Confronting the convective gray zone in the global configuration of the Met Office Unified Model

Lorenzo Tomassini¹, Martin Willett², Alistair Sellar³, Adrian Lock², David Walters², Michael Whitall², Claudio Sanchez², Julian Heming⁴, Paul Earnshaw², José Rodríguez⁵, Duncan Ackerley², Prince Kattanchery xavier², Charmaine Franklin⁶, and Catherine Senior²

¹UK Met Office
²Met Office
³Hadley Centre, Met Office
⁴Met Office, U.K.
⁵Met Office, FitzRoy Road, Exeter EX1 3PB, UK
⁶Australian Bureau of Meteorology

November 28, 2022

Abstract

In atmospheric models with kilometer-scale grids the resolution approaches the scale of convection. As a consequence the most energetic eddies in the atmosphere are partially resolved and partially unresolved. The modeling challenge to represent convection partially explicitly and partially as a subgrid process is called the convective gray zone problem. The gray zone issue has previously been discussed in the context of regional models, but the evolution in regional models is constrained by the lateral boundary conditions. Here we explore the convective gray zone starting from a defined global configuration of the Met Office Unified Model using initialized forecasts and comparing different model formulations to observations. The focus is on convection and turbulence, but some aspects of the model dynamics are also considered. The global model is run at nominal 5km resolution and thus contributions from both resolved and subgrid turbulent and convective fluxes are non-negligible. The main conclusion is that in the present assessment, the configurations which include scale-aware turbulence and a carefully reduced and simplified mass-flux convection scheme outperform both the configuration with fully parameterized convection as well as a configuration in which the subgrid convection parameterization is switched off completely. The results are more conclusive with regard to convective organization and tropical variability than extratropical predictability. The present study thus endorses the strategy to further develop scale-aware physics schemes and to pursue an operational implementation of the global 5km-resolution model to be used alongside other ensemble forecasts to allow researchers and forecasters to further assess these simulations.

Confronting the convective gray zone in the global configuration of the Met Office Unified Model

Lorenzo Tomassini¹, Martin Willett¹, Alistair Sellar¹, Adrian Lock¹, David Walters¹, Michael Whitall¹, Claudio Sanchez¹, Julian Heming¹, Paul Earnshaw¹, José M. Rodriguez¹, Duncan Ackerley¹, Prince Xavier¹, Charmaine Franklin², and Catherine A. Senior³

¹ Met Office, Exeter, UK ² Bureau of Meteorology, Melbourne, Australia	
³ Met Office Hadley Centre, Exeter, UK	

Key Points:

1

2

3

4

5 6

7 8 9

10

11	•	The representation of turbulence and convection in the convective gray zone is in-
12		vestigated at global 5km resolution
13	•	In the examined context a reduced mass-flux convection scheme is beneficial in
14		5km-resolution global model forecasts
15	•	The assessment is more conclusive with regard to convective organization and trop-
16		ical variability than extratropical predictability

Corresponding author: Lorenzo Tomassini, Met Office, FitzRoy Road, Exeter EX1 3PB, United Kingdom, lorenzo.tomassini@metoffice.gov.uk

17 Abstract

In atmospheric models with kilometer-scale grids the resolution approaches the scale of 18 convection. As a consequence the most energetic eddies in the atmosphere are partially 19 resolved and partially unresolved. The modeling challenge to represent convection par-20 tially explicitly and partially as a subgrid process is called the convective gray zone prob-21 lem. The gray zone issue has previously been discussed in the context of regional mod-22 els, but the evolution in regional models is constrained by the lateral boundary condi-23 tions. Here we explore the convective gray zone starting from a defined global config-24 uration of the Met Office Unified Model using initialized forecasts and comparing dif-25 ferent model formulations to observations. The focus is on convection and turbulence. 26 but some aspects of the model dynamics are also considered. The global model is run 27 at nominal 5km resolution and thus contributions from both resolved and subgrid tur-28 bulent and convective fluxes are non-negligible. The main conclusion is that in the present 29 assessment, the configurations which include scale-aware turbulence and a carefully re-30 duced and simplified mass-flux convection scheme outperform both the configuration with 31 fully parameterized convection as well as a configuration in which the subgrid convec-32 tion parameterization is switched off completely. The results are more conclusive with 33 regard to convective organization and tropical variability than extratropical predictabil-34 ity. The present study thus endorses the strategy to further develop scale-aware physics 35 schemes and to pursue an operational implementation of the global 5km-resolution model 36 to be used alongside other ensemble forecasts to allow researchers and forecasters to fur-37 ther assess these simulations. 38

³⁹ Plain Language Summary

A first exploration of kilometer-scale global atmospheric modelling in the context 40 of the current Met Office modelling system, the Met Office Unified Model, is presented. 41 All simulations were performed using a global atmosphere model at nominal 5km res-42 olution. The model resolution thus resides in the so-called convective gray zone where 43 the grid length approaches the scale of turbulence and convection, and contributions from 44 both resolved and subgrid convective and turbulent fluxes are non-negligible. A case study 45 approach has been taken in which various testbed cases are defined and model forecasts 46 are evaluated against observations. The focus is on the representation of convection, tur-47 bulence and dynamics, the key aspects of the model formulation in the convective gray 48 zone. The main finding is that in the present assessment the configurations which in-49 clude a scale-aware representation of turbulence and a carefully reduced and simplified 50 convection scheme outperform both the reference model with fully parameterized con-51 vection as well as a configuration in which the subgrid convection parameterization is 52 switched off completely. An outlook on further work towards the development of an ad-53 equate kilometer-scale resolution global coupled modelling system for use across weather 54 and climate time scales is given. 55

56 1 Introduction

Coarse-resolution atmospheric models do not resolve any of the main energy-containing 57 turbulence scales, and therefore turbulence is essentially unresolved and has to be pa-58 rameterized. Wyngaard (2004) discussed models with resolution in between coarse mesoscale 59 models (grid length 10km and larger) and large-eddy resolving models (grid length 100m 60 and smaller) for which the resolution is similar to the length scale of the most energetic 61 eddies in the atmosphere. The range of values for the length scale of these energetic ed-62 dies and coherent structures depends on the prevailing meteorological situation, but can 63 vary from less than 100m in shallow convection to several tens of kilometers in organized 64 convective systems. This has led to the articulation of the "gray zone" problem, the ques-65 tion about the adequate atmospheric model formulation in situations where model res-66

olution approaches the scale of turbulence and convection (Tomassini et al., 2017; Field et al., 2017; de Roode et al., 2019; Honnert et al., 2020).

A distinction is often made between the gray zone of turbulence and the gray zone 69 of convection, but in practice there is no clear-cut separation and the representation of 70 turbulence has a significant impact on, for example, the modelling of organised deep con-71 vective systems (Honnert et al., 2020). Turbulent eddies and convective motions usually 72 occur simultaneously and across a wide range of scales. Ultimately, all atmospheric scales 73 are deeply intertwined. The terms "convection-permitting" or "turbulence-permitting" 74 75 to designate models operating in the gray zone of turbulence and convection (in contrast to "convection-resolving" and "turbulence-resolving") are adequate as they reflect the 76 fact that convection and turbulence are partly resolved and partly unresolved in these 77 regimes. 78

The term "convection-permitting" does not imply that the subgrid convection scheme 79 is switched off completely in the model (Kendon et al., 2021). On the contrary, it sug-80 gests that part of the convective motion is still unresolved and needs to be parameter-81 ized, either through an adequately formulated (ideally scale-aware) and calibrated con-82 vection scheme or a non-local term in the turbulence parameterization, or both. In one 83 of the earliest explorations of the convective gray zone problem, Roberts (2003) suggested 84 to use a CAPE-dependent (CAPE for Convective Available Potential Energy) CAPE-85 closure time scale in the convection scheme (see also Lean et al. (2008)). The idea was 86 that in high-CAPE environments convective systems are more organized and exhibit a 87 larger spatial extent and that the model should be able to resolve these larger convec-88 tive structures, whereas smaller convection would still need to be parameterized. 89

Gerard and Geleyn (2005) proposed to use a prognostic closure in the representa-90 tion of subgrid convection involving prognostic updraught vertical velocities and a prog-91 nostic fractional area of a model grid box covered by convective clouds to address the 92 gray zone problem. In related earlier work Pan and Randall (1998) argued that the dis-93 tinction between convective and large-scale processes is ambiguous anyway. Arakawa et 94 al. (2011) and Arakawa and Wu (2013) framed the problem in a similar way, and var-95 ious subsequent studies followed an analogous approach (Grell & Freitas, 2014; Sakradz-96 ija et al., 2016; Zheng et al., 2016; Kwon & Hong, 2017; Su et al., 2021; W. Wang, 2022). 97 Moreover, Gerard et al. (2009) and Gerard (2015) highlighted the importance of the in-98 teraction between convection parameterization, microphysics, and the cloud scheme in 99 the context of the gray zone issue. 100

Specific modifications to the convection scheme such as a better coupling between subgrid convection and the resolved mesoscale circulation (Becker et al., 2021), or the representation of low-CAPE convection (McTaggart-Cowan et al., 2020) can also prove important and beneficial in convection-permitting simulations even if these aspects do not address the issue of scale-awareness of the convection parameterization directly from a conceptual perspective.

Convection-permitting regional models have been used with considerable success 107 for more than two decades (Prein et al., 2015; Clark et al., 2016; Kendon et al., 2021), 108 but the dynamics in regional models is constrained and influenced by the prescribed lat-109 eral boundary conditions (Kendon et al., 2010; Radermacher & Tomassini, 2012; Dipankar 110 et al., 2020). In the context of the current modelling system at the Met Office, the Met 111 Office Unified Model, the present study reports on a first thorough exploration of convection-112 permitting global modelling, besides the preliminary study Tomassini (2018). In this work 113 we focus on some of the main and most basic model configuration choices concerning con-114 vection, boundary layer processes, turbulence, and to some extent model dynamics. These 115 choices address the gray zone problem in one way or another and are part of what is some-116 times discussed under the headline of "gray zone model physics". In order to deal with 117

Configuration indicators	Description	Configuration details
GA7	Standard GA7.0 configuration	See Walters et al. (2019)
ShConv15/ShConv25	Scaling of the shallow convective mass flux	Shallow convective mass flux scaling: 0.015 or 0.025 Standard value is 0.03
MidLev	Revised formulation of the midlevel convection scheme and deep convection parameterization switched off	Deep convection parameterization switched off Fixed CAPE timescale of 2700 seconds Prognostic entrainment for midlevel convection Allow to initiate from the top of the boundary layer Time-smoothing of midlevel convective increments Adjustment of midlevel convective momentum transport setting Forced detrainment-weighted CAPE calculation Additional termination condition for convection
RAturb	Turbulence and boundary layer	Blending scheme as in RA1 (Bush et al., 2020), with fraction of maximum diffusion = 0.75 SHARPEST scheme for stable boundary layers Turbulent mixing based on interactively diagnosed turbulent layer depths
MoistCons	Moisture conservation option	Local moisture conservation scheme for advection as in RA1 (Bush et al., 2020)
ConvOff	General convection parameterisation option	No convection parameterisation used

Table 1. Description of different model configuration options. See also main body of text inSection 2 and additional information in Appendix A.

a tractable and well-defined problem, different options concerning the microphysics and
 cloud schemes have not been considered here.

The development of global convection-permitting models is more than just an in-120 cremental enhancement in resolution. Such models allow for, at least partly, resolving 121 fundamental new turbulent and convective phenomena in the atmosphere, studying in-122 teractions between convection and the atmospheric circulation across scales, and thus 123 addressing important and novel science questions (Tomassini, 2020; Senior et al., 2021; 124 Slingo et al., 2022; Tomassini & Yang, 2022). This step change in fidelity and realism 125 of weather and climate models implies also a step change in forecasting severe convec-126 tive storms and in supporting and informing climate change mitigation and adaptation 127 measures for the benefit of societies around the world (Slingo et al., 2021). In the pre-128 sented work, however, the focus is on developing an appropriate tool for these exciting 129 scientific endeavors, and on better understanding the strengths, limitations, and sensi-130 tivities of this tool. 131

The outline of the paper is as follows. In Section 2 the main model configuration 132 options that are studied in the present work are introduced and briefly described. The 133 results from simulations of different forecasts covering varying meteorological conditions 134 are contained in Section 3. The process-based analysis focusses on different locations, 135 times of the year, and particular phenomena. Section 4 highlights some important apects 136 of the convective gray zone problem and discusses a few additional sensitivity experiments. 137 Two whole months, July 2016 and January 2018, were simulated based on a particular 138 convection-permitting configuration with a view towards assessing statistics on climate 139 time scales, and the analysis is showcased in Section 5. Finally, in Section 6 the main 140 findings are summarized and an outlook on future planned work is given. 141

¹⁴² 2 Description of model configurations

In the present study the reference configuration is Global Atmosphere version 7.0 (Walters et al., 2019), *GA7* in short, as used in deterministic forecasts at the Met Office, i.e. with climatological aerosols, without the stochastic physics package, and using 70 levels in the vertical. The simulations are based on the Unified Model code version 11.7. The configurations described in this paper include some of the changes or specifications described in Table 1.

The main model configurations considered are a combination of the configuration options shown in Table 1, namely *GA7*, *MidLevShConv25*, *MidLevShConv25RAturb*, *MidLevShConv15RAturb*, and *ConvOffMoistConsRAturb*. The latter configuration, *Con*- vOffMoistConsRAturb, is chosen in such a way that it shares primary convection, dy namics, and turbulence model formulation features with some of the Unified Model Re gional Atmosphere configurations (Bush et al., 2020). A few important sensitivities and
 additional configuration tests will be discussed in Section 4.

All simulations are run at global N2560 (nominal 5km) resolution unless stated oth-156 erwise. N2560 refers to a latitude-longitude grid with 1920 regular latitude lines between 157 the pole and equator, and 5120 longitude points along each latitude line. The grid length 158 is 5km in latitude direction everywhere, in longitude direction the grid length is about 159 7.8km at the equator, 5km in the midlatitudes, and 4m near the poles. In the following 160 the term "5km-resolution" will be used for the sake of simplicity. Daily varying Oper-161 ational Sea Surface Temperature and Ice Analysis (OSTIA) sea surface temperatures and 162 sea ice are prescribed. The model time step is 90 seconds. Soil moisture is initialised from 163 Met Office Unified Model analysis. 164

In all configurations except GA7, the deep convection parameterization is switched 165 off. In the configurations that include the *MidLev* option the midlevel convection scheme 166 is allowed to start from the top of the boundary layer and not lower. The mixing in the 167 boundary layer is left to the boundary layer turbulence and shallow convection schemes. 168 The rationale is to more realistically represent nonequilibrium convection such as the di-169 urnal cycle of convection and the convection tied to advective boundary layers. The ap-170 proach combines a reduced subgrid convection approach based on a CAPE closure for 171 the free troposphere and a separate representation of boundary layer processes, acknowl-172 edging that free-tropospheric adjustment time scales can differ from boundary layer time 173 scales and that there often is an imbalance between boundary layer heating and deep 174 convective overturning (Bechtold et al., 2014). The choice of a CAPE closure time scale 175 of 2700 seconds is aimed at reducing the subgrid convective mass flux according to the 176 model resolution, allowing convection to be partly explicit, and at the same time to ac-177 count for the fact that vertical motion is not fully resolved by a model with 5km grid 178 spacing. Sensitivity experiments with regard to the CAPE closure time scale are pre-179 sented in Section 4.1. One should stress that even when the deep convection parame-180 terization is switched off in some of the configurations it is still possible for at least some 181 of the surface-based deep convective processes in the model to be handled by the sub-182 grid schemes. 183

The use of a prognostic entrainment rate (Willett & Whitall, 2017) takes into ac-184 count changing convective entrainment in different stages of the life cycle of convective 185 systems and varying degrees of convective organization. Convective increment time-smoothing 186 in addition improves the coupling between subgrid convection and the resolved circula-187 tion, reduces time intermittency of the convection scheme, and undesirable dynamical 188 effects of this intermittency such as spurious gravity waves. One particular advantage 189 of unifying the subgrid convection treatment in the free troposphere is the avoidance of 190 a late afternoon "tea break" in convective rainfall, especially over tropical land. In pre-191 vious model versions an erroneous minimum in modelled convective rainfall featured in 192 the evening when mainly surface-based convection handled by the deep convection scheme 193 was passed over to the midlevel scheme which treated elevated convection. 194

In convection-permitting Unified Model simulations, single grid-column updraughts 195 can become unrealistically intense and persistent because a stagnation point forms at the base of the updraught. Equal convergent velocities at either side of an updraught 197 column mean the SL advection scheme's back-trajectories from the cell centre go straight 198 down, and so fail to advect in what will typically be drier air from neighbouring columns. 199 200 This allows single-point updraught columns to keep creating their own moisture by copying a near-surface moist anomaly upwards. The spuriously created moisture feeds con-201 densation and latent heating, generating stronger ascent and convergence, leading to a 202 positive feedback. To address the problem, all simulations, except GA7 and ConvOff-203 MoistConsRAturb, use the "fountain buster", an implementation of a modification to the 204

semi-Lagrangian (SL) advection scheme, aimed at making a local correction to the lack 205 of conservation. The scheme is called directly after the SL advection increment has been 206 calculated, by interpolation of the virtual potential temperature and moisture variables 207 to departure points. It works by identifying grid points where the horizontal winds (on the grid cell faces) are converging, and adds onto the standard SL increment a simple 209 linear up-wind advection increment arising from just the locally convergent part of the 210 flow. In this way it is adding in the effects of just the convergent inflow that will have 211 been missed by the SL advection (because SL advection interpolates the cell-face winds 212 to the cell centre in order to find the departure point). 213

The label *ShConv* refers to a multiplicative scaling of the shallow convective sub-214 grid mass flux which is included in its default setting also in the reference configuration 215 GA7 and is used in some of the convection-permitting configurations to limit the influ-216 ence of the subgrid shallow convection scheme. The configuration option *RAturb* includes 217 the turbulence blending scheme (Boutle et al., 2014) which blends the one-dimensional 218 turbulence parameterization as used in lower-resolution versions of the global model with 219 a three-dimensional Smagorinsky-Lilly representation of subgrid turbulence as typically 220 applied in large-eddy simulation models. The two schemes are combined via the mix-221 ing length and the key parameter is the ratio of grid length to boundary layer depth. The 222 SHARPEST scheme for stable boundary layers as described in Derbyshire (1997) and 223 used in the Regional Atmosphere configuration (Bush et al., 2020) is employed. The frac-224 tion of the maximum allowed value of the diffusion coefficient is set to 0.75 (Hanley et 225 al., 2015). Moreover, the free atmospheric turbulent mixing in *RAturb* is based on in-226 teractively diagnosed turbulent layer depths throughout the atmospheric column (as in 227 the tropical version of the Regional Atmosphere configuration described in Bush et al. 228 (2020)).229

All simulations, except *GA7* and *ConvOffMoistConsRAturb*, use the parameter choice puns=1.0 in the non-linearity setting for the boundary layer solver in the case of unstable boundary layers which makes the implicit solver more stable (Wood et al., 2007, see Appendix A). All configurations use the multigrid solver (Maynard et al., 2020).

For Met Office Global Model Evaluation and Development (GMED) tickets associated with the different changes and some more information on the tickets see the Appendix A.

²³⁷ **3** Case studies and results

A case study approach is taken in which testbed cases are defined and the model 238 is evaluated against observations and reanalysis data. Five model configurations are fur-239 ther investigated in greater detail: GA7, MidLevShConv25, MidLevShConv25RAturb, Mi-240 dLevShConv15RAturb, and ConvOffMoistConsRAturb. The focus of the study is on model 241 forecasts with lead times of up to 10 days because this way the simulations can be di-242 rectly compared to observations. Note that some errors in the forecast will come from 243 deficiencies in the initial conditions or a lack of predictability not captured due to run-244 ning only one forecast per configuration per initialization time. The different cases aim 245 at covering different, important meteorological conditions and phenomena in different 246 parts of the world and at various times during the year with a focus on high-impact weather. 247

248 249

3.1 African easterly wave case and convection-circulation coupling over Africa

An African easterly wave disturbance is detectable starting from 18:00 UTC on July
7, 2010, over North Africa (Tomassini et al., 2017; Tomassini, 2018). The dynamics of
the wave is rather weak over the first 30 hours after detection, i.e. until about 00:00 UTC
on July 9. Starting around July 9 03:00 UTC a crucial strengthening phase of the wave

occurs, which lasts for about 2 days. Tropical Rainfall Measuring Mission (TRMM) pre-254 cipitation (Huffman et al., 2007) shows distinct organized precipitation ahead of the trough 255 at around 12 to 18 degrees North where the main center of the wave disturbance is lo-256 cated (Figure 1, panel a). The case is of interest because it allows for assessing the cou-257 pling between moist convection and atmospheric circulations over tropical land in the 258 different model configurations. The simulations are initialized on July 7, 2010, 00:00 UTC 259 and on July 11, 2010, 00:00 UTC with European Centre for Medium-Range Weather Fore-260 casts (ECMWF) operational analyses, and two 5-day forecasts are performed. 261

262 $GA\gamma$ is not able to develop and predict the mesoscale convective systems associated with the wave in the early, developing stage (Figure 1, panel b), and also in Con-263 vOffMoistConsRAturb the predictive skill related to the convective systems at and ahead 264 of the wave trough, and the imprint of the diurnal cycle signal, are not satisfactory (Fig-265 ure 1, panel f). In the *MidLevShConv* configurations (panels c, d, and e) the coupling 266 between convection and the wave dynamics is well represented, and moist convection sup-267 ports the dynamical development of the wave (not shown). However, in *MidLevShConv25* 268 the speeds of the convective systems around the wave trough seem to be slightly too slow 269 in the first forecast (Figure 1, panel c), whereas this feature is better represented in Mi-270 dLevShConv25RAturb and MidLevShConv15RAturb (Figure 1, panel d and e, respectively). 271

These characteristics are also confirmed in observed and simulated satellite imagery (Figures 2 and 3), showing reflectance based on the visible channel at $0.8\mu m$ on July 10, 12:00 UTC (Figure 2), and brightness temperature based on the IR channel at $10.8\mu m$ on July 10, 18:00 UTC (Figure 3), i.e., almost 4 days after forecast initialization. The dashed red lines in the plots indicate the approximate position of the wave trough. The satellite simulated fields from the global model simulations are derived using an offline radiative transfer code (Saunders et al., 2018).

In $GA\gamma$ the convective systems over Africa are not represented realistically (Fig-279 ures 2 and 3, top right panel), although it should be mentioned that the satellite sim-280 ulator does not consider cloud water or ice that is held within the convection scheme. 281 MidLevShConv25RAturb and MidLevShConv15RAturb develop well organized and re-282 alistic mesoscale convective systems at the wave trough and over the African continent. 283 The same could be said of *MidLevShConv25*, however, as in the Hovmöller diagrams of 284 Figure 1, in relation to the wave the main convective systems are located behind the trough 285 (Figures 2 and 3, left panel in the middle row), suggesting that they move somewhat too 286 slowly. For *ConvOffMoistConsRAturb* (Figures 2 and 3, bottom right panel), the con-287 vective systems are too scattered and widespread across the land area, a feature that is 288 evident even more clearly in the precipitation field (not shown). Moreover, clouds and 289 rainfall are not very sensitive to the presence of the wave and do not show a distinct re-290 sponse to the wave circulation around the trough in this configuration. 291

In order to better understand the representation of convective systems over the wider 292 Africa-Atlantic region, convective storms are tracked in the area 40W-40E and 25S-25N 293 over the period from July 8 to July 15, 2010, both for the different model configurations 294 and Global Precipitation Measurement (GPM) observations (Huffman et al., 2019). The 295 tracking is based on half-hourly precipitation data, a threshold of 3 mm/hour is chosen 296 to identify convective systems, and an overlap of 50% is required from time stamp to time 297 stamp for propagating systems (Stein et al., 2014; Crook et al., 2019). The histogram 298 of convective storms lifetimes (Figure 4a) indicates that there are overall too many con-299 vective storms in all considered model configurations compared to GPM. GA7 overes-300 timates the number of short-lived storms and underestimates the number of long-lived 301 302 storms. For the other considered, convection-permitting configurations, the long-lived convective storms tend to be too persistent. For the mean size (over the storm life cy-303 cle) all models exhibit too many small storms, and too few very large storms compared 304 to GPM (Figure 4b). This holds true for $GA\gamma$ in particular. For the number of medium-305 sized storms it is hard to make a robust statement. The distributions of mean speeds 306

(over the storm life cycle) suggest that the storms in GA7 move too fast compared to 307 GPM. In the convection-permitting configurations the storms tend to move too slowly, 308 most distinctly so in ConvOffMoistConsRAturb. The distributions confirm the visual im-309 pression obtained from the Hovmöller plots (Figure 1) that the fraction of slower-moving 310 storms, as part of the total number of storms, is slightly larger in *MidLevShConv25* than 311 in MidLevShConv25RAturb. It is to be expected that the characteristics of storms in GPM 312 will have biases over the region, particularly with regard to their areas, and the results 313 should be interpreted with some caution. 314

315 The diurnal cycle of precipitation over the Sahel depends on the exact region considered, and it is not possible to assess the issue very robustly based on just two 5-day 316 forecasts. Nevertheless, choosing the region 1E to 11E and 5N to 10N over the period 317 from July 7 to July 15, 2010, the picture that emerges is plausible (Figure 5). Accord-318 ing to the analysis GA7 shows an erroneous, well-known peak in precipitation around 319 local noon and too little rainfall during night time. The convection-permitting model con-320 figurations fare much better and exhibit an afternoon maximum in agreement with TRMM. 321 In the *MidLev* configurations the rainfall persists somewhat too strongly during the evening, 322 whereas in *ConvOffMoistConsRAturb* rainfall is underestimated during the night. Note 323 that mean rainfall over the course of the day is not generally overpredicted in the model 324 simulations over the area. 325

The latter observation is confirmed also when it comes to the simulation of heavy 326 rainfall over Africa. Box plots for both TRMM observations and model configurations 327 of 3-hourly mean rainfall rates above the threshold of 30mm/hr over the region 20W to 328 45E and 35S to 35N, comprising the whole of the African continent, and the period July 329 7 to July 15, 2010, reveal that the convection-permitting configurations do not overes-330 timate heavy rainfall events over Africa on the 3-hourly time scale (Figure 6). On the 331 contrary, they tend to underestimate extremes compared to TRMM. The box plots show 332 the lower, middle, and upper quartile of the data above the threshold, and the length 333 of the whiskers are 1.5 times the interquartile range. $GA\gamma$ does not have any data points 334 above the threshold when considering 3-hourly mean rainfall rates. 335

Overall the model configurations *MidLevShConv25RAturb* and *MidLevShConv15RAturb* show the best convection-circulation coupling, speed of convective systems, and degree of convective organisation over the African region in the analyses presented in this section. Impressively, the respective simulations can predict the evolution of some individual convective systems several days ahead. And it is these convective systems that impact the livelihoods of people in the area most.

342

3.2 Hurricane Dorian and Typhoon Goni

The simulation and prediction of tropical cyclones are important applications of 343 weather forecasting because of the devastation these phenomena can cause. Furthermore, 344 changes in tropical cyclones under a warming climate may deliver some of the most sig-345 nificant impacts. Here we present results of forecasts of two tropical cyclones, hurricane 346 Dorian initialized on August 30, 2019, 12:00 UTC, and typhoon Goni initialized on Oc-347 tober 28, 2020, at 12:00 UTC. The runs use ECMWF operational analyses as initial con-348 ditions except the operational model shown in Figures 7 and 8. The then operational 349 model is GA6.1 (Walters et al., 2017) in the case of Dorian, and GA7.2 in the case of 350 Goni, both run at 10km resolution and using Met Office analyses as initial conditions. 351

The predictions of central pressure in the simulations of tropical cyclones Dorian and Goni are significantly improved in the *MidLev* model configurations compared to the operational models, *GA7* at 5km resolution, and also *ConvOffMoistConsRAturb* (Figure 7). In particular, spells of rapid intensifications for both tropical cyclones are well captured by the *MidLev* model configurations. The simulations of tropical cyclone tracks, however, are quite consistently degraded in the 5km-resolution simulations compared to the operational models (Figure 8 in the case of Goni, but also in results that are not shown here). This is true, though, also for GA7 at 5km resolution. We do not expect this to be due to the different initial conditions. The reason is not clear at this point and might be related to subtle degradations in regional synoptic scale conditions that affect the steering flow rather than the representation of the local cyclone processes themselves.

364

3.3 EUREC⁴A case over the Atlantic trades

One reason why tropical cyclones tracks are not particularly well simulated by the 365 5km-resolution models might be related to the structure of the atmosphere in the sub-366 tropics which influences the steering flow away from the cyclone center. To get a sense 367 of this issue and to investigate shallow cumulus clouds over the trade wind region, a topic 368 of particular importance in climate change studies, a case during the EUREC⁴A field 369 campaign is considered (Stevens et al., 2021). The models are initialized on February 370 1, 2020, at 00:00 UTC with ECMWF ERA5 reanalysis (Hersbach et al., 2020), and eval-371 uated on February 2. 372

The GOES16 satellite observed and simulated reflectance demonstrates how trade 373 wind cumulus clouds can be controlled in the model by scaling the shallow convective 374 subgrid mass flux (Figure 9). A reduced subgrid mass flux leads to more, and brighter, 375 shallow clouds. A difficulty here is that, as also known from regional models when sim-376 ulating low clouds explicity, the structure and organization of the clouds becomes more 377 realistic but the cloud amount or the optical thickness of the clouds is often overestimated. 378 In the particular snapshot shown in Figure 9 low cloud cover is distinctly underestimated 379 in the GA7 configuration, especially over the Pacific (Figure 9, top right panel). How-380 ever, lower-resolution versions of the GA7 model have been tuned to produce realistic 381 global-mean fluxes, whereas in radiation budget estimates presented in Section 5 MidLevSh-382 *Conv25RAturb* exhibits a distinctly overpredicted global-mean top-of-the-atmosphere out-383 going shortwave radiative flux in the absence of tuning. This shows how seemingly more 384 realistic process representation does not always translate into better agreement with ob-385 servations, at least not in every respect. 386

Observed profiles from HALO research aircraft dropsondes sampled over the EUREC⁴A 387 campaign observation area east of Barbados over a time window of 17:50 to 18:10 UTC 388 (George et al., 2021) show that the increased trade wind cloudiness in the MidLevSh-389 Conv15RAturb configuration is associated with a cooler and moister lower troposphere 390 (Figure 10). The overall structure of the lower troposphere is quite well represented, in-391 cluding by GA7. The slight cold bias in the subtropical trade wind cloud layer is con-392 sistent with anecdotal evidence from other forecasts and might at least partially be due 303 to the somewhat overestimated cloud amount or optical thickness in the lower troposphere 394 of the subtropics. The most notable feature is a dry bias in the subcloud layer of the mod-395 els which is present in all configurations despite the different turbulence schemes used. 396 The moisture profile in the cloud layer is probably most realistic in MidLevShConv15RAturb 397 while $GA\gamma$ tends to be too dry. The case shows quite clearly how none of the configu-398 rations is perfect, and a subjective, informed judgment has to be made when assessing 399 the different model configurations with regard to various aspects and features of the at-400 mosphere. 401

402

3.4 Darwin mesoscale convective systems case

To conclude the analysis of tropical and subtropical cases, advantage is taken of the C-band radar at Darwin on the northern tip of Australia. The models are initialized on February 17, 2014, at 00:00 UTC with ECMWF operational analysis (Franklin et al., 2016).

All models show substantial deficiencies in the profiles of radar reflectivities (Fig-407 ure 11). As with the simulated satellite imagery previously, the simulated radar reflec-408 tivity does not include cloud liquid and ice that are held within the convection scheme, 409 somewhat penalizing the model with parameterized convection $(GA\gamma)$. Most model con-410 figurations tend to underestimate radar reflectivity at upper levels. The configuration 411 with convection switched off, ConvOffMoistConsRAturb, reaches highest, but seems to 412 overestimate high radar relectivities at middle levels. All models simulate an unrealis-413 tically strong outflow around the freezing level. The shortcomings of the models will at 414 least partly be related to the representation of microphysics, an aspect that has not been 415 considered in detail in the present study and could in principle be tuned for a better agree-416 ment with observations. The comparison makes the high-dimensional nature of the model 417 development problem obvious, and the extent of the challenge that is due to the fact that 418 only a limited number of sensitivity experiments and cases can be run and assessed with 419 a high-resolution global model. The profiles in Figure 11 show an area mean over 12 hours, 420 so part of the errors may also be caused by inaccuracies in the exact location of the con-421 vective systems and not only by their structure. 422

Despite the challenges in the assessment, there are robust features across the dif-423 ferent tropical cases, for example the structure of precipitation, here shown around Dar-424 win (Figure 12). Our judgement is that the *MidLev* configurations overall exhibit the 425 most realistic characteristics in terms of structure and organization (Figure 12d). The 426 GPM product likely overestimates the extent of the rainfall areas of the largest clusters. 427 And the model *ConvOffMoistConsRAturb* exhibits a too scattered and blobby rainfall 428 field and misses lighter rainfall. GA7 shows too widespread areas of light rainfall and 429 very little precipitation over land. Similar conlusions can be drawn also when investi-430 gating other cases like the African easterly wave case (Section 3.1) over tropical land (not 431 shown). 432

433

3.5 NAWDEX cases: the Atlantic extratropics

To understand the behavior of the convection-permitting model configurations in the midlatitudes, cases from the North Atlantic Waveguide and Downstream Impact Experiment (NAWDEX) field campaign (Schäfler et al., 2018) are examined. The results from two sets of forecasts are presented here. For one set the models are initialized on September 22, 2016, 00:00 UTC with ECMWF operational analysis, for the other set the initialization time is October 5, 2016, 00:00 UTC, again using ECMWF operational analysis.

The ex-tropical storm Karl reintensified on September 26, 2016, as it approached western Europe. In the subsequent development the jet stream was unusually strong on its southern flank, forming a jet streak that propagated ahead from Karl. The impact on European weather occured through the formation of a new cyclone, Walpurga. Moistureladen air was drawn around the Atlantic subtropical high. A particularly moist boundary layer was observed in this atmospheric rivertype flow that extended to Norway, where it caused heavy, persistent rainfall (Schäfler et al., 2018).

A similar moisture transport pattern was also involved in the second case considered. On October 5, 2016, there was a high-pressure block over Europe. Two days later midlatitude cyclone Sanchez developed over the Atlantic. Nine days after forecast initialisation, on October 13, Sanchez caused heavy rainfall over southeastern France (Schäfler et al., 2018).

⁴⁵³ One could argue that the 5-day rainfall forcast for the September case is most ac-⁴⁵⁴ curate in the GA7 configuration (Figure 13). Although the front-like structure in rain-⁴⁵⁵ fall moved slightly too fast over northern Europe in GA7, the heavy rainfall over Nor-⁴⁵⁶ way is well reproduced. In the convection-permitting MidLevShConv25RAturb config-⁴⁵⁷ uration, the main rainband has moved slightly too slowly, although some of the rainfall reaches the west coast of Norway. *ConvOffMoistConsRAturb* shows more deficiencies and places the main precipitation area further inland.

Both GA7 and MidLevShConv25RAturb reproduce the river-like moisture band from the subtropics towards Scandinavia in the daily-mean vertically integrated water vapor fields (Figure 14, left column for the September case). MidLevShConv25RAturb shows a stronger gradient in moisture in the northern part of Great Britain which could be responsible for the rainfall in this region. The differences in rainfall over the Norwegian coast seem to be related also to differences in the dynamical fields where GA7 exhibits a stronger bend in the upper-level potential vorticity (PV) in the area (not shown).

In the October case, GA7 performs distinctly worse than MidLevShConv25RAturb 467 in terms of rainfall prediction 9 days ahead after initialization (Figure 15). The rainfall 468 field on October 13, when the heavy rainfall event over southern France ocurred (red cir-469 cles in the panels of Figure 15), shows that $GA\gamma$ moves the rainfall too quickly over the 470 continent. Only the *MidLev* configurations predict the location of the heavy rainfall area 471 correctly. Comparing the integrated water vapor fields on October 10 (Figure 14 right 472 column) suggests that *MidLevShConv25RAturb* manages to draw the moisture further 473 north. However, it is not obvious whether this difference is key in this case. 474

It is interesting to examine the PV pattern at the 320K isentrope in the different 475 configurations (Figure 16). It would be difficult to derive the differences in rainfall pre-476 diction from the differences in the upper-level PV field. What is striking is how differ-477 ent the upper-level PV fields are in MidLevShConv25RAturb compared to MidLevShConv25 478 (Figure 16, panels c and d). The two configurations differ only in the representation of 479 turbulence, suggesting that the subgrid turbulence scheme can have a strong impact on 480 upper-level dynamics. However, as noted before, the predictions of rainfall over Europe 481 disagree less, although there are some differences. 482

Based on the two presented cases it is possible that the convection-permitting con-483 figurations tend to move the fronts related to extratropical cyclones somewhat too slowly 484 compared to observations, perhaps due to increased cumulus friction or differences in merid-485 ional geopotential height gradients resulting in changes in geostrophic jet strength. How-486 ever, this does not necessarily translate into worse performing rainfall forecasts compared to the model with parameterized convection. The key features in large-scale dynamical 488 fields that lead to better rainfall prediction can be difficult to identify, and over the whole 489 of the Atlantic area the differences in these fields are typically large in individual fore-490 casts. This poses significant challenges in the assessment of the performance of the var-491 ious configurations over the midlatitudes which will require a much more extensive set 492 of statistics, and the results of such an assessment will depend on the considered quan-493 tity and the specific application.

⁴⁹⁵ 4 Important sensitivities and additional configuration tests

A few additional sensitivity tests regarding model configuration options are briefly described here. They are generally based on a rather limited amount of evidence, but some of them concern important aspects directly related to the convective gray zone problem and are not incidental. Not all sensitivity experiments that were conducted are presented and discussed, though, in this section.

501 4.1 CAPE closure time scale

The CAPE closure time scale of the convection scheme in the *MidLev* configurations is a key parameter (see Table 1) that controls the amount of subgrid convective mass flux relative to the resolved, explicit vertical motion and related rainfall in the model. Figure 17 shows the rainfall Hovmöller plots of total rainfall and convective rainfall for the configuration *MidLevShConv25* in the African easterly wave case discussed in Section 3.1 using three different CAPE closure time scales: 3000s, 2700s, and 2400s. One can see how the amount of convective rainfall increases when reducing the CAPE closure time scale. When the CAPE closure time scale is set to 2400s, the precipitation around the wave trough starts to break up in the early phase of the wave, particularly between July 9 and July 10, because the coupling between convection and wave circulation weakens. This trend continues when further shortening the CAPE closure time scale (not shown).

The effect of changing the CAPE closure time scale can also be identified in global 513 mean values of total rainfall versus convective rainfall. For the MidLevShConv25 con-514 figuration and a mean over the days July 12 to July 15 (the second forecast of the African 515 easterly wave case) the global mean of total precipitation is 3.689, 3.666, and 3.638 mm/day 516 for the simulations with CAPE time scales 3000s, 2700s, and 2400s, respectively. In other 517 words, there is relatively little change in gloabl mean total rainfall. The global mean con-518 vective rainfall, however, increases from 0.909 to 1.032 to 1.178 mm/day, implying that 519 the percentage of convective rainfall as part of the total rainfall increases from 24.7% to 520 28.1% to 32.5% when reducing the CAPE closure time scale. 521

522

4.2 Initial perturbations in the convection scheme

It is common in convection schemes to use initial perturbations in order to repre-523 sent subgrid variability, where some regions will be buoyant enough to trigger convec-524 tion not achievable with the grid-box mean profiles, and this is also the case for the con-525 vection scheme used in the *MidLev* configurations. The original and default option is to 526 apply perturbations in the temperature field, but in the deep convection scheme used 527 in GA7 humidity perturbations are added. Whether temperature or humidity pertur-528 bations are used in the *MidLev* configurations turns out to have quite a substantial ef-529 fect. Using humidity perturbations makes the *MidLev* simulations look somewhat more 530 similar to GA7, the model with fully parameterized convection (Figure 18). The buoy-531 ancy perturbation is fixed when shifting from temperature to humidity perturbations, 532 but the perturbation in terms of moist static energy becomes somewhat larger, a circum-533 stance that could also play a role. 534

In some of the diagnostics used in the African easterly wave case (Section 3.1), for example, one can see that the coupling between rainfall and wave becomes weaker in the early phase of the wave when applying humidity perturbations, and the clouds have a somewhat more blurred appearance compared to the simulations using temperature perturbations (Figure 18). A similar impression, namely that the simulations tend to become a bit more similar to GA7 when using humidity perturbations, results also from the extratropical NAWDEX cases (Section 3.5, not shown).

542

4.3 Sensitivity to turbulence formulation and miscellanea

Given that in Section 3.5 it was shown that the turbulence scheme has a substan-543 tial influence on the large-scale dynamical fields in the troposphere, two more turbulence 544 options were tested and investigated for the NAWDEX cases on top of the *MidLevSh*-545 Conv25RAturb configuration. In one test, instead of relaxing the mixing length towards 546 the Smagorinsky length scale away from turbulent layers, a background mixing length 547 scale of 40m is used if the layer is stable or weakly turbulent. In another test a differ-548 ent stability function is used for unstable Richardson numbers, which produces signif-549 icantly more mixing and which has previously been employed in the tropical version of 550 the Unified Model regional configuration (Bush et al., 2020). In both tests the rainfall 551 forecasts look somewhat degraded. For example, in the October case (see Section 3.5), 552 the rainfall is placed mainly over the Alps in the first test, and mainly east of the Alps 553 in the second test on day nine of the forecast and therefore showing similar deficiencies 554 as GA7 (not shown). 555

A few additional targeted, but limited, sensitivity tests have been conducted also 556 with regard to the fountain buster, the time-smoothing of midlevel convective increments, 557 the convective momentum transport settings, and different moisture conservation schemes 558 (not shown). For example, the impact of the convective increment smoothing and the convective momentum transport settings on the speed of convective systems was assessed 560 in the African easterly wave case, but the impact was found to be small. The use of the 561 fountain buster was judged to be overall beneficial. And the effect of applying different 562 available local moisture conservation schemes was concluded to be minor on the consid-563 ered time scales. 564

565 5 Months of July 2016 and January 2018

To estimate features that are of importance in climate studies, in particular the top-566 of-the-atmosphere energy budget, two full months, July 2016 and January 2018, are cov-567 ered by 5-day forecasts with one overlapping day between forecasts using the convectionpermitting MidLevShConv25RAturb configuration. The first 24 hours of each forecast 569 are not included in the analysis. The first initial times are June 30, 2016, 00:00 UTC. 570 and December 31, 2017, 00:00 UTC, respectively. Due to a model failure, all forecasts 571 that cover January 2018 were run with a so-called polar cap for advection in which the 572 semi-Lagrangian advection is replaced by an interpolation in an area close to the poles 573 (see Appendix A). Low-resolution tests have shown that on time scales of days the im-574 pact of the polar cap scheme on the quality of forecasts is neutral. 575

576

5.1 Mean rainfall and top-of-the-atmosphere radiative fluxes

Figures 19 and 20 show mean top-of-the-atmosphere outgoing shortwave (top row) 577 and longwave (middle row) radiation for both months alongside the corresponding CERES 578 EBAF 4.1 observations (Loeb et al., 2018). Moreover, monthly mean precipitation is de-579 picted for the model simulations and GPM rainfall observations (bottom rows). Over-580 all, the clouds look very realistic in the simulations and it would be difficult to distin-581 guish between simulations and observations by eye. There are, however, some quite dis-582 tinct biases in the precipitation fields, particularly over the tropical oceans. Rather strong 583 rainfall biases can be identified over the tropical Pacific and Indian Ocean during boreal 584 winter (Figure 20) and the southern branch of the ITCZ in the western tropical Pacific 585 during boreal summer (Figure 19). Over tropical land and over the extratropics the bi-586 ases are much less pronounced. Similar tropical rainfall biases have also been described 587 in other convection-permitting global atmosphere models (Caldwell et al., 2021; Hoheneg-588 ger et al., 2022). 589

In terms of the global-mean top-of-the-atmosphere radiation budget, the outgoing longwave radiation (OLR) is fairly close to observations without any tuning, while the model is too reflective when it comes to the top-of-the-atmosphere outgoing shortwave radiative fluxes, especially during boreal summer. The results are summarized in Table 2. Some cloud tuning would be required before coupling the *MidLevShConv25RAturb* configuration to an ocean model.

596 5.2 Madden-Julian Oscillation

A strong Madden-Julian Oscillation (MJO) passed from the Indian Ocean over the Maritime Continent to the West Pacific during the month of January 2018 (Figure 21). The *MidLevShConv25RAturb* model reproduces the propagation of the rainfall very well and the 3-hourly mean rainfall rates are in good agreement with GPM observations. The then-operational GA6.1 global model with parameterized convection shows much weaker eastward propagation with a significant reduction of rainfall when it reaches the Maritime Continent land which is often referred to as the barrier effect (Zhang & Ling, 2017).

	TOA outgoing longwave flux	TOA outgoing shortwave flux
Observed CERES EBAF 4.1 July 2016	244.51 [W/m ²]	92.96 [W/m ²]
Modeled MidLevShConv25RAturb July 2016	243.92 [W/m ²]	97.67 [W/m ²]
Observed CERES EBAF 4.1 January 2018	$237.30 [W/m^2]$	$105.03 [W/m^2]$

 $237.59 \ [W/m^2]$

 $106.47 \ [W/m^2]$

Table 2. Observed and modeled global-mean top-of-the-atmosphere (TOA) radiative fluxes forJuly 2016 and January 2018.

The propagation of the MJO does not appear to suffer from the barrier effect over the Maritime Continent in the convection permitting simulations and its regions of impacts are fairly well captured. Figure 21 also shows evidence of faster eastward Kelvin-like disturbances and westward Rossby or Mixed Rossby-Gravity wave features that are better captured in the convection-permitting model configuration MidLevShConvRAturb compared to the parameterized convection model. The result shows that the convection-permitting model configuration MidLevShConv25RAturb is able to accurately predict MJO-related high impact weather in the region.

5.3 The West Pacific subtropical high

612

Modeled MidLevShConv25RAturb January 2018

The West Pacific subtropical high (WPSH), a distinct anticyclone in the middle 613 and lower troposphere over the western North Pacific, is a key component of the East 614 Asian summer monsoon system. It affects the regional hydrological cycle and has a sig-615 nificant influence on tropical cyclone activity in the Western North Pacific (Yihui & Chan, 616 2005; B. Wang et al., 2013). Correct simulations of the WPSH are important for sea-617 sonal forecasting and climate projections. In low-resolution versions of the Unified Model, 618 as in many other models, the mean-state summer WPSH is too weak and located too 619 far east (Rodríguez et al., 2017). These biases lead to an underestimation of the south-620 westerly monsoon flow over East Asia, affecting its representation of the seasonal wa-621 ter cycle in the area. Studies of systematic errors in climate models have shown that they 622 can develop in the first few days of a simulation, and then persist to climate time scales 623 (Martin et al., 2021). We can therefore gain an understanding of the emergence of such 624 biases by analyzing errors in initialized hindcasts (Rodríguez & Milton, 2019). 625

Here, a comparison of the development of circulation biases associated with the WPSH 626 in global 5km-resolution MidLevShConv25RAturb configuration simulations and low-resolution 627 N216 (about 80km-resolution) GA6 configuration forecasts initialised from Met Office 628 analyses is made for July 2016 (Figure 22). The 850hPa geopotential height fields at 5 629 days lead time indicate that in the 5km-resolution MidLevShConv25RAturb simulations 630 the WPSH is developed more strongly and extends further to the west and south, in bet-631 ter agreement with ERA5 compared to the low-resolution GA6 forecasts (Figure 22, pan-632 els a-c). This is confirmed by comparing three indices, I_{in} , I_W and I_N in ERA5 and the 633 model forecasts at a lead time of 5 days, following Chen et al. (2010) and Lu and Dong 634 (2001). The indices represent, respectively, the intensity, westward extension and north-635 ern edge of the WPSH (Figure 22, panel d). The 5km-resolution MidLevShConv25RAturb 636 forecasts show an improvement in the representation of the WPSH mainly in terms of 637 the intensity and the westward extension, important aspects in the context of the regional 638 weather and climate. However, biases in the region of the Kamchatka Peninsula increase. 639

5.4 Extratropical cyclones

Extratropical cyclones are tracked in the *MidLevShConv25RAturb* configuration simulations for both months, July 2016 and January 2018, using the TRACK (Hodges, 1995) algorithm on the 850hPa relative vorticity field. The same tracking algorithm is also applied to ECMWF analyses and the then-operational Met Office Unified model for
both July 2016 and January 2018. The operational model used by the Met Office in July
2016 and January 2018 was the GA6 configuration with fully parameterized convection
(Walters et al., 2017), at N768 resolution for July 2016 and at N1280 resolution for January 2018.

Cyclones identified in the *MidLevShConv25RAturb* simulation are matched (loca-649 tion and intensity) against those in ECMWF analyses using the method of Froude et al. 650 (2007a, 2007b). Then, for comparison, the matching is re-run on the then-operational 651 GA6.1 forecast data and ECMWF analyses. The mean error in extratropical cyclone lo-652 cation (top row) and intensity (bottom row) for MidLevShConv25RAturb (black lines) 653 and GA6 (red lines) relative to ECMWF analyses, are shown for the NH winter (Fig-654 ures 23a and b) and SH winter (Figures 23c and d). The performance of the MidLevSh-655 Conv25RAturb simulation is almost identical to that of both resolutions of the GA6 op-656 erational model, relative to ECWMF analyses (Figure 23). Therefore, in the context of 657 extratropical cyclone mean location and intensity, there is no improvement from running 658 the MidLevShConv25RAturb configuration at 5km grid spacing compared to the then-659 operational model at 10km resolution, but also no degradation. Nevertheless, the struc-660 ture of individual cyclones (e.g. fronts) may be simulated better at 5km resolution and 661 will be an area of future work. 662

663 6 Summary and conclusions

In the present work the challenges of the convective gray zone are discussed in the 664 context of a 5km-resolution global atmospheric model, and a few fundamental ideas con-665 cerning the representation of turbulence and convection are proposed and assessed. Dif-666 ferent questions regarding the formulation of subgrid turbulence, convection, and model 667 dynamics are considered: should the convection parameterization be switched off at a 668 particular grid length in convection-permitting models or is a more seamless approach 669 to be preferred? Are one-dimensional turbulence parameterizations sufficient in kilometer-670 scale models or is a three-dimensional representation of turbulence beneficial, or a com-671 promise between the two? What are the impacts of non-conservation characteristics of 672 the advection scheme and physical parameterizations? Most of these questions are not 673 in themselves binary in nature (convection parameterization switched on or off). On the 674 contrary, as the model resolution is increased continuously, the resolved convective and 675 turbulent fluxes are supposed to increase continuously and the subgrid, parameterized fluxes 676 should decrease continously. 677

The investigation of the aforementioned questions always takes place in a certain 678 context, with regard to a particular chosen reference configuration and taking specific 679 representations of subgrid processes into account. For example, if the considered repre-680 sentation of subgrid convection is distinctly deficient, then in practice it might be ad-681 visable to switch off such a subgrid convection scheme in kilometer-scale models com-682 pletely, even though in theory convection is not well resolved. Moreover, it is clear that 683 the particular application of the model and the quantities of interest are an important 684 factor in the assessment. There is not one model configuration which performs better 685 than any other configuration in all meteorological situations and geographical locations. The particular metrics used in the evaluation will always influence the conclusions. In 687 numerical weather prediction, for instance, an issue is how to weight more traditional 688 error indices that target large-scale fields in the middle troposphere against more impact-689 based, near-surface measures. 690

Here a case study approach is taken in which selected test cases are compared to observations. A robust conclusion of the assessment in the present study, conditional on the assumed framework, is that the convection-permitting configurations which include scale-aware turbulence and a carefully reduced and simplified mass-flux convection scheme

outperform both the model with fully parameterized convection as well as a configura-695 tion in which the subgrid convection parameterization is switched off completely with 696 regard to the organization of convection and related features of tropical variability. The 697 key idea with regard to the representation of subgrid convection in the convective gray zone is to use a unified scheme in the free troposphere which is allowed to trigger from 699 the top of the boundary layer and not below, and to limit the subgrid convective mass 700 flux, both in the free troposphere as well as the shallow-convective mass flux in the bound-701 ary layer. The assessment in the extratropics is more ambiguous and challenging than 702 in the tropics, and better statistics are needed. The configuration with fully parameter-703 ized convection is more competitive in the midlatitudes, whereas for tropical variabil-704 ity it shows distinct deficiencies. The results are in broad agreement with the study by 705 Judt and Rios-Berrios (2021). Moreover, as already mentioned, the conclusion might dif-706 fer depending on the considered variable and metric, for example rainfall or upper-tropospheric 707 wind and temperature. Also the comparison of longer-term mean statistics can be less 708 straightforward because the model with parameterized convection is rather resolution 709 insensitive and can therefore be tuned using low-resolution versions of the model. This 710 does not apply the same way to the convection-permitting configurations, especially given 711 that physical parameterizations are generally not scale-aware. Nevertheless, for the convection-712 permitting configurations a further effort in cloud tuning will be required, for instance, 713 to bring global-mean top-of-the-atmosphere radiative fluxes more inline with observa-714 tions, an aspect that is important in climate applications. A process-based approach to 715 evaluation, however, as taken in the present study, is certainly a valuable and integral 716 part of any assessment. 717

Based on the presented work the further development of scale-aware physical pa-718 rameterizations is endorsed, particularly with regard to convection (Holloway et al., 2014), 719 not only because of the expected better performance, fidelity of process representation, 720 and agreement with observations, but also because it enables better traceability and seam-721 lessness of the model development process across resolutions, time scales, and applica-722 tions. In the context of the Met Office and its partner institutions, we plan to operationally 723 implement the 5km-resolution global model as an experimental forecast to supplement 724 the main lower-resolution global ensemble. This could contribute to a limited number 725 of operational outputs, likely focussed on near-surface variables and high-impact weather, 726 as well as to the ongoing assessment and development and act as a stepping-stone to-727 wards a future convection-permitting global ensemble. This step-by-step approach al-728 lows for better identifying and understanding the strengths and limitations of a convection-729 permitting global model under various meteorological conditions. Future plans also in-730 clude season- and year-long simulations and the coupling to the NEMO ocean model. 731 But as the radiative flux estimates in the case of the *MidLevShConv25RAturb* config-732 uration show, some cloud tuning will be required as a prerequisite. 733

The Met Office modeling system will be overhauled over the next few years with the introduction of a new dynamical core, a new two-moment microphysics scheme, a new convection scheme, and improvements to the cloud scheme. The present work is therefore not intended to be the final word on convection-permitting global model development. Nevertheless, the results of the study will help to guide some fundamental choices in the future development of an adequate convection-permitting global coupled modeling system for use across weather and climate time scales.

741 Appendix A

Met Office GMED (Global Model Evaluation and Development) tickets associated
with the discussed model configurations are listed and described in some more detail in
Table A1. The information is supplementary to Table 1 and the configuration description in Section 2. A more in-depth description of the tickets is accessible only to regis-

Table A1. Met Office Global Model Evaluation and Development (GMED) tickets and de-scriptions related to the model configurations discussed in Section 2 and Table 1.

Short description	GMED ticket number	Notes
Fountain buster	492	Modification of the semi-Lagrangian advection scheme, see Section 2.
Additional termination condition for convection	417	Addition termination condition for subgrid convection whereby convection terminates when it reaches the level of neutral buoyancy of an undilute parcel.
Adaptation of convective momentum transport settings	487	The pressure gradient term in the Gregory-Kershaw scheme (Gregory et al., 1997) is reduced: cpress_term=0.3 instead of 0.7.
Forced detrainment weighted CAPE calculation	508	The CAPE used in the convection scheme closure is a dilute CAPE, i.e. it is reduced by entrainment. The convection scheme represents an ensemble of plumes through adaptive forced detrainment (Derbyshire et al., 2011). The change modifies the CAPE calculation so that it uses a forced detrainment weighted CAPE.
Time-smoothing of convective increments	191	To improve physics-dynamcs coupling time-damped convective increments to theta and humidity are passed out of the convection scheme rather than instantaneous values. Damping time scale is set to 1800s.
Switch off deep convection parameterization and fix CAPE timescale for midlevel scheme	557	Switch off deep convection scheme and rely on midlevel and shallow convection parameterizations to represent subgrid convection in the model. The CAPE closure time scale in the midlevel parameterization is set to a fixed value.
Setting for implicit boundary layer solver	290	Increase parameter for implicit boundary layer solver in unstable boundary layer columns from 0.5 to puns=1.0. The higher the value the more stable the scheme.
Polar cap for advection	471	The regions close to the poles have very small longitudinal grid lengths at high resolution. This leads to crossing of trajectories in the advection scheme. To avoid these problems polar cap regions are introduced across which advected quantificas are interpolated over the poles from advected values at the edges of the polar caps.
Prognostic entrainment for midlevel convection scheme and allow midlevel convection to initiate from the top of the boundary layer	507	Use of the prognostic entrainment scheme in the midlevel convection parameterization (Willett & Whitall, 2017). Allow the midlevel scheme to start from the top of the boundary layer instand of two model levels above.

tered Unified Model developers and users, but the ticket numbers are nevertheless pro-

vided here as an additional potential reference.

748 Appendix B Open Research

ERA5 data (Hersbach et al., 2020) were downloaded from the Copernicus Climate 749 Change Service (C3S) Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-750 era5-pressure-levels). Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007) 751 data (TMPA/3B43, V7) are available from the Goddard Earth Sciences Data and In-752 formation Services Center (GES DISC), https://disc.gsfc.nasa.gov/datasets/TRMM_3B43_V7/summary. 753 The GPM data (Huffman et al., 2019) were provided by the NASA/Goddard Space Flight 754 Center, which develop and compute the GPM IMERG dataset as a contribution to the 755 GPM project, and archived at the NASA GES DISC (https://disc.sci.gsfc.nasa.gov/). 756 The EUREC⁴A dropsonde dataset (George et al., 2021) is publicly available at AERIS 757 (https://doi.org/10.25326/221). The Darwin radar data can be accessed via the U.S. De-758 partment of Energy Atmospheric Radiation Measurement (ARM) programme data por-759 tal at https://www.arm.gov/data/data-sources/cpol. The daily mean CMsaf water va-760 por path data for the year 2016 can be obtained from the CMsaf team on request through 761 the web page at https://wui.cmsaf.eu/safira/action/viewProduktSearch. The observa-762 tions related to tropical cyclones can be downloaded from the IBTrACS data base of the 763 NOAA data center at https://www.ncdc.noaa.gov/ibtracs/. The CERES Energy Bal-764 anced and Filled (EBAF) TOA and Surface Monthly means data (Loeb et al., 2018) in 765 netCDF Edition 4.1 are available from the NASA Langley Atmospheric Science Data Cen-766 ter DAAC at https://doi.org/10.5067/TERRA-AQUA/CERES/EBAF_L3B.004.1. The 767 model data on which the figures are based are publicly available at 768

⁷⁶⁹ https://doi.org/10.5281/zenodo.7096991.

770 **References**

- 771Arakawa, A., Jung, J.-H., & Wu, C.-M. (2011). Toward unification of the multi-772scale modeling of the atmosphere. Atmos. Chem. Phys., 11, 3731–3742. doi: 10773.5194/acp-11-3731-2011
- Arakawa, A., & Wu, C.-M. (2013). A unified representation of deep moist con vection in numerical modeling of the atmosphere. Part I. J. Atmos. Sci., 70, 1977–1992.
- Bechtold, P., Semane, N., Lopez, P., Chaboureau, J. P., Beljaars, A., & Bormann, N.
 (2014). Representing equilibrium and nonequilibrium convection in large-scale
 models. J. Atmos. Sci., 71, 734–753.

780	Becker, T., Bechtold, P., & Sandu, I. (2021). Characteristics of convective precipita-
781	tion over tropical Africa in storm-resolving global simulations. Q. J. R. Meteo-
782	rol. Soc., 147, 4388–4407.
783	Boutle, I. A., Eyre, J. E. J., & Lock, A. P. (2014). Seamless stratocumulus simula-
784	tion across the turbulent gray zone. Mon. Wea. Rev., 142, 1655–1668.
785	Bush, M., et al. (2020). The first Met Office Unified Model-JULES Regional Atmo-
786	sphere and Land configuration, RAL1. Geosci. Model Dev., 13, 1999–2029.
787	Caldwell, P. M., et al. (2021). Convection-permitting simulations with the E3SM
788	global atmosphere model. J. Adv. Model. Earth Syst., 13, e2021MS002544. doi:
789	10.1029/2021 MS002544
790	Chen, H., Zhou, T., Neale, R. B., Wu, X., & Zhang, G. J. (2010). Performance of
791	the new NCAR CAM3.5 in East Asian summer monsoon simulations: Sensi-
792	tivity to modifications of the convection scheme. Journal of Climate, $23(13)$,
793	3657 - 3675.
794	Clark, P., Roberts, N., Lean, H., Ballard, S. P., & Charlton-Perez, C. (2016).
795	Convection-permitting models: a step-change in rainfall forecasting. Meteo-
796	rological Applications, 23, 165–181.
797	Crook, J., Klein, C., Folwell, S., Taylor, C. M., Parker, D. J., & Stratton, R.
798	(2019). Assessment of the representation of West African storm lifecycles
799	in convection-permitting simulations. Earth and Space Science, 6, 818–835.
800	doi: 10.1029/2018EA000491
801	Derbyshire, S. H. (1997). Recommendations for UM parametrization of stable bound-
802	ary layer (Tech. Rep. No. 38). United Kingdom: MRU Cardington.
803	Derbyshire, S. H., Maidens, A. V., Milton, S. F., Stratton, R. A., & Willett, M. R.
804	(2011). Adaptive detrainment in a convective parametrization. Q. J. R.
805	Meteorol. Soc., 137, 1856–1871.
806	de Roode, S. R., Frederikse, T., Siebesma, A. P., Ackerman, A. S., Field, P. R., Hill,
807	A., Tomassini, L. (2019). Turbulent transport in the grey zone: A large-
808	eddy simulation intercomparison study of the CONSTRAIN cold air outbreak
809	case. J. Adv. Model. Earth Syst., 11, 597–623. doi: 10.1029/2018MS001443
810	Dipankar, A., Webster, S., Sun, X., Sanchez, C., North, R., Furtado, K., Barker,
811	D. (2020). SINGV: A convective-scale weather forecast model for Singapore.
812	Q. J. R. Meteorol. Soc., 146, 4131–4146. doi: 10.1002/qj.3895
813	Field, P. R., Brodzkova, R., Chen, M., Dudhia, J., Lac, C., Hara, T.,
814	McTaggart-Cowan, R. (2017). Exploring the convective greyzone with regional
815	simulations of a cold air outbreak. Q. J. R. Meteorol. Soc., 143, 2537–2555.
816	Franklin, C. N., Protat, A., Leroy, D., & Fontaine, E. (2016). Controls on phase
817	composition and ice water content in a convection-permitting model simulation
818	of a tropical mesoscale convective system. Atmos. Chem. Phys., 16, 8767–8789.
819	doi: 10.5194/acp-16-8767-2016
820	Froude, L. S. R., Bengtsson, L., & Hodges, K. I. (2007a). The predictability of ex-
821	tratropical storm tracks and the sensitivity of their prediction to the observing
822	system. Mon. Wea. Rev., 135, 315–333.
823	Froude, L. S. R., Bengtsson, L., & Hodges, K. I. (2007b). The prediction of extrat-
824	ropical storm tracks by the ECMWF and NCEP ensemble prediction systems.
825	Mon. Wea. Rev., 135, 2545–2567.
826	George, G., et al. (2021). JOANNE: Joint dropsonde Observations of the Atmo-
827	sphere in tropical North atlaNtic meso-scale Environments. Earth Syst. Sci.
828	Data, 13, 5253-5272. doi: $10.5194/essd-13-5253-2021$
829	Gerard, L. (2015). Bulk mass-flux perturbation formulation for a unified approach of
830	deep convection at high resolution. Mon. Wea. Rev., 143, 4038–4063.
831	Gerard, L., & Geleyn, JF. (2005). Evolution of a subgrid deep convection
832	parametrization in a limited area model with increasing resolution. $Q. J.$
833	<i>R. Meteorol. Soc.</i> , 131, 2293–2312. doi: 10.1256/qj.04.72
834	Gerard, L., Piriou, JM., Brokov, R., Geleyn, JF., & Banciu, D. (2009).

835 836 837	Cloud and precipitation parameterization in a meso-gamma scale opera- tional weather prediction model. Mon. Wea. Rev., 137, 3960–3977. doi: 10.1175/2009MWB2750.1
030	Gregory D Kershaw B & Inness P M (1997) Parametrization of momen-
920	tum transport by convection II: Tests in single-column and general circulation
840	models Q. J. B. Meteorol. Soc. 123, 1153–1183
941	Grell G A & Freitas S B (2014) A scale and aerosol aware stochastic convective
842	parameterization for weather and air quality modeling Atmos Chem Phys
843	14. 5233–5250. doi: 10.5194/acp-14-5233-2014
844	Hanley, K. E., Plant, R. S., Stein, T. H. M., Hogan, R. J., Nicol, J. C., Lean, H. W.,
845	Clark, P. A. (2015). Mixing-length controls on high-resolution simu-
846	lations of convective storms. Q. J. R. Meteorol. Soc., 141, 272–284. doi:
847	10.1002/qj.2356
848	Hersbach, H., Bell, B., Berrisford, P., et al. (2020). The ERA5 global reanalysis. Q.
849	J. R. Meteorol. Soc., 146, 1999–2049. doi: 10.1002/qj.3803
850	Hodges, K. I. (1995). Feature tracking on the unit sphere. Mon. Wea. Rev., 123,
851	3458-3465.
852	Hohenegger, C., et al. (2022). ICON-Sapphire: simulating the components of the
853	Earth System and their interactions at kilometer and subkilometer scales.
854	Geosci. Model Dev doi: 10.5194/gmd-2022-171
855	Holloway, C. E., Petch, J. C., Beare, R. J., Bechtold, P., Craig, G. C., Derbyshire,
856	S. H., Woolnough, S. J. (2014). Understanding and representing atmo-
857	spheric convection across scales: recommendations from the meeting held at
858	Dartington Hall, Devon, UK, 28-30 January 2013. Atmos. Sci. Lett., 15,
859	348–353. doi: 10.1002/asl2.508
860	Honnert, R., Efstathiou, G. A., Beare, R. J., Ito, J., Lock, A., Neggers, R., Zhou,
861	B. (2020) . The atmospheric boundary layer and the gray zone of turbulence: a
862	Critical review. J. Geophys. Res., 125. doi: 10.1029/2019JD030317
863	Multinan, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Neikill, E. J., Bowman, K. P.,
864	Oussi-global multi-year combined-sensor precipitation estimates at fine scale
865	J. Hudrometeor 8 38–55
867	Huffman G. J. Stocker E. F. Bolvin D. T. Nelkin E. J. & Tan J. (2019)
868	GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06.
869	Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center
870	(GES DISC). doi: 10.5067/GPM/IMERG/3B-HH/06
871	Judt, F., & Rios-Berrios, R. (2021). Resolved convection improves the representation
872	of equatorial waves and tropical rainfall variability in a global nonhydrostatic
873	model. Geophys. Res. Lett., 48, e2021GL093265. doi: 10.1029/2021GL093265
874	Kendon, E. J., Jones, R. G., Kjellström, E., & Murphy, J. M. (2010). Using and
875	designing GCM-RCM ensemble regional climate projections. J. Climate, 23,
876	6485–6503. doi: 10.1175/2010JCLI3502.1
877	Kendon, E. J., Prein, A. F., Senior, C. A., & Sirling, A. (2021). Challenges and out-
878	look for convection-permitting climate modelling. Phil. Trans. R. Soc. A, 379,
879	20190547. doi: 10.1098/rsta.2019.0547
880	Kwon, Y. C., & Hong, S. Y. (2017). A mass-flux cumulus parameterization scheme
881	across gray-zone resolutions. Mon. Wea. Rev., 145, 583–598.
882	Lean, H. W., Clark, P. A., Dixon, M., Roberts, N. M., Fitch, A., Forbes, R., & Hal-
883	
884	liwell, C. (2008). Characteristics of high-resolution versions of the Met Ooffice
885	liwell, C. (2008). Characteristics of high-resolution versions of the Met Ooffice Unified Model for forecasting convection in the United Kingdom. <i>Mon. Wea.</i>
	 liwell, C. (2008). Characteristics of high-resolution versions of the Met Ooffice Unified Model for forecasting convection in the United Kingdom. Mon. Wea. Rev., 136, 3408–3424. Loch N. C. Doolling, D. R. Wang, H. Su, W. Nguyer, C. Cashett, J. C.
886	 liwell, C. (2008). Characteristics of high-resolution versions of the Met Ooffice Unified Model for forecasting convection in the United Kingdom. Mon. Wea. Rev., 136, 3408–3424. Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Kato, S. (2018). Clouds and the Earths Badiant Energy System (CERES).
886 887 888	 liwell, C. (2008). Characteristics of high-resolution versions of the Met Ooffice Unified Model for forecasting convection in the United Kingdom. Mon. Wea. Rev., 136, 3408–3424. Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Kato, S. (2018). Clouds and the Earths Radiant Energy System (CERES) Energy Balanced and Filled (EBAE) Top-of-Atmosphere (TOA) Edition-4.0
886 887 888 889	 liwell, C. (2008). Characteristics of high-resolution versions of the Met Ooffice Unified Model for forecasting convection in the United Kingdom. Mon. Wea. Rev., 136, 3408–3424. Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Kato, S. (2018). Clouds and the Earths Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 data product. J. Climate, 31, 895–918. doi: 10.1175/JCLI-D-17-0208.1

Lu, R., & Dong, B. (2001). Westward extension of North Pacific subtropical high in 890 Journal of the Meteorological Society of Japan. Ser. II, 79(6), 1229summer. 891 1241. 892 Martin, G. M., Levine, R. C., Rodriguez, J. M., & Vellinga, M. (2021). Understand-893 ing the development of systematic errors in the Asian summer monsoon. Geo-894 scientific Model Development, 14(2), 1007–1035. 895 Maynard, C., Melvin, T., & Müller, E. H. (2020). Multigrid preconditioners for the 896 mixed finite element dynamical core of the LFRic atmospheric model. Q. J. R. 897 Meteorol. Soc., 146, 3917–3936. doi: 10.1002/qj.3880 898 McTaggart-Cowan, R., Vaillancourt, P. A., Separovic, L., Corvec, S., & Zadra, A. 899 A convection parameterization for low-CAPE environments. (2020).Mon. 900 Wea. Rev., 148, 4917-4941. 901 Pan, D. M., & Randall, D. A. (1998). A cumulus parameterization with a prognostic 902 closure. Q. J. R. Meteorol. Soc., 124, 949-981. 903 Prein, A. F., et al. (2015). A review on regional convection-permitting climate mod-904 eling: Demonstrations, prospects, and challenges. Rev. Geophys., 53, 323–361. 905 doi: 10.1002/2014RG000475 906 Radermacher, C., & Tomassini, L. (2012).Thermodynamic causes for fu-907 ture trends in heavy precipitation over Europe based on an ensemble of 908 regional climate model simulations. J. Climate, 25, 7669-7689. doi: 909 10.1175/JCLI-D-11-00304.1 910 Roberts, N. M. (2003). The impact of a change to the use of the convection scheme 911 to high-resolution simulations of convective events (Tech. Rep. No. 407). Met 912 Office. 913 Rodríguez, J. M., & Milton, S. F. (2019).East Asian summer atmospheric mois-914 ture transport and its response to interannual variability of the West Pacific 915 subtropical high: An evaluation of the Met Office Unified Model. Atmosphere, 916 10(8), 457.917 Rodríguez, J. M., Milton, S. F., & Marzin, C. (2017). The East Asian atmospheric 918 water cycle and monsoon circulation in the Met Office Unified Model. Journal 919 of Geophysical Research: Atmospheres, 122(19), 10–246. 920 Sakradzija, M., Seifert, A., & Dipankar, A. (2016). A stochastic scale-aware param-921 eterization of shallow cumulus convection across the convective gray zone. J. 922 Adv. Model. Earth Syst., 8, 786–812. doi: 10.1002/2016MS000634 923 Saunders, R., Hocking, J., Turner, E., Rayer, P., Rundle, D., Brunel, P., ... Lupu, 924 C. (2018). An update on the RTTOV fast radiative transfer model. Geosci. 925 Model Dev., 11, 2717–2737. doi: 10.5194/gmd-11-2717-2018 926 Schäfler, A., et al. (2018). The North Atlantic Waveguide and Downstream Impact 927 Experiment. Bull. Amer. Meteor. Soc., 99, 1607–1637. 928 Senior, C. A., et al. Convection-permitting regional climate-change sim-(2021).929 ulations for understanding future climate and informing decision making in 930 Africa. Bull. Amer. Meteor. Soc., 102, E1206E1223. 931 Slingo, J., Bates, P., Bauer, P., Belcher, S., Palmer, T., Stephens, G., ... Teutsch, 932 Ambitious partnership needed for reliable climate prediction. G. (2022).933 Nature Climate Change, 12, 499–503. 934 Slingo, J., et al. (2021). Next generation climate models: a step change for net zero 935 and climate adaptation (Climate change science and solutions: Briefing 1). 936 United Kingdom: The Royal Society. Retrieved from https://royalsociety 937 .org/-/media/policy/projects/climate-change-science-solutions/ 938 climate-science-solutions-modelling.pdf 939 Stein, T. H. M., Hogan, R. J., Hanley, K. E., Nicol, J. C., Lean, H. W., Plant, R. S., 940 The threedimensional morphology of simulated and observed et al. (2014).941 convective storms over Southern England. Mon. Wea. Rev., 142, 3264–3283. 942 doi: 10.1175/MWRD1300372.1 943 Stevens, B., et al. (2021). EUREC⁴A. Earth Syst. Sci. Data, 13, 4067–4119. doi: 10 944

945	.5194/essd-13-4067-2021
946	Su, CY., Wu, CM., Chen, WT., & Chen, JH. (2021). Implementation of the
947	Unified Representation of deep moist convection in the CWBGFS. Mon. Wea.
948	Rev., 149, 3525-3539.
949	Tomassini, L. (2018). Mesoscale circulations and organized convection in African
950	easterly waves. J. Atmos. Sci., 75, 4357–4381. doi: 10.1175/JAS-D-18-0183.1
951	Tomassini, L. (2020). The interaction between moist convection and the atmospheric
952	circulation in the tropics. Bull. Am. Meteorol. Soc., 101, 1378–1396.
953	Tomassini, L., Field, P. R., Honnert, R., Malardel, S., McTaggart-Cowan, R., Saitou,
954	K., Seifert, A. (2017). The "Grey Zone" cold air outbreak global model in-
955	tercomparison: a cross-evaluation using large-eddy simulations. J. Adv. Model.
956	Earth Syst., 9, 39–64. doi: 10.1002/2016MS000822
957	Tomassini, L., & Yang, GY. (2022). Tropical moist convection an important driver
958	of Atlantic Hadley circulation variability. Q. J. R. Meteorol. Soc doi: 10
959	.1002/qj.4359
960	Walters, D., et al. (2017). The Met Office Unified Model Global Atmosphere $6.0/6.1$
961	and JULES Global Land 6.0/6.1 configurations. Geosci. Model Dev., 10, 1487–
962	1520. doi: 10.5194/gmd-10-1487-2017
963	Walters, D., et al. (2019). The Met Office Unified Model Global Atmosphere 7.0/7.1
964	and JULES Global Land 7.0 configurations. Geosci. Model Dev., 12, 1909–
965	1963. doi: 10.5194/gmd-12-1909-2019
966	Wang, B., Xiang, B., & Lee, JY. (2013). Subtropical high predictability estab-
967	lisnes a promising way for monsoon and tropical storm predictions. <i>Proceedings</i>
968	b) the National Academy of Sciences, $110(8)$, $2118-2122$. Wang W (2022) Economic convertion with a "scale sware" Tightle cumulus particular to the second state of
969	remeterization scheme at kilometer scales. Weather and Forecasting 27, 1401
970	1507
971	Willett M R & Whitell M A (2017) A simple prograstic based connective
972	entrainment rate for the Unified Model: description and tests (Tech Rep
973	No 617) Exeter United Kingdom: Met Office
975	Wood, N., Diamantakis, M., & Staniforth, A. (2007). A monotonically-damping
976	second-order-accurate unconditionally-stable numerical scheme for diffusion. Q.
977	J. R. Meteorol. Soc., 133, 1559–1573.
978	Wyngaard, J. C. (2004). Toward numerical modelling in the "Terra Incognita". J.
979	Atmos. Sci., 61, 1816–1226.
980	Yihui, D., & Chan, J. C. (2005). The East Asian summer monsoon: an overview.
981	Meteorology and Atmospheric Physics, 89(1), 117–142.
982	Zhang, C., & Ling, J. (2017). Barrier effect of the Indo-Pacific Maritime Conti-
983	nent on the MJO: Perspectives from tracking MJO precipitation. J. Climate,
984	30, 3439 – 3459.
985	Zheng, Y., Alpaty, K., Herwhe, J. S., Genio, A. D., & Niyogi, D. (2016). Improving
986	high-resolution weather forecasts using the Weather Research and Forecasting
987	(WRF) model with an updated Kain-Fritsch scheme. Mon. Wea. Rev., 144,
988	833–860. doi: 10.1175/MWR-D-15-0005.1



Figure 1. Hovmöller plots of three-hourly mean precipitation rate for the African easterly wave case. Panel a: TRMM observations; Panel b: *GA7*; Panel c: *MidLevShConv25*; Panel d: *MidLevShConv25RAturb*; Panel e: *MidLevShConv15RAturb*; Panel f: *ConvOffMoistConsRAturb*. The model data are interpolated to the TRMM grid using an area-weighted regridding scheme. The data are averaged over the latitudes 10N to 20N.



Figure 2. Observed (top left) and simulated reflectance at 0.8μ m for the MSG satellite for July 10, 2010, 12:00 UTC, based on the same configurations as in Figure 1, presented in the same order.



Figure 3. Observed (top left) and simulated brightness temperature at 10.8μ m for the MSG satellite for July 10, 2010, 18:00 UTC, based on the same configurations as in Figure 1, presented in the same order.



Figure 4. Results of convective storm tracking over the Africa-Atlantic region (40W-40E, 25S-25N) and the period July 8 to July 15, 2010. Top left panel: histograms of convective stroms life times; Top right panel: histograms of convective storms mean areas; Bottom panel: histograms of convective storms mean areas; Bottom panel: histograms of convective storms mean areas; Bottom panel: histograms area-weighted regridding scheme before the storm tracking is performed.



Figure 5. Diurnal cyle of precipitation for July 7 to July 15, 2010, averaged over the region 1E to 11E and 5N to 10N. The arrangement of the plots is as in Figure 1 with the top left panel showing TRMM observations.



Figure 6. Boxplot of TRMM observations and model simulations of 3-hourly mean rainfall rates above the threshold of 30mm/hr over the region 20W to 45E and 35S to 35N, comprising the whole of the African continent, and the period July 7 to July 15, 2010. The box plots show the lower, middle, and upper quartile of the data above the threshold, and the length of the whiskers are 1.5 times the interquartile range. Individual values are indicated by dots, the points outside the whiskers are also shown as bold dots.



Figure 7. Observed and modelled central pressure developments for tropical cyclones Dorian (top panel) and Goni (bottom panel). The operational model is GA6 in the case of Dorian and GA7 in the case of Goni, both run with fully parameterized convection at 10km resolution. The observations are from the IBTrACS data base of the National Oceanic and Atmospheric Administration (NOAA) National Climate Data Center (NCDC).



Figure 8. Cyclone track predictions for tropical cyclone Goni in the vicinity of the Philippines. The operational model (OP; red lines) is GA7 run with fully parameterized convection at 10km resolution. Black lines are observations and red lines the different 5km-resolution configurations. The observations are from the IBTrACS data base of the National Oceanic and Atmospheric Administration (NOAA) National Climate Data Center (NCDC).



Figure 9. Observed (top left) and simulated reflectance at 0.64μ m for the GOES16 satellite for February 2, 2020, 18:00 UTC, based on the same configurations as in Figure 1, presented in the same order.



Figure 10. Temperature and specific humidity profiles for HALO aircraft dropsondes, ERA5, *GA7, MidLevShConv25RAturb*, and *MidLevShConv15RAturb* on February 2, 2020, 18:00 UTC, averaged over the EUREC⁴A observation circle east of Barbados (Stevens et al., 2021). The observed dropsonde profiles are sampled between 17:50 UTC and 18:10 UTC.



Figure 11. Observed (top left) and simulated radar reflectivity profiles (February 18, 12-24 UTC mean) over the area within reach of the Darwin C-band radar. Color shaded contours represent the frequency of occurrence.



Figure 12. Precipitation rate for February 18, 2014, 18:00 UTC, over the wider Darwin region based on the same configurations as in Figure 1, presented in the same order. Top left is for GPM observations. The model data are interpolated to the GPM grid using an area-weighted regridding scheme.



Figure 13. NAWDEX case precipitation for September 27, 2016, 18:00 UTC, based on the same configurations as in Figure 1, presented in the same order. Top left is for GPM observations. The model data are interpolated to the GPM grid using an area-weighted regridding scheme.



Figure 14. Daily mean integrated water vapor path for September 27, 2016 (left), and October 10, 2016 (right). Top row: Satellite Application Facility on Climate Monitoring (CMsaf) observations. Middle row: *GA7* configuration; bottom row: *MidLevShConv25RAturb*. The model data are interpolated to the CMsaf grid using an area-weighted regridding scheme.



Figure 15. NAWDEX case precipitation for October 13, 2016, 18:00 UTC, based on the same configurations as in Figure 1, presented in the same order. Top left is for GPM observations. The red circle highlights the area of heavy rainfall in southern France. The model data are interpolated to the GPM grid using an area-weighted regridding scheme.



Figure 16. NAWDEX case PV at 320K for October 13, 2016, 18:00 UTC, based on the same configurations as in Figure 1, presented in the same order. Top left is for ERA5.



Figure 17. Model convective rainfall for *MidLevShConv25* with CAPE time scale 3000 seconds (top left), *MidLevShConv25* with CAPE time scale 2700 seconds (top middle), and *Mi-dLevShConv25* with CAPE time scale 2400 seconds (top right). Bottom row: Total precipitation for the same configurations. Although the convective rainfall is only moderately increased in the simulation with CAPE time scale 2400 seconds compared to the other ones, the more active convection scheme starts to disrupt the prediction of the mesoscale convective systems during the early strengthening phase of the African easterly wave. The model data are interpolated to the TRMM grid using an area-weighted regridding scheme



Figure 18. Hovmöller plots of precipitation for the African easterly wave case (top row). Top left: MidLevShConv25RAturb. Top right: the same configuration, but with humidity perturbations in the midlevel scheme instead of temperature perturbations. The model data are interpolated to the TRMM grid using an area-weighted regridding scheme. Middle row: the same comparison, but showing simulated reflectance at 0.8μ m for the MSG satellite for July 10, 2010, 12:00 UTC. Bottom row: the same comparison but showing simulated brightness temperature at 10.8μ m for the MSG satellite for July 10, 2010, 18:00 UTC.



Figure 19. Mean outgoing shortwave radiation for July 2016 from the *MidLevSh-Conv25RAturb* simulation (top left) and CERES EBAF 4.1 (top right). Mean outgoing longwave radiation for July 2016 from the *MidLevShConv25RAturb* simulation (middle left) and CERES EBAF 4.1 (middle right). Mean precipitation for July 2016 based on the *MidLevShConv25RAturb* simulation (bottom left) and GPM (bottom right).



Figure 20. Mean outgoing shortwave radiation for January 2018 from the *MidLevSh-Conv25RAturb* simulation (top left) and CERES EBAF 4.1 (top right). Mean outgoing long-wave radiation for January 2018 from the *MidLevShConv25RAturb* simulation (middle left) and CERES EBAF 4.1 (middle right). Mean precipitation for January 2018 based on the *MidLevSh-Conv25RAturb* simulation (bottom left) and GPM (bottom right).



Figure 21. Right panel: Hovmöller plot of 3-hourly mean rainfall rate based on the *Mi-dLevShConv25RAturb* configuration simulations for January 2018. Middle panel: Hovmöller plot of 3-hourly mean rainfall rate based on the then operational model GA6.1 (Walters et al., 2017) with fully parameterized convection. Right panel: Hovmöller plot of rainfall based on the GPM rainfall data averaged to 3-hourly mean values. For all plots the rainfall was averaged over the latitudes 10 South to 10 North, units are mm/day.



Figure 22. Panel a: July 2016 mean 850hPa geopotential height and horizontal wind in ERA5. Panels b and c: Difference of July 2016 mean 850hPa geopotential height and horizontal wind to ERA5 for the 5km-resolution *MidLevShConv25RAturb* configuration and GA6 at N216 resolution, respectively. Geopotential height is in m and wind in ms⁻¹. WPSH intensity, west extension and northern edge indices in ERA5 and model hindcasts are shown in panel d. Axes in the scatter plot are as follows: abscissa displays values of the west extension index, I_W (°E) and ordinate shows values of the northern edge index, I_N (°N). The radii of scatter circles are proportional to the intensity index, I_{in} . Indices are shown for July 2016 ERA5 (light green), 5km-resolution (blue) and N216-resolution (red). To place 2016 data in context, the following information has been added: dotted lines indicate maximum and minimum values of July ERA5 I_W and I_N for the 1979-2019 period, and a dark green dot shows the period mean value.



Figure 23. Results of extratropical storm tracking for January 2018 in the northern hemisphere (two left panels) and July 2016 in the southern hemisphere (two right panels). Panel a: Mean storm track position error of the global 5km-resolution *MidLevShConv25RAturb* forecasts against ECMWF analysis (black line), and the then-operational Unified Model GA6 at 10km (N1280) resolution against ECMWF analysis (red line) for January 2018 and the northern hemisphere. Panel b: The same models and comparisons as in panel a, but for the mean intensity error of extratropical cyclones. The intensity measure is 850hPa vorticity. Panel c: As for panel a, but for July 2016 and the southern hemisphere. Here the then-operational model is run at N768 resolution. Panel d: The intensity error results for July 2016 over the southern hemisphere.