Optimization of a calibration procedure for weather prediction model

Dr. Antigoni Voudouri, Hellenic National Meteorological Service (HNMS), Greece, PI
Dr. Jean-Marie Bettems, Federal Office of Meteorology and Climatology MeteoSwiss, Co-PI
Dr. Izthak Carmona, Israel Meteorological Service (IMS), Israel, Co-PI

Project team members:

Dr. Evripides Avgoustoglou, Hellenic National Meteorological Service (HNMS), Greece
Dr. Oliver Fuhrer, Federal Office of Meteorology and Climatology MeteoSwiss, Switzerland
Dr. Paola Mercogliano, CIRA-Italian Aerospace Research Center, Capua, Italy
Dr. Edoardo Bucchignani, CIRA-Italian Aerospace Research Center, Capua, Italy

Abstract

This proposal is a follow-up on the Production Project 'Objective calibration of weather prediction models' which took place from 01.04.2015 to 30.09.2016 on Piz Daint[1]. If accepted, this new project will use the same tools which have been developed and installed on Piz Daint in the course of the previous project.

All atmospheric models used for numerical weather prediction (NWP) and climate modeling have inherent uncertainties. Many of them stem from parameterization schemes used for the physical processes within the models, which often include free or poorly confined parameters. Model developers normally calibrate the values of these parameters manually in order to improve the agreement of forecasts with available observations. This 'expert tuning' is typically done once during the development of the model, for a certain target area, and for a certain model configuration, and is often difficult if not impossible to replicate.

A practicable objective multivariate calibration method build on a quadratic meta-model was introduced in our previous project, and successfully applied for the calibration of the COSMO-2 model[2]. The calibration has been performed with six parameters, with Switzerland as focus area, using a one year calibration period. The exceptional computational demand and the substantial complexities regarding the application of a meta-model for a large number of parameters have successfully been tackled leading to an overall improvement of the model performance.

However, due to many technical problems, as described in the final report of the previous project, not all the original goals have been achieved. **In particular, the effect of the soil memory has not been fully evaluated, and the issue of finding a compromise between the computational cost of the method and the quality of the calibration could not be tackled.** Considering that (1) the multivariate calibration method has a huge potential for multiple applications in the NWP community, e.g. when using the model for a different climate or when introducing new parameterizations, and (2) the fact that the previous project has shown that a quadratic meta-model is able to reproduce the dependency of the model quality on unconfined model parameters, the COSMO consortium[3] has decided to define a new priority project (CALMO-MAX) to tackle the remaining issues.

As a demonstration vehicle, we will use the state of the art kilometric configuration of the COSMO model, with unconstrained soil in order to activate the long-term memory of the NWP system. Such a configuration has already been run in the previous project, but only for one month, and without the most recent COSMO developments. Consequently, **a total of 780'000 node hours are requested on the hybrid Piz Daint system at CSCS to consolidate and optimize the objective NWP calibration method introduced in the previous project.**

---

[1] http://www.cosmo-model.org/content/tasks/priorityProjects/calmo/docs/CSCS_Proposal.pdf
[2] http://www.cosmo-model.org/content/tasks/priorityProjects/calmo/docs/CSCS_final_report.pdf,
http://www.cosmo-model.org/content/model/documentation/techReports/docs/techReport31.pdf
[3] http://www.cosmo-model.org/

The main benefits for the COSMO community of a successful CALMO-MAX project will be the provision of a permanent COSMO framework for objective model calibration. This will support the definition of an optimal configuration for the multiple production systems used within the COSMO community, including a possible focus on extreme events, and will also help define an optimal perturbation of the model parameters for the ensemble prediction systems.

The main risk of this project is that the method remains computationally too expensive for regular usage.

Background and significance

Numerical weather prediction (NWP) and climate models use parameterization schemes for physical processes which often include free or poorly confined parameters. Model developers normally calibrate the values of these parameters subjectively to improve the agreement of forecasts with available observations, a procedure referred as expert tuning. A practicable objective multi-variate calibration method has been developed by Bellprat et al. (2012) and implemented for a regional climate model; this method has been shown to be at least as good as expert tuning.

To evaluate the possibility to adapt this calibration method for NWP applications, a COSMO priority project has been defined and conducted (CALMO, 2013-2016)[4]. After a preliminary study on a coarse resolution grid, calibrating only three parameters, a more realistic configuration could be introduced thanks to the Piz Daint production project 'Objective calibration of weather prediction models'. The model resolution was increased from 7km to 2.2km, in addition to the enlargement of the domain size of the simulations, and Northern Italy was also included in the verification area (the same tools are used independently from the verification region, but different sets of observations have to be collected). Besides the daily accumulated precipitation, the daily 2m temperature extremum and the sounding profiles were added to the set of constraining observations. The simulation period was significantly increased from 40 days of 2008 to the entire year of 2013 (a one year simulation period consistently incorporates the weather development on a seasonal basis and should be considered a strong asset toward the operational value of this work). The number of calibrated parameters was increased from 3 to 6. Significant work was also invested in the quadratic meta-model: adding the support of atmospheric profiles, introducing an option not to average temperature extremum over regions, defining a new set of regions, adjusting the RMSE-type performance score and introducing a new COSI performance score (Damrath 2009), introducing a new method for logarithmic transformation of selected parameters, and developing an iterative method to obtain the optimal parameters via convergence in a n-dimensional parameter space of exceptional cardinality.

Following these adaptations, the calibration was performed and the optimal parameters combination was obtained. Using the COSI performance score to quantify the quality of the simulation, which is a combination of root mean square score for continuous fields and equitable thread score for precipitation, a small performance gain of 3-4% was observed. An independent verification of the optimal configuration showed a small reduction of the 2m temperature and precipitation biases, but also a small increase of the 2m dew point bias. This small impact is expected, given that the chosen model configuration is very similar to the model configuration used by the COSMO core development team, which has undergone exceptional expert tuning over a period of almost two decades; arguably this small impact confirms the validity of the calibration method. However, the main learning from the CALMO project is that the meta-model is able to reasonably reproduce the dependency of the model on the unconfined parameters. This is illustrated in Figure 1, where the meta-model prediction is compared to the full model prediction for the minimum daily temperature (left panel) and the daily accumulated precipitation (right panel) at different locations and for different days. To produce these correlations plots, an arbitrary combination of model parameters, not used for building the meta-model, was selected.

---

[4] The CALMO final report will soon be available. Two COSMO Technical Reports (2015 and 2017) and a paper in Atm. Res. (Voudouri et al., 2017) have already been published. A second peer reviewed paper is under preparation. See the references at the end of this document.

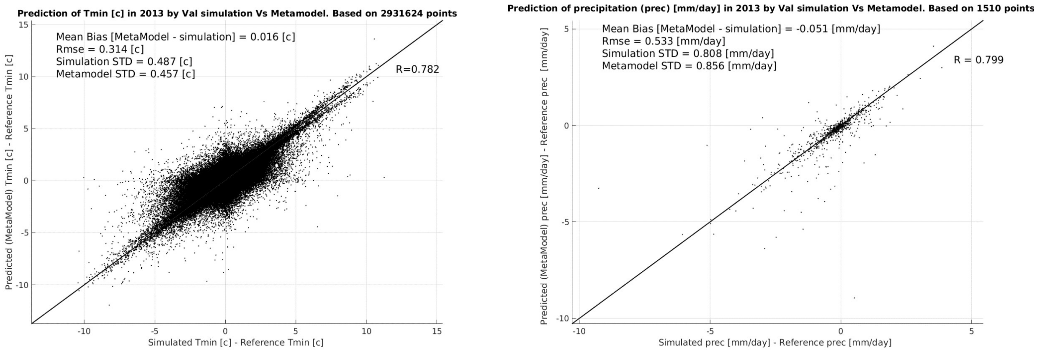

Figure 1: *(**Left panel**) Tmin meta-model prediction for a specific parameter combination, vs COSMO simulation results during the year 2013. X axis presents the simulated Tmin minus the reference simulation. Y axis presents the Meta-Model Tmin minus the reference simulation. Each point is for a different model grid point and for a different day. (**Right panel**) Same as left panel, for daily accumulated precipitation. Each point is for a different region and for a different day.*

Thanks to these developments, the calibration methodology can now readily be applied to a NWP system and the reliability of the calibration results can be trusted. **However, a full assessment of the impact of the soil memory is not available; this is an important issue, because it is expected that the impact of a new set of model parameters can be substantially stronger through the accumulation of heat and humidity in the soil over the full simulation period, as indeed observed in a preliminary experiment with a 1.1 km configuration of the COSMO model** (a performance gain measured by the COSI score exceeding 10% has been observed following a one month calibration). **Furthermore, in order for this method to be used by the COSMO community, it is essential to reduce the computing cost of the calibration**; for that purpose, having a full calibration without any compromises that will serve as basis to optimize the calibration methodology is a pre-requisite.

The proposed objective calibration methodology has the potential to bring a transformative change to atmospheric model development and significantly reduce model development cycle times. The calibration method will be a very useful tool to improve the quality of the multiple configurations of atmospheric models running in Europe and beyond. More specifically, the developed methodology could be used by each COSMO member to define an optimal calibration over the target area of interest, for re-calibration after major model changes (e.g. higher horizontal and / or vertical resolution), as well as for an unbiased assessment of different modules (e.g. parameterization schemes), and for optimal perturbation of parameters when run in ensemble mode. Last but not least, the implementation of the methodology for a specific parameter can clarify the impact of the specific parameter on the overall model performance. Once the meta-model has been fitted to the full COSMO NWP model, both the effect of the parameter setting and parameter space used (i.e., the maximal range of optimal values) can be determined without the use of the full NWP model.

Scientific goals and objectives

The main scientific goals and objectives of the proposed project are in-line with the new follow-up CALMO–MAX research project[5], which has been proposed within the COSMO Consortium and accepted in March 2017. Its main goals are:

- **Optimize the method to find a compromise between the quality of the calibration and the computer costs. Use one year COSMO-1 calibration, with the soil memory switched on, as test bed for this purpose.**
- Establish CALMO as a permanent optimization tool within COSMO. Create and maintain a database of unconfined model parameters.

---

[5] http://www.cosmo-model.org/content/tasks/workGroups/wg3b/docs/cmax_plan_v9.pdf

- Apply the method to different climatology (e.g. Mediterranean). Besides the strong interest of some COSMO partners for a targeted model optimization, this will be used to demonstrate the applicability of the permanent framework.
- Investigate and implement specific calibrations, like a calibration for extreme events or a season dependent calibration. Further consolidate and extend the meta-model and the global model score for this purpose.

The first goal of the CALMO-MAX project requires access to a full calibration, without any compromise, with a non-trivial set of optimum parameters differing as much as possible from the reference set. As a demonstration vehicle, it is planned to use the 1.1 km mesh-size COSMO model version, developed within the COSMO-NExT[6] project at MeteoSwiss, with active soil memory and computed over a full year. This requires a considerable amount of computing resources, and is the main reason to ask for Piz Daint access; in addition, all the tools having already been ported and tested on Piz Daint during the previous project, the settings and the computation of the experiment on this platform will be straightforward.

Once the calibration of the 1.1 km configuration of COSMO is done, the produced data will be used to test different strategies to reduce the cost of the method. Additionally, we will investigate the usage of the method for specific calibrations, like a calibration for extreme events or a season dependent calibration.

One should note that the Piz Daint platform is only requested for the first goal of the CALMO-MAX project. The HPC platform at ECMWF will be used to tackle the other goals.

Description of the research methods, algorithms, and code

In this section, we briefly introduce the code that we will use for the simulations and describe the experimental design for applying the objective calibration method described above.

The code used in this project is the limited area numerical weather prediction (NWP) model COSMO. To our knowledge, it is used as the main computational tool of several other production proposals on the Piz Daint system at CSCS. COSMO employs finite differences to solve the primitive hydro-thermodynamic equations that describe the non-hydrostatic compressible flow of the atmospheric constituents. It contains a comprehensive set of parameterization schemes to represent non-resolved processes. COSMO is suitable for running simulations with grid spacing ranging from O(100km) to O(1km). The COSMO model has been running for many years at CSCS, both for numerical weather prediction and regional climate research, on several generations of CSCS supercomputers. The refactored version of the COSMO model (Lapillonne and Fuhrer, 2013) capable of running on GPU-based hardware architectures (referred to as RC), already applied in our previous project ‘Objective calibration of weather prediction models’, will be used in this project. Note also that the latest release of this model has already been ported and tested on the Piz Daint system after the upgrade to the NVIDIA Tesla P100 graphics processing units.

The model setup will be based on the production configuration that MeteoSwiss is using for the model named COSMO-1. The model is computed on 1158 x 774 grid points in the horizontal, with a grid spacing of 1.1 km, and 80 vertical levels, and spans the greater Alpine region. The simulation domain and the associated topography are shown in Fig. 2. The calibration will be performed for 6 parameters. As learnt in the previous CALMO project, **a robust calibration requires at least [3*N + N * (N-1) / 2 + 1] repetitions of a single model experiment, where N is the number of parameters used, each experiment being characterized by a different set of model parameters**. Each experiment will be computed in hindcast mode for a full year, the model being only constrained by the lateral boundary conditions; the year 2013 has been chosen, because it is representative for a mean climatology over Europe.

---

[6] http://www.meteoswiss.admin.ch/home/research-and-cooperation/international-cooperation/cosmo.subpage.html/en/data/projects/2012/cosmo-next.html

The initial and boundary conditions for all experiments will be taken from the same operational analysis run at 2.2 km resolution. An important issue is the initialization of the soil, since multiple years are typically necessary for the state of the deep soil to adjust to a change in the model climate; for that reason, a spin-up run with a much cheaper standalone soil model will be computed before each calibration.

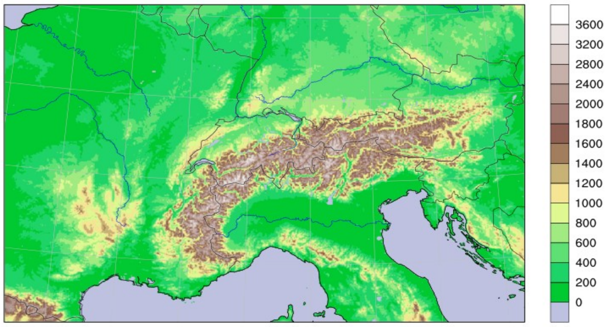

Figure 2: *Orography in meters above sea level and area covered by the 1158 x 774 grid point domain.*

Parallelization approach

The standard COSMO version employs a two-dimensional domain decomposition along the two horizontal directions with a pure MPI parallelization strategy. Inter-node communication is dominated by nearest neighbor communication between adjoining sub-domains in order to update ghost values required for the application of the finite difference operators, which is implemented using a MPI_ISend and MPI_Recv scheme. Communication between accelerators is done using G2G technology avoiding transfers from device to host and vice versa. The G2G transfers are implemented using a MPI_IRecv/MPI_ISend/MPI_Wait strategy and overlap computation with communication wherever possible. The only remaining collective communication patterns are either for computing scalar diagnostics or in the I/O part of the code in order to gather fields for writing to the storage subsystem. Fine-grain, on-node parallelization is done using CUDA threads in the horizontal dimensions, where a minimum of approximately 100 x 100 grid points per accelerator are required for efficient execution (see below).

Since the research in this project requires an ensemble of simulations with different tuning parameters, the ensemble members are completely independent and can be run in parallel. The number of nodes required by each ensemble member is determined by the scalability of the model and the required time-to-solution for the ensemble simulation.

Representative benchmarks and scaling

In the following sections, we present two types of performance benchmarks. First, we present performance measurements for the target application in order to estimate the required computational resources. Piz Daint is a hybrid Cray XC50 high performance computing system with one Intel Xeon E5-2690 v3 @2.6 GHz (12 cores) and one NVIDIA Tesla P100 with 16 GB of fast video memory per node. Second, we present benchmarking results from the old Cray XC30 (Piz Daint) using a representative COSMO simulation at 1.1 km horizontal resolution.

Scalability

The COSMO model has recently been computed on the full Piz Daint system, running a near-global benchmark, showing excellent weak scaling properties on up to 4800 GPU nodes. In terms of strong scalability, the main limiter is the amount of parallelism on a single GPU. Below a threshold of approximately 100 x 100 horizontal grid points, an insufficient amount of parallelism leads to an inefficient usage of the GPUs resources; further increasing the number of nodes to solve a problem of fixed size leads only to a limited decrease in the time-to-solution. This limits the number of GPUs that can be used to solve a problem with a fixed computational domain, such as COSMO-1. Since

COSMO-1 has 1158 x 774 grid points in the horizontal, using more than approximately 100 GPUs per simulation will reduce the parallel efficiency.

Full benchmarks

Table 1 shows the benchmark results on the Piz Daint hybrid nodes for the restructured code (RC) using both the CPU version of the code as well as the GPU version for a simulation representative for one ensemble member with a forecast time of +24 hours, including I/O. The domain size and the model configuration are representative for the runs we plan to execute using the v5.0 of the COSMO code. The code is run in single precision.

							GPU	CPU	GPU	CPU
Target	Nodes	nx	ny	dt(s)	#tracer	time(h)	Wall clock time (min)		Resources (node hours)	
COSMO-1fcst	144	1158	774	10	7	24	**9.1**	**25.1**	**21.93**	**60.20**
COSMO-1fcst	128	1158	774	10	7	24	**10.5**	**31.3**	**22.43**	**66.73**
COSMO-1fcst	96	1158	774	10	7	24	**17.9**	**54.8**	**28.61**	**87.67**

Table 1: *Benchmarks of COSMO-1 (the target configuration we plan to use for this project) on different node configurations for a 1 day forecast.*

For the production simulations of this project, the ratio between wall clock time and model time is of key importance, as it determines the time-to-solution for a given simulation period. In order to be able to calibrate the proposed COSMO-1 configuration within 8 months (see table 3 below), it is necessary to complete a 1 year long simulation within 1 month, assuming we can permanently run 6 ensemble members concurrently (the simulation plan contains an ensemble simulation with completely independent ensemble members) and accounting for 20% down time (maintenance, holidays). For technical reasons related to the COSMO model, a single ensemble member simulation has to be computed in chunks of 24 hour, and a new job has to be submitted, sequentially, for each simulation day; consequently, accounting for queue wait time and failure recovery, one estimates that a ratio between wall clock time and model time of at least 40 must be attained to complete a 1 year long simulation within 1 month. Thus a 24 hour simulation should have a wall clock time of 36 minutes or less. Table 1 shows that this threshold is achieved by all tested node configurations. Furthermore, Table 1 shows that GPU execution is significantly more efficient for all tested node configurations.

For the configuration using 96 nodes and running on GPU, we achieve a total of about 5 wall-clock days for a simulation of 1 model year, not accounting for waiting time between the runs. This is a sufficient time-to-solution for this project and we plan to use this setup for the production runs.

Performance analysis

Table 2 below shows the required performance analysis metrics for the benchmark and has been included for completeness. It should be noted, that some of these performance metrics only make limited sense in the case of GPU execution, but we have included them to comply with CSCS' requirements for production proposals. The benchmark corresponds to a COSMO configuration at 1.1 km resolution integrated for 24 hours with activated performance profiling using CrayPat (pat_build –g mpi,io). The results were obtained on old Piz Daint with the restructured code, on CPU and on GPU, for the target configuration, using 8x8 nodes (see above).

As is typical for weather applications, applying finite difference discretizations for the solution of the underlying governing equations, per process memory usage is very low. MPI communication is dominated by nearest neighbor communication (MPI_Recv) and synchronization. The comparatively large synchronization times, stem from inherent imbalances in the model, mostly caused by modules of differing cost being applied to sea/land grid points and/or cloudy/cloud free regions. The floating point efficiency depends on several factors (local domain size as compared to cache size, vectorization along the first horizontal direction, etc.) and is not straightforward to explain.

Architecture	CPU	GPU
Number of nodes	64 + 1	64 + 1
Number of MPI ranks	512 + 1	64 + 1
Wallclock time [s]	7403.5	3605.1
Memory [MB] / process	710.8	3163.1
MPI (% of total walltime)	11.5%	8.8%
MPI_SYNC (% of total walltime)	6.5%	3.7%
MPI call1 (% of total walltime)	MPI_Wait 6.7%	MPI_Recv 7.4%
MPI call2 (% of total walltime)	MPI_Allreduce 5.5%	MPI_Allreduce 1.5%
%peak (DP)	12.3%	3.3% (not applicable)
PAPI FP OPS / process	2.13E+13	2.87E+12
PAPI L1 DCM / process	2.82E+13	9.64E+12
Write Rate (MB; MB/sec)	66594.6 MB 100.2 MB/s	64602.8 MB 743.14 MB/s

Table 2: *Selected output values from CrayPat for the benchmarks running on 8x8 nodes with the 1.1 km resolution configuration of the restructured COSMO code on Piz Daint*

Project plan: tasks and milestones

This project will be a one year project running from 1.10.2017 to 30.9.2018. The distribution of the individual tasks over the project period is shown in Table 3. During the first eight months, the calibration of the 1.1 km mesh-size COSMO version will be performed, using 6 parameters, to assess the impact of the soil memory and to serve as basis for optimizing the calibration methodology. This requires substantial computing resources (520'000 node hours, see table 4 below). This proposal requests to get these resources in full, in order to generate a complete data set and a robust reference calibration.

Months 9 to 12 are devoted to the refinement and to the optimization of the calibration method, using the reference calibration obtained in task 1, testing different strategies for reducing the computation costs of the calibration (e.g. data thinning in time, domain reduction, iterative calibration in model parameters space).

Task	Description	Months	Milestones
1	Full year COSMO-1 ensemble simulation with different tuning parameters is computed, and the coefficients of the meta-model are derived. The optimum set of tuning parameters is obtained. An independent validation of the calibration is performed.	1-8	A robust calibration of COSMO-1 is available.
2	Refinement and optimization of the calibration method with COSMO-1	9-12	Optimization strategy for reducing the computing cost of the method is defined and validated.

Table 3: *Distribution of tasks and milestones over the year.*

We expect a high interest from the broader research community in the findings of this project and plan to publish the results in a high-profile, peer-reviewed journal.

The project team which will execute the proposed research and simulations has the required skills to perform the tasks in a professional, effective and timely manner. Dr. Voudouri is the project leader of the CALMO-MAX project and will carry the overall responsibility. Dr. Voudouri has significant experience in running production simulation workflows for NWP on the CSCS systems, and in particular on Piz Daint. The Co-PI's (Dr. Bettems and Dr. Carmona) are both CALMO-MAX project members. The team will profit from existing scientific collaboration with Dr Mercogliano, Dr. Bucchignani and Dr. Avgoustoglou, who have experience with high resolution COSMO applications over the Mediterranean. The GPU version of the COSMO model has been developed by a team under the lead of Dr. Fuhrer, who has significant experience with running COSMO on hybrid Piz Daint.

Resource justification

The table 4 below shows the resources needed for the project. The calculations are based on the benchmark and scaling tests from the previous sections, using the 96 GPU nodes configuration, and on the number of simulations required for a robust calibration of N parameters (3*N+N*(N-1)/2+1, plus 10% reserve and 2 additional control runs). We added a margin of 20% to the total nodeh to account for jobs that need to be re-run because of failures of any kind (system problems, data problems etc.); we thus require 13'000 nodeh/year, on GPU, using the single precision, 96 nodes configuration, of the version 5.0 of the COSMO model, at 1.1 km resolution.

Because the methods for reducing the calibration cost have not yet been finalized, only an estimation of the resources required for the second task of the project can be done. One additionally asks for 50% of the resources allocated for task 1.

Simulation type	**Number of simulations**	**nodeh/year**	**Total nodeh**	**Storage needs**
One year simulations with 1.1 km resolution for 6 parameters (including 2 additional control runs)	40 x 1 year (GPU)	**13'000**	**520'000**	**45 TB**
Different types of reduced configurations (thinning in space, in time…)			**260'000**	**22 TB**
Total			**780'000**	**67 TB**

Table 4: *Overview of needs for the allocation period*

The amount of data produced with the simulations will be of the order of 67 TB for the entire amount of simulations. The raw data generated by the 1.1 km configuration of COSMO for the proposed experiments is in fact much larger, but a data thinning policy will be employed. Only 2-dimensional fields for specific variables will be kept for the entire simulation period.

Previous results

This proposal is a follow-up on the Production Project 'Objective calibration of weather prediction models'[7]. For this project, the requested and granted allocations were 1'066'000 nodeh, but we effectively used 317'000 nodeh. The discrepancy between the requested and the used allocations is due to the considerable technical problems met during the project, as discussed in details in the final report of the project[8]. Due to the fact that all tools have now been ported and tested on Piz Daint, and that much experience has been gained on using this platform, it is expected that running the proposed experiments on Piz Daint will now be straightforward.

---

[7] http://www.cosmo-model.org/content/tasks/priorityProjects/calmo/docs/CSCS_Proposal.pdf

[8] http://www.cosmo-model.org/content/tasks/priorityProjects/calmo/docs/CSCS_final_report.pdf

Research publications

Published and submitted papers associated to the previous project, in which CSCS is acknowledged: Avgoustoglou et al. 2017, Khain et al. 2015, Khain et al. 2017, Voudouri et al. 2017, Voudouri et al. 2017b, Voudouri et al. under preparation (see full references below).

References

Avgoustoglou E., A. Voudouri, P. Khain, F. Grazzini and J.M. Bettems, 2017: Design and Evaluation of Sensitivity Tests of COSMO model over the Mediterranean area. Perspectives on Atmospheric Sciences, Vol.1, Springer, pp 49-55

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär. 2012a. Objective calibration of regional climate models. *Journal of Geophysical Research,* **117**, D23115.

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär. 2012b. Exploring perturbed physics ensembles in a regional climate model. *Journal of Climate,* **25**, 4582-4599.

Bianco M. 2013. An interface for halo exchange pattern. http://www.prace-i.eu/IMG/pdf/wp86.pdf.

Khain P., I. Carmona, A. Voudouri, E. Avgoustoglou, J.-M. Bettems, F. Grazzini, 2015: The Proof of the Parameters Calibration Method: CALMO Progress Report. COSMO Technical Report, 25. http://www.cosmo-model.org/content/model/documentation/techReports/default.htm.

Khain P., I. Carmona, A. Voudouri, E. Avgoustoglou, J.-M. Bettems, F. Grazzini, P. Kaufman, 2017: CALMO Progress Report. COSMO Technical Report 31. http://www.cosmo-model.org/content/model/documentation/techReports/default.htm.

Damrath, U., 2009: Long-term trends of the quality of COSMO-EU forecasts. Presentation at the 11th COSMO General Meeting, Offenbach. www.cosmo-model.org/content/consortium/generalMeetings/general2009/wg5-versus2.htm.

Gysi, T., O. Fuhrer, C. Osuna, T. Schulthess, 2014: STELLA: A DSEL library for performance portable implementation of stencil computations on structured grids. *In preparation.*

Lapillonne, X., and O. Fuhrer, 2013: Using compiler directives to port large scientific applications to GPUs: An example from atmospheric science. *Parallel Processing Letters*, in press.

Voudouri A., E. Avgoustoglou and P. Kaufmann, 2017: Impacts of Observational Data Assimilation on Operational Forecasts. Perspectives on Atmospheric Sciences, Vol.1, Springer, pp 143-150

Voudouri A., Khain P., Carmona I., Bellprat O., Grazzini F., Avgoustoglou E., Bettems J.M and Kaufmann P., 2017b: Objective calibration of numerical weather prediction models. Atm. Res. Vol. 190, 128-140

Voudouri A., Khain P., Carmona I., Avgoustoglou E., Bettems J.M., Kaufmann P and Grazzini F.: Optimization of COSMO model performance over a very fine grid (under preparation)