Combined Wind Lidar and Cloud Radar for Wind Profiling

José Dias Neto¹, Louise Nuijens¹, Christine Unal¹, and Steven Knoop²

¹Delft University of Technology
²Royal Netherlands Meteorological Institute

Correspondence: José Dias Neto (J.DiasNeto@tudelft.nl)

Abstract. This paper introduces an experimental setup for retrieving horizontal wind speed and direction profiles by combining wind lidar and cloud radars and the 2 level datasets produced. The experiment took place in Cabauw, the Netherlands, between September 13th and October 3rd 2021. The intermediate processing steps for generating the Level 1 and Level 2 data, such as second trip echos filtering, offset correction, wind retrieval, re-gridding and flagging, are described. In Level 1 (https://doi.org/10.5281/zenodo.6926483, Dias Neto (2022a)), the data from lidar and radars are kept in the original spatial and temporal resolution, while in Level 2 (https://doi.org/10.5281/zenodo.6926605, Dias Neto (2022b)), they are re-gridded to a common spatial and temporal resolution. Statistical analyses of the lidar’s and radar’s wind speed and direction profiles indicate a correlation higher than 0.95 for both variables, and the bias of wind direction and speed are 0.24° and -0.16 m s⁻¹, respectively. Applications of this dataset include numerical model validation, momentum transport studies and spectral analysis for different cloud regimes.

1 Introduction

Wind is an essential component in nearly every weather phenomenon on Earth through its transport of heat, moisture and scalars. How winds blow sets patterns of precipitation on large scales through e.g. atmospheric rivers or monsoons (Gimeno et al., 2014; Zemp et al., 2014; Naakka et al., 2019; Gimeno et al., 2020), but also on small scales by influencing surface heat and moisture fluxes, convection and cloud development. On large-scales, the wind is driven by thermal (pressure) gradients, while on small-scales, winds are modified by a range of small scale processes, including surface drag, momentum transport and gravity waves. The parameterization of those small-scale processes in weather and climate models remains uncertain and persistent wind biases likely related to such processes continue to exist. Wind observations for data assimilation are limited to the surface layer over land where meteorological stations exist, and to single height levels over ocean from satellite scatterometer data and atmospheric motion vectors at selected heights. Recently, wind profiles have become available from the first wind lidar from space (European Space Agency’s Aeolus mission).

In the last decade, due to the current global transition from fossil fuel power plants to systems based on renewable energy sources, the wind energy industry has bloomed, requiring also much more high resolution wind observations to understand the
temporal and spatial scales with which wind varies, including over ocean. In the wind energy community much focus is given
to understanding turbulence and dry convection in the atmospheric boundary layer, as well as topography and surface drag.
Wind turbulence can strongly affect energy production (Elliott and Cadogan, 1990; Peinke et al., 2004; Clifton and Wagner,
2014) or even damage the wind turbines (Kelley et al., 2006). Less attention is given to how the wind is influenced by moist
convection (clouds), even though clouds have long helped visualise and quantify the winds at higher levels (e.g. atmospheric
motion vectors derived from clouds’ motion Velden et al. (2005); Velden and Bedka (2009); Kishtawal et al. (2009); Cordoba
et al. (2017)).

In this study, we attempt to develop a dataset that uses clouds explicitly to make extended wind profiles throughout the
lower atmosphere, and not just in the surface layer, to which meteorological tower and conventional wind lidar measurements
are limited. By measuring winds below and through clouds, we may gain insight into one of the main uncertainties for wind
prediction as highlighted by the numerical weather prediction community: namely, convective momentum transport (CMT). CMT
is broadly defined as convectively-driven transport of momentum through updrafts, downdrafts and the cloud-scale or
even mesoscale circulations that accompany clouds.

Turbulent eddy-resolving models naturally lend themselves to studies of CMT by providing a three-dimensional view of the
multi-scale flows present in the cloud field. However, they are limited by the model’s periodic boundary conditions and the
use of domains smaller than the scales of mesoscale cloud organisation observed in nature. Recent Large-Eddy Simulations
(LES) studies with open boundaries and large domains show that horizontal flows can generate substantial momentum fluxes
(Dixit et al., 2021), which are not present in more traditional (BOMEX, Siebesma and Cuijpers (1995); RICO, vanZanten et al.
(2011)) cases of shallow cumulus convection. They also show that such flows produce so-called counter-gradient transport,
whereby the environmental wind shear is enhanced rather than diminished, as is the case with downgradient (local) turbulent
diffusion. The handful of LES studies focusing on CMT have been carried out for convection over oceans, while studies of
CMT over land are limited.

Recently, Koning et al. (2021) combined 9 years of wind observations from a 200 m tall tower and LES from a commercialised
graphics processing unit (GPU) version of Dutch Atmospheric Large-Eddy Simulation (DALES, Heus et al. (2010))
over Cabauw (The Netherlands) to investigate the relationship between different cloud regimes (clear sky, shallow clouds, non-
convective clouds) and momentum flux and wind shear in the boundary layer. The authors found that clear sky and shallow
clouds days have a similar diurnal cycle of near-surface winds, but the further deepening of the convective boundary layer in the
presence of clouds can lead to a larger daytime increase in near-surface winds. Furthermore, for a similar atmospheric stability
in the surface layer, days with shallow clouds can sustain larger surface momentum fluxes for a given wind gradient. The data
also suggests that more crosswind momentum fluxes are present within the mixed layer - hinting at more organized cloud or
roll cloud structures - compared to days when non-convective clouds are predominant. As part of the Ruisdael Observatory
(https://ruisdael-observatory.nl), the Dutch LES will be run over the entire Netherlands at a 100 m grid daily. High-resolution
wind observations - as the ones introduced herein - will be invaluable in validating these simulations and may even be used for
data assimilation.
Besides meteorological towers (limited to 200 m) or airborne wind measurements (limited in time), ground-based radar and lidars can be used more routinely to measure wind profiles along with cloud profiles. The so-called velocity-azimuth-display (VAD) radar approach for retrieving the wind properties was proposed by Lhermitte (1962); Browning and Wexler (1968); Lhermitte (1969), where the mean Doppler velocity (MDV) as a function of azimuth results in a sine curve. Later, Wilson (1970); Kropfli (1986) used radars to study turbulence in the boundary layer based on the VAD. Based on those previous works, Eberhard et al. (1989) used lidar for studying turbulence for the first time, and since then, lidar has been largely used for this application (Newman et al., 2016; Newman and Clifton, 2017; Mann et al., 2010; Sathe et al., 2015; Smalikho and Banakh, 2017; Bonin et al., 2017). Sathe and Mann (2013) provide a good review of lidar-based experiments. Although radar and lidar have been used for turbulence, few studies combine them to investigate the detailed evolution of wind below and within clouds. One exception is the work from Bühl et al. (2015), where the authors combine radar and lidar observations, but only for retrieving the vertical component of the wind.

This paper introduces an experimental setup where scanning cloud radar and wind lidar observations are combined to retrieve horizontal and vertical wind profiles. By merging radar and lidar observations, this experiment provides, for the first time, continuous profiles of the horizontal wind from near the surface up to cloud tops. The experiment is conducted as part of the Dutch Research Council (NWO) funded Tracing Convective Momentum Transport in Complex Cloudy Atmospheres experiment project (CMTRACE) that targets convective momentum transport under different cloud conditions and across different temporal/spatial scales.

2 Experiment

CMTRACE occurred between September 13th and October 3rd 2021, in Cabauw, The Netherlands. It was carried out by the Delft University of Technology and the Royal Netherlands Meteorological Institute, in coordination with the Ruisdael Land-Atmosphere Interactions Intensive Trace-gas and Aerosol measurement campaign (RITA, https://ruisdael-observatory.nl/the-rita-2021-campaign), which took place at the same site, allowing the synergy of combining observations from in situ measurements (ground-based and airborne) with the CMTRACE remote sensing observations. For this experiment, one wind lidar and two cloud radars were installed close to each other; the wind lidar provided information on the sub cloud layer winds, while the cloud radars provided information on the cloud layer winds. The horizontal distance between them was less than 60 m (Figure 1). This small separation between the instruments was intended to optimize overlap in sampled volume.

2.1 Instruments

The first cloud radar is a dual-frequency (35 and 94 GHz) scanning polarimetric frequency-modulated continuous-wave radar (FMCW) produced by Radiometer Physics GmbH (hereafter CLARA, CLoud Atmospheric Radar). This radar allows setting different configuration parameters for different range intervals (e.g., range resolution, integration time). During the campaign, three range intervals were used, and the radar settings from each range are summarized on Table 1 and Table 3. CLARA was operated performing clockwise and anti-clockwise periodical Plan Position Indicator scans (azimuths: 0-359.9°) with an
The wind lidar is a WindCube-200s scanning lidar produced by Vaisala (hereafter WindCube). It is a pulsed system operating at a wavelength of 1.54 µm, and it is capable of scanning at different azimuths and elevations. This lidar also allows defining different integration times (from 0.1 up to 10 s) and range resolution (from 25 up to 100 m). During the campaign, the WindCube was operated following the six-beam scanning strategy proposed by Sathe et al. (2015), which, according to the authors, provides more information about turbulence than the VAD. The six-beam scanning strategy consists in measuring the radial velocity at five azimuths equally separated by 72° at a specific elevation angle and one additional measurement at 90° elevation (see Sathe et al. (2015) for a complete description). During CMTRACE, the elevation of the slanted measurements was set to 75°, producing a conical sampling volume equivalent to that from CLARA. Table 1 lists additional configuring parameters used by the WindCube.

2.2 Weather characteristics

The 21 days of measurements were characterized by different weather conditions and a diverse cloud cover. For a rapid overview of the weather conditions and cloud coverage, Table 4 provides the daily estimated duration in hours of low-level clouds (LLC), mid-level clouds (MDC), high-level clouds (HLC), deep convective clouds (DCC) and stratiform clouds (SC) (cloud levels are described in Lamb and Verlinde (2011), Ch 1). The duration of the cloud cover was estimated by visual
Table 1. Technical specifications and settings of the lidar and cloud radars operated during CMTRACE.

<table>
<thead>
<tr>
<th>Specifications</th>
<th>MARA&lt;sup&gt;a&lt;/sup&gt;</th>
<th>CLARA&lt;sup&gt;a&lt;/sup&gt;</th>
<th>WindCube</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency [GHz]</td>
<td>94</td>
<td>35</td>
<td>1.94 × 10&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td>Chirp repetition frequency [kHz]</td>
<td>14.0 - 5.8</td>
<td>12.8 - 6.4</td>
<td>20&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Doppler velocity resolution [ms&lt;sup&gt;-1&lt;/sup&gt;]</td>
<td>0.03 - 0.04</td>
<td>0.05 - 0.1</td>
<td>-</td>
</tr>
<tr>
<td>3dB Beam Width [°]</td>
<td>0.55</td>
<td>0.84</td>
<td>-</td>
</tr>
<tr>
<td>Nyquist velocity [±ms&lt;sup&gt;-1&lt;/sup&gt;]</td>
<td>11.2 - 4.6</td>
<td>27.5 - 13.7</td>
<td>41</td>
</tr>
<tr>
<td>Range resolution [m]</td>
<td>22.3 - 37.6</td>
<td>22.3 - 37.6</td>
<td>50</td>
</tr>
<tr>
<td>Temporal sampling [s]</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Time of a complete scan [s]</td>
<td>-</td>
<td>72</td>
<td>22</td>
</tr>
<tr>
<td>Minimum / Maximum range [m]</td>
<td>111 / 11974</td>
<td>111 / 11974</td>
<td>104 / 7176</td>
</tr>
</tbody>
</table>

<sup>a</sup> Chirp repetition frequency, Doppler velocity resolution, Nyquist velocity and range resolution depend on the chirp definition; those values are indicated in Table 2 for MARA and in Table 3 for CLARA.

<sup>b</sup> For the WindCube, this value refers to the pulse repetition frequency.

Table 2. Configuration parameters used by the single frequency radar MARA for each chirp sequence.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration Time [s]</td>
<td>0.034</td>
<td>0.137</td>
<td>0.274</td>
</tr>
<tr>
<td>Range Interval [m]</td>
<td>111.7 - 581.3</td>
<td>620.9 - 1997.9</td>
<td>2033.4 - 11974.8</td>
</tr>
<tr>
<td>Range Resolution [m]</td>
<td>22.3</td>
<td>26.9</td>
<td>37.6</td>
</tr>
<tr>
<td>Nyquist Velocity [±ms&lt;sup&gt;-1&lt;/sup&gt;]</td>
<td>11.2</td>
<td>9.9</td>
<td>4.6</td>
</tr>
<tr>
<td>Doppler velocity resolution [ms&lt;sup&gt;-1&lt;/sup&gt;]</td>
<td>0.04</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Chirp Repetition Frequency [kHz]</td>
<td>14.0</td>
<td>12.4</td>
<td>5.8</td>
</tr>
</tbody>
</table>

Table 3. Configuration parameters used by the dual-band radar CLARA for each chirp sequence.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration Time [s]</td>
<td>0.034</td>
<td>0.137</td>
<td>0.274</td>
</tr>
<tr>
<td>Range Interval [m]</td>
<td>111.7 - 581.3</td>
<td>620.9 - 1997.9</td>
<td>2033.4 - 11974.8</td>
</tr>
<tr>
<td>Range Resolution [m]</td>
<td>22.3</td>
<td>26.9</td>
<td>37.6</td>
</tr>
<tr>
<td>Nyquist Velocity [±ms&lt;sup&gt;-1&lt;/sup&gt;]</td>
<td>27.5</td>
<td>19.2</td>
<td>13.7</td>
</tr>
<tr>
<td>Doppler velocity resolution [ms&lt;sup&gt;-1&lt;/sup&gt;]</td>
<td>0.1</td>
<td>0.07</td>
<td>0.05</td>
</tr>
<tr>
<td>Chirp Repetition Frequency [kHz]</td>
<td>12.8</td>
<td>8.9</td>
<td>6.4</td>
</tr>
</tbody>
</table>
inspection of data recorded by MARA. Table 4 also provides the maximum and mean precipitation rate (RR) measured by an optical disdrometer and the daily mean wind speed (WS200) and direction (WD200) derived from the WindCube observations at 200 m above the surface.

Table 4 indicates that most of the days were non-precipitating with few cloud cover (predominantly LLC). During those days, the mean wind speed at 200 m above the ground ranged between 3.6 and 9.6 ms\(^{-1}\); the wind direction was predominantly from the southwest, but for some days, the wind direction changed to the east. During the precipitating days, DCC and SC clouds were present in addition to LLC, MLC and HLC. The wind speed for these days ranged between 6.7 and 12.9 ms\(^{-1}\), and the wind direction was predominately from the southwest with the exception of days where it changed to the east.

Table 4. Daily characterization of the cloud coverage and precipitation during CMTRACE. The first five columns (LLC, MLC, HLC, DCC, SC) indicate the approximated duration in hours of each class of clouds. RR indicates the maximum and mean precipitation rate in mm/h measured by a nearby optical disdrometer. WS200 and WD200 are the mean wind speed and direction derived from the WindCube observation at 200 m above the ground.

<table>
<thead>
<tr>
<th>date yyyy.mm.dd</th>
<th>LLC [h]</th>
<th>MLC [h]</th>
<th>HLC [h]</th>
<th>DCC [h]</th>
<th>SC [h]</th>
<th>RR [max/mean mm/h]</th>
<th>WS200 [m/s]</th>
<th>WD200 [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>2021.09.13</td>
<td>5</td>
<td>15</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>5.38</td>
<td>86.5</td>
</tr>
<tr>
<td>2021.09.14</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>7.6/1.38</td>
<td>7.38</td>
<td>132.35</td>
</tr>
<tr>
<td>2021.09.15</td>
<td>16</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.4/0.65</td>
<td>6.77</td>
<td>285.65</td>
</tr>
<tr>
<td>2021.09.16</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>7.26</td>
<td>295.25</td>
</tr>
<tr>
<td>2021.09.17</td>
<td>7</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>4.56</td>
<td>238.1</td>
</tr>
<tr>
<td>2021.09.18</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>5.84</td>
<td>110.38</td>
</tr>
<tr>
<td>2021.09.19</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>8.09</td>
<td>109.81</td>
</tr>
<tr>
<td>2021.09.20</td>
<td>7</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>7.57</td>
<td>81.7</td>
</tr>
<tr>
<td>2021.09.21</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>3.66</td>
<td>240.9</td>
</tr>
<tr>
<td>2021.09.22</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>6.39</td>
<td>233.46</td>
</tr>
<tr>
<td>2021.09.23</td>
<td>10</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0.9/0.55</td>
<td>10.22</td>
<td>266.07</td>
</tr>
<tr>
<td>2021.09.24</td>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>8.5</td>
<td>239.73</td>
</tr>
<tr>
<td>2021.09.25</td>
<td>11</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>5.9</td>
<td>182.96</td>
</tr>
<tr>
<td>2021.09.26</td>
<td>0</td>
<td>5</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>6.3/2.01</td>
<td>6.5</td>
<td>199.42</td>
</tr>
<tr>
<td>2021.09.27</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>4.5</td>
<td>0</td>
<td>2.6/0.56</td>
<td>10.63</td>
<td>218.92</td>
</tr>
<tr>
<td>2021.09.28</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0/0.0</td>
<td>9.63</td>
<td>200.22</td>
</tr>
<tr>
<td>2021.09.29</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>47.7/2.55</td>
<td>11.81</td>
<td>243.64</td>
</tr>
<tr>
<td>2021.09.30</td>
<td>6</td>
<td>0</td>
<td>11</td>
<td>1</td>
<td>1.5</td>
<td>-</td>
<td>11.87</td>
<td>225.6</td>
</tr>
<tr>
<td>2021.10.01</td>
<td>4</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>8.3/1.71</td>
<td>12.94</td>
<td>199.79</td>
</tr>
<tr>
<td>2021.10.02</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>11.6/1.89</td>
<td>11.38</td>
<td>188.37</td>
</tr>
<tr>
<td>2021.10.03</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>6.8/1.32</td>
<td>8.09</td>
<td>247.44</td>
</tr>
</tbody>
</table>
3 Data Processing

The CMTRACE dataset is structured in 3 levels according to the processing steps applied. Those different processing levels are designed to facilitate the usage of the CMTRACE dataset, and with that, users will have the possibility to choose the data level that better suits their needs. The processing steps for each level are summarised in Figure 2, and they are described in the following sections.

![Flowchart illustrating the CMTRACE data processing](image)

Figure 2. Flowchart illustrating the CMTRACE data processing: The upper part shows the processing steps applied for generating the Level 1 data, and the lower part shows the steps applied for generating the Level 2 data.

3.1 Processing levels

The original data output from the WindCube, MARA and CLARA is defined as the CMTRACE Level 0 dataset. In Level 0, the variables available in the dataset are related to the scatterer properties (e.g. backscattered signal, MDV and spectrum width), and the data from the WindCube is still affected by noise. Note that neither the lidar nor the radars datasets provide the horizontal wind speed and direction observations at this level.

The Level 1 processing starts with the removal of artefacts present in the Level 0 dataset, as indicated in the flowchart (Figure 2). A filter based on the WindCube’s status variable similar to that described in the manual is applied to the WindCube observations to filter out the noise data and non-realistic MDV values. After this filtering step, the WindCube data is still affected by second trip echoes (STE) produced by clouds from altitudes further than the maximum range sampled by the WindCube. A similar issue was also found in previous experiments (Bonin and Alan Brewer, 2017; Bonin et al., 2017). Panels a and b from Figure 4 show an example of the MDV recorded by the WindCube at 0° azimuth and the equivalent radar reflectivity (Ze) recorded by MARA on September 21st 2021. The MDV values from below 2 km are continuous and...
distributed between -2 and 1 m s\(^{-1}\) (Figure 4-a). In contrast, the MDV rapidly changes from -1 to -7 m s\(^{-1}\) between 11 and 12 UTC at around 2 km. The region with fast velocities extends from 2 up to 6 km in range and appears at different times along the day. In contrast, for the same period, MARA’s Ze does not show any cloud below 6 km, and the only clouds with similar shapes appear at altitudes above 8 km. A filter is then applied to minimise the presence of STE in the dataset (see subsection 3.2). After filtering, the WindCube vertical Doppler velocities are stored as the vertical wind component, and a wind retrieval based on a Fourier transform is applied to the slanted azimuthal MDV observations to retrieve the horizontal wind speed and direction (see subsection 3.3). In addition to the information related to wind, the backscattered data is also included in WindCube Level 1 data.

The Level 1 processing applied to CLARA dataset slightly differs from that applied to the WindCube observations; the radar software internally removes the noise, and for this reason, the artefact removal is skipped. The Fourier transform wind retrieval is also applied to CLARA slanted azimuthal MDV observations to retrieve the speed and direction of the horizontal wind. Surprisingly, after the retrieval, CLARA’s wind profiles also had information from the lowest 2 km, where clouds are not present most of the time; insects are likely the source of information in this region (see subsection 3.4). It was also noticed that an alternating offset affected the wind direction. The magnitude of this offset was estimated using the WindCube wind direction profiles as reference. It was found that each one of the range intervals listed in Table 3 was affected by a different offset (see subsection 3.5). After the offset correction, CLARA’s wind speed and direction are stored as the Level 1 products. Yet, in the Level 1 processing of CLARA’s data, an index is generated to quantify the percentage of invalid data for each complete scan for each height. The invalid index value \(I_{inv}\) is calculated as follows:

\[
I_{inv} = 100 \frac{N_{inv}}{N_{azm}}
\]

Where \(N_{inv}\) is the number of invalids values, and \(N_{azm}\) is the number of azimuths. At the current stage, it is not possible to decouple the vertical wind speed component from the hydrometeors’ fall velocities using MARA’s observations. However, MARA’s vertical MDV is also included in Level 1. MARA’s attenuated Ze is also included in Level 1, which can be used for cloud identification.

At the end of Level 1 processing, the data still have their original temporal and spatial resolution, and the data derived from the different instruments are stored in different datasets. The Level 2 processing merges the wind profiles retrieved from the WindCube and CLARA observations and produces continuous profiles of wind speed and direction from the surface up to the boundary layer or cloud top. The vertical and temporal resolution of Level 2 data is set to 50 m (WindCube range resolution) and to 72 s (duration of CLARA PPI scans), which are the coarser resolutions in the Level 1 data. The radar variables from CLARA and MARA are then interpolated to the new spatial resolution. The temporal resolution from the WindCube and MARA data is adjusted to CLARA’s temporal resolution by averaging all profiles within the time interval of each complete PPI scan from CLARA. This last strategy is used to create profiles that represent equivalent air mass volumes. After that, the processing continues by creating a time-height flag to identify the regions where the lidar and radars provide measurements. Finally, the wind speed and direction profiles from WindCube and CLARA are merged following a hierarchical criterium.
The WindCube data has priority over the CLARA data, meaning that CLARA only provides information in regions where the WindCube data is absent. However, not all data from CLARA are incorporated into Level 2. The distribution of velocity differences as a function of the invalid index \( I_{\text{inv}} \) indicates that a systematic bias becomes apparent for \( I_{\text{inv}} \gt 50\% \) (Figure 3). Therefore, all CLARA’s data flagged with \( I_{\text{inv}} \) equal to 50% or larger are not used in the Level 2 generation. In addition to the wind-related variable, Level 2 also contains the lidar and radar backscattered signals and the vertical Doppler velocities. The complete list of variables available in CMTRACE Level 1 and Level 2 data is given in the Appendix A.

![Figure 3](image-url)  
**Figure 3.** Two-dimensional histogram of the differences between the wind speed retrieved from CLARA’s and WindCube’s observations as a function of the invalid index; the horizontal red line indicates 0 m s\(^{-1}\) and the vertical dashed orange line indicates the invalid index value of 50%.

### 3.2 Second trip echoes filter

In subsection 3.1, it was introduced that the WindCube Doppler velocities observations were affected by the presence of STE produced by clouds. Figure 4-(a) illustrates that the STE can happen at any height. Sometimes, the STE appear in regions close to the surface (hereafter Low-STE), contrasting with the surrounding signal produced by aerosols (e.g. between 11 and 12 UTC below 2 km). However, at other times, the STE appear above the region loaded with aerosols (hereafter High-STE) where there is no clear contrast with the surroundings (e.g. signals above 2 km). In order to minimise the occurrence of Low- and High-STE, two filtering approaches were developed.

The Low-STE filter takes advantage of the contrasting characteristics. This filter is based on the temporal anomaly \( (v'_{\text{azm}}) \) of the MDV from each azimuth angle as indicated by Equation 2; \( v_{\text{azm}} \) is the observed MDV, and \( \bar{v}_{\text{azm}}|\Delta t \) is the mean value calculated within a given time window.

\[
v'_{\text{azm}} = v_{\text{azm}} - \bar{v}_{\text{azm}}|\Delta t
\]  

(2)
Figure 4. Time-height plots: (a) mean Doppler velocity measured by the WindCube, (b) radar equivalent reflectivity measured by MARA, (c) backscattered signal measured by a nearby Ceilometer, note that the negative values are already set as NaN, (d) the same as in (a), but after filtering for second trip echoes. The red curve on (c) indicates the retrieved noise height interface.
The exact size of the time window is arbitrary. However, it should be such that a normal distribution can approximate the distribution of the anomalies (e.g. Figure 5); if this condition is fulfilled, the standard deviation (STD) can be used to characterise the anomaly distribution. The resulting distribution is expected to have the following characteristics: the Low-STE anomalies would populate the edges of the distribution, while the anomaly of the true signal due to turbulence would be close to the centre of the distribution. A too large or too small time window could produce an anomaly distribution that deviates from a normal distribution or a distribution where the anomaly of the Low-STE populates the centre while the true signal anomaly is near the edges.

![Figure 5](https://example.com/figure5.png)

**Figure 5.** Example of the probability distribution of the MDV anomalies calculated using observations at azimuth equal to 0° (the same data from Figure 4-a). The vertical red lines indicate the size of the window used for filtering the data, and the light red areas indicate the values removed.

Once the anomaly STD is calculated, a window $n$ times the STD can be used to discard the data outside the defined window, as illustrated in Figure 5. The exact value of $n$ is arbitrary, but it should be a value that removes most of the STE and preserves most of the valid data. In this work, the values used for $\Delta T$ and $n$ are 3.5 hours and 3, respectively. Those values were found after several trials. Since turbulence during nighttime and daytime is different, only anomalies from between 09:00 and 16:00 UTC were used to calculate the STD. Otherwise, if the nighttime anomalies were included, the STD would be smaller than that from 09:00-16:00 UTC, and then 3-STD would still remove data not affected by STE. Note that this approach may fail in situations when the Low-STE values are comparable to the real data.

The High-STE filtering uses the backscattered signal observations from a nearby ceilometer ($\approx 20$ m away from the Wind-Cube) to estimate the height interface that separates the regions loaded with lidar scatterers from clean regions. The ceilometer...
observations are still affected by noise, and therefore the data from the scatterers free region is dominated by noise (Figure 4-(c)). Basically, the noise height interface (NHI) separates the non-noise region (loaded with scatterers) from the noise dominated region (scatterers free).

Because the ceilometer data did not contain the signal to noise ratio information; an alternative approach was applied to filter out the noise. It was noticed that the noise often reached negative values. This characteristic was then explored to remove the noise from the data. The removal starts by setting all negative values as NaN (not a number). Using a moving window over the time and range coordinates, the NaNs are propagated through the noise region. The NHI is then retrieved as the largest altitude of the heights from the noise-free data in the region between the surface and 4 km. Figure 4-(c) shows that the NHI curve closely follows the region loaded with lidar scatterers. Note that this approach will not work in regions where the noise does not reach negative values.

The detected NHI is then used as a threshold to separate the WindCube data into two regions. One region is below the NHI, where the WindCube data is expected to be predominately originated from lidar-scatterers, and the other is from above the NHI, where clouds and High-STE mainly produced the data. An example of the STE filtered data is shown in Figure 4-(d). One can see that most of the STE visible in Figure 4-(a) are no longer present in Figure 4-(d). Note that due to differences in sensitivity between the WindCube and the Ceilometer, it is possible that using the NHI as a height threshold will possibly remove more data than intended.

3.3 Fourier transform wind retrieval

As described in subsection 2.1, the WindCube scanning strategy used during the campaign produced observations at five azimuthal angles that differ from the four azimuthal angles (0°, 90°, 180°, 270°) used for the Doppler-Beam-Swing strategy (Röttger and Larsen, 1990). Similarly, CLARA produced continuous MDV observations in PPI mode.

In order to retrieve the wind speed and direction profiles from both sets of observations, one can use the velocity azimuth display (VAD) method (Doviak and Zrnic, 2006; Browning and Wexler, 1968). This approach assumes that horizontal wind is uniform within the scanned volume and that the vertical velocity of the scatterers is the same for all azimuths. Under those assumptions, the radial velocity can be described as a Fourier series, but only the first coefficient is used for determining the wind speed and direction (Doviak and Zrnic, 2006; Browning and Wexler, 1968).

In this work, the wind speed and direction profiles are derived using the Fast Fourier Wind Vector Algorithm (FFWVA) proposed by Ishwardat (2017), and a brief description of this method is given below. The FFWVA is similar to the VAD method. However, instead of using the Fourier series, it takes advantage of the currently available Fast Fourier Transform algorithms for digital signal processing to decompose the radial Doppler observations in terms of amplitude and phase of their harmonic frequencies. Note that the unit of these frequencies is 1/deg and not 1/time. The amplitude and phase from the first harmonic are used for calculating the wind speed and direction, and the determination of both quantities is summarised as follows:
\[ a + bi = \text{DFT}(V(\phi))|_{1st} \]  

(3)

\[ \phi_d = -\arctan\left(\frac{b}{a}\right) + 180 \]  

(4)

\[ V_h = \frac{2|a + bi|}{N\cos(\alpha)} \]  

(5)

\( a \) and \( b \) are the real and the imaginary parts from the first harmonic, respectively. \( V(\phi) \) is the azimuthal slanted MDV values from one complete scan, and \( \phi \) is the azimuth. \( \phi_d \) is the wind direction azimuth related to the North. \( N \) is an amplitude correction parameter, and its value is equal to the number of data points used in the transformation. \( \alpha \) is the scanning elevation angle.

### 3.4 Radar observations below 2 km

In addition to the data obtained from the cloud layer, CLARA also received echoes from the sub-cloud layer. Due to CLARA’s polarimetric capability, it was also possible to obtain the differential reflectivity ratio (ZDR) from the sub-cloud layer while executing the PPI scans at 75° elevation. To investigate the origin of those sub-cloud layer echoes and if the retrieved wind information would be biased, the maximum ZDR from each PPI scan is combined with the difference between wind speed derived from the WindCube and CLARA’s data from the entire campaign.

The distribution of maximum ZDR stratified by height shows two main regions (Figure 6). The first region is above 2 km, where the histogram shows one main mode at 0 dB and a broadening of the distribution up to 7 dB. In this region, the temperature is colder than 0 °C suggesting that the ZDR signal is likely produced by pristine ice crystals, snow and supercooled liquid water. In the second region, below 2 km, the distribution shows two main modes, the first at 0 dB and the second at 7 dB. In this region, the temperature is warmer than 0 °C suggesting that the mode at 0 dB is likely to be produced by water droplets. However, what could be producing the second ZDR mode at an elevation of 75°? The high ZDR values below 2 km are likely to be produced by insects. Previous studies have already indicated that insects could produce a strong polarimetric signature in the boundary layer region (Wainwright et al., 2017; Achtemeier, 1991; Wilson et al., 1994; Rennie et al., 2010; Martner and Moran, 2001).

Results from previous studies suggest that insects may actively fly and not only be carried by the wind, and due to it, the derived wind information could be biased (Lhermitte, 1969; Achtemeier, 1991; Wainwright et al., 2017). On the other hand, the study from Wilson et al. (1994) suggests that the wind carried the insects, and no systematic bias was found. To evaluate if CLARA’s retrieved wind speed is biased by the insects, the relationship between the retrieved wind speed difference (CLARA - WindCube) and the ZDR was investigated using data from the lowest 2 km only. Figure 7 shows that the difference in velocity for ZDR larger than 4 dB is similar to that from ZDR around 0 dB, and no systematic bias is found. These results suggest
Figure 6. Two-dimensional histograms of the maximum ZDR for each PPI scan from the entire CMTRACE stratified with height. Note that only data with $I_{inv}$ smaller than 50% is used.

Figure 7. Two-dimensional histograms of the difference between WindCube’s and CLARA’s wind speed from the lowest 2 km as a function of the maximum ZDR. Note that only data with $I_{inv}$ smaller than 50% is used.

that, over Cabauw, the insects are likely to be carried by the horizontal wind, giving the possibility to use them to retrieve information from the horizontal wind. Nevertheless, those insects could be flying up or downward actively (Rennie et al., 2010; Achtemeier, 1991), but it was not investigated in this study.
3.5 Radar wind direction offset estimation

After the campaign, it was noticed that the consecutive wind direction profiles from CLARA were periodically biased. In order to assess this bias, direction profiles from CLARA were compared with the direction profiles retrieved from the WindCube data. The histograms of the differences between both sets of profiles from the entire experimental campaign (Figure 8-(a)) show that most points are distributed below 2 km, and the distribution is much broader in this region if compared with the region above 2 km. The probability distribution of differences (Figure 8-(b)) shows that the maxima are constant within each range interval from each chirp sequence, suggesting that the biases are constant in those regions. It also shows that the biases are positive for some profiles, and for others, the biases are negative.

The biases from each range region were calculated as the mean of all probability distribution maxima within each region. The retrieved values are ±1.4° for the first chirp, ±3.2° for the second, and ±5.4° for the third chirp; the red dashed lines in Figure 8-(b) indicate those values. The negative biases are from profiles when the scans increase in azimuth from 0 to 359°, and the positive biases are from scans when the azimuths decrease from 359 to 0°. Figure 12-(b) shows the result of the offset correction, and one can see that the bias along the range coordinate is close to 0°. We speculate that a possible reason for this range dependent offset is that it takes around 1 s for CLARA to sample the 3 range intervals while CLARA rotates with an angular speed of approximately 5 °s⁻¹. Then, the data from each range interval is stored, and the azimuth from when the sampling started is assigned to them, even though each range interval was sampled at a different azimuth.

![Figure 8](https://example.com/fig8.png)

**Figure 8.** Panel (a) shows the two-dimensional histogram of the wind direction differences (CLARA - WindCube), and Panel (b) shows the probability distribution of the differences. The data used in both panels is from the entire CMTRACE campaign. The red dashed line indicates the mean bias for each range interval.
4 Data Evaluation

4.1 Radiosonde comparison to lidar and radar

In order to evaluate the wind speed and direction retrieved from the WindCube and CLARA observations, profiles of both observables from 34 radiosondes launched in De Bilt were used; the launching site in De Bilt is approximately 25 km away from the remote sensing site in Cabauw. For the evaluation, the Level 1 data from the WindCube and CLARA was used, but the High-STE filter was not applied to the WindCube observations in order to evaluate the agreement between the observations from regions above 4 km. For the evaluation, the WindCube and Clara profiles were averaged within a time window of 10 minutes centred at the launching time. Figure 9 shows an example of those profiles from September 27th, 2021. Even though both sites are far from each other, the WindCube and CLARA profiles surprisingly almost overlap the radiosonde profiles.

A statistical analysis combining all radiosonde profiles suggests that the WindCube and CLARA observations of wind speed and direction are comparable to the radiosonde measurements (Figure 10); it is also supported by the statistical metrics bias, RMSE and correlation listed in Table 5. A more precise evaluation of the WindCube and CLARA data near the surface can be made using the measurements from the KNMI mast tower; however, the mast measurements are not available for this publication.

Table 5. Statistical metrics of the WindCube and CLARA data assuming the radiosonde measurements as a reference. RMSE stands for Root Mean Square Error. The bias and RMSE are in m s$^{-1}$ and degrees for the windspeed and wind direction, respectively.

<table>
<thead>
<tr>
<th></th>
<th>WindCube</th>
<th></th>
<th>Clara</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wind direction</td>
<td>wind speed</td>
<td>wind direction</td>
</tr>
<tr>
<td>bias</td>
<td>0.37</td>
<td>0.52</td>
<td>-0.24</td>
</tr>
<tr>
<td>RMSE</td>
<td>12.62</td>
<td>1.98</td>
<td>14.03</td>
</tr>
<tr>
<td>correlation</td>
<td>0.98</td>
<td>0.92</td>
<td>0.96</td>
</tr>
</tbody>
</table>

4.2 Radar and lidar inter-comparison

In addition to the data evaluation using radiosonde observations, the WindCube and CLARA profiles from the entire dataset were compared against each other. Analogous to the previous subsection 4.1, the High-STE filter was not applied to the WindCube data, in order to assess the quality of the wind profiles in the cloud layer.

The statistical analyses of wind speed (Figure 11-(a)) and direction (Figure 11-(b)) reveal that the observations from both instruments are well correlated with relative small bias, 0.24° for wind direction and -0.16 m s$^{-1}$ for wind speed. The agreement between the data from both instruments is also supported by the statistical metrics listed in Table 6. In addition, Figure 12-(a,b) show the difference between CLARA and WindCube observations of wind speed and direction stratified with range. Those figures show that most of the differences are distributed around 0 for both variables, indicating that the observations provided by both instruments are comparable. They also show that, at 3 km, both difference distributions broaden towards the surface,
Figure 9. Profiles of wind direction (panel a) and wind speed (panel b) from 27.09.2021 at 11:00 UTC. The blue, red, and orange lines indicate the radiosonde, CLARA and WindCube observations, respectively.

Figure 10. Two-dimensional histograms of wind speed (panels a and c) and direction (panels b and d) data retrieved by the WindCube (panels a and b) and CLARA (panels c and d) versus all available radiosonde measurements.
indicating that CLARA observations deviate from the WindCube observations. As suggested in subsection 3.4, insects are likely the source of information in the lowest 2 km. It is possible that insects’ random motion is increasing the uncertainty of the retrieved wind profiles.

Given the good agreement between the observations from both instruments, CLARA observations of wind speed and direction from regions below the NHI were included in the merged Level 2 dataset if the WindCube did not provide them. The impact of the broadening of the difference may not be significant for the Level 2 data because most of the observations in the lowest 3 km are from the WindCube as shown in Figure 13.

![Figure 11](image-url)

*Figure 11.* Two-dimensional histograms from data the data pair WindCube-CLARA for wind speed (Panel a) and wind direction (Panel b); the red line indicates the 1 to 1 line.

Table 6. Statistical metrics of the intercomparison between CLARA and WindCube observations. RMSE stands for root mean square error. The bias and RMSE are in ms$^{-1}$ and degrees for the windspeed and wind direction, respectively.

<table>
<thead>
<tr>
<th>CLARA x WindCube metrics</th>
<th>wind direction</th>
<th>wind speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>bias</td>
<td>0.24</td>
<td>-0.16</td>
</tr>
<tr>
<td>RMSE</td>
<td>12.85</td>
<td>0.93</td>
</tr>
<tr>
<td>correlation</td>
<td>0.98</td>
<td>0.99</td>
</tr>
</tbody>
</table>
Figure 12. Two-dimensional probability distribution of the wind speed (Panel a) and direction (Panel b) differences stratified with height.

Figure 13. Distribution of Level 2 data source stratified with height: WindCube in blue and CLARA in green.
5 Application highlights

The CMTRACE dataset may be used for different applications. Here we would like to highlight two ways in which we are currently using the data: 1) the validation of wind and momentum transport in models; and 2) process studies of convective momentum transport (or transport of scalars through the boundary layer). A third application, which we do not further exemplify here, is the assimilation of wind profiles by weather models.

5.1 Model validation

The DALES model is running in near real-time hindcast mode on a (still) relatively small domain (15 x 15 km) centred at Cabauw with a grid spacing of 75 m. As part of the Ruisdael Observatory, these simulations will, in the future, be running on domains spanning the Netherlands at a grid spacing of ≈100 m. Along with model output at Cabauw from KNMI’s mesoscale weather model HARMONIE, Figure 14 shows an initial comparison of observed and modelled winds on September 29th, 2021. On this day, a frontal system passed over the observational site from the early to late morning. The Ze (Figure 14-(a)) and the MDV (Figure 14-(b)) recorded by MARA show that between 00:00 and 12:00 UTC, precipitating deep clouds with stratiform outflow passed over the site, while after the front passage, shallow cumulus and congestus clouds developed with tops between 4 - 7 km.

The Level 2 wind direction (Figure 15-(a)) and speed (Figure 15-(c)) show that before 06:00 UTC, the wind was mainly southerly throughout the lower and middle troposphere. During the frontal passage (06:00 - 12:00 UTC), the wind direction changed from southerly to westerly, while winds of 10 - 15 m s⁻¹ extended from near the surface up to cloud tops. After 12:00 UTC, the horizontal wind was mainly westerly, and the wind speed in the lower boundary layer was faster at about 15 - 20 m s⁻¹. The DALES model simulates the front passage and post-frontal convection (Figure 15-(b,d)), but some differences are apparent. The frontal passage seems to be slower in the model, judging mainly from the slanted wind direction and wind speed signature. Furthermore, in the period of post-frontal shallow convection (06:00 - 12:00 UTC), wind speeds in the observations appear to reach values up to 20 m s⁻¹ (orange-red) that can extend all the way to the surface while DALES maintains weaker surface layer winds. Comparing winds at 0.1 km above the surface (Figure 16-(a),(b)) reveals that the wind turning associated with the front passage in DALES is well simulated, but the wind speeds at 100 m are evidently too weak. Because the DALES winds showed are averages over the 15 x 15 km domain, while the observations represent wind variability at a single location, DALES is not expected to show the convective variations. However, in the observations, such convective variations would average out over time, and the weak wind bias in DALES appears too persistent throughout the day to be caused by the difference in domain-averaged versus point estimates.

The bias is intriguing because earlier comparisons of DALES to observations revealed a too strong wind bias instead (van Stratum et al., 2019). DALES uses a surface roughness length that underestimates the effective regional roughness length at Cabauw, which leads to too small surface stress. Ongoing investigations also include HARMONIE and employ measurement simulators to ensure a fair comparison of observations to models.

https://doi.org/10.5194/essd-2022-268
Preprint. Discussion started: 19 September 2022
© Author(s) 2022. CC BY 4.0 License.
5.2 Momentum transport on different scales

Focusing on the period of post-frontal shallow convection on the same day (Figure 17), we may illustrate the presence of multi-scale flows in the convective boundary layer. Reynolds averaging (Stull, 2003) is applied to the vertical and horizontal wind using a sliding window to separate flows with scales longer and shorter than 10 minutes (Equation 6). \( v_{\text{obs}} \) is the observed wind, \( \bar{v}_{10m} \) is the 10 minutes averaged wind, and \( v'_{10m} \) is the 10 minutes anomaly.

\[
v'_{10m} = v_{\text{obs}} - \bar{v}_{10m}
\]

Assuming wind speeds of 10-20 m\( \text{s}^{-1} \), a 10-minute average window corresponds to a spatial scale of about 6-12 km, which is in the meso-\( \gamma \) range. Figure 17 a) and b) illustrate the fluctuations in the horizontal and vertical wind speed on these scales, while Figure 17 c) and d) show the presence of convective up- and downdrafts that carry different winds. Evidently, during the frontal passage, before 12:00 UTC, the vertical motion is on average downward (red) in the presence of precipitating convection, even though individual updrafts and downdrafts are visible. MARA’s observations from the same period and height show an intensification of Ze and MDV (Figure 14-a, b), which suggests the occurrence of precipitation. It is possible that \( \bar{w}_{10m} \) not only contains information from downward winds but also from the terminal fall velocity of raindrops (Aoki et al., 2016).
Figure 15. Time-height plot of wind direction (panels: a and b) and speed (panels: c and d) from the CMTRACE Level 2 (panels: a and c) and from DALES (panels: b and d) from September 29th, 2022.
Figure 16. Time series of wind direction (panel a) and speed (panel b) from September 29th, 2022 at 0.1 km above the surface. The blue curve is from the lowest available observation in the CMTRACE Level 2; the red is from DALES.

After 12 UTC, mesoscale fluctuations in both the vertical and horizontal wind are evident, where downward motion tends to be accompanied by stronger horizontal winds extending from the top of the boundary layer to the surface and upward motion to weaker horizontal winds extending upward from the surface. Qualitatively comparing the horizontal and vertical speed observation with MARA's Ze suggests that periods of faster horizontal winds and downward motion correlate to periods of congestus and precipitation. These mesoscale fluctuations may be coupled to an organization in the cloud field, and it is a subject of ongoing research in a community that studies the interaction of clouds with circulations. They may also contribute non-negligibly to total momentum transport. Past LES modelling studies, including with DALES, have not been able to accurately represent these mesoscale flows due to their use of small domain sizes and periodic boundary conditions. As DALES will be running on much larger domains with open boundary conditions, the ability to validate such flows with observations is invaluable. In our current ongoing work, we use spectral analysis to derive the contribution of convective and mesoscale fluctuations to total momentum transport.
Figure 17. Time-height plots from September 29th, 2022: (a) CMTRACE Level 2 horizontal wind speed main component (10 minutes averaged), (b) CMTRACE Level 2 lidar vertical mean Doppler velocity (10 minutes averaged, negative means towards the surface), (c) CMTRACE Level 2 horizontal wind speed anomaly, (d) CMTRACE Level 2 lidar vertical mean Doppler velocity anomaly.
6 Conclusions

This manuscript introduces an experimental setup for retrieving horizontal wind speed and direction profiles from near-surface to cloud top, taking advantage of the synergy of using wind lidars to retrieve wind profiles in the boundary layer and cloud radars to retrieve wind in the cloud layer. This first CMTRACE campaign was conducted in Cabauw, the Netherlands and lasted 21 days. During the experiment, most days were non-precipitating with the presence of few clouds (mainly low level clouds). The winds at 200 m were mainly southwesterly, with speeds between 3.6 and 12.9 m s$^{-1}$.

The CMTRACE dataset was processed in two levels. In Level 1, the processing minimizes the presence of second trip echoes that affect the WindCube measurements, and reduces the bias in wind direction profiles derived from the radar observations. The Level 1 observations from each instrument are kept in their original spatial and temporal resolution. In Level 2, the lidar and radar differences in the sampled volume are reduced, and the horizontal wind profiles from both instruments are merged, creating a single profile with a flag to identify the measuring instrument.

Statistical analyses are used to assess the level of confidence of the CMTRACE dataset. When correlating the CMTRACE data with radiosonde measurements, the results show that this correlation is higher than 0.9 for wind speed and wind direction. These results also indicated that the absolute wind speed bias is smaller than 0.55 m s$^{-1}$, and the absolute wind direction bias is smaller than 0.4° (Table 5). The surprising possibility of using CLARA for retrieving wind profiles in the boundary layer, likely due to the presence of insects, allowed the agreement between lidar and radar measurements to be evaluated. The intercomparison between lidar and radar reveals that for wind speed and direction, the correlation is higher than 0.95 (Table 6), and the bias of wind direction and speed are 0.24° and -0.16 m s$^{-1}$, respectively.

Possible applications of the CMTRACE dataset include model validation (e.g., DALES and HARMONIE), spectral analysis of the scales of motion that accompany diverse (organized) cloud fields, process studies of momentum transport, and transport of scalars (e.g., air pollution). The subsequent CMTRACE campaigns are intended to last longer than the first to capture a diverse cloud regime and periods when congestus clouds are predominant. An extension of the experiment to other regions (e.g., in the tropics) is also planned.

7 Data availability

The CMTRACE Level 1 and Level 2 datasets are publicly available on the ZENODO platform. CMTRACE Level 1 can be downloaded from https://doi.org/10.5281/zenodo.6926483 (Dias Neto, 2022a) and Level 2 from https://doi.org/10.5281/zenodo.6926605 (Dias Neto, 2022b). The Level 0 can be requested from the corresponding author.
Appendix A: List of variables

The files from the Level 1 and Level 2 data are structured differently. As described in section 3, Level 1 data from each instrument is kept with its original spatial and temporal resolution. The Level 1 data from the instrument was resampled to a common spatial and temporal resolution and used to generate the Level 2 data. The variables available in Level 1 and Level 2 data are listed in Table A1.

Author contributions. JDN wrote the paper, developed the data processing, and designed the experiment. LJ supervised and designed the experiment and reviewed the manuscript. CU managed the radars, designed the experiment and reviewed the manuscript. SK managed the lidar, designed the experiment and reviewed the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This publication is part of the NWO Talent Scheme Vidi project CMTRACE with project number 192.050, financed by the Dutch Research Council (NWO). We acknowledge the Ruisdael Observatory for providing the infrastructure for this experiment. We acknowledge the Royal Netherlands Meteorological Institute for providing the radiosonde and ceilometer observations (https://dataplatform.knmi.nl/) and support at the Cabauw site. We thank Saverio Guzzo from the Delft University of Technology for managing the network and the data transfer. We also thank Robert Mackenzie from the Delft University of Technology for mounting the instruments on the experimental site and operating it during the experiment.
## Table A1.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>WindCube</th>
<th>Clara</th>
<th>Mara</th>
<th>merged dataset</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>horizontal_wind_direction</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>horizontal wind direction retrieved using the FFT method with respect to true north</td>
</tr>
<tr>
<td>horizontal_wind_speed</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>horizontal wind speed retrieved using the FFT method</td>
</tr>
<tr>
<td>zonal_wind</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>zonal wind retrieved using the FFT method</td>
</tr>
<tr>
<td>meridional_wind</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>meridional wind retrieved using the FFT method</td>
</tr>
<tr>
<td>vertical_wind_speed</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td>observed vertical wind speed (negative towards the ground)</td>
</tr>
<tr>
<td>lidar_relative_beta</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>Attenuated relative backscatter coefficient from the vertical beam</td>
</tr>
<tr>
<td>start_scan</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>starting time of each scan used for generating each profile</td>
</tr>
<tr>
<td>end_scan</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>end time of each scan used for generating each profile</td>
</tr>
<tr>
<td>nan_percentual</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>percentual of invalid data per complete PPI scan</td>
</tr>
<tr>
<td>max_radar_differential_reflectivity</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>maximum ZDR from each complete PPI scan</td>
</tr>
<tr>
<td>radar_vertical_doppler Velocity</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>Doppler velocity observed by a vertically pointing 94 GHz radar (negative towards the ground)</td>
</tr>
<tr>
<td>radar_equivalent_reflectivity</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>attenuated radar linear equivalent reflectivity factor observed by a vertically pointing 94 GHz radar</td>
</tr>
<tr>
<td>radar_spectrum_width</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>Measure of the dispersion of radial velocity within the radar measurement volume</td>
</tr>
<tr>
<td>radar_spectrum_skewness</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
<td>Skewness of the velocity distribution within the radar measurement volume</td>
</tr>
<tr>
<td>rain_rate</td>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
<td>rain fall rate measured by a near by PARSIVEL (optical disdrometer)</td>
</tr>
<tr>
<td>data_flag</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td>instrument identification flag: 0: no data, 1: WindCube lidar, 2: CLARA 35 GHz radar</td>
</tr>
</tbody>
</table>
References


