ECMWF COPERNICUS REPORT

Copernicus Climate Change Service

Proposed method to assess the usability of the RCP-GCM-RCM modelling setup and description of the first results

D34b_Lot2.1.4.1

Issued by: SMHI/ Jon Stark

Date: 31/5/2020

Ref: C3S_D34b_Lot2.1.4_Proposed method to assess the usability of the RCP-GCM-RCM modelling setup and description of the first results _V1

Contributors

DMI
Ole B. Christensen
Fredrik Boberg

SMHI
Grigory Nikulin
Erik Kjellström

ETHZ
Marie-Estelle Demory

CNRS
Robert Vautard

MF
Samuel Somot
Lola Corre

ICTP
Erika Coppola

OGS
James Ciarlo
Cosimo Solidoro

MOHC
Erasmo Buonomo
Richard Jones

HZG
Claas Teichmann

KNMI
Erik van Meijgaard

Table of Contents

Executive Summary

This report outlines the proposed method to assess the usability of the experimental setup for the choice of forcing scenarios (RCPs), global (GCMs) and regional climate models (RCMs) in C3S_34b_Lot2. The project aims at partly filling a matrix of RCPxGCMxRCM combinations in a rational and efficient way. As the matrix will not be completely filled it is of interest to evaluate the choice of the strategy and to investigate to which extent the partly filled matrix resembles the full matrix. This evaluation may be used as guidance for design of future similar exercises including, for instance, CMIP6 results.

We present results from a study of Analysis of Variance (ANOVA) that has been performed on a completely filled 5x4 GCMxRCM matrix, which has been reported on in Christensen and Kjellström (2020), and discuss how an ANOVA-based technique can be used to fill holes in that matrix when single GCM-RCM combinations are removed. We find that a sparsely filled matrix may be sufficient for replicating the most important features of a fully populated matrix if designed properly. These results apply for seasonal mean features of temperature, precipitation and average 10m wind speed. In the report it is also outlined how this analysis can be extended in the final year of the project, which involves investigating other variables and other higher-order variability terms including extremes. It is also planned to investigate the role of internal variability for the matrix design.

The ANOVA analysis reveals that climate change, as opposed to mean climate, can be approximated well through a sum of a GCM-only and an RCM-only term. This gives hope that an emulation technique may show good results when used to analyse climate change, such that an incomplete matrix may be filled out and therefore give more equal weight to each GCM and to each RCM than it is the case for simple ensemble-of-opportunity averages of available models.

1. Introduction

This report describes the evaluation of the strategy generally followed in the planning of regional downscaling simulations in the COPERNICUS C3S_D34b_Lot2 (PRINCIPLES) project as described in D34b_Lot2.1.1.1 ("Experimental design for the GCM/RCM matrix"). The project aims to "fill" the 3-dimensional (RCP x GCM x RCM) matrix of EURO-CORDEX simulations that existed at the onset of the project. Since there are too many (several hundred) combinations to enable a complete matrix filling, a strategy has to be devised in order to partly fill the matrix in a rational way. The procedure, which has been adopted in collaboration with the ECMWF, has been to select sub-matrices, or slices, in the matrix and as far as possible fill those. This decision at the beginning of the project was made through expert judgement and can now be tested when most of the simulations are completed. With this technique the well filled sub-matrices resulting from the simulation effort in this project can be filled out with emulated values, and the accuracy can be investigated.

In this report we will outline a plan to conclude on the usefulness of this approach regarding the ability to draw conclusions about climate change from such a matrix, and to which degree uncertainties can be inferred from an incompletely filled matrix. Such conclusions may be useful for design of future experiments and considerations about what kind of matrix to provide for climate service purposes.

2. Results and Discussion

2.1 ANOVA

In order to distinguish between the influences of the various dimensions of the matrix, an Analysis of Variance (ANOVA) analysis has been performed on one matrix slice. Only the RCP8.5 scenario is considered, but 5 GCMs and 4 RCMS (Table 1) have been used in all possible mutual combinations. The work has been published as Christensen and Kjellström (2020) with data from a time when only 19 of the 20 simulations in question were available; most of the results presented below are further elaborated in this paper. Nine of the twenty simulations have been produced within C3S_34b_Lot 2.

Table 1. The GCMs and RCMs participating in C3S_34b_Lot 2, which have been analyzed in Christensen and Kjellström (2020).

GCMs	CNRM-CM5 (Voldoire et al., 2013)	EC-EARTH (Hazeleger et al., 2012)	HadGEM2-ES (Collins et al., 2011)	MPI-ESM-LR (Giorgetta et al., 2013)	NorESM1-M (Bentsen et al., 2013)
RCMs	HIRHAM5 (Christensen et al., 2006)	REMO2015 (Jacob et al., 2012)	RACMO22E (van Meijgaard et al., 2008, 2012)	RCA4 (Samuelsson et al., 2011)	

In this ANOVA approach, we look at the influence from the choice of period (present or future, chosen as 1981-2010 and 2071-2100, respectively), of GCM, and of RCM in a statistical way, writing a result Y_{ijkl} from period *i*, GCM *j,* RCM *k* and year *l* as

$$Y_{ijkl} = M + S_i + G_j + R_k + SG_{ij} + SR_{ik} + GR_{jk} + SGR_{ijk} + Z_{ijkl} \quad (1)$$

i.e., as a sum of a mean (*M* over both periods and all GCM-RCM combinations), an average climate change contribution (*S*; really half the climate change, as it is defined as the deviation from two-period mean of each period), an average contribution from each individual GCM (G_j) relative to the average and a corresponding RCM contribution (R_k), (half the) GCM and RCM contributions to climate change (SG_{1j} and SR_{1k}), a cross term for an individual simulation's deviation from the before-mentioned averaged terms (*GR*), a corresponding climate change cross term (*SGR*), and finally a contribution from inter-annual variability (Z_l). Each term is defined to sum to zero over any index.

Such an analysis has been performed for each grid point, each season, and for seasonal averages of temperature, precipitation, and mean wind speed. One main general conclusion is that cross terms are important for mean climate (*GR*), but generally negligible for climate change (*SGR*). When the average characteristics of a GCM (the *G* term) and the average characteristics of an RCM (the *R* term) are significant, but the cross terms are not, it means that the signal from a particular model does not depend much on the model it is combined with, just on average quantities; in other words, the climate change response of an entire matrix could, to a strong degree, be emulated if just the averaged terms *G* and *R* were known, approximated from an incompletely filled matrix. In Table 2 (from Christensen and Kjellström, 2020) we illustrate this, listing the percentage of points in the area, where each of the terms in Eq. 1 are formally statistically significant. We note that while the *GR* term is generally statistically significant in more than 60% of the domain, the *SGR* term show this behaviour in less than 5% of the grid points. So, speaking generally, we cannot estimate the average climate of a particular simulation from a sum of a GCM and an RCM contribution over more than at most half the area. But the climate change is very well estimated this way.

Table 2. Percentage of grid points where each term is significant at a 95% level for temperature, precipitation, and wind speed, winter and summer. From Christensen and Kjellström (2020).

	S	G	R	SG	SR	GR	SGR
T winter	100	100	100	100	50	83	5
T summer	100	100	99	100	40	73	3
pr winter	91	99	96	80	27	62	2
pr summer	76	99	98	72	33	74	5
w10m winter	80	100	100	82	36	63	3
w10m summer	84	100	99	86	43	75	2

The ANOVA analysis gives a clean separation of climate and climate change into GCM and RCM influence. As an example we show in Fig. 1 the various terms for winter temperature. The typical winter climate change signal with stronger temperature increase in the north is clearly seen in the figure (second row from the top); the two panels in this row are the same, with opposite sign; the sum of the two is the average climate change. The strong imprint of the GCMs and RCMs is also clear. For example, it is clear that the CNRM-CM5 model is colder than average while the NorESM1-M model is warmer over much of Europe (third row). Similarly, RACMO22E tends to be colder than the other RCMs particularly over southern Europe while HIRHAM5 is warmer than average over most of the domain (fourth row). Apart from these strong features of the model climates the choice of model also influences the climate change signal. For example we note a stronger than average change in the HadGEM2-ES and a weaker than average change in NorESM1-M (fifth row). The RCMs also show some imprint on the climate change signal with REMO2015 showing stronger than average temperature increase in parts of northern Europe while RCA4 shows stronger signal over the Alps (bottom row).

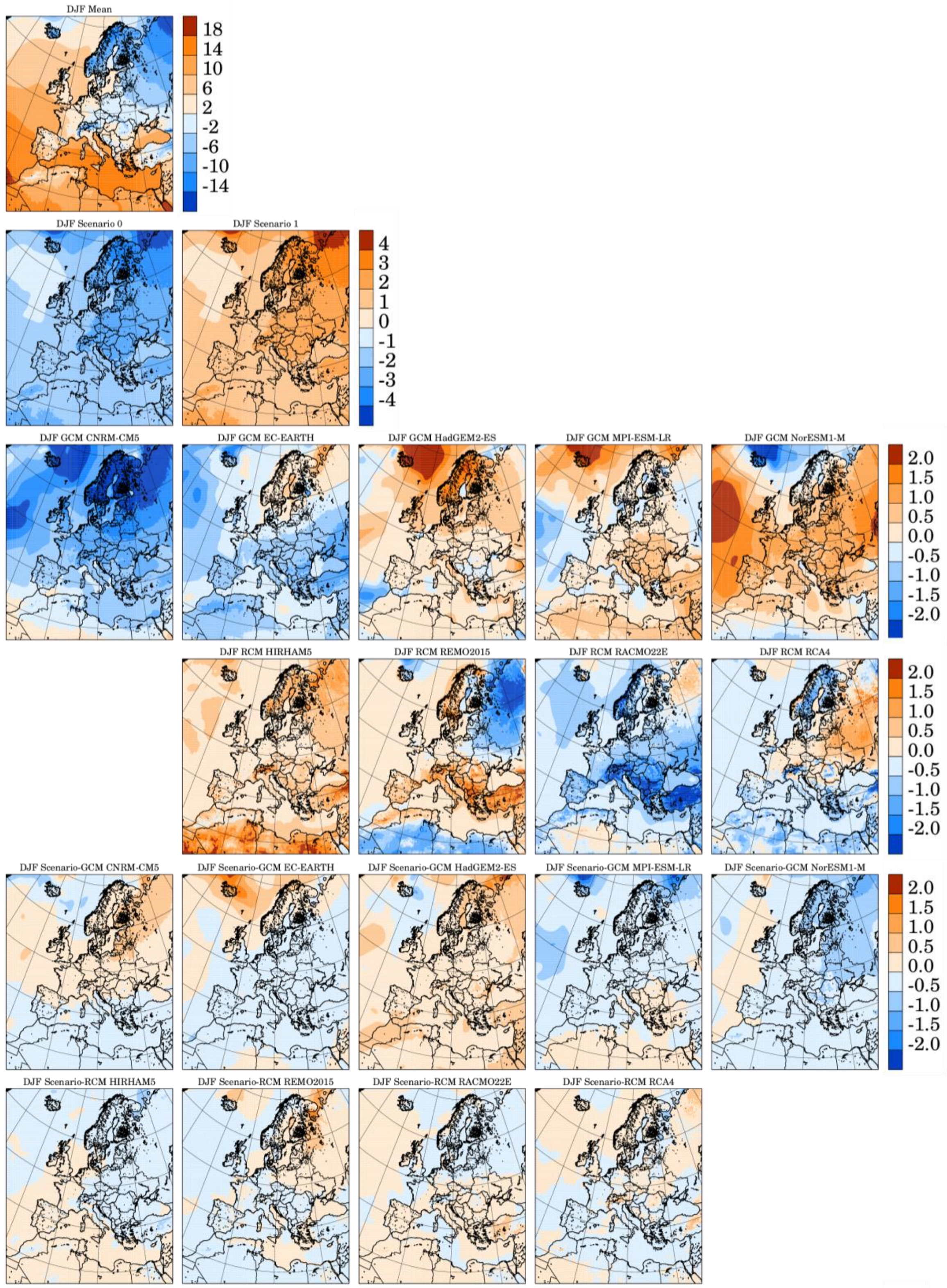

Figure 1 Individual terms from ANOVA analysis of DJF temperature; modified from Christensen and Kjellström 2020. Top row: M; second row S_1 and S_2, the deviation of control climate and scenario climate from the total mean; third row G_j for the 5 GCMs; fourth row R_k for the 4 RCMs. Fifth row: SG_{2j} sixth row: SR_{2k}.; same color scale as in the fifth row.

2.2 Effects of Matrix Holes

Based on an ANOVA analysis it is possible to "fill holes" in a matrix, i.e., calculate emulated values for model combinations, which have not yet been filled by an actual simulation. Given trustworthy values of the linear (single-index) terms in an ANOVA analysis, emulated values of an entry Y_{ijk} can be calculated from the equation $S_{ijk} = 0$ for a hole corresponding to GCM j and RCM k and valid for both periods $i=1,2$; this corresponds to the explicit equation

$$0 = Y_{ijk} - Y_{ij.} - Y_{i.k} - Y_{.jk} + Y_{i..} + Y_{.j.} + Y_{..k} - Y_{...} \qquad (2)$$

where the dots indicate averaging over a dimension. Since both total means and single-RCM or single-GCM means enter this equation, a fully coupled linear system of equations will result from a situation with several holes. This equation system is not always solvable. Let us examine the GCMxRCM population matrix (1 if a simulation exists, otherwise 0). Examples of population matrices where the procedure does not work are instances with no simulations at all for a specific GCM or a specific RCM.

Also situations where the matrix can be split into two non-connected sub-matrices are invalid; see examples in Fig. 2. Conceptually: If one sub-matrix has generally much higher values than a remaining disconnected sub-matrix, there is no way to know if the set of GCMs or the set of RCMs are the reason; this is reflected in a redundancy in the set of equations, which makes the solution non-unique. For the current situation of 5 GCMs and 4 RCMs the maximum number of holes turns out to be 12, i.e., having performed 8 simulations. It means that for a 5x4 GCM-RCM matrix we need at least 8 simulations to have a chance to build the entire matrix with this ANOVA technique. Of course it does not mean that the reconstruction of the missing element would be satisfactory in practice, but that is a minimum number of simulations. Even for more simulations, only some matrix configurations allow for this method to be applied.

1	1			
	1			1
		1	1	
			1	1

1	1			1
	1			
		1	1	
			1	1

1	1			
1	1			
		1	1	
			1	1

1	1	1		1
		1	1	
			1	1

Figure 2 Examples of GCMxRCM matrices with 12 holes. The two in the top are solvable. The bottom 2 not: The first can be split into 2 sub-matrices as indicated by yellow and green colours; the second has a row without simulations (red).

We assume that both present and future periods either both exist or both do not exist.

An example of the kind of equation system to solve is the following, where we have two holes, $Y_{ij'k'}$ and $Y_{ij''k''}$ and assume that $j' \neq j''$ and $k' \neq k''$, i.e., that the holes have different GCMs as well as different RCMs. Let's look at the control period *i=1.* In eq. 2 only the last term, the total ensemble mean, connects the two holes. The 2-dimensional system can then be expressed in the form

$$\begin{bmatrix} (N_G-1)(N_R-1) & 1 \\ 1 & (N_G-1)(N_R-1) \end{bmatrix} \begin{bmatrix} Y_{1j'k'} \\ Y_{1j''k''} \end{bmatrix} = \begin{bmatrix} B_{1j'k'} \\ B_{1j''k''} \end{bmatrix}$$

where the number of GCMs is $N_G = 5$, the number of RCMs is $N_R = 4$, and the B terms are shorthands for expressions, which do not depend on the hole values, only on various averages of existing simulation data for the point, season and field in question. The exact definitions can be determined from Eq. 2.

We want to study to which degree sparsely filled matrices where holes are synthetically filled can replicate features of the full matrix. An example of a metric is shown in Fig. 3, where we look at the mean climate averaged over both periods ($Y_{.jk}$) and the mean climate change ($Y_{2jk} - Y_{1jk}$) separately at a seasonal basis. Here, 1000 different matrices with a given number of holes have been generated in a bootstrap (for 1 and 2 holes we take all possible choices: 20 possibilities for one hole, and 20*19/2=190 for two holes, corresponding to the number of different GCMxRCM matrices with 1 or 2 holes out of 20 simulations). Only solvable matrices (cf. Fig. 2) have been considered. For each matrix, the holes have been filled as described above. Since any specific peculiarities of an individual simulation (*GR* term) are impossible to emulate, we compare emulated values at each point *jk* with the value emulated when only *jk* was missing, i.e., the development of emulated values as more and more existing simulations are removed.

For a number *n* of holes, we look at each *jk* combination in turn and find the matrices where *jk* is one of the holes, and calculate the average squared deviation from the 1-hole emulated value across all grid points. We average this quantity over all relevant matrices and after that over all *jk* combinations. In the end we have, for each field and season investigated, a measure of the mean squared deviation from the best emulated value, as a function of *n*. Taking the square root we are left with a measure of the deviation from the situation with only one hole. The unit is the same as that of the quantity examined.

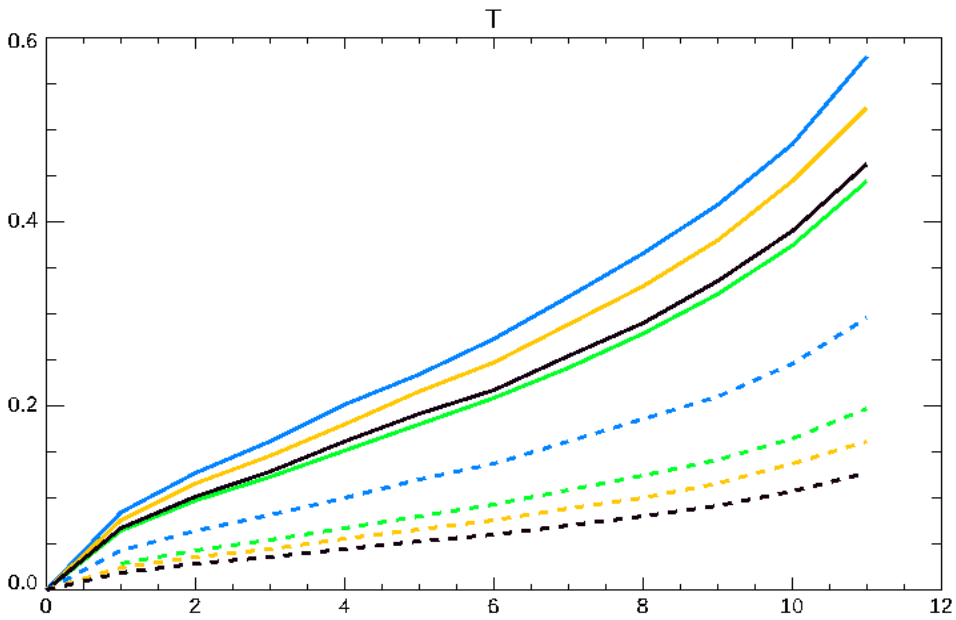

Figure 3 Average deviation (deg. C) from one-hole emulated values of seasonal mean temperature as a function of the number of excess holes (the number of holes more than the one hole we compare to, for each hole in each configuration). Up to 1000 configurations have been examined for each number of holes (see text). More than 11 excess holes, i.e., a total of 8 existing simulations, cannot be treated with this method. Solid lines: mean climate. Dashed lines: Future minus mean, i.e., 50% of the climate change signal. DJF blue, MAM green, JJA orange, SON black.

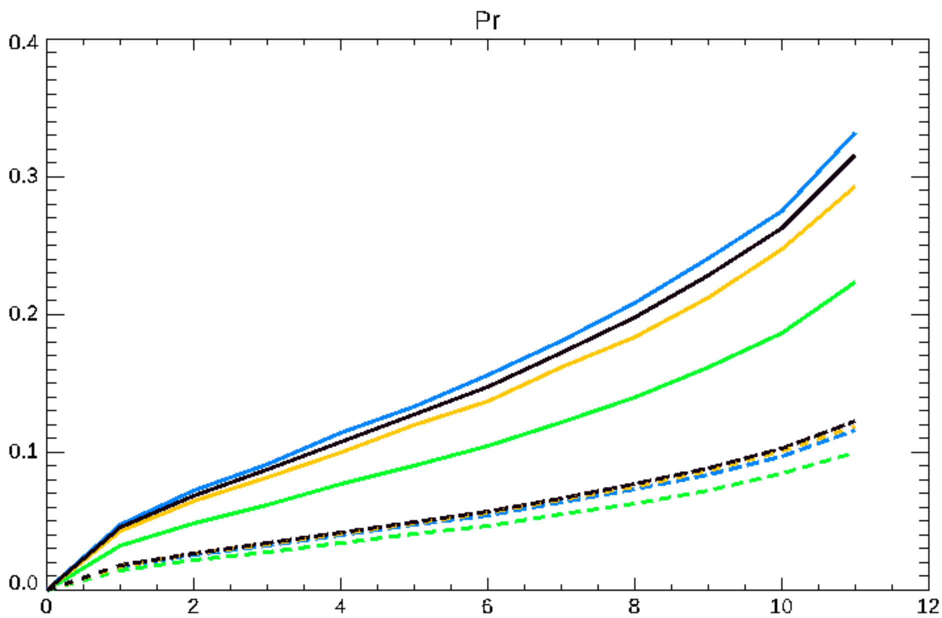

Figure 4 Like Fig. 3, but for precipitation (mm/d).

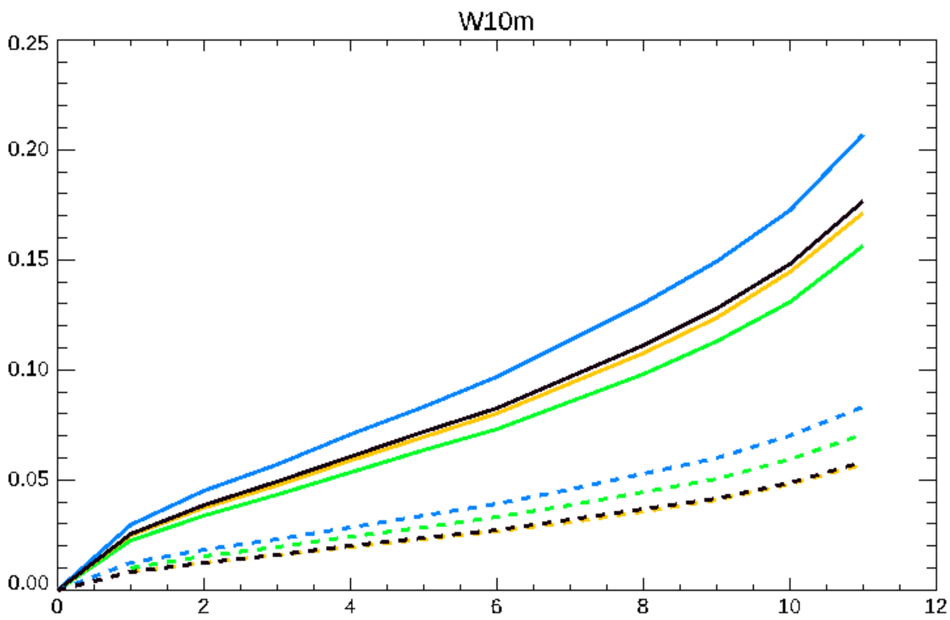

Figure 5 Like Fig. 3, but for average 10-m wind speed (m/s).

The shapes of these curves are quite similar: A large jump in deviation happens when 1 extra hole is introduced. After that there is a slight upwards curve as a function of the number of holes, until we reach the maximum possible number of 11 extra holes where we note that the steepness of the curves increases. The average deviation over the full domain is of the order 5-10 times larger for 11 extra holes compared to 1 extra hole both for the mean and for the climate change signal. We also note that there are differences in how large the errors are between the seasons. Winter stands out with the largest error for the mean climate for all three variables. Conversely, spring shows the smallest error. For the climate change signal the difference between the seasons are less consistent. However, winter stands out also here with larger errors for temperature and wind speed than in any of the other seasons.

A different way to approach the issue is to look at how to best estimate the complete matrix mean including all 20 simulations from a set of fewer available simulations. This complete matrix mean is not necessarily closer to the physical truth than a simple average; it does, however, introduce a more democratic weighting of both the RCMs and the GCMs involved. In a pure ensemble of opportunity, each GCM will have an effective weight corresponding to the number of times it has been downscaled, and correspondingly for each RCM; with the technique being developed here, there will be equal weight between the GCMs chosen to be represented in the matrix and also between the RCMs. We have therefore analysed two strategies for approximating this quantity from an incomplete matrix: The simple ensemble-of-opportunity "direct" average of the existing simulations in the incomplete matrix, and the mean obtained from matrix filling with emulated values filled into holes as outlined above. In Fig. 6 we see an example for one arbitrary 5x4-member matrix with 12 holes and only 8 simulations. It is clear that the direct 8-member ensemble-of-opportunity mean is mostly much farther from the complete-matrix "truth" than it is possible to achieve with the matrix-filling method used here. Further, we note relatively poor performance for the direct-average method over sea in general, where the emulation technique gives equal weight to the SST values of each GCM, just as in the true full-matrix average. A notable exception is the sea ice covered areas north of Russia and of Iceland, where the matrix filling technique is further from the true average than the direct average. The emulated results are worse in some northern sea-ice covered areas, where Christensen and Kjellström (2020) saw large difference between individual simulations and their emulated counterparts, i.e., a large role of specific GCM-RCM combinations. This is probably related to specifics of sea ice description in the 8 models used, compared to the total ensemble. This needs further investigation.

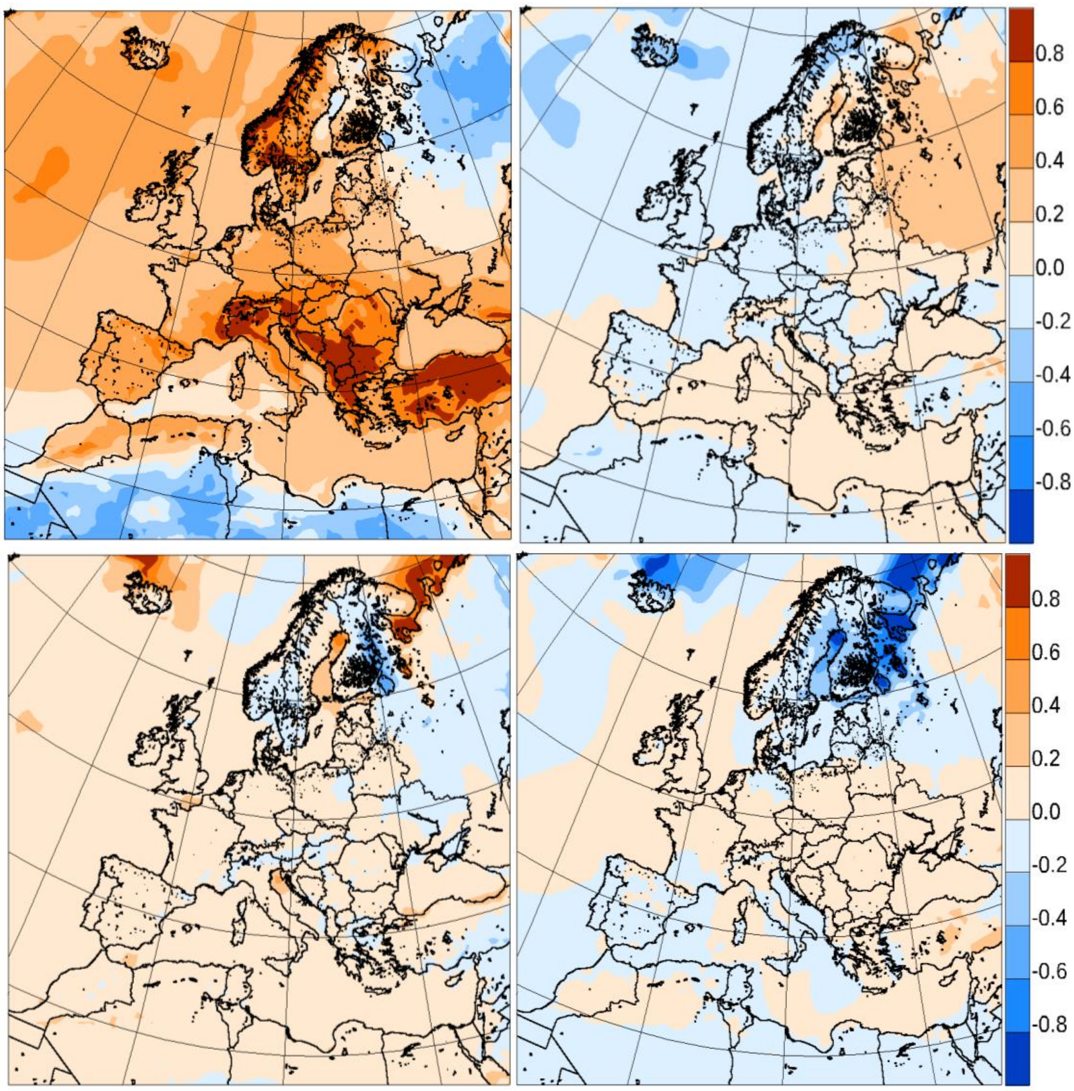

Figure 6 Top: Deviation of one direct 8-model average DJF temperature from true 20-model average (deg. C). Bottom: Deviation of emulated 20-value average based on the same 8 models from true 20-model average. The left column shows differences in mean climate, the right column shows differences in (full) climate change.

To systematise this also for other number of holes and other seasons, we plot in Fig. 7 the RMS average over all points and all bootstrapped matrices of the deviations seen in Fig. 6 as a function of the number of holes. It is clear from the figure that the matrix-filling procedure creates a matrix that is much more similar to the original full matrix compared to the direct mean of any ensemble of opportunity consisting of fewer members. This is true for all numbers of holes investigated both for the mean climate (top panel) and for the climate change signal (bottom panel).

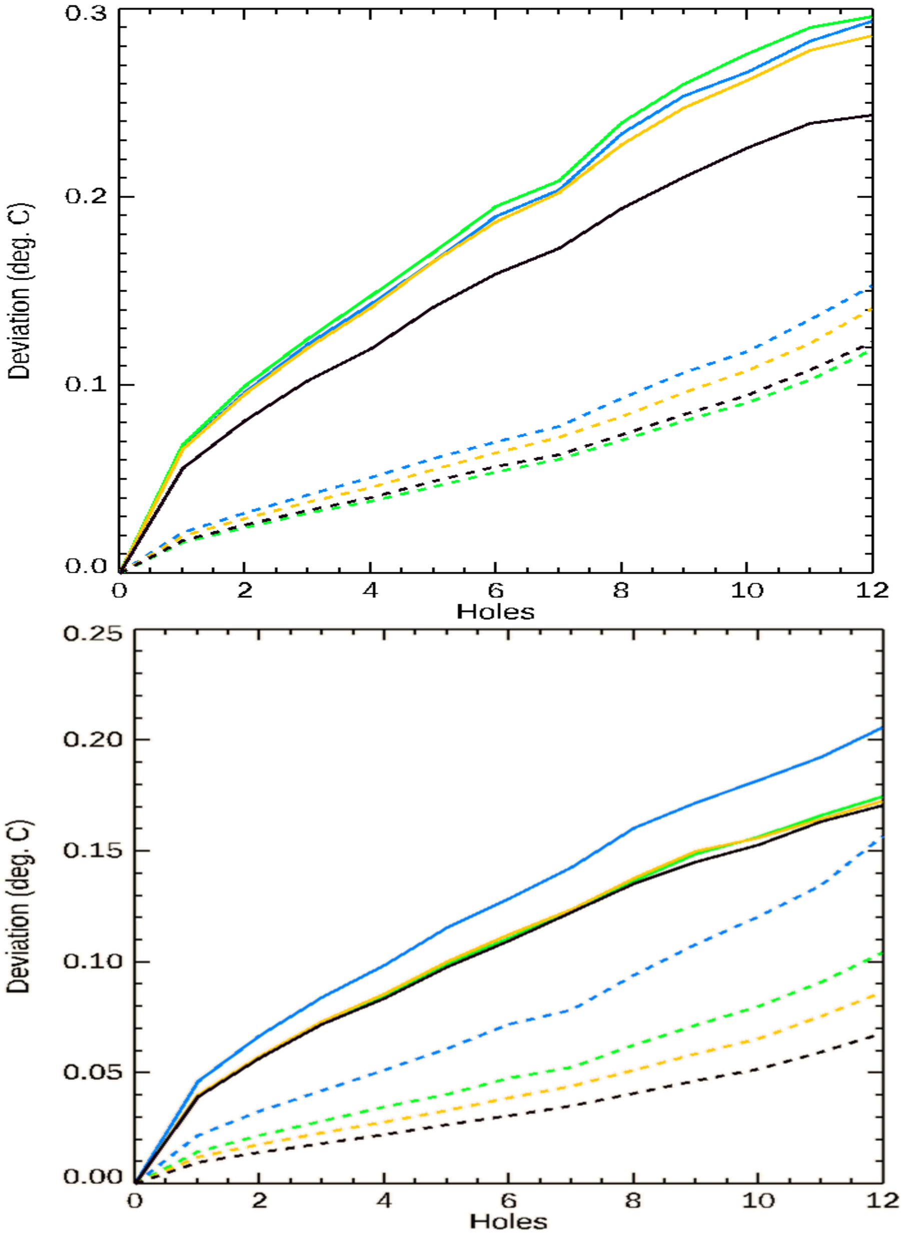

Figure 7 RMS deviations over points and bootstrapped simulations of deviation from true 20-model complete matrix average seasonal mean temperatures as a function of the total number of holes. Full lines: Deviation of direct mean of existing ensembles-of-opportunity from 20-model truth; dashed lines: Deviation of means over emulated full matrix from 20-model truth. Top panel: Mean climate. Bottom panel: climate change. DJF blue, MAM green, JJA orange, SON black.

3. Perspectives and Plans

Current work focuses on a systematic investigation of effects of making the matrix sparser and sparser. This will allow a quantification of how much information may be gained by adding new simulations to existing, sparse real-world simulation matrices. Of course, we can only aim for an emulation of filled matrices; it is an additional challenge to ascertain that the GCMs and RCMs in the matrix as far as possible are representative for larger multi-model ensembles.

In the remainder of the project, we will extend the analysis to investigate also what happens when not just multi-annual seasonal averages are being studied, but also, e.g., extremes; both extreme seasons within a simulation and daily extremes. Apart from that, it is obviously possible to look at other fields than the current set consisting of temperature, precipitation, and 10-m wind speed.

In addition, we will investigate if it is possible to go beyond the current model-only world and learn something about biases. One obvious step would be to make a missing-simulation analysis of bias, i.e., investigating to which extent the biases of individual simulations can be written as the sum of a GCM-specific part and an RCM-specific part. This would supplement the current analysis of mean fields and of climate change and also supplementing the evaluation of the entire ensemble performed by Vautard et al. (2020).

A different perspective, which will also be pursued in the project, is to put these results into perspective through further analyses of the role of internal variability, particularly of the GCM, in significance determination. Even when looking at 30-year averages, longer-time variations exist in GCM simulations, the details of which can be studied through downscaling of different ensemble members of the same GCM. In this project, several different RCMs have been applied to downscaling of 3-member GCM ensembles for two different CMIP5 GCMs.

Finally, when the matrix-filling technique has been analysed, it will be applied to the entire existing Euro-CORDEX GCMxRCM matrix, and a comparison between direct averages and emulated-matrix averages should be made and studied in order to possibly achieve a more trustworthy estimate of the information contained in the very large but also very expensive model matrix.

4. References

Bentsen M, Bethke I, Debernard JB, Iversen T, Kirkevåg A, Seland Ø, Drange H, Roelandt C, Seierstad IA, Hoose C, Kristjánsson JE (2013) The Norwegian Earth System Model, NorESM1-M—Part 1: description and basic evaluation of the physical climate. *Geosci Model Dev*. 6:687–720. https://doi.org/10.5194/gmd-6-687-2013

Christensen OB, Drews M, Christensen JH, Dethloff K, Ketelsen K, Hebestadt I, Rinke A (2007) The HIRHAM Regional Climate Model. Version 5 (beta). Danish Climate Centre, Danish Meteorological Institute. Denmark. Danish Meteorological Institute. Technical Report, No. 06-17

Christensen OB, and Kjellström E (2020) Partitioning uncertainty components of mean climate and climate change in a large ensemble of European regional climate model projections. *Clim. Dyn.* https://doi.org/10.1007/s00382-020-05229-y

Christensen JH, and Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climatic Change* **81**, 7-30.

Collins WJ, Bellouin N, Doutriaux-Boucher M, Gedney N, Halloran P, Hinton T, Hughes J, Jones CD, Joshi M, Liddicoat S, Martin G, O'Connor F, Rae J, Senior C, Sitch S, Totterdell I, Wiltshire A, Woodward S (2011) Development and evaluation of an Earth-System model – HadGEM2. *Geosci Model Dev*. 4:1051–1075, https://doi.org/10.5194/gmd-4-1051-2011

Giorgetta MA, et al. (2013) Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *J Adv Model Earth Syst* 5:572–597. https://doi.org/10.1002/jame.20038

Hazeleger W, Wang X, Severijns C, Ştefănescu S, Bintanja R, Sterl A, Wyser K, Semmler T, Yang S van den Hurk B, van Noije T, van der Linden E, van der Wiel K (2012) EC-Earth V2.2: description and validation of a new seamless earth system prediction model. *Clim Dyn* 39: 2611. https://doi.org/10.1007/s00382-011-1228-5

Jacob D, Elizalde A, Haensler A, Hagemann S, Kumar P, Podzun R, Rechid D, Remedio AR, Saeed F, Sieck K, Teichmann C, Wilhelm C (2012) Assessing the transferability of the regional climate model REMO to different coordinated regional climate downscaling experiment (CORDEX) regions. *Atmosphere* 3:181–199. https://doi.org/10.3390/atmos3010181

Kjellström E, Thejll P, Rummukainen M, Christensen JH, Boberg F, Christensen OB, Fox Maule C (2013) Emerging regional climate change signals for Europe under varying large-scale circulation conditions, *Clim. Res.*, **56**, 103–119, doi: 10.3354/cr01146.

Samuelsson P, Jones C, Willén U, Ullerstig A, Gollvik S, Hansson U, Jansson C, Kjellström E, Nikulin G, Wyser K (2011) The Rossby Centre Regional Climate Model RCA3: model description and performance. *Tellus* A 63:4–23. https://doi.org/10.1111/j.1600-0870.2010.00478.x

van Meijgaard E, van Ulft LH, van den Berg WJ, Bosveld FC, van den Hurk BJJM, Lenderink G, Siebesma AP (2008) The KNMI regional atmospheric climate model RACMO version 2.1. KNMI Tech. Rep. TR-302, 43 pp

van Meijgaard E, van Ulft LH, Lenderink G, de Roode SR, Wipfler L, Boers R and Timmermans RMA (2012) Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe, Climate changes Spatial Planning publication: KvR 054/12, the Programme Office Climate changes Spatial Planning, Nieuwegein, the Netherlands, ISBN/EAN 978-90-8815-046-3, 44 pp. http://climexp.knmi.nl/publications/FinalReport_KvR-CS06.pdf

Vautard R, Kadygrov N, Iles C, Boberg F, Buonomo E, Bülow K, Coppola E, Corre L, van Meijgaard E, Nogherotto R, Sandstad M, Schwingshackl C, Somot S, Aalbers E, Christensen OB, Ciarlo JM, Demory M-E, Giorgi F, Jacob D, Jones RG, Keuler K, Kjellström E, Lenderink G, Levavasseur G, Nikulin G, Sillmann J, Solidoro C, Sørland SL, Steger C, Teichmann C, Warrach-Sagi K, Wulfmeyer V (2020) Evaluation of the large EURO-CORDEX regional climate model ensemble. *J. Geoph. Res*. submitted

Voldoire A, Sanchez-Gomez E, Salas y Mélia D et al. (2013) The CNRM-CM5.1 global climate model: description and basic evaluation. *Clim Dyn* 40:2091. https://doi.org/10.1007/s00382-011-1259-y

Copernicus Climate Change Service

ECMWF - Shinfield Park, Reading RG2 9AX, UK

Contact: info@copernicus-climate.eu

climate.copernicus.eu copernicus.eu ecmwf.int