

RESEARCH ARTICLE

10.1002/2016MS000822

Key Points:

- Models successfully simulate stratocumulus-to-cumulus transition but underestimate cloud liquid water and cloud ice
- Profiles are largely resolution-independent when convection parameterizations are enabled, but do not converge to results of high-resolution convection-off simulations
- Convection and boundary layer parameterizations interact; a unified approach for scale-aware parameterizations is needed

Correspondence to:

L. Tomassini,
lorenzo.tomassini@metoffice.gov.uk

Citation:

Tomassini, L., P. R. Field, R. Honnert, S. Malardel, R. McTaggart-Cowan, K. Saitou, A. T. Noda, and A. Seifert (2017), The “Grey Zone” cold air outbreak global model intercomparison: A cross evaluation using large-eddy simulations, *J. Adv. Model. Earth Syst.*, 9, 39–64, doi:10.1002/2016MS000822.

Received 26 SEP 2016

Accepted 8 DEC 2016

Accepted article online 20 DEC 2016

Published online 20 JAN 2017

The “Grey Zone” cold air outbreak global model intercomparison: A cross evaluation using large-eddy simulations

Lorenzo Tomassini ¹, Paul R. Field ¹, Rachel Honnert ², Sylvie Malardel ³, Ron McTaggart-Cowan ⁴, Kei Saitou ⁵, Akira T. Noda ⁶, and Axel Seifert ⁷

¹Met Office, Exeter, UK, ²CNRM, Météo France/CNRS, Toulouse, France, ³European Centre for Medium-Range Weather Forecasts, Reading, UK, ⁴Environment and Climate Change Canada, Dorval, Quebec, Canada, ⁵Japan Meteorological Agency, Tokyo, Japan, ⁶Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan, ⁷Deutscher Wetterdienst, Offenbach, Germany

Abstract A stratocumulus-to-cumulus transition as observed in a cold air outbreak over the North Atlantic Ocean is compared in global climate and numerical weather prediction models and a large-eddy simulation model as part of the Working Group on Numerical Experimentation “Grey Zone” project. The focus of the project is to investigate to what degree current convection and boundary layer parameterizations behave in a scale-adaptive manner in situations where the model resolution approaches the scale of convection. Global model simulations were performed at a wide range of resolutions, with convective parameterizations turned on and off. The models successfully simulate the transition between the observed boundary layer structures, from a well-mixed stratocumulus to a deeper, partly decoupled cumulus boundary layer. There are indications that surface fluxes are generally underestimated. The amount of both cloud liquid water and cloud ice, and likely precipitation, are under-predicted, suggesting deficiencies in the strength of vertical mixing in shear-dominated boundary layers. But also regulation by precipitation and mixed-phase cloud microphysical processes play an important role in the case. With convection parameterizations switched on, the profiles of atmospheric liquid water and cloud ice are essentially resolution-insensitive. This, however, does not imply that convection parameterizations are scale-aware. Even at the highest resolutions considered here, simulations with convective parameterizations do not converge toward the results of convection-off experiments. Convection and boundary layer parameterizations strongly interact, suggesting the need for a unified treatment of convective and turbulent mixing when addressing scale-adaptivity.

1. Introduction

The simulation and analysis of cloud transitions have been a major focus in recent research efforts [e.g., Albrecht *et al.*, 1995; Sandu and Stevens, 2011; van der Dussen *et al.*, 2013]. Understanding the morphological evolution of clouds over subtropical and tropical oceans is central when identifying and quantifying cloud feedbacks under climate change [Zhang *et al.*, 2013]. Cloud transitions are also particularly relevant in the context of convection and boundary layer parameterization development because of their tight connection to the representation of scale growth in the boundary layer and convective organization. With global weather and climate models now operating at resolutions that reach from planetary into convection-permitting scales, the question of parameterization adaptiveness has become pressing. Are parameterizations in global numerical models able to adjust to situations in which the model resolution approaches the depth of the boundary layer and the scale of convection?

For this reason the Working Group on Numerical Experimentation (WGNE) of the World Climate Research Programme (WCRP) has initiated the “Grey Zone” project, which was endorsed by GASS (Global Atmospheric System Studies). The aim of the project is to gain insight into how models behave at resolutions from one to around ten kilometers with and without convection parameterizations and to provide guidance for the design of scale-adaptive subgrid parameterizations. The first case study within the “Grey Zone” project is a cold air outbreak over the North Atlantic Ocean in January 2010. Cold air outbreaks are not typical cases of

© 2016. The Authors.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

either the pure deep convective grey zone [Arakawa and Wu, 2013] or the turbulence grey zone [Honnert *et al.*, 2011] but could be called a case of shallow convective grey zone, which has not yet been extensively investigated. The term “grey zone” here denotes a situation in which convective or turbulent vertical transport processes are partly resolved by the dynamics of the numerical model, and partly subgrid-scale and therefore parameterized. Technically speaking the “grey zone” problem relates to the question of how to appropriately define the scale separation between resolved flow and subgrid-scale processes in the context of atmospheric turbulence and convection [Holloway *et al.*, 2014; Yano *et al.*, 2015].

Past studies of cloud transitions mainly focussed on subtropical and tropical transitions [Krueger *et al.*, 1995; Teixeira *et al.*, 2011], extra-tropical cloud transitions, such as those observed in cold air outbreaks, have been less considered. Moreover, the advantage of selecting the cold air outbreak case to study the grey zone problem is that the case involves a genuine interaction between convection and dynamics on a wide range of scales and does not encompass only surface-forced convection as in many tropical and subtropical situations. Maritime cold air outbreaks are characterized by equatorward flow and stratocumulus clouds in the early stage. As air advects over warmer seas the stratocumulus transitions into mixed-phase cumulus clouds [Dorman *et al.*, 2004; Kolstad *et al.*, 2009].

Cold air outbreaks are interesting in their own right and present challenges to the representation of cloud physics in cold environments [Morrison *et al.*, 2012]. Many of the major biases in global climate models are related to processes that are central in cold air outbreaks, such as vertical mixing in the boundary layer and mixed-phase cloud microphysics [Tsuchida *et al.*, 2006; Gettelman *et al.*, 2010; Pithan *et al.*, 2014; Tan *et al.*, 2016].

The present paper summarizes the global model intercomparison conducted as part of the “Grey Zone” project. The case study was designed around the CONSTRAIN aircraft-based field campaign conducted by the UK Met Office [McBeath *et al.*, 2014], but the present study also employs the results from a large-eddy simulation (LES) to augment the observational data set as a baseline for the global model intercomparison. Although none of the participating global models in the present study employ convection parameterizations that are scale-aware by design, the implicit interaction between resolved dynamics and subgrid-scale parameterizations could in principle result in a scale-aware behavior. Moreover, in practice global models are run at “grey zone” resolutions more readily, irrespective of the suitability of their parameterizations. It is therefore legitimate and indeed appropriate to investigate the degree of scale-awareness of global model parameterizations in the framework of a model intercomparison.

The paper is structured as follows. In section 2, we describe the “Grey Zone” global model case and observations, while section 3 covers the simulation with the LES model. In section 4, the global model reference simulations and major common model biases are discussed, while section 5 examines the dependence of the simulations on resolution when the convection schemes are enabled. Section 6 investigates the role of the convection parameterizations and their interaction with the boundary layer schemes based on simulations with convection parameterizations turned off. Finally, section 7 centers on the role of precipitation formation and its possible relation to cloud organization in the cold air outbreak case. A discussion and conclusion section, section 8, completes the paper.

2. Global Model Cold Air Outbreak Simulations and Observations

The case considered in the “Grey Zone” model intercomparison is a cold air outbreak which crossed from the Norwegian Sea into the Atlantic Ocean on 31 January in 2010. A strong northerly flow extended from latitudes higher than 70° N to the southern tip of the British Isles. A more detailed description of the synoptic conditions can be found in Field *et al.* [2014]. The global models are initialized on 30 January at 12 UTC using the European Centre for Medium-Range Weather Forecasts (ECMWF) analysis, except ARPEGE (which uses Météo France analysis), and run for 36 h. The first 12 h are regarded as spin-up time and will not be considered in the subsequent analysis.

Observations from satellites (Figure 1) indicate that in the northern part of the cold air outbreak a closed stratocumulus deck prevailed. As the air mass swept southward over increasingly warm sea surface temperatures the boundary layer deepened, the stratocumulus layer broke up, and the cloud morphological structure transformed into open cellular convection. The global model simulations employed in this study will

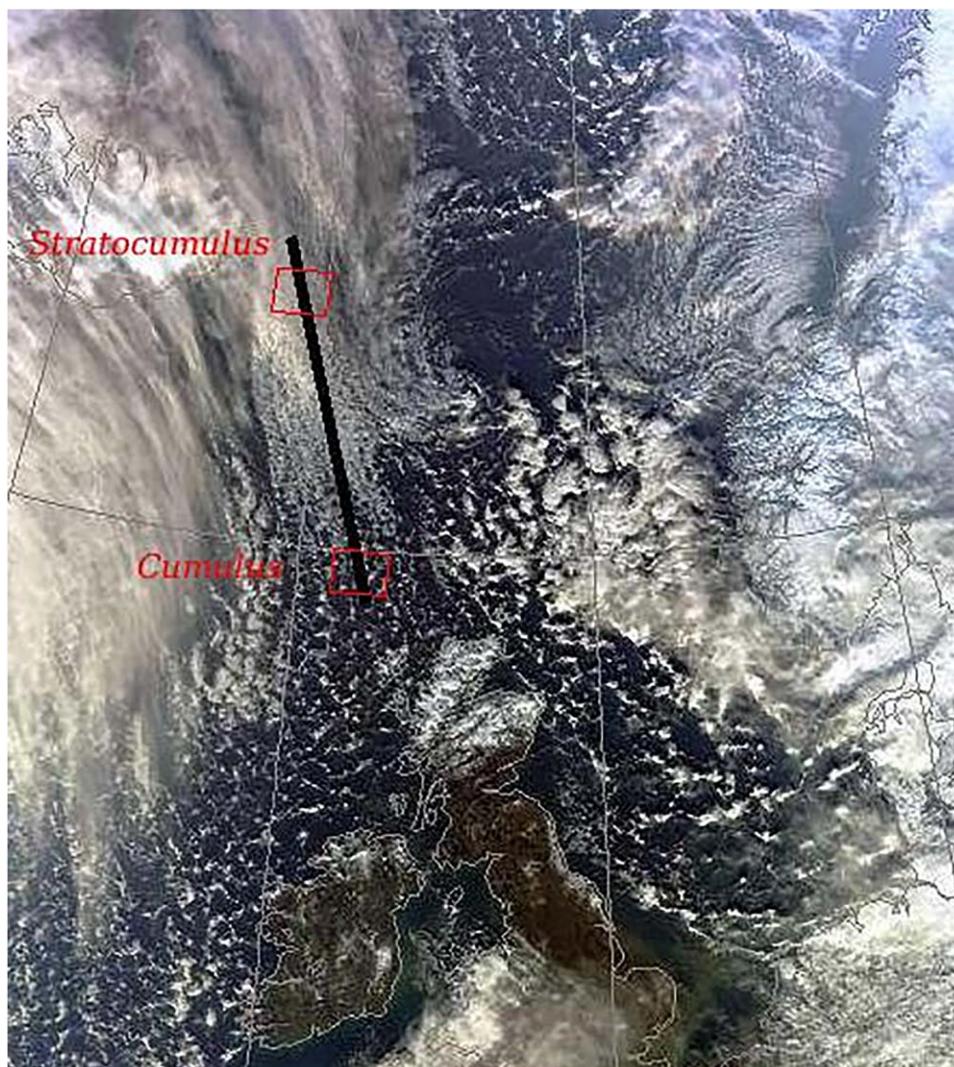


Figure 1. MODIS satellite image for 12:53 UTC on 31 January 2010 (image courtesy of NEODAAS/University of Dundee). The black line indicates the cold air outbreak track (see text and Table 1 for details).

mainly be assessed in a Lagrangian manner, in an attempt to approximately follow the air trajectory. The considered track is based on high-resolution simulations with the Met Office Unified Model and is indicated in Figure 1 as a bold black line, defined more precisely in Table 1. Averages are computed over boxes of sizes 100 km by 100 km along the track. For the subsequent analysis two specific time points and locations, named *stratocumulus* and *cumulus*, are picked. These are representative of the stratocumulus and the cumulus stages of the cloud and boundary layer development, respectively, namely 2 UTC (10.4827° W, 64.9470° N) and 12 UTC (8.27347° W, 59.2450° N) on 31 January 2010.

For evaluation purposes we use observations collected by a FAAM (Facility for Airborne Measurement) BAe-146 aircraft flight from 10 to 15 UTC on 31 January in a region approximately bounded by 8–6° West and 58–60° North [Field *et al.*, 2014]. The observations are considered to be representative for the *cumulus* region. Unfortunately, no observations are available for the *stratocumulus* area.

The global models that participated in the intercomparison are ARPEGE (Météo France), GEM (Environment Canada), GSM (Japan Meteorological Agency), ICON (German Weather Service and Max Planck Institute for Meteorology), IFS (European Centre for Medium-Range Weather Forecasts), NICAM (Japan Agency for Marine Earth Science and Technology/AORI/AICS/NIES), and the Unified Model (UK Met Office). More information on the models can be found in the Appendix. The global model data were interpolated to 22 levels

Table 1. Definition of the Cold Air Outbreak Track

Time (UTC)	Central Box Longitude ^a	Central Box Latitude ^a
00:00	-11.0000	66.0000
00:30	-10.8780	65.7479
01:00	-10.7501	65.4873
01:30	-10.6176	65.2216
02:00	-10.4827	64.9470
02:30	-10.3413	64.6767
03:00	-10.1978	64.4028
03:30	-10.0526	64.1156
04:00	-9.91296	63.8315
04:30	-9.78121	63.5379
05:00	-9.65895	63.2414
05:30	-9.54565	62.9391
06:00	-9.43723	62.6458
06:30	-9.33045	62.3451
07:00	-9.23094	62.0552
07:30	-9.13290	61.7578
08:00	-9.03786	61.4627
08:30	-8.94444	61.1705
09:00	-8.85194	60.8858
09:30	-8.76048	60.6048
10:00	-8.66814	60.3271
10:30	-8.57708	60.0531
11:00	-8.48368	59.7778
11:30	-8.38252	59.5138
12:00	-8.27347	59.2450
12:30	-8.15987	58.9635
13:00	-8.04626	58.6821

^aAverages are taken over boxes of sizes 100 km by 100 km

large-eddy simulation (UCLA-LES) model [Stevens *et al.*, 1999, 2005]. In the UCLA-LES, the parameterization of cloud microphysical processes rests upon the two-moment bulk microphysics scheme formulated by Seifert and Beheng [2006].

The set-up of the “Grey Zone” cold air outbreak large-eddy simulation case adopts a quasi-Lagrangian perspective and is based on high-resolution simulations with the limited area version of the UK Met Office Unified Model. The forcing data used for the LES case are derived from the same track as the one utilized for the global model evaluation and described in section 2. The case includes interactive radiation, forcing of the large-scale vertical velocity to simulate subsidence, and nudging of the meridional wind toward the geostrophic meridional wind profile. No horizontal advective tendencies are specified. The LES is “quasi-Lagrangian” in the sense that the change in the forcing data (of the traditional Eulerian LES set-up) represents the advection of the domain southward over warmer water. There is a spin-up of 1.5 h with environmental conditions corresponding to time (and location) 0 UTC of the track. After the spin-up the LES is run for 13 h. Therefore hour 1.5 of the LES is identified with 0 UTC hereafter.

In the following, the results from a reference simulation are shown the domain of which spans an area of 100 km × 100 km. The horizontal grid spacing is 250 m, with a vertical grid spacing of 25 m. A cloud droplet number concentration of 50 per cm³ is assumed. The LES results depend on the chosen values for

in the vertical between 1000 hPa and 100 hPa except for the Unified Model. In the latter case the data were processed on the native 70 vertical model levels. Table 2 contains an overview of the experiments conducted with the different models. In the reminder of the paper “convection off” experiments are understood to have all convection parameterization turned off completely, including the shallow convection scheme. With the ARPEGE model, experiments were undertaken in which only the deep convective parameterization was switched off, but they were virtually identical to the convection-on experiments and are therefore not included in the present paper. The special case of GEM which includes a “convective mixing” description within the boundary layer scheme is discussed in more detail in section 5.

3. Large-Eddy Simulation of the Cold Air Outbreak Case

In order to set the global model simulations in context with cloud-resolving simulations, and since there are no observations available for the stratocumulus stage of the cold air outbreak, the global model hindcasts are compared to large-eddy simulations with the University of California, Los Angeles

Table 2. Overview Over the Global Model Experiments

Model	Institution	Resolutions	Convection
ARPEGE	Météo France	16 km, 8 km, 4 km, 2 km	On/Off
GEM	Environment Canada	10 km	On/Off/All Off
GSM	Japan Meteorological Agency	TL959 (≈ 20 km)/TL1919 (≈ 10 km)	On/Off
ICON	German Weather Service/Max Planck Institute for Meteorology	10 km, 5 km, 1 km	On/Off
IFS	European Centre for Medium-Range Weather Forecasts	T1279 (≈ 12 km)/T3999 (≈ 5 km)	On/Off (T3999)
NICAM	JAMSTEC/University of Tokyo Atmosphere and Ocean Research Institute/Advanced Institute for Computational Science/National Institute for Environmental Studies	14 km	Off
UM	UK Met Office	N512 (≈ 25 km)	On/Off

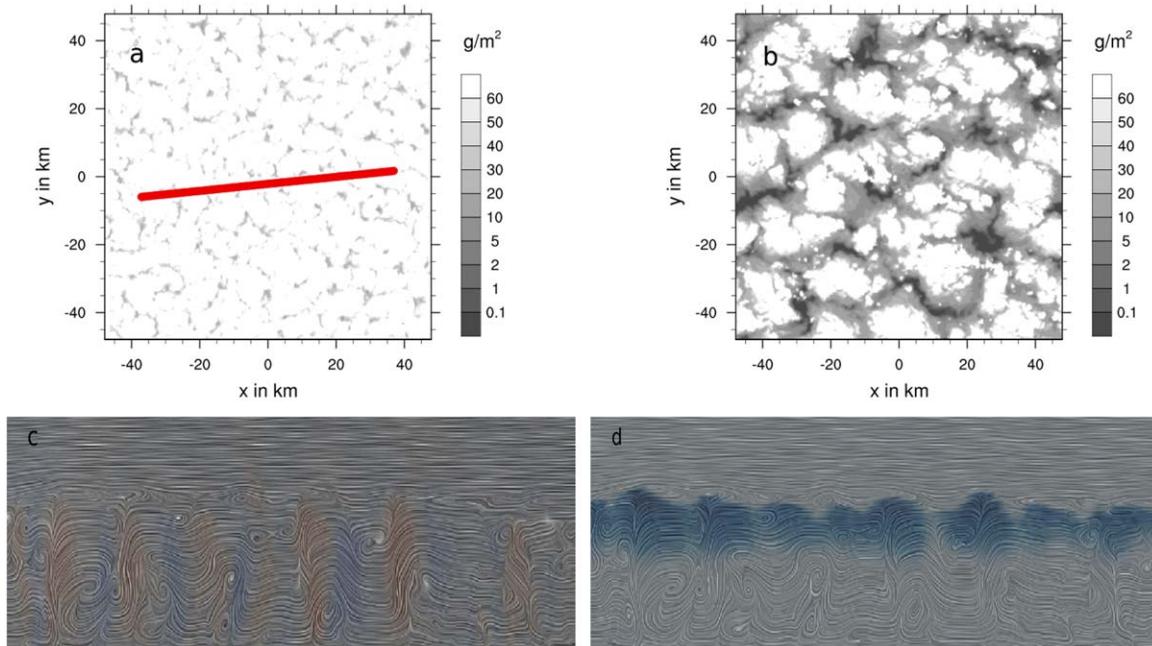


Figure 2. (top row) The sum of liquid and ice water path for the reference UCLA-LES simulation at (a) 2 UTC and (b) 12 UTC. The red line in Figure 2a indicates the position of the cross section used for the plots shown in Figures 2c and 2d. (bottom row) Cross section of streamlines approximately perpendicular to the main wind direction at 00:30 UTC. Colors in Figure 2c indicate upward (red) and downward (blue) vertical motion, blue color in Figure 2d indicates cloud liquid water. Only the lower 2000 m of the cross section are shown.

ice-microphysical parameters related to ice particle shape, fall velocity, and ice particle mass distribution. The parameters for the reference simulation are selected in such a way that the profiles of cloud liquid water and cloud ice are close to the aircraft observations shown in section 4 and at the same time are consistent with observed values reported in *Heymsfield and Kajikawa [1987]*.

The LES integration shows an essentially closed stratocumulus deck at 2 UTC (Figure 2a) and broken cumulus clouds at 12 UTC (Figure 2b). However, the structure of the cumulus clouds does not show distinct open cells which enclose cloud-free areas.

One might expect indications of cloud streaks in the early phase of the LES simulation. Mesoscale circulations are present in the LES, as shown by cross sections approximately perpendicular to the main wind direction at 00:30 UTC (bottom row of Figure 2), but they are not organized in cloud streets (the position of the cross section is indicated by a red line in Figure 2a). Liquid water is formed on top of the strongest updrafts (Figure 2d) whereas cloud ice is present mainly in areas of modest (both positive and negative) vertical velocities (not shown). Boundary layer rolls organized in streaks are typically observed for moderately unstable stratifications and start to dissipate when

$$-\frac{H}{L} \leq c \quad (1)$$

Here H is the boundary layer height, L the Monin-Obukhov length, and c a critical value. *Deardoff [1972]* found a critical value of about 5. However, the critical value itself depends on the magnitude of wind shear, surface heat fluxes, and boundary layer height (*Huang et al. [2009]*, who give a range for c of 5–45). For the cold air outbreak LES, $-\frac{H}{L} = 8.3$ at 00:30 UTC. Since surface buoyancy fluxes are rather large in the present case, the critical value might be rather high. However, some of the global model simulations of the present cold air outbreak case develop quite distinct cloud streets (see section 4).

4. Reference Simulations and Common Model Biases

In this section, we discuss the results of the intercomparison of the global models in what we call their reference configurations (i.e., resolutions are close to their respective operational configurations and convection parameterizations turned on, with the exception of NICAM), and point to a few major common model

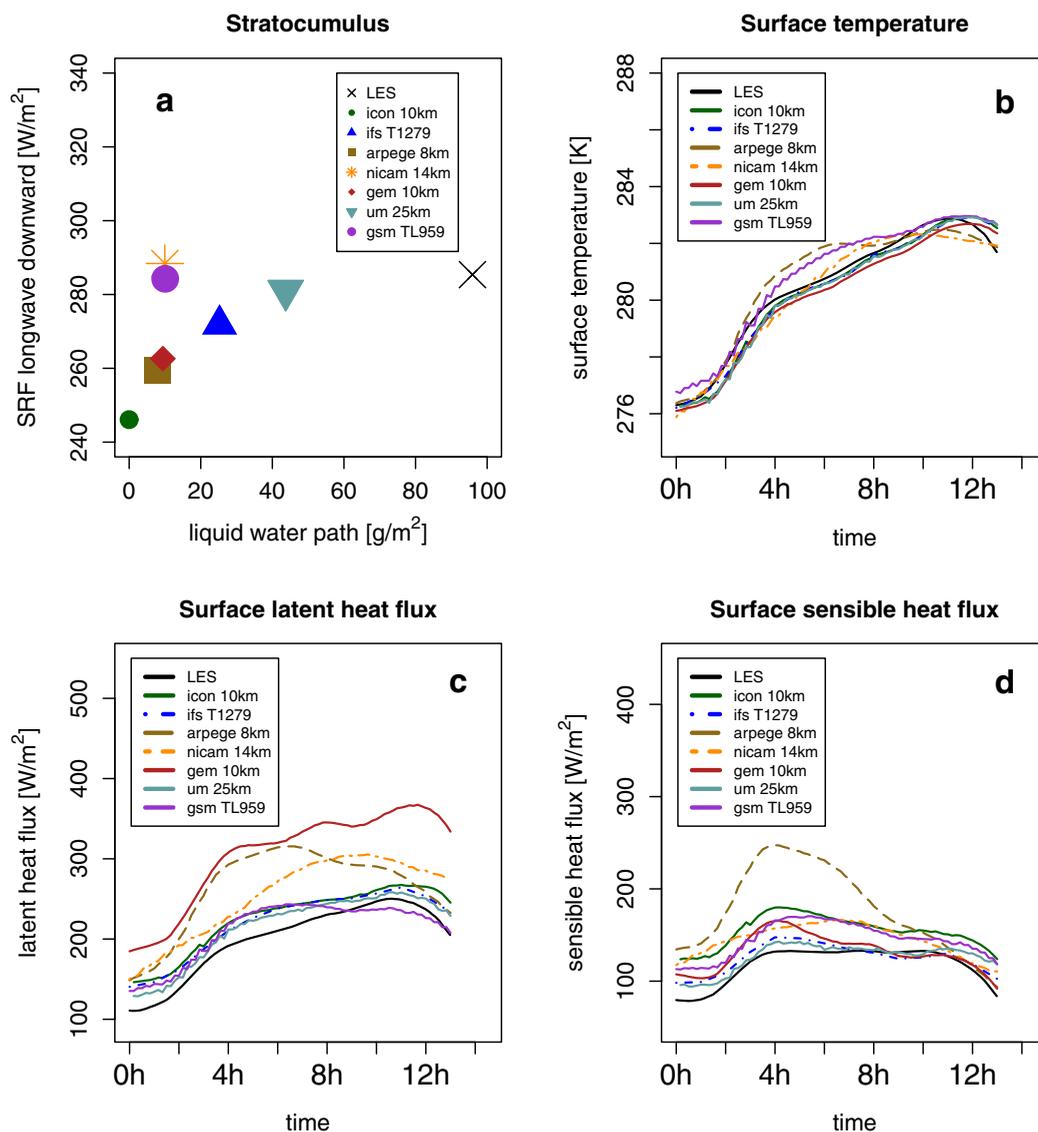


Figure 3. (a) relation between liquid water path and surface (SRF) longwave downward radiative flux at 2 UTC (i.e., the stratocumulus situation) across the ensemble. (b) surface temperature for the UCLA-LES and the global models in their reference configurations along the cold air outbreak track. (c and d) As Figure 3a but for surface latent heat flux and surface sensible heat flux, respectively.

deficiencies in the simulation of the cold air outbreak case. The identification of the model biases will guide the discussion of the role of the convection and boundary layer schemes in the model simulations in subsequent sections.

The amount of condensate, and particularly liquid water, in the model experiments correlates rather strongly with radiative fluxes at the surface and top of the atmosphere. The surface longwave downward flux exhibits an expected log-like relationship with the liquid water path as the emissivity saturates for high liquid water path (Figure 3a). In order to avoid the complication of having to assess differences in radiative transfer calculations, in the remainder of the paper we will focus on cloud liquid and ice water content in the various experiments and not discuss radiative fluxes any further.

In the global model hindcast experiments, the sea surface temperatures (SSTs) are supposed to be prescribed. However, not all models specified the same SSTs in their simulations. Additionally, in some models the surface temperature is allowed to adjust to some degree to the surface fluxes to represent changes in the ocean mixed layer. As a result, SSTs are not identical in all global model simulations (Figure 3b).

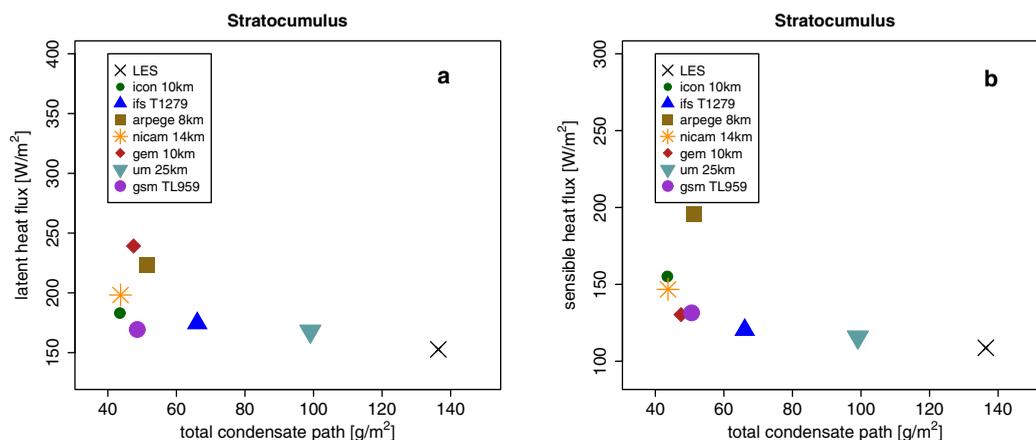


Figure 4. Relation between vertically integrated total condensate mixing ratio and (a) surface latent heat flux and (b) surface sensible heat flux in the *stratocumulus* situation for global models in their standard configurations.

Surface fluxes, both latent and sensible heat fluxes, vary substantially across the model ensemble (Figures 3c and 3d). However, the qualitative behavior of the various models is not attributable to differences in surface fluxes apart from the fact that both ARPEGE and GEM tend to show larger heat fluxes and warmer, more realistic, temperature profiles (see below). Figure 4 shows the relationship between surface heat fluxes and total condensate amount in the *stratocumulus* situation and demonstrates that surface fluxes are not a major control of atmospheric condensate load in the cold air outbreak case.

A key issue in the context of the cold air outbreak simulation is the question of whether the models sufficiently mix heat and moisture in the vertical within shear-dominated boundary layers as encountered in the present case [Field *et al.*, 2014]. Figure 5 displays profiles of liquid water potential temperature and total water mixing ratio for the reference configurations of the models, both for the *stratocumulus* time (i.e., 2 UTC) and location (upper row), and the *cumulus* time (i.e., 12 UTC) and location (bottom row). Each red cross in the bottom row indicates the mean over multiple measurements at a particular height along a leg of the aircraft flight in the *cumulus* region of the cold air outbreak.

The general boundary layer structures are reasonably well reproduced by the models. However, the strength of the inversion in the global models is too weak compared to the UCLA-LES. This is likely partly due the relatively low vertical resolution of the pressure-interpolated data used for the plots, except in the case of the Unified Model. Some models, e.g., ICON, IFS, and GSM, show indications of too weak vertical mixing and too strong a degree of decoupling in the *cumulus* case. Most models are too cold in the sub-cloud layer, except models with strong heat fluxes like ARPEGE and GEM in the *cumulus* case, suggesting that surface fluxes might be generally underestimated. NICAM has a distinctly lower boundary layer height than all other models in the *stratocumulus* situation. Here and in the following the cloud layer is defined by the presence of either cloud liquid water or cloud ice, the top of the boundary layer is considered to be the top of the cloud layer. The boundary layer height of the UCLA-LES also appears to be too low and subcloud layer temperatures too cold when compared to observations, suggesting that either the surface heat fluxes or the cloud-top entrainment of overlying warmer air is under-predicted [Skylingstad and Edson, 2009].

Figure 6 shows profiles of liquid water and total ice mixing ratios for the reference configurations, and observations from the aircraft flight in the *cumulus* case (bottom row of Figure 6). Again each red cross indicates the mean over multiple measurements at a particular height along a leg of the flight. Note that model mean profiles are calculated over an area of 100 km × 100 km, and therefore the observations are comparable to the model results to only a limited extent. Total cloud ice, which includes all cloud ice categories, is shown here because the distinction between various cloud ice categories differs among models.

The models appear to underestimate cloud liquid water in the *stratocumulus* area, certainly when assessed against the UCLA-LES. Field *et al.* [2014] speculate that the *stratocumulus* deck consists mainly of liquid water. The models agree well with regard to the height of the cloud layer for the *stratocumulus* deck, except for NICAM. Also the amount of total ice is quite consistent across models.

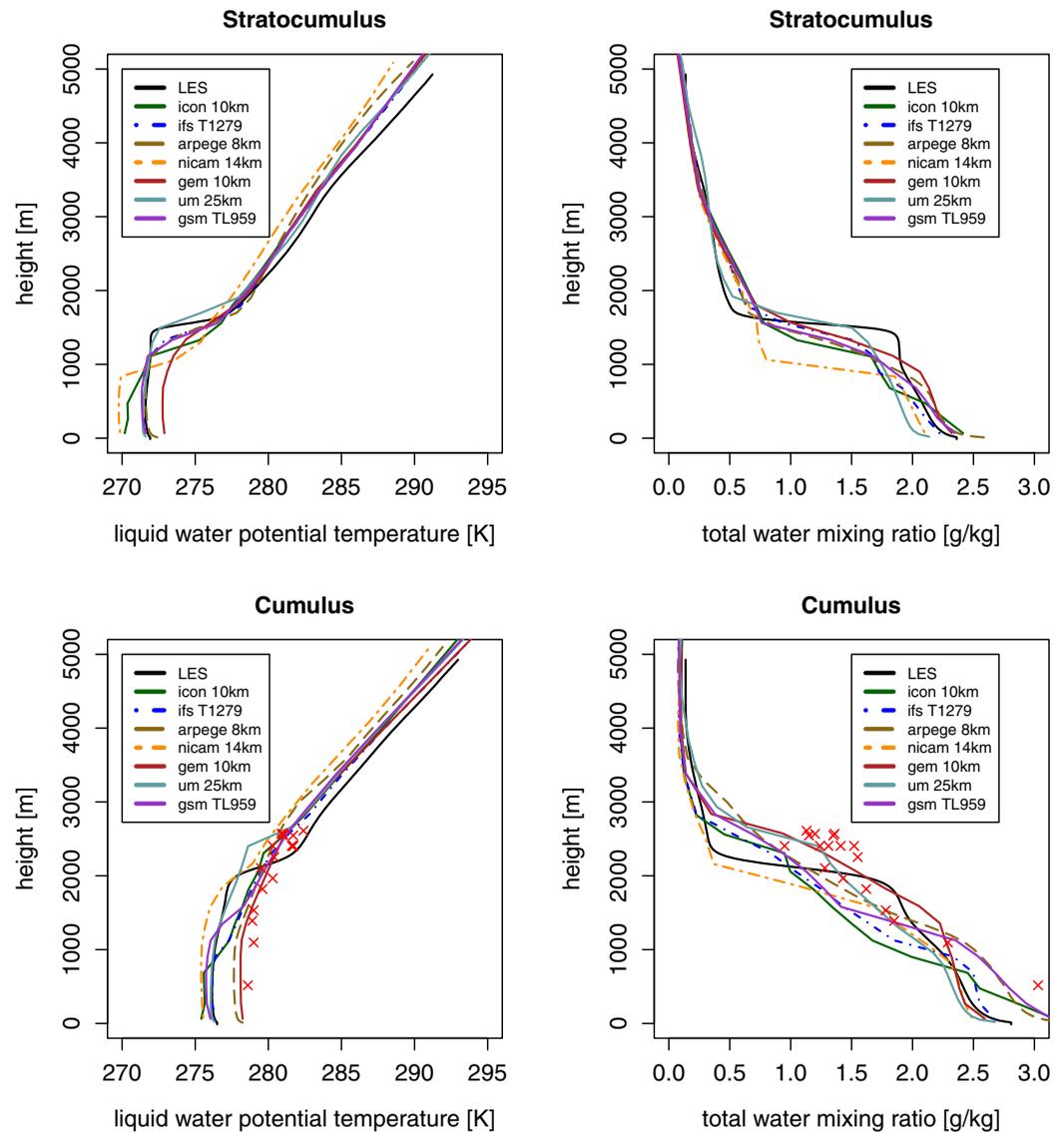


Figure 5. Profiles of liquid water potential temperature (left column) and total water mixing ratio (right column) for the UCLA-LES and the global models in their reference configurations for (top row) the *stratocumulus* situation and (bottom row) the *cumulus* situation. For the *cumulus* case observations from the FAAM aircraft flight are indicated by red crosses. Each cross represents a mean over a number of probes taken during one leg of the flight at a particular height.

The aircraft observations in the *cumulus* case reveal that all models, including the UCLA-LES, significantly underestimate cloud condensate and particularly cloud ice. Moreover, it is again apparent that the UCLA-LES underestimates the boundary layer depth.

Inspecting higher resolution simulations and simulations with convection schemes turned off confirms that underestimation of total cloud condensate, especially in the stratocumulus region, is a common model deficiency across the ensemble. Figure 7 shows the sum of the liquid water and the ice water paths at 12 UTC for three models in their highest resolution configurations, with and without convection parameterization. Black polygons indicate the cold air outbreak track. The cloud morphological structures are quite diverse across the different model simulations. The models generally succeed in representing the broken-up cumulus features in the southern part of the cold air outbreak. However, an apparent common model deficiency is the insufficient southward extent of the closed stratocumulus deck in the northern part, a shortcoming also noted by *Field et al.* [2014]. The models generally break up the stratiform cloud too easily along the cold air outbreak track as a result of the warming SSTs, and instead exhibit broken convective clouds.

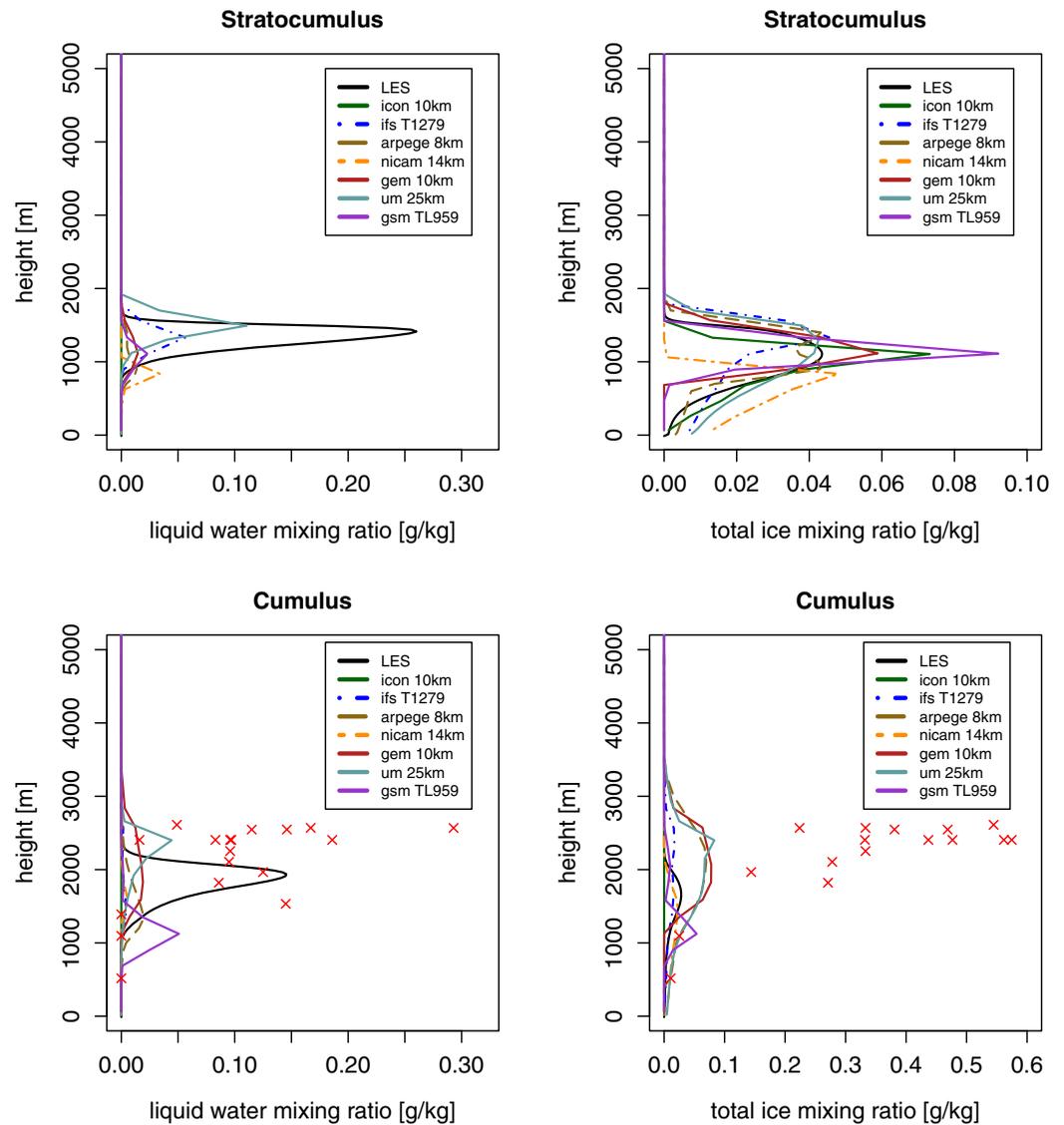


Figure 6. Profiles of liquid water mixing ratio (left column) and total ice mixing ratio (right column) for the UCLA-LES and the global models in their reference configurations for (top row) the *stratocumulus* situation and (bottom row) the *cumulus* situation. For the *cumulus* case observations from the FAAM aircraft flight are indicated by red crosses. Each cross represents a mean over a number of probes taken during one leg of the flight at a particular height.

The ARPEGE model exhibits excessive roll vortex formation that is accentuated by the convection parameterization. Other models do not form cloud streets, or like IFS and to some degree ICON with convection parameterizations turned on, even produce rolls transverse to the large-scale wind direction. Although these differences are interesting in themselves, it is beyond the scope of the present paper to investigate the various aspects of mesoscale boundary layer roll formation in greater depth.

5. Resolution Dependence and the Scale-Adaptivity of Convection Schemes

In this section, we address the question of the degree to which the results of the simulations depend on model resolution when the convection schemes are enabled, mainly in terms of profiles of cloud liquid water and total ice mixing ratios. The behavior of the convective parameterization tendencies and their potential scale-adaptivity is explored. All simulations considered in this section include convection parameterizations.

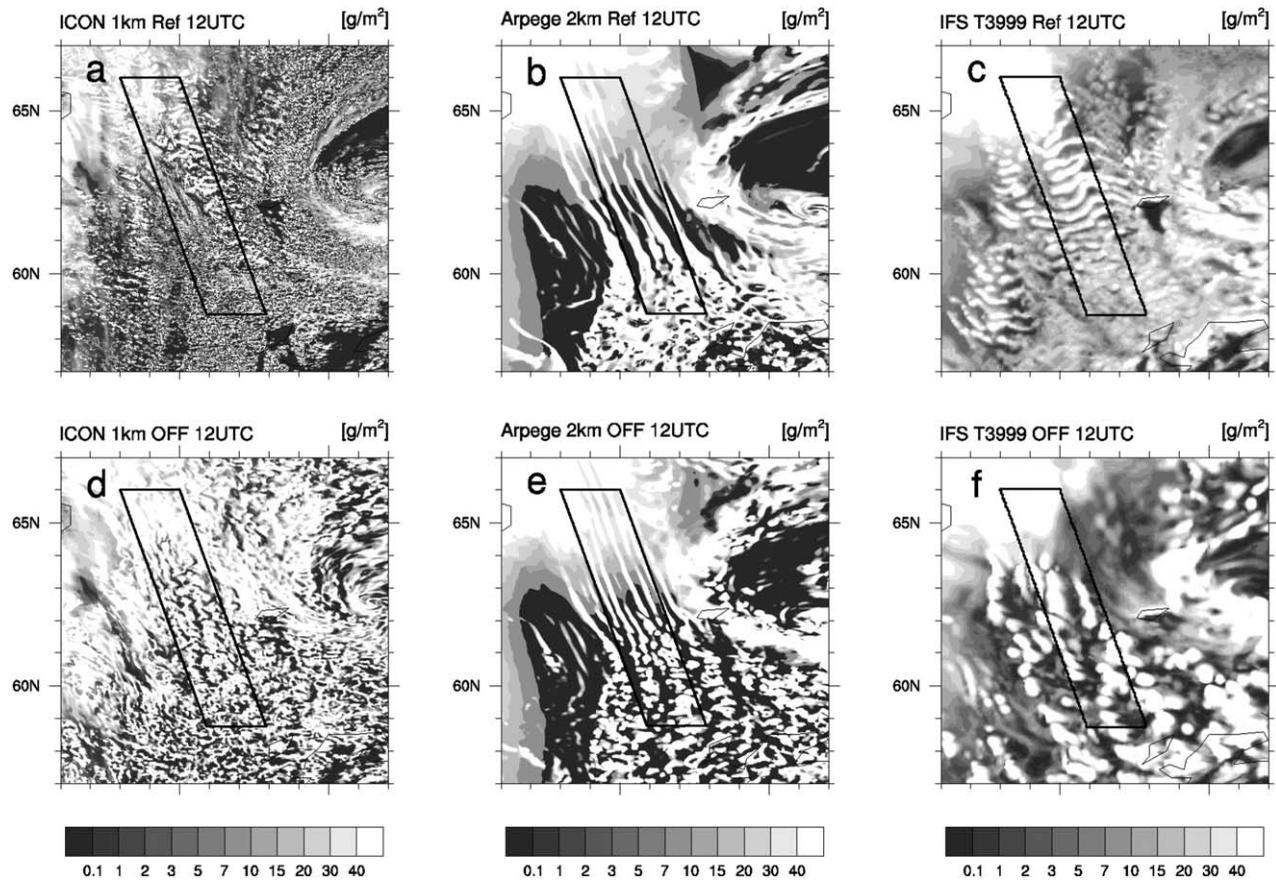


Figure 7. The sum of liquid and ice water path at 12 UTC for three different global models: (a and d) ICON, (b and e) ARPEGE, and (c and f) IFS at their respective highest resolutions (1 km for ICON, 2 km for ARPEGE, and T3999 for IFS). (a–c) The top row shows simulations with convection parameterizations, (d–f) the bottom row simulations with convection parameterizations turned off. The black polygons indicate the cold air outbreak track (see Table 1).

Figures 8 and 9 show cloud liquid water and total ice mixing ratios for a subset of the models with simulations at different resolutions for the *stratocumulus* situation (Figure 8) and the *cumulus* state (Figure 9), all including convective parameterizations. It is striking that, in terms of the profiles, there is hardly any resolution dependence. In the case of ARPEGE, the 2-km and 4-km resolution simulations are very similar, as with 8-km and 16-km simulations, especially for the total ice mixing ratio in the *stratocumulus* case. So for this model a resolution of around 5 km seems to have a certain significance as a threshold, below which the results improve to some degree for the *stratocumulus* boundary layer.

The resolution-independence of the profiles is in principle a desirable property of the models. But what is the underlying reason for this scale-independence? Do the convective parameterizations behave in a scale-adaptive way? For the ICON model, subgrid tendencies from the convection parameterization and from the boundary layer scheme are available separately across a wide range of resolutions for the present study. Inspecting the tendencies from the convection parameterization in isolation reveals that they are essentially independent of the model resolution (Figure 10). The convection scheme removes instability independently of the model grid spacing, preventing a larger proportion of the vertical transport from being resolved explicitly at higher resolutions.

The convection scheme in ICON is not scale-aware by design, and therefore the result is not necessarily surprising. However, in principle the possibility exists that as resolution increases, a larger part of the vertical transport is resolved and the parameterized contribution diminishes as a reaction to the atmospheric profile experienced by the convection parameterization. But the results suggest that the interplay between resolved dynamics and parameterized transport in numerical models is not correctly controlled purely by the fact that the two parts of the model components are called sequentially in the code. Thus, convection

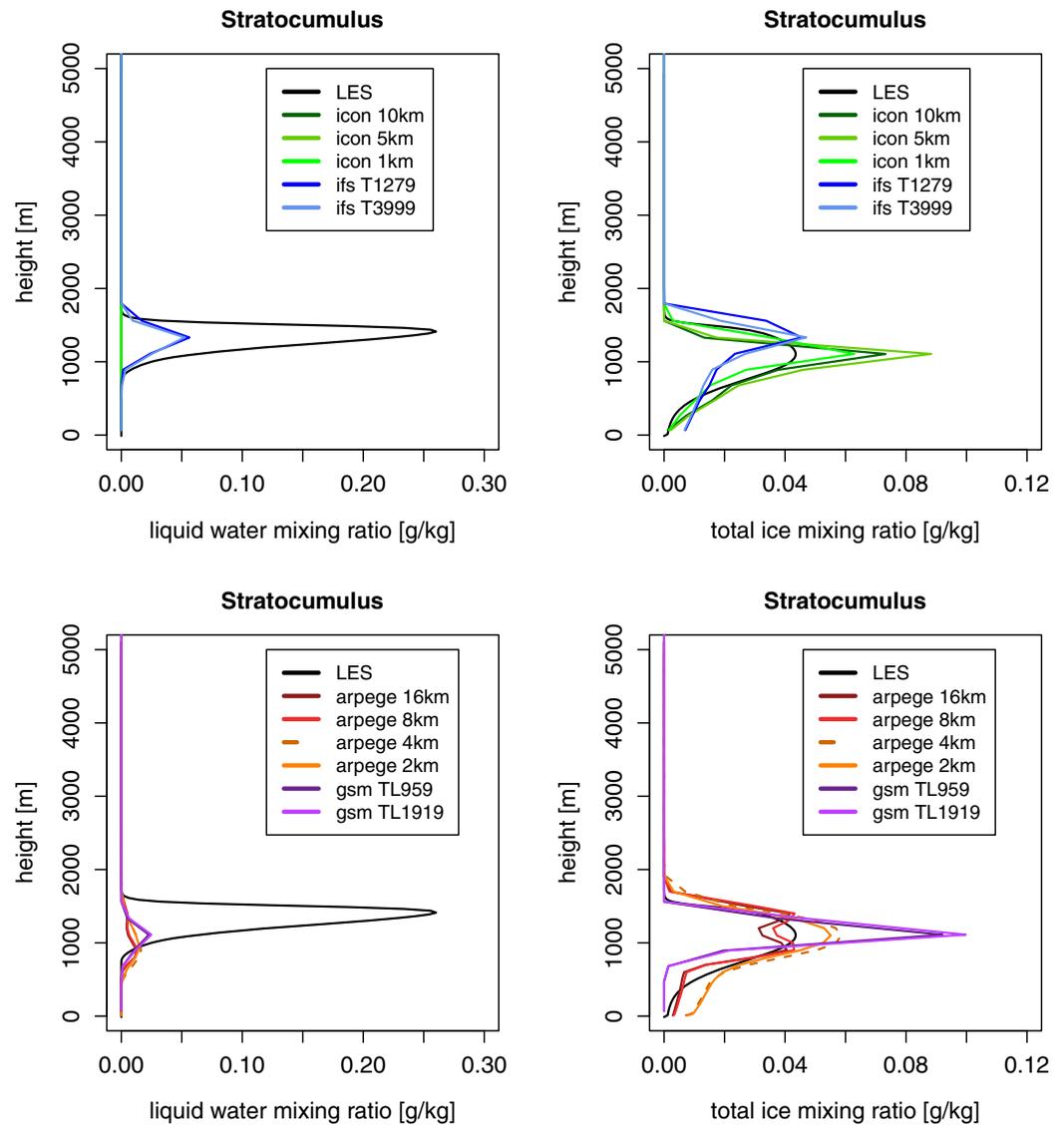


Figure 8. Profiles of (left column) liquid water mixing ratio and (right column) total ice mixing ratio for the UCLA-LES and global models across different resolutions as indicated in the legends for the stratocumulus situation. All simulations, except the UCLA-LES, include convective parameterizations.

parameterizations need to be designed explicitly to behave in a scale-aware manner [Arakawa and Wu, 2013; Grell and Freitas, 2014]. Another aspect of the problem is that the convection parameterization and the boundary layer parameterization may interact. This issue will be examined, among others, in the next section where experiments with convective parameterizations will be compared to simulations with convection schemes turned off.

6. Role of the Convection Parameterization

In the present section we investigate the role of the convection parameterization in more detail and include simulations with convection parameterizations turned off in the discussion. We come back to the question of resolution dependence, focussing on convection-off experiments, and examine the degree to which the convection and boundary layer parameterizations interact across resolutions.

In Figure 11 convection-on and convection-off experiments are compared for selected models at resolutions typical for current global numerical weather prediction models. The convection parameterization

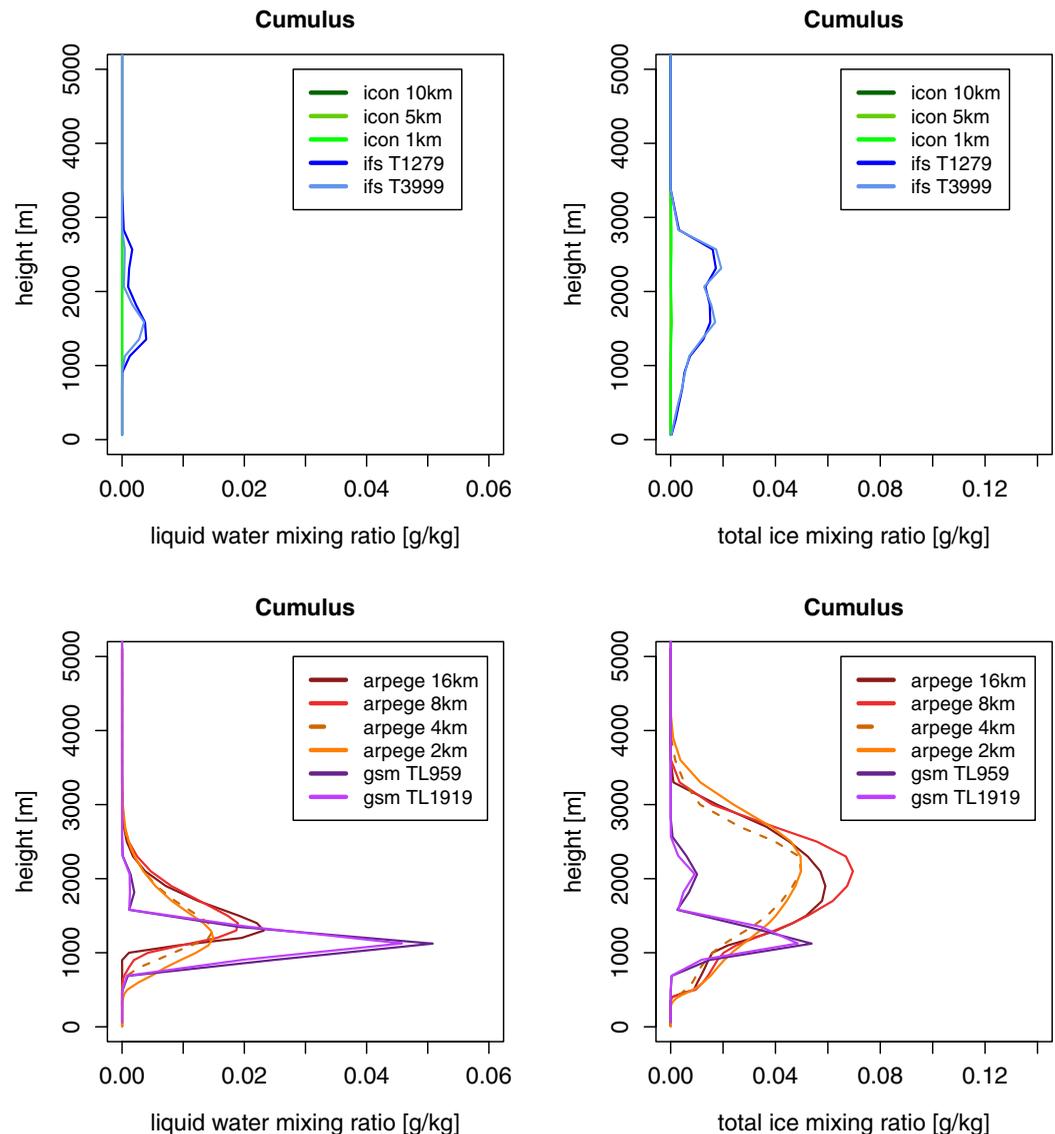


Figure 9. Profiles of (left column) liquid water mixing ratio and (right column) total ice mixing ratio for the UCLA-LES and global models across different resolutions as indicated in the legends for the *cumulus* situation. All simulations, except the UCLA-LES, include convective parameterizations.

insensitivity of some models in the *stratocumulus* situation is mainly due to the fact that in those models the boundary layer parameterization does most of the work. In the *cumulus* case, most models show a rather distinct difference between the simulations with and without convection parameterization indicating that the convection scheme is active in this situation and that it plays a distinct role. The group of models which, in the *stratocumulus* case, shows weak sensitivity to the activation of the convection parameterization includes ARPEGE, the Unified Model, and GSM.

The group of models that shows a clear difference between convection-on and convection-off experiments in the *stratocumulus* case includes ICON, IFS, and GEM. GEM has a separate convection parameterization as well as a radiatively inactive shallow convection adjustment within the boundary layer parameterization in the reference configuration. In the “OFF” simulation the separate convection parameterization is turned off, while in the “ALL OFF” experiment the convection parameterization in the boundary layer scheme is also disabled. We cannot identify the exact cause of the various behaviors for all models in detail here, but in the case of ICON (and possibly other models) an important aspect of the issue is of microphysical nature rather than related to convective or turbulent mixing. When the convection scheme is turned off, all the

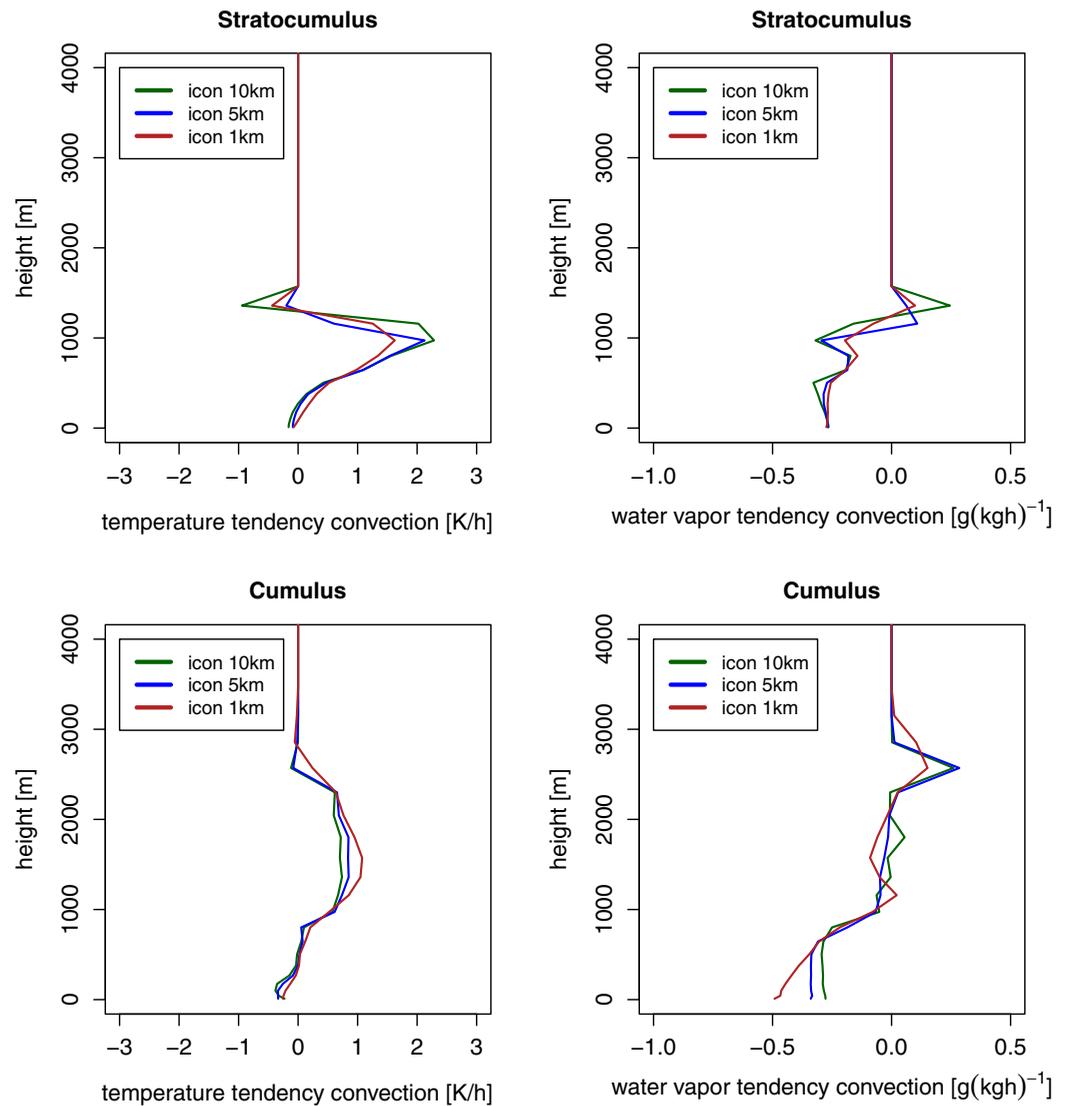


Figure 10. (left column) Temperature tendencies and (right column) water vapor tendencies from the convection parameterization for ICON at different resolutions for (top row) the *stratocumulus* situation and (bottom row) the *cumulus* situation.

microphysics calculations are done by the large-scale microphysics scheme. In the simulation with convection parameterization the simplified microphysics treatment within the convection scheme is active. This simplified convective microphysics is different from the large-scale microphysics in ICON, and in many models.

In section 5 it was established that, in terms of profiles, there is hardly any resolution dependence in the simulations that include convection parameterizations. Is the resolution dependence larger in the simulations with convection parameterization turned off? Figure 12 compares convection-off experiments across resolutions for selected models. In the convection-off experiments the resolution dependence remains modest in the *stratocumulus* case (Figure 12, first row). In the *cumulus* case (Figure 12, second row), however, the dependence is more noticeable. For ARPEGE, the difference between 2 km and 4 km on the one hand, and 8 km and 16 km grid spacing on the other hand, is much more pronounced than in the corresponding convection-on experiments (compare to Figure 9). Also ICON exhibits quite distinct resolution dependence in the liquid water profile.

The stronger resolution sensitivity of the convection-off simulations compared to the convection-on experiments is apparent in the overall amount of cloud liquid water and total ice across the ensemble (Figure 13)

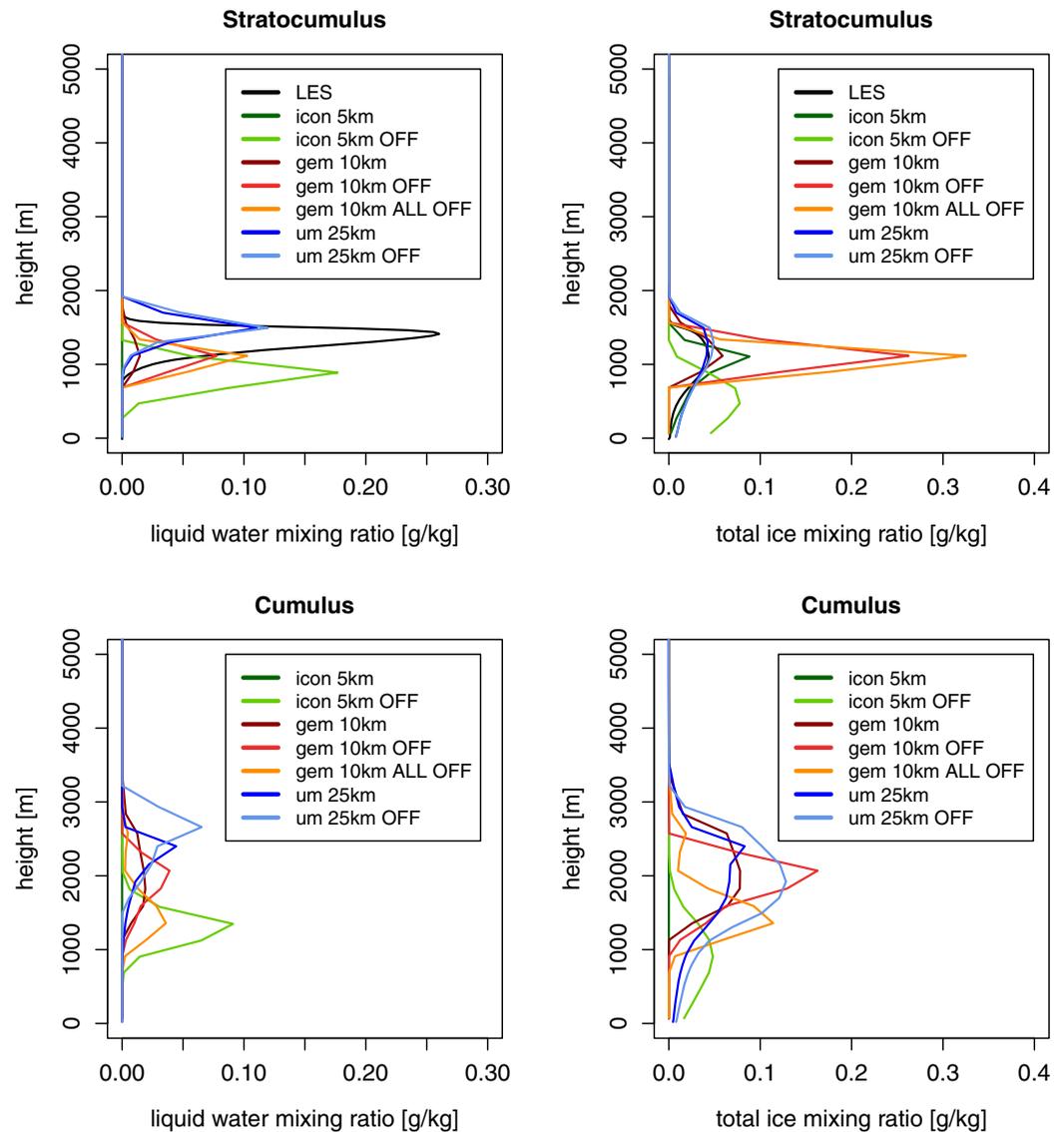


Figure 11. Profiles of (left column) liquid water mixing ratio and (right column) total ice mixing ratio for the UCLA-LES and selected global models with convection parameterizations turned on and off according to the legends, both for (top row) the *stratocumulus* and (bottom row) the *cumulus* situation. In the “OFF” simulation of the GEM model the convection parameterization is turned off, while in the “ALL OFF” experiment the additional representation of convective mixing in the boundary layer scheme is also disabled.

as well. Solid circles in Figure 13 relate to the convection-on experiments, and open circles indicate convection-off simulations. It is notable that increasing the resolution does not have a systematic effect on the amount of condensate, but the sensitivity of the amount of condensate to resolution differs from model to model. In the *cumulus* case (second row of Figure 13), for ARPEGE with convection disabled the total condensate amount increases with enhanced resolution, while for ICON with convection off the total condensate amount decreases with enhanced resolution.

In order to shed more light on the role of the convection parameterization and its interaction with the boundary layer scheme, tendencies from the convection and boundary layer parameterizations are investigated. First, results from the ICON model are presented for which information about the convection and boundary layer parameterizations is available separately across a range of resolutions. Figure 14 shows water vapor and cloud water tendencies of the boundary layer parameterization for convection-on and convection-off experiments. They clearly demonstrate that the convection parameterization and boundary layer parameterization interact: the tendencies depend strongly on whether the convection scheme is

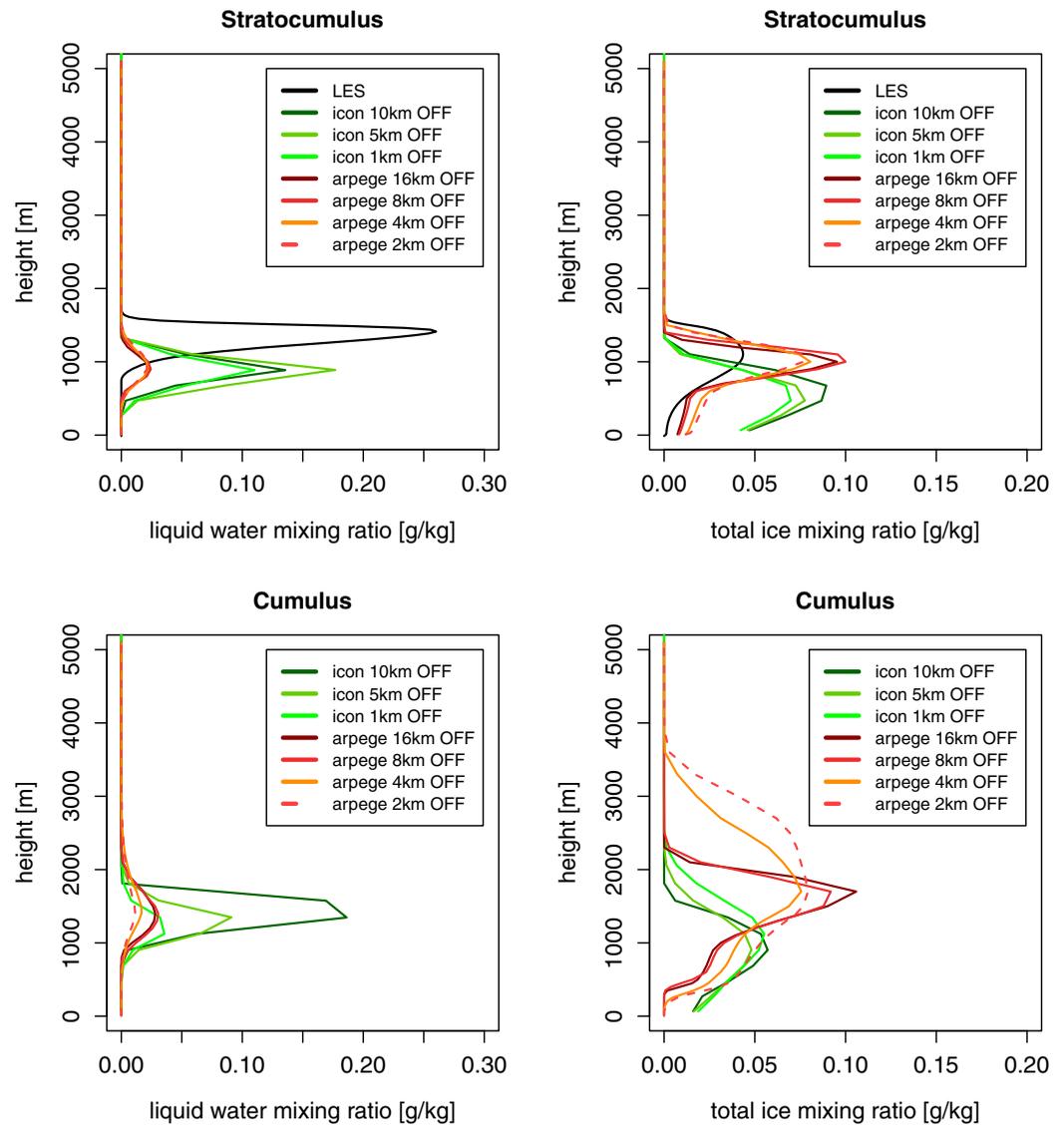


Figure 12. Profiles of (left column) liquid water mixing ratio and (right column) total ice mixing ratio for the UCLA-LES and selected convection-off global model experiments across different resolutions as indicated in the legends, for (first row) the stratocumulus situation and (second row) the cumulus situation.

active. Without the convection parameterization, the boundary layer scheme transports water vapor from the subcloud layer into the cloud layer, and evaporates cloud liquid water in the cloud layer. When the convection scheme is turned on, the boundary layer scheme counteracts the drying tendency of the convection scheme (see Figure 10) in the subcloud layer where it adds water vapor. Since there is no cloud liquid water when the convection scheme is switched on, the turbulence tendencies in the cloud layer vanish when the convection scheme is turned on.

The reason for the absence of cloud water when the convection scheme is turned on is of microphysical origin. When the convection scheme is turned off, the cloud water is produced by the large-scale microphysics scheme. The large-scale microphysics scheme does not produce cloud ice, only some snow (not shown). With the convection scheme turned on the microphysics is handled mainly by the convection scheme. Convective microphysics produces cloud ice (not shown), and hardly any cloud water, which is passed on to the large-scale prognostic variables and appears in the profile. The main reason for the distinct difference between convection-on and convection-off experiment is therefore rooted in the fact that the microphysics is treated differently in the two configurations.

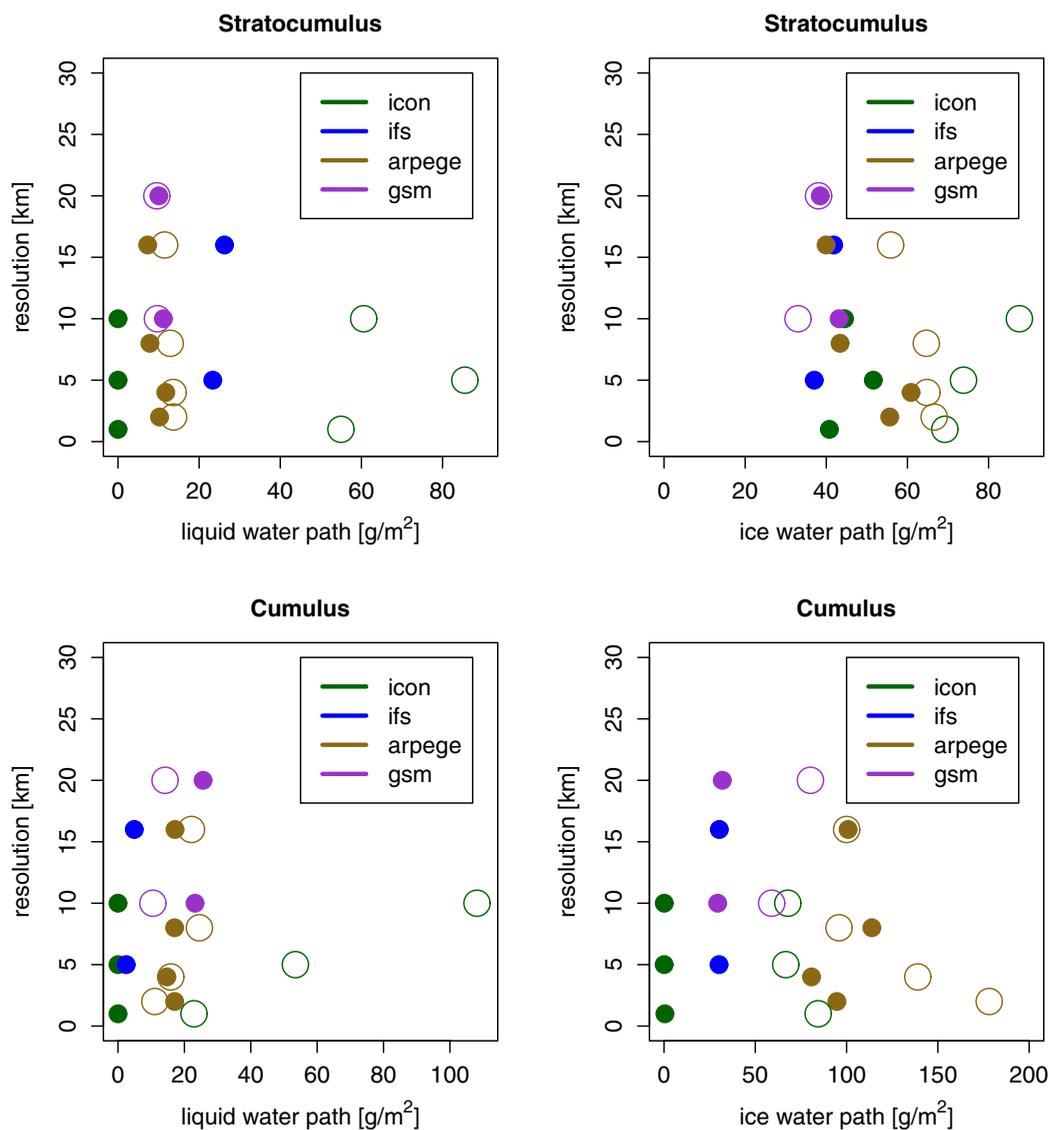


Figure 13. Dependence of the vertically integrated liquid water (left column) and total ice (right column) mixing ratio on the model resolution for (top row) the *stratocumulus* and (bottom row) the *cumulus* situation. Solid circles denote convection-on, open circles convection-off experiments.

From Figure 14 it is also apparent that there is a certain degree of scale-awareness in the boundary layer parameterization when the convection scheme is turned off, especially in the cumulus case where the model resolutions reach into the convective “grey zone.”

Turning our attention to a wider set of models, Figure 15 compares the sum of convection and boundary layer tendencies for convection-on experiments to tendencies in convection-off experiments. As already pointed out, in the case of some models the convection scheme is hardly active in the *stratocumulus* situation. The Unified Model and ARPEGE are examples for this behaviour.

In the *cumulus* case, where convection is more prominent, for many models the sum of the convection and boundary layer tendencies are smaller than the boundary layer tendency alone (with convection scheme turned off) indicating a degree of compensation between turbulent and convective processes when the convection parameterization is active. Also the overall boundary layer height, judged from the cloud top heights indicated by the condensate profiles, is not reduced when the convection scheme is turned off for most models, suggesting that including the convection scheme does not imply stronger vertical mixing of heat and moisture. One might also note that the Unified Model, the profiles of which probably compare

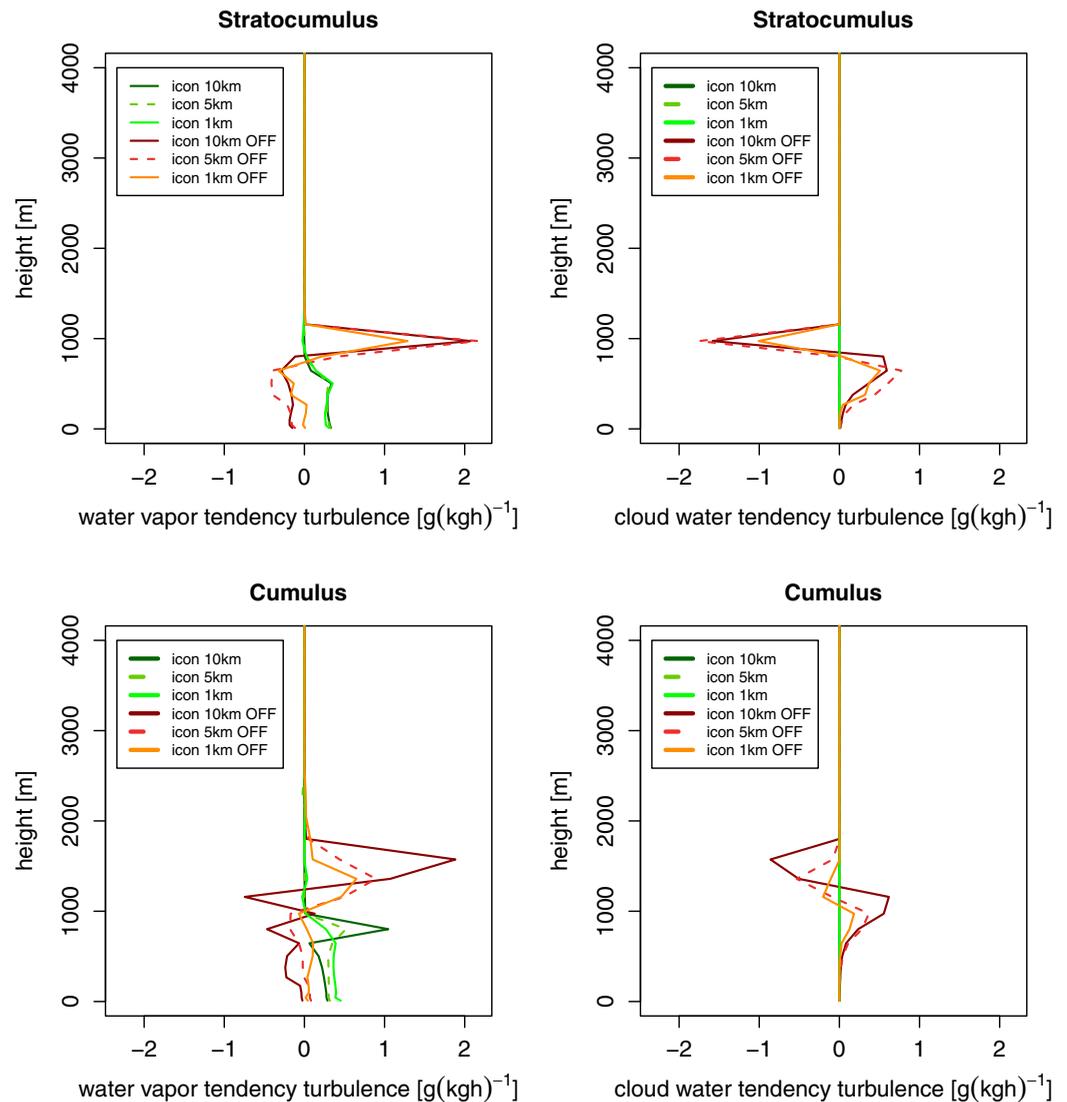


Figure 14. Boundary layer parameterization tendencies of (left column) water vapor and (right column) cloud liquid water from ICON for convection-on and convection-off experiments, for (top row) the *stratocumulus* and (bottom row) the *cumulus* situation. The plots show that the boundary layer scheme is scale-adaptive to some degree, and that the boundary layer tendencies depend on whether the convection scheme is enabled or not.

most favorably with observations (Figure 6), exhibits rather strong subgrid mixing of water from the sub-cloud layer into the cloud layer.

7. Cloud Organization and Regulation by Precipitation

The results so far suggest that the models do not grossly underestimate vertical mixing, and that differences in profiles of condensate mixing ratio can not be fully explained by surface fluxes or vertical mixing strength. This is indicative of the fact that microphysical processes, in particular regulation by precipitation, are an important factor in the cold air outbreak case. In the present section we therefore examine the role of precipitation and its possible relation to cloud organization.

Under nonequilibrium conditions it is difficult to pin down the impact of the efficiency of precipitation formation: the amount of available cloud condensate in the atmosphere that might be depleted by precipitation processes may depend on many factors. The early stage of the cold air outbreak case can, assuming a Lagrangian perspective, more likely be described as an equilibrium situation than the later stage. In the

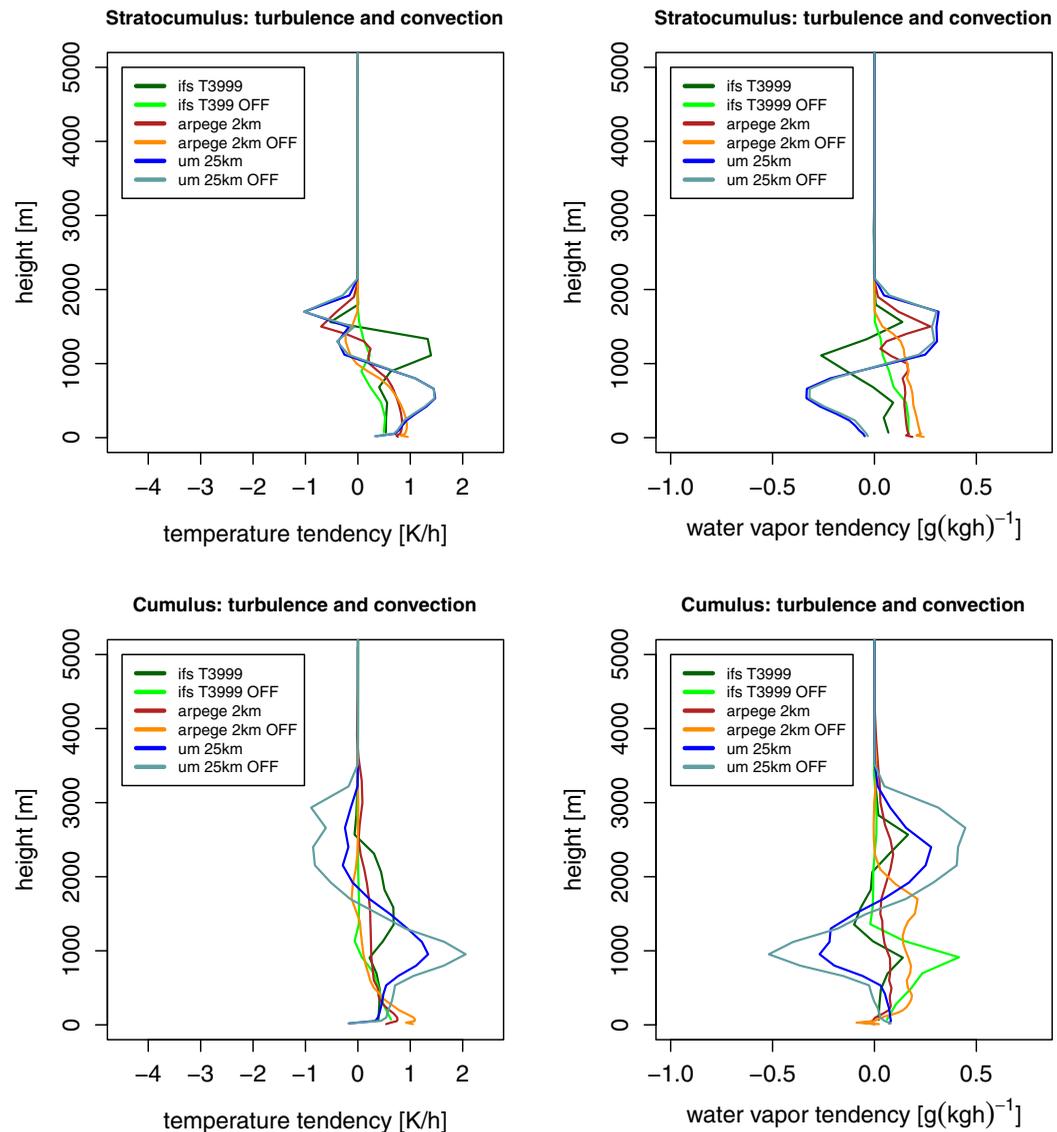


Figure 15. Total tendencies of (left column) temperature and water vapor (right column) from the boundary layer and convection schemes in convection-on and convection-off experiments for selected global models, for (top row) the *stratocumulus* and (bottom row) the *cumulus* situation.

cumulus situation the air has experienced rapidly changing environmental conditions, whereas in the more stationary *stratocumulus* situation the impact of varying precipitation efficiencies in different models may more easily be detected. Figure 16 shows the snowfall rate versus vertically integrated cloud ice mixing ratio in the *stratocumulus* situation for convection-on experiments (Figure 16a) and convection-off simulations (Figure 16b). As discussed in section 6, the microphysics may be handled by different schemes in the convection-on and convection-off experiments. However, in both cases a dependence of the amount of cloud ice on the strength of snowfall is clearly present.

The *cumulus* case is very much a transient situation, and it is more difficult to identify a relationship between precipitation rate and the amount of cloud condensate in the atmosphere across the ensemble. However, looking at individual models across a range of resolutions one might expect a consistent behavior. As discussed in section 6, different models show a different dependence of the total condensate amount in the atmosphere on resolution (see Figure 13). In the *cumulus* case (second row of Figure 13), for ARPEGE with convection off the total condensate amount increases with increased resolution, while for ICON with convection off the total condensate amount decreases with increased resolution. Is this difference behavior

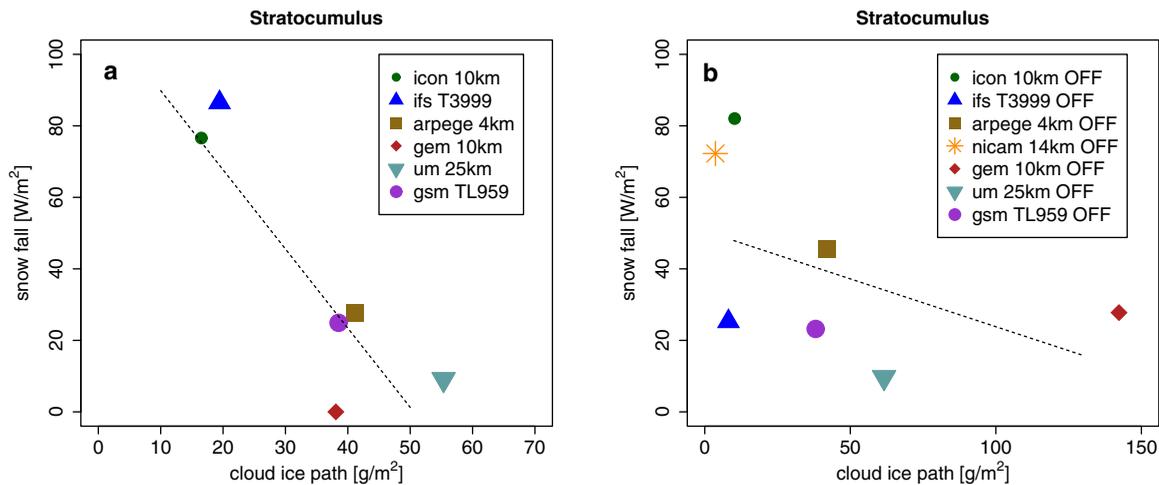


Figure 16. Relation between vertically integrated cloud ice mixing ratio and surface snowfall rate in the *stratocumulus* situation for global models with (a) convection schemes turned on and (b) convection parameterizations switched off. Linear regression lines are indicated by dashed lines.

related to regulation by precipitation? Figure 17 shows a scatter plot of the total condensate path against total precipitation for the ICON and ARPEGE convection-off experiments across resolutions for the *cumulus* situation. One can see that there is no consistent dependence of the atmospheric condensate amount on precipitation. However, higher resolution versions of the respective models consistently show higher precipitation rates, indicating that precipitation strengthens as vertical motion is increasingly resolved by the model. Whether this actually leads to more or less condensate in the atmosphere might again depend on microphysical aspects. Note that for ARPEGE the condensate predominantly consists of cloud ice, while for ICON liquid water dominates over ice (bottom row of Figure 13).

The question whether the strength of precipitation is linked to the form and degree of cloud organization in the simulations suggests itself. By visual inspection of Figure 7 one can see that with regard to cloud organization and scales, the high-resolution simulations exhibit small-scale structures that can not be resolved by coarser models. However, the high-resolution simulations also show organized cloud clusters which exhibit similar scales as the cloud structures apparent in the low-resolution simulations. In the compar-

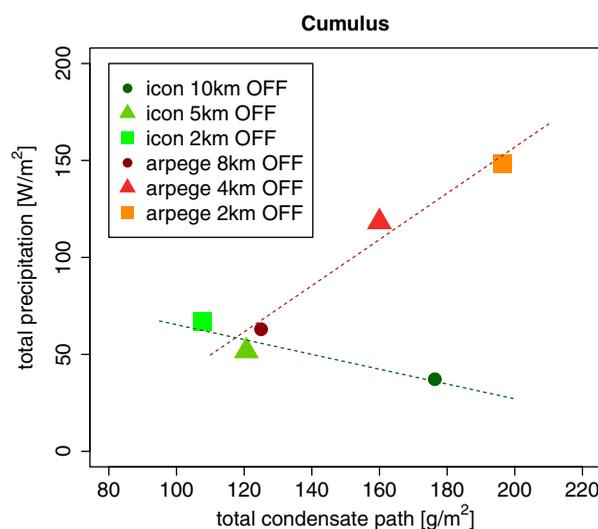


Figure 17. Total precipitation versus total condensate path for convection-off experiments with ICON and ARPEGE in the *cumulus* situation. There is a consistent relation between precipitation and resolution, but not between precipitation and total condensate path. Linear regression lines are indicated by dashed lines.

ison of convection-on versus convection-off simulations, the picture is also not very clear-cut. On the one hand convection-off simulations tend to look “blobbier” (Figure 7, bottom row), while on the other hand the parameterization may enhance cloud organization. In order to obtain a more quantitative view, a spectral decomposition of the variance of the sum of vertically integrated cloud liquid water and cloud ice is performed for ARPEGE and ICON on a domain which is representative of the *cumulus* stage of the cloud development, namely 13° to 5° West and 58° to 62° North. The area comprises about 440×440 km. A Discrete Cosine Transform is used as described in Denis et al. [2002]. The total power of the spectra, i.e., the total variance, is normalized to one.

For ICON the coarser resolution simulations have more power in the larger scales compared to the higher resolution simulations

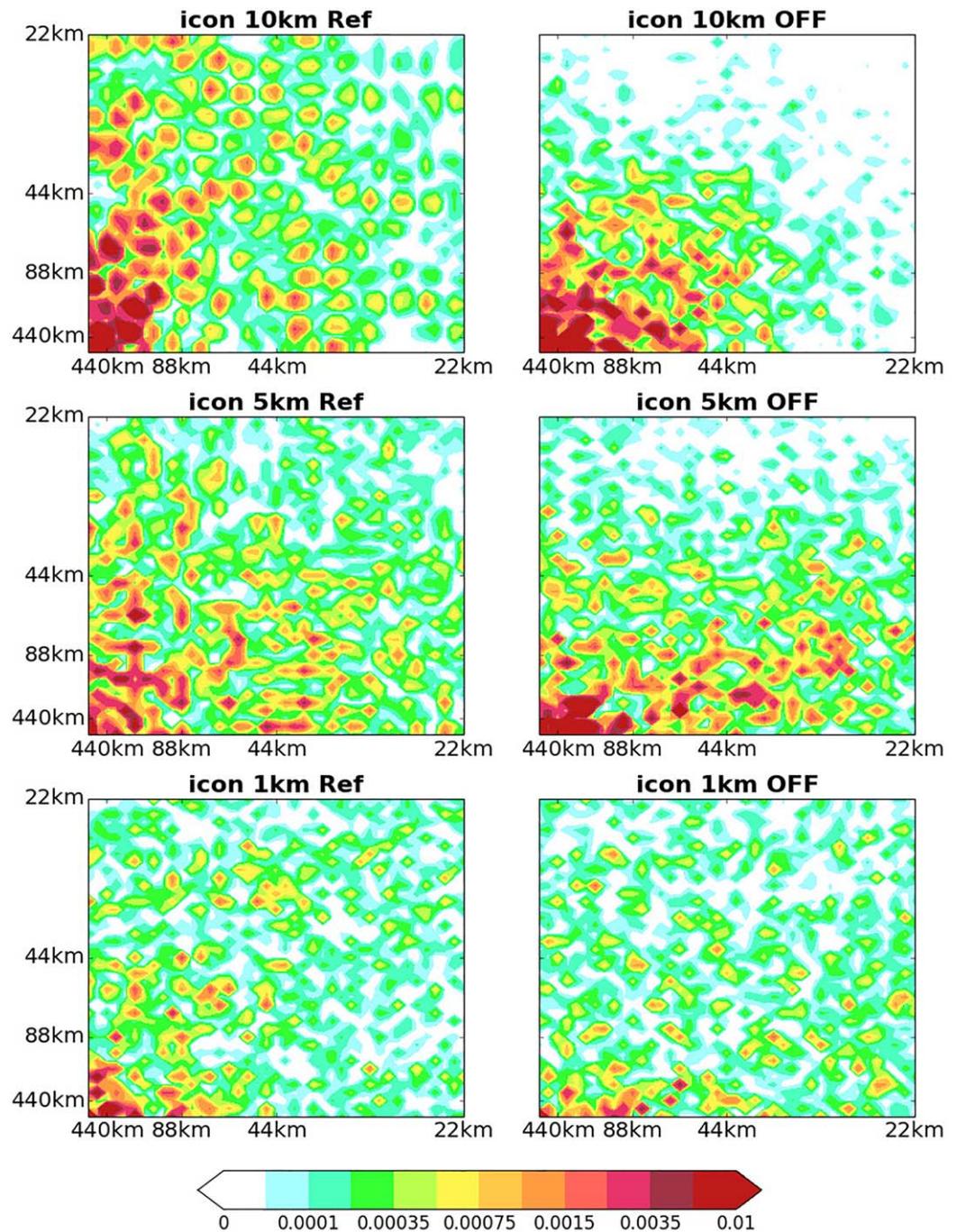


Figure 18. Spectral decomposition of the variance of the sum of liquid and ice water path at 12 UTC for ICON at (top row) 10 km, (middle row) 5 km, and (bottom row) 1 km resolution on the domain 13° to 5° West and 58° to 62° North using a Discrete Cosine Transform. The left column shows the spectral analysis for simulations with convection parameterization, the right column for simulations with convection parameterization turned off. The axes labeling refers to wavelengths. In all cases the total variance is normalized to one, and only wavelengths larger than 22 km are shown.

(Figure 18). Only scales larger than 22 km are shown in the figure since scales smaller than the effective model resolution do not correspond to real features of the simulations and are therefore difficult to interpret. Interestingly, while the convection-on simulations tend to have a concentration of power in large scales along the x-direction, the convection-off experiments exhibit a predominance of power in large scales along the y-direction, the latter being more consistent with cloud street development (see also

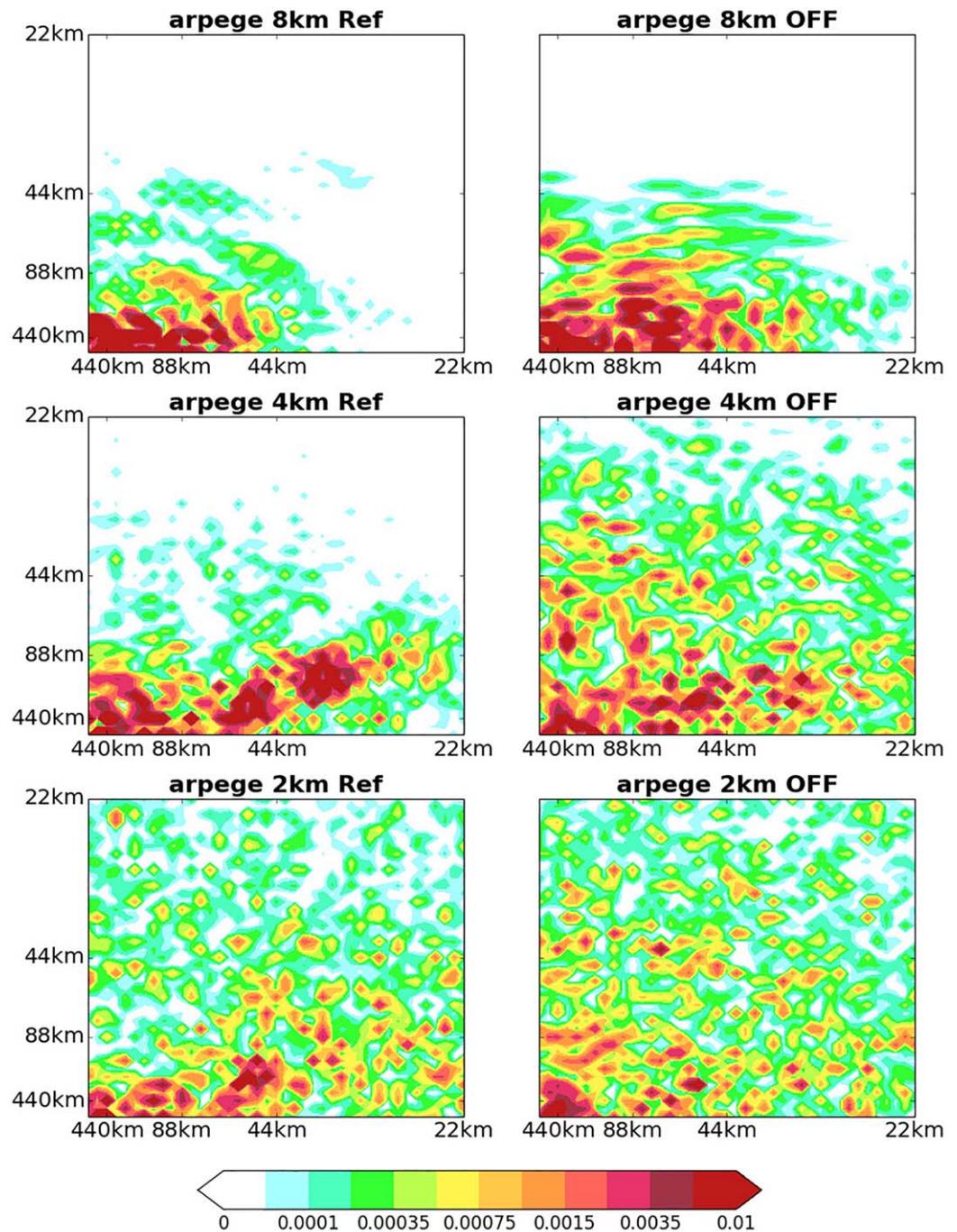


Figure 19. Spectral decomposition of the variance of the sum of liquid and ice water path at 12 UTC for ARPEGE at (top row) 8 km, (middle row) 4 km, and (bottom row) 2 km resolution on the domain 13° to 5° West and 58° to 62° North using a Discrete Cosine Transform. The left column shows the spectral analysis for simulations with convection parameterization, the right column for simulations with convection parameterization turned off. The axes labeling refers to wavelengths. In all cases the total variance is normalized to one, and only wavelengths larger than 22 km are shown.

Figure 7). Generally for ICON the simulations with convection parameterization turned on have more power in the smallest scales compared to simulations with convection parameterization switched off, although less so at 1-km resolution. There are important differences in the behavior of ARPEGE (Figure 19). Again only scales larger than 22 km are shown in the figure. The dominance of large scales in y-direction, the strong longitudinal rolls, is apparent in both convection-on and convection-off experiments. The smallest

scales apparently have more power in the convection-off simulations compared to convection-on. In ARPEGE it appears that the convection parameterization tends to support the organization of convection and clouds in the cold air outbreak case.

So in summary, regarding the convection-off experiments, for both ICON and ARPEGE the higher resolution simulations show stronger precipitation rates and more power in the smaller scales. However, no consistent impact of the convection parameterization could be identified either with regard to precipitation or with regard to cloud organization.

The radar of the United Kingdom Weather Radar Network at Drium-A-Starraig, at the northern tip of Scotland, suggests a mean precipitation rate of around 174 W/m^2 (the radar does not distinguish between rain and snow) for an area of about $100 \text{ km} \times 100 \text{ km}$, bounded by $(-8.2\text{E}, 59.2\text{N})$, $(-8.124\text{E}, 58.355\text{N})$, $(-5.988\text{E}, 59.293\text{N})$, and $(-6.506\text{E}, 58.42\text{N})$, during most of 31 January 2010. The region is representative of the *cumulus* situation. The global models show a spread in total precipitation rate of 0 W/m^2 (GEM) to about 120 W/m^2 (ICON) in the *cumulus* case. Although the area is more than 100 km away from the radar location and therefore a possible error of the order of 50% has to be taken into account, the measurement suggests that the models, in addition to underestimating cloud liquid water and cloud ice, also tend to underestimate precipitation in the *cumulus* area.

8. Discussion and Conclusions

A stratocumulus-to-cumulus cloud transition observed in a cold air outbreak on 31 January 2010 over the North Atlantic Ocean has been explored in simulations with an LES model and seven global climate and numerical weather prediction models as part of the WGNE “Grey Zone” project. The models succeed in representing the general stratocumulus and cumulus boundary layer structures of the northern and southern areas of the cold air outbreak, respectively. However, the closed cloud decks do not extend southward enough in the simulations, and the models consistently underestimate both cloud liquid water as well as cloud ice amounts.

The spread in simulated surface heat fluxes is large, but it does not explain the differences in atmospheric condensate across the ensemble. Turbulent surface heat fluxes seem to be underestimated considering the subcloud-layer cold bias in most models compared to observations. Too weak cloud-top entrainment might also be a cause; however, inversion strengths are too small when compared to the LES. Deficiencies in vertical mixing of heat and moisture are identifiable, with some models showing too markedly a tendency toward decoupling. As discussed in *Field et al.* [2014], the presence of strong wind shear might prevent the turbulence schemes from efficiently mixing across the whole depth of the boundary and cloud layer, especially in the northern stratocumulus area of the cold air outbreak. But also in the more southern cumulus region, the model with largest tendencies from subgrid mixing, the Unified Model, best agrees with observations of atmospheric liquid water and cloud ice content.

Profiles of cloud liquid water and cloud ice are not resolution-dependent in the case when convection parameterizations are enabled in the models. This suggests that the convection parameterizations remove instability in the atmosphere regardless of the model resolution, highlighting the need for parameterizations to be scale-aware by design.

For some models in the *stratocumulus* situation, the convection scheme is essentially not triggered. In the Unified Model this was achieved deliberately by adding a dynamical constraint in order to suppress the triggering of the convection scheme in shear-dominated boundary layers [*Field et al.*, 2014]. For other models in the *stratocumulus* case, and for all models in the *cumulus* situation, convection-on experiments differ substantially from convection-off simulations even at the highest model resolutions considered. This implies that the behavior of the convection parameterizations is not scale-adaptive in the sense that their tendencies would become small in the high-resolution limit. In principle such a scale-dependent behavior could occur even when the schemes are not scale-aware by design, purely due to the fact that the parameterizations experience different profiles depending on the model resolution. However, the present study demonstrates that the convection parameterizations tend to remove atmospheric instability too easily not allowing the dynamics to resolve larger parts of the vertical mixing as the model grid spacing decreases. Moreover, a comparison of parameterization tendencies from convection-on and convection-off experiments reveals that

convection and boundary layer parameterizations do not operate independently, but crucially impact each other. Turning on the convection scheme provokes the boundary layer parameterization to counteract the convective drying tendency in the subcloud layer. This also leads to the fact that overall vertical mixing from the sum of convection and turbulence parameterizations is smaller than the vertical mixing from the turbulence scheme alone for most models.

A somewhat more distinct resolution dependence of the profiles can be observed when the convection parameterizations are switched off. The boundary layer tendencies exhibit a certain level of scale-awareness, clearly more so than the convection parameterizations. Indeed, while none of the convection parameterizations considered here is scale-aware by design, both the turbulent fluxes and the lower-atmospheric mixing length in the boundary layer schemes typically depend on the low-level wind speed which may be a function of resolution. As resolution is changed, the unresolved component of the wind changes, leading to a certain degree of scale-awareness.

In the ensemble of simulations considered, resolution per se does not have a systematic impact on the amount of condensate in the atmosphere, something which is also the case in the absence of convective parameterizations. This may seem surprising as one might expect that high-resolution simulations are better able to resolve vertical mixing processes. However, resolution seems to correlate with the rate of precipitation, at least for the two models for which convection-off simulations are available at a range of resolutions. For both of the models, high-resolution simulations show larger precipitation rates than coarser model configurations.

Given that models with well-mixed boundary layers do not necessarily exhibit more cloud condensate it has to be concluded that the degree of vertical mixing is not the only, or even main, cause for differences in condensate profiles. The representation of cloud microphysical processes exerts an important control on cloud liquid water and cloud ice profiles in the cold air outbreak case. This is also one of the reasons why distinct differences between simulations with convective parameterizations turned on or off can be observed. When the convection parameterization is switched on, the microphysics is handled to a substantial degree by the convection scheme, while with convection parameterization turned off the large-scale microphysics scheme takes over fully.

The present study shows that the question of scale-adaptivity is intimately linked to the problem of conceptualizing the interaction between convective and turbulent mixing in the boundary layer, i.e., the interaction between the convection scheme and the boundary layer parameterization. The importance of this issue was also pointed out by *Zhang et al.* [2013] in the context of subtropical stratocumulus-to-cumulus transitions and questions related to the simulation of climate change. Thus for a satisfactory formulation of scale-adaptivity in subgrid parameterizations it is necessary to adopt a unified approach which comprises both the convection as well as the boundary layer scheme.

Appendix A: Global Models

The global models used in the present study are described in the following paragraphs. They are all run on a global domain. However, ARPEGE employs a stretched grid configuration, and the ICON model setup is based on one-way nested grid refinements on subdomains for the 5 km and the 1 km resolution simulations. Given the short length of the simulations and the large extent of the subdomains the fact that the ICON high-resolution experiments are effectively run on a limited area will not affect the results.

A1. The ARPEGE Model

ARPEGE (Action de Recherche Petite Echelle Grande Echelle) is the global spectral model operationally used at Météo France [Courtier *et al.*, 1991]. The deep convection is represented by a mass-flux scheme based on Bougeault [1985]. The boundary-layer parameterizations is an eddy-diffusivity mass-flux scheme combining a prognostic turbulent kinetic energy based turbulence scheme [Cuxart *et al.*, 2000] and a mass-flux scheme [Bechtold *et al.*, 2001].

A2. Global Environmental Multiscale (GEM) Model

The Canadian Global Environmental Multiscale (GEM) model employs an implicit, semi-Lagrangian dynamical core with a staggered log-hydrostatic-pressure vertical coordinate [Girard *et al.*, 2014]. The results

presented in this study were generated using the Yin-Yang grid configuration, in which two overlapping limited area latitude-longitude domains cover the globe with a minimum in grid deformation [Qaddouri and Lee, 2011]. Deep moist convection is parameterized following the approach proposed by Kain and Fritsch [1990], while penetrating shallow cumulus are represented using the Kuo-transient scheme described by Bélair *et al.* [2005]. The planetary boundary layer parameterization employs a 1.5 order (prognostic turbulent kinetic energy) closure that conserves potential temperature and specific humidity, with counter-gradient corrections applied to represent cumulus-topped boundary layers [McTaggart-Cowan and Zadra, 2015].

A3. The Global Spectral Model (GSM)

The Global Spectral Model (GSM) of the Japan Meteorological Agency (JMA) is based on the hydrostatic primitive equations with a shallow atmosphere assumption. Time integration rests upon a two-time-level, semiimplicit semi-Lagrangian scheme [Japan Meteorological Agency, 2013]. A prognostic Arakawa-Schubert scheme based on Arakawa and Schubert [1974] is adopted as the cumulus convection scheme. The Mellor and Yamada [1974] level-2 closure is used in the boundary layer scheme.

A4. The ICON (ICOsahedral Nonhydrostatic) Model

The ICON model, here used in the version 1.6.0, is based on a nonhydrostatic dynamical core which is formulated on an icosahedral-triangular Arakawa C grid [Zängl *et al.*, 2015]. The boundary layer scheme employs a 1.5 order prognostic turbulent kinetic energy closure scheme [Raschendorfer, 2001]. The convection parameterization is based on the original scheme of Tiedtke [1989] with modifications described in Bechtold *et al.* [2008]. In addition to prognostic cloud liquid and cloud ice there are diagnostic cloud water and cloud ice variables within the model which take subgrid-scale variability of cloud water and cloud ice into account, mainly in situations where relative humidity is close to, but below, 100%, but also cloud water and cloud ice produced within the convection scheme is considered. It is those diagnosed quantities that are passed to the radiation scheme for the computation of radiative fluxes.

A5. The Integrated Forecasting System (IFS) Model

The dynamical core of the Integrated Forecasting System (IFS) model is hydrostatic, two-time-level, semiimplicit, semi-Lagrangian and applies spectral transforms between grid-point space (where the physical parameterizations and advection are calculated) and spectral space. The moist convection scheme is based on Tiedtke [1989] and represents deep, shallow, and mid-level convection. It has evolved over time and amongst many changes includes a modified entrainment formulation [Bechtold *et al.*, 2008]. The boundary layer scheme uses a K-diffusion turbulence closure everywhere, except for unstable boundary layers where an Eddy-Diffusivity Mass-Flux (EDMF) framework is applied to represent the nonlocal boundary layer eddy fluxes [Köhler *et al.*, 2011].

A6. The NICAM Model

The Nonhydrostatic Icosahedral Atmospheric Model (NICAM) has been jointly developed at the Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Atmosphere and Ocean Research Institute, the University of Tokyo, RIKEN Advanced Institute for Computational Science, and the National Institute for Environmental Studies [Tomita and Satoh, 2004; Satoh *et al.*, 2014]. NICAM is generally run without using a cumulus parameterization. The subgrid-scale turbulent closure scheme used in NICAM simulations is a level-2 Mellor-Yamada model [Mellor and Yamada, 1982; Nakanishi and Niino, 2006; Noda *et al.*, 2010].

A7. Met Office Unified Model (UM)

The Met Office Unified Model Global Atmosphere configuration 6.0 is used in the present study [Walters *et al.*, 2016]. The model dynamical core employs a semiimplicit, semi-Lagrangian formulation to solve nonhydrostatic, fully compressible equations of deep-atmospheric motion [Davies *et al.*, 2005]. Convection is represented by a mass flux scheme based originally on Gregory and Rowntree [1990] but with further developments, including detrainment [Derbyshire *et al.*, 2011]. The boundary layer parameterization is the first order turbulence closure mixing scheme of Lock *et al.* [2000] with modifications described in Lock [2001] and Brown [2008]. Based on the cold air outbreak case and the CONSTRAIN field campaign, an additional dynamical constraint was added to the boundary layer scheme that would suppress the triggering of the convection scheme in shear-dominated boundary layers and instead use the nonlocal diffusion profiles of

the boundary layer scheme to parameterize the mixing right up to cloud top [Field et al., 2014; Walters et al., 2014].

Acknowledgments

We thank Adrian Hill and Stephan de Roode for defining the cold air outbreak LES case, Pier Siebesma for coordinating the WGNE “Grey Zone” project, Cathy Hohenegger and Bjorn Stevens for their endorsement of the project and inspiring discussions, and Tobias Göcke and Martin Köhler for help with the ICON simulations. L.T. and A.S. received partial funding through the Hans-Ertel Centre for Weather Research of the German Weather Service and the Max Planck Institute for Meteorology. Support by Sean Milton and David Walters is gratefully acknowledged. The data used in the present study are available from the authors upon request (lorenzo.tomassini@metoffice.gov.uk).

References

- Albrecht, B. A., C. S. Bretherton, D. Johnson, W. H. Schubert, and A. S. Frisch (1995), The Atlantic stratocumulus transition experiment—ASTEX, *Bull. Am. Meteorol. Soc.*, *76*, 889–904.
- Arakawa, A., and W. H. Schubert (1974), Interaction of a cumulus cloud ensemble with the large-scale environment, Part I, *J. Atmos. Sci.*, *29*, 937–949.
- Arakawa, A., and C.-M. Wu (2013), A unified representation of deep moist convection in numerical modeling of the atmosphere. Part I, *J. Atmos. Sci.*, *70*, 1977–1992.
- Bechtold, P., E. Bazile, F. Guichard, P. Mascart, and E. Richard (2001), A mass-flux convection scheme for regional and global models, *Q. J. R. Meteorol. Soc.*, *127*, 869–886.
- Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo (2008), Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales, *Q. J. R. Meteorol. Soc.*, *134*, 1337–1351.
- Bélair, S., J. Mailhot, C. Girard, and P. Vaillancourt (2005), Boundary layer and shallow cumulus clouds in a medium-range forecast of a large-scale weather system, *Mon. Weather Rev.*, *137*, 1938–1960.
- Bougeault, P. (1985), A simple parameterization of the large-scale effects of deep cumulus convection, *Mon. Weather Rev.*, *113*, 2108–2121.
- Brown, A. R. (2008), Large-eddy simulation and parametrization of the baroclinic boundary layer, *Q. J. R. Meteorol. Soc.*, *122*, 1779–1798.
- Courtier, P., C. Freydier, J. F. Geleyn, F. Rabier, and M. Rochas (1991), The ARPEGE project at METEO-FRANCE, in *ECMWF Seminar Proceedings*, vol. II, pp. 193–231.
- Cuxart, J., P. Bougeault, and J. L. Redelsperger (2000), A turbulence scheme allowing for mesoscale and large-eddy simulations, *Q. J. R. Meteorol. Soc.*, *126*, 1–30.
- Davies, T. M., J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood (2005), A new dynamical core for the Met Office’s global and regional modeling of the atmosphere, *Q. J. R. Meteorol. Soc.*, *131*, 1759–1782.
- Deardoff, J. W. (1972), Numerical investigation of neutral and unstable planetary boundary layers, *J. Atmos. Sci.*, *29*, 91–115.
- Denis, B., J. Côté, and R. Laprise (2002), Spectral decomposition of two-dimensional atmospheric fields on limited-area domains using the Discrete Cosine Transform, *Mon. Weather Rev.*, *130*, 1812–1829.
- Derbyshire, S. H., A. V. Maidens, S. F. Milton, R. A. Stratton, and M. R. Willett (2011), Adaptive detrainment in a convective parametrization, *Q. J. R. Meteorol. Soc.*, *137*, 1856–1871.
- Dorman, C. E., R. C. Beardsley, N. A. Dashko, C. A. Friehe, D. Kheif, K. Cho, R. Limeburner, and S. M. Varlamov (2004), Winter marine atmospheric conditions over the Japan Sea, *J. Geophys. Res.*, *109*, C12011, doi:10.1029/2001JC001197.
- Field, P. R., R. J. Cotton, K. McBeath, A. P. Lock, S. Webster, and R. P. Allan (2014), Improving a convection-permitting model simulation of a cold air outbreak, *Q. J. R. Meteorol. Soc.*, *140*, 124–138.
- Gettelman, A., X. Liu, S. J. Ghan, H. Morrison, S. Park, A. J. Conley, S. A. Klein, J. Boyle, D. L. Mitchell, and J.-L. F. Li (2010), Global simulations of ice nucleation and ice supersaturation with an improved cloud scheme in the Community Atmosphere Model, *J. Geophys. Res.*, *115*, D18216, doi:10.1029/2009JD013797.
- Girard, C., et al. (2014), Staggered vertical discretization of the Canadian Environmental Multiscale (GEM) model using a coordinate of the log-hydrostatic-pressure type, *Mon. Weather Rev.*, *142*, 1183–1196.
- Gregory, D., and P. R. Rowntree (1990), A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure, *Mon. Weather Rev.*, *118*, 1483–1506.
- Grell, G. A., and S. R. Freitas (2014), A scale and aerosol aware convection scheme stochastic convective parameterization for weather and air quality modeling, *Atmos. Chem. Phys.*, *14*, 5233–5250.
- Heymsfield, A. J., and M. Kajikawa (1987), An improved approach to calculating terminal velocities of plate-like crystals and graupel, *J. Atmos. Sci.*, *44*, 1088–1099.
- Holloway, C. E., et al. (2014), Understanding and representing atmospheric convection across scales: Recommendations from the meeting held at Dartington Hall, Devon, UK, 28–30 January 2013, *Atmos. Sci. Lett.*, *15*, 348–353, doi:10.1002/asl2.508.
- Honnert, R., V. Masson, and F. Couvreux (2011), A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometric scale, *J. Atmos. Sci.*, *68*, 3112–3131.
- Huang, Q., J. H. Marsham, D. J. Parker, W. Tian, and T. Weckwerth (2009), A comparison of roll and nonroll convection and the subsequent deepening moist convection: An LEM case study based on SCMS data, *Mon. Weather Rev.*, *137*, 350–365.
- Japan Meteorological Agency (2013), Outline of the operational numerical weather prediction at the Japan Meteorological Agency. [Available at www.jma.go.jp/jma/jma-eng/jma-center/nwp/outline2013-nwp/index.htm.]
- Kain, J. S., and J. M. Fritsch (1990), A one-dimensional entraining/detraining plume model and its application in convective parameterization, *J. Atmos. Sci.*, *47*, 2784–2802.
- Köhler, M., M. Ahlgrimm, and A. Beljaars (2011), Unified treatment of dry convective and stratocumulus-topped boundary layers in the ECMWF model, *Q. J. R. Meteorol. Soc.*, *137*, 43–57.
- Kolstad, E. W., T. J. Bracegirdle, and I. A. Seierstad (2009), Marine cold-air outbreaks in the North-Atlantic: Temporal distribution and associations with large-scale atmospheric circulations, *Clim. Dyn.*, *33*, 187–197, doi:10.1007/s00382-008-0431-5.
- Krueger, S. K., G. T. McLean, and Q. Fu (1995), Numerical simulation of the stratus-to-cumulus transition in the subtropical marine boundary layer. Part I: Boundary layer structure, *J. Atmos. Sci.*, *52*, 2839–2850.
- Lock, A. P. (2001), The numerical representation of entrainment in parametrizations of boundary layer turbulent mixing, *Mon. Weather Rev.*, *129*, 1148–1163.
- Lock, A. P., A. R. Brown, M. R. Bush, G. M. Martin, and R. N. B. Smith (2000), A new boundary layer mixing scheme, Part I: Scheme description and single-column model tests, *Mon. Weather Rev.*, *128*, 3187–3199.
- McBeath, K., P. R. Field, and R. J. Cotton (2014), Using operational weather radar to assess high-resolution numerical weather prediction over the British Isles for a cold air outbreak case-study, *Q. J. R. Meteorol. Soc.*, *140*, 225–239.
- McTaggart-Cowan, R., and A. Zadra (2015), Representing Richardson number hysteresis in the NWP boundary layer, *Mon. Weather Rev.*, *143*, 1232–1258.
- Mellor, G. L., and T. Yamada (1974), A hierarchy of turbulence closure models for planetary boundary layers, *J. Atmos. Sci.*, *31*, 1791–1806.
- Mellor, G. L., and T. Yamada (1982), Development of a turbulent closure model for geophysical fluid problems, *Rev. Geophys.*, *20*, 851–875.

- Morrison, H., G. de Boer, G. Feingold, J. Harrington, M. D. Shupe, and K. Sulia (2012), Resilience of persistent Arctic mixed-phase clouds, *Nat. Geosci.*, *5*, 11–17.
- Nakanishi, M., and H. Niino (2006), An improved Mellor-Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog, *Boundary Layer Meteorol.*, *119*, 397–407.
- Noda, A. T., K. Oouchi, M. Satoh, H. Tomita, S. Iga, and Y. Tsushima (2010), Importance of the subgrid-scale turbulent moist process: Cloud distribution in global cloud-resolving simulations, *Atmos. Res.*, *96*, 208–217.
- Pithan, F., B. Medeiros, and T. Mauritsen (2014), Mixed-phase clouds cause climate model biases in Arctic wintertime temperature inversions, *Clim. Dyn.*, *43*, 289–303, doi:10.1007/s00382-013-1964-9.
- Qaddouri, A., and V. Lee (2011), The Canadian Global Multiscale Model on the Yin-Yang grid system, *Q. J. R. Meteorol. Soc.*, *137*, 1913–1926.
- Raschendorfer, M. (2001), The new turbulence parametrization of LM, *COSMO Newsl.*, *1*, 89–97.
- Sandu, I., and B. Stevens (2011), On the factors modulating the stratocumulus to cumulus transitions, *J. Atmos. Sci.*, *68*, 1865–1881.
- Satoh, M., et al. (2014), The Non-hydrostatic Icosahedral Atmospheric Model: Description and development, *Prog. Earth Planet. Sci.*, *1*, doi:10.1186/s40645-014-0018-1.
- Seifert, A., and K. D. Beheng (2006), A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description, *Meteorol. Atmos. Phys.*, *92*, 45–66, doi:10.1007/s00703-005-0112-4.
- Skyllingstad, E. D., and J. B. Edson (2009), Large-eddy simulation of moist convection during a cold air outbreak over the Gulf Stream, *J. Atmos. Sci.*, *66*, 1274–1293.
- Stevens, B., C. Moeng, and P. Sullivan (1999), Large-eddy simulations of radiatively driven convection: Sensitivities to the representation of small scales, *J. Atmos. Sci.*, *56*, 3963–3984.
- Stevens, B., et al. (2005), Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus, *Mon. Weather Rev.*, *133*, 1443–1462.
- Tan, I., T. Storelvmo, and M. D. Zelinka (2016), Observational constraints in mixed-phase clouds imply higher climate sensitivity, *Science*, *352*, 224–227.
- Teixeira, J., et al. (2011), Tropical and subtropical cloud transitions in weather and climate prediction models: The GCSS/WGNE pacific cross-section intercomparison (GPCI), *J. Clim.*, *24*, 5223–5256.
- Tiedtke, M. (1989), A comprehensive mass flux scheme for cumulus parameterization in large-scale models, *Mon. Weather Rev.*, *117*, 1779–1800.
- Tomita, H., and M. Satoh (2004), A new dynamical framework of nonhydrostatic global model using the icosahedral grid, *Fluid Dyn. Res.*, *34*, 357–400.
- Tsushima, Y., S. Emori, T. Ogura, M. Kimoto, M. J. Webb, K. D. Williams, M. A. Ringer, B. J. Soden, B. Li, and N. Andronova (2006), Importance of the mixed-phase cloud distribution in the control climate for assessing the response of clouds to carbon dioxide increase: A multi-model study, *Clim. Dyn.*, *27*, 113–126, doi:10.1007/s00382-006-0127-7.
- van der Dussen, J. J., et al. (2013), The GASS/EUCLIPSE model intercomparison of the stratocumulus transition as observed during ASTEX: LES results, *J. Adv. Model. Earth Syst.*, *5*, 483–499, doi:10.1002/jame.20033.
- Walters, D. N., et al. (2014), The Met Office Unified Model Global Atmosphere 4.0 and JULES Global Land 4.0 configurations, *Geosci. Model Dev.*, *7*, 361–386, doi:10.5194/gmd-7-361-2014.
- Walters, D. N., et al. (2016), The Met Office Unified Model Global Atmosphere 6.0/6.1 and JULES Global Land 6.0/6.1 configurations, *Geosci. Model Dev. Discuss.*, 1–52, doi:10.5194/gmd-2016-194.
- Yano, Y.-I., et al. (2015), Basic concepts for convection parameterization in weather forecast and climate models: COST action ES0905 final report, *Atmospheres*, *6*, 88–147, doi:10.3390/atmos6010088.
- Zängl, G., D. Reinert, P. Ripodas, and M. Baldauf (2015), 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*, 563–579.
- Zhang, M., et al. (2013), CGILS: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models, *J. Adv. Model. Earth Syst.*, *5*, 826–842, doi:10.1002/2013MS000246.