Asia-Pacific Journal of Atmospheric Sciences (2025) 61:16
https://doi.org/10.1007/s13143-025-00398-4

ORIGINAL ARTICLE

Predicting Convectively Induced Turbulence With Regionally Convection-Permitting Simulations

Haoming Chen[1] · Christy Yan-yu Leung[2] · Ping Cheung[2] · Haolin Liu[1] · Sai Tick Chan[2] · Xiaoming Shi[1,3]

Received: 27 May 2024 / Revised: 31 March 2025 / Accepted: 3 April 2025

Abstract

Convectively induced turbulence (CIT) is a severe aviation hazard. It is challenging to forecast CIT because low-resolution models cannot explicitly resolve convective motions at kilometer scales. In this study, we used the Model for Prediction Across Scales (MPAS) to simulate CIT cases with convection-permitting resolution (∼1 km) in the region of the CIT events and coarse resolution in other parts of the globe. We developed a method to estimate the eddy dissipation rate (EDR) using the resolved wind field of the MPAS simulations. The method is based on explicit filtering and reconstruction in the turbulence modeling for large-eddy simulations (LES). It estimates turbulence kinetic energy (TKE), which is then used to derive EDR. The new method produces different turbulence distribution and intensity than previous methods based on second-order structure functions and convective gravity wave drag, with higher accuracy and better correlation with observations for CIT cases tested in this study. The 1-km resolution simulation generates more accurate EDR and improves spatial patterns, but it is computationally demanding. The 3-km resolution can get benefits from reasonable accuracy and affordable computational cost. Because convection-permitting resolutions are in the gray zone for simulating convection, we evaluated the sensitivity of the prediction to the variations in physical and numerical schemes. Varying cumulus convection parameterization and monotonicity of numerical schemes are identified as practical approaches to generate beneficial ensemble spread. However, the physical perturbation-based ensemble has limitations, and initial condition perturbations are still necessary to encompass uncertainties in the development of convection.

Keywords Turbulence · Ensemble simulations · Subfilter-scale reconstruction · Convection

1 Introduction

Aviation turbulence is the primary weather-related factor contributing to aviation incidents, causing numerous injuries, occasional fatalities, and structural damage each year. Furthermore, schedule delays, air traffic management problems, and operational costs to airlines are usually related to turbulence (Tvaryanas 2003; Sharman et al. 2012; Kim and Chun 2016; Sharman and Lane 2016). Convectively induced turbulence (CIT) is a type of aviation turbulence and a challenge for aviation safety. CIT can be generated from the following physical mechanisms: 1) convection penetrates the upper troposphere and enhances the background wind shear, 2) buoyancy gradients at the cloud boundary causes flow deformation, and 3) convection triggers gravity waves breaking above itself (Lane et al. 2003).

Extended author information available on the last page of the article

Atmospheric models are used to simulate the relevant weather conditions and forecast CIT. Many turbulence prediction products rely on empirical indices related to measures of gravity waves or atmospheric instability (Endlich 1964; Dutton 1980; Vogel and Sampson 1996; Ellrod and Knox 2010; Muñoz-Esparza and Kosović 2018). Numerical resolution is a critical barrier when atmospheric models are used for forecasting aviation turbulence. Convection-permitting (∼1 km) resolution can help a model explicitly resolve convection, but it is computationally expensive. Additionally, eddies span a spectrum of sizes from 100 kilometers to centimeters in the atmosphere. Aircraft bumpiness is most pronounced when the size of the turbulent eddies encountered is about the size of the aircraft (Vinnichenko 2013). For commercial aircraft, this would correspond to eddy sizes of a few hundred to a few thousand meters, which is infeasible for operational numerical weather prediction (NWP) to resolve.

Previous studies have developed practical algorithms to forecast non-convectively induced turbulence. Sharman et al. (2006) developed the Graphical Turbulence Guidance system version 2 (GTG2) to forecast aviation turbulence with the simulated fields from an NWP model at 20-km horizontal resolution. GTG2 utilizes many turbulence diagnostics to improve forecasting performance, and it is recognized that NWP model resolution is one factor that hampers more accurate results. In its latest version, GTG3, Sharman and Pearson (2017) used simulation results from 13-km resolution Weather Research and Forecasting Rapid Refresh and acknowledged the necessity for higher resolutions (grid spacing less than 3 km or 1 km) for some upper-level turbulence events. In addition, GTG3 was also applied at a higher resolution (3 km) and machine learning was applied to improve the EDR forecast (Muñoz-Esparza et al. 2020). The Global-Korean aviation turbulence guidance (Lee et al. 2022) is a product similar to GTG3,and it uses the results with 10-km resolution from Global Data Assimilation and Prediction System (GDAPS) to provide turbulence intensity forecasts. It shows including the near-cloud turbulence diagnostics can improve the performance skill, especially during the summer (Kim et al. 2021). However, overpredicting still occurs at all altitudes due to the coarse mesh of the NWP (Lee et al. 2022). These products can be applied to meshes with different resolutions to generate forecast by different large-scale turbulence diagnostics and parameterized wave breaking about turbulent mixing. If we have convection-permitting resolutions and resolve convection, can we directly estimate CIT intensity from one variable based on simulated fields? This question is the overarching theme of our study.

Previous research has employed high-resolution simulations to study CIT and shown promising results. Barber et al. (2019) used the Weather Research and Forecast (WRF) model with nested domains to capture turbulence in the Gulf of Mexico, and their finest resolution was 3 km. The simulation utilized in Barber et al. (2019) captured CIT successfully and showed that developing convection can generate more substantial turbulence than mature convection. Lane and Sharman (2014) used a large-eddy simulation with 75-m resolution in the horizontal. They found that the position of the most intense turbulence is outside of convective clouds, and CIT extends to 50 km away from the cloud boundary, beyond the Federal Aviation Administration (FAA) guidelines. Other studies also found that the high-resolution models can help us understand the life cycle of CIT (Lane et al. 2009; Trier et al. 2010; Trier and Sharman 2016).

The eddy dissipation rate to the one-third power (EDR, unit: $m^{2/3}s^{-1}$) is valuable for comparing high-resolution model prediction and observed aviation turbulence. EDR has been adopted as a standard turbulence indicator that the International Civil Aviation Organization reports. It represents the kinetic energy transfer rate from large-scale eddies to small scales (Ahmad and Proctor 2012). Some previous studies used the second-order structure functions (2ndSF) to calculate EDR from high-resolution model output, which is a very useful statistic and measures the kinetic energy of all vortex structures at a given scale (Kolmogorov 1991; Frehlich and Sharman 2004; Sharman et al. 2006; Barber et al. 2019). The calculation of EDR from 2ndSF on a rectilinear grid mesh is a well-established technique, and some studies have calculated best-fit functions based on a statistical analysis of physical quantities in the middle and upper atmosphere, such as wind speed, pressure and potential temperature (Frehlich and Sharman 2004, 2010; Lindborg 1999). Barber et al. (2019) used the turbulence kinetic energy (TKE) from the planetary boundary layer (PBL) scheme to compute EDR and made a comparison with the result from the 2ndSF. They suggest that the 2ndSF is more reliable than the PBL TKE. Moreover, convective gravity wave drag (CGWD) can also be used to calculate the EDR by estimating the impact due to gravity wave breaking (Kim et al. 2019).

In this research, we evaluate the potential of the Model for Prediction Across Scales (MPAS) in predicting CIT with several reported CIT incidents near Hong Kong. Section 2 lists the incidents' details, model configurations, and methods to estimate EDR. A new method to estimate EDR based on subfilter-scale reconstruction(SFSR) (Chow et al. 2005) is developed. Section 3 shows the performance of MPAS in simulating convection. Different EDR estimation methods and the influence of resolution are presented in section 4. Section 5 discusses how the influence of different physical parameterization or numerical scheme options might change the simulated convection. Section 6 evaluates the performance of those methods with more cases and discusses the statistical properties of those methods. Section 7 summarizes and discusses the implications of our study.

2 Experimental Design and Methods

2.1 MPAS Setup

This study used MPAS version 7 to conduct regionally convection-permitting simulations. MPAS features a non-hydrostatic dynamical core that utilizes unstructured Voronoi meshes and C-grid discretization (Skamarock et al. 2012). The global variable-resolution mesh can have finer resolutions in interested areas. In recent years, MPAS has been extensively used to investigate significant weather phenomena that depend on resolution, such as clouds, extreme precipitation events, and atmospheric rivers. O'Brien et al. (2013); Landu et al. (2014); Yang et al. (2014); Hagos et al. (2015); Sakaguchi et al. (2015); Zhao et al. (2016).

This study focuses on five CIT incidents reported near Hong Kong. Table 1 lists those cases' time, altitudes, the

Table 1 Time (UTC), altitude, flight stage, and maximum turbulence intensity of the five CIT cases.

Case	Time (UTC)	Altitude (m)	Stage	Max EDR ($m^{2/3}s^{-1}$)@Time
1	2020-05-21	10000	cruising	0.465@01:49
2	2020-06-06	9450	cruising	0.623@09:35
3	2020-06-08	4500	landing/taking off	0.493@04:24
4	2020-08-26	6600	cruising	0.687@13:27
5	2021-06-27	3900	landing/taking off	0.516@23:48

stage of airplane, and in-situ EDR record (Takacs et al. 2005). EDR and location data were recorded by aircraft and shown in Fig. 1 .These cases have both high EDR records and weather phenomena that can potentially induce aviation turbulence that can be observed through satellites or radar, such as severe convection. However, distinguishing the source of the high EDR in landing/take-off is not a simple qualitative task, which may require airlines to provide more detailed records and research. Unless specified otherwise, our numerical experiments are conducted with a 3–60 km resolution mesh. Figure 2 shows the mesh configuration, which has a 3-km resolution in South China and the South China Sea and a transition to a 60-km resolution away from this region. To examine the impact of resolution on the results of this study, other refined resolutions, 1 km, 9 km, and 18 km, are used for the refined region in some simulations (Fig. 1.).

The model is configured to have 55 vertical layers, with the top of the model at 30 km above the surface. The "Base" experiment uses the Grell-Freitas (GF) convection parameterization, which is modified to work across grid spacings from mesoscale to convective scales (Grell and Freitas 2014), the MPAS microphysics suite, which uses the Thompson scheme (Thompson et al. 2008) for grid cells smaller than 10 km and the WRF Single-Moment 6-class (WSM6) scheme (Hong et al. 2006) for other cells, the planetary boundary layer scheme suite, which uses the Yonsei University (YSU) (Hong 2010) at the coarser resolution and the Mellor-Yamada-Nakanishi-Niino (MYNN) (Nakanishi and Niino 2009) at the finer resolution. The Noah land surface scheme (Chen and Dudhia 2001), and the RRTMG short and longwave radiation schemes (Mlawer et al. 1997; Iacono et al. 2000) are used in all simulations.

Because the convection-permitting resolution is in the gray zone for convection, the choices of relevant parameterization are subject to uncertainty. Therefore, we evaluate the potential of physics-based ensembles by varying physical or numerical options, one at a time, in our simulations. Table 2 lists the difference between each experiment and the "Base" run. We varied the choices for microphysics, which has influence on hydrometeors as well as the convection (Mohan et al. 2019); cumulus convection parameterization since we have high resolution to simulate the convection directly; monotonic limiter in scalar advection, turning which off may allow more numerical noise at small scales; and the Smagorinsky coefficient which can influence the turbulent viscosity for horizontal turbulence mixing. SMAG-S and

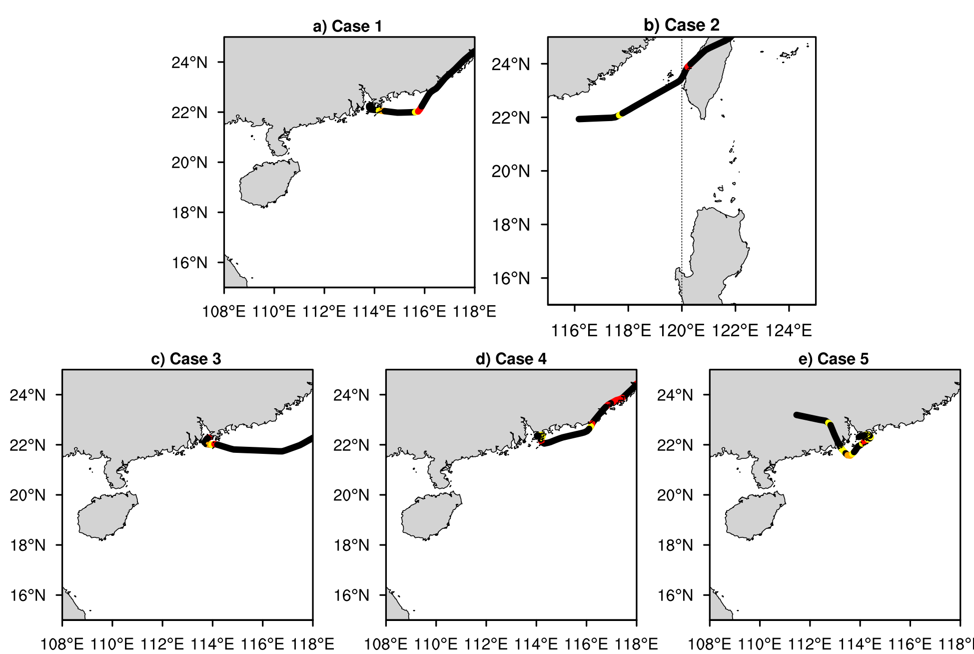

Fig. 1 Trajectories and turbulence from different cases in Table 1. a) Case 1, b) Case 2, c) Case 3, d) Case 4 and e) Case 5. The line represents the route of the airplanes, while red indicates the EDR ($m^{2/3}s^{-1}$) is higher than 0.3, orange for 0.2–0.3, yellow for 0.1–0.2, black for 0–0.1

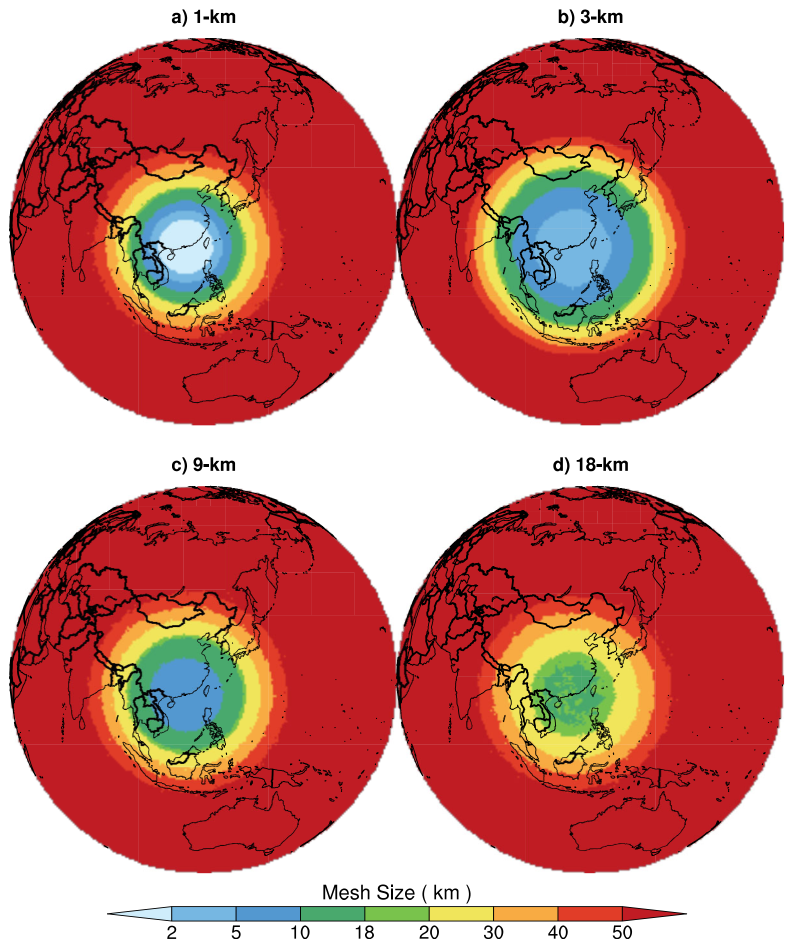

Fig. 2 Global variable-resolution mesh size in the variable-resolution a) 1 ~ 60 km, b) 3 ~ 60 km, c) 9 ~ 60 km, and d) 18 ~ 60 km experiments

SMAG-L represent two experiments with small (0.025) and large (0.5) Smagorinsky coefficients. Overall, in the gray zone, the selections of those schemes are controversial and can influence small-scale motions, which impact the estimation of EDR. And we do not intend to demonstrate that some schemes are superior to others. It represents the limitations, instead of improvements, of our modeling techniques.

The initial conditions are derived from the European Centre for Medium-Range Weather Forecast (ECMWF) fifth-generation reanalysis (ERA5) data at a 0.25° horizontal grid spacing and 37 vertical levels (Bell et al. 2021). The initialization time of our simulations is approximately 12 hours before the occurrence of the maximum EDR observation in each case. We tested experiments with initialization six hours before the CIT incidents, but those simulations produced less accurate predictions, probably resulting from the need for model spinup.

2.2 Estimation of EDR

Here, we briefly describe the three methods of computing EDR, which we compared in this study. They are based on (1) 2ndSF, (2) SFSR, and (3) CGWD. The second is a method we developed based on the explicit and reconstruction modeling of turbulence, which has been used in large-eddy simulations (LES) and gray-zone simulations.

Table 2 Model parameterizations used in simulations

Experiments	Physics/Numerics	Default options ("Base")	Experiment choice
WSM6	Microphysics	Thompson	WSM6
NoCU	Cumulus convection	Grell-Freitas	None
NoML	Monotonic limiter	On	Off
SMAG-S	Smagorinsky coefficient	0.125	0.025
SMAG-L	Smagorinsky coefficient	0.125	0.5

Each member has one modification in its configuration compared to "Base"

2.2.1 Second-Order Structure Functions

In this method, turbulence can be described by longitudinal and transverse structure functions, which are defined by

$$D_{LL}(r) = \left\langle [u_L(x) - u_L(x+r)]^2 \right\rangle \tag{1}$$

$$D_{NN}(r) = \left\langle [u_N(x) - u_N(x+r)]^2 \right\rangle \tag{2}$$

respectively. They measure the kinetic energy of all vortex structures with a scale less than or equal to the length r. Here the u_L is the velocity component along the position vector $\boldsymbol{r} = (x, y, z)$, and u_N is the transverse component, r is a separation distance expressed in units of spatial grid steps, and the angle brackets indicate the average in a spherical surface with radius $|\vec{r}|$. In Kolmogorov's model, based on universal equilibrium hypotheses, when the length scale is in the inertial subrange, the structure functions and EDR can be linked by

$$D_{LL}(r) = C_k \varepsilon^{2/3} r^{2/3} \approx 2\varepsilon^{2/3} r^{2/3} \tag{3}$$

$$D_{NN}(r) = \frac{4}{3} C_k \varepsilon^{2/3} r^{2/3} \approx \frac{8}{3} \varepsilon^{2/3} r^{2/3} \tag{4}$$

where C_k is set to 2 and $\varepsilon^{1/3}$ is the EDR. The difference in the coefficients in two directions is deduced by Monin and Yaglom (2013). In our calculation, r is seven grid spacings because it should represent the spectral resolution of the advection scheme, which is 7 to 10 Δx (Skamarock 2004; Muñoz-Esparza et al. 2018; Barber et al. 2019).

For the resolution of convection-permitting simulations, it is difficult to apply the same horizontal separation length (in our mesh, 30 km) to the vertical because of the relatively shallow depth of the troposphere. Many previous studies consider the horizontal velocities and gradients only to calculate the structure functions (Barber et al. 2019; Frehlich and Sharman 2004), and this approximation is also what we adopted here.

2.2.2 Sub-Filter-Scale Reconstruction (SFSR)

This method estimates EDR by computing TKE first. Here, we adopt the idea of explicit filtering and reconstruction in turbulence parameterization (Chow et al. 2005). This framework separates subfilter scales into resolvable subfilter scales (RSFS) and subgrid scales (SGS). We compute both the RSFS part and SGS part to provide EDR with similar statistical characteristics when the resolution is changed. Following (Chow et al. (2005)), we first reconstruct RSFS velocity through deconvolution. The reconstructed RSFS velocity

$$\widetilde{u}_i^* = \overline{\widetilde{u}}_i + (I - G)\overline{\widetilde{u}}_i + (I - G)(I - G)\overline{\widetilde{u}}_i + \cdots \tag{5}$$

where the overline denotes the filter, the tilde denotes discretization, $\overline{\widetilde{u}}_i$ is, therefore, the grid variable from MPAS, I is the identity operator, and G is the filter. In this study, the explicit filter is a top-hat filter (1-2-1 filter) applied to all three dimensions. The corresponding cutoff wavelength is $2\Delta x$. This filter is the recommendation from Chow et al. (2005); Gullbrand and Chow (2003). Keeping $\overline{\widetilde{u}}_i$ is the zero-order reconstruction and is what we adopted ($\widetilde{u}_i^* = \overline{\widetilde{u}}_i$). Including more terms on the right side of Eq. 5 generates higher-order reconstruction, which is not used in this study because it may occasionally generate negative TKE.

After obtaining RSFS velocities, the RSFS TKE is

$$\tau_{RSFS}^{ij} = (\overline{\widetilde{u}_i^* \widetilde{u}_i^*} - \overline{\widetilde{u}_i^*}\,\overline{\widetilde{u}_i^*})/2 \tag{6}$$

The SGS TKE is calculated based on Shi and Wang (2022) and this method improved the calculation of SGS term with the nonlinear backscatter and anisotropy (NBA) model, which allows both energy forward scatter and backscatter.

$$\begin{aligned}\tau_{SGS}^{ij} = &-C_s^2 l^2 [2(2S_{mn}S_{mn})^{1/2} S_{ij} \\ &+ C_1(S_{ik}S_{kj} - S_{mn}S_{mn}\delta_{ij}/3) \\ &+ C_2(S_{ik}R_{kj} - R_{ik}S_{kj})]\end{aligned} \tag{7}$$

Here the overline still denotes the top-hat filter. S is the resolved strain-rate tensor $S_{ij} = \frac{1}{2}(\frac{\partial \overline{\widetilde{u}}_i}{\partial x_j} + \frac{\partial \overline{\widetilde{u}}_j}{\partial x_i})$, R is the resolved rotation-rate tensor $R_{ij} = \frac{1}{2}(\frac{\partial \overline{\widetilde{u}}_i}{\partial x_j} - \frac{\partial \overline{\widetilde{u}}_j}{\partial x_i})$, $C_s = [8(1 + C_b)/27\pi^2]^{1/2}$, $C_1 = C_2 = 960^{1/2} C_b/7(1 + C_b)S_k$, $S_k = 0.5$, C_b=0.36, and $l = \Delta x \Delta y \Delta z$ (Mirocha et al. 2009). Assuming in the inertial subrange, the EDR is the following (Schumann 1991),

$$\mathrm{TKE} = \tau_{RSFS} + \tau_{SGS} \tag{8}$$

$$\varepsilon^{1/3} = (\mathrm{TKE}^{3/2}/L)^{1/3} \tag{9}$$

where $L = (\lambda \Delta x \Delta y \Delta z)^{1/3}$ is the integral scale of the turbulence, Δx, Δy and Δz are grid spacings. MPAS data were interpolated to 0.04° × 0.04° rectangular grid before applying the above equations by using Earth System Modeling Framework library through the NCAR Command Language (Brown et al. 2012) with bilinear method, which is widely used in MPAS hexagon mesh data regridding (Li et al. 2022; Xu et al. 2021). Therefore, Δx and Δy are approximately

4.5 km for the region near Hong Kong. In addition, interpolating the results to a coarser resolution than the $0.04° \times 0.04°$ can cause the turbulence regions to expand, while the EDR values in the severe turbulence regions will decrease. Therefore, it is important to select an appropriate resolution to which the raw model output is interpolated. The Δz is 500 m, which is the grid spacing in the middle troposphere. λ is a flow dependent quantity and difficult to obtain (Barber et al. 2019; Sharman et al. 2012). This value can be calculated from boundary layer parameterization schemes (Ahmad and Proctor 2012), but this method does not work in our high-altitude cases. We acknowledged that this problem is difficult to solve immediately, and we selected a constant value, $\lambda = 8$, for our calculation because of the cutoff wavelength ($2\Delta x$) in our filter.

Applying filters on the original MPAS grid to calculate the RSFS EDR is also possible, Allen (2005) developed filters for hexagonal grids. Our evaluation using the filtering technique described in Allen (2005) yielded results similar in spatial distribution and magnitude to our calculation using data interpolated to the latitude-longitude grids (See Appendix Fig. 13). However, due to regional refinement, MPAS mesh has some grid cells with five or seven edges, which will make the method based on hexagonal grids have to skip these pentagons or heptagons, resulting in gaps in the horizontal distribution of EDR; this problem becomes more severe in coarser mesh such as 9–60 km mesh since the number of pentagons or heptagons increases.

2.2.3 CGWD-Based Estimation

CGWD parameterization was proposed by Chun and Baik (1998). Kim et al. (2019) used their parameterization to calculate EDR. The wave stress above the cloud is calculated by:

$$\tau = -\left[\rho\,|U|^2 \,/\, (N \bigtriangleup x)\right] U c_1 c_2^2 \mu^2 \tag{10}$$

Here, ρ is the air density, U is the basic-state wind, N is Brunt-Väisälä frequency, $\triangle x$ is the horizontal grid spacing. The parameter c_1 is a constant related to diabatic forcing $c_1 = \pi ln\left[(a_1 + a_2)^2 / 4a_1 a_2\right]$, a_1 is the half width of heating $a_1 = \alpha\Delta x/2$, α is assumed as 0.4, and $a_2 = 5a_1$. The c_2 has correlation with the stability and the position of cloud, it is defined as $c_2 = cos(\frac{Nz_t}{U}) - cos(\frac{Nz_b}{U})$, z_t represents the top of the cloud while z_b represents the bottom of the cloud. We also consider clouds on different levels and clear sky between clouds as a whole within a horizontal column when multiple cloud layers appear. This means that we consider the bottom of the lowest level cloud as the cloud bottom, and the top of the highest level cloud as the cloud top. μ represents a nonlinearity factor of thermally induced internal gravity waves, it is defined as $\mu = gQ_0 a_1/(c_p T N U^2)$, Q_0 is the column-maximum heating rate inside of the cloud, in MPAS model, it is the maximum of the summation of the tendency of potential temperature from cumulus convection and cloud microphysics in a column, and c_p is the specific heat of air at a constant pressure. Here, the cloud boundary is defined by the summation of the heating rate with a threshold of 10^{-5}K/s.

Minimum Richardson number $\mathbf{Ri}_{min} \approx \mathbf{Ri}\frac{1-\mu|c_2|}{1+\mu\mathbf{Ri}^{1/2}|c_2|^2}$ can identify levels of wave breaking by a value smaller than 0.25. The calculation of the saturated wave stress and related quantities below will be changed as follows (Lindzen 1981). However, if the value of $\mathbf{Ri}_{min}$ is greater than 0.25, the wave stress is the same as the values that are lower than this altitude since we assume that there is no wave breaking.

$$\tau = -\left[\rho\,|U|^3 \,/\, (N \bigtriangleup x)\right] c_1 c_2^2 \mu_s^2 \tag{11}$$

$$\mu_s = |c_2|^{-1}(2\sqrt{2 + \mathbf{Ri}^{-1/2}} - 2 - \mathbf{Ri}^{-1/2}) \tag{12}$$

Then, CGWD can be given by the gradient of stress,

$$\mathrm{CGWD} = -\frac{1}{\rho}\frac{\partial \tau}{\partial z} \tag{13}$$

and the diffusion coefficient is

$$K_{\mathrm{CGWD}} = \left|\mathrm{CGWD}\frac{c - U}{N^2}\right| \tag{14}$$

where c is the horizontal phase speed, which is set to zero by assuming that CGW is stationary relative to the convection system, U is the basic-state wind, N is the Brunt-Väisälä frequency, and ρ is the density of the air. The TKE in this method is

$$\mathrm{TKE} \approx \left(C_d^{-1}\frac{K_{\mathrm{CGWD}}}{L}\right)^2 \tag{15}$$

and the EDR is

$$\mathrm{EDR} \approx \left(\frac{C_\varepsilon \mathrm{TKE}_{\mathrm{CGWD}}^{3/2}}{L}\right)^{1/3} \tag{16}$$

where C_d is set to 0.1 (Lane and Sharman 2008) and the C_ε is set to 0.93 (Moeng and Wyngaard 1988), here L is a length scale set as the vertical grid spacing.

3 Large-scale Environmental Conditions

Figure 3 shows the spatial distribution of brightness temperature for Case 1 (Table 1) from an infrared channel of Himawari-8 satellite observation and the experiments with different physics or numerics options (Table 2). The multi-scale structural similarity (MSSSIM, a higher value close to 1 indicates higher similarity between the two images) is used to compare the similarity between experiments and observation. The MSSSIM values are shown in respective figure titles. The brightness temperature for MPAS model data is simulated with the Radiative Transfer for TIROS Operational Vertical Sounder (RTTOV). The overall spatial distribution of clouds is similar in those simulations, with one intense convective system in the northern part of the South China Sea and another overland in the Guangdong province of China. However, comparing the Base run with satellite images, we can find that the pattern of the over-land convection is not entirely the same. In the satellite image, there is a gap (clear-sky area) between the convective systems over land and ocean, but in the Base run, the two are partially connected, and clouds partially cover the coastal line. NoCU simulation is the only run exhibiting clear-sky conditions along the coastal line. However, there is a deviation between NoCU and observation in this clear-sky area, so its MSSSIM is not the highest. WSM6 displays notably higher cloud tops and less anvil cloud, so it has the lowest MSSSIM, and the value is significantly different from the other five experiments. The other three simulations, NoML, SMAG-S, and SMAG-L, appear to have minimal changes to the Base because their influences are at small scales, at least for this case and the infrared channel. Nevertheless, compared to satellite data, all six experiments have successfully simulated the approximate location and intensity of convection at large scales without any significant biases.

4 EDR Estimation in Convection-Permitting Simulations

In this section, we evaluate the performance of different EDR estimation methods for convection-permitting simulations of Case 1 (Table 1) with the Base configuration of MPAS.

Figure 4 shows the results from the three methods and that from GTG3, which uses several indices related to upper-level turbulence, such as Frontogenesis function (isentropic coordinates), Ri (Richardson Number on dry air or moist air), and $|\text{Deformation}|^2/\text{Ri}$, and adopts a dynamic weighting method to obtain a comprehensive forecast. The GTG3 data source is from the World Area Forecast System (WAFS) of the National Weather Service, United States, with a resolution of 0.25°. Data at 00:00 UTC on May 21, 2020, are used because the closest GTG3 prediction is at 00:00 UTC.

SFSR and 2ndSF predict significant turbulence at the location of the CIT incident (red segment of the flight trajectory in

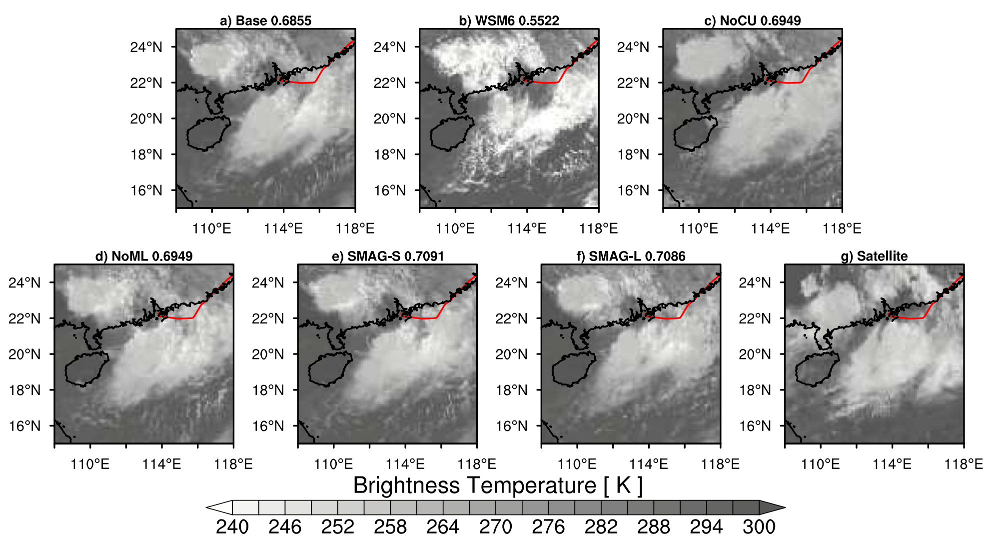

Fig. 3 Spatial distributions of the brightness temperature simulated by RTTOV for different experiments and observed by Himawari-8 for its Channel 7 on May 21, 2020, at 01:50 UTC and their corresponding MSSSIM values to observation. (a) Base, (b) WSM6, (C) NoCU, (d) NoML, (e) SMAG-S, (f) SMAG-L, and (g) satellite observation. The red line represents the route of the airplane in Case 1 in Table 1, the details about the turbulence are in Fig. 1

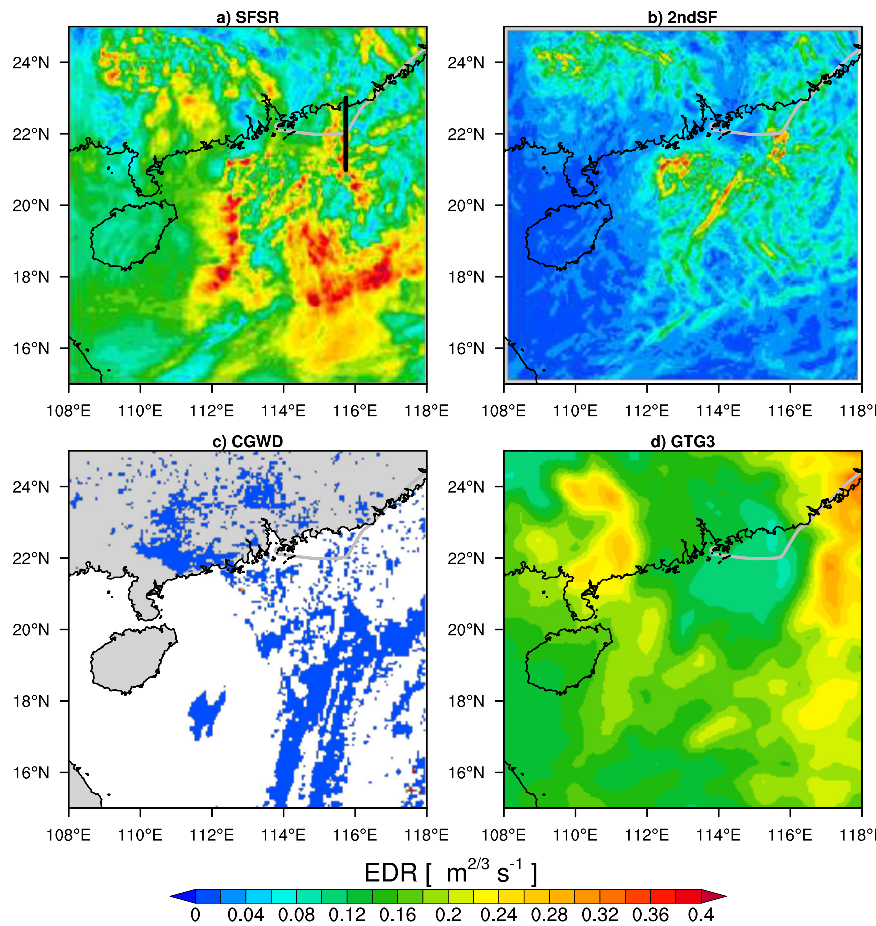

Fig. 4 Spatial distributions of the EDR calculated with different methods at 10 km altitude, on May 21, 2020, at 00:00 UTC. (a) is based on the subfilter-scale reconstruction, (b) second-order structure functions, (c) CGWD, and (d) GTG3 Forecast (27 km resolution). The lead time of GTG3 forecast is 12 hours. The source of the data is https://aviationweather.gov/wifs/. The gray line represents the trajectory of the airplane in Case 1 (Table 1). The details about the turbulence are in Fig. 1. The black line (21° to 23° N, 115.75° E) in (a) represents the cross-section in Fig. 5. and Fig. A1

Fig. 1), the result based on the 2ndSF (Fig. 4b) underestimates turbulence intensity. Because CGWD (Fig. 4c) is designed to be active above the cloud top but not in the cloud layer, it does not compute some high-EDR regions in the cloud and shows sparse data in the background. But it obtained some severe turbulence south of the coast, which is consistent with other methods. The SFSR method (Fig. 4a) yields the EDR most close to observation for EDR exceeding 0.1, while overestimating for lower observed EDR. The north-south EDR section in Fig. 5 (black line in Fig. 4a) shows both SFSR and 2ndSF capture strong EDR regions extending from 6 km to 12 km. As for CGWD, it does not calculate the EDR of the corresponding regions because many regions are located below the clouds. But it still captures turbulence

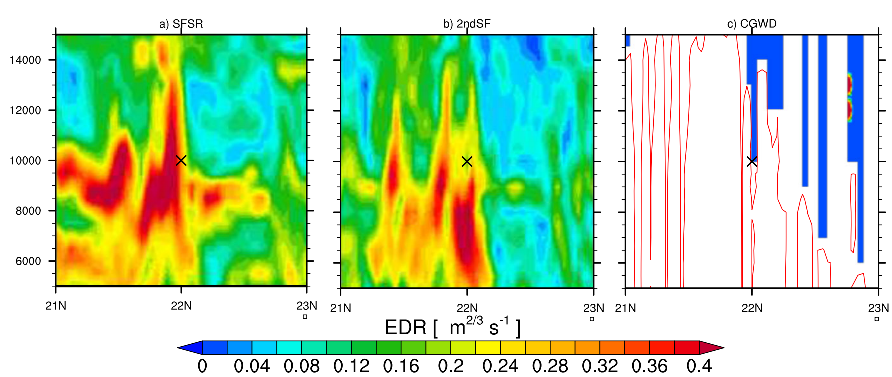

Fig. 5 Cross-section (black line in Fig. 4a, 21° to 23° N, 115.75° E) of the EDR calculated with different methods on May 21, 2020, at 00:00 UTC. (a) is based on the subfilter-scale reconstruction, (b) second-order structure functions, (c) CGWD. The "X" notation represents the position of the Case 1. The red lines represent the boundaries of the cloud. The white regions in (c) mean the CGWD can not be calculated because it is below the cloud top or there is no cloud in this horizontal grid

activities outside the clouds especially there are severe turbulence regions at slightly higher altitudes. Figure 12 shows the two-dimensional distribution of the heating rate and the minimum Richardson number in this section. At higher altitudes, although the area with a minimum Richardson number of less than 0.25 is relatively small, it still triggers severe turbulence due to wave breakings.

Compared with GTG3 results, all three methods can not calculate the turbulence area in the northeast quadrant of this domain compared to GTG3. The reason could be the inconsistencies in the prediction for the rapidly evolving convection between WAFS and our MPAS results. The GTG3 results using WAFS with different lead times show an apparent difference in the location of turbulence in this area. In contrast, the turbulence area in the southeast quadrant consistently exists among different WAFS experiments. SFSR calculates the southeast quadrant turbulence because it is related to vertical wind shear (See Appendix Fig. 15), which is not directly considered by the CGWD and 2ndSF. The differences in the spatial distribution of EDR shows that sensitivity of different methods to turbulence is inconsistent, and in practice, it may be necessary to combine them for use.

It should be noted that the separation in 2ndSF may be below the effective resolution of MPAS due to the implicit diffusion of the advection scheme (Skamarock et al. 2014). Therefore, we also did sensitivity tests on the separation length of the 2ndSF. The spatial distributions of the EDR with the variations of separation lengths from $7\Delta x$ to $15\Delta x$ in the 3 km mesh have similar large-scale patterns. Applying linear regression to equations (3) and (4) with varying separation distances seems to be a better method in this situation. However, drastic numerical changes in each cell can lead to many negative values and overestimation. Thus, we still used a separation length of $7\Delta x$ for our 2ndSF calculations.

Because the components of SFSR method rely on resolved velocities that can be changed with the variations of resolution, it is necessary to examine what resolution is sufficient for the EDR estimation and the performance under different resolutions. The calculations are identical for different resolutions, while the integral scale, L, should be varied based on the resolution. Figure 6 shows the spatial distributions of EDR calculated from the simulations with different resolutions in the refined region, from 1 km to 18 km. Those results exhibit a remarkable difference in EDR intensities and coverings, with the 18-km mesh simulation showing the most substantial turbulence areas and the 1-km mesh simulation showing the smallest strong EDR areas. As the resolution decreases, the RSFS component also decreases, while SGS will increase more, so the EDR value and turbulence areas will gradually increase. The maximum EDR values for the 1, 3, and 9-km mesh are approximately 0.55, while the maximum EDR value for the 18-km mesh reaches 0.75. This suggests that although calculating both SGS and RSFS components can improve the method's robustness under different kilometer-scale resolution meshes, for lower resolutions such as 18-km, the SFSR method may become inappropriate because overestimation will significantly occur.

It is tempting to suggest that those differences between Fig. 6a), b) and c) in intensity can be calibrated by adjusting the factor λ above. However, careful examination reveals that such tuning would not yield the same EDR pattern. For instance, in Fig. 6a, the high EDR region in the southeast quadrant is organized in a triangular shape with some wave patterns, but in Fig. 6c, the turbulence area found in the southeast quadrant has higher EDR values in the further east region. After all, when the resolution is coarsened, MPAS cannot resolve convection anymore, and systematic misrepresentation of convective weather systems could occur in the simulations.

The different EDR values at different resolutions also remind us of the turbulence intensity thresholds other researchers used previously. Sharman and Pearson (2017) defined EDR range of 0.15–0.22 as light turbulence, 0.22–0.34 as moderate, and $>$ 0.34 as severe for mid-size aircraft. However, based on observations, flights between Hong Kong and nearby destinations, such as Taipei, often report light turbulence when EDR reaches 0.1. This discrepancy with previous studies is probably due to the medium sizes of the aircraft on those routes (Sharman et al. 2014). Therefore, our study modified the threshold for light turbulence to $0.1\sim0.22$ (Sharman et al. 2014; ICAO 2010). Moreover, our EDR estimations need to be calibrated when there is sufficient data for the remapping or calibration.

Lastly, although the EDR calculation yields results closer to observations at the 1-km resolution, it demands much more computational resources (Table 3.). In our 1-km resolution simulation, integrating one time step (6 seconds model time) takes about 12 seconds wall-clock time when using 480 cores, and outputing the large files is equally time-consuming. The resulting wall-clock time for integrating MPAS for one hour is 1 hours. By contrast, the 3-km resolution simulation takes only 15 minutes of wall-clock time for the same task when using only 240 cores.

5 Sensitivity to Gray Zone-Related Parameterizations

Because the limited predictability of convection implies the need for ensemble forecast, here we evaluate how sensitive the CIT prediction is to the variation of physics and numerical schemes, which arguably represent potential sources of uncertainty other than initial conditions (Bouttier et al. 2012). Previous studies indeed suggested that in the gray zone, turbulence and convection representations could significantly

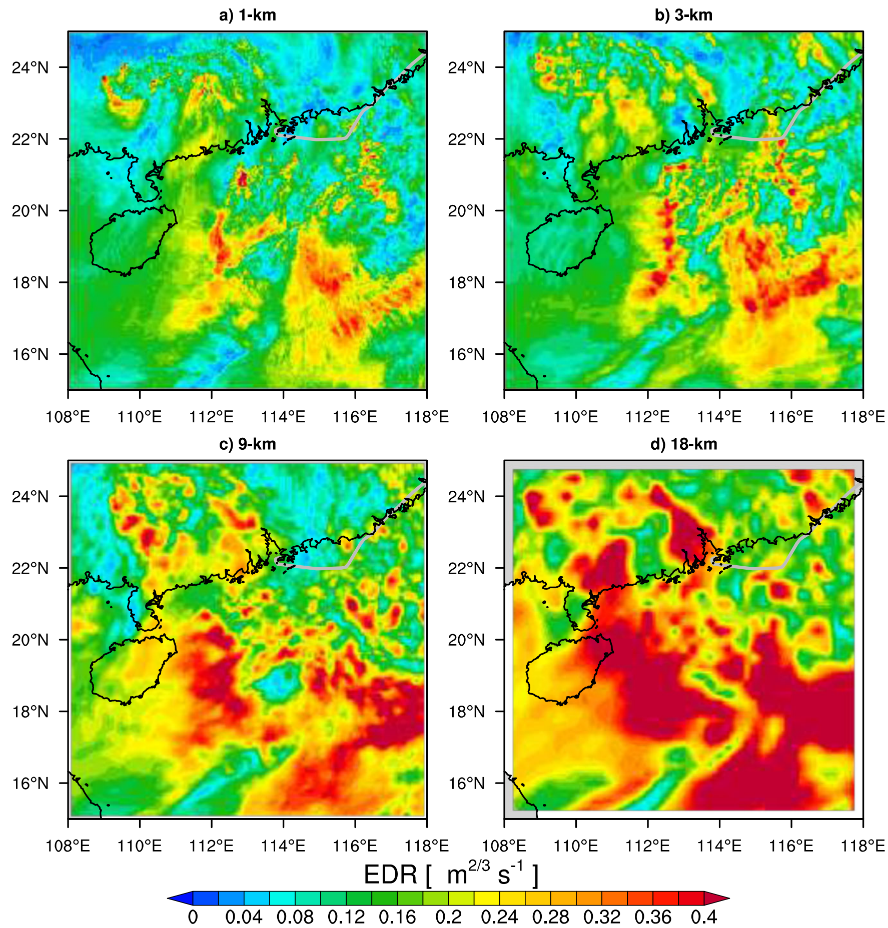

Fig. 6 Spatial distributions of the EDR at the altitude of 10 km, on May 21, 2020, at 01:50 UTC, calculated with the subfilter-scale reconstruction method for MPAS simulations with different resolutions: (a) 1–60 km, (b) 3–60 km, (c) 9–60 km, and (d) 18–60 km. The gray line represents the route of the airplane in Case 1 in Table 1. The details about the turbulence are in Fig. 1

change the development and intensity of convection (Shi et al. 2019; Shi and Wang 2022).

Figure 7 shows the spatial distribution of the EDR for Case 1 in those different experiments, with the EDR calculated using the SFSR method. The WSM6 simulation exhibits more intense turbulence at some locations, but the overall pattern differs from other simulations. In the southeast corner of some simulations, a band of high EDR oriented from southwest to northeast exists. However, in the WSM6 simulation, this band appears significantly smaller and weaker. Though having a pattern similar to most others, SMAG-S exhibits higher EDR, probably because the smaller eddy viscosity prevented strong dissipation due to the parameterized turbulence mixing. NoCU simulation exhibits some pretty localized regions of EDR maximum values near the coast, which better matches the airplane report of the CIT incident.

Figure 8 shows the evolution of horizontally averaged radar reflectivity factor in the whole plotted domain (15°N to 25°N, 108°E to 118°E) to include active convection in this refined region. All experiments show the development of convection to its peak intensity and then gradually decline. The WSM6 simulation has an earlier triggering of deep convection before 20:00 UTC on May 20; by 01:49 UTC on May 21, its convection has started decaying. Both the NoCU and Base experiments exhibit delayed convection triggering and peaking intensity. Notably, the NoCU experiment demonstrates more vigorous convection, resulting in more intense turbulence near the coastline, as illustrated in Fig. 7. WSM6 exhibits a notably lower reflectivity, with convection occurring at lower heights. This difference may be attributed to WSM6 underestimating precipitation particles, which results from a higher melting level (Min et al. 2015).

Table 3 Computational resources consumption for 15 hours model time integration using different MPAS meshes

Meshes(km)	Cells	Cores	Time (min)
1 ~ 60	2,827,196	480	890
3 ~ 60	835,586	240	210
9 ~ 60	293,533	240	37
18 ~ 60	207,915	240	29

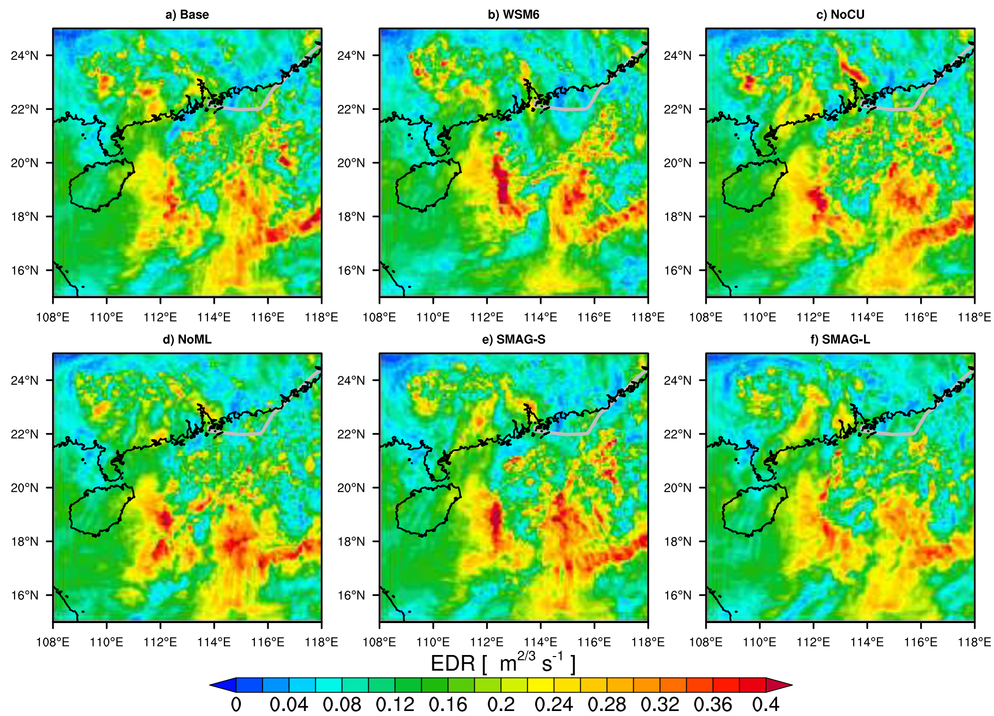

Fig. 7 Spatial distributions of the EDR calculated on different experiments at the altitude of 10 km, on May 21, 2020, at 01:50 UTC (a) Base, (b) WSM6, (C) NoCU, (d) NoML, (e) SMAG-S and (f) SMAG-L. The EDR here is calculated using the subfilter-scale reconstruction method. The gray line represents the route of the airplane in Case 1 in Table 1, the detailed locations of the encountered turbulence are in Fig. 1

This explanation also aligns with the WSM6's lower brightness temperature and the production of larger cloud areas at higher altitudes in Appendix Fig. 14.

To further illustrate the intensity and evolution of deep convection and its influence on turbulence, Fig. 9 shows the horizontal distribution of vertical velocity and the time series of the area-averaged EDR. Intense vertical velocity regions can indicate the location of strong convection, as shown in Fig. 3. From the satellite images (Fig. 3g), there is a wide spatial distribution of strong convection over the

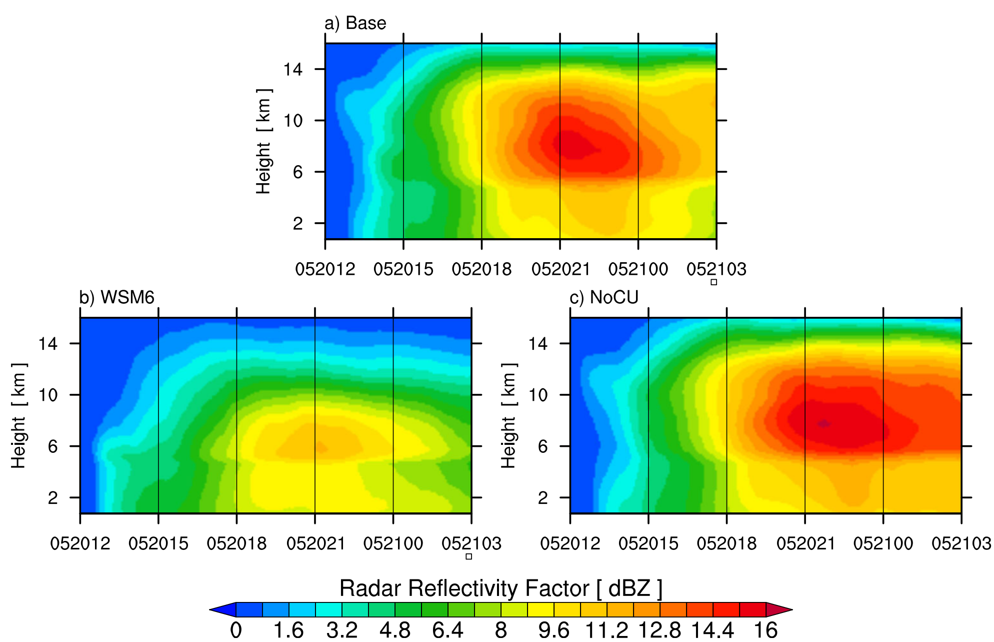

Fig. 8 Time series of radar reflectivity factor, averaged from 15°N to 25°N, 108°E to 118°E, from May 20, 2020, 12:00 UTC to May 21, 2020, 03:00 UTC, for a) Base, b) WSM6, and c) NoCU

ocean to the south of Hong Kong, which coincides with the turbulence distribution shown in Fig. 7. Notably, a convection system is present in the Base and NoCU experiments, as well as in the satellite images within the blue box in Fig. 9a and c. However, this convection system is absent in the WSM6 experiment, and consequently, the corresponding turbulence is also lower in WSM6. In addition, in the EDR-time series plot in Fig. 9d, WSM6 reaches its peak turbulence intensity earlier. The peaking time is consistent with the time of the strongest convection. The Base and NoCU also have this property during peak times. In terms of intensity, the results from Base (0.127) and WSM6 (0.126) are very similar, while NoCU (0.130) is higher than them. This magnitude relationship is also consistent with the intensity of convection.

It is worth noting that the excessive cloud in WSM6 at higher altitudes compared to Base and NoCU impacts practical CIT prediction. It is crucial to identify the turbulence outside the clouds for aviation turbulence. Figure 9e compares the fraction of out-of-cloud (10^{-5}kg/kg of total cloud condensate is set as the cloud boundary) turbulence, and Appendix Fig 14. shows the height of clouds in these experiments. At lower altitudes, the cloud spatial distributions of the three experiments are consistent. However, the clouds in the other two experiments almost disappear at higher altitudes, while in WSM6, they cover a larger area. As a result, out-of-cloud turbulence dominates at upper levels, with a fraction close to 100% in the Base and NoCU experiments. In contrast, the WSM6 simulation has only about 30% out-of-cloud turbulence at those levels. Thus, the choice of microphysics could produce qualitatively different CIT predictions in operational use.

6 Evaluation With Other Cases

6.1 EDR Distribution of Five Cases

The relatively accurate prediction of CIT in Case 1 presented above is not necessarily generalizable because different convective systems have different predictability challenges. In Fig. 10, the EDR distribution along the flight route is shown for all the five cases listed in Table 1 and the six experiments simulating each case with varied physics and numerics configurations. Since the CGWD method can not be applied below or inside the clouds because this method assumes that there is no convection-induced gravity in these regions. Therefore, the CGWD method's estimation is not comparable with the other two methods in the distribution of EDR and we only show the results from 2ndSF and RSFS.

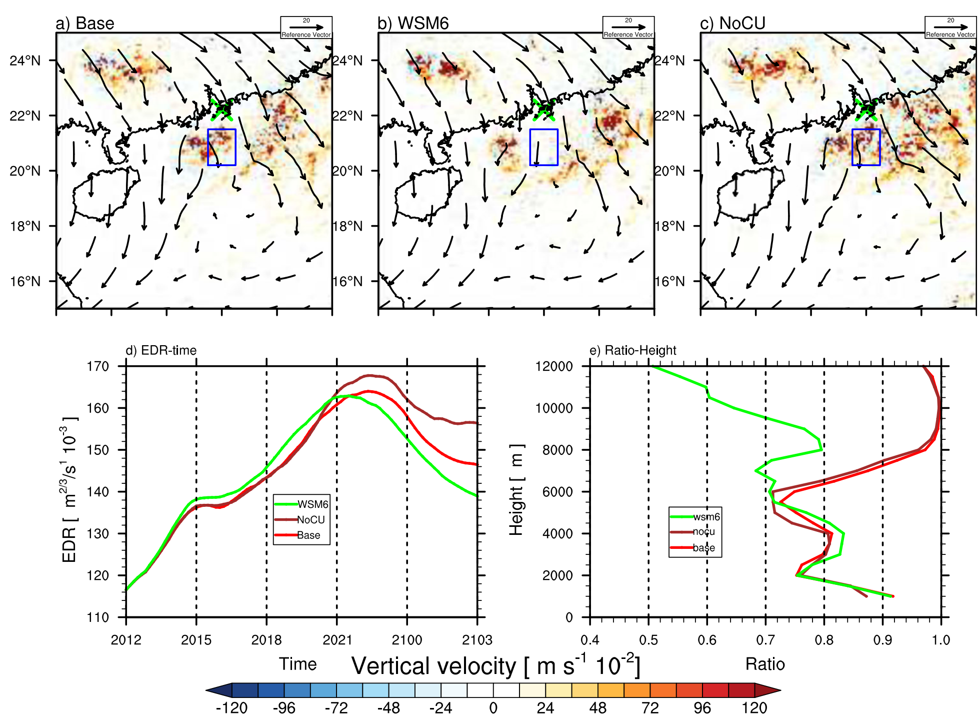

Fig. 9 Spatial distributions of the EDR at the altitude of 10 km, on May 21, 2020, at 01:50 UTC a) Base, b) WSM6, c) NoCU. d) Time series of EDR at the altitude of 10 km, averaged from 15°N to 25°N, 108°E to 118°E, from May 20, 2020, 12:00 UTC to May 21, 2020, 01:50 UTC. The green "X" notation represents the position of Hong Kong. The numbers in the X axis represent the date and hour. e) Vertical profile of the ratio between the area of turbulence that happens out of the cloud and the area of all the turbulence, including half an hour before and after the reporting time of Case 1. The thresholds of turbulence and cloud in e) are $0.10\mathrm{m}^{2/3}\mathrm{s}^{-1}$ of EDR and 10^{-5}kg/kg of cloud water mixing ratio and ice mixing ratio

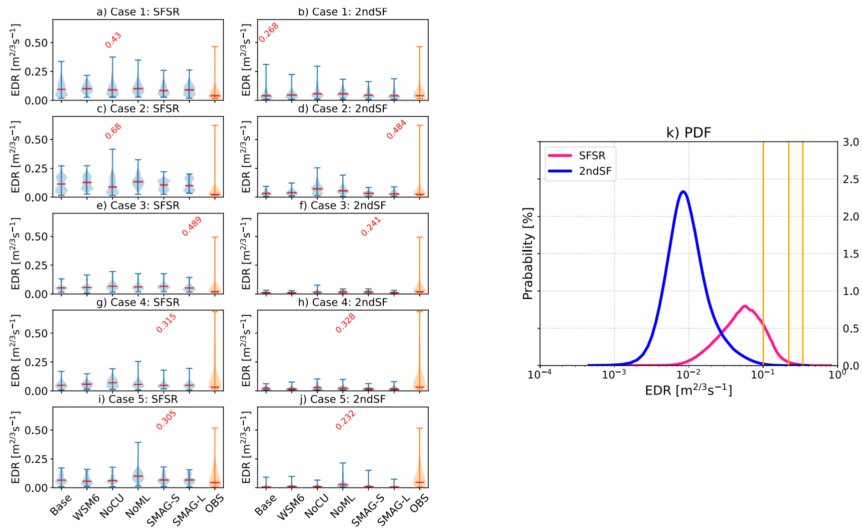

Fig. 10 Violin plot of the EDR within 10 km of the flight route in each case, including half an hour before and half an hour after the reporting time of CIT incidents, and collected according to airplanes' locations at the respective time. The rows correspond to the individual cases, from Case 1 to Case 5, while the columns represent the different methods, subfilter-scale reconstruction (SFSR) and second-order structure functions (2ndSF). In a panel, Different positions represent experiments with varied physics or numerics options. The observation distribution for the corresponding time is shown as the last violin in each panel. The red horizontal line in a violin represents the median of EDR in experiments or observations. The red numbers above the violins are the lowest Kolmogorov-Smirnov statistic among experiments when they are compared with observations. The altitude change of airplanes is considered in the sampling. k) Density plots of the EDR ($\mathrm{m}^{2/3}\mathrm{s}^{-1}$) from SFSR and 2ndSF. The data is from all experiments, from 15°N to 25°N, 105°E to 125°E and all altitudes. Three orange lines represent the thresholds 0.1, 0.22, and 0.34 for light, moderate, and severe turbulence, respectively

Firstly, the EDR data distributions of different methods are evaluated. Regarding 2ndSF and SFSR, their kernel density estimations resemble the observations, with a higher frequency of low EDR values and sporadic high EDR values. These distributions are similar even for Cases 3 and 4, where the forecasting performance is not good. Those poor performances in Cases 3 and 4 are probably not due to the methods' drawbacks. Instead, the errors are likely caused by biases in positions and morphology of convection.

When measuring the performance of different methods with the Kolmogorov-Smirnov (KS) statistic. The 2ndSF method appears slightly better than the SFSR method in that its best-member KS statistic values are lower than those of SFSR in three out of five cases since SFSR overestimates the EDR values which are lower than 0.1. However, the EDR results from the SFSR method have maxima closer to the observed maximum EDR values. The overall range of EDR estimated from the SFSR method is larger than that of 2nSF. Therefore, the SFSR method is likely to have a higher probability of detection for CIT events, and statistical evaluations for a longer-term period are required to confirm this. That being said, if we choose ensemble members producing higher EDR maxima (e.g., NoCU, NoML, and Base), the prediction for maximum EDR appears acceptable for Cases 1, 2, and 5, but substantial underestimation exists in the other two cases.

NoCU usually exhibits more substantial turbulence than others, probably because the GF convection scheme, though scale-aware, still stabilizes the atmosphere too much and thereby weakens resolved convective motions, lowering EDR estimation in the SFSR. NoML can also generate relatively high EDR in some cases. The monotonic limiter helps advection schemes avoid generating new local extremes due to numerical errors but can also attenuate physical extremes. Thus, turning it off seems beneficial in some regimes.

6.2 Distribution of EDR and Thresholds

Sharman and Pearson (2017) suggest that a lognormal distribution is an essential characteristic of a diagnostic to be included in GTG3, and EDR should also be a lognormal distribution in the nature (Nastrom and Gage 1985). Although there may still be bias because the results are from specific locations and time (Sharman and Pearson 2017), we used the EDR results from our 30 simulations at all altitudes and time points in the convection-permitting area to obtain distributions for SFSR and 2ndSF. Figure 10k reveals that 2ndSF results exhibit a right-skewed lognormal distribution, while SFSR results show a left-skewed one. The results of 2ndSF have a smaller variance and a higher peak probability density, and their median and mean are smaller than that of SFSR. The probability of EDR values below 0.1 is approximately 67% in SFSR. This probability is lower than previous statistics in North America from United Airlines and Delta Air Lines (Sharman et al. 2014) due to the slightly overestimation. For EDR higher than 0.5, the probability is 4×10^{-5}. This probability has the same magnitude with the previous results

(2×10^{-5}.) Sharman et al. (2014), indicating that for extreme turbulence, the EDR obtained based on the SFSR method with 3-km mesh can capture the statistical characteristics of the observed EDR. The SFSR's probability density function shape is closer to the previous observation results (Sharman and Lane 2016). Although both the SFSR and 2ndSF methods have approximately a lognormal distribution, it is evident that SFSR has better statistical characteristics.

6.3 Probability of Detection and False Alarms

The Probability of Detection (POD) and False Alarm Rate (FAR) are significant metrics in evaluating a prediction method. To further quantify the performance of the SFSR method, we made scatter plots between EDR data using different methods and observations. Muñoz-Esparza et al. (2018) indicated skilled pointwise forecasts are not expected from NWP. Since we have high-resolution mesh and observations have limited spatial coverage, we coarsened the mesh and extracted data within 50 km distance from the aircraft. However, extracting the average value from coarsened cells will decrease the numerical values of EDR. So when half of the cells in the region within the 50-km radius can reach the value of turbulence (EDR > 0.1), the maximum values are selected; otherwise, the average values are selected. For each time of a reported turbulence incident, the best member EDR (closest to observation) (Fig. 11a) or the average of members (Fig. 11b) is selected from six members. Moreover, only Cases 1, 2, and 5, for which MPAS has good performance, are used in the analysis. Figure 11 shows the results. Overall, SFSR estimation has a higher correlation coefficient than the 2ndSF results. In the low EDR region (0–0.1), the overall values of SFSR are higher. In the region with turbulence (EDR$>$ 0.1), the EDR values from SFSR are closer to the observation. Both SFSR and 2ndSF exhibit low bias for extreme turbulence, yet SFSR has higher values than the 2ndSF. The apparent higher correlation coefficient of SFSR in predicting turbulence results also supports this conclusion. In Fig. 11b, the values of higher EDR are decreased because we took the unweighted average, while overestimation is evident in the low EDR region. Hence, the correlation coefficients in both methods are low.

The results also show the difference in POD and FAR between SFSR and 2ndSF. The relative operating characteristic (ROC) curves in Fig. 11d suggest that these two methods are superior to random guesses when the best member of the ensemble is used (i.e., optimal in Fig. 11), and SFSR has a relatively better performance with a larger area under the curve (AUC). Meanwhile, similar to Fig. 11b, the AUCs of both methods in Fig. 11e significantly decrease when we use averaged EDR from the physical perturbation-based ensemble of each case. Although SFSR performs better, its AUC is only slightly higher than 0.5, which is not very valuable in predicting turbulence. This result indicates that averaging ensemble members' estimations without weighting may not be suitable for predicting aviation turbulence. As for the

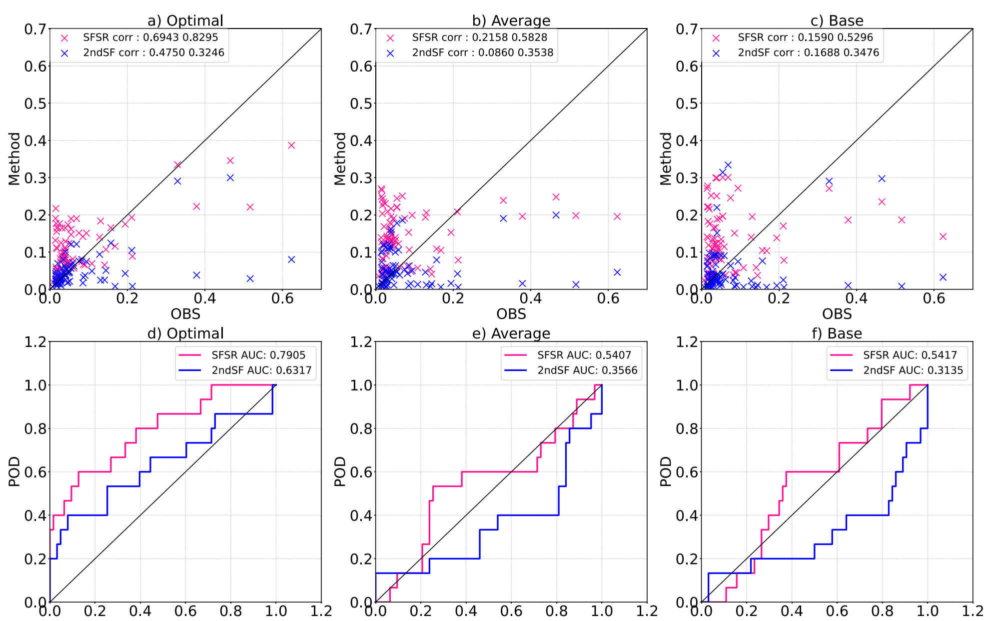

Fig. 11 Scatter plots of the EDR between observation (X-axis) and two methods (Y-axis) and ROC curves from Cases 1, 2, and 5. Curves are constructed based on two methods with an observational threshold of EDR $= 0.1\mathrm{m}^{2/3}\mathrm{s}^{-1}$. When half of the cells in the coarse cell reach turbulence, the maximum values are selected, and if there is no report, the average value is selected. At the time of a report, the optimal members from the physical perturbation-based ensemble of each case were selected from six members in a) and d), while the averages of members were used in b) and e). The results from Base were used in c) and f)

results of the Base experiment selected in Fig. 11c and f, its indices are close to the average results, which implies ensemble forecasting is required to provide sufficient spread to capture aviation turbulence.

7 Summary and Discussion

This study uses the non-hydrostatic, variable-resolution MPAS to predict convectively induced turbulence. The MPAS mesh has a refined 3-km resolution region that covers the South China continent and the South China Sea and is centered in Hong Kong. The convection-permitting mesh is beneficial for predicting aviation turbulence. Instead of using large-scale diagnostic quantities, we can calculate EDR by computing the TKE from the resolved motions. We compared three methods to calculate EDR from convection-permitting simulation output. One of those methods is developed in this study and named the sub-filter scale reconstruction (SFSR). It employs the framework of explicit filtering and reconstruction in turbulence parameterization and estimates resolvable subfilter-scale TKE and subgrid scales TKE, which are then used to calculate EDR. The SFSR method is designed different from 2ndSF and CGWD and it is more sensitive to the turbulence triggered from vertical shear. In MPAS experiments, the SFSR can capture some turbulence cases which are observed by the airplanes.

From the perspective of the process of EDR calculation and results, it can be found that different methods can capture turbulence based on different mechanisms. 2ndSF is more sensitive to turbulence near strong convection, it can also capture the turbulence where vertical motions are negligible near active convection (Barber et al. 2019), CGWD is more localized and adept at evaluating turbulence generated by gravity wave breaking above clouds (Kim et al. 2019), and SFSR captures most of the energy in turbulent flow in the convection-permitting mesh (Carati et al. 2001). In the analysis of Case 1, all three methods can capture turbulence near strong convection, but only SFSR can predict turbulence with strong vertical shear. In Case 5, the aircraft was searching for a landing route below or inside of the cloud at low altitudes. CGWD is not effective in this situation, while SFSR and 2ndSF can predict turbulence and SFSR can provide EDR values closer to observations. In general, CGWD is designed to predict certain sources of CIT so that other clear-air turbulence (CAT) and mountain-wave turbulence (MWT) methods are used together to capture the aviation turbulence (Kim et al. 2019). 2ndSF is frequently used for CAT, although (Barber et al. 2019) applied it to CIT successfully, other indices based on CAT and MWT are still combined for forecasting (Sharman and Pearson 2017). This demonstrates the more generic application potential of SFSR in forecasting CIT. However, the potentials of this method still needs to be tested in cases of MWT and CAT.

Because the SFSR method relies on the resolved velocity field to estimate TKE, we assessed its dependency on MPAS resolutions. Testing with refined region resolution of 1, 3, 9, and 18 km shows that higher resolution simulations provide better EDR estimation regarding intensity and spatial coverage and reduce the overestimation. The reason is that the SFSR method was originally designed for large eddy simulation, and some assumptions and parameters are best suited to large eddy simulation scales. Therefore, we still need to optimize those parameters within the SFSR method under different resolution conditions in our future work. Although the results from SFSR are better at higher resolutions, increasing the resolution also substantially increases computational costs. The 3-km resolution is a balance between accuracy and computational resource demand. Thus, we use it for other evaluations in our study.

Convection-permitting resolutions are in the gray zone for turbulence and convection parameterization. Therefore, the choices of relevant physical parameterizations and numerical schemes comprise a fundamental source of uncertainty in predicting convectively induced turbulence. We examined some scheme variations available in MPAS and found that such a physics and numerics perturbation-based ensemble effectively captures some convection stochasticity. However, among those variations, the choice of microphysics and cumulus convection schemes exhibit more impact on the predicted convection. Compared to the Thompson microphysics scheme, the initiation and intensification of convection occurred earlier in the simulation using WSM6, which also caused higher cloud tops. Switching off the scale-aware GF convection scheme resulted in more intense turbulence and a prolonged convective system. Furthermore, we observed a strong correlation between the intensity and evolution of turbulence with convection, emphasizing the necessity of accurate simulations in convection systems for turbulence forecasting.

Further testing with more CIT cases showed that for both the distribution and maximum values of EDR, the SFSR method can provide results closer to observations and statistics properties similar to the previous studies. However, the overestimation of EDR with lower values also needs to be addressed, in other words, SFSR needs to become more localized like CGWD. More observations should be included to enable EDR remapping to reduce avoidance bias, which is probably the cause of the statistical differences between SFSR estimation and observation (Sharman and Pearson 2017). For some convective systems, significant location bias exists in the convection-permitting simulations. The physics perturbation-based ensemble has its limitation in generating enough ensemble spread. We will test the effectiveness of combining physics and initial condition perturbation-based ensemble and evaluate related post-processing methods for providing accurate predictions in the future.

Appendix A: Heating Rate and Ri_{min}

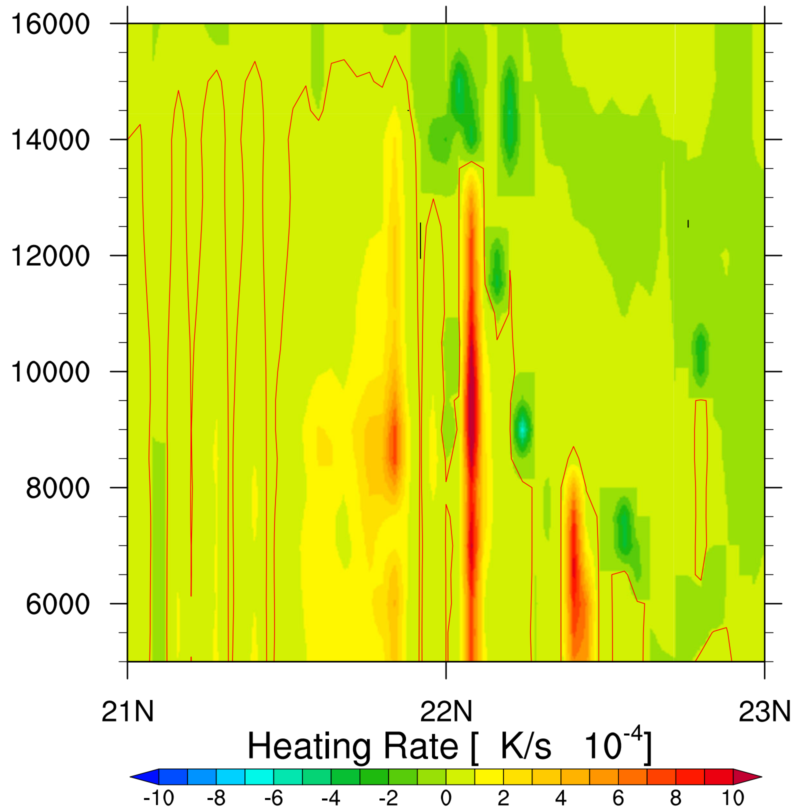

Fig. 12 Cross-sections (black line in Fig. 4a, 21° to 23° N, 115.75° E) of the summation of the heating rates and $\mathbf{Ri}_{min}$ on May 21, 2020, at 00:00 UTC. The red lines represent the boundaries of the cloud. The black lines represent the areas where the $\mathbf{Ri}_{min}$ is smaller than 0.25

Appendix B: Filtering on Hexagon Mesh

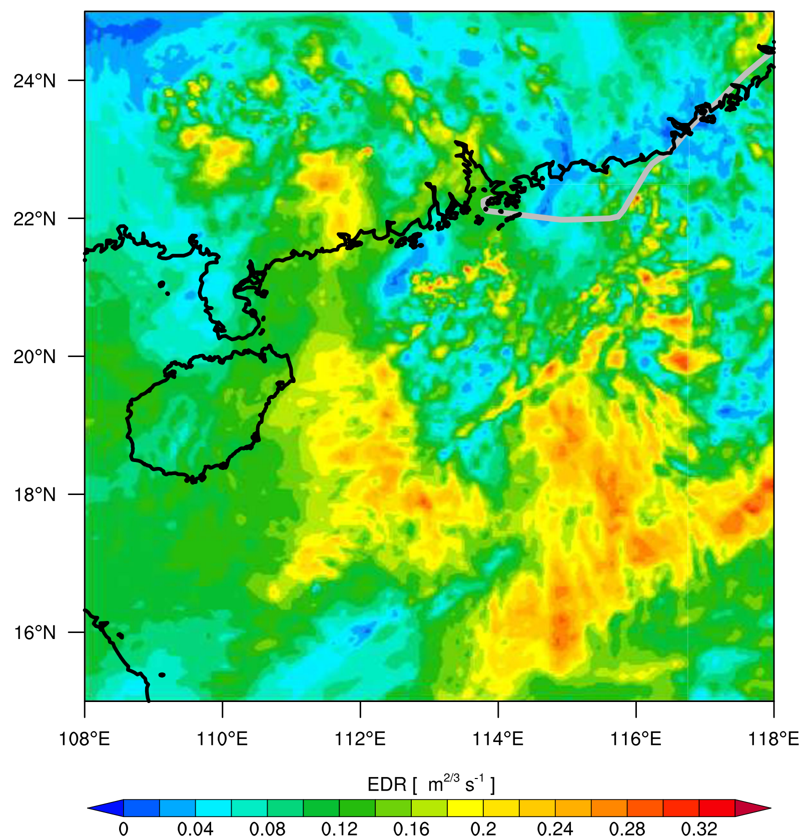

Fig. 13 Spatial distributions of the SFSR-EDR of Case 1 at May 21, 2020 01:50 a.m at the altitude of 10 km. Applying the subfilter-scale reconstruction method in original hexagon mesh. The gray line represents the route of the airplane in Case 1 in Table 1, the details about the turbulence are in Fig. 1

Appendix C: The Influence of WSM6 to Cloud Top Height

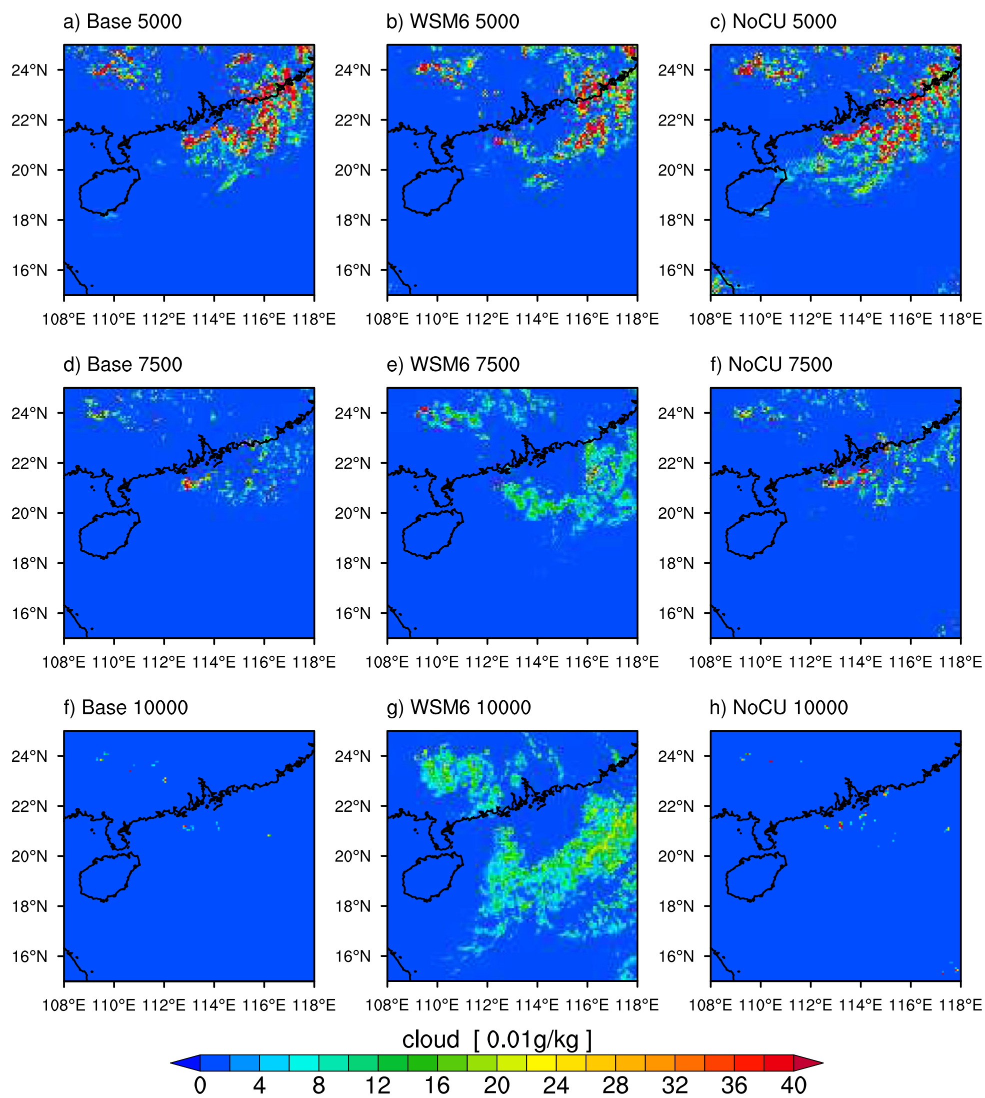

Fig. 14 Spatial distributions of the cloud in different altitudes with different options at May 21, 2020 01:50 a.m

Appendix D: Vertical Shear in Case1

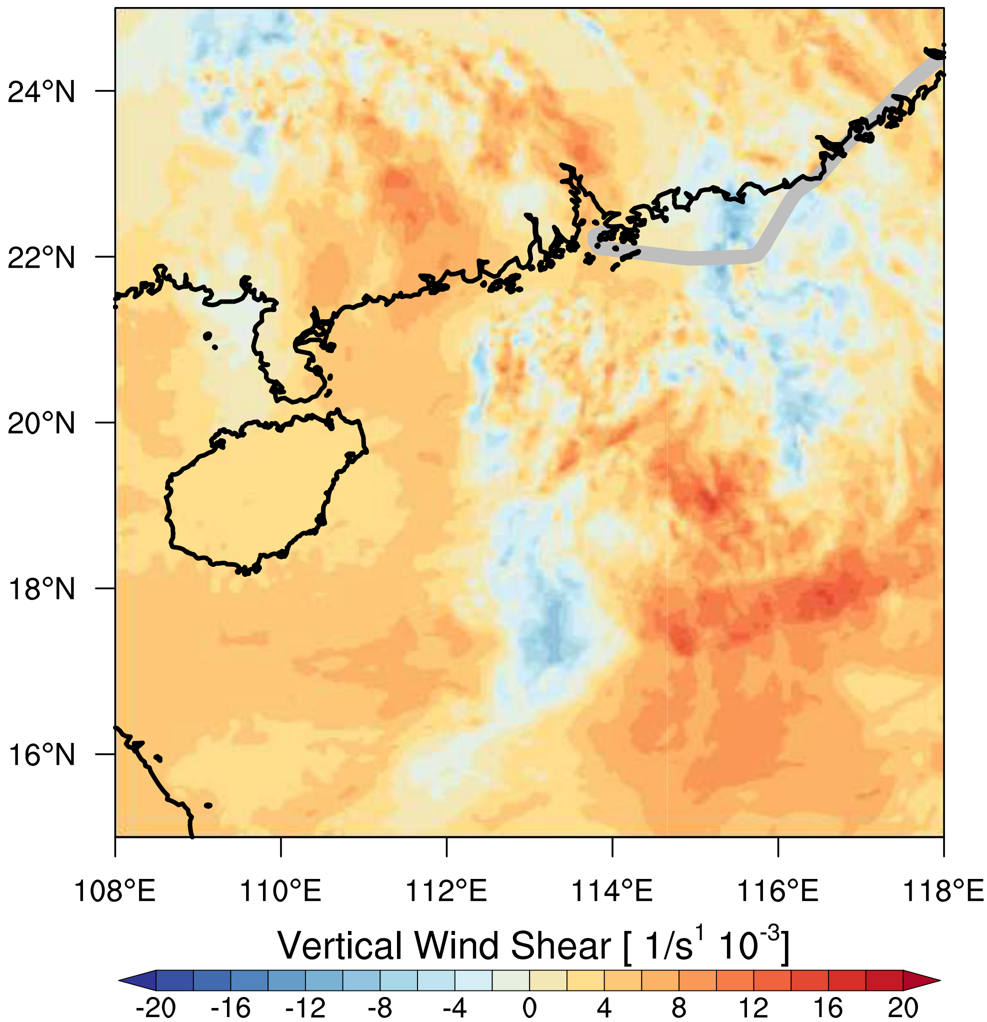

Fig. 15 Spatial distributions of the veritical wind shear at 10,000 m at May 21, 2020 00:00 UTC

Acknowledgements The authors thank anonymous reviewers and the editor, Dr. Hyeyum Hailey Shin, for their valuable comments and suggestions, which substantially improved our manuscript. The project is part of the Aviation Research and Development Project Phase 2 (AvRDP2), supported by the World Meteorological Organization. H.C. and X.S. were partially supported by the Research Grant Council (RGC) projects HKUST-16301322 and AoE/P-601/23-N. The authors thank HKUST Fok Ying Tung Research Institute and National Supercomputing Center in Guangzhou Nansha sub-center for providing high-performance computational resources.

Funding Open access funding provided by Hong Kong University of Science and Technology.

Data Availability The technology to generate the meshes can be found in http://mpas-dev.github.io/MPAS-Tools/stable/mesh_creation.html#building-a-jigsaw-mesh. The ERA5 data can be downloaded on https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.bd0915c6?tab=overview. The namelist file and codes to calculate the EDR with different methods can be found in https://doi.org/10.5281/zenodo.8092926.

References

Ahmad, N., Proctor, F.: Estimation of eddy dissipation rates from mesoscale model simulations (2012). https://doi.org/10.2514/6.2012-429

Allen, J.D.: Perfect Reconstruction Filter Banks for the Hexagon Grid. In: 2005 5th International conference on information communications & signal processing, pp. 73–76. IEEE, Bangkok, Thailand (2005). https://doi.org/10.1109/ICICS.2005.1689007

Barber, K.A., Deierling, W., Mullendore, G., Kessinger, C., Sharman, R., Muñoz-Esparza, D.: Properties of convectively induced turbulence over developing oceanic convection. Mon. Weather Rev. **147**(9), 3429–3444 (2019). https://doi.org/10.1175/MWR-D-18-0409.1

Bell, B., Hersbach, H., Simmons, A., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Soci, C., Villaume, S., Bidlot, J.-R., Haimberger, L., Woollen, J., Buontempo, C., Thépaut, J.-N.: The ERA5 global reanalysis: preliminary extension to 1950. Q. J. R. Meteorol. Soc. **147**(741), 4186–4227 (2021). https://doi.org/10.1002/qj.4174

Bouttier, F., Vié, B., Nuissier, O., Raynaud, L.: Impact of stochastic physics in a convection-permitting ensemble. Mon. Weather Rev. **140**(11), 3706–3721 (2012). https://doi.org/10.1175/MWR-D-12-00031.1

Brown, D., Brownrigg, R., Haley, M., Huang, W.: NCAR Command Language (NCL). UCAR/NCAR - Computational and Information Systems Laboratory (CISL) (2012). https://doi.org/10.5065/D6WD3XH5

Carati, D., Winckelmans, G.S., Jeanmart, H.: On the modelling of the subgrid-scale and filtered-scale stress tensors in large-eddy simulation. J. Fluid Mech. **441**, 119–138 (2001)

Chen, F., Dudhia, J.: Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity. Monthly Weather Rev. **129**(4), 569–585 (2001). https://doi.org/10.1175/1520-0493(2001)129h0569:CAALSHi2.0.CO;2

Chow, F.K., Street, R.L., Xue, M., Ferziger, J.H.: Explicit filtering and reconstruction turbulence modeling for large-eddy simulation of neutral boundary layer flow. J. Atmos. Sci. **62**(7), 2058–2077 (2005). https://doi.org/10.1175/JAS3456.1

Chun, H.-Y., Baik, J.-J.: Momentum flux by thermally induced internal gravity waves and its approximation for large-scale models. J. Atmos. Sci. **55**(21), 3299–3310 (1998)

Dutton, M.: Probability forecasts of clear-air turbulence based on numerical model output. probability forecasts of clear-air turbulence based on numerical model output (1980)

Ellrod, G.P., Knox, J.A.: Improvements to an operational clear-air turbulence diagnostic index by addition of a divergence trend term. Weather Forecast. **25**(2), 789–798 (2010). https://doi.org/10.1175/2009WAF2222290.1

Endlich, R.M.: The Mesoscale Structure of Some Regions of Clear-Air Turbulence. J. Appl. Meteorol. **3**(3), 261–276 (1964). https://doi.org/10.1175/1520-0450(1964)003<0261:TMSOSR>2.0.CO;2

Frehlich, R., Sharman, R.: Estimates of turbulence from numerical weather prediction model output with applications to turbulence diagnosis and data assimilation. Mon. Weather Rev. **132**(10), 2308–2324 (2004). https://doi.org/10.1175/1520-0493(2004)132h2308:EOTFNWi2.0.CO;2

Frehlich, R., Sharman, R.: Climatology of velocity and temperature turbulence statistics determined from rawinsonde and ACARS/AMDAR data. J. Appl. Meteorol. Climatol. **49**(6), 1149–1169 (2010). https://doi.org/10.1175/2010JAMC2196.1

Grell, G.A., Freitas, S.R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmos. Chem. Phys. **14**(10), 5233–5250 (2014). https://doi.org/10.5194/acp-14-5233-2014

Gullbrand, J., Chow, F.K.: The effect of numerical errors and turbulence models in large-eddy simulations of channel flow, with and without explicit filtering. Journal of Fluid Mechanics **495**, 323–341 (2003). https://doi.org/10.1017/S0022112003006268

Hagos, S., Leung, L.R., Yang, Q., Zhao, C., Lu, J.: Resolution and dynamical core dependence of atmospheric river frequency in global model simulations. J. Clim. **28**(7), 2764–2776 (2015). https://doi.org/10.1175/JCLI-D-14-00567.1

Hong, S.-Y.: A new stable boundary-layer mixing scheme and its impact on the simulated East Asian summer monsoon. Q. J. R. Meteorol. Soc. **136**(651), 1481–1496 (2010). https://doi.org/10.1002/qj.665

Hong, S.-Y., Kim, J.-H., Lim, J.-O., Dudhia, J.: The WRF single moment microphysics scheme (WSM). J. Korean Meteorological Soc. **42**, 129–151 (2006)

Iacono, M.J., Mlawer, E.J., Clough, S.A., Morcrette, J.-J.: Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR community climate model, CCM3. J. Geophys. Res.: Atmos. **105**(D11), 14873–14890 (2000). https://doi.org/10.1029/2000JD900091

ICAO: Meteorological service for international air navigation. Int. Civ (2010)

Kim, S.H., Chun, H.Y., Lee, D.B., Kim, J.H., Sharman, R.D.: Improving numerical weather prediction-based near-cloud aviation turbulence forecasts by diagnosing convective gravity wave breaking. Weather Forecasting (5), 36 (2021)

Kim, S.-H., Chun, H.-Y.: Aviation turbulence encounters detected from aircraft observations: spatiotemporal characteristics and application to Korean Aviation Turbulence Guidance: Aviation turbulence

encounters detected from aircraft observations. Meteorol. Appl. **23**(4), 594–604 (2016). https://doi.org/10.1002/met.1581
Kim, S.-H., Chun, H.-Y., Sharman, R.D., Trier, S.B.: Development of near-cloud turbulence diagnostics based on a convective gravity wave drag parameterization. J. Appl. Meteorol. Climatol. **58**(8), 1725–1750 (2019). https://doi.org/10.1175/JAMC-D-18-0300.1
Kolmogorov, A.N.: The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. Proc.: Math. Phys. Sci. **434**(1890), 9–13 (1991). arXiv:51980
Landu, K., Leung, L.R., Hagos, S., Vinoj, V., Rauscher, S.A., Ringler, T., Taylor, M.: The dependence of ITCZ structure on model resolution and dynamical core in aquaplanet simulations. J. Clim. **27**(6), 2375–2385 (2014). https://doi.org/10.1175/JCLI-D-13-00269.1
Lane, T.P., Sharman, R.D., Clark, T.L., Hsu, H.-M.: An investigation of turbulence generation mechanisms above deep convection. J. Atmos. Sci. **60**(10), 1297–1321 (2003). https://doi.org/10.1175/1520-0469(2003)60(1297:AIOTGM)2.0.CO;2
Lane, T.P., Sharman, R.D.: Some influences of background flow conditions on the generation of turbulence due to gravity wave breaking above deep convection. J. Appl. Meteorol. Climatol. **47**(11), 2777–2796 (2008)
Lane, T.P., Sharman, R.D.: Intensity of thunderstorm-generated turbulence revealed by large-eddy simulation. Geophys. Res. Lett. **41**(6), 2221–2227 (2014). https://doi.org/10.1002/2014GL059299
Lane, T.P., Doyle, J.D., Sharman, R.D., Shapiro, M.A., Watson, C.D.: Statistics and dynamics of aircraft encounters of turbulence over Greenland. Mon. Weather Rev. **137**(8), 2687–2702 (2009). https://doi.org/10.1175/2009MWR2878.1
Lee, D.-B., Chun, H.-Y., Kim, S.-H., Sharman, R.D., Kim, J.-H.: Development and evaluation of global korean aviation turbulence forecast systems based on an operational numerical weather prediction model and in situ flight turbulence observation data. Weather Forecast. **37**(3), 371–392 (2022)
Li, G., Chen, H., Xu, M., Zhao, C., Zhong, L., Li, R., Fu, Y., Gao, Y.: Impacts of topographic complexity on modeling moisture transport and precipitation over the Tibetan Plateau in summer. Advan. Atmos. Sci. **39**, 1151–1166 (2022). https://doi.org/10.1007/s00376-022-1409-7
Lindborg, E.: Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? J. Fluid Mech. **388**, 259–288 (1999). https://doi.org/10.1017/S0022112099004851
Lindzen, R.S.: Turbulence and stress owing to gravity wave and tidal breakdown. J. Geophys. Res. Oceans **86** (1981)
Min, K.-H., Choo, S., Lee, D., Lee, G.: Evaluation of WRF cloud microphysics schemes using radar observations. Weather Forecasting **30**(6), 1571–1589 (2015). https://doi.org/10.1175/WAF-D-14-00095.1. Chap. Weather and Forecasting
Mirocha, J.D., Lundquist, J.K., Kosović, B.: Implementation of a nonlinear subfilter turbulence stress model for large-eddy simulation in the advanced research wrf model. Monthly Weather Rev. **138**(11), 4212–4228 (2009)
Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. J. Geophys. Res.: Atmos. **102**(D14), 16663–16682 (1997). https://doi.org/10.1029/97JD00237
Moeng, C.-H., Wyngaard, J.C.: Spectral analysis of large-eddy simulations of the convective boundary layer. J. Atmos. Sci. **45**(23), 3573–3587 (1988)
Mohan, P.R., Srinivas, C.V., Yesubabu, V., Baskaran, R., Venkatraman, B.: Tropical cyclone simulations over Bay of Bengal with ARW model: sensitivity to cloud microphysics schemes. Atmos. Res. **230**, 104651 (2019). https://doi.org/10.1016/j.atmosres.2019.104651
Monin, A.S., Yaglom, A.M.: Statistical Fluid Mechanics, Volume II: Mechanics Of Turbulence. Courier Corporation (2013)
Muñoz-Esparza, D., Sharman, R.D., Deierling, W.: Aviation turbulence forecasting at upper levels with machine learning techniques based on regression trees. J. Appl. Meteorology Climatology **59**(11), 1883–1899 (2020). https://doi.org/10.1175/JAMC-D-20-0116.1. Chap. Journal of Applied Meteorology and Climatology
Muñoz-Esparza, D., Kosović, B.: Generation of inflow turbulence in large-eddy simulations of nonneutral atmospheric boundary layers with the cell perturbation method. Mon. Weather Rev. **146**(6), 1889–1909 (2018). https://doi.org/10.1175/MWR-D-18-0077.1
Muñoz-Esparza, D., Sharman, R., Sauer, J., Kosović, B.: Toward low-level turbulence forecasting at eddy-resolving scales. Geophys. Res. Lett. **45**(16), 8655–8664 (2018). https://doi.org/10.1029/2018GL078642
Nakanishi, M., Niino, H.: Development of an improved turbulence closure model for the atmospheric boundary layer. J. Meteorological Soc. Japan. Ser. II **87**(5), 895–912 (2009). https://doi.org/10.2151/jmsj.87.895
Nastrom, G.D., Gage, K.S.: A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. J. Atmos. Sci. **42**(9), 950–960 (1985). https://doi.org/10.1175/1520-0469(1985)042<0950:ACOAWS>2.0.CO;2
O'Brien, T.A., Li, F., Collins, W.D., Rauscher, S.A., Ringler, T.D., Taylor, M., Hagos, S.M., Leung, L.R.: Observed scaling in clouds and precipitation and scale incognizance in regional to global atmospheric models. J. Clim. **26**(23), 9313–9333 (2013). https://doi.org/10.1175/JCLI-D-13-00005.1
Sakaguchi, K., Leung, L.R., Zhao, C., Yang, Q., Lu, J., Hagos, S., Rauscher, S.A., Dong, L., Ringler, T.D., Lauritzen, P.H.: Exploring a multiresolution approach using AMIP simulations. J. Clim. **28**(14), 5549–5574 (2015). https://doi.org/10.1175/JCLI-D-14-00729.1
Schumann, U.: Subgrid length-scales for large-eddy simulation of stratified turbulence. Theoret. Comput. Fluid Dyn. **2**(5), 279–290 (1991). https://doi.org/10.1007/BF00271468
Sharman, R., Lane, T. (eds.): Aviation Turbulence. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-23630-8
Sharman, R.D., Pearson, J.M.: Prediction of energy dissipation rates for aviation turbulence. Part I: forecasting nonconvective turbulence. Journal of Applied Meteorology and Climatology **56**(2), 317–337 (2017). https://doi.org/10.1175/JAMC-D-16-0205.1
Sharman, R., Tebaldi, C., Wiener, G., Wolff, J.: An integrated approach to mid- and upper-level turbulence forecasting. Weather Forecast. **21**(3), 268–287 (2006). https://doi.org/10.1175/WAF924.1
Sharman, R.D., Doyle, J.D., Shapiro, M.A.: An investigation of a commercial aircraft encounter with severe clear-air turbulence over Western Greenland. J. Appl. Meteorol. Climatol. **51**(1), 42–53 (2012). https://doi.org/10.1175/JAMC-D-11-044.1
Sharman, R.D., Cornman, L.B., Meymaris, G., Pearson, J., Farrar, T.: Description and derived climatologies of automated In Situ Eddy-Dissipation-Rate reports of atmospheric turbulence. J. Appl. Meteorol. Climatol. **53**(6), 1416–1432 (2014). https://doi.org/10.1175/JAMC-D-13-0329.1
Shi, X., Wang, Y.: Impacts of cumulus convection and turbulence parameterizations on the convection-permitting simulation of typhoon precipitation. Monthly Weather Rev. **150** (2022). https://doi.org/10.1175/MWR-D-22-0057.1
Shi, X., Wang, Y.: Impacts of cumulus convection and turbulence parameterizations on the convection-permitting simulation of typhoon precipitation. Mon. Weather Rev. **150**(11), 2977–2997 (2022). https://doi.org/10.1175/MWR-D-22-0057.1
Shi, X., Chow, F.K., Street, R.L., Bryan, G.H.: Key elements of turbulence closures for simulating deep convection at kilometer-scale

resolution. J. Advan. Model. Earth Syst. **11**(3), 818–838 (2019). https://doi.org/10.1029/2018MS001446

Skamarock, W.C., Park, S.-H., Klemp, J.B., Snyder, C.: Atmospheric Kinetic Energy Spectra from Global High-Resolution Nonhydrostatic Simulations. J. Atmos. Sci. **71**(11), 4369–4381 (2014). https://doi.org/10.1175/JAS-D-14-0114.1. Chap. Journal of the Atmospheric Sciences

Skamarock, W.C.: Evaluating mesoscale NWP models using kinetic energy spectra. Mon. Weather Rev. **132**(12), 3019–3032 (2004). https://doi.org/10.1175/MWR2830.1. Chap. Monthly Weather Review

Skamarock, W.C., Klemp, J.B., Duda, M.G., Fowler, L.D., Park, S.-H., Ringler, T.D.: A multiscale nonhydrostatic atmospheric model using centroidal voronoi tesselations and c-grid staggering. Mon. Weather Rev. **140**(9), 3090–3105 (2012). https://doi.org/10.1175/MWR-D-11-00215.1

Takacs, A., Holland, L., Hueftle, R., Brown, B., Holmes, A.: Using in situ eddy dissipation rate (EDR) observations for turbulence forecast verification (2005)

Thompson, G., Field, P.R., Rasmussen, R.M., Hall, W.D.: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: implementation of a new snow parameterization. Monthly Weather Rev. **136**(12), 5095–5115 (2008). https://doi.org/10.1175/2008MWR2387.1

Trier, S.B., Sharman, R.D.: Mechanisms influencing cirrus banding and aviation turbulence near a convectively enhanced upper-level jet stream. Mon. Weather Rev. **144**(8), 3003–3027 (2016). https://doi.org/10.1175/MWR-D-16-0094.1

Trier, S.B., Sharman, R.D., Fovell, R.G., Frehlich, R.G.: Numerical simulation of radial cloud bands within the upper-level outflow of an observed mesoscale convective system. J. Atmos. Sci. **67**(9), 2990–2999 (2010). https://doi.org/10.1175/2010JAS3531.1

Tvaryanas, A.P.: Epidemiology of Turbulence-Related Injuries in Airline Cabin Crew, 1992–2001. Aviation Space Environ. Med. **74**(9), 970–976 (2003)

Vinnichenko, N.: Turbulence in the free atmosphere. Springer Science & Business Media (2013)

Vogel, G., Sampson, C.: Clear air turbulence indices derived from U.S. navy numerical model data: a Verification study, 34 (1996)

Xu, M., Zhao, C., Gu, J., Feng, J., Hagos, S., Leung, L., Luo, Y., Guo, J., Li, R., Fu, Y.: Convection-permitting hindcasting of diurnal variation of Mei-yu rainfall over East China with a global variable-resolution model. J. Geophys. Res.: Atmos. **126**, 1 (2021). https://doi.org/10.1029/2021JD034823

Yang, Q., Leung, L.R., Rauscher, S.A., Ringler, T.D., Taylor, M.A.: Atmospheric moisture budget and spatial resolution dependence of precipitation extremes in aquaplanet simulations. J. Clim. **27**(10), 3565–3581 (2014). https://doi.org/10.1175/JCLI-D-13-00468.1

Zhao, C., Leung, L.R., Park, S.-H., Hagos, S., Lu, J., Sakaguchi, K., Yoon, J., Harrop, B.E., Skamarock, W., Duda, M.G.: Exploring the impacts of physics and resolution on aqua-planet simulations from a nonhydrostatic global variable-resolution modeling framework. J. Advan. Model. Earth Syst. **8**(4), 1751–1768 (2016). https://doi.org/10.1002/2016MS000727

Authors and Affiliations

Haoming Chen[1] · Christy Yan-yu Leung[2] · Ping Cheung[2] · Haolin Liu[1] · Sai Tick Chan[2] · Xiaoming Shi[1,3]

✉ Xiaoming Shi
shixm@ust.hk

Haoming Chen
hchenda@connect.ust.hk

Christy Yan-yu Leung
yyleung@hko.gov.hk

Ping Cheung
pcheung@hko.gov.hk

Haolin Liu
hliudd@connect.ust.hk

Sai Tick Chan
stchan@hko.gov.hk

1 Division of Environment and Sustainability, The Hong Kong University of Science and Technology, Hong Kong 999077, Hong Kong, China

2 Hong Kong Observatory, Hong Kong 999077, Hong Kong, China

3 Center for Ocean Research in Hong Kong and Macau, The Hong Kong University of Science and Technology, Hong Kong, China