

Impact of model resolution and turbulence scheme on the representation of mountain waves and turbulence

Roshny Siri Jagan¹ and Juerg Schmidli¹

¹Goethe University, Frankfurt am Main, Germany

Correspondence: Juerg Schmidli (schmidli@iau.uni-frankfurt.de)

Abstract. Simulating mountain waves and associated turbulence in the upper troposphere and lower stratosphere (UTLS) remains a challenge in numerical weather prediction (NWP). We investigate how the representation of mountain-wave dynamics and turbulence in the ICOSahedral Nonhydrostatic (ICON) model depends on model resolution and turbulence parameterization. ICON simulations were performed in NWP mode (ICON-NWP) with varying horizontal (2 km, 1 km, 500 m) and vertical (400 m, 200 m, 100 m) resolutions, using the operational turbulent kinetic energy scheme (ICON-TKE) and the newly developed two-energy turbulence scheme (ICON-2TE). The simulations were evaluated against high-frequency in situ observations from the Deep Propagating Gravity Wave Experiment (DEEPWAVE) over New Zealand on 12 July 2014, as well as a nested large-eddy simulation (ICON-LES) at 130 m resolution. The results show reasonable agreement with observations: ICON-LES more closely captures wavelength and phase, while ICON-NWP better reproduces wave amplitude. Near-convergence of local wave and turbulence structures requires horizontal grid spacings of 1 km or finer and vertical spacings in the UTLS of 200 m or finer. A key finding is that both turbulence schemes yield similar wave structures, despite large differences in simulated turbulent kinetic energy. This discrepancy is attributed to the empirical horizontal-shear source term of the operational TKE-scheme configuration, which produces spurious TKE at km-scale resolution. In terms of bulk measures, the area-averaged gravity-wave momentum flux approaches convergence already at 1 km. These results provide guidance on the resolution and turbulence representation needed for reliable simulations of small-scale mountain waves and turbulence in the UTLS.

1 Introduction

The upper troposphere and lower stratosphere (UTLS), spanning roughly 6 to 25 km altitude (Riese et al., 2012), plays a central role in climate dynamics by regulating the distribution of greenhouse gases. As a transition zone between the relatively well-mixed troposphere and the stably stratified lower stratosphere, the UTLS facilitates tracer transport through a range of dynamical processes operating across multiple spatial and temporal scales (Scherllin-Pirscher et al., 2021; Riese et al., 2012). These processes include the large-scale Brewer–Dobson circulation (months to years), Rossby wave breaking and baroclinic cyclones (several days), deep convection on shorter timescales (Schäfler et al., 2023), and small-scale turbulence. Turbulence in the UTLS is particularly important for stratosphere–troposphere exchange of chemical constituents. Unlike the planetary boundary layer, where turbulence is sustained, turbulence aloft is intermittent and localized, producing short-lived but intense

25 mixing events, especially near the tropopause (Mirza et al., 2024). Acting on horizontal scales of a few hundred meters to several kilometers, UTLS turbulence can significantly influence large-scale circulation patterns (Rogel et al., 2023).

Upper-air turbulence in the UTLS can be broadly categorized by its origin into three types: clear-air turbulence, often associated with jet streams; mountain-wave turbulence (MWT), generated by flow over topography; and convectively induced turbulence, linked to thunderstorms (Kim et al., 2023; Ren and Lynch, 2024). Among these, MWT is of particular importance
30 because it poses substantial challenges for aviation safety, flight planning, and route optimization (Mahalov, 2016). Many long-haul flight routes cross major mountain ranges, where MWT occurs frequently and often with high intensity. Understanding mountain-wave dynamics and their associated turbulence is therefore not only relevant for weather and climate research, but also for operational forecasting and aviation risk management.

Mountain waves are a specific type of gravity wave that arises when stably stratified air is forced to flow over topography,
35 producing oscillations with gravity as the restoring force. As these waves propagate upward, their amplitude increases with decreasing air density. When the intrinsic phase speed of the wave approaches the background wind speed, the wave amplitude reaches a maximum and momentum is transferred to the mean flow. Beyond this point, the wave can become unstable and generate turbulence. This occurs either through catastrophic wave breaking, in which the wave is completely destroyed, or through wave saturation, where energy is dissipated into turbulence that limits further amplitude growth. In the latter case, the
40 wave only partially breaks but still produces localized turbulent mixing. These processes illustrate the close coupling of gravity waves and turbulence, which is central to understanding energy and momentum transfer in the UTLS (Arbor, 1999; Schneider et al., 2017; Dörnbrack et al., 1995; Sharman et al., 2012; Fritts and Alexander, 2003).

Mountain waves play a major role in atmospheric energetics by transporting energy and momentum away from their source regions. They are typically generated in the troposphere and can propagate both vertically into the stratosphere and horizontally
45 over long distances. The structure and propagation of a mountain wave depend strongly on the size and shape of the underlying topography as well as on the vertical profiles of wind, temperature, and moisture in the flow (Durran, 2015). Orographic gravity waves often dominate the global gravity-wave spectrum, with momentum fluxes several times larger than those of non-orographic waves and concentrated over regions of complex terrain (Holt et al., 2017).

Mountain waves have been investigated extensively through ground-based instruments, in situ aircraft measurements, satellite observations, and dedicated field campaigns (Heale et al., 2022; Grubišić and Lewis, 2004; Grubišić et al., 2008; Bougeault et al., 1993; Kuettner and O'Neill, 1981; Bougeault et al., 2001; Wratt et al., 1996; Fritts et al., 2016; Schäfler et al., 2018; Jackson et al., 2018; Rapp et al., 2021). These studies have identified several “hotspots” of gravity-wave activity, such as regions downstream of major mountain ranges (Heale et al., 2022). Despite this rich observational record, forecasting mountain waves and associated turbulence in numerical weather prediction (NWP) and climate models remains highly challenging.
55 The difficulty arises from the multiscale interactions between waves and turbulence, and from the fact that a large part of the gravity-wave spectrum lies below the typical grid resolution of global models and thus must be parameterized (Plougonven et al., 2020; Achatz et al., 2024; Voelker, 2024). Recent advances in computational power have enabled mesoscale nested simulations with horizontal resolutions of about 1 km or finer, offering new opportunities to explicitly resolve small-scale processes

while still accounting for the large-scale environment provided by coarser parent domains (Muños-Esparza et al., 2020; Rogel
60 et al., 2023).

Large-eddy simulation (LES) and direct numerical simulation (DNS) provide powerful tools for investigating turbulence associated with mountain waves at very high resolution. These approaches can resolve small-scale structures that remain inaccessible to operational NWP models. For example, Bramberger et al. (2020) used a two-dimensional model with 200 m horizontal and 100 m vertical grid spacing to simulate a mountain-wave breaking event over Iceland. Their results demonstrated
65 that the observed turbulence could be attributed directly to breaking mountain waves. Such high-resolution studies highlight the potential of LES and DNS to improve physical understanding of mountain-wave turbulence, while also providing valuable benchmarks for evaluating parameterizations in coarser-resolution models.

Several recent modeling studies have addressed the resolution required to capture mountain-wave dynamics and associated processes with fidelity. Lund et al. (2020) and Fritts et al. (2021) investigated the nonlinear evolution of mountain waves over
70 the southern Andes using simulations at 500 m resolution, revealing the complex dynamics that emerge with increasing wave forcing. Extending this work, Fritts et al. (2022) examined how model resolution affects the representation of mountain waves propagating into the thermosphere. Their results showed that coarser resolutions quickly degrade the ability to simulate key wave dynamics, leading to systematic biases in global models that influence both weather and climate predictions. Through simulations of wintertime flow over the southern Andes at horizontal resolutions ranging from 0.5 to 8 km, they demonstrated
75 that resolutions of about 2 km or finer are necessary to adequately capture mountain-wave responses, particularly in the mesosphere.

Most turbulence parameterizations used in numerical models were originally developed for the atmospheric boundary layer (ABL), but they are not necessarily suited to conditions in the UTLS. The ABL is characterized by a solid lower boundary and strongly influenced by diurnal cycles of radiative heating and cooling, leading to turbulence generation mechanisms that
80 differ fundamentally from those aloft. In contrast, turbulence in the UTLS occurs in a strongly stratified environment, is often patchy and anisotropic, and appears as sudden, localized bursts of velocity and temperature fluctuations that deviate from classical Kolmogorov scaling (Rodríguez Imazio et al., 2023; Kleerorin et al., 2019; Mahalov, 2016; Paoli et al., 2014). The challenges of applying ABL-based schemes at km-scale resolutions in the UTLS have been demonstrated by Muños-Esparza et al. (2020), who showed that turbulence parameterization significantly influences gravity-wave activity and vertical mixing in
85 high-resolution weather simulations. Using a case of widespread turbulence over the U.S. Great Plains, they found that a 1 km model tends to overestimate vertical velocities, wave fluxes, and kinetic energy compared to a 250 m LES. These findings underline the need for turbulence parameterizations adapted specifically to UTLS conditions in order to capture wave–turbulence interactions more realistically.

In this study, we evaluate the performance of two turbulence schemes implemented in the ICOSahedral Nonhydrostatic
90 (ICON) limited-area model for simulating mountain waves and associated turbulence in the UTLS. Specifically, we compare the operational turbulent kinetic energy scheme (ICON-TKE) with the newly developed two-energy turbulence scheme (ICON-2TE). Sensitivity experiments are conducted with varying horizontal and vertical resolutions to determine the model configuration required to capture key wave and turbulence characteristics. The simulations are qualitatively evaluated against

high-frequency in situ observations from the Deep Propagating Gravity Wave Experiment (DEEPWAVE), a field campaign
95 conducted over New Zealand in 2014, and against a nested large-eddy simulation (ICON-LES) at 130 m resolution. The ob-
servational comparison is used primarily to verify that both schemes produce realistic flow structures rather than to provide an
exhaustive validation. The remainder of the paper is organized as follows: Sect. 2 describes the ICON model setup, including
the two turbulence schemes, the case study, and the analysis methodology. Sect. 3 presents the evaluation of a reference simu-
lation and the results of the sensitivity experiments, including a discussion of their implications for simulating mountain waves
100 and turbulence in the UTLS. Finally, Sect. 4 summarizes the main findings and provides recommendations for future work.

2 Experimental setup

2.1 Case study and observational data

The case analyzed in this study is a mountain-wave event over New Zealand on 12 July 2014 during the DEEPWAVE campaign,
with the topography and study domain shown in Fig. 1. This case is useful because it combines a clearly observed mountain
105 wave with relatively rare high-frequency in situ measurements near the tropopause. Earlier analyses inferred possible signatures
of turbulence impacts on tracer distributions in the UTLS region (Lachnitt et al., 2023). Although direct observational evidence
of wave breaking was not available, the event provides targeted in situ observations that make it suitable for evaluating how
turbulence parameterizations and model resolution affect the simulation of mountain waves in the UTLS.

The synoptic situation during the event was characterized by a trough to the west of New Zealand and a weak low-pressure
110 system to the south of the islands. This configuration produced moderate northwesterly flow across the South Island, i.e., about
10-15 m/s at 5 km, oriented perpendicular to the Southern Alps. At tropopause level (around 10 km altitude), winds were also
from the northwest, reaching moderate speeds of about 20 m/s along the DEEPWAVE flight path. Overall, weak directional
shear created favorable conditions for the generation of vertically propagating mountain waves over the South Island.

The DEEPWAVE research flight FF09 was carried out with the DLR Falcon aircraft on 12 July 2014 between 17:15 and
115 20:15 UTC. The aircraft followed a clockwise rectangular flight track across the South Island to capture the structure of the
mountain-wave event (see Fig. 1). To sample vertical variability, the flight pattern consisted of two stacked legs separated by
approximately 75 min, with the lower level at 360 hPa (7.9 km) and the upper level at 230 hPa (10.9 km). In the southern
segment of the track, the lower leg was flown between 17:30 and 17:50 UTC and the corresponding upper leg between 18:47
and 19:07 UTC. In the northern segment, the lower leg took place from 18:12 to 18:32 UTC and the upper leg from 19:27 to
120 19:47 UTC. These stacked cross sections provide in situ measurements of mountain waves and turbulence signatures near the
tropopause and serve as the observational reference for the model evaluation. Importantly, the simulated primary wave during
this period is quasi-stationary, with only weak temporal variations in wave amplitude and vertical structure. The observational
comparison therefore focuses on the time of the flight transect, which is representative of the modeled flow regime.

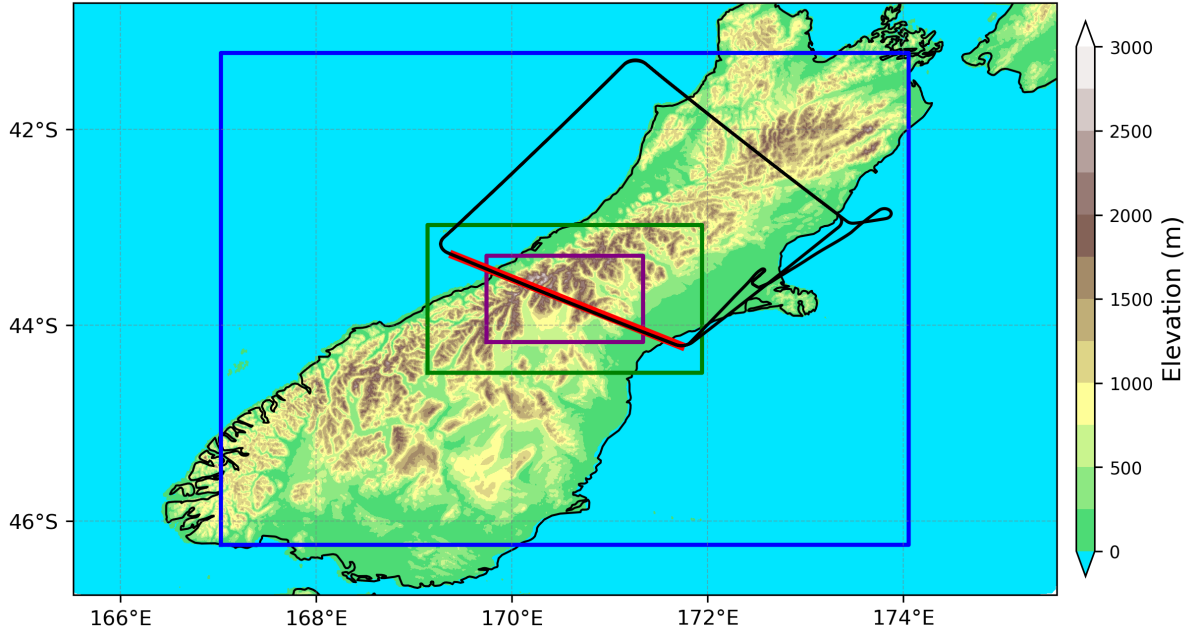


Figure 1. Orography (color shading) and simulation domains for the nested ICON setup: 1 km (full region), 500 m (blue box), and 130 m for the southern flight segment (green box). The red line marks the cross-section along the flight path used for the main analysis, and the purple box indicates the averaging region for calculating bulk properties of the GW field (Sect. 3.4). The black solid line shows the full flight path of research flight FF09 of the Falcon aircraft on 12 July 2014 (17:15–20:15 UTC) during the DEEPWAVE campaign.

2.2 Model and turbulence schemes

125 The numerical experiments were performed with the ICOSahedral Nonhydrostatic (ICON) model version *icon-2024.01-dwd-2.0*, operated in both NWP and LES modes. ICON is a fully compressible atmospheric model jointly developed by the German Weather Service and the Max Planck Institute for Meteorology. It employs an icosahedral-triangular grid constructed using geodesic Delaunay triangulation, a height-based vertical coordinate system, and C-grid staggering (Zängl et al., 2015). The physical parameterizations include the Tiled TERRA land surface scheme (Schulz et al., 2016), a one-moment cloud micro-
130 physics scheme with graupel (Seifert, 2008), the low-level flow blocking component of the subgrid-scale orographic drag scheme (Lott and Miller, 1997), and the ecRad radiation scheme (Hogan and Bozzo, 2018).

A central aim of this study is to assess how the simulation of mountain waves and turbulence in ICON depends on both the choice of turbulence parameterization and the model resolution. Two turbulence schemes are considered: the operational turbulent kinetic energy scheme (TKE scheme; Raschendorfer, 2001) and the more recently developed two-energy turbu-
135 lence scheme (2TE scheme; Bařták Ďurán et al., 2022). The TKE scheme predicts a single turbulent kinetic energy (TKE) and accounts for turbulence generation beyond vertical wind shear through additional empirical source terms for large-scale horizontal shear (HS term) and breaking subgrid-scale mountain waves (SSO term) (Goecke and Machulskaya, 2021). Its de-

velopment and tuning prioritize operational NWP applications, including aviation turbulence products such as eddy dissipation rate (EDR). In contrast, the 2TE scheme predicts two prognostic turbulence energies, enabling a more flexible representation of anisotropic and stably stratified turbulence. The scheme avoids prescribed minimum diffusion coefficients, dynamically adjusts the turbulence length scale, and employs an assumed probability density function (APDF) method to represent buoyancy production. Through this formulation, the 2TE scheme provides a unified framework for representing turbulence, shallow convection, and subgrid-scale cloud processes (Bašták Ďurán et al., 2018; Singh et al., 2025). Further details on the two turbulence schemes are provided in Appendix A.

While the TKE scheme has long been used operationally and the 2TE scheme is currently under development, their performance in simulating mountain waves and associated turbulence in the UTLS has not been systematically evaluated. A further aim of this study is therefore to provide a first evaluation of the 2TE scheme in this regime, to inform its ongoing development. In addition, the sensitivity of these simulations to horizontal and vertical resolution is not well constrained, despite growing use of km-scale NWP. The case is also demanding in its own right: the Southern Alps of New Zealand present three-dimensional, multi-scale terrain, with several ridges of different heights and orientations, in contrast to the isolated or quasi-two-dimensional ridges of many earlier studies, so that waves from several sources interact. This study therefore investigates how the two turbulence schemes, in combination with different resolutions, represent mountain-wave dynamics, turbulence characteristics, and momentum fluxes during the DEEPWAVE case. The different ICON configurations and sensitivity experiments are introduced in the following subsection.

2.3 Model configurations and simulations

Based on the two turbulence schemes described above, we carried out a set of numerical experiments with varying horizontal and vertical resolutions, complemented by a nested LES and targeted sensitivity runs. The experiments are summarized in Table 1. Limited-area ICON simulations in NWP mode were conducted at horizontal grid spacings of approximately 2km (I20-2TE), 1 km (I10-2TE, I10-TKE), and 0.5 km (I05-2TE). All configurations employed 137 vertical levels with spacing increasing from about 20 m near the surface to finer than 200 m up to 14 km, a model top at 30 km, and Rayleigh damping above 20 km. The simulations were initialized from IFS analyses at 00:00 UTC on 12 July 2014 and integrated for 24h, with outputs every 30 min and every 5 min between 17:00–20:00 UTC. Among these, the I10-2TE simulation was selected as the reference case for the analysis.

A online nested LES configuration at 130 m (I01-LES) targeted the southern flight segment. The LES was realized with four nested domains down to 130 m, with the outermost nest run using the 2TE scheme and the inner nests employing a Smagorinsky closure.

To isolate parameterized sources of turbulence production in the TKE scheme, we additionally performed targeted sensitivity runs at 1 km resolution: one with only subgrid-scale orography activated (I10-TKE-SSO), one with only horizontal-shear production activated (I10-TKE-HS), and one with both terms deactivated (I10-TKE-none).

Table 1. Overview of ICON model experiments used in this study. The table lists the turbulence scheme, horizontal resolution, and grid setup for each configuration, including NWP runs, nested LES, and targeted ICON-TKE sensitivity experiments. SSO refers to subgrid-scale orography, HS refers to horizontal shear, and GP denotes grid points.

Experiment	Resolution	Turbulence setup	# of GPs	Domain Size (km x km)
I20-2TE	2076 m	2TE	127466	677 x 812
I10-2TE	1038 m	2TE	502719	677 x 812
I05-2TE	519 m	2TE	1167450	558 x 564
I01-LES	130 m	Smagorinsky ^a	2227352	167 x 225
I10-TKE	1038 m	TKE, with SSO & HS	502719	677 x 812
I10-TKE-SSO	1038 m	TKE, with SSO	502719	677 x 812
I10-TKE-HS	1038 m	TKE, with HS	502719	677 x 812
I10-TKE-none	1038 m	TKE, no extra terms	502719	677 x 812

^a The LES are run as an online nest consisting of four domains: 1 km, 519 m, 260 m and 130 m.

170 2.4 Bulk characteristics of the GW field

In addition to local characteristics, we also evaluate bulk characteristics of the simulated gravity wave field over the Southern Alps, namely the area-averaged vertical fluxes of momentum induced by the waves. For this purpose, the model output on terrain-following levels was interpolated to geometric height levels, and averages were taken over the analysis domain shown in Fig. 1. The vertical fluxes of horizontal momentum are defined as

$$175 \quad \overline{\rho u' w'} = \overline{\rho(u - \bar{u})(w - \bar{w})}, \quad \overline{\rho v' w'} = \overline{\rho(v - \bar{v})(w - \bar{w})}, \quad (1)$$

where overbars denote horizontal means at a given altitude and primes the corresponding perturbations. The total flux magnitude is then given by

$$\tau = \sqrt{(\overline{\rho u' w'})^2 + (\overline{\rho v' w'})^2}. \quad (2)$$

The variance of vertical velocity, $\overline{w'^2}$, follows as a special case of these definitions. Along with the subgrid-scale TKE, these diagnostics characterize the bulk properties of the GW field and parameterized turbulence across different resolutions and turbulence schemes. In particular, area-averaged momentum fluxes provide a measure of convergence with resolution and are an important target quantity for parameterizations in coarse-resolution models.

3 Results and Discussion

We first examine the I10-2TE simulation to establish the baseline characteristics of the wave event. This case serves as the reference for assessing the impact of resolution and turbulence parameterizations in the subsequent subsections. We chose

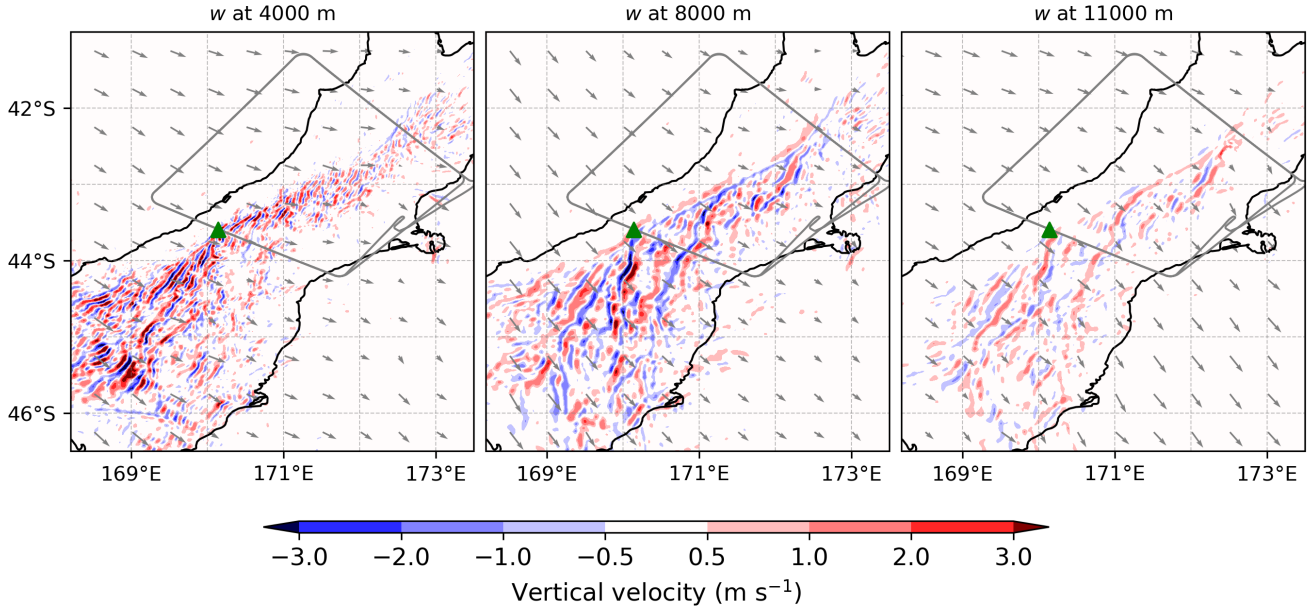


Figure 2. Horizontal cross-sections of vertical wind at three heights, with the upper two corresponding to the lower and upper flight legs, from the I10-2TE simulation at 18:00 UTC on 12 July 2014. Horizontal wind vectors are overlaid. The full flight track is shown by the grey line, and the green marker denotes the location of Mount Cook, New Zealand.

I10-2TE as the reference because, as will be shown, the standard TKE configuration produces excessive and likely unrealistic TKE, making it less consistent as a baseline.

3.1 Reference simulation of the wave event

The reference simulation I10-2TE reproduces a distinct mountain-wave pattern over the South Island at flight time (18:00 UTC). Horizontal cross-sections of vertical velocity show alternating bands of ascent and descent downstream of the main ridge of the Southern Alps, extending from the mid-troposphere into the lower stratosphere (Fig. 2). The wave signal is most pronounced near the southern leg and south of the DEEPWAVE flight track. It is stronger in the troposphere (4 and 8 km) than in the lower stratosphere (11 km), consistent with vertically propagating gravity waves. Note the dominant waves also increase in horizontal scale between 4 and 8 km.

A vertical cross-section along the southern segment of the flight path at 18:00 UTC highlights the primary wave response above the topography (Fig. 3c). The transect passes near Mount Cook, New Zealand's highest mountain, at about $x = 68$ km. Further southeast, at about $x = 125$ km, it crosses the Two Thumb Range, the last major ridge in this section. A strong mountain wave, with vertical velocities exceeding 3 m/s, is triggered by the Mount Cook region, while a weaker wave develops above the Two Thumb Range. Both waves propagate into the lower stratosphere and are associated with leeside descent, low-level steepening of the isentropes, and the production of turbulence (subgrid-scale TKE) as seen in Fig. 3b. Significant turbulence is

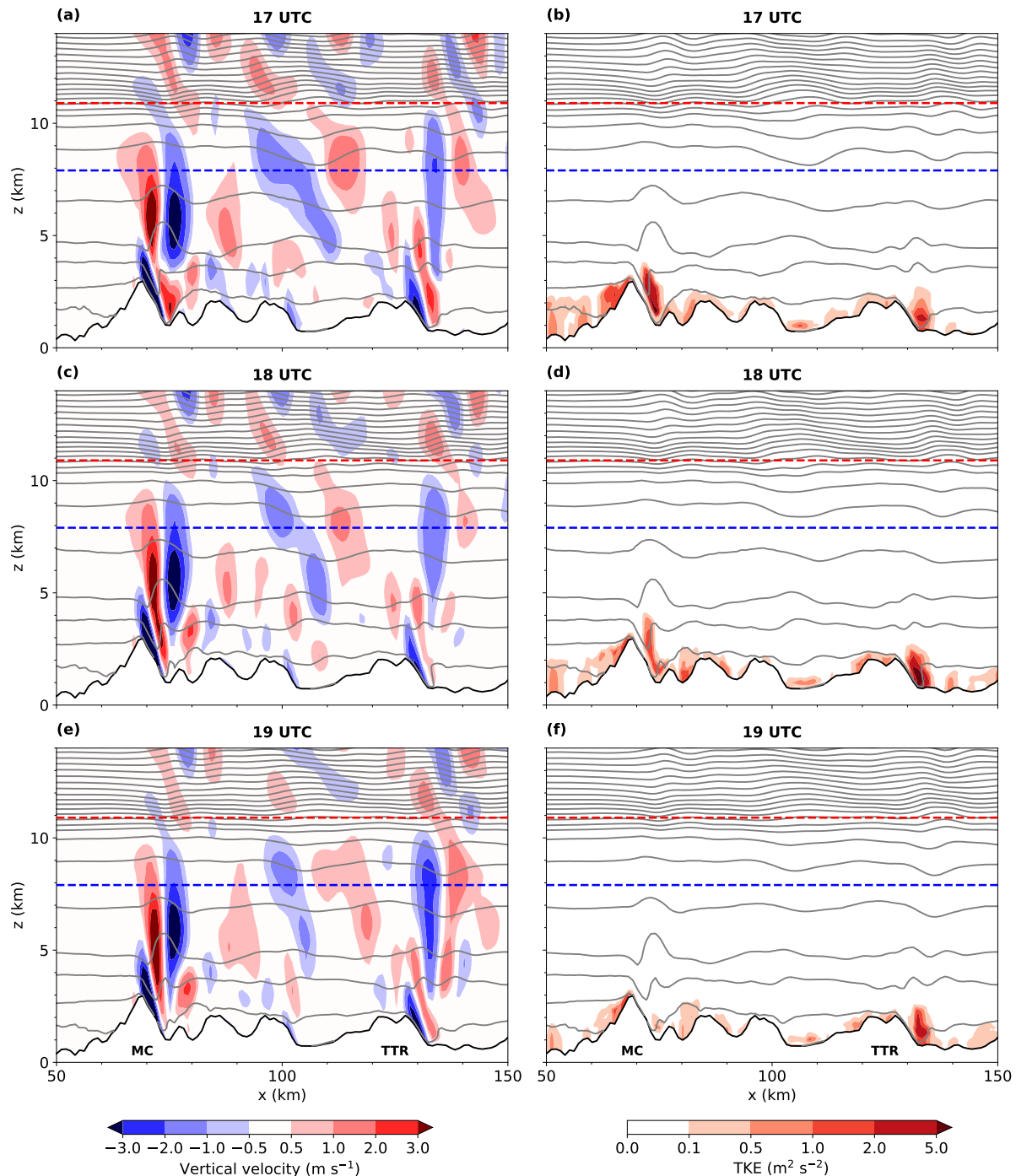


Figure 3. Vertical cross-sections of vertical wind (left column) and parameterized TKE (right column) over the southern segment of the flight track, simulated with I10-2TE at 17:00 UTC (a, b), 18:00 UTC (c, d), and 19:00 UTC (e, f). The similarity between the three times illustrates the quasi-stationary nature of the simulated wave field during the DEEPWAVE flight period. The blue dashed line denotes the lower flight leg, and the red dashed line the upper flight leg. Grey contours show potential temperature at 5 K spacing. The locations of Mount Cook (MC) and the Two Thumb Range (TTR) are indicated.

confined to the ABL and to regions of steep isentropes associated with lee-side flow separation. Note that non-orographic BL turbulence is weak, as expected for the early morning hours (06:00 local time).

To assess the stationarity of the simulated wave field, the same cross-section is shown at 17:00 and 19:00 UTC (Fig. 3a,e). At upper levels, the primary wave over Mount Cook remains largely unchanged, whereas noticeable differences appear further downstream above the Two Thumb Range. Near the surface, the flow over Mount Cook varies between the three times, with associated variations in the parameterized TKE. In summary, the large-scale wave pattern over the Mount Cook region is quasi-stationary, whereas smaller-scale features and the low-level TKE show less stationarity. This temporal variability at lower levels has little effect on the wave maxima near 6 km or the flight-leg altitudes.

Comparison of I10-2TE with observations (Fig. 4) shows that the simulated wave amplitude, in terms of vertical velocity perturbations, agrees reasonably well with the observations on the lower flight leg but it is underestimated on the upper leg. A spatial phase mismatch between simulated and observed waves is also apparent, which is to be expected. While the simulated wave field is almost stationary, particularly for the strong wave triggered above Mount Cook, the real atmosphere is likely much less stationary due to multi-scale interactions among a spectrum of gravity waves and turbulent motions. Additional variability is also expected from temporal changes in the upstream profile. Note that the observed wave phases differ between the lower and upper flight legs. These differences are unlikely to reflect a vertical discontinuity; instead, they most likely result from the roughly one-hour time difference between the two legs and the non-stationary nature of the real gravity-wave field.

The horizontal wind along both legs is compared with the observations in Fig. 5 for the reference simulation I10-2TE at 17, 18, and 19 UTC. The simulated wind direction on the lower leg changes by about 10° per hour, and the simulation agrees most closely with the observations at the output times nearest each crossing (18 UTC for the lower leg, 19 UTC for the upper leg). The upper-level wind-speed bias (up to about 10 m/s) is common to all resolutions and to both turbulence schemes (not shown, as the winds are largely insensitive to scheme and resolution), so the upper-leg differences reflect the represented flow rather than the turbulence parameterization, consistent with the nearly identical resolved variances of the two schemes (Sect. 3.4). Part of the observed wind fluctuation is the wave itself (e.g., wind-speed perturbations of about 5 m/s and wind-direction perturbations of $10\text{--}15^\circ$ on the upper leg).

In summary, the reference simulation captures a strong vertically propagating mountain wave but shows no evidence of significant wave breaking or turbulence at flight altitudes associated with the dominant wave, consistent with the observations. It provides the baseline for assessing the sensitivity to model resolution and turbulence parameterization presented in the following subsections.

3.2 Sensitivity to horizontal and vertical resolution

Figure 6 presents the wave field in terms of vertical velocity and potential temperature. The simulated wave field is strongly affected by horizontal resolution, with substantially more wave activity at finer grid spacing. At 2 km resolution, the horizontal wavelength increases and the gravity wave exhibits a more hydrostatic character, as shorter scales are no longer resolved. This is particularly evident for the wave triggered by the Two Thumb Range. At finer resolution (0.5 km), the waves become more non-hydrostatic, with vertical velocity perturbations more vertically aligned. Importantly, the number of resolved waves

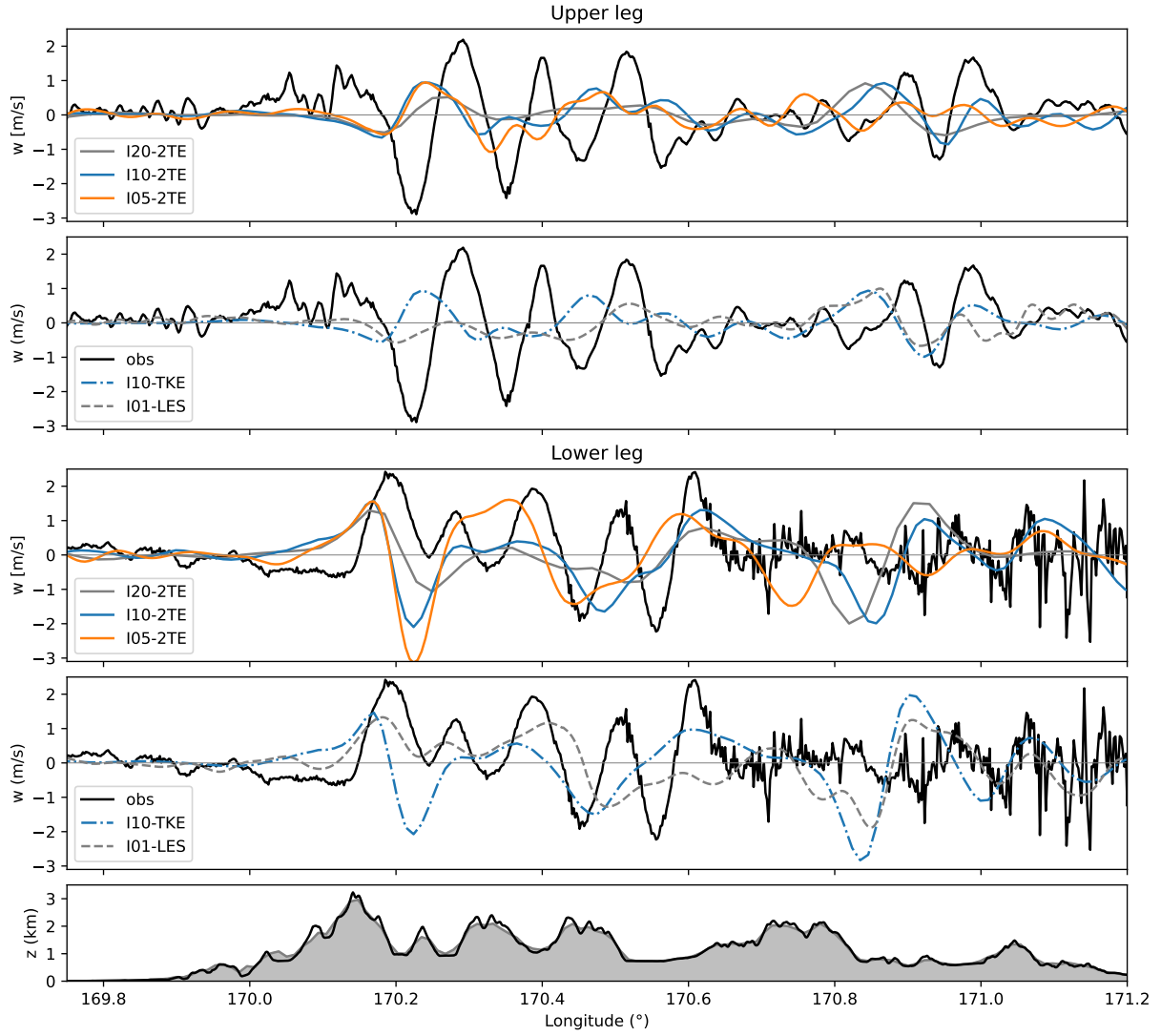


Figure 4. Comparison of vertical velocity from ICON simulations at different horizontal resolutions at 18:00 UTC with in situ observations along the southern segment of the flight track. Results are shown separately for the lower and upper flight legs, together with the underlying terrain along the cross-section, shown for I10-2TE (shading) and I01-LES (black line). Note that 0.2° on the horizontal axis corresponds to 16 km.

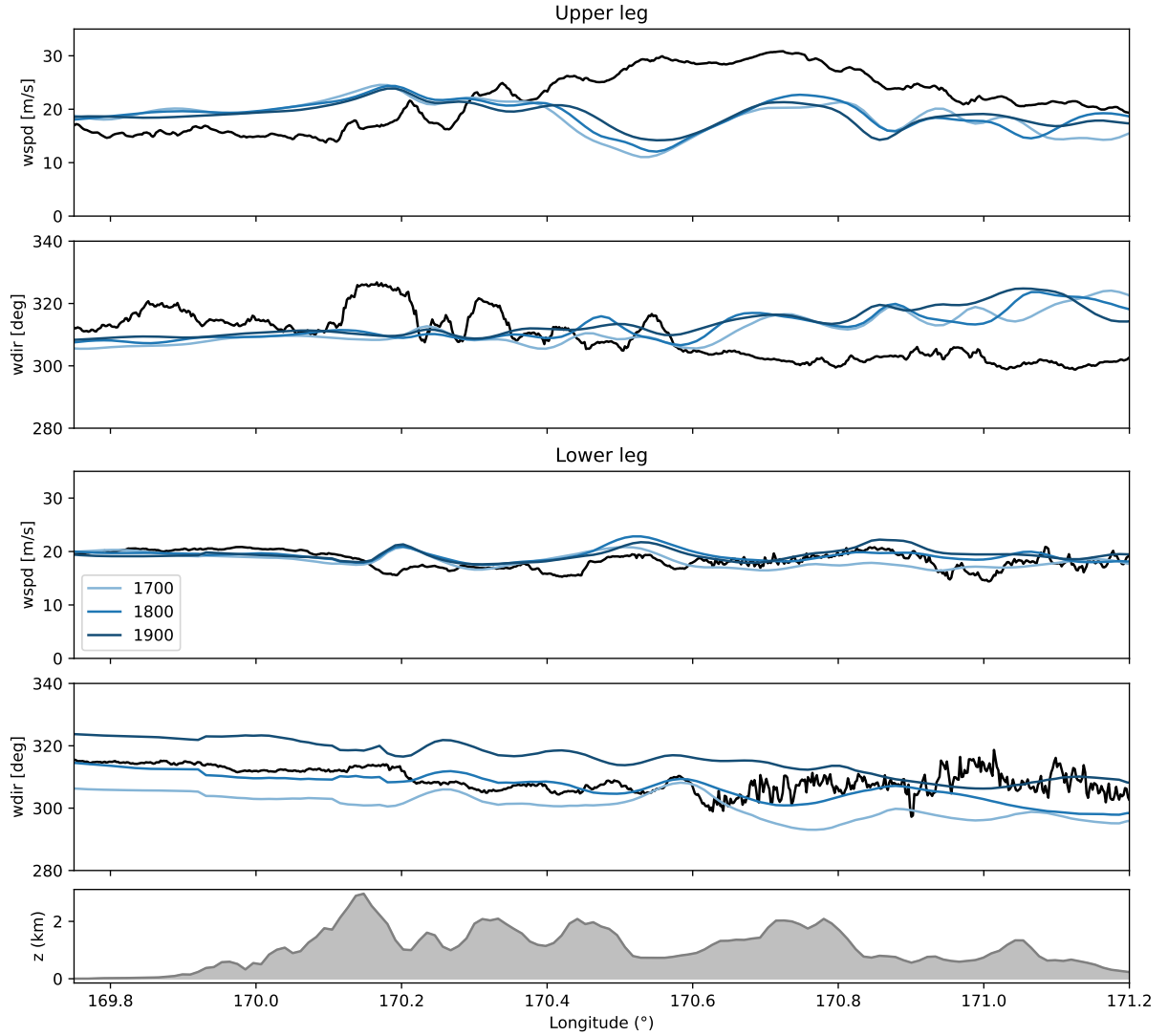


Figure 5. Comparison of horizontal wind speed and direction from the reference simulation I10-2TE at 17:00, 18:00, and 19:00 UTC with in situ observations along the southern segment of the flight track (lower leg: 17:30-17:50 UTC, upper leg: 18:47-19:07 UTC). Results are shown separately for the lower and upper flight legs, together with the underlying terrain along the cross-section.

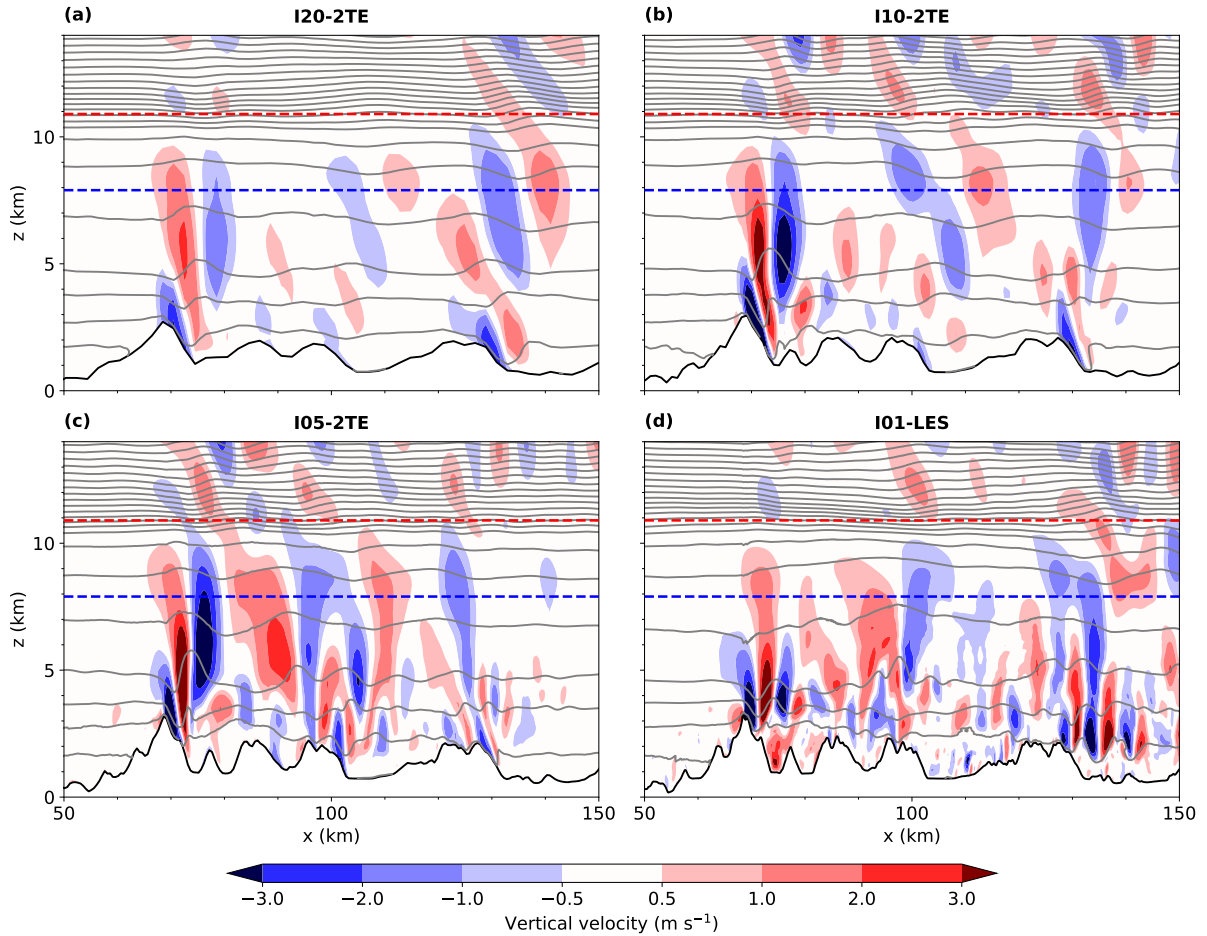


Figure 6. Comparison of vertical velocity along the southern flight-path transect at 18:00 UTC for 2TE simulations and LES at different horizontal resolutions: (a) 2 km, (b) 1 km, (c) 500 m, and (d) 130 m.

increases as the model begins to capture shorter wavelengths. This effect is even more pronounced in the LES at 130 m. The LES further shows that the shortest wavelengths are trapped in the layer of enhanced stability below about 5 km, with a clear example of trapped lee waves in the lee of the Two Thumb Range. Another notable difference between the NWP simulations and the LES is the representation of the flow structure immediately downstream of the mountains (Mount Cook, an intermediate ridge at ~ 100 km, and the Two Thumb Range). In the NWP runs, the downslope flow penetrates to lower levels than in the LES, which has a significant effect on mountain-wave triggering. These differences are likely due not only to resolution, but also to the different representation of the atmospheric boundary layer and its coupling with the mountain wave (e.g. Jiang et al., 2008). For completeness, horizontal cross-sections of the flow at different resolutions are shown in Appendix B, confirming the sensitivity of dominant gravity-wave scales to model resolution.

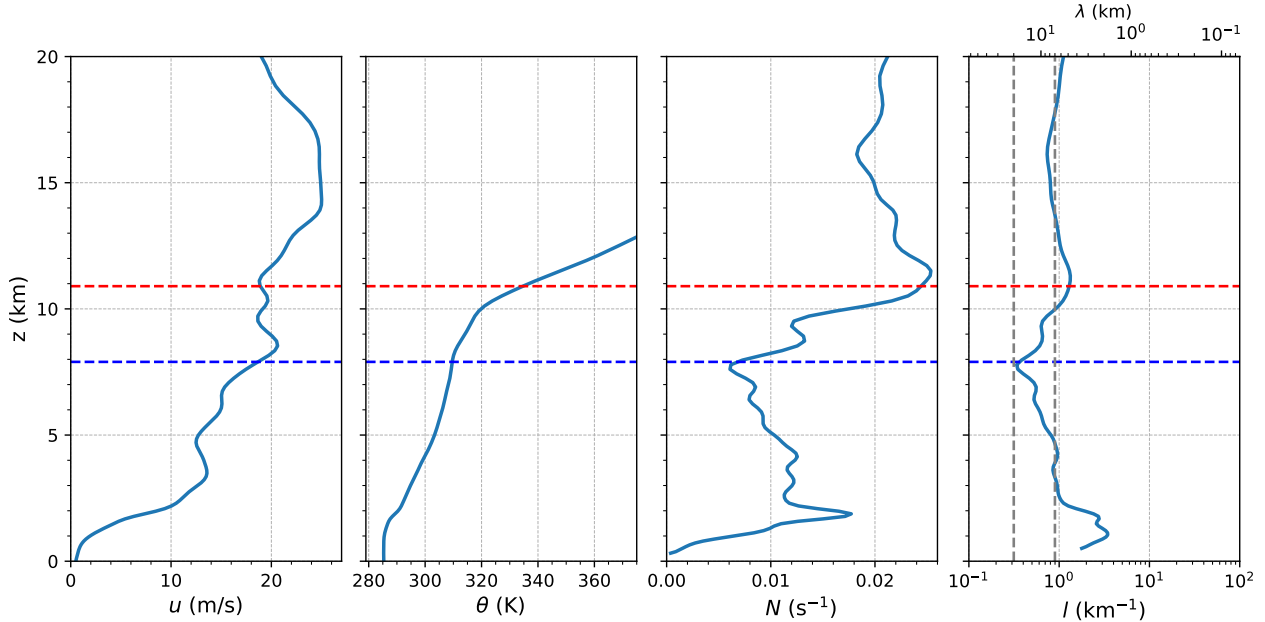


Figure 7. Upstream profiles of wind speed parallel to the southern flight leg (u), potential temperature (θ), Brunt–Väisälä frequency (N), and Scorer parameter (l) from the I10-2TE simulation. Profiles are taken at 10 km along the red transect line shown in Fig. 1 at 18:00 UTC. Wind speed and potential temperature are smoothed vertically using a Gaussian filter with a standard deviation of 1.5 vertical grid points. The grey vertical dashed lines indicate a Scorer parameter values corresponding to wavelengths of 7 and 20 km, respectively.

To gain further insight into the wave environment, upstream profiles of wind speed, potential temperature, Brunt–Väisälä frequency, and the Scorer parameter are shown in Fig. 7. Wind speed increases approximately linearly from about 10 m/s at ridge height to nearly 20 m/s just below the tropopause at 9 km. In terms of stability, the lower troposphere (below 5 km) is markedly more stable than the upper troposphere above. These two factors combine to produce a decrease of the Scorer parameter with height through much of the troposphere, followed by a recovery to slightly larger values in the lowermost stratosphere. The Scorer parameter profile implies a cut-off wavelength of about 7 km just above ridge level, increasing to roughly 20 km near the tropopause. This vertical structure explains the generation of trapped lee waves in the more stable layer below 5 km and is likely responsible for the reduced simulated wave amplitude at the upper flight leg. In the observations, by contrast, the primary wave keeps a nearly constant horizontal wavelength of about 8 km and a strong signal on the upper leg, indicating that it propagates more freely than the model’s Scorer profile would suggest. This difference may in part reflect a bias in the simulated upstream wind and stability profile, which sets the Scorer parameter. Comparison of these simulations with in situ observations along the southern leg (Fig. 4) shows that the wave amplitude is reproduced reasonably well on the lower flight leg but underestimated on the upper leg. Among the NWP simulations, the 0.5 km run (I05-2TE) agrees best with the observed amplitudes at the lower level. The LES better reproduces the wavelength and phase than the NWP simulations, but underestimates the amplitude, particularly on the upper leg. These discrepancies likely reflect, among other factors, the

non-stationarity of the real wave field compared to the quasi-stationary model simulations. For example, in the 0.5 km run, the simulated wave above Mount Cook at the altitude of the lower flight leg compares reasonably well with the observed wave at the altitude of the upper leg. The underestimation of the simulated amplitude aloft is most likely a matter of wave propagation (as discussed above), not of turbulent dissipation. Given the three-dimensional, multi-scale orography, exact agreement is not expected, in particular downstream of the primary wave, where waves from different ridges interfere.

The sensitivity to vertical resolution was assessed with additional 2TE runs in which the maximum grid spacing below 14 km was set to 400 m, 200 m, and 100 m, respectively. The coarsest grid (400 m) shows clear differences from the reference run, particularly in the stratosphere, whereas the 200 m and 100 m results are very similar, indicating near-convergence at 200 m (not shown). This finding motivated the choice of a maximum vertical spacing of 200 m as the default in the NWP setup used here.

3.3 Sensitivity to turbulence parameterization

3.3.1 Comparison of the two turbulence schemes

At 1 km resolution, the two turbulence schemes produce rather similar wave patterns, particularly for the wave field above Mount Cook (Figs. 8 a,c), as also seen in the horizontal cross-sections (Fig. B2). In contrast, the simulated TKE differs markedly between the schemes (Figs. 8 b,d). The TKE scheme produces substantially higher TKE values throughout the wave field, whereas the 2TE scheme confines turbulence mainly to the ABL and the strongest low-level mountain waves. The elevated TKE values above the ABL and outside regions of strong gravity-wave activity appear questionable and warrant further investigation. A key difference between the two schemes is the inclusion of additional empirical source terms in the TKE scheme (see Appendix A). This motivated further experiments to isolate the impact of these terms.

3.3.2 Effect of optional TKE source terms in the TKE scheme

Sensitivity experiments with individual source terms switched off are shown in Fig. 9. These targeted experiments allow us to disentangle the effects of the horizontal shear and subgrid-scale orography source terms in the TKE equation (see Appendix A). While there are striking differences in the TKE fields, the resolved wave structures are almost identical across all four simulations. In I10-TKE and I10-TKE-HS, which both include the horizontal shear term, substantial TKE appears not only in the ABL and low-level flow-separation regions but also in patches throughout the troposphere, with weaker pockets even extending into the lower stratosphere. In contrast, I10-TKE-none and I10-TKE-SSO, which exclude the horizontal shear term, produce TKE distributions more similar to those obtained with the 2TE scheme, with a modest increase in TKE at low levels for I10-TKE-SSO. The occurrence of large TKE values in regions without strong wave activity suggests a breakdown of the empirical horizontal-shear parameterization at km-scale resolution. The parameterization is a two-dimensional Smagorinsky-type closure (Appendix A) that relates horizontal-shear production to the deformation of the resolved flow. Its physical interpretation assumes that deformation near the model cutoff is associated, on average, with a net downscale transfer of kinetic energy to unresolved motions. This interpretation is most plausible when near-grid variability is dominated by

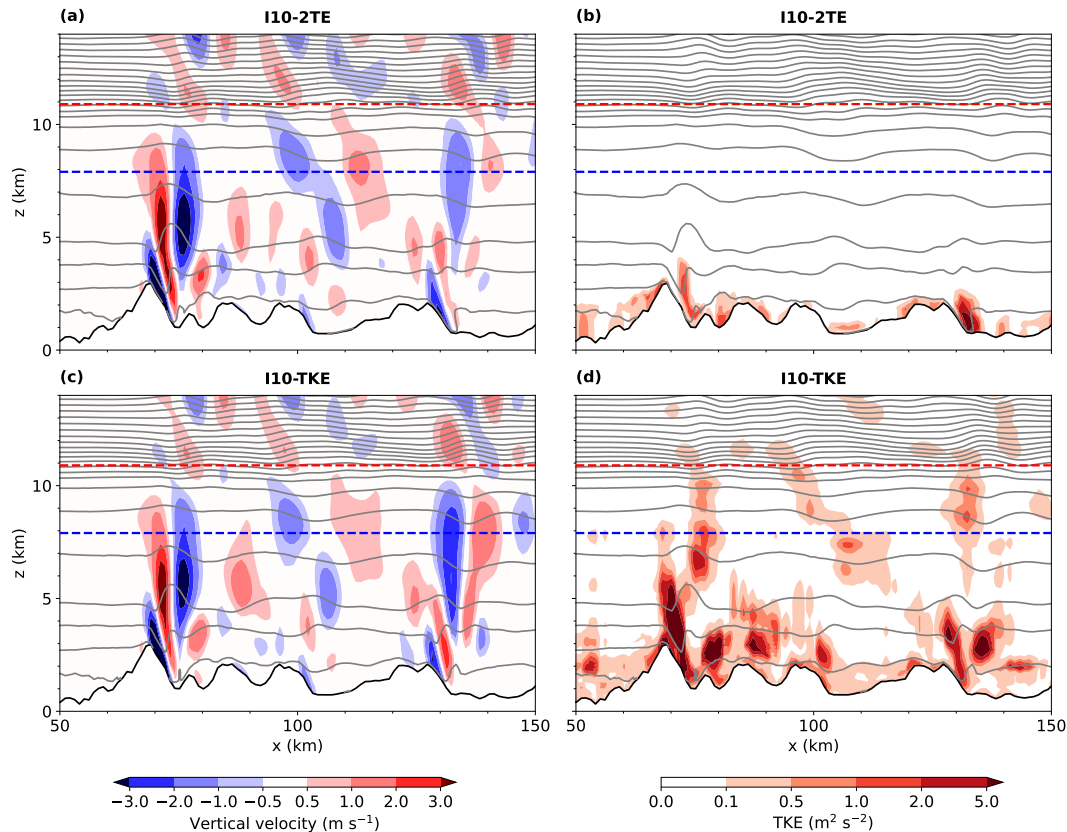


Figure 8. Comparison of vertical velocity and TKE along the southern flight-path transect at 18:00 UTC for ICON simulations at 1 km resolution using (a, b) the 2TE scheme and (c, d) the TKE scheme.

stratified mesoscale turbulence, as may occur in strongly deformational frontal and jet regions. In the present simulations, however, much of the grid-scale deformation is associated with resolved, coherent gravity waves. The closure cannot distinguish this largely reversible wave deformation from strain associated with an irreversible turbulent cascade and therefore converts wave-related deformation into parameterized TKE production. This explains why the excess TKE is spatially correlated with the resolved wave field rather than with regions of active turbulence. This finding is consistent with Bramberger et al. (2020), who showed that the Graphical Turbulence Guidance tool, which uses a similar empirical formulation, overpredicted turbulence magnitude over a large area during a breaking mountain-wave event over Iceland. This does not imply that low TKE is realistic everywhere. Two distinct deficiencies coexist at these resolutions: the horizontal shear term over-produces TKE where the observations show no turbulence (e.g., smooth spatial variation of w over the upstream half of the flight segment in Fig. 4), whereas genuine small-scale turbulence (e.g., downstream of the Two Thumb Range, and at the sub-km scales identified by Lachnitt et al., 2023) is under-resolved and hence missed by all configurations, including the LES.

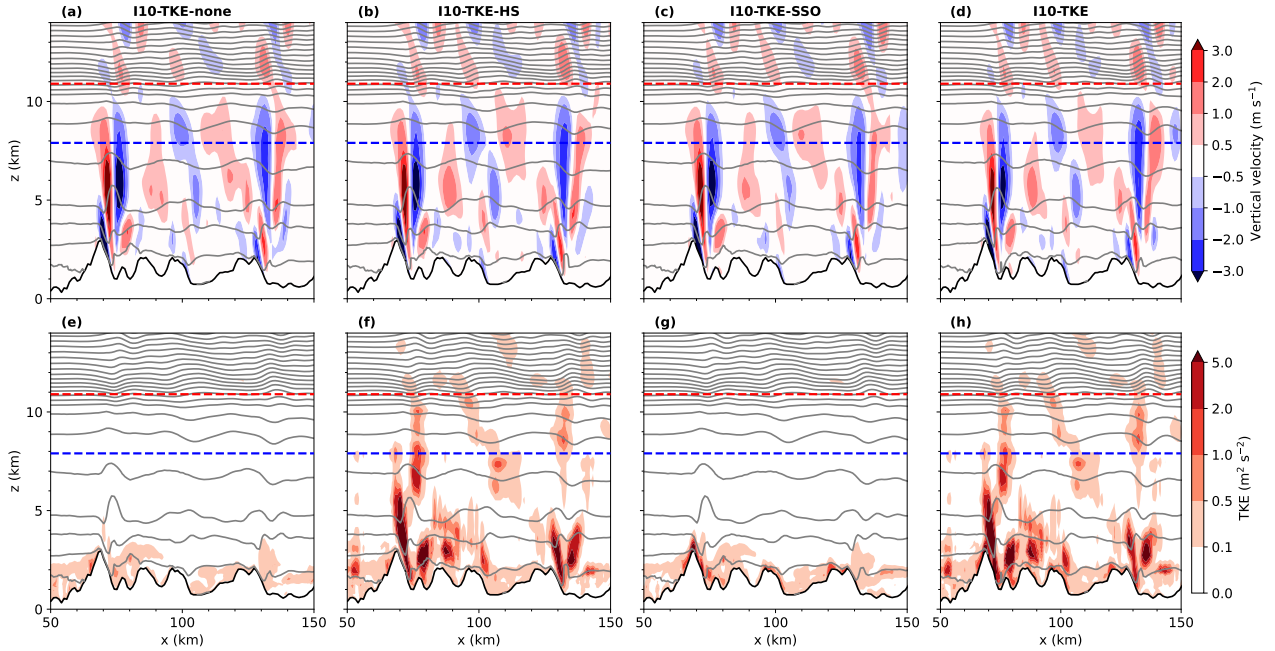


Figure 9. Comparison of vertical velocity and TKE along the southern flight-path transect at 18:00 UTC for ICON simulations using the TKE scheme: (a, e) both horizontal shear (HS) and subgrid-scale orography (SSO) deactivated, (b, f) HS activated, (c, g) SSO activated, and (d, h) both HS and SSO activated (standard operational configuration).

3.4 Bulk properties of the GW field

3.4.1 Vertical velocity variance and TKE

Domain-averaged profiles of resolved vertical velocity variance ($\overline{w'^2}$) and parameterized TKE highlight important differences in how resolution and turbulence schemes represent small-scale energy (Fig. 10). Among the 2TE simulations, the 0.5 km run performs particularly well: its variance profile aligns closely with the LES in the lower troposphere, indicating that fine-scale vertical motions are reasonably captured at this resolution. In contrast, coarser runs at 1 km and 2 km systematically underestimate the variance, with the largest discrepancy below 5 km in the layer containing the trapped lee waves. This highlights that quasi-LES resolutions of a few hundred meters are required to reproduce the vertical distribution of wave-induced resolved variance for this mountain wave event, which included low-level trapped lee waves.

Above the layer of trapped lee waves, the two simulations I10-2TE and I05-2TE agree very well, indicating near convergence with respect to the resolved variance associated with the dominant mountain waves. While the LES may be a more realistic reference at lower heights due to its improved representation of short-scale trapped lee waves, this is likely less true in the upper troposphere, where comparison with observations indicate that the LES underestimates the magnitude of the vertical velocity fluctuations. This is consistent with an over-dissipation of the near-grid-scale wave amplitude by the LES subgrid closure at

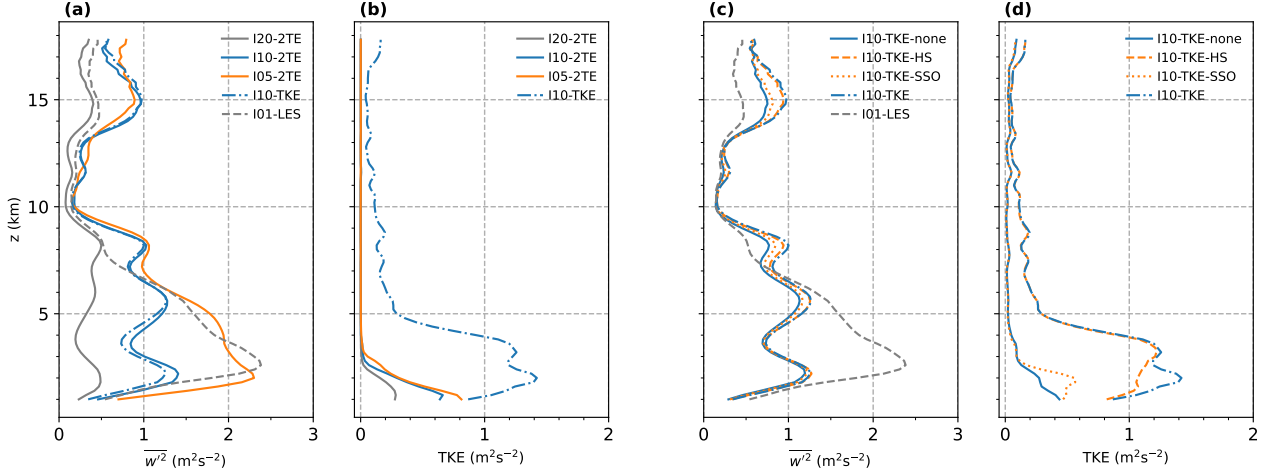


Figure 10. Profiles of resolved vertical velocity variance ($\overline{w'^2}$) and parameterized TKE, averaged over the analysis subdomain shown in Fig. 1, at 18:00 UTC for (a, b) different horizontal resolutions using the 2TE scheme and I10-TKE and (c, d) different TKE-scheme configurations, together with the LES.

marginal resolution: the LES resolves the wavelength and phase but damps the amplitude of the shortest resolved waves. The LES is therefore not automatically the more realistic reference.

At 1 km resolution, the resolved variances ($\overline{w'^2}$) are broadly similar between the two turbulence schemes, but the TKE fields differ substantially. Simulations with the TKE scheme generate larger TKE values than those with the 2TE scheme, particularly when the horizontal shear term is active (the default setup of the scheme). These inflated values contrast with the more realistic profiles from the 2TE simulations and from the TKE scheme without the empirical shear term. Together, these findings emphasize that, while resolved variances converge toward LES at sufficiently fine resolution, the representation of TKE remains highly sensitive to the choice of turbulence formulation.

3.4.2 Wave momentum fluxes

Domain-averaged profiles of vertical momentum flux at three consecutive times illustrate how resolution and turbulence schemes influence the representation of wave-induced transport (Fig. 11). In the 2TE simulations, the low-level momentum flux increases with increasing resolution as the orography is better resolved. With the exception of the last time (19:00 UTC), the profiles for the two higher-resolution simulations, I10-2TE and I05-2TE, are similar, indicating near convergence of the area-averaged momentum flux at about 1 km resolution. The momentum flux profile for I10-TKE is very close to that for I10-2TE for all three times, despite the large differences in TKE between the two schemes. This suggests that the empirical horizontal shear term in the TKE scheme inflates the TKE but has little influence on the bulk momentum flux. Because the area-averaged momentum flux, and hence its vertical divergence (the wave drag), is nearly insensitive to the horizontal shear term (Fig. 11), the additional TKE is not accompanied by a corresponding momentum deposition. Turbulence production asso-

ciated with wave breaking would leave a signature in the drag; the horizontal-shear TKE does not. We interpret this decoupling
335 as further evidence that the additional TKE is spurious.

Overall, the results highlight that area-averaged quantities, such as the area-averaged momentum flux, generally exhibit a
more robust and monotonic behavior with increasing resolution than local quantities such as vertical velocity or local TKE
fields. Local fields are strongly influenced by phase shifts, small-scale variability, and non-stationarity, making convergence
assessments more difficult and less meaningful. In contrast, spatial averaging reduces this sensitivity and therefore provides
340 a clearer indication of resolution effects, making momentum flux a more robust diagnostic for evaluating mountain-wave
simulations.

4 Conclusions

This study investigated the simulation of mountain waves with the ICON model, focusing on the influence of horizontal and
vertical resolution and the representation of turbulence through different parameterization schemes. We examined the perfor-
345 mance of the newly implemented two-energy scheme (2TE scheme) in reproducing mountain-wave structures and associated
turbulence during a case from the DEEPWAVE campaign over New Zealand on 12 July 2014. Model simulations were com-
pared with high-frequency aircraft observations and a nested LES run for detailed visual and statistical evaluation.

Our results demonstrated that both vertical and horizontal resolutions play a critical role in accurately capturing the structure
and behavior of mountain waves. We find that a vertical grid spacing of less than 200 m is necessary to adequately resolve the
350 wave field in the UTLS region. Insufficient vertical resolution damps the wave amplitude, leading to underestimation of wave
strength. With respect to horizontal resolution, simulations at 0.5 km approached the LES in terms of wavelength, while coarser
runs (1-2 km) underestimated the number and detail of shorter-scale waves. Differences between NWP and LES simulations
also point to the importance of how the ABL is represented, as ABL coupling affects lee-side flow separation, and thus the
triggering of mountain waves.

355 Comparison with observations shows, however, that the LES is not automatically more realistic. While the LES may be a
more realistic reference at lower heights due to its improved representation of orography and short-scale trapped lee waves,
this is likely less true in the upper troposphere, where comparison with observations indicate that the LES underestimates the
magnitude of the vertical velocity fluctuations. This may reflect that turbulence and wave breaking remain under-resolved even
at 200 m resolution, leading to excessive damping of wave amplitudes.

360 The two turbulence schemes produced very similar resolved wave fields and comparable area-averaged momentum-flux
profiles, but differed strongly in parameterized TKE. This difference originates in the empirical horizontal-shear source term
of the operational TKE-scheme configuration, which produces spurious TKE at km-scale resolution, as shown by the controlled
source-term experiments (Fig. 9) and by the scaling of the term (Appendix A). The 2TE scheme reproduces the resolved wave
field as well as the operational scheme, remains numerically well behaved in the strongly stratified UTLS, and does not exhibit
365 this artifact, which makes it a viable alternative for such applications. We do not, however, claim that the 2TE scheme is more
accurate in representing turbulence. That a substantial part of the gravity-wave field in this case is non-breaking, including the

primary wave at flight level, is what makes the spurious horizontal-shear TKE clearly identifiable: no significant turbulence is expected in such regions, and none is observed.

370 Although differences in the resolved flow fields are comparatively small, an accurate representation of TKE is critical for many applications, including aviation turbulence forecasting and the transport and dispersion of tracers such as greenhouse gases and air pollutants. Spurious or overly damped turbulence can directly affect estimates of mixing and exchange between atmospheric layers.

Hence, our findings suggest that careful reevaluation of turbulence parameterizations is required when using high-resolution NWP models in mountainous terrain. For km-scale resolutions, model setups should consider disabling or tuning the horizontal
375 shear and SSO terms in the TKE scheme to avoid spurious TKE production. Alternatively, schemes such as 2TE provide a promising framework for turbulence representation at these scales.

The results have important implications for the configuration of high-resolution regional weather models and for advancing the understanding of terrain-induced gravity waves and associated turbulent processes. Future work will extend this analysis to additional DEEPWAVE cases and other campaigns to test the robustness of these conclusions. A further priority is coupling the
380 2TE scheme with gravity-wave parameterizations to bridge explicitly resolved and parameterized wave–turbulence interactions, particularly in global and climate modeling frameworks.

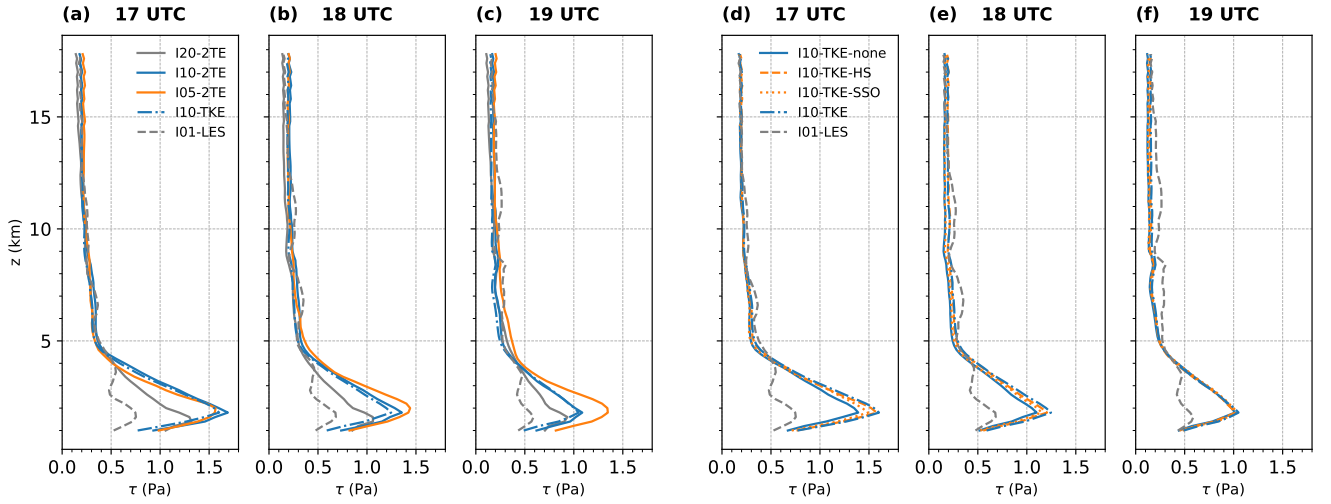


Figure 11. Profiles of area-averaged momentum flux, τ , for (a–c) different horizontal resolutions using the 2TE scheme and I10-TKE, and (d–f) different TKE-scheme configurations at three consecutive times, together with the LES. Fluxes are calculated within the analysis subdomain shown in Fig. 1 after interpolation to regular vertical levels with 200 m spacing.

Appendix A: Overview of the TKE and 2TE turbulence schemes in ICON

A1 TKE budget and the operational ICON TKE scheme

The general form of the TKE budget equation in a coordinate system aligned with the mean wind, under horizontal homogeneity and negligible subsidence, is

$$\underbrace{\frac{\partial \bar{e}}{\partial t}}_{\text{I}} = \underbrace{\frac{g}{\theta_v} (\overline{w'\theta'_v})}_{\text{II}} - \underbrace{\overline{u'w'} \frac{\partial U}{\partial z}}_{\text{III}} - \underbrace{\frac{\partial (\overline{w'e})}{\partial z}}_{\text{IV}} - \underbrace{\frac{1}{\bar{\rho}} \frac{\partial (\overline{w'p'})}{\partial z}}_{\text{V}} - \underbrace{\varepsilon}_{\text{VI}} \quad (\text{A1})$$

where term I represents the local tendency of TKE; term II the buoyancy production or destruction; term III the shear production; term IV the turbulent transport term; term V the pressure transport; and term VI the viscous dissipation rate (Stull, 1988).

The operational turbulence scheme in ICON (Raschendorfer, 2001) is based on this classical TKE framework but augments it by explicitly separating turbulence from other small-scale motions such as gravity-wave-induced momentum tendencies and horizontal deformation. The corresponding prognostic equation for TKE is

$$\begin{aligned} \frac{\partial (q^2/2)}{\partial t} = & \underbrace{\overline{\beta w'\theta'_v}}_{\text{Buoyancy}} - \underbrace{\left(\overline{u'_1 w'} \frac{\partial U_1}{\partial z} + \overline{u'_2 w'} \frac{\partial U_2}{\partial z} \right)}_{\text{Vertical Shear}} + \underbrace{\frac{\partial}{\partial z} \left[\alpha_T \lambda q \frac{\partial (q^2/2)}{\partial z} \right]}_{\text{Turbulent Transport}} - \underbrace{\frac{q^3}{\alpha_M \lambda}}_{\text{Dissipation}} \\ & + \underbrace{U_1 \frac{\partial U_1}{\partial t} \bigg|_{\text{SSO}} + U_2 \frac{\partial U_2}{\partial t} \bigg|_{\text{SSO}}}_{\text{Subgrid-scale Orography}} + \underbrace{S_H}_{\text{Horizontal Shear}}. \end{aligned} \quad (\text{A2})$$

Here, the subscripts 1 and 2 denote the two horizontal wind components (U_1, U_2) and $q = \sqrt{2\bar{e}}$ is the turbulent velocity scale. The additional terms account for (i) contributions from the SSO parameterization, which represents drag and gravity-wave momentum deposition from unresolved topography; and (ii) turbulence generation associated with large-scale horizontal shear, deformation, and divergence (an empirical source term). The purpose of these extensions is to enable the turbulence closure to operate also in stable layers above the boundary layer and to provide diagnostic quantities such as the eddy dissipation rate (EDR), which is used in aviation turbulence applications (Goecke and Machulskaya, 2021).

The horizontal-shear source term follows a two-dimensional Smagorinsky-type closure (Goecke and Machulskaya, 2021) that relates the horizontal-shear production to the deformation of the resolved flow, under the assumption that this deformation is associated, on average, with a net downscale transfer of kinetic energy to unresolved turbulence,

$$S_H = l_h^2 \sqrt{C^3 \alpha_M} \left(\sqrt{\text{DEF}^2 + c_1^2 \text{DIV}^2} - c_1 \text{DIV} \right)^3, \quad l_h = \frac{a_h l}{\text{Re}^{2/3}}, \quad a_h = 0.2,$$

where DEF is the horizontal deformation, DIV the horizontal divergence, l the grid spacing, and C a tuning constant. The source therefore scales approximately as $S_H \sim l^2 |\nabla_h U|^3 \text{Re}^{-4/3}$, i.e. as a grid-scale two-dimensional Smagorinsky production whose length scale is set by the grid spacing rather than by a physical turbulence scale. The applicability of this interpretation at km-scale resolution is discussed in Sect. 3.3.

Because these additional turbulence sources arise from parameterized processes rather than directly from the resolved flow, the resulting TKE tendencies may not always reflect the local dynamical response in situations dominated by wave propagation and breaking.

A2 Two-energy (2TE) scheme formulation

The 2TE scheme (Bařták řurán et al., 2022) provides a unified framework for representing turbulence, shallow convection, and subgrid-scale cloud processes. It introduces two prognostic energy equations: one for TKE and one for the total turbulence energy (TTE). The prognostic TKE equation is

$$\frac{\partial e_k}{\partial t} = \underbrace{\frac{g}{\theta_0} \overline{\theta'_v w'}}_{\text{Buoyancy (APDF)}} - \underbrace{\left(\overline{u'_1 w'} \frac{\partial U_1}{\partial z} + \overline{u'_2 w'} \frac{\partial U_2}{\partial z} \right)}_{\text{Vertical Shear}} + \underbrace{\frac{\partial}{\partial z} \left(K_{e_k} \frac{\partial e_k}{\partial z} \right)}_{\text{Turbulent Transport}} - \underbrace{\frac{2e_k}{\tau_k}}_{\text{Dissipation}} \quad (\text{A3})$$

and the accompanying equation for TTE is

$$\frac{\partial e_s}{\partial t} = \left(\overline{u'_1 w'} \frac{\partial U_1}{\partial z} + \overline{u'_2 w'} \frac{\partial U_2}{\partial z} \right) + \frac{\partial}{\partial z} \left(K_{e_s} \frac{\partial e_s}{\partial z} \right) - \frac{2e_s}{\tau_s}, \quad (\text{A4})$$

where e_k is the turbulent kinetic energy (TKE), e_s the total turbulence energy (TTE), and τ_k , τ_s and K_{e_k} , K_{e_s} are the corresponding dissipation time scales and diffusion coefficients.

A key feature of the scheme is the quasi-prognostic stability parameter Ri_f^* , defined as a combination of a non-local and a local estimate of the flux Richardson number,

$$Ri_f^* = a Ri_f^{TE} + (1 - a) Ri_f^{GR}, \quad (\text{A5})$$

where Ri_f^{TE} is calculated from the two prognostic turbulence energies. This nonlinearity and non-locality help suppress the oscillatory behavior often encountered in classical TKE closures. In contrast to the TKE scheme, the mixing length in the 2TE scheme is dynamically adjusted based on the diagnosed ABL height and the bulk state of the ABL (convective versus stable). Another important aspect is the formulation of the stability functions, which do not rely on a critical gradient Richardson number. This guarantees realizability of the scheme over the full range of stratification conditions.

A notable technical property of the 2TE formulation is that it yields numerically stable solutions without requiring a prescribed minimum diffusion coefficient. This avoids the need for ad hoc lower bounds on vertical diffusivity that are often introduced to maintain numerical stability in classical TKE closures. Allowing diffusivities and mixing lengths to approach small values organically preserves the intended behavior of turbulence in weakly turbulent or strongly stratified layers.

The 2TE scheme is coupled to the APDF framework for diagnosing buoyancy production. The APDF uses a double-Gaussian joint PDF of vertical velocity, liquid water potential temperature, and total water content. Although originally developed for cloud-topped boundary layers, this coupling ensures a consistent representation of buoyancy effects based on the diagnosed turbulent state. In non-cloudy regimes such as the UTLS, buoyancy production is generally small; however, the APDF coupling allows the scheme to represent vertical-velocity skewness and associated flux asymmetries, which may occur in wave-breaking regions.

Focusing on the UTLS region, the most relevant characteristics of the 2TE scheme compared to the TKE scheme are: (i) the local–nonlocal stability parameter, which strongly reduces the occurrence of oscillations and thereby allows smaller exchange coefficients without risking numerical instability, thus preventing the overmixing that can erode thin stably stratified layers; and (ii) stability functions that guarantee realizability over the full range of gradient Richardson numbers.

Appendix B: Additional horizontal cross sections

Code and data availability. The airborne measurement data from the DEEPWAVE campaign are available through the HALO database (<https://doi.org/10.17616/R39Q0T>; HALO-DB, 2022). We thank Hans-Christoph Lachnitt for providing the post-processed data in a convenient format. The ICON model is distributed under an open-source BSD-3C license (<https://www.icon-model.org>) and is publicly available at <https://gitlab.dkrz.de/icon/icon-model/-/releases/icon-2025.04-1-public> (GitLab, 2025). The ICON-2TE implementation is included in the ICON source code and can be made available on request. The simulation output, processing resources, and analysis scripts used to generate the figures in this paper are available from the corresponding author upon request.

Author contributions. RS: conceptualization, model implementation and simulations, data curation, formal analysis, visualization, writing — original draft, and writing — review and editing. JS: conceptualization, methodology, model scheme development, funding acquisition, project administration, supervision, software, and writing — review and editing.

Competing interests. The authors have declared that there are no competing interests.

Acknowledgements. The authors thank the German Research Foundation (DFG) for financial support through CRC 301 TPChange (Project-ID 428312742, sub-projects B06 “Impact of small-scale dynamics on UTLS transport and mixing”). This work used resources of the Deutsches Klimarechenzentrum (DKRZ), granted by its Scientific Steering Committee (WLA) under project ID bb1096. We thank Hans-Christoph Lachnitt and Peter Hoor for providing the processed observational data for the case study from the HALO database, and Tobias Goecke for constructive discussions. The first author acknowledges the initial technical support of Shweta Singh for running the ICON model.

References

- 460 Achatz, U., Alexander, M. J., Becker, E., Chun, H.-Y., Holt, L., Plougonven, R., Polichtchouk, I., Sato, K., Sheshadri, A., Stephan, C. C., Niekerk, A. V., and Wright, C. J.: Atmospheric Gravity Waves: Processes and Parameterization, *JOURNAL OF THE ATMOSPHERIC SCIENCES*, 81, 2024.
- Arbor, A.: A Numerical Study of Gravity Wave Breaking and Impacts on Turbulence and Mean State, *JOURNAL OF THE ATMOSPHERIC SCIENCES*, 56, 1999.
- 465 Bašták Ďurán, I., Geleyn, J.-F., Váňa, F., Schmidli, J., and Brožková, R.: A Turbulence Scheme with Two Prognostic Turbulence Energies, *Journal of the Atmospheric Sciences*, 75, 3381–3402, <https://doi.org/10.1175/JAS-D-18-0026.1>, 2018.
- Bašták Ďurán, I., Sakradzija, M., and Schmidli, J.: The Two-Energies Turbulence Scheme Coupled to the Assumed PDF Method, *Journal of Advances in Modeling Earth Systems*, 14, e2021MS002922, <https://doi.org/10.1029/2021MS002922>, 2022.
- Bougeault, P., Jansa, A., Attié, J.-L., Beau, I., Benesch, B., Benoit, R., Bessemoulin, P., Caccia, J., Campins, J., Carissimo, B., Champeaux, J.-L., Crochet, M., Druilhet, A., Durand, P., El Khalfi, A., Flamant, P., A, G., Georgelin, M., Hoinka, K., and Tannhauser, D.: The atmospheric momentum budget over a major mountain range: First results of the PYREX field program, *Annales Geophysicae*, 11, 395–418, 1993.
- 470 Bougeault, P., Binder, P., Buzzzi, A., Dirks, R., Houze, R., Kuettner, J., Smith, R. B., Steinacker, R., and Volkert, H.: The MAP Special Observing Period, *Bulletin of the American Meteorological Society*, 82, 433–462, [https://doi.org/10.1175/1520-0477\(2001\)082<0433:TMSOP>2.3.CO;2](https://doi.org/10.1175/1520-0477(2001)082<0433:TMSOP>2.3.CO;2), publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, 2001.
- 475 Bramberger, M., Ewald, F., and Sharman, R.: Mountain-Wave Turbulence Encounter of the Research Aircraft HALO above Iceland, *JOURNAL OF APPLIED METEOROLOGY AND CLIMATOLOGY*, 59, 2020.
- Durran, D.: MOUNTAIN METEOROLOGY Lee Waves and Mountain Waves, in: *Encyclopedia of Atmospheric Sciences*, edited by Gerald R. North, John Pyle, and Fuqing Zhang, vol. 4, pp. 95–102, Academic Press, 2nd edn., ISBN 978-0-12-382225-3, <https://doi.org/10.1016/B978-0-12-382225-3.00202-4>, <https://www.sciencedirect.com/science/article/pii/B9780123822253002024>, 2015.
- 480 Dörnbrack, A., Gerz, T., and Schumann, U.: Turbulent breaking of overturning gravity waves below a critical level, *Applied Scientific Research*, 54, 163–195, <https://doi.org/10.1007/BF00849114>, 1995.
- Fritts, D. C. and Alexander, M. J.: Gravity wave dynamics and effects in the middle atmosphere, *Reviews of Geophysics*, 41, 2001RG000106, <https://doi.org/10.1029/2001RG000106>, 2003.
- 485 Fritts, D. C., Smith, R. B., Taylor, M. J., Doyle, J. D., Eckermann, S. D., Dörnbrack, A., Rapp, M., Williams, B. P., Pautet, P.-D., Bossert, K., Criddle, N. R., Reynolds, C. A., Reinecke, P. A., Uddstrom, M., Revell, M. J., Turner, R., Kaifler, B., Wagner, J. S., Mixa, T., Kruse, C. G., Nugent, A. D., Watson, C. D., Gisinger, S., Smith, S. M., Lieberman, R. S., Laughman, B., Moore, J. J., Brown, W. O., Haggerty, J. A., Rockwell, A., Stossmeister, G. J., Williams, S. F., Hernandez, G., Murphy, D. J., Klekociuk, A. R., Reid, I. M., and Ma, J.: The Deep Propagating Gravity Wave Experiment (DEEPWAVE): An Airborne and Ground-Based Exploration of Gravity Wave Propagation and Effects from Their Sources throughout the Lower and Middle Atmosphere, *Bulletin of the American Meteorological Society*, 97, 425–453, <https://doi.org/10.1175/BAMS-D-14-00269.1>, 2016.
- 490 Fritts, D. C., Lund, T. S., Wan, K., and Liu, H.-L.: Numerical Simulation of Mountain Waves over the Southern Andes. Part II: Momentum Fluxes and Wave–Mean-Flow Interactions, *Journal of the Atmospheric Sciences*, 78, 3069–3088, <https://doi.org/10.1175/JAS-D-20-0207.1>, 2021.

- 495 Fritts, D. C., Lund, A. C., Lund, T. S., and Yudin, V.: Impacts of Limited Model Resolution on the Representation of Mountain Wave and Secondary Wave Dynamics in Local and Global Models: 2. Mountain Wave and Secondary Wave Evolutions in the Thermosphere, *Journal of Geophysical Research: Atmospheres*, 127, e2021JD036035, <https://doi.org/10.1029/2021JD036035>, 2022.
- Goecke, T. and Machulskaya, E.: Aviation Turbulence Forecasting at DWD with ICON: Methodology, Case Studies, and Verification, *Monthly Weather Review*, 149, 2115–2130, <https://doi.org/10.1175/MWR-D-19-0383.1>, 2021.
- 500 Grubišić, V. and Lewis, J. M.: SIERRA WAVE PROJECT REVISITED: 50 Years Later, *Bulletin of the American Meteorological Society*, 85, 1127–1142, <https://doi.org/10.1175/BAMS-85-8-1127>, 2004.
- Grubišić, V., Doyle, J. D., Kuettner, J., Mobbs, S., Smith, R. B., Whiteman, C. D., Dirks, R., Czyzyk, S., Cohn, S. A., Vosper, S., Weissmann, M., Haimov, S., De Wekker, S. F. J., Pan, L. L., and Chow, F. K.: THE TERRAIN-INDUCED ROTOR EXPERIMENT: A Field Campaign Overview Including Observational Highlights, *Bulletin of the American Meteorological Society*, 89, 1513–1534, <https://doi.org/10.1175/2008BAMS2487.1>, 2008.
- 505 Heale, C. J., Bossert, K., and Vadas, S. L.: 3D Numerical Simulation of Secondary Wave Generation From Mountain Wave Breaking Over Europe, *Journal of Geophysical Research: Atmospheres*, 127, <https://doi.org/10.1029/2021JD035413>, e2021JD035413, 2022.
- Hogan, R. J. and Bozzo, A.: A Flexible and Efficient Radiation Scheme for the ECMWF Model, *Journal of Advances in Modeling Earth Systems*, 10, 1990–2008, <https://doi.org/10.1029/2018MS001364>, 2018.
- 510 Holt, L. A., Alexander, M. J., Coy, L., Liu, C., Molod, A., Putman, W., and Pawson, S.: An evaluation of gravity waves and gravity wave sources in the Southern Hemisphere in a 7 km global climate simulation, *Quarterly Journal of the Royal Meteorological Society*, 143, 2481–2495, <https://doi.org/10.1002/qj.3101>, 2017.
- Jackson, D. R., Gadian, A., Hindley, N. P., Hoffmann, L., Hughes, J., King, J., Moffat-Griffin, T., Moss, A. C., Ross, A. N., Vosper, S. B., Wright, C. J., and Mitchell, N. J.: The South Georgia Wave Experiment: A Means for Improved Analysis of Gravity Waves and Low-Level Wind Impacts Generated from Mountainous Islands, *Bulletin of the American Meteorological Society*, 99, 1027–1040, <https://doi.org/10.1175/BAMS-D-16-0151.1>, 2018.
- 515 Jiang, Q., Smith, R. B., and Doyle, J. D.: Impact of the Atmospheric Boundary Layer on Mountain Waves, *Journal of the Atmospheric Sciences*, 65, 592–608, <https://doi.org/10.1175/2007JAS2376.1>, 2008.
- Kim, S.-H., Kim, J.-H., Chun, H.-Y., and Sharman, R. D.: Global response of upper-level aviation turbulence from various sources to climate change, *npj Climate and Atmospheric Science*, 6, 92, <https://doi.org/10.1038/s41612-023-00421-3>, 2023.
- 520 Kleeorin, N., Rogachevskii, I., Soustova, I. A., Troitskaya, Y. I., Ermakova, O. S., and Zilitinkevich, S.: Internal gravity waves in the energy and flux budget turbulence-closure theory for shear-free stably stratified flows, *Physical Review E*, 99, 063106, <https://doi.org/10.1103/PhysRevE.99.063106>, 2019.
- Kuettner, J. P. and O'Neill, T. H. R.: ALPEX The GARP Mountain Subprogram, *Bulletin of the American Meteorological Society*, 62, 793–805, <https://doi.org/10.1175/1520-0477-62.6.793>, 1981.
- 525 Lachnitt, H.-C., Hoor, P., Kunkel, D., Bramberger, M., Dörnbrack, A., Müller, S., Reutter, P., Giez, A., Kaluza, T., and Rapp, M.: Gravity-wave-induced cross-isentropic mixing: a DEEPWAVE case study, *Atmospheric Chemistry and Physics*, 23, 355–373, <https://doi.org/10.5194/acp-23-355-2023>, 2023.
- Lott, F. and Miller, M. J.: A new subgrid-scale orographic drag parametrization: Its formulation and testing, *Quarterly Journal of the Royal Meteorological Society*, 123, 101–127, <https://doi.org/10.1002/qj.49712353704>, 1997.
- 530

- Lund, T. S., Fritts, D. C., Wan, K., Laughman, B., and Liu, H.-L.: Numerical Simulation of Mountain Waves over the Southern Andes. Part I: Mountain Wave and Secondary Wave Character, Evolutions, and Breaking, *Journal of the Atmospheric Sciences*, 77, 4337–4356, <https://doi.org/10.1175/JAS-D-19-0356.1>, 2020.
- Mahalov, A.: Turbulence and Waves in the Upper Troposphere and Lower Stratosphere, in: *Aviation Turbulence*, edited by Sharman, R. and Lane, T., pp. 407–424, Springer International Publishing, Cham, ISBN 978-3-319-23629-2 978-3-319-23630-8, https://doi.org/10.1007/978-3-319-23630-8_20, 2016.
- Mirza, A. K., Dacre, H. F., and Lo, C. H. B.: A case study analysis of the impact of a new free tropospheric turbulence scheme on the dispersion of an atmospheric tracer, *Quarterly Journal of the Royal Meteorological Society*, 150, 1907–1925, <https://doi.org/10.1002/qj.4681>, 2024.
- 540 Muñoz-Esparza, D., Sharman, R. D., and Trier, S. B.: On the Consequences of PBL Scheme Diffusion on UTLS Wave and Turbulence Representation in High-Resolution NWP Models, *MONTHLY WEATHER REVIEW*, 148, 2020.
- Paoli, R., Thouron, O., Escobar, J., Picot, J., and Cariolle, D.: High-resolution large-eddy simulations of stably stratified flows: application to subkilometer-scale turbulence in the upper troposphere–lower stratosphere, *Atmospheric Chemistry and Physics*, 14, 5037–5055, <https://doi.org/10.5194/acp-14-5037-2014>, 2014.
- 545 Plougonven, R., de la Cámara, A., Hertzog, A., and Lott, F.: How does knowledge of atmospheric gravity waves guide their parameterizations?, *Quarterly Journal of the Royal Meteorological Society*, 146, 1529–1543, <https://doi.org/10.1002/qj.3732>, <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.3732>, 2020.
- Rapp, M., Kaifler, B., Dörnbrack, A., Gisinger, S., Mixa, T., Reichert, R., Kaifler, N., Knobloch, S., Eckert, R., Wildmann, N., Giez, A., Krasauskas, L., Preusse, P., Geldenhuys, M., Riese, M., Woiwode, W., Friedl-Vallon, F., Sinnhuber, B.-M., Torre, A. D. L., Alexander, P.,
- 550 Hormaechea, J. L., Janches, D., Garhammer, M., Chau, J. L., Conte, J. F., Hoor, P., and Engel, A.: SOUTHTRAC-GW: An Airborne Field Campaign to Explore Gravity Wave Dynamics at the World’s Strongest Hotspot, *Bulletin of the American Meteorological Society*, 102, E871–E893, <https://doi.org/10.1175/BAMS-D-20-0034.1>, 2021.
- Raschendorfer, M.: The new turbulence parameterization of LM, *COSMO Newsletter*, 1, 89–97, <http://www.cosmo-model.org>, 2001.
- Ren, D. and Lynch, M. J.: Changes in Global Aviation Turbulence in the Remote Sensing Era (1979–2018), *Remote Sensing*, 16, <https://doi.org/10.3390/rs16112038>, 2024.
- 555 Riese, M., Ploeger, F., Rap, A., Vogel, B., Konopka, P., Dameris, M., and Forster, P.: Impact of uncertainties in atmospheric mixing on simulated UTLS composition and related radiative effects, *Journal of Geophysical Research: Atmospheres*, 117, 2012JD017751, <https://doi.org/10.1029/2012JD017751>, 2012.
- Rodriguez Imazio, P., Mininni, P. D., Godoy, A., Rivaben, N., and Dörnbrack, A.: Not All Clear Air Turbulence Is Kolmogorov—The Fine-Scale Nature of Atmospheric Turbulence, *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037491, <https://doi.org/10.1029/2022JD037491>, 2023.
- 560 Rogel, L., Ricard, D., Bazile, E., and Sandu, I.: Effects of subgrid-scale turbulence parametrization on the representation of clear-air turbulence using kilometre- to hectometre-scale numerical simulations, *Quarterly Journal of the Royal Meteorological Society*, 149, 3301–3322, <https://doi.org/10.1002/qj.4557>, 2023.
- 565 Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J. D., McTaggart-Cowan, R., Methven, J., Rivière, G., Ament, F., Boettcher, M., Bramberger, M., Cazenave, Q., Cotton, R., Crewell, S., Delanoë, J., Dörnbrack, A., Ehrlich, A., Ewald, F., Fix, A., Grams, C. M., Gray, S. L., Grob, H., Groß, S., Hagen, M., Harvey, B., Hirsch, L., Jacob, M., Kölling, T., Konow, H., Lemmerz, C., Lux, O., Magnusson, L., Mayer, B., Mech, M., Moore, R., Pelon, J., Quinting, J., Rahm, S., Rapp, M., Rautenhaus, M., Reitebuch, O., Reynolds, C. A., Sodemann,

- H., Spengler, T., Vaughan, G., Wendisch, M., Wirth, M., Witschas, B., Wolf, K., and Zinner, T.: The North Atlantic Waveguide and Downstream Impact Experiment, *Bulletin of the American Meteorological Society*, 99, 1607–1637, <https://doi.org/10.1175/BAMS-D-17-0003.1>, 2018.
- Scherllin-Pirscher, B., Steiner, A. K., Anthes, R. A., Alexander, M. J., Alexander, S. P., Biondi, R., Birner, T., Kim, J., Randel, W. J., Son, S.-W., Tsuda, T., and Zeng, Z.: Tropical Temperature Variability in the UTLS: New Insights from GPS Radio Occultation Observations, *Journal of Climate*, 34, 2813–2838, <https://doi.org/10.1175/JCLI-D-20-0385.1>, 2021.
- 575 Schneider, A., Wagner, J., Faber, J., Gerding, M., and Lübken, F.-J.: Case study of wave breaking with high-resolution turbulence measurements with LITOS and WRF simulations, *Atmospheric Chemistry and Physics*, 2017.
- Schulz, J.-P., Vogel, G., Becker, C., Kothe, S., Rummel, U., and Ahrens, B.: Evaluation of the ground heat flux simulated by a multi-layer land surface scheme using high-quality observations at grass land and bare soil, *Meteorologische Zeitschrift*, 25, 607–620, <https://doi.org/10.1127/metz/2016/0537>, 2016.
- 580 Schäfler, A., Sprenger, M., Wernli, H., Fix, A., and Wirth, M.: Case study on the influence of synoptic-scale processes on the paired H₂ O–O₃ distribution in the UTLS across a North Atlantic jet stream, *Atmospheric Chemistry and Physics*, 23, 999–1018, <https://doi.org/10.5194/acp-23-999-2023>, 2023.
- Seifert, A.: A Revised Cloud Microphysical Parameterization for COSMO-LME, COSMO Newsletter Proceedings from the 8th COSMO General Meeting in Bucharest, 2006, Consortium for Small-Scale Modelling., pp. 25–28, <http://www.cosmo-model.org>, 2008.
- 585 Sharman, R. D., Trier, S. B., Lane, T. P., and Doyle, J. D.: Sources and dynamics of turbulence in the upper troposphere and lower stratosphere: A review, *Geophysical Research Letters*, 39, 2012GL051996, <https://doi.org/10.1029/2012GL051996>, 2012.
- Singh, S., Schmidli, J., Bašták Ďurán, I., and Westerhuis, S.: Impact of the Turbulence Parameterization on Simulations of Fog Over Complex Terrain, *Journal of Geophysical Research: Atmospheres*, 130, e2024JD042610, <https://doi.org/10.1029/2024JD042610>, 2025.
- Stull, R. B., ed.: *An Introduction to Boundary Layer Meteorology*, Springer Netherlands, Dordrecht, ISBN 978-90-277-2769-5 978-94-009-3027-8, <https://doi.org/10.1007/978-94-009-3027-8>, 1988.
- 590 Voelker, G. S.: MS-GWaM: A Three-Dimensional Transient Gravity Wave Parametrization for Atmospheric Models, *JOURNAL OF THE ATMOSPHERIC SCIENCES*, 81, 2024.
- Wratt, D. S., Ridley, R. N., Sinclair, M. R., Larsen, H., Thompson, S. M., Henderson, R., Austin, G. L., Bradley, S. G., Auer, A., Sturman, A. P., Owens, I., Fitzharris, B., Ryan, B. F., and Gayet, J.-F.: The New Zealand Southern Alps Experiment, *Bulletin of the American Meteorological Society*, 77, 683–692, [https://doi.org/10.1175/1520-0477\(1996\)077<0683:TNZSAE>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0683:TNZSAE>2.0.CO;2), 1996.
- 595 Zängl, G., Reinert, D., Rípodas, P., and Baldauf, M.: The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, *Quarterly Journal of the Royal Meteorological Society*, 141, 563–579, <https://doi.org/10.1002/qj.2378>, [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2378](https://onlinelibrary.wiley.com/doi/pdf/10.1002/qj.2378), 2015.

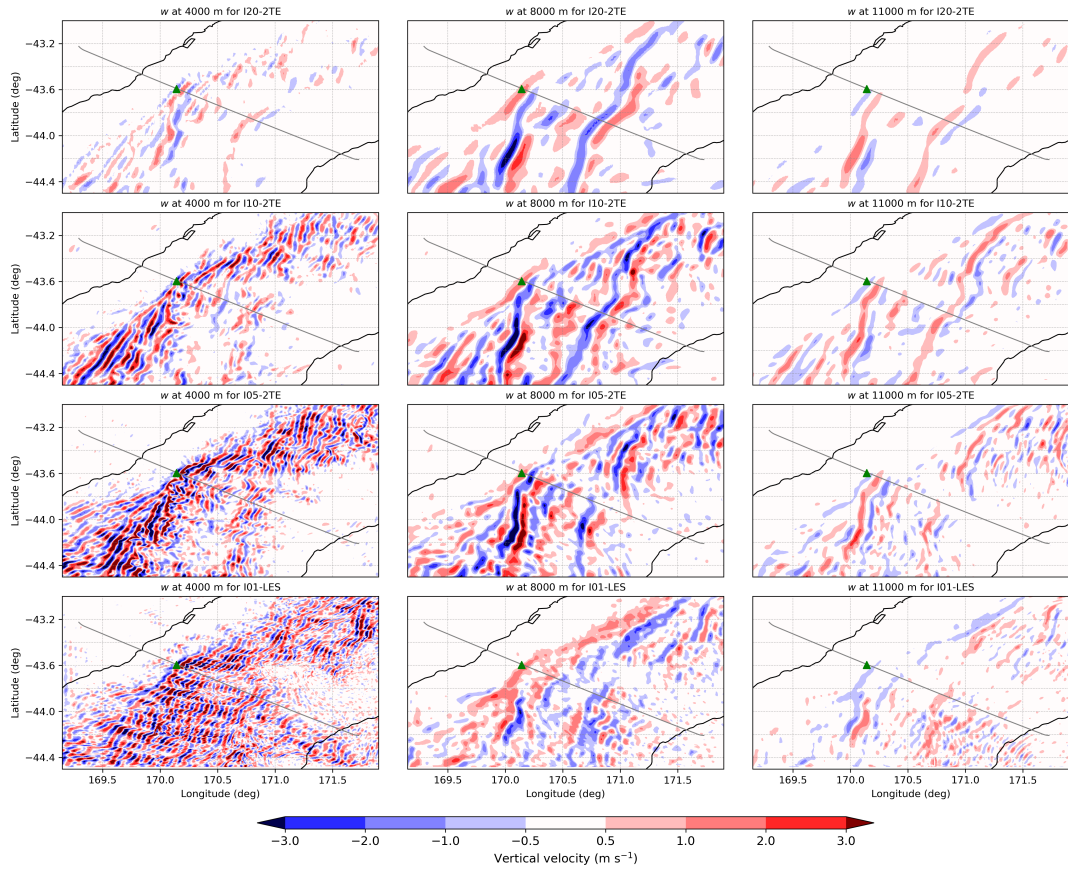


Figure B1. Horizontal cross-sections of vertical velocity at 8000 m and 11 000 m, corresponding to the lower and upper flight legs, for ICON simulations using the 2TE turbulence scheme at different horizontal resolutions, together with the LES.

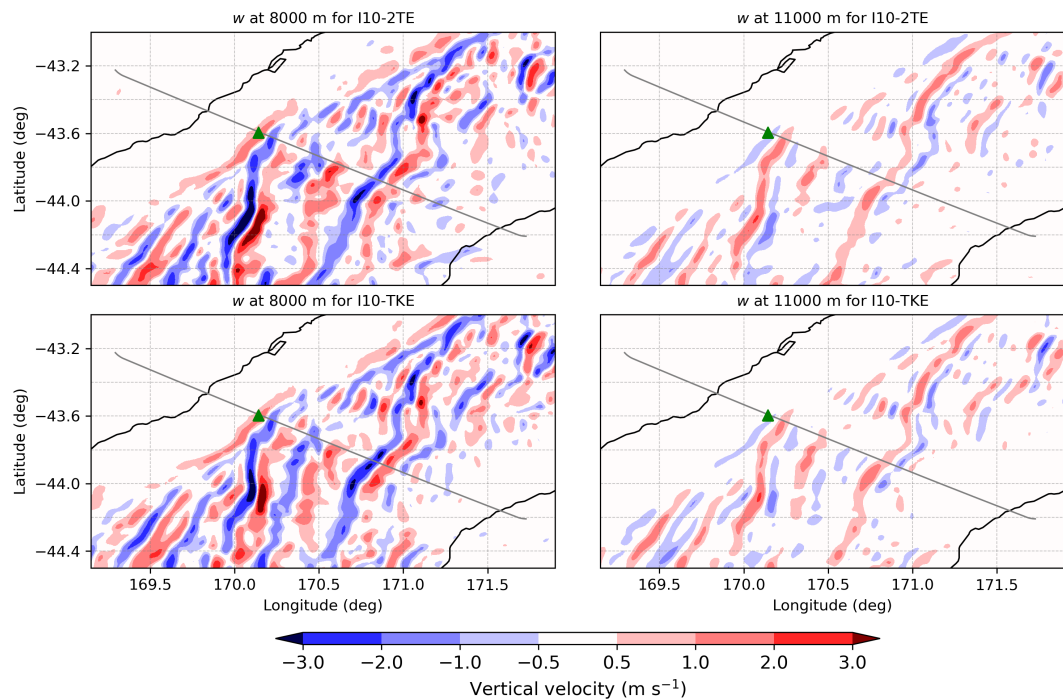


Figure B2. Horizontal cross-sections of vertical velocity at 8000 m and 11 000 m, corresponding to the lower and upper flight legs, for the ICON simulations I10-2TE and I10-TKE. The two simulations produce broadly similar wave structures, indicating that differences between the schemes are primarily reflected in the parameterized TKE rather than in the resolved vertical velocity patterns.