SPECIAL PROJECT FINAL REPORT

This template is available at:
http://www.ecmwf.int/en/computing/access-computing-facilities/forms

All the following mandatory information needs to be provided.

Project Title:	SPHERA – PRE (Special Project: High rEsolution ReAnalysis over Italy – PREliminary study)
Computer Project Account:	spitcere
Start Year - End Year :	2017 - 2017
Principal Investigator(s)	Ines Cerenzia, Tiziana Paccagnella
Affiliation/Address:	ARPAE Emilia-Romagna, SIMC, Viale Silvani, 6 Bologna, Italy
Other Researchers (Name/Affiliation):	Davide Cesari (ARPAE Emilia-Romagna, SIMC), Andrea Montani (ARPAE Emilia-Romagna, SIMC), Chiara Marsigli (ARPAE Emilia-Romagna, SIMC)

The following should cover the entire project duration.

Summary of project objectives
(10 lines max)

This late special project aimed at accomplishing the preliminary tests necessary for the development of SPHERA, the high resolution regional reanalysis over Italy, based on the COSMO model, under development at ARPAE Emila-Romangna, SIMC. Another Special Project (named SPHERA) is ongoing and it deals with the second part of the preliminary tests, the production and evaluation of the archive. Objective of the SPHERA-PRE project were: (i) the setup of the SPHERA production suite on the ECMWF supercomputer (HPC), (ii) some first tests aimed at identifying a suitable model configuration for the reanalysis production (with respect e.g. to the nesting modality into the forcing model, to the bottom boundary condition to assign to the soil, to the dataset of observations to assimilate), and (iii) the development of a verification tool

Summary of problems encountered
(If you encountered any problems of a more technical nature, please describe them here.)

Some of the preliminary steps for the definition of the proper SPHERA setup had been underestimated (e.g. the development of the verification tool box was slower than expected due to technical issues with the verification code). Moreover, the performed preliminary tests proceeded slower than originally foreseen. Indeed, the amount of time needed to perform the model integration had been underestimated at the time of the project submission. The data assimilation was erroneously not active and the elaboration of the observations required for the assimilation was not considered in the suite used for the estimation. Because of these are time-consuming steps, the actual time needed to perform a 24h run was about 50% higher than estimated.

Experience with the Special Project framework
(Please let us know about your experience with administrative aspects like the application procedure, progress reporting etc.)

We are fine with it.

Summary of results
(This section should comprise up to 10 pages and can be replaced by a short summary plus an existing scientific report on the project.)

Introduction

ARPAE Emilia-Romagna, SIMC is developing a high resolution atmospheric regional reanalysis over Italy, SPHERA, performed with the COSMO non-hydrostatic Limited Area Model. COSMO is developed in the framework of the COSMO (COnsortium for Small scale MOdelling, Schättler et al., 2011)
consortium cooperation. It is used in the operational NWP suites in Italy, as well as in several other ECMWF Member States (Switzerland, Germany, Greece) and Co-operating States (Romania, Israel).

SPHERA is performed by means of a dynamical downscaling of a global reanalysis and by employing observational nudging during the model integration. SPHERA will cover 25 years and will produce three-dimensional hourly model output.

At the time of submitting this project, the idea was to feed SPHERA with the initial and boundary conditions from COSMO-REA6 reanalysis archive: a regional reanalysis dataset covering Europe with a 6km resolution, based on the COSMO model and forced by Era-Interim. However, at the first stage of the special project, it was decided to force SPHERA with ERA5, the global reanalysis currently under production at ECMWF. The intent was to provide SPHERA with a more complete, accurate and up-to-date set of initial and boundary conditions . It was hypothesised that ERA5 could provide more accurate information than Era-Interim (up-to-date IFS code, newly reprocessed observation dataset that could not be ingested in Era-Interim, 31km horizontal resolution, hourly output) and even more precise and consistent than a regional reanalysis archive based on Era-Interim (COSMO-REA6 is based on a COSMO version of 2012). Furthermore, the timetable of ERA5 production was quite coherent with the one of the SPHERA production. Therefore, the activity was in part reviewed in order to follow this new project development.

The results are divided among the three main branches of this late special project.

1) Set up of the SPHERA production suite

This section aimed at developing an operational automatic system for the reanalysis production to run on the ECMWF supercomputer (HPC). For simplicity, the simulation was structured in 24h-long runs. Essentially, the steps performed in the suite were:

1. extract the observation from MARS and elaborate them for the ingestion in COSMO
2. extract the initial and boundary conditions (coming from ERA5) from MARS
3. run the pre-processing model to COSMO (named INT2LM)
4. elaborate the INT2LM output in order to setup the warm initialization of the atmosphere, the soil, the lakes and the snow, as well as to configure the daily refresh and interpolation of the sea surface temperature from ERA5
5. run the COSMO model
6. post-process and compress the COSMO output for the storage on ECFS
7. transfer and storage on ECFS

Apart from the COSMO simulation, the points 2, 6 and 7 required a large amount of time to be completed for each 24h simulation. Therefore, specific suites were developed for each of these points and run in parallel to the main suite (that included all the other points).
Despite this reduction, the time period needed to accomplish a 24h run resulted approximately 50% higher than the one specified in the project request. This was because of the lack of some time-consuming passages in the script used for the estimate: indeed the data assimilation was erroneously not active and the observation elaboration was not considered at the initial stage. The SBU resulted higher than the estimated ones as well, by a value of 16%.
The system of routines was managed by the ECFLOW package.

2) Preliminary tests for SPHERA

The most relevant test (in term of computational time consumption) built up and partially performed during SPHERA-PRE regarded the selection of the modality by which COSMO is nested into the driving dataset ERA5. As general practice, high resolution runs are nested in coarser resolution integration of the same model, in order to ensure a ratio of spatial resolution between 2:1 and 5:1 (e.g. Warner et al. 1997, Denis et al. 2001). However, some recent studies (Marsigli et al. 2013) and experiences in the operational chain building-up (Arpagaus, MeteoSwiss, pers. comm.) demonstrated a neutral or improved performance of the high resolution run, when the intermediate step with the coarser resolution model was avoided. One of the potential reasons is that the coarser resolution model generally runs in the "grey zone", where deep convection is partially resolved and partially subgrid. By avoiding this intermediate step, the deep convection process is integrally resolved by the high resolution model without any additional perturbation.

Regarding this choice, two options had been considered for SPHERA:

- 2step: COSMO-I2 was one-way nested in COSMO-10M (a COSMO model configuration with horizontal resolution of 10km, domain covering the whole Mediterranean Sea and convection parameterized by Tiedtke scheme, Tiedtke, 1989), which in turn was one-way nested in ERA5. The ratios of spatial resolutions between COSMO-I2, COSMO-10M and ERA5 were respectively 5:1and 3 :1, in agreement with the traditional practice.

- 1step: COSMO-I2 was directly one-way nested in ERA5, with a resolution step of 15:1.

These two configurations were tested on two parallel suites over one year (2015), plus 6 months of initialization used to spin up the model soil fields. Their comparison was part of the SPHERA special project and the results are reported in detail in the respective report. However, it can be anticipated that the 1step configuration shows a clear improvement over the 2step case, in term of precipitation and temperature at 2m.

A second test regarded the type of observations to introduce in the assimilation procedure. Observations are downloaded from MARS database. The final list of data type that has been identified was: SYNOP, SHIP, AIREP, PILOT and TEMP. It resulted that the METAR reports caused a number of model crashes in several occasions. Therefore, METAR reports were always neglected. Sporadically, also the PILOT data caused a model crash. Therefore, it has been introduced a security check that removes these data and re-launches the model, any time the COSMO model crashed.

A third test dealt with the bottom boundary condition to assign to the soil, in particular for temperature. In general, the deep soil temperature in the regional reanalysis is interpolated from the deepest temperature of the driver model. However, while the deepest soil level in ERA5 is located at 1.954m of depth, the lowest one in COSMO goes down up to 14.58m. Therefore, the direct interpolation of ERA5 deep temperature in SPHERA would cause an error in the amplitude of the seasonal temperature variation (which is dumped by depth) and in the temporal shift of the surface temperature signal (which is delayed by depth). The question is reported here for completeness, but the main part of activity has been performed within the SPHERA special project. Details about it are reported in the SPHERA report.

3) Development of a verification tool box

Three types of quality checks of the reanalysis performance have been implemented:
1. verification of the surface variables against observations
2. control of the observations ingested by data assimilation
3. control of the time trend of the surface variables

In the following they are described in detail.

Verification of the surface variables against observations

Precipitation:
The verification of precipitation is performed by aggregating the model and observation data on 0.2°x0.2° boxes and comparing the maximum of the daily accumulated precipitation. The contingency table scores are used to plot the performance diagrams (Figure 1). Several precipitation thresholds are applied to discriminate among the rainfall intensity. Observations are extracted from

DEWETRA platform, the monitoring tool of the Italian Civil Protection (http://www.mydewetra.org/)

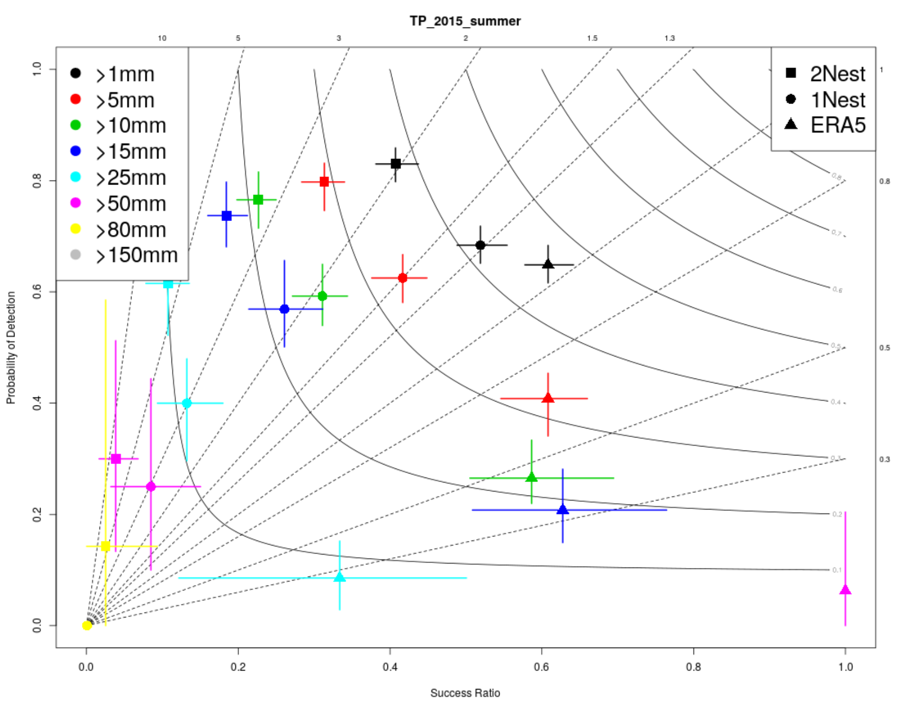

Figure 1. Performance diagram of the daily accumulated precipitation at different thresholds aggregated over 0.2°x0.2° boxes for SPHERA-1step (labelled "1Nest"), SPHERA-2step (labelled "2Nest") and ERA5.

Temperature at 2m:

The verification of 2m temperature is performed using the nearest point method and calculating the bias and RMSE scores. The procedure is applied on a three-hourly subset of the archive. Height altitude correction has been performed if the difference between model grid point and observation was lower than 500m, while the data was discarded if the delta was larger than 500m. Observations include data from the high resolution monitoring network (which densely covers Northern Italy) and from the SYNOP stations. These observations have not be ingested by the data assimilation procedure. Some examples of the performed plots are reported in Figure 2

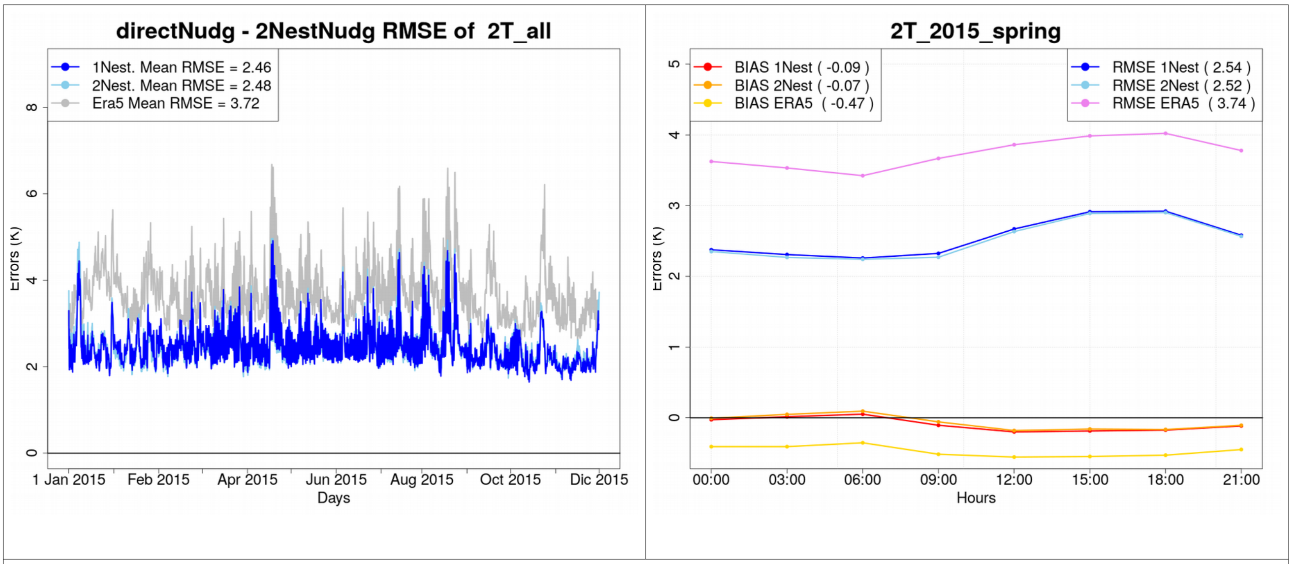

Figure 2. RMSE of tempearature at 2m in SPHERA-1step (labelled "1Nest"), SPHERA-2step (labelled "2Nest") and ERA5 along time (left) and averaged on the time range for the spring months 2015 (right).

Control of the observations ingested by data assimilation

June 2018

The number of observations ingested by data assimilation and their status (active, passive, rejected, neglected) are valuable indicators of the correct performance of the data assimilation scheme and in general of the simulation itself. Indeed a constant or incrementing number of data assimilated with status active (in proportion to the total number) along months shows that the reanalysis gets close to the observations, which is a desirable behaviour. Vice-versa, a reduction of the number of ingested data with status active can indicate that the background is deviating from the observation space, or, especially if the reduction is sudden, a problem in the assimilation scheme. An example of this type of monitoring plot is reported in Figure 3.

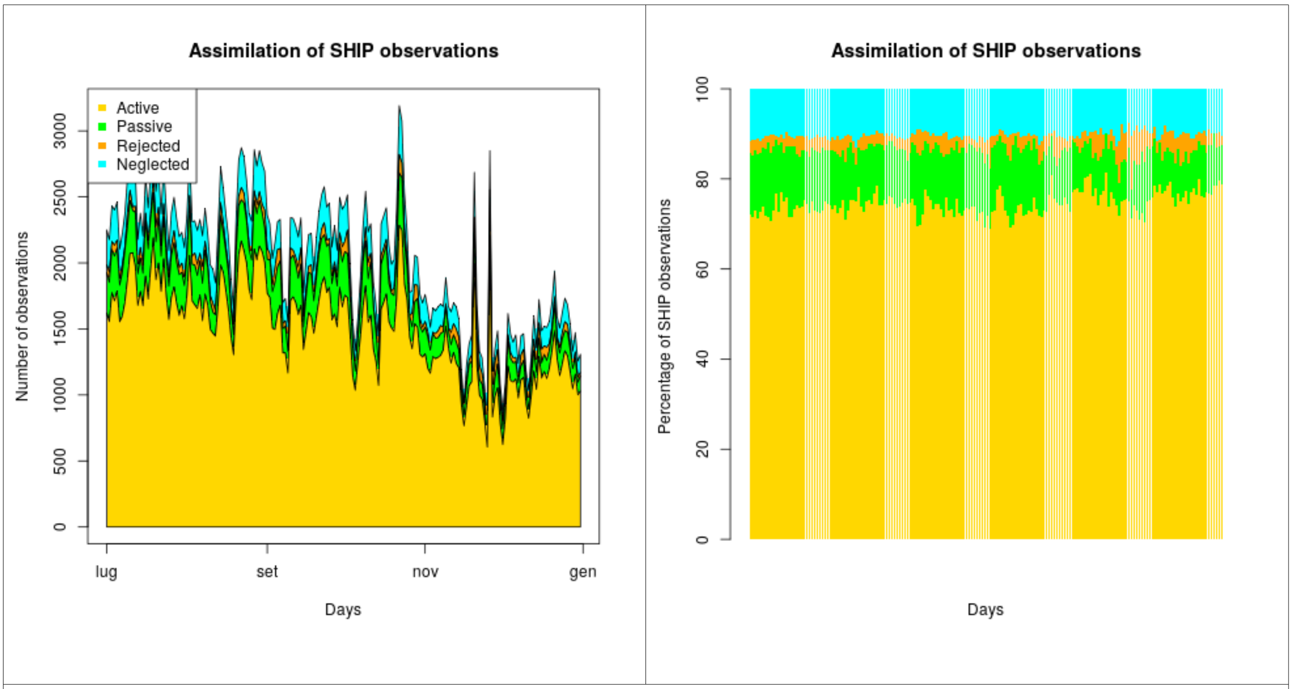

Figure 3. Number of SHIP observations ingested in the period July-December 2015 discriminated among the assimilation status (left) and the proportion of the number of data with a specific status to the total SHIP data (right).

<u>Control of the time trend of the surface variables</u>

The domain average of some surface variables (daily accumulated precipitation, maximum and minimum temperature at 2m, mean sea level pressure and surface pressure) is plotted against time in order to pinpoint time trends over the years or anomalous deviations along time. An example of this type of monitoring plot is reported in Figure 4.

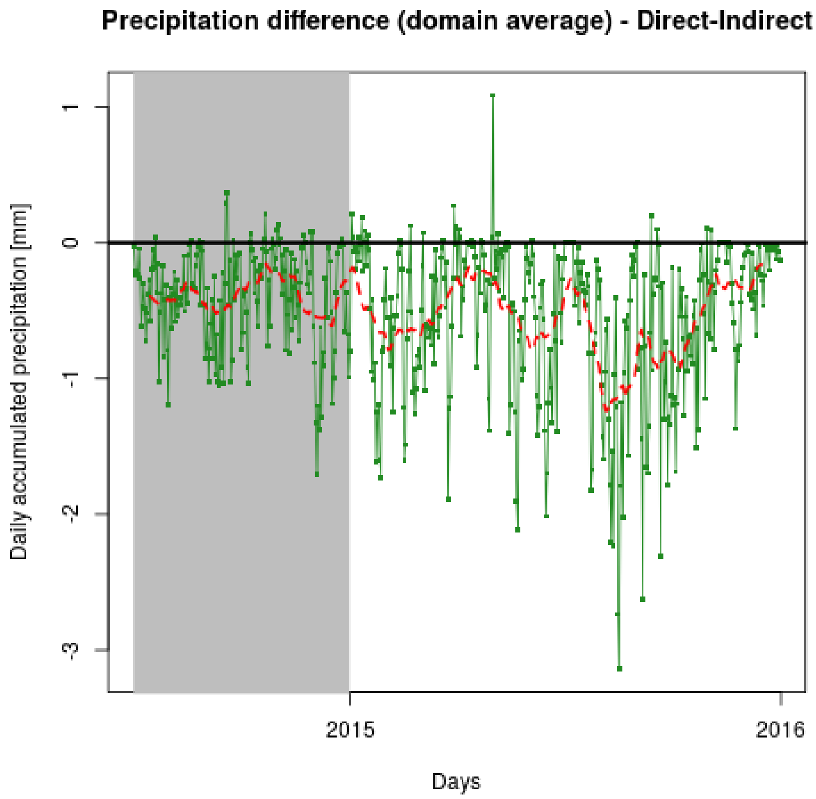

Figure 4. Difference of the domain-averaged daily accumulated precipitation simulated by SPHERA-1step and SPHERA 2-step, plotted along the full integration time. The shaded box indicates the spin-up period.

Conclusions

This late special project achieved to develop the required technical tools for the SPHERA production: i.e. the HPC suite and the verification tool box. Moreover, through this project it was possible to test and define a suitable configuration of COSMO (Table1), despite few points still need to be defined. The first one regards the nesting modality into ERA5. Part of the experiment that tackles this issue, has been performed during SPHERA-PRE. The remaining part of it, as well as the second point to be defined (deep soil temperature), have been performed within SPHERA special project.

	SPHERA setup
Initial condition	ERA5
Boundary condition	ERA5, updated every hour
Nesting modality	To be defined (see SPHERA report)
Sea Surface Temperature	Interpolated from ERA5 every day
Deep soil temperature	To be defined (see SPHERA report)
Observations assimilated	SYNOP (not temperature at 2m and precipitation), SHIP (not temperature at 2m and precipitation), TEMP, PILOT and AIREP
Code version	INT2LM 2.04 (pre-processing) COSMO 5.03 in single precision (to be upgraded to COSMO 5.05)
Domain	38N, 5.7W- 53N,18.2W
Resolution	2.2km horizontal, 65vertical levels (0-22km), 7 soil level (0-14.58m)
Physical schemes:	
Radiation	δ two-stream scheme after Ritter and Geleyn, 1992

Turbulence	Prognostic turbulent kinetic energy closure at level 2.5 including effects from subgrid-scale condensation and from thermal circulation (Rashendorfer 2001)
Transfer	Surface layer scheme coupled with the turbulence scheme (Rashendorfer)
Land-Surface	Multi-layer soil after Jacobsen and Heise (1982)
Convection	Only shallow convection (reduced Tiedtke 1989)
Microphysics	Grid scale cloud and precipitation scheme (3 categories ice scheme) and a statistical scheme for sub-grid clouds (Sommeria and Deardorff, 1977)
Subgrid scale Orography	Lott and Miller, 1997
Lake	Two-layer bulk model after Mironov (2008)
External parameters	
Orography	GLOBE
Land cover	Global Landcover 2000 Database
Soil type	Digital Soil Map of the World (FAO/UNESCO)

Table 1. SPHERA setup.

List of publications/reports from the project with complete references

Poster with title “SPHERA: High rEsolution ReAnalysis over Italy. Plan and setup” (Ines Cerenzia , Tiziana Paccagnella, Andrea Montani, Arpae-Emilia Romagna, HydroMeteoClimate Service, Bologna, Italy) presented at the 5th International Conference on Reanalysis

Future plans

(Please let us know of any imminent plans regarding a continuation of this research activity, in particular if they are linked to another/new Special Project.)

This late special project dealt with the preliminary studies required for the development of the multi-decadal reanalysis archive SPHERA. Another special project (indeed named SPHERA) is ongoing and it will last up to 2020. Please refer to the section future plans of the report of SPHERA for more details.