

Is there potential added value in COSMO–CLM forced by ERA reanalysis data?

Claus-Jürgen Lenz¹  · Barbara Früh¹ · Fatemeh Davary Adalatpanah^{1,2}

Received: 8 April 2016 / Accepted: 30 January 2017 / Published online: 1 March 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract An application of the potential added value (PAV) concept suggested by Di Luca et al. (Clim Dyn 40:443–464, 2013a) is applied to ERA Interim driven runs of the regional climate model COSMO–CLM. They are performed for the time period 1979–2013 for the EURO–CORDEX domain at horizontal grid resolutions 0.11°, 0.22°, and 0.44° such that the higher resolved model grid fits into the next coarser grid. The concept of the potential added value is applied to annual, seasonal, and monthly means of the 2 m air temperature. Results show the highest potential added value at the run with the finest grid and generally increasing PAV with increasing resolution. The potential added value strongly depends on the season as well as the region of consideration. The gain of PAV is higher enhancing the resolution from 0.44° to 0.22° than from 0.22° to 0.11°. At grid aggregations to 0.88° and 1.76° the differences in PAV between the COSMO–CLM runs on the mentioned grid resolutions are maximal. They nearly vanish at aggregations to even coarser grids. In all cases the PAV is dominated by at least 80% by its stationary part.

Keywords Potential added value · COSMO–CLM · Regional climate model

1 Introduction

Increasing computing power and storage capacities allow the development and application of regional climate models (RCMs) with increasing spatial resolution. Compared to global reanalysis and global climate projections the aim of RCMs is to provide more detailed information due to the higher numerical grid resolution and hence to a better representation of the surface characteristics as well as the meteorological processes. Moreover, the RCMs can be based on non-hydrostatic equation systems and allow to use more sophisticated physical parameterizations or even to calculate these physical processes directly.

These more detailed information provided by the RCMs are usually called “added value”. The “added value” is often used to show the advantages of RCM simulations, but scientifically sound investigations only started recently. A comprehensive summary about the different kinds of understanding of “added value” as well as the current state of knowledge and research in this subject can be found in Di Luca et al. (2015).

Feser (2006) presents the application of a spatial filter, which is used to decompose the driving data and the RCM results into different spatial scales. For these spatial scales the added value is shown visually by comparing maps, which represent the deviation of both, the driving NCP reanalysis data and the Regional Model REMO results, from high resolution analysis data originating from the “Europa Model” of Germany’s National Meteorological Service (DWD). Added value can be seen in the REMO results especially in the medium-scaled temperature pattern comparing time series of spatial correlation coefficients of NCEP and REMO results with the “Europa Model” analysis data, respectively.

✉ Claus-Jürgen Lenz
claus-juergen.lenz@live.de

Barbara Früh
barbara.frueh@dwd.de

¹ Deutscher Wetterdienst, Frankfurter Straße 135,
63067 Offenbach, Germany

² Brandenburgische Technische Universität Cottbus, Lehrstuhl
Umweltmeteorologie, Burger Chaussee 2, 03044 Cottbus,
Germany

In Feser et al. (2011), the authors give an overview on the work done in exploring the added value of regional atmospheric models. According to the paper the added value depends on the region of interest, e.g. added value can be located in orographically heterogeneous regions and near coastlines. Further, added value can predominantly be found in mesoscale phenomena as polar lows and tropical cyclones, which cannot be resolved sufficiently by the global model grids.

In their recent paper, Lucas–Picher et al. (2016) present the added value of simulations performed with the Canadian Regional Climate Model (CRCM5) for North America. The authors show the benefits of the RCM simulations done on model grids with different horizontal resolution for some North American regions affected by characteristic weather phenomena as snow bands, orographical rainfall, and wind channeling. The added value increases with increasing RCM grid resolution. Whereas the added value in the RCM simulations at a 0.44° grid is limited compared to the driving data, the highest gain in added value is obtained enhancing the grid resolution from 0.44° to 0.22° .

In Winterfeldt and Weisse (2009), the authors come to the conclusion that in the near surface wind pattern over open sea no added value exists. In the coastal area it can be found only in frequency distributions.

Most papers on added value deal with the horizontally extremely heterogeneous precipitation. Especially the influence of directly calculated (grid width <3 km) versus parameterized deep convection (grid widths >5 km) is subject of investigation. Kendon et al. (2012) present results of a climate version of the Met Office Unified Model for the southern part of the UK at a 1.5 km grid and compare them with the outcome of a 12 km RCM with respect to precipitation intensity, duration, and pattern. A comparison of the precipitation results of COSMO–CLM simulations at a 12 km and a convection resolving 2.2 km grid for a 10 year period is performed in Ban et al. (2014) for the region of the European Alps. Prein et al. (2013) elaborate the added value of convection permitting simulations for COSMO–CLM, MM5, and WRF at different physical and numerical settings for summer 2007 and winter 2007/2008 in eastern Austria. All three papers confirm that the daily cycle of precipitation is more realistically represented in the high-resolution simulations. Secondly, in the model runs on the finer grids the precipitation intensity, size of precipitation area, and precipitation sum have a better agreement with measurements.

Dosio et al. (2015) showed that COSMO–CLM simulations over Africa driven by four different GCMs cannot enhance the forecast skill of the large scale fields. The results of COSMO–CLM depend strongly on the driving global model. But the probability density function of precipitation, the duration of dry and humid periods, and the

frequency of high precipitation events are more realistic in the RCM. In addition, COSMO–CLM is able to simulate the bimodal frequency of precipitation especially of the western African monsoon.

Prömmel et al. (2010) investigate the added value of the 2 m air temperature with a regional model REMO run at a $1/6^\circ$ grid for the European Alps. The authors compare the results of the REMO run and the ERA 40 forcing data with station data for some Alpine subregions using the root mean square error and the reduction of error. Added value is found in the flatter regions around the Alps only in winter, but for the inner Alpine region in the whole year.

Di Luca et al. (2012, 2013a, b) propose a method to consider if there is any potential for added value in high-resolution models and named their concept the “potential added value”. In this concept the results of a high-resolution model are not compared with results from another coarser grid model or measurements, but with the model results themselves aggregated to coarser grid cells similar to those of a virtual driving model. After this aggregation the results on the original grid and the virtually created coarser grids are analysed to calculate the benefit of the higher resolution for different meteorological parameters. Because the potential added value concept bases on one single model run, this method is not influenced by different physical parameterizations as used often in models covering different spatial scales. The “potential added value” concept takes only into account the effects of the representation of the topography on the original and the aggregated coarser grids.

In the present paper, the potential added value (PAV) concept of Di Luca et al. (2013a) is applied to annual, seasonal, and monthly means of the 2 m air temperature resulting from COSMO–CLM runs with different grid resolution driven by ERA Interim data. After a short description of the data used for this study in Sect. 2, and an introduction into the PAV concept of Di Luca et al. (2013a), in Sect. 3 the results of the PAV study are presented, and discussed in Sect. 4. In Sect. 5, a summary with the main conclusion is provided.

2 Data used

2.1 Model data

The base for this study are results of the regional climate model COSMO–CLM, which is the climate version of the nonhydrostatic weather forecast model COSMO and further developed by the CLM-Community (Böhm et al. 1996; Rockel et al. 2008).

Three of the considered COSMO–CLM simulations are performed on the EURO–CORDEX domain (Jacob 2014) at horizontal grid resolutions of 0.44° , 0.22° , and 0.11° .

The numerical grids are chosen such that 2×2 grid cells of the 0.11° grid exactly fit into one grid cell of the 0.22° grid and 4×4 grid cells into one of the 0.44° coarse grid. This choice allows a direct comparison of aggregated finer grid cells with results calculated on the next coarser numerical grid. According to the EURO–CORDEX configuration the model domain contains 40 vertical model levels. The COSMO–CLM runs are abbreviated in the following as CLM44 (0.44°), CLM22 (0.22°), and CLM11 (0.11°).

CLM22 and CLM44 are performed for January 1959 to December 2013. For 01 January 1959, the atmosphere as well as the soil and the soil surface parameters are initialized by ERA 40 (Uppala et al. 2005). For 1959–1978, the lateral boundary conditions from ERA 40 are used switching to ERA Interim (Dee et al. 2011) on 01 January 1979 until the end of the simulation.

The atmospheric and soil parameters of CLM11 are initialized on 01 January 1979 by data from CLM22 interpolated to the 0.11° grid. As for the other runs at coarser model grids ERA Interim reanalysis forcing is used until the end of 2013.

In addition, a fourth COSMO–CLM run is performed on a model domain containing central Europe at a model grid of 0.025° horizontal resolution for the time period from 1995 to 2012. The initialization and the driving data for this run were taken from an ERA Interim driven COSMO–CLM run performed in the context of EURO–CORDEX (Kotlarski et al. 2014). Whereas in all COSMO–CLM simulations at 0.11° and coarser the deep convection is parameterized using the Tiedtke mass flux scheme, this parameterization is switched off in the simulation at the 0.025° grid due to the assumption that deep convection is calculated directly on such a numerically high-resolving model grid. In contrast to the model runs CLM11, CLM22, and CLM44 the number of vertical model layers in this run is 50. It will be named CLM025 in the following. All COSMO–CLM runs have been performed on a rotated model grid.

2.2 Observation data

For comparison purposes in addition to the COSMO–CLM simulation results a high-resolution gridded observational data set is used and investigated in the same manner as the model results. The purpose of this investigation is an evaluation of the model results. It is expected that an application of the PAV concept on a high resolution observational data set can simply be performed and should yield results comparable to those achieved for the model results. In this sense the application of the PAV concept on a high-resolution gridded observational data set can be considered as a kind of qualitative validation of both, the model results and the observational data set.

As observational data the HYRAS data set (Rauthe et al. 2013; Brienen et al. 2016) is selected. The HYRAS data set contains daily sums of precipitation, daily means of 2 m air temperature, and air humidity at a 5 km grid covering all river basins in Germany as well as parts of neighbouring countries with shared river catchments (e.g. Rhine, Danube) for the time period from 1951 to 2006. Due to this high spatial resolution the HYRAS data set can be used for the evaluation of simulations of regional climate models performed at high-resolution grids.

3 Definition of the potential added value

In contrast to considerations of the (real) added value, which is usually investigated by comparing simulation results of numerical climate models with a reference data set originating from an external source, e.g. the driving data, observations, reanalysis data, or data from other models, the potential added value is calculated from the inspected model results only. The advantage of this procedure is to avoid differences in the bias of the datasets to be compared, e.g. due to different parameterizations or to different basic climate states in the models and/or data involved.

The concept of the potential added value of Di Luca et al. (2013a) is based on a perfect model approach for the determination of the importance of fine-scale features in the model (and in this paper observational) fields. Since this concept is very important for the paper, its main points are summarized here: To calculate the potential added value the high-resolution model results are aggregated to coarser grid widths. These newly generated coarse grid models are considered to be the virtual driving model or the “virtual global climate model (VGCM)” (Di Luca et al. 2013a). The statistics of these newly generated fields are assumed to behave similar to those from a real driving coarser grid model.

To uncover the differences between the regional climate model and its corresponding VGCM for the 2 m air temperature $T_{i,k}$ the perturbation theory is applied, this is separating the perturbation $RCM_{i,k}$ from the basic state $VGCM_k$:

$$T_{i,k} = VGCM_k + RCM_{i,k} \quad (1)$$

with i, k indices representing space and time, respectively. $VGCM_k$ is the arithmetic average of field $T_{i,k}$ over the aggregated grid boxes and represents the virtual global climate model part of the high-resolution results at any time k . $RCM_{i,k}$ is the spatial fluctuation of field $T_{i,k}$ in the aggregated grid boxes and corresponds to the virtual regional climate model part.

The effect of this decomposition is illustrated in Fig. 1 for the month of January 1979. Figure 1a shows

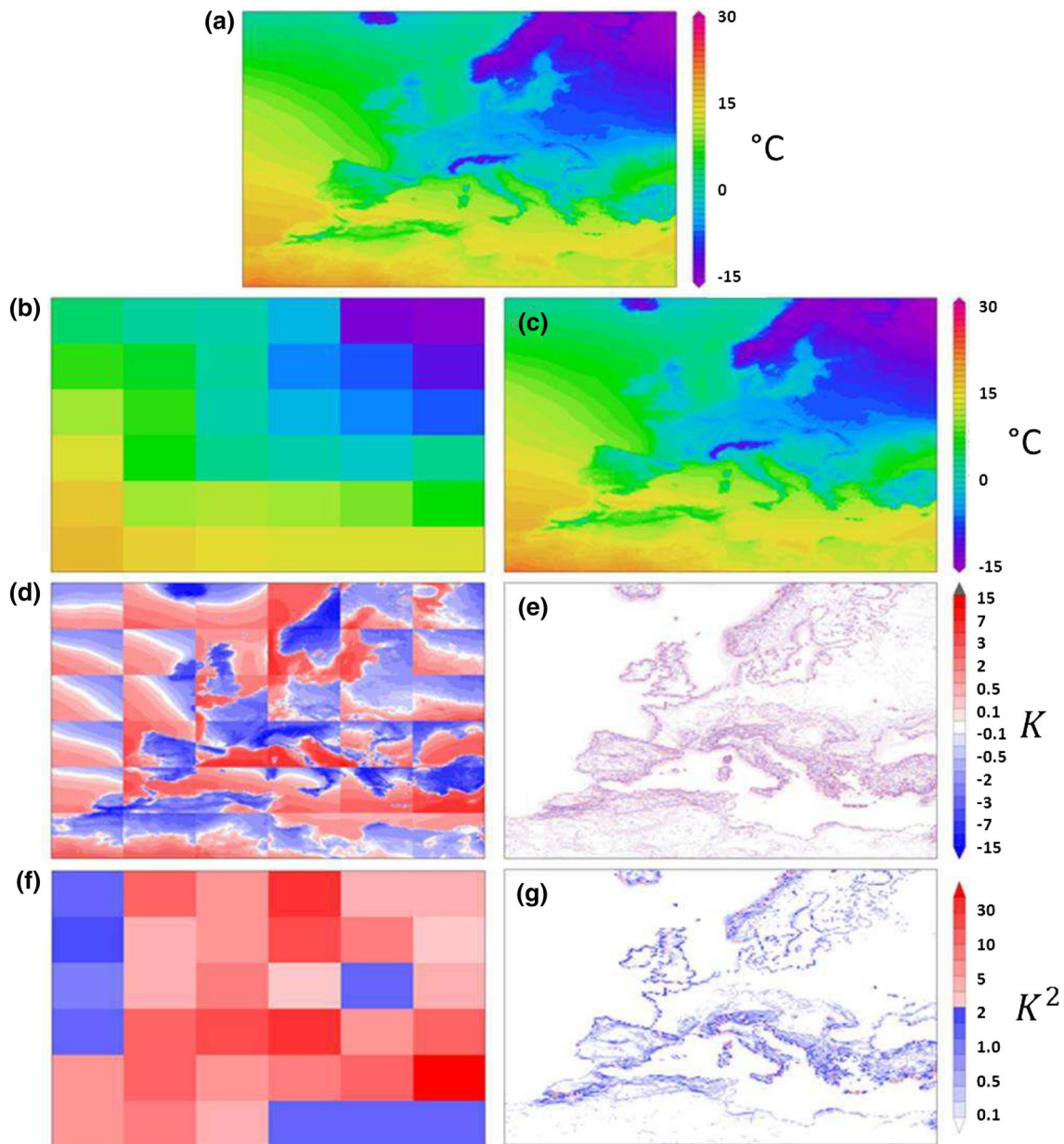


Fig. 1 Monthly mean 2 m temperature for January 1979 from CLM11 as an example of $T_{i,k}$ (a), decomposed into the virtual global climate model part $VGCM_k$ aggregated at 7.04° (b), and 0.22° (c), and

the virtual regional climate model part $RCM_{i,k}$ after aggregation at 7.04° (d), and 0.22° (e). Potential added value PAV of CLM11 at 7.04° (f) and 0.22° (g)

the monthly mean 2 m air temperature from CLM11, which represents $T_{i,k}$. In Fig. 1b, c $VGCM_k$ at 7.04° and 0.22° are shown, respectively. The differences between the 2 m temperature field at 0.11° grid and the aggregated field at 0.22° are small and visible only by a slightly lower sharpness of the temperature field in the coastal and mountainous regions. $VGCM_k$ at 7.04° does not show any information on coastlines and mountain ranges any more. A resolution of 7.04° is much coarser than the grids used in present global climate models, whose horizontal

resolution for the atmospheric model component is usually 1° to 2° (e.g. Flato et al. 2013). This resolution has been added in the figures and the further considerations for its clearness only. In Fig. 1d, e the $RCM_{i,k}$ is displayed for 7.04° and 0.22° , respectively. At the coarser grid all information on coastlines and mountains is contained in the regional fraction of the temperature field. It is interesting to note that the temperature gradient on the water surfaces of the Atlantic Ocean is visible in Fig. 1d. In contrast $RCM_{i,k}$ at 0.22° is small (Fig. 1e). Non zero

values occur only adjacent to the coastlines and in the areas with steep orography.

$VGCM_k$ and $RCM_{i,k}$ can be further separated into a stationary and a time dependent fraction, respectively (see Di Luca et al. 2013a):

$$VGCM_k = sVGCM + tVGCM_k, \tag{2}$$

$$RCM_{i,k} = sRCM_i + tRCM_{i,k}, \tag{3}$$

with $sVGCM$ and $sRCM_i$ the long-time mean of $VGCM_k$ and $RCM_{i,k}$, and $tVGCM_k$ and $tRCM_{i,k}$ being the time dependent part of $VGCM_k$ and $RCM_{i,k}$, respectively. In the potential added value study of Di Luca et al. (2013a) the variance decomposition is applied to 3-hourly, daily, and 16-day time series. In the present work this decomposition is used for long-time annual, seasonal, and monthly time series. $tVGCM_k$ indicates the temporal deviation of the VGCM from its long-time annual, seasonal, or monthly mean $sVGCM$ and thus the anomaly from the climatology of $VGCM$. $sRCM_i$ describes the stationary effects due to a higher spatial resolution and representation of mountain ranges and coastlines in $RCM_{i,k}$ compared to its VGCM. $tRCM_{i,k}$ represents the anomaly from the climatology of $RCM_{i,k}$. These regional anomalies result e.g. from deviations of the temperature difference between sea and land surfaces from its long-time climatological mean difference in coastal areas. In the case of mountainous regions $tRCM_{i,k}$ result from differences between a monthly/seasonal/annual mean of the vertical temperature gradient and its long-time climatological mean leading to differential deviations of the temperature from the climatology depending on the height level of the orography.

Finally the total variance σ^2 of $T_{i,k}$ can be written as

$$\sigma^2 = \sigma_{tVGCM_k}^2 + \sigma_{sRCM_i}^2 + \sigma_{tRCM_{i,k}}^2 + \overline{(tVGCM_k \cdot tRCM_{i,k})}^i. \tag{4}$$

The overbars denote averaging over space and time indicated by the indices i and k , respectively. Di Luca et al. (2013a) define the potential added value (PAV) from Eq. (4) as the sum of all terms containing components of $RCM_{i,k}$ after the variance decomposition of the field $T_{i,k}$, except the last covariance term in Eq. (4) Table 1. Despite this term

can be temporally and regionally quite large (see Sect. 4.4), in a long-time and spatial average it is at least ten times smaller than the other variance terms and can be omitted. Finally the potential added value PAV is defined as

$$PAV \approx \sigma_{sRCM_i}^2 + \sigma_{tRCM_{i,k}}^2 = \overline{(sRCM_i)^2}^i + \overline{(tRCM_{i,k})^2}^i \tag{5}$$

In Fig. 1f, g PAV is shown for the aggregation to the 7.04° and 0.22° grid, respectively. PAV is considerably high in the case of the aggregation to the coarse grid and comparably low for the 0.22° grid. In the latter case PAV is pronounced only along the coastlines and the mountainous regions as already seen in the $RCM_{i,k}$.

4 Results

The model runs CLM44, CLM22, and CLM11 are performed on the EURO–CORDEX domain (indicated by the black box in Fig. 2). Due to the aggregation procedure the Europe domain for the PAV investigation is slightly smaller (solid red border box). The data for these runs are available for the time period from 1979 to 2013. All analyses for the Europe domain refer to this time range. Due to the choice of the numerical grids, the aggregation is executed on the rotated grid of the COSMO–CLM runs.

The domain surrounded by the dotted red line in Fig. 2 indicates the Central Europe high resolution (CE) domain, where the CLM025 run and HYRAS are additionally investigated. CLM025 covers mainly flat and hilly terrain in most parts of its domain, only the southernmost part is dominated by the high mountain range of the Alps. Therefore, this domain can be considered as representative for the continental areas of Europe. Due to data availability the common time period covers 1995–2006 only. In addition CLM025 and HYRAS data are provided on different grids compared to CLM44, CLM22, and CLM11. Therefore all COSMO–CLM results are interpolated to the kilometer based grid of the HYRAS data before calculating PAV for the CE domain.

Table 1 Summary of the abbreviations used in the equations and in the text

Name	Equations	Description/explanation
$VGCM_k$	(1)	Total virtual global climate model part of $T_{i,k}$, Time series of spatial means
$sVGCM$	(2)	Stationary part of spatial means $VGCM_k$
$tVGCM_k$	(2)	Time dependent or transient part of spatial means $VGCM_k$
$RCM_{i,k}$	(1)	Time series of spatial deviations of $T_{i,k}$ from $VGCM_k$
$sRCM_i$	(3)	Stationary part of spatial deviations $RCM_{i,k}$
$tRCM_{i,k}$	(3)	Time dependent or transient part of spatial deviations $RCM_{i,k}$
PAV	(5)	Potential added value

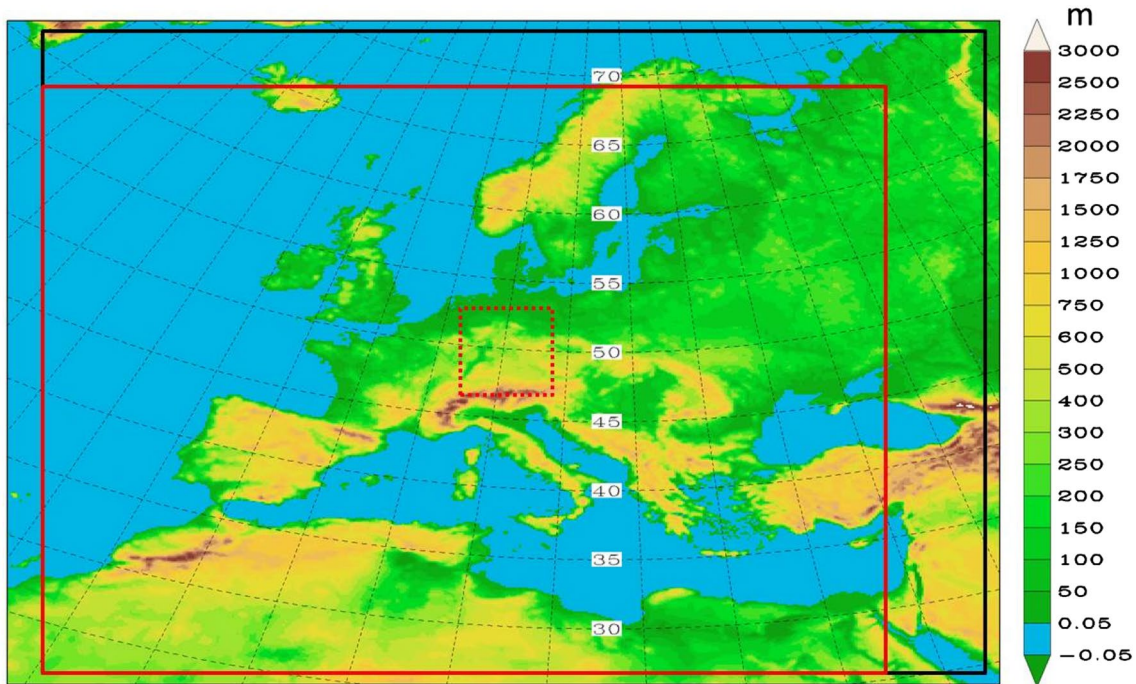


Fig. 2 Considered domains: *black rectangle* EURO-CORDEX model domain, *solid red rectangle* Europe domain of PAV considered for CLM44, CLM22, and CLM11 only, *dotted red rectangle* Central

Europe high resolution domain of PAV considered additionally for CLM025 and HYRAS data

Both domains, the Europe as well as the CE domain, are located in a sufficient distance to the sponge layers of the model areas.

4.1 Validation of COSMO-CLM against HYRAS by the stationary part of virtual global climate model results *sVGCM*

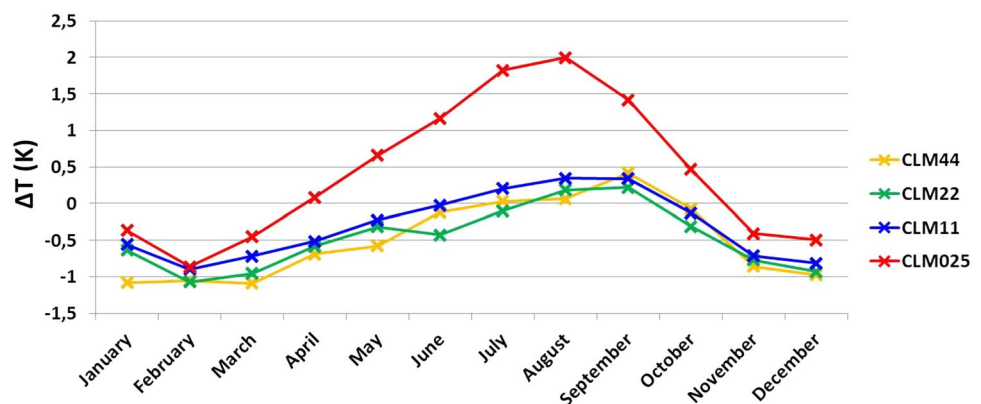
4.1.1 Monthly mean 2 m temperature, Central Europe domain

In Fig. 3, the annual cycle of the area averaged monthly mean deviations ΔT of *sVGCM* for the COSMO-CLM

simulations with respect to the HYRAS data averaged over 1996–2005 are shown.

The results for CLM44, CLM22, and CLM11 show a similar annual cycle of ΔT with negative values of about -1 K in winter and up to $+0.5$ K in late summer. These deviations represent the mean bias of the model with respect to the observational data. This means that the winter months in these COSMO-CLM runs are calculated slightly too cold and the late summer months slightly too warm in the considered region and time period, which is in agreement with the results shown in Fig. 2 in Kotlarski et al. (2014). These characteristics seem to be only weakly dependent on the resolution of the pristine numerical grids

Fig. 3 Annual cycle of the area averaged monthly deviations ΔT of *sVGCM* between the COSMO-CLM runs (CLM44, CLM22, CLM11 and CLM025) and the HYRAS data for the CE domain



used, but with the lowest bias in CLM11 in the annual mean. CLM025 shows a similar behaviour in winter, but a strong positive bias of up to +2 K in the summer half year. During summer time this run is about 1.5 K warmer than the COSMO–CLM runs with parameterized deep convection, whereas the differences in winter are quite low. In total, the annual mean of CLM025 is about 0.6 K higher than in CLM11 and hence CLM025 is the only run with a positive annual bias. These results are in accordance with Ban et al. (2014) who found a positive 2 m air temperature bias of about +2 K in southern Germany in summer in the convection resolving run and only +0.5 K in the run using the mass–flux Tiedtke scheme. Ban et al. (2014) explain this effect by the lower cloud coverage in the high-resolution run and consequently a higher global solar radiation. In winter, the stratification of the near-surface atmosphere is more stable and the role of convection is greatly reduced resulting in smaller differences between the model runs with and without parameterized convection than in summer.

4.2 Potential added value PAV

As shown in Eq. (5) the potential added value *PAV* can be interpreted as a measure of variance and variability of a considered field quantity inside the aggregated grid boxes.

The absolute values of *PAV* cannot be interpreted and judged quantitatively. The *PAV* strongly depends on the region, time period, and season under consideration. It is only qualitatively meaningful in the context with other *PAV* values calculated for other comparable regions, time periods, seasons, and different degrees of grid aggregation. Di Luca et al. (2013a) suggest to use a normalized form of the *PAV*, but this normalization changes and strongly weakens the annual cycle of the *PAV* due to the additional component $tVGCM_k$, which is needed for the normalization.

4.2.1 Seasonal mean 2 m temperature, Europe domain

In Fig. 5, the potential added value *PAV* averaged over the Europe domain is shown for the winter (DJF) and summer (JJA) as a function of the aggregation steps. The shape of the curves seems to exponentially increase with increasing aggregation grid width. Hence, considering seasonal means of the 2 m air temperature *PAV* can be found in the COSMO–CLM results on all considered grids.

The *PAV* in winter is about twice as high as in summer. The contributions to the *PAV* are, as can be seen in Fig. 1, mostly coming from the higher and more realistic resolution of the coastlines and mountain ranges and the corresponding better representation of processes related to these formations. The majority of the *PAV* is coming from the coastal effects and these seem to be strongest during

winter and less in summer, when the temperature differences between the water and the adjacent land surfaces are often less than in winter.

The *PAV* is generally highest for CLM11 and lowest for CLM44. The differences in *PAV* are maximal for the aggregations on 0.88° and 1.76°, whereas they nearly vanish in the case of an aggregation to a very coarse grid (7.04°). The *PAV* quantities of CLM11 and CLM22 differ less than between CLM22 and CLM44. Hence, the gain of *PAV* enhancing the grid resolution from 0.44° to 0.22° is higher than from 0.22° to 0.11°. Similar results concerning the real added value have been found by Lucas–Picher et al. (2016): As can be seen in Fig. 5 in this paper the differences in orographical structures between the 0.44° and 0.22° runs are much higher than between the 0.22° and the 0.11° runs. Consequently the gain in real added value concerning precipitation is higher changing the horizontal grid resolution from 0.44° to 0.22° than from 0.22° to the even finer resolution of 0.11°. Lucas–Picher et al. (2016) explain their results with the similarity of the orography in the 0.44° run and the driving model, which does not allow to add really much more information to the RCM results compared to the driving data. First the orography on the 0.22° and finer grids show more detailed characteristics along the coastlines and in the mountainous regions, which are needed for the simulation of smaller scale phenomena.

4.2.2 Seasonal mean 2 m temperature, Central Europe domain

Figure 5 shows a similar analysis but limited to the CE domain for the winter (DJF) and summer (JJA). In addition to the curves shown in Fig. 4 the *PAV* of CLM025 as well as the HYRAS data are included. As described in Sect. 2.2 the results for HYRAS can be considered as a kind of qualitative validation for the corresponding model results.

The shape of the *PAV* curves is similar to the corresponding graphs for the Europe domain. But the increase of *PAV* with increasing degree of aggregation is weaker. In contrast to the Europe domain, in the CE region the *PAV* for the winter months is smaller than for summer. The difference between the summer and winter *PAV* increases with growing aggregation steps. The largest differences between summer and winter can be seen in CLM025, where the summer *PAV* is more than twice the winter value. The reason for this change in annual *PAV* cycle behaviour may be the absence of coastlines in the CE region, so that the *PAV* in this domain only originates from the effect due to the resolution of mountain ranges. In summer, the lapse rate is maximal and hence in aggregated grid cells containing steep and variable orography the aggregation is followed by a high variance of temperature leading to higher *PAV* values as in winter. This effect is in addition depending

Fig. 4 Potential added value *PAV* area averaged over the Europe domain in summer (JJA, *solid lines*) and winter (DJF, *dashed lines*) for CLM44 (yellow), CLM22 (green), and CLM11 (blue)

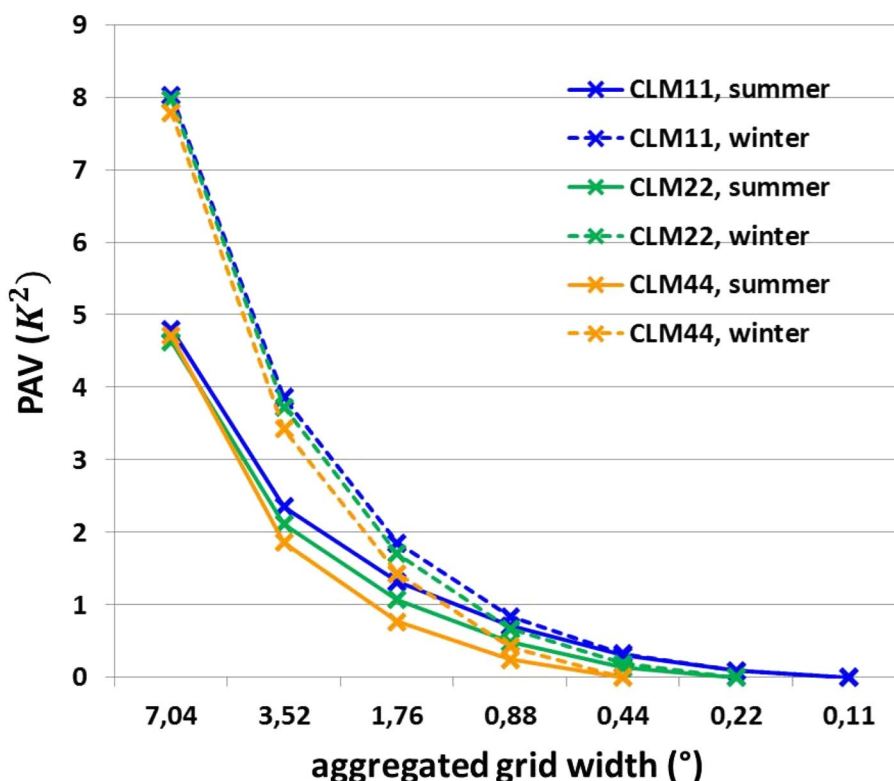
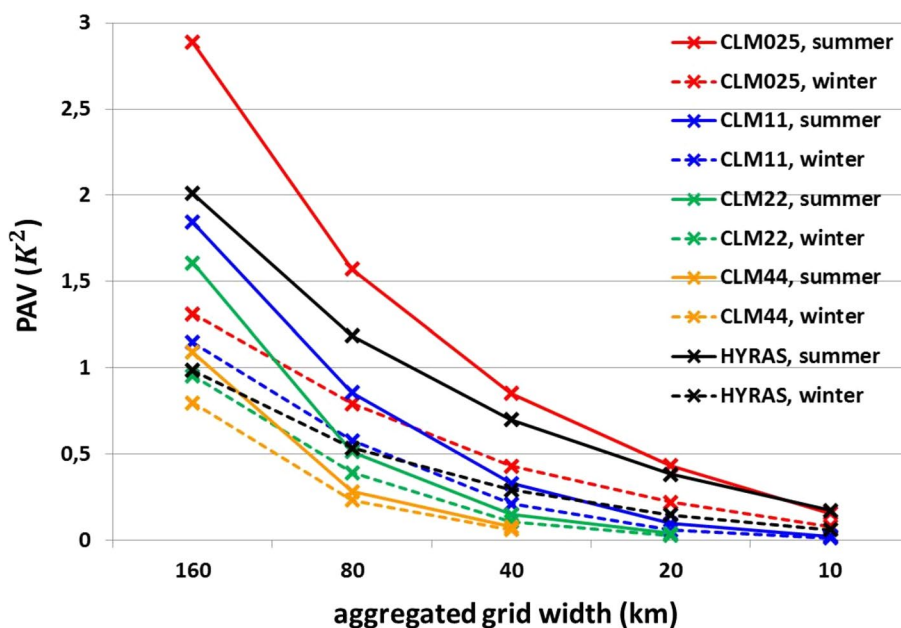


Fig. 5 Potential added value *PAV* area averaged over the Central Europe domain in summer (JJA, *solid lines*) and winter (DJF, *dashed lines*) for CLM44 (yellow), CLM22 (green), CLM11 (blue), and CLM025 (red), and the HYRAS (black) data



on the grid resolution: increasing grid resolution causes a more realistic representation of the real orographical features in the model orography and hence a higher variance of the temperature in the aggregated grid cells. In the CLM025 run the lapse rate is even stronger as in the runs with parameterized convection due to the lower cloud cover and higher shortwave insolation leading to a *PAV* increase

caused by the model physics in addition to the effect of grid resolution.

Considering the CE domain the differences in *PAV* between the runs CLM11–CLM22 and CLM22–CLM44 are not as clear as in the Europe domain. The potential added value in CLM025 is considerably higher than in CLM11 due to the effect of higher grid resolution and in

summer due to the above explained additional effect of the changed model physics.

The *PAV* calculated in the same manner for the gridded observational HYRAS data shows similar characteristics as for the COSMO-CLM runs concerning the increase with increasing aggregation and the annual cycle. But the lines for the HYRAS data are less steep than the *PAV* curves for the model results. They show *PAV* quantities similar to CLM025 for the aggregation to 10 km, but much lower values at aggregations to coarser grids. A reason for these different characteristics might be the calculation of the data, which is done by numerical integration of prognostic equations on a numerical grid in the model and by horizontal interpolation of irregular distributed observations on a grid in the HYRAS data (Rauthe et al. 2013).

4.2.3 Monthly mean 2 m temperature, Europe domain

In Fig. 6, the annual cycle of the *PAV* is shown deduced from monthly means area averaged over the Europe domain for CLM11 for different aggregation steps. For CLM22 and CLM44 the annual cycle of the *PAV* is quite similar to CLM11. As seen in the previous paragraphs the *PAV* increases with increasing number of aggregation steps. A noteworthy annual cycle in *PAV* can only be seen for the aggregation on grid sizes larger than 0.44° width. The maximum of the annual cycle is in the winter months from November to February, whereas from April to October there is an extended minimum. This minimum corresponds to the time period when the difference between land and water temperatures is lower than during winter.

4.2.4 Monthly mean 2 m temperature, Central Europe domain

Figure 7 shows the annual cycle of the *PAV* for all COSMO-CLM runs performed on the numerical grids

with different horizontal resolutions and the observational HYRAS data aggregated to 160 km over the CE domain in the time period 1995–2006 based on monthly means of *PAV*. The CLM11, CLM22, and CLM44 runs tend to a similar annual cycle with the absolute maximum in April and a broad maximum lasting until September. The remaining autumn and the winter months are dominated by a broad minimum. This annual cycle is totally different from the corresponding results for the Europe domain, where the maximum is in winter (see Fig. 6). This may be due to the absence of the *PAV* originating from the coastlines in the CE domain. Nevertheless, the absolute values of the *PAV* in the Europe domain at the aggregation to 1.76° and the CE domain at the aggregation to 160 km (which are nearly analog) are of the same range of about $1.5\text{--}2\text{ K}^2$.

A similar annual cycle can be found in the aggregated HYRAS data, which confirms the results found in the COSMO-CLM runs.

The *PAV* quantities of CLM025 are distinct in range and temporal development. They are generally higher than the *PAV* deduced from the other COSMO-CLM results and from the HYRAS data. The annual cycle shows a maximum level of *PAV* extending from April until September with the maximum values calculated for July and August, which corresponds with the summer maximum of *PAV* in Fig. 5.

In winter, the lapse rate is generally lower than in summer, hence the variance of temperature in the aggregated grid cells is lower resulting in smaller *PAV* values. In the run with non-parameterized deep convection generally the cloud cover is lower resulting in a higher global radiation, stronger heating in the planetary boundary layer (see Fig. 3), and a higher lapse rate, which is followed by a higher temperature variance and *PAV* in the aggregated grid cells. Due to the yearly cycle of deep convection this effect is strongest in summer and minimal in winter.

Fig. 6 Monthly mean potential added value *PAV* area averaged over the Europe domain for CLM11 aggregated to different resolutions, time period 1979–2013

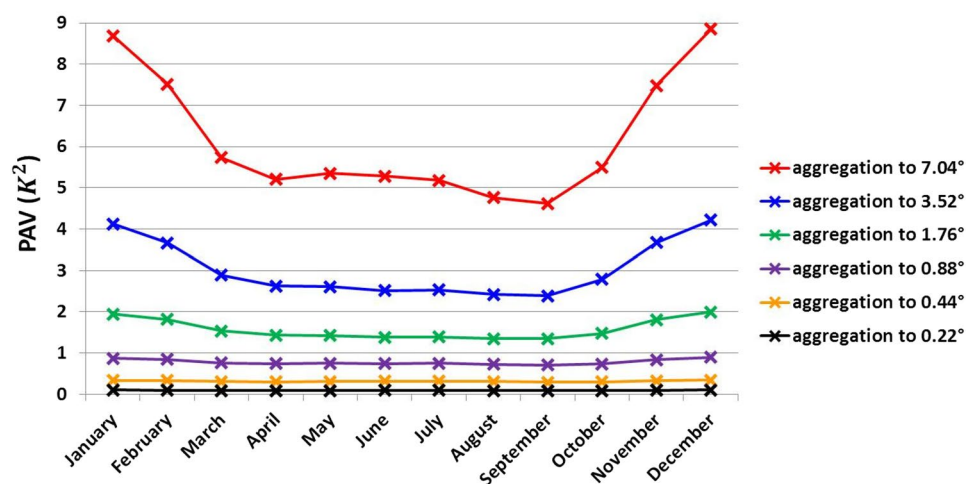
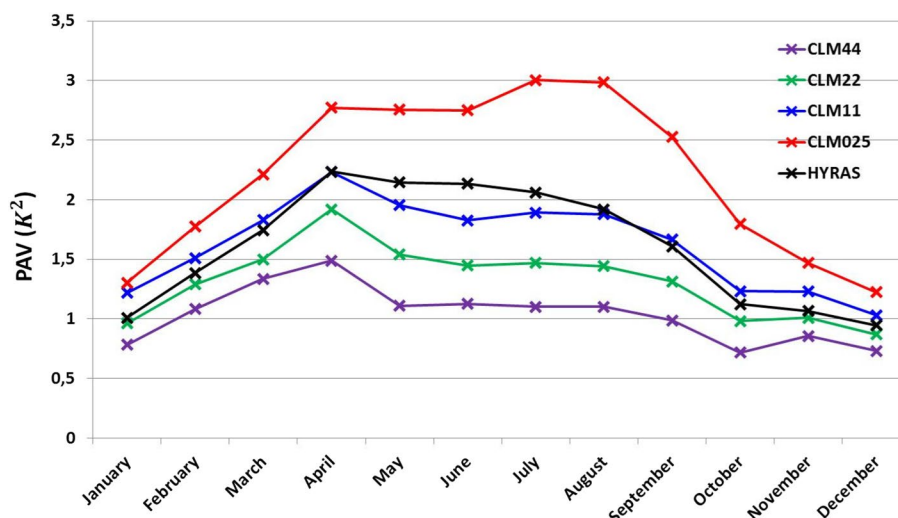


Fig. 7 Monthly mean potential added value PAV area averaged over the CE domain for CLM44, CLM22, CLM11, CLM025, and HYRAS aggregated to 160 km, time period 1995–2006



4.3 Components of the potential added value $\sigma_{sRCM_i}^2$ and $\sigma_{tRCM_{i,k}}^2$

The potential added value PAV can be subdivided into the variances of the stationary $sRCM_i$ and the time dependent part $tRCM_{i,k}$ (see Eq. 5). The subdivision of PAV into its stationary and transient fraction allows to estimate if the origin of the PAV is predominantly coming from its stationary part or if it is produced by transient synoptic conditions.

4.3.1 Annual mean 2 m temperature, Central Europe domain

Considering the PAV components $\sigma_{sRCM_i}^2$ and $\sigma_{tRCM_{i,k}}^2$ on the basis of annual means the stationary component dominates the total PAV by more than 99%. This percentage is independent of the number of aggregation steps (not shown here).

4.3.2 Monthly mean 2 m temperature, Central Europe domain

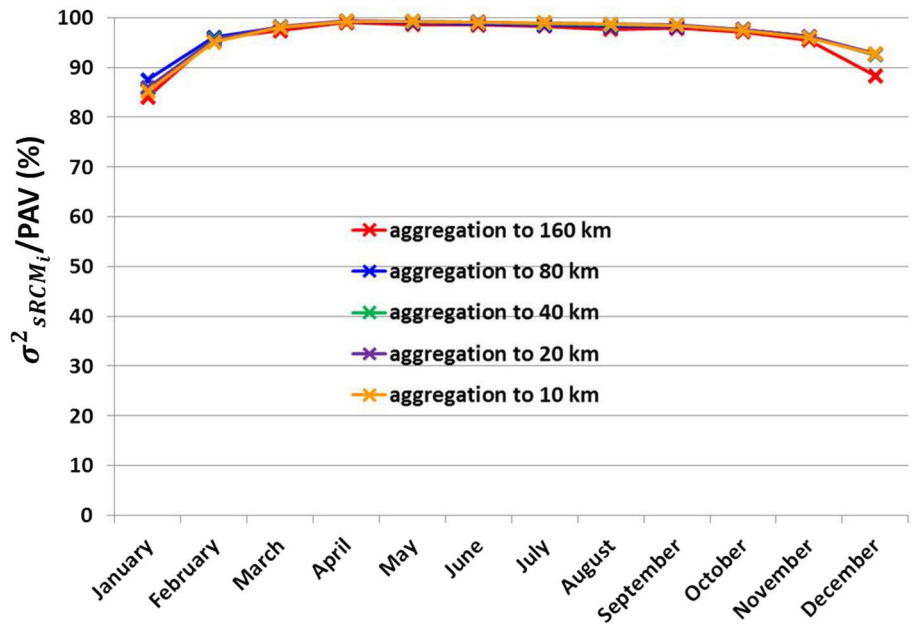
The subdivision of the total PAV in a stationary and a time dependent component yield a very high percentage of the stationary part from March to September of more than 98% (see Fig. 8). Due to the absence of coastal effects in the CE domain both PAV components should be induced by the orographical effect only. From March to September, the vertical stratification of the atmosphere in the lower and middle parts of the troposphere seems to be quite uniform and the variations within the months are relatively small resulting in a very small transient fraction $\sigma_{tRCM_{i,k}}^2$ of only about 1–2%. The high percentage of $\sigma_{sRCM_i}^2$ on PAV is only

slightly dependent on the degree of aggregation with the lower percentages at aggregation to the coarser grids.

In the winter months from November to February the percentage of the stationary part $\sigma_{sRCM_i}^2$ is considerably lower than in summer, but nevertheless very high with all values exceeding 84%. This means that the transient part $\sigma_{tRCM_{i,k}}^2$ is much higher and amounts up to 16% compared to a maximum of 2% in the summer months. The reason for this might be the higher variability of the vertical stratification in the CE domain during winter. This variability strongly depends on the type of large-scale synoptic situations. During windy conditions the temperature usually decreases with height, whereas during calm wind situations the near-surface layers cool and a strong stable stratification may develop below the lowest 1000–1500 m resulting in an only moderate temperature decrease or even increase with increasing height. The dependency of the subdivision of PAV into its stationary and time dependent components on the aggregation is higher in winter than in summer, especially in December and January.

The percentage of the stationary and transient part of the PAV is strongly dependent on the time scales considered. The order of both components is similar based on 3-hourly data indicating the influence of synoptic conditions on the daily temperature cycle and its variation in space and time (Di Luca et al. 2013a). The percentage of $\sigma_{tRCM_{i,k}}^2$ deduced from monthly means is much smaller than the percentage calculated from 3-hourly data. The reason for this smaller percentage may be the fact that on this long temporal scale only variations in the synoptic conditions lasting for at least several days can influence the monthly means and the effect on the daily temperature cycle is not included.

Fig. 8 Percentage of stationary fraction $\sigma^2(sRCM_i)$ from total PAV on the base of area averaged monthly mean 2 m temperature over the CE domain for CLM11 on different degrees of aggregation, time period 1995–2006



4.3.3 Monthly mean 2 m temperature, Europe domain

Considering the annual cycle of the percentage of the stationary part $\sigma^2_{sRCM_i}$ of the total PAV for the entire Europe domain (Fig. 9) reveals differences in the results compared to the CE domain: Generally the percentage of $\sigma^2_{sRCM_i}$ is between 93 and 97% and hence very high throughout the year. Especially in winter the percentage of $\sigma^2_{sRCM_i}$ is higher in the Europe than in the CE domain, probably due to the presence of coastal areas in the Europe domain, which yield a high stationary PAV due to the climatological mean temperature difference between

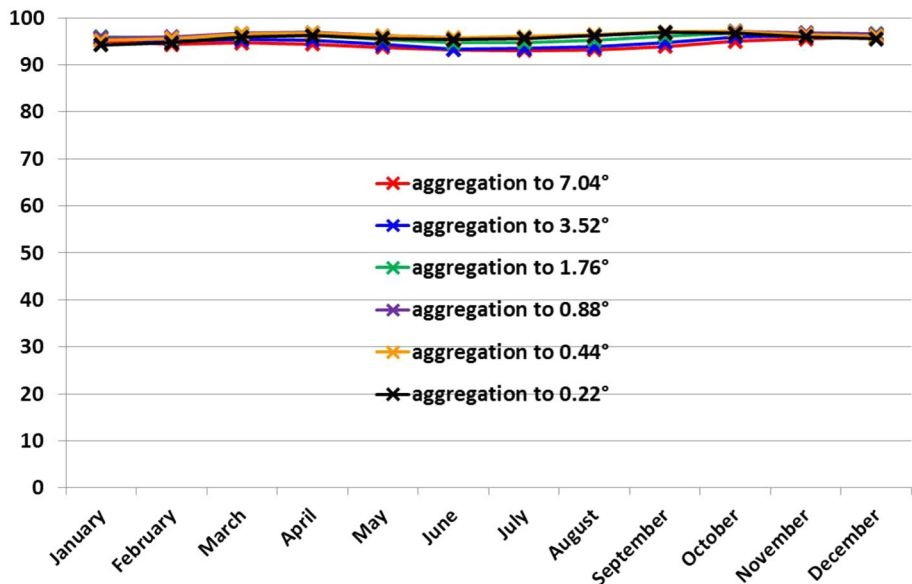
land and nearby sea surfaces in addition to the above described PAV resulting from orographical effects.

4.4 Covariance term $\overline{(tVGCM_k \cdot tRCM_{i,k})^2}^i$

4.4.1 Monthly mean 2 m temperature, Europe domain

In the definition of the potential added value from Eqs. 4 to 5 the covariance term $\overline{(tVGCM_k \cdot tRCM_{i,k})^2}^i$ is

Fig. 9 Percentage of stationary fraction $\sigma^2(sRCM_i)$ from total PAV based on area averaged monthly mean 2 m temperature over the Europe domain for CLM11 on different degrees of aggregation, time period 1979–2013



omitted according to Di Luca et al. (2013a). This term gives a description of the interaction between the temporal deviations of $VGCM_k$ and $RCM_{i,k}$.

Averaging this term over a larger domain as e.g. the CE or even the Europe domain yields very small contribution of this term to the PAV , which is on the order of at least 10,000 smaller than $\sigma_{sRCM_i}^2$ and about 1000 smaller as $\sigma_{tRCM_{i,k}}^2$. Considering annual, seasonal, and monthly area averages neglecting this term is justified in the present case.

Figure 10 shows the spatial distribution of the covariance term averaged over January during the time period 1979–2013. Considering each aggregated grid box it seems that the area mean of the covariance term in each grid cell is close to 0 K^2 , and hence the summation over all aggregated grid boxes will result in a very small contribution of this term to the total PAV . But on local scale the covariance term can be quite large, up to ten times larger than the area averaged monthly means of PAV . The covariance term is positive where the virtual global anomaly $tVGCM_k$ is enhanced by the temporal fraction of the potential added value $tRCM_{i,k}$, and negative where $tRCM_{i,k}$ weakens $tVGCM_k$. This can be seen in Fig. 10 e.g. over the Baltic Sea, where the balancing effect of the sea water damps the negative as well as the positive virtual global anomalies towards 0 K . On the surrounding land masses negative $tVGCM_k$ are usually linked to low cloud coverage conditions and consequently to a strong cooling in these regions in January due to radiation energy loss. These negative anomalies are obviously enhanced by the transient part of the PAV and lead to a strong positive covariance term.

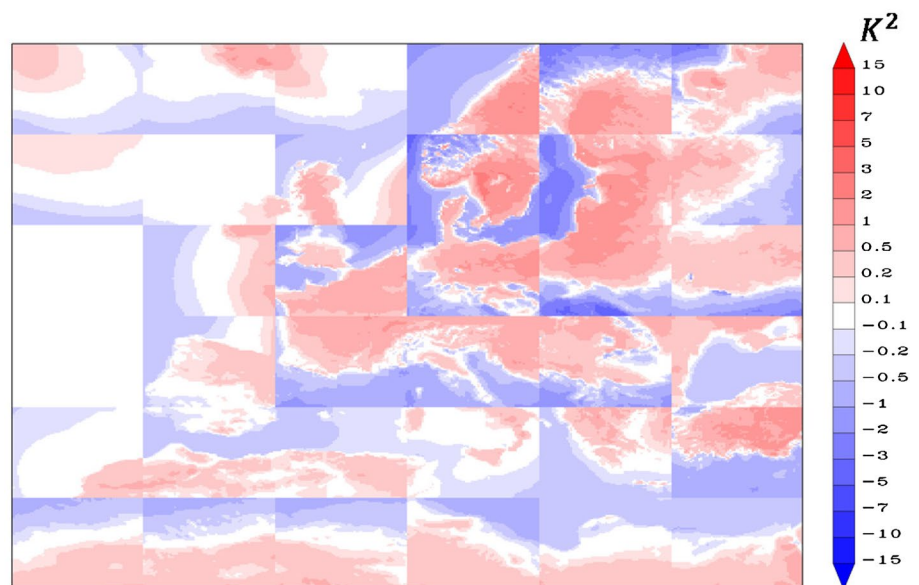
5 Summary and conclusions

In the present work, the potential added value concept of Di Luca et al. (2013a) is applied to monthly, seasonal, and annual means of the 2 m air temperature resulting from ERA Interim forced COSMO–CLM runs performed on numerical grids with different grid widths. The PAV analysis is performed on two domains: one domain covers large parts of (1) Europe and the other one (2) Central Europe (CE) only. The time period investigated on the Europe domain is 1979–2013 whereas the period for the CE region extends from 1995 to 2006.

The validation of the stationary part of the virtual global climate model $sVGCM$ shows generally no dependence on the aggregation steps. In the CE domain the area averaged COSMO–CLM results have a negative bias (about -1 K) during the winter and a positive bias (up to $+2\text{ K}$) in summer. The summer bias for CLM44, CLM22, and CLM11 is only about $+0.5\text{ K}$ and hence these results confirm the outcome of Kotlarski et al. (2014). The highest positive summer bias is resulting from the model run with directly calculated deep convection due to a lower cloud cover and a higher shortwave radiation. It coincides with the results of Ban et al. (2014).

The results show that PAV is contained in all COSMO–CLM simulations presented. For each run the PAV increases with increasing aggregation step for annual, seasonal, and monthly mean values. The gain of PAV is slightly higher by increasing the resolution of the numerical model grid from 0.44° to 0.22° than from 0.22° to 0.11° in the temporal mean, which is consistent with the findings

Fig. 10 Map of the covariance term $(tVGCM_k \cdot tRCM_{i,k})^2\text{ K}^2$ of the 2 m temperature resulting from CLM11 aggregated at the 7.04° grid for January during the time period 1979–2013 in K^2 over the Europe domain



of Lucas–Picher et al. (2016) for the real added value with respect to precipitation. Lucas–Picher et al. (2016) explain their results with the similarity of the orography in the 0.44° run and the driving model. The orography on the 0.22° and finer grids show more detailed characteristics along the coastlines and in mountainous regions, which are needed for the simulation of smaller scale phenomena. Regarding the Europe domain *PAV* in winter is generally about twice as high as in summer, whereas for the CE domain without the effects of the coastlines the maximum of *PAV* occurs in the summer season. The *PAV* is highest in the CLM025 run. These results are not only a consequence of the high resolution model grid, but in addition an effect of the direct calculation of deep convection. In winter convection processes are on its minimum and hence the influence of the physics on the *PAV* is minimal leading to only slightly higher *PAV* values than in CLM11. But in summer, the *PAV* in CLM025 is much higher than in CLM11 showing the additional effect of the changed deep convection treatment in the model run, which is on its maximal contribution. The differences in the annual cycle of *PAV* in the two domains may be due to the absence of coastlines in the CE domain as well as unequal time periods considered.

Due to its definition [see Eq. (5)] the *PAV* is the sum of the variance of the stationary and the time dependent parts of the spatial deviation $RCM_{i,k}$ from the arithmetic average of the considered quantity and hence can be understood as variability in the aggregated grid cells. In the definition and the calculation of the *PAV* neither data from observations nor from a driving coarser scale model are involved. So there is no direct link from the *PAV* to the real added value and no statement about real added value can be deduced from the *PAV*. The presence of *PAV* does not imply that there is some real added value in the model results. But the absence of *PAV* would be a proof for the absence of variability in an aggregated grid and hence for the lack of real added value.

The *PAV* considerations do not give a strict rule which horizontal grid resolution would be the best. It shows only that downscaling makes sense, but that the gain of *PAV* due to a selection of a higher grid resolution decreases with increasing resolution. An increase in grid resolution for RCM calculations is very demanding concerning the computational resources, especially for computing power and data storage. Each user has to decide by himself if the gain of *PAV* and perhaps real added value due to a finer grid resolution is possible and worth the strongly increasing computational costs.

The *PAV* has been calculated besides the COSMO–CLM results in addition for the gridded observational HYRAS data set. The *PAV* of a gridded observational data set is expected to be in the same order as the *PAV* of RCM results, despite the calculation of gridded observational data sets is

based on totally different concepts than the calculation of RCM data. Hence a different behavior of *PAV* with respect to grid aggregation can be seen in the results. In the future, the *PAV* could be perhaps used as a measurement for variability in the creation of gridded observational data sets.

The percentage of the stationary part of *PAV* is generally higher than 80%. The annual cycle of the percentage of both *PAV* components strongly depends on the domain and the time period of the data base used for the *PAV* calculation.

The covariance term $\overline{(tVGCM_k \cdot tRCM_{i,k})^2}^i$, which represents a temporally mean and has been omitted by Di Luca et al. (2013a), can locally be of considerable amount, but averages nearly to 0 K^2 in each grid box of the aggregated grid and hence in the entire domain. Therefore the disregard of this term is justified in the present cases.

Acknowledgements We thank the German Federal Ministry of Education and Research (BMBF) for supporting this study within the framework of the research programme ‘Mittelfristige Klimaprognosen’ (MiKlip) (<http://www.fona-miklip.de/en/>) under Grant Number 01LP1154A and two unknown reviewers for their useful comments. Further we thank the German climate computing center (DKRZ) for providing computing time for some of the COSMO–CLM runs and the ERA reanalysis data for initialization and driving. We appreciate Dr. Susanne Brienens for providing the data for the 0.025 degrees COSMO–CLM run and Dr. Monika Rauthe for the gridded observational HYRAS data set.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ban N, Schmidli J, Schär C (2014) Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J Geophys Res Atmos* 119:7889–7907. doi:10.1002/2014JD021478
- Böhm U, Kücken M, Ahrens W, Block A, Hauffe D, Keuler K, Rockel B, Will A (1996) CLM—the climate version of LM: brief description and long-term applications. *COSMO Newsl* 6:225–235
- Brienens S, Früh B, Walter A, Trusilova K, Becker P (2016) A Central European precipitation climatology—Part II: application of the high-resolution HYRAS data for COSMO–CLM evaluation. *Met Z* 25:195–214. doi:10.1127/metz/2016/0617
- Dee DP et al (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Q J R Meteorol Soc* 137:553–597. doi:10.1002/qj.828
- Di Luca A, de Elía R, Laprise R (2012) Potential for added value in precipitation simulated by high-resolution nested regional climate models and observations. *Clim Dyn* 38:1229–1247. doi:10.1007/s00382-011-1068-3

- Di Luca A, de Elía R, Laprise R (2013a) Potential for added value in temperature simulated by high-resolution nested RCMs in present climate and in the climate change signal. *Clim Dyn* 40:443–464. doi:[10.1007/s00382-012-1384-2](https://doi.org/10.1007/s00382-012-1384-2)
- Di Luca A, de Elía R, Laprise R (2013b) Potential for small scale added value of RCM's downscaled climate change signal. *Clim Dyn* 40:601–618. doi:[10.1007/s00382-012-1415-z](https://doi.org/10.1007/s00382-012-1415-z)
- Di Luca A, de Elía R, Laprise R (2015) Challenges in the quest for added value of regional climate dynamical downscaling. *Curr Clim Change Rep* 1:10–21. doi:[10.1007/s40641-015-0003-9](https://doi.org/10.1007/s40641-015-0003-9)
- Dosio A, Panitz H-J, Schubert-Frisius M, Lüthi D (2015) Dynamical downscaling of CMIP5 global circulation models over CORDEX-Africa with COSMO-CLM: evaluation over the present climate and analysis of the added value. *Clim Dyn* 44:2637–2661. doi:[10.1007/s00382-014-2262-x](https://doi.org/10.1007/s00382-014-2262-x)
- Feser F (2006) Enhanced detectability of added value in limited-area model results separated into different spatial scales. *Mon Weather Rev* 134:2180–2190. doi:[10.1175/MWR3183.1](https://doi.org/10.1175/MWR3183.1)
- Feser F, Rockel B, von Storch H, Winterfeldt J, Zahn M (2011) Regional climate models add value to global model data. *Bull Am Meteorol Soc* 92:1181–1192. doi:[10.1175/2011BAMS3061.1](https://doi.org/10.1175/2011BAMS3061.1)
- Flato G et al (2013) Evaluation of climate models. In: Stocker TF et al (eds) *Climate Change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA: 2013, pp 741–866
- Jacob D et al (2014) EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg Environ Change* 14:563–578. doi:[10.1007/s10113-013-0499-2](https://doi.org/10.1007/s10113-013-0499-2)
- Kendon EJ, Roberts NM, Senior CA, Roberts MJ (2012) Realism of rainfall in a very high-resolution regional climate model. *J Clim* 25:5791–5806. doi:[10.1175/JCLI-D-11-00562.1](https://doi.org/10.1175/JCLI-D-11-00562.1)
- Kotlarski S et al (2014) Regional climate modelling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci Model Dev Discuss* 7:217–293. doi:[10.5194/gmdd-7-217-2014](https://doi.org/10.5194/gmdd-7-217-2014)
- Lucas-Picher P, Laprise R, Winger K (2016) Evidence of added value in north American regional climate model hindcast simulations using ever-increasing horizontal resolutions. *Clim Dyn*. doi:[10.1007/s00382-016-3227-z](https://doi.org/10.1007/s00382-016-3227-z)
- Prein AF et al (2013) Added value of convection permitting seasonal simulations. *Clim Dyn* 41:2655–2677. doi:[10.1007/s00382-013-1744-6](https://doi.org/10.1007/s00382-013-1744-6)
- Prömmel K, Geyer B, Jones JM, Widmann M (2010) Evaluation of the skill and added value of a reanalysis-driven regional simulation for Alpine temperature. *Int J Climatol* 30:760–773. doi:[10.1002/joc.1916](https://doi.org/10.1002/joc.1916)
- Rauthe M, Steiner H, Riediger U, Mazurkiewicz A, Gratzki A (2013) A central European precipitation climatology - Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS). *Meteorol Z* 22:235–256. doi:[10.1127/0941-2948/2013/0436](https://doi.org/10.1127/0941-2948/2013/0436)
- Rockel B, Will A, Hense A (2008) The regional climate model COSMO-CLM (CCLM). *Meteorol Z* 8:347–348. doi:[10.1127/0941-2948/2008/0309](https://doi.org/10.1127/0941-2948/2008/0309)
- Uppala SM et al (2005) The ERA-40 re-analysis. *Q J R Meteorol Soc* 131:2961–3012. doi:[10.1256/qj.04.176](https://doi.org/10.1256/qj.04.176)
- Winterfeldt J, Weisse R (2009) Assessment of value added for surface marine wind speed obtained from two regional climate models. *Mon Weather Rev* 137:2955–2965. doi:[10.1175/2009MWR2704.1](https://doi.org/10.1175/2009MWR2704.1)