Increasing resolution and resolving convection improves the simulation of cloud-radiative effects over the North Atlantic

Fabian Senf[1,1], Aiko Voigt[2,2], Nicolas Clerbaux[3,3], Hartwig M Deneke[1,1], and Anja Hünerbein[4,4]

[1]Leibniz Institute for Tropospheric Research
[2]Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Department Troposphere Research; Lamont-Doherty Earth Observatory, Columbia University
[3]Royal Meteorological Institute of Belgium
[4]Leibniz-Institute for Tropospheric Research

November 30, 2022

Abstract

Clouds interact with atmospheric radiation and substantially modify the Earth's energy budget. Cloud formation processes occur over a vast range of spatial and temporal scales which make their thorough numerical representation challenging. Therefore, the impact of parameter choices for simulations of cloud-radiative effects is assessed in the current study. Numerical experiments were carried out using the ICOsahedral Nonhydrostatic (ICON) model with varying grid spacings between 2.5 and 80 km and with different subgrid-scale parameterization approaches. Simulations have been performed over the North Atlantic with either one-moment or two-moment microphysics and with convection being parameterized or explicitly resolved by grid-scale dynamics. Simulated cloud-radiative effects are compared to products derived from Meteosat measurements. Furthermore, a sophisticated cloud classification algorithm is applied to understand the differences and dependencies of simulated and observed cloud-radiative effects. The cloud classification algorithm developed for the satellite observations is also applied to the simulation output based on synthetic infrared brightness temperatures, a novel approach that is not impacted by changing insolation and guarantees a consistent and fair comparison. It is found that flux biases originate equally from clear-sky and cloudy parts of the radiation field. Simulated cloud amounts and cloud-radiative effects are dominated by marine, shallow clouds, and their behaviour is highly resolution dependent. Bias compensation between shortwave and longwave flux biases, seen in the coarser simulations, is significantly diminished for higher resolutions. Based on the analysis results, it is argued that cloud-microphysical and cloud-radiative properties have to be adjusted to further improve agreement with observed cloud-radiative effects.

Increasing resolution and resolving convection improves the simulation of cloud-radiative effects over the North Atlantic

Fabian Senf[1], Aiko Voigt[2,3], Nicolas Clerbaux[4], Anja Hünerbein[1], Hartwig Deneke[1]

[1]Leibniz Institute for Tropospheric Research, Leipzig
[2]Institute for Meteorology and Climate Research - Department Troposphere Research, Karlsruhe Institute of Technology, Karlsruhe
[3]Lamont-Doherty Earth Observatory, Columbia University, New York, USA
[4]Royal Meteorological Institute of Belgium, Brussels

Key Points:

- biases in cloud-radiative effects become smaller as grid spacing is decreased, especially from 80 to 20 km
- refinements down to 2.5 km substantially reduce shortwave CRE biases only when the convection scheme is disabled
- compensating biases between longwave and shortwave become smaller for finer resolutions leading to more realistic radiation fluxes

Corresponding author: Fabian Senf, `senf@tropos.de`

Abstract

Clouds interact with atmospheric radiation and substantially modify the Earth's energy budget. Cloud formation processes occur over a vast range of spatial and temporal scales which make their thorough numerical representation challenging. Therefore, the impact of parameter choices for simulations of cloud-radiative effects is assessed in the current study. Numerical experiments were carried out using the ICOsahedral Nonhydrostatic (ICON) model with varying grid spacings between 2.5 and 80 km and with different subgrid-scale parameterization approaches. Simulations have been performed over the North Atlantic with either one-moment or two-moment microphysics and with convection being parameterized or explicitly resolved by grid-scale dynamics. Simulated cloud-radiative effects are compared to products derived from Meteosat measurements. Furthermore, a sophisticated cloud classification algorithm is applied to understand the differences and dependencies of simulated and observed cloud-radiative effects. The cloud classification algorithm developed for the satellite observations is also applied to the simulation output based on synthetic infrared brightness temperatures, a novel approach that is not impacted by changing insolation and guarantees a consistent and fair comparison. It is found that flux biases originate equally from clear-sky and cloudy parts of the radiation field. Simulated cloud amounts and cloud-radiative effects are dominated by marine, shallow clouds, and their behaviour is highly resolution dependent. Bias compensation between shortwave and longwave flux biases, seen in the coarser simulations, is significantly diminished for higher resolutions. Based on the analysis results, it is argued that cloud-microphysical and cloud-radiative properties have to be adjusted to further improve agreement with observed cloud-radiative effects.

Plain Language Summary

Clouds are a major challenge for climate science and their effects are difficult to quantify. Clouds scatter sunlight back into space and thus prevent the Earth from warming up. But clouds also hold back heat radiation upwelling from the surface. Both effects typically compensate each other and thus lead to the net cloud-radiative effect. Computer programs that are used to simulate the climate - so-called climate models - often use very coarse grid-box sizes in their computational mesh. Cloud processes and their effects are represented in them in a very simplified way, which leads to problems. For this reason, this study deals with the question to what extent the simulations of cloud-radiative effects can be improved by choosing more precise descriptions of the cloud processes. To investigate this, different configurations of more realistic models were taken to simulate cloud formation over the North Atlantic. The resulting simulation data were compared to satellite observations. It could be shown that problematic biases of the coarser climate models are reduced if, as is usual in weather models, one switches to smaller grid-box sizes and improved descriptions of the cloud processes.

1 Introduction

Clouds are very effective in cooling the Earth. Clouds scatter sunlight back to space before it can be absorbed by the Earth's surface. They also trap longwave radiation originating from the warm surface and thus induce a counter-acting greenhouse effect (Ramanathan et al., 1989). In the global mean, the shortwave effect of clouds (-46 to $-48\ \mathrm{W\,m^{-2}}$) dominates over their longwave effect (26 to 28 $\mathrm{W\,m^{-2}}$) in the top-of-the-atmosphere (TOA) radiation budget, leading to a net negative cloud-radiative effect (CRE) of -18 to $-20\ \mathrm{W\,m^{-2}}$ (Arking, 1991; G. L. Stephens et al., 2012; Henderson et al., 2013; Zelinka et al., 2017). The magnitude of net radiative effects becomes even larger and more important for cloud systems over the mid-latitude oceans, where the net CRE is more than twice the global average (see e.g. Zelinka et al., 2017).

Cloud feedbacks, i.e. the impact of changes in clouds on the TOA radiation budget, remain a major source of uncertainty in future climate projections (Boucher et al., 2013; Ceppi et al., 2017). Recent work indicates that the global-mean cloud feedback to global warming is likely positive, i.e., cloud changes will lead to an additional warming (Ceppi et al., 2017). This is largely attributed to a reduction in low-level cloud amount and a rise of high-level clouds (Zelinka et al., 2017). Yet, significant uncertainties remain in the parameterization of clouds and their radiative effects, in particular regarding the treatment of cloud microphysical processes in climate models (Gettelman & Sherwood, 2016). Understanding clouds and their radiative changes is also relevant for regional climate change, as the simulated response of the atmospheric circulation to global warming is strongly shaped by clouds (Voigt & Shaw, 2015; Voigt et al., 2019; Ceppi & Shepherd, 2017).

The steady increase in computational power and advent of a new generation of models that can harness this power has begun to allow for global atmospheric simulations with horizontal grid spacings on the order of a few kilometers (e.g. Satoh et al., 2018; Stevens et al., 2019). In these high-resolution simulations, clouds and the atmospheric flow interact much more naturally than in current low-resolution models typically run horizontal grid spacings of around 50 km. The explicit simulation of at least part of the cloud-scale circulations in fact provides a physical link between the resolved atmospheric flow and the parameterized cloud-microphysical processes (Satoh et al., 2019; Stevens et al., 2020). Moreover, and importantly, high-resolution models and satellite observations probe the atmosphere on similar spatial and temporal scales, allowing for a meaningful comparison between simulation and observations that helps model evaluation as well as the interpretation of observations (Satoh et al., 2019). As such, high-resolution modelling might break the so-called cloud parameterization "deadlock" (Randall et al., 2003) and promises to lead to more reliable simulations of cloud and precipitation responses to future climate change (Roberts et al., 2018; Collins et al., 2018; Stevens et al., 2020).

Motivated by these advances, we consider the radiative effects of mid-latitude cloud systems in simulations with a large range of horizontal resolutions, with three different treatments of atmospheric convection, and with two different treatments of cloud microphysics in this study. This creates a hierarchy of simulations that at the one end resembles current low-resolution climate models with parameterized convection and relatively simple cloud microphysics, and at the other end resembles the next-generation high-resolution models with explicit convection and more detailed cloud microphysics. Through this approach we investigate how a sequential reduction in model grid spacing from climate-model scales of 80 km down to 2.5 km affects, and hopefully improves, the simulation of cloud-radiative effects. Furthermore, we investigate the impact of subgrid-scale parameterization choices regarding convection (fully explicit convection vs. parameterized shallow convection vs. parameterized convection) and cloud microphysics (one-moment scheme vs. two-moment scheme) on cloud-radiative effects and the radiation budget. To this end we analyze simulations with the ICOsahedral Nonhydrostatic (ICON) model (Zängl et al., 2014) over a large domain of the North Atlantic. Our work contributes to recent efforts to understand the sensitivity of climate simulations with respect to horizontal resolution and convection parameterization (Webb et al., 2015; Haarsma et al., 2016; Evans et al., 2017; Maher et al., 2018; Thomas et al., 2018; Vannière et al., 2019). We expand these efforts by bridging the gap between current climate models and and convection-permitting models.

The focus region of this study is the mid-latitude North Atlantic. This is motivated on the one hand by its importance for current and future European weather, and on the other hand by the difficulties of current coarse-resolution global climate models to represent the radiative effects of mid-latitude clouds (Bodas-Salcedo et al., 2014; Voigt et al., 2019) and their coupling to the circulation (Grise & Polvani, 2014). Cloud-radiative effects in the mid-latitudes feed back onto circulations. As such, they are essential to anticipated poleward shift and strengthening of the eddy-driven jet streams under global warming (Voigt & Shaw, 2016; Albern et al., 2019; Ceppi & Hartmann, 2016; Li et al., 2019), and they also can impact mid-latitude weather on time-scales of days (Schäfer & Voigt, 2018; Grise et al., 2019)

Biases in simulated mid-latitude CREs appear to be primarily due to deficiencies in parameterized physics of clouds and convection (Ceppi & Hartmann, 2015). These physics strongly depend on cloud type. Analysis of data from space-born imaging radiometers has shown that low-level clouds over the oceans provide the largest contribution to the net TOA CREs because reflection of sunlight dominates over the trapping of longwave radiation (Hartmann et al., 1992; Ockert-Bell & Hartmann, 1992; Chen et al., 2000). The traditional cloud classification approaches have been revised to assess the importance of cloud regimes as a whole using clustering techniques (Oreopoulos & Rossow, 2011; Oreopoulos et al., 2016; McDonald & Parsons, 2018) and the vertical structure of cloud fields based on active satellite sensors (G. Stephens et al., 2018; L'Ecuyer et al., 2019). The latter showed that clouds are predominantly organized in multiple layers, which is typically not resolved by passive imagery. Because active satellite observations are very sparse in time and space, we here nevertheless rely on the traditional cloud classification approach to separate cloud-cover and CRE model biases into contributions from different cloud types. The comparison is based on instantaneous and high-resolution geostationary satellite data. We follow modern model evaluation standards and sequentially derive synthetic satellite observations using a satellite simulator (similar to Bodas-Salcedo et al., 2011; Pincus et al., 2012; Matsui et al., 2019) and cloud products with an advanced cloud classification software. For the latter step, we apply the cloud classification consistently for the full diurnal cycle (including nighttime). This improves the attribution of instantaneous CREs to different cloud types.

The paper is organized as follows: In section 2, the setup of the ICON model simulations and sensitivity studies is described. Sect. 2 also provides information on the observed and synthetic narrow-band satellite radiances that are forwarded into the cloud classification software and on our method for deriving TOA radiation fluxes from Meteosat observations. Sect. 3 presents the main results. We first consider domain-averaged radiation fluxes and CREs, and then split cloud cover and radiative effects into contributions from different cloud types. A summary and conclusions are given in section 4. A more detailed description of the modifications of the cloud classification software and supporting information is provided in the supplement.

2 Data and Methods

2.1 Overview of the Analyses Workflow

Before we provide more details, Fig. 1 presents an overview of the workflow and analyses steps for observations (black) and simulations (blue). Used acronyms are listed in Tab. 1. The diagram is to be read from top to bottom. The input data from Meteosat SEVIRI (see Sect. 2.2) and ICON (see Sect. 2.3) are provided in the first row. From these, observed and simulated cloud types (Fig. 1a) and CREs (Fig. 1b) are derived, as shown in the last row. Importantly, this workflow makes sure that observations and simulations are directly comparable to each other.

For cloud classification, ICON simulations are translated into observation space using the SynSat forward operator (Sect. 2.3). Based on observed and synthetic infrared BTs, cloud types are derived with the help of the NWCSAF v2013 software (Sect. 2.4). For the assessment of CREs, Meteosat SEVIRI data are processed to obtain GERB-like all-sky radiation fluxes at the top of the atmosphere (Sect. 2.2). The observed all-sky fluxes are supplemented by simulated clear-sky fluxes, which are corrected with a scal-

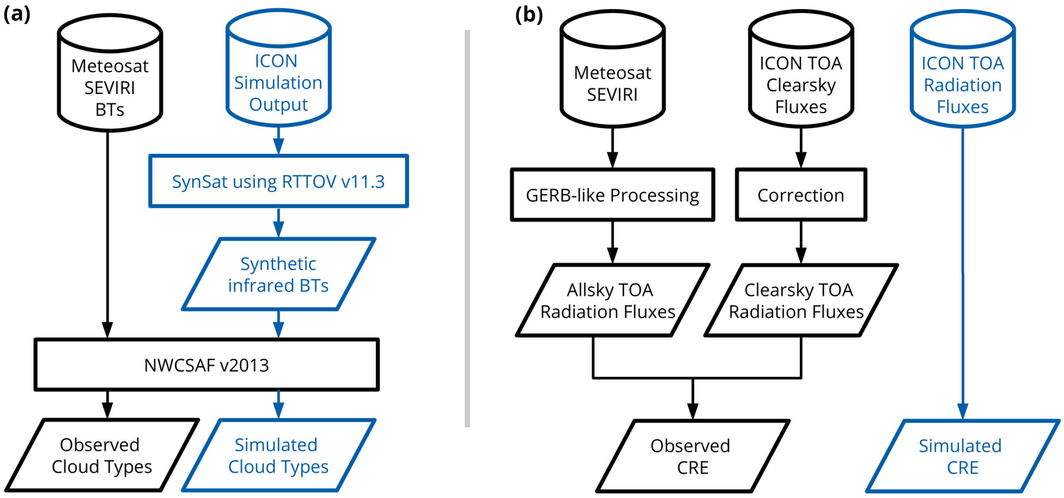

Figure 1. Overview of the workflow for (a) the calculation of a consistent cloud classification and (b) the derivation of CREs. Two parallel paths for observations (black) and the simulations (blue) are shown. The symbols in the top row visualize the input data (either satellite data archive or simulation output). Final data are shown in the last row. Rectangles denote processing methods further discussed in the text, and slanted parallelograms correspond to intermediate and final data.

Table 1. List of most important acronyms.

Acronym	Explanation
BT	Brightness Temperature
CRE	Cloud-Radiative Effect
GERB	Geostationary Earth Radiation Budget
ICON	ICOsahedral Nonhydrostatic
NAWDEX	North Atlantic Waveguide and Downstream impact EXperiment
NWCSAF	Satellite Application Facility in support to NoWCasting and very short range forecasting
RTTOV	Radiative Transfer for Television infrared observation satellite Operational Vertical sounder
RRTM	Rapid Radiation Transfer Model
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SynSat	synthetic satellite imagery
TOA	Top-Of-the-Atmosphere

ing factor in the shortwave and a constant additive offset in the longwave part to correct for biases in simulated ocean surface properties (Sect. 2.5).

2.2 Meteosat Observations

Observations are provided by measurements of the imaging radiometer SEVIRI (Spinning Enhanced Visible and InfraRed Imager) on board the geostationary satellites of the Meteosat Second Generation (MSG) series operated by EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites). We utilize multi-spectral data from SEVIRI's operational prime service located at a nominal longitude of zero degrees and a nadir resolution of $3{\times}3$ km^2 (Schmetz et al., 2002). An example of upwelling thermal radiation measured at 10.8 μm is provided in Fig. 2 (top row). In the atmospheric window at 10.8 μm, atmospheric gases are relatively transparent and thermal emission mainly originates from the Earth surface, from clouds or from a combination of the two (in case of semi-transparent or fractional clouds). High BTs typically represent clear regions, whereas low temperatures represent emission from high cirrus clouds. In the scene of Fig. 2, a low-pressure system is located in the Atlantic ocean. Its frontal cloud system, seen by the low BTs, extends towards the south and approaches the British Islands. In the western part of this low-pressure system, cold and rather dry air is advected southwards together with marine, low-level clouds that formed within the cold sector.

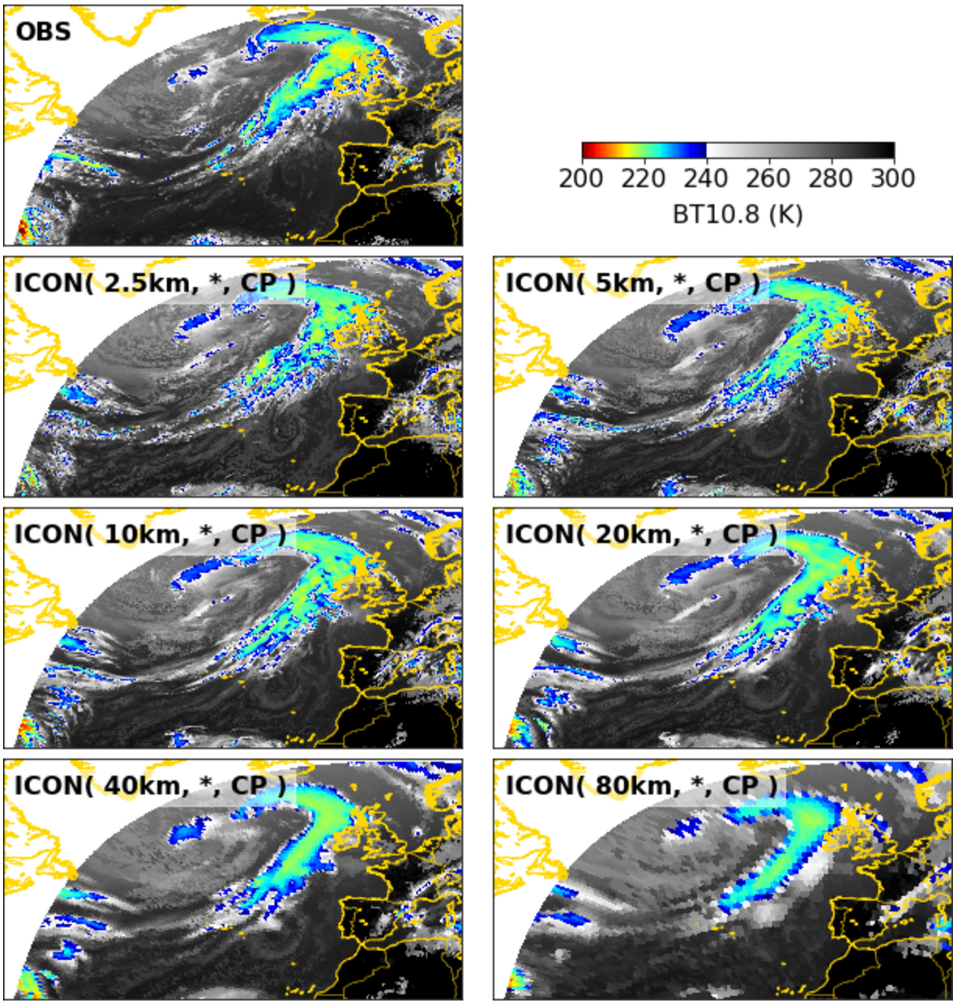

Figure 2. Overview of observed and simulated BTs from Meteosat SEVIRI's window channel at 10.8 μm for 1200 UTC 23 Sept 2016. Observations are compared to ICON simulations with increasing grid spacing (left to right and downwards, from 2.5 to 80 km). Only the subset of simulation experiments with one-moment microphysics and fully parameterized convection is chosen for visualization. A special color scheme is used to highlight observed and simulated features. BTs over land are also shown to improve anticipation of the cloud scenery. Further analysis however only considers the Atlantic ocean region.

The Meteosat satellites also carry the broadband radiometer GERB (Harries et al., 2005) for accurate measurements of all-sky TOA radiation fluxes. Unfortunately, during the period of our analysis GERB was in "safe mode" to protect its sensors. We therefore base our TOA radiation flux estimates on SEVIRI data. So-called GERB-like radiation flux products are derived as internal products in the Royal Meteorological Institute of Belgium (RMIB) GERB processing system which have been retrieved from the RMIB archive for our study. All GERB-like processing steps are explained in detail in Dewitte et al. (2008) and updates on the calibration of SEVIRI data are given in Meirink et al. (2013). The accuracy of the applied narrowband-to-broadband conversion is 3.5% for shortwave fluxes F_{sw} and 0.7% for longwave fluxes F_{lw} (Clerbaux et al., 2005). For a particular scene type, this error must be considered as a systematic error. For estimates of downwelling shortwave fluxes, temporal variations in the total solar irradiance are taken into account as described in Mekaoui and Dewitte (2008). Throughout the paper, we use a positive-upward convention so that upwelling fluxes are positive and downwelling fluxes are negative (following G. L. Stephens, 2005).

2.3 ICON Simulations

We analyze simulations with the ICON model in limited-area setup performed over a large area of the North Atlantic (model version icon-2.1.00 with bug fixes for two-moment cloud microphysics). The simulations were already described in Stevens et al. (2020) (see their Fig. 3) and were performed in support to the NAWDEX field campaign of fall 2016 (Schäfler et al., 2018). The domain extends from 78°W to 40°E in longitudinal direction, and from 23°N to 80°N in latitudinal direction. ICON is used with the numerical weather prediction physics package in a setup that largely follows the tropical Atlantic setup of Klocke et al. (2017). ICON is initialized from the Integrated Forecast System (IFS) analysis data of the European Center for Medium-Range Weather Forecasts (ECMWF) at 0 UTC. The lateral boundary data are taken from IFS at 3-hourly resolution. When available, i.e. at 0 UTC and 12 UTC, IFS analysis data are used. In between 3-hr, 6-hr and 9-hr IFS forecast data are used. The continually updated analysis and forecast data ensure that the model stays close to the actual meteorology over the simulation period over several days (see below). The IFS data is retrieved at the highest available resolution in space ($\sim$ 9 km horizontal grid spacing). 11 days are analyzed in total. These result from 4 simulation sets that each cover a time span of 3 or 4 days, and for which the first day is disregarded as spin-up. The simulations are listed in Tab. 2.

Table 2. List of days simulated with ICON during the period of the NAWDEX field campaign in fall 2016. N_{sim} is the number of simulations as a result of testing for the sensitivity with respect to horizontal resolution and the treatment of cloud microphysics and convection.

	Simulation period	Analyzed days	N_{sim}
Set 1	Sep 20:0UTC - Sep23:0UTC	Sep 21, 22	14
Set 2	Sep 22:0UTC - Sep26:0UTC	Sep 23, 24, 25	20
Set 3	Sep 29:0UTC - Oct02:0UTC	Sep 30, Oct 01, 02	14
Set 4	Oct 02:0UTC - Oct06:0UTC	Oct 03, 04, 05	14

The simulations are performed for six horizontal grid spacings of 80, 40, 20, 10, 5 and 2.5 km. In the vertical, always the same set of 75 levels is used. The thickness of the lowest model layer above ground is 20 m. The model layer thickness increases to $\approx$ 100 m at 1 km altitude above ground up to 1200 m at the model top of 30 km. Sweeping through the horizontal resolution allows us to cover both the horizontal resolution of present-day global climate models, which typically run at 50-100 km, as well as the resolution of existing convection-permitting regional climate simulations (Prein et al., 2015) and upcoming global simulations (Stevens et al., 2019), which run at 2-5 km. Depending on horizontal resolution, subgrid-scale convection is parameterized following Bechtold et al. (2008) based on the scheme of Tiedtke (1989). When fully enabled, the convection scheme interactively decides on the type of convection to be activated, either deep, mid-level or shallow convection. For the finest resolution of 2.5 km the convection parameterization scheme is switched off either fully or partly. In the latter setup, only shallow convection is parameterized, whereas mid-level and deep convection are explicitly represented (ICON Model Tutorial April 2018). The setup with only shallow convection parameterization has emerged as the standard setup for 2.5km-ICON simulations at the German Weather Service (pers. comm. A. Seifert). For resolutions of 5 km and coarser, the convection scheme is fully enabled and takes care of shallow as well as mid-level and deep convection. In addition, for a three-day subset (Sep 22, 23, and 24), the 2.5 km simulations are repeated with fully enabled convection parameterization, and the 5 and 10 km simulations with fully disabled convection parameterization. This allows us to compare the impact of the convection scheme with respect to changes in resolution. Besides assessing the impact of resolution and representing convection in an explicit or parameterized manner, we study the impact of representing cloud microphysics. To this end, all simulations are performed with the one-moment cloud microphysical scheme with graupel described in Baldauf et al. (2011) as well as with the two-moment cloud microphysical scheme of Seifert and Beheng (2006). The one-moment scheme is currently used operationally by the German Weather Service; the two-moment scheme is used in large-eddy mode simulations with ICON (Heinze et al., 2017).

To indicate the model setup in the plots and tables, the following nomenclature is used. For instance `ICON( 10km, *, CP )` refers to ICON simulations with 10 km grid spacing, one-moment microphysics and fully enabled convection parameterization. In contrast, `ICON( 2.5km, ** )` refers to ICON simulations with 2.5 km grid spacing, two-moment microphysics and fully disabled convection parameterization - a setup that is called "simulation with explicit convection" in the following. Lastly, `ICON( 2.5km, **, sCP )` refers to a simulation in which only the shallow convection parameterization is enabled. Tab. 3 summarizes the model setups.

Table 3. Overview of different treatment of convection for the four sets of simulations (see Tab. 2). sCP means that only the shallow convection scheme is active. CP means that convection is fully parameterized. A notation example is given in the last row for simulations with 2.5 km grid spacing and one-moment cloud microphysics (indicated by *; two-moment cloud microphysics are indicated by **).

	explicit convection	sCP	CP
Set 1, 3, 4	2.5 km	2.5 km	5 - 80 km
Set 2	2.5, 5, 10 km	2.5 km	2.5 - 80 km
Notation example	`ICON( 2.5km, * )`	`ICON( 2.5km, *, sCP )`	`ICON( 2.5km, *, CP )`

Radiative transfer is calculated by the global model version of the Rapid Radiation Transfer Model, RRTMG (Mlawer et al., 1997). RRTMG uses a reduced number of g-points (g is the relative rank of the atmospheric absorption coefficient within a wave length interval) for the correlated k-method to mitigate some of the computational burden of the parent RRTM model. 14 bands are used in the shortwave, 16 bands are used in the longwave. The solar constant is set to 1361.4 $\mathrm{Wm^{-2}}$. The diffuse ocean albedo is set to a constant value, $\alpha_{\mathrm{dif}} = 0.07$. The direct ocean albedo follows the radiation scheme of Ritter and Geleyn (1992) and is a function of the diffuse albedo and the solar zenith angle, μ_0,

$$\alpha_{\mathrm{dir}} = \frac{1 + 0.5\cos\mu_0\,(\alpha_{\mathrm{dif}}^{-1} - 1)}{(1 + \cos\mu_0\,(\alpha_{\mathrm{dif}}^{-1} - 1))^2}. \tag{1}$$

The maximum value allowed for α_{dir} is 0.999. The diffuse and the direct ocean albedos are independent of wavelength do not depend on surface roughness and wind speed. For cloud overlap, the generalized maximum-random overlap scheme of Hogan and Illingworth (2000) is used, with a vertical decorrelation length scale of 2 km. Ozone is specified according to the Global and regional Earth system Monitoring using Satellite and in situ data (GEMS) climatology (Hollingsworth et al., 2008), and aerosol according to the climatology of Tegen et al. (1997). Only aerosol-radiation-interactions are considered, aerosol-cloud interactions are not taken into account. The cloud droplet number used in the radiation for the effective radius of droplets and crystals follows a prescribed vertical profile taken from the global atmosphere model ECHAM6 (Stevens et al., 2013). Cloud optical properties, i.e., single scattering albedo, extinction coefficient and asymmetry factor, are also specified as in ECHAM6. Radiation is called every 12 minutes. The radiation fields are output every hour and are always consistent with the simulated cloud field, insolation, solar zenith angle and the state of the atmosphere and surface. Simulated radiation fluxes were re-gridded onto the observational grid (Sect. 2.2). The analysis is restricted to ocean areas free from sea ice, which avoids complications from differences in surface albedo. As such, the analysis domain includes the North Atlantic and connected water bodies, including the North sea and the Baltic sea (see e.g. Fig. 2 and Fig. 4). The southern boundary is at 28.3°N and is determined by the boundary nudging zone of the 80 km grid. A maximum satellite zenith angle of 75° marks the northern boundary of the domain.

For a fair comparison between observations and simulations, the simulated data have to be transformed into the observational space using forward operators (or sometimes called instrument simulators). This has become a standard approach in the last decades (Morcrette, 1991; Roca et al., 1997; Chaboureau et al., 2000) and is especially important when such ambiguous variables like cloud cover and cloud types are taken into consideration (e.g. Pincus et al., 2012). For our study, we apply the so-called SynSat operator after Keil et al. (2006) and Senf and Deneke (2017) to derive synthetic satellite images with the sensor characteristics of MSG SEVIRI. The SynSat operator prepares vertical profiles of atmospheric temperature, humidity, condensate content and subgrid-scale cloud cover as well as several surface variables to perform single-column radiative transfer calculations with the RTTOV model (Saunders et al., 1999; Matricardi et al., 2004), here version 11.3. Radiative transfer calculations are performed for different streams per vertical column which are combined using the maximum-random overlap assumption. We apply a standard configuration that has been operationally employed by the German Weather Service for several years and utilized for ICON simulations in previous studies (Heinze et al., 2017; Senf et al., 2018; Pscheidt et al., 2019). For this, diagnostic subgrid-scale cloud condensate content is added to its grid-scale counterpart, and ice and snow masses are simply combined to a frozen condensate content. Radiative properties of frozen condensate are estimated using relations for randomly-oriented hexagonal columns after Fu (1996) and McFarquhar et al. (2003). The derivation of synthetic BTs is impacted by uncertainties in the formulation of microphysical and radiative hydrometeor properties. A complicating fact is that different model parameterization handle hydrometeor properties differently leading to model-internal inconsistencies as additional cause for uncertainties in the forward calculations. Considering these issues and typical parameter variations, Senf and Deneke (2017) showed that uncertainties in BTs are in the order of a few Kelvin and largest for semi-transparent cirrus clouds with low cloud-top temperatures and with emissitivities close to 0.5.

Fig. 2 also provides a sequence of synthetic BTs for different model grid spacings from 2.5 to 80 km. As expected, the simulations capture the general cloud scenery and the synoptic-scale features very well. All simulations show the frontal cloud band that approaches the European continent and the upper-level trough located upstream in the North Atlantic. The coarser the resolution, the less detail can be seen in the synthetic

BT-fields. However, no abrupt quality changes appear to happen with increased grid spacing.

2.4 Cloud Classification

A cloud classification is derived from simulation and satellite data with the NWCSAF software version 2013. As input, the NWCSAF software expects multi-spectral data of MSG SEVIRI in its native data format. Using a set of several multi-spectral tests, a categorical classification is derived for all pixels classified as cloudy (Derrien & Le Gléau, 2005). The applied thresholds mainly depend on the illumination, the viewing geometry, the geographical location and numerical forecast data describing the moisture and thermodynamic structure at coarser resolution. For the latter, short-term IFS forecasts are supplied.

Cloud types are mainly distinguished by their cloud-top height and opacity similar to the ISCCP-approach (International Satellite Cloud Climatology Project, see e.g. Rossow and Schiffer (1999)). No further distinction between convective and stratiform cloud structures is performed. The typical properties of the NWCSAF cloud types are shown in Fig. 3 and contrasted to the categorization after Hartmann et al. (1992). For practical reasons, we consider planetary albedo instead of cloud-optical thickness as measure of cloud opacity. Clouds are divided into different height classes: very low, low, mid-level, high and very high clouds are approximately separated by cloud-top altitudes of 2, 3.5, 6.5 and 9.5 km. These values correspond to pressure levels of 800, 650, 450 and 300 hPa and to environmental temperatures of $+8$, 0, -18 and -40°C. Therefore, very low and low clouds are purely liquid clouds, mid-level and high cloud categories might contain a mixture of hydrometeor phases, and very high clouds are completely glaciated at cloud top. As shown in Fig. 3, the high and very high clouds are further subdivided by different opacity levels and called: semi-transparent (semi.) thin, semi. moderately thick, semi. thick cirrus as well as high and very high opaque clouds. We call all these categories together "cirrus clouds". The very high opaque clouds might also contain deep convective cores and parts of anvils close to upper-level convective outflow. An additional class is used for fractional clouds for which multi-spectral signatures of clouds and underlying surface are identified. Fractional clouds are typically made of small boundary-layer cumuli. The separation between this and the very-low cloud category is rather artificial. We therefore combine these two classes and end up with eight cloud types that will be utilized for further analysis. No undefined class exists, i.e. satellite pixels are either classified as cloud-free ($k = 0$) or cloudy ($k > 0$). Therefore, the total domain-average cloud cover can be estimated from the sum of fractions of the individual cloud types.

For very low / fractional clouds ($k = 1$ and $k = 9$ in Fig. 3), very low albedo values (close to the clear-sky albedo of $\sim$0.1) are most probable. This cloud type mainly consists of shallow clouds with low geometrical and optical thicknesses especially due to high sub-pixel variability and considerable clear-sky contributions. For more opaque clouds with higher cloud tops, averaged albedo shifts to higher values. These cloud types have larger vertical and horizontal extent, and thus higher cloud-optical thicknesses. A similar shift to higher albedo values is found for semi-transparent cirrus going from semi. thin ($k = 6$) to semi. moderately thick ($k = 7$) to semi. thick ($k = 8$). Cloud-spatial structures and sub-pixel variability might be also an important factor for the albedo of semi-transparent cloud categories.

The NWCSAF software has undergone more than a decade of development and is highly adjusted to the needs of operational forecasters and nowcasting applications. It tries to account for as much information as available to derive a comprehensive and instantaneous classification of the cloud field. Changes in solar illumination can lead to changes in product quality and systematic differences, especially between day- and night-

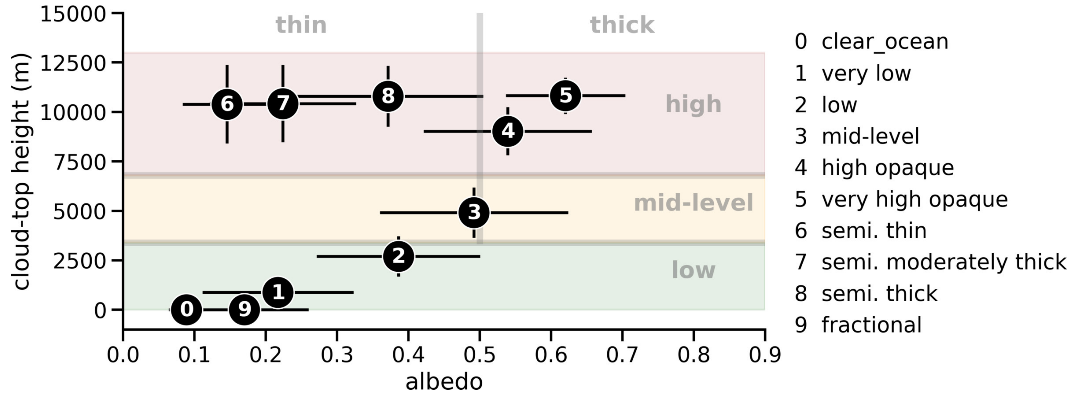

Figure 3. Planetary albedo versus cloud-top height for the different NWCSAF classes. The circles represent averages and the error bars give the standard deviation of clear-sky or cloud properties. Data have been taken from the observed scenery shown in Fig. 2 and 4. Numbers $k = \{0 \ldots 9\}$ refer to the different classes listed in the legend. Note that the cloud classes "fractional" and "very low" (which are shown separately here) are combined in the following analysis. For comparison, a second categorization after Hartmann et al. (1992) is provided as background image. It separates cloud amounts into three height categories (low, mid-level and high) as well as into two opacity levels (thin and thick clouds).

time, are inevitable in the standard setup of the NWCSAF cloud classification. To mitigate these problems and to build a time-consistent cloud classification, we implemented a modification to the cloud product generation chain. The NWCSAF software has been set up to run in permanent-night conditions at which only infrared radiation of terrestrial origin is utilized. We developed an algorithm which reads in infrared SEVIRI radiances from a selected scene and thereafter outputs these data into a template valid for the same day, but for 0 UTC. The template files, including the embedded satellite radiances, are supplied to the NWCSAF software which generates a cloud classification in night-mode. To keep the software itself unmodified, we provide simple estimates of radiances at 3.9 μm which are mandatory, but contaminated with sunlight during day-time (further explained in the supplement). Beyond time consistency, there is an other major advantage of our approach: It also allows to exchange real observations with synthetic observations. In our case, we utilized synthetic radiances derived from all the different simulations with the SynSat method (see Sect. 2.3) and provide these data to the NWCSAF software. In this way, a cloud classification is obtained for all simulations that is directly comparable to its observational counterpart.

An example scenery of an instantaneous and high-resolution cloud classification is shown in Fig. 4. The scene is similar to the one shown in Fig. 2, but here the focus is on 2.5 km simulations with different treatment of convection and cloud microphysics. A frontal cloud band extends from the British Island to the open Atlantic. West of this cold front, marine clouds of type "low" and "very low / fractional" propagate towards the European continent. In the subtropical areas, Meteosat observations show a rather low fraction of low and very low / fractional marine clouds. The amount of these cloud types, which appear in large patches of marine stratocumulus, is strongly overestimated. This is a common bias in all considered ICON simulations at 2.5 km, especially in the variants with explicit convection (see also (Senf et al., 2018)), and might reflect weaknesses in the setting and coupling of the convection scheme and planetary-boundary layer scheme.

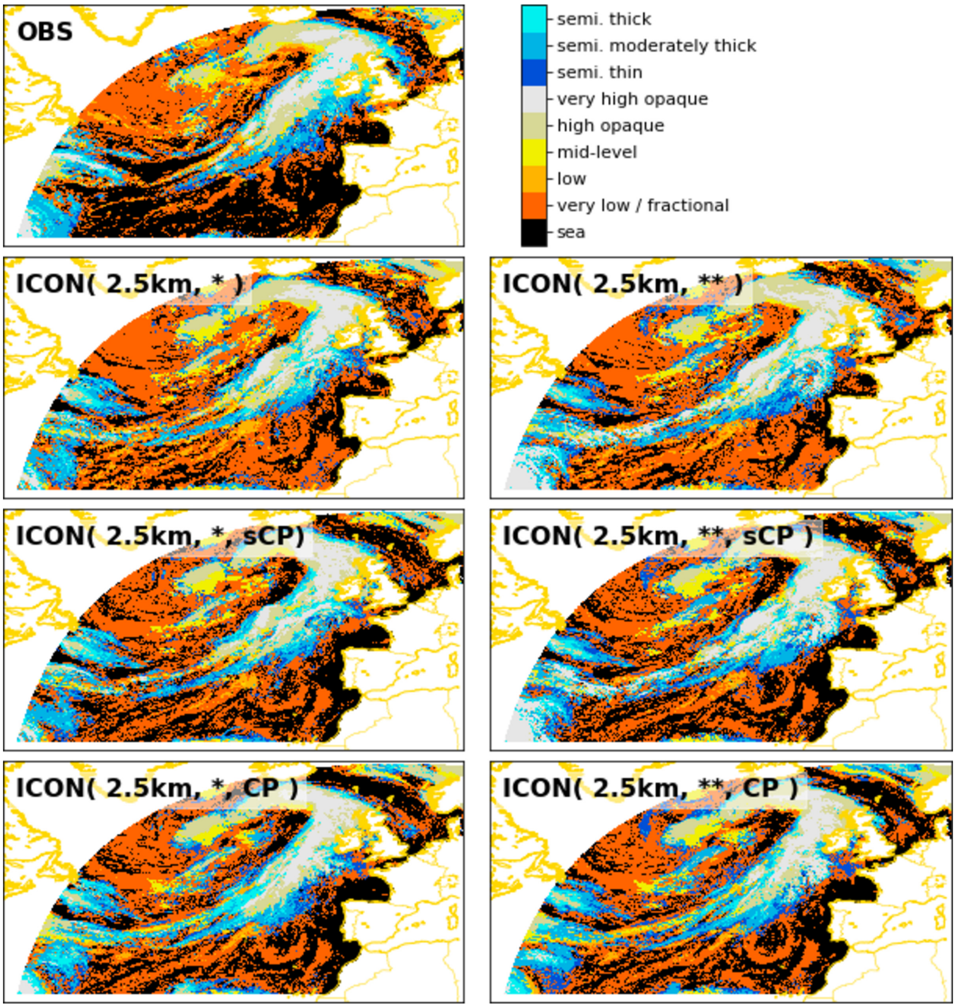

Figure 4. Example of observed and simulated cloud types for 1200 UTC 23 Sept 2016 as derived from Meteosat SEVIRI observations (top left) and ICON simulations with 2.5 km horizontal resolution. The left column is for simulations with one-moment cloud microphysics (*), the right column for simulations with two-moment microphysics (**). The second row is for fully explicit convection, the third row for simulations with a shallow convection scheme (sCP), and the fourth row for simulations with fully parameterized convection (CP).

2.5 Estimation of Observed Clear-Sky Radiation Fluxes

We are interested in the cloud impact on broadband shortwave and longwave radiation fluxes. This impact is commonly measured in terms of cloud-radiative effects (CREs),

$$\mathrm{CRE}_{\mathrm{net}} = \overline{F_{\mathrm{net,clear}}} - \overline{F_{\mathrm{net}}}\,, \tag{2}$$

which are defined as time-average difference between hypothetical clear-sky fluxes that would occur in the absence of clouds and cloud-affected all-sky fluxes. We follow the sign convention of G. L. Stephens (2005) and remind the reader that we defined upwelling all-sky and clear-sky fluxes as positive. Positive CREs indicate a gain of radiative energy and a warming effect of clouds, negative CREs indicate a loss of radiative energy and a cooling effect. Note that CREs are the net result of different cloud types; the radiative impact of individual cloud types is analyzed later in Sect. 3.2.

The ICON simulations provide all-sky and clear-sky fluxes, where the latter are calculated via a second radiation call with cloud fields set to zero. Simulated CREs follow directly from the application of eq. (2). Deriving clear-sky fluxes for the observations is more difficult. Observational clear-sky fluxes could be estimate from all-sky fluxes in regions classified as cloud-free, but these might contain undetected clouds and could be biased toward drier and more stable atmospheric conditions (Sohn et al., 2010). For our analysis the situation is even more challenging because (i) the North Atlantic is very cloudy, and (ii) we are interested in instantaneous high-resolution radiation fluxes and CREs, for which the clear-sky fluxes cannot be derived by temporal and spatial aggregation (as done in, e.g., Futyan and Russell (2005)). We therefore apply the following recipe to estimate observational clear-sky fluxes (clear-sky path in Fig. 1b):

(i) Clear-sky fluxes are taken from simulations as first guess (similar to Allan, 2011). The `ICON( 10km, *, CP )` experiment has been chosen as reference.
(ii) A bias correction is applied to simulated clear-sky fluxes under the constraint that the *radiative effects of undetected clouds have similar magnitudes in observations and simulations*.

The second step is based on the fact that for ICON simulations, differences between clear-sky and all-sky radiation fluxes are also available for regions that are classified as cloud-free ($k = 0$). As shown in Fig. 5, these differences are not zero and are caused by undetected clouds. We thus need to distinguish between all-sky and clear-sky fluxes in cloud-free regions. Therefore, a distinction between "cloud-free" and "clear-sky" is made throughout the rest of the paper.

The radiative effects of undetected clouds help us to establish a bias correction to translate simulated clear-sky fluxes into observational estimates (see also supplement) and to assess the quality of the NWCSAF cloud detection (modified by us to run in night-mode). For a perfect cloud classification, all values should be at zero. This is not the case, however, and this demonstrates that a small amount of clouds remains undetected. Undetected clouds from the simulations contribute around 3 $\mathrm{W\,m^{-2}}$ of additional shortwave reflection in cloud-free regions (Fig. 5a). In the longwave, simulated flux differences are between 1 and 2 $\mathrm{W\,m^{-2}}$ in cloud-free regions (Fig. 5b) and result from the reduced emission temperature of undetected clouds. The shortwave and longwave effects of undetected clouds partially cancel. When weighted by the fraction of cloud-free areas of around 25%, we conclude that CREs of undetected clouds have negligible impact on the total domain-average radiation budget.

Fig. 5 additionally shows two observational estimates of the effects of undetected clouds: one just takes uncorrected (first-guess) ICON clear-sky fluxes (gray symbols) and the other one uses bias-corrected ICON clear-sky fluxes (black symbols). It can be seen that the bias correction brings the observational estimates close to the simulations. The bias correction reduces the first-guess clear-sky fluxes by 4 to 6 $\mathrm{W\,m^{-2}}$ in the shortwave and by 2 $\mathrm{W\,m^{-2}}$ is the longwave. We believe the overestimation in the shortwave results from a too bright ocean surface albedo in ICON. Additional support for this interpretation comes from independent internal investigations by the German Weather Service (pers. comm. A. Seifert). Moreover, simulated ocean surface seems to be too warm causing an overestimation of outgoing longwave clear-sky fluxes that adds to the shortwave bias.

Technically, an offset of 2 $\mathrm{W\,m^{-2}}$ is subtracted from $F_{\mathrm{lw,clear}}$ as simple bias correction in the longwave. For the shortwave, it is more appropriate to apply a scaling factor to the upwelling flux $F_{\mathrm{sw,up,clear}}$ (see Fig. 6). A scaling factor of 0.88 brings the ICON curve approximately down to the observational curve. Fig. 6 also shows that all ICON simulations lie together closely. It is therefore of minor importance which ICON experiment is chosen as reference. After correction, the simulated clear-sky fluxes are used to-

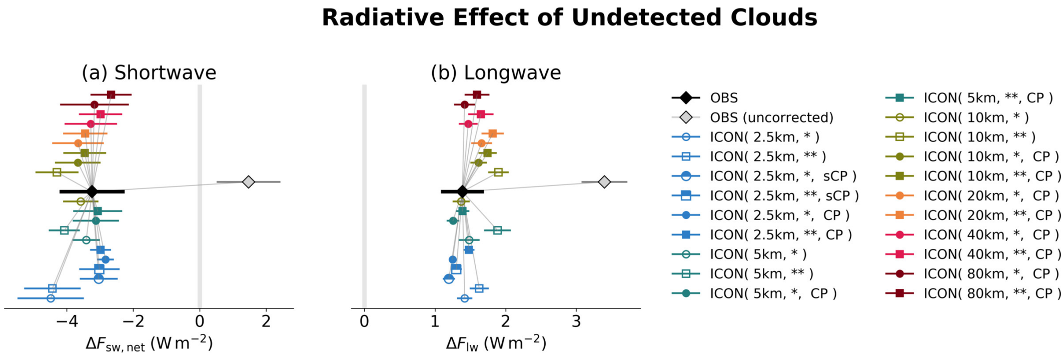

Figure 5. The radiative effect of undetected clouds in areas classified as cloud-free (i.e. $k = 0$). All data points show the average difference between clear-sky and all-sky fluxes for (a) shortwave $\Delta F_{\mathrm{sw,net}} = F_{\mathrm{sw,net,clear}} - F_{\mathrm{sw,net}}$ and (b) longwave $\Delta F_{\mathrm{lw}} = F_{\mathrm{lw,clear}} - F_{\mathrm{lw}}$. The bars give an robust estimate of the standard error of the daily-average values over all simulation sets, thus provide a confidence interval. For this, the difference between the 84-th and 16-th percentile has been calculated to approximate twice the multi-day standard deviation 2σ which was further divided by $\sqrt{N}$ with $N = 11$ for the all experiments except the additional runs from simulation set 2 (see Tab. 3). Colored symbols represent different simulations which have been vertically stacked to improve visibility. The gray symbols show the uncorrected observational estimate where the all-sky fluxes are based on Meteosat, but the clear-sky fluxes are directly taken from `ICON( 10km, *, CP )`. The black symbols show the corrected observational values with a scale factor applied to the shortwave and a constant additive offset to the longwave part of clear-sky fluxes taken from `ICON( 10km, *, CP )`. Thin gray lines connect all other symbols to the observation for improved interpretation. The clear-sky bias of the simulations is directly obtained from the difference between black and gray symbols.

gether with observed all-sky fluxes for the calculation of observed CREs using eq. (2). In summary, the applied strategy for cloud classification is extremely helpful to establish a consistent bias correction of instantaneous clear-sky fluxes estimated from simulations.

3 Results

3.1 Domain and Time-Averaged Radiation Fluxes and Cloud-Radiative Effects

We begin with a comparison of observed and simulated radiation fluxes averaged over the North Atlantic domain and all days (Fig. 7). The observed net flux is around 25 $\mathrm{W\,m^{-2}}$ and directed outward (Fig. 7a), implying that in this time of the year the North Atlantic region looses more radiative energy than it gains. All simulations show larger net fluxes, indicating that they overestimate the loss of radiative energy. Simulations with partly or fully parameterized convection have a net flux of around 30 $\mathrm{W\,m^{-2}}$, with the coarsest resolution showing the smallest deviation with respect to observations. Furthermore, simulations with fully parameterized convection have net fluxes slightly closer to the observation when using one-moment microphysics instead of two-moment microphysics. This might reflect previous model tuning that was done for one-moment but not for two-moment microphysics. Simulations with parameterized shallow convection show net fluxes very similar to simulations with fully parameterized convection. Much stronger deviations occur, however, for simulations with explicit convection, for which the net flux reaches

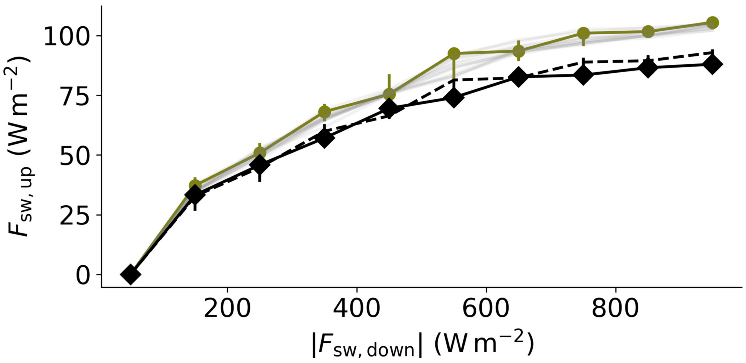

Figure 6. Simulated and observed upwelling versus downwelling shortwave fluxes in cloud-free areas. The upwelling flux is calculated for 10 bins of the downwelling flux. Symbols denote conditional median values and error bars show the inter-quartile range. Simulations are shown in gray, with the simulations for `ICON( 10km, *, CP )` shown in olive green. Observations are shown by the black diamonds and the black sold line. The dashed black line shows the upwelling flux from `ICON( 10km, *, CP )` rescaled by a factor of 0.88.

about 40 $\mathrm{W\,m^{-2}}$. We note that the deviations in the net flux are not simply a result of differences in the downwelling shortwave flux, which amount to 1 $\mathrm{W\,m^{-2}}$ due to slight differences in the solar constant in the simulations and observations.

The better agreement in terms of the net flux for low-resolution simulations and for simulations with (partly) parameterized convection results from compensating biases in outgoing longwave fluxes and upwelling shortwave fluxes (Fig. 7b and d). These compensating radiation flux biases are a known problem of a large number of climate models where tuning was aimed in particular at the net TOA energy balance (Klein et al., 2013). With one exception, the simulations overestimate outgoing longwave radiation (Fig. 7b), which corresponds to a too high effective emission temperature. The longwave bias increases with increasing grid spacing, with the largest bias found for the coarsest simulation at 80 km resolution. Simulations with fully parameterized convection underestimate upwelling shortwave radiation, which corresponds to a too low planetary albedo. Similar to the longwave bias, the shortwave bias is stronger for the coarser simulations. The better agreement in the net flux found for the coarser simulations is thus achieved for the wrong reason: a systematic bias compensation between longwave and shortwave fluxes that increases when a coarser resolution is used. Put differently, this also means that bias compensation becomes smaller as the resolution is made finer - an encouraging signature of convergence with increasing resolution. Similarly, Hohenegger et al. (2020) found that net shortwave TOA radiation shows a continuous improvement for successive grid refinements in their global ICON simulations with explicit convection.

For the highest resolution simulations at 2.5 km the outgoing longwave flux improves when the shallow-convection scheme is disabled so that convection becomes fully explicit. This is in particular the case for two-moment microphysics, which agrees best with observations in terms of the longwave flux (Fig. 7b). However, the simulations with fully explicit convection strongly overestimate the upwelling shortwave flux. As a result, the overall most satisfying agreement is found for simulations that combine two-moment microphysics and parameterized shallow convection. The shallow-convection parameterization avoids the strong overestimation of upwelling shortwave flux found for fully explicit convection.

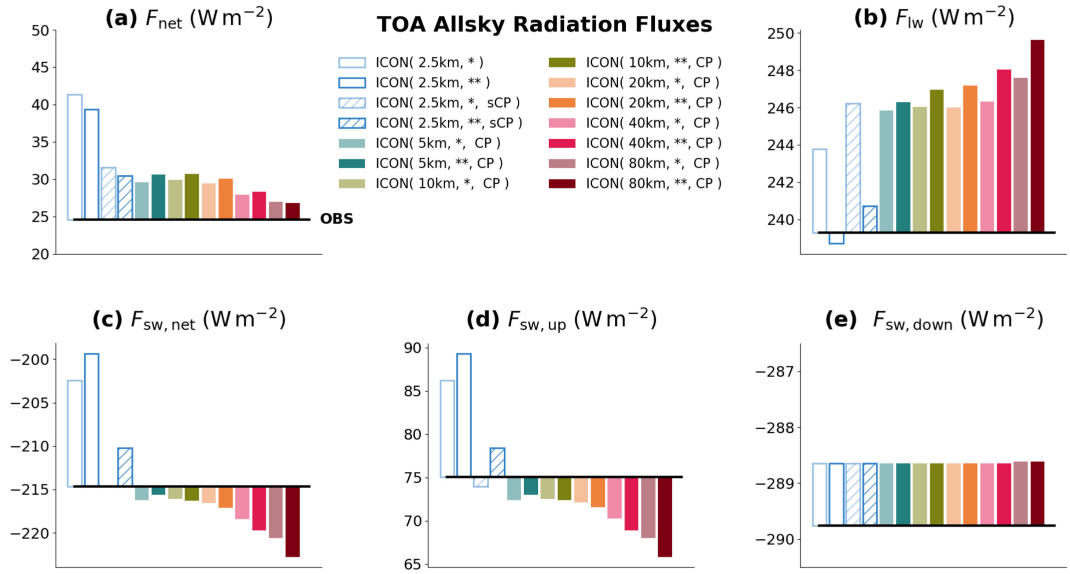

Figure 7. Domain and time-averaged all-sky radiation fluxes: (a) total net flux, (b) outgoing longwave flux, (c) net shortwave flux, (d) upwelling shortwave flux, and (e) downwelling shortwave flux. Observations are shown by the black horizontal lines. The deviations of simulated fluxes with respect to observations are shown by colored bars.

The simulation of domain- and time-averaged CREs and cloud cover is analyzed in Fig. 8. For the observations, CREs are around -41 W m^{-2} in the shortwave and around 27 W m^{-2} in the longwave, with a net cooling effect of clouds of -14 W m^{-2}. These CRE values are in the same range as global and long-term averaged observations. However, in the seasonal mean, twice as large CRE values would be found for the North Atlantic region (Zelinka et al., 2017). Simulated shortwave and longwave CREs are negatively correlated, with more positive longwave CREs obtained for more negative shortwave CREs (Fig. 8a). Simulations with fully parameterized convection lie in the upper left quadrant of Fig. 8a and thus underestimate the magnitude of both longwave and shortwave CREs. Although these simulations show some improvement with decreasing grid spacing, none of the simulations approaches the observed CREs, and the impact of resolution appears to saturate at grid spacings between 10 and 20 km. This indicates that even if the grid spacing was further reduced, the simulations would be unable to approach the observations if convection is fully parameterized. This idea is supported by Fig. S5 (supplementary material).

In contrast, simulations with shallow-convection scheme or with fully explicit convection are scattered around the observations (Fig. 8a). In these simulations, the impact of grid-scale cloud microphysics is also much more pronounced. This is because less or no subgrid-scale cloud condensate is produced by the convection parameterization, which has its own and much simpler convection microphysics description. Overall, this suggests a clear benefit from (partly) disabling the convection scheme. In fact, simulations with shallow-convection scheme and two-moment microphysics show a remarkable match with observed longwave and shortwave CREs.

Fig. 8b-d further shows the relation between CREs and cloud cover. In the observations, cloud cover is around 73%. Cloud cover is a primary control on CREs (e.g. Dolinar et al., 2015). Unsurprisingly this is visibile in the simulations, which show a near-linear relation between cloud cover and the CREs. In part, this clear relation is due to the fact that the analyses were only made for one particular model, the ICON model. Greater spread would be expected for the comparison of several models with different parameterizations (see Nam et al., 2012). For our analysis, the observations do not fall onto the simulation-based relationship. This leads to a dilemma: For none of the simulations do CREs and cloud cover at the same time match the observations. Cloud cover is better simulated for coarser grid spacings, whereas CREs improve as the grid spacing is refined. This indicates that the distribution of cloud-optical thicknesses and, associated with this, the vertical cloud structure is insufficiently represented in ICON.

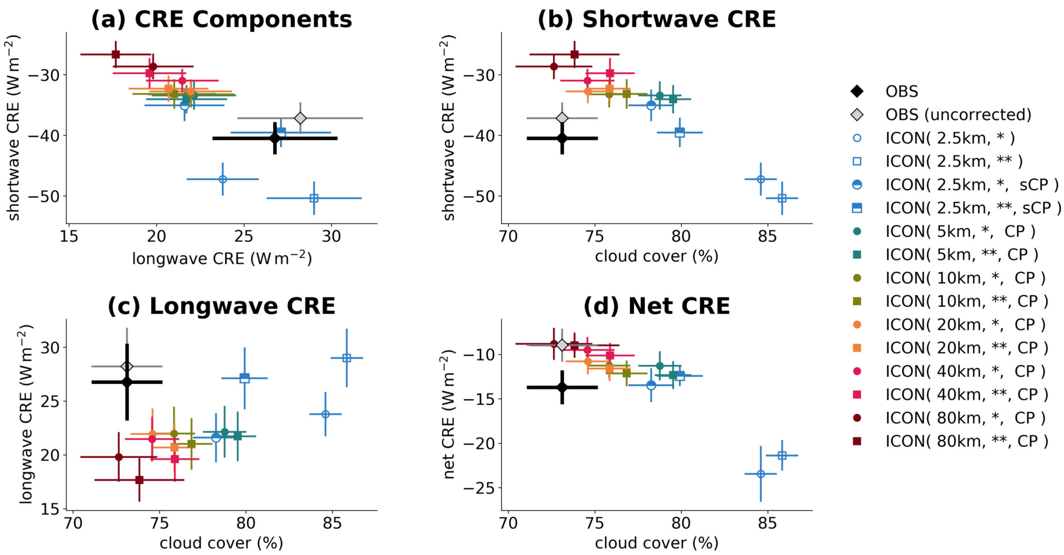

Figure 8. Comparison of domain- and time-averaged cloud-radiative effects and cloud cover: (a) longwave CRE vs. shortwave CRE. Cloud cover vs. (b) shortwave CRE, (c) longwave CRE, and (d) net CRE. Similar to Fig. 5, symbols denote average values and error bars provide confidence intervals. Please note the differences in the y-axis range.

Using Eq. (2) the radiation flux biases of the ICON simulations with respect to observations can be written as the sum of clear-sky and CRE biases, i.e.,

$$\delta\overline{F} = \overline{F_{\mathrm{ICON}}} - \overline{F_{\mathrm{OBS}}} = \delta\overline{F_{\mathrm{clear}}} - \delta\mathrm{CRE}. \tag{3}$$

The results of this decomposition are collected in Fig. 9, with net flux biases shown in the left column, shortwave flux biases in the middle column, and longwave flux biases in the right column. The matrix presentation of Fig. 9 allows for two implicit summing rules: the left column is the sum of the middle and right columns, and the first row is the sum of 2nd and 3rd rows. The second row of Fig. 9 shows that net biases are to a substantial extent due to clear-sky biases, which are independent of the simulation setup and amount to ~ 7.4 W m^{-2}. The biases in simulated clear-sky fluxes have already been identified in Sect. 2.5 where a correction for observational clear-sky estimates was constructed. The clear-sky bias mostly arises from the shortwave (~ 5.6 W m^{-2}), with a smaller longwave contribution (~ 1.8 W m^{-2}). The magnitude of the clear-sky shortwave bias is somewhat surprising, and likely reflects an imperfect representation of ocean surface albedo in the ICON simulations.

The dependence of all-sky flux biases on resolution and the treatment of convection and cloud microphysics results entirely from CREs (Fig. 9, third row). The net CRE bias counteracts the clear-sky bias and thus reduces the net all-sky bias for simulations with fully parameterized convection. For simulations with fully explicit convection, the net CRE bias adds to the clear-sky bias and therefore increases the net all-sky radiation bias. For simulations with parameterized shallow convection, the CRE biases depend on cloud microphysics. With one-moment microphysics, the CRE biases are similar to the biases found for fully parameterized convection. In contrast, with two-moment microphysics there is essentially no CRE bias, neither in the shortwave, longwave or net. The net flux bias of the two-moment simulation with parameterized shallow convection is therefore entirely due to clear-sky biases, which could be decreased by adjusting the ocean albedo.

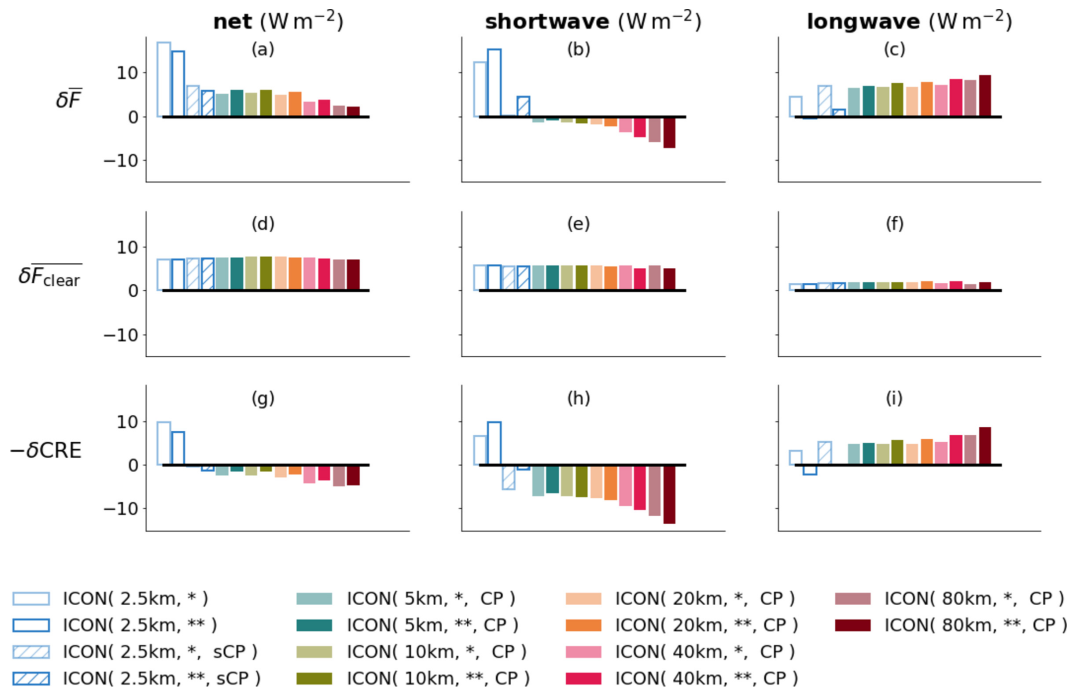

Figure 9. Decomposition of domain- and time-averaged biases for net (left), shortwave (middle) and outgoing longwave (right) radiation fluxes. The all-sky bias (1st row) is the sum of clear-sky (2nd row) and CRE (3rd row) biases. The clear-sky biases are calculated with respect to the bias-corrected clear-sky fluxes of `ICON( 10km, *, CP )`, which serves as observational reference.

The above analyses have already shown that CRE biases become smaller when the spatial resolution of ICON is refined. This effect is quantified more precisely in Fig. 10 which shows the resulting CRE biases and their changes for simulation set 2 (see Tab. 3). For set 2, additional simulations are available, which allow to assess the effect of grid refinement on CRE biases simulated with explicit convection. The magnitudes of shortwave CRE biases become larger for increasing grid spacing from 2.5 to 10 km and explicit convection (Fig. 10a). The sign of the longwave CRE bias depends on the choice of the microphysics scheme. For a detailed assessment of the resolution impact, simulation pairs were formed in which one simulation has half the grid spacing of the other simulation. Microphysics and convection parameterization were chosen identically. Absolute values of the CRE biases were subtracted from each other in such a way that a positive value indicates an improvement by grid refinement. It can be seen that refining resolution always improves shortwave CRE biases (Fig. 10c). The improvement is less pronounced for grid spacings less than 20 km and fully parameterized convection. This saturation can be overcome when fully explicit convection is used for which refinement down to 2.5 km provides substantial reduction of shortwave CRE biases. For the longwave, the behavior is different (Fig. 10d). Simulations with fully parameterized convection and two-moment microphysics experience continuous improvement with each refinement step down to 2.5 km. In contrast, longwave CRE biases simulated with explicit convection and one-moment microphysics even become worse when horizontal resolution is refined. As a further analysis, simulation pairs were formed which have the same resolution and convection parameters, but differ in terms of microphysics. Positive changes in CRE biases indicate improvements when switching to two-moment microphysics (Fig. 10e and f). For coarse resolutions, switching to two-moment microphysics leads to worse CRE biases in the longwave and in the shortwave. For smaller grid spacing and partly or fully parameterized convection, the sign changes and switching to two-moment microphysics can now lead to substantial improvements. For simulations with fully explicit convection, these improvement of CRE biases are only found in the longwave whereas switching microphysics causes unexpectedly increased biases in the shortwave. The clarification of the exact causes for the parameter dependencies found here requires further investigations.

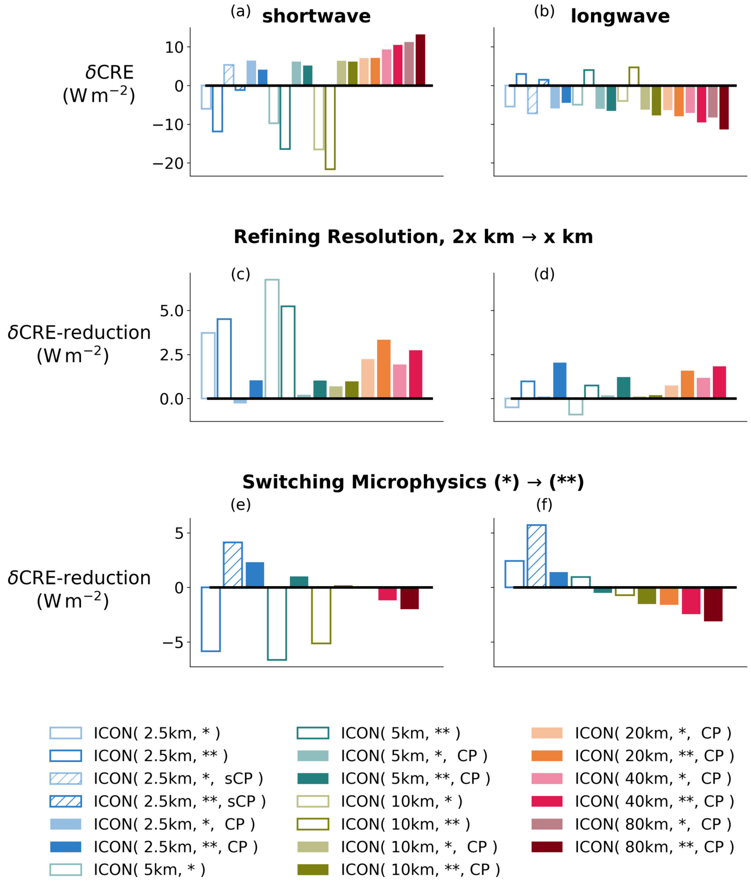

Figure 10. Impact of resolution and microphysics on CRE biases. Similar to Fig. 9h-i, CRE biases are shown for shortwave (left) and longwave (right), but only for simulation set 2 (see Tab. 3). Additionally, it is shown how CRE biases are reduced when resolution is refined (middle row) and microphysics is switched from the one-moment scheme to the two-moment scheme (bottom row). All other parameters were set equal and the ICON experiment, to which it is switched, is indicated in colored bars. A reduction of the CRE bias is shown with positive and increase with negative values.

3.2 Dependence of Cloud-Radiative Effects and Cloud Cover on Cloud Type

We now explore the origins of the domain- and time-averaged cloud-cover and CRE biases in the ICON simulations. To this end we use the cloud classification outlined in Sect. 2.4, which allows us to quantify the biases as a function of cloud type. This is done by writing the instantaneous domain-averaged net flux, F_{net}, as a sum of contributions from the K cloud types of the cloud classification,

$$F_{\mathrm{net}} = \sum_{k=0}^{K} f_k \, F_{\mathrm{net},k} \, , \tag{4}$$

where f_k is the fractional cloud cover of a certain cloud type k and $F_{\mathrm{net},k}$ is the instantaneous net flux averaged over the area covered by cloud type k. Areas classified as cloud-free are included at $k = 0$. As before a positive sign is taken for upwelling fluxes. Instantaneous domain- and time-averaged CREs are decomposed analogously,

$$\mathrm{CRE}_{\mathrm{net}} = -\sum_{k=0}^{K} \overline{f_k \left(F_{\mathrm{net},k} - F_{\mathrm{net,clear},k}\right)} \, , \tag{5}$$

where the cloud type-separated instantaneous net fluxes are averaged over time. This yields to a CRE decomposition into contributions from different cloud types. Note that clear-sky and cloud-free fluxes are not equal, $F_{\mathrm{net},0} \neq F_{\mathrm{net,clear},0}$, because of clouds that are undetected by the cloud classification (cf. Fig. 5).

Fig. 11 presents the cloud-type separation of total cloud cover. In the observations, cloud cover is dominated by very low / fractional clouds, which contribute around 30% to the total observed cloud cover of 73%. The three cloud types "low", "high opaque" and "semi. moderately thick" clouds each provide around 10%. The remaining cloud types are less important. From a qualitative point of view, all simulations capture the cloud cover of the different cloud types rather well. A few features of simulated cloud types, however, stand out:

(i) The cloud cover of very low / fractional clouds strongly depends on resolution and is better simulated in coarse-resolution simulations with grid spacings between 10 and 80 km. Finer-resolution simulations substantially overestimate very low / fractional cloud cover, with a more severe overestimation as the grid spacing is decreased. The largest overestimation is found for simulations with shallow of fully explicit convection.

(ii) Most simulations underestimate the low cloud cover and overestimate the cloud cover of semi-transparent clouds. These biases are less resolution dependent and become smaller when convection is fully explicit.

(iii) The choice of the microphysics scheme (one-moment vs. two-moment scheme) has a dominant impact on the cloud cover of cirrus clouds, which are represented by the five cloud types "high" and "very high opaque" as well as "semi. thin", "semi. moderately thick" and "semi. thick". The effect is evident for high and very high opaque clouds, for which the two-moment scheme produces smaller cloud cover than the one-moment scheme for fully parameterized convection but higher cloud cover for very high opaque clouds and parameterized shallow convection. At the same time, the two-moment scheme leads to increased cloud cover and cloud-cover biases for semi. thin and moderately thick clouds independent of the treatment of convection.

An overestimation of marine shallow cloud cover has already been observed in Senf et al. (2018), where ICON simulations were performed at 2.5 km grid spacing and with fully explicit convection. This persistent bias can also be found here and is a problem

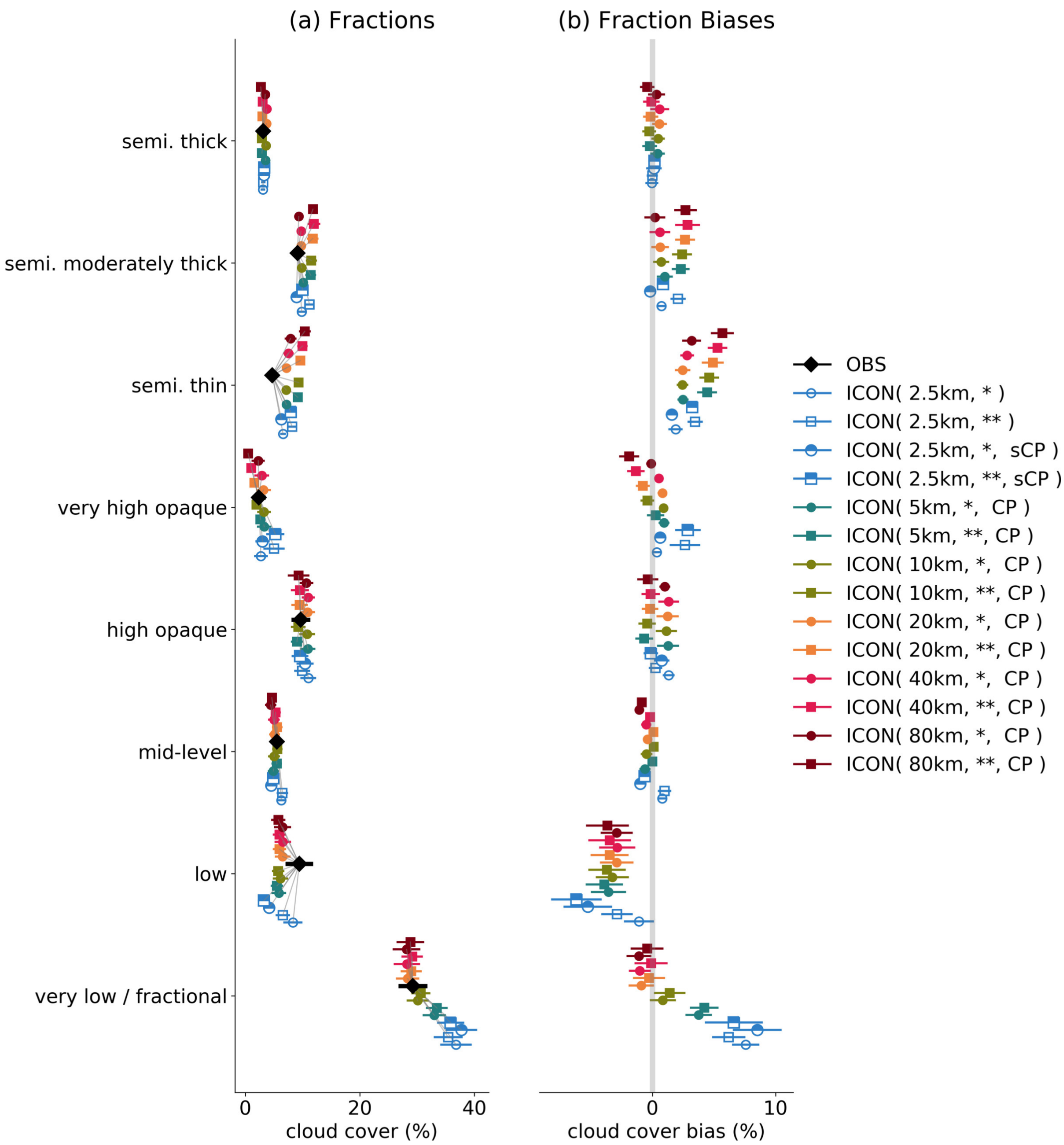

Figure 11. Observed and simulated cloud cover as a function of cloud type (a) as well as cloud cover biases of the simulations with respect to Meteosat observations (b). Similar to Fig. 5, symbols denote average values and error bars provide confidence intervals.

especially for simulated cloud coverage in the subtropical regions (see Fig. 4). A grid spacing of 2.5 km is still too coarse, so that the cloud-scale circulations are not sufficiently resolved. As result, too large and too regular structures of marine stratocumulus appear in the simulations.

To understand the microphysical sensitivity of the simulated cirrus clouds, it must be considered that the microphysics scheme in ICON was inherited from the weather model of the Consortium for Small-scale Modeling (COSMO). For COSMO a systematic overestimation of the cirrus cover was found (Böhme et al., 2011; Senf & Deneke, 2017). In order to eliminate this error, adjustments were made in the description of ice microphysics which reduce the optical thickness of cirrus clouds (Eikenberg et al., 2015; Köhler & Seifert, 2015). In the ICON simulations presented here, this may lead to a situation where semi-transparent cirrus is more overestimated by the two-moment scheme.

The domain- and time-averaged shortwave CRE depends on the typical albedo of a certain cloud type (see Fig. 3). This relation is further illustrated by Fig. 12a where CREs have been calculated for a hypothetical overcast situation in which the radiative effect of each cloud type was considered separately assuming a total coverage of 100%. Based on observations, very low / fractional clouds induce a rather low shortwave overcast CRE of -30 W m^{-2}. The shortwave overcast CRE increases reaching -140 W m^{-2} for very high, opaque clouds. The concurrent increase of albedo and cloud-top height also leads to increases in longwave overcast CREs. The imperfect compensation between short- and longwave CREs causes net effects that have different signs for observed opaque and observed semi-transparent cirrus clouds. All opaque clouds induce a net cooling due to their negative net CREs in the observation. For observed low and mid-level clouds, the magnitudes of net overcast CREs are largest with -50 W m^{-2}. The warming effect of observed semi-transparent clouds is less pronounced and is largest for semi. thick clouds with 15 W m^{-2}. Theses numbers are consistent with the findings of Chen et al. (2000) who attribute the largest negative shortwave CRE to their deep convective cloud type (comparable with our opaque very high category) and who also find a positive net CRE for their cirrus cloud type (comparable to our semi-transparent thin category).

The comparison of observed overcast CREs with their simulated counterparts helps to assess how good the different simulation setups represent the individual cloud-type specific radiation fluxes (independently of the fractional cloud cover of each type). On a qualitative level, all simulations perform very well showing the observed dependence of overcast CREs on cloud type. Most remarkably, none of the simulated semi-transparent cloud types causes significant positive net CREs (except for `ICON(2.5 km, **, sCP)`), i.e. hardly any of the ICON simulations induce a net domain-average warming from semi-transparent cirrus (see Fig. 12b). For all simulated semi-transparent cirrus cloud types, the longwave CREs and thus their thermal cloud emissitivies are underestimated (see Fig. 12a).

The dependence of all-sky CREs on cloud type is presented in Fig. 12c-d. Following eq. (5), all-sky CREs are calculated by weighting the difference between overcast and clear-sky radiation fluxes by the cloud cover of each cloud type. The relative amount of each cloud type determines the importance of this cloud type and its CREs for the domain- and time-average. Thus, simulated biases in all-sky CREs can arise from biases in (i) the radiative properties of a given cloud type, and (ii) the cloud cover of a given cloud type. Biases in radiative properties result from a misrepresentation of the distribution of cloud-optical thickness which is directly linked to the representation of vertical structure of the cloud type. Cloud-cover biases provide information on the misrepresentation of the horizontal extent of the respective cloud type. From Fig. 12d, we infer that mainly the four cloud types "very low / fractional", "low", "mid-level" and "high opaque" (with decreasing importance) contribute to the observed negative net all-sky CREs. The remaining four cloud types either have near zero net overcast CREs or too little cloud cover. For simulations with fully parameterized convection, the magnitudes of net all-sky CREs for very low / fractional and low clouds are severely underestimated. The discrepancy is much reduced for simulations with shallow convection at 2.5 km grid spacing, especially for one-moment microphysics. In contrast, the net all-sky CREs of very low / fractional clouds are overestimated in simulations with fully explicit convection. The all-sky net CREs of mid-level clouds are better represented for simulations with either shallow or full convection scheme than in simulations with fully explicit convection. In addition, semi. moderately thick clouds have too negative all-sky net CREs in all simulations, with the largest bias for simulations with fully explicit convection.

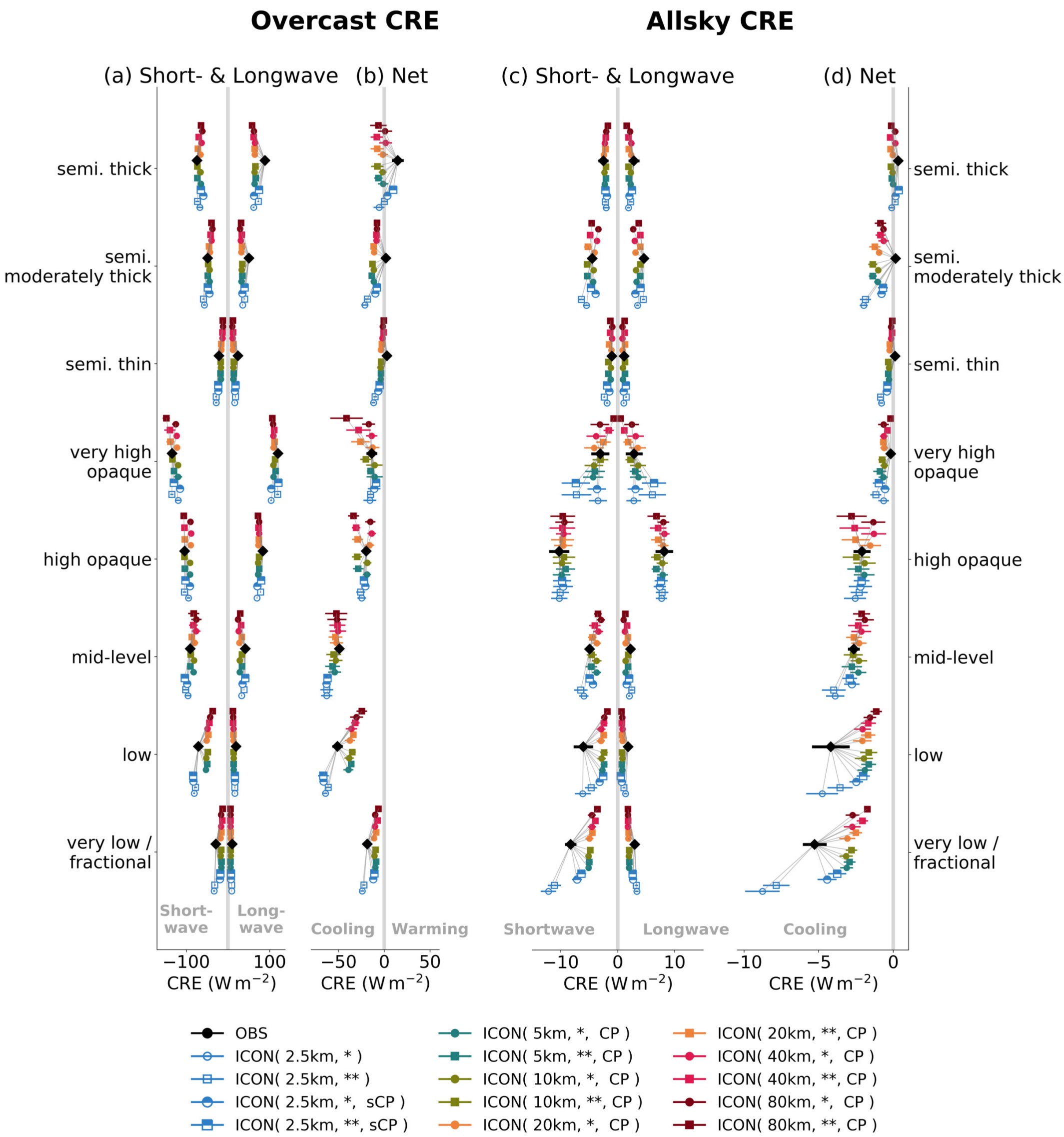

Figure 12. Observed and simulated (a,b) overcast CREs and (c,d) all-sky CREs for different cloud types. Overcast CREs are calculated assuming a hypothetical cloud cover of 100%. All-sky CREs include weighting by the cloud-type's specific cloud cover. Similar to Fig. 5, symbols denote average values and error bars provide confidence intervals.

To separate the effects of cloud type-dependent cloud cover and radiative properties on biases of simulated all-sky CREs, we apply a bias decomposition to eq. (5),

$$\begin{aligned}\delta\mathrm{CRE}_{\mathrm{net}} = & \underbrace{-\sum_{k=0}^{K}\overline{\delta f_k\,(F_{\mathrm{net},k} - F_{\mathrm{net,clear},k})}}_{\text{cloud cover}} \underbrace{-\sum_{k=0}^{K}\overline{f_k\,\delta(F_{\mathrm{net},k} - F_{\mathrm{net,clear},k})}}_{\text{radiative properties}} \\ & \underbrace{-\sum_{k=0}^{K}\overline{\delta f_k\,\delta(F_{\mathrm{net},k} - F_{\mathrm{net,clear},k})}}_{\text{co-variation}}\,.\end{aligned} \tag{6}$$

The first term results from a misrepresentation of cloud cover, the second term from a misrepresentation of radiative properties and overcast CREs, and the third term from the co-variation between the two factors. The "cloud cover" term shows how well the horizontal extend is simulated by cloud type. The "radiation flux" term is related to the vertical structure of a cloud type. As before, cloud-free contributions are included at $k = 0$. The decomposition holds for the all-sky net CREs as well as its shortwave and longwave components.

Fig. 13 summarizes biases in the domain- and time-averaged CREs and their decomposition. As discussed in Sect. 3.1, net CREs are biased negative for simulations with explicit convection, i.e. clouds cool too much, but biased positive for simulations with shallow-convection scheme and fully parameterized convection (except for `ICON(2.5km, *, sCP)`), i.e. clouds cool too little. For the latter simulations, net CRE biases become smaller as the grid spacing is decreased. The compensation of CRE biases originating in the longwave and shortwave is very apparent for fully convection-parameterized simulations (Fig. 13a-c).

The bias compensation between shortwave and longwave CREs leads to different roles of cloud cover and radiative properties, depending on whether one looks at net CREs or their shortwave and longwave components. For net CREs, cloud cover biases dominate. They are responsible for around half of the positive bias for fully parameterized convection (Fig. 13d). For simulations with fully explicit convection, in contrast, biases in radiative properties clearly control the net CRE biases. For the shortwave and longwave CRE components, biases in radiative properties dominate in general. A pronounced compensation between shortwave and longwave CRE biases is apparent. We thus find that the earlier discussed compensation of shortwave and longwave flux biases directly traces back to a misrepresentation of cloud-radiative properties. Switching from one-moment to two-moment microphysics has different effects on cloud-cover and radiative-properties related CRE biases. It is found for nearly all simulations that the shortwave and longwave CRE biases due to radiative properties become smaller. For the coarser simulations, the resulting improvement is more than compensated by biases in the "cloud cover" term. Thus, the CRE biases become larger when switching to the two-moment scheme in these coarser ICON experiments (see also Fig. 10e and f). The simulations with shallow-convection parameterization possess smaller biases than the fully parameterized simulations. The simulations with fully explicit convection show acceptable results for the longwave bias due to radiative properties. Their worse net performance originates from the missing compensation by shortwave biases which are also negative for these simulations.

The interpretation of CRE biases is further supported by Fig. 14 which provides a detailed bias decomposition separated by cloud type. We see that not only the compensation between shortwave and longwave CRE biases is important, but also the compensation of biases originating from different cloud types (Klein et al., 2013). For the net CRE biases (Fig. 14c), mainly cloud types "very low / fractional" and "low" contribute to the positive bias of simulations with fully parameterized convection. This is partially compensated by a negative net CRE bias from semi. moderately thick clouds. When split by cloud type, the net CRE bias of simulations with fully parameterized convection is dominated by CRE biases due to radiative properties.

For shortwave and longwave CRE biases (Fig. 14a,b), it is found that the resolution dependence of CRE biases not only originates from very low / fractional and low clouds, but also from very high opaque clouds. This cloud type is connected to deep convection which representation significantly improves for decreasing grid spacing. Especially, some simulations with two-moment microphysics show a rather poor performance for the very high opaque clouds. The coarse simulation at 80 km underestimates the fractional coverage of this cloud type, in contrast the simulation with shallow convection parameterization at 2.5 km overestimates the fractional coverage of very high opaque clouds (see also Fig. 11b). The spatial representation of this cloud type needs to be addressed

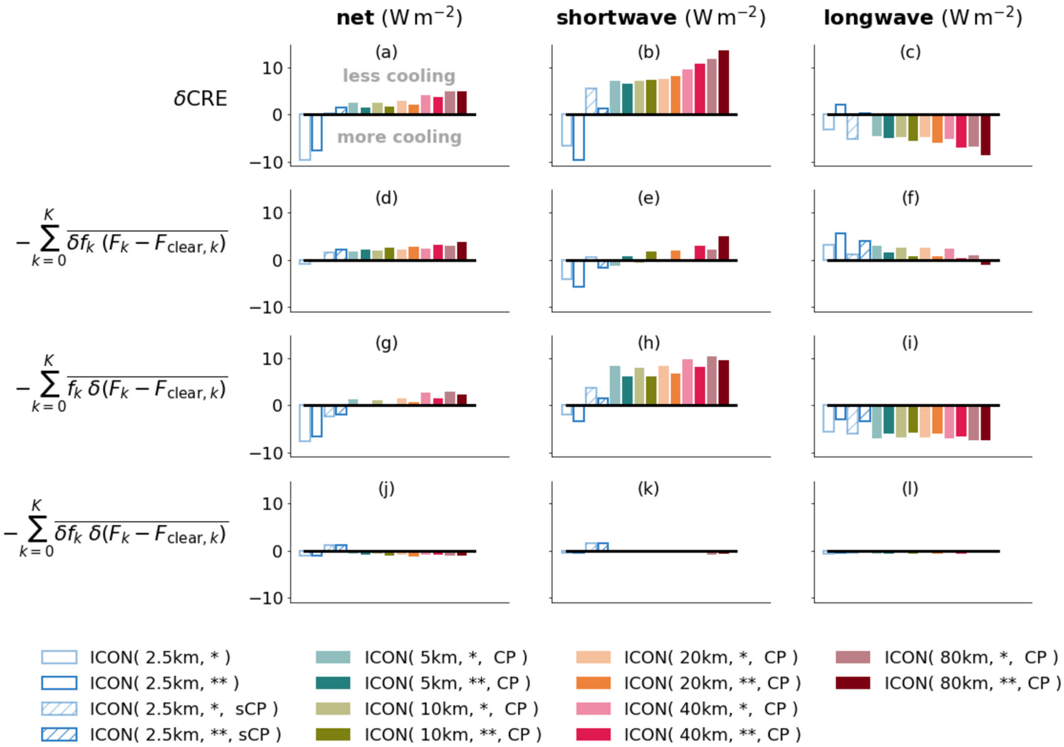

Figure 13. Decomposition of CRE biases (1st row) into contributions from biases in cloud cover (2nd row) and cloud-radiative properties (3rd row). Co-variations between biases in cloud cover and radiative properties are shown in the 4th row. The net CRE biases (left column) are decomposed into shortwave and longwave (middle and right columns) contributions.

in future. In the shortwave, the positive CRE bias of simulations with fully parameterized convection mainly comes from very low / fractional and low clouds. For the former, biases in radiative properties dominate whereas for the latter CRE biases due to cloud cover also contribute. Switching from one-moment to two-moment scheme, we find improvements in the representation of shortwave components of individual radiative properties (see Fig. 13h and Fig. 14g) which indicate that the vertical structure of clouds in terms of optical thicknesses has improved. These improvements are partially masked by worse cloud cover biases (see Fig. 13e). In the longwave, many cloud types simulated with fully parameterized convection show a negative bias originating from the bias in radiative properties. The magnitudes of the individual longwave biases are much smaller for simulations with explicit convection.

In summary, the above analysis showed that future model development should equally concentrate on improvements of simulated clear-sky and cloud-affected TOA radiation fluxes. For the former, we recommend to revise the formulation of ocean albedo to reach better consistency with observations. For CREs, strategies for further improvement depend on the choice of the convection scheme, especially at kilometer-scale resolutions. For simulations with fully parameterized convection, radiation is typically too weakly interacting with clouds, i.e. clouds appear too dark and too warm, especially for low and very low / fractional clouds. Hence, in contrast to the well known "too few, too bright" low-cloud problem of several climate models (Nam et al., 2012; Klein et al., 2013), low and very low clouds in ICON at coarse resolutions need to become brighter. This could be achieved by improving radiative properties of these cloud types, either from a macrophysical or a microphysical point of view. Specifically, in the used ICON version the ef-

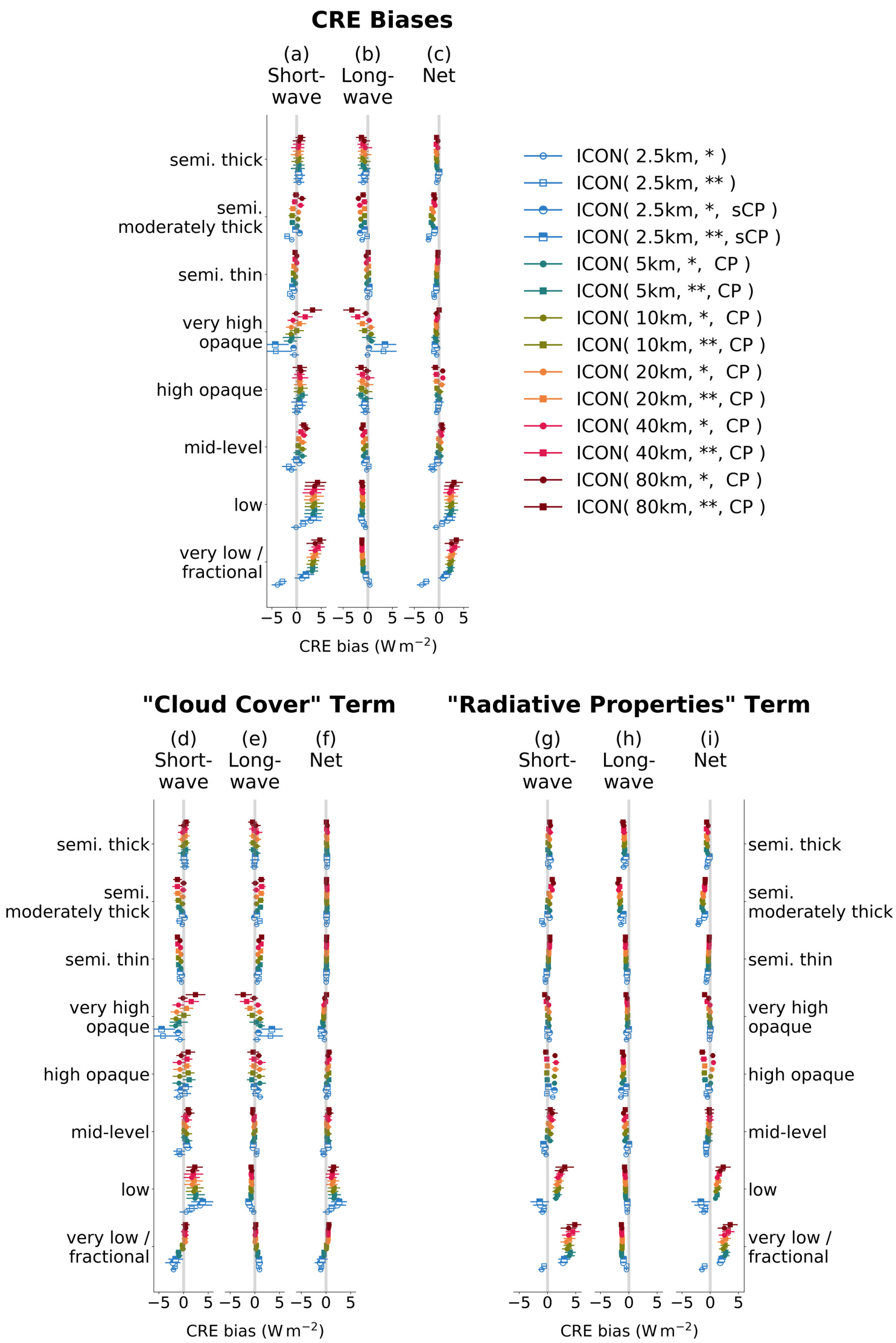

Figure 14. CRE biases and their decomposition for different cloud types. Following eq. (6), (top row) biases in CREs are separated into (bottom row) contribution from (left) cloud-cover biases and (right) radiation-flux biases. The split into (a, d, g) shortwave and (b, e, h) longwave components that sum up to the (c, f, i) net CRE bias is also provided in the different sub-panels. Similar to Fig. 5, symbols denote average values and error bars provide confidence intervals.

fective radius of cloud particles taken in the radiative transfer follows from a prescribed number concentration of cloud particles and is unaware of the number concentration simulated by the two-moment microphysics scheme. Adjusting this inconsistency might help to correct the CRE biases, e.g. the negative biases in longwave CREs of semi-transparent cirrus. For simulations with only shallow or fully explicit convection, the radiative properties of clouds show signs of improvement. However, ICON with shallow convection simulates very low clouds, which still appear too dark and too similar to the clouds from the fully parameterized convection simulations. To improve the representation of this cloud type, new parameterization approaches need to be explored, such as those using stochastic sampling (Sakradzija & Klocke, 2018). The simulations with explicit convection show a promising convergence, which should be investigated by further refinements down to the hectometer-scale (Stevens et al., 2020).

4 Conclusions and Outlook

Clouds regulate Earth's energy budget (Ramanathan et al., 1989). Shallow low-level clouds are efficient scatterers of shortwave radiation and, in combination with their small thermal contrast to Earth's surface, they have strong negative cloud-radiative effects and cool the Earth. In contrast, the cloud-radiative effects of high-level cirrus clouds also include longwave effects so that depending on cirrus-optical properties these clouds can either have a near zero or a warming effect (G. L. Stephens, 2005).

In mid-latitude environments, cyclones lead to the formation of frontal cloud bands with a complicated mixture of stratiform and convective clouds, possibly including multi-layer structures and embedded convection. Realistically representing such complex cloud structures and their radiative effects poses a challenge to numerical models, especially over oceans where extended shallow boundary-layer cloud fields occur in addition. Furthermore, the radiative impact of clouds on the mid-latitude circulation might depend on cloud type. We therefore investigated the ability of a specific numerical weather prediction - the ICON model (Zängl et al., 2014) - to represent cloud cover and cloud-radiative effects for selected days of the NAWDEX field campaign in boreal autumn 2016 over a large North Atlantic domain. Using a comprehensive set of sensitivity simulations that vary horizontal grid spacing between 2.5 and 80 km, we identified sensitivities with respect to model resolution. Moreover, we studied the impact of different choices regarding the parameterization of cloud microphysics (one-moment versus two-moment scheme) and convection (fully parameterized, shallow-convection only, fully explicit). This allowed us to identify strengths and weaknesses of the different model setups, in particular with respect to top-of-atmosphere radiation fluxes and cloud-radiative effects.

To assess the ICON model we made use of multi-spectral observations from the geostationary Meteosat satellite in two ways. First, we analyzed observational estimates of instantaneous top-of-atmosphere radiation. Second, we derived a detailed multi-spectral cloud classification from the Meteosat observations. For a consistent comparison between the ICON simulations and the observations, the simulation data were forwarded to a satellite forward operator performing radiative transfer calculations to derive synthetic infrared satellite images. This transfer of the simulations to observation space allowed us to subject simulations and observations to the same cloud classification software, and to analyze and compare observed and simulated cloud-type fields within the same framework.

In observations, the average net TOA radiation flux over the North Atlantic region and for the selected analysis days is around $+25\ \mathrm{W\,m^{-2}}$, indicating a net energy loss (remember that we adopted a positive-upward convention for radiation fluxes). Clouds substantially contribute to the energy loss and are responsible for a net cooling of $-14\ \mathrm{W\,m^{-2}}$. Major contributors to the net CRE are shallow clouds of the cloud type "very low / fractional" and "low", which both contribute around $-5\ \mathrm{W\,m^{-2}}$ to the total net CRE. The shallow clouds also account for around half of the total cloud cover of 73%.

The main results of our comparison between observed and ICON simulated radiation fluxes and cloud fields are as follows:

(i) For all model setups, the domain- and time-averaged net TOA radiation flux is larger than in the observations, independent of resolution and the treatment of cloud microphysics and convection. The ICON model thus overestimates the TOA loss of radiative energy. Simulations with fully parameterized convection underestimate TOA shortwave reflection and overestimate outgoing longwave radiation, i.e. seen from space they are too dark and too warm.

(ii) There is a systematic bias compensation between shortwave reflection and outgoing longwave radiation. The compensation is stronger for coarse-resolution simulations and becomes smaller for finer resolutions. Clear-sky and CRE biases have similar magnitudes, but only CRE biases are sensitive to horizontal resolution and in fact decrease with finer resolution. For fully parameterized-convection simulations, clouds are too weakly interacting with the radiation field leading to positive CRE biases in the shortwave and negative CRE biases in the longwave which partially compensate each other.

(iii) For none of the ICON setups, a simultaneous match between observed and simulated CREs and total cloud cover is achieved. Cloud cover compares better to observations for coarse resolutions, whereas CREs compares better to observations for finer resolutions.

(iv) The cloud cover of shallow clouds (types: "very low / fractional" and "low") strongly depends on resolution. It compares well with observations for coarser resolutions of 10-80 km, but finer resolutions and explicit convection severely overestimate it by up to 50% relative to observations. For simulations with fully parameterized convection, net CRE-biases of shallow clouds are dominated by positive shortwave biases in radiative properties. Biases in shortwave and net CREs are reduced when only shallow convection parameterization is applied. Using explicit convection even switches the sign of the shortwave CRE-biases leading to too bright shallow clouds and too large cloud-induced reflection.

(v) The choice of the microphysics scheme has dominant impact on cloud cover of cirrus clouds leading to smaller cloud cover for high opaque and very high opaque clouds and larger cloud cover for semi. thin and semi. moderately thick clouds. No pronounced net warming effect is found for simulated semi-transparent clouds. The net CRE bias of semi-transparent clouds is negative and caused by a misrepresentation of cirrus radiative properties, especially in the longwave.

In summary, our analysis shows that refining horizontal resolution allows the ICON model to more accurately represent cloud-radiative effects over the North Atlantic. We found substantial bias compensation between top-of-atmosphere shortwave and longwave radiation fluxes as well as between clear-sky fluxes and cloud-radiative effects. An acceptable net performance of a selected model setup is not at all a guarantor of realistic individual contributions. The best representation of the domain-average longwave and shortwave CREs is achieved when ICON is configured with two-moment cloud microphysics, a shallow-convection scheme (explicit treatment of mid-level and deep convection) and a horizontal resolution of 2.5 km.

Starting with climate model resolution of 80 km, the improvement from increasing resolution are gradually up to a resolution of 10 to 20 km, at which point a further increase in resolution only leads small and insufficient improvements of the simulated shortwave CREs. Instead, at finer resolutions, the saturation is overcome when the convection scheme is disabled so that the model is allowed to represent convection in an explicit manner. If convection treated explicitly, the simulation of CREs is even improved by refinements at (and possibly beyond) the kilometer scale. However, a resolution of 2.5 km is still too coarse to resolve the shallow clouds and circulation in the marine boundary layer, because of which the best simulation of average CREs at 2.5 km is achieved with an explicit treatment of mid-level and deep convection but a parameterized treatment of shallow convection. This simulation setup can represent the radiative-properties term in the CRE decomposition in a satisfactory manner for all cloud types except for very low clouds. For this cloud type, improvements in the simulation of cloud-optical thickness and thus vertical structure is needed. Moreover, the 2.5-km setup with parameterized shallow convection shows some deficits with regard to the fractional coverage of cloud types "very high opaque" and "low" which could be an indication that the linking between resolved and parameterized convection has weaknesses in this setup. Compared to fully explicit convection, the use of a shallow-convection scheme mitigates the otherwise too high fractional coverage of very low clouds and too strong cloud shortwave reflection, and at the same time does not affect longwave CRE, which are dominated by high-level clouds. A deeper understanding of the spatial distribution of the CRE biases is needed. A promising approach would be the analysis of the cloud distribution and its radiative effects as a function of meteorological conditions, e.g. cloud controlling factors depending on large-scale circulation and vertical velocity regimes.

Acknowledgments

FS and AV are supported by the German Ministry of Education and Research (BMBF) and FONA: Research for Sustainable Development (www.fona.de). This work contributes to the WCRP's Grand Challenge on Clouds, Circulation, and Climate Sensitivity and the BMBF-funded project HD(CP)^2 : High Definition Clouds and Precipitation for Advancing Climate Prediction. FS acknowlegdes funding under respective grants 01LK1507C and 01LK1503F, AV is supported under Grant Agreement 01LK1509A. The ICON simulations were performed by AV at the DKRZ in Hamburg, Germany, which is thanked for its support. We also thank EUMETSAT for producing the SEVIRI data, which have been obtained from the TROPOS satellite data archive.

Concerning data availability: The GERB-like data is made freely available to the user community via the RMIB OnLine Shortterm Service (ROLSS, see ftp://gerb.oma.be) server, after registration. The primary data of the ICON simulations (run scripts, namelists, scripts for lateral boundary data) will be published at KITopen of Karlsruhe Institute of Technology. Analysis data have been collected at the long-term archive (LTA) of DKRZ and can be assessed under `https://cera-www.dkrz.de/WDCC/ui/cerasearch/entry?acronym=DKRZ_LTA_834_ds00048`.

Open science: The analysis source code has been made freely available to improve reproducibility of our results. Basic analysis tools are written in Python and published at http://doi.org/10.5281/zenodo.3657387. The final plots for our paper were done with Jupyter Notebooks which are hosted at https://github.com/fsenf/nbook.CRE-2020-paper-plots.

References

Albern, N., Voigt, A., & Pinto, J. G. (2019). Cloud-radiative impact on the regional responses of the midlatitude jet streams and storm tracks to global warming. *J. Adv. Model. Earth Syst.*, *11*(7), 1940-1958.

Allan, R. P. (2011). Combining satellite data and models to estimate cloud radiative effect at the surface and in the atmosphere. *Meteorological Applications*, *18*(3), 324-333.

Arking, A. (1991). The radiative effects of clouds and their impact on climate. *Bull. Amer. Meteor. Soc.*, *72*(6), 795-814.

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., & Reinhardt, T. (2011). Operational convective-scale numerical weather prediction

with the cosmo model: Description and sensitivities. *Mon. Wea. Rev.*, *139*(12), 3887-3905.
Bechtold, P., Köhler, M., Jung, T., Doblas-Reyes, F., Leutbecher, M., Rodwell, M. J., ... Balsamo, G. (2008). Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *134*(634), 1337–1351.
Bodas-Salcedo, A., Williams, K. D., Ringer, M. A., Beau, I., Cole, J. N. S., Dufresne, J. L., ... Yokohata, T. (2014, Jan). Origins of the Solar Radiation Biases over the Southern Ocean in CFMIP2 Models*. *J. Climate*, *27*(1), 41-56. doi: 10.1175/JCLI-D-13-00169.1
Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., ... et al. (2011). COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc.*, *92*(8), 1023-1043.
Böhme, T., Stapelberg, S., Akkermans, T., Crewell, S., Fischer, J., Reinhardt, T., ... van Lipzig, N. (2011, April). Long-term evaluation of COSMO forecasting using combined observational data of the GOP period. *Meteor. Z.*, *20*, 119-132. doi: 10.1127/0941-2948/2011/0225
Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... et al. (2013). Clouds and aerosols. In *Climate change 2013: The physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* (pp. 571–657). Cambridge University Press.
Ceppi, P., Brient, F., Zelinka, M. D., & Hartmann, D. L. (2017). Cloud feedback mechanisms and their representation in global climate models. *Wiley Interdiscip. Rev. Clim. Change*, *8*(4), e465.
Ceppi, P., & Hartmann, D. L. (2015). Connections between clouds, radiation, and midlatitude dynamics: a review. *Curr. Clim. Change Rep.*, *1*(2), 94-102.
Ceppi, P., & Hartmann, D. L. (2016). Clouds and the atmospheric circulation response to warming. *J. Climate*, *29*(2), 783-799.
Ceppi, P., & Shepherd, T. G. (2017, Nov). Contributions of Climate Feedbacks to Changes in Atmospheric Circulation. *J. Climate*, *30*(22), 9097-9118. doi: 10.1175/JCLI-D-17-0189.1
Chaboureau, J.-P., Cammas, J.-P., Mascart, P., Pinty, J.-P., Claud, C., Roca, R., & Morcrette, J.-J. (2000). Evaluation of a cloud system life-cycle simulated by the Meso-NH model during FASTEX using METEOSAT radiances and TOVS-3I cloud retrievals. *Quart. J. Roy. Meteor. Soc.*, *126*(566), 1735-1750.
Chen, T., Rossow, W. B., & Zhang, Y. (2000). Radiative effects of cloud-type variations. *J. Climate*, *13*(1), 264-286.
Clerbaux, N., Bertrand, C., Caprion, D., Depaepe, B., Dewitte, S., Gonzalez, L., & Ipe, A. (2005). Narrowband-to-broadband conversions for seviri. In *Proc. of the 2005 eumetsat meteorological satellite conference* (p. 351-357).
Collins, M., Minobe, S., Barreiro, M., Bordoni, S., Kaspi, Y., Kuwano-Yoshida, A., ... others (2018). Challenges and opportunities for improved understanding of regional climate dynamics. *Nature Climate Change*, *8*(2), 101.
Derrien, M., & Le Gléau, H. (2005). MSG/SEVIRI cloud mask and type from SAFNWC. *Int. J. Remote Sens.*, *26*, 4707-4732. doi: 10.1080/01431160500166128
Dewitte, S., Gonzalez, L., Clerbaux, N., Ipe, A., Bertrand, C., & Paepe, B. D. (2008). The geostationary earth radiation budget edition 1 data processing algorithms. *Adv. Space Res.*, *41*(11), 1906 - 1913.
Dolinar, E. K., Dong, X., Xi, B., Jiang, J. H., & Su, H. (2015). Evaluation of cmip5 simulated clouds and toa radiation budgets using nasa satellite observations. *Climate Dyn.*, *44*(7-8), 2229–2247.
Eikenberg, S., Khler, C., Seifert, A., & Crewell, S. (2015, Apr). How microphysical

choices affect simulated infrared brightness temperatures. *Atmos. Res.*, *156*, 6779. Retrieved from `http://dx.doi.org/10.1016/j.atmosres.2014.12.010` doi: 10.1016/j.atmosres.2014.12.010

Evans, S., Marchand, R., Ackerman, T., Donner, L., Golaz, J.-C., & Seman, C. (2017). Diagnosing cloud biases in the gfdl am3 model with atmospheric classification. *J. Geophys. Res. Atmos.*, *122*(23), 12,827-12,844.

Fu, Q. (1996). An Accurate Parameterization of the Solar Radiative Properties of Cirrus Clouds for Climate Models. *J. Climate*, *9*, 2058-2082.

Futyan, J. M., & Russell, J. E. (2005). Developing clear-sky flux products for the geostationary earth radiation budget experiment. *J. Appl. Meteor.*, *44*(9), 1361-1374.

Gettelman, A., & Sherwood, S. C. (2016). Processes responsible for cloud feedback. *Curr. Clim. Change Rep.*, *2*(4), 179-189.

Grise, K. M., Medeiros, B., Benedict, J. J., & Olson, J. G. (2019). Investigating the influence of cloud radiative effects on the extratropical storm tracks. *Geophys. Res. Lett.*, *46*(13), 7700-7707.

Grise, K. M., & Polvani, L. M. (2014, Aug). Southern Hemisphere Cloud-Dynamics Biases in CMIP5 Models and Their Implications for Climate Projections. *J. Climate*, *27*(15), 6074-6092. doi: 10.1175/JCLI-D-14-00113.1

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., ... von Storch, J.-S. (2016). High resolution model intercomparison project (highresmip v1.0) for cmip6. *Geosci. Model Dev.*, *9*(11), 4185–4208.

Harries, J. E., Russell, J. E., Hanafin, J. A., Brindley, H., Futyan, J., Rufus, J., ... Ringer, M. A. (2005). The geostationary earth radiation budget project. *Bull. Amer. Meteor. Soc.*, *86*(7), 945-960.

Hartmann, D. L., Ockert-Bell, M. E., & Michelsen, M. L. (1992). The effect of cloud type on earth's energy balance: Global analysis. *J. Climate*, *5*(11), 1281-1304.

Heinze, R., Dipankar, A., Carbajal Henken, C., Moseley, C., Sourdeval, O., Trömel, S., ... Quaas, J. (2017). Large-eddy simulations over germany using ICON: a comprehensive evaluation. *Quart. J. Roy. Meteor. Soc.*, *143*(702), 69–100.

Henderson, D. S., L'Ecuyer, T., Stephens, G., Partain, P., & Sekiguchi, M. (2013). A multisensor perspective on the radiative impacts of clouds and aerosols. *J. Appl. Meteor. Climatol.*, *52*(4), 853–871.

Hogan, R. J., & Illingworth, A. J. (2000). Deriving cloud overlap statistics from radar. *Quart. J. Roy. Meteor. Soc.*, *126*(569), 2903-2909.

Hohenegger, C., Kornblueh, L., Klocke, D., Becker, T., Cioni, G., Engels, J. F., ... Stevens, B. (2020). Climate statistics in global simulations of the atmosphere, from 80 to 2.5 km grid spacing. *J. Meteor. Soc. Japan*, *98*(1), 73-91. doi: 10.2151/jmsj.2020-005

Hollingsworth, A., Engelen, R. J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., ... Consortium, T. G. (2008, 08). Toward a Monitoring and Forecasting System for Atmospheric Composition: The GEMS Project. *Bull. Amer. Meteor. Soc.*, *89*(8), 1147-1164. Retrieved from `https://doi.org/10.1175/2008BAMS2355.1` doi: 10.1175/2008BAMS2355.1

Keil, C., Tafferner, A., & Reinhardt, T. (2006). Synthetic satellite imagery in the Lokal-Modell. *Atmos. Res.*, *82*, 19-25.

Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., & Gleckler, P. J. (2013). Are climate model simulations of clouds improving? an evaluation using the isccp simulator. *J. Geophys. Res. Atmos.*, *118*(3), 1329-1342. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/jgrd.50141` doi: 10.1002/jgrd.50141

Klocke, D., Brueck, M., Hohenegger, C., & Stevens, B. (2017). Rediscovery of the doldrums in storm-resolving simulations over the tropical Atlantic. *Nature Geoscience*, *10*(12), 891-896.

Köhler, C. G., & Seifert, A. (2015, Feb). Identifying sensitivities for cirrus modelling using a two-moment two-mode bulk microphysics scheme. *Tellus B*, *67*(0). Retrieved from http://dx.doi.org/10.3402/tellusb.v67.24494 doi: 10.3402/tellusb.v67.24494

L'Ecuyer, T. S., Hang, Y., Matus, A. V., & Wang, Z. (2019). Reassessing the effect of cloud type on earth's energy balance in the age of active spaceborne observations. part i: Top of atmosphere and surface. *J. Climate*, *32*(19), 6197-6217.

Li, Y., Thompson, D. W. J., Bony, S., & Merlis, T. M. (2019, Feb). Thermodynamic Control on the Poleward Shift of the Extratropical Jet in Climate Change Simulations: The Role of Rising High Clouds and Their Radiative Effects. *J. Climate*, *32*(3), 917-934. doi: 10.1175/JCLI-D-18-0417.1

Maher, P., Vallis, G. K., Sherwood, S. C., Webb, M. J., & Sansom, P. G. (2018). The impact of parameterized convection on climatological precipitation in atmospheric global climate models. *Geophys. Res. Lett.*, *45*(8), 3728-3736.

Matricardi, M., Chevallier, F., Kelly, G., & Thépaut, J.-N. (2004). An improved general fast radiative transfer model for the assimilation of radiance observations. *Quart. J. Roy. Meteor. Soc.*, *130*, 153-173.

Matsui, T., Dolan, B., Rutledge, S. A., Tao, W.-K., Iguchi, T., Barnum, J., & Lang, S. E. (2019). Polarris: A polarimetric radar retrieval and instrument simulator. *J. Geophys. Res. Atmos.*, *124*(8), 4634-4657. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD028317 doi: 10.1029/2018JD028317

McDonald, A. J., & Parsons, S. (2018). A comparison of cloud classification methodologies: Differences between cloud and dynamical regimes. *J. Geophys. Res. Atmos.*, *123*(19), 11,173-11,193.

McFarquhar, G. M., Iacobellis, S., & Somerville, R. C. J. (2003). SCM Simulations of Tropical Ice Clouds Using Observationally Based Parameterizations of Microphysics. *J. Climate*, *16*(11), 1643-1664.

Meirink, J. F., Roebeling, R. A., & Stammes, P. (2013). Inter-calibration of polar imager solar channels using seviri. *Atmos. Meas. Tech.*, *6*(9), 2495–2508.

Mekaoui, S., & Dewitte, S. (2008). Total Solar Irradiance Measurement and Modelling during Cycle 23. *Sol. Phys.*, *247*(1), 203-216.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., & Clough, S. A. (1997). Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res. Atmos.*, *102*(D14), 16663-16682.

Morcrette, J.-J. (1991). Evaluation of Model-generated Cloudiness: Satellite-observed and Model-generated Diurnal Variability of Brightness Temperature. *Mon. Wea. Rev.*, *119*(5), 1205-1224.

Nam, C., Bony, S., Dufresne, J.-L., & Chepfer, H. (2012). The "too few, too bright" tropical low-cloud problem in cmip5 models. *Geophys. Res. Lett.*, *39*(21).

Ockert-Bell, M. E., & Hartmann, D. L. (1992). The effect of cloud type on earth's energy balance: Results for selected regions. *J. Climate*, *5*(10), 1157-1171.

Oreopoulos, L., Cho, N., Lee, D., & Kato, S. (2016). Radiative effects of global modis cloud regimes. *J. Geophys. Res. Atmos.*, *121*(5), 2299-2317.

Oreopoulos, L., & Rossow, W. B. (2011). The cloud radiative effects of international satellite cloud climatology project weather states. *J. Geophys. Res. Atmos.*, *116*(D12).

Pincus, R., Platnick, S., Ackerman, S. A., Hemler, R. S., & Patrick Hofmann, R. J. (2012). Reconciling simulated and observed views of clouds: Modis, isccp, and the limits of instrument simulators. *J. Climate*, *25*(13), 4699-4720.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., ... Leung, R. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, *53*(2), 323-361.

Pscheidt, I., Senf, F., Heinze, R., Deneke, H., Trömel, S., & Hohenegger, C. (2019). How organized is deep convection over germany? *Quart. J. Roy. Meteor. Soc.*,

145(723), 2366-2384.
Ramanathan, V., Cess, R., Harrison, E., Minnis, P., Barkstrom, B., Ahmad, E., & Hartmann, D. (1989). Cloud-radiative forcing and climate: Results from the earth radiation budget experiment. *Science*, *243*(4887), 57–63.
Randall, D., Khairoutdinov, M., Arakawa, A., & Grabowski, W. (2003). Breaking the cloud parameterization deadlock. *Bull. Amer. Meteor. Soc.*, *84*(11), 1547-1564.
Ritter, B., & Geleyn, J.-F. (1992). A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. *Mon. Wea. Rev.*, *120*(2), 303-325.
Roberts, M. J., Vidale, P. L., Senior, C., Hewitt, H. T., Bates, C., Berthou, S., ... Wehner, M. F. (2018). The benefits of global high resolution for climate simulation: Process understanding and the enabling of stakeholder decisions at the regional scale. *Bull. Amer. Meteor. Soc.*, *99*(11), 2341-2359.
Roca, R., Picon, L., Desbois, M., Le Treut, H., & Morcrette, J.-J. (1997). Direct comparison of meteosat water vapor channel data and general circulation model results. *Geophys. Res. Lett.*, *24*(2), 147-150.
Rossow, W. B., & Schiffer, R. A. (1999). Advances in understanding clouds from isccp. *Bull. Amer. Meteor. Soc.*, *80*(11), 2261-2287.
Sakradzija, M., & Klocke, D. (2018). Physically constrained stochastic shallow convection in realistic kilometer-scale simulations. *J. Adv. Model. Earth Syst.*, *10*(11), 2755-2776. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001358` doi: 10.1029/2018MS001358
Satoh, M., Noda, A. T., Seiki, T., Chen, Y.-W., Kodama, C., Yamada, Y., ... Sato, Y. (2018). Toward reduction of the uncertainties in climate sensitivity due to cloud processes using a global non-hydrostatic atmospheric model. *Prog. Earth Planet. Sci.*, *5*(1), 67.
Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Düben, P. (2019, 17). Global cloud-resolving models. *Curr. Clim. Change Rep.*.
Saunders, R., Matricardi, M., & Brunel, P. (1999). An improved for assimilation of satellite radiance observations. *Quart. J. Roy. Meteor. Soc.*, *125*(556), 1407-1425.
Schäfer, S. A. K., & Voigt, A. (2018, Mar). Radiation Weakens Idealized Midlatitude Cyclones. *Geophys. Res. Lett.*, *45*(6), 2833-2841. doi: 10.1002/2017GL076726
Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J. D., McTaggart-Cowan, R., ... Zinner, T. (2018). The north atlantic waveguide and downstream impact experiment. *Bull. Amer. Meteor. Soc.*, *99*(8), 1607-1637.
Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., & Ratier, A. (2002). An introduction to Meteosat Second Generation (MSG). *Bull. Amer. Meteor. Soc.*, *83*(7), 977-992.
Seifert, A., & Beheng, K. D. (2006, 01). A two-moment cloud microphysics parameterization for mixed-phase clouds. part 1: Model description. *Meteor. Atmos. Phys.*, *92*(1), 45–66.
Senf, F., & Deneke, H. (2017). Uncertainties in synthetic meteosat seviri infrared brightness temperatures in the presence of cirrus clouds and implications for evaluation of cloud microphysics. *Atmos. Res.*, *183*, 113-129.
Senf, F., Klocke, D., & Brueck, M. (2018). Size-resolved evaluation of simulated deep tropical convection. *Mon. Wea. Rev.*, *146*(7), 2161-2182.
Sohn, B. J., Nakajima, T., Satoh, M., & Jang, H. S. (2010, December). Impact of different definitions of clear-sky flux on the determination of longwave cloud radiative forcing: NICAM simulation results. *Atmos. Chem. Phys.*, *10*(23), 11641-11646. doi: 10.5194/acp-10-11641-2010
Stephens, G., Winker, D., Pelon, J., Trepte, C., Vane, D., Yuhas, C., ... Lebsock,

M. (2018). Cloudsat and calipso within the a-train: Ten years of actively observing the earth system. *Bull. Amer. Meteor. Soc.*, *99*(3), 569-581.

Stephens, G. L. (2005). Cloud feedbacks in the climate system: A critical review. *J. Climate*, *18*(2), 237-273.

Stephens, G. L., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., … Andrews, T. (2012). An update on earth s energy balance in light of the latest global observations. *Nat. Geosci.*, *5*(10), 691.

Stevens, B., Acquistapace, C., Hansen, A., , & Coauthors incl. Senf, F. (2020). Large-eddy and storm resolving models for climate prediction the added value for clouds and precipitation. *J. Meteor. Soc. Japan*. doi: 10.2151/jmsj.2020-021

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., … Roeckner, E. (2013). Atmospheric component of the mpi-m earth system model: Echam6. *J. Adv. Model. Earth Syst.*, *5*(2), 146-172.

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., … Klocke, D. (2019). Dyamond: the dynamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Prog. Earth Planet Sci.*, *6*(1), 61.

Tegen, I., Hollrig, P., Chin, M., Fung, I., Jacob, D., & Penner, J. (1997). Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. *J. Geophys. Res. Atmos.*, *102*(D20), 23895-23915.

Thomas, M. A., Devasthale, A., Koenigk, T., Wyser, K., Roberts, M., Roberts, C., & Lohmann, K. (2018). A statistical and process oriented evaluation of cloud radiative effects in high resolution global models. *Geosci. Model Dev. Discuss.*, *2018*, 1–30. doi: 10.5194/gmd-2018-221

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, *117*(8), 1779–1800.

Vannière, B., Demory, M.-E., Vidale, P. L., Schiemann, R., Roberts, M. J., Roberts, C. D., … Senan, R. (2019). Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution. *Climate Dyn.*, *52*(11), 6817–6846.

Voigt, A., Albern, N., & Papavasileiou, G. (2019, May). The Atmospheric Pathway of the Cloud-Radiative Impact on the Circulation Response to Global Warming: Important and Uncertain. *J. Climate*, *32*(10), 3051-3067. doi: 10.1175/JCLI-D-18-0810.1

Voigt, A., & Shaw, T. A. (2015, Feb). Circulation response to warming shaped by radiative changes of clouds and water vapour. *Nat. Geosci.*, *8*(2), 102-106. doi: 10.1038/ngeo2345

Voigt, A., & Shaw, T. A. (2016, Dec). Impact of Regional Atmospheric Cloud Radiative Changes on Shifts of the Extratropical Jet Stream in Response to Global Warming. *J. Climate*, *29*(23), 8399-8421. doi: 10.1175/JCLI-D-16-0140.1

Webb, M. J., Lock, A. P., Bretherton, C. S., Bony, S., Cole, J. N. S., Idelkadi, A., … Zhao, M. (2015). The impact of parametrized convection on cloud feedback. *Phil. Trans. R. Soc. A*, *373*(2054), 20140414.

Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2014). The ICon (ICosahedral non-hydrostatic) modelling framework of dwd and MPI-m: Description of the non-hydrostatic dynamical core. *Q.J.R. Meteorol. Soc*, *141*(687), 563-579.

Zelinka, M. D., Randall, D. A., Webb, M. J., & Klein, S. A. (2017). Clearing clouds of uncertainty. *Nature Climate Change*, *7*(10), 674.

Supporting Information for "Increasing Resolution and Resolving Convection Improves the Simulation of Cloud-Radiative Effects over the North Atlantic"

Fabian Senf[1], Aiko Voigt[2,3], Nicolas Clerbaux[4], Anja Hünerbein[1], Hartwig Deneke[1]

[1]Leibniz Institute for Tropospheric Research, Leipzig
[2]Institute for Meteorology and Climate Research - Department Troposphere Research, Karlsruhe Institute of Technology, Karlsruhe
[3]Lamont-Doherty Earth Observatory, Columbia University, New York, USA
[4]Royal Meteorological Institute of Belgium, Brussels

1 Cloud Classification with NWCSAF v2013

1.1 Adjustments for Permanent Night Mode

For cloud classification, we apply the NWCSAF software v2013 (Derrien & Le Gléau, 2005). We keep the software itself unmodified and implement all changes via an interface that controls the input files and the execution of the software package. Within that interface, observed and synthetic infrared BTs are read from disk space and written into Meteosat SEVIRI HRIT template files (the native data format distributed by EUMETSAT). The template files themselves are valid for 0 UTC, but data embedded into the template files can have any time stamp. The NWCSAF software retrieves night-time cloud classifications independent of the actual time stamp of the embedded input data.

The Metetosat SEVIRI imager measures radiances, expressed in terms of brightness temperatures (BTs), in several channels. One of them, the 3.9 μm channel, is affected by shortwave as well as longwave radiation (see e.g. Lindsey et al., 2006). During night-time, the use of the 3.9-μm brightness temperature is beneficial for detecting clouds and their microphysical characteristics at their top (Lensky & Rosenfeld, 2003). Therefore, the 3.9 μm channel is mandatory for the NWCSAF cloud classification at night-time. Because we aim to feed the NWCSAF software with both night-time and day-time scenes, 3.9-μm radiances can be contaminated by sunlight, which might lead to erroneous cloud classifications by the NWCSAF software.

We mitigate this problem in the simplest possible way: we estimate 3.9-μm BT from BTs of the other infrared channels by means of a linear regression derived from a least-squares fit,

$$T_{3.9} = T_{10.8} + a_0 + a_1 \Delta T_{8.7-10.8} + a_2 \Delta T_{10.8-12.0} + a_3 \Delta T_{13.4-10.8} \,. \tag{1}$$

T_i is the infrared BT of a SEVIRI channel with central wave length i μm, and ΔT_{i-j} is the BT difference of two SEVIRI channels centered at i and j μm. The regression was applied to observed SEVIRI data over the North Atlantic analysis domain and each 0 UTC time slot within the whole NAWDEX period. The resulting average regression parameters of $(a_0,\ a_1,\ a_2,\ a_3) = (3\text{ K},\ 1.8,\ 1.5,\ 0.12)$ are then used to estimate the 3.9-μm BT, which is then fed into the NWCSAF software instead of observed or simulated values. The approximation of the 3.9 μm channel is done for both the SEVIRI observations and the synthetic observations derived from the ICON simulations with the SynSat forward operator.

Corresponding author: Fabian Senf, `senf@tropos.de`

The linear regression gives acceptable results, as described in more detail in the following section. Testing against observed 3.9-μm BTs at 0 UTC, explained variances are always above 99%, average biases are $\sim$ 0.1 K and average RMSEs are below 2 K. We note that the current method is limited to ocean regions. For surfaces with a more heterogeneous surface emissivity, e.g., in the Saharan region, a more sophisticated approach would be needed.

1.2 Evaluation of NWCSAF Cloud Classification in Night Mode

As described above, we estimate 3.9-μm BTs from other channels' BTs. The extent to which this degrades the NWCSAF cloud classification is analysed below for 0 UTC at night time (see Tables S1 and S2). We use a pixel-based comparison and a binary verification concept in which a certain cloud type is considered to be present or not. Using the 2×2 contingency tables (see Wilks, 2006, p.260), five verification metrics are considered: proportion correct (PC, Wilks eq. 7.7), critical success index (CSI, Wilks eq. 7.8), BIAS (Wilks eq. 7.9), probability of detection (POD, Wilks eq. 7.12) and false alarm rate (FAR, Wilks eq. 7.13). The verification performs best if PC, CSI, BIAS and POD have values of 1m and FAR is zero. For each verifcation metric, the largest absolute deviation from these optimal values is marked in bold. In general, the performance of NWCSAF run in night-mode is very good. The degradation is strongest for fractional, very low and semi-transparent thin clouds, but even for these the performance is satisfactory. We conclude that the modified NWCSAF software will provide a robust cloud classification that can be used to assess differences between observations and simulations.

A comparison between cloud classification by our degraded NWCSAF night-time approach mode and the standard day-time NWCSAF approach is shown in Tab. S3 and S4. For the latter also solar SEVIRI channels have been used. The comparison thus shows the overall information loss when only thermal infrared BTs can be used. The verification scores are substantially worse than above. This means (i) there is a significant change in detection quality during the course of the day (which we tried to minimize with our permanent night-mode setup), and (ii) the solar channels help a lot during daytime. Again, the most affected cloud type is "fractional", followed by "very low" and "semi-transparent". The "semi. above" class is not assigned during night-time.

Table S1. Evaluation table for NWCSAF cloud masking at 0 UTC. The evaluation is based on 25 days from 20 September to 14 October, 2016.

Cloud Mask Type	f_{degraded} (%)	f_{standard} (%)	PC	CSI	BIAS	POD	FAR
clear	33.4	32.6	**0.967**	**0.905**	1.024	0.961	**0.030**
partially cloudy	18.7	19.5	0.989	0.945	**0.960**	**0.952**	0.002
cloudy	47.9	47.9	0.977	0.953	1.000	0.976	0.022

Table S2. Evaluation table for NWCSAF cloud masking at 0 UTC. The evaluation is based on 25 days from 20 September to 14 October, 2016.

Cloud Type	f_{degraded} (%)	f_{standard} (%)	PC	CSI	BIAS	POD	FAR
very low	19.5	19.4	0.975	0.881	1.002	0.937	**0.016**
low	10.3	9.8	0.980	0.817	1.047	0.921	0.014
mid-level	4.0	4.2	0.995	0.881	0.958	0.917	0.002
high opaque	6.8	6.8	1.000	1.000	1.000	1.000	0.000
very high opaque	1.2	1.2	1.000	1.000	1.000	1.000	0.000
semi. thin	5.0	5.1	0.986	0.759	0.984	0.856	0.007
semi. moderately thick	9.2	8.9	0.991	0.907	1.035	0.967	0.007
semi. thick	2.0	2.0	1.000	1.000	1.000	1.000	0.000
semi. above	0.0	0.0	1.000	-	-	-	0.000
fractional	8.6	9.9	**0.961**	**0.654**	**0.866**	**0.738**	0.014

Table S3. Evaluation table for NWCSAF cloud masking at 12 UTC. The evaluation is based on 25 days from 20 September to 14 October, 2016.

Cloud Mask Type	f_{degraded} (%)	f_{standard} (%)	PC	CSI	BIAS	POD	FAR
clear	26.4	34.9	**0.892**	**0.701**	**0.758**	**0.724**	0.018
partially cloudy	21.7	20.9	0.967	0.857	1.040	0.941	0.026
cloudy	51.5	44.2	0.919	0.845	1.165	0.992	**0.138**

Table S4. Evaluation table for NWCSAF cloud typing at 12 UTC. The evaluation is based on 25 days from 20 September to 14 October, 2016.

Cloud Type	f_{degraded} (%)	f_{standard} (%)	PC	CSI	BIAS	POD	FAR
very low	17.6	16.6	0.919	0.617	1.063	0.787	0.055
low	8.6	12.6	0.959	0.678	0.685	0.681	0.001
mid-level	3.8	4.7	0.991	0.814	0.815	0.814	0.000
high opaque	6.6	6.7	0.999	0.986	0.987	0.986	0.000
very high opaque	1.4	1.4	1.000	0.995	0.999	0.997	0.000
semi. thin	5.4	2.7	0.960	0.342	1.955	0.753	0.034
semi. moderately thick	9.1	3.0	0.939	0.329	**3.038**	1.000	**0.063**
semi. thick	2.3	2.2	0.999	0.955	1.047	1.000	0.001
semi. above	0.0	4.3	0.957	-	-	-	-
fractional	10.3	19.0	**0.825**	**0.252**	0.540	**0.310**	0.054

2 On the Bias in Simulated Clear-Sky Radiation Fluxes

Here, we provide further and more detailed information on the bias correction of simulated clear-sky radiation fluxes. It has been discussed in the main part that it is challenging to derived accurate estimates for observed clear-sky fluxes, especially due to the high cloud coverage and the rather low cloud-free fraction found in our analysis domain. For that reason we decided to use simulated clear-sky radiation fluxes as substitute for observed clear-sky fluxes. However, systematic biases in simulated fluxes need to characterized and corrected.

In the following, we consider longwave fluxes and skip the subscript "lw", but the same also applies to upwelling shortwave fluxes with the subscript "sw, up". We assume that simulated radiation fluxes have a systematic bias B and a random error ε, i.e.

$$F_{\mathrm{ICON,clear}} = F_{\mathrm{OBS,clear}} + B + \varepsilon \tag{2}$$

After statistical averaging, the contribution of the random error ε gets smaller and less important. Thus, the bias B can be estimated from the average difference between simulated and observed clear-sky fluxes. The observed clear-sky flux $F_{\mathrm{OBS,clear}}$ is however unknown. Combining cloud detection (or detection of cloud-free regions) and observational flux estimates, all-sky fluxes in cloud-free regions $F_{\mathrm{OBS,0}}$ can derived. In addition to the clear-sky information these fluxes contain the radiative effect of undetected clouds, i.e.

$$F_{\mathrm{OBS,clear}} = F_{\mathrm{OBS,0}} + \Delta F_{\mathrm{OBS}} \tag{3}$$

The term ΔF_{OBS} characterizes our cloud detection capabilities. It is thus a characteristic property of the cloud classification algorithm. As we also derive a cloud classification based on simulations (in a very consistent way), we are able to estimate the average magnitude of the radiative effect of undetected clouds in ICON simulations as

$$F_{\mathrm{ICON,clear}} = F_{\mathrm{ICON,0}} + \Delta F_{\mathrm{ICON}}. \tag{4}$$

Both, $F_{\mathrm{ICON,clear}}$ and $F_{\mathrm{ICON,0}}$ are known and ΔF_{ICON} can be derived (see Fig. 5 in the main part). If we assume that the radiative effects of undetected clouds have similar magnitudes in simulations and observations, i.e. $\Delta F_{\mathrm{OBS}} \approx \Delta F_{\mathrm{ICON}}$, a bias correction

$$B = F_{\mathrm{ICON,clear}} - (F_{\mathrm{OBS,0}} + \Delta F_{\mathrm{SIM}}) \tag{5}$$
$$= F_{\mathrm{ICON,0}} - F_{\mathrm{OBS,0}} \tag{6}$$

can be derived. This means that if a bias correction is found that adjusts differences in observed and simulated all-sky fluxes in cloud-free regions, this is equivalent to a bias correction that adjusts the radiative effects of undetected clouds in simulations and observations. For upwelling shortwave clear-sky fluxes, we applied a scaling factor and for longwave clear-sky fluxes an offset is added.

3 Additional Data Overview

We provide three additional overview plots to supplement the figures shown in the main part of the manuscript. BTs from window channel at 10.8 μm are shown in Fig. S1, BTs from the water vapor channel at 6.2 μm are shown in Fig. S2 and the dependence of cloud typing on grid spacing is visualized in Fig. S3.

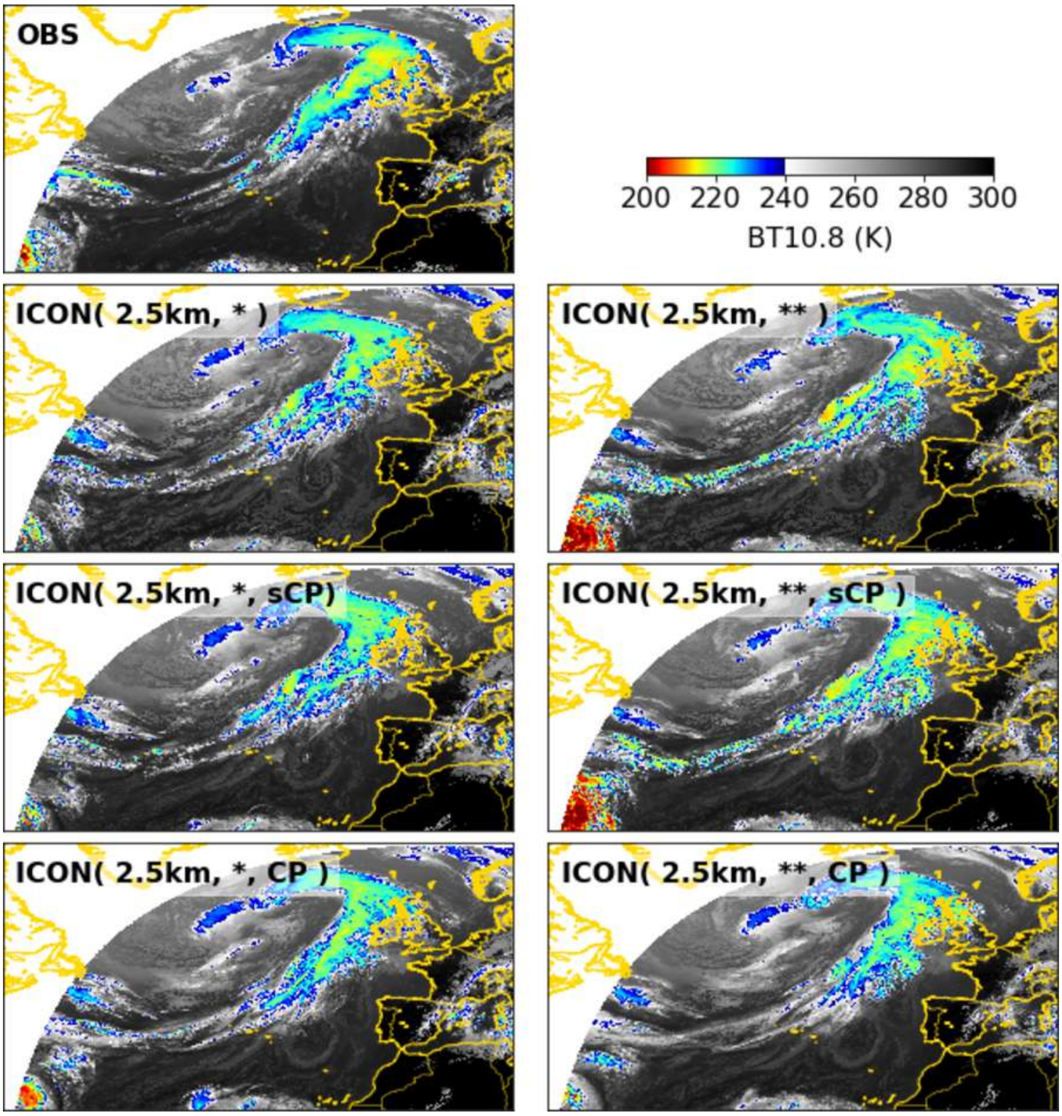

Figure S1. Overview of observed and simulated BTs from Meteosat SEVIRI's window channel at 10.8 μm for 1200 UTC 23 Sept 2016. Meteosat SEVIRI observations (top left) are compared ICON simulations with 2.5 km horizontal resolution. The left column is for simulations with one-moment cloud microphysics (*), the right column for simulations with two-moment microphysics (**). The second row is for fully explicit convection, the third row for simulations with a shallow convection scheme (sCP), and the fourth row for simulations with fully parameterized convection (CP). A special color scheme is used to highlight observed and simulated features. BTs over land are also shown to improve anticipation of the cloud scenery.

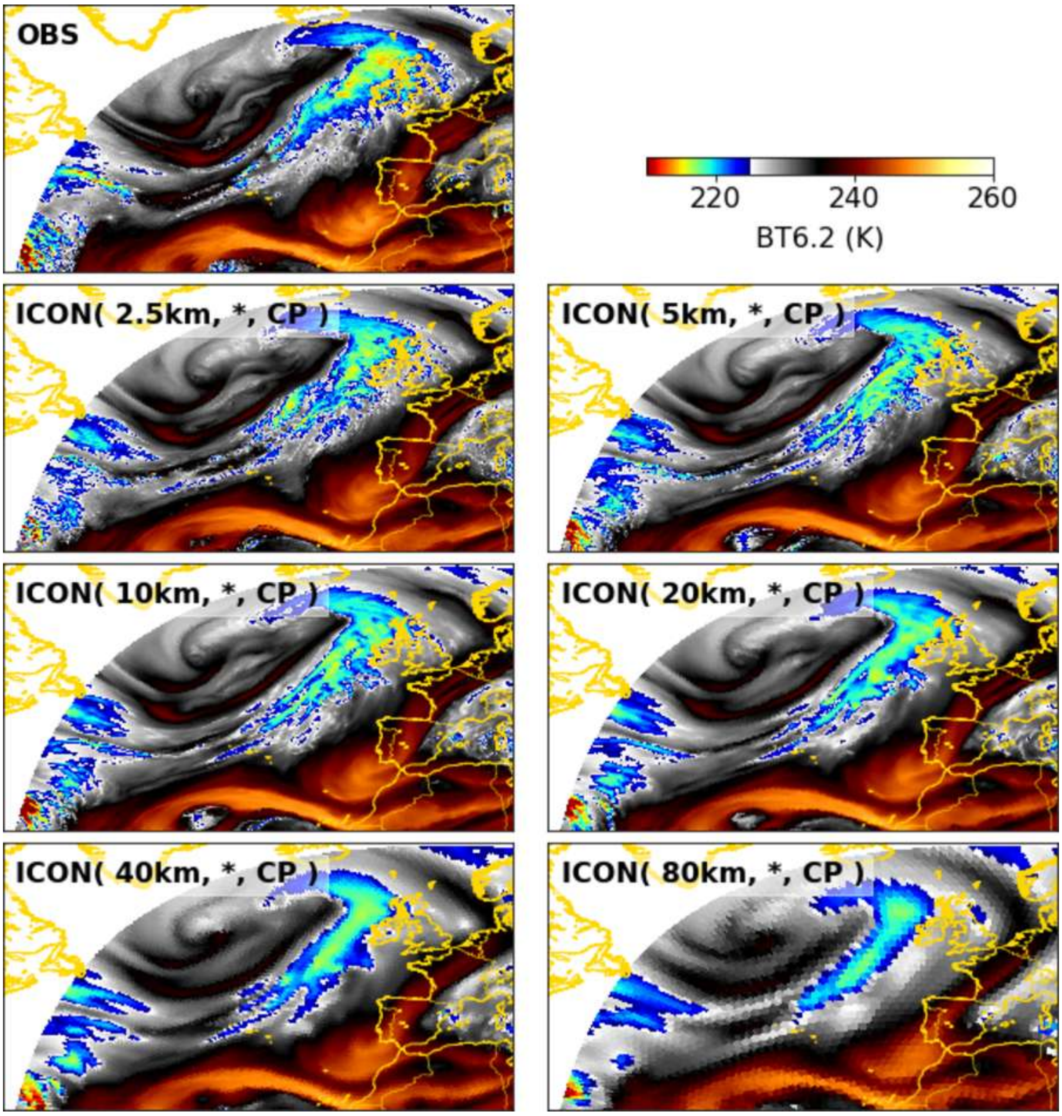

Figure S2. Overview of observed and simulated BTs from Meteosat SEVIRI's water vapor channel at 6.2 μm for 1200 UTC 23 Sept 2016. Observations are compared to ICON simulations with increasing grid spacing (left to right and downwards, from 2.5 to 80 km). Only the subset of simulations with one-moment microphysics and fully-parameterized convection is chosen for visualization. A special color scheme is used to highlight observed and simulated features. BTs over land are also shown to improve anticipation of the cloud scenery.

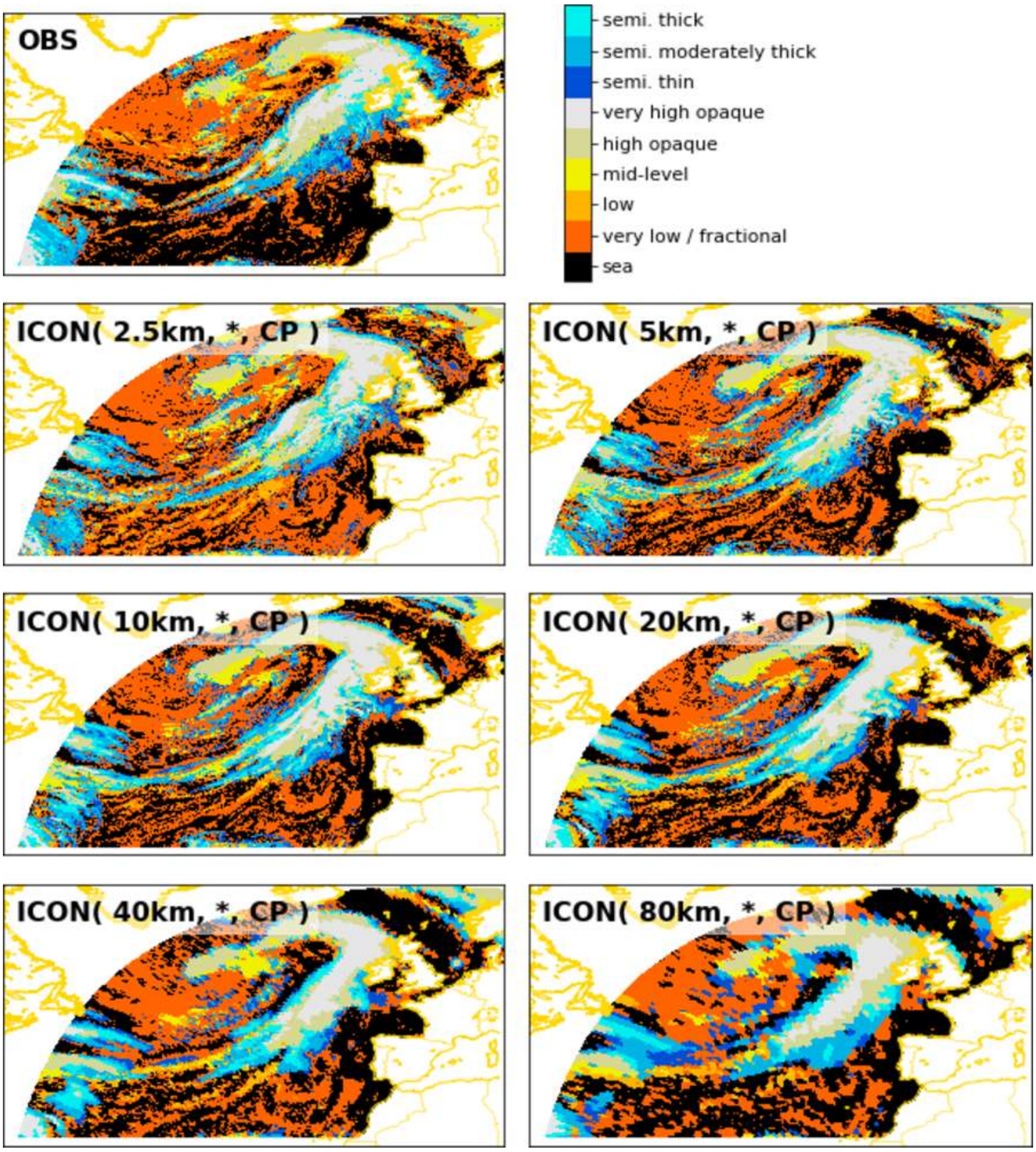

Figure S3. Overview of observed and simulated cloud types for 1200 UTC 23 Sept 2016. Cloud classification based on Meteosat SEVIRI (top row) is compared to cloud classification based on synthetic radiances derived from the ICON simulations and RTTO. The same simulations are shown as in Fig. S2.

3.1 Supplementary Data Analysis with Special Emphasis on Set 2

The following Figs. S4-S9 provides plots from additional analysis which support the arguments and conclusion made in the main part of the manuscript. Six additional numerical experiments have been performed in simulation set 2. These include ICON simulations with 2.5 km grid spacing and fully enabled convection parameterization and ICON simulations with 5 and 10 km grid spacing with only explicit convection (fully disabled convection scheme). Two main conclusion can be derived from the additional data analysis:

(i) For simulations with fully explicit convection, biases in radiation fluxes and cloud-radiative effects are reduced when the grid spacing is sequentially brought down from 10 to 2.5 km. The simulations at 2.5 km do not seem to have reached a stage where signatures of convergence can be identified. Further reduction in grid spacing is needed.

(ii) The simulation with parameterized convection and 2.5 km has similar error characteristics then its coarser counterparts. This means that difference in e.g. `ICON( 2.5km, *)` and `ICON( 5km, *, CP)` which are discussed in the main part of the manuscript are not due to difference in grid spacing.

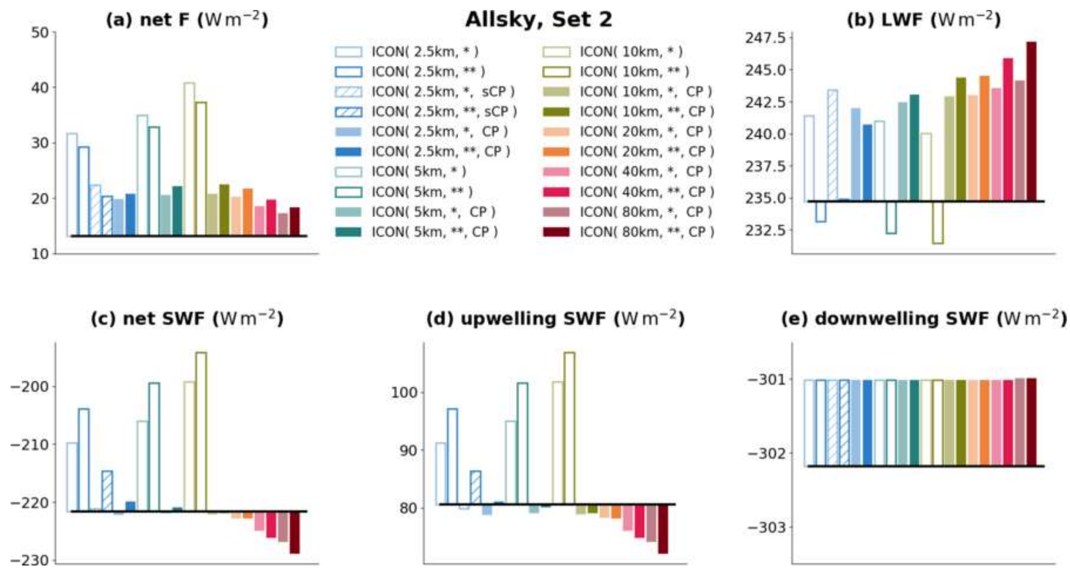

Figure S4. Analysis of domain-average allsky radiation fluxes: (a) total net flux, (b) emitted longwave flux, (c) net shortwave flux, (d) upwelling shortwave flux, and (e) downwelling shortwave flux. The Meteosat observations (black line) were chosen as reference, and deviation of simulated fluxes are shown with colored bars. The simulation experiments differ with regard to horizontal grid spacing (2.5 5, 10, 20, 40 and 80 km), and parameterization choice (one-moment vs. two-moment microphysics, with vs. without convection parameterization scheme). All values represent time averages over 3 days from simulation set 2.

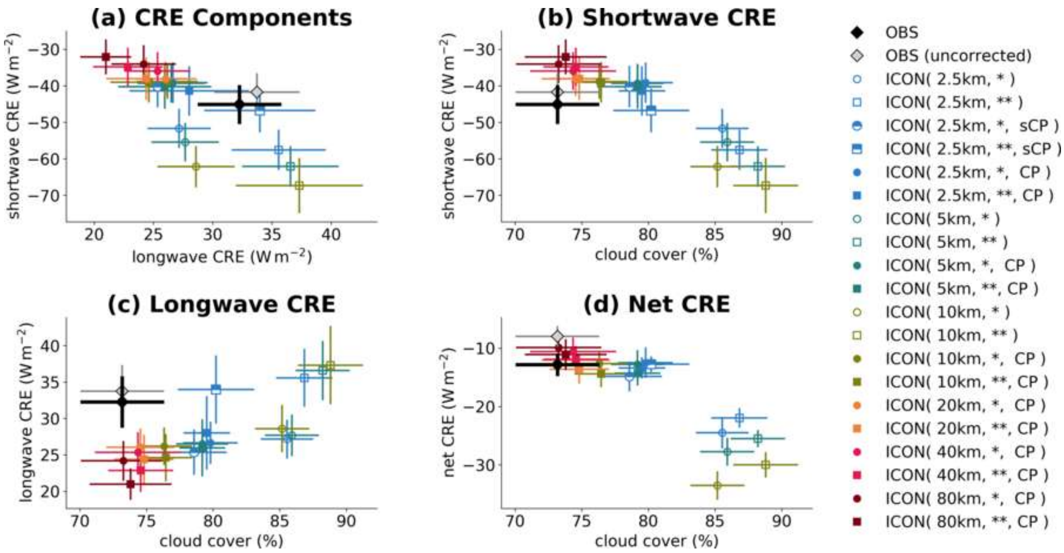

Figure S5. Comparison of domain-average allsky cloud-radiative effects and total cloud cover: (a) longwave CRE vs. shortwave CRE, and cloud cover vs. (b) shortwave CRE, (c) longwave CRE, and (d) net CRE. Symbols and error bars represent average and estimates of standard errors, respectively. With different colors and symbols styles different simulations experiments are distinguished. Please note the differences in the y-axis ranges. All values represent time averages over 3 days from simulation set 2.

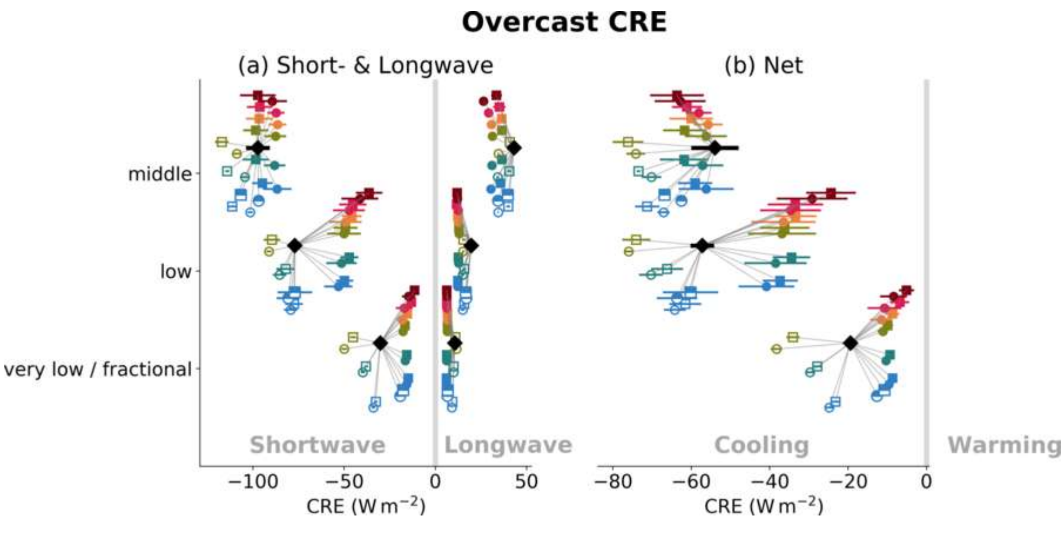

Figure S6. Overcast CRE for different shallow cloud types. A legend for color and symbols can be found in Fig. S5. All values represent time averages over 3 days from simulation set 2.

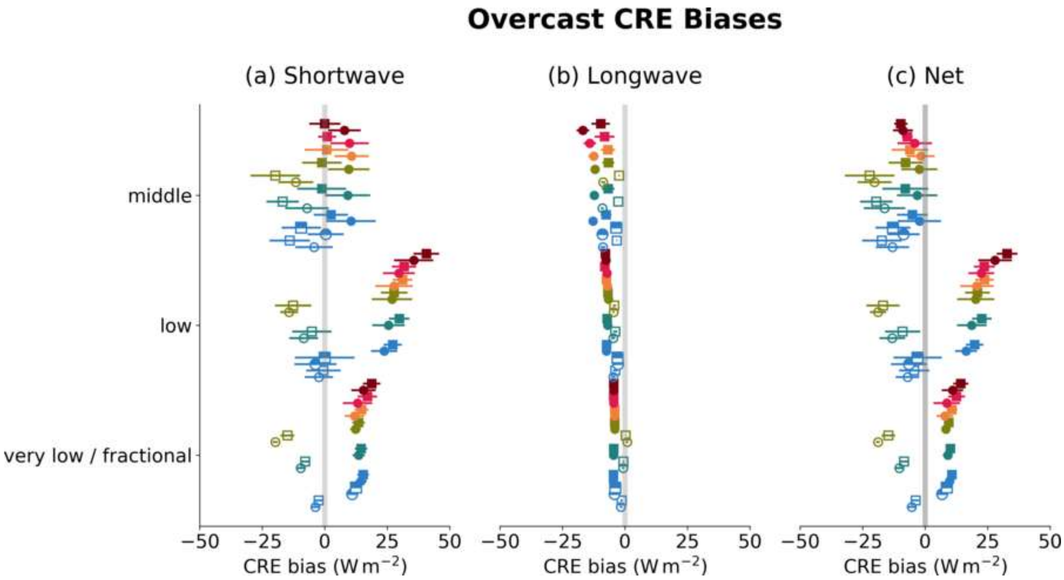

Figure S7. Biases in overcast CRE for different shallow cloud types. A legend for color and symbols can be found in Fig. S5. All values represent time averages over 3 days from simulation set 2.

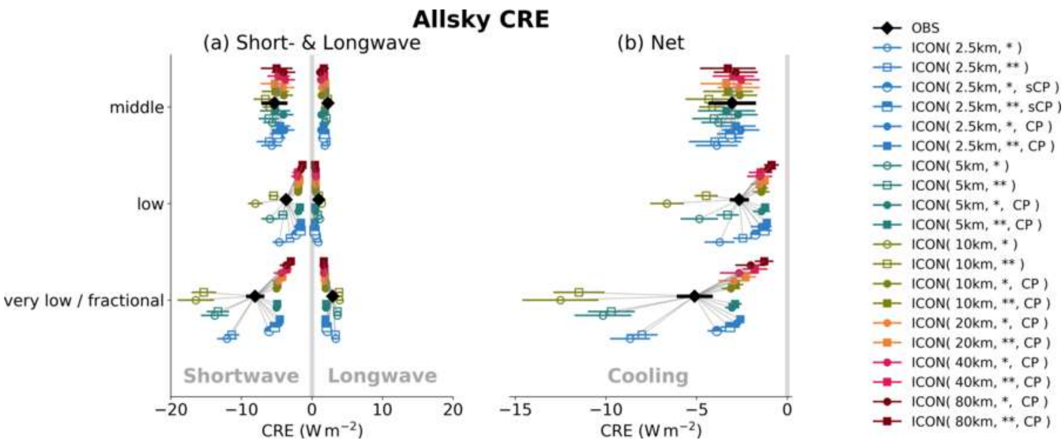

Figure S8. Allsky CRE for different shallow cloud types. A legend for color and symbols can be found in Fig. S5. All values represent time averages over 3 days from simulation set 2.

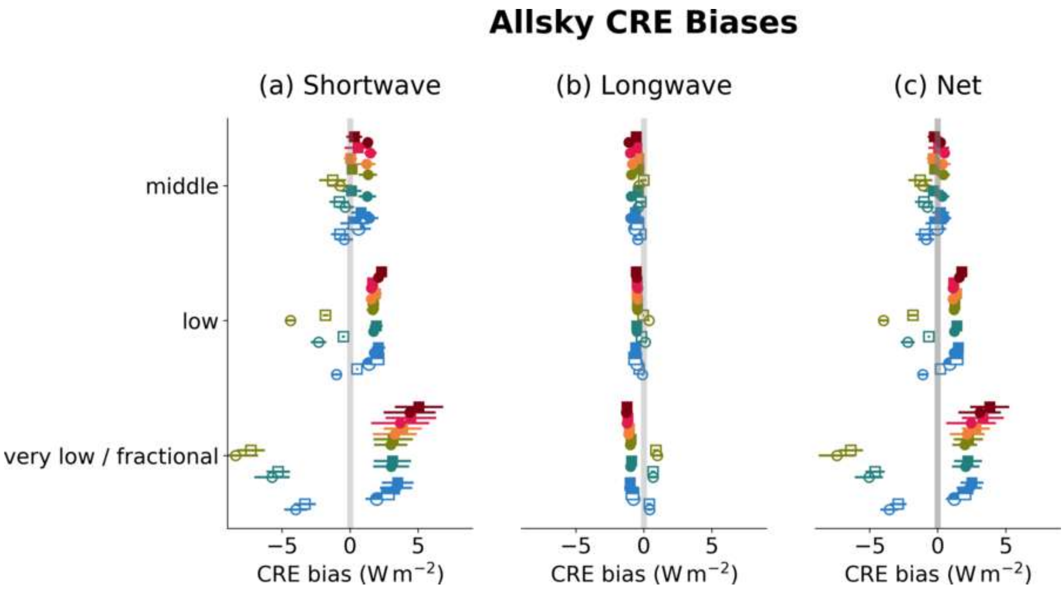

Figure S9. Biases in allsky CRE for different shallow cloud types. A legend for color and symbols can be found in Fig. S5. All values represent time averages over 3 days from simulation set 2.

References

Derrien, M., & Le Gléau, H. (2005). MSG/SEVIRI cloud mask and type from SAFNWC. *Int. J. Remote Sens.*, *26*, 4707-4732. doi: 10.1080/01431160500166128

Lensky, I. M., & Rosenfeld, D. (2003). Satellite-based insights into precipitation formation processes in continental and maritime convective clouds at nighttime. *J. Appl. Meteor.*, *42*(9), 1227 - 1233.

Lindsey, D. T., Hillger, D. W., Grasso, L., Knaff, J. A., & Dostalek, J. F. (2006). GOES Climatology and Analysis of Thunderstorms with Enhanced 3.9-μm Reflectivity. *Mon. Wea. Rev.*, *134*(9), 2342-2353.

Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic Press.