

RESEARCH ARTICLE

The evaluation of boundary-layer turbulence in high-resolution numerical weather prediction simulations using Doppler lidar

Natalie J. Harvey¹  | Helen F. Dacre¹  | Kirsty E. Hanley²  |
Humphrey W. Lean² | Christopher J. Walden^{3,4} | Michael Baidu⁵  |
Steven Boeing⁵  | Andrew N. Ross⁵

¹Department of Meteorology, University of Reading, Reading, UK

²MetOffice@Reading, Brian Hoskins Building, University of Reading, Reading, UK

³Science and Technology Facilities Council, Chilbolton Observatory, Stockbridge, UK

⁴National Centre for Atmospheric Science, Leeds, UK

⁵Institute for Climate and Atmospheric Science, University of Leeds, Leeds, UK

Correspondence

Natalie J. Harvey, Department of Meteorology, University of Reading, Berkshire, RG6 7BE, UK.
Email: n.j.harvey@reading.ac.uk

Funding information

Natural Environment Research Council, Grant/Award Numbers: NE/X018555/1, NE/Z503964/1, NE/X018539/1

Abstract

This study introduces a new generalisable methodology for evaluating the representation of boundary-layer turbulence in sub-kilometre numerical weather simulations through direct comparison with high-resolution Doppler lidar and sonic anemometer measurements. We derive key boundary-layer parameters, such as aerosol height, cloud-base height, turbulent mixing height, sensible heat flux, and boundary-layer classification, using an established observational algorithm and similar diagnostics from newly available high temporal resolution outputs from UKV (1.5 km) and WMV (300 m) simulations of the Unified Model. For the case study considered here, there is agreement in the temporal evolution and magnitude of bulk boundary-layer characteristics for both simulations, including the timing of the morning transition from stable to unstable surface conditions. However, significant differences are found in both the shape and magnitude of vertical velocity variance profiles. In the UKV simulations, peak vertical velocity variance is underestimated by up to a factor of 5 compared with observations, while differences in the WMV simulations are smaller (up to a factor of 3). In the UKV simulations, turbulence is predominately unresolved (parametrized). By contrast, in the WMV over 90% of the vertical velocity variance is resolved explicitly in the middle of the boundary layer once the convective regime is established. These results demonstrate that, for this case study, parametrized turbulence is underestimated in both numerical weather prediction simulations. More broadly, this work highlights the value of high-resolution observations in diagnosing model performance and provides a transferable evaluation framework applicable at any site with high-frequency simulation data and Doppler lidar observations.

KEYWORDS

convective boundary layer, Doppler lidar, turbulence parametrization, vertical velocity variance

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 Crown copyright. Met Office and The Author(s). *Quarterly Journal of the Royal Meteorological Society* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society. This article is published with the permission of the Controller of HMSO and the King's Printer for Scotland.

1 | INTRODUCTION

With advances in computational capabilities, it is now feasible to run operational global weather forecasts at kilometre-scale resolution (e.g., Rackow *et al.*, 2025; Wedi *et al.*, 2020; Zeman *et al.*, 2021). At regional scales, numerical weather prediction (NWP) simulations can also be run (mostly in research or demonstration mode) with grid lengths of order 100 m (e.g., Blunn *et al.*, 2022; Boutle *et al.*, 2014, 2016; Hanley *et al.*, 2016; Hanley & Lean, 2024; Jayakumar *et al.*, 2025; Joe *et al.*, 2018; Pereira Filho *et al.*, 2018; Price *et al.*, 2011; Valkonen *et al.*, 2020; World Meteorological Organisation, 2024). Although these simulations typically improve the representation of the atmosphere, they have a high computational cost. They also come with other challenges: there are issues with coupling the dynamical core to the model physical parametrizations at this scale, due to increased temporal and spatial resolution of the simulation, plus existing parametrizations for phenomena that would have been subgrid in coarser models, but are now fully or partially resolved, may not be suitable (e.g., Malardel, 2013).

One such phenomenon is convection. Convection is important, as it can evolve into large organized mesoscale convective systems, leading to high-impact severe weather and precipitation events that cause flash flooding and risk to life (e.g., Li *et al.*, 2023; Rasmussen *et al.*, 2015). These events are expected to become more common in a warmer climate (e.g., Houze Jr. *et al.*, 2019). Therefore, having simulations that can predict the onset of these convective storms better is essential. Note that a more complete description of the challenges and opportunities of hectometric modelling can be found in Lean *et al.* (2024).

Previous studies have shown that some aspects of convection are improved by using grid lengths less than 1 km. For example, rainfall rate distribution, updraft size and strength, precipitation onset and organisation (e.g., Hanley *et al.*, 2015, 2016; Hanley & Lean, 2024; Nicol *et al.*, 2015), plus orographic enhancement of precipitation (Bain *et al.*, 2022) have seen improvements in sub-kilometre simulations. Unfortunately performing simulations on a finer grid does not solve all the deficiencies in sub-kilometre forecasting, due to the grid length used in the simulations being comparable with the dominant turbulence length-scales in the atmospheric boundary layer. Sub-kilometre simulations using the Met Office Unified Model (UM) typically simulate too many small precipitating showers and tend to have clouds that are too deep (e.g., Bergemann *et al.*, 2022; Hanley & Lean, 2024; Jucker *et al.*, 2020). This result points to the need for a boundary-layer turbulence scheme that is better suited to sub-kilometre resolutions.

However, the systematic evaluation of the turbulent mixing in sub-kilometre NWP models compared with observations is rare and often limited to case studies or near-surface variables such as 2-m temperature (e.g., Bauer *et al.*, 2020, 2023; Lean *et al.*, 2019; Matějka & Láská, 2022). Goger and Dipankar (2024) and Sakradzija *et al.* (2025) assess the vertical structure of the atmosphere using lidar, but for a limited number of case studies within the Cross-valley flow in the Inn Valley investigated by dual-Doppler lidar measurements (CROSSINN) (Adler *et al.*, 2021) and Field Experiment on submesoscale spatio-temporal variability in Lindenberg (FESST-Val) (Hohenegger *et al.*, 2023) field campaigns, respectively. A recent study by Ahlgrimm and Päsche (2026) assessed winds, turbulent kinetic energy, and eddy dissipation rate in the first 600 m of the boundary layer using Doppler lidar for a five-day summer case study in Lindenberg, Germany. Their focus was to compare two different turbulence schemes for a range of horizontal resolutions and to link this to applications in wind energy and dispersion modelling.

There is a larger body of literature that compares sub-kilometre simulations with different boundary-layer schemes with large-eddy case-study simulations as truth (e.g., Doubrawa & Muñoz-Esparza, 2020; Shin & Dudhia, 2016). These studies typically find that different schemes perform better at different resolutions and also conclude the need for a universal boundary-layer parametrization that is scale-aware and works for all prognostic variables. This view is also supported strongly in Honnert *et al.* (2020). The methodology developed in this study has been designed so that it can be applied to long-term observational datasets to enable statistically robust conclusions to be drawn for a variety of atmospheric regimes, rather than focusing purely on case studies.

Evaluation of sub-km simulations for moist convection typically focuses on comparisons of precipitation structures such as the size of convective cells, number of cells, and precipitation onset time (e.g., Hanley & Lean, 2024; Stein *et al.*, 2014) rather than the boundary-layer turbulent structures themselves, which influence the time of initiation of convection by modifying the vertical distribution of temperature and humidity. This suggests there is a need to evaluate the pre-convective environment alongside analysing the result of the convection initiated. One such method to evaluate vertical mixing in sub-kilometre simulations is to use high-frequency output from a simulation to compare simulated vertical velocity with observations from Doppler lidar more directly. These observations are ideally suited to probing the atmospheric boundary layer, due to their high temporal and vertical resolution, plus these instruments tend to be very stable and can be used

over an extended period time, not just for case studies (e.g., Harvey *et al.*, 2015).

This study presents a new methodology for comparing Doppler lidar observations of the turbulent characteristics of the boundary layer with the same features in sub-kilometre NWP simulations using high-frequency simulation output. This approach enables the determination of turbulence diagnostics (e.g., vertical velocity variance) from the simulations that can be compared directly with observations. The boundary layer is classified into different types, which allows seamless comparison across horizontal grid spacings as well as testing the underlying physics that links the boundary layer with the surface and cloud layers. Here, a case-study day is presented to illustrate the key aspects of this new approach. The horizontal grid spacings considered here are 1.5 km and 300 m, as these represent typical values for operational or quasi-operational regional (e.g., Hanley & Lean, 2024; World Meteorological Organisation, 2024) NWP models. Therefore understanding the representation of turbulence in these models is scientifically valuable. However, the same approach could be applied to simulations with finer horizontal resolution (e.g., 100 m), where the atmospheric boundary layer would be resolved better, but these simulations are much more computationally expensive and not feasible for operational use.

The article is organized as follows. Section 2 outlines the case-study day considered here, June 12, 2023. The observational measurements and Met Office model configuration used in this article, including model diagnostics, are described in Section 3, with a focus on the boundary-layer scheme. Section 4 contains the main results comparing the sub-kilometre simulations and the observations. Finally, the conclusions are presented in Section 5.

2 | CASE-STUDY DESCRIPTION: JUNE 12, 2023

The case study analysed in this article is June 12, 2023. This is the first intensive observation period of the Wessex Convection Experiment (WesCon)–Observing the Evolving Structures of Turbulence (WOEST) field campaign that took place in the south of England in the summer of 2023 (full details of the campaign can be found in Barr *et al.*, 2025).

The day was dominated by high pressure with southeasterly flow over southern England (see Figure 1). Showers, associated with upper-level trough features, initiated over southern Wales and the Midlands at approximately 1230 UTC. These showers continued to develop and merge into larger mesoscale convective systems (Figure 2). These

systems evolved further into an organized convective line orientated meridionally, concentrated over the West Midlands, extending into the southern part of northwest England. Forecasts of these mesoscale convective systems and associated heavy precipitation led to an amber weather warning being issued by the National Severe Weather Warning Service for North Wales and the Midlands for the late afternoon and early evening (National Severe Weather Warning Service, 2023). There was local flash flooding in northwest London and Bedfordshire and operations were interrupted at several UK airports (Slow & Rhoden-Paul, 2023).

The boundary-layer evolution at Chilbolton on June 12, 2023 can be seen in Figure 3. Cumulus cloud first appears at approximately 0700 UTC and caps the boundary layer through most of its development (boundary-layer type VI). The sensible heat flux increases by $100 \text{ W} \cdot \text{m}^{-2}$ at approximately 1100 UTC at the same time as a large non-precipitating cloud moves away and the boundary layer is classified as well-mixed and cloud-free (type IIIa). The growth of the convective boundary layer throughout the morning is evident in the backscatter (Figure 3a), raw vertical velocity (Figure 3c), and standard deviation of the vertical velocity (Figure 3b). At 1300 UTC, the turbulence driven by surface heating is associated with a skewness value of 1 (Figure 3d) and a maximum vertical velocity standard deviation greater than $1 \text{ m} \cdot \text{s}^{-1}$ (Figure 3b). The sensible heat flux peak is approximately $225 \text{ W} \cdot \text{m}^{-2}$ at 1100 UTC (Figure 3e). At 1900 UTC, there is a stratocumulus cloud identified using the negative skewness at cloud base. This coincides with the sensible heat-flux observations becoming negative, indicating a stable boundary layer (type II). After this the cloud disperses. No precipitation is recorded on the six rain gauges located at Chilbolton on this day.

3 | METHODS AND DATA

This section outlines the boundary-layer observations and simulations used in this study. Figure 4 shows schematically the evolution of the convective boundary layer and highlights the key boundary characteristics that are used to evaluate the simulations: aerosol height (height to which aerosols extend vertically in the atmosphere, h_{aer}), cloud-base height (h_{cld}), turbulent mixing height (height to which turbulent eddies extend vertically into the atmosphere, h_{turb}), and sensible heat flux.

3.1 | Doppler lidar observations

The instruments used in this article are located at the Chilbolton Observatory in southern England ($51^{\circ}09'N$,

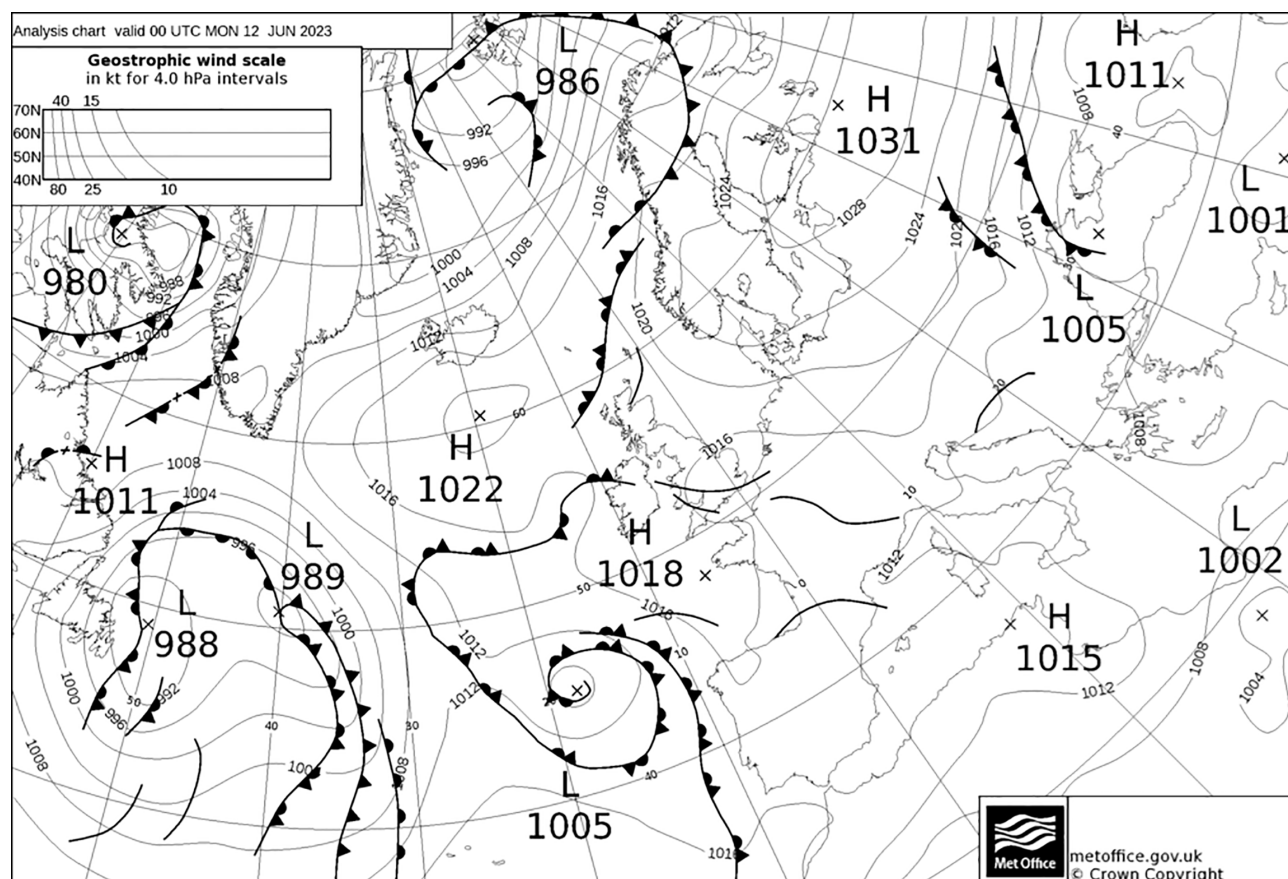


FIGURE 1 UK Met Office surface analysis chart at 0000 UTC on June 12, 2023. Mean sea-level pressure isobars overlaid with active surface fronts (solid black lines with filled symbols) and upper-level troughs (solid black lines).

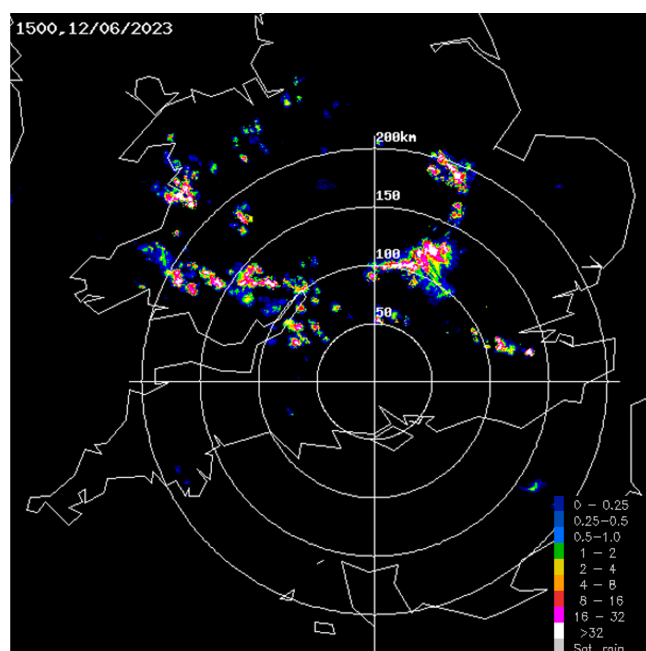


FIGURE 2 Rain rate ($\text{mm}\cdot\text{h}^{-1}$) from the Met Office radar composite at 1500 UTC on June 12, 2023. Chilbolton is located at the centre of the large white cross, with 50-km rings out to 200 km.

$01^{\circ}26'W$). The lidar is a HALO photonics Streamline Pro $1.55\text{-}\mu\text{m}$ heterodyne Doppler lidar (Pearson *et al.*, 2009), which records vertical profiles of attenuated backscatter coefficient, β , and radial Doppler velocity for each 30m range gate every 10s. Calibration of β is achieved using the methodology described in Westbrook *et al.* (2010). A scaling factor is applied to ensure the integrated backscatter through a completely attenuating stratocumulus cloud matches $1/2S$, where $S = 20$ is the expected extinction-to-backscatter ratio. The lidar is installed on a flat roof 3 m above ground level and is configured primarily in zenith-staring mode to monitor vertical velocity w and derive turbulence in the boundary layer, but it also performs conical step-stare velocity-azimuth display (VAD) scans to derive horizontal wind profiles once every 10 min. Each VAD scan takes approximately 120 s and comprises a sequence of 24 stares 15° off-zenith spaced by 15° in azimuth.

The sonic anemometer is on a mast located in a grassy area of the Chilbolton Observatory. It is mounted at a height of 5 m above the ground and measures the three components of the wind (u, v, w) and the sonic temperature at a rate of 20 Hz. Hourly mean sensible

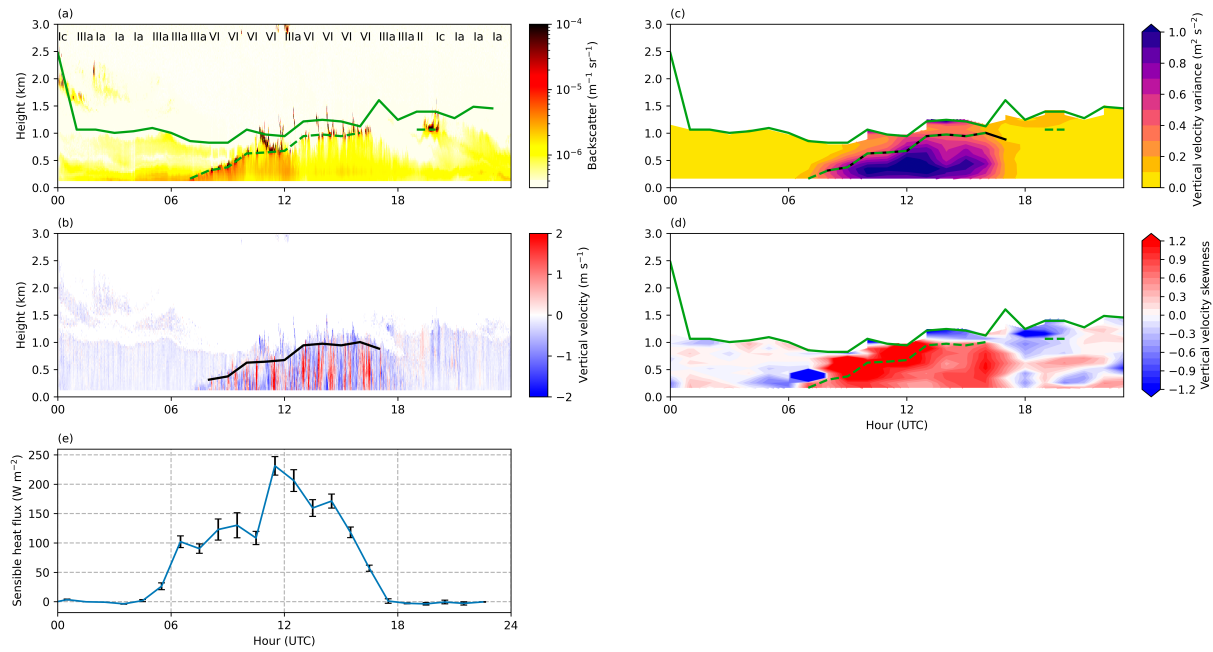


FIGURE 3 Observations taken on June 12, 2023: (a) attenuated lidar backscatter coefficient with the most probable boundary-layer type shown in Roman numerals, (b) hourly mean vertical velocity variance, (c) raw vertical velocity, (d) hourly mean skewness of vertical velocity, and (e) hourly mean surface sensible heat flux with error bars showing the associated standard error. In (a), (b), and (d), the solid green line indicates the diagnosed aerosol top, h_{aer} , the green dashed line indicates the diagnosed cloud base, h_{cld} , and the solid black line indicates the turbulent mixing height, h_{turb} .

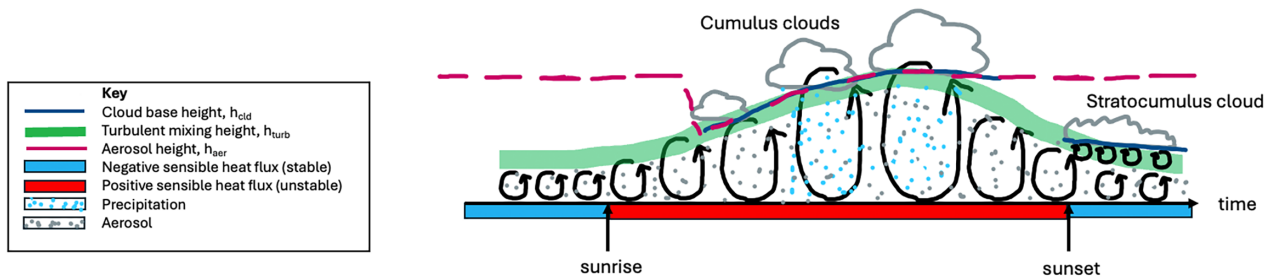


FIGURE 4 Schematic of convective boundary-layer development with parameters that can be determined using the observations outlined in this article. Note that Doppler lidar cannot penetrate through clouds, so the aerosol height and turbulent mixing height are limited to cloud-base height when clouds are present. Arrows indicate turbulent eddies.

heat flux, H , and its standard error are estimated using standard eddy-correlation techniques through the following equation:

$$H = \rho C_p \overline{w'T'}, \quad (1)$$

where ρ is the density of air, C_p is the specific heat capacity of dry air, w' is the fluctuation of vertical velocity from its detrended hourly-mean value, and T' is the fluctuation of sonic temperature (equivalent to virtual temperature, not true air temperature) from its detrended hourly-mean value.

The sonic anemometer and Doppler lidar observations are combined using the algorithm described in Harvey

et al. (2013) to create a dataset of hourly boundary-layer classification, aerosol height, turbulent mixing height, surface stability, and cloud-base height (described below). The algorithm classifies the boundary layer into nine types using the above information. The nine types diagnosed are based on the six types described in Lock *et al.* (2000), with their type I (stable) split into three types (Ia, Ib, and Ic) and their type III (well-mixed) split into two types (IIIa and IIIB), based on the presence of cloud and the number of cloud layers present. A description of these types is outlined in Table 1. Vertical velocity spectra and vertical profiles of vertical velocity variance, vertical velocity skewness, are used to evaluate numerical weather prediction models.

TABLE 1 Boundary-layer type descriptions used in this study, as in Harvey *et al.* (2013).

Boundary-layer type	Description	Diagnosis
Ia	Stable, no cloud	Negative sensible heat flux, no cloud
Ib	Stable, stratus topped	Negative sensible heat flux, cloud present No significant cloud-base turbulence
Ic	Stable, forced cumulus under stratocumulus	Negative sensible heat flux, two cloud layers present Negative cloud-base vertical velocity skewness
II	Stable, stratocumulus topped	Negative sensible heat flux, one cloud layer present Negative cloud-base vertical velocity skewness
IIIa	Unstable, no cloud	Positive sensible heat flux, no cloud
IIIb	Well-mixed, cloud topped	Positive sensible heat flux, cloud present
IV	Decoupled stratocumulus	Positive sensible heat flux, cloud present Negative cloud-base vertical velocity skewness
V	Cumulus under stratocumulus	Positive sensible heat flux, two cloud layers present Positive cloud-base vertical velocity skewness
VI	Cumulus topped	Positive sensible heat flux, one cloud layer present Positive cloud-base vertical velocity skewness

3.1.1 | Aerosol Height, h_{aer}

There are many different definitions of boundary-layer depth in the literature and methods to determine it using remote-sensing instruments (Barlow *et al.*, 2011; Bonin *et al.*, 2018; Davies *et al.*, 2007; Davis *et al.*, 2000; Emeis *et al.*, 2008; Flamant *et al.*, 1997; Hennemuth & Lammert, 2006; Kotthaus *et al.*, 2018, 2020; Pearson *et al.*, 2010; Philibert *et al.*, 2024; Steyn *et al.*, 1999; Vakkari *et al.*, 2015). As in Harvey *et al.* (2013), this study considers the aerosol height, h_{aer} , as the lidar is sensitive to aerosol. This height is determined using the attenuated backscatter coefficient. It is defined as the lowest height at which 80% of the lidar profiles within an hour have no detectable backscatter. This is similar to using a gradient method (e.g., Flamant *et al.*, 1997). During the day, h_{aer} and other measures of boundary-layer height can be equivalent. However, during the night h_{aer} gives a depth representative of the residual layer rather than a measure of the depth of the stable boundary layer. Also, note that the lidar beam is rapidly attenuated by cloud, therefore in the presence of thick cloud, h_{aer} will be at cloud base (see dashed pink contour in Figure 4).

3.1.2 | Cloud presence, height h_{cld} , and type

As in Harvey *et al.* (2013), all clouds with their base at or below h_{aer} are considered to be generated by boundary-layer processes and are therefore included in the

diagnosis of the boundary-layer type. Cloud is determined using a backscatter threshold of $5 \times 10^{-5} \text{ m}^{-1} \cdot \text{sr}^{-1}$. This threshold is consistent with that used in previous studies (e.g., Hogan *et al.*, 2004; Westbrook *et al.*, 2010). Cloud is considered to be present during a given hour if a cloud is detected for more than 5% of the hour-long window. Cloud-base height, h_{cld} , is defined as the hour-mean lowest height that exceeds the backscatter threshold for cloud presence. For broken cloud, the backscatter information is also used to determine the number of cloud layers present. The cloud-base heights within each hour are binned into intervals of 108 m (three lidar range gates) and a histogram of these heights is created. Independent cloud layers are identified as peaks in the histogram separated by at least one bin where the probability of cloud is less than 5%. This threshold is used to ensure that ascending or descending layer clouds (e.g., in the vicinity of fronts) are not diagnosed as multiple layers.

3.1.3 | Turbulent mixing height, h_{turb}

In this study, turbulent mixing height is defined as the maximum height of the hour mean vertical velocity variance being greater than $0.1 \text{ m}^2 \cdot \text{s}^{-2}$. Where there is cloud present, the turbulent mixing height is set to equal cloud base, as above this the lidar experiences high levels of extinction and is unreliable (Figure 4). This is consistent with other approaches in the literature (e.g., Barlow *et al.*, 2011; Lean *et al.*, 2019). As in Harvey *et al.* (2013), the

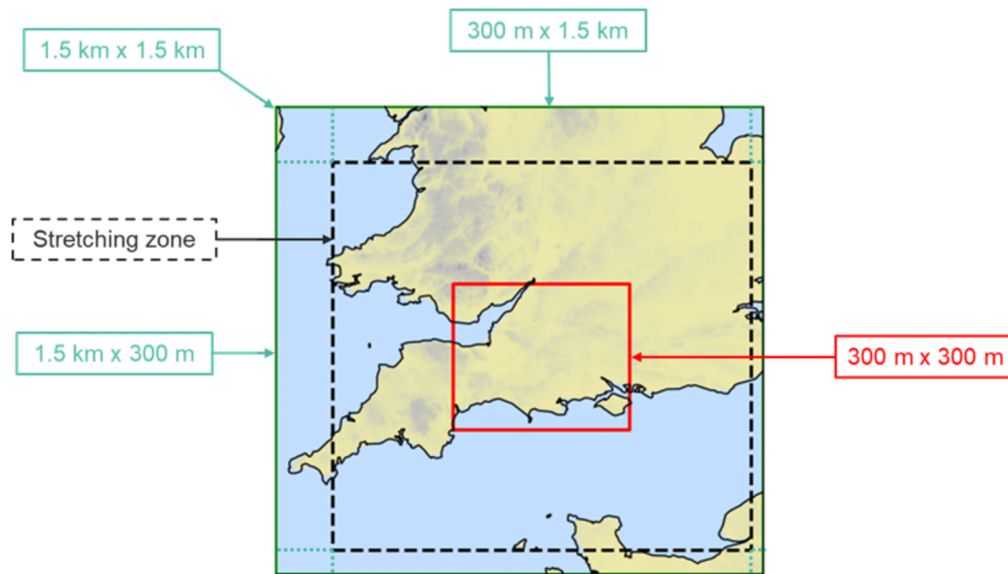


FIGURE 5 The variable-resolution 300-m Wessex model (WMV) domain. The red line shows the inner 300-m grid length region, the region between the red line and the dashed black line shows the stretching zone, and beyond the dashed black line lies the 1.5-km grid length region. Taken from Hanley and Lean (2025).

vertical velocity variance is calculated using spatial means over each set of three adjacent range gates (covering 90 m). Time means are one-hour averages centred on the hour of interest.

Cloud type is determined using the sign of the minimum of the vertical velocity skewness at cloud base. The sign indicates whether stratocumulus or cumulus cloud is present. Negative skewness indicates that the turbulence is driven from cloud-top cooling and therefore stratocumulus cloud is present. Positive skewness indicates surface-driven turbulence associated with cumulus cloud (following (Hogan *et al.*, 2009; LeMone, 1990; Lothon *et al.*, 2009; Moeng & Rotunno, 1990).

3.1.4 | Surface stability

The stability of the surface layer is determined using the sign of the sensible heat flux, H . The probability of H being negative, that is, the surface layer being stable, is found by calculating the sampling error of H using an autocorrelation method (Wilks, 2019). Full details of this calculation can be found in Harvey *et al.* (2013). Typically, H is negative at night and positive during the day (Figure 4).

3.1.5 | Boundary-layer type

The vertical velocity variance measurements, cloud type, number of cloud layers, and surface stability are combined,

as in Harvey *et al.* (2013), to determine the boundary-layer types described in Table 1.

3.2 | Model description

The simulations were performed using version 13.0 of the UM (Davies *et al.*, 2005; Wood *et al.*, 2013) and used the Regional Atmosphere and Land (RAL) version 3.2 science configuration (Bush *et al.*, 2025). The sub-kilometre simulations were performed using the variable resolution 300-m Wessex model (WMV: Hanley & Lean, 2025). The WMV domain, shown in Figure 5, consists of three regions: a high-resolution fixed-grid-length inner region with a grid length of 300 m, a variable-resolution stretching zone, which uses an inflation factor of 1% to increase the grid length from 300 m to 1.5 km, and a fixed-grid-length outer region of 1.5 km. The WMV was one-way nested inside the 1.5-km grid length UKV (Tang *et al.*, 2012), with initial data provided by the 0000 UTC operational UKV analysis and lateral boundary conditions updated every 15 min. Both the UKV and WMV domains have 70 vertical levels on a stretched grid, with 23 levels in the lowest 2 km. The UKV was run using a time step of 60 s while the WMV used a 12-s time step. Both models included the Met Office–Reading Urban Surface Exchange Scheme (MORUSES) urban scheme (Porson *et al.*, 2010) used operationally in the UKV; however, the WMV land use was derived from the 300-m resolution European Space Agency (ESA) Climate Change Initiative (CCI) land-cover dataset (Swinnen *et al.*, 2023), whereas

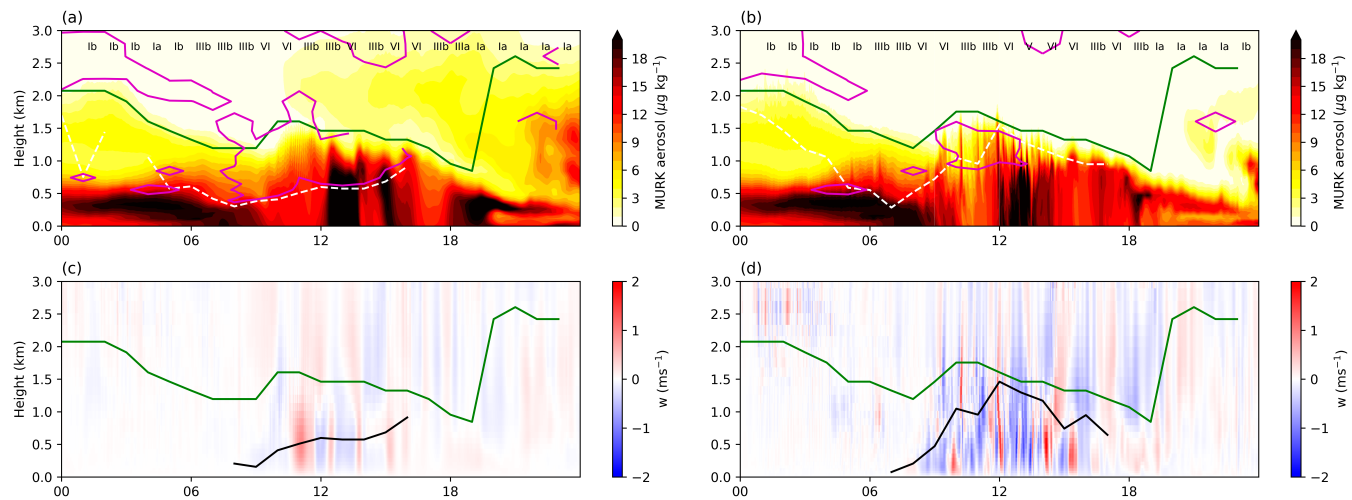


FIGURE 6 Total aerosol with the most probable boundary-layer type shown in Roman numerals for (a) UKV and (b) WMV simulations for June 12, 2023. The magenta contour shows where cloud liquid water is greater than $0.012 \text{ g} \cdot \text{kg}^{-1}$. Vertical velocity is shown in (c) for UKV and (d) for WMV. The solid green line indicates the diagnosed aerosol top, h_{aer} , the dashed white line indicates the diagnosed cloud base, h_{cld} , and the solid black line indicates the diagnosed turbulent mixing height, h_{turb} .

the UKV land use was derived from the 1990 25-m resolution Institute of Terrestrial Ecology (ITE) land-cover dataset. MURK aerosol (a single prognostic aerosol tracer used primarily for visibility prediction, Clark *et al.*, 2008) and soil moisture are both downscaled from the UKV analysis. This setup is similar to that used to produce the 300-m ensemble forecasts that were used to inform the selection of intensive observation periods during the WesCon-WOEST field campaign (Hanley & Lean, 2025) and in the Met Office London Model (Hanley & Lean, 2024).

A full description and evaluation of RAL3 is provided in Bush *et al.* (2025). This configuration uses both the new bimodal cloud scheme (Van Weverberg *et al.*, 2021) and the Cloud AeroSol Interacting Microphysics (CASIM) multimoment microphysics scheme using double-moment capability (Field *et al.*, 2023). The boundary-layer scheme in RAL3.2 blends 1D vertical mixing together with a Smagorinsky subgrid turbulence scheme (Boutle *et al.*, 2014). The blending uses a weighting that tends to 1 (i.e., the 1D boundary-layer mixing length) outside turbulent layers. Within turbulent layers, the blending depends on the ratio of the boundary-layer depth to the simulation grid size (Boutle *et al.*, 2014). The blending weight is based on the turbulent kinetic energy partitioning within the boundary layer given by Honnert *et al.* (2011). RAL3.2 also includes changes to remedy issues with the semi-Lagrangian advection scheme, which leads to non-conservation, and other updates to bring together the science configurations used in regional and global configurations of the UM (see Bush *et al.*, 2025).

3.3 | Model boundary-layer diagnostics

In the absence of a forward model for lidar backscatter, MURK aerosol (Clark *et al.*, 2008) is used as a proxy for backscatter and a simple percentile threshold method is used to determine h_{aer} in the simulations. For each simulation output time, the lowest height where the MURK aerosol falls below 20% of the value in the lowest model level is found. h_{aer} heights are then averaged over each hour to give h_{aer} as in the observational algorithm (Figure 6a,b). There are other more complex algorithms (e.g., Glazer *et al.*, 2025) that could be used to determine h_{aer} , but this approach is chosen as it mimics most closely what is done in the observational algorithm. The value of h_{aer} determined here is largely insensitive to small changes in the percentage threshold chosen.

Cloud presence and cloud base are determined using the high-frequency diagnostic of cloud base for cloud cover greater than 0.1 oktas. This diagnostic is averaged over each hour to give an hourly mean value. Over each hour, the number of cloud layers is also determined. This is calculated by determining whether the distribution of cloud-base heights for each hour is bimodal. As in the observational algorithm, independent cloud layers are identified as peaks in the distribution separated by at least 150 m (Figure 6a,b). This separation is chosen as, in the height range of interest, it is always greater than the vertical grid spacing.

The simulated turbulent mixing height, h_{turb} , is determined using a vertical velocity variance threshold of $0.1 \text{ m}^2 \cdot \text{s}^{-2}$ as in the observations (Figure 6c,d). The total vertical velocity variance is composed of unresolved and

TABLE 2 The contingency table comparing the sign of the cloud-base skewness calculated using the raw Doppler lidar observations of vertical velocity and an estimate of cloud-base skewness using the gradient of the vertical velocity variance profile when the surface is considered unstable for June, July, and August 2023.

		Observed skewness	
		Positive	Negative
Diagnosed skewness	Positive	109	111
	Negative	189	494

resolved components. The resolved component is calculated directly from the high-frequency vertical velocity data on all 70 model levels. This is added to the unresolved component from the subgrid parametrization within the UM, which is an instantaneous diagnostic available hourly.

The sign of the vertical velocity skewness at cloud base is used to determine whether a cloud is cumulus or stratocumulus. As there is no diagnostic available in the UM for unresolved vertical velocity skewness, it is estimated using the negative gradient of the vertical velocity variance (following Moeng & Wyngaard, 1989). Using an estimate of the skewness from the resolved vertical velocity is not appropriate here, especially in the UKV, as so much of the vertical mixing is done by the parametrization scheme and therefore would not be representative (see Section 4.1). This approach has been tested using three months of vertical velocity data from Chilbolton over June, July, and August 2023. Table 2 shows the contingency table comparing the sign of vertical velocity skewness at cloud base, calculated directly from the observed vertical velocities from the Doppler lidar, with an estimate of the sign of the skewness using the negative of the gradient of vertical velocity variance when the surface is unstable. Although this is not a one-to-one relationship, the relationship holds approximately 67% of the time. This increases to 73%

when only including hours where the absolute magnitude of the observed skewness is greater than 0.3 (not shown). Figure 7 shows the sign of the vertical velocity skewness for the case-study day of June 12, 2023 for the UKV and WMV simulations, respectively.

The determination of whether a cloud is decoupled from the surface is performed in the same way as in the observational algorithm. The cloud is considered as decoupled if the sign of the second derivative of the vertical velocity at the height of maximum curvature in the lower half of the boundary layer is negative. Note that the vertical resolution in the simulations is greatly reduced compared with the lidar vertical resolution, with only 12 levels in the first kilometre.

The simulated surface stability is determined using the high-frequency sensible heat flux diagnostic. The data are averaged to give an hourly mean, with the uncertainty estimated from its standard error over the hour of interest (Section 4.2). Finally, the boundary-layer classification for each hour is determined by combining surface stability with vertical velocity variance and skewness in the same way as the observational algorithm to give one of nine possible types given in Table 1 (see roman numerals in Figure 6a,b).

3.4 | Spectral analysis of resolved vertical velocity

Comparing the spectra of resolved vertical velocity between observations and the model simulations provides insight into how the resolved turbulence is represented across different scales. Figure 8 shows a comparison of the resolved vertical velocity spectra at 555 m for the whole of June 12, 2023 for the observations and simulations. Note that the spectral energy density has been multiplied by the frequency to highlight the most energetic scales. The UKV and WMV resolved vertical velocity spectra are created from 60-s and 12-s outputs, respectively. This model time

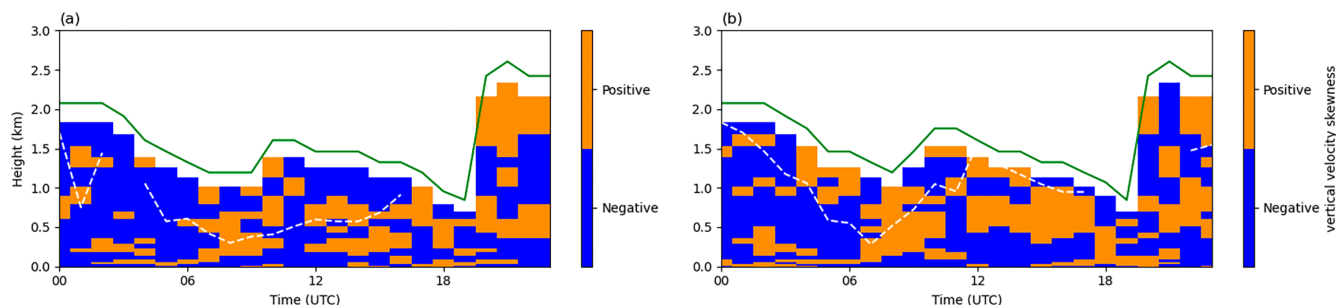


FIGURE 7 Sign of the estimated simulated vertical velocity skewness for (a) UKV simulation and (b) WMV simulation for June 12, 2023. The solid green line indicates the diagnosed aerosol top, h_{aer} , and the dashed white line indicates the diagnosed cloud base, h_{cld} .

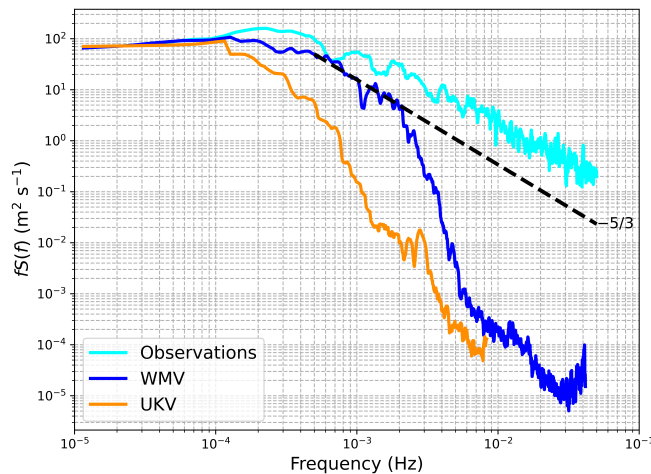


FIGURE 8 Resolved vertical velocity spectra at 555 m for June 12, 2023 for the observations (cyan), the WMV simulation (blue), and the UKV simulation (orange). All spectra have been smoothed using a 20-point moving average. The dashed black line indicates a $-5/3$ slope.

step output is similar to the 10-s temporal resolution of the vertical velocity from the Doppler lidar.

The observations show a continuous energy cascade, capturing the full turbulent spectrum, following a $-5/3$ slope in the inertial subrange. The UKV and WMV simulations have been bias-corrected so that the three datasets have the same magnitude at the lowest frequencies. However, at intermediate frequencies the simulations start diverging from observations, with the UKV simulation diverging with a slope closer to $-7/2$. The WMV simulation stays close to the observations, following a $-5/3$ slope until a frequency of approximately 1×10^{-3} , and then drops off with a slope close to $-13/2$. At high frequencies, the observations continue with a power-law slope but the simulations continue to drop off, indicating that they cannot resolve small-scale turbulent motions. Although the WMV simulation is better at reproducing spectral energy than the UKV simulation, it still misses some higher-frequency variability and therefore is not resolving the full range of turbulent motion observed fully. The UKV simulation is able to resolve structures on a time-scale of approximately 2.5 hours, whereas this time-scale is approximately 15 minutes for the WMV. Similar behaviour is seen at all heights throughout the atmospheric boundary layer (up to approximately 1500 m). The missing high-frequency energy in the model simulations corresponds to unresolved subgrid turbulence, which affects vertical mixing, boundary-layer development, and the formation of boundary-layer clouds. We investigate the model parametrized turbulence further in the next section.

4 | COMPARISON OF OBSERVED AND SIMULATED BOUNDARY LAYERS

4.1 | Evaluation of resolved and unresolved boundary-layer turbulence

The structure of the total vertical velocity field is similar in both the UKV and WMV simulations (Figure 9a,b). The variance is greater than $0.1 \text{ m}^2 \cdot \text{s}^{-2}$ between 0700 and 1800 UTC, with peak values of vertical velocity variance greater than $1.0 \text{ m}^2 \cdot \text{s}^{-2}$ at 1100 UTC. The peak at 1100 UTC is likely to be due to mixing within the cumulus clouds that are present at that time. In the UKV simulations, the vertical velocity variance is dominated by the unresolved component even when the boundary layer is fully developed (Figure 9e,g). In the WMV simulation, the resolved and unresolved contributions have more similar magnitudes but have different structures (Figure 9d,f). The resolved variance extends throughout the depth of the boundary layer, but is only significant between 0800 and 1600 UTC (Figure 9d), and at some times and heights there is no contribution from the unresolved component (Figure 9h). The unresolved variance has a contribution near the surface, which dominates at the beginning and end of the day, and a contribution near the top of the aerosol layer likely to be linked to entrainment across the interface between the boundary layer and the free troposphere (Figure 9f,h).

4.2 | Evaluation of boundary-layer characteristics

Figure 10 compares boundary-layer diagnostics from the 1.5-km UKV and 300-m WMV model time series and the observations for June 12, 2023. The observed aerosol height (Figure 10a) ranges between 1000 and 1500 m (after an initial height of 2500 m related to a high-level cloud). The aerosol height in both simulations has a larger range (750–2000 m) than in the observations. The closest match between observations and the simulation is between 1300 and 1700 UTC, once the convective boundary layer is fully developed. The discrepancies seen between the simulation and observations are likely due to the differences in the diagnostic and associated thresholds used to determine aerosol height, although there may be prediction errors in the simulations and the simulated prediction of cloud cover. The observations are also impacted by the extinction of the lidar beam by hydrometeors, which limits the height to which aerosol can be detected. h_{aer} and h_{cld} vertical resolution matches the resolution of the lidar, which is 30 m. h_{turb} vertical resolution is 90 m, due to the determination

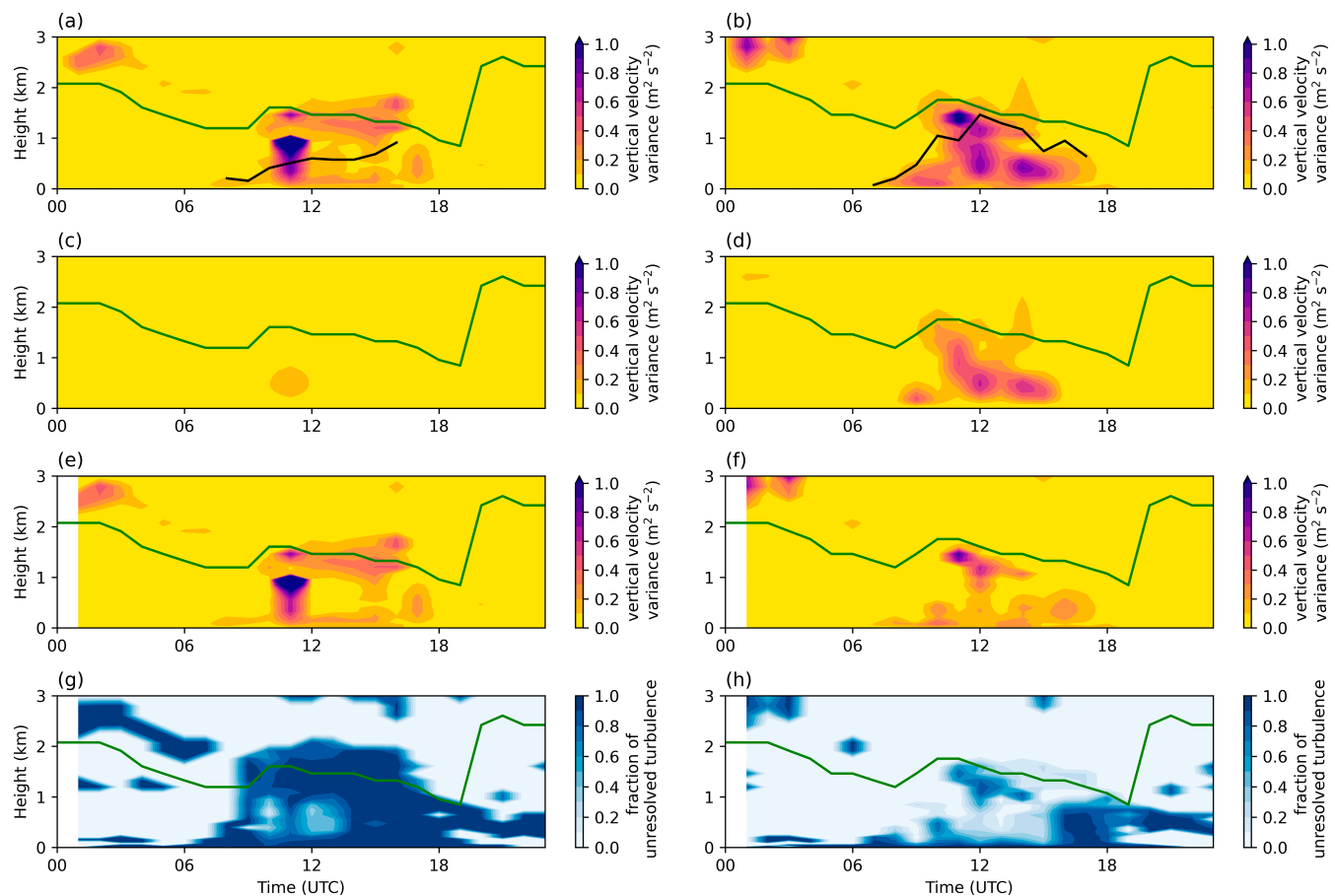


FIGURE 9 Total simulated vertical velocity variance evolution from the (a) UKV and (b) WMV simulations, resolved vertical velocity variance calculated from the (c) UKV and (d) WMV simulation time series, and unresolved vertical velocity variance from the (e) UKV and (f) WMV simulations (UM diagnostic) for Chilbolton on June 12, 2023. The fraction of unresolved vertical velocity variance is shown in (g) for the UKV and (h) for the WMV. The solid green line indicates the aerosol top, h_{aer} , diagnosed from the time series of MURK aerosol and the solid black line indicates the turbulent mixing height, h_{turb} .

of the observed vertical velocity variance. The differences between model levels in the simulated boundary layer (below 2 km) range between 16 and 170 m. It is possible to get a closer match between the observations and simulations by further changing the thresholds (not shown) but the aerosol heights presented here use thresholds that give subjectively closest to what would be diagnosed by eye.

Cloud is more prevalent in the simulations than in the observations between 0000 and 0700 UTC. However, during the development of the convective boundary layer (0700–1200 UTC), the cloud-base height (Figure 10b) and turbulent mixing height (Figure 10c) have a very similar evolution in both simulations and observations, with the WMV having slightly higher values (100–750 m) than the observations despite having lower peak sensible heat flux values, likely due to differences in cloud cover or advection of clouds from elsewhere in the domain. Brooke and Osborne (2022) found that, for a similar site in the south of England, the diagnosed boundary layer developed too

early in the UKV compared with observations, leading to the onset of convection occurring on average 67 min too early and the overprediction of boundary-layer heights of up to 450 m. For this case study, differences of this magnitude are not seen for the UKV simulation; however, the boundary layer in the WMV simulation does grow more quickly than both the observations and UKV simulation. Using the same definition of onset as in Brooke and Osborne (2022), that is, the time at which the vertical velocity variance at a height of 200 m is greater than $0.12 \text{ m} \cdot \text{s}^{-1}$, this onset difference is not seen in this case study, as the onset time in both the simulations and observations is within the hour centred on 0900 UTC (not shown).

The evolution of the surface sensible heat flux is also very similar for the observations and the 300-m WMV simulation, but the 1.5-km UKV simulation has a much flatter evolution (Figure 10e). During the first five hours of the case-study day, the sensible heat flux is only just stable in the observations and both simulations (in fact, at

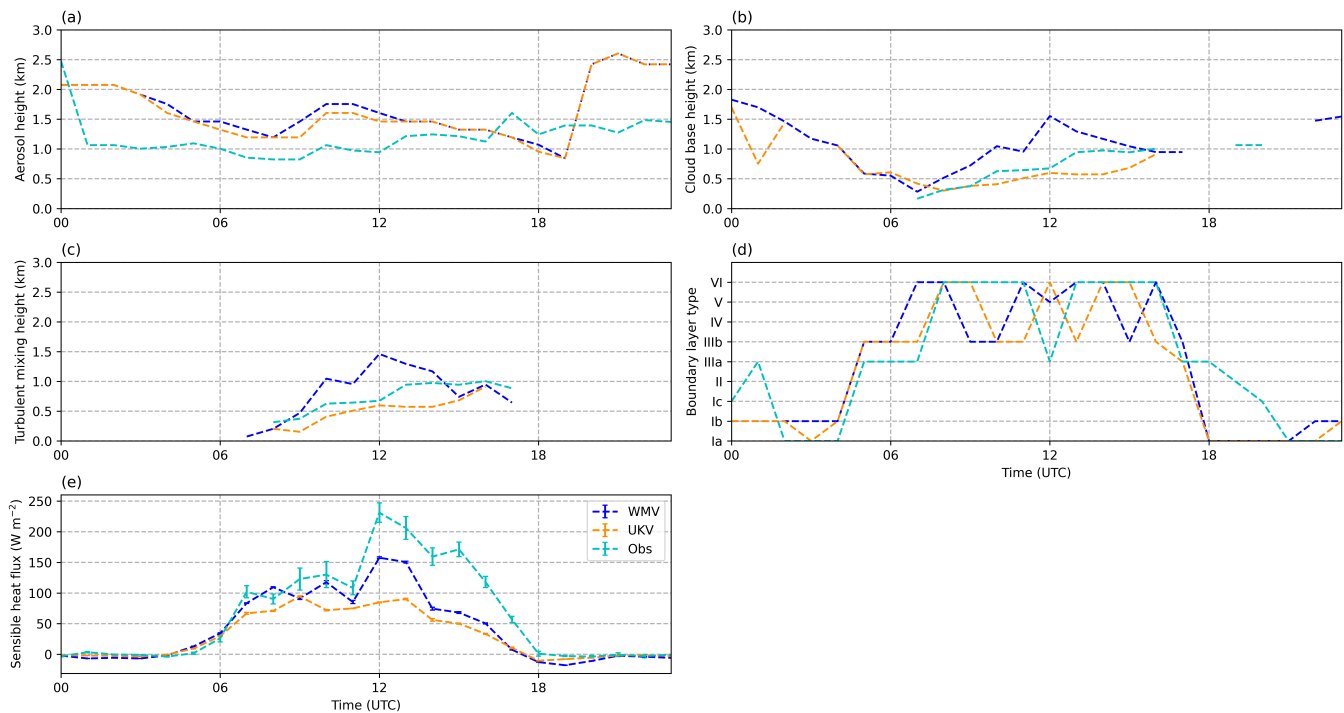


FIGURE 10 Observed (cyan), WMV simulation (blue), and UKV simulation (orange) (a) aerosol height, h_{aer} , (b) cloud-base height, h_{cld} , (c) turbulent mixing height, h_{turb} , (d) boundary-layer type, and (e) surface sensible heat flux for June 12, 2023.

0100 UTC the observations are slightly positive, indicating an unstable surface). The transition from stable to unstable occurs at the same time (0500 UTC) in all datasets. The values of sensible heat flux in the observations and WMV simulation increase in a similar way until 1200 UTC, when the observations have a peak value that is over $50 \text{ W} \cdot \text{m}^{-2}$ higher than in the simulation. However, in the UKV simulation the peak value occurs at 0900 UTC at approximately $100 \text{ W} \cdot \text{m}^{-2}$. This suppressed evolution is likely due to the presence of more cloud in the UKV simulation than in the observations and WMV simulation. It is also possible that the evolution is being influenced by surface properties such as soil moisture, which are not investigated here. The transition from unstable back to stable occurs at approximately 1800 UTC in both simulations. A previous study has found that the UM with a 1.5-km grid length has the transition time from stable to unstable too early by 18 minutes and the transition back to stable 55 min too early on average (Brooke & Osborne, 2022). This is consistent with the findings here, which have a time resolution of 1 hour.

The close agreement between the bulk boundary-layer parameters, h_{aer} , h_{turb} , h_{cld} , and sensible heat flux, is also reflected in the evolution of boundary-layer classification. The WMV simulation starts as type Ib (stable with cloud) and then transitions to type IIIb (unstable with cloud) before becoming type VI (cumulus-capped) at 0700 UTC.

At 0900 UTC, the type changes back to IIIb and then back to VI (cumulus-capped) until 1800 UTC, when it changes to type Ia due to the surface becoming stable again, and the clouds have dispersed. The UKV simulation transitions from stable to unstable and back to stable again at the same time as the WMV simulation. However, the transitions between IIIb and VI throughout the unstable period are not at the same time, due to the different cloud structures present in the simulations. Due to the observations having less cloud at the beginning of the case-study period, the observed boundary-layer classification initially transitions between Ia (stable, no cloud) and IIIa (unstable, no cloud) based on the sign of the sensible heat flux. During the development of the convective boundary layer, the classification oscillates between VI (cumulus-capped) and IIIa based on the presence of cloud. The observations have cloud that persists after the sensible heat flux changes from unstable to stable. This is reflected in the cloudy-stable types (Ic, II) being diagnosed.

Although the overall development of key boundary properties between the observations and simulations follows a similar evolution, there are differences in the evolution of the vertical velocities and their associated statistics. This is highlighted when inspecting the vertical profiles of the vertical velocity variance (Figure 11). To account for differences in the boundary-layer depth, the height axis

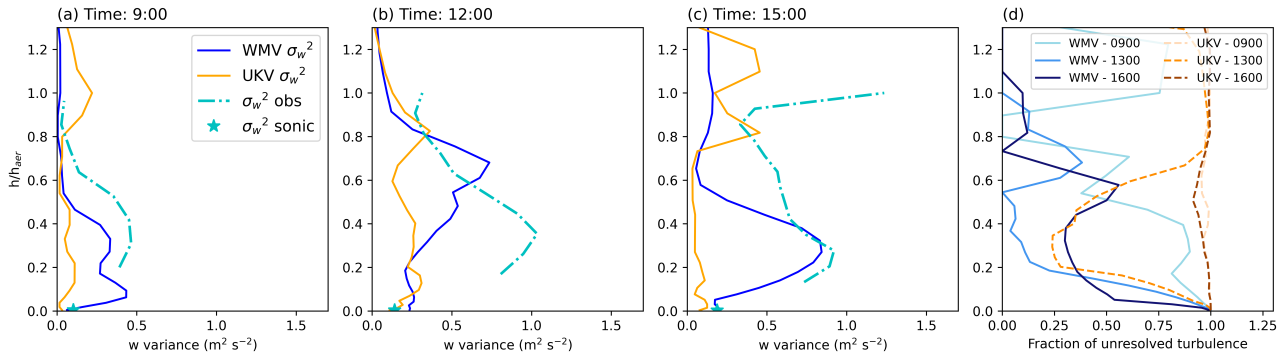


FIGURE 11 Vertical velocity variance ($\text{m}^2 \cdot \text{s}^{-2}$) profiles at (a) 0900, (b) 1200, and (c) 1500 UTC for the UKV and WMV at 0900, 1200, and 1500 UTC. The dot-dashed cyan line is from the lidar observations and the cyan star is vertical velocity variance calculated from the sonic anemometer. The blue solid line is the vertical velocity variance from the WMV simulation and the solid orange line is from the UKV simulation. Panel (d) shows the fraction of unresolved variance as a function of height for both the UKV (dashed lines) and WMV (solid lines) simulations at 0900, 1200, and 1500 UTC.

has been scaled by h_{aer} for both the simulations and the observations.

At 0900 UTC, the lidar has a peak vertical velocity variance value of $0.46 \text{ m}^2 \cdot \text{s}^{-2}$ at $h/h_{\text{aer}} = 0.35$, larger in magnitude than both the UKV and WMV simulations, which have peak values of 0.11 and $0.43 \text{ m}^2 \cdot \text{s}^{-2}$, respectively. The simulations both have a vertical profile that has a double peak not seen in the lidar observations. This structure may be present in the real world, but the lidar observations are not reliable below 200 m. There is good agreement between the near-surface sonic anemometer estimate of the variance and the WMV simulation. The underestimation of vertical velocity variance magnitude and vertical distribution is similar to that presented in Lean *et al.* (2019) and potentially contributes to the early onset of convection seen in previous studies (e.g., Hanley & Lean, 2024, 2025).

By 1200 UTC, the simulated peak variances are both less than the peak value in the observations, which has a maximum value of $1.03 \text{ m}^2 \cdot \text{s}^{-2}$ at $h/h_{\text{aer}} = 0.37$. The simulations both still have vertical profiles with double peaks, with the peak near the top of the boundary layer being more pronounced ($0.68 \text{ m}^2 \cdot \text{s}^{-2}$ at $h/h_{\text{aer}} = 0.69$ in the WMV and $0.36 \text{ m}^2 \cdot \text{s}^{-2}$ at $h/h_{\text{aer}} = 0.84$ in the UKV). At this time, the UKV simulation matches most closely the variance calculated using the sonic anemometer. Note that at this time there is cloud present in both the observations and the simulations, which may be influencing the shape of the variance profiles.

At 1500 UTC, the convective boundary layer is well developed, and this is the time at which the WMV profile matches the observations best in both magnitude and shape. The peak magnitude in the observations within the boundary layer is $0.9 \text{ m}^2 \cdot \text{s}^{-2}$ at $h/h_{\text{aer}} = 0.27$, compared with $0.84 \text{ m}^2 \cdot \text{s}^{-2}$ at $h/h_{\text{aer}} = 0.28$ in the WMV. The UKV profile is much less peaked and has a maximum value of $0.45 \text{ m}^2 \cdot \text{s}^{-2}$ at $h/h_{\text{aer}} = 0.82$, with variance values below

$0.1 \text{ m}^2 \cdot \text{s}^{-2}$ through most of the boundary layer. At this time the WMV simulation matches most closely the variance calculated using the sonic anemometer. The evolution of the fraction of unresolved vertical velocity variance is shown in Figure 11d. This again highlights that the turbulence in the UKV simulation is mostly represented by the subgrid turbulence scheme, with the fraction of unresolved turbulence being close to 1 apart from when the convective boundary layer is fully developed (e.g., 1200 UTC), when the fraction of unresolved turbulence can be as low as 0.25. In the WMV simulation, much more of the turbulence is resolved, especially once the convective boundary layer has developed, when the fraction of unresolved turbulence is less than 0.1 (and as low as zero at some heights and times).

Despite the magnitude of the vertical velocity variance being almost a factor of two larger in the observations at the scaled height of the observational peak than in the simulations, cloud initiates earlier, plus the cloud base is higher, in the sub-kilometre WMV simulation. This is potentially due to the simulated boundary-layer moisture being higher than in reality, due to reduced entrainment of relatively dry free tropospheric air. Increased boundary-layer moisture may also be related to errors in the soil moisture. It has been shown that increased soil moisture can impact boundary-layer evolution significantly during the morning transition (e.g., Brooke *et al.*, 2024). Alternatively, cloud may be initiated non-locally and advected over Chilbolton.

5 | CONCLUSIONS

In this article, we have demonstrated how the representation of turbulence in sub-kilometre numerical weather simulations can be evaluated using high-frequency

model output and comparison with continuously operating Doppler lidar and sonic anemometer observations. Designing the simulation output data to match the observational dataset as closely as possible allows the boundary-layer parametrization scheme to be verified in a direct and consistent way. It also enables an assessment of resolved and unresolved vertical velocity variance.

Firstly, the case study, June 12, 2023, used to illustrate the methodology is presented. This is followed by a brief description of how the observations are used to derive the boundary-layer quantities of aerosol height, cloud-base height, turbulent mixing height, sensible heat flux, and boundary-layer category. The same treatment is then applied to new high-frequency output of vertical velocity and aerosol fields from the sub-kilometre WMV simulation to determine analogous quantities. Comparisons between these fields for the case study, June 12, 2023, are favourable for the broad boundary-layer characteristics of aerosol height, cloud-base height, and turbulent mixing height for both the UKV and WMV simulations. For this case, the WMV simulation has a higher cloud-base height and turbulent mixing height than both UKV and observations. Despite this good agreement on bulk characteristics, there is a discrepancy in the magnitude of the peak sensible heat flux of approximately $100 \text{ W} \cdot \text{m}^{-2}$ for the UKV and $50 \text{ W} \cdot \text{m}^{-2}$ for the WMV, respectively, but the times of transition from stable to unstable and back again happen in the same hour in both simulations and observations. The boundary-layer classification evolution is also similar, with small timing differences in the presence of cumulus clouds, which is expected.

The simulated and observed vertical velocity variance profiles have large differences in magnitude and vertical structure. The UKV simulations have much lower variance values than the WMV and observations. The majority of the turbulence in the UKV simulation is unresolved, with more turbulence being resolved when the convective boundary layer is established. The WMV variance peak values are also lower than the observations, but the shape of the variance profile is more similar to the observations than the UKV. During the times when the convective boundary layer is well developed, more than 90% of the turbulence is resolved in the WMV. The application of this methodology to a year-long dataset of simulations of varying grid lengths (1.5 km to 100 m) and observations will be the focus of future work. This will enable the identification of model bias and areas that should be focused on for model development and improvement. Conditionally sampling by time of day, boundary-layer type, and time of year will allow the identification of which part of the boundary-layer parametrization is leading to any

systematic errors. The impact of the representation of turbulence on the development of convective clouds and precipitation will also be assessed using the year-long dataset. Furthermore, the 1D boundary-layer scheme, which is part of the blended scheme used here in this study, is also used in global NWP and climate simulations, so there is potential to give insights into the systematic biases in those too (e.g., initiation time in tropical deep convection Hohenegger & Stevens, 2013 and warm/dry bias in regional climate simulations over land (Maraun et al., 2021; Van Weverberg et al., 2018). Evaluating model grid points at nearby observation sites could also enable an assessment of the predictability of the onset of convection and aid comparisons with ensemble NWP simulations.

The method presented here highlights the usefulness of Doppler lidar for the evaluation of boundary-layer turbulence schemes. It could easily be applied to new turbulence parametrization schemes or other NWP simulations. Similar analysis could also be performed where long-term Doppler lidar observations are available (e.g., from the Cloudnet Data Portal).¹ Improvements in the representation of boundary-layer turbulence are wide-ranging and have implications for many applications such as renewable energy and air-quality forecasts.

ACKNOWLEDGMENTS

The authors acknowledge the National Centre for Atmospheric Science (NCAS) and the staff of the Chilbolton Observatory for providing the Doppler lidar, sonic anemometer, and rain-gauge data used in this article. Other observations were taken as part of the WesCon–Observing the Evolving Structures of Turbulence project (WOEST, NERC NE/X018539/1). We also thank Peter Clark and Daniel Shipley for useful discussions about this work.

Natalie J. Harvey, Helen F. Dacre, Andrew N. Ross, Michael Baidu, and Steven Boeing are funded through the UMBRELLA project (NE/X018555/1) as part of the UKRI–Met Office ParaChute programme (ParaChute–Representing turbulence in models–Met Office) and the UKRI–NERC funded TECTTO project (NE/Z503964/1). This work used Monsoon2, a collaborative High-Performance Computing facility funded by the Met Office and the Natural Environment Research Council. This work used JASMIN, the UK's collaborative data analysis environment (<https://www.jasmin.ac.uk>).

CONFLICTS OF INTEREST STATEMENT

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the article; or in the decision to publish the results.

DATA AVAILABILITY STATEMENT

The lidar data used in this study are freely available from the Cloudnet Data Portal: <https://cloudnet.fmi.fi/>. The simulation data that support the findings of this study are available from the corresponding author upon reasonable request.

ENDNOTE

¹<https://cloudnet.fmi.fi/>.

ORCID

Natalie J. Harvey  <https://orcid.org/0000-0003-0973-5794>

Helen F. Dacre  <https://orcid.org/0000-0003-4328-9126>

Kirsty E. Hanley  <https://orcid.org/0000-0002-1002-7667>

Michael Baidu  <https://orcid.org/0000-0003-3305-5633>

Steven Boeing  <https://orcid.org/0000-0003-3794-2563>

REFERENCES

- Adler, B., Gohm, A., Kalthoff, N., Babić, N., Corsmeier, U., Lehner, M. et al. (2021) Crossinn: A field experiment to study the three-dimensional flow structure in the inn valley, Austria. *Bulletin of the American Meteorological Society*, 102, E38–E60.
- Ahlgrimm, M. & Päsche, E. (2026) Wind and turbulence evaluation of the icon model (icon-2024.01-1) at sub-kilometer scales using Doppler Lidar observations. *EGU sphere*, 2026, 1–44.
- Bain, C.L., Bush, M., Buxmann, J.C., Davies, P., Finnenkoetter, A., Flack, D.L.A. et al. (2022) The met Office winter testbed 2020/2021: experimenting with an on-demand 300-m ensemble in a real-time environment. *Meteorological Applications*, 29, e2096. Available from: <https://doi.org/10.1002/met.2096>
- Barlow, J.F., Dunbar, T.M., Nemitz, E.G., Wood, C.R., Gallagher, M.W., Davies, F. et al. (2011) Boundary layer dynamics over London, UK, as observed using Doppler Lidar during REPAREE-II. *Atmospheric Chemistry and Physics*, 11, 2111–2125. Available from: <https://acp.copernicus.org/articles/11/2111/2011/>
- Barr, S., Neely, I.I.R.N., Schultz, D.M. & Stein, T.H. (2025) The Wessex convection experiment (WesCon)-observing the evolving structures of turbulence (WOEST): coordinated aircraft and ground-based observations of turbulence and convection in southwest England during summer 2023.
- Bauer, H.-S., Muppa, S.K., Wulfmeyer, V., Behrendt, A., Warrach-Sagi, K. & Späth, F. (2020) In: Tellus, A. (Ed.) *Multi-Nested WRF Simulations for Studying Planetary Boundary Layer Processes on the Turbulence-Permitting Scale in a Realistic Mesoscale Environment*, Vol. 72. Dynamic Meteorology and Oceanography, 1761740.
- Bauer, H.-S., Späth, F., Lange, D., Thundathil, R., Ingwersen, J., Behrendt, A. et al. (2023) Evolution of the convective boundary layer in a WRF simulation nested down to 100 m resolution during a cloud-free case of LAPE, 2017 and comparison to observations. *Journal of Geophysical Research: Atmospheres*, 128, e2022JD037212. Available from: <https://doi.org/10.1029/2022JD037212>
- Bergemann, M., Lane, T.P., Wales, S., Narsey, S. & Louf, V. (2022) High-resolution simulations of tropical island thunderstorms: does an increase in resolution improve the representation of extreme rainfall? *Quarterly Journal of the Royal Meteorological Society*, 148, 3303–3318. Available from: <https://doi.org/10.1002/qj.4360>
- Blunn, L.P., Coceal, O., Nazarian, N., Barlow, J.F., Plant, R.S., Bohnenstengel, S.I. et al. (2022) Turbulence characteristics across a range of idealized urban canopy geometries. *Boundary-Layer Meteorology*, 182, 275–307.
- Bonin, T.A., Carroll, B.J., Hardesty, R.M., Brewer, W.A., Hajny, K., Salmon, O.E. et al. (2018) Doppler Lidar observations of the mixing height in Indianapolis using an automated composite fuzzy logic approach. *Journal of Atmospheric and Oceanic Technology*, 35, 473–490. Available from: <https://journals.ametsoc.org/view/journals/atot/35/3/jtech-d-17-0159.1.xml>
- Boutle, I.A., Eyre, J.E.J. & Lock, A.P. (2014) Seamless stratocumulus simulation across the turbulent gray zone. *Monthly Weather Review*, 142, 1655–1668. Available from: <https://journals.ametsoc.org/view/journals/mwr/142/4/mwr-d-13-00229.1.xml>
- Boutle, I.A., Finnenkoetter, A., Lock, A.P. & Wells, H. (2016) The London model: forecasting fog at 333 m resolution. *Quarterly Journal of the Royal Meteorological Society*, 142, 360–371. Available from: <https://doi.org/10.1002/qj.2656>
- Brooke, J.K. & Osborne, S.R. (2022) Observations and numerical weather forecasts of land-surface and boundary-layer evolution during an unusually dry spring at a site in central England. *Quarterly Journal of the Royal Meteorological Society*, 148, 15–34. Available from: <https://doi.org/10.1002/qj.4190>
- Brooke, J.K., Best, M.J., Lock, A.P., Osborne, S.R., Price, J., Cuxart, J. et al. (2024) Irrigation contrasts through the morning transition. *Quarterly Journal of the Royal Meteorological Society*, 150, 170–194. Available from: <https://doi.org/10.1002/qj.4590>
- Bush, M., Flack, D.L.A., Lewis, H.W., Bohnenstengel, S.I., Short, C.J., Franklin, C. et al. (2025) The third met Office unified model-JULES regional atmosphere and land configuration, RAL3. *Geoscientific Model Development*, 18, 3819–3855. Available from: <https://gmd.copernicus.org/articles/18/3819/2025/>
- Clark, P.A., Harcourt, S.A., Macpherson, B., Mathison, C.T., Cusack, S. & Naylor, M. (2008) Prediction of visibility and aerosol within the operational met Office unified model. i: model formulation and variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, 134, 1801–1816.
- Davies, T., Cullen, M., Malcolm, A., Mawson, M., Staniforth, A., White, A. et al. (2005) A new dynamical core for the met Office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 131, 1759–1782. Available from: <https://doi.org/10.1256/qj.04.101>
- Davies, F., Middleton, D. & Bozier, K. (2007) Urban air pollution modelling and measurements of boundary layer height. *Atmospheric Environment*, 41, 4040–4049.
- Davis, K.J., Gamage, N., Hagelberg, C., Kiemle, C., Lenschow, D. & Sullivan, P. (2000) An objective method for deriving atmospheric structure from airborne lidar observations. *Journal of Atmospheric and Oceanic Technology*, 17, 1455–1468.
- Doubrawa, P. & Muñoz-Esparza, D. (2020) Simulating real atmospheric boundary layers at gray-zone resolutions: how do currently available turbulence parameterizations perform? *Atmosphere*, 11, 345. Available from: <https://www.mdpi.com/2073-4433/11/4/345>
- Emeis, S., Schafer, K. & Munkel, C. (2008) Surface-based remote sensing of the mixing-layer height-a review. *Meteorologische Zeitschrift*, 17, 621–630.

- Field, P.R., Hill, A., Shipway, B., Furtado, K., Wilkinson, J., Miltenberger, A. et al. (2023) Implementation of a double moment cloud microphysics scheme in the UK met Office regional numerical weather prediction model. *Quarterly Journal of the Royal Meteorological Society*, 149, 703–739.
- Flamant, C., Pelon, J., Flamant, P.H. & Durand, P. (1997) Lidar determination of the entrainment zone thickness at the top of the unstable marine atmospheric boundary layer. *Boundary-Layer Meteorology*, 83, 247–284.
- Glazer, R.H., Grimmond, S., Blunn, L., Fenner, D., Lean, H., Christen, A. et al. (2025) Hectometric-scale modelling of the urban mixed layer evaluated with a dense LiDAR-ceilometer network. *EGU sphere*, 2025, 1–26. Available from: <https://egusphere.copernicus.org/preprints/2025/egusphere-2025-2064/>
- Goger, B. & Dipankar, A. (2024) The impact of mesh size, turbulence parameterization, and land-surface-exchange scheme on simulations of the mountain boundary layer in the hectometric range. *Quarterly Journal of the Royal Meteorological Society*, 150, 3853–3873. Available from: <https://doi.org/10.1002/qj.4799>
- Hanley, K.E. & Lean, H.W. (2024) The performance of a variable-resolution 300-m ensemble for forecasting convection over London. *Quarterly Journal of the Royal Meteorological Society*, 150, 3737–3756. Available from: <https://doi.org/10.1002/qj.4794>
- Hanley, K. & Lean, H. (2025) Urban-scale modelling for WesCon. Available from: https://digital.nmla.metoffice.gov.uk/IO_b240621a-8ebb-431a-92a7-79b36e6a7086/. accessed: 2025-07-09
- Hanley, K.E., Plant, R.S., Stein, T.H.M., Hogan, R.J., Nicol, J.C., Lean, H.W. et al. (2015) Mixing-length controls on high-resolution simulations of convective storms. *Quarterly Journal of the Royal Meteorological Society*, 141, 272–284. Available from: <https://doi.org/10.1002/qj.2356>
- Hanley, K.E., Barrett, A.I. & Lean, H.W. (2016) Simulating the 20 may 2013 Moore, Oklahoma tornado with a 100-metre grid-length NWP model. *Atmospheric Science Letters*, 17, 453–461. Available from: <https://doi.org/10.1002/asl.678>
- Harvey, N.J., Hogan, R.J. & Dacre, H.F. (2013) A method to diagnose boundary-layer type using Doppler Lidar. *Quarterly Journal of the Royal Meteorological Society*, 139, 1681–1693. Available from: <https://doi.org/10.1002/qj.2068>
- Harvey, N., Hogan, R. & Dacre, H. (2015) Evaluation of boundary-layer type in a weather forecast model utilizing long-term Doppler LIDAR observations. *Quarterly Journal of the Royal Meteorological Society*, 141, 1345–1353.
- Hennemuth, B. & Lammert, A. (2006) Determination of the atmospheric boundary layer height from radiosonde and LIDAR backscatter. *Boundary-Layer Meteorology*, 120, 181–200.
- Hogan, R.J., Behera, M.D., O'Connor, E.J. & Illingworth, A.J. (2004) Estimate of the global distribution of stratiform supercooled liquid water clouds using the lite LIDAR. *Geophysical Research Letters*, 31.
- Hogan, R.J., Grant, A.L., Illingworth, A.J., Pearson, G.N. & O'Connor, E.J. (2009) Vertical velocity variance and skewness in clear and cloud-topped boundary layers as revealed by Doppler LIDAR. *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography*, 135, 635–643.
- Hohenegger, C. & Stevens, B. (2013) Controls on and impacts of the diurnal cycle of deep convective precipitation. *Journal of Advances in Modeling Earth Systems*, 5, 801–815.
- Hohenegger, C., Ament, F., Beyrich, F., Löhnert, U., Rust, H., Bange, J. et al. (2023) Festival: the field experiment on sub-mesoscale spatio-temporal variability in lindenbergl. *Bulletin of the American Meteorological Society*, 104, E1875–E1892. Available from: <https://journals.ametsoc.org/view/journals/bams/104/10/BAMS-D-21-0330.1.xml>
- Honnert, R., Masson, V. & Couvreur, F. (2011) A diagnostic for evaluating the representation of turbulence in atmospheric models at the kilometric scale. *Journal of the Atmospheric Sciences*, 68, 3112–3131. Available from: <https://journals.ametsoc.org/view/journals/atsc/68/12/jas-d-11-061.1.xml>
- Honnert, R., Efstathiou, G.A., Beare, R.J., Ito, J., Lock, A., Neggers, R. et al. (2020) The atmospheric boundary layer and the “gray zone” of turbulence: A critical review. *Journal of Geophysical Research: Atmospheres*, 125, e2019JD030317. Available from: <https://doi.org/10.1029/2019JD030317>
- Houze, R.A., Jr., Wang, J., Fan, J., Brodzik, S. & Feng, Z. (2019) Extreme convective storms over high-latitude continental areas where maximum warming is occurring. *Geophysical Research Letters*, 46, 4059–4065. Available from: <https://doi.org/10.1029/2019GL082414>
- Jayakumar, A., Anurose, T.J., Bhati, S., Hendry, M.A., Hayman, G., Gordon, H. et al. (2025) Development of an integrated modeling framework for visibility and air quality forecasting in Delhi. *Bulletin of the American Meteorological Society*, 106, E261–E274. Available from: <https://journals.ametsoc.org/view/journals/bams/106/2/BAMS-D-24-0194.1.xml>
- Joe, P., Belair, S., Bernier, N., Bouchet, V., Brook, J., Brunet, D. et al. (2018) The environment Canada pan and parapan American science showcase project. *Bulletin of the American Meteorological Society*, 99, 921–953.
- Jucker, M., Lane, T.P., Vincent, C.L., Webster, S., Wales, S.A. & Louf, V. (2020) Locally forced convection in subkilometre-scale simulations with the unified model and WRF. *Quarterly Journal of the Royal Meteorological Society*, 146, 3450–3465. Available from: <https://doi.org/10.1002/qj.3855>
- Kotthaus, S., Halios, C.H., Barlow, J.F. & Grimmond, C. (2018) Volume for pollution dispersion: London’s atmospheric boundary layer during Clearflo observed with two ground-based lidar types. *Atmospheric Environment*, 190, 401–414. Available from: <https://www.sciencedirect.com/science/article/pii/S1352231018304333>
- Kotthaus, S., Haeffelin, M., Drouin, M.-A., Dupont, J.-C., Grimmond, S., Haeffelin, A. et al. (2020) Tailored algorithms for the detection of the atmospheric boundary layer height from common automatic lidars and ceilometers (ALC). *Remote Sensing*, 12, 3259.
- Lean, H.W., Barlow, J.F. & Halios, C.H. (2019) The impact of spin-up and resolution on the representation of a clear convective boundary layer over London in order 100 m grid-length versions of the met Office unified model. *Quarterly Journal of the Royal Meteorological Society*, 145, 1674–1689.
- Lean, H.W., Theeuwes, N.E., Baldauf, M., Barkmeijer, J., Bessardon, G., Blunn, L. et al. (2024) The hectometric modelling challenge: gaps in the current state of the art and ways forward towards the implementation of 100-m scale weather and climate models. *Quarterly Journal of the Royal Meteorological Society*, 150, 4671–4708. Available from: <https://doi.org/10.1002/qj.4858>
- LeMone, M.A. (1990) Some observations of vertical velocity skewness in the convective planetary boundary layer. *Journal of Atmospheric Sciences*, 47, 1163–1169.

- Li, P., Song, F., Chen, H., Li, J., Prein, A.F., Zhang, W. et al. (2023) Intensification of mesoscale convective systems in the east Asian rainband over the past two decades. *Geophysical Research Letters*, 50, e2023GL103595. Available from: <https://doi.org/10.1029/2023GL103595>
- Lock, A.P., Brown, A.R., Bush, M.R., Martin, G.M. & Smith, R.N.B. (2000) A new boundary layer mixing scheme. Part i: scheme description and single-column model tests. *Monthly Weather Review*, 128, 3187–3199. Available from: https://journals.ametsoc.org/view/journals/mwre/128/9/1520-0493_2000_128_3187_anblms_2.0.co_2.xml
- Lothon, M., Lenschow, D.H. & Mayor, S.D. (2009) Doppler LIDAR measurements of vertical velocity spectra in the convective planetary boundary layer. *Boundary-Layer Meteorology*, 132, 205–226.
- Malardel, S. (2013) Physics/Dynamics Coupling at Very High Resolution. In: *Proceedings of ECMWF workshop on recent developments in Numerical Methods for Atmosphere and Ocean Modelling*. Reading, UK: ECMWF. Available from: <https://www.ecmwf.int/sites/default/files/elibrary/2014/10948-physicsdynamics-coupling-very-high-resolution-permitted-versus-parametrized-convection.pdf>
- Maraun, D., Truhetz, H. & Schaffer, A. (2021) Regional climate model biases, their dependence on synoptic circulation biases and the potential for bias adjustment: A process-oriented evaluation of the austrian regional climate projections. *Journal of Geophysical Research: Atmospheres*, 126, e2020JD032824. Available from: <https://doi.org/10.1029/2020JD032824>
- Matějka, M. & Láská, K. (2022) Impact of the selected boundary layer schemes and enhanced horizontal resolution on the weather research and forecasting model performance on James Ross island, Antarctic peninsula. *Czech Polar Reports*, 12, 15–30.
- Moeng, C.-H. & Rotunno, R. (1990) Vertical-velocity skewness in the buoyancy-driven boundary layer. *Journal of Atmospheric Sciences*, 47, 1149–1162.
- Moeng, C.-H. & Wyngaard, J.C. (1989) Evaluation of turbulent transport and dissipation closures in second-order modeling. *Journal of the Atmospheric Sciences*, 46, 2311–2330. Available from: https://journals.ametsoc.org/view/journals/atsc/46/14/1520-0469_1989_046_2311_eottad_2_0_co_2.xml
- National Severe Weather Warning Service. (2023) Severe weather warning: 12 june 2023. Available from: https://digital.nmla.metoffice.gov.uk/index.php?name=SO_1950f37e-410b-47ae-8a18-b3dbeb7af84f
- Nicol, J.C., Hogan, R.J., Stein, T.H.M., Hanley, K.E., Clark, P.A., Halliwell, C.E. et al. (2015) Convective updraught evaluation in high-resolution NWP simulations using single-Doppler radar measurements. *Quarterly Journal of the Royal Meteorological Society*, 141, 3177–3189. Available from: <https://doi.org/10.1002/qj.2602>
- Pearson, G., Davies, F. & Collier, C. (2009) An analysis of the performance of the UFAM pulsed Doppler LIDAR for observing the boundary layer. *Journal of Atmospheric and Oceanic Technology*, 26, 240–250. Available from: https://journals.ametsoc.org/view/journals/atot/26/2/2008jtecha1128_1.xml
- Pearson, G., Davies, F. & Collier, C. (2010) Remote sensing of the tropical rain forest boundary layer using pulsed Doppler LIDAR. *Atmospheric Chemistry and Physics*, 10, 5891–5901. Available from: <https://acp.copernicus.org/articles/10/5891/2010/>
- Pereira Filho, A.J., Vemado, F., Saito, K., Seko, H., Rojas, J.L.F. & Abi KARAM, H. (2018) ARPS simulations of convection during TOMACS. *Journal of the Meteorological Society of Japan. Series II*, 96A, 247–263.
- Philibert, A., Lothon, M., Amestoy, J., Meslin, P.-Y., Derrien, S., Bezombes, Y. et al. (2024) CALOTRITON: a convective boundary layer height estimation algorithm from ultra-high-frequency (UHF) wind profiler data. *Atmospheric Measurement Techniques*, 17, 1679–1701. Available from: <https://amt.copernicus.org/articles/17/1679/2024/>
- Porson, A., Clark, P.A., Harman, I.N., Best, M.J. & Belcher, S.E. (2010) Implementation of a new urban energy budget scheme in the MetUM. Part i: description and idealized simulations. *Quarterly Journal of the Royal Meteorological Society*, 136, 1514–1529. Available from: <https://doi.org/10.1002/qj.668>
- Price, J.D., Vosper, S., Brown, A., Ross, A., Clark, P., Davies, F. et al. (2011) COLPEX: Field and numerical studies over a region of small hills. *Bulletin of the American Meteorological Society*, 92, 1636–1650. Available from: https://journals.ametsoc.org/view/journals/bams/92/12/2011bams3032_1.xml
- Rackow, T., Pedruzo-Bagazgoitia, X., Becker, T., Milinski, S., Sandu, I., Aguridan, R. et al. (2025) Multi-year simulations at kilometre scale with the integrated forecasting system coupled to FESOM2.5 and NEMOV3.4. *Geoscientific Model Development*, 18, 33–69. Available from: <https://gmd.copernicus.org/articles/18/33/2025/>
- Rasmussen, K., Hill, A., Toma, V., Zuluaga, M., Webster, P. & Houze, R., Jr. (2015) Multiscale analysis of three consecutive years of anomalous flooding in Pakistan. *Quarterly Journal of the Royal Meteorological Society*, 141, 1259–1276.
- Sakradzija, M., Ahlgrim, M., Beyrich, F., Heerwaarden, C.v., Päsche, E., Görsdorf, U. et al. (2025) Where numerical weather prediction meets large-eddy simulations in a convective boundary layer during the FESSTVaL field experiment. *Quarterly Journal of the Royal Meteorological Society*, 152, e70037. Available from: <https://doi.org/10.1002/qj.70037>
- Shin, H.H. & Dudhia, J. (2016) Evaluation of PBL parameterizations in WRF at subkilometer grid spacings: turbulence statistics in the dry convective boundary layer. *Monthly Weather Review*, 144, 1161–1177. Available from: <https://journals.ametsoc.org/view/journals/mwre/144/3/mwr-d-15-0208.1.xml>
- Slow, O. & Rhoden-Paul, A. (2023) Storms bring flash flooding and travel disruption. Available from: <https://www.bbc.co.uk/news/uk-65879069>. accessed: 2025-07-08
- Stein, T.H.M., Hogan, R.J., Hanley, K.E., Nicol, J.C., Lean, H.W., Plant, R.S. et al. (2014) The three-dimensional morphology of simulated and observed convective storms over Southern England. *Monthly Weather Review*, 142, 3264–3283. Available from: <https://journals.ametsoc.org/view/journals/mwre/142/9/mwr-d-13-00372.1.xml>
- Steyn, D.G., Baldi, M. & Hoff, R. (1999) The detection of mixed layer depth and entrainment zone thickness from LIDAR backscatter profiles. *Journal of Atmospheric and Oceanic Technology*, 16, 953–959.
- Swinnen, E., Van der Tol, C., Blessing, S., Camacho, F., Giering, R., Jolivet, D. et al. (2023) ESA vegetation parameters climate change initiative (vegetation parameters cci) - LAI and fAPAR, version 1.0. Available from: <https://doi.org/10.5285/34e4bfe402c048c783e64eac0f0bca37>
- Tang, Y., Lean, H. & Bornemann, J. (2012) The benefits of the met Office variable resolution NWP model for forecasting

- convection. *Meteorological Applications*, 20, 417–426. Available from: <https://doi.org/10.1002/met.1300>
- Vakkari, V., O'Connor, E.J., Nisantzi, A., Mamouri, R.E. & Hadjimitsis, D.G. (2015) Low-level mixing height detection in coastal locations with a scanning Doppler LIDAR. *Atmospheric Measurement Techniques*, 8, 1875–1885. Available from: <https://amt.copernicus.org/articles/8/1875/2015/>
- Valkonen, T., Stoll, P., Batrak, Y., Koltzow, M., Schneider, T.M., Stigter, E.E. et al. (2020) Evaluation of a sub-kilometre NWP system in an Arctic fjord-valley system in winter. *Tellus A: Dynamic Meteorology and Oceanography*, 72, 1–21.
- Van Weverberg, K., Morcrette, C., Petch, J., Klein, S., Ma, H.-Y., Zhang, C. et al. (2018) Causes: attribution of surface radiation biases in nwp and climate models near the us southern great plains. *Journal of Geophysical Research: Atmospheres*, 123, 3612–3644.
- Van Weverberg, K., Morcrette, C.J., Boutle, I., Furtado, K. & Field, P.R. (2021) A bimodal diagnostic cloud fraction parameterization. Part I: motivating analysis and scheme description. *Monthly Weather Review*, 149, 841–857. Available from: <https://journals.ametsoc.org/view/journals/mwre/149/3/MWR-D-20-0224.1.xml>
- Wedi, N.P., Polichtchouk, I., Dueben, P., Anantharaj, V.G., Bauer, P., Boussetta, S. et al. (2020) A baseline for global weather and climate simulations at 1 km resolution. *Journal of Advances in Modeling Earth Systems*, 12, e2020MS002192. Available from: <https://doi.org/10.1029/2020MS002192>
- Westbrook, C., Illingworth, A., O'Connor, E. & Hogan, R. (2010) Doppler lidar measurements of oriented planar ice crystals falling from supercooled and glaciated layer clouds. *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography*, 136, 260–276.
- Wilks, D. (2019) *Statistical Methods in the Atmospheric Sciences*. Oxford, UK: Elsevier.
- Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M. et al. (2013) An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Quarterly Journal of the Royal Meteorological Society*, 140, 1505–1520. Available from: <https://doi.org/10.1002/qj.2235>
- World Meteorological Organisation. (2024) Paris olympics will advance research into weather forecasting and urban meteorology. Available from: <https://wmo.int/media/news/paris-olympics-will-advance-research-weather-forecasting-and-urban-meteorology>
- Zeman, C., Wedi, N.P., Dueben, P.D., Ban, N. & Schär, C. (2021) Model intercomparison of COSMO 5.0 and IFS 45r1 at kilometer-scale grid spacing. *Geoscientific Model Development*, 14, 4617–4639. Available from: <https://gmd.copernicus.org/articles/14/4617/2021/>

How to cite this article: Harvey, N.J., Dacre, H.F., Hanley, K.E., Lean, H.W., Walden, C.J., Baidu, M. et al. (2026) The evaluation of boundary-layer turbulence in high-resolution numerical weather prediction simulations using Doppler lidar. *Quarterly Journal of the Royal Meteorological Society*, e70230. Available from: <https://doi.org/10.1002/qj.70230>