CrossMark

How well can a convection-permitting climate model reproduce decadal statistics of precipitation, temperature and cloud characteristics?

Erwan Brisson[1,2] · Kwinten Van Weverberg[3,4] · Matthias Demuzere[1] ·
Annemarie Devis[1] · Sajjad Saeed[1] · Martin Stengel[5] · Nicole P. M. van Lipzig[1]

Abstract Convection-permitting climate model are promising tools for improved representation of extremes, but the number of regions for which these models have been evaluated are still rather limited to make robust conclusions. In addition, an integrated interpretation of near-surface characteristics (typically temperature and precipitation) together with cloud properties is limited. The objective of this paper is to comprehensively evaluate the performance of a 'state-of-the-art' regional convection-permitting climate model for a mid-latitude coastal region with little orographic forcing. For this purpose, an 11-year integration with the COSMO-CLM model at Convection-Permitting Scale (CPS) using a grid spacing of 2.8 km was compared with in-situ and satellite-based observations of precipitation, temperature, cloud properties and radiation (both at the surface and the top of the atmosphere). CPS clearly improves the representation of precipitation, in especially the diurnal cycle, intensity and spatial distribution of hourly precipitation. Improvements in the

representation of temperature are less obvious. In fact the CPS integration overestimates both low and high temperature extremes. The underlying cause for the overestimation of high temperature extremes was attributed to deficiencies in the cloud properties: The modelled cloud fraction is only 46 % whereas a cloud fraction of 65 % was observed. Surprisingly, the effect of this deficiency was less pronounced at the radiation balance at the top of the atmosphere due to a compensating error, in particular an overestimation of the reflectivity of clouds when they are present. Overall, a better representation of convective precipitation and a very good representation of the daily cycle in different cloud types were demonstrated. However, to overcome remaining deficiencies, additional efforts are necessary to improve cloud characteristics in CPS. This will be a challenging task due to compensating deficiencies that currently exist in 'state-of-the-art' models, yielding a good representation of average climate conditions. In the light of using the CPS models to study climate change it is necessary that these deficiencies are addressed in future research.

✉ Erwan Brisson
  erwan.brisson@gmail.com

1 Physical and Regional Geography Research Group, Katholieke Universiteit Leuven, Leuven, Belgium

2 Institute for Atmospheric and Environmental Sciences, Goethe University Frankfurt, Frankfurt am Main, Germany

3 Met Office, Atmospheric Processes and Parametrisations, Exeter, UK

4 Université catholique de Louvain, Louvain-la-neuve, Belgium

5 Deutscher Wetterdienst (DWD), Offenbach, Germany

1 Introduction

Global circulation models (GCMs) are generally used to obtain present-day and future climate change projections for a broad range of applications such as city management, agriculture, biodiversity, geopolitical studies, etc. Although GCMs model large-scale atmospheric present-day features accurately (Hanssen-Bauer and Foerland 2001; Nieto et al. 2004; Hazeleger et al. 2010), their resolution remains too

coarse to provide detailed information at the local scale. In order to overcome this drawback, regional climate models (RCMs) are nested in a GCM (Giorgi and Bates 1989) over a limited area. The benefits of using such RCMs are twofold: First, the use of a restricted area reduces the computational cost of the model. At equal cost it is, therefore, possible to use more complex and sophisticated physical schemes. Second, because RCMs use a refined resolution, a more detailed description of surface parameters (orography, land use, etc) is provided. This may result in improvements, particularly in the representation of spatial variability.

The need for finer representations of the spatial variability has been one of the driving factors for constantly increasing RCMs' resolution, the latter being possible due to the continuous growth of computing resources. In the latest internationally coordinated projects, the typical size of RCM grid-mesh ranged from 50 to 25 km [PRUDENCE and ENSEMBLES respectively (Christensen and Christensen 2007; van der Linden and Mitchell 2009)]. This resolution increase was found to improve the representation of the precipitation probability function and more specifically the representation of small-scale convective precipitation due to a refinement of surface characteristics (Boberg et al. 2009). More recently the EURO-CORDEX program started climate integrations over Europe characterised by a 12 km grid mesh (Kotlarski et al. 2014). However, increasing the resolution does not necessarily correct for all deficiencies observed in RCMs. For example the diurnal cycle of precipitation is still characterised by a maximum occurring too early in the afternoon when refining RCM resolution from 50 to 12 km (Walther et al. 2013; Clark et al. 2007). This is mainly due to deficiencies in the convective parametrisation which does not represent some of the physical processes such as cold pools or gravity waves that organise convection on the mesoscale. This issue is improved when the convection is partly resolved which is occurring in simulations with grid mesh sizes as fine as at least ~4 km (Weisman et al. 1997). These simulations are often referred to as convective permitting scale (CPS) simulations.

A few recent studies have shown the multiple benefits of performing CPS simulations for modelling precipitation (Kendon et al. 2012; Prein et al. 2013; Ban et al. 2014; Prein et al. 2015). These studies have shown that the added value mostly occurs during convective-prone periods and for shorter time-scales (e.g., hourly or daily). On hourly time-scales, Prein et al. (2013), Chan et al. (2012) and Fosser et al. (2015) found an improved representation of precipitation diurnal cycle, intensity and occurrence in CPS compared to non-CPS simulations. Interestingly, studies do not agree on whether or not the climate sensitivity of hourly summer precipitation at CPS may differ significantly from non-CPS projections (Kendon et al. 2014; Banet al. 2015). The precipitation objects are generally smaller and more peaked in

agreement with observations (Prein et al. 2013). On daily time-scales and for summer, most of the CPS-precipitation intensity distribution does not significantly diverge from non-CPS simulations. Only the highest quantiles (i.e., above 90th or 95th) are found to improve, especially over mountainous areas (Ban et al. 2014; Chan et al. 2012). In winter, the representation of daily precipitation intensity is found to improve for CPS compared to non-CPS. However this improvement mainly results from orographic forcing and is, therefore, appearing only in areas with complex topography (Prein et al. 2013; Chan et al. 2012). In addition, Brisson et al. (2015) show that the structure of precipitation objects is improved over flat and hilly areas on daily time-scales, while for longer time-scales (i.e., monthly or longer), the added value of CPS is mostly averaged out.

Hohenegger et al. (2009), Prein et al. (2013) and Ban et al. (2014) indicate that the improved description of orography is beneficial for the spatial representation and the diurnal evolution of temperature over the Alps. Prein et al. (2013), however, found that using a simple height bias correction for a non-CPS model has about the same added value as performing CPS simulations, pointing at the limited added value of CPS simulations for temperature representation. Very little information is available in the literature for the added value of CPS for temperature over plain areas, although over these areas the differences in the added values of CPS is not dominated by the improved description of the orography. Investigating CPS simulations over such areas is therefore of great interest as possible improvements resulting from other benefits of using CPS (e.g., more detailed surface parameters, explicit resolved convection, etc.) are likely to be easier to identify.

Another benefit of using CPS lies in the representation of clouds. Non-CPS simulations often produce too large cloud fractions and too large TOA outgoing radiation (Kothe et al. 2011) and fail at reproducing the diurnal cycle of clouds (Pfeifroth et al. 2012; Langhans et al. 2013). High cloud amount are particularly overestimated in non-CPS compared to CPS (Böhme et al. 2011). Cloud fraction often decreases and incoming radiation increases at CPS (Fosser et al. 2015; Prein et al. 2013). Similarly the diurnal cycle of cloudiness improves (Langhans et al. 2013). For West Africa, mesoscale convective systems and associated cold pools are much better represented by CPS than non-CPS models, influencing the sensitivity of CPS models to changes in vegetation (Lauwaet et al. 2010). This improved representation of cold pool dynamics was also found to affect the sensitivity of mesoscale convective systems to water bodies and thereby the rainfall patterns nearby Lake Chad (Lauwaet et al. 2012). Note that most studies aimed at the evaluation of clouds in CPS are limited to a few days and to Alpine areas.

This study primarily aims at identifying the benefits and trade-offs of a CPS versus a non-CPS model effort using

an 11-year simulation over Belgium. This allows for the investigation of the added value of a CPS simulation in precipitation, temperature and clouds in areas with weak orographic forcing and for different seasons. A unique feature of this study is the comprehensive approach to evaluate the interconnection of cloud properties, TOA and surface radiation and their role in the creation of a surface temperature bias.

2 Data and methods

2.1 The COSMO-CLM model

All simulations of this study were performed using the Consortium for Small-scale Modelling in climate mode (COSMO-CLM) model. The COSMO-CLM model is a non-hydrostatic limited area climate model. This model is based on the COSMO model (Steppeler et al. 2003), designed by the Deutsche Wetterdienst (DWD) for operational weather prediction. In order to perform climate integrations with the COSMO model, the climate limited-area modelling (CLM) community provided extensions such as dynamic surface boundaries, a more complex soil model and the possibility to use various CO_2 concentration values (Böhm et al. 2006; Rockel et al. 2008). In this study, we use the 3rd order Runga-Kutta split-explicit time stepping scheme (Wicker and Skamarock 2002), the lower boundary fluxes provided by the TERRA model (Doms et al. 2011) and the radiative scheme after Ritter and Geleyn (1992).

Other settings are based on a previous study by Brisson et al. (2015) who provide recommendations for performing climate simulations at CPS. The one-moment microphysical parametrisation includes a representation of graupel hydrometeors in the finest resolution nest. In addition, the domain size of this simulation is large enough to ensure that the spatial spin-up described in Brisson et al. (2015) remains outside of the evaluation domain. The three-step nesting strategy shown in Fig. 1, is used in this study. The ERA-Interim reanalysis (grid mesh of ∼ 0.75° and 60 vertical levels) is used as initial and boundary conditions to nest a 100 × 100 grid points domain with a 0.22° (∼ 25 km) grid mesh size and 32 vertical levels. The resulting three hourly outputs are employed to nest a 0.0625° (∼ 7 km) domain. Finally, the hourly outputs of the latter nest, characterised by 150 × 150 grid points and 40 vertical levels, are used as input for the 0.025°(∼ 2.8 km) simulation on a 192 × 175 grid points domain and 40 vertical levels. The different simulations with the resolutions 25, 7 and 2.8 km are, hereafter, respectively referred to as C25, C7 and C3. The C25 and C7 do not explicitly resolve convection within the grid-scale and hence use the convection scheme after Tiedtke (1989) while the C3 dynamically resolves deep convection.
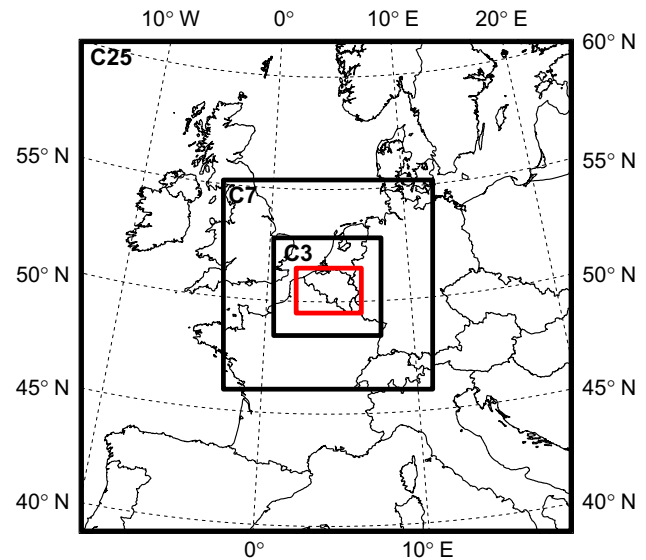

Fig. 1 Map of the three COSMO-CLM simulation nests (*black*) and the evaluation domain (*red*)

2.2 Evaluation period and domain

A 12-year period (i.e., 1999–2010) is simulated with this configuration. The first year of this simulation is used as spin-up period for the TERRA model (i.e. soil model component of COSMO-CLM). The resulting evaluation period is therefore 11-year long. This 11-year period is used for all analyses unless stated otherwise. The ERA-Interim reanalysis is used to drive the boundaries of the C25, ensuring that the large-scale forcing in the simulations stays close to the observed forcing. This allows for a detailed comparison of all simulations against observations.

Figure 1 depicts the simulation domains (in black) and the evaluation domain (in red). To investigate the added value of CPS by comparing CPS and non-CPS simulations only one single evaluation domain is used. This domain is located within the 0.025° simulation domain and does not encompass any lateral boundaries or spatial spin-up areas [as referred to in Brisson et al. (2015)]. This prevents evaluation biases due to the relaxation zone processes or deficiencies in the representation of convective events.

2.3 Observational datasets

2.3.1 Surface datasets

Figure 2 shows the locations of the numerous observations. (i) Daily (black points in Fig. 2) and hourly (blue points in Fig. 2) precipitation observations are available. Daily values are derived from the Royal Meteorological Institute (RMI) of Belgium and from the Global Historical
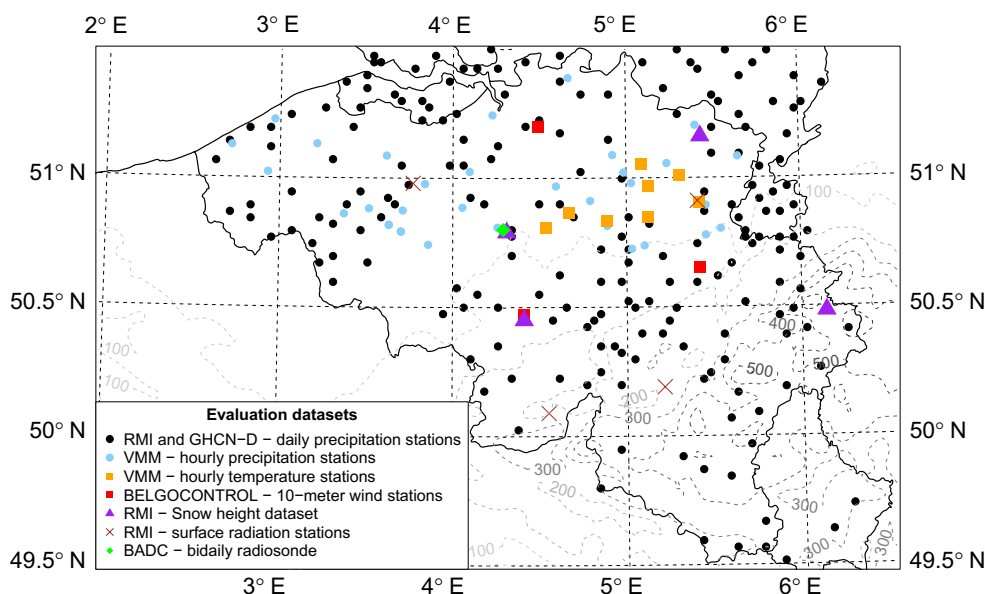
Fig. 2 Map of the evaluation domain (*red*) together with the locations of the observational datasets. The *black points* indicate the RMI and GHCN-D stations (i.e., daily observations) while the *orange squares* and *blue points* respectively indicate the locations of the precipitation stations extracted from the VMM dataset (i.e., hourly observations). 10-m wind speed stations are shown with *red squares*, snow height measurements with *purple triangles*, surface radiation with *brown crosses* and the radiosonde launching location with a *green diamond*. In addition, *dashed contour lines* show the orographical features of the evaluation domain

Climatology Network-Daily (GHCN-D) dataset (Menne et al. 2012) with a total of 199 stations covering the full simulation period (2000–2010). Hourly values are derived from the Vlaamse Milieumaatschappij (VMM) dataset. In total 37 stations are available with an averaged time-coverage of about 58 % of the simulation period. (ii) Wind speeds and directions are obtained as 10-min mean values from the Meteorological Services of Belgocontrol. The stations (red squares) are located in Antwerp, Liège and Charleroi at respective surface altitudes of 12, 187 and 265 m. (iii) Snow measurements were derived from the 3-hourly station network of RMI. Four meteorological stations namely Elsenborn, Gosselies, Uccle and Kleine-Brogel (shown as purple triangle) with respective altitudes of 570, 187, 101 and 64 m, were selected. The modelled snow depth data was extracted from the COSMO-CLM grid cells, which encompasses the coordinates of the stations location. (iv) Finally, radiation hourly measurements were derived from the RMI network. Four stations, namely Diepenbeek, Dourbs, Humain and Melle (shown as brown crosses) with respective altitudes of 39, 233, 296 and 15 m, were selected. The modelled surface radiative fluxes were extracted from the COSMO-CLM grid cells, which encompasses the coordinates of the stations locations.

In addition to these stations, the 0.25° resolution E-OBS [v10.0—(Haylock et al. 2008)] gridded dataset is used in this evaluation. Both the spatial and daily variability of E-OBS precipitation is underestimated compared to the

RMI and GHCN-D datasets. In addition the lower precipitation intensities are generally overestimated in E-OBS compared to the two other datasets while the highest intensities are underestimated (Fig. S1). These differences arises for two different reasons. First, only few stations (16) are used in E-OBS compared to the RMI and GHCN-D datasets. Second, E-OBS is a gridded product; each grid point describes values (e.g., precipitation accumulations) over a much larger area than for stations values. Therefore, the use of precipitation extracted from the E-OBS dataset is (in this study) restricted to the evaluation of spatial patterns or statistics on scales equal or greater than a month. The VMM dataset is found to match the RMI dataset fairly well for the higher daily intensity quantiles. However, for the lowest quantiles, the VMM dataset shows lower precipitation intensities than the RMI dataset (Fig. S1). This is in line with Willems et al. (2014). In addition, although only neighbouring stations are compared (i.e., distance smaller than ∼ 3 km), the distance between stations from different datasets may be large enough to result in random discrepancies for local events.

2.3.2 Radiosonde datasets

Observed radio-sounding profiles for Uccle [Fig. 2, see also Van Malderen and De Backer (2010)] have been retrieved from the British Atmospheric Data Centre (BADC). For the summer months (June, July and August) during the

period 2000–2010, approximately 360 soundings of (dew point) temperature, pressure and wind speed are available at both 00 and 12 h, while the number of valid observations varies between 237 and 360 depending on the variable and level of interest. Specific humidity is derived from the observed (dew-point) temperature and pressure values at each level. Due to the low number of observations (<50) in the upper atmosphere and the poor quality of radio soundings at these levels (Dee et al. 2011; Moradi et al. 2013), humidity observations above 300 hPa are excluded. For comparison, also the vertical profiles for temperature, specific humidity and wind speed have been retrieved from the grid cell covering Uccle in the European Center for Medium-Range Weather Forecast (ECMWF) Era Interim $0.75 \times 0.75°$ re-analysis product (Dee et al. 2011). This data is available on 37 fixed pressure levels between 1000 and 1 hPa and is selected for the time steps at which Uccle radio-soundings are available. For each available time step of the observed profiles, corresponding profiles are selected from C3, C7 and C25. The modelled profiles are spatially averaged per model level, representing an area of 25×25 km^2. That means that for the C3, 9×9 neighbouring grid cells are used to have one value per variable on one specific pressure level. Before the observed profiles are compared with modelled profiles for day and night separately, both COSMO-CLM and BADC radio-sounding variables are linearly interpolated onto the Era Interim pressure levels.

2.3.3 Satellite datasets

A. CMSAF cloud properties: Satellite retrievals of cloud properties were provided by the European Organisation for the Exploitation of Meteorological Satellites (EUMET-SAT) Satellite Application Facility on Climate Monitoring (CM SAF), i.e. hourly data of the CLAAS (CLoud property dAtAset using SEVIRI) dataset (Stengel et al. 2014) for the period 2004–2010. Cloud variables used in this study are cloud optical thickness (COT) and cloud top pressure (CTP), which are available on hourly resolution for pixels identified to contain clouds. The data is initially on native SEVIRI (Spinning Enhanced Visible and InfraRed Imager) instrument resolution with pixel sizes of about 4 km by 6 km over Central Europe. The CLAAS COT and CTP data was used to compose two-dimensional (2D) histograms as reported for the International Satellite Cloud Climatology Project (ISCCP) in Rossow and Schiffer (1999). Their cloud type classifications based on these 2D histograms is used in this study as well. Three COT intervals are used to discriminate between thin, intermediately thick and thick clouds with COT-thresholds of 3.6 and 23 respectively. Similarly, three CTP intervals are used to separate low, mid-level and high clouds, with CTP thresholds of

680 and 440 hPa respectively. From high, thin clouds to low, thick clouds, these classes include cirrus, cirrostratus, deep convection, altocumulus, altostratus, nimbostratus, cumulus, stratocumulus and stratus (Rossow and Schiffer 1999). It should be noted that low level clouds can be mis-classified during the construction of this fraction when the low clouds are overlaid by mid-level and/or high clouds. However, since the same methodology was applied on the simulated cloud fields, this comparison is considered to be fair. Since the visible channel is required for the COT retrievals, only daytime hours are included in the analysis. Uncertainties associated with CMSAF retrieved COT are well described, for example by Bugliaro et al. (2011). A correlation coefficient of 0.79 and a mean standard deviation of 0.92 were obtained between the CMSAF and their ground-truth COTs. Liquid-water clouds showed the best agreement, while a slight overestimation was present in ice clouds with $COT > 2$ and mixed-phase clouds. Stengel et al. (2014) report CLAAS COT comparisons against MODIS with a standard deviation of 6.2 and a bias of 1.8. Generally, optically thick clouds, (i.e., clouds with optical thicknesses above 100) are those with the highest uncertainties in COT retrievals due to the saturation in reflectances in the visible channels for these very bright clouds. Concerning the location of the cloud top, standard deviations of 2.5 km and biases of -1.0 km are found when comparing CLAAS cloud top height to the Cloud–Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument (Winker et al. 2009). When translating these values to cloud top pressure the standard deviation is 154 hPa and the bias 29 hPa for CLAAS CTP retrievals, and the correlations is approximately 0.85. In general, these uncertainties have a limited impact on the ISCCP classification of cloud types. The CTP of optically opaque, single-layer clouds can be determined with quite some confidence. For more details on the cloud property uncertainties please refer to Stengel et al. (2014) and Kniffka et al. (2013).

The Top-of-the-atmosphere (TOA) outgoing shortwave (OSW) and longwave (OLW) radiances are also used in the evaluation. These data were available at a slightly coarser resolution of about 9×18 km.

B. COSMO-CLM cloud properties: The comparison between satellite retrievals and simulated cloud fields is a trade-off between staying close to the original radiative transfer code of the model and still making sure that assumptions made in the satellite retrievals are taken into account, in order to make an apples-to-apples comparison. To make a fair comparison between the simulated and observed cloud optical properties, modelled COT and CTP were calculated offline, using the original Ritter and Geleyn (1992) scheme. The absorption and scattering optical depth in each model layer are calculated exactly as in the radiation scheme, using

three visible bands (Ritter and Geleyn 1992) and taking into account the optical effects of cloud liquid and ice, but ignoring any contribution from other hydrometeor species, such as snow. These calculations also account for the diagnosed grid-scale bulk cloud fraction, based on total relative humidity, as well as the shallow-cumulus cloud fraction. It should be mentioned that the original radiation code in the COSMO-CLM does not take into account the forward scattered peak for the calculation of the COT. Since these forward scattered photons eventually reach the surface, this is a good approximation for a numerical model. However, for the satellite observations, the forward scattered photons are lost from the beam and are hence not accounted for when retrieving the COT (Meirink, J.F., personal communication). Hence, in the off-line calculations for simulated COT, the forward scattered peak was taken into account, to be consistent with the satellite observations.

To derive a vertically integrated cloud optical depth, we follow Schroeder et al. (2006):

$$\tau = -\mu \ln \left(\prod_j \left(1 - b_j + b_j \exp \left(-\frac{\tau_j}{\mu_j} \right) \right) \right) \tag{1}$$

where j is the layer index, μ is the cosine of the solar zenith angle, τ_j is the optical depth and b_j is the bulk cloud fraction of layer j, normalised by the column cloud fraction. The latter is determined using a maximum-random overlap, following Oreopoulos and Khairoutdinov (2003). The simulated CTP has been estimated following Pincus et al. (2012) as the mean extinction-weighted pressure of the first visible optical depth [see their Eq. (2)], starting from the top of the atmosphere. For consistency with this value, only visible COT values >1 have been considered as cloudy in the all the remaining analysis. Since the satellite signal saturates for large values of COT, all satellite and simulated COT-values higher than 50 where thresholded to 50. The values of these lower and upper thresholds are somewhat arbitrary, but the same value is applied to the satellite and model fields. A lower value would detect more cloud, but retrieved COT at lower values becomes more uncertain because of increased complications presented by the variable land-surface albedo and the increased chance of interpreting multi-layer clouds. This approach has been followed before by e.g., Van Weverberg et al. (2012).

C. Regridding and averaging of cloud properties: Although the SEVIRI instrument and the COSMO-CLM model have fairly similar horizontal resolutions, much of the analysis involves regridding of data. Since the TOA radiances are provided on a coarser grid, all other properties are regridded to the TOA-grid spacing of 9×18 km. While this is fairly straightforward for properties like the CTP, it is

nontrivial for highly nonlinear properties like COT. We follow Schroeder et al. (2006) to perform the spatial aggregation of COT in the model and the observations to the TOA-grid:

$$\tau = -\mu \ln \left(\frac{1}{N} \sum_i \exp \left(\frac{-\tau_i}{\mu_i} \right) \right) \tag{2}$$

where μ is the cosine of the solar zenith angle, τ is the column COT (derived from Eq. 1) for the grid cell and N is the number of grid cells being aggregated over.

Using collocated COT and CTP values (regridded to the CMSAF-grid), each grid box could be assigned to one of the nine ISCCP cloud types, identical to the definitions applied to the satellite retrievals. This allows for a fair comparison between the simulated and observed cloud types.

3 Results

The large-scale dynamics determines to a large extent meteorological variables like near-surface characteristics (temperature and precipitation) and cloud properties. When a model is not able to represent large-scale atmospheric conditions, it will fail to adequately represent these meteorological variables. Therefore, an evaluation of the large-scale dynamics in the COSMO-CLM model (Sect. 3.1) is performed, notably to ensure that the three-step strategy does not deteriorate the solution of the C3. Then, in-depth comprehensive evaluations of precipitation (Sect. 3.2), temperature (Sect. 3.3), radiation and cloud properties (Sect. 3.4) are perfromed.

3.1 Large-scale forcing

The large-scale forcing is evaluated using wind speed, temperature and specific humidity which are the main variables used to force the COSMO-CLM at its boundary. The left panel in Fig. 3 reveals small relative differences between the summertime modelled and reanalysis large-scale wind speeds, being less than 10 % for atmospheric levels above 850 hPa. The comparison with the BADC radiosonde measurements shows similar results. Both the gridded and radiosonde wind speeds close to the surface tend to be inaccurate due to respectively its limited spatial scale (Decker et al. 2012) and inaccuracies caused by swinging sensors in an unstable surface layer (Genthon et al. 2010). Consequently, the C3 near-surface wind speeds is compared to three 10 m ground-based wind speed observations (see Fig. 2) by using the dimensionless Perkins' skill score (Perkins et al. 2007; Devis et al. 2013, 2014). The Perkins' skill score represents the common area between the probability density functions of the observed and modelled values.
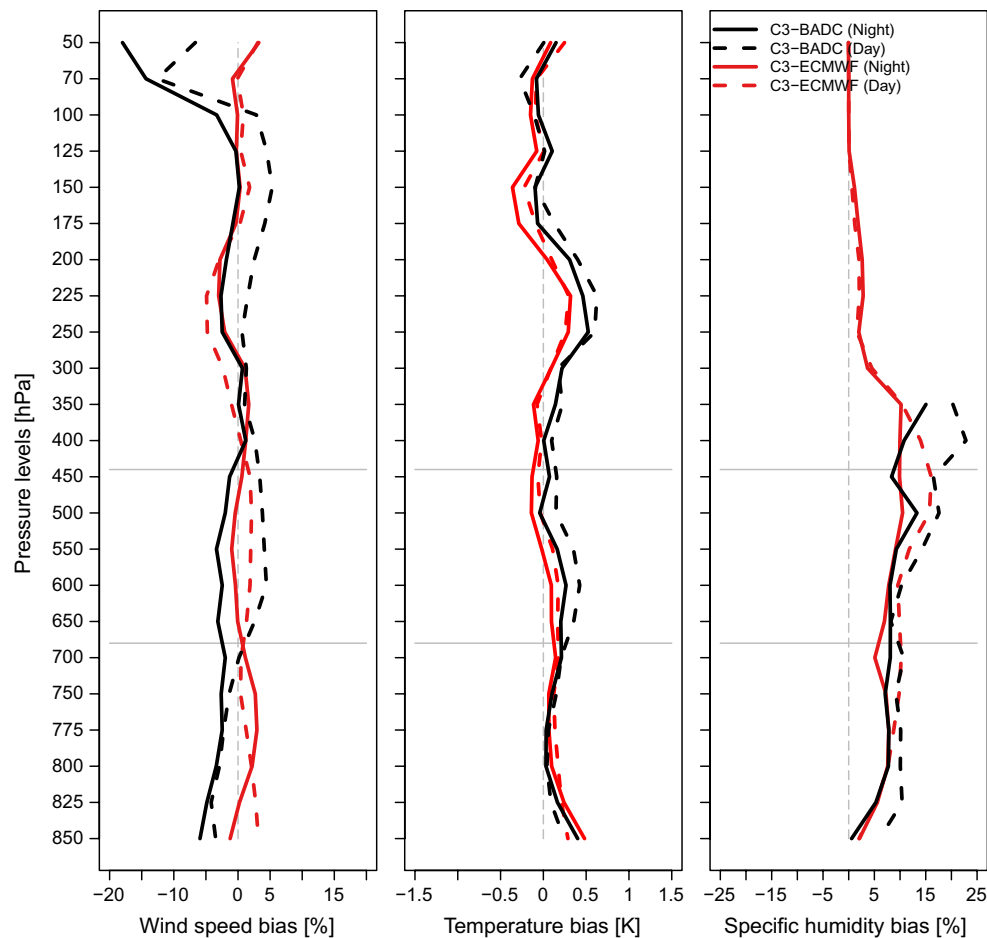
Fig. 3 Averaged summer bias for wind speed (*left*), temperature (*middle*) and specific humidity (*right*) for radio-soundings available during the period 2000–2010 for Uccle (WMO ID: 06447). *Black* and *red colors* refer to C3 minus BADC and C3 minus ECMWF respectively while night and day are indicated by respectively full and *dashed lines*. The perfect model *line* is indicated by the dashed *light-grey line*, while the ISCCP clouds levels are indicated by the solid *light-grey lines*. As mentioned in Sect. 2.3.2, levels below 850 hPa and above 300 hPa (for humidity only) are excluded from the analysis

Modelled and observed winds speeds for the location of Antwerp station strongly agree with a skill score of 0.85 (Fig. S2). However, the variance is underestimated, especially at night. Results for Liège and Charleroi are of similar quality, with skill scores of 0.91 for both stations (histograms not shown). Figure S3 indicates that there is also a strong correspondence between the observed and the modelled wind speed directions during day and night. Again, similar results are obtained for the Liège and Charleroi stations (not shown).

In addition, the model performance with respect to vertical temperature and humidity profiles are shown in Fig. 3, middle and right panel respectively. Similar as for wind speed, the C3 simulation performs well for temperature above 850 hPa, with differences generally smaller than 1 Kelvin. Closer to the surface, air temperatures are overestimated both during the day and the night, with a largest bias of 1.5 K for the former. This will be further discussed in Sect. 3.3. The humidity

bias generally increases with height in comparison with both BADC observations and ECMWF data. However, it is difficult to make firm conclusions, as the bias in the upper troposphere is in line with the instruments uncertainties; the instruments used in this analysis consist of two radiosonde types, namely RS80 and RS92, which tend to underestimate humidity (Miloshevich et al. 2004, 2009; Vömel et al. 2007; Van Malderen and De Backer 2010). This underestimation increases with decreasing temperature (and therefore with increasing height) and can reach up to 30 % (Miloshevich et al. 2006) at night and 50 % during the day (Vömel et al. 2007). Such underestimation results in uncertainties that are in the range of the biases observed in Fig. 3. In addition, these radiosoundings are assimilated in the production process of the ERA-Interim reanalysis dataset. This explains why C3 shows a bias of similar amplitude when compared to ECMWF or BADC radiosoundings. Other model nests (C25 and C7) show similar behavior (not shown).

As discussed in Sect. 2.3.2, observed humidity values above 300 hPa are of poor quality and infrequently sampled and as such not shown here. At these levels, C3 exactly follows the ECMWF forcing which is a model result since no humidity increments are allowed at this level (Dee et al. 2011). Note also that both the temperature and humidity bias profiles for C3-BADC and C3-ECMWF are very similar. This is not surprising since the Uccle station (ID 06447) is used in ERA-Interim's data assimilation scheme.

Based on these results, it is most likely that the systematic deviations from the forcing in the CPS simulations are rather limited. As such, the differences in the representation of precipitation, temperature and clouds, observed between C25, C7, C3 and the observations are not caused by deficiencies in the large-scale dynamics.

3.2 Representation of precipitation

The precipitation daily cycle, as described by the average over all VMM stations, ranges from ~0.07 to ~0.09 mm/h (black line in Fig. 4). A peak, related to convective activity, occurs from ~3 pm till ~9 pm (local time). Although peaks are modelled in C25 and C7s, the maximum intensity of these peaks occurs around 1 PM. In the C3, the timing of the convective peak is modelled accurately. The representation of the precipitation diurnal range is also improved in the C3 (Fig. 4); while the diurnal precipitation range is underestimated by 25 and 52 % in C7 and C25 respectively, this turned into a slight overestimation (13 %) in the C3.

In addition to this improved description of the diurnal cycle of precipitation, C3 is showing superior skill to model the distribution of the hourly precipitation accumulations compared to C7 (Fig. 5a). For these two simulations
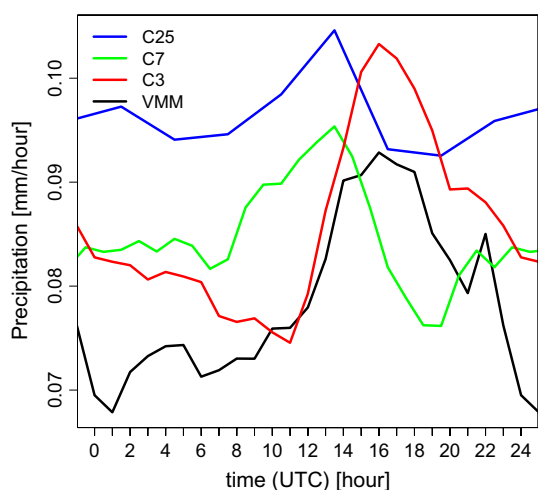

Fig. 4 Daily cycle of hourly precipitation for the period 2000–2010. C25, C7, C3 and the VMM datatset are respectively shown in *blue*, *green*, *red* and *black*

(i.e., C3 and C7), Perkins' skill scores (PSS) (Perkins et al. 2007) of respectively 0.95 and 0.93 are found for the hours with a precipitation accumulation above 0.1 mm. Aggregating C3 to 25 km does not result in a deterioration of the representation of the hourly distribution (PSS equal to 0.95). This indicates that the added value of the CPS simulation to represent the hourly distribution probably lies in the explicit representation of deep convection more than in the ability to produce a precipitation dataset at finer resolution. On longer timescales, the benefit of CPS over non-CPS simulations is reduced (Fig. 5b, c). This is notably reflected in the PSS of C3 and C7 which while it differs by 0.024 at the hourly scale, differs by only 0.003 at the daily timescale. This feature is also found over mountainous area Ban et al. (2014) (Table 1).

Improving the timing and the intensity of the most extreme precipitation events are not the only added values of the CPS simulation. Due to the local character of convective precipitation, the spatial structure of the precipitation pattern can be very detailed. High precipitation spatial variance is, therefore, expected during convective events. The spatial variance of daily precipitation at CPS outperforms the two non-CPS simulations (Fig. 6). Although the time-averaged variance from the C3 is still lower than the observed variance by up to 5 mm^2/day^2, it is largely improved compared to the C25 and the C7 simulations on all spatial scales (Fig. 6a). The highest variance quantiles, that characterise local and intense precipitation typical for convective events are also largely improved in the CPS simulation. Indeed the 95th percentile of the variance on a spatial scale of 50 km is underestimated by more than 55 % for C25 and C7 while this underestimation is reduced to 15 % for C3 (Fig. 6b).

A refinement of the grid also allows for more extreme precipitation in a single grid point for a similar amount of water over a given area. Because ground-based stations measure precipitation over surface areas no larger than square decimeters, highest precipitation quantiles are expected to be improved in the highest resolution simulations. To understand the contribution of the grid refinement compared to that of partially resolving convection, the CPS-model output is aggregated to the grid of the C25. Figure 6 does not show large differences of the variance for the aggregated C3 compared to the original C3, and the aggregated output is more realistic than those of both the C25 and the C7. This result leads to two conclusions. First, the C3 does not represent correctly small scale spatial variability of precipitation as shown in Fig. 6. Second, the improved representation of precipitation on CPS is not due to the different grid on which the analysis is performed.

Refining the model grid does not only allow for larger precipitation extremes over a grid-point. Decreasing
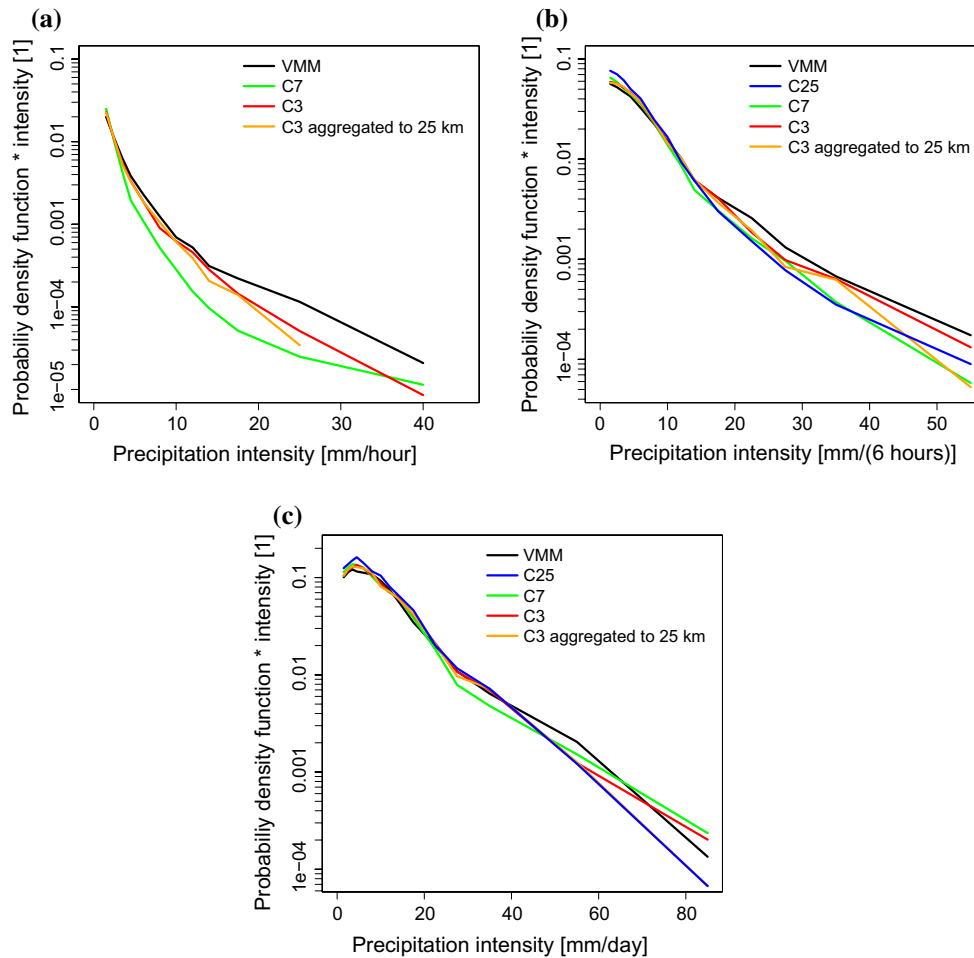
Fig. 5 Probability distribution of the hourly (**a**), 6-hourly (**b**) and 24-hourly (**c**) precipitation accumulations multiplied by precipitation intensity. C25, C7, C3, C3 aggregated to 25 km and the VMM datatset are respectively shown in *blue*, *green*, *red*, *orange* and *black*. C25 is not available on an hourly time-scale and is therefore not shown in **a**

Table 1 Daily precipitation intensity average, 95th and 99th quantiles for the observation and the different simulations for the full evaluation period and the summer period only

Dataset	Full period			Summer		
	Average	95th	99th	Average	95th	99th
RMI (obs)	**2.45**	**12.00**	**23.00**	**2.73**	**14.30**	**27.70**
C25	2.4	11.8	21.5	2.2	11.2	22.6
C7	2.2	11.0	20.9	1.9	10.3	23.0
C3	2.2	11.2	21.7	2.1	11.6	26.6
C3 aggregated to ∼25 km	2.2	11.0	20.6	2.1	11.1	24.4

the grid spacing also allows for the use of more accurate external parameters. The representation of precipitation in mountainous areas was, notably, found to improve at higher resolution due to the improved representation of orography (Prein et al. 2013). The processes inherent to the interaction of mountains or hills with air masses—such as the condensation of water during a forced ascent of an air mass and the triggering of convection—are two examples of processes that may benefit from an improved

description of orography. To illustrate this benefit, the temporally averaged daily precipitation over the period 2000–2010 is shown in Fig. S4. Although some small resolution dependencies appear in the flat areas (i.e., western part of the domain), the main differences occur in the hilly area in the South-East where precipitation amounts reach up to 3.4 mm/day. Indeed, the spatial extent of this area is largely overestimated in the coarser simulation (i.e., C25) compared to the C7 and C3. This difference in
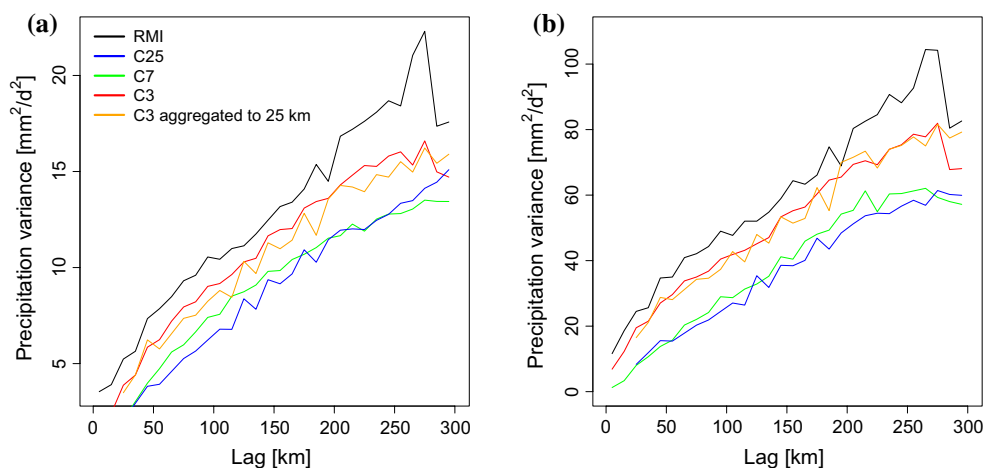
Fig. 6 Precipitation variogram for RMI observations (*black line*), C25 (*blue line*), C7 (*green line*), C3 (*red line*) and C3 aggregated to 25 km (*orange line*). Both the temporal mean (**a**) and the 95th quantile (**b**) are shown
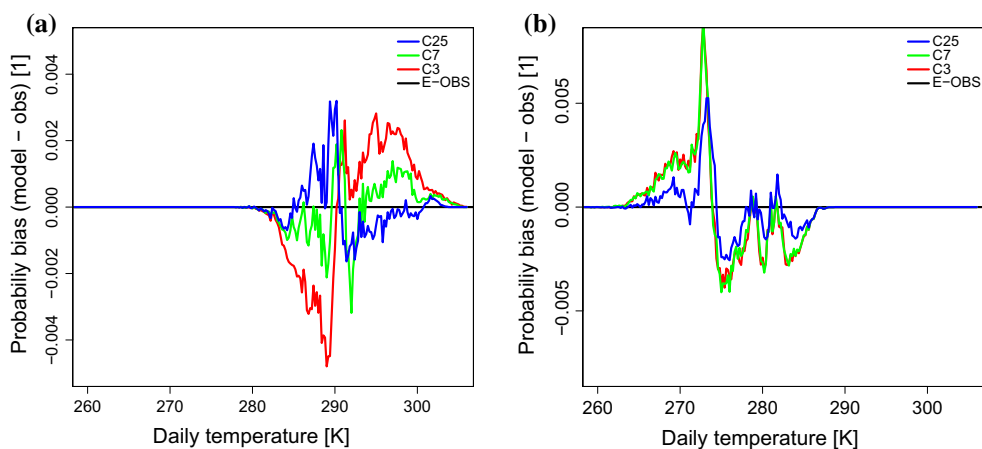

Fig. 7 Empirical probability density function of daily temperature for the summer months (**a**) and the winter months (**b**). The observed probability are subtracted to the modelled ones so that the best fit is indicated with the 0-line

precipitation depth between the C25 and the C3 is significantly correlated ($R^2 = 0.62$) to the difference in surface altitude between the two simulations. Such a positive correlation is likely related to the impact of orography description in the model on the representation of precipitation. In addition, a yearly cycle in this correlation is found with a maximum R^2 of 0.69 occurring in winter and a minimum R^2 of 0.12 in summer (not shown). Such yearly cycle was expected because the triggering of convective events due to orography is rather limited in Belgium (Goudenhoofdt and Delobbe 2013). Therefore an improvement of the orography description is mainly affecting stratiform precipitation. While summer precipitation is composed of both convective and stratiform precipitation events, winter precipitation mainly consists of stratiform precipitation events. This results in higher sensitivity of precipitation to the

description of orography in Belgium in winter compared to summer.

3.3 Representation of temperature

Temperature also benefits from the use of CPS. However, these benefits are mostly related to the refinement of the orography. The extend of the area with lower temperature (i.e., hilly area) is larger in C25 compared to C3 and the E-OBS dataset (Fig. S5). Similarly to the findings of Sect. 3.2, the coarse representation of orography in C25 is likely to be responsible for this deficiency.

Although temperature averages are fairly well reproduced by these models with a time-averaged bias equal or lower than ~0.5 K, the temperature range in C3 is significantly overestimated compared to the observed range (not

Table 2 Domain- and time-averaged TOA outgoing shortwave radiation (OSR), surface incoming shortwave radiation (SFC ISR), TOA outgoing longwave radiation (OLR) and cloud fraction (CF) for summer (JJA) 2004–2010

	CMSAF	C3
TOA OSR (W/m^2)	308	291
SFC ISR (W/m^2)	412	442
TOA OLR (W/m^2)	239	239
Total CF (%)	65	46

OSR, OLR and CF are calculated from hourly values in the satellite and the observations, for daytime hours only (zenith angle < 65°)

Observed and simulated SFC ISR are averaged over the locations of the surface radiation stations depicted in Fig. 1 and for consistency are for the same daytime hours as the satellite radiation data only. OSR and OLR include clear-sky and cloudy regions and cloud fractions are for regions with COT >1

shown). In summer, the probability of having cold days is underestimated in C3 while it is overestimated in the C25 (Fig. 7). On the contrary, the frequency of warmest summer days is overestimated in the finest simulation and underestimated in the coarsest simulation (Fig. 7a). It is hypothesised that these differences may arise from changes in the radiative forcing. This hypothesis is further explored in Sect. 3.4. It should be noted that C3 overestimation of warm days in summer is likely to result in an overall too warm and therefore too high planetary boundary layer (PBL). This has strong implications for the triggering of convective precipitation as the PBL may reach the level of free convection either too early or on the wrong location. However, based on the results from Sect. 3.2, the modelled precipitation diurnal cycle and precipitation highest quantiles are in fair agreement with observations. This fair representation of precipitation could be due to a unknown compensating bias or to the small impact of the process described above (i.e., impact on convection triggering of the growth of the PBL due to an overestimation of temperature). Further analyses are therefore necessary, but due to the lack of relevant observation and model simulations, these analyses are outside the scope of this study.

Another bias, observed for all resolutions, although larger in C3, is the overestimation of the occurrence of temperatures lower than ∼275 K in winter with a larger bias around 273K (Fig. 7b). This overestimation is also reflected in an overestimation of snow cover which results in too high short-wave radiation reflection. Indeed, the histograms of the observed and simulated snow depth reveal an overestimation of the number of snow episodes in model simulations compared to observations (Fig. S6). This overestimation is more evident for events with snow-depth lower than 5 cm. For larger snow-depths the model simulations does not show large differences compared to observations,

although this could be related to low occurrence probability for such events at the observation locations.

3.4 Radiative and cloud properties

Clouds play a key role in the energy balance of climate models. In addition, convection has a large impact on the development of clouds and the representation of convection in climate models is therefore likely to influence the model performance in terms of clouds. As an example, a consistent misrepresentation of the diurnal cycle was found by Langhans et al. (2013) and Pfeifroth et al. (2012) between precipitation and clouds in non-CPS models. Hence, it is useful to explore the added value of CPS to the representation of clouds and their impact on the radiation balance. This evaluation is limited to C3 due the unavailability of the corresponding outputs for the parent nests.

The evaluation of clouds and radiation is performed using MSG-satellite and surface station data, as presented in Sect. 2.3.3. Geostationary satellites clearly have an advantage over ground-based point measurements in terms of their spatial and temporal coverage and hence can provide a more detailed evaluation of cloud properties than surface stations. Nevertheless, one should keep in mind that satellite measurements are associated with a number of limitations and uncertainties as listed in Sect. 2.3.3.

From Table 2 the domain- and time-averaged Top-Of-the-Atmosphere (TOA) Outgoing Shortwave Radiation (OSR) is slightly underestimated by the C3 simulation, compared to the CMSAF retrieved radiation. The outgoing longwave radiation (OLR) is much better captured. The magnitude of these biases compares well with earlier studies (e.g., Kothe et al. 2011). It should be noted that it would be desirable for any climate model to have the seasonal averaged TOA radiation well captured, in order to avoid a drift of the model to a different climatology. Note however that the numbers in Table 2 are for daytime only (not including the nighttime and low sun angles). As such, these numbers can not provide a complete picture of the TOA radiation budget in COSMO-CLM.

An underestimation of the daytime OSR can be caused by a lack of cloudiness in the simulations, or by wrong optical properties (i.e. too large transmission) of the simulated clouds, or by a combination of both.

Apart from an overall negative bias in OSR, the distribution of the OSR values is broader in the C3 experiment than in the CMSAF retrievals (Fig. 8). Indeed, low (<250 W/m^2) and high values (>600 W/m^2) of OSR are more prevalent in C3 than in the satellite observations. The distribution of OLR seems fairly well represented. Low values of OSR are typically related to cloud-free areas, while high values of OSR can be related to very reflective clouds. Hence, the too broad and too skewed distribution of the OSR in
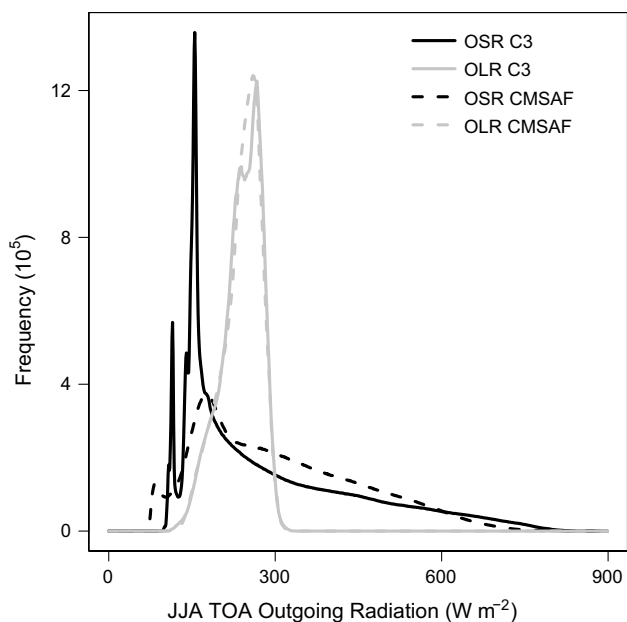
Fig. 8 Histograms of TOA OSR and TOA OLR for C3 and as obtained by the CMSAF during Summer (JJA) 2004–2010 (bins of 1 W/m^2). Provided is the total frequency for the full analysis domain and for all output times available

C3 suggests that the slight underestimation of domain- and time averaged OSR (Table 2) masks a large overestimation of clear sky conditions, partly offset by too frequent reflective clouds when they are present.

Based on the satellite-retrieved and simulated cloud optical thickness (COT; Sect. 2.3.3), it is possible to obtain an ad-hoc, but consistent measure of the total cloud fraction in the observations and the simulations. The cloud fraction in Table 2 is based on the occurrence of clouds thicker than COT >1 and shows an important underestimation of the total cloud cover in the C3 experiment. To further explore the nature of this underestimation, Fig. 9 shows 2D-histograms of the frequency of clouds, binned by COT and cloud top pressure (CTP), using the ISCCP framework (Sect. 2.3.3). COT in the model is calculated by following as closely as possible the formulations from the radiation scheme of Ritter and Geleyn (1992) and data are regridded to the coarser TOA-grid using Eq. 2. Cloudy grid cells are defined as grid cells with COT > 1. An overview of the domain- and time-averaged cloud cover for the 9 distinct cloud types, defined as in the ISCCP framework, is provided in Table 3.

From Fig. 9 and Table 3, it is mainly the high and intermediately thick cloud cover (cirrus (Ci), cirrostratus (Cs) and altostratus (As)) that is under-represented in the C3 simulation. Clouds with CTP < 300 hPa only occur about 25 % as frequently in the C3 simulation than in the CMSAF. Conversely, thin, low clouds (Cumulus—Cu), as well as very thick, low clouds (COT > 40; Stratus—St) have almost twice the observed cover in the C3. Such biases are well beyond measurement's uncertainties. It should be pointed out that the very important underestimation of total cloud cover, even for fairly thick clouds, results in only a modest underestimation of TOA OSR. This indicates that
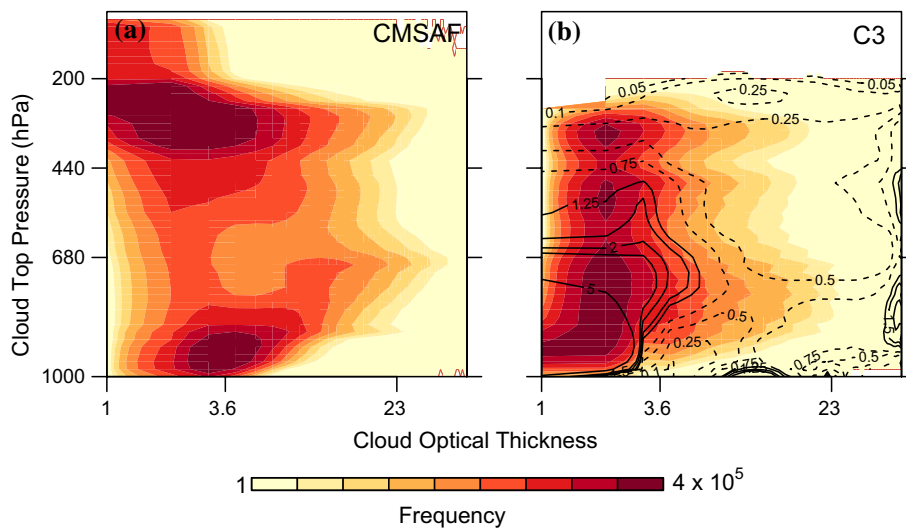

Fig. 9 2D-histograms of the absolute frequency of occurrence of clouds as retrieved by the CMSAF (**a**) and as simulated in C3 (**b**) for Summer (JJA) 2004–2010. Clouds are binned by Cloud Optical Thickness (abscissa—bins of 1) and Cloud Top Pressure (ordinate—bins of 20 hPa), as in the ISCCP framework. Note that the abscissa has a logarithmic scale and that the colour scale is non-linear (each level is 1.5 times smaller than the level above). The contours on panel (**b**) denote the relative bias between the absolute frequencies in C3 and the CMSAF. *Solid (dashed) lines* denote a positive (negative) bias in the simulations. Only daytime hours (zenith angle < 65°) are included

Table 3 Frequencies of the ISCCP cloud types (in %) as obtained by the CMSAF and as simulated by C3 for summer (JJA) 2004–2010. See text for the full description of the cloud acronyms

%	CMSAF	C3
Ci CF	11	5
Cs CF	14	5
Cb CF	3	2
Ac CF	3	7
As CF	11	6
Nb CF	4	3
Cu CF	5	8
Sc CF	12	9
St CF	2	3

Only daytime hours (zenith angle < 65°) and regions with COT > 1 are included

the underestimated cloud cover should be partly compensated by too reflective clouds, as suggested by the right-end tail in the histograms in Fig. 8.

Indeed, Fig. 10 shows the biases in the TOA radiative fluxes for clouds, binned in the COT-CTP parameter space, as in Fig. 9. While the model captures the increase of TOA OSR with increasing COT, the OSR is largely overestimated for all cloudy grid cells, pointing to too reflective clouds. For fairly low and intermediately thick clouds, the overestimation amounts to more than 25 %. For high-level clouds, the overestimation is more modest. The C3 experiment also captures the general decrease of TOA OLR with decreasing CTP, but is biased low (Fig. 10c, d) for all clouds. The bias becomes larger for high clouds, but the magnitude of the bias remains smaller than for OSR. It is clear that the bias in OSR partly offsets the underestimated cloud cover to produce seasonally-averaged TOA OSR that is only slightly negatively biased compared to the observations. Similarly, the low bias in OLR for cloudy grid cells conspires with the too frequent clear skies (that are associated with large values of OLR) to produce a seasonal average and a distribution close to the observations (Table 2; Fig. 8).

Fig. 10 2D-histograms of the TOA outgoing radiative fluxes as retrieved by the CMSAF (*left*) and as simulated in C3 (*right*) for Summer (JJA) 2004–2010. All fluxes are binned by Cloud Optical Thickness (abscissa—bins of 1) and Cloud Top Pressure (ordinate - bins of 20 hPa), as in the ISCCP framework. Provided are the outgoing shortwave radiation (*top*) and outgoing longwave radiation (*bottom*). Note that the *colour scale* denotes different values in the *top* and *bottom panels* and that the abscissa has a logarithmic scale. The contours on the *right-hand panels* denote the relative bias between the radiation in C3 and the CMSAF. *Solid* (*dashed*) *lines* denote a positive (negative) bias in the simulations. Only daytime hours (zenith angle < 65°) are included
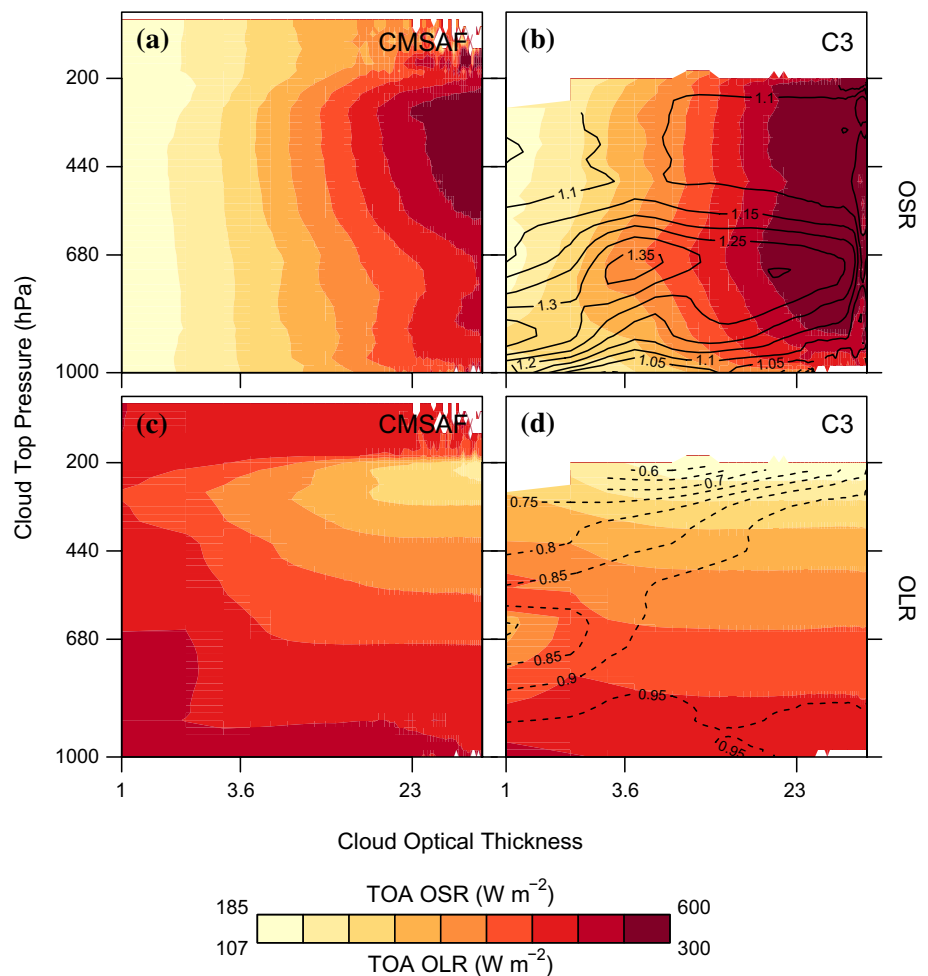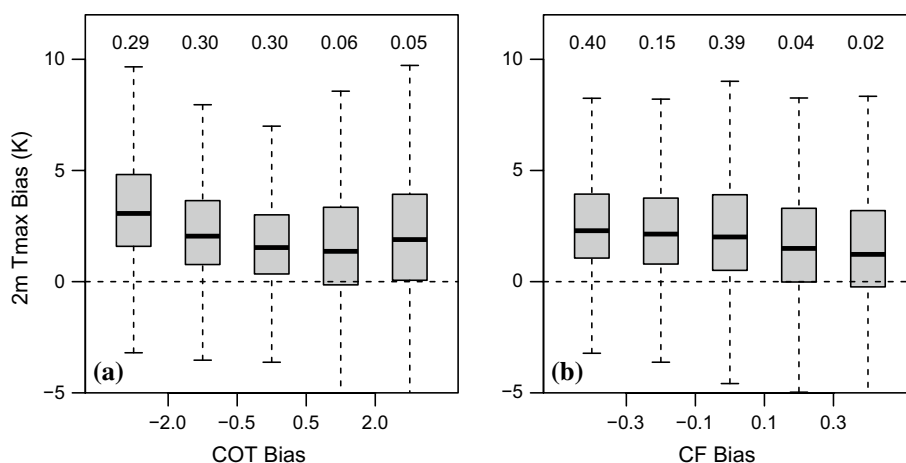
Fig. 11 *Box plots* of the domain- and time-averaged Summer (JJA) 2004–2010 mean 2 m-maximum temperature bias composited by its coincident bias in cloud optical thickness (COT—*left*) and cloud fraction (CF—*right*). Temperature bias *boxes* and *whiskers* are provided for respectively 5 coincident COT and CF intervals, indicated by the numbers on the X-axes. The *box* and *whisker* lengths denote the 25th and 75th percentiles and the 5th and 95th percentiles respectively. The fraction of data points in each *box-whisker plot* is indicated by *the number above each box*. The 2 m maximum temperature bias is the difference between C3 (regridded to EOBS) and EOBS. The COT bias is the difference between C3 (regridded to EOBS) and CMSAF (regridded to EOBS) and the cloud fraction bias is the difference between the fraction of the EOBS grid that has COT >1 according to C3 (regridded to CMSAF) and according to CMSAF. Cloud fractions and COT are averaged over the time period 9–12 UTC and all averaging is done using the definition of extinction coefficient (see text)

It should be mentioned that most previous studies using the COSMO-CLM at scales that still required a convective parametrisation (e.g., Kothe et al. 2011; Pfeifroth et al. 2012) provide evidence for too large cloud cover and too large OSR (Kothe et al. (2011)). Studies that employed the COSMO-CLM in a convection-permitting setting often found lower cloud cover than in coarse-scale simulations, but did not show the large negative bias as shown in our study (Prein et al. 2013; Langhans et al. 2013). It should be stressed that previous studies were usually focused on the Alpine region and for shorter time-scales or case studies specifically tailored at convective events only (e.g., Langhans et al. 2013).

It is very likely that the negative bias in TOA OSR, due entirely to a lack of high-topped and/or intermediately thick clouds, is the cause of the warm temperature bias during summer in the C3 experiment. Since absorption of shortwave radiation by clouds is negligible, the low bias in OSR leads to too much shortwave radiation reaching the surface, as shown in Table 2, averaged for the surface radiation stations available. To further support this hypothesis, Fig. 11 shows box-whisker plots of the domain-averaged 2 m maximum temperature (i.e., maximum instantaneous temperature) bias for 5 intervals of coincident COT bias (left) and CF bias (right). COT was re-gridded to the EOBS-grid following Schroeder et al. (2006), using the relation between transmission and optical depth (see Figure caption for more details on the regridding). The COT and cloud fractions that are confronted with the 2 m maximum temperature bias in this figure are averaged over a time window from 9 to 12UTC. This assumes that cloud properties during this time period have the most significant impact on maximum temperature. We used the EOBS temperature data in this comparison given that these are gridded data and hence are more straightforward to compare against the gridded COT and CF data. From Fig. 11a, the better the COT is captured, the smaller the maximum temperature bias. While there remains a slight warm bias even when the COT is well captured (−0.5 <COT bias< 0.5) or even overestimated (COT bias > 0.5), this is a clear indication that clouds are a major contributor to the maximum temperature bias in the model. A COT bias < −2 leads to a median temperature bias of about 3 K, whereas days with a well-simulated COT (−0.5 < COT bias < 0.5) have a median bias of only 1 K. Figure 11b shows the correlation between the 2 m maximum temperature bias and the domain- and time averaged daytime cloud fraction bias. The relation between the cloud fraction bias and the temperature bias is far less clear, probably due to the fact that biases in cloud fractions are often associated with high, thin clouds (Fig. 9), that are transparent to shortwave radiation. COT is a more direct link between clouds and radiation and hence a more appropriate measure to estimate the link between clouds and the surface temperature.

While the above analysis encompasses an evaluation of daytime mean cloud and radiation fields, the diurnal variability of these fields could produce biases in the maximum 2 m-temperatures as well (even if the daily mean values were captured well). Pfeifroth et al. (2012) and Langhans et al. (2013) found that apart from the diurnal cycle in
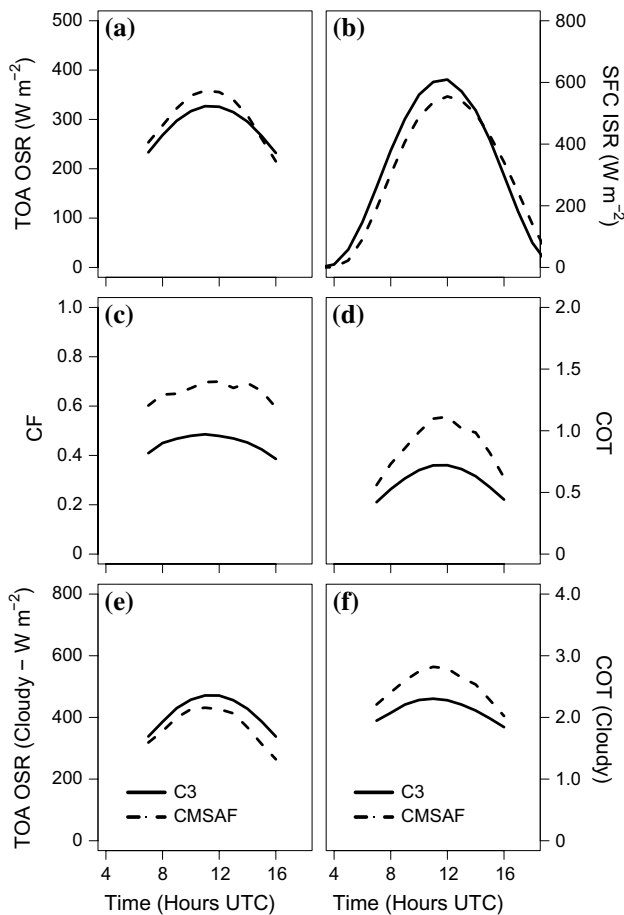
Fig. 12 Overview of domain- and time averaged diurnal cycles as observed and as in the C3 experiment for Summer (JJA) 2004–2010. Provided are the total TOA outgoing shortwave radiation (**a**), station-averaged incoming shortwave radiation (**b**), total cloud fraction (**c**), total cloud optical thickness (**d**), TOA outgoing shortwave radiation for cloudy regions only (**e**) and cloud optical thickness for cloudy regions only (**f**). The observations in all panels are obtained from the CMSAF, except for panel (**b**), which are the surface radiation stations denoted in Fig. 2. Only daytime hours (zenith angle < 65°) are included for the panels involving CMSAF-data

precipitation, coarser-scale simulations have more difficulties capturing the diurnal cycle of cloudiness than CPS-simulations. In this section, the diurnal cycles of cloud properties in the 11-year convection-permitting climate simulation are explored. Figure 12a compares the JJA mean diurnal cycle in OSR in the model and the observations. From this figure, OSR seems to be mainly underestimated early in the day and during noon, and becomes better captured or even overestimated in the late afternoon. The larger underestimated TOA OSR in the morning hours in the C3 experiment is also reflected in the positive bias in shortwave radiation (Fig. 12b) reaching the surface during the morning. This means that the surface mainly receives too much shortwave radiation before and around noon, leading to a too fast increase in surface temperatures during

this time and exacerbated maximum temperatures shortly thereafter. By the late afternoon (after 14 UTC), TOA OSR in C3 is overestimated and the shortwave radiation reaching the surface underestimated. From Fig. 12c, the different behaviour before and after noon is unrelated to a shift in the total cloud fraction bias. Indeed, in contrast to many coarse-scale simulations (e.g., Pfeifroth et al. 2012; Jaeger et al. 2008), the diurnal cycle of cloud fraction is very well captured in this convection-permitting simulation. Despite the overall lack of cloudiness, this is a clear benefit of a CPS configuration.

Figure 12d shows the diurnal cycle of domain-total COT and Fig. 12e and f show the diurnal cycles of TOA OSR and COT for cloudy grid cells only. During the afternoon the positive bias in OSR for cloudy grid cells grows larger and the COT overestimation decreases. Hence, it seems that while the cloud fraction bias remains fairly constant during the afternoon, the biases in the optical properties of the simulated clouds become quite different. Clouds in the late afternoon are much more reflective in the C3 than in the observations while their COT becomes more similar. A more detailed picture of the diurnal cycle of cloud properties is painted in Fig. 13, showing the diurnal cycles of each individual ISCCP cloud type. While the diurnal cycles of many individual cloud types are fairly well represented in the C3, the decline of thick and reflective clouds types in the afternoon [e.g., stratocumulus (sc) and nimbostratus (Ns)] is a lot slower in the C3 than in the observations. Conversely, the increase of non-reflective and optically thin clouds during the afternoon [e.g., cirrus (ci) and altocumulus (ac)] is a lot slower in the C3 than in the observations as well. While the cloud fraction bias remains fairly constant throughout the day, clouds in the afternoon become more reflective and hence produce a smaller bias in the total TOA OSR and the radiation reaching the surface. It should be stressed that this is for the wrong reasons however, since the cloudy TOA OSR becomes even more overestimated, while the cloud fraction bias is not improved.

4 Discussion and conclusion

The main goal of this study is to identify the benefits of convection-permitting scale (CPS) simulations over Belgium using an 11-year simulation. Unlike many previous studies, the region of interest in this study has very weak orographic forcing. The goals of this paper are twofold. Firstly, this study examines CPS-benefits associated with the explicit representation of the convection and the larger detail of the orography and land surface. Secondly, this study aims at understanding remaining biases at CPS. To do so, biases in temperature are linked to deficiencies in the representation of radiation and clouds.

Fig. 13 Diurnal cycles of the cloud fraction in each of the ISCCP cloud categories as obtained from he CMSAF and as simulated in C3. The figure shows high-level clouds towards the top of the figure and thick clouds towards the *right hand side* of the figure. See text for the full description of the cloud acronyms. Only daytime hours (zenith angle < 65°) are included
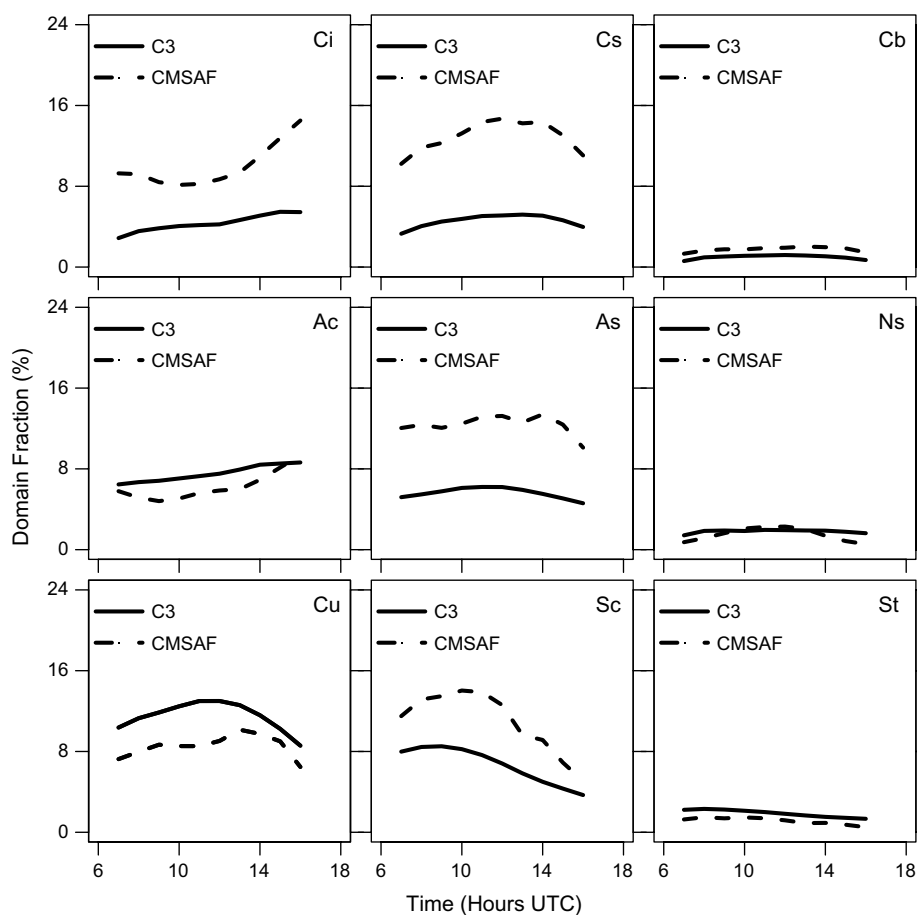
A first step in this analysis is to ensure that the large-scale forcing is well represented. The representation of the large-scale forcing in the model was evaluated by means of the 10-m wind speed and radio-soundings. Generally, the large-scale dynamics are well represented by the CPS simulations in terms of wind speed and direction. On vertical levels above 850 hPa, the temperature bias is generally smaller than 1 K, while the humidity bias is not larger than the instruments uncertainties. These results increase our confidence that biases in the precipitation, temperature and cloud fields are not likely to result from systematic deviations from the large-scale forcing.

Generally, the added value of CPS simulations found for complex orography (Ban et al. 2014; Fosser et al. 2015) is also found for moderate orography (i.e., Belgium). The timing of the mid-afternoon precipitation peak is better captured when convection is partly resolved. In addition, the hourly rain rates are also improved in the CPS compared to the non-CPS simulations, similarly to findings by Prein et al. (2013). However, on longer time-scales (i.e., 12 h, daily), this added value vanishes. Aggregating the CPS simulations to the coarsest non-CPS simulation grid does not significantly degrade the hourly and daily distribution. These findings indicate that the added value of CPS

simulations probably result from either the better description of the surface in the model or from the explicit treatment of deep convection.

For temperature, the benefits of using CPS simulations are much smaller than for precipitation. The spatial variability of temperature is also similar in the different simulations. Large improvements are only observed for the representation of the spatial distribution of temperature over hilly areas. Similarly to previous studies (e.g., Prein et al. 2013), we found that these improvements are likely to result from an improved description of the orography in the high-resolution models. This grid refinement also results in increased biases in the distribution of daily temperatures. Too frequent days with temperature values around 273 K are correlated with too frequent days with snow cover in the model compared to the observations. In addition, in summer, a warm bias is observed in the CPS simulation while the coarsest simulation is characterised by a cold bias.

A detailed evaluation of the simulated cloud fields at CPS against satellite information reveals a modest low bias in the the top-of-the-atmosphere (TOA) outgoing shortwave radiation, mainly during summer. This low bias originates from a significant over-representation of clear-sky conditions in the model, partly offset by too reflective

clouds when they are present. Previous studies also pointed out that cloud cover is generally reduced at CPS compared to non-CPS (Kothe et al. 2011; Pfeifroth et al. 2012; Langhans et al. 2013; Prein et al. 2013; Ban et al. 2014). Hence, too much shortwave radiation reaches the surface, which was found to correlate with the occurrence of the warm-temperature bias in summer.

Previous studies show that non-CPS usually fail to capture the diurnal cycle of cloudiness (Pfeifroth et al. 2012; Jaeger et al. 2008). In contrast, for most cloud types in our CPS simulations, the diurnal cycle of cloud fraction is well captured despite an underestimation of total cloudiness. However, thick and reflective cloud types are too persistent in the afternoon compared to the observations. Conversely, the afternoon increase in thin and transparent clouds is not well-captured by the simulations.

The results of this study show the potential of CPS simulations to increase the confidence in RCM, not only for areas with steep orographic gradients but also for regions with moderate orographic forcing. Although the added value of CPS for the representation of precipitation, temperature and clouds are primarily found in the short spatio-temporal scales, these are crucial to many impact studies such as hydrology or soil erosion studies. However, this study also point to a deterioration in temperature partly explained by deficiencies of cloud processes. Further investigations of the parametrisation of cloud processes remains, therefore, necessary. In addition, it was also shown that even if the representation of some fields are fairly well represented, it may results from error compensation. This was notably true for the seasonally-averaged representation of well-captured TOA outgoing radiances, which result from compensating errors of too low cloud fractions and too reflective clouds. It should be further investigated what the specific reasons are for the lack of cloudiness and the too reflective clouds in the simulations and for the change in behaviour of the cloud properties in the afternoon. One obvious shortcoming of the current radiation scheme is that it assumes the snow and graupel species to be transparent to radiation. It is likely that much larger cloud fractions would occur if the snow species would be taken into account. Furthermore, assumptions about the cloud overlap between the vertical model levels are very arbitrary and have been shown before to have a significant impact on the radiative transfer (e.g., Neggers and Siebesma 2013). However, this analysis also shows that any attempt to increase the overall cloudiness in the COSMO-CLM model, should be concerted with an improvement of the too reflective clouds. If not, it is likely that the TOA OSR becomes too large, possibly leading to a cool bias in the model. It should be further investigated why clouds are too reflective in the COSMO-CLM. A better link between the radiation scheme and the model microphysics might be required to better capture the different cloud drop sizes in

different regimes. It is likely that a more advanced approach for microphysics, i.e. a two-moment particle size distribution, better captures the subtle changes in particles sizes. In a framework of climate projections, it is essential to investigate in details and correct this error compensation to prevent a biased representation of a changing climate.

References

Ban N, Schmidli J, Schaer C (2014) Evaluation of the convection-resolving regional climate modelling approach in decade-long simulations. J Geophys Res Atmos 119(13):7889–7907

Ban N, Schmidli J, Schär C (2015) Heavy precipitation in a changing climate: does short-term summer precipitation increase faster? Geophys Res Lett 42(4):1165–1172

Boberg F, Berg P, Thejll P, Gutowski WJ, Christensen JH (2009) Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ENSEMBLES models. Clim Dyn 35(7–8):1509–1520

Böhm U, Kücken M, Ahrens W, Block A, Hauffe D, Keuler K, Rockel B, Will A (2006) CLM—the climate version of LM: brief description and long-term applications. COSMO Newsl 6:225–235

Böhme T, Lipzig NV, Delobbe L, Goudenhoofdt E, Seifert A (2011) Evaluation of microphysical assumptions of the COSMO model using radar and rain gauge observations. Meteorol Zeitschrift 20(2):133–144

Brisson E, Demuzere M, Lipzig NPMV (2015) Modelling strategies for performing convection-permitting climate simulations. Meteorol, Zeitschrift (published online)

Brisson E, Demuzere M, Willems P, van Lipzig N (2015) Assessment of natural climate variability using a weather generator. Clim Dyn 44(1–2):495–508

Bugliaro L, Zinner T, Keil C, Mayer B, Hollman R, Reuter M, Thomas W (2011) Validation of cloud property retrievals with simulated satellite radiances: a case study for SEVIRI. Atmos Chem Phys 11(12):5603–5624

Chan SC, Kendon EJ, Fowler HJ, Blenkinsop S, Ferro CT, Stephenson DB (2012) Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? Clim Dyn 41(56):1475–1495

Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of this century. Clim Change 81(S1):7–30

Clark AJ, Gallus W, Chen T-C (2007) Comparison of the diurnal precipitation cycle in convection-resolving and non-convection-resolving mesoscale models. Mon Weather Rev 135(10):3456–3473

Decker M, Brunke MA, Wang Z, Sakaguchi K, X Z, G BM (2012) Evaluation of the reanalysis products from GSFC, NCEP, and ECMWF using flux tower observations. J Clim 25(6):1916–1944

Dee DP, Uppala SM, Simmons J, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda M, Balsamo G, Bauer P, Bechtold P, Beljaars CM, van de Berg L, Bidlot J, Bormann N, Delsol C, Dragani R, Fuentes M, Geer J, Haimberger L, Healy SB, Hersbach H, Hólm EV, Isaksen L, Kå llberg P, Köhler M, Matricardi M, McNally P, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thépaut J-N, Vitart F (2011) The ERA-interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137(656):553–597

Devis, van Lipzig N, Demuzere M (2014) A height dependent evaluation of wind and temperature over Europe in the CMIP5 Earth System Models. Clim Res 61:41–56

Devis A, Van Lipzig NPM, Demuzere M (2013) A new statistical approach to downscale wind speed distributions at a site in northern Europe. J Geophys Res Atmos 118:2272–2283

Doms G, Forstner J, Heis E, Herzog HJ, Raschendorfer M, Reinhardt T, Ritter B, Schrodin R, Schulz JP, Vogel G (2011) A description of the nonhydrostatic regional COSMO model part II: physical parameterization. Technical report, September

Fosser G, Khodayar S, Berg P (2015) Benefit of convection permitting climate model simulations in the representation of convective precipitation. Clim Dyn 44(1–2):45–60

Genthon C, Town MS, Six D, Favier V, Argentini S, Pellegrini A (2010) Meteorological atmospheric boundary layer measurements and ECMWF analyses during summer at Dome C, Antarctica. J Geophys Res Atmos 115(D5)

Giorgi F, Bates GT (1989) The climatological skill of a regional model over complex terrain. Mon Weather Rev 117(11):2325–2347

Goudenhoofdt E, Delobbe L (2013) Statistical characteristics of convective storms in Belgium derived from volumetric weather radar observations. J Appl Meteorol Climatol 52(4):918–934

Hanssen-Bauer I, Foerland E (2001) Verification and analysis of a climate simulation of temperature and pressure fields over Norway and Svalbard. Clim Res 16:225–235

Haylock MR, Hofstra N, Klein Tank MG, Klok EJ, Jones PD, New M (2008) A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. J Geophys Res 113(D20):11

Hazeleger W, Severijns C, Semmler T, Ştefănescu S, Yang S, Wang X, Wyser K, Dutra E, Baldasano JM, Bintanja R, Bougeault P, Caballero R, Ekman AML, Christensen JH, van den Hurk B, Jimenez P, Jones C, Kå llberg P, Koenigk T, McGrath R, Miranda P, Van Noije T, Palmer T, Parodi J, Schmith T, Selten F, Storelvmo T, Sterl A, Tapamo H, Vancoppenolle M, Viterbo P, Willén U (2010) EC-Earth: a seamless Earth-system prediction approach in action. Bull Am Meteorol Soc 91(10):1357–1363

Hohenegger C, Brockhaus P, Bretherton CS, Schär C (2009) The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. J Clim 22:5003–5020

Jaeger EB, Anders I, Lüthi D, Rockel B, Schär C, Seneviratne SI (2008) Analysis of ERA40-driven CLM simulations for Europe. Meteorol Zeitschrift 17(4):349–367

Kendon EJ, Roberts NM, Fowler HJ, Roberts MJ, Chan SC, Senior C (2014) Heavier summer downpours with climate change revealed by weather forecast resolution model. Nat Clim Chang 4(7):570–576

Kendon EJ, Roberts NM, Senior C, Roberts MJ (2012) Realism of rainfall in a very high-resolution regional climate model. J Clim 25(17):5791–5806

Kniffka A, Stengel M, Lockhoff M, Meirink JF (2013) Validation report SEVIRI cloud products edition 1, EUMETSAT satellite application facility on climate monitoring. www.cmsaf.eu. SAF/CM/DWDI/VAL/SEV/CLD. Technical report

Kothe S, Dobler A, A B, B A (2011) The radiation budget in a regional climate model. Clim Dyn 36(5–6):1023–1036

Kotlarski S, Keuler K, Christensen OB, Colette, Déqué M, Gobiet, Goergen K, Jacob D, Lüthi D, van Meijgaard E, Nikulin G, Schär C, Teichmann C, Vautard R, Warrach-Sagi K, Wulfmeyer V (2014) Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. Geosci Model Dev 7(4):1297–1333

Langhans W, Schmidli J, Fuhrer O, Bieri S, Schär C (2013) Long-term simulations of thermally-driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. J Appl Meteorol Climatol 1:130117155938002 (1984)

Lauwaet D, van Lipzig NPM, Kalthoff N, de Ridder K (2010) Impact of vegetation changes on a mesoscale convective system in West Africa. Meteorol Atmos Phys 107:109–122

Lauwaet D, van Lipzig NPM, Van Weverberg K, De Ridder K, Goyens C (2012) The precipitation response to the desiccation of Lake Chad. Q J R Meteorol Soc 138(April):707–719

Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG (2012) An overview of the Global historical climatology network-daily database. J Atmos Ocean Technol 29(7):897–910

Miloshevich LM, Paukkunen A, Vömel H, Oltmans SJ (2004) Development and validation of a time-lag correction for Vaisala radiosonde humidity measurements. J Atmos Ocean Technol 21:1305–1327

Miloshevich LM, Vömel H, Whiteman DN, Leblanc T (2009) Accuracy assessment and correction of Vaisala RS92 radiosonde water vapor measurements. J Geophys Res Atmos 114:1–23

Miloshevich LM, Vömel H, Whiteman DN, Lesht BM, Schmidlin FJ, Russo F (2006) Absolute accuracy of water vapor measurements from six operational radiosonde types launched during AWEX-G and implications for AIRS validation. J Geophys Res Atmos 111(9):1–25

Moradi I, Soden B, Ferraro R, Arkin P, Vömel H (2013) Assessing the quality of humidity measurements from global operational radiosonde sensors. J Geophys Res Atmos 118(14):8040–8053

Neggers R, Siebesma A (2013) Constraining a system of interacting parameterizations through multiple-parameter evaluation: tracing a compensating error between cloud vertical structure and cloud overlap. J Clim 26(17):6698–6715

Nieto S, Dolores Frías M, Rodríguez-Puebla C (2004) Assessing two different climatic models and the NCEP-NCAR reanalysis data for the description of winter precipitation in the Iberian Peninsula. Int J Climatol 24(3):361–376

Oreopoulos L, Khairoutdinov M (2003) Overlap properties of clouds generated by a cloud-resolving model. J Geophys Res Atmos 108(15):4479

Perkins SE, Pitman J, Holbrook NJ, McAneney J (2007) Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. J Clim 20(17):4356–4376

Pfeifroth U, Hollmann R, Ahrens B (2012) Cloud cover diurnal cycles in satellite data and regional climate model simulations. Meteorol Zeitschrift 21(6):551–560

Pincus R, Platnick S, Ackerman SA, Hemler RS, Hoffman RJP (2012) Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators. J Clim 25(13):4699–4720

Prein F, Gobiet A, Suklitsch M, Truhetz H, Awan NK, Keuler K, Georgievski G (2013) Added value of convection permitting seasonal simulations. Clim Dyn 41(9–10):2655–2677

Prein AF, Holland GJ, Rasmussen RM, Done J, Ikeda K, Clark MP, Liu CH (2013) Importance of regional climate model grid spacing for the simulation of heavy precipitation in the colorado headwaters. J Clim 26(13):4848–4857

Prein AF, Langhans W, Fosser G, Ferrone A, Ban N, Goergen K, Keller M, Tölle M, Gutjahr O, Feser F, Brisson E, Kollet S, Schmidli J, van Lipzig NPM, Leung R (2015) Demonstrations, prospects,

and challenges. Rev. Geophys, A review on regional convection-permitting climate modeling

Ritter B, Geleyn J-F (1992) A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. Mon Weather Rev 120(2):303–325

Rockel B, Will A, Hense A (2008) The regional climate model COSMO-CLM (CCLM). Meteorol Zeitschrift 17(4):347–348

Rossow WB, Schiffer R (1999) Advances in understanding clouds from ISCCP. Bull Am Meteorol Soc 80(11):2261–2287

Schroeder M, van Lipzig NPM, Ament F, Chaboureau J-P, Crewell S, Fischer J, Matthias V, van Meijgaard E, Walther A, Willen U (2006) Model predicted low-level cloud parameters part II: comparison with satellite remote sensing observations during the BALTEX bridge campaigns. Atmos Res 82(1–2):83–101

Stengel MS, Kniffka K, Meirink JFM, Lockhoff ML, Tan JT, Hollmann RH (2014) CLAAS: the CM SAF cloud property data set using SEVIRI. Atmos Chem Phys 14(8):4297–4311

Steppeler J, Doms G, Schättler U, Bitzer HW, Gassmann A, Damrath U, Gregoric G (2003) Meso-gamma scale forecasts using the nonhydrostatic model LM. Meteorol Atmos Phys 82(1–4):75–96

Tiedtke M (1989) A comprehensive mass flux scheme for cumulus parameterization in large-scale models. Mon Weather Rev 117(8):1779–1800

van der Linden P, JFB Mitchell (eds) (2009) Climate change and its impacts: summary of research and results from the ENSEMBLES project. Technical report, Met Office Hadley Centre, FitzRoy Road, Exeter EX13PB, UK

Van Malderen R, De Backer H (2010) A drop in upper tropospheric humidity in autumn 2001, as derived from radiosonde measurements at Uccle. Belgium J Geophys Res Atmos 11(November 2009):1–18

Van Weverberg K, van Lipzig N, Delobbe L, Vogelmann A (2012) The role of precipitation size distributions in km-scale NWP simulations of intense precipitation: evaluation of cloud properties and surface precipitation. Q J R Meteorol Soc 138(669B):2163–2181

Vömel H, Selkirk H, Miloshevich L, Valverde-Canossa J, Valdés J, Kyrö E, Kivi R, Stolz W, Peng G, Diaz J (2007) Radiation dry bias of the Vaisala RS92 humidity sensor. J Atmos Ocean Technol 24:953–963

Walther A, Jeong J-H, Nikulin G, Jones C, Chen D (2013) Evaluation of the warm season diurnal cycle of precipitation over Sweden simulated by the Rossby Centre regional climate model RCA3. Atmos Res 119:131–139

Weisman ML, Skamarock WC, Klemp JB (1997) The resolution dependence of explicitly modeled convective systems. Mon Weather Rev 125(4):527–548

Wicker LJ, Skamarock WC (2002) Time-splitting methods for elastic models using forward time schemes. Mon Weather Rev 130(8):2088–2097

Willems P, Van den Zegel B, Tran Q, Akurut M, Verhoest N, Vanhaute W-J (2014) Onderzoek naar het gebruik van neerslaggeneratoren voor het aanmaken van overstromingsrisicokaarten-2. Deelproces hydrologie. Eindrapport, KU Leuven & U.Gent, voor TWOL onderzoeksproject L 2011 T 0005 / Vlaamse Milieumaatschappij—Afdeling Operatione. Technical report

Winker DM, Vaughan M, Omar A, Hu Y, Powell K, Liu Z, Hunt WH, Young S (2009) Overview of the CALIPSO mission and CALIOP data processing algorithms. J Atmos Ocean Technol 26(11):2310–2323