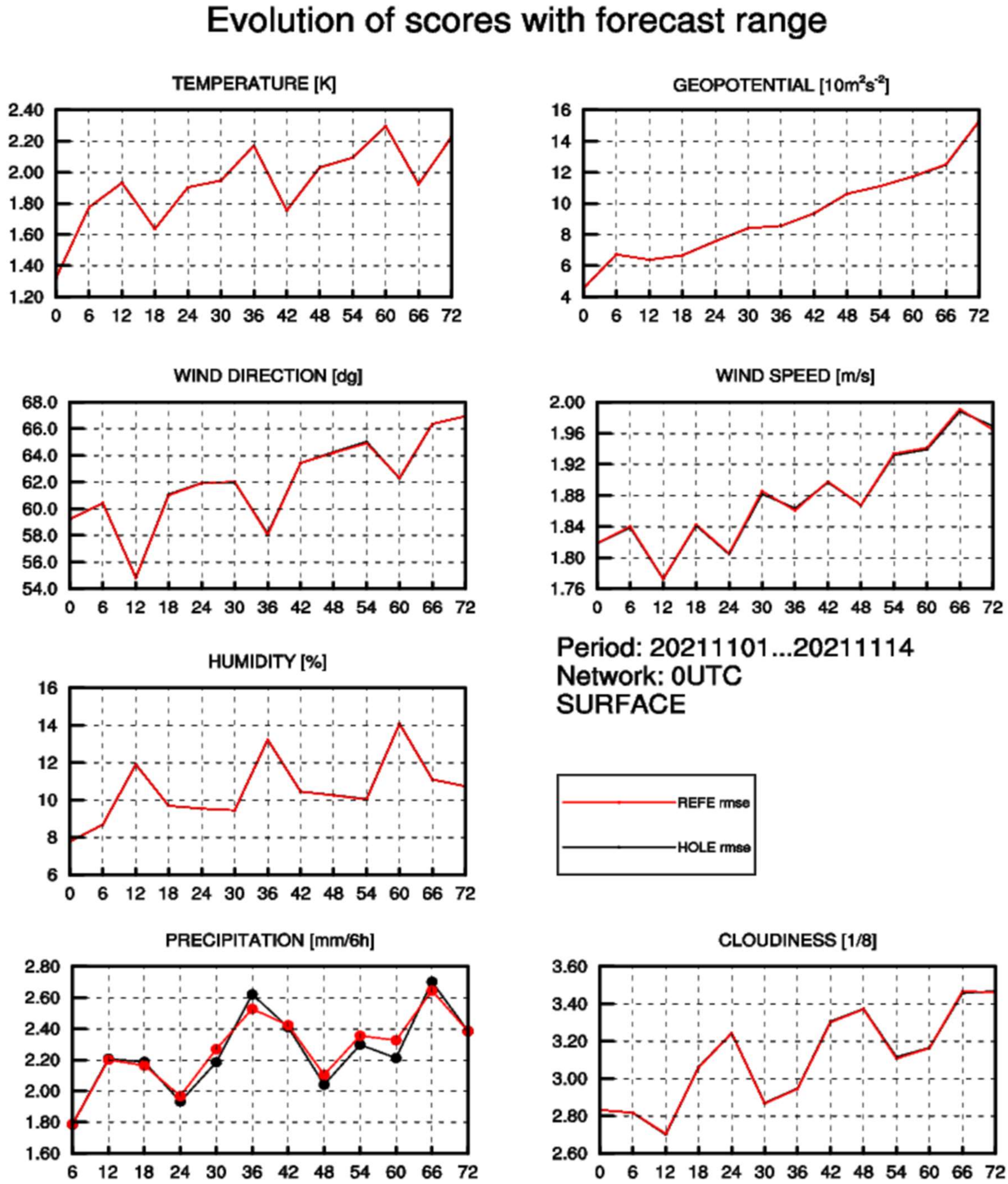

Figure 6: Temporal evolution of the forecast RMSE for various meteorological fields on the surface. REFE (red) means the reference forecast, HOLE (black) is the forecast produced from hollow LBC files. The leadtime is on the x-axis.

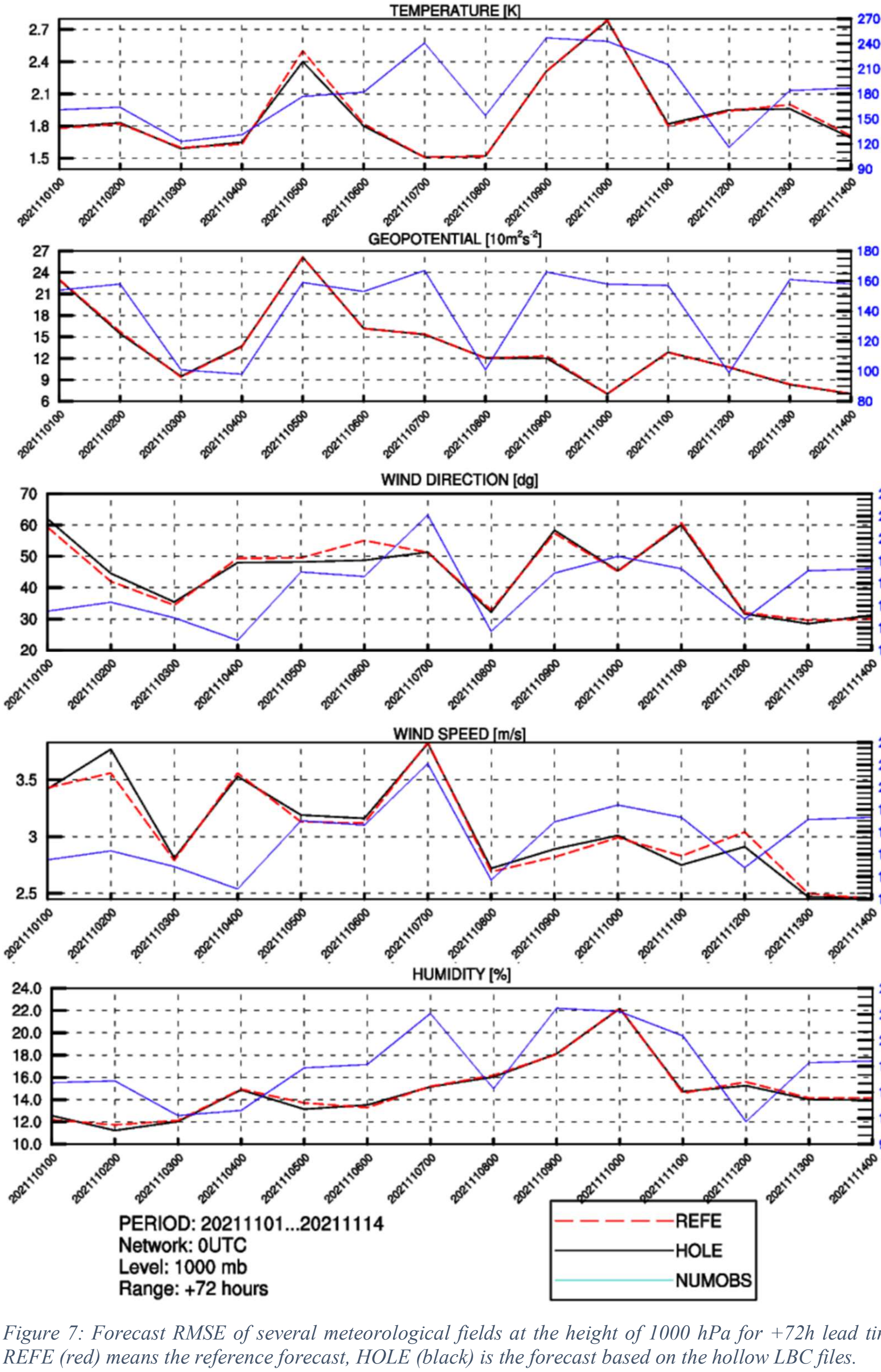

Figure 7: Forecast RMSE of several meteorological fields at the height of 1000 hPa for +72h lead time. REFE (red) means the reference forecast, HOLE (black) is the forecast based on the hollow LBC files.

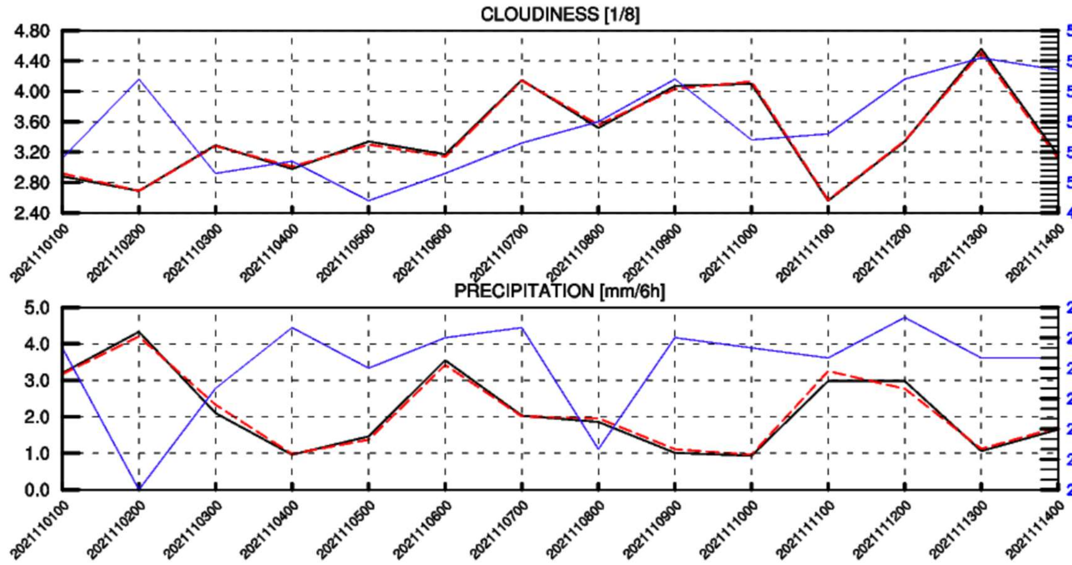

Figure 8: The same as Fig. 7, but for cloudiness and precipitation on the surface.

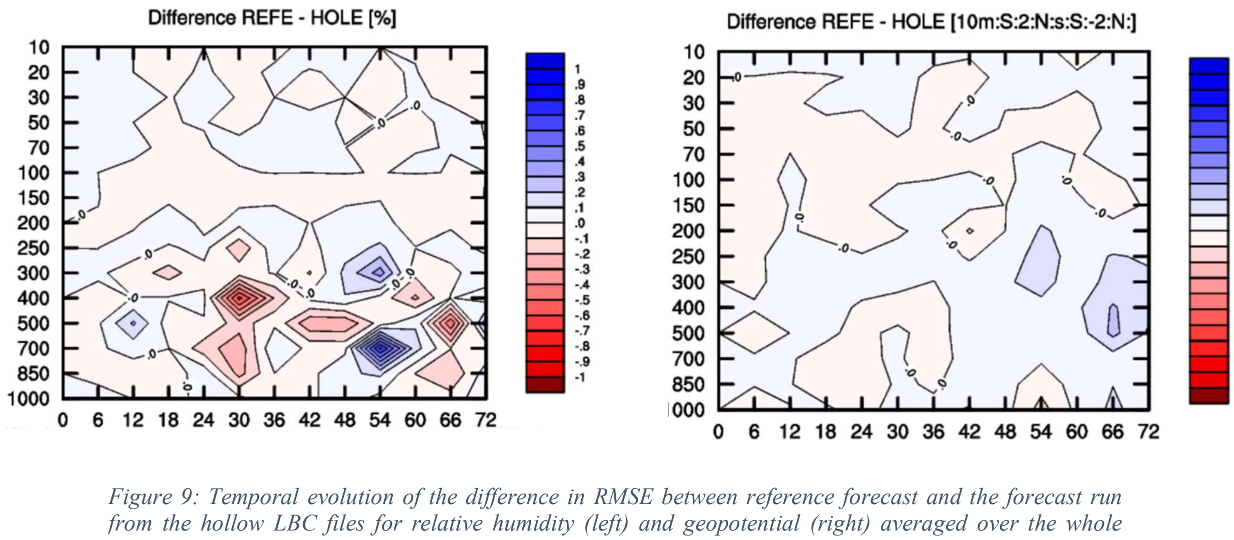

Figure 9: Temporal evolution of the difference in RMSE between reference forecast and the forecast run from the hollow LBC files for relative humidity (left) and geopotential (right) averaged over the whole horizontal domain. Vertical pressure levels are on the y-axis, the lead time on the x-axis.

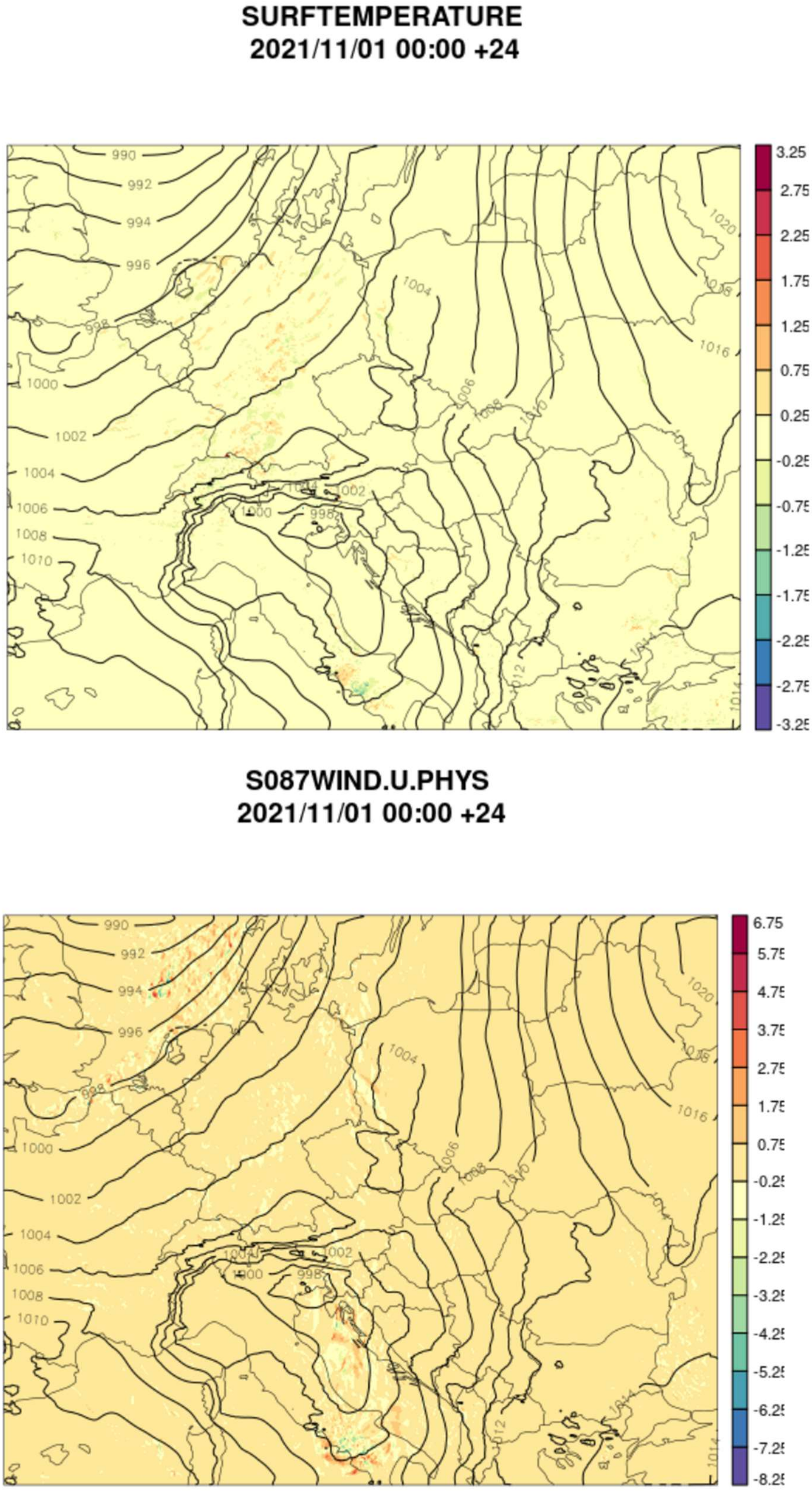

Figure 10 The difference between the reference forecast and the forecast run from the hollow LBC files for the surface air temperature (top) and the zonal wind component (bottom). The isolines show MSLP (hPa).

4. Compression

Size reduction estimate

We first estimate the potential size reduction obtained by hollowing an array. The size of an array is calculated as number of reals saved in a FA file according to formulas in Tab. 2. At this stage, we do not take packing into account, keeping its default value (NFPGRIB=2 in the namelist NAMFPC). The calculations are done for one horizontal array, which is one vertical level of an upper air field (Tab. 3). If we compare orange-colored values [1,2] we can see, that the size of a hollow field is only 36% of the reference one. It means that the size is reduced by 64%. Taking into consideration the fact that 37 surface fields are not the subject of hollowing, the reduction for the whole file is slightly lower (58,5%), but still significant.

*Table 2 The size of a spectral and a grid-point array in different LBC files. The size of the extension zone is 11. Notations: Na – NMSMAX, Nb - NSMAX, a – NDLON, b – NDGL, z=11+2*NBZONL (27 in our case).*

	Reference file	Hollow file
spectral array	$\pi N_a N_b = \pi \frac{(a-3)(b-3)}{9}$	-
grid-point array	ab	$a(z+6) + bz - z(z+6)$

Table 3 The size of a spectral and grid-point array in different LBC files given as the number of reals saved.

	Reference file	Hollow file
Na	127	107
Nb	119	89
NBZONL	8	8
a	384	324
b	360	270
spectral array	47 479 [2]	-
grid-point array	138 240	17 091 [1]

Compression methods

We made an estimate of the size of a hollow array. But with the methods used so far (default packing) the size of the LBC files is not reduced. We need to choose a compression method to really obtain a smaller FA file. We adopt two methods for reducing size of a hollow LBC file. The first one uses different packing options available in fullpos namelist under NFPGRIB parameter. It should be mentioned that as fullpos was not sensitive to changing only NFPGRIB value, we had to change also NVGRIB parameter, which belongs to NAMFA namelist. For testing purposes, we made a hole only in one field S001TEMPERATURE and filled it with a constant value T=200K. The file was then processed by fullpos, but no change in domain geometry was made – we tested only various NFPGRIB/NVGRIB parameters. By adding TFP_[X]%LLGP=T in the namelist NAMAFN, we forced fullpos to produce output in gridpoint representation. So basically, we intended to obtain on output the same array which was on input, but lesser in size. However, we obtained on output some noise near the inner border of the frame. The noise occurred even when no size compression was applied (NFPGRIB=2). We found out that the problem does not exist for surface variables, so we did a trick which involved manually renaming S001TEMPERATURE to SURFTEMPERATURE. After that modification, fullpos output is correct. A possible explanation for this might be that GMV model variables are transformed to spectral space even if only fullpos is run, while GFL variables (surface fields) are not. We realize that this trick cannot be applied as an operational solution, but we did it just to carry out an experiment with different packing options.

Another method we tried was an external compression tool available commonly in Linux environment called *gzip*. The experiment was made on a single-field file, extracted from the main file using fortran-based tool available at CHMI. In this case, the input size is simply the size of this single-field file, while the output size is the size of the compressed file .gz.

Compression results

We mention first the results obtained with different packing methods available in the fullpos. It turned out that applying level 2 second order-packing (NFPGRIB=140) results in a compression by a factor of 2.2 (Tab. 4) compared to the default (NFPGRIB=2). Gzip was able to compress it even more - by around a factor of 7.6. It offers various levels of compression, however the differences are in the order of decimal parts of percent. It should be added that gzip can perform that successfully only if the hole is filled with a constant value. For a full file, the compression is almost negligible (around 1.04) . For this reason, in order to fully benefit from the above described size reduction, the step of filling hole with constant value is crucial and cannot be skipped. In Table 4 we express the compression in percents of the size of the original single-field LBC file.

Table 4 Comparison of the total file size for two different packing methods.

	input size	output size	compression [%]
fullpos with NFPGRIB=140	270 336 B	122 880 B	54,5%
gzip (with default compression level)	270 336 B	35 612 B	86,8%

5. Summary

To apply the frame approach on the coarse resolution LBC files is a method which allows for possibility to transfer smaller size LBC files while keeping the same level of precision of the coupling and therefore appears to be an effective solution. Obviously, national weather services can benefit from it only if the process of framing will be carried out on the side of the partner who prepares the global forecast and produces the LBC input (MF) already before the transfer. What should be also taken into consideration at this stage is that every national weather services has its own local domain. So the process of preparing and transferring LBC files would have to be more “personalized”. Having received the framed files, national weather services would fill the hole with interpolated values from the inner edges of the frame. After that, LBC files would be ready to be downscaled to a finer resolution with e927 configuration and processed further as usual.

No progress on this subject would have been made, had it not been for my supervisor Petra Smolíková, whom I would like to thank for great support and guidance. I am also grateful to other colleagues from CHMI for tips they gave me.

References

[GS] G. Stachura, Frame approach in coupling, RC LACE, report from the stay, 2021

[MF] Hollow and core coupling with AROME, report of Météo-France, available upon request